technische universität
dortmund

MAX-PLANCK-GESELLSCHAFT

Dissertation

Makes caterpillars floppy toxins Mcf1 and Mcf2 from *Photorhabdus luminescens*

Vorgelegt von

Philipp Heilen

Zur Erlangung des akademischen Grades des Doktors der Naturwissenschaft

(Dr. rer. nat.)

Vorgelegt der Fakultät Chemie und Chemische Biologie

Der Technischen Universität Dortmund

Angefertigt am

Max-Planck-Institute für Molekulare Physiologie Dortmund

Abteilung für Strukturbiochemie

Dortmund, 04.06.2025

1st referee: Prof. Dr. Stefan Raunser

Department of Structural Biochemistry, Max-Planck-Institute of Molecular Physiology, Dortmund

Faculty of Chemistry and Chemical Biology, Technical University Dortmund

2nd referee: Prof. Dr. Rasmus Linser

Department of Physical Chemistry

Faculty of Chemistry and Chemical Biology, Technical University Dortmund

Date of submission: 04.06.2025

Table of Contents

1. List of Abbreviations

aa	Amino acids
ABD	Activator binding domain
ABH	Alpha/Beta hydrolase
ACD	Actin crosslinking domain
ADP	Adenosine diphosphate
AF3	Alphafold 3
Arf3	ADP-ribosylation factor 3
ATP	Adenosine triphosphate
Bad	Bcl2-anatgonist of cell death
Bak	Bcl2-antagonist/killer
Bax	Bcl2-associated X protein
Bcl2	B-cell lymphoma/leukemia 2
Bid	BH3 interacting domain death agonist
BH3 domain	Bcl2-homology 3-like domain
Blf1	Burkholderia lethal factor 1
bp	Base pair
BT	*Bacillus thuringiensis*
CIS	Contractile injection system
CPD	Cysteine protease domain
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CROPS	Combined repetitive oligopeptide sequences
Cryo-EM	Electron cryogenic microscopy
CSPG4	Chondroitin sulfate proteoglycan 4
DMSO	Dimethyl sulfoxide
DNA	Desoxyribonucleic acids
dNTPs	Deoxynucleotide triphosphate
dsDNA	Double stranded DNA
DUF1	Domain of unknown function
EEL	Electron energy-loss
FZD	Frizzled
g	gravity
GC content	Guanine Cytosine content

GDP	Guanosine diphosphate
gp96	Glycoprotein 96
GTD	Glycosyltransferase domain
GTP	Guanosine triphosphate
h, min, sec	hour, minute, second
HEK	Human Embryonic Kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HR	Hypersensitive response
Hrk	Activator of apoptosis harakiri
ICP-MS	Inductively coupled plasma mass spectrometry
IP6	Inositol hexakisphosphate
IPTG	Isopropyl-β-D-thiogalactopyranoside
ISAC	Iterative stable alignment and clustering approach
ISAS	Leibnitz-Institut für analytische Wissenschaften
ITC	Isothermal titration calorimetry
kDa, mDa	kilo Dalton, mega Dalton
LB	Lysogeny broth
LC-MS/MS	Liquid chromatography-mass spectrometry
LCT	Large clostridial toxins
LDLR	Low-density lipoprotein receptor
LN	Liquid nitrogen
MARTX	Multifunctional-autoprocessing repeats-in-toxins
Mcf	Makes caterpillars floppy
MS	Mass spectrometry
MW	Molecular weight
NAD	Nicotinamide adenine dinucleotide
NanoDSF	Nano differential scanning fluorometry
NED	N-terminal effector domain
Ni-IDA	Nickel iminodiacetic acid
OD	Optical density
PARP	Poly(ADP-ribose)polymerase
PCR	Polymerase chain reaction
pdb	Protein data base
PED	Protease effector domain

PEG	Polyethylene glycol
pI	Isoelectric point
Pir	*Photorhabdus* insect-related
PVC	*Photorhabdus* virulence cassette
PVDF	Polyvinylidene Fluoride
Rab	Ras related protein in brain
RBD	Receptor binding domain
RID	Rho GTPase-inactivation domain
rpm	Revolutions per minute
RT	Room temperature
RTX	Repeats in toxin
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Sf9	*Spodoptera frugiperda*
SI	Sucrase-isomaltase
SPA	Single particle analysis
Tc	Toxin complex
TcdA/B	*Clostridioides difficile* toxin A/B
TD	Translocation domain
TH	Transmembrane helix
TrisHCl	Tris(hydroxymethyl)aminomethane hydrochloride
UDP	Uridine diphosphate
UV-Vis	Ultraviolet -visible
V-ATPase	Vacuolar-type ATPase
VPP	Volta phase plate
X-FEG	Extreme-brightness Field Emission Gun
YNB	Yeast nitrogen base
YPD	Yeast peptone dextrose
µl, ml, l	microliter, milliliter, liter

2. Zusammenfassung

Die ‘Makes caterpillars floppy’ (Mcf) Toxine 1 und 2 sind Virulenzfaktoren des entomopathogenen Bakteriums *Photorhabdus luminescens*. Dieses lebt in Symbiose mit Nematoden der Gattung *Heterohabditis* und verursacht den Tod des Wirtes der als Nahrungsquelle für beide Organismen dient. Während Mcf1 Apoptose im Mitteldarm des Insektenwirtes induziert, ist der Wirkmechanismus von Mcf2 gänzlich unbekannt. Beide Toxine teilen eine konservierte zentrale und C-terminale Region, unterscheiden sich allerdings in ihrem N-terminalen Effektor. Trotz ihres Potenzials für den biologischen Pflanzenschutz und für ihre Anwendung in der Biomedizin und Biochemie, sind die Struktur und der Wirkmechanismus dieser Toxine weitgehend unerforscht.

Im Rahmen dieser Arbeit habe ich die Kryoelektronenmikroskopie genutzt um die Struktur von Mcf1 (3.6 Å) und Mcf2 zu erhalten. Allerdings stellte sich Mcf2 im Vergleich zu Mcf1 als instabiler und flexibler heraus, was die Strukturaufklärung erschwerte. Dennoch konnte ich die Struktur beider Toxine bestimmen, welche sich als Seepferdchen-ähnlich beschreiben lässt. Sie besteht aus einem flexiblen Kopf, der mit einem länglichen Schwanz verbunden ist. Eine genaue Analyse der Struktur von Mcf2 zeigte, dass der N-terminale Effektor anders positioniert ist als in Mcf1 und unterschiedliche Konformationen annimmt.

Die Kopfregion beider Toxine besteht aus einer Protease-Effektor-Domäne (PED), einer Aktivator-bindenden-Domäne (ABD) und einer N-terminalen-Effektor-Domäne (NED), die allerdings nicht konserviert ist. Der Schwanz umfasst zwei mutmaßliche Translokationsdomänen (TD) und drei mutmaßliche Rezeptor-bindende Domänen (RBD). Diese konservierte Anordnung der Domänen ermöglichte es mir die Mcf Toxine als Teil der ABCD-Toxin-Familie zu klassifizieren.

Basierend auf biochemischen Experimenten und struktureller Homologie, konnte ich weiterhin aufzeigen, dass die Translokation von der Ansäuerung der Endosomen abhängt und dass die Translokationsdomänen den ‘Large clostridial toxins’ (LCT) ähneln. Somit sind auch die Mcf Toxine ein Teil der LCT-Superfamilie.

Zusätzlich habe ich die GTPase Arf3 als zytosolischen Aktivator der Autoproteolyse von Mcf identifiziert, der die Freisetzung der Effektoren nach erfolgreicher Translokation initiiert. Die Struktur des Mcf1ΔC15-Arf3 Komplexes zeigte, dass GTP-gebundenes Arf3 an der ABD bindet um die Kaskade proteolytischer Reaktionen auszulösen. Des Weiteren

bildete die Struktur die Grundlage für eine Reihe von Mutanten mit deren Hilfe ich einen Aktivierungsmechanismus formulieren konnte.

Nach der Translokation der Kopfregion in das Zytosol, bindet Arf3 die ABD an einem Helix-Bündel, das zuvor von dem C-terminus von Mcf1 blockiert wurde. Diese Interaktion wird vermutlich während der Translokation unterbrochen und stellt sicher, dass Mcf1 nicht verfrüht aktiviert wird. Die Bindung von Arf3 begradigt das Helix-Bündel und lässt die ABD nach außen schwingen, geleitet von einer zentralen Haken-Schleife-Interaktion zwischen ABD und PED. Diese orchestrierte Bewegung zieht die Schnittstelle auf dem NED-ABD Linker zum aktiven Zentrum der Protease und initiiert die Proteolyse.

Hierbei konnte ich zeigen, dass während einer in vitro Aktivierung mit Arf3, Mcf1 an drei Positionen geteilt wird, was zu den folgenden vier Fragmenten führt: NED, ABD, PED und Schwanz-Fragment. Im Gegensatz dazu wird Mcf2 nur zweimal gespalten, was vermutlich an der veränderten Position der NED und ihrer erhöhten Flexibilität liegt. Somit wird bei Mcf2 entweder die gesamte Kopfregion freigesetzt oder die NED zusammen mit einem nicht separierten ABD/PED Fragment. Die Expression der unterschiedlichen Fragmente in Hefezellen führte zur Entdeckung von zwei toxischen Effektoren. Zusätzlich zur NED hat auch die PED toxische Eigenschaften, was ihr eine faszinierende Doppelfunktion verleiht: Auf der einen Seite verursacht die PED die Proteolyse und Freisetzung der Fragmente, auf der anderen ist sie selbst ein toxischer Effektor, der zur hohen Potenz der Mcf Toxine beiträgt.

Zusammenfassend geben meine Entdeckungen einen detaillierten Einblick in die modulare Architektur, den Aktivierungsmechanismus und die Effektordiversität der Mcf Toxine von *P. luminescens*. Die vergleichende Analyse von Mcf1 und Mcf2 veranschaulicht ihre Vielseitigkeit unterschiedlich große N-terminale Effektoren zu transportieren (37-105 kDa) und hebt die daraus resultierenden Veränderungen in der Aktivierung und Proteolyse hervor. Damit legt meine Arbeit das Fundament für die Weiterentwicklung der Mcf Toxine zur Anwendung in der Schädlingsbekämpfung und als neuartige biochemische Werkzeuge.

3. Abstract

The ‘Makes caterpillars floppy’ (Mcf) toxins 1 and 2 are major virulence factors of the entomopathogenic *Photorhabdus luminescens*. They kill the insect hosts of *P. luminescens*, resulting in a floppy morphology, to ensure the survival of the bacterium and its symbiotic nematodes from the *Heterohabditis* genus. While Mcf1 is known to induce apoptosis in the insect midgut, the mechanism by which Mcf2 exerts its toxicity is entirely unknown. Both toxins share a conserved central and C-terminal region, but differ in their putative N-terminal effector. Despite their potential for biological pest control and use as biomedical or biochemical tools, the structure and mechanism of these toxins remains unknown.

In this work, I applied electron cryogenic microscopy (cryo-EM) to solve the structures of Mcf1 (3.6 Å) and Mcf2. However, compared to Mcf1, Mcf2 displayed an increased instability as well as conformational flexibility, which impeded high-resolution structure determination and yielded a partial and distorted cryo-EM density. Despite these challenges, both toxins were found to adopt a characteristic seahorse-like shape consisting of a flexible head and an elongated tail. A detailed analysis of the Mcf2 head further revealed, that the N-terminal effector is positioned differently than in Mcf1 and adopts different conformations.

In both toxins, the head region contains a protease effector domain (PED), an activator binding domain (ABD), and the N-terminal effector domain (NED), which however is not conserved between Mcf1 and Mcf2. The tail encompasses two putative translocation (TD) and three putative receptor-binding domains (RBD). This conserved domain organization enabled me to classify the Mcf toxins as part of the ABCD toxin family.

Based on biochemical experiments and structural homology, I could demonstrate that translocation depends on endosomal acidification and that the translocase is similar to that of the large clostridial toxins (LTC), thereby revealing Mcf toxins as members of the LCT superfamily.

Furthermore, I identified Arf3, a small GTPase, as the cytosolic activator of Mcf that causes the effector release into the cytosol after translocation. The structure of the Mcf1ΔC15-Arf3 complex revealed that GTP bound Arf3 engages the ABD to trigger a cascade of autoproteolytic cleavage reactions. It moreover enabled me to perform a series of mutations and propose an activation mechanism.

After translocation of the head into the cytosol, Arf3 binds to the ABD on a helix bundle that was previously occupied by the C-terminus of Mcf1. This interaction is likely lost during translocation and ensures that Arf3 triggered activation can only occur after translocation. The Arf3 binding straightens the helix bundle and causes a swinging motion of the ABD that is guided by a central hook-loop interaction of ABD and PED. This organized movement pulls the cleavage site on the NED-ABD linker towards the active protease and initiates proteolysis.

I could demonstrate by in vitro activation with Arf3 that Mcf1 undergoes three cleavage events, yielding four fragments: NED, ABD, PED, and the tail. In contrast, Mcf2 exhibited a different cleavage pattern, likely due to the altered position of the NED and increased head flexibility. This resulted in two cleavage reactions releasing either the complete head or the NED together with a single ABD/PED fragment.

By employing a yeast viability assay and expressing the different fragments, I could further uncover the presence of two toxic effectors. In addition to the NED, the PED also exerts toxicity giving it an intriguing dual function: On one hand it causes the proteolytic cleavage and effector release, on the other it is a toxic effector itself, contributing to the potency of Mcf toxins.

Collectively, my findings provide a detailed insight into the modular architecture, activation mechanism, and effector diversity of the Mcf toxins of *P. luminescens*. The comparative analysis of Mcf1 and Mcf2 underscores their versatility to translocate variable sized N-terminal effectors (37-105 kDa), and highlights the resulting changes in activation and autoproteolysis. My work lays a foundation for the rational engineering of Mcf toxins for use in pest control and as novel biochemical tools.

4. Introduction

Since its origin, life had to fight for survival, not only against environmental influences but also against competition within. As part of the kingdom of Monera [1], bacteria therefore forged an arsenal of virulence factors to defend their biological niche and gain access to resources [2]. Since virulence factors have to perform sophisticated tasks for the bacteria, these intricate molecular machines are also of interest for a wide range of applications such as medical delivery systems for gene and cancer therapy [3][4], biological pest control to overcome emerging resistances [5][6] and biotechnological tools that exploit their mechanism to manipulate the molecules of life [7].

4.1 *Photorhabdus luminescens*

One example of such bacteria is the entomopathogenic gram-negative *Photorhabdus luminescens*, which lives in symbiosis with nematodes of the genus *Heterohabditis* [8]. The nematodes can be found ubiquitously in the soil all around the globe, with the exception of antarctica [9], where they prey on a large variety of hosts like beetle larvae (e.g. *Otiorhynchus sulcatus*) and caterpillars (*Korscheltellus lupulina*) [10]. Its complex life cycle switches between a symbiotic and pathogenic phase when encountering the insect larvae, which makes it an attractive tool for targeted pest control [11][12], [13].

Other *Photorhabdus* species are *Photorhabdus temperata* which is also entomopathogenic and *Photohabdus asymbiotica* that was isolated from human specimen, demonstrating its ability to also infect humans [11][14].

4.2 Lifecyle of entomopathogenic nematodes and *P. luminescens*

The *Heterohabditis* nematodes start their lifecycle as dauer larvae, also called infective juveniles, which is a non-feeding stage characterized by a closed stoma [15]. They track down host larvae by gustation, olfactory-, thermo- and humidity sensing [16]. As soon as they reach their host, they invade the insect larvae through natural openings or by penetrating the cuticle and locate to the hemolymph [17]. At this point the parasitic nematodes are unable to kill the host on their own. Therefore, they switch to the pathogenic phase by regurgitating *Photorhabdus* bacteria from their pharynx into the hemolymph. In contrast to the nematode, *Photorhabdus luminescens* wields an arsenal of virulence factors

that are secreted and kill the larvae within 24h [18] while suppressing an immune response of the host [19]. This allows both the nematode and bacterium to feed off the cadaver and to reproduce inside the larvae for two or more generations [20] depending on the size of the larva. Once the food source is depleted, new infective juveniles are formed that take up *P. luminescens* and search for a new host (Fig. 1).

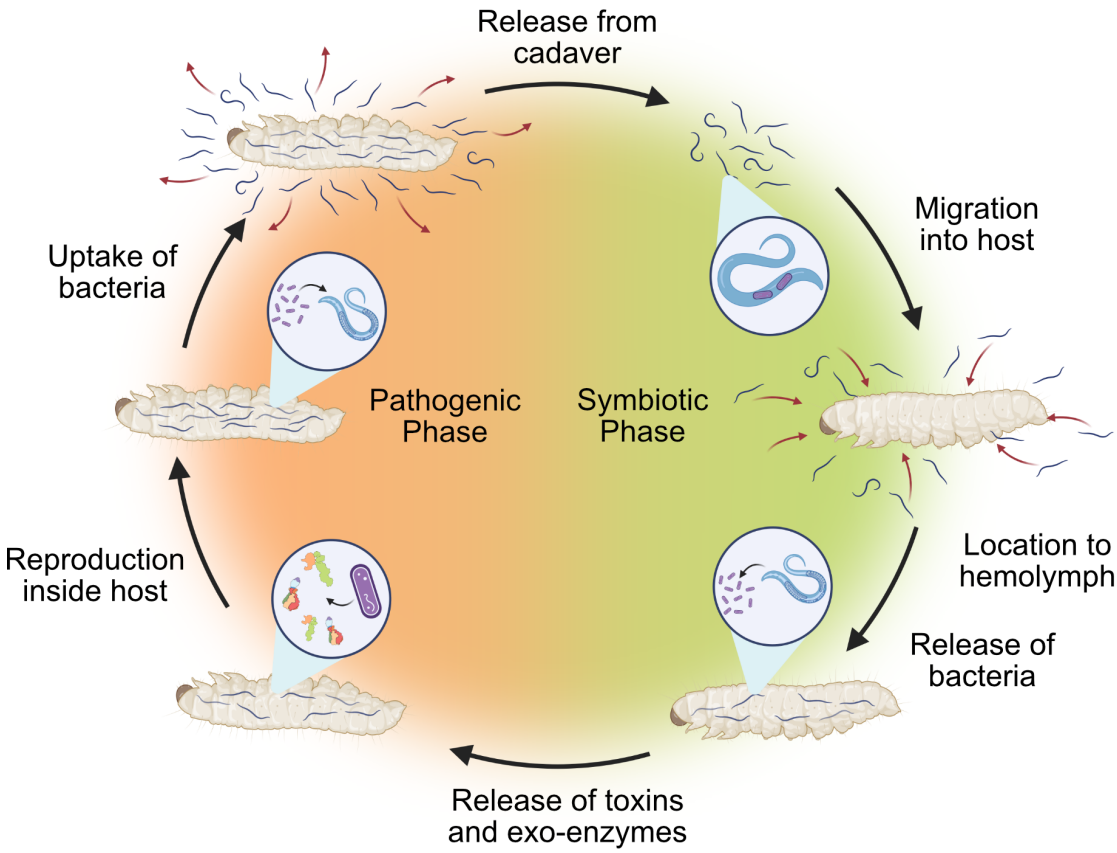


Figure 1 The Lifecycle of *Heterohabditis* nematodes: The nematodes start their lifecycle as infective juveniles inside the soil, where they sense and migrate into their host larvae through natural openings. Afterwards, they locate to the hemolymph and switch to their pathogenic phase, where *P. luminescens* bacteria (violet) are released. The bacteria secrete toxins (Tc and Mcf1 toxin depicted) and other exo-enzymes that kill the host and allow the feeding and reproduction of *P. luminescens* and *Heterohabditis*. Once the nutrients are depleted, the nematodes take up the symbiotic bacteria and search for a new host. Figure adapted from [13] and created with Biorender.

In addition to virulence factors, *Photorhabdus* also produces antibiotics to defend the resources against other microorganisms. This effect was first observed during the 'American Civil War' by soldiers, whose wounds began to glow in the dark. The eponymous luminescence of *Photorhabdus* bacteria was caused by a luciferase and correlated with a faster recovery of the soldiers leading to its name 'angles glow'. Now it is known that this increased recovery rate was caused by the antibiotics produced by *Photorhabdus* [21][22]. To this day the bacterial genus is a source of new antibiotics like darobactin which was discovered in the year 2019 [23].

Taken together the antimicrobial and targeted entomopathogenic properties are the reason why *P. luminescens* is deployed in new pest control strategies to reduce the use of less precise chemical pesticides [24]. Furthermore, the potential usage of its toxins to create insect resistant transgenic plants is explored in the light of the increasing number of *Bacillus thuringiensis* (BT) resistant pests that feed on BT crops [25].

4.3 Toxins of *P. luminescens*

To utilize and engineer the virulence factors of *Photorhabdus luminescens*, a detailed insight into their mechanism of intoxication is required and an understanding on how these molecular machines adapted to the challenges a toxin encounters while killing a host cell. Namely, the targeting of specific cells followed by pore formation or the delivery of a toxic payload through the membrane into the cell where it is activated. Furthermore, the whole process needs to ensure the stability of the toxin and prevent uncontrolled activation [26][27].

Some examples of virulence factors that overcame those challenges and are found in *Photorhabdus* are: The 1.7 MDa large toxin complexes (Tc) which deliver a toxic cargo incapsulated inside a cocoon via a syringe like mechanism [28]. The *Photorhabdus* insect related toxin (Pir), which is a binary toxin consisting of PirA (45 kDa) and PirB (14 kDa) and suggested to function similar to the Cry toxin of *Bacillus thuringiensis* [29]. The XaxAB-like toxin, another binary toxin that was shown to induce necrosis in insect midgut cells [30]. The *Photorhabdus* virulence cassette (10 MDa, PVC), which is a contractile injection system (CIS) that injects various toxic cargos into the target cells [31][32]. And finally, the ‘makes caterpillars floppy’ toxins (Mcf), which encompasses both Mcf1 (324 kDa) and Mcf2 (262 kDa) who apart from an initial characterization [33][34][35] still need to be investigated in more detail to understand their mechanism of toxicity.

4.4 Mcf1 toxin makes caterpillars floppy

The Mcf1 toxin was discovered when expressing a *P. luminescens* subsp. *akhurstii* (W14) genomic library in *Escherichia coli* cells and subsequently injecting them in *Manduca sexta* larvae. One gene was observed to cause a loss of body turgor 12h after infection ultimately leading to the death of the larvae within 24h (Fig. 2 a-c). Therefore, the corresponding gene was named ‘Makes caterpillars floppy’ [34] after its floppy morphology.

A sequence comparison revealed several regions of homology and provided first insights into a possible mechanism of the toxin (Fig. 2 d). Starting at the N-terminus, a short sequence of 15 amino acids shows similarity to a conserved region of the pro-apoptotic Bcl homology domain 3 (BH3) [36], henceforth called BH3-like domain. Continuing towards the center, a region follows with similarity to the 50 kDa MCF-like effector from the multifunctional-autoprocessing repeats-in-toxin (MARTX) toxins of *vibrio vulnificus* [37][38]. Furthermore, 567 amino acids resemble (20% identity) the translocation machinery of *Clostridioides difficile* toxin A and B (TcdA/B) [39]. And lastly, a C-terminal region, ranging from amino acids 2,791–2,925, is similar to putative receptor binding regions of the RTX-like toxin from *Actinobacillus pleuropneumoniae* [40].

A more detailed analysis of larvae infected with Mcf1 expressing *E. coli* led to the conclusion that Mcf1 destroys the insect midgut cells by disintegrating the actin skeleton and inducing apoptosis possibly due to the BH3-like domain. In addition, Mcf1 also kills hemocytes, preventing the encapsulation of the *E. coli* cells and their clearance by the immune system. The disintegration of the midgut and malpighian tubules of the larvae are the cause for the floppy morphology and death of the larvae [34].

When topically applied, Mcf1 also kills mammalian cell lines (COS-7, NIH3T3, HeLa) by inducing apoptosis as the intoxication causes caspase 3 activity, PARP cleavage and DNA laddering within 6h. These effects can be prevented by the caspase inhibitor zVAD-fmk. To investigate whether the BH3-like domain induces apoptosis, the N-terminal 1280 residues of Mcf1, which contain the BH3-like domain, were expressed in NIH3T3 cells. Interestingly, they localize to the Golgi and kill the cells accompanied by cell shrinkage towards the nucleus which did not show fragmentation. In addition, the cells did not exhibit a clear apoptotic phenotype [33].

Although being initially described in the *akhurstii* (W14) subspecies, homologs to Mcf1 could also be identified in *P. temperata* [41] and *P. asymbiotica* [34] as well as in the subspecies *laumondii* (TT01). In regard to a possible release mechanism, parts of a type IV secretion system were found upstream of the Mcf1 gene in *P. luminescens* [34][42].

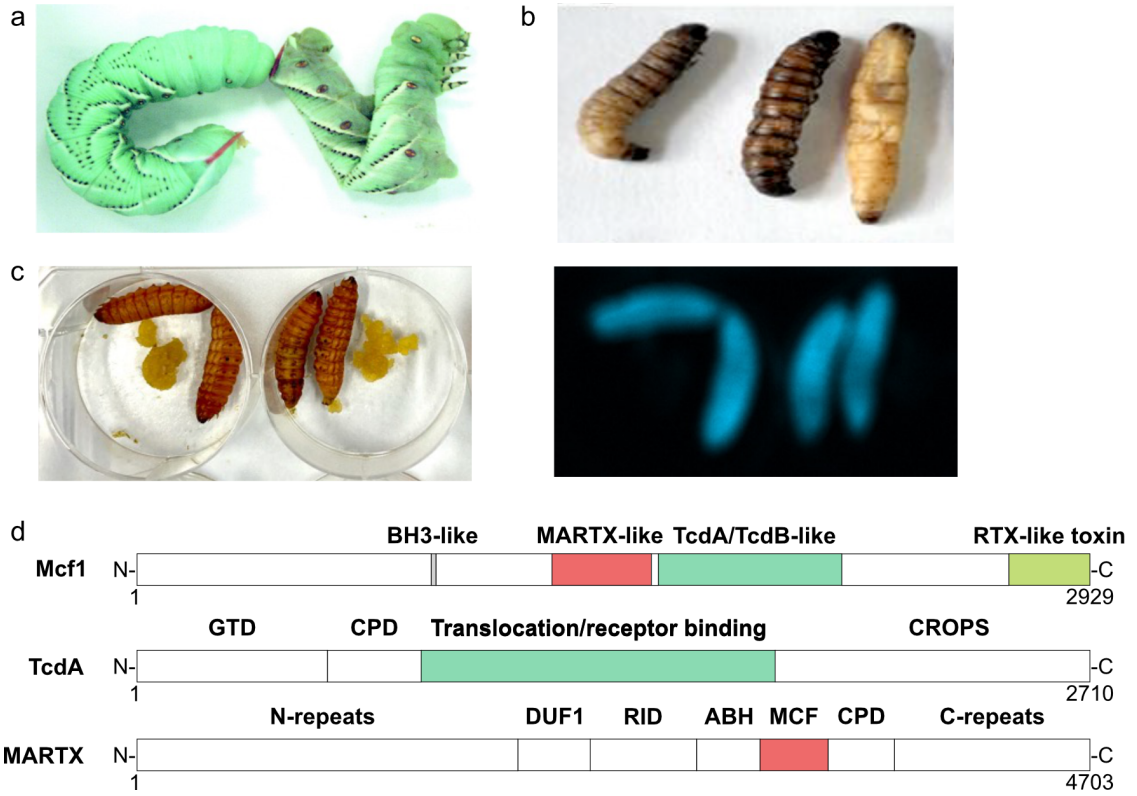


Figure 2 Mcf1 intoxication and sequence homology: **a**) *Manduca sexta* larvae 24h after injection with *E. coli* producing Mcf1 (right) isolated from *P. luminescens* and control (left) [34]. **b**) *Galleria mellonella* larvae 24h after injection with *E. coli* producing Mcf1 isolated from *P. temperata* [41]. **c**) *Galleria mellonella* larvae 24h after injection with *P. luminescens* TT01 at an OD of 1 (left) show luminescence (right, Images provided by Dr. Njenga Ng`Ang`A). **d**) Regions of sequence homology in Mcf1 compared with TcdA and MARTX toxin. Figure d) adapted from [43].

4.4.1 BH3-like domain mutations reduce toxicity

To validate the possible connection between induction of apoptosis and the BH3-like domain, further studies were performed by Dowling et al [44].

In general, the B-cell lymphoma/leukemia 2 (Bcl2) protein family is responsible for the regulation of apoptosis both inhibitory and activating. A conserved key feature of those proteins are the Bcl homology domains (BH1-4). Family members that induce apoptosis contain up to three BH3 domains (e.g. Bad, Bid, Bax, Bak), opening the possibility for Mcf1 to mimic a single BH3 domain containing protein based on their sequence similarity (Fig. 3) [45][46][47][48].

To initiate apoptosis, BH3-only proteins include a single 9-16 amino acid conserved region that functions as an early sensor of apoptosis. They translocate to the mitochondria where they inhibit anti-apoptotic Bcl2 family members by inserting the BH3 domain into a hydrophobic groove consisting of BH1, 2 and 3 domains of the antiapoptotic family

members [49][50]. Furthermore, it was suggested that they are involved in the oligomerization of tBid and Bax which form pores in the outer mitochondrial membrane [51]. However, the detailed mechanism of BH3-only induced pore formation of Bax and Bak remains unclear. Nonetheless, their pore formation leads to the release of cytochrome c into the cytosol, setting of a cascade of apoptotic events [52].

Mcf1 induces apoptosis in mammalian cells accompanied by all the expected markers like DNA laddering, activation of caspase 3, PARP cleavage [33], loss of mitochondrial membrane potential and release of cytochrome c [44]. In addition, cells overexpressing anti-apoptotic Bcl-x_L as well as cells without Bax are resistant to Mcf1 intoxication. Dowling et al. introduced mutations in the BH3-like domain, demonstrating that a double mutant (L915E G919R) reduced caspase 3 activity and the number of cells entering apoptosis. This confirmed a connection between the induction of apoptosis and the BH3-like domain. Notably the double mutant was not able to completely prevent apoptosis [44].

BH3 consensus	**L---LA-VGDGFESR**		**L---LA-VGDGFESR**
P. luminescens	**L**KAG**L**TS**VGDGFE**P**R**	**mHrk**	TALR**L**QAL**GD**ELHRR
P. temperata	**L**KAG**LA**S**VGDGFESR**	**hHrk**	TAAR**L**RAL**GD**ELHQR
P. asymbiotica	**L**KAG**LA**S**VGDRFE**P**R**	**mBak**	VGRQ**L**ALI**GD**DINRR
		hBak	VGRQ**L**AII**GD**DINRR
		mBaxA	LSEC**L**RRIS**D**ELDSN
		hBaxA	LSEC**L**KRI**GD**ELDSN
		mBad	YGRE**L**RRM**GD**EFEGS
		hBad	YGRE**L**RRMS**D**EFVDS

Figure 3 BH3-like domain: Sequence alignment of the BH3-like region of Mcf1 with a BH3 consensus sequence as well as a comparison of the consensus sequence to eukaryotic Hrk, Bak, Bax and Bad. Figure adapted from [44].

4.4.2 MCF like effector from MARTX toxins

The multifunctional-autoprocessing repeats-in-toxins (MARTX) toxin is the main virulence factor of *Vibrio vulnificus,* a gram-negative marine bacterium, that is often consumed with seafood and can cause wound and gastrointestinal infections [53]. The toxin is formed by a single polypeptide chain that combines a cysteine protease domain (CPD) with multiple effectors that are flanked by N- and C-terminal glycine rich repeats, thereby reaching 3500-5300 amino acids in size (Fig. 2d). After binding and translocation into the target cell by an unknown mechanism that involves the glycine rich repeats, the CPD is activated by inositol hexakisphosphate (IP6) and the different effectors are released into the cytosol (Fig. 4) [53][54][55]. The MARTX toxins contain up to five effectors

from a pool of nine that are combined in a modular fashion to form a viable toxin. One of the most abundant effector modules is the Makes Caterpillars Floppy-like effector (MCF) which shows the previously mentioned sequence similarity to the central part of the Mcf toxins [56][57]. This 376 amino acid long effector is positioned either after the Actin Cross-linking Domain (ACD) or the Alpha/Beta Hydrolase (ABH) domain [53]. It stays connected to them even after cleavage of the MARTX toxin by the CPD as their connection lacks a cleavage recognition site. The ACD or ABH are released in a second autoproteolytic cleavage of MCF that occurs inside the cell once the ACD/ABH-MCF fragment encounters ADP ribosylation factors (ARFs) in their active, GTP bound form [38]. These guanosine triphosphatases (GTPases) regulate eukaryotic cell membrane trafficking by cycling between an inactive GDP bound state in the cytosol and an active GTP bound state located on the plasma membrane, Golgi or endosomes [58].

Subsequent to the second cleavage, MCF binds phosphatidylinositol-5-phosphate (PtdIns5P) at membranes and is acetylated at glycine 16. At this point it can encounter its actual target, Rab (Ras-related protein in brain) GTPases. 50% of all cellular Rabs are cleaved or degraded by MCF, which binds them with the same interface that was previously occupied by Arfs. The misdirection of cleaved Rab disrupts their signaling and causes damage of the cellular organelles, finally activating the apoptotic pathway accompanied by a loss of mitochondrial membrane potential and Golgi fragmentation [59]. Interestingly, despite the clear connection between MCFs CPD and the toxic effect, toxicity could only be reduced by mutating amino acids of the Cys-His-Asp (CHD) catalytic triad and was still not completely abolished [60].

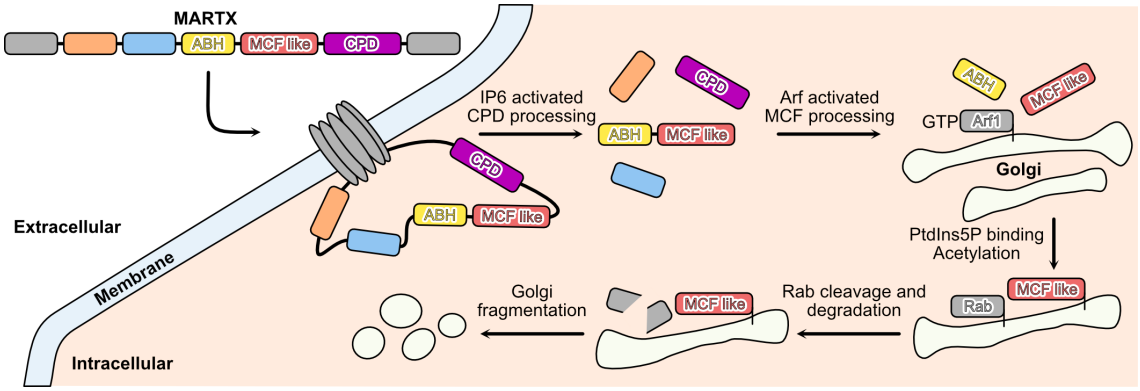


Figure 4 MARTX mechanism of intoxication: MARTX toxin consisting of N- and C-terminal glycine rich repeats (GRR, grey), domain of unknown function (DUF1, orange), Rho GTPase-inactivation domain (RID, blue), Alpha/Beta Hydrolase (ABH, yellow), Makes Caterpillars Floppy like effector (MCF, red) and cysteine protease domain (CPD, violet) is translocated into the target cell by the GRR. The CPD is activated by inositol hexakisphosphate (IP6) and releases the effectors. The ABH-MCF fragment is further cleaved

after activation by GTP-bound Arfs and binds phosphatidylinositol-5-phosphate (PtdIns5P) on the membrane. MCF is acetylated and either cleaves or degrades Rabs causing Golgi fragmentation. Figure adapted from [59].

4.4.3 *Clostridioides diffcile* toxin A and B (TcdA/B)

The sequence analysis of Mcf1 presented similarities to the translocation machinery of *Clostridioides difficile* toxin A (TcdA, 308 kDa) and B (TcdB 270 kDa). Both toxins are members of the large clostridial toxins (LCT) [61] and the major virulence factor of a *C. difficile* infection. The spore forming bacteria is known to colonize the colon after an antibiotic treatment disrupted the gut microbiome and is therefore often encountered in hospital associated infections [62]. It can cause fatal colonic perforation, pseudomembranous colitis and toxic megacolon [63].

The toxins share a similar domain architecture comprising the following domains from N- to C-terminus: A glucosyltransferase domain (GTD), cysteine protease domain (CPD), delivery domain for receptor binding, pore formation and translocation, and lastly a combined repetitive oligopeptide sequences (CROPS) domain [64]. When encountering their target cells, the CROPS domains are suggested to bind glycans on the cell surface while additional cell surface receptors bind the delivery domain (Fig. 5) [65][66].

In case of TcdA, binding was described to the low-density lipoprotein receptor (LDLR), sulfated glycosaminoglycans [67], glycoprotein 96 (gp96) [68] and sucrase-isomaltase (SI) [69], while TcdB was shown to interact with Frizzled (FZD1, 2, 7) [70], chondroitin sulfate proteoglycan 4 (CSPG4) [71] and Nectin 3 [72]. After binding to the cell surface, the toxins undergo endocytosis during which the acidification of the endosomes triggers an unknown conformational change in the delivery domain that leads to pore formation and translocation of the N-terminal CPD and GTD [73].

Once in the cytosol, the CPD is activated by binding of IP6 and releases the GTD with a single proteolytic cleavage. Afterwards the GTD can glycosylate the ‘switch I region’ of Rho GTPases (e.g. Rac1 and Cdc42) with UDP-glucose [74][75] leading to its inactivation and causing actin depolymerization, cell collapse and death [76]. Notably, despite TcdA and B containing a GTD, it was indicated that they target different GTPases [77].

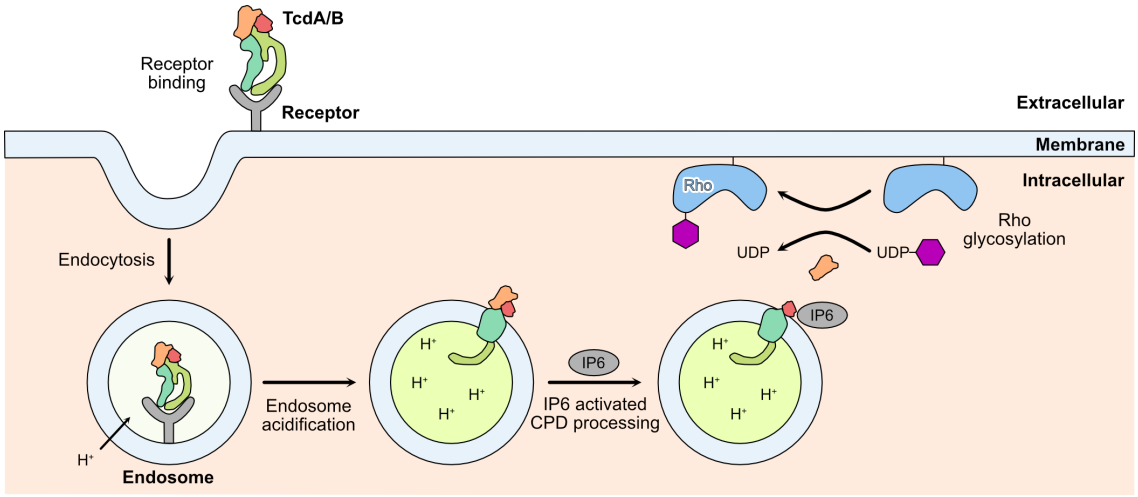


Figure 5 TcdA/B mechanism of intoxication: TcdA/B bind to cell surface receptors (TcdA: glycans, glycosaminoglycans, low-density lipoprotein receptor; TcdB: chondroitin sulfate proteoglycan 4, Nectin 3, Frizzled 1, Frizzled 2, Frizzled 7) and undergo endocytosis. The acidification of the endosome triggers pore formation and translocation of the N-terminal glucosyltransferase domain (GTD, orange) and cysteine protease domain (CPD, red) into the cytosol. The latter one is activated by inositol hexakisphosphate (IP6, grey) and releases the GTD that glycosylates Rho GTPases causing their inactivation. Adapted from [64].

4.4.4 Translocase of Large Clostridial Toxins (LCTs)

The delivery domain of large clostridial toxins, like TcdA/B, is over 100 kDa large and at high pH values consists of an α-helical domain that is entangled with β-sheets, that make up the majority of the domain [78]. These β-sheet-rich regions are also involved in receptor binding as describe above. A recent study revealed the presence of homologs to this translocase in 1104 sequences, of which 769 were outside of the Clostridia genus and part of for example *Pseudomonas* (419 sequences), *Vibrio* (72 sequences) and *Providencia* (67 sequences) [79]. Especially the presence in known pathogens like *Pseudomonas* (419 sequences), *Vibrio* (72 sequences), *Photorhabdus* (44 sequences) and *Xenorhabdus* (36 sequences) indicate a general role of this translocase not only across different genera but also in different environments like gut, soil, wastewater, marine and aquatic. Apart from their similarity, the genes are also located close to secretion systems (type I, III, IV, and VI) and associated to glucosyltransferases (~30%), cysteine proteases (~10%) or both (~20%) suggesting them as possible translocated effectors.

Based on the similarities of the delivery domain, Orrell et al. were able to identify an evolutionary conserved translocase, spanning residues 851-1473 of TcdB, that is sufficient to form pores and translocate N-terminal cargo. A key feature of this region are three

conserved regions sharing a common hydrophobicity pattern and parts of which are predicted to form transmembrane helices (I: 956–1019, II: 1029–1078, II: 1090–1110) [79]. They furthermore contain amino acids essential for a successful translocation [80].

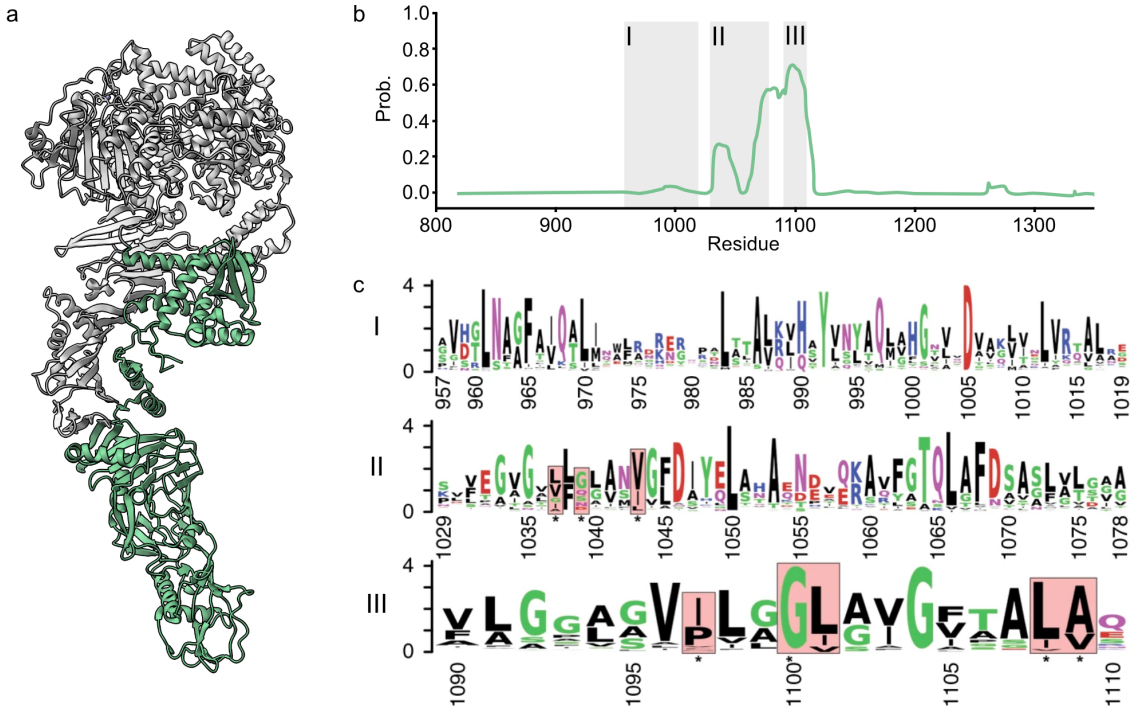


Figure 6 Evolutionary conserved translocase: **a**) Residues 851-1473 of the conserved translocase on TcdA (green, pdb: 4R04 [78]). **b**) Prediction of potential transmembrane helices in the conserved translocase of TcdA with the three regions of sequence conservation (grey). **c**) Consensus sequence of the three conserved regions. Amino acids essential for translocation are boxed in red. Figure adapted from [79].

4.4.5 Repeats in toxins (RTX) toxins

Repeats in toxin (RTX) toxins showed similarity to the C-terminus of Mcf1 and are named after their characteristic feature of glycine and aspartate-rich nonapeptide repeats at the C-terminus of the protein. These repeats create Ca^{2+} binding sites by forming a parallel β-roll, which is a right-handed helix of parallel β-sheets, that binds the Ca^{2+} at its turns with aspartic acid. Due to the lower Ca^{2+} concentration inside the cell (<100 nM) [81] these repeats are unstable and only form outside of the cell, after the RTX toxin was exported through a type I secretion system (TISS). Here, the Ca^{2+} promotes the folding of the functional toxin [82].

Apart from toxins, a wide range of proteins contain RTX repeats including proteases and lipases that assist the virulence factors or bacteriocins that serve as defense for the secreting bacteria [83]. So far over 1000 RTX family members were identified containing <10 to >40 nonapeptide repeats [82].

One example is the RTX toxin (ApxIVA) of *Actinobacillus pleuropneumoniae* that shows C-terminal similarity to Mcf1 and is genetically distant from the other RTX toxins. *A. pleuropneumoniae* is the causative agent of porcine contagious pleuropneumonia, a respiratory disease in swine, that is caused by four different RTX toxins. The presence of ApxIVA was shown to be required to exert full virulence in swine, its mechanism however is unknown [84].

4.5 Mcf2 toxin

The Mcf2 toxin was first described for *P. luminescens* subspecies *akhurstii* (W14), the same strain that led to the discovery of Mcf1, but is also found in subspecies *laumondii* (TT01). It is located in a 7167 bp open reading frame that encodes a 262 kDa large protein. Since the initial discovery of the toxin in 2003 by Waterfield et al. the literature on Mcf2 remained very limited and is mostly restricted to the aforementioned publication [35].

The similarity to Mcf1 results in a familiar homology pattern (Fig. 7 a) with the central region resembling the MCF-like effector of the *V. vulnificus* MARTX toxin (30% identity) and residues 1015-1548 being 20% identical to residues 867-1368 of *C. difficile* TcdBs translocation region.

Differences to Mcf1 are a shorter N-terminus comparable to the HrmA effector (Hypersensitive response and pathogenicity outer protein A) of the plant pathogen *Pseudomonas syringae* (40% identity) and the absence of a BH3 as well as RTX-like domain. An alignment of the Mcf toxins further revealed that apart from the shorter N-terminus, the C-terminus of Mcf2 extends 72 amino acids further. Mcf2 is located in proximity to three genes of the type I secretion system, suggesting its release by this secretion mechanism.

In regard to the function, Mcf2 expressing *E. coli* were able to kill *Manduca sexta* larvae within a similar time and exhibiting a similar floppy morphology as Mcf1. Expressing the toxins in *E.coli* confirmed their size difference but also revealed laddering of Mcf2, hinting at its degradation inside the cells (Fig. 7 b). Furthermore, the first 359 amino acids that resemble HrmA localize to the nucleus in mammalian NIH3T3 cells and cause cell collapse within 24h, suggesting it as the N-terminal effector of Mcf2.

It seems contradicting to have both Mcf1 and Mcf2 present in the same *Photorhabdus* strains with each of them being sufficient to ensure the survival of the bacterium by killing the host. However, there are several possibilities that would explain this phenomenon.

First, the functional redundancy could be a safety mechanism to ensure the death of the insect larvae as the survival of *Photorhabdus* depends on an efficient intoxication. Second, the Mcf toxins could target different insect hosts that might be more susceptible to either of the two. However, an initial study showed expression of both Mcf toxins when exposed to *Galleria mellonella* larvae suspension [85]. Lastly, the point of attack within the host might be different. While Mcf1 is known to target the insect midgut and hemocytes, the mechanism of toxicity of Mcf2 is unknown.

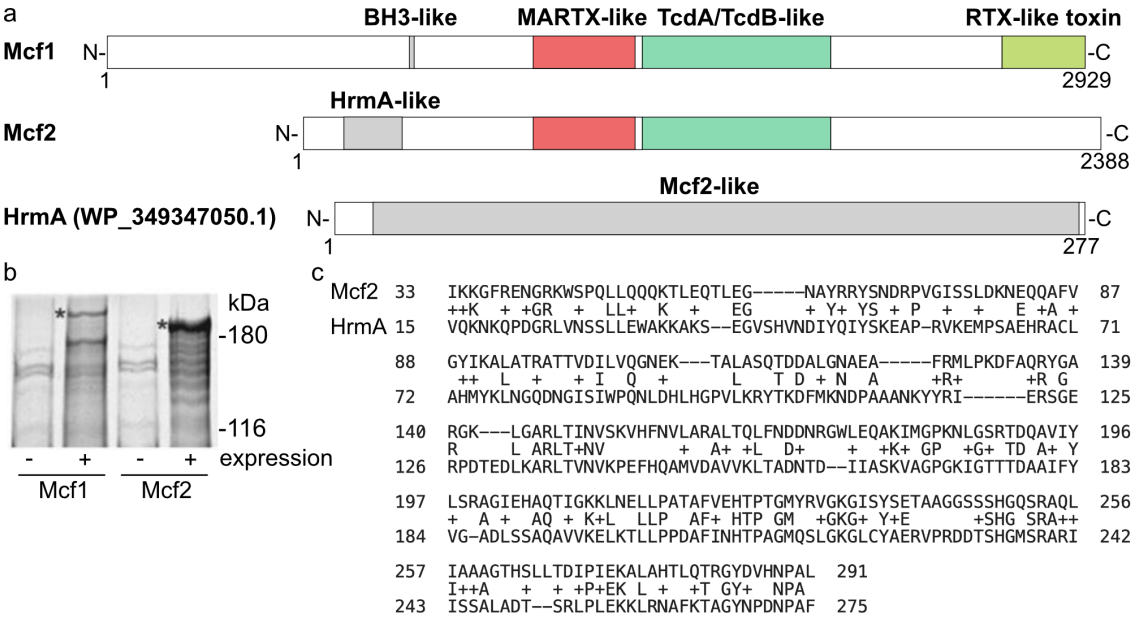


Figure 7 Mcf2 sequence homology and expression: **a**) Mcf2 sequence comparison to Mcf1 shows a comparable architecture of the central region. The N-terminus is shortened and similar to HrmA of the plant pathogen *Pseudomonas syringae*, which it covers nearly completely. The C-terminus is 72 amino acids longer than Mcf1. **b**) SDS-PAGE of Mcf1 and 2 expressing *E. coli* cells. **c**) Sequence alignment of Mcf2 (W14) with HrmA (*Sequence ID: WP_349347050.1*). The HrmA sequence was selected for the largest coverage with an identity of 30%. Figure 7 a) was adapted from [43] and b) from [35].

4.5.1 HrmA (former HopA1)

Pseudomonas syringae is a plant pathogen that can cause different reactions upon infecting a plant. When encountering the host plant, the bacteria colonize the leaf intercellular space and cause necrosis, slowly proliferating and progressing throughout the leaf. In cases where it infects a non-host or a host that acquired resistance, the infection will trigger a defensive hypersensitive response (HR), which is the programmed cell death at the location of infection which prevents further spread of the invader [86].

Both responses depend on hrp/hrc (HR and pathogenicity/HR and conserved) genes that encode a type III secretion system and avr (avirulence) as well as hop (Hrp-dependent

outer protein) genes which contain the effectors that are translocated through this system [87].

HopPsyA (HrmA) is one of these effectors that was identified in *P. syringae* pathovar *syringae* (*Psy* 61) and demonstrated to trigger cell death as part of the hypersensitive response. When compared to a *P. syringae* control that did not express HrmA, it caused faster cell death preventing the progression of the infection over several days ultimately leading to a reduced colonization of the leaf and halted disease progression (Figure 8). Furthermore, expressing HrmA inside tobacco cells also resulted in their death [86].

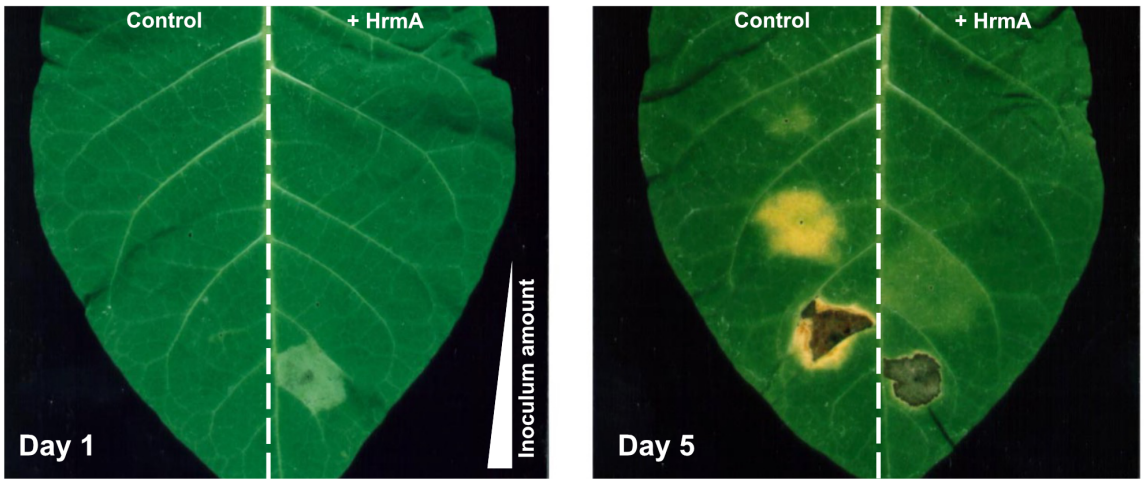


Figure 8: Tobacco leaves intoxication by HrmA: Tobacco leaves were intoxicated with increasing amounts of *P. syringae* either as control (left half of the leave) or expressing HrmA (right half of the leave). The infection with HrmA expressing cells caused hypersensitive response preventing colonization of the leave, while the infection progressed for the control. Figure adapted from [88].

4.6 Electron cryogenic microscopy to solve structures of bacterial toxins

Bacterial toxins are complex molecular machines that adopt a range of different conformations to target cells and deliver a toxic effector across the membrane into the host cell. Structures of these dynamic processes are challenging to resolve. However, electron cryogenic microscopy (cryo-EM) has emerged as a powerful tool that enables us to obtain different states of the toxins and understand their molecular mechanism.

Unlike X-ray crystallography, which usually depends on large amounts of sample and the formation of well-diffracting crystals, cryo-EM captures structures in a more native, hydrated state and excels at solving structures of large proteins and complexes. Due to the fast and preserving nature of the plunge freezing process, where the sample is vitrified in liquid ethane, flexible protein regions and heterogeneity are preserved. The short time

between protein purification and data collection allows it to investigate fast and dynamic processes as well as fragile proteins and complexes, which makes the technique highly attractive to investigate bacterial toxins and their conformational states [89].

These different states and flexible regions can then be identified and analyzed by an increasing amount of processing algorithms like 2D classification [90][91], 3D classification [92], 3D variability analysis [93] and 3D flex refinements [94]. Such dynamic information cannot be obtained from crystals and is crucial for understanding protein movements and formulating molecular mechanisms.

Recent technological advances like the development of direct electron detectors improved the quality of the obtainable reconstructions [95]. They increased the detection quantum efficiency and allowed for a high readout speed that enabled the recording of movies. Those movies can be motion corrected to improve the signal to noise ratio of the micrograph [96]. Furthermore, the increased acquisition speed in combination with automated data acquisition [97] and accessible processing software caused a steep increase in cryo-EM structures [98], culminating in the currently highest resolution reconstruction of mouse apoferritin at 1.22 Å from the year 2020 [99]. In addition, the imaging of small proteins and other samples with poor contrast can be improved by the Volta phase plate (VPP), that increases the phase contrast of the sample and helps to improve the signal to noise ratio of the vitrified samples [100].

Taken together, cryo-EM is a powerful tool that allows the structural study of bacterial toxins throughout their intoxication pathways to understand their function and molecular mechanism.

5. Aims and motivation

The Mcf toxins are key virulence factors of *P. luminescens*, that enable the bacteria to kill the insect larvae and feed off the cadaver. Their ability to also allow *E. coli* to survive in the host, demonstrates their function as autonomous toxic entities that target cells to translocate a so far unknown toxic effector and underlines their potential as tunable molecular machines.

The central aim of my PhD project was to elucidate the mechanism of toxicity of Mcf1 to understand this barely characterized member of the *P. luminescens* toxin repertoire. Insights into the structure and function of Mcf1 would contribute to a broader comprehension of the arsenal of *P. luminescens*, and support the engineering of this toxin for an environmentally sustainable pest control in the future.

Previous studies have shown that Mcf1 induces apoptosis in host cells and it has been proposed that this is caused by a N-terminal BH3-like domain. However, the underlying mechanism remained unclear. To address this, I sought to express Mcf1 in sufficient quality to enable electron cryogenic microscopy (cryo-EM) analysis. Such an approach would allow the assignment of the conserved regions on the structure and enable the exploration of the undescribed regions of Mcf1. By performing a structural homology search I aimed to uncover other functional domains to better understand the modular architecture of the toxin.

In addition to structural studies, I planned to integrate biochemical and microbiological assays to validate the functionality of identified domains with a special interest in locating the actual toxic effector and the components necessary for its targeted delivery. Resolving the molecular organization and domain interactions within Mcf1, the overarching goal was to formulate a plausible and experimentally supported mechanism by which Mcf1 intoxicates target cells.

Beyond the application in pest control, these mechanistic insights may also provide the necessary understanding of clinically relevant pathogens such as *C. difficile* whose TcdA and B toxins employ a similar translocation machinery.

A complementary objective of my work was to investigate Mcf2, a toxin sharing the general architecture of Mcf1 while exhibiting distinct differences. Mcf2 possesses a shorter N-terminal region lacking the BH3-like domain and resembling HrmA. In addition, the

C-terminus is slightly extended. The availability of two closely related toxin variants provided the unique opportunity to distinguish conserved elements from variable regions that may confer functional specialization and investigate their impact on the mechanism with respect to engineering them in the future.

To achieve this, I aimed to determine the structure of Mcf2 by applying the previously established cryo-EM workflow for Mcf1. The following structural analysis of Mcf2 would enable the identification of core components of the intoxication machinery and allow for comparative studies to assess how specific differences influence the function of the toxin.

Ultimately, these insights on Mcf1 and Mcf2 would lay the foundation for rational engineering Mcf toxins, facilitating their optimization as tools for ecological pest control and expanding our understanding of bacterial toxins of *P. luminescens*.

6. Materials

6.1 Instruments

Equipment	Manufacturer
Autoclav LaM-3-20-MCS-J	SANOclav
BIO-RAD C1000 Touch Thermal Cycler	Bio-Rad
Centrifuge 5424	Eppendorf
Centrifuge 5427 R	Eppendorf
Centrifuge Avanti J-15R	Beckman Coulter
Centrifuge Avanti J-26S XP	Beckman Coulter
Column - monoQ	Sigma-Aldrich
Countess II FL Automated Cell Counter	ThermoFisher
Electron microscope Talos Arctica	ThermoFisher Scientific
Electron microscope Titan Krios	ThermoFisher Scientific
Fluorescence microscope EVOS M7000	ThermoFisher Scientific
GE Aekta prime	GE Healthcare
GloQube Plus	Quorum Technologies Ltd
Heat shock bath	TW2 JULABO
Homogenizer	Branson sonifier 450
Incubation shaker Innova 44/44R	New Brunswick Scientific
Incubator (37 °C)	BINDER
Incubator (30 °C)	Thermo Scientific Heraeus
Mini-Protean tetra cell	Bio-Rad
Molecular Imager Gel Doc™ XR System	Bio-Rad
Mono Q 5/50 column	Cytiva
Mono S 5/50 column	Cytiva
pH-Meter FiveEasy FE20-basic	Mettler Toledo
Power Pac 300	Bio-Rad
Shaker (37 °C, Excella E24 Incubator)	New Brunswick Scientific
Shaker poymax 1040	Heidolph
Shaker KS250 basic	IKA Labortechnik
Spectrophotometer DS 11	DeNovix
Thermomixer comfort	Eppendorf

Trans-Blot Turbo Transfer System	Bio-Rad
Vitrobot cryo plunger	ThermoFisher
Vortex-Genie 2	Scienetific Industries

6.2 Chemicals

Chemicals	**Manufacturer**
Acetic acid	PanReac AppliChem
Adenine	Roth
Agarose	Biozym Scientific GmbH
Bafilomycin A1	Sidma-Aldrich
Bromphenolblue	Serva
c0mplete Mini, EDTA-free	Roche
D(+)-Galactose	Roth
D(+)-Glucose monohydrate	Merck
Dithiotreitol (DTT)	Sigma-Aldrich
DMSO	Sigma-Aldrich
EDTA	Roth
Glycerol	Fisher scientific
Glycine	Roth
GTP	Jena Bioscience
HEPES	VWR chemicals
Histidine	Roth
Hydrochloric acid 1 mol/l	VWR chemicals
Imidazole	Sigma-Aldrich
IPTG	Merck
Leucine	Sigma-Aldrich
$MgCl_2$	PanReac AppliChem
Midori green Advanced DNA stain	Nippon genetics
Ni-IDA resin	Macherey Nagel
PEG 3350	Sigma-Aldrich
Peptone	Gibco
Powdered milk	Roth
Sodium chloride	Roth

Sodium dodecyl sulfate (SDS)	Boehringer
Sodium hydroxide (2M)	J.T.Baker Avantor
Tris	Roth
Tryptone	Gibco
Tryptophane	Roth
Tween 20	Roth
Uracil	Sigma-Aldrich
Uranyl formate	Polysciences Inc.
Yeast extract	MP Biomedicals
Yeast nitrogen base	Sigma-Aldrich

6.3 Media and buffer solutions

Media	**Composition**
LB medium (1 l)	10 g bacterial tryptone
	5 g yeast extract
	10 g NaCl

Buffers	**Composition**
100x Amino acid stocks for yeast	
Leucine	400 mg/40 ml
Tryptophane	80 mg/40 ml
Histidine	80 mg/40 ml
Uracil	80 mg/40 ml
Adenine	160 mg/ ml
Elution buffer	20 mM Tris-HCl (pH 8)
	100 mM NaCl
	250 mM imidazole
20x Galactose (for yeast)	40% Galactose in miliQ (sterile)
20x Glucose (for yeast)	40% Glucose in miliQ (sterile)
4x Laemmli sample buffer	250 mM Tris-HCl (pH 8)
	100 mM DTT
	6% SDS
	40% glycerol

	0.02% Bromophenolblue
10x LiAc	1M LiAc
Lysis buffer	20 mM Tris-HCl (pH 8)
	100 mM NaCl
10x SDS running buffer	250 mM Tris-HCl (pH 8)
	1.9 M glycine
	2.8% SDS
TAE (Agarose gels)	40 mM Tris
	20 mM Acetic acid
	1 mM EDTA
TBS tween	20 mM Tris-HCl (pH 8)
	150 mM NaCl
	0.05% Tween-20
TE buffer (pH 8)	10 mM Tris
	1 mM EDTA
TSS buffer (20 ml)	0.3 ml $MgCl_2$ (2M)
	1 ml DMSO
	2 g PEG 3350
	LB medium to 20 ml
YNB yeast nitrogen base	6.7 g/l yeast nitrogen base
YPD yeast peptone dextrose	10 g/l yeast extract
	20 g/l peptone
	20 g/l glucose

6.4 Consumables

Consumable	Manufacturer
10 cm petri-dish, sterile	Sarstedt
C-flat 2/1 grids	Plano
Eppendorf tubes (0.5 ml, 1.5 ml, 2 ml, 5 ml)	Sarstedt
DMEM medium	Gibco
Dialysis bags	SpectrumLabs
Eppendorf tubes (1.5 ml, 2 ml, 5 ml)	Sarstedt
Falcon tubes (15 ml, 50 ml)	Sarstedt

FreeStyle 293 expression medium	Gibco
GeneRuler 1kb plus	ThermoFisher Scientific
PageRuler unstained broad range protein ladder	ThermoFisher Scientific
Parafilm	Pechiney Plastic Packaging
Pipette tips (10 µl, 200 µl, 1.25 ml)	Sarstedt
SDS-Gels Mini PROTEAN TGX Stain-Free Precast Gels (4-15%, 15 wells)	Bio-Rad
Serological pipettes, sterile	Sarstedt
Sf-900 III SFM	Gibco
Spectra HR ladder	ThermoFisher Scientific
Spreaders, T-shaped	VWR
Trans-Blot Turbo Transfer Pack	Bio-Rad
Ultra Clear 24 well crystallization plate	Crystalgen
UltrAuFoil 1.2/1.3	Quantifoil Micro Tools
Well plates (6, 12, 24, 96 wells)	Sarstedt
Western Lightning Plus-ECL	Perkin Elmer

Kitsystems	**Manufacturer**
GeneMorph II Kit	Agilent
QIAprep Spin Miniprep Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen

6.5 Enzymes and antibodies

Enzymes and Antibodies	**Manufacturer**
anti-FLAG M2 #F3165	Sigma-Aldrich
anti-His #H1029	Sigma-Aldrich
anti-Mcf1 (rabbit)	Cambridge research biochemicals
anti-mouse HRP #1706516	Bio-Rad
anti-Myc 9B11 #2276	Cell Signaling Technologies
Anti-pan-ADP-ribose binding reagent MABE1016	Merck

anti-Rab3A #ab3335	Abcam
anti-Rab4 #2167	Cell Signaling Technology
anti-Rab5A, #2143	Cell Signaling Technology
anti-Rab7 #2094	Cell Signaling Technology
anti-Rab11 #3539	Cell Signaling Technology
anti-rabbit HRP #1706515	Bio-Rad
anti-RPS9 serum (rabbit)	gift of Prof. S. Rospert
FastDigest Enzymes	ThermoFisher scientific
Phusion High-Fidelity DNA-Polymerase	ThermoFisher scientific

6.6 Cell lines

Strain	Description
E. coli DH5α	F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^- (Invitrogen)
E. coli BL21 (DE3) CodonPlus RIPL	F^- *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal*λ(DE3) *endA* I [*argU proL* Cam^r] [*argU ileY leuW* Strep/$Spec^r$] (Agilent)
HEK293 GnTI$^-$	Immortalized kidney cell line not expressing N-acetylglucosaminyltransferase I (GnTI)
HEK293T	Derivative of human embryonic kidney 293 cells containing the SV40 T-antigen
S. cerevisiae MH272-3fα	"Wild-type" strain, ura3, leu2, his3, trp1, ade2
Sf9	Immortalized insect cell line from ovarian tissue of *Spodoptera frugiperda* Sf21

6.7 Software

Software	Publisher
Chimera	University of California
Coot	MRC Laboratory of Molecular Biology
cryoSPARC	Structura Biotechnology Inc.

ExPASy	SIB Swiss Institute of Bioinformatics
PHENIX	PHENIX Industrial Consortium
Relion	MRC LMB (Cambridge, UK)
SPHIRE	UTHealth Medical School, MPI Dortmund
trRosetta	Yang Lab

7. Methods

7.1 Molecular biology

7.1.1 Restriction cloning

Most constructs used in this study were obtained by restriction cloning. Where the insert was produced by PCR and ligated into the cut vector.

To obtain the insert, primers were designed including the corresponding restriction sites of the vector and a PCR was performed. 1 µl of template DNA was mixed with 1 µl dNTPs (4 mM), 0.2 µl of each primer (100 µM), 4 µl HF-Phusion buffer (5-fold) and adjusted to 20 µl with sterile miliQ-H_2O before 0,2 µl of Phusion polymerase were added. In case of primers with a high GC content (> ~70%), GC-buffer was used and DMSO (0.6 µl) was added if required. The PCR protocol began with an initial denaturing for 4 min at 98 °C followed by 30 cycles of 15 sec denaturing, 30 sec annealing and extension at 72 °C with 15s/kbp. The final extension took 10 min at 72 °C before the sample was cooled down to 4 °C. Afterwards, the PCR product was analyzed by an agarose gel (1%) and purified from it with the QIAquick Gel Extraction Kit according to the standard protocol.

The purified PCR product and the vector were then cut with the corresponding restriction enzymes (10 µl of DNA, 2 µl FastDigest buffer, 6 µl sterile miliQ-H_2O, 1 µl fast-digest enzyme each, 20 min at 37 °C) and again analyzed and purified from an agarose gel. Finally, the cut vector and insert were ligated with T4 DNA ligase for 1h at RT (5 µl vector, 10 µl insert, 4 µl T4-DNA-ligasebuffer, 1 µl T4 DNA ligase).

7.1.2 Transformation in chemical competent DH5 α *E. coli*

90 µl of chemical competent DH5 α *E. coli* cells were added to the ligation mix and kept on ice for 15 min. Afterwards, transformation was induced with a heat shock in a 42°C water bath for 75 sec, followed by a 1 min incubation on ice. 1 ml of LB media was added and the cells were incubated at 37 °C for 1h. To obtain single, transformed colonies, the cells were spun down in a table top centrifuge at 6000 rpm for 2 min, the supernatant was removed, the cells were streaked out on a pre-warmed LB-Agar plate containing the corresponding antibiotic and incubated at 37 °C overnight.

7.1.3 Plasmid purification

Single colonies were picked and incubated in 6 ml LB medium with the corresponding antibiotic at 37 °C and 180 rpm overnight. On the next day the cells were spun down for 1 min at max rpm and the contained DNA was purified with the Quiagen Miniprep Kit according to the standard protocol with minor adjustments.

7.1.4 Plasmid validation

To confirm the presence of the insert, a sample of the plasmid was cut with restriction enzymes and the fragment size was assessed by an agarose gel. In case the fragments were of the right size, 300-500 ng of plasmid DNA were mixed with 0.2 µl sequencing primer (100 µM), filled up to 15 µl with sterile miliQ-H_2O and send for sequencing at Microsynth Seqlab.

7.1.5 Gibson assembly

In cases where restriction cloning was not feasible a Gibson assembly was performed, where dsDNA fragments with overlapping ends of 20-40 bp are chewed back, annealed, extended, repaired and ligated in a single step.

The overlapping DNA fragments for insert and vector were produced by PCR as described above. Afterwards they were combined to a total vol of 5 µl with an equimolar ratio and mixed on ice with 5 µl of vortexed GA HiFi 1-Step master mix (2-fold). The mixture was spun down and incubated for 1h at 50 °C before transformation into XL10 Gold cells as described above.

7.2 Protein biochemistry

7.2.1 Protein purification

Mcf1

The purified plasmids were transformed into *E. coli* BL21-CodonPlus (DE3)-RIPL cells for expression. A single colony was picked to inoculate LB media (kanamycin 50 µg/ml) and the cells were incubated to an OD of ~1 at 37 °C and 160 rpm. The expression was induced with 20 µM IPTG overnight at 22 °C. On the next day, the cells were spun down

at 4 °C with 4000 rpm to be resuspended lysis buffer containing 20 mM TrisHCl pH 8 and 100 mM NaCl. After lysis by ultrasonication, large cell debris was removed by two rounds of centrifugation (5 and 10 min) in a table top centrifuge at 4 °C with max speed. The soluble fraction was then applied to a column with Ni-IDA resin, washed with 5 column volumes of lysis buffer and eluted with lysis buffer supplemented with 250 mM imidazole. Afterwards the eluate was loaded on a mono Q 5/50 column and further purified with a linear gradient from 100 to 500 mM NaCl in 20 mM TrisHCl pH 8. Fractions were collected according to the chromatogram and analyzed by stain-free SDS-PAGE. The protein was either directly used for cryo-EM sample preparation or stored in aliquots at -20 °C.

Mcf2

Mcf2 was purified in the same way as Mcf1. For the optimized cryo-EM sample 100 mM of HEPES pH 7 was used instead of TrisHCl. Similar to Mcf1, cryo-EM samples were prepared on the day of purification.

Arf3

Arf3 Q71L with an N-terminal deletion of 17 amino acids was purified in the same way as Mcf1. After purification it was supplemented with GTP (1 mM) and $MgCl_2$ (10 mM) and stored in aliquots at -20 °C.

TccC3

TccC3 was purified following the protocol of Mcf1 with the following adjustments. Protein expression was induced with 0.1 mM IPTG and the lysis buffer contained 20 mM TrisHCl pH 8 and 500 mM NaCl. The protein was dialyzed to 20 mM TrisHCl pH 8 and 150 mM NaCl after elution from the Ni-IDA resin.

7.2.2 SDS-PAGE

Semi-denaturing SDS-PAGE with gradient gels (4-15% Mini-PROTEAN TGX stain-free) was used to assess protein quality after purifications and to analyze the cleavage pattern of Mcf toxins.

The protein samples were mixed with 4-fold Laemmli buffer and incubated for 5 min at 95 °C. After a short centrifugation, they were loaded on the gel and a constant voltage of

160V was applied for 43 min. Finally, the gel was imaged in a ChemiDoc Imaging system with the stein-free program.

7.2.3 Western blot

In case a western blot was performed after SDS-PAGE, a colored protein ladder (Spectra HR) was used while running the gel. Afterwards the gel was sandwiched on a PVDF membrane of the Trans-Blot Turbo Transfer Pack and the protein was transferred on the membrane with the Trans-Blot Turbo system at 1.3 A over 7 min. To block the free area of the membrane, it was incubated in TBS-tween supplemented with 5% milk powder while shaking at RT for 1h. The membrane was subsequently washed three times in TBS-tween before applying the primary antibody in TBS-tween overnight at 4 °C while shaking (anti-Mcf 1:10000 rabbit polyclonal antibodies custom-made by Cambridge research biochemicals; anti-Myc 1:2000, 9B11 #2276 Cell Signaling Technologies; anti-FLAG 1:5000, clone M2, #F3165 SigmaAldrich; anti-His 1:3000, #H1029 SigmaAldrich). On the next day the membrane was washed three times with TBS-tween and incubated with the secondary antibody (anti-mouse HRP 1:3000, #1706516, Bio-Rad; anti-rabbit HRP 1:3000, #1706515, Bio-Rad) for 3-4h while shaking at RT. After three final washes, the membrane was imaged using Western Lightning Plus (1 ml) in the ChemiDoc Imaging system with the program blots/high resolution as well as wight light image to annotate the colored protein ladder on the western blot.

7.2.4 Coomassie staining

As a preparation for the mass spectrometry (MS) measurements to identify the Mcf1 cleavage sites, in-house MS measurements were performed, which required Coomassie staining. Therefore, Imperial^{TM} protein stain (thermo scientific) was added to the gels and they were warmed for a few seconds in the microwave. After incubation in the stain for 1h, the staining solution was replaced with a 10% acetic acid, 5% ethanol solution and additional paper towels to accumulate the stain overnight. The gel was imaged on the following day in the ChemiDoc Imaging system.

7.2.5 Mcf1 cleavage assays

Mcf1 (4 µg in 6 µl) in 20 mM TrisHCl pH8 and 100 mM NaCl was mixed with 4 µg Arf3Q71L in 20 mM TrisHCl pH8, 100 mM NaCl supplemented with 10 mM $MgCl_2$ and

1 mM GTP. The addition of $MgCl_2$ and GTP ensures loading of the Arf3 Q71L mutant with GTP, keeping it in its active state that was described to activate MARTX MCF-like toxin[59]. The mixture was incubated at RT for 90 min, while samples were taken at 0, 1, 3, 10, 30 and 90 min. 4-fold Laemmli buffer was added to halt the reaction and the samples were analyzed by SDS-PAGE and western blot (anti-FLAG 1:5000, clone M2, #F3165 SigmaAldrich; anti-His 1:3000, #H1029 SigmaAldrich). The same experiment was performed for the Mcf1ΔC15 mutant with Mcf1C1397A as a control without acquiring western blots.

7.2.6 Mcf2 cleavage assays

The Mcf2 cleavage analysis was carried out as described above with the following alterations. Two constructs of Mcf2 were used, either with a N- or C-terminal Myc-tag. The incubation of Mcf2 with Arf3Q71L took place over a total of 12 h and samples were taken after 40min, 3h and 12h. Furthermore, an additional sample of Mcf2 incubating 12h without Arf3Q71L was acquired to observe the instability of Mcf2. To confirm that the observed cleavage of unactivated Mcf2 is linked to its own protease activity, two samples of Mcf2C791A incubating for 12h with and without Arf3Q71L were taken, which showed no cleavage. For western blot analysis (anti-Mcf 1:10000, anti-Myc 1:2000, 9B11 #2276 Cell Signaling Technologies), anti-Myc antibodies were applied to trace the N- and C-terminus, while anti-Mcf1 antibody was able to also track all tail fragments of Mcf2 that are similar to Mcf1, proving it a useful tool for staining against Mcf2. The increased incubation time for Mcf2 was chosen to increase the fragment amount, making it possible to track them on the SDS-PAGE.

7.2.7 Mcf1 cleavage in *E. coli*

Arf3Q71L and Mcf1 were co-expressed from the same plasmid in *E. Coli* BL21-CodonPlus (DE3)-RIPL cells by inducing them with 20 µM IPTG for 1 or 3h at 21 °C. Afterwards the cells were spun down and lysed with 4-fold Laemmli buffer. Lastly, the lysed samples were analyzed by SD-PAGE and anti-Myc western blot (anti-Myc 1:2000, 9B11 #2276 Cell Signaling Technologies).

The same approach was carried out for the Mcf1L911AK912A mutant and McfC1397A.

7.2.8 Co-expression of Mcf1-Arf3 for MS sample preparation

To obtain high amounts of pure 105 kDa NED fragment, Mcf1 and Arf3Q71L were co-expressed overnight from the same plasmid that was induced by 20 µM IPTG. Afterwards, the cells were lysed in 20 mM HEPES pH 7.5, 100 mM NaCl buffer and loaded on Ni-IDA beads for washing and elution with 250 mM imidazole. The sample was diluted twice with 20 mM HEPES pH 7.5. To separate the NED from the other Mcf1 fragments, the difference in the pI of the NED (8.6) to full size Mcf1 (6.1) was used as a means of separation on a mono S (5/50) column, on which a NaCl gradient was applied. Lastly, the fractions were analyzed by SDS-PAGE and the bands corresponding to the 105 kDa NED were cut out and quadruplicates were sent for LC-MS/MS analysis at Proteome Factory AG.

7.2.9 ADP ribosylation test in yeast

Yeast cells expressing the Mcf1 1-881 fragment were grown on galactose-containing media and mixed with 1 ml of YNB (Yeast nitrogen base) to an OD of 1. They were spun down in a table top centrifuge at max speed for 2 min, their supernatant was removed and the cells were vortexed and incubated in 0.2 M NaOH for 5 min. Afterwards, they were spun down again and the supernatant was removed to add 50 µl of Laemmli buffer. The sample was heated to 95 °C for 5 min and 10 µl were applied to an SDS-PAGE after a short centrifugation. Finally, a western blot was performed with ADP-ribose binding reagent (Anti-pan-ADP-ribose binding reagent, 1:3000, MABE1016, Merck) and the PVDF membrane was imaged in the ChemiDoc Imaging system.

7.2.10 ADP ribosylation test on insect cell lysate

To examine the ADP-ribosylating capabilities of the Mcf1 NED, 3 µg NED or 2 µg TccC3 ART were incubated with 30 µl Sf9 cell lysate (3 mg/ml) in the presence of 16.7 µM biotinylated NAD for 30 min at 30 °C. The samples were mixed with 4-fold Laemmli buffer and applied on a SDS-PAGE for the following western blot analysis with HRP-streptavidin (1:5000, Pierce #21130).

7.2.11 Rab cleavage assay in yeast

Yeast expressing the Mcf1 fragment 1151-1638 was lysed as described above (SECTION) and applied on a SDS-PAGE for western blot analysis against myc, Rab3a, Rab4, Rab5A, Rab7 and Rab11 (anti-Rab3A, #ab3335, Abcam; anti-Rab4, #2167, Cell Signaling Technology; anti-Rab5A, #2143 Cell Signaling Technologies; anti-Rab7, #2094, Cell Signaling Technologies; Rab11, #3539, Cell Signaling Technologies). The Rab antibodies were kindly provided by Nathalie Bleimling.

7.2.12 Rab cleavage assay in vitro

Rab1b was incubated together with Mcf1 and Arf3Q71L as described for the Mcf1 cleavage assays. To ensure the release of the NED and its availability for digesting Rab1b, the easily activated Mcf1ΔC15 variant was included in the experiment.

In case of Rab23, either Mcf1 or Mcf1C1397A as a negative control were incubated with the GTPase. To ensure the activity of Arf3Q71L, an additional sample was supplemented with GTP (1 mM) and $MgCl_2$ (10 mM).

Rab1b was kindly provided by Nathalie Bleimling, while Rab23 was expressed in *E. Coli* BL21-CodonPlus (DE3)-RIPL cells and purified on Ni-IDA resin.

7.2.13 Random mutagenesis

In order to identify amino acids essential for the toxicity of Mcf1 NED, loss of function mutants were screened from a random mutagenesis assay.

The GeneMorph II Kit (Agilent) was used to acquire mutated versions of the NED. After an initial optimization, 0.1 ng DNA were used as matrix for 25 cycles of PCR to obtain on average 15 mutations per construct. Subsequently, the PCR product was cut by BamH1 and Xho1 to be ligated into the YEpGal555 vector cut by the same restriction enzymes. The plasmids were transformed into DH5 α *E. coli* cells and distributed on several LB plates with ampicillin to grow up to 10000 colonies. All colonies were washed off the plate to purify their containing plasmid DNA with several Minipreps, that were afterwards combined and transformed into *S. cerevisiae* MH272-3fα cells. The yeast was grown in YNB agar plates containing galactose (GHLTU) to induce expression of the mutated NED by the Gal1 promotor.

Single yeast colonies were picked and spread out again on GHLTU plates to increase their cell mass over 2 days at 30 °C. The cells were picked and used to validate successful expression of the full size NED by anti-Mcf1 western blot (Section Yeast viability assay). Colonies that did not express the NED and mutations that introduced a stop codon were excluded from the following steps, while the DNA of the remaining colonies was purified from the yeast cells with a miniprep column. Afterwards, every single obtained construct was transformed into DH5 α *E. coli* again to increase the plasmid amount for sanger sequencing (Section Plasmid validation). Finally, all purified and sequenced plasmids were screened in a yeast viability assay (section yeast viability assay) with a clean genetic background and the mutations that led to a loss of toxicity were analyzed.

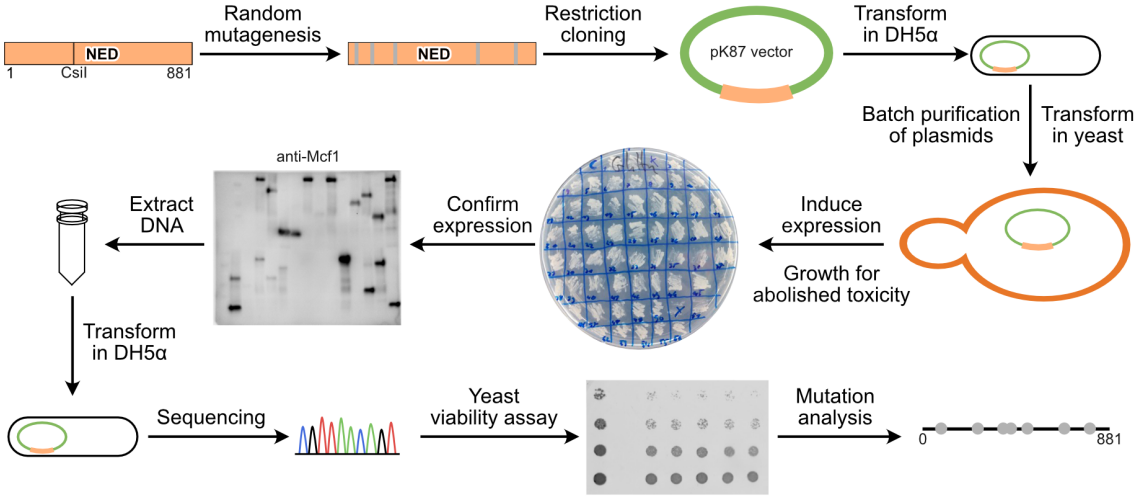


Figure 9 Random mutagenesis scheme: The Mcf1 effector (amino acids 1-881) underwent random mutagenesis and was screened for a loss of toxicity in a yeast viability assay to identify amino acids essential for its function.

Due to the high number of mutations in each construct and the gain of toxicity for all single mutations that were derived from the analysis, the total number of mutations per construct was reduced. To achieve this, the CsiI restriction site within the NED was used to combine a mutated half with a non-mutated one (972 bp/324 aa + 1671 bp/557aa) for all constructs that shared two hits (Fig. 9).

7.3 Analytical

7.3.1 Isothermal titration calorimetry (ITC)

The Mcf1 NED as well as TccC3 from *P. luminescens* were purified as described above in large scale (1l culture). Due to the high expression levels of the proteins, they were

dialyzed against 150 mM NaCl and 20 mM TrisHCl pH 8 buffer right after elution from the Ni-IDA resin. The dialysis was performed overnight from the same buffer stock for both proteins, to eliminate buffer related differences during the measurement, which was carried out with a MicroCal PEAQ-ITC (Malvern Panalytical). TccC3 served as a positive control for binding of nicotinamide dinucleotide (NAD^+). Both proteins were titrated at a concentration of 25 µM with 2.2 mM NAD^+. For data analysis a one-site-binding fit was calculated with the MicroCal PEAQ-ITC Analysis software.

7.3.2 Inductively coupled plasma mass spectrometry ICP-MS

Mcf1 was expressed and purified as described in 7.2.1. After elution from the ion-exchange column, the protein was dialyzed overnight at 4 °C against ICP buffer (20 mM TrisHCl pH 8, 50 mM NaCl) with a reduced NaCl concentration, to prevent the strong Na^+ signal from overlapping with other ions. The same buffer from dialysis was also used as a blank for the ICP-MS measurements.

7.3.3 UV-Vis Spectroscopy

The UV-Vis spectra were measured on a Fluoromax 4 (HORIBA). The emission spectra were measured from 780-900 nm (slit 5) with an excitation wavelength of 764 nm (slit 20) applying undiluted as well as 4-fold diluted samples. The experiment was performed with different batches of Mcf1, Mcf1ΔC15, Mcf1RRR1208-1209-1214AAA, each time showing slight differences for the emission around 820 nm with no clear pattern. Therefore, the batch-to-batch variance was attributed to differences in buffer composition. The excitation spectra was recorded from 700-800 nm (slit 1) while measuring the emission at 820 nm.

7.3.4 Nano differential scanning fluorimetry (nano DSF)

Melting curves were corded on by nanoDSF on a Prometheus Panat (nanoTemper). The protein was spun down at 4 °C and max speed for 5 min in a table top centrifuge to remove protein aggregates. Afterwards the undiluted sample (concentrations from 0.5 – 2 mg/ml) was transferred into the capillaries for measurement according to the standard protocol provided by the manufacturer

7.3.5 Mass photometry

Mass photometry measurements were carried out by the protein quality center with a TwoMP (Refeyn).

7.4 Cell biology

7.4.1 Cultivation of competent DH5 α *E. coli*

A few cells were scraped off a frozen aliquot of competent DH5 α *E. coli* with a sterile inoculation loop and spread on a LB plate without an antibiotic. After growing over night at 37 °C, a single colony was picked to inoculate 200 ml of LB media. The cells were grown to an OD of 0.3-0.5 over 3-4 h and spun down at 4°C and 2700 g for 10 min. Finally, the cells were resuspended in 20 ml TSS buffer and stored in aliquots at -80 °C.

7.4.2 Sf9 cell intoxication – Bafilomycin

Sf9 cells were grown to a density of 1 million/ml in Sf-900 III SFM medium at 27 °C and 120 rpm. Once they reached the right density, they were distributed on a 96 well plate with 100 µl per well and left to settle down for 30 min. Afterwards, the cells were exposed to 100 pM bafilomycin A1 for 15 min before Mcf was added at the given concentrations. All added proteins were sterile filtered (0.2 µm) with a syringe before measuring their concentration and mixing them with the cells. The experiment was performed in triplicates and the cells were kept at 28 °C for 24 h without shaking, before imaging at an EVOS M7000.

7.4.3 HeLa and Hek293 GnTI- cell intoxication

HeLa as well as Hek293 GnTI- cells were grown in 96 well plates to the desired cell density and intoxicated with sterile filtered Mcf toxin (0.2 µm) at the given concentration. The cells were incubated at 37°C with 5% CO_2 for 24h and imaged with an EVOS M7000.

7.4.4 Sf9 cell intoxication – Mcf cleavage

To observe the cleavage of Mcf1 and 2 in vivo, 5 ml of Sf9 cells at a density of 4-5 million/ml were exposed to 5 nM of toxin for 30 min, 1.5 h, and 4 h. The following steps were all carried out at 4 °C. The cells were spun down for 2 min at 1000 rpm and washed

three times with lysis buffer (20 mM TrisHCl pH8, 100 mM NaCl) to remove residual toxin. Afterwards, they were lysed by passing them ten times through a 26-gauge needle and centrifuged for 2 min at 1000 rpm to clear the lysate. Finally, the lysate was mixed 1:1 with 4-fold Laemmli buffer and analyzed by SDS-PAGE and western blot.

7.4.5 Yeast competent cells

S. cerevisiae MH272-3fα cells were thawn from the -80 °C freezer and grown in 2-5 ml yeast peptone dextrose (YPD) media at 175 rpm and 30 °C overnight. On the next day, the cells were diluted to obtain 40 ml with an OD of 0.1 and continued to grow for 5-6 h to an OD of ~0.4. Once the OD was reached, the cells were cooled down on ice and all following steps were also carried out on ice or in cooled systems. The cells were spun down for 5 min at 4.5k rpm at 4°C to remove the supernatant and resuspended with 40 ml cold miliQ. Afterwards, they were spun down again for 2 min at 7k rpm and 4 °C to be washed with transformation mix (800 µl miliQ, 100 µl TE buffer (10-fold), 100 µl LiAc 1M). Finally, 400 µl of the transformation mix were added for every 40 ml of culture with an OD of 0.4 and the cells were stored on ice at 4 °C for transformation.

7.4.6 Yeast transformation

Transformation was performed with the lithium-acetate method (R.D. Gietz 2002). For each transformation, 300 µl of transformation solution were prepared (240 µl 50% PEG, 30 µl 1M LiAc, 30 µl TE buffer 10-fold) and a 50 µl drop of yeast was carefully layered on top. 800-1000 ng of plasmid DNA were injected into the yeast drop before vortexing. Afterwards, the sample was incubated for 30 min at 30 °C while being vortexed shortly every 10 min. This was followed by another incubation of 15 min at 42 °C without vortexing. Finally, the sample was washed three times with 1 ml of YNB medium to remove the PEG. After the last washing step, the supernatant was mostly removed and the cells were plated with the residual media on YNB plates (YNB, 1-2% agar) supplemented with DHLTU (D:glucose (20 g/l), H:histidine (20 µg/ml), L:leucine (100 µg/ml), T:tryptophane (20 µg/ml) and U:uracil (20 µg/ml)). The plates were incubated overnight at 30 °C in a plastic bag to prevent them from drying out. Successful transformation was confirmed on the next day when comparing against a control without plasmid that did not grow.

Lastly, the transformed yeast cells were either used directly for viability assays or a generous amount of cells was resuspended in 1 ml of YPD with 20% glycerol for storage at -80 °C.

7.4.7 Yeast viability assay

Yeast cells transformed with the plasmid of interest were spread out on a YNB DHLTU plate to increase their cell mass overnight at 30 °C. On the next day the dilutions of OD_{600} normalized yeast cells were prepared as follows. For each construct 300 µl of YNB were mixed with enough cells to reach an OD 4. 50 µl of this mixture were subsequently added to 200 µl YNB in a 96 well plate and transferred further to create different dilution steps. Lastly, 3 µl drops of each dilution were placed on a DHLTU and a GHLTU (G:galactose 20g/l) containing YNB agar plate and incubated in a plastic bag at 30°C. The yeast growth was monitored over the following 3-5 days and images were acquired with a ChemiDoc Imaging system [101]. To confirm, that the yeast cells cultivated on the galactose containing media were expressing the protein of interest, 1 ml of these cells (OD 1) was collected in YNB medium. The cells were spun down at max speed for 2 min and resuspended in 200 µl of 0.2 M NaOH by vortexing to disrupt the cell wall. After a 5 min incubation, the cells were spun down again and the supernatant was replaced with 50 µl Laemmli buffer[102]. The samples were heated to 95 °C for 5 min and 10 µl were applied on a SDS-PAGE for analysis by western blot. The applied antibodies and their dilutions are listed below:

Anti-Myc (1:10000, clone 9B11 #2276, Cell Signaling Technology), anti-Mcl1 (1:10000, rabbit polyclonal antibodies custom-made by Cambridge research biochemicals), Anti-pan-ADP-ribose binding reagent (1:3000, MABE1016, Merck) and anti-RPS9 serum (1:10000, polyclonal rabbit antibodies were a generous gift of Prof. S. Rospert); secondary anti-mouse HRP (1:3000, #1706516, Bio-Rad), anti-rabbit HRP (1:3000, #1706515, Bio-Rad)

7.5 Structural biology

7.5.1 X-ray crystallography

Mcf1 was expressed and purified as described previously. Afterwards the protein was concentrated with a concentrator (Amicon Ultra, 0,5 ml, MW cutoff 100 kDa) to a concentration of 8 mg/ml. Based on previous crystallizations of Mcf1 by Dr. Philine Hagel[103], a range of crystallization buffers were prepared varying the PEG 3350 and propanediol concentration for hanging drop vapor diffusion crystallization (Supplementary Fig. 7 b). The reservoir of a pre-greased Ultra Clear 24 well crystallization plate (Crystalgen) was filled with 500 µl of crystallization buffer and a 5 µl drop of protein sample mixed with crystallization buffer (2.5 µl + 2.5 µl) was positioned on a coverslip. Subsequently, the coverslip was placed on the well drop facing down and stored at 20 °C. Crystals were checked for protein contend under UV-light and their growth was monitored regularly until no further growth was observed. At this point, the crystals were fished with a loop and stored at LN temperature. The measurements of X-ray fluorescence were carried out by Dr. Raphale Gasper-Schoenenbruecher at the SLS synchrotron in Villingen, Switzerland.

7.5.2 SPA cryo-EM preparation

All Mcf1 samples were prepared with glow discharged (GloQube, Quorum) C-flat 2/1 400 mesh grids (emsdiasum) operating a Vitrobot Mark IV (Thermo Fisher Scientific) at 100% humidity and 13 °C. The samples were blotted for 3 sec with a blotting force of -3 to +2 without drain time to reduce the preferred orientation of the protein particles and plunged into liquid ethane. After clipping, the grids were screened in a 200 kV Talos Arctica Microscope (Thermo Fisher) with a Falcon III detector (Thermo Fisher Scientific). The general ice quality was judged based on the atlas recoded with the EPU software (Thermo Fisher Scientific) and holes in squares of different ice thickness were chosen for high resolution imaging. To determine the best setup for the following data collection, particle distribution inside the hole was evaluated by acquiring micrographs at the edge as well as in the center of the hole. Finally, grids with satisfying particle quality were stored in auto grid boxes at liquid nitrogen temperature until dataset collection.

Mcf1

Mcf1 was purified on the day of sample preparation. 1.1 mg/ml Mcf1 were mixed with Tween-20 to a final concentration of 0.002% by careful pipetting. The Tween-20 was added to prevent protein aggregation on the grid bars and facilitate an equal particle distribution inside the holes. After mixing, 3 µl of the sample were applied on the grid, which was blotted and plunged as described above.

Mcf1 ΔC15 C1397A

Mcf1 ΔC15 C1397A was also purified on the day of plunging. The cysteine mutant of Mcf1 was chosen to prevent autoproteolysis of the truncated and therefore more reactive mutant. 3 µl of 0.7 mg/ml with a Tween-20 concentration of 0.008% were plunged as described above.

Mcf1 ΔC15 C1397A + Arf3 Q71L

To obtain a structure of the Mcf1-Arf3 complex, the freshly purified, truncated and C1397A mutant of Mcf1 (2 mg/ml) was mixed with a 5-fold excess of Arf3 Q71L. The sample was diluted to a final Mcf1 ΔC15 C1397A concentration of 0.5 mg/ml and mixed with Tween-20 (0.006%) before plunge freezing.

Mcf2

The first batch of Mcf2 samples was prepared following the protocol of Mcf1 grids preparation with freshly purified Mcf2 (0.8 mg/ml) supplemented with 0.005% Tween-20 on C-flat 2/1 400 mesh grids. The initial screening at the 200 kV Talos Arctica Microscope showed a good particle distribution and shape. However, after collecting and processing a larger dataset (Supplementary Fig. 3), Mcf2 was revealed to be fragmented.

To prevent the autoproteolysis of Mcf2 that produces the fragments, the following sample preparations were carried out with freshly purified Mcf2C791A, the cysteine mutant of the PED. In addition, a screen with TritionX-100 as detergent at 0.002% was performed, however no particles could be observed in the hole. Furthermore, a volta phase plate [100] and UltrAuFoil [104] grids (Quantifoil) were used to increase the contrast and better judge the particle quality already during screening. For the UltrAu 1.2/1.3 grids, 0.4 mg/ml Mcf2 supplemented with 0.005% Tween-20 were blotted with a blot force of -3 for 3.5 sec without drain time. The protein concentration was reduced to better distinguish the

particles for screening and particle picking with crYOLO. Despite these changes, a dataset of Mcf2C791A resulted mostly in protein fragments (Supplementary Fig. 4).

In an effort to increase the stability of Mcf2, intensive nanoDSF (nano-differential scanning fluorimetry) and mass photometry measurements were performed, finally resulting in the following preparation protocol (Supplementary Fig. 5). 3.5 µl of Mcf2C791A (0.4 mg/ml) in a buffer of 50 mM NaCl and 100 mM HEPES pH 7 were applied to UltrAu 1.2/1.3 grids in the Vitrobot Mark IV at 13 °C and 100% humidity. After incubation for 30 sec, the drop. Was blotted away by hand to leave only a thin film of liquid on the grid. A new drop of 3.5 µl was applied swiftly and the grid was immediately blotted with a blot force of -3 for 3.5 sec without drain time. After clipping and screening in the Talos Arctica Microscope, a sample with a good ice gradient and particle distribution was stored for data collection.

7.5.3 Data collection

All Mcf1 datasets were recorded on a Titan Krios G3 with a 300 kV X-FEG (Thermo Fisher Scientific) and a K3 direct electron detector (Gatan) at a magnification of 105,000x in super resolution mode (physical pixel size 0.9, super resolution 0.45) with a defocus range from -1 to -2.5 µm. For Mcf2, a Titan Krios G2 with a 300 kV X-FEG, a Cs corrector, a K3 direct electron detector and a post-column energy filter (BioQuantum) was operated at a magnification of 81,000x if not stated otherwise (Supplementary Table 1, 2). The micrographs were recorded using the EPU software (Thermo Fisher scientific) and processed on the fly with TranSPHIRE [97] to perform motion correction with Motioncor2 [96] and CTF estimation with CTFFIND4 [105].

Mcf1

Two datasets were collected from Mcf1 at a Titan Krios with a post-column energy filter (BioQunatum) set to a slit width of 20 eV. The first dataset comprised 7907 micrographs with an exposure time of 3 sec over 60 frames leading to a dose of 71 $e^-/Å^2$. The second dataset had a dose of 69 $e^-/Å^2$ and consisted of 15000 micrographs, of which the last 5955 were acquired at a stage tilt of 30° to reduce the preferred orientation of the particles. Before processing, all micrographs were manually inspected and micrographs with large contaminations, drift or a resolution limit above 6 Å were removed to reduce the amount of data.

Mcf1 ΔC15 C1397A

A single dataset weas recorded on a Titan Krios with a post-column energy filter set to a slit width of 15 eV. The micrographs consisted of 60 frames with an exposure time of 3.5 seconds resulting in a total dose of 60.85 $e^-/Å^2$. During data transfer, all 9032 micrographs were pre-processed with TranSPHIRE, Motioncor2 and CTFFIND4. The manual inspection reduced the micrograph number to 6987.

Mcf1 ΔC15 C1397A + Arf3 Q71L

The Mcf1 ΔC15 C1397A + Arf3 Q71L sample was imaged with a Titan Krios equipped with a C_S corrector and a post-column energy filter with a slit with of 15 eV. 18671 micrographs were recorded with 60 frames and an exposure time of 2 sec leading to a total dose of 62.6 $e^-/Å^2$. Again, the data was pre-processed with TranSPHIRE, Motioncor2, CTFFIND4 and manually cleaned resulting in 15693 final micrographs.

Mcf2

All Mcf2 datasets were collected with a 300 kV Titan Krios with an in-column C_S corrector, a K3 direct electron detector and a post-column energy filter. The pre-processing was either carried out on the fly with TranSPHIRE or after acquisition with CryoSPARC [106].

7.5.4 Data processing

Mcf1

The Mcf1 model is based on a composite map originating from the Mcf1 and Mcf1 ΔC15 dataset. This was necessary since the initial reconstruction from the Mcf1 dataset was missing the NED which was better resolved in the ΔC15 variant (Supplementary Fig. 1).

After manual cleaning of the first dataset, Mcf1 particles were picked on 10 micrographs by hand to train a picking model with crYOLO [107]. Subsequently, this model was employed to pick and extract 787549 particles with a box size of 300 pixels. However, due to the elongated shape of the particles, many picks were off-centered. To re-center them, an initial 3D refinement was performed in MERIDIEN [108] (SPHIRE) coarsely shifting all particles (search range 25, step range 5 pixels). The alignment parameters were used to re-extract now properly centered particles with a box size of 320 pixels.

Bad particles were removed by 2D classification with the iterative stable alignment and clustering approach ISAC (SPHIRE) [91] without pre-alignment and a translation search range of 0 as the particles were already aligned for re-extraction. The remaining 363044 particles resulted in a 4.3 Å reconstruction of Mcf1, which was further improved by three iterative rounds of polishing in RELION 3.1 [109] and 3D refinement in MERIDIEN, resulting in a 3.6 Å reconstruction of Mcf1.

The local resolution of the map showed flexibility of the tip of Mcf1, which made it difficult to trace the peptide backbone. Signal subtraction was applied to this region, which shifted the center of alignment closer to the tip and thereby improved the density. Both the initial 3.6 Å map as well as the signal subtracted version were used for model building.

In addition to the tip, the local resolution also worsened towards the head of Mcf1, which was further emphasized by its weaker density which only appeared at lower visualization thresholds. One possible reason for this is heterogeneity in the density, which can be sorted for in 3D to select only the particles containing the full head density. In order to perform such a 3D classification, the particles were imported into RELION 3.1 and classified without image alignment as the projection parameters were also included from the previous reconstruction. To make sure, that only changes in the head region of Mcf1 were considered for classification, a soft mask and a low-pass filtered reconstruction were provided. The 3D classification revealed that a majority of particles (87%) did not contain the NED and only 47522 particles (13%) showed an additional density. This demonstrates that the head and especially the NED are more flexible than the tail of Mcf1. In addition, the dataset showed a preferred orientation which could be the result from the Mcf1 head attaching to the air-water interface, which would also impede its reconstruction.

To obtain more particles with the full head density, a larger second dataset was recorded, which included 5955 micrographs that were acquired with a 30° tilt in order to counteract the preferred orientation. The same processing scheme as for the first dataset was applied and led to a 4.6 Å reconstruction after 2D classification. Although it still suffered from preferred orientation, the anisotropy improved which was evident by the full head density being present already at this step of the processing pipeline. Since the reconstruction of the tail from the first dataset was sufficient for model building, the following processing focused only on the head density.

Again, 3D classification was performed in RELION with a mask covering the head. This time 219318 particles contained the full head. This increase from 13% to 30% of full-head particles is likely the result of tilting the stage. Combining the 3D classes of both datasets led to 266840 particles that were used to reconstruct the head region.

After a first reconstruction of the combined particles, signal subtraction was performed to focus the refinement on the head of Mcf1. To further sort out possible heterogeneity of the NED, another 3D classification was performed, this time using a mask solely focusing on the flexible NED. The resulting 2 classes with the most complete NED density were combined (222562 particles) for a 3D refinement and subsequent signal subtraction which yielded the final reconstruction of the head with a resolution of 4.1 Å. Although the overall resolution was worse than full-size Mcf1, the completeness of the density allowed modelbuilding of the full head, except for a central density in the NED. The reconstruction of Mcf1 ΔC15 C1397A was used later on to fill this gap.

Mcf1 ΔC15 C1397A

The dataset of Mcf1 ΔC15 C1397A was recorded as a negative control to properly evaluate the changes that occur upon binding of Arf3 (Supplementary Fig. 21). 2202784 particles were picked from 6987 micrographs with crYOLO and extracted in RELION 3.1 with a pixel size of 1.8 Å and a box size of 150 pixels. Binning the particles to a larger pixel size reduces the achievable resolution but drastically speeds up the initial steps of processing as the load of data decreases. For further processing the particles were imported into CryoSPARC, where three independent 2D classifications were performed with different number of classes (150, 175, 200) to sample the data at different levels. For each 2D classification an ab initio model was reconstructed and the most complete one consisting of 544954 particles was used for non-uniform 3D refinement leading to a 3.7 Å reconstruction of the head. To further enhance the resolution the particles were unbinned to a pixel size of 0.9 and reextracted. Two rounds of particle polishing in RELION 3.1 with non-uniform 3D refinements led to the final local refinement focused on the NED and PED with a resolution of 3.1 Å.

To obtain a structure that also includes the ABD, a crYOLO model was trained on the particles that led to the previous reconstruction. Picking with this model gave rise to 2181255 particles that underwent a similar processing pipeline as before. The particles were extracted binned to a pixel size of 1.8 Å and extracted with a box size of 150 pixels.

After the transfer to CryoSPARC, 2D classifications were performed to remove bad particle which resulted in 904685 particles and a map from non-uniform 3D refinement which still displayed a weaker density for the ABD. This decreased density is caused by the absence of the ABD in some of the particles that went into the refinement. To identify only the particles that contain the additional ABD density, an alignment free 3D classification into ten classes was performed. Class 7 showed the most pronounced ABD density and the containing 82866 particles were unbinned for two cycles of polishing and non-uniform 3D refinement. Afterwards, the particles were sorted one last time by alignment free 3D classification to select 53952 particles, which finally gave rise to a 3.46 Å map containing the NED, PED and ABD. The central section of the NED of this map was used to create the composite map of full size Mcf1.

Mcf1 ΔC15 C1397A + Arf3 Q71L

After cleaning the micrographs, 1491851 particles were picked with crYOLO, binned to 1.8 Å pixel size and extracted with a box size of 150 pixels for further processing in CryoSPARC (Supplementary Fig. 22). Similar to the Mcf1ΔC15C1397A pipeline, three 2D classifications were performed with different final numbers of classes (100, 150, 200) to create ab initio models. The best model contained 770049 particles that were used for a non-uniform 3D refinement resulting in a 4.3 Å map. To identify particles that contained an additional density compared to the Mcf1ΔC15C1397A reconstruction, alignment free 3D classification into ten classes was performed, which revealed sever classes with a clear extra density. Of those, the largest class (101942 particles) was selected for non-uniform 3D refinement. To further push the resolution, the particles were unbinned and recentered for another non-uniform 3D refinement which allowed particle polishing. The final non-uniform 3D refinement led to a reconstruction of 4.0 Å.

Mcf2

The processing of all Mcf2 datasets was performed in CryoSPARC (Supplementary Fig. 3, 4, 6). For the first Mcf2 dataset, the crYOLO picking model of Mcf1 was used to pick 4057158 particles, that were binned to a pixel size of 1.76 Å and extracted with a box size of 288 pixels. The following 2D classification was used to remove bad particles and select 34113 particles that could have originated from full size Mcf2 with different orientations. Afterwards four ab initio reconstructions were created to select the most complete maps of Mcf2 which corresponded to 16494 particles. The following non-uniform refinement

generated a reconstruction of only 8.8 Å which is far away from the Nyquist limit of 3.5 Å. This is likely a result of the poor quality and low number of particles. In an attempt to increase the number of healthy particles, a new crYOLO model was trained on particles from the best ab initio model. A subsequent 2D classification showed no improvement in particle quality.

Mcf2 C781A

The second Mcf2 dataset was recorded on the Mcf2C791A mutant that lacks the catalytic cysteine of the NED which should reduce the number of fragmented particles (Supplementary Fig. 4). Again, the Mcf1 picking model was applied to judge the sample quality. This time, however two different confidence thresholds were tested to allow the picking of less similar particles and improve the model later on by training on the wanted particles. After extracting the particles with a pixel size of 1.76 Å and a box size of 288 pixels, however, in both cases the 2D classification revealed particles either representing highly tilted projections or fragments of Mcf2, which was further highlighted by the calculated ab initio reconstructions. Following the hypothesis that complete Mcf2 particles must be present in the data based on the mutation of the Mcf2 protease domain, extensive training of crYOLO picking models was performed. Micrographs with good contrast and particle distribution were chosen to pick particles with a different strictness of selection. For the strictest model, only particles were chosen that resembled full size Mcf2. For the less strict selection, particles which could relate to tilted orientations of Mcf2 were also included. The least strict model was trained by picking all particles in the micrograph independent of their shape. Corresponding to the training strategy, different numbers of particles were picked by crYOLO. However, subsequent 2D classification showed that none of the models resulted in complete Mcf2 particles.

Mcf2 C781A double blotting

After manual cleaning of the micrographs, 347971 particles were picked with crYOLO based on the Mcf1 picking model and extracted with a pixel size of 1.76 Å and a box size of 192 pixels (Supplementary Fig. 6). These are 10 times more particles than for the first Mcf2 dataset indicating, that the changes in sample preparation increased the dataset quality. The following 2D classification revealed a large amount of Mcf2 like particles that were subjected to an ab initio reconstruction to create three different classes. After improving their resolution with a heterogeneous refinement, the most complete first two

classes (257784 particles) were combined to perform a non-uniform refinement which resulted in a 4.6 Å resolution map of full size Mcf2. At this point the resolution did not improve further even when unbinning the particles, performing per particle CTF refinement, and cleaning the dataset further by removing overrepresented views or 2D classes with a lower estimated resolution.

Therefore, it was decided to increase the number and orientation of particles that crYOLO might have missed due to an imperfect initial model by retraining on particles of the latest refinement, but also by creating a template for the template picker. The latter option was chosen based on the assumption that certain views were missed by using the Mcf1 model for picking and therefore they are also not included in the latest refinement and not included in the retraining data of crYOLO. A newly created template, however will sample all projections of this refinement and give access to previously unknown particles with the disadvantage of being less precise. In the following only the most successful strategy will be described.

As expected, the template picker increased the particle amount drastically to 24486624 which were all extracted at a pixel size of 1.76 Å and a box size of 192 pixels. During 2D classification, only obvious artifacts were removed to prevent a biased selection by the user and instead a processing pipeline was chosen that relies more on sorting in 3D. To achieve this, all particles were subjected to a heterogeneous refinement with a proper Mcf2 and three junk templates. This reduced the particle amount to 4227733 which were subsequently cleaned further by 2D classification and subjected to a non-uniform refinement leading to a reconstruction of 3.74 Å consisting of 1136182 particles. An additional round of 2D classification followed by reextraction of the particles further improved the resolution to 3.6 Å. At this point the map presented features similar to Mcf1. The tip of the tail was not fully resolved and the head appeared very flexible. Therefore, the same approach of signal subtraction was used to focus the refinement on these challenging areas. The head reached a resolution of 5.7 Å and the tail refined to 3.4 Å. Lastly the map interpretability was improved by DeepEMhancer.

7.5.4.1 3D variability analysis

A comparison of the head to Mcf1 as well as the fitting of the model revealed the absence of the NED of Mcf2. As the 2D classes and maps at lower thresholds indicated the pres-

ence of a flexible region this was further investigated. From the binned non-uniform refinement (3.7Å, Supplementary Fig. 6 c) the particles were down sampled to a box size of 96 pixels and subjected to a 3D variability analysis split into ten classes. The analysis was focused on the head region with a wide and soft mask to capture all possible movements of the NED. The classes depict a wide range of NED movement and a broad distribution of particles that could not be refined to a better resolution when using the single classes of the 3D variability analysis or classes from a similar 3D classification.

7.5.5 Model building

Mcf1

The Mcf1 model is based on the initial reconstruction at 3.6 Å, its signal subtraction with an improved quality of the tip (3.6 Å), the focused reconstruction of the head based on the combined datasets at 4.1 Å and the Mcf1ΔC15C1397A reconstruction at a resolution of 3.46 Å. The final map depicted in the figures is the composite map that was further polished with DeepEMhancer [110].

To obtain a starting point for model building, the map_to_model function of Phenix [111] was used to assign three helices of the TD1 and the long helix TD2. Based on these anchor points and with the help of a secondary structure prediction of Mcf1 by ExPASy [112], the model was manually built in coot following the peptide backbone. Smaller domains and flexible loops were predicted in trRosetta [113][114] as Alphafold [115] was not yet available at that time. The predictions were then positioned into the map by rigid body fit in Chimera and further refined manually in Coot. Lastly, the model was further refined in Isolde and Phenix to improve model parameters such as the Clashscore and remove Ramachandran outliers.

Mcf1 ΔC15 C1397A

The head section of the previous Mcf1 model was aligned with the map by a rigid body fit in UCSF Chimera [116]. The missing regions of the NED and PED were built by hand in Coot [117] and further refined in Isolde [118], trRosetta and Phenix.

Mcf1 ΔC15 C1397A + Arf3 Q71L

The model was created by performing a rigid body fit of the Mcf1ΔC15C1397A model and Arf3 (pdb 6II6 [38]) into the refinement. To further improve the fit, flexible fitting in

the iMODFIT Chimera plugin [119] was applied. Finally, the model was optimized in Isolde and Phenix.

Mcf2

Due to the poor quality of the reconstruction, no detailed modelbuilding based on side chain information was possible and in certain regions the peptide backbone could also not be traced. To still allow for a general interpretation of the map, alphafold predictions of single domains were rigid body fitted into the density in ChimeraX [120].

8. Results and discussion

All work related to Mcf2 was carried out by P. Heilen as well as the data processing, model building and structure analysis of Mcf1. Furthermore, the Mcf intoxication assays, all experiments regarding the identification of the HSE and the random mutagenesis and in-depth analysis of Mcf1 NED were also performed by P. Heilen. The optimization of the purification protocol and plunging conditions of Mcf1, the cleavage analysis and formulation of the Arf3 activation mechanism was done together with Dr. A. Belyy, who also processed the datasets of Mcf1ΔC15 and the Mcf1ΔC15-Arf3 complex.

8.1 Mcf1 architecture

Mcf1 had previously proven to be challenging to resolve by Dr. P. Hagel as it demonstrated an inherent flexibility and preferred orientation [103], which limited the achievable resolution (4.8 Å). However, it was still possible to describe Mcf1 as an elongated protein consisting of several domains. In addition, mutational studies suggested the presence of a cysteine protease domain and an analysis of the hydrophobicity pattern of the TcdA/TcdB like region revealed similarities to *C. difficile* TcdB.

A possible reason for the limited resolution could be that the purification took nearly 2 days which could have further curtailed the protein quality. In order to ensure the best protein quality possible, the purification and plunging conditions were optimized, shortening the time from cell lysis to cryo-EM screening to about 9 hours. To overcome the preferred orientation, we recorded an additional dataset including micrographs at a stage tilt of 30°. These changes in combination with an upgraded direct electron detector from K2 to K3 (Gatan Inc.) and an intensive processing pipeline were key to obtaining the structure of Mcf1 at 3.6 Å resolution (Supplementary Table 1, 3; Supplementary Fig. 1).

8.1.1 Mcf toxin is a member of the ABCD subfamily

The structure revealed Mcf1 to be an elongated, seahorse-like protein that can be divided into two distinct structural parts, a tail and a head (Fig. 10 a-c). While the head consists of the N-terminal effector domain (NED), activator binding domain (ABD) and protease effector domain (PED), the tail comprises the translocation domains 1 and 2 (TD1 and 2) that are linked by the transmembrane helix (TH) and the putative receptor binding domains 1-3 (RBD1-3). To assess whether this separation into head and tail correlates with

functional differences, we compared the structures of the single domains with the protein database (pdb) using the DALI server [121] (Fig. 10 c, d, Supplementary Fig. 2).

Starting at the N-terminal effector domain (NED), a central region of 114 out of 912 amino acids showed structural similarities to bacterial ADP-ribosyltransferases RhsP2 (7RT7 [122]) and Rhs1 (7ZHM [123]) from *Pseudomonas aeruginosa* and *Salmonella enterica* respectively, suggesting a role as toxic effector. Next, the activator binding domain (ABD), which is connected to the NED by an unresolved linker of 14 amino acids, showed high similarities to the deamidase *Burkholderia* lethal factor 1 (Blf1) (3TU8 [124]), including it too as a potential effector.

The following protease effector domain (PED) on the other hand is similar to cysteine protease domains (CPDs) of TcdA (7POG [125]) and TcdB (7V1N [126]) of *Clostridioides difficile* as well as the MCF-like domain of the multifunctional-autoprocessing repeats-in-toxin (MARTX) toxin from *Vibrio vulnificus* [38], confirming the sequence similarity on a structural level. In case of TcdA and TcdB, the CPD is responsible for the release of the N-terminal effector inside the host cell [61], while the CPD of the MCF-like domain was shown to induce apoptosis by cleaving Rab proteins [59], providing two possible functions for the PED.

As the last domain of the head of Mcf1, the PED is linked to the translocation domains 1 and 2 (TD1 and 2) of the tail, that are interconnected by a linker predicted to form a transmembrane helix (TH, section 8.3). Both TD1/2 are structurally similar to the translocation apparatus of the Tcd toxins [79], indicating that they employ a similar machinery to translocate N-terminal positioned effectors through the membrane of the host cell, again confirming the previously described sequence similarity.

Lastly, the Mcf1 tail folds back in on itself and forms the putative receptor binding domains 1-3 (RBD1-3), which primarily consist of β-sheets. The RBD3 and the previous TD2 represent a fold similar to the β-roll of RTX toxins. However, they are missing the characteristic calcium binding sites [82][127]. In addition, RBD3 is once again similar to a part of TcdA and B with a proposed function as a receptor binding domain.

Due to the intricate fold of the tail that protects the translocation domains one and two by folding the putative receptor binding domains around them, the C-terminus of Mcf1 finds itself beneath the head, only 2 nm away from the N-terminus.

Taken together, the homologies of the Mcf1 domains with the aforementioned toxins allow us to assign Mcf1 as a member of the ABCD bacterial toxin subfamily, where the essential functions of the toxin are distributed as modules across the protein. These functions are assigned as follows: A, biological activity (NED, ABD); B, binding (ABD 1-3); C, cutting (PED); D, delivery (TD1, TH, TD2).

To investigate whether this organization also holds true for Mcf2, we performed a sequence comparison of both toxins (Fig. 10 e). It revealed the NED as the only major difference between the two proteins, while the rest showed a high alignment score with an overall identity of 77 %. Such a high level of conservation strongly suggests that Mcf2 is organized in the same way and allows us to allocate both Mcf toxins as part of the ABCD bacterial toxin subfamily.

With the previously described functions in mind, we can draw the conclusion of a possible mechanism of Mcf toxins, where the tail is responsible for targeting of the host cells and translocating the head through the membrane, while the head harbors the toxic effector(s), as well as the machinery necessary for its release inside the host cytosol.

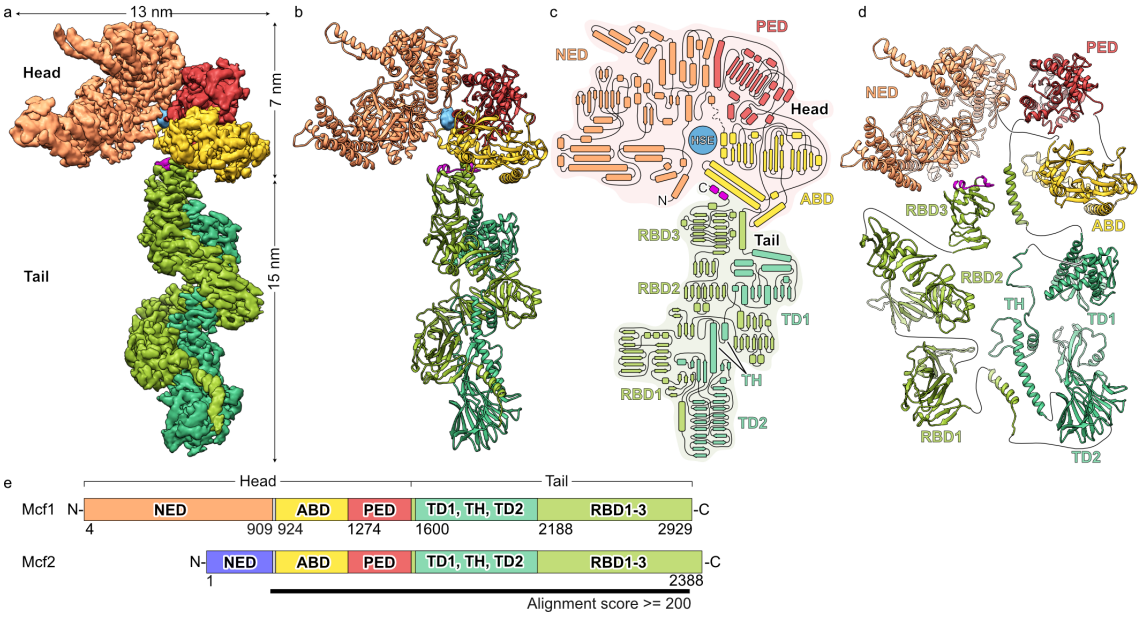


Figure 10 Structure of the Mcf1 toxin: **a**) Composite cryo-EM density separated in head and tail region and colored according to different domains (orange, N-terminal effector domain NED; yellow, activator binding domain ABD; red, protease effector domain PED; dark green, translocation domains 1,2 TD1,2 and transmembrane helix TH; light green, putative receptor binding domains 1-3 RBD1-3; blue, head stabilizing element HSE). **b**) Atomic model of Mcf1 with a scheme of the secondary structure (**c**) and separated single domains (**d**). **e**) Diagram showing the locations of the domains on Mcf1 and aligning them with Mcf2 based on sequence similarity. Figure adapted from [43].

8.2 Towards the structure of Mcf2

To evaluate whether this mechanism can be applied generally to both Mcf toxins, a detailed structure of Mcf2 is needed. This would enable us to confirm the presence of the previously identified key elements in Mcf2 and assess whether changes to them are altering the activation.

Initially, we prepared the Mcf2 sample in the same manner as Mcf1, which resulted in micrographs displaying Mcf like particles. However, it is difficult to assess Mcf particles purely by the micrographs as its elongated shape results in vastly different projections, depending on its orientation. Therefore, the true quality of a vitrified sample can only be assessed after particle picking and 3D reconstruction. The initial reconstructions confirmed the presence of Mcf2 particles, but also revealed the presence of Mcf2 fragments mostly consisting of the tail (Supplementary Fig. 3 b). Even when re-training the picking model [107] on only complete particles, the number of particles and resolution could not be improved from 16500 particles refining to a 9 Å reconstruction.

Based on the hypothesis that Mcf2 instability is likely caused by its own PED cleaving the protein, we exchanged the catalytic cysteine (C791A) of the PED. Furthermore, we also screened the grids with a Volta phase plate (VPP) to better assess the sample quality based on the particles inside the micrographs with an increased contrast (Supplementary Fig. 4 c). Again, despite extensive training of different particle picking models in crYOLO, mostly fragmented particles could be identified or a strong preferred orientation prevented the identification of full-size particles (Supplementary Fig. 4 b).

Although one can apply a vast range of processing tools during single particle analysis to sort and clean the data (2D classification, 3D classification, re-training picking models, 3D variability analysis), a good dataset ultimately still relies on sample quality with stable particles oriented differently in the vitreous ice layer. As this seemed not to be the case, we investigated the properties of the Mcf2 sample applied on the cryo-EM grid to further optimize the sample preparation.

When we incubated the Mcf2C791A mutant together with Tween-20, we observed no fragmentation on an SDS-PAGE and only a small shift in melting temperature (Supplementary Fig. 5 a, b). This changed however once we performed mass photometry experiments, revealing the presence of two fragments (257 kDa and 62 kDa) in the Tween-20

containing sample. Originally, we added Tween-20 to prevent the Mcf particles from preferentially binding on the grid film, instead of localizing inside the electron transparent holes. One option to avoid the use of detergent and still get particles into the holes is by double blotting [128]. For this method, the protein sample is applied twice on the grid, the first time to saturate the preferential binding to the film and the second time to position particles into the holes. In addition, we performed a buffer screen and increased the stability of the sample by optimizing the melting temperature from 40 °C to 49 °C in a buffer of 100 mM HEPES pH 7 (Supplementary Fig. 5 c).

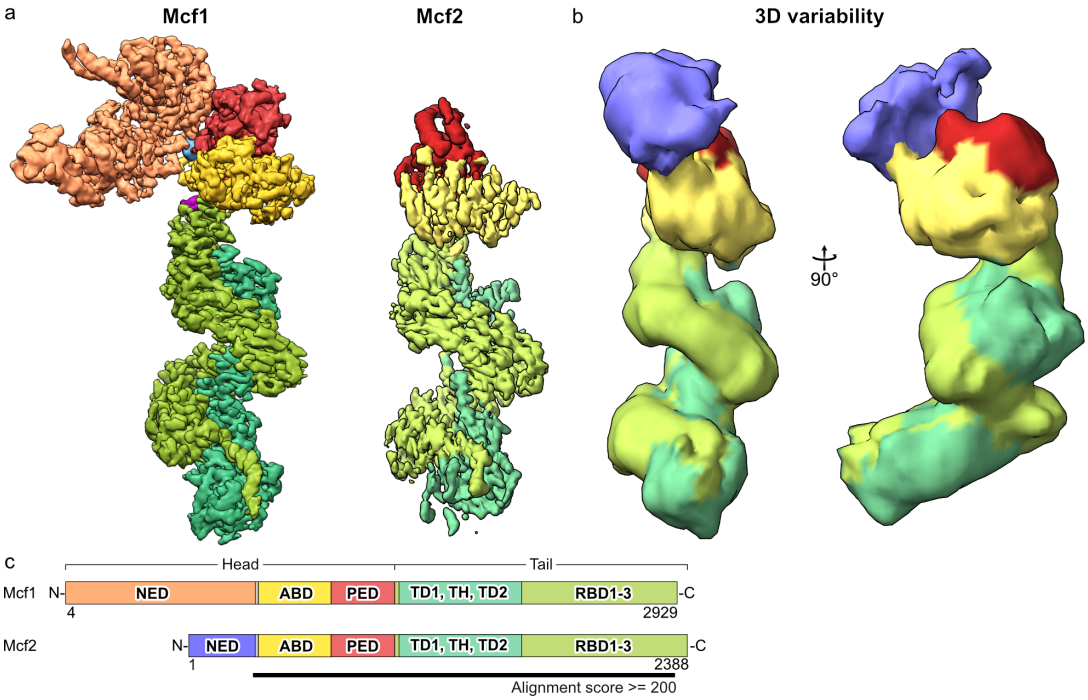


Figure 11 Structure of Mcf2: **a**) Comparison of the composite cryo-EM density of Mcf1 and the density of Mcf2 colored according to domains (orange, N-terminal effector domain NED; yellow, activator binding domain ABD; red, protease effector domain PED; dark green, translocation domains 1,2 TD1,2 and transmembrane helix TH; light green, putative receptor binding domains 1-3 RBD1-3; blue, head stabilizing element HSE). **b**) Density resulting from a 3D variability analysis of the Mcf2 dataset demonstrating the position and flexibility of the NED of Mcf2 (violet). **c**) Diagram of the Mcf1 and Mcf2 architecture. Figure adapted from [43].

Based on these modifications we prepared new samples by double botting Mcf2C791A in 50 mM NaCl and 100 mM HEPES pH 7. This procedure resulted in a dataset of 18600 micrographs with more complete particles resembling Mcf2 (Supplementary Fig. 5 d, 6) and a structure could be determined to a resolution of 4.6 Å. This was further improved to a reconstruction at 3.6 Å (Fig. 11) by including particles in all possible orientations and sorting them in an unbiased manner in 3D. However, due to preferred orientation and inherent flexibility of Mcf2, the resolution was heterogeneous leading to interrupted helices and β-sheets merging together which prevented the tracing of the peptide backbone.

Similar to Mcf1, we applied signal subtraction focused on the head or tail region to further refine the maps, which finally yielded in a 3.7 Å reconstruction of the head and 3.4 Å reconstruction of the tail (Fig. 12). These reconstructions allowed for model fitting, but lacked resolution in many places for detailed sidechain information and suffered from anisotropy as well as distinct densities for constructing β-sheets (Supplementary Fig. 7). This made tracing the backbone challenging.

One reason for the strong artifacts in the reconstruction could be the double blotting by hand during sample preparation, which is hard to properly control and reproduce. To improve the data, Mcf2 could be crosslinked by glutaraldehyde [129] to reduce the inherent flexibility of the protein and prevent cleavage in presence of Tween-20, which would allow us to use the same plunging conditions as for Mcf1.

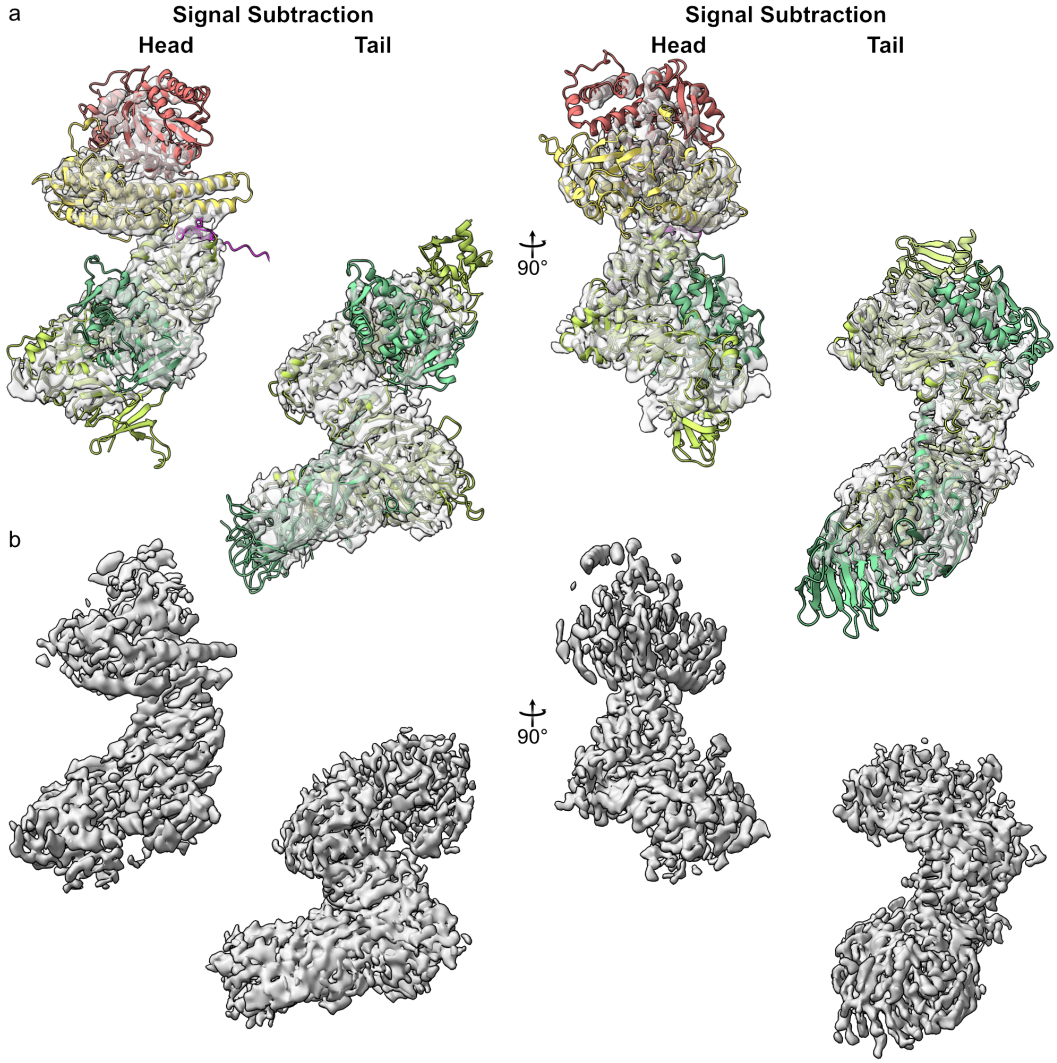


Figure 12 Signal subtraction of Mcf2: **a**) Mcf2 head and tail signal subtracted EM-densities with a rigid body fit of the single alphafold3 predicted domains as well as the densities alone (**b**).

8.2.1 Mcf2 has similar architecture with different effector

The general architecture of Mcf2 is similar to Mcf1 and comprises an elongated protein separated into a head and a tail region (Fig. 11). However, for Mcf2 the effector and parts of the PED are neither resolved in the full reconstruction nor in the signal subtractions (Fig. 11 a, 12). As the 2D classes already showed a high degree of flexibility of the head and a fuzzy density which might correspond to the effector (Supplementary Fig. 8 b), we performed 3D variability analysis to resolve the different states of Mcf2 (Fig. 11 b, Supplementary Fig. 6 c). Indeed, a density of the effector appeared at the head of Mcf2. In addition to its high degree of flexibility, the effector is not positioned at the same interface as the NED of Mcf1 but shifted by 90° (Supplementary Fig. 8 a). This could be caused by its smaller size of 37 kDa which reduces the area of the interface between NED and PED and weakens the interaction thereby allowing for more movement of the NED. In addition, the small NED of Mcf2 is also not stabilized by the HSE, which is missing in the reconstruction. The repositioning of the NED of Mcf2 and the increased flexibility of the head might be the reason for its instability and explain the difficulties in obtaining intact particles. A detailed analysis of the differences to Mcf1 is however hindered by the poor local resolution of the reconstruction. Nonetheless, alpahfold predictions of the domains can be fitted into the density, which confirms that Mcf2 has the same general architecture as Mcf1 and therefore likely employs the same mechanism.

8.3 The tail region reveals Mcf as part of the LCT superfamily

Given the accumulation of similarities between Mcf1 and TcdA/B, we performed an in-depth comparison focused on the tail of Mcf1.

The Mcf toxins and TcdA/B resemble each other not only in terms of general architecture (Fig. 13 a) and structural similarities of the CPD, TD1, TD2 and RBD3 (Supplementary Fig. 2), but also on a sequential level. A recent analysis of the translocation machinery of large clostridial toxins (LCTs) found over 1100 homologs, containing a conserved machinery [79]. One key feature of this evolutionarily conserved translocase is a fragment of three sequence-conserved regions, whose consensus sequence was derived from all members of the LCT family. When comparing this consensus sequence with Mcf1 and 2 we found a striking resemblance that is even more prominent than the one to TcdA (Fig.

13 b). In addition, the last two sequence-conserved regions are located on the linker between TD1 and TD2. This linker is predicted to form transmembrane helices [130] (Fig. 13 c) and we therefore named it transmembrane helix (TH) (Supplementary Fig. 2).

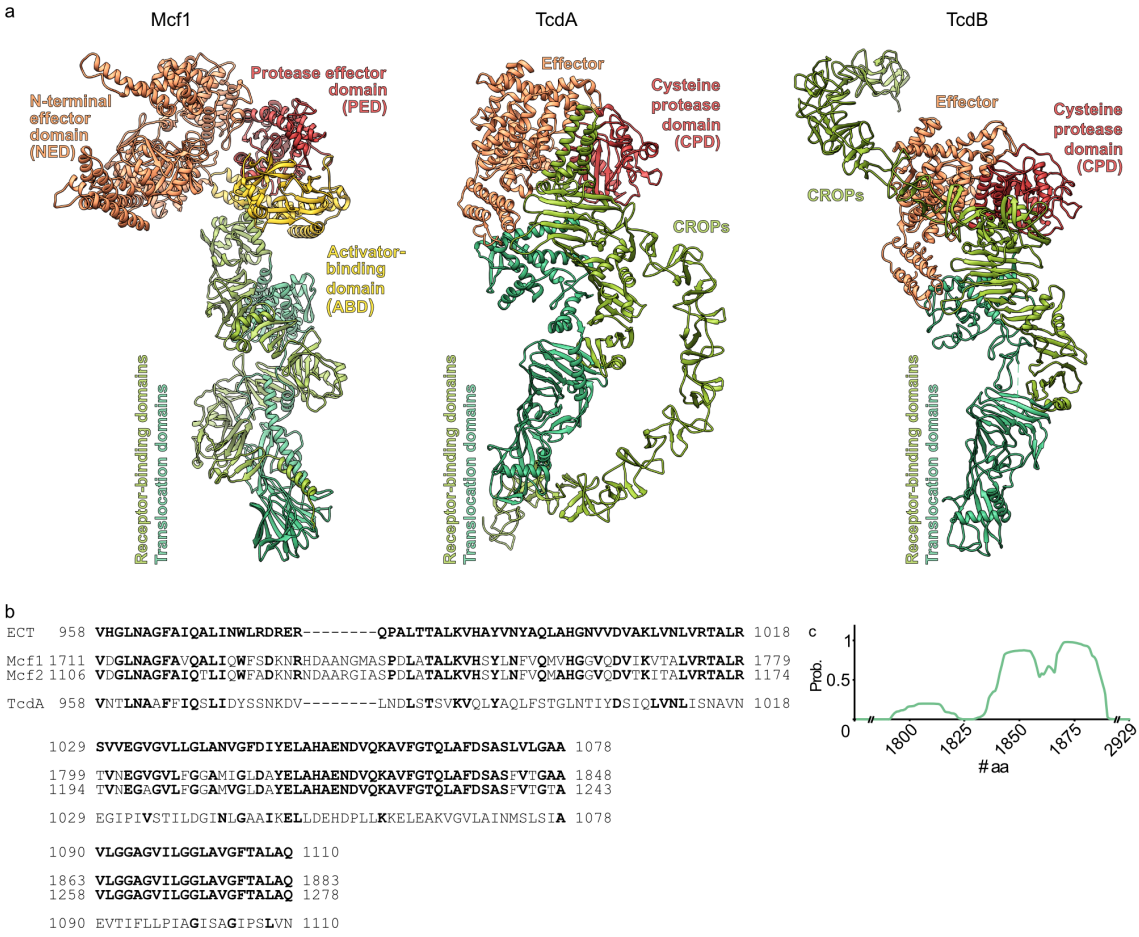


Figure 13 Comparison of Mcf1 with TcdA/B of *C. difficile*: **a**) Comparison of the structure and domain organization of Mcf1 with TcdA (7POG [125]) and TcdB (7V1N [126]). **b**) Sequence alignment of the three conserved regions of the LCT translocase with Mcf1 and Mcf2 as well as the comparison to TcdA. **c**) Probability for transmembrane helices as predicted by TMHMM 2.0 [130]. Figure adapted from [43].

Following these findings, we wanted to investigate whether the translocation of Mcf toxins is carried out in the same manner as the LCTs. In their case, the toxin binds to the host cell surface and undergoes endocytosis. Subsequently, the acidification of the endosomes triggers a yet unknown conformational change in the translocase to transport the N-terminal part of the toxin through the membrane into the host cell cytosol. A common means to block translocation is via the inhibition of the V-ATPase, a proton pump responsible for endosome acidification, with bafilomycin A1 [131], which we used to validate the translocation mechanism of the Mcf toxins.

Intoxication of Sf9 cells with Mcf1 leads to mortality within 24h. The dead cells appear as masses of vesicles and cell debris (Fig. 14 a). However, in the presence of bafilomycin,

the cells survive and only show slight membrane deformations caused by the inhibition of endocytosis and the lysosome. The same is true for Mcf2 (Fig. 14 b), albeit with a reduced initial toxicity in the absence of bafilomycin. Overall, these observations identify the Mcf toxins as members of the LCT superfamily of bacterial toxins.

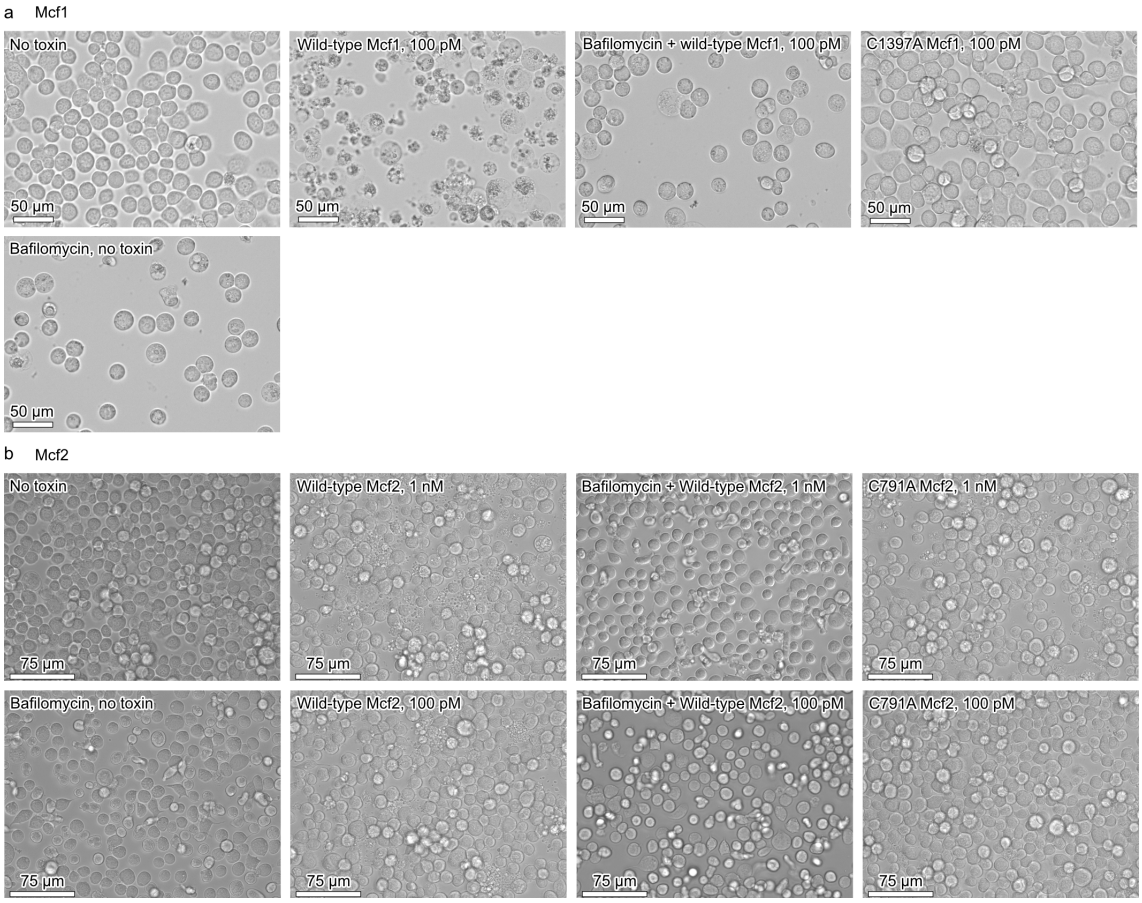


Figure 14 Bafilomycin inhibits Mcf toxicity: Sf9 cell intoxication with either Mcf1 (**a**) or Mcf2 (**b**) at the indicated concentrations causes cell death. In presence of bafilomycin (100 pM) or when intoxicating with a protease deficient mutant (Mcf1 C1397A, Mcf2 C791A), the cells are able to survive. Figure a) adapted from [43].

Based on the hypothesis that the Mcf toxins and TcdA are translocated by a similar mechanism, we compared their translocation domains in further detail (Fig. 15 a). Strikingly, the orientation of TD1 and 2 towards each other is different between Mcf1 and TcdA. In case of Mcf1, the TD2 is rotated with respect to the TD1 when compared to the translocase of TcdA. This can be observed when superimposing either of the two domains which results in the remaining translocase not aligning (Fig. 15 b). The varying orientation is caused by different conformations of the TH domain that connects the TDs and is further influenced by the interacting RBD1-3.

These two distinct conformations of the translocase demonstrate that translocation can occur from two different initial states of the translocation machinery and further highlights the major conformational rearrangements likely happening during translocation. They furthermore hint at a possible role of the TH domain as a key regulator, that could cause rearrangements of the translocase upon membrane insertion leading to formation of a pore. This is also reflected in the AF3 prediction of Mcf2, where the TH domain adopts an additional exposed conformation, resulting in the TD2 to be dislodged from the other domains of the tail (Fig. 15 a). In general, the versatility and flexibility of the LCT translocase could be a reason why its detailed mechanism remains unknown. Nonetheless, solving this mystery is highly important due to the presence of the translocase in pathogenic bacteria such as *Clostridia* and *Vibrio*.

Possible strategies to tackle this challenge should closely follow the process of intoxication, particularly the toxin-receptor interaction prior to endocytosis, which has proven difficult to imitate using in vitro reconstitutions. This was also shown by Dr. A. Belyy who attempted to purify Mcf1 from endosomes or recreate a pore by acidification of Mcf1 in presence of liposomes or detergents. Despite the difficulties, resolving structures along this trajectory could give insights into the changes that prime the translocase for pore-formation and help to monitor the transition in a controlled manner for cryo-EM analysis.

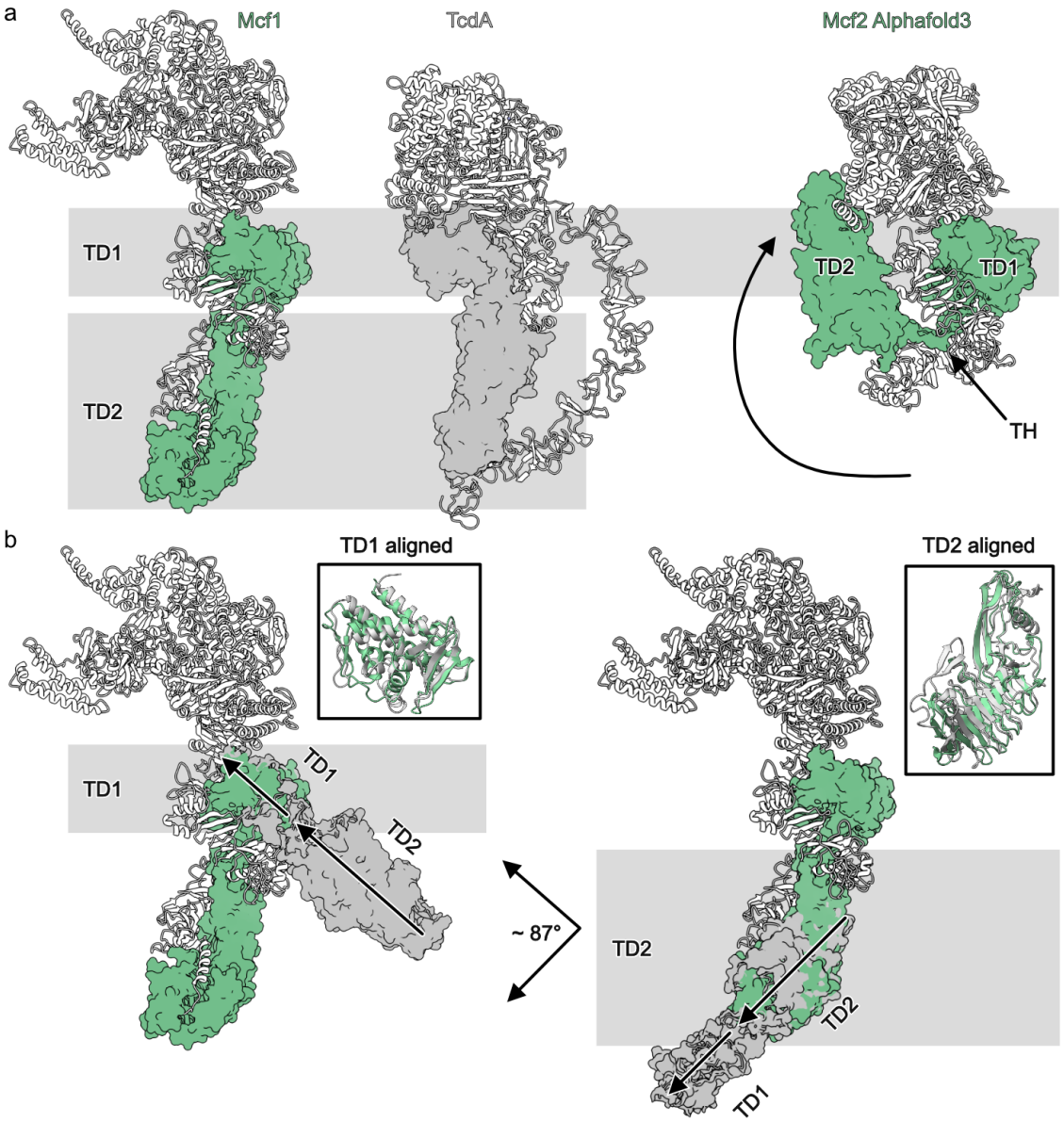


Figure 15 Different initial states of the LCT translocase of Mcf toxins and TcdA: **a**) Surface representation of the translocase (TD1, TH, TD2) of Mcf1, Mcf2 Alphafold3 prediction (green) and TcdA (grey, pdb 7POG [125]). The reorientation of the TD2 from Mcf1 to the Mcf2 prediction is indicated by an arrow. **b**) Superimposition of the TD1 (left) or TD2 (right) of Mcf1 and TcdA reveals a ~87° rotation of the TD2 in the Mcf1 structure.

8.4 Receptor binding of Mcf toxin

Independent of the translocation mechanism, the initial step to ensure that Mcf1 can target the right cells and undergo endocytosis is the binding to the cell surface. This could be either directly on the membrane, or to proteins- and/or glycan-receptors on the surface. Based on the similar architecture of Mcf1 and TcdA/B as well as domain homology, the putative RBDs 1-3 are likely responsible for the selectivity of the toxin. An analysis of the electrostatic Coulomb potential on the surface of Mcf1 showed no apparent charged

patches, that would orient or bind Mcf1 to the membrane in a preferred manner (Suppelementary Fig. 9). In addition, we were able to intoxicate HeLa as well as Hek293 GnTI- cells, which lack complex N-linked glycans on their surface (Supplementary Fig. 10). This led us to the conclusion that binding does not depend on complex N-linked glycans and indicated that Mcf1 likely binds to a receptor present on both cell types.

Furthermore, apart from the different NED, varying affinity to the same receptors or binding to completely different receptors is a possible explanation for the divergent toxicity of Mcf1 and 2 on Sf9 cells (Fig. 14 a, b). To determine whether these changes in receptor binding are at the structural level, we highlighted amino acids that are different in Mcf2 using the structure of Mcf1 as template. However, there were no apparent regions with accumulated changes in the RBDs on the Mcf1 structure as well as on the Mcf2 AF3 prediction (Supplementary Fig. 11), indicating that the sequence differences if present are more subtle or might be induced upon receptor binding. For a more detailed analysis and comparison of the RBD surface, a sidechain quality model of Mcf2 would be necessary.

8.4.1 Mcf receptor CRISPR knock-out screen

To obtain candidates for the cell surface receptors that the Mcf toxins bind with their putative RBDs, we sent out protein samples to Creative Biogene for CRISPR library screening (Human Genome-wide CRISPR Lentiviral Library Cat.# CCLV0001). During their initial tests on HEK293T cells, they confirmed our intoxication experiments that demonstrated Mcf1 is more potent than Mcf2 (Fig. 14 a, b; Supplementary Fig. 12 a). By using positive selection, they identified genes that conveyed resistance to the toxins. The most meaningful hits are discussed below.

In general, the genes from the Mcf1 screening had a higher score (Supplementary Table 4, 5, Supplementary Fig. 12 b), which is probably related to its increased toxicity that allowed for a more sensitive readout. For Mcf1, two hits are a V-ATPase subunit [132] (ATP6V0C) and its assembly factor [133] (CCDC115), which validated the dependence of Mcf intoxication on the acidification of endosomes. Another hit, EMC6 is located in the endoplasmic reticulum and assists ER membrane insertion of new membrane proteins as part of the endoplasmic reticulum membrane protein complex (EMC) [134][135], hinting towards the binding of Mcf1 to a membrane protein. STT3A on the other hand is a catalytic subunit of the N-oligosaccharyltransferase (OST) complex [136][137], that cat-

alyzes the initial step of N-glycosylation, which is the transfer of a lipid linked oligosaccharide to the asparagine of the nascent protein containing the Asn-X-Ser/Thr motif, with X not being proline [138]. This suggests that Mcf1 intoxication depends on the presence of N-linked glycans that are attached to this motif and not further modified by the N-acetylglucosaminyltransferase I (GnTI), as Hek293 GnTI deficient cells were still susceptible to Mcf1.

Interestingly, the V-ATPase subunits do not appear in the Mcf2 top 10 hits, which might indicate that Mcf2 is less dependent on the acidification of the endosomes or depends on different factors inside the endosomes to trigger its translocation. To confirm whether the hits directly interact with Mcf and how they fit into the intoxication mechanism, detailed binding studies need to be carried out either with the purified proteins or on the knock out cell lines themselves. This could also give insights into the steps that follow the initial binding of Mcf toxins on the cell surface.

8.5 Mcf1 head region is supported by C-terminus and head stabilizing element (HSE)

The C-terminus of Mcf1 is not only a possible receptor binder as part of the RBD3, but also interacts intimately with the head region. Several amino acids form hydrophobic interactions with a helix bundle of the ABD and the backbone helix connecting the head with the tail (Fig. 16 a, c), collectively arranging the neck region of the toxin. In light of the previously discussed rearrangements upon translocation (8.3), this network of interactions will probably be lost during poreformation, which suggests an important regulatory role for the very C-terminus.

This hypothesis is further supported by a comparison to Mcf2, where most of the hydrophobic amino acids responsible for the interaction network are conserved (Fig. 16 a-c). Interestingly, the C-terminus of Mcf2 continues 72 amino acids further, which might also cause differences in toxicity (Fig. 14) and affect the stability. To gain more insight into this suggested regulatory element, the structure of Mcf2 would be needed at better quality, as a clear assignment of the C-terminal amino acids is not possible (Supplementary Fig. 13) and the AF3 prediction leaves the last 72 amino acids unstructured (Supplementary Fig. 14), which is expected as it was probably trained on Mcf1.

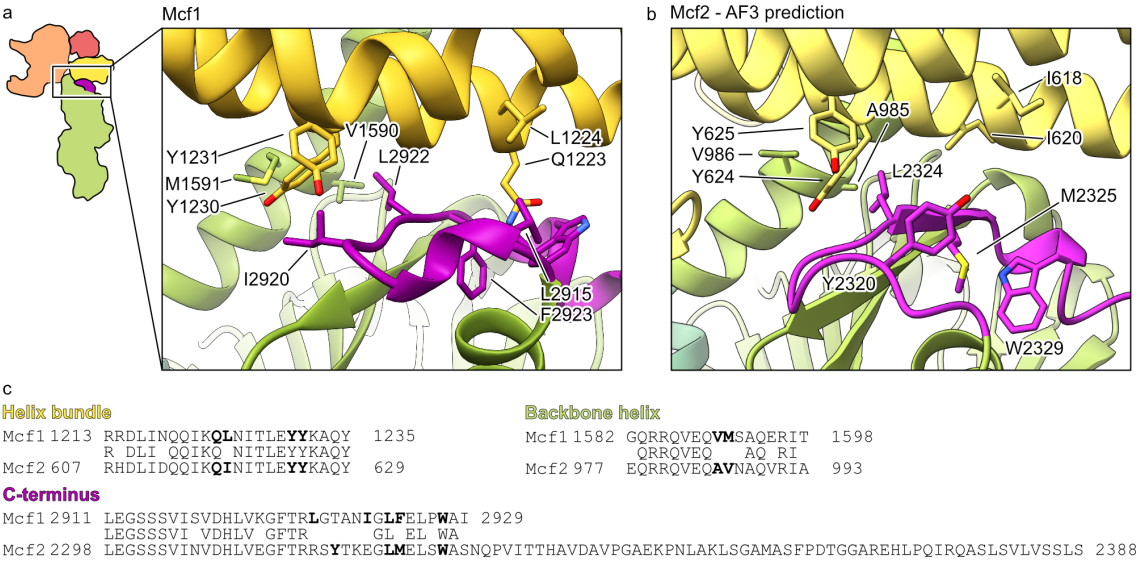


Figure 16 C-terminus binds head with hydrophobic interface: Atomic model of Mcf1 (**a**) and alphafold 3 prediction of Mcf2 (**b**) with a focus on the interaction between C-terminus and helix bundle of the ABD, highlighting amino acids that form hydrophobic interactions. **c**) Sequence alignment of the hydrophobic interaction network of Mcf1 and Mcf2. Figure a) adapted from [43].

In addition to the head-tail interaction network, the head is also internally stabilized at the interface of the N-terminal effector domain (NED) to the Protease effector domain (PED) and the activator binding domain (ABD). The upper section interacts by matching electrostatic interactions between NED and PED (Fig. 17 b), while the lower interface of all three domains is organized around a spherical density in the cryo-EM map (Fig. 17 c), henceforth called head stabilizing element (HSE). The HSE is surrounded by five hydrophilic arginine residues and fragments into three smaller volumes at a higher map threshold (Fig. 17 c-e). Their presence at such high thresholds, where the surrounding amino acid density is no longer visible, indicates the presence of atoms that strongly interact with the electron beam like metals or phosphate. Such a stabilizing element is not present in Mcf2 as the NED is likely not large enough to form the coordination environment of the HSE which contributes to the structural changes observed for the Mcf2 head.

In an effort to characterize the HSE, we performed inductively coupled plasma mass spectrometry (ICP-MS) measurements at the ISAS with the help of Ingo Feldmann. Briefly, ICP-MS applies an argon plasma to heat the sample (7000 K), causing the atoms to form ions that are directed to a mass spectrometer for analysis. In our setup we tested for Mn, Fe, Ni, Cu, and Zn (Supplementary Fig. 15 a). Unsurprisingly, we encountered Ni^{2+} ions that were probably left from the Ni-IDA resin used for protein purification. In addition, we confirmed the presence of Zn^{2+} ions in our sample. When compared to a standard, the intensity resulted in a ratio of Zn:Mcf1 of 1:2. This contradicts the observation of three

distinct densities in the HSE (Fig. 17 c) which would suggest the presence of three metal ions per Mcf1. One possible explanation for this discrepancy is that the ions were lost during ESI-MS sample preparation, which included an overnight dialysis of the protein. Another possibility is that a single ion is located in different positions of the interface and that these positions were averaged to the large spherical HSE during 3D refinement.

To further confirm the presence of zinc by another method, we grew Mcf1 crystals (Supplementary Fig. 15 b) based on previous work by Dr. Philine Hagel [103] to record X-ray fluorescence spectra at a synchrotron with the help of Dr. Raphale Gasper-Schoenenbruecher. The spectrum showed very low counts at the absorption edge of zinc (Supplementary Fig. 15 c), again possibly caused by a loss of the metal during crystallization.

Lastly, we also recorded UV-vis spectra of Mcf1 and compared them to two mutants which do not contain the HSE (Mcf1ΔC15, Mcf1RRR1208-1209-1214AAA). The first one, Mcf1ΔC15, is a 15 amino acid truncated version of Mcf1 and the second version was mutated in three arginine residues to no longer bind the HSE. However, no differences in absorption or emission could be observed (Supplementary Fig. 16)

In general, the identification of unknown densities in cryo-EM maps of biological samples is challenging. However, a promising approach is led by Dr. Bonnie Murphy that aims to perform electron energy-loss (EEL) measurements at single atom resolution [139]. In the future this method might help to identify the HSE.

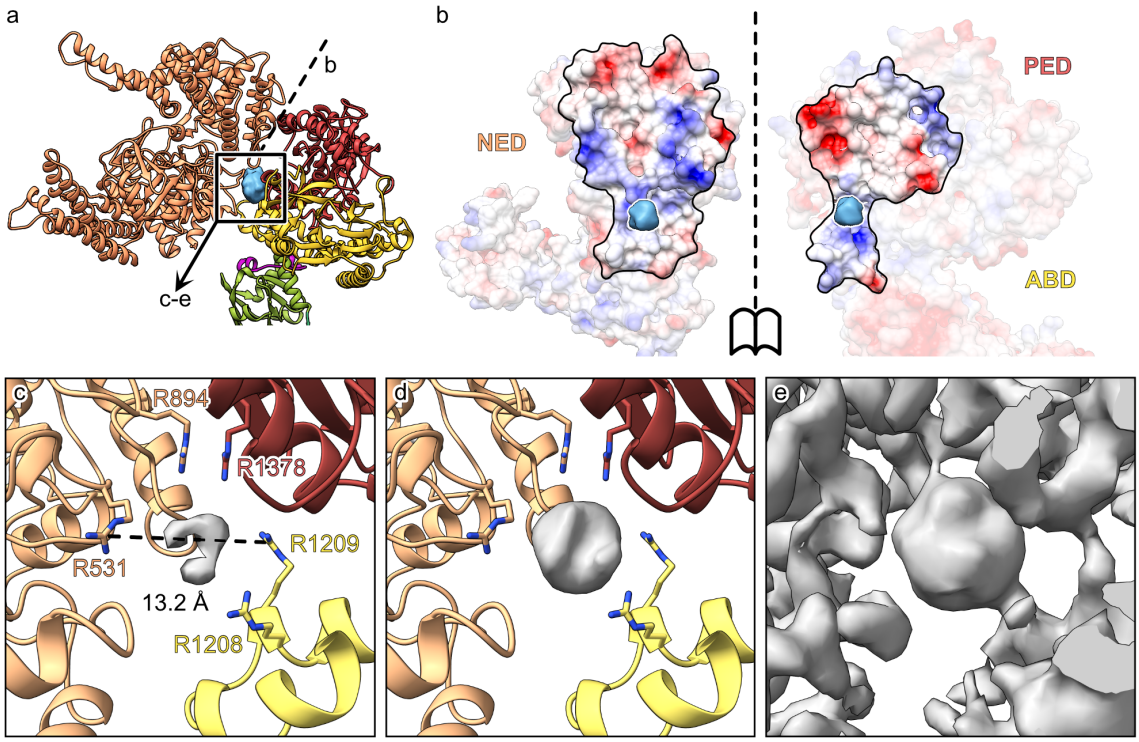


Figure 17 Head stabilizing element (HSE): **a**) Head stabilizing element (HSE, blue) within the atomic model of Mcf1. **b**) Surface representation of the N-terminal effector domain (NED) – protease effector domain (PED) interface colored according to electrostatic Coulomb potential (-10 kcal mol^{-1}, red; +10 kcal mol^{-1}, blue). Atomic model of Mcf1 with the composite cryo-EM density at high (**c**) and medium (**d**) threshold presenting the HSE as three densities and globular volume, respectively. **e**) Composite cryo-EM density at a threshold used for model building showing connections between the HSE and surrounding amino acids. Figure adapted from [43].

8.6 Arf3 activates Mcf autoproteolysis

The previous structural analysis of Mcf1 and 2 suggests the translocation of the N-terminal effector into the host cell, where it is released by the PED and ultimately causes the death of the host. This theory is supported by a loss-of-function mutation of the PED (Mcf1C1397A, Mcf2C791A), which prevents effector release and drastically reduces the toxicity of both Mcf1 and Mcf2 (Fig. 14). To further validate the cleavage of the Mcf toxins, we tracked the effector release by intoxicating Sf9 cells with N- or C-terminal myc-tagged toxin. This strategy allows us to trace the N- or C-terminal fragment size throughout the intoxication of the cells and gave insights into the underlying mechanism when combined with our structural knowledge.

The western blot analysis revealed, that Mcf1 releases a NED of 105 kDa while Mcf2 releases a 37 kDa NED, aligning with our structure-based assignment of the NED (Fig. 18 a, b). The C-terminal fragments on the other hand showed size differences between Mcf1 and Mcf2. While the N- and C-terminal fragments of Mcf2 add up to roughly the

full size of Mcf2, this is not the case for Mcf1, where a gap of about 75 kDa remains, indicating that Mcf1 is cleaved further. This diverging cleavage pattern could be related to the difference in the C-terminal regulatory element or the changed NED from Mcf1 to Mcf2.

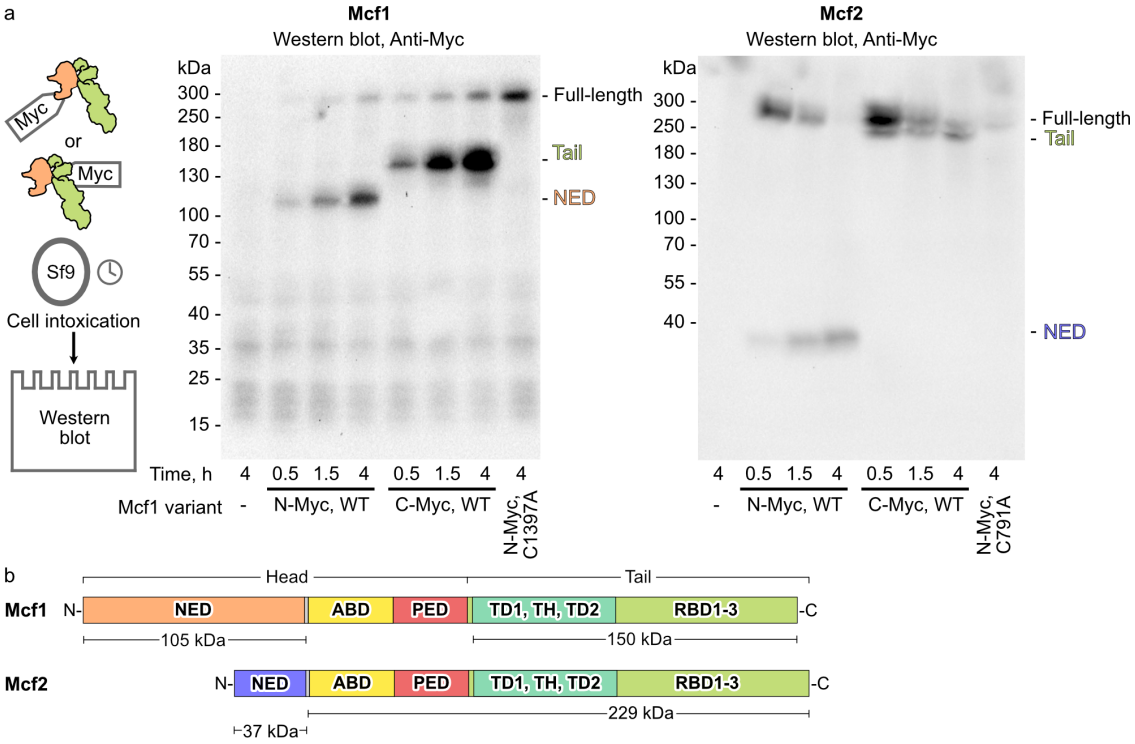


Figure 18 In vivo cleavage of Mcf1 and Mcf2: **a**) Western blot analysis of Sf9 cells intoxicated with Mcf1 or Mcf2 as well as their protease deficient mutant (Mcf1 C1397A, Mcf2 C791A). The toxins were Myc-tagged on the N- or C-terminus to trace the respective fragments. **b**) Diagram aligning the fragments to the domains of Mcf1 and Mcf2. Figure adapted from [43].

Since the purified Mcf1 toxin was stable enough for structure determination and is only cleaved after intoxication, we set out to identify the cellular activator of the PED. Considering the similarity to the MCF-like effector of the MARTX toxin from *Vibrio vulnificus*, we hypothesized that both are activated by a similar mechanism. In case of the MCF-like effector, Lee et al. showed binding and activation by ADP-ribosylation factors (Arf) in their active, GTP bound, state [38]. To validate this hypothesis, we purified Arf3 as an active GTP-bound mutant (Arf3Q71L) and incubated it together with either of the two Mcf toxins to assess the proteolytic activity. Arf3 is normally located on the Golgi and endosomal membranes [140] where it could encounter the translocated PED in context of the Mcf1 intoxication mechanism.

8.6.1 Mcf1 is processed by three autocleavage reactions

Indeed, Mcf1 showed a time dependent autoproteolysis in the presence of Arf3, resulting in the previously observed NED fragment (Fig. 19 a). In contrast to TcdA/B, additional cleavage reactions were observed leading to several C-terminal fragments instead of a single one. Furthermore, the cleavage efficiency was weak in comparison to the in vivo activation (Fig. 18 a), raising the question whether an additional factor is required to accelerate the activation. To identify such an additional eukaryotic factor, we co-expressed Mcf1 and Arf3Q71L in the prokaryote *E. coli* and observed complete cleavage of Mcf1 one hour after induction of expression (Fig. 19 c). This implied that the participation of an additional eukaryotic activator is unlikely.

Another possible explanation are the different states in which Mcf encounters Arf3. In the slow in vitro setup, Mcf1 is in the pre-translocation state which would not encounter Arf3 in the biological context. In the latter, Mcf1 would only be exposed to Arf3 after translocation through the membrane, which would involve major rearrangements of the structure. As discussed previously (8.3, 8.5), the neck region would probably undergo major transformations leading to the loss of the C-terminal head interaction. To mimic this re-arrangement, we removed the last 15 amino acids (Mcf1ΔC15) to allow for more flexibility of the head of Mcf1. As soon as we expose Mcf1ΔC15 to Arf3Q71L, the cleavage efficiency was greatly increased (Fig. 19 b), enabling us to track the fragments and map them on the structure (Fig. 19 d). According to our analysis, Mcf1 is cleaved at three positions, giving rise to the NED, ABD, PED and Mcf1 tail fragments. Furthermore, the increased activation upon removal of the C-terminus, highlights its role as an important regulatory element, that prevents Mcf from a premature activation before encountering the host and undergoing translocation.

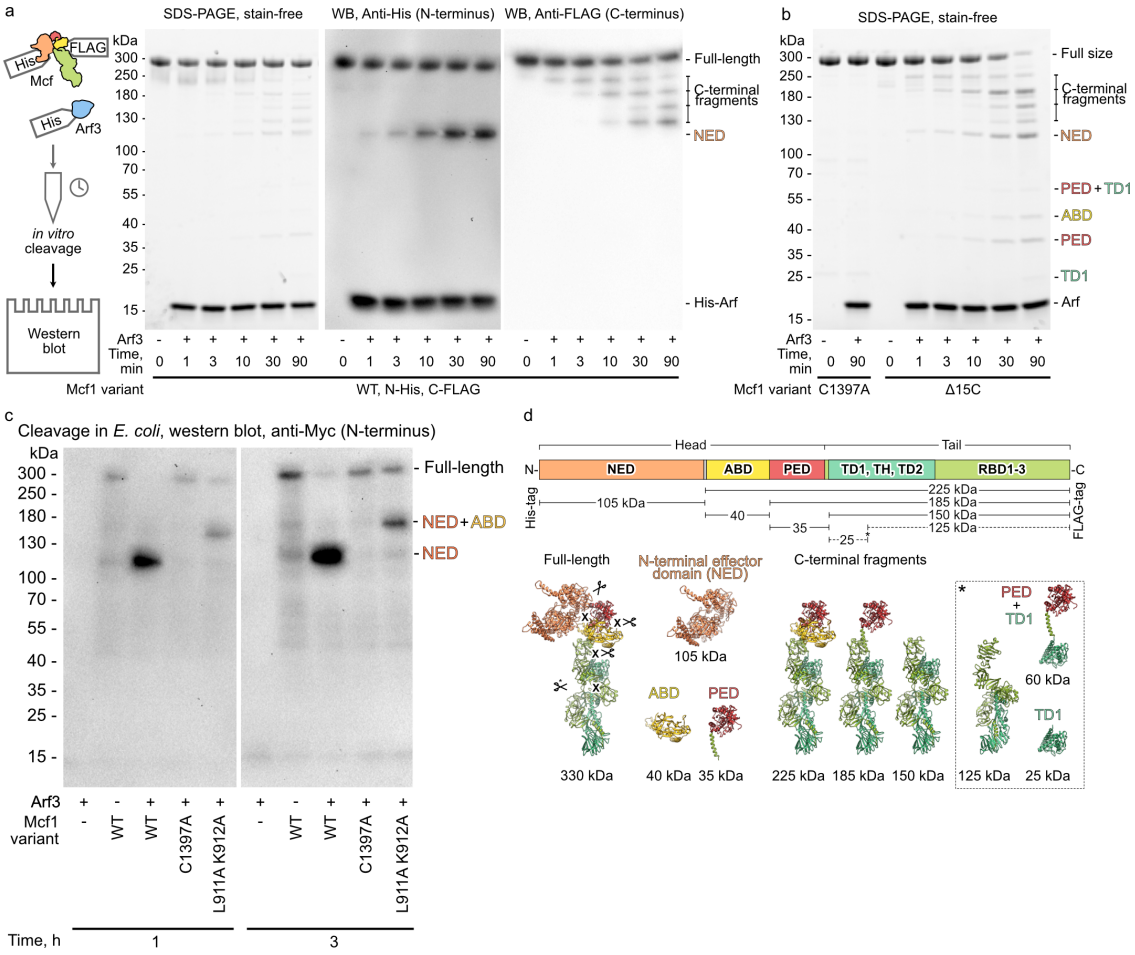


Figure 19 In vitro cleavage of Mcf1 into four fragments: **a**) SDS-PAGE and western blot analysis following the fragmentation of N-His- or C-FLAG-tagged Mcf1 during incubation with Arf3Q71L in the presence of GTP (1 mM) and $MgCl_2$ (10mM). **b**) SDS-PAGE of the same fragmentation experiment with the truncated Mcf1ΔC15 variant. **c**) Co-expression of Mcf1, Mcf1C1397A or Mcf1L911A K912A with Arf3Q71L in *E. coli* over 1 or 3h. **d**) Diagram of the fragmentation pattern resulting from the analysis of the SDS-PAGE and western blots with the according structures. Figure adapted from [43].

Based on the cleavage pattern of Mcf1ΔC15 we set out to identify the exact cleavage site and pattern that is recognized by the PED of Mcf1. To this end, we used the N-terminal His-tag to purify the NED of Mcf1 from co-expression cultures with Arf3. The purified NED was sent for mass spectrometry analysis, which identified the cleavage site between lysine 912 and alanine 913 (Supplementary Fig. 17). To further confirm cleavage at this position, we mutated both amino acids to an alanine and observed a shift in the N-terminal fragment to higher molecular weight corresponding to the NED+ABD fragment when intoxicating Sf9 cells (Fig. 20 a) or co-expressing the mutant with Arf3Q71L in *E. coli* (Fig. 19 c). This ultimately confirmed the L912/K913 cleavage site. Furthermore, intoxication of Sf9 cells with the mutant demonstrated, that cleavage is still possible at the other cleavage sites suggesting that the effector release is not dependent on a particular order of cleavage reactions.

Since we cannot purify the central fragments of Mcf1 by using tags and the C-terminal fragments had proven to be too difficult to separate, we mutated candidates for the other cleavage sites based on the Mcf1 structure and their similarity to the L912/K913 site. This strategy allowed us to also narrow down the second cleavage site to the 1271-IQGG-1274 motif (Fig. 20 a, b).

Interestingly, we encountered an additional C-terminal fragment of 125 kDa (Fig. 19 d), which only appears in the in vitro activation and results from the additional removal of the TD1 from the tail. This fragment did not appear when intoxicating Sf9 cells (Fig. 18 a) and is likely an artifact of the in vitro conditions, where Mcf1 does not follow the normal translocation of the head, which would probably leave the TD1 bound to the membrane and therefore inaccessible for the PED. This difference is another indication, that C-terminal rearrangement and translocation of the head happen prior to the activation of the PED by Arf.

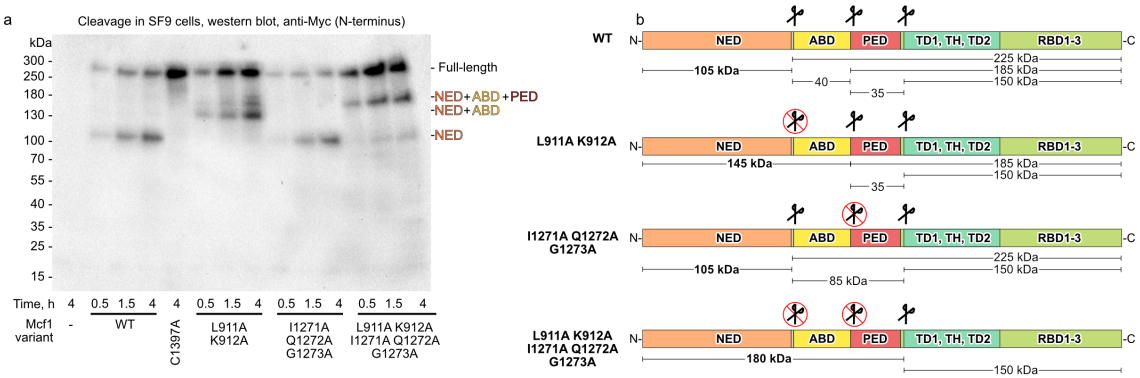


Figure 20 Cleavage site mutations of Mcf1: a) Western blot of Sf9 cells intoxicated with Myc-tagged Mcf1 mutated at the different cleavage sites, resulting in a shift of the cleavage pattern (Mcf1, Mcf1 C1397A, Mcf1 L912A K913A, Mcf1 I1271A Q1272A G1273A). B) Diagram of the constructs highlighting the targeted cleavage site(s). Figure adapted from [43].

8.6.2 Mcf2 is processed by two autocleavage reactions

Similar to Mcf1, Mcf2 is also activated by Arf3Q71L in a time-dependent manner (Fig. 21 a). However, 12h incubation of Mcf2 alone also showed cleavage activity, revealing Mcf2 as more unstable. This is also reflected in the clear fragmentation pattern of Mcf2 compared to the slow activation of Mcf1 (Fig. 19 a, 21 a). Furthermore, the instability of Mcf2 was already suggested by Waterfield et al. in 2003 (Fig. 7 b). Applying a similar tagging strategy as for Mcf1, we could identify the different fragments corresponding to the NED and several C-terminal fragments. In addition, the previously identified cleavage

site responsible for NED release is conserved for Mcf2 (Fig. 21 c) highlighting the modularity of the Mcf toxins.

Strikingly, Mcf2 also showed the presence of a larger 96 kDa N-terminal fragment comprising the complete head (NED-ABD-PED) which was not found for Mcf1 (Fig. 21 a). This further points to a greater flexibility in Mcf2 activation. However, this is only the case for the in vitro experiment, since Sf9 cell intoxication only led to the detection of the 37 kDa NED fragment (Fig. 18 a). Furthermore, unlike Mcf1, Mcf2 did not show separate fragments for ABD and PED, therefore Mcf2 is likely processed via only two cleavage reactions that release either the full head or the NED and a single ABD/PED fragment (Fig. 21 b). Possible reasons for the missing second ABD/PED cleavage site are the absence of the identified 1271-IQGG-1274 motif in Mcf2 (Fig. 21 c) or that the site is not accessible for the protease.

The observed changes in activation are likely caused by the differences of the NED and C-terminus of Mcf2. Compared to Mcf1, the missing HSE and the smaller NED of 37 kDa reduce the interaction interfaces within the head (Fig. 17) and lead to an increased flexibility of Mcf2. This facilitates the activation and enhances the observed cleavage pattern (Fig. 21). In addition, we do not know whether the longer C-terminus of Mcf2 destabilizes the head-tail interaction and thereby weakens its ability to keep Mcf2 in the stable pre-translocation state. A chimera of Mcf2 with the Mcf1 effector might allow to compare the cleavage efficiencies and draw conclusions on the underlying structural mechanism and the role of the C-terminus.

Furthermore, it is not clear what impact the structural rearrangements during translocation would have on the cleavage pattern of Mcf2, since we observe only the release of the NED in vivo and none of the other described fragments of the in vitro experiment. To properly correlate the different cleavage patterns of Mcf1 vs 2 as well as in vivo vs in vitro, more insights into the translocation mechanism are needed, that would shed light on the availability of the identified cleavage sites after translocation.

Together, these observations are an example of how changes to the regulatory elements of Mcf on a structural level change the activation pattern. Therefore, it is important to proceed with caution during the engineering process of the toxin, with particular attention

to the NED-PED/ABD interface and C-terminal head interaction. Introducing a high degree of flexibility or repositioning of the effector might make the toxin architecture unstable, initiating a premature activation.

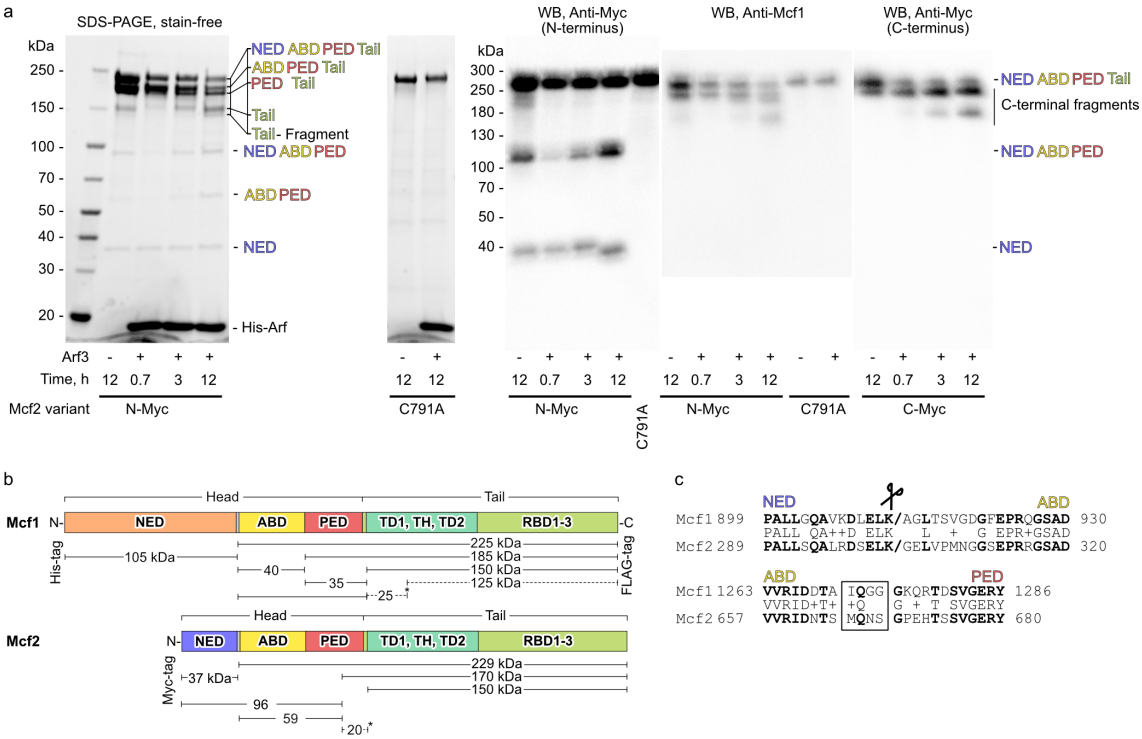


Figure 21 In vitro cleavage of Mcf2 into three fragments: **a**) SDS-PAGE and western blot analysis following the fragmentation of N- or C-terminal Myc-tagged Mcf2 during incubation with Arf3Q71L in the presence of GTP (1 mM) and $MgCl_2$ (10mM). **b**) Diagram comparing the fragmentation pattern of Mcf1 and Mcf2. Mcf1 is cleaved into NED (105 kDa), ABD (40 kDa), PED (35 kDa) and tail (150 kDa). Mcf2 is cleaved into NED (37 kDa), ABD+PED (59 kDa) and tail (150 kDa) or the head remains intact leading to a head (96 kDa) and tail (150 kDa) fragment. **c**) Sequence alignment of the NED/ABD cleavage site of Mcf1 and Mcf2. Figure b) adapted from [43].

8.7 The Mcf head contains two effectors

Despite the differences between in vitro cleavage and effector release after translocation, it is evident, that the Mcf toxins release several fragments into the host cell cytosol, where they cause apoptotic blebbing and fragmentation of the nuclei [33] (Fig. 14). To understand which fragments are causing the cytotoxic effect, we expressed them in *Saccharomyces cerevisiae* and analyzed their toxicity [101][141][142].

8.7.1 Mcf1

The 105 kDa NED fragment of Mcf1 showed high toxicity, preventing any yeast growth upon expression. Strikingly, the toxin is so potent that yeast growth was also reduced for the uninduced control due to leaky expression of the construct (Fig. 22 b). Based on the

similarity of the central domain to ADP-ribosyltransferases (Fig. 22 a, Supplementary Fig. 2), we performed western blots on the NED expressing yeast as well as intoxicated Sf9 cell lysate to identify possible substrates with anti-ADP-ribose binding protein (Fig. 22 b, Supplementary Fig. 18 a). Interestingly, no ADP-ribosylation was detected when comparing the NED sample with a negative control or the ADP-ribosylating TccC3. However, this was explained by isothermal titration calorimetry (ITC) measurements, where we detected no affinity of the NED towards NAD^+ (Supplementary Fig. 18 b), the substrate of ADP-ribosylation, which concludes that the homology is not close enough and the NED exerts toxicity by a different mechanism.

The 40 kDa ABD fragment did not show any toxicity in yeast, which was surprising considering the similarity to the Burkholderia lethal factor 1 (Blf1) (Fig. 22 a, c). However, a closer inspection of the active site revealed the absence of the catalytic cysteine in the ABD. This raises the possibility that Mcf1 once contained an additional toxic effector, which by now lost its function and only serves as a structural component of the toxin. In an attempt to introduce toxicity to the ABD, we mutated the catalytic site back to a cysteine, but observed no toxicity (Fig. 22 a, c), indicating that the ABD is still too dissimilar to Blf1. In the future it might be possible to further increase the toxicity of Mcf1 by exchanging larger parts of the ABD with Blf1 and to thereby introduce an additional toxic effector.

Previous work also suggested that the Mcf1 toxicity originates from a BH3-like domain (amino acids 911-925) that would induce apoptosis [44] and is located on the N-terminus of the ABD fragment as part of the identified cleavage site (Fig. 22 e). Our experiment suggests though, that this is not the case, as the ABD is not toxic at least in yeast cells. One reason could be the low similarity between the BH3-like domain of the Mcf toxins and Ybh3p, the BH3 protein of yeast [143]. However, the resemblance to the BH3 consensus is also limited and it is possible that the introduced mutations by Dowling et al. [44] prevented effector cleavage, which caused the reduced toxicity in their experiment.

The last 56 kDa fragment we expressed in yeast consisted of the PED including parts of the NED and backbone helix (Fig. 22 a). Surprisingly it also showed toxicity in yeast, that was dependent on its protease activity as mutations of the active site reduced toxicity (Fig. 22 d). Furthermore, we observed cleavage of the N-terminal ABD fraction that was still connected to the PED resulting in a 40 kDa PED fragment. This demonstrates an

intriguing dual function of the PED. On one hand, it is responsible for NED release and Mcf1 processing upon activation by Arf3. On the other hand, it also releases itself as a toxic effector into the cells. A similar protease domain from the *Photorhabdus asymbiotica* toxin, PaTox, was shown to also have a positive effect on toxicity [144], however it was purely based on the cleavage and the PaTox protease domain demonstrated no toxicity itself. This indicates that the dual role of the PED might be a feature specific to Mcf toxins of *P. luminescens*.

Based on the similarity between the PED and the MCF-like effector from the MARTX toxin, we sought to validate whether the PED exerts its toxicity in the same way by cleaving Rab proteins which disturbs their localization and interrupts cellular signaling ultimately leading to apoptosis [59]. Therefore, we analyzed the lysate of PED expressing yeast with a western blot against all Rab-antibodies available in the lab to test for YPT1 cleavage, which is a Rab family GTPase of *S. cerevisiae* [145][146]. The only antibody that showed a signal on the western blot was anti-Rab3a. However, no cleavage of the stained protein could be observed (Supplementary Fig. 19 a).

In a different approach, we directly incubated activated Mcf1 with some of the reported targets of the MCF-like effector (Rab23 and Rab1b) as purified proteins. In case of Rab23, the C-terminus was shown to be cleaved off, while Rab1b was completely degraded by the MCF-like effector. However, no cleavage of either Rab1b or Rab23 could be observed by activated Mcf1 (Supplementary Fig. 19 b,c). Furthermore, the proteolytic sites of the MCF-like effector system are different to the ones of Mcf and it was reported that the structure of the targeted Rab determines binding rather than a specific sequence [59].

Taken together, these findings indicate that the Mcf1 PED has a different target than the MCF-like effector or cleaves a Rab GTPase, that we did not test. A similar approach to Lee et al. with affinity purification mass spectrometry could be applied to identify the targets of the Mcf1 PED.

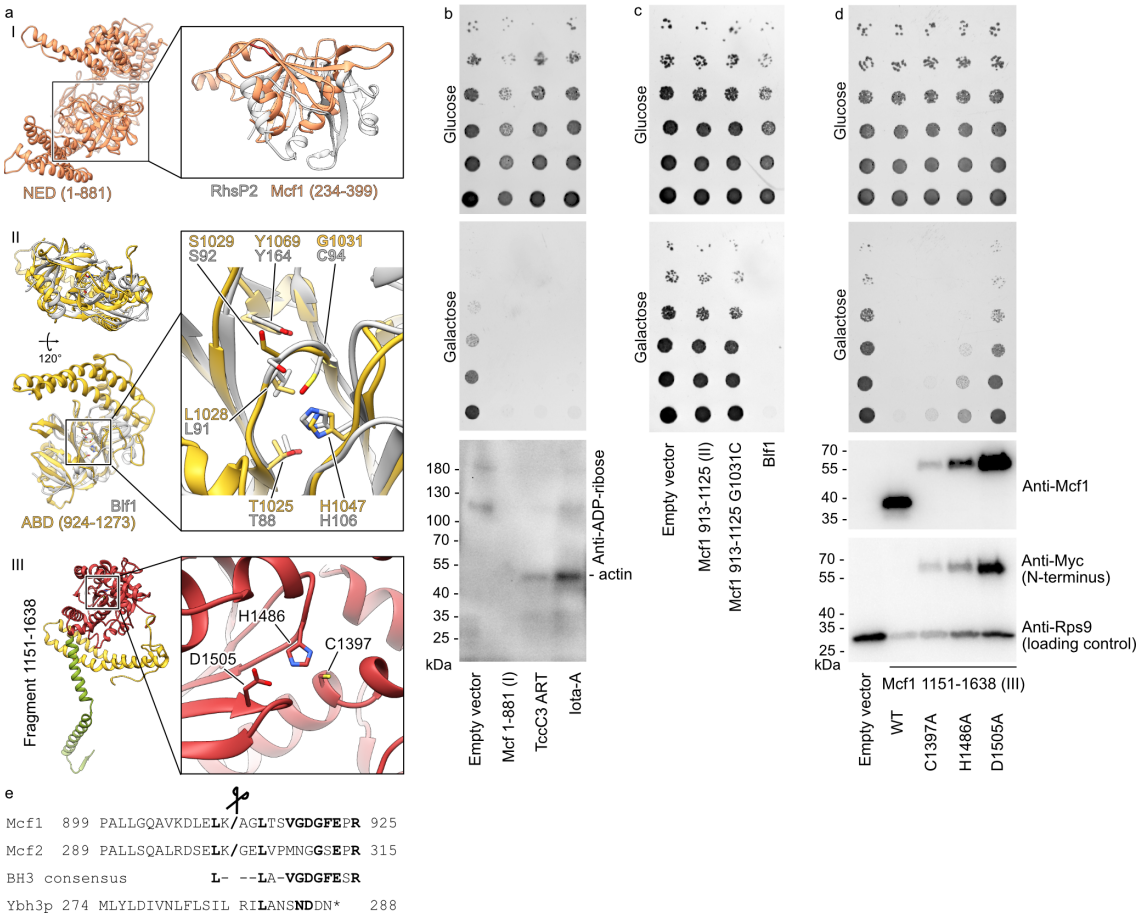


Figure 22 Yeast toxicity assay on Mcf1 head fragments: **a**) Structures of NED and ABD superimposed with structural homologs RhsP2 (7RT7 [122]) and Blf1 (3TU8 [124]) respectively, as well as the structure of the PED. **b**) Yeast toxicity assay and anti-ADP-ribose western blot for NED expression. **c**) Yeast toxicity assay for the ABD compared to Blf1. **d**) Yeast toxicity assay of the PED with different mutants of the active site. **e**) Sequence alignment of the BH3-like domain of Mcf1 and Mcf2 against the BH3 consensus sequence [44] as well as the yeast BH3 (Ybh3p). Figure adapted from [43].

8.7.2 Mcf2

To verify whether the previously described toxicity of the NED and PED are also observed in Mcf2 and therefore are a general feature of Mcf toxins, we applied the same strategy of expressing the different fragments in *S. cerevisiae*.

As expected, all fragments containing the 36.7 kDa NED exerted strong toxicity (Fig. 23). In addition, the PED was also toxic, although to a lower extent than the NED. This resemblance to Mcf1 confirms the double function of the PED as a general feature of Mcf toxins. In the same way, the ABD of Mcf2 was also non-toxic in yeast. Interestingly, the Mcf2C791A mutant was still toxic, revealing that a functioning PED is not required for toxicity if the full toxin is expressed inside the target cell. Taken together, we could show

that the HrmA-like effector of Mcf2 is toxic in Sf9 and yeast cells, however, the exact mechanism of toxicity remains to be investigated.

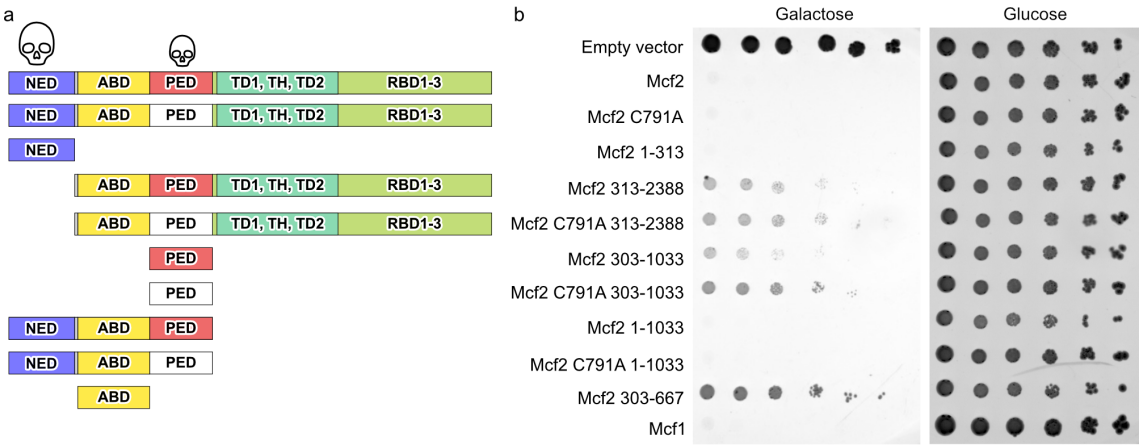


Figure 23 Yeast toxicity assay on Mcf2 fragments: **a**) Diagram of the constructs expressed in the yeast toxicity assay (**b**) (white PED is C791A mutant). Toxic domains are marked with a skull.

8.8 Mcf1 NED is a single 105 kDa effector of unknown function

Considering the unknown mechanism of action for both NEDs of Mcf1 and Mcf2, we focused on the identification of the NED of Mcf1, since it showed no sequence similarity to known effectors and the structural homology to ADP-ribosyltransferases could not be confirmed on a functional level (Fig. 22 b).

Due to the NEDs unusually large size for a toxic effector and its organization into several domains, we wanted to validate whether single domains (Fig. 24 a, b) or N- and C-terminal truncations (Fig. 24 c) would remain toxic in yeast. After induction of expression with galactose, only the full NED fragment (aa 1-881) exerted a drastic toxic effect in yeast cells. Strikingly, expression of the single domains and any truncation on either end of the NED resulted in an immediate loss of toxicity, presenting the NED as a single toxic entity (Fig. 24 c).

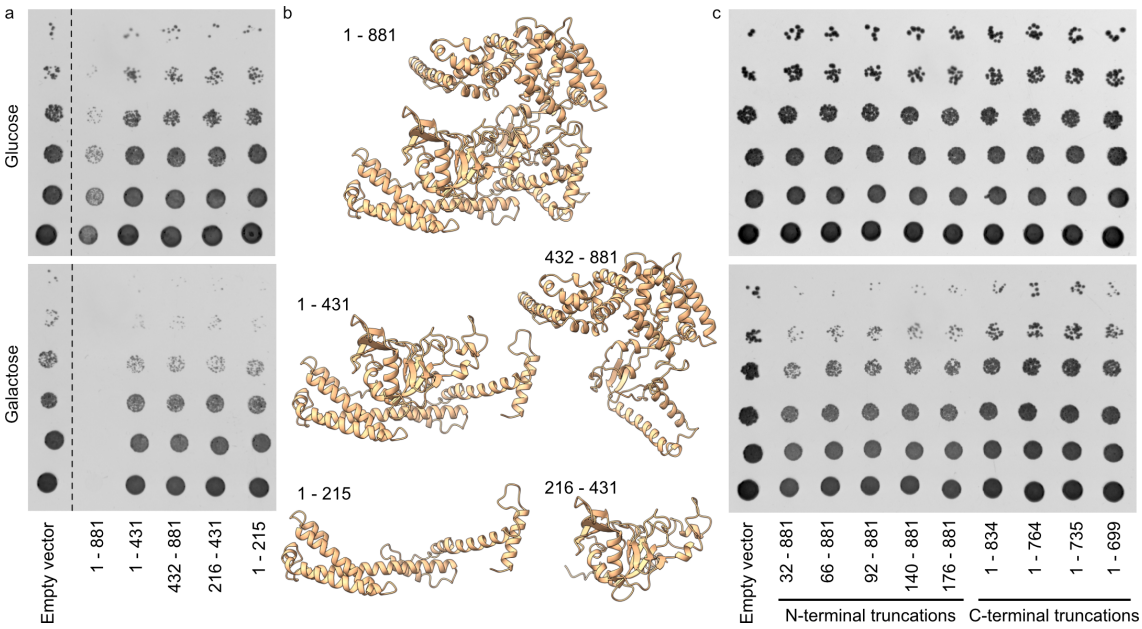


Figure 24 Yeast toxicity assay on Mcf1 NED truncations: **a**) Yeast toxicity assay of the NED and its single domains with their structure (**b**). **c**) Yeast toxicity assay of N- and C-terminal truncated versions of the NED. Figure adapted from [43].

Following this result, we wanted to identify key amino acids within this 105 kDa effector that are linked to the toxicity by performing random mutagenesis and subsequently screening the resulting mutants in our *S. cerevisiae* assay for a loss of toxicity. Due to the large size of the effector we had to introduce a rather high number of mutations per construct (~10 mutations) in order to ensure a proper coverage, which did not allow for direct conclusions based on single mutations. A comparison of all non-toxic mutants, however, revealed hotspots that were mutated in two to four different constructs (Fig. 25 a). Of those constructs, the triple and quadruple hit locations were chosen to introduce point mutations, to validate if they would remain non-toxic. However, this was not the case and all point mutants regained their toxicity (Fig. 25 b).

Following the assumption that more than a single mutation is required to render the NED non-toxic we wanted to reduce the number of mutations per construct not to a single one but to a more moderate number. To this end, we used restriction enzymes to combine one half of the mutated plasmids of the double hit group (aa 1-324) with the other half of unmutated NED (325-881) and vice versa (methods 7.2.13). This gave rise to two constructs, which still showed a reduced toxicity (Fig. 25 c). A closer inspection revealed that the first L284P mutant introduced a proline in the center of the NED, likely preventing proper folding and thereby reducing the toxicity. The other mutant introduced a bulky

hydrophilic arginine residue into a hydrophobic interface, also interrupting the fold of the NED.

All in all, the mechanism of the Mcf1 NED remains elusive, but of high interest due to its unique features that show little similarity to known effectors. A possible strategy to identify the intracellular targets and the function of the NED are pull-down experiments or proximity labeling with the NED as bait.

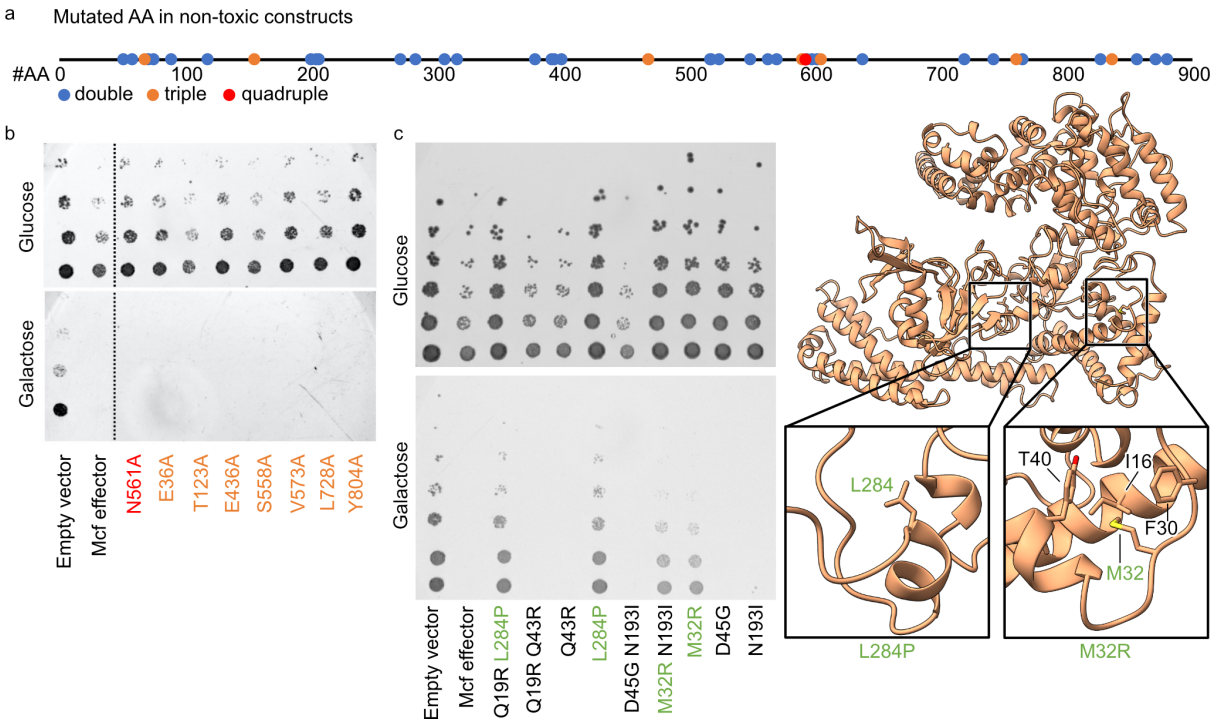


Figure 25 Random mutagenesis of the Mcf1 NED: **a**) Diagram of the mutations that were present in two (blue), three (orange) or four (red) different non-toxic mutants. **b**) Yeast toxicity assay of point mutations introduced at those triple and quadruple hit positions. **c**) Yeast toxicity assay of double- and their resulting single-mutants as well as the structure of the NED highlighting mutations that resulted in reduced toxicity.

8.9 Arf3 binding shifts cleavage site towards PED

A major question that still needs to be addressed is the structural basis for the cleavage mechanism that leads to the release of the effector fragments. A similar mechanism as for the MCF-like effector that is also activated by Arf [38] is not possible, as the reported Arf binding site is occupied by the NED of Mcf1 (Supplementary Fig. 20 a, b). Furthermore, the closest cleavage site of the NED is 25 Å away from the catalytic center of the PED, suggesting major changes upon Arf binding (Supplementary Fig. 20 c).

In order to understand those changes, we set out to create a stable Mcf1-Arf3 complex and obtain its structure. Since we already demonstrated that the activation efficiency is increased for the Mcf1ΔC15 mutant (Fig. 19 b), we concluded that the affinity to Arf3 is

higher for this mutant and should create a more stable complex for structure determination. Furthermore, this construct should resemble the pore-state of Mcf1 closer. Therefore, we acquired the structures of Mcf1ΔC15 with and without Arf3 to properly evaluate the changes.

It immediately became apparent that the removal of the last 15 amino acids destabilizes the whole tail region of Mcf1, as the reconstruction of Mcf1ΔC15 yielded a structure of only the head region (Fig. 27 a). The close interaction between the NED and PED (Fig. 17 b) keeps them well aligned, while the ABD shows an increased flexibility after losing the stabilizing interaction with the HSE and the C-terminus, which complicated the structure determination of the full head. However, intensive 3D classification finally resulted in a 3.5 Å reconstruction of the full head without Arf3 (Supplementary Fig. 21, Supplementary table S1, 3).

In contrast to Mcf1ΔC15 the 4 Å Mcf1ΔC15-Arf3 complex showed density for the ABD in all 3D classes (Supplementary Fig. 22 b) with an additional density for Arf3 bound to the ABD, resulting in its stabilization. A detailed analysis of the model revealed that Arf3 binds to the same helix bundle, that was previously occupied by the Mcf1 C-terminus (Fig. 16 a, 26 a-c). This explains the increased activation when removing the last 15 amino acids for the Mcf1ΔC15 mutant which no longer compete with Arf3 for biding to the ABD. Together with the previously described fold of the C-terminus, these findings prove the regulatory role of the C-terminus in ensuring that activation of the PED happens only after translocation, during which its interaction with the ABD is likely disrupted and leaves the ABD open for binding of Arf3.

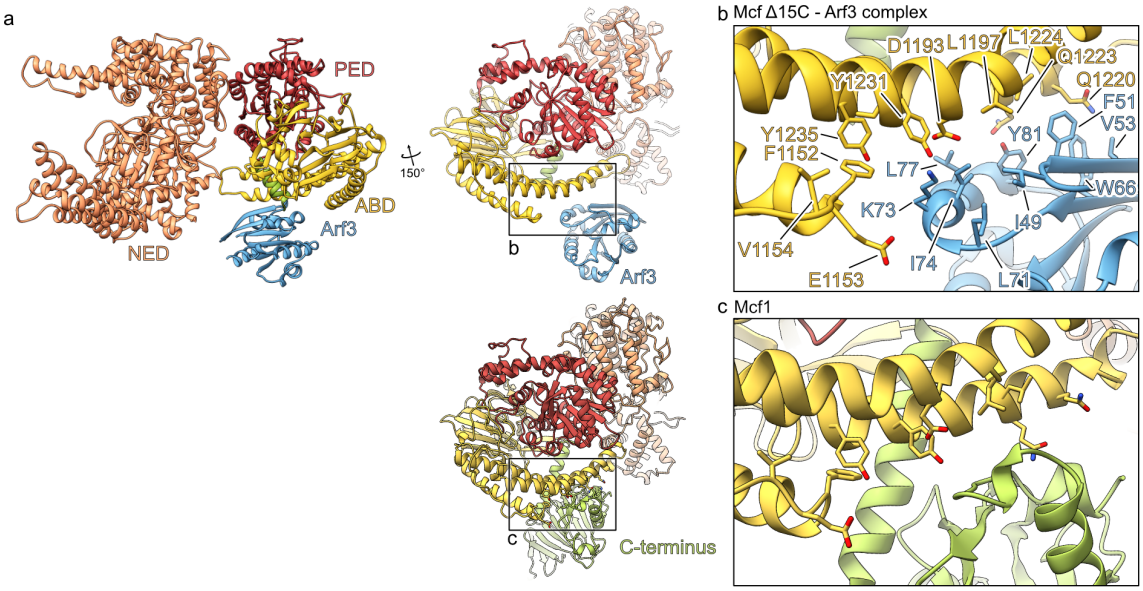


Figure 26 Mcf1ΔC15-Arf3 complex reveals Arf3 binding site: **a**) Structure of Mcf1ΔC15-Arf3 in comparison to the Mcf1 including highlighted areas of the ABD-Arf3 (**b**) and ABD-C-terminus (**c**) interactions. Figure adapted from [43].

However, the structure of the Mcf1ΔC15-Arf3 complex still displayed no change in the PED that would explain the initiation of the cleavage reactions. To find such an explanation we had to perform a more detailed analysis of the data. When inspecting the 3D classes of the Mcf1ΔC15-Arf3 complex, we noticed a swinging movement of the ABD (Fig. 27 a) that is directed by the binding of Arf3 and stretches the cleavable linker towards the active site of the PED (Fig. 27 b, c, f). This observation highlights the power of cryo-EM processing as it is able to resolve such movements that would not be visible in a crystal structure and enables us to formulate the following mechanism of force transduction from Arf3 binding to the stretching of the linker (Fig. 27).

Arf3 binding to the helix bundle of the ABD straightens them out to increase the interaction interface to Arf3. This results in an outward pull on the ABD, that moves it away from the NED, whose interaction with the PED holds it in position and prevents it from giving in to the pulling force. Therefore, the linker connecting the rigid NED and the outward swinging ABD is stretched and moved closer to the active site of the PED. The directionality of the pulling motion is conducted by the only interaction between the ABD and the PED: an arginine 957 hook that docks into a loop of the PED (aa 1290-1297, Supplementary Fig. 23 a). Upon pulling on the ABD, this hook-loop interaction acts like a hinge, directing the outward movement of the ABD and the subsequent rearrangement of the cleavable linker.

To support this mechanism, we created a series of structure-based mutations. First, we removed the arginine hook (R957A) which abolished any cleavage. This indicates that the loss of directionality of the induced movement does not rearrange the linker towards the active site and thereby prevents the cleavage (Supplementary Fig. 23 b). If we on the other hand increased the loop size by interrupting the tyrosine 1286 (T1286S) allowing for a wider opening of the loop, the linker can be pulled easier, drastically increasing the cleavage speed (Supplementary Fig. 23 a, b). Furthermore, this mutant highlights the order of cleavage reactions, with the NED being cleaved off first, followed by the release of ABD and PED. Together, these two mutations demonstrate the importance of the hook-loop interaction for triggering the cleavage reaction.

Since the linker is not resolved in our structures, we wanted to test whether the observed changes in the ABD indeed pull on the cleavable linker as proposed. The deletion of one or two amino acids to shorten the linker (ΔK907, ΔD908, Supplementary Fig. 23 c, d) showed a reduced cleavage efficiency dependent on the linker length, indicating that it is harder to pull the linker to the active site.

Lastly, we destabilized the base that anchors the interaction of ABD and PED and serves as the rigid platform on which all the rearrangements are organized. For this we introduced two glycine in a helix of the PED (L1359G, R1360G), which also reduced the cleavage efficiency (Supplementary Fig. 23 c, e) suggesting that the movement cannot be transduced properly to the linker.

All in all, both the cryo-EM data as well as the mutations support our mechanism of activation, where Arf3 binds on the ABD helix bundle that was previously occupied by the Mcf1 C-terminus and introduces a pulling motion, which shifts the cleavable linker towards the active site of the PED and initiates the cleavage cascade.

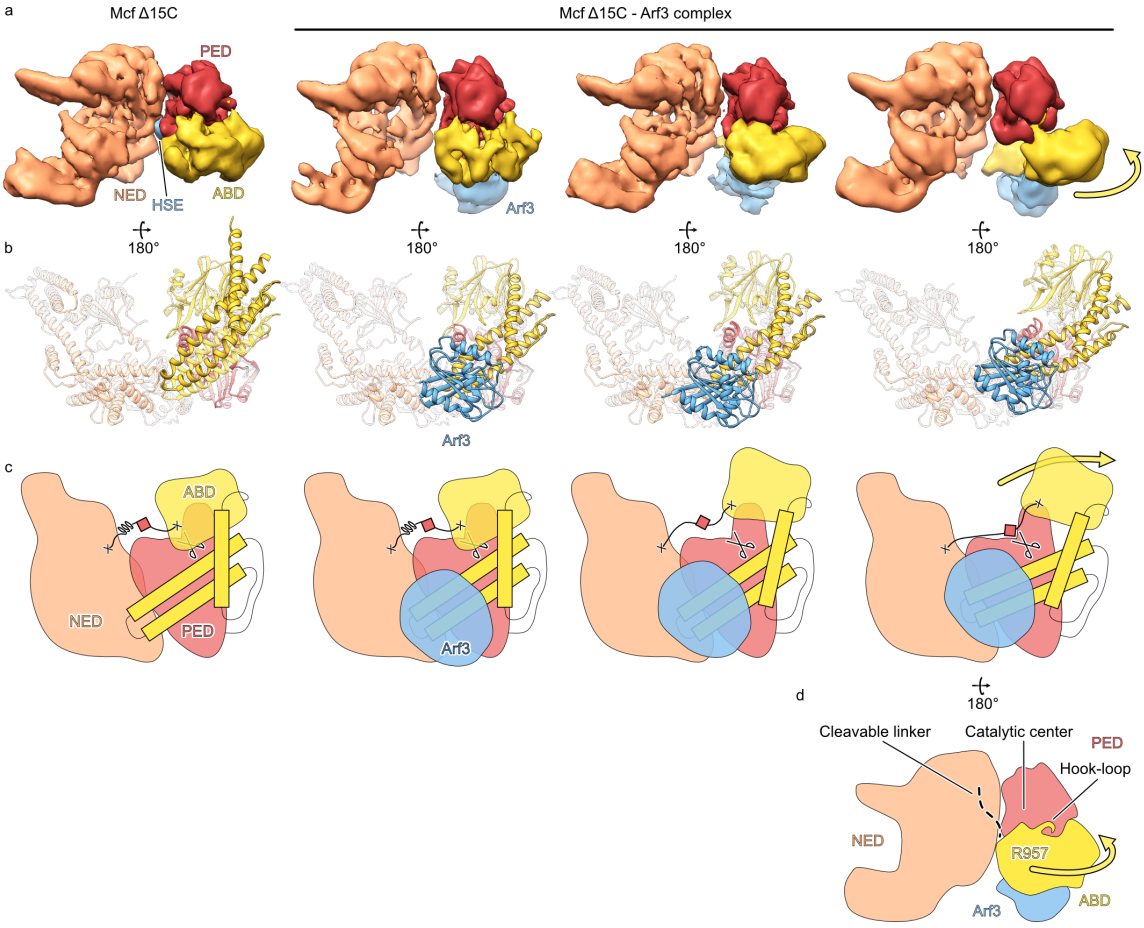


Figure 27 Activation mechanism of Mcf1: **a**) 3D class averages of Mcf1ΔC15 and the Mcf1ΔC15-Arf3 complex revealing the loss of the HSE and a swinging movement of the ABD. **b**) Models corresponding to the Class averages rotated by 180° to depict the Arf binding site on the helix bundle of the ABD. **c**) Activation scheme of Mcf1 upon Arf3 binding from bottom and front (**d**). Figure adapted from [43].

Due to the absence of a complex structure of Mcf2 with Arf, we can only speculate based on the previously described results how the changes in Mcf2 would affect the mechanism. The repositioning of the NED of Mcf2 will also result in a different positioning of the cleavable linker towards the protease domain, which might explain the differences in the cleavage pattern between Mcf1 and Mcf2. Furthermore, it also opens the possibility that Arf3 might bind to the PED in a similar fashion as for the activation of the MCF-like effector of MARTX (Supplementary Fig. 24), instead of binding the helix bundle of the ABD as we could show for Mcf1. This could also explain why Mcf2 was cleaved properly without interrupting the C-terminal interaction (Fig. 19), since activation could happen by Arf interacting with a different surface. The possibility that Mcf2 can be activated by a different or additional interaction site, would also increase the speed of activation inside the host cell. To verify this interaction, mutations to the respective binding interfaces could be introduced and the structure of Mcf2 in complex with Arf could be solved.

8.10 Mechanism of Mcf toxins

Taking all our findings into account we were able to formulate the following intoxication mechanism for the Mcf toxins.

As a member of the ABCD subfamily of bacterial toxins, the Mcf toxins first engage their target cells by binding to an unknown receptor protein and/or glycan on the cell membrane with the putative receptor binding domains 1-3. After establishing this first contact, the toxins are taken up by endocytosis. At this point, the acidification of the endosomes triggers the translocation domains 1 and 2 which are connected by a predicted transmembrane helix, to form a pore analogous to the other large clostridial toxins and translocate the head into the cytosol. Notably, the stabilizing C-terminal interaction with the ABD of the head is lost during translocation, which makes it accessible for Arf3 binding. In case of Mcf1, we could show that Arf3 binds to the now available helix bundle of the ABD, whose subsequent straightening pulls on the whole ABD and positions the cleavable linker that connects NED and ABD at the active site of the PED. Here, the effector release is initiated, causing three cleavage reactions and the release of the NED, ABD and PED into the cytosol, where NED and PED exert their toxicity. In case of Mcf2, our cleavage assays suggest only two cleavage reactions, leading to the release of the NED and a single ABD/PED fragment. The Sf9 intoxication however only confirms the release of the NED. In addition, the exact binding site of Arf3 on Mcf2 needs to be confirmed to consolidate the activation mechanism. Nonetheless, for both Mcf toxins, the NED and PED are toxic and ultimately cause the death of the host cell.

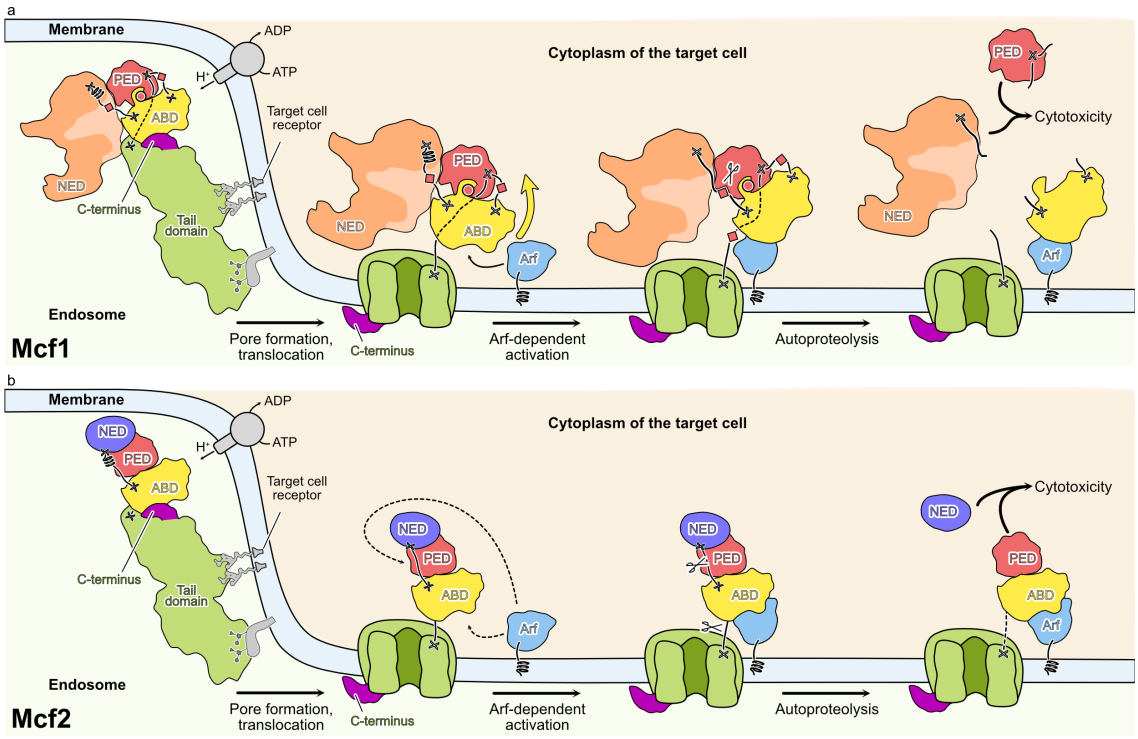


Figure 28 Mechanism of activation of Mcf1 and Mcf2: **a**) Mcf1 binds the target cell surface via proteins and/or glycans to undergo endocytosis. The acidification triggers poreformation and translocation of the head into the cytosol, which interrupts the C-terminal-ABD interaction. Arf3 in the active, GTP-bound, state binds the ABD to initiate three cleavage reactions that release the two toxic effectors, NED and PED, into the cytosol. **b**) The Mcf2 uptake and translocation of the head follows the same mechanism as Mcf1. The translocated head is activated by active Arf3, that might interact with the same binding site as for Mcf1 or a different one corresponding to the MCF-like effector of the MARTX toxin. This causes the release of the NED in Sf9 cells and the additional cleavage of an ABD-PED fragment in vitro. Similar to Mcf1, both the NED and PED are toxic effectors. Figure a) adapted from [43].

9. References

[1] U. Kutschera, "Haeckel's 1866 tree of life and the origin of eukaryotes," *Nature Microbiology 2016 1:8*, vol. 1, no. 8, pp. 1–1, Jul. 2016, doi: 10.1038/nmicrobiol.2016.114.

[2] A. Casadevall and L. A. Pirofski, "Virulence factors and their mechanisms of action: the view from a damage–response framework," *J Water Health*, vol. 7, no. S1, pp. S2–S18, Aug. 2009, doi: 10.2166/WH.2009.036.

[3] D. Roderer, E. Schubert, O. Sitsel, and S. Raunser, "Towards the application of Tc toxins as a universal protein translocation system," *Nature Communications 2019 10:1*, vol. 10, no. 1, pp. 1–11, Nov. 2019, doi: 10.1038/s41467-019-13253-8.

[4] J. Kreitz *et al.*, "Programmable protein delivery with a bacterial contractile injection system," *Nature*, vol. 616, no. 7956, pp. 357–364, 2023, doi: 10.1038/s41586-023-05870-7.

[5] A. Bravo and M. Soberón, "How to cope with insect resistance to Bt toxins?," *Trends Biotechnol*, vol. 26, no. 10, pp. 573–579, Oct. 2008, doi: 10.1016/J.TIBTECH.2008.06.005.

[6] Y. Xiao and K. Wu, "Recent progress on the interaction between insects and Bacillus thuringiensis crops," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 374, no. 1767, p. 20180316, Mar. 2019, doi: 10.1098/RSTB.2018.0316.

[7] G. Schmidt, "Some Examples of Bacterial Toxins as Tools," *Toxins 2024, Vol. 16, Page 202*, vol. 16, no. 5, p. 202, Apr. 2024, doi: 10.3390/TOXINS16050202.

[8] N. J. Tobias *et al.*, "Photorhabdus-nematode symbiosis is dependent on hfq-mediated regulation of secondary metabolites," *Environ Microbiol*, vol. 19, no. 1, pp. 119–129, Jan. 2017, doi: 10.1111/1462-2920.13502.

[9] A. H. Bhat, A. K. Chaubey, and T. H. Askary, "Global distribution of entomopathogenic nematodes, Steinernema and Heterorhabditis," *Egypt J Biol Pest Control*, vol. 30, no. 1, pp. 1–15, Dec. 2020, doi: 10.1186/S41938-020-0212-Y/TABLES/2.

[10] "Heterorhabditis bacteriophora – Nematodes for Beetle and Caterpillar Control." Accessed: Apr. 29, 2025. [Online]. Available: https://www.koppert.com/crop-protection/biological-pest-control/beneficial-nematodes/heterorhabditis-bacteriophora/

[11] S. Eckstein and R. Heermann, "Regulation of Phenotypic Switching and Heterogeneity in Photorhabdus luminescens Cell Populations," *J Mol Biol*, vol. 431, no. 23, pp. 4559–4568, Nov. 2019, doi: 10.1016/J.JMB.2019.04.015.

[12] R. H. ffrench-Constant, A. Dowling, and N. R. Waterfield, "Insecticidal toxins from Photorhabdus bacteria and their potential use in agriculture," *Toxicon*, vol. 49, no. 4, pp. 436–451, 2007, doi: 10.1016/j.toxicon.2006.11.019.

[13] E. Gerdes, "Photorhabdus Luminescens: Virulent Properties and Agricultural Applications," *American Journal of Agriculture and Forestry*, vol. 3, no. 5, p. 171, 2015, doi: 10.11648/j.ajaf.20150305.12.

[14] J. G. Gerrard and R. P. Stevens, "A Review of Clinical Cases of Infection with Photorhabdus Asymbiotica," in *The Molecular Biology of Photorhabdus Bacteria*, R. H. ffrench-Constant, Ed., Cham: Springer International Publishing, 2017, pp. 179–191. doi: 10.1007/82_2016_56.

[15] L. E. Vlaar *et al.*, "On the role of dauer in the adaptation of nematodes to a parasitic lifestyle," *Parasit Vectors*, vol. 14, no. 1, p. 554, 2021, doi: 10.1186/s13071-021-04953-6.

[16] S. S. Gang and E. A. Hallem, "Mechanisms of host seeking by parasitic nematodes," *Mol Biochem Parasitol*, vol. 208, no. 1, pp. 23–32, Jul. 2016, doi: 10.1016/J.MOLBIOPARA.2016.05.007.

[17] S. Subramanian and M. Muthulakshmi, "Entomopathogenic Nematodes," *Eco-friendly Pest Management for Food Security*, pp. 367–410, Jan. 2016, doi: 10.1016/B978-0-12-803265-7.00012-9.

[18] S. Forst, B. Dowds, N. Boemare, and E. Stackebrandt, "Xenorhabdus and Photorhabdus spp.: bugs that kill bugs," *Annu Rev Microbiol*, vol. 51, pp. 47–72, 1997, doi: 10.1146/ANNUREV.MICRO.51.1.47.

[19] E. Kenney, J. M. Hawdon, D. O'Halloran, and I. Eleftherianos, "Heterorhabditis bacteriophora Excreted-Secreted Products Enable Infection by Photorhabdus luminescens Through Suppression of the Imd Pathway," *Front Immunol*, vol. 10, p. 483369, Oct. 2019, doi: 10.3389/FIMMU.2019.02372/BIBTEX.

[20] J. J. Ryder and C. T. Griffin, "Phased infectivity in Heterorhabditis megidis: the effects of infection density in the parental host and filial generation," *Int J Parasitol*, vol. 33, no. 10, pp. 1013–1018, 2003, doi: https://doi.org/10.1016/S0020-7519(03)00131-0.

[21] A. Dutta, A. R. Flores, P. A. Revell, and L. Owens, "Neonatal Bacteremia and Cutaneous Lesions Caused by Photorhabdus luminescens: A Rare Gram-Negative Bioluminescent Bacterium," *J Pediatric Infect Dis Soc*, vol. 7, no. 3, pp. e182–e184, 2018, doi: 10.1093/jpids/piy064.

[22] K. H. Nealson and J. W. Hastings, "Bacterial bioluminescence: Its control and ecological significance," *Microbiol Rev*, vol. 43, no. 4, pp. 496–518, 1979, doi: 10.1128/mmbr.43.4.496-518.1979.

[23] Y. Imai *et al.*, "A new antibiotic selectively kills Gram- negative pathogens," *Nature*, vol. 576, no. December, 2019, doi: 10.1038/s41586-019-1791-1.

[24] F. L. Inman, S. Singh, and L. D. Holmes, "Mass Production of the Beneficial Nematode Heterorhabditis bacteriophora and Its Bacterial Symbiont Photorhabdus luminescens," *Indian J Microbiol*, vol. 52, no. 3, pp. 316–324, 2012, doi: 10.1007/s12088-012-0270-2.

[25] R. H. Ffrench-Constant and D. J. Bowen, "Novel insecticidal toxins from nematode-symbiotic bacteria," *Cellular and Molecular Life Sciences*, vol. 57, no. 5, pp. 828–833, 2000, doi: 10.1007/s000180050044.

[26] N. A. Aleksandrova, S. G. Roche, Y. S. Low, and M. J. Landsberg, "Recent insights into mechanisms of cellular toxicity and cell recognition associated with the ABC family of pore-forming toxins," *Biochem Soc Trans*, vol. 51, no. 3, pp. 1235–1244, Jun. 2023, doi: 10.1042/BST20221409.

[27] T. Jank and K. Aktories, "Structure and mode of action of clostridial glucosylating toxins: the ABCD model," *Trends Microbiol*, vol. 16, no. 5, pp. 222–229, 2008, doi: 10.1016/j.tim.2008.01.011.

[28] P. N. Ng'ang'a *et al.*, "Multistate kinetics of the syringe-like injection mechanism of Tc toxins," *Sci Adv*, vol. 11, no. 1, Jan. 2025, doi: 10.1126/SCIADV.ADR2019.

[29] H. C. Wang, S. J. Lin, H. C. Wang, R. Kumar, P. T. Le, and J. H. Leu, "A bacterial binary toxin system that kills both insects and aquatic crustaceans: Photorhabdus insect-related toxins A and B," *PLoS Pathog*, vol. 19, no. 5, p. e1011330, May 2023, doi: 10.1371/JOURNAL.PPAT.1011330.

[30] X. Zhang *et al.*, "XaxAB-like binary toxin from Photorhabdus luminescens exhibits both insecticidal activity and cytotoxicity," *FEMS Microbiol Lett*, vol. 350, no. 1, pp. 48–56, 2014, doi: 10.1111/1574-6968.12321.

[31] I. Vlisidou, A. Hapeshi, J. R. J. Healey, K. Smart, G. Yang, and N. R. Waterfield, "The photorhabdus asymbiotica virulence cassettes deliver protein effectors directly into target eukaryotic cells," *Elife*, vol. 8, Sep. 2019, doi: 10.7554/ELIFE.46259.

[32] F. Jiang *et al.*, "Cryo-EM Structure and Assembly of an Extracellular Contractile Injection System," *Cell*, vol. 177, no. 2, pp. 370-383.e15, Apr. 2019, doi: 10.1016/J.CELL.2019.02.020.

[33] A. J. Dowling, P. J. Daborn, N. R. Waterfield, P. Wang, C. H. Streuli, and R. H. Ffrench-Constant, "The insecticidal toxin Makes caterpillars floppy (Mcf) promotes apoptosis in mammalian cells," *Cell Microbiol*, vol. 6, no. 4, pp. 345–353, 2004, doi: https://doi.org/10.1046/j.1462-5822.2003.00357.x.

[34] P. J. Daborn, N. Waterfield, C. P. Silva, C. P. Y. Au, S. Sharma, and R. H. Ffrench-Constant, "A single Photorhabdus gene, makes caterpillars floppy (mcf), allows Escherichia coli to persist within and kill insects," *Proc Natl Acad Sci U S A*, vol. 99, no. 16, pp. 10742–10747, 2002, doi: 10.1073/pnas.102068099.

[35] N. R. Waterfield, P. J. Daborn, A. J. Dowling, G. Yang, M. Hares, and R. H. Ffrench-Constant, "The insecticidal toxin Makes caterpillars floppy 2 (Mcf2) shows similarity to HrmA, an avirulence protein from a plant pathogen," *FEMS Microbiol Lett*, vol. 229, no. 2, pp. 265–270, 2003, doi: 10.1016/S0378-1097(03)00846-2.

[36] R. C. Budd, "Activation-induced cell death," *Curr Opin Immunol*, vol. 13, no. 3, pp. 356–362, 2001.

[37] A. Herrera *et al.*, "N-terminal autoprocessing and acetylation of multifunctional-autoprocessing repeats-in-toxins (MARTX) Makes Caterpillars Floppy-like effector is stimulated by adenosine diphosphate (ADP)-Ribosylation Factor 1 in advance of Golgi fragmentation," *Cell Microbiol*, vol. 22, no. 2, p. e13133, 2020, doi: https://doi.org/10.1111/cmi.13133.

[38] Y. Lee, B. Sik, S. Choi, E. Lee, S. Park, and J. Hwang, "Makes caterpillars floppy-like effector-containing MARTX toxins require host ADP-ribosylation factor (ARF) proteins for systemic pathogenicity," *Pnas*, vol. 116, no. 36, 2019, doi: 10.1073/pnas.1905095116.

[39] F. Hofmann, C. Busch, U. Prepens, I. Just, and K. Aktories, "Localization of the Glucosyltransferase Activity of Clostridium difficile Toxin B to the N-terminal Part of the Holotoxin*," *Journal of Biological Chemistry*, vol. 272, no. 17, pp. 11074–11078, 1997, doi: https://doi.org/10.1074/jbc.272.17.11074.

[40] A. Schaller *et al.*, "Characterization of apxIVA, a new RTX determinant of Actinobaci//us pleuropneumoniae."

[41] I. Ullah *et al.*, "Identification and Characterization of the Insecticidal Toxin 'Makes Caterpillars Floppy' in Photorhabdus temperata M1021 Using a Cosmid Library," *Toxins (Basel)*, vol. 6, no. 7, pp. 2024–2040, 2014, doi: 10.3390/toxins6072024.

[42] P. J. Christie and J. P. Vogel, “Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells,” *Trends Microbiol*, vol. 8, no. 8, pp. 354–360, Aug. 2000, doi: 10.1016/S0966-842X(00)01792-3.

[43] A. Belyy, P. Heilen, P. Hagel, O. Hofnagel, and S. Raunser, “Structure and activation mechanism of the Makes caterpillars floppy 1 toxin,” *Nat Commun*, vol. 14, no. 1, p. 8226, 2023, doi: 10.1038/s41467-023-44069-2.

[44] A. J. Dowling, N. R. Waterfield, M. C. Hares, G. Le Goff, C. H. Streuli, and R. H. Ffrench-Constant, “The Mcf1 toxin induces apoptosis via the mitochondrial pathway and apoptosis is attenuated by mutation of the BH3-like domain,” *Cell Microbiol*, vol. 9, no. 10, pp. 2470–2484, 2007, doi: 10.1111/j.1462-5822.2007.00974.x.

[45] T. Igaki *et al.*, “Drob-1, a Drosophila member of the Bcl-2/CED-9 family that promotes cell death,” *Proceedings of the National Academy of Sciences*, vol. 97, no. 2, pp. 662–667, 2000, doi: 10.1073/pnas.97.2.662.

[46] A. Kelekar and C. B. Thompson, “Bcl-2-family proteins: the role of the BH3 domain in apoptosis,” *trens in CELL BIOLOGY*, vol. 8, pp. 324–349, 1998, doi: 10.1016/S0962-8924(98)01321-X.

[47] Y. Tsujimoto and S. Shimizu, “Bcl-2 family: life-or-death switch,” *FEBS Lett*, vol. 466, no. 1, pp. 6–10, Jan. 2000, doi: 10.1016/S0014-5793(99)01761-5.

[48] T. Igaki and M. Miura, “Role of Bcl-2 family members in invertebrates,” *Biochim Biophys Acta Mol Cell Res*, vol. 1644, no. 2–3, pp. 73–81, Mar. 2004, doi: 10.1016/j.bbamcr.2003.09.007.

[49] H. Zha, C. Aimé-Sempé, T. Sato, and J. C. Reed, “Proapoptotic protein Bax heterodimerizes with Bcl-2 and homodimerizes with Bax via a novel domain (BH3) distinct from BH1 and BH2,” *Journal of Biological Chemistry*, vol. 271, no. 13, pp. 7440–7444, Mar. 1996, doi: 10.1074/jbc.271.13.7440.

[50] M. Sattler *et al.*, “Structure of Bcl-x(L)-Bak peptide complex: Recognition between regulators of apoptosis,” *Science (1979)*, vol. 275, no. 5302, pp. 983–986, Feb. 1997, doi: 10.1126/SCIENCE.275.5302.983/ASSET/5DFF2525-FA10-4DD1-90CC-6458526C5A34/ASSETS/GRAPHIC/SE0674748003.JPEG.

[51] M. C. Wei *et al.*, “tBID, a membrane-targeted death ligand, oligomerizes BAK to release cytochrome c,” *Genes Dev*, vol. 14, no. 16, pp. 2060–2071, Aug. 2000, doi: 10.1101/GAD.14.16.2060.

[52] M. C. Wei *et al.*, “Proapoptotic BAX and BAK: A requisite gateway to mitochondrial dysfunction and death,” *Science (1979)*, vol. 292, no. 5517, pp. 727–730, Apr. 2001, doi: 10.1126/SCIENCE.1059108/ASSET/6F75DA57-D445-4561-839C-7C1F1C55CAB9/ASSETS/GRAPHIC/SE1619369004.JPEG.

[53] K. J. F. Satchell, “Multifunctional-autoprocessing repeats-in-toxin (MARTX) Toxins of Vibrios,” *Microbiol Spectr*, vol. 3, no. 3, pp. 10.1128/microbiolspec.VE-0002–2014, Jun. 2015, doi: 10.1128/MICROBIOLSPEC.VE-0002-2014.

[54] H. E. Gavin and K. J. F. Satchell, “MARTX toxins as effector delivery platforms,” *Pathog Dis*, vol. 73, no. 9, Dec. 2015, doi: 10.1093/FEMSPD/FTV092.

[55] B. S. Kim, “The Modes of Action of MARTX Toxin Effector Domains,” *Toxins (Basel)*, vol. 10, no. 12, Dec. 2018, doi: 10.3390/TOXINS10120507.

[56] S. Agarwal, S. Agarwal, M. Biancucci, and K. J. F. Satchell, “Induced autoprocessing of the cytopathic Makes caterpillars floppy-like effector domain of the Vibrio vulnificusMARTX toxin,” *Cell Microbiol*, vol. 17, no. 10, pp. 1494–1509, 2015, doi: 10.1111/cmi.12451.

[57] S. Agarwal, Y. Zhu, D. R. Gius, and K. J. F. Satchell, "The Makes Caterpillars Floppy (MCF)-Like Domain of Vibrio vulnificus Induces Mitochondrion-Mediated Apoptosis," *Infect Immun*, vol. 83, no. 11, pp. 4392–4403, 2015, doi: 10.1128/IAI.00570-15.

[58] L. A. Volpicelli-Daley, Y. Li, C. J. Zhang, and R. A. Kahn, "Isoform-selective effects of the depletion of ADP-ribosylation factors 1-5 on membrane traffic," *Mol Biol Cell*, vol. 16, no. 10, pp. 4495–4508, Oct. 2005, doi: 10.1091/MBC.E04-12-1042/ASSET/IMAGES/LARGE/ZMK0100573140009.JPEG.

[59] A. Herrera, M. M. Packer, M. Rosas-Lemus, G. Minasov, J. H. Brummel, and K. J. F. Satchell, "Vibrio MARTX toxin processing and degradation of cellular Rab GTPases by the cytotoxic effector Makes Caterpillars Floppy", doi: 10.1101/2023.04.19.537381.

[60] S. Agarwal *et al.*, "Autophagy and endosomal trafficking inhibition by Vibrio cholerae MARTX toxin phosphatidylinositol-3-phosphate-specific phospholipase A1 activity," *Nat Commun*, vol. 6, Oct. 2015, doi: 10.1038/NCOMMS9745.

[61] D. E. Voth and J. D. Ballard, "Clostridium difficile toxins: Mechanism of action and role in disease," *Clin Microbiol Rev*, vol. 18, no. 2, pp. 247–263, 2005, doi: 10.1128/CMR.18.2.247-263.2005/ASSET/DE4F1FEA-2169-4FA5-A3C0-945A5AC796BF/ASSETS/GRAPHIC/ZCM0020521290003.JPEG.

[62] A. Y. Guh *et al.*, " Trends in U.S. Burden of Clostridioides difficile Infection and Outcomes ," *New England Journal of Medicine*, vol. 382, no. 14, pp. 1320–1330, Apr. 2020, doi: 10.1056/NEJMOA1910215/SUPPL_FILE/NEJMOA1910215_DISCLOSURES.PDF.

[63] W. K. Smits, D. Lyras, D. B. Lacy, M. H. Wilcox, and E. J. Kuijper, "Clostridium difficile infection," *Nature Reviews Disease Primers 2016 2:1*, vol. 2, no. 1, pp. 1–20, Apr. 2016, doi: 10.1038/nrdp.2016.20.

[64] S. L. Kordus, A. K. Thomas, and D. B. Lacy, "Clostridioides difficile toxins: mechanisms of action and antitoxin therapeutics," *Nature Reviews Microbiology 2021 20:5*, vol. 20, no. 5, pp. 285–298, Nov. 2021, doi: 10.1038/s41579-021-00660-2.

[65] L. E. Hartley-Tassell *et al.*, "Lectin activity of the TcdA and TcdB toxins of Clostridium difficile," *Infect Immun*, vol. 87, no. 3, Mar. 2019, doi: 10.1128/IAI.00676-18/SUPPL_FILE/IAI.00676-18-S0005.PDF.

[66] K. D. Tucker and T. D. Wilkins, "Toxin A of Clostridium difficile binds to the human carbohydrate antigens I, X, and Y," *Infect Immun*, vol. 59, no. 1, pp. 73–78, 1991, doi: 10.1128/IAI.59.1.73-78.1991.

[67] L. Tao *et al.*, "Sulfated glycosaminoglycans and low-density lipoprotein receptor contribute to Clostridium difficile toxin A entry into cells," *Nature Microbiology 2019 4:10*, vol. 4, no. 10, pp. 1760–1769, Jun. 2019, doi: 10.1038/s41564-019-0464-z.

[68] X. Na, H. Kim, M. P. Moyer, C. Pothoulakis, and J. T. LaMont, "gp96 is a human colonocyte plasma membrane binding protein for Clostridium difficile toxin A," *Infect Immun*, vol. 76, no. 7, pp. 2862–2871, Jul. 2008, doi: 10.1128/IAI.00326-08/ASSET/04116951-DE51-4527-B3AE-BFC125DAD723/ASSETS/GRAPHIC/ZII0070873810007.JPEG.

[69] C. Pothoulakis *et al.*, "Rabbit sucrase-isomaltase contains a functional intestinal receptor for Clostridium difficile toxin A.," *J Clin Invest*, vol. 98, no. 3, pp. 641–649, Aug. 1996, doi: 10.1172/JCI118835.

[70] L. Tao *et al.*, "Frizzled proteins are colonic epithelial receptors for C. difficile toxin B," *Nature*, vol. 538, no. 7625, pp. 350–355, 2016, doi: 10.1038/nature19799.

[71] M. E. LaFrance, M. A. Farrow, R. Chandrasekaran, J. Sheng, D. H. Rubin, and D. B. Lacy, "Identification of an epithelial cell receptor responsible for Clostridium difficile TcdB-induced cytotoxicity," *Proc Natl Acad Sci U S A*, vol. 112, no. 22, pp. 7073–7078, Jun. 2015, doi: 10.1073/PNAS.1500791112/SUPPL_FILE/PNAS.201500791SI.PDF.

[72] P. Yuan *et al.*, "Chondroitin sulfate proteoglycan 4 functions as the cellular receptor for Clostridium difficile toxin B," *Cell Res*, vol. 25, no. 2, pp. 157–168, Feb. 2015, doi: 10.1038/cr.2014.169.

[73] S. Genisyuerek, P. Papatheodorou, G. Guttenberg, R. Schubert, R. Benz, and K. Aktories, "Structural determinants for membrane insertion, pore formation and translocation of Clostridium difficile toxin B," *Mol Microbiol*, vol. 79, no. 6, pp. 1643–1654, Mar. 2011, doi: 10.1111/J.1365-2958.2011.07549.X.

[74] I. Just, J. Selzer, M. Wilm, C. Von Eichel-Streiber, M. Mann, and K. Aktories, "Glucosylation of Rho proteins by Clostridium difficile toxin B," *Nature 1995 375:6531*, vol. 375, no. 6531, pp. 500–503, Jun. 1995, doi: 10.1038/375500a0.

[75] I. Just, J. Selzer, C. Von Eichel-Streiber, and K. Aktories, "The low molecular mass GTP-binding protein Rho is affected by toxin A from Clostridium difficile.," *J Clin Invest*, vol. 95, no. 3, pp. 1026–1031, Mar. 1995, doi: 10.1172/JCI117747.

[76] S. Hippenstiel *et al.*, "Rho protein inactivation induced apoptosis of cultured human endothelial cells," *Am J Physiol Lung Cell Mol Physiol*, vol. 283, no. 4 27-4, 2002, doi: 10.1152/AJPLUNG.00467.2001/ASSET/IMAGES/LARGE/H51021072008.JPEG.

[77] C. Quesada-Gómez *et al.*, "Analysis of TcdB proteins within the hypervirulent clade 2 reveals an impact of RhoA glucosylation on Clostridium difficile proinflammatory activities," *Infect Immun*, vol. 84, no. 3, pp. 856–865, Mar. 2016, doi: 10.1128/IAI.01291-15/ASSET/A1FB0D2A-B8B0-4738-BC2F-1AA6E7D9E183/ASSETS/GRAPHIC/ZII9990916150009.JPEG.

[78] N. M. Chumbler *et al.*, "Crystal structure of Clostridium difficile toxin A," *Nat Microbiol*, vol. 1, no. 1, 2016, doi: 10.1038/nmicrobiol.2015.2.

[79] K. E. Orrell, M. J. Mansfield, A. C. Doxey, and R. A. Melnyk, "The C. difficile toxin B membrane translocation machinery is an evolutionarily conserved protein delivery apparatus," *Nat Commun*, vol. 11, no. 1, pp. 1–11, 2020, doi: 10.1038/s41467-020-14306-z.

[80] Z. Zhang *et al.*, "Translocation domain mutations affecting cellular toxicity identify the Clostridium difficile toxin B pore," *Proc Natl Acad Sci U S A*, vol. 111, no. 10, pp. 3721–3726, 2014, doi: 10.1073/pnas.1400680111.

[81] P. Gangola and B. P. Rosen, "Maintenance of intracellular calcium in Escherichia coli.," *Journal of Biological Chemistry*, vol. 262, no. 26, pp. 12570–12574, Sep. 1987, doi: 10.1016/S0021-9258(18)45243-X.

[82] I. Linhartová *et al.*, "RTX proteins: a highly diverse family secreted by a common mechanism," *FEMS Microbiol Rev*, vol. 34, no. 6, pp. 1076–1112, Nov. 2010, doi: 10.1111/J.1574-6976.2010.00231.X.

[83] M. Sugawara *et al.*, "Rhizobitoxine production in Agrobacterium tumefaciens C58 by Bradyrhizobium elkanii rtxACDEFG genes," *FEMS Microbiol Lett*, vol. 269, no. 1, pp. 29–35, Apr. 2007, doi: 10.1111/J.1574-6968.2006.00590.X.

[84] J. Liu *et al.*, "In vivo induced RTX toxin ApxIVA is essential for the full virulence of Actinobacillus pleuropneumoniae," *Vet Microbiol*, vol. 137, no. 3–4, pp. 282–289, Jun. 2009, doi: 10.1016/J.VETMIC.2009.01.011.

[85] A. Münch, L. Stingl, K. Jung, and R. Heermann, "Photorhabdus luminescens genes induced upon insect infection," *BMC Genomics*, vol. 9, no. 1, pp. 1–17, May 2008, doi: 10.1186/1471-2164-9-229/FIGURES/3.

[86] J. R. Alfano *et al.*, "The Pseudomonas syringae Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants," *Proc Natl Acad Sci U S A*, vol. 97, no. 9, pp. 4856–4861, Apr. 2000, doi: 10.1073/PNAS.97.9.4856/ASSET/74034E1F-4205-4D52-ABBE-8376A6706FB2/ASSETS/GRAPHIC/PQ0800694005.JPEG.

[87] J. R. Alfano and A. Collmer, "The type III (Hrp) secretion pathway of plant pathogenic bacteria: Trafficking harpins, Avr proteins, and death," *J Bacteriol*, vol. 179, no. 18, pp. 5655–5662, 1997, doi: 10.1128/JB.179.18.5655-5662.1997/ASSET/96BBF821-6FDC-4841-827B-5DB6B8A6AAE7/ASSETS/JB.179.18.5655-5662.1997.FP.PNG.

[88] J. R. Alfano, H.-S. Kim, T. P. Delaney, and A. Collmer, "Evidence That the Pseudomonas syringae pv. syringae hrp-Linked hrmA Gene Encodes an Avr-Like Protein That Acts in an hrp-Dependent Manner Within Tobacco Cells," *Mol Plant Microbe Interact*, vol. 580, no. 5, pp. 580–588, 1997.

[89] D. Quentin and S. Raunser, "Electron cryomicroscopy as a powerful tool in biomedical research," *J Mol Med*, vol. 96, no. 6, pp. 483–493, 2018, doi: 10.1007/s00109-018-1640-y.

[90] Z. Yang, J. Fang, J. Chittuluru, F. J. Asturias, and P. A. Penczek, "Iterative stable alignment and clustering of 2D transmission electron microscope images," *Structure*, vol. 20, no. 2, pp. 237–247, Feb. 2012, doi: 10.1016/j.str.2011.12.007.

[91] F. Schöenfeld, M. Stabrin, T. R. Shaikh, T. Wagner, and S. Raunser, "Accelerated 2D Classification With ISAC Using GPUs," *Front Mol Biosci*, vol. 9, p. 919994, Jul. 2022, doi: 10.3389/FMOLB.2022.919994/BIBTEX.

[92] S. H. W. Scheres, "Chapter Six - Processing of Structurally Heterogeneous Cryo-EM Data in RELION," in *The Resolution Revolution: Recent Advances In cryoEM*, vol. 579, R. A. B. T.-M. in E. Crowther, Ed., Academic Press, 2016, pp. 125–157. doi: https://doi.org/10.1016/bs.mie.2016.04.012.

[93] A. Punjani and D. J. Fleet, "3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM," *J Struct Biol*, vol. 213, no. 2, p. 107702, Jun. 2021, doi: 10.1016/J.JSB.2021.107702.

[94] A. Punjani and D. J. Fleet, "3DFlex: determining structure and motion of flexible proteins from cryo-EM," *Nature Methods 2023 20:6*, vol. 20, no. 6, pp. 860–870, May 2023, doi: 10.1038/s41592-023-01853-8.

[95] G. McMullan, A. R. Faruqi, D. Clare, and R. Henderson, "Comparison of optimal performance at 300 keV of three direct electron detectors for use in low dose electron microscopy," *Ultramicroscopy*, vol. 147, pp. 156–163, Dec. 2014, doi: 10.1016/J.ULTRAMIC.2014.08.002.

[96] S. Q. Zheng, E. Palovcak, J. P. Armache, K. A. Verba, Y. Cheng, and D. A. Agard, "MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy," *Nat Methods*, vol. 14, no. 4, pp. 331–332, 2017, doi: 10.1038/nmeth.4193.
[97] M. Stabrin, F. Schoenfeld, T. Wagner, S. Pospich, C. Gatsogiannis, and S. Raunser, "TranSPHIRE: Automated and feedback-optimized on-the-fly processing for cryo-EM," 2020, doi: 10.1101/2020.06.16.155275.
[98] W. Chiu, M. F. Schmid, G. D. Pintilie, and C. L. Lawson, "Evolution of standardization and dissemination of cryo-EM structures and data jointly by the community, PDB, and EMDB," *J Biol Chem*, vol. 296, p. 100560, Jan. 2021, doi: 10.1016/J.JBC.2021.100560.
[99] T. Nakane *et al.*, "Single-particle cryo-EM at atomic resolution," 2020.
[100] R. Danev, B. Buijsse, M. Khoshouei, J. M. Plitzko, and W. Baumeister, "Volta potential phase plate for in-focus phase contrast transmission electron microscopy," *Proc Natl Acad Sci U S A*, vol. 111, no. 44, pp. 15635–15640, Nov. 2014, doi: 10.1073/PNAS.1418377111/SUPPL_FILE/PNAS.1418377111.SM02.AVI.
[101] A. Belyy, I. Tabakova, A. E. Lang, T. Jank, Y. Belyi, and K. Aktories, "Roles of Asp179 and Glu270 in ADP-Ribosylation of Actin by Clostridium perfringens Iota Toxin," *PLoS One*, vol. 10, no. 12, p. e0145708, Dec. 2015, doi: 10.1371/JOURNAL.PONE.0145708.
[102] V. V Kushnirov and V. V Kushnirov, "Yeast Functional Analysis Report Rapid and reliable protein extraction from yeast", doi: 10.1002/1097-0061(20000630)16:9.
[103] P. Hagel, "Structural insights into multidomain toxins - Mcf1 (Pl) and TcdA&TcdB (Cd)," Oct. 2018.
[104] C. J. Russo and L. A. Passmore, "Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens," *J Struct Biol*, vol. 193, no. 1, pp. 33–44, Jan. 2016, doi: 10.1016/J.JSB.2015.11.006.
[105] A. Rohou and N. Grigorieff, "CTFFIND4: Fast and accurate defocus estimation from electron micrographs," *J Struct Biol*, vol. 192, no. 2, pp. 216–221, Nov. 2015, doi: 10.1016/J.JSB.2015.08.008.
[106] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker, "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination," *Nature Methods 2017 14:3*, vol. 14, no. 3, pp. 290–296, Feb. 2017, doi: 10.1038/nmeth.4169.
[107] T. Wagner *et al.*, "SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM," *Commun Biol*, vol. 2, no. 1, pp. 1–13, 2019, doi: 10.1038/s42003-019-0437-z.
[108] T. Moriya *et al.*, "High-resolution single particle analysis from electron cryo-microscopy images using SPHIRE," *Journal of Visualized Experiments*, vol. 2017, no. 123, pp. 1–11, 2017, doi: 10.3791/55448.
[109] J. Zivanov *et al.*, "New tools for automated high-resolution cryo-EM structure determination in RELION-3," *Elife*, vol. 7, pp. 1–22, 2018, doi: 10.7554/eLife.42166.
[110] R. Sanchez-Garcia, J. Gomez-Blanco, A. Cuervo, J. Carazo, C. Sorzano, and J. Vargas, "DeepEMhancer: a deep learning solution for cryo-EM volume post-processing," 2020, doi: 10.1101/2020.06.12.148296.

[111] T. C. Terwilliger, P. D. Adams, P. V. Afonine, and O. V. Sobolev, “A fully automatic method yielding initial models from high-resolution electron cryo-microscopy maps,” *Nat Methods*, vol. 15, no. 11, p. 905, Nov. 2018, doi: 10.1038/S41592-018-0173-1.
[112] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel, and A. Bairoch, “ExPASy: The proteomics server for in-depth protein knowledge and analysis,” *Nucleic Acids Res*, vol. 31, no. 13, pp. 3784–3788, Jul. 2003, doi: 10.1093/NAR/GKG563.
[113] Z. Du *et al.*, “The trRosetta server for fast and accurate protein structure prediction,” *Nature Protocols 2021 16:12*, vol. 16, no. 12, pp. 5634–5651, Nov. 2021, doi: 10.1038/s41596-021-00628-9.
[114] J. Yang, I. Anishchenko, H. Park, Z. Peng, S. Ovchinnikov, and D. Baker, “Improved protein structure prediction using predicted interresidue orientations,” *Proc Natl Acad Sci U S A*, vol. 117, no. 3, pp. 1496–1503, 2020, doi: 10.1073/pnas.1914677117.
[115] J. Jumper *et al.*, “Highly accurate protein structure prediction with AlphaFold,” *Nature 2021 596:7873*, vol. 596, no. 7873, pp. 583–589, Jul. 2021, doi: 10.1038/s41586-021-03819-2.
[116] E. F. Pettersen *et al.*, “UCSF Chimera--a visualization system for exploratory research and analysis,” *J Comput Chem*, vol. 25, no. 13, pp. 1605–1612, Oct. 2004, doi: 10.1002/JCC.20084.
[117] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, “Features and development of Coot,” *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. Pt 4, pp. 486–501, 2010, doi: 10.1107/S0907444910007493.
[118] T. I. Croll, “ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps,” *Acta Crystallogr D Struct Biol*, vol. 74, no. 6, pp. 519–530, Jun. 2018, doi: 10.1107/S2059798318002425/IC5101SUP2.MP4.
[119] J. R. Lopéz-Blanco and P. Chacón, “iMODFIT: efficient and robust flexible fitting based on vibrational analysis in internal coordinates,” *J Struct Biol*, vol. 184, no. 2, pp. 261–270, Nov. 2013, doi: 10.1016/J.JSB.2013.08.010.
[120] E. C. Meng *et al.*, “UCSF ChimeraX: Tools for structure building and analysis,” *Protein Science*, vol. 32, no. 11, p. e4792, Nov. 2023, doi: 10.1002/PRO.4792.
[121] L. Holm, A. Laiho, P. Törönen, and M. Salgado, “DALI shines a light on remote homologs: One hundred discoveries,” *Protein Science*, vol. 32, no. 1, p. e4519, Jan. 2023, doi: 10.1002/PRO.4519.
[122] N. P. Bullen *et al.*, “An ADP-ribosyltransferase toxin kills bacterial cells by modifying structured non-coding RNAs,” *Mol Cell*, vol. 82, no. 18, pp. 3484-3498.e11, Sep. 2022, doi: 10.1016/J.MOLCEL.2022.08.015.
[123] D. Jurėnas, M. Rey, D. Byrne, J. Chamot-Rooke, L. Terradot, and E. Cascales, “Salmonella antibacterial Rhs polymorphic toxin inhibits translation through ADP-ribosylation of EF-Tu P-loop,” *Nucleic Acids Res*, vol. 50, no. 22, pp. 13114–13127, Dec. 2022, doi: 10.1093/NAR/GKAC1162.
[124] A. Cruz-Migoni *et al.*, “A Burkholderia pseudomallei toxin inhibits helicase activity of translation factor eIF4A,” *Science (1979)*, vol. 334, no. 6057, pp. 821–824, Nov. 2011, doi: 10.1126/SCIENCE.1211915/SUPPL_FILE/1211915.CRUZ.SOM.REVISED.PDF.
[125] A. Aminzadeh, C. E. Larsen, T. Boesen, and R. Jørgensen, “ High-resolution structure of native toxin A from Clostridioides difficile ,” *EMBO Rep*, vol. 23,

no. 1, Jan. 2022, doi: 10.15252/EMBR.202153597/SUPPL_FILE/EMBR202153597-SUP-0002-TABLEEV1.DOCX.

[126] J. Luo *et al.*, "TFPI is a colonic crypt receptor for TcdB from hypervirulent clade 2 C. difficile," *Cell*, vol. 185, no. 6, pp. 980-994.e15, Mar. 2022, doi: 10.1016/J.CELL.2022.02.010.

[127] L. Motlova, N. Klimova, R. Fiser, P. Sebo, and L. Bumba, "Continuous Assembly of β-Roll Structures Is Implicated in the Type I-Dependent Secretion of Large Repeat-in-Toxins (RTX) Proteins," *J Mol Biol*, vol. 432, no. 20, pp. 5696–5710, Sep. 2020, doi: 10.1016/J.JMB.2020.08.020.

[128] J. Snijder *et al.*, "Vitrification after multiple rounds of sample application and blotting improves particle density on cryo-electron microscopy grids," *J Struct Biol*, vol. 198, no. 1, p. 38, Apr. 2017, doi: 10.1016/J.JSB.2017.02.008.

[129] H. Stark, "GraFix: Stabilization of Fragile Macromolecular Complexes for Single Particle Cryo-EM," *Methods Enzymol*, vol. 481, no. C, pp. 109–126, Jan. 2010, doi: 10.1016/S0076-6879(10)81005-5.

[130] A. Krogh, B. Larsson, G. Von Heijne, and E. L. L. Sonnhammer, "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes," *J Mol Biol*, vol. 305, no. 3, pp. 567–580, Jan. 2001, doi: 10.1006/JMBI.2000.4315.

[131] M. Qa'Dan, L. M. Spyres, and J. D. Ballard, "pH-Induced Conformational Changes in *Clostridium difficile* Toxin B," *Infect Immun*, vol. 68, no. 5, pp. 2470–2474, May 2000, doi: 10.1128/IAI.68.5.2470-2474.2000.

[132] L. Wang, D. Wu, C. V. Robinson, H. Wu, and T. M. Fu, "Structures of a Complete Human V-ATPase Reveal Mechanisms of Its Assembly," *Mol Cell*, vol. 80, no. 3, pp. 501-511.e3, Nov. 2020, doi: 10.1016/J.MOLCEL.2020.09.029.

[133] A. L. Miles, S. P. Burr, G. L. Grice, and J. A. Nathan, "The vacuolar-ATPase complex and assembly factors, TMEM199 and CCDC115, control HIF1α prolyl hydroxylation by regulating cellular iron levels," *Elife*, vol. 6, Mar. 2017, doi: 10.7554/ELIFE.22693.

[134] A. Guna, N. Volkmar, J. C. Christianson, and R. S. Hegde, "The ER membrane protein complex is a transmembrane domain insertase," *Science (1979)*, vol. 359, no. 6374, pp. 470–473, Jan. 2018, doi: 10.1126/SCIENCE.AAO3099.

[135] M. J. Shurtleff *et al.*, "The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins," *Elife*, vol. 7, May 2018, doi: 10.7554/ELIFE.37018.

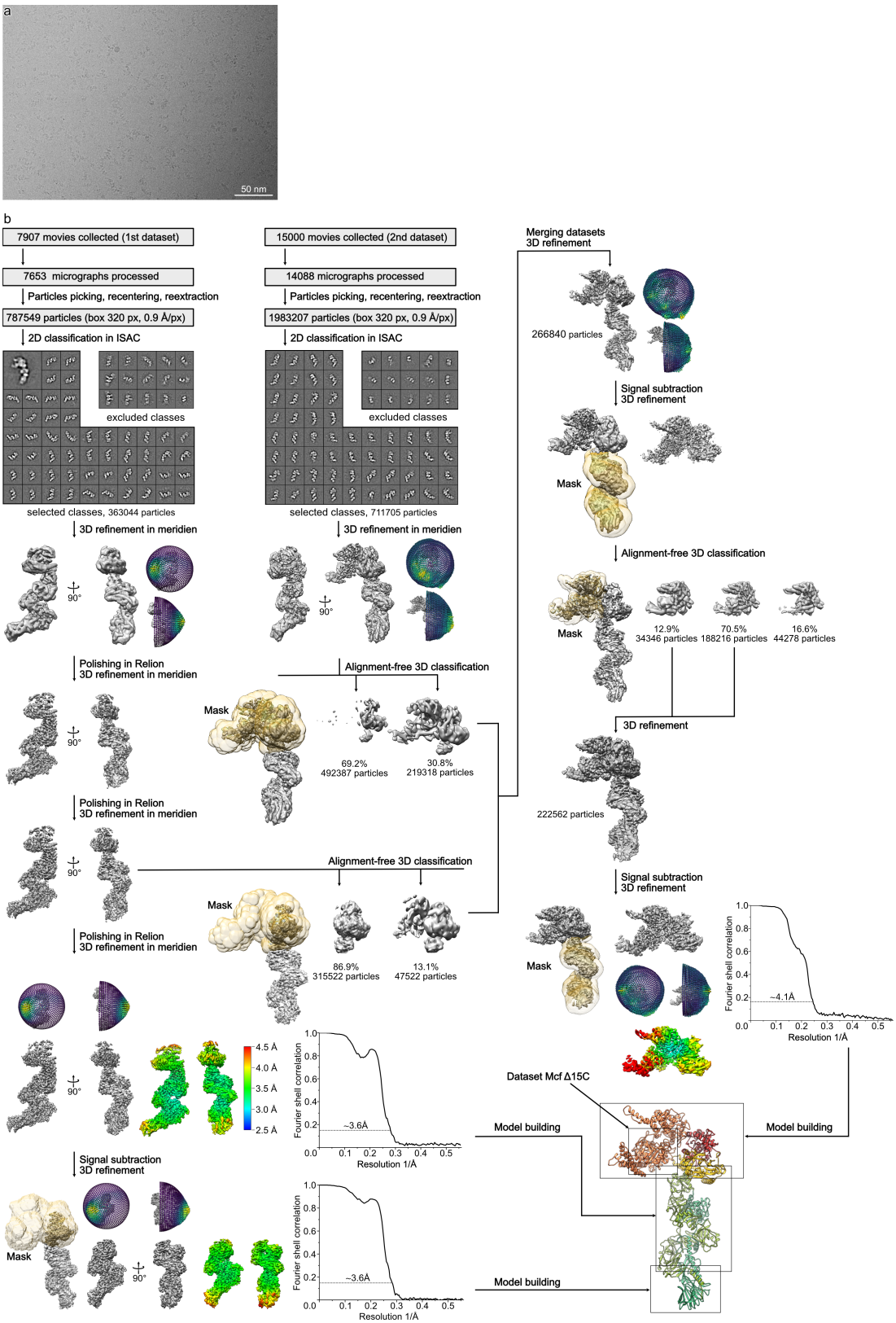
[136] A. S. Ramírez, J. Kowal, and K. P. Locher, "Cryo-electron microscopy structures of human oligosaccharyltransferase complexes OST-A and OST-B," *Science (1979)*, vol. 366, no. 6471, pp. 1372–1375, Dec. 2019, doi: 10.1126/SCIENCE.AAZ3505.

[137] M. P. Wilson *et al.*, "Active site variants in STT3A cause a dominant type I congenital disorder of glycosylation with neuromusculoskeletal findings," *Am J Hum Genet*, vol. 108, no. 11, pp. 2130–2144, Nov. 2021, doi: 10.1016/J.AJHG.2021.09.012.

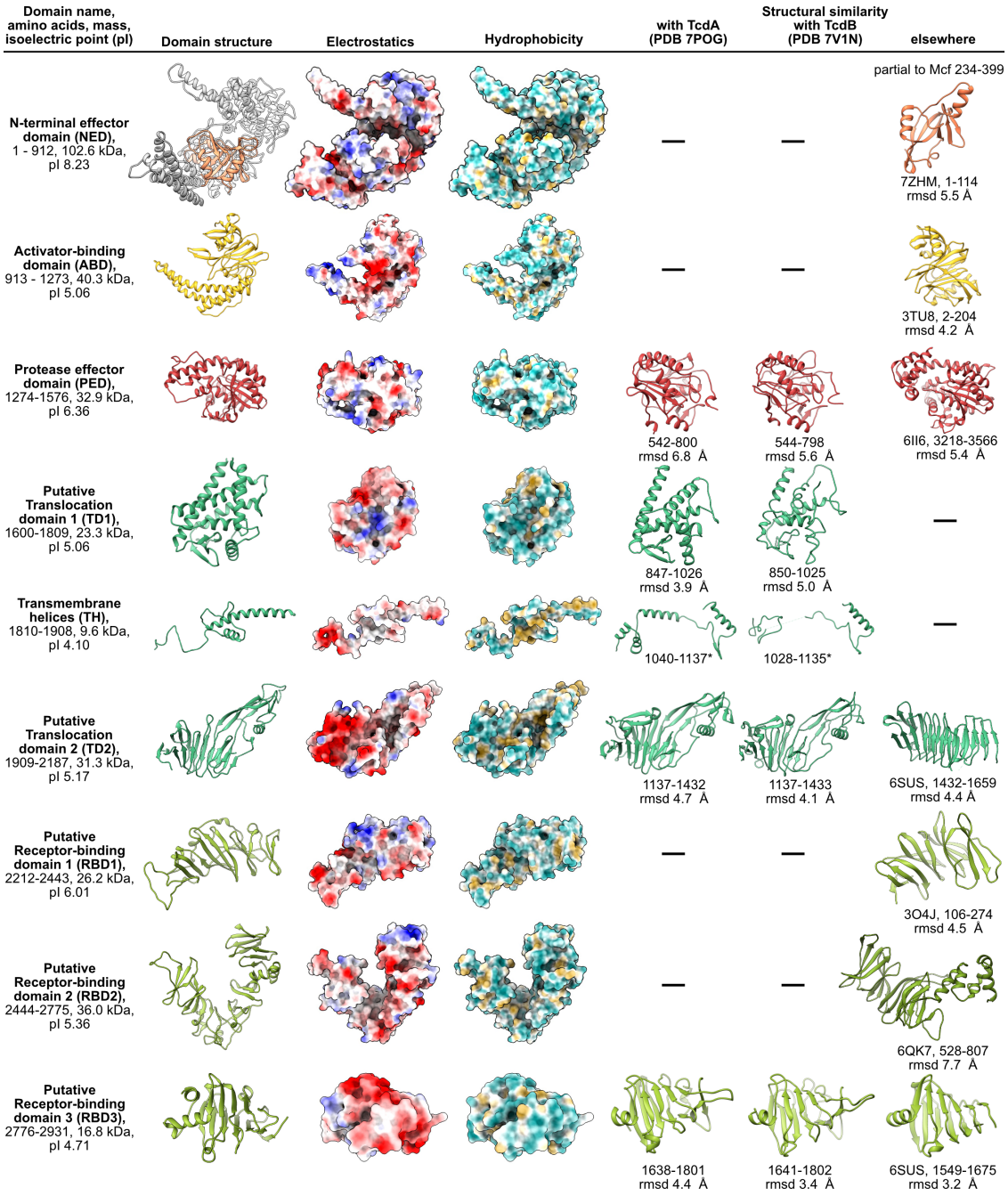
[138] S. Mohanty, B. P. Chaudhary, and D. Zoetewey, "Structural Insight into the Mechanism of N-Linked Glycosylation by Oligosaccharyltransferase," *Biomolecules 2020, Vol. 10, Page 624*, vol. 10, no. 4, p. 624, Apr. 2020, doi: 10.3390/BIOM10040624.

[139] O. Pfeil-Gardiner *et al.*, "Elemental mapping in single-particle reconstructions by reconstructed electron energy-loss analysis," *Nature Methods 2024 21:12*, vol. 21, no. 12, pp. 2299–2306, Oct. 2024, doi: 10.1038/s41592-024-02482-5.
[140] C. Casalou, A. Ferreira, and D. C. Barral, "The Role of ARF Family Proteins and Their Regulators and Effectors in Cancer Progression: A Therapeutic Perspective," *Front Cell Dev Biol*, vol. 8, Apr. 2020, doi: 10.3389/FCELL.2020.00217.
[141] A. Belyy *et al.*, "Mechanism of threonine ADP-ribosylation of F-actin by a Tc toxin," *Nat Commun*, vol. 13, no. 1, Dec. 2022, doi: 10.1038/S41467-022-31836-W.
[142] A. Belyy *et al.*, "Actin activates Pseudomonas aeruginosa ExoY nucleotidyl cyclase toxin and ExoY-like effector domains from MARTX toxins," *Nature Communications 2016 7:1*, vol. 7, no. 1, pp. 1–14, Dec. 2016, doi: 10.1038/ncomms13582.
[143] S. Büttner *et al.*, "A yeast BH3-only protein mediates the mitochondrial pathway of apoptosis," *EMBO Journal*, vol. 30, no. 14, pp. 2779–2792, Jul. 2011, doi: 10.1038/EMBOJ.2011.197/SUPPL_FILE/EMBJ2011197.REVIEWER_COMMENTS.PDF.
[144] X. Bogdanovic *et al.*, "A cysteine protease–like domain enhances the cytotoxic effects of the Photorhabdus asymbiotica toxin PaTox," *Journal of Biological Chemistry*, vol. 294, no. 3, pp. 1035–1044, 2019, doi: 10.1074/jbc.RA118.005043.
[145] D. Gallwitz, C. Donath, and C. Sander, "A yeast gene encoding a protein homologous to the human c-has/bas proto-oncogene product," *Nature*, vol. 306, no. 5944, pp. 704–707, 1983, doi: 10.1038/306704A0.
[146] M. Miaczynska, S. Lorenzetti, U. Bialek, R. M. Benito-Moreno, R. J. Schweyen, and A. Ragnini, "The yeast Rab escort protein binds intracellular membranes in vivo and in vitro," *J Biol Chem*, vol. 272, no. 27, pp. 16972–16977, Jul. 1997, doi: 10.1074/JBC.272.27.16972.
[147] V. Harmat *et al.*, "Structure and Catalysis of Acylaminoacyl Peptidase: CLOSED AND OPEN SUBUNITS OF A DIMER OLIGOPEPTIDASE," *Journal of Biological Chemistry*, vol. 286, no. 3, pp. 1987–1998, Jan. 2011, doi: 10.1074/JBC.M110.169862.
[148] M. I. Dauden *et al.*, "Molecular basis of tRNA recognition by the Elongator complex," *Sci Adv*, vol. 5, no. 7, 2019, doi: 10.1126/SCIADV.AAW2326/SUPPL_FILE/AAW2326_SM.PDF.
[149] Schrödinger LLC, "The PyMOL Molecular Graphics System, Version~1.8," Nov. 2015.

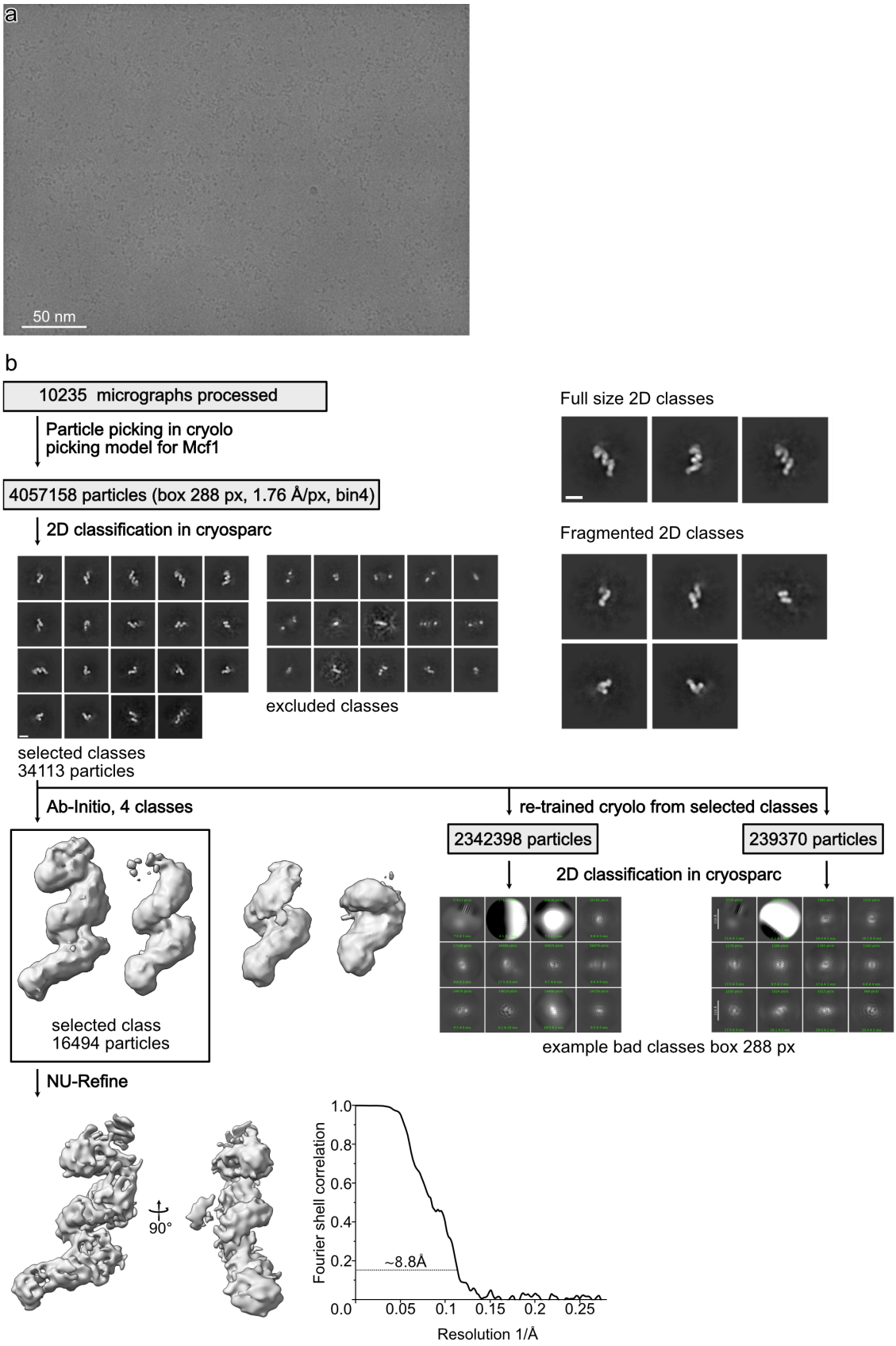
10. Appendix



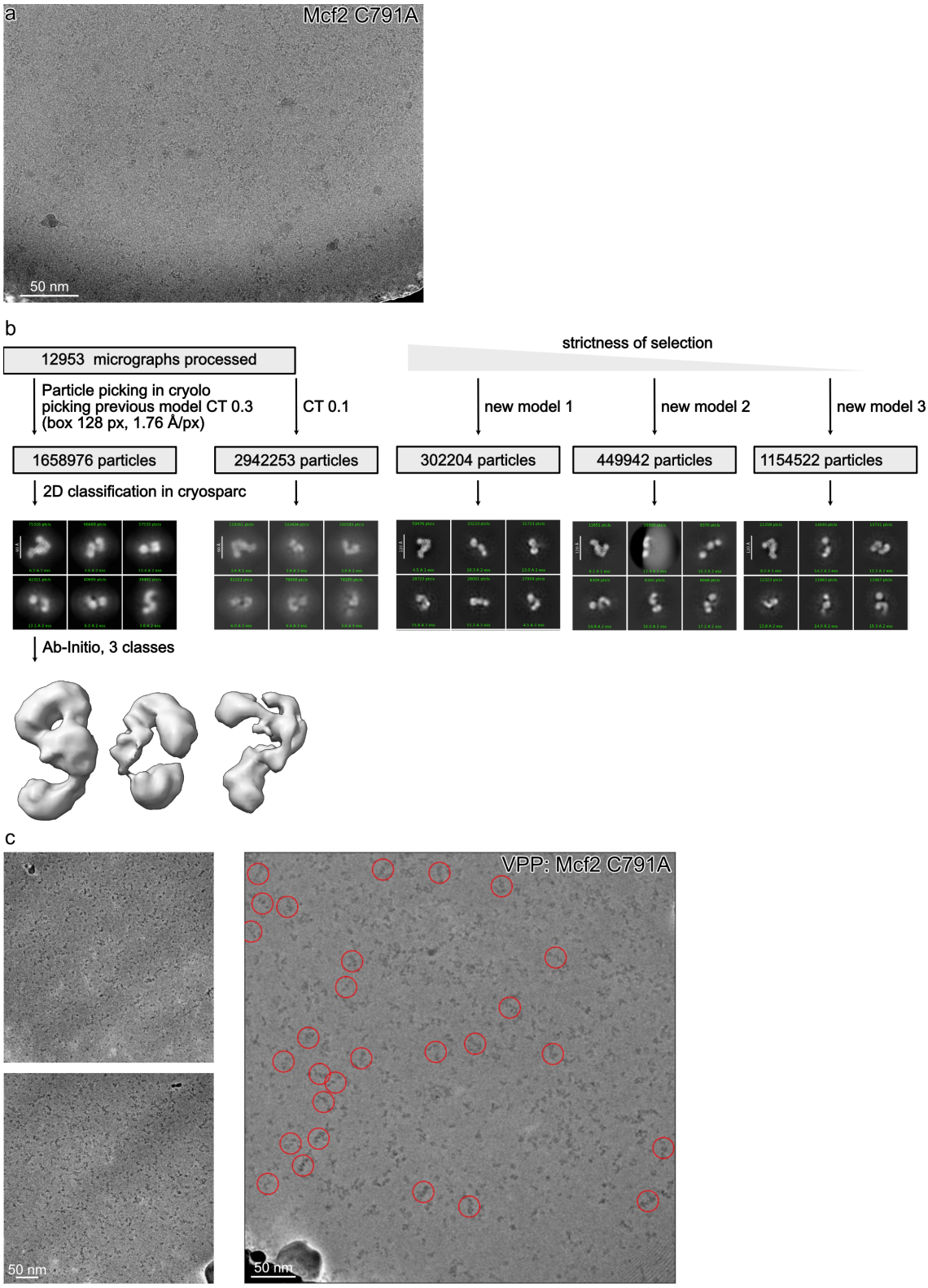
Supplementary Figure 1 Processing scheme of Mcf1: **a**) Micrograph of the first dataset. **b**) Detailed processing scheme of both merged datasets leading to all four maps that were used for model building. Figure from [43].



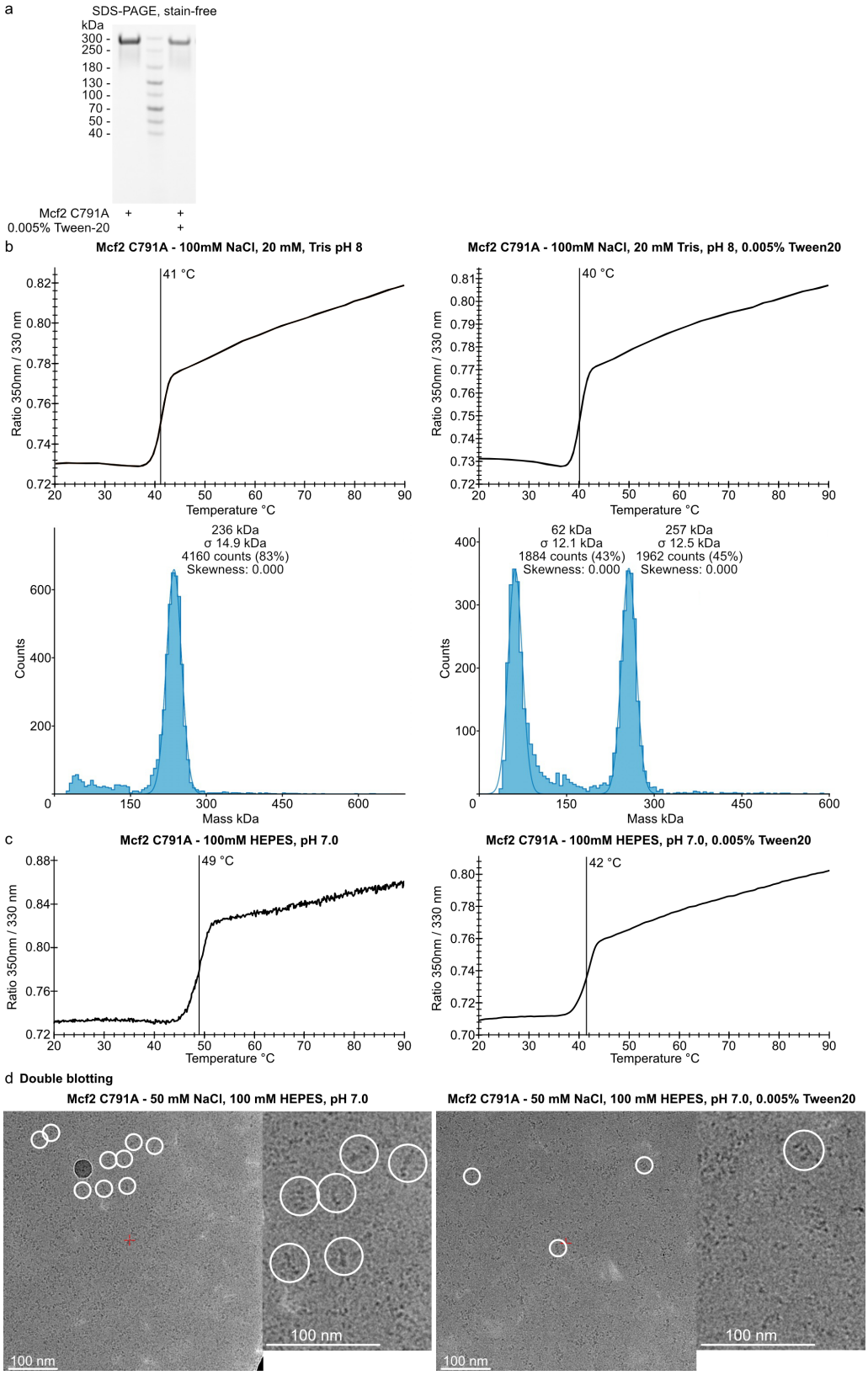
Supplementary Figure 2 Structural homologs to Mcf1: Single Mcf1 domains and their structural homologs as determined by a DALI search[121] (*Salmonella enterica* Rhs1 – pdb 7ZHM[123]; Burkholderia lethal factor (Blf1) – pdb 3TU8[124]; *C. difficile* TcdA – pdb 7POG[125]; *C. difficile* TcdB – pdb 7V1N[126]; *Bordella pertussis* RTX domain of adenylate cyclase toxin – pdb 6SUS[127]; *Aeropyrum pernix* Acylaminoacyl peptidase – pdb 3O4J[147]; *S. cerevisiae* Elongator complex – pdb 6QK7[148]). Rmsd was calculated by cealign in PyMOL[149]. Figure from [43].



Supplementary Figure 3 Processing scheme of the first Mcf2 dataset: **a**) Exemplary micrograph of the Mcf2 dataset and the processing scheme (**b**) highlighting 2D classes depicting Mcf2 fragments.

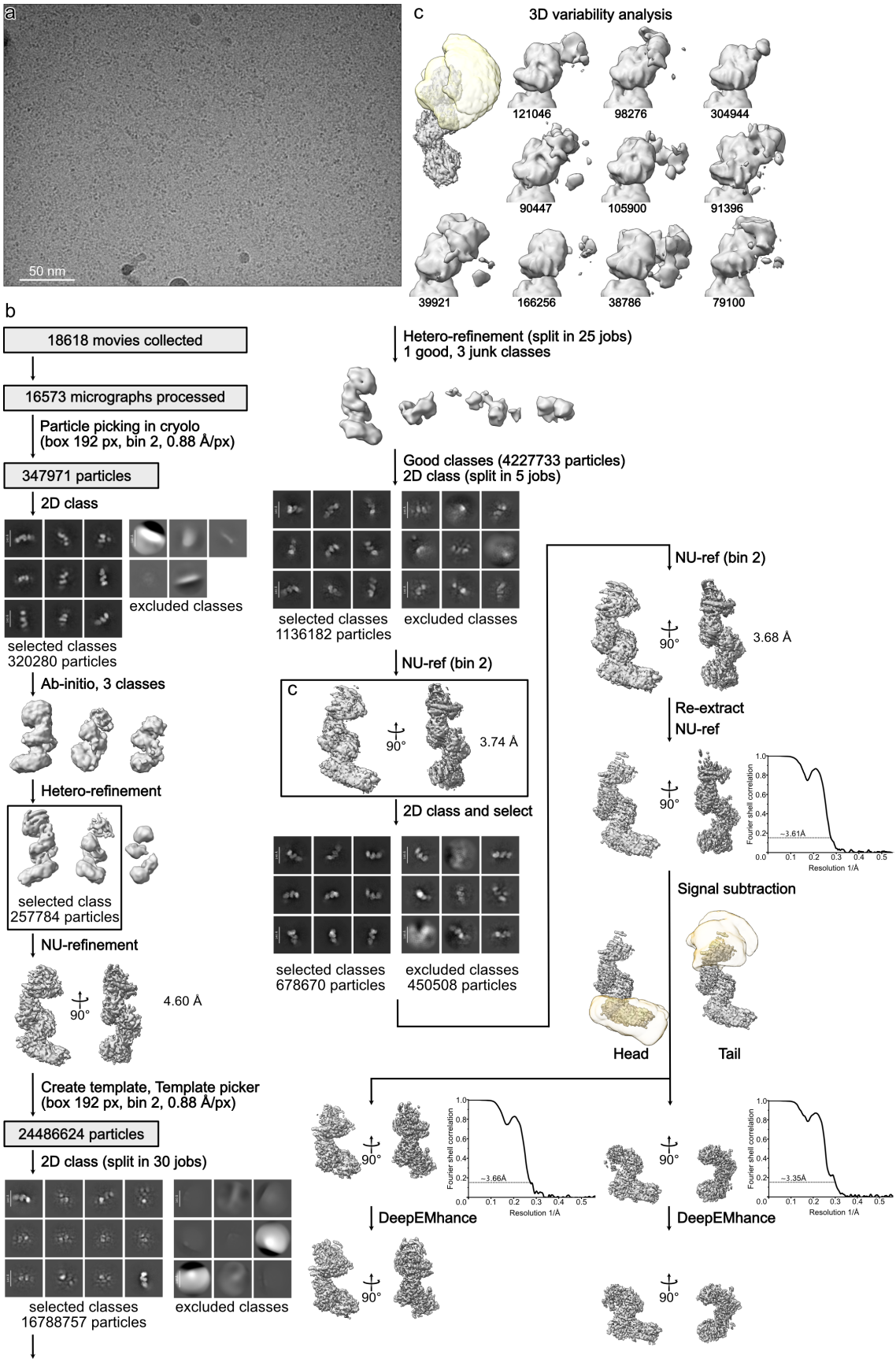


Supplementary Figure 4 Processing of the Mcf2 C791A dataset: a) Exemplary micrograph and processing scheme focused on the extensive particle picking strategies. C) Micrograph of Mcf2 C791A recorded with the Volta phase plate (VPP) to increase the contrast, highlighting full size Mcf2 particles.

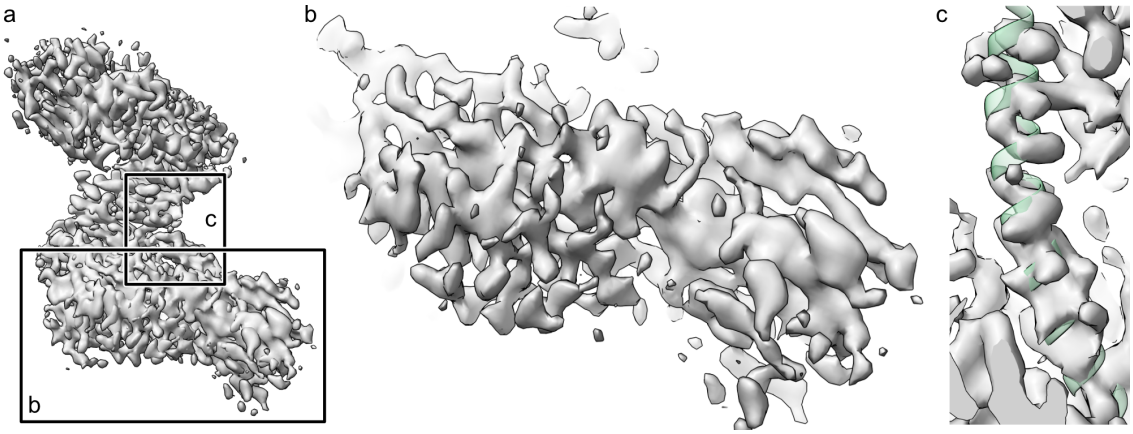


Supplementary Figure 5 Sample optimization Mcf2: **a**) SDS-PAGE of Mcf2 C791A in presence and absence of Tween-20 (0,005%) to observe stability in plunging conditions. **b**) Melting curves recorded with nanoDSF (top) and mass photometry measurements (bottom) of Mcf2 C791A in absence (left) and presence (right) of 0,005% Tween20. **c**) NanoDSF measurements of the improved buffer after performing a buffer screen in absence (left) and presence (right) of 0,005% Tween20. **d**) Exemplary micrographs of Mcf2

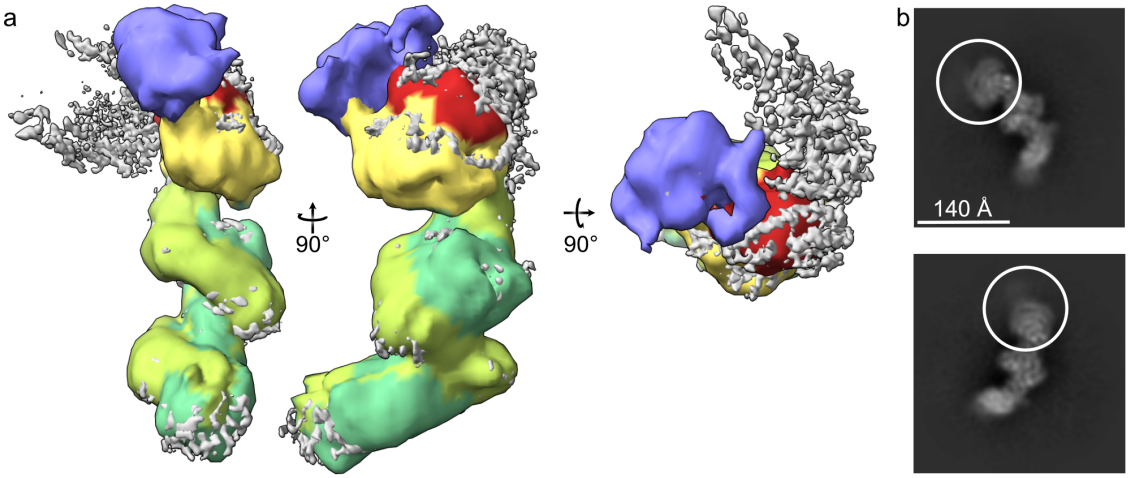
C971A in optimized buffer conditions without tween on grids prepared by double blotting. White circles highlight full size Mcf2 particles whose number is drastically reduced in the presence of tween.



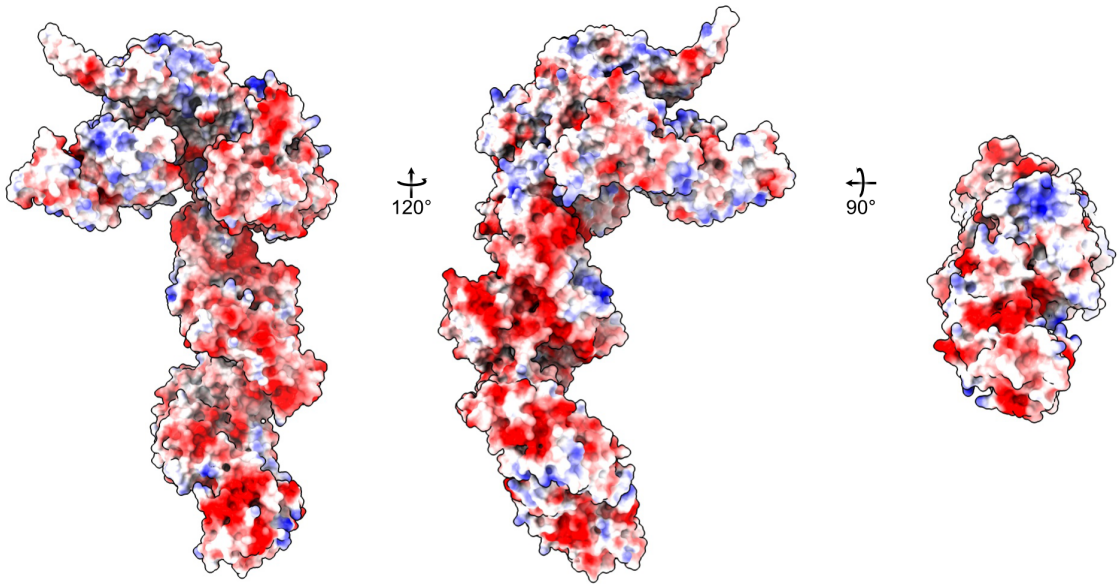
Supplementary Figure 6 Processing scheme of buffer optimized and double blotted Mcf2 C791A: **a**) Exemplary micrograph and processing scheme (**b**) of the Mcf2 C791A double blot dataset. **c**) 3D variability analysis highlighting the flexibility of the N-terminal effector domain in 10 different classes.



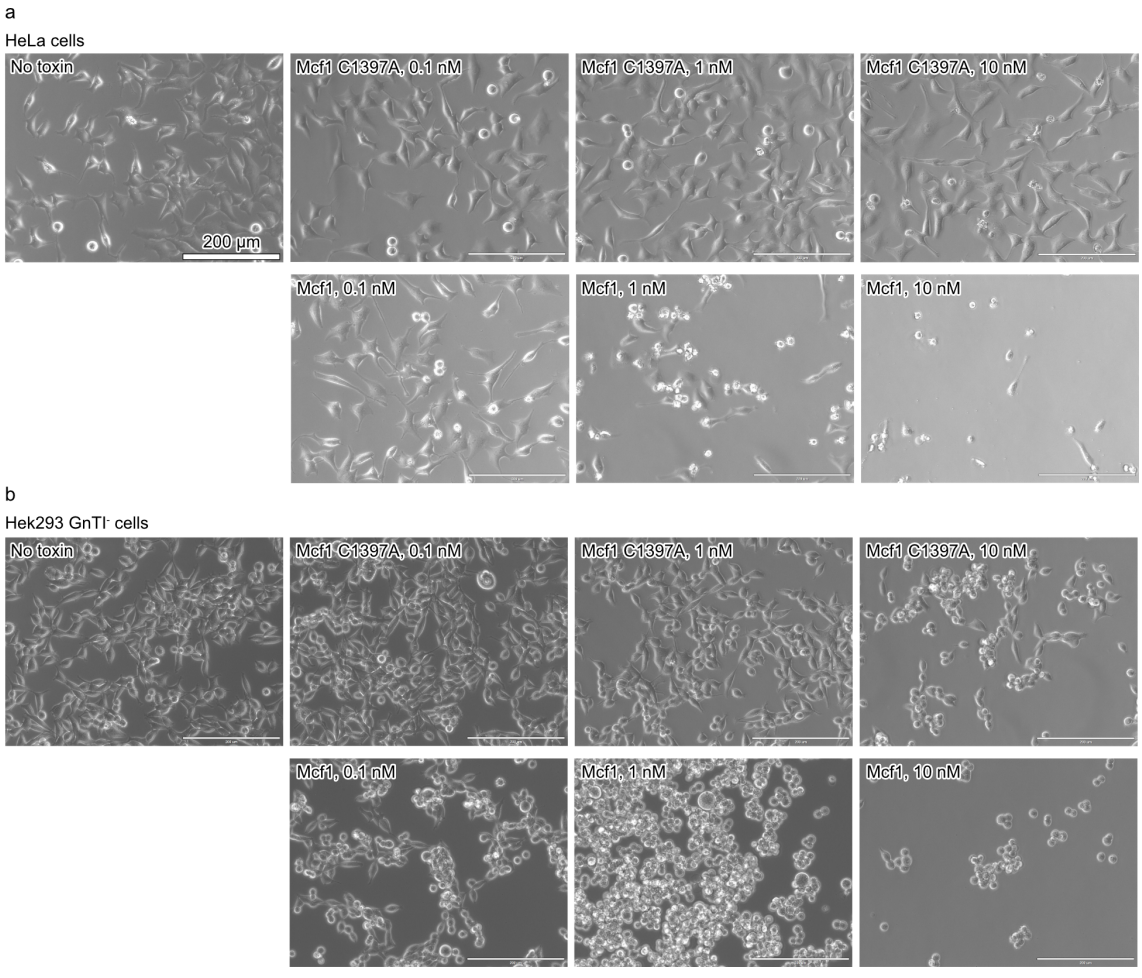
Supplementary Figure 7 Details of the signina subtracted tail cryo-EM density: **a**) Overview of the density and highlights depicting merged β-sheets (**b**) and an interrupted α-helix (**c**).



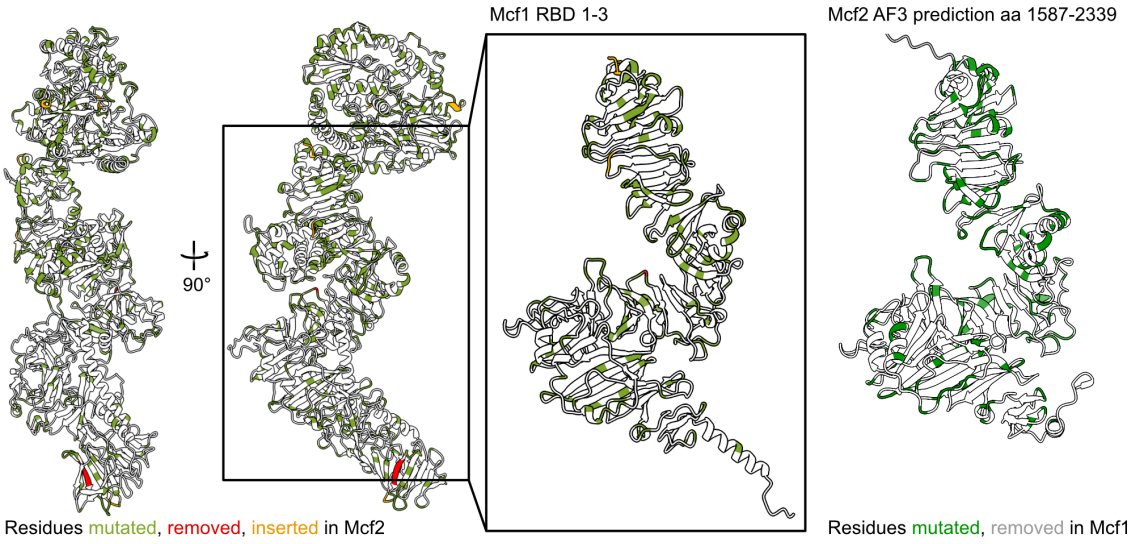
Supplementary Figure 8 Difference in N-terminal effector domain (NED) between Mcf1 and Mcf2: **a**) Composite cryo-EM map of Mcf1 (grey) superimposed with a Mcf2 (colored) density from the 3D variability analysis in different orientations to visualize the change position of the Mcf2 NED (violet). **b**) 2D class averages showing the flexibility of the head and NED of Mcf2 (white circle) as a fuzzy signal.



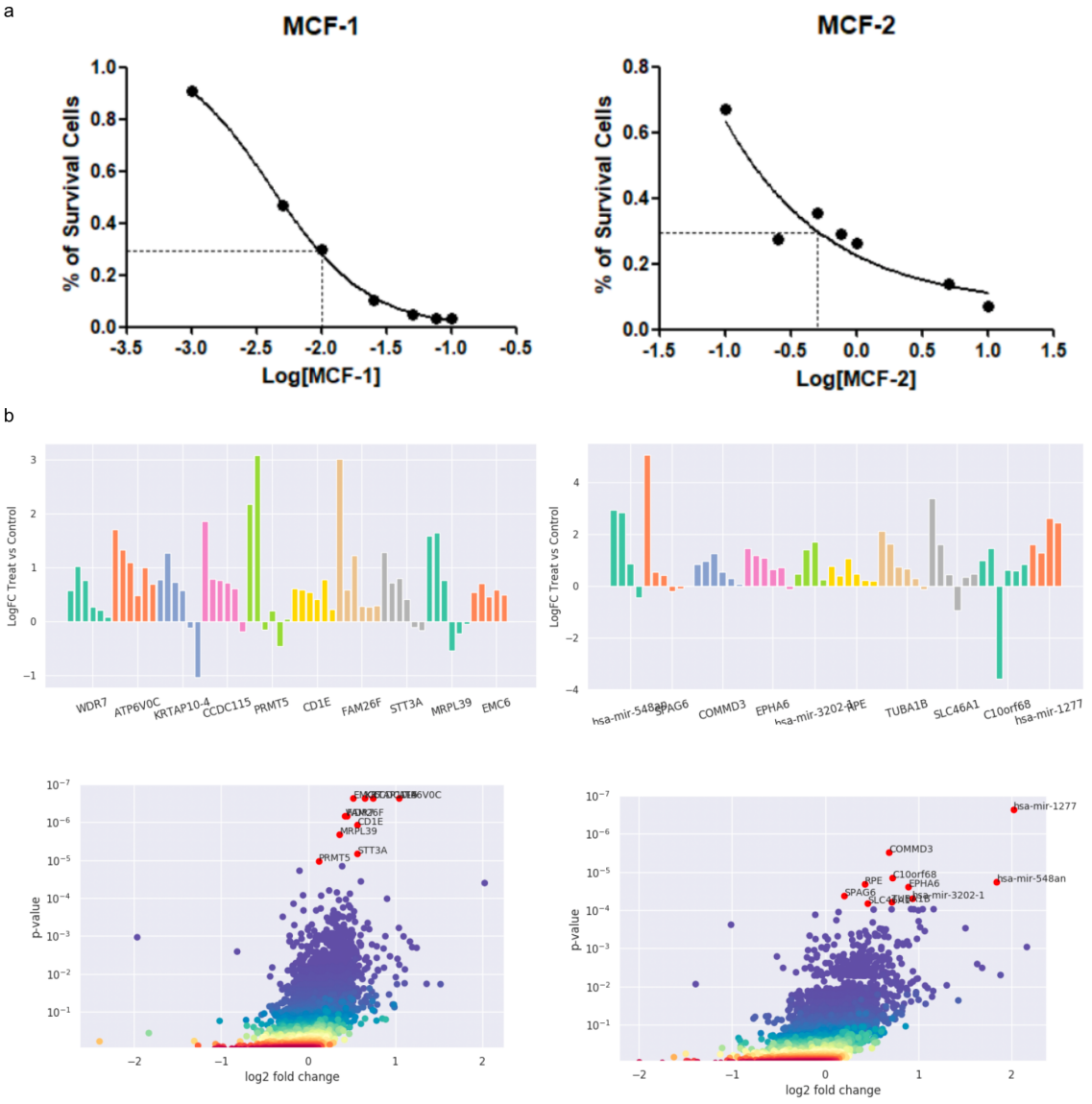
Supplementary Figure 9 Coulomb surface of Mcf1: Surface of the Mcf1 model colored according to the electrostatic Coulomb potential (-10 kcal mol^{-1}, red; +10 kcal mol^{-1}, blue). Figure from [43].



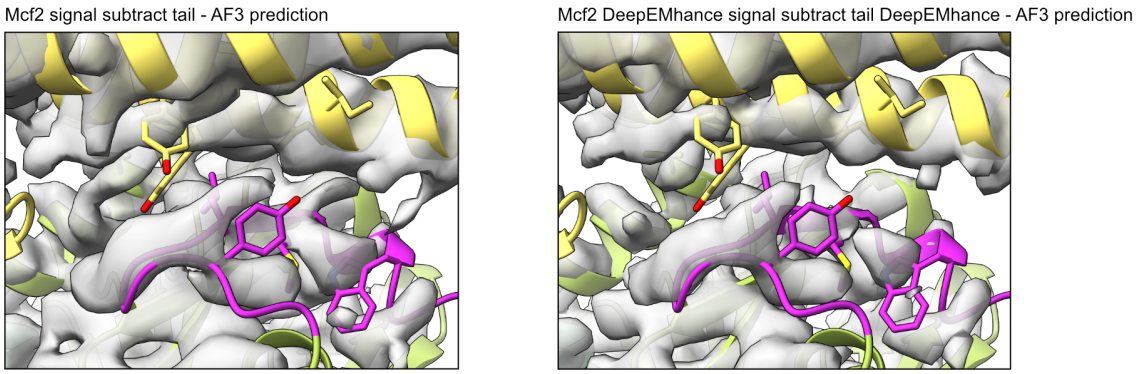
Supplementary Figure 10 Mcf1 cell intoxication: HeLa and Hek 293 GnTI- cells intoxicated with Mcf1 at the indicated concentration after 24 h.



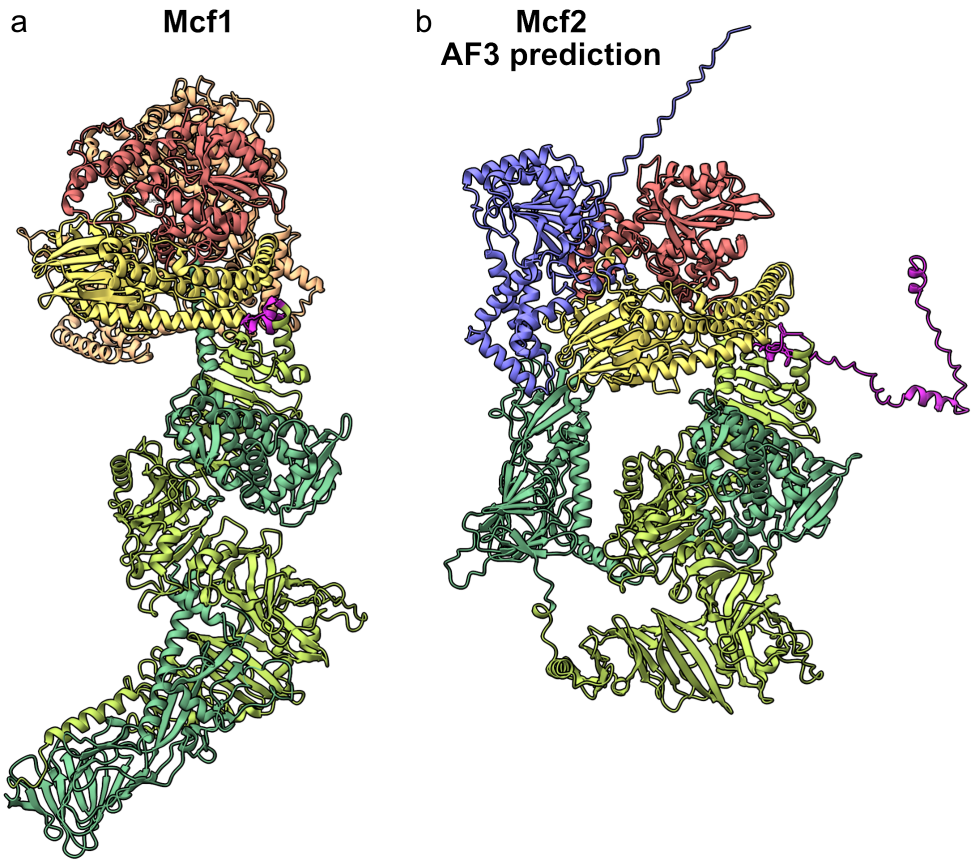
Supplementary Figure 11 Sequence differences of the receptor binding domains of Mcf1 and Mcf2: Differences in between the Mcf1 and Mcf2 sequence mapped to the structure of Mcf1 or alphafold 3 prediction of Mcf2 (green - mutated; red - deleted; yellow - inserted amino acids).



Supplementary Figure 12 Mcf1 and Mcf2 CRISPR screen: **a**) Toxicity of Mcf1 and Mcf2 on Hek293T cells. **b**) Positive selection results for the CRISPR screen of Mcf1 (left) and Mcf2 (right). Figures provided by Creative Biogene.



Supplementary Figure 13 Mcf2 C-terminus: Rigid body fit of the C-terminal (magenta) region of the Mcf2 alphafold3 prediction into the signal subtracted map (left) and the corresponding DeepEMhanced version (right).



Supplementary Figure 14 Mcf1 structure (a) and Mcf2 alphafold 3 (AF3) prediction colored according to their domains (orange, N-terminal effector domain NED of Mcf1; violet NED of Mcf2; yellow, activator binding domain ABD; red, protease effector domain PED; dark green, translocation domains 1,2 TD1,2 and transmembrane helix TH; light green, putative receptor binding domains 1-3 RBD1-3; magenta, C-terminus).

a

ICP-MS Measurement 1

Isotope	Intensity AVG	Intensity STD	Intensity RSD	Concentration AVG	Concentration STD	Concentration RSD
	[cps]	[cps]	[%]	[ng/mL]	[ng/mL]	[%]
Mn55(LR)	29927	4549	15,2	-1,1	0,10	8,55
Fe54(LR)	121538	783	0,64	-35,7	0,19	0,53
Ni58(LR)	1425714	43388	3,04	41,9	1,86	4,45
Ni60(LR)	513817	14094	2,74	41,6	1,55	3,73
Ni62(LR)	14180120	89606	0,63	-23,2	2,33	10,03
Cu63(LR)	13958373	175057	1,25	-27,7	10,38	37,43
Cu65(LR)	75668	900	1,19	4,3	0,09	2,05
Zn64(LR)	500181	3924	0,78	**73,8**	0,63	0,85
Zn66(LR)	264854	6868	2,59	**76,5**	2,03	2,66
Zn68(LR)	180392	3357	1,86	**74,4**	1,43	1,92

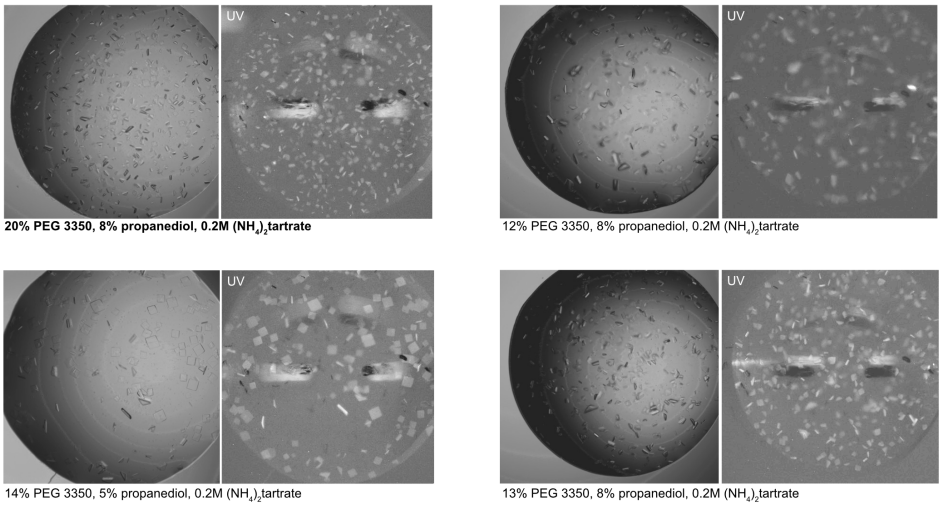
ICP-MS Measurement 2

Isotope	Intensity AVG	Intensity STD	Intensity RSD	Concentration AVG	Concentration STD	Concentration RSD
	[cps]	[cps]	[%]	[ng/mL]	[ng/mL]	[%]
Mn55(LR)	21814	2399	11	-1,3	0,05	3,91
Fe54(LR)	119517	2590	2,17	-36,2	0,63	1,73
Ni58(LR)	1434269	101364	7,07	42,3	4,35	10,3
Ni60(LR)	524099	41905	8	42,8	4,62	10,81
Ni62(LR)	13629241	743928	5,46	-8,9	19,32	217,44
Cu63(LR)	13766569	1353662	9,83	-16,4	80,27	490,63
Cu65(LR)	73080	2057	2,82	4,0	0,20	4,97
Zn64(LR)	524126	33451	6,38	**77,6**	5,33	6,87
Zn66(LR)	272110	17244	6,34	**78,7**	5,11	6,49
Zn68(LR)	185240	12663	6,84	**76,4**	5,38	7,04

With an average Zn conc of 76.2 ng/ml (M=65,38 g/mol)
With a Mcf1 conc of 0.72 mg/ml (M=329100 g/mol)
c(Zn) = 1.165 µM
c(Mcf1) = 2.188 µM
Zn : Mcf1
1 : 1.9

b

Hanging drop crystallization



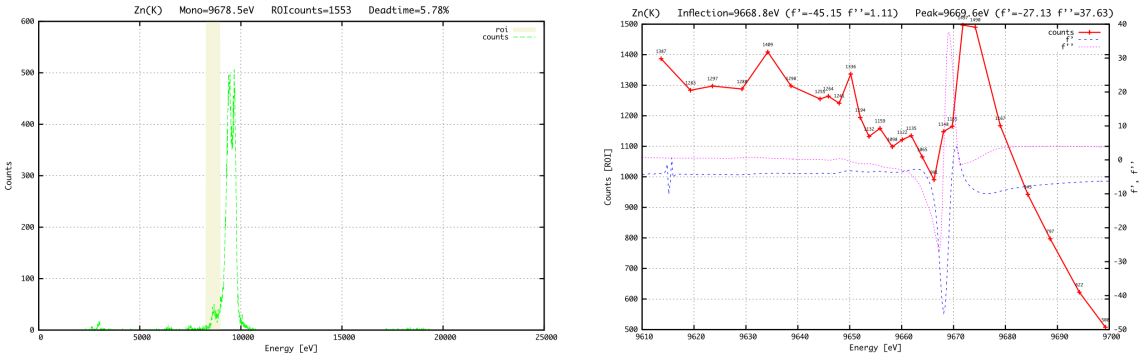
20% PEG 3350, 8% propanediol, 0.2M $(NH_4)_2$tartrate

12% PEG 3350, 8% propanediol, 0.2M $(NH_4)_2$tartrate

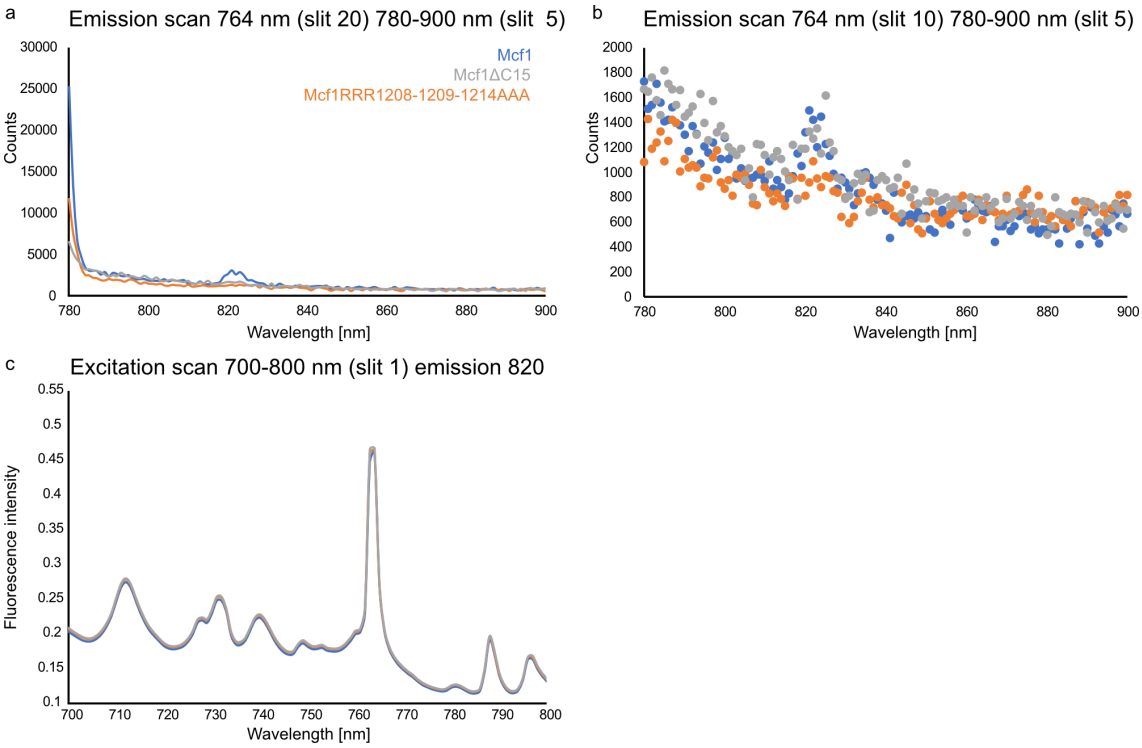
14% PEG 3350, 5% propanediol, 0.2M $(NH_4)_2$tartrate

13% PEG 3350, 8% propanediol, 0.2M $(NH_4)_2$tartrate

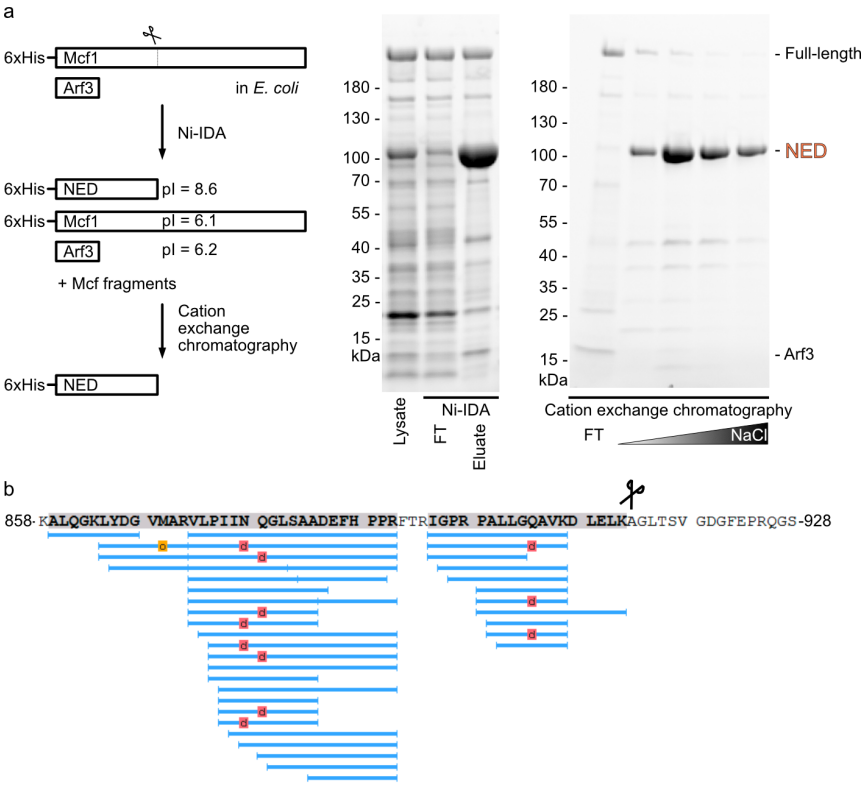
c



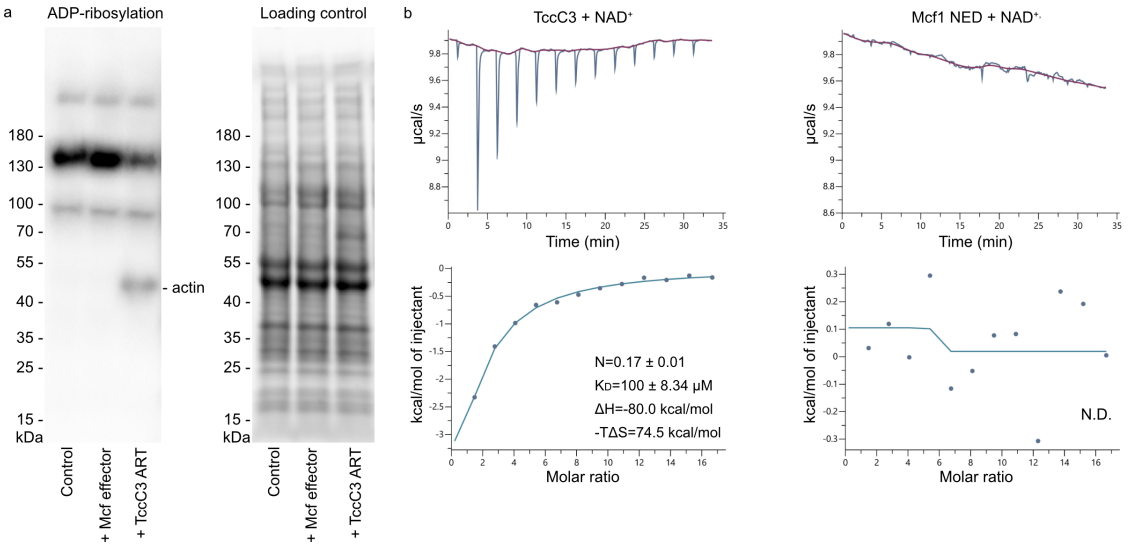
Supplementary Figure 15 Characterization of the head stabilizing element (HSE): **a**) Results of two independent ICP-MS measurements detecting the presence of zinc. **b**) Crystals obtained by hanging drop crystallization and (**c**) the X-ray fluorescence spectra (recorded by Dr. Raphale Gasper-Schoenenbruecher).



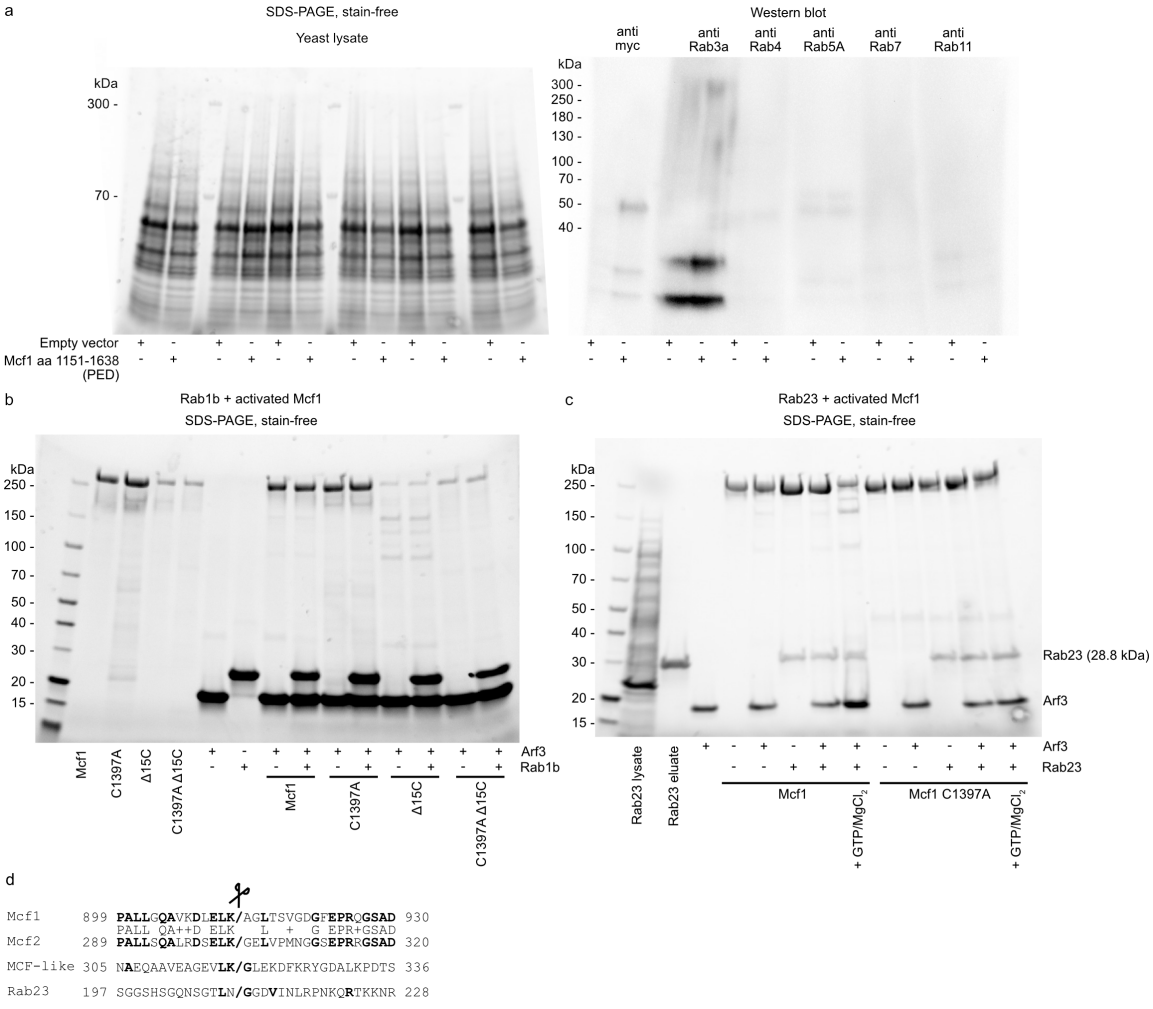
Supplementary Figure 16 UV-Vis spectroscopy of the HSE: (**a**) (**b**) Emission and (**c**) excitation scans of Mcf1 and Mcf1 mutants missing the HSE (Mcf1ΔC15, Mcf1RRR1208-1209-1214AAA).



Supplementary Figure 17 Mcf1 cleavage site mass spectrometry: **a**) His-Mcf1 and His-Arf3 were co-expressed in *E. coli* to purify the NED fragment using the His-tag and a cation exchange column. SDS-PAGE analysis of the sample during the purification process. Gel pieces were cut and sent for mass spectrometry analysis. **b**) Sequence coverage of MS analysis revealing the C-terminal end of the NED fragment. Figure from [43].

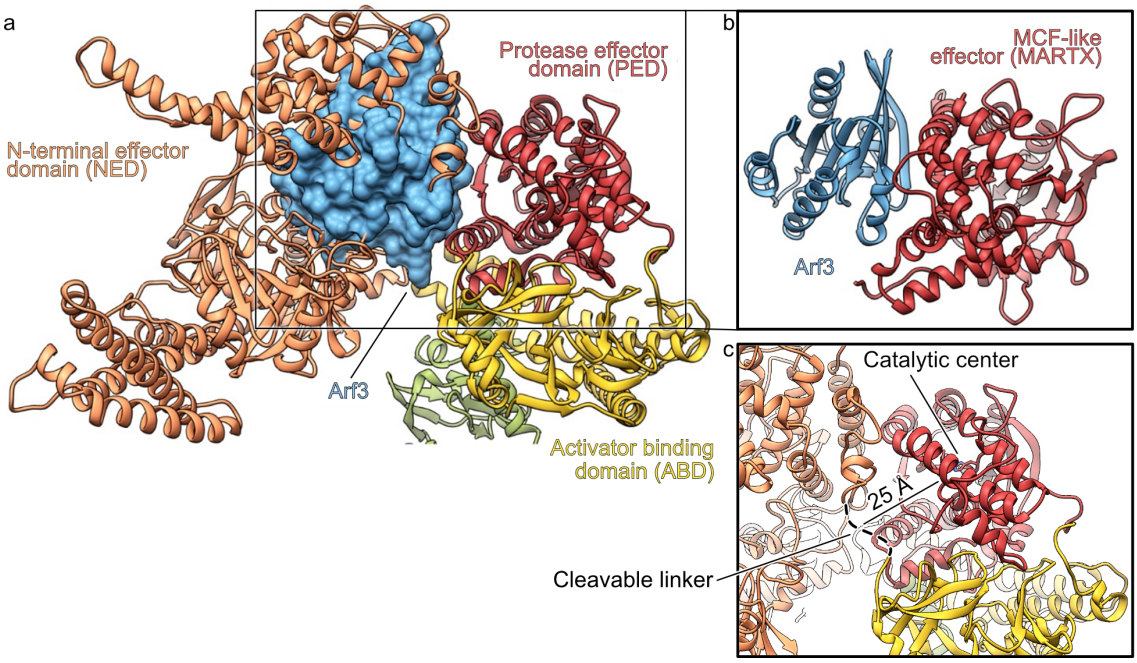


Supplementary Figure 18 ADP-ribosylating properties of the NED: **a**) Anti-ADP-ribose western blot of insect cell lysate incubated with the N-terminal effector domain (NED) or the ADP-ribosyltransferase TccC3. **b**) ITC measurements of TccC3 and Mcf1 to determine their affinity to NAD^+. Figure adapted from [43].

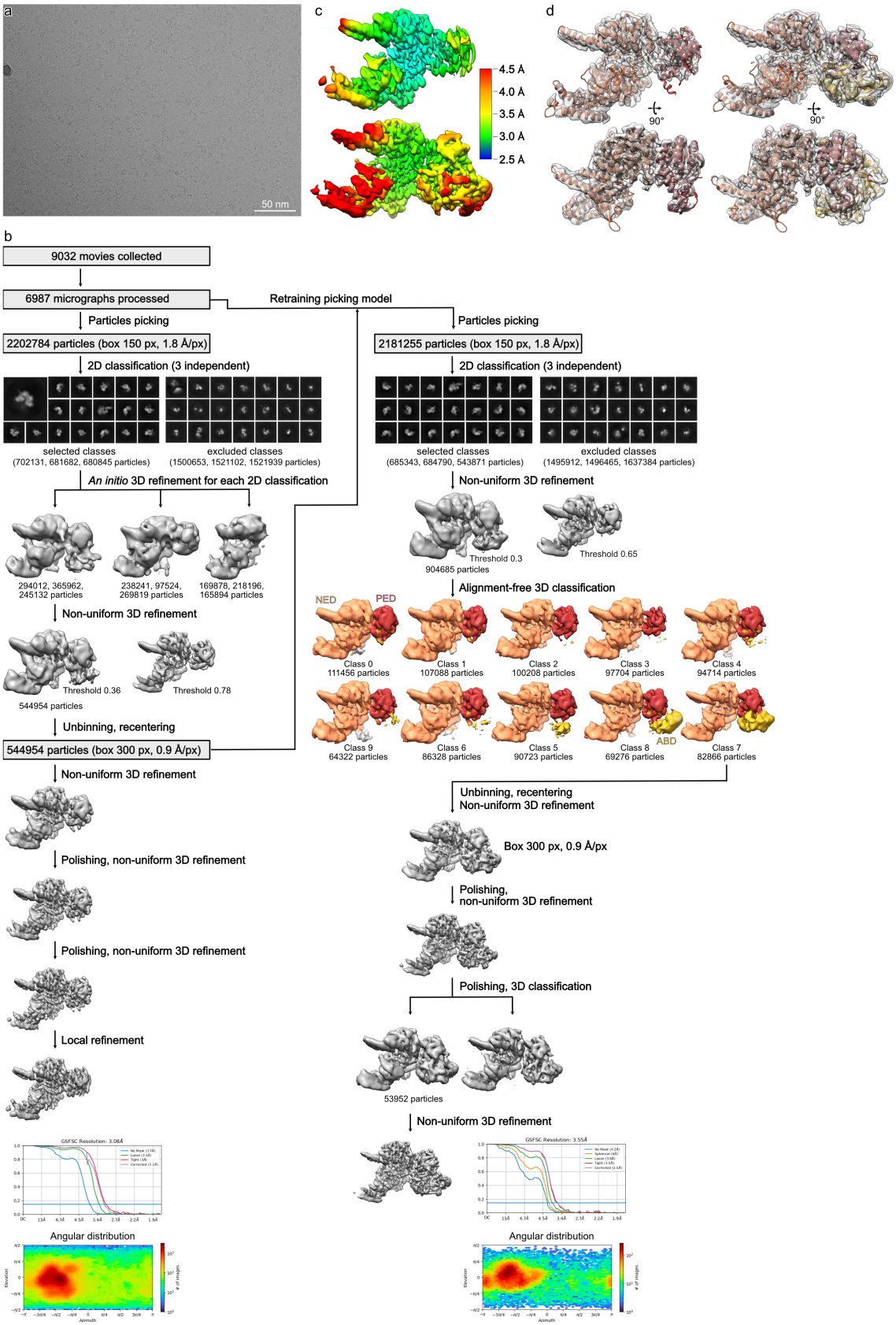


Supplementary Figure 19 Rab protein cleavage analysis: **a**) Yeast lysate of NED expressing yeast cells and western blot against Rab proteins (anti-Rab3a, -Rab4, -Rab5A, -Rab7, -Rab11). **b**) SDS-PAGE of Rab1b incubated with Arf3 activated Mcf1, Mcf1C1397A, Mcf1ΔC15 or Mcf1ΔC15 C1397A. **c**) SDS-PAGE of Rab23 incubated with Arf3 activated Mcf1 or Mcf1C1397A. **d**) Sequence alignment of the Mcf1 and Mcf2

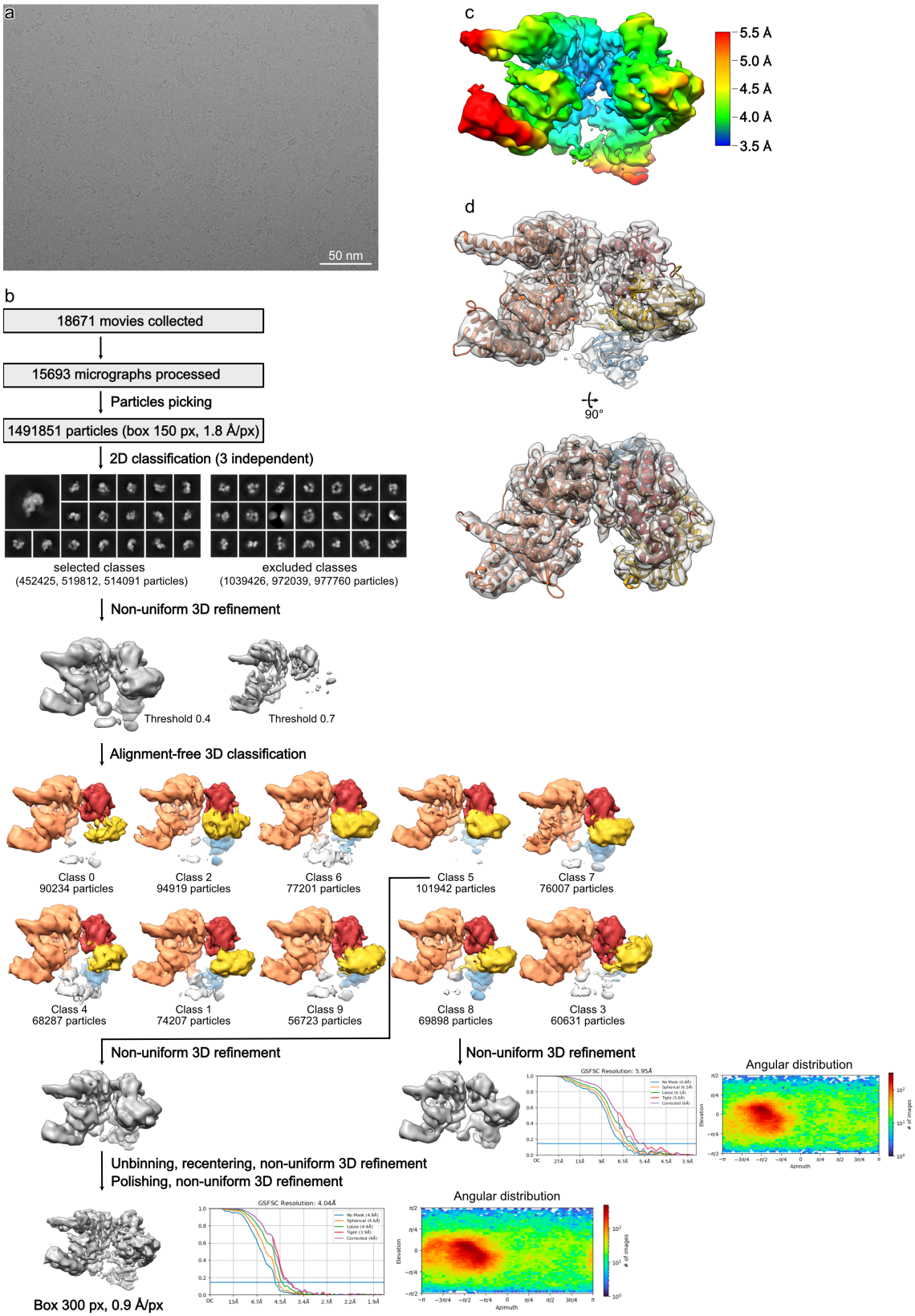
NED/ABD cleavage site with the internal cleavage site of the MCF-like effector and its targeted cleavage site on Rab23 [59].



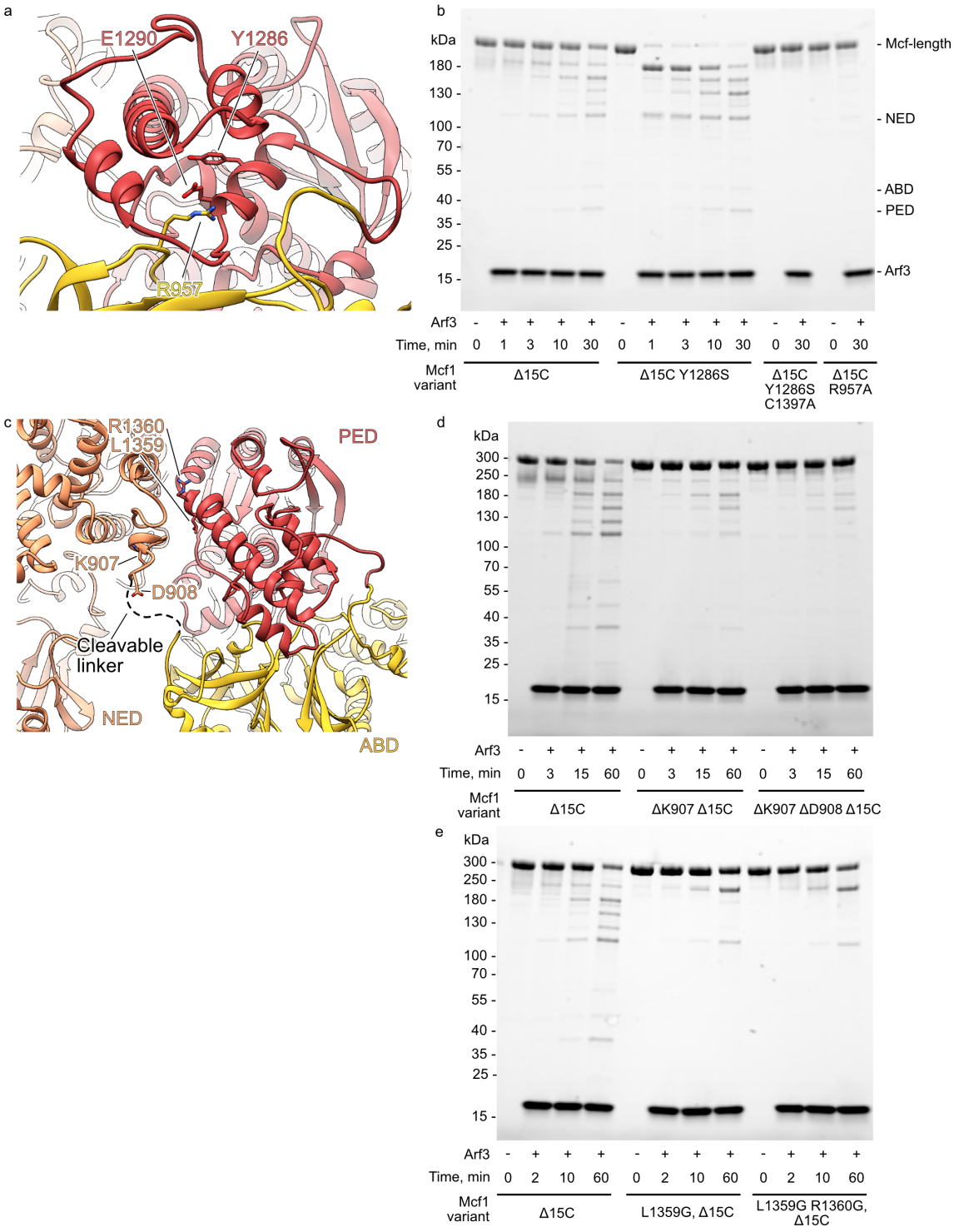
Supplementary Figure 20 Arf3 binding sites of the MCF-like effector compared to Mcf1: **a**) Superimposition of the Mcf1 PED with the structure (**b**) of the MCF-like effector in complex with Arf3 (blue surface, pdb 6II6[38]) showing the clash between NED and Arf3 when binding at the same site. **c**) Close-up of the Mcf1 head depicting the distance between the active site of the PED and the biochemically determined NED/ABD cleavage site. Figure adapted from [43].



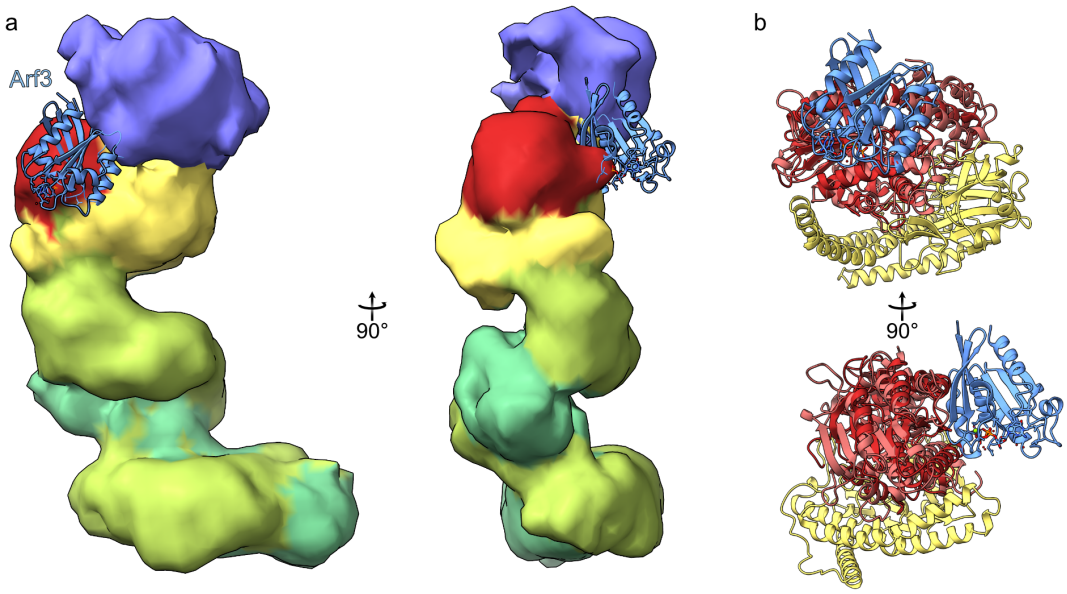
Supplementary Figure 21 Processing scheme of Mcf1ΔC15 C1397A: **a**) Exemplary micrograph of the Mcf1ΔC15 dataset. **b)** Processing workflow and depiction of the local resolution (**c**) as well as the model inside the cryo-EM density (**d**). The Mcf1ΔC15 C1397A mutant was chosen to increase the stability of the particles. Figure from [43].



Supplementary Figure 22 Processing scheme of Mcf1ΔC15 C1397A-Arf3: **a**) Exemplary micrograph of the Mcf1ΔC15 C1397A-Arf3 dataset. **b**) Processing workflow and depiction of the local resolution (**c**) as well as the model inside the cryo-EM density (**d**). Figure form [43].



Supplementary Figure 23 Mutations of the Mcf1 activation mechanism: Structure of the hook-loop interaction (**a**) highlighting residues that were mutated to observe their cleavage efficiency on SDS-PAGE (**b**). **c**) Structure of the cleavable linker between (NED/ABD) and the PED highlighting residues that were used to truncate the linker (**d**) or destabilize the helix that anchors the hook-loop interaction. To imitate the post-translocation state, all mutations were introduced to the truncated Mcf1ΔC15. Figure adapted from [43].



Supplementary Figure 24 Arf binding site of MARTX MCF-like effector is accessible for Mcf2: **a**) 3D variability density of Mcf2 superimposed with the Arf3 structure (blue) at the binding site of the MCF-like effector. **b**) Structure of Mcf ABD-PED superimposed with the MCF-like effector-Arf3 complex (6II6 [38]) that was fitted into the 3D variability density.

Table S1 **Cryo-EM data collection and processing statistics – Mcf1**

Sample	**Mcf1**	**Mcf1ΔC15C1397A**	**Mcf1ΔC15C1397A + Arf3**
Microscope	Titan Krios G3		
Voltage (kV)	300		
Defocus range (μm)	-1 to -2.5		
Camera	Gatan K3 in super resolution mode		
Pixel size (Å)	physical pixel size 0.9, super resolution mode 0.45		
Total electron dose (e/Å^2)	71/69	60.85	62.6
Exposure time (sec)	3	3.5	2
Frames per movie	60	60	60
Number of movies	21741	9032	18671
3D Refinement			
Number of particles	363044 for the tail region; 219318 for the head region	53952	101942
Final resolution (Å)	3.6	3.5	4.0

Table S2: **Cryo-EM data collection and processing statistics – Mcf2**

Sample	**Mcf2**	**Mcf2 C791A**	**Mcf2 C791A double blot**
Microscope	Titan Krios G2		
Voltage (kV)	300		
Defocus range (μm)	-1 to -2.9		
Camera	Gatan K3 in super resolution mode		
Pixel size (Å)	physical pixel size 0.9, super resolution mode 0.45		

Total electron dose (e/Å^2)	63.9	82.4	59.7
Exposure time (sec)	3	2.4	4
Frames per movie	60	60	60
Number of movies	10253	12953	18618

Table S3 **Model statistics**

Sample	**Mcf1**	**Mcf1ΔC15C1397A**	**Mcf1ΔC15C1397A + Arf3**
Non-hydrogen atoms	22783	12619	13714
Molprobity score	1.52	1.51	1.68
Clashscore	3.44	4.21	6.41
Cβ deviations >0.25Å	0	0	0
Bad bonds	0	0	0
Bad angles	0	0	0
Poor rotamers (%)	0	0	0
Favored rotamers (%)	99.71	99.47	99.66
Ramachandran favored (%)	94.40	95.61	95.32
Ramachandran allowed (%)	5.57	4.39	4.68
Ramachandran outliers (%)	0.03	0	0

Table S4 **CRISPR knock-out screen positive selection Mcf1**

Gene	**pos|score**	**pos|pvalue**	**pos|fdr**	**pos|rank**	**pos|lfc**	**pos|good grnas**
ATP6V0C	7.23E-09	2.28E-07	0.001238	1	1.0452	6
EMC6	5.99E-07	2.28E-07	0.001238	2	0.51615	6
CCDC115	8.35E-07	2.28E-07	0.001238	3	0.74051	5
KRTAP10-4	1.05E-06	2.28E-07	0.001238	4	0.65116	4
FAM26F	1.72E-06	6.85E-07	0.002475	5	0.44181	6

WDR7	1.90E-06	6.85E-07	0.002475	6	0.4243	6
CD1E	3.22E-06	1.14E-06	0.003536	7	0.56297	6
MRPL39	5.60E-06	2.05E-06	0.005569	8	0.36095	3
STT3A	8.70E-06	6.62E-06	0.015952	9	0.56206	4
PRMT5	1.36E-05	1.07E-05	0.023267	10	0.1196	4

Table S5 **CRISPR knock-out screen positive selection Mcf2**

Gene	**pos|score**	**pos|pvalue**	**pos|fdr**	**pos|rank**	**pos|lfc**	**pos|good grnas**
hsa-mir-1277	1.19E-09	2.28E-07	0.00495	1	2.0186	4
COMMD3	1.70E-06	2.97E-06	0.032178	2	0.68629	6
C10orf68	1.06E-05	1.39E-05	0.088119	3	0.72133	5
RPE	1.31E-05	2.03E-05	0.088119	4	0.42111	6
EPHA6	1.54E-05	2.44E-05	0.088284	5	0.89303	5
hsa-mir-548an	1.59E-05	1.80E-05	0.088119	6	1.8403	3
SPAG6	2.53E-05	4.18E-05	0.128422	7	0.20404	3
hsa-mir-3202-1	3.04E-05	4.81E-05	0.128422	8	0.93345	4
TUBA1B	3.85E-05	6.05E-05	0.128422	9	0.70943	5
SLC46A1	3.98E-05	6.46E-05	0.128422	10	0.45677	5

Table S6 **List of plasmids for protein expression in *E. coli* Mcf1**

Name	**Description**
pETDuet	*E. coli* vector expressing two genes
1330 pET28 Iota-A	Clostridium perfringens Iota-A with N-terminal His-tag in pET28 vector
pB656 pET28 MBP-TcHVR ΔN100 (TcART)	TcHVR with the N-terminal deletion in fusion with MBP tag in pET28 vector

pB475 pET28 His-Arf3 Q71L	Human ADP-ribosylating factor 3 (Arf3; Uniprot P61204) with N-terminal deletion of 17 amino acids and mutation Q71L codon-optimized for expression *E. coli*
pB394 pET28 His-Mcf1	Mcf1 gene (Uniprot Q8KT65) was PCR amplified from *Photorhabdus luminescens* strain akhurstii
pB397 pET28 His-Myc-Mcf1-FLAG	N-terminal Myc and C-terminal FLAG tagged Mcf1
pB446 pET28 His-Myc-Mcf1_C1397A-FLAG	N-terminal Myc and C-terminal FLAG tagged Mcf1 with mutation C1397A
pB852 pETDuet Arf3 Q71L	Arf3 Q71L from pB475 in pETDuet vector
pB858 pETDuet Arf3 Q71L His-Myc-Mcf1-FLAG	Mcf1 and Arf3 Q71L in pETDuet vector
pB891 pET28 His-Mcf1-Myc pB914 His-Mcf1Δ15C	Mcf1 with C-terminal Myc tag
pB915 His-Mcf1 C1397A Δ15C	Mcf1 with C-terminal deletion of 15 amino acids
pB916 pETDuet His-Myc-Mcf1-FLAG	Mcf1 from pB397 in pETDuet vector
pB917 pETDuet Arf3 Q71L His-Myc-Mcf1_C1397A-FLAG	Mcf1 C1397A with N-terminal myc and C-terminal FLAG-tag together with Arf3 Q71L in pETDuet vector
pB973 pET28 His-Mcf1 Y1286S Δ15C	Mcf1 with C-terminal deletion of 15 amino acids and Y1286S mutation
pB987 pET28 His-Mcf1 R957A Δ15C	Mcf1 with C-terminal deletion of 15 amino acids and R957A mutation
pB989 pET28 His-Mcf1 Y1286S C1397A Δ15C	Mcf1 with C-terminal deletion of 15 amino acids and Y1286S C1397A mutations
pB992 pET28 His-Mcf1 L1359G Δ15C	Mcf1 with C-terminal deletion of 15 amino acids and L1359G mutation
pB993 pET28 His-Mcf1 L1359G R1360G Δ15C	Mcf1 with C-terminal deletion of 15 amino acids and L1359G R1360G mutations
pB996 pET28 His-Mcf1 L911A K912A Δ15C	Mcf1 with C-terminal deletion of 15 amino acids and L911A K912A mutations
pB1000 pETDuet His-Myc-Mcf1-FLAG L911A K912A	Mcf1 from pB397 with L911A K912A mutations in pETDuet vector
pB1004 pET28 His-Myc-Mcf1-FLAG L911A K912A I1271A Q1272A G1273A	Mcf1 with L911A K912A I1271A Q1272A G1273A mutations

pB1006 pETDuet Arf3 Q71L His-Myc-Mcf1-FLAG L911A K912A	His-Myc-Mcf1-FLAG with mutations L911A K912A together with Arf3 Q71L in pETDuet vector
pB1008 pETDuet Arf3 Q71L His-Myc-Mcf1-FLAG L911A K912A I1271A Q1272A G1273A	His-Myc-Mcf1-FLAG L911A K912A I1271A Q1272A G1273A together with Arf3 Q71L in pETDuet vector
pB1012 pET28 His-Myc-Mcf1-FLAG I1271A Q1272A G1273A	Mcf1 with N-terminal myc and C-terminal FLAG-tag and I1271A Q1272A G1273A mutations
pB1022 pET28 His-Mcf1 ΔK907 Δ15C	Mcf1 with ΔK907 deletion and C-terminal deletion of 15 amino acids
pB1023 pET28 His-Mcf1 ΔK907 ΔD908 Δ15C	Mcf1 with ΔK907 ΔD908 deletion and C-terminal deletion of 15 amino acids
pH93 pET28a His-Myc-Mcf1-FLAG LKAGL_VGNGL_toAla	Mcf1 with N-terminal myc and C-terminal FLAG-tag and LKAGL and VGNGL mutated to alanine
pH94 pET28a His-Myc-Mcf1-FLAG LKAGL_LKAAA	Mcf1 with N-terminal myc and C-terminal FLAG-tag and LKAGL to LKAAA mutation
pH95 pET28a His-Myc-Mcf1-FLAG LKAGL_LAAAL	Mcf1 with N-terminal myc and C-terminal FLAG-tag and LKAGL to LAAAL mutation
pH96 pET28a His-Myc-Mcf1-FLAG LKAGL_LKAAL	Mcf1 with N-terminal myc and C-terminal FLAG-tag and LKAGL to LKAAL mutation
pH97 pET28a His-Myc-Mcf1-FLAG LKAGL_LAAGL	Mcf1 with N-terminal myc and C-terminal FLAG-tag and LKAGL to LAAGL mutation

Table S7 **List of plasmids for protein expression in *E. coli* Mcf2**

Name	**Description**
2473 YEpGal555	E. coli/S. cerevisiae shuttle vector [ADE2] with Gal1 promoter
pH148 pET28a Myc-Mcf2	Mcf2 gene (NCBI Reference Sequence: WP_214040153.1) was PCR amplified from *Photorhabdus luminescens* strain akhurstii
pH149 pET28a His-Myc-Mcf2 C810A	Mcf2 with N-terminal myc tag and C810A mutation
pH150 pET28a His-Mcf2-Myc	Mcf2 with with C-terminal myc tag
pH151 pET28a His-Myc-Mcf2 Δ14C	Mcf2 with N-terminal myc tag and deletion of the last 14 amino acids

pH152 pET28a His-Myc-Mcf2 Δ14C C810A	Mcf2 with N-terminal myc tag, a deletion of the last 14 amino acids and C810A mutation
pH161 pET28a His-Myc-Mcf2 Effector 1-321	Mcf2 effector fragment (amino acid 1-321) with N-terminal myc tag
pH167 pET28a His-Myc-Mcf2 L313K314_AA	Mcf2 with N-terminal myc tag and L313K314 to alanine mutation
pH173 pET28a His-N-AcGFP-Myc-Mcf2	Fusion protein of N-AcGFP and Mcf2

Table S8 **List of plasmids for protein expression in *S. cerevisiae* Mcf1**

Name	**Description**
2473 YEpGal555	E. coli/S. cerevisiae shuttle vector [ADE2] with Gal1 promoter
pB380 YEpGal555 Mcf1 fragment 6	Mcf1 fragment of amino acids 1133-1619
pB408 YEpGal555 Mcf1 fragment 6 D1505A	Mcf1 fragment of amino acids 1133-1619 with a D1505A mutation
pB409 YEpGal555 Mcf1 fragment 6 C1397A	Mcf1 fragment of amino acids 1133-1619 with a C1397A mutation
pB419 YEpGal555 Mcf1 fragment 6 H1486A	Mcf1 fragment of amino acids 1133-1619 with a H1486A mutation
pB631 YEpGal555 TcART	TccC3 with 100 amino acid deletion at N-terminus
pB798 YEpGal555 Mcf1 1-881	Mcf1 fragment of amino acids 1-881
pB799 YEpGal555 Mcf1 1-431	Mcf1 fragment of amino acids 1-431
pB800 YEpGal555 Mcf1 432-881	Mcf1 fragment of amino acids 432-881
pB801 YEpGal555 Mcf1 216-431	Mcf1 fragment of amino acids 216-431
pB802 YEpGal555 Mcf1 1-215	Mcf1 fragment of amino acids 1-215
pB827 YEpGal555 Mcf1 1-980	Mcf1 fragment of amino acids 1-980
pB828 YEpGal555 Mcf1 1-1080	Mcf1 fragment of amino acids 1-1080
pB829 YEpGal555 Mcf1 1-1180	Mcf1 fragment of amino acids 1-1180
pB830 YEpGal555 Mcf1 1-1280	Mcf1 fragment of amino acids 1-1280

pB831 YEpGal555 Mcf1 1-1380	Mcf1 fragment of amino acids 1-1380
pB832 YEpGal555 Mcf1 1-1480	Mcf1 fragment of amino acids 1-1480
pB833 YEpGal555 Mcf1 1-1580	Mcf1 fragment of amino acids 1-1480
pB834 YEpGal555 Mcf1 1-1630	Mcf1 fragment of amino acids 1-1480
pB990 YEpGal555 Burkholderia lethal factor 1	*S. cerevisiae* codon optimized gene of *Burkholderia pseudomallei* lethal factor 1 (Blf1, Uniprot Q63UP7)
pH009 YEpGal555 Mcf1 32-881	Mcf1 fragment of amino acids 32-881
pH010 YEpGal555 Mcf1 66-881	Mcf1 fragment of amino acids 66-881
pH011 YEpGal555 Mcf1 92-881	Mcf1 fragment of amino acids 92-881
pH012 YEpGal555 Mcf1 140-881	Mcf1 fragment of amino acids 140-881
pH013 YEpGal555 Mcf1 176-881	Mcf1 fragment of amino acids 176-881
pH014 YEpGal555 Mcf1 1-834	Mcf1 fragment of amino acids 1-834
pH016 YEpGal555 Mcf1 1-764	Mcf1 fragment of amino acids 1-764
pH017 YEpGal555 Mcf1 1-735	Mcf1 fragment of amino acids 1-735
pH018 YEpGal555 Mcf1 1-699	Mcf1 fragment of amino acids 1-699
pH023 YEpGal555 Mcf1 913-1125	Mcf1 fragment of amino acids 913-1125
pH024 YEpGal555 Mcf1 913-1125 G1031C	Mcf1 fragment of amino acids 913-1125 with a G1031C mutation
pH26-pH82 YEpGal555 Mcf1 frag	Mcf1 fragments resulting from the random mutagenesis
pH85 YEpGal555 Mcf1 1-912 N571A	Mcf1 fragment of amino acids 1-912 with point mutation of quadruple hit N571A (numbering including tag)
pH86 YEpGal555 Mcf1 1-912 E46A	Mcf1 fragment of amino acids 1-912 with point mutation of triple hit E46A (numbering including tag)
pH87 YEpGal555 Mcf1 1-912 T133A	Mcf1 fragment of amino acids 1-912 with point mutation of triple hit T133A (numbering including tag)
pH88 YEpGal555 Mcf1 1-912 E446A	Mcf1 fragment of amino acids 1-912 with point mutation of triple hit E446A (numbering including tag)
pH89 YEpGal555 Mcf1 1-912 S568A	Mcf1 fragment of amino acids 1-912 with point mutation of triple hit S568A (numbering including tag)
pH90 YEpGal555 Mcf1 1-912 V583A	Mcf1 fragment of amino acids 1-912 with point mutation of triple hit V583A (numbering including tag)

pH91 YEpGal555 Mcf1 1-912 L738A	Mcf1 fragment of amino acids 1-912 with point mutation of triple hit L738A (numbering including tag)
pH92 YEpGal555 Mcf1 1-912 Y814A	Mcf1 fragment of amino acids 1-912 with point mutation of triple hit Y814A (numbering including tag)
pH98 YEpGal555 Mcf1 1-912 pH47mut	pH47 with reduced number of mutations by combining with unmutated Mcf1 at CsiI restriction site
pH99 YEpGal555 Mcf1 1-912 pH47mut	pH47 with reduced number of mutations by combining with unmutated Mcf1 at CsiI restriction site
pH100 YEpGal555 Mcf1 1-912 pH47mut	pH47 with reduced number of mutations by combining with unmutated Mcf1 at CsiI restriction site
pH101 YEpGal555 Mcf1 1-912 single pH47mut_1	pH47 with reduced number of mutations to a single one
pH102 YEpGal555 Mcf1 1-912 single pH47mut_2	pH47 with reduced number of mutations to a single one
pH103 YEpGal555 Mcf1 1-912 pH69mut	pH69 with reduced number of mutations by combining with unmutated Mcf1 at CsiI restriction site
pH104 YEpGal555 Mcf1 1-912 pH69mut	pH69 with reduced number of mutations by combining with unmutated Mcf1 at CsiI restriction site
pH105 YEpGal555 Mcf1 1-912 pH69mut	pH69 with reduced number of mutations by combining with unmutated Mcf1 at CsiI restriction site
pH106 YEpGal555 Mcf1 1-912 single pH69mut_1	pH69 with reduced number of mutations to a single one
pH107 YEpGal555 Mcf1 1-912 single pH69mut_2	pH69 with reduced number of mutations to a single one
pH108 YEpGal555 Mcf1 1-912 single pH69mut_3	pH69 with reduced number of mutations to a single one
pH141 pK87 Mcf1 fragment 6 IQGG_AAAA	Mcf1 fragment of amino acids 1133-1619 with mutation IQGG to AAAA to identify second cleavage site
pH142 pK87 Mcf1 fragment 6 GGKQ_AAAA	Mcf1 fragment of amino acids 1133-1619 with mutation GGKQ to AAAA to identify second cleavage site
pH143 pK87 Mcf1 fragment 6 QRTD_AAAA	Mcf1 fragment of amino acids 1133-1619 with mutation QRTD to AAAA to identify second cleavage site
pH144 pK87 Mcf1 fragment 6 DSVG_AAAA	Mcf1 fragment of amino acids 1133-1619 with mutation DSVG to AAAA to identify second cleavage site

Table S9 **List of plasmids for protein expression in *S. cerevisiae* Mcf2**

Name	**Description**
2473 YEpGal555	*E. coli*/*S. cerevisiae* shuttle vector [ADE2] with Gal1 promoter
pH153 YEpGal555 myc-Mcf2	Full size Mcf2

pH154 YEpGal555 myc-Mcf2 C810A	Full size Mcf2 with C810A mutation
pH155 YEpGal555 myc-Mcf2 Effector 1-313	Effector fragment of Mcf2 (amino acids 1-313)
pH156 YEpGal555 myc-Mcf2 C810A 313-2388	Mcf2 without effector (amino acids 313-2388) as C810A mutant
pH157 YEpGal555 myc-Mcf2 313-end	Mcf2 without effector (amino acids 313-2388)
pH162 YEpGal555 myc-Mcf2 CPD	CPD fragment of Mcf2 (amino acids 303-1033)
pH163 YEpGal555 myc-Mcf2 CPD C810A	CPD fragment of Mcf2 (amino acids 303-1033) as C810A mutant
pH164 YEpGal555 myc-Mcf2 head	Full head fragment of Mcf2 (amino acids 1-1033)
pH165 YEpGal555 myc-Mcf2 head C810A	Full head fragment of Mcf2 (amino acids 1-1033) as C810A mutant
pH166 YEpGal555 myc-Mcf2 ABD	Activator binding domain fragment of Mcf2 (amino acids 303-667)

11. Table of figures

11.1 Main figures

11.2 Supplementary figures

12. Acknowledgements

I would like to thank Prof. Dr. Stefan Raunser for the opportunity and trust in me to pursue this exiting project that already began during my master. His scientific guidance and personal support during my PhD journey were essential in completing this work. Furthermore, I would like to express my appreciation to the members of my thesis advisory committee Prof. Dr. Roger Goody and Prof. Dr Rasmus Linser for the insightful discussions and encouraging feedback on my work.

I owe unlimited gratitude to my direct supervisor, Dr. Alexander Belyy, who was a guiding light throughout the early years of the PhD and whom I regard as one of my scientific mentors. From cloning to protein purification, cryo-EM imaging to data processing, planning experiments to executing them – he taught me everything I know and greatly influenced who I am today.

All members of the department that I met during my time contributed to a fun and stimulating atmosphere. Everyone was always ready to help and greatly supported my work. It is impossible to not become friends when working close together with the most amazing people one could wish for. Thank you very much for the great time!

A special thanks goes out to Dr. Philine Hagel, my project predecessor, and the entire toxin group for helping me to develop the project and find my home in the world of bacterial toxins. Therefore, I thank: Raphael Bader, Dr. Oleg Sitsel, Dr. Peter Njenga Ng`Ang`A, Dr. Patrick Günther, Dr. Heng Zhou, and Dr. Jai Prakash Singh – the sunshine of our office.

I am grateful to Dr. Amrita Rai for providing the Arf3 plasmid and initial protein samples.

I would also like to emphasize and acknowledge the excellent technical assistance, especially from Sandra Bergbrede as well as also from Nathalie Bleimling, Marion Hülseweh, Karin Vogel-Bachmayra, Diana Ludwig, and Dr. Wei Zhou, whose efforts keep the lab running like a smooth engine.

The cryo-EM facility would have long collapsed if not for the constant and meticulous care of Dr. Oliver Hofnagel and Dr. Daniel Prumbaum, who introduced me to the instruments without sparing any details on their history and operation. I thank them for their insights and constant availability in case of questions and troubles.

My work also greatly benefited from the ‘mass spectrometry’ and ‘crystallography and biophysics’ service at the MPI, as well as the ICP-MS measurements performed by Ingo Feldmann at the ISAS.
A PhD contains a lot of meetings, retreats and bureaucracy which was mostly organized in the background by Meike Schulte and Monika von Manger and kept my days free for research, which I am grateful for.
It has been an honor to be part of the thriving IMPRS LM PhD program organized by Christa and Lucia who manage to create and foster a supportive and outgoing community, while still taking care of each individual student.
For a fresh view on scientific problems and afterwork conversations, I am grateful to the “girls office” – Vicki, Elisa, Sree, Alicia, and their ever-present companion Baptiste – though one had to be careful to not end up with a drink in your hand.
I thank Alex (Mingjun Xu) for our jamming sessions in the basement and shared concert experiences. The travels and music we enjoyed together helped me to relax and recharge. This was also done by all members of our short-lived DnD group – Rapha, Adri, Gavin and Mika – where we escaped reality for a few evenings and on trips to the medieval market.
Furthermore, I would like to express my gratitude to our bike group for tours either around Dortmund or all the way to Renesse. Special thanks to Wout for being a source of constant positivity and extensive processing knowledge as a ‘true scientist’, Thorsten and Markus for their IT support and a different perspective on scientific problems, and to Raphael who journeyed through the PhD together with me and shared in my hardships.
Lastly, I will be forever thankful to my family and closest friends – Kai, Caro, Thomas, and Jonas – who accompanied me already during my studies at the university and continue to be a bedrock especially in stressful and troubled times.
And last but not least, thanks to Mcf for making not only caterpillars but also my brain floppy at times.

13. Eidesstattliche Versicherung

Eidesstattliche Versicherung (Affidavit)

Name, Vorname
(Surname, first name)

Matrikel-Nr.
(Enrolment number)

Belehrung:

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

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

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

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

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

The submission of a false affidavit is punishable.

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

I have taken note of the above official notification.

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.
Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

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

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

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

Ort, Datum
(Place, date)

Unterschrift
(Signature)