

Dissertation for the award of the degree
“Doctor rerum naturalium”
of the Georg-August-Universität Göttingen

Crystal Structure & Characterization of the SCOC Coiled Coil Domain

submitted by

Caroline Anna Julie Behrens

from Hamburg–Altona

within the doctoral program

Biomolecules: Structure – Function – Dynamics
of the Georg-August University School of Science (GAUSS)



Göttingen 2013

Members of the Thesis Committee

Dr. Karin Kühnel (1st Reviewer)

Department of Neurobiology, Max Planck Institute for Biophysical Chemistry

Prof. Dr. Michael Thumm (2nd Reviewer)

Department of Biochemistry II, University of Göttingen

Prof. Dr. Markus Zweckstetter

*Research Group Structure Determination of Proteins Using NMR,
Max Planck Institute for Biophysical Chemistry*

Prof. Dr. Marina Rodnina

Department of Physical Biochemistry, Max Planck Institute for Biophysical Chemistry

Date of the oral examination: August 7th, 2013

Affidavit

Herewith I declare that I have written this Dissertation independently and without any sources or aids other than quoted.

Göttingen, June 14th, 2013

(Caroline Anna Julie Behrens)

Abstract

Autophagy is a conserved eukaryotic process for the degradation of cytosolic content. An expanding double layered membrane takes up part of the cytoplasm and matures to an autophagosomal vesicle. The autophagosome fuses with the lysosome, where its content is degraded. The short coiled coil protein (SCOC) forms a complex with fasciculation and elongation protein zeta 1 (FEZ1). Their complex is involved in autophagy regulation. Here, I present the crystal structure of the functional important human SCOC coiled coil domain (residues 78–159), which was determined at 2.8 Å resolution. SCOC forms a parallel left handed coiled coil dimer. I observed two distinct dimers in the crystal structure with a 3-molecule asymmetric unit, showing that SCOC is conformational flexible. Importantly, several of the canonical hydrophobic a/d-heptad core positions are occupied by well conserved, polar and charged residues. I assumed that these non-canonical residues might have a strong effect on the stability and oligomerization state of SCOC. The influence of the residues was investigated by characterising mutant SCOC constructs. The double core mutations E93V/K97L and N125L/N132V led to a change from dimer to either trimer or tetramer formation and the thermostabilities of these mutants were dramatically increased. Moreover, I present the formation of a stable homogeneous complex of SCOC with the coiled coil domain of FEZ1. A minimal FEZ1 region (residues 227–290) was sufficient for interaction with SCOC. Complex formation with the SCOC double core mutants was impaired, showing that dimerization of SCOC is essential for interaction with FEZ1. In addition I identified SCOC surface residue R117 as important for binding.

The human pathogen *Shigella flexneri* escapes autophagy by a sophisticated mechanism involving the *Shigella flexneri* proteins VirG and IcsB–IpgA. VirG is an autotransporter protein, consisting of a signal sequence, a passenger domain (residues 52–758) and a transmembrane domain. I established a purification protocol for the VirG (52–758) under denaturing conditions. First crystallization trials of refolded VirG passenger domain yielded spherulites and microcrystalline structures. Initial studies of IcsB–IpgA indicate, that IcsB contains flexible and presumably unfolded regions, while its chaperone IpgA is well structured and stable.

Acknowledgements

First of all, I would like to thank my supervisor Dr. Karin Kühnel for her advice and support during my thesis. Her door was always open for discussion and I am very thankful for her guidance into the field of protein crystallography. Furthermore, I would like to thank Prof. Dr. Reinhard Jahn for financial support during the first year of my PhD, for the infrastructure in his department that allowed me to pursue my research, and for general scientific and personal advice. Next, I would like to express my gratitude towards Dr. John Chua, Beyenech Binotti and Brigitte Barg-Kues for the tremendous efforts they made during our collaboration on the SCOC–FEZ1 interaction. Dagmar Diezmann and Michaela Hellwig have provided crucial and excellent technical support, which I appreciated a lot.

I am very thankful to all those who have helped me with several methods that I conducted during the course of my thesis. My thanks goes to Dr. Matias Hernandez for teaching me how to use the SEC-MALLS equipment. I am very grateful for help with nucleotide analysis that was made possible by Dr. Corinna Pohl and Prof. Dr. Marina Rodnina. My gratitude goes to Prof. Dr. Markus Zweckstetter and Dr. Stefan Bibow for two-dimensional NMR analysis. Dr. Danilo Meyer and Prof. Dr. Kai Tittmann provided advice and equipment for the ITC measurements, for which I am very thankful. A special thanks goes to Dr. Vlad Pena and Dr. Ashwin Chari for the extensive discussions and helpful advice on crystallography methods and X-ray data collection. I would also like to express my gratitude towards the staff of beamline X10SA at the SLS.

Thanks a lot to all my friends and colleagues in the Jahn department for a nice and supportive working atmosphere, and to Frieder for routine and quick daily support with lab equipment and the coffee machine. Janina and Saskia, working with you was a professional and personal gain. Andreea, thanks for being the most helpful bench partner I can imagine. Thanks to Birk for constant coffee delivery during the writing process!

Last but not least I would like to thank GGNB for the financial support during my thesis, and moreover also for the many possibilities I was able to experience by learning new methods, skills and friends in graduate school.

Finally, I would like to thank my parents and my siblings for just being as awesome as you are! Our family support system has been the greatest gift through the last year(s). Thank you, Andreas—not only for Linux and L^AT_EX assistance—but for everything else.

Contents Overview

1	Introduction	1
1.1	General introduction to Autophagy	1
1.2	SCOC and its interaction partners	7
1.3	<i>Shigella flexneri</i> and Autophagy	11
2	Materials and Methods	17
2.1	Materials	17
2.2	Methods	32
3	SCOC and its interaction partners	57
3.1	Characterization & structure determination of the SCOC coiled coil domain	57
3.2	Interaction of SCOC's coiled coil domain with Arl1 and FEZ1	99
3.3	Discussion	110
3.4	Outlook	115
4	<i>Shigella flexneri</i> and Autophagy	117
4.1	Characterization of VirG	117
4.2	Characterization of IcsB–IpgA	124
4.3	Discussion and Outlook	130
A	Appendix	133
A.1	DNA sequences of synthetic genes	133
A.2	Crystallization conditions of SCOC ccd constructs	135
A.3	Supplementary Figures	143
A.4	Computational Methods	145
A.5	Mass spectrometry results	160
	Bibliography	163

Contents

1	Introduction	1
1.1	General introduction to Autophagy	1
1.1.1	Subtypes of Autophagy	1
1.1.2	The autophagic core machinery	4
1.2	SCOC and its interaction partners	7
1.2.1	SCOC: an Arl1 effector	7
1.2.2	SCOC interaction with Arl1	7
1.2.3	SCOC interaction with FEZ1	8
1.2.4	Coiled coil domains: features, stability and oligomerization state	10
1.2.5	Aims	10
1.3	Shigella flexneri and Autophagy	11
1.3.1	Selective Autophagy: Xenophagy	11
1.3.2	Escape of intracellular Shigella flexneri from autophagy	12
1.3.3	The autotransporter protein VirG	13
1.3.4	The secretion protein IcsB–IpgA	15
1.3.5	Aims	15
2	Materials and Methods	17
2.1	Materials	17
2.1.1	Enzymes	17
2.1.2	Kits	17
2.1.3	Columns for Chromatography	18
2.1.4	Antibodies	18
2.1.5	Buffers, media and antibiotics	19
2.1.6	DNA constructs	23
2.1.7	Oligonucleotides	28
2.2	Methods	32

-
- 2.2.1 Molecular cloning 32
 - 2.2.1.1 PCR 32
 - 2.2.1.2 DNA agarose gel electrophoresis and gel extraction purification 34
 - 2.2.1.3 Restriction digest 34
 - 2.2.1.4 Ligations 34
 - 2.2.1.5 Transformation 35
 - 2.2.1.6 Plasmid purification 35
 - 2.2.1.7 Mutagenesis 35
 - 2.2.2 Expression and Purification 36
 - 2.2.2.1 Expression analysis 36
 - 2.2.2.2 Minimal expression for selenomethione-labelling 37
 - 2.2.2.3 Minimal expression for ¹⁵N-labelling 37
 - 2.2.2.4 Expression and purification of Strep-tagged SCOC (78–159) domain and mutants 37
 - 2.2.2.5 Expression and purification of His-tagged SCOC constructs 38
 - 2.2.2.6 Expression and purification of GST-SCOC fusion proteins 40
 - 2.2.2.7 Expression and purification of Arl1 (15–181) Q79L 41
 - 2.2.2.8 Coexpression and copurification of SCOC–FEZ1 complexes 42
 - 2.2.2.9 Expression and denaturing purification of VirG (353–758) 43
 - 2.2.2.10 Expression and denaturing purification of VirG passenger domain (52–758) 45
 - 2.2.2.11 Expression and purification of IcsB–IpgA complexes 46
 - 2.2.3 Biochemical methods 47
 - 2.2.3.1 SDS-PAGE and Western blot 47
 - 2.2.3.2 Limited proteolysis 48
 - 2.2.3.3 N-terminal sequencing 48
 - 2.2.3.4 GST-Pulldown 49
 - 2.2.3.5 Nucleotide-dependent Arl1-Pulldown 49
 - 2.2.3.6 Isothermal titration calorimetry 50
 - 2.2.3.7 Circular Dichroism spectroscopy 50
 - 2.2.3.8 SEC-MALLS 51
 - 2.2.3.9 Crystallization 51

2.2.3.10	Flash cooling of protein crystals	53
2.2.4	Computational methods	53
2.2.4.1	Structure determination of SCOC ccd	53
2.2.4.2	Sequence alignment of SCOC	55
3	SCOC and its interaction partners	57
3.1	Characterization & structure determination of the SCOC coiled coil domain	57
3.1.1	Generation of SCOC constructs for crystallization	57
3.1.2	Expression and Purification of SCOC constructs	60
3.1.2.1	Expression and Purification of SCOC (78–159)	60
3.1.2.2	Expression and Purification of SCOC (78–132) and SCOC (112–159)	62
3.1.2.3	Expression and Purification of SCOC (78–141) and SCOC (78–151)	64
3.1.2.4	Expression and Purification of GST-tagged SCOC constructs	66
3.1.3	Biochemical and biophysical characterization of SCOC coiled coil domain	66
3.1.3.1	Limited proteolysis of SCOC coiled coil domain	66
3.1.3.2	Pull-down experiments for the identification of novel SCOC binding partners	71
3.1.3.3	Analysis of secondary structure and stability of SCOC coiled coil domain by CD spectroscopy	72
3.1.3.4	Characterization of oligomerization state of SCOC coiled coil domain by SEC-MALLS measurements	73
3.1.4	Crystallization of SCOC coiled coil domain	73
3.1.4.1	Crystallization screening of SCOC (78–159)	76
3.1.4.2	Refinement of the crystallization conditions of SCOC (78–159) crystals	76
3.1.4.3	Screening of other SCOC ccd constructs	80
3.1.4.4	Crystallization of selenomethionine-labelled SCOC crystals	80
3.1.4.5	Crystals grown with <i>in situ</i> proteolysis	80

3.1.5	Structure determination	81
3.1.6	Structure of SCOC's coiled coil domain	85
3.1.7	Biophysical characterization of SCOC's mutants	91
3.2	Interaction of SCOC's coiled coil domain with Arl1 and FEZ1	99
3.2.1	Insights into SCOC–FEZ1 complex formation	99
3.2.1.1	Approaches to find a minimal FEZ1 binding domain	99
3.2.1.2	Coexpression and copurification of SCOC–FEZ1 ccd complexes	101
3.2.1.3	Characterization of SCOC–FEZ1 complexes	102
3.2.2	Interaction of SCOC with Arl1	104
3.2.2.1	Expression, purification and nucleotide exchange of Arl1	104
3.2.2.2	Arl1-Pulldown experiments	106
3.2.2.3	Analytical gel filtration	107
3.2.2.4	ITC	108
3.3	Discussion	110
3.3.1	Characterization and structure of SCOC ccd	110
3.3.2	Interaction with FEZ1 and Arl1	111
3.4	Outlook	115
4	Shigella flexneri and Autophagy	117
4.1	Characterization of VirG	117
4.1.1	Expression and purification of VirG (353–758)	117
4.1.2	Expression and purification of VirG (52–758)	120
4.1.3	Analysis of refolding and secondary structure by CD spectroscopy	120
4.1.4	Crystallization of VirG	121
4.2	Characterization of IcsB–IpgA	124
4.2.1	Expression and purification of IcsB–IpgA	124
4.2.2	Limited proteolysis of IcsB–IpgA	126
4.2.3	Crystallization of IcsB–IpgA	127
4.2.4	Analysis of IcsB–IpgA by heteronuclear NMR spectroscopy	127
4.3	Discussion and Outlook	130
A	Appendix	133
A.1	DNA sequences of synthetic genes	133

A.1.1	SCOC isoform 1	133
A.1.2	Arl1 full length Q79L	133
A.1.3	VirG passenger domain (52–758)	133
A.1.3.1	part 1	133
A.1.3.2	part 2	134
A.2	Crystallization conditions of SCOC ccd constructs	135
A.3	Supplementary Figures	143
A.3.1	Elution Profiles	143
A.3.2	Purification of SCOC ccd–FEZ1 ccd complexes	143
A.4	Computational Methods	145
A.4.1	COILS results	145
A.4.2	HHpred results	149
A.4.3	XDS input file	153
A.4.4	Systematic absences by CORRECT.LP	154
A.4.5	List of Se positions by Phenix	158
A.4.6	T-COFFEE alignment across species	158
A.5	Mass spectrometry results	160

Bibliography

List of Figures

1.1	Autophagosome formation	2
1.2	Different subtypes of autophagy and possible membrane sources for membrane biogenesis	3
1.3	Chronology of autophagosome biogenesis in mammals	5
1.4	Complexes involved in initiation and nucleation of autophagy	6
1.5	Crystal structure of GTP-Arl1 (PDB 1UPT)	8
1.6	Helical wheel scheme of a parallel dimeric coiled coil	11
1.7	Mechanism of <i>Shigella flexneri</i> 's escape from xenophagy	13
1.8	Secretion mechanism of type Va autotransporters	14
1.9	Crystal structure of VirG (591–758)	15
3.1	Overview of SCOC constructs used for crystallization	58
3.2	Disorder prediction for SCOC Isoform 1 by IUPred [101, 102]	58
3.3	Purification of SCOC (78–159)	61
3.4	Purification of SCOC (78–132)	62
3.5	Purification of SCOC (112–159)	63
3.6	Purification of SCOC (78–141)	64
3.7	Purification of SCOC (78–151)	65
3.8	Purification of GST-SCOC full length	67
3.9	Purification of GST-SCOC (78–159)	68
3.10	Limited proteolysis of SCOC (78–159)	69
3.11	Limited proteolysis of SCOC (78–159) by Trypsin and Subtilisin	70
3.12	GST-Pulldown of GST-SCOC (78–159)	71
3.13	CD measurements of SCOC (78–159)	73
3.14	SEC-MALLS experiments of SCOC (78–159)	74
3.15	Exemplary crystallization hits of SCOC ccd constructs from 96-well plate screening	75
3.16	Effect of streak seeding	76
3.17	Crystals in hanging drops grown in 24-well format	77
3.18	Crystals obtained from refinement with grid and random screens	79

3.19 Crystals grown by <i>in situ</i> proteolysis	82
3.20 Phenix Xtriage analysis	86
3.21 Polygon plot by Phenix	87
3.22 Ramachandran plot of all non Pro/Gly residues of SCOC L105M (78–159)	87
3.23 Structure of SCOC coiled coil	88
3.24 Two distinct SCOC dimers in the crystal structure	90
3.25 Surface charge of SCOC	91
3.26 SCOC conservation degree on the surface	92
3.27 Sequence alignment of human SCOC isoforms and homologues	92
3.28 Helical wheel diagram of the SCOC dimer	93
3.29 Stabilizing interactions of non-canonical core positions	94
3.30 CD spectra of SCOC mutants	95
3.31 Thermal unfolding CD curves	96
3.32 Analytical gel filtration elution profiles of SCOC mutants	97
3.33 Elution profiles and corresponding molecular weights determined by SEC-MALLS are shown for all SCOC mutants	98
3.34 FEZ1 domain	99
3.35 Test coexpression of Strep-SCOC (78–159) with His-FEZ1 ccd constructs	100
3.36 Copurification of wt Strep-SCOC (78–159) with His-FEZ1 (226–290)	101
3.37 Coexpression of Strep-SCOC (78–159) constructs with His-FEZ1 (227–290)	102
3.38 Analytical gel filtration of SCOC (78–159)—FEZ1 (227–290) complex	103
3.39 SEC-MALLS measurements of wt SCOC (78–159)—FEZ1 (227–290) complex	103
3.40 Purification of Q79L Arl1 (15–181)	105
3.41 Analysis of bound nucleotide	106
3.42 His-Arl1 pulldown with SCOC (78–159)	106
3.43 Analytical gel filtration	107
3.44 ITC titration of SCOC (615 μ M) into Arl1 (50 μ M)	109
3.45 Co-immunoprecipitation of full length FEZ1 tagged N-terminally with a V5 epitope and various EmGFP-SCOC (78–159) variants from transfected 293 cell lysates using an anti-GFP antibody	113
3.46 SCOC complexes in autophagy and beyond	114

4.1	Purification of VirG (353–758)	118
4.2	Purification of VirG (52–758)	119
4.3	CD spectrum of VirG (353–758) at 20 °C	121
4.4	CD spectrum of VirG (52–758) in 1 M urea at 20 °C	121
4.5	96-well plate screening of VirG (353–758)	122
4.6	96-well screening of VirG (52–758)	123
4.7	Purification of IcsB–IpgA	125
4.8	Limited proteolysis of IcsB–IpgA by Trypsin and Chymotrypsin	126
4.9	Two-dimensional HSQC spectrum of IcsB–IpgA	128
4.10	Two-dimensional TROSY spectra of IcsB–IpgA and IpgA	129
A.1	Elution profiles of SCOC–FEZ1 copurification	143
A.2	Schägger gels of copurifications of mutant SCOC ccd–FEZ1 ccd	144

List of Tables

2.1	Enzymes	17
2.2	Kits	17
2.3	Columns	18
2.4	Antibodies	19
2.5	Composition of 1 L minimal media for ¹⁵ N-labelling (modified from Studier [90])	21
2.6	Composition of 1 L minimal media for selenomethionine-labelling (modified from Studier [90])	22
2.7	Composition of 1 L ZYM-50502 media [90]	23
2.8	DNA constructs used in this study	24
2.9	Oligonucleotides	29
2.10	PCR reaction 50 µL	32
2.11	PCR temperature cycling	33
2.12	Restriction digest reaction setup	34
2.13	Mutagenesis PCR reaction (25 µL) setup	35
2.14	Mutagenesis PCR Temperature cycling	36
2.15	Composition of Arl1-Pulldown reactions	49
3.1	N-terminal sequencing results	70
3.2	SCOC peptides detected by mass spectrometry in band 5	72
3.3	Exemplary crystallization conditions of SCOC ccd constructs from 96-well plate screening	74
3.4	Crystallization conditions in 24 well plate screens	77
3.5	Exemplary crystallization conditions of Random and Grid screens	78
3.6	Exemplary crystallization conditions of <i>in situ</i> proteolysis	83
3.7	Data collection	83
3.8	Diffraction data and refinement statistics	84
3.9	Summary of biophysical characterization of SCOC mutants	93
3.10	FEZ1 fragments comprising the C-terminal ccd	100
3.11	Summary of SCOC ccd crystallization	111

3.12	Chrystallization probability of yeast FEZ1 homologues	116
4.1	Crystallization conditions of VirG (353–758) from 96-well plate screens	122
4.2	Crystallization conditions of VirG (52–758) from 96-well screens	124

List of Abbreviations

aa	Amino acid
ccd	coiled coil domain
CD	Circular Dichroism
CV	Column volume
cvt	cytoplasm to vacuole
fwd	Forward
GdnHCl	Guanidinium Hydrochloride
GDP	guanosine diphosphate
GDPNHP	Guanosine-5'-[(β , γ)-imido]triphosphate
GST	glutathione-S-transferase
GTP	guanosine triphosphate
HSQC	Heteronuclear single quantum coherence spectroscopy
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
kb	Kilo base
MAD	Multi wavelength anomalous diffraction or dispersion
MCS	multiple cloning site
MWCO	molecular weight cut off
NaAc	Sodium Acetate

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein DataBase
PEG	poly ethylene glycol
rev	Reverse
rpm	Rounds per minute
RT	Room temperature, 22–25 °C
SAD	Single wavelength anomalous diffraction or dispersion
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC-MALLS	Size Exclusion Chromatography coupled to Multi Angle Laser Light Scaterer
TROSY	Transverse relaxation optimized spectroscopy
wt	wild-type
(v/v)	Volume per volume
(w/v)	Weight per volume

1 Introduction

1.1 General introduction to Autophagy

Autophagy is a fundamental eukaryotic cellular pathway for the degradation of cellular components such as proteins and membrane organelles. The evolutionary conserved self-eating mechanism occurs at a low basal level and becomes activated under stress conditions. It serves as an adaptive catabolic process and provides energy for the cell.

Autophagy was first described in mammals during the 1950s [1–3], however the term “autophagy” was established by Christian De Duve in 1963 [4], when he observed vesicles with a double membrane engulfing degraded cytoplasmic content in EM images. Although autophagy was initially investigated in mammals, milestones for autophagy-related research were genetic screens in yeast in the 1990s [5–8]. To date, 36 autophagy related genes (Atg) were identified [9], (reviewed e.g. by Mizushima *et al.* [10]).

During macroautophagy biogenesis, the best characterized autophagic subtype, a growing isolation membrane, also called phagophore, enwraps its cytoplasmic content. Fusion of its ends leads to the formation of an autophagosome, whose outer membrane then fuses with the lysosome in animal cells. Contents are degraded by lysosomal hydrolases (Figure 1.1). In yeast, the autophagosome initiates at the phagophore assembly site (PAS) next to the vacuole, to which Atg proteins are initially recruited. The emerging autophagosome fuses with the vacuole.

1.1.1 Subtypes of Autophagy

There are specific and non-specific subtypes of autophagy (see Figure 1.2). *Macroautophagy*, hereafter referred to as autophagy, connotes the bulk degradation of cytoplasmic content [10, 12]. It occurs at a basal level and is important for mediating cellular homeostasis by degrading protein aggregates and damaged or excess organelles which are too large to be processed by the proteasome. Thus autophagy is crucial for maintaining quality control of es-

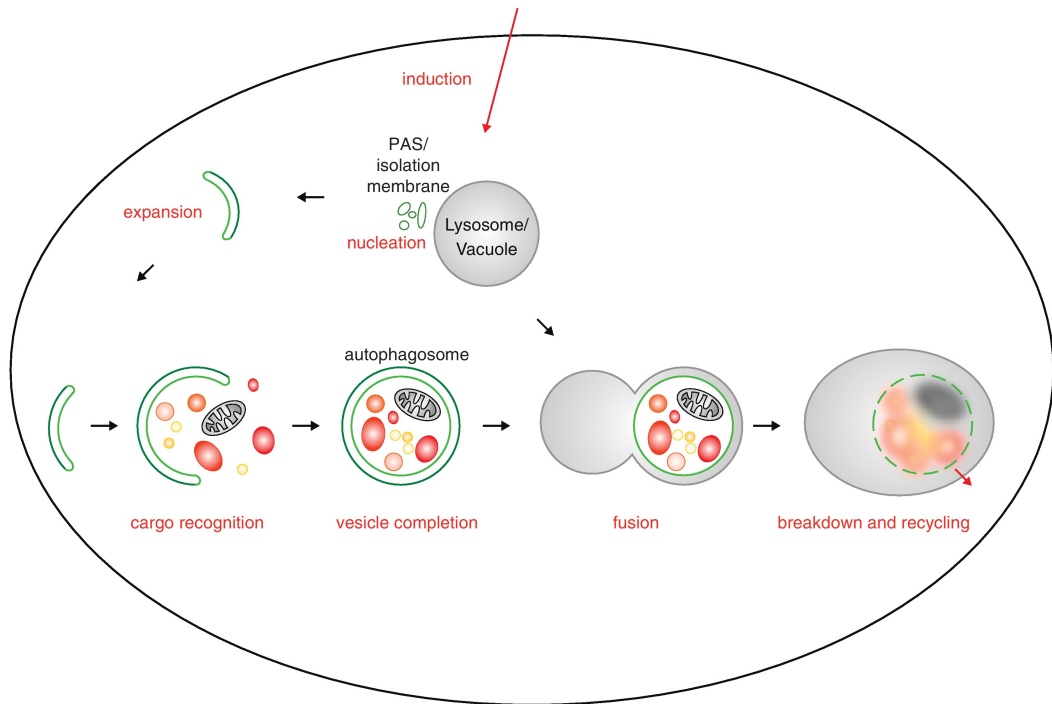


Figure 1.1: Autophagosome formation

After induction and nucleation of autophagy, the isolation membrane is expanding. Cytoplasmic content is engulfed and the autophagosome matures. Upon fusion with the lysosome, its content is degraded. Figure is cited from Kraft & Martens [11].

essential cell components [11]. Impaired autophagy has been linked to cancer and to aging, furthermore to a number of neurodegenerative diseases caused by protein aggregates, such as Parkinson, Alzheimer and Huntington's disease [12, 13]. Selective types of autophagy include the targeted degradation of mitochondria (*mitophagy* [14]), of ribosomes (*ribophagy* [15]) and the removal of peroxisomes (*pexophagy* [9]). *Xenophagy*, a defense mechanism against invading pathogens, will be discussed in more detail in Section 1.3.1. The yeast-specific *cytoplasm-to-vacuole targeting pathway* (*Cvt*) is a biosynthetic pathway and selectively delivers a-mannosidase and proaminopeptidase I to the vacuole [16]. Proaminopeptidase becomes activated in the vacuole through processing. Selective autophagy pathways are mediated by specific adaptors, e.g. p62, Optineurin or Atg36 [17].

In higher eukaryotic species *chaperone-mediated autophagy* occurs, where

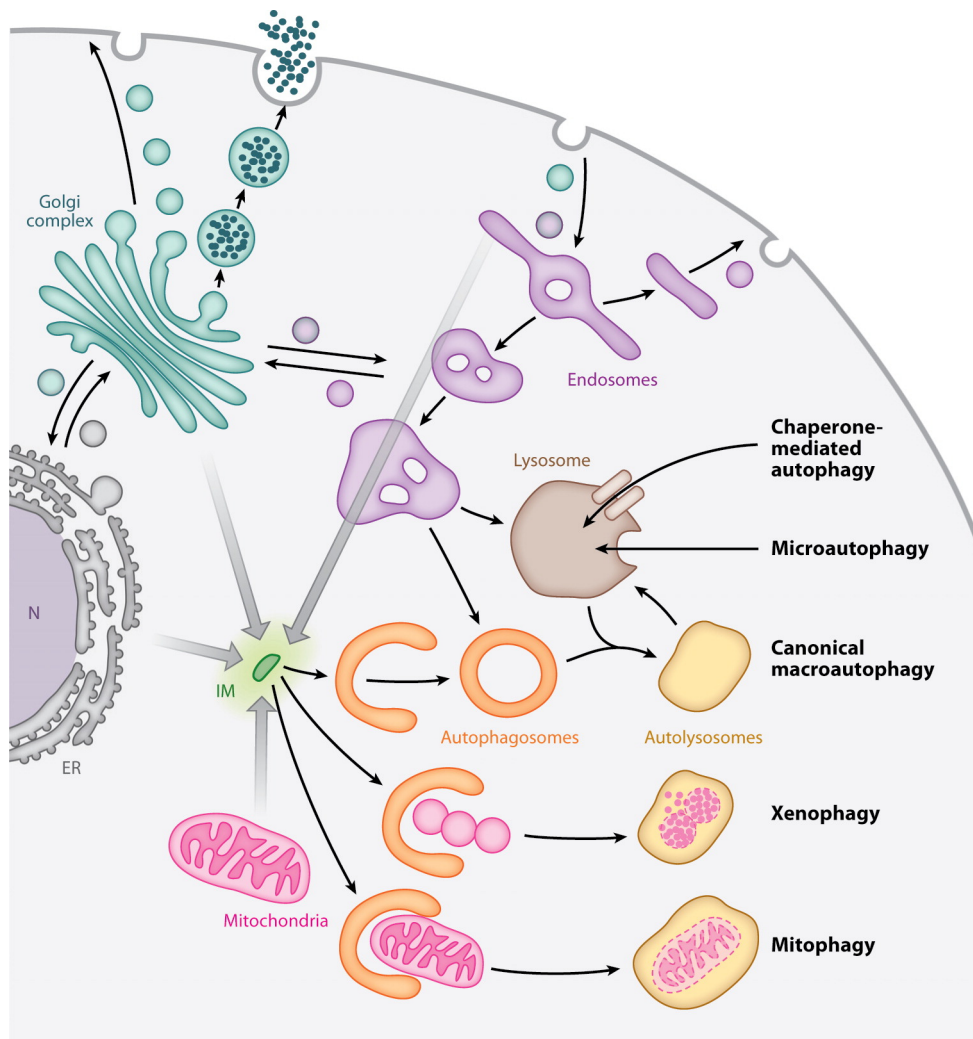


Figure 1.2: Different subtypes of autophagy and possible membrane sources for membrane biogenesis

Subtypes of autophagy are depicted, e.g. canonical macroautophagy, xenophagy and mitophagy. Delivery of cytoplasmic content through invagination of the lysosomal membrane into the lumen is called microautophagy, whereas in chaperone-mediated autophagy, cargo is specifically recognized by chaperones and translocated across the lysosomal membrane. Lipid sources include the ER, Golgi, plasma membrane and mitochondria. Figure is cited from Mizushima *et al.* [10].

cargo is specifically recognized by chaperones via the consensus motif KFERQ. Chaperone-substrate complexes bind to a receptor at the membrane. The unfolded cargo is transported across the lysosomal membrane and degraded [18].

During *microautophagy*, cargo is directly engulfed by invaginations of the lysosomal membrane [19].

The source of lipids for the growing isolation membrane has been debated in the field for some time [20–22]. Several cellular compartments have been identified as the membrane source. Autophagosomes originate close to the ER and EM tomograph images showed direct connections between isolation membrane and ER [23–25]. Nevertheless, other sources of membrane precursors are reported, among them mitochondria [26], the Golgi apparatus [27] and the plasma membrane [28]. The membrane source might depend on the specific conditions, e.g. autophagy subtype and how autophagy is induced. Cells may be able to use a combination of compartments for autophagosomal biogenesis [11].

1.1.2 The autophagic core machinery

All autophagy subtypes share a common molecular machinery for the induction, nucleation and maturation of autophagosomes. The autophagic core machinery consists of four main groups [12]:

- (I) unc-51-like kinase (ULK) complexes (Atg1 in yeast)
- (II) ubiquitin-like conjugation systems Atg8 and Atg12
- (III) class III phosphatidylinositol 3-kinase (PtdIns3K) (yeast Vps34) complex including Beclin1 (yeast Atg6) and ultraviolet irradiation resistance-associated gene (UVRAG) (yeast Vps38)
- (IV) the transmembrane protein Atg9 and associated proteins.

A chronology of events during mammalian autophagosome biogenesis is depicted in Figure 1.3. Autophagy is initiated through the activation of ULK1 kinase after it dissociates from the target of rapamycin (TOR) complex 1 (TORC1) (Figure 1.4). Several phosphorylation events occur in the ULK1 complex, their biological implications are not yet completely understood. The next key event in autophagosome biogenesis is the recruitment of the Beclin1-PtdIns3K complex (Vps34, p150, Beclin1 and Atg14). UVRAG, which is also part of different Beclin1 complexes [30–34], is a positive regulator of autophagy

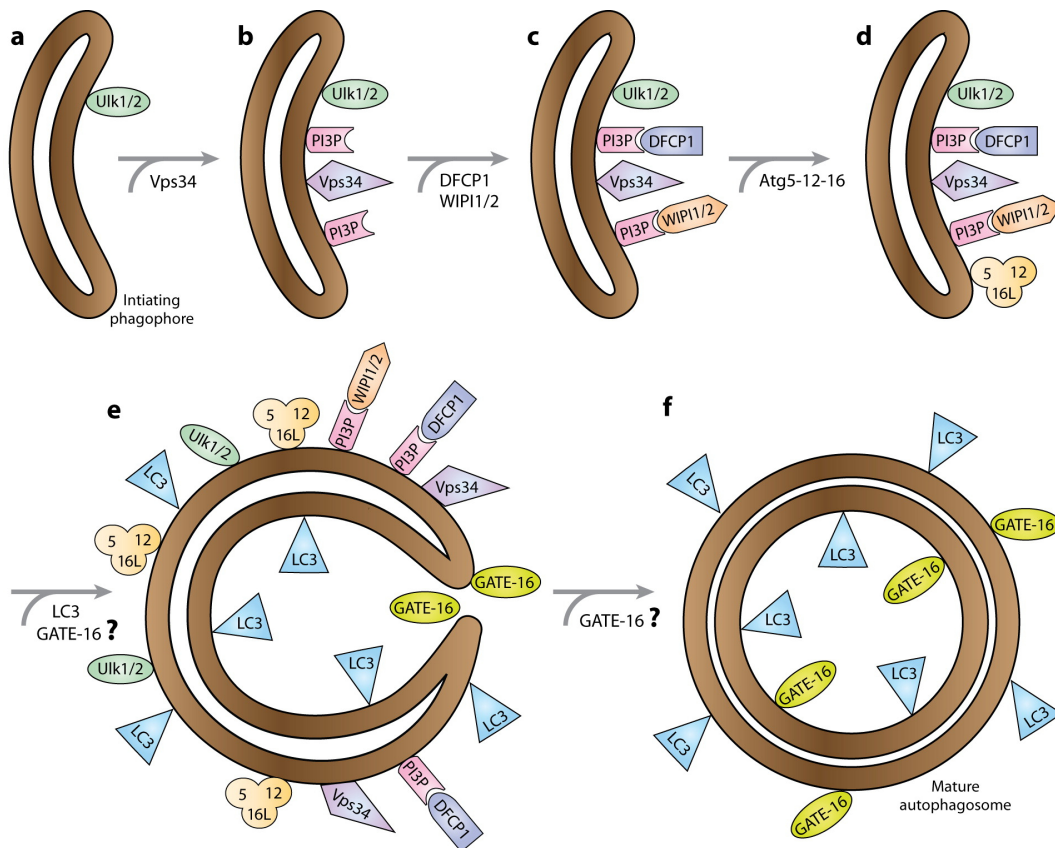


Figure 1.3: Chronology of autophagosome biogenesis in mammals

Autophagy proteins are recruited in a hierarchical manner to the isolation membrane and most of them associate only transiently with it. Figure is cited from Weidberg *et al.* [29].

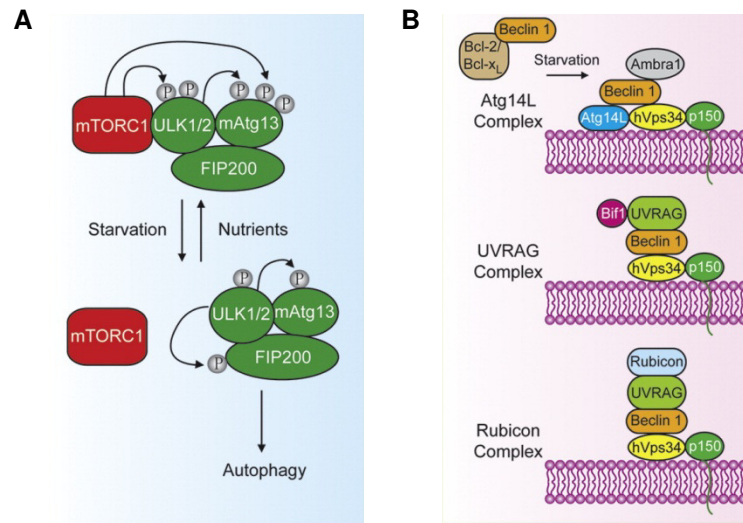


Figure 1.4: Complexes involved in initiation and nucleation of autophagy

(A) ULK1 complexes involved in initiation and regulation of autophagy. Under nutrient starvation, ULK1 and mAtg13 are dephosphorylated, mTOR dissociates and FIP200 and Atg13 are phosphorylated by ULK1. (B) PtdIns3K complexes comprising Beclin1 and UVRAG. PtdIns3K complexes are involved in clearance (Rubicon complex) or formation of autophagosomes (Atg14L and UVRAG complex). Figures were modified from Yang & Klionsky [12].

(Figure 1.4). Phosphatidylinositol-3-phosphate is produced by PtdIns3K complex. Further expansion of the isolation membranes depends on the two ubiquitin-like conjugation systems. The Atg12–Atg5–Atg16L conjugate is generated by E1 and E2-like enzymes Atg7 and Atg10 forming a bond between Atg12 and Atg5, which further associates with Atg16L. LC3, the mammalian Atg8 homologue, is covalently linked to phosphatidylethanolamine by the action of E1- and E2-like enzymes Atg7 and Atg3, requiring also C-terminal processing of LC3 by Atg4 [29, 35]. Lipidated LC3 localizes to both sides of the isolation membrane, whereas the Atg12–Atg5–Atg16L complex associates with the outer side of the isolation membrane. Atg9 plays another key role in autophagy. It shuttles in vesicles between autophagosomes and Golgi/endosomes, thereby supplying membrane precursors to the autophagosome [12]. Shortly before or after completion of autophagosome formation Atg12–Atg5–Atg16L dissociates from the surface. The “naked” autophagosome then fuses with the

lysosome where its content is degraded and recycled.

1.2 SCOC and its interaction partners

1.2.1 SCOC: an Arl1 effector

Human short coiled coil protein (SCOC) is an effector of the Golgi resident Arf-like GTPase 1 (Arl1) [36] and endogenous SCOC colocalizes with Arl1 at the Golgi [36]. In addition, SCOC was recently identified as a positive regulator of autophagy in a genome-wide siRNA screen [35]. SCOC is widely expressed in the human body, most abundantly in the brain, heart and skeletal muscle [36]. Su *et al.* [37] observed enrichment of human SCOC mRNA in fetal brain, indicating a role for SCOC in mammalian nervous system development.

At least four different human isoforms share the well conserved C-terminal coiled coil domain of SCOC, but they differ in their N-termini resulting from alternative splicing (see Figure 3.27). The coiled coil domain (ccd) is identical for isoforms 1, 2 and 3, whereas isoform 4 is missing residues 85–112. Homologues of SCOC are found in many eukaryotes with the ccd being conserved across species (Section A.4.6). The *C. elegans* orthologue of SCOC, UNC-69a, and the yeast orthologue, Slo1p, have been described in literature [37, 38]. However, the *S. cerevisiae* homologue is less well conserved than other eukaryotes [37]. SCOC does not seem to be a conserved GTPase effector, since yeast Slo1p interacts with Arl3p in a nucleotide-independent manner and not with Arl1p [38]. A deletion of Slo1p did not affect viability of the cells or impair the known function of Arl3p [38]. UNC-69 does not interact with any of the *C. elegans* Arl homologues [37].

1.2.2 SCOC interaction with Arl1

Arl1 is a small GTPase (181 residues) that is involved in membrane trafficking. Arl1 is essential for the translocation of GRIP-domain containing golgins, large coiled coil proteins important for Golgi structure and trafficking, to the Golgi [39]. In yeast, activated Arl1p recruits the golgin Imh1p via interaction with its GRIP domain to the trans-Golgi network. GTP is hydrolyzed upon recycling of Imh1p to the cytosol [40]. Arl1 is also involved in the activation of Arf1 by targeting the guanine nucleotide exchange factors BIG1 and BIG2 to the Golgi [41].

Three crystal structures of human Arl1 are deposited in the Protein Data

Bank (PDB) (PDB accession numbers 1UPT [39], 1R4A [42], 4DCN [43]). Arl1 was crystallized in complex with the GRIP domain of Golgin-245 (1UPT and 1R4A), and in complex with the Arfaptin2 BAR domain. The crystal structure in complex with the Golgin-245 GRIP domain (1UPT) is depicted in Figure 1.5.

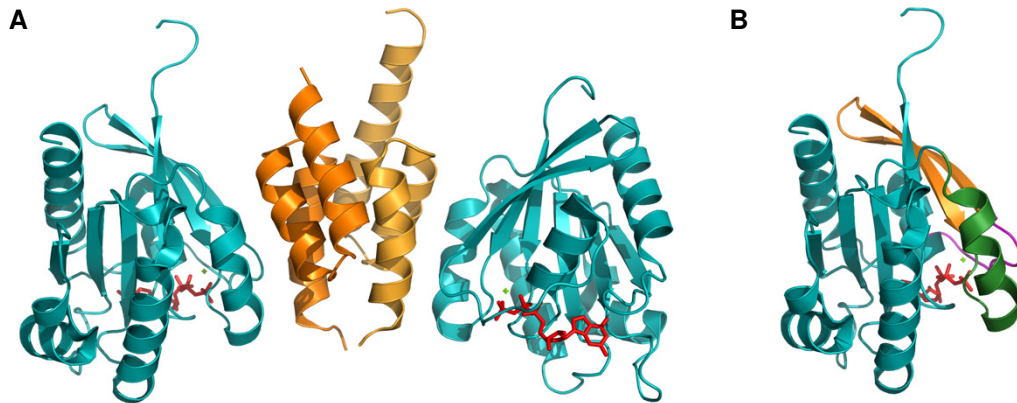


Figure 1.5: Crystal structure of GTP-Arl1 (PDB 1UPT)

(A) Structure of GTP-Arl1 in complex with the GRIP domain of Golgin-245. The homodimer consisting of two Arl1 molecules (cyan) and two GRIP domains (light orange and orange) is depicted with two GTP molecules (red) and $MgCl_2$ ions (light green) (B) Detailed view of Arl1. GTP is coloured red, switch 1 region purple, switch region 2 green and the interswitch region orange.

The GRIP domain oligomerizes to a homodimer with each molecule binding to an Arl1 molecule. Arl1 features a fold typical for Arf family GTPases: six β -strands are enwrapped by five α -helices. The molecule has two switch and one interswitch region which change their conformation between nucleotide binding states. Interfaces between Arl1 and the GRIP domain occur between α -helices of the GRIP domain and the switch regions, explaining specificity for GTP-bound Arl1 [39].

1.2.3 SCOC interaction with FEZ1

Mammalian SCOC interacts with fasciculation and elongation protein zeta 1 (FEZ1) [36, 44], a protein that is essential for kinesin-1 mediated transport along axons. *C. elegans* orthologue of SCOC UNC-69 interacts with UNC-76,

the FEZ1 orthologue. Deletion of UNC-69 resulted in defects of axon growth, guidance and their fasciculation, and an abnormal presynaptic organization was observed, implying a function of the UNC-69–UNC-76 complex in axonal transport of vesicles [37].

Human FEZ1 (392 residues) is a mainly natively unfolded protein with a conserved coiled coil domain in the C-terminal half of the protein [45].

FEZ1 acts as an adaptor in kinesin-1 mediated axonal transport to nerve terminals by binding to both the heavy chain of the motor protein kinesin-1 [46, 47] and its cargo, for example as recently shown for Syntaxin 1a and Munc18 containing transport vesicles [48]. Phosphorylation of FEZ1 regulates cargo [49] and kinesin binding [48]. Mutations of the *C. elegans* FEZ1 orthologue UNC-76 lead to severe defects in axon growth and fasciculation [50], similar to the phenotype when its binding partner UNC-69 was deleted. Importantly, UNC-69 function was rescued when human SCOC gene was expressed under the UNC-69 promoter, implying that the function is conserved [37]. *In vitro* interaction of both proteins was confirmed by GST-pulldown experiments. Point mutation of L287 or deletion of UNC-76 residues 281–299 abolished the interaction, proving that the conserved coiled coil of UNC-76 is responsible for binding [37]. Similarly, McKnight *et al.* [35] show interaction of SCOC ccd with FEZ1 ccd by pulldown experiments and co-immunoprecipitation. Mutation of leucines in the C-terminal coiled coil of FEZ1 led to disruption of the SCOC–FEZ1 complex in the pulldown experiments.

A distinct regulatory role in autophagy has also been attributed to SCOC and FEZ1 [35]. Axonal transport in fruit fly is regulated by the phosphorylation of UNC-76 by UNC-51/Atg1 [49, 51]. Human orthologues of FEZ1 and ULK1 also interact, and their complex formation is modulated by SCOC [35]. FEZ1 interacts with both the kinase and the proline-serine rich domain of ULK1. Their interaction is nutrient-independent, but a depletion of SCOC leads to reduced complex formation [35]. The FEZ1–ULK1 complex inhibits autophagy induction, and ULK1 is released upon binding of SCOC to FEZ1 [35]. SCOC–FEZ1 also forms a complex with UVRAG [35]. Under starvation conditions, this interaction is reduced, but can be stabilized in the presence of FEZ1.

Summing up, SCOC–FEZ1 complex formation is mediated through the coiled coil domains of SCOC and FEZ1 [35, 44]. Together they interact with ULK1 and UVRAG, implying a regulatory function of SCOC in the crosstalk between these two essential autophagy complexes.

1.2.4 Coiled coil domains: features, stability and oligomerization state

α -helical coiled coil domains are one of the most common protein structures found in diverse protein families [52]. Despite their rather simple fold, proteins containing coiled coils exhibit a large variety of functions, among them transcription factors, proteins involved in vesicular transport or scaffolding proteins. Besides facilitating protein homo-oligomerization, coiled coils are also very important for mediating protein-protein interactions. The coiled coil interaction network in *S. cerevisiae* was characterized through yeast two-hybrid assays. In this study 3495 pairwise interactions were identified among 598 predicted coiled coil regions in 453 proteins, which are extensively involved in the organization of the cellular machinery [53].

Coiled coils are formed by at least two α -helices that are wound around each other forming a superhelical structure reviewed in Lupas & Gruber [54]. The sequence pattern underlying the supercoiled structure is characterized by a heptad repeat pattern $(a, b, c, d, e, f, g)_n$ [54, 55], where positions a and d are occupied by mostly apolar amino acids like leucine, valine and isoleucine. These residues form the hydrophobic core of coiled coils, determining the fold into an α -helix with a hydrophobic site, due to the winding of ~ 3.6 residues per turn. Oligomerization of two or more such helices results in energetically favored burial of the hydrophobic a/d residues. Complementary packing of a and d residues in the hydrophobic core combined with electrostatical interactions of e and g residues determines structural stability and specificity (see Figure 1.6) [56]. Hence, in the ideal case, e and g are charged residues, and the remaining three residues b,c and f, facing the surface-exposed sites of the helix, are preferably hydrophilic [57].

Although the prediction of coiled coils has become a reliable bioinformatical tool [58], it is still difficult to predict the oligomerization state based on sequence data. The nature of the amino acids at the a/d-positions is important for determining the oligomerization state of a coiled coil protein [55, 56, 59].

1.2.5 Aims

SCOC and FEZ1 interact via their coiled coil domains. Their complex has been implicated in the regulation of the first steps of autophagy. FEZ1 interaction with ULK1 is modulated by SCOC, and SCOC itself interacts with UVRAG. Moreover, an analogous complex in *C. elegans* is involved into the kinesin-dependent transport along axons. Also, SCOC does interact with

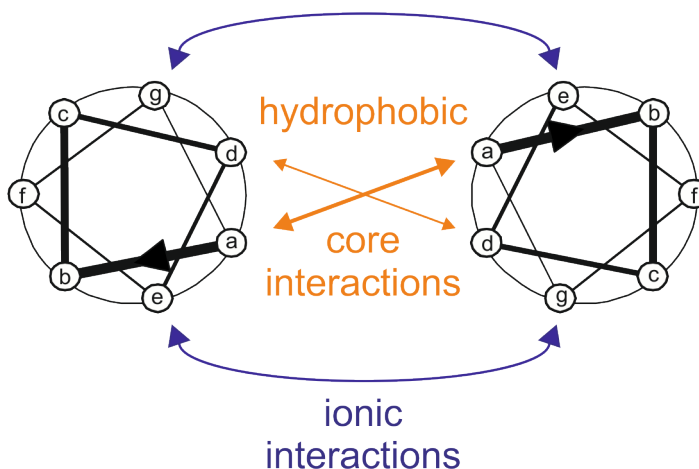


Figure 1.6: Helical wheel scheme of a parallel dimeric coiled coil

Hydrophobic interactions between residues a and d stabilize the core. Ionic interactions between residues e and g give specificity to the coiled coil.

Golgi-resident Arl1. How can such a small protein exhibit so many different functions? SCOC contains only one functional domain, thus, the answer to the diversity of interactions and functions must rely in the nature of SCOC coiled coil. Hence, it was my aim to determine the SCOC ccd structure by X-ray crystallography. This included the expression and purification of SCOC constructs suitable for crystallography. These constructs were further characterized by biophysical and biochemical methods, providing insights into the stability and oligomerization state of SCOC. In addition, the interactions of SCOC with FEZ1 and Arl1 were analyzed with different methods.

1.3 *Shigella flexneri* and Autophagy

1.3.1 Selective Autophagy: Xenophagy

Autophagy is of great medical interest since it is not only a major target of cancer research, but also plays an important role in the innate and adaptive immunity response of higher eukaryotes (reviewed in Levine *et al.* [60], Deretic & Levine [61]). Xenophagy is, like other selective autophagy pathways, mediated through autophagy adaptors. Adaptors function by binding both the autophagic cargo and to a LC3 protein family member by a LC3-interacting

region (LIR). Sequestosome 1 (SQSTM1/p62) [62], nuclear dot protein 52 kDa (NDP52) [63–65], optineurin (OPTN) [66, 67], and neighbor of BRCA1 gene 1 (NBR1) [68] are adaptor proteins involved in the autophagic clearance of pathogens [69, 70]. All of these adaptors contain a ubiquitin binding domain, by which they recognize their ubiquitinated substrate [70].

Xenophagy plays a role in intracellular infections with various bacteria, e.g. *Shigella*, *Mycobacteria*, *Salmonella*, *Listeria*, and *Legionella* [69]. The escape mechanisms of pathogens as well as the host cell’s defensive mechanism are diverse. For example, *Mycobacterium tuberculosis* survives in phagosomal compartments after invading the host cell by arresting fusion of the phagosome with the lysosome [71, 72]. Virulent *Mycobacteria* strains can also resist and inhibit autophagy, but this inhibition can be overcome by the induction of autophagy through various stimuli [73, 74].

In *Salmonella* infection, the bacterium survives in a *Salmonella*-containing vacuole, in which *S. typhimurium* can replicate [75]. It secretes several effectors through its type III secretion system (TTSS), which results in bacteria invading the cytosol. *S. typhimurium* in the cytosol are rapidly polyubiquitinated and then recognized by the respective cargo adaptor NDP52 [63, 69].

1.3.2 Escape of intracellular *Shigella flexneri* from autophagy

Shigella flexneri is a human pathogen causing bacillary dysentery *Shigellosis*. This mucosal bacterium has versatile instruments that circumvent the host cell immune response (reviewed in Ashida *et al.* [76]). It disrupts the initial vacuolar membrane surrounding the bacterium, it multiplies in epithelial cells, invading cells by exploiting actin polymerization. It manipulates the host cell death and signalling pathways. It has adapted a sophisticated mechanism to escape autophagy (see Figure 1.7). *Shigella*’s outer membrane protein VirG, which is crucial for its actin-based motility [77, 78], has a binding site for the host protein Atg5, by which autophagic destruction of the pathogen is triggered. However, *Shigella* secretes IcsB through its TTSS, which masks the binding site of Atg5 on VirG [78]. Mutant *Shigella* bacteria lacking IcsB were enwrapped by multilamellar structures positive for LC3 more frequently, as observed by EM. *In vitro* interactions of VirG-IcsB and VirG-Atg5 were confirmed by pulldown assays. Both VirG and IcsB originate from *Shigella flexneri*’s large virulence plasmid. VirG triggers autophagy, whereas IcsB did not reduce overall autophagy levels [78].

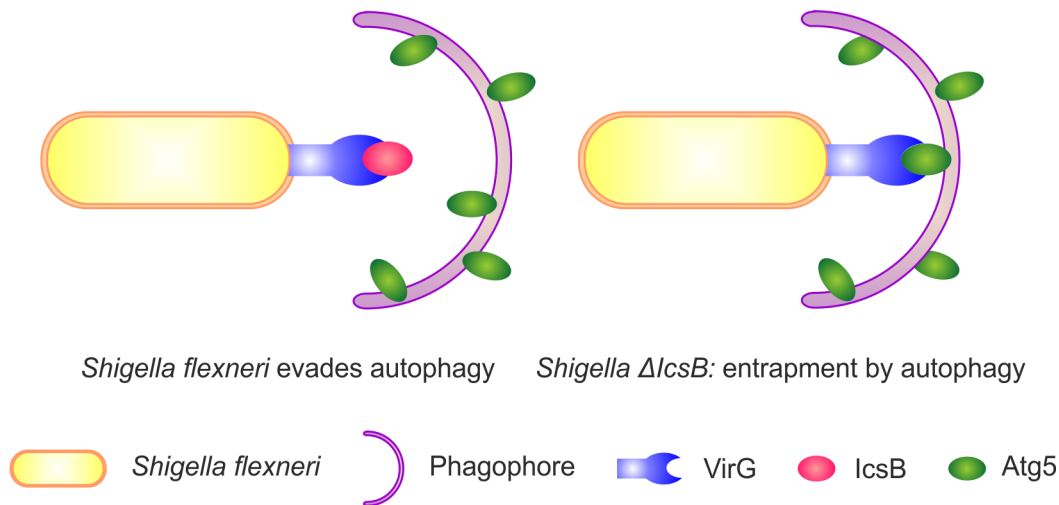


Figure 1.7: Mechanism of *Shigella flexneri*'s escape from xenophagy

VirG on the outer membrane of *Shigella* contains a common binding site for *Shigella*'s virulence effector IcsB and the host cell's protein Atg5. IcsB masks the binding site, protecting *Shigella* from degradation by autophagy. In Δ IcsB strains, VirG is recognized by the autophagic machinery through interaction with Atg5 and entrapped by autophagosomes.

In addition to this escape mechanism, *Shigella* also induces autophagy via Shiga toxins, resulting in cell death in an autophagy-dependent manner [79]. Remnants of the disrupted vacuolar membrane are targeted to autophagy via ubiquitination and interaction with p62 and LC3 [80]. Recently, also the cytoskeleton has been involved in the host cell's response to *Shigella*. Cytosolic *Shigella* are trapped in septin cages [81] and *Shigella* are targeted to an actin and septin-dependent autophagic pathway, which requires p62 and NDP52 [64].

1.3.3 The autotransporter protein VirG

VirG, also connoted as IcsA, is an autotransporter protein with the typical domain structure of a type Va autotransporter. It features a N-terminal signal sequence (1–52), a passenger domain (53–758) and a transmembrane porin domain (759–1102). The secretion mechanism of type Va autotransporters is depicted in Figure 1.8. In the bacterial cytosole, VirG is stabilized by chaperones, e.g. DnaK [83]. VirG is translocated through the inner bacterial membrane by the Sec machinery. The signal sequence is cleaved in the periplasm,

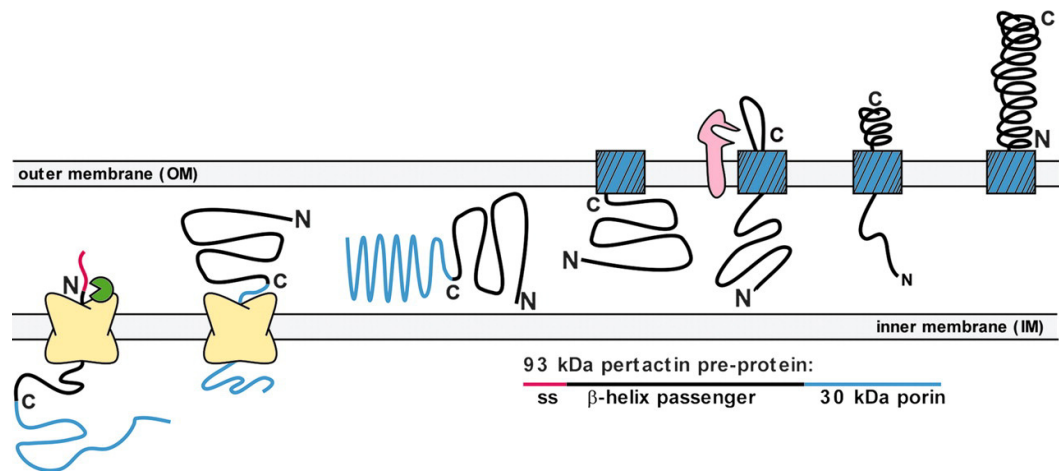


Figure 1.8: Secretion mechanism of type Va autotransporters

Figure was modified from Junker *et al.* [82]. Copyright (2006) National Academy of Sciences, USA

where VirG is chaperoned by Skp [84]. The transmembrane porin structure is inserted into the membrane and VirG's passenger domain is secreted through the porin structure.

Native folding occurs at the outer bacterial membrane, with the autochaperone region comprising residues 591–758 serving as a template and platform for correct folding of the entire passenger domain [82, 85, 86], however partial folding in the periplasm has also been discussed [84]. The autochaperone region has been shown to be essential for folding by mutational analysis [87].

A fragment of VirG comprising the autochaperone region has been crystallized by Dr. K. Kühnel (PDB 3ML3, Figure 1.9). The VirG fragment folds into two coils of a right handed parallel β -helix, with the last two antiparallel β -sheets covering the hydrophobic core. β -helical fold is typical for autotransporters [82].

VirG localizes to the pole of *Shigella*, where it recruits factors important for *Shigella*'s actin-based motility. VirG hijacks the Cdc42-controlled molecular machinery essential for actin assembly. First, IcsA binds to N-WASP and activates it. A ternary complex with Arp2/3 is formed, which stimulates actin assembly and polymerization. N-WASP and the Arp2/3 complex are crucial for *Shigella*'s ability to move and replicate within the host cell [77].

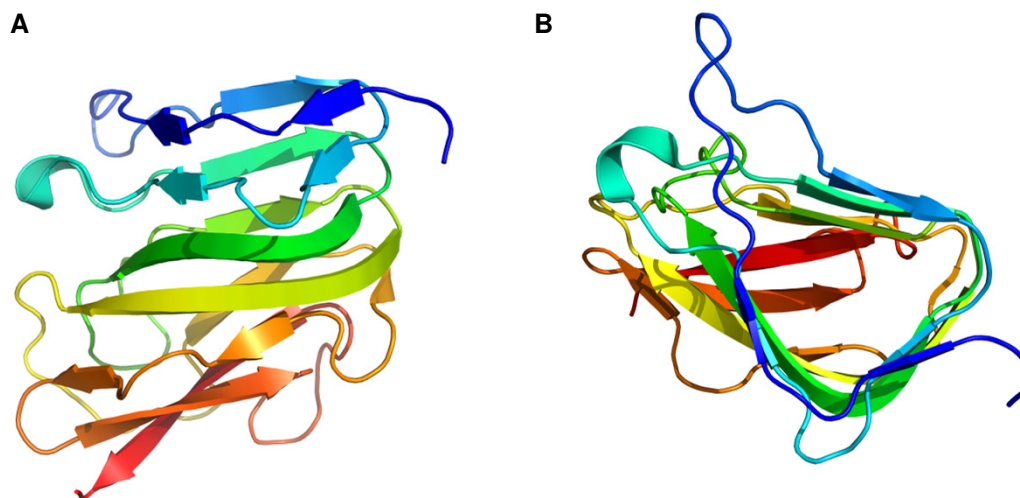


Figure 1.9: Crystal structure of VirG (591–758)

Side and top view of rainbow-coloured VirG (591–758) from blue (N-terminus) to red (C-terminus)

1.3.4 The secretion protein IcsB–IpgA

IcsB was described first by Allaoui *et al.* [88] as a virulence factor located on the *Shigella* virulence plasmid in 1992. Initial studies on a IcsB defective strain indicated, that IcsB was not crucial for invasiveness of bacteria, but caused invading bacteria to be trapped by “protrusions surrounded by two membranes”. This was probably the first observation of IcsB’s role in protecting the bacterium from autophagy. IcsB-defective bacteria were therefore not able to spread across cells [88]. IcsB is secreted via *Shigella*’s TT3S *in vivo* and *in vitro*. It is chaperoned by IpgA, the protein originating from the gene downstream of IcsB. The stop codon between the two of them is transient, so that they can be translated and secreted together as a translational fusion protein. IcsB’s middle domain was found to be involved in the interaction [89].

1.3.5 Aims

VirG and IcsB are part of a mechanism exploiting an evolutionary niche by hijacking the host cell’s defense mechanism. In the long term, structural insights into the protein structure and their interactions could contribute to the development of specific drugs against *Shigellosis*.

Hence, the aim of this project was to purify and structurally characterize *Shigella flexneri*'s proteins VirG and IcsB–IpgA. This involved the development of a suitable purification strategy for VirG. The VirG passenger domain is a membrane-attached protein, which folds into its native state at the outer membrane of the bacterium. Hence, it would be challenging to find suitable conditions for native folding to occur. Furthermore, protein interactions should be characterized by suitable biochemical methods.

2 Materials and Methods

2.1 Materials

2.1.1 Enzymes

Enzymes were used according to the manufacturers instructions. Table 2.1 lists the enzymes used in this study.

Table 2.1: Enzymes

Enzyme	Manufacturer
Restriction enzymes	NEB
T4 DNA Ligase	NEB
Alkaline Phosphatase	NEB
Trypsin	Hampton Research
Subtilisin	Hampton Research
DNaseI	Applichem
Lysozyme	Roth

2.1.2 Kits

Table 2.2 lists the kits used in this study.

Table 2.2: Kits

Kit	Manufacturer
NucleoSpin Plasmid kit	Macherey & Nagel
NucleoSpin Extract II kit	Macherey & Nagel
NucleoBond Xtra	Macherey & Nagel

(...)

Kits (*continued*)

Kit	Manufacturer
Phusion High-Fidelity PCR Kit	NEB
Proti-Ace	Hampton Research
Proti-Ace II	Hampton Research
QuickChange Lightning Site-Directed Mutagenesis Kit	Agilent Technologies (Stratagene)

2.1.3 Columns for Chromatography

Table 2.3 lists the columns used in this study.

Table 2.3: Columns

Column	Manufacturer
HiTrap Q FF column (5 mL)	GE Healthcare
His-Trap FF column (1 and 5 mL)	GE Healthcare
Strep-Trap column (5 mL)	GE Healthcare
GSTrap column (5 mL)	GE Healthcare
HiLoad 16/60 Superdex 200 prep grade	GE Healthcare
HiLoad 16/60 Superdex 75 prep grade	GE Healthcare
HiTrap Q FF column (5 mL)	GE Healthcare
Protino® GST/4B (5 mL)	Macherey and Nagel
Protino® Ni-NTA Column (5 mL)	Macherey and Nagel
Superdex 200 10/300 GL	GE Healthcare

2.1.4 Antibodies

Table 2.4 lists the antibodies used in this study.

Table 2.4: Antibodies

Antibody	Manufacturer
GST (monoclonal)	Clontech
Hexa-Histidine tag (DIA900)	Dianova
murine Strep-tag II	IBA GmbH
goat polyclonal mouse IgG (HRP labeled)	BioRad Laboratories GmbH
goat polyclonal rabbit IgG (HRP labeled)	BioRad Laboratories GmbH
SCOC coiled coil domain	generated for this Study by SySy
FEZ1	gift from John Chua

2.1.5 Buffers, media and antibiotics

1 x PBS

150 mM NaCl, 20 mM NaH₂PO₄ pH 7.4

1x TBS

50 mM Tris, 150 mM NaCl, pH 7.5
for TBS-T 0.1 % (v/v) Tween 20 was added.

3 x sample buffer for Schagger gel electrophoresis

2.5 mL 1M Tris, 6 g SDS, 15 mg Serva Blue, 4.33 g Glycerol, H₂O to 50 mL
make 10 mL aliquots, add 200 μ L β -Mercaptoethanol

10 x Anode buffer for Schagger gel electrophoresis

2M Tris to pH 8.9 with HCl

10x Cathode Buffer Schagger gel electrophoresis

1 M Tris, 1 M Tricin

10 x SDS-PAGE electrophoresis buffer

30.3 g Tris, 142.6 g Glycine, 10 g SDS pH ~8.4

10 x TBE

121 g/L Tris, 61.8 g/L boric acid, 7.5 g/L EDTA

1000 x Antibiotic stocks

Ampicillin 100 µg/ml (w/v)

Kanamycin 30 µg/ml (w/v)

Blotto

5 % (w/v) Milkpowder, 0.1 % (v/v) Tween 20 in TBS

Coomassie Blue staining solutions

A: 500 mg CoomassieR, 650 mL H₂O, 250 mL isopropanol, 100 mL acetic acid

B: 50 mg CoomassieR, 800 mL H₂O, 100 mL isopropanol, 100 mL acetic acid

C: 20 mg CoomassieR, 900 mL H₂O, 100 mL acetic acid

D: 900 mL H₂O, 100 mL acetic acid

Gel buffer for Schagger gel electrophoresis

3 M Tris pH 8.45, 0.3 % SDS

Lower gel buffer for SDS gel electrophoresis

1.5 M Tris pH 8.8, 0.4 % SDS

Upper gel buffer for SDS gel electrophoresis

0.5 M Tris pH 6.8 with HCl, 0.4 g SDS

Transfer buffer

25 mM Tris, 192 mM glycine, 0.04% w/v SDS, 20% v/v methanol

Luria Bertani (LB) media and plates

10 g/L (w/v) tryptone, 5 g/L (w/v) yeast extract, 10 g/L (w/v) NaCl

Plates contained 18 g/L (w/v) Agar.

SOC-media

2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 20 mM glucose

After autoclaving 10 mM MgCl₂, 10 mM Mg₂SO₄ were added.

M9 minimal media for ¹⁵N-labelling

See Table 2.5, modified from Studier [90].

**Table 2.5: Composition of 1 L minimal media for ^{15}N -labelling
(modified from Studier [90])**

Volume (mL)	Component	composition
200	5 x M9	37.66 g/L (w/v) $\text{NaH}_2\text{PO}_4 \times 2 \text{H}_2\text{O}$ 15 g/L (w/v) KH_2PO_4 2.5 g/L (w/v) NaCl
2.5	NH_4Cl (0.2 g/mL)	
1	1 M MgSO_4	
100	4 % glucose	
0.1	thiamine vitamin	
0.1	1 M CaCl_2	
0.2	1000x trace metals mixture	50 mL FeCl_3 in $\sim 0.12 \text{ M HCl}$ 2 mL 1 M CaCl_2 1 mL 1 M $\text{MnCl}_2 \times 4 \text{H}_2\text{O}$ 1 mL 1 M $\text{ZnSO}_4 \times 7 \text{H}_2\text{O}$ 1 mL 0.2 M $\text{CoCl}_2 \times 6 \text{H}_2\text{O}$ 2 mL 0.1 M $\text{CuCl}_2 \times 2 \text{H}_2\text{O}$ 1 mL 0.2 M $\text{NiCl}_2 \times 6 \text{H}_2\text{O}$ 2 mL 0.1 M $\text{Na}_2\text{MoO}_4 \times 2 \text{H}_2\text{O}$ 2 mL 0.1 M $\text{Na}_2\text{SeO}_3 \times 5 \text{H}_2\text{O}$ 2 mL 0.1 M H_3BO_3
700	H_2O	

M9 minimal media for selenomethionine-labelling

See Table 2.6, modified from Studier [90].

Table 2.6: Composition of 1 L minimal media for selenomethionine-labelling (modified from Studier [90])

Volume (mL)	Component	composition
200	5 x M9	37.66 g/L (w/v) NaH ₂ PO ₄ x 2 H ₂ O 15 g/L (w/v) KH ₂ PO ₄ 5 g/L (w/v) NH ₄ Cl 2.5 g/L (w/v) NaCl
1	1 M MgSO ₄	
100	4 % glucose	
0.1	thiamine vitamin	
1	Fe(II) ₂ SO ₄	
10	Aminosäuremix	Lysin 10 g/L Phenylalanine 10 g/L Threonin 10 g/L Isoleucine 5 g/L Leucine 5 g/L Valine 5 g/L
700	H ₂ O	

Terrific broth (TB) media

12 g/L tryptone, 24 g/L yeast extract, 0.4 % glycerol, 2.31 g/L KH₂PO₄ , 12.54 g/L K₂HPO₄

ZYM-5052 media (Autoinducible media)

See Table 2.7, Studier [90].

Table 2.7: Composition of 1 L ZYM-50502 media [90]

Volume (mL)	Component	composition
950	ZY	10 g/L N-Z-Amine AS (Sigma) 5 g/L yeast extract-B (QBIOgene)
20	50x5052	250 g/L (w/v) glycerol 25 g/L (w/v) glucose 100 g/L (w/v) α -lactose x H ₂ O 730 mL H ₂ O
20	50xM	222.5 g/L (w/v) Na ₂ HPO ₄ x 2 H ₂ O 170 g/L (w/v) KH ₂ PO ₄ 134 g/L (w/v) NH ₄ Cl 35.5 g/L Na ₂ SO ₄ 800 mL H ₂ O
1	2 M MgSO ₄	
0.2	1000x trace metals mixture	50 mL FeCl ₃ in ~0.12 M HCl 2 mL 1 M CaCl ₂ 1 mL 1 M MnCl ₂ x 4 H ₂ O 1 mL 1 M ZnSO ₄ x 7 H ₂ O 1 mL 0.2 M CoCl ₂ x 6 H ₂ O 2 mL 0.1 M CuCl ₂ x 2 H ₂ O 1 mL 0.2 M NiCl ₂ x 6 H ₂ O 2 mL 0.1 M Na ₂ MoO ₄ x 2 H ₂ O 2 mL 0.1 M Na ₂ SeO ₃ x 5 H ₂ O 2 mL 0.1 M H ₃ BO ₃

2.1.6 DNA constructs

Table 2.8 lists the DNA constructs used in this study.

Table 2.8: DNA constructs used in this study

gene	vector	tag	restriction sites	resistance
Arl1 (15–181) Q79L	pET-22b	6 x His	NdeI XhoI	Amp C. Behrens
Arl1 (15–181) wt	pET-22b	6 x His	NdeI XhoI	Amp C. Behrens
FEZ1 full length (John Chua)	pASK-IBA37	6 x His, StrepII		Amp
FEZ1 (226–290)	pET-22b	6 x His	NdeI XhoI	Amp C. Behrens
FEZ1 (226–295)	pET-22b	6 x His	NdeI XhoI	Amp C. Behrens
FEZ1 (227–290)	pET-22b	6 x His	NdeI XhoI	Amp C. Behrens
FEZ1 (227–295)	pET-22b	6 x His	NdeI XhoI	Amp C. Behrens
FEZ1 (225–295)	pET-22b	6 x His	NdeI XhoI	Amp C. Behrens
IcsB (235–494) (MCSI) IpgA full length (MCSII)	pETDuet-1	6 x His (I)	BamHI NotI (MCSI) NdeI XhoI (MCSII)	Amp C. Behrens
IcsB full length (MCSI) IpgA full length (MCSII)	pETDuet-1	6 x His (I)	BamHI NotI (MCSI) NdeI XhoI (MCSII)	Amp C. Behrens
IcsB full length (MCSI) IpgA full length (MCSII)	pETDuet-1	6 x His (I) StrepII C-terminal (MCSII)	BamHI NotI (MCSI) NdeI XhoI (MCSII)	Amp C. Behrens

(...)

DNA constructs used in this study (*continued*)

gene	vector	tag	restriction sites	resistance
IcsB full length (MCSI) IpgA full length (MCSII)	pETDuet-1	6 x His (I) StrepII N-terminal (MCSII)	BamHI NotI (MCSI) XhoI (MCSII)	Amp C. Behrens
IcsB full length (MCSI) (K. Kühnel)	pETDuet-1	6 x His (I)	BamHI NotI	Amp K. Kühnel
IcsB (216–494) (MCSI) IpgA full length (MCSII)	pETDuet-1	6 x His (I)	BamHI NotI (MCSI) XhoI (MCSII)	Amp C. Behrens
IcsB (293–494) (MCSI) IpgA full length (MCSII)	pETDuet-1	6 x His (I)	BamHI NotI (MCSI) XhoI (MCSII)	Amp C. Behrens
IcsB (185–494) (MCSI) IpgA full length (MCSII)	pETDuet-1	6 x His (I)	BamHI NotI (MCSI) XhoI (MCSII)	Amp C. Behrens
SCOC (78–159)	pET-28a	Strep II	NcoI XhoI	Kan C. Behrens
SCOC (78–159)	pGEX-4T1	GST	BamHI XhoI	Amp C. Behrens
SCOC (78–159) E93V/K97L	pET-28a	Strep II	NcoI XhoI	Kan C. Behrens
SCOC (78–159) L105M	pET-28a	Strep II	NcoI XhoI	Kan C. Behrens
SCOC (78–159) N125L/N125V	pET-28a	Strep II	NcoI XhoI	Kan C. Behrens

(...)

DNA constructs used in this study (*continued*)

gene	vector	tag	restriction sites	resistance
SCOC (78–159) R17E	pET-28a	Strep II	NcoI XhoI	Kan C. Behrens
SCOC (78–159) R99E	pET-28a	Strep II	NcoI XhoI	Kan C. Behrens
SCOC (78–159) R99E/R117E	pET-28a	Strep II	NcoI XhoI	Kan C. Behrens
SCOC (78–141)	pET-22b	6 x His	NdeI XhoI	Amp C. Behrens
SCOC (78–151)	pET-22b	6 x His	NdeI XhoI	Amp C. Behrens
SCOC (78–132)	pET-28a	6 x His	NdeI XhoI	Kan C. Behrens
SCOC (112–159)	pET-28a	6 x His	NdeI XhoI	Kan C. Behrens
SCOC full length	pGEX-4T1	GST	BamHI XhoI	Amp C. Behrens
VirG (52–758)	pET-28a	6 x His	NdeI XhoI	Kan C. Behrens
VirG (353–758)	pET-28a	6 x His with linker, Strep II C-terminal	NdeI XhoI	Kan C. Behrens
VirG (353–758)	pET-28a	Strep II C-terminal	NdeI XhoI	Kan C. Behrens
VirG (353–758)	pET-28a	6 x His with linker	NdeI XhoI	Kan C. Behrens
	pET-22b	6 x His		Amp Novagen
	pET-28a	6 x His, T7		Kan Novagen
	pETDuet-1	6 x His, S-Tag		Amp Novagen

(...)

DNA constructs used in this study (*continued*)

gene	vector	tag	restriction sites	resistance
	pGEX-4T1	GST		Amp K. Kühnel

2.1.7 Oligonucleotides

Table 2.9 lists the oligonucleotides used in this study.

Table 2.9: Oligonucleotides

Name	Sequence 5' to 3'
Arl1 Q181 NotI noStop rev	ATAGTTTAGCGGCCCGCCTGGCGACTTTTCAGTGTTTC
Arl1 T15 BamHI fwd	CGGGATCCTATGACACGTGAAATGCGTATCCTGATTC
Arl1 T15 NdeI fwd	GGAATTCATATGACACGTGAAATGCGTATCCTGATTCTGGGGC
FEZ1 ccd H226 NdeI fwd	GGAATCCCCATATGCACATGTCTGGG
FEZ1 ccd L290 Strep Stop XhoI rev	CCGCTCGAGTTATTTTTCGAACTGCGGGTGGCTCCAGCTGCCGCGGGCA CCAGCTTTTTCATCAGTTC
FEZ1 ccd L290 XhoI noStop rev	CCGCTCGAGGCTGCCGCGGGCACCAGCTTTTTCATCAGTTCTCGCTG
FEZ1 ccd L295 Strep Stop XhoI rev	CCGCTCGAGTTATTTTTCGAACTGCGGGTGGCTCCAGCTGCCGCGGGCA CCAGTTTCTCTTCCGCCTCTT
FEZ1 ccd L295 XhoI noStop rev	CCGCTCGAGGCTGCCGCGGGCACCAGTTTCTCTTCCGCCTCTTTTT
FEZ1 ccd M227 NdeI fwd	GGAATCCCCATATGTCTGGGTCTGAG
FEZ1 ccd R225 NdeI fwd	GGAATCCCCATATGAGGCACATGTCT
IcsB A235 BamHI fwd	CGGGATCCCGCTAATCAAAAAAAGACCCCTATG
IcsB I494 NotI rev	ATAAGAATGCGGCCCGCCTATATATTAGAATGAGAGTTATTCA
IcsB N216 BamHI fwd	CGGGATCCCACTTATTAAATTCAAAACAAGATCAAAATAACAACAAAA
IcsB Q293 BamHI fwd	CGGGATCCCAATTGGCAAATTATAAATTAATCAGTAAATCTGAAAA
IcsB S185 BamHI fwd	CGGGATCCCTCAGGTTACAGTGTCGATAC
SCOC ccd 112E NdeI fwd	GGAATCCCCATATGACACTGGAAGATC

(...)

Oligonucleotides (*continued*)

Name	Sequence 5' to 3'
SCOC ccd 132N XhoI rev	CCGCTCGAGTTAATTCTCCTCTTTTACG
SCOC ccd 159K XhoI rev	CCGCTCGAGTTATTTACGTTTGGATTTGGT
SCOC ccd M78 NcoI fwd	GAATTCATATGATGAATGCCGACATGGATG
SCOC ccd M78 NdeI fwd	GGAATCCCCATATGATGAATGCCGAC
SCOC ccd M78 pGex4T1 fwd	CGGGGATCCATGATGAATGCCGACATGGATGCC
SCOC ccd N141 XhoI rev	CCGCTCGAGTTAATTTTCGATATACTGGC
SCOC ccd T151 XhoI rev	CCGCTCGAGTTAGGTTTCAAAGACGCT
SCOC E93V fwd	TGCCGAAAATCAGGTGGTACTGGAGGAAAAAACC
SCOC E93V/K97L fwd	GCCGAAAATCAGGTGGTACTGGAGGAATTAACCCGTCTGATCAAC
SCOC E93V/K97L rev	GTTGATCAGACGGGTTAATTCCTCCAGTACCACCTGATTTTCGGC
SCOC E93V rev	GGGTTTTTTCCTCCAGTACCACCTGATTTTCGGCA
SCOC fl pGex4ti BamHI fwd	CGGGGATCCATGCGTCGCCGTGTGTTTTCT
SCOC fl pGex4ti stop XhoI rev	CCGCTCGAGTTATAATTTACGTTTGGATTTGGTATCGGTGG
SCOC Gateway ccd M78 BamHI fwd	CACCATGATGAATGCCGACATGGATGC
SCOC Gateway rev	TTTACGTTTGGATTTGGTATCGGTGGTTTGA
SCOC L105M fwd	ACCCGTCTGATCAACCAAGTTATGGAGCTGCAG
SCOC L105M rev	CTGCAGCTCCATAACTTGGTTGATCAGACGGGT
SCOC N125L fwd	GTGTCGATGCCGTAAGAGGAGCTACTGAAACTGAAAAGTGAGAATCA

(...)

Oligonucleotides (*continued*)

Name	Sequence 5' to 3'
SCOC N125L rev	TGATTCTCACTTTTCAGTTTCAGTAGCTCCTCTTTTACGGCATCGACAC
SCOC N132V fwd	GAATCTGAAACTGAAAAGTGAGGTTCAAGTGCTGGGCCAGTATATC
SCOC N132V rev	GATATACTGGCCAGCACTTGAACCTCACTTTTTCAGTTTCAGATTTC
SCOC R117E fwd	ACTGGAAGATCTGTCTGCCGAGGTCGATGCCGTAAAAGAGG
SCOC R117E rev	CCTCTTTTACGGCATCGACCTCGGCAGACAGATCTTCCAGT
SCOC R99E fwd	GTGGAAGTGGAGGAAAAAACCAGCTGATCAACCAAGTTCTGGAG
SCOC R99E rev	CTCCAGAACTTGGTTGATCAGCTCGGTTTTTTTCTCCAGTCCAC
SCOC Strep rev 1	CTCCAGCTGCCGCGCGGCACCAGTTTACGTTTGG
SCOC Strep XhoI rev 2	CCGCTCGAGTTATTTTTTCGAACTGCGGGTGGCTCCAGCTG
VirG R758 Strep XhoI rev	CCGCTCGAGTTATTTTTTCGAACTGCGGGTGGCTCCACCGCCGCGACTAC TCATTTGAGTAGACTCTTGATTG
VirG R758 XhoI rev	CCGCTCGAGTTAGCGACTACTCATTTGAGTAGACTCTTGATTG
VirG S353 NdeI fwd	GGAATTCATATGTCATCCATTCTGAAAATTATCAACAATGATTACACTG
VirG S353 TEV linker NdeI fwd	GGAATTCATATGCGTAAACGTGAAAATTTATATTTTCAGGGCTCATCCA TTCTGAAAATTATC
VirG synthGene A52 NdeI fwd	CTGGAATTCATATGGCTACACCTCTGAGTG
VirG synthGene R758 XhoI rev	CCGCTCGAGTTAACGACTACTCATCTGTGTC

2.2 Methods

2.2.1 Molecular cloning

2.2.1.1 PCR

Cloning was performed with standard methods. For PCR, Phusion High-Fidelity PCR kit was used and the manufacturer's manual was followed as below (see Table 2.10). Template DNA concentration was 10 ng/ μ L.

Table 2.10: PCR reaction 50 μ L

Component	volume (μ L)
water	32.5
5x HF	10
10 mM dNTP	1
10 μ M P fwd	2.5
10 μ M P rev	2.5
template	1
Phu	0.5

PCRs with primer pairs of \sim 59 $^{\circ}$ C melting temperatures were performed with basic temperature cycling. Long primers or primer pairs with larger differences in melting temperatures required touchdown temperature cycling. Temperature cycling comprised following steps:

1. initial denaturation
2. denaturation
3. annealing
4. elongation
5. final extension
6. hold

For standard procedure, steps 2–4 were repeated for 30 cycles. Annealing temperature was chosen according to primers. In case of Touchdown temperature cycling, the annealing temperature changed with every 5th step.

Table 2.11: PCR temperature cycling

Touchdown				Standard			
Step	T ($^{\circ}\text{C}$)	Length (s)	Repeat	Step	T ($^{\circ}\text{C}$)	Length (s)	Repeat
1	98	120	–	1	98	120	–
2	98	15		2	98	15	
3	70	30	4 x	3	57–62	30	30 x
4	72	30 s/kb		4	72	30 s/kb	
2	98	15		5	72	420	–
3	68	30	4 x	6	16	∞	
4	72	30 s/kb					
2	98	15					
3	66	30	4 x				
4	72	30 s/kb					
2	98	15					
3	64	30	4 x				
4	72	30 s/kb					
2	98	15					
3	62	30	4 x				
4	72	30 s/kb					
2	98	15					
3	60	30	4 x				
4	72	30 s/kb					
2	98	15					
3	58	30	4 x				
4	72	30 s/kb					
5	72	420	–				
6	16	∞					

2.2.1.2 DNA agarose gel electrophoresis and gel extraction purification

PCR products and DNA plasmids were visualized by 1 % and 0.8 % agarose gel electrophoresis, respectively. Samples for electrophoresis were mixed with 6 x loading dye (Fermentas). 1 x TBE was used as running buffer. Agarose gels were stained with GelGreen™ Nucleic Acid Gel Stain and sizes of DNA products were assessed with Fermentas DNA ladders (GeneRuler 1 kb or GeneRuler 100 bp Plus). Gel extraction was conducted with the NucleoSpin Extract II kit according to the manufacturer's instructions.

2.2.1.3 Restriction digest

Restriction digest reactions were carried out for one hour at 37 °C with enzymes from NEB in 40 µL reactions (see Table 2.12).

Table 2.12: Restriction digest reaction setup

digest	PCR	vector	control
Component	volume (µL)		
water	3.6	to 40 µL	20.7
PCR product	30	–	–
vector (~ 1µg/µL)	–	4 µg	–
plasmid	–	–	5
10 x buffer	4	4	3
enzyme 1	1	1	0.5
enzyme 2	1	1	0.5
BSA (if required)	0.4	0.4	0.3

2.2.1.4 Ligations

Ligations were performed with 100 ng of vector combined with 5–10-fold excess of insert. A final reaction volume of 10 µL was set up with 1 µL of T4 DNA Ligase and 1 µL of 10 x buffer. Ligation was conducted at room temperature for 30 min, or, if unsuccessful, over night at 4 °C.

2.2.1.5 Transformation

Heat shock Transformation Chemocompetent *E. coli* (100 μ L XL1-blue or BL21 (DE3)) cells were incubated with DNA for 20 min on ice. Cells were heat shocked at 42 °C for 45 s, then 900 μ L prewarmed LB or SOC medium was added. Bacteria recovered for 50 min at 37 °C, before they were spun down for 2 min at 5000 rpm in a table top centrifuge. Cells were plated out on LB agar plates containing appropriate antibiotics.

Electroporation Transformation For transformation of two plasmids, electrocompetent BL21 (DE3) were incubated on ice in 2 mm gap electroporation cuvettes for 20 min. 2 μ L of 100 ng/ μ L of each plasmid DNA were added. Cells were shocked with 2.5 kV in a Biorad Gene Pulser II. 900 μ L prewarmed SOC media was added and cells were recovered and plated out following the same protocol as above.

2.2.1.6 Plasmid purification

The NucleoSpin Plasmid Kit was utilized for small scale preparation of plasmid DNA. Plasmids were checked for successful cloning by control restriction digest (Section 2.2.1.3) or colony PCR.

For sequencing, mid scale plasmid purifications were conducted with the NucleoBond® Xtra Midi Kit according to the manufacturer’s instructions. Concentration of DNA was determined spectroscopically using a Nanodrop spectrophotometer (ND-1000 from NanoDrop Technologies Inc.).

2.2.1.7 Mutagenesis

Side-directed mutagenesis of SCOC was performed with the QuickChange Lightning site-directed mutagenesis kit. See table Table 2.13 and Table 2.14 for reaction setup and temperature cycling.

Table 2.13: Mutagenesis PCR reaction (25 μ L) setup

Component	volume (μ L)
water	17.18
10x buffer	2.5
dNTP mix	1

(...)

Mutagenesis PCR reaction (25 μ L) setup (*continued*)

Component	volume (μ L)
100 ng P fwd	1
100 ng P rev	1
template 50 ng	1.32
enzyme	1

Table 2.14: Mutagenesis PCR Temperature cycling

Step	Temperature ($^{\circ}$ C)	Length (s)	Repeat
1	98	120	
2	98	15	
3	57–62	30	30 x
4	72	30 s/kb	
5	72	420	
6	16	∞	

After PCR, parental DNA was digested by incubation with additional 1 μ L DpnI. Cells were transformed with DpnI-treated DNA according to the provided manual.

2.2.2 Expression and Purification**2.2.2.1 Expression analysis**

Expression conditions were tested in LB, TB and autoinducible media at different temperatures (18 $^{\circ}$ C, 25 $^{\circ}$ C, 30 $^{\circ}$ C and 37 $^{\circ}$ C). 200 mL cultures were inoculated with an over night culture and grown at 37 $^{\circ}$ C until OD \sim 0.5 at 600 nm. Expression in LB and TB was induced with 1 mM IPTG. After this, cultures were distributed to different temperatures. Samples for SDS-PAGE gel electrophoresis were taken after appropriate incubation times. 200 μ L of culture were spun down at maximum speed for 1 min in a microcentrifuge, resuspended in 50 μ L H₂O and augmented with SDS sample buffer. At the

last time point, 50 mL of the culture were harvested, resuspended in 5 mL buffer and lysed by sonication. After 10 min centrifugation in a table top microcentrifuge, 2 mL of the supernatant were incubated for 1 hour with appropriate beads. Samples for SDS-PAGE gel electrophoresis were taken of pellet, supernatant and beads.

2.2.2.2 Minimal expression for selenomethione-labelling

A colony from a freshly streaked plate was picked and grown overnight at 37 °C in LB. Minimal media expression cultures for incorporation of selenomethionine were set up according to Table 2.6 on page 22 and prewarmed at 37 °C. For inoculation the overnight culture was harvested at low speed and gently resuspended in H₂O. Minimal media cultures were grown until OD 0.3–0.5 at 600 nm. Then, the cultures were augmented with amino acid mix and solid selenomethionine (50 mg/L). After another 20 min incubation, expression was induced with 1 mM IPTG. Expression was carried out at 25 °C over night.

2.2.2.3 Minimal expression for ¹⁵N-labelling

A colony from a freshly streaked plate was picked and grown overnight at 37 °C in LB. Minimal media expression cultures for ¹⁵N-labelling were set up according to Table 2.5 on page 21 and prewarmed at 37 °C. For inoculation the overnight culture was harvested at low speed and gently resuspended in H₂O. Minimal media cultures were grown until OD 0.3–0.5 at 600 nm. Expression was induced with 1 mM IPTG and carried out at 25 °C over night.

2.2.2.4 Expression and purification of Strep-tagged SCOC (78–159) domain and mutants

A synthetic gene of human SCOC isoform 1 was ordered from Mr Gene (DNA sequence Section A.1). Strep-tagged SCOC ccd constructs were cloned with NcoI and XhoI restriction sites into vector pET-28a (Novagen) using the full length synthetic gene as template for PCR. The StrepII-Tag was added by a 2-step PCR strategy with two overlapping reverse primers. Single and double mutants of SCOC's ccd M78–K159 Strep-SCOC-pET28a were created by using the QuickChange Lightning site-directed mutagenesis kit. M78–K159 Strep-SCOC-pET28a constructs were transformed into BL21(DE3) by heat-shock protocol.

9 L of LB culture supplemented with 30 mg/L kanamycin were grown until OD ~0.6 at 600 nm. Expression was induced with 1 mM IPTG and cells were

harvested after 3 h incubation at 37 °C. Bacteria were spun down with a JS-4.2 rotor in a Beckman J6-MI centrifuge at 4,000 rpm, 4 °C for 20 min. The pellet was resuspended into 500 mM NaCl, 50 mM HEPES pH 7.5 or 500 mM NaCl, 30 mM HEPES pH 7.5. Cell suspension was frozen at –20 °C or directly used for protein purification. After thawing (if required), a PI tablet, DNa-seI, lysozyme, and 1 M MgCl₂ at a final concentration of 1 mM were added. The mixture was left with gentle stirring for 20 min at room temperature. Cells were lysed with three repetitions in a microfluidizer M-110L (Microfluidics Corporation) according to the provided manual. Cell debris was pelleted by 45 min centrifugation with a SS34 rotor in a Du Pont Sorvall centrifuge at 14000 rpm and 4 °C. The supernatant was filled in a 50 mL Superloop (GE Healthcare) and applied to a 5 mL StrepTrap column (GE Healthcare) connected to the Äkta Prime FPLC system or Äkta Purifier FPLC system. Supernatant was loaded onto the column at a flow rate of 1 mL/min, then the column was washed with 7 CV buffer at 1.5 mL/min. Affinity-purified protein complexes were eluted at 1 mL/min with 6 CV of buffer supplemented with 2.5 mM desthiobiotin. Elution fractions containing the protein were collected and pooled.

Pooled protein fractions were concentrated to ~ 4 mL with a concentrator (Vivaspin 20, 10000 MWCO; Sartorius) and spun for 5 min at maximum speed in a microcentrifuge at 4 °C. The sample was applied on a 5 mL sample loop and loaded on an equilibrated Superdex 75 16/60 column connected to the Äkta FPLC system, using 250 mM NaCl, 20 mM HEPES pH 7.5 as running buffer at a flow rate of 1 mL/min. The size exclusion buffer contained additional 5 mM TCEP when selenomethionine-labelled protein was purified. Samples for SDS-PAGE gel electrophoresis were taken of pellet, supernatant and flowthrough fractions and of all elution fractions containing protein. The concentrated protein was aliquoted, flash frozen in liquid nitrogen and stored at –80 °C.

2.2.2.5 Expression and purification of His-tagged SCOC constructs

Expression and purification of SCOC (78–132) and SCOC (112–159)

Overlapping parts of SCOC's ccd were cloned into pET-28a with NdeI and XhoI restriction sites using the synthetic SCOC gene as template for PCR. Due to time constraints, both constructs were expressed in minimal media with selenomethionine incorporation only.

9 L of M9 minimal media culture supplemented with 30 mg/L kanamycin were grown until OD ~0.5 at 600 nm. After addition of amino acid mix and

solid selenomethionine, expression was induced with 1 mM IPTG. Cells were harvested after over night incubation at 25 °C. Bacteria were spun down with a JS-4.2 rotor in a Beckman J6-MI centrifuge at 4,000 rpm, 4 °C for 20 min. The pellet was resuspended into buffer A (50 mM HEPES pH 7.5, 400 mM NaCl, 15 mM imidazole). A PI tablet, DNaseI, lysozyme, and 1 M MgCl₂ at a final concentration of 1 mM were added. The mixture was left with gentle stirring for 20 min at room temperature. Cells were lysed on ice by sonication with a Branson Sonifier 450 sonicator. Duty cycle was set to 50 and output control to the microtip limit, which resulted in an output of 20–30 %. Cycles of 30 s on/off were performed summing up to 2 min total sonication time. Cell debris was pelleted by 45 min centrifugation with a SS34 rotor in a Du Pont Sorvall centrifuge at 14000 rpm and 4 °C. The supernatant was filled in a 50 mL Superloop (GE Healthcare) and applied to a 5 mL HisTrap column (GE Healthcare) connected to the Äkta Prime FPLC system or Äkta Purifier FPLC system. Supernatant was loaded onto the column at a flow rate of 1 mL/min, then the column was washed with 30 CV buffer A at 1.5 mL/min. Affinity-purified protein was eluted with a gradient towards 50 % buffer B (50 mM HEPES pH 7.5, 400 mM NaCl, 1 M imidazole) at 1 mL/min. Elution fractions containing the protein were collected and pooled.

Pooled protein fractions were concentrated to ~4 mL with a concentrator (Vivaspin 20, 3000 MWCO; Sartorius) and then spun for 5 min at maximum speed in a microcentrifuge at 4 °C. The sample was applied on a 5 mL sample loop and loaded on an equilibrated Superdex 75 16/60 column connected to the Äkta FPLC system, using gf buffer (300 mM NaCl, 20 mM HEPES pH 7.5, 5 mM TCEP) at a flow rate of 1 mL/min. Samples for SDS-PAGE gel electrophoresis were taken of pellet, supernatant and flowthrough fractions and of all elution fractions containing protein. The concentrated protein was aliquoted, flash frozen in liquid nitrogen and stored at –80 °C.

Expression and purification of SCOC (78–141) and SCOC (78-K151)

C-terminally truncated constructs of SCOC's ccd were cloned into pET-22b with NdeI and XhoI restriction sites using the synthetic SCOC gene as template for PCR. The constructs were expressed differently but purified via the same protocol.

For expression of SCOC ccd (78–141), 9 L of LB culture supplemented with 100 mg/L ampicillin were grown until OD ~0.5 at 600 nm. Expression was induced with 1 mM IPTG. Cells were harvested after 3 h incubation at 37 °C.

For expression of SCOC (78–151), 9 L of autoinducible culture supplemented with 100 mg/L ampicillin were grown until OD ~0.7 at 600 nm. Cultures were

shifted to 18 °C and left for expression over night.

Bacteria were spun down with a JS-4.2 rotor in a Beckman J6-MI centrifuge at 4,000 rpm, 4 °C for 20 min. The pellet was resuspended into buffer A (50 mM HEPES pH 7.5, 400 mM NaCl, 15 mM imidazole). A PI tablet, DNaseI, lysozyme, and 1 M MgCl₂ at a final concentration of 1 mM were added. The mixture was left with gentle stirring for 20 min at room temperature. Cells were lysed with three repetitions in a microfluidizer M-110L (Microfluidics Corporation) according to the provided manual. The debris was pelleted by 45 min centrifugation with a SS34 rotor in a Du Pont Sorvall centrifuge at 14000 rpm and 4 °C. The supernatant was filled in a 50 mL Superloop (GE Healthcare) and applied to a 5 mL HisTrap column (GE Healthcare) connected to the Äkta Prime FPLC system or Äkta Purifier FPLC system. Supernatant was loaded onto the column at a flow rate of 1 mL/min, then the column was washed with 30 CV buffer A at 1.5 mL/min. Affinity-purified protein was eluted with a gradient towards 50 % B (50 mM HEPES pH 7.5, 400 mM NaCl, 1 M imidazole) at 1 mL/min. Elution fractions containing the protein were collected and pooled.

Pooled fractions were dialyzed against gf buffer (250 mM NaCl, 20 mM HEPES pH 7.5, 2 mM MgCl₂) over night at 4 °C. 200 µL Thrombin (1 U/µL) was added to the dialysis tube for cleavage of the His-Tag. On the next day, Thrombin activity was inhibited by the addition of Pefabloc to a final concentration of 0.4 mM. The protein solution was loaded onto a 5 mL GStap column (GE Healthcare) at 1 mL/min. The flowthrough containing the cleaved protein was fractionated and collected. The flowthrough fractions were concentrated to ~ 4 mL with a concentrator (Vivaspin 20, 3000 MWCO; Sartorius) and then spun for 5 min at maximum speed in a microcentrifuge at 4 °C. The sample was applied on a 5 mL sample loop and loaded on an equilibrated Superdex 200 16/60 column connected to the Äkta FPLC system, using gf buffer (250 mM NaCl, 20 mM HEPES pH 7.5, 2 mM MgCl₂) at a flow rate of 1 mL/min. Samples for SDS-PAGE gel electrophoresis were taken of pellet, supernatant and flowthrough fractions and of all elution fractions containing protein. The concentrated protein was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

2.2.2.6 Expression and purification of GST-SCOC fusion proteins

GST-tagged full length SCOC and GST-tagged SCOC (78–159) were cloned into pGEX-4T1 with BamHI and XhoI restriction sites using the synthetic SCOC gene as template for PCR.

For expression of both GST fusion proteins, 9 L of LB culture supplemented with 100 mg/L ampicillin were grown until OD ~0.5 at 600 nm. Expression was induced with 1 mM IPTG. Cells were harvested after 3 h incubation at 37 °C. Bacteria were spun down with a JS-4.2 rotor in a Beckman J6-MI centrifuge at 4,000 rpm, 4 °C for 20 min. The pellet was resuspended into buffer A (50 mM HEPES pH 7.5, 300 mM NaCl). A PI tablet, DNaseI, lysozyme, and 1 M MgCl₂ at a final concentration of 1 mM were added. The mixture was left with gentle stirring for 20 min at room temperature. Cells were lysed with three repetitions in a microfluidizer M-110L (Microfluidics Corporation) according to the provided manual. The debris was pelleted by 45 min centrifugation with a SS34 rotor in a Du Pont Sorvall centrifuge at 14000 rpm and 4 °C. The supernatant was filled in a 50 mL Superloop (GE Healthcare) and applied to a 5 mL Protino® GST/4B column (Macherey and Nagel) connected to the Äkta Prime FPLC system. Supernatant was loaded onto the column at a flow rate of 1 mL/min, then the column was washed with 10 CV buffer A at 1.5 mL/min. Affinity-purified protein was eluted with 12 CV buffer B (50 mM HEPES pH 7.5, 300 mM NaCl, 10 mM glutathione) at 1 mL/min. Elution fractions containing the protein were collected and pooled.

Pooled fractions were concentrated to ~ 4 mL with a concentrator (Vivaspin 20, 30000 MWCO; Sartorius) and then spun for 5 min at maximum speed in a microcentrifuge at 4 °C. The sample was applied on a 5 mL sample loop and loaded on an equilibrated Superdex 200 16/60 column connected to the Äkta FPLC system, using gf buffer (250 mM NaCl, 20 mM HEPES pH 7.5) at a flow rate of 1 mL/min. Samples for SDS-PAGE gel electrophoresis were taken of pellet, supernatant and flowthrough fractions and of all elution fractions containing protein. The concentrated protein was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

2.2.2.7 Expression and purification of Arl1 (15–181) Q79L

A synthetic gene of human Arl1 with a dominant active mutation Q79L was ordered from Mr Gene (DNA sequence in Section A.1). His-tagged Arl1 (15–181) was cloned with NdeI and XhoI restriction sites into vector pET-22b using the full length synthetic gene as template for PCR.

9 L of autoinducible media supplemented with 100 mg/L ampicillin were grown until OD ~0.7 at 600 nm. Expression was conducted over night at 18 °C. Bacteria were spun down with a JS-4.2 rotor in a Beckman J6-MI centrifuge at 4,000 rpm, 4 °C for 20 min. The pellet was resuspended into buffer A (50 mM NaH₂PO₄ pH 7.5, 300 mM NaCl, 10 mM imidazole, 5 mM MgCl₂).

Cell suspension was frozen at $-20\text{ }^{\circ}\text{C}$ or directly used for protein purification. After thawing (if required), a PI tablet, DNaseI and lysozyme were added. The mixture was left with gentle stirring for 20 min at room temperature. Cells were lysed on ice by sonication with a Branson Sonifier 450 sonicator. Duty cycle was set to 50 and output control to the microtip limit, which resulted in an output of 20–30 %. Cycles of 30 s on/off were performed summing up to 2 min total sonication time. Cell debris was pelleted by 60 min centrifugation with a SS34 rotor in a Du Pont Sorvall centrifuge at 14000 rpm and $4\text{ }^{\circ}\text{C}$. The supernatant was filled in a 150 mL superloop (GE Healthcare) and applied to a 5 mL HisTrap column (GE Healthcare) connected to the Äkta Prime FPLC system. Supernatant was loaded onto the column at a flow rate of 1 mL/min, then the column was washed with 30 CV buffer A at 1.5 mL/min. Affinity-purified protein was eluted with a gradient towards 50 % buffer B (50 mM NaH_2PO_4 pH 7.5, 300 mM NaCl, 1 M imidazole, 5 mM MgCl_2) at 1 mL/min.

Pooled protein fractions were concentrated to ~ 4 mL with a concentrator (Vivaspin 20, 10000 MWCO; Sartorius) and then spun for 5 min at maximum speed in a microcentrifuge at $4\text{ }^{\circ}\text{C}$. The sample was applied on a 5 mL sample loop and loaded on an equilibrated Superdex 200 16/60 column connected to the Äkta FPLC system running with gf A buffer (50 mM NaH_2PO_4 pH 7.5, 300 mM NaCl, 5 mM MgCl_2) at a flow rate of 1 mL/min. Pooled protein fractions were concentrated to ~ 8 mg/mL.

2 mL of the protein solution were mixed with excess EDTA (125 μL of a 0.2 M EDTA stock), 16 μL of shrimp alkaline phosphatase (1 U/ μL) and diluted with 20 mM HEPES pH 7.5 to 2.5 mL. GDPNHP, a non-hydrolyzable GTP derivate, was added to a final concentration of 10 mM. The sample was incubated in the cold room over night for nucleotide exchange. On the next day, excess MgCl_2 was added. GDPNHP-loaded Arl1 T15-Q181 Q79L was diluted to 5 mL with gf B buffer (30 mM Tris pH 8.0, 300 mM NaCl, 5 mM MgCl_2 , 5 mM DTT) and purified again over Superdex 200 16/60 column as before. Samples for SDS-PAGE gel electrophoresis were taken of pellet, supernatant and flowthrough fractions and of elution fractions containing protein. The concentrated protein was mixed with 20 % glycerol, aliquoted, flash frozen in liquid nitrogen and stored at $-80\text{ }^{\circ}\text{C}$.

2.2.2.8 Coexpression and copurification of SCOC–FEZ1 complexes

Human FEZ1 ccd constructs were cloned with NdeI and XhoI restriction sites into vector pET-22b (Novagen) using full length FEZ1 as template for PCR (gift by John Chua). Both FEZ1-pet22b and SCOC-pET-28a (78–159) were

co-transformed into BL21 by electroporation.

3 L of LB culture supplemented with 100 mg/L ampicillin and 30 mg/L kanamycin were grown until OD ~0.6 at 600 nm. Expression was induced with 1 mM IPTG and cells were harvested after 3 h incubation at 37 °C. Bacteria were spun down with a JS-4.2 rotor in a Beckman J6-MI centrifuge at 4,000 rpm, 4 °C for 20 min. The pellet was resuspended into 250 mM NaCl, 50 mM HEPES pH 7.5. Cell suspension was frozen at –20 °C or directly used for protein purification. After thawing (if required), a PI tablet, DNaseI, lysozyme, and 1 M MgCl₂ at a final concentration of 1 mM were added. The mixture was left with gentle stirring for 20 min at room temperature. Cells were lysed with three repetitions in a microfluidizer M-110L (Microfluidics Corporation) according to the provided manual. Cell debris was pelleted by 45 min centrifugation with a SS34 rotor in a Du Pont Sorvall centrifuge at 14000 rpm and 4 °C. The supernatant was filled in a 50 mL Superloop (GE Healthcare) and then applied to a 5 mL StrepTrap column (GE Healthcare) connected to the Äkta Prime FPLC system or Äkta Purifier FPLC system. Supernatant was loaded onto the column at a flow rate of 1 mL/min, then the column was washed with 7 CV buffer at 1.5 mL/min. Affinity-purified protein complexes were eluted at 1 mL/min with 6 CV of buffer supplemented with 2.5 mM desthiobiotin. Elution fractions containing the protein were collected and pooled.

For SEC-MALLS analysis, the protein complex was further purified by size exclusion chromatography. Pooled protein fractions were concentrated to ~ 4 mL with a concentrator (Vivaspin 20, 3000 MWCO; Sartorius) and spun for 5 min at maximum speed in a microcentrifuge at 4 °C. The sample was applied on a 5 mL sample loop and loaded on an equilibrated Superdex 75 16/60 column connected to the Äkta Prime FPLC system in the cold room, using 250 mM NaCl, 50 mM HEPES pH 7.5 as running buffer at a flow rate of 1 mL/min. Samples for SDS-PAGE gel electrophoresis were taken of pellet, supernatant and flowthrough fractions and of all elution fractions containing protein. The concentrated complex, either affinity- or gel-filtration purified, was prone to aggregation and precipitation, and was therefore freshly used for different applications.

2.2.2.9 Expression and denaturing purification of VirG (353–758)

Several methods and constructs were tried for expression and purification of VirG (353–758). Due to the formation of inclusion bodies, all purification protocols involved denaturation and refolding of the protein. VirG (353–758) with

different tags was cloned into pET-28a using the pWR100 virulence plasmid of *Shigella flexneri* (gift from Prof. A. Zychlinsky) as template for PCR.

GdnHCl-denaturing purification of C-terminal Strep-tagged VirG (353–758) with stepwise refolding

9 L of LB culture supplemented with 30 mg/L kanamycin were grown until OD ~0.6 at 600 nm. Expression was induced with 1 mM IPTG and cells were harvested after incubation over night at 30 °C. Bacteria were spun down with a JS-4.2 rotor in a Beckman J6-MI centrifuge at 4,000 rpm, 4 °C for 20 min. The pellet was resuspended into sonication buffer 20 mM NaH₂PO₄ pH 7.5, 300 mM NaCl, 1 mM EDTA, 2 mM benzamidine. Cell suspension was frozen at –20 °C or directly used for protein purification. After thawing (if required), a PI tablet, DNaseI, lysozyme and 1 M MgCl₂ at a final concentration of 1 mM were added. The mixture was left stirring in the cold room for 20 min.

Cells were lysed on ice by sonication with a Branson Sonifier 450 sonicator. Duty cycle was set to 50 and output control to the microtip limit, which resulted in an output of 20–30 %. Cycles of 30 s on/off were performed summing up to 2 min total sonication time. Insoluble fractions were harvested by centrifugation for 20 min in a SS34 rotor at 14000 rpm, 4 °C in a Du Pont Sorvall centrifuge. The pellet was resuspended in 300 mL sonication buffer and 1% Triton-X-100 was added. After 5 minutes of stirring, the pellet was again harvested and resuspended as before. Inclusion bodies were solubilized with 6 M GdnHCl by thorough homogenization. At 4 °C, the solution was slowly and drop-wise diluted with sonication buffer with three times its own volume of sonication buffer to 2 M GdnHCl while gently stirred. Debris and precipitate were pelleted as before. The remaining supernatant was dialyzed over night against 4 L sonication buffer supplemented with 1.5 M GdnHCl at 4 °C. On the next day, the solution was dialyzed for 5 h against 1 M and then 0.5 M GdnHCl sonication buffer. Final dialysis occurred over night against 0 M GdNHCl. Between all the dialysis steps, precipitate was removed by centrifugation as before.

The final supernatant was filtered through a 0.45 µm filter, filled into a 150 mL Superloop (GE Healthcare) and applied to a 5 mL StrepTrap column (GE Healthcare) connected to the Äkta Prime FPLC system. Supernatant was loaded onto the column at a flow rate of 1 mL/min, then the column was washed with 7 CV buffer at 1.5 mL/min. Affinity-purified protein was eluted at 1 mL/min with 6 CV of sonication buffer supplemented with 2.5 mM desthio-biotin. Elution fractions containing the protein were collected and pooled. Pooled protein fractions were concentrated to ~ 5 mL with a concentrator (Vi-

vaspin 20, 30000 MWCO; Sartorius) and spun for 5 min at maximum speed in a microcentrifuge at 4 °C. The sample was applied on a 5 mL sample loop and loaded on an equilibrated Superdex 200 16/60 column connected to the Äkta FPLC system, using 300 mM NaCl, 30 mM NaH₂PO₄ pH 7.5, 1 mM EDTA, 2 mM benzamidine and 2 mM DTT as running buffer at a flow rate of 1 mL/min. Samples for SDS-PAGE gel electrophoresis were taken of all pellet, supernatant and flowthrough fractions and of all elution fractions containing protein. The concentrated protein was aliquoted, flash frozen in liquid nitrogen and stored at −80 °C.

2.2.2.10 Expression and denaturing purification of VirG passenger domain (52–758)

VirG (52–758) was cloned into pET-28a from a synthetic gene optimized for *E. coli* expression. Due to the expression in inclusion bodies, the purification protocol involved denaturation and refolding of the protein. 4.5 L of LB culture supplemented with 30 mg/L kanamycin were grown until OD ~0.5 at 600 nm. Expression was induced with 1 mM IPTG and cells were harvested after incubation for 5 h at 37 °C. Bacteria were spun down with a JS-4.2 rotor in a Beckman J6-MI centrifuge at 4,000 rpm, 4 °C for 20 min. The pellet was resuspended into 30 mM Tris pH 7.4, 500 mM NaCl, 15 mM imidazole, 5 mM benzamidine. Cell suspension was frozen at −20 °C or directly used for protein purification.

After thawing (if required), a PI tablet, DNaseI, lysozyme and 1 M MgCl₂ at a final concentration of 1 mM were added. The mixture was left stirring in the cold room for 20 min. Cells were lysed on ice by sonication with a Branson Sonifier 450 sonicator. Duty cycle was set to 50 and output control to the microtip limit, which resulted in an output of 20–30 %. Cycles of 30 s on/off were performed summing up to 2 min total sonication time. Urea was added to a final concentration of 6 M urea. The sample was sonicated again with the same settings as above. Cell debris was pelleted by 60 min centrifugation with a SS34 rotor in a Du Pont Sorvall centrifuge at 14000 rpm and 4 °C. The supernatant was filtered through a 0.45 µm filter, filled into a 150 mL Superloop (GE Healthcare) and applied to a Protino® Ni-NTA Column 5 mL (Macherey and Nagel) connected to the Äkta Prime FPLC system. Supernatant was loaded onto the column at a flow rate of 1 mL/min, then the column was washed with 30 CV buffer A (30 mM Tris pH 7.4, 500 mM NaCl, 4 M urea, 15 mM imidazole, 5 mM benzamidine) at 2.5 mL/min. Affinity-purified protein was eluted with a gradient towards 40 % buffer B (30 mM Tris pH 7.4, 500 mM

NaCl, 2 M urea, 10 mM imidazole, 5 mM benzamidine) at 1 mL/min. Elution fractions containing the protein were pooled and collected.

VirGs passenger domain was further purified via anion exchange chromatography. Over night, pooled fractions were dialyzed against buffer AX A (100 mM NaCl, 30 mM Tris pH 8.0, 1 M urea, 1 mM EDTA, 5 mM benzamidine). The protein solution was loaded at 0.5 mL/min onto a 5 mL HiTrapQ FF column (GE Healthcare) via a 50 mL superloop. The column was washed with 5 CV buffer AX A. A gradient from 0–40 % buffer AX B (1 M NaCl, 30 mM Tris pH 8.0, 1 M urea, 1 mM EDTA, 5 mM benzamidine) at 1 mL/min eluted the protein from the column. Finally, VirG A52-R758 was polished in a size exclusion purification step. Pooled protein fractions were dialyzed against gf buffer (30 mM Tris pH 8.0, 400 mM NaCl, 500 mM urea, 1 mM EDTA, 5 mM benzamidine) at 4 °C, concentrated to ~ 5 mL with a concentrator (Vivaspin 20, 30000 MWCO; Sartorius) and then spun for 5 min at maximum speed in a microcentrifuge at 4 °C. The sample was applied on a 5 mL sample loop and loaded on an equilibrated Superdex 200 16/60 column connected to the Äkta Prime FPLC system at a flow rate of 1 mL/min. Samples for SDS-PAGE gel electrophoresis were taken of all pellet, supernatant and flowthrough fractions and of all elution fractions containing protein. The concentrated protein was supplemented with 20 % glycerol, aliquoted, flash frozen in liquid nitrogen and stored at –80 °C.

2.2.2.11 Expression and purification of IcsB–IpgA complexes

The virulence factor IcsB and its chaperone IpgA were cloned into the two multiple cloning sites (MCS) of pETDuet-1 with the *Shigella* virulence plasmid as template for PCR. IcsB was cloned with restriction sites BamHI and NotI into MCS I, IpgA was cloned with NdeI and XhoI sites into MCS II. C- or N-terminal StrepII-Tag was added to IpgA by PCR with the reverse primer. Complexes were affinity-purified via the His-Tag on IcsB and furthermore by size exclusion chromatography.

9 L of autoinducible media supplemented with 100 mg/L ampicillin were grown until OD ~0.7 at 600 nm. Cultures were incubated over night for expression at 18 °C. Bacteria were spun down with a JS-4.2 rotor in a Beckman J6-MI centrifuge at 4,000 rpm, 4 °C for 20 min. The pellet was resuspended into buffer A (50 mM NaH₂PO₄ pH 7.5, 300 mM NaCl, 30 mM imidazole). Cell suspension was frozen at –20 °C or directly used for protein purification. After thawing (if required), a PI tablet, DNaseI, lysozyme and 1 M MgCl₂ at a final concentration of 1 mM were added. The mixture was left stirring in the

cold room for 20 min. Cells were lysed on ice by sonication with a Branson Sonifier 450 sonicator. Duty cycle was set to 50 and output control to the microtip limit, which resulted in an output of 20–30 %. Cycles of 30 s on/off were performed summing up to 2 min total sonication time. Cell debris was pelleted by 60 min centrifugation with a SS34 rotor in a Du Pont Sorvall centrifuge at 14000 rpm and 4 °C. The supernatant was filtered through a 0.45 µm filter, filled into a 150 mL Superloop (GE Healthcare) and applied to a 5 mL HisTrap column (GE Healthcare) connected to the Äkta Prime FPLC system. Supernatant was loaded onto the column at a flow rate of 1 mL/min, then the column was washed with 30 CV buffer A at 1.5 mL/min. Affinity-purified protein was eluted with a gradient towards 50 % buffer B (50 mM NaH₂PO₄ pH 7.5, 300 mM NaCl, 1 M imidazole) at 1 mL/min. Elution fractions containing the protein were collected.

Pooled fractions were dialyzed against gf buffer (300 mM NaCl, 30 mM HEPES pH 7.5, 2 mM DTT) at 4 °C, concentrated to ~ 5 mL with a concentrator (Vivaspin 20, 30000 MWCO; Sartorius) and spun for 5 min at maximum speed in a microcentrifuge at 4 °C. The sample was applied on a 5 mL sample loop and loaded on an equilibrated Superdex 200 16/60 column connected to the Äkta Prime FPLC system at a flow rate of 1 mL/min. Samples for SDS-PAGE gel electrophoresis were taken of pellet, supernatant and flowthrough fractions and of all elution fractions containing protein. The concentrated protein was aliquoted, flash frozen in liquid nitrogen and stored at –80 °C.

For the complexes with shortened IcsB constructs (IcsB (235–494), (216–494), (293–494) and (185–494)), the same protocol was followed except that the gf buffer contained only 200 mM NaCl (30 mM HEPES pH 7.5, 200 mM NaCl, 2 mM DTT).

2.2.3 Biochemical methods

2.2.3.1 SDS-PAGE and Western blot

For routine applications, such as quality control of purification protocols, *SDS-PAGE* was used. Gels were made up of a resolving and a separating gel. The protocol for 8–15 % SDS-PAGE gels was adapted from Laemmli [91]. For analysis of smaller proteins, gels were prepared by a protocol from Schägger & von Jagow [92]. Composition of all buffers and solutions can be found in Section 2.1.5. Samples were mixed with 3 x sample buffer and heated to 95 °C for 3–5 min. SDS-PAGE gels were run at a constant current of 180 V until the blue front was leaking through. For electrophoresis of Schägger gels, a current

of 90 V was applied until the protein front reached the separating gel. Then the current was adjusted to 120 V.

Gels were stained with Coomassie Brilliant Blue with a fast staining protocol. The gel was placed in a microwavable plastic container, well covered with solution A and heated for 1 min at maximum power in the microwave. Solution A was replaced by solution B and heated again. The same procedure was followed with solution C and D. Gels were left shaking in destaining solution D for 5 min.

For *western blotting* with the semi-dry method [93], the gel and a nitrocellulose membrane were washed in water and then transfer buffer. Four filter papers were also soaked in transfer buffer. The sandwich of two filter papers, gel, membrane and two filter papers was applied to 45 mA. Transfer time varied depending on the protein size. Next, the membrane was washed with TBS-T and then blocked for 20 min with Blotto. Incubation with the first antibody diluted in Blotto occurred over night or for one hour. Again, the membrane was washed three times with TBS-T. The membrane was incubated with the second antibody for 45 min, thoroughly washed with TBS-T before detection with ECL solution (Perkin Elmer). For detection and analysis, a Imageready LAS-1000 CCD camera (Fujifilm) and AIDA software (Fujifilm) were used.

2.2.3.2 Limited proteolysis

Limited proteolysis was performed in search of stable fragments or domains of proteins. For this purpose, the Proti-Ace I + II (Hampton Research) Kits were utilized and the manufacturers instructions were followed. The kits contain twelve different proteases (α -Chymotrypsin, Trypsin, Elastase, Papain, Subtilisin, Endoproteinase Glu-C, Proteinase K, Endoproteinase Arg-C, Pepsin, Thermolysin, Bromelain, Actinase) that were diluted to a working concentration of 0.01 mg/mL in the provided buffer (10 mM HEPES pH 7.5, 500 mM NaCl). Proteases were mixed with the same volume of protein solution (1–10 mg/mL). Time course of the degradation process was followed by taking samples after different time points (e.g. 5, 10, 15, 10, 30, 45 and 60 min). Reactions in the samples were stopped by the immediate addition of sample buffer. Samples were analyzed by SDS-PAGE.

2.2.3.3 N-terminal sequencing

Stable fragments found by limited proteolysis were sent to SeqLab (Sequence Laboratories Göttingen GmbH). The first five amino acids of a fragment were

determined by Edman's degradation. For preparation of the fragment, the manufacturers instructions were followed. The protocol included western blotting of the fragment on a PVDF membrane. The PVDF membrane was activated for 20 s in methanol before washing it in transfer buffer. For transfer, CAPS buffer (10 mM CAPS pH 11, 10 % methanol) was used. After blotting, the membrane was washed with water and methanol and then stained with Coomassie Blue R-250 (0.1 % (w/v) Coomassie Blue R-250, 1 % acetic acid, 40 % methanol). Bands were visualized by destaining with 50 % methanol. The membrane was washed with water and dried. Relevant protein bands were cut out with a scalpel and sent to SeqLab in an Eppendorf tube.

2.2.3.4 GST-Pulldown

Pulldown of immobilized GST-SCOC-ccd was applied to find novel interaction partners from HEK lysate. The pulldown procedure was modified from a Nature Methods protocol by Einarson based on Einarson [94]. Two reactions of 50 μ L of glutathione SepharoseTM 4 Fast Flow beads (GE Healthcare), same amounts of cleared HEK lysate and 10 μ g GST or GST-SCOC-ccd, respectively, were incubated at 4 °C in a head over end shaker. The two samples were spun down for 5 min at 5000 rpm in a microcentrifuge. Beads were washed 4 times with ice-cold PBS buffer and centrifuged again. Samples of supernatant, wash fractions and beads were analyzed by SDS-PAGE gel electrophoresis. The gel with bands of pulled protein was given to the mass spectrometry facility for further analysis.

2.2.3.5 Nucleotide-dependent Arl1-Pulldown

Table 2.15: Composition of Arl1-Pulldown reactions

component			control	GTP	GTP γ s	GDP
His-Arl1	Q79L	(15–181)	200 μ L	200 μ L	200 μ L	200 μ L
			–	20 μ L	20 μ L	20 μ L
Strep-SCOC		(78–159)	30 μ L	30 μ L	30 μ L	30 μ L
buffer			770 μ L	750 μ L	750 μ L	750 μ L

To check whether SCOC's ccd is interacting with Arl1 in a nucleotide-dependent manner, a pulldown experiment was performed. Four reactions were set up according to Table 2.15.

Arl1 was preincubated with an excess of the respective nucleotide on ice for 20 min. Then, SCOC (78–159) was added and the reactions were left on a rotary shaker at 4 °C over night. After the addition of 50 μ L of Ni²⁺-sepharose beads to each tube, the mixtures were left for another two hours for binding on the rotary shaker. Beads were pelleted by 2 min centrifugation at 2000 rpm in a microcentrifuge. The flowthrough was collected and beads were washed twice with 2 mL ice-cold buffer. Protein complexes were eluted with buffer B. Samples of supernatant, flowthrough, wash and beads were analyzed by SDS-PAGE.

2.2.3.6 Isothermal titration calorimetry

Isothermal titration calorimetry can reveal insights into thermodynamic features of an interaction of a protein and an interaction partner. Binding affinities, stoichiometry and enthalpy can be determined.

The ITC200 (Microcal) was used for measurements of SCOC and Arl1. This instrument holds a cell of 200 μ L volume and a syringe of 40 μ L. The concentration of the component in the cell should be ~10 times higher than in the cell. SCOC (78–159) in the syringe was applied at a concentration of 615 μ M and GDPNHP-loaded Arl1 (15–181) at 50 μ M. Both proteins were diluted from stock solutions of higher concentration with ITC buffer (20 mM HEPES pH 7.5, 150 mM NaCl). Stock solutions were dialyzed in the same buffer over night before. Standard settings of the ITC200 were used. 25 total injections of 2 μ L were performed with 120 s delay. Temperature was set to 20 °C, stirring speed to 300 rpm and reference power to 7 μ cal/s.

2.2.3.7 Circular Dichroism spectroscopy

Circular Dichroism spectroscopy (CD) was used to analyze secondary structure and stability of proteins.

CD protein samples were dialyzed to buffers containing NaF instead of NaCl. For SCOC ccd constructs, 20 mM NaH₂PO₄ pH 7.5, 250 mM NaF, for VirG S353-R758 30 mM NaH₂PO₄ pH 7.4, 300 mM NaF were used. Measurements of VirG's passenger domain were performed in the latter buffer supplemented with 0.5 M urea. For each spectrum, three repetitions and a buffer baseline were collected. Protein concentration was usually between 15–30 μ M.

Measurements were done with a Chirascan Circular Dichroism spectrometer (Applied Photophysics) using a Hellma quartz cuvette with a path length of 0.1 cm. Far UV CD spectra of SCOC were recorded between 200 and 260 nm with a step size of 0.5 nm, a band width of 1.5 nm and an averaging time of 5 s at either 20 °C or 93 °C. Thermal melts were carried out from 20 °C to 93 °C at 208 nm with a heating rate of 0.5 °C/min. Bandwidth was 1.5 nm and the averaging time was 1.5 s. For analysis of VirG's secondary structure, bandwidth was set to 2 nm, thermal melts were carried out at 216 nm. Data were analyzed with the manufacturer's ProView Software. Melting curves were fitted to a sigmoid shape and melting points were determined as the maxima of the first derivative of the function $f(x) = Ab + ((At - Ab)/(1 + \exp((x_0 - x)/w))) + m$.

2.2.3.8 SEC-MALLS

The oligomerization state of SCOC's ccd and its mutants were analyzed by Size exclusion chromatography coupled to a Multi Angle Laser Light Scattering (SEC-MALLS). All proteins loaded on the column were affinity and gel filtration purified before.

A Superdex 10/300 GL was connected to an Eclipse 2 system from Wyatt Technology with a DAWN EOS multi-angle light scattering setup and an Agilent 1100 series HPLC pump. The instruments were equilibrated with buffer (20 mM HEPES pH 7.5, 250 mM NaCl, 0.1 μ M filtered) until the baseline of the MALLS detector perpendicular to the laser (detector 11) was stable. For determination of the current protein concentration, a variable wavelength detector (Agilent) was connected to the MALLS instrument. The wavelength was set to 280 nm and the instrument was balanced (autozero function). 500 μ L of \sim 1.5 mg/mL protein sample were injected manually into a 500 μ L sample loop. Data collection was started, then the sample loop valve was turned manually from "load" to "inject". Scattering data and absorption at 280 nm were recorded for 60 min at a flow rate of 0.5 mL/min. Scattering data were analyzed with the manufacturer's ASTRA software. Molecular weights were determined from three SEC-MALLS experiments.

2.2.3.9 Crystallization

Screening

Initial screening for crystallization conditions were done in 96-well format with purchased crystallization screens. Sitting drops with two different pro-

tein concentrations were pipetted into the two wells of MRC plates (Hampton Research) by the Cartesian Mycosys robot (Cartesian Dispensing Systems). Routinely, a method setting 200 nL drops of 100 nL protein solution and 100 nL crystallization buffer was used. After complete dispensing of the robot, plates were rapidly covered with transparent sealing tape to prevent drying of the crystal drops. Plates were stored at 20 °C for SCOC screening plates and at 4 °C for VirG and IcsB proteins in a Formulatrix robot for automated imaging. Available crystallization screens included $(\text{NH}_4)_2\text{SO}_4$, Anions, Cations, ClassicLite, Classics I+II, ComPas, MbClass I+II, JCSG+, PACT, PEGI+II, pH-clearI+II, ProComplex (all Qiagen), SaltRx and Index (Hampton Research).

Refinement of crystallization conditions

First hits were refined with different strategies.

With the Tecan robot, strategies for refinement with *96-well format Grid Screens and Random Screens* were done. The Tecan robot dispensed newly customized screens from stock solutions into master blocks with one 1 mL reservoirs. Stock solutions were either purchased (Hampton Research or Qiagen) or self made from ultrapure chemicals. 96-well MRC plates were filled with the new crystallization conditions. With this method, rapid refinement of pH and precipitant concentration is possible, but with random screens also the necessity of the presence of several components can be addressed.

Larger crystals for data collection at the synchrotron were usually grown in *24-well format*. In 24-well grid screens, pH and precipitant concentration of a known crystallization condition were varied along smaller increments. Hanging drops were set up in self greased 24-well Linbro plates (Jena Bioscience) with Bayer medium viscosity silicon grease (Jena Bioscience) or in pre-greased Crystalgen SuperClearT Plates (Jena Bioscience). 500–1000 μL reservoirs were set up from self-made screening buffers from ultrapure chemicals or stock solutions purchased from Hampton Research. Up to 4 crystal drops made up of 1 μL protein and 1 μL reservoir solution were pipetted on a siliconized cover slide (22 mm diameter). Slides were turned upside down and fitted tightly onto the grease above a reservoir.

Strike seeding was used for optimization of crystal growth and diffraction quality. A crystal or a crystalline structure like a cluster was smashed with an acupuncture needle in mother liquor. A newly set drop was streaked through with the acupuncture needle to provide nucleation sites for crystal growth.

***In situ* proteolysis**

Crystals for the structure determination of SCOC's ccd were obtained by *in situ* proteolysis. Crystal drops were supplemented with small amounts of protease to cleave flexible parts of the protein in order to improve diffraction quality.

For *in situ* proteolysis crystallization, 3 mg/ml of selenomethionine-labelled L105M SCOC (78–159) were mixed with subtilisin in a 1:2000 (w/w) ratio and kept on ice until setting of the crystallization plates. Crystal drops were streak seeded and set at 3 and 1.5 mg/mL. Crystals were grown in hanging drops using Linbro plates at 20 °C by mixing 1 μ L protein with 1 μ L of the precipitant containing 20 mM sodium acetate pH 4.7, 1 M 1,6-hexanediol and 10 mM CoCl₂. Crystals were soaked in mother liquor supplemented with 20 % PEG 400 and flash cooled in liquid nitrogen.

2.2.3.10 Flash cooling of protein crystals

X-ray diffraction data of protein crystals are collected at 100 K in order to protect the crystal from radiation damage. Hence, crystals are flash cooled in liquid nitrogen before data collection. The formation of ice in a protein crystal leads to loss of diffraction quality. Soaking of crystals in cryoprotectants prevents ice formation. Cryoprotectants used in this study are e.g. glycerol, ethylene glycol, sugars (xylitol and sucrose) and PEG 400. Cryoprotectant mixes were composed in a way that the cryoprotectant replaced water from the initial crystallization condition, so that buffer and precipitant concentration remained constant during soaking. Crystals fished from their drops were first soaked in a fresh drop of mother liquor, then in a drop mixed 1:1 of mother liquor and cryoprotectant and finally in a cryoprotectant drop. Crystals were fished out of the cryoprotectant and rapidly flash cooled.

2.2.4 Computational methods

2.2.4.1 Structure determination of SCOC ccd

X-ray diffraction data were collected at 100 K at beamline X10SA (Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland). Data were processed and scaled with the XDS software package [95]. The package consists of three programs, XDS, XSCALE and XDSCONV. XDS performs eight subroutines. XYCORR, INIT and COLSPOT are used for determination and correction of spots versus the background of the dataset. Spots are indexed by IDXREF and integrated by INTEGRATE. During the final CORRECT step, unit cell

dimensions are refined and the space group determined). A XDS input file specific for the Pilatus detector at beamline X10SA can be found in the Appendix Section A.4.3. Values that need to be provided include detector specifications as well as unit cell dimensions and a space group—if unknown, CORRECT identifies for a space group. Furthermore, resolution cutoff and other values can be customized according to the respective dataset. XSCALE is used for scaling and merging of data. XDSCONV converts to output formats specific for the following programs, e.g. SHELX, CNS or CCP4, CCP4_F or CCP4_I format. Processed data was converted to .mtz file format.

With the experimental data, intensities of the reflections are provided. The goal for successful structure determination is the calculation of an electron density map which can be obtained by Fourier synthesis methods. A function of the electron density is given as:

$$\rho(x, y, z) = \frac{1}{V} \sum_{-h}^h \sum_{-k}^k \sum_{-l}^l |F_{hkl}| \exp[-2\pi i(hx + ky + lz - \alpha_{hkl})] \quad (2.1)$$

Structure factors F_{hkl} are experimentally obtained via the measured intensities of the reflection spots. Phases α_{hkl} cannot be measured directly but must be derived. The *phase problem* can be solved via direct methods (suitable for small proteins with excellent diffraction data above 1.2 Å), molecular replacement or experimental phasing. Molecular replacement requires the availability of a structurally similar model. Ideally at least 30 % sequence identity of the target and model is required. In the case of SCOC, no structure with sufficient sequence identity is deposited in the PDB, hence experimental phasing by single anomalous diffraction (SAD) was chosen as method structure determination. In a SAD experiment, differences between Friedel pairs are measured at a specific wavelength where anomalous diffraction occurs. This requires the presence of an anomalous scatterer, in this case selenium, which was incorporated as selenomethionine. From these differences, the positions of the anomalous atom positions can be determined by Patterson methods or direct methods. Phenix [96]), which was utilized in this study, contains the program AutoSol for finding anomalous atom positions. AutoSol uses hybrid substructure search (HYSS), which is a combination of both Patterson and direct methods. After determination of the Se positions, PHASER and RESOLVE, two other subprograms of Phenix, were used for refining and density modification of the initial map. Initial model building was performed with AutoBuild. Successive manual model building was done with Coot8 [97].

During refinement, which was performed with Phenix, the atomistic model was compared to the observed diffraction data, while taking into account the geometrical restraints. Rotamer outliers of side chains and other geometric constraints were identified and corrected with Coot. Figures were prepared with PyMOL [98].

2.2.4.2 Sequence alignment of SCOC

For the analysis of SCOC ccd conservation, a BLAST search was performed with SCOC isoform 1. Several sequences from 30 % identity upwards were picked and aligned with T-Coffee. Furthermore, human isoform 1, 2 and 3, and a close mammalian homologue (rat SCOC) were included. Analysis of the degree of conservation was done with AMAS [99]. Figures were prepared with PyMOL [98].

3 SCOC and its interaction partners

3.1 Characterization & structure determination of the SCOC coiled coil domain

3.1.1 Generation of SCOC constructs for crystallization

Structure determination of SCOC ccd was a major goal of this study. In general, proteins for crystallization can be engineered at two levels: at the *DNA level* by modifying the recombinant DNA sequence, or at the *protein level*, adding modifications to the protein by (bio)chemical methods such as limited proteolysis. This chapter deals with strategies utilized to engineer the protein on the DNA level, while protein modifications and optimization of crystallization conditions will be explained in the following chapters. The design and cloning of truncated versions of SCOC's ccd was one main pillar to solve the SCOC ccd structure. Figure 3.1 gives an overview of the constructs used for crystallization.

SCOC is a small protein with a well structured, highly conserved ccd (see Section 3.1.6). I performed initial bioinformatic analysis to characterize the SCOC ccd. When analyzing SCOC isoform 1 with the coiled coil prediction program COILS [100], a C-terminal coiled coil domain was confirmed, starting around residue 80 and ending latest with residue 144 (see Section A.4.1 for complete COILS prediction results). Further analysis with IUpred [101, 102] strengthened the prediction: A ordered region of unknown function from amino acid 89 to 149 was located (see Figure 3.2). IUpred also detected another ordered region at the N-terminus (residues 1–54). Between the two regions, SCOC isoform 1 was predicted to contain a unstructured stretch. The extend of assumed flexibility differed between long and close distance disorder tendency. In addition, also the very C-terminus from residue 150 to 159 is predicted to be unstructured.

Flexible or disordered parts of a protein lead to a variety of energetically

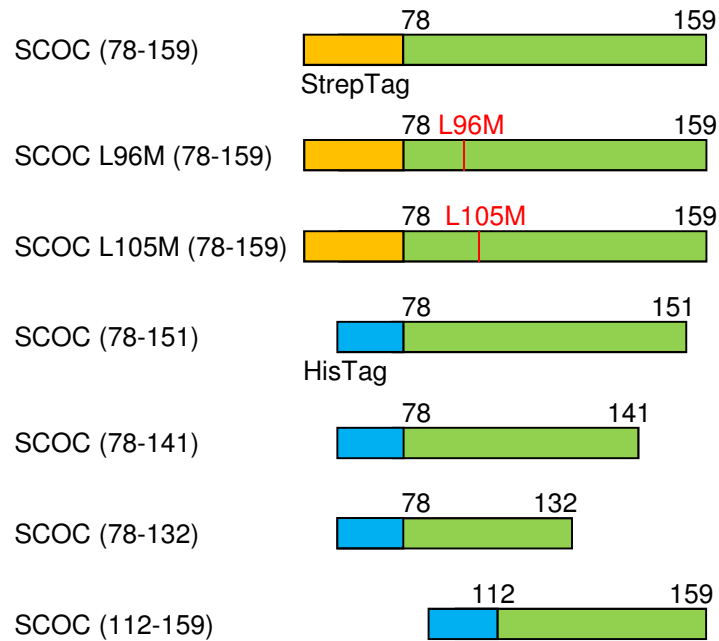


Figure 3.1: Overview of SCOC constructs used for crystallization

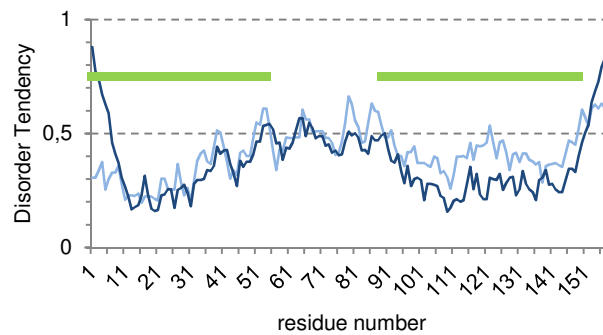


Figure 3.2: Disorder prediction for SCOC Isoform 1 by IUpred [101, 102]

Domains of unknown functions domains are labelled green. The dark blue line shows the disorder tendency for long disorder (~30 residues) and light blue for short disorder at a threshold of 0.5.

equally favored conformational states, that the protein can adopt. Heterogeneity of conformational states in a protein sample can cause problems with crystallization. As a result, the protein might not crystallize at all or formed crystals diffract poorly. The homogeneity of a protein's conformational state is a crucial determinant for successful protein crystallization. Therefore, I aimed to crystallize two C-terminally truncated ccd constructs, SCOC (78–151) and SCOC (78–141). Both of them were designed on the basis of IUpred predictions, to avoid problems arising from the more flexible C-terminus.

Although SCOC ccd is a small protein, the elongated form of the a long ccd and potential conformational flexibility hamper crystallization. Therefore I also created two constructs comprising overlapping halves of the coiled coil domain SCOC (78–132) and SCOC (112–159) in order to facilitate crystallization. Structure determination of the two overlapping halves of the coiled coil domain would allow assembly of the complete coiled coil domain.

When solving a protein crystal structure by X-ray crystallography, every crystallographer comes across the *phase problem* (see Section 2.2.4.1 for more details). In order to obtain an electron density map, structure factor amplitudes F_{hkl} and the phase angles a_{hkl} are needed. Structure factors are directly obtained by the diffraction experiment, as they are proportional to the square root of the measured intensities of reflections spots. Phases, however, cannot be directly determined from a simple diffraction experiments but must be supplied via specific *experimental methods*, like single- or multi wavelength anomalous diffraction, or by *molecular replacement*, where a similar known structure with a minimum of 30 % sequence can be used as a search model for determination of the new protein structure.

I subjected the sequence of SCOC (78–159) to analysis through HHpred server [103–105]. HHpred looks for homologues of the search model using hidden Markov models, searching through sequence alignment databases such as Pfam or SMART instead of sequence databases like UniProt. Protein Data Bank (PDB) numbers of homologues with known structure are returned in the results, making HHpred a powerful tool when looking for models to solve structures by *molecular replacement*. The closest protein with solved structure found by HHpred shared only 19 % identity with SCOC's ccd (see Section A.4.2 for a text file of HHpred results). This was not sufficient for molecular replacement. Hence, it became obvious that the phases for structure solvation had to be obtained through experimental methods. SCOC (78–159) contains four methionines which can be labelled with Se by minimal expression with selenomethionine, making *single and multi-wavelength anomalous diffraction (SAD and MAD)* a good option for obtaining the phases. However, three of the four me-

thionines are located within the first ten amino acids of the construct. Thus, they might be flexible residues, giving only little or no anomalous signal. Two mutants were created with additional methionine sites, SCOC L96M (78–159) and L105M (78–159), to facilitate structure determination by SAD or MAD.

TarO, a bioinformatic tool for the prediction of likeliness of crystallization [106], classified SCOC (78–159) as “recalcitrant” with a rather low crystallization index of $1.76e+6$. Nevertheless, due to its highly ordered structure and its small size, the ccd seemed to be an ideal target for crystallization at first glance.

3.1.2 Expression and Purification of SCOC constructs

3.1.2.1 Expression and Purification of SCOC (78–159)

Full-length SCOC isoform 1 with any other tag than GST proved to express only in low amounts and to be very unstable, prone to aggregation and degradation. Therefore, I aimed for expressing and purification of the ccd shared by all SCOC isoforms.

I successfully cloned SCOC (78–159) into pET-28a, thereby cleaving the plasmid’s His-Tag and adding a StrepTag II by a two-step PCR protocol. By this approach I obtained a Strep-tagged protein in a plasmid with kanamycine resistance, which was suitable for coexpression and proper analysis with His-tagged Arl1/Fez1 expressed from ampicillin-resistant pET-22b plasmids.

Expression behavior was analyzed in LB, TB and autoinducible media at different temperatures. SCOC (78–159) was expressed as soluble protein in large amounts in LB medium at 37 °C. After expression for 3 hours, cells were harvested and lysed. SCOC (78–159) was purified in two steps via affinity chromatography and size exclusion chromatography (see Figure 3.3). As StrepTagII-affinity purification is very specific, the protein was already pure after the first purification step (see Figure 3.3 C). A polishing step with size exclusion chromatography was performed to ensure the absence of protein aggregates for crystallization.

SCOC (78–159) concentrated to high concentrations (30 mg/mL) without precipitation. For crystallization, it was used at concentrations of 1–10 mg/mL. In general, SCOC (78–159) purification proved to be very robust, producing high yields of very pure protein.

SCOC (78–159) single mutants (L96M, L105M, R99E, R117E) and double mutants (R99E/R117E, E93V/K97L, N125V/N132L) were expressed and purified via the same protocol. Expression behavior, stability and purity after

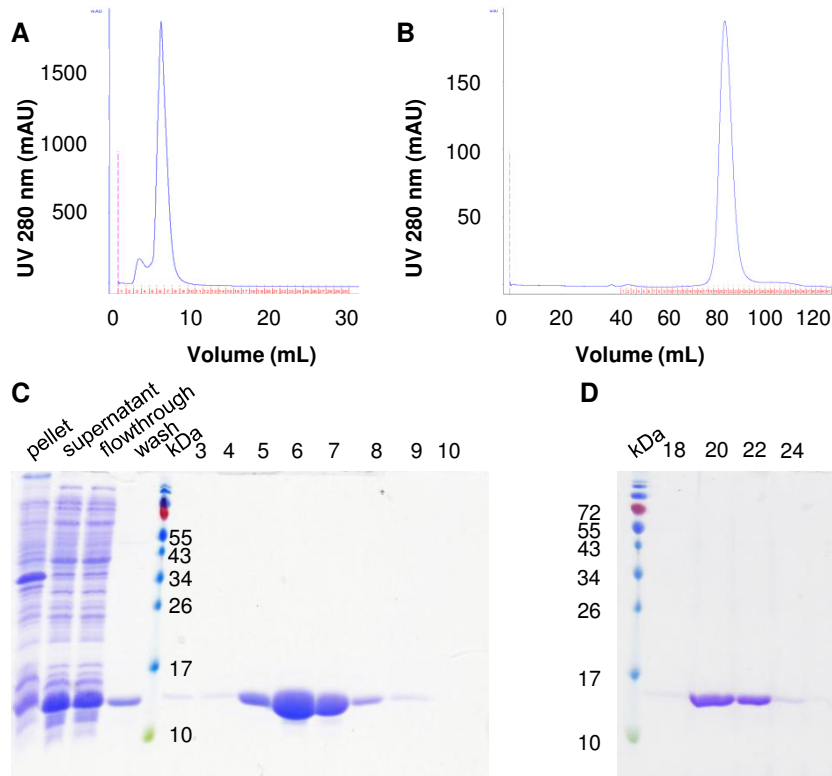


Figure 3.3: Purification of SCOC (78–159)

(A) Elution profile of affinity chromatography via StrepTrap column
(B) elution profile of size exclusion chromatography via HiLoad 16/60 S75 column
(C) SDS-PAGE of affinity chromatography
(D) SDS-PAGE of size exclusion chromatography

purification steps remained the same.

3.1.2.2 Expression and Purification of SCOC (78–132) and SCOC (112–159)

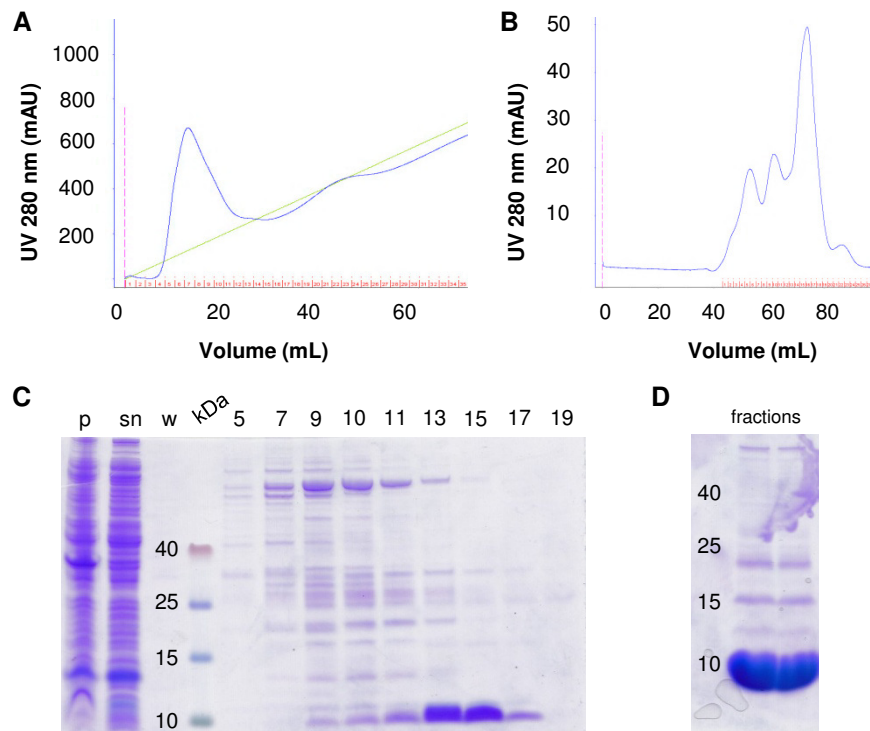


Figure 3.4: Purification of SCOC (78–132)

(A) Elution profile of affinity chromatography via HisTrap column (B) elution profile of size exclusion chromatography via HiLoad 16/60 S200 column (C) Schagger gel of affinity chromatography (D) concentrated fractions from size exclusion chromatography

Two overlapping fragments of SCOC ccd were cloned into pET-28a. The fragments (78–132) and (112–159) were expressed in minimal media for selenomethionine incorporation. After overnight expression at 25 °C, cells were harvested and lysed by sonication. SCOC fragments were purified in two steps via HisTrap chromatography and a polishing step with size exclusion chromatography. Elution profiles and SDS-PAGE analysis of both purification steps are depicted in Figure 3.4 for SCOC (78–132) and in Figure 3.5 for

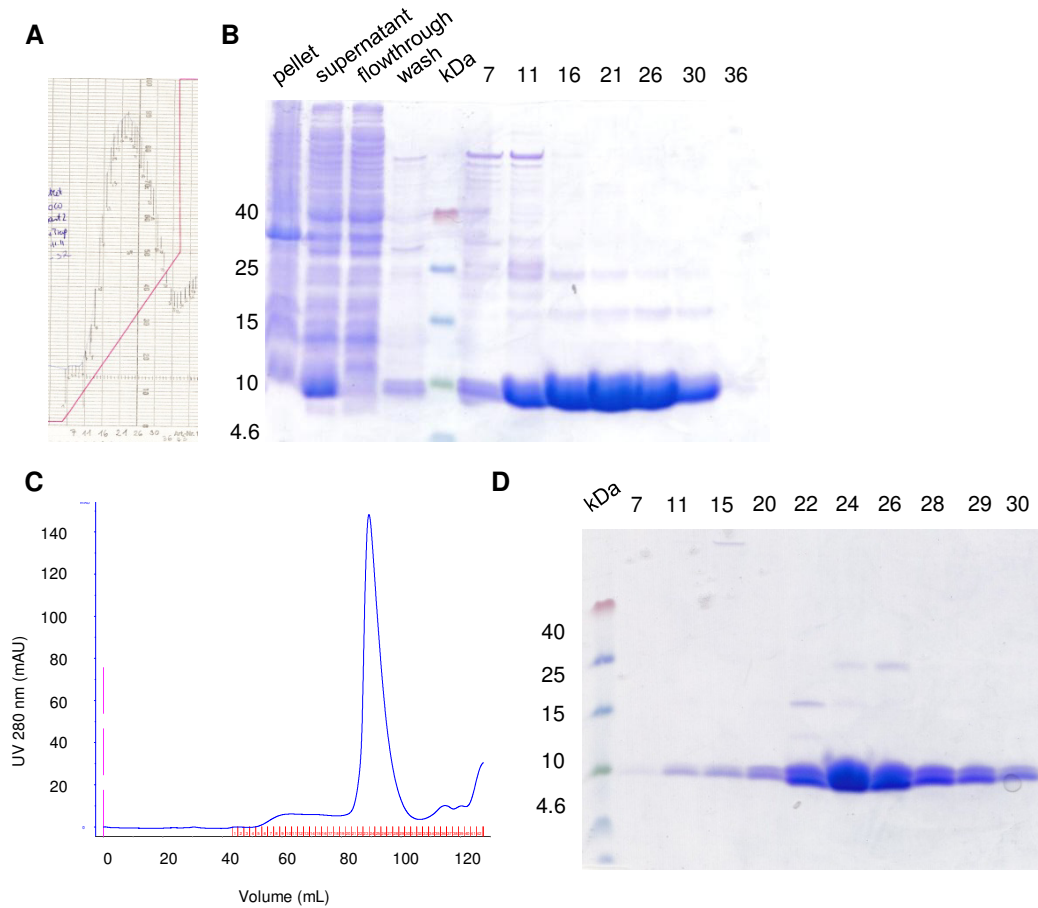


Figure 3.5: Purification of SCOC (112–159)

(A) Elution profile of affinity chromatography via HisTrap column
 (B) Schagger gel of affinity chromatography (C) elution profile of size exclusion chromatography via HiLoad 16/60 S200 column (D) Schagger gel from size exclusion chromatography

SCOC (112–159). Both proteins were concentrated and used for crystallization screening.

3.1.2.3 Expression and Purification of SCOC (78–141) and SCOC (78–151)

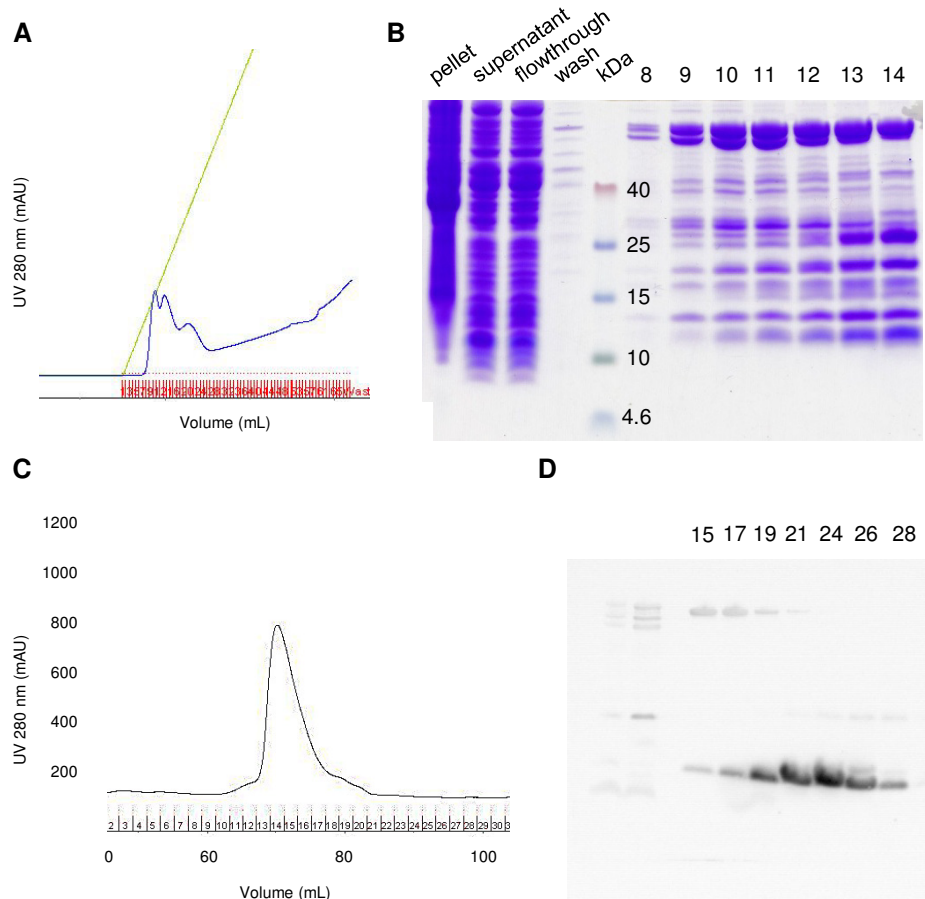


Figure 3.6: Purification of SCOC (78–141)

- (A) Elution profile of affinity chromatography via HisTrap column
 (B) Schagger gel of fractions from affinity chromatography (C) elution profile of size exclusion chromatography via HiLoad 16/60 S200 column
 (D) anti-His western blot of size exclusion chromatography

C-terminally truncated constructs were cloned into pET-22b. SCOC (78–141) was expressed in LB at 37 °C, SCOC (78–151) was expressed in au-

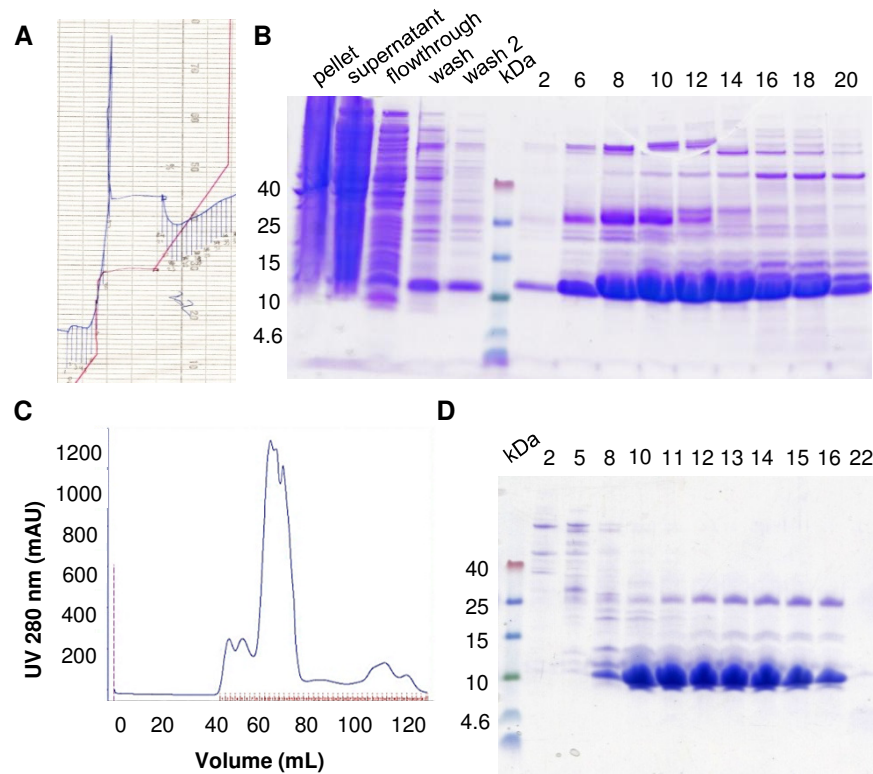


Figure 3.7: Purification of SCOC (78–151)

(A) Elution profile of affinity chromatography via HisTrap column
(B) Schagger gel of affinity chromatography (C) elution profile of size exclusion chromatography via HiLoad 16/60 S200 column (D) Schagger gel of size exclusion chromatography

toinducible medium at 18 °C. After lysis of the cells, proteins were purified via HisTrap chromatography, yielding contaminated proteins in large amounts. The HisTag was cleaved by Thrombin in an overnight incubation and removed by a second step of HisTrap chromatography. Protein aggregates were removed in a polishing step via size exclusion chromatography. Peaks observed in SCOC (78–141) elution profiles are less significant as the protein did not contain aromatic residues for UV absorption. Elution profiles and Schagger gel analysis can be found in Figure 3.6 for SCOC (78–141) and Figure 3.7 for SCOC (78–151). Pure proteins were concentrated to ~15 mg/mL for crystallography.

3.1.2.4 Expression and Purification of GST-tagged SCOC constructs

GST-tagged full length SCOC isoform 1 and GST-tagged SCOC (78–159) were cloned into pGEX-4T1. Both fusion proteins were expressed in LB at 37 °C. After lysis of the cells, proteins were purified via Protino® GST/4B column chromatography. GST-Tag was cleaved by Thrombin in an overnight incubation and removed by a second step of Protino® GST/4B chromatography. Protein aggregates were removed in a polishing step via size exclusion chromatography. Elution profiles and Schagger gel analysis can be found in Figure 3.8 for GST-SCOC full length and Figure 3.9 for GST-SCOC (78–159). Full length GST-SCOC yielded less protein amounts with contaminations of 40 and 70 kDa, presumably chaperones.

3.1.3 Biochemical and biophysical characterization of SCOC coiled coil domain

SCOC (78–159) was characterized by biochemical and biophysical methods.

3.1.3.1 Limited proteolysis of SCOC coiled coil domain

Well structured, compact domains lacking flexible regions and/or residues are more likely to crystallize. **Limited proteolysis** was performed in search of stable fragments of SCOC ccd.

The protein was digested by twelve different proteases from the Proti-Ace Kits 1 & 2 (Hampton Research). Progress of proteolysis was assessed by taking samples after different time points. For comparison, also undigested protein sample (time = 0 min) was loaded on gels. Analysis by Schagger gels showed, that SCOC (78–159) is stable (see Figure 3.10). Most of the proteases left the

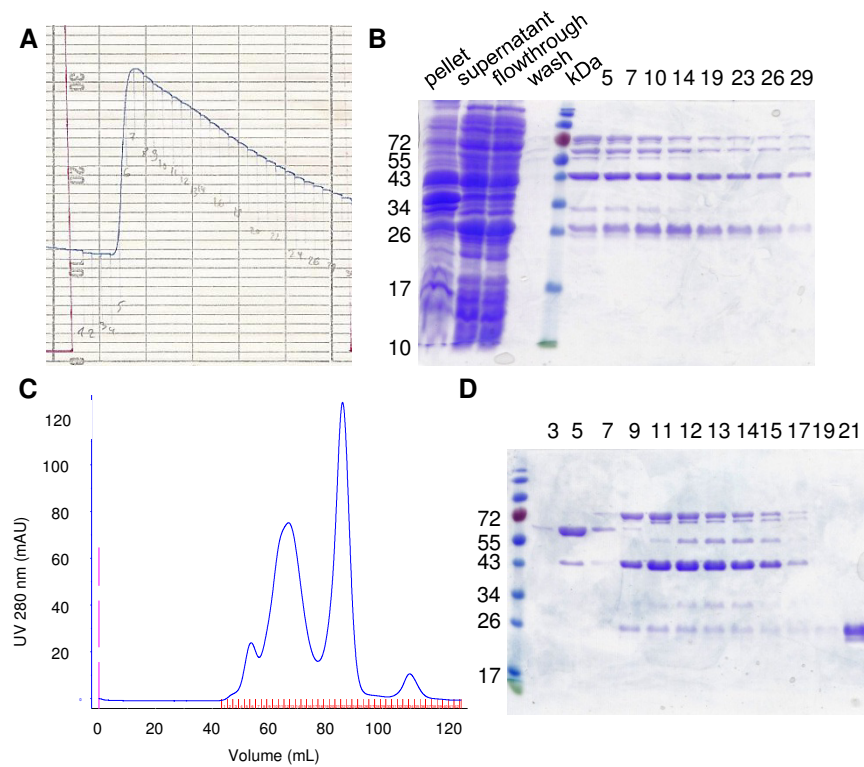


Figure 3.8: Purification of GST-SCOC full length

(A) Elution profile of affinity chromatography via GST column (B) SDS-PAGE of affinity chromatography (C) elution profile of size exclusion chromatography via HiLoad 16/60 S200 column (D) SDS-PAGE of size exclusion chromatography

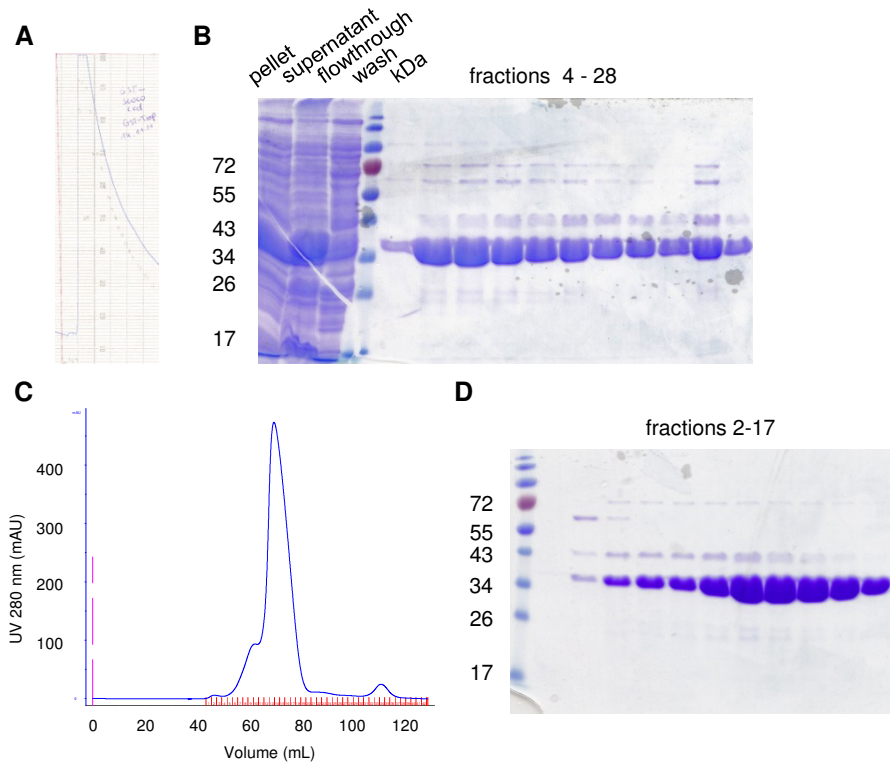


Figure 3.9: Purification of GST-SCOC (78–159)

(A) Elution profile of affinity chromatography via GST column (B) SDS-PAGE of affinity chromatography (C) elution profile of size exclusion chromatography via HiLoad 16/60 S200 column (D) SDS-PAGE of size exclusion chromatography

protein almost intact, showing that the chosen fragment of the ccd is fairly structured with a compact fold. Three proteases, Trypsin, Proteinase K and Subtilisin, produced different proteolytic fragments. Trypsin cleaved only a few residues, yielding a fragment that only a little smaller than the full-length construct. Proteinase-K and Subtilisin trimmed the protein more aggressively, creating at least two smaller fragments at ~ 5 and ~ 10 kDa.

Proteolysis by Subtilisin and Trypsin was analyzed more thoroughly to obtain the fragments which would be sent for N-terminal sequencing. The experiment was repeated at a larger scale with both of the proteases. Samples were taken at time points after 5, 10, 20, 30 and 45 minutes. The proteolysis fragments were reproduced (see Figure 3.11). The fragment at ~ 10 kDa (band 1 in Figure 3.11) appeared with both proteases. Subtilisin also yielded a slightly

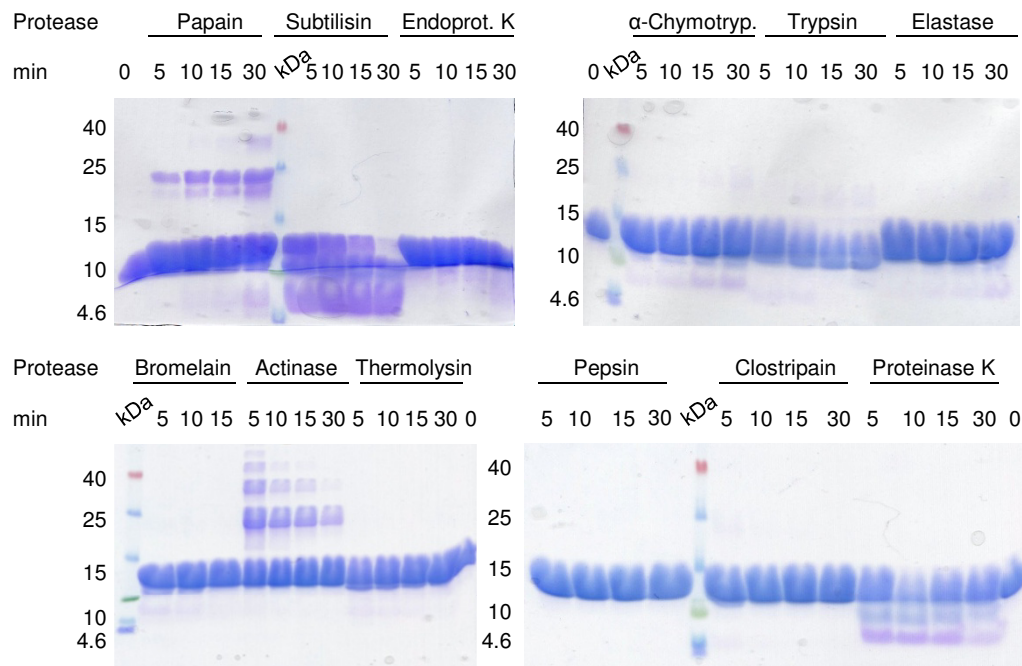


Figure 3.10: Limited proteolysis of SCOC (78–159)

SCOC (78–159) was digested by twelve proteases at RT. Samples at different time points were analyzed on Schagger gels.

smaller size fragment (band 2 in Figure 3.11) and a ~ 5 kDa fragment (band 3 in Figure 3.11). The three bands were blotted on a PVDF membrane and sent for N-terminal sequencing to Seqlab.

The four N-terminal amino acids were sequenced. Results are shown in Table 3.1. Sequence matches of fragments and target sequence are indicated with bold letters.

Comparison with the sequence of the full-length construct revealed that fragments 1 and 2 remained with intact N-terminus. This means, that Subtilisin created two C-terminally truncated fragments slightly smaller than the full-length protein. As the digest was performed with C-terminally Strep-tagged protein, I assumed that Subtilisin cleaved the StrepTag and a few additional residues from the C-terminus of SCOC, since the calculated size of SCOC (78–159) without StrepTag is 9.4 kDa. This result is in agreement with the prediction made by IUPred (see Figure 3.2), that the last few residues of SCOC are disordered. However, the exact C-terminal end of the fragment cannot be concluded from the height of the band in the gel.

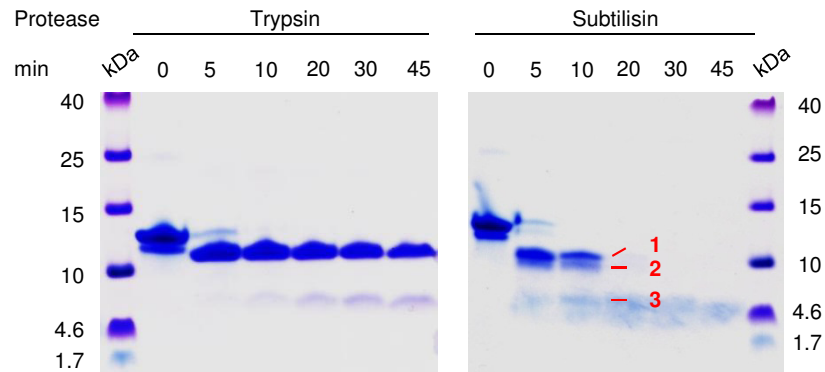


Figure 3.11: Limited proteolysis of SCOC (78–159) by Trypsin and Subtilisin

Bands indicated with red numbers were sent for N-terminal sequencing.

Table 3.1: N-terminal sequencing results

fragment	approx. molecular weight (kDa)	sequence
1	10	AMMN
2	<10	MMN
3	5	(D/E)TKS
full-length (<i>Strep Tag</i>)	11	<u>MM</u> NADMDAVDAENQVELEEKTRLINQVLELQHTLEDLSARV DAVKEENLKLKSENQVLGQYIENLMSASSVFQTTDT <u>TKSK</u> RK LVPRGSWSHPQFEK

Fragment number 3 was identified as the C-terminal part of the fragment, namely the StrepTag with linker plus SCOC's seven C-terminal residues — although the calculated molecular weight of only 2.5 kDa does not correspond to the migrating length of the fragment in the gel. This construct was not of interest for crystallization. Yet I concluded that the C-Terminus of SCOC is flexible, and assumed that residues from amino acid D153 onwards are likely to be disordered.

3.1.3.2 Pull-down experiments for the identification of novel SCOC binding partners

A **GST-Pulldown** experiments were conducted to identify novel interaction partners of SCOC ccd. GST-SCOC (78–159) was immobilized on beads and incubated lysate from HEK cells or lysate from starved, autophagy activated HEK cells. Also, GST protein alone was incubated with HEK cells as control experiment. Samples were analyzed by SDS-PAGE (see Figure 3.12).

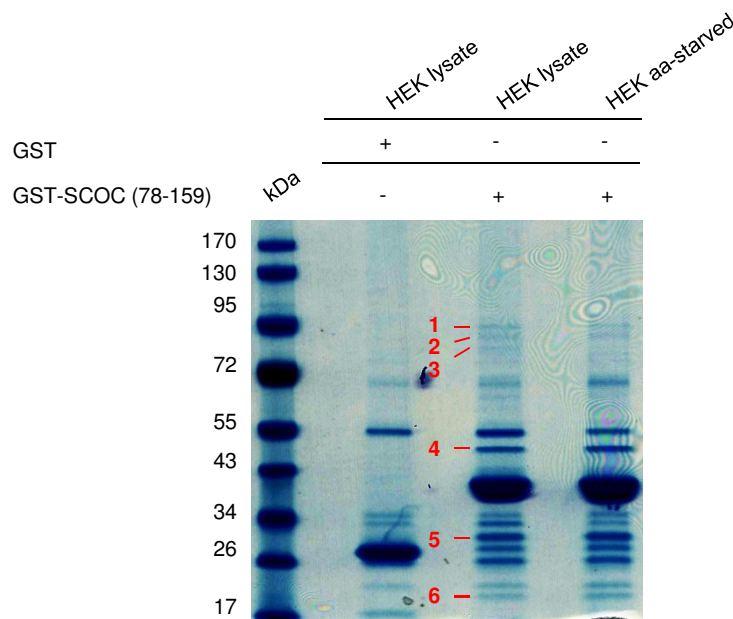


Figure 3.12: GST-Pulldown of GST-SCOC (78–159)

GST-SCOC was incubated with lysate from HEK cells/starved HEK cells. Bands indicated by red numbers were given to mass spectrometry for analysis.

Several proteins were pulled by GST-SCOC (78–159), but there was no difference observable between pull-down with HEK or starved HEK cells. False positives were excluded by comparison with the bands of a pulled-down with GST alone, leaving six bands for further analysis. The indicated bands were digested with Trypsin and the peptides were analyzed by mass spectrometry. Identified peptides were screened with a human protein library. Mostly peptides of GST, Keratin and some chaperones were detected (see Section A.5 for mass spectrometry results). Only band number five contained relevant peptides: five peptides unique for SCOC were obtained with significant protein score (Table 3.2).

Table 3.2: SCOC peptides detected by mass spectrometry in band 5

Peptide	Sequence
1	VDAVKEENLK
2	ASGWDGMGFFSR
3	LINQVLELQHTLEDLSAR
4	SENQVLGQYIENLMSASSVFQTTDTK
5	LKSENQVLGQYIENLMSASSVFQTTDTK
full-length isoform 1 (<i>Strep Tag</i>)	MRRRVFSSQDWR ASGWDGMGFFSR RRTFCGRSGRSCRGQLVQVSR PEVSAGSLLLPAQAEHDSSRILYPRPKSLLPKMMNADMDAVIDA ENQVELEEKTRLINQVLELQHTLEDLSARVDAVKEENLKLKSE NQVLGQYIENLMSASSVFQTTDTKSKRK

Peptide 2 was not only unique for SCOC, but specific for isoform 1. This result indicated, that SCOC is interacting with itself via the ccd in HEK cells.

3.1.3.3 Analysis of secondary structure and stability of SCOC coiled coil domain by CD spectroscopy

Secondary structure of SCOC was investigated by **CD spectroscopy**. CD spectra were measured from 200 to 250 nm at 20 °C. The spectrum showed two minima around 208 and 223 nm, typical for proteins with α -helical structure (Figure 3.13 A). This confirmed the folding of SCOC in α -helices. Stability of SCOC (78–159) was assessed by thermal unfolding from RT to 95 °C measured at 209 nm by CD (Figure 3.13 B). The melting point of 48 °C was defined by fitting the unfolding curve to a sigmoidal function. After heating

the sample to 95 °C, it was cooled to RT. Another spectrum was collected of the same sample, showing that SCOC refolds after thermal denaturation.

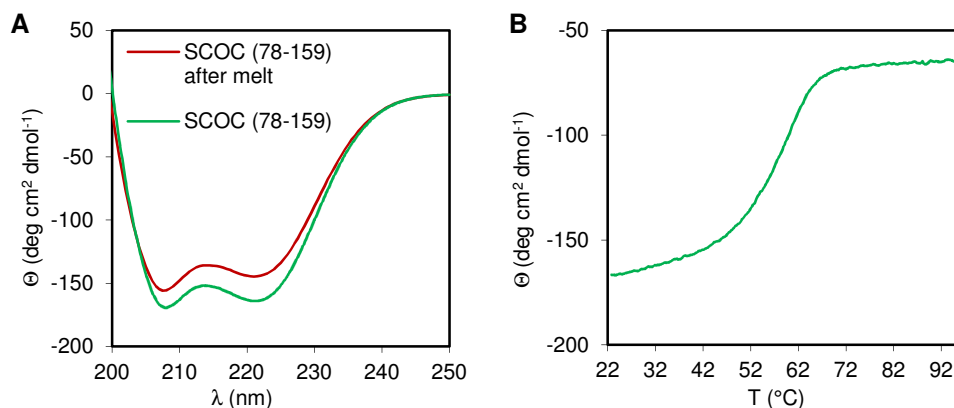


Figure 3.13: CD measurements of SCOC (78–159)

(A) CD spectra (B) Melting curve

3.1.3.4 Characterization of oligomerization state of SCOC coiled coil domain by SEC-MALLS measurements

Oligomerization state of SCOC was investigated by SEC-MALLS. MALLS, in contrast to analytical gel filtration, yields a molecular weight of a protein independent of shape and can therefore be used for the determination of molecular weights of non-globular proteins. The protein samples were run on a Superdex 10/300 GL before analysis with the MALLS detector. Molecular concentrations were determined with a multi-wavelength spectrophotometer. Molecular weights were determined from three runs (see Figure 3.14).

The calculated molecular weight of Strep-tagged SCOC is 10.5 kDa. The measured molecular weight of wild-type SCOC is 26.3 ± 0.3 kDa, which is consistent with a dimer.

3.1.4 Crystallization of SCOC coiled coil domain

The structure determination of SCOC ccd was a major milestone for this study. To achieve this goal, I pursued several strategies for cloning and crystallization of SCOC ccd constructs, which were introduced in Section 3.1.1. Hits from screening of these constructs with 200 nL sitting drops at 20 °C are shown in Figure 3.15, the respective crystallization conditions are listed in Table 3.3.

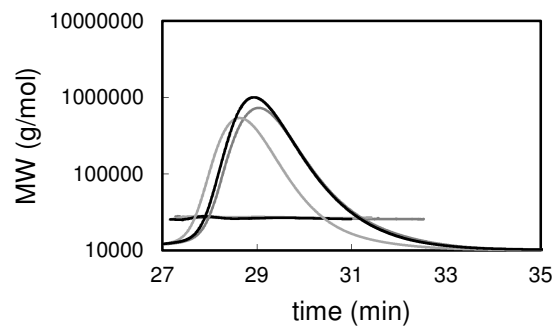


Figure 3.14: SEC-MALLS experiments of SCOC (78–159)

Table 3.3: Exemplary crystallization conditions of SCOC ccd constructs from 96-well plate screening

condition	construct	buffer composition
A	SCOC (78–159)	0.085 M NaH_2PO_4 0.085 M KH_2PO_4 0.085 M MES pH 6.5 1.7 M NaCl 15 % (v/v) glycerol
B	SCOC (78–159) L105M	0.01 M CoCl_2 0.1 M NaAc pH 4.6 1.0 M 1,6-hexanediol
C	SCOC (78–151)	0.1 M BIS-TRIS pH 5.5 17 % w/v PEG 10,000
D	SCOC (78–141)	0.4 M Li_2SO_4 12 % (w/v) PEG 8000 20 % (v/v) glycerol
E	SCOC (78–132)	0.1 M TRIS pH 8.5 25 % (w/v) PEG 8000
F	SCOC (112–159)	0.1 M BICINE pH 9.0 2.0 M NaCl

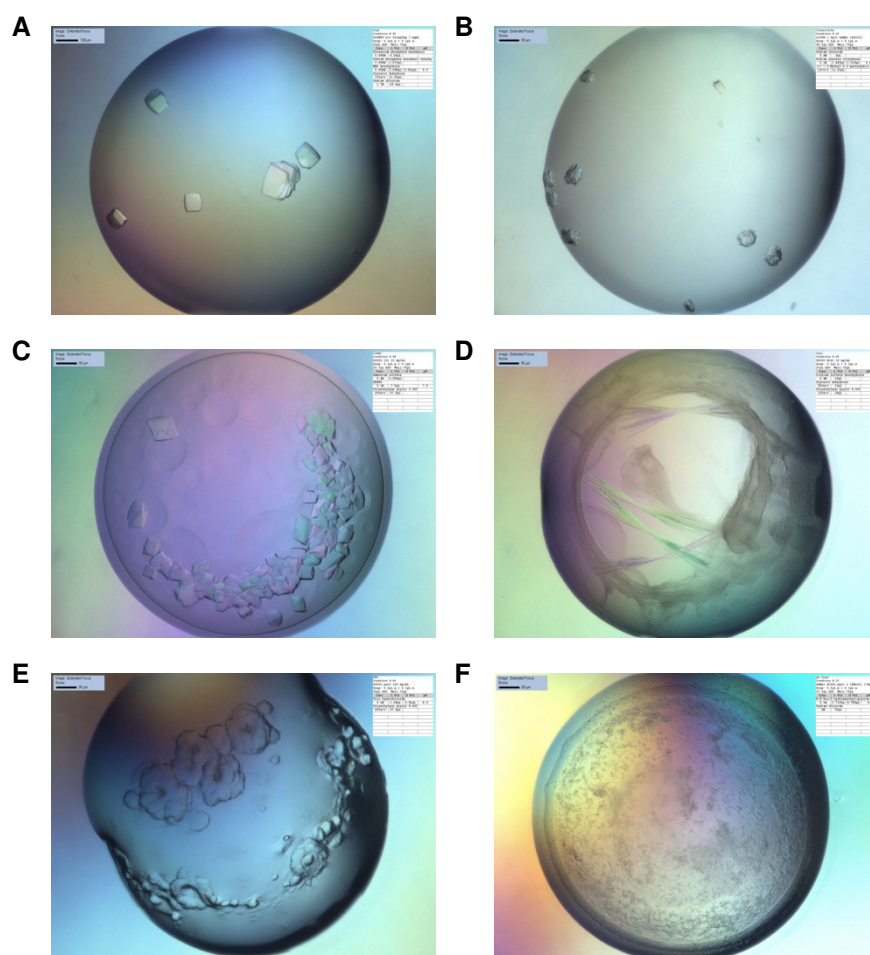


Figure 3.15: Exemplary crystallization hits of SCOC ccd constructs from 96-well plate screening

(A) SCOC (78-159) (B) SCOC L105M (78-159) (C) SCOC (78-151) (D) SCOC (78-141) (E) SCOC (78-132) (F) SCOC (112-159). Crystallization conditions are listed in Table 3.3.

3.1.4.1 Crystallization screening of SCOC (78–159)

The first construct subjected to several crystallization screens was **SCOC ccd (78–159)**. It yielded crystals in more than 100 initial conditions (see Figure 3.15 and Table 3.3). A table of the crystallization conditions from 96-well screens can be found in Section A.2. Thorough analysis of the table revealed that the construct showed a tendency to crystallize with low-molecular weight PEGs or sodium salts as precipitants. The pH for most initial hits was limited to acidic to neutral conditions—a lot of hits contained NaAc (Sodium acetate) around pH 4.6 or MES and HEPES buffer between pH 6.0–7.5. SCOC crystals exhibited different shapes—among them cubic (Figure 3.15), star-like (Figure 3.16) and also bipyramidal shapes (Figure 3.17). Many crystal shapes featured snowflake or flower-like forms, hinting towards multiple crystals growing from one nucleation site. SCOC (78–159) crystallized best from low protein concentrations, mostly between 1–7 mg/mL. Above 10 mg/mL, the protein rather precipitated than crystallized.

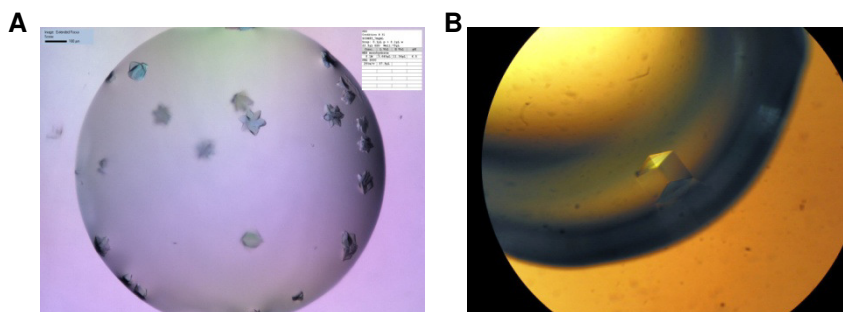


Figure 3.16: Effect of streak seeding

(A) Star-shaped, multiple crystals in a sitting drop in a 96-well plate screen. (B) Single, cubic crystal in a sitting drop set up with streak seeding in a 24-well plate grid screen. Both crystals were grown in 0.1 M MES pH 6.5, 20 mM MgCl₂ and PEG 3350 as precipitant (A: 19% (w/v), B: 25% (w/v) PEG 3350).

3.1.4.2 Refinement of the crystallization conditions of SCOC (78–159) crystals

First crystallization conditions were optimized with **grid screens in 24-well format**. Usually, pH was varied along rows in increments of one pH unit for first optimization. Precipitant concentration was varied along columns.

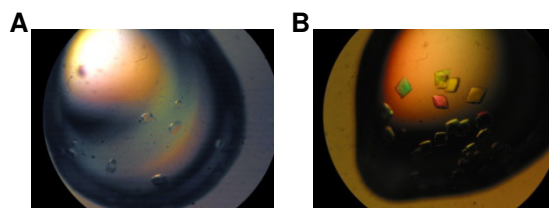


Figure 3.17: Crystals in hanging drops grown in 24-well format

(A) Crystals from a grid screen varying pH (MES 5.5 – MES 6.5) against NaAc concentration (B) Crystals from a grid screen varying pH (MES 5.5 – HEPES 7.5) against PEG 3350 concentration

Figure 3.17) shows crystals obtained by two grid screens (see Table 3.4 for crystallization conditions). In general, 24 well plates yielded larger crystals suitable for flash cooling and data collection. The range of protein concentration tested remained the same for 24-well plates. In order to obtain single, slow growing crystals instead of multiple, fast growing crystals, I also tested crystallization at 4 °C. However, this did not yield any crystals. I further tried to control the vapor diffusion rate to obtain slower growing crystals. In 24-well sitting drop plates, 5 μL (2.5 μL protein plus 2.5 μL crystallization buffer) drops were covered with silicon oil. Only phase separation but no crystals were observed.

Table 3.4: Crystallization conditions in 24 well plate screens

condition	construct	buffer composition
A	SCOC (78–159)	0.1 M MES pH 6.5 1.4 M NaAc
B	SCOC (78–159)	0.1 M MES pH 6.5 19 % (w/v) PEG 3350 0.01 M MgCl_2

When refining the crystallization conditions in 24 well plates, I applied **streak seeding**, providing nucleation sites with the small crushed crystal fragments added by the streak in order to avoid multiple crystals. Figure 3.16 demonstrates the effect of this method: Panel A shows star-shaped crystals grown in a 96-well plate. Panel B shows a crystal with almost the same buffer condition, grown from a streak seeded drop in a 24-well plate.

Crystals from 24-well and 96-well plates were soaked stepwise in a variety of cryoprotectants, and flash cooled. Diffraction quality was tested at 100 K at the synchrotron (Beamline X10SA, Paul Scherrer Institute, Swiss Light Source Villigen, Switzerland) and eventually data were collected. Crystals from PEG conditions diffracted only up to 6 Å. Other samples from salt conditions diffracted to better resolution (4 Å), but the lattice was of insufficient quality with multiple, smeary spots and high mosaicity. The best crystal from initial screening was taken out of a drop from the Cryo screen condition D6 (see Figure 3.15). It diffracted up to 3.6 Å, but with split and smeary spots.

Next, **Random and Grid Screens** were designed and dispensed with the Tecan robot, taking the Cryo screen condition D6 yielding the best crystals so far as basis. Several Grid Screens were set up, varying NaCl concentration in small increments against the pH of MES buffer between 5.2–6.9 or HEPES buffer between 6.8–8.0. As the Cryo condition contained several components, including two different phosphate salts and glycerol, a Random screen was designed to evaluate their effects on crystallization. Crystals obtained from these screens are shown in Figure 3.18, the respective buffer compositions are shown in Table 3.5.

Table 3.5: Exemplary crystallization conditions of Random and Grid screens

condition	construct	buffer composition
A	SeMet SCOC (78–159)	0.085 M MES pH 6.77 0.085 M NaH ₂ PO ₄ 0.085 M KH ₂ PO ₄ 1.43 M NaCl 15 % (v/v) glycerol
B	SCOC (78–159) w/o Tag	0.085 M MES pH 6.45 0.0122 M NaH ₂ PO ₄ 0.147 M KH ₂ PO ₄ 1.77 M NaCl
C	SCOC (78–151)	0.1 M MES pH 5.55 2.79 M NaCl
D	SCOC (78–141)	0.1 M HEPES pH 7.31 2.68 M NaCl

The designed screens yielded many crystals, most of them multiple or single

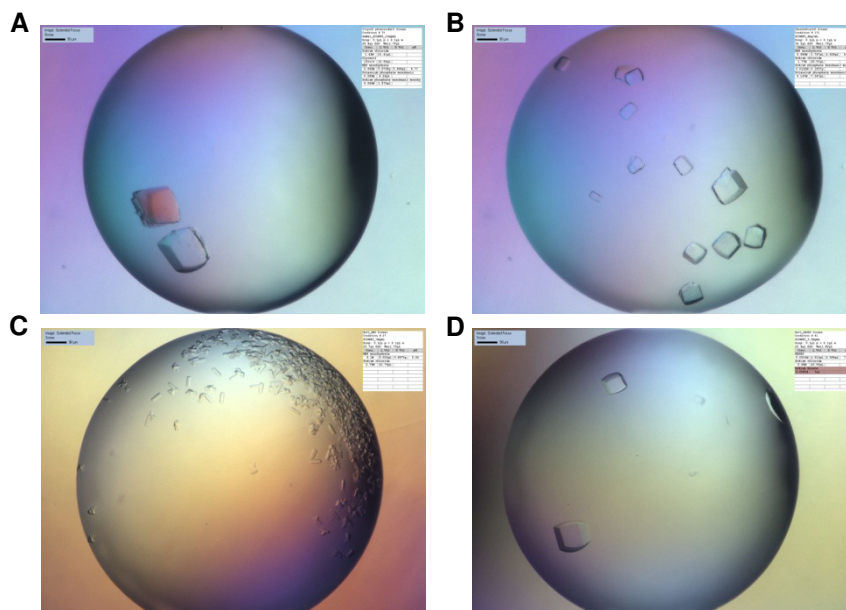


Figure 3.18: Crystals obtained from refinement with grid and random screens

(A) selenomethionine-labelled SCOC (78–159) crystals from a random Screen (B) SCOC (78–159) crystals from a random screen, crystallized without StrepTag (C) SCOC (78–159) crystals from a grid screen (D) SCOC (78–159) crystals from a grid screen

crystals in elongated cuboid shapes. In addition, the screens were utilized to test the crystallization behaviour of SCOC (78–159) variants without the StrepTag (cleaved with Thrombin with a protocol according to Section 2.2.2.5) (see Figure 3.18 B) or selenomethionine-labelled SCOC (see Figure 3.18 A) (see Section A.2 for a table summarizing crystallization conditions from random and grid screens). Large and single crystals were soaked in cryoprotectant (only if necessary, as many conditions already contained glycerol as cryoprotectant), and their diffraction was tested. However, neither diffraction resolution nor lattice quality was significantly improved.

Hence, although SCOC (78–159) crystallized easily from a large variety of conditions, the diffraction quality even after extensive screening was not sufficient to solve the structure.

3.1.4.3 Screening of other SCOC ccd constructs

In parallel, I cloned and purified **C-terminally truncated constructs SCOC (78–141) and SCOC (78–151)**. Crystallization screens were set up at 20 °C with 200 nL sitting drops. SCOC (78–151) at 15 and 7.5 mg/mL protein concentration yielded mostly crystal clusters (Figure 3.15 C). SCOC (78–141) at 12 and 6 mg/mL was easily crystallized, forming needles or needle clusters in a few conditions (Figure 3.15 D) (see Table 3.3 for crystallization conditions). The crystal needles were very fragile and required careful handling during soaking and flash cooling. The ccd fragment crystals diffracted to only 5–6 Å resolution.

Another attempt to achieve better diffracting crystals was made by cloning and purifying two **overlapping halves of SCOC ccd, SCOC (78–132) and SCOC (112–159)**. 96-well plates sitting drops were dispensed at 20 °C at 40 and 20 mg/mL (SCOC 78–132) and 17 and 8.5 mg/mL (SCOC 112–159). Both constructs yielded only spherulites, gel or cluster-like structures in 96-well screening (Figure 3.15 E and F).

Summing up, the purpose-built constructs—some of them missing presumably unstructured parts of the ccd, others designed to pack better in the crystal—did not result in better diffracting crystals compared to SCOC (78–159).

3.1.4.4 Crystallization of selenomethionine-labelled SCOC crystals

Selenomethionine labelled crystals (SeMet SCOC (78–159) in Figure 3.18 A) grew predominantly in conditions similar to the native protein. The crystals showed similar diffraction quality as native crystals, however, the anomalous signal dropped rapidly between 5 and 6 Å, making structure determination from single or multi-wavelength anomalous diffraction (SAD or MAD) very challenging. Hence, two mutant constructs, SCOC L96M (78–159) and SCOC L105M (78–159), were cloned, purified and subjected to crystallization screens. Initial screening of SCOC L96M (78–159) did not yield promising conditions. SCOC L105M (78–159) showed similar crystallization behavior as wt selenomethionine crystals and was subjected to *in situ* proteolysis (see below).

3.1.4.5 Crystals grown with *in situ* proteolysis

After conducting limited proteolysis experiments (see Section 3.1.3) of SCOC (78–159), the overall stability but also the presence of flexible parts was demonstrated. Trypsin and Subtilisin were two proteases which were able to cleave

SCOC's ccd. It proved very challenging to isolate fragments of the ccd obtained by limited proteolysis. Hence, I aimed to mimic limited proteolysis in the crystal drop by conducting *in situ* proteolysis. Subtilisin and Trypsin were added to the protein stock in small amounts (1:1000 w/w Trypsin and 1:2000 w/w Subtilisin) just before crystallization. Screening of SCOC (78–159) was performed as described before. Again, 96-well screening yielded a lot of crystals, some of them shown in Figure 3.19, see Table 3.6 for buffer composition (see Section A.2 for a table of crystallization conditions obtained by *in situ* proteolysis). Obtained crystals differed in crystal shape and quality of diffraction. In high NH_4NO_3 concentration (Figure 3.19 C), round-shaped crystal clusters as well as cuboid crystals grew. Also, multiple crystals grew from the condition that had yielded crystals diffracting up to 3.6 Å before (Figure 3.19 B). Comparison with the conditions found by conventional screening indicated a few novel crystallization conditions, one of them yielding crystal rods shown in Figure 3.19 A. Native crystals of this condition were flash cooled with a broad range of cryoprotectants and tested for diffraction. The use of 15–30 % PEG 400 as cryoprotectant improved diffraction quality tremendously, showing distinct, well-defined diffraction spots up to a resolution of 3.0 Å. Crystals of similar rod shape were obtained by 24-well grid screening of selenomethionine-labelled SCOC L105M (78–159) at 3 and 1.5 mg/mL, varying 1,6-hexanediol concentration against pH of NaAc buffer (see Figure 3.19 D). Flash cooling with 20 % PEG 400 resulted in high-quality diffraction patterns of SCOC L105M (78–159) crystals grown in 0.01 $\text{CoCl}_2 \times 6 \text{H}_2\text{O}$, 0.1 M NaAc pH 4.63, 0.7 M 1,6 hexanediol with streak seeding. The crystals yielded datasets 2.8 Å resolution, which was sufficient for structure determination.

3.1.5 Structure determination

X-ray diffraction data were collected at 100 K at beamline X10SA (Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland). Data for structure determination were collected from a selenomethionine SCOC L105M (78–159) crystal grown in 0.01 CoCl_2 , 0.1 M NaAc pH 4.63, 0.7 M 1,6 hexanediol, cryosoaked in mother liquor supplemented with 20 % PEG 400. First, a crystal of the same purification batch was used for a measurement of a fluorescence spectrum and a scan to define the exact Se K absorption edge. Inflection, Peak, and remote 1 and remote 2 energies were determined with AUTOCHOOCH [107]. After mounting of the crystal for data collection, four test shots were taken for indexing of the crystal. Unit cell and collection strategies were determined with go.com (local software written by Dr. M. Wang).

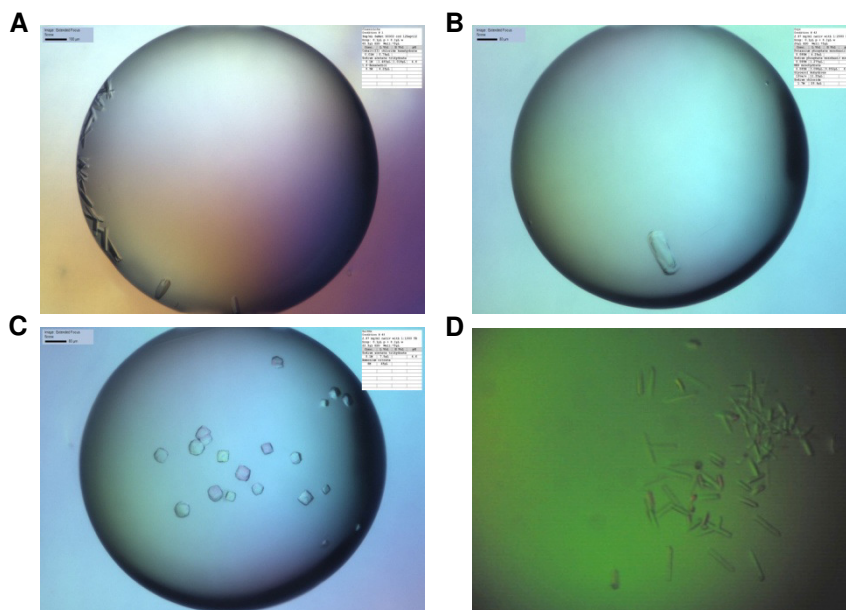


Figure 3.19: Crystals grown by *in situ* proteolysis

(A) Crystal rods in a drop from ClassicLite screen (Qiagen) (B) Multiple crystal from in a drop from Cryo screen (Qiagen) (C) Spherulites and cuboid crystals in a drop from SaltRX screen (Hampton Research) (D) Crystal rods in a streak-seeded drop from 24-well grid screening

A SAD dataset of 180° was collected at Se peak wavelength (details of data collection are noted in Table 3.7.)

Data were processed with input values specific for the Pilatus detector and FRIEDEL'S LAW = FALSE with XDS [95] (see Section A.4.3 for XDS.INP file). The orthorhombic space group 20 ($C222_1$) was confirmed by the expected absences of reflections along $(0,0,2n)$ (see Section A.4.4 for a list of systematic absences as expected by CORRECT.LP). Dimensions of the unit cell, data and refinement statistics are listed in Table 3.8. With XDSCONV, scaled data was converted to unmerged CCP4F format.

Phenix was used for phasing, density modification, initial model building and refinement [96]. For structure determination, the unmerged diffraction data in CCP4F format was supplied as .mtz-file, together with the sequence as .fasta-file into Phenix. Moreover, wavelength was set to 0.979 \AA , and $f'(-6.9)$ and $f''(4.802)$ were put in as determined by the fluorescence scan. Model auto-building was activated with thoroughness set to "thorough". Xtrriage package

Table 3.6: Exemplary crystallization conditions of *in situ* proteolysis

condition	construct	composition
A	SeMet SCOC (78–159) 1:2000 SU	0.01 CoCL ₂ x 6 H ₂ O 0.1 M NaAc pH 4.6 0.5 M 1,6 hexanediol
B	SCOC (78–159) 1:2000 SU	0.085 M MES pH 6.45 0.085 M NaH ₂ PO ₄ 0.085 M KH ₂ PO ₄ 1.7 M NaCl 15 %(v/v) Glycerol
C	SCOC (78–151) 1:1000 TR	0.1 M NaAC pH 4.6 6 M NH ₄ NO ₃
D	SCOC L105M (78–141) 1:2000 SU	0.01 CoCL ₂ x 6 H ₂ O 0.1 M NaAc pH 4.63 0.7 M 1,6 hexanediol

Table 3.7: Data collection

Detector distance	520 mm
$\varphi/\Delta\varphi$	0/0.5 °
Exposure time	0.4 s
Beam intensity	0.35
λ	0.979
No of frames	360
Space group	C222 ₁

Table 3.8: Diffraction data and refinement statistics

SCOC L105M (78–159)	
Data collection	
Space group	C222 ₁
Cell dimensions	
a, b, c (Å)	70.7, 114.5, 93.1
α, β, γ	90, 90, 90
Resolution (highest res. shell) (Å)	48.87–2.70 (2.87–2.70)
R _{meas}	6.9 (44.9)
no. of observed/unique reflections	61581/20013 (9535/3184)
I/ σ (I)	13.3 (2.7)
Completeness (%)	99.1 (97.9)
Wilson B factor (Å ²)	27.00
Molecules/AU	3
Refinement	
Resolution (Å)	2.7
R _{work} /R _{free}	23.1/26.9
No of atoms	
Residues included in model (number of protein atoms)	A: 86–146 (477) B: 88–147 (462) C: 86–146 (440)
Water	43
B-factors (Å²)	
Overall	22.7
Protein	23.7
Water	21.6
r.m.s.d.	
Bond lengths	0.08
Bond angles	1.015
PDB ID	

by Phenix determined cutoff of the anomalous signal at 4.4 Å; whereas the high resolution limit for the dataset was 2.7 Å (see Figure 3.20). Phenix found ten Se sites, three of them with occupancies above 0.8 (see Section A.4.5 for a list of Se sites). The obtained experimental electron density map phasing model yielded an overall score of 39.01. The positions of three Se sites localized to correct positions in the electron density maps. The initial model contained three molecules A, B and C in the asymmetric unit, for two of them already large parts of the amino acid sequence were built in. Manual model building was done with Coot [97]. From the initial model, molecule B was rebuilt manually, as well as large parts of molecule C. Rotamer outliers of side chains and geometric constraints were checked and corrected with Coot. 29 Cycles of refinement with Phenix against the initial electron density map resulted in a final structure refined to $R_{\text{work}}/R_{\text{free}}$ of 23.1/26.9 %. Validation of the structure was done with Phenix. A polygon plot showing the refinement statistics of the solved final structure in comparison to structures in the PDB with similar resolution cutoff, is presented in Figure 3.21. All values lie within the permitted limits of other structures submitted to the PDB. The Ramachandran plot for assessing the correct geometry of the peptide backbone of the refined structure is shown in Figure 3.22. The residues are all located in the allowed regions.

3.1.6 Structure of SCOC's coiled coil domain

X-ray structure of SCOC L105M (78–159) revealed that SCOC is a parallel left-handed coiled coil dimer. The structure comprises eight heptad repeats and has a length of about 80 Å. Intriguingly, I observed two distinct dimers in the crystal structure (see Figure 3.23). The asymmetric unit contains three SCOC molecules A, B and C. Molecules A and B form dimer AB. The second dimer CC' contains an internal two-fold symmetry, since the dimer is composed of molecule C and a symmetry related copy of C. In the crystal structure, the C-terminus of the CC'-dimer is embedded between the N-termini of symmetry related copies of AB-dimers.

Dimer CC' is a regular coiled coil with tight intertwined helices. Chains A and B are not as tightly packed because molecule A is bent with a bulge around residues A116 (Figure 3.24 A). Overlay of the two dimers (Figure 3.24 B) showed that molecules B and C adopt a very similar conformation and superimpose with a r.m.s.d. of 1.0 Å for the C_{α} atoms. In contrast, the differences between A and either B or C are more pronounced with r.m.s.d. values of 2.9 Å and 2.3 Å, respectively. The coiled coil pitch of dimer AB (residues 97–143) is 129 Å and 151 Å for dimer CC' as calculated with the program

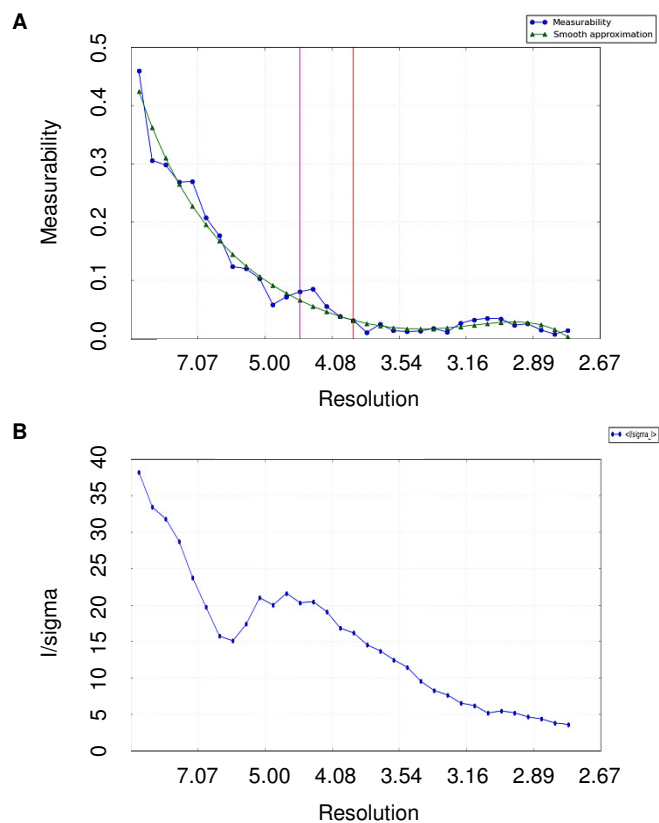


Figure 3.20: Phenix Xtriage analysis

(A) Anomalous signal. Recommended resolution cutoff for heavy-atom search: 4.44 Å (purple line). Optimistic resolution cutoff: 3.993 Å (red line). (B) Signal to noise I/σ for SAD data set.

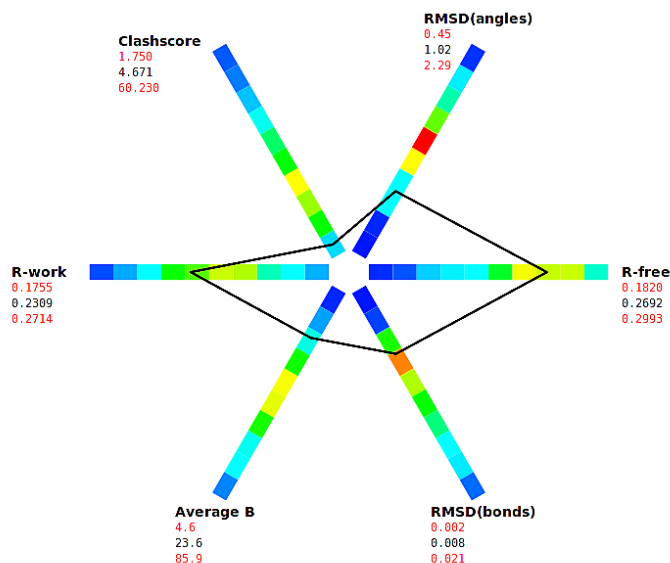


Figure 3.21: Polygon plot by Phenix

Selected statistics across 716 PDB entries of similar resolution are compared to the current structure, with range indicated by red numbers.

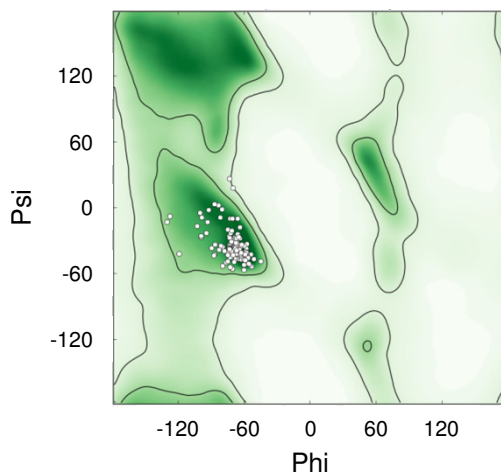


Figure 3.22: Ramachandran plot of all non Pro/Gly residues of SCOC L105M (78-159)

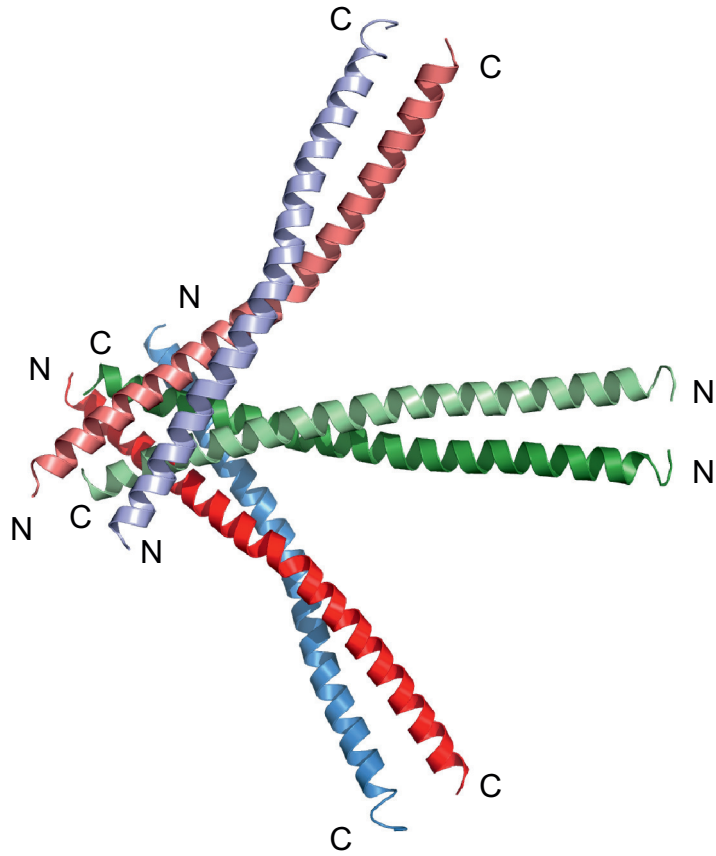


Figure 3.23: Structure of SCOC coiled coil

Two asymmetric units with six molecules A (red), A' (palered), B (blue), B' (paleblue), C (green) and C' (palegreen) are displayed.

TWISTER [108]. Both values are similar to the coiled coil pitch of 135 Å for the regular left-handed parallel coiled coil GCN4 leucine zipper dimer (pdb entry: 2ZTA). The SCOC coiled coil domain (residues 78–146) has a pI of 4.3 and the overall surface charge of the molecule is negative (Figure 3.25). There are only a few conserved positively charged patches present, which includes residues R99 and R117. Conservation of SCOC was analyzed across various species with T-Coffee [109] and displayed on the molecule’s surface (Figure 3.26, see Section A.4.6 for alignment). In general, the surface and core of SCOC are highly conserved, but residues at both ends and near the bulge of molecule A are more variable across species (see Figure 3.26 and Figure 3.27).

Coiled coils are characterized by the heptad repeat *'abcdefg'* featuring hydrophobic core residues on positions a and d (see chapter 3). Heptad assignment of SCOC sequence is shown in the alignment (Figure 3.27) and in the helical wheel representation (Figure 3.28). Remarkably, half of the a heptad positions at the core of the ccd are occupied by asparagine residues and lysine K97. Non-canonical polar pairings at a-positions are found at the N-terminal end of the coiled coil (N90 and K97) and close to the bulge of molecule A (N125 and N132). Additionally, there is one charged d-residue (E93), whereas the other d-positions are occupied by leucines and V121 (Figure 3.28).

The charged and polar residues at the coiled coil interface are stabilized through hydrogen bonds and salt bridges (Figure 3.29). N125 forms a net of hydrogen bonds with E124 and K129 (Fig 1 F). In dimer AB, there is also a water molecule stabilizing N125. Polar interactions of N132 with E131 are very similar in both dimers. E93, however, forms both intra- and intermolecular salt bridges with K97 in dimer AB and, whereas it forms an intramolecular hydrogen bond with N90 in dimer CC'. Importantly, all polar and charged core residues are highly conserved among human isoforms and other species (Figure 3.27 and Figure 3.26). I expected that these residues result in weakened core interactions, creating a destabilized dimer, which would explain the observed conformational flexibility of SCOC. In order to test the influence of these polar core residues on the stability of the protein two double core mutants E93V/K97L and N125L/N132V were created. Additionally, a second set of mutants was prepared to probe SCOC–FEZ1 complex formation. SCOC is known to interact with the coiled coil domain of FEZ1 [37], (see Section 3.2.1, Figure 3.26 for further details). The FEZ1 coiled-coil domain (residues 227–290) is negatively charged with a pI of 4.7. I therefore speculated that SCOC residues R99 and R117, which are conserved and surface exposed might be important for SCOC–FEZ1 complex formation and prepared the R99E and R117E SCOC mutants.

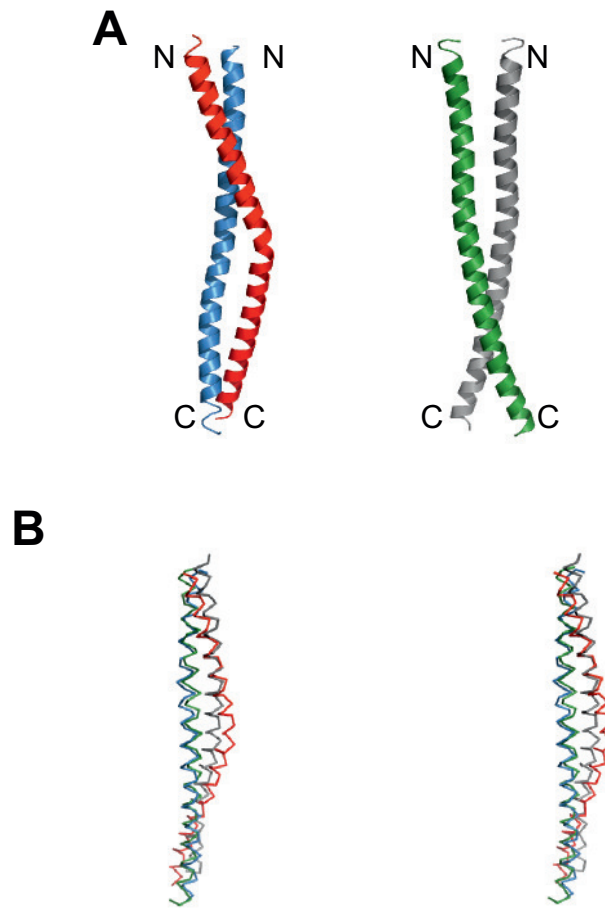


Figure 3.24: Two distinct SCOC dimers in the crystal structure

(A) Cartoon representation of the two SCOC dimers AB and CC'. Molecule A is colored red, B blue, C green and C' grey. (B) Stereo view of dimers overlay. Same coloring as in (A).

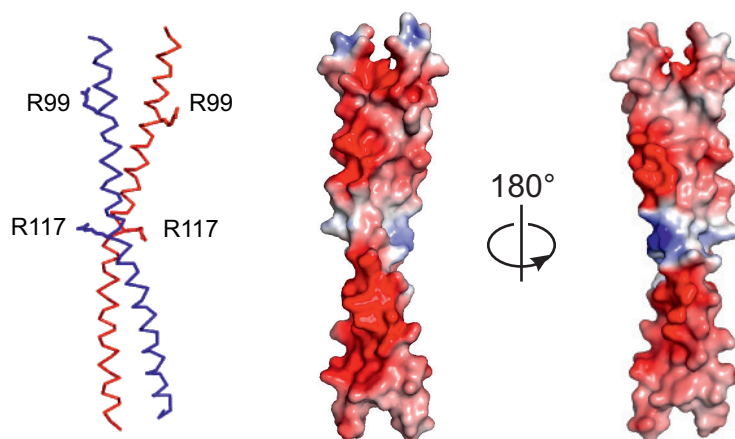


Figure 3.25: Surface charge of SCOC

Left panel shows a ribbon representation of dimer AB. Residues R99 and R117 are drawn as stick models. Electrostatic surface potential is shown for dimer AB in two orientations. Blue shows positive charge and red corresponds to negative charges. The first figure of the electrostatic surface potential is shown in the same orientation as the ribbon representation.

3.1.7 Biophysical characterization of SCOC's mutants

I assumed that the substitution of polar residues with hydrophobic amino acids in the coiled core would result in enhanced stability. CD was used to analyze secondary structure and the stabilities of SCOC mutants. All SCOC mutants exhibit α -helical secondary structure at 20 °C as the wt protein (see Figure 3.30 A). Thermal unfolding curves were measured between 20 °C and 93 °C at a wavelength of 208 nm. Wild-type SCOC (78–159) unfolds at 48 °C (see Section 3.1.3). Both arginine surface mutations had no significant effect on stability of the proteins (see Figure 3.31 and Table 3.9 for a summary of the biophysical characterization of SCOC mutants). In contrast, both double core mutants are much more stable than the wild-type protein. The melting temperature of E93V/K97L is increased by almost 30 °C. Strikingly, the N125L/N132V mutant is extremely stable. The protein remains fully folded at even 93 °C (see Figure 3.30 B).

Non-polar core position of coiled coils influence the oligomerization state [59]. The dramatic differences in stability of both core double mutants indicated a change of oligomerization state. First, I analyzed the size of mutant SCOC oligomers by analytical gel filtration (Figure 3.32). Elution profiles showed

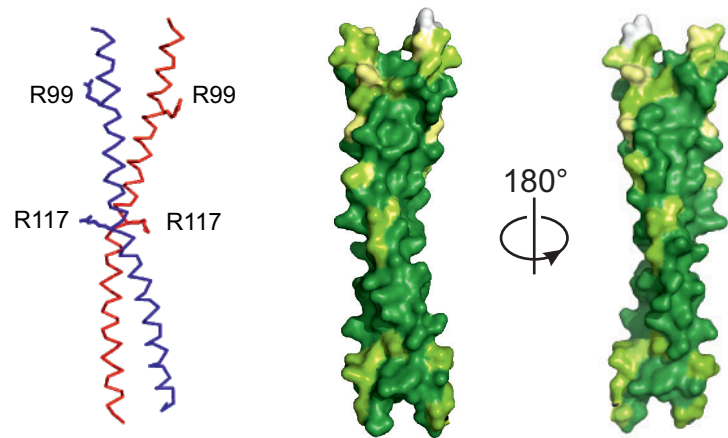


Figure 3.26: SCOC conservation degree on the surface

Left panel shows a ribbon representation of dimer AB. Residues R99 and R117 are drawn as stick models. The two right panels show the surface conservation of SCOC. Dark green correspond to strong conservation and lighter green shades represent less conserved regions. The first conservation figure is shown in the same orientation as the ribbon representation. Sequence alignments were done with T-Coffee and analysis of the degree of conservation was done with AMAS [99].

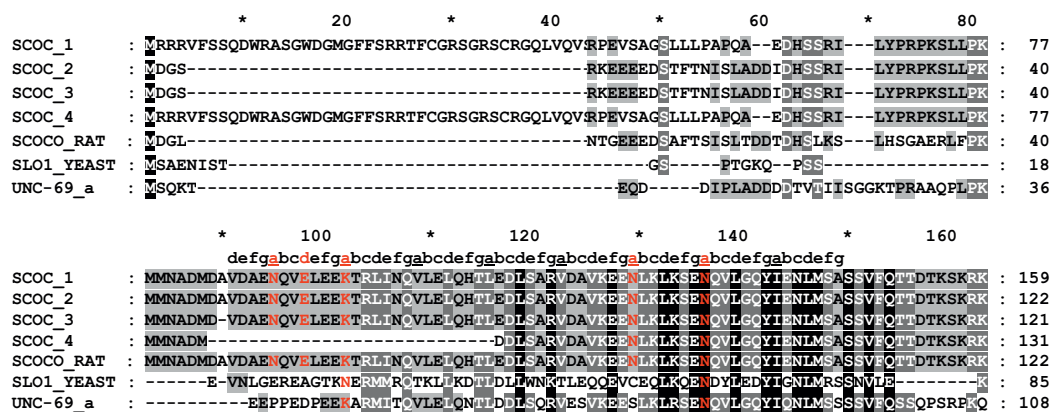


Figure 3.27: Sequence alignment of human SCOC isoforms and homologues

Alignment was done with T-Coffee. Coiled coil heptad positions were assigned with TWISTER. Non-canonical polar amino acids at core positions a and d are colored red.

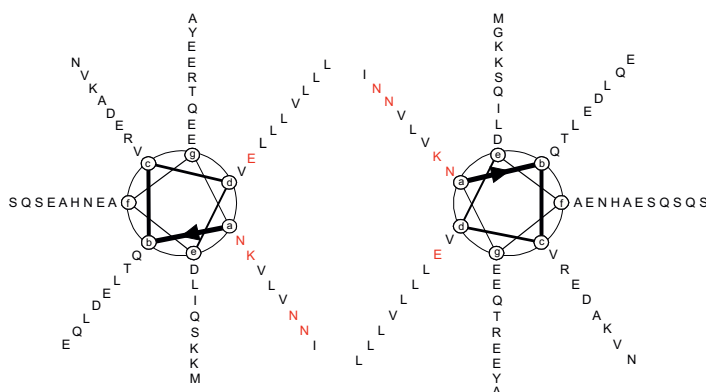


Figure 3.28: Helical wheel diagram of the SCOC dimer

Polar core residues are marked red.

Table 3.9: Summary of biophysical characterization of SCOC mutants

SCOCO	T _m , °C	Mw, kDa (error)	n (error)
wt	48.0	26.29 (0.3)	2.40 (0.1)
E93V/K97L	75.0	34.80 (0.1)	3.18 (0.1)
N125L/N132V	—	42.07 (0.1)	3.85 (0.1)
R99E	42.0	27.00 (0.6)	2.47 (0.1)
R117E	51.0	23.31 (2.1)	2.13 (0.2)

similar elution volumes of ~ 15 mL, except for the N125L/N132V mutant that eluted around 14 mL. This indicated a higher oligomerization state. However, the non-globular shape of elongated coiled coils allows no accurate determination of molecular weight by analytical gel filtration. Therefore, I used SEC-MALLS to measure the molecular weights of the mutant proteins. Molecular weights were averaged from three SEC-MALLS measurements for each mutant (Figure 3.33) and are summarized in Table 3.9. The sequence based molecular weight of Strep-tagged SCOC is 10.5 kDa. The stoichiometry (n) is the ratio of the measured molecular weight and the MW of a single Strep-tagged SCOC molecule (10538 Da). Measured molecular weight of wild-type SCOC is 26.3 ± 0.3 kDa (see Section 3.1.3). The molecular weight determined for N125L/N132V is 42.1 ± 0.1 kDa, which corresponds to a tetramer. The

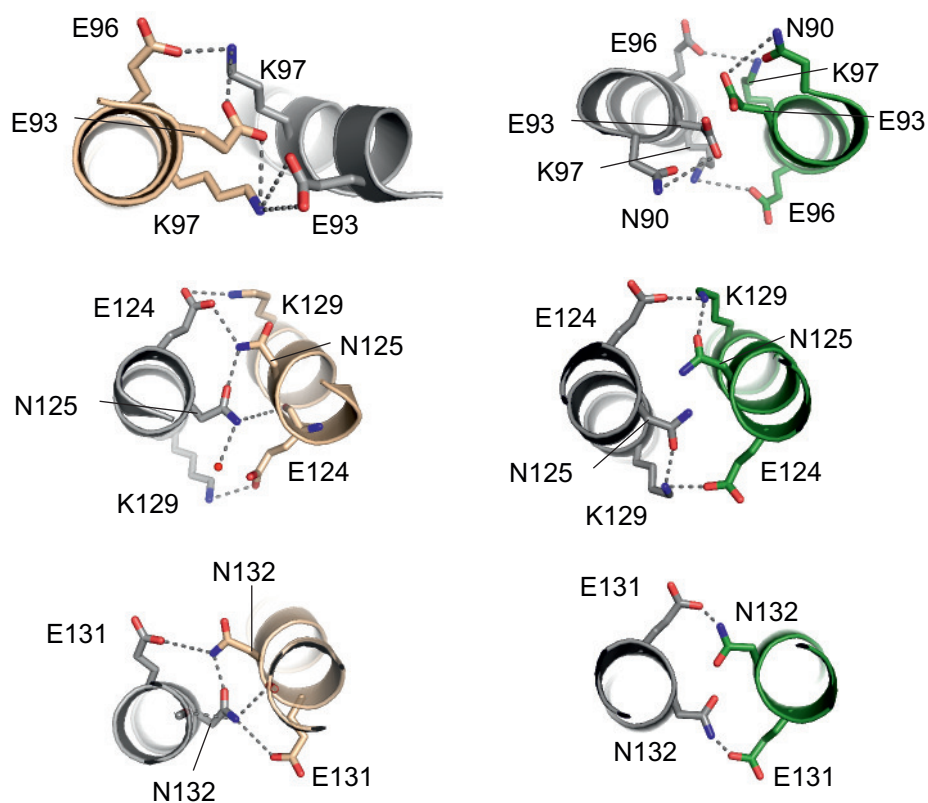


Figure 3.29: Stabilizing interactions of non-canonical core positions

Detailed molecular interactions of the polar core positions of dimer AB (left panel) and dimer CC' (right panel). Core residue E93, K97, N125 and N132, which were used for mutagenesis studies are stabilized through a network hydrogen bonds and salt bridges as shown. Figures were prepared with PyMol.

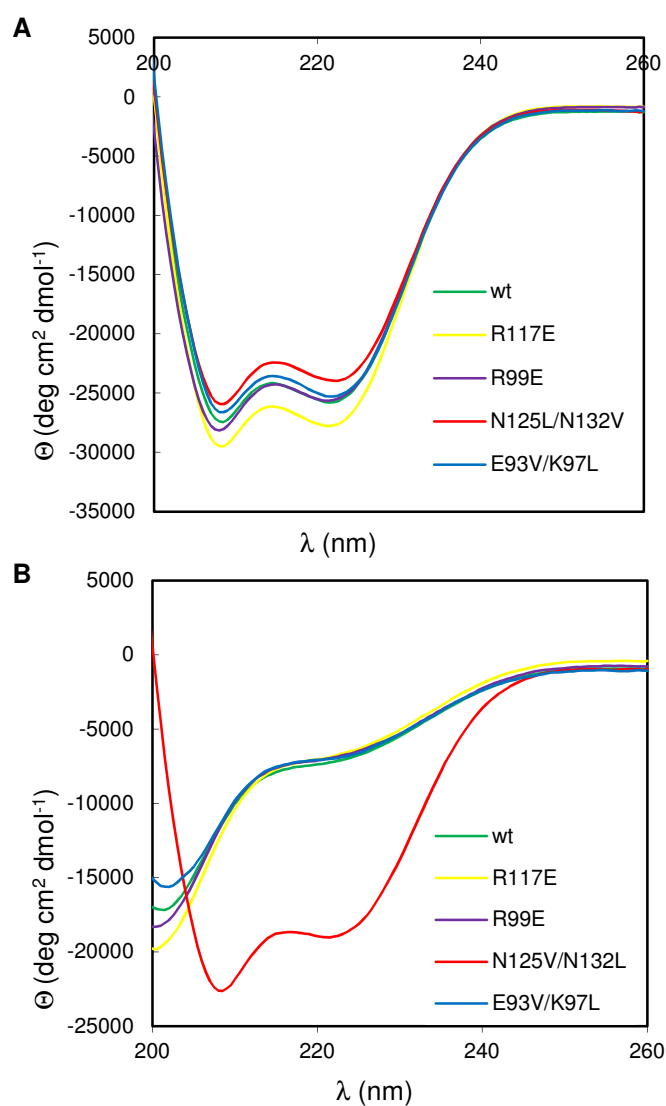


Figure 3.30: CD spectra of SCOC mutants

(A) CD spectra of SCOC mutants at 20 °C from 200 to 260 nm. (B) CD spectra of SCOC mutants at 93 °C from 200 to 260 nm.

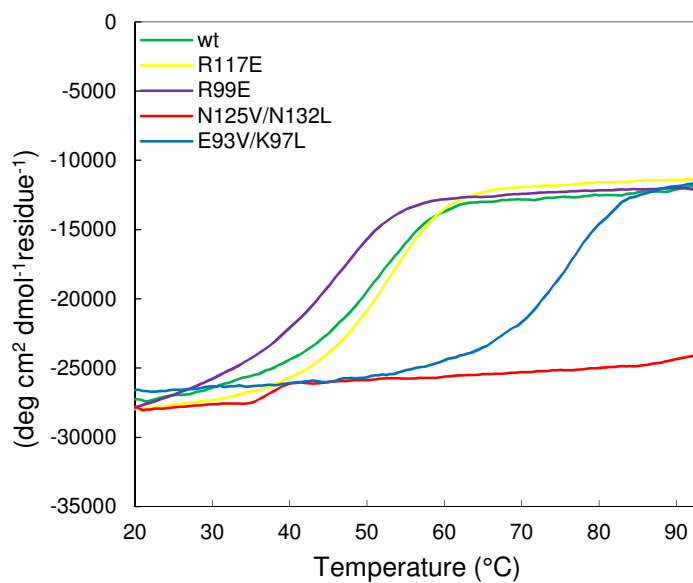


Figure 3.31: Thermal unfolding CD curves

Melting curves were recorded at 208 nm from 20 to 93 °C at a protein concentration of approx. 15 μ M.

E93V/K97L double mutant forms a trimer (34.8 ± 0.1 kDa). Both R99E and R117E surface mutants are dimers as the wt protein.

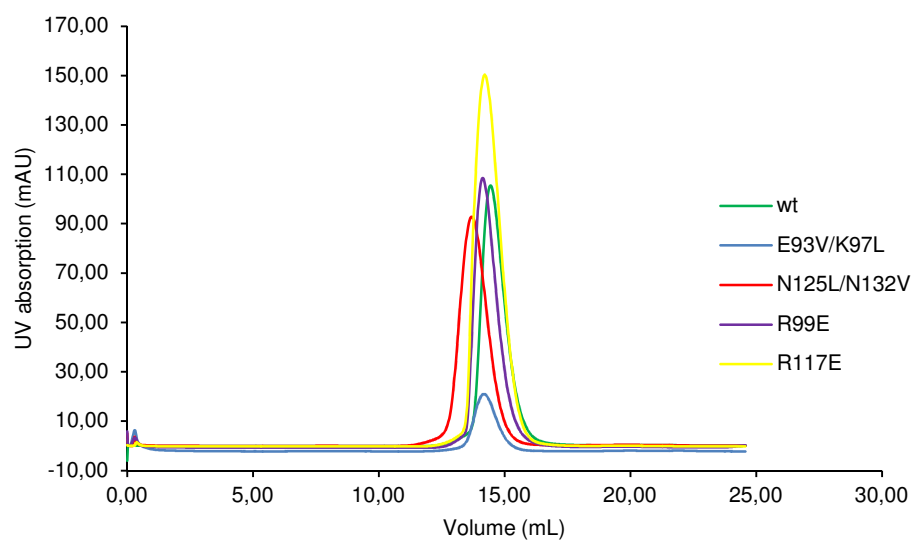


Figure 3.32: Analytical gel filtration elution profiles of SCOC mutants

Proteins at 2 mg/mL were separated via Superdex 10/300 GL chromatography.

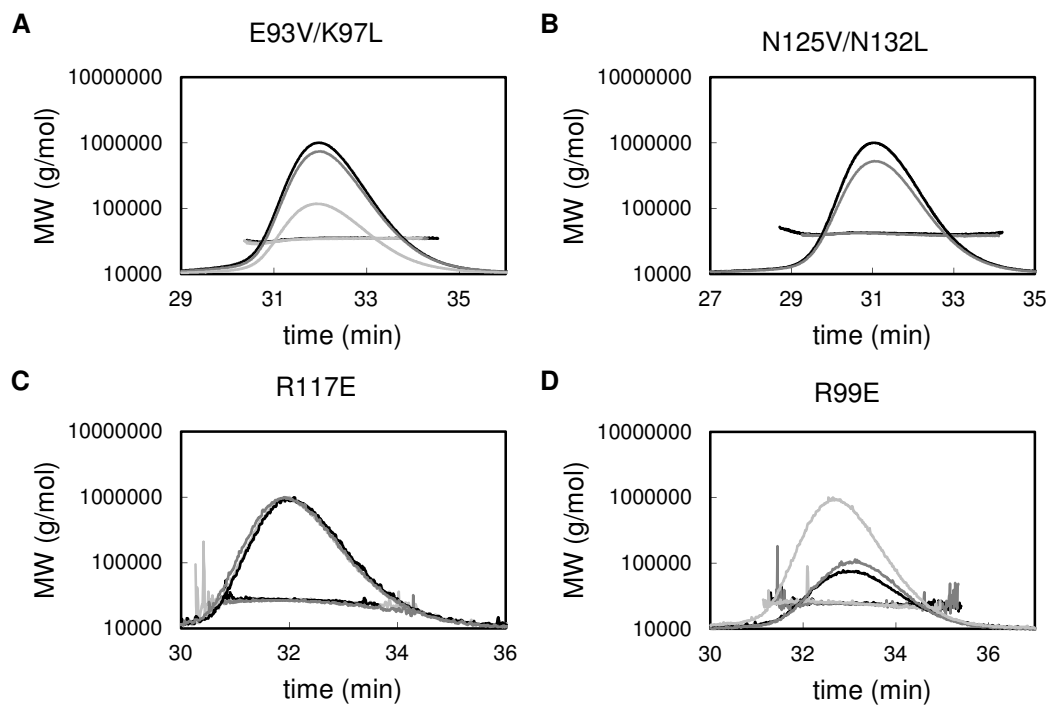


Figure 3.33: Elution profiles and corresponding molecular weights determined by SEC-MALLS are shown for all SCOC mutants

3.2 Interaction of SCOC's coiled coil domain with Arl1 and FEZ1

3.2.1 Insights into SCOC–FEZ1 complex formation

3.2.1.1 Approaches to find a minimal FEZ1 binding domain



Figure 3.34: FEZ1 domain

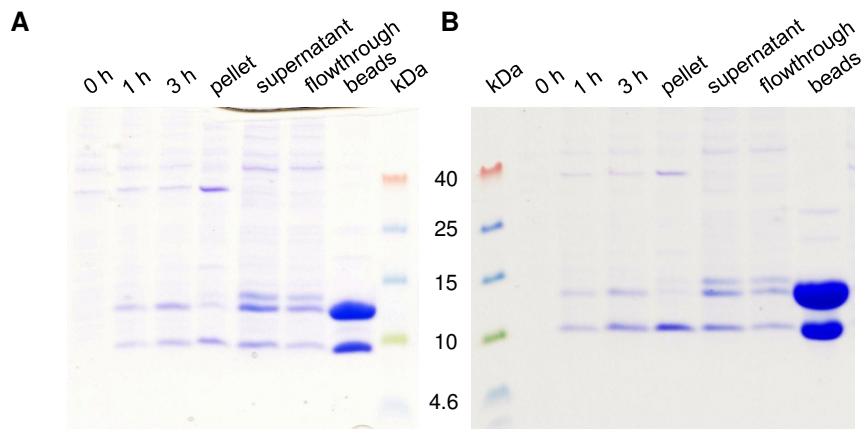
Human FEZ1 (392 residues) has properties of a natively unfolded protein. It contains three glutamate rich regions and a conserved coiled coil domain in the C-terminal half of the protein [45, 110], see Figure 3.34. Coiled coils of both proteins are involved in the interaction of SCOC and FEZ1 [35].

As a protein with large intrinsically disordered parts, FEZ1 is a difficult target for crystallography. At first, I aimed to purify a complex of full-length FEZ1 with SCOC. However, coexpression in *E.coli* BL21 led to insoluble proteins, probably due to formation of inclusion bodies. I therefore prepared several FEZ1 constructs comprising parts of the coiled coil domain (see Table 3.10). According to COILS predictions, the ccd of FEZ1 starts around residue 230 and ends at residues 290–300. Noticeable, this part of FEZ1 contains a lot of negatively charged residues and SCOC ccd also has a negative net charge. Hence, I cloned FEZ1 fragments with HisTag to add positive charge, so that SCOC and FEZ1 ccd would not repel each other in solution. FEZ1 ccd constructs were cloned with NdeI and XhoI restriction sites into vector pET22b (Novagen) using full length FEZ1 as template for PCR.

I expressed Strep-tagged SCOC (78–159) in combination with each of the FEZ1 constructs in *E.coli* BL21. Test expression was performed with Strep-tactin beads. The expression experiments revealed, that all of the His-tagged FEZ1 fragments were enriched together with Strep-tagged SCOC (78–159) in varying amounts. The minimal FEZ1 region interacting with SCOC (781–59) comprises residues 227–290 (Figure 3.35 A). Coexpression of FEZ1 (226–290) with SCOC (78–159) yielded a complex with equivalent amounts of both pro-

Table 3.10: FEZ1 fragments comprising the C-terminal ccd

N-terminal aa	C-terminal aa	length	net charge	net charge with HisTag	pI with His Tag
M227	L290	64	-5	-4/+2	5.94
M227	L295	69	-2	-1	6.65
H226	L290	65	-5/-4	-4/-3	6.03
H226	L295	70	-2/-1	-1/0	6.71
R225	L290	66	-4/-3	-3/-2	6.22
R225	L295	71	-1/0	0/+1	7.12
SCOC (78-159)		98	-5		4.98

**Figure 3.35: Test coexpression of Strep-SCOC (78-159) with His-FEZ1 ccd constructs**

(A) Coexpression of Strep-SCOC (78-159) and His-FEZ1 (227-290) in LB medium at 37 °C (B) Coexpression of Strep-COC (78-159) and His-FEZ1 (226-290) in LB medium at 37 °C

teins (Figure 3.35 B). Therefore, this complex was chosen for further investigation of the interaction.

3.2.1.2 Coexpression and copurification of SCOC–FEZ1 ccd complexes

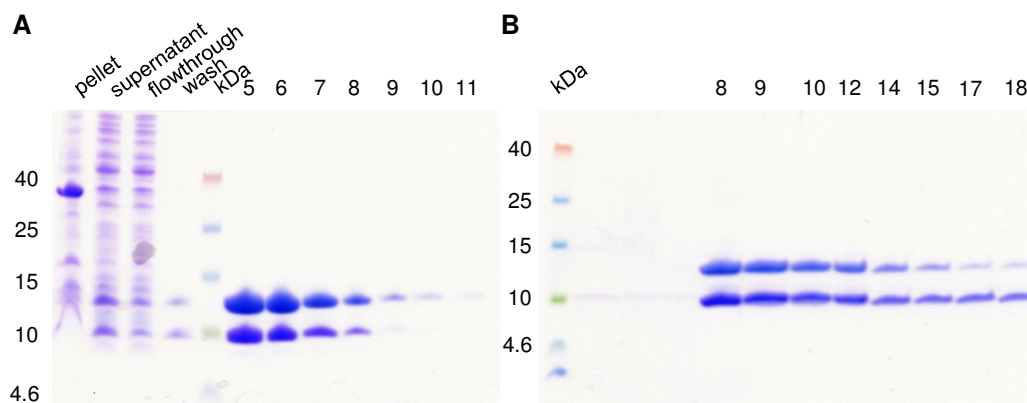


Figure 3.36: Copurification of wt Strep-SCOC (78–159) with His-FEZ1 (226–290)

(A) Schagger gel of affinity chromatography with StrepTrap column
(B) Schagger gel of size exclusion chromatography

Both FEZ1 (226–290) pET-22b and SCOC (78–159) pET-28a were co-transformed into BL21 by electroporation. The complex was expressed in LB medium at 37 °C. Cells were harvested and lysed. Affinity purification was conducted via StrepTrap chromatography. The protein complex eluted from the column in the same fractions 5–11 as SCOC (78–159) alone as shown in Figure 3.3 and Figure 3.36 B (see Section A.3.1 for elution profiles). The FEZ1 SCOC complex was further purified via size exclusion for SEC-MALLS (see Figure 3.36 A, C). The complex eluted from Superdex 75 16/60 column in a single peak. For further analysis, the complex was concentrated and used without freezing or further incubation, as it showed a tendency to precipitate and aggregate when protein samples were stored at 4 °C or –80 °C.

I assumed that oligomerization state and basic residues of SCOC would influence the interaction with FEZ1. Hence, SCOC double core mutants E93V/K97L and N125L/N132V, and arginine mutants R99E and R117E were also coexpressed and affinity purified with FEZ1 (226–290) via the same protocol as above (see Section A.3.2 for gels of affinity purification). Figure 3.37

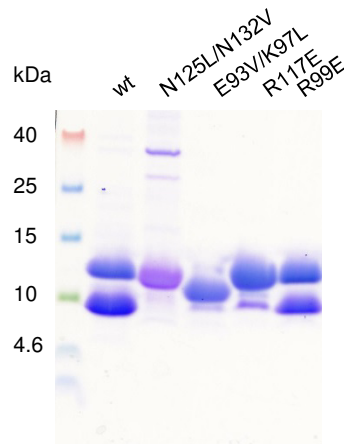


Figure 3.37: Coexpression of Strep-SCOC (78–159) constructs with His-FEZ1 (227–290)

Schägger gel with samples eluting from a StrepTrap column

shows copurification samples of all SCOC mutants with FEZ1 ccd. The tetrameric N125L/N132V and trimeric E93V/K97L mutants did not bind FEZ1 ccd showing that SCOC dimerization is crucial for SCOC–FEZ1 ccd complex formation. R117 is required for FEZ1 ccd interaction, since binding of R117E to FEZ1 ccd was almost completely abolished. The second arginine R99 is not important for complex formation, as R99E still interacted with FEZ1 ccd.

3.2.1.3 Characterization of SCOC–FEZ1 complexes

The formation of SCOC–FEZ1 ccd complex was analyzed by analytical gel filtration (Figure 3.38). The elution profile of the complex shows a considerable shift to shorter retention time compared to SCOC ccd alone. This indicates the formation of a stable complex with much larger molecular weight. Comparison to retention time of FEZ1 (227–190) was not possible, as the fragment was not expressed as soluble protein from *E. coli*, and could therefore not be purified.

The molecular weight of wt SCOC–FEZ1 ccd complex was analyzed with SEC-MALLS. The measurements yielded a molecular weight of 120.2 ± 4 kDa for the SCOC–FEZ1 ccd complex and showed that the complex is homogeneous (Figure 3.39). FEZ1 is a dimer in solution [44, 45, 111]. Assuming that both proteins are dimers and that they interact with a 1:1 stoichiometry, there would be six copies of each protein in the SCOC–FEZ1 complex.

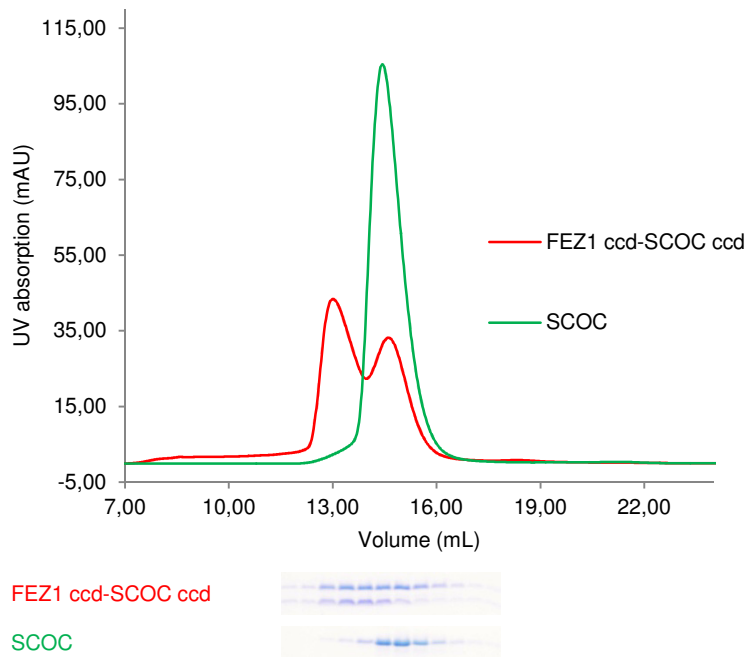


Figure 3.38: Analytical gel filtration of SCOC (78–159)—FEZ1 (227–290) complex

Elution profiles of SCOC (78–159)—FEZ1 (227–290) complex and SCOC (78–159) and the respective SDS-PAGE gels of the elution fractions

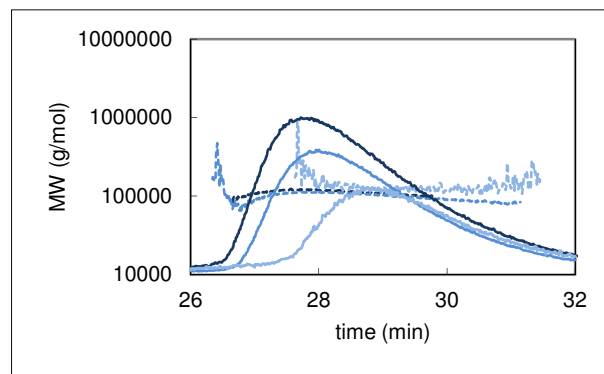


Figure 3.39: SEC-MALLS measurements of wt SCOC (78–159)—FEZ1 (227–290) complex

Elution profiles and molecular weights from three measurements are shown for the complex.

3.2.2 Interaction of SCOC with Arl1

3.2.2.1 Expression, purification and nucleotide exchange of Arl1

SCOC interacts with the small, Golgi-resident GTPase Arl1 in a nucleotide dependent manner [36]. For investigation of the interaction, I cloned dominant active Q79L Arl1 (15–181) into pET-22b. The protein was expressed overnight from autoinducible media. Cells were harvested and lysed. All buffers contained MgCl_2 for stable nucleotide binding of Arl1. In a first purification step, Arl1 was affinity purified via HisTrap chromatography (Figure 3.40 A, B). Minor contaminations were then removed by size exclusion chromatography (Figure 3.40 C, D).

SCOC is known to bind Arl1 in its GTP-bound state. Therefore, Arl1 had to be loaded with GTP or a non-hydrolyzable GTP derivative by **nucleotide exchange**. Several protocols are known for nucleotide exchange of GTPases [112]. For dominant active Arl1, a protocol with buffers supplemented with GTP was reported by Panic *et al.* [39]. Wu *et al.* [42] utilize wt Arl1 and convert it to GTP-bound state by adding EDTA, an alkaline phosphatase and excess of the non-hydrolyzable GDPNP. After His-Arl1 Q79L, which was purified in presence of GTP γ s, failed to interact with SCOC, I utilized a modified protocol providing excess GDPNP. I performed nucleotide exchange of Arl1 with a protocol utilizing SAP for hydrolyzing nucleotides. Small amounts of Arl1 were incubated with excess EDTA for chelating of Mg_2^+ , SAP and an excess of the non-hydrolyzable GTP derivate GDPNP. After incubation over night at 4 °C, the reaction was stopped with excess MgCl_2 . Protein samples were then subjected to another gel filtration purification step for the removal of excess nucleotide (Figure 3.40 E, F). Arl1 eluted from the column in two peaks. The first peak represented Arl1 which oligomerized via cysteine bridges, as the DTT concentration in the size exclusion buffer was not sufficient. The purification step of the fractions of the first peak was repeated with increased DTT concentration.

The nature of nucleotide-bound state was analyzed by **reversed-phase chromatography** runs of the protein and reference nucleotides. The measurements were conducted by Dr. Corinna Pohl (Department for Biophysical Chemistry, MPI-BPC). Analyses was done via a LiChrosorb column HiBar using a phosphate Acetonitril buffer (for a short description of the method, see Smith & Rittinger [112]). As references, equal amounts of GDP, GDPNP and GTP were used. Also, EF-Tu, a GDP-binding protein, was loaded on the column for comparison (Figure 3.41). GDPNP and the Arl1 sample eluted

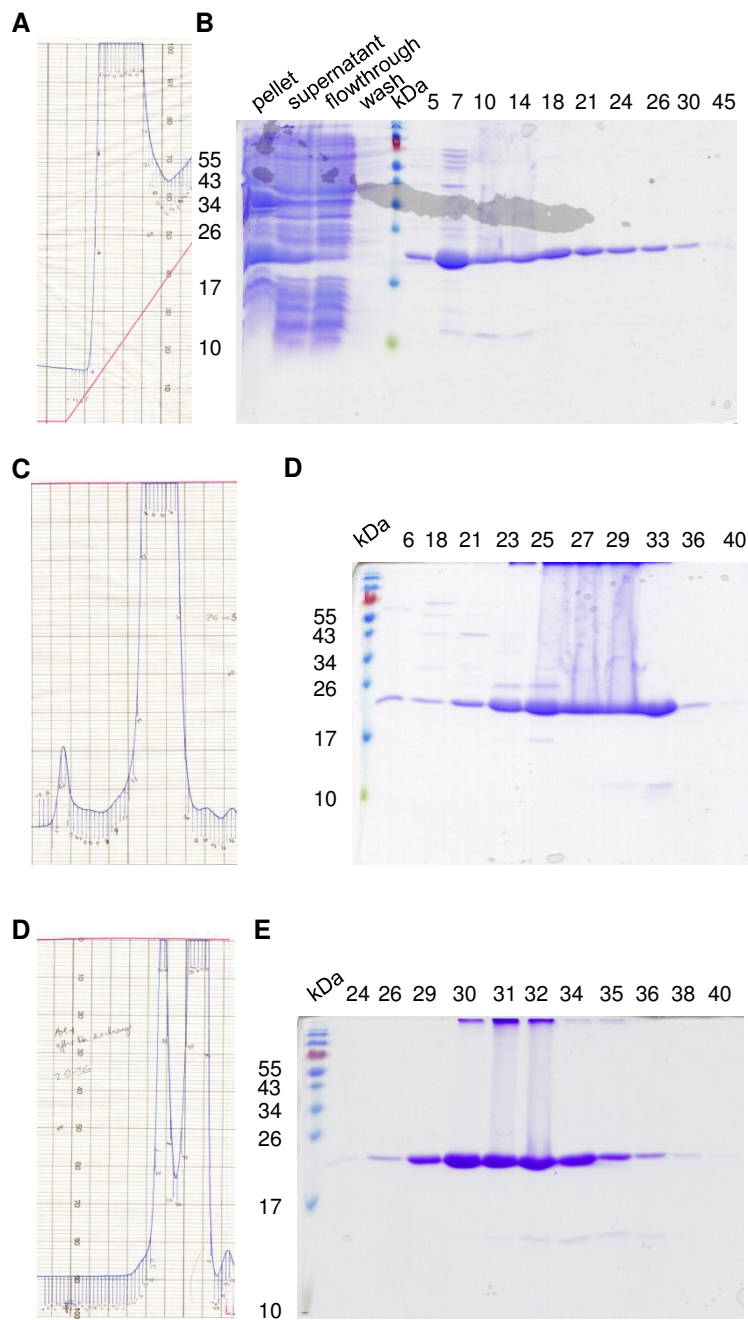


Figure 3.40: Purification of Q79L Arl1 (15–181)

(A) Elution profile of affinity chromatography (B) SDS-PAGE gel of affinity chromatography (C) Elution profile of first size exclusion step (D) SDS-PAGE gel of first size exclusion step (E) Elution profile of second size exclusion step (F) SDS-PAGE gel of second size exclusion step

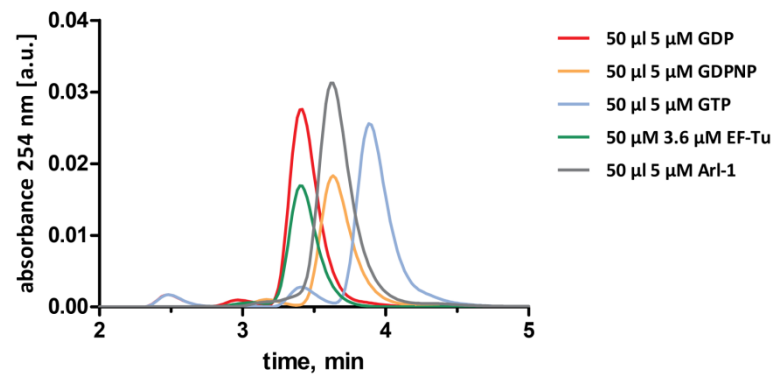


Figure 3.41: Analysis of bound nucleotide

Arl1, EF-Tu and reference nucleotides were separated via reversed-phase chromatography. Figure was produced by C. Pohl.

at same retention times, confirming that the nucleotide bound to Arl1 was GDPNP.

3.2.2.2 Arl1-Pulldown experiments

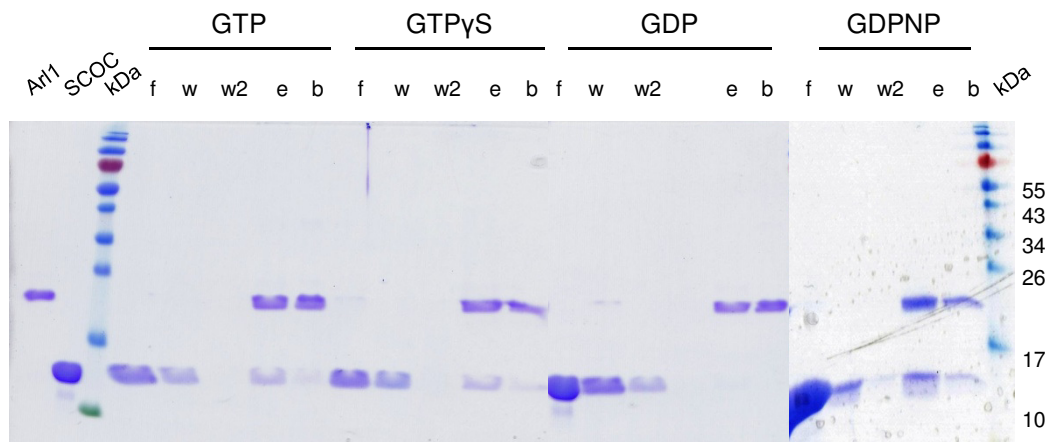


Figure 3.42: His-Arl1 pulldown with SCOC (78–159)

Arl1 was preincubated with an excess of GTP, GTP γ S, GDP and GDPNP respectively. SCOC (78–159) and Ni²⁺-sepharose beads were added for overnight incubation.

SCOC ccd interacts with Arl1 preferably in it's GTP-bound state, as shown

by yeast to hybrid screen and pulldown experiments [36]. I performed pull-down experiments with immobilized His-Arl1 (15–181) loaded with different nucleotides and SCOC (78–159) for confirmation of this interaction (Figure 3.42). Recombinantly expressed and purified proteins were incubated with Ni^{2+} -sepharose beads and an excess of the respective nucleotide. Flowthrough and wash fractions, and proteins bound to the beads were analyzed by SDS-PAGE.

GDP-bound Arl1 did not bind to SCOC. GTP and GTP γ s loaded Arl1 pulled SCOC's ccd in small amounts, GDPNP-Arl1 enriched SCOC in larger amounts on elution and beads fractions.

3.2.2.3 Analytical gel filtration

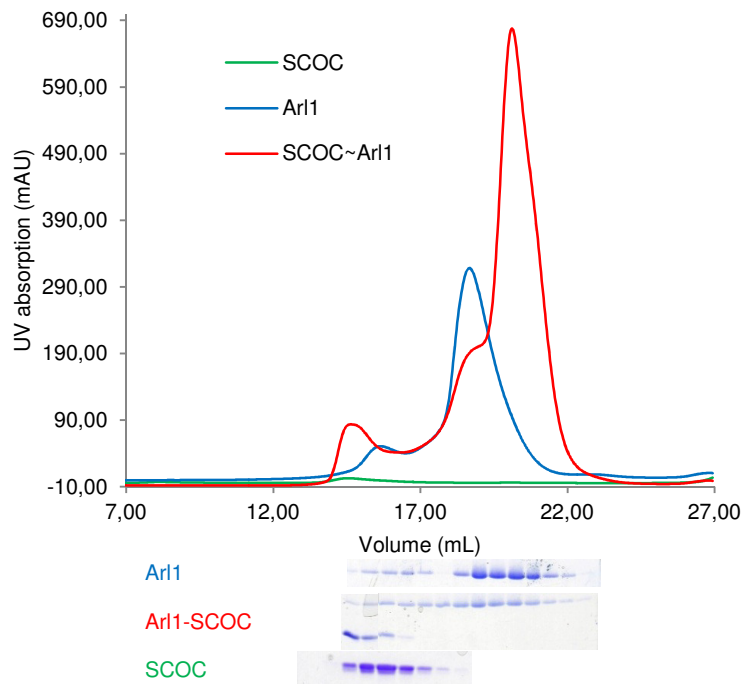


Figure 3.43: Analytical gel filtration

Elution profiles of SCOC (78–159), GDPNP-Arl1 (15–181) and 1:1 mixture of Arl1 and SCOC with GDPNP and the respective SDS-PAGE gels of the elution fractions

For further investigation of the interaction, the proteins were analyzed with

analytical gel filtration via Superdex 10/300 GL chromatography. A shift of retention volumes to larger molecular weight would indicate complex formation. SCOC (78–159), GDPNP-bound Arl1 and a 1:1 mixture of the proteins were separated on the size exclusion column (Figure 3.43). The protein mixture was preincubated together with an excess of GDPNP overnight at 4 °C. SDS-PAGE shows that SCOC elutes at ~ 15 mL, Arl1 elutes in two peaks at ~ 16 and 19 mL. The protein mixture sample shows three peaks in the elution profile. At ~ 14 mL, SCOC elutes from the column, a shoulder at ~19 mL indicates Arl1 elution. The large peak at ~ 20 mL corresponds to excess GDPNP.

This experiment showed, that the complex of GDPNP-Arl1 SCOC was not stable in a gel filtration experiment, as no peak shift in the elution profile was observed. This might indicate that the interaction of SCOC and Arl1 has a low affinity.

3.2.2.4 ITC

ITC experiments were conducted to prove whether SCOC and Arl1 interact with low affinity. In an ITC setup, the two interaction partners are slowly and stepwise mixed by titrating one protein in the syringe into the other protein in the cell. This occurs in an adiabatic jacket, so that small differences in temperature arising from the endo- or exothermic mixing process can be recorded. SCOC (615 μ M) was titrated into Arl1 (50 μ M). As a control, SCOC was also titrated into ITC buffer, and ITC buffer was titrated into Arl1, to exclude errors resulting from dilution heat (Figure 3.44 A). Titration curves showed, that SCOC and Arl1 interact in an exothermic manner. Also, a small dilution energy of SCOC was observed, and larger endothermic reaction for dilution of Arl1. However, analysis of the raw ITC data (Figure 3.44 B) for determination of affinity and stoichiometry could not be achieved. An accurate measurement of enthalpy of the reaction was not possible, as the curve was missing an initial plateau at the beginning of the titration curve.

Hence, it can be concluded that Arl1 and SCOC do interact in a GTP-dependent manner. Nevertheless, the detailed characterization of this interaction requires further experiments, including thorough analysis of Arl1's nucleotide state and optimization of ITC conditions.

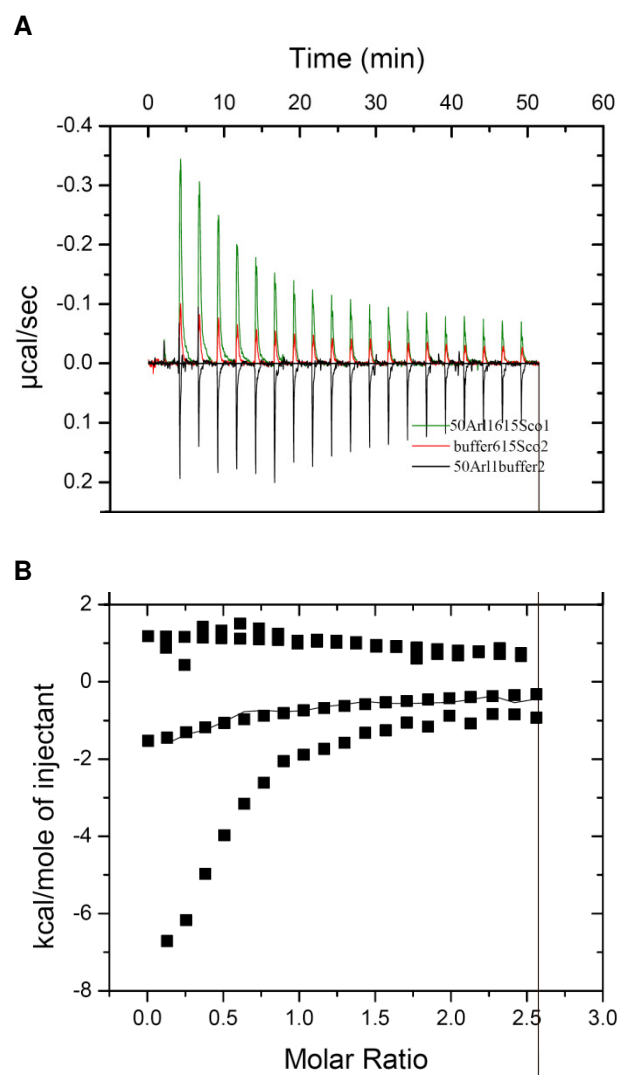


Figure 3.44: ITC titration of SCOC (615 μM) into Arl1 (50 μM)

(A) Raw data of SCOC titration into Arl1 (green line), SCOC into buffer (red line), and buffer into Arl1 (black line). (B) ITC analysis of raw data

3.3 Discussion

3.3.1 Characterization and structure of SCOC ccd

Here, I presented the crystal structure of SCOC ccd. The model comprises residues 86 – 147 and shows that SCOC is a left-handed coiled coil dimer. Heptad assignment of the coiled coil reveals several non-canonical, polar residues in the hydrophobic core, which are crucial for stability and oligomerization state of the coiled coil. Moreover, the *in vitro* interaction of SCOC to a minimal binding region of FEZ1 is also dependent on its oligomerization state as shown by copurification of both proteins.

Several SCOC constructs were used in this study in order to determine the crystal structure of SCOC ccd. Number of hits from screening and their best diffraction resolution limits are summed up in Table 3.11. A Strep-tagged SCOC ccd construct (78–159) covering the whole ccd yielded crystallization hits in more than 100 initial screening conditions. Although first crystals of this construct were easily obtained, it proved to be difficult to improve their diffraction quality beyond 3.2 |A| with low mosaicity. Limited proteolysis and structure prediction indicated flexible residues at the C-terminus of the ccd. Nevertheless, constructs with truncated C-termini did not result in better diffracting crystals. Optimized conditions yielded crystals diffracting to higher resolution, but problems with high mosaicity remained. Finally, *in situ* proteolysis of SCOC L105M (78–159) resulted in crystal rods diffracting to 2.8 Å.

From a SAD dataset collected from a single selenomethionine labeled SCOC L105M (78–159) crystal, I determined the structure of the SCOC coiled coil domain. Remarkably, the asymmetric unit contains 3 molecules, which form two distinct dimers with conformational differences. This plasticity is due to the polar and charged residues at the a/d-heptad positions. Imperfect coiled coils with polar residues in the hydrophobic coiled coil core have been observed before. For example, the autophagy protein Beclin1 also features an imperfect ccd with polar residues R and E in the hydrophobic core [113]. Mutation of these residues to canonical amino acids also led to increased stability. However, the crystal structure of Beclin1 features only one conformation of the coiled coil. The nature of the amino acids at the a/d-positions determines the oligomerization state of a coiled coil protein [55]. The influence of different amino acids at a/d heptad positions on the oligomerization state of a coiled coil protein was studied with the GCN4 leucine zipper in a landmark publication [59]. The GCN4 leucine zipper contains a single polar core residue (Asn16)

Table 3.11: Summary of SCOC ccd crystallization

construct	screening hits	resolution (Å)
SCOC (78–159)	> 100	3.6
24-well grid		3.2
96-well grid and random		3.6
w/o Tag	~ 80	3.0
SCOC (78–159) L105M		
<i>in situ</i> proteolysis		2.8
SCOC (78–151)	~ 40	5
SCOC (78–141)	~ 30	6
SCOC (78–132)	spherulites	–
SCOC (112–159)	spherulites	–

at an a-position and the authors concluded that Asn16 imposed specificity for dimerization of the GCN4 leucine zipper at the expense of its stability. My results are in agreement with these data. The well conserved, non-canonical a/d heptad residues in SCOC were replaced with either leucines or valines. Double core mutations E93V/K97L and N125L/N132V led to a change from dimer to either trimer or tetramer formation and the thermostabilities of these core mutants were dramatically increased. I did not create single mutation variants of the core positions, therefore I cannot conclude whether the mutation of a single core residue would have been sufficient for change of the oligomerization state, as observed for GCN4. As expected, single mutations of conserved arginine sites on c and g positions at the surface of the protein (R99 and R117), which might be involved in ionic interactions, did not influence stability and oligomerization state.

3.3.2 Interaction with FEZ1 and Arl1

The SCOC–FEZ1 complex plays a regulatory role for induction and progression of starvation induced autophagy [35]. Complex formation is mediated through the coiled coil domains of SCOC and FEZ1 [35, 37, 44]. Here I demonstrated that SCOC ccd and FEZ1 ccd form a stable homogeneous complex with a molecular weight of 120 kDa, which would correspond to six copies of each

molecule assuming a 1:1 stoichiometry. A complex of full-length GFP-FEZ1 with FLAG-SCOC was analyzed with blue native gels by McKnight *et al.* [35]. A molecular weight of 300 kDa was detected, corresponding to a 2:2 complex. Hence, the stoichiometry of a complex formed under *in vivo* conditions cannot be conclusively evaluated. I tested whether oligomerization state of SCOC influenced complex formation *in vitro* by copurification experiments. Indeed, complex formation was abolished for both double core mutants N125L/N132V and E93V/K97L, forming a tetramer and a trimer, respectively. As these residues are buried in the hydrophobic core, it is very unlikely that they are directly involved in intermolecular contacts to FEZ1. Dimeric oligomerization state is thus crucial for the interaction *in vitro*. Also, the single arginine mutation R117E showed decreased binding to FEZ1 ccd. I conclude that R117 is essential for SCOC–FEZ1 interaction *in vitro*, probably mediating ionic interactions. Co-immunoprecipitation experiments with full-length FEZ1 and EmGFP-SCOC constructs were performed by Dr. J. Chua (Department of Neurobiology) in order to gain insights into the complex formation *in vivo*. For this purpose, full length V5-FEZ1 and EmGFP SCOC wt and mutant constructs were cotransfected in HEK cells. Co-immunoprecipitation was carried out with anti-GFP antibodies. The tetrameric core mutant N125L/N132V did not bind FEZ1 (Figure 3.45). However, in contrast to the copurification studies both R117E and the trimeric SCOC mutant bound full-length FEZ1 in the co-immunoprecipitation experiments. R117E is involved in interactions of both coiled coil domains, but the abolishing effect of a single mutation *in vitro* is *in vivo* rescued by other factors stabilizing the interaction. Based on these results I cannot exclude that the N-terminal region of FEZ1 might also be involved in SCOC binding, which was in fact reported earlier for NEK1 (Nima-related kinase 1). The coiled coil region of NEK1 comprising residues 497–555 interacts with both the coiled-coil region of FEZ1 and the N-terminal region of FEZ1 [114].

Interestingly, SCOC wt co-immunoprecipitated with three distinct FEZ1 bands co-immunoprecipitated with SCOC wt (Figure 3.45), corresponding to three differently phosphorylated FEZ1 variants. Four phosphorylation sites are known, S58, S301, S314 and S316 [48]. Binding of FEZ1 to Kinesin-1 and Munc18 is regulated by phosphorylation of the major phosphorylation site S58 [48]. SCOC bound preferably to the lowest FEZ1 band, indicating a preference for FEZ1 in its multi-phosphorylated state. How phosphorylation of FEZ1 might interplay with the recruitment of the SCOC–FEZ1 complex to the phagophore remains to be elucidated.

McKnight *et al.* [35] showed, that SCOC and FEZ1 together interact with

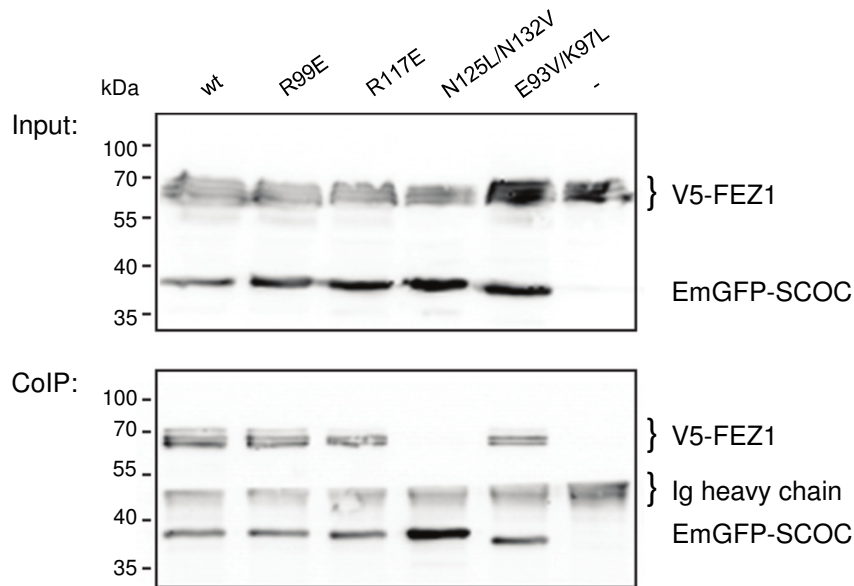


Figure 3.45: Co-immunoprecipitation of full length FEZ1 tagged N-terminally with a V5 epitope and various EmGFP-SCOC (78–159) variants from transfected 293 cell lysates using an anti-GFP antibody

ULK1 and UVRAG (see Section 1.2.1) to regulate induction and progression of macroautophagy. The crosstalk of Beclin1- and ULK1-complexes is mediated by phosphorylation: ULK1 phosphorylates AMBRA1, a member of a Beclin1 complex, which promotes autophagy[115]. Moreover, it was shown recently, that ULK1 phosphorylates also Beclin1 itself, thereby activating the Vps 34 complex for autophagy[116]. ULK1- and Beclin1-complexes and their interactions are discussed in Wirth *et al.* [117]. The interplay between these two essential autophagy complexes is not completely understood. Several functions of SCOC were proposed by McKnight *et al.* [35]:

1. upon aa starvation, SCOC binds to FEZ1–ULK1, which releases ULK1 complex
2. upon aa starvation, UVRAG dissociates from FEZ–SCOC, thereafter forming an UVRAG–Beclin1–PtdIns3K complex
3. SCOC might also play a role as membrane-proximal scaffolding protein, maybe in interplay with Arl1 at the Golgi.

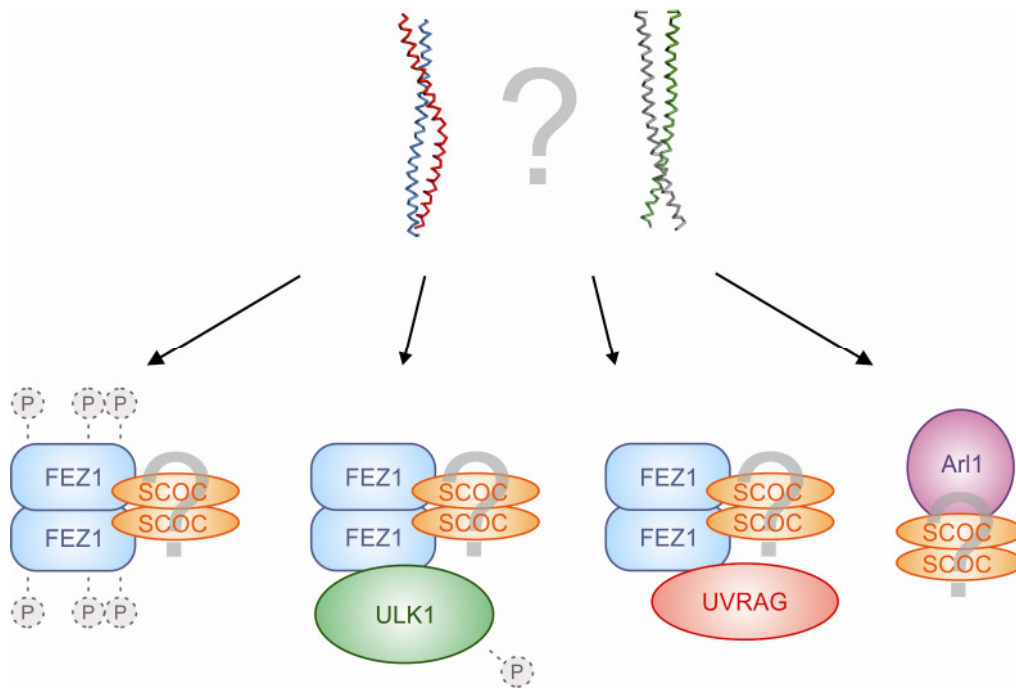


Figure 3.46: SCOC complexes in autophagy and beyond

Functional implications of distinct SCOC dimers remain unclear

Complexes which contain SCOC are depicted in Figure 3.46. With our current knowledge, it remains unclear whether the distinct SCOC dimers observed in the crystal structure also exhibit distinct physiological functions. SCOC might change its conformation upon binding to one or the other interaction partner. Also (de)phosphorylation of FEZ1 or ULK1 might change the conformational state, maybe triggering dissociation of the complex. One can also speculate that SCOC functions as a FEZ1 effector, locking the intrinsically disordered protein in a conformation in which it is less affine for binding ULK1 and UVRAG. How FEZ1 and SCOC are recruited to the ULK1 complex or the phagophore is still unknown.

In this work, I also analyzed the interaction of SCOC with Arl1. The physiological function of this interaction is less understood than those of FEZ1-SCOC complexes. SCOC colocalizes with Arl1 at the Golgi [36], but upon aa starvation, it is recruited to the phagophore, partially colocalizing with Atg9 and LC3 [35]. I was able to confirm interaction of SCOC with Arl1 by pulldown experiments, which showed that SCOC binds to Arl1 preferably in its GTP-

bound state. Initial ITC measurements confirmed an exothermic interaction between Arl1 and SCOC, but more experiments are needed to gain insights into the stability, structure and function of this complex.

Further structural characterization of the SCOC–FEZ1 complex, combined with functional analysis of the larger assemblies with ULK1 or UVRAG, will help us to gain new insights into how this small protein fulfills its functions in complexes with diverse interaction partners.

3.4 Outlook

Currently it is unclear, whether the two distinct dimers observed in the crystal structure are functionally relevant and whether they are also present in solution. SCOC structure in solution can be further analyzed by NMR spectroscopy. A two dimensional HSQC spectrum of ^{15}N -labelled SCOC could display the presence of residues that exist in more than one conformation. That would be a strong hint for the physiological importance of the two dimers. The crystallization and structure determination of the FEZ1–SCOC ccd complex would reveal, which of the dimeric SCOC conformations interacts with FEZ1. The complex can be purified in large amounts and is homogenous as shown by SEC-MALLS. It might therefore be suitable for crystallization. Phasing of data would be possible by molecular replacement using the SCOC structure as a search model. My data show that FEZ1 ccd is sufficient but not exclusively responsible for interaction with SCOC. In our lab, a purification protocol of full-length FEZ1 was established by Dr. J. Chua.

A stable full length complex of recombinantly purified proteins can be subjected to biochemical and biophysical analysis. For example, by using SEC-MALLS information the stoichiometry of functional complex can be obtained. Proteolytic digest could lead to the identification of other FEZ1 regions interacting with SCOC. Alternatively, also detailed knowledge of the molecular structure of FEZ1 would be helpful to understand the interplay of both ccds. Human FEZ1, a mainly natively unfolded protein, has a presumably low success rate for crystallization. However, recent alignment studies revealed that there are also FEZ1 homologues in yeast. I subjected several FEZ1 homologues from yeast to analysis by the crystallography prediction software TarO (see Table 3.12).

Several yeast homologues yielded “high score” or “amenable” for crystallization likeliness, among others a homologue from *Lachancea thermotolerans*, a thermophilic yeast. Screening of one of these constructs might yield crystals

Table 3.12: Crystallization probability of yeast FEZ1 homologues

Sequence ID	Organism	crystallization propability	length
C5DHV8	<i>Lachancea thermotolerans</i>	high-scoring	240
Q6FVD2	<i>Candida glabrata</i>	high-scoring	251
C8Z3Y6	<i>Saccharomyces cerevisiae</i>	amenable	280
B5VDW9	<i>Saccharomyces cerevisiae</i>	amenable	280
Q6CTM1	<i>Kluyveromyces lactis</i>	amenable	261

suitable for structure determination. Importantly, all of the yeast species have a SCOC homologue.

The interaction of SCOC with Arl1 should be thoroughly analyzed by ITC. Initial ITC experiments confirmed the interaction, but due to the missing plateau at the beginning of the titration curve, specific values for affinity and stoichiometry could not be derived. The protein concentration of Arl1 needs to be adjusted for further experiments. There might still be a fraction of Arl1 not loaded with GDPNP but GDP. Either, nucleotide exchange has to be optimized, or Arl1 has to be more concentrated. Besides ITC studies, the protein complex could also be structurally characterized. Based on my results, it seems that the interaction has a low affinity. A gel-filtration stable Arl1–SCOC complex was not obtained so far. This circumstance, paired with the likeliness of SCOC to crystallize on its own, might make structural investigation by X-ray crystallography difficult. Again, two dimensional NMR experiments could provide further insights. Titration experiments with stepwise addition of Arl1 to a ^{15}N -labelled SCOC sample could reveal amide peaks of amino acids involved in the interaction.

For a complete picture of SCOC’s interaction with FEZ1, and their common interplay with complexes involved in autophagy, further experiments will be needed. We assume, that upon aa starvation, SCOC does not localize to the Golgi anymore, but is recruited together with FEZ1 to the phagophore. How is this recruitment regulated? What role does phosphorylation of FEZ1 and ULK1 play? Does the interplay of FEZ1–SCOC with different autophagy key players involve conformational changes of SCOC? Further research combining structural biology methods with biochemistry and cell biology will be needed for detailed understanding of SCOC’s function.

4 Shigella flexneri and Autophagy

4.1 Characterization of VirG

4.1.1 Expression and purification of VirG (353–758)

The overexpression in *E. coli* and purification of VirG was extremely challenging. No fragment used in this study could be expressed as a soluble protein from *E. coli*. All purification protocols involved denaturation and refolding to the native structure of the protein. Finding conditions for successful refolding of VirG *in vitro* has been the main challenge of this project. VirG (353–758) originated from a limited proteolysis experiment of VirG (52–758) conducted by Dr. K. Kühnel. The tryptic digest yielded VirG (353–758) as stable fragment.

The fragment was cloned from *Shigella flexneri*'s virulence plasmid with a C-terminal StrepTag into pET-28a. Expression of VirG (353–758) was tested under various conditions with different media and expression strains. Still no expression conditions yielding soluble protein were found.

VirG was expressed at 37 °C in LB for 3 h. Cells were harvested and lysed, and after the first centrifugation step the supernatant was discarded. The pellet was resuspended and washed with Triton-X. Then, inclusion bodies were solubilized with 6 M GdnHCl. After dropwise dilution to 2 M GdnHCl, further refolding was performed by stepwise dialysis summing up to ~ 36 h dialysis time. Between dialysis steps, large amounts of precipitate were removed. After final dialysis affinity StrepTrap chromatography was conducted. The protein eluted together with a ~ 26 kDa degradation product (Figure 4.1 A). Next, size exclusion chromatography was performed, by which the degradation product was partially removed (Figure 4.1 B). Addition of EDTA and benzamidine as protease inhibitors only confined the problem. Due to degradation and precipitation, yields of the purification protocol varied and were unreproducible.

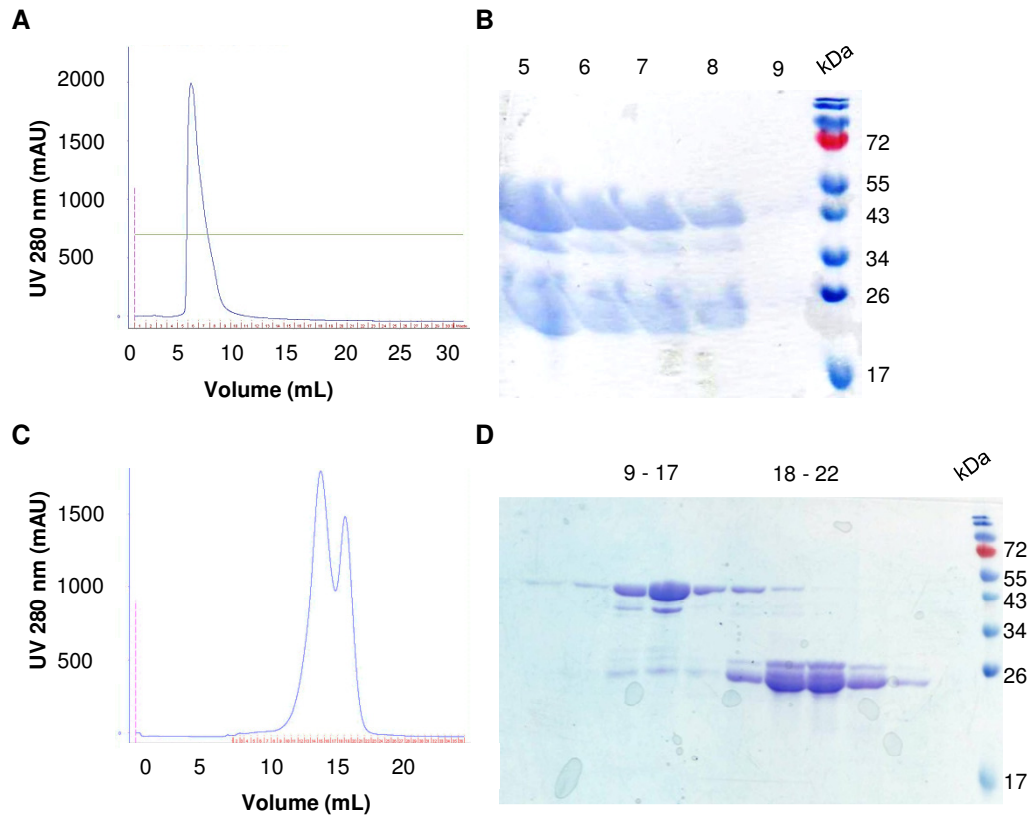


Figure 4.1: Purification of VirG (353–758)

(A) Elution profile of affinity chromatography (B) SDS-PAGE gel of affinity chromatography (C) Elution profile of size exclusion chromatography (D) SDS-PAGE gel of size exclusion chromatography

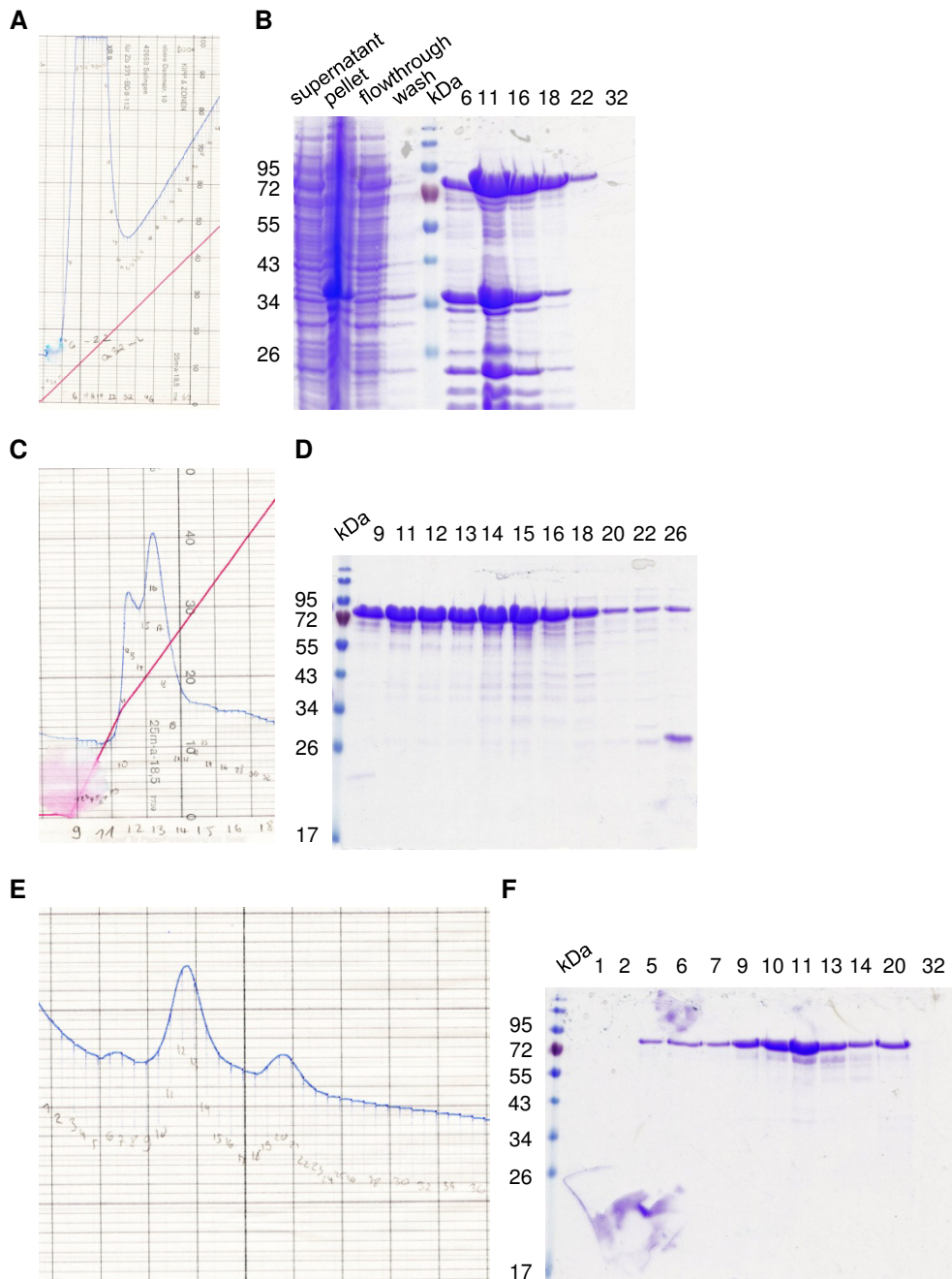


Figure 4.2: Purification of VirG (52–758)

(A) Elution profile of affinity chromatography (B) SDS-PAGE gel of affinity chromatography (C) Elution profile of anion exchange chromatography (D) SDS-PAGE gel of anion exchange chromatography (E) Elution profile of size exclusion chromatography (D) SDS-PAGE gel of size exclusion chromatography

4.1.2 Expression and purification of VirG (52–758)

VirG's passenger domain was cloned from a synthetic gene optimized for *E. coli* expression with a N-terminal His-Tag in pET-28a. Expression conditions were tested extensively. I tested different media (LB, TB, autoinducible and minimal media), a variety of *E. coli* expression strains (BL21, Rosetta, Lemo (NEB) and SoluBL21 (amsbio)), and a range of temperatures from 16 °C – 37 °C and also C- or N-terminal affinity tags. Nevertheless, the protein remained in the membrane fraction. Therefore, VirG (52–758) was expressed in LB at 37 °C, presumably in inclusion bodies. After harvesting and lysing the cells, the sample was denatured with 6 M urea. Affinity chromatography (Figure 4.2 A & B) was carried out with a gradient from 4 to 2 M urea. VirG (52–758) eluted from the Protino® Ni-NTA Column 5 mL in large amounts, but also with some contamination (Figure 4.2 B). Refolding was further performed via dialysis to 1 M urea. Contaminants were removed with the second purification step through anion exchange chromatography (Figure 4.2 C & D). The protein precipitated partially while concentrating for size exclusion chromatography. Also, removal of urea below 1 M resulted in precipitation. Size exclusion chromatography was carried out with 1 M urea. Despite addition of 2 mM DTT, VirG (52–758) eluted in two peaks from the gel filtration column, both containing the passenger domain (Figure 4.2 E & F). Degradation of the passenger domain was inhibited by the addition of benzamidine and EDTA as protease inhibitors. The purification yielded between 1.2 – 1.8 mg protein at a concentration of 10 – 15 mg/mL.

4.1.3 Analysis of refolding and secondary structure by CD spectroscopy

Refolding and secondary structure of VirG fragments was assessed by CD spectroscopy. The crystal structure of the C-terminal autochaperone region showed a β -barrel, which is typical for autotransporter proteins [85].

The CD spectrum of VirG (353–758) was recorded from 190 to 250 nm with a bandwidth of 2 nm. It shows a minimum at 216 nm, confirming a β -strand rich secondary structure (Figure 4.3).

The CD spectrum for VirG (52–758) was measured from 200 to 250 nm in an NaF buffer supplied with 1 M urea. Due to the urea concentration, the signal becomes noisy towards lower wavelength (Figure 4.4). Nevertheless, the spectrum also clearly shows a minimum at 216 nm, indicating successful refolding for the passenger domain.

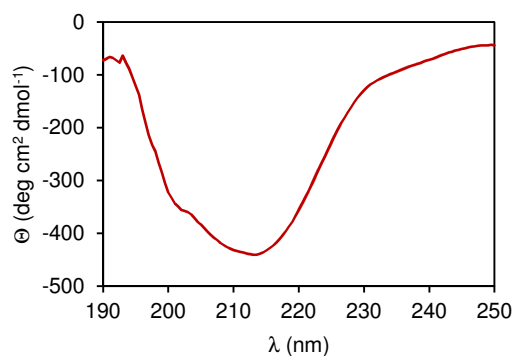


Figure 4.3: CD spectrum of VirG (353–758) at 20 °C

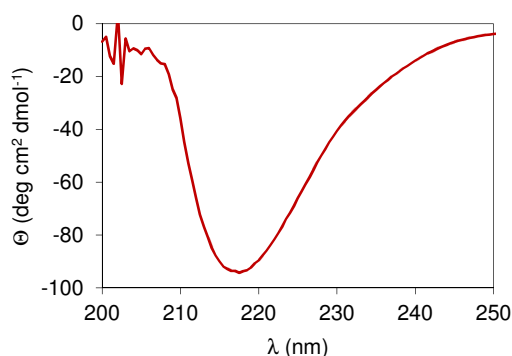


Figure 4.4: CD spectrum of VirG (52–758) in 1 M urea at 20 °C

4.1.4 Crystallization of VirG

96-well plate screens were set up with VirG (353–758) at 4 °C with protein concentrations of 7.5 and 15 mg/mL. The screens yielded no crystals, but some promising conditions with spherulites or microcrystalline structures (Figure 4.5). Crystallization conditions can be found in Table 4.1. Spherulites grew from clear drops (Figure 4.5 D) or precipitate (Figure 4.5 B) with crystallization conditions containing PEG 4000 or 8000 and Li_2SO_4 . The spherulites were stained to confirm the protein content (Figure 4.5 A). Microcrystalline structures were observed for one condition with Li_2SO_4 as precipitant (Figure 4.5 C).

Due to the low yield from purification, refinement in 24-well plate format was not possible. Also, the spherulites proved to be difficult to reproduce and

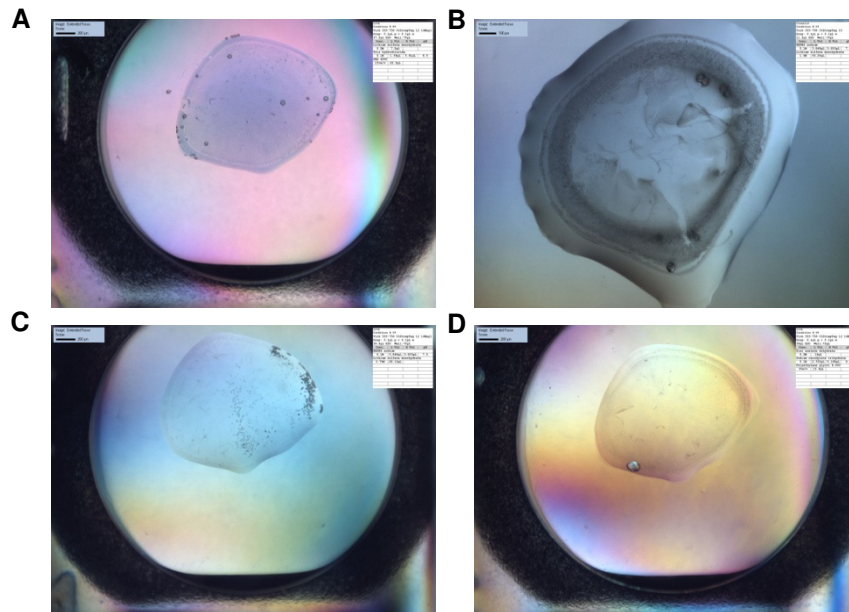


Figure 4.5: 96-well plate screening of VirG (353–758)

(A) Stained spherulites (B) Spherulites growing from protein precipitate
(C) microcrystals and small spherulites (D) spherulites

Table 4.1: Crystallization conditions of VirG (353–758) from 96-well plate screens

condition	buffer composition
A	0.2 M Li_2SO_4 0.1 M Tris pH 8.5 15% w/v PEG 4000
B	0.1 M HEPES pH 7.5 1.5 M $\text{Li}_2\text{SO}_4 \times \text{H}_2\text{O}$
C	0.1 M HEPES pH 7.5 0.75 M $\text{Li}_2\text{SO}_4 \times \text{H}_2\text{O}$
D	0.2 M $\text{ZnAc}_2 \times 2 \text{H}_2\text{O}$ 0.1 M Na cacodylate $\times 3 \text{H}_2\text{O}$ pH 6.5 9% w/v PEG 8000

could not be improved.

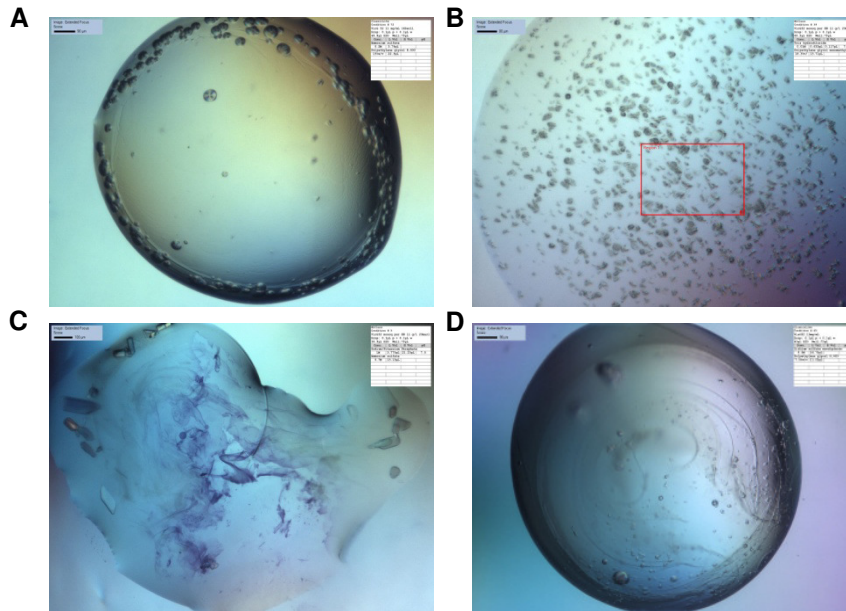


Figure 4.6: 96-well screening of VirG (52–758)

(A) Spherulites (B) Spherulites or microcrystals (C) Salt crystals (D) Microcrystals and spherulites

Drops were also set with VirG (52–758) at 4 and 20 °C with protein concentrations of 11 and 15 mg/mL. The protein buffer contained 1.0 M urea, so that final concentration of urea in the 1:1 sitting drops was 0.5 M urea. The fragment also yielded several conditions with almost crystalline structures (Figure 4.6), buffer compositions can be found in Table 4.2. Spherulites in different size and microcrystals were observed (Figure 4.6 A, B and D). The crystals in Figure 4.6 C turned out to be salt crystals, as they were not stained with protein dye.

Table 4.2: Crystallization conditions of VirG (52–758) from 96-well screens

condition	buffer composition
A	0.2 M NH_4SO_4 15% PEG 8000
B	0.01 Tris pH 7.5 18.3% w/v PEG MME
C	0.7 M $(\text{NH}_4)_2\text{SO}_4$
D	0.5 M Li_2SO_4 7,5% PEG 8000 1 M $(\text{Na/K})_3\text{PO}_4$

4.2 Characterization of IcsB–IpgA

4.2.1 Expression and purification of IcsB–IpgA

IcsB is translocated via the TTSS together with its chaperone IpgA. *In vivo*, the stop codon between the two genes on the virulence plasmid of *Shigella* is transient and can be read-through, creating a translational fusion protein [89]. Therefore, IcsB was cloned together with IpgA into the two cloning sites of pETDuet-1 for coexpression. For limited proteolysis experiments, I prepared two plasmids comprising N- or C-terminally Strep-tagged IpgA. IcsB was His-tagged via the vector.

Expression tests showed that IcsB–IpgA was expressed best in autoinducible media with overnight incubation at 18 °C. Cells were harvested and lysed. Affinity purification was carried out via HisTrap chromatography. It yielded a complex IcsB–IpgA with minor contaminations (Figure 4.7 A). The complex, which was prone to precipitation above 4 °C, was slowly concentrated and then subjected to size exclusion chromatography. IcsB–IpgA eluted together in a single peak from HiLoad 16/60 Superdex 200 column (Figure 4.7 B).

Furthermore, I also created four plasmids with shortened IcsB fragments IcsB (126–494), IcsB (185–494), IcsB (235–494) and IcsB (293–494) in MCSI and IpgA in MCSII. The stable fragments originated from a limited proteolysis experiment conducted by Dr. K. Kühnel. Shortened constructs of IcsB were expressed together with IpgA under same conditions. I followed the same

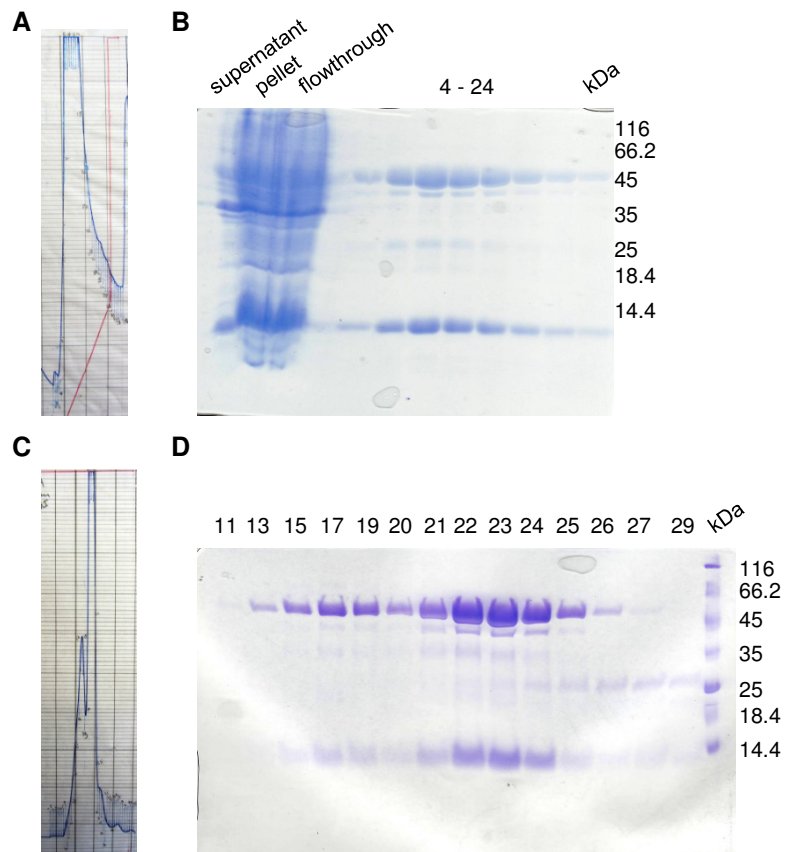


Figure 4.7: Purification of IcsB-IpgA

(A) Elution profile of affinity chromatography (B) SDS-PAGE gel of affinity chromatography (C) Elution profile of size exclusion chromatography (D) SDS-PAGE gel of size exclusion chromatography

purification protocols. Purification revealed that shorter IcsB fragments bound IpgA to less extend, indicating that the N-Terminal region of IcsB is involved in IcsB–IpgA interaction.

4.2.2 Limited proteolysis of IcsB–IpgA

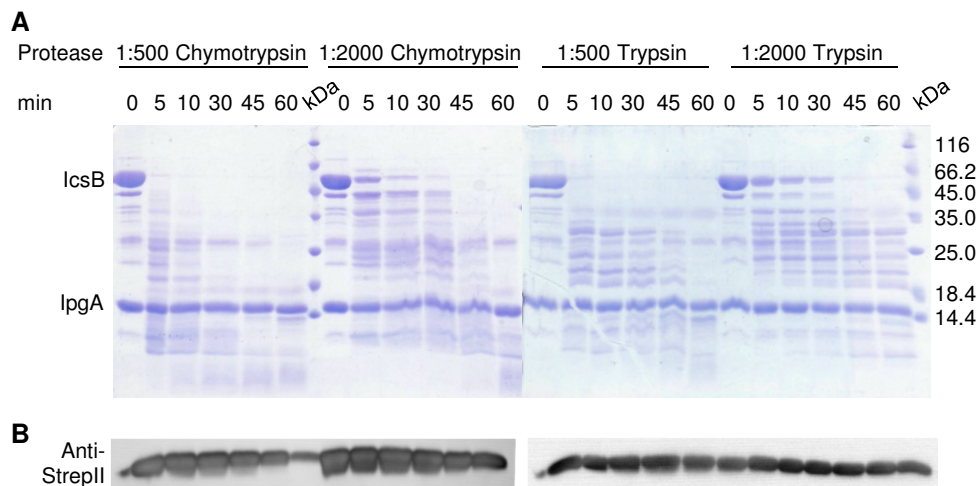


Figure 4.8: Limited proteolysis of IcsB–IpgA by Trypsin and Chymotrypsin

(A) SDS-PAGE analysis of indicated time points (B) Anti-Strep western blot showing Strep-IpgA bands

Limited proteolysis was performed in search of stable fragments of IcsB and IpgA, or smaller, more stable complexes IcsB–IpgA. For proteolytic digest, a complex of His-tagged IcsB with N-terminal Strep-tagged IpgA was utilized. The protein complex was subjected to digest by Trypsin and Chymotrypsin at 1:500 and 1:2000 w/w ratios. Progress of proteolysis was assessed by taking samples after different time points. For comparison, also undigested protein sample (time = 0 min) was loaded on gels. Analysis by SDS-PAGE gels showed, that IcsB is rapidly digested by both proteases (Figure 4.8). Also smaller fragments created by the proteases are no more resistant than the full-length protein. In contrast, no proteolytic products of IpgA were observed. Western blot against the StrepTag of IpgA confirmed the integrity of the protein (Figure 4.8 B). Hence, I concluded, that IpgA is well structured and stable, whereas IcsB is more flexible and might comprise disordered regions.

4.2.3 Crystallization of IcsB-IpgA

Extensive crystallization screening was performed with IcsB-IpgA complexes. 1:1 drops were set in 96-well plates at 4 °C with various protein concentrations between 10 – 50 mg/mL. I conducted screening with all routinely available screens for the full-length complex, and to less extent also with the shortened IcsB fragments IcsB (126–494), IcsB (185–494), IcsB (235–494) and IcsB (293–494). So far, the screens did not return a hit.

4.2.4 Analysis of IcsB-IpgA by heteronuclear NMR spectroscopy

Considering the results from limited proteolysis of IcsB-IpgA and the lack of the crystallization hits, I speculated that IcsB is a protein with natively unfolded, flexible regions. The lack of secondary structure can in general be addressed by CD spectroscopy. However, the formation of the very stable IcsB-IpgA complex was likely to prevent any conclusions from a CD spectrum, as it would contain overlaid signals from both proteins. Hence, I aimed to analyze the structural features of IcsB-IpgA by NMR spectroscopy in collaboration with Dr. S. Bibow and Prof. M. Zweckstetter.

For this purpose, I expressed and purified ¹⁵N labelled IpgA, and ¹⁵N labelled IcsB-IpgA. A heteronuclear single quantum coherence (HSQC) experiment of the complex was performed. In a resulting (optimal) spectrum, every proton attached to a nitrogen atom would give a signal, hence displaying all protons in peptide bonds as well as protons from side chains containing nitrogen. The HSQC spectrum of IcsB-IpgA shows some well-dispersed peaks, but also an overlap of several peaks in the middle of the spectrum (Figure 4.9). This indicates that large parts of the protein complex are flexible.

In addition, also transverse relaxation optimized spectroscopy experiments with both the complex and IpgA alone were conducted (Figure 4.10). The IpgA spectrum shows mostly well-dispersed peaks. Comparison of the spectrum of IpgA with the spectrum of IcsB-IpgA indicates that the overlapping peaks arising from 30–40 flexible residues must emerge from IcsB. Data collection was generally hampered by the fact that the complex precipitated at 20 °C, lowering the signal of data collection within 12 h acquisition time.

I concluded that IcsB contains flexible regions of unknown size and might not be a suitable target for crystallography.

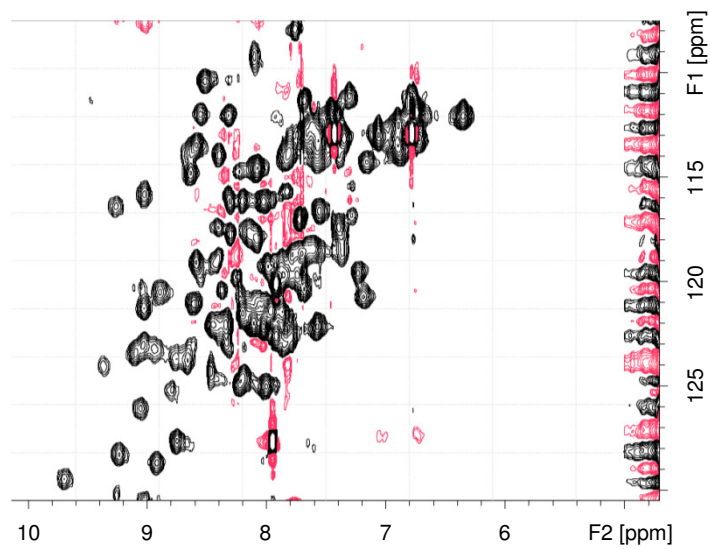


Figure 4.9: Two-dimensional HSQC spectrum of IcsB-IpgA

Spectrum was collected by Dr. S. Bibow at 5 °C with 16 scans, 300 increments in the indirect ^{15}N dimension and an acquisition time of 3h. Figure was produced by Dr. S. Bibow.

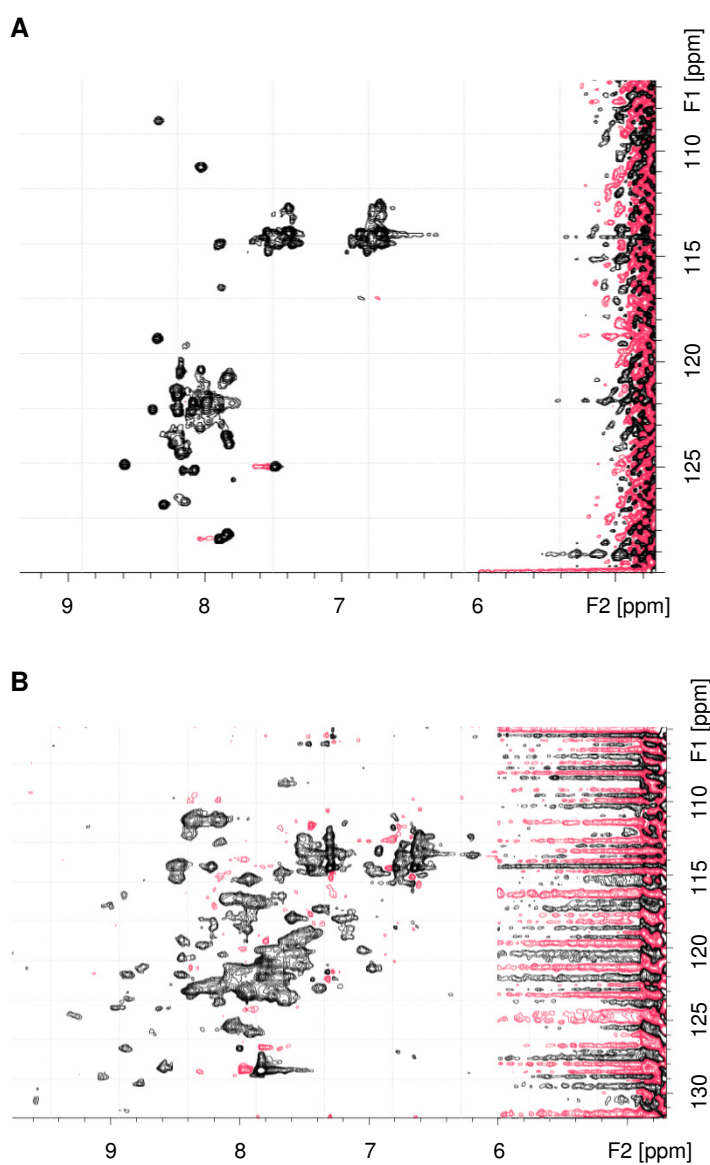


Figure 4.10: Two-dimensional TROSY spectra of IcsB-IpgA and IpgA

(A) IcsB-IpgA TROSY spectrum was collected by Dr. S. Bibow at 20 °C with 128 scans, 256 increments and an acquisition time of 12 h. (B) IpgA TROSY spectrum was collected by Dr. S. Bibow at 15 °C with 32 scans, 105 increments and an acquisition time of 2.5 h. Figure was produced by Dr. S. Bibow.

4.3 Discussion and Outlook

I aimed to structurally characterize two of *Shigella flexneri* virulence factors, VirG and IcsB, which are both part of *Shigella*'s sophisticated mechanism to escape xenophagy (see Section 1.3.2). None of the VirG proteins used in this study was expressed as soluble protein. The C-terminal autochaperone region (591–758), which was purified and crystallized by Dr. K. Kühnel, is so far the only solubly expressed fragment of VirG [85]. I was able to successfully purify two constructs of VirG.

VirG (353–758) was successfully refolded via a stepwise protocol including several dialysis steps from 2 – 0 M GdnHCl. Refolding was confirmed by CD spectroscopy. Although this progress was a milestone for the project, the purification protocol is still problematic. The entire purification spans five days, after which characterization by CD spectroscopy and other experiments with the fresh protein can only be started. I tried to confine degradation of the protein by the addition of protease inhibitors, and performed the entire purification at 4 °C in order to keep self-proteolysis to a minimum. However, the problem remained and resulted in very little and unreproducible protein yields. Moreover, the protocol also included the preparation of 12 L dialysis buffer, containing ultra-pure GdnHCl, which made the purification quite expensive. Hence, I tried to optimize the purification by using urea instead of GdnHCl, and aiming to reduce dialysis steps. I also tested purification of His-tagged VirG (353–758), which would allow refolding on the affinity column. However, all attempts to purify the construct via any other method than described above, failed. Nevertheless, I was able to set up a few crystallization screens which yielded spherulites. The refinement of these conditions was tedious due to the unreproducibility of both purification and crystallization results. Hence, although first steps towards a successful crystallization of VirG (353–758) were made, I continued optimization with VirG (52–758), which featured a simplified and less demanding purification protocol.

VirG (52–758) was refolded while eluting from the affinity column and dialyzed to 1 M final urea concentration. Under these conditions, the protein was fairly stable at 4 °C. Crystallization screens yielded a few promising conditions, mostly comprising PEG 8000 or Li₂SO₄ as precipitant. First refinement attempts did not yield improved crystallization hits. More experiments are needed to optimize the crystallization of VirG (52–758).

IcsB–IpgA and its shortened fragments were subjected to extensive crystallization screening without any hit. Initial characterization of IcsB–IpgA and IpgA by two-dimensional NMR spectroscopy revealed, that IcsB contains flex-

ible residues. I concluded that IcsB comprises natively unfolded regions and might thus not be a suitable target for crystallography.

Further insights into the interaction of VirG and IcsB might be difficult to obtain at the current stage of the project. First attempts of *in vitro* complex formation by analytical gel filtration with buffer containing 1 M urea showed no interaction of both proteins. Several reasons are possible. Maybe, IcsB is already considerably unfolded in 1 M urea, hence it cannot interact with VirG. Another reason might be, that VirG passenger domain did not refold to its native fold, but adopted a misfolded structure. Also, the interaction of both proteins might feature low affinity, which could be addressed by ITC. ITC measurements however will be very demanding due to the large protein amounts needed. In addition, VirG (52–758) precipitates below 0.5 M urea, and IcsB–IpgA precipitates above 4 °C, further constraining possible experimental conditions.

IcsB has been shown to bind to cholesterol in the host cell’s membrane [118]. Upon deletion of IcsB’s cholesterol binding domain, *Shigella*’s evasion of autophagy is severed. VirG’s passenger domain is responsible for *Shigella*’s actin-based motility (see Section 1.3.3, Egile *et al.* [77], May & Morona [87], Suzuki *et al.* [119]) and also membrane-attached. Hence, interaction of both proteins might also require proximity of the outer bacterial membrane or the host cell’s membrane.

The involvement of septins into the entrapment of *Shigella*, and interference with the host cell’s inflammatory response by polyubiquitination of NEMO further complicate the picture on how *Shigella* is interacting with the host cell (see Section 1.3.2, Mostowy *et al.* [81], Ashida *et al.* [120]). Additional factors might modulate IcsB–VirG interaction. Structural insights into VirG’s passenger domain will help us to understand the interaction between the host cell and the microbe, and might in the long term help to find vaccines or antibiotics against bacillary dysentery.

Appendix

A.1 DNA sequences of synthetic genes

A.1.1 SCOC isoform 1

atgCGTcgccgtgtgttttctagccaggattggcgtgcttcaggatgggacgggatgggctttt
ttagtcgtcgtaccttctgtggtcgttcaggcgttcttgtcgtggacagctggttcaagttc
tcgcccggaaagtttctgctggttctctgctgctgcctgctccacaagccgaagatcatagcagt
cgtattctgtatcctcgccaaaatccctgctgcccgaatgatgaatgccgacatggatgccg
ttgatgccgaaatcagggtggaactggaggaaaaaacccgtctgatcaaccaagttctggagct
gcagcatacactggaagatctgtctgcccgtgtcgatgccgtaaaaggagaatctgaaactg
aaaagtgagaatcaagtgtgggcccagtatatcgaaaatctgatgagcgcctctagcgtctttc
aaaccaccgatacctaaatccaaacgtaataa

A.1.2 Arl1 full length Q79L

atgggtggctttttttcgagcatcttctcctccctgtttggtacacgtgaaatgcgtatcctga
ttctggggctggatgggtgctggcaaaaccacaatcctgtatcgtctgcaggttgggtgaagttgt
tacgacaatcccgacaatcggctttaatggtgaaaccgtgacctataaaaacctgaaattccag
gtgtgggacctgggaggtctgacatcaattcggccgtattggcgtgctattattctaaccagg
acgccgtgatttatgttggtgactcctgtgaccgtgaccgtattggcatcagcaaatccgaact
ggtggcaatgctggaagaagaggaactgcgcaaagcaattctggtggtattcgctaataaaca
gacatggagcaggctatgacaagctctgagatggccaactcactgggtctgccggctctgaaag
accgtaaatggcaatctttaaactcggccacgaaaggcacgggtctggacgaagcaatgga
gtggctggtagaaactgaaaagtcgccagtaa

A.1.3 VirG passenger domain (52–758)

A.1.3.1 part 1

tgctacacctctgagtggaacccaagaactgcatttttagcaggacaactatgagaaactgctg
acccgggtgatggctctgtctcctctgggagctggatgaagatggatggatgcttggtatatca
cctccagcaatccgagtcctgtctctctgtacaaaactgcgattaatagcgacatcatgattag

cgccgggtcatggtggagccgggtgacaacaacgatggtaacagctgtggtggtaacgggtggagat
tctatcacaggcagcgatctgtcgattatcaatcaagggatgatcctgggcggtagtggtggat
ctggggctgatcacaatggagatggtggcgaagccgtgacaggggacaacctgtttatcattaa
cggggaaatcatttcaggtggccacgggggtgactcttattcagactctgatggcggaaatggg
ggggatgctgtaaccggtgtgaatctgcctatcatcaataaaggaaccattagcgggtggcaatg
gcggtaacaattatggcgagggtgacggggggaatggcggggatgctatcacagggagtagcct
gagcgtgattaacaaaggcacctttgccgggggaaatggtggtgccgcctatggttatggctat
gacggctatggcgggaatgccattactggtgacaatctgtccgtgattaacaatggggcgattc
tgggcggaaatggaggacattgggggtgacgctatcaacggtagcaatatgactatcgccaatag
cgggtacattatcagcggtaagaggacgacgggtacacaaaatggtgccggaatgctatccac
attacgggtgggaacaattctctgatcctgcacgaaggttctgttattaccggagatgtccagg
tgaacaattccagcattctgaaaatcattaacaacgactatacgggtacaacaccgactattga
aggcgatctgtgtgctggtgattgtacaaccgtagcctgtcaggcaacaattcaccgtgagc
ggagatgtttcttttggggagaactcatcactgaacctggccggcattagtagcctggaagcaa
gcggcaatatgagctttggcaataacgtgaaagtggaagcgattattaacaactgggcacagaa
agactataaactgctgtccgccgataaaggcattaccggcttttagcgtgagcaatatcagtatt
atcaatccgctgctgaccactggtgctatcgattatacgaatcctatatctctgaccagaata
aactgatctatggtctgtcctggaacgataccgatggggattccatggcgagtttaacctgaa
agagaacgccgaactgaccgtcagtacaattctggccgataatctgagccatcacaacatcaac
tcgtgggatggcaaatctctgacgaaaagcgggtgagggtacactgattctg

A.1.3.2 part 2

agcctgacaaaatccggtgaaggaacgctgattctggccgagaaaaaacacctatagcggcttta
ccaatatcaatgccggcatcctgaaaatgggtacggttgaggcaatgactcgtacagcaggcgt
tatcgtgaataaagggtgccacgctgaacttttagcggatgaaccagaccgtgaacacactgctg
aatagtggtacagtgctgatcaacaacatcaacgctccgtttctgcctgatcctgtgattgta
cgggcaacatgaccctggaaaaaatggacacgtgattctgaataatagtagtagcaacgtggg
acagacctatgtacaaaaaggcaactggcatggcaaaaggggtattctgagcctgggagctggt
ctgggtaatgacaactccaaaaccgatcgtctggagattgctggacatgcttccggatcactt
atgtggccgtgaccaatgaaggcggtagtggggacaaaaccctggaggagtagacagattatcag
taccgatagcagcgacaaaaacgctttcatccagaaaggccgtattgtagcaggttcgtatgat
tatcgtctgaaacagggcacagtgagtggactgaacaccaacaatggtatctgaccagccaaa
tggacaaccaggagtctaaacaatgagcaaccaagagtcgacacagatgagtagtctg

A.2 Crystallization conditions of SCOC ccd constructs

(see next page)

Crystallisation conditions of SCOC (78-159) with StrepTag, native						
No_screem_well	Salt 1	Salt 2	Buffer	Precipitant 1	Precipitant 2	crystal characterisation
43_Classics_D4			0.1 M Sodium cacodylate pH 6.5	1.4 M Sodium acetate		big crystals
43_Classics_H4	0.2 M Magnesium chloride		0.1 M TRIS.HCl pH 8.5	30 %(w/v) PEG 4000		big crystals
44_PEGs_E5		0.2 M Magnesium chloride		20 %(w/v) PEG 3350		big crystals
44_PEGs_F5		0.2 M Sodium nitrate		20 %(w/v) PEG 3350		big crystals
45_pHClear_A8			0.1 M Citric acid 5.0	2.0 M Sodium chloride		big crystals
52_Anions_A11			0.1 M MES pH 6.5	0.6 M Sodium fluoride		big crystals
84_pHClearII_B6		1.0 M Lithium chloride	0.1 M BICINE 9.0	20 %(w/v) PEG 6000		big crystals
43_Classics_H5	0.2 M Lithium sulfate		0.1 M TRIS.HCl pH 8.5	30 %(w/v) PEG 4000		crystals
44_PEGs_F8		0.2 M Magnesium formate		20 %(w/v) PEG 3350		crystals
44_PEGs_G1		0.2 M Magnesium acetate		20 %(w/v) PEG 3350		crystals
44_PEGs_G9		0.2 M Sodium sulfate		20 %(w/v) PEG 3350		crystals
44_PEGs_H8		0.2 M di-Ammonium phosphate		20 %(w/v) PEG 3350		crystals
45_pHClear_B3			0.1 M MES 6.0	3.0 M Sodium chloride		crystals
45_pHClear_B4			0.1 M HEPES 7.0	3.0 M Sodium chloride		crystals
45_pHClear_D7			0.1 M Citric acid 4.0	30 %(w/v) PEG 6000		crystals
45_pHClear_G1			0.1 M Citric acid 4.0	10 %(v/v) MPD		crystals
52_Anions_E12			0.1 M MES pH 6.5	0.5 M Sodium succinate		crystals
52_Anions_F6			0.1 M HEPES pH 7.5	0.5 M Sodium succinate		crystals
52_Anions_H1			0.1 M HEPES pH 7.5	0.75 M Sodium sulfate		crystals
52_Anions_H4			0.1 M HEPES pH 7.5	0.6 M K/Na tartrate		crystals
53_Cations_B1			0.1 M HEPES pH 7.5	3.5 M Ammonium chloride		crystals
53_Cations_H8			0.1 M TRIS.HCl pH 8.5	2.0 M Potassium acetate		crystals
84_pHClearII_A12		1.0 M Lithium chloride	0.1 M BICINE 9.0	10 %(w/v) PEG 6000		crystals
84_pHClearII_C1			0.1 M Citric acid 7.5	5 %(v/v) Isopropanol		crystals
84_pHClearII_E5			7.5	0.8 M Na/K phosphate		crystals
84_pHClearII_E11			7.5	1.0 M Na/K phosphate		crystals
44_PEGs_F3		0.2 M Lithium nitrate		20 %(w/v) PEG 3350		large cluster
52_Anions_H10			0.1 M TRIS.HCl pH 8.5	0.6 M K/Na tartrate		large cluster
53_Cations_C8			0.1 M MES pH 6.5	2.25 M Ammonium acetate		large cluster
53_Cations_D8			0.1 M TRIS.HCl pH 8.5	2.25 M Ammonium acetate		large cluster
43_Classics_H1	0.2 M Ammonium sulfate		0.1 M Sodium acetate pH 4.6	25 %(w/v) PEG 4000		microcrystals

44_PEGs_F4		0.2 M Magnesium nitrate		20 %(w/v) PEG 3350		microcrystals
45_pHClear_C2			0.1 M Citric acid 5.0	5 %(w/v) PEG 6000		microcrystals
45_pHClear_C7			0.1 M Citric acid 4.0	10 %(w/v) PEG 6000		microcrystals
45_pHClear_D1			0.1 M Citric acid 4.0	20 %(w/v) PEG 6000		microcrystals
52_Anions_F5			0.1 M HEPES pH 7.5	1.0 M Sodium succinate		microcrystals
52_Anions_G5			0.1 M Sodium acetate pH 4.6	2.4 M Sodium thiocyanate		microcrystals
52_Anions_G10			0.1 M MES pH 6.5	0.6 M K/Na tartrate		microcrystals
53_Cations_E2			0.1 M Sodium acetate pH 4.6	1.1 M Calcium chloride		microcrystals
53_Cations_F9			0.1 M TRIS.HCl pH 8.5	3.2 M Sodium chloride		microcrystals
84_pHClearII_B5		1.0 M Lithium chloride	0.1 M TRIS 8.0	20 %(w/v) PEG 6000		microcrystals
84_pHClearII_G7			5.0	1.0 M Sodium malonate		microcrystals
43_Classics_F8	0.2 M Magnesium acetate		0.1 M Sodium cacodylate pH 6.5	20 %(w/v) PEG 8000		multiple crystals
43_Classics_H6	0.2 M Sodium acetate		0.1 M TRIS.HCl pH 8.5	30 %(w/v) PEG 4000		multiple crystals
44_PEGs_C7			0.1 M MES pH 6.5	25 %(w/v) PEG 3000		multiple crystals
44_PEGs_C8			0.1 M MES pH 6.5	25 %(w/v) PEG 4000		multiple crystals
44_PEGs_E7		0.2 M Calcium chloride		20 %(w/v) PEG 3350		multiple crystals
44_PEGs_F12		0.2 M Lithium acetate		20 %(w/v) PEG 3350		multiple crystals
44_PEGs_G8		0.2 M Magnesium sulfate		20 %(w/v) PEG 3350		multiple crystals
44_PEGs_G11		0.2 M Ammonium sulfate		20 %(w/v) PEG 3350		multiple crystals
45_pHClear_E12			0.1 M BICINE 9.0	1.6 M Ammonium sulfate		multiple crystals
52_Anions_B10			0.1 M TRIS.HCl pH 8.5	0.6 M tri-Sodium citrate		multiple crystals
52_Anions_F1			0.1 M HEPES pH 7.5	3.5 M Sodium nitrate		multiple crystals
53_Cations_H3			0.1 M HEPES pH 7.5	2.2 M Potassium chloride		multiple crystals
52_Anions_B2			0.1 M HEPES pH 7.5	1.25 M Sodium acetate		multiple crystals
53_Cations_C4			0.1 M Sodium acetate pH 4.6	1.25 M Lithium acetate		single crystals
53_Cations_C6			0.1 M Sodium acetate pH 4.6	1.0 M Magnesium acetate		single crystals
43_Classics_E11			0.1 M HEPES sodium salt pH 7.5	1.5 M Lithium sulfate		small crystals
43_Classics_F7	0.2 M Calcium acetate		0.1 M Sodium cacodylate pH 6.5	18 %(w/v) PEG 8000		small crystals
43_Classics_G8				10 %(w/v) PEG 1000	10 %(w/v) PEG 8000	small crystals
43_Classics_G11	0.2 M Ammonium sulfate		0.1 M Sodium acetate pH 4.6	30 %(w/v) PEG 2000 MME		small crystals
44_PEGs_A2			0.1 M Sodium acetate pH 4.6	30 %(v/v) PEG 300		small crystals
44_PEGs_B6			0.1 M Sodium HEPES pH 7.5	25 %(w/v) PEG 2000 MME		small crystals
44_PEGs_G6		0.2 M Ammonium acetate		20 %(w/v) PEG 3350		small crystals

45_pHClear_D10			0.1 M HEPES 7.0	30 %(w/v) PEG 6000		small crystals
52_Anions_B8			0.1 M TRIS.HCl pH 8.5	1.25 M Sodium acetate		small crystals
52_Anions_G4			0.1 M Sodium tartrate pH 5.6	0.6 M K/Na tartrate		small crystals
53_Cations_D2			0.1 M HEPES pH 7.5	2.25 M Ammonium acetate		small crystals
53_Cations_E1			0.1 M Sodium acetate pH 4.6	2.2 M Calcium chloride		small crystals
43_Classics_C1				0.4 M Ammonium phosphate		small multiple crystals
44_PEGs_F9		0.2 M Sodium formate		20 %(w/v) PEG 3350		small multiple crystals
44_PEGs_H9		0.2 M tri-Lithium citrate		20 %(w/v) PEG 3350		small multiple crystals
44_PEGs_H10		0.2 M tri-Sodium citrate		20 %(w/v) PEG 3350		small multiple crystals
45_pHClear_D3			0.1 M MES 6.0	20 %(w/v) PEG 6000		small multiple crystals
52_Anions_B4			0.1 M HEPES pH 7.5	0.6 M tri-Sodium citrate		small multiple crystals
53_Cations_D3			0.1 M HEPES pH 7.5	2.5 M Lithium acetate		small multiple crystals
84_pHClearII_E7			5.0	1.0 M Na/K phosphate		small multiple crystals
43_Classics_A5			0.1 M HEPES sodium salt pH 7.5	10 %(v/v) Isopropanol	20 %(w/v) PEG 4000	spherulites
43_Classics_F2			0.1 M MES pH 6.5	1.6 M Magnesium sulfate		spherulites
45_pHClear_B5			0.1 M TRIS 8.0	3.0 M Sodium chloride		spherulites
45_pHClear_B6			0.1 M BICINE 9.0	3.0 M Sodium chloride		spherulites
45_pHClear_E6			0.1 M BICINE 9.0	0.8 M Ammonium sulfate		spherulites
45_pHClear_G2			0.1 M Sodium acetate 5.0	10 %(v/v) MPD		spherulites
84_pHClearII_E6				8,20 0.8 M Na/K phosphate		spherulites

Cryo D6 Grid Screen: crystallisation conditions of SCOC (78-159) variants

well	SCOC ccd	SCOC ccd w/o Tag	SeMet SCOC ccd	Precipitant 1	Glycerol	pH: MES 0.085 M	KH ₂ PO ₄	NaH ₂ PO ₄
D1			big crystals	NaCl 1.3 M	15 %v/v	6.386	0.085 M	0.085 M
D3		crystals	big crystals	NaCl 1.427 M	15 %v/v	6.386	0.085 M	0.085 M
F3	small crystals		big crystals	NaCl 1.427 M	15 %v/v	6.643	0.085 M	0.085 M
G1		crystals	big crystals	NaCl 1.3 M	15 %v/v	6.771	0.085 M	0.085 M
G2	small crystals		big crystals	NaCl 1.364 M	15 %v/v	6.771	0.085 M	0.085 M
G3		big crystals	big crystals	NaCl 1.427 M	15 %v/v	6.771	0.085 M	0.085 M
H1			big crystals	NaCl 1.3 M	15 %v/v	6.9	0.085 M	0.085 M
F1	crystals	cluster	cluster	NaCl 1.3 M	15 %v/v	6.643	0.085 M	0.085 M
D2			crystals	NaCl 1.364 M	15 %v/v	6.386	0.085 M	0.085 M
B1			microcrystals	NaCl 1.3 M	15 %v/v	6.129	0.085 M	0.085 M
B3			microcrystals	NaCl 1.427 M	15 %v/v	6.129	0.085 M	0.085 M
C1		small crystals	microcrystals	NaCl 1.3 M	15 %v/v	6.257	0.085 M	0.085 M
G6			microcrystals	NaCl 1.618 M	15 %v/v	6.771	0.085 M	0.085 M
G7			microcrystals	NaCl 1.682 M	15 %v/v	6.771	0.085 M	0.085 M
E3			multiple crystals	NaCl 1.427 M	15 %v/v	6.514	0.085 M	0.085 M
H2		microcrystals	multiple crystals	NaCl 1.364 M	15 %v/v	6.9	0.085 M	0.085 M
H3			multiple crystals	NaCl 1.427 M	15 %v/v	6.9	0.085 M	0.085 M
E2			small crystals	NaCl 1.364 M	15 %v/v	6.514	0.085 M	0.085 M
G5			small crystals	NaCl 1.555 M	15 %v/v	6.771	0.085 M	0.085 M
C3			spherulites	NaCl 1.427 M	15 %v/v	6.257	0.085 M	0.085 M
D4			spherulites	NaCl 1.491 M	15 %v/v	6.386	0.085 M	0.085 M
D5		crystals	spherulites	NaCl 1.555 M	15 %v/v	6.386	0.085 M	0.085 M
F2	crystals		spherulites	NaCl 1.364 M	15 %v/v	6.643	0.085 M	0.085 M
F4	small crystals		spherulites	NaCl 1.491 M	15 %v/v	6.643	0.085 M	0.085 M
F5	small crystals		spherulites	NaCl 1.555 M	15 %v/v	6.643	0.085 M	0.085 M
F6			spherulites	NaCl 1.618 M	15 %v/v	6.643	0.085 M	0.085 M
G4	small crystals		spherulites	NaCl 1.491 M	15 %v/v	6.771	0.085 M	0.085 M
H4			spherulites	NaCl 1.491 M	15 %v/v	6.9	0.085 M	0.085 M
F8	small crystals			NaCl 1.745 M	15 %v/v	6.643	0.085 M	0.085 M

Cryo D6 Random Screen: crystallisation conditions of SCOC (78-159)							
well	screen_plate	crystal characteris	Buffer	Precipitant	Cryoprotectant	Salt 1	Salt 2
C5	87/1	big crystals	MES 0.085 M 6.265	NaCl 1.461 M		NaH2PO4 0.06122 M	KH2PO4 0.1714 M
C7	87/1	big crystals	MES 0.085 M 7	NaCl 1.494 M	Glycerol 14.29 %v/v	NaH2PO4 0.04898 M	KH2PO4 0.1673 M
F8	87/1	big crystals	MES 0.085 M 6.102	NaCl 1.314 M	Glycerol 15.1 %v/v	NaH2PO4 0.06122 M	KH2PO4 0.08163 M
B5	87/2	big crystals	MES 0.085 M 6.898	NaCl 1.706 M	Glycerol 19.59 %v/v	NaH2PO4 0.08163 M	KH2PO4 0.09388 M
F5	87/2	big crystals	MES 0.085 M 6.857	NaCl 1.592 M		NaH2PO4 0.1878 M	KH2PO4 0.02041 M
G3	87/2	big crystals	MES 0.085 M 6.449	NaCl 1.771 M		NaH2PO4 0.01224 M	KH2PO4 0.1469 M
A6	87/1	crystal	MES 0.085 M 6.878	NaCl 1.755 M	Glycerol 11.02 %v/v		KH2PO4 0.1265 M
D10	87/1	crystal	MES 0.085 M 6.755	NaCl 1.837 M	Glycerol 6.939 %v/v		KH2PO4 0.1429 M
G5	87/1	crystal	MES 0.085 M 6.857	NaCl 1.69 M	Glycerol 12.24 %v/v	NaH2PO4 0.1755 M	
E3	87/2	crystal	MES 0.085 M 6.551	NaCl 1.396 M	Glycerol 16.33 %v/v	NaH2PO4 0.1959 M	
G10	87/2	crystal	MES 0.085 M 6.469	NaCl 1.886 M			KH2PO4 0.1184 M
G2	87/2	crystal	MES 0.085 M 6.673	NaCl 1.69 M			KH2PO4 0.151 M
E5	87/1	crystal clusters	MES 0.085 M 6.592	NaCl 1.363 M	Glycerol 12.24 %v/v	NaH2PO4 0.1878 M	
E6	87/1	crystal clusters		NaCl 1.853 M		NaH2PO4 0.05714 M	
E7	87/1	crystal clusters	MES 0.085 M 6.061	NaCl 1.347 M		NaH2PO4 0.1714 M	
E9	87/1	crystal clusters	MES 0.085 M 6.796	NaCl 2 M		NaH2PO4 0.1143 M	KH2PO4 0.06939 M
H10	87/1	crystal clusters	MES 0.085 M 6.612	NaCl 1.412 M		NaH2PO4 0.1143 M	KH2PO4 0.1061 M
H11	87/1	crystal clusters		NaCl 1.494 M			KH2PO4 0.0449 M
H4	87/1	crystal clusters	MES 0.085 M 6.51	NaCl 1.739 M		NaH2PO4 0.06531 M	KH2PO4 0.1224 M
D12	87/2	crystal clusters			Glycerol 0.4082 %v/v		KH2PO4 0.0449 M
D8	87/2	crystal clusters	MES 0.085 M 6.571	NaCl 1.396 M		NaH2PO4 0.1673 M	KH2PO4 0.07347 M
E9	87/2	crystal clusters	MES 0.085 M 6.082	NaCl 1.314 M		NaH2PO4 0.09796 M	KH2PO4 0.1306 M
A8	87/1	microcrystals	MES 0.085 M 6.714	NaCl 1.918 M	Glycerol 16.73 %v/v	NaH2PO4 0.1633 M	KH2PO4 0.02041 M
D9	87/1	microcrystals	MES 0.085 M 6.673	NaCl 1.869 M	Glycerol 2.449 %v/v		KH2PO4 0.1714 M
E8	87/1	microcrystals	MES 0.085 M 6.388	NaCl 1.461 M		NaH2PO4 0.1102 M	KH2PO4 0.1388 M
H7	87/2	microcrystals	MES 0.085 M 6.388	NaCl 1.82 M	Glycerol 5.714 %v/v	NaH2PO4 0.1633 M	
D7	87/1	multiple crystals	MES 0.085 M 6.082	NaCl 1.249 M	Glycerol 3.265 %v/v		KH2PO4 0.1347 M
G12	87/1	multiple crystals	MES 0.085 M 6.49	NaCl 1.249 M		NaH2PO4 0.1959 M	
A4	87/1	spherulites	MES 0.085 M 6.816	NaCl 1.82 M	Glycerol 15.92 %v/v		KH2PO4 0.1755 M
C3	87/1	spherulites	MES 0.085 M 6.041	NaCl 1.918 M		NaH2PO4 0.1265 M	
D8	87/1	spherulites	MES 0.085 M 6.51	NaCl 1.984 M	Glycerol 12.65 %v/v	NaH2PO4 0.04082 M	KH2PO4 0.1102 M
F4	87/1	spherulites	MES 0.085 M 6.265	NaCl 1.869 M	Glycerol 12.65 %v/v	NaH2PO4 0.08163 M	KH2PO4 0.08163 M
G1	87/1	spherulites	MES 0.085 M 6.286	NaCl 1.739 M	Glycerol 17.55 %v/v	NaH2PO4 0.03673 M	KH2PO4 0.1061 M
G11	87/1	spherulites	MES 0.085 M 6.347	NaCl 1.641 M	Glycerol 13.47 %v/v	NaH2PO4 0.1837 M	
G8	87/1	spherulites	MES 0.085 M 6.388	NaCl 1.38 M	Glycerol 15.51 %v/v	NaH2PO4 0.1633 M	KH2PO4 0.06939 M
C9	87/2	spherulites	MES 0.085 M 6.694	NaCl 1.2 M	Glycerol 19.18 %v/v	NaH2PO4 0.08571 M	KH2PO4 0.1714 M
D5	87/2	spherulites	MES 0.085 M 6.653	NaCl 1.804 M	Glycerol 16.73 %v/v	NaH2PO4 0.1265 M	KH2PO4 0.06122 M
F1	87/2	spherulites	MES 0.085 M 6.041	NaCl 1.739 M	Glycerol 17.14 %v/v		KH2PO4 0.1061 M
F10	87/2	spherulites	MES 0.085 M 6.082	NaCl 1.673 M	Glycerol 0 %v/v	NaH2PO4 0.1878 M	
F6	87/2	spherulites	MES 0.085 M 6.51	NaCl 1.755 M		NaH2PO4 0.06531 M	KH2PO4 0.1878 M
F7	87/2	spherulites	MES 0.085 M 6.592	NaCl 1.837 M		NaH2PO4 0.151 M	KH2PO4 0.07347 M

In situ proteolysis crystallization conditions of SCOC (78-159)

plate_well	Protease	Native	precipitant	buffer	pH	salt	salt 2	precipitant 2
122_Cryo_D6	TR 1:1000	crystals	1.7 M Sodium chloride	0.085 M MES pH 6.5	6.5	0.085 M Sodium phosphate	0.085 M Potassium phosphate	15 %(v/v) Glycerol
122_Cryo_D7	TR 1:1000	crystals	3.655 M Sodium chloride	0.085 M HEPES pH 7.5	7.5			15 %(v/v) Glycerol
122_Cryo_D11	TR 1:1000	crystals	1.4 M Sodium formate	0.07 M Sodium acetate pH 4.6	4.6			30 %(v/v) Glycerol
122_Cryo_G11	TR 1:1000	crystals	25.5 %(w/v) PEG 2000 MME	0.085 M Sodium acetate pH 4.6	4.6	0.17 M Ammonium sulfate		15 %(v/v) Glycerol
123_Classics_A6	TR 1:1000	crystals	20 %(v/v) Isopropanol	0.1 M Sodium acetate pH 4.6	4.6	0.2 M Calcium chloride		
123_Classics_D6	TR 1:1000	crystals	2.0 M Sodium chloride	0.1 M MES pH 6.5	6.5	0.1 M Sodium phosphate	0.1 M Potassium phosphate	
123_Classics_H1	TR 1:1000	crystals	25 %(w/v) PEG 4000	0.1 M Sodium acetate pH 4.6	4.6	0.2 M Ammonium sulfate		
123_Classics_H2	TR 1:1000	crystals	30 %(w/v) PEG 4000	0.1 M Sodium acetate pH 4.6	4.6	0.2 M Ammonium acetate		
124_Salt_RX_D1	TR 1:1000	crystals	2.5 M Ammonium nitrate	0.1 M Sodium acetate	4.6			
124_Salt_RX_D4	TR 1:1000	crystals	6.0 M Ammonium nitrate	0.1 M Sodium acetate	4.6			
124_Salt_RX_D5	TR 1:1000	crystals	6.0 M Ammonium nitrate	BIS-TRIS propane	7.0			
124_Salt_RX_D7	TR 1:1000	crystals	1.5 M Sodium nitrate	0.1 M Sodium acetate	4.6			
124_Salt_RX_G7	TR 1:1000	crystals	0.7 Ammonium tartrate dibasic	0.1 M Sodium acetate	4.6			
124_Salt_RX_H5	TR 1:1000	crystals	0.5 M Potassium thiocyanate	0.1 M Sodium acetate	4.6			
125_Cryo_B10	SU 1:2000	crystals	59.5 %(v/v) MPD	0.085 M HEPES pH 7.5	7.5			15 %(v/v) Glycerol
125_Cryo_C6	SU 1:2000	crystals	1.5 M Ammonium sulfate	0.075 M TRIS.HCl pH 8.5	8.5			25 %(v/v) Glycerol
125_Cryo_D6	SU 1:2000	crystals	1.7 M Sodium chloride	0.085 M MES pH 6.5	6.5	0.085 M Sodium phosphate	0.085 M Potassium phosphate	15 %(v/v) Glycerol
125_Cryo_D11	SU 1:2000	crystals	1.4 M Sodium formate	0.07 M Sodium acetate pH 4.6	4.6			30 %(v/v) Glycerol
125_Cryo_D12	SU 1:2000	crystals	3.6 M Sodium formate					10 %(v/v) Glycerol
125_Cryo_E2	SU 1:2000	crystals	1.36 M Ammonium sulfate	0.085 M MES pH 6.5	6.5		8.5 %(v/v) Dioxane	15 %(v/v) Glycerol
125_Cryo_E5	SU 1:2000	crystals	1.275 M Ammonium sulfate	0.085 M TRIS pH 8.5	8.5		10.2 %(v/v) Glycerol	15 %(v/v) Glycerol
125_Cryo_F8	SU 1:2000	crystals	16 %(w/v) PEG 8000	0.08 M Sodium cacodylate pH 6.5	6.5	0.16 M Magnesium acetate		20 %(v/v) Glycerol
125_Cryo_F10	SU 1:2000	crystals	25.5 %(w/v) PEG 8000	0.085 M Sodium cacodylate pH 6.5	6.5	0.17 M Ammonium sulfate		15 %(v/v) Glycerol
125_Cryo_G12	SU 1:2000	crystals	5.6 %(w/v) PEG 4000	0.07 M Sodium acetate pH 4.6	4.6			30 %(v/v) Glycerol
126_pHClear_A2	TR 1:1000	crystals	1.0 M Sodium chloride	0.1 M Citric acid	5.0			
127_Classics_A6	SU 1:2000	crystals	20 %(v/v) Isopropanol	0.1 M Sodium acetate pH 4.6	4.6	0.2 M Calcium chloride		
127_Classics_D6	SU 1:2000	crystals	2.0 M Sodium chloride	0.1 M MES pH 6.5	6.5	0.1 M Sodium phosphate	0.1 M Potassium phosphate	
127_Classics_H1	SU 1:2000	crystals	25 %(w/v) PEG 4000	0.1 M Sodium acetate pH 4.6	4.6	0.2 M Ammonium sulfate		
128_JCSG_D12	SU 1:2000	crystals	16 %(w/v) PEG 8000				0.04 M Potassium phosphate	20 %(v/v) Glycerol
128_JCSG_H7	SU 1:2000	crystals	25 %(w/v) PEG 3350	0.1 M BIS-TRIS pH 5.5	5.5	0.2 M Ammonium sulfate		
129_ComPas_E2	SU 1:2000	crystals	10 %(w/v) Isopropanol	0.1 M Sodium acetate pH 4.6	4.6		0.2 M Calcium chloride	

129_ComPas_G9	SU 1:2000	crystals	1.0 M Sodium formate	0.1 M Sodium acetate pH 4.6	4.6		1.0 M Sodium formate	
130_Classics_A1	SU 1:2000	crystals	1.0 M 1,6-Hexanediol	0.1 M Sodium acetate pH 4.6	4.6	1.0 M 1,6-Hexanediol	0.01 M Cobalt chloride	
130_Classics_B3	SU 1:2000	crystals	3.0 M Sodium chloride	0.1 M MES	6.0			
130_Classics_B4	SU 1:2000	crystals	30 %(v/v) MPD	0.1 M Sodium acetate pH 4.6	4.6	30 %(v/v) MPD	0.2 M Sodium chloride	
128_JCSG_G12	SU 1:2000	crystals cluster	3 M Sodium chloride	0.1 M BIS-TRIS pH 5.5	5.5			
122_Cryo_H2	TR 1:1000	crystals clusters	25.5 %(w/v) PEG 4000	0.085 M Sodium acetate pH 4.6	4.6	0.17 M Ammonium acetate		15 %(v/v) Glycerol
122_Cryo_D5	TR 1:1000	microcrystals	1.7 M Sodium chloride	0.085 M Sodium acetate pH 4.6	4.6			15 %(v/v) Glycerol
124_Salt_RX_B1	TR 1:1000	microcrystals	3.2 M Sodium chloride	0,1 M BIS-TRIS propane	7.0			
124_Salt_RX_A11	TR 1:1000	microcrystals	2.2 M Sodium chloride	0,1 M Tris	8.5			
124_Salt_RX_F7	TR 1:1000	microcrystals	0.8 M Lithium sulfate	0,1 M Sodium acetate	4.6			
125_Cryo_C2	SU 1:2000	microcrystals	0.7 M Ammonium phosphate	0.07 M tri-Sodium citrate pH 5.6	5.6			30 %(v/v) Glycerol
125_Cryo_G1	SU 1:2000	microcrystals	1.7 %(v/v) PEG 400	0.085 M HEPES sodium salt pH 7.5	7.5	1.7 M Ammonium sulfate		15 %(v/v) Glycerol
126_pHClear_A8	TR 1:1000	microcrystals	2.0 M Sodium chloride	0.1 M Citric acid	5.0			
126_pHClear_B3	TR 1:1000	microcrystals	3.0 M Sodium chloride	0.1 M MES	6.0			
126_pHClear_B9	TR 1:1000	microcrystals	4.0 M Sodium chloride	0.1 M MES	6.0			
127_Classics_C1	SU 1:2000	microcrystals	0.4 M Ammonium phosphate					
127_Classics_C2	SU 1:2000	microcrystals	1.0 M Ammonium phosphate	0.1 M tri-Sodium citrate pH 5.6	5.6			
128_JCSG_C3	SU 1:2000	microcrystals	20 %(w/v) PEG 3350			0.2 M Ammonium nitrate		
128_JCSG_H9	SU 1:2000	microcrystals	25 %(w/v) PEG 3350	0.1 M BIS-TRIS pH 5.5	5.5	0.2 M Lithium sulfate		
128_JCSG_D1	SU 1:2000	microcrystals	24 %(w/v) PEG 1500					20% (w/v) Glycerol
130_Classics_C6	SU 1:2000	microcrystals	2.0 M Ammonium sulfate	0.1 M TRIS.HCl pH 8.5	8.5	0.1 M TRIS.HCl pH 8.5		
122_Cryo_E2	TR 1:1000	multiple crystals	1.36 M Ammonium sulfate	0.085 M MES pH 6.5	6.5		8.5 %(v/v) Dioxane	15 %(v/v) Glycerol
122_Cryo_H4	TR 1:1000	quasicrystalline	24 %(w/v) PEG 4000	0.08 M TRIS.HCl pH 8.5	8.5	0.16 M Magnesium chloride		20 %(v/v) Glycerol
122_Cryo_H5	TR 1:1000	quasicrystalline	25.5 %(w/v) PEG 4000	0.085 M TRIS.HCl pH 8.5	8.5	0.17 M Lithium sulfate		15 %(v/v) Glycerol
122_Cryo_H6	TR 1:1000	quasicrystalline	25.5 %(w/v) PEG 4000	0.085 M TRIS.HCl pH 8.5	8.5	0.17 M Sodium acetate		15 %(v/v) Glycerol
122_Cryo_H8	TR 1:1000	quasicrystalline	25.5 %(w/v) PEG 5000 MME	0.085 M MES pH 6.5	6.5	0.17 M Ammonium sulfate		15 %(v/v) Glycerol
122_Cryo_C8	TR 1:1000	spherulites	1.36 M Ammonium sulfate	0.085 M HEPES pH 7.5	7.5	0.085 M Sodium chloride		15 %(v/v) Glycerol
122_Cryo_E5	TR 1:1000	spherulites	1.275 M Ammonium sulfate	0.085 M TRIS pH 8.5	8.5		10.2 %(v/v) Glycerol	15 %(v/v) Glycerol
123_Classics_H10	TR 1:1000	spherulites	2.0 M Sodium chloride			10 %(w/v) PEG 6000		
124_Salt_RX_G10	TR 1:1000	spherulites	1.0 Ammonium tartrate dibasic	0.1 M Sodium acetate	4.6			
131_pHClear_B3	SU 1:2000	spherulites	3.0 M Sodium chloride	0.1 M MES	6.0			

A.3 Supplementary Figures

A.3.1 Elution Profiles

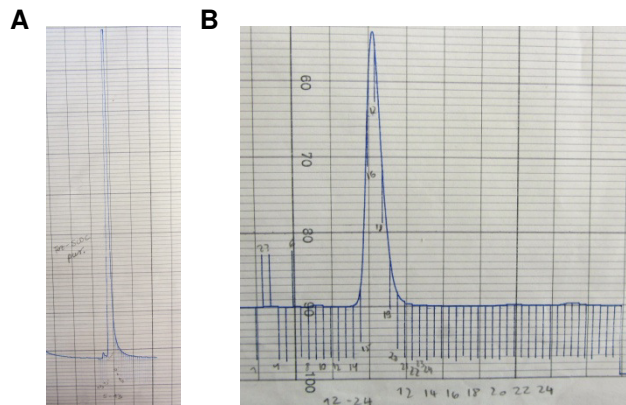


Figure A.1: Elution profiles of SCOC-FEZ1 copurification

(A) Elution profile of StrepTrap chromatography (B) Elution profile of size exclusion chromatography

A.3.2 Purification of SCOC ccd-FEZ1 ccd complexes

(see next page)

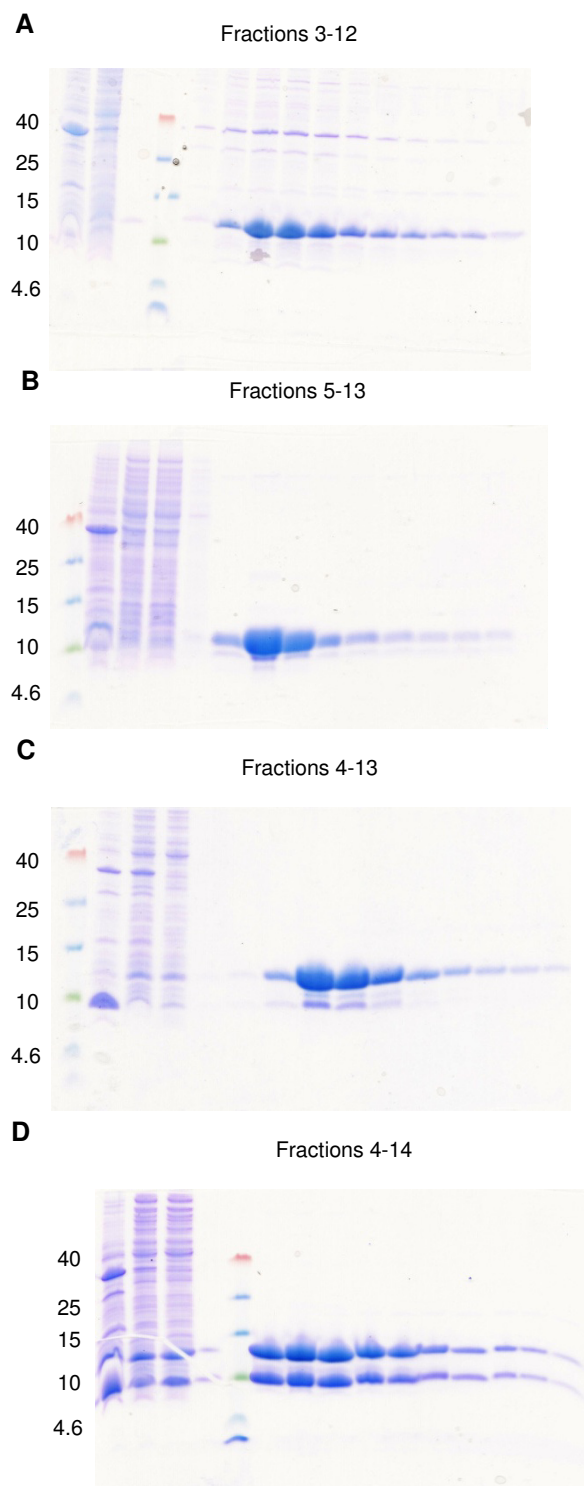


Figure A.2: Schägger gels of copurifications of mutant SCOC ccd-FEZ1 ccd

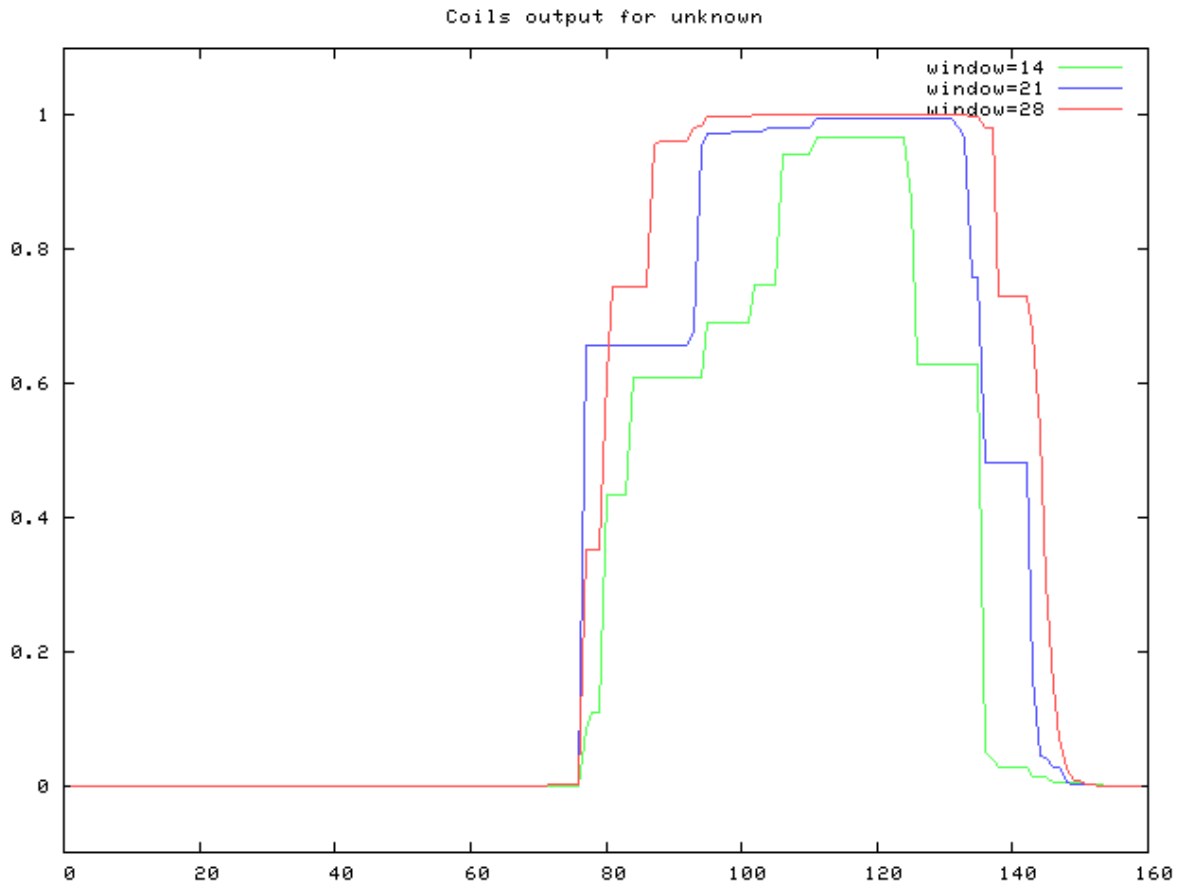
(A) Schägger gel of SCOC N125V/N132L with FEZ1 (B) Schägger gel of SCOC E93L/K97V with FEZ1 (C) Schägger gel of SCOC R117E with FEZ1 (D) Schägger gel of SCOC R99E with FEZ1

A.4 Computational Methods

A.4.1 COILS results

(see next page)

COILS of SCOC Isoform1:



```
# NCOILS version 1.0
# using MTIDK matrix
# no weights
# Input file is ../wwtmp/.COILS.25785.6837.seq
#
```

1	M	d	0.000	d	0.000	d	0.000
2	R	e	0.000	e	0.000	e	0.000
3	R	f	0.000	f	0.000	f	0.000
4	R	g	0.000	g	0.000	g	0.000
5	V	a	0.000	a	0.000	a	0.000
6	F	b	0.000	b	0.000	b	0.000
7	S	c	0.000	c	0.000	c	0.000
8	S	d	0.000	d	0.000	d	0.000
9	Q	e	0.000	e	0.000	e	0.000
10	D	f	0.000	f	0.000	f	0.000
11	W	g	0.000	g	0.000	g	0.000
12	R	a	0.000	a	0.000	a	0.000
13	A	b	0.000	b	0.000	b	0.000
14	S	c	0.000	c	0.000	c	0.000
15	G	c	0.000	c	0.000	c	0.000
16	W	d	0.000	d	0.000	d	0.000
17	D	e	0.000	e	0.000	e	0.000
18	G	f	0.000	f	0.000	f	0.000
19	M	g	0.000	g	0.000	g	0.000
20	G	a	0.000	a	0.000	a	0.000
21	F	b	0.000	b	0.000	b	0.000
22	F	c	0.000	c	0.000	c	0.000
23	S	f	0.000	f	0.000	f	0.000
24	R	g	0.000	g	0.000	g	0.000
25	R	a	0.000	a	0.000	a	0.000
26	T	b	0.000	b	0.000	b	0.000
27	F	c	0.000	c	0.000	c	0.000
28	C	d	0.000	d	0.000	d	0.000
29	G	e	0.000	e	0.000	e	0.000
30	R	f	0.000	f	0.000	f	0.000
31	S	g	0.000	g	0.000	g	0.000

32	G	a	0.000	a	0.000	a	0.000
33	R	b	0.000	b	0.000	b	0.000
34	S	c	0.000	c	0.000	c	0.000
35	C	d	0.000	d	0.000	d	0.000
36	R	e	0.000	e	0.000	e	0.000
37	G	f	0.000	f	0.000	f	0.000
38	Q	g	0.000	g	0.000	g	0.000
39	L	a	0.000	a	0.000	a	0.000
40	V	b	0.000	b	0.000	b	0.000
41	Q	c	0.000	c	0.000	c	0.000
42	V	d	0.000	d	0.000	d	0.000
43	S	e	0.000	e	0.000	e	0.000
44	R	f	0.000	f	0.000	f	0.000
45	P	b	0.000	b	0.000	b	0.000
46	E	c	0.000	c	0.000	c	0.000
47	V	d	0.000	d	0.000	d	0.000
48	S	e	0.000	e	0.000	e	0.000
49	A	f	0.000	f	0.000	f	0.000
50	G	b	0.000	b	0.000	b	0.000
51	S	c	0.000	c	0.000	c	0.000
52	L	d	0.000	d	0.000	d	0.000
53	L	e	0.000	e	0.000	e	0.000
54	L	f	0.000	f	0.000	f	0.000
55	P	g	0.000	g	0.000	g	0.000
56	A	a	0.000	a	0.000	a	0.000
57	P	b	0.000	b	0.000	b	0.000
58	Q	c	0.000	c	0.000	c	0.000
59	A	d	0.000	d	0.000	d	0.000
60	E	e	0.000	e	0.000	e	0.000
61	D	f	0.000	f	0.000	f	0.000
62	H	g	0.000	g	0.000	g	0.000
63	S	a	0.000	a	0.000	a	0.000
64	S	b	0.000	b	0.000	b	0.000
65	R	c	0.000	c	0.000	c	0.000
66	I	d	0.000	d	0.000	d	0.000
67	L	e	0.000	e	0.000	e	0.000
68	Y	a	0.000	a	0.000	a	0.000
69	P	b	0.000	b	0.000	b	0.000
70	R	c	0.000	c	0.000	c	0.000
71	P	d	0.000	d	0.000	d	0.000
72	K	e	0.000	e	0.002	e	0.003
73	S	f	0.000	f	0.003	f	0.003
74	L	g	0.000	g	0.003	g	0.003
75	L	a	0.000	a	0.003	a	0.003
76	P	f	0.000	f	0.003	f	0.003
77	K	g	0.084	g	0.657	g	0.351
78	M	a	0.110	a	0.657	a	0.351
79	M	b	0.110	b	0.657	b	0.351
80	N	c	0.434	c	0.657	c	0.612
81	A	d	0.434	d	0.657	d	0.745
82	D	e	0.434	e	0.657	e	0.745
83	M	f	0.434	f	0.657	f	0.745
84	D	g	0.610	g	0.657	g	0.745
85	A	a	0.610	a	0.657	a	0.745
86	V	b	0.610	b	0.657	b	0.745
87	D	c	0.610	c	0.657	c	0.956
88	A	d	0.610	d	0.657	d	0.961
89	E	e	0.610	e	0.657	e	0.961
90	N	f	0.610	f	0.657	f	0.961
91	Q	g	0.610	g	0.657	g	0.961
92	V	a	0.610	a	0.657	a	0.961
93	E	b	0.610	b	0.680	b	0.982
94	L	c	0.610	c	0.953	c	0.985
95	E	f	0.690	f	0.973	f	0.998
96	E	g	0.690	g	0.973	g	0.998
97	K	a	0.690	a	0.973	a	0.998
98	T	b	0.690	b	0.973	b	0.998
99	R	c	0.690	c	0.975	c	0.998
100	L	d	0.690	d	0.975	d	0.998
101	I	e	0.690	e	0.975	e	0.998
102	N	f	0.748	f	0.975	f	1.000
103	Q	g	0.748	g	0.975	g	1.000
104	V	a	0.748	a	0.982	a	1.000
105	L	b	0.748	b	0.982	b	1.000
106	E	c	0.943	c	0.982	c	1.000
107	L	d	0.943	d	0.982	d	1.000
108	Q	e	0.943	e	0.982	e	1.000
109	H	f	0.943	f	0.982	f	1.000

110	T	g	0.943	g	0.982	g	1.000
111	L	a	0.968	a	0.996	a	1.000
112	E	b	0.968	b	0.996	b	1.000
113	D	c	0.968	c	0.996	c	1.000
114	L	d	0.968	d	0.996	d	1.000
115	S	e	0.968	e	0.996	e	1.000
116	A	f	0.968	f	0.996	f	1.000
117	R	g	0.968	g	0.996	g	1.000
118	V	a	0.968	a	0.996	a	1.000
119	D	b	0.968	b	0.996	b	1.000
120	A	c	0.968	c	0.996	c	1.000
121	V	d	0.968	d	0.996	d	1.000
122	K	e	0.968	e	0.996	e	1.000
123	E	f	0.968	f	0.996	f	1.000
124	E	g	0.968	g	0.996	g	1.000
125	N	a	0.871	a	0.996	a	1.000
126	L	b	0.628	b	0.996	b	1.000
127	K	c	0.628	c	0.996	c	1.000
128	L	d	0.628	d	0.996	d	1.000
129	K	e	0.628	e	0.996	e	1.000
130	S	f	0.628	f	0.996	f	1.000
131	E	g	0.628	g	0.996	g	1.000
132	N	a	0.628	a	0.982	a	1.000
133	Q	b	0.628	b	0.967	b	1.000
134	V	c	0.628	c	0.758	c	0.998
135	L	d	0.628	d	0.758	d	0.998
136	G	e	0.050	e	0.482	e	0.982
137	Q	f	0.038	f	0.482	f	0.982
138	Y	g	0.027	g	0.482	g	0.731
139	I	a	0.027	a	0.482	a	0.731
140	E	b	0.027	b	0.482	b	0.731
141	N	c	0.027	c	0.482	c	0.731
142	L	d	0.027	d	0.482	d	0.731
143	M	b	0.012	b	0.163	b	0.677
144	S	c	0.012	c	0.044	c	0.529
145	A	d	0.012	d	0.042	d	0.303
146	S	e	0.005	e	0.028	e	0.148
147	S	f	0.004	f	0.028	f	0.069
148	V	g	0.004	g	0.005	g	0.023
149	F	a	0.004	a	0.001	a	0.007
150	Q	b	0.004	b	0.001	b	0.007
151	T	f	0.001	f	0.001	f	0.001
152	T	g	0.001	g	0.001	g	0.001
153	D	e	0.001	e	0.000	e	0.000
154	T	f	0.000	f	0.000	f	0.000
155	K	g	0.000	g	0.000	g	0.000
156	S	a	0.000	a	0.000	a	0.000
157	K	b	0.000	b	0.000	b	0.000
158	R	c	0.000	c	0.000	c	0.000
159	K	d	0.000	d	0.000	d	0.000

A.4.2 HHpred results

(see next page)

HHPred results

Query	SCOCccd (78-159)	Prob	E-value	P-value	Score	SS	Coils	Query	HMM	Template	HMM
No Hit											
3tso_C	RAB11 family-interactin	97.3	0.00074	2.5E-08	43.5	6.0	62	6-81	2-68	(75)	
1ci6_A	Transcription factor AT	95.6	0.083	2.8E-06	31.0	7.1	46	20-65	18-63	(63)	
2gzd_C	RAB11 family-interactin	95.3	0.012	3.9E-07	40.0	2.8	47	35-81	49-100	(107)	
1hjb_A	Ccaat/enhancer binding	94.0	0.29	9.7E-06	30.4	6.9	49	20-68	31-79	(87)	
1gu4_A	CAAT/enhancer binding p	94.0	0.13	4.4E-06	31.4	5.2	46	20-65	31-76	(78)	
1gd2_E	Transcription factor PA	93.9	0.12	3.9E-06	30.6	4.7	39	21-59	32-70	(70)	
3s9g_A	Protein hexim1; cyclin	93.5	0.26	8.5E-06	32.9	6.3	40	19-58	31-84	(104)	
1t2k_D	Cyclic-AMP-dependent tr	93.5	0.62	2.1E-05	27.0	7.3	43	20-62	17-59	(61)	
1jnm_A	Proto-oncogene C-JUN; B	93.0	0.25	8.3E-06	28.5	5.0	43	20-62	17-59	(62)	
2wt7_B	Transcription factor MA	92.9	0.67	2.2E-05	29.0	7.2	49	17-65	40-88	(90)	
3mq7_A	Bone marrow stromal ant	92.1	0.65	2.2E-05	32.0	6.9	57	19-75	65-121	(121)	
3q8t_A	Beclin-1; autophagy, AT	91.3	1.4	4.6E-05	26.0	7.4	65	1-65	1-65	(96)	
2wt7_A	Proto-oncogene protein	90.5	1.6	5.4E-05	25.3	7.8	46	19-64	17-62	(63)	
1a92_A	Delta antigen; leucine	90.4	1.6	5.2E-05	26.0	6.4	46	20-65	2-47	(50)	
2fxo_A	Myosin heavy chain, car	88.5	2.8	9.3E-05	25.2	8.3	49	17-65	75-123	(129)	
1mlj_A	Fibrinogen alpha subuni	86.8	5.3	0.00018	28.9	8.6	52	17-68	110-161	(491)	
3ghg_A	Fibrinogen alpha chain;	86.3	4	0.00013	30.6	7.9	48	21-68	113-160	(562)	
2xdj_A	Uncharacterized protein	85.9	5	0.00017	25.4	7.7	45	21-65	9-53	(83)	
1svf_A	Protein (fusion glycopr	85.5	4.4	0.00015	25.1	6.5	45	25-69	14-58	(64)	
2jee_A	YIU; FTSZ, septum, coi	84.2	4.8	0.00016	26.0	6.5	38	23-60	25-62	(81)	
3ghg_A	Fibrinogen alpha chain;	83.2	3.6	0.00012	30.9	6.5	41	26-66	50-91	(562)	
2dgc_A	Protein (GCN4); basic d	83.1	2.4	8.1E-05	24.8	4.5	33	31-63	29-61	(63)	
1gd2_E	Transcription factor PA	82.8	5.2	0.00017	23.2	6.2	46	20-65	24-69	(70)	
3ghg_B	Fibrinogen beta chain;	82.2	5.8	0.00019	31.0	7.5	46	20-65	83-128	(461)	
3okq_A	BUD site selection prot	81.5	10	0.00033	25.5	7.7	45	21-65	24-80	(141)	
3q4f_C	DNA repair protein XRCC	81.0	4	0.00013	29.5	5.8	34	27-60	149-182	(186)	
2fyz_A	Fusion protein, fusion	80.9	6	0.0002	24.5	5.8	42	25-66	17-58	(63)	
1dh3_A	Transcription factor CR	80.6	3.7	0.00012	23.3	4.6	30	31-60	21-50	(55)	
1gu4_A	CAAT/enhancer binding p	80.1	7	0.00023	23.5	6.0	35	31-65	35-69	(78)	
1mlj_B	Fibrinogen beta chain;	80.0	8.6	0.00029	30.1	7.9	46	20-65	88-133	(464)	
3s9g_A	Protein hexim1; cyclin	80.0	11	0.00036	24.9	7.3	43	18-60	51-93	(104)	
3ghg_C	Fibrinogen gamma chain;	79.1	12	0.00041	28.5	8.3	49	17-65	83-131	(411)	
3hnw_A	Uncharacterized protein	78.7	12	0.00039	24.6	8.3	49	17-65	67-115	(138)	
3mud_A	DNA repair protein XRCC	78.5	15	0.00048	26.0	8.0	47	21-67	123-170	(175)	
3q8t_A	Beclin-1; autophagy, AT	78.0	8.2	0.00027	22.5	8.1	48	18-65	4-51	(96)	
4gif_A	Polycystic kidney disea	76.7	9	0.0003	22.3	5.7	39	15-53	4-42	(45)	
1deq_A	Fibrinogen (alpha chain	76.6	14	0.00047	26.4	7.6	56	11-66	77-133	(390)	
2vfv_A	Alanyl-tRNA synthetase;	75.9	4.2	0.00014	26.5	4.3	61	9-69	1-62	(171)	
2v4h_A	NF-kappa-B essential mo	75.6	12	0.0004	23.2	8.1	54	14-67	51-104	(110)	
3w03_C	DNA repair protein XRCC	75.2	7.5	0.00025	28.0	5.8	34	27-60	140-173	(184)	
3oa7_A	Head morphogenesis prot	74.7	5.7	0.00019	28.8	5.1	50	18-67	9-58	(206)	
2hv8_D	RAB11 family-interactin	74.3	5.4	0.00018	24.7	4.2	25	42-66	26-50	(64)	
1nkp_A	C-MYC, MYC proto-oncoge	74.3	12	0.0004	22.6	5.9	39	21-59	48-86	(88)	
1deq_A	Fibrinogen (alpha chain	74.0	11	0.00036	27.1	6.4	39	29-67	57-95	(390)	
1go4_E	MAD1 (mitotic arrest de	73.6	11	0.00036	24.8	5.8	34	31-64	11-44	(100)	

21d3_A Myosin VI; molecular mo	73.0	6	0.0002	25.3	4.4	49	20-68	30-78	(88)
1nlw_A MAD protein, MAX dimeri	72.4	9.7	0.00032	22.6	5.0	35	24-58	46-80	(80)
3oa7_A Head morphogenesis prot	71.9	22	0.00072	25.7	7.6	52	19-70	24-75	(206)
2v66_B Nuclear distribution pr	71.6	13	0.00043	21.8	8.3	49	17-65	55-103	(111)
2wt7_B Transcription factor MA	71.2	9.1	0.0003	23.7	4.8	39	21-59	51-89	(90)
2v66_B Nuclear distribution pr	71.2	13	0.00044	21.7	8.3	49	17-65	34-82	(111)
1jnm_A Proto-oncogene C-JUN; B	71.0	11	0.00038	21.2	4.9	35	31-65	21-55	(62)
2v4h_A NF-kappa-B essential mo	68.3	19	0.00062	22.3	7.2	47	14-60	58-104	(110)
3u59_A Tropomyosin beta chain;	67.6	16	0.00052	21.2	7.8	63	3-65	1-63	(101)
2p90_A Hypothetical protein CG	67.3	3.7	0.00012	30.5	2.7	44	22-65	219-262	(319)
1gk7_A Vimentin; intermediate	66.5	15	0.00049	20.6	4.5	39	16-58	1-39	(39)
3hnw_A Uncharacterized protein	65.7	26	0.00085	22.9	8.3	49	17-65	81-129	(138)
1mlj_C Fibrinogen gamma chain;	65.2	33	0.0011	26.2	7.8	49	17-65	83-131	(409)
2oqg_A Transcription factor HY	64.9	18	0.00059	20.8	5.4	36	29-64	7-42	(42)
1ses_A Seryl-tRNA synthetase;	64.8	46	0.0015	25.5	9.1	63	18-80	35-113	(421)
1nkp_B MAX protein, MYC proto-	64.6	11	0.00038	22.4	4.1	36	24-59	46-81	(83)
3u0c_A Invasin IPAB, 62 kDa an	64.4	26	0.00085	22.5	7.4	56	10-65	144-199	(201)
3ghg_C Fibrinogen gamma chain;	63.4	18	0.00061	27.6	6.0	46	17-62	90-135	(411)
3caz_A BAR protein; thermo-aci	63.3	20	0.00068	22.6	5.4	60	6-65	6-65	(294)
3vrp_A Signal transduction pro	62.7	7.4	0.00025	30.6	3.7	52	30-81	52-115	(331)
2dq0_A Seryl-tRNA synthetase;	62.5	51	0.0017	25.6	8.5	63	18-80	38-118	(455)
1x79_B RAB GTPase binding effe	62.5	12	0.0004	22.0	4.0	43	17-59	62-104	(112)
3he5_B Synzip2; heterodimeric	62.5	22	0.00072	20.9	4.9	29	37-65	8-36	(52)
1avy_A Fibrin; gpwac M; bact	62.2	17	0.00058	23.0	4.8	38	23-60	6-43	(74)
3lpx_A Spectrin alpha chain, e	61.8	27	0.00089	21.8	5.7	56	12-67	82-141	(161)
1hjb_A Ccaat/enhancer binding	61.5	25	0.00084	21.4	6.0	35	31-65	35-69	(87)
3bux_B E3 ubiquitin-protein li	61.3	18	0.00059	28.3	5.6	40	30-70	62-105	(329)
1mlj_B Fibrinogen beta chain;	61.1	51	0.0017	25.7	8.3	49	17-65	99-147	(464)
2yy0_A C-MYC-binding protein;	60.7	22	0.00074	21.2	4.8	28	26-53	20-47	(53)
4e17_B Catenin alpha-1; four h	60.7	16	0.00052	20.8	3.9	23	18-40	12-34	(40)
1wt6_A Myotonin-protein kinase	60.5	31	0.001	22.1	8.3	52	14-65	20-71	(81)
3bas_A Myosin heavy chain, str	60.0	25	0.00083	20.8	7.4	47	17-63	41-87	(89)
2fxo_A Myosin heavy chain, car	59.6	25	0.00084	20.8	8.3	49	17-65	40-88	(129)
1clg_A Tropomyosin; contractil	59.0	31	0.001	21.6	7.6	60	6-65	1-60	(284)
3nud_A DNA repair protein XRCC	58.9	26	0.00088	24.6	5.8	34	27-60	123-156	(175)
2b9c_A Striated-muscle alpha t	58.9	24	0.00081	20.4	7.7	45	21-65	76-120	(147)
2zgm_A Prefoldin beta subunit	58.8	28	0.00095	21.1	7.0	46	25-70	70-115	(117)
2akf_A Coronin-1A; coiled coil	58.0	21	0.00069	19.3	4.3	25	22-46	3-27	(32)
4gkw_A Spindle assembly abnorm	57.3	28	0.00095	20.7	7.5	61	5-65	19-79	(167)
3q0x_A Centriole protein; cent	57.3	51	0.0017	23.6	8.3	52	14-65	160-211	(228)
1lk9_A DNA repair protein XRCC	57.0	48	0.0016	23.1	7.1	42	26-67	119-160	(213)
3ghg_B Fibrinogen beta chain;	56.8	54	0.0018	25.5	7.8	49	17-65	87-135	(461)
2xdj_A Uncharacterized protein	56.7	34	0.0011	21.4	6.2	49	22-70	24-72	(83)
2xzz_A Immunoglobulin-binding	56.6	29	0.00096	20.5	6.0	41	27-67	1-41	(114)
3o0z_A RHO-associated protein	55.8	31	0.001	20.6	8.3	49	17-65	82-130	(168)
4e61_A Protein BIM1; EBI-like	55.2	41	0.0014	21.9	6.1	36	22-57	8-43	(106)
4gkw_A Spindle assembly abnorm	54.8	32	0.0011	20.5	7.7	45	21-65	115-159	(167)
2efr_A General control protein	54.5	31	0.001	20.2	8.6	51	17-67	13-63	(155)
1dip_A Delta-sleep-inducing pe	53.8	22	0.00073	22.7	4.2	28	35-62	18-45	(78)

A.4.3 XDS input file

```

! Settings for PILATUS6M @ SLS generated by GO.COM
! job control
JOB= ALL
!JOB=XYCORR INIT COLSPOT IDXREF DEFPIX INTEGRATE CORRECT
MAXIMUM_NUMBER_OF_JOBS=4
MAXIMUM_NUMBER_OF_PROCESSORS=12
!MINUTE=0 !how long to wait (minutes), implies only one JOBS

! for this experiment:
ORGX= 1252
ORGY= 1238
DETECTOR_DISTANCE= 520.19
OSCILLATION_RANGE= 0.5
X-RAY_WAVELENGTH= 0.97936
NAME_TEMPLATE_OF_DATA_FRAMES=../../dts6_8_?????.cbf CBF
DATA_RANGE= 1 360
SPOT_RANGE= 1 29
BACKGROUND_RANGE= 1 9

SPACE_GROUP_NUMBER=20
UNIT_CELL_CONSTANTS=70.99 114.8 93.27 90 90 90
!REIDX= 0 0 -1 0 0 -1 0 0 -1 0 0 0

REFINE(IDXREF)=BEAM AXIS ORIENTATION CELL !DISTANCE
REFINE(INTEGRATE)=DISTANCE BEAM ORIENTATION CELL !AXIS
REFINE(CORRECT)=DISTANCE BEAM ORIENTATION CELL AXIS

FRIEDEL'S_LAW=FALSE
STRICT_ABSORPTION_CORRECTION=FALSE ! but read Tips_and_Tricks in XDSwiki

! parameters with changes wrt default values:
TRUSTED_REGION=0.00 1.15
VALUE_RANGE_FOR_TRUSTED_DETECTOR_PIXELS=8000. 30000.
!MINIMUM_ZETA=0.05
CORRECTIONS=DECAY MODULATION ABSORP !default value
INCLUDE_RESOLUTION_RANGE=50 2.7
STRONG_PIXEL=6

! parameters specifically for this detector and beamline:
DETECTOR=PILATUS NX=2463 NY=2527 QX=0.172 QY=0.172 !PILATUS 6M
MINIMUM_VALID_PIXEL_VALUE=0 OVERLOAD=1048500
DIRECTION_OF_DETECTOR_X-AXIS=1 0 0
DIRECTION_OF_DETECTOR_Y-AXIS=0 1 0
INCIDENT_BEAM_DIRECTION=0 0 1
ROTATION_AXIS=1 0 0
FRACTION_OF_POLARIZATION=0.99
POLARIZATION_PLANE_NORMAL=0 1 0
SENSOR_THICKNESS=0.32
!UNTRUSTED_RECTANGLE= 487 495 1 2527
!UNTRUSTED_RECTANGLE= 981 989 1 2527
!UNTRUSTED_RECTANGLE=1475 1483 1 2527
!UNTRUSTED_RECTANGLE=1969 1977 1 2527
!UNTRUSTED_RECTANGLE= 1 2463 195 213
!UNTRUSTED_RECTANGLE= 1 2463 407 425
!UNTRUSTED_RECTANGLE= 1 2463 619 637
!UNTRUSTED_RECTANGLE= 1 2463 831 849
!UNTRUSTED_RECTANGLE= 1 2463 1043 1061
!UNTRUSTED_RECTANGLE= 1 2463 1255 1273
!UNTRUSTED_RECTANGLE= 1 2463 1467 1485
!UNTRUSTED_RECTANGLE= 1 2463 1679 1697
!UNTRUSTED_RECTANGLE= 1 2463 1891 1909
!UNTRUSTED_RECTANGLE= 1 2463 2103 2121
!UNTRUSTED_RECTANGLE= 1 2463 2315 2333
NUMBER_OF_PROFILE_GRID_POINTS_ALONG_ALPHA/BETA=13 !used by: INTEGRATE

```

A.4.4 Systematic absences by CORRECT.LP

(see next page)

```

*****
STATISTICS OF SAVED DATA SET "XDS_ASCII.HKL" (DATA_RANGE= 1 360)
FILE TYPE: XDS_ASCII MERGE=FALSE FRIEDEL'S_LAW=FALSE
*****
REFLECTIONS OF TYPE H,0,0 0,K,0 0,0,L OR EXPECTED TO BE ABSENT (*)
-----
H K L RESOLUTION INTENSITY SIGMA INTENSITY/SIGMA #OBSERVED
0 0 3 31.091 0.1063E+01 0.7320E+00 1.45 2*
0 0 4 23.318 0.9329E+03 0.2888E+02 32.30 2
0 0 5 18.654 -0.9618E+00 0.1158E+01 -0.83 2*
0 0 6 15.545 0.1544E+04 0.4774E+02 32.34 2
0 0 7 13.325 -0.2947E+01 0.1584E+01 -1.86 2*
0 0 8 11.659 0.2685E+04 0.8286E+02 32.40 2
0 0 9 10.364 0.1320E+01 0.2049E+01 0.64 2*
0 0 10 9.327 0.4313E+03 0.1437E+02 30.01 2
0 0 11 8.479 0.2488E+01 0.2420E+01 1.03 2*
0 0 12 7.773 0.3143E+03 0.1115E+02 28.20 2
0 0 13 7.175 0.1370E+01 0.4084E+01 0.34 1*
0 0 14 6.662 -0.1066E+01 0.3203E+01 -0.33 2
0 0 15 6.218 0.2823E+00 0.3355E+01 0.08 2*
0 0 16 5.829 0.5158E+03 0.2482E+02 20.78 1
0 0 17 5.487 0.5630E+01 0.3666E+01 1.54 2*
0 0 18 5.182 0.2051E+04 0.9115E+02 22.50 1
0 0 19 4.909 0.5728E+01 0.4087E+01 1.40 2*
0 0 20 4.664 0.1208E+04 0.3895E+02 31.02 2
0 0 21 4.442 0.4415E+01 0.4595E+01 0.96 2*
0 0 22 4.240 0.9103E+02 0.6727E+01 13.53 2
0 0 23 4.055 0.2575E+01 0.5479E+01 0.47 2*
0 0 24 3.886 0.9100E+03 0.4330E+02 21.02 1
0 0 25 3.731 -0.8265E+01 0.5884E+01 -1.40 2*
0 0 26 3.587 0.4981E+03 0.2853E+02 17.46 1
0 0 27 3.455 -0.3966E+01 0.7675E+01 -0.52 2*
0 0 28 3.331 0.7266E+01 0.7965E+01 0.91 2
0 0 29 3.216 -0.3841E+01 0.7302E+01 -0.53 2*
0 0 30 3.109 0.1818E+01 0.7215E+01 0.25 2
0 0 31 3.009 0.6106E+01 0.7896E+01 0.77 2*
0 0 32 2.915 0.8847E+02 0.1167E+02 7.58 1
0 0 33 2.826 -0.6238E+01 0.7833E+01 -0.80 2*
0 0 34 2.743 -0.2290E+02 0.1318E+02 -1.74 1
0 4 0 28.688 0.3090E+04 0.9461E+02 32.66 2
0 6 0 19.125 0.1290E+04 0.3979E+02 32.41 2
0 8 0 14.344 0.9439E+03 0.2939E+02 32.11 2
0 10 0 11.475 0.3734E+03 0.1712E+02 21.81 1
0 12 0 9.563 0.5466E+03 0.2485E+02 22.00 1
0 14 0 8.197 0.7812E+02 0.3643E+01 21.45 2
0 16 0 7.172 0.4355E+02 0.2993E+01 14.55 2

```

0	18	0	6.375	0.7077E+02	0.3815E+01	18.55	2
0	20	0	5.738	0.3751E+01	0.3077E+01	1.22	1
0	24	0	4.781	0.4514E+02	0.5097E+01	8.86	2
0	28	0	4.098	0.2303E+03	0.1080E+02	21.32	2
0	30	0	3.825	0.8876E+03	0.3087E+02	28.75	2
0	32	0	3.586	0.1593E+03	0.9846E+01	16.18	2
0	34	0	3.375	0.1155E+03	0.9101E+01	12.69	2
0	36	0	3.188	0.8166E+02	0.8434E+01	9.68	2
0	38	0	3.020	0.2137E+02	0.7004E+01	3.05	2
0	40	0	2.869	0.7636E+01	0.7082E+01	1.08	2
0	42	0	2.732	0.1131E+02	0.7408E+01	1.53	2
2	0	0	35.492	0.1314E+03	0.4316E+01	30.44	2
4	0	0	17.746	0.2984E+04	0.9153E+02	32.60	2
6	0	0	11.831	0.3087E+04	0.1343E+03	22.99	1
8	0	0	8.873	0.3279E+05	0.1002E+04	32.72	2
10	0	0	7.098	0.1233E+03	0.6181E+01	19.95	2
12	0	0	5.915	0.1314E+03	0.7656E+01	17.16	2
14	0	0	5.070	0.9808E+02	0.7777E+01	12.61	2
16	0	0	4.436	0.1451E+04	0.4758E+02	30.49	2
18	0	0	3.944	0.1169E+02	0.9609E+01	1.22	2
20	0	0	3.549	-0.6686E+01	0.1242E+02	-0.54	2
22	0	0	3.227	0.4901E+02	0.1849E+02	2.65	1
24	0	0	2.958	0.9941E+02	0.1277E+02	7.78	2
26	0	0	2.730	0.2548E+02	0.1413E+02	1.80	1

AVERAGE INTENSITY FOR 31 REFLECTIONS WHICH SHOULD
BE SYSTEMATICALLY ABSENT IS 0.1% OF MEAN INTENSITY

COMPLETENESS AND QUALITY OF DATA SET

R-FACTOR
observed = (SUM(ABS(I(h,i)-I(h)))/(SUM(I(h,i))))
expected = expected R-FACTOR derived from Sigma(I)

COMPARED = number of reflections used for calculating R-FACTOR
I/SIGMA = mean of intensity/Sigma(I) of unique reflections
(after merging symmetry-related observations)
Sigma(I) = standard deviation of reflection intensity I
estimated from sample statistics

R-meas = redundancy independent R-factor (intensities)
Diederichs & Karplus (1997), Nature Struct. Biol. 4, 269-275.

CC(1/2) = percentage of correlation between intensities from
random half-datasets. Correlation significant at
the 0.1% level is marked by an asterisk.
Karplus & Diederichs (2012), Science 336, 1030-33

Anomal = percentage of correlation between random half-sets
 Corr of anomalous intensity differences. Correlation
 significant at the 0.1% level is marked.
 SigAno = mean anomalous difference in units of its estimated
 standard deviation ($|F(+)-F(-)|/\text{Sigma}$). F(+), F(-)
 are structure factor estimates obtained from the
 merged intensity observations in each parity class.
 Nano = Number of unique reflections used to calculate
 Anomal_Corr & SigAno. At least two observations
 for each (+ and -) parity are required.

NOTE: Friedel pairs are treated as different reflections.

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE ≥ -3.0 AS FUNCTION OF RESOLUTION													
RESOLUTION	NUMBER OF REFLECTIONS	COMPLETENESS	R-FACTOR	R-FACTOR COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomal	SigAno	Nano			
LIMIT	OBSERVED	UNIQUE POSSIBLE	OF DATA	observed	expected			Corr					
8.01	2582	764	779	98.1%	2.7%	3.0%	2576	37.44	3.3%	99.9*	77*	2.333	302
5.70	4645	1369	1377	99.4%	3.9%	3.9%	4635	24.54	4.6%	99.8*	71*	2.011	598
4.67	6184	1757	1769	99.3%	4.0%	3.9%	6167	25.58	4.7%	99.8*	51*	1.475	785
4.05	7137	2078	2097	99.1%	4.0%	4.0%	7103	25.74	4.7%	99.7*	42*	1.233	923
3.62	8293	2370	2379	99.6%	5.4%	5.5%	8275	19.65	6.5%	99.5*	20*	0.994	1088
3.31	8870	2614	2624	99.6%	8.7%	9.2%	8821	12.47	10.3%	99.0*	9	0.861	1177
3.06	10159	2830	2838	99.7%	15.8%	17.2%	10123	7.71	18.6%	97.2*	2	0.778	1305
2.87	10666	3065	3072	99.8%	23.9%	25.8%	10607	5.30	28.3%	93.5*	-2	0.740	1395
2.70	10768	3200	3251	98.4%	45.9%	48.9%	10652	2.78	54.7%	79.8*	-2	0.728	1422
total	69304	20047	20186	99.3%	5.7%	6.0%	68959	14.30	6.8%	99.9*	28*	1.043	8995

NUMBER OF REFLECTIONS IN SELECTED SUBSET OF IMAGES	70473
NUMBER OF REJECTED MISFITS	1138
NUMBER OF SYSTEMATIC ABSENT REFLECTIONS	31
NUMBER OF ACCEPTED OBSERVATIONS	69304
NUMBER OF UNIQUE ACCEPTED REFLECTIONS	20047

A.4.5 List of Se positions by Phenix

```
CRYST1 70.984 114.753 93.273 90.00 90.00 90.00 C 2 2 21
HETATM 1 SE SE Z 1 18.124 -8.556 -68.382 1.33 24.22 SE
HETATM 2 SE SE Z 2 29.566 -46.380 -18.613 0.78 24.98 SE
HETATM 3 SE SE Z 3 31.938 -5.731 -64.424 0.81 24.88 SE
HETATM 4 SE SE Z 4 22.513 -18.294 -9.542 0.18 20.19 SE
HETATM 5 SE SE Z 5 25.560 -30.104 -36.916 0.15 20.41 SE
HETATM 6 SE SE Z 6 34.482 12.241 -63.311 0.15 20.13 SE
HETATM 7 SE SE Z 7 9.465 -14.196 -38.506 0.23 20.64 SE
HETATM 8 SE SE Z 8 31.491 -0.764 -16.316 0.20 20.50 SE
HETATM 9 SE SE Z 9 19.459 -14.048 -15.544 0.22 20.06 SE
HETATM 10 SE SE Z 10 18.011 6.285 -28.540 0.21 21.05 SE
TER
END
```

A.4.6 T-COFFEE alignment across species

(see next page)


```

          *           20           *           40           *
gi|168803996   G.gallus:  ----MMNADMDAVEAENQ-VELEEKTRLINQVLELQHTLE--DLSARVDAVKKEENL : 49
gi|21357487   D.melanogaster: SSTDSSSHNSLEA-MEMAQ-DDREEKARLITQVLELQNTID--DLSQRVDSVKKEENL : 53
gi|290462705   L.salmonis:  ----SLMSDPQAVDEENDSEEQEEKARLIAQVLELQNTID--DLSQRVDNVKKEENL : 51
gi|291400856   O.cuniculus:  CLLTKMMSADTDAVGAVNQ-VELEEKTRLIHQVLELQHTLE--DLFARVETVKKEENM : 54
gi|332373254   D.ponderosae:  SSSPQ--HNSLDQ--DASP-DEQEEKARLISQVLELQNTID--DLSQRVDSVKKEENL : 50
sp|Q08BG7|     D.rerio:  EQYS-TMNCEDG-DMENQ-VEQEEKTRLINQVLELQHTLE--DLSARVDAVKKEENL : 52
sp|Q5RJZ6|     R.norvegicus:  RLFPKMMNADMDAVDAENQ-VELEEKTRLINQVLELQHTLE--DLSARVDAVKKEENL : 54
sp|Q9UIL1-2|   H.sapiens:  SLLPKMMNADMDAVDAENQ-VELEEKTRLINQVLELQHTLE--DLSARVDAVKKEENL : 54
sp|Q9UIL1-3|   H.sapiens:  SLLPKMMNADMD-VDAENQ-VELEEKTRLINQVLELQHTLE--DLSARVDAVKKEENL : 53
sp|Q9UIL1|     H.sapiens:  SLLPKMMNADMDAVDAENQ-VELEEKTRLINQVLELQHTLE--DLSARVDAVKKEENL : 54
tr|B4NF72|     D.willistoni:  SSTDSSSHNSIDA-LEMAQ-DDREEKARLITQVLELQNTID--DLSQRVDSVKKEENL : 53
tr|F1PFR7|     C.familiaris:  SLLPEMNADMDAVDAENQ-VELEEKTRLINQVLELQHTLEDTDLSARVDAVKKEENL : 56
tr|G5EDQ5|     C.elegans b:  GKTFRAA-QPLP--KEEPP-EDPEEKARMITQVLELQNTID--DLSQRVESVKKEESL : 51
tr|Q3T068|     B.taurus:  SLLPKMMNADMDAVDAENQ-VELEEKTRLINQVLELQHTLE--DLSARVDAVKKEENL : 54
tr|Q9U377|     C.elegans c:  GKTFRAA-QPLP--KEEPP-EDPEEKARMITQVLELQNTID--DLSQRVESVKKEESL : 51

```

```

          60           *           80           *           100           *
gi|168803996   G.gallus:  KLKSENQVLGQYIENLMSASSVFQTTDT-----KSKR-----K : 82
gi|21357487   D.melanogaster:  KLRSENQVLGQYIENLMSASSVFQSTSP-----SAAKK-----K : 87
gi|290462705   L.salmonis:  RLRSENQVLGQYIENLMSASSVFQPANNEASCVKSPITSIK-----K : 93
gi|291400856   O.cuniculus:  KLKSENQVPRHYLRNLMSSSEVVFHYT-----KNKR-----T : 85
gi|332373254   D.ponderosae:  KLRSENQVLGQYIENLMSASSVFQSTSP-----K-ISK-----K : 83
sp|Q08BG7|     D.rerio:  KLKSENQVLGQYIENLMSASSVFQTTDT-----KSKR-----K : 85
sp|Q5RJZ6|     R.norvegicus:  KLKSENQVLGQYIENLMSASSVFQTTDT-----KSKR-----K : 87
sp|Q9UIL1-2|   H.sapiens:  KLKSENQVLGQYIENLMSASSVFQTTDT-----KSKR-----K : 87
sp|Q9UIL1-3|   H.sapiens:  KLKSENQVLGQYIENLMSASSVFQTTDT-----KSKR-----K : 86
sp|Q9UIL1|     H.sapiens:  KLKSENQVLGQYIENLMSASSVFQTTDT-----KSKR-----K : 87
tr|B4NF72|     D.willistoni:  KLRSENQVLGQYIENLMSASSVFQSTNA-----NATKK-----K : 87
tr|F1PFR7|     C.familiaris:  KLKSENQVLGQYIENLMSASSVFQTTDT-----KSKR-----K : 89
tr|G5EDQ5|     C.elegans b:  KLRSENQVLGQYIQNLMSSSVFQSSQP-----S-REKQDDSVHDDDFGEYEY : 98
tr|Q3T068|     B.taurus:  KLKSENQVLGQYIENLMSASSVFQTTDT-----KSKR-----K : 87
tr|Q9U377|     C.elegans c:  KLRSENQVLGQYIQNLMSSSVFQSSQP-----S-RPK-----Q : 84

```

A.5 Mass spectrometry results

(see next page)

Search title	C_Behrens_211211_5					
Search Parameters	-----					
Taxonomy filter Homo sapiens (human)					
Enzyme	Trypsin					
Maximum Missed Cleavages	2					
Instrument type	ESI-QUAD-TOF					
Isotope error mode	0					
Protein hits	-----					
prot_hit_num	prot_acc	prot_desc	prot_score	prot_mass	pep_seq	pep_res_after
1	gij189054178	unnamed p	214	65980	SLVNLGGSK	S
1	gij189054178	unnamed p	214	65980	TLLEGEESR	M
1	gij189054178	unnamed p	214	65980	TNAENEFVTIK	K
1	gij189054178	unnamed p	214	65980	LALDLEIATYR	T
1	gij189054178	unnamed p	214	65980	SLDLDSIIAEVK	A
2	gij11139093	GrpE-like p	143	24133	ALADTENLR	Q
2	gij11139093	GrpE-like p	143	24133	TLRPALVGVVK	E
2	gij11139093	GrpE-like p	143	24133	DLLEVADVLEK	A
3	gij435476	cytokeratin	117	62092	TLLDIDNTR	M
3	gij435476	cytokeratin	117	62092	HGVQELEIELQSQLSKK	A
4	gij226528280	short coiled	63	18034	VDAVKEENLK	L
4	gij226528280	short coiled	63	18034	ASGWDGMGFFSR	R
4	gij226528280	short coiled	63	18034	LINQVLELQHTLEDLSAR	V
4	gij226528280	short coiled	63	18034	SENQVLGQYIENLMSASSVFQTTD	S
4	gij226528280	short coiled	63	18034	SENQVLGQYIENLMSASSVFQTTD	S
4	gij226528280	short coiled	63	18034	SENQVLGQYIENLMSASSVFQTTD	S
4	gij226528280	short coiled	63	18034	LKSENQVLGQYIENLMSASSVFQT	S

Bibliography

1. Clark, S. L. “Cellular Differentiation In the Kidneys of Newborn Mice Studied With the Electron Microscope”. *Journal of Biophysical and Biochemical Cytology* **3**, 349 (1957) (cit. on p. 1).
2. Novikoff, A. B. “The Proximal Tubule Cell In Experimental Hydronephrosis”. *Journal of Biophysical and Biochemical Cytology* **6**, 136 (1959) (cit. on p. 1).
3. Ashford, T. P. & Porter, K. R. “Cytoplasmic Components In Hepatic Cell Lysosomes”. *Journal of Cell Biology* **12**, 198 (1962) (cit. on p. 1).
4. Deduve, C. & Wattiaux, R. “Functions of Lysosomes”. *Annual Review of Physiology* **28**, 435 (1966) (cit. on p. 1).
5. Tsukada, M. & Ohsumi, Y. “Isolation and Characterization of Autophagy-defective Mutants of *Saccharomyces-cerevisiae*”. *FEBS Letters* **333**, 169–174 (1993) (cit. on p. 1).
6. Titorenko, V. I., Keizer, I., Harder, W. & Veenhuis, M. “Isolation and Characterization of Mutants Impaired In the Selective Degradation of Peroxisomes In the Yeast *Hansenula-polymorpha*”. *Journal of Bacteriology* **177**, 357–363 (1995) (cit. on p. 1).
7. Harding, T. M., Morano, K. A., Scott, S. V. & Klionsky, D. J. “Isolation and Characterization of Yeast Mutants In the Cytoplasm To Vacuole Protein Targeting Pathway”. *Journal of Cell Biology* **131**, 591–602 (1995) (cit. on p. 1).
8. Thumm, M. *et al.* “Isolation of Autophagocytosis Mutants of *Saccharomyces-cerevisiae*”. *FEBS Letters* **349**, 275–280 (1994) (cit. on p. 1).
9. Motley, A. M., Nuttall, J. M. & Hettema, E. H. “Pex3-anchored Atg36 tags peroxisomes for degradation in *Saccharomyces cerevisiae*”. *The EMBO Journal* **31**, 2852–2868 (2012) (cit. on pp. 1, 2).
10. Mizushima, N., Yoshimori, T. & Ohsumi, Y. “The Role of Atg Proteins in Autophagosome Formation”. *Annual Review of Cell and Developmental Biology* **27**, 107–132 (2011) (cit. on pp. 1, 3).
11. Kraft, C. & Martens, S. “Mechanisms and regulation of autophagosome formation”. *Current Opinion in Cell Biology* **24**, 496–501 (2012) (cit. on pp. 2, 4).
12. Yang, Z. & Klionsky, D. J. “Mammalian autophagy: core molecular machinery and signaling regulation”. *Current Opinion in Cell Biology* **22**, 124–131 (2010) (cit. on pp. 1, 2, 4, 6).

13. Mizushima, N., Levine, B., Cuervo, A. M. & Klionsky, D. J. “Autophagy fights disease through cellular self-digestion”. *Nature* **451**, 1069–1075 (2008) (cit. on p. 2).
14. Rambold, A. S. & Lippincott-Schwartz, J. “Mechanisms of mitochondria and autophagy crosstalk”. *Cell Cycle* **10**, 4032–4038 (2011) (cit. on p. 2).
15. Kraft, C., Deplazes, A., Sohrmann, M. & Peter, M. “Mature ribosomes are selectively degraded upon starvation by an autophagy pathway requiring the Ubp3p/Bre5p ubiquitin protease”. *Nature Cell Biology* **10**, 602–610 (2008) (cit. on p. 2).
16. Hutchins, M. U. & Klionsky, D. J. “Vacuolar localization of oligomeric alpha-mannosidase requires the cytoplasm to vacuole targeting and autophagy pathway components in *Saccharomyces cerevisiae*”. *Journal of Biological Chemistry* **276**, 20491–20498 (2001) (cit. on p. 2).
17. Behrends, C. & Fulda, S. “Receptor proteins in selective autophagy.” *International Journal of Cell Biology* **2012**, 673290 (2012) (cit. on p. 2).
18. Arias, E. & Cuervo, A. M. “Chaperone-mediated autophagy in protein quality control”. *Current Opinion in Cell Biology* **23**, 184–189 (2011) (cit. on p. 4).
19. Li, W.-w., Li, J. & Bao, J.-k. “Microautophagy: lesser-known self-eating”. *Cellular and Molecular Life Sciences* **69**, 1125–1136 (2012) (cit. on p. 4).
20. Rubinsztein, D. C., Shpilka, T. & Elazar, Z. “Mechanisms of Autophagosome Biogenesis”. *Current Biology* **22**, R29–R34 (2012) (cit. on p. 4).
21. Longatti, A. & Tooze, S. A. “Vesicular trafficking and autophagosome formation”. *Cell Death and Differentiation* **16**, 956–965 (2009) (cit. on p. 4).
22. Tooze, S. A. & Yoshimori, T. “The origin of the autophagosomal membrane”. *Nature Cell Biology* **12**, 831–835 (2010) (cit. on p. 4).
23. Hayashi-Nishino, M. *et al.* “A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation”. *Nature Cell Biology* **11**, 1433 (2009) (cit. on p. 4).
24. Yla-Anttila, P., Vihinen, H., Jokita, E. & Eskelinen, E.-L. “3D tomography reveals connections between the phagophore and endoplasmic reticulum”. *Autophagy* **5**, 1180–1185 (2009) (cit. on p. 4).
25. Axe, E. L. *et al.* “Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum”. *Journal of Cell Biology* **182**, 685–701 (2008) (cit. on p. 4).
26. Hailey, D. W. *et al.* “Mitochondria Supply Membranes for Autophagosome Biogenesis during Starvation”. *Cell* **141**, 656–667 (2010) (cit. on p. 4).
27. Ohashi, Y. & Munro, S. “Membrane Delivery to the Yeast Autophagosome from the Golgi-Endosomal System”. *Molecular Biology of the Cell* **21**, 3998–4008 (2010) (cit. on p. 4).

28. Ravikumar, B., Moreau, K., Jahreiss, L., Puri, C. & Rubinsztein, D. C. "Plasma membrane contributes to the formation of pre-autophagosomal structures". *Nature Cell Biology* **12**, 747 (2010) (cit. on p. 4).
29. Weidberg, H., Shvets, E. & Elazar, Z. "Biogenesis and Cargo Selectivity of Autophagosomes". *Annual Review of Biochemistry, Vol 80* **80**, 125–156 (2011) (cit. on pp. 5, 6).
30. Itakura, E., Kishi, C., Inoue, K. & Mizushima, N. "Beclin 1 Forms Two Distinct Phosphatidylinositol 3-Kinase Complexes with Mammalian Atg14 and UVRAG". *Molecular Biology of the Cell* **19**, 5360–5372 (2008) (cit. on p. 4).
31. Matsunaga, K. *et al.* "Two Beclin 1-binding proteins, Atg14L and Rubicon, reciprocally regulate autophagy at different stages". *Nature Cell Biology* **11**, 385 (2009) (cit. on p. 4).
32. Zhong, Y. *et al.* "Distinct regulation of autophagic activity by Atg14L and Rubicon associated with Beclin 1-phosphatidylinositol-3-kinase complex". *Nature Cell Biology* **11**, 468 (2009) (cit. on p. 4).
33. Takahashi, Y. *et al.* "Bif-1 interacts with Beclin 1 through UVRAG and regulates autophagy and tumorigenesis". *Nature Cell Biology* **9**, 1142–1151 (2007) (cit. on p. 4).
34. Liang, C. *et al.* "Beclin1-binding UVRAG targets the class C Vps complex to coordinate autophagosome maturation and endocytic trafficking". *Nature Cell Biology* **10**, 776–787 (2008) (cit. on p. 4).
35. McKnight, N. C. *et al.* "Genome-wide siRNA screen reveals amino acid starvation-induced autophagy requires SCOC and WAC". *The EMBO Journal* **31**, 1931–1946 (2012) (cit. on pp. 6, 7, 9, 99, 111–114).
36. Van Valkenburgh, H., Shern, J. F., Sharer, J. D., Zhu, X. J. & Kahn, R. A. "ADP-ribosylation factors (ARFs) and ARF-like 1 (ARL1) have both specific and shared effectors: Characterizing ARL1-binding proteins". *Journal of Biological Chemistry* **276**, 22826–22837 (2001) (cit. on pp. 7, 8, 104, 107, 114).
37. Su, C.-W. *et al.* "The short coiled-coil domain-containing protein UNC-69 cooperates with UNC-76 to regulate axonal outgrowth and normal presynaptic organization in *Caenorhabditis elegans*." *Journal of Biology* **5**, 9 (2006) (cit. on pp. 7, 9, 89, 111).
38. Panic, B., Whyte, J. R. C. & Munro, S. "The ARF-like GTPases Arl1p and Arl3p act in a pathway that interacts with vesicle-tethering factors at the Golgi apparatus". *Current Biology* **13**, 405–410 (2003) (cit. on p. 7).
39. Panic, B., Perisic, O., Veprintsev, D. B., Williams, R. L. & Munro, S. "Structural basis for Arl1-dependent targeting of homodimeric GRIP domains to the Golgi apparatus". *Molecular Cell* **12**, 863–874 (2003) (cit. on pp. 7, 8, 104).

40. Chen, K.-Y., Tsai, P.-C., Liu, Y.-W. & Lee, F.-J. S. “Competition between the golgin Imh1p and the GAP Gcs1p stabilizes activated Arl1p at the late-Golgi”. *Journal of Cell Science* **125**, 4586–4596 (2012) (cit. on p. 7).
41. Christis, C. & Munro, S. “The small G protein Arl1 directs the trans-Golgi-specific targeting of the Arf1 exchange factors BIG1 and BIG2”. *Journal of Cell Biology* **196**, 327–335 (2012) (cit. on p. 7).
42. Wu, M. S., Lu, L., Hong, W. J. & Song, H. W. “Structural basis for recruitment of GRIP domain golgin-245 by small GTPase Arl1”. *Nature Structural & Molecular Biology* **11**, 86–94 (2004) (cit. on pp. 8, 104).
43. Nakamura, K. *et al.* “Structural Basis for Membrane Binding Specificity of the Bin/Amphiphysin/Rvs (BAR) Domain of Arfaptin-2 Determined by Arl1 GTPase”. *Journal of Biological Chemistry* **287**, 25478–25489 (2012) (cit. on p. 8).
44. Assmann, E. M., Alborghetti, M. R., Camargo, M. E. R. & Kobarg, J. “FEZ1 dimerization and interaction with transcription regulatory proteins involves its coiled-coil region”. *Journal of Biological Chemistry* **281**, 9869–9881 (2006) (cit. on pp. 8, 9, 102, 111).
45. Lanza, D. C. F. *et al.* “Human FEZ1 has characteristics of a natively unfolded protein and dimerizes in solution”. *Proteins-structure Function and Bioinformatics* **74**, 104–121 (2009) (cit. on pp. 9, 99, 102).
46. Blasius, T. L., Cai, D., Jih, G. T., Toret, C. P. & Verhey, K. J. “Two binding partners cooperate to activate the molecular motor Kinesin-1”. *Journal of Cell Biology* **176**, 11–17 (2007) (cit. on p. 9).
47. Gindhart, J. G. *et al.* “The kinesin-associated protein UNC-76 is required for axonal transport in the Drosophila nervous system”. *Molecular Biology of the Cell* **14**, 3356–3365 (2003) (cit. on p. 9).
48. Chua, J. J. E. *et al.* “Phosphorylation-regulated axonal dependent transport of syntaxin 1 is mediated by a Kinesin-1 adapter”. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 5862–5867 (2012) (cit. on pp. 9, 112).
49. Toda, H. *et al.* “UNC-51/ATG1 kinase regulates axonal transport by mediating motor-cargo assembly”. *Genes & Development* **22**, 3292–3307 (2008) (cit. on p. 9).
50. Bloom, L. & Horvitz, H. R. “The Caenorhabditis elegans gene unc-76 and its human homologs define a new gene family involved in axonal outgrowth and fasciculation”. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 3414–3419 (1997) (cit. on p. 9).
51. Mochizuki, H. *et al.* “Unc-51/ATG1 Controls Axonal and Dendritic Development via Kinesin-Mediated Vesicle Transport in the Drosophila Brain”. *PLoS ONE* **6**, e19632 (2011) (cit. on p. 9).

52. Rackham, O. J. L. *et al.* “The Evolution and Structure Prediction of Coiled Coils across All Genomes”. *Journal of Molecular Biology* **403**, 480–493 (2010) (cit. on p. 10).
53. Wang, Y. *et al.* “Coiled-coil networking shapes cell molecular machinery”. *Molecular Biology of the Cell* **23**, 3911–3922 (2012) (cit. on p. 10).
54. Lupas, A. N. & Gruber, M. “The structure of alpha-helical coiled coils”. *Fibrous Proteins: Coiled-coils, Collagen and Elastomers* **70**, 37 (2005) (cit. on p. 10).
55. Woolfson, D. N. “The design of coiled-coil structures and assemblies”. *Fibrous Proteins: Coiled-coils, Collagen and Elastomers* **70**, 79 (2005) (cit. on pp. 10, 110).
56. Grigoryan, G. & Keating, A. E. “Structural specificity in coiled-coil interactions”. *Current Opinion in Structural Biology* **18**, 477–483 (2008) (cit. on p. 10).
57. Mason, J. M. & Arndt, K. M. “Coiled coil domains: Stability, specificity, and biological implications”. *Chembiochem* **5**, Boehringer Ingelheim Fonds (2004) (cit. on p. 10).
58. Gruber, M., Soeding, J. & Lupas, A. N. “Comparative analysis of coiled-coil prediction methods”. *Journal of Structural Biology* **155**, 140–145 (2006) (cit. on p. 10).
59. Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. “A Switch Between 2-stranded, 3-stranded and 4-stranded Coiled Coils In Gen4 Leucine-zipper Mutants”. *Science* **262**, 1401–1407 (1993) (cit. on pp. 10, 91, 110).
60. Levine, B., Mizushima, N. & Virgin, H. W. “Autophagy in immunity and inflammation”. *Nature* **469**, 323–335 (2011) (cit. on p. 11).
61. Deretic, V. & Levine, B. “Autophagy, Immunity, and Microbial Adaptations”. *Cell Host & Microbe* **5**, 527–549 (2009) (cit. on p. 11).
62. Lamark, T., Kirkin, V., Dikic, I. & Johansen, T. “NBR1 and p62 as cargo receptors for selective autophagy of ubiquitinated targets”. *Cell Cycle* **8**, 1986–1990 (2009) (cit. on p. 12).
63. Thurston, T. L. M., Ryzhakov, G., Bloor, S., von Muhlinen, N. & Randow, F. “The TBK1 adaptor and autophagy receptor NDP52 restricts the proliferation of ubiquitin-coated bacteria”. *Nature Immunology* **10**, 1215 (2009) (cit. on p. 12).
64. Mostowy, S. *et al.* “p62 and NDP52 Proteins Target Intracytosolic Shigella and Listeria to Different Autophagy Pathways”. *Journal of Biological Chemistry* **286**, 26987–26995 (2011) (cit. on pp. 12, 13).
65. Von Muhlinen, N. *et al.* “LC3C, Bound Selectively by a Noncanonical LIR Motif in NDP52, Is Required for Antibacterial Autophagy”. *Molecular Cell* **48**, 329–342 (2012) (cit. on p. 12).

66. Korac, J. *et al.* "Ubiquitin-independent function of optineurin in autophagic clearance of protein aggregates". *Journal of Cell Science* **126**, 580–592 (2013) (cit. on p. 12).
67. Wild, P. *et al.* "Phosphorylation of the Autophagy Receptor Optineurin Restricts Salmonella Growth". *Science* **333**, 228–233 (2011) (cit. on p. 12).
68. Kirkin, V. *et al.* "A Role for NBR1 in Autophagosomal Degradation of Ubiquitinated Substrates". *Molecular Cell* **33**, 505–516 (2009) (cit. on p. 12).
69. Jo, E.-K., Yuk, J.-M., Shin, D.-M. & Sasakawa, C. "Roles of autophagy in elimination of intracellular bacterial pathogens." *Frontiers in Immunology* **4**, 97 (2013) (cit. on p. 12).
70. Shaid, S., Brandts, C. H., Serve, H. & Dikic, I. "Ubiquitination and selective autophagy". *Cell Death and Differentiation* **20**, 21–30 (2013) (cit. on p. 12).
71. Harriff, M. J., Purdy, G. E. & Lewinsohn, D. M. "Escape from the Phagosome: The Explanation for MHC-I Processing of Mycobacterial Antigens?" *Frontiers in Immunology* **3**, 40 (2012) (cit. on p. 12).
72. Huynh, K. K., Joshi, S. A. & Brown, E. J. "A delicate dance: host response to mycobacteria". *Current Opinion in Immunology* **23**, 464–472 (2011) (cit. on p. 12).
73. Deretic, V. *et al.* "Mycobacterium tuberculosis inhibition of phagolysosome biogenesis and autophagy as a host defence mechanism". *Cellular Microbiology* **8**, 719–727 (2006) (cit. on p. 12).
74. Basu, J., Shin, D.-M. & Jo, E.-K. "Mycobacterial signaling through toll-like receptors." *Frontiers in Cellular and Infection Microbiology* **2**, 145 (2012) (cit. on p. 12).
75. Bakowski, M. A., Braun, V. & Brumell, J. H. "Salmonella-Containing Vacuoles: Directing Traffic and Nesting to Grow". *Traffic* **9**, 2022–2031 (2008) (cit. on p. 12).
76. Ashida, H. *et al.* "Shigella are versatile mucosal pathogens that circumvent the host innate immune system". *Current Opinion in Immunology* **23**, 448–455 (2011) (cit. on p. 12).
77. Egile, C. *et al.* "Activation of the CDC42 effector N-WASP by the Shigella flexneri IcsA protein promotes actin nucleation by Arp2/3 complex and bacterial actin-based motility". *Journal of Cell Biology* **146**, 1319–1332 (1999) (cit. on pp. 12, 14, 131).
78. Ogawa, M. *et al.* "Escape of intracellular Shigella from autophagy". *Science* **307**, 727–731 (2005) (cit. on p. 12).
79. Lee, M.-S. *et al.* "Shiga toxins induce autophagy leading to differential signalling pathways in toxin-sensitive and toxin-resistant human cells". *Cellular Microbiology* **13**, 1479–1496 (2011) (cit. on p. 13).

80. Dupont, N. *et al.* “Shigella Phagocytic Vacuolar Membrane Remnants Participate in the Cellular Response to Pathogen Invasion and Are Regulated by Autophagy”. *Cell Host & Microbe* **6**, 137–149 (2009) (cit. on p. 13).
81. Mostowy, S. *et al.* “Entrapment of Intracytosolic Bacteria by Septin Cage-like Structures”. *Cell Host & Microbe* **8**, 433–444 (2010) (cit. on pp. 13, 131).
82. Junker, M. *et al.* “Pertactin beta-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins”. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4918–4923 (2006) (cit. on p. 14).
83. Janakiraman, A., Fixen, K. R., Gray, A. N., Niki, H. & Goldberg, M. B. “A Genome-Scale Proteomic Screen Identifies a Role for DnaK in Chaperoning of Polar Autotransporters in Shigella”. *Journal of Bacteriology* **191**, 6300–6311 (2009) (cit. on p. 13).
84. Wagner, J. K., Heindl, J. E., Gray, A. N., Jain, S. & Goldberg, M. B. “Contribution of the Periplasmic Chaperone Skp to Efficient Presentation of the Autotransporter IcsA on the Surface of Shigella flexneri”. *Journal of Bacteriology* **191**, 815–821 (2009) (cit. on p. 14).
85. Kühnel, K. & Diezmann, D. “Crystal Structure of the Autochaperone Region from the Shigella flexneri Autotransporter IcsA”. *Journal of Bacteriology* **193**, 2042–2045 (2011) (cit. on pp. 14, 120, 130).
86. Benz, I. & Schmidt, M. A. “Structures and functions of autotransporter proteins in microbial pathogens”. *International Journal of Medical Microbiology* **301**, 461–468 (2011) (cit. on p. 14).
87. May, K. L. & Morona, R. “Mutagenesis of the Shigella flexneri autotransporter IcsA reveals novel functional regions involved in IcsA biogenesis and recruitment of host neural Wiscott-Aldrich syndrome protein”. *Journal of Bacteriology* **190**, 4666–4676 (2008) (cit. on pp. 14, 131).
88. Allaoui, A., Mounier, J., Prevost, M. C., Sansonetti, P. J. & Parsot, C. “icsB: a Shigella-flexneri Virulence Gene Necessary For the Lysis of Protrusions During Intercellular Spread”. *Molecular Microbiology* **6**, 1605–1616 (1992) (cit. on p. 15).
89. Ogawa, M., Suzuki, T., Tatsuno, I., Abe, H. & Sasakawa, C. “IcsB, secreted via the type III secretion system, is chaperoned by IpgA and required at the post-invasion stage of Shigella pathogenicity”. *Molecular Microbiology* **48**, 913–931 (2003) (cit. on pp. 15, 124).
90. Studier, F. W. “Protein production by auto-induction in high-density shaking cultures”. *Protein Expression and Purification* **41**, 207–234 (2005) (cit. on pp. 20–23).
91. Laemmli, U. K. “Cleavage of Structural Proteins During Assembly of Head of Bacteriophage-t4”. *Nature* **227**, 680 (1970) (cit. on p. 47).

92. Schagger, H. & von Jagow, G. "Tricine Sodium Dodecyl-sulfate Polyacrylamide-gel Electrophoresis For the Separation of Proteins In the Range From 1-kda To 100-kda". *Analytical Biochemistry* **166**, 368–379 (1987) (cit. on p. 47).
93. Towbin, H., Staehelin, T. & Gordon, J. "Electrophoretic Transfer of Proteins From Polyacrylamide Gels To Nitrocellulose Sheets: Procedure and Some Applications". *Proceedings of the National Academy of Sciences of the United States of America* **76**, 4350–4354 (1979) (cit. on p. 48).
94. Einarson, M. B. "Detection of protein-protein interactions using the GST fusion protein pull-down technique". *Nature Methods* **1**, 275–276 (2004) (cit. on p. 49).
95. Kabsch, W. "Automatic Processing of Rotation Diffraction Data From Crystals of Initially Unknown Symmetry and Cell Constants". *Journal of Applied Crystallography* **26**, 795–800 (1993) (cit. on pp. 53, 82).
96. Adams, P. D. *et al.* "PHENIX: a comprehensive Python-based system for macromolecular structure solution". *Acta Crystallographica Section D* **66**, 213–221 (2010) (cit. on pp. 54, 82).
97. Emsley, P. & Cowtan, K. "Coot: model-building tools for molecular graphics". *Acta Crystallographica Section D* **60**, 2126–2132 (2004) (cit. on pp. 54, 85).
98. Schrödinger, LLC. *The PyMOL Molecular Graphics System, Version 1.3r1*. 2010 (cit. on p. 55).
99. Livingstone, C. D. & Barton, G. J. "Protein-sequence Alignments—A Strategy For the Hierarchical Analysis of Residue Conservation". *Computer Applications in the Biosciences* **9**, 745–756 (1993) (cit. on pp. 55, 92).
100. Lupas, A., van Dyke, M. & Stock, J. "Predicting Coiled Coils From Protein Sequences". *Science* **252**, 1162–1164 (1991) (cit. on p. 57).
101. Dosztányi, Z., Csizmók, V., Tompa, P. & Simon, I. "The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins". *Journal of Molecular Biology* **347**, 827–839 (2005) (cit. on pp. 57, 58).
102. Dosztányi, Z., Csizmók, V., Tompa, P. & Simon, I. "IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content". *Bioinformatics* **21**, 3433–3434 (2005) (cit. on pp. 57, 58).
103. Remmert, M., Biegert, A., Hauser, A. & Söding, J. "HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment". *Nature Methods* **9**, 173–175 (2012) (cit. on p. 59).
104. Söding, J. "Protein homology detection by HMM-HMM comparison". *Bioinformatics* **21**, 951–960 (2005) (cit. on p. 59).
105. Söding, J., Biegert, A. & Lupas, A. N. "The HHpred interactive server for protein homology detection and structure prediction". *Nucleic Acids Research* **33**, W244–W248 (2005) (cit. on p. 59).

106. Overton, I. M. *et al.* “TarO: a target optimisation system for structural biology”. *Nucleic Acids Research* **36**, W190–W196 (2008) (cit. on p. 60).
107. Evans, G. & Pettifer, R. F. “CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra”. *Journal of Applied Crystallography* **34**, 82–86 (2001) (cit. on p. 81).
108. Strelkov, S. V. & Burkhard, P. “Analysis of alpha-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation”. *Journal of Structural Biology* **137**, 54–64 (2002) (cit. on p. 89).
109. Notredame, C., Higgins, D. G. & Heringa, J. “T-Coffee: A novel method for fast and accurate multiple sequence alignment”. *Journal of Molecular Biology* **302**, 205–217 (2000) (cit. on p. 89).
110. Suzuki, T. *et al.* “Identification of FEZ1 as a protein that interacts with JC virus agnoprotein and microtubules: Role of agnoprotein-induced dissociation of FEZ1 from microtubules in viral propagation”. *Journal of Biological Chemistry* **280**, 24948–24956 (2005) (cit. on p. 99).
111. Alborghetti, M. R. *et al.* “Human FEZ1 Protein Forms a Disulfide Bond Mediated Dimer: Implications for Cargo Transport”. *Journal of Proteome Research* **9**, 4595–4603 (2010) (cit. on p. 102).
112. Smith, S. J. M. & Rittinger, K. “Preparation of GTPases for structural and biophysical analysis.” *Methods in molecular biology (Clifton, N.J.)* **189**, 13–24 (2002) (cit. on p. 104).
113. Li, X. *et al.* “Imperfect interface of Beclin1 coiled-coil domain regulates homodimer and heterodimer formation with Atg14L and UVRAG”. *Nature Communications* **3**, 662 (2012) (cit. on p. 110).
114. Lanza, D. C. F. *et al.* “FEZ1 interacts with CLASP2 and NEK1 through coiled-coil regions and their cellular colocalization suggests centrosomal functions and regulation by PKC”. *Molecular and Cellular Biochemistry* **338**, 35–45 (2010) (cit. on p. 112).
115. Di Bartolomeo, S. *et al.* “The dynamic interaction of AMBRA1 with the dynein motor complex regulates mammalian autophagy”. *Journal of Cell Biology* **191**, 155–168 (2010) (cit. on p. 113).
116. Russell, R. C. *et al.* “ULK1 induces autophagy by phosphorylating Beclin-1 and activating VPS34 lipid kinase”. *Nature Cell Biology*. advance online publication (2013) (cit. on p. 113).
117. Wirth, M., Joachim, J. & Tooze, S. A. “Autophagosome formation—The role of ULK1 and Beclin1–PI3KC3 complexes in setting the stage”. *Seminars in Cancer Biology*. in press (cit. on p. 113).
118. Kayath, C. A. *et al.* “Escape of intracellular Shigella from autophagy requires binding to cholesterol through the type III effector, IcsB”. *Microbes and Infection* **12**, 956–966 (2010) (cit. on p. 131).

119. Suzuki, T. *et al.* “Neural Wiskott-Aldrich syndrome protein (N-WASP) is the specific ligand for Shigella VirG among the WASP family and determines the host cell type allowing actin-based spreading”. *Cellular Microbiology* **4**, 223–233 (2002) (cit. on p. 131).
120. Ashida, H. *et al.* “A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKK gamma to dampen the host NF-kappa B-mediated inflammatory response”. *Nature Cell Biology* **12**, 66 (2010) (cit. on p. 131).