



MAX-PLANCK-GESELLSCHAFT

**Structural Investigations on Cholesterol binding Membrane Proteins
SREBP cleavage-activating Protein (Scap) and Patched1
by Cryo-EM**

Dissertation

zur Erlangung des akademischen Grades des Doktors der Naturwissenschaft
(Dr. rer. nat.)

Vorgelegt an der Fakultät Chemie und Chemische Biologie der
Technischen Universität Dortmund

Angefertigt am Max-Planck-Institut für molekulare Physiologie in Dortmund

Vorgelegt von
Birte Weyers (geb. Siebolds)



First referee: **Prof. Dr. Stefan Raunser**
Department of Structural Biochemistry, Max Planck Institute of Molecular Physiology,
Dortmund
Faculty of Chemistry and Chemical Biology, Technical University Dortmund

Second referee: **Prof. Dr. Daniel Summerer**
Faculty of Chemistry and Chemical Biology, Technical University Dortmund

Date of submission: 12th of January 2021

Eidesstattliche Versicherung

Weyers, Birte
Name, Vorname

148549
Matrikelnummer

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die obenstehende Belehrung habe ich zur Kenntnis genommen:

Dortmund, 12.01.2021

Ort, Datum

Unterschrift

Titel der Dissertation:

Structural Investigations on Cholesterol binding Membrane Proteins SREBP cleavage-activating Protein (Scap)

and Patched1 by Cryo-EM.

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.

Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

Dortmund, 12.01.2021

Ort, Datum

Unterschrift

Acknowledgements

In this section I would like to take the opportunity to thank all the people which continuously supported me during my PhD time.

First of all, I thank Prof. Dr. Stefan Raunser for giving me the great opportunity to conduct my PhD thesis in his group at the Max Planck Institute of molecular Physiology in Dortmund. For me it was an honor to be part of an international research group at one of the best research institutions. Personally, I have learned a lot - but most important I have learned being an independent researcher with lots of opportunities. I always enjoyed his supportive attitude throughout my time in his group that include the continuous profitable discussions. I am thankful for his confidence in me and my competences working on complex research projects. Especially, I enjoyed working on such interesting and challenging projects on my own and in a great team.

Furthermore, I want to thank Prof. Dr. Daniel Summerer for reviewing my thesis and the continuous support throughout my whole PhD period as an active member of my TAC.

Moreover, I would like to thank the IMPRS-LM coordinators Christa Hornemann and Dr. Lucia Sironi for creating a great PhD graduate program at the institute. I especially enjoyed the workshops they have continuously organized, the chance of organizing an international student symposium and the continuous support and help. Additionally, I want to thank my whole TAC, Prof. Dr. Stefan Raunser, Prof. Dr. Daniel Summerer and Prof. Dr. Michael Ehrmann. I always appreciated the great ideas and supportive feedback from all members.

Thanks to the whole AG Raunser and department III of the institute for a friendly and motivating atmosphere in the lab and the office. A special thanks to the technical assistants, who always helped a lot in the daily lab work as well as Dr. Oliver Hofnagel and Dr. Daniel Prumbaum for the 24/7 support while data acquisitions. I would like to thank my former supervisor Dr. Arne Bothe for his continuous support especially in the beginning of my PhD period, in which he helped me in getting started. Next to him, I want to thank my office members for bearing me laughing, complaining and cursing especially when working with the terminal - it was great sharing the office with you. I personally appreciate the support of the whole membrane protein group during the time from which I specifically would like to thank for the support during lab work and thesis writing. A special thank you for you guys: Claudia Antoni, Patrick Günther, Dr. Philine Hagel, Dr. Franziska Leidreiter, Pascal Lill, Dr. Tobias Raisch and Dr. Sebastian Tacke - I will never forget working with you!

I would like to thank my family and friends for the constant motivation and support throughout this time. Especially my parents always supported me in being independent and following my own ideas and interests. My special thanks to my husband Fabian Weyers for his trust, personal support and encouragement – I could not wish for a better partner than you at my side.

List of Content

ACKNOWLEDGEMENTS.....	VI
LIST OF CONTENT	VII
LIST OF ABBREVIATIONS.....	XI
1 ABSTRACT	1
1.1 ABSTRACT	1
1.2 ZUSAMMENFASSUNG	2
2 INTRODUCTION.....	4
2.1 BIOLOGICAL MEMBRANES.....	4
2.1.1 <i>Architecture of biological membranes</i>	4
2.1.2 <i>Cholesterol in biological membranes</i>	7
2.1.3 <i>Proteins in biological membranes (membrane proteins)</i>	8
2.2 CHOLESTEROL AND ITS IMPACT IN SEVERE DISEASES.....	10
2.2.1 <i>Cholesterol biosynthesis</i>	12
2.2.2 <i>Cholesterol homeostasis</i>	13
2.2.3 <i>Statins</i>	16
2.3 SCAP.....	17
2.3.1 <i>Molecular features</i>	17
2.3.2 <i>Structural investigations</i>	19
2.3.3 <i>Synthetic antagonists</i>	23
2.4 PATCHED1	24
2.4.1 <i>Hedgehog signalling pathway</i>	24
2.4.2 <i>Structural investigations</i>	27
3 OBJECTIVE.....	31
4 MATERIALS AND INSTRUMENTATION.....	32
4.1 CHEMICALS	32
4.2 CONSUMABLES	34
4.3 ELECTRONIC DEVICES.....	35
4.4 NON-ELECTRONIC DEVICES.....	37
4.5 MEDIA	37
4.6 CHROMATOGRAPHY COLUMNS	37
4.7 KITSYSTEMS.....	37
4.8 ENZYMES AND ANTIBODIES.....	38
4.9 CELL LINES	38
4.10 PLASMIDS	39
5 METHODS	41

5.1	MOLECULAR BIOLOGY.....	41
5.1.1	<i>Agarose gel electrophoresis</i>	41
5.1.2	<i>Nucleic acid quantification</i>	41
5.1.3	<i>Restriction/Ligation cloning</i>	41
5.1.3.1	Polymerase chain reaction.....	41
5.1.3.2	Digestion of plasmid DNA	42
5.1.3.3	Digestion of amplified PCR products	42
5.1.3.4	DNA ligation.....	43
5.1.4	<i>Site-directed mutagenesis</i>	43
5.1.5	<i>Chemical transformation of E. coli</i>	43
5.1.6	<i>Electroporation of E. coli</i>	44
5.1.7	<i>Blue-white screen</i>	44
5.1.8	<i>Transfection of insect cells and BacMam virus production</i>	45
5.1.9	<i>Transfection of human cell lines</i>	45
5.1.10	<i>DNA sequencing</i>	46
5.2	CELL BIOLOGY	46
5.2.1	<i>Eukaryotic cell maintenance</i>	46
5.2.2	<i>Virus plaque assay</i>	47
5.3	BIOCHEMICAL METHODS.....	47
5.3.1	<i>Semi-denaturing SDS-PAGE</i>	47
5.3.2	<i>Western Blot</i>	48
5.3.3	<i>Heterologous expression of proteins in adherent mammalian cells</i>	49
5.3.4	<i>Detergent screen</i>	49
5.3.5	<i>Fluorescence-detection size-exclusion chromatography</i>	51
5.3.6	<i>Heterologous expression of proteins in suspension mammalian cells</i>	51
5.3.6.1	Heterologous expression of Scap and Patched1 constructs.....	51
5.3.6.2	Heterologous expression of Scap and SREBP complex.....	52
5.3.6.3	Heterologous expression of Scap, SREBP and Insig1 complex.....	52
5.3.7	<i>Cell lysis of mammalian cells</i>	53
5.3.8	<i>Membrane preparation</i>	53
5.3.9	<i>Solubilization</i>	53
5.3.10	<i>Affinity chromatography</i>	54
5.3.10.1	Immobilized metal ion affinity chromatography of poly-Histidine modified Proteins	54
5.3.10.2	Immunoaffinity chromatography of Flag modified proteins	55
5.3.10.3	Affinity chromatography of StrepII modified proteins.....	56
5.3.11	<i>Detergent exchange</i>	56
5.3.12	<i>Reconstitution into lipid nanodiscs</i>	57
5.3.13	<i>Reconstitution into amphipols</i>	58
5.3.14	<i>Size-exclusion chromatography</i>	58

5.4	BIOPHYSICAL METHODS.....	60
5.4.1	<i>Tandem mass spectrometry</i>	60
5.4.2	<i>Nano differential scanning fluorimetry</i>	61
5.4.3	<i>Fluorescence microscopy</i>	61
5.5	STRUCTURAL BIOLOGY METHODS.....	62
5.5.1	<i>Negative stain electron microscopy</i>	62
5.5.1.1	Sample preparation.....	62
5.5.1.2	Data collection	63
5.5.1.3	Single particle analysis and data processing.....	63
5.5.2	<i>Cryo-electron microscopy</i>	63
5.5.2.1	Sample preparation.....	64
5.5.2.2	Data acquisition	65
5.5.2.3	Single particle analysis.....	66
6	RESULTS AND DISCUSSION.....	68
6.1	PATCHED PROTEIN	68
6.1.1	<i>Homolog screen of Patched proteins</i>	68
6.1.2	<i>Truncated human Patched1 constructs</i>	73
6.1.3	<i>Heterologous expression and purification strategy of Patched1</i>	80
6.1.4	<i>Negative staining electron microscopy</i>	83
6.1.5	<i>Conclusion</i>	86
6.2	SCAP.....	90
6.2.1	<i>Validations for a stable Scap construct</i>	90
6.2.2	<i>Heterologous expression and purification strategy of porcine Scap</i>	97
6.2.3	<i>Scap reconstituted into lipid nanodiscs</i>	105
6.2.4	<i>Scap exchange into Digitonin or GDN</i>	108
6.2.1	<i>Conclusion</i>	116
7	REFERENCES	120
8	SUPPLEMENTARY INFORMATION	130
9	PUBLICATIONS AND CONFERENCE CONTRIBUTIONS.....	142

List of Abbreviations

2D	two-dimensional
3D	three-dimensional
AA	amino acid
AMP	adenosine mono phosphate
Anzergent 3-10	<i>n</i> -decyl-N,N-dimethyl-3-ammonio-1-propanesulfonate
Anzergent 3-12	<i>n</i> -dodecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate
<i>AOX1</i>	alcohol oxidase
Apo A	apolipoprotein A
ATP	adenosine 5'-triphosphate
BacMam	Baculovirus transduction of mammalian cells
BSA	bovine serum albumin
cAMP	cyclic adenosine mono phosphate
CHS	cholesteryl hemisuccinate
CMC	critical micelle concentration
CoA	coenzyme A
COPII	coat protein complex II
CPU	central processing unit
(Cryo-)EM	(cryo) electron microscopy
C-terminus	carboxy terminus
CTF	contrast transfer function
CV	column volume
Cymal-4/6	4/6-Cyclohexyl-1-heptyl- β -D-maltoside
ddH ₂ O	double distilled water
Dhh	desert hedgehog
DDM	<i>n</i> -dodecyl β -D-maltopyranoside
DED	direct electron detector
DM	<i>n</i> -decyl β -D-maltopyranoside
(ds)DNA	(double stranded) deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
DTT	dithiothreitol
(e)BFP	(enhanced) blue fluorescent protein
ECD	extracellular domain
ED	electron diffraction
EDTA	ethylenediaminetetraacetic acid
(e)GFP	(enhanced) green fluorescent protein
EM	electron microscopy
ER	endoplasmic reticulum

List of Abbreviations

FBS	fetal bovine serum
FPP	farnesyl pyrophosphate
(F)SEC	(fluorescence-assisted) size exclusion chromatography
FU	fluorescence unit
GDN	glyco-diosgenin
GGPP	geranyl-geranyl pyrophosphate
Gli2/3	glioma-associated oncogene homolog 2/3
GnTI	N-acetyl-glucosaminyl transferase I
GPR161	G-protein coupled receptor 161
GPU	graphical processing unit
HDL	high density lipoprotein
HABA	4'-hydroxyazobenzene-2-carboxylic acid
HEK	human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hh	hedgehog
His ₈	octahistidine
HMG-CoA	3-hydroxy-3-methylglutaryl coenzyme A
HMGCR	3-hydroxy-3-methylglutaryl coenzyme A reductase
HMGCS	β -hydroxy β -methylglutaryl-coenzyme A synthase
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
Ihh	Indian hedgehog
IMAC	immobilized metal ion affinity chromatography
IMP	integral membrane protein
Insig-1	insulin induced gene 1
IPTG	isopropyl β -D-1-thiogalactopyranoside
LDAO	lauryldimethylamine-N-oxide
LDL	low density lipoprotein
LiAc	lithium acetate
LMNG	lauryl maltose-neopentyl glycol
MOI	multiplicity of infection
MP	membrane protein
MS	mass spectrometry
MSP	membrane scaffold protein
MW	molecular weight
m/z	mass-to-charge ratio
nanoDSF	nano differential scanning fluorimetry
Ni ²⁺ -NTA	Nickel ²⁺ - nitrilotriacetic acid
Ni ²⁺ -IDA	Nickel ²⁺ - iminodiacetic acid

List of Abbreviations

NPC-1	Niemann-pick disease type C1
OD ₆₀₀	optical density at 600 nm
OG	n-octyl- β -D-glucopyranoside
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDE4	phosphodiesterase 4
pEG	plasmid Eric Gouaux
PKA	protein kinase A
PMSF	phenylmethyl sulfonylfluoride
POI	protein of interest
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
PTM	post-translational modification
QMS	quadrupole mass analyzer
RND	resistance-nodulation division
RT	room temperature
SAMS	statin associated muscle symptoms
Sap A	Saposin A
SDS	sodium dodecyl sulfate
Scap	SREBP cleavage-activating protein
SEC	size-exclusion chromatography
Shh	Sonic hedgehog
SMALPs	styrene maleic acid co-polymer lipid particles
Smo	Smoothened
SNR	signal-to-noise ratio
SOC	super optimal broth
SPA	scintillation proximity assay
SREBP	sterol regulatory element-binding protein
SSD	sterol sensing domain
TAE	Tris-acetate-EDTA
TCEP	Tris(2-carboxyethyl)phosphine
TEM	transmission electron microscopy
TFA	trifluoroacetic acid
T _m	melting temperature
Tris	tris(hydroxymethyl)aminomethane
VIPER	Validation of Individual Parameter Reproducibility
VLDL	very low density lipoprotein
X-FEG	high brightness field-emission gun

List of Abbreviations

X-Gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside
YFP	yellow fluorescent protein

1 Abstract

1.1 Abstract

Membrane proteins are essential for cell viability, as they are responsible for numerous regulatory processes like cell-cell communication, signal transduction or the transport of molecules across lipid bilayers. (1) Furthermore, mutations in their amino acid sequence or altered three-dimensional (3D) conformations can lead to severe diseases. Based on their significant role, membrane proteins represent the majority of all drug targets (more than 60%). (2) For a complete understanding of their function, the 3D structure of the protein is necessary and provides a first insight on a molecular level that furthermore enables the research for new efficient drugs. In this thesis, structural investigations of two polytopic membrane proteins, called Patched and sterol regulatory element binding protein- (SREBP-) cleavage activating protein (SCAP), were performed. Both membrane proteins bind cholesterol and contain the sterol sensing domain (SSD) at the transmembrane helix 2 to 6 which is highly conserved upon membrane proteins involved in cholesterol metabolism. (3)

The first membrane protein, Patched, is one of the key proteins in Hedgehog (Hh) signalling, which is one of the main pathways in tissue patterning events during embryonic development and stem cell biology. (4) Patched suppresses Smoothed (Smo) activity by an unknown mechanism that leads to an inactivation of the Hh signalling cascade which is lifted upon the complex formation of Patched and Hh proteins. Here, I show the expression and purification of a C-terminally truncated Patched1 construct with a yield of 0.01 mg protein per liter of mammalian cell culture which was sufficient and suitable for structural analysis by transmission electron microscopy (EM). However, before I had the chance to structurally characterize Patched1, several cryo-EM structures of mammalian Patched1 alone and in complex with Sonic Hh (Shh) were reported by other groups in the last two years. (5-10) The membrane protein Scap regulates the cholesterol homeostasis via escorting the transcriptional regulator SREBP from the endoplasmic reticulum (ER) to the Golgi in a sterol dependent manner. (11) At increased cholesterol concentrations, the sterol is bound to Scap and prevents the Scap-SREBP complex movement to the Golgi, as well as the proteolytic processing of SREBP. Although the different complex states of Scap are of great interest, the structure of the full-length Scap is still remaining. In this thesis, I present the first cryo-EM reconstruction of the monomeric full-length mammalian Scap at sub-nanometer resolution. A previously determined crystal structure of the C-terminal WD40 domain from fission yeast Scap could be unambiguously docked into this reconstruction. I therefore predict a similar structure of the mammalian WD40 repeat domain, which forms an eight-bladed β -propeller. These data provide a first insight into the architecture of the full-length mammalian membrane protein Scap and can be very valuable for structure analysis of Scap complex formations.

1.2 Zusammenfassung

Membranproteine sind für das Überleben von Zellen essentiell, da sie für die Regulation von zahlreichen Prozessen hauptverantwortlich sind. So regulieren sie beispielsweise verschiedenste Signalweiterleitungen, die Kommunikation zwischen verschiedenen Zellen oder aber auch den Transport von Molekülen über die Lipiddoppelschicht. (1) Damit einhergehend können vereinzelte Veränderungen, wie zum Beispiel Mutationen in der Aminosäuresequenz, zu schweren Erkrankungen führen. Mit über 60% stellen Membranproteine die größte Proteingruppe dar, die heute von Medikamenten adressiert werden, was wiederum die Relevanz der Strukturaufklärung von Membranproteinen unterstreicht. (2) Denn die Strukturaufklärung von Proteinen ist nicht nur für das Verständnis der Funktionen von Proteinen erforderlich, sondern liefert zudem einen ersten Überblick über deren Aufbau auf molekularer Ebene um die Forschung von neuen, effizienteren Medikamenten zu erlauben. In dieser Arbeit wurden zwei Membranproteine, Patched und Scap, strukturell untersucht. Beide Membranproteine sind in der Lage Cholesterol zu binden und besitzen eine SSD, die zwischen den Transmembranhelices 2 und 6 lokalisiert ist. Die SSD ist eine hochkonservierte Domäne, die in vielen Membranproteinen vorzufinden ist und an dem Cholesterolmetabolismus beteiligt sind. (3)

Das erste Membranprotein, Patched, ist eines der Schlüsselproteine im Hh Signalweg, der unter anderem für die korrekte Ausbildung von Stammzellen verantwortlich gemacht wird. (4) Patched unterdrückt die Aktivität von Smo durch einen noch unbekanntem Mechanismus, welcher weiterführend zu einer Inaktivierung der Hh Signalkaskade führt. Sowohl die Hh Signalkaskade, als auch Smo können erst wieder aktiviert werden, wenn Patched durch seinen Liganden Hh inaktiviert vorliegt. In dieser Arbeit konnte ich die Expression und Aufreinigung eines C-terminal verkleinerten Patched Konstruktes mit einer Ausbeute von 0,01 mg Protein pro Liter Zellkultur aufzeigen, welche für weiterführende strukturelle Analysen mittels Transmissionselektronenmikroskopie genutzt werden konnten. Bevor ich das Membranprotein Patched näher strukturell analysieren konnte, wurden mehrere hochauflösende 3D Rekonstruktionen von Patched alleine und im Komplex mit Shh gelöst und veröffentlicht. (5-10)

Das Membranprotein Scap hingegen reguliert das Gleichgewicht von Cholesterol, in dem es einen Komplex mit einem transkriptionsregulierenden Protein, SREBP, bildet und dieses bei geringer Sterolkonzentration von dem ER zum Golgi geleitet. Am ER wird der Transkriptionsfaktor von SREBP proteolytisch abgespalten, sodass dieser die Expression von Genen aktiviert, die die Sterolsynthese regulieren. (11)

Obwohl die verschiedenen Zustände von Scap von großem Interesse sind, fehlt bis heute die volllängen Struktur des Membranproteins um auf molekularer Ebene die Signalweiterleitung zu verstehen. In dieser Arbeit konnte ich somit die erste cryo-EM Rekonstruktion des monomeren, volllängen Scap in einer Sub-Nanometer Auflösung aufklären. Die zuvor veröffentlichte Struktur einer löslichen WD40 Domäne des Hefepilzhomologs konnte ich zudem erfolgreich in einen Bereich der Rekonstruktion

docken, sodass ich davon ausgehe, dass das Homolog aus *S. scrofa* eine ähnliche 3D-Struktur aufweist. Die hier aufgeführten Ergebnisse können weiterhin in Bezug auf eine hochauflösende 3D Struktur von Scap alleine, sowie im Proteinkomplex mit seinen Bindungspartnern verwendet werden und liefern somit nützliche Informationen in Betracht auf den Aufbau des Membranproteins.

2 Introduction

2.1 Biological membranes

One of the most fundamental structures that enable the function of cells are membranes. Biological membranes protect, stabilize and organize every single cell. They function as a spatial boundary to the external environment, where they build a barrier to every other ion or molecule either to protect them from the periphery or keep them separated. In addition to that, lipid membranes are selectively permeable and thereby able to control the import and export of soluble molecules across the membrane. This remarkable feature allows a communication with the periphery including other cells. Next to the plasma membrane, biological membranes exist within eukaryotic cells. Here, they build separated spaces, called compartments, which are the basis of the life of eukaryotic cells. The compartmentalisation inside a cell allows a large variety of metabolic reactions as biological membranes can control the import and export of ions or larger molecules.

The different functions of membranes represent their diversity which is further described in this section. (12)

2.1.1 Architecture of biological membranes

Biological membranes are composed of lipids, proteins and carbohydrates, whereby lipids, like phospholipids, glycolipids or cholesterol, represent the largest proportion. (13)

Lipids are amphiphilic molecules and consist of a hydrophilic (polar) head and a hydrophobic (non-polar) hydrocarbon tail. The hydrophobic tails are mainly composed of two long hydrocarbon chains which are mainly fatty acid rests. These rests can be either saturated (linear tail) or unsaturated (non-linear tail). In the latter status of a fatty acid, two or more carbon atoms are bound via a double bond. The polar head groups of lipids are composed of a high variation of different molecules, which enables a diverse range of lipids, like phospholipids that can be often found in biological membranes of mammals. (12) Its characteristic feature is the polar head group, which is composed of a polar molecule, like ethanolamine, choline or inositol, which is bound to a glycerol molecule via phosphate. The polar head and the non-polar tails are connected via two ester bonds between the fatty acid carboxyl groups and two hydroxy groups of the glycerol (see Figure 2.1). Other lipids, like sphingolipids contain the sphingosine back bone and can have an additional fatty acid at their hydrophobic tail. The hydrophobic tail is linked via the sphingosine backbone to the polar group which is composed of similar polar head groups that come across phospholipids (see Figure 2.1). In contrast to phospholipids, glycolipids contain several additional carbohydrates that are bound to the lipids polar head group.

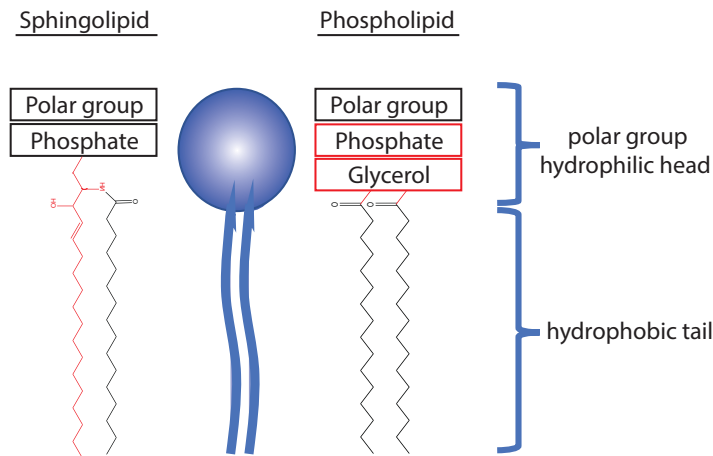


Figure 2.1: Setup of lipids that can be found in biological membranes.

Lipids are amphiphilic molecules composed of a hydrophilic head, the polar group, and a hydrophobic tail, the nonpolar tail. Sphingolipids (left) and Phospholipids (right) are two major molecules of mammalian biological membranes. The polar group is connected to a phosphate that is either directly bound to the hydrophobic tail (sphingolipids) or linked via a glycerol molecule to the hydrophobic tail (phospholipids). The individual backbones of a sphingolipid and a phospholipid are highlighted in red.

Amphiphilic molecules, such as lipids, spontaneously form larger, highly symmetrical aggregates when surrounded by water molecules. The formation of these symmetrical aggregates is based in the lipids structure. On the basis of their structure, lipids form lipid bilayers due to the hydrophobic effect. In lipid bilayers, the polar head is exposed to the surrounding aqueous solution whereas the nonpolar tail is localized inside the bilayer. Other molecules like membrane proteins or cholesterol are embedded in the lipid bilayer either via their hydrophobic areas (steroids and integral membrane proteins) or peripherally bound to the lipids (peripheral membrane proteins). In 1972, Singer and Nicholson schematically described these characteristics of a biological membrane in the 'fluid mosaic model' (see Figure 2.2). (14)

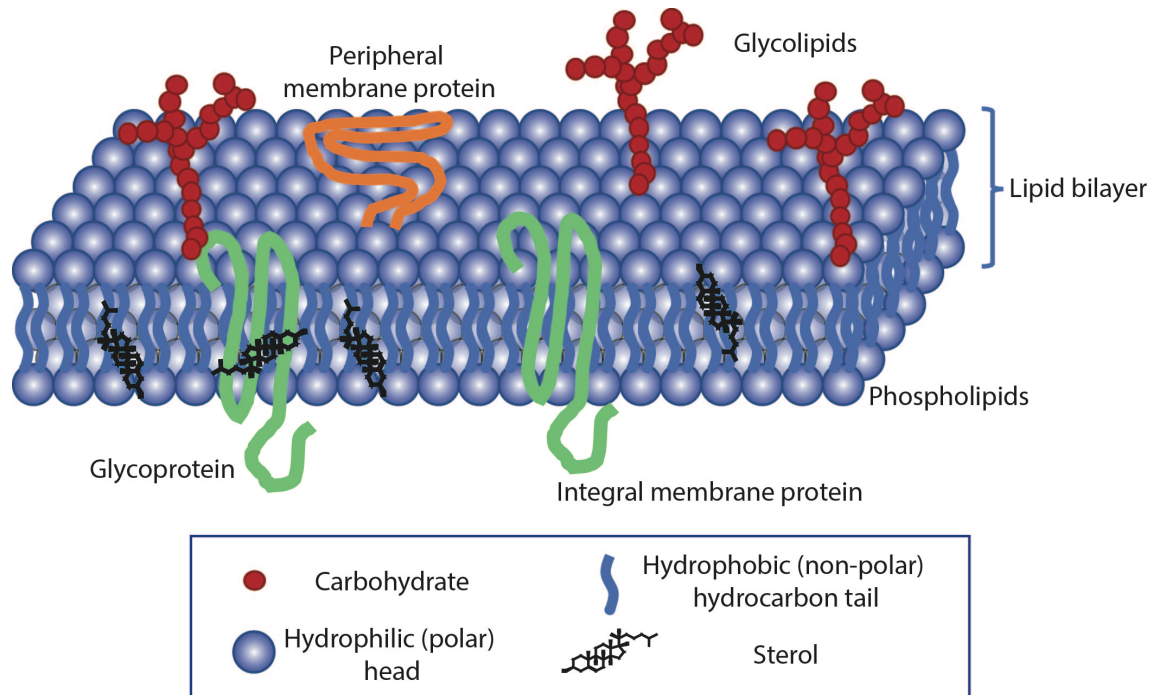


Figure 2.2: Schematic representation of a biological membrane according to the fluid mosaic model first described by Singer and Nicholson in 1972. (14)

In the fluid mosaic model, a biological membrane is mainly composed of lipids. Lipids contain a hydrophilic (polar) head and a hydrophobic (non-polar) hydrocarbon tail ranging into the inner leaflet of a spontaneously formed lipid bilayer when lipids are surrounded of an aqueous solution. Biological membranes contain different kinds of proteins, like integral or peripheral membrane proteins. Furthermore, sterols are also part of biological membranes as well as carbohydrates that are linked to proteins (glycoproteins) or lipids (glycolipids).

In addition to the complex composition of biological membranes as described before, one of the main assumptions of the fluid mosaic model, is the mobility of every single molecule. The lipid molecules allow lateral as well as transversal diffusion of molecules. The single molecules of biological membranes contain a relatively high lateral mobility. On average one lipid molecule moves 2 μm per second. (14, 15) The ability of lateral diffusion was therefore captured in the first term ‘fluid’. The variety of different lipids that are part of biological membranes was furthermore described with the second term ‘mosaic’. (12) The steroid cholesterol is also part of biological membranes in the fluid mosaic model and stabilises membranes via its rigid characteristics. Cholesterol interacts with the hydrocarbon chains of the other lipids via hydrophobic interactions which are then limited in their rotation. Due to the limited rotation of the long hydrocarbon chains the lipid bilayer is more rigid and less dynamic. (14)

In 1997, Simons and Ikonen identified special sterol rich compartments in eukaryotic cells and called them ‘lipid rafts’. (16) They showed that these functional regions mainly serve as signalling components. Initially, lipid rafts were introduced as “glycolipid rich apical membranes of epithelial cells” by Simons and van Meer. (17) These regions contain higher concentrations of sterols and a different lipid and protein composition compared to other membrane regions. Today, ‘lipid rafts’ are known as “heterogenous, highly dynamic, sterol and sphingolipid enriched domains” that in contrast

to the fluid mosaic model demonstrate the heterogeneity of biological membranes. The assembly of membrane components to complexes is based on either strong binding interactions or lower interactions with low on/off rates and highlights the dynamic and heterogeneity of the lipid bilayers. (18, 19)

2.1.2 Cholesterol in biological membranes

Chemically, cholesterol is a derivative of the steroid backbone gonane. At position 3 of the steroid backbone it contains a hydroxy group and an aliphatic chain at position 17 (see Figure 2.3). Based on its structural features, cholesterol is a lipid and embedded in biological membranes. The sterol is bound to the hydrophobic fatty acid tails with its hydrocarbon chain and steroid backbone via hydrophobic interactions. The hydrophilic hydroxyl group faces to the outside of the lipid bilayer and interacts with the hydrophilic head groups of the other lipids and the surrounding aqueous medium (see Figure 2.3).

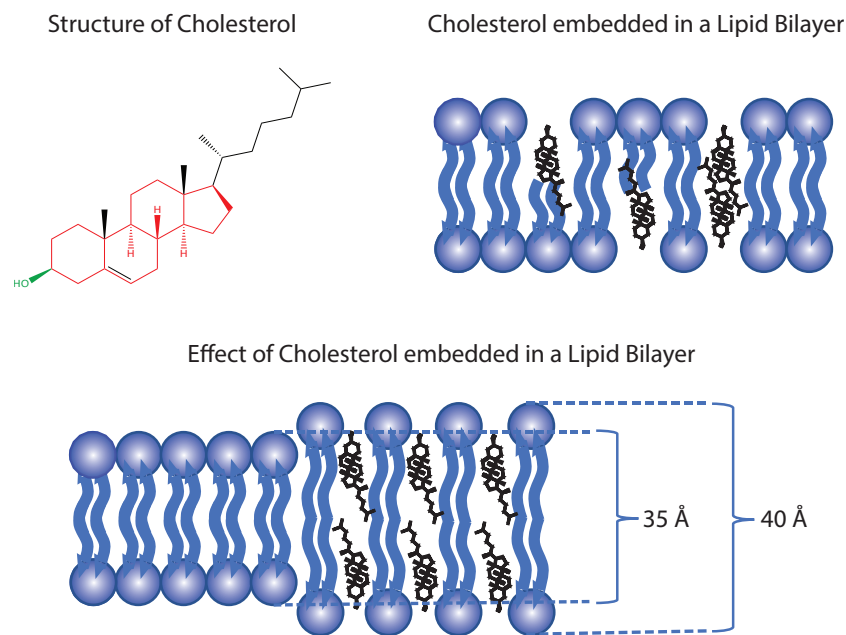


Figure 2.3: Structure of cholesterol and its effect in biological membranes.

Cholesterol is chemically derived from the steroid core gonane (red) and carries an additional hydroxy group at position 3 (green) as well as a hydrocarbon chain at position 17 (black). When embedded in a membrane, the polar hydroxyl group faces towards the outside, while the hydrocarbon tail points towards the hydrophobic core of the membrane. Cholesterol is accumulated alongside the hydrophobic fatty acid tails of the lipids. The fatty acid tails are elongated when cholesterol is accumulated and cause a 5 Å thicker phospholipid bilayer.

Cholesterol is one of the major components of biological membranes in animals and acts as a fluidity buffer. In presence of cholesterol lipid bilayers are more fluid at low temperatures and less fluid at higher temperatures compared to membranes without cholesterol. (20) Particularly the fluidity is one of the key factors of biological membranes because of its physiological relevance. Depending on the

temperature, pressure and the percentage of lipid molecules, lipid bilayers undergo a variety of phase transitions. The lipids in a model bilayer are coordinated in a sub-gel phase at low temperatures (below room temperature) and transfer into the gel phase around room temperature. At increasing temperatures, the lipids are first coordinated in an undulated gel phase and switch to the fluidic phase. The main difference of these phases is the free space of the long hydrocarbon chains that are either well-ordered in all-trans conformations at limited space (gel phase and undulated gel phase) or disordered in many different gauche conformations (fluid phase). The lipids in the sub-gel phase are able to rotate and are coordinated orthorhombic. The optimal state of a biological membrane for a complete physiological function is the fluidic phase. In this state, the lateral diffusion coefficient and furthermore, the movement of the molecules is the highest. The cholesterol concentration in biological membranes ranges from 4 – 50 mol% and is dependent on the environment and special need. At these molar ratios, the phase transition (T_m) from the gel phase to the fluid phase is shifted and from 50 mol% on not detectable anymore. At low temperatures ($<T_m$) cholesterol changes the well-ordered symmetry of the gel phase and in parallel increases the rigidity of the hydrocarbon chains at higher temperatures ($>T_m$) of the fluidic phase. (20) The cells of animals can be exposed to extreme temperatures and pressures for example when they live in the deep-sea. With the integration of additional cholesterol these organisms are able to keep the fluidic state of their membranes.

2.1.3 Proteins in biological membranes (membrane proteins)

In addition to lipids, membrane proteins play a crucial role in biological membranes. Lipid membranes pose a large energy barrier for transport of hydrophilic and large compounds from one side to the other, and furthermore efficiently block signalling. To overcome this restriction, different classes of membrane proteins serve as channels and transporters for molecules and relay signalling information across the membrane. (21)

Membrane proteins are divided into two classes depending on the nature of their association with the membrane. First, peripheral membrane proteins are not integrated into the lipid bilayer, but instead localized to the surface of the membrane. This class comprises all proteins that are either covalently attached to lipids, interact with the membrane surface, or bind to other membrane proteins. In contrast, integral membrane proteins insert at least partially into the membrane bilayer. They are further subdivided by their topology. Monotopic membrane proteins are inserted into the membrane but do not completely pass the bilayer, while bitopic (single-span) and polytopic (multi-span) membrane proteins stretch across the membrane by one or multiple α -helices, respectively (see Figure 2.4). The transmembrane helices of such proteins are enriched in hydrophobic amino acids like leucine or isoleucine. Those residues are exposed to the hydrophobic core of the lipid bilayer and interact with the lipid's fatty acid tails via van der Waals forces. (21, 22)

Additionally to membrane proteins containing transmembrane helices, another class of proteins spans the membrane by using β -barrels. For example, bacterial porins often features multiple β -sheets that are highly ordered in a cylindrical shape that forms β -barrels. (13)

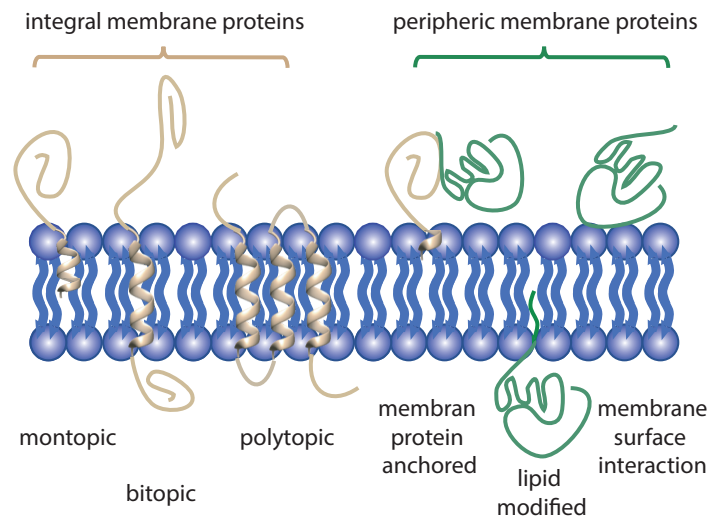


Figure 2.4: The integral and peripheral types of membrane proteins.

In total, there are six different types of membrane proteins that are classified into two main classes. First, the integral membrane proteins (highlighted in beige), containing all proteins that are embedded into the lipid bilayer in a monotopic, bitopic or polytopic manner, respectively. Monotopic proteins are incorporated via a hydrophobic area that is inserted into the membrane but does not completely span the bilayer, bitopic membrane proteins contain one single-span region and the polytopic membrane proteins feature several lipid bilayer spanning regions (multiple-span). The second class, named peripheral membrane proteins (highlighted in green), describes all proteins that are anchored to the membrane without penetrating it. Peripheral membrane proteins can be anchored to other membrane proteins, covalently attached to a lipid (lipid modified) or interacting with the lipid bilayer (membrane surface interaction). (21)

The insertion of membrane proteins into the lipid bilayer is dependent on the membrane protein class. Peripheral membrane proteins are mainly post-translational modified and either attached to the membrane or inserted by a hydrophobic anchor (see Figure 2.4). In contrast to that, most integral membrane proteins are co-translational inserted into the lipid bilayer by a special translocation pathway (see Figure 2.5). The nascent protein is translated at the ribosome and contains a signal sequence which can be identified and recognized by a protein named signal recognition particle (SRP). First, the SRP binds to the ribosome and the signal sequence that subsequently binds to its receptor. The SRP receptor is localized in the ER membrane which results in a movement of the ribosome-nascent protein-SRP protein complex to the ER. Here, the nascent protein gets inserted into the ER membrane by the translocon that is bound to the SRP receptor. The translated nascent peptide passes the translocon, enters into the ER and gets inserted into the ER membrane. In a final step, a signal peptidase removes the signal sequence if necessary. Additional ER membrane localized proteins allow a correct folded transmembrane protein. (23)

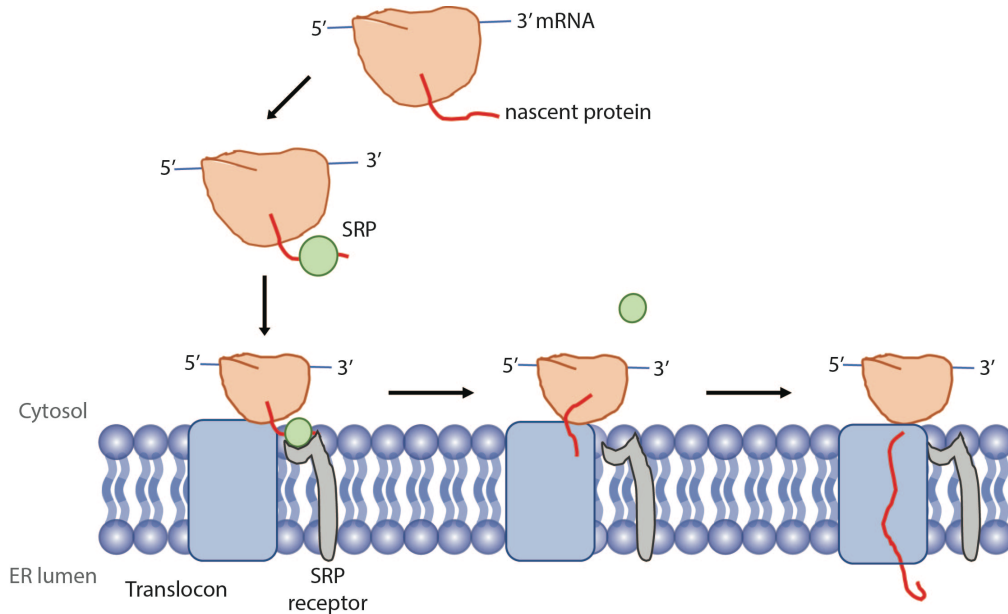


Figure 2.5: Co-translational membrane insertion of integral membrane proteins adapted from Rodwell *et al.* (24)

Most of the integral membrane proteins get inserted via a co-translational mechanism into the ER membrane. The nascent membrane protein contains a N-terminal signal sequence that the SRP (signal recognition particle) directly identifies and binds to. Its receptor, the SRP receptor, is localized at the ER membrane and brings the nascent protein-ribosome-SRP protein complex into close proximity with the translocon. The translocon enables the unfolded, nascent protein, to cross the ER membrane and enters the ER lumen. The translocon is a protein complex composed of several proteins that furthermore enables the membrane proteins correct secondary structure and insertion into the lipid bilayer. (24)

In 2007, a highly conserved and essential cytosolic ATPase, called TRC40, was identified which recognises membrane proteins containing a transmembrane region that is not long enough for a proper recognition by the SRP. The TRC40 protein especially identifies transmembrane domains that are located at the C-terminal regions of the nascent proteins. The TRC40 protein binds its receptor that is localized in the ER membrane. (25) This receptor gets stimulated by the ATPase activity of TRC40 and the transmembrane protein inserts into the membrane after TRC40 release. Nevertheless, the ways of membrane insertion, either co-translational or post-translational are not competitive nor interfering with each other. (23)

2.2 Cholesterol and its impact in severe diseases

In chapter 2.1.2. the structural features of cholesterol and its effect in biological membranes were described in detail. This chapter focuses on its additional features, like how it is related to severe diseases, the cholesterol biosynthesis, cholesterol homeostasis and available therapies addressing cholesterol related diseases.

In mammals, approximately 80% of the cholesterol is synthesised in the liver, while the remaining 20% are taken up with nutrition. Notably, the endoplasmic reticulum (ER) membranes of hepatocytes, which is the primary location of cholesterol biosynthesis, contain up to 5 mol% cholesterol. (26) At the same time, 25% of the total amount of cholesterol is localized at the neurons in order to properly form and maintain axons, dendrites and synaptic connections, thus representing the second large reservoir of cholesterol in the body. (27-33)

In addition to the function as a major component of biological membranes where it serves as the regulator of rigidity, fluidity and permeability, cholesterol is the precursor of all steroid hormones, several vitamins and bile acids. Furthermore, cholesterol serves as an important cofactor of several membrane proteins that are for example sterol transport proteins. In two known cases, cholesterol is covalently bound through an ester bond to the membrane proteins, either localized at the total C-terminus (sonic hedgehog) or to an aspartic acid (Smoothed receptor), both having an important function in the Hh signalling cascade. The cholesterylation of Hh mainly provokes its spatial distribution, whereas the lipidation of Smoothed (Smo) might be an additional event of Hh signal transduction as the signalling is inhibited by Patched1 and stimulated by Hh. (34-36)

Due to its immense importance as a membrane building block and a precursor for a large range of hormones and further biomolecules, cholesterol levels have to be tightly controlled. A dysregulation of the cholesterol homeostasis can lead to cardiovascular diseases, several congenital human diseases and many types of cancer. (37-39) The primary cause of many of these diseases are rising cholesterol concentrations in the blood, causing an increased blood pressure. Diseases like hyperlipidemia or hypercholesterolemia describe such disfunctions that are the main cause of coronary artery diseases. Cardiovascular diseases are the leading causes of death globally in the last 20 years. (40, 41)

Atherosclerosis is a disease that arises by an inflammation of the walls of arteries. In 1910, Adolf Windaus already showed a 20-fold increased amount of cholesterol in these regions compared to healthy aortas. (42) The course of this disease starts when blood vessels contain deposits of fats that later on evolve into atherosclerotic plaque. The process of atherosclerosis is mainly characterized by three stages: the formation of fatty streaks, the atheroma formation and last atherosclerotic plaque formations.

A first hint of atherosclerosis formation are fatty acid streaks (see Figure 2.6). These streaks arise due to an increase of lipoproteins in the arteries. The low-density lipoproteins (LDL) are cholesterol rich and usually balanced in their presence in the plasma and intracellular region. However, at this stadium LDL is accumulated inside the vascular intima, and this accumulation furthermore leads to an oxidation process of these particles. In their oxidized state, the particles (LDL) activate endothelial cells. Finally, leukocytes get activated that have an early state impact in atherosclerosis. The accumulation of adhesion molecules in the intima leads to an immigration of monocytes which differentiate to macrophages. Macrophages are able to absorb cholesterol and differentiate to foam cells that die due to apoptosis. The apoptosis of foam cells leads to a formation of a necrotic nucleus from which arises atheroma, that describes the mature atherosclerosis plaque formation. Ultimately, the massive injuries in this region and the growth of plaque lead to the formation of a thrombus. (43)

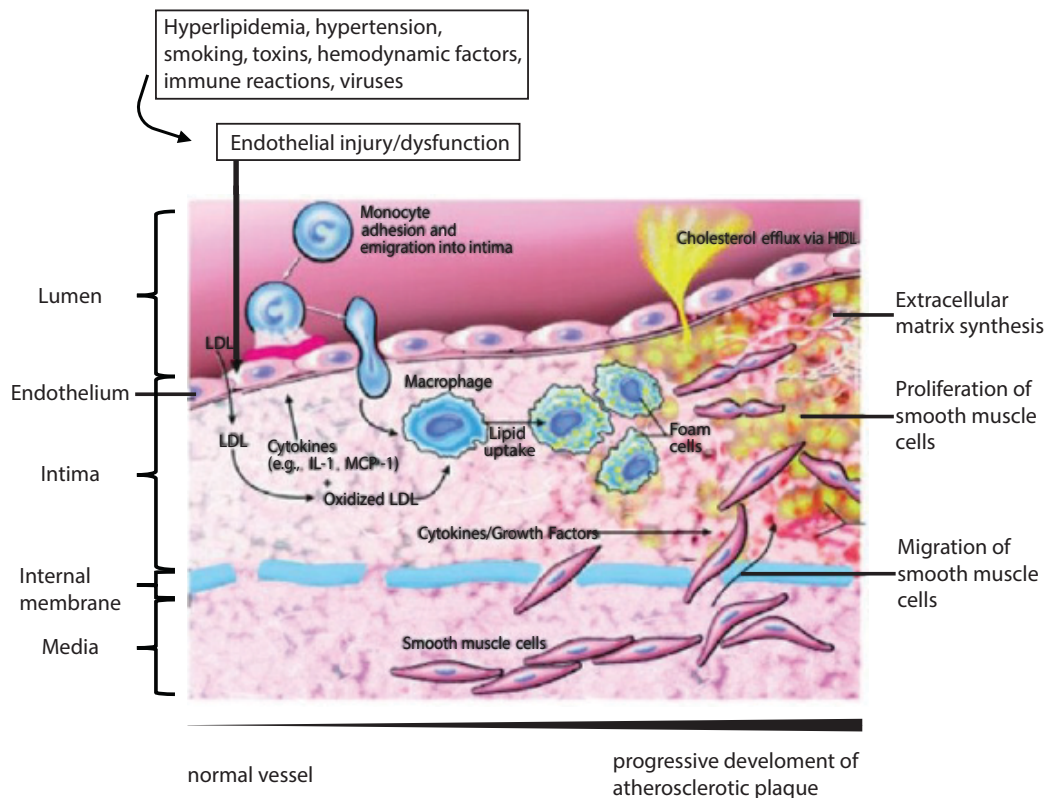


Figure 2.6: Representation of the progressive development of atherosclerotic plaque by M. Rafieian-Kopaei, *et al.* (43)

The initial factors of an endothelial injury or dysfunction can arise from hyperlipidemia, hypertension, smoking, toxins, hemodynamic factors, different immune reactions or viruses. At this stage, the formation of fatty streaks starts by an increasing amount of LDL (low-density lipoproteins) containing cholesterol inside the intima that usually is well balanced. In addition, adhesion molecules accumulate inside the intima that furthermore lead to the formation of macrophages in this area. Macrophages absorb the cholesterol and differentiate to foam cells which rupture due to apoptosis. From this stage on, atherosclerotic plaque formation is set. (43)

2.2.1 Cholesterol biosynthesis

As mentioned above, nearly 80% of the cholesterol in mammalian cells is synthesized at the ER of liver cells. To prevent the severe diseases occurring at an imbalanced cholesterol homeostasis (see chapter 2.2), the concentration of cholesterol needs to be highly controlled. Several checkpoints exist in mammalian cells that regulate the uptake of cholesterol and the *de novo* biosynthesis, as well as the transport inside the body and its degradation.

In total, the cholesterol biosynthesis contains 10 to 12 reactions to synthesize the triterpene squalene, which cyclized to lanosterol and additional 18 steps to convert lanosterol to cholesterol. The initial reactant is acetyl-coenzyme A (Acetyl-CoA) that stemmed from β -oxidation of fatty acids, amino acid degradation and glycolysis. The second reaction is the synthesis of mevalonic acid from β -hydroxy β -methylglutaryl-coenzyme A (HMG-CoA) that is enzymatically controlled by the HMG-CoA reductase

(HMGCR) (see Figure 2.7). (44) This reaction is the rate limiting step in the biosynthesis of cholesterol and therefore one of the most important checkpoints of cholesterol homeostasis. (45)

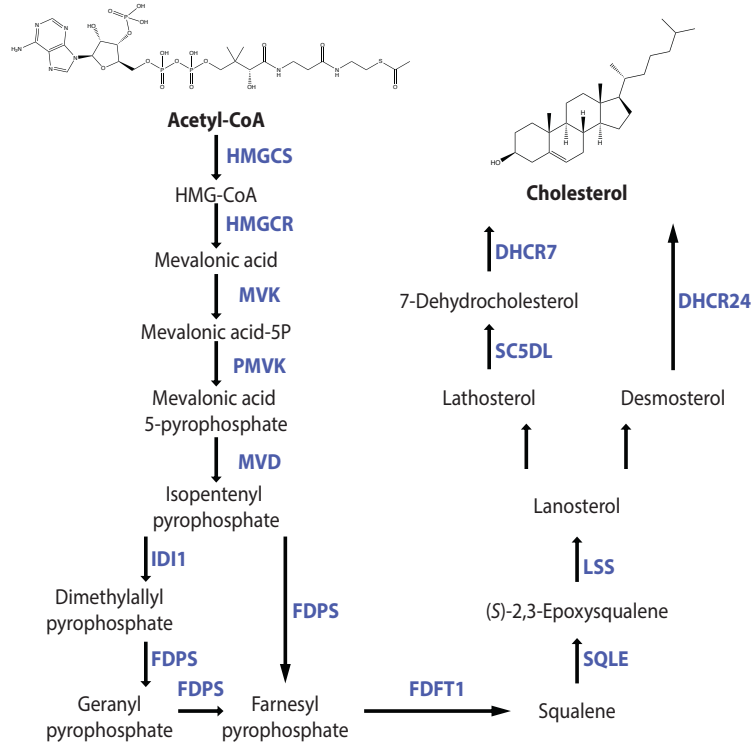


Figure 2.7: Biosynthesis pathway of cholesterol.

The reactant of the *de novo* synthesis is Acetyl-Coenzyme A (Acetyl-CoA) that gets acetylated to β -hydroxy β -methylglutaryl-coenzyme A (HMG-CoA) by the HMG-CoA synthase (HMGCS). Subsequently, the HMG-CoA gets reduced to mevalonic acid which is described as the rate limiting step of the pathway and enzymatically controlled by the HMG-CoA reductase (HMGCR). From the intermediate, mevalonic acid, there are additional 9 to 12 reactions necessary until the final product cholesterol is synthesized. All enzymes involved in this pathway are highlighted in blue.

2.2.2 Cholesterol homeostasis

In order to keep the cholesterol synthesis balanced, the regulation of the enzymatic activity of the HMGCR is highly regulated amongst mammals. For example, the reductase can be regulated by adenosine monophosphate (AMP) kinases. Here, the AMP kinase phosphorylates the C-terminal Ser-871 which leads to a significantly decreased enzymatic activity of the HMGCR. (45) Another regulation mechanism of the HMGCR activity is a feedback loop initiated by high concentrations of downstream products of the cholesterol biosynthesis pathway. The insulin induced gene 1 protein (Insig-1), localized at the ER membrane, binds in a lanosterol- and 24,25-dihydrolanosterol dependent manner to the reductase. This binding event triggers a ubiquitination cascade, ultimately leading to the proteasomal degradation of HMGCR and thus a downregulation of cholesterol synthesis. (46)

Another central regulation mechanism of cholesterol homeostasis is the sterol regulatory element-binding protein- (SREBP-) signalling pathway, whose importance is underlined by the Nobel prize in 1985. The prize was awarded to M. S. Brown and J. L. Goldstein for its discovery of this pathway and the identification of all involved proteins. The three involved membrane proteins are localized in the outer ER membrane where initially a complex is formed between SREBP and the SREBP cleavage-activating protein (Scap). Both proteins bind via their C-terminal regions independently of the sterol concentration. (5)

The SREBP-Scap dimer can be subjected to coat protein II- (COPII-) mediated trafficking in a manner that is controlled by the sterol concentration in the ER membrane: At low concentrations (below 5 mol%) of cholesterol and 25-hydroxy cholesterol, Scap is in an open conformation which allows the binding of the COPII at a cytosolic loop region consisting of six amino acids (MELADL) (see Figure 2.8, A). This allows a COPII vesicular trafficking from the ER membrane to the Golgi apparatus membrane. At the Golgi membrane, the C-terminal region is processed by the serine endopeptidase S1P and subsequently, the transcriptionally active N-terminus is cleaved off by the zinc metalloprotease S2P. The released soluble N-terminal domain of SREBP enters the nucleus where it activates the transcription of genes encoding for several proteins and enzymes that are necessary for cholesterol *de novo* biosynthesis (HMGCR; see chapter 2.2.1) and the import of cholesterol into the cells (LDL) (see Figure 2.8, A). (37)

At high sterol concentrations in the outer ER membrane from at least 5 mol%, the SREBP pathway is inhibited by a two-step mechanism (see Figure 2.8 B): First, Scap binds sterols, that causes a conformational change. Second, Insig-1 binds to this sterol-bound Scap/SREBP complex. As a consequence, the MELADL sequence is in an inaccessible conformation and cannot be recognized by COPII. Thus, the Scap/SREBP complex and especially the transcription factor domain of SREBP is retained in the ER membrane, preventing the upregulation of cholesterol biosynthesis and its uptake. (see Figure 2.8, B). (5)

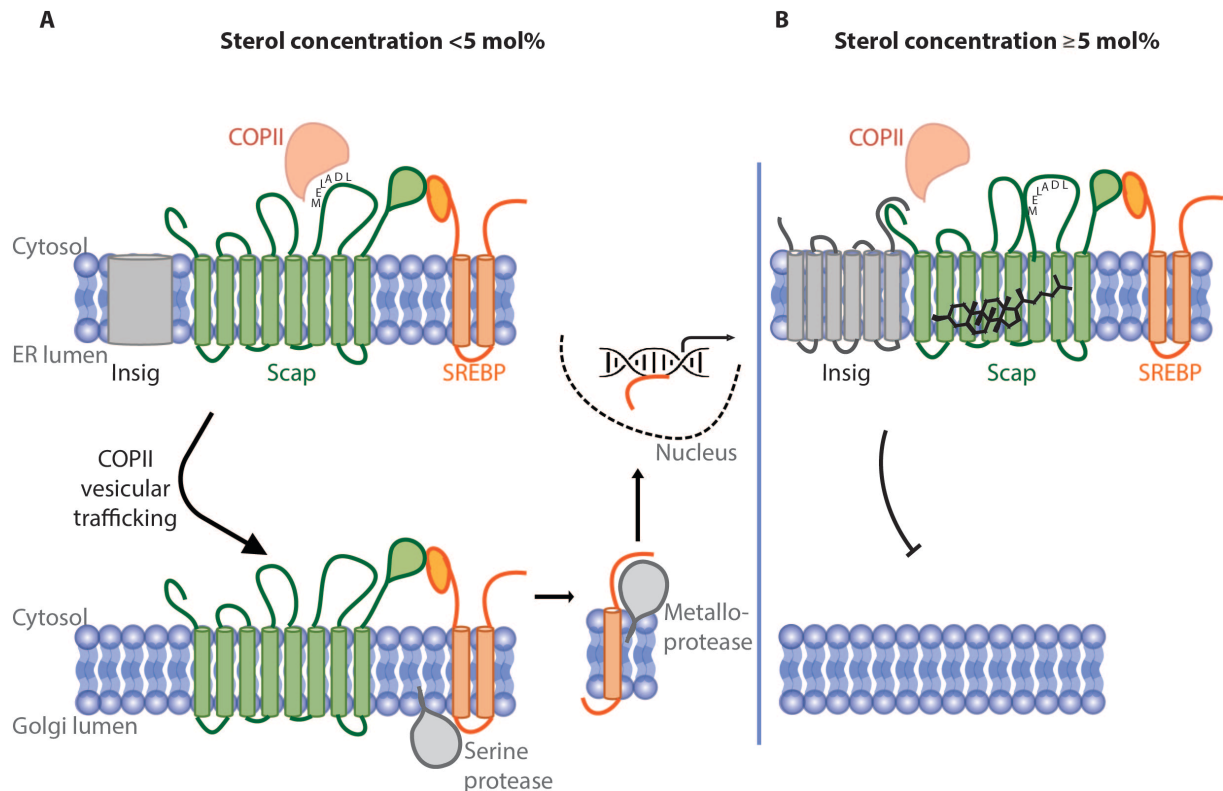


Figure 2.8: Schematic representation of the SREBP signalling pathway adapted from M. S. Brown and J. L. Goldstein. (5)

(A) Signalling cascade at low sterol concentrations below 5 mol% at the ER membrane. SREBP (orange) and Scap (green) form a heterocomplex with the COPII protein complex (red) that trafficks Scap and SREBP into the Golgi apparatus membrane. SREBP gets proteolytically processed by two endopeptidases (grey) and the soluble N-terminus enters the nucleus (red). In the nucleus, the transcription of genes that are encoded for proteins involved in the biosynthesis of cholesterol is activated. (B) SREBP pathway at higher sterol concentrations (≥ 5 mol%). Sterols (black) bind to Scap (green) and changes its conformation. These conformational changes induce a complex formation of Insig/Scap/SREBP and an inaccessible MELADL sequence. The complex recruitment to the Golgi apparatus membrane is suppressed and cholesterol biosynthesis is downregulated. (5)

In addition to the regulation of the cholesterol biosynthesis, the amount of cellular cholesterol is regulated by the absorption of cholesterol from the plasma via LDL as mentioned before. The 20% of cholesterol that is taken up with nutrition gets transported into the periphery by the absorption into the intestine and its solubilization in micelles. Here, the sterol is packed into lipoprotein particles, called chylomicrons, thus enabling the transport of the cholesterol into hepatocytes by binding to apolipoprotein E receptors. From these receptors, cholesterol can be either directly used for bile acid synthesis, esterified, used for secretion into circulation or absorbed by very low density lipoproteins (VLDL). The cholesterol containing VLDL can be transformed into LDL particles that furthermore transport cholesterol into the periphery. Via receptor-mediated endocytosis, cholesterol is embedded in membranes of cells that require the sterol. (47) By an accumulation of cholesterol-containing LDL particles in the blood plasma, the cells express its receptor preferentially. The LDL particles at the LDL receptor are absorbed into cells via endocytosis and get digested or released to the periphery where the cholesterol undergoes an esterification process. (48) The amount of LDL containing cholesterol in

the plasma is furthermore controlled by the return to the liver by high density lipoproteins (HDL) particles. (44)

2.2.3 Statins

Statins are synthetic small-molecule drugs which are often used as a treatment of an imbalanced (upregulated) cholesterol homeostasis because they lower the concentration of cholesterol-containing LDL in the plasma. Mechanistically, they bind into the active site of the HMGCR and thereby competitively inhibit the conversion of HMG-CoA to mevalonate (see Figure 2.7). Consequently, cholesterol biosynthesis is stopped and the concentration of cholesterol in the plasma is reduced. Most statins contain rigid and hydrophobic parts which are structurally reminiscent of HMG-CoA. Two major sub groups were classified by Istvan and Deisenhofer in 2001. The first class (type 1 statins) describes all statins containing the mimicking decalin-ring area of mevastatin (lactone). Type 2 statins comprise all statins with a β -hydroxy as HMG-CoA mimicking molecule (see Figure 2.9). (49)

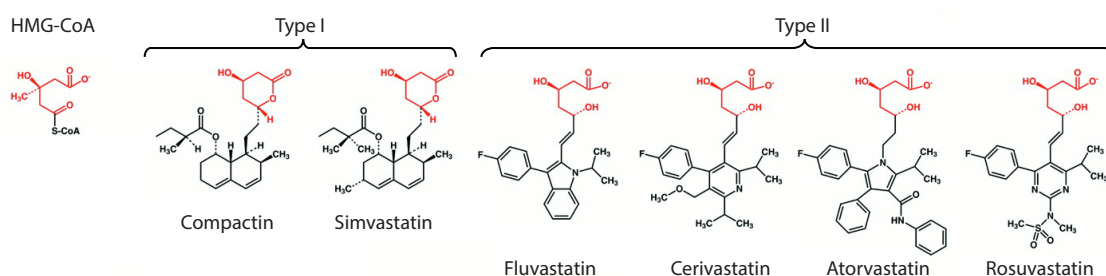


Figure 2.9: A variety of statins structures classified into type 1 and 2 statin molecules.

Structural features of HMG-CoA and six different statins that are further classified into type I and II statins. All statins contain a mimicking structure element (highlighted in red) of the HMG-CoA substrate that is binding into the active site of its converting enzyme (HMGCR). The two different types of statins are classified into type I for containing a lactone and type II when featuring a β -hydroxy group. (49)

The excessive consumption of statins can lead to a lot of different side effects as they inhibit an early reaction step of the biosynthesis of cholesterol. One of the main side effects from statins are associated with a reduced skeletal muscle function. These diseases are known as statin-associated muscle symptoms (SAMS) and occur in up to every fifth person taking statins. (50) The activation of the phosphoinositide 3-kinase (PI3K)/Akt pathway is one of the leading causes for a dysfunction of muscles, because it is related to muscle hypertrophy or muscle atrophy. A decreasing concentration of farnesyl pyrophosphate (FPP) (see Figure 6) leads to a suppression of prenylated proteins phosphorylating Akt in order to phosphorylate another protein and finally inhibit the expression of genes that are encoded for atrogenic proteins. (51) In addition to FPP, the synthesis of geranylgeranyl pyrophosphate (GGPP) is suppressed by a treatment with statins (see Figure 6). Both molecules are intermediates of the cholesterol *de novo* biosynthesis and activate many guanosine triphosphates binding regulatory proteins via prenylation. (52)

As the usage of statins for treating cardiovascular diseases is controversial discussed nowadays, the medical need of drugs that specifically address an imbalanced cholesterol homeostasis is pivotal. A strategy of an effective and specifically treating an imbalanced cholesterol homeostasis could be addressing the SREBP signalling pathway described in chapter 2.2.2.

2.3 Scap

2.3.1 Molecular features

The eight-transmembrane-helix protein was first discovered in the laboratory of M. S. Brown and J. L. Goldstein. Here, they extracted its cDNA from Chinese hamster ovarian cell lines containing mutations against “sterol-mediated suppression of cholesterol synthesis and uptake.” (53) Furthermore, they identified several transmembrane spanning regions that contain features of a so-called sterol-sensing domain (SSD), extending over transmembrane helices 2-6. The SSD is present in seven other integral membrane proteins, including the HMGCR, Niemann-Pick C1 (NPC-1), Patched and Dispatched proteins. The domain is evolutionarily highly conserved amongst proteins that are involved in cholesterol metabolism and its signalling. (54, 55) Six years after the discovery of Scap, they determined the cytosolic localization of both termini (N- and C-terminus) at the cytosol by using special protease protection and glycosylation assays. In the same study, they showed Scap containing eight transmembrane helices. (56) By co-immunoprecipitation studies, they showed an interaction between the C-terminal regions of Scap and its binding partner, SREBP. (57, 58) Additional studies on the SREBP processing showed an important ER to Golgi movement most probably via the COPII vesicular trafficking. By fluorescence microscopy they observed a movement of fluorescently labelled Scap when cells expressing the fusion protein were supplemented with cholesterol or 25-hydroxycholesterol. Under sterol-depletion the fusion protein of Scap was retained at the ER membrane. (59) Later on, they were able to show conformational changes at the luminal region of loop 7 in a cholesterol dependent manner. (60) In the same year (2002), their laboratory has worked on the second binding partner of Scap, called Insig-1. Insig-1 is an extremely hydrophobic membrane protein localized in the ER membrane. The binding of Insig-1 to Scap is dependent on the concentration of sterol in the ER membrane. (61)

In the past two decades, the group of the two scientists Brown and Goldstein were focusing on biochemical and structural features mainly of the hamster Scap as described before. They proposed an overall topology model of hamster Scap primarily based on their studies (see Figure 2.10). (55)

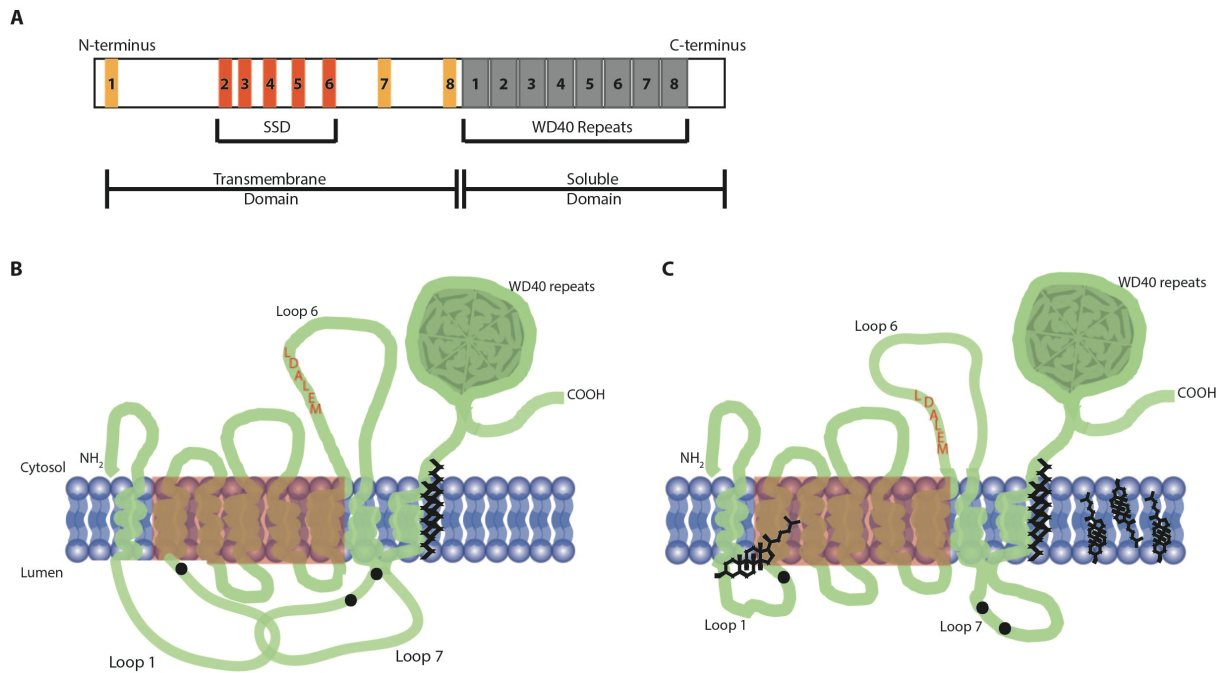


Figure 2.10: Molecular features of the eight transmembrane protein Scap and its conformational changes upon varying sterol concentration inside membranes based on analysis of hamster Scap. (55)

(A) Graphical representation of the sequence features and motifs of Scap. Scap is a transmembrane protein containing eight membrane spanning helices (shaded) that are localized at the N-terminal region. The C-terminal region is mainly soluble and contains eight WD40 repeats (grey). (B and C) Schematic representation of Scap (green) localized either inside the ER membrane or Golgi membrane. The N- and C-terminal regions are localized in the cytosol. In total eight membrane spanning helices are linked to the seven loops ranging either into the cytosol or the lumen of ER or Golgi apparatus. The first and seventh loop (at the luminal area) contain several post-translational glycosylations (black). From helix number 2 until number 6 a sterol-sensing domain is localized (red). Two palmitoylation sites are localized at the cytosolic area of the C-terminal region (black). At loop 6, the COPII protein complex binding motif (MELADL) is highlighted (orange). Scap changes its conformation in a sterol-dependent manner. (B) Conformation of Scap at sterol concentrations below 5 mol%. Loop 1 is bound to loop 7, and the MELADL sequence is accessible for COPII binding. (C) Conformation of Scap at high sterol concentrations (≥ 5 mol%). Cholesterol is bound to loop 1, and as a consequence loop 7 adopts a different conformation. This causes loop 6 to adopt a different conformation and the MELADL sequence is buried as well as inaccessible for binding to COPII protein complexes. (55)

The first loop extends into the ER lumen and interacts with loop 7 under low sterol concentrations. In addition, it contains a binding site for cholesterol. As mentioned before, the SSD is localized from helix number 2 to helix number 6 and contains the Insig-1 binding site. The Insig-1 binding site was identified in the labs of Goldstein and Brown by introducing three point mutations in this region. These mutations prevented the complex formation of Insig-1 and Scap. Subsequently, the movement of Scap from the ER membrane to the Golgi membrane was blocked. (62) Loop 6 is localized on the cytosolic site and contains the six amino acids MELADL that are the recognition site of the COPII protein complex. This region is accessible at sterol concentrations in the membrane below 5 mol%, and sterically inaccessible for the COPII protein complex at higher sterol concentrations. As a consequence of these conformational changes, loop 7 is also released from binding to loop 1 under high sterol concentrations

as loop 1 binds cholesterol (see Figure 2.10 B and C). This hydrophobic part of the protein is followed by a soluble C-terminal region which features several WD40 repeats. The cytosolic C-terminal region binds to SREBP as soon as Scap is co-translational inserted into the ER membrane. (55)

Scap contains several post-translational modifications, like three glycosylated amino acids localized in loop 1 and 7 and additional two palmitoylated amino acids rest facing into the cytosol (see Figure 2.10 B and C).

The eight-membrane spanning protein Scap is an escort protein and functions as a sensory protein of cholesterol homeostasis. Due to its high variety of functions and main function in cholesterol homeostasis its 3D structure and its structure in complex with its binding partners Insig-1 and SREBP is of great interest. As described before, new drugs treating an imbalanced homeostasis of cholesterol are needed in order to heal these diseases more specifically without side-effects.

2.3.2 Structural investigations

In the past 25 years after the discovery of Scap its overall topology is known mainly based on experiments with hamster Scap (see chapter 2.3.1). (56) Nevertheless, in 2015 X. Gong *et al.* published the first 3D crystal structure of the WD40 repeats domain of a Scap ortholog from fission yeast called Scp1. (63) In *Schizosaccharomyces pombe* the protein Scp1 is a transmembrane protein and shares high similarity features with the mammalian Scap (see Figure 2.11). The fission yeast Scp1 is roughly 200 amino acids smaller in size but features all the domains available in mammalian Scap. Both are composed of eight transmembrane helices out of which 5 helices (helix 2 until 6) contain the SSD (red in Figure 10). Furthermore, both orthologues are made up of a long cytosolic and soluble C-terminal region which is nearly as long as the transmembrane domain. The soluble domain of the Scap proteins features eight WD40 repeats (grey in Figure 2.11). (55) Although, the amino acid sequence identity of the yeast and mammalian Scap is very low. In *S. pombe* a similar but simplified signalling cascade of the SREBP signalling pathway was found in 2005 by A. L. Hughes, *et al.* (64) In the yeast system, all three main members of the pathway were found and named *sre1+* (SREBP), *scp1+* (Scap) and *ins1+* (Insig-1). They postulate a similar cascade as *sre1* is proteolytically cleaved and activated with Scp1 in a sterol dependent manner. Furthermore, they observe that the sterol synthesis is dependent on the oxygen level. These findings show that the SREBP pathway in yeast responds to the levels of sterol as a secondary response of oxygen supply. (64, 65)

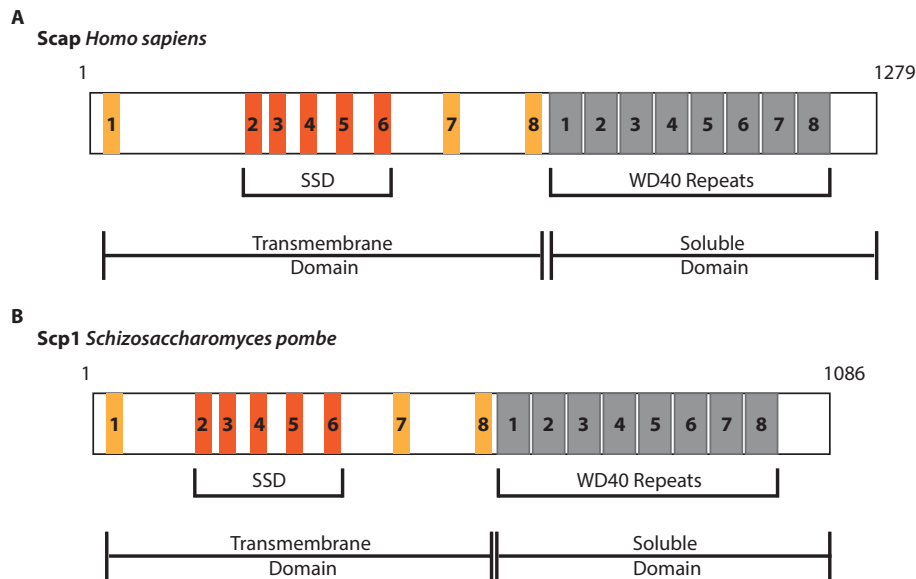


Figure 2.11: Comparison of the overall topology of Scap *Homo sapiens* and *Schizosaccharomyces pombe*.

(A) The Topology model of human Scap containing 1279 amino acids. The human Scap features eight transmembrane domains (coloured) out of which helix 2 until 6 form the highly conserved sterol sensing domain (SSD, red) (highly conserved upon membrane proteins that have an impact on cholesterol movement and homeostasis). Its soluble domain forms eight WD40 repeats (grey) and an additional soluble C-terminus. (B) The topology model of the fission yeast Scp1 shares a high similarity to the mammalian Scap. It contains 1086 amino acids but all the other features, like eight transmembrane domains (coloured), the SSD (red) and eight WD40 repeats (grey) are part of both Scap orthologues. (63)

X. Gong *et al.* solved the structure of the soluble WD40 repeats domain from Scp1 at 2.1 Å resolution by x-ray crystallography. They were able to show its activity by binding to the C-terminal region of the SREBP ortholog from Sre1 via *in vitro* studies. (63) The crystal structure features an eight-bladed β -propeller that was purified out of co-expression experiments in *Escherichia coli*. Via binding studies with a variety of truncated versions of the soluble C-terminal domain they observed different binding affinities to Sre1 in order to analyse the specific recognition and binding areas of the heterodimeric complex (see Figure 2.12). The *in vitro* reconstituted complex was furthermore tested successfully for its activity *in vivo* with a fission yeast strain expressing an epitope tagged *scp1* WT from *scp1*⁺ promoter. They observed that residues 1055 – 1085 are required for the complex formation and the activation of the Sre1 processing. (63)

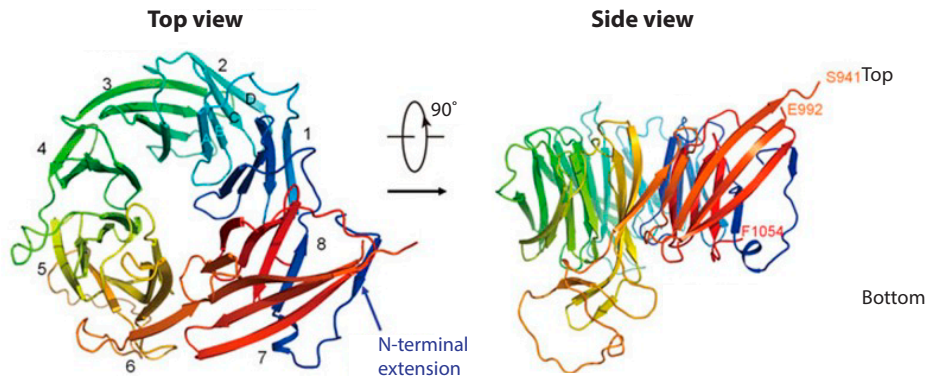


Figure 2.12: Crystal structure of the WD40 repeats from the fission yeast Scp1. (PDB: 4YHC)

The structure of the Scp1 ortholog from *S. pombe* contains eight blades that are individually highlighted in a different colour and consecutively numbered. The centre of the soluble domain contains a hole that is surrounded by the β -propeller formation. (63)

In 2016 the complex formation of the soluble C-termini from the fission yeast orthologues Sca1 and Sre1 was solved by single particle cryo-EM. They observed an oligomeric formation of an octamer composed of four Sca1 WD40 domains (3.5 Å) and four molecules of the Sre1 C-terminal domain (5.4 Å). The oligomeric complex is based on the formation of a dimer of two molecules (dimer) of Sre1 which builds the centre. Each Sre1 protein is furthermore bound to Scp1 WD40 domain that is ranged to the outside (see Figure 2.13). (66) Gong *et al.* analysed the complex activity *in vitro* and *in vivo*. They have observed in total three regions in Sre1 C-terminal domain that are required for the complex formation with Sca1 *in vitro*. The tetrameric formation of the C-terminal domains of Sca1 are furthermore required in *S. pombe* for its complete activation. For this issue they have performed the same *in vivo* analysis as before for the activity of soluble WD40 repeat domain of Sca1. They have used a fission yeast strain expressing an epitope tagged *scp1* WT from *sre1*⁺ promoter in order to observe an impact of the expression level at low oxygen levels. By this study, they have furthermore shown that the oligomeric protein complex is active *in vivo*. (66)

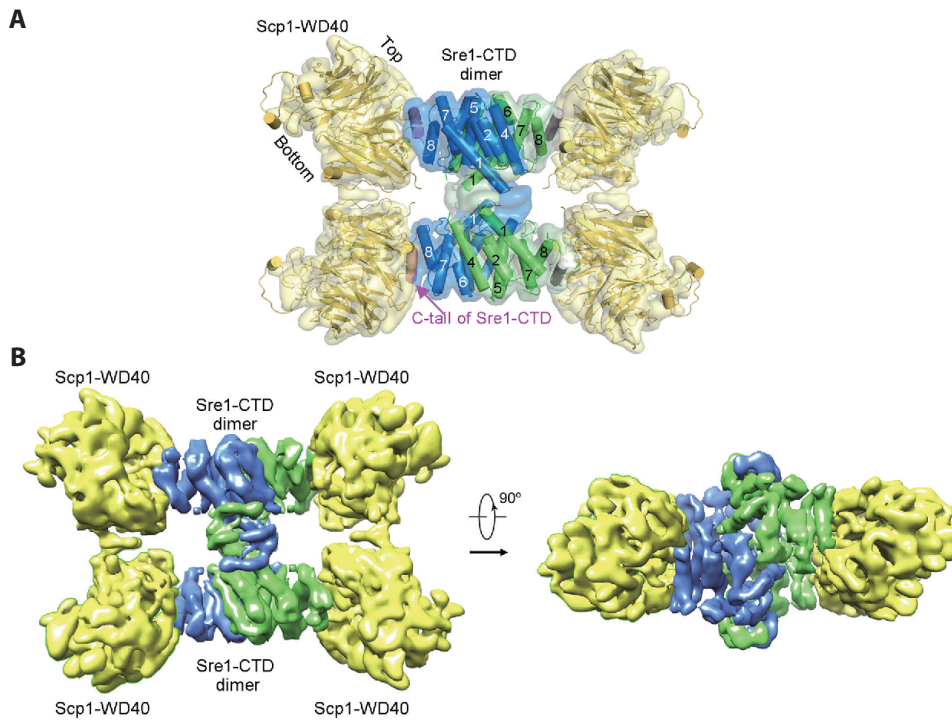


Figure 2.13: Cryo-EM reconstruction of the *S. pombe* Sre1-Sca complex from their C-terminal domains. (PDB: 5GRS)

(A) The overall cryo-EM map of the WD40 repeats of Sre1 bound to the C-terminal domain of Sca1 from *S. pombe* showing the dimer formation of the two Sre1 proteins in the middle (coloured in blue and green) bound to a second pair of Sre1 dimer formation. Each Sre1 domain is bound to one WD40 domain of Sca1. The previously solved crystal structure of Sca1 was fitted into the density map (yellow). (B) Each Sre1 monomeric C-terminal protein is coloured in blue or green which forms a dimer of a dimer. Each dimeric complex formation additionally features two Scp1 WD40 repeats domain (yellow) ranging into the outside and additionally form a complex with another Scp1 protein from the second dimeric complex. (66)

As described above and shown in Figure 2.11, the overall topology of the fission yeast orthologs Sca1 and Sre1 share high similarity. Furthermore, a similar SREBP signalling pathway is present in *S. pombe* including a SREBP ortholog activating enzymes for biosynthesis of sterols, like ergosterol. (66) Nevertheless, the amino acid sequence of each protein from the fission yeast compared to the mammalian does not feature an identity level that would be necessarily needed for an insight into the 3D structure of Scap protein nor the complex build by Scap and SREBP. In addition to that, the transmembrane domains of both proteins are missing in both structure analysis. For a complete understanding of the binding to the third membrane protein Insig-1 and the sterols, the structure of the transmembrane domain including the cytosolic and luminal regions is of particular importance. Furthermore, the complex formation of the proteins from *S. pombe* might differ from the mammalian orthologs because of a low similarity in their amino acid sequence. In order to understand the signalling cascade on a molecular level and designing new drugs against an imbalanced cholesterol homeostasis, the 3D structure of the full-length mammalian proteins Scap, SREBP and Insig is of interest. Structural information of Scap alone and in complex with its binding partners are important for giving an insight into the molecular machine that measures and controls the cholesterol concentrations of membranes.

2.3.3 Synthetic antagonists

Today, there are several drugs available treating an imbalanced cholesterol homeostasis that leads to atherosclerotic disorders. Most of them belong to the functional molecular class of statins. As described in chapter 2.2.3, statins have a lot of side effects that roughly every fifth person taking statins has been reported. (50) Statins are competitive inhibitors of the HMGCR that catalyses the rate limiting step of *de novo* cholesterol synthesis in cells. By the inhibition of the reaction from HMG-CoA to mevalonate the synthesis of additional intermediates of this pathway are blocked leading to severe side effects. To date, there are three inhibitors available, that more specifically treat these diseases by directly addressing the complex formation of Scap and SREBP (see Figure 2.14).

The first inhibitor is called Fatostatin and known to bind to Scap and initiates its retention in the ER membrane. The binding of Fatostatin was successfully verified by binding analysis where an interaction between Fatostatin and the Amino-terminal domain of Scap could be identified. The synthetic inhibitor stabilizes the complex formation of Scap and Insig proteins by binding to Scap. Thereby, the dissociation of Scap from Insig for the COPII vesicular trafficking from the ER to the Golgi membrane is hindered. (67, 68)

Betulin is the second small molecule actively addressing the Scap-SREBP complex. Similar to Fatostatin, Betulin binds Scap and the Scap/Insig complex is actively stabilized. The signalling cascade is blocked by the permanent localization of the hetero-trimeric complex Insig/Scap/SREBP in the ER membrane. Further analysis showed a clear decrease in the promoter activity of target genes that are encoded for proteins that are necessary for cholesterol synthesis like HMGCR. (67, 69)

The last known synthetic compound, Xanthohumol, was originally extracted from hops (Flavonoid) and afterwards prenylated. (70, 71) The compound was identified as an antagonist of SREBP and specifically inhibits its activation. It is suppressing the ER-Golgi movement, by binding to a special region of the COPII protein complex that usually interacts with Scap upon low sterol concentrations in the ER membrane. In addition, Xanthohumol actively inhibits the synthesis of triglyceride that causes the inhibition of SREBP maturation and gene transcription of its target genes. (67, 72)

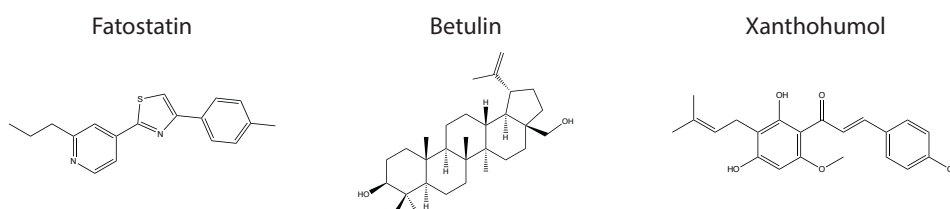


Figure 2.14: Molecular structures of three synthetic antagonists of the SREBP signalling pathway adapted from Lee *et al.* (67)

Fatostatin, Betulin and Xanthohumol are small molecules addressing the complex formation of Scap/SREBP. (67)

2.4 Patched1

Patched is a polytopic membrane protein localized in the cell membrane of a large variety of cell types, like brain, lung, liver and skin cells. Similar to Scap, Patched proteins contain the highly conserved SSD (amongst proteins related to cholesterol binding or cholesterol homeostasis). Furthermore, Patched has a significant role in signalling like Scap and binds cholesterol. (73)

Originally, there was found one Patched gene for *Drosophila* and later on two paralogs (termed Patched1 and 2) in vertebrates. Both proteins share a sequence identity of 57%. Patched2 lacks the long N- and C-terminal regions of Patched1 and differs in its expression and function compared to Patched1. The human Patched1 is composed of 1447 amino acids and human Patched2 of 1203 amino acids. Patched1 protein acts as the receptor of processed Hedgehog (Hh) proteins and as a transcription inhibitor by suppressing Smo activity in the Hh signaling pathway. (74) Due to its pivotal role in the signaling cascade of Hh proteins (Hh signaling pathway) Patched1 belongs to one of the new drug targets. In this chapter, the function of Patched1 in the Hh signaling pathway and its structural features are highlighted.

2.4.1 Hedgehog signalling pathway

The Hh signalling pathway is one of the main pathways in tissue patterning events during embryonic development and stem cell biology. (74) Dysregulation in this signalling cascade leads to several types of cancer and especially birth defects, which are the reasons for its essential need of knowledge on a molecular basis. (75-79)

In 1980, Nüsslein-Volhard and Wieschaus discovered the Hedgehog proteins that all function as an intracellular signal transducer. (80) In total, three different Hh proteins were detected and later on called Desert Hedgehog (Dhh), Indian Hedgehog (Ihh) and Sonic Hedgehog (Shh). All three proteins share a high similarity with their orthologue expressed in *Drosophila*. (81) Nevertheless, besides the co-expression of Ihh and Shh in gut cells and Ihh and Dhh in adult ovarian cells there are no overlapping expression patterns found for the three Hh family members. (82, 83) The gene *Dhh* is proposed to play a fundamental role in the nerve sheath formation and is expressed in Schwann cells. Additionally, Dhh is a key protein for the maintenance of the spermatogenesis. (84-86) Compared to this, Ihh is essential in development, especially for the chondrocyte development. (87) Shh is the best and most analyzed protein member of the Hh family. Shh is the regulator of stem cell homeostasis and is involved in the maintenance of the neural progenitor cells. (88) The Hh proteins undergo several auto processing steps at the ER before it binds to a membrane protein called Dispatched, that transfers the protein from the cytosolic area into the extracellular region. The nascent Hh proteins release its processing domain by an internal cleavage and adds a cholesterol molecule at the termini via an ester bond to the glycine rest. Due to the internal modification, the glycine is exposed. The nascent protein furthermore contains a signal sequence and undergoes a proteolytically cleavage on the other terminus (signal

peptide cleavage) which results in a cysteine at the total N-terminus. An *O*-acyltransferase finally adds a palmitate to the N-terminus that builds the final post-translational modified Hh protein with a decreased molecular weight of 19 kDa. The auto-processed Hh protein is more stable compared to its pre-cursor version, because of the post-translational modified Hh proteins that build higher oligomeric structures afterwards. In this oligomeric state, the protein is stabilized and a long-time activity is given in contrast to the un-modified (nascent) Hh protein. (89) The mature protein is then released of the cell by Dispatched. (90)

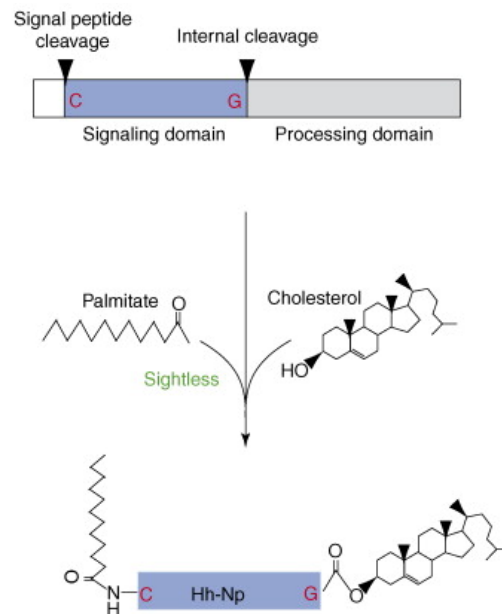


Figure 2.15: Auto processing of the nascent Hh protein from I. Guerrero and C. Chiang. (89)

The precursor Hh proteins undergo several post-translational modifications. First, the large C-terminal domain (grey) is cleaved by an internal modification at a glycine residue which is coupled with an additional cholesterol modification step. Due to the internal rearrangement the glycine (red) residue is exposed and the sterol is added to the glycine residue (red) via an ester bond. Second, the Hh proteins gets palmitoylated at the N-terminal cysteine. This reaction is catalysed by an *O*-acyltransferase, named Slightless (green). (89)

Dispatched and SCUBE2 proteins bind to the mature Hh proteins and Dispatched releases the protein into the extracellular area by an unknown mechanism. (91) A structure of Dispatched bound to Shh was solved by single particle cryo-EM at an overall resolution of 4.8 Å. This revealed a first molecular insight into the binding of this heteromeric interaction but we are still lacking information on the mechanisms of adopting Hh from inside the cell and release into the extracellular milieu. (92)

In mammalian cells, the released mature Hh protein bind to its receptor Patched1 which is localized close to the primary cilium (see Figure 2.16). The primary cilium is a cell organelle that dynamically extend from mammalian cells in order to receive and transmit information from the outside to the cell itself. These organelles are microtubule-based and extend from the basal body. (93) In the absence of mature Hh proteins, the Hh signalling pathway is inactivated by the repressed activator Smoothed (Smo). Smo is repressed by Patched1 (see Figure 2.16 (A)). Until today, the way of suppression is still unknown but several hypotheses are postulated. The most popular hypothesis of suppression includes

another ligand, that could be the cholesterol moiety of the mature Shh. (94) Furthermore, there are studies clearly showing a dislocated Smo protein from the primary cilium when repressed by Patched1. (95, 96) In presence of the double-lipidated Hh proteins, Smo is freed from Patched1 repression and activates the Hh signalling cascade. The Hh protein binds to its receptor Patched1 most probably via the post-translational added lipidation moiety at its N-terminus. Several structures of the complex formation either without Shh lipidations, with only the palmitate and with both lipidations were published in 2018 and 2019. (5-10)

In its active form (when freed from Patched1 repression), Smo accumulates to the primary cilium and activates the down-stream signalling cascade of the Hh signalling pathway (see Figure 2.16 (B)). The signalling cascade starts with the activation of the glioma-associated oncogene homologs 2 and 3 (GLI2/3). By the maintenance of the full-length transcription factors GLI2/3 due to the simultaneous degradation of cAMP by the phosphodiesterase 4 (PDE4) and the release of the G-protein coupled receptor 161 (GPR161) from the primary cilium, the active forms of GLI2/3 enter the nucleus. GLI2/3 are transcription factors that induce in their full-length version the expression of Hh target genes (see Figure 2.16 B). (96)

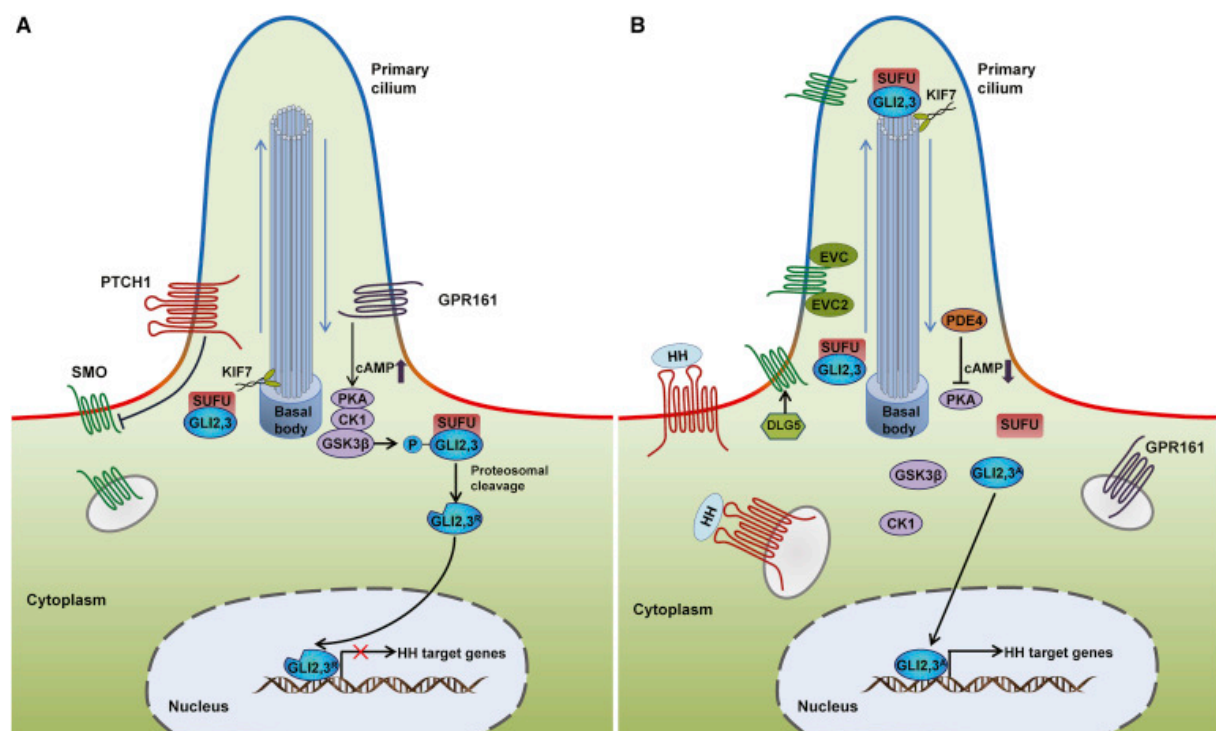


Figure 2.16: Schematic illustration of the Hh signalling pathway from F. Wu, *et al.*. (96)

(A) The Hh signalling cascade in the absence of the mature Hh protein. Patched1 (red) is localized at the primary cilium where it actively represses Smo (green). The signalling is blocked and the expression of Hh target genes is inactive. (B) The Hh signalling pathway in the presence of the fully lipidated Hh protein. Hh is bound to its receptor Patched1 and Smo is freed from Patched1 repression. The Hh signalling is active. The glioma-associated oncogene homolog 2 and 3 (GLI2 and 3, blue) maintain their full-length and phosphorylation site, when the protein kinase A (PKA, purple) is inhibited by the decreasing level of cyclic AMP (cAMP) that was induced by the release of G-protein coupled receptor 161 (GPR161) from primary cilium. The transcription factor of GLI2 and 3 enters the nucleus and activates the expression of the Hh target genes. (96)

2.4.2 Structural investigations

The mammalian Patched1 is an integral membrane protein containing 12 transmembrane domains. Both N- and C-termini are located into the cytosol of a cell as Patched1 proteins are localized in cell membranes. The protein has two extracellular domains (ECD) that face the extracellular region. ECD 1 is localised between the first and second transmembrane helices and ECD 2 between helix number 7 and helix number 8. (10) As mentioned before, Patched1 contains the SSD which is highly conserved upon membrane proteins that are involved in either cholesterol homeostasis, direct binding to cholesterol or both (see Figure 2.17). (55, 56)

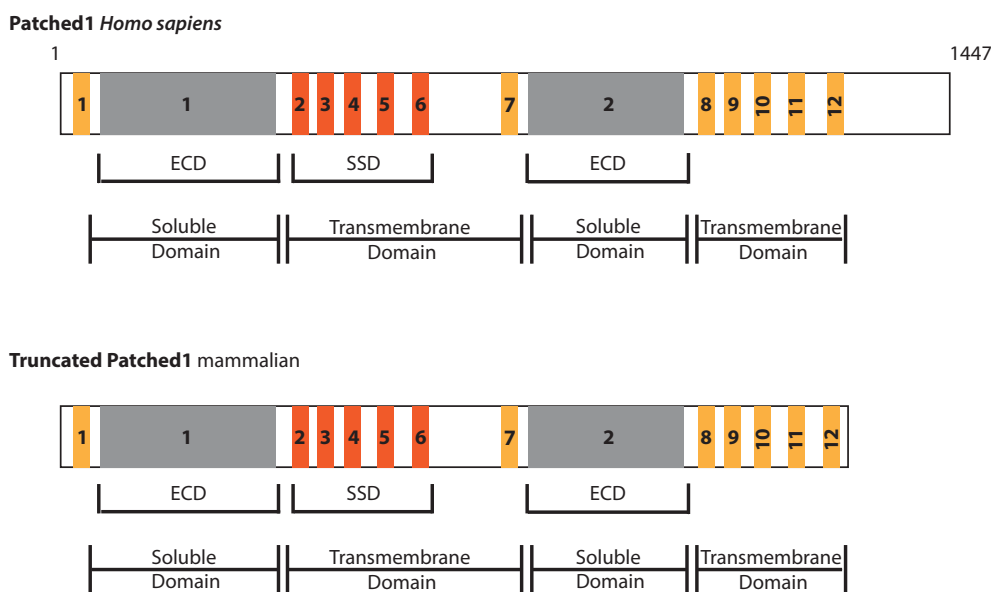


Figure 2.17: Overall topology models of human Patched1 and the truncated Patched1 used for structural analysis.

The human Patched1 protein contains 1447 amino acids and is composed of 12 transmembrane domains (coloured). From helix number 2 until helix number 6, the SSD (red) is localized. Between the first two transmembrane domains and helix number 7 and 8, Patched1 contains two large soluble domains that are localized into the extracellular region (ECD, grey). Both N- and C-termini are localized into the cytosol of a cell. The truncated Patched1 homolog that was used for a variety of structural analysis with and without its ligand Shh is mainly missing the complete extended version of the C-terminus. Some structures were solved by using a truncated N- and C-terminus of the mammalian Patched1. (5-10)

Furthermore, Patched1 shares high sequence identity with another SSD containing membrane protein named NPC1. In the SSD, both Patched1 and NPC1 share a sequence identity of 33%. NPC1 belongs to the Resistance-Nodulation-Division (RND) family that are known to either transport or bind sterols and related chemical structures. (6, 95, 97) As Patched1 is the receptor of lipidated Hh proteins and its published structure shows several additional densities that would fit for additional cholesterol related

molecules (cholesteryl hemisuccinate; CHS) there is strong evidence of supporting and furthermore transport cholesterol molecules. (5-10, 98)

As mentioned before, in 2018 and 2019 several structures were published of a full length and truncated mammalian (either murine or human) Patched1 alone and in complex with lipidated or unmodified Shh proteins. (5-10) The first published structure refers to the murine Patched1 homolog. Zhang *et al.* solubilized and purified a Patched1 construct that was missing an internal cytosolic loop and the flexible C-terminus. Via single-particle cryo-EM they were able to solve the structure at an overall resolution of 3.6 Å. In contrast to the findings, that the internal cytosolic domain would initiate an oligomerization of Patched1 they were still observing a homodimeric structure without this feature. (6) Via a Gli-dependent luciferase assay they successfully showed its activity of suppressing Hh signalling, although they were working with a truncated construct of the murine Patched1. Furthermore, they found three extra cavities that range from the extracellular region to the inside of the transmembrane domains. It is mainly discussed that sterols, like cholesterol molecules, would most probably fit into these extra densities. In addition to the analysis of Patched1 structure, Zhang *et al.* analysed the effect of Patched1 on the distribution of cholesterol in the outer and inner leaflet of the lipid bilayer. It was shown, when Shh was absent the amount of cholesterol was reduced in a Patched1 dependent manner. These findings indicate an activation of Smo by cholesterol retaining in the inner leaflet of the membrane which furthermore supports the hypothesis of a regulation of Smo by a Patched1 dependent cholesterol transport. (96, 99) Several weeks after the first published structure, another two reconstructions of Patched1 with and without Shh were published by X. Gong *et al.*. In contrast to the first publication, they were able to solve the structure of the human Patched1 containing the internal cytosolic domain, but lacking the C-terminus, too. With these findings, this structure stands in contrast to the hypothesis of an oligomerisation triggered by the internal cytosolic domain, which was furthermore refuted by the structures solved by Y. Zhang *et al.*, X. Qi *et al.* as well as H. Qian, *et al.* (6, 7, 9-11) By *in vivo* experiments where they measured a similar binding affinity of their truncated construct compared to full length protein, they were able to prove its activity. In comparison with the first structure, they have found additional densities in a pocket build by the two ECDs and at the transmembrane region, each would fit to one sterol molecule. When analysing the structure out of the truncated Patched1 and the non-modified Shh, a mainly interactions of charged amino acid residues between both proteins is observed. Shh is bound to both ECD1 and ECD2 regions and initiates an increasing interface between these domains. (10) One month later, X. Qi *et al.* published additional two structures of a similar truncated Patched1 construct from human (compared to the first structure) and in complex with the complete lipid-modified Shh. They found a monomeric and trimeric formation of the human Patched1 construct but were focusing on the monomeric formation as Patched1 was postulated be in an oligomeric state only upon presence of the cytosolic domain that was missing. Furthermore, they observed an increased stability of the monomer compared to the trimer during purification. They solve the structure by single-particle cryo-EM at overall resolutions of 3.5 Å (for the human Patched1) and 3.8 Å (in the case of Patched1 bound to Shh). In addition, they have observed extra densities in which sterol molecules most likely would fit into. (7)

In 2019, the same group of the second published structure of Patched1 and first published Patched1-Shh structure by Gong *et al.* published a third reconstructed structure of the complex build of Patched1 and palmitoylated Shh by Qian *et al.*. This time, they have reached a resolution of roughly 4 Å with the palmitoylated Shh in complex that furthermore induced an oligomerization of a 4:2 stoichiometric ratio of Patched1:Shh. As they have found a dimer of a 2:1 stoichiometry they were processing the data with an adaptive mask for this ratio that improved the overall resolution up to 3.6 Å. With the additional palmitoylation of the Shh they were able to show the extra density created by the palmitate. The palmitate was extensively interacting with the ECD1 region and the rest of Shh was bound to the ECD1 domain with less affinity. (9) The last structure of Patched1 in complex with a lipid-mimicry Shh (Shh_{C24II}) was solved in the end of 2019 at a resolution of 3.4 Å by C. Qi *et al.*. Similar to most of the reconstructions, they observed a monomeric Patched1 protein bound to one Shh_{C24II} molecule. In contrast to all the other structures available of Patched1 in complex with different Shh versions, C. Qi *et al.* had a large amount of extra sterol-like densities. They furthermore explained these additional densities in their reconstruction as “the missing clues for linking the putative sterol translocation pathway from the inner to the outer leaflet of the SSD.” Their double-mutated Shh construct showed median activity compared to the recombinant human mature Shh, which they have analysed via the Gli-dependent Hh pathway activity assay. (10)

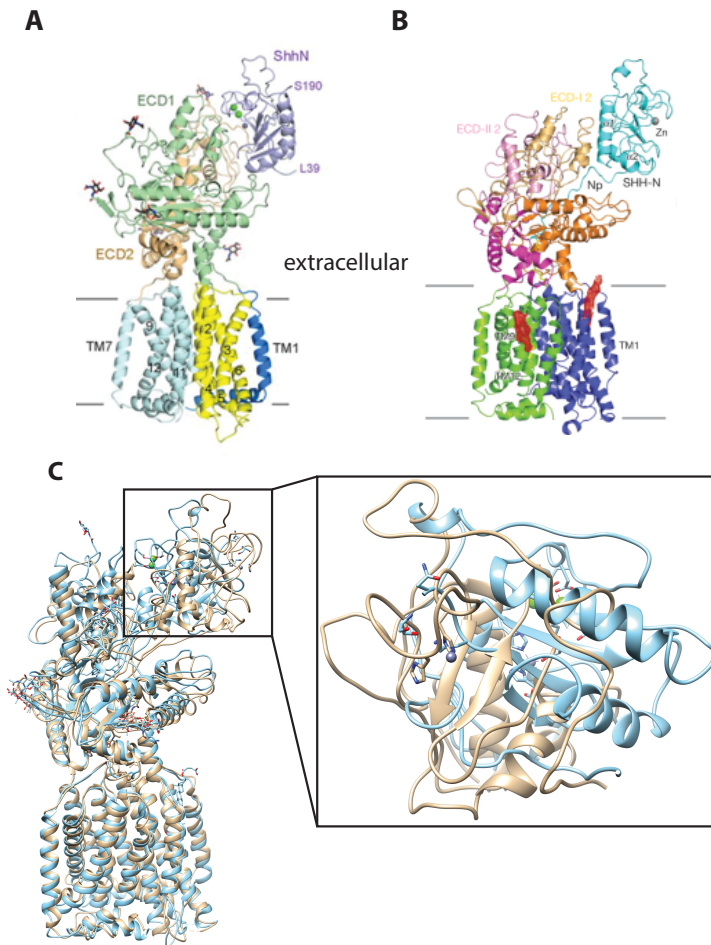


Figure 2.18: Comparison of structures of human Patched1 in complex with Shh with and without lipidations.

(A) Structure of the human C-terminal truncated Patched1 protein in complex with Shh at an overall resolution of 3.6 Å. (PDB code: 6DMB) (B) Structure of the human Patched1 protein in complex with lipidated Shh at 3.8 Å resolution. (PDB code: 6OEU) (7, 10) (C) An overlay of both human Patched1 constructs in complex with Shh and a close-up of the bound Shh state. The blue strand refers to structure 6DMB and the beige one refers to the structure of 6OEU.

All in all, the data of the previous two years highlight the importance of understanding the complex formation between Patched1 and its ligand Shh as well as the effort on working on such membrane proteins interacting with cholesterol in order to fulfil the understanding on a molecular basis and the elucidation of new drugs treating related diseases. Nevertheless, it remains still unclear which stoichiometry is built upon Patched1 ligand binding and which oligomeric state might be an artefact or not. Furthermore, the binding side of the cholesterol moiety that is part of the mature Shh remains unclear and is one of the biggest enigmas as X. Qi *et al.* published structure does not feature a density of the cholesterol from the double lipidated Shh they have used. (5-10)

3 Objective

Membrane proteins are pivotal regulators in a variety of processes for cell viability like cell-cell communication, signal transduction or the transport of molecules across lipid bilayers. (1, 2) A complete understanding of their function requires the elucidation of three-dimensional (3D) structures. A 3D structure of the target protein enables an insight into its binding motif and complex formations in transduction processes on a molecular level. However, less than 7% of the published structures belong to membrane proteins in 2020 highlighting the complex study of membrane proteins. (100, 101)

The overall goal of this thesis was the analysis of two membrane proteins on a molecular level via cryo-EM. Here, the polytopic transmembrane proteins, Patched and Scap, were characterised due to their key functions in signalling. Both, Patched1 and Scap, contain a sterol sensing domain (SSD) and bind cholesterol. (3) Patched is the receptor of processed Hh proteins and functions as a transcription inhibitor by suppressing Smoothed (Smo) activity that serve as the starting point in the Hh signalling cascade. (74) A dysregulated Hh signalling pathway leads to several types of cancer and birth defects. (75-79) Scap regulates cholesterol homeostasis via escorting SREBP from the ER to the Golgi membrane in a sterol dependent manner. An imbalanced cholesterol homeostasis often causes cardio vascular diseases which is one of the leading causes of death. (5)

For the elucidation of structural insights of both, Patched and Scap, the first aim is the establishment of a large-scale expression followed by an efficient purification strategy. For this purpose, the validation of a stable membrane protein construct is required. Here, different paralogs and orthologs need to be expressed in small-scale first and screened for its stability when solubilized in a detergent afterwards. Furthermore, the protein constructs' native conformations need to be confirmed by fluorescence microscopy in the case of truncations or other modifications that would be inevitable for the design of a stable membrane protein. Negative stain electron microscopy will be used for the design of a first 3D model and the verification of a homogeneous protein sample that is required for structure analysis by cryo-EM. In accordance with recently published data, a truncated construct of Patched1 was successfully used for structural analysis via cryo-EM and furthermore gave a first insight of the complex formation with its receptor Shh. (5-10) Since several high-resolution cryo-EM structures of Patched1 alone and with (lipidated) Shh were published in the past two years, I focused working on Scap as we are still lacking structural information on full-length mammalian Scap alone and in its bound conformation with Insig and SREBP. The monomeric as well as the heterodimeric and heterotrimeric structure analysis would help visualizing the conformational changes upon protein complex formation and its binding interfaces on a molecular level.

4 Materials and Instrumentation

4.1 Chemicals

Table 4.1: Chemicals.

Chemical	Manufacturer
1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC)	Anatrace
2-(4-Hydroxyphenylazo)benzoic acid (HABA)	Sigma-Aldrich Chemie GmbH
3x FLAG Peptide (25 mg)	APExBIO
4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES)	GERBU Biotechnik GmbH
4-Cyclohexyl-1-heptyl- β -D-maltoside (Cymal-4)	Anatrace
6-Cyclohexyl-1-heptyl- β -D-maltoside (Cymal-6)	Anatrace
Acetonitrile	Thermo Fisher Scientific
Acetic acid	Sigma-Aldrich Chemie GmbH
Adenosine 5'-triphosphate (ATP) disodium salt hydrate	Sigma-Aldrich Chemie GmbH
Agarose	Biozym Scientific GmbH
Ammonia (25%)	Fischer Scientific
Ammonium bicarbonate (NH_4HCO_3)	Riedel-de Haën
Amphipol A ₈₋₃₅	Anatrace
Aprotinin	Carl Roth GmbH + Co. KG
Biobeads M2	Bio-Rad Laboratories
Boric acid	Merck KGaA
Bromophenol blue	SERVA Electrophoresis GmbH
Calcium chloride	Fluka Chemicals
Calcium sulfate dihydrate	Carl Roth GmbH + Co. KG
Calpain inhibitor III	Cayman Chemical Comp.
CellLight™ ER-GFP, BacMam 2.0	Invitrogen™ Thermo Fisher Scientific
CHAPS	Anatrace
CHAPSO	Anatrace
Cholesterol	Sigma-Aldrich Chemie GmbH
Cholesteryl hemisuccinat (CHS)	Sigma-Aldrich Chemie GmbH
Cobalt(II) chloride hexahydrate	Carl Roth GmbH + Co. KG
Collodion solution	Sigma-Aldrich Chemie GmbH
Coomassie-brilliant blue R250	SERVA Electrophoresis GmbH
Copper(II) sulfate pentahydrate	Merck KGaA
D-(+)-biotin	Merck KGaA
D-desthiobiotin	Sigma-Aldrich Chemie GmbH

D-glucose (dextrose)	Carl Roth GmbH + Co. KG
Digitonin	Sigma-Aldrich Chemie GmbH
Dipotassium phosphate	Carl Roth GmbH + Co. KG
Dithiothreitol (DTT)	Carl Roth GmbH + Co. KG
Ethanol	Thermo Fisher Scientific
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth GmbH + Co. KG
Fos-Choline 11	Anatrace
Fos-Choline 13	Anatrace
FuGENE® transfection reagent	Promega
Glycine	Carl Roth GmbH + Co. KG
Glyco-diosgenin (GDN)	Sigma-Aldrich Chemie GmbH
Hydrochloric acid (1 M)	VWR International
Hydrochloric acid (25%)	AppliChem GmbH
iBind™ Solution Kit	Thermo Fisher Scientific
Iodoacetamide	Sigma-Aldrich Chemie GmbH
Iron(II) sulfate heptahydrate	Fluka Chemicals
Lauryl dimethylamine-N-oxide (LDAO)	Santa Cruz Biotechnology
Lauryl maltose neopentyl glycol (LMNG)	Anatrace
Leupeptin hemisulfate	Cayman Chemical Comp.
Lipofectamine® 3000 (reagent and additive)	Thermo Fisher Scientific
Lithium acetate (LiAc)	FLUKA
Lithium chloride	Carl Roth GmbH + Co. KG
Magnesium sulfate heptahydrate ($\text{MgSO}_4 \times 7\text{H}_2\text{O}$)	J. T. Baker
Manganese(II) sulfate monohydrate	Carl Roth GmbH + Co. KG
Methanol	Sigma-Aldrich Chemie GmbH
Midori green	NIPPON Genetics
n-Dodecyl β -D-maltoside (DDM)	Sigma-Aldrich Chemie GmbH
n-Dodecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate (Anzergent 3-12)	Anatrace
n-Decyl β -D-maltoside (DM)	Anatrace
n-Decyl-N,N-dimethyl-3-ammonio-1-propanesulfonate (Anzergent 3-10)	Anatrace
n-Octyl- β -D-Glucopyranoside (OG)	Anatrace
PANFect A (Transfection Reagent and buffer)	PAN™ Biotech GmbH
Pefabloc	Sigma-Aldrich Chemie GmbH
Pepstatin A	Carl Roth GmbH + Co. KG
Phenylmethanesulfonyl fluoride (PMSF)	SERVA Electrophoresis GmbH
Phosphoric acid (85%)	Sigma-Aldrich Chemie GmbH

Plus ECL (enhanced substrate)	chemiluminescence	Perkin Elmer
Sodium hydroxide (2M)		Waldeck
Sodium iodide		Carl Roth GmbH + Co. KG
Sodium molybdate dihydrate		Merck KGaA
Sulfuric acid		FLUKA
Trifluoroacetic acid (TFA)		FLUKA
Tris(2-carboxyethyl)phosphine (TCEP)		Biosynth
Tris PUFFERAN		Carl Roth GmbH + Co. KG
Trypan blue		FLUKA
Uranyl formate		Polysciences Inc.
Xylene cyanol		Sigma-Aldrich Chemie GmbH

4.2 Consumables

Table 4.2: Consumables.

Consumable	Manufacturer
96-well deep well plates 2.2 ml	VWR International
96-well plate, flat-bottom microplate, black (MicroWell™)	Thermo Fisher Scientific
6/12-well plates (sterile)	Sarstedt AG & Co. KG
Amicon® Ultra-4 and 15 Centrifugal Filter Unit (Millipore)	Merck KGaA
Blotting Paper (Ø55/20 mm, grade 595)	Electron Microscopy Sciences
Econo-Pac® Chromatography Columns	Bio-Rad Laboratories
Copper grids (for negative stain EM)	Plano GmbH
Cuvettes (10, 4, 45 mm)	Sarstedt AG & Co. KG
Disposable pipette tips (10 µl, 200 µl, 1.25 µl)	Sarstedt AG & Co. KG
Durapore® Membrane Filter, 0.22 µm (47 mm membrane)	Millipore
Eppendorf Tubes (0.5 µl, 1.5 µl, 2 µl and 5 ml)	Sarstedt AG & Co. KG
Falcon™ Standard- Tissue Culture Dishes	Falcon
Falcon Tubes (15 ml and 50 ml)	Sarstedt AG & Co. KG, Greiner
Filtropur V50 250/500 ml, 0.22 mm	Sarstedt AG & Co. KG
GF-1.2/1.3-3Au-45nm-50	Protochips Inc.
Heat-Inactivated Fetal Bovine Serum (FBS)	Sigma-Aldrich Chemie GmbH
iBind™ Cards	Thermo Fisher Scientific
nanoDSF Grade High Sensitivity Capillaries	NanoTemper Technologies
Ni ²⁺ -IDA resin (Protino)	MACHERY-NAGEL GmbH & Co. KG

PageRuler™ Prestained protein ladder	Thermo Fisher Scientific
PureCube Ni ²⁺ -NTA Agarose	Cube Biotech
QUANTIFOIL R 1.2/1.3 Cu (200 or 300)	Quantifoil Micro Tools GmbH
SDS-Gels mini-PROTEAN TGX Stain Free Precast gels (4-15%, 15 well)	Bio-Rad Laboratories
Serological Rotalibo® pipettes (5 ml, 10 ml, 25 ml)	Carl Roth GmbH + Co. KG
Serological pipettes, sterile (5 ml, 10 ml, 25 ml)	Sarstedt AG & Co. KG
Spectra/Por membrane tubing (MWCO 12-14 kDa, 30 kDa, 100 kDa)	Carl Roth GmbH + Co. KG
Strep-Tactin® Superflow® 50% suspension	iba
TALON® beads	Clontech Laboratories
Trans-Blot® Turbo™ Transfer Pack PVDF	Bio-Rad Laboratories
UltrAuFoil Holey Gold Films R1.2/1.3 (200 and 300 mesh)	Quantifoil Micro Tools GmbH

4.3 Electronic devices

Table 4.3: Electronic devices.

Device	Manufacturer
Äktapurifier	GE Healthcare
Autoklav LaM-4-20-MCS-J	SANOclav
Autosampler C-96	Bio-Rad Laboratories
Balance ME4002T	Mettler Toledo
Carbon evaporator Leica EM ACE600	Leica Microsystems
Centrifuge 5424	Eppendorf
Centrifuge Allegra® X-15R	Beckman Coulter GmbH
Centrifuge Avanti® J-26XP	Beckman Coulter GmbH
ChemiDoc MP Imaging System	Bio-Rad Laboratories
Concentrator plus	Eppendorf
Countess II FL Automated Cell Counter	Thermo Fisher Scientific
DiluPhotometer™ OD ₆₀₀	Implen
Electron microscope JEM-1400	JEOL GmbH
Electron microscope Spirit	FEI, Thermo Fisher Scientific
Electron microscope Titan Krios	FEI, Thermo Fisher Scientific
Electron microscope Talos Arctica	FEI, Thermo Fisher Scientific
Elektroporator Gene Pulser Xcell™ System	Bio-Rad Laboratories
Fluorescence detector RF-20A	Shimadzu
Fluorescence microscope (EVOSFL Auto)	Life Technologies

GloCube® Glow Discharge System for TEM grids and surface modification	Quorum Technologies Ltd
High Vacuum Coating System MED 020	Bal-Tec
Incubator (37 °C)	BINDER
Laminar flow hood CellGard ES NU-480	Nuaire
Magnetic Stirrer MR3000	Heidorf
Microfluidizer Model 110S	Microfluidics Corporation
Mini membrane vacuum pump Laboport®	KNF
Mini-Sub Cell-GT	Bio-Rad Laboratories
Molecular Imager® GelDoc™ XR System	Bio-Rad Laboratories
nanoDSF Prometheus NT.48	nanoTemper Technologies
NGC Discover™ 10 Chromatography System	Bio-Rad Laboratories
Pestle for 1.5 ml microtube	Argos
pH-Meter FE20 Five Easy™	Mettler Toledo
PowerPac 300	Bio-Rad Laboratories
PowerPac HV Power Supply	Bio-Rad Laboratories
Shaker (insect and mammalian cell culture)	INFORS HT
Shaker Excella E24	Eppendorf / New Brunswick
Shaker Innova 44	Eppendorf / New Brunswick
Shaker S4li	Eppendorf
Spectrophotometer DS-11	DeNovix
Sonicator-Sonifier® - Branson	VWR International
ThermoStat Plus	Eppendorf
Table top centrifuge 5417R	Eppendorf
Trans-Blot® Turbo™ Transfer System	Bio-Rad Laboratories
Ultrasonic Bath Sonorex Digitec	Bandelin
Ultracentrifuge Optima™ XPN-80	Beckman Coulter GmbH
Vortex-Genie 2	Scientific Industries
Vitrobot™ cryo plunger	FEI
Water bath TW2	JULABO
Zeiss LSM 800 microscope with Airyscan	Carl Zeiss Microscopy GmbH

4.4 Non-electronic devices

Table 4.4: Non-electronic devices.

Device	Manufacturer
Dounce Homogenizer (2 ml, 5 ml, 15 ml, 40 ml)	Kontes Glass & Co
iBind™ Western Device	Thermo Fisher Scientific
Manual pipettes PIPETMAN (0.2-2 µl, 1-10 µl, 2-20 µl, 20-200 µl, 100-1000 µl)	Gilson
Mini-PROTEAN® Tetra-Cell	Bio-Rad Laboratories
Pipetus®	Hirschmann Laborgeräte
Rotors for centrifugation (JLA 8.100, JA-25.50)	Beckman Coulter GmbH
Rotors for Ultracentrifugation (Ti70, Ti45, TLA-120.1)	Beckman Coulter GmbH

4.5 Media

Table 4.5: Media used for insect and mammalian cell culture.

Medium	Manufacturer
Freestyle™ 293 expression medium (mammalian cell culture)	Gibco™
Opti-MEM reduced serum media	Gibco™
Sf-900 III SFM (insect cell culture)	Gibco™

4.6 Chromatography columns

Table 4.6: Columns for size-exclusion chromatography.

Column	Diameter/Height/Total Volume	Resin	Manufacturer
Superdex 200 Increase 10/300	10 mm/300 mm/24 ml	Superdex 200	GE Healthcare
Superose 6 10/300	10 mm/300 mm/24 ml	Superose 6	GE Healthcare
Superose 6 Increase 10/300	10 mm/300 mm/24 ml	Superose 6	GE Healthcare

4.7 Kitsystems

Table 4.7: Kitsystems for nucleic acid extraction and purification.

Kit	Manufacturer
QIAprep Spin Miniprep Kit	QIAGEN
QIAquick Gel Extraction Kit	QIAGEN
QIAquick PCR Purification Kit	QIAGEN
Plasmid Midi Kit	QIAGEN

4.8 Enzymes and Antibodies

Table 4.8: Antibodies and agarose-bound antibodies.

Antibody	Manufacturer
Anti-Flag M2 (mouse) - monoclonal	Sigma-Aldrich Chemie GmbH
Anti-Flag M2 affinity gel - monoclonal	Sigma-Aldrich Chemie GmbH
Anti-poly histidine (mouse) - monoclonal	Sigma-Aldrich Chemie GmbH
Goat anti-mouse horse radish peroxidase (HRP)	Bio-Rad Laboratories
StrepMAP-Classic (mouse) – monoclonal	Iba

Table 4.9: Enzymes.

Enzyme	Manufacturer
Benzonase Nuclease (≥ 250 units/ μ l)	Sigma-Aldrich Chemie GmbH
HRV 3C protease	Dortmund Protein Facility
PreScission Protease	Dortmund Protein Facility
T4 DNA Ligase	New England Biolabs (NEB)
Trypsin	Sigma-Aldrich Chemie GmbH
FastDigest restriction enzymes	Sigma-Aldrich Chemie GmbH

4.9 Cell lines

Table 4.10: Bacterial, human and insect cell lines.

Organism	Name	Genotype/Remarks
<i>Escherichia coli</i>	Top10F'	ϕ {proAB, lacI ^Q , lacZ Δ M15, Tn10 (Tet ^R)} mcrA, Δ (mrr-hsdRMS-mcrBC), ϕ 80lacZ Δ M15, Δ lacX74, deoR, recA1, λ ⁻ araD139, Δ (ara- leu)7697, galU, galK, rpsL(Str ^R), endA1, nupG
<i>Escherichia coli</i>	BL21(DE3)	F ⁻ dcm ompT hsdS(r – m –) gal λ (DE3)
<i>Escherichia coli</i>	DH10EmBacY	Δ lacX74 recA1 endA1 araD139 Δ (ara, leu)7697 galU galK λ ⁻ rpsL nupG/ bMON14272/pMON7124
<i>Homo sapiens</i>	HEK 293GnT1 ⁻	Adherent human embryonic kidney cell line
<i>Homo sapiens</i>	HEK 293GnT1 ⁻	Suspension human embryonic kidney cell line
<i>Homo sapiens</i>	HEK FreeStyle-293F	Adherent human embryonic kidney cell line
<i>Homo sapiens</i>	HEK FreeStyle-293F	Suspension human embryonic kidney cell line
<i>Spodoptera frugiperda</i>	Sf-9	Immortalised ovarian insect cell line

4.10 Plasmids

Table 4.11: Cloning and expressions vectors (plasmids).

Internal number	Plasmid	Manufacturer/Reference
1952	pACEMam1_2xStrepII-eGFP-His ₈ -Scap_human	This work
1953	pACEMam1_2xStrepII-eGFP-His ₈ -Scap_human	This work
1422	pcDNA3.1 C-eGFP Scap_cat	GenScript Biotech
1421	pcDNA3.1 C-eGFP Scap_chicken	GenScript Biotech
1419	pcDNA3.1 C-eGFP Scap_fruitfly	GenScript Biotech
1423	pcDNA3.1 C-eGFP Scap_pig	GenScript Biotech
1424	pcDNA3.1 C-eGFP Scap_rat	GenScript Biotech
1420	pcDNA3.1 C-eGFP Scap_zebrafish	GenScript Biotech
1442	pcDNA3.1 C-ter eGFP_8xHis PTCH Drosophila	GenScript Biotech
1444	pcDNA3.1 C-ter eGFP_8xHis PTCH Rat	GenScript Biotech
1445	pcDNA3.1 C-ter eGFP_8xHis PTCH1 Zebrafish	GenScript Biotech
1369	pcDNA3.1 N-eGFP Scap_cat	GenScript Biotech
1368	pcDNA3.1 N-eGFP Scap_chicken	GenScript Biotech
1366	pcDNA3.1 N-eGFP Scap_fruitfly	GenScript Biotech
1735	pcDNA3.1 N-eGFP Scap_hamster	This work
1931	pcDNA3.1 N-eGFP Scap_human	GenScript Biotech
1932	pcDNA3.1 N-eGFP Scap_mouse	GenScript Biotech
1370	pcDNA3.1 N-eGFP Scap_pig	GenScript Biotech
1371	pcDNA3.1 N-eGFP Scap_rat	GenScript Biotech
1367	pcDNA3.1 N-eGFP Scap_zebrafish	GenScript Biotech
1389	pcDNA3.1 N-ter eGFP_8xHis PTCH Drosophila	GenScript Biotech
1391	pcDNA3.1 N-ter eGFP_8xHis PTCH Rat	GenScript Biotech
1392	pcDNA3.1 N-ter eGFP_8xHis PTCH1 Zebrafish	GenScript Biotech
1628	pENTR233.1 1_PTCH1 (Hs)	Harvard PlasmID Database
1629	pENTR233.1 1_PTCH2 (Hs)	Harvard PlasmID Database
1639	pENTR233.1 1_PTCH2 (Mm)	Harvard PlasmID Database
1281	pEG BacMam	Eric Gouaux (102)
1641	pEG BacMam_His8-eGFP-3C-PTCH2_Mm	This work
1642	pEG BacMam_PTCH2_Mm-3C-eGFP-His8	This work
1643	pEG BacMam_His8-eGFP-3C-PTCH1_HS	This work
1644	pEG BacMam_PTCH1_HS-3C-eGFP-His8	This work
1645	pEG BacMam_His8-eGFP-3C-PTCH2_HS	This work
1646	pEG BacMam_PTCH2_HS-3C-eGFP-His8	This work
1647	pEG BacMam_His8-eGFP-3C-PTCH2_XI	This work

1648	pEG BacMam_PTCH2_Xl-3C-eGFP-His8	This work
1733	pEG BacMam_Scap(rat)_N-eGFP	This work
1734	pEG BacMam_Scap(pig)_C-eGFP	This work
1779	pEG BacMam_StrepII-His-Scap(pig)-eGFP	This work
1780	pEG BacMam_Scap(pig)-eGFP-His-StrepII	This work
1781	pEG BacMam_FLAG-His-Scap(pig)-eGFP	This work
1782	pEG BacMam_Scap(pig)-eGFP-His-FLAG	This work
1802	pEG BacMam_PTCH2_Mm-3C-eGFP-His8-StrepII	This work
1803	pEG BacMam_PTCH1_HS-3C-eGFP-His8-StrepII	This work
1824	pEG BacMam_PTC1_HS(1-1415)_eGFP-His-StrepII	This work
1825	pEG BacMam_PTC1_HS(1-1365)_eGFP-His-StrepII	This work
1826	pEG BacMam_PTC1_HS(1-1210)_eGFP-His-StrepII	This work
1827	pEG BacMam_PTC1_HS(1-1178)_eGFP-His-StrepII	This work
1828	pEG BacMam_PTC1_HS(53-1447)_eGFP-His-StrepII	This work
1829	pEG BacMam_PTC1_HS(53-1415)_eGFP-His-StrepII	This work
1830	pEG BacMam_PTC1_HS(53-1365)_eGFP-His-StrepII	This work
1831	pEG BacMam_PTC1_HS(53-1210)_eGFP-His-StrepII	This work
1832	pEG BacMam_PTC1_HS(1-1178)_eGFP-His-StrepII	This work
1813	pEG BacMam_StrepII-His-Scap(pig)	This work
1954	pEG BacMam-2xStrepII-eGFP-His ₈ -Scap_human	This work
1955	pEG BacMam-2xStrepII-eGFP-His ₈ -Scap_human	This work
1031	pMSP1D1	Addgene, Stephen Sligar (103)
1032	pMSP1E3D1	Addgene, Stephen Sligar (104)

5 Methods

5.1 Molecular biology

5.1.1 Agarose gel electrophoresis

The separation of DNA fragments of diverse sizes was performed by Agarose gel electrophoresis. Gels containing 1% Agarose (w/v) in 1x Tris-acetate-EDTA (TAE-) buffer were used and supplemented with 0.004% (v/v) Midori Green. The samples were treated with 6x DNA-sample buffer and loaded into the pockets of the gel. A voltage of 80 V for 30 min was applied. The nucleic acid fragments were detected via the dye Midori Green at the excitation range of UV wavelengths.

50x TAE buffer	2 M Tris, 1 M acetic acid, 50 mM EDTA (pH 8)
6x DNA-sample buffer	30% (v/v) glycerol, 0.125% (w/v) bromophenol blue in TAE buffer, 0.125% (w/v) xylene cyanol in TAE buffer

5.1.2 Nucleic acid quantification

The quantification of double stranded DNA (dsDNA) was determined with the spectrophotometer DS-11 (Denovix). 1 μ l of the sample containing nucleic acid fragments was applied and analysed at a wavelength of 260 nm. Here, a concentration of 50 μ g/ml dsDNA correspond to an absorption of 1.

5.1.3 Restriction/Ligation cloning

5.1.3.1 Polymerase chain reaction

The amplification of DNA was achieved by polymerase chain reaction (PCR) corresponding to Mullis *et al.* (105) 25-50 ng of template DNA, 0.5 μ M oligo nucleotides (primer), 1 mM dNTPs and 4 μ l 5x Phusion GC buffer (Thermo Fisher Scientific) were mixed and filled up with ddH₂O to a final reaction volume of 20 μ l. In a final step, 1 U Phusion DNA-polymerase (Thermo Fisher Scientific) was added. The reaction was performed with 30 cycles of denaturation, primer hybridization and elongation (see table 5.1).

Table 5.1: PCR conditions for DNA amplification.

Step	Temperature / °C	Time / s
Initial denaturation	98	120
Denaturation	98	45
Primer hybridization	60	30
Elongation	72	45/kb
Final elongation	72	300
End	8	∞

The amplified DNA fragments were verified using agarose gel electrophoresis (see chapter 5.1.1). The nucleic acid fragments of interest were purified via the QIAquick PCR Purification Kits (QIAGEN).

5.1.3.2 Digestion of plasmid DNA

The restriction reaction of plasmid DNA was performed with FastDigest restriction endonucleases (Thermo Fisher Scientific). 15-20 µg of DNA were incubated with 1 µl of the first restriction enzyme and filled up with 1x FastDigest buffer (Thermo Fisher Scientific) to 30-40 µl for 3 h at 37 °C. 1 µl of the second restriction enzyme was added to the reaction mixture and incubated for further 3 h at 37 °C. A second restriction enzyme was used to prevent self-ligation. The enzymes were inactivated by increasing the temperature to 67 °C for 5 min. The linearized DNA was isolated using the QIAquick PCR Purification Kit (QIAGEN) and verified via agarose gel electrophoresis (see chapter 5.1.1).

5.1.3.3 Digestion of amplified PCR products

The digestion of amplified and purified PCR products was carried out with both FastDigest restriction endonucleases (Thermo Fisher Scientific). The purified PCR products were supplemented with 1 µl of each restriction enzymes in 1x FastDigest Green Buffer (Thermo Fisher Scientific) in 40 µl total volume. The reaction mixture was incubated for 3 h at 37 °C. The digested PCR product was further analysed via agarose gel electrophoresis (see chapter 5.1.1) and isolated via the QIAquick Gel Extraction Kit (QIAGEN).

5.1.3.4 DNA ligation

150 ng of purified and linearized plasmid DNA were added to 300 ng of the purified PCR products and filled up to 10 μ l with 1x T4 ligase buffer. 1 U T4 DNA ligase (NEB) was finally added to start the reaction for 16 h at 25 °C. The ligation product was chemically transformed into competent *E. coli* Top10F' cells (see Chapter 5.1.5).

Ligase buffer 50 mM Tris, 1 mM MgCl₂, 0.1 mM ATP, 1 mM DTT, pH 7.5 (HCl)

5.1.4 Site-directed mutagenesis

The mutation of 1-12 nucleobases was achieved using site-directed mutagenesis. The plasmid DNA was added to each oligo nucleotide (primer) separately. 50 ng of template DNA was mixed with 0.5 μ M oligo nucleotide, 1 mM dNTPs, 8 μ l 5x Phusion HF buffer (Thermo Fisher Scientific) and filled up with water to a final volume of 40 μ l. In a last step, 1 U Phusion DNA-polymerase (Thermo Fisher Scientific) was added. The fragments of the plasmid DNA were amplified using PCR with 30 cycles of denaturation, primer hybridization and elongation (see table 5.2).

Table 5.2: PCR conditions for DNA amplification.

Step	Temperature / °C	Time / s
Initial denaturation	98	120
Denaturation	98	45
Primer hybridization	60	30
Elongation	72	45/kb
Final elongation	72	300
end	8	∞

Both reactions were combined and 1 μ l *DpnI* (FastDigest, Thermo Fisher Scientific) was added to digest methylated nucleic acid fragments. The digestion of plasmid DNA (starting material) was incubated for 3 h at 37 °C and inactivated afterwards by heating the reaction to 96 °C for 5 min. The PCR products were transformed into chemically competent *E. coli* Top10F' cells (see chapter 5.1.5).

5.1.5 Chemical transformation of *E. coli*

The insertion of plasmid DNA into bacteria can be performed via chemical transformation. The bacteria were treated with calcium chloride to make them chemically competent. (106) This method is based on the bivalent cations Ca²⁺ that interact with the negatively charged DNA backbone and with the negatively charged head groups of the lipids from the outer membrane and thereby spatialize the

plasmid DNA with the membranes of the gram-negative bacteria. Finally, the DNA can enter the cell via a heat-shock of the bacteria suspension.

The transformation was carried out with 80 µl chemical competent cells. Here, *E. coli* Top10F' or BL21(DE3) were used. The cells were thawed on ice for 5 min and transferred to 25 ng of plasmid DNA and incubated for additional 10 min on ice. The heat-shock was performed for 45 sec at 42 °C. The cell suspension was cooled for 2 min on ice and finally grown in 0.5 ml Super Optimal Broth-medium treated with 20 mM glucose (SOC) for 1 h at 37 °C. The suspension was spread on LB Agar plates containing the appropriate antibiotic (50 µg/ml kanamycin or 125 µg/ml carbenicillin) and incubated for 16 h at 37 °C.

SOC-medium	2% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.06% (w/v) NaCl, 0.02% (w/v) KCl, 0.2% (v/v) MgCl ₂ x 6 H ₂ O, 0.3% (v/v) MgSO ₄ x 7 H ₂ O, 4% (v/v) 50% glucose solution
LB-agar plates	0.5% (w/v) yeast extract, 1% (w/v) tryptone, 1% (w/v) NaCl, 1.5% (w/v) bacto-agar (pH 7.4)

5.1.6 Electroporation of *E. coli*

Another method to insert plasmid DNA into cells is called electroporation and was used to insert bacmid DNA into electrocompetent bacteria. By the application of a voltage, the cell membrane is destabilized and porous. Furthermore, negatively charged DNA fragments accelerate in the electric field and can enter the cells. (107, 108)

The electroporation of competent *E. coli* (DH10EmBacY) with bacmid DNA was performed according to Wu and Letchworth. (109) 80 µl of electrocompetent cells were thawed for 5 min on ice and incubated with 2.5 µg bacmid DNA for 10 min at 4 °C. The cell suspension was transferred into an electroporation cuvette (Bio-Rad) and an electric pulse of 1.8 kV with 25 µFD and 200 Ω was applied. 500 µl LB medium was added to the reaction mixture and incubated for 16 h at 37 °C. The cells were spread onto LB agar plates containing 50 µg/ml kanamycin, 10 µg/ml tetracyclin, 7 µg/ml gentamycin, 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 100 µg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and incubated for 48 h at 37 °C.

5.1.7 Blue-white screen

The selection of a recombinant generated bacmid was done with the blue-white screening. This method is based on the expression of an enzyme called β-galactosidase. The enzyme hydrolyses X-Gal to galactose and 5-bromo-4-chloro-3-hydroxyindole. In a final step, 5-bromo-4-chloro-3-hydroxyindole gets oxidised by oxygen to 5,5'-dibromo-4,4'-dichloro-indigo which is a blue bye. The encoding

sequence for the β -galactosidase, *lacZ*, is interrupted by a successful integration of the gene of interest (bacmid DNA). The expression of this enzyme is then inhibited and the cells will appear white. Cells without the integration of bacmid DNA will turn blue because of the intact *lacZ* codon. (110)

The LB agar plates were screened for white single cell colonies after incubation of two days. A control LB-agar plate (50 μ g/ml kanamycin, 10 μ g/ml tetracyclin, 7 μ g/ml gentamycin, 1 mM IPTG and X-Gal) was spread with white single cell colonies and a negative control, a blue single cell colony. After an incubation of 48 h at 37 °C the bacmid DNA was isolated with the Plasmid Midi Kit (QIAGEN).

5.1.8 Transfection of insect cells and BacMam virus production

Viruses were generated using the EmBacY system. The genes of interest were cloned into one of the multiple cloning sites (MCS) of the pEG-BacMam vector. The plasmid DNA was then transformed into DH10EmBacY competent cells via electroporation and the bacmid DNA was produced and isolated as described in chapter 4.1.7.

The bacmid DNA was transiently transfected into Sf9 cells with the transfection reagent FuGENE[®] HD (Promega). 1×10^6 cells/per well were filled in a 6-well microtiter plate, filled up with FreeStyle[™] 293 medium supplemented with 5% FBS and settled down for 25 min at 27 °C. Meanwhile, 2 μ g of the isolated bacmid DNA, 10 μ l of FuGENE[®] HD (Promega) were mixed with 100 μ l of Sf-900 III SFM medium and incubated for 10 min at RT. After the incubation times, the transfection mixture was added dropwise to the adherent cells. The transfected insect cells were incubated for four days at 27 °C with minimal light exposure. The virus (P1) was harvested by separation of the virus particles from the insect cells by centrifugation at 1.500 x g for 10 min. The virus (supernatant) was filtered through 0.22 μ m Durapore[®] Membrane Filter (47 mm membrane, Millipore) and stored at 4 °C until further usage. The virus amplification (generation P2 and P3) was performed with 50 to 500 ml of Sf9 cells (1×10^6 cells/ml). The insect cells were transduced with a multiplicity of infection (MOI) of 0.01 with P1 and incubated for up to three days at 27 °C until all cells showed yellow fluorescence protein (YFP) fluorescence. The viruses P2 and P3 were separated from the cells by centrifugation at 1.500 x g for 10 min, filtered through 0.22 μ m Filtropur V50 (Sarstedt) and stored at 4 °C with minimal light exposure until further usage. Viruses P2 and P3 were used for the transfection of large-scale protein expression in suspension mammalian cells after the determination of their titer (see chapter 5.2.2).

5.1.9 Transfection of human cell lines

The insertion of plasmid DNA into adherent mammalian cells were performed by transient transfection with lipofectamine[®] 3000 (Invitrogen, Thermo Fisher Scientific). In an aqueous solution the lipid molecules of the lipofectamines are forming liposomes. By the addition of lipofectamine with DNA,

the nucleic acid molecules can insert into the inside of the liposomes. Furthermore, the liposomes contain a cationic surface by which they can interact with the negatively charged plasma membrane of the mammalian cells.

0.5×10^6 HEK293 F or GnT1⁻ cells were plated in a 10 cm petri dish and incubated for 24 hours at 37 °C with 5% CO₂. The medium was exchanged with Dulbecco's modified Eagle's medium (DMEM, Gibco™) supplemented with 5% FBS. 10 µg plasmid DNA were mixed with 125 µl Opti-MEM (Sigma-Aldrich) and added to a reaction mixture of 10 µl lipofectamine[®] 3000 additive (Invitrogen, Thermo Fisher Scientific), 7 µl lipofectamine[®] 3000 reagent (Invitrogen, Thermo Fisher Scientific) and 125 µl Opti-MEM. The transfection mixture was incubated for 15 min at RT and added dropwise to the adherent mammalian cells.

5.1.10 DNA sequencing

The new generated plasmid DNA, either via restriction/ligation cloning (see chapter 5.1.3) or mutagenesis PCR (see chapter 5.1.4), was analysed using sequencing reaction according to Sanger. (111) Fluorescence labelled dideoxynucleotides interrupt the elongation of the DNA polymerase due to the absence of 3'-hydroxy group at the synthesized strand. The sequence will be determined by the electrophoretic separation of the DNA fragments and the detection of the last fluorophore coupled nucleobase.

200 ng of plasmid DNA were mixed with 0.03 pmol of each primer (forward and reverse) and filled up with ddH₂O to 15 µl and analysed at Microsynth Seqlab (Single-Tube Sequencing - Economy Run).

5.2 Cell biology

5.2.1 Eukaryotic cell maintenance

S. frugiperda (Sf9) cells were maintained in Sf900 III serum free medium (Gibco™) at 27 °C while shaking with minimal light exposure. The cells were cultured between $0.8 - 4 \times 10^6$ cells/ml. When the cells reached a maximum density of 4×10^6 cells/ml they were divided and diluted (splitted) to 0.8×10^6 cells/ml.

Adherent HEK293S T and GnT1⁻ cells were cultured in DMEM (Gibco™) containing 10% FBS at 37 °C with 5% CO₂ and minimal light exposure. The adherent mammalian cells were plated to 70 % confluency in petri dishes (10 cm) or 6-well culture dishes and splitted every 48 to 72 h with fresh medium.

The suspension HEK293 cells were maintained in FreeStyle™ 293 medium supplemented with 5% FBS according to the manufacturer's protocol. The suspension cells were grown at 37 °C with continuous 8% CO₂ while shaking with minimal light exposure. After 48 to 72 h the cells were splitted to 0.8×10^6 cells/ml with fresh medium.

The cell vitality was verified every 48 to 72 h by mixing 1:1 (v/v) of cell suspension with Trypan blue (FLUKA). The cells were counted and analyzed with the Countess II FL Automated Cell Counter (Thermo Fisher Scientific).

Adherent eukaryotic cells were furthermore analyzed by a conventional light microscope (EVOSFL Auto - Life Technologies) using the bright field channel. The cells were analysed by changes in their morphology and confluency.

5.2.2 Virus plaque assay

The virus titer was determined by using the plaque assay. The plaque assay allows a direct quantification of infectious virions via counting of discrete plaques in cell culture. (112) 1×10^5 Sf9 cells were seeded into each well of a 96 well-plate (black, Thermo Fisher Scientific) and attached to the bottom of the well for 25 min at 27 °C. Meanwhile the dilution series of the virus was performed in triplicate ranging from 10^{-1} to 10^{-8} (360 μ l medium + 40 μ l virus). The supernatant of each well was removed and 100 μ l of each virus dilution was added. The viruses were incubated for 72 hours at 27 °C and analysed with a fluorescence microscope (EVOSFL Auto - Life Technologies). Here, the amount of foci was screened for each well and the dilutions with less than 10 foci were counted. The viral titer was calculated using equation 1.

Equation 1: $(\text{average number of foci}) \times (\text{dilution factor}) \times 10 = \text{virus titer}$

5.3 Biochemical Methods

5.3.1 Semi-denaturing SDS-PAGE

The sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to separate proteins according to their molecular weight. The method is based on SDS which is added to the samples to denature the proteins that move in an electric field through a porous gel to the anode. SDS is an anionic tenside that unfolds proteins due to hydrophobic interactions with the hydrophobic dodecyl rest of the SDS and the hydrophobic areas of a protein that are located inside a folded protein. The protein-SDS conjugates are negatively charged and move to the anode. The separation of the proteins is feasible because of the constant mass-to-charge ratio of their specific molecular weight. In contrast to the standard denaturing SDS-PAGE, the semi-denaturing SDS-PAGE was used for protein samples containing a fluorescent reporter tag, like the enhanced green fluorescent protein (eGFP). The protein samples were treated with less SDS-concentrated sample buffer and not heated before loading onto the gel. Here, the proteins keep their folded structure in order to allow an analysis of the fusion proteins.

The separation of proteins via semi-denaturing SDS-PAGE was performed with the 4 – 15% Mini-PROTEAN TGX Stain Free Precast Gels (Bio-Rad) according to Lämmli. (113) The samples were treated with 4x SDS sample buffer and 20 µl were loaded onto the gels. The electrophoresis was started in a Mini-PROTEAN Tetra system (Bio-Rad) filled with SDS-PAGE running buffer for 40 min at 200 V. The gels were analysed with a ChemiDoc MP imaging system (Bio-Rad). The fluorescent labelled proteins were excited with light containing the appropriate wavelength and the emitted light was detected using the correct filters (see table 5.3).

Table 5.3: Imaging settings of fluorescent labelled proteins.

Fluorophore	Excitation	Emission
Blue fluorescent protein (BFP)	Red Epi Illumination	695/55 filter
EGFP	Blue Epi Illumination	530/28 filter
MCherry	Green Epi Illumination	605/50 filter

Non-fluorescent labelled proteins were detected either via Coomassie staining or the stain free imaging technology. The latter technology is based on the activation of trihalo compounds that are added into the gels. By a short photoactivation trihalo enhances the fluorescence of the amino acid tryptophan that makes the proteins fluorescent in the gel. The fluorophores are covalently bound to the proteins which allows a repeatedly imaging without staining and destaining steps.

The second visualisation method is based on the dye Coomassie brilliant blue R-250 that attaches to the side chains of the basic amino acids. The gels were stained by incubating the gels under gentle agitation for 1 hour in the Coomassie staining solution. Subsequently, the gels were transferred into destaining solution to remove excess dye.

4x SDS sample buffer	0.25 M Tris/HCl, 6% (w/v) SDS, 40% (v/v) glycerol, 0.1 M DTT, 0.5% (w/v) bromophenol blue
SDS running buffer	0.25 M Tris, 1 M glycine, 1% (w/v) SDS
Coomassie staining solution	50% (v/v) ethanol, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie brilliant blue R-250
Destaining solution	5% (v/v) ethanol, 10% (v/v) acetic acid

5.3.2 Western Blot

For qualitative analysis of a protein of interest (POI) that was previously analysed via SDS-PAGE, western blotting was performed. This method is based on a transfer of proteins from a gel to an immobilized membrane and can be visualised via immunodetection. The proteins are able to move from the SDS-gel to a membrane by application of an electric potential. Finally, the proteins can be detectable by using specific antibodies. (114) Here, the visualisation was performed by using secondary

antibodies conjugated to an enzyme, the horseradish peroxidase (HRP). This enzyme catalyses the oxidation of luminol so that light will be emitted and the proteins can be visualised.

The transfer of the proteins from the SDS-gel (see chapter 5.3.1) to the PVDF membrane was performed with the Trans-Blot[®] Turbo[™] Transfer system (Bio-Rad) for 7 minutes at 25 V (turbo protocol). The blocking, washes and antibody incubation steps were performed by using the iBind Western Device (Thermo Fisher Scientific) according to the manufacturer's protocol in 2.5 hours at RT. Finally, 2 ml of Plus ECL (Perkin Elmer) were added on top of the membrane and incubated for 1 min. The detection was done with the ChemiDoc MP imaging system (Bio-Rad).

5.3.3 Heterologous expression of proteins in adherent mammalian cells

The plasmid DNA, containing the gene of interest and a fluorescence reporter tag, was transiently transformed into adherent HEK293 cells (see chapter 5.1.9). After the transfection, the cells were grown for up to 72 hours at 37 °C with 5% CO₂. The expression of the POI was observed and analysed with a fluorescence microscope (EVOSFL Auto - Life Technologies) every 24 hours. Here, the cells were screened for general health using the bright field channel and for the amount of over-expressed proteins using the appropriate fluorescence channel. Depending on that, the cells were harvested by centrifugation at 1000 x g for 5 minutes at 4 °C. The cell pellet was washed three times with 5 ml cold phosphate-buffered saline (PBS) buffer, shock frozen in liquid nitrogen (N₂) and stored at -20 °C until further usage.

PBS buffer 3 mM KCl, 150 mM NaCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7.4

5.3.4 Detergent screen

The solubilization, that describes the release of the membrane protein from its native environment, the lipid bilayer, is one of the most challenging steps in structure analysis of membrane proteins. Nevertheless, the POI needs to be in high purity as well as in an aqueous solution for structural studies and the solubilization step has remained indispensable. Integral membrane proteins that are integrated into the lipid bilayer can get solubilized by using detergents. Detergents are amphiphilic molecules composed of a hydrophilic head and a hydrophobic tail. The latter is able to bind to the hydrophobic areas of a transmembrane protein and the hydrophilic head is protruded to the outside and attached to the aqueous environment. The amphiphilic molecules build micelles in aqueous solutions when used in concentrations above their critical micelle concentration (CMC). By using detergents with concentrations above their CMC, the detergent molecules first get integrated into the lipid bilayer. Then, detergent-detergent interactions destabilize the bilayer and can solubilize the integral membrane protein. (115, 116) Detergents differ in their length of the hydrophobic tail and the

head group, which can be polar (ionic, isoelectric) or nonpolar. Depending on this, they differ in their solubilization efficiency but can also disrupt the proteins native conformation.

Therefore, different kinds of detergents were screened for an efficient solubilization of the membrane protein of interest (see table 5.4).

Table 5.4: Detergents used in the detergent screen, their CMC and solubilization concentration (C_{final}).

Detergent	CMC / % (w/v)	C_{final} / % (w/v)
n-Decyl β -D-maltopyranoside (DM)	0.0870	1.0
n-Dodecyl β -D-maltopyranoside (DDM)	0.0087	1.0
DDM/ Cholesteryl hemisuccinat (CHS)	-	1.0:0.1
n-Octyl β -D-glucopyranoside (OG)	0.5300	2.0
n-Undecyl-phosphocholine (Fos-Choline 11)	0.0620	1.0
Fos-Choline 13	0.0270	1.0
4-Cyclohexyl-1-heptyl- β -D-maltoside (Cymal-4)	0.3700	1.0
6-Cyclohexyl-1-heptyl- β -D-maltoside (Cymal-6)	0.0099	1.0
Lauryl dimethylamine-N-oxide (LDAO)	0.0230	1.0
Lauryl maltose neopentyl glycol (LMNG)	0.0010	0.7
3-[(3-Cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS)	0.4900	0.7
n-Decyl-N,N-dimethyl-3-ammonio-1-propanesulfonate (Anzergent 3-10)	1.2000	2.0
n-Dodecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate (Anzergent 3-12)	0.0940	1.0
Buffer (without detergent)	-	-

The adherent mammalian HEK293 cells expressing the membrane protein of interest were thawed on ice and directly treated with 5 ml PBS. The cell lysis was performed with a cell douncer by shear forces (Kontes Glass & Co). 250 μ l of the cell suspension was mixed with 250 μ l of each detergent buffer and incubated for 1 hour at 4 °C while rotation. The isolation of the solubilized fraction was done by ultracentrifugation for 1h at 100.000 xg at 4 °C (TLA-120.1). The efficiency of the solubilization was analysed by semi-denaturing SDS-PAGE (see chapter 5.3.1) and fluorescence-detection size-exclusion chromatography (FSEC) (see chapter 5.3.5) of the supernatant.

Solubilization buffer 200 mM NaCl, 30 mM HEPES, 5% glycerol, 1 mM TCEP (pH 7.5)

5.3.5 Fluorescence-detection size-exclusion chromatography

In a final step, the fluorescence tagged membrane proteins were analysed by FSEC. Each step was performed at 4 °C using the NGC Discover™ 10 chromatogram system equipped with an autosampler C-96 (Bio-Rad) and a fluorescence detector RF-20A (Shimadzu). For FSEC, a Superose 6 increase 10/300 GL (GE Healthcare) was used and equilibrated with 30 ml SEC running buffer with a flow rate of 0.5 ml/min. 400 µl of the samples were injected automatically using the autosampler when multiple samples were analysed. For each run 2 ml of equilibration, 30 ml elution with a flow rate of 0.35 ml/min and 2 ml washing in reverse flow was set up for each run. The detection of the POI was observed via measuring the absorbance at 280 nm (for all proteins) and the appropriate wavelength for the fluorescent fusion proteins eGFP, mCherry and eBFP (see table 5.5). The chromatograms were analysed for their monodispersity and a single gaussian-shaped peak.

Table 5.5: Excitation and Emission wavelengths for the detection of eBFP, eGFP and mCherry fusion proteins at FSEC runs.

Fluorophore	Excitation wavelength / nm	Emission wavelength / nm
eBFP	383	445
eGFP	475	510
mCherry	587	610

SEC running buffer 200 mM NaCl, 30 mM HEPES, 5% glycerol, 1 mM TCEP, 0.05% (w/v) DDM (pH 7.5)

5.3.6 Heterologous expression of proteins in suspension mammalian cells

5.3.6.1 Heterologous expression of Scap and Patched1 constructs

For expression cultures, suspension F or GnTI⁻ cells were amplified in free style medium supplemented with 2% FBS (Gibco™) between 2 to 2.5 x 10⁶ cells/ml.

First, a test expression was performed for each membrane protein of interest. Here, both mammalian cell lines were infected with either 4.5% (v/v) or 9% (v/v) of P2 or P3 virus (see chapter 5.1.8) and incubated for 6 hours at 37 °C with 8% CO₂ in a shaker (130 rpm, INFORS HT). Then 10 mM sodium butyrate were added for induction of gene expression for up to 72 hours at 30 °C with 8% CO₂ in a shaker (130 rpm, INFORS HT). Every 24 hours 1 ml of the cell suspension was harvested and analysed for gene expression. First, 10 µl were used for the validation of cell vitality and amount of expressed fluorescent reporter gene if used via the Countess II FL Automated Cell Counter (Thermo Fisher Scientific). The cells were isolated from the supernatant by centrifugation for 5 min at 1500 xg. The pellet was flash-frozen in N₂(l) and stored at -20 °C until further usage or directly used for an analysis by semi-denaturing SDS-PAGE (see chapter 5.3.1). Here, the pellet was homogenized in PBS buffer and

the cells were lysed with a pestle (Argos) by shearing forces. The lysed and homogenized samples were then applied onto SDS-Gels mini-PROTEAN TGX Stain Free Precast gels (4-15%, 15 well, Bio-Rad). The expression of Scap and Patched1 constructs were then successively increased up to 20 l expression cultures of HEK293F cells. The cells were infected with 9% (v/v) of P2 or P3 and incubated for 6 hours at 37 °C with 8% CO₂ at 130 rpm (INFORS HT shaker). Then 10 mM sodium butyrate were added for induction of gene expression for 48 hours post-induction at 30 °C with 8% CO₂ in a shaker (130 rpm, INFORS HT). The cells were harvested at 1500 xg for 5 min at 4 °C, isolated from the medium, flash-frozen in N₂(l) and stored at -80 °C until further usage.

PBS buffer 3 mM KCl, 150 mM NaCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7.4

5.3.6.2 Heterologous expression of Scap and SREBP complex

The expression of the heterodimer Scap and SREBP was successively increased up to 20 l expression cultures of HEK293GnT1⁻ cells. The cells were amplified between 2 to 2.5 x 10⁶ cells/ml in free style medium supplemented with 2% FBS (Gibco™) and infected with 4.5% (v/v) P2 virus expressing Scap and 4.5% (v/v) P2 virus expressing SREBP. After 6 hours of cultivation at 37 °C with 8% CO₂ at 130 rpm (shaker INFORS HT) the gene expression was induced by the addition of 10 mM sodium butyrate and the temperature was decreased to 30 °C. The cells were harvested at 1500 xg for 5 min at 4 °C after 48 hours of gene expression of Scap and SREBP. The cell pellet was flash-frozen in N₂(l) and stored at -80 °C.

5.3.6.3 Heterologous expression of Scap, SREBP and Insig1 complex

The gene expression of the heterotrimer composed of Scap, SREBP and Insig1 was expressed in 2 to 2.5 x 10⁶ cells/ml HEK293GnT1⁻ cells. The cells were grown in free style medium supplemented with 2% FBS (Gibco™) and infected with 3% (v/v) P2 virus expressing Scap, 3% (v/v) P2 virus expressing SREBP and 3% P2 virus expressing Insig1. The complex was expressed for 6 hours at 37 °C with 8% CO₂ at 130 rpm (shaker INFORS HT) and the expression was induced with 10 mM sodium butyrate. After induction, the temperature was decreased to 30 °C and the genes were expressed for 48 hours. The cells were harvested using centrifugation at 1500 xg for 5 min at 4 °C and flash-frozen with N₂(l) and stored at -80 °C until further usage.

5.3.7 Cell lysis of mammalian cells

The lysis of suspension HEK293 F or GnT1⁻ cells was performed by sonication. Sonication physically disrupt the cells by using pulsed, high frequency sound waves. The cells were thawed on ice and homogenized with lysis buffer 1:1 (cell pellet:lysis buffer) (w/v) and incubated for 30 min at 4 °C while smooth rotation. The sonication was performed with a sonicator (Sonifier[®] – Branson, VWR) and solid horn (19 mm diameter) 3 times at 40% duty cycle, 3.5 output control for 2 minutes with 2 minutes pause in between at 4 °C. The lysate was isolated from the cell garbage by low speed centrifugation at 10.000 rpm for 20 minutes at 4 °C (centrifuge Avanti[®] J-26XP, JA 25.50 - Beckman Coulter).

Lysis buffer Scap	150 mM NaCl, 30 mM Tris (pH 8.4), 5% glycerol, 1 mM TCEP, 1 μM Calpain III, 2 μM Leupeptin, 1 μM Pepstatin A, 0.1 mM Pefabloc, 0.05 μM Aprotinin, Benzonase (Nuclease HC)
Lysis buffer (complexes)	150 mM NaCl, 30 mM Tris (pH 8.0), 5% glycerol, 1 mM TCEP, 1 μM Calpain III, 2 μM leupeptin, 1 μM pepstatin A, 0.1 mM Pefabloc, 0.05 μM Aprotinin, Benzonase (Nuclease HC)
Lysis buffer (Patched1)	150 mM NaCl, 30 mM HEPES (pH 7.5), 5% glycerol, 1 mM TCEP, 2 μM leupeptin, 1 μM pepstatin A, 0.1 mM Pefabloc, 0.05 μM Aprotinin, Benzonase (Nuclease HC)

5.3.8 Membrane preparation

The membranes of the isolated cell lysate (see chapter 5.3.7) were separated by ultracentrifugation (Optima™ XPN-80, Beckman Coulter) at 40.000 rpm (Ti 45, Beckman Coulter) or 60.000 rpm (Ti70, Beckman Coulter) for 1 hour at 4 °C. The membrane fraction (the pellet) was homogenized in 50 ml lysis buffer using a dounce homogenizer (Kontes Glass & Co). Please refer to chapter 5.3.7 for the compositions of the lysis buffers for Scap, Scap containing complexes and Patched1 samples, respectively. The membranes were either directly used for further membrane protein purification or flash-frozen in N₂(l) and stored at -80 °C until further usage.

5.3.9 Solubilization

The solubilization of the membrane proteins was performed using the detergent that was most efficient in the detergent screen (see chapter 5.3.4). Therefore, Scap, Scap/SREBP, Scap/SREBP/Insig1 and Patched1 proteins were solubilized in a detergent mixture of DDM and CHS (10:1) or LMNG/CHS. The homogenized membranes (see chapter 5.3.8) were either thawed on ice or directly used for solubilization. The membrane proteins and protein complexes were solubilized for one hour at 4 °C in

detergent buffer containing DDM/CHS or LMNG/CHS under gentle rotation in a 1:1 ratio (membranes:solubilization buffer, (v/v)). The solubilized membrane proteins were isolated from the membrane fraction by ultracentrifugation for one hour at 4 °C with 40.000 rpm (Ti 45, Beckman Coulter) or 60.000 rpm (Ti70, Beckman Coulter) in an Optima™ XPN-80 (Beckman Coulter). The supernatant was used for further purification of the membrane protein (complex) of interest via affinity chromatography (see chapter 5.3.10).

Solubilization buffer Scap (DDM/CHS)	150 mM NaCl, 30 mM Tris (pH 8.4), 5% glycerol, 1 mM TCEP, 2% DDM/0.2% CHS (w/v)
Solubilization buffer Scap (LMNG/CHS)	150 mM NaCl, 30 mM Tris (pH 8.4), 5% glycerol, 1 mM TCEP, 2% LMNG/0.2% CHS (w/v)
Solubilization buffer complexes (LMNG/CHS)	150 mM NaCl, 30 mM Tris (pH 8.0), 5% glycerol, 1 mM TCEP, 2% LMNG/0.2% CHS (w/v)
Solubilization buffer Patched1 (DDM/CHS)	150 mM NaCl, 30 mM HEPES (pH 7.5), 5% glycerol, 1 mM TCEP, 2% DDM/0.2% CHS (w/v)

5.3.10 Affinity chromatography

The affinity chromatography is a method used for the isolation of specified molecules from a complex mixture. The method is based on a highly specific and mostly reversible interaction of two molecules. One of the interaction partners, mostly the affinity ligand, is part of a stationary phase and the molecule of interest is part of a mobile phase. (117)

5.3.10.1 Immobilized metal ion affinity chromatography of poly-Histidine modified Proteins

At an immobilized metal ion affinity chromatography (IMAC), the POI including the affinity tag is retained at the column resin by an interaction with an immobilized metal ion. Here, the POI is conjugated with a poly-Histidine tag that forms complexes with the chelated metals, the divalent Ni^{2+} . The divalent cations are furthermore bound to the nitrilotriacetic acid (NTA) and builds the stationary phase with the Ni^{2+} ions.

The Ni^{2+} -NTA beads were washed three times each with two column volumes (CV) ddH₂O and two CV of the general buffer. The equilibrated beads were applied to 10 CV of the solubilized fraction (see chapter 5.3.9) and incubated for one hour at 4 °C under gentle rotation. From this step on each flow, including sample application, washing and elution steps were performed at 4 °C by gravity flow. In a next step, the beads and the immobilized solubilisate were applied to an Econo-Pac® Chromatography Column (Bio-Rad). The immobilized cell components were isolated from the non-immobilized part by

gravity flow. Unspecific bound cell components were separated by the addition of 10 times 1 CV general buffer containing 30 mM imidazole. A second washing step was performed to elute other contaminants by an ATP-wash with 5 times 1 CV ATP washing buffer. In a final step, the immobilized proteins were eluted by the addition of 6 times 0.5 CV general buffer supplemented with 150 mM imidazole. All eluted fractions were analysed by semi-denaturing SDS-PAGE (see chapter 5.3.1) for presence of the POI. All eluted fractions containing the purified POI were combined and used for further purification steps, like a second affinity chromatography (see chapter 5.3.10.2) or size-exclusion chromatography (see chapter 5.3.15).

General buffer (Scap)	150 mM NaCl, 30 mM Tris (pH 8.4), 5% glycerol, 1 mM TCEP, 0.05% DDM/0.005% CHS (w/v)
General buffer (Patched1)	150 mM NaCl, 30 mM HEPES (pH 7.5), 5% glycerol, 1 mM TCEP, 0.05% DDM/0.005% CHS (w/v)
ATP washing buffer (Scap)	100 mM NaCl, 30 mM Tris (pH 8.4), 40 mM KCl, 10 mM MgSO ₄ x 7H ₂ O, 2 mM ATP, 5% glycerol, 1 mM TCEP, 0.05% DDM/0.005% CHS (w/v)

5.3.10.2 Immunoaffinity chromatography of Flag modified proteins

The POI complexes containing a Flag purification tag were purified by immunoaffinity chromatography using an anti-Flag affinity gel that represents the stationary phase.

All steps for the immunoaffinity chromatography were performed at 4 °C. The Anti-Flag M2 affinity gel beads (Sigma-Aldrich Chemie GmbH) were separated from the storage buffer by centrifugation at 1000 xg for 3 minutes (Allegra® X-15R, Beckman Coulter GmbH) and equilibrated by the application of two times 2 CV general buffer. The solubilisate was incubated with the equilibrated beads for 1 hour at 4 °C under gentle rotation and separated by applying the Flag beads and solubilization mixture to an Econo-Pac® Chromatography Column (Bio-Rad). Unspecific bound cell compartments were removed from the stationary phase by 10 washing steps with 1 CV general buffer. The purified POI complexes were eluted by the addition of 6 times 0.5 CV general buffer containing 100 µg/ml 3xFLAG peptide (APEX-BIO). All fractions were analysed by semi-denaturing SDS-PAGE (see chapter 5.3.1) for the presence of purified POI complexes. The fractions containing the purified POI complexes were selected and combined. The combined fractions were used for a further affinity chromatography to isolate a proper POI complex (see chapter 5.3.10.3).

General buffer (complexes)	150 mM NaCl, 30 mM Tris (pH 8.0), 5% glycerol, 1 mM TCEP, 0.05% DDM/0.005% CHS (w/v) or 0.03% LMNG/0.03% CHS (w/v)
----------------------------	--

5.3.10.3 Affinity chromatography of StrepII modified proteins

StrepII modified membrane proteins were isolated from contaminations by affinity chromatography. Here, the stationary phase is composed of Strep-Tactin that specifically binds proteins containing a StrepII affinity tag.

All steps of the affinity chromatography were performed at 4 °C. The Strep-Tactin® Superflow® 50% suspension beads (iba) were separated from the storage buffer by centrifugation at 1000 xg for 3 minutes (Allegra® X-15R, Beckman Coulter GmbH) and washed three times with 2 CV ddH₂O. The beads were equilibrated two times with 2 CV of the appropriate general buffer. The eluted fractions of chapter 4.3.10.1 or 4.3.10.2 were applied to the equilibrated beads and incubated for 1 hour at 4 °C while gentle rotation. The beads including the immobilized fraction were isolated from the buffer and contaminations by gravity flow after applying the fractions to an Econo-Pac® Chromatography Column (Bio-Rad). Unspecific bound compartments were separated by a washing step using 10 times one CV general buffer. In a final step, the POI was eluted by using 6 times 0.5 CV general buffer supplemented with 2.5 mM desthiobiotin. The fractions of each step were collected and analysed by semi-denaturing SDS-PAGE (see chapter 5.3.1) and the fractions were combined when purified POI was present. The combined fractions were concentrated using an Amicon Ultra-1 (Merck Millipore) up to 500 µl total volume and used for size-exclusion chromatography (see chapter 5.3.15).

General buffer (Scap)	150 mM NaCl, 30 mM Tris (pH 8.4), 5% glycerol, 1 mM TCEP, 0.05% DDM/0.005% CHS (w/v)
General buffer (Patched1)	150 mM NaCl, 30 mM HEPES (pH 7.5), 5% glycerol, 1 mM TCEP, 0.05% DDM/0.005% CHS (w/v)
General buffer (complexes)	150 mM NaCl, 30 mM Tris (pH 8.0), 5% glycerol, 1 mM TCEP, 0.05% DDM/0.005% CHS (w/v)

5.3.11 Detergent exchange

For structural studies of membrane proteins, the choice of detergent can be the pivotal factor for success. The detergent that is used for solubilization might have an impact on the proteins native conformation and therefore it can destabilize the POI while purification. Furthermore, detergents are amphipathic molecules that are also known as surfactants due to their ability of decreasing the surface tension of water. This has a major impact on the vitrification process of the sample during sample preparation for cryo-EM. These two important key facts need to be considered while purification of a membrane protein for structural studies. Therefore, an exchange of detergent during purification was performed in order to stabilize the POI and to obtain a homogeneous ice and sample distribution while vitrification process of the protein sample.

Here, the POI and the complexes were solubilized using the most efficient detergent that was verified by the detergent screen (see chapter 5.3.4) and exchanged to a detergent during affinity chromatography if necessary (see chapter 5.3.10). The solubilized POI was applied onto the affinity beads (either Ni²⁺-NTA or Anti-Flag M2 affinity gel beads) and after an incubation of 1 hour at 4 °C while gentle rotation the detergent exchange was performed. From this step on all washing and elution buffers were supplemented with the appropriate exchange detergent (see table 5.6) and the affinity chromatography was executed as described in the affinity chromatography chapter, respectively (see chapter 5.3.10.1 for poly-Histidine modified proteins and 5.3.10.2 for Flag modified proteins).

Table 5.6: Solubilization and exchange detergents for Scap that were used during membrane protein purification.

POI	Solubilization detergent	Exchange detergent
Scap	DDM/CHS	Digitonin
		LMNG/CHS
		GDN

5.3.12 Reconstitution into lipid nanodiscs

A nearly native environment for membrane proteins can be mimicked by their reconstitution into lipid nanodiscs after solubilisation in detergent. Lipid nanodiscs are small discs composed of lipids that are surrounded by a protein, the membrane scaffold protein (MSP), that acts like a belt and holds the disc together. According to the number of transmembrane helices the size of MSPs can be easily adjusted. For the reconstitution of Scap alone the scaffold proteins Msp1D1 (Addgene, 20061) and Msp1E3D1 (Addgene, 20066) in combination with POPC and different cholesterol and CHS concentrations (see table 5.7) were used. (103, 104)

Table 5.7: Lipid composition for reconstitution of Scap and the Scap/Insig1 complex into lipid nanodiscs with the scaffold proteins Msp1D1 and Msp1E3D1.

POI	MSP	Lipid composition
Scap	Msp1D1 (103)	POPC/5 mol% cholesterol
		POPC/10 mol% cholesterol
		POPC/10 mol% CHS
	Msp1E3D1 (104)	POPC/5 mol% cholesterol
		POPC/10 mol% cholesterol
		POPC/10 mol% CHS

The lipids were mixed with the appropriate cholesterol or CHS concentration and furthermore added to the scaffold proteins Msp1D1 or Msp1E3D1 in 100:1 molar ratio (450 µM lipids and 45 µM MSP). The POI was finally added in a molar ratio of 1:1 to the reaction mixture and incubated for 2 hours at

4 °C while gentle rotation. The detergent was either removed by dialysis or by the addition of pre-equilibrated Biobeads SM2 (Bio-Rad). The dialysis was performed in a Spectra/Por membrane tubing (MWCO 100 kDa, Carl Roth) over night at 4 °C in 5 l of the appropriate general buffer under gentle rotation. Another method for detergent removal was the addition of Biobeads. The Biobeads SM2 (Bio-rad) were pre-equilibrated with the appropriate general buffer and added to the reaction mixture in a ratio of 30:1 (w/w) to the used detergent. The removal of the detergent via Biobeads was performed for 16 hours at 4 °C while gentle rotation. The reconstituted protein samples were concentrated up to 500 µl total volume using an Amicon Ultra-15 (MWCO 100.000) at 4 °C and separated from Biobeads SM2 via low speed centrifugation at 1000 xg for 5 min at 4 °C.

General buffer (Scap)	150 mM NaCl, 30 mM Tris (pH 8.4), 5% glycerol, 1 mM TCEP
General buffer (Scap/SREBP)	150 mM NaCl, 30 mM Tris (pH 8.0), 5% glycerol, 1 mM TCEP

5.3.13 Reconstitution into amphipols

Amphipols are amphiphilic polymers that contain a hydrophilic backbone with hydrophobic chains. In addition to lipid nanodiscs, amphipols provide a more native environment for solubilized membrane proteins compared to detergents.

For the reconstitution of the detergent solubilized POI and protein complexes into amphipols the affinity chromatography purified protein samples were used. Here, the protein samples were concentrated up to a total volume 3 ml as described before. The amphipols A₈₋₃₅ (Bio-Rad) were added to the protein sample in a 5:1 (A₈₋₃₅:protein) molar ratio and incubated for two to three hours at 4 °C while gentle rotation. Following, the detergents were removed either via the addition of Biobeads SM2 (Bio-Rad) or by dialysis (see chapter 5.3.12).

General buffer (Scap)	150 mM NaCl, 30 mM Tris (pH 8.4), 5% glycerol, 1 mM TCEP
General buffer (Scap/SREBP)	150 mM NaCl, 30 mM Tris (pH 8.0), 5% glycerol, 1 mM TCEP
General buffer (Patched1)	150 mM NaCl, 30 mM HEPES (pH 7.5), 5% glycerol, 1 mM TCEP

5.3.14 Size-exclusion chromatography

The size-exclusion chromatography (SEC) is based on the different diffusion characteristics of biomolecules through a porous matrix. Here, the molecules get sorted dependent on their hydrodynamic radius. Molecules with a small hydrodynamic radius diffuse into the porous matrix and

get retarded compared to molecules with a larger hydrodynamic radius. The porous matrix is not accessible for molecules with a larger hydrodynamic radius. (118)

The SEC was performed as a final step for a purification of a membrane protein or a membrane protein complex. All fractions containing the POI or the protein complex of interest were combined and concentrated up to a final volume of 500 μ l using an Amicon Ultra-15 (MWCO 100.000) at 4 °C. All steps for the SEC were performed at 4 °C with a Superpose 6 Increase 10/300 GL column (GE Healthcare). The column was washed in reverse flow with 50 ml of ddH₂O at a flow rate of 0.5 ml/min and equilibrated with 30 ml of the SEC buffer at a flow rate of 0.4 ml/min. The 500 μ l of the concentrated sample were injected automatically with 1 ml of SEC buffer at a flow rate of 0.3 ml/min. The elution of the POI was performed at a flow rate of 0.35 ml/min with 30 ml of SEC buffer. The detection of the POI was observed via measuring the absorbance at 280 nm (for all proteins) and the appropriate wavelength for the fluorescent fusion proteins eGFP, mCherry and eBFP (see table 5.8). The chromatograms were analysed for their monodispersity and a single gaussian-shaped peak. The fractions were analysed for the presence of the POI via semi-denaturing SDS-PAGE (see chapter 5.3.1). All fractions containing the POI were combined and concentrated up to 5 mg/ml using an Amicon Ultra-15 (MWCO 100.000) at 4 °C. The purified samples were used for structural and biophysical analysis.

Table 5.8: Excitation and emission wavelengths for the detection of eBFP, eGFP and mCherry fusion proteins at SEC runs.

Fluorophore	Excitation wavelength / nm	Emission wavelength / nm
eBFP	383	445
eGFP	475	510
mCherry	587	610

SEC buffer (Scap)	150 mM NaCl, 30 mM Tris, 1 mM TCEP, 0.05% (w/v) DDM/0.005% (w/v) CHS or 0.05% (w/v) GDN (pH 8.4)
SEC buffer (Scap /SREBP)	150 mM NaCl, 30 mM Tris, 1 mM TCEP, 0.05% (w/v) DDM/0.005% (w/v) CHS or 0.05% (w/v) GDN (pH 8.0)
SEC buffer (Patched1)	150 mM NaCl, 30 mM HEPES, 1 mM TCEP, 0.05% (w/v) DDM/0.005% (w/v) CHS (pH 7.5)

5.4 Biophysical methods

5.4.1 Tandem mass spectrometry

The tandem mass spectrometry (MS/MS) is a highly selective and sensitive method for the detection of biomolecules and was used for a qualitative analysis of a POI and for an analysis of protein contaminations. First, the proteins get enzymatically digested by trypsin to peptides which can be detected via mass spectrometry (MS). Here, a high-performance liquid chromatography is used for the separation of the peptides. Subsequently, an electrospray ionisation charges the eluted peptides that get separated by a quadrupole mass analyzer (QMS). The method of a QMS is based on collision induced dissociation by a collision cell (middle quadrupole) in which the peptides gets fragmented by their mass-to-charge ratio. (119)

The protein samples of interest and 5 pmol of BSA (as a reference) were separated by SDS-PAGE and stained with Coomassie brilliant blue R-250 (see chapter 5.3.1). The protein bands were cut into 1-2 mm³ gel pieces and fixed by the addition of 500 µl solution A for 16 hours at 23 °C. The gel pieces were destained with 200 µl of washing solution 1 and 2 successively as described in table 5.9.

Table 5.9: Washing steps and composition of washing solutions used for destaining.

Solution	Composition	Temperature / °C	Incubation time / min
Solution 1	25 mM NH ₄ HCO ₃ in acetonitrile (ratio 3:1)	37	30
Solution 2	25 mM NH ₄ HCO ₃ in acetonitrile (ratio 1:1)	37	15

The proteins were deoxidised by 200 µl of solution B for 45 min at 37 °C and alkylated with 200 µl of solution C in the dark. The gel slides were dried by the addition of 100 µl acetonitrile. The tryptic digest was performed with 0.01 mg/ml trypsin in 25 mM ammonium bicarbonate for 16 hours at RT. Finally, 10 µl of 10% (v/v) TFA were added to stop the digestion reaction. The extraction of the peptides out of the gel slides, the samples were sonicated for 30 min at 4 °C. The supernatant was removed and 75 µl acetonitrile were added and incubated for 15 min. The supernatant, that contains the peptides, was transferred into a second vessel and 75 µl acetonitrile were mixed with the gel slides for 15 min. The supernatants were combined and the peptides were concentrated in the Concentrator plus (Eppendorf) for 4 hours at RT. The analysis by mass spectrometry were performed at an in-house service.

Solution A	40% (v/v) ethanol, 10% (v/v) acetic acid
Solution B	50 mM DTT in 25 mM ammonium bicarbonate
Solution C	55 mM 2-iodoacetamide in 25 mM ammonium bicarbonate

5.4.2 Nano differential scanning fluorimetry

The thermal stability of a POI was analysed with the nano differential scanning fluorimetry (nanoDSF) method which is based on the thermal unfolding of proteins that can be measured by their intrinsic fluorescence of the amino acids tryptophan and tyrosine. The fluorescence of these aromatic amino acids is dependent on their environment that varies in a protein folded and unfolded structure.

The stability of the POI was analysed with the Prometheus NT.48 (NanoTemper Technologies). 15 μ l of the protein sample with a final concentration of 0.1 mg/ml were soaked into nanoDSF Grade high sensitivity capillaries (NanoTemper Technologies). The stability was screened in a range of 15 – 90 °C (1 °C/min) and the intrinsic fluorescence was detected at 330 nm and 350 nm at a laser intensity of 10 – 20%. The melting temperature T_m (at which 50% of the protein population is unfolded) were calculated and displayed using the internal software.

5.4.3 Fluorescence microscopy

Fluorescence microscopy was used on a regular basis for the analysis of the virus production and amplification (see chapter 5.1.8), the plaque assay (see chapter 5.2.2) or mammalian cells expressing proteins with a fluorescence reporter tag (4.3.3) by using the fluorescence microscope EVOSFL Auto (Life Technologies).

For localization studies of adherent mammalian cells (HEK293 F or GnTI) expressing Patched1 or Scap membrane proteins containing an eGFP reporter tag the LSM 800 confocal microscope with Airyscan (Carl Zeiss Microscopy GmbH) was used at 40-fold magnification. Here, the adherent mammalian cells were either transfected with the plasmid DNA encoding the Scap or Patched1 proteins by using lipofectamine[®] 3000 (Invitrogen, Thermo Fisher Scientific) (see chapter 5.1.9), or infected with 1% (v/v) P2 viruses for expression of the membrane protein of interest and incubated for 24 hours at 37 °C (8% CO₂).

5.5 Structural biology methods

5.5.1 Negative stain electron microscopy

A first three-dimensional (3D) model of the POI and the sample quality can be obtained via negative stain electron microscopy (EM). Here, the POI is embedded in a layer of a dried heavy metal solution and fixed on a copper grid. The copper grids contain several holes (squares) and are covered with a plastic film which is additionally coated with a continuous layer of carbon. Lastly, the carbon coated grids will get glow-discharged in order to adsorb the protein sample on the grid. (120)

A high voltage electron beam is used in a transmission electron microscope (TEM) to create an image of the specimen. Most of the electrons scatter at the heavy metal that surrounds the protein specimen, and provides a negative image of the protein sample. (121)

5.5.1.1 Sample preparation

The negative staining grids were carbon coated manually by using bare copper grids (G2400C, Plano GmbH) that were coated with a polymer film out of collodion (Sigma-Aldrich Chemie GmbH). The coating process was performed in a glass beaker filled with ddH₂O. The water surface was cleaned by the addition of 5 drops of the collodion solution two times. Finally, three drops of the collodion solution were added onto the calm water surface. The copper grids were then put on top of the film after evaporation of the solvent. The coated grids were then removed by using a blotting paper and were transferred into a petri dish. The grids were dried for 16 hours at room temperature and coated with 6-8 nm carbon in the carbon evaporator Leica EM ACE600 (Leica Microsystems).

The carbon coated copper grids were cleaned and glow-discharged in a plasma cleaner GloCube (Quorum Technologies Ltd) for 60 sec at 10 mA. The 0.75% uranyl formate staining solution was prepared freshly by transferring 0.0375 g uranyl formate into a beaker that were stirred for five minutes with pre-heated 5 ml ddH₂O. 3.5 µl NaOH (5 M) were added and stirred for another five minutes. The solution was filtered through 0.22 µm pores and stored at room temperature with minimal light exposure. The purified protein samples were diluted with staining buffer to a final concentration of 0.005 – 0.01 mg/ml. 4 µl of the protein sample were applied onto the glow-discharged grid and incubated for 45 seconds at RT. The excess liquid was blotted with a Whatman filter paper (No. 4) and washed by the addition of 10 µl of the appropriate staining buffer immediately. The excess fluid was removed as described before and the sample was washed additional two times with each 10 µl of the staining buffer. 10 µl of the 0.75% uranyl formate staining solution were added on top of the grid and the excess solution was removed with a filter paper. Immediately, a 10 µl of the staining solution were added and incubated for 45 seconds at RT. A final blotting step was performed before the grid was air-dried at RT for five minutes.

Staining buffer (Scap)	150 mM NaCl, 30 mM Tris (pH 8.4)
Staining buffer (Patched1)	150 mM NaCl, 30 mM HEPES (pH 7.5)
Staining buffer (Scap/Insig1)	150 mM NaCl, 30 mM Tris (pH 8.0)

5.5.1.2 Data collection

The negative stain images were collected with a JEOL JEM-1400Flash Electron microscope (JEOL GmbH) equipped with a LaB₆ cathode at 120 kV voltage. The images were recorded using a 4k x 4k CMOS TemCam (F416, TVIPS, Gauting GER) with one second exposure time at a nominal magnification of 50,000x. The EMMENU software (TVIPS) was used for full control of the camera by computer interface. The defocus was set between 1.5 and 2 μm .

5.5.1.3 Single particle analysis and data processing

The SPHIRE software package was used for single particle analysis and data processing of negative stain data sets. (122)

In single particle analysis the particles need to be selected manually, semi-automatically or automatically. The particles were automatically picked via the particle picking software crYOLO, which is part of the SPHIRE software package. CrYOLO is a deep-learning object detection system. The detection system is based on the *You Only Look Once* (YOLO) system that was first described in 2016 and employs deep convolutional neural networks (CNNs). (123) Here, a small data set of 20 micrographs was selected for manual picking and training of the network, which was performed in e2boxer (EMAN2). (124) The picked particles were extracted as images and classified into two-dimensional (2D) classes. First, the images showing the extracted particles were aligned by in-plane rotation and shifted in x/y direction in the 2D space by using ISAC, the iterative stable alignment and clustering approach, which is part of SPHIRE. (125) The 2D class averages were selected manually and used for further calculations of an *ab initio* 3D model by the Validation of Individual Parameter Reproducibility (VIPER). (126) The 3D volumes were visualized using UCSF Chimera. (127)

5.5.2 Cryo-electron microscopy

In cryo-electron microscopy (cryo-EM) the sample is vitrified by a plunge freezing process in liquid ethane or propane. The proteins are fixed in their hydrated state and the proteins conformation is in

a near-native state. Therefore, cryo-EM enables the determination of high-resolution structures compared to negative stain EM.

5.5.2.1 Sample preparation

For sample preparation of cryo-EM grids holey carbon grids were used. The grids contain several squares (holes) and vary in their mesh size and space between the holes. The grids are composed of copper, nickel or gold and furthermore covered with additional graphene oxide.

The protein sample is surrounded by a thin layer of vitrified ice, which should be as thin as possible in order to obtain a high signal-to-noise ratio (SNR). Nevertheless, the ice should be thick enough to prevent a specific orientation of the protein particles or displacement of the particles.

5.5.2.1.1 Sample preparation in pure ice

The Vitrobot was used for the vitrification process of all samples that were plunged on freshly glow-discharged holey Carbon Quantifoil 1.2/1.3 grids or with an additional gold film either on Quantifoil (UltraAuFoil® R 1.2/1.3) or C-flat™ (GF-1.2/1.3Au) grids. 3.5 µl of the purified protein sample were applied on the grid in a chamber with 100% humidity at 4 °C. After sample application, the grid was blotted automatically for 3 seconds with a blotforce of -3 and 0 seconds drain time. For the blotting process pre-cooled (4 °C) filter papers (Ø55/20 mm, grade 595) were used. Immediately afterwards, the grid was plunge-frozen into liquid ethane which was cooled to -160 °C by a surrounded chamber filled with liquid nitrogen.

5.5.2.1.2 Sample preparation on graphene grids

In cryo-EM it is necessary that the protein enters the holes of a carbon grid. The addition of a continuous film can be used to support the protein entering the holes. Here, a continuous supportive film such as additional carbon or graphene oxide are used.

First, a graphene-on-copper layer was produced by cutting 2x2 cm pieces of graphene-on-copper that were flattened. The flattened graphene-on-copper pieces were placed on top of a collodion coated water surface and incubated for 5 minutes. Excessive collodion was removed and the grids were incubated for 2 hours in 10% (w/v) sodium persulfate to etch the copper. The grids were washed three times by the exchange of the sodium persulfate solution with ddH₂O. The excess water was removed by using a Whatman paper No. 5 and furthermore the grids were dried for 16 hours at RT. The collodion layer was dissolved by rinsing the grid into ethyl acetate and washing in ethylene glycol, chloroform, acetone and isopropanol. The grids were either directly used for the plunge-freezing step with the protein of interest or first treated with plasma in a helium atmosphere according to Naydenova *et al.* with amylamine in order to get a positive charge. (128) In the latter case, the grids were treated with

plasma for 30 seconds at 20 mA. 3.5 μ l of the purified protein sample were applied on top of the grids covered with graphene oxide in a chamber with 100% humidity at 4 °C and incubated for 3 minutes. The grid was plunge-frozen with a Vitrobot into liquid ethane which was cooled to -160 °C after blotting the excess water with filter papers (\varnothing 55/20 mm, grade 595) for 3 seconds with a blotforce of -3 and 0 seconds drain time.

5.5.2.2 Data acquisition

Single-particle cryo-EM datasets of porcine Scap protein, either reconstituted into amphipols (A₈₋₃₅) or solubilized in detergent, were recorded on a spherical aberration (Cs) corrected FEI Titan Krios electron microscope. The microscope was operated at 300 kV and equipped with a high brightness field-emission gun (X-FEG) and a FEI Volta phase plate. (129, 130) Equally dosed frames of the datasets with Scap reconstituted into amphipols or solubilized in Digitonin were acquired using a K2 Summit (Gatan) direct detector at 1.09 Å/pixel (reconstituted into A₈₋₃₅) or 1.07 Å/pixel (solubilized in Digitonin). Furthermore, a quantum energy filter set to 20 eV was used for both datasets.

The micrographs of the dataset for the 3D reconstruction of Scap protein solubilized in glyco-diosgenin (GDN) were acquired at a FEI Titan Krios electron microscope that was equally equipped as described above without a Volta phase plate. The images were detected with a K3 camera (Gatan) direct detector using super-resolution mode at 0.9 Å/pixel in combination with a quantum energy filter set to 20 eV. Further details of the settings for the data acquisitions were illustrated in table 5.10.

The data acquisition was performed automatically using the EPU software (FEI) and as a pre-processing pipeline TransPHIRE was used during data set collection for monitoring the quality of acquired data.

Table 5.10: Settings for data acquisition of porcine Scap in different detergents or reconstituted in Amphipols (A₈₋₃₅) on a Cs corrected FEI Titan Krios electron microscope operated at 300 kV and equipped with X-FEG.

Settings	Scap in A ₈₋₃₅	Scap in Digitonin	Scap in GDN
Detection camera	K2 Summit (Gatan)	K2 Summit (Gatan)	K3 Super-resolution (Gatan)
Voltage (kV)	300	300	300
Magnification	105k	130k	105k
Pixel size (Å/pixel)	1.09	1.07	(2 x 0.45)=0.9
Total electron dose (e ⁻ /Å ²)	73.5	69	68
Number of frames/image	60	50	60
Exposure time (sec)	15	15	3
Estimated (adjusted) defocus range (μm)/phase shift range	0-180 ° (30-140 °)	0-170 ° (30-140 °)	-1.2 - -4.0 (-1.4 - -2.8)
Micrographs	7711	6871	7964

5.5.2.3 Single particle analysis

Three-dimensional high-resolution reconstructions of proteins can be obtained by single particle analysis of vitreous ice-embedded samples. In single particle analysis the protein particles get classified by an alignment of several similar images (of particles) resulting in an image with higher signal-to-noise ratio. After the refinement, the superposed particles can result in a high-resolution 3D model of the protein of interest. (131, 132) Initially, the collected data needed to be pre-processed.

The pre-processing steps were performed during data acquisition with TranSPHIRE as mentioned before (see chapter 5.5.2.2). First, the correction for beam-induced motion was performed with MotionCor2. Due to the doming effect of the electron beam on top of the vitreous ice layer, the frames need to be corrected for motion. This results in a significantly improved Thon ring quality and furthermore 3D reconstruction resolution. (133)

Second, the compensation and correction for the contrast transfer function (CTF) has to be performed prior starting processing. The CTF mathematically describes the imaging process in the EM expressed in the Fourier space. An accurate estimation of the CTF is essential for a high-resolution reconstruction due to the image formation in an electron microscope. The two major parameters are the defocus and the aberration of the lens that effect the CTF of micrographs formation. Here, the CTF was determined with CTFFIND4 within TranSPHIRE while data acquisition on the non-dose weighted aligned micrographs. (134) Not correct aligned frame averages were manually evaluated and removed.

The third pre-processing step is the estimation of the phase shift when using the Volta phase plate for data collection (see table 5.10). The variability of the phase shift of scattered relative to the unscattered electrons over time need be considered and therefore, the phase shift needs to be estimated while data acquisition. The estimation of the phase shift is implemented in CTFFIND4 which was used for the datasets collected with the Volta phase plate.

Finally, the particles were selected in the validated micrographs. Here, the automated particle picker crYOLO was used, as described before (see chapter 5.5.1.3). A small data set of 50 micrographs was selected for manual picking and training of the network from a vitrified sample. For this purpose, the image processing suite e2boxer (EMAN2) was used. (124)

After pre-processing and validation for suitable micrographs, the cryo-EM data set can be processed. The aim of the processing is a final 3D model that can be built based on extracted 3D information. For the data sets porcine Scap in amphipols (A_{8-35}) and Scap in Digitonin the software package SPHIRE was used for processing. (122) The processing of the collected data set of porcine Scap in GDN was performed with RELION and SPHIRE. (122, 135)

The automatically picked particles of all three data sets were extracted as images. The reference-free two-dimensional alignments and classifications of the Scap in amphipols and Digitonin data sets, were performed using ISAC, the iterative stable alignment and clustering approach, within the SPHIRE package. (120, 122) The 2D classification for the data set of Scap in GDN was performed within both software packages SPHIRE (ISAC) and RELION (maximum-likelihood). Both approaches are based on an

iterative comparison of images from at least one reference in different orientations with every single extracted particle image. The averages of each particle build the reference of the next iteration. In a final step, the best orientation and class for each particle is calculated. Furthermore, ISAC permits the alignment and clustering of a heterogenous set of EM projection images via a new clustering algorithm (EQK-means). EQK-means is an improved k-means algorithm, in which the reproducibility of the 2D classes is screened.

In all three data sets, the 2D class averages were manually selected and contaminations were discarded. After several rounds of 2D classification using ISAC (for Scap in A₈₋₃₅ and Scap in Digitonin), the selected 2D class averages were used for a reproducible *ab initio* 3D structure determination using the validation of individual parameter reproducibility (RVIPER), which is part of SPHIRE. (122, 125) In contrast to that, the 2D classification of the data set from Scap in GDN was performed once using ISAC. The best 2D class averages were selected for further 2D classification using maximum-likelihood which is implemented in RELION. The best 2D class averages were selected and used for further processing in SPHIRE in order to determine an *ab initio* 3D structure by RVIPER. (122)

Subsequently, the 3D refinement was performed in SPHIRE using MERIDIEN with no internal symmetry (C1) imposed. In all three datasets MERIDIEN was operated with a mask containing the detergent micelle (Digitonin/GDN) or amphipol (A₈₋₃₅) and the soluble part of Scap in order to improve the results. Here, the application of a mask filters the noise surrounding the membrane protein of interest.

After combining the resulting two half-maps and using the PostRefiner in SPHIRE, the maps were filtered to 9.4 Å (Scap in A₈₋₃₅), 7.6 Å (Scap in Digitonin) and 6.6 Å (Scap in GDN) using a sharpening B-factor of 300 Å.

All three overall resolutions were estimated with the gold standard Fourier-Shell-Correlation (FSC) = 0.143 criterion of the two half maps. The gold standard FSC describes the cross-correlation of the two half maps that were created during MERIDIEN (3D refinement). The two half maps were generated from the splitted dataset into two subsets that were furthermore independently refined. (136)

The analysis and visualisation of all three protein structures was performed in UCSF Chimera. (127)

6 Results and Discussion

6.1 Patched protein

In order to analyse the transmembrane protein Patched on a molecular level, I evaluated its heterologous expression in a near native environment. Several published results indicated a non-native conformation when human Patched proteins were expressed in yeast or insect cells. (137, 138) In both expression systems the membrane protein was expressed at high levels and could be purified but failed in binding to its ligand Hh when solubilized in detergent. This indicated a non-native conformation of Patched when expressed in yeast or insect cells. The membrane composition is highly variable in cells from different organisms, which represents the environment of a membrane protein. Dependent on its composition of the membrane, the membrane proteins native conformation can be changed and affect its stability and activity. Therefore, I used the mammalian cell line HEK293 as expression host for mammalian and insect homologs and paralogs of the Patched protein.

6.1.1 Homolog screen of Patched proteins

I selected six different homologs of the human Patched protein that coincide from at least 75% query coverage and share a high sequence identity. Based on this sequence alignment study (see Figures 8.1 and 8.2), I used the drosophila Patched homolog, two Patched1 and four Patched2 proteins (see table 6.1).

Table 6.1: Homologs of Patched protein sharing high sequence identity with the human Patched protein that were used for the homology screen.

Protein	Organism	Sequence identity [%]	Query coverage [%]
Patched	<i>Drosophila melanogaster</i>	38.61	88
Patched1	<i>Danio rerio</i>	57.63	99
Patched1	<i>Homo sapiens</i>	100	100
Patched2	<i>Homo sapiens</i>	100	100
Patched2	<i>Mus musculus</i>	91.02	98
Patched2	<i>Rattus norvegicus</i>	91.53	98
Patched2	<i>Xenopus laevis</i>	68.08	92

In addition to that, I have either added a N-terminal eGFP reporter tag and C-terminal poly-Histidine (His₆) purification tag or vice versa. In total, I performed the homolog screen with 14 different Patched constructs. A schematic representation of the construct's setup is shown in Figure 6.1.

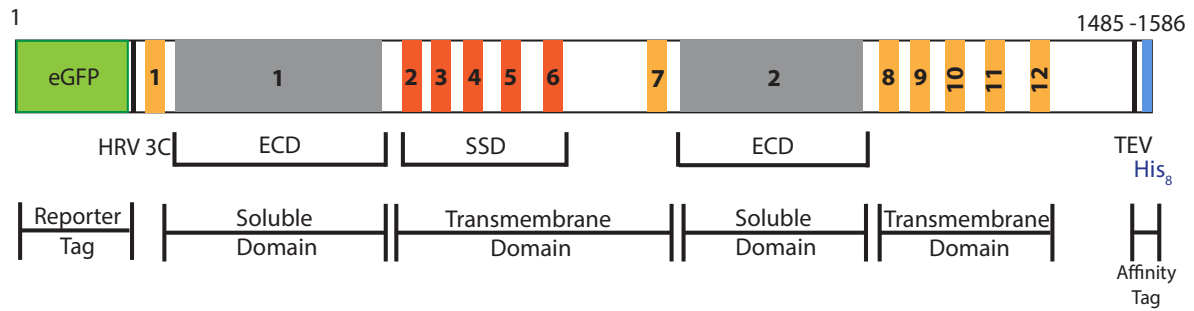
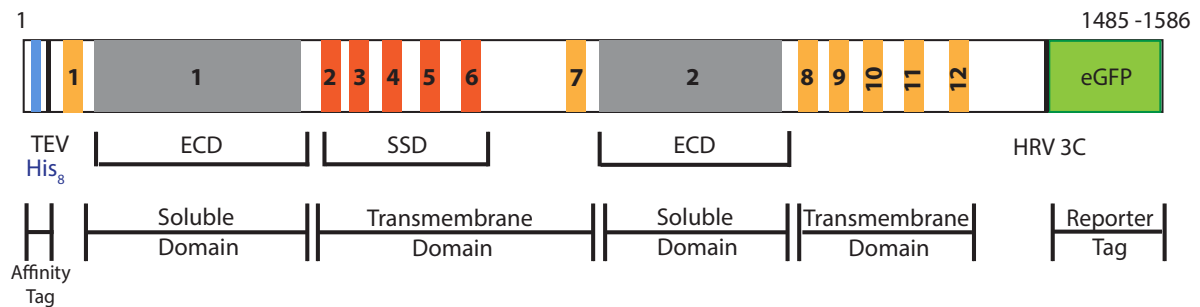
N-terminal eGFP Patched constructs**C-terminal eGFP Patched constructs**

Figure 6.1: General setup of Patched constructs used for the homolog screen.

The Patched constructs are composed of a N-terminal eGFP (green) reporter Tag and C-terminal poly-Histidine (His₈) (blue) affinity purification tag (top) or vice versa (bottom). All constructs shared high sequence identity and query cover with the human Patched protein and were supposed to be composed of 12 transmembrane domains (red and orange) and two larger extracellular domains (ECD) (grey). The sterol sensing domain is highlighted in red and located from helix 2 until helix 6.

The encoding DNA sequences for these protein constructs, were cloned into the plasmid Eric Gouaux (pEG) for baculovirus transduction of mammalian cells (pEG BacMam) via restriction digest and ligation cloning. (102) The plasmids were transformed into the adherent HEK293 T cell line by transient transfection (see chapter 5.1.9). The cells were incubated according to the manufacturer's guidance and the expression was traced by light and fluorescence microscopy (see chapter 5.2.1). The correct localization and the expression yield of the protein constructs were analysed via fluorescence microscopy after 48 h of expression before cell harvest (see Figure 6.2). Here, the overall fluorescence signals are enriched at the plasma membranes of the cells indicating a correct folded Patched protein in most of the cases. A low number of cells were not attached to the dish surface and appeared more rounded. In addition, these cells were completely fluorescently labelled that implied that the strong overexpression of the construct of interest led to cell death in these few cases. All in all, cells expressing Patched2 constructs showed highest amounts of fluorescent cells compared to Patched1 and the drosophila homolog. In addition to that, the number of fluorescent cells is enriched when expressing C-terminal modified Patched constructs compared to N-terminal modified proteins (see Figure 6.2).

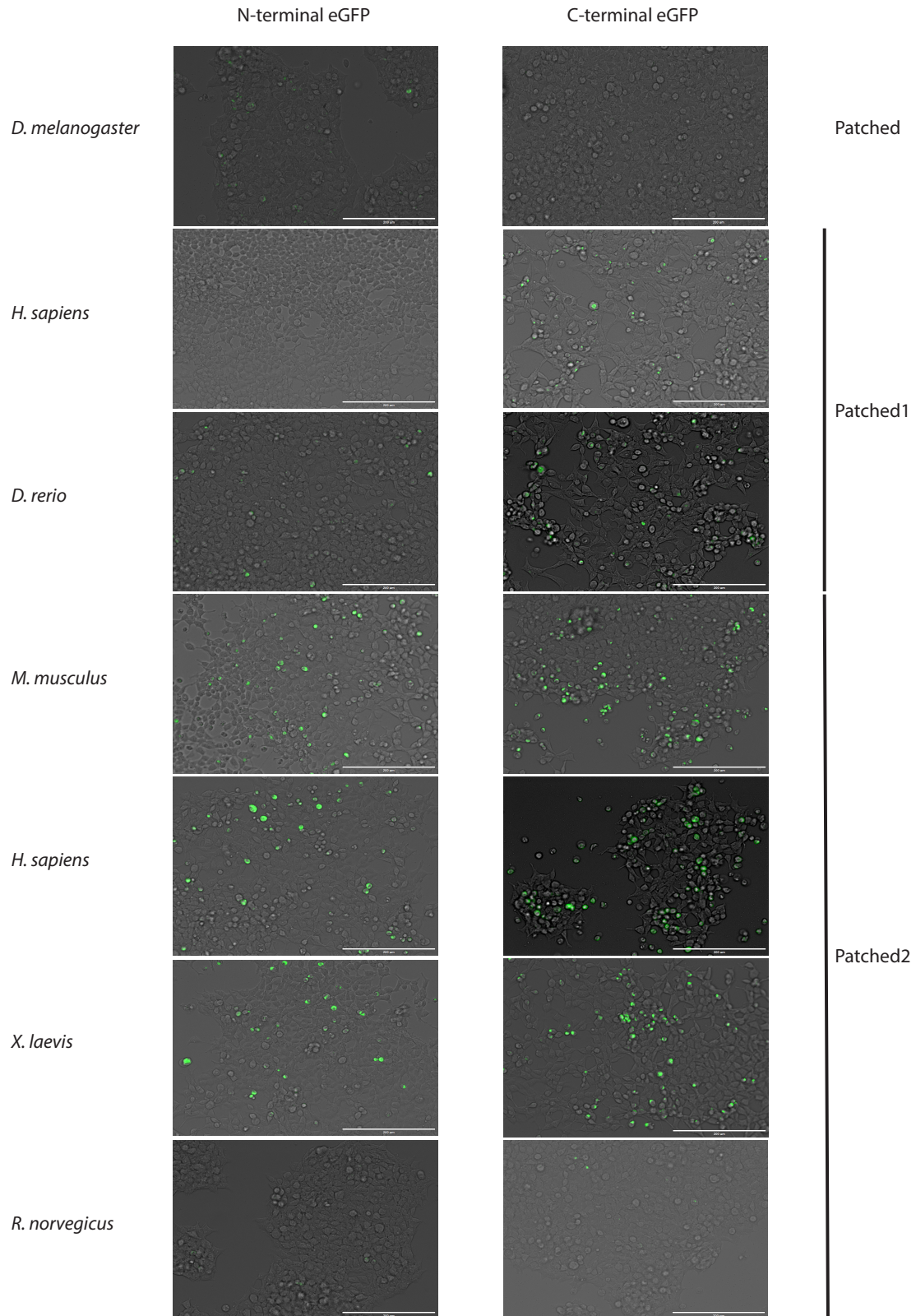


Figure 6.2: Fluorescence images of transiently transfected adherent HEK293 T cells after 48 h expression of Patched constructs.

Merged images of the eGFP fluorescence channel (green) and the bright field (grey) of all screened Patched constructs used for the homolog screen taken with the EVOSFL Auto Fluorescence microscope (Life Technologies) either containing N-(left column) or C-terminal (right column) reporter tags. Scale bars: 200 µm.

The cells were lysed in the appropriate buffer and used for the whole-lysate-solubilization in DDM containing buffer (with a final DDM concentration of 1% (v/v)) as described earlier in chapter 5.3.4. The solubilize was used for an analysis of the expression, solubilization efficiency, monodispersity via FSEC and semi-denaturing SDS-PAGE (see chapter 5.3.1 and 5.3.5) (see Figure 6.3).

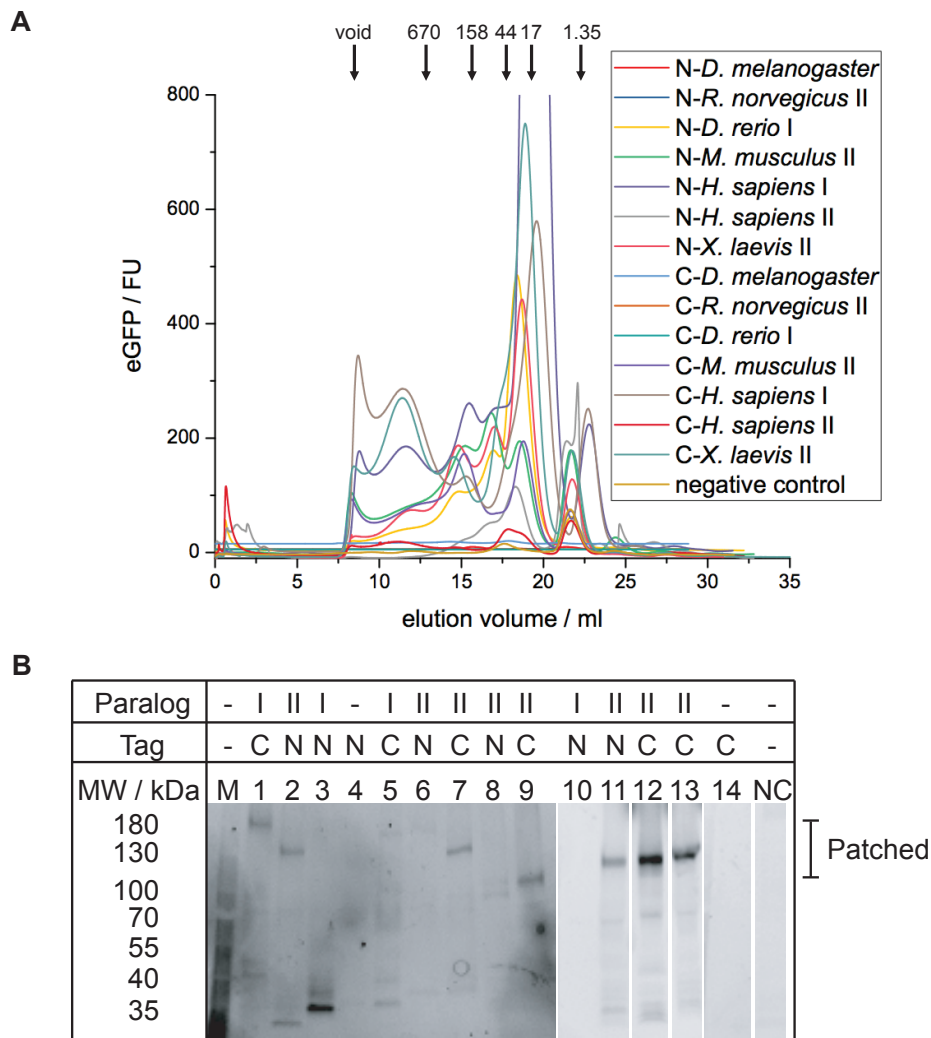


Figure 6.3: FSEC chromatograms and SDS-PAGE-gels of the Patched homolog screen.

(A) FSEC chromatograms of the 14 different Patched constructs after solubilization in buffer containing 1% DDM and a negative control. (B) Fluorescence image of the SDS-PAGE-gel from the DDM-solubilized 14 Patched homologs and paralogs and the negative control (NC) taken with the ChemiDoc MP imaging system (Bio-Rad). M: PageRuler Prestained Protein Ladder, 1: *D. rerio* (Patched1-eGFP), 2: *M. musculus* (eGFP-Patched2), 3: *H. sapiens* (eGFP-Patched1), 4: *D. melanogaster* (eGFP-Patched), 5: *H. sapiens* (Patched1-eGFP), 6: *X. laevis* (eGFP-Patched2), 7: *X. laevis* (Patched2-eGFP), 8: *R. norvegicus* (eGFP-Patched2), 9: *R. norvegicus* (Patched2-eGFP), 10: *D. rerio* (eGFP-Patched1), 11: *H. sapiens* (eGFP-Patched2), 12: *H. sapiens* (Patched2-eGFP), 13: *M. musculus* (Patched2-eGFP), 14: *D. melanogaster* (Patched-eGFP) and the scale bar (right) indicates the range of molecular weights of the Patched proteins.

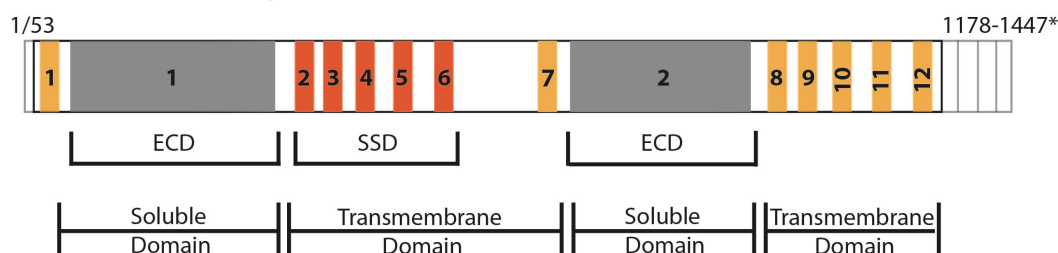
The chromatograms of the 14 FSEC runs show five main peaks that appear in all screened homologs and paralogs of Patched constructs. The first peak at around 8.5 ml refers to the void volume that

contains all aggregated but still fluorescent DDM-solubilized particles. A second peak could be observed at 12 ml that might refer to higher oligomers of the eGFP labelled protein constructs as 670 kDa proteins elute later at 12.5 ml elution volume and one Patched protein is around 250 kDa including the DDM detergent micelle. (141) A third peak fraction appears after an elution volume of 15 ml that correlated with one monomeric Patched construct solubilized in DDM. A fourth and fifth peak fraction could be observed after 18 ml and between 22 - 23 ml elution volume which refers to an elution of solubilized molecules of ≤ 20 kDa containing the eGFP fluorescent tag and indicated degradation of the POI. These results were in accordance with the semi-denaturing SDS-PAGE analysis, in which the solubilized protein samples showed various different degradation products for some homologs and paralogs (see Figure 6.3 A and B). Noticeable, Patched2 constructs showed a separated and sharp third eGFP peak at around 15 ml and less degradation products by semi-denaturing SDS-PAGE analysis indicating a stable monomeric Patched2 protein construct (see Figure 6.3). Contrary to the stable Patched2 constructs, the Patched1 proteins appeared less stabilized in its monomeric conformation. By FSEC the Patched1 paralog constructs show an increased amount of the void volume compared to the protein constructs of the Patched2 paralogs (see Figure 6.3 A). In addition to that, various DDM-solubilized Patched1 constructs featured eGFP fused degradation products of the POI at semi-denaturing SDS-PAGE (see Figure 6.3 B and C).

To conclude, the FSEC results were in accordance with the eGFP expression pattern that could have been observed in the fluorescence images represented in Figure 6.2. The homolog screen, indicated a less stable Patched1 than Patched2 paralog while over-expression in HEK293 cells and DDM-solubilization. The main differences of both membrane proteins are the enriched N- and C-termini of the Patched1 protein compared to Patched2. These findings could indicate that the soluble N- and C-terminal regions of Patched1 have a high flexibility and an impact on its native conformation during over-expression in mammalian cells and solubilization in the detergent DDM. Furthermore, previous studies with the wild-type murine Patched1 protein show degradation of the protein that could have been clarified later when a binding site for the Nedd4 E3 ubiquitin ligase (HECT) was identified at the C-terminal region. (6, 138, 142) An additional Gli-dependent activity assay and a plasma membrane-localization study revealed a functional C-terminal truncated Patched1 protein. Furthermore, it has been shown that the C-terminal domain is solely responsible for turnover which indicate less degradation of the protein of interest without the full-length C-terminal region. (138, 142, 143) Therefore, I designed N- and C-terminal truncated Patched1 protein constructs in order to improve the protein's stability while expression and solubilization.

6.1.2 Truncated human Patched1 constructs

Based on the results of the Patched homolog screen, I selected nine different truncated human Patched1 constructs containing a C-terminal purification and reporter tag. I designed five N-terminal and four C-terminal truncated Patched1 variants (see table 6.2 and figure 6.4).

Truncated Patched1 *Homo sapiens*

Patched1 constructs	Amino acid sequence	Molecular weight / kDa (including tag)
1	1-1415	188.85
2	1-1365	183.74
3	1-1210	167.00
4	1-1178	163.61
5	53-1447*	187.55
6	53-1415	183.78
7	53-1365	178.68
8	53-1210	161.90
9	53-1178	158.54

Figure 6.4: Schematic setup and details of the truncated human Patched1 constructs used for stability screen.

The truncated protein constructs either contain the total N-terminus (M1) or started from amino acid P53 and a variety of C-terminal endings ranging from F1178 to N1447*. The transmembrane domains (coloured) as well as both ECD (grey) were maintained in all truncated protein constructs.

The truncations were implemented based on the sequence alignment of the human Patched1 and Patched2 (see Figure 8.3). The protein sequence alignment clearly showed a strong coverage in conservation, quality, consensus and occupancy of both amino acid sequences starting from amino acid P53 until F1180 (Patched1)/I1180 (Patched2). Additional information based on the transmembrane prediction calculator TMHMM which is based on the hidden Markov model (HMM) proposed a first transmembrane domain starting from amino acid F101 (Patched1)/V101 (Patched2) and a last transmembrane helix until amino acid P1175. (144, 145) Due to these results, I designed truncated human Patched1 constructs either starting from M1 or P53 in order to validate an effect of the shortened N-terminus. The effect of an additional truncated C-terminus was analysed by the

elimination of this domain by 269, 237, 82 or 32 amino acids from the total C-terminus N1447* (see Figure 6.4).

The expression pattern and stabilization of the shortened human Patched1 proteins were analysed by fluorescence microscopy based on its plasma membrane localization (see Figure 8.4) and FSEC screening (see chapter 5.3.5). These experiments were performed by Julia Kotschy under my guidance. (146)

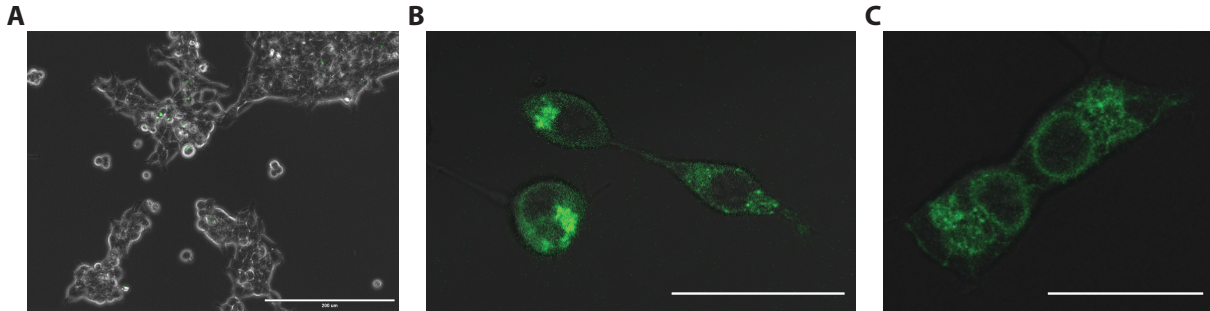


Figure 6.5: Fluorescence images of transiently transfected adherent HEK293 cells after 48 h expression of truncated human Patched1 variant number 3 (M1-1210).

Merged images of the eGFP fluorescence channel (green) and the bright field (grey) taken with the (A) EVOSFL Auto Fluorescence microscope (Life Technologies, scale bar: 200 μm) and (B) recorded with the Zeiss LSM 800 microscope by using Airyscan at 40-fold magnification (Carl Zeiss Microscopy GmbH, scale bar: 50 μm). (C) Reference-fluorescence image of CellLight ER-GFP (BacMam 2.0, Invitrogen, Thermo Fisher Scientific) infected adherent HEK293 cells taken with the Zeiss LSM 800 microscope at 40x magnification (Airyscan, Carl Zeiss Microscope GmbH, scale bar: 30 μm).

Figure 6.5 represents the expression pattern of the truncated human Patched1 construct 3 (1-1210, A and B) and a GFP fused ER marker (C, as a reference) after 48 h expression in adherent HEK293 cells. The fluorescence images of all shortened human Patched1 proteins show a membrane localization and medium expression level with around 30% eGFP fluorescence signal (see figure 8.4 and 6.5). The eGFP fluorescence signal is observable at the cell membrane and additional membranes inside the cells, like in the ER membranes or Golgi membranes (see Figure 6.5 B and C). Compared to the human wild-type Patched1 construct, the expression pattern is similar which includes the amount of eGFP expressing cells and its localization in the cells (see Figure 6.2).

Apparently, the overexpression of the cell membrane protein lead to an altered distribution of Patched1 localization to various lipid bilayers as these are necessary to form the correct folded membrane protein. Moreover, an overexpression of the membrane protein of interest could have been induced an accumulation in the ER membrane while translation. As described before, membrane proteins get inserted into the ER membrane during translation process by the translocon that could explain an increased amount of Patched1 proteins in the ER membrane when overexpressed. (24) A few cells show extensive expression of eGFP indicating the formation of aggregates out of degradation products that still contain the fluorophore or a partially folded full-length POI. In addition to that, an overexpression of an active full-length Patched1 could entail different effects on cells per se. As mentioned before, Patched1 inhibits the Hh signalling cascade by suppression of another membrane

protein called Smo. An abnormal activation of this pathway can lead to different kind of cancers and therefore can have an effect on the cells overexpressing Patched1. (147, 148)

In order to exclude the expression of degradation products and evaluate the stability of the new designed truncated Patched1 protein constructs a FSEC screen and additional SDS-PAGE analysis were performed (see Figure 6.6).

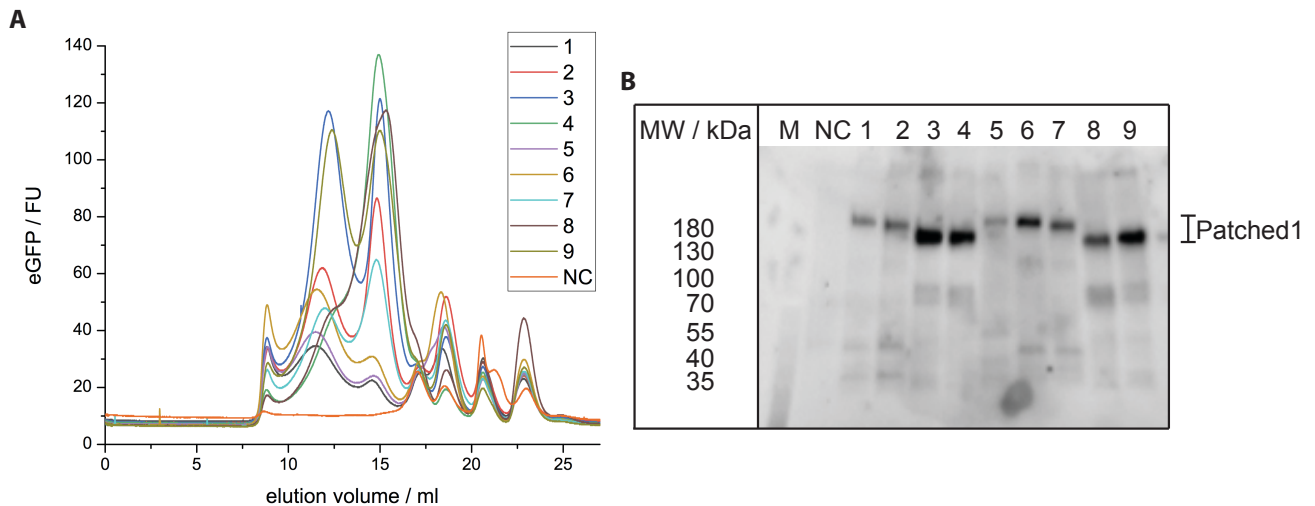


Figure 6.6: Stability screen of the nine truncated human Patched1 constructs.

(A) Nine chromatograms of the N- and C-terminal truncated Patched proteins after solubilization in DDM visualized as followed: 1 (black), 2 (red), 3 (blue), 4 (green), 5 (purple), 6 (beige), 7 (cyan), 8 (brown), 9 (olive) of the DDM-solubilized fraction and one negative control (NC) without overexpressed Patched protein construct (orange). (B) Fluorescence image of the semi-denaturing SDS-PAGE gel of the nine shortened protein constructs 1-9, the negative control (NC) and the PageRuler Prestained Protein Ladder (M) taken with the ChemiDoc MP imaging system (Bio-Rad). The scale bar indicates the molecular weight range of the screened Patched1 truncations.

The results of the stability screen are presented in Figure 6.6 that include the FSEC chromatograms (A) and the SDS-PAGE analysis (B) of the DDM-solubilized fractions of the Patched1 expressing HEK293 cells that were illustrated before (see Figures 6.5 and 8.4). All nine chromatograms contain three main peaks between 8.5 ml and 17 ml elution volume. A first fluorescence peak appears after 8.5 ml of elution, the second at around 12 ml elution volume and a last main peak at around 15 ml. The first peak refers to the void volume in which aggregated particles elute. The second and third elution peaks presumably refer to the membrane protein of interest as they do not appear in the negative control (DDM-solubilized fraction without POI overexpression). The additional elution peaks from 17 ml on most likely appear due to the elution of intrinsic fluorescent biomolecules that were either solubilized in DDM or could not be separated by ultra-centrifugation and were summed up as contaminants (see Figure 6.6 A). In the second main peak at around 12 ml, solubilized particles with a hydrodynamic radius of around 450 kDa elute. These findings would indicate the existence of a DDM-solubilized homodimer or homotrimer of two or three Patched1 proteins with a molecular weight of roughly 350-500 kDa. Taken the DDM detergent micelle of around 70 kDa into account this would be in accordance with the molecular weight (450 kDa) that was calculated previously by using a protein standard on the

same size-exclusion column (see Figure 6.3 A). The third main peak at around 15 ml refers to fluorescent protein particles with a molecular weight of 170 kDa and indicates the elution of the monomeric truncated Patched1 proteins. For protein construct 4 and 8 these two main peaks do not appear separated which implies less stable monomeric Patched1 constructs compared to the other screened proteins. Notably, the ratio between the oligomeric (second main) peak and the monomeric (third main) peak can be observed with a shortened C-terminus. Here, the second main peak in which the oligomeric Patched1 molecules elute is less distinct for constructs with a shorter C-terminus compared to the constructs containing longer C-termini and the monomer eluting (third main) peak is increasing simultaneously. These observations indicate a stabilized Patched1 monomer upon truncated C-termini that are in accordance with the findings that Patched1 oligomerization is mediated by its C-terminal region. (142, 149)

The SDS-PAGE gel shows one main band for each of the new designed Patched1 constructs in the range of 160 to 190 kDa (see Figure 6.6 B) and for most proteins additional negligible bands below the calculated molecular weights of the POI. These indicate a slight degradation of the POI as these bands appear after the eGFP fluorescence excitation. Nevertheless, the protein bands around 170 kDa show most intense fluorescence and implies a stable DDM solubilized protein constructs. In conjunction with the result of the FSEC screen the solubilizes clearly showed the successful expression and DDM solubilization of all human truncated Patched1 constructs.

Taken everything into account, the expression pattern, FSEC screen and SDS-PAGE analysis indicate an increased stability of Patched1 construct 3 compared to all other truncations as well as the full-length human Patched1. Here, the stability could be improved as this construct showed separated homomultimeric and monomeric gaussian-like peaks eluting during SEC after DDM-solubilization and less aggregated protein particles as the void volume peak was less distinctive (see Figure 6.6 A). These findings could be confirmed by SDS-PAGE analysis in which construct 3 showed less degradation products and the strongest protein band compared to the other truncated POI. In comparison with the full-length human Patched1 containing the C-terminal eGFP reporter tag, the truncated version (1-1210) showed an increased stabilization of monomers that could be separated from higher oligomers and degradation products via SEC. The truncated construct 3 lacks 237 amino acids at the C-terminus which equals around 16% of the total protein and was used for further studies of human Patched1 (see Figure 6.7).

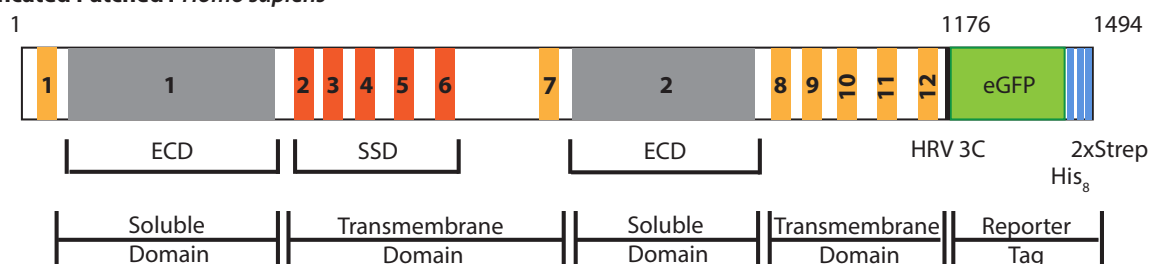
Truncated Patched1 *Homo sapiens*

Figure 6.7: Setup of the truncated human Patched1 variant used for large scale expression and structural studies.

The truncated protein construct was composed of the human Patched1 1-1210 followed by a C-terminal HRV 3C protease cleavage site (site), the eGFP reporter tag (green), an octa-Histidine- (His₈-) and double StrepII affinity purification tag (blue).

In membrane protein purification the solubilization step is one of the most critical steps and requires an extended attention in particular. The solubilization of a membrane protein is described as its extraction out of the lipid bilayer in order to keep it soluble in an aqueous solution. Detergents are amphipathic molecules and commonly used for the solubilization of membrane proteins. They can be either non-ionic, ionic or zwitterionic in their structure and therefore varying in their ability of stabilizing a membrane protein of interest. In order to improve the solubilization step, twelve different detergents and detergent mixtures that included zwitterionic, ionic and non-ionic detergents were tested. The solubilization efficiency was verified via FSEC after the solubilization in different detergents and performed by Julia Kotschy under my guidance. (146)

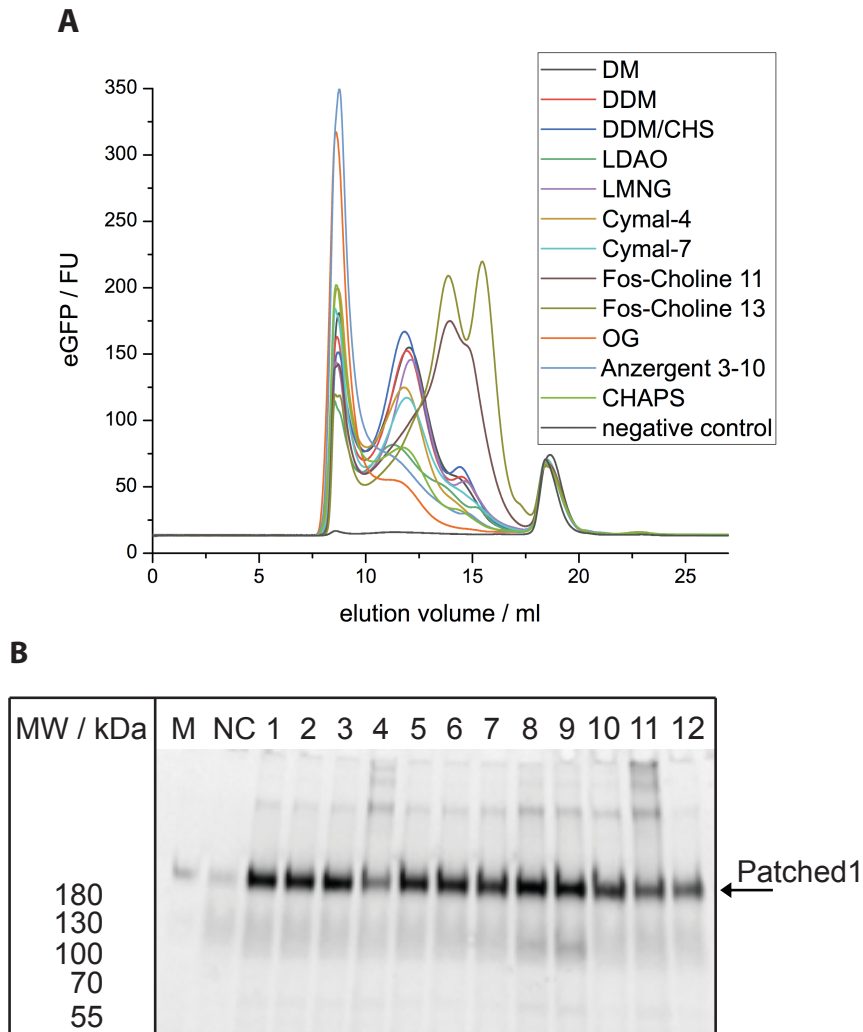


Figure 6.8: Detergent screen of the C-terminal truncated human Patched1 construct 3 (1-1210).

(A) Chromatograms of the detergent solubilized Patched1 construct 3 coloured as follows: (1) DM (black), (2) DDM (red), (3) DDM/CHS mixture (blue), (4) LDAO (green), (5) LMNG (purple), (6) Cymal-4 (beige), (7) Cymal-7 (cyan), (8) Fos-Choline 11 (brown), (9) Fos-Choline 13 (olive), (10) OG (orange), (11) Anzergent 3-10 (light blue), (12) CHAPS (light green), negative control (grey). (B) Fluorescence image of the semi-denaturing SDS-PAGE gel taken with the ChemiDoc MP imaging system (Bio-Rad) of the 12 tested detergents for solubilization, one negative control and the protein standard (M) PageRuler Prestained Protein Ladder. The arrow indicates the molecular weight of C-terminal eGFP-tagged Patched1.

Figure 6.8 represents the chromatograms (A) and the fluorescence image of the semi-denaturing SDS-PAGE gel (B). All twelve chromatograms contain four main peaks as well as an additional peak at around 18 ml which could be verified as fluorescently solubilized contaminants of the cells since this fractions was present under all conditions including the negative control that does not contain overexpressed proteins (see Figure 6.8, A, black chromatogram). The first peak appears after a retention volume of 8.5 ml and contains all aggregated fluorescent particles. Except from Fos-Choline 11 and 13, the second main fraction elutes after 12 ml retention volume and corresponds to particles with a molecular weight of around 450 kDa as mentioned before (see Figure 6.3). A third fraction elutes at around 14 ml of retention volume which correlates to fluorescent particles with a molecular weight

of roughly 180 kDa. As described before, oligomeric truncated Patched1 proteins most probably elute in the second fraction at 12 ml as this matches the calculated molecular weight of a homotrimer as well as homodimer including the detergent micelle of roughly 400-500 kDa. The third main peak refers to the monomeric truncated Patched1 protein particles of around 200 kDa as this coincides with the size estimated by SEC of a protein standard (see Figure 6.3 A). In contrast to these results, the solubilization with the zwitterionic detergents Fos-Choline 11 (brown) and 13 (olive) show diverse SEC characteristics compared to the other ten screened detergents. Here, only a broad shoulder starting from 12 ml retention volume was detectable and two non-separated peak fractions elute at around 14 to 16 ml. This indicates the elution of a monomeric Patched1 construct and degradation products with a molecular weight of up to 45 kDa (see Figure 6.3). As these elution fractions could not be properly separated by SEC, a specific, non-degraded formation of construct 3 could not be efficiently solubilized nor stabilized by Fos-Choline 11 or 13. Both detergents differ in their length of hydrocarbon chain but share the same zwitterionic head group which could disrupt the intramolecular protein-protein interactions of the POIs native conformation. Furthermore, it was shown before by T. E. Cleveland IV *et al.* that Fos-Choline 12 solubilized mouse Patched1 failed in interacting with Hh proteins indicating the solubilization of a non-native Patched1 conformation when using Fos-Choline as solubilization detergent. (137) For some other detergents, like OG (orange) or Anzergent 3-10 (light blue) the void volume (first main elution fraction) appeared very strong in comparison to the other peak fractions. These findings indicate a weak stabilization of the POIs native conformation as the amount of aggregated fluorescent biomolecules exceed the oligomeric or monomeric conformation of Patched1. Similar to Fos-Choline, Anzergent contains a zwitterionic head group which could interfere with intramolecular interactions and disrupt the protein's native structure. The other detergents screened for solubilization of the truncated Patched1 construct 3 seem to stabilize oligomeric conformations as good as the monomeric state when taking the ratio of homotrimer or homodimer to one monomer protein particle (1:3 or 1:2) into account. The analysis of the solubilized protein samples via semi-denaturing SDS-PAGE show one main protein band for all tested detergents at around 170 kDa in the fluorescence channel (see Figure 6.8 B). Additional fluorescent protein bands appear for the POI solubilized samples with LDAO, Fos-Choline 11/13, OG and Anzergent 3-10 at a molecular weight of around 90 kDa and one band above 180 kDa. These findings indicate the stabilization of a mixed population of oligomeric, monomeric and degraded fluorescent POI when using these detergents for the extraction of truncated construct 3 of human Patched1. Furthermore, the monomeric POI band of 170 kDa is less intense for LDAO, OG and Anzergent 3-10 treated samples compared to all other conditions.

In combination with the results of the FSEC screening and, the detergent mixture composed of DDM and CHS, a cholesterol derivative, seems to solubilize the truncated Patched1 most efficiently. These include the extraction efficiency out of the lipid bilayer on itself but most important the stabilization of the protein's native conformation during SEC as this could be visualized by a sharp gaussian-like elution peak. DDM is a mild detergent and in combination with CHS a detergent micelle that apparently

mimics the proteins native environment most efficiently as cholesterol is a well-known binding partner of Patched1. (137, 138)

For further studies on the human Patched1 membrane protein the truncated protein construct containing the first 1210 amino acids of the wild type human Patched followed by C-terminal HRV 3C protease cleavage site, the eGFP reporter tag and the affinity purification tags His₈ as well as StrepII (see Figure 6.7) was used. Furthermore, the most efficient detergent for solubilizing the membrane protein construct 3 was verified by an analogous FSEC screen using different kinds of detergents including ionic, zwitterionic and non-ionic detergents. Here, the detergent mixture composed of DDM and CHS showed the most stabilized protein after extraction out of the membrane and was used in the following solubilization and purification steps.

6.1.3 Heterologous expression and purification strategy of Patched1

The expression of the stabilized human Patched1 protein construct 3 was first tested in suspension mammalian cells with a potent P3 BacMam virus (see chapter 5.1.8). Here, the mammalian cell lines HEK293 F and GnTi⁻ cells as well as 4.5% and 9% (v/v) virus infection were tested, screened and validated for up to three days by Julia Kotschy under my guidance. For further information please refer to her master thesis. (144) The expression of the truncated Patched1 protein was scaled-up to 3.5 l of suspension HEK293 F cells at 30 °C and harvested after 48 hours post-induction. The cells were lysed via sonication as described in chapter 5.3.7 and the protein solubilization was performed with the detergent mixture DDM/CHS as this most efficiently solubilized the POI (see chapter 6.1.2 and Figure 6.8). Subsequently, a purification strategy for the isolation of the POI without additional tags was established by Julia Kotschy under my guidance. (146)

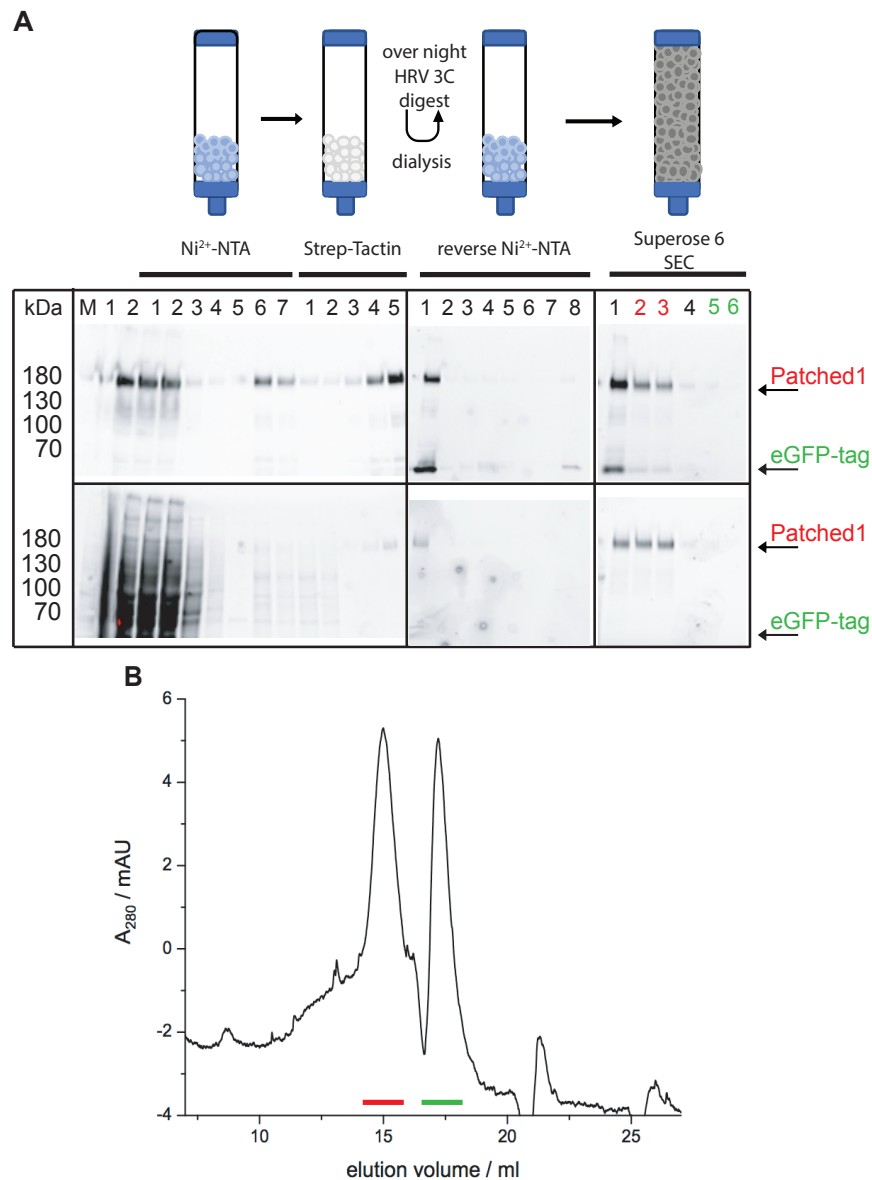


Figure 6.9: Purification results of the truncated human Patched1 construct 3.

(A) Scheme of the purification strategy and three exemplary semi-denaturing SDS-PAGE gels of eluting fractions in the fluorescence channel (top) and the stain-free view (bottom) taken with the ChemiDoc MP imaging system (Bio-Rad). The left SDS-gel shows the collected fractions of the solubilization before ultracentrifugation (1), after ultracentrifugation (2) and the tandem-affinity chromatography by first the Ni²⁺-NTA including the flow-through (1), four washing steps with 30 mM imidazole (2-4) and the elution fractions (5-7) with 150 mM imidazole followed by the Strep-Tactin affinity chromatography out of the flow-through (1), the combined washes (2) and three elution fractions by using buffer containing 2.5 mM desthiobiotin (3-5). The middle SDS-gel comprises the samples of all collected reverse Ni²⁺-NTA chromatography fractions of the over-night HRV 3C protease digested POI with the input (1), the flow-through (2), five washing steps with 30 mM imidazole (3-7) and the elution of the digested eGFP-tag (8). The right SDS-gel contains selected fractions of the eluted peaks collected while SEC in the following order (1) input, (2) after 14.5 ml, (3) at 15.5 ml, (4) after 17 ml, (5) at 18.5 ml and (6) after 22 ml elution volume. The arrows indicated the Patched1 (red) and the digested eGFP-tag (green) levels. (B) Exemplary chromatogram of the tandem purified and HRV 3C digested construct 3 (POI) solubilized in DDM/CHS including two main peak fractions composed of the construct 3 and the digested POI (red) and the digested eGFP-tag (green).

The purification strategy as well as the purification results of the human Patched1 construct 3 were represented in Figure 6.9. Starting from the isolation of the DDM/CHS-solubilized fraction via ultracentrifugation a tandem-affinity chromatography was performed. By semi-denaturing SDS-PAGE the amount of solubilized POI and the presence of the POI in all collected fractions of the affinity chromatographies were verified. Here, one main fluorescence protein band of roughly 170 kDa could be observed after solubilization and ultracentrifugation steps that coincides with the calculated MW of the POI of 167 kDa. A less intense fluorescent protein band of the same size was present in the fractions of the flow-through and the first washing step of the Ni²⁺-NTA affinity chromatography. These results indicate either weak interactions between the mobile phase and the POI or a saturated state of the mobile phase by the POI or contaminants since the SDS-gel in the stain free channel shows a decreased variety of additional protein bands from the first purification step on (see Figure Figure 6.9 A, left). With an increased volume of buffer containing 30 mM imidazole the variety of contaminants is decreased in the stain-free SDS-gel view. The fluorescent protein band of 170 kDa is present after the application of 150 mM imidazole containing buffer which demonstrates the elution of the POI. Compared to the sample before IMAC, the elution fractions feature less additional protein bands in the SDS-gel of the stain-free channel, showing the efficiency of the first affinity chromatography by Ni²⁺-NTA. Nevertheless, a tandem-affinity strategy was necessary as a small proportion of contaminants co-eluted with the POI. Therefore, a second affinity chromatography was performed via the double C-terminal StrepII tag. The flow-through and washing fractions showed weak fluorescence protein bands at 170 kDa whereas the elution fractions 3-5 contain one intense fluorescence band at 170 kDa which could be analysed as the POI which was additionally present in the stain-free channel (see Figure 6.9 A, left). The isolated POI sample was digested by the addition of HRV 3C protease that cleaves the C-terminal eGFP-purification-tag off the truncated Patched1 protein. A sample that was taken after the over-night incubation while dialysis showed two fluorescence bands at 170 kDa and 65 kDa with equal intensity (see Figure 6.9 A, centre). These results conclude that around 50% of the POI was digested as the C-terminal tag is around 35 kDa in MW. Nevertheless, half of the proportion was non-digested and should be isolated from the digested POI via reverse IMAC. The collected fractions of the reverse IMAC did not show any protein band although the fluorescent as well as stain-free image of the SDS-gel were considered (see Figure 6.9 A, centre). As the digested POI should not be immobilized, the flow-through as well as the first washing step (fraction 2 and 3 in centre gel, Figure 6.9 A) were combined and concentrated up to 500 µl and applied onto SEC in order to separate the POI from rest contaminations and aggregated protein particles. Here, two main peak fractions eluted at approximately 15 ml and 17.5 ml elution volume (see Figure 8.6 B). The first peak fractions contained a mixture of the Patched1 constructs as two main fluorescent protein bands at 170 kDa and 35 kDa were observed. Nevertheless, in the stain-free vimage only the band at 150 kDa appeared, indicating a low proportion of the full-length protein construct 3 and cleaved C-terminal tag, compared to the digested POI of around 140 kDa. The second peak fractions were supposed to be composed of a C-terminal eGFP-purification tag and the protease mixture as both are around 35 kDa and a protein

band at around this MW was detected in the fluorescence as well as in the stain-free channel (see Figure 6.9 A, right).

An improved purity of the POI could be obtained by the design of another protein construct without the reporter tag eGFP in order to avoid the non-digested contaminations in the POI peak fractions at a final purification step. Here, the construct is composed of the truncated human Patched1 protein from 1-1210 followed by the C-terminal purification tags His₈ and the double StrepII. This new protein construct should enable the purification by first tandem-affinity chromatography and a final SEC without the contamination of a non-digested protein fraction.

Nevertheless, a stable POI could be separated from contaminants by tandem-affinity chromatography followed by an over-night proteolytic digest for the C-terminal eGFP-containing purification tag removal as well as a reverse Ni²⁺-NTA and a final SEC. The tandem-affinity chromatography included the purification first by Ni²⁺-NTA via the octa-Histidine (His₈) tag and second by Strep-Tactin via the double StrepII modifications at the C-terminus. The digested POI was additionally isolated from the added HRV 3C protease and the removed tag by a reverse Ni²⁺-NTA affinity chromatography since the protease and the tag were immobilized via their poly-Histidine tag. A small amount of non-digested Patched1 construct³ was co-eluted in the final SEC step which should be insignificant for structural investigations by electron microscopy. The 3.5 l expression cultures of the human truncated Patched1 resulted in 0.03 mg purified protein that could be used for structural analysis by negative stain electron microscopy.

6.1.4 Negative staining electron microscopy

The fraction of the retention volume at 15 ml with a protein concentration of 0.03 mg/ml (see Figure 6.9, B red) was used for an analysis via negative stain by Julia Kotschy under my guidance. (146) Subsequently, the purity and stability of the POI was analysed by negative stain EM

A representative negative stain electron micrograph showing the digested and truncated human Patched1 is shown in Figure 6.10 A and B.

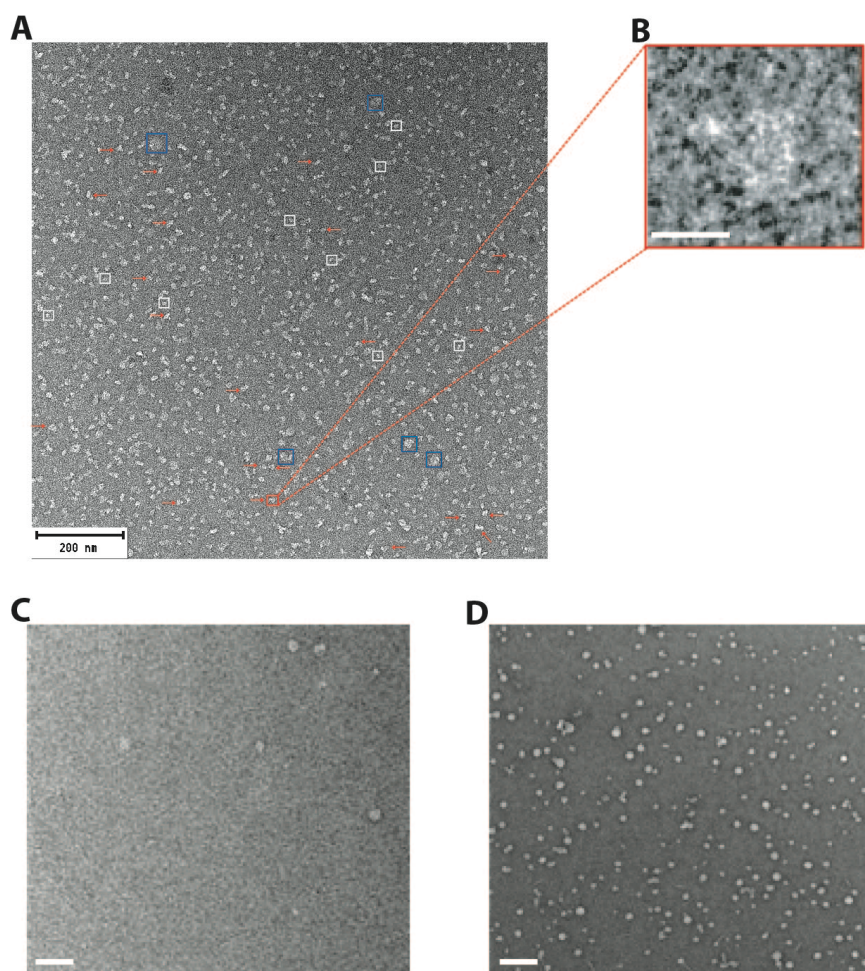


Figure 6.10: Comparison of negative stain electron micrographs of the truncated human Patched1 solubilized in DDM/CHS and DDM/CHS containing buffer.

Representative electron micrograph of the 15 ml peak fraction after SEC of the digested Patched1 protein construct 3 (A) and a close-up of a protein particle (B). The peak fraction shows three main populations, first small spherical particles of <10 nm diameter (white square), second globular particles of around 12-15 nm (red arrows) and large aggregated-like structures of >20 nm (blue square) under negative stain conditions. Black scale-bar: 200 nm and white scale bar: 15 nm. Electron micrograph of the solubilization buffer containing DDM/CHS at 0.05%/0.005% (v/v) (C) and 1%/0.1% (v/v) (D) concentrations. The spherical DDM/CHS detergent micelles differ in their diameter of around 15 to 20 nm which are less present at Patched1 buffer concentrations of 0.05%/0.005% (v/v). Scale-bars: 100 nm.

The analysed protein sample features a mixed protein population of particles with a diameter between 8 to 25 nm. Mainly three different populations were identified under negative stain conditions (see Figure 6.10 A). The first population features small globular shaped particles of <10 nm (white squares), whereas a second and most abundant particle fraction exhibits a spherical shape between 12-15 nm in diameter and a prominent extra density extracting out of the sphere (red arrows, see Figure 6.10 A and B). The third population is composed of round particles larger than 20 nm (blue square, see Figure 6.10 A). Furthermore, a representative negative stain micrograph of the buffer containing DDM/CHS, that was used for the final purification step, is shown in Figure 6.10 C and for the solubilization step in Figure 6.10 D, respectively. Here, globular shaped particles with a diameter of above 20 nm could be observed. These observations indicate a mixed population of empty detergent micelles and membrane

protein incorporated micelles in the sample of fraction 15 represented in Figure 6.10 A. The less abundant small particles of <10 nm could represent degraded protein particles or other smaller buffer compositions that could be visible under negative stain conditions. In addition, the protruding extra density of the particles with a diameter of around 12-15 nm suggest the incorporation of the membrane protein of interest into the detergent since the protein's structure is supposed to feature two enlarged ECD of about 300 amino acids, respectively (see Figure 6.10 B). (10)

Nevertheless, the overall distribution indicates a globular protein shape and furthermore, a membrane protein that is incorporated into a detergent micelle (see Figure 6.10 B). An additional analysis of the buffer containing DDM/CHS was considered in order to exclude an analysis of an empty detergent micelle. In accordance with the results of T. Gewering *et al.*, the used detergent mixture composed of DDM and CHS features globular particles above 20 nm under negative stain conditions when used at concentrations of 1% DDM/0.1% CHS (v/v) which is above its CMC of 0.009% (0.17 mM, see Figure 6.10 A). (150, 151) Unfortunately, the varying particle diameter and globular shape of the protein particles could hinder a proper analysis of the POI's structure as these are limiting factors for a 3D reconstruction, especially via negative stain electron microscopy as the resolution is limited to 20 Å. (152) Furthermore, the presence of the non-digested Patched1 construct 3 in the sample could raise problems during single-particle analysis since the protein's molecular weights differ in around 35 kDa which could be challenging to distinguish at these resolutions. As mentioned before, another protein construct without the eGFP reporter tag could improve the heterogeneity of the particle's distribution under negative stain electron microscopy. Here, the additional over-night protease digestion-step could be omitted and improve the protein's purity and stability. Furthermore, a detergent exchange of DDM/CHS after solubilization could improve the stability of the membrane protein's native state since this was not considered before but for example was applied for the structure analysis of the human NPC1 protein that shares high sequence identity with the transmembrane domain of the human Patched1. (153, 154) A FSEC screening after an exchange of DDM/CHS with a broad range of different detergents could be performed in order to overcome this obstacle. These sample improvements would be essential for a Patched1 structure analysis on a molecular level by cryo-EM.

6.1.5 Conclusion

The structures of the human membrane protein Patched1 alone and in complex with the mature Shh on a molecular basis would be inevitable for a complete understanding of the mechanism by which the signalling cascade of the Shh pathway is regulated as an imbalanced Shh signalling pathway can lead to severe diseases and birth defects. (74-78) Here, I present a truncated and stabilized human Patched1 membrane protein construct compared to the full-length protein featuring higher expression yield and less degradation as well as aggregation during purification. In total, 0.01 mg of the truncated Patched1 protein could be purified per liter cell culture that could be used for structural investigations by negative stain EM. At this stage of my research, several 3D reconstructions of mammalian Patched1 alone and in complex with its ligand Shh were published in the last two years ranging between 3.3 to 3.9 Å resolution. (5-10)

For structure analysis of the transmembrane protein, I started to identify a stabilized version of the Patched1 protein since several studies failed in the purification of an active mouse Patched1 protein. (137, 138) I tested the expression and solubilization in DDM of different Patched homologs including Patched1 and Patched2 via FSEC. Here, the C-terminal tagged human Patched2 was the most prominent construct. By a multi-sequence alignment, I visualized a high protein sequence identity of Patched1 and 2 but differences in their N- and C-termini which are elongated for Patched1 proteins (see Figure 8.3). Additionally, an ubiquitin ligase binding site at the C-terminus was identified in mouse Patched1 which is also present in the human homolog by Kim *et al.*. (143) Based on these results, I designed nine different N- and C-terminal truncated human Patched1 constructs and screened for a stabilized Patched1 construct by FSEC (see Figure 6.4). The construct 3 (1-1210) was most promising because of an increased expression level and less aggregation as well as degradation when solubilized in DDM. These results were in accordance with all published data as all studies have used a truncated mammalian Patched1 construct for structure determination. (5-10) Furthermore it was shown, that the N-terminal deletion of Patched1 did not improve the protein's stability since all protein constructs were composed of the full-length N-terminus. (5-10) Nevertheless, Y. Zhang *et al.*, X. Qi *et al.* as well as C. Qi *et al.* additionally modified the intracellular region localized between helix 6 and 7. (5-7) Most notably, Y. Zhang *et al.* have shown an increased expression level by a truncation and boosted stabilization via FSEC because of a second E3 ubiquitin ligase interaction site in this region. (5) In contrast, H. Qian *et al.* and X. Gong *et al.* have used a C-terminal truncated human Patched1 protein (1-1305) for a 3D reconstruction of the membrane protein alone and in complex with palmitoylated Shh on a molecular basis. (8-10)

The truncated protein construct 3 showed a correct localization in the biological membrane after over-expression in HEK293 cells by fluorescence microscopy indicating a correctly folded membrane protein (see Figure 6.5). Additionally, X. Qi *et al.*, H. Qian *et al.* and X. Gong *et al.* verified the proteins correct activity by a luciferase activity assay when working with a similar C-terminal truncated Patched1

construct. (6, 8, 9) Based on these data, I predict a similar activity for the truncated protein construct 3.

One of the most critical steps when working with transmembrane proteins is the extraction of the POI out of the biological membrane which is called solubilization. An improved solubilization efficiency could be shown when using the detergent mixture composed of DDM and CHS that I validated by a FSEC screen after solubilization with a wide range of non-ionic, ionic or zwitter-ionic detergents (see Figure 6.8). The detergent DDM is a mild non-ionic detergent that was successfully used for solubilization of various membrane proteins besides Patched1, like NPC-1. (6, 10, 153, 154) NPC-1 shares a high sequence homology with Patched1 proteins and equally to Patched1 comprises an SSD in its transmembrane region. (153, 154) The high homology of the transmembrane region could indicate that DDM is an efficient detergent for the solubilization. Nevertheless, the addition of CHS to DDM enables a detergent mixture that is most likely mimicking the native environment of mammalian Patched1 proteins most efficient. Via the FSEC detergent screen, an increased amount of stabilized Patched1 construct 3 could be solubilized by DDM/CHS which is consistent with the results of the experimental data received by Y. Zhang *et al.*, C. Qi *et al.*, H. Qian *et al.* as well as X. Gong *et al.*. (5, 7-9)

Finally, I established a purification strategy of a stabilized human Patched1 construct by tandem affinity chromatography, over-night eGFP-tag removal by protease digest, followed by a reverse IMAC and a final SEC step. By negative stain electron microscopy and semi-denaturing SDS-PAGE analysis a small proportion of co-eluted non-cleaved eGFP containing protein with the POI could be verified (see Figures 6.9 and 6.10). Here, a modified protein construct with an absent eGFP reporter tag could improve purity, stabilization and the purification yield since the presence of a non-digested protein as well as two purification steps could be omitted. This conclusion could be confirmed by all published structures of the mammalian Patched1 alone. (5, 6, 9) Y. Zhang *et al.*, X. Qi *et al.* and X. Gong *et al.* were able to purify the POI without the eGFP reporter tag for structure analysis by cryo-EM. (5, 6, 9) In contrast to the structural investigations by negative electron microscopy, the 3D reconstruction of the POI was not solved in DDM/CHS since all exchanged the detergent mixture to either Amphipols (PYMAL-C8) or Digitonin. (5, 6, 9) These findings could reveal that an exchange from DDM/CHS was necessary in order to improve the homogeneity of the protein particles and the stabilization of Patched1 for structure analysis by cryo-EM.

Surprisingly, Y. Zhang *et al.*, X. Qi *et al.*, X. Gong *et al.* and H. Qian *et al.* showed the presence of a mixed population of formed complexes after Patched1-Shh complex assembly. By SEC several eluted peak fractions could be observed indicating the stabilization of either two Patched1 molecules bound to one Shh protein (2:1) by Y. Zhang *et al.*, X. Qi *et al.* and H. Qian *et al.* or even higher oligomeric complex formations like H. Qian *et al.* composed of four Patched1 proteins and two Shh. (5, 6, 8-10) Comparatively, the formation of a higher oligomeric complex composed of Patched1 and the double lipidated mature Shh (4:2) is rather unlikely of natural occurrence since this was stabilized by crosslinking for structural investigations. (8) Besides these observations, C. Qi *et al.* published the structure of a heterodimer protein complex composed of one Patched1 protein and one Shh protein

(1:1). (7) in comparison with the other used Protein constructs, C. Qi *et al.* used a human Patched1 protein with a single mutation Y645A as this is known to be less dedicated for ubiquitination and might have an impact on a higher oligomeric complex formation compared to the other protein constructs. (5-10)

Taken together, several published 3D reconstructions of the complex formation of a mammalian Patched1 and Shh protein showing different orientations of the Shh-bound conformations (see Figure 2.18). All data show an active mammalian Patched1 protein construct that slightly varied in its molecular weight. Nevertheless, all structures revealed a similar binding interface that is located between the enlarged ECD 1 and 2. (5-10) In the case of palmitoylated Shh proteins, the binding site of this post-translational modification could be well resolved and was localized to the inner conduit of Patched1 by binding to ECD 1. (6, 8-10) Furthermore, the electron density of the cholesterol moiety in the structures solved by Y. Zhang *et al.* as well as X. Qi *et al.* could not be visualized and remains unclear to date. (5, 6, 10) Since all complex structures composed of Shh and a truncated mammalian Patched1 protein feature different stoichiometries as well as high variations in their conformation and binding motifs, a clear 3D reconstruction that shows the native state is remaining. Presumably, the differences in the reconstructions could occur due to different Patched1 constructs as well as different modified Shh ligands or indicate a highly flexible protein complex. (5-10)

Although many differences could be observed when comparing all published reconstructions of the mammalian Patched1 protein, all structures visualized the presence of several additional densities that would indicate the presence of putative binding sites of sterols. Besides X. Qi *et al.*, all Patched1 proteins were solubilized in CHS containing buffer that could be fitted into these extra densities. The assumption of sterol binding sites can be enabled as CHS is a known cholesterol analogue. (5-10) Consequently, all publications describe a putative mode of action of the mammalian Patched1 protein that functions as a sterol sensor that enables the efflux of cholesterol molecules from the inner membrane leaflet to the outer leaflet and regulates the cholesterol concentrations in these areas. (5-10) These results would confirm the hypothesis that a second messenger, like cholesterol in this case, would control the activity of the receptor Patched1 and furthermore the downstream activity of Smo. (5) It was postulated before that a second messenger would be inevitable for the signal transfer, as Smo and Patched1 failed in binding. (155, 156) Y. Zhang *et al.* therefore hypothesised that Patched1 in vertebrates regulates the cholesterol concentration in the inner cell membrane region by acting as a cholesterol transporter from the inner to the outer membrane part in its active and Shh unbound state. Upon ligand binding, Patched1 is inactivated since its inner conduit is blocked by mature Shh. At high cholesterol concentrations in the inner membrane layer Smo is bound to cholesterol and activates downstream transduction of the Hh signalling pathway. (5, 157, 158) This hypothesis coincides with the results of all published structures of the mammalian Patched1 in complex with its ligand Shh. (6-10) Nevertheless, the presence of an extra density in the outer transmembrane area remains unclear since cholesterol is hydrophobic and would provide the presence of a carrier molecule for keeping the cholesterol soluble in the aqueous solution. (5-10) This postulation was furthermore addressed by C. Kowatsch *et al.* which represented two possibilities that included the presence of such a carrier

molecule for cholesterol and a third one in which Patched1 acts like a flippase for cholesterol without the need of a carrier. (157) The carrier molecule could either directly bind cholesterol and transfer it from Patched1 to Smo or most unlikely adapt the cholesterol from Smo and transfer it to Patched1. The latter scenario could be excluded as several structures contain extra densities for a sterol in the inner transmembrane region which would be not required here. (157)

All published structures of the Patched1 protein alone and in complex with its ligand Shh shed light in the recognition and binding interface on a molecular level and furthermore the regulation of Smo activation by Patched1 inactivation and the downstream Hh signalling cascade. Here, the structures of mammalian Patched1 reveal evidence of a cholesterol transporter function from the inner membrane to the outer membrane layer which is blocked by Shh complex formation that hinders Patched1 suppressing Smo activity. (5-10, 157)

6.2 Scap

Since the discovery of Scap in 1996 in the lab of J. L. Goldstein and M. Brown, many structural features and its overall topology are known mainly based on the analysis of the hamster Scap homolog. (55) The membrane protein Scap was either over-expressed in *E. coli* or Sf9 insect cells for structural analysis of the soluble fission yeast WD40 repeats domain or for cholesterol binding studies of the transmembrane domain. (63, 66, 138, 139) Nevertheless, we are still lacking a structure of a mammalian full-length Scap. For this purpose, I decided to use the mammalian cell line HEK293 for the expression of Scap membrane proteins in order to keep the environment as native as possible. One of the main obstacles when working with membrane proteins are their special hydrophobic native environment compared to soluble proteins which make them more unique and beyond that less stable. By using a near native expression host the membrane protein of interest should be stabilized during expression in contrast to insect cells or bacteria like *E. coli* which might be a reason for the lack of structural information on mammalian full-length Scap nowadays. In addition to that, the structure analysis of many different multi-spanning trans-membrane proteins, like Patched1 or Dispatched-1 revealed a stabilized protein while expression, solubilization and purification by using the human HEK293 cell line as expression system. (5-10, 92)

6.2.1 Validations for a stable Scap construct

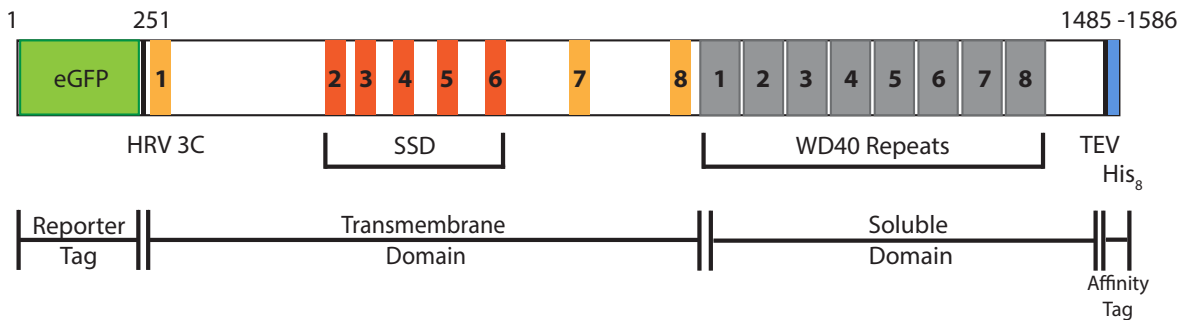
Similar to the approach for the evaluation of a stable Patched protein, I started to screen for the most stable Scap homolog via FSEC. Based on a multi-sequence alignment (see Figure 8.5) I selected in total nine homologs that share a high sequence identity and query coverage with the human Scap (see Table 6.2).

Table 6.2: Homologs of Scap proteins sharing high sequence identity with the human Scap homolog that were used for the homology screen.

Protein	Organism	Sequence identity [%]	Query coverage [%]
Scap	<i>Cricetulus griseus</i>	92.03	100
Scap	<i>Danio rerio</i>	68.78	100
Scap	<i>Drosophila melanogaster</i>	32.74	87
Scap	<i>Felis catus</i>	92.66	100
Scap	<i>Gallus gallus</i>	74.70	100
Scap	<i>Homo sapiens</i>	100	100
Scap	<i>Mus musculus</i>	91.95	100
Scap	<i>Rattus norvegicus</i>	92.49	100
Scap	<i>Sus scrofa</i>	92.27	100

All nine different homologs were cloned into the pEG BacMam vector for the expression of the protein of interest in mammalian cells. The homologs containing either N-terminal or C-terminal modified eGFP tags were screened for their expression level, expression pattern, solubilization and stability in a mild detergent DDM via fluorescence microscopy, FSEC and SDS-PAGE analysis. A scheme of the proteins topology either containing N-(top) or C-terminal reporter tags (bottom) are shown in Figure 6.11.

N-terminal eGFP-Scap constructs



C-terminal eGFP-Scap constructs

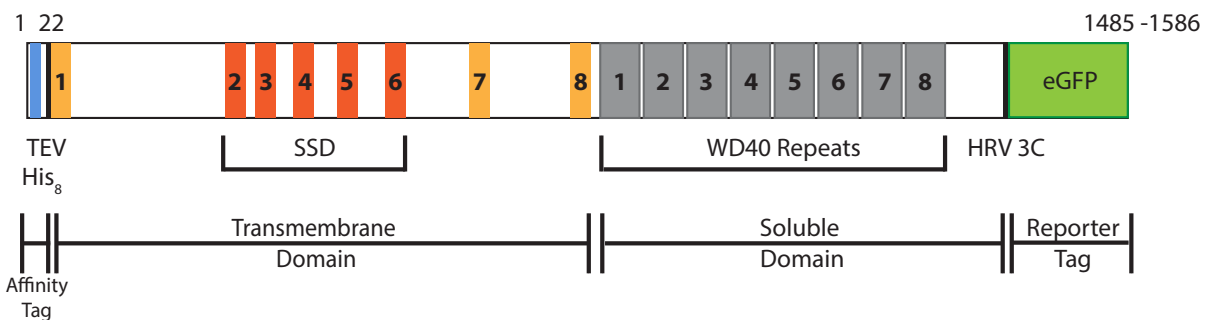


Figure 6.11: General setup of Scap constructs used for the homolog screen.

In total nine different Patched homologs were used containing a N-terminal eGFP (green) reporter Tag and C-terminal poly-Histidine (His₈) (blue) affinity purification tag (top) or vice versa (bottom). All constructs shared high sequence identity and query cover with the human Scap homolog and are supposed to be composed of 8 transmembrane domains (red and orange) and eight WD 40 repeat domains (grey). The sterol sensing domain is highlighted in red and located from helix 2 until helix 6.

First, the expression level and pattern were analysed via the eGFP fluorescence intensity and its localization in biological membranes by fluorescence microscopy. The cells were transiently transfected with the DNA encoding the 18 different protein constructs that were over-expressed for 48 h (see Figure 8.6). Selected fluorescence images of mammalian cells expressing either N- (left column) or C-terminal (right column) eGFP modified *H. sapiens*, *M. musculus*, *R. norvegicus* and *S. scrofa* Scap homologs are shown in Figure 6.12.

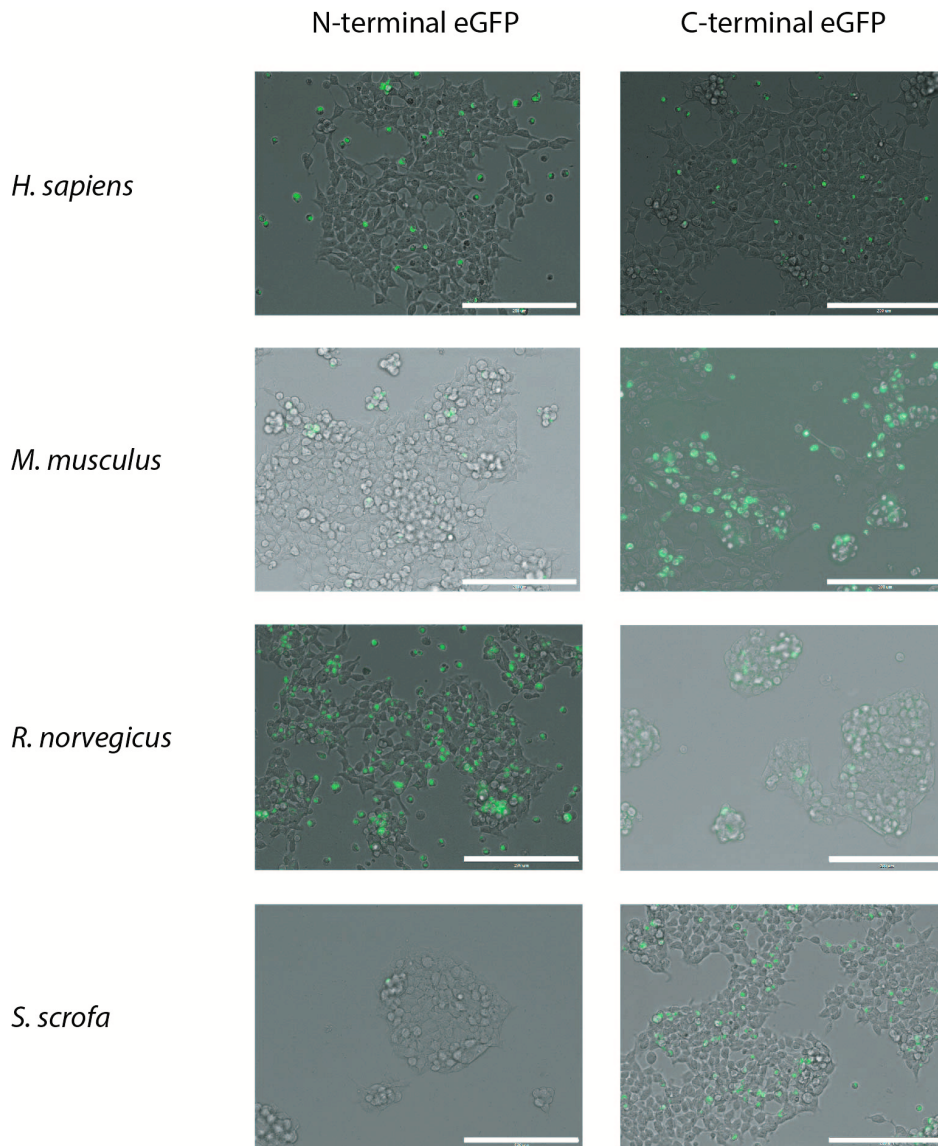


Figure 6.12: Selected fluorescence images of transiently transfected adherent HEK293 T cells after 48 h expression of Scap homologs.

Merged images of the eGFP fluorescence channel (green) and the bright field (grey) of selected Scap constructs used for the homolog screen taken with the EVOSFL Auto Fluorescence microscope (Life Technologies) either containing N-(left column) or C-terminal (right column) reporter tags. Scale bars: 200 μm .

For all tested eGFP-modified Scap homologs, eGFP fluorescence could be observed after 48 h of expression in adherent HEK293 T cells (see Figures 6.12 and 8.6). The number of eGFP-expressing cells for the Scap homologs *C. griseus*, *D. rerio*, *D. melanogaster*, *F. catus* as well as *G. gallus* was lower compared to the homologs of *H. sapiens*, *M. musculus*, *R. norvegicus* and *S. scrofa* (see Figures 6.12 and 8.6). Furthermore, all Scap constructs show fluorescence at the cell membrane, whereby the *M. musculus* and *R. norvegicus* Scap homologs showed additional fluorescence inside the cells. In comparison with all screened Scap homologs the C-terminal modified *M. musculus* and N-terminal modified *R. norvegicus* homologs show the highest amount of eGFP expression (see Figure 6.12).

Therefore, these results show that eGFP-containing proteins were expressed and furthermore correctly localized at membrane regions in the mammalian HEK293 cell line. The localization of most

of the eGFP fluorescence at membrane regions indicate the expression of the membrane protein Scap. Some high expressing eGFP containing proteins like the C-terminal eGFP-modified *M. musculus* or the N-terminal eGFP-modified *R. norvegicus* homologs additionally featured whole cell distributed eGFP fluorescence signals that were not attached to the surface. In these few cases, the high over-expression of the protein constructs led to cell death including aggregation of eGFP-containing proteins. The Scap homologs from *C. griseus*, *D. rerio*, *D. melanogaster* and *G. gallus* share less sequence identity to the human version which would explain the decreased expression levels compared to the *H. sapiens*, *M. musculus*, *R. norvegicus* and *S. scrofa* Scap homologs since the expression was performed in human embryonic kidney (HEK) cells.

These results either demonstrate the expression of the full-length Scap homologs or the expression of degradation products including the eGFP-reporter tag. In order to exclude the expression of degradation products and verify the proteins stability as well as the stability when solubilized in a detergent micelle, the FSEC screen and a subsequent semi-denaturing SDS-PAGE analysis of all 18 constructs were performed. The chromatograms and the SDS-gels were represented in Figure 6.13.

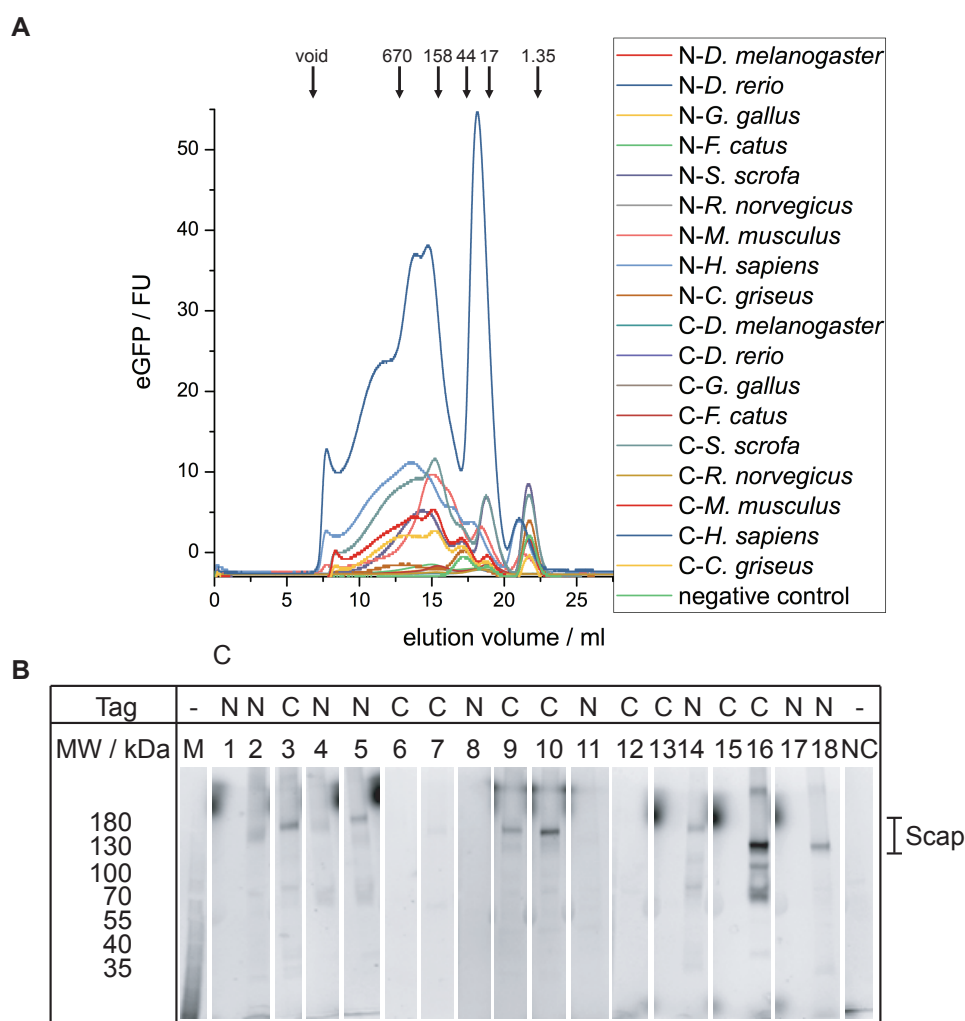


Figure 6.13: FSEC chromatograms and SDS-PAGE-gels of the Scap homolog screen.

(A) FSEC chromatograms of the 18 different Scap constructs and a negative control (NC, without expression of Scap) after solubilization in buffer containing 1% DDM. (B) Fluorescence image of the SDS-PAGE-gel from the DDM-solubilized 9 Scap homologs either containing N- (N) or C- (C) terminal eGFP modifications and one NC (green) taken with the ChemiDoc MP imaging system (Bio-Rad) as followed: M: PageRuler Prestained Protein Ladder, 1: *D. melanogaster* (N-eGFP), 2: *D. rerio* (N-eGFP), 3: *M. musculus* (C-eGFP), 4: *M. musculus* (N-eGFP), 5: *R. norvegicus* (N-eGFP), 6: *D. rerio* (C-eGFP), 7: *F. catus* (C-eGFP), 8: *F. catus* (N-eGFP), 9: *R. norvegicus* (C-eGFP), 10: *S. scrofa* (C-eGFP), 11: *G. gallus* (N-eGFP), 12: *G. gallus* (C-eGFP), 13: *D. melanogaster* (C-eGFP), 14: *S. scrofa* (N-eGFP), 15: *C. griseus* (C-eGFP), 16: *H. sapiens* (C-eGFP), 17: *C. griseus* (N-eGFP), 18: *H. sapiens* (N-eGFP) and NC: negative control without Scap over-expression. The scale bar indicates the range of the molecular weights of the POI.

The chromatograms of the FSEC analysis show mainly three peaks in the range of 8 to 16 ml elution volume. The first elution fraction at 8 ml is for most of the protein homologs not properly separated from the second peak at around 10 to 13 ml elution volume. Furthermore, the second peak turns into the third peak fraction at around 15 ml. The overall fluorescence intensity of the first main peak at around 8 ml is the lowest compared to all other peaks and refers to the void volume in which aggregated and fluorescently labelled molecules elute. The second peak fraction elutes at about 3 ml and comprises the elution of fluorescent biomolecules with a molecular weight of ≥ 600 kDa. These findings indicate the elution of higher oligomers consistent of at least homotrimers incorporated into

the DDM detergent micelle. For most of the Scap homologs, this peak is not separated from the final last main peak fraction at 15 ml elution volume. In the third peak fraction, the elution of particles with a molecular weight of around 250 kDa could be observed that would coincide with the calculated MW of 180 kDa of the protein constructs and the additional 70 kDa DDM micelle. (140) Later eluting peak fractions indicate the presence of biomolecules that were fluorescent and smaller in their hydrodynamic radius compared to the full-length POI constructs, which could indicate the degradation of the eGFP modified terminal region. Nevertheless, the third main peak fraction of all screened homolog constructs show the highest fluorescence intensity as well as a sharp and gaussian-like peak indicating the stabilisation of the monomeric protein structure when solubilized in DDM. These results were in accordance with the semi-denaturing SDS-PAGE analysis (see Figure 6.13 B). Most of the DDM solubilized Scap constructs featured a high intense fluorescent protein band at roughly 180 kDa. The N-terminal eGFP-modified *S. scrofa* (sample 14) or *R. norvegicus* (sample 16) as well as the C-terminal eGFP-modified *H. sapiens* (sample 5) Scap homologs featured additional fluorescent protein bands at around 100 and 70 kDa (see Figure 6.13 B). The additional fluorescent bands indicate the degradation of the POI while DDM solubilization and furthermore a destabilized full-length protein construct. Contrary to that, some of the POI constructs were not solubilizable by DDM as they did not feature any fluorescence elution peak between 12 to 15 ml during FSEC screening nor fluorescence protein band in the semi-denaturing SDS-PAGE gel after solubilization, e.g. the C-terminal modified *D. rerio* (sample 6), the *G. gallus* (sample 12) or the N-terminal modified *D. melanogaster* (sample 1) Scap homolog (see Figure 6.13). On the other side, the porcine Scap homolog containing a C-terminal eGFP reporter tag showed less degradation products by SDS-PAGE and FSEC analysis as well as a gaussian-like monodisperse elution peak including the highest fluorescence intensity at 15 ml elution volume (sample 10, see Figure 6.13). Therefore, the *S. scrofa* Scap was identified as most suitable homolog construct for structure analysis. Based on these results, a potent baculovirus was amplified and used for the heterologous over-expression of porcine Scap in suspension HEK293 cells. (102)

The correct folding of the porcine Scap homolog was analysed after infection of adherent HEK293 T cells via its localization in the ER membrane by fluorescence microscopy (see Figure 6.14).

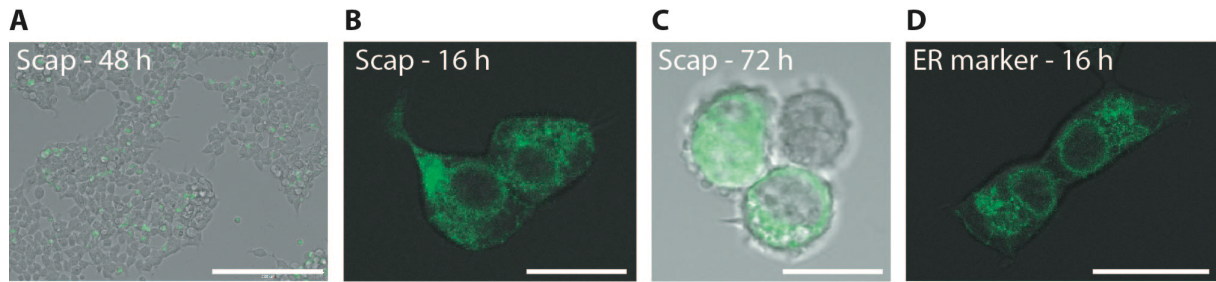


Figure 6.14: Fluorescence microscopy of adherent HEK293 T cells over-expressing C-terminal eGFP-modified porcine Scap.

Merged images of the GFP-fluorescence channel (green) and the bright field channel (grey) (A) taken with the EVOSFL Auto Fluorescence microscope (Life Technologies) or (B-D) with the Zeiss LSM 800 microscope by using Airyscan at 40-fold magnification (Carl Zeiss Microscopy GmbH). (A) Transiently transfected adherent HEK293 T cells after 48 h expression of porcine Scap showing 30% eGFP fluorescently labelled cells. Scale bar: 200 μm . (B) Adherent HEK293 T cells 16 h post-infection of C-terminal eGFP-modified porcine Scap. The POI is localized at the ER membrane and in additional biological membranes of the cells. Scale bar: 20 μm . (C) Adherent HEK293 T cells 72 h post-infection of C-terminal eGFP-modified porcine Scap. The infected cells feature a globular shape with an increased amount of eGFP expression indicating cell death at 72 h post-infection. Scale bar: 20 μm . (D) Adherent HEK293 T cells after 16 h baculovirus infection using CellLight ER-GFP (BacMam 2.0, Invitrogen, Thermo Fisher Scientific). Scale bar: 30 μm .

Around 30% of all transiently transfected adherent HEK293 T-cells showed eGFP expression mainly localized at biological membranes like in the plasma membrane (see Figure 6.14). At higher magnification, the eGFP localization could be analysed more concise. At 40-fold magnification the expression pattern of the porcine Scap construct featured preferred localization in ER-membranes (see Figure 6.14, B) that could be verified via the comparison with the CellLight ER-GFP (BacMam 2.0, Invitrogen, Thermo Fisher Scientific) marker that is shown in Figure 6.14, D. Due to over-expression of the POI additional eGFP fluorescence could be observed at further biological membranes like the plasma membrane (see Figure 6.14, B). After 72 h post-infection, the over-expression led to an accumulation of eGFP containing proteins in the whole cell and furthermore induced cell death as recorded at 40x magnification in Figure 6.14, C. As mentioned before, the vertebrate Scap is expressed in the ER membrane where it controls the movement of SREBP via the COPII vesicular trafficking from the ER membrane to the Golgi membrane in a sterol dependent manner. (55, 59) Therefore, an increased number of Scap molecules besides the ER membrane was expected since the cells were grown under sterol deprivation that induces the COPII vesicular trafficking of the heterodimeric protein complex composed of Scap and SREBP to the Golgi membrane. (59) All in all, these data show the expression of a correct folded C-terminal eGFP modified porcine Scap when expressed for up to 48 h.

Taken the analysis of the homolog screen via FSEC and the localization studies by fluorescence microscopy into account, the porcine C-terminal eGFP-modified Scap construct was identified as most suitable for structural analysis via electron microscopy. The protein construct showed a high expression in the mammalian cells as well as a membrane localized expression pattern that fit the protein's localization at either ER or Golgi membranes. (55, 59) The solubilization by a mild detergent

DDM was most efficient for the porcine Scap and furthermore showed a stable protein construct as low degradation products and a monodisperse, gaussian-like sharp peak could be observed by FSEC.

6.2.2 Heterologous expression and purification strategy of porcine Scap

After the identification of a stable Scap homolog, the heterologous expression in mammalian suspension cells was established. The mammalian human embryonic kidney cell line HEK293 was used as the expression system features a near native membrane composition for the POI. Especially for the production of membrane proteins, the composition of the biological membranes should be nearly native in order to transcribe a correctly folded membrane protein. Previous experiments by a former lab member failed in an efficient solubilization in a mild detergent of a mammalian Scap homolog when expressed in insect cells. Here, the hamster Scap could be invariably solubilized by the ionic detergent Fos-Choline which most probably destabilized the protein's native state. (158) Furthermore, the mammalian cell line was successfully recombinantly expressing other membrane proteins that could be used for structure analysis via cryo-EM including Patched1, NPC1 as well as Dispatched. (5-10, 153, 154, 160) These membrane proteins share a high sequence identity in their transmembrane region as all contain the five transmembrane spanning SSD. (55)

The expressions of either 4.5% or 9% (v/v) P2 virus infected HEK293 F or GnTI⁻ cell lines were screened for up to 72 h. A sample was taken after 24 h of expression of each condition and used for an analysis via fluorescence microscopy and semi-denaturing SDS-PAGE, respectively (see Figure 6.15).

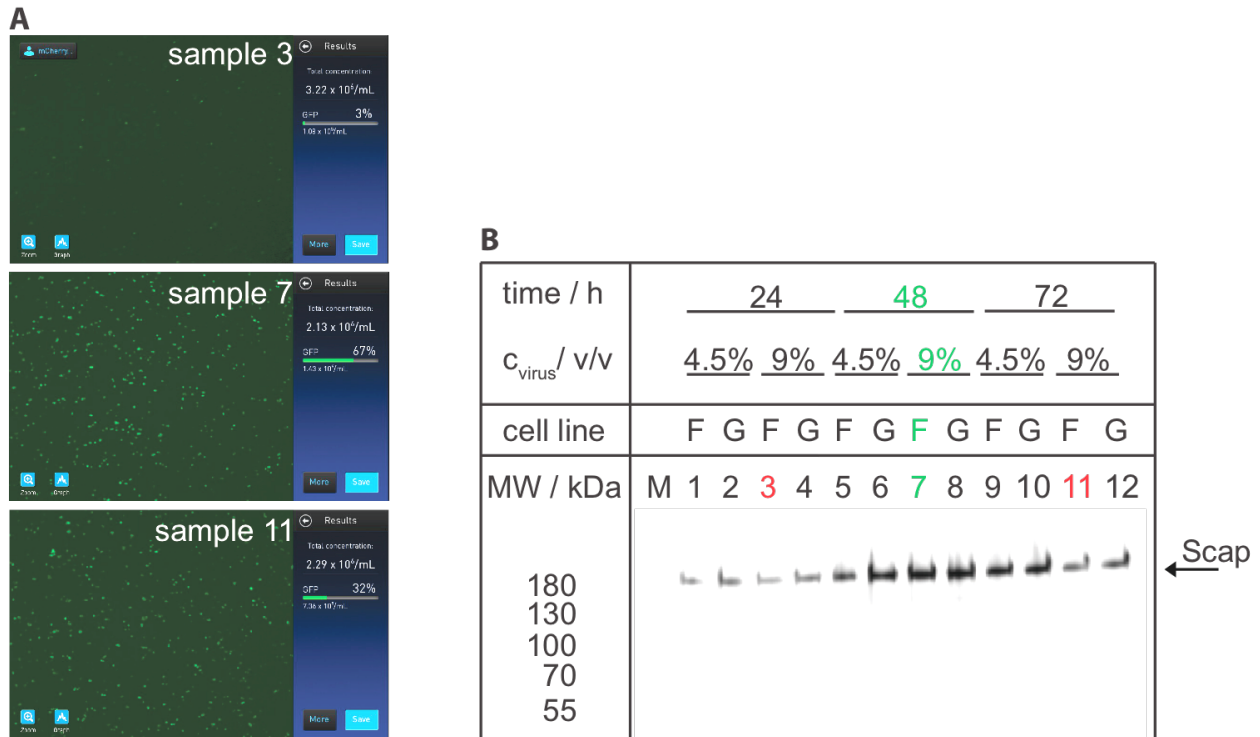


Figure 6.15: Test-expression of the porcine eGFP-Scap homolog in suspension HEK293 F or GnTI⁻ cells. The HEK293 F or GnTI⁻ cells were infected with 4.5% or 9% potent P2 virus and expressed for up to 72 h at 30 °C after 10 mM sodium butyrate induction. The expression was screened and validated by (A) eGFP expressing cells via the Countess II FL Automated Cell Counter (Thermo Fisher Scientific) representatively showing the expression of 9% infected F cells after 24 h (sample 3, top), 48 h (sample 7, center) and 72 h (sample 11, bottom) and (B) the expression of the eGFP-modified full length construct by semi-denaturing SDS-PAGE in the fluorescence channel taken with the ChemiDoc MP imaging system (Bio-Rad). Samples were applied as followed: M: PageRuler Prestained Protein Ladder, 1: 4.5% virus infected F cells after 24 h expression, 2: 4.5% virus infected GnTI⁻ cells after 24 h expression, 3: 9% virus infected F cells after 24 h expression, 4: 9% virus infected GnTI⁻ cells after 24 h expression, 5: 4.5% virus infected F cells after 48 h expression, 6: 4.5% virus infected GnTI⁻ cells after 48 h expression, 7: 9% virus infected F cells after 48 h expression, 8: 9% virus infected GnTI⁻ cells after 48 h expression, 9: 4.5% virus infected F cells after 72 h expression, 10: 4.5% virus infected GnTI⁻ cells after 72 h expression, 11: 9% virus infected F cells after 72 h expression, 12: 9% virus infected GnTI⁻ cells after 72 h expression. The arrow indicates the MW of the POI Scap.

The recombinant expression of the POI after 9% (v/v) infected F-cells was illustrated in Figure 6.15, A. Here, the percentage of eGFP expressing cells increased from 3% after 24 h post-infection up to 67% at 48 h post-infection and decreased to 32% of fluorescently labeled cells after 72 h expression. In order to exclude the expression and validation of degradation products containing eGFP, a semi-denaturing SDS-PAGE was performed subsequently (see Figure 6.15, B). The expression of the full-length recombinant porcine Scap could be confirmed by SDS-PAGE as a prominent fluorescent protein band at above 180 kDa could be observed that fit with the calculated MW of around 175 kDa. The intensity of this protein band is increased after 48 hours post-infection compared to the corresponding sample after 24 h expression. Nevertheless, the expression the protein band at around 180 kDa was less intense for all tested conditions after 72 h expression and additional fluorescent protein bands appeared at around 100 kDa and 70 kDa in comparison with the samples taken after 48 h, respectively

(see Figure 6.15, B). All in all, these results indicate a degradation of the POI between 48 h and 72 h post-infection for the 4.5% and 9% (v/v) P2 virus infected F as well as GnT1⁻ cell expression. The over-expression of Scap can lead to an imbalanced cholesterol homeostasis as this is highly regulated by Scap. At up-regulated Scap expression levels an increase of COPII vesicular trafficking from the ER- to the Golgi-membrane could occur that furthermore can lead to an up-regulated sterol synthesis. (55) The cells might activate different degradation pathways in order to lower the Scap level and keep the sterol level balanced. Nevertheless, this would indicate the expression of an active porcine Scap under all tested expression conditions. By comparing the expression level of both tested cell lines, the HEK293 F as well as GnT1⁻ cells yielded in similar expression of the recombinant POI. Due to several known Asparagine modified glycosylation sites at residues 263, 590 and 641, a preferred recombinant expression of the native porcine Scap in F cells was expected as these are able to build complex N-glycans compared to the GnT1⁻ cell line, that lacks in an active N-acetylglucosaminyltransferase I. (55) Therefore, the expression in HEK293 F cells after 48 h post-infection of 9% P2 virus was validated as most efficient since the POI showed highest expression level as well as lowest degradation (see sample 7, Figure 6.15). Furthermore, the expression of the porcine Scap protein construct could be up-scaled to 25 l of cell expression without loss of full-length recombinant protein.

In order to identify the most efficient detergent for the extraction of Scap from its membranes a FSEC analysis was performed after solubilization with a high variety of detergents. Here, detergents that were either ionic, zwitterionic or non-ionic were screened and validated for their efficiency by FSEC and semi-denaturing SDS-PAGE analysis (see Figure 6.16).

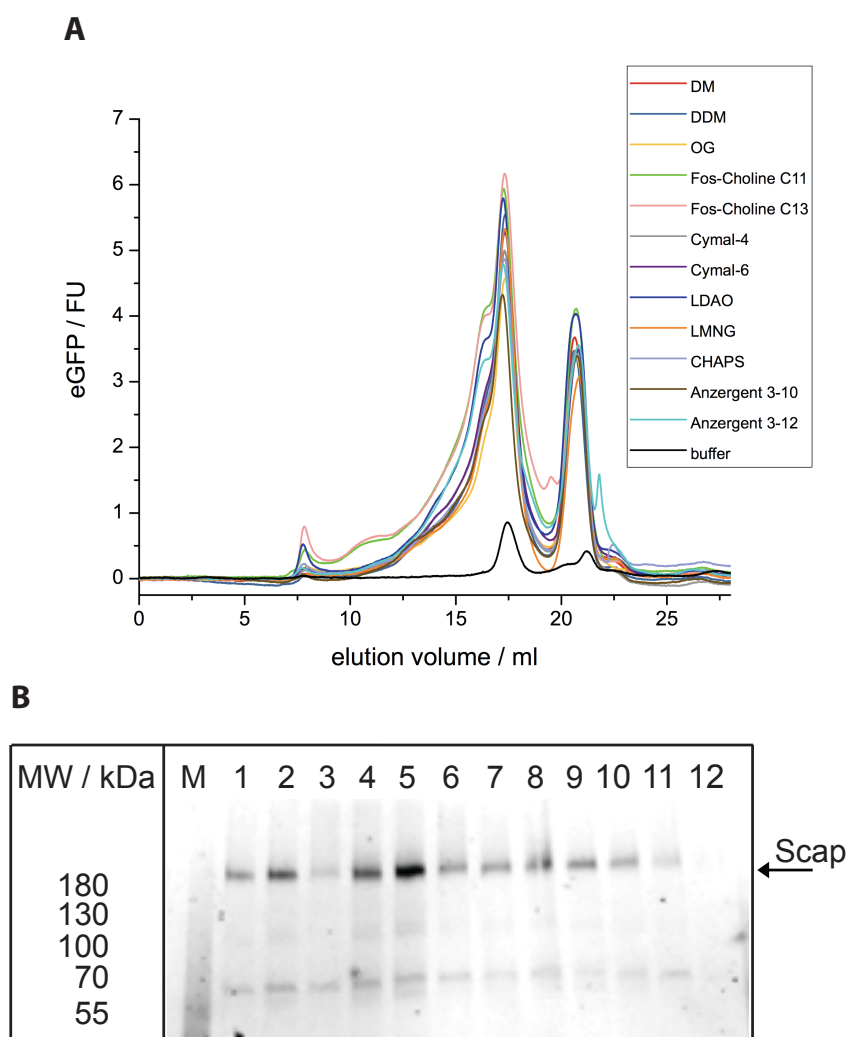


Figure 6.16: Detergent screen of the porcine C-terminal eGFP-modified Scap.

(A) The 13 chromatograms of the tested detergents were depicted as followed: red: DM (sample 1), blue: DDM (sample 2), yellow: OG (sample 3), green: Fos-Choline C11 (sample 4), pink: Fos-Choline C13 (sample 5), grey: Cymal-4 (sample 6), purple: Cymal-6 (sample 7), dark blue: LDAO (sample 8), orange: LMNG (sample 9), light blue: CHAPS (sample 10), brown: Anzergent 3-10 (sample 11), cyan: Anzergent 2-12 (sample 12), black: buffer without the addition of detergent (negative control) and feature three main peaks, including the void volume at 8 ml, one brought peak at around 12 ml and a third eluting between 15 and 17 ml including a small shoulder peak fraction that is not properly separated. (B) Fluorescence image of the semi-denaturing SDS-PAGE gel taken with the ChemiDoc MP imaging system (Bio-Rad) showing first M: PageRuler Prestained Protein Ladder followed by samples 1-12. The arrow indicates the MW of the full-length Scap construct.

The chromatograms of the detergent screen feature three main peaks in the range of 8 to 20 ml. A low intense first peak fraction elutes at around 8 ml and refers to the elution of aggregated fluorescent protein particles. The second peak is most dominant for the zwitterionic detergents Fos-Choline C11 (green) and C13 (pink) at roughly 12 ml of elution volume. Here, eGFP containing particles with a MW of more than 700 kDa elute. The main peak fraction elutes between 15 to 17 ml, in which particles with a MW of about 160 kDa are supposed to elute when compared with the protein standard shown in Figure 6.13, A. For all screened detergents, the third main fraction co-elutes with a second peak

fraction (shoulder) at earlier retention volumes of 15 ml which most probably refers to the elution of the monomeric full-length POI. Furthermore, these results indicate the stability of various oligomeric protein states that could not be separated by SEC (see Figure 6.16, A). The SDS-gel depicted in Figure 6.16 B shows two striking fluorescent protein bands at 180 kDa and a less intense band at 60 kDa. These results show the solubilization of the full-length POI by all tested detergents and most probably the solubilization of a C-terminal degradation product (see Figure 6.16, B). Notably, the zwitterionic detergents Fos-Choline C11 and C13 show the highest fluorescence intensity of the third peak fraction by FSEC as well as a protein band at 180 kDa at semi-denaturing SDS-PAGE (see Figure 6.16). Nevertheless, the void volume as well as the amount of degradation products appeared more prominent compared to all other screened detergents that implies a strong solubilization efficiency including the disruption of intramolecular protein-protein interactions. These results are consistent with the results of T. E. Cleveland IV *et al.*, who used Fos-Choline for the solubilization of a vertebrate membrane protein Patched1 that failed in interacting with its ligand and indicated strong protein-protein interactions, preferentially induced by Fos-Choline that modified its native conformation. (138) Therefore, DDM (blue) was identified as most efficient solubilization detergent because this featured an intense full-length band at 180 kDa at semi-denaturing SDS-PAGE analysis and a high intense main peak fraction at 15 to 17 ml including a low intense void volume (first peak, 8 ml elution volume) (see Figure 6.16). Furthermore, DDM was successfully used for solubilization of a wide range of membrane proteins including SSD containing proteins like Patched1, NPC1 and Dispatched that could indicate a stabilization of the five transmembrane spanning region SSD by DDM. (5-10, 92, 153, 154, 160) Nevertheless, the results of the detergent screen showed that the protein construct could not be separated from contaminants and furthermore need to be improved in order to increase its stability. Therefore, a new porcine Scap construct was designed containing the full-length protein sequence and an additional N-terminal purification tag for tandem affinity chromatography (octa-Histidine and double StrepII tag) (see Figure 6.17, A). The new protein construct show similar protein expression yields in the mammalian HEK293 F-cells and was supposed to be as active as the porcine Scap containing the C-terminal eGFP reporter tag since all functional domains were maintained. The purification of the new protein construct by first tandem affinity chromatography including IMAC (Ni²⁺-NTA) and Strep-Tactin and a final SEC was established (see Figure 6.17, B). The results of a purification of the porcine Scap was representatively shown in Figure 6.17 B and C. The new designed porcine Scap construct was solubilized in a detergent mixture composed of DDM/CHS as the addition of CHS was supposed to mimic the presence of cholesterol and therefore featured in increased similarity to the native environment of Scap. The solubilization of DDM/CHS was as efficient as DDM alone and therefore used for further purifications.

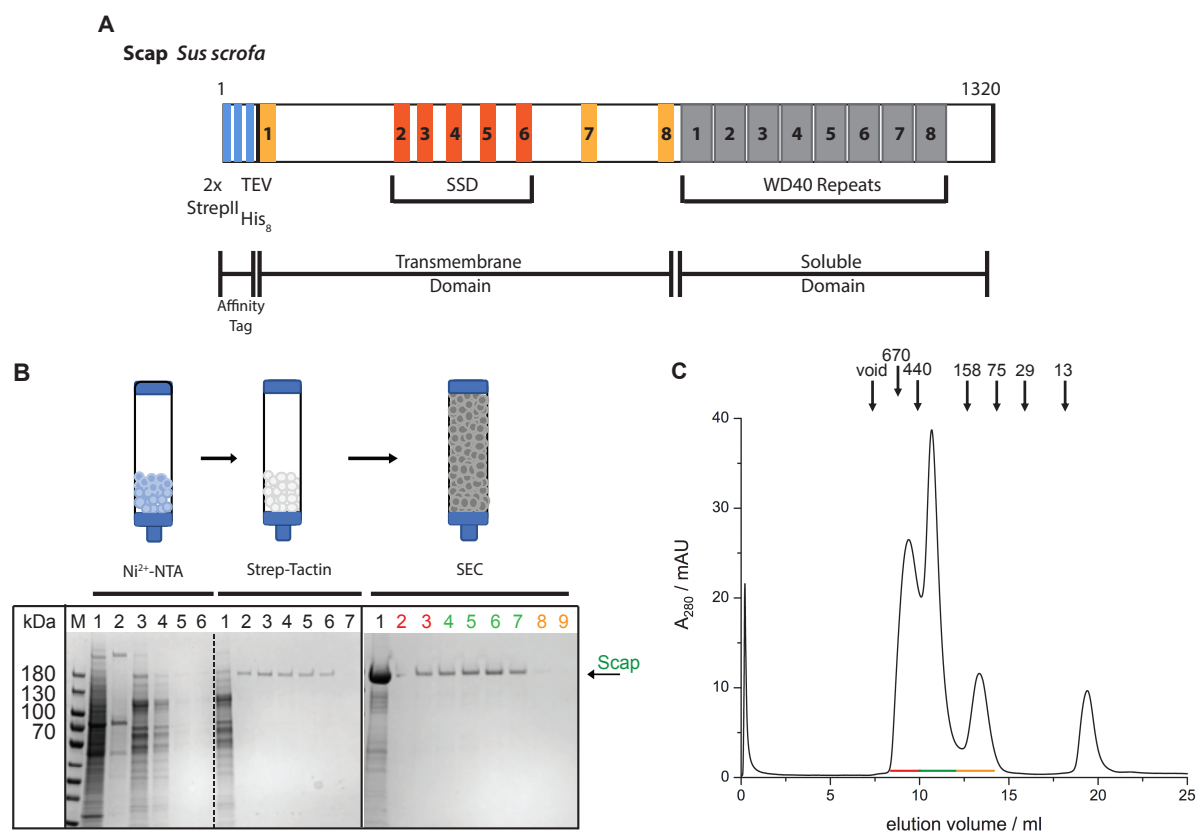


Figure 6.17: Final porcine Scap construct and a representative purification.

(A) Scheme of the porcine Scap construct that was used for further structure analysis by cryo-EM. The protein construct contains the full-length protein sequence and the additional N-terminal purification tags: double StrepII (2x StrepII) and octa-Histidine (His₈). (B) The final purification strategy including the tandem affinity chromatography as well as a final SEC and the representative analysis of elution fractions by semi-denaturing SDS-PAGE after Coomassie-staining taken with the ChemiDoc MP imaging system (Bio-Rad) showing first M: PageRuler Prestained Protein Ladder followed by samples of the Ni²⁺-NTA column 1: Flow-through, 2: ATP-wash, 3-6: washing with 30 mM imidazole, the Strep-Tactin column 1: Flow-through, 2: washing, 3-7: elution fractions and the SEC 1: input after concentration to 500 μ l volume, 2-3: first peak fraction (red), 4-7: second peak fraction (green) and 8-9: third peak fraction. The arrow indicates the MW of the full-length Scap (green) construct. (C) Representative chromatogram of SEC using Superdex 200 (GE Healthcare) after DDM/CHS solubilization and purification via tandem affinity chromatography of 10 I HEK293 F-cell expressing porcine Scap. The sample elutes in mainly three peak fractions. A first fraction elutes at 9 ml representing the void volume (red), which is followed by a second peak at 11 ml in which the POI elutes (green) and a third fraction at around 13 ml that could not be analysed by SDS-PAGE (yellow).

All elution fractions of the chromatography steps were collected and analysed by semi-denaturing SDS-PAGE for its presence or contaminants of the POI. A first purification of the POI was performed by IMAC. A lot of DDM/CHS solubilized biomolecules could be separated by the application of the solubilize to Ni²⁺-NTA agarose beads as the flow-through fractions contains a high variety of different protein bands (see Figure 6.17, B left SDS-gel sample 1). Additional proteins with a MW of above 180 kDa, 70 kDa and 40 kDa could be isolated from the POI by an additional ATP-wash since they were verified as heat-shock proteins by MS/MS (see Figure 6.17, B left SDS-gel sample 2 and Table 8.2). Additional washing steps removed low-binding contaminants and the final elution fraction was used for a second affinity chromatography by Strep-Tactin (see Figure 6.17, B centre). The flow-through

sample features lots of contaminants that could be isolated from the POI. Contrary to that, the washing fraction showed only one single protein band at 160 kDa indicating the elution of the POI and no contaminants (see Figure 6.17, B centre sample 2). These results indicated a saturated state of the Strep-Tactin beads so that some POI could not immobilized at the stationary phase. Nevertheless, the elution fractions feature all a single protein band at 160 kDa which is in accordance with the calculated size of 145 kDa as membrane proteins can behave different compared to soluble proteins which are used in the protein standard (see sample M in Figure 6.17, B). The POI containing elution fractions were combined and concentrated up to 500 μ l volume and applied on a SEC column Superdex 200 in order to isolate potential aggregated proteins. A representative chromatogram was depicted in Figure 6.17 C that features the elution of four peak fractions which were furthermore analysed by semi-denaturing SDS-PAGE analysis (see Figure 6.17, B right gel). The first peak fraction elutes after 9 ml elution volume and could not properly separated from the second peak at 11 ml elution volume (see Figure 6.17, C). The third peak fraction was well separated and appeared after 13 ml as well as the last fraction at 19 ml.

The first peak fraction refers to the elution of higher oligomeric POI as the fraction featured one visible protein band at 160 kDa (see Figure 6.17, B and C, samples 2/3 red). In the peak fraction of around 11 ml, biomolecules with a MW of about 250 kDa elute, that would fit to the monomeric porcine Scap incorporated into the detergent mixture DDM/CHS as one protein band with a MW of 160 kDa could be observed via SDS-PAGE (see Figure 6.17 B and C, green highlighted samples). (141) The elution fractions at higher retention volumes from 13 to 19 ml did not contain any significant protein band and could not be analysed by SDS-PAGE (see Figure 6.17 B and C). Nevertheless, all contaminants could be well isolated from the POI as the main peak fraction, highlighted in green, was comprised of one protein with a MW of 160 kDa. Additional MS/MS analysis of the prominent protein band verified the elution of the porcine Scap (see Table 8.1 and Figure 8.7).

In order to validate the protein's homogeneity and analyse a first 3D reconstruction of the POI, the first (red) and second (green) peak fractions were collected and analysed by negative stain EM (see Figure 6.18). Under negative stain conditions the first peak fraction contained a mixture of protein particles ranging between 10 to 30 nm. Nevertheless, most of the particles contain a globular shaped particle of roughly 15 nm in its diameter, which furthermore could be visualized in the second peak fraction (see Figure 6.18, A and B). Here, the particles appeared more homogeneous compared to the sample of the first peak fraction indicating the separation of oligomeric states by SEC (see Figures 6.17 and 6.18). These protein particles were composed of a globular shaped base and a protruding density. These results could indicate that the protein is mainly incorporated into the detergent which is globularly shaped. The protruding density could show the long C-terminus including the WD40 repeat domain (see Figures 6.18 B-D and 6.17 A). For a first 3D reconstruction at low resolution a dataset of 150 micrographs was recorded at 1.84 $\text{\AA}/\text{pixel}$ and 11,976 particles were picked automatically using crYOLO within the software package Sphire. (122, 161)

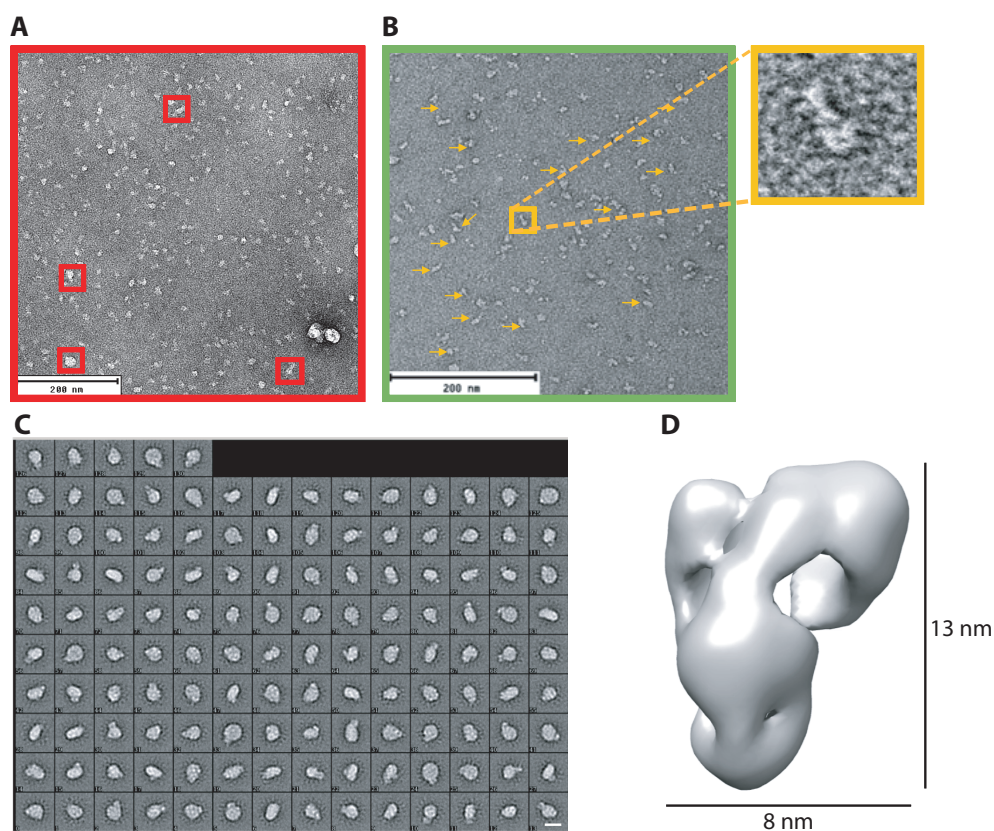


Figure 6.18: Negative stain EM analysis of the porcine Scap in DDM/CHS.

Representative negative stain micrographs of the first (red) (A) and second (green) (B) peak fractions from SEC of the porcine Scap solubilized in the detergent mixture DDM/CHS. Both peak fractions feature globular shaped protein particles mainly with a diameter of 15 nm. (A) The first peak fraction of the SEC additionally showed the presence of higher oligomers (highlighted in red squares) with a diameter of up to 30 nm. (B) The second peak fraction was composed of particles with an increased homogeneity compared to the first peak fraction. The main population featured a globular shape with an additional protruding density (highlighted in yellow). Scale bars: 200 nm. (C) 2D class averages of around 11,976 particles automatically picked from 150 micrographs that were taken with the second peak fraction using the software package SPHIRE. (121) Scale bar: 12 nm. (D) First 3D reconstruction of the porcine Scap construct. Scale bar height: 13 nm and width: 8 nm.

The 2D class averages of roughly 12,000 particles were mainly composed of the globular shaped basis and an additional density sticking out of this basis which could be observed on the negative stain micrographs (see Figure 6.18 B and C). For some class averages, the additional protruding density was missing either indicating an empty DDM/CHS 2D class or a bottom/top view of the particle incorporated into the detergent (see classes 99 and 113, Figure 6.18 C). Therefore, all classes depicted in Figure 6.18 C were used for a reconstruction of a first 3D model of the porcine Scap in DDM/CHS (see Figure 6.18 D). After the 3D refinement of this negative stain dataset, a final resolution of about 25 Å. The 3D model calculated by RVIPER of Scap featured a “L-shaped” protein with an 8 nm basis and 13 nm lengths. (124) Since the protein particles are composed of the detergent mixture DDM/CHS and the porcine Scap an assumption of the particle’s arrangement could be predicted at this low-resolution information. The mammalian Scap is supposed to be composed of eight transmembrane helices that comprise roughly 70% of the full-length protein construct including the transmembrane

connecting loop regions and the soluble rest of the protein construct (30%) which is about 45 kDa in MW (see Figure 6.17, A). (55) The protein particles base of 8 nm in length most was predicted to represent the detergent micelle including the transmembrane region of Scap and the soluble C-terminal domain is protruding outside the membrane region.

To conclude, the heterologous expression of the porcine Scap construct and its purification strategy was established and yielded in 0.05 mg POI per liter cell culture. The protein construct showed a correctly localized protein under fluorescence microscopy when over-expressed in HEK293 cells. By solubilization in a mild detergent mixture DDM/CHS, a tandem-affinity chromatography and a final SEC step, the porcine Scap was isolated from most contaminants. Furthermore, the sample quality and quantity were sufficient for a first low-resolution reconstruction of the mammalian full-length Scap by negative stain EM. Here, the protein particles were composed of a globular base and a protruding extra density with a length of 13 nm.

In order to achieve high-resolution information on the porcine Scap the protein sample was analysed by cryo-EM. Therefore, the protein sample needs to be vitrified in a thin layer of ice that was unfeasible with the buffer composition of DDM/CHS. An improvement of the buffer was furthermore validated by the protein's exchange from the detergent DDM/CHS to lipid nanodiscs, Amphipols, Digitonin or GDN.

6.2.3 Scap reconstituted into lipid nanodiscs

Lipid nanodiscs provide a near native environment for membrane proteins after their solubilization of the biological membrane. They are composed of lipids of choice which are embedded by MSPs. The scaffold proteins act like a belt and keep the lipid in a disc-shape when surrounded by aqueous solutions. The diameter of a nanodisc can be modified by the number of transmembrane helices of the MSP. I therefore used the first 3D reconstruction of the negative stain data set in order to adjust a lipid nanodisc of at least 8 nm diameter since the negative stain data showed a diameter of roughly 8 nm when surrounded by the detergent (see Figure 6.18 D). The scaffold proteins Msp1D1 (Addgene, 20061) and Msp1E3D1 (Addgene, 20066) are supposed to form lipid nanodiscs with a diameter of 9.1 nm and 12.9 nm, respectively and were used for further reconstitution experiments. Next to the scaffold protein, the nanodiscs are composed of lipids that need to mimic the proteins native environment as good as possible. Scap is known to be expressed and mainly localized in the ER-membrane that furthermore is composed of more than 50% phosphatidylcholine and around 5 mol% cholesterol. (55, 162) I therefore designed lipid nanodiscs with a lipid mixture of POPC and varying cholesterol or CHS concentrations between 5 to 10 mol% (see table 6.3).

Table 6.3: Composition of lipid nanodiscs for reconstitution of Scap.

POI	MSP	Lipid composition
Scap	Msp1D1	POPC/5 mol% cholesterol
		POPC/10 mol% cholesterol
		POPC/10 mol% CHS
	Msp1E3D1	POPC/5 mol% cholesterol
		POPC/10 mol% cholesterol
		POPC/10 mol% CHS

First, the porcine Scap was solubilized and purified as described in chapter 6.2.2. Subsequently, the isolated POI was mixed with the pre-equilibrated and assembled nanodiscs composed of POPC and cholesterol or CHS as described in chapter 5.3.12.

The reconstitution of the porcine Scap could be successfully performed with the Msp1D1 scaffold protein that is supposed to form 9 nm nanodiscs. Here, the lipid composition of POPC and 5 mol% featured most reconstituted POI after detergent removal via biobeads (see Figure 6.19).

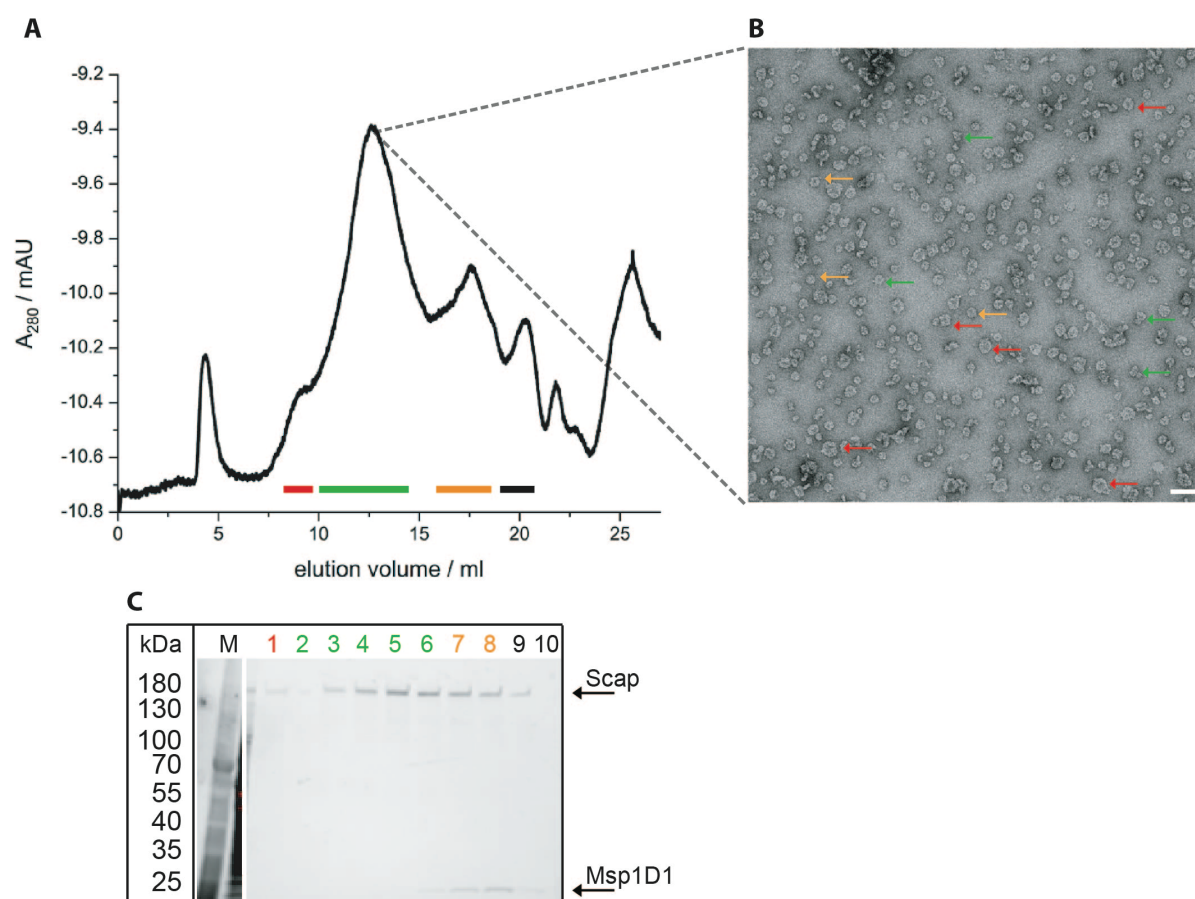


Figure 6.19: Reconstitution of porcine Scap from DDM/CHS into lipid nanodiscs Msp1D1.

The solubilized and purified porcine Scap construct in DDM/CHS was reconstituted into lipid nanodiscs composed of Msp1D1 and a lipid mixture of POPC and 5 mol% cholesterol. The POI/lipid/nanodisc reaction mixture was applied to biobeads for detergent removal for 16 h at 4 °C while gentle rotation. (A) The chromatogram of the final SEC features four main peaks that are highlighted in red, green, yellow and black. (B) The representative negative stain micrograph of the green peak fraction showing

some reconstituted protein particles which could be verified by semi-denaturing SDS-PAGE analysis. Scale bar: 200 nm (C) The SDS-gel in stain-free channel taken with the ChemiDoc MP imaging system (Bio-Rad) showing first M: PageRuler Prestained Protein Ladder followed by the peak fractions collected by SEC: sample 1: red peak fraction, samples 2-6: green peak fractions, sample 7 - 8: yellow peak fractions and samples 9 – 10: black peak fractions. The top arrow shows the MW of the porcine Scap and the bottom arrow indicates the MW of Msp1D1.

In Figure 6.19 the chromatogram as well as the analysis of collected SEC fractions by semi-denaturing SDS-PAGE and a representative negative stain micrograph illustrate the reconstitution of the DDM/CHS solubilized porcine Scap into lipid nanodiscs. After DDM/CHS removal via biobeads the POI/Msp1D1/lipid reaction mixture was applied onto a SEC column for the isolation of successfully reconstituted particles. The chromatogram features four main peak fractions. The first fraction eluting at 8 ml (highlighted in red) represents the void volume that could be not properly separated from the main peak fraction from 11 to 13 ml elution volume (highlighted in green) (see Figure 6.19, A). Two additional peak fractions could be separated from the main fraction via SEC by eluting at 17 (yellow) and 20 ml (black), respectively (see Figure 6.19, A). The analysis of the elution fractions by SDS-PAGE verified the presence of the POI and the Msp1D1 scaffold protein in late elution fractions of the second (green) peak at 14 ml retention volume (see Figure 6.19, A and C, sample 6). This fraction was furthermore analysed by negative stain EM (see Figure 6.19, B). The negative stain micrograph showed the presence of a mixture of small liposomes (black arrows), empty lipid nanodiscs (yellow arrows) and a few reconstituted proteins in nanodiscs (green arrows). Around 5% of all particles feature a protein incorporated into a lipid nanodisc indicating a low efficient incorporation of the POI into the Msp1D1 forming nanodiscs filled with POPC and 5 mol% cholesterol.

The Msp1E3D1 containing lipid nanodiscs were unfeasible for the reconstruction of the DDM/CHS solubilized porcine Scap since all tested conditions resulted in either degradation or aggregation of the POI (see Table 6.3 and Figure 8.8). Most probably the scaffold protein Msp1E3D1 forms a loose scaffold around the porcine Scap in order to properly reconstitute the POI into the nearly 13 nm nanodiscs. Since the first negative stain 3D model of the porcine Scap in DDM/CHS features a diameter of around 8 nm in its transmembrane region, a reconstitution into the 5 nm larger nanodisc is less frequent (see Figure 6.18 D). Furthermore, the POI tends to either degrade by a long-lasting detergent removal via dialysis (see Figure 8.8, A) or aggregate by a fast detergent removal via biobeads (see Figure 8.8, A). After various trials of reconstitution, the purified porcine Scap could be only reconstituted into the lipid nanodiscs composed of the Msp1D1 and 5 mol% cholesterol. Nevertheless, only 5% of the particles appeared incorporated into the nanodiscs indicating a less stabilized protein conformation compared to the detergent solubilized version. The fixed size of the assembled nanodiscs may hinder a proper reconstitution since the Msp1D1 scaffold proteins form 9 nm nanodiscs. Taken the lipids in the inner disc and the transmembrane helices into account, the Msp1D1 nanodisc forms too tight nanodiscs for the porcine Scap.

In order to improve the diameter of a lipid nanodisc, the self-assembled scaffold protein Saposin A could improve the reconstitution efficiency. The diameter of nanodiscs assembled by Saposins is

variable as they adopt their size based on the lipids and the reconstituted membrane protein. Previously published data showed the reconstitution of membrane proteins in Salipros, the “saposin-lipoprotein nanoparticle system”. (163, 164) Furthermore, the successful exchange of DDM/CHS to Digitonin or its analogue GDN featured suitable buffer conditions for the structure analysis of mammalian Patched1 as well as Dispatched-1 by cryo-EM. (5-10, 92, 160) These membrane proteins feature high similarity in the transmembrane region since Scap, Patched1 as well as Dispatched-1 contain the five transmembrane helices spanning SSD. (55) Therefore, I assumed a similar sample improvement by a detergent exchange to Digitonin or GDN.

6.2.4 Scap exchange into Digitonin or GDN

In order to improve the sample quality for structure analysis by cryo-EM, the porcine Scap construct was solubilized by DDM/CHS and exchanged into either Digitonin or GDN amphiphile while IMAC at the first affinity chromatography step. GDN is a known analogue of Digitonin with a similar chemical structure but features an increased stability and is less toxic compared to its analogue (see Figure 6.20, D and H). (165) A representative purification after detergent exchange into Digitonin as well as GDN is illustrated in Figure 6.20.

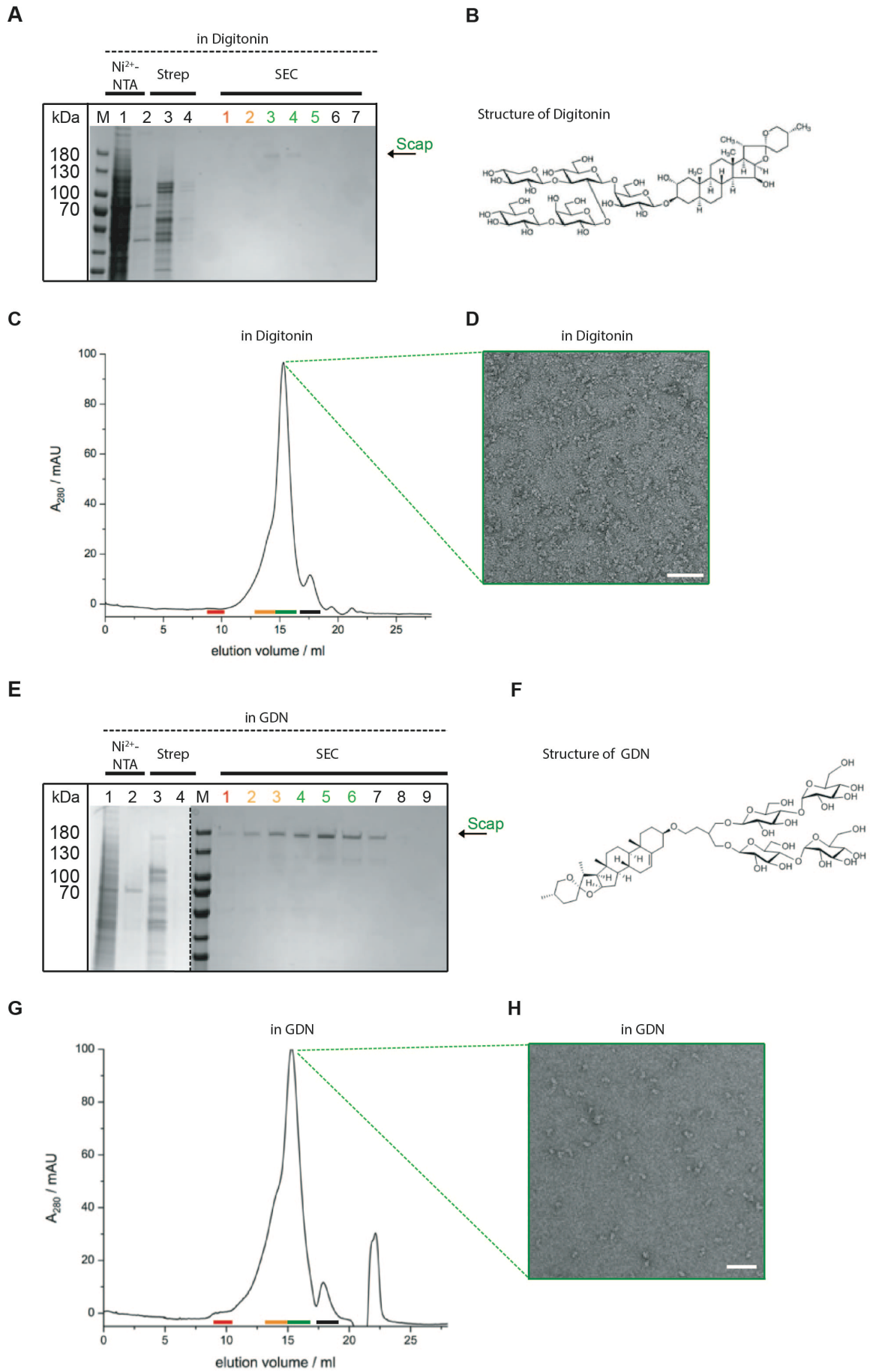


Figure 6.20: Representative purification of porcine Scap exchanged into Digitonin or GDN.

The DDM/CHS solubilized porcine homolog was exchanged to the amphiphiles Digitonin (A-D) or GDN (E-H) at tandem-affinity chromatography. (A) Coomassie stained SDS-gel of the affinity chromatography (1-4) and SEC elution fractions (1-7) after Digitonin exchange taken with the ChemiDoc MP imaging system (Bio-Rad) as followed M: PageRuler Prestained Protein Ladder, 1: Ni²⁺-NTA wash by 30 mM imidazole, 2: Ni²⁺-NTA ATP-wash, 3: Strep-Tactin flow-through, 4: Strep-Tactin wash and the peak fractions collected by SEC: 1: red peak fraction, 2: orange peak fraction, 3-5: green peak fraction and 6-7: black peak fraction. (B) The structural formular of the amphiphile Digitonin. (C) Representative chromatogram of porcine Scap after DDM/CHS exchange to Digitonin featuring one main peak fraction (highlighted in green) in which the purified POI elutes and (D) Representative negative stain micrograph of the POI in Digitonin. (E) Coomassie stained SDS-gels of the affinity chromatography (1-4) and SEC elution fractions (1-9) after exchange to GDN taken with the ChemiDoc MP imaging system (Bio-Rad) as followed, 1: Ni²⁺-NTA wash by 30 mM imidazole, 2: Ni²⁺-NTA ATP-wash, 3: Strep-Tactin flow-through, 4: Strep-Tactin wash, M: PageRuler Prestained Protein Ladder and the peak fractions collected by SEC: 1: red peak fraction, 2-3: orange peak fraction, 4-6: green peak fraction and 7-9: black peak fraction. (F) The structural formular of the amphiphile GDN. (G) Representative chromatogram of porcine Scap after DDM/CHS exchange to GDN featuring one main peak fraction (highlighted in green) in which the purified POI elutes and (H) representative negative stain micrograph of the POI in GDN.

In Figure 6.20 the SDS-gels of the semi-denaturing SDS-PAGE analysis from elution fractions of the tandem affinity chromatography as well as SEC at either Digitonin- or GDN-exchange show similar protein bands for each fraction, respectively (see Figure 6.20, A and E). Furthermore, the fractions of the flow-through and washings steps collected during tandem affinity chromatography show several protein bands that were not in the MW-range of the full-length POI (145 kDa) indicating the separation of contaminants as well as the immobilization of full-length porcine Scap. This could be verified by SEC (Figure 6.20, C and G) and the analysis of its elution fractions by semi-denaturing SDS-PAGE (Figure 6.20, A and E, right). One main peak fraction with an absorption at 280 nm of 100 mAU after 15 ml elution volume could be observed (see Figure 6.20, C and G highlighted in green). These data indicate the stabilization of monomeric state of the porcine Scap when exchanged from DDM/CHS to Digitonin or GDN. The quality of the main peak fraction (highlighted in green) was furthermore analysed by negative stain EM. The representative micrographs show a homogeneous particles of mainly globular protein particles (see Figure 6.20, D and H). The protein particles were composed of a globular base and the additional protruding density, that was furthermore observed when solubilized in DDM/CHS under negative stain conditions (see Figure 6.20 and 6.18). In comparison with the data of the DDM/CHS solubilized protein, the exchange into Digitonin and GDN stabilized the monomeric conformation of the porcine Scap. Additionally, the POI appeared stable in a buffer without the supplement Glycerol, which is advantageous for structure analysis by cryo-EM. (166) These results would confirm the assumption that the exchange of DDM/CHS to GDN or Digitonin stabilize membrane proteins containing the SSD since the structure of the mammalian Patched1, NPC1 as well as Dispatched were exchanged to Digitonin or GDN for cryo-EM sample preparation. (5-10, 92, 153, 154, 160)

Finally, the porcine Scap construct could be exchanged from DDM/CHS into Digitonin and GDN during tandem affinity chromatography. The POI featured improved stabilization when solubilized in the

amphiphilic molecules and furthermore show a homogeneous particle distribution under negative stain conditions (see Figure 6.20). Based on similar structural features of both amphiphile, there was no difference in purity, purification yield or stability of the porcine Scap validated. I therefore tested both purified samples for further analysis by cryo-EM. Dependent on the protein concentration and grid-type, I screened for a high variety of different plunging conditions in order to receive the protein particles vitrified in a thin and homogeneous ice layer that enables high resolution data acquisition of the POI (see Table 6.4). Additional parameters like variations of sample volume, incubation time on the grid, blotting force, drain time, humidity as well as temperature did not feature rising differences and were maintained as described in Chapter 5.5.2.1.

Table 6.4 Plunging conditions for porcine Scap in Digitonin or GDN.

Grid Type	Concentration / mg ml ⁻¹
QF 1.2/1.3 Cu200/300	5
	3.5
	2.5
	1.75
	1.25
C-flat 1.2/1.3 Cu50	3.5
	2.5
	1.75
	1.25
QF 1.2/1.3 Cu200 +2 nm carbon	3.5
	2.5
	1.75
	1.25
QF 2.1 Au (graphene)	0.8
	0.3
UltrAuFoil 1.2/1.3 Au300	5
	3.5
	2.5

The screening of the listed conditions was performed at the electron microscope Talos Arctica (FEI, Thermo Fisher Scientific). Preferably, the POI was attracted by the graphene layer surrounding the holes in which the protein molecules should enter when vitrified in ice. At rising protein concentrations from 2.5 mg/ml until 5 mg/ml the protein molecules distributed homogeneously in the ice layer. Furthermore, the ice thickness of the sample applied onto QF 1.2/1.3 Cu300 appeared most feasible for high resolution information on porcine Scap when solubilized in GDN since this allowed data acquisition without VPP in a defocus range of -1.2 - -4.0 μm .

An overview of information about the data acquisition and the processing were represented in Table 6.5.

Table 6.5: Settings and information on the data set of porcine Scap in GDN on a Cs corrected FEI Titan Krios electron microscope operated at 300 kV and equipped with X-FEG.

Settings	Scap in GDN
Detection camera	K3 Super-resolution (Gatan)
Voltage (kV)	300
Magnification	105k
Pixel size (Å/pixel)	0.9
Total electron dose ($e^-/\text{Å}^2$)	68
Number of frames/image	60
Exposure time (sec)	3
Estimated defocus range (μm)	-1.2 - -4.0
Micrographs	7964
Initial particle number	840,356
Final particle number	133,395
Symmetry imposed	C1
Resolution (FSC at 0.143, Å)	6.6

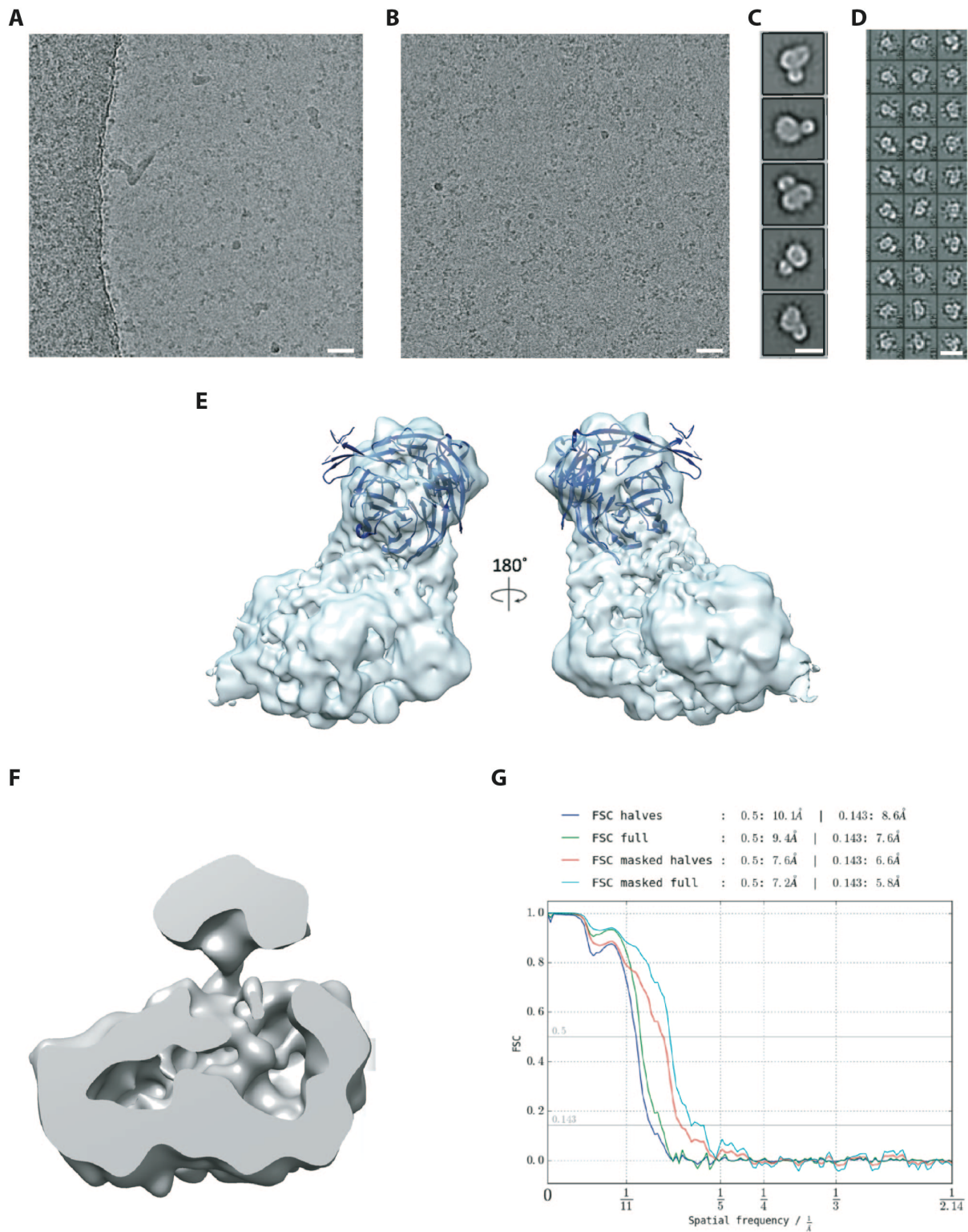


Figure 6.21: Cryo-EM of full-length porcine Scap in GDN.

Representative cryo-EM micrographs of Scap in GDN at $-2.5 \mu\text{m}$ defocus (A) at the carbon edge and (B) in the centre of a grid hole. Scale bars: 20 nm. (C) Representative 2D class averages generated using (C) RELION and (D) SPHIRE. Scale bars: 10 and 15 nm. (121, 133) (E) Refined 3D reconstruction of Scap including the fitted crystal structure of the fission yeast WD40 repeats (PDB: 4YHC). The fission yeast homolog features a prolonged WD40 repeat domain of 100-200 amino acids compared to the mammalian Scap homologs. (F) Refined 3D reconstruction of porcine Scap in the center of the electron density map that highlights the filled detergent micelle by potent transmembrane helices. (G) Fourier Shell Correlation (FSC) plot. The 0.143 criterion of the masked half-maps indicates a resolution of 6.6 Å.

In total 8000 micrographs were collected and 7000 micrographs were selected for data processing. In Figure 6.21 A and B selected micrographs showing the porcine Scap in GDN vitrified in a thin ice layer of around 20 – 30 nm were shown. The micrographs illustrate the presence of protein particles distributed homogeneously in the grid holes at a defocus of $-2.5\ \mu\text{m}$ without the usage of the VPP. By crYOLO 840 k particles were picked automatically within the software package SPHIRE. (161, 122) Here, 450 k particles were extracted and selected after a first 2D classification via ISAC. (125) After three iterations of 2D classification within RELION about 140 k particles were selected for a final 2D classification in ISAC within SPHIRE (see Figure 6.21, C and D). (122, 134) The 2D class averages created by ISAC as well as SPHIRE feature a globular shape including a protruding density. Furthermore, additional features inside the globular shaped base became visible indicating the transmembrane helices that could be incorporated into the detergent micelle GDN (see Figure 6.21, C and D). Compared to the 2D class averages from the negative stain data set, the 2D class averages of cryo-EM analysis feature higher resolution information since secondary structure elements like the transmembrane helices become visible here (see Figures 6.21, C and D and Figure 6.18, C). The refined 3D reconstruction of the porcine Scap was calculated with a final particle stack of around 100 k particles (see Figure 6.21, E and F). The diameter of the initial 3D reconstruction is 13.5 nm which is in accordance with the data obtained by the analysis of the negative stain data set (see Figure 6.18, C). These data imply a similar arrangement of the transmembrane helices within the DDM/CHS and GDN detergent micelles. In both reconstructions one porcine Scap would fit into the model which could be verified by SEC (see Figure 6.20, F). A close-up of the final 3D reconstruction, showed the presence of additional densities in the interior of the globular shaped base (see Figure 6.21, F). These findings could indicate the presence of transmembrane helices in this region which could be already observed at some 2D class averages (see Figure 6.21, C and D). The resolution estimated by the 0.143 criterion of the masked half-maps estimated an overall resolution of 6.6 Å. Nevertheless, separated helices could not be identified in the final 3D model indicating a resolution of about 10 Å. Additionally I was able to dock the crystal structure of the Scap WD40 repeats (PDB: 4YHC) into the protruding globular electron density (see Figure 6.21, E). In terms of shape and dimension the crystal structure fits into the globular density, only two loops do not fit properly due to the limited resolution and the extended (100-200 amino acids) WD40 domain in fission yeast Scap compared to mammalian Scap. (167) The hypothesis postulated by Gong *et al.*, that mammalian WD40 repeats might form a similar 3D structure like the fission yeast Scap homolog, could be confirmed. (63)

The high flexibility of the membrane protein inserted into the digitonin-micelle might be problematic for correct alignment and further processing in order to reach atomic resolution. For obtaining a reconstruction of the full-length Scap, it is necessary to stabilize the protruding soluble C-terminal domain of Scap in order to achieve a higher resolution of the initial model for building an atomic model. Furthermore, I assume a stabilization of Scap upon complex formation with its binding partner SREBP, since the complex is formed immediately after the localization in the ER-membrane of Scap and independent on the sterol concentration. In order to improve the flexibility and fix the protruding soluble domains compared to the detergent micelle, the analysis of the Scap/SREBP complex by cryo-

EM could be performed. X. Gong, *et al.* solved the structure of the fission yeast homolog complex of Scap/SREBP by cryo-EM in 2016. Here, they were able to stabilize the complex formation by using the GraFix approach that furthermore implies a weak interaction of complex formation that needed additional fixation reagents in order to keep Scap and SREBP in its bound-state. For structure analysis of the mammalian complex formed by Scap and SREBP the usage of GraFix might be inevitable in order to improve the complex stability as well. (66)

It has been previously reported that the addition of binding partners like Fab fragments specifically binding the membrane protein of interest, were successfully used to improve the flexibility of the membrane protein inserted into a lipid mimicking system, like detergents. Thereby, they were able to overcome the problematic flexibility and improved the alignment. (168, 169) Additional improvement could be expected by the addition of the BRIL domain that acts like a “fiducial marker” for structure analysis by single-particle cryo-EM. (170, 171) By the incorporation of a thermostabilized version of b₅₆₂RIL into the membrane proteins GPCR at intracellular loop regions or termini a less flexible protein could be generated which improved the resolution. Since strong binding Fab fragments are available against the BRIL domain, they could increase the MW by the addition of the Fab fragment and an additional generated nanobody by around 65 kDa. Furthermore, the BRIL domain is known to stabilize GPCRs that would improve the flexibility of the protruding soluble region. (170, 171)

In summary, I show the first electron density of a full-length mammalian Scap at sub-nanometer resolution. Here, I observed two entities that are supposed to be the transmembrane domain, containing the SSD, and the soluble large WD40 repeats. The crystal structure of the fission yeast WD40 repeats features an eight-bladed β -propeller that was docked into the density which is supposed to be the mammalian WD40 domain and showed a similar shape (see Figure 6.21, E). (63) I therefore hypothesize a similar 3D structure of the mammalian WD40 repeats.

6.2.1 Conclusion

For a complete understanding of the highly regulated cholesterol homeostasis on a molecular level, the 3D structure of one of the key regulators Scap is necessary. (55) Since the discovery of Scap by J. L. Goldstein and M. Brown in 1996, its structural features could be verified mainly based on the basis of hamster Scap. (55) Nevertheless, we are still lacking molecular information of a full-length Scap.

I therefore started with a Scap homolog screen in order to identify the most suitable Scap homolog for structure analysis. Based on its expression pattern, expression yield, and solubilization efficiency by a mild detergent DDM, I identified the porcine Scap to be the most suitable protein homolog (see Figures 6.12 and 6.13). I observed a preferred localization of the porcine Scap construct in the ER membrane and in additional membrane regions of the cells by fluorescence microscopy (see Figure 6.14). These results indicated the expression of a native conformation of Scap, because the membrane protein is supposed to be localized into the ER-membrane and at increasing sterol concentrations in the Golgi membrane. (5, 55) Furthermore, I could increase the expression level by improving the quality of the baculovirus. I successfully scaled up the expression up to 25 l of suspension HEK293 F cells without reduction of the protein yield. Additionally, I identified the detergent mixture composed of DDM and CHS as most efficient detergent for the solubilization of the porcine Scap construct (see Figure 6.16). DDM is a mild detergent and composed of a non-ionic head group and a long hydrocarbon chain (dodecyl rest). Based on its structural features, I assume that the micelle formed out of the detergent DDM/CHS has a similar diameter such as the biological membrane of the ER and the Golgi which are about 36 Å to 37 Å. (172) Furthermore, the detergent was used for the solubilization of other membrane proteins featuring an SSD, such as Patched1, NPC1 as well as Dispatched. (5-10, 92, 153, 154, 160) These data could imply the correct folded SSD of the porcine Scap after solubilization, which comprises five of eight transmembrane helices. (55)

I designed a porcine protein construct without the C-terminal eGFP tag in order to improve the purity (see Figure 6.17, A). I observed similar expression yields as well as solubilization efficiency of the new construct compared to the eGFP-modified version. By negative stain EM, I observed particles of different sizes indicating degradation and aggregation of the new protein construct after the final SEC. Based on these results, I assume that the detergent DDM/CHS was most efficient for solubilization but for the purification process. Therefore, I performed additional experiments in order to prevent degradation and aggregation of Scap. First, I used lipid nanodiscs in order to mimic the lipid bilayer of Scap. In order to build nearly native lipid nanodiscs, I used a mixture of phosphatidylcholine (POPC) and cholesterol because the biological membranes of the mammalian ER and Golgi are mainly composed of phosphatidylcholine and cholesterol. (162) About 5% of the total protein population could be incorporated into the scaffold protein Msp1D1 that assembles nanodiscs of 9 nm diameter (see Figure 6.19). Additional formations into lipid nanodiscs of an increased diameter of 12.1 nm either resulted in aggregation or degradation of the POI (see Figure 8.8). These results indicate that the

diameter of the transmembrane region of the porcine Scap is below 9 nm because a small protein population could be reconstituted. The fixed size of the scaffold protein might hinder a proper reconstitution into the lipid nanodiscs formed by Msp1D1. Recently published data show variable nanodiscs by a self-assembling scaffold protein, called Saposin A. By using Saposin A, the diameter of the lipid nanodiscs is variable as it adopts the size based on the lipids as well as the POI and furthermore, could improve the reconstitution efficiency into lipid nanodiscs. (163, 164) Strikingly results could be obtained by the exchange of DDM/CHS to Digitonin or its analogue GDN for structure analysis of Patched1 alone and in complex with its ligand as well as Dispatched-1. (6-11, 92, 160) Based on these data, I established the purification strategy of the new membrane protein construct after a detergent exchange of DDM/CHS into Digitonin or GDN. Here, the POI could be isolated from contaminants, by tandem affinity chromatography, including first an IMAC (by the N-terminal His₈-tag) and second a Strep-Tactin (by the double N-terminal StrepII-tags) chromatography and a final SEC (see Figure 6.20). Based on the chromatogram of the GDN or Digitonin solubilized porcine Scap, I validated the presence of mainly monomeric POI when solubilized in GDN or Digitonin (see Figure 6.20). During SEC, I observed an additional small shoulder before the main peak fraction, that might correspond to the presence of an higher oligomeric state of Scap (see Figure 6.20, C and G). It has been reported, when working with either the trans-membrane domain or the WD40 repeats of hamster and fission yeast Scap homologs, Scap forms higher-oligomers. The hamster homologs were solubilized in Fos-Choline 13 detergent which can cause conformational changes due to its ionic head group. The ionic detergent is known to disrupt intra- and intermolecular interactions that can lead to stabilization of artefacts. Furthermore, the oligomeric states of the fission yeast Scap homolog alone and in complex with SREBP needed to be stabilized by GraFix. (10, 63, 66, 173) Based on these data, I assume that the porcine Scap is mainly monomeric.

The purity of the GDN and Digitonin solubilized porcine Scap construct could be identified as well as a homogeneous particle population by negative stain-EM (see Figure 6.20, C and G). Since Digitonin as well as GDN share similar chemical structures and were both used for structure determination of mammalian Patched1 constructs, I tested different plunging conditions for both protein samples. (5-10) Nevertheless, a dataset of the GDN solubilized porcine Scap featured an increased homogeneous thin ice layer and a protein sample that could be visualized without using the VPP. GDN contains, in contrast to Digitonin, an improved stability and a 25-fold lower CMC of 18 μ M. (174-177) Due to the decreased CMC, I could use a decreased detergent concentration of 0.01% (v/v) instead of 0.1% (v/v). Furthermore, by the elimination of glycerol in the final SEC the signal-to-noise ratio could be improved, as organic molecules, like glycerol, reduce the signal-to-noise ratio of the images that is necessary for an accurate assignment of the angles during image processing. (166) Collecting a cryo-EM data set without using the VPP might contain increased high-resolution information of the full-length pig Scap. This presumably has an impact on the resolution of the collected data sets of porcine Scap taken with the VPP (see Table 5.10). This improved sample quality led to a final 3D reconstruction of the pig homolog at an overall resolution of 6.6 Å which was calculated by the 0.143 criterion of the masked half maps (see Figure 6.21, G). The reconstruction featured additional densities in the interior of the

detergent micelle which could imply the localization of some transmembrane helices in these areas. In addition to these findings, some 2D class averages showed the presence of secondary structure elements in the globular shaped base. Since the porcine Scap construct is supposed to be composed of eight transmembrane α -helices, these should be visible here. Usually, single alpha helices can be observed separated from starting at resolutions of about 8 Å, indicating a resolution of about 10 Å in reality.

Nevertheless the, the crystal structure of the soluble WD40 repeat domain of a fission yeast Scap homolog could be successfully docked into the density protruding from the globular shaped base (see Figure 6.21, E). The shape as well as the dimension of the crystal structure fits into the density that was supposed to build the soluble C-terminal domain. Since the protein construct used for the crystal structure determination feature an extended domain of about 100-200 amino acids compared to mammalian Scap, a small region of the crystal structure was projected out of the 3D density (see Figure 6.21, E). (63) I therefore assume that the mammalian WD40 repeat domain might form a similar 3D structure like the fission yeast homolog that was postulated by X. Gong, *et al.*. (63)

To conclude, the data presented in this work show the first 3D electron density map of a full-length mammalian Scap at around 10 Å resolution. The porcine Scap homolog appeared most stable at a homolog screen as well as the heterologous expression could be established via tandem-affinity chromatography and a final SEC. Furthermore, an increased purity and stability could be achieved by the elimination of the eGFP reporter tag and exchange from DDM/CHS to GDN and Digitonin. In comparison with the data from the recently published cryo-EM reconstructions of Patched1, NPC-1 as well as Dispatched that were successfully exchanged from DDM/CHS to GDN or Digitonin, all membrane proteins share similarity in their transmembrane region. These membrane proteins contain the SSD at helix 2 to 6 that is conserved at a variety of membrane proteins regulating cholesterol metabolism and and signaling. I therefore assume by using DDM/CHS for solubilization and GDN or Digitonin for further purification the porcine Scap protein construct could be stabilized in its native state. Furthermore, I predict a similar 3D reconstruction of the SSD that comprises 5 trans-membrane helices of 8 in mammalian Scap proteins. (5-10, 92, 153, 154, 160)

The cryo-EM density map could not be improved for further structure determination of full-length porcine Scap. The main obstacle is the flexibility of the protruding soluble region out of the globular shaped detergent base. For the elucidation of a high-resolution reconstruction a correct alignment is necessary. An improved protein for structure analysis by single-particle cryo-EM could be obtained by stabilization of the protein with its binding partners in the future. It has been previously reported that the addition of binding partners like Fab fragments specifically binding the membrane protein of interest, were successfully used to improve the flexibility of the membrane protein. (168, 169) Additional improvement could be expected by the addition of the BRIL domain that acts like a “fiducial marker” for structure analysis by single-particle cryo-EM. (170, 171) By the incorporation of a thermostabilized version of b₅₆₂RIL into the membrane proteins GPCR at intracellular loop regions or termini a less flexible protein could be generated which improved the resolution. (170, 171)

7 References

- (1) S. Tan, *et al.*: Membrane proteins and membrane proteomics. *Proteomics*, **2008**, 8: 3924-3932.
- (2) J. P. Overington, B. Al-Lazikani and A. L. Hopkins: How many drug targets are there? *Nature Reviews Drug Discovery*, **2006**, 5: 993-996.
- (3) M. Brown and J. Goldstein: Cholesterol feedback: from Schoenheimer's bottle to Scap's MELADL. *Journal of Lipid Research*, **2009**, 50: 15-27.
- (4) C. C. Hui and S. Angers: Gli proteins in development and disease. *Annu. Rev. Cell Dev. Biol.*, **2011**, 27: 513-537.
- (5) Y. Zhang, *et al.*: Structural Basis for cholesterol transport-like activity of the Hedgehog receptor Patched. *Cell*, **2018**, 175: 1352-1364.
- (6) X. Qi, *et al.*: Structures of human Patched and its complex with native palmitoylated sonic hedgehog. *Nature*, **2019**, 560: 128-132.
- (7) C. Qi, *et al.*: Structural basis of sterol recognition by human Hedgehog receptor PTCH1. **2019**, *Sci Adv*, 5: eaaw6490.
- (8) H. Qian, *et al.*: Inhibition of tetrameric Patched1 by Sonic Hedgehog through an asymmetric paradigm. *Nature Communications*, **2019**, 10.
- (9) X. Gong, *et al.*: Structural basis for the recognition of Sonic Hedgehog by human Patched1, *Science*, **2018**, 361: eaas8935.
- (10) X. Qi, *et al.*: Two Patched molecules engage distinct sides on Hedgehog yielding a signaling-competent complex. *Science*. **2018**, 362: eaas8843.
- (11) J. Goldstein and M. Brown: A century of cholesterol and coronaries: from plaques to genes to statins. *Cell*, **2015**, 161: 161-172.
- (12) J. Lombard: Once upon a time the cell membranes: 175 years of cell boundary research, *Biol Direct.*, **2014**, 932, 1-35.
- (13) P. L. Yeagle: The membranes of cells. *Academic press (Elsevier)*, **2016**, 3rd edition.
- (14) S. J. Singer, G. L. Nicholson: The fluid mosaic model of the structure of cell membranes, *Science*, **1972**, 175, 720-731.
- (15) G. C. K. Roberts: Encyclopedia of biophysics, Chapter: Lipid lateral diffusion. *Springer*, Edt. **2013**.
- (16) K. Simons and E. Ikonen: Functional rafts in cell membranes. *Nature*, **1997**, 387: 569-572.
- (17) K. Simons and G. van Meer: Lipid sorting in epithelial cells. *Biochemistry*, **1988**, 27: 6197-6202.
- (18) L. J. Pike: Rafts defined: a report on the Keystone Symposium on Lipid Rafts and Cell Function. *J Lipid Res.*, **2006**, 47: 1597-1598.
- (19) L. J. Pike: The challenge of lipid rafts. *J Lipid Res.*, **2009**, 50: 323-328.
- (20) L. Redondo-Morata, *et al.*: Influence of Cholesterol on the Phase Transition of Lipid Bilayers: A Temperature-Controlled Force Spectroscopy Study. *Langmuir*, **2012**, 28: 12851-12860.
- (21) H. Lodish, A. Berk, S. L. Zipursky, *et al.*: Molecular Cell Biology 4th edition. *W. H. Freeman and Company*, **2000**, 3.4, Figure 3-32.

- (22) S. Tan, *et al.*: Membrane proteins and membrane proteomics. *Proteomics*, **2008**, 8: 3924-3932.
- (23) S. Shao and R.S. Hedge: Membrane protein insertion at the endoplasmic reticulum. *Ann. Rev. Cell Dev. Biol.*, **2011**, 27: 25-56.
- (24) V. W. Rodwell, *et al.*: Harper's illustrated biochemistry, McGraw-Hill Education, New York, Chicago, San Francisco, Athens, London, Madrid, Mexico City, Milan, New Delhi, Singapore, Sydney, Toronto, **2015**.
- (25) S. Stefanovic and R. S. Hedge: Identification of a targeting factor for posttranslational membrane protein insertion into the ER. *Cell*, **2007**, 128: 1147-1159.
- (26) J. M. Dietschy, S. D. Turley & D. K. Spady: Role of liver in the maintenance of cholesterol and low density lipoprotein homeostasis in different animal species, including humans. *Journal of lipid research*. **1993**, 34: 1637-1659.
- (27) K. Nieweg, H. Schaller and F. W. Pfrieger: Marked differences in cholesterol synthesis between neurons and glial cells from postnatal rats. *J Neurochem.*, **2009**, 109: 125-134.
- (28) A. M. Heacock *et al.*: Cholesterol synthesis and nerve regeneration. *J. Neurochem.*, **1984**, 42: 98-993.
- (29) H. Hayashi *et al.*: Glial lipoproteins stimulate axon growth of central nervous system neurons in compartmented cultures. *J. Biol. Chem.*, **2004**, 279: 14009-14015.
- (30) Q.-W. Fan *et al.*: Cholesterol-dependent modulation of dendrite outgrowth and microtubule stability in cultured neurons. *J. Neurochem.*, **2002**, 80: 178-190.
- (31) C. Goritz *et al.*: Multiple mechanisms mediate cholesterol-induced synaptogenesis in a CNS neuron. *Mol. Cell. Neurosci.*, **2005**, 29: 190-201.
- (32) D. H. Mauch *et al.*: CNS synaptogenesis promoted by glia-derived cholesterol. *Science*, **2001**, 294: 1354-1357.
- (33) F. W. Pfrieger: Role of cholesterol in synapse formation and function. *Biochim. Biophys. Acta*, **2003**, 1610: 271-280.
- (34) R. K. Mann and P. A. Beachy: Novel lipid modifications of secreted protein signals. *Annu Rev Biochem.*, **2004**, 73: 891-923
- (35) J. A. Porter, K. E. Young and P. A. Beachy: Cholesterol modification of hedgehog signaling proteins in animal development. *Science*, **1996**, 274: 255-259.
- (36) X. Xiao, *et al.*: Cholesterol modification of smoothened is required for hedgehog signalling, *Molecular cell*, **2017**, 66: 154-162.
- (37) T. J. McIntosh: The effect of cholesterol on the structure of phosphatidylcholine bilayers. *Biochim Biophys Acta*, **1978**, 513: 43-58.
- (38) Dong, Tang, Chen: Dual functions of Insig proteins in cholesterol homeostasis. *Lipids in Health and Disease*, **2012**, 11: 173-180.
- (39) O. F. Kuzu, M. A. Noory and G. P. Robertson: The role of cholesterol in cancer. *Cancer Res.*, **2016**, 76: 2063-2070.
- (40) H. Ritchie and M. Roser: Causes of Death, *Ourworldindata.org.*, **2018 (updated 2019)**.
- (41) <https://www.amnesty.org/en/what-we-do/death-penalty/> (time span **2007-2016**).

- (42) A. Windaus: Über den Gehalt normaler und atheromatöser Aorten an Cholesterin und Cholesterinestern. *Zeitschrift Physiol. Chemie*, **1910**, 67: 174.
- (43) M. Rafieian-Kopaei, *et al.*: Atherosclerosis: Process, Indicators, Risk Factors and New hopes. *Int J Prev med.*, **2014**, 5: 927-946.
- (44) S. Sitaula and T. P. Burris: Cholesterol and other steroids., *Encyclopedia of Cell Biology, Elsevier Inc.*, **2016**, 173-179.
- (45) P. R. Clarke and D. G. Hardie: Regulation of HMG-CoA reductase: identification of the site phosphorylated by the AMP-activated protein kinase in vitro and in intact rat liver. *EMBO J.*, **1990**, 9: 2439-2446.
- (46) B.-L. Song, N. B. Javitt and R. A. DeBose-Boyd: Insig-mediated degradation of HMG CoA reductase stimulated by lanosterol, an intermediate in the synthesis of cholesterol. *Cell Metab.*, **2005**, 1: 179-189.
- (47) G. W. Go and A. Mani: Low-density lipoprotein receptor (LDLR) family orchestrates cholesterol homeostasis. *Yale Journal of Biology and Medicine*, **2012**, 85: 19-28.
- (48) E. Ikonen: Cellular cholesterol trafficking and compartmentalization. *Nature Reviews Molecular Cell Biology*, **2008**, 9: 125-138.
- (49) E. S. Istvan and J. Deisenhofer: Structural mechanism for statin inhibition of HMG-CoA reductase. *Science*, **2001**, 292: 1160-1164.
- (50) J. D. Cohen, E. A. Brinton and M. K. Ito, *et al.*: Understanding Statin Use in America and Gaps in Patient Education (USAGE): an internet-based survey of 10,138 current and former statin users. *J Clin Lipidol*, **2012**, 6: 208-215.
- (51) P. Cao, J. Hanai, P. Tanksale, *et al.*: Statin-induced muscle damage and atrogen-1 induction is the result of a geranylgeranylation defect. *FASEB J.*, **2009**, 23: 2844-2854.
- (52) P. D. Thompson, *et al.*: Statin-associated side effects. *J of the American college of cardiology*, **2016**, 67: 2395-2410.
- (53) X. Hua, A. Nohturfft, J. L. Goldstein, M. S. Brown: Sterol resistance in CHO cells traced to point mutation in SREBP cleavage activating protein (SCAP). *Cell*, **1996**, 87: 415-426.
- (54) P. Kuwabara, M. Labouesse: The sterol-sensing domain: multiple families, a unique role? *Trends Genet.*, **2002**, 18: 193-201.
- (55) M. S. Brown, A. Radhakrishnan and J. L. Goldstein: Retrospective on cholesterol homeostasis: the central role of Scap. *Annu. Rev. Biochem*, **2018**, 87: 783-807.
- (56) A. Nohturfft, M. S. Brown, J. L. Goldstein: Topology of SREBP cleavage-activating protein, a polytopic membrane protein with a sterol-sensing domain. *J Biol Chem.*, **1998**, 273: 17243-17250.
- (57) J. Sakai, *et al.*: Identification of complexes between the COOH-terminal domains of sterol regulatory element-binding proteins (SREBPs) and SREBP cleavage-activating protein (SCAP). *J Biol Chem.*, **1997**, 272: 20213-20221.
- (58) J. Sakai, *et al.*: Cleavage of sterol regulatory element binding proteins (SREBPs) at site-1 requires interaction with SREBP cleavage-activating protein. Evidence from in vivo competition studies. *J Biol Chem.*, **1998**, 273: 5785-5793.

- (59) A. Nohturfft, *et al.*: Regulated step in cholesterol feedback localized to budding of SCAP from ER membranes. *Cell*, **2000**, 102: 315-253.
- (60) A. J. Brown, *et al.*: Cholesterol addition to ER membranes alters conformation of SCAP, the SREBP escort protein that regulates cholesterol metabolism. *Mol Cell*, **2002**, 10: 237-245.
- (61) T. Yang, *et al.*: Crucial step in cholesterol homeostasis: sterols promote binding of SCAP to INSIG-1, a membrane protein that facilitates retention of SREBPs in ER. *Cell*, **2002**, 110: 489-500.
- (62) J. D. Feramisco, *et al.*: Intramembrane aspartic acid in SCAP protein governs cholesterol-induced conformational change. *PNAS*, **2005**, 102: 3242-3247.
- (63) X. Gong, *et al.*: Structure of the WD40 domain of Scap from fission yeast reveals the molecular basis for SREBP recognition. *Cell Res.*, **2015**, 25: 401-411.
- (64) A. L. Hughes, B. L. Todd, P. J. Espenshade: SREBP pathway responds to sterols and functions as an oxygen sensor in fission yeast. *Cell*, **2005**, 120: 831-842.
- (65) T. F. Osborne, P. J. Espenshade: Evolutionary conservation and adaptation in the mechanism that regulates SREBP action: what a long, strange tRIP it's been. *Genes Dev.*, **2009**, 23: 2578-2591.
- (66) X. Gong, *et al.*: Complex structure of the fission yeast SREBP-SCAP binding domains reveals an oligomeric organization, *Cell Res.*, **2016**, 26: 1197-1211.
- (67) S. H. Lee, *et al.*: The cellular function of Scap in metabolic signalling. *Experimental & Molecular Medicine*, 2020, 52: 724-729.
- (68) Kamisuki, S. *et al.*: A small molecule that blocks fat synthesis by inhibiting the activation of SREBP. *Chem. Biol.*, **2009**, 16: 882-892.
- (69) J. J. Tang, *et al.*: Inhibition of SREBP by a small molecule, betulin, improves hyperlipidemia and insulin resistance and reduces atherosclerotic plaques. *Cell Metab.*, **2011**, 13: 44-56.
- (70) X. Cheng, J. Li and D. Guo: SCAP/SREBPs are central players in lipid metabolism and novel metabolic targets in cancer therapy. *Curr. Top. Med. Chem.*, **2018**, 18: 484-493.
- (71) M. Liu, *et al.*: Xanthohumol, a prenylated chalcone from beer hops, acts as an alpha-glucosidase inhibitor in vitro. *J. Agric. Food Chem.*, **2014**, 62: 5548-5554.
- (72) S. Miyata, *et al.*: Xanthohumol improves diet-induced obesity and fatty liver by suppressing sterol regulatory element-binding protein (SREBP) Activation. *J. Biol. Chem.*, **2015**, 290: 20565–20579.
- (73) M. Brown and J. Goldstein: Cholesterol feedback: from Schoenheimer's bottle to Scap's MELADL. *Journal of Lipid Research*, **2009**, 50: 15-27.
- (74) J. Briscoe and P. Théron: The mechanism of Hedgehog signalling and its roles in development and disease. *Nature Reviews*, **2013**, 14: 416-429.
- (75) F. Wu, *et al.*: Hedgehog Signaling: From Basic Biology to Cancer Therapy. *Cell Chem Biol.*, **2017**, 24: 252-280.
- (76) J. Taipale and P. A. Beachy: The Hedgehog and Wnt signalling pathways in cancer. *Nature*, **2001**, 411: 349-354.
- (77) P. A. Beachy and S.S. Karhadkar, D. M. Berman: Tissue repair and stem cell renewal in carcinogenesis. *Nature*, **2004**, 432: 324-331.

- (78) L. L. Rubin and F. J. de Sauvage: Targeting the Hedgehog pathway in cancer. *Nat. Rev. Drug Discov.*, **2006**, 5: 1026-1033.
- (79) J. Xie, *et al.*: Activating Smoothed mutations in sporadic basal-cell carcinoma. *Nature*, **1998**, 391: 90-92.
- (80) C. Nüsslein-Volhard and E. Wieschaus: Mutations affecting segment number and polarity in *Drosophila*, *Nature*, **1980**, 287, 795.
- (81) Y. Echelard, *et al.*: Sonic-Hedgehog, A Member Of A Family Of Putative Signaling Molecules, Is Implicated In The Regulation Of Cns Polarity. *Cell*, **1993**, 75: 1417-1430.
- (82) G. R. van den Brink, *et al.*: Indian Hedgehog is an antagonist of Wnt signaling in colonic epithelial cell differentiation. *Nat Genet.*, **2004**, 36: 277-282.
- (83) L. J. Spicer, *et al.*: pathway and function in the mammalian ovary: a novel role for hedgehog proteins in stimulating proliferation and steroidogenesis of theca cells. *Reproduction*. **2009**, 138: 329-339.
- (84) A. Szczepny, *et al.*: Expression of hedgehog signalling components in adult mouse testis. *Dev Dyn.*, **2006**, 235: 3063-3070.
- (85) R. Mirsky, *et al.*: Schwann cell-derived desert hedgehog signals nerve sheath formation. *Ann N Y Acad Sci.*, **1999**, 883: 196-202.
- (86) E. Parmantier, *et al.*: Schwann cell-derived Desert hedgehog controls the development of peripheral nerve sheaths. *Neuron.*, **1999**, 23: 713-724.
- (87) A. Vortkamp, *et al.*: Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein. *Science*, **1996**, 273: 613-622.
- (88) J. Jiang and C.-C. Hui: Hedgehog signaling in development and cancer, *Developmental cell*, **2008**, 15, 801-812.
- (89) I. Guerrero and C. Chiang: A conserved mechanism of Hedgehog gradient formation by lipid modifications. *TRENDS in Cell Biology*, **2007**, 17: 1-5.
- (90) A. Pan, *et al.*: A review of hedgehog signaling in cranial bone development. *Front. Physiol.*, **2013**, 4, 61.
- (91) H. Tukachinsky, *et al.*: Dispatched and Scube mediate the efficient secretion of the cholesterol-modified hedgehog ligand, *Cell reports*, **2012**, 2: 308-320.
- (92) H. Chen, Y. Liu and X. Li: Structure of human Dispatched-1 provides insights into Hedgehog ligand biogenesis., *Life Sci Alliance*, **2020**, 3, 8: e202000776.
- (93) D. Venkatesh: Primary cilia. *J oral Maxillofac Pathol.*, **2017**, 21: 8-10.
- (94) L. Milenkovic, *et al.*: Single-molecule imaging of Hedgehog pathway protein Smoothed in primary cilia reveals binding events regulated by Patched1, *PNAS*, **2015**, 112, 8320-8325.
- (95) <http://www.uniprot.org/>, August **2020**.
- (96) F. Wu, *et al.*: Hedgehog Signaling: From Basic Biology to Cancer Therapy, *Cell chemical biology*, **2017**, 24: 252-280.
- (97) H. Nikaido and Y. Takatsuka: Mechanisms of RND multidrug efflux pumps. *Biochim. Biophys. Acta*, **2009**, 1794: 769-781.

- (98) M. Bidet, *et al.*: The hedgehog receptor patched is involved in cholesterol transport. *PLoS One*, **2011**, 6: e23834.
- (99) B. R. Myers, *et al.*: Rapid, direct activity assays for Smoothed reveal Hedgehog pathway regulation by membrane cholesterol and extracellular sodium. *PNAS*, **2017**, 114: E11141-E11150.
- (100) <https://www.rcsb.org/stats/growth/growth-released-structures> (14.09.2020 10:15 a.m.).
- (101) <https://blanco.biomol.uci.edu/mpstruc/> (14.09.2020 10:15 a.m.).
- (102) A. Goehring, *et al.*: Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat. Protoc.*, **2014**, 9:2574-2585.
- (103) I. G. Denisov, *et al.*: Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. *J. Am. Chem. Soc.*, **2004**, 126: 3477-3487.
- (104) I. G. Denisov, *et al.*: Cooperativity in cytochrome P450 3A4: linkages in substrate binding, spin state, uncoupling, and product formation. *J. Biol. Chem.*, **2007**, 282:7066-7076.
- (105) K. Mullis, *et al.*: Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Biotechnology*, **1992**, 24: 17-27.
- (106) C. Mülhardt: Der Experimentator, Molekularbiologie/Genomics. *Spektrum Akademischer Verlag Heidelberg*, **2009**, 6. Ep.: 160.
- (107) J. F. Miller, W. J. Dower and L. S. Tompkins: High-voltage electroporation of bacteria: genetic transformation of *Campylobacter jejuni* with plasmid DNA. *Proc. Natl. Acad. Sci. U S A*, **1988**, 85: 856.
- (108) W. J. Dower, J. F. Miller, C. W. Ragsdale: High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.*, **1988**, 16: 6127.
- (109) S. Wu, G. J. Letchworth: High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol. *Biotechniques*, **2004**, 36: 152-154.
- (110) M. Koenen, *et al.*: Immunoenzymatic detection of expressed gene fragments cloned in the lac Z gene of *E. coli*. *EMBO J.*, **1982**, 1: 509-512.
- (111) F. Sanger, *et al.*: DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U S A*, **1977**, 74: 5463.
- (112) A. Baer and K. Kehn-Hall: Viral concentration determination through plaque assays: using traditional and novel overlay systems. *J. Vis. Exp.* **2014**, 4: e52065.
- (113) U. K. Laemmli: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **1970**, 227: 680-685.
- (114) W. N. Burnette: „Western Blotting“: electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Analyt. Biochem.*, **1981**, 112: 195-203.
- (115) M. le Maire, P. Champeil and P. J. Booth: Interaction of membrane proteins and lipids with solubilizing detergents. *Biochem. Biophys. Acta*, **2000**, 1508: 86-111.
- (116) M. Almgren: Mixed micelles and other structures in the solubilization of bilayer lipid membranes by surfactants. *Biochem Biophys. Acta*, **2000**, 1508: 146-163.

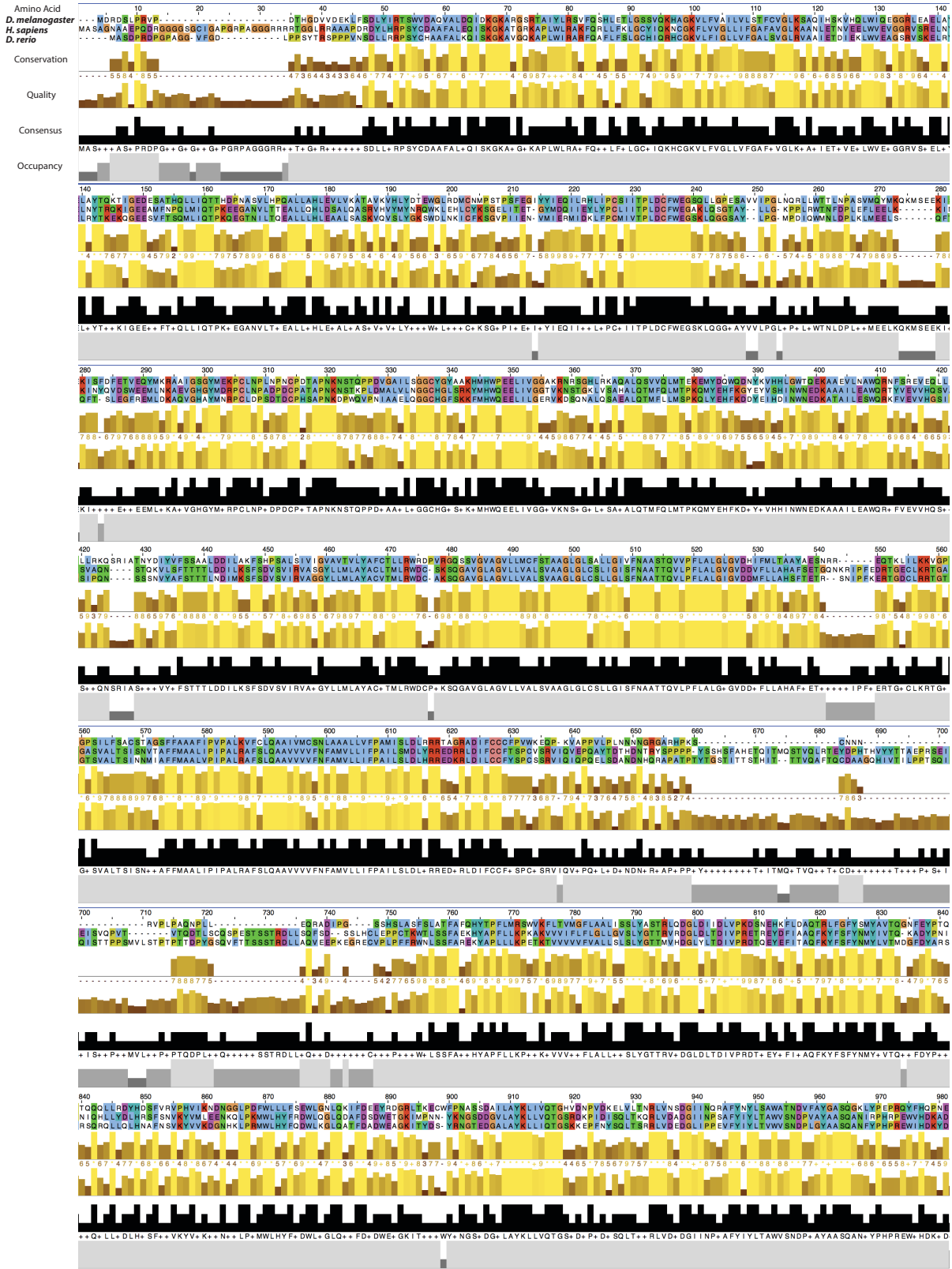
- (117) M. Urh, D. Simpson and K. Zhao: Affinity chromatography: general methods, *Methods Enzymol.*, **2009**, 463: 417-438.
- (118) J. Wen, T. Arakawa und J. S. Philo: Size-exclusion chromatography with on-line light-scattering, absorbance and refractive index detectors for studying proteins and their interactions. *Analyt. Biochem.*, **1996**, 240: 155-166.
- (119) J. O. Becker and A. N. Hoofnagle: Replacing immunoassays with tryptic digestion-peptide immunoaffinity enrichment and LC-MS/MS. *Bioanalysis*, **2012**, 4: 281-290.
- (120) M. Ohi, *et al.*: Negative Staining and image classification - powerful tools in modern electron microscopy, *Biol Proced Online*, **2004**, 6: 23-34.
- (121) D. S. Booth, *et al.*: Visualizing proteins and macromolecular complexes by negative stain EM: from grid preparation to image acquisition. *J. Vis. Exp.*, **2011**, 22: 3227.
- (122) T. Moriya, *et al.*: High-resolution single Particle Analysis from Electron Cryo-microscopy Images using SPHIRE, *J Vis Exp.*, **2017**, 123.
- (123) J. Redmon, *et al.*: You Only Look Once: Unified, real-time object detection. *IEEE Conf. Comput. Vis. Pattern Recognition (CVPR)*. **2016**, 779-788.
- (124) G. Tang, *et al.*: EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol.*, **2007**, 157: 38-46.
- (125) Z. Yang, *et al.*: Iterative Stable Alignment and Clustering of 2D Transmission Electron Microscope Images. *Structure*, **2012**, 20: 237-247.
- (126) P. A. Penczek, R. A. Grassucci, J. Frank: The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy*, **1994**, 53: 251-270.
- (127) E. F. Petterson *et al.*: UCSF Chimera - A visualization system for exploratory research and analysis, *J Comput Chem.*, **2004**, 25: 1605-1612.
- (128) K. Naydenova, *et al.*: Multifunctional graphene supports for electron cryomicroscopy. *PNAS*, **2019**, 116: 11718-11724.
- (129) R. Danev and W. Baumeister: Cryo-EM single particle analysis with the volta phase plate. *eLife*, 2016, 5: e13046.
- (130) R. Danev, D. Tegunov, W. Baumeister: Using the volta phase plate with defocus for cryo-EM single particle analysis, *eLife*, **2017**, 6: e23006.
- (131) J. Frank, A. Verschoor and M. Boublik: Computer averaging of electron micrographs of 40S ribosomal subunits, *Science*, **1981**, 214: 1353-1355.
- (132) M. van Heel and J. Frank: Use of multivariate statistics in analyzing the images of biological macromolecules. *Ultramicroscopy*, **1981**, 6: 187-194.
- (133) S. Zheng, *et al.*: MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods*, **2017**, 14: 331-332
- (134) A. Rohou and N. Grigorieff: CTFIND4: Fast and accurate defocus estimation from electron micrographs, *J. Structural Biology*, **2015**, 192: 216-221.

- (135) S. H. Scheres: RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, **2012a**, 180: 519-530.
- (136) S. H. Scheres and S. Chen: Prevention of overfitting in cryo-EM structure determination. *Nature Methods*, **2012b**, 9: 853-854.
- (137) O. Joubert, *et al.*: Functional studies of membrane-bound and purified human Hedgehog receptor Patched expressed in yeast. *Biochimica et Biophysica Acta*, **2009**, 1788: 1813-1821.
- (138) T. E. Cleveland IV, J. M. McCabe and D. J. Leahy: Detergent-solubilized Patched purified from Sf9 cells fails to interact strongly with cognate Hedgehog or Ihog homologs. *Protein Expr. Purif.*, **2014**, 104: 92-102.
- (139) A. Radhakrishnan, *et al.*: Direct Binding of Cholesterol to the Purified Membrane Region of SCAP: Mechanism for a Sterol-Sensing Domain. *Molecular Cell*, **2004**, 15: 259-268.
- (140) P. J. Espenshade, *et al.*: Sterols block binding of COPII proteins to SCAP, thereby controlling SCAP sorting in ER. *PNAS*, **2002**, 99: 11694-11699.
- (141) P. Strop and A. T. Brunger: Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci.*, **2005**, 14: 2207-2211.
- (142) X. Lu, S. Liu and T. B. Kornberg: The C-terminal tail of the Hedgehog receptor Patched regulates both localization and turnover. *Genes & Dev.*, **2006**, 20: 2539-2551.
- (143) J. Kim, *et al.*: The role of ciliary trafficking in Hedgehog receptor signaling. *Science signaling*, **2015**, 8: ra55.
- (144) A. Krogh, *et al.*: Predicting transmembrane protein topology with hidden Markov model: application to complete genomes. *J Mol Biol.*, **2001**, 305: 567-580.
- (145) <http://www.cbs.dtu.dk/services/TMHMM/> (TMHMM Server v. 2.0)
- (146) J. Kotschy: Heterologous purification of the human membrane protein Patched 1 for analysis of structure function relations. **2018**, RFWU Bonn, Master Thesis.
- (147) S. Teglund and R. Toftgård: Hedgehog beyond medulloblastoma and basal cell carcinoma. *Biochim Biophys. Acta*, **2010**, 1805: 181-208.
- (148) M. Gerling, *et al.*: Stromal Hedgehog signaling is downregulated in colon cancer and its restoration restrains tumor growth. *Nat. Commun.*, **2016**, 7, 12321.
- (149) A. Fleet, *et al.*: Activities of the Cytoplasmic Domains of Patched-1 Modulate but are not essential for the Regulation of canonical Hedgehog Signaling. *J. Biol. Chem.*, **2016**, 291: 17557-17568.
- (150) T. VanAken, *et al.*: Alkyl glycoside detergents: Synthesis and applications to the study of membrane proteins. *Methods Enzymol.*, **1986**, 125: 27-35.
- (151) T. Gewering, *et al.*: Know your detergents: A case study on detergent background in negative stain electron microscopy. *Journal of Struct. Biol.*, **2018**, 203: 242-246.
- (152) C. A. Scarff, *et al.*: Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems. *J. Vis. Exp.*, **2018**, 132: 57199.
- (153) X. Li, *et al.*: Structure of human Niemann-Pick C1 protein. *PNAS*, **2016**, 113: 8212-8217.

- (154) X. Li, *et al.*: 3.3 Å structure of human Niemann-Pick C1 protein reveals insights into the function of the C-terminal luminal domain in cholesterol transport. *PNAS*, **2017**, 114: 9116-9121.
- (155) N. Deneff, *et al.*: Hedgehog induces opposite changes in turnover and subcellular localization of patched and smoothed. *Cell*, **2000**, 102: 521-531.
- (156) J. Taipale, *et al.*: Patched acts catalytically to suppress the activity of Smoothed. *Nature*, **2002**, 418: 892-897.
- (157) C. Kowatsch, *et al.*: Structures of vertebrate Patched and Smoothed reveal intimate links between cholesterol and Hedgehog signaling. *Curr. Opin. Struct. Biol.*, **2019**, 57: 204-214.
- (158) M. K. Cooper, *et al.*: A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. *Nat. Genet.*, **2003**, 33: 508-513.
- (159) D. Prumbaum: Strukturanalyse und funktionelle Bindungsstudie von SCAP. Dissertation, TU Dortmund, Max-Planck-Institut für molekulare Physiologie, Abteilung für Strukturbiochemie, 2015.
- (160) F. Cannac, *et al.*: Cryo-EM structure of the Hedgehog release protein Dispatched. *Science Advances*, **2020**, 6: eaay7928.
- (161) T. Wagner, *et al.*: SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Comm. Biology*, **2019**, 2: 2018.
- (162) D. Casares, *et al.*: Membrane Lipid Composition: Effect on Membrane and Organelle Structure, Function and Compartmentalization and Therapeutic Avenues. *Int. J. Mol. Sci.*, **2019**, 20: 2167.
- (163) J. Frauenfeld, *et al.*: A saposin-lipoprotein nanoparticle system for membrane proteins. *Nature Methods*, **2016**, 13: 345-351.
- (164) A. Flayhan, *et al.*: Saposin Lipid Nanoparticles: A High Versatile and Modular Tool for Membrane Protein Research. *Structure*, **2018**, 26: 345-355.e5.
- (165) P. S. Chae, *et al.*: A New Class of Amphiphiles Bearing Rigid Hydrophobic Groups for Solubilization and Stabilization of Membrane Proteins. *Chemistry*, **2012**, 18: 9485-9490.
- (166) R. F. Thompson, *et al.*: An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*, **2016**, 100: 3-15.
- (167) C. Stirnimann, *et al.*: WD40 proteins propel cellular networks. *Trends Biochem Sci*, **2010**, 35: 565-574.
- (168) J. D. Brunner and S. Schenck: Production and Application of Nanobodies for Membrane Protein Structural Biology. *Methods Mol. Biol.*, **2020**, 2127: 167-184.
- (169) H. Kaur, *et al.*: Identification of conformation-selective nanobodies against the membrane protein insertase BamA by an integrated structural biology approach. *J. Biomol NMR*, **2019**, 73: 375-384.
- (170) N. Tsutsumi, *et al.*: Structure of human Frizzled5 by fiducial-assisted cryo-EM supports a heterodimeric mechanism of canonical Wnt signaling. *eLife*, 2020, 9:e58464.
- (171) S. Mukherjee, *et al.*: Synthetic antibodies against BRIL as universal fiducial marks for single-particle cryo-EM structure determination of membrane proteins. *Nature Comm.*, **2020**, 11:1598.
- (172) K. Mitra, *et al.*: Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proc. Natl. Acad. Sci. U S A*, **2004**, 101: 4083-4088.

- (173) A. Radhakrishnan, *et al.*: Direct binding of cholesterol to the purified membrane region of SCAP: mechanism for a sterol-sensing domain. *Mol Cell*, **2004**, 15: 259-268.
- (174) M. Baker, *et al.*: Single-particle cryo-EM of the ryanodine receptor channel in an aqueous environment. *Eur. J. Transl. Myol*, **2015**, 25: 35–48.
- (175) P.D. Abeyrathne, *et al.*: Analysis of 2-D Crystals of Membrane Proteins by Electron Microscopy. In *Comprehensive Biophysics*; Elsevier: Amsterdam, The Netherlands, **2012**; Chapter 119–1.15, pp. 277–310.
- (176) A. Stetsenko, A. Guskov: An Overview of the Top Ten Detergents used for membrane protein crystallization. *Crystals*, **2017**, (7)197.
- (177) P. Chae, *et al.*: A new class of amphiphiles bearing rigid hydrophobic groups for solubilisation and stabilization of membrane proteins. *Chemistry*, **2012**, 18: 9485-9490.
- (178) F. Sievers, *et al.*: Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **2011**, 7: 539.
- (179) A. M. Waterhouse, *et al.*: Jalview version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **2009**, 25: 1189-1191.

8 Supplementary Information



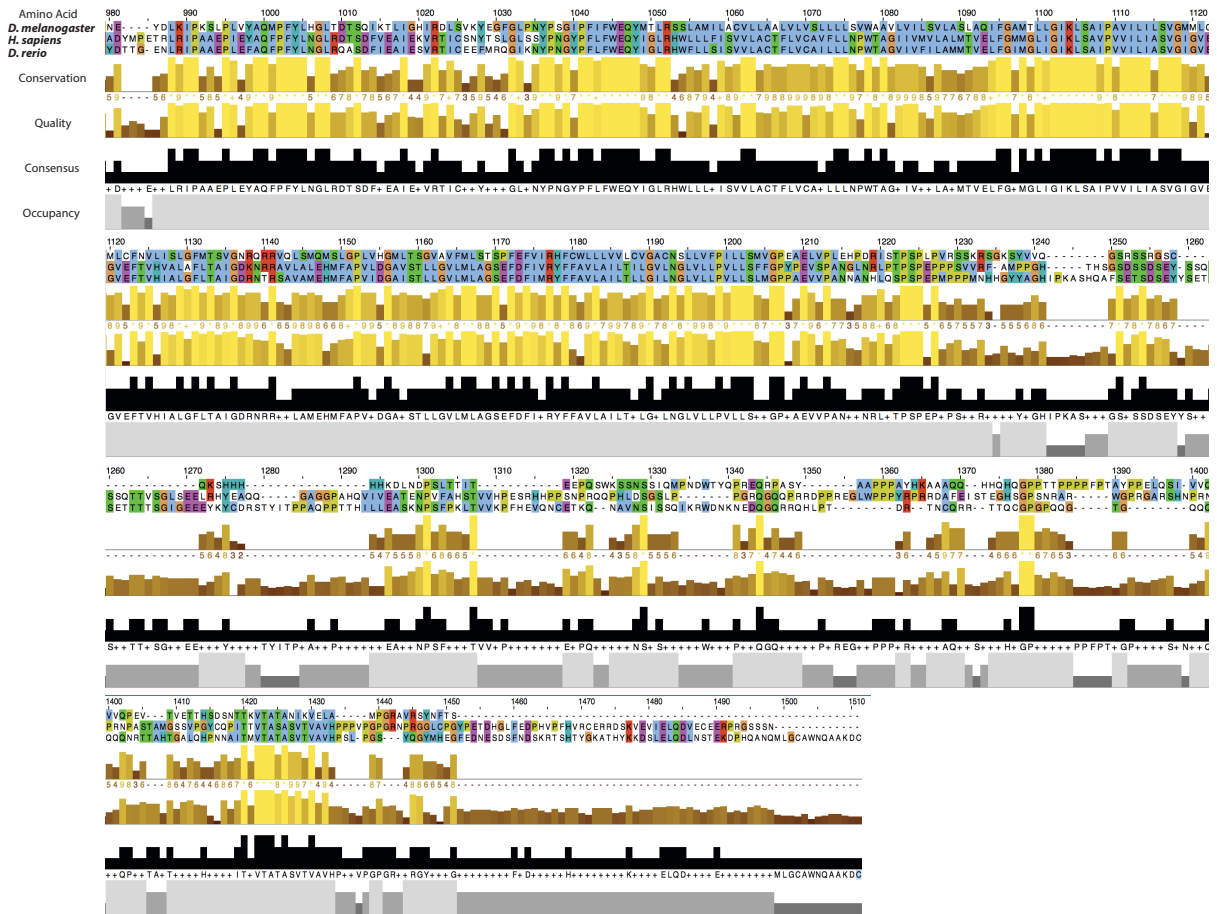
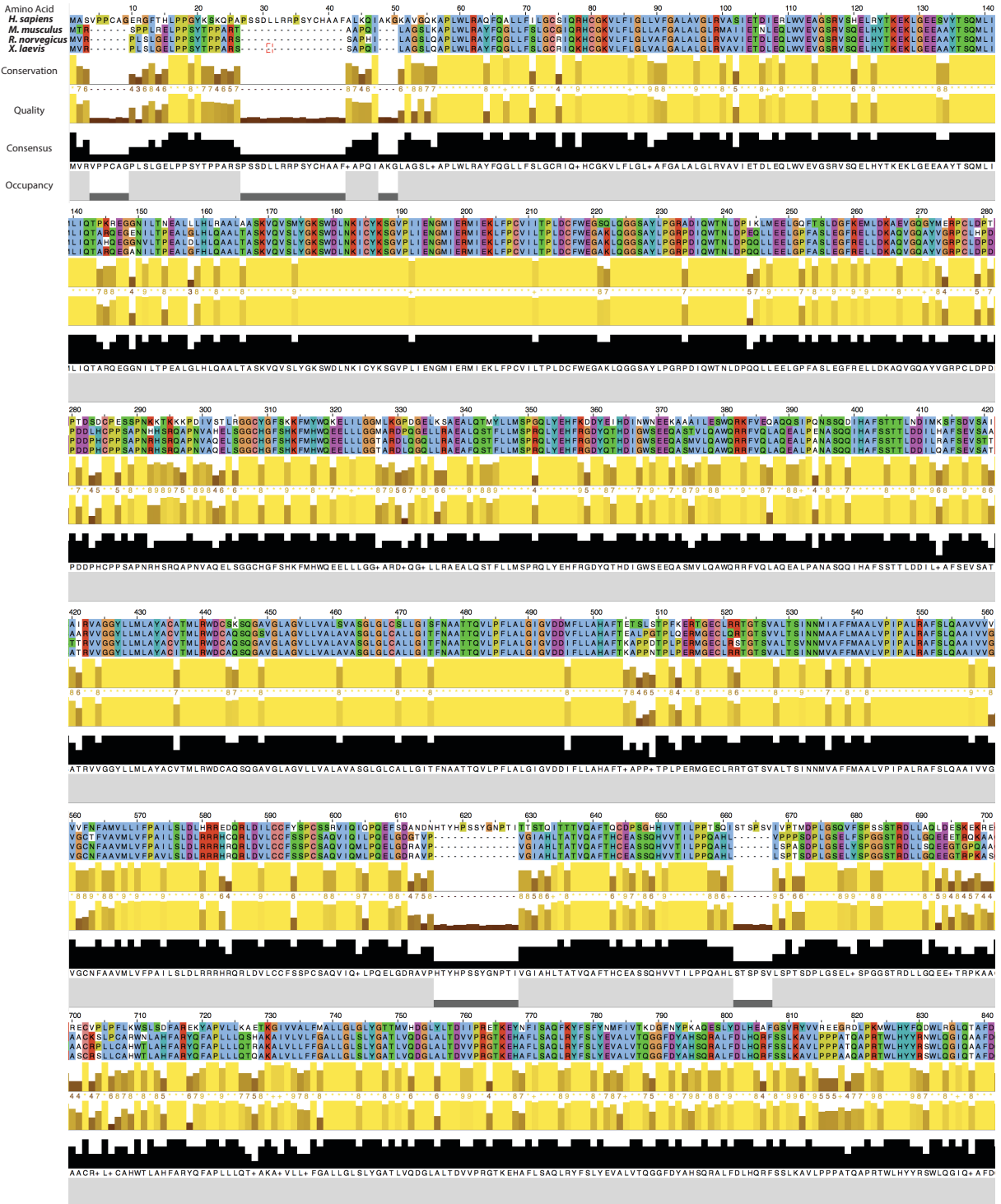


Figure 8.1: Multi sequence alignment of Patched1 homologs used for homolog screen (see chapter 6.1.1).

The alignment was calculated with Clustal Omega and visualized using Jalview 2.11.1.2. The colours were selected dependent on conservation, quality, consensus and occupancy (ClustalX). (178, 179)

Supplementary Information



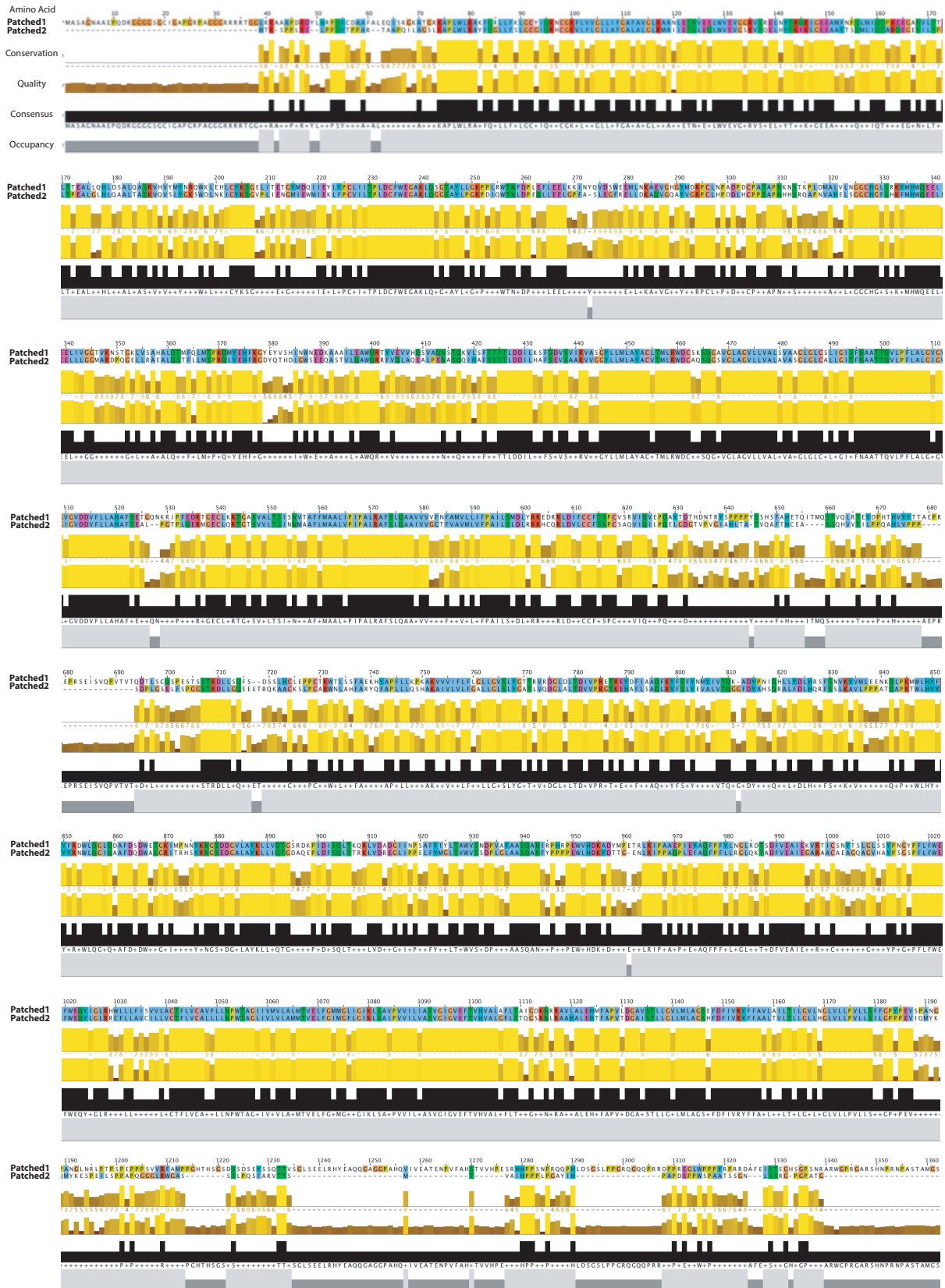
Supplementary Information



Figure 8.2: Multi sequence alignment of Patched2 homologs used for homolog screen (see chapter 6.1.1).

The alignment was calculated with Clustal Omega and visualized using Jalview 2.11.1.2. The colours were selected dependent on conservation, quality, consensus and occupancy (ClustalX). (178, 179)

Supplementary Information



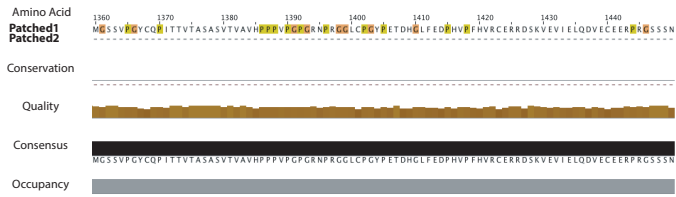


Figure 8.3: Protein sequence alignment of human Patched1 and Patched2 (see chapter 6.1.2). The alignment was calculated with Clustal Omega and visualized using Jalview 2.11.1.2. The colours were selected dependent on conservation, quality, consensus and occupancy (ClustalX). (178, 179)

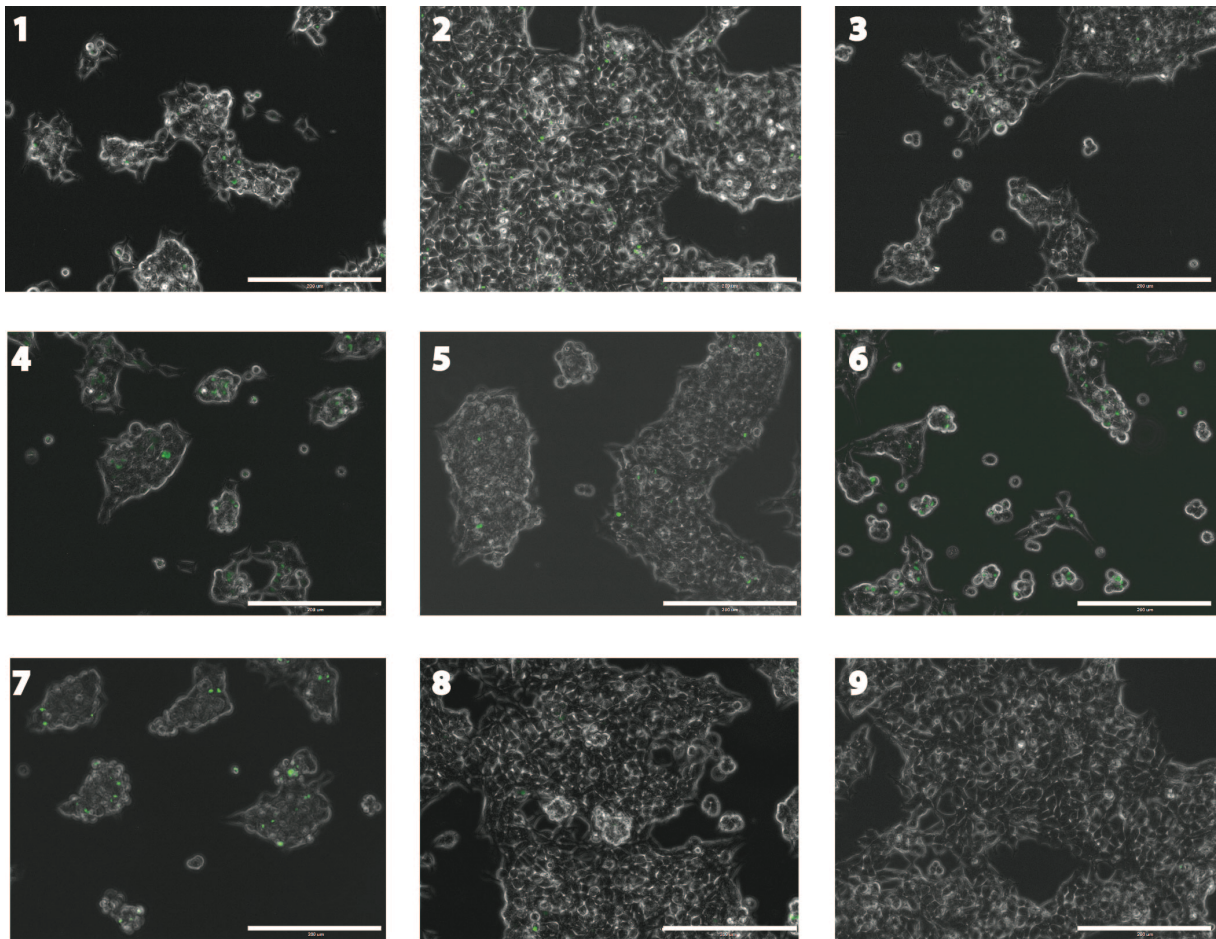
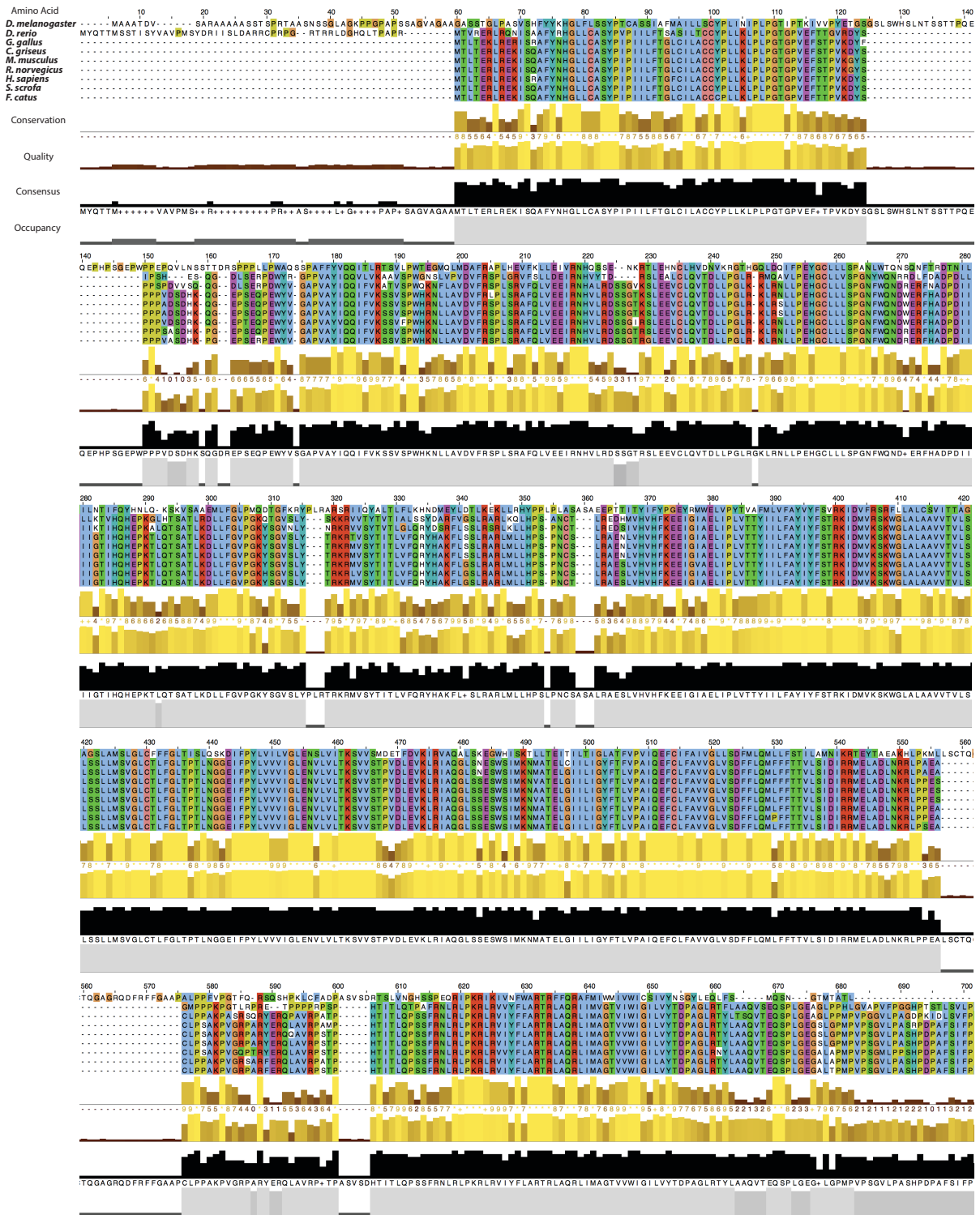


Figure 8.4: Fluorescence images of transiently transfected adherent HEK293 cells after 48 h expression of truncated human Patched1 variants (see chapter 6.1.2). Merged images of the eGFP fluorescence channel (green) and the bright field (grey) of Patched constructs 1-9 for the identification of a stabilized human Patched1 taken with the EVOSFL Auto Fluorescence microscope (Life Technologies). Scale bars: 200 µm.

Supplementary Information



Supplementary Information

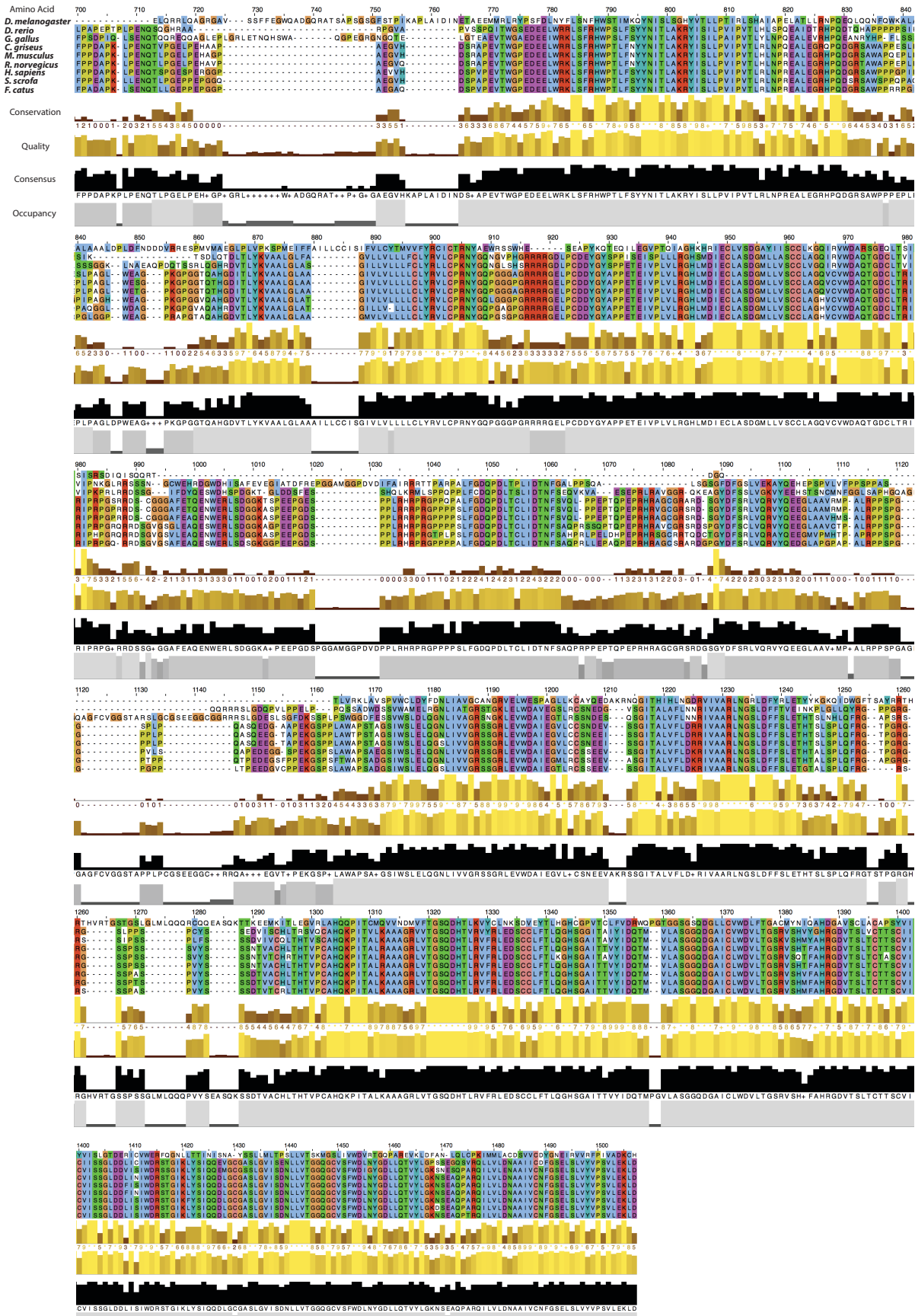
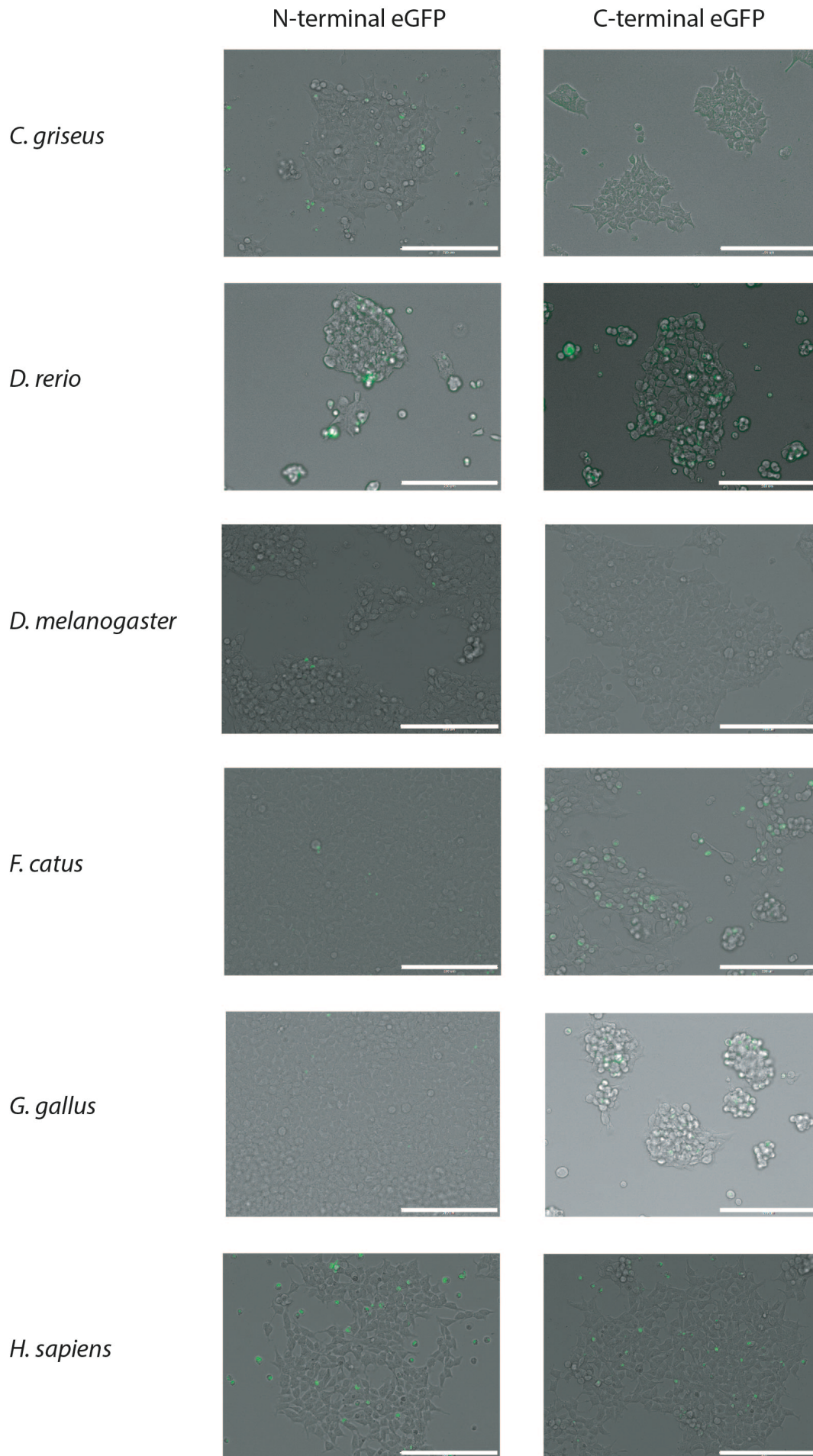


Figure 8.5: Multi-sequence alignment of Scapv for the homolog screen (see chapter 6.2.1.1). The alignment was calculated with Clustal Omega and visualized using Jalview 2.11.1.2. The colours were selected dependent on conservation, quality, consensus and occupancy (ClustalX). (178, 179)



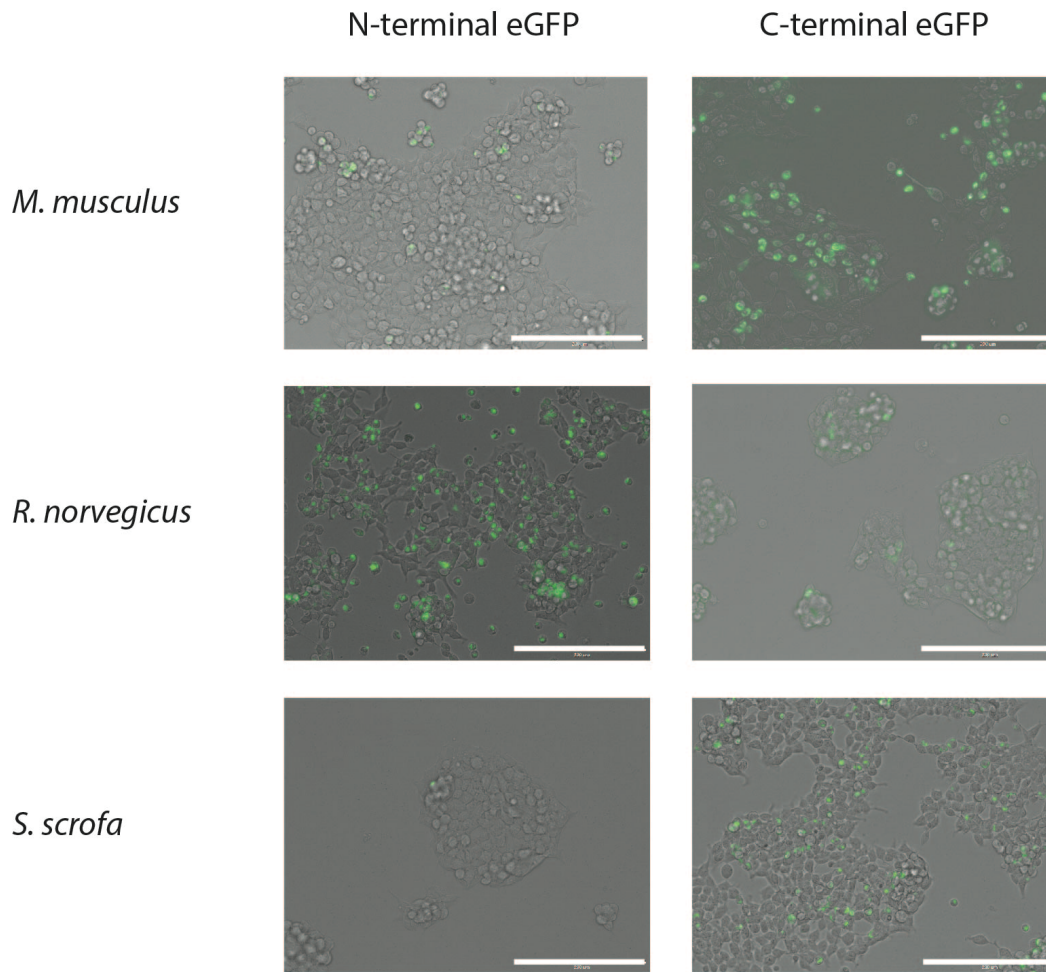


Figure 8.6: Fluorescence images of all Scap homologs expressed in adherent HEK293 T cells (see chapter 6.2.1.1).

Merged images of the fluorescence (green) and bright field channel (grey) of adherent HEK293 T cells over-expressing the nine different Scap homologs (*C. griseus*, *D. rerio*, *D. melanogaster*, *F. catus*, *G. gallus*, *H. sapiens*, *M. musculus*, *R. norvegicus* and *S. scrofa*) containing either N- (left column) or C-terminal (right column) eGFP reporter tags taken with the EVOSFL Auto Fluorescence microscope (Life Technologies). Scale bars: 200 μ m.

Table 8.1: Information on the purified porcine Scap verified by MS/MS analysis.

Gene names	Razor + unique peptides	Unique peptides	Mol. weight [kDa]	Sequence coverage [%]	Intensity	iBAQ	Potential contaminant
KRT1	5	2	66,038	6,8	1,9E+08	6,0E+06	+
SCAP	4	4	139,73	2,6	1,0E+08	2,1E+06	

MWSHPQFEKSSGWSHPQFEKSSGHHHHHHHHGGSENLYFQGGTLTERLREKISQAFYNHGLLCASYPIPIILFTGLCILACCYPLLLKLPPLPGTGPVEFTTPVKDYSPPPSASDHKPGPESEQPEWYVVGAPVAYIQQIFVKSSVSPWHKNL LAVDVFRSPLSRFAQLVEEIRNHVLRDSSGTRSLEEVCLQVTDLLPGLRKLNRNIPPEHGCLLLSPGNFQNDRERFHADPDIIGTIHQHEPKTLQTSATLTKDLLFGVPGKHSGVSLYTRKRLVSYTITLVFQHYHAKFLGSLRRLMLLH PSPNCSLRAESLVHVHFKEEIGIAELIPLVTTYIILFAYIYFSTRKIDMVKSKWGLALAAVTVLSSLLMSVGLCTLFGTPTLNGGEIFPYLVVVIGLENVLVLTKSVVSTPVDLEVKLRIAQGLSSESWSIMKNMATELGIILIGYFT LVPAIQEFCLFAVVGVLVSDFFLQMPFFTTVLSIDIRMEELADLNKRLPPEACLPPAKPVGRSARFERQLAVRPAT PHTITLQPSFRNLRLPKRLRVIYFLARTRLAQRLIMAGTVVWIGILAYTDPAGLRTYLAAQVTEQSPLEGALAMPVPSGVLPAHPDPAFSIFPPEAPKLEENQTLPGEPPEPGGQAEGVHDSPAPEVTWGPEDDEELWRKLSFRHWP TLFSSYINITLAKRYISLLPVIPTLRLNPREALEGRHPQDSRSAWSPQPAQGGGLWDAGPKGPGVAQAHRDVTLY KVAALGLATGILLVLLCLYRVLCPRNYQPAGAGPGRRRRGELPCDDYGYAPPETEIVPLVLRGHLMDIECLASD GMLLVSCCLAGHVCVWDAQTGDCLTRIPHPGRQRDSGVSLEAQQENWERLSDGGKASPEEPGDSPLRHRPRG TPLPSLFGDQPDLTCLIDTNFSAHPRLELDHPEPRHRSGCRRTQDCTGYDFSRLVQRAYQEEGMVPMHTPAPRP PSPGPTPPQTPEDEGSFPPEKGSPTSFTWAPSADGSIWSLELQGNLIVVGRSSGRLEVWDAIEGMLRCSSEEVASG ITALVFLDKRIVAARLNGSLDFFSLEHTALSPLQFRGAPGRGSSPTSVPYSSSDTVVCHLTHTVCAHQKPITA LKAAAGRLVTGSQDHTLRVFRLEDSCLFTLQGHSGAITTVYIDQTMVLASGGQDGAICLWDVLTGSRVSHMFAH RGDVTSLTCTTSCVISSGLDDLISIWDRSTGKLYSIQQDLGCGASLGVISDNLLVTGGQCVSFWDLNYGDLLQ TVYLGKDSEAQPARQILVLDNAAIVCNFGSELSLVYVPSVLEKLDLEVLVLFQGGGSMKGEELFTGVVPIILVELD GDVNGHKFSVSGEGEGDATYGKLTGKFICTTGKLPVPWPTLVTTLTGYVQCFSTRYPDHMKQHDFFKSAPPEGYVQ ERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIKDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRH NIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDMVLLLEFVTAAGITHGMDELYK

Figure 8.7: MS/MS analysis of porcine Scap.

The green highlighted peptides could be verified by MS/MS analysis after tryptic digest.

Table 8.2: MS/MS analysis of heat-shock protein HSP70.

Gene names	Peptides	Razor + unique peptides	Unique peptides	Mol. weight [kDa]	Sequence coverage [%]	Intensity	iBAQ
HSP70	53	53	13	70,051	84,4	3,4E+11	1,1E+10
-	7	7	7	24,409	31,6	4,5E+10	4,1E+09
PCCA	62	62	62	80,058	75	1,6E+11	3,4E+09
-	60	60	42	172,74	53	4,3E+10	7,2E+08

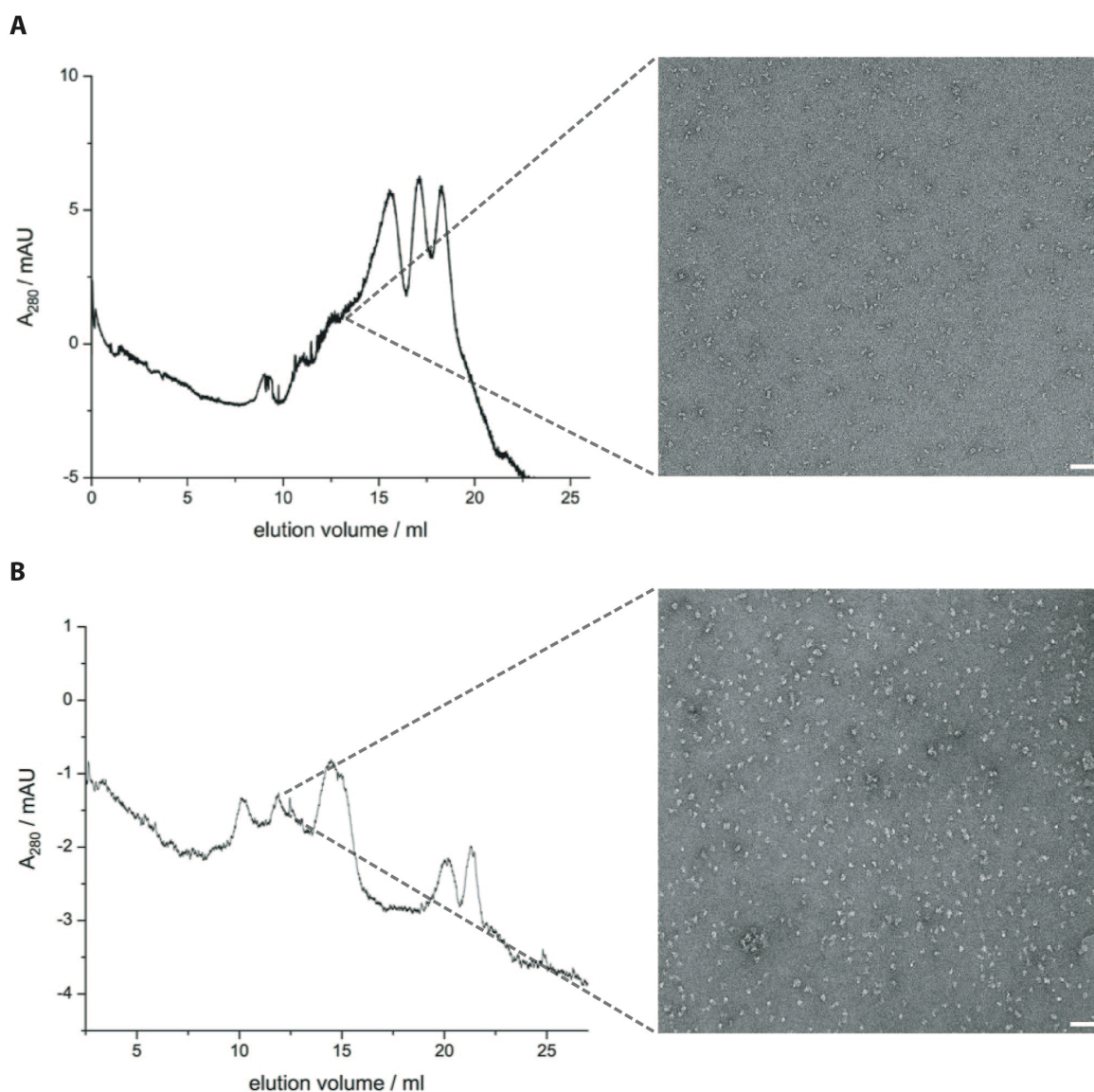


Figure 8.8: Reconstitution of porcine Scap into lipid nanodiscs composed of Msp1E3D1, POPC and 5 mol% cholesterol.

(A) Chromatogram of SEC (left) after reconstitution of porcine Scap solubilized in DDM/CHS into lipid nanodiscs composed of Msp1E3D1, POPC and 5 mol% cholesterol after detergent removal via dialysis. The chromatogram as well as the representative negative stain micrograph (right) of the elution fraction at 12.5 ml feature mainly degradation products of Scap. (B) Chromatogram of SEC (left) after reconstitution of porcine Scap solubilized in DDM/CHS into lipid nanodiscs composed of Msp1E3D1, POPC and 5 mol% cholesterol after detergent removal via biobeads. The chromatogram as well as the representative negative stain micrograph (right) of the elution fraction at 12.5 ml feature mainly degradation products of Scap. Furthermore, additional peak fractions at earlier retention volumes indicate the protein's aggregation at SEC (left).

9 Publications and Conference Contributions

Publications

T. Wagner, F. Merino, M. Stabrin, T. Moriya, C. Antoni, A. Apelbaum, P. Hagel, O. Sitsel, T. Raisch, D. Prumbaum, D. Quentin, D. Roderer, Sebastian Tacke, **B. Siebolds**, E. Schubert, T. R. Shaikh, P. Lill, C. Gatsogiannis and S. Raunser: SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Communications Biology*, 2019, 2, 218.

Conferences and Courses

Juni 2018

Workshop: 4th International SPHIRE Workshop: Single Particle Analysis of High Resolution Cryo-EM-Data.

Max Planck Institute of molecular Physiology, Dortmund

September 2018

Conference: Chemistry at the Interface of Biology and Medicine

Oral Presentation: A general approach towards Membrane protein Production for Structural Studies.
University of Crete, Institute of Organic Chemistry

November 2018

Symposium: Unravelling Cellular Intricacies: From Experiment to Insight

Poster Presentation: A general approach towards Membrane protein Production for Structural Studies.
Max Planck Institute of molecular Physiology, Dortmund