

# Dissertation

## Structural Investigation of Cholesterol Homeostasis and Bacterial Toxins

---

Vorgelegt von

**Patrick Günther**

Zur Erlangung des akademischen Grades des Doktors der Naturwissenschaft

(Dr. rer. nat.)

Vorgelegt der Fakultät Chemie und Chemische Biologie  
der Technischen Universität Dortmund

Angefertigt am

Max-Planck-Institut für Molekulare Physiologie Dortmund  
Abteilung für Strukturbiochemie

Prüfungsdatum: 22.08.2023, Dortmund









**First Referee:**

Prof. Dr. Stefan Raunser

Department of Structural Biochemistry, Max-Planck Institute for 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



# Contents

<b>1</b>	<b>Abbreviations</b> .....	<b>IV</b>
<b>2</b>	<b>Abstract</b> .....	<b>VI</b>
<b>3</b>	<b>Zusammenfassung</b> .....	<b>VIII</b>
<b>4</b>	<b>Table of Figures</b> .....	<b>X</b>
<b>5</b>	<b>Introduction</b> .....	<b>1</b>
5.1	Membrane proteins in biological membranes .....	1
5.2	Biological role of cholesterol and its production .....	2
5.3	Biosynthetic pathway of cholesterol .....	3
5.4	HMGCR and its regulation .....	6
5.5	Pharmaceutical inhibition of the HMGCR .....	8
5.6	Feedback mechanisms to control the activity of the HMGCR .....	10
5.7	Structural studies on the HMGCR-UBIAD1 complex .....	14
5.8	Cholesterol homeostasis by the SREBP signaling pathway .....	16
5.9	Structural studies of the SCAP-Insig complex .....	18
5.10	Bacterial secretion systems .....	23
5.11	The Type VI secretion system .....	24
5.12	Discovery of Rhs-repeat containing proteins .....	28
5.13	Structures of Rhs-repeat containing proteins .....	30
<b>6</b>	<b>Aims and motivation</b> .....	<b>34</b>
<b>7</b>	<b>Materials</b> .....	<b>36</b>
7.1	Instruments .....	36
7.2	Chromatographic Columns .....	37
7.3	Chemicals .....	37
7.4	Consumables .....	39
7.5	Plasmids .....	40
7.6	Enzymes and antibodies .....	42
7.7	Cell Lines .....	42
7.8	Buffer Solutions .....	43
<b>8</b>	<b>Methods</b> .....	<b>45</b>
<b>8.1</b>	<b>Molecular biology</b> .....	<b>45</b>
8.1.1	Polymerase chain reaction (PCR) .....	45
8.1.2	Circular polymerase extension cloning (CPEC) .....	45
8.1.3	Site-directed mutagenesis .....	45
8.1.4	Chemical transformation of <i>Escherichia coli</i> cells .....	46
8.1.5	Plasmid isolation .....	46
8.1.6	Sanger sequencing .....	46
<b>8.2</b>	<b>Cell biology</b> .....	<b>46</b>
8.2.1	Culturing and transfection of human adherent cell lines .....	46
8.2.2	Culturing of human suspension cells lines .....	47
8.2.3	Transient transfection of human adherent cells .....	47
8.2.4	Electroporation of DH10EmBacY <i>E. coli</i> cells .....	47
8.2.5	Bacmid isolation from DH10EmBacY cells .....	48
8.2.6	Transfection of <i>Sf9</i> insect cells – baculovirus generation and amplification .....	48
8.2.7	Endpoint dilution assay .....	48
8.2.8	Transduction of HEK293 GnTI <sup>-</sup> suspension cells .....	49
<b>8.3</b>	<b>Protein biochemistry</b> .....	<b>49</b>
8.3.1	Semi-denaturing SDS-PAGE .....	49

8.3.2	Western blot.....	50
8.3.3	Blue native PAGE.....	50
8.3.4	FSEC-TS with whole cell lysates .....	50
8.3.5	Detergent-screening .....	51
8.3.6	Protein purification of SCAP-Insig and HMGCR-UBIAD1 complexes .....	52
8.3.7	Protein purification HMGCR .....	52
8.3.8	Enzymatic activity assay for the HMGCR.....	53
8.3.9	Reconstitution of membrane proteins into nanodiscs .....	53
8.3.10	Reconstitution of membrane proteins into liposomes .....	54
<b>8.4</b>	<b>Structural biology .....</b>	<b>54</b>
8.4.1	Transmission electron microscopy – negative staining .....	54
8.4.2	Cryo-electron microscopy.....	55
<b>8.5</b>	<b>Methods specific to RhsA .....</b>	<b>57</b>
8.5.1	Bacterial strains and culture conditions.....	57
8.5.2	DNA manipulation and plasmid construction.....	58
8.5.3	Generation of <i>P. protegens</i> mutants .....	58
8.5.4	<i>Pseudomonas</i> growth competition assays.....	58
8.5.5	Toxicity assays.....	59
8.5.6	Protein expression and purification .....	59
8.5.7	Purification of RhsA $\Delta$ TMD and VgrG1 for cryo-EM.....	59
8.5.8	Western blot analyses .....	60
8.5.9	<i>De novo</i> protein sequencing via LC-MS/MS.....	60
8.5.10	Negative stain electron microscopy.....	60
8.5.11	Sample vitrification and data collection .....	60
8.5.12	Cryo-EM image processing.....	61
8.5.13	Model building, refinement and validation .....	62
<b>9</b>	<b>Results and discussion .....</b>	<b>64</b>
<b>9.1</b>	<b>Structural characterization of the Scap-Insig complex .....</b>	<b>64</b>
9.1.1	Small-scale expression of the human SREBP, Scap and Insig complex .....	65
9.1.2	Initial expression and purification of the Scap-Insig complex .....	68
9.1.3	Optimization of the purification workflow of the Scap-Insig complex .....	72
9.1.4	Stabilization of the Scap-Insig complex .....	74
9.1.5	Cryo-EM screening and data collection of the Scap-Insig complex in GDN.....	77
9.1.6	Improved sample preparation of the Scap-Insig complex in LMNG.....	79
9.1.7	Cryo-EM screening and data collection of the Scap-Insig complex in LMNG.....	83
9.1.8	Nanodisc reconstitution of the Scap-Insig complex .....	85
9.1.9	Conclusion Scap-Insig complex .....	90
<b>9.2</b>	<b>Structural characterization of the HMG-CoA reductase.....</b>	<b>94</b>
9.2.1	Small-scale expression of the human HMGCR.....	94
9.2.2	Large scale expression and purification of the HMGCR.....	96
9.2.3	Biochemical characterization of the HMGCR.....	97
9.2.4	Structural characterization of the full-length HMGCR .....	99
9.2.5	Identification of a stable HMGCR-UBIAD1 complex .....	102
9.2.6	Optimization of the HMGCR-UBIAD1 complex purification .....	108
9.2.7	HMGCR-UBIAD1 complex in LMNG-CHS .....	115
9.2.8	Conclusion HMGCR-UBIAD1 complex.....	119
<b>9.3</b>	<b>Concluding Discussion.....</b>	<b>124</b>
9.3.1	Structural determination of small membrane proteins using cryo-EM.....	124
<b>9.4</b>	<b>Structure of a bacterial Rhs effector exported by the type VI secretion system .....</b>	<b>128</b>
9.4.1	Abstract.....	129
9.4.2	RhsA forms a cocoon-like structure that undergoes N- and C-terminal autocleavage..	129
9.4.4	RhsA possesses a unique plug domain at its N-terminus .....	135
9.4.5	Comparison between Rhs proteins of known structure .....	138
9.4.6	Architecture of the pre-firing complex .....	140
9.4.7	Model of RhsA firing events and toxin release .....	141

<b>10</b>	<b>References .....</b>	<b>145</b>
<b>11</b>	<b>Appendix .....</b>	<b>166</b>
<b>12</b>	<b>Acknowledgements .....</b>	<b>182</b>
<b>13</b>	<b>Eidesstaatliche Versicherung .....</b>	<b>184</b>

## 1 Abbreviations

°C	Degree Celsius
μl, ml, l	Microliter, milliliter, liter
24,25-DHL	24,25-dihydrolanosterol
25-HC	25-hydroxy cholesterol
aa	Amino acid
AcMNPV	<i>Autographa californica</i> multiple nucleopolyhedrosis virus
BacMam	Baculovirus transduction of mammalian cells
bHLH-LZ	Basic-helix-loop-helix leucine zipper
BRIL	Thermostabilized apocytochrome b562RIL
CHS	Cholesteryl hemi succinate
CoA	Coenzyme A
cryo-EM	Cryo-electron microscopy
cryo-ET	Cryo-electron tomography
CTD	Carboxy-terminal domain
Da, kDa	Dalton, kilo Dalton
DDM	<i>n</i> -dodecyl-β-D-maltopyranoside
DM	<i>n</i> -decyl-β-D-maltopyranoside
DNA	Desoxyribonucleic acid
Eag	Effector-associated gene
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enhanced green-fluorescent protein
EM	Electron microscopy
FSEC	Fluorescence-detection size-exclusion chromatography
FSEC-TS	Fluorescence-detection size-exclusion chromatography-based thermostability assay
GDN	Glyco diosgenin
GnTI <sup>-</sup>	N-acetyl-glucosaminyl transferase I negative
h, min, s	Hour, minute, second
HEK	Human embryonic kidney
HMG-CoA	3-hydroxy-3-methylglutaryl-CoA
HMGCR	HMG-CoA-reductase
HRV 3C Protease	Recombinant type 14 3C protease from human rhinovirus
IMP	Integral membrane protein

---

LB	Luria-Broth
LMNG	Lauryl maltose neopentyl glycol
MSP	Membrane scaffold protein
MW	Molecular weight
NADPH	Nicotinamide adenine dinucleotide phosphate
ND	Nanodisc
nm	Nanometer
NTA	Nitrilotriacetic acid
PAAR	Cone-shaped proline-alanine-alanine-arginine
Rhs	Rearrangement hotspot
SapA	Saposin A
Scap	SREBP cleavage-activating protein
SEC	Size-exclusion chromatography
<i>Sf9</i>	<i>Spodoptera frugiperda</i> insect cell line
SGS1	Salivary gland surface protein 1
SMALPs	styrene maleic acid co-polymer lipid particles
SNR	Signal-to-noise ratio
SPA	Single particle analysis
SREBP	Sterol regulatory-element binding protein
T6SS	Type VI secretion system
TagBFP	Blue fluorescent protein
TEM	Transmission electron microscopy
TM	Transmembrane
TMD	Transmembrane domain
UBIAD1	UbiA prenyltransferase domain-containing protein 1
VgrG	Valine glycine repeat G protein
YFP	Yellow fluorescent protein

## 2 Abstract

Membrane proteins regulate a variety of processes that are critical for living organisms. They participate in cell-cell communication, catalyze reactions in or at the membrane, are involved in transmitting signals from the environment into the cell, and can transport molecules across membranes. Approximately 60% of all clinically approved drugs target membrane proteins, underscoring their importance. In order to understand the function of membrane proteins and to design more targeted drugs, determining their precise three-dimensional structures is required. In this PhD project, I aimed to structurally characterize two membrane protein complexes involved in the regulation of cholesterol homeostasis and the type VI secretion system (T6SS) effector RhsA.

Cholesterol is a lipid molecule and is important to regulate membrane fluidity but also serves as precursor for the production of hormones or bile acids. It can be synthesized in a biosynthetic cascade involving more than 30 steps. Increasing cholesterol levels in the membranes are sensed by proteins of the SREBP pathway, SREBP, Scap and Insig, which control the expression of enzymes required for cholesterol uptake and *de novo* synthesis. Here I investigated the interaction of human Scap with Insig1. After co-expression, I showed that Scap and Insig interact and can be co-purified. I established a complete purification protocol to routinely obtain several hundred micrograms per liter expression culture of the pure complex. I determined biochemically via SEC-MALS that Scap and Insig form a 1:1 complex. Sufficient amounts of the 1:1 complex of Scap and Insig were purified and subjected to single particle cryo-EM analysis. However, the quality of the reconstruction was limited due to the flexibility of the WD40 of Scap. Since several other groups published structures of Scap and the Scap-Insig complex during my PhD, I stopped working on this project.

Instead, I continued working on the HMG-CoA reductase (HMGCR) which catalyzes the rate-limiting step in the *de novo* synthesis of cholesterol. It catalyzes the conversion of HMG-CoA to mevalonate, which is used for the synthesis of a plethora of sterol and non-sterol-based molecules. I initially expressed the full-length human HMGCR in mammalian cells and established a complete purification protocol. The purified enzyme was catalytically active with kinetic parameters consistent with those previously reported for the isolated reductase domain. However, I was able to show for the first time that the enzyme kinetics for the full-length enzyme followed curves obtained for substrate-inhibited enzymes such as phosphofructokinase. I attempted to structurally characterize the purified HMGCR, but the particles displayed a heterogeneous size distribution on EM grids, suggesting the co-existence of different oligomers. Instead, I decided to study the interaction of the transmembrane domain of the



HMGCR with the regulatory protein UBIAD1. I succeeded in co-expressing optimized constructs containing the TMDs of both proteins in mammalian cells and found that they co-migrate on size exclusion chromatography after solubilization, indicating direct interaction. I successfully scaled up the expression and established a tandem affinity purification protocol. Milligram amounts of the pure complex were obtained, which migrated as a single gaussian-shaped peak in SEC. I subjected this complex to cryo-EM SPA analysis and obtained a reconstruction with a resolution of 7.3 Å. Prior to further optimization, the structure of the hamster HMGCR-UBIAD1 complex was reported by another group. Therefore, I also abandoned this project and worked instead on bacterial effectors that are secreted by the T6 secretion system (T6SS).

The T6SS is a widely distributed secretion device conferring an advantage of bacteria in their biological niche. This apparatus is mostly used to directly deliver toxic effector proteins into competing bacteria. One such effector is RhsA. Secondary structure predictions revealed that RhsA contains repetitive sequences known as rearrangement hotspot (Rhs) regions. Such repeats have been found in other protein families, such as the BC components of Tc toxins and teneurins. In these structures, the Rhs repeats form a large hollow beta-barrel cocoon that can encapsulate a cargo. However, when I started to work with it, the structure of a bacterial T6SS effector was unknown. Therefore, I set out to solve the structure of RhsA from *Pseudomonas protegens*. I determined the structure of RhsA and its cognate spike protein VgrG1, both to a resolution of 3.3Å. The Rhs repeats of RhsA form a large anticlockwise spiral that encapsulates the toxic effector. I found that RhsA undergoes two cleavages at the N- and C-termini, which are required for interbacterial competition. Combining the biochemical data from my collaborators and previous knowledge, we were able to propose a model for RhsA mounting, delivery, and toxin release. This provides a framework for determining T6SS Rhs effector structures and sheds light on the similarities and differences of Rhs repeat-containing proteins. My PhD work showcases that biochemical studies combined with structural determination by cryo-EM provides valuable insights into molecular processes that occur in or at the membrane and is of utmost pharmacological interest.

### 3 Zusammenfassung

Membranproteine sind an der Regulierung einer Vielzahl von Prozessen beteiligt. So spielen sie zum Beispiel eine Rolle für die Weiterleitung von Signalen in die Zelle oder katalysieren Reaktionen an oder in der Membran. Pharmakologisch gesehen sind sie ebenfalls von großer Bedeutung, da etwa 60 % aller klinisch zugelassenen Medikamente die Aktivität von Membranproteinen modulieren. Daher ist es unabdingbar ihre dreidimensionale Struktur zu bestimmen, um daraus Schlüsse auf ihre Funktion und Arbeitsweise zu machen. Das Ziel dieser Arbeit war es zwei Membranproteinkomplexe, die an der Regulierung der menschlichen Cholesterin-Homöostase beteiligt sind, sowie den Typ VI Sekretionssystems (T6SS) Effektor RhsA strukturell zu charakterisieren.

Cholesterin ist ein Lipidmolekül, das sich in Membranen einlagert und ihre Fluidität bestimmt. Außerdem fungiert es als Ausgangsmolekül für die Biosynthese von Sexualhormonen und Gallensäuren. Ein Anstieg der Cholesterinkonzentration in der Membran wird von Proteinen des SREBP-Signalweges, SREBP, Scap and Insig, wahrgenommen, die daraufhin die Expression von Proteinen der Cholesterinsynthese steuern. Ist genug Cholesterol in den Membranen der Zellen vorhanden, wird die körpereigene Synthese sowie die Aufnahme aus dem Blut gestoppt. Um diese Regulation auf molekularer Ebene zu verstehen, sind hochauflösende Strukturen der beteiligten Proteinkomplexe notwendig. In dieser Arbeit wurde die Interaktion von dem menschlichen Scap mit Insig1 mittels Kryo-EM untersucht. Dazu habe ich beide Proteine koexprimiert und in kleinem Maßstab aufgereinigt. Ich konnte schrittweise die Aufreinigungsstrategie optimieren und ein Aufreinigungsprotokoll etablieren. Mittels SEC-MALS habe ich festgestellt, dass Scap und Insig einen 1:1 Heterodimer bilden. Dieser biochemisch reine Komplex wurde dann mittels Einzelpartikel-Kryo-EM-Analyse untersucht. Die Auflösung der Rekonstruktion war jedoch aufgrund der Flexibilität der WD40 Domäne von Scap limitiert. Gleichzeitig erschienen Publikationen anderer Gruppen, die die Struktur von Scap oder des Komplexes aus Scap und Insig beschrieben, so dass ich meine Arbeit an diesem Projekt einstellte.

Stattdessen habe ich meinen Fokus auf die HMG-CoA Reduktase gelegt, die den geschwindigkeitsbegrenzenden Schritt der körpereigenen Cholesterinsynthese katalysiert. Diese wandelt HMG-CoA in Mevalonat um, einen wichtigen Baustein für die weitere Synthese einer Vielzahl von Molekülen, darunter auch Cholesterin. Zunächst habe ich ein Volllängenkonstrukt der humanen HMGR exprimiert und anschließend gereinigt. Ich konnte die enzymatische Aktivität des gereinigten Proteins messen, die im gleichen Bereich lag, wie zuvor für die katalytische Untereinheit berichtet. Allerdings zeigte das Volllängenprotein eine

Reduktion der Reaktionsgeschwindigkeit bei hohen Substratkonzentrationen. Eine solche Substratinhibierung wurde bisher nur für andere Enzyme des Stoffwechsels beschrieben und konnte hier erstmals für die HMGCR gezeigt werden. Jedoch gelang es mir nicht, die Struktur der HMGCR bestimmen. Stattdessen habe ich die Membrandomäne der HMGCR mit UBIAD1 koexprimiert, um die HMGCR zu stabilisieren. Ich konnte die Interaktion der beiden Proteine nachweisen und den Komplex anschließend in hoher Reinheit aufreinigen. Außerdem konnte ich mittels Einzelpartikel-Kryo-EM-Analyse eine dreidimensionale Rekonstruktion mit einer Auflösung von 7,3 Å bestimmen. Bevor ich die Probe weiter optimieren konnte, wurde eine Struktur von der HMGCR im Komplex mit UBIAD1 des Hamsters veröffentlicht. Aus diesem Grund habe ich meine Arbeit an diesem Projekt zu diesem Zeitpunkt eingestellt und mich auf bakterielle Effektoren konzentriert, die über das Typ 6 Sekretionssystem exportiert werden.

Das T6SS ist weit verbreitet in Bakterien. Es wird hauptsächlich dazu genutzt giftige Effektoren in benachbarte und rivalisierende Bakterien zu befördern, was einen Überlebensvorteil darstellt. Der Effektor RhsA von *Pseudomonas protegens* kodiert repetitive Sequenzen die als *rearrangement hotspot (Rhs)* Regionen bekannt sind. Andere Proteine, die diese Sequenzen beinhalten formen große Kokons, die eine toxische Proteinkomponente in ihrem Inneren einschließen. Die Struktur eines Rhs-Effektors, der durch T6SS transportiert wird, war zu diesem Zeitpunkt jedoch noch nicht gelöst. In meiner Arbeit konnte ich die Struktur von RhsA und die des T6SS Spike-Proteins VgrG1 mit einer Auflösung von 3,3 Å bestimmen. Die Struktur zeigt, dass auch RhsA einen Kokon bestehend aus einer Spirale von *Rhs-repeats* bildet. Im inneren dieses Kokons ist der toxische Teil des Effektors eingeschlossen und beschützt somit das produzierende Bakterium. Außerdem konnte ich feststellen, dass RhsA sich autokatalytisch an den N- und C-termini schneidet, was wichtig für die toxische Aktivität des Effektors ist. Das erlaubte uns, mit Einbeziehung vorheriger Erkenntnisse meiner Kooperationspartner, einen Mechanismus aufzustellen wie RhsA sekretiert wird und wie möglicherweise das Toxin aus dem Inneren freigesetzt wird. Zusammenfassend erweitert meine Arbeit das Verständnis bakterieller Rhs-Effektoren und bildet die Grundlage für die Bestimmung weiterer Strukturen von T6SS Rhs-Effektoren.

Insgesamt stellt meine Arbeit die Stärke von biochemischen und strukturbestimmenden Methoden heraus, um Prozesse, die an oder in der Membrane stattfinden heraus und ist von großem pharmakologischem Nutzen.

## 4 Table of Figures

Figure 1: Overview of different types of membrane proteins.....	1
Figure 2: Structure of steroids and cholesterol. ....	2
Figure 3: Cholesterol biosynthetic pathway.....	4
Figure 4: Structure of the catalytic portion of the HMGCR. ....	8
Figure 5: Structural formulas of the first statin in its next generation derivatives.....	10
Figure 6: Sterol-accelerated degradation of the HMGCR via the ERAD pathway. ....	12
Figure 7: Model for the role of UBIAD1-mediated stabilization of the HMGCR. ....	13
Figure 8: Structures of the HMGCR-UBIAD1 complex solved by cryo-EM. ....	15
Figure 9: Control of cholesterol homeostasis by the SREBP signaling pathway. ....	17
Figure 10: Cryo-EM structures of Scap and Scap-Insig complex reveal their molecular architecture.....	19
Figure 11: Schematic model of the cSCAP conformational changes upon Insig binding.....	22
Figure 12: Schematic and structural representation of secretion systems in Gram-negative bacteria. ....	24
Figure 13: Tomographic slices of an intact T6SS from <i>Myxococcus xanthus</i> .....	26
Figure 14: Model for assembly and secretion of PAAR-domain containing effectors.....	27
Figure 15: Structure of a TcB-TcC cocoon and Teneurin2. ....	31
Figure 16: Cryo-EM structure of the salivary gland surface protein 1 (SGS1). ....	32
Figure 17: Small-scale expression of complexes of the SREBP signaling pathway. ....	67
Figure 18: Initial one-step affinity purification of the Scap-Insig complex. ....	69
Figure 19: Preliminary negative stain analysis of the potential Scap-Insig complex. ....	71
Figure 20: Tandem-affinity purification of the Scap <sup>D428A</sup> -Insig complex in DDM-CHS.....	73
Figure 21: Tandem-affinity purification of the Scap <sup>D428A</sup> -Insig complex in GDN. ....	75
Figure 22: Single particles analysis of the Scap-Insig complex in GDN.....	78
Figure 23: Optimized purification procedure of the Scap-Insig complex in LMNG-CHS. ....	80
Figure 24: Molecular mass analysis of the Scap-Insig complex.....	82
Figure 25: Cryo-EM data collection and processing of the Scap-Insig complex in LMNG-CHS. ....	85
Figure 26: Nanodisc reconstitution of the Scap-Insig complex.....	87
Figure 27: Cryo-EM screening and data collection of the nanodisc reconstituted Scap-Insig complex.....	88
Figure 28: Detergent-screen of the human HMGCR.....	95
Figure 29: Large-scale purification of the human HMGCR. ....	96

---

Figure 30: Final purification step of the HMGCR via SEC. ....	97
Figure 31: Michaelis-Menten kinetics of the full-length HMGCR. ....	98
Figure 32: Negative stain analysis of the full-length HMGCR in DDM micelles. ....	100
Figure 33: BN-PAGE of the purified HMGCR in DDM micelles. ....	101
Figure 34: Cryo-EM analysis of the full-length HMGCR. ....	102
Figure 35: Identification of HMGCR-UBIAD1 interaction by FSEC. ....	104
Figure 36: Single-step affinity purification of the HMGCR-UBIAD1 complex. ....	106
Figure 37: SEC profiles of HMGCR-UBIAD1 complexes. ....	107
Figure 38: Negative stain EM analysis of HMGCR-UBIAD1 complexes. ....	108
Figure 39: Alphafold prediction of the HMGCR <sup>TMD</sup> -BRIL construct. ....	110
Figure 40: Purification of the HMGCR <sup>TMD</sup> -BRIL-UBIAD1 <sup>G186R</sup> complex. ....	111
Figure 41: Characterization of the HMGCR <sup>TMD</sup> -BRIL-UBIAD1 <sup>G186R</sup> -BAG2 complex. ....	112
Figure 42: Cryo-EM analysis of the HMGCR <sup>TMD</sup> -BRIL-UBIAD1 <sup>G186R</sup> -BAG2 complex. ....	114
Figure 43: Characterization of the HMGCR <sup>TMD</sup> -BRIL-UBIAD1 <sup>G186R</sup> complex in LMNG-CHS. ....	115
Figure 44: Cryo-EM analysis of the HMGCR <sup>TMD</sup> -BRIL-UBIAD1 <sup>G186R</sup> complex in LMNG-CHS. ....	117
Figure 45: High-resolution structure of the T6SS effector RhsA. ....	130
Figure 46: Autoproteolysis of RhsA occurs at its N- and C-terminus. ....	133
Figure 47: A unique plug domain seals the Rhs barrel of RhsA. ....	136
Figure 48: Structural comparison of Rhs repeat containing proteins. ....	139
Figure 49: High-resolution structure of <i>P. protegens</i> VgrG1. ....	140
Figure 50: Model of T6SS-dependent delivery of RhsA into the cytoplasm of a susceptible bacterial cell. ....	143



## 5 Introduction

### 5.1 Membrane proteins in biological membranes

Within eukaryotic cells, closed internal compartments are separated by single or double membranes. This allows the generation of distinct chemical reaction spaces and thus different chemical milieus within the same cell. Additionally, membranes play an important role in maintaining the concentration gradients between two reaction spaces, which, for example, are linked to energy production or propagating signals from the outside into internal signals. However, this may imply that specific proteins are required for the transmission of signals from the outside of the cell into the interior. Or they might be involved in the transport of molecules across the membrane.

These proteins are known as membrane proteins and are classified into two main classes: integral (IMP) or peripheral membrane proteins (Figure 1). They can be further subdivided based on their architecture, orientation, whether they span only one or both leaflets of the membrane, or whether they bind peripherally via lipid modifications or hydrophobic sites. However, membrane proteins are not the only component of biological membranes.

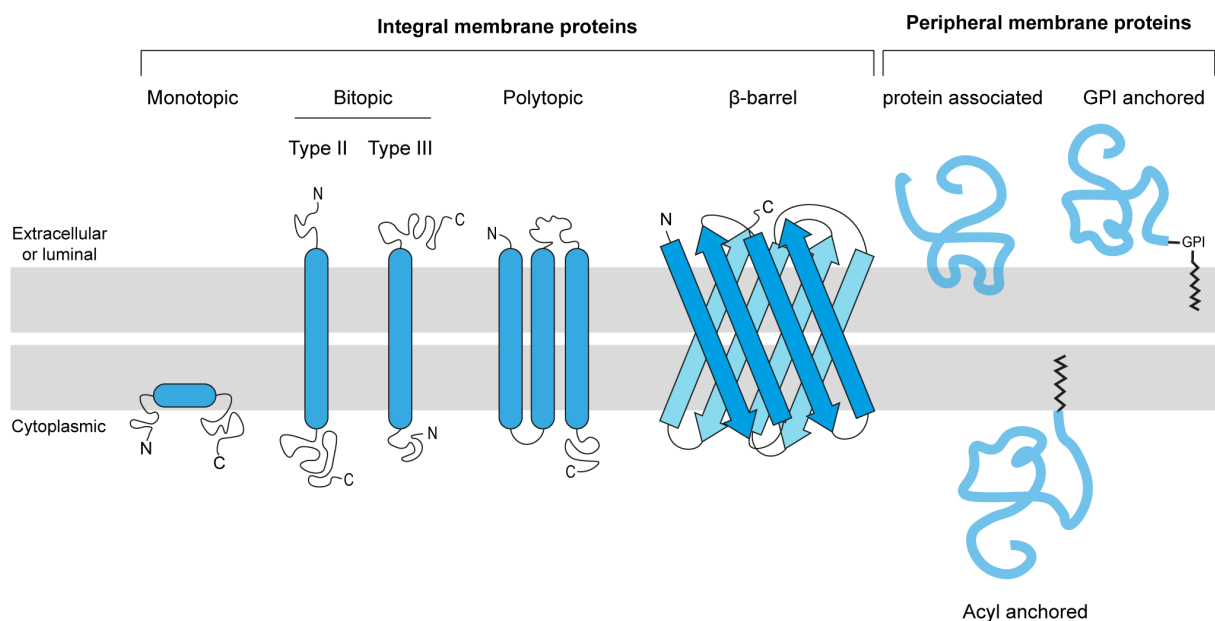


Figure 1: Overview of different types of membrane proteins

Membrane proteins can be divided into integral and peripheral membrane proteins. IMPs can be further divided into mono-, bi- or polytopic membrane proteins. A separate class is represented by  $\beta$ -barrel proteins. Lipid modifications or hydrophobic patches on the protein surface allow peripheral membrane proteins to localize to the membrane.

## 5.2 Biological role of cholesterol and its production

Biological membranes of animals are mainly composed of phospholipids and sphingolipids as well as the essential molecule cholesterol.<sup>1</sup> Cholesterol was the first discovered steroid which has been isolated from gall stones. The simplest steroid backbone consists of 17 carbon atoms arranged in four bonded rings, three six-membered cyclohexyl rings and a five-membered cyclopentane ring. Sterols are typically modified versions of this simple backbone and contain a hydroxyl group at position 3 and an eight-carbon side chain at position 17 (Figure 2). These molecules are highly hydrophobic which results in their exclusive partition into biological membranes.

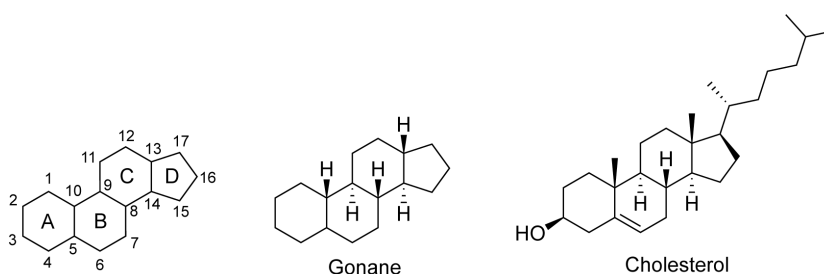


Figure 2: Structure of steroids and cholesterol.

The simplest steroid is gonane. The numbering of the carbon atoms of the steroids are indicated on the left. Moreover, the structure of the first discovered sterol, cholesterol, is shown.

Incorporated into the membrane, the hydrophobic steroid backbone of cholesterol is oriented parallel to the acyl chains of the lipid molecules. The hydroxyl group of cholesterol faces the aqueous environment and can interact with the head groups of lipids.<sup>2</sup> The successive incorporation of cholesterol into biological membranes broadens, or at high concentrations even eliminates, the gel to liquid crystalline phase transition of lipids.<sup>3</sup> More specifically, cholesterol induces the formation of *trans-gauche* conformers of the acyl chains of lipids. This prevents the clustering of lipid molecules at low temperatures, making the membrane more fluid. However, this makes the lipid bilayer stiffer at higher temperatures. In addition, the induction of *trans-gauche* conformers also increases the mechanical strength of the bilayer and generally reduces the permeability of the membrane to smaller ions and molecules.<sup>4</sup>

Since cholesterol can affect the biophysical properties of biological membranes, it can also modulate the function of integral membrane proteins which are embedded in the bilayer.<sup>5</sup> Microdomains that form in the membrane by clustering of cholesterol and sphingolipids were proposed which were termed “lipid rafts”.<sup>6</sup> It was proposed that lipid rafts can serve as hubs to initiate signal transduction cascades by selectively enriching a subset of proteins for example



receptors in these cholesterol rich patches.<sup>7</sup> However, their existence *in vivo* remains under debate.

Cholesterol levels can also vary in different parts of a cell and are not evenly distributed. The lowest cholesterol levels (<10 mol%) have been described in the endoplasmic reticulum and Golgi apparatus.<sup>8,9</sup> Higher levels of cholesterol have been found in the plasma membrane of cells ranging from 10 to 30 mol%.<sup>10</sup> The highest amount of cholesterol with up to 50 mol% is found in the plasma membrane of red blood cells.<sup>11</sup>

Cholesterol itself not only has unique biophysical properties that are beneficial to living cells, but it can also serve as a precursor molecule. Processed cholesterol derivatives include steroid hormones such as testosterone, estrogen, and cortisol, as well as oxysterols, which serve as important by-products and can further regulate lipid metabolism.<sup>12</sup> Bile acids, another product of cholesterol metabolism, are necessary for the digestion and absorption of dietary fats in the digestive tract. In addition, cholesterol is required for the synthesis of vitamins such as vitamin D.<sup>13</sup>

Cholesterol also represents an important post-translational modification. In two instances, covalent post-translational modifications of proteins by cholesterol were found to be crucial for their biological activity. Sonic hedgehog (Shh) signaling is a determinant of vertebrate morphogenesis and limb development.<sup>14</sup> C-terminal cholesteroylation of Shh is required for long range signaling where mutants that are unable to be modified were not capable of creating an extracellular gradient of the morphogen.<sup>15,16</sup> Another protein involved in Hedgehog (Hh) mediated signaling, the G-protein coupled receptor smoothed (SMO), is inhibited by Patched 1 (Ptch1) in the absence of Hh. Once Hh is bound to Ptch1, repression of SMO is reversed and will activate a series of signaling events and downstream gene modulation.<sup>17</sup> SMO is also post-translationally modified by a cholesteryl entity at an internal residue which is crucial for proper Hh signaling.<sup>18</sup>

Giving the diverse functions of cholesterol for biological membranes as well as serving as precursor molecule a constant supply by *de novo* synthesis must be guaranteed for metazoan organisms.

### 5.3 Biosynthetic pathway of cholesterol

The majority of cholesterol (~80%) is synthesized *de novo* by the body itself from the simple building block acetyl-CoA, which mainly occurs in the endoplasmic reticulum (ER) membrane of the liver.<sup>19</sup> Interestingly, only a small percentage (~20%) of the total cholesterol in the human

body is taken up by the diet. Thus, the balance between *de novo* synthesis and uptake from the diet must be tightly controlled.

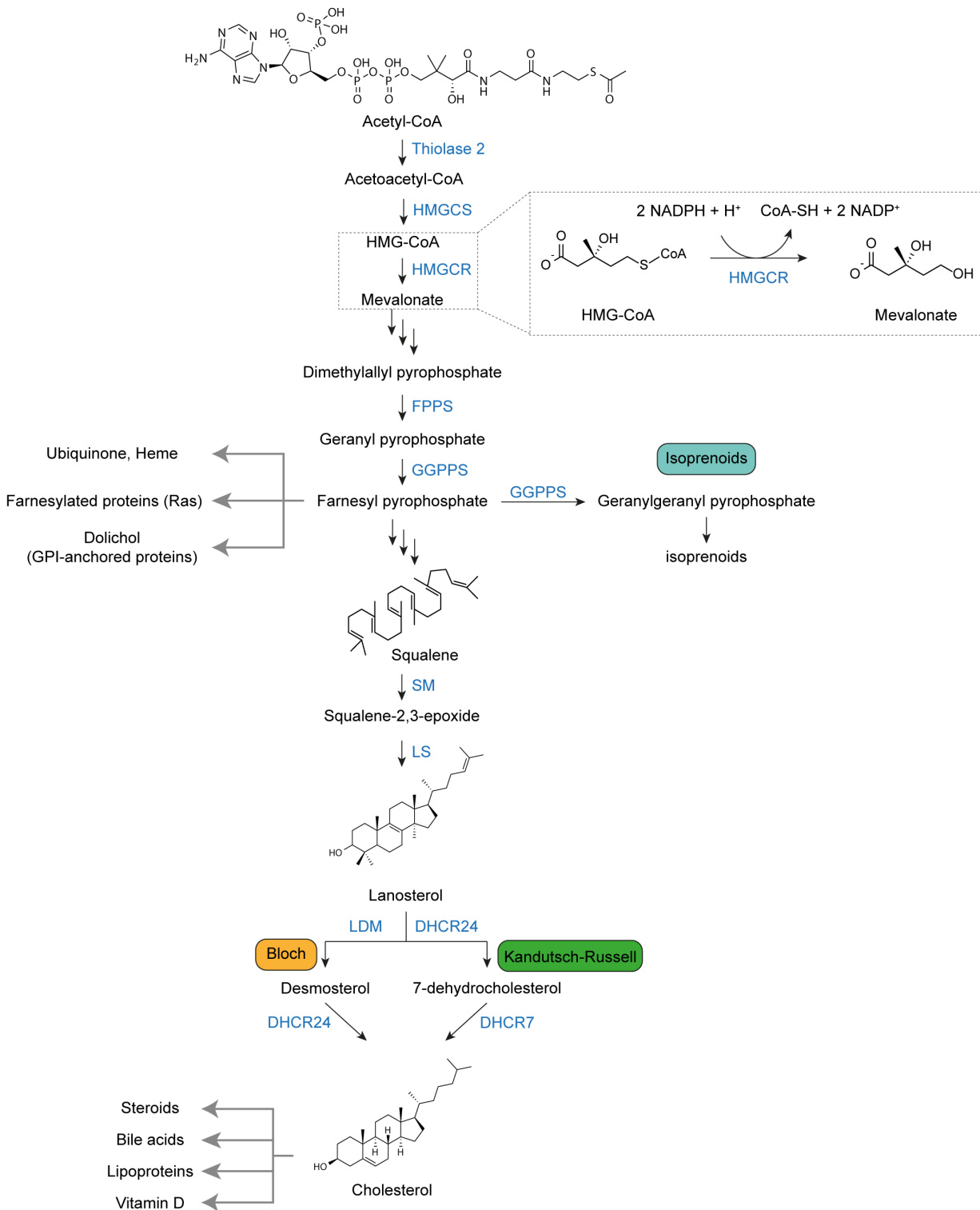


Figure 3: Cholesterol biosynthetic pathway.

Cholesterol is synthesized *de novo* from the water-soluble acetyl-CoA. Condensation reactions of three acetyl-CoA molecules lead to HMG-CoA that is the substrate of the HMGCR. The HMGCR catalyzes the conversion of HMG-CoA into mevalonate which is then further processed in multiple enzymatic reactions to form geranyl pyrophosphate (GPP). Condensation of GPP and isopentyl-5-pyrophosphate (IPP) leads to farnesyl pyrophosphate. This serves as an important precursor molecule. Further condensations lead to generation of

lanosterol via epoxidation and cyclization of squalene. Lanosterol can be converted via the Bloch or Kandutsch-Russell pathways into cholesterol. Enzymes are depicted in blue. Multiple arrows indicate multiple reaction steps.

The biosynthesis of cholesterol involves nearly 30 reaction steps catalyzed by more than 20 enzymes that can be divided into two major parts (Figure 3). The first part is the mevalonate pathway that produces hydrophobic lanosterol from the water-soluble precursor acetyl-CoA. The second part is the conversion of lanosterol to cholesterol.

In the first step of cholesterol biosynthesis, acetoacetyl-CoA is formed by condensation of two molecules of acetyl-CoA. Acetyl-CoA stems from other metabolic routes such as glycolysis,  $\beta$ -oxidation of fatty acids or degradation of amino acids. Then, condensation of a third molecule of acetyl-CoA with acetoacetyl-CoA produces HMG-CoA which is catalyzed by the HMG-CoA synthase (HMGCS). As mentioned above, the reduction of HMG-CoA by the HMGCR as major rate-limiting steps of the whole cholesterol synthesis pathway leads to formation of mevalonate.<sup>20</sup> This step consumes two reductive equivalents of NADPH. A series of phosphorylations and decarboxylations generates isopentyl-5-pyrophosphate (IPP), a basic 5-carbon structure crucial to isoprenoid synthesis. IPP can isomerize to dimethylallyl pyrophosphate (DMAPP), which allows condensation with another IPP molecule to form geranyl pyrophosphate (GPP). This 10-carbon isoprenoid can react with another IPP molecule to yield the 15-carbon farnesyl pyrophosphate (FPP).

FPP is not only a precursor for the downstream cholesterol synthesis, but also serves to produce non-sterol isoprenoids required for normal cellular homeostasis. For example, FPP can be attached to small GTPases such as Ras proteins or G-proteins as post-translational modification important for their cellular localization and thus their function. Other non-sterol products of FPP include dolichol, which is involved in the modification of glycoproteins, or ubiquinone which is a component of the electron transport chain.

The next steps in the cholesterol biosynthetic pathway involve the condensation of two FPP molecules leading to the formation of the 30-carbon molecule squalene catalyzed by the squalene synthase. Epoxidation and complex cyclization reaction as well as leads to the formation of lanosterol. These steps are catalyzed by the squalene monooxygenase and squalene epoxidase.

For the conversion of lanosterol to cholesterol two pathways, namely the Bloch<sup>21</sup> and Kandutsch-Russell<sup>22</sup> pathway, were shown to be used in cells. Both pathways catalyze a series of intramolecular rearrangements, demethylations and reductions to ultimately form cholesterol. They both utilize the same set of enzymatic reactions but differ in their order of appearance. In principle, each substrate from the Bloch pathway can be converted into the

analogous substrate of the Kandutsch-Russell pathway by reducing the  $\Delta 24$  double bond. Tissues like testis or the adrenal gland were shown to exclusively use the Bloch pathway whereas in most other tissues the main route is the Kandutsch-Russell pathway.<sup>23</sup> The penultimate product in the Bloch pathway before cholesterol is desmosterol which is converted by DHCR24 into cholesterol. The penultimate product of the Kandutsch-Russell pathway is 7-dehydrocholesterol which is converted by the 7-dehydrocholesterol reductase (DHCR7) into cholesterol. Reduced cholesterol production can have serious consequences for humans. Mutations in the terminal enzyme DHCR7 lead to a reduction in cholesterol, provoking the phenotypic effects of the Smith-Lemli-Opitz syndrome (SLOS), such as microcephaly or mental retardation.<sup>24,25</sup>

To balance cholesterol uptake and synthesis, several feedback mechanisms control the activity and expression levels of the proteins involved. These feedback mechanisms occur at various layers including the regulation of transcription, translation, post-translational modifications and proteasomal degradation.<sup>20</sup> Ultimately, they culminate in the tight regulation of the key players 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase (HMGCR), which catalyzes the rate-limiting step of cholesterol synthesis<sup>20</sup>, the low-density lipoprotein (LDL) receptor (LDLR), which is the receptor for LDL particles that are transported in the bloodstream<sup>26</sup>, and the cholesterol 7 $\alpha$ -hydroxylase, which is an enzyme catalyzing the first catabolic step of cholesterol breakdown into primary bile acids or cholic acid that can be secreted via the gall bladder<sup>27</sup>.

### **5.4 HMGCR and its regulation**

The HMGCR is the main rate-limiting enzyme of the whole cholesterol biosynthesis cascade. Human HMGCR is an integral membrane protein consisting of 888 amino acids that can be divided into two principal domains. The N-terminal domain is a transmembrane domain (TMD) that results in anchoring of the protein to the ER membrane. The TMD consists of 8 TM helices separated by short and hydrophilic loop regions.<sup>28</sup> TM helices 2-6 form a sterol-sensing domain (SSD), a fold that is abundant in several other integral membrane proteins involved in the many aspects of the complex orchestration of cholesterol homeostasis, such as Niemann-Pick disease type C1 (NPC1), Patched, Dispatched, SREBP-cleavage activating protein (SCAP), DHCR7 and Smoothed.<sup>29,30</sup> However, while the SSD fold is quite conserved between different members of the family, their sequences deviate quite substantially from one another. For example, the SSD of HMGCR and SCAP share only 26% sequence identity and 49% sequence similarity. The large cytosolic domain is of roughly equal size as the TMD with approximately

50 kDa and is connected to TMD via an unstructured linker region consisting of approximately 120 residues. Its cytosolic C-terminal domain (EC 1.1.1.34) converts 3-hydroxy-3-methylglutaryl coenzyme A into mevalonate by using two NADPH molecules as reductive equivalents. Mevalonate is used for production of sterols as well as non-sterol isoprenoids such as farnesyl and geranylgeranyl. The TMD of the enzyme is dispensable in order to carry out the catalytic activity but is required for regulatory function in response to increase sterol levels in the membrane which will be discussed below.<sup>31</sup>

Several crystal structures of the cytosolic domain have been solved in different states with bound substrates, cofactors and inhibitors, but structures of the full-length protein still remain elusive.<sup>32-34</sup> The catalytic portion of the HMGCR was found as tetramer in the asymmetric unit of the crystals (Figure 4). In-solution experiments confirmed the formation of tetramers.<sup>33</sup> However, it is not clear if the full-length protein can also form a tetramer. The crystal structures reveal that the active site of the enzyme is formed by two protomers and no contacts from the dimer-dimer interface are required. Therefore, it is conceivable that the full-length protein may also exist as a dimer with the potential to oligomerize into a tetramer.

The catalytic activity of the C-terminal domain of the HMGCR can be modulated by posttranslational modification, specifically by phosphorylation of serine 872 (S872). Phosphorylation at this position attenuates the activity of the HMGCR.<sup>35,36</sup> The same effect can be seen by mutation of S872 to aspartate.<sup>35</sup> The degree of phosphorylation is controlled by interplay of AMP-activated protein kinase and the HMG-CoA phosphorylase.<sup>35</sup> The exact consequence of the phosphorylation onto the structure is not fully understood but most likely is that the negative charges of the phosphoryl group contribute to charge-repulsion with NADPH and thus reduces its affinity for the cofactor slowing down the reaction.<sup>33</sup>

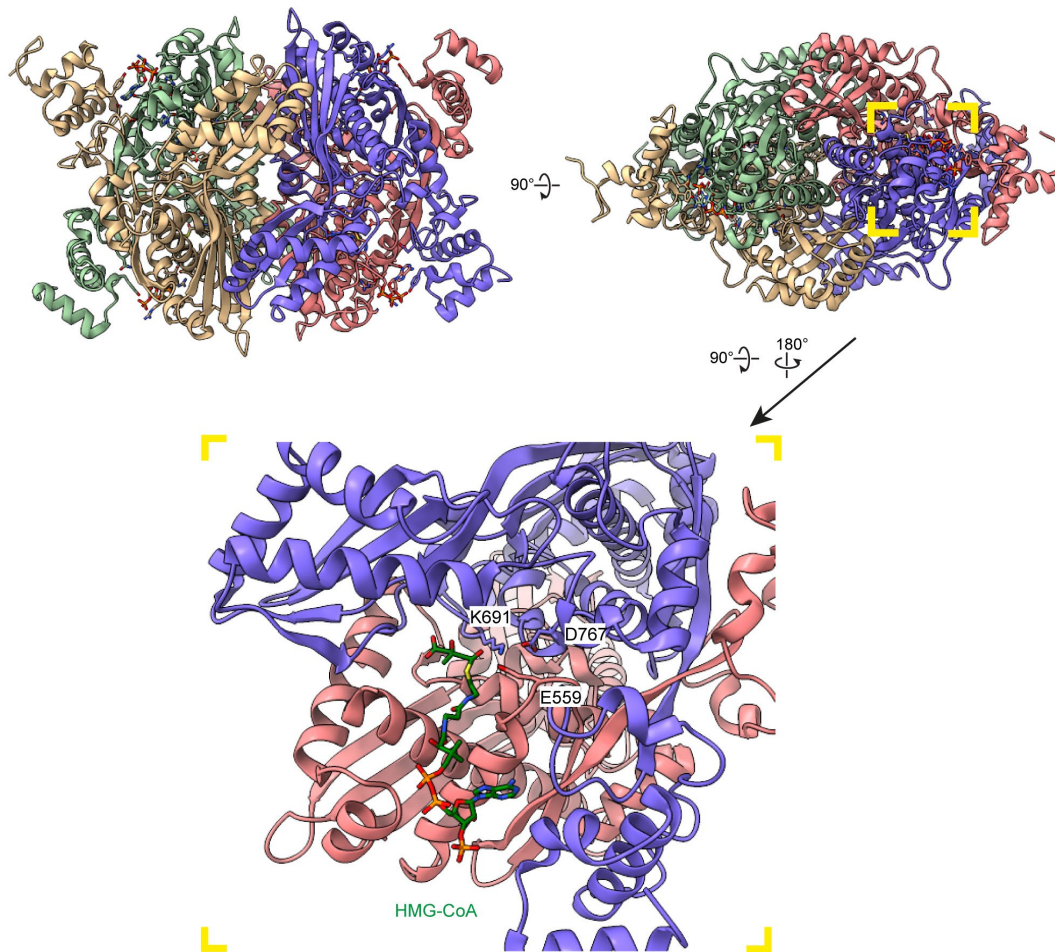


Figure 4: Structure of the catalytic portion of the HMGCR.

The catalytic portion of the HMGCR forms a tetramer in the asymmetric unit of the crystal as well as in solution. Each protomer of the tetrameric assembly is colored by chain. The yellow inset highlights the catalytic center formed by two protomers consisting of the residues lysine 691 (K691), aspartate 767 (D767), glutamate 559 (E559) and histidine 866 (H866). The latter is not resolved in the crystal structure. (PDB ID: 1DQ9)

## 5.5 Pharmaceutical inhibition of the HMGCR

A major risk factor for cardiovascular diseases (CVDs), such as atherosclerosis that can cause stroke or heart attack and account for more than 30% of all death worldwide, are elevated cholesterol levels.<sup>37</sup> Particularly, elevated levels of low-density lipoprotein (LDL) particles, often publicly referred to as “bad cholesterol” is of major concern for CVDs. LDL particles contain a hydrophobic core consisting of fatty acids and esterified cholesterol molecules which are solubilized by one copy apolipoprotein B (apoB). Typically, LDL particles are removed from the bloodstream by the presentation of the LDL receptor (LDLR) on the surface of hepatocytes, which recognize apoB and in turn endocytose LDL particles.<sup>38</sup>

Dysfunctional endothelial cells in the walls of arteries are thought to be initiators of the formation of atherosclerotic plaques.<sup>39</sup> LDL particles in the bloodstream can penetrate through these lesions and accumulate in the wall of the arteries. The proteinaceous component apoB and lipid molecules can be oxidized in this environment which is sensed by macrophages.<sup>40</sup> Large amounts of cholesteryl esters are ingested by the macrophages for their degradation. In case these macrophages take up more cholesterol than they can clear, cholesterol accumulates as lipid droplets in the cytoplasm and turns them into foam cells.<sup>41</sup> The foam cells that accumulate in the arterial plaque can secrete a variety of cytokines that will attract immune cells such as T-cells to this site. In turn, release of cytokines and accumulation of T-cells result in inflammatory reactions causing inducing smooth muscles cells surrounding this site to produce collagen in the arterial intima.<sup>42</sup> Collagen production leads to a sealing of the plaque from the bloodstream.<sup>43</sup> Until it ruptures, the plaque can stably build up and increase in size.<sup>44</sup> If the plaque eventually ruptures, blood clotting is initiated that seals the blood vessel causing stroke or heart attack.

Isolated fungal metabolites ML-236A and ML-236B (termed compactin<sup>45</sup>) were discovered in 1976 to competitively inhibit the HMGCR<sup>46,47</sup> *in vivo* and *in vitro*.<sup>48</sup> This molecule, marketed as Mevastatin, was the first statin administered to patients to inhibit the HMGCR. Its cholesterol-lowering effect is due to the inhibition of HMGCR, which blocks *de novo* cholesterol synthesis and in turn leads to the expression of LDLR, which as a result clears LDL particles from the bloodstream.

The discovery of cholesterol-lowering drugs laid the foundation for generation of a variety of molecules to pharmaceutically inhibit HMGCR (Figure 5). All statins share structural similarity to the natural substrate HMG-CoA. They contain an HMG-CoA-like part that is covalently attached to a hydrophobic rigid part. Type I statins are administered as prodrugs that are converted *in vivo* from the  $\delta$ -lactone form to the carboxylate form<sup>49,50</sup> which represents the active form of the compound that can bind and inhibit the HMGCR.<sup>46</sup> Type II statin inhibitors contain larger hydrophobic groups that are bound to the HMG-CoA-like part that is already in the active hydroxy-acid form.

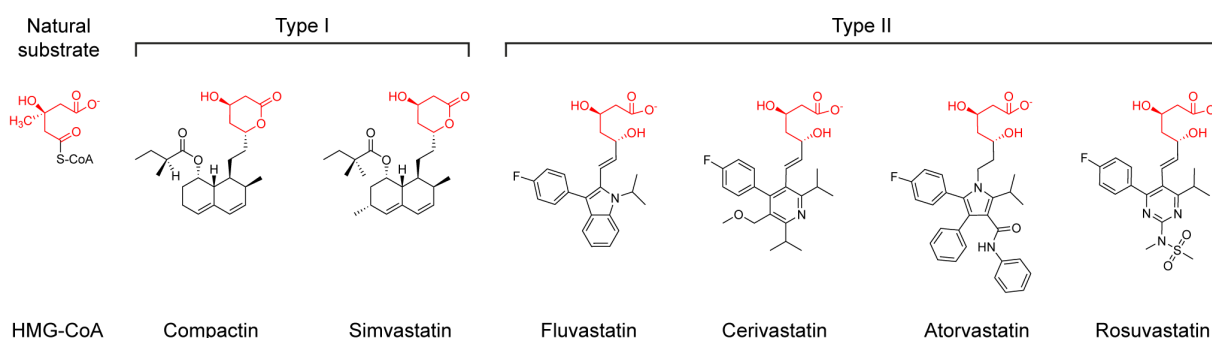


Figure 5: Structural formulas of the first statin in its next generation derivatives.

All statins share an HMG-CoA like moiety which is depicted in red and are competitive inhibitors. Compactin was the first statin that was isolated. Type I statins are in a prodrug form. The  $\delta$ -lactone ring will be converted to the active  $\beta$ -hydroxy acid form. Type II statins are administered in the active hydroxy-acid form. They contain larger hydrophobic groups compared to Type I statins.

In spite of the positive effects of statins in lowering LDL cholesterol levels in the blood, statins can also cause severe side effects. One of the most commonly observed side effects of statin treatment are myopathies.<sup>51</sup> Myalgia, a form of myopathy, is a benign side effect that occurs in 1-10% of the statin-treated patients. The most serious but rare adverse effect of statins is rhabdomyolysis, potentially causing acute renal failure and disseminated intravascular coagulation, which will ultimately lead to death. The approximate occurrence of rhabdomyolysis is 0.15 per 1 million statin prescribed patients.<sup>52</sup> Since statins inhibit one of the very early steps of cholesterol synthesis intermediate metabolites of the cholesterol biosynthetic pathway are also produced at a lower levels such as FPP or GPP. These intermediate building blocks are required for prenylation of i.e., small GTPases. The lipid modification is needed for cellular localization and proper activity of these enzymes. It was hypothesized that dysfunction of small GTPases is one of the causative effects of statin-induced myopathies.<sup>53</sup>

## 5.6 Feedback mechanisms to control the activity of the HMGCR

Patients that receive cholesterol-lowering drugs such as statins initially show strongly reduced blood LDL levels due to inhibition of *de novo* synthesis of cholesterol by blocking the HMGCR. However, shortly after they receive treatment by statins increased expression of HMGCR can be detected which attenuates the effect of statins. Sterols and non-sterol isoprenoids govern a multivalent feedback mechanism first detected in cultured fibroblasts.<sup>54</sup> Increased amounts of HMGCR up to ~200-fold after statin administration were then also discovered in mice<sup>55,56</sup>, rats<sup>57,58</sup> and humans<sup>59</sup>, weakening their effect.

Multiple feedback mechanisms converge to regulate the activity of HMGCR, which is essential for maintaining a steady supply of products and metabolites involved in cholesterol



biosynthesis. The accumulation of HMGCR in the endoplasmic reticulum (ER) membrane, induced by statin treatment, primarily arises from impaired ER-associated degradation (ERAD) of HMGCR. Studies have demonstrated that the addition of lanosterol, geranylgeraniol, 25-hydroxycholesterol (25-HC) and 24,25-dihydrolanosterol (24,25-DHL-) - all intermediates in the cholesterol synthetic pathway - to cultured skin fibroblasts enhances the degradation of HMGCR (Figure 6).<sup>54</sup>

The catalytic activity of the HMGCR is mediated solely by the cytosolic domain, independent of the transmembrane domain (TMD). However, the TMD of the HMGCR has been thought to be crucial in a process called sterol-accelerated degradation. In this process, intermediate products of the cholesterol synthetic pathway, which are highly hydrophobic, enrich in the membrane and trigger HMGCR degradation. This hypothesis was supported by the observation that C-terminal domain of HMGCR truncated by the TMD was found to be resistant to degradation in the presence of sterols.<sup>60</sup> Opposingly, an artificial construct was designed carrying the TMD of the HMGCR fused C-terminally to  $\beta$ -galactosidase to replace the cytosolic domain. This construct exhibited the same turnover as the full-length HMGCR.<sup>61</sup> Blocking the 26S proteasome by specific inhibitors led to an increased accumulation of the ubiquitinated form of the reductase in the ER membrane.<sup>62,63</sup>

The degradation machinery is recruited to the HMGCR in a sterol-dependent manner, especially after an increase in lanosterol in the membranes (Figure 6). Two proteins are crucial for the ubiquitination of the reductase, the Insulin-induced gene 1 (Insig1) protein and Insig2.<sup>64,65</sup> Insig binds in a sterol-dependent manner to the YIYF tetrapeptide motif to recognize the TMD of the HMGCR. This YIYF motif is also present in SCAP, another known binding partner of Insig, which will be discussed below. Insigs can bind to the known membrane-bound RING domain-containing E3 ubiquitin ligases called gp78<sup>66</sup>, Trc8<sup>67</sup> or RNF145<sup>68</sup>. Gp78 has binding motifs for Ufd1 that can increase the ubiquitination activity of gp78.<sup>69</sup> Moreover, gp78 has binding sites for UBE2G2<sup>69</sup>, a cytosolic E2 ligase, and VCP<sup>70</sup> (Valosin-containing protein), an AAA ATPase facilitating the extraction of ubiquitinated membrane proteins ultimately delivering the substrate to the proteasome. This complex associates with the HMGCR in a sterol-dependent manner to ubiquitinate the reductase at cytosolic lysine 89 and 248 (Figure 6).<sup>64</sup> This marks the reductase for proteasomal degradation via the ERAD, ultimately shutting down cholesterol *de novo* synthesis.

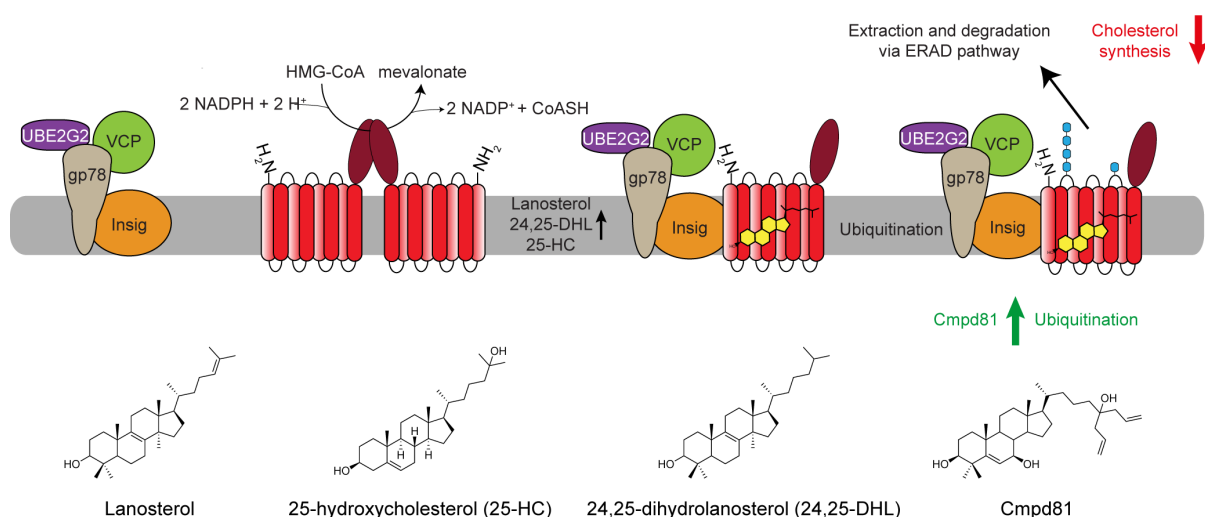


Figure 6: Sterol-accelerated degradation of the HMGCR via the ERAD pathway.

The HMGCR is integrated into the ER membrane converting HMG-CoA into mevalonate. Mevalonate is an important precursor molecule for the production of sterol and non-sterol molecules. A degradation machinery is recruited to the transmembrane domain of the HMGCR upon increasing concentrations of lanosterol, 25-HC and 24,25-DHL in the membranes. This will lead to polyubiquitination of the reductase at lysine residues 89 and 248 and degradation via the ERAD pathway. A synthetic compound Cmpd81<sup>71</sup> increases the ubiquitination of the HMGCR.

Pharmaceutically, this natural degradation mechanism may offer a platform to design drugs to promote reductase turnover and potentially overcome statin-induced accumulation of the enzyme. Natural products with high efficacy in inducing the degradation of HMGCR are the sterols 24,25-DHL or 25-HC. However, administration of these molecules as a direct treatment option is not a valid strategy because they normally enter the Bloch or Kandutsch-Russell pathway, which would mean their conversion to cholesterol. These molecules have, however, been used as a starting point for the generation of other oxysterol derivatives that do not enter the normal biosynthetic pathways but retain their ability to recruit the degradation machinery to the HMGCR. One such derivative, named Cmpd 81, has been synthesized and shown to induce effective degradation of HMGCR in cell culture experiments.<sup>71</sup> To test *in vivo* activity in mice, Cmpd81 was administered in combination with lovastatin. After 20 weeks of treatment, reduced serum LDL levels were observed in mice.<sup>71</sup>

Insigs are essential for sterol-induced degradation of the HMGCR. However, a protein that counteracts the destabilizing role governed by Insigs and ultimately protects the HMGCR from degradation is called UBIAD1 (UbiA prenyltransferase domain-containing protein 1). UBIAD1 is an integral membrane protein with 9 predicted TM helices with a predicted fold commonly expected for prenyltransferases. It catalyzes the geranylgeranylation of menadione using geranylgeranyl phosphate (GGPP) to form menaquinone-4 (MK-4), a derivative of

vitamin K<sub>2</sub>.<sup>72</sup> UBIAD1 exerts its stabilizing function by binding directly to the TMD of the HMGCR in conditions where cholesterol abundance is low and compete with binding of Insig (Figure 7).<sup>73,74</sup> However, when membrane levels of GGPP in particular rise due to the stabilization of reductase and increased production of this downstream product, the protective effect regulated by UBIAD1 is attenuated.<sup>73</sup> It is thought that GGPP binds to UBIAD1, resulting in dissociation from the TMD of the reductase. After dissociation, UBIAD1 translocates to the Golgi apparatus, thereby increasing ERAD mediated by Insig.<sup>75</sup> Thus, both proteins coordinate the stabilization and turnover of the reductase to ensure a continuous and homeostatic production of mevalonate.

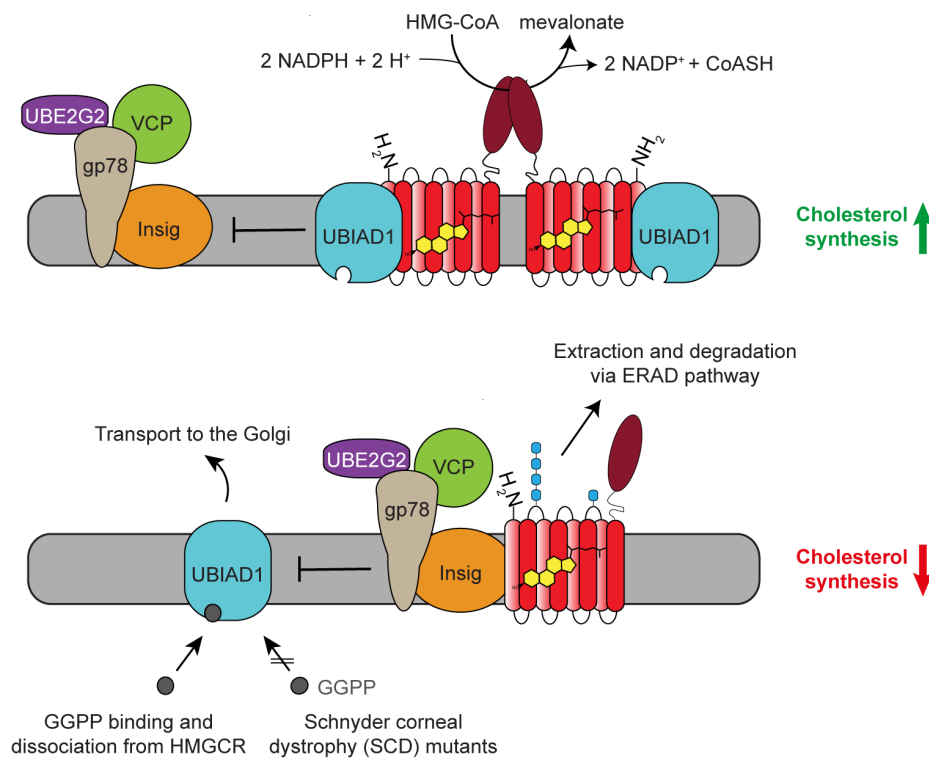


Figure 7: Model for the role of UBIAD1-mediated stabilization of the HMGCR.

During conditions of sterol depletion, UBIAD1 plays a crucial role in stabilizing HMGCR within the endoplasmic reticulum (ER) membrane. This stabilization is essential as it ensures the availability of mevalonate, a vital precursor for the synthesis of both sterol and non-sterol products. However, when the levels of the intermediate molecule geranylgeranylpyrophosphate (GGPP) increase in the cell membranes, GGPP binds to UBIAD1. As a result, UBIAD1 dissociates from the HMGCR, leading to its translocation from the ER to the Golgi apparatus. This translocation triggers the recruitment of the ER-associated degradation (ERAD) machinery to the transmembrane domain (TMD) of HMGCR, facilitated by Insig. Consequently, this recruitment of the ERAD machinery serves to blunt the cholesterol biosynthesis pathway.

Mutations in UBIAD1 have been found in patients suffering from Schnyder corneal dystrophy (SCD). SCD is an autosomal dominant disorder of the eye that manifests as corneal opacity due to the accumulation of cholesterol.<sup>76-78</sup> To date, 24 disease mutations in UBIAD1 gene have

been described that are linked to cholesterol accumulation in the eye.<sup>79,80</sup> This is not understood at the molecular level. However, the underlying rationale for these mutations is either the failure of UBIAD1 to bind to GGPP, which normally results in dissociation from the HMGCR, or the mutations result in increased binding affinity to the HMGCR. Both effects result in sequestration of UBIAD1 in the ER membrane and increased protection of the HMGCR, leading to overproduction of cholesterol.

Insig and UBIAD1 bind competitively to the TMD of HMGCR but under different conditions. When the abundance of cholesterol and midway precursors of the synthetic cascade is low, UBIAD1 preferentially binds to the TMD of HMGCR. This leads to a stabilization of the HMGCR and production of mevalonate which subsequently leads to synthesis of the former mentioned products. As their levels in the membrane increase, especially that of GGPP, this causes dissociation of UBIAD1 from the HMGCR. After synthesis of cholesterol and modified cholesterol derivatives such as oxysterols, Insig binding to the HMGCR is promoted, leading to degradation and shutdown of *de novo* synthesis of these molecules. In contrast, UBIAD1 requires substrates that are generated much earlier in the cholesterol synthesis pathway. This difference in binding requirements between Insigs and UBIAD1 provides the basis for a constant and finely regulated flow of mevalonate, which is essential for the production of both sterol and non-sterol products.

### **5.7 Structural studies on the HMGCR-UBIAD1 complex**

To understand regulation of HMGCR modulated by UBIAD1 and Insig at a molecular detail, structures of these proteins, ideally as complexes, are required. As described above, the crystal structure of the catalytic part of the HMGCR was solved more than 20 years ago in 2000. Crystal structures of the archaeal UbiA, UBIAD1 homologs have been determined in both apo- and substrate-bound states.<sup>81,82</sup> These structures elucidated the overall fold of the protein and identified residues critical for the catalytic function of the protein. Structures of the full-length human or mammalian proteins remained elusive.

In 2022, cryo-EM structures of the TMDs of HMGCR complexed with UBIAD were determined, revealing different oligomeric assemblies and conformational states (Figure 8A).<sup>83</sup> Several construct boundaries and stabilizing point mutations of the proteins were explored. Homogeneous preparations of the HMGCR-UBIAD1 complex were obtained with the hamster HMGCR by truncating the C-terminal domain and introducing two point mutations that render the protein resistant to ubiquitination (K89R, K248R). Hamster UBIAD1 was truncated at the first 40 N-terminal residues, which are thought to be flexible and cytosolic. In addition, a point

mutation N1202 was introduced, which attenuates the enzymatic activity of UBIAD1. First, a thermostabilized apocytochrome b562RIL (BRIL)<sup>84</sup> was introduced into the cytosolic loop between TM 8 and 9 of UBIAD as a fiducial for particle alignment during image processing. However, this approach resulted in a low resolution of the TM helices in the complex, likely due to flexibility between the fiducial marker and TMDs. Instead, a Fab fragment (Fab<sup>15B2</sup>) derived from mouse immunization and subsequent screening of hybridoma clones was used. The Fab fragment bound specifically to the complex and allowed the determination of multiple cryo-EM structures resolved to 3.3 Å resolution (Figure 8A).

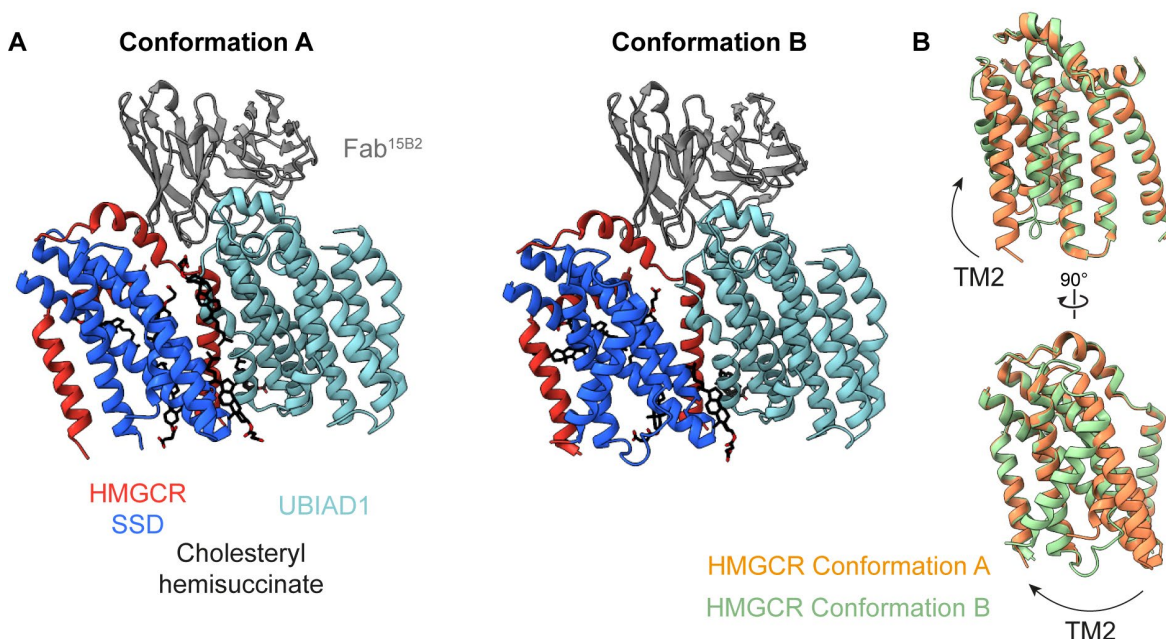


Figure 8: Structures of the HMGCR-UBIAD1 complex solved by cryo-EM.

(A) Two structures of the HMGCR-UBIAD1 complex were solved representing two distinct conformations of the HMGCR-TMD. The structures were solved with the Fab fragment Fab<sup>15B2</sup> to guide particles alignment. (B) Superimposition of the HMGCR-TMD shows the tilting of the TM2 of 45° from Conformation A to B.

The structures reveal the HMGCR-UBIAD complex exists as a monomeric and dimeric complex with stoichiometric copies of each protein. The Fab fragment is bound to the dimer interface of both TMDs. TMs 1-7 of HMGCR and all 9 TMs of UBIAD1 were resolved. The structure of UBIAD1 exhibits structural similarity to the previously reported archaeal UbiA structures with an RMSD of 3.1Å. The interface between the two proteins of the complex involves hydrophilic and hydrophobic interactions between TMs 5 and 7 of HMGCR and TMs 2 and 4 of UBIAD1. Mutations in several interface residues resulted in decreased complex formation and enhanced ERAD of HMGCR. Interestingly, several cholesteryl hemisuccinate (CHS) molecules were resolved at the interface of the complex and also bound to the TMD of the HMGCR (Figure 8A, black).

Comparing the structures of HMGCR SSD reveals the existence of two distinct conformations which are depicted as Conformation A and Conformation B.<sup>83</sup> In conformation A, TM2 is oriented perpendicular to the membrane. TM4 unwinds into two half-helices, S4a and S4b. Such a configuration of the S4 helix has been observed for Scap in a complex with Insig, which will be discussed below. In Conformation B, TM2 of the HMGCR undergoes a major conformational change, resulting in a 45° tilt of TM2 and an intact S4 helix that is not unwound (Figure 8B). Conformation B is further stabilized by interactions between the TM helices of the HMGCR. Considering that the crystal structure of the catalytic portion of the HMGCR was found as a tetramer both in the asymmetric unit of the crystal and in solution, it is tempting to speculate that the Conformation B of the SSD may stabilize an oligomeric form of the HMGCR. Based on the analysis of both SSD conformations of the HMGCR, the authors of the study, after comparing the SSD conformations with other proteins known to contain an SSD, proposed that Conformation A likely promotes binding to Insig, leading to ERAD of the HMGCR. Whereas in Conformation B, the SSD adopts a conformation that is resistant to bind Insig. The significant rotation of TM2 within the membrane from Conformation A to B is thought to potentially reorient the YIYF motif, required for interaction with Insig. The potential interface would be blocked in conformation B by the steric hindrance of residue F80 of the YIYF motif. However, structures of the HMGCR-Insig complex need to be determined in the future to identify the complex interface. Furthermore, how sterols regulate the different conformations of the SSD of the HMGCR remains to be elucidated.

While the modulation of reductase half-life in response to post-mevalonate products represents a significant regulatory mechanism, it is just one facet of its tight regulation. Transcriptional control of the genes required for cholesterol uptake and synthesis, among which the transcription of the *HMGCR* gene is controlled, plays a crucial role in maintaining cholesterol homeostasis.

### **5.8 Cholesterol homeostasis by the SREBP signaling pathway**

Cholesterol, taken up by the diet or synthesized endogenously, will cause end-product inhibition of biosynthetic pathways. The protein sensors that orchestrate cholesterol homeostasis are three proteins and are part of the sterol-regulatory element binding protein (SREBP) signaling pathway that constitute a negative-feedback loop: Scap, Insig and SREBP<sup>85,86</sup>.

SREBPs (SREBP-1a, -1c and 2) are synthesized as membrane bound transcription factors. They have a N-terminal domain that spans 480 amino acids and contains a basic-helix-loop-helix leucine zipper (bHLH-LZ) transcription factor. This N-terminal domain is connected to two transmembrane helices, bridged by a short 30 amino acid luminal loop. The C-terminal domain (CTD), which extends into the cytosol, comprises approximately 600 amino acids.<sup>87,88</sup> The CTD of SREBPs interact with the C-terminal WD40 domain of the SREBP cleavage-activating protein (Scap) and thus form a complex that resides in the ER membrane.<sup>89,90</sup>

If the abundance of cholesterol is low (<5 mol%)<sup>91</sup>, Scap facilitates the transport of SREBP from the endoplasmic reticulum (ER) to the Golgi apparatus through COPII vesicular transport (Figure 9). In low cholesterol regimens, Scap is in a presumed open conformation, exposing a cytosolic loop that contains a MELADL sequence motif. This motif can be recognized by the coat protein complex II (COPII) machinery.<sup>92</sup> Once at the Golgi, SREBP is activated by two sequential proteolytic cleavages. First, the luminal loop is cleaved by Site-1 protease (S1P), a serine protease, followed by cleavage in the TM-1 helix by Site-2 protease (S2P), a zinc protease located in the membrane.<sup>93–95</sup> These proteolytic events liberate the active N-terminal transcription factor that is no longer bound to the membrane. Consequently, the released N-terminal transcription factor domain of SREBP forms homodimers and is then imported into the nucleus with the help of importin proteins.<sup>96,97</sup> In the nucleus, the active and dimerized transcription factor can initiate the transcription of genes encoding proteins responsible for cholesterol biosynthesis, such as the *HMGCR* gene, and cholesterol uptake, for example the *LDLR* gene.<sup>98</sup>

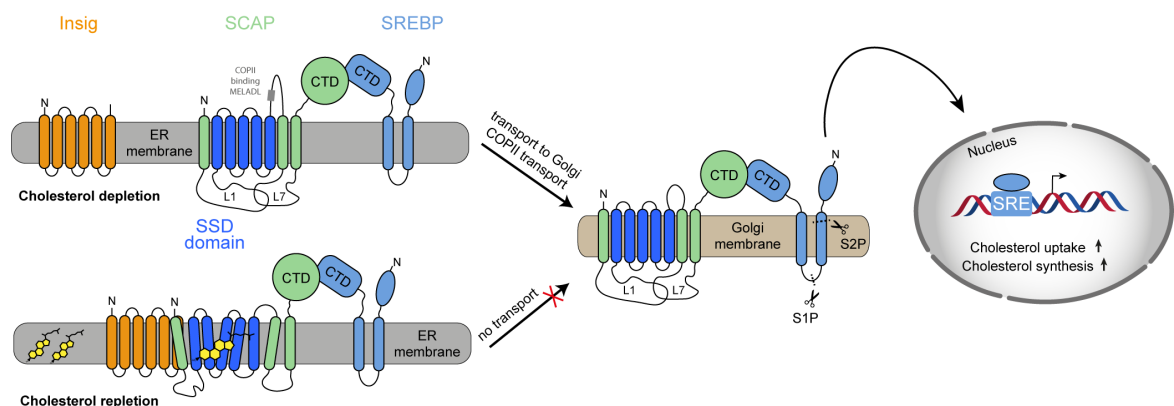


Figure 9: Control of cholesterol homeostasis by the SREBP signaling pathway.

After genes have been activated to increase cholesterol synthesis or uptake, these products cause end-product inhibition. Once the levels of cholesterol rise in the membranes (>5 mol%), cholesterol can bind to the SSD of Scap causing a conformational change that masks the COPII

recognition sequence.<sup>99</sup> Consequently, this change promotes binding to either Insig1 or Insig2 leading to a retention of the ternary complex in the ER (Figure 9) which in turn shuts down cholesterol synthesis and uptake.<sup>100–102</sup> Insig binding to Scap is triggered by oxygenated sterol derivatives such as 25-HC and not by cholesterol whereas Scap cannot bind oxysterols.<sup>104</sup> This dual requirement for both cholesterol and its derivatives emphasizes their crucial role in regulating cholesterol homeostasis and emphasizes the need for high-resolution structures of these proteins in order to understand their complex regulation at a molecular level.

## 5.9 Structural studies of the SCAP-Insig complex

Scap is an eight-pass TM protein with helices 2-6 constituting an SSD, a domain also found in other proteins that are known to be involved in the regulation of cholesterol homeostasis. Scap contains three large loops, one of which faces the cytosolic site and contains the crucial MELADL sequence responsible for COPII-mediated trafficking. Besides, SCAP possesses two large luminal loops (L1 and L7) that exert a regulatory function. L1 can bind to L7 in conditions where cholesterol abundance is low<sup>103</sup> referred as open and active conformation that can escort SREBP.<sup>104</sup> This interaction is disrupted if cholesterol levels rise due to the observation that L1 was measured to bind cholesterol which might force Scap to adopt an inactive and closed conformation.<sup>105</sup> Two mutations in L1 (Y234A) and L7 (Y640S) were found to disrupt their mutual interaction and stabilize Scap in a cholesterol-bound conformation also when sterols are not present.<sup>104,106</sup> In addition, the ability of Scap to bind to Insig can be artificially modulated by a mutation in the SSD. This mutation replaces aspartate 428 with alanine and displays binding to Insig without the requirement of sterols.<sup>107</sup>

Despite the discovery of this pathway more than 20 years ago, structural characterization of the involved proteins has remained elusive for a long time. However, structural studies have successfully elucidated homologous structure. For instance, that of a mycobacterial Insig homolog MvINS.<sup>108</sup> In addition, crystal structures and cryo-EM structures of the CTDs of fission yeast SREBP and Scap have been solved.<sup>109,110</sup> The transmembrane domain truncated CTDs of Scap and SREBP homologs were expressed and purified as complexes. The cryo-EM structure reveals that the CTDs form a higher oligomer of dimer of dimers with a stoichiometric ratio of 4:4 (Figure S 1). However, it remains elusive whether the full-length complexes anchored to the membrane can also adopt this oligomeric assembly.

A number of factors have contributed to the successful determination of multiple structures of the Scap-Insig complex as well as Scap alone.<sup>111–113</sup> These advances include the optimization of expression and purification protocols for eukaryotic membrane proteins, the improvements



in software and generation of image processing workflows, and the use of specific binders that act as fiducials during image analysis. Structures were determined in the cholesterol-free and bound forms of both the Scap-Insig complex and Scap alone. These structures provide a detailed understanding of Scap-mediated regulation to maintain cholesterol homeostasis.

The first structure of the human Scap-Insig complex was determined in the beginning of 2021 to an average resolution of 3.7 Å (Figure 10).<sup>111</sup> The Scap construct was truncated by the flexible WD40 and coexpressed with a modified version of full-length human Insig2. The oxysterol 25-HC was added to all steps, including expression and purification, necessary to maintain a stable complex association. This pure and stable complex was subjected to cryo-EM SPA analysis. In the complex structure, both TMDs of Scap and Insig, with the exception of one TM helix of Scap, were resolved at local resolutions ranging from 3.3 to 3.9 Å. The domain formed by the luminal loops L1 and L7 was resolved at a lower resolution, preventing the faithful construction of an atomic model. This domain was observed as a globular domain protruding from the detergent micelle.

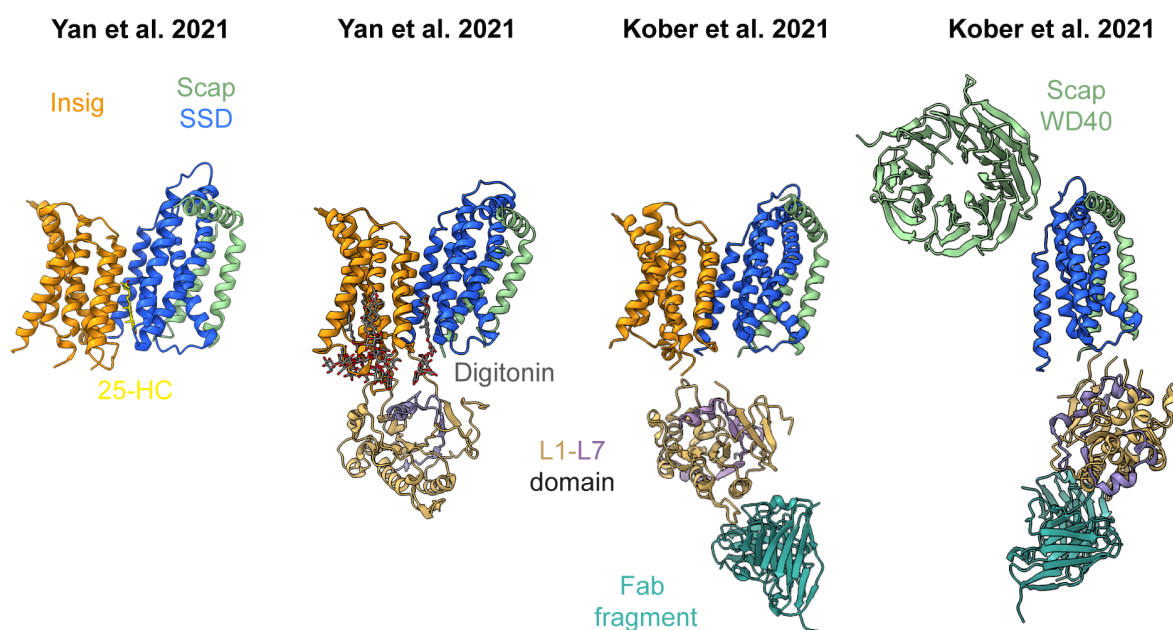


Figure 10: Cryo-EM structures of Scap and Scap-Insig complex reveal their molecular architecture.

Several cryo-EM structures of Scap and the Scap-Insig complex were determined in 2021. The first structure (PDB ID: 6M49) revealed a stoichiometric 1:1 complex between Scap and Insig, with a 25-hydroxycholesterol (25-HC) molecule located at their interaction interface. In the second structure (PDB ID: 7ETW), the luminal domain of Scap was resolved, consisting of intertwined loops L1 and L7. In addition, digitonin molecules were identified and shown in gray. Later in 2021, further structures of the Scap-Insig complex and of Scap alone were solved. A specific Fab fragment (shown in green) guided the particle alignment during image processing. The resolution of the latter reconstructions was modest, allowing only rigid-body fitting of the TMDs, while the L1-L7 domain was resolved to a higher resolution of 2.9 Å for Scap-Insig and 3.5 Å for Scap alone (PDB ID: 7LKF & 7LKH). Scap

has been kept in the same orientation in this figure, which illustrates conformational changes of Scap and the L1-L7 domain.

The structure reveals that both proteins form a complex between their TMDs (eight helices of Scap and six helices of Insig). Notably, an elliptical density is sandwiched between the S4-S6 helices of the SSD of Scap and the TM helices 3 and 4 of Insig2, which has been assigned to a 25-HC molecule. The S4 helix of Scap is discontinuous and unwound at its midpoint, which are depicted as the half-helices S4a and S4b. The unwinding of the S4 helix induces a tilt in the S4a segment, exposing hydrophilic residues to the hydrophobic environment of the membrane at the kink. This unfavorable conformation provides a binding interface for Insig2. The oxysterol 25-HC is accommodated by the kinked S4a helix and Insig. Moreover, it has been proposed that the energetically unfavorable process of S4 unwinding is compensated by interaction of the 25-hydroxy group of 25-HC with hydrophilic residues at kink of the S4 helix. Certain mutations (D428A) in Scap promote the interaction with Insig even without the requirement of sterols or oxysterols. This residue project towards the kink in the unwound S4 helix in the Insig-bound state. Introducing a more hydrophobic residue at this position, such as alanine, can potentially lower the energy barrier for unwinding of the S4 helix into two half-helices for the mutant D428A-Scap, which is normally overcome by 25-HC for the wild-type proteins.

Shortly thereafter, another cryo-EM structure of the Scap-Insig complex was determined. This structure also resolved at high resolution the luminal domain formed by L1 and L7 of Scap. (Figure 10).<sup>112</sup> The D428A mutation was introduced into Scap resulting in both TMDs to adopting the same conformation as in the previously solved structure of WT Scap-Insig. The complex was isolated in the steroidal detergent digitonin. Multiple copies of digitonin were observed bound to the Scap-Insig interface as well as to the TMD of Insig (Figure 10). However, no sterol molecule was found bound to Scap other than the one present at the interface.

The luminal loops L1 and L7 form a tightly intertwined globular domain that is oriented parallel to the membrane and interacts with the L1 loop of Insig. This domain shows high similarity to the extracellular domain (ECD) of NPC1, a known protein involved in the transport of LDL-derived cholesterol.<sup>114</sup> Similar structural features are also found in Ptch1, Dispatched, NPC1L1 (NPC1-like) and NCR1 (NPC1-related), whose activities are modulated by cholesterol or are involved in cholesterol uptake. These proteins feature an SSD-domain as part of their TMD and two extracellular domains.<sup>115-119</sup> However, this structure of the Scap-Insig complex does not explain the mechanism by which the L1-L7 luminal domain contributes to cholesterol homeostasis.

One month later, in 2021, further cryo-EM structures of the chicken orthologue WT Scap alone and the complex of a mutant Scap with Insig were reported (Figure 10).<sup>113</sup> The structure of the chicken Scap (cScap) alone in its cholesterol-free form was determined in GDN micelles assisted by the Fab fragment 4G10<sup>Fab</sup> which guided image processing. The structure revealed that the Fab fragment specifically bound to the L1-L7 luminal domain which could be resolved to a local resolution of 2.9 Å. However, the TMD and the WD40 of cScap could only be resolved to a resolution 4.1 Å preventing *de novo* model building. Nevertheless, the maps obtained were of sufficient quality to build a backbone model for six resolved helices of the TMD and to fit a homology model of the WD40 domain to the observed density. It was shown that purified cScap retained specific and saturable binding to cholesterol, but not to 25-HC or lanosterol. This indicates that cScap has been purified in a physiologically relevant form and demonstrates its specificity for cholesterol over its derivatives.<sup>120</sup>

The architecture of the L1-L7 domain of Scap exhibits a close resemblance to the previously determined structure with an RMSD of 1.3 Å. However, the higher resolution of the reconstruction visualized more features leading to the identification of a disulfide bond formed between C147 and C169 of L1 loop. When these cysteines were mutated, it resulted in diminished Scap-dependent SREBP2 processing in cells suggesting their relevance for Scap function.

The complex structure of chicken Scap-Insig1 was resolved albeit at low resolution for the TMD at 4.6 Å which permitted *de novo* model building. Again, a backbone model of the TMDs for Scap and Insig was placed into the density. The backbone model demonstrates good agreement with the structure previously solved of the Scap-Insig2<sup>111</sup> complex with an (RMSD) of 2.8 Å. TM7 and TM8 of cScap could not be resolved indicating their flexibility. Moreover, the WD40 domain of cScap also remained unresolved. The domain consisting of L1-L7 was resolved in this structure to a resolution of 3.5 Å, which had a similar conformation to the wt cScap alone with an RMSD of 0.77 Å.

Comparing the structures of cScap alone and in complex with Insig reveals significant differences in conformation of cScap. In the structure of cScap alone the density for TM2 is not observable, as opposed to cScap in the complex with Insig where the TM7 and TM8 are unresolved. In the structure of cScap alone TM8 is tightly packed against the SSD bringing the WD40 domain near the SSD (Figure 11A). Whereas in the complex structure with Insig this intramolecular interaction of the TM8 with the SSD is blocked by interaction with Insig. This potentially might explain why the WD40 is resolved in the cScap structure alone but not in complex with Insig.

Another significant change is observed in the position of the L1-L7 luminal domain. A comparison between the cScap structure alone and the complex structure with Insig reveals that the L1-L7 domain undergoes a rotation of approximately  $215^\circ$  parallel to the membrane. As a result, there is a shift in its positioning from below the SSD in cScap alone to the opposite side of the complex interface where Insig is bound (Figure 11B). This relocation positions the L1-L7 domain below the lipid bilayer suggesting its reorientation upon binding of Insig.

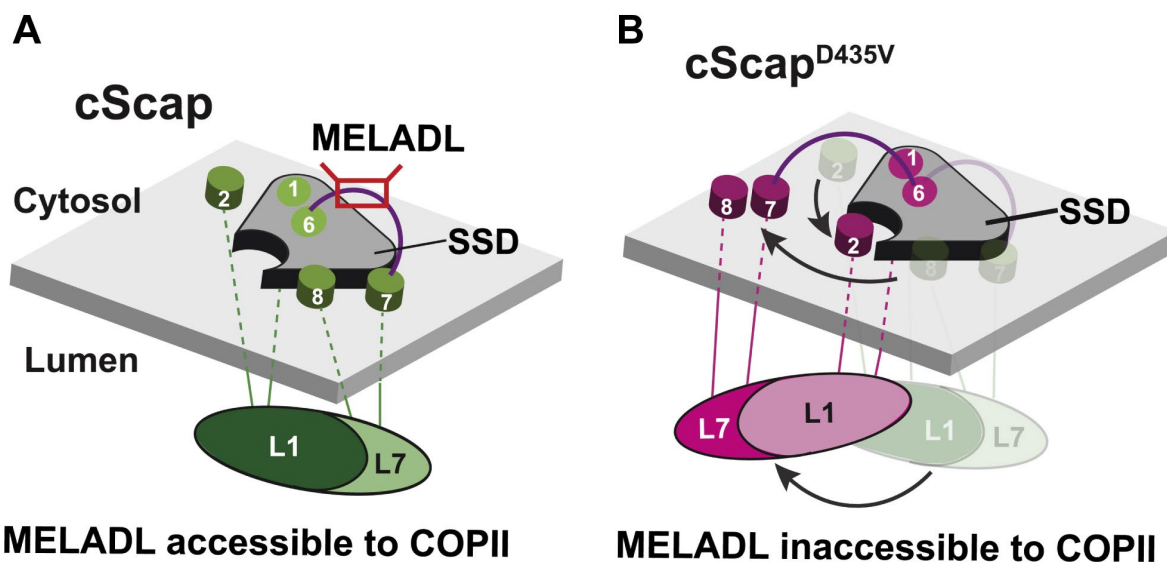


Figure 11: Schematic model of the cScap conformational changes upon Insig binding.

(A) cScap not in complex with Insig adopts a conformation where the L1-L7 domain is directly positioned below the SSD. The MELADL sequence between the loop that connects TM 6 and TM7 is exposed and can be recognized by the COPII machinery. (B) Upon Insig binding, the L1-L7 domain undergoes a huge conformational change and is positioned below the lipid bilayer and not below the complex interface. This brings TM2 into a position to create and interaction interface with Insig. Insig is not shown in this model. Additionally, the TM7 and TM8 are moved away from the SSD. The rearrangement results in a concealment of the MELADL motif preventing recognition of Scap by the COPII machinery. This figure has been adapted and slightly modified from ref.<sup>113</sup> and reprinted with permission from Elsevier.

The concerted conformational change and rearrangement of the TMD of Scap upon Insig binding along with the repositioning of the L1-L7 domain potentially results in a concealment of the MELADL motif. Consequently, Scap cannot be recognized by the COPII machinery for transport of SREBP to the Golgi apparatus shutting down transcription of genes required for cholesterol uptake and biosynthesis. Interestingly, it was not found that the interaction between L1 and L7 is disrupted during the conformational change as previously proposed.

Despite demonstrating saturable binding of radiolabeled cholesterol to Scap *in vitro* with expecting affinity and specificity, none of the determined structures have visualized the binding of cholesterol. Further research is needed to unveil the precise location and mechanism by which cholesterol binds to Scap, shedding light on the molecular basis of cholesterol

recognition and regulation by Scap. Such findings would provide valuable insights into the intricate processes of cholesterol homeostasis and advance the understanding of Scap-mediated cholesterol metabolism. Ultimately, this knowledge could serve as a foundation for design of targeted drugs that modulate the activity of these proteins with the goal to treat cholesterol-related disorders.

## 5.10 Bacterial secretion systems

Bacteria interact with their environment by secretion of a plethora of molecules, including small molecules, DNA, and proteins into the surroundings. Among these molecules, secreted proteins can be essential for fulfilling functions required for normal bacterial growth. But secreted proteins can also be toxic effectors that exert an inhibiting effect on the proliferation of competing bacterial cells in a polymicrobial environment which confers a fitness advantage in their biological niche.

More general secretion routes are governed by the general secretion (Sec) and twin-arginine translocation (Tat) pathway, which is found in eukarya, prokarya and archaea.<sup>121</sup> Client proteins of the Sec pathway contain cleavable N-terminal signal sequences that are usually removed after their transport. The transport of an unfolded nascent polypeptide chain is facilitated by SecYEG translocon in bacteria which is located in the inner membrane.<sup>122,123</sup> Contrary, the Tat pathway is responsible for the transport of folded proteins which do not contain specific sequences but rather patches of clustered arginine often referred as twin-arginine motif.<sup>124</sup> These pathways are typically used to transport cargoes to the periplasm of bacteria and in some instances in a second step into the extracellular space. In general, this transport is not very targeted and is used for the transport of a broad range of proteins. The direct transport of toxic effectors into competing bacteria therefore requires the use of more specialized secretion apparatuses.

For this purpose, Gram-negative bacteria have developed specialized nanomachines which are megadalton-sized and are called bacterial secretion systems (SS). They are used to transport molecules and virulence factors across the inner and outer membrane of the host cell into the environment or even directly into the cytosol of a prokaryotic or eukaryotic recipient cell. Typically, these systems are anchored to the IM and OM, but in some cases they may be located only in the OM. (Figure 12). Some can even cross a third membrane which is in this case the membrane of the recipient cell. They are classified from type I to type X secretion systems (T1SS to T10SS) based on the order of their discovery. Each of these SSs features differences in size, composition, and architecture. In addition, each secretion system differs in whether it

transports cargo in a folded, partially folded, or unfolded state. One such system that delivers folded effector proteins into target cells is the Type VI secretion system (T6SS). The T6SS is mainly used for competition among bacteria but can also deliver in some cases effector proteins into eukaryotic cells.<sup>125</sup>

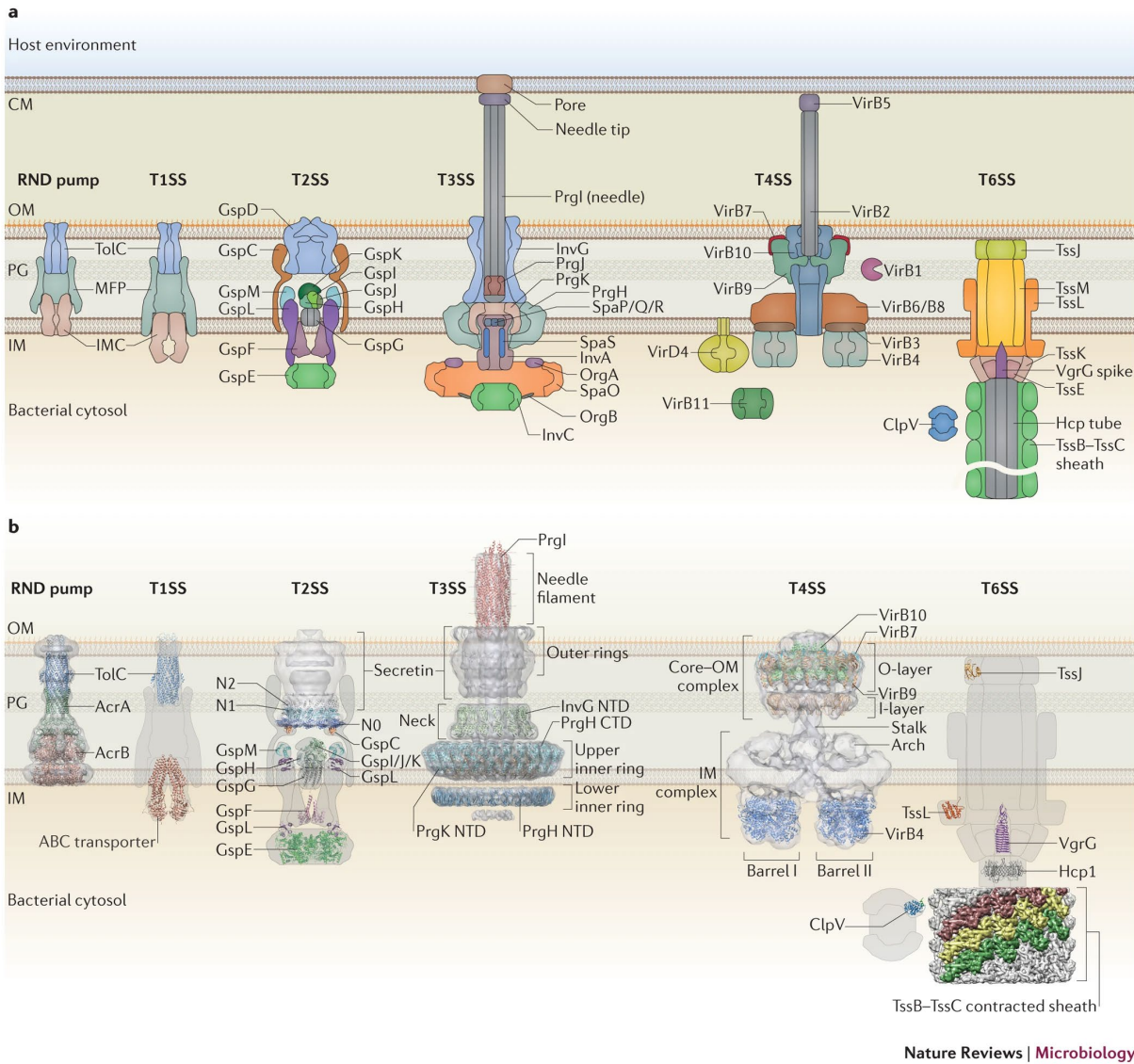


Figure 12: Schematic and structural representation of secretion systems in Gram-negative bacteria.

(A) Compositional organization of the bacterial secretion systems. These systems are anchored in the outer (OM) and inner membrane (IM) for transport of molecules across both membranes. Some of them such as the T3SS, T4SS and after a fire event also the T6SS, can also contact the membrane of the host cell. (B) Atomic models solved from crystal structures of individual components were docked into low-resolution cryo-EM maps, if available. This figure has been adapted from ref.<sup>126</sup> and reprinted with permission from Springer Nature.

### 5.11 The Type VI secretion system

The T6SS has been first discovered in the human-pathogenic bacterial strains *Pseudomonas aeruginosa* and *Vibrio cholerae* in 2006.<sup>127,128</sup> In the following years the list of bacteria that

poses at least one T6SS increasingly enlarged and became evident that T6SSs is not exclusive to human pathogens but can be found in a variety of environmental bacteria.<sup>129</sup> In 2010, the first antibacterial T6SS was described in *P. aeruginosa*<sup>130</sup> and shortly after also in *V. cholerae* and *Serratia marcescens*.<sup>131,132</sup> Guided by bioinformatic analysis, it has been identified that most of environmental bacteria as well a human pathogenic bacteria harbor at least one locus encoding all components for a functional T6SS.<sup>129</sup> Such a T6SS locus encodes for minimally 14 individual proteins that are required to assemble a functional core of the secretion system. These core subunits are highly conserved among all bacteria.<sup>129</sup> In most cases, additional accessory proteins are encoded within these gene clusters. Such genes are mostly not conserved between species and range from 15 to 20 genes in total.

When cells expressing a functional T6SS are visualized by cryo-electron tomography (cryo-ET), the T6SS appears as long and tubular structure in the cytosol of bacterial cells. The tubular structures correspond to the T6SS which can exist in two distinct states: an extended state with an approximate length of 450 nm and a contracted state with an approximate length of 280 nm (Figure 13A & B).<sup>133</sup> The ability of T6SSs to contract is reminiscent of the contractile behavior observed in bacteriophages. Not only the mechanism is conserved, but also the high degree of structural similarity of individual components to homologs of contractile bacteriophages.<sup>134</sup> The core subunits of the T6SS form the basic apparatus, which can be divided into three subcomplexes: a membrane complex spanning the cell envelope that serves as platform for assembly of the cytoplasmic baseplate complex and the elongated sheath-tube complex that bears similarity to the contractile sheath of bacteriophages (Figure 13C).<sup>133,135</sup> Multiple copies of the hexameric and ring-shaped hemolysin co-regulated protein (Hcp) stack on top of each other to form a large hollow tube which can project several hundred nanometers into the cytosol of the bacterial cell.<sup>136,137</sup> This tube is surrounded by multiple copies of VipA and VipB to complete the sheath-complex. The end of this tube that will eventually puncture the recipient cell is decorated with the homotrimeric valine-glycine repeat protein G (VgrG). The tip of VgrG typically associates with proteins of the proline-alanine-alanine-arginine (PAAR) family also known as PAAR-domain containing effectors.<sup>138</sup> The T6SS is not constitutively active but can dynamically and rapidly assemble upon certain triggers such as direct cell-cell contact with neighboring bacteria. This contact drives the assembly and contraction of the system and propels the sharpened tip, formed by the VgrG-PAAR spike complex, forward into the recipient cell.<sup>139</sup>



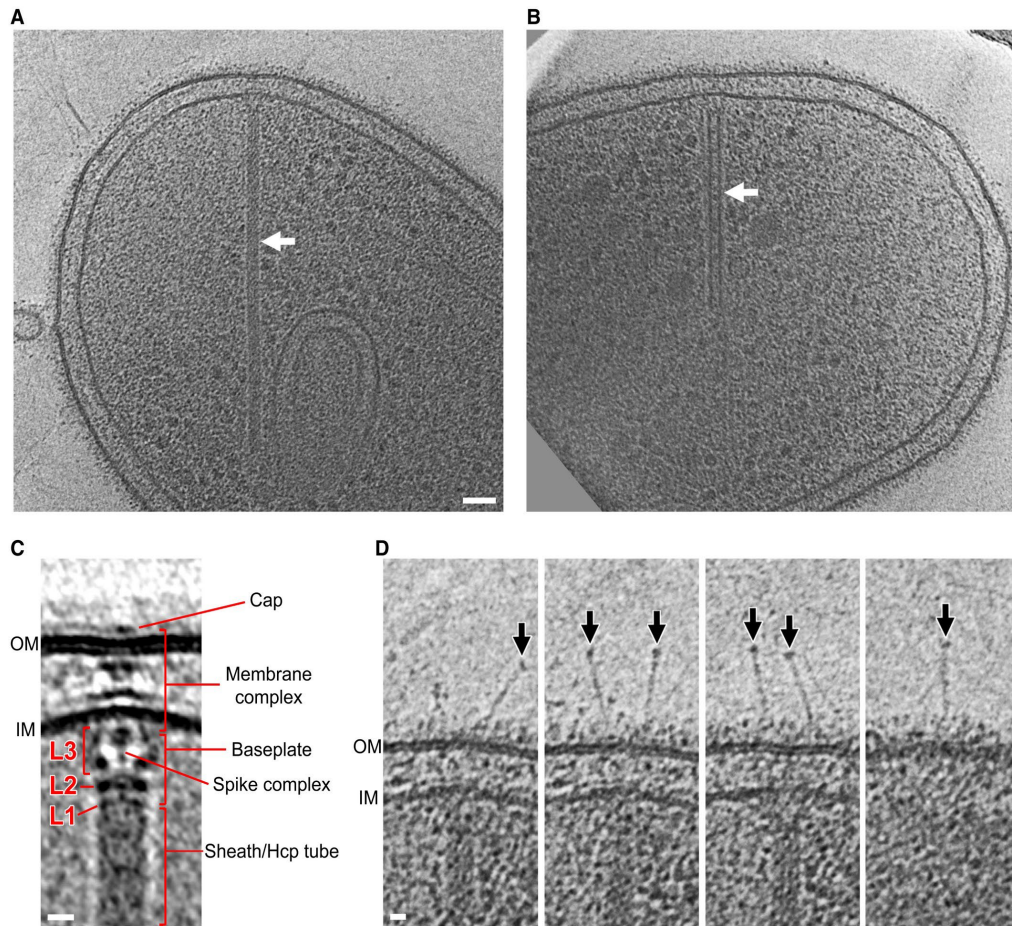


Figure 13: Tomographic slices of an intact T6SS from *Myxococcus xanthus*.

The assembled T6SS was found in a (A) extended and (B) contracted form. The arrow indicates the large tube formed by T6SS components. Scale bar 50 nm. (C) Central slice of an extended T6SS sub-tomogram average. Individual components are labelled. Scale bar 10 nm. (D) Tomographic slice showing extracellular bacteriophage tail fiber-like antennae indicated by the arrows. Scale bar 10 nm. The figure has been adapted from ref.<sup>133</sup> under license CC BY 4.0.<sup>140</sup>

Along with Hcp and VgrG, toxic effectors located on the tip of the tube-spike complex are injected into recipient cells and typically cause the attacked cell to die. The T6SS apparatus gene clusters also encode proteins that confer self-protection but also protect sister bacterial cells in a cell population from the effectors. These proteins are called immunity proteins and are co-expressed with the effector or produced in sister bacterial cells. The mechanism of protection occurs either via direct binding of the cognate immunity protein to the toxic part of the effector thus neutralizing it or by converting toxic products generated by effectors into harmless substances.<sup>141,142</sup> The toxic effectors themselves exhibit diverse activities such as functioning as nucleases, metallopeptidases, ADP-ribosyltransferases or deaminases.<sup>143–146</sup>

On the basis of their size and the presence of distinct domains, the effectors are recruited to different positions of the T6SS tip. Effectors typically smaller than 50 kDa can be



accommodated within the hollow Hcp tube that is approximately 40Å in diameter.<sup>147</sup> Oppositely, larger effectors typically contain additional domains that allow them to localize to the tip of the secretion system and interact with VgrG directly or mediated via adaptor proteins. The phospholipase Tle1 from the enteroaggregative *Escherichia coli* is recruited to the tip of the T6SS by direct interaction of the C-terminal helix of VgrG and the effector itself. Notably, Tle1 does not possess a PAAR domain.<sup>148,149</sup>

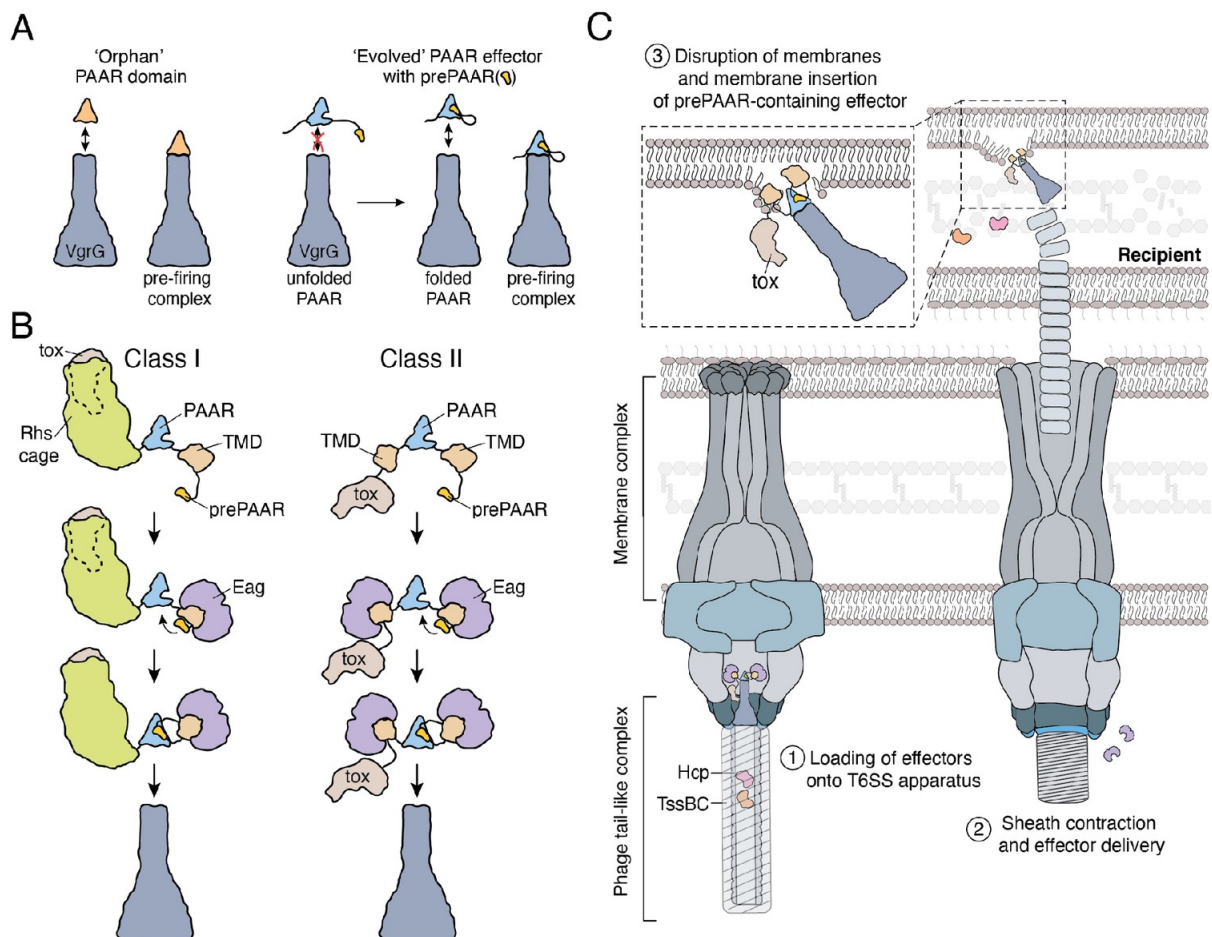


Figure 14: Model for assembly and secretion of PAAR-domain containing effectors.

(A) PAAR-domain containing effectors can be further classified based on presence of a prePAAR motif. Orphan PAAR effectors lack a prePAAR motif and directly interact with VgrG. PrePAAR effectors are multidomain effector domains. The prePAAR motif usually completes the fold of the PAAR domain. (B) These prePAAR effectors can be further divided into Class I and II. Class I effectors contain a single TMD and in most instances an Rhs cage which serves as encapsulation device of the actual effector. Class II effectors harbor two TMDs and contain other toxic effectors than an Rhs cage. (C) Proposed model for secretion of prePAAR-containing effectors into the recipient cell. The TMDs of the effector likely insert and disrupt the membrane of the recipient cell allowing effector delivery into the cytosol. The figure has been adapted from ref.<sup>150</sup> under license of CC BY 4.0.<sup>151</sup>

Effectors containing a PAAR domain are found in a wide range of bacteria. They can be further classified based on additional sequence motifs that are encoded within the polypeptide chain of the effector (Figure 14A-B). One subfamily of PAAR effectors, known as Class II effectors,

possesses an additional N-terminal motif which was defined as prePAAR. This domain was hypothesized to be required to complete the fold of the PAAR domain. In addition, class II effectors contain at least one N-terminal TMD with one to three TM helices. It is believed that these TM helices insert into the membrane of the recipient cell, thereby facilitating the translocation of the toxic effector i.e. from the periplasm into the cytosol (Figure 14C).<sup>150</sup>

To prevent premature insertion of the TM helices in the producing bacterium, the TMDs of the effector must be protected and solubilized. Three families of chaperones were identified which can bind to the TMDs of T6SS effector proteins, namely the DUF2169, DUF4123 and DUF1795 chaperone family. Chaperones of the DUF2169 and DUF4123 family bind their respective effectors as well as a separate protein containing a PAAR domain.<sup>152</sup> Whereas DUF1795 chaperones, termed effector-associated gene (Eag) chaperones, recognize and solubilize the TMDs of evolved PAAR effectors.<sup>153,154</sup> These chaperones associate with the effector TMDs before loading and secretion from the producing cell.<sup>150,153</sup> However, before the TMDs can insert into the membrane of the recipient cell, it requires the removal of the Eag chaperone from the TM helices. The TMDs of the effector Tse6 from *P. aeruginosa* are stabilized by binding to the cognate chaperone EagT6. When this complex was incubated with liposomal membranes *in vitro*, spontaneous insertion of the complex into the membrane was observed, suggesting the release of the protective chaperone.<sup>155</sup> At what point the chaperone is removed during this process or even during a firing event *in vivo* is not yet understood.

Substantial progress has been made in the structural characterization of various components of the T6SS including the *in situ* tomographic analysis of the assembled T6SS in the extended and contracted state.<sup>133</sup> Moreover, high resolution cryo-EM structures of the sheath and membrane-spanning complex have been determined.<sup>156,157</sup> In terms of effectors, a Class II effector called Tse6 from *P. aeruginosa* has been structurally characterized. This includes the elucidation of the architecture of complex composed of VgrG-Tse6-EagT6.<sup>155</sup> However, the structural characterization of a Class I effector remains elusive so far. Class I effectors were found to contain multiple copies of rearrangement hotspot (Rhs) elements originally discovered in *E. coli*.

### 5.12 Discovery of Rhs-repeat containing proteins

Two major sources of repetition were found in the genome of the *E. coli* K-12 strain: rRNA operons and Rhs elements.<sup>158</sup> The function of rRNA in bacteria is well understood, while the function of Rhs elements has remained enigmatic. A region in the K12 chromosome, the *glyS* gene, was found to be regularly duplicated in laboratory conditions. This gene is flanked by

two highly repetitive sequences termed Rhs regions *rhsA* and *rhsB*, because recombination frequently takes place between these sequences. Subsequently, more *rhs* elements were identified in *E. coli* numerically named in order of their discovery ranging from *rhsA* to *rhsH*. These elements share homology towards each other and are capable to efficiently recombine.<sup>159–161</sup>

In the genome of *E. coli*, the Rhs elements contain a conserved core region spanning 3.7 kilobase pairs (kbp). This core section typically encodes approximately 1200 amino acids and is characterized by repetitive tyrosine-aspartate (YD) repeats following a consensus sequence of GxxxxYxYDxxGRL(I/T).<sup>162,163</sup> Adjacent and downstream of the core region, there is a smaller sequence spanning 0.4-0.6 kbp that does not exhibit homology to other parts of the genome.<sup>158</sup> The difference is also reflected by the GC-content, that is of about 60% in the core region but only ~40% in the core extension part.<sup>158,159</sup> The termination point of the core region is marked by a conserved motif called PxxxxDPxGW/L/F.<sup>164</sup> These findings may suggest that the evolution of the Rhs core region and the Rhs core extension (Rhs-CT) proceeded independently. Recombination events among the core regions with different C-terminal core extensions were frequently observed. These extensions are termed hypervariable region (HVR).<sup>164,165</sup>

The core extensions are markedly shorter and encode proteins ranging from 130 to 177 amino acids exhibiting diverse predicted functions.<sup>161,164,165</sup> Some of these extensions show sequence similarity with CdiA, a protein involved in contact-dependent growth inhibition (CDI) in *E. coli* or *Dickeya dadantii*.<sup>166,167</sup> CDI serves as an intercellular competition mechanism, mediated by the action of CdiA. This large effector protein is secreted to the cell surface and delivered to adjacent cells through direct cell contacts, exerting antibacterial activity. Intriguingly, the C-terminal extension of CdiA is solely responsible for its antibacterial effect and is cleaved from the conserved core region by a conserved autoproteolytic aspartyl protease.<sup>167</sup> The producing or sister bacterial cells are protected by expression of an immunity protein CdiI.<sup>167</sup>

Based on the initial discovery of Rhs elements in *E. coli*, more representatives were found across a wide range of Proteobacteria. One example are the BC components of Tc toxins, which are found in the insect and human pathogen *Photorehabdus luminescens* and *Yersinia pestis*.<sup>168,169</sup> Tc toxins are multiprotein complexes that are assembled from three different subunits TcA, TcB and TcC. The TcA component forms homopentamers with a size of 1.4MDa. This pentamer contains several receptor binding domains but also functions as membrane permeation device.<sup>170,171</sup> A pH shift, e.g. after endocytosis of the toxin from the plasma membrane, triggers a conformational change of the TcA subunit, resulting in a puncture

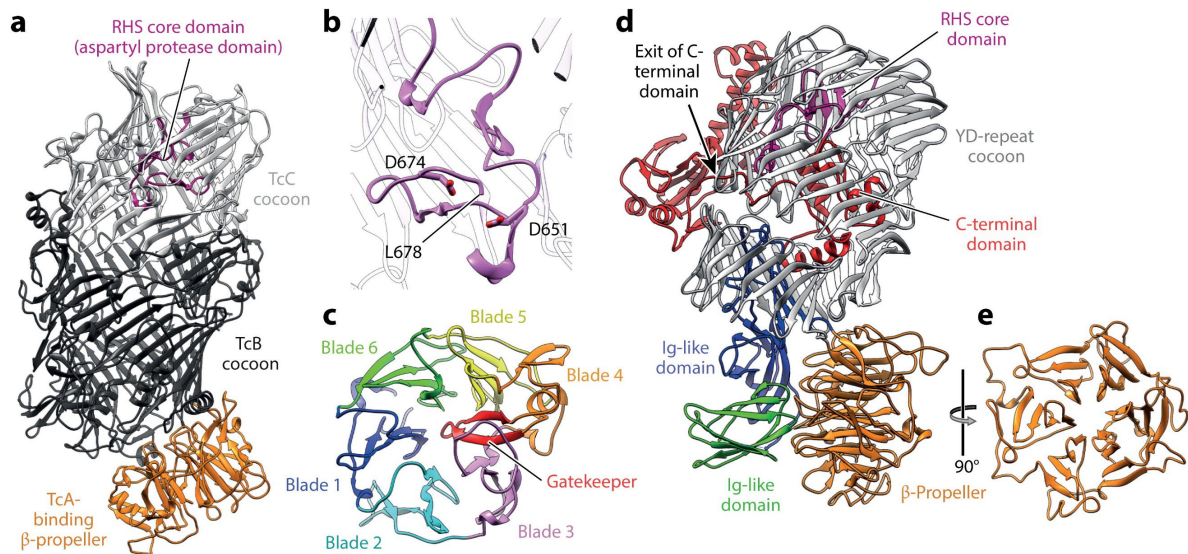
of the cell membrane of the attacked cell. The toxin is encapsulated in the interior of the hollow cocoon formed by the TcB-TcC subunit and is eventually threaded through the translocation channel of the TcA subunit into the cytosol of the attacked cell.<sup>171</sup>

Rhs-repeat containing proteins have been also identified in *Aedes aegypti*, the mosquito species responsible for transmitting the malaria parasite *Plasmodium falciparum*.<sup>172</sup> In mosquitos Rhs-repeats are found in the salivary gland surface protein 1 (SGS1). SGS1 is the most abundant protein in the saliva of this mosquito species where it is thought to be involved in modulation of the host immune system during blood feeding.<sup>173</sup> Phylogenetic analysis has been suggested that SGS was acquired by a lateral gene transfer from prokaryote.<sup>172,174,175</sup> But Rhs repeat-containing proteins can be also found in vertebrates including humans. Teneurins contain Rhs repeats and are a class of transmembrane receptors with a large extracellular domain. They are involved in a multitude of neurobiological processes including the mediation of cell-cell contacts between neurons required for axon guidance<sup>176</sup> and neuronal development<sup>177</sup>. The broad distribution of Rhs repeat containing across different species and kingdoms of life underscores the versatility of being adapted for their specific needs. In order to gain a deeper understanding of their similarities and differences, their three-dimensional structure is of utmost importance.

### 5.13 Structures of Rhs-repeat containing proteins

The first solved structures of Rhs-repeat containing proteins were obtained of BC-components of insecticidal targeting Tc toxin.<sup>171,178</sup> The TcB and TcC subunits form complexes that naturally occur as split subunits but can also be found as natural fusion between the two subunits.<sup>179</sup> Thus, complexes for structural investigations were obtained by genetically fusing the subunits for overexpression and purification leading to the determination of several crystal structures of the TcB-TcC complex from *P. luminescens*<sup>171,180,181</sup>. All structures share a common fold of antiparallel  $\beta$ -sheets representing the Rhs-repeats that assemble into a continuous counterclockwise spiral to form a large hollow cocoon with external dimensions of approximately 110 x 60 Å (Figure 15A). The actual toxin is encoded in the most C-terminal part of TcC, known as the hypervariable region (HVR), which is autoproteolytically cleaved by an aspartyl protease (Figure 15B). The HVR is encapsulated in the interior of the cocoon which provides an environment that maintains the HVR in an unfolded state. This is required, because the HVR is threaded out of the cocoon in an unfolded state through the  $\beta$ -propeller and into the translocation channel of the TcA subunit (Figure 15C).

One side of the cocoon is closed by a distorted six-bladed  $\beta$ -propeller. The propeller is inclined at  $45^\circ$  to the longitudinal axis of the cocoon. The other side of the cocoon is sealed by a conserved Rhs core domain that functions also as an aspartyl protease to cleave off the HRV after its encapsulation. The BC-component recognizes and interacts with the TcA subunit through the  $\beta$ -propeller. Binding to the TcA subunit induces a refolding and opening of the  $\beta$ -propeller enabling the threading of the toxic effector into the translocation channel of the TcA subunit which eventually translocates the effector into the cytoplasm of the targeted cell.<sup>180</sup>




 Roderer D, Raunser S. 2019. *Annu. Rev. Microbiol.* 73:247–65

Figure 15: Structure of a TcB-TcC cocoon and Teneurin2.

(A) The crystal structure of the TcB-TcdB2-TccC3 cocoon from *P. luminescens*. The  $\beta$ -propeller mediates binding to TcA. The Rhs-repeats form a large hollow cocoon to encapsulate the toxic effector which is not resolved in the structure (PDB ID: 4O9X). (B) The aspartyl protease is part of the Rhs core domain that also functions as sealing of one side of the cocoon. (C) The distorted six-bladed propeller is in a closed conformation when not bound to TcA. (D) Structure of the chicken teneurin 2 (Ten2) (PDB ID: 6FB3). (E) Top-view of the  $\beta$ -propeller of Ten2 shows that it is also six-bladed propeller but in a canonical conformation and not distorted. This figure has been adapted from ref.<sup>182</sup> and reprinted with permission from Annual Reviews, Inc.

Several crystal and cryo-EM structures of human and chicken teneurins have been solved providing insight in their shared fold and architecture.<sup>183,184</sup> The Rhs repeats of Teneurins assemble into a large hollow cocoon with outer dimensions of  $\sim 80 \times 50$  Å (Figure 15D & E). Multiple domains are located on the outside of the Rhs shell which can function as receptor binding domains such as the epidermal growth factor-like (EGF) repeats, a Fibronectin domain (FN-plug), an NHL  $\beta$ -propeller and a C-terminal domain. The cocoon is sealed on the one side by a conserved YD-shell plug while the other side of the cocoon is capped by the FN-plug. In contrast to BC components of Tc toxins, teneurins encapsulate smaller cargoes. A linker domain

consisting of 90 amino acids is situated in the interior which covalently connects the Rhs shell to the C-terminal domain. The linker is tightly packed against the Rhs repeats in the inside and exits the shell through a small cavity between the Rhs repeats. The C-terminal domain is packed tightly against the outer wall of the cocoon.

More recently, the cryo-EM structure SGS1 has been determined. SGS1 is a protein found in high quantities in the saliva of mosquitos.<sup>173,185</sup> The protein was isolated from salivary gland extract due to difficulties in recombinant expression. The crude extract was applied to cryo-EM grids, which allowed structural determination.<sup>173</sup>

The structure reveals that SGS1 is a large, cocoon-shaped protein with outer dimensions of 210 x 115 Å (Figure 16). Like BC-components and teneurins, the Rhs-repeats of SGS1 assemble into a large and counterclockwise spiraling hollow cocoon with almost four complete turns. Between the first and second spiral layer, there is a large opening of the cocoon which is sealed by a 3-helix turn. The external part of the cocoon is highly decorated with five potential receptor binding domains. In the inside of the cocoon, several putative TM helices are encapsulated traversing through the interior of the cocoon. These TM helices are not completely folded and primarily kept in an unfolded state. Similar to teneurins, the C-terminus of the encapsulated cargo emerges from a cavity located between the Rhs-repeats.

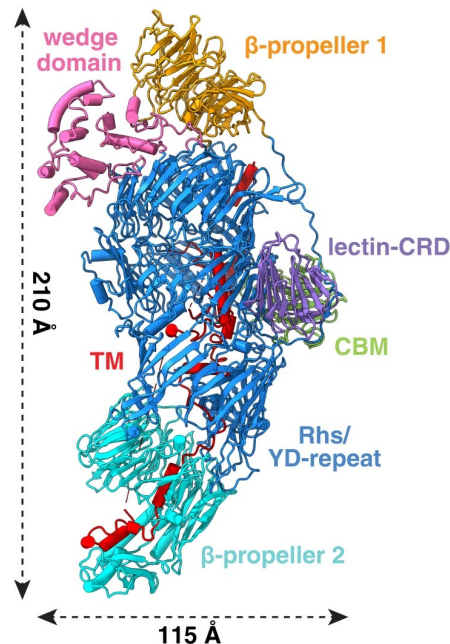


Figure 16: Cryo-EM structure of the salivary gland surface protein 1 (SGS1).

Cartoon representation of the atomic model of SGS1. Each domain is colored and labelled. TM: transmembrane helices, CBM (carbohydrate-binding module), lectin-CRD (lectin carbohydrate-recognition domain). This figure has been adapted and slightly modified from ref.<sup>173</sup> under the license CC BY 4.0.<sup>151</sup>

The structural analysis of proteins mainly composed of  $\beta$ -sheet poses a challenge for particle alignment during image processing. This is due to close spacing and sequence conservation of  $\beta$ -sheet which made it difficult in the past to obtain high-resolution structures.<sup>186</sup> However, advances in software development and image processing workflows have enabled the determination of structures of Rhs-repeat containing proteins. Despite these advances, the successful structural characterization of a bacterial T6SS Rhs effector has remained elusive so far. Obtaining a structure of a bacterial T6SS Rhs effector would be of great interest, as it would permit valuable knowledge about the architecture and functional mechanisms of Rhs repeat-containing proteins in the context of the T6SS. Such knowledge would greatly enhance our understanding of this unique family of proteins and their roles in bacterial systems.



## 6 Aims and motivation

Biological membranes are essential for a wide range of physiological processes, including energy production, cell-to-cell communication and as a protective barrier. Therefore, processes in or at the membrane are of great interest.

For my PhD, I set out to study medically relevant integral membrane proteins that control cholesterol homeostasis, thereby helping to better understand this important regulatory network and lay the foundation for the next generation of rationally designed drugs to treat cardiovascular diseases. I also wanted to gain a better understanding of the bacterial Rhs effector RhsA, which is transported across bacterial membranes for intercellular competition.

Proteins of the SREBP pathway orchestrate cholesterol homeostasis in humans and other higher species. I wanted to understand the underlying mechanism of how these proteins are regulated by cholesterol and the oxysterol 25-HC. Despite their importance, structural knowledge was scarce and limited to soluble portions or mycobacterial homologs, with no insight into the full-length integral membrane proteins. Therefore, I wanted to fill this gap and understand their function by solving high-resolution structures of the proteins alone or in complex. Initially, I set out to structurally characterize subcomplexes composed of the full-length human proteins SREBP, Scap, and Insig. I planned to co-express SREBP-Scap-Insig and Scap-Insig as complexes in mammalian cells and study their interaction through small scale purifications. From there, the expression would be upscaled and a full purification protocol would need to be established. Finally, a stable complex in sufficient quantities should be structurally determined by cryo-EM analysis.

The HMGCR catalyzes the overall rate-limiting step in cholesterol synthesis. I wanted to understand the underlying principles of how sterols accelerate the degradation of HMGCR mediated by the TMD of this protein. The structure of the catalytic portion of this enzyme has been solved. However, the structure of the full-length protein remains elusive. To fill this gap, I wanted to determine the structure of the full-length human HMGCR, ideally in different states and bound to its interaction partners that mediate or prevent degradation of the HMGCR. My initial plan was to overexpress the human HMGCR in mammalian cells and establish a purification protocol. Before proceeding with structural determination, I wanted to measure enzyme kinetics to verify that I had purified the protein in a physiological form. As a next step, I planned to proceed with the structural determination using cryo-EM. As a next step, I would co-express the HMGCR with its interaction partners, Insig1 and UBIAD1, and test their interaction on a small scale. After verifying that they form a complex, I would scale up the expression and determine the complex structure using cryo-EM SPA analysis.



The T6SS transports folded cargoes across several membranes into recipient bacterial cells, including RhsA. RhsA is known to contain repetitive sequences called Rhs repeats. I wanted to understand the function of RhsA by solving a high-resolution cryo-EM structure. I also wanted to compare the overall architecture and fold with other known structures of Rhs repeat-containing proteins. First, I attempted to use cryo-EM analysis to determine the structure of a complete pre-firing complex (PFC) consisting of RhsA, VgrG1, and EagR1. Previously, it was noted that some of these PFCs exhibit a high degree of effector flexibility relative to VgrG, which may prevent the determination of a high-resolution reconstruction as a whole complex. An alternative approach would be to determine the structure of the effector alone.

## 7 Materials

### 7.1 Instruments

<b>Instrument</b>	<b>Manufacturer</b>
ÄKTA purifier	GE Healthcare
Analytical balance PT1200	Sartorius
Autoklav LaM-4-20-MCS-J	SANOclav
Autosampler C-96	Bio-Rad
C1000 Touch Thermal Cycler	Bio-Rad
Carbon evaporator Leica EM ACE600	Leica Microsystems
Centrifuge 5424	Eppendorf
Centrifuge Allegra® X-15R	Beckman Coulter
Centrifuge Avanti® J-26XP	Beckman Coulter
Centrifuge Rotors (JLA 8.100, JA-25.50, JA-10)	Beckman Coulter
ChemiDoc MP Imaging System	Bio-Rad
Countess II FL Automated Cell Counter	ThermoFisher
Dounce Homogenizer	Kontes Glass & Co
Electron microscope Talos Arctica	Thermo Fisher Scientific
Electron microscope Titan Krios	Thermo Fisher Scientific
Electroporator Gene Pulser Xcell™System	Bio-Rad
FEI Tecnai Spirit	Thermo Fisher Scientific
Fluorescence detector RF-20A	Shimadzu
Fluorescence microscope (EVOSFL Auto)	Life Technologies
GloCube Plus	Quorum Technologies Ltd
Heat shock bath	TW2 JULABO
iBind Western Device	Thermo Fisher Scientific
Incubator (37 °C)	BINDER
JEOL JEM-1400	JEOL
Magnetic stirrer MR3000	Heidolph
Microfluidizer Model 110S	Microfluidics Corporation
Microwave Continent MW 800 G	IKA Labortechnik
Mini-membrane-vacuum pump Laboport®	KNF
Mini-PROTEAN® Tetra-Cell	Bio-Rad
Molecular Imager® Gel Doc™ XR System	Bio-Rad
NanoDrop Spectrophotometer ND-1000	PEQLAB
NGC Discover™ 10 Chromatography System	Bio-Rad
pH-Meter FE20 Five Easy™	Mettler Toledo
Pipettes PIPETMAN (0,2-2 µl, 1-10 µl, 2-20 µl, 20-200 µl, 100-1000 µl)	Eppendorf
Pipetting device	Hirschmann Laborgeräte
PowerPac 300	Bio-Rad
PowerPac HV Power Supply	Bio-Rad
Shaker	INFORS HT
Shaker Excella E24	Eppendorf

Shaker Innova	Eppendorf
Sonicator-Sonifier-Branson	VWR International
Tabletop centrifuge 5417R	Eppendorf
ThermoMixer C	Eppendorf
ThermoStat Plus	Eppendorf
Trans-Blot Turbo Transfer System	BioRad Laboratories
Ultracentrifugation rotors (Ti70, Ti45, TLA-120.1, TLA-55)	Beckman Coulter
Ultracentrifuge Optima™ XPN-80	Beckman Coulter
Ultrasonic bath	Sonorex Digitec Bandelin
Vacuum pump	Ilmvac
Vitrobot cryo plunger	Thermo Fisher Scientific
Vortex-Genie 2	Scientific Industries
Zeiss LSM800 microscope with Airyscan	Carl Zeiss Microscopy GmbH

## 7.2 Chromatographic Columns

Column Type	Manufacturer
Superose 6 5/150 Increase	GE Healthcare
Superose 6 10/300	GE Healthcare
Superose 6 10/300 Increase	GE Healthcare
Superdex 200 10/300 Increase	GE Healthcare
HiLoad 16/600 Superose 6 pg	GE Healthcare

## 7.3 Chemicals

Name	Supplier
2-(4-Hydroxyphenylazo)benzoic acid (HABA)	Sigma-Aldrich
25-hydroxy cholesterol	Cayman Chemicals
3x FLAG peptide	APEX-BIO
Acetic acid (glacial)	Sigma-Aldrich
AEBSF	Sigma-Aldrich
Agarose	Biozym Scientific GmbH
Amphiphol A8-A35	Anatrace
Aprotinin	Sigma-Aldrich
Brain Polar Lipids (BPL)	Avanti Polar Lipids
Bromophenolblue	SERVA Electrophoresis GmbH
Cholesterol	Sigma-Aldrich
Cholesteryl hemisuccinate (CHS)	Sigma-Aldrich
Collodion solution	Sigma-Aldrich
Coomassie-Brilliantblue R250	SERVA Electrophoresis GmbH

## Materials

---

CYMAL-6	Anatrace
D-(+)-Biotin	Sigma-Aldrich
DDM-CHS 10:1 solution	Anatrace
Desthiobiotin	Sigma-Aldrich
Di-potassiumhydrogenphosphate	Roth
DIBMA	Anatrace
Digitonin	Sigma-Aldrich
DMSO	Sigma-Aldrich
Ethanol	Thermo Fisher Scientific
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth
Fos-Choline 13	Anatrace
FuGENE® transfection reagent	Promega
GDN	Anatrace
Glycerol	GERBU
Glycine	Roth
HEPES	VWR
HMG-CoA	Sigma-Aldrich
Hydrochloride acid	VWR
iBind Solution Kit	Thermo Fisher Scientific
Imidazole	Sigma-Aldrich
IPTG	Sigma-Aldrich
Isopropanol	Carl Roth
Lanosterol	Sigma-Aldrich
Lauryl dimethylamine-N-oxide (LDAO)	Anatrace
Lauryl maltose neopentyl glycol (LMNG)	Anatrace
Leupeptin hemisulfate	Roth
Lipofectamine® 3000 transfection reagent	Thermo Fisher Scientific
Methanol	J.T. Baker
Midori green	NIPPON Genetics
Milk powder	Roth
<i>n</i> -decyl- $\beta$ -D-maltopyranoside (DM)	Anatrace
<i>n</i> -dodecyl- $\beta$ -D-maltoside (DDM)	Carl Roth
<i>n</i> -nonyl- $\beta$ -D-glucopyranoside (NG)	Anatrace
NADPH	Sigma-Aldrich
POPC	Avanti Polar Lipids
POPE	Avanti Polar Lipids
POPG	Avanti Polar Lipids
Sodium chloride	Roth
Sodium dodecylsulfate (SDS)	Roth
Sodium hydroxide	Waldeck
Soy Polar Lipids	Avanti Polar Lipids
Tris PUFFERAN	Roth
Tris(2-carboxyethyl)phosphine hydrochloride (TCEP)	FLUKA
Triton X-100	Anatrace
Trypan blue	FLUKA

Tween-20	Anatrace
Uranyl formate	Polysciences Inc.

## 7.4 Consumables

<b>Material</b>	<b>Supplier</b>
10 cm petri-dish, sterile	Sarstedt
6,12,24 & 96-well plates (sterile)	Sarstedt
96-well deep well plates 2,2 ml	VWR
Amicon®Ultra-4 und 15 Concentrator	Merck Millipore
Anti-FLAG M2 affinity gel	Signal-Aldrich
Copper grids (for negative stain)	Plano
DMEM/F-12 medium	Thermo Fisher Scientific
Econo-Pac® chromatography columns (50 ml)	BioRad
Eppendorf tubes (0,5 ml, 1,5 ml, 2 ml und 5 ml)	Sarstedt
Falcon tubes (15 ml and 50 ml)	Sarstedt
Falcon tubes (50 ml)	Greiner
Fetal Bovine Serum (FBS)	Thermo Fisher Scientific
Filter paper Whatman No.4	GE Healthcare
Filtropur V50 250/500ml, 0.22µm	Sarstedt
Freestyle 293 expression medium	Gibco
iBind cards	Thermo Fisher Scientific
Ni-NTA superflow resin	QIAGEN
Opti-MEM reduced serum media	Gibco
PageRuler Prestained protein ladder	Thermo Fisher Scientific
Parafilm	Pechiney Plastic Packaging
Pipette tips (10 µl, 200 µl, 1.25 ml)	Sarstedt
QIAprep Spin Miniprep Kit	QIAGEN
QIAquick Gel Extraction Kit	QIAGEN
QIAquick PCR Purification Kit	QIAGEN
QUANTIFOIL R 1.2/1.3 Au (300 mesh)	Quantifoil Micro Tools
QUANTIFOIL R 1.2/1.3 Cu (200 or 300 mesh)	Quantifoil Micro Tools
QUANTIFOIL R 2/1 Cu (300 mesh)	Quantifoil Micro Tools
QUANTIFOIL R 2/1 Cu + 2nm C Layer (300 mesh)	Quantifoil Micro Tools
RPMI 1640 medium	Thermo Fisher Scientific
SDS-Gels Mini-PROTEAN TGX Stain Free Precast Gels (4-15%, 15 wells)	Bio-Rad
Serological Pipettes, sterile (5 ml, 10 ml und 25 ml)	Sarstedt
Serological Rotilabo®-Pipettes (5 ml, 10 ml und 25 ml)	Roth
Sf-900 III SFM	Gibco
Spectra HR ladder	Thermo Fisher Scientific
Strep-Tactin® Superflow® high capacity 50% suspension	IBA
Trans-Blot Turbo Transfer Pack PVDF	Bio-Rag Laboratories
Trypsin solution	Thermo Fisher Scientific
UltrAuFoil Holey Gold Films R 1.2/1.3 (300 mesh)	Quantifoil Micro Tools

## 7.5 Plasmids

Name	Manufacturer/Reference
1281_pEG_BacMam	Eric Gouaux
2079_pMDC_SREBP_Human_Strep	
2080_pMDC_SREBP_Mouse_Strep	
2081_pMDC_SREBP_Human_FLAG	
2082_pMDC_SREBP_Mouse_FLAG	
2129_Crelox_Human_Scap_Srebp_bacmam_vector	
2130_Crelox_Mouse_Scap_Srebp_bacmam_vector	
2152_SREBP_Human_3xFLAG_mCherry_HRV	
2153_SREBP_Mouse_3xFLAG_mCherry_HRV	
2159_pcDNA3.1_Flag_HMGCR_Xenopus-tropicalis	
2189_TagBFP_Rab7_Fusion	
2190_pEG_BacMam_3xFLAG_TagBFP	
2282_pEG_BacMam_DrInsig1_56-251_Linkers_mCherry	
2283_pEG_BacMam_HsInsig1_78-278_Linkers_mCherry	
2284_pEG_BacMam_XtInsig1_56-251_Linkers_mCherry	
2286_pEG_BacMam_DrInsig1_StrepII_HRV_55-251	
2287_pEG_BacMam_HsInsig1_StrepII_HRV_78-278	
2288_pEG_BacMam_XtInsig1_StrepII_HRV_55-251	
2320_CreLoxP_Coexpression_HMGCR-eGFP_Insig1-TagBFP	
2321_CreLoxP_Coexpress_HMGCR(K89A,K248A)-eGFP_Insig1-TagBFP	
2282_pEG_BacMam_DrInsig1_56-251_Linkers_mCherry	
2283_pEG_BacMam_HsInsig1_78-278_Linkers_mCherry	
2284_pEG_BacMam_XtInsig1_56-251_Linkers_mCherry	
2286_pEG_BacMam_DrInsig1_StrepII_HRV_55-251	
2287_pEG_BacMam_HsInsig1_StrepII_HRV_78-278	
2288_pEG_BacMam_XtInsig1_StrepII_HRV_55-251	
2324_pEG_BacMam_HMGCR_E700R_eGFP_Strep	
2325_pEG_BacMam_HMGCR_E700R_E709K_eGFP_Strep	
2329_pEG_BacMam_HMGCR_E700R_E709K_R641A_eGFP	
2330_pEG_BacMam_HMGCR_E700R_E709K_R641A_A695E_eGFP	
2347_pEG_BacMam_HsInsig_78-277_mMBP_Strep_mCherry_HRV_Dela	
2348_pEG_BacMam_HsInsig_78-277_mMBP_Strep_mCherry_HRV_Linkers_A	
2351_pEG_BacMam_HsInsig_78-277_mMBP_Strep_mCherry_HRV_Linkers_AA	
2371_pEG_BacMam_Strep_mCherry_HRV_Insig2_1-225	
2372_pEG_BacMam_Strep_mCherry_HRV_Insig2_20-225	
2373_pEG_BacMam_Strep_mCherry_HRV_mMBP_Insig2_20-225	
2383_pEG_BacMam_StrepII_mCherry_SREBP2_HS_R519A	
2392_pEG_BacMam_SCAP(D428A)_Human	
2393_pEG_BacMam_SCAP(D428A)_Mouse	
2324_pEG_BacMam_HMGCR_E700R_eGFP_Strep	
2325_pEG_BacMam_HMGCR_E700R_E709K_eGFP_Strep	
2329_pEG_BacMam_HMGCR_E700R_E709K_R641A_eGFP	
2330_pEG_BacMam_HMGCR_E700R_E709K_R641A_A695E_eGFP	
2347_pEG_BacMam_HsInsig_78-277_mMBP_Strep_mCherry_HRV_Dela	
2348_pEG_BacMam_HsInsig_78-277_mMBP_Strep_mCherry_HRV_Linkers_A	
2351_pEG_BacMam_HsInsig_78-277_mMBP_Strep_mCherry_HRV_Linkers_AA	
2371_pEG_BacMam_Strep_mCherry_HRV_Insig2_1-225	
2372_pEG_BacMam_Strep_mCherry_HRV_Insig2_20-225	
2373_pEG_BacMam_Strep_mCherry_HRV_mMBP_Insig2_20-225	
2383_pEG_BacMam_StrepII_mCherry_SREBP2_HS_R519A	
2392_pEG_BacMam_SCAP(D428A)_Human	
2393_pEG_BacMam_SCAP(D428A)_Mouse	
2324_pEG_BacMam_HMGCR_E700R_eGFP_Strep	
2464_pEG_BacMam_StrepII-mCherry_HRV_SREBP2wt_Xenopus	
2465_pEG_BacMam_StrepII-mCherry_HRV_SREBP2wt_rabbit	
2466_pEG_BacMam_StrepII-mCherry_HRV_SREBP2wt_fruitfly	
2467_pEG_BacMam_StrepII-mCherry_HRV_SREBP2R482A_Xenopus	
2468_pEG_BacMam_StrepII-mCherry_HRV_SREBP2R482A_rabbit	
2469_pEG_BacMam_StrepII-mCherry_HRV_SREBP2R486A_fruitfly	
2464_pEG_BacMam_StrepII-mCherry_HRV_SREBP2wt_Xenopus	
2465_pEG_BacMam_StrepII-mCherry_HRV_SREBP2wt_rabbit	

2466_pEG_BacMam_StrepII-mCherry_HRV_SREBP2wt_fruitfly	
2467_pEG_BacMam_StrepII-mCherry_HRV_SREBP2R482A_Xenopus	
2522_pEG_BacMam_3xFLAG_BFP_HsINS1_S207A	
2523_pEG_BacMam_3xFLAG_BFP_HsINS1_S207E	
2612_pEG_BacMam_StrepII_3C_Cytosolic_SREBP_Human	
2613_pEG_BacMam_StrepII_mcherry_3C_FL-Frizzled2	
2614_pEG_BacMam_Frizzled2_3C_eGFP_StrepII	
2615_pEG_BacMam_3xFLAG_YFP_3C_HsINS1	
2617_pEG_BacMam_HMGCR-Cter(342-888)_3C_Strep	
2618_pEG_BacMam_HMGCR_Cter(428-888)_3C_StrepII	
2619_pEG_BacMam_StrepII_His8_eGFP_3C_HMGCR	
2620_pEG_BacMam_StrepII_3C_HMGCR-Cter(342-888)	
2621_pEG_BacMam_StrepII_3C_HMGCR_Cter(428-888)	
2622_pEG_BacMam_Strep_linker_HMGCR-human	
2624_pCDNA3.1_UBIAD1_cDNA_clone_Genescript	
2626_pEG_BacMam_Strep_mCherry_3C_UBIAD1	
2627_pEG_BacMam_Strep_mCherry_3C_UBIAD1(G186A)	
2635_pET28a_His7_TEV_SREBP1a(568-1147)_C-term	
2636_pET19_Strep_3C_SREBP1a(569-1147)_C-term	
2637_pET19_10His_SREBP1a(653-1147)	
2638_pET28a_7His_TEV_SREBP1a(653-1147)	
2639_pEG_BacMam_StrepII_mCherry_3C_SREBP1a(569-1147)_C-term	
2640_pEG_BacMam_StrepII_mCherry_3C_SREBP1a(653-1147)_C-term	
2641_pET28_His6_MBP_3C_SREBP1a(568-1147)_C-term	
2658_del_BAK5	Kosiakoff Lab
2659_del_BAG2	Kosiakoff Lab
2660_pHFT_BRIL	Kosiakoff Lab
2661_pEG_BacMam_FLAG_YFP_INS1(BRIL-TM3/4)	
2662_pEG_BacMam_FLAG_YFP_INS1(A-BRIL-A-TM3/4)	
2663_pEG_BacMam_FLAG_YFP_INS1(AA-BRIL-AA-TM3/4)	
2664_pEG_BacMam_FLAG_YFP_INS1(ARRQL-BRIL-ARSTL-TM3/4)	
2665_pEG_BacMam_FLAG_YFP_INS1(AAA-BRIL-AAA-TM3/4)	
2522_pEG_BacMam_3xFLAG_BFP_HsINS1_S207A	
2523_pEG_BacMam_3xFLAG_BFP_HsINS1_S207E	
2669_pET28_His6_MBP_3C_SREBP1a(653-1147)_C-term	
2678_pEG_BacMam_3xFLAG_eGFP_INS1_human	
2697_pFL_His-FLAG_SREBP2(795-1070)_soluble-domain	
2698_pET28a_His-MBP_3C_SREBP2(795-1070)_soluble-domain	
2699_pEG_BacMam_3xFLAG_eYFP_3C_INS1(78-278)_BRIL	
2700_pEG_BacMam_3xFLAG_eYFP_3C_INS1(78-278)_BRIL-A	
2701_pEG_BacMam_3xFLAG_eYFP_3C_INS1(78-278)_BRIL-AA	
2702_pEG_BacMam_3xFLAG_eYFP_3C_INS1(78-278)_BRIL-AAA	
2703_pEG_BacMam_3xFLAG_eYFP_3C_INS1(78-278)_BRIL-AARAFAAA	
2697_pFL_His-FLAG_SREBP2(795-1070)_soluble-domain	
2698_pET28a_His-MBP_3C_SREBP2(795-1070)_soluble-domain	
2699_pEG_BacMam_3xFLAG_eYFP_3C_INS1(78-278)_BRIL	
2711_pEG_BacMam_StrepII_eYFP_3C_Ins1(1-278)_BRIL(TM4/5)	
2712_pEG_BacMam_StrepII_eYFP_3C_Ins1(1-278)_BRIL(TM4/5)_AAA	
2713_StrepII_eGFP_3C_BRIL_SCAP(D428A_20-1279)_BRIL	
2711_pEG_BacMam_StrepII_eYFP_3C_Ins1(1-278)_BRIL(TM4/5)	
2751_pEG_BacMam_3xFLAG_3C_UBIAD1	
2752_pEG_BacMam_3xFLAG_3C_UBIAD1(G186R)	
2756_pEG_BacMam_HMGCR(X_tropicalis)wt_HRV_eGFP_StrepII	
2757_pEG_BacMam_Strep_His8_eGFP_3C_HMGCR(X_tropicalis)wt	
2758_pEG_BacMam_HMGCR(X_tropicalis)_K89A_HRV_eGFP_StrepII	
2759_pEG_BacMam_HMGCR(X_tropicalis)_K89A_K249A_HRV_eGFP_StrepII	
2760_pEG_BacMam_HMGCR_K89A_K248A_A695C_StrepII_eGFP	
2761_pEG_BacMam_HMGCR_K89A_K248A_E700C_StrepII_eGFP	
2762_pEG_BacMam_N-ter_BRIL_HMGCR(human)_K89A_K248A_StrepII_eGFP	
2763_pEG_BacMam_loop1_BRIL_HMGCR(human)_K89A_K248A_StrepII_eGFP	
2764_pEG_BacMam_loop2_BRIL_HMGCR(human)_K89A_K248A_StrepII_eGFP	
2765_pEG_BacMam_loop3_BRIL_HMGCR(human)_K89A_K248A_StrepII_eGFP	
2830_pEG_BacMam_C-ter_BRIL(ARRQL-linker)_HMGCR(1-341)_K89A_K248A_StrepII_eGFP	
2831_pEG_BacMam_C-ter_BRIL(ARRQL-linker)_HMGCR(1-340)_K89A_K248A_StrepII_eGFP	
2832_pEG_BacMam_C-ter_BRIL(ARRQL-linker)-HMGCR(1-339)_K89A_K248A_StrepII_eGFP	

2830\_pEG\_BacMam\_C-ter\_BRIL(ARRQL-linker)\_HMGCR(1-341)\_K89A\_K248A\_StrepII\_eGFP  
 2850\_pEG\_BacMam\_HMGCR\_K89A\_K248A(1-346)\_StrepII\_eGFP  
 2858\_pEG\_BacMam\_3xFLAG\_eYFP\_3C\_UBIAD1(G186R)\_Avi-Tag  
 2859\_pEG\_BacMam\_3xFLAG\_eYFP\_3C\_Avi-Tag\_UBIAD1(G186R)

## 7.6 Enzymes and antibodies

Enzymes and antibodies	Manufacturer
Anti-FLAG antibody (mouse) [H-5]	Santa Cruz Biotechnology
Anti-His antibody (mouse)	Sigma-Aldrich
Anti-mouse antibody Alexa488 conjugate (goat)	Abcam
Anti-mouse antibody Alexa568 conjugate (goat)	Abcam
Anti-mouse antibody-HRP conjugate (goat)	Bio-Rad
Benzonase	Sigma-Aldrich
BSA	Thermo Fisher Scientific
DpnI FD	Thermo Fisher Scientific
HRV 3C protease	In-house production
Phusion High-Fidelity DNA Polymerase	Thermo Fisher Scientific
TEV protease	In-house production

## 7.7 Cell Lines

Organism	Name	Description
<i>Escherichia coli</i>	TOP10F'	F' [proAB, lacI <sub>q</sub> , lacZΔM15, Tn10(Tet <sub>R</sub> )] mcrA, Δ(mrr-hsdRMS-mcrBC), φ80lacZΔM15, ΔlacX74, deoR, recA1, λ-
<i>Escherichia coli</i>	XL10 Gold	Tet <sup>r</sup> Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacI <sup>q</sup> ZDM15 Tn10 (Tet <sup>r</sup> ) Amy Cam <sup>r</sup> ]
<i>Escherichia coli</i>	BL21(DE3) RIPL	<i>E. coli</i> B F <sup>-</sup> ompT hsdS(r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) dcm <sup>+</sup> Tet <sup>r</sup> gal λ(DE3) endA Hte [argU proL Cam <sup>r</sup> ] [argU ileY leuW Strep/Spec]
<i>Escherichia coli</i>	BL21(DE3)	F <sup>-</sup> dcm ompT hsdS(r – m –) gal λ(DE3)
<i>Escherichia coli</i>	DH10EmBacY	F-mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ (ara, leu)7697 galU galK λ-rpsL nupG/ bMON14272/ pMON7124
<i>Homo sapiens</i>	HEK293 GnTI <sup>-</sup>	Immortalized human kidney cancer cell line, N-acetylglucosaminyltransferase I–negative
<i>Homo sapiens</i>	HEK293T	Adherent human embryonic kidney cells
<i>Spodoptera frugiperda</i>	Sf9	Immortalized insect ovarian cell line



## 7.8 Buffer Solutions

### Media

SOC-Media (1 l)	20 g tryptone, 5 g yeast-extract, 0.58 g NaCl 0.19 g KCl, 2.03 g MgCl <sub>2</sub> · 6 H <sub>2</sub> O, 2.46 g MgSO <sub>4</sub> · 7 H <sub>2</sub> O, 50% glucose solution
LB-Media (1 l)	10 g bacterial tryptone, 5 g yeast extract, 10 g NaCl

### General Buffers

10x PBS (500 ml)	40 g NaCl, 1 g KCl, 9.02 g Na <sub>2</sub> HPO <sub>4</sub> · 2 H <sub>2</sub> O 1.2 g K <sub>2</sub> PO <sub>4</sub> , adjust to pH 7.4
10x HABA solution (100 ml)	24.22 mg HABA, dissolve it in 99 ml mQ- H <sub>2</sub> O, add 1 ml 1N NaOH
Regeneration Buffer Strep-Column (500 ml)	1x HABA solution, 100 mM Tris-HCl (pH 8), 150 mM NaCl, 0.5 mM EDTA, adjust to pH 8
SDS sample buffer (4x)	250 mM Tris-HCl (pH 8), 100 mM DTT, 6% SDS, 40% glycerol, 0.02% Bromophenolblue
10x SDS running buffer	250 mM Tris-HCl (pH 8), 1.9 M glycine, 2.8% SDS
Coomassie staining solution	0.15% Coomassie Brilliant Blue R250, 12% Acetic acid, 44% Ethanol
Coomassie destaining solution	10% Acetic acid
TBS	150 mM NaCl, 20 mM Tris, pH 8.0
TBS-T	150 mM NaCl, 20 mM Tris, 0.1% (v/v) Tween-20, pH 8.0
Blocking solution	5% (w/v) milk powder dissolved in TBS
Anode buffer	25 mM imidazole pH 7.0
Dark blue cathode buffer	50 mM tricine, 7.5 mM imidazole, 0.02% (w/v) Coomassie blue G-250
Light blue cathode buffer	50 mM tricine, 7.5 mM imidazole, 0.002% (w/v) Coomassie blue G-250

### FSEC buffers

FSEC solubilization buffer	25 mM HEPES pH 7.4, 150 NaCl, 5 µg/ml AEBSF, 5 µg/ml Leupeptin, 1 mM TCEP, 2% (w/v) respective detergent
FSEC running buffer	25 mM HEPES pH 7.4, 150 mM NaCl, 0.05% DDM, 0.005% CHS, 1 mM TCEP

**Purification Buffers SCAP-Insig complex**

Freezing Buffer (Buffer A1)	1x PBS
Homogenization Buffer (Buffer B1)	25 mM HEPES pH 7.4, 150 mM NaCl 1 mM TCEP, 5 µg/ml AEBSF, 5 µg/ml Aprotinin, 5 µg/ml Leupeptin, 1 mM EDTA, 500 U Benzonase, 1 µg/ml cholesterol, 10 µg/ml 25-HC
Wash Buffer Strep-Column (Buffer C1)	25 mM HEPES pH 7.4, 150 mM NaCl, 1 mM TCEP, 1 mM EDTA, 5 µg/ml AEBSF, 5µg/ml Aprotinin, 5µg/ml Leupeptin, 0.01 % LMNG, 0.002% CHS, 1 µg/ml cholesterol, 10 µg/ml 25-HC
Elution Buffer (Buffer D1)	25 mM HEPES pH 7.4, 150 mM NaCl, 1 mM TCEP, 0.01 % LMNG, 0.002% CHS, 1 µg/ml cholesterol, 10 µg/ml 25-HC
SEC Buffer DDM (Buffer E1)	25 mM HEPES pH 7.4, 150 mM NaCl, 1 mM TCEP, 0.001 % LMNG, 0.0002% CHS, 1 µg/ml cholesterol, 10 µg/ml 25-HC

**Purification Buffers HMGCR-UBIAD1 complex**

<b>Freezing Buffer (Buffer A1)</b>	<b>1x PBS</b>
Homogenization Buffer (Buffer B1)	25 mM HEPES pH 7.4, 150 mM NaCl 1 mM TCEP, 5 µg/ml AEBSF, 5 µg/ml Aprotinin, 5 µg/ml Leupeptin, 1 mM EDTA, 500 U Benzonase, 1 µg/ml lanosterol, 10 µg/ml 25-HC
Wash Buffer Strep-Column (Buffer C1)	25 mM HEPES pH 7.4, 150 mM NaCl, 1 mM TCEP, 1 mM EDTA, 5 µg/ml AEBSF, 5µg/ml Aprotinin, 5µg/ml Leupeptin, 0.01 % LMNG, 0.002% CHS, 1 µg/ml lanosterol, 10 µg/ml 25-HC
Elution Buffer (Buffer D1)	25 mM HEPES pH 7.4, 150 mM NaCl, 1 mM TCEP, 0.01 % LMNG, 0.002% CHS, 1 µg/ml lanosterol, 10 µg/ml 25-HC
SEC Buffer DDM (Buffer E1)	25 mM HEPES pH 7.4, 150 mM NaCl, 1 mM TCEP, 0.001 % LMNG, 0.0002% CHS, 1 µg/ml lanosterol, 10 µg/ml 25-HC

## 8 Methods

### 8.1 Molecular biology

#### 8.1.1 Polymerase chain reaction (PCR)

A polymerase chain reaction with flanking primers was performed to amplify DNA. For this, 50 ng of template containing the gene of interest were mixed with 0.5  $\mu\text{M}$  of oligonucleotides, 1 mM dNTPs, 1x HF Phusion buffer and adjusted with mQ-H<sub>2</sub>O to 50  $\mu\text{l}$ . Then, 1 U of purchased HF Phusion Polymerase was added and the PCR product was amplified using standard PCR protocols. Briefly, the initial template was denatured for 30s at 98°C. Amplification was achieved by performing 30 cycles consisting of 15s at 98°C, primer annealing for 30s at 60°C and extension at 72°C for 30s/kbp. The reaction was completed by an extension of 10min at 72°C. To check for presence of the desired PRC product, 10  $\mu\text{l}$  of the crude mixture were mixed with 2  $\mu\text{l}$  of 6x DNA loading dye and resolved in an 1% agarose gel. The crude PCR mixture was then further purified with the QIAquick PCR Purification Kit or QIAquick Gel Extraction Kit according to manufacturer's guidelines.

#### 8.1.2 Circular polymerase extension cloning (CPEC)

To generate new plasmids or constructs the CPEC method was used as previously described.<sup>187</sup> Briefly, vectors were linearized by PCR as described above in chapter 8.1.1. The insert was amplified with flanking primers containing complementary overhangs to the linearized vector. Crude PCR mixtures were analyzed via agarose gel electrophoresis. PCR products with the correct size were isolated via the QIAquick Gel Extraction Kit (QIAGEN) and eluted in the supplemented Elution buffer. Then, 200ng of the linearized vector were mixed with equimolar amounts of the complementary insert in a total volume of 20 $\mu\text{l}$  containing all components for a PCR as described in 8.1.1. The mixture was denatured for 30s at 98°C, annealed for 30s at 55°C and extended for 1min/kb according to longest piece in the mixture at 72°C. In total two extension cycles were performed. A final extension period was added for 5min at 72°C. 10 $\mu\text{l}$  of the mixture were transformed into chemically competent XL-10 Gold cells.

#### 8.1.3 Site-directed mutagenesis

Site-directed mutagenesis as previously described<sup>188</sup> was used to introduce point mutations, insert or delete up to 15 amino acids into the constructs. Amplification with mutagenic primers was performed as described in chapter 8.1.1 without subsequent PCR purification. Instead, 1 $\mu\text{l}$  of Fast-digest DpnI was added directly to the crude reaction mixture in order to digest the

template DNA. The mixture was incubated for 1h at 37°C followed by 20min at 80°C to heat-inactivate the enzyme. The reaction was placed for 15min at 4°C and 10µl were used for transformation into chemically competent XL-10 Gold *E. coli* cells.

### **8.1.4 Chemical transformation of *Escherichia coli* cells**

50 ng of plasmid DNA (or as much as indicated above) were mixed with 50 µl suspension of chemically competent *E. coli* cells. The tube was incubated for 30 minutes on ice, followed by heat-shock at 42 °C for 45s and additional incubation on ice for two minutes. Then 1 ml of pre-warmed SOC media was added and the cells were incubated at 37 °C while shaking at 800 rpm for one hour. Subsequently, 50 µl of the cell suspension were applied onto an agar plate containing an appropriate antibiotic. Single colonies were obtained after an over-night incubation at 37 °C.

### **8.1.5 Plasmid isolation**

A single colony was picked from an overnight agar plate and transferred to a tube containing 5 ml of LB medium supplemented with the appropriate antibiotic (e.g., 125 µg/ml ampicillin). The tube was incubated overnight at 37°C in an orbital shaker at 180 rpm. The next day, the cells were pelleted by centrifugation at 4000 rpm. Plasmid DNA was isolated from the pellet using the Qiagen Miniprep Kit according to the manufacturer's instructions.

### **8.1.6 Sanger sequencing**

300 µg of plasmid DNA was mixed with 1.5 nmol of sequencing primer in a total volume of 15 µl. The tube was sent to the external company Microsynth Seqlab for sequencing by the Sanger method.

## **8.2 Cell biology**

### **8.2.1 Culturing and transfection of human adherent cell lines**

The adherent human cell lines HEK293T and HEK293 GnTI<sup>-</sup> were cultured in DMEM/F12 medium supplemented with 10% FBS and incubated at 37°C with 5% CO<sub>2</sub>. The cells were initially seeded at a density of 1 x 10<sup>6</sup> cells/ml. When the cells reached 70-90% confluency, the cells were detached from the cell culture plate by trypsin treatment and transferred to a culture plate at a density of 1 x 10<sup>6</sup> cells/ml. Cells were analyzed by conventional light microscopy for obvious morphological changes or contaminations.

### 8.2.2 Culturing of human suspension cells lines

HEK293 GnTI<sup>-</sup> and HEK293F cells were cultured in Freestyle 293 Expression Medium supplemented with 2% FBS at 37°C, 5% CO<sub>2</sub> and 8% humidity while shaking at 130 rpm. Cells were regularly maintained at a cell density of 0.5-5 x 10<sup>6</sup> cells/ml. Cell density was measured with the Countess II FL Automated cell counter.

### 8.2.3 Transient transfection of human adherent cells

Cells were seeded at a density of 1 x 10<sup>6</sup>/ml in culture dishes/plates and grown to 70-90% confluency. On the day of transfection, a specified amount of DNA (see Table 1) was mixed with 200µl Opti-MEM medium in a sterile tube. The diluted DNA was then combined with 200 µl Opti-MEM medium containing FuGENE reagent. The mixture was incubated for 15 min at room temperature before being added to the confluent cells. The plate was then incubated at 37°C with 5% CO<sub>2</sub> for 24h. Transfection success was determined by fluorescence microscopy, with cells checked for eGFP or mCherry expression.

Table 1: DNA and FuGENE amount for transfection of adherent human cells.

Well-plate or dish	DNA amount	FuGENE reagent
6-well	3 µg	9 µl
12-well	1 µg	3 µl
24-well	500 ng	1.5 µl
96-well	200 ng	0.6 µl
3.5 cm dish	1 µg	3 µl
10 cm dish	10 µg	30 µl

### 8.2.4 Electroporation of DH10EmBacY *E. coli* cells

100 µl electro-competent DH10EMBacY cells were gently mixed with 50 ng of plasmid DNA. The cell suspension was transferred into an electroporation cuvette with 2 mm sample width. The cuvette was placed in the electroporator and an electric pulse of 2.5 kV was applied for 5 ms. The cells were immediately recovered using 1 ml of pre-warmed SOC-media and incubated overnight at 37°C while shaking at 650rpm. The next morning, 100µl of the cell suspension were spread on an agar plate containing 10 µg/ml tetracycline, 50 µg/ml kanamycin, 10 µg/ml gentamycin, 1mM IPTG and 100 µg/ml X-Gal. After incubation of the plate at 37°C for 48h, the plate was screened for single white colonies which represent successful integration of the GOI into the bacterial artificial genome.

### **8.2.5 Bacmid isolation from DH10EmBacY cells**

A single white colony was inoculated into 5ml LB medium containing 10 µg/ml tetracycline, 50 µg/ml kanamycin and 10 µg/ml gentamycin. The cells were incubated at 37°C overnight while shaking at 160rpm. To isolate the Bacmid DNA, the cells were collected by centrifugation at 4000x g at room temperature. Cell lysis was achieved by the P1, P2 and N3 buffers of the Plasmid Miniprep Kit (Qiagen). The lysate was spun for 5min at 15,000x g at 4°C from which 720µl of soluble supernatant were transferred to a fresh tube. Next, 500µl of isopropanol were added to the soluble lysate. The precipitated DNA was collected by centrifugation at 16,900x g for 10min at RT and the resulting supernatant was carefully removed. The pellet was washed once with 200µl ice-cold 70% ethanol and resuspended in 30µl sterilized water.

### **8.2.6 Transfection of *Sf9* insect cells – baculovirus generation and amplification**

For generation of the virus stocks, *Sf9* cells were transfected with the isolated bacmid (section 8.2.5). To this end, 100µl of Sf-900 III SFM medium were added to the 30µl bacmid solution which were mixed with 10µl FuGENE HD Transfection Reagent (Promega). The mixture was incubated for 30min at RT. Meanwhile, 3 ml of *Sf9* insect cells at a density of  $0.7 \times 10^6$  cell/ml were allowed to adhere to the bottom of a 6-well plate at 27°C. Then, the transfection mixture containing the bacmid was dropwise added to the cells. P1 virus stocks were harvested after 4 days of incubation at 27°C in the dark from the supernatant of the culture and supplemented with 2% (v/v) FBS for storage.

For generation of viruses with higher potency further amplification was performed. P2 viruses were generated by infecting 200 ml *Sf9* cells at a density of  $1 \times 10^6$  cells/ml with 0.2% (v/v) P1 virus from the previous step. After 24h of incubation at 27°C while shaking at 100rpm, 200 ml fresh Sf-900 III SFM medium was added. Again 24h later, 100 ml fresh medium was added to the culture. The P2 virus was harvested from the supernatant after a total of five days. The cells were separated from the virions by centrifugation at 4000x g for 10min at RT. The resulting supernatant was sterile filtered through a 0.22µm filter, supplemented with 2% (v/v) FBS and stored at 4°C. The potency of the resulting P2 virus was determined using the endpoint dilution assay.

### **8.2.7 Endpoint dilution assay**

The potency of the virus assayed using the endpoint dilution assay as described previously<sup>189</sup>. Briefly, a 10 ml stock of *Sf9* insect cells with a density of  $0.7 \times 10^6$  cells/ml was prepared. 100 µl

of cell suspension were transferred using a multichannel pipette into the wells of a 96-well black plate and allowed to attach to the bottom of the wells for 20 minutes. Meanwhile, a 10-fold serial dilution of the virus was prepared in sterile deep-well 8-strip clusters resulting in  $10^1$  to  $10^{-8}$  stocks of the viruses (40  $\mu$ l virus + 360  $\mu$ l Sf-900 III SFM medium). Then, the medium of the wells was replaced with 100  $\mu$ l of the virus dilutions. Each dilution was measured in triplicate. The cells were incubated for 72 hours at 27°C in the dark. Using a fluorescence microscope, the number of red foci was evaluated. Foci were manually counted from wells with fewer than 10 foci which were used to calculate the viral titer:

$$(\text{Average \# foci}) \cdot \text{dilution factor} \cdot 10 = \text{pfu/ml}$$

### 8.2.8 Transduction of HEK293 GnTI<sup>-</sup> suspension cells

For expression of the target protein in mammalian cells, HEK293 GnTI<sup>-</sup> or HEK293F cells were grown in Freestyle medium supplemented with 2% FBS in an orbital shaker at 37 °C with 8% CO<sub>2</sub> to a density of 2-3 x 10<sup>6</sup> cells/ml. P2 virus was added up to 10% (v/v) of the initial culture volume. In case of coexpression of multiple proteins, an equal amount of the different P2 viruses were added. The flask was returned to 37°C with 8% CO<sub>2</sub> for 8h while shaking at 130rpm. Then, sodium butyrate was added to a final concentration of 10 mM and the temperature was lowered to 30°C. After a total of 48h after transduction, cells were harvested by centrifugation at 1500x g for 5min at 4°C. The pellets were resuspended in 1x PBS buffer (25 ml buffer/L expression culture), flash frozen in liquid N<sub>2</sub> and stored at -80°C till the day of purification.

## 8.3 Protein biochemistry

### 8.3.1 Semi-denaturing SDS-PAGE

Protein samples were mixed with 4x SDS sample buffer. Membrane protein samples were directly loaded into the wells of a 4-15% Mini-PROTEAN TGX Stain-Free (semi-denaturing) gel. Other samples were optionally heated at 95°C for 5min and then loaded into the wells of the gel. Proteins were separated by applying a voltage of 180V for 30min. The gels were examined by stain-free imaging in a ChemiDoc Imaging system. If proteins were tagged with fluorescent proteins in-gel fluorescence was measured for three channels in parallel (Blue Epi Illumination, 530/28 filter; Green Epi Illumination, 605/50 filter; Red Epi Illumination, 695/55 filter). In order to stain gels by conventional Coomassie blue stain, the gels were incubated 30min in Coomassie staining solution during slow agitation. The excess staining solution was

removed and replaced by Destaining solution and further agitated for 1h. Image of the Coomassie stained gel were recorded in a ChemiDoc Imaging system.

### **8.3.2 Western blot**

The different protein samples were separated via SDS-PAGE as described in section 8.3.1. The gel was placed onto a PVDF membrane and proteins were transferred using a Trans-Blot Turbo Transfer system (Biorad) by applying 2.5A for 3min. Afterwards, the membrane was blocked for 1h at RT with blocking solution. The membrane was washed thoroughly with TBS-T buffer and then incubated with a 1:500-1:1000 dilution of the primary antibody overnight at 4°C during slow agitation. The solution was removed and the membrane was again washed with TBS-T buffer. Then, 1:2000 dilution of the secondary HRP-conjugated antibody was added to membrane and gently agitated for 1h at RT. The solution was removed and the membrane was washed extensively with TBS-T. Then, 5 ml Western Lightning Plus-ECL (Perkin Elmer) solution was freshly prepared according to manufacturer's guidelines and added to the membrane. Protein bands were detected after 1min incubation at RT by transferring the membrane into a ChemiDoc MP Imaging system.

### **8.3.3 Blue native PAGE**

To probe for the oligomeric state of a membrane protein or membrane protein complex Blue native PAGE was performed. Glycerol to a final concentration of 10% (v/v) was added to the purified protein sample. Coomassie G-250 Brilliant blue was added from a 5% (w/v) stock solution to achieve a detergent:dye ratio of 8:1 (w/w) and directly loaded into the wells of an 4-15% Mini-PROTEAN TGX Stain-Free (semi-denaturing) gel. Initially, dark blue cathode buffer was initially added in the inner chamber. Anode buffer was added to the outer chamber. Electrophoresis was performed at 4°C at 100V till about one-third of the running front entered the gel. At that point, electrophoresis was stopped and the cathode buffer was replaced by the light blue cathode buffer and electrophoresis was resumed for 120 at 150V at 4°C. The gel was immediately stained and destained with Coomassie staining solution as described in section 8.3.1.

### **8.3.4 FSEC-TS with whole cell lysates**

For small-scale FSEC-TS analysis, HEK293 GnTI<sup>-</sup> grown in 10cm dishes were harvested 24h after transient transfection. The cells were washed once with 1 ml ice-cold 1xPBS and resuspended in FSEC solubilization buffer. Samples were rotated at 4°C for 2h to achieve



solubilization of the target membrane protein. Insoluble material was pelleted at 20,000x g for 30min at 4°C. 100µl aliquots were divided into thin-walled PCR tubes. After 10-min incubation of the samples at 4, 20, 30, 40, 50 and 60°C in a Thermocycler, samples were filtered through a 0.22µm centrifugal filter unit. The filtrate containing the solubilized membrane protein of interest was injected into a Superose 6 10/300 column equilibrated in FSEC running buffer run at a flowrate of 0.45 ml/min. A fluorometer was connected to the SEC column in order to identify the fluorescently labelled target protein [(TagBFP, excitation: 402 nm, emission: 457 nm); (GFP, excitation: 475 nm, emission, 510 nm); (eYFP, excitation: 513 nm, emission: 527 nm); (mCherry, excitation: 587 nm, emission: 610 nm)]. The fluorescence intensities of the peak of the eluted protein were normalized relatively to the height of the peak of the sample incubated at 4°C which was set to a value of 1. Data was visualized in Origin and fitted by dose-response curve fit to determine the melting temperature ( $T_m$ ). This general workflow was used to compare the thermal stability of the target protein under different buffer conditions, detergents, supplemented lipids, and ligands.

### 8.3.5 Detergent-screening

For membrane proteins, finding a detergent that extracts the target protein with high efficiency but still maintains it in a native-like state is crucial. However, the appropriate detergent has to be identified for each protein individually. FSEC was used to identify detergents that on the one hand solubilize the target protein efficiently but on the other hand give rise to a gaussian shape peak during chromatographic separation in a SEC column.

Here, 1 ml of P2-virus transduced HEK293 GnTI<sup>-</sup> cells expressing the target protein were harvested by centrifugation at 1500x g for 5min at 4°C. The cell pellet was resuspended in 800µl FSEC lysis buffer. Subsequently a final of 2% (w/v) of the detergent to be tested was added to the resuspended cells. The sample was incubated for 2h at 4°C during slow rotation. Insoluble material was removed by centrifugation at 20,000x g for 30min at 4°C and the supernatant was filtered through a 0.22µm centrifugal filter. Then, 150µl of the cleared lysate were injected into a Superose 6 10/300 column pre-equilibrated in FSEC running buffer with a flowrate of 0.4 ml/min. A connected fluorometer was used to identify the fluorescently labelled target protein [(TagBFP, excitation: 402 nm, emission: 457 nm); (GFP, excitation: 475 nm, emission, 510 nm); (eYFP, excitation: 513 nm, emission: 527 nm); (mCherry, excitation: 587 nm, emission: 610 nm)]. Complete fluorescent traces were recorded and plotted in Origin.

### 8.3.6 Protein purification of SCAP-Insig and HMGCR-UBIAD1 complexes

On the day of purification, a cell pellet corresponding to 2L expression culture was thawed at RT. The cell suspension was mixed with an equal volume of homogenization buffer (B1). Cell clumps were disrupted by ten passages through a class douncer. Then, LMNG-CHS (5:1) solution and sodium cholate were added to a final concentration of 1% (w/v) and 0.2% (w/v), respectively. For the SCAP-Insig complex, additionally 10 µg/ml cholesterol and 1 µg/ml 25 hydroxycholesterol were added to each buffer during the purification. Membrane proteins were solubilized for 2h at 4°C during slow rotation after which insoluble material was removed by centrifugation at 40,000rpm for 1h at 4°C in a Ti45 rotor. The cleared supernatant was loaded to anti-FLAG M2 resin, rinsed with 10CVs of wash buffer and eluted with wash buffer supplement with 200 µg/ml 3x FLAG peptide. The eluate was then loaded to Strep-Tactin high-capacity beads. The resin was washed with 10CVs of wash buffer and fractionated with elution buffer containing 10mM D-Desthiobiotin. Each fraction was analyzed by SDS-PAGE for presence of the target proteins. Fractions containing the proteins were pooled and concentrated to 500µl. After addition of 100µg of in-house produced HRV 3C protease samples were slowly rotated over night at 4°C. The next morning, the samples were spun at 15,000x g for 5min at 4°C and injected into a Superose 6 10/300 increase column at a flowrate of 0.4 ml/min pre-equilibrated in SEC buffer. Fractions were again confirmed and analyzed by SDS-PAGE. Protein complexes were concentrated to 1 mg/ml and immediately used for structural characterization.

### 8.3.7 Protein purification HMGCR

On the day of purification, a cell pellet corresponding to 1L expression culture was thawed at RT. The cell suspension was mixed with an equal volume of homogenization buffer (B1). Cell clumps were disrupted by ten passages through a class douncer. Then, DDM-CHS (10:1) solution was added to a final concentration of 1% (w/v) DDM and 0.1% (w/v) CHS. Solubilization was completed for 1h at 4°C during slow rotation after which insoluble material was removed by centrifugation at 40,000rpm for 1h at 4°C in a Ti45 rotor. The cleared supernatant was loaded to Strep-Tactin beads. The resin was rinsed with 10CVs of wash buffer and fractionated with elution buffer containing 10mM D-Desthiobiotin. Fractions containing the target protein were confirmed by SDS-PAGE and pooled and further concentrated to 500µl. The protein was applied to a Superose 6 10/300 increase column run at a flowrate of 0.4 ml/min pre-equilibrated in SEC buffer. Fractions were again confirmed and analyzed by SDS-PAGE, pooled and concentrated to 4 mg/ml. The protein was aliquoted and flash frozen in liquid N<sub>2</sub>.

Aliquots were stored at  $-80^{\circ}\text{C}$  and used for activity assays. For structural characterization, the protein was immediately used without a freezing step.

### 8.3.8 Enzymatic activity assay for the HMGCR

The activity assay was previously established in the laboratory of Arne Bothe. The enzymatic activity of the HMGCR was followed by measuring the consumption of NADPH, which is converted to  $\text{NADP}^+$ . NADPH has an additional absorption maximum at 340 nm compared to  $\text{NADP}^+$ . In a volume of 200  $\mu\text{l}$ , 120 nM purified HMGCR in DDM-CHS micelles was mixed with 800  $\mu\text{M}$  NADPH, 0.05% DDM, 0.005% CHS in a buffer containing 25 mM HEPES pH 7.4 and 150 mM NaCl. This was used to blank the system prior to the addition of HMG-CoA (0-400  $\mu\text{M}$ ). The UV/Vis absorbance at 340 nm was recorded over a period of 5 min, from which the initial reaction rates ( $v_0$ ) were calculated and plotted according to Michaelis-Menten.

### 8.3.9 Reconstitution of membrane proteins into nanodiscs

Lipid nanodiscs are small lipidic patches engulfed by membrane scaffold protein (MSP). They offer a near-to-native environment to study membrane proteins in aqueous buffers. They also allow easy modification of the type of lipid or even lipid-mixtures and i.e., addition of cholesterol and cholesterol derivatives. Here, a general protocol is described which was used for different membrane proteins throughout this thesis.

Lipids were purchased from Avanti Polar Lipids dissolved in chloroform at 25 mg/ml. The following lipids were used for nanodisc reconstitution: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylglycerol (POPG), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE), and soybean polar lipid extract. Cholesterol and 25-hydroxycholesterol were dissolved in pure ethanol at 10 mg/ml. When lipid mixtures were used for reconstitution, they were first mixed in chloroform to ensure proper mixing.

Single lipids or lipid mixtures dissolved in chloroform were transferred to a new glass vial. The chloroform was evaporated using a gentle argon stream to form a thin lipid film. The lipid film was dried overnight in a vacuum chamber and solubilized at 50 mM lipid concentration in nanodisc buffer containing 100 mM sodium cholate. Lipid stocks were stored at  $-20^{\circ}\text{C}$  until further use.

For nanodisc reconstitution, the purified membrane protein, MSP and lipids were mixed at a molar ratio of 1:5:250 for MSP1D1 and 1:5:500 for MSP1E3D1. The reconstitution mixture was incubated for 1h on ice. The nanodisc formation was initiated by the addition of 250 mg

activated SM-2 Biobeads. After slow rotation for 1h at 4°C a second batch of fresh SM-2 Biobeads were added and slowly rotated overnight. Biobeads were removed by filtration through a 0.22µm centrifugal filter and further purified via SEC in a Superose 6 10/300 increase column run at a flowrate of 0.4 ml/min pre-equilibrated in detergent-free SEC buffer. Reconstitution was assessed by SDS-PAGE and negative-stain EM.

### **8.3.10 Reconstitution of membrane proteins into liposomes**

Lipid stocks were prepared as described in section 8.3.9 and mixed with the purified target protein with a 2000-fold molar excess of the lipids over the membrane protein. The mixture was incubated for 1h on ice. Liposome formation was initiated by the addition of 250 mg activated SM-2 Biobeads. After slow rotation for 1h at 4°C a second batch of fresh SM-2 Biobeads were added and slowly rotated overnight at 4°C. Biobeads were removed from the mixture by centrifugation at 15,000x g for 10min at 4°C.

## **8.4 Structural biology**

### **8.4.1 Transmission electron microscopy – negative staining**

#### **8.4.1.1 Sample preparation and data collection**

The 0.75% uranyl formate staining solution was freshly prepared before use. To this end, 37.5 mg of uranyl formate were dissolved in 5 ml of boiling milliQ-water. The solution was allowed to stir for 5 minutes in the dark. Then, 4.5 µl 5 M NaOH was added and the solution was stirred for 5 additional minutes. The solution was subsequently filtered using a 0.22 µm syringe filter and stored plastic test tubes covered with aluminum foil up to three days.

Four microliters of the sample, at a concentration of 0.004 mg/mL, were applied to freshly glow-discharged, carbon-coated copper grids. The sample was incubated for 90s on the grid and excess protein solution was blotted away with Whatman No. 4. The grid was washed twice with four microliters mQ-H<sub>2</sub>O and once with 0.75% (w/v) uranyl formate. A second batch of staining solution was incubated on the grid for 90s before excess was again blotted away. The grids were air-dried and imaged on a JEOL JEM-1400 or FEI Tecnai Spirit microscope. The microscope was equipped with a LaB6 cathode and a 4k × 4k CMOS detector F416 (TVIPS), operating at 120 kV. Typically, images were recorded on the FEI Tecnai Spirit at a nominal magnification of 52,000x corresponding to a pixel size of 2.13 Å/pix. Images were recorded with a defocus in a range from -1.5 to -2 µm.

### 8.4.1.2 Single particle analysis and image processing

The negative stain micrographs were processed using the SPHIRE<sup>190</sup> software package. Particle coordinates were detected using the general model for negative stain of crYOLO<sup>191</sup> or they were manually selected. The particles were extracted from the micrographs and classified in 2D using ISAC<sup>192</sup>. The obtained class averages were further manually inspected and selected.

## 8.4.2 Cryo-electron microscopy

### 8.4.2.1 Sample preparation and data collection

The Vitrobot Mark IV was used for all grid preparations set at 8°C and 100% humidity. The sample preparations and plunging conditions were for each sample optimized in several rounds of screening. Samples were screened regularly on a Talos Arctica operated at 200kV and equipped with a Falcon 3 detector.

All datasets were monitored on-the-fly with TranSPHIRE<sup>193</sup> for quality control i.e. to monitor defocus values or particle motion. Moreover, TranSPHIRE was also used to perform pre-processing steps such as beam-induced motion correction by MotionCor2<sup>194</sup>, estimating CTF parameter using CTFFIND4<sup>195</sup> and automated detection of particle location coordinates by the general cryo-EM model of crYOLO<sup>191</sup>.

#### 8.4.2.1.1 Scap-Insig complex

Typically, a 4µl aliquot of the SCAP-Insig complex in LMNG was applied to freshly glow-discharged UltrAuFoil R1.2/1.3 300 grids (Quantifoil). After removing excess liquid by blotting for 3.5s (blot force -3), grids were immediately plunge-frozen in liquid ethane cooled by liquid nitrogen. The grid was then transferred to a Cs-corrected Titan Krios operated at 300 kV equipped with a GIF BioQuantum energy filter set to a slit width of 20eV and a K3 Summit Detector. Image stacks were recorded in super-resolution mode with a nominal magnification of 81,000x, resulting in a super-resolution pixel size of 0.45Å/pix. A total exposure time of 4s with a dose of 89 e<sup>-</sup>/Å<sup>2</sup> was fractionated over 60 frames. Data was collected in automated fashion using EPU with a defocus range from -0.8 to -2.2 µm.

#### 8.4.2.1.2 HMGCR-UBIAD1 complex

A 2.5µl aliquot of the HMGCR-UBIAD1 complex in LMNG at a concentration of 1 mg/ml was applied to freshly glow-discharged holey carbon gold grids (Quantifoil Au R1.2/1.3 300 mesh). After removing excess liquid by blotting for 3.5s (blot force 0), grids were immediately plunge-

frozen in liquid ethane cooled by liquid nitrogen. The grid was then transferred to a  $C_s$ -corrected Titan Krios operated at 300 kV equipped with a GIF BioQuantum energy filter set to a slit width of 15eV and a K3 Summit Detector. Image stacks were recorded in super-resolution mode with a nominal magnification of 81,000x, resulting in a super-resolution pixel size of 0.44Å/pix. A total exposure time of 2s with a dose of 60  $e^-/\text{Å}^2$  was fractionated over 60 frames. Data was collected in automated fashion using EPU with a defocus range from -0.8 to -2.2  $\mu\text{m}$ .

#### **8.4.2.2 Image processing**

All datasets were processed by iteratively change processing software as needed and indicated from SPHIRE, RELION<sup>196</sup> and cryoSPARC<sup>197</sup>. Initially, micrographs with a resolution limit lower than 10Å were manually excluded using the graphical CTF assessment tool in SPHIRE. Dose-weighted micrographs were then imported for particle extraction into SPHIRE.

##### **8.4.2.2.1 Scap-Insig1 complex**

The best dataset collected for the SCAP-Insig1 complex was composed of 7204 good micrographs. Particle coordinates were determined with the general cryo-EM model of crYOLO and subsequently extracted in SPHIRE with a box size of 300 x 300 pixels. The 1,750,773 particles were classified in 2D using the reference-free clustering algorithm ISAC. ISAC was performed at a pixel size of 3.36Å/pix. A subset of particles showing clearly features of the protein complex were manually selected from representative class averages and subjected once more to classification via ISAC. Good particles were extracted from classes showing features of the protein complex. Then, 732,902 good particles were imported into RELION for 2D classifications to further clean-up the dataset and remove junk particles (ignore CTF till first peak: true). From 2D classifications in RELION 125,458 good particles were selected and subjected to an *ab-initio* initial model generation in cryoSPARC (3 models). The best model from cryoSPARC was imported into RELION and used for a 3D classification without mask (K=2, T=4). Particles from the class showing higher resolution were selected and subjected to a Gold-standard 3D refinement in RELION using a soft-mask. The final map was postprocessed with a soft mask and a B-factor of -130  $\text{Å}^2$ .

##### **8.4.2.2.2 HMGCR-UBIAD1**

8,901 micrographs were collected from which 8,692 good micrographs were selected with the criteria described above. Particles coordinates were determined using the general cryoEM model of crYOLO (confidence threshold 0.1) and subsequently extracted in SPHIRE with a

box size of 360 x 360 pixels binned three times to a boxsize of 120 x 120. The binned 4,351,845 particles were imported into cryoSPARC for multiple rounds of 2D classification. A subset of 876,936 good particles were selected for an *Ab initio* model generation which served as references for heterogenous refinements (1 good model, 2 junk models). After multiple rounds of heterogenous refinement, a subset of 190,851 good particles was selected and reextracted using a box size of 360 x 360 pixels. The good particles were iteratively refined by non-uniform refinement and local refinement in cryoSPARC until the resolution and the quality of the map did not further improve. The final map was postprocessed with a soft mask and a B-factor of -100 Å<sup>2</sup>.

## 8.5 Methods specific to RhsA

Methods related to RhsA have been published previously<sup>198</sup>. For ease of reading, method descriptions were integrated and reprinted in this part of the thesis. All molecular biology methods including work regarding bacterial culturing, construct design, generation of mutants, competition assays, western blot analysis and protein purification have been performed by Shehryar Ahmad, Kartik Sachar and John C. Whitney. Cryo-EM analysis has been performed with the help of Dennis Quentin and Christos Gatsogiannis. The method description of section 8.5.1 to 8.5.5 has been adapted, slightly adjusted and reprinted from Ref.<sup>150</sup> in the original publication<sup>198</sup> and therewith in this thesis under the CC-BY 4.0 license<sup>151</sup>.

Patrick Günther\*, Dennis Quentin\*, Shehryar Ahmad, Kartik Sachar, Christos Gatsogiannis, John C. Whitney, und Stefan Raunser. “Structure of a bacterial Rhs effector exported by the type VI secretion system”. *PLoS Pathogens* 18, Nr. 1 (2022): 1–25.

Online version: <https://doi.org/10.1371/journal.ppat.1010182>.

### 8.5.1 Bacterial strains and culture conditions

*Pseudomonas protegens* Pf-5 (Table S 1) was grown in Lysogeny Broth (LB) (10 g L<sup>-1</sup> NaCl, 10 g L<sup>-1</sup> tryptone, and 5 g L<sup>-1</sup> yeast extract) at 30°C or on solid LB containing 1.5% or 3% agar. Media were supplemented with gentamicin (30 µg mL<sup>-1</sup>) and irgasan (25 µg mL<sup>-1</sup>) as needed. *Escherichia coli* strains XL-1 Blue, SM10 and CodonPlus (DE3) were used for plasmid maintenance and toxicity experiments, conjugative transfer and protein overexpression, respectively (S1 Table). All *E. coli* strains were grown at 37°C in LB medium. Unless otherwise noted, media was supplemented with 150 µg mL<sup>-1</sup> carbenicillin, 50 µg mL<sup>-1</sup> kanamycin, 200 µg

mL<sup>-1</sup> trimethoprim, 15 µg mL<sup>-1</sup> gentamicin, 0.25-1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), 0.1% (w/v) rhamnose or 40 µg mL<sup>-1</sup> X-gal.

### 8.5.2 DNA manipulation and plasmid construction

All primers used in this study were synthesized by Integrated DNA Technologies (IDT). Molecular biology reagents (Phusion polymerase, restriction enzymes and T4 DNA ligase) were obtained from New England Biolabs (NEB). Sanger sequencing was performed by Genewiz Incorporated.

Heterologous expression plasmids: pETDuet-1, pET29b and pSCrhaB2-CV. Splicing by overlap-extension PCR was used to make mutant constructs. Standard restriction enzyme-based cloning procedures were subsequently used to ligate wild-type or mutant PCR products into the plasmid of interest.

### 8.5.3 Generation of *P. protegens* mutants

In-frame chromosomal deletion mutants in *P. protegens* were made using the pEXG2 plasmid as described previously for *Pseudomonas aeruginosa*<sup>199</sup>. Briefly, ~500 bp upstream and downstream of target gene were amplified by standard PCR and spliced together by overlap-extension PCR. The resulting DNA fragment was ligated into the pEXG2 allelic exchange vector using standard cloning procedures (Table S 2). Deletion constructs were transformed into *E. coli* SM10 and subsequently introduced into *P. protegens* via conjugal transfer. Merodiploids were directly plated on LB (lacking NaCl) containing 5% (w/v) sucrose for *sacB*-based counter-selection. Deletions were confirmed by colony PCR in strains that were resistant to sucrose, but sensitive to gentamicin. Chromosomal point mutations or epitope tags were constructed similarly with the constructs harboring the mutation or tag cloned into pEXG2. Sucrose-resistant and gentamicin-sensitive colonies were confirmed to have the mutations of interest by Sanger sequencing of appropriate PCR amplicons.

### 8.5.4 *Pseudomonas* growth competition assays

Recipient *P. protegens* strains contained a  $\Delta$ *pppA* mutation to stimulate T6SS effector secretion and induce a ‘counterattack’ from *P. protegens* donor strains<sup>200</sup>. Recipient strains were also marked with a tetracycline-resistant, *lacZ*-expression cassette at a neutral phage site (*attB*) to differentiate from unlabeled donor strains.

Stationary-phase overnight cultures of *P. protegens* donors and recipients were mixed in a 1:1 (v/v) ratio and relative abundance of donor:recipient was determined by plating part of the



competition mixture on LB plates with 40  $\mu\text{g mL}^{-1}$  X-gal. Ten microlitres of each competition mixture was then spotted on a 0.45  $\mu\text{m}$  nitrocellulose membrane that was overlaid on a 3% LB agar plate and incubated face up at 30 °C. Competition spots were harvested after 20-25 hours by resuspending in LB and counting CFU by plating on LB agar with 40  $\mu\text{g mL}^{-1}$  X-gal. The final ratio of donor:recipient colony forming units were then normalized to the initial ratios of donor and recipient strains and reported as the  $\log_{10}$  of the competitive index.

### 8.5.5 Toxicity assays

Wild-type RhsA and the various RhsA truncations and site-specific mutants used in this study were cloned into the rhamnose-inducible pSCrhaB2-CV vector<sup>201</sup>. RhsI was cloned into the IPTG-inducible vector pPSV39<sup>202</sup>. The various RhsA expressing pSCrhaB2-CV plasmids were co-transformed into *E. coli* XL-1 Blue with pPSV39::*rhsI*. Stationary-phase overnight cultures containing these plasmids were serially diluted  $10^{-6}$  in 10-fold increments and each dilution was spotted onto LB agar plates containing 0.1% (w/v) L-rhamnose, 500  $\mu\text{M}$  IPTG, trimethoprim 250  $\mu\text{g mL}^{-1}$  and 15  $\mu\text{g mL}^{-1}$  gentamicin. Photographs were taken after overnight growth at 37°C.

### 8.5.6 Protein expression and purification

Wild-type RhsA, its N-terminal (C538A, D322A, H530A, D318-320A) and C-terminal autoproteolysis mutants (D1346A, D1364N) were cloned into MCS-1 of pETDuet-1 and co-expressed with RhsI, which was cloned into MCS2. Plasmids were co-transformed into *E. coli* BL21 Codon Plus alongside a pET29b vector expressing EagR1.

### 8.5.7 Purification of RhsA $\Delta$ TMD and VgrG1 for cryo-EM

RhsA $\Delta$ TMD was co-expressed with RhsI using pETDuet-1 (see Table S 2 for details). VgrG1 was expressed in isolation using pET29b. Plasmids were individually expressed in *E. coli* BL21 Codon Plus. Strains harboring pETDuet-1 expressing RhsA $\Delta$ TMD-RhsI or pET29b expressing VgrG1 were inoculated in separate flasks containing 100 mL LB with selection and incubated overnight in a shaking incubator at 37°C. Following 14-18hr incubation, the culture was subinoculated (1/50 dilution) into four flasks each with 1 litre of LB and appropriate antibiotic selection. Cultures were initially incubated at 37°C until the culture reached an OD of ~0.3. The incubator was subsequently cooled to 18°C and each culture induced with 1 mM IPTG upon reaching an OD of 0.6-0.7. Cultures were harvested by centrifugation at 9,800 g for 10 minutes. Pellets were resuspended in 30 mL lysis buffer (50 mM Tris-HCl, 300 mM NaCl, 10

mM imidazole) and lysed by sonication (6 x 30 second pulses, amplitude 30%) and then spun at 39,000 g. Cleared lysates were applied to a Ni-NTA gravity flow column equilibrated using lysis buffer. The column was washed with the lysis buffer three times and the samples were eluted in 3 mL of elution buffer (lysis buffer with 400 mM imidazole). The samples were applied to a HiLoad 16/600 Superdex 200 column equilibrated in 20 mM Tris-HCl pH 8.0 150 mM NaCl and collected in the same buffer. The sample was reinjected into a Superose 6 5/150 column using the same buffer for EM analysis and *de novo* protein sequencing.

### 8.5.8 Western blot analyses

Western blot analyses were performed using rabbit anti-RhsA (diluted 1:5,000, Genscript, custom polyclonal<sup>150</sup>), rabbit anti-FLAG (diluted 1:5,000, Sigma), rabbit anti-VSV-G (diluted 1:5,000; Sigma) and mouse anti-His<sub>6</sub> (diluted 1:5,000, Genscript) and detected with anti-rabbit or anti-mouse horseradish peroxidase-conjugated secondary antibodies (diluted 1:5,000; Sigma). Development of western blots was completed using chemiluminescent substrate (Clarity Max, Bio-Rad) and imaged with the ChemiDoc Imaging System (Bio-Rad).

### 8.5.9 *De novo* protein sequencing via LC-MS/MS

RhsA $\Delta$ TMD was purified as described above. A concentrated sample was applied onto an SDS-PAGE. The band corresponding to the N- and C-terminally autoproteolyzed Rhs<sub>stage</sub> was excised from the gel and analysed by LC-MS/MS at the Proteome Factory AG company.

### 8.5.10 Negative stain electron microscopy

Four microliters of sample at a concentration of 0.005 mg/ml were applied to freshly glow-discharged carbon-coated copper grids. After 90s incubation time, excess sample was blotted away with Whatman No. 4, then washed twice with four microliters purification buffer and once with 0.75% (w/v) uranyl formate. A second batch of staining solution was incubated on the grid for 90s before excess was again blotted away. Grids were air-dried and imaged on a JEOL JEM-1400 microscope, equipped with a LaB<sub>6</sub> cathode and 4k x 4k CMOS detector F416 (TVIPS), operating at 120 kV.

### 8.5.11 Sample vitrification and data collection

Three microliters of the EagR1-RhsA-VgrG1 complex, at a concentration of 0.1 mg/ml, were applied to a freshly glow-discharged holey carbon grid (QF 2/1 200-mesh). The grid was blotted for 3s (blot force -5, drain time 0.5 s, 8 °C, 100 % humidity) and immediately plunged into

nitrogen cooled liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). Data collection was performed on a  $C_s$ -corrected Titan Krios (Thermo Fisher Scientific) operating at 300 kV in an automated fashion using EPU (Thermo Fisher Scientific). Movies were recorded on a Falcon 3 detector in linear mode at a nominal magnification of 59,000x with a calibrated pixel size of 1.1 Å/pixel. Image stacks were acquired in a defocus range from -1.2 to -2.2  $\mu\text{m}$  with an accumulated dose of  $90 \text{ e}^-/\text{Å}^2$  fractionated over 40 frames with a total exposure time of 1.5 s.

For RhsA $\Delta$ TMD, three microliters of sample, at a concentration of 4 mg/ml, were applied to a freshly glow-discharged holey carbon grid (QF 1.2/1.3 200-mesh). The grid was blotted for 3s (blot force -5, drain time 0.5 s, 8 °C, 100 % humidity) and immediately plunged into liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). The grid was transferred to a Titan Krios (Thermo Fisher Scientific) operating at 300 kV equipped with a GIF BioQuantum energy filter (Gatan), set to a slit width of 20 eV, and K3 Summit Detector (Gatan). Movies were recorded in counting mode at a nominal magnification of 105,000x with a calibrated pixel size of 0.91 Å/pixel in an automated fashion using EPU (Thermo Fisher Scientific). Image stacks were acquired in a defocus range from -0.8 to -2.2  $\mu\text{m}$  with an accumulated dose of  $61 \text{ e}^-/\text{Å}^2$  fractionated over 20 frames with a total exposure time of 2 s.

Both datasets were monitored live with TranSPHIRE<sup>193</sup> to evaluate i.e. the defocus range and astigmatism. Pre-processing was performed on-the-fly in TranSPHIRE including drift-correction and dose-weighting with MotionCor2<sup>194</sup>, CTF estimation on dose-weighted micrographs with CTFFIND4<sup>195</sup> and picking using the general model of crYOLO<sup>191</sup>.

#### 8.5.12 Cryo-EM image processing

After preprocessing in TranSPHIRE all processing steps were carried out in the SPHIRE<sup>190</sup> software package unless otherwise stated. Images with a resolution limit less than 6 Å were unselected using the graphical CTF assessment tool in SPHIRE for both datasets.

In case of the PFC, particles were extracted with box size of 408 x 408 pixels from 1250 good micrographs. Reference-free 2D classification and cleaning of the dataset was performed with the iterative stable alignment and clustering approach ISAC<sup>192</sup> implemented in SPHIRE. ISAC was performed at pixel size of 6.29 Å/pix. Using the Beautifier tool, the original pixel size was restored creating sharpened 2D class averages showing high-resolution details. A subset of particles showing clear high-resolution details were selected for structure refinement. 3D refinement was performed in MERIDIEN imposing  $C_3$  symmetry with a 25 Å lowpass-filtered

reference of our previously determined VgrG1 structure (EMD-0136). The two half-maps were combined with the PostRefiner tool in SPHIRE using a soft mask and automatic estimation of B-factors.

For RhsA $\Delta$ TMD, particles were extracted from 13,090 good micrographs with a final box size of 288 x 288 pixels. Particles were subjected to ISAC which was performed at pixel size of 3.3 Å/pix. A subset of particles showing high-resolution features were selected from beautified class averages. An initial reference for refinement was generated from the beautified averages using RVPER. All refinements and classifications were performed with implied  $C_2$  symmetry. Subsequent 3D refinement of the good particles with MERIDIEN yielded a map of overall 3.6 Å resolution but showed clear resolution anisotropy in the peripheries of the barrels. We imported the particle stack into RELION 3.1.0<sup>196</sup> with projection parameters obtained from MERIDIEN. Particles were classified into 4 classes without image alignment (T=4,  $C_2$  symmetry, soft mask) using the map from MERIDIEN as reference low-pass filtered to 30 Å. The particles belonging to the class displaying the highest resolution were selected for another round of MERIDIEN after removal of duplicated particles (minimum inter-particle distance threshold 100 Å). Particles were further CTF-refined and polished in Relion. The final refinement was performed with MERIDIEN. The final map was evaluated using the 3D FSC tool<sup>203</sup>. Map for Figure 45 and was postprocessed with the DeepEMhancer<sup>204</sup> using the high-resolution model. Local resolution was estimated with a normal postprocessed map in SPHIRE.

### 8.5.13 Model building, refinement and validation

The previously obtained VgrG1 structure from *P. aeruginosa* (PDB ID: 6H3L) was docked into the map as rigid body in UCSF Chimera<sup>205</sup>. The *P. protegens* VgrG1 sequence was manually adjusted in Coot<sup>206</sup> and iteratively refined in Phenix<sup>207</sup> and ISOLDE<sup>208</sup>. Model validity was assessed in Phenix with MolProbity<sup>209</sup>. Final model statistics are given in Table 4.

For RhsA $\Delta$ TMD, initial backbone traces were identified with automated model building software Buccaneer<sup>210</sup>. The model was manually adjusted to completion in Coot<sup>206</sup> using both the DeepEMHanced and a normal postprocessed map to exclude any bias. Stabilized  $\beta$ -strands of the toxin domain were identified via secondary structure prediction and manually placed inside the densities. The single interpretable  $\alpha$ -helix was not predicted by secondary structure predictions. Instead, we manually check the rest of the sequence to determine the sequence register. We placed residues 289-302 in this density guided by AlphaFold<sup>211</sup> predictions. This

was moreover guided by the matched hydrophobicity towards the interacting Rhs repeats. To identify residues corresponding to the linker density which complements the plug domain, we searched in the vicinity of the helix sequence for residues encoding bulky side chains. F271 served as anchor point to build residues 268-275 guided by AlphaFold<sup>211</sup> predictions.

Iterative refinement of the model in Phenix<sup>207</sup> and ISOLDE<sup>208</sup> was performed until convergence. Model validity was assessed in Phenix with MolProbity<sup>209</sup>. Final model statistics are given in Table 4.

Figures were prepared in Chimera X<sup>212</sup>. Multiple sequence alignments and secondary structure predictions were calculated using the MPI Bioinformatics toolkit<sup>213,214</sup> and visualized for creation of figures in Jalview<sup>215</sup>.

The cryo-EM maps of VgrG1 and RhsA have been deposited in the Electron Microscopy Data Bank (EMDB) under the accession codes of EMD-13843 and EMD-13867, respectively. The refined models for VgrG1 and RhsA were uploaded in the PDB and the entries have the IDs 7Q5P and 7Q97, respectively.

## 9 Results and discussion

### 9.1 Structural characterization of the Scap-Insig complex

The SREBP signaling pathway has been discovered more than 20 years ago by Brown and Goldstein<sup>85</sup> for which they were awarded the Nobel Prize in Medicine/Chemistry. However, structural determination of the membrane proteins of this pathway has remained challenging. One of the reasons might be, that even though there have been tremendous advances in the structural determination of membrane proteins using cryo-EM, which is reflected by the number of new structures solved, the protein production and sample preparation are still major bottlenecks. Handling mammalian membrane proteins outside their native environment and in aqueous detergent solutions proves to be particularly challenging. Additionally, many membrane proteins from higher organisms necessitate post-translational modifications or specific chaperones to achieve proper folding.

In the last decade, the use of eukaryotic expression systems to overexpress membrane proteins from higher organisms such as mammalian membrane proteins has increased.<sup>189</sup> However, technical challenges remain often for membrane protein targets below 150 kDa in size. Such proteins often lack large and rigid extramembranous domains that can facilitate the particle alignment during cryo-EM image processing.

The number of structures solved of proteins of the SREBP signaling pathway is lagging behind compared to other molecular systems. This is explained by multiple aspects, including challenges in protein overexpression, purification, grid preparation and image processing. During my doctoral thesis, I employed state-of-the-art methods to overcome these challenges and to lay the foundation for structural determination of Scap, Insig and SREBP. The ultimate goal would be to capture the structure of the tripartite complex composed of the three proteins or subcomplexes composed of two of these proteins. Prior attempts to determine the structures of these proteins alone and not in complex have been attempted by myself and other members of the group, which have proven difficult due to their relatively small size. Thus, I sought to overcome the size limitation and increase the size. I exploited the natural capability of these proteins to form hetero-oligomeric assemblies. Specifically, I tried to co-express these proteins. This allowed for complex formation within the native membrane environment prior to purification, thereby increasing the chances of successful structure determination via cryo-EM.

### 9.1.1 Small-scale expression of the human SREBP, Scap and Insig complex

Since human proteins are most relevant from a medical perspective, I directly chose constructs for overexpression in HEK293 GnTI<sup>-</sup> cells of the human full-length proteins SREBP, Scap and Insig.

Several approaches exist for overexpression of complexes in mammalian cells. Some complexes rely on the precise stoichiometric expression of individual proteins to ensure correct assembly. The MultiBac system<sup>216</sup> has been useful for the assembly of expression plasmids. This system involves assembling expression vectors from donor and acceptor vectors, each carrying their respective gene of interest, as well as promoter and terminator regions. Additionally, these vectors contain specific recombination sites that allow the hierarchical assembly of an expression vector encompassing the genetic information for all subunits of a complex. The expression vector incorporates a cassette enabling the generation of a baculovirus capable of transducing mammalian cells. This should allow the stoichiometric expression of all components of complexes within the same cell.

For the purpose of overexpression, I have assembled expression vectors comprising various combinations of the proteins of the SREBP signaling pathway. Vectors were designed to accommodate either two proteins (Scap-Insig and Scap-SREBP) or all three proteins (SREBP-Scap-Insig) in a single vector. To facilitate later purification procedures each protein of interest was fused in-frame with a purification tag. Distinct fluorescent proteins were also included in the constructs allowing the detection during expression and purification.

The obtained vectors were used for BacMam virus generation and subsequent transduction of HEK293 GnTI<sup>-</sup> cells aiming to achieve high levels of overexpression. However, during preliminary expression tests, the expression rates of all proteins were very low which could not be optimized. Moreover, coexpression of Scap-SREBP resulted in significant degradation of SREBP which could not be effectively inhibited. This is likely due to enhanced escort of SREBP to the Golgi apparatus in the absence of Insig. Furthermore, it is common that constructs require several rounds of optimization which includes generation of truncations, introduction of point mutations and change of purification tags. Several steps are involved in the hierarchical assembly of expression vectors using this procedure. It requires first modification of the donor and acceptor vectors and then the assembly of the final expression vector which is overall time consuming.

A more flexible approach involves monocistronic BacMam virus generation for each construct and subsequently mixing different viruses in stoichiometric ratio for transduction of mammalian cells. Such an approach has been successfully employed in insect and mammalian cells.<sup>217</sup> By

using this approach, it becomes feasible to rapidly modify the expression construct. Here, I created BacMam viruses that carry genetic information for overexpression of full-length human SREBP1a and Insig1 (hereafter only referred to as SREBP and Insig). SREBP was N-terminally fused to a StrepII-mCherry tag for purification and detection purposes followed by a 3C cleavage site that allows removal of the purification tag after its purification. Insig was N-terminally in-frame fused to a 3xFLAG-TagBFP tag followed by a 3C protease site. The construct for the human Scap included an N-terminally fused StrepII-His<sub>6</sub>-eGFP tag followed by a 3C site. This construct was kindly provided by Birte Weyers.

The viruses containing each a monocistronic expression cassette of SREBP, Scap and Insig were used to transduce HEK293 GnTI<sup>-</sup> suspension cells in small-scale in different combinations. Combinations involved mixtures of viruses to probe complex formation of Scap-Insig and SREBP-Scap-Insig. SREBP showed strong degradation upon coexpression of Scap, as described above. Therefore, I focused my work on complexes that involve Insig. Sterols, especially cholesterol and 25-HC, were shown to induce interaction of Scap and Insig.<sup>218</sup> Thus, cholesterol and 25-HC were added to the expression culture in final concentrations of 10µg and 1µg, respectively. The cells were harvested after 48h of transduction and directly solubilized by two of the most common mild detergents which are used for extraction and solubilization of membrane proteins, DDM and digitonin. The solubilized proteins were separated from insoluble material and applied to the respective affinity beads to enrich proteins and isolate potential complexes. Samples at each step of the small-scale purification were collected and loaded onto a semi-denaturing SDS-PAGE. The fluorescent proteins allowed detection in the gel of the SDS-PAGE. Additionally, proteins were detected using stain-free imaging which relies on fluorescence emitted by aromatic side chains, and further confirmed by Western blots against the purification tags. Only mCherry (SREBP) and eGFP (Scap) could be detected with the setup of the imaging system used in this study. Insig was detected by Western blot and stain-free imaging.

The solubilized fractions show presence of all three proteins Scap, Insig and SREBP judged by in-gel fluorescence and Western blot in the respective experiment (Figure 17). Scap and SREBP migrate both at 170 kDa in an SDS-PAGE due to their small size difference. Insig migrates at approximately 55kDa which matches the expected size. Differences can be seen for the solubilization efficiency for the used detergents DDM and digitonin. DDM generally showed stronger fluorescent bands, indicating that more material was solubilized compared to digitonin.



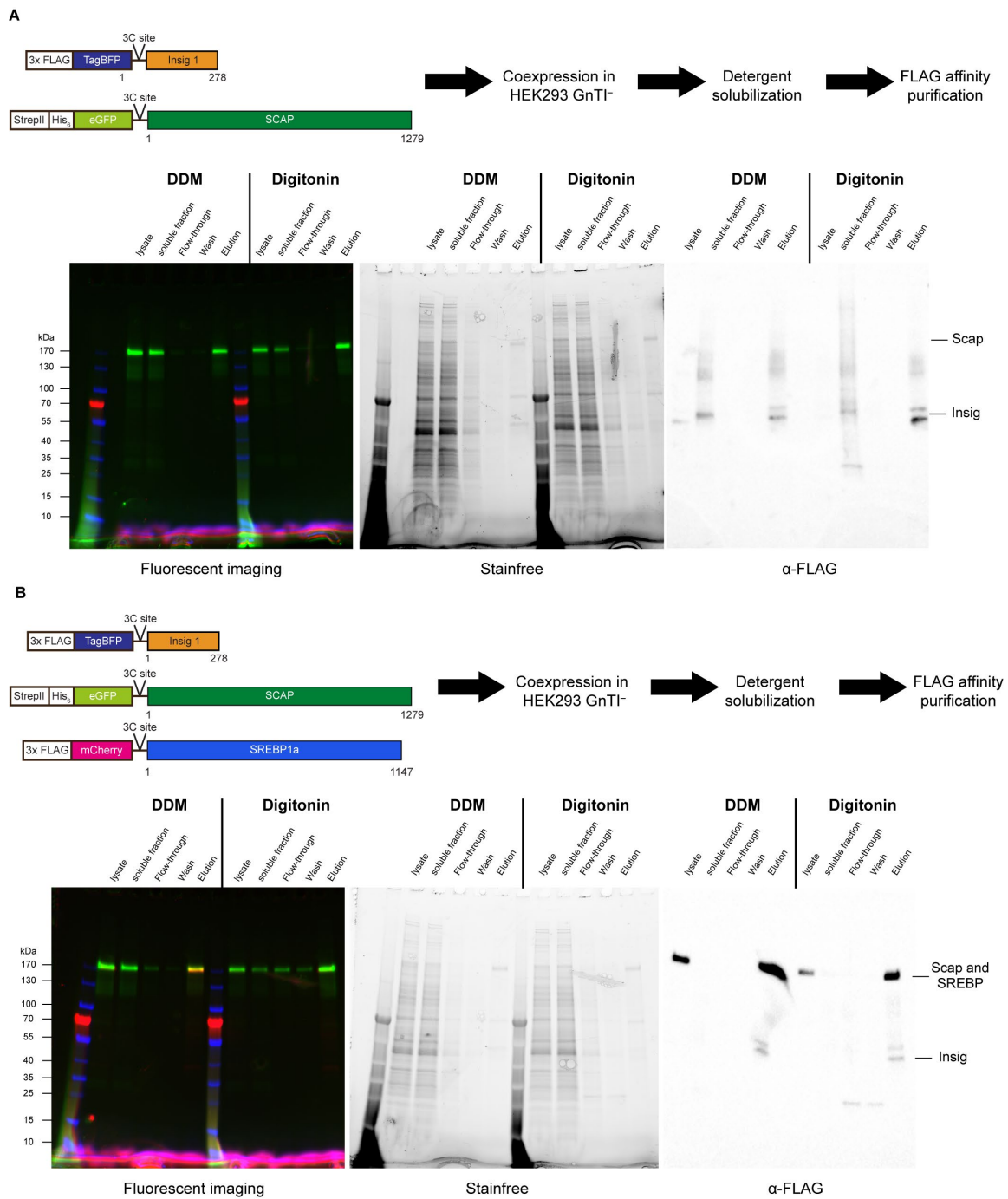


Figure 17: Small-scale expression of complexes of the SREBP signaling pathway.

Baculovirus-mediated overexpression in HEK293 GnTI<sup>-</sup> cells and small-scale purification of potential (A) Scap-Insig or (B) SREBP-Scap-Insig complexes. The gels in (A) and (B) show each the detection of protein bands by in-gel fluorescence, stain free detection and Western blot against the Flag tag. The cells were harvested after 48h and subsequently disrupted by solubilization with DDM or digitonin, as indicated. Insoluble material was removed and the soluble fraction was applied to FLAG-M2 affinity beads. The flow-through was discarded and beads were washed with two column volumes of wash buffer. The bound proteins were eluted with 3XFLAG peptide. Samples of each step of the purification were loaded onto a semi denaturing SDS-PAGE and visualized as indicated. For Western blot, proteins were transferred to a PVDF membrane. FLAG-tagged proteins were detected by HRP-mediated luminescence.

Interestingly, Scap could specifically be enriched by overexpression with Insig and subsequent Insig-mediated pull-down (Figure 17A). Very low amounts of Scap could be detected in the flow-through of DDM solubilized samples indicating that this detergent is mild enough to maintain intact Scap-Insig complexes isolated from overexpressed membranes. Digitonin was lower in solubilization efficiency and lead to more Scap that could be detected in the flow-through and washing steps.

When Scap was overexpressed in conjunction with Insig and SREBP, it co-precipitated in pull downs using the FLAG-tag present on both SREBP and Insig (Figure 17B). Scap was observed in the elution fractions indicated by the green fluorescent band that migrates at the same size as SREBP in an SDS-Page. The presence of SREBP was confirmed via a FLAG specific western blot. In the case of both detergents, a minimal amount of Scap was detected in the flow-through, with a slightly higher amount when solubilized in digitonin. As both SREBP and Insig possess a FLAG-tag in this experimental setup, it is challenging to determine whether Scap bound to each protein individually or formed a complex involving both proteins. Moreover, due to the single step purification it might be that potentially different subcomplexes formed such as SREBP-Scap, Scap-Insig and SREBP-Scap-Insig. Giving the specific isolation of Scap mediated by Insig, the focus of this work was shifted towards scaling up the expression and purification of the Scap-Insig complex.

### **9.1.2 Initial expression and purification of the Scap-Insig complex**

After an initial demonstration that Scap specifically bound to Insig, the expression was scaled up and it was tried to get a homogeneous complex for structural characterization. For this purpose, commonly two liters of HEK293 GnTI<sup>-</sup> suspension cells were transduced with viruses encoding for the human Scap and Insig as described above. Furthermore, cholesterol and 25-HC were included already in the expression step and were supplemented in all buffers during the purification. Expression showed strongest fluorescence signals of both proteins after 48h post-infection at which point cells were harvested. Whole cells were lysed and solubilized by addition of 1.5% (w/v) DDM-CHS mixture. This detergent mixture has been shown in the previous chapter to efficiently solubilize both proteins, and it furthermore stabilized the complex intact during purification. A single affinity purification step was performed with FLAG-M2 affinity beads. Samples at each step of the purification were taken and examined by semi-denaturing SDS-PAGE.

The SDS-PAGE in Figure 18 shows good and efficient solubilization of Scap judged by green fluorescence. Scap migrates at 170 kDa matching its expected size. Insig could not be detected

by in-gel fluorescence due to the imaging set-up. Scap could also be detected in the flow-through fraction, indicating that not all Scap molecules were bound to Insig. Small amounts of Scap were also detected in the wash fraction. The elution fractions show the presence of three major protein bands as well as some minor bands, whereas the latter likely correspond to unspecific contaminations. Two protein bands corresponding to Scap and Insig which migrate at 170 and 55 kDa, respectively, were observed in stoichiometric amounts. Another protein band migrating at approximately 70 kDa could be also observed. This band corresponds to the heat-shock protein 70 (HSP70). HSP70 was found to be a common co-purified protein during the purification of Scap alone. The identification of HSP70 was performed by Birte Weyers using LC-MS/MS analysis.

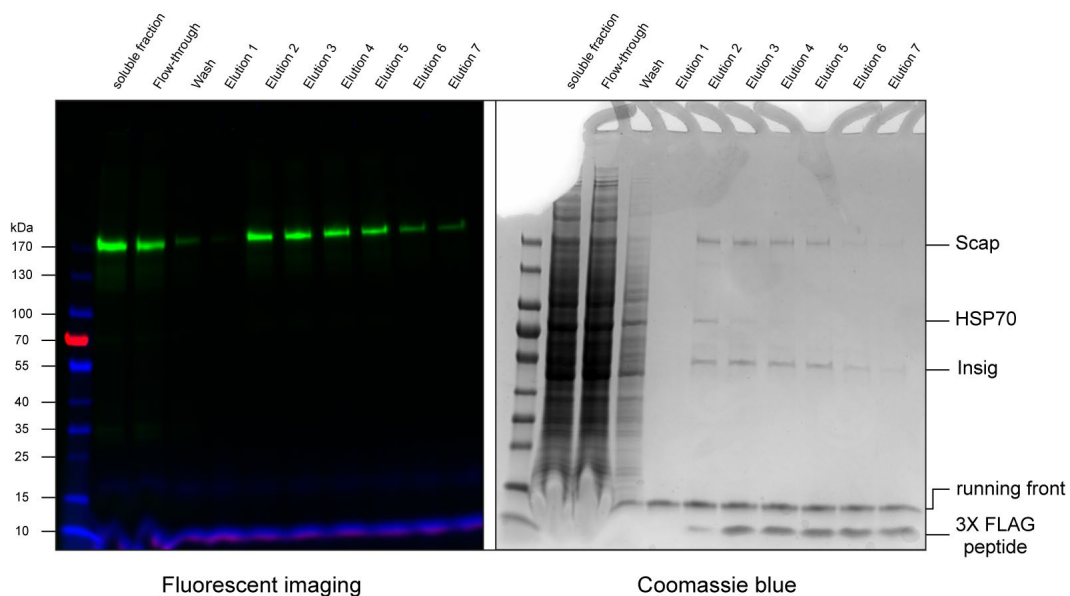


Figure 18: Initial one-step affinity purification of the Scap-Insig complex.

The proteins Scap and Insig were overexpressed in HEK293 GnTI<sup>-</sup> suspension cells and membrane proteins were solubilized with the detergent DDM-CHS. Insoluble material was separated and the soluble fraction was applied to FLAG-M2 affinity beads. The beads were washed and bound proteins were eluted in seven fractions with the 3xFLAG peptide. Protein bands were visualized by in-gel fluorescence or Coomassie blue staining.

Nonetheless, sufficient amounts of Scap were pulled down by Insig providing evidence for their interaction that persisted during detergent extraction and isolation from overexpressed cells. To further characterize the isolated complex and remove unspecific contaminations and aggregates, the elution fractions from the single affinity chromatography step were combined, concentrated and injected into a Superose 6 column. Fluorescent proteins and purification tags were not removed allowing fluorescent detection during size-exclusion chromatography (SEC). The fluorescent traces of the SEC run are shown in Figure 19A. A void volume which shows fluorescent signal for both proteins is observed at approximately 8 ml. Then, multiple peaks

over an elution volume from 10-17.5 ml are observed with different intensity signals for Scap and Insig. Excess Insig can be detected at approximately 16 ml elution volume. Since the complex was purified via a single FLAG affinity chromatography step, excess amounts of Insig are expected. When Insig purified alone, Insig elutes as bell-shaped peak at 16 ml (Figure S 2A). When Scap is purified alone, it shows a gaussian-shaped peak at 15 ml elution volume (Figure S 2B).

Interestingly, a fluorescent signal with comparable intensity for Scap and Insig was found at an elution volume of 12 ml. It is worth mentioning that that neither of the proteins were detected at this elution volume when purified alone, suggesting the formation of a higher molecular weight complex. Subsequent SDS-PAGE analysis of the fraction at 12 ml shows the presence of both proteins Scap and Insig (Figure 19B). An additional band corresponding to HSP70 was also detected. But also other contaminants are visible. Overall, the yield of the complex was rather low at this stage with approximately 0.05 mg per liter of expression culture.

In order to evaluate the sample quality further, the peak fraction at 12 ml which contains Scap and Insig was analyzed by negative stain EM (Figure 19C). The visualized particles are heterogenous in size and shape. Smaller particles can be observed at around 5-10 nm in size. These particles could correspond to empty detergent micelles as well as to particles representing the individual proteins co-purified or from complexes fallen apart during the purification. A larger particle species is observed ranging from 20-25 nm in diameter. The larger particles display various shapes including mostly globular particles. Based on their size, these assemblies could harbor multiple copies of Scap and Insig. To further validate that Scap is part of these larger particles, the His-tagged Scap was labelled with 5 nm Ni-NTA gold particles allowing the visualization in negative stain EM (Figure 19D). Indeed, the larger particles exhibited binding to one to four gold particles, suggesting the presence of various complex species with varying copies of Scap. However, given the limited resolution obtained with negative stain EM, it is not possible to determine the precise amount of Scap copies present in the particles. Rather, it can be qualitatively confirmed that Scap is indeed part of the larger sized particles.

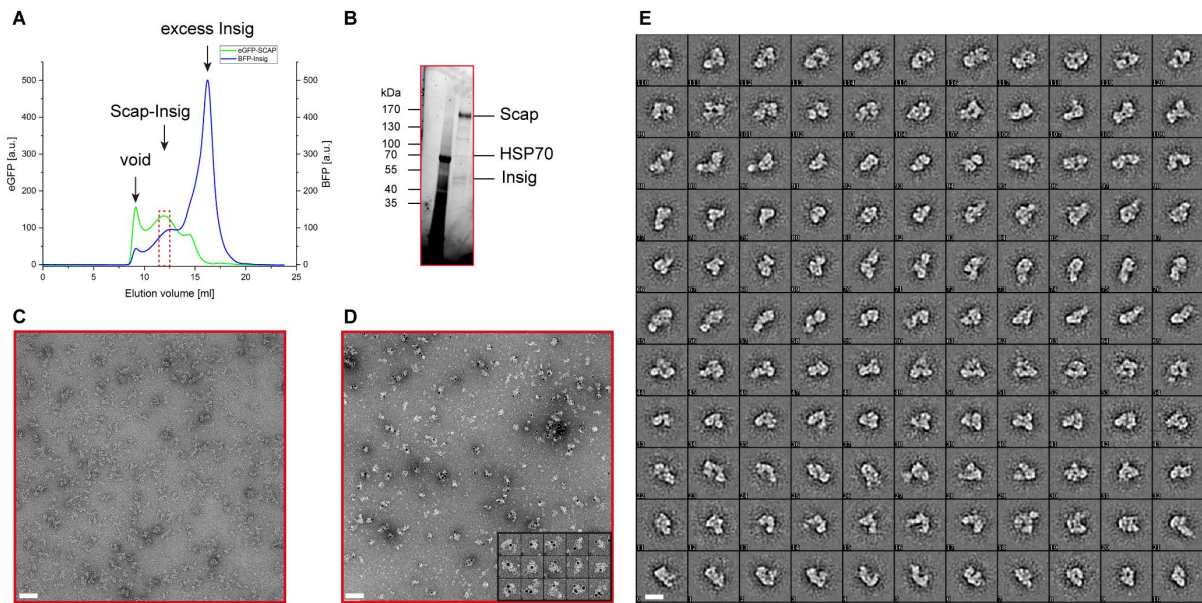


Figure 19: Preliminary negative stain analysis of the potential Scap-Insig complex.

(A) Fluorescence-detection size-exclusion chromatography of the purified Scap-Insig complex. The red-dotted fraction was loaded onto an (B) SDS-PAGE. The gel shows multiple weak bands that include Scap, Insig and HSP70. Other contaminants are also present but are not labelled. Protein bands were visualized by stain-free detection. (C) Negative stain micrograph of the peak fraction containing Scap and Insig. Scale bar, 50 nm. (D) Negative stain micrograph of the 5 nm Ni-NTA gold labelled sample. The sample was incubated with the gold particles and then applied to a negative stain grid. The inset shows enlarged particle views that are labelled with one or multiple gold particles. Scale bar of the micrograph, 50 nm. (E) Representative 2D class averages of the sample shown in (C). Scale bar, 20 nm.

To gain a better understanding of the overall shape of the particles and increase the signal-to-noise ratio, a small negative stain dataset was collected. Particle coordinates were manually selected and used for image processing. Particles were extracted and classified in two-dimensional space using ISAC<sup>192</sup> integrated in SPHIRE<sup>190</sup>. The 2D class averages show increased contrast compared to the raw particles and confirm the heterogenous shape of the particles (Figure 19E). Multiple different shapes can be observed which might represent different views of the same complex or different oligomeric assemblies. A reliable 3D model could not be generated due to the heterogeneity among the class averages.

In order to proceed with the structural determination with cryo-EM, optimization of the sample quality was required to increase the homogeneity. Moreover, initially I performed only a single-step affinity purification. A tandem-affinity purification workflow pulling on both sides of the complexes should drastically improve the homogeneity of the complex.

### 9.1.3 Optimization of the purification workflow of the Scap-Insig complex

With the aim of improving the sample quality, I explored multiple different conditions to stabilize the Scap-Insig complex. Despite successfully confirming the stable interaction between Scap and Insig during small-scale and large-scale purifications, there remained a concern that the complex could potentially disintegrate when handled in an aqueous detergent solution after the proteins were solubilized from the membrane. To overcome this problem, all purification buffers were supplemented with cholesterol and 25-HC. However, their solubility even in detergent solution is low. Therefore, I thought to introduce a well-known mutation into the SSD of Scap (D428A), that displays interaction with Insig even in the absence of sterols<sup>107</sup> and thus increases the stability of the Scap-Insig complex. This mutation was introduced in the vector encoding full-length Scap (termed Scap<sup>D428A</sup> but hereafter only referred as Scap for simplicity) and overexpressed as described above with WT Insig (Figure 20A). The complex was solubilized and purified using the detergent mixture DDM-CHS by tandem-affinity chromatography and SEC. Purification tags were not cleaved for ease of detection. The fractions containing Scap and Insig coeluting in SEC were analyzed by SDS-PAGE and negative stain EM.

The SEC profile shows three main peaks, a void peak at 8 ml elution volume, a broad gaussian-shaped peak at 15 ml with a shoulder on the side towards higher molecular weight, and an additional peak at 19 ml corresponding to D-desthiobiotin which was used for elution from the last affinity column (Figure 20B). Fluorescence-detection SEC reveals that the elution profile of Insig appears as sharp gaussian shaped peak at 15 ml (Figure 20C). Opposingly, the peak for Scap elution is detected at a slightly earlier elution volume of 14 ml and is much wider. The fluorescent traces do show some overlap indicating the Scap and Insig partially coeluted. But the shape of the fluorescent traces also indicate that the complex might have fallen apart during the SEC or that different oligomeric species are present in the preparation that could not be separated from each other. The SDS-PAGE shows that three main protein bands can be detected in the main peak over an elution volume from 12-16 ml (Figure 20D). These protein bands migrate at 50, 70 and 170 kDa corresponding to Insig, HSP70 and Scap, respectively.



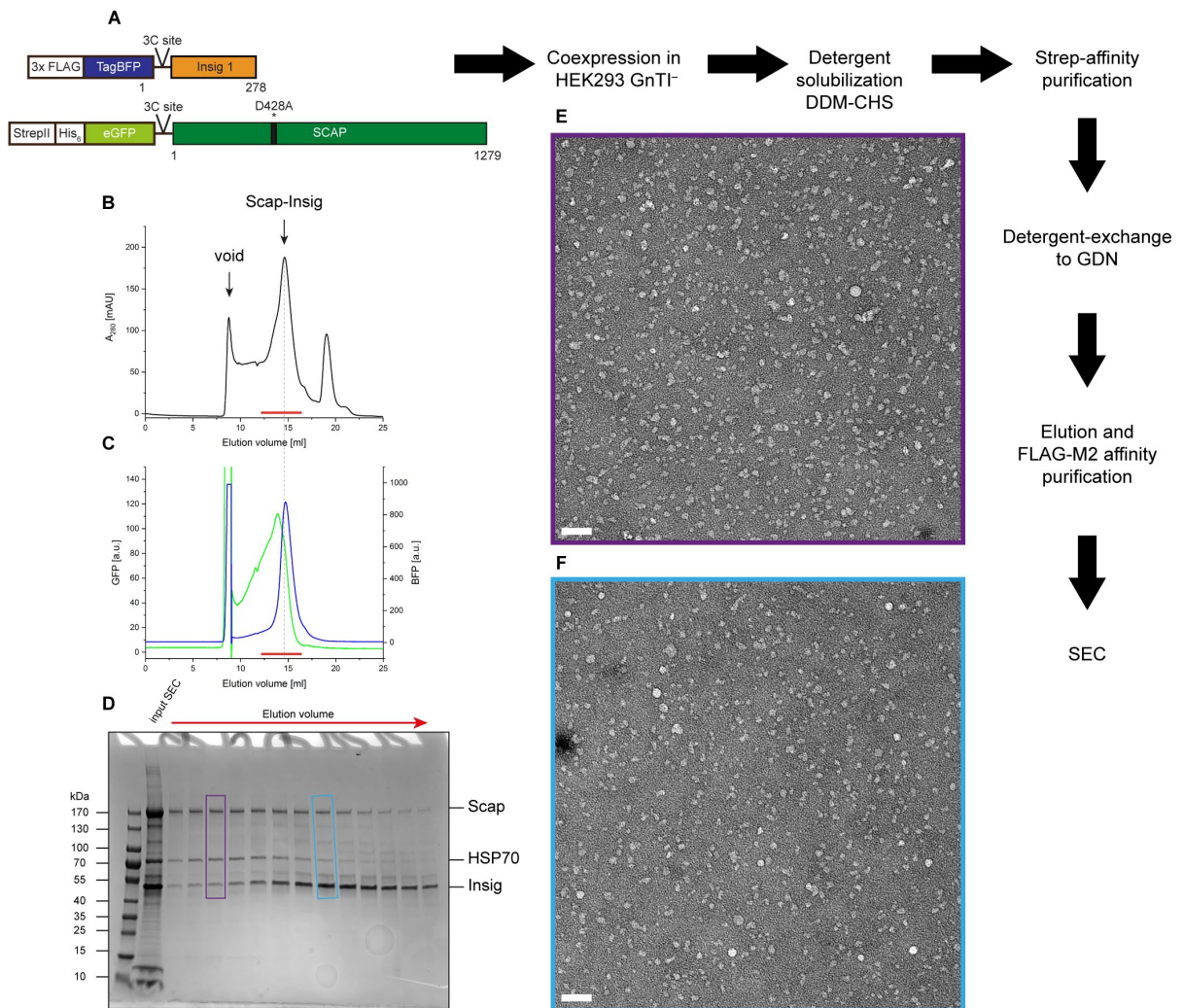


Figure 20: Tandem-affinity purification of the Scap<sup>D428A</sup>-Insig complex in DDM-CHS.

(A) The constructs including the mutant Scap<sup>D428A</sup> and WT Insig were coexpressed in HEK293 GnTI cells. The complex was solubilized and purified in DDM-CHS. The SEC profiles for the Scap-Insig complex in DDM-CHS were recorded by (B) A<sub>280</sub> absorption and (C) fluorescence. The green trace records eGFP fluorescence emanating from Scap. The blue trace represents the fluorescence detected from Insig. (D) The elution fractions indicated by the red bar were loaded onto a semi-denaturing SDS-PAGE and protein bands were visualized by Coomassie blue staining. (E) & (F) Negative stained EM micrographs of the fractions indicated on the SDS-PAGE in (D). Scale bar, 50 nm.

To further investigate the sample heterogeneity, two fractions from different elution volumes of the main peak where both Scap and Insig are co-eluted were taken and analyzed by negative stain EM. The sample taken from earlier elution volumes (13 ml) shows different sized and heterogeneous particles with a diameter ranging from 15-25 nm (Figure 20E). The second sample taken from later elution volume of the main peak (15 ml) shows a more homogenous size of the particles with an average diameter of the main species ranging from 12-16 nm (Figure 20F). However, also larger and smaller particles can be observed. Larger particles likely correspond to aggregated particles that could not be removed by SEC. Smaller particles likely

correspond to empty DDM micelles. DDM must be supplemented above the cmc to maintain membrane proteins in a soluble state in aqueous buffer. Hence, empty DDM micelle can be commonly visualized in negative stain EM when supplemented to the buffer.<sup>219</sup>

These results indicate that the overall yield and stability of the complex could be drastically improved by overexpressing a mutant Scap (Scap<sup>D428A</sup>) with Insig. The sample quality could be optimized by tandem-affinity purification pulling on both sides of the complex. However, heterogeneity remains observed by the rather broad fluorescent SEC elution trace of Scap indicating the presence of potential different sized oligomers of Scap either bound or unbound to Insig. The broad peak could also be an indication of dissociated complexes that fall apart during SEC. Furthermore, HSP70 appears to be consistently co-purified with the Scap-Insig complex in stoichiometric amounts. These results strongly suggest that further optimization of sample quality is required to increase the success rate of cryo-EM structural determination.

### 9.1.4 Stabilization of the Scap-Insig complex

Since the yield of the complex increased by overexpression of the mutant Scap<sup>D428A</sup>, I pursued stabilization of the complex by biochemical means instead of further changing the constructs. Before starting to purify membrane proteins, the ideal choice of the detergent cannot be predicted and has a drastic effect on the stability of the protein. The identification of the appropriate detergent for solubilization and purification for each protein or complex requires empirical testing and cannot be predicted beforehand. Moreover, detergents that work good for solubilization may not be the optimal choice for downstream applications such as structural determination. Thus, I explored the use of different detergents for solubilization and purification including detergent exchanges during the purification.

DDM-CHS was initially successful in solubilization and purification of the complex. However, DDM has a rather high critical micellar concentration (cmc) compared to other detergents such as lauryl maltose neopentyl glycol (LMNG), glyco diosgenin (GDN) or digitonin. Higher cmc detergents can produce a stronger background in cryo-EM imaging since more detergent needs to be supplemented which has been observed in negative stain EM (section 9.1.3). This can decrease the overall contrast of the images and may negatively affect the outcome of the structural determination. Therefore, I tried to solubilize the complex initially in DDM-CHS and then exchange the detergent at later stages during the purification which included detergent exchange during the wash on the first affinity column (Figure 21A). Among the different detergents tested for detergent exchange, GDN resulted in a pure and well behaving complex.



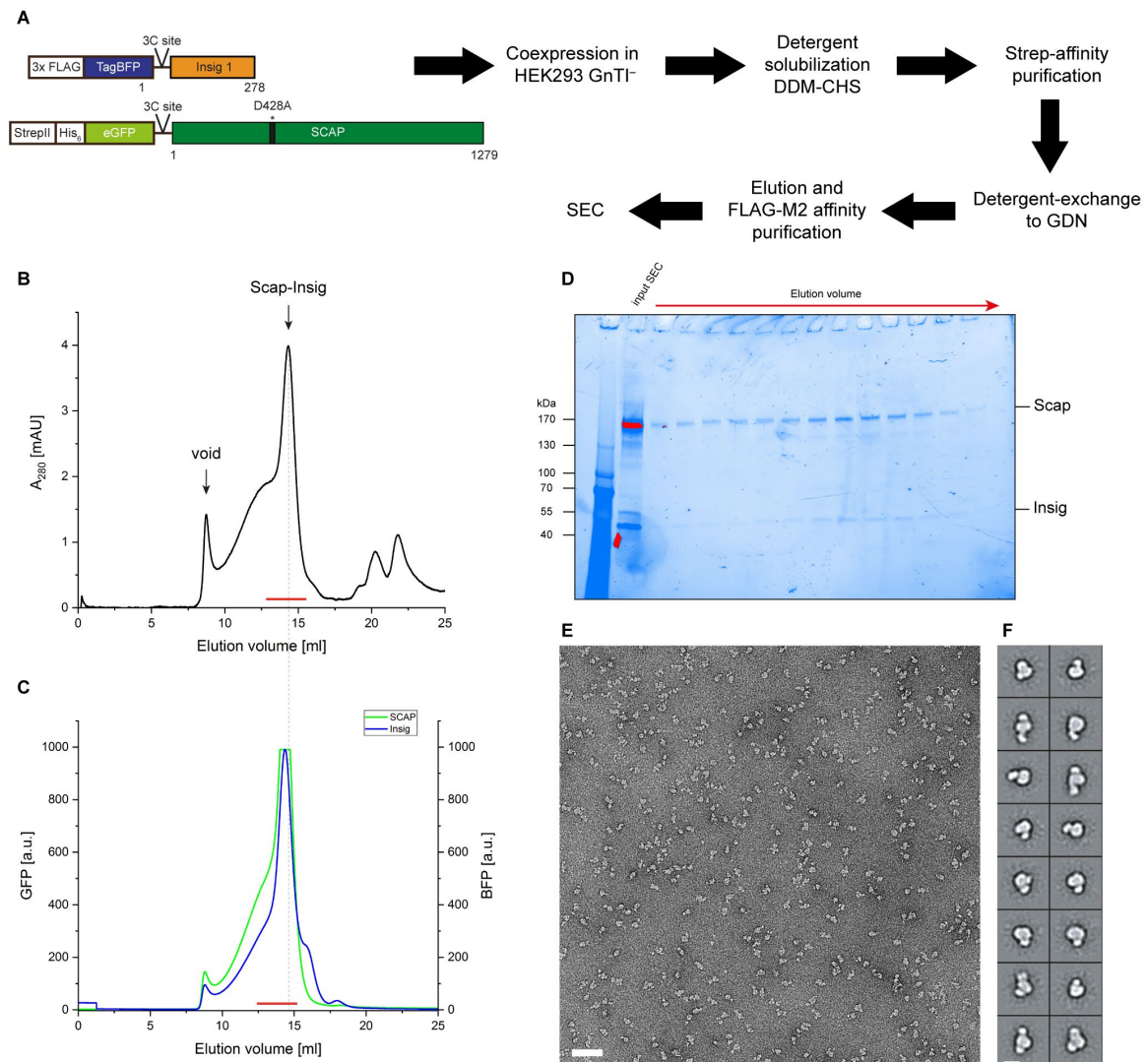


Figure 21: Tandem-affinity purification of the Scap<sup>D428A</sup>-Insig complex in GDN.

(A) The constructs including the mutant Scap<sup>D428A</sup> and WT Insig were coexpressed in HEK293 GnTI<sup>-</sup> cells. The complex was solubilized in DDM-CHS and exchanged to GDN during the first affinity purification step. The SEC profiles for the Scap-Insig complex in GDN were recorded by (B)  $A_{280}$  absorption and (C) fluorescence detection. The green trace records eGFP fluorescence emanating from Scap. The blue trace represents fluorescence detected emanating from Insig. (D) The elution fractions indicated with the red bar were loaded onto a semi-denaturing SDS-PAGE and proteins bands were visualized by stain-free detection. Saturated pixels are indicated in red. (E) Exemplary negative stain EM micrograph of the peak fraction containing the Scap-Insig complex. Scale bar, 50 nm. (F) Representative 2D class averages generated in SPHIRE<sup>190</sup>. Scale bar, 12 nm.

The complex was first purified by FLAG-M2 affinity beads. During the washing step DDM-CHS was replaced by the addition of GDN in the washing buffer. Proteins were eluted by the 3xFLAG peptide and directly applied to Strep beads. The immobilized material was again washed and eluted from the beads supplemented with GDN in the buffer. Eluted fractions were pooled, concentrated to 500 $\mu$ l and further purified via SEC using a Superose 6 column. Fractions of the SEC run were analyzed by SDS-PAGE and negative stain.

Two major peaks could be detected during SEC (Figure 21B). The first peak elutes at 8 ml and corresponds to the void volume in which aggregates elute. The second peak located at an elution volume of 14 ml displays a majorly gaussian shape with a notable shoulder towards higher molecular weights. Peaks observed at 20 ml and 22 ml correspond to buffer components. Fluorescence-detection SEC reveals signal for Scap and Insig coincided with an elution volume of 14 ml (Figure 21C). This was also confirmed by loading samples of the elution fractions onto a semi-denaturing SDS-PAGE (Figure 21D). The presence of the shoulder towards higher elution volumes suggests the existence of higher molecular weight species of Scap and Insig that cannot be separated from each other. Interestingly, HSP70 could no longer be detected suggesting that a detergent exchange to GDN increased the stability of the Scap-Insig complex.

A sample of the main peak eluting at 14 ml which contains both proteins was taken and analyzed by negative stain EM. A representative micrograph is shown in Figure 21E and displays a homogenous and uniform particle distribution. The particles appear globular shaped with a diameter of 12-14 nm. A smaller density can be always seen projecting away from the globular density. This particle likely corresponds to the complex embedded in the detergent micelle with additional non-membranous domains projecting out of the micelle. To confirm this, a small negative stain dataset was recorded. Particle locations were manually selected and subsequently extracted. The extracted particles were classified in 2D in order to increase the signal-to-noise ratio. Representative 2D class averages are shown in Figure 21F and reveal a homogenous particle species with an average diameter of 12-14 nm. The class averages feature globular or disc-shaped density corresponding to the detergent micelle engulfing the complex composed of Insig and Scap. In most of the classes an additional dot-like density located outside of the micelle can be observed. Since I used the full-length version of Scap, this domain likely corresponds to the extramembranous WD40 domain of Scap which would match the observed size. Interestingly, in some of the classes a second blob-like density can be observed on the opposing side of the micelle. This density must also be part of Scap, since the mass contributed by Insig exclusively originates from the TM helices and a small N-terminal but unstructured stretch composed of 78 amino acids. The N-terminal domain of Insig would be too small to fit the observed dimensions of the second blob-like density. Moreover, predictions regarding the orientation of both proteins in the membrane suggest that the N-terminus of Insig is facing the cytosol as well as the WD40 domain of Scap. Upon closer inspection of the domain organization of Scap, it is evident that Scap possesses two large loops L1 and L7 which are composed of 238 and 135 amino acids, respectively. These two loops were shown to interact

with each other under certain conditions.<sup>103</sup> A globular domain formed by interaction of these two loops would fit the observed size of a second blob-like density protruding out of the micelle. To unambiguously confirm the arrangement of Scap and Insig in the complex, a high-resolution structure is required. The improved sample preparation workflow and the increased sample quality has paved the way for structural characterization of the complex via cryo-EM.

### 9.1.5 Cryo-EM screening and data collection of the Scap-Insig complex in GDN

Several plunge-freezing conditions and grid types were tested in order to obtain a grid suitable for cryo-EM data collection. The best particle distribution was obtained when the complex was concentrated to 1.4 mg/ml while using QF 1.2/1.3 200-mesh copper grids. A representative cryo-EM micrograph is shown in Figure 22A revealing a homogenous particle distribution. The particles are vitrified in a thin monolayer of vitreous ice with good contrast. The ice thickness was measured to be in a range of 25-30 nm which is ideal for a particle with a size of 12-14 nm. This grid was selected for cryo-EM data collection using a Titan Krios G2 equipped with a K2 summit direct electron detector. In total, 5601 micrographs were collected in super-resolution mode with a pixel size of 0.84 Å/pix after two-fold Fourier-cropping. Particle coordinates were automatically detected with the general model of crYOLO<sup>191</sup> and used for subsequent particle extraction. The dataset was cleaned in two-dimensional space by 2D classification using ISAC embedded in SPHIRE<sup>190</sup> (Figure 22B). Particles present in 2D class averages that show clear protein-like densities were further selected for an initial model generation and subsequent 3D refinement.

The 2D class averages formed during image processing confirm the overall shape of the complex as seen in negative stain. Here, the resolution that is obtained is considerably higher compared to negative stain allowing the visualization of more features. A disc-like density can be observed which corresponds to the detergent-belt solubilizing the TMDs of the complex. Another globular shaped density protruding from the micelle can be observed which likely corresponds to the WD40 domain of Scap. Similarly, as seen in negative stain, less than half of the 2D classes show a second blob-like density protruding from the micelle which is always on the opposite site of the micelle relative to the other blob-like density. As mentioned above, this domain likely corresponds to the domain formed by loops L1 and L7.

The particles corresponding to the classes showing clear protein-like features were further selected for a 3D refinement. After 2D classification 50,693 particles were retained and refined against an initial reference. The final map was resolved to a resolution of approximately 15 Å (Figure 22C). At this resolution, no high-resolution information can be seen nor could any

densities corresponding to TMDs of both proteins be revealed. Accordingly, the micelle appears hollow from the inside. Overall, the 3D reconstruction confirms the observation from 2D classification of a disc-like density corresponding to micelle as well as an additional blob projecting out of the micelle. This blob likely represents the WD40 domain of Scap. In half of the 2D classes a second blob on the other side of the micelle could be observed. This domain is not resolved in the 3D reconstruction potentially indicating flexibility of this domain.

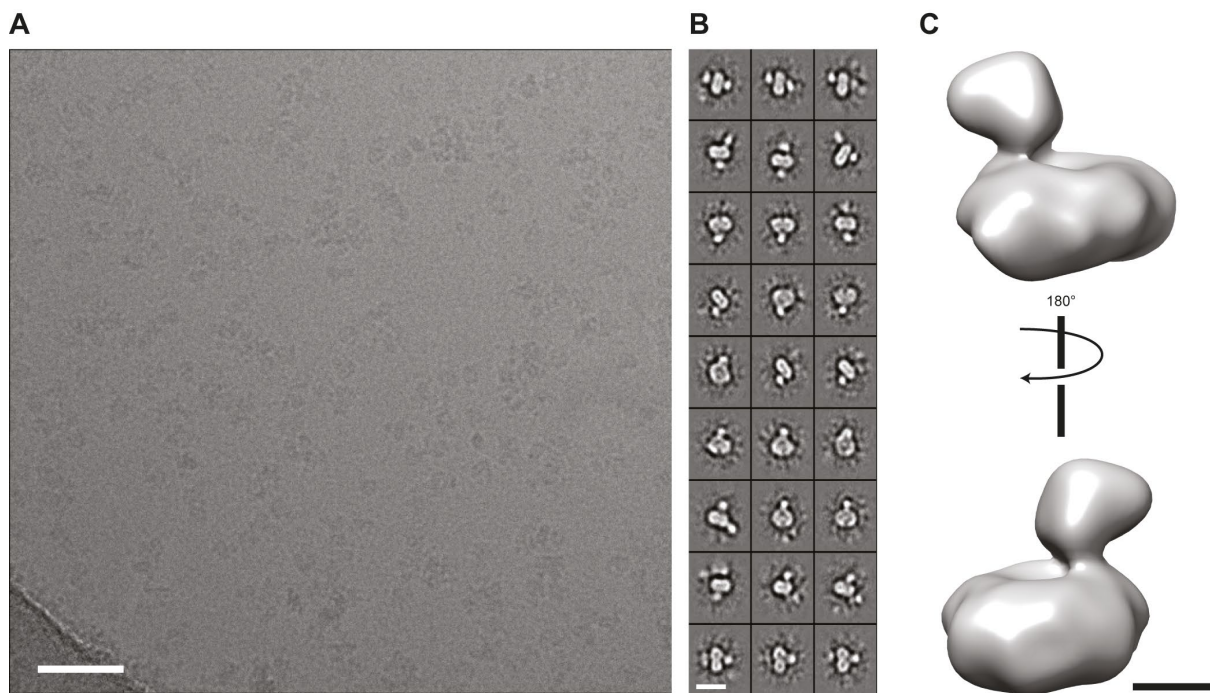


Figure 22: Single particles analysis of the Scap-Insig complex in GDN.

(A) Representative cryo-EM micrograph of the Scap-Insig complex purified in GDN. The complex was vitrified in a thin pure ice layer using QF 1.2/1.3 200-mesh grids. Scale bar, 50 nm. (B) Representative 2D class averages generated using ISAC integrated in SPHIRE<sup>190</sup>. Scale bar, 12 nm. (C) Final reconstruction of the Scap-Insig complex calculated with MERIDIEN integrated in SPHIRE<sup>190</sup>. Scale bar, 5 nm.

The limited resolution of the 3D reconstruction could be due to the limited number of good particles. To overcome this bottleneck, I collected a second dataset containing more than 9000 micrographs. Particles from both datasets were combined which resulted in 200K particles that were used to calculate a 3D reconstruction. Unfortunately, the resolution of the reconstructed map did not improve and still no higher resolution features could be observed. Therefore, I thought to further optimize the sample quality, which is most often the major bottleneck in SPA analysis of small membrane proteins.

### 9.1.6 Improved sample preparation of the Scap-Insig complex in LMNG

The sample quality was improved by the workflow that involved solubilization in DDM-CHS and detergent exchange to GDN. Although DDM-CHS was consistently utilized for the solubilization, I sought to further optimize the purification protocol by testing different detergents than DDM-CHS for solubilization. Ideally, detergent exchange during the purification can be circumvented to further ensure proper complex assembly. Another mild detergent commonly used for solubilization and structural determination of membrane proteins is LMNG. Here, I used a detergent mixture composed of the cholesterol derivative CHS, termed LMNG-CHS. LMNG has a much lower cmc than DDM. Thus, much less detergent can be supplemented to the purification buffer potentially leading to enhanced contrast in cryo-EM micrographs.

The complex composed of the mutant Scap (Scap<sup>D428A</sup>) was coexpressed with WT Insig. The purification was performed as discussed in chapter 9.1.5 with the modification that LMNG-CHS was used in all steps without the need for exchange of the detergent. After tandem-affinity purification and 3C treatment in order to remove the purification tags, the complex was further purified by injection into a Superose 6 column in buffer supplemented with 0.001%/0000.1% (w/v) LMNG/CHS. The fractions of the main peak were analyzed by semi-denaturing SDS-PAGE.

The SEC profile shows the appearance of one main peak at approximately 15 ml elution volume (Figure 23A) which contains Scap and Insig judged by SDS-PAGE (Figure 23B). Consistent with purifications described before, a shoulder towards higher molecular weight species overlapping with the main peak can be still observed. However, solubilization and purification in LMNG-CHS results in a better separation of the shoulder the shoulder from the main peak and elutes as individual peak at approximately 13 ml. Another peak can be observed at approximately 18 ml elution volume. This peak corresponds to the cleaved off fluorescent purification tags which were separated successfully from the main peak containing the Scap-Insig complex. A void volume can be detected at 8 ml but the height of the peak is negligibly low.

A sample of the main peak eluting at 15 ml was analyzed by negative stain EM. A representative is shown in Figure 23C and reveals a homogenous and uniform particle distribution. The particles range between 10-12 nm in diameter and show a mostly globular and disc-shaped particle with an additional blob-like domain projecting away. These particles are reminiscent to the observed particles when the complex was solubilized in DDM-CHS and exchanged to GDN. However, the particle distribution looks much more homogenous and uniform. To further



confirm this, a negative stain dataset was recorded and 2D class averages were calculated. The class averages show a disc-shaped particle with one or sometimes two small blob-like densities protruding (Figure 23D). The second blob-like part of the particle is always found on the opposite side of the disc-shaped part relative to the other blob-like domain.

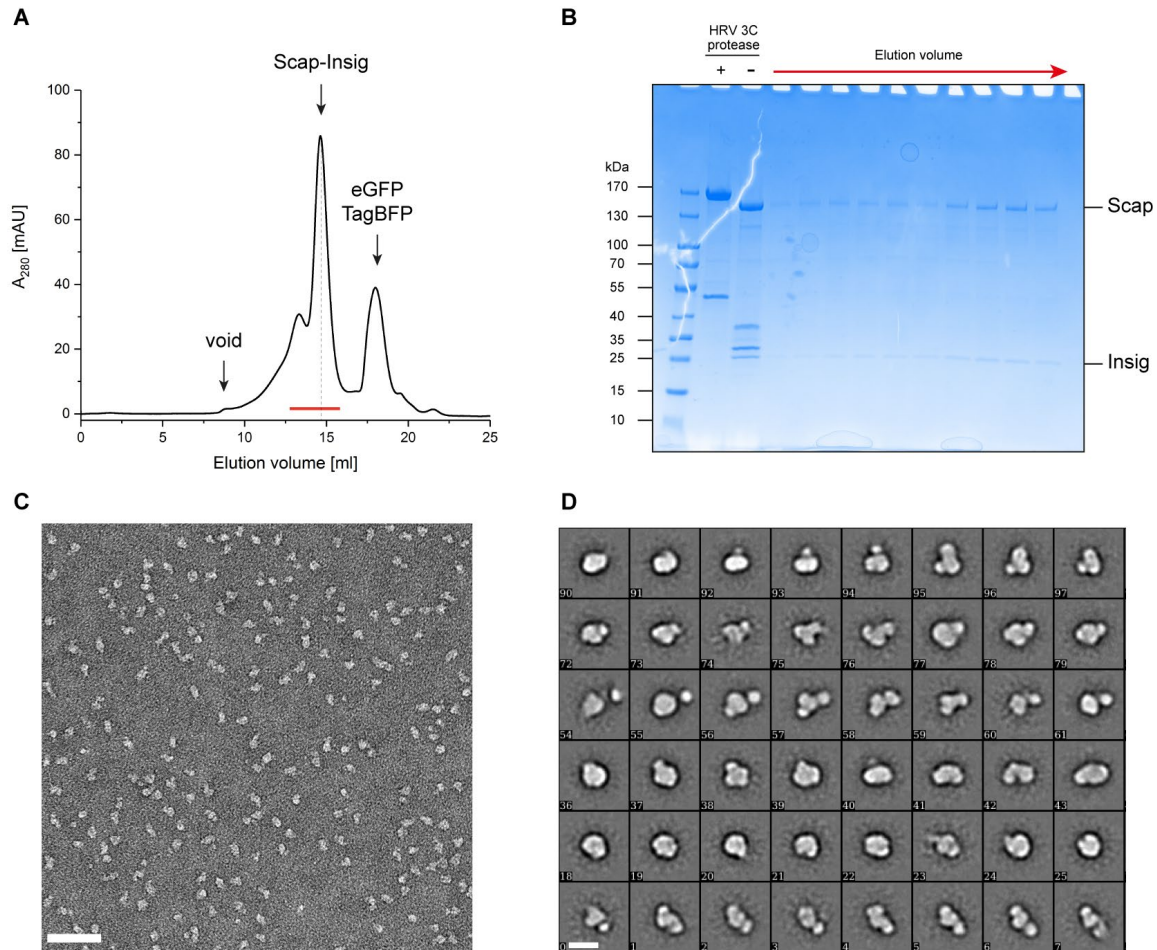


Figure 23: Optimized purification procedure of the Scap-Insig complex in LMNG-CHS.

(A) The SEC profile of the tandem-affinity purified Scap-Insig complex shows one gaussian-shaped main peak eluting at 15 ml. A small but separated shoulder is observed to earlier elution volumes right before the main peak. Cleaved off purification tags (eGFP, TagBFP) were successfully separated from the Scap-Insig complex. (B) SDS-PAGE shows the purity of the Scap-Insig complex. Both proteins coeluted at an elution volume of 15 ml. Protein bands were visualized by Coomassie blue staining. (C) Representative negative stain EM micrograph revealing a homogenous particle distribution. Scale bar, 50 nm. (D) Representative 2D class averages of the Scap-Insig complex. Scale bar, 12 nm.

To validate the sample quality and estimate the stoichiometry of the complex, I sought to determine the native mass of the purified complex. In SDS-PAGE, a harsh detergent is added to the purified components and the sample is heated up. This results in an unfolding of the proteins and dissociation of complexes. During gel electrophoresis, proteins can be separated based on their charge/mass ratio which is proportional to their size. This is efficient to analyze

the purity of a given sample but gives little information about the stoichiometry and oligomeric state. One such method to detect the oligomeric state and size of native membrane protein complexes is called blue native PAGE (BN-PAGE).<sup>220</sup> Following purification in an uncharged detergent, charge is added to the native complex by addition of Coomassie blue G-250. The hydrophobic and negatively charged dye will also bind to membrane proteins allowing their migration to the anode during gel electrophoresis. Proteins and protein complex are then separated based on their size during a gel electrophoresis.

A sample of the purified Scap-Insig complex in detergent was mixed with Coomassie G-250 and loaded into the wells of native PAGE. Electrophoresis was performed at 4° to prevent heating of the samples. The gel was recovered and proteins were fixed and stained with additional Coomassie stain and subsequently destained.

The gel of the BN-PAGE shows multiple bands (Figure 24A). The most prominent band migrates at an approximate size of 350 kDa. Several other but weaker bands are detected corresponding to higher molecular weight species. The observed mass accounts for contribution from proteins, the detergent micelle and the amount of accumulated dye molecules and are thus a rough estimate of the total complex size. Based on this limited analysis, roughly a 1:1 complex was estimated with 170 kDa stemming from Scap, 30 kDa from Insig, and 150 kDa coming from the micelle and dye molecules.

To gain further insights into the mass of the complex, SEC coupled to multi-angle light scattering (SEC-MALS) was performed with a sample taken from the main peak and reinjected for the analysis. The SEC-MALS measurements were kindly performed by Raphael Gasper. MALS is a post-column detection method commonly used in conjunction with SEC. In this technique, eluted fractions from SEC are passed through a post-column MALS detector. The light scattering is recorded from multiple angles and is proportional to the square of the molar mass, concentration and refractive index. From this, the molecular mass and monodispersity of fractions eluting from a SEC column can be determined.

The SEC-MALS profile reveals a main peak at 15ml with a shoulder towards higher molecular weight species as described above (Figure 24B). Analysis of the molar ratio confirms the heterogeneity in the shoulder of the main peak which likely contains higher molecular weight species that cannot be completely separated from the main peak. Interestingly, a monodisperse species can be detected in the main peak with a calculated mass of approximately 353 kDa. This is in rough agreement with the observed mass of the complex in BN-PAGE. Again, this mass matches with a potential 1:1 stoichiometric complex of Scap and Insig covered by a detergent micelle with a mass of approximately 150 kDa. A study investigating the aggregation

number of LMNG molecules with the membrane protein BmrA quantified that 157 LMNG molecules associated with the protein.<sup>221</sup> This would match to an observed micellar size of approximately 150 kDa. Another study using mass photometry calculated that the *E. coli* *bo3* oxidase binds 146 LMNG molecules that would correspond to a micelle size of 146 kDa also matching the observation of a 1:1 complex of Scap-Insig.<sup>222</sup>

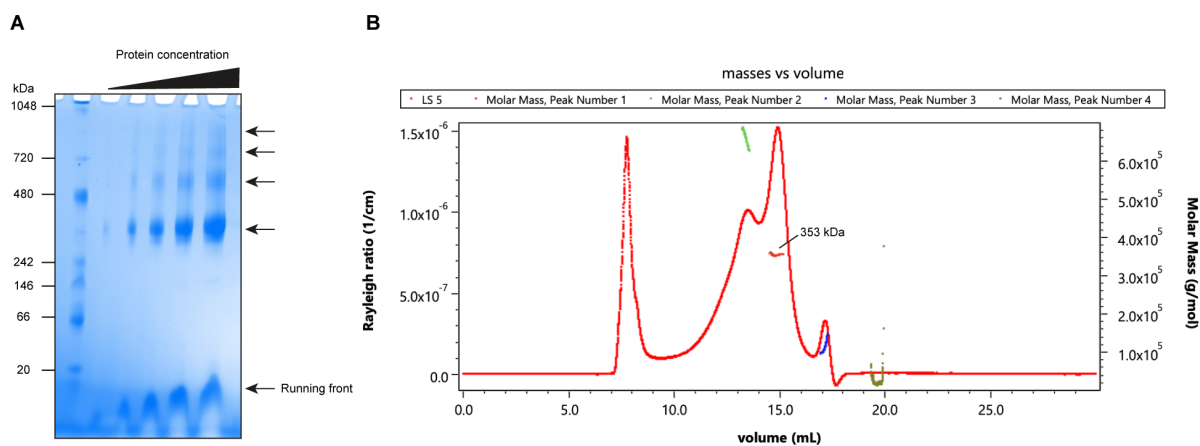


Figure 24: Molecular mass analysis of the Scap-Insig complex.

(A) A sample of the isolated Scap-Insig complex in LMNG-CHS was separated in a BN-PAGE along with NativeMark™ ladder. An increasing protein amount was loaded into different wells of the cell as indicated. The most prominent migrates at approximately 350 kDa most likely corresponding to a 1:1 complex of Insig and Scap. Higher oligomeric assemblies are also observed and indicated with the arrows. (B) SEC-MALS analysis of the Scap-Insig complex. The complex was separated by a Superose 6 column and subsequently analyzed by MALS. The light scattering chromatogram is shown with insets of molecular mass analysis. The measurement and analysis were kindly performed by Raphael Gasper.

These results confirm that the purification procedure could be optimized by solubilization and purification of the Scap-Insig complex in LMNG-CHS. The SEC profile shows that the main peak can be better separated from the shoulder appearing towards higher molecular weight species. This peak was not further analyzed since it shows particles of different sizes and shapes. Also, SEC-MALS analysis indicates heterogeneity among the particles eluting in the shoulder peak. Likely, higher oligomeric species are eluted in this peak with varying stoichiometric assemblies that cannot be separated from each other. The gaussian-shaped main peak eluting at 15 ml elution volume shows high purity and monodispersity. SDS-PAGE analysis confirmed that Scap and Insig are co-eluting suggesting they form a tight complex after tandem-affinity purification and SEC. Negative stain EM analysis reveals that indeed the particles corresponding to the Scap-Insig complex display a homogenous particle distribution. The particles have on average a diameter of 12 nm with a large disc-shaped part. This corresponds to the detergent micelle surrounding the TMDs of the complex. All classes show



at least one additional density outside of the detergent micelle which is blob-like and globular. This domain likely corresponds to WD40 of Scap. Sometimes a second domain on the opposite side of the micelle can be observed which could correspond to a domain formed by L1 and L7 of Scap. However, to answer this question and to confirm the estimated 1:1 stoichiometry of the complex, determination of a higher resolution structure is required.

### 9.1.7 Cryo-EM screening and data collection of the Scap-Insig complex in LMNG

Several plunge-freezing conditions and grid types were tested to find suitable conditions that would allow the structural determination of the Scap-Insig complex. Initially, I obtained a good particle distribution using QF 1.2/1.3 200-mesh copper grids when the complex was purified in GDN. I sought to find parameters that increase the contrast of the images which would be beneficial for structural determination. It was shown that grids made of conventional amorphous carbon (holey carbon grids) undergo a strong beam induced motion during irradiation.<sup>223</sup> In principle, this can be corrected computationally by motion correction algorithms. However, reliably correcting the beam-induced motion when small targets are imaged becomes tremendously important. Subtle inaccuracies in motion correction for larger particles is typically not an issue. However, for smaller particles this could potentially pose a problem by hindering accurate alignment of particles in both 2D and 3D. One effective solution is the use of gold-support grids, which directly mitigate the motion induced by the beam, enhancing the image contrast.<sup>224</sup>

Here, the Scap-Insig complex purified in LMNG-CHS was concentrated to 1 mg/ml. A portion of the complex was applied to UltrAuFoil 1.2/1.3 grids. A grid with an average ice thickness of 25 nm was transferred to a Titan Krios equipped with a K3 direct detector for data collection. An overview of the parameters of the data collection and image processing are listed in Table 2. A detailed image processing workflow is shown in Figure S 3.

The micrographs show a uniform and homogeneous particle distribution (Figure 25A). The contrast is slightly enhanced by using holey gold grids. 2D class averages show an increased signal-to-noise ratio compared to previous dataset. The typical particle shape of the Scap-Insig complex is observed with a spherical micelle and a large globular domain protruding out (Figure 25B). The class averages show for the first time high-resolution features such as stripes in the detergent region that likely correspond TMDs of the proteins. Moreover, clear secondary structural features in the 2D classes for the WD40 domain could be restored where individual blades of the propeller become visible (Figure 25C). The good particles from 2D classification were used to generate an initial model that can be used as reference for 3D classifications.

Roughly 65K particles were selected from 3D classification for a high-resolution 3D refinement. The overall achieved resolution was limited to 12Å and could not be reduced by adjusting the image processing workflow or by increasing the dataset size.

Table 2: Statistics of the cryo-EM data collection and image processing of the Scap-Insig complex purified in LMNG-CHS.

	<b>Scap-Insig in LMNG</b>
<b>Data collection</b>	
Microscope	Titan Krios (X-FEG, Cs-corrected)
Magnification	81,000
Voltage (kV)	300
Defocus range (µm)	-0.8 to -2.2
Camera	K3 super-resolution mode
Pixel size (Å/pixel)	0.45 Å/pix
Total electron dose (e <sup>-</sup> /Å <sup>2</sup> )	89
Exposure time (s)	4
Number of images	7204
<b>Image processing</b>	
Number of picked particles	1,750,773 (crYOLO)
Number of particles after ISAC	732,902
Number of final particles	67,554
Final resolution (Å)	12 Å
Symmetry	<i>C</i> <sub>1</sub>
Map sharpening B factor (Å <sup>2</sup> )	-130

The reconstruction shape closely reassembles the previous reconstruction (Figure 25D). The elliptical region corresponds to the detergent micelle that covers the TMDs. The domain outside of the micelle becomes in this reconstruction a little bit more featured and is disc-shaped. The dimensions of the additional domain outside of the micelle is slightly more defined than the previous reconstructions and align well with the fitted crystal structure of the WD40 of Scap (Figure 25D). However, no high-resolution features were discernable that would unambiguously confirm its identity.

In some 2D classes high-resolution features likely representing the TM helices were resolved. However, this signal could not be restored during 3D refinement. Lowering the threshold and cutting through the detergent micelle reveals a low-resolution density inside the micelle (Figure 25E). The density is not clearly resolved and individual helices cannot be seen at this resolution.

The limited resolution can be due to multiple reasons, among which flexibility of the extramembranous domain relative to the detergent micelle can hamper the proper alignment of the particle during image processing. Moreover, membrane proteins handled in detergent solution are kept outside of their natural environment. Thus, embedding of the complex in a more lipid-like environment could facilitate structural determination to a higher resolution.

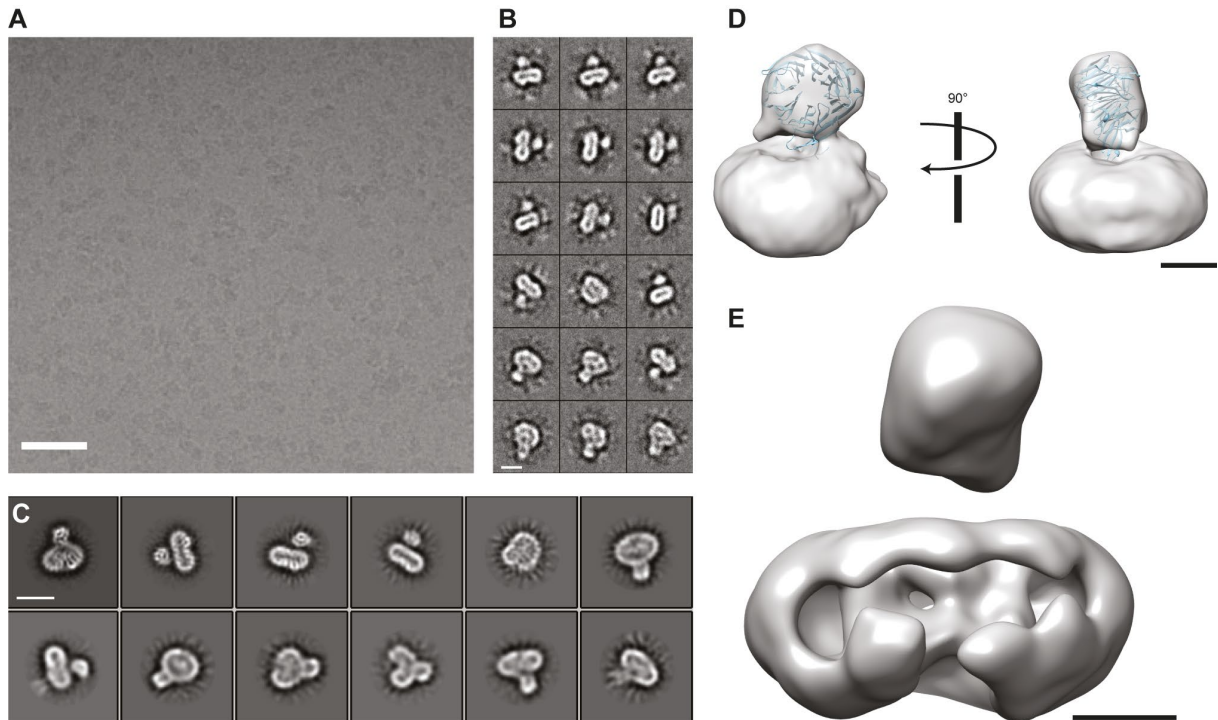


Figure 25: Cryo-EM data collection and processing of the Scap-Insig complex in LMNG-CHS.

(A) Representative cryo-EM micrograph of the Scap-Insig complex that was applied to UltrAuFoil 1.2/1.3 grids. Scale bar, 50 nm. (B) 2D class averages generated in SPHIRE show the characteristic shape of the complex with an elliptical density corresponding to the detergent micelle and a small extramembranous globular domain. Scale bar, 12 nm. (C) 2D class averages generated in Relion show for the first time high-resolution information for the Scap-Insig complex. Stripes in the detergent micelle can be seen corresponding to the TMDs of both proteins. Furthermore, individual blades of the WD40 propeller of Scap can be resolved. Scale bar, 12 nm. (D) Final reconstruction of the Scap-Insig complex with the fitted WD40 crystal structure of Scap (PDB ID: 4YHC). Scale bar, 5 nm. (E) Higher surface threshold of the map presented in (D) clearly shows the density inside the detergent micelle likely corresponding to the TMDs of the complex. Scale bar, 5 nm.

### 9.1.8 Nanodisc reconstitution of the Scap-Insig complex

Lipid nanodiscs are small discoidal lipid bilayers which are maintained in a soluble state in aqueous buffer solutions without the need of detergent. Two copies of the amphipathic membrane scaffold protein (MSP) wraps around a small lipidic patch and allows the handling without further addition of detergent.<sup>225</sup> The MSPs are derived from apolipoprotein A-I (Apo-AI), a naturally occurring protein that solubilizes lipid- and cholesterol-filled vesicles. Several

engineered variants of the original Apo-AI exist with varying lengths, forming nanodiscs of varying diameters ranging from 8-16 nm. This flexibility allows the reconstitution of different sized membrane proteins in a lipid environment.

As described above, the detergent micelle surrounding the TMDs of the Scap-Insig complex has an average diameter of around 12 nm. Therefore, I intentionally chose an MSP scaffold protein that would allow the formation of nanodiscs with a diameter close to 13 nm. To achieve an efficient reconstitution, different lipids and lipid mixtures were tested including the addition of cholesterol and 25-HC.

Eventually I found good reconstitution conditions when using the scaffold protein MSP1E3D1 with the lipid POPC without addition of additional cholesterol or 25-HC. For reconstitution, the Scap-Insig complex purified in DDM-CHS was mixed with MSP1E3D1 and POPC in a molar ratio of 1:5:500, respectively. This ensures excess of empty nanodiscs are present which can be removed from the protein of interest containing nanodiscs via SEC. A short incubation step on ice of this mixture followed addition of Biobeads to initiate detergent removal. The mixture containing both empty nanodiscs and nanodisc reconstituted Scap-Insig complex as well as potential aggregates was further purified by injection into a Superdex 200 column. Analysis of the fractions was conducted by SDS-PAGE and negative stain EM.

The SEC profile shows one broad peak with multiple shoulders (Figure 26A). The main peak containing Scap, Insig and MSP1E3D1 elutes over a range from 10-14 ml (Figure 26B). Additional peaks are observed at later elution volumes likely corresponding to free scaffold protein. Fractions of the main peak were analyzed by negative stain revealing different sized particles in each fraction (Figure 26C).

The first fraction (Figure 26, red) shows heterogenous particle distribution with smaller and larger sized particles ranging from 5-40 nm. The last fraction (Figure 26, blue) contains more homogenous particles with a smaller size. The average diameter of the particle in this fraction is approximately 8-9 nm which would be too small for nanodiscs accommodating the Scap-Insig complex. These small particles may potentially encapsulate individual proteins from complexes that have dissociated are represented empty nanodiscs.

Interestingly, the fraction collected from the middle of the peak (Figure 26, green) consists of homogenous sized particles with an average diameter of 13 nm. The particles are disc-shaped and contain a small globular density projecting outward. The disc-shaped part of the particles represents the typical shape of a nanodisc. This observation of particle shape is reminiscent of the observations made for the Scap-Insig complex purified in detergent.

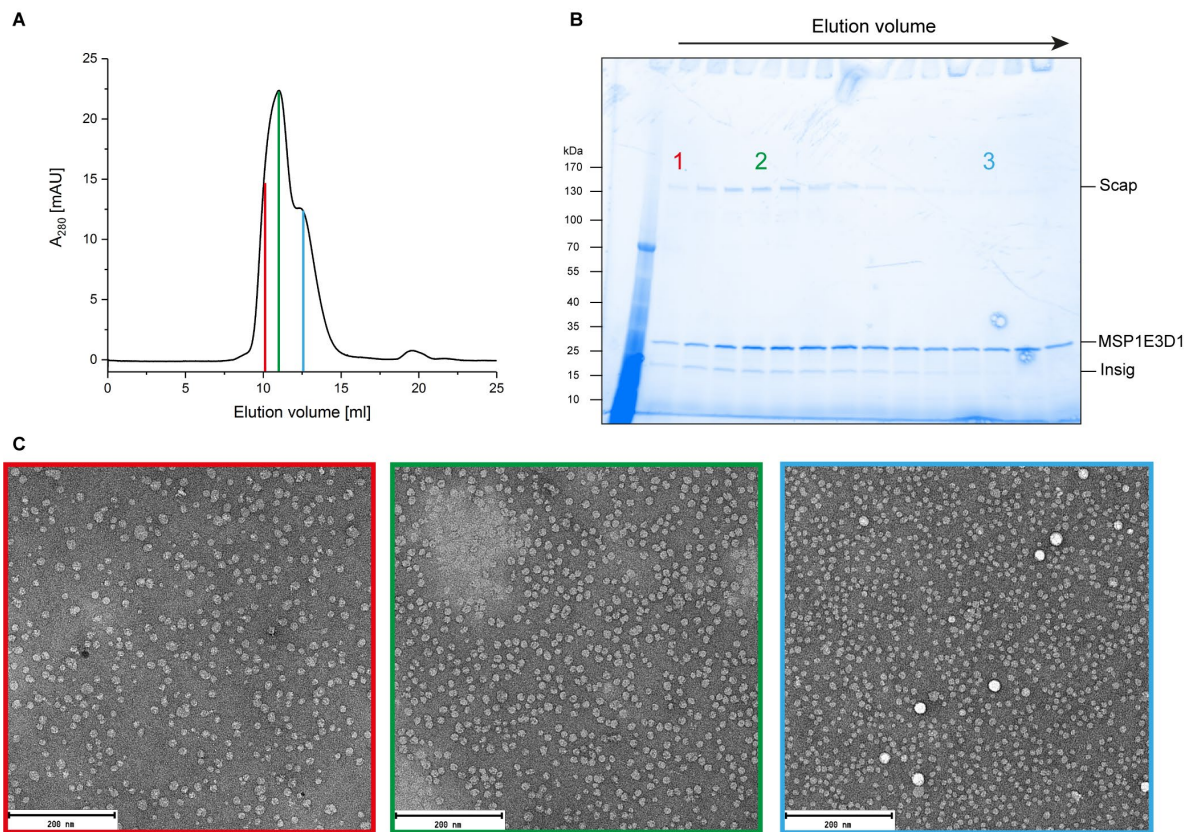


Figure 26: Nanodisc reconstitution of the Scap-Insig complex.

(A) SEC profile of the reconstituted Scap-Insig complex into MSP1E3D1 lipid nanodiscs. A Superdex 200 increase column was used. (B) The elution fractions indicated with the red bar were loaded onto a semi-denaturing SDS-PAGE and proteins bands were visualized by stain-free detection. The colored (A) bars and (B) colored numbers represent the fractions that were analyzed by (C) negative stain EM. Scale bar, 200 nm.

To achieve a higher resolution of the complex embedded in nanodiscs, I concentrated the sample containing the reconstituted Scap-Insig complexes and applied an aliquot to cryo-EM grids. These grids were rapidly frozen in liquid ethane and screened for optimal freezing conditions featuring a thin monolayer of particles. A grid with a good particle distribution and thin ice layer was selected for data collection on a Titan Krios 300 kV microscope equipped with a K3 direct electron detector. A dataset comprising approximately 5000 images was recorded in super-resolution mode at a pixel size of  $0.34 \text{ \AA}/\text{pixel}$ . Following data collection, particle picking and extraction was performed to calculate 2D class averages. From these class averages, a subset of 37,000 particles exhibiting clear features of the complex were selected for further refinement through a 3D reconstruction process.

The cryo-EM micrographs reveal a heterogeneous particle distribution with particles of different size and shape (Figure 27A). Additionally, aggregates of various sizes can be observed. Nonetheless, a substantial number of particles exhibit a clear disc-like shape which corresponds to side-views of the nanodiscs reconstituted complex were seen. The 2D classes

clearly provide conformation of the particle shape (Figure 27B). However, it is also evident that the size of the elliptical density, here the nanodiscs, is highly heterogeneous in diameter ranging from 12-20 nm. In some of the classes the extramembranous domain can be seen protruding from the nanodisc. Given the strong heterogeneity of the nanodisc size, calculating a high-resolution reconstruction proved challenging (Figure 27C).

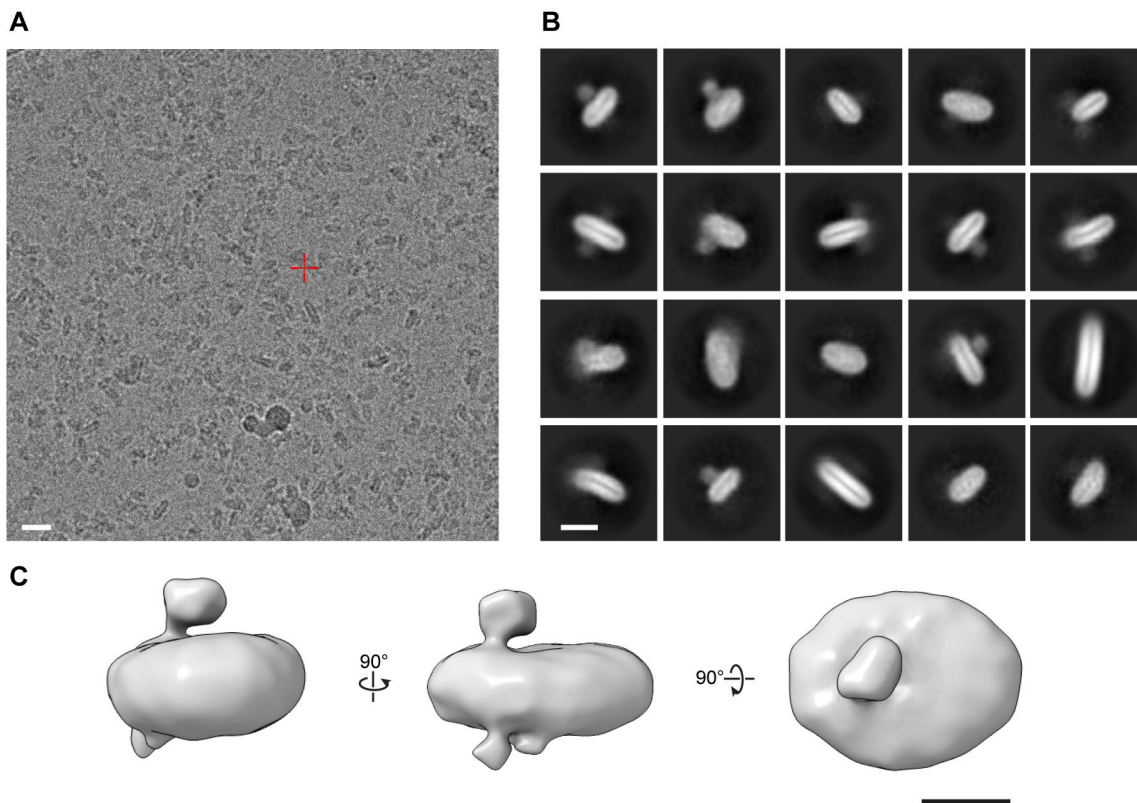


Figure 27: Cryo-EM screening and data collection of the nanodisc reconstituted Scap-Insig complex.

(A) Representative cryo-EM micrograph of the Scap-Insig complex in MSP1E3D1 nanodisc. Scale bar, 20 nm. (B) 2D class averages show the heterogeneity of the complex embedded in the nanodisc. The elliptical density corresponds to the nanodisc hosting the TMDs. A small extramembranous globular domain can be seen, which probably corresponds to the WD40 domain of Scap. Scale bar, 12 nm. (C) Final reconstruction of good particles selected for 3D refinement. Reconstruction was computed with 37K particles. Scale bar, 5 nm.

I tried to reduce the size heterogeneity of the nanodisc by using a self-circularized cMSP1E3D1 which should have a narrower size distribution, as recently reported.<sup>226</sup> However, all attempts to reconstitute the complex into these circularized MSPs failed. Alternative approaches for encapsulation of membrane proteins into a membrane mimicking environment are lipid nanoparticles formed by Saposin A.<sup>227</sup> Saposin A is an amphiphilic protein and plays a role in sphingolipid catabolism. In the absence of lipids, Saposin A adopts a closed and monomeric soluble form. The presence of lipids triggers a conformational change that leads to oligomerization of multiple copies of the protein engulfing highly ordered lipid bilayers,



reminiscent to nanodiscs. A notable advantage is that the Saposin A is a relatively small protein exhibiting modular assembly which drives automatic adaption of the nanoparticle size given by the size of the target protein.<sup>228</sup> I conducted numerous reconstitution attempts with exploration of different conditions including the use of different lipids, varying ratios of Saposin A and different strategies for detergent removal. To sum up, none of these strategies led to the successful reconstitution of the Scap-Insig complex in Saposin lipoparticles. Also, other detergent mimetics such as the amphiphilic polymers Amphipols, were tested for reconstitution and stabilization of the complex but were unsuccessful as well (Figure S 4). Consequently, I concluded that the most promising conditions for structural determination would involve the complex being purified in detergent combined with the use gold grids for plunge freezing and data collection.

### 9.1.9 Conclusion Scap-Insig complex

The proteins and mechanisms that regulate cholesterol homeostasis in vertebrates were discovered more than 20 years ago. These proteins include Scap, Insig and SREBP. Despite their discovery very little structural knowledge was gathered, especially of their TMDs. To address this knowledge gap, my research focused on the structural characterization of subcomplexes formed by these intricate membrane proteins.

Initially, I conducted small-scale experiments to overexpress and purify various subcomplexes consisting of SREBP, Scap, and Insig. Among the subcomplexes studied, I successfully demonstrated the interaction between WT Scap and WT Insig1 in the presence of cholesterol and 25-HC. To facilitate structural studies, I aimed to optimize the stability of the complex and develop a purification protocol suitable for such investigations.

I found that introducing a point mutation (D428A) in Scap's SSD strongly enhanced the stability of the complex. This mutation was reported before enabling Scap to interact with Insig even in the absence of sterols. I successfully scaled up the expression of the optimized Scap construct with Insig. As a result, I consistently obtained several hundred micrograms of purified Scap-Insig complex per liter of the expression culture.

First, I tried to explore the use of different detergents to solubilize and purify the complex. Eventually, I found that the complex was most stable when solubilized and purified in LMNG-CHS. Employing tandem-affinity purification and a final SEC led to a pure complex composed of Scap and Insig. To assess the sample quality negative stain EM was performed revealing a uniform particle distribution. On average, the particles exhibited a diameter of 12 nm featuring a large elliptical shaped part representing the detergent micelle, along with one, occasionally two, globular domains protruding from the micelle. I hypothesized that the first globular domain corresponded to the WD40 domain of Scap, while the second domain, situated on the opposite side of the micelle, potentially constituted a domain formed by the L1 and L7 loops of Scap, or vice versa. To further confirm the sample quality SEC-MALS was conducted. The results demonstrate that the primary species of the complex exhibited monodispersity with an observed molecular mass of 353 kDa. This suggests that Scap and Insig form a complex with a 1:1 stoichiometry.

To gain deeper insights into the arrangement of the Scap-Insig complex, I conducted screening experiments to identify optimal conditions for plunge freezing. By utilizing holey gold grids, which provided improved image contrast due to a reduced beam induced motion during imaging, I successfully identified conditions that enabled the embedding of the detergent-purified complex within a thin monolayer of vitreous ice. Subsequent image processing



techniques allowed me to observe high-resolution features of the complex for the first time. Secondary structure features such as stripes belonging to TM helices of the TMDs as well as individual blades of the WD40 propeller of Scap became visible in 2D class averages. However, despite these advances, I encountered challenges during 3D refinement which made it difficult to restore these features. The resolution of the reconstructions was always limited to resolutions of around 12Å. As reference, features such as  $\alpha$ -helices can be resolved at resolutions below 6Å. Nevertheless, within the detergent micelle, I managed to discern a weak but poorly resolved density, likely corresponding to the TMDs of the complex. I performed a rigid body fitting of the crystal structure of the WD40 domain into the extramembranous blob-like density which aligned with the observed dimensions. To further optimize the sample quality and stabilize the complex, different reconstitutions techniques using various membrane mimetics such as nanodiscs, Saposins and Amphipols were explored. These attempts did not yield the desired results.

Although the complex purified in LMNG-CHS displayed high-resolution features in 2D, translating these features into 3D space proved challenging. I suspected that difficulties during the alignment of the particle arise leading to difficulties in achieving a high-resolution structure. One reason could be the dominant high frequency but low-resolution signal of the detergent micelle. Most of the mass of the complex is embedded in the detergent micelle and only a small portion is extramembranous. During image alignment, the contribution of micelle could be so strong that alignment on internal features, such as the TMDs of the complex, is prevented. Additionally, the domains outside of the micelle might have exhibited excessive flexibility relative to the micelle, further complicating particle alignment.

To overcome these alignment issues, I decided to explore approaches to increase the size of the complex outside of the detergent micelle in order to facilitate particle alignment. Antigen binding fragments (Fab fragments) are widely used as specific binders to increase the particle size and overcome the alignment difficulty. However, much smaller in size and single chained is the VHH domain of camelid antibodies, termed nanobodies. They are between 12-15kDa in size, highly soluble and can bind cryptic and smaller antigen binding sites compared to Fab fragments.<sup>229</sup> A collaboration was started with Dirk Görlich (MPI Göttingen) to identify nanobodies that specifically bind to the Scap-Insig complex. The aim was to administer the complex to alpacas and generate an immune response that would allow for the identification of binders through *in vitro* screenings. Unfortunately, it appeared that the complex administered to the alpacas lacked sufficient stability or immunogenicity to induce a robust immune response

and facilitate the identification of binders to the complex. Furthermore, attempts to generate nanobodies against Scap alone were also unsuccessful.

During that time of my research, I came across several papers that reported the high-resolution structure of the Scap-Insig complex and Scap alone.<sup>111-113</sup> The first structure reported by Yan et al.<sup>111</sup> in 2021 showed that Scap and Insig formed a stoichiometric 1:1 complex through their TMDs. Within the complex interface, an elongated density was found attributed to 25-HC. The structure was determined by truncating the WD40 of Scap and coexpressing it with Insig2. Three mutations were introduced into Insig (C14S, C90S and C215S) likely mitigating issues with unspecific aggregation. 25-HC was supplemented during the expression extraction to maintain a stable complex. The complex was solubilized in LMNG-CHS and purified via tandem-affinity purification and SEC performed in the presence of LMNG-CHS. Interestingly, Yan et. al reported a small globular domain outside of the detergent micelle despite truncating the Scap construct by the WD40 domain. However, the resolution of this domain in the study was limited.

Another high-resolution structure of the Scap-Insig complex appeared several months later by Yan et al.<sup>112</sup> addressing the previous limitation by resolving the globular domain to higher resolution. This time, the complex was purified in digitonin. The structure reveals that the previously globular domain corresponds to a domain formed by the large loops L1 and L7 of Scap. The two loops form a tightly intertwine globular domain which has high structural similarity to the ECD of NPC1.<sup>114</sup>

One month later, more structures of the Insig-Scap complex and Scap alone were solved by Kober et al.<sup>113</sup>. The structures could resolve the domain formed by L1-L7 to high resolution. The authors generated Fab fragments that specifically to this domain to guide particle alignment. However, the resolution of the TMDs was comparatively lower than in the previous structures, allowing only for the fitting of homology models. The structure of full-length Scap alone was determined as well as the Scap-Insig complex. To form a stable complex, a mutation was introduced into Scap (D435V) which exhibited a similar effect to the D428A mutation. Interestingly, a significant conformational change of the L1-L7 domain was observed upon Insig binding causing this domain to rotate 215° perpendicular to the membrane. This conformational change results in a rearrangement of the TMD of Scap eventually concealing the MELADL motif and blocking Scap-mediated transport of SREBP to the Golgi apparatus ultimately blunting cholesterol uptake and *de novo* synthesis.

Furthermore, the concerted conformational change of the SSD of Scap also led to a repositioning of TM8 upon Insig binding. This helix is directly connected to the WD40 domain of Scap. While the reconstruction of Scap alone revealed a resolved WD40 domain, in the context of the complex, TM8 moved away from the SSD and the interface of the complex, introducing flexibility to this domain.

These determined structures shed light on the complex structural dynamics of Scap in response to Insig binding. Moreover, they provide insights into the challenges I encountered in resolving the TMDs of this complex. The WD40 domain exhibits increased flexibility upon Insig binding since its position relative to the SSD of Scap is altered. This flexibility hinders proper particle alignment during image processing. The first high-resolution structures were solved by circumventing this issue by truncating the WD40 domain which reduced the flexibility of the complex. In contrast in the structure of Scap alone the WD40 domain is closer positioned to the SSD and likely less flexible allowing to be resolved. Furthermore, these structures provide a possible explanation for the requirements of 25-HC for the formation of the Scap-Insig complex over cholesterol. An elliptical density was found sandwiched between the interaction interface of Scap and Insig, oriented parallel to the TM helices. However, based on the local environment of this binding pocket, cholesterol would be accepted due to its increased hydrophobicity. In contrast, 25-HC carries a hydroxy group at position 25 that can form hydrophilic interactions with Insig and Scap.

In summary, I was able to independently identify the interaction between Scap and Insig, and I was able to establish a complete purification protocol that allowed visualization of secondary structural features in 2D classes for the first time. I found that a stabilizing mutation greatly increased the yield. However, the flexibility introduced by the WD40 caused problems in particle alignment of the complex, which hindered high-resolution structural determination. Overall, these structural insights will pave the way for further investigations of the dynamic interplay between these proteins, and hopefully in the future, the structures of SREBP and Scap bound to SREBP.

## 9.2 Structural characterization of the HMG-CoA reductase

The HMGCR catalyzes the rate-limiting step in the *de novo* synthesis of cholesterol and other non-sterol intermediates. A class of molecules called statins is widely prescribed to treat imbalances in cholesterol production. The central role of this protein in cholesterol production, but also as a target for pharmaceutical intervention, highlights its high medical relevance and the need for a detailed molecular structure. Structures of the catalytic part of this protein have been determined, revealing the binding mode for its natural substrate, but also for statins. However, little structural information has been obtained about the transmembrane domain of this protein. The TMD of the protein is central to a feedback mechanism that degrades the reductase upon accumulation of sterols in the membrane, which is necessary to shut down *de novo* synthesis. Given the high medical importance but lack of structures of the full length protein or the TMD, I decided to close this gap.

The full-length human HMGCR was previously studied in our group by Arne Bothe. Recombinant expression was achieved in Hi5 insect cells with subsequent purification attempts. Although the protein was enzymatically active, it eluted only in the void volume during a final SEC step, indicating aggregated and not properly folded protein. Negative staining of the void peak confirmed this observation by the presence of only aggregated particles. To increase solubility, the HMGCR was fused to an N-terminal mCherry. Negative staining of this construct revealed homogeneous but monomeric particles. However, at least a dimeric form of the protein is required to form the active site. During my research, I switched to a more native-like expression system and a different construct for overexpression and purification. The results will be divided into two parts, where the first part will describe the structural characterization of the HMGCR alone. The second part will discuss progress in studying the HMGCR in complex with native interaction partners.

### 9.2.1 Small-scale expression of the human HMGCR

As a first step, I wanted to screen the expression of the human HMGCR in mammalian cells on a small scale. For this purpose, I designed a construct for overexpression of the HMGCR in HEK293 GnT1<sup>-</sup> cells. The construct contains the full-length human HMGCR fused to a cleavable C-terminal eGFP-StrepII tag (Figure 28A). A BacMam virus was generated and used for small scale transduction of mammalian cells. The highest expression rates, indicated by green fluorescence, were obtained 48 hours after infection (Figure S 5). To gain more information about the potential oligomeric state of the protein and the ability to solubilize the protein from overexpressed membranes, a detergent screen was performed. Whole cells

overexpressing the HMGCR were harvested and aliquoted for solubilization with different detergents. After solubilization, insoluble material was removed and a sample of the supernatant containing the solubilized membrane proteins was analyzed by FSEC and SDS-PAGE.

In total, 12 different detergents were used for solubilization. The SDS-PAGE reveals efficient solubilization for all detergents (Figure 28B). A protein band migrating at 130 kDa can be detected corresponding to the eGFP-tagged HMGCR. No degradation products of the HMGCR were detected. The elution profiles of the FSEC runs show for all samples a void volume eluting at 8 ml of varying height (Figure 28C). A second bell-shaped peak around 13 ml elution volume can be detected which corresponds to the specific elution of the fluorescently tagged HMGCR. Depending on the detergent, this peak is more gaussian shaped, indicating a more homogenous solubilized protein species. Increased width of the peak may indicate formation of inseparable oligomers or aggregates.

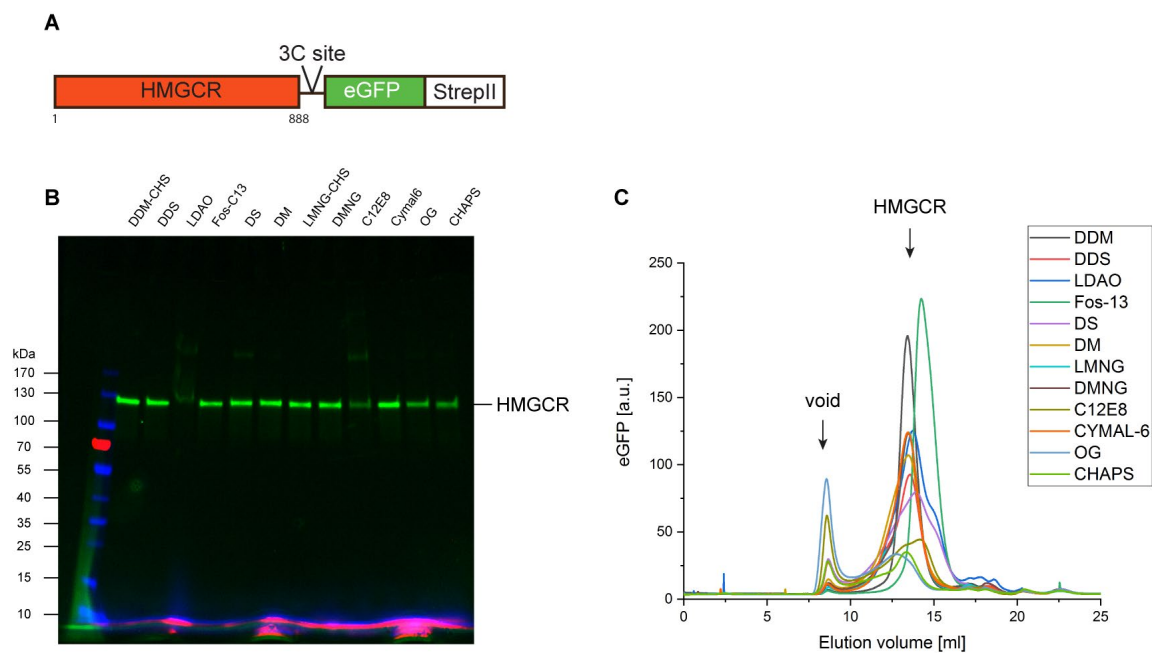


Figure 28: Detergent-screen of the human HMGCR

(A) Construct of the full-length human HMGCR fused to a cleavable eGFP-StrepII tag. (B) SDS-PAGE analysis of detergent-solubilized cell lysates from HEK293  $GnTI^{-}$  overexpressing the HMGCR. Detergents are indicated at the top of the gel. No band is visible for the LDAO sample because proteins did not enter the gel due to detergent charge interference. The band migrating at 130 kDa belongs to HMGCR. (C) FSEC profiles after detergent solubilization. Samples were injected onto a Superose 6 column.

Based on this analysis, Fos-Choline-13 (Fos-C13) produces a gaussian-shaped peak with good solubilization efficiency. However, Fos-C13 is a rather harsh detergent which serves a positive control for extraction but is not suitable for downstream applications. LMNG-CHS, which was

used for the purification of the Scap-Insig complex as described above, displayed a good solubilization efficiency as the other detergent tested here. However, most of the HMGCR can be detected in the void volume indicating increased destabilization of the protein upon extraction. Here, the best mild detergent for solubilization that produces a gaussian-shaped peak is DDM-CHS. As described above, DDM-CHS is the most successful detergent mixture for determining membrane protein structures and such an ideal candidate for usage in large scale solubilization and purification of the HMGCR. Other detergents showed peak of increased sized and are thus less suitable for purification of the reductase.

### 9.2.2 Large scale expression and purification of the HMGCR

The expression culture volume was scaled up to 800 ml and transduced with the BacMam virus encoding the full-length HMGCR. Cells were harvested 48 h post-infection and whole cells were solubilized by addition of DDM-CHS. Insoluble material was removed via centrifugation and the soluble fraction containing the HMGCR was applied to Strep Tactin beads. Several washes were performed to remove non-specific binders and contaminants. The reductase was eluted by addition of D-desthiobiotin to the elution buffer. Elution fractions were analyzed by SDS-PAGE, pooled and concentrated and further purified via SEC in the presence of DDM-CHS.

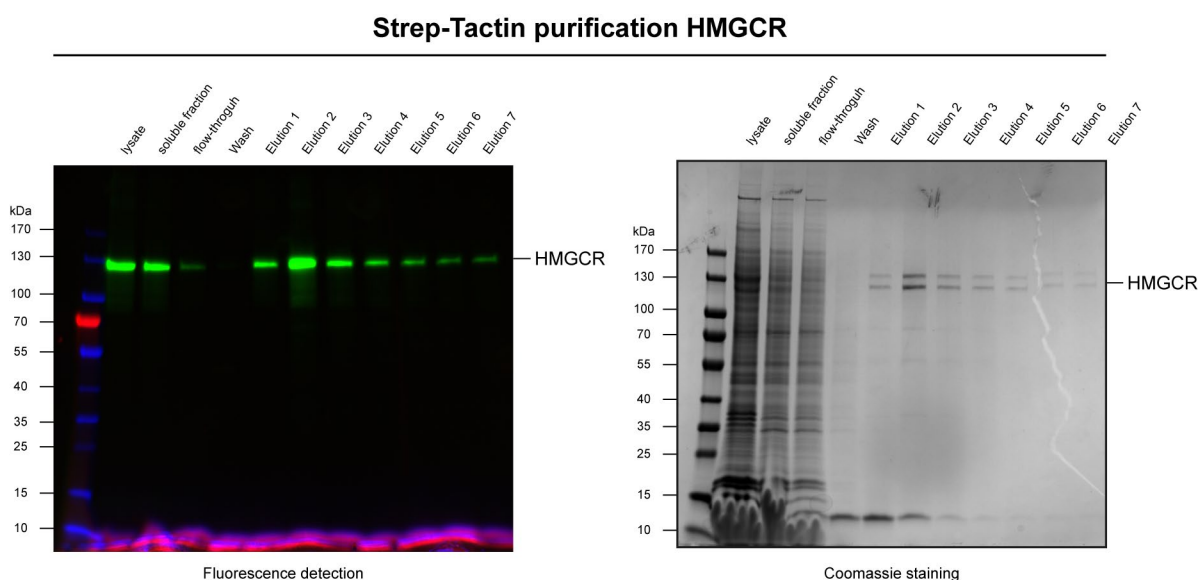


Figure 29: Large-scale purification of the human HMGCR.

The SDS-PAGE analysis (Figure 29) of the purification confirms the high solubilization efficiency using DDM-CHS indicated by little decrease of fluorescence intensity comparing the uncleared lysate and soluble fraction. When applied to Strep Tactin beads, little reductase can be found in the flow through. Elution using D-desthiobiotin shows a nice elution pattern and

presence of the reductase in all elution fractions. Coomassie staining reveals high purity of the band migrating at 130 kDa. A second band is observed slightly above the main band which does not display any fluorescence. This band also corresponds to the HMGCRC but with an unfolded eGFP. The elution fractions were pooled and further purified using SEC. For ease of detection, the eGFP tag was initially not cleaved off.

The SEC profile shows a single gaussian shaped peak at 13 ml with a small shoulder towards higher molecular weights (Figure 30A). Buffer components can be detected at 20 ml. SDS-PAGE analysis of the elution fractions from SEC reveal a high purity of the HMGCRC (Figure 30). All minor contaminations visible in the concentrated sample prior to SEC could be separated from the reductase.

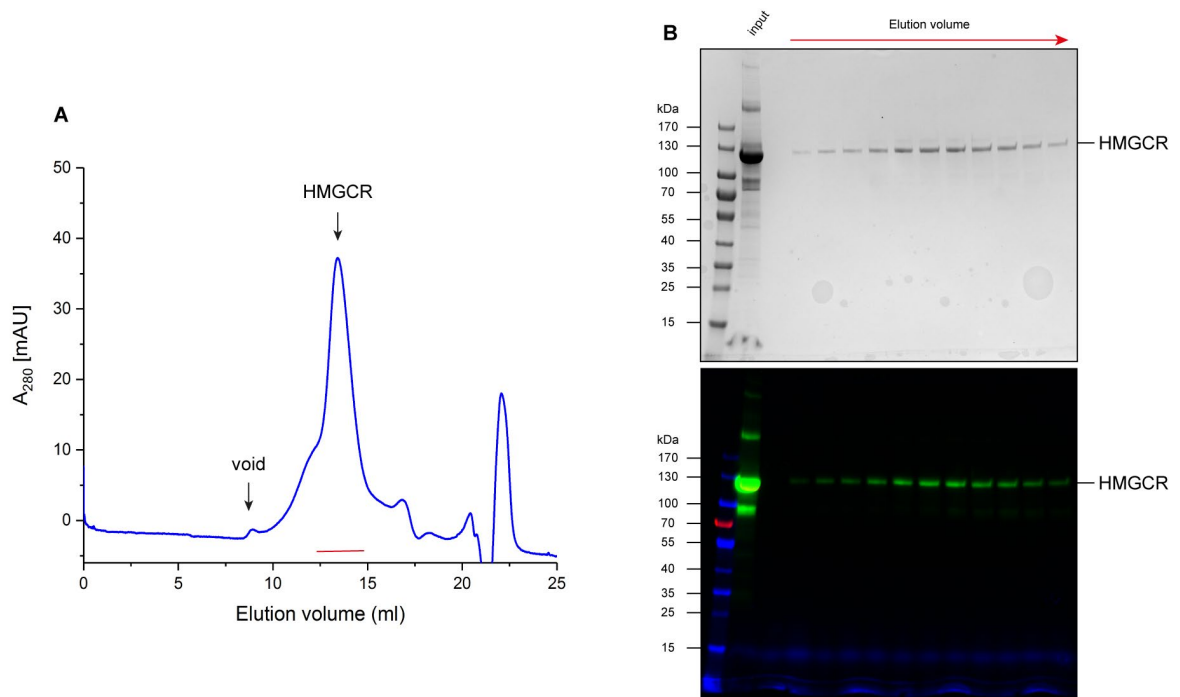


Figure 30: Final purification step of the HMGCRC via SEC.

(A) SEC profile of the purified eGFP-tagged HMGCRC. The protein was purified in DDM-CHS and injected into a Superose 6 column for the final purification step. (B) SDS-PAGE analysis of the indicated elution fractions of the SEC run. Protein bands were detected by Coomassie blue staining or fluorescence detection. The gels show a high purity of the HMGCRC after SEC.

### 9.2.3 Biochemical characterization of the HMGCRC

To verify that the reductase was purified in a relevant physiological conformation, the enzymatic activity was monitored. As fast assay monitoring the consumption of NADPH was already established by a previous member of the group Arne Bothe. However, the activity was always checked in microsomes but not with purified protein. Therefore, I used this assay to characterize the kinetic parameters of the purified HMGCRC.

The assay is a spectroscopic assay that monitors the consumption of NADPH. Two equivalents of the cofactor NADPH are used to convert one equivalent HMG-CoA into mevalonate (Figure 31A). NADPH and the reaction product  $\text{NADP}^+$  both share an absorption maximum at 260 nm which stems from absorption by the adenine moiety. However, NADPH has an additional absorption maximum at 340 nm which is absent for  $\text{NADP}^+$ . Thus, the enzymatic reaction can be followed by measuring the decrease of absorption at 340 nm.

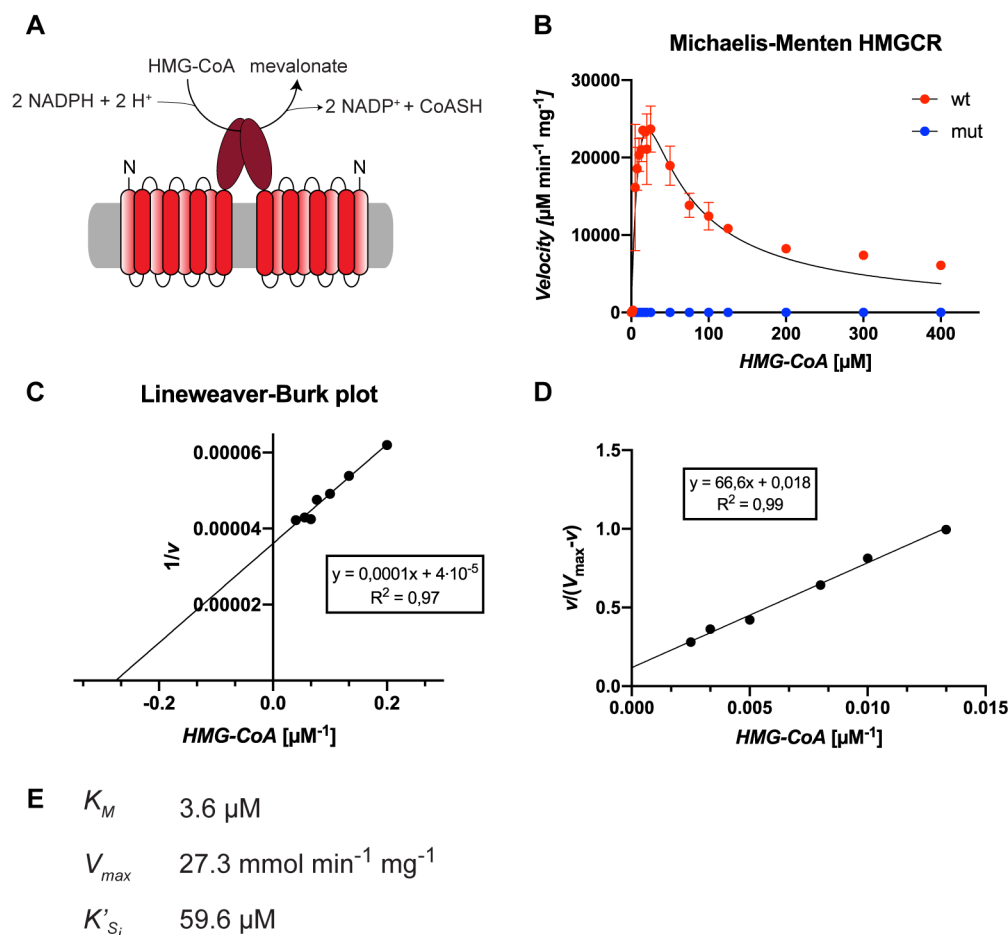


Figure 31: Michaelis-Menten kinetics of the full-length HMGCR.

(A) Schematic representation of the enzymatic reaction of the HMG-CoA reductase. Two equivalents of NADPH are used to convert one equivalent HMG-CoA to mevalonate. (B) Michaelis-Menten plot of the full-length HMGCR (red data points). Four amino acids were mutated generating either a steric clash or removal of stabilizing salt bridges at the tetramer interface (R641A, A695E, E700R, E709K; blue data points). Measurements were performed in technical replicates. (C) Lineweaver-Burk plot of initial reaction velocities of the wt. (D) Plot for determination of the apparent substrate inhibition constant  $K'_{Si}$ . (E) Kinetic values of the WT HMGCR.

Here, the protein was purified in DDM/CHS micelles via affinity purification and SEC. The assay was set up by using a fixed concentration of reductase at 120 nm with varying HMG-CoA concentration between 0-400  $\mu\text{M}$ . An excess amount of NADPH at 800  $\mu\text{M}$  was provided. The consumption of NADPH was measured at a wavelength of 340 nm over 5 minutes from



which the initial reaction rates were calculated. According to Michaelis-Menten, the initial reaction velocities ( $v_0$ ) were plotted against the substrate concentration (Figure 31B).

Here, at low substrate concentrations the reaction rate follows exponential growth with increased substrate concentrations. The maximum reaction velocity ( $V_{\max}$ ) is reached at  $\sim 25 \mu\text{M}$  of substrate. Further increase in substrate concentration leads to a reduced reaction velocity possibly reaching a lower plateau at high substrate concentrations. Interestingly, the kinetic measured here follows those typically observed for substrate-inhibited enzymes. Substrate-inhibition is common for many metabolic proteins i.e., phosphofructokinase or malate dehydrogenase which is meant to provide a constant flow of products.<sup>230</sup> Substrate-inhibition has not been observed in activity assays using only the C-terminal portion of HMGCR. The kinetics of the TMD-deleted catalytic portion follows a normal logarithmic growth curve until a plateau at high concentrations is reached.<sup>231</sup>

To further derive kinetic parameters, the data was plotted according to Lineweaver-Burk. At low substrate concentrations, plotting the reciprocal substrate concentration against the reciprocal reaction velocity revealed a linear relationship (Figure 31C).  $K_M$  and  $V_{\max}$  values could be determined to  $3.6 \mu\text{M}$  and  $27.7 \text{ mmol min}^{-1} \text{ mg}^{-1}$ , respectively, which are in agreement with the values determined for the catalytic domain.<sup>232</sup> At higher substrate concentrations the plot of the reciprocal substrate concentration against  $v/(V_{\max}-v)$  results in a linear relationship (Figure 31D) as described elsewhere.<sup>233</sup> The apparent substrate inhibitory constant  $K'_{Si}$  could be determined to  $59.6 \mu\text{M}$  which corresponds to the substrate concentration at which the reaction velocity is reduced to the half maximum analogously to the definition of the  $K_M$  value. Till now, the oligomeric state of the full-length protein is not known. The cytosolic domain was found as a tetramer in the asymmetric unit of the crystal structure. Analytical ultracentrifugation experiments also supported the existence of an in-solution tetramer.<sup>234</sup> Therefore, I introduced four mutations which generate either steric clashes or remove stabilizing salt bridges at the tetramer interfaces but do not interfere in formation of the dimer. Interestingly, the mutant does not show any enzymatic activity (Figure 31B, blue data points). This might point to the presence of an *in vivo* tetramer as catalytic active species. However, further experiments need be performed in the future to prove the existence of a tetramer as relevant form. In this context, a high-resolution structure of the full-length HMGCR is of utmost importance.

#### 9.2.4 Structural characterization of the full-length HMGCR

The enzyme purified in DDM-CHS micelles exhibit catalytic activity that matches the previously determined values for the catalytic portion. Here, the full-length protein was purified

in an enzymatically active form. This purified construct was subjected to negative stain EM in order to characterize the sample quality. The purified HMGCR in DDM micelles tagged with eGFP was applied to negative stain grids and particles were subsequently visualized. Negative stain micrographs display a heterogeneous particle distribution (Figure 32A). Smaller particles with approximately 5 nm diameter likely correspond to empty detergent micelles as described above. Then, larger particles of different diameter and shapes can be observed. The diameter varies between 10-20 nm. To confirm this, all particles were manually picked and classified in 2D using ISAC (Figure 32B). Indeed, the class averages reveal different shaped particles. Smaller particles appear roundish and featureless. Those likely correspond to empty detergent micelles. Larger particles also feature a disc-shaped part with a small globular domain protruding. Based on the results discussed above, these particles likely correspond to the HMGCR where the disc-shaped part represents the detergent micelle and the small globular domain projecting out corresponds to the C-terminal extramembranous domain. However, precise prediction of the oligomeric assembly of the HMGCR cannot be made at this resolution.

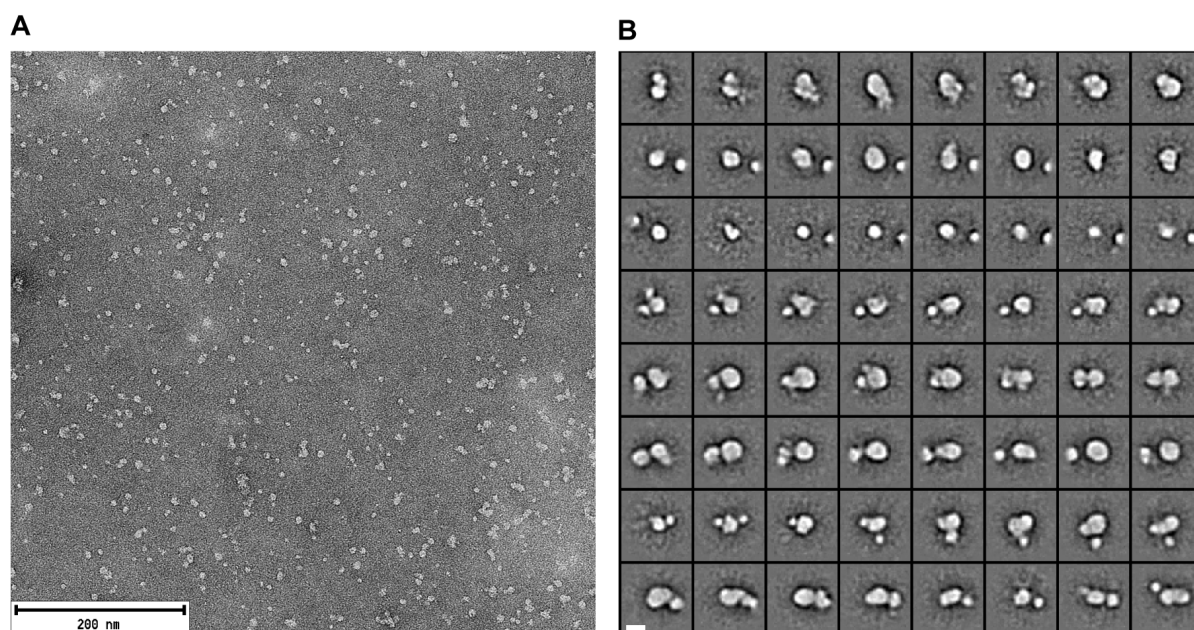


Figure 32: Negative stain analysis of the full-length HMGCR in DDM micelles.

(A) Representative negative stain EM micrograph of the human eGFP-tagged HMGCR purified in DDM-CHS. Scale bar, 200 nm. (B) The 2D class averages confirm the heterogeneous distribution of particle shapes and size. Scale bar, 10 nm.

The eGFP-tagged HMGCR was further biochemically characterized by BN-PAGE to assess the native sizes of these oligomers. Three bands were observed migrating at approximately 400, 480 and 720 kDa (Figure 33). The band migrating at 480 kDa is the most prominent. The size would roughly fit a dimeric form where each protomer consists of a molecular weight of 130

kDa engulfed by a micelle approximately 220 kDa in size. The band migrating at 720 kDa could correspond to the tetrameric form. An attempt was made to determine the native mass of the complex by SEC-MALS. However, a more homogeneous preparation is required. Mixtures containing oligomeric species of different sizes eluting in the same peak during SEC are not suitable for SEC-MALS analysis.

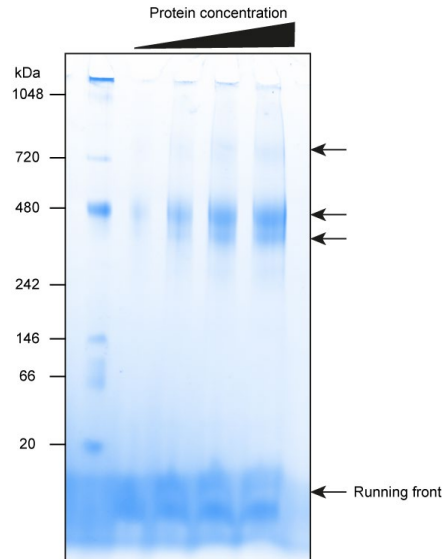


Figure 33: BN-PAGE of the purified HMGCR in DDM micelles.

Varying amounts of SEC-purified eGFP-tagged full-length HMGCR were applied to BN-PAGE. Protein bands were fixed and visualized using Coomassie blue stain. The arrows indicate prominent bands besides the running front which contains excess Coomassie blue dye.

Based on the size distribution of the particles in negative stain EM micrographs, the preparation likely contains a mixture of different oligomeric forms that are either in dynamic equilibrium, rapidly associating and dissociating, or different oligomers were directly extracted and cannot be separated during purification.

To confirm the presence of different oligomeric forms of the protein, the sample of HMGCR purified in DDM-CHS micelles was subjected to cryo-EM analysis. Freezing conditions were optimized by screening different conditions in a Talos Arctica 200kV electron microscope. Then, a good grid was used to set up the data set collection. A typical cryo-EM micrograph confirms the heterogeneity of the sample (Figure 34A). Particles of different sizes and shapes were observed. In addition, the particles tend to cluster and aggregate on the grid, indicating that sample optimization is required. This is confirmed by the low quality of the 2D class averages, which do not provide high resolution information. The particles appear disc-shaped in the 2D classes with varying sizes ranging from 6-14 nm in diameter (Figure 34B). Some of the particles also have a small globular density protruding. However, the quality of the data prevented the calculation of a three dimensional model.

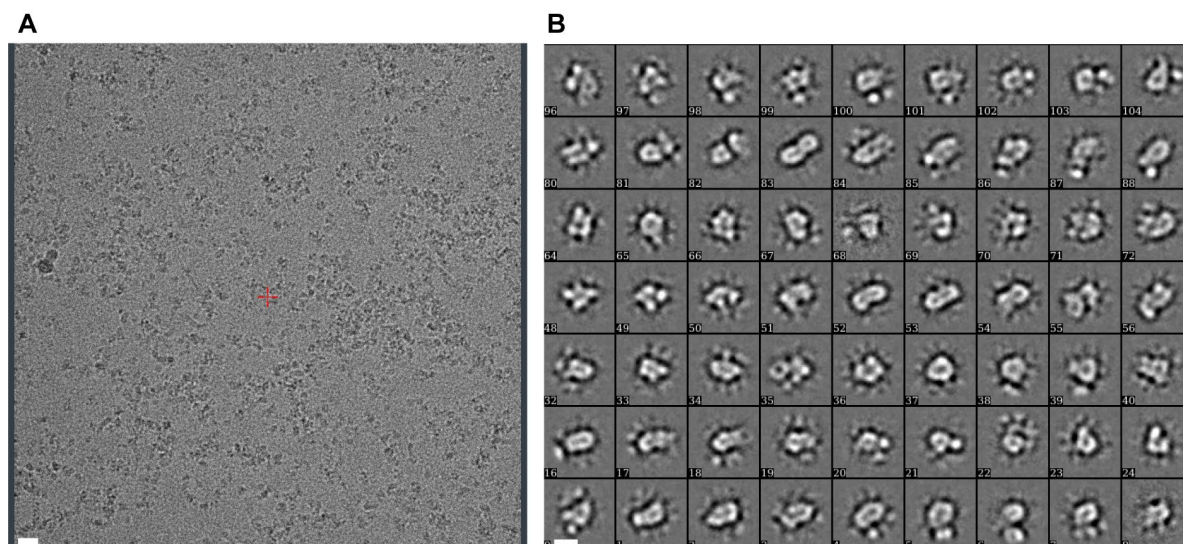


Figure 34: Cryo-EM analysis of the full-length HMGCR.

(A) Representative cryo-EM micrograph of the human full-length HMGCR purified in DDM-CHS. The image was taken using a Talos Arctica 200 kV microscope. Scale bar, 20 nm. (B) Representative 2D class averages generated in SPHIRE using ISAC. Scale bar, 10 nm.

Altogether, a lot of optimization was required to increase the homogeneity of the preparation. I explored different construct boundaries, stabilizing mutations, buffer conditions, addition of cofactors or inhibitors at different steps of purification, use of different detergents, and HMGCR orthologues from different species to obtain a more homogeneous sample before proceeding to cryo-EM. In short, none of these efforts resulted in a more homogeneous sample quality suitable for structural determination. Thus, other approaches to potentially stabilize the HMGCR were needed.

### 9.2.5 Identification of a stable HMGCR-UBIAD1 complex

Many conditions and construct boundaries were explored to obtain a sample suitable for the structural determination of the full-length HMGRC. However, none of them resulted in a homogeneous particle distribution as observed in negative stain EM and cryo-EM micrographs, which is a prerequisite for solving its structure. Based on the progress described above in investigating the Scap-Insig complex, I thought I might be able to stabilize the HMGCR by adding another binding partner. Central to the regulation of the HMGCR is the sterol-dependent recruitment of a degradation machinery to the TMD of the reductase mediated by Insig. I tried to co-express the HMGCR together with Insig. When both wild-type proteins Insig and HMGCR were coexpressed, no detectable levels of reductase were observed at any time point. It is likely that the HMGCR is directly degraded due to an excess of Insig in the membranes. To circumvent this, two point mutations were introduced that render the reductase resistant to

ubiquitination. Mutation of the two ubiquitination sites (K89, K248) to alanine restored expression of the HMGCR while co-expressing Insig. However, after initial purification of a potential complex, no interaction of these two proteins could be detected. One reason might be that the ubiquitination sites are close to the interaction interface. However, both sites are located on cytosolic loops whereas the interaction interface is presumably exclusive residing in the membrane. Another possibility for the lack of detectable interaction could be that specific sterols or other factors are required to induce the interaction. However, since WT reductase presumably is immediately degraded due to overexpression of Insig, it suggests that all necessary factors are present in the expression host. Also, the addition of the sterols lanosterol and 25-HC during expression and purification did induce their interaction. Another possibility is that the interaction between Insig and the reductase is very transient and does not survive extraction from the membrane and purification. During my studies, I was not able to further optimize the conditions to detect the interaction of both proteins and to purify a stable complex. Further attempts to reconstitute this interaction will be part of future studies and will require more rigorous exploration of the conditions.

Given that Insig is a potent accelerator of HMGCR degradation, proteins that exert the opposing effect are described. A known stabilizer of the reductase is UbiA prenyltransferase domain-containing protein 1 (UBIAD1), which competes with Insig for binding to the TMD of the HMGCR. Disease-associated mutations of UBIAD1 increase the binding affinity to the HMGCR, leading to increased cholesterol production observed in patients suffering from Schnyder corneal dystrophy (SCD). Therefore, UBIAD1 might be an ideal candidate to stabilize the TMD of the reductase for structural determination.

To test whether the HMGCR and UBIAD1 form a complex, the ubiquitination-resistant form of the reductase with two mutations (K89A, K248A) was coexpressed with UBIAD1. The HMGCR was fused to a C-terminal eGFP-StrepII tag and UBIAD1 was fused to an N-terminal StrepII-mCherry tag, allowing fluorescent detection during SEC. A second UBIAD1 construct was generated with a disease-relevant mutation (G186R), which has been shown to increase association with the HMGCR.<sup>74</sup> After coexpression of the constructs, cells were harvested and membranes were solubilized using DDM-CHS. Insoluble material was removed by centrifugation, and soluble supernatants containing the solubilized membrane proteins were injected into an SEC column and complete fluorescence traces were recorded. As a control, UBIAD1 constructs were expressed alone, solubilized and also injected into an SEC column with fluorescence detection. Fluorescence traces were compared to detect potential higher oligomeric species upon coexpression of UBIAD1 with the HMGCR.

Fluorescence traces of both UBIAD1 constructs show the presence of two major peaks corresponding to the void volume eluted at 8 ml and a single gaussian-shaped peak at approximately 15 ml. The peak at 15 ml corresponds to the elution of UBIAD1 in its uncomplexed form. Interestingly, specific fluorescence from UBIAD1 can be detected at earlier elution volumes when co-expressed with the HMGCR. Additional peaks corresponding to UBIAD1 were detected at approximately 11 and 13 mL, which were not seen when UBIAD1 was expressed alone. In addition, the fluorescence of UBIAD1 aligns well with the detected green fluorescence corresponding to the eGFP-tagged HMGCR. This suggests that UBIAD1 and the HMGCR form a complex upon co-expression.

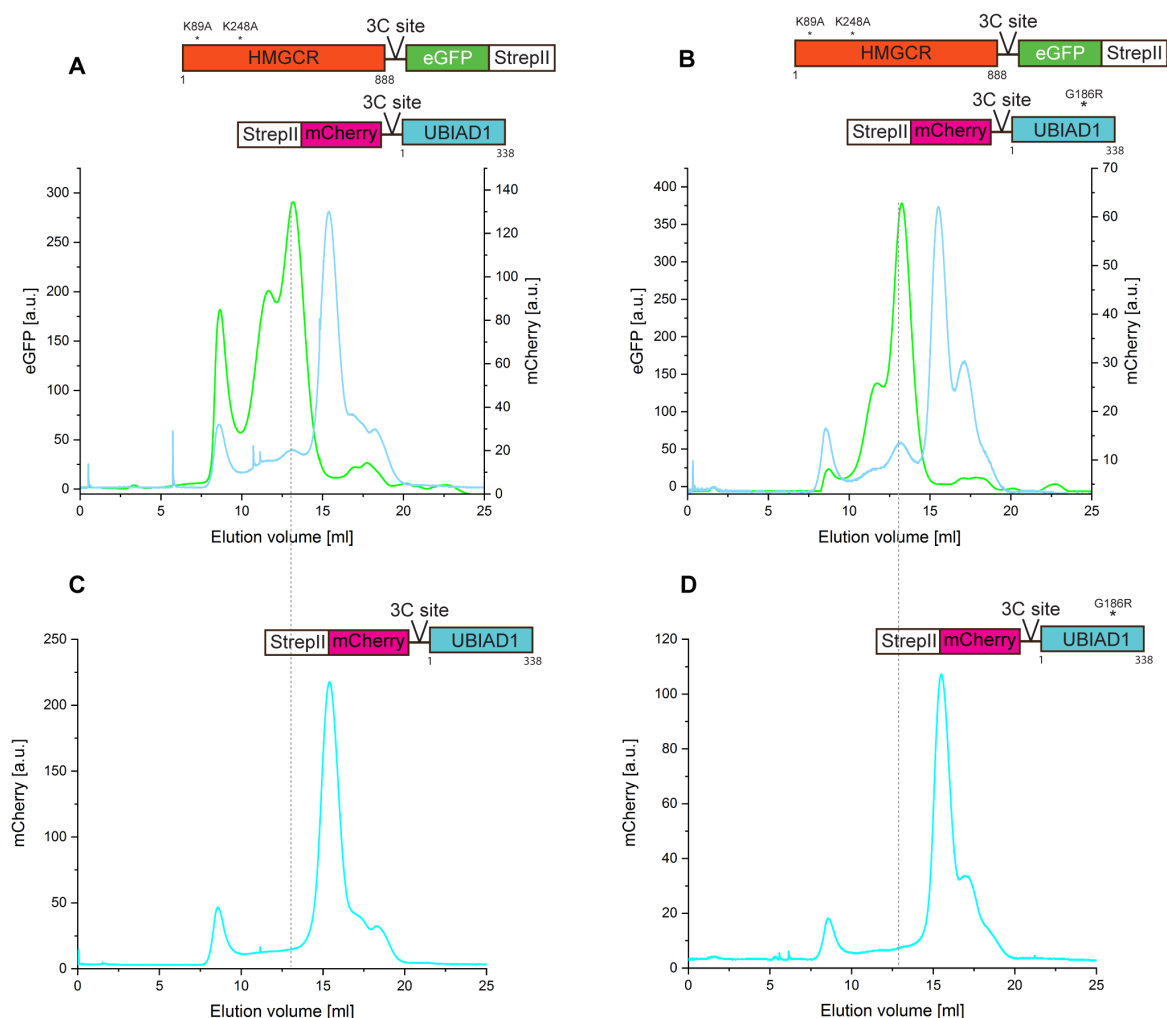


Figure 35: Identification of HMGCR-UBIAD1 interaction by FSEC.

(A) - (D) FSEC analysis of the interaction between HMGCR and UBIAD1. The indicated constructs were coexpressed (A & B) with the HMGCR containing two mutations of the ubiquitination sites. (C & D) UBIAD1 constructs were expressed alone and prepared for FSEC analysis. The gray dotted line indicates the position where the respective HMGCR-UBIAD1 complexes elute. These peaks are absent when UBIAD1 is expressed alone.

To confirm this finding, both proteins were expressed on a larger scale and purified by a single affinity purification step. After overexpression of HMGCR with both UBIAD1 constructs, cells were harvested and membranes were solubilized by the addition of DDM-CHS. Insoluble material was removed by centrifugation, and the soluble fraction containing the solubilized membrane proteins was applied to amylose resin previously incubated with an MBP-fused anti-eGFP nanobody. The MBP-fused eGFP nanobody construct was kindly provided by Oleg Sitsel. The bound proteins were washed and eluted by adding D-maltose to the elution buffer. Samples were taken at each step of the purification and analyzed by semi-denaturing SDS-PAGE. Since the gel is semi-denaturing, a shift of the HMGCR is expected after elution from the amylose beads. The MBP-fused anti-eGFP nanobody will remain bound to the eGFP-fused HMGCR, causing a shift in the gel toward higher molecular weights.

The gel shows that in both experiments, overexpressing WT and mutant (G186R) UBIAD1 along with the HMGCR, both proteins could be efficiently solubilized (Figure 36). An excess of unbound UBIAD1 could be detected in the flow-through. However, slightly more unbound UBIAD1 could be detected for the WT protein. Unbound HMGCR was also detected in the flow-through potentially corresponding to not properly folded protein where the tag is inaccessible. When maltose was added for elution of bound proteins, large quantities of UBIAD1 (WT and mutant) could be detected in the elution fractions coeluting with the HMGCR. This might indicate that potential complexes with the reductase were isolated. Eluted fractions were combined, concentrated and injected into a Superose 6 column to further purify the potential complex.

For both complexes containing WT UBIAD1 or the mutant form, the SEC profiles exhibited a similar pattern (Figure 37). In the case of WT UBIAD1, multiple peaks were observed (Figure 37A). A void volume was detected eluting at 8 ml. Additional peaks at 15 ml and 17 ml were observed likely corresponding to uncomplexed proteins or buffer components. Interestingly, a gaussian-shaped peak with a large shoulder towards higher elution volumes was detected at 12.5 ml elution volume. SDS-PAGE analysis of fractions of this peak reveals coelution of the HMGCR and UBIAD1. This elution volume matches the observation from small-scale screening where additional peaks for UBIAD1 were observed at 13 ml upon coexpression.



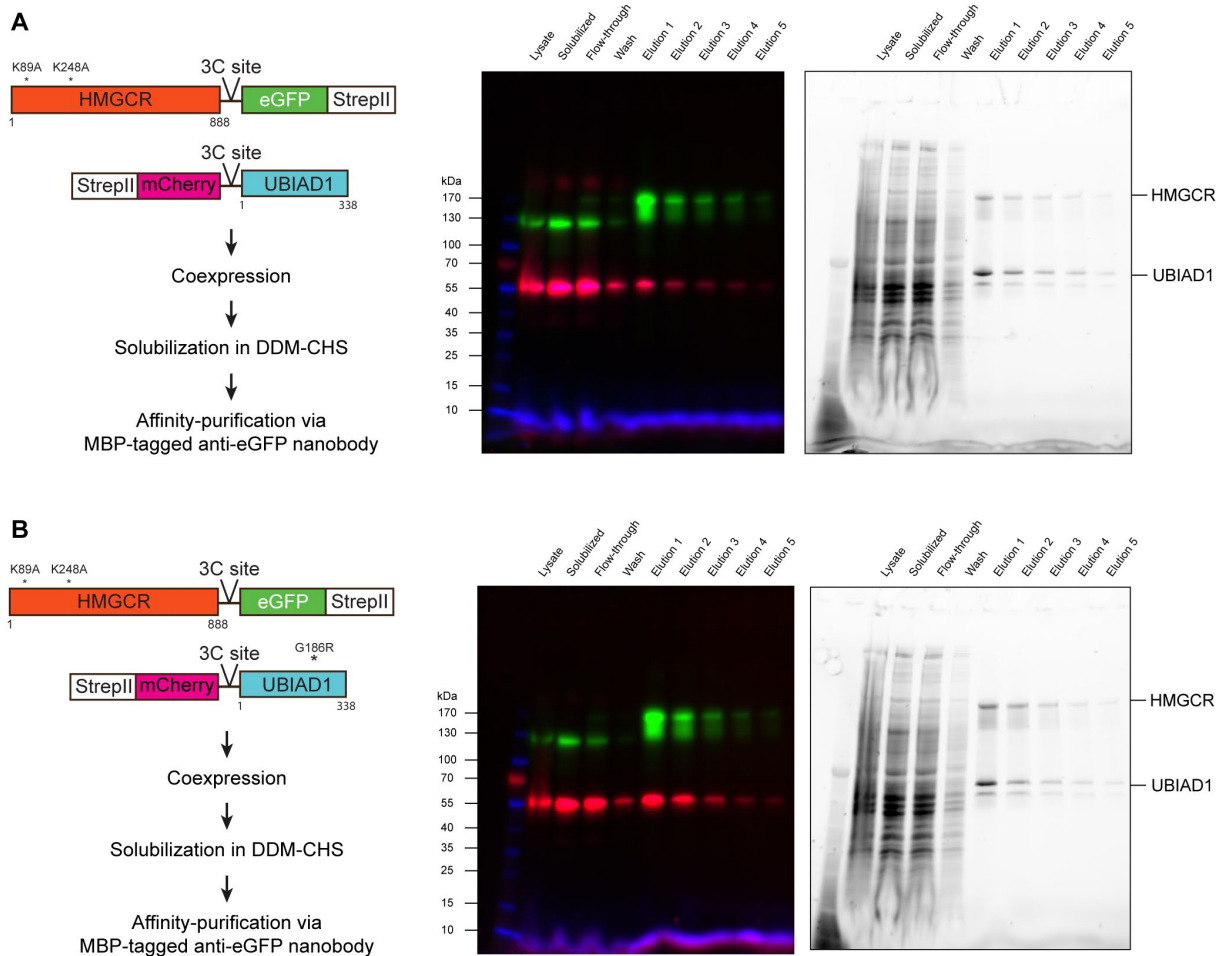


Figure 36: Single-step affinity purification of the HMGCR-UBIAD1 complex.

(A) & (B) The shown UBIAD1 constructs (WT and mutant) were coexpressed with a full-length ubiquitination-deficient HMGCR construct. Both proteins were fused to a fluorescent protein for easy detection. A single affinity purification step was performed using an MBP-labeled anti-eGFP nanobody immobilized on amylose beads. Bound proteins were eluted with D-maltose. Aliquots were loaded onto semi-denaturing SDS-PAGE and protein bands were visualized by in-gel fluorescence and stain-free detection. A shift of the HMGCR-specific band in the elution fractions is observed due to binding of the MBP-labeled anti-eGFP nanobody to the eGFP fused to the HMGCR.

The SEC profile of the UBIAD1<sup>G186R</sup> copurified with the HMGCR shows a similar shape to the WT UBIAD1 (Figure 37B). However, the height of the void volume is considerably lower. Furthermore, the main peak which elutes at 12.5 ml is more gaussian-shaped compared to the WT UBIAD1 but also has a shoulder towards earlier elution volumes. The SDS-PAGE analysis indicates that both proteins elute at this elution volume, possibly forming a complex. To characterize the sample homogeneity, a sample of the main peak was taken from each preparation and analyzed by negative staining EM.



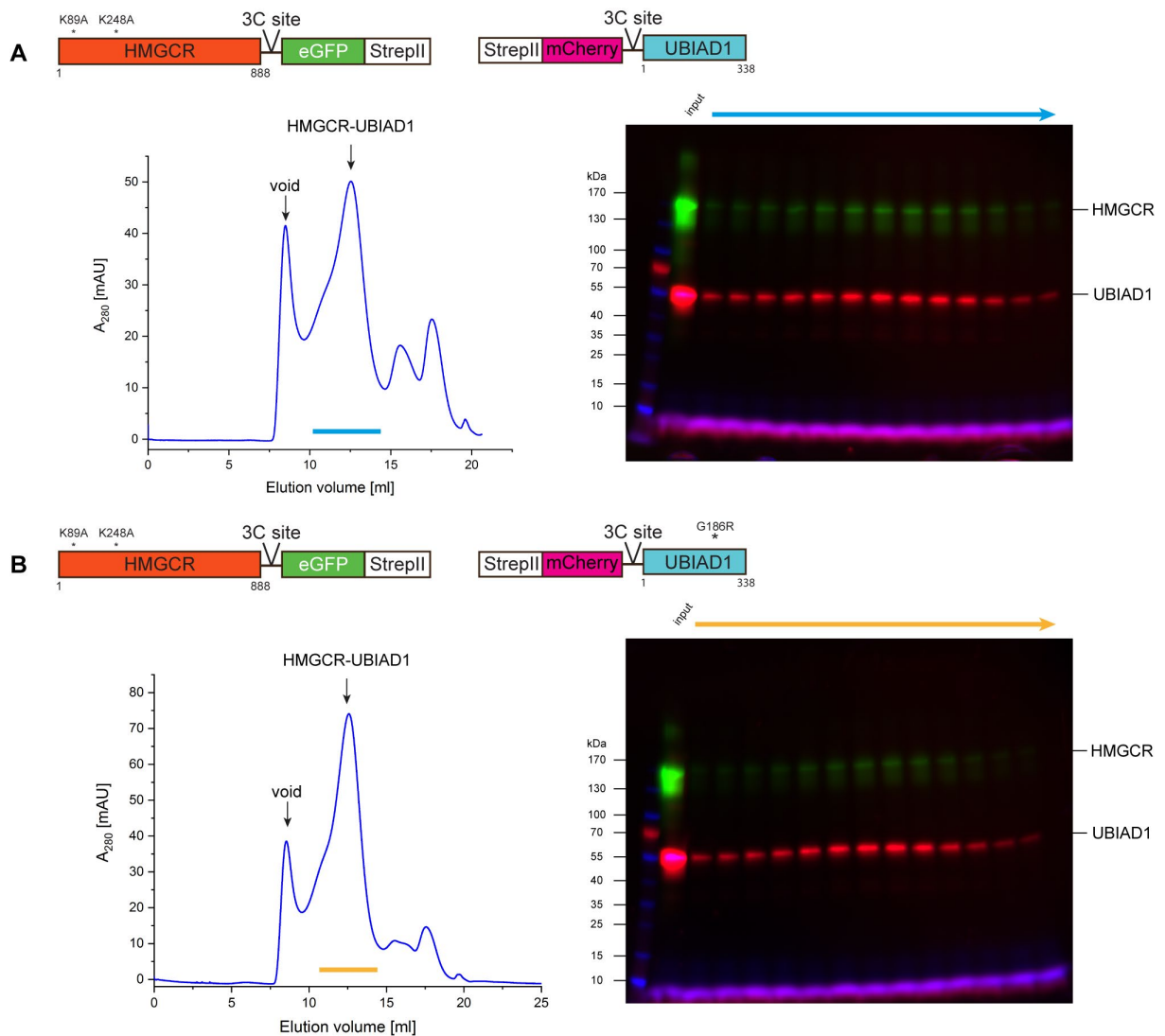


Figure 37: SEC profiles of HMGCR-UbiAD1 complexes.

The samples of the complexes with the HMGCR with (A) WT UbiAD1 and (B) UbiAD1<sup>G186R</sup> were concentrated after the single-step affinity purification, concentrated and injected into a Superose 6 column. Fractions of the main peak were analyzed by SDS-PAGE and protein bands were visualized by in-gel fluorescence.

For both constructs, a heterogeneous particle distribution was observed in negative stain EM micrographs (Figure 38). As described above, smaller particles with no features and a diameter of 5 nm likely correspond to empty detergent micelles. Larger particles range from 10-20 nm with a disc-shaped portion and a protruding globular domain. To further characterize particle shape, small data sets were acquired. Particle positions were manually selected and used for extraction and subsequent 2D classification using ISAC. The class averages confirm the observation of a heterogeneous distribution of particle size and shape. Smaller particles are observed with a disk-shaped region. These particles could correspond to the individual proteins of complexes that have fallen apart. Larger particles could correspond to larger oligomers of HMGCR alone or of complexes formed with UbiAD1. Since a single-step affinity purification

was performed, a mixture of different subcomplexes and uncomplexed particles was expected. Therefore, the next step would be to optimize the constructs and the purification protocol to obtain a more homogeneous preparation.

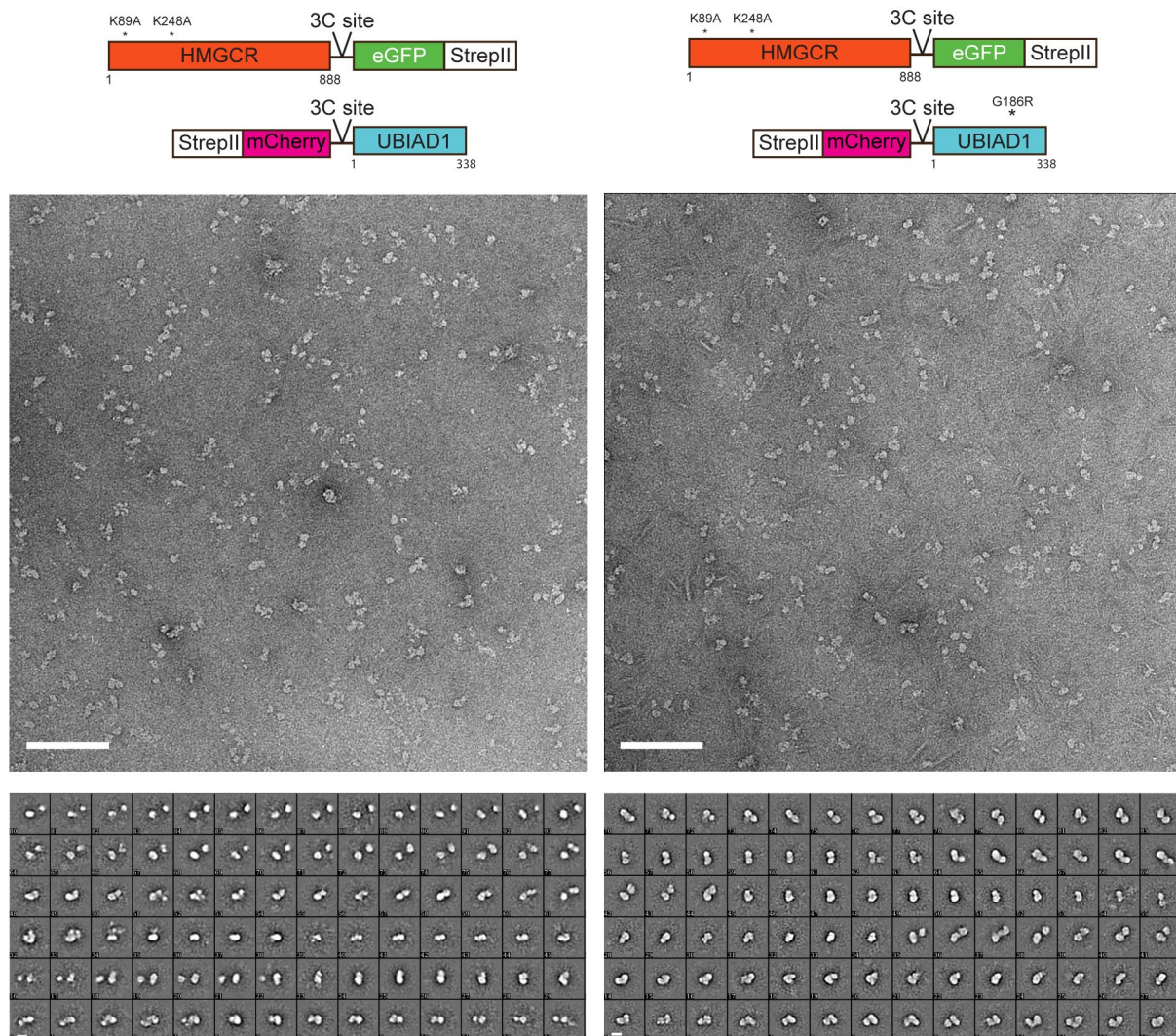


Figure 38: Negative stain EM analysis of HMGCR-UBIAD1 complexes.

The indicated samples were purified via a single-step affinity purification and SEC in DDM-CHS. The samples from SEC were applied to negative stain grids and subsequently visualized. A representative micrograph is shown. Datasets were recorded and 2D class averages using ISAC were calculated. Scale bar micrograph, 100 nm. Scale bar 2D class averages, 10 nm.

### 9.2.6 Optimization of the HMGCR-UBIAD1 complex purification

In order to increase the sample homogeneity, the UBIAD1 construct was modified. The StrepII-mCherry tag was replaced by a 3XFLAG-eYFP tag. Important is the change to a 3XFLAG tag in order to be able to perform a tandem-affinity purification. Based on the SEC profiles above (Figure 37) it seems that the UBIAD1<sup>G186R</sup> is better suited for the complex purification due to the observation that less of a void volume can be detected compared to the complex with the

WT UBIAD1. Hence, the mutant UBIAD1<sup>G186R</sup> was further selected to study the interaction with the HMGCR.

High heterogeneity was observed when the HMGCR was purified alone, as described above. The HMGCR probably formed different oligomeric assemblies and the C-terminal cytosolic portion may be the main factor for oligomerization. To solve the structures of the Scap-Insig complex, the deletion of the flexible CTD of Scap helps with a more rigid particle alignment focusing on the TMD of the complex during image processing. Here, the interaction between UBIAD1 and the HMGCR is solely due to their TMDs. Thus, the C-terminal domain appears to be indispensable for complex formation and has been deleted. The last TM helix of the HMGCR ends after residue 339. Thus, residues 340-888 were deleted and the construct was named HMGCR<sup>TMD</sup>. Both proteins, HMGCR<sup>TMD</sup> and UBIAD1, are exclusively embedded in the membrane without additional extramembranous domains. This could pose a problem during image processing to guide faithful particle alignment. Fab fragments have been reported to successfully add an ordered and sufficiently large handle to small membrane proteins for accurate particle alignment.<sup>235</sup> However, they require immunization of an animal and subsequent screening of hybridoma libraries to identify binders to the target. A universal approach has been described by fusion of an engineered version of apocytochrome b562 RIL (BRIL), which is grafted into the protein construct.<sup>84</sup> Ideally, the BRIL is fused to two adjacent TM helices, replacing the connecting loop but can also be fused to the termini of the protein. Universal anti-BRIL Fab fragments have been generated that can bind to the extramembranous BRIL domain, which can be used as a fiducial for particle alignment.<sup>84</sup> This approach has been successful in the determination of a cryo-EM of Frizzled 5.<sup>236</sup> Here, several constructs varying the position of the BRIL insertion in loops of the HMGCR have been screened. However, insertion into the loops of the HMGCR resulted in low expression yields and undesirable FSEC profiles. Eventually, a stably expressing and well behaving construct was obtained when BRIL was fused to the C-terminus of HMGCR<sup>TMD</sup>, termed HMGCR<sup>TMD</sup>-BRIL (Figure S 6). In addition, the rigid fusion of the BRIL domain to the TMD of HMGCR was validated by calculating an AlphaFold<sup>211</sup> model, which shows that the BRIL is tightly fused to the end of the TM8 helix of the reductase (Figure 39).

To assess if the modified constructs form a stable complex, they were coexpressed in HEK293 GnTI<sup>-</sup> cells by transduction of two BacMam viruses each carrying the respective construct (Figure 40A). The cells were harvested after 48h of infection and directly used for solubilization with DDM-CHS. Insoluble material was removed and solubilized complexes were enriched by tandem-affinity purification using FLAG-M2 and Strep-Tactin beads. The purified complex

was mixed with the universal Fab fragment BAG2 which binds to BRIL and further purified via SEC using a Superose 6 column. The constructs for BRIL and BAG2 were kindly provided by the Kossiakoff lab.<sup>84</sup> Samples were taken from each step of the purification and analyzed via semi denaturing SDS-PAGE.

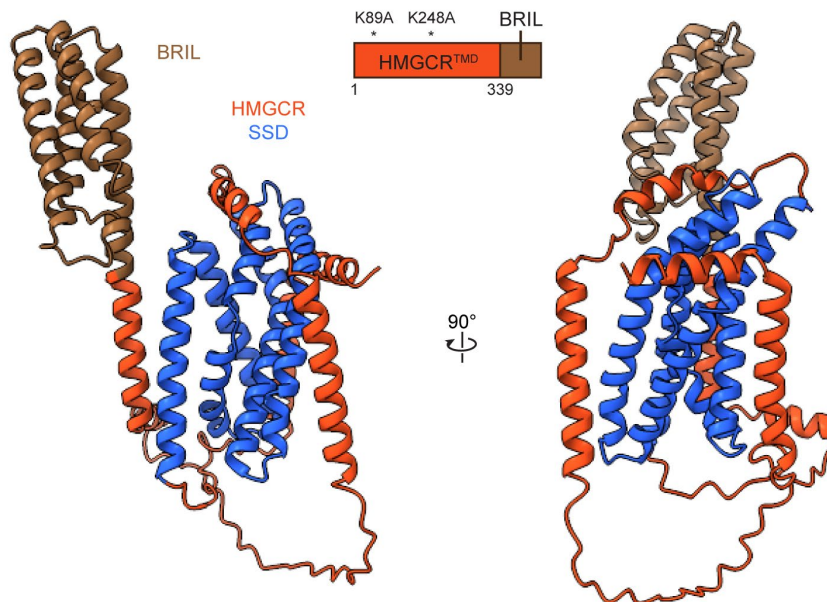


Figure 39: AlphaFold prediction of the HMGCR<sup>TMD</sup>-BRIL construct.

The SDS-PAGE displays the presence of two major bands after elution from the second affinity column migrating at 55 and 70 kDa, respectively, corresponding to the HMGCR<sup>TMD</sup>-BRIL and UBIAD1<sup>G186R</sup> (Figure 40B). Both proteins copurified after tandem-affinity purification and no major contaminations could be detected indicating a pure preparation of the complex.



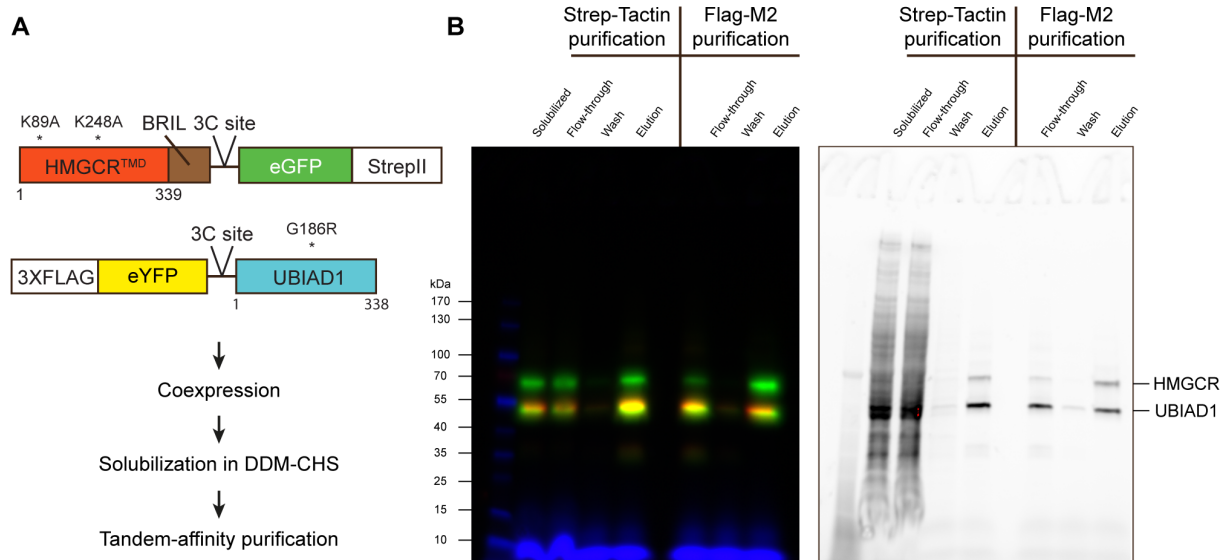


Figure 40: Purification of the HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup> complex.

(A) The HMGCR construct was truncated by the C-terminal cytosolic domain and C-terminally fused to a BRIL domain. This construct was coexpressed with the Flag-tagged UBIAD1 construct and purified via tandem-affinity purification. (B) The SDS-PAGE of the tandem-affinity purification reveals isolation of a stable and pure HMGCR-UBIAD1 complex. Protein bands were visualized by in-gel fluorescence and stainfree detection.

For further analysis of sample quality and complex characterization, the elution fractions were pooled, concentrated and treated overnight with 3C protease to remove purification tags. A twofold excess of BAG2 was added to form a ternary complex with the HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup> complex. The sample was injected into a Superose 6 column to separate the cleaved purification tags from the complex as well as unbound Fab fragments. The SEC profile shows three main peaks as well as a void volume (Figure 41A). A peak containing the cleaved purification tags is detected at 17.5 mL. An excess of the Fab fragment BAG2 is detected at an elution volume of 19 ml. Interestingly, a gaussian-shaped main peak eluting at 14 ml is observed with a shoulder towards higher molecular weight species. This peak may contain the ternary complex of HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup>-BAG2. To further analyze the sample and verify that the Fab fragment was bound, a fraction was taken from the main peak and analyzed by negative staining EM.

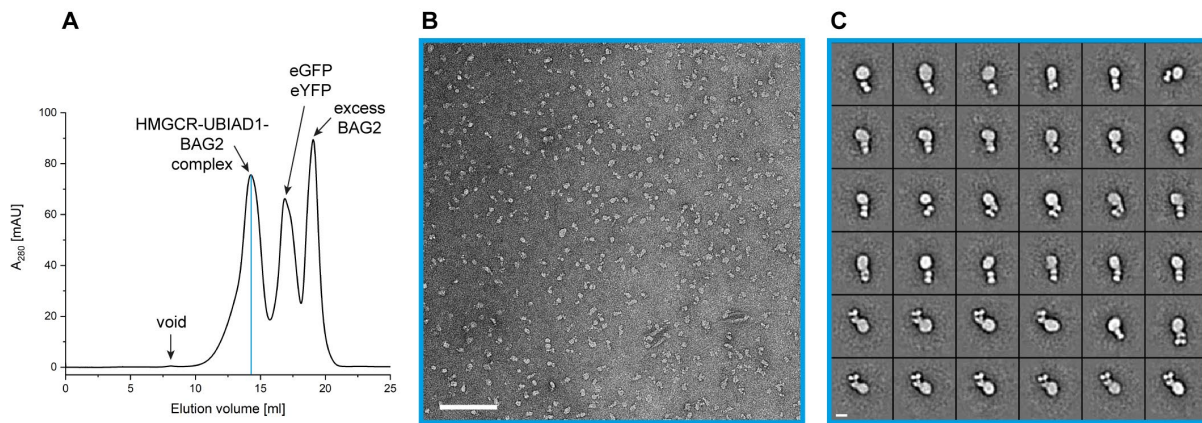


Figure 41: Characterization of the HMGCRTMD-BRIL-UBIAD1<sup>G186R</sup>-BAG2 complex.

(A) SEC profile of HMGCRTMD-UBIAD1-BAG2 complex. The HMGCRTMD-UBIAD1 complex was purified via tandem-affinity purification. Purification tags were cleaved off and BAG2 was added with two-fold excess to reconstitute a ternary complex composed of HMGCRTMD-BRIL, UBIAD1<sup>G186R</sup> and BAG2. The complex was injected into a Superose 6 column. (B) The indicated fraction from SEC was analyzed by negative stain EM revealing a homogenous particle distribution. Scale bar, 100 nm. (C) Representative 2D class averages of the ternary complex reveal high particle homogeneity. Scale bar, 10 nm.

The micrographs reveal a homogenous particle distribution with an average diameter of 12 nm (Figure 41B). All particles are characterized by a roundish disk-shaped portion and a thin elongated density protruding from the disk. To further analyze the particle shape, a negative stain EM dataset was recorded. Particle locations were manually selected and coordinates were used for extraction. 2D class averages were computed using ISAC implemented in SPHIRE. Indeed, the class averages reveal a homogenous and uniform particle shape (Figure 41C). The particles have a disk-shaped portion with a protruding density. The protruding density corresponds to the bound Fab fragment, as revealed by its characteristic shape generated by the heavy and light chains, which can be observed even at low resolution. However, higher resolution is required to confirm the presence of both proteins in the complex.

The homogenous sample was used to prepare cryo-EM grids with a thin vitreous ice layer embedding the particles. Here, QF 1.2/1.3 grids gave the best particle distribution. Grids were screened in a Talos Arctica 200 kV microscope. Initial images confirmed the homogenous particle distribution in a thin layer of vitreous ice without the presence of aggregates. A good grid was transferred to a Titan Krios 300 kV microscope for data collection. Several thousand micrographs were acquired in super-resolution mode using a K3 direct electron detector with an unbinned pixel size of 0.44 Å/pix. Particle locations were automatically detected using the general model of crYOLO. The coordinates were used to extract the particles and calculate 2D class averages using ISAC embedded in SPHIRE.

The 2D class averages confirm the shape observed for particles in negative stain EM. A disc-shaped part, which is the micelle, of varying diameter is observed with a small density protruding out, which belongs to the Fab fragment. Different views representing side and top views of the complex are observed indicating a uniform orientation on the grid. However, it is also evident from the 2D classes that the position of the Fab fragment ensembles different positions suggesting that it might be flexibly linked to the TMD of the HMGCR. In some classes, BAG2 is oriented almost perpendicular to the micelle, whereas in other classes, the average orientation is almost parallel to the micelle. In addition, some classes do not reveal the additional density of the Fab fragment, indicating that not all complexes were saturated with the Fab fragment or that the Fab fragment is indeed flexible. The good class averages with clear density for BAG2 were selected and an initial model was calculated which served as reference for 3D refinement. The final reconstruction shows a limited resolution, where only the disc-shaped part corresponding to the detergent micelle can be reconstructed. No features appear inside the micelle. A density corresponding to the Fab fragment could not be resolved, indicating the flexibility of the BRIL-BAG2 domain relative to the TMDs of the complex. The fiducial will only aid in particle alignment if it is held rigidly in place. In the future, other BRIL designs including the insertion into candidate positions in UBIAD1.

As discussed above, the first structure of the Scap-Insig complex, composed only of the TMDs, was solved by truncation of the cytosolic domain of Scap without addition of a fiducial. It was suspected that a fiducial may not necessarily be needed to determine the structure of the HMGCR-UBIAD1 complex. Hence, I purified the complex without addition of the Fab fragment.

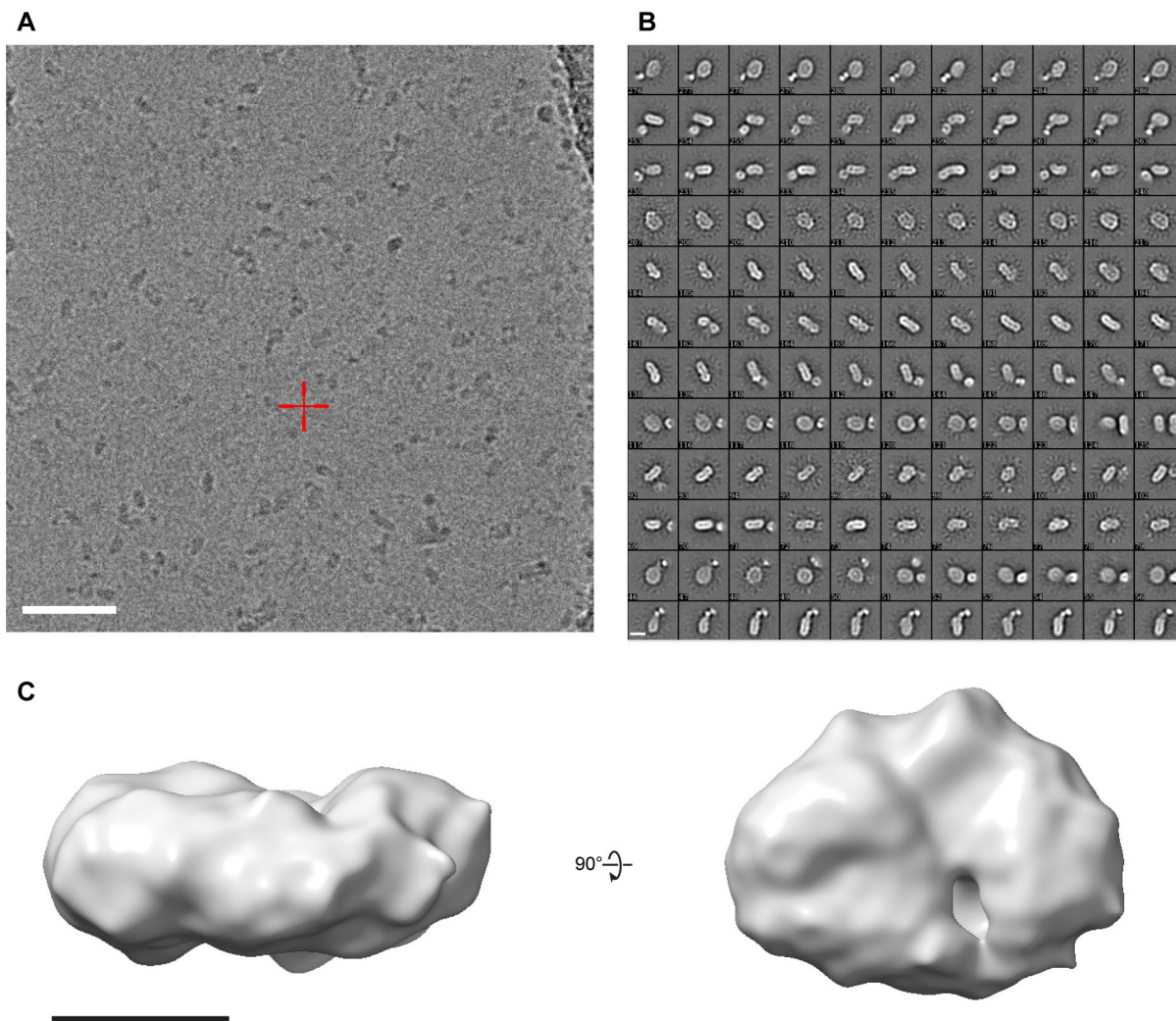


Figure 42: Cryo-EM analysis of the HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup>-BAG2 complex.

(A) Representative cryo-EM micrograph during optimization of freezing conditions, imaged on a Talos Arctica 200 kV instrument. A grid of thin ice was transferred to a Titan Krios 300 kV microscope for data collection. Scale bar, 50 nm. (B) Representative 2D class averages generated with ISAC integrated into SPHIRE. The class averages show that the particles in the vitreous ice layer are randomly oriented. A disk-shaped density with the protruding Fab fragment BAG2 is observed. However, the Fab fragment shows flexibility. Scale bar, 10 nm. (C) Final reconstruction of the complex. The Fab fragment is not resolved, indicating the flexibility of this part of the complex. The overall resolution was limited to 15Å. Scale bar, 5 nm.



### 9.2.7 HMGCR-UBIAD1 complex in LMNG-CHS

The HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup> complex was purified as described above. Different purification conditions were tested to further stabilize the complex, including the use of different detergents for solubilization and purification. Finally, LMNG-CHS was used for solubilization and purification, resulting in a pure preparation after tandem affinity purification. The complex was concentrated, treated with 3C protease to remove purification tags, and injected into a Superose 6 column as a final purification step. The SEC profile shows two major peaks with no detectable void volume (Figure 43A). Two peaks were observed eluting at 15 and 17.5 ml. The latter peak contains cleaved purification tags. The main peak elutes at 15 ml and is gaussian-shaped with a small shoulder at earlier elution volumes. Negative staining EM analysis of this fraction reveals a homogeneous particle distribution with particle diameters ranging from 10-14 nm (Figure 43B).

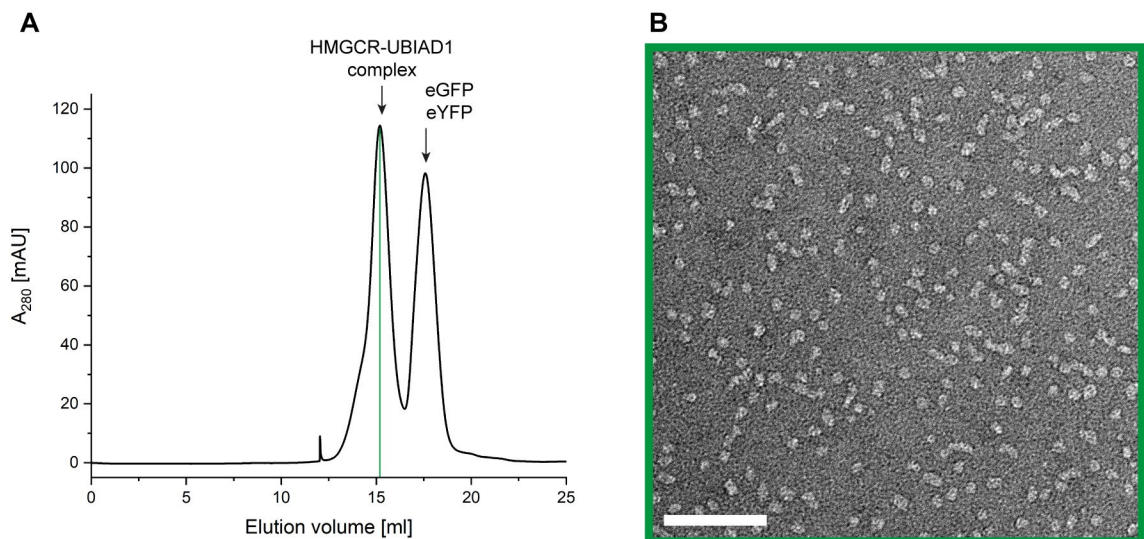


Figure 43: Characterization of the HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup> complex in LMNG-CHS.

(A) The complex solubilized in LMNG-CHS and tandem-affinity purified. The purification tags were cleaved off by 3C treatment and separated from the complex via SEC. (B) Negative stain EM micrograph of the indicated fraction containing the HMGCR-UBIAD1 complex. Scale bar, 100 nm.

The complex was concentrated to 1.5 ml/ml and screened for a homogenous particle distribution in a thin vitreous ice layer. Again, QF 1.2/1.3 grids produced the best particle distribution with a single layer of particles. The grid was transferred to a Titan Krios 300 kV microscope equipped with a K3 direct electron detector for data collection (Figure 44A). An overview of the parameters of the data collection and image processing are listed in Table 3. A detailed image processing workflow is shown in Figure S 7.

Table 3: Statistics of the cryo-EM data collection and image processing of the HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup> complex purified in LMNG-CHS.

	<b>Scap-Insig in LMNG</b>
<b>Data collection</b>	
Microscope	Titan Krios (X-FEG, Cs-corrected)
Magnification	81,000
Voltage (kV)	300
Defocus range ( $\mu\text{m}$ )	-0.8 to -2.2
Camera	K3 super-resolution mode
Pixel size ( $\text{\AA}/\text{pixel}$ )	0.44 $\text{\AA}/\text{pix}$
Total electron dose ( $\text{e}^-/\text{\AA}^2$ )	60
Exposure time (s)	2
Number of images	8901
<b>Image processing</b>	
Number of picked particles	4,351,845 (crYOLO)
Number of particles after 2D classification	876,936
Number of final particles	190,851
Final resolution ( $\text{\AA}$ )	7.3 $\text{\AA}$
Symmetry	$C_1$
Map sharpening B factor ( $\text{\AA}^2$ )	-100

Particle locations were automatically detected using the general model of crYOLO. The coordinates were used to extract the particles and calculate 2D class averages in cryoSPARC. The class averages show a disc-shaped particle corresponding to side views and top views indicating that the complex is vitrified in all orientations (Figure 44B). They reveal high-resolution features that can be seen as stripes inside the micelle corresponding to the TMDs of the complex. Furthermore, 2D classification reveals that at least two different species of varying size are present on the grid. One species is composed of smaller sized particles with an average diameter of 10 nm. Based on the results obtained for the Scap-Insig complex, this assembly likely represents a 1:1 complex. Interestingly, particles with a larger but varying diameter are observed which likely correspond to higher oligomeric assemblies of the complex.

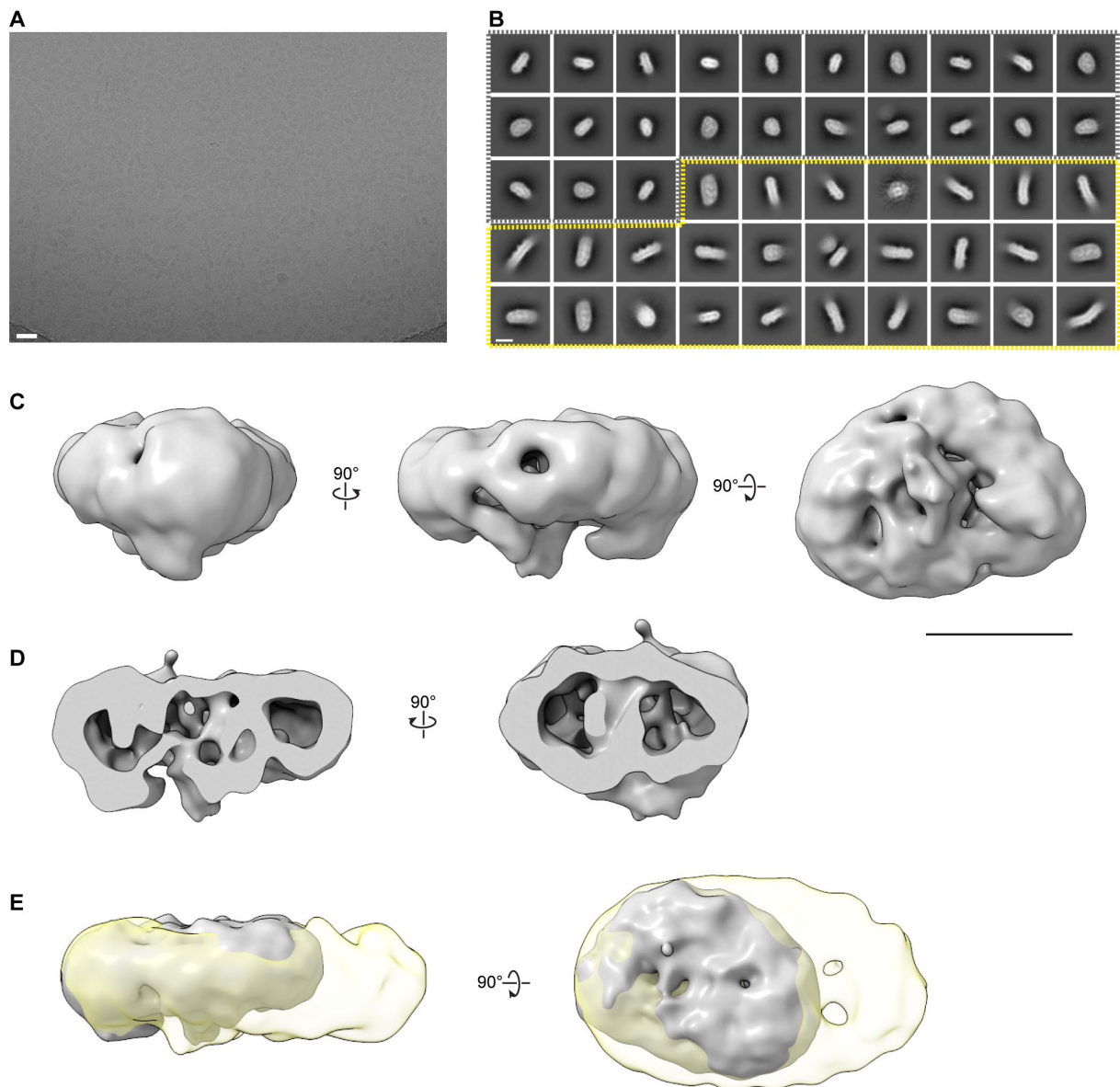


Figure 44: Cryo-EM analysis of the HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup> complex in LMNG-CHS.

(A) Representative MotionCor2-corrected micrograph showing a homogeneous particle distribution of the HMGCR-UBIAD1 complex in pure ice. Scale bar, 20 nm. (B) Representative 2D class averages showing the distribution of smaller, potentially monomeric 1:1 complex of HMGCR and UBIAD1 in gray encircled. Larger particles were also observed, highlighted in yellow. Class averages were calculated in cryoSPARC. Scale bar, 10 nm. (C) The final 3D reconstruction of the smaller particles shows the limited resolution of a disk-shaped or elliptical particle. Small parts protrude from the micelle, likely representing the TMDs of the complex. Scale bar, 5 nm. (D) Cut-open view of the micelle shows an unstructured density in the interior. (E) Comparison of the final reconstruction obtained for the smaller and larger sized particles.

The particles were separated for an initial model generation and subsequent refinement. Larger sized particles were processed separately but the final 3D reconstruction was limited in resolution. Thus, it was focused on the smaller more homogeneous particles.

The smaller sized particles were selected from 2D class averages and an initial model was generated. After multiple rounds of heterogeneous refinement and cleaning of the dataset, a final subset of 190K particles was refined. The final reconstruction nominally displayed a resolution of 7.3Å (Figure 44C). However, based on manual inspection of the map, this resolution is likely overestimated by refinement algorithm. The reconstruction matches the expected shape of an elliptical shaped particle. Even though the BRIL domain was still included in the HMGCR construct it is not resolved in the reconstruction. As described above, the BRIL domain likely exhibits flexibility hindering its reconstruction. Small densities protrude out of the micelle. These small bumps likely represent the TMDs of the complex which are embedded in the micelle. A closer look inside the detergent micelle shows the presence of an additional but unstructured density (Figure 44D). These densities correspond to TMDs of both proteins. Again, based on the size of the particle a 1:1 complex of HMGCR and UBIAD1 would match the observed dimensions of the reconstruction. The features of the TMDs could be recovered to some extent in the 2D classes but could not be restored in the final 3D reconstruction. Other processing strategies were explored but could not increase the overall resolution of the reconstruction.

Interestingly, when comparing the reconstruction of the smaller and larger size particles reveals that the larger sized particles are nearly double the size of the potentially 1:1 complex (Figure 44E). Thus, it seems conceivable that higher oligomers of the HMGCR-UBIAD1 complex are formed and evenly distributed. Above, I speculated that the cytosolic domain of the HMGCR might be the main driving force for formation of higher oligomers. But given the observation of higher oligomeric species of unknown stoichiometry, it seems conceivable that the TMD of HMGCR might be also capable to drive formation of higher oligomers. However, to unambiguously confirm this hypothesis a higher resolution structure of these complexes is required.

### 9.2.8 Conclusion HMGCR-UBIAD1 complex

The HMGCR catalyzes the rate-limiting step in the de novo synthesis of cholesterol. Several crystal structures of the catalytic portion of the enzyme were solved more than 20 years ago, explaining the catalytic mechanism of the enzyme as well as the mode of action of statins. Statins are potent inhibitors that bind to HMGCR and block de novo cholesterol synthesis. They are among the most widely prescribed drugs worldwide to lower blood cholesterol levels. However, shortly after the start of treatment with statins elevated reductase levels can be detected in the liver of patients and attenuate the effect of statins. These increased levels of reductase after statin treatment mainly arise from a perturbed feedback system converging to control the amount of HMGCR that is present in the membranes. These feedback systems control reductase levels at multiple levels, including control of HMGCR gene transcription by the SREBP signaling pathway, post-translational modification that reduces catalytic activity, and ubiquitination of reductase leading to proteasome-mediated degradation via ERAD. Studies have shown that statin-induced accumulation of the HMGCR in ER membranes is mainly due to sequestration of UBIAD1 in the ER. UBIAD1 binds to the TMD of the HMGCR and acts as a stabilizer. In contrast, Insigs counteracts the stabilizing role of UBIAD1 by recruiting a degradation machinery to the TMD of the reductase marking it for degradation by ERAD. However, the molecular basis of binding these regulatory mechanisms remains elusive. Therefore, my research aims to fill this gap by solving high-resolution structures of the full-length HMGCR or its complexes with interacting partners.

First, I cloned the full-length human HMGCR into vectors that allow overexpression in a more native expression system rather than insect cells. I successfully overexpressed the full-length human HMGCR in HEK293 GnTI- and was able to set up a complete purification protocol. I have performed a small scale detergent screen to find that the reductase is best solubilized in DDM-CHS. I then scaled up the expression and purified the membrane protein using the DDM-CHS detergent mixture. To test that the enzyme was purified in an active and relevant form, I performed Michaelis-Menten enzyme kinetics. I found that the  $K_M$  and  $V_{max}$  values were consistent with the parameters reported for the catalytic portion of the enzyme. Notably, at high substrate concentrations, the full-length HMGCR exhibited a behavior known as substrate inhibition, which has not been previously reported for this enzyme. At low substrate concentrations, an exponential increase in reaction rate is observed, reaching  $V_{max}$ . However, further addition of substrate leads to a decrease in reaction velocity. Such behavior is often observed for other metabolic enzymes such as phosphofructokinase. Rapid supply of products is sought when they are in shortage. However, after an initial burst of product production,

slower reaction rates result in a constant but continuous flow of reaction products. This mechanism helps to maintain a homeostatic level of production without depletion or accumulation of downstream products. Therefore, such a mechanism for HMGCR seems very conceivable and needs to be further investigated in the future. Given that an active form of the HMGCR was purified, I proceeded with the structural characterization via negative stain EM and cryo-EM. A plethora of purification conditions were tested, including buffer optimization, use of different detergents, addition of cofactors or inhibitors. However, upon visualization using negative stain EM and cryo-EM analysis, a very heterogeneous sample preparation was observed. At this point it was concluded that potentially the presence of a stabilizing interaction partner might help in order to obtain a sample suitable for structural characterization of the HMGCR.

As mentioned above, the HMGCR is the target of multiple feedback mechanisms that control the turnover of this critical enzyme. Two key players counteract each other and compete to bind to the TMD of the reductase to either recruit a degradation machinery or stabilize it. Insigs are membrane-embedded oxysterol sensors that recruit the E3 ligase gp78 to the TMD of the HMGCR when sterol levels in the membrane are elevated. This in turn ubiquitinates the HMGCR and marks it for degradation via ERAD, shutting down cholesterol synthesis. This is counteracted by UBIAD1, which protects the HMGCR from degradation when cholesterol levels are low. I started to study the interaction of Insig1 and HMGCR by co-expressing both proteins followed by small-scale purification attempts. Coexpression of WT HMGCR with Insig resulted in no detectable expression of the HMGCR, probably due to increased ERAD caused by co-expression of Insig. Instead, two mutations rendering the HMGCR resistant to ubiquitination were introduced and coexpressed with Insig1. This time, expression of both proteins was detected. However, I could not detect any significant interaction when I tried to co-purify both proteins. This may indicate that other factors are required to form the correct complex. However, when both WT proteins were overexpressed, the HMGCR was presumably efficiently degraded by Insig, showing that potentially all factors are present in the expression host to form a complex. Another possibility is that the complex is very transient and not stable enough during purification. More rigorous testing, including screening for stabilizing mutations, is needed in the future to find the underlying principles for purifying a stable HMGCR-Insig complex suitable for structural determination.

I then shifted my focus to the interaction of the HMGCR with UBIAD1. To study their interaction, I coexpressed the WT UBIAD together with the HMGCR in mammalian cells. In addition, I designed a second UBIAD1 construct that harbored a well-known point mutation

commonly found in SCD patients that increased the affinity for the HMGCR. This construct was named UBIAD1<sup>G186R</sup>. I first coexpressed both UBIAD1 constructs separately with the full-length HMGCR and screened their interaction on a small scale using FSEC. Indeed, when UBIAD1 constructs were coexpressed with HMGCR, UBIAD1 could be detected at earlier elution volumes compared to UBIAD1 expressed without HMGCR, indicating the formation of potential complexes. To verify this finding, I scaled up the expression and established a purification protocol using single-step affinity purification and SEC. Indeed, both proteins co-purified and formed a complex as judged by SEC and SDS-PAGE analysis. The sample quality of the purified complex was verified by negative staining EM. Inspection of the micrographs and calculation of 2D class averages from negative stain images revealed a heterogeneous particle distribution as seen for HMGCR only before. It is possible that a mixture of oligomers was purified, resulting in a suboptimal sample for structural determination. Most likely, such oligomers were induced by the C-terminal domain of the HMGCR, which is flexibly anchored to the TMD. The linker connecting the catalytic domain to the TMD consists of more than 100 amino acids with no strong predicted secondary structure. Moreover, this linker adopts an unstructured conformation when calculated by Alphafold models, which could induce a high degree of flexibility of the C-terminal domain. Therefore, I designed a HMGCR construct truncated by the flexible linker and the C-terminal catalytic portion. This construct contains only the TMD of the HMGCR, termed HMGCR<sup>TMD</sup>. In addition, a fiducial marker was introduced by fusing a BRIL domain to the C-terminal helix of HMGCR<sup>TMD</sup>. This construct was named HMGCR<sup>TMD</sup>-BRIL and was used to study the interaction with UBIAD1.

During several purification optimization steps, I found that a slightly more stable complex was purified using the disease mutant G186R of UBIAD1. In addition, a 3XFLAG tag was fused to UBIAD1, which allowed tandem affinity purification and improved the homogeneity of the preparation. This construct of UBIAD1 was coexpressed with HMGCR<sup>TMD</sup>-BRIL and subsequently purified by tandem affinity purification and SEC. Prior to SEC, a ternary complex was formed by addition of the universal Fab fragment binding to BRIL, BAG2. The HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup>-BAG2 complex showed high homogeneity when examined by negative staining and cryo-EM. A disk-shaped particle was observed corresponding to the detergent micelle with the Fab fragment protruding. Unfortunately, the fiducial exhibited flexibility that did not help to properly align the particles during image processing.

Instead, the purification was optimized by screening more detergents and skipping the addition of the Fab fragment. Finally, the complex consisting of HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup> was solubilized and purified in LMNG-CHS. The complex migrated as a single gaussian-shaped

peak in SEC and showed high homogeneity in negative stain EM and when vitrified in pure ice. Unfortunately, cryo-EM images of the complex lacked appropriate features to resolve the complex at high resolution. Despite these efforts, I obtained a reconstruction with a resolution of 7.3Å. Based on the overall dimensions of the reconstruction, a 1:1 complex seemed likely. However, a higher resolution to confirm this hypothesis could not be obtained. Image processing without a stable and rigid fiducial proved to be very challenging. Interestingly, during the analysis, another species was identified on the grid with almost twice the size of the potential 1:1 complex. This suggests that the complex may also exist in larger oligomeric forms, possibly as a dimer. However, the resolution of this particle was also limited, preventing its characterization.

During this time of my research, I came across a publication reporting the high-resolution structure of the hamster HMGCR-UBIAD1 complex in two different conformations.<sup>83</sup> The authors reported several modifications of the constructs to obtain a stable complex. The HHMGCR was modified by truncation of the catalytic C-terminal portion. In addition, two mutations were introduced into the TMD by replacing K89 and K248 with arginine. UBIAD1 was truncated at amino acids 1-40, which are presumably flexible and cytosolic. A mutation N102S was introduced to attenuate the enzymatic activity. This is also a known SCD mutation that enhances the association of UBIAD1 with the HMGCR. Both proteins were coexpressed and purified as a complex by a single step affinity purification and SEC. But when imaged using cryo-EM and subsequent analysis, the lack of extramembranous fiducials hampered the correct particle alignment and limited the resolution. A BRIL domain was between TM 8 and 9 of UBIAD1 replacing a short cytosolic loop. This UBIAD1 construct was coexpressed with the TMD of the HMGCR and purified as complex. Addition of the anti-BRIL Fab and an anti-Fab nanobody facilitated the structural determination to a resolution of 3.6 Å. The structure shows that the TMDs of the HMGCR and UBIAD1 form a 1:1 complex. However, the resolution in the TMD region was limited and prevented the construction of an atomic model.

To overcome this hurdle, a specific Fab fragment was generated that specifically binds to the TMDs of the complex, resulting in the successful structural determination of the TMDs to a resolution of 3.3 Å. Notably, the HMGCR-UBIAD1 complex was observed in two states, a monomeric or a dimeric 1:1 complex. However, the authors commented that the dimeric form is likely an artifact of detergent solubilization and is unlikely to exist *in vivo*. Interestingly, different conformations of the HMGCR-TMD were observed, termed conformation A and B. While most of the TMD, especially the SSD, adopted the same conformation in both states, a



drastic rotation of TM2 is observed between conformation A and B. The YIYF motif, crucial for the interaction with Insig, is located on TM2. Comparing the TMD of HMGCR in different conformations with the Insig-bound form of Scap, a rotation of TM2 is required to either expose or hide the YIFY motif for interaction with Insig.

These structures shed light on the conformational dynamics of the TMD of the HMGCR and provide a rationale for the molecular requirements that allow interaction with Insigs. Despite these advances, it remains unclear how sterols modulate the conformation of the HMGCR and how increased levels of GGPP in the membrane lead to dissociation of UBIAD1 from the HMGCR, allowing interaction with Insigs.

I have successfully established a complete purification protocol of the human HMGCR<sup>TMD</sup>-BRIL-UBIAD<sup>G186R</sup> complex. From the initial identification of the interaction, several optimizations were required to obtain a homogeneous preparation suitable for structural characterization. Data collection and subsequent 2D classification revealed a mixture of different oligomers within the purified complex. Unfortunately, the resulting reconstruction had limited resolution, likely due to the lack of extramembranous fiducials to guide particle alignment. The next step would have been to generate fiducial markers that are tightly bound to the TMDs of the complex, such as Fab fragments or nanobodies. These efforts will pave the way for the structural determination of more structures of the human complex in different conformations and under different conditions, i.e. loaded with different oxysterols or in the Insig-bound form. Ultimately, these structures will contribute to a comprehensive understanding of the sterol-mediated regulation of the HMGCR and may provide a potential starting point for the design of the next generation of cholesterol-lowering drugs.

## 9.3 Concluding Discussion

### 9.3.1 Structural determination of small membrane proteins using cryo-EM

Membrane proteins fulfill crucial functions in living systems, including transport of molecules across membranes, involvement in energy production or inter-cellular communication. The human genome encodes approximately 20,000 proteins, of which approximately 25% are membrane proteins.<sup>237</sup> Their importance is further underscored by the fact that 50-60% of all FDA-approved drugs target membrane proteins, with approximately 25% of these drugs targeting G-protein coupled receptors (GPCRs).<sup>238</sup> However, there is a clear mismatch between their importance and the number of solved structures deposited in the Protein Data Bank (PDB). As of July 2023, 206,656 entries of solved protein structures are listed in the PDB, but only 7128 are represented by membrane proteins, of which 1608 entries are unique membrane protein structures according to the "Membrane Proteins of Known 3D Structure" database. This discrepancy is even more dramatic when looking at membrane proteins with a molecular weight below 100 kDa. For this class of membrane proteins, only 147 entries are listed in the PDB with structures determined below a resolution of 5 Å.

Cryo-EM has been the method of choice to determine membrane protein structures in the recent years. The “resolution revolution” in cryo-EM<sup>239</sup> has led to an exponential growth of determined structures using this method compared to X-ray protein crystallography and nuclear magnetic resonance (NMR) spectroscopy. Advances in structural determination using cryo-EM include the development of direct electron detectors<sup>240,241</sup> which allow the recording of movies which help to account for beam-induced motion. But they also include development of better microscopes and software algorithms for data processing. Recent advances that ease the structural analysis and model building include the development of AlphaFold2<sup>211</sup> which provides calculated three-dimensional protein models with high accuracy. However, experimental structures are still key in order to understand the functions and mechanisms that are fulfilled by these proteins which will be supported by computational methods. With all these advances, the remaining challenge at present and in the future will be the sample preparation and quality of the purified protein that is applied to the grids as said by Werner Kühlbrandt in 2022: *“Therefore, when all other technical and computational issues are settled, protein biochemistry is likely to remain as the hardest part of single-particle cryoEM, because each protein is different, details matter, and there are no shortcuts.”*<sup>242</sup>

This mismatch between the importance of membrane proteins but limited number of solved structures is due to several difficulties during all stages of the structural determination pipeline. Membrane proteins are still a difficult target when it comes to their expression and purification.

Membrane proteins are generally expressed at much lower yields than soluble proteins. In addition, eukaryotic membrane proteins cannot be expressed routinely in easy-to-use bacterial expression systems; successful membrane protein production often necessitates protein expression in eukaryotic cell lines. Advances have been made in this regard with a routine workflow to routinely produce sufficient amounts of eukaryotic membrane proteins using transient transfection with expression plasmids<sup>243</sup> or transduction with recombinant baculoviruses<sup>189</sup> of mammalian cells in suspension culture. After overexpression, they must be extracted from the membrane to allow for their purification. Typically, detergents are used to efficiently solubilize the membrane protein of interest and to keep it in a soluble state when handled in aqueous buffers during purification. However, detergents generally reduce the contrast of cryo-EM images, which can be more problematic for small membrane proteins. In addition, some enzymes or transporters require stabilization outside of their native environment other than detergents. Such approaches include the use of Amphipols<sup>244</sup>, Saposins<sup>228</sup>, styrene maleic-acid lipid particles (SMALPs)<sup>245</sup>, nanodiscs<sup>246</sup>, Peptidisc<sup>247</sup> or even proteoliposomes<sup>248</sup>. There are a plethora of structures determined using these alternative approaches but the most successful approach in determined cryo-EM structures is still the use of DDM as detergent.<sup>249</sup> The major challenge for successful structure determination is still the size and stability of the target. The detergent micelle is heterogeneous in composition and varies from particle to particle. It masks the signal of the TM helices, making them difficult to align. Most membrane proteins below 100 kDa lack domains outside the detergent micelle, which typically alleviate the problem of particle alignment. There are examples of sub-100kDa proteins which were solved without bearing features outside of the micelle such as the high-resolution cryo-EM structure of the human glucose transporter 4 (GLUT4).<sup>250</sup> This required rigorous testing of different detergents, ideal sample preparation, and a stable and rigid conformation of the protein. In general, however, these are still the exceptions, and most small membrane proteins cannot be resolved without features outside the micelle. To overcome these inherent problems of small membrane proteins, several approaches and strategies have been developed to provide an additional ordered mass outside the detergent micelle that can be used for particle alignment. For example, the attractive pharmacological targets of GPCRs naturally interact with their cognate G proteins in initiating downstream signaling cascades of which a first structure of a GPCR-G protein complex was solved in 2017 using cryo-EM SPA.<sup>251</sup> Engineering that G protein and generating an alpaca derived nanobody, Nb35, facilitated the structural determination of a plethora of GPCR-G protein structures. Moreover, GPCR structures were solved bound to various ligands, binding partners and in different conformational states proving

cryo-EM as the method of choice nowadays for their structural determination.<sup>252–255</sup> Other classes of proteins such as the human proton-coupled monocarboxylate transporter 1 (MCT1) are on their own to small and featureless targets for structural determination through cryo-EM. But they naturally form complexes with auxiliary proteins, here Basigin-2 or Embigin, that themselves bear extramembranous domains which facilitated the high-resolution structural determination of the complexes.<sup>256,257</sup>

However, there are many membrane proteins for which this approach of forming a complex between interaction partners is difficult to establish or does not exist. In this context, several strategies have been established to add an additional but ordered mass outside the micelle. In X-ray protein crystallography, antigen-binding fragments of antibodies (Fabs) have been successfully used to promote a sufficient handle for crystal formation. These Fab fragments have also been successfully transferred to cryo-EM, where they provide a rigid fiducial for structural determination of small membrane proteins.<sup>235</sup> Much smaller, typically 12-15 kDa in size, but equally suited are camelid derived single-domain antibody fragments, called nanobodies.<sup>258</sup> However, they can bind much better to cryptic epitopes due to their high variability of the complementarity-determining region 3 (CDR3).<sup>259</sup> Strategies have been developed to further artificially increase the size of nanobodies by grafting other rigid molecules onto them such as megabodies<sup>260</sup>, legobodies<sup>261</sup> or a synthetic Fab fragment that can bind to nanobodies, termed NabFabs<sup>262</sup>. However, all require animal immunization and intensive screening to identify stable binders. Synthetic libraries containing Fab fragments<sup>263</sup> and artificial nanobodies, known as sybodies<sup>264</sup>, have been developed to bypass animal immunization. However, they still require the production of the target protein to select potential binders. A more universal approach has been introduced by grafting a BRIL fragment into the target protein by replacing loop regions.<sup>84</sup> Fab fragments that specifically bind to BRIL fragment are available and provide a handle for particle alignment. This approach was successful used to determine the structure of the GPCR Frizzled 5.<sup>236</sup> In this regard, GPCR are well known targets and the site of insertion has been extensively characterized in the past. For new protein targets of initially unknown fold are more difficult to engineer and this strategy needs to be further tested in the future.

In conclusion, there has been a tremendous development of approaches and strategies for the determination of small membrane proteins by cryo-EM SPA in the last decade. These advances include optimized expression, stabilization of target proteins after extraction from the membrane, and the use of fiducials to guide particle alignment. All of these approaches have lowered the size limit for cryo-EM analysis. As a result and combined with continued hardware

and software developments in the field of cryo-EM, it is expected that a variety of small membrane protein structures representing pharmacologically relevant targets will be routinely determined by cryo-EM in the future. In particular, the visualization of bound small molecules will be critical for the optimization of the small-molecule of interest using structure-based drug design.

## 9.4 Structure of a bacterial Rhs effector exported by the type VI secretion system

The results of this section were originally published in *Plos pathogens* in 2022:

Patrick Günther\*, Dennis Quentin\*, Shehryar Ahmad, Kartik Sachar, Christos Gatsogiannis, John C. Whitney, und Stefan Raunser. “Structure of a bacterial Rhs effector exported by the type VI secretion system”. *PLoS Pathogens* 18, Nr. 1 (2022): 1–25.

Online version: <https://doi.org/10.1371/journal.ppat.1010182>.

The original publication has been reformed in order to be adapted in this thesis. The introduction of the original publication was not reprinted in this thesis. Instead, a broader introduction was included in the introduction part of this thesis. One paragraph from the introduction of the original publication has been moved into the results section of this thesis describing the previous discovery of the operon encoding RhsA in *Pseudomonas protogens*. Moreover, Figure 1 of the original publication has been moved into the results section of this thesis for ease of reading. All methods regarding the results of this section have been reprinted in section 8.5 of this thesis. Work regarding biochemical methods has been performed by Shehryar Ahmad, Kartik Sachar and John C. Whitney. Moreover, the protein for cryo-EM analysis has been kindly provided by Shehryar Ahmad, Kartik Sachar and John C. Whitney. The cryo-EM analysis has been performed with help from Dennis Quentin and Christos Gatsogiannis. Supplementary figures and tables have been moved to the Appendix section of this thesis. Supplementary videos and mass spectrometry datasets are not included in this thesis but can be downloaded from the online version of the publication.

Parts of the original publication including text and figures are reused in this thesis under the CC-BY 4.0 license<sup>265</sup>. Section 9.4.5 has been slightly rephrased and extended to discuss the SGS1 structure. Moreover, Figure 48 has been adapted and modified to also include the structure of SGS1.

### 9.4.1 Abstract

The type VI secretion system (T6SS) is a widespread protein export apparatus found in Gram-negative bacteria. The majority of T6SSs deliver toxic effector proteins into competitor bacteria. Yet, the structure, function, and activation of many of these effectors remains poorly understood. Here, we present the structures of the T6SS effector RhsA from *Pseudomonas protegens* and its cognate T6SS spike protein, VgrG1, at 3.3 Å resolution. The structures reveal that the rearrangement hotspot (Rhs) repeats of RhsA assemble into a closed anticlockwise  $\beta$ -barrel spiral similar to that found in bacterial insecticidal Tc toxins and in metazoan teneurin proteins. We find that the C-terminal toxin domain of RhsA is autoproteolytically cleaved but remains inside the Rhs ‘cocoon’ where, with the exception of three ordered structural elements, most of the toxin is disordered. The N-terminal ‘plug’ domain is unique to T6SS Rhs proteins and resembles a champagne cork that seals the Rhs cocoon at one end while also mediating interactions with VgrG1. Interestingly, this domain is also autoproteolytically cleaved inside the cocoon but remains associated with it. We propose that mechanical force is required to remove the cleaved part of the plug, resulting in the release of the toxin domain as it is delivered into a susceptible bacterial cell by the T6SS.

### 9.4.2 RhsA forms a cocoon-like structure that undergoes N- and C-terminal autocleavage

Previously, our group showed that the T6SS of the soil bacterium *Pseudomonas protegens* secretes the prePAAR and PAAR-containing effector RhsA<sup>150</sup>. The *rhsA* operon contains genes encoding its cognate VgrG spike and Eag chaperone proteins (Figure 45A), both of which are required for the delivery of this effector into susceptible competitors<sup>150</sup>. Previous negative stain electron microscopy experiments suggest that EagR1, RhsA and VgrG assemble to form a complex necessary for T6SS function (Figure 45B)<sup>150</sup>. The N-terminus of RhsA consists of a prePAAR motif as well as two predicted transmembrane helices within its TMD. The prePAAR motif is proposed to contribute sequence elements that ‘complete’ the PAAR domain and enable its interaction with the tip of VgrG1<sup>150</sup> whereas the transmembrane helices likely play a role in target cell penetration<sup>266</sup>.

We previously demonstrated that a fragment of RhsA lacking its N-terminal prePAAR motif and TMD (Figure 45C, residues 74-1486), RhsA $\Delta$ TMD, is stable in the absence of its EagR1 chaperone and can be purified to homogeneity as a soluble protein when over-expressed in *Escherichia coli*<sup>150</sup>. We therefore expressed and purified RhsA $\Delta$ TMD (Figure S 8) and used it for cryo-EM and single particle analysis. The raw images and 2D class averages suggest that

RhsA $_{\Delta TMD}$  forms a stable dimer in solution, which facilitated structure determination due to the increased size of the particles. Subsequent image processing imposing  $C_2$  symmetry resulted in a reconstruction with a resolution of 3.3 Å from 454,740 particles (Figure 45D, Table 4). The high quality of the map allowed us to build an atomic model of 75% of the protein, comprising residues 268-275, 289-302, 305-366, 371-386, 395-871, 891-1039 and 1051-1350 (Figure 45E). The 25% of remaining unresolved regions correspond to the PAAR domain (residues 75-267) and the toxin domain (residues 1351-1486), indicating that these regions likely exhibit a high degree of flexibility.

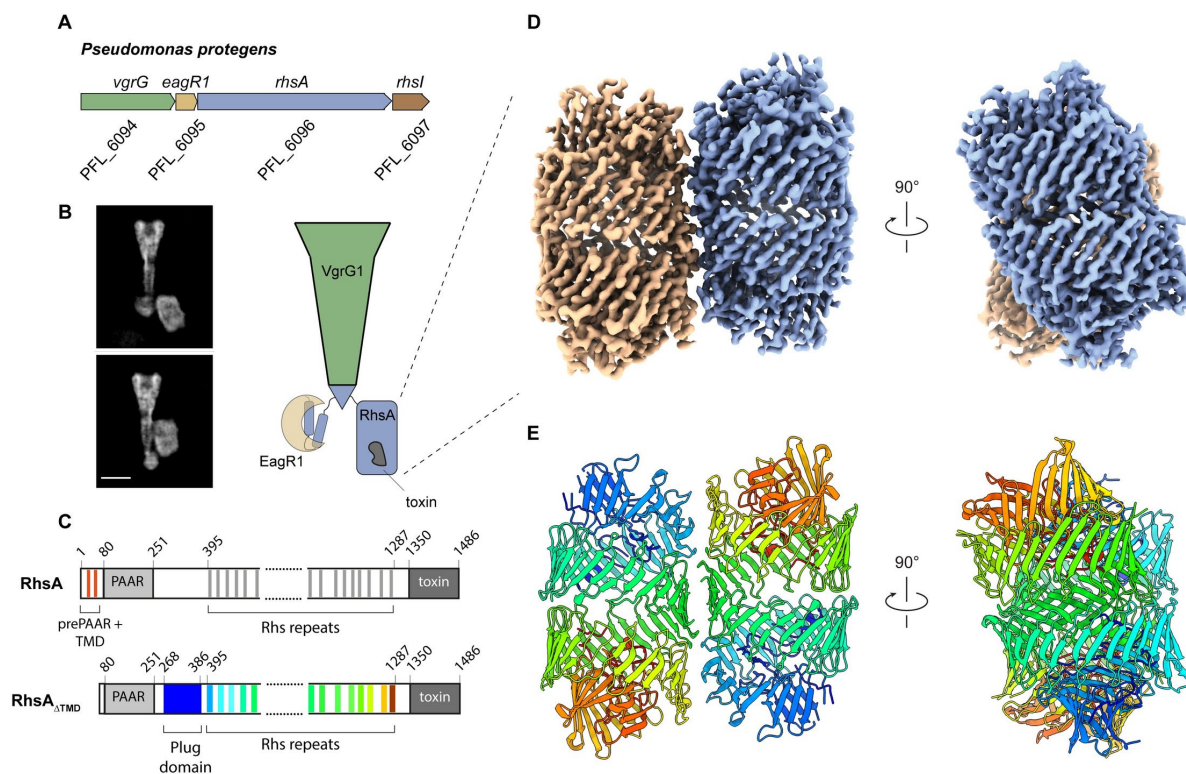


Figure 45: High-resolution structure of the T6SS effector RhsA.

(A) Genomic context of *rhsA* (PFL\_6096, blue) in *Pseudomonas protegens* Pf-5. Upstream genes encoding the cognate VgrG protein, *vgrG1* (PFL\_6094, green) and RhsA-specific chaperone, *eagR1* (PFL\_6095, light brown) are shown. Self-protection against RhsA is conferred via expression of the downstream immunity-encoding gene, *rhsI* (dark brown). (B) Representative cryo-EM 2-D class averages of the assembled pre-firing complex composed of VgrG1, RhsA and EagR1 (left) and a schematic representation of each of the components that comprise this complex (right). Scale bar, 10 nm. (C) Full-length RhsA contains a prePAAR motif and a TMD comprising two transmembrane helices upstream of its PAAR domain. The RhsA $_{\Delta TMD}$  truncation of RhsA was used in this study to determine the high-resolution structure of the RhsA cocoon. (D) Cryo-EM density of RhsA $_{\Delta TMD}$  displayed perpendicular to the central symmetry axis of the barrel and rotated 90° clockwise (map postprocessed with DeepEMhancer). (E) Cartoon representation of the atomic model of RhsA colored in rainbow from N-terminus (blue) to C-terminus (red). This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.



Table 4: Statistics of cryo-EM data collection, image processing and model validation.

	<b>VgrG1</b>	<b>RhsA<sub>ΔTMD</sub></b>
<b>Data collection</b>		
Microscope	Titan Krios (X-FEG, Cs-corrected)	Titan Krios (X-FEG, Cs 2.7 mm)
Magnification	59,000	105,000
Voltage (kV)	300	300
Defocus range (μm)	-1.2 to -2.2	-0.8 to -2.2
Camera	F3 linear	K3 counting
Pixel size (Å/pixel)	1.1	0.91
Total electron dose (e <sup>-</sup> /Å <sup>2</sup> )	90	62
Exposure time (s)	1.5	2
Number of images	1250	13,090
<b>Refinement</b>		
Number of final particles	423,980	454,740
Final resolution (Å)	3.3	3.3
Symmetry	C <sub>3</sub>	C <sub>2</sub>
Map sharpening B factor (Å <sup>2</sup> )	-174.5	DeepEMhancer
<b>Model composition</b>		
Non-hydrogen atoms	14868	17070
Protein residues	1881	2094
RMSD bond	0.007	0.006
RMSD angles	0.722	0.570
Model-to-map fit, CC Mask	0.77	0.77
<b>Validation</b>		
MolProbity	1.85	1.42
Clashscore	6.86	4.15
EMRinger score	2.77	4.70
Poor rotamers (%)	0.00	0.45
<b>Ramachandran</b>		
Favored (%)	92.46	96.59
Allowed (%)	7.22	3.31
Outliers (%)	0.32	0.10

The two molecules in the RhsA $\Delta$ TMD dimer interact along their longitudinal axes and are slightly tilted with respect to one another. The dimer interface is comprised of several complementary hydrophilic surfaces indicating that it is mostly electrostatic in nature (Figure S 9A-B). The 76 antiparallel  $\beta$ -strands of RhsA $\Delta$ TMD spiral in an anticlockwise manner resulting in a large hollow cocoon-shaped structure with outer dimensions of 86 x 65 Å. The overall structure of the Rhs cocoon resembles that of other YD-repeat containing proteins, such as the BC component of Tc toxins<sup>267,268</sup> and human teneurin2<sup>269</sup>. We also identified the conserved catalytic center of an aspartyl autoprotease that was first identified in Tc toxins<sup>267,268</sup>. In line with functioning to self-cleave the C-terminal toxin domain, we observed that the density of the cryo-EM map stops abruptly after tryptophan 1350 (Figure 46A). This position is in agreement with cleavage sites in other Rhs-related toxins that share the same PxxxxDPxG W/L/F consensus sequence found in RhsA (Figure S 10A), indicating that the C-terminal toxin domain is likely autoproteolytically cleaved similar to the HVR found in Tc toxins. To experimentally test the proteolytic activity of this motif, we mutated either aspartate 1324 or aspartate 1346, which are the RhsA residues in the equivalent position to the catalytic dyad of Tc toxin aspartyl autoproteases, to asparagine. Consistent with their proposed role in autocleavage and toxin release, these RhsA variants did not undergo autoproteolysis at their C-terminus (Figure 46B) and were substantially less toxic when overexpressed in *E. coli* compared to the wild-type protein (Figure S 10B). We next introduced these mutations into the chromosome of *P. protegens* to test whether C-terminal autoproteolytic cleavage is indeed required for interbacterial competition. In contrast to a strain expressing wild-type RhsA, strains expressing the D1324N or D1346N variants of the protein were unable to outcompete RhsA-sensitive recipient bacteria indicating that C-terminal autocleavage is required for T6SS-dependent killing by this effector (Figure 46C). These findings mirror what has been observed for T6SS-exported Rhs proteins in *Aeromonas dhakensis* and *Enterobacter cloacae* suggesting that toxin domain liberation is a universal property of Rhs effectors<sup>270,271</sup>.

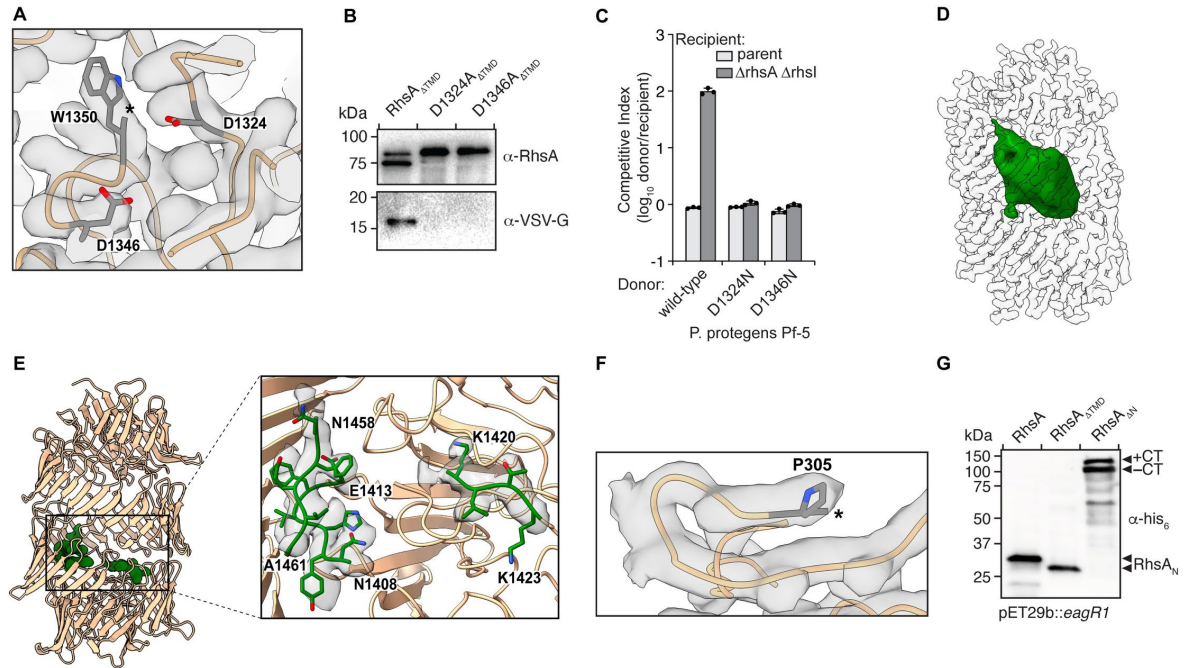


Figure 46: Autoproteolysis of RhsA occurs at its N- and C-terminus.

(A) RhsA is autoproteolytically cleaved at its C-terminus at position W1350. The end of the connected density is indicated with an asterisk. Catalytic aspartates D1324 and D1346, are shown in stick representation. (B) Western blot analysis shows proteolytic cleavage of the C-terminal toxin domain. Mutation of either D1324 or D1346 to asparagine prevents autoproteolytic cleavage of the C-terminal RhsA toxin domain. Blots were performed against the Rhs barrel ( $\alpha$ -RhsA) and against a C-terminal VSV-G epitope tag ( $\alpha$ -VSV-G). (C) Outcome of intraspecific growth competition assays between the indicated *P. protegens* donor and recipient strains. Donor strains were competed against recipient strains that either contain (light grey) or lack the *rhsA-rhsI* effector-immunity pair (dark grey). The recipient strains also lack *pppA* to stimulate type VI secretion in donors<sup>152</sup>. Data are mean  $\pm$  s.d. for  $n = 3$  biological replicates and are representative of two independent experiments. (D) The toxin domain of RhsA is encapsulated by its cocoon-shaped Rhs repeat-containing domain. Difference map of the encapsulated toxin at extremely high-density threshold and low pass-filtered to 20 Å (green) is shown. The Rhs cocoon is depicted using a transparent space-filling representation at normal map threshold. (E) Regions of the toxin domain are stabilized inside the cocoon through interactions with the C-terminal autoprotease domain. The densities appear at the same density threshold as the rest of the map. The atomic models are shown in stick representation. (F) RhsA undergoes N-terminal cleavage at proline 305. The end of the density is indicated with an asterisk. (G) N-terminal cleavage occurs in RhsA and a mutant lacking the prePAAR-TMD (RhsA $_{\Delta$ TMD}, residues 74-CT) but not in a mutant lacking the entire N-terminal region (RhsA $_{\Delta$ N}, residues 297-CT). The indicated RhsA constructs were purified from *E. coli* and subject to Western blot and detected using an N-terminal His<sub>6</sub>-tag antibody ( $\alpha$ -his<sub>6</sub>). This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

We also identified additional density inside the Rhs cocoon corresponding to the C-terminal toxin domain, which is predicted to function as a DNase based on its sequence similarity to several characterized endonucleases and homology to other characterized Rhs effectors<sup>272</sup>. The density filling the cocoon is, for the most part, not well defined and is apparent only at a much lower density threshold compared to the rest of the map (Figure 46D). Nevertheless, we could build three  $\beta$ -strands of this domain, comprising residues 1408-1413, 1420-1423 and 1458-1461, respectively, all of which interact with the inner surface of the cocoon structure (Figure 46E). The interface between the  $\beta$ -strands and the cocoon is stabilized by both hydrophilic and hydrophobic interactions (Figure 46E, Figure S 11). In sum, these structural data indicate that most of the toxin domain is either only partially folded or very flexible.

Interestingly, the density inside the Rhs cocoon not only ends after tryptophan 1350 but also before proline 305 (Figure 46F), suggesting that RhsA is proteolytically cleaved at this position as well. However, the N-terminal region is presumably still associated with the cocoon structure because the protein was purified by affinity chromatography using an N-terminal His-tag. Therefore, we heat-denatured RhsA<sub>WT</sub> and RhsA $\Delta$ TMD prior to SDS-PAGE to determine if the N-terminal domain was indeed cleaved. Consistent with our structural data, we observed bands that migrate at molecular weights consistent with the loss of  $\sim$ 304 and  $\sim$ 230 amino acids, respectively, corresponding to the 304 N-terminal residues of RhsA<sub>WT</sub> and 230 N-terminal residues of RhsA $\Delta$ TMD (Figure 46G). To unambiguously determine the position of the cleavage site we performed *de novo* protein sequencing via LC-MS/MS. We found that the N-terminal peptide of the Rhs-cocoon detected by mass spectrometry exactly matches the proposed cleavage site at proline 305 (Figure S 12A). In addition, we further validated cleavage at this position by generating a non-cleavable version of RhsA in which both residues H304 and P305 were mutated to alanine. This RhsA variant was unable to undergo autoproteolysis at its N-terminus (Figure S 12B). Collectively, these data demonstrate that the protein is indeed cleaved between residues 304 and 305. By contrast, a truncation of RhsA lacking this entire N-terminal region, RhsA $\Delta$ N, did not undergo N-terminal proteolysis (Figure 46G). A similar observation was recently made for the T6SS effector TseI from *A. dhakensis*, which is also N-terminally cleaved even though it differs from RhsA at its N-terminus in that it lacks prePAAR, PAAR and a TMD<sup>270</sup>. This cleavage event was shown to be essential for the activity of TseI after its secretion by the T6SS. Two glutamate residues at the +7 and +8 position relative to the N-terminal cleavage site are responsible for the autoproteolysis of this effector<sup>270</sup>. However, these glutamates are not conserved in RhsA and are replaced by alanine (A312) and lysine (K313), which cannot act as catalytic center (Figure S 12C-E). Therefore, we examined the direct

vicinity of P305 in our structure and found a cysteine residue (C538) flanked by two histidines (H530 and H555) that protrude from the wall of the Rhs barrel and are located near proline 305 (Figure S 12F). Unlike the previously described glutamate residues, this site is conserved in class I prePAAR T6SS Rhs effectors, but does not exist in Tc toxins (Figure S 12G)<sup>270</sup>. In addition, we identified three consecutive aspartates (D318-D320) in close proximity to the cleavage site.

To test the hypothesis of whether cysteine 538 or the other proximal residues are involved in N-terminal cleavage, we mutated these residues individually and performed a Western blot analysis against the N-terminus of the overexpressed constructs (Figure S 12H). Contrary to our assumption, we found that mutation of C538 to alanine did not abolish the N-terminal cleavage. Mutations of other conserved residues near the cleavage site had also little effect. Only the triple mutant (D318-D320 to alanine) led to a minimal reduction of N-terminal cleavage. However, we believe that this could be an indirect effect induced by the mutationally altered conformation in this region similar to the construct where we mutated H530 to alanine. Since there are no other obvious potential catalytically active residues besides the tested ones, we speculated that the encapsulated C-terminal toxin domain could be involved in catalyzing the N-terminal cleavage reaction. To test this, we generated a truncated form of RhsA lacking its C-terminal toxin domain (RhsA $\Delta$ tox) and examined N-terminal cleavage via Western blot. This analysis clearly showed that RhsA $\Delta$ tox is still able to undergo autoproteolysis at the N-terminus (Figure S 12I). Therefore, the N-terminal cleavage must occur via an as yet unknown mechanism.

#### 9.4.4 RhsA possesses a unique plug domain at its N-terminus

Our structural data show that the C-terminal aspartyl protease domain of RhsA seals the cocoon-shaped structure at one end, while the other end is capped by an N-terminal domain that adopts a smaller structure. This N-terminal domain of the barrel, formed by residues 268-386, caps the cocoon structure in a manner that is reminiscent of how a cork is used to plug a champagne bottle and thus we refer to it as the N-terminal plug domain (RhsA $\Delta$ plug) (Figure 47A-E). The interface between the plug domain and the Rhs repeats is mainly stabilized by hydrophobic interactions (Figure 47A-C) and a few hydrophilic interactions (Figure S 13A). In addition to the cork structure (residues 305-386), the plug domain contains an anchor helix (residues 289-302) and an N-terminal seal peptide (residues 268-275). The seal peptide fills a small opening in the cork as it leaves the cocoon and connects the plug domain to the unmodelled PAAR domain (Figure 47D-E, Figure S 14A-B). In doing so, the seal peptide,

together with the cork, closes off the cocoon entirely (Figure S 14B). The anchor helix is amphipathic and stabilized by interactions with a hydrophobic patch of the Rhs repeats in this region ( Figure 47E-F, Figure S 14C-E). The position of the anchor helix in our structure suggests its function is twofold. On the one hand it holds the N-terminal cleavage site in place. On the other hand, it ensures that the N-terminal plug domain remains stably attached to the cocoon, so that it remains sealed even after N-terminal cleavage.

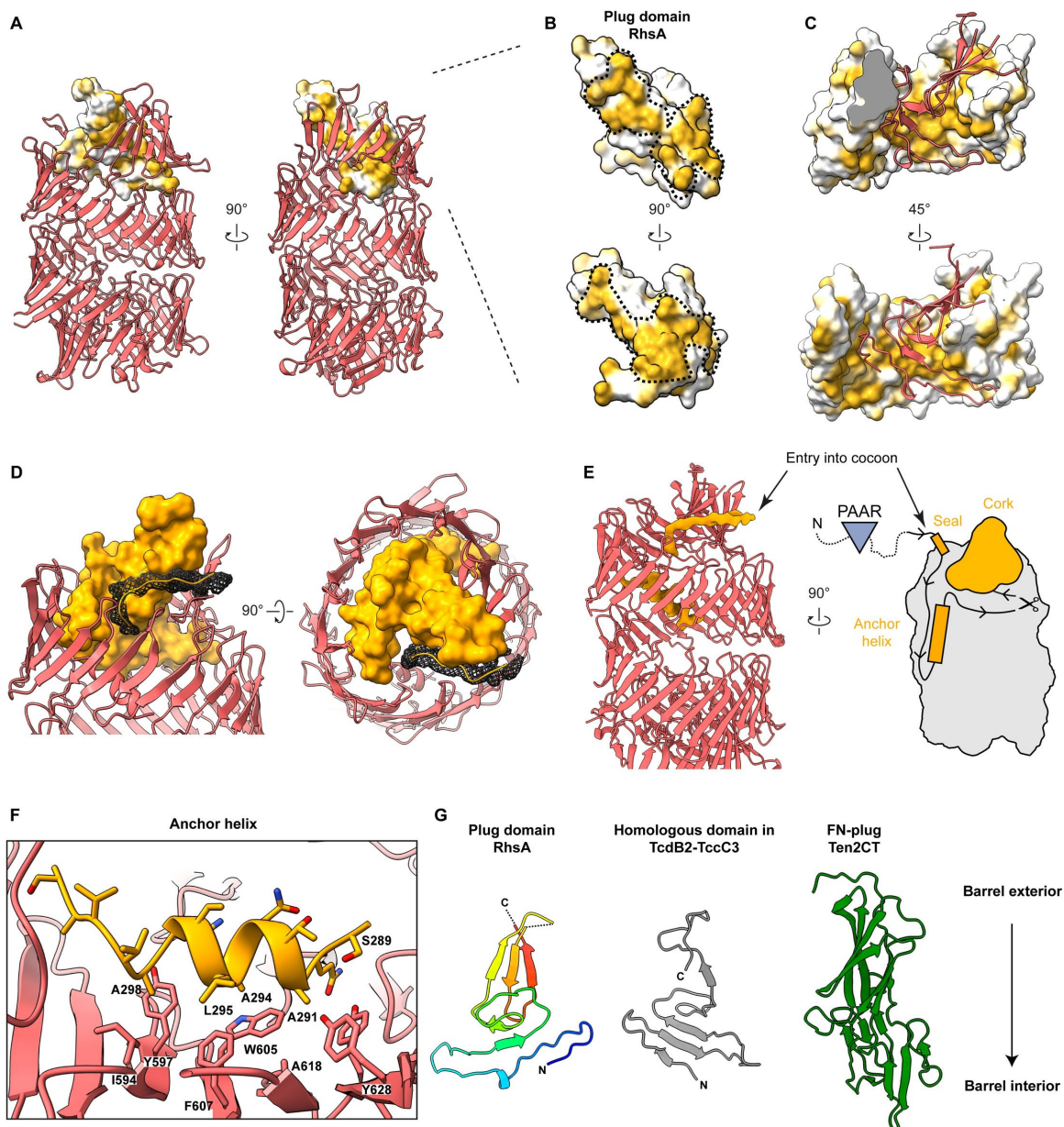


Figure 47: A unique plug domain seals the Rhs barrel of RhsA.

(A) Surface representation of the cork domain of RhsA. The molecular surface is colored according to hydrophobicity where ochre and white indicate hydrophobic and hydrophilic regions, respectively. The Rhs barrel is shown as a cartoon. (B) Enlarged view of the cork domain of RhsA. The hydrophobic surface spirals around the domain as indicated by the black dashed line. (C) The upper Rhs repeats of RhsA possess complementary hydrophobic patches to those found on its plug domain (shown in cartoon representation, red). (D) The cocoon

structure of RhsA is closed off by an N-terminal plug comprised of a ‘cork’ domain (orange, density representation), a ‘seal’ peptide (mesh), and an anchor helix (E, F). The seal and the cork density together form a cap structure. (E) The seal peptide, which also functions as the linker to the unmodelled PAAR domain, not only complements the shape of the cork but is also the entry point of the N-terminal part of the protein into the inside of the cocoon. The cocoon remains stably bound to the cleaved N-terminal region, including the PAAR domain, due to the anchor helix inside the cocoon. (F) The anchor helix of RhsA is stabilized by hydrophobic interactions with the inner wall of the Rhs repeats. (G) Cartoon representation of the cork region of the RhsA plug domain and a comparison with the homologous N-terminal plug domain of *Phototribadus luminescens* TcdB2 (grey, PDB ID: 6H6G) and the unique FN-plug domain of human teneurin2 (green, PDB ID: 6FB3). The lower part of each depicted plug domain inserts into each of their respective Rhs barrels. The cork domain of RhsA is colored in rainbow. This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

The plug domain of RhsA possesses sequence and structural similarity to a domain found in the BC components of TcdB2-TccC3<sup>267</sup> and YenBC<sup>268</sup> ( Figure 47G, Figure S 15A-C) (18% sequence identity to TcdB2; 22% sequence identity to YenB). Interestingly, in Tc toxins this domain does not act as a plug that prevents the release of the toxin, but instead forms a negatively charged constriction through which the toxin domain is threaded prior to its translocation into a target cell<sup>273</sup>. A plug domain has also been described in teneurins (called fibronectin-plug, FN-plug)<sup>269,274</sup>, however, it does not bear sequence or structural similarity to the plug domain of RhsA described herein. The FN-plug is narrower than the plug domains of RhsA and Tc toxins and extends further into the Rhs cocoon structure ( Figure 47G) making numerous hydrophilic interactions with residues lining the inside of the YD-shell<sup>269</sup>.

Our structure suggests that in contrast to Tc toxins the plug domain of RhsA tightly seals its Rhs cocoon. Because the plug domain appears to strongly interact with the Rhs barrel, mechanical removal of this entire domain after translocation of the cocoon into the target cell cytosol seems unlikely. Instead, we propose that only the N-terminal peptide seal of the RhsA plug is pulled out during T6SS-dependent delivery of RhsA into a susceptible bacterial cell. The seal peptide is part of the N-terminal region that is proteolytically cleaved, so we speculate that it would be more easily removed compared to the entire plug domain, which is held in place by the anchor helix that exists downstream of the N-terminal cleavage site. Penetration of the outer membrane and peptidoglycan layer as well as translocation of RhsA through the target cell inner membrane could provide the mechanical force to remove the seal peptide, creating a channel through which the unfolded toxin domain could exit. Since the plug domain resembles the constriction site in Tc toxins that effectors must pass during the initial translocation process, the same process is conceivable for the release of the C-terminal toxin domain of RhsA.

#### 9.4.5 Comparison between Rhs proteins of known structure

A common feature of all Rhs proteins that have been structurally characterized to date is the cocoon structure formed by the Rhs repeats. While the Rhs cocoons from RhsA and teneurin2 are comprised of three  $\beta$ -helical turns of Rhs repeats, the cocoons of Tc toxins and SGS1 have four turns and therefore have bigger overall dimensions and an internal cavity with larger volume. Consistent with this observation, the effector domains inside Tc toxin cocoons and SGS1 cocoon (~30 kDa) are larger than the toxin domain of RhsA (~15 kDa).

All characterized Rhs proteins contain a conserved C-terminal region. In Tc toxins, SGS1 and RhsA, this domain functions as an aspartyl autoprotease, whereas in teneurins it functions as a YD shell plug that is not cleaved (Figure 48 blue). Interestingly, this autoproteolytic event in SGS1 is only detected when the protein is transported from the salivary gland into the saliva. It is thought that salivary transport activates the aspartyl protease and allows the cleavage of the C-terminal transmembrane helices.<sup>173</sup> What is conserved in all cases, however, is the architecture of the C-terminal domain in which the last 14 Rhs repeats spiral into the interior of the cocoon and seal it at one end.

The plug domains that close off the N-terminal end of the cocoon structures are more variable than their C-terminal counterparts (Figure 48 orange). While the barrel of teneurins is sealed with an FN-plug that is mainly held in place by hydrophilic interactions with the interior of the barrel, TcdB2 from *Photorhabdus luminescens* and other BC components of Tc toxins close the cocoon using a  $\beta$ -propeller domain that acts as gatekeeper for toxin release. In SGS1, two  $\beta$ -propellers are located on the outside of the cocoon, with  $\beta$ -propeller 2 having a function similar to that of Tc toxins. In Tc toxins, the  $\beta$ -propeller is in a distorted conformation that refolds to a canonical arrangement upon binding to the TcA subunit. This in turn causes the propeller to open, allowing the toxin to thread into the translocation channel of TcA. The SGS1 propeller has nine blades, which are also in a distorted conformation. Therefore, a similar rearrangement of the propeller upon binding to the target receptor is conceivable for SGS1. RhsA uses an N-terminal plug domain, which is homologous to the domain that forms the constriction site in Tc toxins. Both TcBC and RhsA encapsule a toxic effector, whereas teneurins act as scaffolding proteins with an empty cocoon (Figure 48red). Nevertheless, teneurins encode a C-terminal ancient toxin component that sits outside of the barrel and is inactive (ABD Tox-GHH, Figure 48 red). While RhsA contains two autoprotease sites, namely an aspartyl protease that is responsible for the cleavage of the C-terminal toxic effector and an unknown protease site that



cleaves the N-terminal 304 residues of the protein, Tc toxins and SGS1 only contain the former site and teneurins none at all.

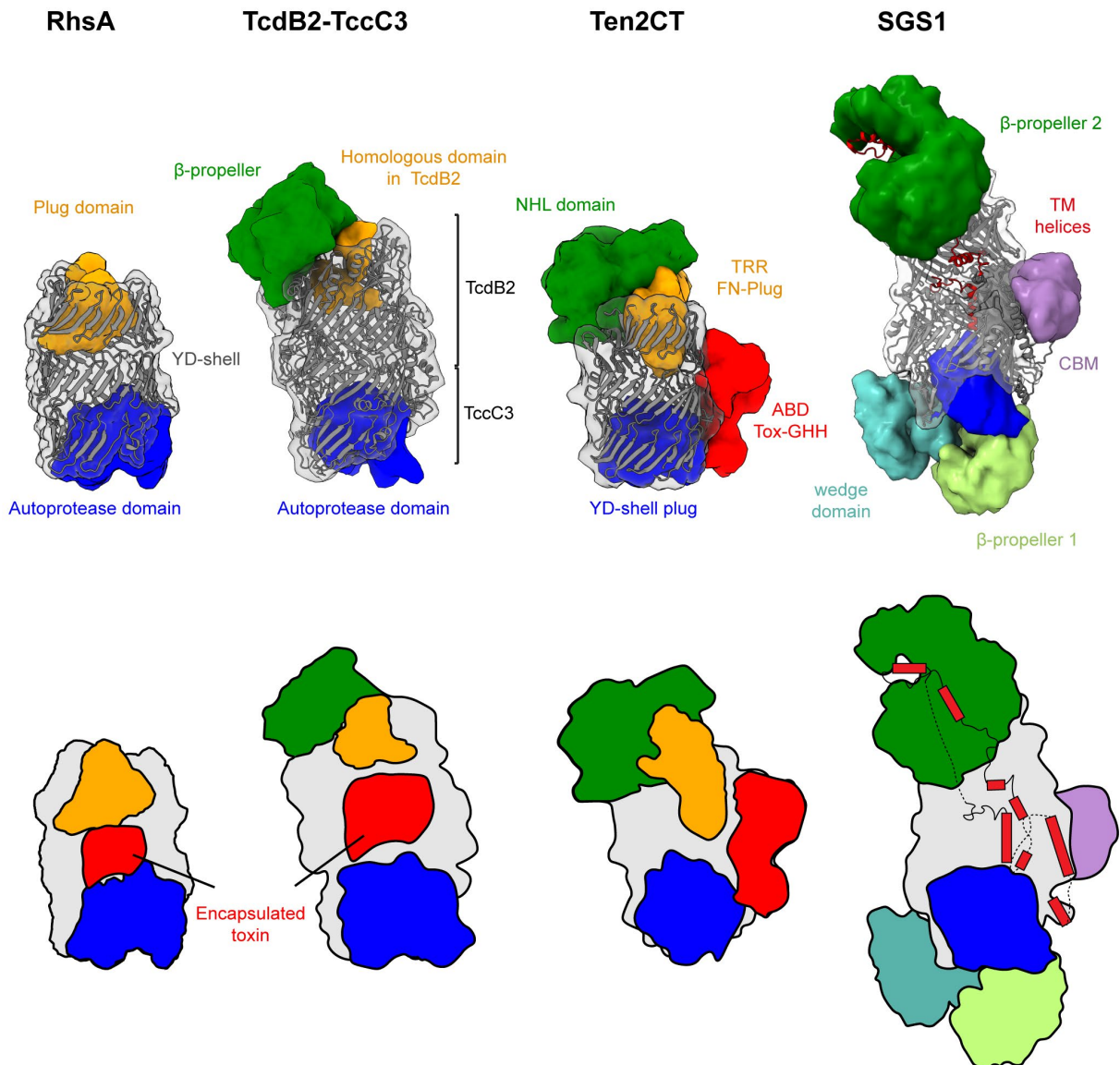


Figure 48: Structural comparison of Rhs repeat containing proteins.

The C-terminal autoproteolysis domain or YD-shell plug is structurally conserved among diverse Rhs proteins (blue). RhsA, BC components of Tc toxin complexes and SGS1 encapsulate their effector domains (red) whereas in teneurin proteins the toxin domain is appended to the outside of the Rhs barrel (red). A distinguishing feature of these Rhs proteins is the unique N-terminal plug domain for each protein family (orange) even absent in SGS1. RhsA is capped by a cork-like plug domain that seals the Rhs barrel (orange). In Tc toxins, the homologous plug domain acts as constriction site and the cocoon is sealed off by a  $\beta$ -propeller domain (green). The  $\beta$ -propeller 2 seals the cocoon of SGS1. Teneurin proteins are capped with a non-homologous FN-plug (orange) that is stabilized by an NHL domain (green). SGS1 contains additional domains outside of the cocoon potentially acting as receptor binding domains (CBM: carbohydrate-binding module). The lower row shows schematic representations of the domain organizations. This figure has been modified and extended from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

In conclusion, based on its unique plug domain and its fusion to an N-terminal PAAR domain, we propose that our RhsA structure represents the founding member of a fourth structural class of Rhs repeat containing proteins (Figure 48).

#### 9.4.6 Architecture of the pre-firing complex

To investigate how RhsA is mounted onto VgrG1 we set out to determine the structure of the secretion competent pre-firing complex (PFC) comprising VgrG1 in complex with full-length RhsA and EagR1. We purified the complex as described previously<sup>150</sup> and examined it by single particle cryo-EM. The 2D class averages enabled us to unequivocally characterize the arrangement of the subunits within the complex (Figure 49A). Expectedly, both EagR1 and RhsA are located at the tip of VgrG1. Interestingly, a large fraction of PFCs contained RhsA dimers instead of monomers (Figure S 16), similar to what was observed in our analysis of RhsA $\Delta$ TMD alone. This demonstrates that the loading of two RhsA molecules onto one VgrG1 is sterically possible although it is unlikely to occur *in vivo* given that current data indicates a single VgrG homotrimer caps the Hcp tube<sup>275</sup>.

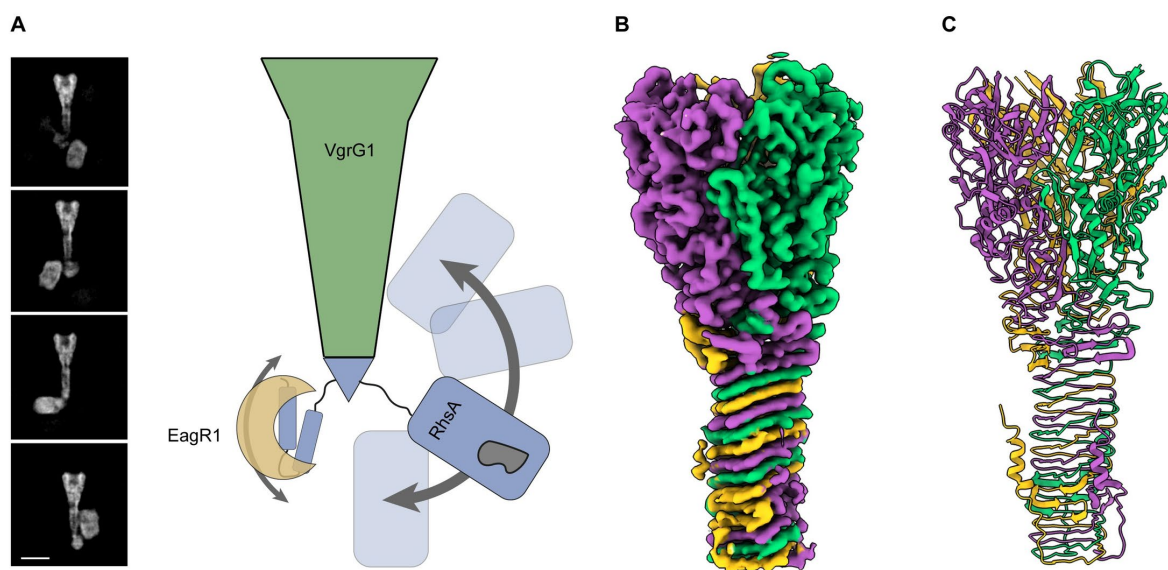


Figure 49: High-resolution structure of *P. protegens* VgrG1.

(A) The RhsA barrel of the assembled pre-firing complex (PFC) displays high positional flexibility relative to VgrG1. Representative cryo-EM 2-D class averages depicting flexibility are shown. Scale bar, 10 nm. (B) Cryo-EM density and (C) ribbon representation of the molecular model of the *P. protegens* VgrG1 trimer viewed along the long axis of the protein. Each protomer is colored differently to highlight their positions within the homotrimeric VgrG1 spike. This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

As is the case for the previously characterized VgrG1-EagT6-Tse6 complex from *P. aeruginosa*<sup>266</sup>, T6SS effectors can adopt multiple positions relative to their cognate VgrG spike

protein. Unfortunately, this conformational heterogeneity prevented us from obtaining a high-resolution 3D reconstruction of the entire complex. Instead, we applied  $C_3$  symmetry during processing to determine the three-dimensional structure of *P. protegens* VgrG1 (Figure 49B). In applying this symmetry operator, the RhsA and EagR1 components of the complex, which do not adopt this symmetry, are averaged out during image processing. The cryo-EM map of VgrG1 reached a resolution of 3.3 Å and allowed us to build almost the complete atomic model of the protein comprising residues 8-643 (Figure 49C, Table 4). As expected, given its high sequence homology, the structure of VgrG1 from *P. protegens* is nearly identical (71% sequence identity, 81% sequence similarity) to our previously determined VgrG1 structure from *P. aeruginosa* (r.m.s.d of 1.035 between 544 pruned C $\alpha$  atoms; r.m.s.d of 1.428 across the complete structure, Figure S 16B-C)<sup>266</sup> even though their respective effectors, RhsA and Tse6, bear no sequence or structural similarity to one another beyond their N-terminal prePAAR and PAAR domains. Intriguingly, we identified two spherical densities in the center of the  $\beta$ -sheet prism of the VgrG1 trimer (Figure S 16E). Since we also observed these densities previously in the VgrG1 structure from *P. aeruginosa*<sup>266</sup>, we speculate that this may be a common feature of VgrG1 proteins. Based on the exclusive clustering of positively charged residues around this density and its overall size, we hypothesize that it corresponds to either a phosphate or sulfate anion, which probably helps to stabilize the core of VgrG1 given that in the absence of this anion, the presence of positively charged residues in the core of the protein would be energetically unfavorable.

#### 9.4.7 Model of RhsA firing events and toxin release

Based on the collective structural and functional data presented in this work, combined with the findings of other recently published<sup>270,276</sup> work on T6SS-exported Rhs proteins, we propose a model for cytoplasmic delivery of RhsA and suggest a possible release mechanism for the toxin domain of the effector (Figure 50).

First, RhsA is expressed and autoproteolytically cleaved at defined N- and C-terminal positions. The N-terminal domain comprising the prePAAR motif, TMD region, PAAR domain and the linker to the Rhs barrel likely all simultaneously interact with the EagR1 chaperone and this complex remains associated with the RhsA cocoon. The C-terminal domain comprising the detached toxin domain is partially unfolded and remains inside the cocoon.

EagR1 then shields the transmembrane domains of RhsA from the aqueous milieu as the effector is loaded onto VgrG1 in the cytoplasm of the T6SS-containing cell resulting in a mature PFC. The exact location of T6SS effector delivery in the target cell remains unclear and may

differ depending on the T6SS tail spike complex being exported<sup>277,278</sup>. Nonetheless, most characterized PAAR effectors exert their catalytic activity in the cytosol of the target cell, i.e. (p)ppApp synthetases, ADP-ribosyl transferases, DNases and NAD<sup>+</sup>/NADP<sup>+</sup> hydrolases<sup>272,279–281</sup>. We propose that the RhsA-loaded VgrG1 tip is delivered to the periplasm as we have previously suggested for the Tse6 effector and that the PFC spontaneously enters the target cell inner membrane<sup>266</sup>. It remains unclear at which step the chaperone is stripped off; however, Coulthurst and colleagues detected a secreted T6SS Rhs effector in *Serratia* by mass spectrometry and did not detect its cognate Eag chaperone<sup>282</sup>. This finding supports a model whereby EagR1 dissociates from the spike complex during the loading event of the PFC onto an Hcp tube inside the lumen of the T6SS baseplate. But given the observation that the TMD helices are unstable without their protective chaperones<sup>150</sup>, it is also conceivable that EagR1 is removed shortly before insertion into the membrane of the target cell. As proposed previously for Tse6, we suggest that the transmembrane helices of the TMD region spontaneously enter the inner membrane of the target cell<sup>266</sup>.

Given the rigid structure of RhsA, we propose that the barrel remains intact during the firing and translocation events. Two different scenarios are conceivable for the release of the toxin domain from the cocoon and its delivery into the cytosol of the target cell if the VgrG spike only protrudes into the periplasm. In one scenario, translocation of the linker between the PAAR domain and the plug domain would lead to the removal of the seal peptide. This would result in the opening of a passageway through which the toxin domain could be threaded in an unfolded or partially unfolded state and subsequently translocated, with the assistance of the N-terminal transmembrane helices, into the cytosol. In this case, the RhsA barrel would remain in the periplasm but interact directly with the inner membrane of the target cell. This model has similarity to the currently proposed model for translocation of diphtheria toxin from *Corynebacterium diphtheriae*. In this case, two  $\alpha$ -helical hairpins insert into the endosomal membrane and form a translocation channel through which the toxin is delivered into the cytosol of the intoxicated cell<sup>283</sup>. The refolding of diphtheria toxin in the target cell cytosol is thought to drive the translocation through the loosened-up membrane structure created by the two inserted transmembrane helices. Moreover, it was previously shown<sup>270</sup> that the two VgrG-interacting domains of the T6SS effector TseI from *A. dhakensis*, denoted as VIRN and VIRC, which are equivalent to the cleaved N- and C-terminal domains of RhsA, interact directly with one other even in the absence of the Rhs shell. Hence, it seems plausible that once the seal peptide is removed from RhsA, the unfolded C-terminal toxin domain will be threaded out of the barrel via interactions with its N-terminal domain.

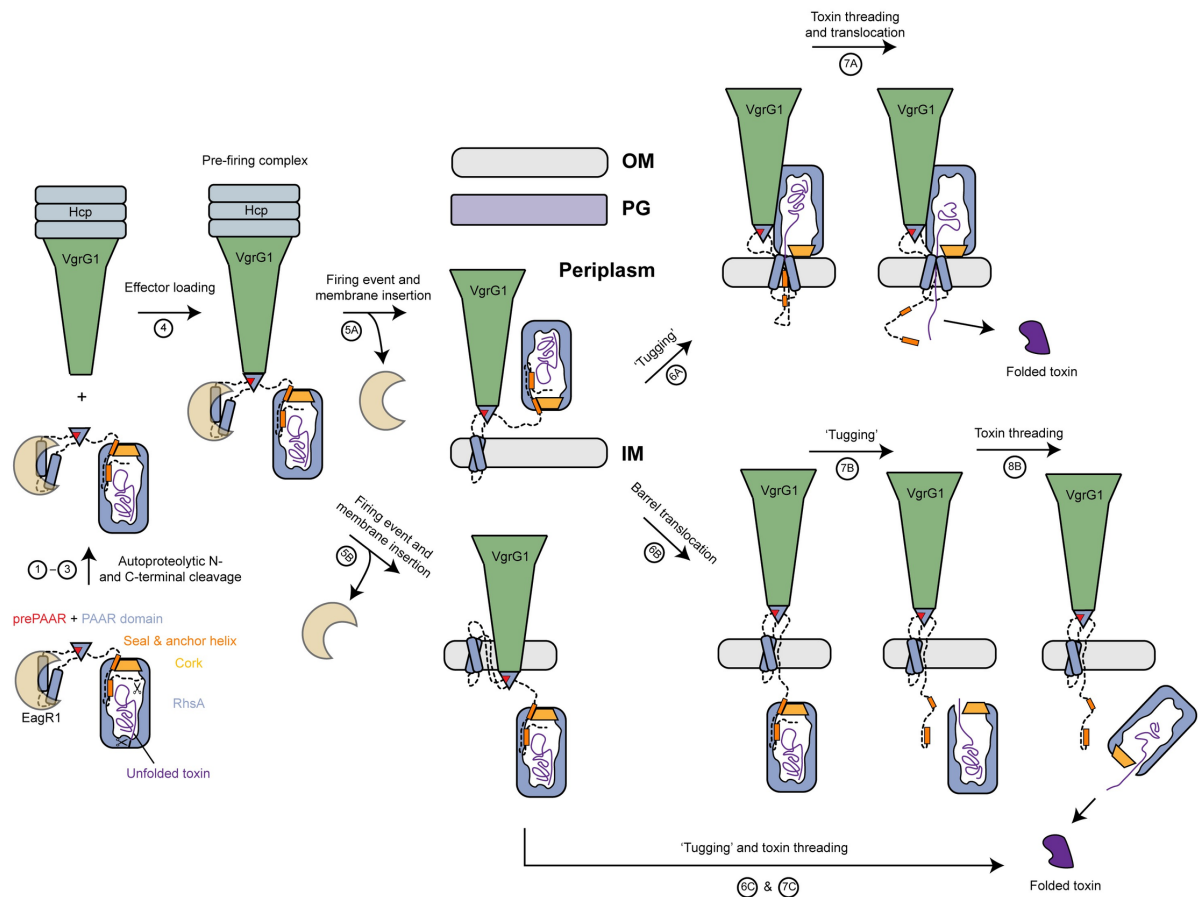


Figure 50: Model of T6SS-dependent delivery of RhsA into the cytoplasm of a susceptible bacterial cell.

(1-3) RhsA undergoes N- and C-terminal autoproteolytic processing. The prePAAR motif ‘completes’ the PAAR domain fold. The EagR1 chaperone solubilizes the two transmembrane helices of RhsA to facilitate loading onto its cognate cytoplasmic VgrG1 (via the prePAAR + PAAR domain). (4) The secretion competent RhsA effector is loaded onto the VgrG1 spike. (5) During a firing event the EagR1 chaperone is dissociated from the complex and the T6SS injects the PFC into target cells where it crosses the peptidoglycan (PG) layer and inserts into the inner membrane (IM) (5A). Alternatively, the tip of the VgrG1 spike could be directly delivered into the cytosol along with RhsA (5B). (6-8) Two different scenarios (6A or 6B, C) are proposed as possible mechanisms for toxin domain release from the cocoon into the cytosol of the target cell. Either the toxin domain alone (mechanism 6A) or the entire RhsA barrel is translocated across the inner membrane (mechanism 6B). In both cases, the seal of the cocoon is likely removed by translocation-induced pulling and the energy required for release and translocation of the toxin domain out of the Rhs cage is probably driven by its spontaneous refolding in the prey bacterium’s cytosol. This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

In a second scenario, the TMD region would facilitate the translocation of the entire RhsA protein. Mechanical forces during translocation in combination with hydrophobic interactions in the membrane could lead to ‘tugging’ of the N-terminal domain and a release of the toxin domain from the cocoon. However, since the barrel is large (86 x 65 Å), its transport across the membrane is relatively unlikely. In both scenarios, the energy required for the release and

translocation of the toxin domain into the target cell cytosol would be driven by spontaneous toxin refolding<sup>284</sup>.

An alternative route would be the direct delivery of the barrel into the cytosol. This would require that the PFC not only penetrates the outer membrane, but also the inner membrane to reach the cytosol of the target cell. In this case, the cocoon would be located in the cytosol of the target cell, and the toxin domain would be released analogously to above-described scenario B. However, it is unclear what the role of the TMD would be in this alternative delivery mechanism.

Besides the proposed scenarios, other pathways for RhsA delivery and toxin release that have not been described are also conceivable. For example, periplasmic or inner membrane proteins found in the target cell could be involved in the active transport of the toxin from the periplasm into the cytosol.

Overall, this study provides detailed molecular insights into the autoproteolytic processing of Rhs effectors and its importance for toxin release inside the target cell. It not only enhances our knowledge about Rhs effector function, but also lays a foundation for a mechanistic understanding of how the T6SS machinery functions. The unique ability of the T6SS to mediate contact-dependent killing of a wide range of bacteria may enable development of novel therapeutics for selective depletion of drug-resistant bacterial pathogens.

## 10 References

1. Berg, J., Tymoczko, J. L. & Stryer, L. *Biochemie*. (2014).
2. Ohvo-Rekilä, H., Ramstedt, B., Leppimäki, P. & Peter Slotte, J. Cholesterol interactions with phospholipids in membranes. *Progress in Lipid Research* **41**, 66–97 (2002).
3. Howard, F. & Ferris, H. The function of sterols in membranes. *Biochimica et Biophysica Acta* **457**, 109–132 (1976).
4. Yeagle, P. L. Cholesterol and the cell membrane. *BBA - Reviews on Biomembranes* **822**, 267–287 (1985).
5. Yeagle, P. L. Modulation of membrane function by cholesterol. *Biochimie* **73**, 1303–1310 (1991).
6. K., Simons. & E., Ikonen. Functional rafts in cell membranes. *Nature* **387**, 569–572 (1997).
7. Simons, K. & Toomre, D. Lipid rafts and signal transduction. *Nature Reviews Molecular Cell Biology* **1**, 31–39 (2000).
8. Schroeder, F. *et al.* Recent advances in membrane cholesterol domain dynamics and intracellular cholesterol trafficking. *Proc Soc Exp Biol Med* **213**, 150–177 (1996).
9. Bretscher, M. S. & Munro, S. Cholesterol and the Golgi apparatus. *Science* **261**, 1280–1281 (1993).
10. Van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: Where they are and how they behave. *Nature Reviews Molecular Cell Biology* **9**, 112–124 (2008).
11. Spector, A. A. & Yorek, M. A. Membrane lipid composition and cellular function. *J Lipid Res* **26**, 1015–1035 (1985).
12. Töröcsik, D., Szanto, A. & Nagy, L. Oxysterol signaling links cholesterol metabolism and inflammation via the liver X receptor in macrophages. *Molecular Aspects of Medicine* **30**, 134–152 (2009).
13. ASBMB Today. *American Society for Biochemistry and Molecular Biology* **12**, 46 (2013).
14. Hammerschmidt, M., Brook, A. & McMahon, A. P. The world according to hedgehog. *Trends in Genetics* **13**, 14–21 (1997).
15. Porter, J. A. *et al.* Hedgehog patterning activity: Role of a lipophilic modification mediated by the carboxy-terminal autoprocessing domain. *Cell* **86**, 21–34 (1996).
16. Lewis, P. M. *et al.* Cholesterol modification of sonic hedgehog is required for long-range signaling activity and effective modulation of signaling by Ptc1. *Cell* **105**, 599–612 (2001).

17. Jiang, J. & Hui, C. chung. Hedgehog Signaling in Development and Cancer. *Developmental Cell* **15**, 801–812 (2008).
18. Xiao, X. *et al.* Cholesterol Modification of Smoothed Is Required for Hedgehog Signaling. *Molecular Cell* **66**, 154-162.e10 (2017).
19. Turley, S. D. & Dietschy, J. M. The Intestinal Absorption of Biliary and Dietary Cholesterol as a Drug Target for Lowering the Plasma Cholesterol Level. *Preventive Cardiology* **6**, 29–33 (2003).
20. Goldstein, J. L. & Brown, M. S. Regulation of the mevalonate pathway. *Nature* **434**, 425–430 (1990).
21. Bloch, K. The Biological Synthesis of Cholesterol. *Science (New York, N.Y.)* **150**, 19–28 (1965).
22. Kandutsch, A. A. & Russell, A. E. Preputial Gland Tumor Sterols. *The Journal of biological chemistry* **235**, 2256–2261 (1960).
23. Mitsche, M. A., McDonald, J. G., Hobbs, H. H. & Cohen, J. C. Flux analysis of cholesterol biosynthesis in vivo reveals multiple tissue and cell-type specific pathways. *eLife* **4**, 1–21 (2015).
24. Correa-Cerro, L. S. & Porter, F. D. 3 $\beta$ -Hydroxysterol  $\Delta$ 7-reductase and the Smith-Lemli-Opitz syndrome. *Molecular Genetics and Metabolism* **84**, 112–126 (2005).
25. Fitzky, B. U. *et al.* Mutations in the Delta7-sterol reductase gene in patients with the Smith-Lemli-Opitz syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 8181–6 (1998).
26. Brown, M. S. & Goldstein, J. L. Receptor-mediated endocytosis: insights from the lipoprotein receptor system. *Proceedings of the National Academy of Sciences* **76**, 3330–3337 (1979).
27. Boyd, G. S., Hattersley, N. G., Ozasa, S. & Suckling, K. E. CHOLESTEROL 7 $\alpha$ -HYDROXYLASE IN RAT LIVER MICROSOMES - CHARACTERISTICS, SOLUBILISATION, RESOLUTION, AND SUBSTRATE SPECIFICITY. in *Microsomes, Drug Oxidations and Chemical Carcinogenesis* (eds. Coon, M. J. et al.) 699–708 (Academic Press, 1980). doi:10.1016/B978-0-12-187702-6.50019-0.
28. Roitelman, J., Olender, E., Bar-Nun, S., Dunn, W., Jr & Simoni, R. Immunological evidence for eight spans in the membrane domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase: implications for enzyme degradation in the endoplasmic reticulum. *Journal of Cell Biology* **117**, 959–973 (1992).



29. Hua, X., Sakai, J., Brown, M. S. & Goldstein, J. L. Regulated Cleavage of Sterol Regulatory Element Binding Proteins Requires Sequences on Both Sides of the Endoplasmic Reticulum Membrane (\*). *Journal of Biological Chemistry* **271**, 10379–10384 (1996).
30. Kuwabara, P. E. & Labouesse, M. The sterol-sensing domain: multiple families, a unique role? *Trends in Genetics* **18**, 193–201 (2002).
31. Liscum, L. *et al.* Domain structure of 3-hydroxy-3-methylglutaryl coenzyme A reductase, a glycoprotein of the endoplasmic reticulum. *Journal of Biological Chemistry* **260**, 522–530 (1985).
32. Istvan, E. S., Palnitkar, M., Buchanan, S. K. & Deisenhofer, J. Crystal structure of the catalytic portion of human HMG-CoA reductase: Insights into regulation of activity and catalysis. *EMBO Journal* **19**, 819–830 (2000).
33. Istvan, E. S. & Deisenhofer, J. The structure of the catalytic portion of human HMG-CoA reductase. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1529**, 9–18 (2000).
34. Istvan, E. S. & Deisenhofer, J. Structural Mechanism for Statin Inhibition of HMG-CoA Reductase. *Science* **292**, 1160–1164 (2002).
35. Omkumar, R. V., Darnay, B. G. & Rodwell, V. W. Modulation of Syrian hamster 3-hydroxy-3-methylglutaryl-CoA reductase activity by phosphorylation. Role of serine 871. *Journal of Biological Chemistry* **269**, 6810–6814 (1994).
36. Clarke, P. R. & Hardie, D. G. 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* **9**, 2439–2446 (1990).
37. Benjamin, E. J. *et al.* Heart Disease and Stroke Statistics—2017 Update: A Report From the American Heart Association. *Circulation* **135**, e146–e603 (2017).
38. Goldstein, J. L. & Brown, M. S. Binding and Degradation of Low Density Lipoproteins by Cultured Human Fibroblasts: COMPARISON OF CELLS FROM A NORMAL SUBJECT AND FROM A PATIENT WITH HOMOZYGOUS FAMILIAL HYPERCHOLESTEROLEMIA. *Journal of Biological Chemistry* **249**, 5153–5162 (1974).
39. Bonetti, P. O., Lerman, L. O. & Lerman, A. Endothelial Dysfunction. *Arteriosclerosis, Thrombosis, and Vascular Biology* **23**, 168–175 (2003).
40. Steinberg, D. & Witztum, J. L. Oxidized Low-Density Lipoprotein and Atherosclerosis. *ATVB* **30**, 2311–2316 (2010).
41. Schaffner, T. *et al.* Arterial Foam Cells With Distinctive Immunomorphologic and Histochemical Features of Macrophages. **100**, (1980).

42. Hansson, G. K. & Jonasson, L. The Discovery of Cellular Immunity in the Atherosclerotic Plaque. *Arteriosclerosis, Thrombosis, and Vascular Biology* **29**, 1714–1717 (2009).
43. Ross, R. & Glomset, J. A. The Pathogenesis of Atherosclerosis. *N Engl J Med* **295**, 369–377 (1976).
44. Arroyo, L. H. & Lee, R. T. Mechanisms of plaque rupture: mechanical and biologic interactions. *Cardiovascular Research* **41**, 369–375 (1999).
45. ENDO, A. A historical perspective on the discovery of statins. *Proc Jpn Acad Ser B Phys Biol Sci* **86**, 484–493 (2010).
46. Endo, A., Kuroda, M. & Tanzawa, K. Competitive inhibition of 3-hydroxy-3-methylglutaryl coenzyme a reductase by ML-236A and ML-236B fungal metabolites, having hypocholesterolemic activity. *FEBS Letters* **72**, 323–326 (1976).
47. Brown, A. G., Smale, T. C., King, T. J., Hasenkamp, R. & Thompson, R. H. Crystal and molecular structure of compactin, a new antifungal metabolite from *Penicillium brevicompactum*. *J. Chem. Soc., Perkin Trans. 1* 1165–1170 (1976) doi:10.1039/P19760001165.
48. Endo, A., Tsujita, Y., Kuroda, M. & Tanzawa, K. Inhibition of Cholesterol Synthesis in vitro and in vivo by ML-236A and ML-236B, Competitive Inhibitors of 3-Hydroxy-3-methylglutaryl-Coenzyme A Reductase. *European Journal of Biochemistry* **77**, 31–36 (1977).
49. Fleming, C. D. *et al.* Structural Insights into Drug Processing by Human Carboxylesterase 1: Tamoxifen, Mevastatin, and Inhibition by Benzil. *Journal of Molecular Biology* **352**, 165–177 (2005).
50. Corsini, A. Pharmacology of competitive inhibitors of HMG-CoA reductase. *Pharmacological Research* **31**, 9–27 (1995).
51. Joy, T. R. & Hegele, R. A. Narrative Review: Statin-Related Myopathy. *Ann Intern Med* **150**, 858–868 (2009).
52. Kashani, A. *et al.* Risks associated with statin therapy: a systematic overview of randomized clinical trials. *Circulation* **114**, 2788–2797 (2006).
53. Cao, P. *et al.* Statin-induced muscle damage and atrogen-1 induction is the result of a geranylgeranylation defect. *FASEB J* **23**, 2844–2854 (2009).
54. Brown, M. S. & Goldstein, J. L. Multivalent feedback regulation of HMG CoA reductase, a control mechanism coordinating isoprenoid synthesis and cell growth. *J Lipid Res* **21**, 505–517 (1980).

55. Schonewille, M. *et al.* Statins increase hepatic cholesterol synthesis and stimulate fecal cholesterol elimination in mice. *Journal of Lipid Research* **57**, 1455–1464 (2016).
56. Kita, T., Brown, M. S. & Goldstein, J. L. Feedback Regulation of 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase in Livers of Mice Treated with Mevinolin, a Competitive Inhibitor of the Reductase. *J Clin Invest* **66**, 1094–1100 (1980).
57. Ness, G. C., Chambers, C. M. & Lopez, D. Atorvastatin action involves diminished recovery of hepatic HMG-CoA reductase activity. *Journal of Lipid Research* **39**, 75–84 (1998).
58. Bilhartz, L. E., Spady, D. K. & Dietschy, J. M. Inappropriate hepatic cholesterol synthesis expands the cellular pool of sterol available for recruitment by bile acids in the rat. *J Clin Invest* **84**, 1181–1187 (1989).
59. Reihner, E. *et al.* Influence of Pravastatin, a Specific Inhibitor of HMG-CoA Reductase, on Hepatic Metabolism of Cholesterol. *New England Journal of Medicine* **323**, 224–228 (1990).
60. Gil, G., Faust, J. R., Chin, D. J., Goldstein, J. L. & Brown, M. S. Membrane-bound domain of HMG CoA reductase is required for sterol-enhanced degradation of the enzyme. *Cell* **41**, 249–258 (1985).
61. Skalnik, D. G., Narita, H., Kent, C. & Simoni, R. D. The membrane domain of 3-hydroxy-3-methylglutaryl-coenzyme A reductase confers endoplasmic reticulum localization and sterol-regulated degradation onto beta-galactosidase. *Journal of Biological Chemistry* **263**, 6836–6841 (1988).
62. Ravid, T., Doolman, R., Avner, R., Harats, D. & Roitelman, J. The Ubiquitin-Proteasome Pathway Mediates the Regulated Degradation of Mammalian 3-Hydroxy-3-methylglutaryl-coenzyme A Reductase \*. *Journal of Biological Chemistry* **275**, 35840–35847 (2000).
63. Inoue, S., Bar-Nun, S., Roitelman, J. & Simoni, R. D. Inhibition of degradation of 3-hydroxy-3-methylglutaryl-coenzyme A reductase in vivo by cysteine protease inhibitors. *Journal of Biological Chemistry* **266**, 13311–13317 (1991).
64. Sever, N. *et al.* Insig-dependent Ubiquitination and Degradation of Mammalian 3-Hydroxy-3-methylglutaryl-CoA Reductase Stimulated by Sterols and Geranylgeraniol \*. *Journal of Biological Chemistry* **278**, 52479–52490 (2003).
65. Lee, P. C. W., Sever, N. & DeBose-Boyd, R. A. Isolation of Sterol-resistant Chinese Hamster Ovary Cells with Genetic Deficiencies in Both Insig-1 and Insig-2 \*. *Journal of Biological Chemistry* **280**, 25242–25249 (2005).

66. Song, B.-L., Sever, N. & DeBose-Boyd, R. A. Gp78, a Membrane-Anchored Ubiquitin Ligase, Associates with Insig-1 and Couples Sterol-Regulated Ubiquitination to Degradation of HMG CoA Reductase. *Molecular Cell* **19**, 829–840 (2005).
67. Jo, Y., Lee, P. C. W., Sguigna, P. V. & DeBose-Boyd, R. A. Sterol-induced degradation of HMG CoA reductase depends on interplay of two Insigs and two ubiquitin ligases, gp78 and Trc8. *Proc Natl Acad Sci U S A* **108**, 20503–20508 (2011).
68. Jiang, L.-Y. *et al.* Ring finger protein 145 (RNF145) is a ubiquitin ligase for sterol-induced degradation of HMG-CoA reductase. *J Biol Chem* **293**, 4047–4055 (2018).
69. van de Weijer, M. L. *et al.* Multiple E2 ubiquitin-conjugating enzymes regulate human cytomegalovirus US2-mediated immunoreceptor downregulation. *J Cell Sci* **130**, 2883–2892 (2017).
70. Ye, Y., Meyer, H. H. & Rapoport, T. A. The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol. *Nature* **414**, 652–656 (2001).
71. Jiang, S.-Y. *et al.* Discovery of a potent HMG-CoA reductase degrader that eliminates statin-induced reductase accumulation and lowers cholesterol. *Nat Commun* **9**, 5138 (2018).
72. Nakagawa, K. *et al.* Identification of UBIAD1 as a novel human menaquinone-4 biosynthetic enzyme. *Nature* **468**, 117–121 (2010).
73. Schumacher, M. M., Elsabrouty, R., Seemann, J., Jo, Y. & DeBose-Boyd, R. A. The prenyltransferase UBIAD1 is the target of geranylgeraniol in degradation of HMG CoA reductase. *eLife* **4**, e05560 (2015).
74. Jiang, S.-Y. *et al.* Schnyder corneal dystrophy-associated UBIAD1 mutations cause corneal cholesterol accumulation by stabilizing HMG-CoA reductase. *PLOS Genetics* **15**, e1008289 (2019).
75. Schumacher, M. M., Jun, D.-J., Johnson, B. M. & DeBose-Boyd, R. A. UbiA prenyltransferase domain-containing protein-1 modulates HMG-CoA reductase degradation to coordinate synthesis of sterol and nonsterol isoprenoids. *Journal of Biological Chemistry* **293**, 312–323 (2018).
76. Klintworth, G. K. Corneal dystrophies. *Orphanet J Rare Dis* **4**, 7 (2009).
77. Weiss, J. S. Schnyder corneal dystrophy. *Current Opinion in Ophthalmology* **20**, 292 (2009).
78. Weiss, J. S. *et al.* Mutations in the UBIAD1 Gene on Chromosome Short Arm 1, Region 36, Cause Schnyder Crystalline Corneal Dystrophy. *Investigative Ophthalmology & Visual Science* **48**, 5007–5012 (2007).

79. Nowinska, A. K. *et al.* Phenotype–Genotype Correlation in Patients With Schnyder Corneal Dystrophy. *Cornea* **33**, 497 (2014).
80. Nickerson, M. L. *et al.* The UBIAD1 Prenyltransferase Links Menaquinone-4 Synthesis to Cholesterol Metabolic Enzymes. *Human Mutation* **34**, 317–329 (2013).
81. Cheng, W. & Li, W. Structural Insights into Ubiquinone Biosynthesis in Membranes. *Science* **343**, 878–881 (2014).
82. Huang, H. *et al.* Structure of a Membrane-Embedded Prenyltransferase Homologous to UBIAD1. *PLOS Biology* **12**, e1001911 (2014).
83. Chen, H. *et al.* Regulated degradation of HMG CoA reductase requires conformational changes in sterol-sensing domain. *Nat Commun* **13**, 4273 (2022).
84. Mukherjee, S. *et al.* Synthetic antibodies against BRIL as universal fiducial marks for single-particle cryoEM structure determination of membrane proteins. *Nat Commun* **11**, 1598 (2020).
85. Brown, M. S. & Goldstein, J. L. The SREBP Pathway : Regulation of Cholesterol Metabolism by Proteolysis of a Membrane-Bound Transcription Factor. *Cell* **89**, 331–340 (1997).
86. Goldstein, J. L., DeBose-Boyd, R. A. & Brown, M. S. Protein sensors for membrane sterols. *Cell* **124**, 35–36 (2006).
87. Yokoyama, C. *et al.* SREBP-1 , a Basic-Helix-Loop-Helix-Leucine Zipper Protein That Controls Transcription of the Low Density Lipoprotein Receptor Gene. *Cell* **75**, 187–197 (1993).
88. Hua, X. *et al.* SREBP-2 , a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 11603–11607 (1993).
89. Sakai, J. Identification of complexes between the COOH-terminal domains of sterol regulatory element-binding proteins (SREBPs) and SREBP cleavage-activating protein. *J Biol Chem* **272**, 20213–20221 (1997).
90. Hua, X., Nohturfft, A., Goldstein, J. L. & Brown, M. S. Sterol resistance in CHO cells traced to point mutation in SREBP cleavage-activating protein. *Cell* **87**, 415–426 (1996).
91. Radhakrishnan, A., Goldstein, J. L., McDonald, J. G. & Brown, M. S. Switch-like Control of SREBP-2 Transport Triggered by Small Changes in ER Cholesterol : A Delicate Balance. *Cell Metabolism* **8**, 512–521 (2008).

92. Sun, L., Li, L., Goldstein, J. L. & Brown, M. S. Insig Required for Sterol-mediated Inhibition of Scap / SREBP Binding to COPII Proteins in Vitro. *The Journal of biological chemistry* **280**, 26483–26490 (2005).
93. Sakai, J. *et al.* Sterol-regulated release of SREBP-2 from cell membranes requires two sequential cleavages, one within a transmembrane segment. *Cell* **85**, 1037–1046 (1996).
94. Rawson, R. B. *et al.* Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs. *Mol Cell* **1**, 47–57 (1997).
95. Sakai, J. *et al.* Molecular identification of the sterol-regulated luminal protease that cleaves SREBPs and controls lipid composition of animal cells. *Mol Cell* **2**, 505–514 (1998).
96. Lee, S. J., Sekimoto, T. & Yamashita, E. The Structure of Importin-beta bound to SREBP-2: nuclear import of a transcription factor. *Science* **302**, 1571–1576 (2003).
97. Nagoshi, E. M. I. & Yoneda, Y. Dimerization of Sterol Regulatory Element-Binding Protein 2 via the Helix-Loop-Helix-Leucine Zipper Domain Is a Prerequisite for Its Nuclear Localization Mediated by Importin beta. *Molecular and Cellular Biology* **21**, 2779–2789 (2001).
98. Horton, J. D. *et al.* SREBPs : activators of the complete program of cholesterol and fatty acid synthesis in the liver Find the latest version : synthesis in the liver. *Journal of Clinical Investigation* **109**, 1125–1131 (2002).
99. Brown, A. J., Sun, L., Feramisco, J. D., Brown, M. S. & Goldstein, J. L. Cholesterol Addition to ER Membranes Alters Conformation of SCAP, the SREBP Escort Protein that Regulates Cholesterol Metabolism. *Molecular Cell* **10**, 237–245 (2002).
100. Yang, T. *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* **110**, 489–500 (2002).
101. Yabe, D., Brown, M. S. & Goldstein, J. L. Insig-2, a second endoplasmic reticulum protein that binds SCAP and blocks export of sterol regulatory element-binding proteins. *Proc Natl Acad Sci U S A* **99**, 12753–12758 (2002).
102. Yabe, D., Xia, Z.-P., Adams, C. M. & Rawson, R. B. Three mutations in sterol-sensing domain of SCAP block interaction with insig and render SREBP cleavage insensitive to sterols. *Proc Natl Acad Sci U S A* **99**, 16672–16677 (2002).
103. Zhang, Y. *et al.* Direct Demonstration That Loop1 of Scap Binds to Loop7. *J Biol Chem* **291**, 12888–12896 (2016).

104. Zhang, Y., Motamed, M., Seemann, J., Brown, M. S. & Goldstein, J. L. Point Mutation in Luminal Loop 7 of Scap Protein Blocks Interaction with Loop 1 and Abolishes Movement to Golgi \*. *Journal of Biological Chemistry* **288**, 14059–14067 (2013).
105. Motamed, M. *et al.* Identification of Luminal Loop 1 of Scap Protein as the Sterol Sensor That Maintains Cholesterol Homeostasis. *J Biol Chem* **286**, 18002–18012 (2011).
106. Gao, Y., Zhou, Y., Goldstein, J. L., Brown, M. S. & Radhakrishnan, A. Cholesterol-induced conformational changes in the sterol-sensing domain of the Scap protein suggest feedback mechanism to control cholesterol synthesis. *Journal of Biological Chemistry* **292**, 8729–8737 (2017).
107. Feramisco, J. D. *et al.* Intramembrane aspartic acid in SCAP protein governs cholesterol-induced conformational change. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3242–3247 (2005).
108. Ren, R. *et al.* Crystal structure of a mycobacterial Insig homolog provides insight into how these sensors monitor sterol levels. *Science* **349**, 187–191 (2015).
109. Gong, X. *et al.* Structure of the WD40 domain of SCAP from fission yeast reveals the molecular basis for SREBP recognition. *Nature Publishing Group* **25**, 401–411 (2015).
110. Gong, X. *et al.* Complex structure of the fission yeast SREBP-SCAP binding domains reveals an oligomeric organization. *Cell Research* **26**, 1197–1211 (2016).
111. Yan, R. *et al.* A structure of human Scap bound to Insig-2 suggests how their interaction is regulated by sterols. *Science* **371**, eabb2224 (2021).
112. Yan, R. *et al.* Structural basis for sterol sensing by Scap and Insig. *Cell Reports* **35**, 109299 (2021).
113. Kober, D. L. *et al.* Scap structures highlight key role for rotation of intertwined luminal loops in cholesterol sensing. *Cell* **184**, 3689-3701.e22 (2021).
114. Li, X., Saha, P., Li, J., Blobel, G. & Pfeffer, S. R. Clues to the mechanism of cholesterol transfer from the structure of NPC1 middle luminal domain bound to NPC2. *Proc Natl Acad Sci U S A* **113**, 10079–10084 (2016).
115. Winkler, M. B. L. *et al.* Structural Insight into Eukaryotic Sterol Transport through Niemann-Pick Type C Proteins. *Cell* **179**, 485-497.e18 (2019).
116. Qian, H. *et al.* Structural Basis of Low-pH-Dependent Lysosomal Cholesterol Egress by NPC1 and NPC2. *Cell* **182**, 98-111.e18 (2020).
117. Huang, C.-S. *et al.* Cryo-EM structures of NPC1L1 reveal mechanisms of cholesterol transport and ezetimibe inhibition. *Sci Adv* **6**, eabb1989 (2020).

118. Gong, X. *et al.* Structural basis for the recognition of Sonic Hedgehog by human Patched1. *Science* **361**, eaas8935 (2018).
119. Cannac, F. *et al.* Cryo-EM structure of the Hedgehog release protein Dispatched. *Sci Adv* **6**, eaay7928 (2020).
120. Radhakrishnan, A., Sun, L.-P., Kwon, H. J., Brown, M. S. & Goldstein, J. L. Direct binding of cholesterol to the purified membrane region of SCAP: mechanism for a sterol-sensing domain. *Mol Cell* **15**, 259–268 (2004).
121. Papanikou, E., Karamanou, S. & Economou, A. Bacterial protein secretion through the translocase nanomachine. *Nat Rev Microbiol* **5**, 839–851 (2007).
122. Veenendaal, A. K. J., van der Does, C. & Driessen, A. J. M. The protein-conducting channel SecYEG. *Biochim Biophys Acta* **1694**, 81–95 (2004).
123. Chen, Y., Shanmugam, S. K. & Dalbey, R. E. The Principles of Protein Targeting and Transport Across Cell Membranes. *Protein J* **38**, 236–248 (2019).
124. Palmer, T. & Berks, B. C. The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol* **10**, 483–496 (2012).
125. Cianfanelli, F. R., Monlezun, L. & Coulthurst, S. J. Aim, Load, Fire: The Type VI Secretion System, a Bacterial Nanoweapon. *Trends in Microbiology* **24**, 51–62 (2016).
126. Costa, T. R. D. *et al.* Secretion systems in Gram-negative bacteria: Structural and mechanistic insights. *Nature Reviews Microbiology* **13**, 343–359 (2015).
127. Mougous, J. D. *et al.* A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* **312**, 1526–1530 (2006).
128. Pukatzki, S. *et al.* Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc Natl Acad Sci U S A* **103**, 1528–1533 (2006).
129. Boyer, F., Fichant, G., Berthod, J., Vandenbrouck, Y. & Attree, I. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? *BMC Genomics* **10**, 104 (2009).
130. Hood, R. D. *et al.* A Type VI Secretion System of *Pseudomonas aeruginosa* Targets a Toxin to Bacteria. *Cell Host & Microbe* **7**, 25–37 (2010).
131. MacIntyre, D. L., Miyata, S. T., Kitaoka, M. & Pukatzki, S. The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proc Natl Acad Sci U S A* **107**, 19520–19524 (2010).
132. Murdoch, S. L. *et al.* The opportunistic pathogen *Serratia marcescens* utilizes type VI secretion to target bacterial competitors. *J Bacteriol* **193**, 6057–6069 (2011).



133. Chang, Y., Rettberg, L. A., Ortega, D. R. & Jensen, G. J. In vivo structures of an intact type VI secretion system revealed by electron cryotomography. *EMBO reports* **18**, 1090–1099 (2017).
134. Leiman, P. G. *et al.* Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc Natl Acad Sci U S A* **106**, 4154–4159 (2009).
135. Silverman, J. M., Brunet, Y. R., Cascales, E. & Mougous, J. D. Structure and regulation of the type VI secretion system. *Annual Review of Microbiology* **66**, 453–472 (2012).
136. Cianfanelli, F. R., Monlezun, L. & Coulthurst, S. J. Aim, Load, Fire: The Type VI Secretion System, a Bacterial Nanoweapon. *Trends in Microbiology* **24**, 51–62 (2016).
137. Ho, B. T., Dong, T. G. & Mekalanos, J. J. A view to a kill: The bacterial type VI secretion system. *Cell Host and Microbe* **15**, 9–21 (2014).
138. Shneider, M. M. *et al.* PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature* **500**, 350–353 (2013).
139. Lossi, N. S. *et al.* The HsiB1C1 (TssB-TssC) complex of the *Pseudomonas aeruginosa* type VI secretion system forms a bacteriophage tail sheathlike structure. *Journal of Biological Chemistry* **288**, 7536–7548 (2013).
140. CC-BY license cell.
141. English, G. *et al.* New secreted toxins and immunity proteins encoded within the type VI secretion system gene cluster of *Serratia marcescens*. *Molecular Microbiology* **86**, 921–936 (2012).
142. Ting, S. Y. *et al.* Bifunctional Immunity Proteins Protect Bacteria against FtsZ-Targeting ADP-Ribosylating Toxins. *Cell* **175**, 1380-1392.e14 (2018).
143. Pei, T. T. *et al.* Intramolecular chaperone-mediated secretion of an Rhs effector toxin by a type VI secretion system. *Nature Communications* **11**, (2020).
144. Iyer, L. M., Zhang, D., Rogozin, I. B. & Aravind, L. Evolution of the deaminase fold and multiple origins of eukaryotic editing and mutagenic nucleic acid deaminases from bacterial toxin systems. *Nucleic Acids Research* **39**, 9473–9497 (2011).
145. Russell, A. B. *et al.* Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475**, 343–349 (2011).
146. Ma, J. *et al.* PAAR-Rhs proteins harbor various C-terminal toxins to diversify the antibacterial pathways of type VI secretion systems. *Environmental Microbiology* **19**, 345–360 (2017).

147. Silverman, J. M. *et al.* Haemolysin Coregulated Protein Is an Exported Receptor and Chaperone of Type VI Secretion Substrates. *Molecular Cell* **51**, 584–593 (2013).
148. Flaugnatti, N. *et al.* A phospholipase A1 antibacterial Type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery. *Molecular Microbiology* **99**, 1099–1118 (2016).
149. Flaugnatti, N. *et al.* Structural basis for loading and inhibition of a bacterial T6SS phospholipase effector by the VgrG spike. *EMBO J* **39**, e104129 (2020).
150. Ahmad, S. *et al.* Structural basis for effector transmembrane domain recognition by type VI secretion system chaperones. *eLife* **9**, 1–29 (2020).
151. CC BY 4.0.
152. Burkinshaw, B. J. *et al.* A type VI secretion system effector delivery mechanism dependent on PAAR and a chaperone-co-chaperone complex. *Nature Microbiology* **3**, 632–640 (2018).
153. Whitney, J. C. *et al.* An Interbacterial NAD(P)<sup>+</sup> Glycohydrolase Toxin Requires Elongation Factor Tu for Delivery to Target Cells. *Cell* **163**, 607–619 (2015).
154. Alcoforado Diniz, J. & Coulthurst, S. J. Intraspecies competition in *Serratia marcescens* is mediated by type VI-secreted Rhs effectors and a conserved effector-associated accessory protein. *Journal of Bacteriology* **197**, 2350–2360 (2015).
155. Quentin, D. *et al.* Mechanism of loading and translocation of type VI secretion system effector Tse6. *Nature Microbiology* **3**, 1142–1152 (2018).
156. Durand, E. *et al.* Biogenesis and structure of a type VI secretion membrane core complex. *Nature* **523**, 555–560 (2015).
157. Ge, P. *et al.* Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states. *Nat Struct Mol Biol* **22**, 377–382 (2015).
158. Hill, C. W. Large genomic sequence repetitions in bacteria: lessons from rRNA operons and Rhs elements. *Research in Microbiology* **150**, 665–674 (1999).
159. Wang, Y. D., Zhao, S. & Hill, C. W. Rhs elements comprise three subfamilies which diverged prior to acquisition by *Escherichia coli*. *Journal of Bacteriology* **180**, 4102–4110 (1998).
160. Hill, C. W. *et al.* Correlation of Rhs elements with *Escherichia coli* population structure. *Genetics* **141**, 15–24 (1995).
161. Zhao, S. & Hill, C. W. Reshuffling of Rhs components to create a new element. *Journal of Bacteriology* **177**, 1393–1398 (1995).

162. Feulner, G. *et al.* Structure of the rhsA locus from Escherichia coli K-12 and comparison of rhsA with other members of the rhs multigene family. *Journal of Bacteriology* **172**, 446–456 (1990).
163. Minet, A. D., Rubin, B. P., Tucker, R. P., Baumgartner, S. & Chiquet-Ehrismann, R. Teneurin-1, a vertebrate homologue of the Drosophila pair-rule gene ten-m, is a neuronal protein with a novel type of heparin-binding domain. *Journal of Cell Science* **112**, 2019–2032 (1999).
164. Jackson, A. P., Thomas, G. H., Parkhill, J. & Thomson, N. R. Evolutionary diversification of an ancient gene family (rhs) through C-terminal displacement. *BMC Genomics* **10**, 1–16 (2009).
165. Hill, C. W., Sandt, C. H. & Vlazny, D. A. Rhs elements of Escherichia coli: a family of genetic composites each encoding a large mosaic protein. *Molecular Microbiology* **12**, 865–871 (1994).
166. Poole, S. J. *et al.* Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (cdi) and rearrangement hotspot (rhs) systems. *PLoS Genetics* **7**, (2011).
167. Aoki, S. K. *et al.* A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* **468**, 439–442 (2010).
168. Bowen, D. *et al.* Insecticidal Toxins from the Bacterium *Photorhabdus luminescens*. *Science* **280**, 2129–2132 (1998).
169. *The genus Yersinia: from genomics to function.* (Springer, 2007).
170. Gatsogiannis, C. *et al.* A syringe-like injection mechanism in *Photorhabdus luminescens* toxins. *Nature* **495**, 520–523 (2013).
171. Meusch, D. *et al.* Mechanism of Tc toxin action revealed in molecular detail. *Nature* **508**, 61–65 (2014).
172. Klasson, L., Kambris, Z., Cook, P. E., Walker, T. & Sinkins, S. P. Horizontal gene transfer between Wolbachia and the mosquito *Aedes aegypti*. *BMC Genomics* **10**, 33 (2009).
173. Liu, S., Xia, X., Calvo, E. & Zhou, Z. H. Native structure of mosquito salivary protein uncovers domains relevant to pathogen transmission. *Nat Commun* **14**, 899 (2023).
174. Korochkina, S. *et al.* A mosquito-specific protein family includes candidate receptors for malaria sporozoite invasion of salivary glands. *Cellular Microbiology* **8**, 163–175 (2006).
175. Woolfit, M., Iturbe-Ormaetxe, I., McGraw, E. A. & O'Neill, S. L. An Ancient Horizontal Gene Transfer between Mosquito and the Endosymbiotic Bacterium *Wolbachia pipientis*. *Molecular Biology and Evolution* **26**, 367–374 (2009).

176. Tucker, R. P. Teneurins: Domain Architecture, Evolutionary Origins, and Patterns of Expression. *Frontiers in Neuroscience* **12**, (2018).
177. Hong, W., Mosca, T. J. & Luo, L. Teneurins instruct synaptic partner matching in an olfactory map. *Nature* **484**, 201–207 (2012).
178. Busby, J. N., Panjikar, S., Landsberg, M. J., Hurst, M. R. H. & Lott, J. S. The BC component of ABC toxins is an RHS-repeat-containing protein encapsulation device. *Nature* **501**, 547–550 (2013).
179. The Role of TcdB and TccC Subunits in Secretion of the *Photobacterium* Tcd Toxin Complex | PLOS Pathogens. <https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1003644>.
180. Gatsogiannis, C. *et al.* Tc toxin activation requires unfolding and refolding of a  $\beta$ -propeller. *Nature* **563**, 209–233 (2018).
181. Roderer, D., Schubert, E., Sitsel, O. & Raunser, S. Towards the application of Tc toxins as a universal protein translocation system. *Nature Communications* **10**, 1–11 (2019).
182. Roderer, D. & Raunser, S. Tc Toxin Complexes: Assembly, Membrane Permeation, and Protein Translocation. *Annu. Rev. Microbiol.* **73**, 247–265 (2019).
183. Jackson, V. A. *et al.* Structures of Teneurin adhesion receptors reveal an ancient fold for cell-cell interaction. *Nature Communications* **9**, 1–9 (2018).
184. Li, J. *et al.* Structural Basis for Teneurin Function in Circuit-Wiring: A Toxin Motif at the Synapse. *Cell* **173**, 735–748.e15 (2018).
185. King, J. G., Vernick, K. D. & Hillyer, J. F. Members of the Salivary Gland Surface Protein (SGS) Family Are Major Immunogenic Components of Mosquito Saliva \*. *Journal of Biological Chemistry* **286**, 40824–40834 (2011).
186. Si, D. & He, J. Modeling Beta-Traces for Beta-Barrels from Cryo-EM Density Maps. *BioMed Research International* **2017**, e1793213 (2017).
187. Quan, J. & Tian, J. Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways. *PLOS ONE* **4**, e6441 (2009).
188. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC biotechnology* **8**, 91 (2008).
189. Goehring, A. *et al.* Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nature Protocols* **9**, 2574–2585 (2014).
190. Moriya, T. *et al.* High-resolution single particle analysis from electron cryo-microscopy images using SPHIRE. *Journal of Visualized Experiments* **2017**, 1–11 (2017).

191. Wagner, T. *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Communications Biology* **2**, 1–13 (2019).
192. Yang, Z., Fang, J., Chittuluru, J., Asturias, F. J. & Penczek, P. A. Iterative stable alignment and clustering of 2D transmission electron microscope images. *Structure* **20**, 237–247 (2012).
193. Stabrin, M. *et al.* TranSPHIRE: automated and feedback-optimized on-the-fly processing for cryo-EM. *Nature Communications* **11**, 1–14 (2020).
194. Zheng, S. Q. *et al.* MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* **14**, 331–332 (2017).
195. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology* **192**, 216–221 (2015).
196. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, 1–22 (2018).
197. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**, 290–296 (2017).
198. Günther, P. *et al.* Structure of a bacterial Rhs effector exported by the type VI secretion system. *PLoS Pathogens* **18**, 1–25 (2022).
199. Hmelo, L. R. *et al.* Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange. *Nature Protocols* **10**, 1820–1841 (2015).
200. Basler, M., Ho, B. T. & Mekalanos, J. J. Tit-for-tat: Type VI secretion system counterattack during bacterial cell-cell interactions. *Cell* **152**, 884–894 (2013).
201. Cardona, S. T. & Valvano, M. A. An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia*. *Plasmid* **54**, 219–228 (2005).
202. Silverman, J. M. *et al.* Haemolysin Coregulated Protein Is an Exported Receptor and Chaperone of Type VI Secretion Substrates. *Molecular Cell* **51**, 584–593 (2013).
203. Zi Tan, Y. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods* **14**, 793–796 (2017).
204. Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Communications Biology* 1–8 (2020) doi:10.1101/2020.06.12.148296.
205. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605–1612 (2004).
206. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **66**, 486–501 (2010).

207. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213–221 (2010).
208. Croll, T. I. ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallographica Section D: Structural Biology* **74**, 519–530 (2018).
209. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science* **27**, 293–315 (2018).
210. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D: Biological Crystallography* **62**, 1002–1011 (2006).
211. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
212. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science* **30**, 70–82 (2021).
213. Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *Journal of Molecular Biology* **430**, 2237–2243 (2018).
214. Gabler, F. *et al.* Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Current Protocols in Bioinformatics* **72**, 1–30 (2020).
215. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
216. Bieniossek, C., Richmond, T. J. & Berger, I. MultiBac: Multigene Baculovirus-Based Eukaryotic Protein Complex Production. *Current Protocols in Protein Science* **51**, 5.20.1-5.20.26 (2008).
217. Sokolenko, S. *et al.* Co-expression vs. co-infection using baculovirus expression vectors in insect cell culture: Benefits and drawbacks. *Biotechnol Adv* **30**, 766–781 (2012).
218. Adams, C. M. *et al.* Cholesterol and 25-hydroxycholesterol inhibit activation of SREBPs by different mechanisms, both involving SCAP and insigs. *Journal of Biological Chemistry* **279**, 52772–52780 (2004).
219. Gewering, T., Janulienė, D., Ries, A. B. & Moeller, A. Know your detergents: A case study on detergent background in negative stain electron microscopy. *Journal of Structural Biology* **203**, 242–246 (2018).
220. Wittig, I., Braun, H.-P. & Schagger, H. Blue native PAGE. *Nat Protoc* **1**, 418–428 (2006).

221. Chaptal, V. *et al.* Quantification of Detergents Complexed with Membrane Proteins. *Sci Rep* **7**, 41751 (2017).
222. Olerinyova, A. *et al.* Mass Photometry of Membrane Proteins. *Chem* **7**, 224–236 (2021).
223. Wright, E. R., Iancu, C. V., Tivol, W. F. & Jensen, G. J. Observations on the behavior of vitreous ice at approximately 82 and approximately 12 K. *J Struct Biol* **153**, 241–252 (2006).
224. Russo, C. J. & Passmore, L. A. Ultrastable gold substrates for electron cryomicroscopy. *Science* **346**, 1377–1381 (2014).
225. Denisov, I. G., Grinkova, Y. V., Lazarides, A. A. & Sligar, S. G. Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. *J Am Chem Soc* **126**, 3477–3487 (2004).
226. Miehl, J., Goricanec, D. & Hagn, F. A Split-Intein-Based Method for the Efficient Production of Circularized Nanodiscs for Structural Studies of Membrane Proteins. *ChemBioChem* **19**, 1–8 (2018).
227. Popovic, K., Holyoake, J., Pomès, R. & Privé, G. G. Structure of saposin A lipoprotein discs. *Proc Natl Acad Sci U S A* **109**, 2908–2912 (2012).
228. Frauenfeld, J. *et al.* A saposin-lipoprotein nanoparticle system for membrane proteins. *Nat Methods* **13**, 345–351 (2016).
229. Brunner, J. D. & Schenck, S. Production and Application of Nanobodies for Membrane Protein Structural Biology. in *Expression, Purification, and Structural Biology of Membrane Proteins* (eds. Perez, C. & Maier, T.) 167–184 (Springer US, 2020). doi:10.1007/978-1-0716-0373-4\_12.
230. Reed, M. C., Lieb, A. & Nijhout, H. F. The biological significance of substrate inhibition: A mechanism with diverse functions. *BioEssays* **32**, 422–429 (2010).
231. Jawaid, S. *et al.* Human hydroxymethylglutaryl-coenzyme a reductase (HMGCR) and statin sensitivity. *Indian Journal of Biochemistry and Biophysics* **47**, 331–339 (2010).
232. Jawaid, S. *et al.* Human hydroxymethylglutaryl-coenzyme a reductase (HMGCR) and statin sensitivity. *Indian Journal of Biochemistry and Biophysics* **47**, 331–339 (2010).
233. Yoshino, M. & Murakami, K. Analysis of the substrate inhibition of complete and partial types. *SpringerPlus* **4**, (2015).
234. Istvan, E. S. & Deisenhofer, J. The structure of the catalytic portion of human HMG-CoA reductase. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1529**, 9–18 (2000).
235. Wu, S. *et al.* Fabs enable single particle cryoEM studies of small proteins. *Structure* **20**, 582–592 (2012).

236. Tsutsumi, N. *et al.* Structure of human Frizzled5 by fiducial-assisted cryo-EM supports a heterodimeric mechanism of canonical Wnt signaling. *eLife* **9**, e58464 (2020).
237. Almén, M. S., Nordström, K. J., Fredriksson, R. & Schiöth, H. B. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biology* **7**, 50 (2009).
238. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat Rev Drug Discov* **16**, 19–34 (2017).
239. Kühlbrandt, W. The resolution revolution. *Science* **343**, 1443–1444 (2014).
240. Milazzo, A.-C. *et al.* Active pixel sensor array as a detector for electron microscopy. *Ultramicroscopy* **104**, 152–159 (2005).
241. McMullan, G., Chen, S., Henderson, R. & Faruqi, A. R. Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy* **109**, 1126–1143 (2009).
242. Kühlbrandt, W. Forty years in cryoEM of membrane proteins. *Microscopy* **71**, i30–i50 (2022).
243. Jordan, M., Köhne, C. & Wurm, F. M. Calcium-phosphate mediated DNA transfer into HEK-293 cells in suspension: Control of physicochemical parameters allows transfection in stirred media: Transfection and protein expression in mammalian cells. *Cytotechnology* **26**, 39–47 (1998).
244. Tribet, C., Audebert, R. & Popot, J. Amphipols : Polymers that keep membrane proteins soluble. *Proc. Natl. Acad. Sci. USA* **93**, 15047–15050 (1996).
245. Knowles, T. J. *et al.* Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer. *J. Am. Chem. Soc.* **131**, 7484–7485 (2009).
246. Bayburt, T. H. & Sligar, S. G. Membrane protein assembly into Nanodiscs. *FEBS Letters* **584**, 1721–1727 (2010).
247. Carlson, M. L. *et al.* The Peptidisc, a simple method for stabilizing membrane proteins in detergent-free solution. *eLife* **7**, e34085 (2018).
248. Yao, X., Fan, X. & Yan, N. Cryo-EM analysis of a membrane protein embedded in the liposome. *Proc Natl Acad Sci U S A* **117**, 18497–18503 (2020).
249. Harrison, P. J., Vecerkova, T., Clare, D. K. & Quigley, A. A review of the approaches used to solve sub-100 kDa membrane proteins by cryo-electron microscopy. *Journal of Structural Biology* **215**, 107959 (2023).
250. Yuan, Y. *et al.* Cryo-EM structure of human glucose transporter GLUT4. *Nat Commun* **13**, 2671 (2022).



251. Liang, Y.-L. *et al.* Phase-plate cryo-EM structure of a class B GPCR-G-protein complex. *Nature* **546**, 118–123 (2017).
252. Velazhahan, V. *et al.* Structure of the class D GPCR Ste2 dimer coupled to two G proteins. *Nature* **589**, 148–153 (2021).
253. García-Nafria, J. & Tate, C. G. Cryo-EM structures of GPCRs coupled to Gs, Gi and Go. *Mol Cell Endocrinol* **488**, 1–13 (2019).
254. Lee, Y. *et al.* Molecular basis of  $\beta$ -arrestin coupling to formoterol-bound  $\beta$ 1-adrenoceptor. *Nature* **583**, 862–866 (2020).
255. Chen, Q. *et al.* Structures of rhodopsin in complex with G-protein-coupled receptor kinase 1. *Nature* **595**, 600–605 (2021).
256. Wang, N. *et al.* Structural basis of human monocarboxylate transporter 1 inhibition by anti-cancer drug candidates. *Cell* **184**, 370–383.e13 (2021).
257. Xu, B. *et al.* Embigin facilitates monocarboxylate transporter 1 localization to the plasma membrane and transition to a decoupling state. *Cell Reports* **40**, 111343 (2022).
258. Hamers-Casterman, C. *et al.* Naturally occurring antibodies devoid of light chains. *Nature* **363**, 446–448 (1993).
259. Henry, K. A. & MacKenzie, C. R. Antigen recognition by single-domain antibodies: structural latitudes and constraints. *MAbs* **10**, 815–826 (2018).
260. Uchański, T. *et al.* Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nat Methods* **18**, 60–68 (2021).
261. Wu, X. & Rapoport, T. A. Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *Proc Natl Acad Sci U S A* **118**, e2115001118 (2021).
262. Bloch, J. S. *et al.* Development of a universal nanobody-binding Fab module for fiducial-assisted cryo-EM studies of membrane proteins. *Proc Natl Acad Sci U S A* **118**, e2115435118 (2021).
263. Fellouse, F. A. *et al.* High-throughput Generation of Synthetic Antibodies from Highly Functional Minimalist Phage-displayed Libraries. *Journal of Molecular Biology* **373**, 924–940 (2007).
264. Zimmermann, I. *et al.* Synthetic single domain antibodies for the conformational trapping of membrane proteins. *eLife* **7**, 1–32 (2018).
265. CC BY 4.0 Plos pathogens.
266. Quentin, D. *et al.* Mechanism of loading and translocation of type VI secretion system effector Tse6. *Nature Microbiology* **3**, 1142–1152 (2018).

267. Meusch, D. *et al.* Mechanism of Tc toxin action revealed in molecular detail. *Nature* **508**, 61–65 (2014).
268. Busby, J. N., Panjikar, S., Landsberg, M. J., Hurst, M. R. H. & Lott, J. S. The BC component of ABC toxins is an RHS-repeat-containing protein encapsulation device. *Nature* **501**, 547–550 (2013).
269. Jackson, V. A. *et al.* Structures of Teneurin adhesion receptors reveal an ancient fold for cell-cell interaction. *Nature Communications* **9**, 1–9 (2018).
270. Pei, T. T. *et al.* Intramolecular chaperone-mediated secretion of an Rhs effector toxin by a type VI secretion system. *Nature Communications* **11**, (2020).
271. Donato, S. L. *et al.* The  $\beta$ -encapsulation cage of rearrangement hotspot (Rhs) effectors is required for type VI secretion. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 33540–33548 (2020).
272. Tang, J. Y., Bullen, N. P., Ahmad, S. & Whitney, J. C. Diverse NADase effector families mediate interbacterial antagonism via the type VI secretion system. *Journal of Biological Chemistry* **293**, 1504–1514 (2018).
273. Gatsogiannis, C. *et al.* Tc toxin activation requires unfolding and refolding of a  $\beta$ -propeller. *Nature* **563**, 209–233 (2018).
274. del Toro, D. *et al.* Structural Basis of Teneurin-Latrophilin Interaction in Repulsive Guidance of Migrating Neurons. *Cell* **180**, 323-339.e19 (2020).
275. Renault, M. G. *et al.* The gp27-like Hub of VgrG Serves as Adaptor to Promote Hcp Tube Assembly. *Journal of Molecular Biology* **430**, 3143–3156 (2018).
276. Donato, S. L. *et al.* The  $\beta$ -encapsulation cage of rearrangement hotspot (Rhs) effectors is required for type VI secretion. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 33540–33548 (2020).
277. Ho, B. T., Fu, Y., Dong, T. G. & Mekalanos, J. J. *Vibrio cholerae* type 6 secretion system effector trafficking in target bacterial cells. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 9427–9432 (2017).
278. Vettiger, A. & Basler, M. Type VI Secretion System Substrates Are Transferred and Reused among Sister Cells. *Cell* **167**, 99-110.e12 (2016).
279. Ting, S. Y. *et al.* Bifunctional Immunity Proteins Protect Bacteria against FtsZ-Targeting ADP-Ribosylating Toxins. *Cell* **175**, 1380-1392.e14 (2018).
280. Ahmad, S. *et al.* An interbacterial toxin inhibits target cell growth by synthesizing (p)ppApp. *Nature* **575**, 674–678 (2019).

- 
281. Mok, B. Y. *et al.* A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. *Nature* **583**, 631–637 (2020).
282. Cianfanelli, F. R. *et al.* VgrG and PAAR Proteins Define Distinct Versions of a Functional Type VI Secretion System. *PLoS Pathogens* **12**, 1–27 (2016).
283. Murphy, J. R. Mechanism of diphtheria toxin catalytic domain delivery to the eukaryotic cell cytosol and the cellular factors that directly participate in the process. *Toxins* vol. 3 294–308 Preprint at <https://doi.org/10.3390/toxins3030294> (2011).
284. Roderer, D., Schubert, E., Sitsel, O. & Raunser, S. Towards the application of Tc toxins as a universal protein translocation system. *Nature Communications* **10**, 1–11 (2019).
285. Silverman, J. M. *et al.* Haemolysin Coregulated Protein Is an Exported Receptor and Chaperone of Type VI Secretion Substrates. *Molecular Cell* **51**, 584–593 (2013).
286. Ahmad, S. *et al.* Structural basis for effector transmembrane domain recognition by type VI secretion system chaperones. *eLife* **9**, 1–29 (2020).
287. Tang, J. Y., Bullen, N. P., Ahmad, S. & Whitney, J. C. Diverse NADase effector families mediate interbacterial antagonism via the type VI secretion system. *Journal of Biological Chemistry* **293**, 1504–1514 (2018).

## 11 Appendix

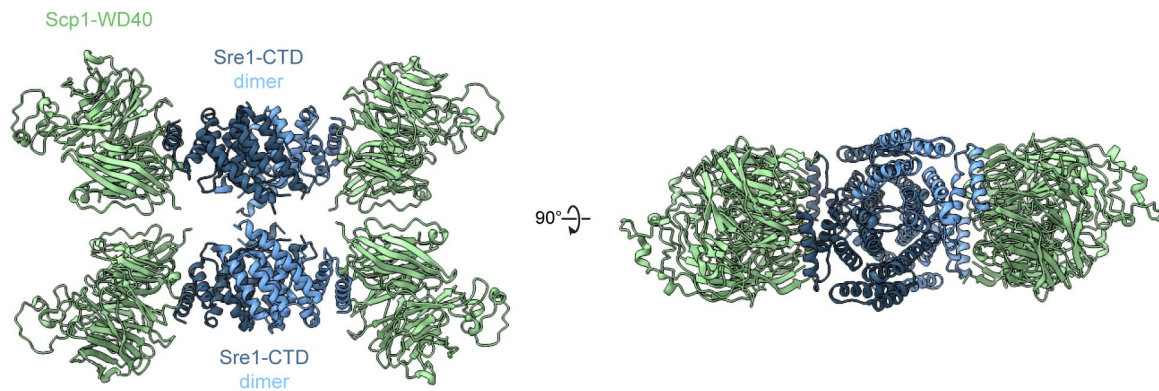


Figure S 1: Cryo-EM structure of the fission yeast Sre1-CTD and Scp1-WD40 complex.

The C-terminal domains (CTD) of the *S. pombe* SREBP and Scap homologs Sre1 and Scp1 form a higher oligomeric complex. The CTDs of Sre1 form dimer of dimers each decorated with one WD40 domain of Scp1. (PDB ID: 5GRS).

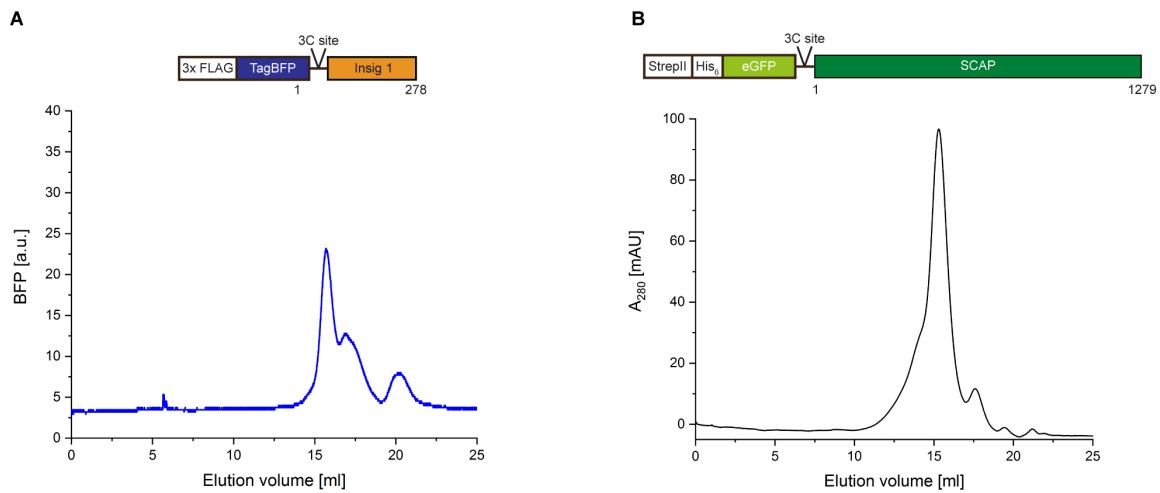


Figure S 2: Size-exclusion runs of purified (A) Insig and (B) Scap. The constructs are indicated above the SEC profile. The chromatogram of Scap was kindly provided by Birte Weyers.

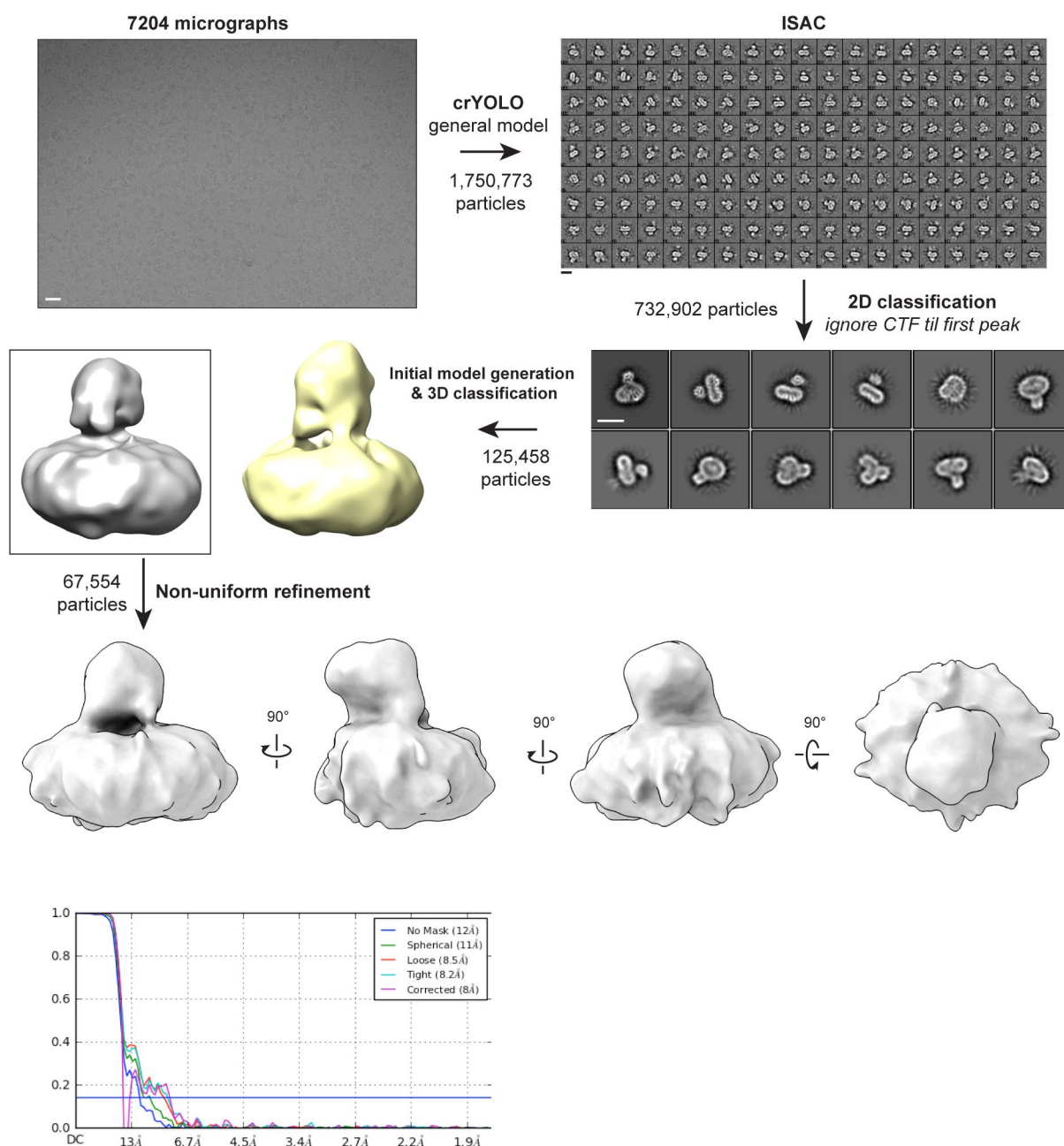


Figure S 3: Image processing workflow for the Scap-Insig complex in LMNG-CHS.

7204 good micrographs were recorded. Particle coordinates were selected using the general model of crYOLO. Subsequently, particles were extracted and cleaned in 2D by ISAC integrated in SPHIRE. Particles belonging to classes showing clear protein features were selected and transferred to Relion. Further 2D classification was used to separate out bad particles. Classes showing high-resolution features were selected for *Ab initio* model generation and 3D classification against two references. The class containing more features inside the micelle was selected for a final non-uniform refinement.

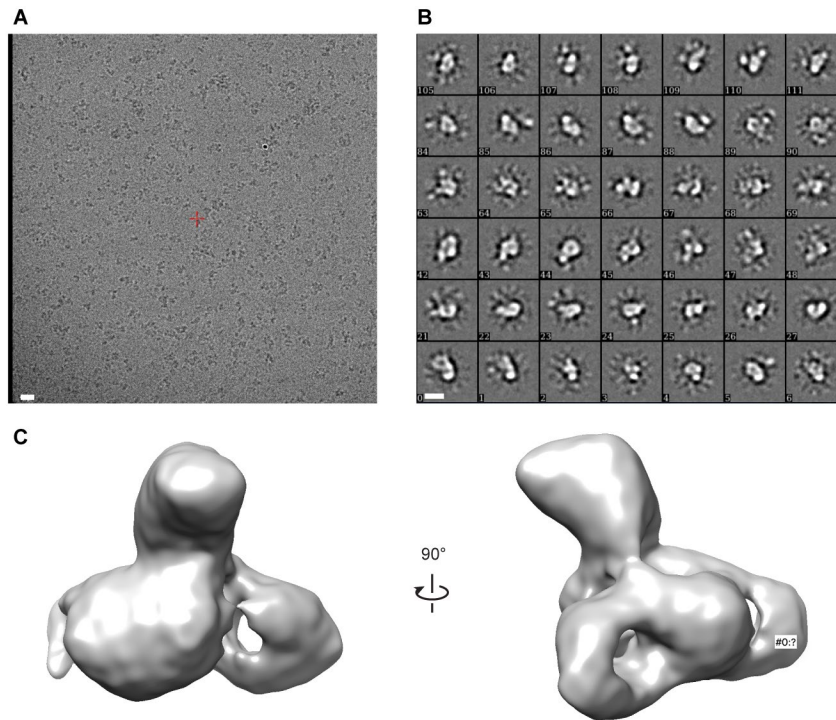


Figure S 4: Cryo-EM of the Scap-Insig complex in Amphipols.

(A) Representative micrograph of the Scap-Insig complex reconstituted in Amphipols revealing a heterogeneous particle distribution. Scale bar, 20 nm. (B) 2D class averages of the complex in Amphipols. Scale bar, 12 nm. (C) The reconstruction shows a fragmented density of the Amphipol belt covering the TMDs of both proteins.

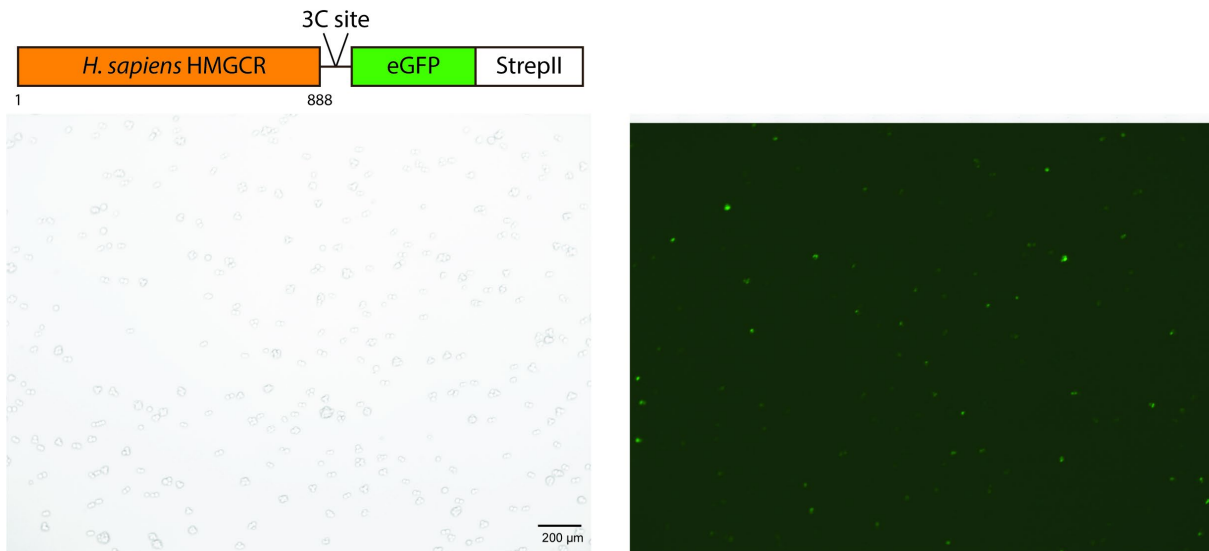


Figure S 5: Overexpression of the human HMGCR in HEK293 GnTI<sup>-</sup> cells.

The cells were harvested 48h post-infection and examined for green fluorescence indicated overexpression. Scale bar, 200 μm.

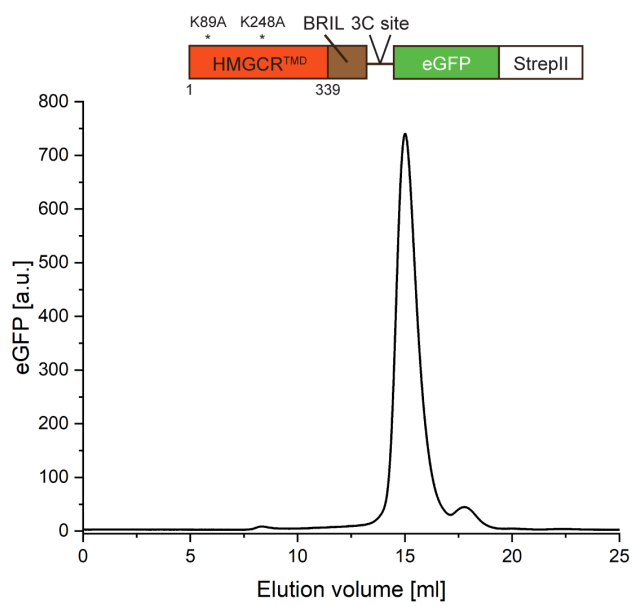


Figure S 6: FSEC analysis of the HMGCR<sup>TMD</sup>-BRIL construct.

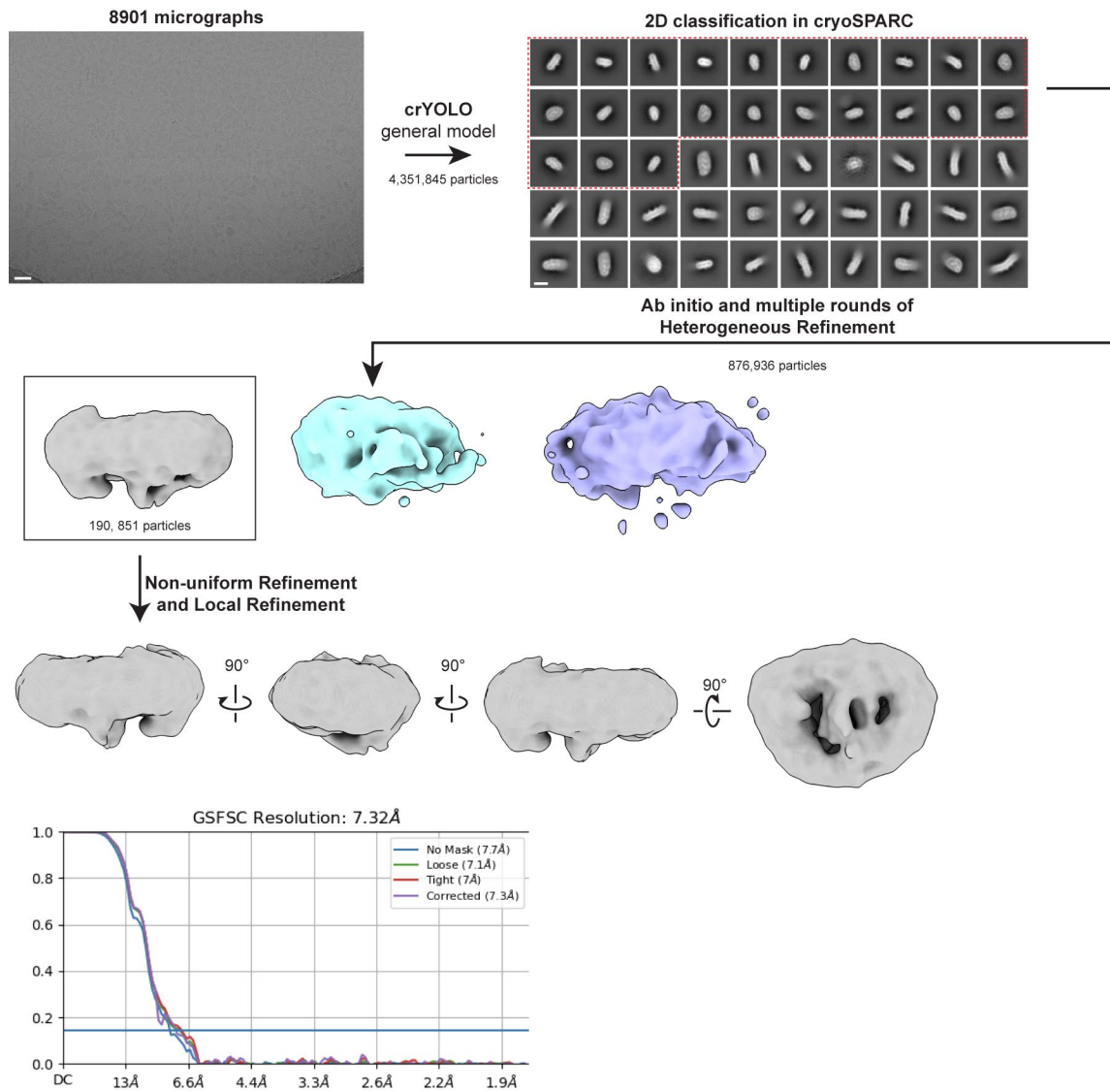


Figure S 7: Image processing workflow for the HMGCR<sup>TMD</sup>-BRIL-UBIAD1<sup>G186R</sup> complex in LMNG-CHS. 8901 micrographs were recorded. Particle coordinates were selected using the general model of crYOLO. Subsequently, particles were extracted and cleaned by 2D classification in cryoSPARC. Classes encircled in the red dotted box showing high-resolution features were selected for *Ab initio* model generation and subsequent heterogeneous refinement. Finally, a particle subset of 190K particles was refined to a resolution of 7.3Å using non-uniform and local refinement.



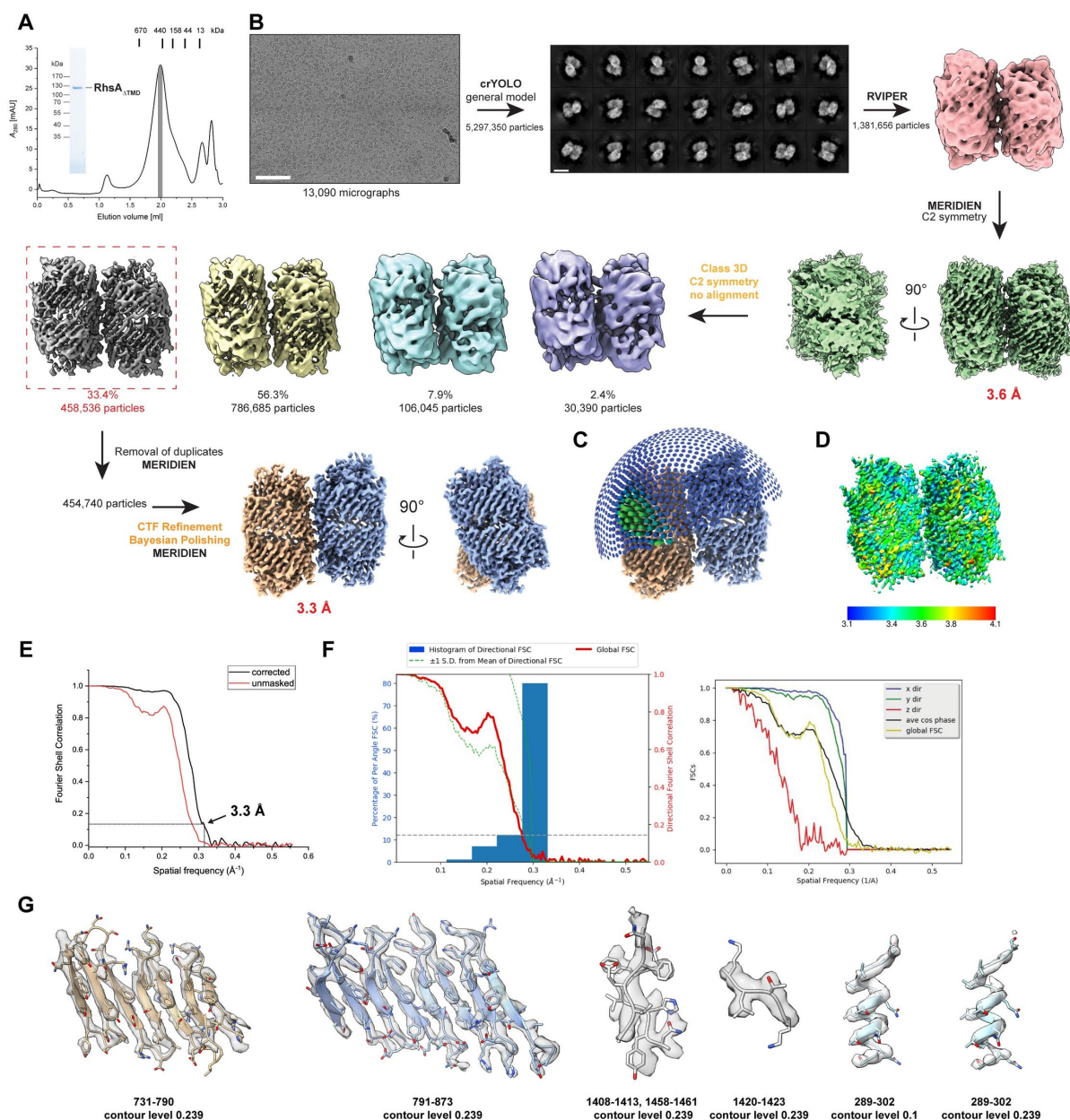


Figure S 8: Cryo-EM processing workflow used to obtain the structure of RhsA<sub>ATMD</sub>.

(A) Purification of RhsA<sub>ATMD</sub> via size-exclusion chromatography using a Superose 6 5/150 increase column. Molecular weight standards are indicated at their respective elution volumes. The grey bar denotes pooled and concentrated fractions used for cryo-EM analysis. The same material was analyzed for purity via semi-denaturing SDS-PAGE imaged with a stain-free filter. (B) Representative cryo-EM micrograph of RhsA<sub>ATMD</sub> used for structural determination. Scale bar EM micrograph, 100 nm. Particles were picked with the general model of crYOLO. The rest of the processing workflow is indicated and summarized in the methods section. Scale bar 2D class averages, 10 nm. Used software packages are highlighted. Orange font depicts steps carried out in Relion. The final map was calculated using MERIDIEN and postprocessed with DeepEMhancer. (C) Angular distribution plot of the final reconstruction. (D) Local resolution estimates visualized on a map postprocessed in SPHIRE. (E) Fourier shell correlation plot calculated from two independently processed maps. Resolution estimation is reported

at the gold standard cutoff of 0.143. (F) Resolution anisotropy was assessed with the 3DFSC online server tool. (G) Selected regions of the map are shown as transparent surface with the built atomic models as stick representations. This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

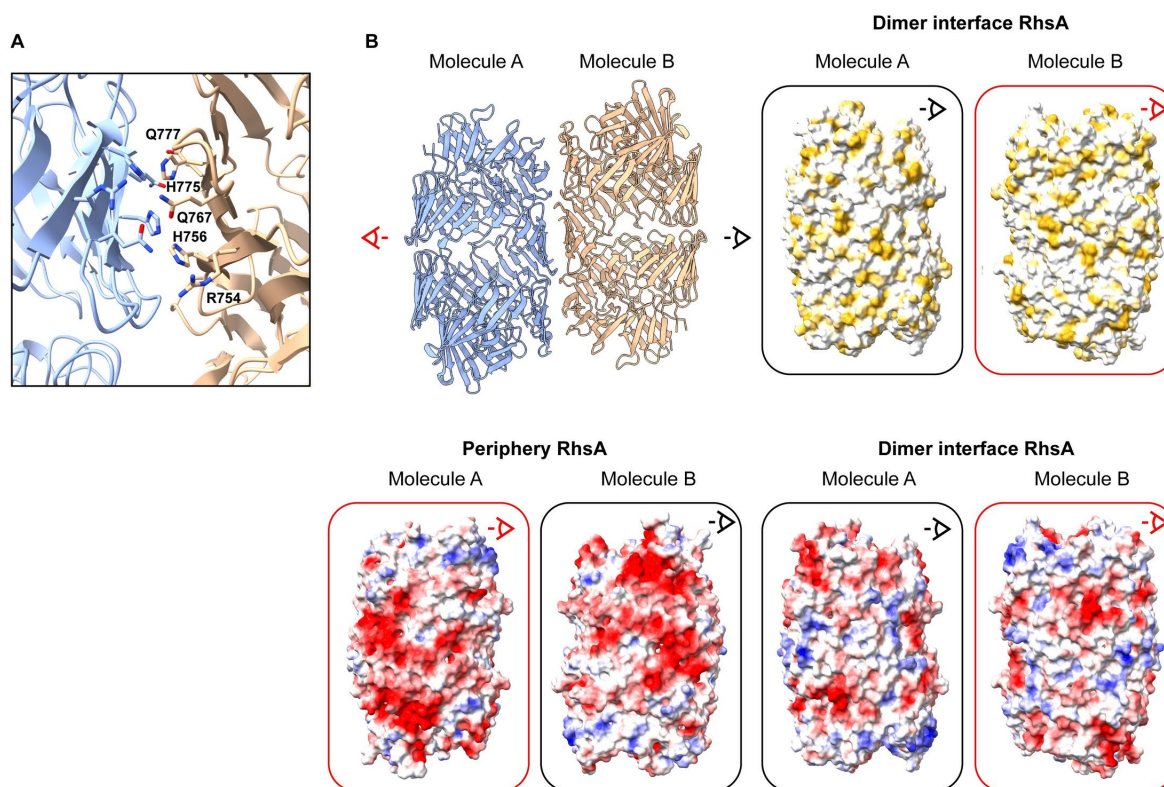


Figure S 9: Structural analysis of the dimer interface of RhsA.

(A) Close up view of the dimer interface of the two RhsA molecules at the central symmetry axis. Potential candidate residues engaging in stabilizing interactions are labelled and shown in stick representation. (B) Surface properties of the dimer interface and the periphery of RhsA. The surface of the interface is colored according to its Coulomb potential indicating positively (red, -20) and negatively (blue, +20) charged areas. The second representation shows the same interface but colored according to hydrophobicity. Ochre indicates hydrophobic and white indicates hydrophilic regions. The area facing the outer peripheries of both barrels shows that these regions would electrostatically repel each other and explain why only two barrels can interact at the same time. This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.





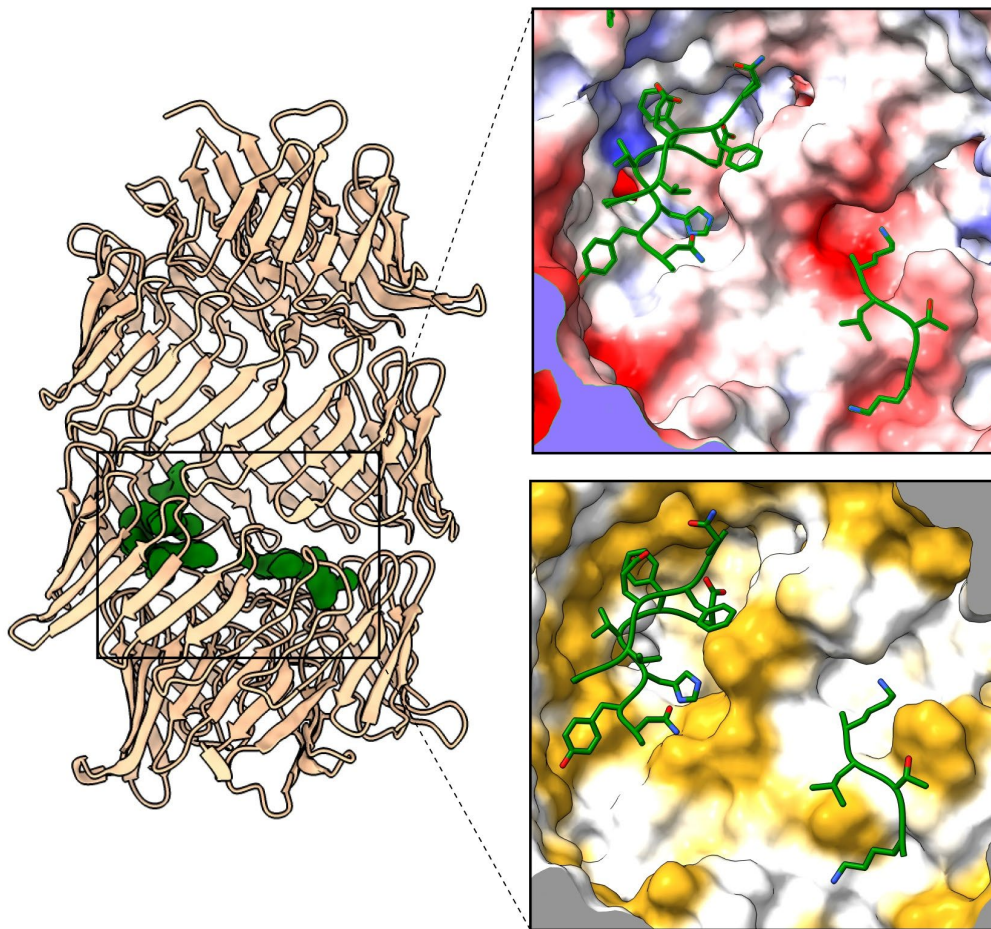


Figure S 11: Local environment of the ordered toxin fragments inside the Rhs cocoon.

Interaction of three  $\beta$ -strands belonging to the toxin domain of RhsA with the Rhs barrel. The toxin fragments are stabilized by the interaction with hydrophobic and hydrophilic surfaces inside the Rhs barrel. The molecular surface is colored according to hydrophobicity with ochre and white indicating hydrophobic and hydrophilic regions, respectively. The electrostatic representation is colored according to Coulomb potential, which depicts positively (red, -20) and negatively (blue, +20) charged areas. This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

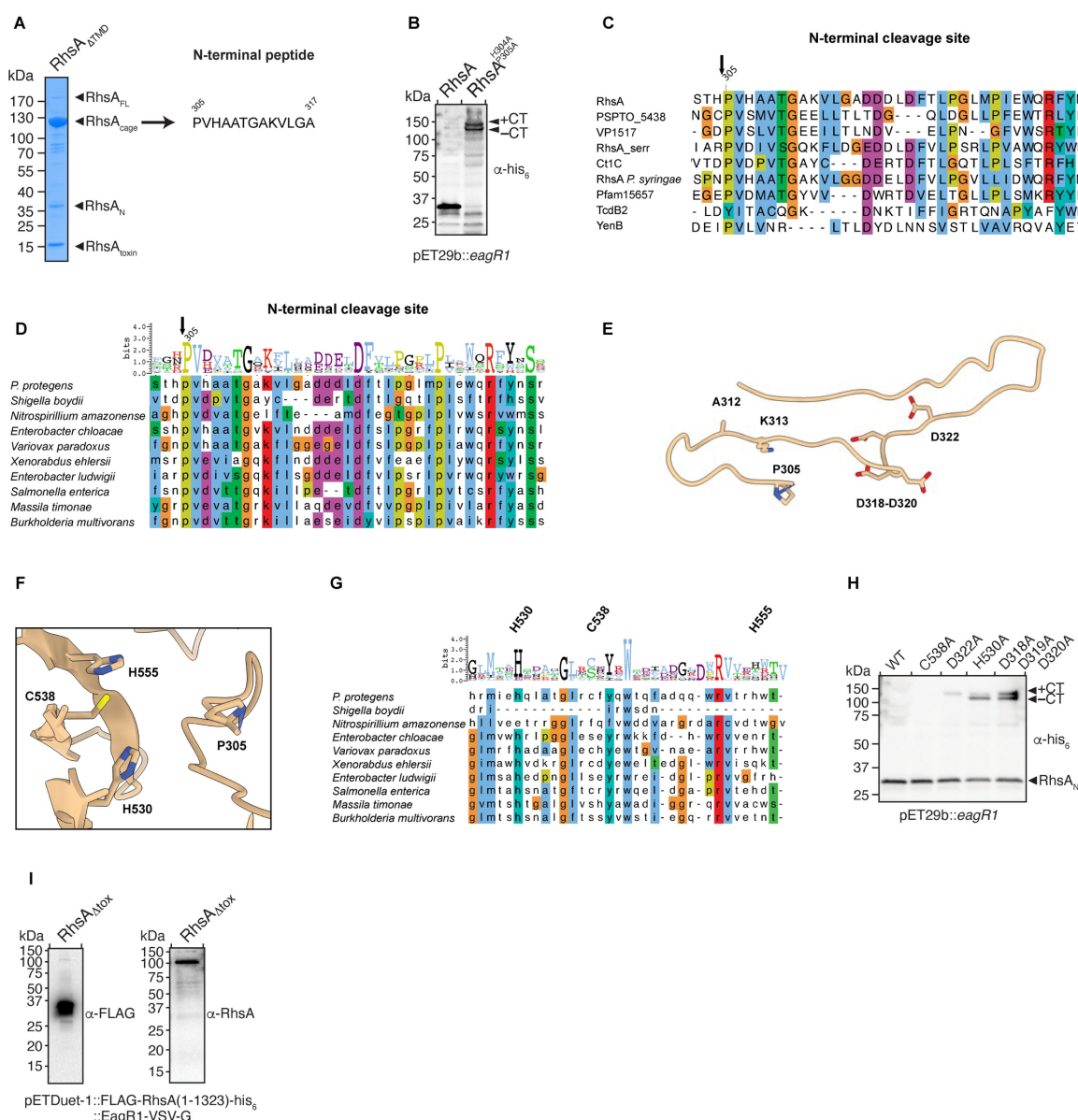


Figure S 12: Characterization of the N-terminal autoproteolysis site of RhsA.

(A) The N-terminal cleavage site was identified by protein sequencing. The band containing the Rhs<sub>cage</sub> was excised and analyzed by *de novo* protein sequencing using LC-MS/MS. (B) Western blot analysis of wild-type or a N-terminal cleavage resistant mutant (H304A/P305A) confirming the location of the cleavage site at P305. (C) Multiple sequence alignment of the N-terminal cleavage site. The cleavage site is indicated by a black arrow. Coloring is according to the ClustalW color code. (D) Sequence alignment of the T6SS class I prePAAR effectors highlighting the N-terminal cleavage site. Residue conservation is depicted as Weblogo. The Weblogo represents residue conservation. (E) Cartoon representation of the N-terminal cleavage site in RhsA. Residues A312 and K313 as well as conserved and potential catalytically active residues are shown in stick representation. Residues A312 and K313 correspond to the catalytically active glutamates in TseI<sup>242</sup>. (F) Hypothesized cysteine protease motif which is near the N-terminal cleavage site P305. (G) Multiple sequence alignment highlighting the conservation of the hypothesized cysteine protease motif among class I prePAAR effectors. The Weblogo represents residue conservation. (H) Western blot analysis of potential residues involved in N-terminal cleavage and generation of the cleavage product RhsA<sub>N</sub>. Impaired cleavage was assessed by appearance of the full-length

RhsA chain which is a mixture of both species, the C-terminally cleaved fragment (-CT) and the C-terminally uncleaved fragment (+CT). RhsA was coexpressed with its cognate chaperone EagR1. The blot was performed against N-terminal His<sub>6</sub>-tagged proteins ( $\alpha$ -His). (I) Western blot analysis to examine the potential involvement of RhsA's C-terminal toxin domain in the autoproteolysis of its N-terminal region. An RhsA truncation lacking the C-terminal toxin domain (RhsA $\Delta$ <sub>tox</sub>, residues 1-1323) was expressed and examined for N-terminal cleavage using antibodies that recognize its N-terminus ( $\alpha$ -FLAG) or its Rhs cocoon ( $\alpha$ -RhsA). This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

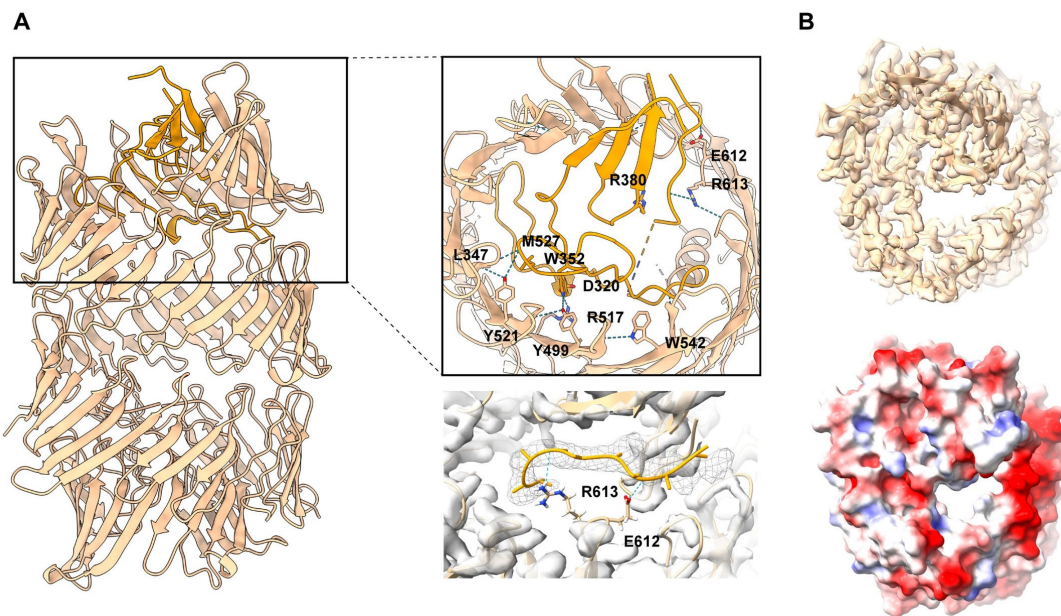


Figure S 13: Hydrophilic interactions of the plug domain with the Rhs barrel.

(A) Only a few hydrophilic interactions stabilize the plug domain. Residues participating in hydrogen bonds are labeled and shown in stick representation. The cork and the seal are colored in orange whereas the Rhs repeats is colored in beige. (B) The barrel would not be closed without the observed density corresponding to the seal. The model for the seal was manually removed to visualize the opening though which toxin is threaded into the target cell after removal of the seal and the anchor helix. This figure reprinted from ref.<sup>200</sup> under the license CC BY 4.0<sup>237</sup>.



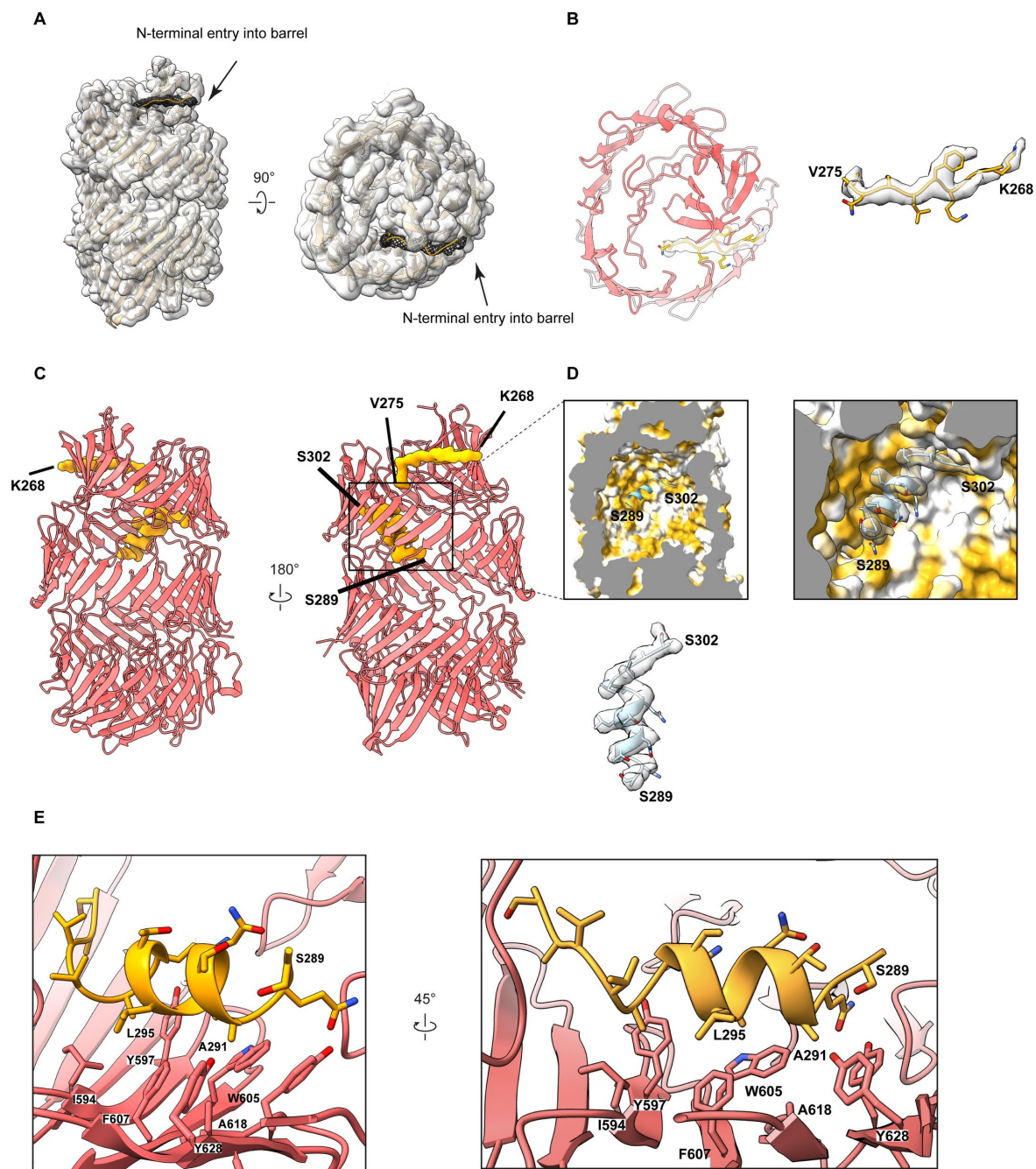


Figure S 14: Stabilization of the N-terminal seal and anchor helix of RhsA.

(A-B) A small density, corresponding to seal (residues K268-V275) of RhsA enters the barrel from the top (mesh). This results in a complete sealing of the cocoon. (C-E) The seal leads further down into an amphipathic helix which strongly interacts with the inner surface of the cocoon and thus serves as an anchor point for the N-terminal domain. The amphipathic helix is stabilized by hydrophobic interactions with Rhs repeats. This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

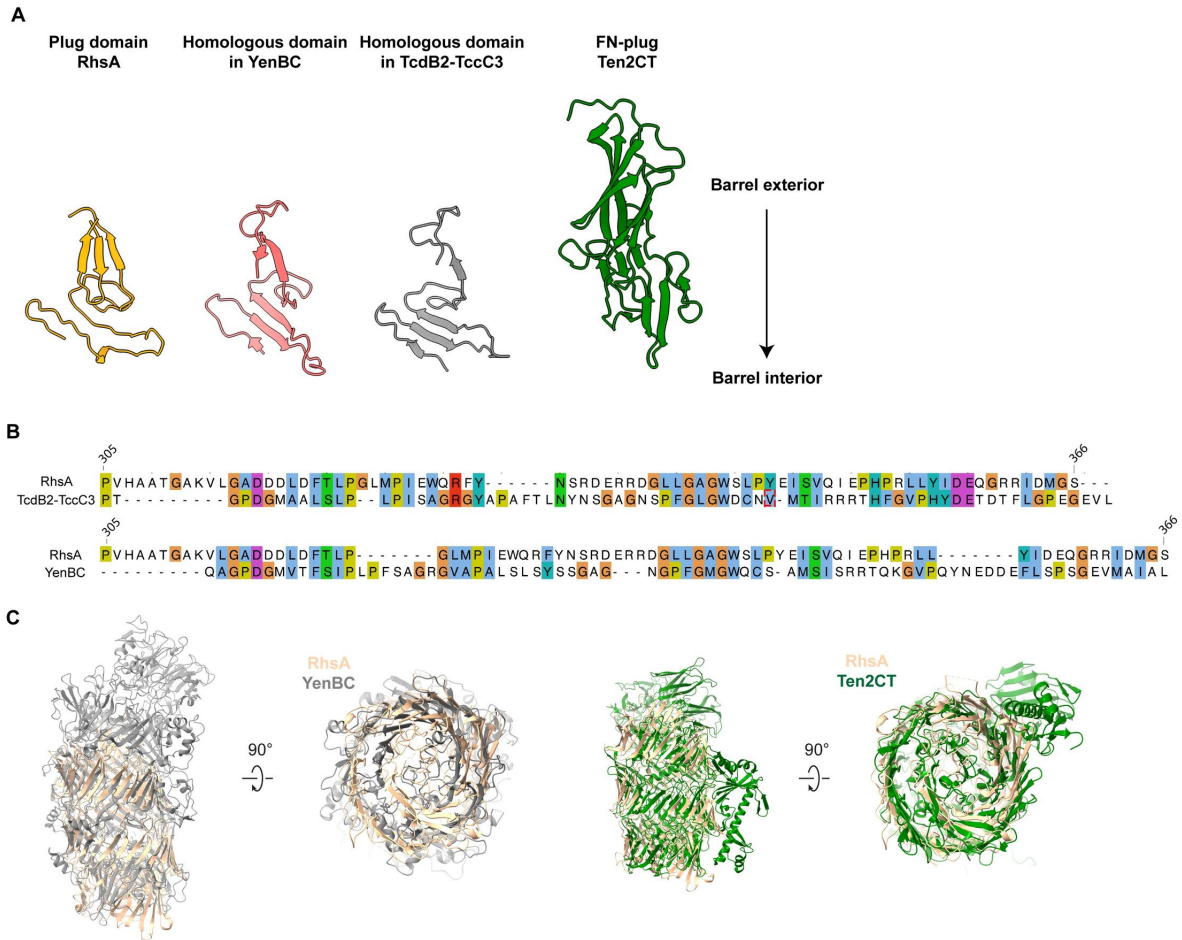


Figure S 15: Comparison of Rhs repeat containing proteins of known structure.

(A) Comparison of the plug domains of RhsA (orange) with YenBC (PDB ID: 4IGL), TcdB2-TccC3 (PDB ID: 6H6G) and Ten2CT (PDB ID: 6FB3). (B) Sequence alignment of the plug domain of RhsA with the sequences of the homologous domains found in TcdB2-TccC3 (top) and YenBC (bottom) are shown. Residues are colored according to the ClustalW color code. (C) Structural overlay of RhsA with YenBC (left) and Ten2CT (right). This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.



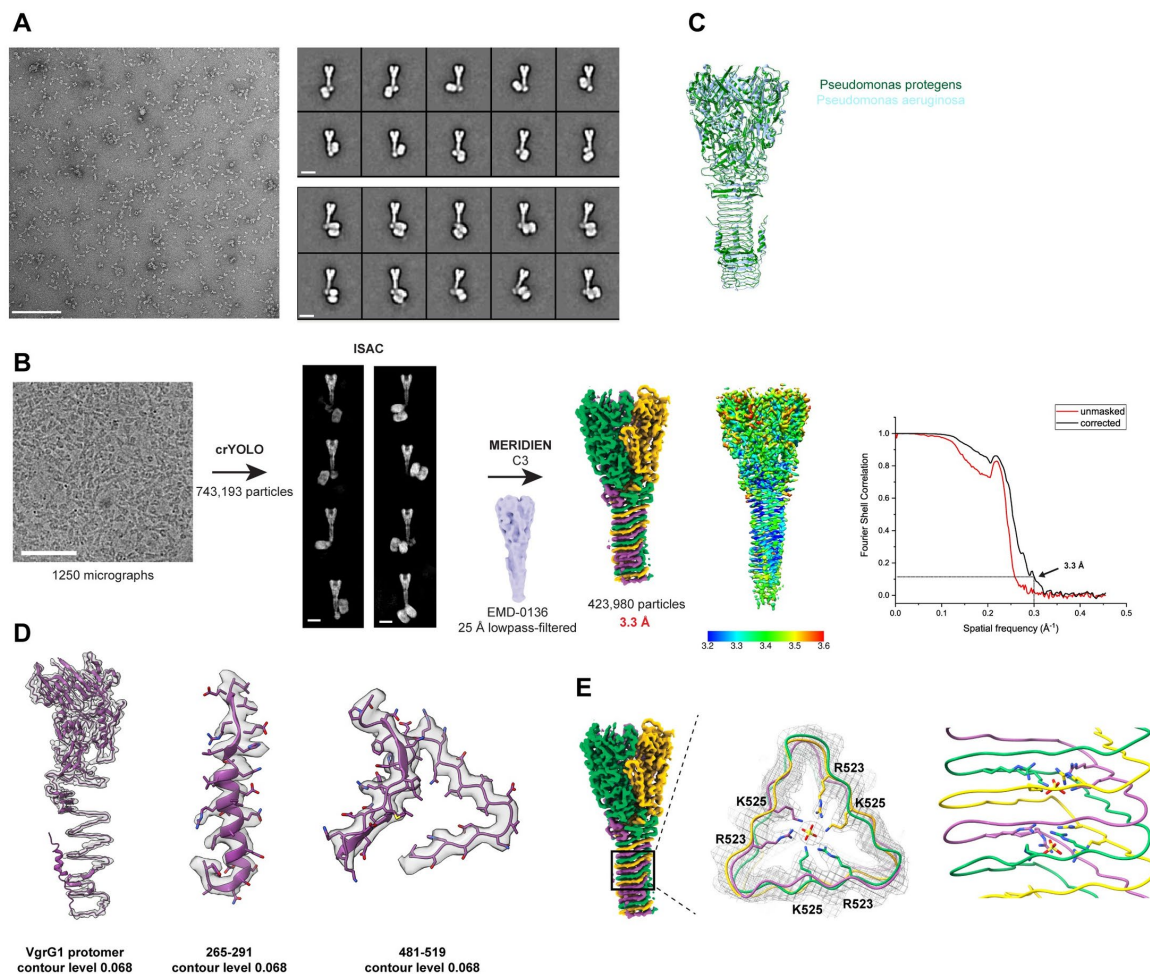


Figure S 16: Cryo-EM of the PFC.

(A) Representative negative stain micrograph and 2D class averages of the intact PFC consisting of VgrG1, RhsA and EagR1. Scale bar micrograph, 100 nm. Scale bar 2D class averages, 10 nm. (B) Cryo-EM micrograph of the PFC and image processing workflow. Particles were picked with the general model of crYOLO and classified with ISAC. Refinement in MERIDIEN led to a reconstruction of 3.3 Å (FSC=0.143 criterion). The final reconstruction was either colored individually for each protomer or according to local resolution. Fourier shell correlation was plotted according to two independent refined maps. Dashed line indicates the gold standard FSC criterion of 0.143. Scale bar micrograph, 100 nm. Scale bar 2D class averages, 10 nm. (C) Overlay of VgrG1 from *P. protegens* (green) with VgrG1 from *P. aeruginosa* (blue, PDB ID: 6H3N). (D) Map quality of selected parts of the structure (transparent surface) with built atomic models (cartoon representation). (E) Potential ion binding site in the middle part of the  $\beta$ -prism of *P. protegens* VgrG1. Sulfate ions were modelled into the spherical densities. Coordinating residues are displayed in stick representation. This figure reprinted from ref.<sup>198</sup> under the license CC BY 4.0<sup>265</sup>.

Table S 1: Bacterial strains used for results related to RhsA.

Organism	Genotype	Description	Reference
<i>P. protegens</i> Pf-5	wild-type		Paulsen et al., Nature Biotechnology, 2005 <sup>257</sup>
	PFL_6096_D1324N	Expresses RhsA D1324N	This study
	PFL_6096_D1346N	Expresses RhsA D1346N	This study
	$\Delta$ PFL_6079 attB::lacZ, TetR	<i>pppA</i> deletion, constitutive <i>lacZ</i> expression, Tet <sup>R</sup>	Ahmad et al., eLife, 2020 <sup>152</sup>
	$\Delta$ PFL_6079 $\Delta$ PFL_6096 $\Delta$ PFL_6097 attB::lacZ, TetR	<i>pppA rhsA rhsI</i> deletion, constitutive <i>lacZ</i> expression, Tet <sup>R</sup>	Ahmad et al., eLife, 2020 <sup>152</sup>
<i>E. coli</i> SM10 $\lambda$ pir	<i>thi thr leu tonA lac Y supE recA::RP4-2-Tc::Mu</i>	Conjugation strain	BioMedal LifeScience
<i>E. coli</i> XL-1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacI<sup>R</sup> Z<math>\Delta</math>M15 Tn10</i> (Tet <sup>R</sup> )]	Cloning strain	Novagen
<i>E. coli</i> BL21 (DE3) CodonPlus	F' <i>ompT gal dcm lon hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) <math>\lambda</math>(DE3) pLysS(cm<sup>R</sup>)</i>	Protein expression strain	Novagen

Table S 2: Plasmids used for RhsA-specific results.

Plasmid	Relevant features	Reference
pETDuet-1	Co-expression vector with <i>lacI</i> , T7 promoter, N-terminal His <sub>6</sub> tag in MCS-1, Amp <sup>R</sup>	Novagen
pET29b	Expression vector with <i>lacI</i> , T7 promoter, C-terminal His <sub>6</sub> tag, Kan <sup>R</sup>	Novagen
pEXG2	Allelic exchange vector containing <i>sacB</i> , Gm <sup>R</sup>	Hmelo et al., Nature Protocols, 2015 <sup>199</sup>
pScRhaB2-CV	Expression vector with <i>PrhaB</i> , Tmp <sup>R</sup>	Cardona et al., Plasmid, 2005 <sup>201</sup>
pPSV39-CV	Expression vector with <i>lacI</i> , <i>lacUV5</i> promoter, C-terminal VSV-G tag, Gm <sup>R</sup>	Silverman, Cell, 2013 <sup>285</sup>
pETDuet-1::PFL_6096 ::PFL_6097	Co-expression vector for RhsA and RhsI	Ahmad et al., eLife, 2020 <sup>286</sup>
pETDuet-1::PFL_6096_74-CT ::PFL_6097	Co-expression vector for RhsA <sub>ANT</sub> and RhsI	Ahmad et al., eLife, 2020 <sup>286</sup>
pET29b::PFL_6094	Expression vector for VgrG1	This study

pETDuet-1::PFL_6096_74-CT_VSV-G ::PFL_6097	Co-expression vector for C-terminally VSV-G tagged RhsA <sub>ANT</sub> and RhsI	This study
pETDuet-1::PFL_6096_74-CT_D1324N_VSV-G ::PFL_6097	Co-expression vector for C-terminally VSV-G tagged RhsA <sub>ANT</sub> D1324N and RhsI	This study
pETDuet-1::PFL_6096_74-CT_D1346N_VSV-G ::PFL_6097	Co-expression vector for C-terminally VSV-G tagged RhsA <sub>ANT</sub> D1346N and RhsI	This study
pETDuet-1::PFL_6096_C538A ::PFL_6097	Co-expression vector for RhsA <sub>C538A</sub> and RhsI	This study
pETDuet-1::PFL_6096_D322A ::PFL_6097	Co-expression vector for RhsA <sub>D322A</sub> and RhsI	This study
pETDuet-1::PFL_6096_H530A ::PFL_6097	Co-expression vector for RhsA <sub>H530A</sub> and RhsI	This study
pETDuet-1::PFL_6096_D318A_D319A_D320A ::PFL_6097	Co-expression vector for RhsA <sub>D318A, D319A, D320</sub> and RhsI	This study
pETDuet-1::PFL_6096_H304A_P305A ::PFL_6097	Co-expression vector for RhsA <sub>H304A, P305A</sub> and RhsI	This study
pETDuet-1::PFL_6096_290-CT ::PFL_6097	Co-expression vector for RhsA <sub>AN</sub> and RhsI	This study
pETDuet-1::PFL_6096_1-1323 ::PFL_6097	Co-expression vector for RhsA <sub>Atox</sub> and RhsI	This study
pET29b::PFL_6095	Expression vector for EagR1	Ahmad et al., eLife, 2020 <sup>150</sup>
pEXG2::PFL_6096_D1324N	Allelic exchange construct for generating <i>rhsA</i> _D1324N	This study
pEXG2::PFL_6096_D1346N	Allelic exchange construct for generating <i>rhsA</i> _D1346N	This study
pSCRhaB2-CV:: PFL_6096_74-CT	Expression vector for RhsA <sub>ATMD</sub>	This study
pSCRhaB2-CV:: PFL_6096_74-CT_D1324N	Expression vector for RhsA <sub>ATMD</sub> D1324N	This study
pSCRhaB2-CV:: PFL_6096_74-CT_D1346N	Expression vector for RhsA <sub>ATMD</sub> D1346N	This study
pPSV39-CV::PFL_6097	Expression vector for RhsI	Tang et al, JBC, 2018 <sup>287</sup>

## 12 Acknowledgements

First of all, I would like to thank my supervisor, Prof. Dr. Stefan Raunser, for giving me the opportunity to work on these amazing projects as a master's student and for trusting me to continue as a PhD student in his laboratory. Throughout my PhD, he has continuously supported me and given me the opportunity to develop my personal and scientific skills. He is a great mentor and has always supported me to be an independent scientist, for which I am very grateful.

Furthermore, I would like to thank Prof. Dr. Daniel Summerer for willingness to be my second referee of my thesis. Moreover, for being part of my TAC committee and his valuable suggestions and the support throughout my studies.

Many thanks to my colleagues in our vibrant research department. I have had the opportunity to meet so many people over the years for which I am grateful. Barathy introduced me to various membrane protein-related workflows in the lab, where he was a great support, especially in getting started in the lab, but also during our epic badminton sessions. Tobias is not only a good “Wander-Tobi” but also a great scientist and good friend. He always had good ideas to push the boundaries of my projects and was an emotional support. Thanks also to Wouter for useful suggestions on my projects and funny moments outside the lab. A big thank you also goes to Marion for the technical support in growing large amounts of HEK cultures. I would also like to thank all past and present members of the membrane protein group for their support and discussions at all the membrane protein meetings.

I would also like to thank my office mates Pascal and Birte for their support, but also for the many funny moments and their friendship. I would also like to thank my current office mates Astrid, my dearest Italian friend Adriana, and dearest American friend Gavin, for the funny and nice environment and the support over the years.

I would also like to thank Christos Gatsogiannis for always having an open door to help with any questions related to processing single particles.

I would also like to thank the entire toxin group for their critical questions and ideas on how to tackle our toxins related projects. Also, thank you to Karin Vogel-Bachmayr and Nathalie Bleimling for providing technical support.

A big thanks to Daniel Prumbaum and Oliver Hofnagel for their continuous and excellent EM support even at late hours or during the weekend. And of course for the fun during our badminton sessions.

I would also like to express my gratitude to Lucia Sironi and Christa Hornemann for their constant support throughout the whole of our IMPRS program.

For the fruitful discussions and successful collaboration, I would like to thank my collaborators Shehryar Ahmad, Kartik Sachar and John C. Whitney.

On a personal note, I would like to thank my friends, family and girlfriend for their constant help and emotional support during my PhD and studies. I have always been able to rely on them, for which I am very grateful.

## 13 Eidesstaatliche Versicherung

### Eidesstattliche Versicherung (Affidavit)

Günther, Patrick

Name, Vorname  
(Surname, first name)

166642

Matrikel-Nr.  
(Enrolment number)

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 oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

Ort, Datum  
(Place, date)

Unterschrift  
(Signature)

Titel der Dissertation:  
(Title of the thesis):

---

### Structural Investigation of Cholesterol Homeostasis and Bacterial Toxins

---

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.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.\*

\*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

Ort, Datum  
(Place, date)

Unterschrift  
(Signature)