

Rockefeller University

Digital Commons @ RU

---

Student Theses and Dissertations

---

2022

## Structure and Mechanism of PCAT1, a Polypeptide Processing and Secretion Transporter

Virapat Kieuvongngam

Follow this and additional works at: [https://digitalcommons.rockefeller.edu/student\\_theses\\_and\\_dissertations](https://digitalcommons.rockefeller.edu/student_theses_and_dissertations)



Part of the Life Sciences Commons

---



**Structure and mechanism of PCAT1, a polypeptide processing and secretion transporter**

A Thesis Presented to the Faculty of  
The Rockefeller University  
in Partial Fulfillment of the Requirements for  
the degree of Doctor of Philosophy

by  
Virapat Kieuvongngam  
June 2022



## Structure and mechanism of PCAT1, a polypeptide processing and secretion transporter

Virapat Kieuvongngam, Ph.D.  
The Rockefeller University 2022

ATP-binding cassette (ABC) transporters utilize the energy from ATP to transport substrates across biological membranes. Various ABC transporters perform diverse biological functions across all forms of life ranging from importing essential nutrients to exporting toxic drugs (Ford and Beis 2019; ter Beek, Guskov, and Slotboom 2014).

Bacterial cells utilize a class of ABC transporters for exporting proteins or peptides. Unlike the Sec translocon machinery, the ABC peptide exporters are dedicated to specific peptides (Fath and Kolter 1993). These peptides function as quorum sensing peptides, biofilms, or antimicrobial peptides. Among these ABC peptide exporters are Peptidase Containing ABC Transporters (PCATs) that perform dual functions of peptide maturation through proteolytic cleavage and peptide export (Gebhard 2012). As the name connotes, these transporters contain an accessory cysteine protease domain that interacts with the core ABC transporter. This structural feature is essential for the function and is unique among ABC transporters, making it a biologically interesting target for investigation.

Although PCATs, which were first described 20 years ago, are essential to prokaryotic life, structural and functional studies of these proteins have been lacking. Until recently, only soluble parts of the proteins have been crystallized. The first full-length structure of PCATs was described in 2015 (Lin, Huang, and Chen 2015). However, structures of PCATs in complex with their substrates are needed to understand how PCATs recognize and transport their peptide substrates.

To understand how PCATs work at the atomic level, I mainly took a structural approach. To this end, I decided to use cryo-electron microscopy (cryo-EM) and single-particle reconstruction techniques to obtain high-resolution structures of PCATs in various conformational states in complex with their substrates. These structures, together with the biochemical evidence, give us a clearer mechanistic picture of PCATs.

First, to understand how PCATs bind substrate, I determined the structure of a PCAT from *Clostridium thermocellum* (abbreviated as PCAT1) in the substrate-bound inward-facing conformation. I have identified structural features that enable the substrate to bind, translocate into the transmembrane cavity, and orient properly for cleavage.

Next, to understand how substrate binding coordinates with ATP binding, I have determined three structures of PCAT1 in the active-turnover condition, where PCAT1 is allowed to transition freely through the transport cycle. In addition, I have determined the structure of a clear outward-facing structure PCAT1 trapped in the  $Mg^{2+}$  condition that elucidates how the behavior of the core transporter affects the accessory peptidase domain. These structures together enable us to propose a mechanism of how the ATP binding and hydrolysis cycle is synchronized with substrate binding and processing, a unique feature crucial for strict coupling of cleavage and translocation.

In addition to the cryo-EM work, I collaborated with Dr. Paul Dominic Olinares in Professor Brian T. Chait's laboratory to study the stoichiometry of the PCAT1-substrate complex and intermediates of PCAT1 in the transport cycle. This work allows us to delineate the steps along the transport cycles that are short-lived and cannot be captured using structural study.

Dedicated to my mother, my father, my two brothers, and my husband

## ACKNOWLEDGEMENTS

First, I would like to thank my advisor Dr. Jue Chen for her guidance throughout my graduate study. She has devoted a tremendous amount of time training me to be a scholar. I look up to her as a model of a well-rounded successful working professional.

I would also like to thank my collaborators, Dr. Paul Dominic Olinares and Dr. Brian Chait for their expert contribution to our works. I am truly indebted to them for the time they dedicated to teaching me. I have learned enormously from both.

I would like to thank my thesis committee members, Dr. Jue Chen, Dr. Brian Chait, Dr. Shixin Liu, and Dr. Gregory Alushin as they have pushed me to improve my research as well as giving me career advice over the years. In addition, I would like to thank Dr. Nikolaus Grigorieff for being my external examiner and sharing his expertise on cryo-electron microscopy.

I would also like to thank all the administrative staff at the graduate program including Marta Delgado, Cristian Rosario, Kristen Cullen, Emily Harms, Andrea Morris, and the dean of Rockefeller graduate program, Dr. Sidney Strickland for ensuring that I had access to educational resources to help me grow as a scientist.

I would like to thank the staff at the Evelyn Gruss Lipper Cryo-Electron Microscopy center including Mark Ebahim, Johanna Sotiris, and Honkit Ng for their technical support which made my experiments possible.

I feel genuinely grateful to all the former and current lab members of the Chen lab with whom I have developed deep and meaningful friendships. I would like to thank Donna Tallent for proofreading this manuscript. They make a wonderful team and they have made the Chen lab a warm welcoming place for me.

I would like to express my deepest gratitude to my mother Somjai, my father Mitchai, and my younger brothers who have supported my education and instilled in me a great love of learning. Finally, I would like to express my greatest appreciation to my husband Dr. Braden Purcell who encouraged me to push through challenges and reach my goals while never forgetting to enjoy the journey.

## TABLE OF CONTENTS

<b>Lists of Figures</b>	<b>vi</b>
<b>Lists of Tables</b>	<b>vii</b>
<b>Chapter 1: Introduction</b>	
Overview of ABC transporters	1
Structural diversity of ABC transporters	5
General knowledge of Peptidase Containing ABC Transporters (PCATs)	7
<b>Chapter 2: Cryo-EM structure of the substrate-bound PCAT1</b>	
Summary	12
Stoichiometry determination of PCAT1-substrate complex	12
Structural determination of PCAT1-substrate complex	16
Investigating the proteolysis mechanism of PCAT1	21
Mechanistic details of substrate recognition	23
Asymmetric positioning of the cargos	25
Conformational changes upon substrate binding	27
Discussion	29
<b>Chapter 3: Structure of PCAT1 under the active turnover and Mg<sup>2+</sup> free condition reveals the rate-limiting steps and the role of ATP in PCAT1 transport cycle</b>	
Summary	36
Structure determination of <i>wt</i> PCAT1 in the ATP-bound, pre-hydrolysis conformation	36
Structure determination of <i>wt</i> PCAT1 under active turnover condition	42
Conformational dynamic in the active turnover condition	49
Discussion	51
<b>Chapter 4: Additional native mass spectrometric study of PCAT1</b>	
Introduction	59
Summary	60
General workflow of the time-resolved nMS-based reaction monitoring	60
Time course of substrate cleavage revealed by nMS reaction monitoring	64
Transient ATP hydrolysis intermediate captured by time-resolved native mass spectrometry	68
Determining the coupling of proteolysis cleavage function and ATP hydrolysis of PCAT1	70
Discussion	74
<b>Chapter 5: Conclusion</b>	
Concluding remarks	80
Future directions	81
<b>References</b>	<b>83</b>

## List of Figures

<b>Figure 1.1</b>	The general architecture of an ABC transporter and the mechanism of ATP binding and hydrolysis	3
<b>Figure 1.2</b>	The crystal structures of <i>wt</i> PCAT1 in <i>apo</i> inward facing conformation and E648Q PCAT1 in ATP $\gamma$ S-bound occluded conformation	9
<b>Figure 1.3</b>	Sequence alignments	10
<b>Figure 2.1</b>	Gel filtration chromatography profile of PCAT1	13
<b>Figure 2.2</b>	Gel filtration chromatography profile of CtA	13
<b>Figure 2.3</b>	Native Mass spectrometry analysis of <i>wt</i> PCAT1 and C21A PCAT1-CtA complex	15
<b>Figure 2.4</b>	Summary of Cryo-EM image processing workflow	18
<b>Figure 2.5</b>	Local resolution estimation of the PCAT1-CtA complex density map	19
<b>Figure 2.6</b>	The cryo-EM structure of PCAT1-CtA complex	20
<b>Figure 2.7</b>	Substrate-peptidase domain interaction	22
<b>Figure 2.8</b>	Common PCAT residues for substrate recognition	24
<b>Figure 2.9</b>	Positions of the two cargos	26
<b>Figure 2.10</b>	Conformational changes upon CtA binding	28
<b>Figure 2.11</b>	The alternating access mechanism of PCAT1	31
<b>Figure 3.1</b>	Summary of cryo-EM image processing of PCAT1 in Mg <sup>2+</sup> free condition	39
<b>Figure 3.2</b>	The cryo-EM structure of PCAT1 in the OF conformation	40
<b>Figure 3.3</b>	Comparison of conformational differences between PCAT1 in the OF and occluded conformations	41
<b>Figure 3.4</b>	Summary of cryo-EM image processing of PCAT1 under active turnover condition	45
<b>Figure 3.5</b>	Quality of the cryo-EM density map of <i>wt</i> PCAT1 in the IF conformations	46
<b>Figure 3.6</b>	The cryo-EM structure of PCAT1 IF conformations	47
<b>Figure 3.7</b>	Substrate cleavage under active turnover condition	48
<b>Figure 3.8</b>	Conformational variability in the active turnover dataset	50
<b>Figure 3.9</b>	Peptide transport cycle model of PCAT1	52
<b>Figure 3.10</b>	Kinetic model diagram summarizing the cryo-EM experiments	55
<b>Figure 4.1</b>	Workflow for nMS-based continuous reaction monitoring	63
<b>Figure 4.2</b>	nMS-based monitoring of the PCAT1 peptidase activity	67
<b>Figure 4.3</b>	nMS-based monitoring of the PCAT1 ATP hydrolysis activity	69
<b>Figure 4.4</b>	nMS-based monitoring of PCAT1 peptidase and ATP hydrolysis activity in the absence of Mg <sup>2+</sup>	72
<b>Figure 4.5</b>	nMS-based monitoring of the peptidase and ATP hydrolysis activity of PCAT1	73
<b>Figure 4.6</b>	Reaction monitoring of PCAT1 peptidase activity in nMS-compatible detergent C <sub>8</sub> E <sub>4</sub> and detergent used for structural study UDM	76

**List of Tables**

<b>Table 2.1</b>	Mass measurement from native MS analysis of PCAT1	14
<b>Table 2.2</b>	Summary of cryo-EM data and structure refinement statistics for PCAT1-CtA complex	17
<b>Table 3.1</b>	Summary of Cryo-EM data and structure refinement statistics for PCAT1 under Mg <sup>2+</sup> free condition	38
<b>Table 3.2</b>	Summary of Cryo-EM data and structure refinement statistics for PCAT1 under active turnover condition	44
<b>Table 4.1</b>	Mass measurements from the nMS-based reaction monitoring of PCAT1 activities	66

## **CHAPTER 1: Introduction**

### **1.1 Overview of ABC transporters**

#### **1.1.1 Function of ABC transporters**

Cells are the basic functional units of life. Cells consist of cytoplasm enclosed within biological membranes. Composed of phospholipid bilayers, the cellular membranes establish boundaries of cells and the environment. The membranes maintain cellular homeostasis needed for metabolism.

Embedded within the lipid bilayers are membrane transport proteins that function as pumps to selectively pump in essential nutrients and pump out toxic chemicals. These transporters work to maintain the concentration of chemical substrates against their electrochemical gradients by harnessing the energy either directly from ATP or indirectly from electrochemical gradients of other molecules.

Among these transporters are ATP-binding cassette (ABC) transporters. This superfamily of transporters share unique structural and functional characteristics in that functionally they harness the energy from ATP to unidirectionally transport the substrates across membranes. Structurally, this task is accomplished through a protein domain called the nucleotide-binding domain (NBD) which bind ATP and couple the binding to transport function (Srikant and Gaudet 2019; Binet et al. 1997; Locher 2016).

ABC transporters are found in all domains of life. Both importers and exporters are found in the prokaryotes (Davidson and Chen 2004). Prokaryotic cells use ABC exporters to secrete molecules performing diverse biological functions, including cell-to-cell communication, quorum sensing, biofilm formation, and antimicrobial activity (Fath and Kolter 1993). The secreted molecules are chemically diverse ranging from small molecule drugs, to short peptides, to large proteins.

#### **1.1.2 General structural features of an ABC transporter**

A typical ABC transporter is a dimer of two half transporters forming a conserved core component of the transporter. Each half transporter contains two domains: the transmembrane domain (TMD) and the nucleotide binding domain (NBD) (Ford and Beis 2019) (Fig. 1.1 A). The TMD consists mainly of  $\alpha$  helices spanning the lipid bilayers, forming a cavity in the membrane. The NBDs are in the cytosolic side attached to the TMD. With the ATP binding site facing each other, the NBDs can form a dimer when ATP is bound. The assembly of ABC transporters can either be a homodimer; that is to say, the two half-transporters are encoded in one gene, or a heterodimer, when the two half transporters are encoded on different genes or as different domains in one gene. These protein domains can be encoded as separate genes that assemble to form functional transporters or fused as domains connected through flexible linkers.

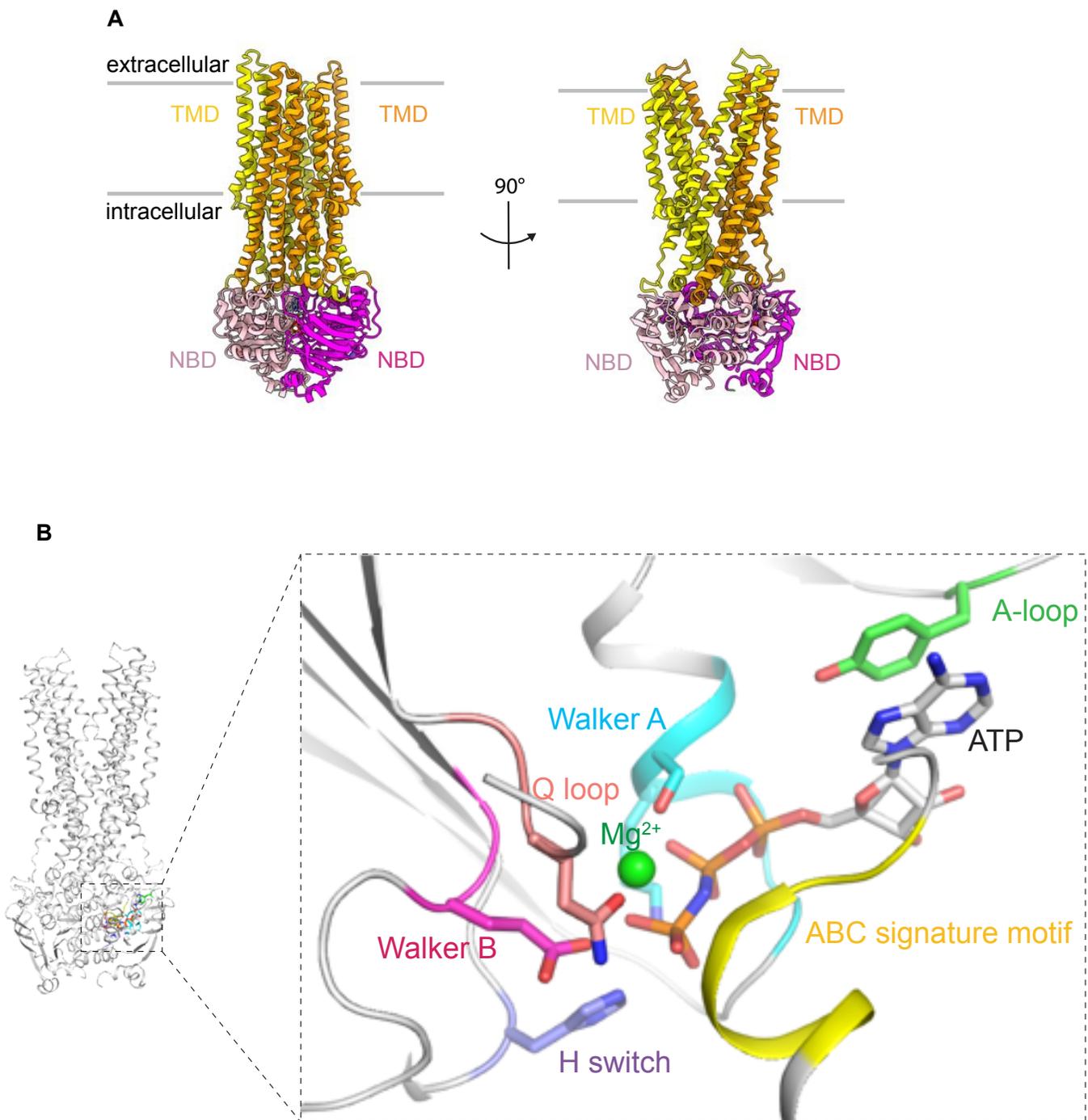
### 1.1.3 General features of ATP-powered pump in ABC transporters

All ABC transporters share a highly conserved nucleotide binding domain consisting of a core catalytic subdomain, ABC-specific three-stranded antiparallel  $\beta$ -sheet (ABC $\beta$ ), and an  $\alpha$ -helical subdomain (ABC $\alpha$ ). The core catalytic subdomain is RecA-type ATPase. ABC signature motif (also known as LSGGQ motif) is located at the end of an ABC $\alpha$  subdomain (Locher 2016; Rees, Johnson, and Lewinson 2009).

Nucleotide binding is stabilized by a Walker A motif on the RecA subdomain (Fig. 1.1 B). The Walker A motif provides amide NH groups binding to the  $\beta$  and  $\gamma$  phosphate of the nucleotides. The serine/threonine residue of the Walker A motif is involved in coordinating the  $Mg^{2+}$  ligand that further coordinates  $\beta$  and  $\gamma$  phosphate. The A loop provides an aromatic residue forming a  $\pi$ - $\pi$  stacking interaction with the base ring of the nucleotide. The Walker B motif helps coordinate the  $Mg^{2+}$  ion. The catalytic glutamate responsible for polarizing the attacking water molecule is located near the  $\gamma$  phosphate. The conserved histidine of the H switch motif is responsible for stabilizing the catalytic transition state. (C. Thomas and Tampé 2020; Locher 2016).

NBD dimerizes in a head-to-tail arrangement such that each of the nucleotide binding sites are formed by a Walker A/B motif of one NBD and the signature motif of the opposite NBD. Two nucleotides are sandwiched at the dimer interface, stabilizing the dimer (J. Chen et al. 2003). The signature motif makes a contact along the nucleotide molecule and orients the nucleotide for hydrolysis. After hydrolysis, the inorganic phosphate is released, allowing the NBD to disengage (Oldham and Chen 2011).

The contact along the NBD and the coupling helix of the TMD translates motion of NBD dimerization to conformational changes on the TMDs. This links the cycle of NBD dimerization and ATP hydrolysis to the cycle of TMD motion from inward-facing to outward-facing conformation. This cycle is the basis of a transport cycle. The thermodynamic role of ATP binding and hydrolysis is to provide an irreversible process that drives the transport cycle in one direction; thereby the substrate transport process can build up the substrate concentration against its concentration gradient.



**Fig. 1.1 The overall architecture and mechanism of a canonical ABC transporter**

(A) Cartoon ribbon diagram representing overall architecture of a canonical ABC transporter, Sav1866 (PDB: 2hyd). The domain is color-coded and labeled. Magenta, NBD; orange, TMD; green, PEP.

(B) A zoom-in view of the nucleotide binding site of Sav1866 showing the mechanism of ATP binding and hydrolysis. Motifs participating in ATP binding and hydrolysis are color-coded and labeled. Green, A-loop; yellow, ABC signature motif; light pink, Q loop; blue, Walker A; magenta, Walker B; purple, H switch.

### 1.1.4 General transport mechanism of ABC transporters

ABC transporters mediate transport processes of diverse molecules, ranging from nutrients to xenobiotics, to peptides. However, based on the structural and biophysical studies, the mechanism of transport of a canonical ABC transporter is consistent with the alternating access model (Khare et al. 2009; Jardetzky 1966). The alternating access model proposes that transporters conformationally alternate between an inward-facing conformation and an outward-facing conformation. In the inward-facing conformation, the transmembrane cavity is open to the cytosolic side. On the contrary, the transmembrane cavity of the outward-facing conformation is open to the extracellular side. The model proposes that the transport substrate is captured at the binding site in one conformation. The conformational change allows the bound substrate to be exposed and be released from the binding site at the opposite side of the membrane. In the case of an exporter, the substrate binds to the binding site opening to the cytosol in the inward-facing conformation. The substrate is enclosed within the transmembrane cavity. After the conformational transition to the outward-facing conformation, the substrate is exposed to the extracellular side and is released.

The alternating access model of transporters makes three predictions on the structures of the transporters. These predictions are:

1. transporters contain a transmembrane cavity in the lipid membrane that is large enough to host the substrate
2. transporters can assume two different conformations such that the transmembrane cavity is open to one side of the membrane in one conformation, and to other opposite of the membrane in the other conformation.
3. the transmembrane cavity contains a binding site for the transported substrate. The affinity for the transport substrates is high in one conformation and low in the other conformation.

There are many exceptions where the alternating access model may not account for how a particular ABC transporter works (Locher 2016). For example, transport of a protein substrate too large to be accommodated inside the transmembrane cavity is not consistent with the model. The notable example of this case is the hemolysin secretion mechanism by the type 1 secretion system in *E. coli* (S. Thomas, Holland, and Schmitt 2014; Binet et al. 1997; Michiels et al. 2001; Lecher et al. 2012).

In the next section, I shall discuss a few notable studies of ABC transporters to shed light on how the various ABC transporters employs the alternating access mechanism. A special emphasis is placed on structural diversity in ABC transporters correlated with variations in the alternating access mechanism and how they contribute to functional diversity among ABC transporters.

## 1.2 Structural diversity of ABC transporters

ABC transporters share a common domain architecture consisting of TMD and NBD. Nevertheless, sequence homology and structural studies have categorized ABC transporters based on direction of transport, domain folds, and symmetry of nucleotide binding sites (Rees, Johnson, and Lewinson 2009). In the following paragraphs I shall discuss representative ABC transporters for classes that are relevant to my work from which we have sufficient structural information.

### 1.2.1 A case study of homodimeric ABC exporters: ABCG2

ABCG2 is a representative transporter for a homodimeric exporter where both nucleotide binding sites are ATP-hydrolysis competent (consensus sites). ABCG2 functions as broad specificity efflux pumps of xenobiotic molecules and is implicated in multi-drug resistance in cancers. ABCG2 adopts an inward-facing *apo* state with the transmembrane cavity accessible from the inner membrane leaflets (Taylor et al. 2017). The substrate-bound structure is shown to exist in an inward-facing state with flat polycyclic compounds bound to a hydrophobic binding pocket in the cavity (Manolaridis et al. 2018). In the ATP-bound state, ABCG2 adopts an outward-facing state with an external cavity facing extracellular space, and the hydrophobic binding pocket is collapsed. The collapsed binding pocket is no longer able to bind substrates. Therefore, the conformational changes upon ATP binding is found to be consistent with the classic alternating access mechanism (Manolaridis et al. 2018; Q. Yu et al. 2021). It is hypothesized that upon inward-facing to outward-facing conformational changes, the hydrophobic substrate is squeezed out through a peristaltic movement of the cavity. ATP binding alone is important for substrate extrusion while ATP hydrolysis is necessary to reset the transporter back to an inward-facing state. It is curious to see if two ATP must be hydrolyzed or simply one ATP hydrolysis is sufficient to reset the transporter back to an inward-facing state.

Recent cryo-EM study of ABCG2 under an active turnover condition reveals two distinct inward-facing classes: turnover-1 and turnover-2. Transport substrates are found at the binding pocket in both conformations indicating that these are pre-translocation states. Moreover, ATP is found binding to the nucleotide binding sites in both conformations, suggesting that ATP binding per se, does not induce conformational changes from inward-facing to outward-facing states. Rather, ATP binding plays a role in stabilizing NBD dimerized states, or primes NBDs for full dimerization. The result from the active turnover experiment indicates that the full NBD dimerization is a short-lived state and is not observed in the active turnover condition. This implies that NBD dimerization is the rate limiting step in the transport cycle.

### 1.2.2 A case study of ABC transporters with asymmetric nucleotide binding sites: MRP1

MRP1 is a well-studied ABC exporter representing a class of asymmetric transporters with one consensus nucleotide binding site, and one ATP-hydrolysis deficient (degenerate) site. MRP1 functions as a xenobiotic molecule pump and is implicated in drug resistance in cancers. MRP1 adopts an inward-facing *apo* state with the transmembrane cavity accessible from the cytosol (Johnson and Chen 2017). The binding pocket is located in the transmembrane cavity and is composed of two parts: a positively charged region and a hydrophobic pocket. The binding

site configuration explains the ability of MRP1 to recognize large amphipathic substrates. Upon substrate binding, the binding pocket accommodates the substrate, bringing two halves of the TM bundles closer together via rigid body rotation. Consequently, NBD moves about 12 Å closer together. Therefore, the probability of NBD dimerization through thermal fluctuation is higher. Thus, the conformational change upon substrate binding primes NBD for dimerization and hydrolysis of ATP (Johnson and Chen 2017).

In the ATP-bound state, MRP1 adopts an outward-facing conformation with the substrate-binding site rearranged into a low-affinity state (Johnson and Chen 2018). This suggests that the substrate is released upon inward-facing to outward-facing conformational changes. These structures together provide conclusive evidence for the alternating access mechanism of transport.

Recently, a single molecule kinetic study of MRP1 establishes that the rate limiting step of the MRP1 transport cycle is the transition from outward-facing state to inward-facing state, suggesting that MRP1 spends the majority of its time in the OF conformation under physiological condition (Wang et al. 2020). Furthermore, the cryo-EM structure of MRP1 under active turnover condition is in the outward-facing post-hydrolysis conformation where ADP is found at the consensus site and ATP is found at the degenerate site. The overall conformation in the post-hydrolysis state is essentially identical to the ATP-bound conformation, suggesting the release of inorganic phosphate from the consensus site does not induce any conformational change.

The nucleotide binding site configuration under active turnover condition suggests a hypothesis that the asymmetry of the nucleotide binding site may play an important role in modulating the rate of OF-IF transition, shifting the rate limiting step to NBD dissociation. My analysis of PCAT1 under active turnover condition in chapter 3 will further expand on the distinction that asymmetric nucleotide binding sites play in setting the rate limiting step.

### **1.2.3 A case study of ABC transporters lacking substrate binding sites in the TMD: BtuCD**

Unlike the previous two case studies, BtuCD is an ABC importer responsible for cobalamin import in bacteria. BtuCD forms a homodimer consisting of two subunits of BtuC and two subunits of BtuD in the inner membrane. BtuCD employs a substrate-binding protein BtuF that binds cobalamin in the periplasmic space and feed it to the transmembrane cavity of the TMDs. BtuCD forms an outward-facing conformation in the absence of BtuF (Locher, Lee, and Rees 2002). This observation is consistent with the hypothesis that an ABC exporter in the *apo* state adopts an inward-facing conformation, while an ABC importer adopts an outward-facing conformation. The substrate-loaded BtuF docks onto the binding site on the TMDs in the ATP-bound conformation. BtuF releases the ligand into the transmembrane cavity of BtuCD (Korkhov, Mireku, and Locher 2012b). After ATP hydrolysis, release of inorganic phosphate causes the cytosolic gate to open allowing the substrate to leave the transmembrane cavity.

Unlike the cases of ABCG2 or MRP1, no substrate binding site is observed in the transmembrane cavity of BtuCD. The surface lining of BtuCD shows no significant affinity for

the substrate, likened to a “Teflon coating” pathway merely providing a conduit traversing the lipid bilayer (Korkhov, Mireku, and Locher 2012b). Here the third requirement of the alternating access mechanism is provided by the high cobalamin affinity BtuF which remains in the periplasmic space, and absence of any cytosolic cobalamin binding site on BtuCD. My analysis of PCAT1 in the substrate-bound conformation will further discuss how PCAT1 can transport substrates even though it lacks the substrate binding sites in the TMDs.

### **1.3 General knowledge of Peptidase Containing ABC Transporters (PCATs)**

#### **1.3.1 Discovery of PCATs**

A particular family of ABC transporters found in gram-positive bacteria functions to export a short peptide or a small protein out of the cells. The peptides function as quorum-sensing molecules, biofilms, or antimicrobial peptides. Notably, the ABC transporters are found to contain an accessory protein domain that are cysteine proteases (Havarstein, Diep, and Nes 1995). Because of this, the family of ABC transporters are named Peptidase Containing ABC Transporters, or PCATs (Lin, Huang, and Chen 2015). The accessory domains are fused to the core transporters or located in the same operon and assembled with the core to form a functional transporter.

Furthermore, the peptide substrates of PCATs contain two domains: the N-terminal leader peptide domain and the C-terminal cargo domain. The N-terminal leader peptide is important for binding to the PCATs, and the C-terminal cargo domain is the domain that carries the function.

#### **1.3.2 Cellular functions of PCATs**

Biochemical studies establish that PCATs serve a dual function, both as a proteolytic maturation for the peptide substrates and an exporter pumping the peptide substrates (Nishie et al. 2011). The proteolytic maturation occurs when the peptidase domain of PCATs cleave off N-terminal domain of the substrates at a conserved cleavage site. This conserved cleavage site contains a conserved double-glycine motif (van Belkum, Worobo, and Stiles 1997). The peptidase domain cleaves the N-terminal leader peptide from the C-terminal cargos. Concomitantly, the C-terminal cargos are transported to the outside of the cells (Havarstein, Diep, and Nes 1995). The requirements for cleavage before transport may evolve as a safety mechanism to prevent active cargo some of which are toxins from harming the secreting cells.

### **1.4 Prior studies of PCATs**

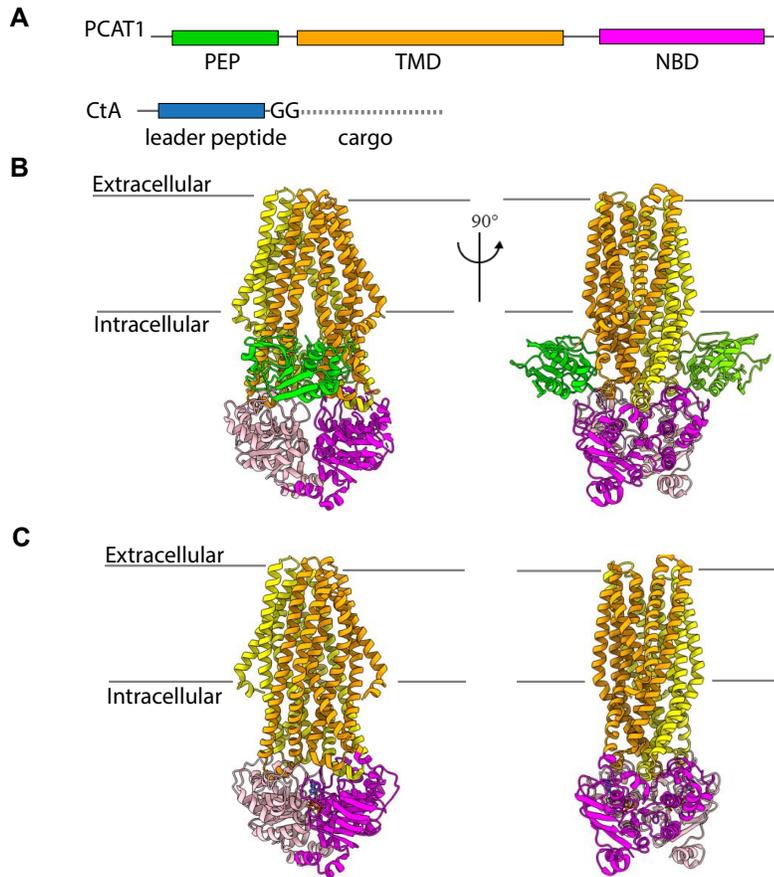
#### **1.4.2 Structural studies of PCATs**

Due to the difficulty of obtaining crystal structures of membrane proteins, early structural studies mainly investigated the soluble domains of PCATs. For instance, a peptidase domain of ComA system from *Streptococcus* has been crystallized (Kotake et al. 2008; Ishii, Yano, and Hayashi 2006; Ishii et al. 2010). Similarly, a peptidase domain of the LahT system from *Lachnospiraceae* has also been crystallized (Bobeica et al. 2019). These studies have established that the peptidase domains of PCATs are structurally and catalytically similar to the papain-like cysteine proteases

as predicted from sequence alignment (Drenth et al. 1968). The study has identified a catalytic active site and a plausible substrate recognition site. Therefore, it has shown that PEPs are important for recognition and cleavage of the substrate.

The first full-length structural study of PCATs is that of PCAT1 from *Clostridium thermocellum*, a Gram-positive thermophilic bacterium (Lin, Huang, and Chen 2015). Two crystal structures are reported: *apo wt*PCAT1 in the inward-facing conformation and ATP hydrolysis deficient mutant E648Q ATP $\gamma$ S-bound occluded conformation (Fig. 1.1). The inward-facing conformation shows TM helices of PCAT1 forming a large  $\alpha$  helical barrel that may be able to accommodate its large 90 amino acid residue peptide. Moreover, the peptidase domain is found to locate adjacent to the cytosolic opening of the transmembrane cavity. This evidence suggests the possibility that perhaps PCATs transport the peptide substrate via the alternating access mechanism. Sequence alignment of PCATs and their substrates suggest that even though the cargo domains are diverse, the N-terminal leader domain of the substrates are well conserved. The double glycine motif at the most C-terminal end of the leader peptide is all conserved among PCAT substrates suggesting a common cleavage mechanism.

Unlike the *apo* inward-facing conformation, the E648Q mutant ATP $\gamma$ S-bound occluded conformation shows dimerized NBD, and the transmembrane cavity is closed off from both cytosol and extracellular space. The peptidase domain cannot be observed in the occluded conformation suggesting that the peptidase domain is flexible relative to the core transporter upon NBD dimerization.



**Fig. 1.2 The crystal structures of *wt*PCAT1 in *apo* inward facing conformation and E648Q PCAT1 in ATP $\gamma$ S-bound occluded conformation**

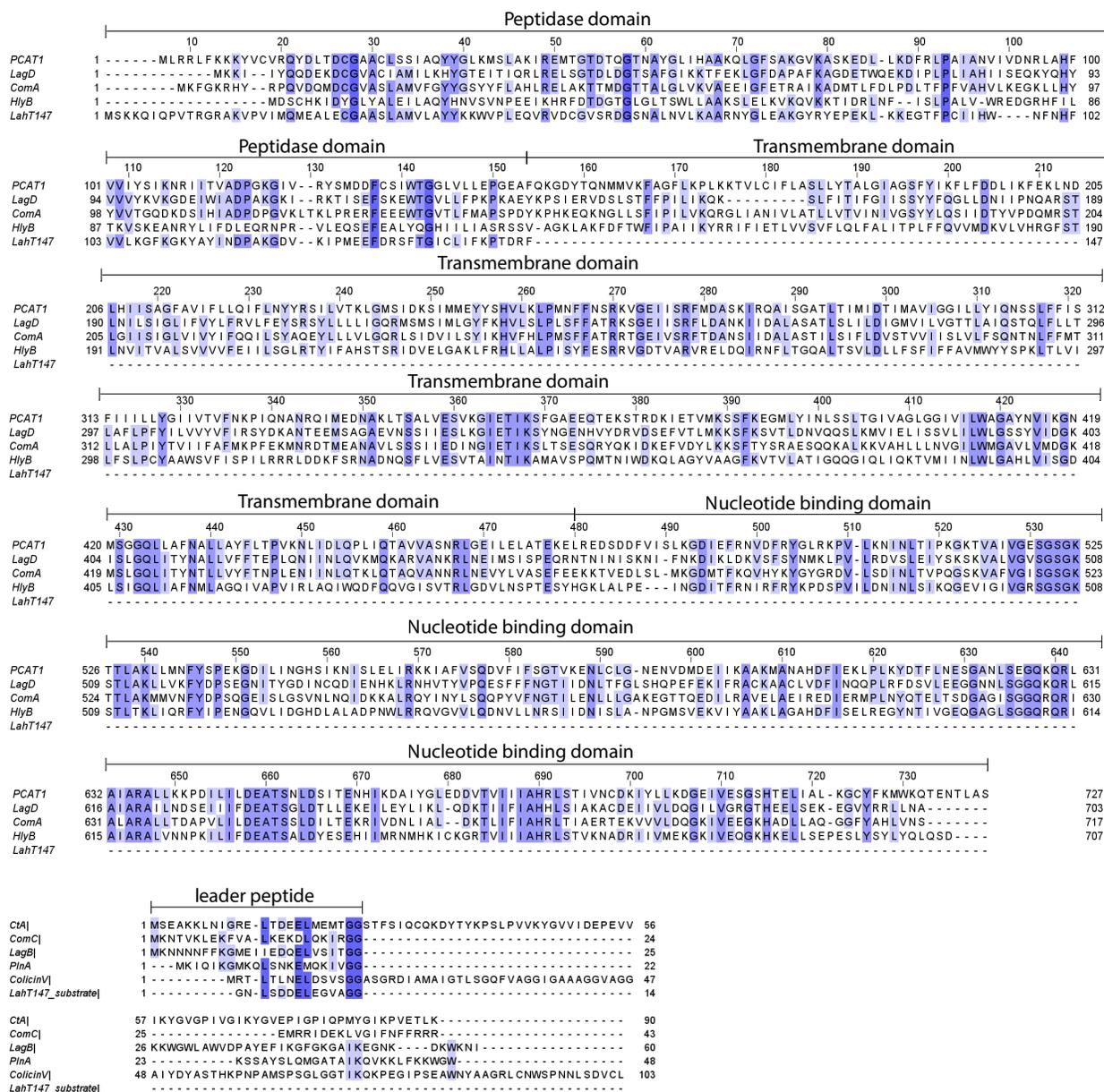
(A) Cartoon illustration depicting a general domain organization of PCAT1 and its substrate CtA

(B) Two orthogonal views of the *apo* PCAT1 structure (PDB: 4ry2)

The cartoon is color-coded by domains. Magenta, NBD;

yellow, TMD; green, PEP.

(C) Two orthogonal views of the ATP $\gamma$ S-bound occluded structure (PDB: 4s0f).



**Fig. 1.3 Domain organization and sequence alignment of PCATs and their peptide substrates**  
Sequence alignments of PCATs and their substrates.  
PCAT1 from *Clostridium thermocellum*, LagD from *Lactococcus lactis*,  
ComA from *Streptococcus pneumoniae*, HlyB from *Escherichia coli*,  
and lahT147 from Lachnospiraceae bacterium C6A11;  
and CtaI from *Clostridium thermocellum*, ComC from *Streptococcus pneumoniae*,  
LagB from *Lactococcus lactis*, plantaricin-A (plnA) from *Lactobacillus plantarum*,  
ColicinV from *Escherichia coli*, and the leader sequence peptide of LahT147. % identity is color-coded in blue.

### 1.3 Dissertation summary

The present dissertation includes all my work in the Chen lab investigating the structure and function of PCAT1.

Chapter 1 contains a general introduction to ABC transporters aiming at laying a background knowledge of ABC transporters relevant for future discussions. I place particular emphasis on the PCAT1 from *Clostridium thermocellum*. This work was performed by Dr. David Lin from the Chen lab and serves as a background for my present work.

Chapter 2 contains the structural and biochemical study of the interaction between PCAT1 and its substrate. I shall present briefly the stoichiometric study of the PCAT1-substrate complex. This question is answered through a native mass spectrometry project performed by my collaborator, Dr. Dominic Olinares, in the Brian Chait Laboratory. Further, I will give a detailed description of the cryo-EM work on the PCAT1-substrate complex. This work includes the study of catalytic proteolytic mechanisms of PCATs, the substrate recognition mechanism, and the substrate orientation in the translocation cavity.

Chapter 3 contains the structural study of PCAT1 aiming at capturing conformational states of PCAT1 in the active transport cycle. I shall focus on the cryo-EM study of PCAT1 in the ATP-regenerating active turnover condition. This work attempts to investigate how PCATs couple substrate proteolytic cleavage to ATP hydrolysis. Furthermore, I will describe how the conformational changes in PCAT1 enable strict coupling between substrate cleavage and transport.

Chapter 4 explores the biophysical study of PCAT1 using native mass spectrometry technique. This project aims at capturing the real-time dynamics of PCAT1 proteolysis and ATP hydrolysis activities. The mass spectrometric work was performed by Dr. Dominic Olinares from the Brian Chait Laboratory.

Chapter 5 discusses the conclusions that the thesis work has advanced. I shall compare the work on PCAT1 in the context of structural studies of other ABC transporters. In the end, I shall discuss a plausible research direction that flow from my work, what future advancements in the field may look like, and what limitations a future investigation may encounter.

## CHAPTER 2: Cryo-EM structure of the substrate-bound PCAT1

### 2.1 Summary

The first full-length structural study of PCAT1 in 2015 yielded two valuable structures: the *apo* inward-facing conformation, and the ATP $\gamma$ S bound occluded conformation (Lin, Huang, and Chen 2015). Despite these structural data, it was not clear how the peptide substrates interact with PCAT1. Therefore, obtaining the structure of PCAT1-substrate complex is critical. In this project, I solved the structure of PCAT1-substrate complex to 3.3 Å resolution by cryo-EM single particle reconstruction technique. Overall, transporters resemble the *apo* inward-facing conformation. In addition, we found the substrate bind to the peptidase domain. However, only one substrate is oriented for translocation. Through native mass spectrometry (native MS) experiments, we validated the stoichiometry of the complex. Disulfide cross linking experiments confirmed the substrate is translocated in an unfolded state. This work enables us to propose the transport mechanism of PCAT1 where two peptide substrates are bound, yet one is transported.

This chapter is adapted from my work co-authored with Paul Dominic B. Olinares, Anthony Palillo, Michael L. Oldham, Brian T. Chait, and Jue Chen. The work was published in *eLife* on January 14, 2020 (Kieuvongngam et al. 2020).

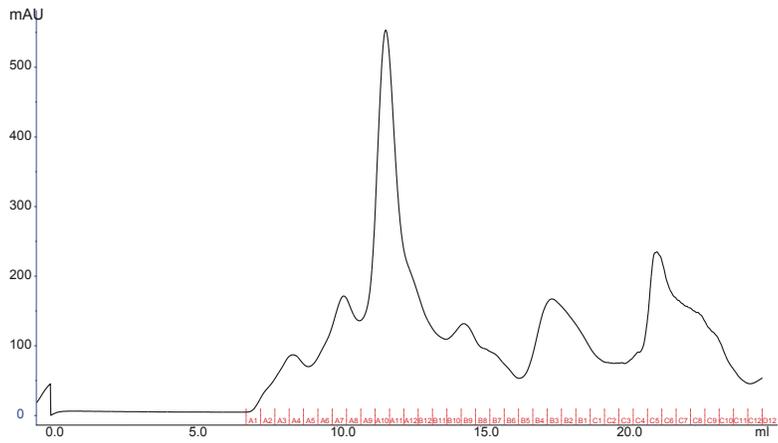
### 2.2 Results

#### 2.2.1 Stoichiometry determination of PCAT1-substrate complex

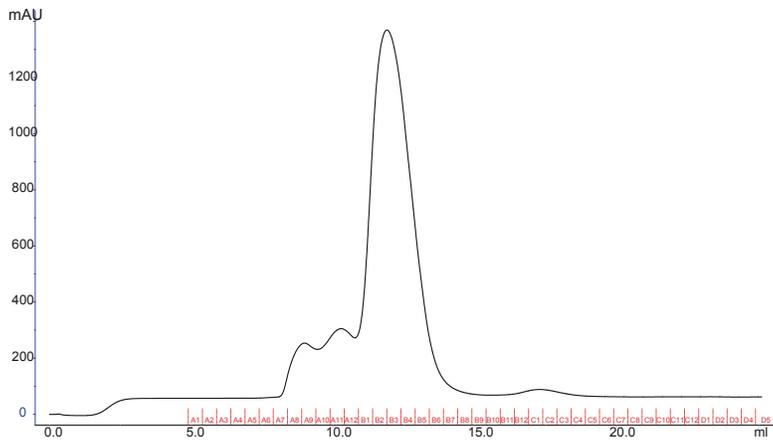
PCAT1 contains two identical PEP domains, raising the question of whether two copies of the substrate can bind and be translocated simultaneously. To address this question, we used native MS to determine the maximum binding capacity of PCAT1.

To obtain a stable PCAT1-CtA complex, we used a proteolytic-deficient mutant (C21A) that can bind but does not cleave the substrate. We purified C21A PCAT1 using Glutathione affinity column followed by gel filtration chromatography (Fig. 2.1). We purified CtA from inclusion bodies using 6xHis affinity column under 8M urea condition to unfold the aggregates. CtA is then refolded overnight in urea-free buffer. The final gel filtration chromatography is used to further purify CtA (Fig. 2.2).

Two-fold molar excess of the substrate was mixed with PCAT1 prior to MS analysis. In both detergents, we observed two main assemblies of 172.3 and 182.5 kDa, corresponding to dimeric PCAT1 bound to one or two CtA, respectively. The result indicates that PCAT1 can bind up to 2 substrates, most likely one on each peptidase domain (Table 2.1, Fig. 2.3).

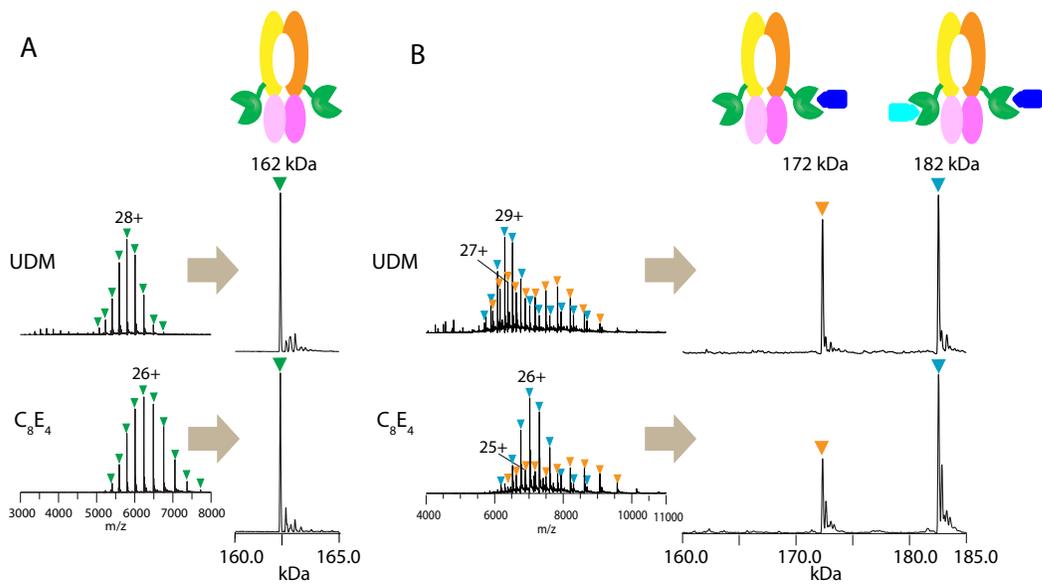


**Fig. 2.1** Gel filtration chromatography profile of PCAT1



**Fig. 2.2** Gel filtration chromatography profile of CtA



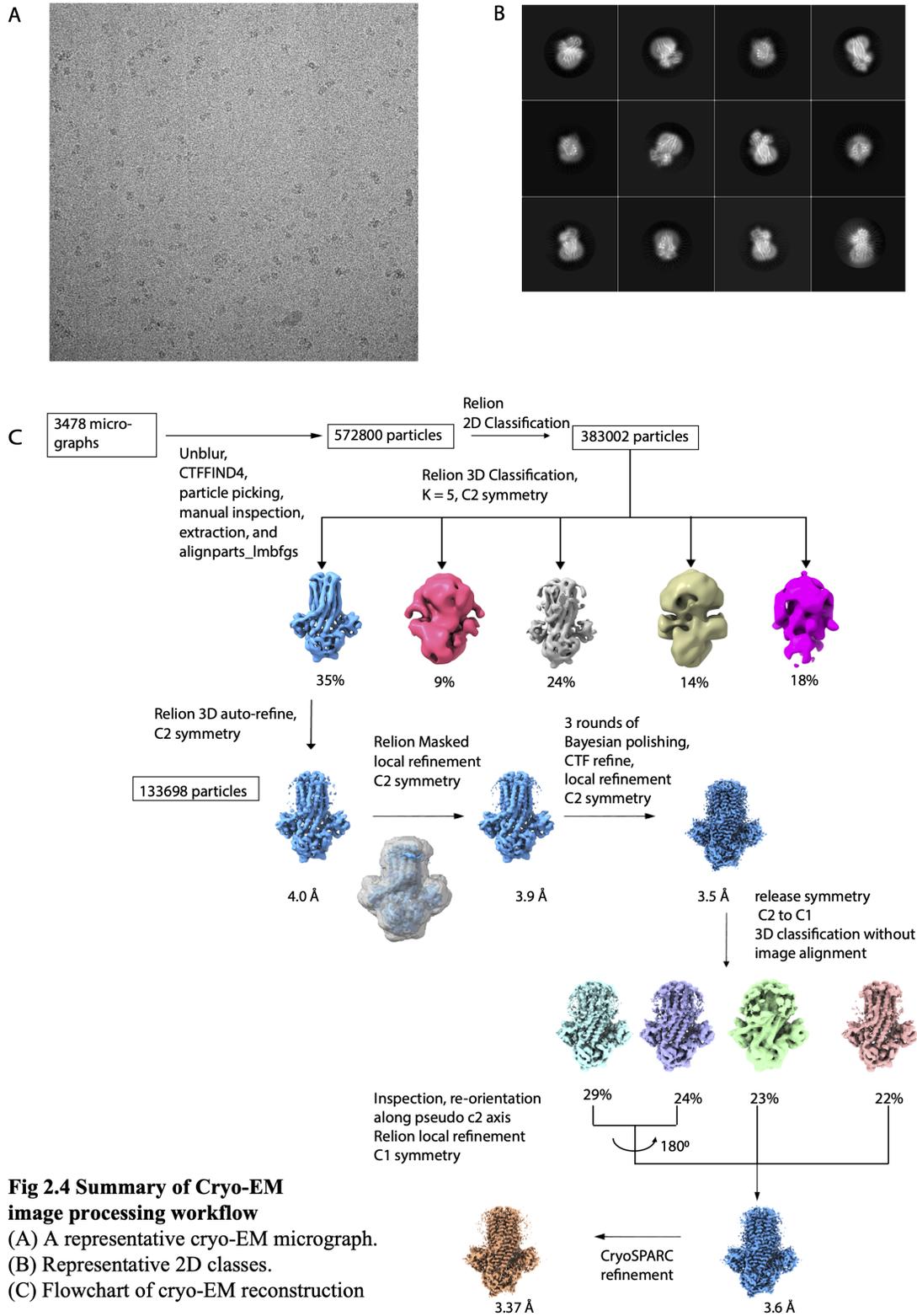


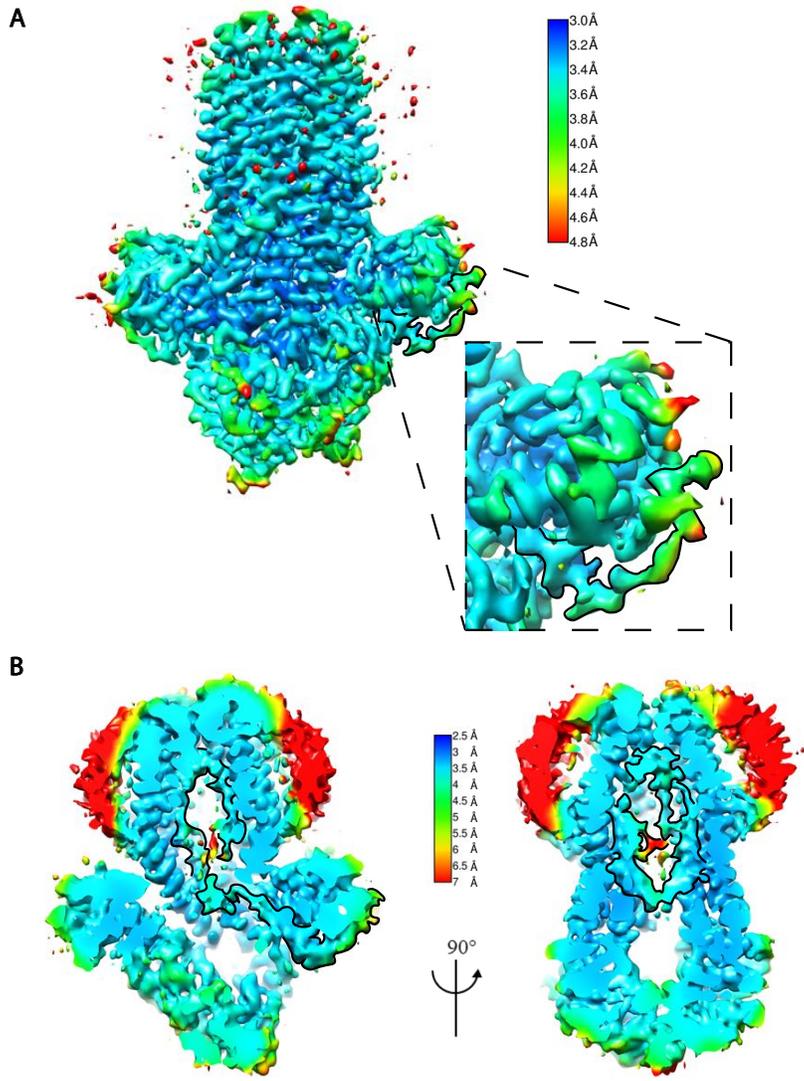
**Fig. 2.3 Native mass spectrometry analysis of *wt*PCAT1 and C21A PCAT1-CtA complex**  
 Samples were analyzed in 200 mM ammonium acetate containing  $C_8E_4$ . The mass of the protein complex corresponding to PCAT1 complex is indicated. (A) *wt*PCAT1. (B) C21A PCAT1 in the presence of 2x molar ratio CtA.

### 2.2.2 Structural determination of PCAT1-substrate complex

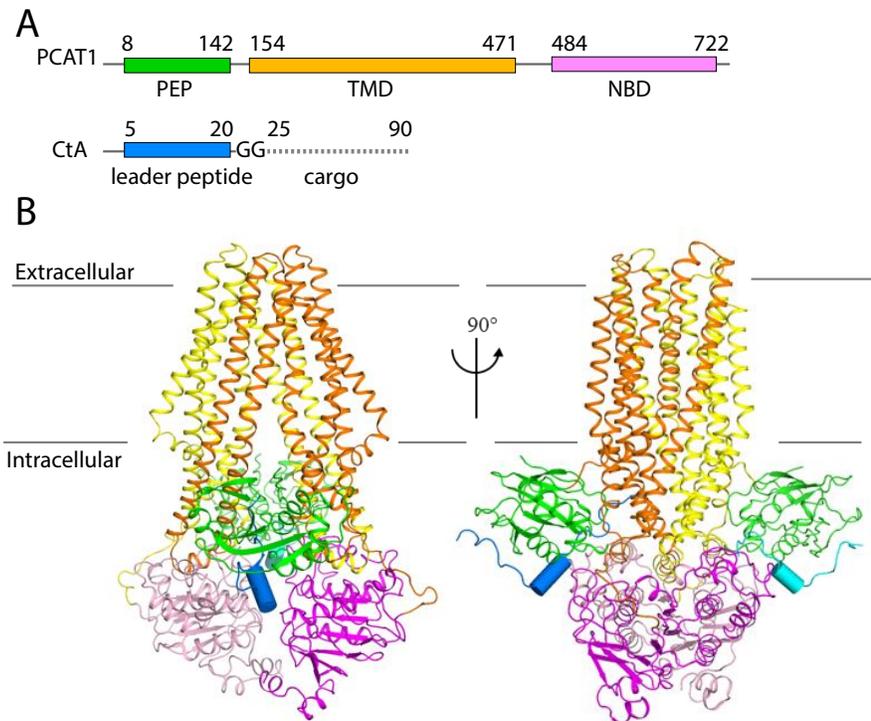
To obtain the samples of PCAT1-substrate complex, we incubated excess substrates to proteolytic-deficient mutant (C21A) PCAT1. The samples were plunge-frozen in liquid ethane for cryo-EM data collection. Using single-particle reconstruction software Relion and CryoSPARC (Zivanov et al. 2018; Kimanius et al. 2016; Punjani et al. 2017), we were able to obtain a cryo-EM density map of the PCAT1-substrate complex without symmetry imposed (Fig. 2.4, Table 2.2). The local resolution from the Bsoft package ranges from 3.0 to 5.5 Å (Fig. 2.5) (Heymann and Belnap 2007). The TMDs are much more well-resolved than NBDs and PEPs, indicating the local flexibility of the NBDs and PEPs. The extra-density belonging to the substrate N-terminal leader peptide is observed on both binding grooves of the PEPs, consistent with the native MS analysis. In this substrate-bound ATP-free conformation, the TM cavity is continuous with the cytosol and closed off to the extracellular space. The two PEP domains, each binding a CtA molecule, interact with the transporter core at the TMD/NBD interface. The structure is two-fold symmetric, except for the C-terminal regions of the substrate, a point that we will discuss in detail below (Fig 2.6).







**Fig. 2.5 Local resolution estimation of the PCAT1-CtA complex density map**  
 (A) The local resolution estimation of the cryo-EM density map. The box shows the magnified view of the substrate density highlighted in the thick boundaries.  
 (B) Cross sectional view of the translocation pathway.



**Fig. 2.6 The cryo-EM structure of PCAT1-CtA complex**

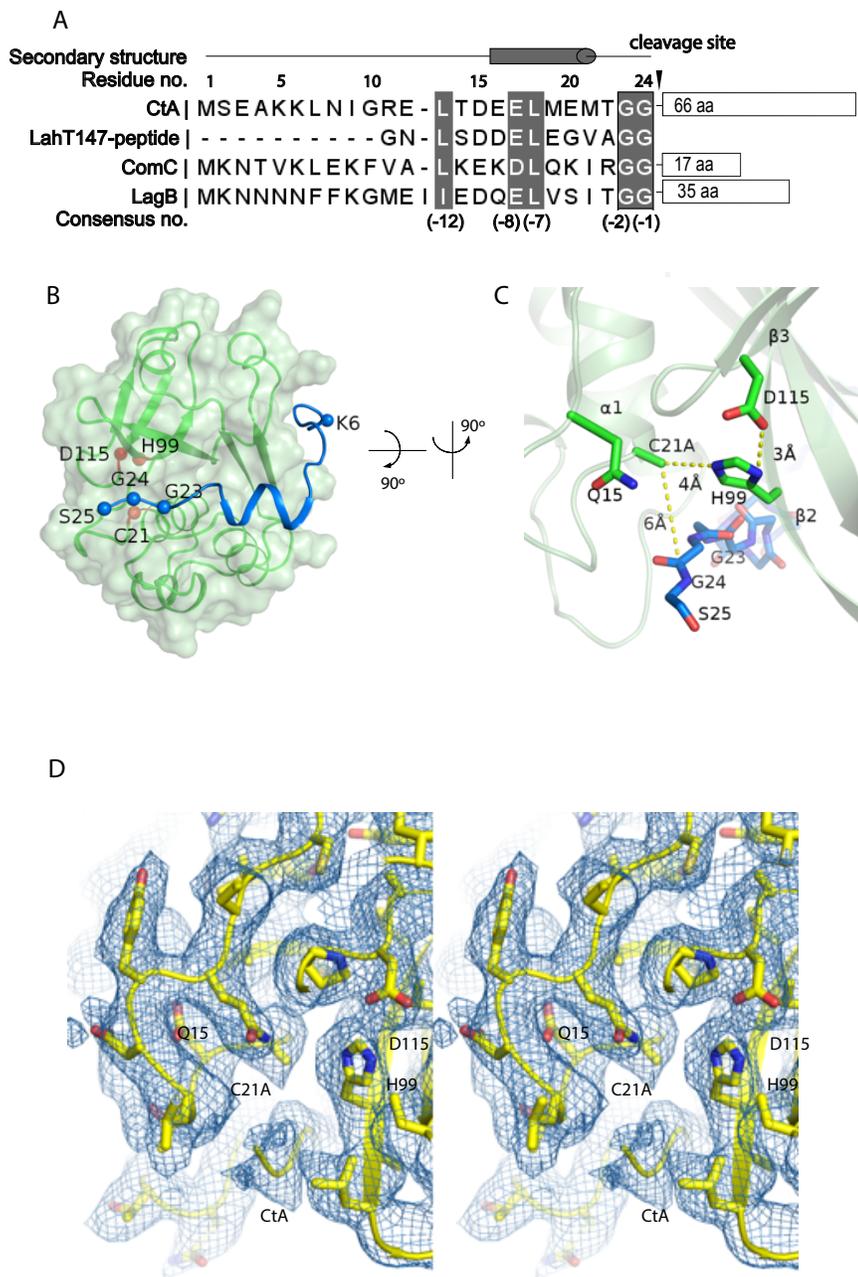
(A) Cartoon illustration depicting domain organization of PCAT1 and its substrate, CtA. The symbol GG denotes the double glycine motif. The dotted line represents the unstructured cargo region.

(B) Two orthogonal views of the PCAT1-CtA complex structure, the leader peptide helix of CtA is shown as a cylinder. The cartoon is color-coded by domains. Blue-cyan, CtA; magenta-pink, NBD; yellow-orange, TMD; green, PEP.

### 2.2.3 Investigating the proteolysis mechanism of PCAT1

The sequence of the CtA leader peptide is homologous to that of other PCAT substrates in Gram-positive bacteria. The cleavage site contains a consensus sequence of L(-12)XXXE(-8)L(-7)XXXXG(-2)G(-1). Residues 8–25 of the leader peptide form a L-shaped structure that wraps around the PEP domain, with the N-terminus in the cytosol and the C-terminal end at the lateral opening of the translocation pathway. Residues 15–21 form a short helix, docking onto a hydrophobic groove on PEP. The double-glycine motif (G23 and G24), inserted to the active site of PEP, interacts with catalytic residues C21A and H99 (Fig. 2.7). The total buried surface of the leader peptide by the PEP domain is approximately 740 Å<sup>2</sup>.

The substrate-binding cleft at the cleavage site is very narrow, which explains the specificity for the double-glycine motif in proteolysis. The C $\beta$  atom of C21 (mutated to alanine) is about 5.5 Å away from G24, consistent with C21 serving as a nucleophile to attack the substrate backbone (Schechter and Berger 1967). H99 orients the N atom of its imidazole ring towards the C21A, consistent with its role to polarize the attacking C21 (Husain and Lowe 1968; K. H. Wu and Tai 2004). The side chain of D115 forms a hydrogen bond with H99, maintaining H99 in an electronegative state and a catalytically favorable position (Vernet et al. 1995). In addition, the highly conserved Q15 residue is in the vicinity, suggesting that it may function as an oxyanion hole that stabilizes the tetrahedral intermediate (Ménard et al. 1991). The configuration of the active site is typical of a cysteine protease, indicating that PCAT1 processes the CtA substrate via the common catalytic thiol mechanism (Drenth et al. 1968; Kamphuis et al. 1984).



**Fig. 2.7 Substrate-peptidase domain interaction**

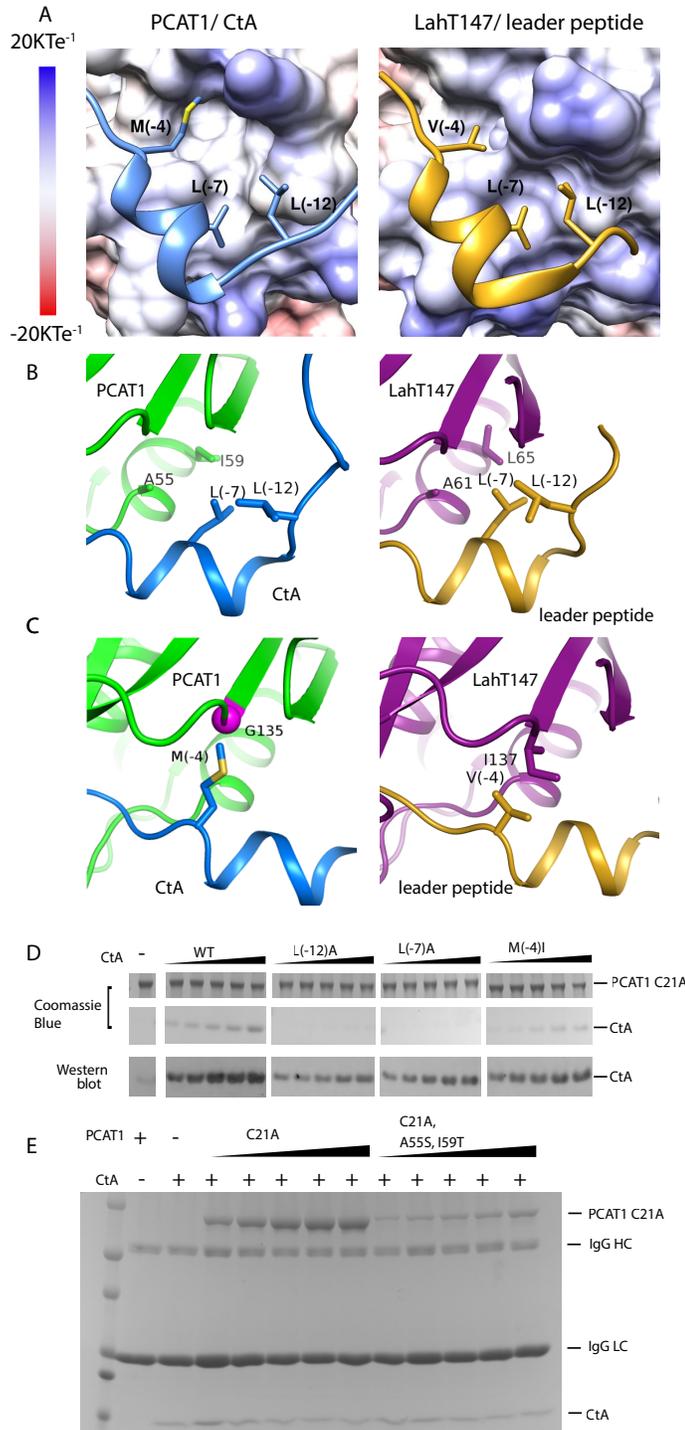
(A) Sequence alignment of different leader peptides. Consensus residues are highlighted in gray. (B) Docking of the leader peptide (blue) on the PEP surface (green). The catalytic triad residues are indicated as red spheres and the CtA double glycine motif are indicated as blue spheres. (C) A zoomed-in view of the catalytic site. (D) A stereoview of the catalytic site. The cryo-EM map is shown as mesh.

## 2.2.4 Mechanistic details of substrate recognition

The structure of the PEP domain is highly conserved among different PCATs. The overall root-mean-square deviations (RMSDs) among the structures of ComA (PDB: 3K8U) (Ishii et al. 2010), LahT147 (PDB: 6MPZ) (Bobeica et al. 2019), and the PEP domain of PCAT1 are approximately 1 Å. In addition, substrates of the different PEPs share a consensus leader peptide sequence, yet each PEP specifically recognizes one or a small subset of substrates.

What are the common substrate recognition motifs and what are the unique features that confer specificity? The leader peptides of CtA and the LahT147 substrate both contain a two-turn helix consisting of residues from positions  $-10$  to  $-4$ , docking into a shallow groove on the surface of the PEP. A hydrophobic knot, formed by the two leucine residues at positions  $-7$  and  $-12$  of the leader peptide and two hydrophobic residues on PEP, is observed in both structures (Fig 2.8 A-C). To test the importance of this interaction, we mutated the conserved leucine residues in CtA and estimated its affinity for PCAT1(C21A) using a pull-down assay (Fig. 2.8 D-E). Compared to the *wt* CtA, the L( $-7$ )A and L( $-12$ )A mutants have much lower affinity for PCAT1. The pulled down PCAT1 could only be detected by Western blot even at the highest concentration of CtA tested. We found hydrophobic residues A55 and I59 on PCAT1 that are within the Van der Waals radius of L( $-7$ )A and L( $-12$ )A, and are conserved in LahT147 as well. Furthermore, reciprocal mutations introduced to residues A55 and I59 of PCAT1 lowered its affinity for CtA by more than eightfold, consistent with the role that these residues play on the CtA recognition. This hydrophobic interaction is likely a common feature shared by different PEP domains. Alanine substitution of either one of the two leucine residues in the ComA substrate decreased its affinity by approximately twentyfold (Kotake et al. 2008). Similarly, LahT147 does not process precursor peptides with substitutions at L( $-7$ ) or L( $-12$ ) (Bobeica et al. 2019).

Mutational data at the  $-4$  position indicate that this residue is also important for the PEP-substrate interaction; however, in contrast to the highly conserved L( $-7$ ) and L( $-12$ ) residues, the amino acid identity at position  $-4$  varies among different substrates and appears to correlate with the interacting residue on PEP. For example, M ( $-4$ ) of CtA packs against G135 of PCAT1. In LahT147, a much smaller residue V( $-4$ ) of the substrate interacts with a larger residue I137 on PEP. Therefore, it seems that the leader peptide position  $-4$  confers specificity among different PCATs.



**Fig. 2.8 Common PCAT residues for substrate recognition**

(A) Insertion of three hydrophobic residues onto a hydrophobic groove on the surface of the PEP domain, represented as electrostatic potential surfaces. CtA is shown in blue and the LahT147 substrate in gold.

(B) A conserved hydrophobic knot. (C) Residue at position -4 of the substrate conferring specificity.

(D) Pull-down of the *wt* or mutant CtA using PCAT1(C21A) conjugated resin. The CtA constructs contain a 3x Flag tag at their C-terminus. Western blot was performed using the anti-Flag M2 antibody.

(E) Pull-down of PCAT1(C21A) using the M2 antibody resin against the Flag-tagged CtA.

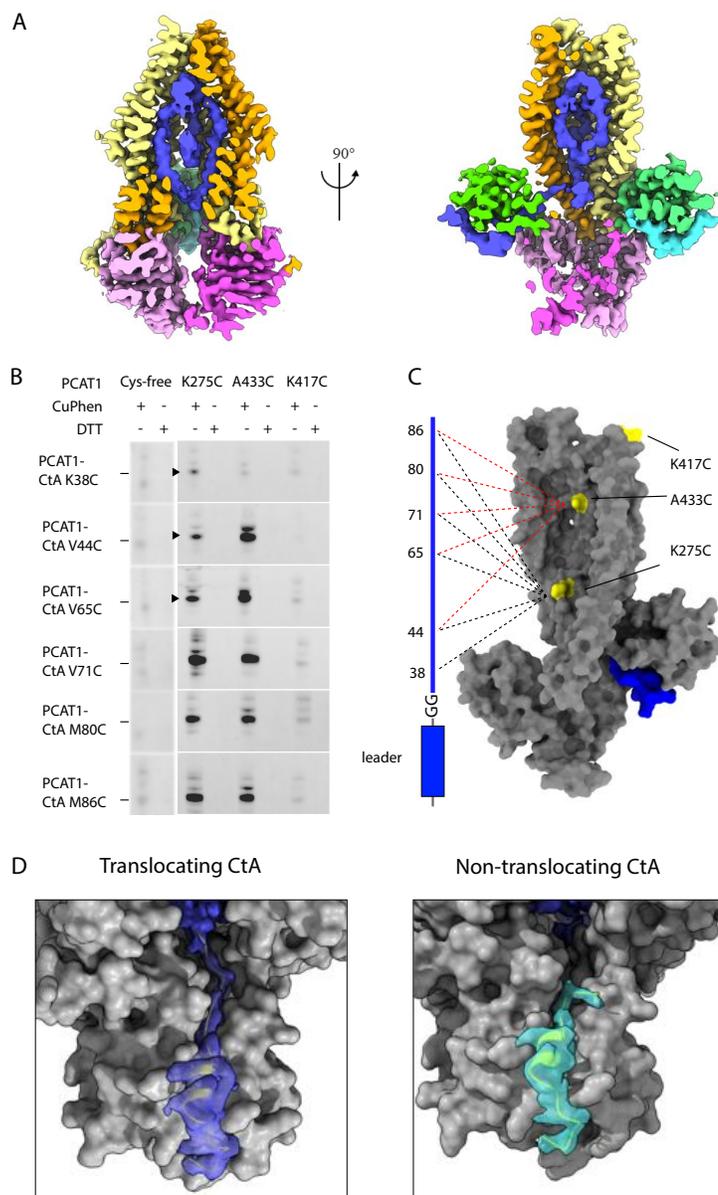
Mutation of hydrophobic residue A55 and I59 were introduced to PCAT1(C21A) background.

IgG HC and IgG LC denote the heavy chain and the light chain of the M2 antibody, respectively.

## 2.2.5 Asymmetric positioning of the cargos

In contrast to the leader peptide, the cargo region (residues 25–90) is flexible and less defined. Inside the translocation pathway, we observe density of elongated shape packed along the interior surface of the cavity shown in the composite cryo-EM map of PCAT1 and CtA (Fig 2.9 A). To test if this density corresponds to the C-terminal region of CtA, single cysteines were introduced at six positions downstream of the leader peptide. We also placed single cysteine, in three position, in the otherwise cysteine-free PCAT1: K275C and A433C on the interior surface of the TM cavity and K417C in an extracellular loop outside the translocation pathway. The proximity of the cysteine pairs was analyzed by mixing CtA with PCAT1 in the presence of the oxidizing agent CuPhen or the reducing agent DTT, and then using Western blot to detect the appearance of PCAT1-CtA crosslinked product (Fig. 2.9 B-C). Mass spectrometry analysis was performed on three samples excised from the SDS-PAGE gel. In each case, the expected cross-linked peptide containing the engineered disulfide bond was observed. Cysteines at multiple positions along CtA can be crosslinked to K275C and A433C on PCAT1, suggesting that the cargo inside the cavity is unstructured and flexible, consistent with the amorphous nature of the EM density. The cysteine at position 38 of CtA can be weakly crosslinked to K275C but not to K433C, which is consistent with the structure. K38 is near the leader peptide at the cytoplasmic side of the membrane and is not able to reach the extracellular ceiling of the cavity prior to proteolytic cleavage. No crosslinking product was observed for the cysteine placed outside the TM cavity (K417C), indicating that the reaction is specific.

The density inside the TM cavity appears to be connected to only one of the two leader peptides, raising the question whether one or two cargos are inside the TM cavity. We asked if the TM cavity is large enough to accommodate two cargos. Using the 3V server (Voss and Gerstein 2010), the volume of the TM cavity is estimated to be 15,000 Å<sup>3</sup>. Considering the average density of a small protein is around 1.35 g/cm<sup>3</sup> (Fischer, Polikarpov, and Craievich 2009), the size of a CtA cargo is approximately 9000 3Å<sup>3</sup>, a conservative estimate for an unstructured protein. Thus, it is likely that only one cargo is inserted into the TM cavity. We termed the cargo inside the TM cavity the ‘translocating CtA’ as it is positioned for cleavage and translocation (Fig. 2.9 D). The density for the cargo region is weak and amorphous, suggesting that there is no specific interaction between the cargo and transporter, and the substrate specificity is conferred through the leader peptide only. For the other substrate, the non-translocating CtA, no density is observed beyond S26, two residues beyond the double glycine motif. Furthermore, G23, G24, and S25 form a kink, orienting the C-terminus towards the cytoplasm. There is also biochemical evidence that the C-terminus of CtA is accessible: the PCAT1-CtA complex can be pulled down using an antibody against a tag placed at the C-terminus of CtA. Based on these observations, we suggest that the non-translocating cargo is located in the cytosol, flexibly linked to the leader peptide.



### Fig 2.9 Positions of the two cargos

(A) Two orthogonal views of the the translocation pathway. A composite cryo-EM map displays PCAT1 density, the CtA density, and the density inside the translocation pathway. The cryo-EM density inside the TM cavity (shown in blue) likely corresponds to the C-terminal region of one CtA molecule.

(B) Disulfide crosslinking experiments between PCAT1 and CtA. Crosslinked PCAT1-CtA products were detected by Western blot using an anti-HA antibody against HA-tagged CtA. Three bands (indicated by arrows) were excised from the SDS-PAGE and analyzed by mass spectrometry. In all three cases, peptide fragments with the correct disulfide bond were identified.

(C) Summary of the crosslinking results. A PCAT1 monomer is shown as grey surface and CtA is represented by a cartoon. The dotted lines illustrate the crosslinked pairs (black, 275C crosslink pairs; red, 433C crosslinked pairs).

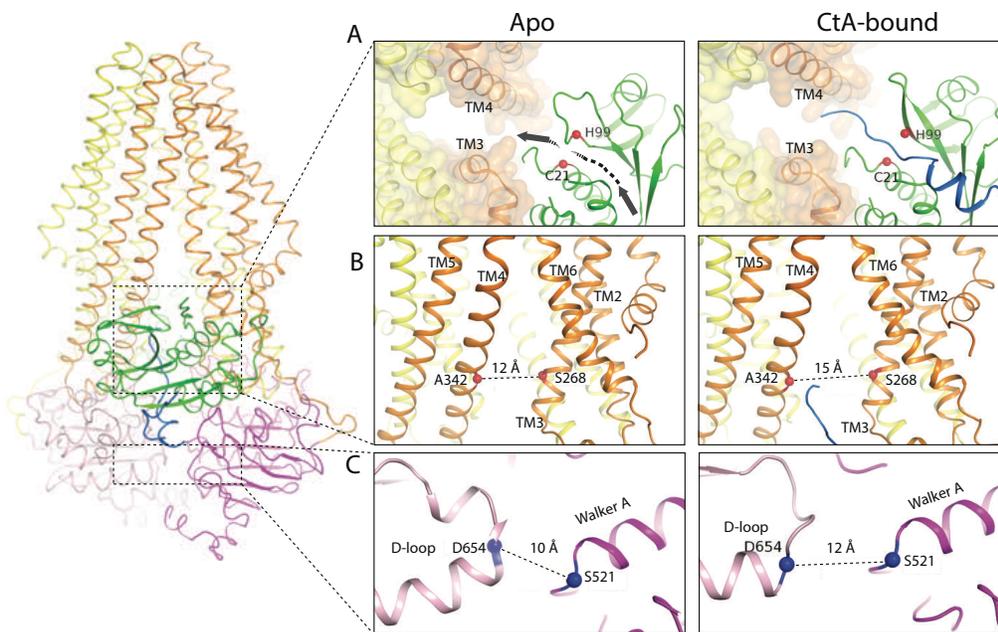
(D) Comparison of the cryo-EM density at the two catalytic sites. The density is displayed as surface. Dark blue, the translocating CtA; light blue, the non-translocating CtA. The TMD and PEP are shown as grey surfaces and the NBDs are omitted for clarity. The leader peptides are shown as ribbons.

## 2.2.6 Conformational changes upon substrate binding

Comparing the structures of PCAT1 in the presence and absence of CtA shows that in the substrate-bound conformation, the intracellular opening of the TM cavity (formed from TM3 and TM4) is approximately 3 Å wider and the two NBDs are further apart (Fig 2.10).

The conformational change of PCAT1 is opposite to that of the multidrug transporter MRP1 (Johnson and Chen 2017), in which substrate binding brings the two NBDs closer. MRP1, like many other ABC transporters, has higher ATPase activity in the presence of its substrate (Jin et al. 2012a; M. Chen, Abele, and Tampé 2003). In contrast, addition of CtA reduces the ATPase activity of PCAT1 by about 10% for the wild-type protein and 50% for the cleavage-incompetent C21A mutant. As ATP hydrolysis requires NBD dimerization, the structural observation correlates well with the distinct biochemical property of this protein-conducting ABC transporter. The insertion of a large cargo into the TM cavity would likely decrease the rate of NBD dimerization thereby slowing down ATP hydrolysis.

Binding of the leader peptide also induces local conformational changes within the PEP domain. The largest displacement, approximately 5 Å, occurs in a loop (residues 91–98) preceding the catalytic residue H99. As a result, in the substrate-bound structure, a path between the catalytic site and the TM cavity opens, permitting insertion of the C-terminal cargo into the translocation pathway at the lateral opening between TM3 and TM4.



**Fig. 2.10 Conformational changes upon CtA binding**

(A,B) Widening of the lateral opening of the TM pathway. The arrow in the apo structure delineates the CtA binding groove. The catalytic residues C21 and H99 are labeled. The  $C\alpha$  distances between residues S268 and A342 that line the lateral opening are shown. (C) The NBDs are further apart in the CtA-bound structure. The  $C\alpha$  distances between S521 in the Walker A motif and D654 in the D-loop of the other NBD are indicated.

## 2.3 Discussion

### 2.3.1 Location of the substrate recognition site

The current understanding of substrate recognition of ABC exporters is that the substrates interact with the core transporters directly through TMDs. The structures of ABC exporters, such as the multidrug transporters MRP1 (Johnson and Chen 2017), ABCG2 (Taylor et al. 2017), P-glycoprotein (Jin et al. 2012b), the Lipid A flippase MsbA (Mi et al. 2017), LPS transporter LptB<sub>2</sub>FG (Tang et al. 2019), and the promiscuous exporter TmrAB (Nöll et al. 2017) all reveal a well-defined substrate-binding site at the interface of the two TMDs.

The structure of PCAT1-substrate complex presents a different mechanism by which the polypeptide substrates are recognized. Instead of the TMDs, the cytosolic peptidase domain binds substrate through its N-terminal leader peptide sequence. The common feature shared by PCATs is a hydrophobic knot formed by two residues of the peptidase domain and residues at positions -7 and -12 of the leader peptide. Position -4 of the leader peptide likely confers substrate specificity for different PCATs.

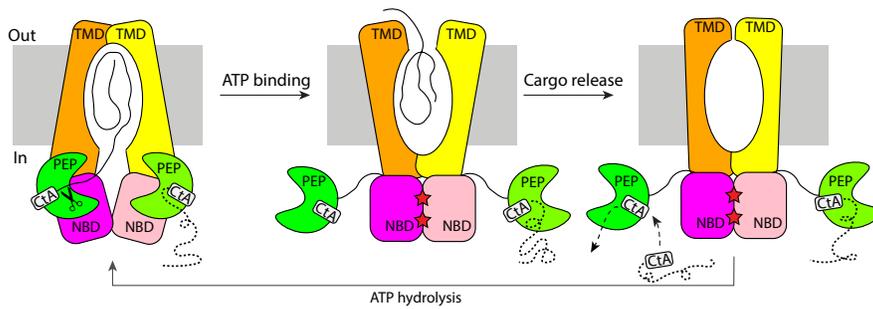
### 2.3.2 Flexibility of polypeptide cargoes in the translocation cavity

Unlike most ABC exporters, the TMDs of PCAT1 do not provide any specific binding region for the cargo. Instead, they merely provide a large conduit to transverse the lipid bilayer. This observation explains two phenomena: First, the PCAT system can be used to secrete different cargo proteins that are tethered to the same secretion signal. Second, the cargo is flexible inside the TM cavity. The absence of a high-affinity binding site in the TMDs is reminiscent of the *E. coli* cobalamin importer BtuCD, where the substrate is recruited by the periplasmic binding protein and the TMDs provide a Teflon-like pathway for conduction (Korkhov, Mireku, and Locher 2012a; Korkhov et al. 2014).

### 2.3.3 Proposal of PCAT1 substrate binding, processing, and transport

The structure of the substrate-bound PCAT1 enables us to expand the previous working model of how polypeptides are processed and transported by PCATs in Gram-positive bacteria (Fig. 2.11). PCAT1 adopts an inward-facing conformation in the absence of ATP, in which the NBD interface is open and each PEP domain docks onto a lateral opening of the TM pathway. Two substrates can bind to PCAT simultaneously, but only one substrate is positioned for cleavage and translocation. The translocating substrate inserts its C-terminal cargo into the TM cavity and the corresponding PEP domain cleaves the substrate to free the cargo. The other non-translocating cargo is located in the cytosol. ATP binding brings the two NBDs into close contact, orienting the translocation pathway towards the extracellular space. As the TM pathway confers no specific binding of the substrate, the cargo will be released and, subsequently, the TMDs will isomerize to form the occluded conformation where both the extracellular gate and the cytoplasmic gate are closed (PDB: 4S0F). Formation of the closed NBD dimer disengages the PEP domains from the core transporter permitting release of the leader peptide and binding of a new substrate. ATP hydrolysis and ADP release reset the transporter to the inward-facing conformation allowing the PEP to dock back and position one of the two substrates for cleavage and translocation.

A key aspect of this model is the strict coupling of substrate cleavage to translocation. In the inward-facing conformation where two substrates are bound, the non-translocating substrate adopts a kink at the double-glycine motif, preventing uncoupled cleavage. In the ATP-bound conformation, the disengaged PEP domains have little proteolytic activity as the transporter core is necessary for substrate cleavage.



**Fig. 2.11 The alternating access mechanism of PCAT1**

CtA is recruited and cleaved in the inward-facing conformation. ATP binding stabilizes the outward-facing conformation in which the PEP domains are disengaged. After cargo release, TMDs isomerize to form an occluded cavity. ATP hydrolysis resets PCAT1 back to the inward-facing conformation, allowing PEP to dock into the TMD-NBD interface.

## 2.4 Materials and Methods

### 2.4.1 Native mass spectrometry analysis

The purified PCAT1 samples were buffer exchanged into native MS solution (200 mM ammonium acetate pH 7.5 with either 0.058% UDM or 0.5% C<sub>8</sub>E<sub>4</sub>) using Zeba microspin desalting columns (Thermo Scientific) with a 40 kDa molecular weight cut-off (MWCO). The typical concentrations used for native MS analysis were 4 μM PCAT1 monomer for the PCAT1-only sample and 4 μM PCAT1(C21A) monomer + 8 μM CtA (a two-fold excess of CtA). An aliquot (2–3 μL) of the sample was loaded into a gold-coated quartz capillary that was fabricated in-house. The sample was then electrosprayed into an Exactive Plus EMR instrument (Thermo Fisher Scientific) using a static nanospray source. The MS parameters used include: spray voltage, 1.0–1.4 kV; capillary temperature, 125°C; S-lens RF level, 200; resolving power, 8750 or 17,500 at *m/z* of 200; AGC target,  $1 \times 10^6$ ; number of microscans, 5; maximum injection time, 200 ms; injection flatapole, 8 V; interflatapole, 4 V; bent flatapole, 4 V; ultrahigh vacuum pressure,  $7\text{--}10 \times 10^{-10}$  mbar; total number of scans, 100. The in-source dissociation (ISD) and higher-energy collisional dissociation (HCD) parameters were varied and optimized accordingly. Mass calibration was performed using cesium iodide. The acquired MS spectra were visualized using Thermo Xcalibur Qual Browser (version 3.0.63) and deconvolution was performed either manually or using UniDec version 3.2 (Marty et al. 2015; Reid et al. 2019a). The deconvolved spectra from UniDec were plotted using the *m/z* software (Proteometrics LLC). Experimental masses were reported as the average mass ± standard deviation (S.D.) across all calculated mass values obtained within the observed charge state series.

### 2.4.2 Expression and purification of PCAT1

The *wt* and mutant PCAT1 genes were subcloned into the pMCSG20 vector with an N-terminal glutathione-S-transferase (GST) tag and a Tobacco Etch Virus (TEV) protease cleavage site. *E. coli* strain BL21(DE3) codon plus (RIL) cells expressing PCAT1 were lysed and solubilized in buffer containing 1% *n*-dodecyl-β-D-maltoside (DDM; Anatrace), 50 mM Tris pH 7.0, 500 mM NaCl, 10% glycerol, and 5 mM DTT. PCAT1 was enriched on Glutathione Sepharose 4B resin, washed with buffer containing 50 mM Tris pH 7.0, 500 mM NaCl, 10% glycerol, 5 mM DTT, and 2 mM *n*-undecyl-β-D-maltopyranoside (UDM; Anatrace). The GST tag was removed by cleavage with TEV protease, and PCAT1 was further purified using a Superdex 200 increase column (GE Healthcare) in a buffer containing 50 mM Tris pH 7.0, 150 mM NaCl, 2 mM UDM.

### 2.4.3 Expression and purification of PCAT1 substrate

The gene encoding CtA was subcloned into the pMCSG7 vector with an N-terminal TEV-cleavable 6x His tag and a C-terminal 3x Flag or HA-tag. RIL cells expressing CtA were resuspended in lysis buffer (50 mM Tris pH 7.0, 150 mM NaCl, and 10% glycerol), lysed by three passes through a high-pressure homogenizer (Emulsiflex C-3; Avestin), and centrifuged at 80,000 g for 40 min to isolate inclusion bodies. The pellet was washed extensively with lysis buffer plus 1% Triton X-100, then with lysis buffer alone, before solubilizing in 8 M urea. The denatured protein was purified on cobalt affinity resin (Clontech Laboratories) and refolded via dialysis in lysis buffer plus 5 mM DTT. The His-tag was removed by cleavage with TEV

protease and the protein were further purified by gel-filtration chromatography (Superdex 75 HiLoad 16/60, GE Healthcare).

#### 2.4.4 Cryo-EM sample preparation and data collection

Purified PCAT1(C21A) (5 mg/ml, 62  $\mu$ M monomer) was mixed with 62  $\mu$ M refolded CtA and incubated on ice for 30 mins. About 3  $\mu$ l of sample was applied onto glow-discharged holey carbon grids (Quantifoil gold R1.2–1.3), incubated for 20 s at 100% humidity, and blotted with filter paper for 3 s before being plunge-frozen into liquid ethane using a Vitrobot Mark IV (FEI). A dataset of 3,478 movies was collected on the Titan Krios Transmission Electron Microscope (FEI) outfitted with a K2 Summit direct electron detector (Gatan) with a super-resolution pixel size of 0.545 Å using Leginon (Suloway et al. 2009). The electron dose rate was eight electrons/pixel/s with a total exposure time of 12 s resulting in a total electron dose of 80 electrons/Å<sup>2</sup> over 60 frames.

#### 2.4.5 Single particle reconstruction

The procedure for image processing is summarized in Fig. 2.4. Movie frames were corrected using a gain reference and binned by a factor of two to a pixel size of 1.09 Å. Movie frames were aligned using Unblur (Grant and Grigorieff 2015) and the effective contrast transfer function (CTF) was determined from frame-summed micrographs using CTFFIND4 (Rohou and Grigorieff 2015). Templates for auto-picking were generated from 2D classes generated from 5,000 manually picked particles in RELION. After manual inspection to remove false positives, 572,800 automated picked particles were extracted with a box size of 300 pixels and subjected to drift correction using alignparts\_lmbfgs (Rubinstein and Brubaker 2015). The resulting particles were 2D-classified into 150 classes after which 383,002 particles were selected. An *ab initio* 3D model with C2 symmetry, generated from CryoSPARC, was low-pass filtered to 60 Å and used as an initial model for 3D Classification in RELION. The most populated class was further refined in RELION to 4 Å resolution. A smoothed mask, excluding the detergent micelle, was created and used for 3D classification without alignment and subsequent local refinement in RELION. The final 3D reconstruction with C2 symmetry yielded a 3.9 Å map.

The movie frames were also motion-corrected using MotionCor2 software (Zheng et al. 2017), and CTF estimation was calculated using gCTF (Zhang 2016) both implemented in Relion 3 (Zivanov et al. 2018). Subsequently, the final set of selected particles were re-extracted and subjected to three iterations of Bayesian particle polishing, CTF refinement, and local refinement with C2 symmetry in Relion 3 (Zivanov, Nakane, and Scheres 2019). One round of refinement in C1 was performed to release the symmetry from C2 to C1. Subsequent 3D classification without image alignment was performed. The 3D classes were manually inspected for differences in local asymmetry. The particles belonging to a 3D class that appeared to have an opposite orientation were rotated 180° along the symmetry axis. One round of local refinement with C1 symmetry yielded 3.6 Å resolution in Relion. The final round of refinement was performed using CryoSPARC producing a map of 3.37 Å resolution (Punjani et al. 2017). The resolution was estimated using Fourier Shell Correlation (FSC) with 0.143 cutoff implemented in the 3DFSC web application (Zi Tan et al. 2017). Local resolution estimation from two CryoSPARC half maps was performed in CryoSPARC.

#### 2.4.6 Model building and refinement

The crystal structure of PCAT1 (PDB:4RY2) was placed into a sharpened cryo-EM map (sharpening factor,  $-80 \text{ \AA}^2$ ) using rigid body fitting in Chimera (Pettersen et al. 2004) followed by manual adjustments in Coot (Emsley and Cowtan 2004). The final model consists of residues 9–722 of PCAT1, residues 8–29 of the translocating CtA, and residues 8–25 of the non-translocating CtA.

The model was initially refined against one working half map in real-space by PHENIX (Adams et al. 2010), followed by rounds of refinement in reciprocal space using REFMAC (Brown et al. 2015), with secondary structure and reference restraints derived from ProSMART (Nicholls et al. 2014). The quality of the final model was evaluated by MolProbity (V. B. Chen et al. 2010). To assess the degree of overfitting, we calculated the FSC curves between the model and working half map, the non-working half map, and the full map using SPIDER (Shaikh et al. 2008; Frank et al. 1996).

#### 2.4.7 Site-directed mutation analysis of substrate binding

To assess residues on CtA for PEP binding, PCAT1 (C21A)-bound GST sepharose resin was used to pull down *wt* or mutant CtA. CtA titration was performed by mixing 20  $\mu\text{l}$  of Glutathione resin with 0.2, 0.4, 0.8, 1.6, or 3.2 nmol of CtA in 100  $\mu\text{l}$  reaction buffer. The reaction mixture was incubated for 30 min on ice, after which the resin was washed twice with 400  $\mu\text{l}$  wash buffer (50 mM Tris pH 7.0, 500 mM NaCl, 10% glycerol, 2 mM UDM, 5 mM DTT). The samples were analyzed by SDS-PAGE as well as Western with M2 anti-flag mouse antibody as primary antibody (1:5000) and goat anti-mouse Alexa Fluor 680 antibody as a secondary antibody (1:1000).

To assess the reciprocal binding residues on PCAT1, site-directed mutagenesis was performed to introduce mutation to PCAT1 (C21A), after which the pulldown experiment was performed using the C-terminal 3xFlag tagged CtA as a bait. For each reaction, 0.6  $\mu\text{M}$  of PCAT1 was incubated on ice for 10 min with 1.2, 2.4, 4.8, 9.6, or 19.2  $\mu\text{M}$  of CtA in 100  $\mu\text{l}$  reaction buffer (50 mM HEPES pH 7.0, 150 mM NaCl, 2 mM UDM, 5 mM DTT). To capture the PCAT1 (C21A)-CtA complex, 15  $\mu\text{l}$  of Anti-Flag M2 agarose affinity gel (Sigma-Aldrich) was added to the reaction and incubated for 30 min. The agarose affinity gel was washed twice with 400  $\mu\text{l}$  reaction buffer. Samples analyzed by SDS-PAGE.

#### 2.4.8 Disulfide cross-linking analysis of substrate translocation

To perform disulfide cross-linking accessibility assay, we constructed a cysteine-free PCAT1 by replacing all nine cysteines in PCAT1 with serine. Single cysteine substitutions at positions 275, 417, or 433 were introduced to the cysteine-free PCAT1 construct by site-directed mutagenesis. Mutant CtA (0.6  $\mu\text{M}$ ) containing an introduced cysteine residue and a HA tag was mixed with equimolar single-cysteine PCAT1 in buffer containing 50 mM HEPES pH 7.0, 150 mM NaCl, and 2 mM UDM. The reaction mixture was incubated at room temperature for 10 min

in the presence of 15  $\mu$ M Cu-Phenanthroline or 50 mM DTT before being analyzed by SDS-PAGE. For Western blotting, the cross-linked product was visualized with a primary anti-HA mouse monoclonal antibody (1:10000) and a secondary goat anti-mouse Alexa Fluor 680 antibody (1:1000). In addition, we have verified the identity of the cross-linked products by mass spectrometry for the following pairs: PCAT1 K275C-CtA K38C, PCAT1 K275C-CtA K44C, and PCAT1 K275C-CtA K65C. MS analysis was performed by The Rockefeller University Proteomics Resource Center.

## **CHAPTER 3: Structure of PCAT1 under the active turnover and Mg<sup>2+</sup> free condition reveals the rate-limiting steps and the role of ATP in PCAT1 transport cycle**

### **3.1 Summary**

The work described in the previous chapter proposed a hypothesis regarding the transport cycle of PCAT1. The hypothesis details a plausible mechanism for strict coupling between proteolytic activity of PEP and ATP hydrolysis. However, despite the wealth of biochemical evidence indicating the dual-functional coupling unique to PCATs, the structural mechanism remains poorly understood. To obtain direct structural information, I examine the conformational states of PCAT1 in the full functional transport cycle. I use cryo-EM single particle reconstruction techniques to reconstruct 3 structures of PCAT1 under an active turnover condition where functional PCAT1 are allowed to go through the transport cycle. I also reconstructed a structure of PCAT1 trapped in the outward-facing conformation using Mg<sup>2+</sup> free condition. The rigid-body movement of NBD is consistently observed across all structures indicating the intrinsic conformational heterogeneity of PCAT1. The rigid body movement of PEP relative to the TMD-NBD, is observed in the outward-facing state. The PEP movement is consistent with the strict coupling phenomenon and mechanistically explains how inter-domain motions coordinate the hydrolysis and proteolytic cleavage. In the turnover condition, only a small fraction of PCAT1 adopts the NBD-dimerized conformation, indicating that ATP hydrolysis and the subsequent NBD dissociation occur faster than NBD dimerization.

The data from the Mg<sup>2+</sup> free and ATP turnover condition support two mechanistic conclusions:

- (1) the ATP-bound NBD-dimerized conformation is the lowest energy state.
- (2) The rate-limiting step in PCAT1 is the formation of the NBD dimer.

Comparing these results to other studies suggest that whereas the thermodynamic property is universal among ABC transporters, the kinetic bottleneck varies from transporter to transporter. To my best knowledge this is the first example of using cryo-EM data to understand thermodynamic and kinetic properties of an active transporter.

### **3.2 Results**

#### **3.2.1 Structure determination of *wt*PCAT1 in the ATP-bound, pre-hydrolytic conformation**

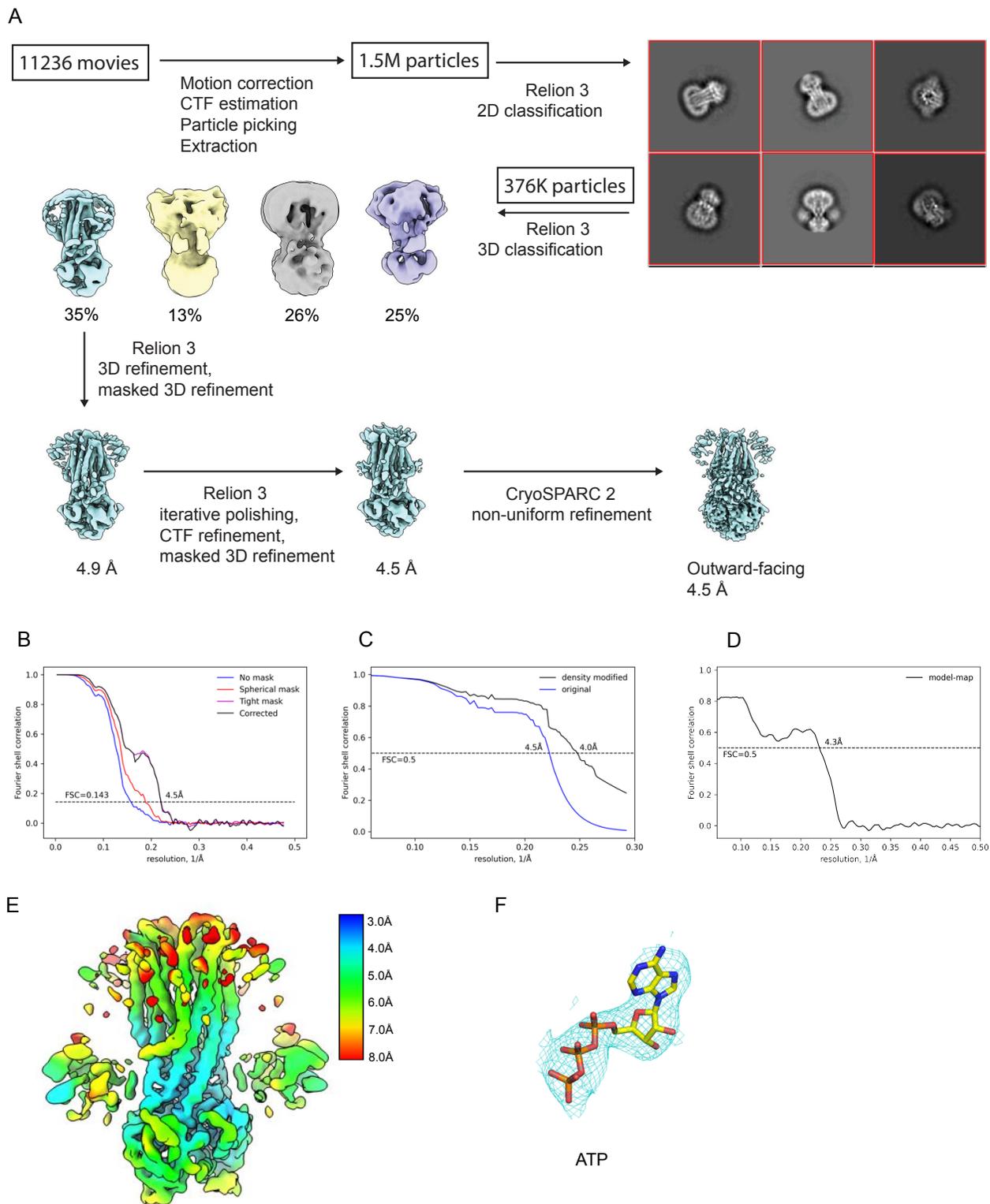
Previous biochemical studies established that addition of ATP in the absence of Mg<sup>2+</sup> inhibits substrate cleavage of PCAT1 (Lin, Huang, and Chen 2015 ). To study the structure of the ATP-bound, proteolytically inactive form, 6 mg/ml *wt*PCAT1 was vitrified in the presence of the substrate CtA (2x molar ratio) and ATP (10 mM). Mg<sup>2+</sup> ion was omitted to prevent ATP hydrolysis. Single particle reconstruction resulted one predominant conformation, with an overall resolution of 4.5 Å (Fig. 3.1). Density modification further improved the map, showing side-chain densities for the NBDs and the intracellular regions of the TMDs (Fig. 3.1E, Fig. 3.2). Prominent density at the ATPase sites is consistent with ATP without Mg<sup>2+</sup> ion (Fig. 3.1 F). The

TMDs show well-defined density in the intracellular region; the outer leaflet region exhibits clear secondary feature with little sidechain information. Densities corresponding to the two PEP domains are visible as two amorphous blobs attached to the transporter core. No density corresponds to the substrate was observed, possibly due to the limited resolution. The structure of ATP-bound PCAT1 was built based on the crystal structure of apo PCAT1(PDB: 4ry2). Polyalanine models were placed where sidechain densities were invisible. The final model consisting of 1418 residues and 2 ATP molecules was refined to good stereochemistry (Table 3.1).

In the presence of ATP without  $Mg^{2+}$ , *wt*PCAT1 adopt outward-facing (OF) conformations in which the NBDs form a closed dimer, and the translocation pathway opens to the extracellular space (Fig. 3.2). The intracellular, lateral opening of the translocation pathway observed in the IF conformations is completely closed (Lin, Huang, and Chen 2015). The lower resolution of the outer leaflet region of TMD indicating its flexibility (Fig 3.2 B), as observed in other ABC transporters such as the multidrug transporter P-glycoprotein (Kim and Chen 2018). The PEP domains are flexibly attached to the core transporter, with no defined interactions with TMDs nor NBDs (Fig. 3.2 B).

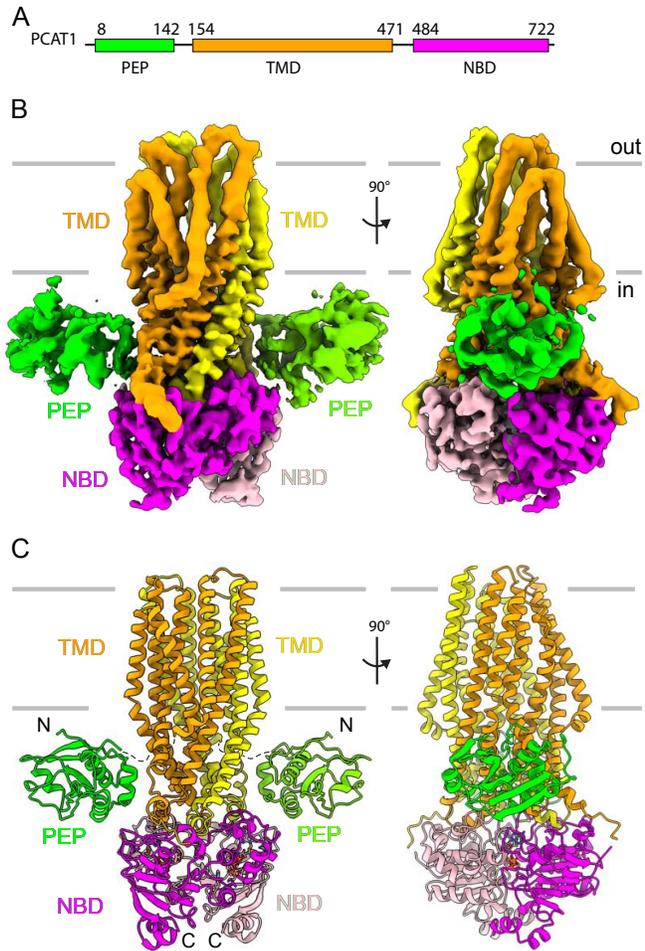
Structural comparison with the ATP $\gamma$ S-bound, E648Q mutant PCAT1 determined in the absence of substrate reveals two major differences (Fig. 3.3). First, in the crystal structure of the ATP hydrolysis deficient E648Q mutant (PDB: 4s0f), the PEP domains are entirely invisible (Lin, Huang, and Chen 2015). Second, the *wt*PCAT1 is in an outward-facing conformation, containing a TM cavity open to the extracellular space and closed off to the cytosol (Fig. 3.3 A). In contrast, the E648Q-PCAT1 is in an occluded conformation where the TM cavity is closed to both sides of the membrane (Fig. 3.3 B). It is possible that these differences are due to the energetic differences introduced by the mutation, the  $Mg^{2+}$ , the substrate, or the influence of crystal packing.





**Fig 3.1 Summary of cryo-EM image processing of PCAT1 in  $Mg^{2+}$  free condition**

- (A) Flowchart summarizing the cryo-EM image processing  
 (B) Fourier shell correlation (FSC) curve for the two half maps used in the final reconstruction  
 (C) Estimated resolution dependent map quality of the original and the density modified map  
 (D) Model validation of the PCAT1 in the ATP-bound outward-facing conformation  
 (E) Local resolution estimation of the PCAT1 in the ATP-bound outward-facing conformation  
 (F) Cryo-EM density for the nucleotide in the OF conformation



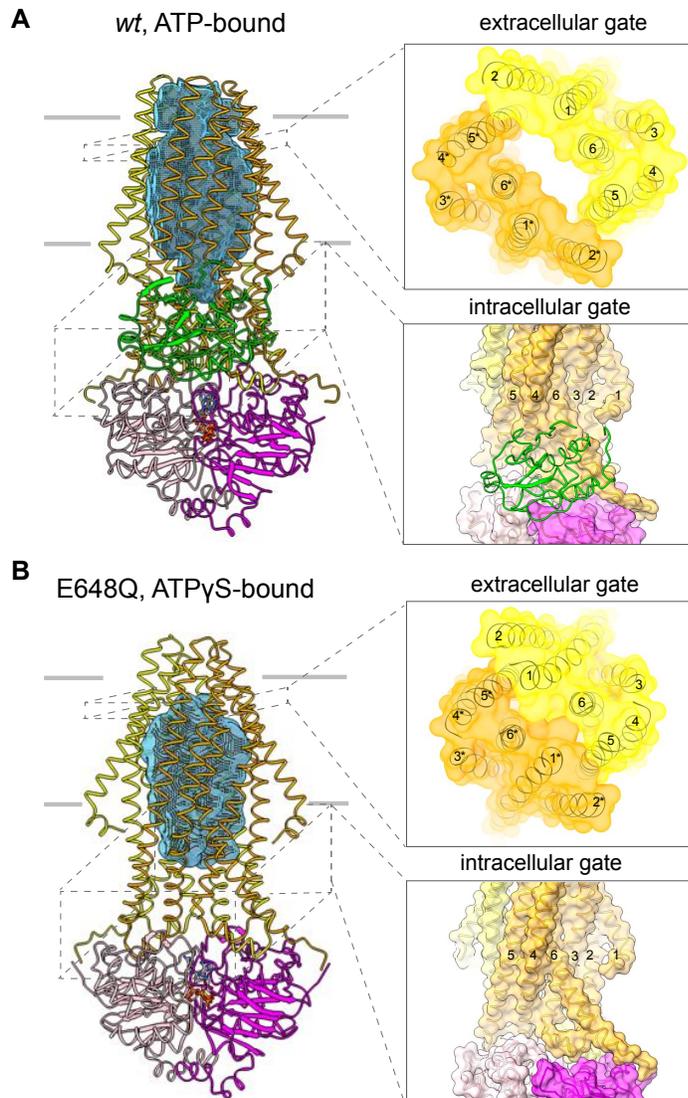
**Fig. 3.2 The cryo-EM structure of PCAT1 in the OF conformation**

(A) Cartoon illustration depicting domain organization of PCAT1.

(B) Two orthogonal views of cryo-EM density map of PCAT1 in OF conformation.

The density is color-coded by domains. Magenta, NBD; orange, TMD; green, PEP.

(C) The ribbon representation of PCAT1 in OF conformation color-coded by domains.



**Fig. 3.3 Comparison of conformational differences between PCAT1 in OF and occluded conformations**

Orthogonal views of PCAT1 shown as ribbon cartoons. The volume inside the TM cavity generated using a probe of 3 Å radius is shown as a blue mesh. The panel shows PCAT1 from the front view. The zoomed-in panel shows the extracellular gate viewed from extracellular space and the intracellular gate. TM helices forming the gates are labeled.

(A) *wt*PCAT1 in ATP-bound OF conformation

(B) E648Q PCAT1 in ATP $\gamma$ S-bound occluded conformation

### 3.2.2 Structure determination of PCAT1 under active turnover condition

Next, we included  $Mg^{2+}$  to investigate the conformational states of PCAT1 under active turnover condition. The cryo-EM sample was prepared by vitrifying the *wt* PCAT1 with 2-fold molar excess of the substrate CtA, 10 mM ATP- $Mg^{2+}$ , and the creatine phosphate ATP-regenerating system. In marked contrast to the structure determined in the absence of  $Mg^{2+}$ , single particle reconstruction delineates four distinct conformations (Fig. 3.4 and Table 3.2). Among the 671,000 particles analyzed by 3D classification, about 83% exhibit IF conformations with different degrees of NBD separation and only 14% are in an NBD-dimerized configuration. Subsequent refinement resulted in three IF structures, referred to as IF wide (IF<sub>W</sub>), intermediate (IF<sub>I</sub>) and narrow (IF<sub>N</sub>), respectively (Fig. 3.4, 3.5). The overall resolutions of IF<sub>W</sub>, IF<sub>I</sub> and IF<sub>N</sub> structures are 4.1, 3.7, and 3.7 Å, respectively (Fig. 3.5 A). A fourth structure, showing a full NBD-dimerized configuration, could only be refined to 6.6 Å suggesting that particles within this class are heterogenous (Fig. 3.4).

Different from the OF reconstruction, all three IF structures show strong densities in the TMDs, enabling unambiguous residue assignment (Fig. 3.5, 3.6). The densities for NBD and PEP are weaker, suggesting that these domains are relatively dynamic. Extra densities are observed in the NBDs near the Walker A motif in all three IF reconstructions (Fig. 3.6 C). The shape and size of the density are consistent with ATP and a  $Mg^{2+}$  ion, indicating that ATP are bound to the IF conformations.

The three IF structures share a similar overall conformation, with different degrees of separation of the intracellular regions (Fig. 3.6 B). Similar to the structure of PCAT1-CtA complex that was obtained with a proteolytic-deficient C21A mutant (PDB: 6V9Z) (Kieuvongngam et al. 2020), the two PEP domains dock onto the intracellular openings of the TM cavity, each bound with a leader peptide. No density was observed beyond the double-glycine cleavage site (G23-24), suggesting that cleavage has already taken place (Fig. 3.7 A-B). Cleavage experiments using the same cryo-EM sample indicates that a fraction of the substrate has been processed within the time window of sample vitrification (Fig. 3.7 C).

Inside the TM cavity, we observe amorphous densities which may correspond to the cargo peptide (Fig. 3.6 D). Recently an electron paramagnetic resonance study argues for direct interaction between the cargo domain and the core transporter (Rahman and Mchaourab 2020). In contrast, we do not observe any EM density that could be interpreted as specific interactions between the substrate cargo and residues lining translocation pathway in all three IF structures. The cryo-EM structures support the previous conclusion that substrate is recruited by the PEP domain and the TM cavity serves as a “Teflon-like” conduit for the cargo (Kieuvongngam et al. 2020).

The discrepancy between the electron paramagnetic resonance study (Rahman and Mchaourab 2020) and the present cryo-EM results is worth considering. First the study uses a bimane-labeled substrate at the native cysteine 31. Upon incubation with detergent solubilized PCAT1, a blue shift in fluorescence emission is observed consistent with changes of environment to a more hydrophobic environment, compared to the detergent micelles negative control. In addition, a methanethiosulfonate spin label substrate is used to probe the interaction. A broadening of EPR

line is observed consistent with a restricted motion of the spin label. The authors argue that this evidence indicate that the substrate is translocated into the hydrophobic transmembrane cavity and thus the mobility of spin label on the substrate is restricted. We argue that this observed interaction could be protein-protein interactions but can still be non-specific i.e. non-specific promiscuous binding of substrates to a membrane protein. A better control should guard against this possibility would be to use other small-molecule ABC transporters as a negative control. Furthermore, an in-vitro pulldown experiment using the Flag-tagged substrate to pull down the core transporter showed that the substrate cannot pull down the core transporter (Lin, Huang, and Chen 2015). Therefore, if there is any direct interaction, the affinity is too low to be detected in a bulk pulldown assay.

In summary, under turnover condition, the majority of PCAT1 adopt IF conformations. One of the substrate has likely been cleaved and released to the TM cavity. Two ATPs are bound to the separated NBDs. The presence of ATP-bound IF states suggests that ATP binding, per se, does not induce conformational changes required for NBD dimerization. Rather, the role of ATP is to stabilize the NBD dimers.

**Table 3.2 Summary of cryo-EM data and structure refinement statistics for *wf*PCAT1 under active turnover condition**

<b>Data collection</b>			
Microscope	Titan Krios (FEI)		
Voltage (kV)	300		
Detector	K2 (Gatan)		
Pixel size (Å)	1.03		
Defocus range (μM)	0.8-2.2		
Movies	11966		
Frames/movie	50		
Dose rate (electron/pixel/s)	8		
Total dose (electrons/Å <sup>2</sup> )	75		
<b>Conformation</b>	<b>IF<sub>W</sub></b>	<b>IF<sub>I</sub></b>	<b>IF<sub>N</sub></b>
<b>Refinement</b>			
Number of particles	134772	88549	80156
Resolution (Å)	4.1	3.7	3.7
<b>Model composition</b>			
Non-hydrogen atoms	11460	11460	11444
Protein residues	1458	1458	1458
ATP	2	2	2
Mg <sup>2+</sup>	2	2	2
<b>Validation</b>			
<b>RMS deviation</b>			
Bond lengths (Å)	0.0028	0.0025	0.0027
Bond angles (°)	1.29	1.22	1.26
<b>Ramachandran</b>			
Favored (%)	94.95	94.89	96.05
Allowed (%)	5.05	5.11	3.95
<b>Rotamer</b>			
Favored (%)	87.3	96.63	95.34
Allowed (%)	9.73	3.13	4.10

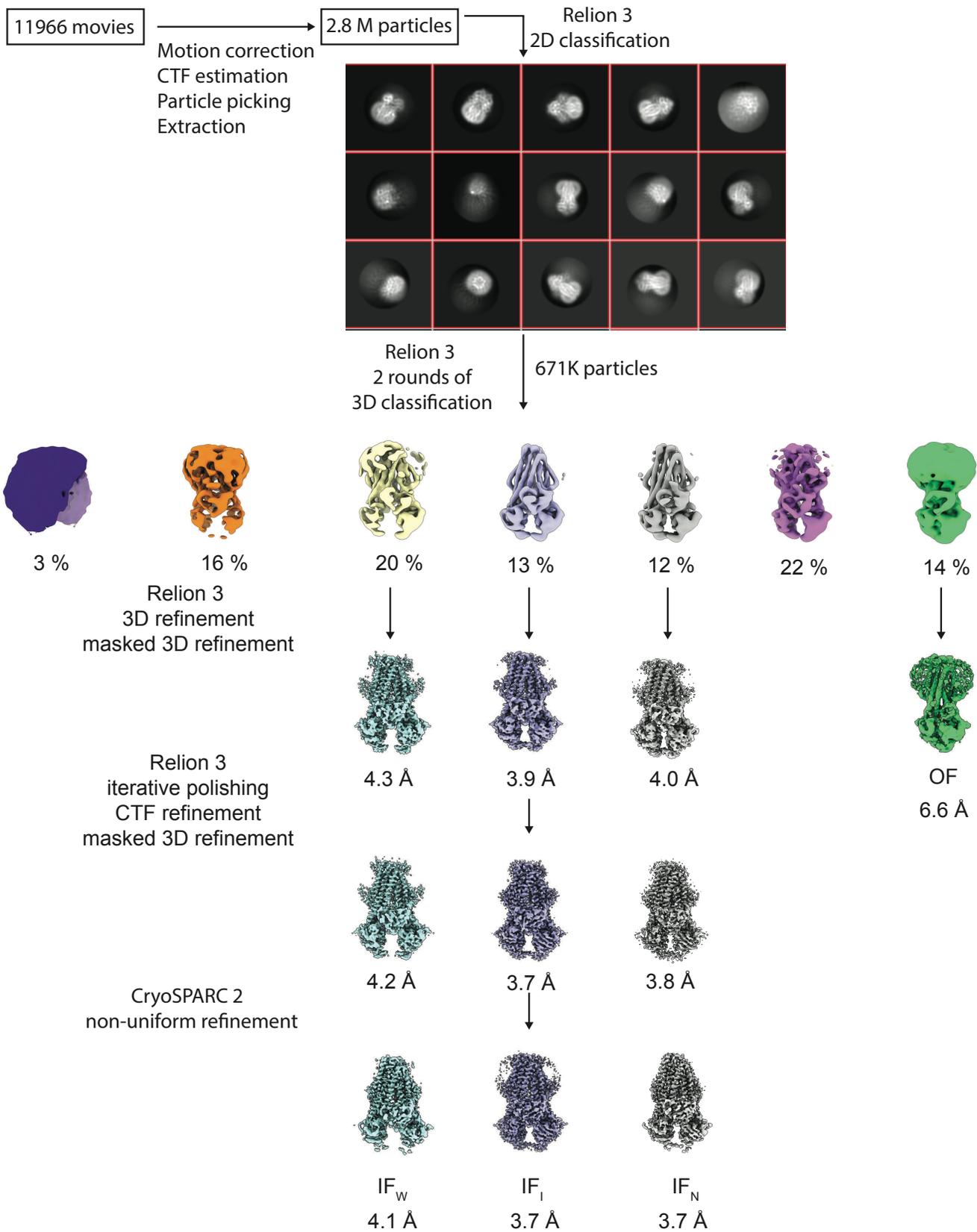
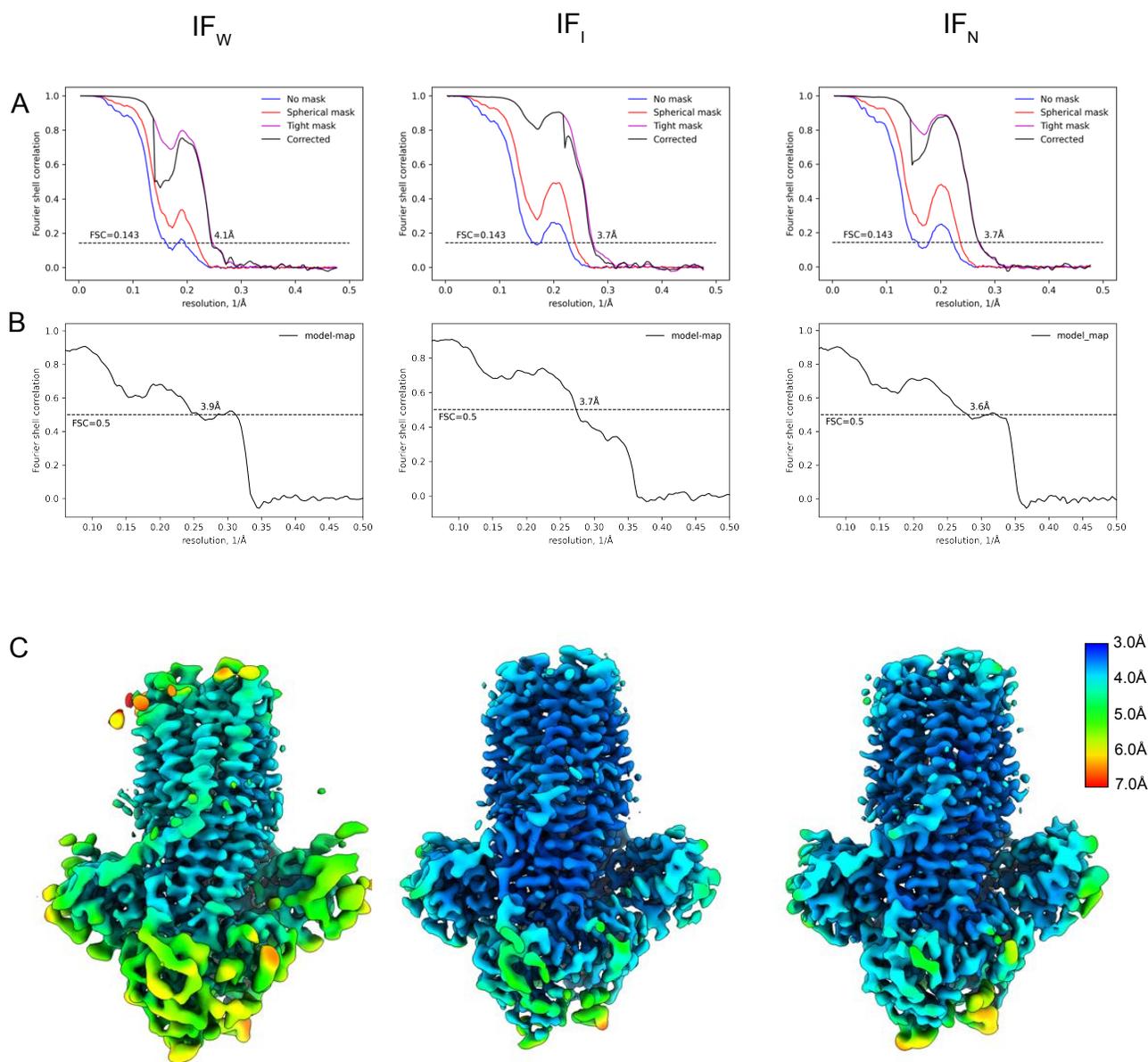
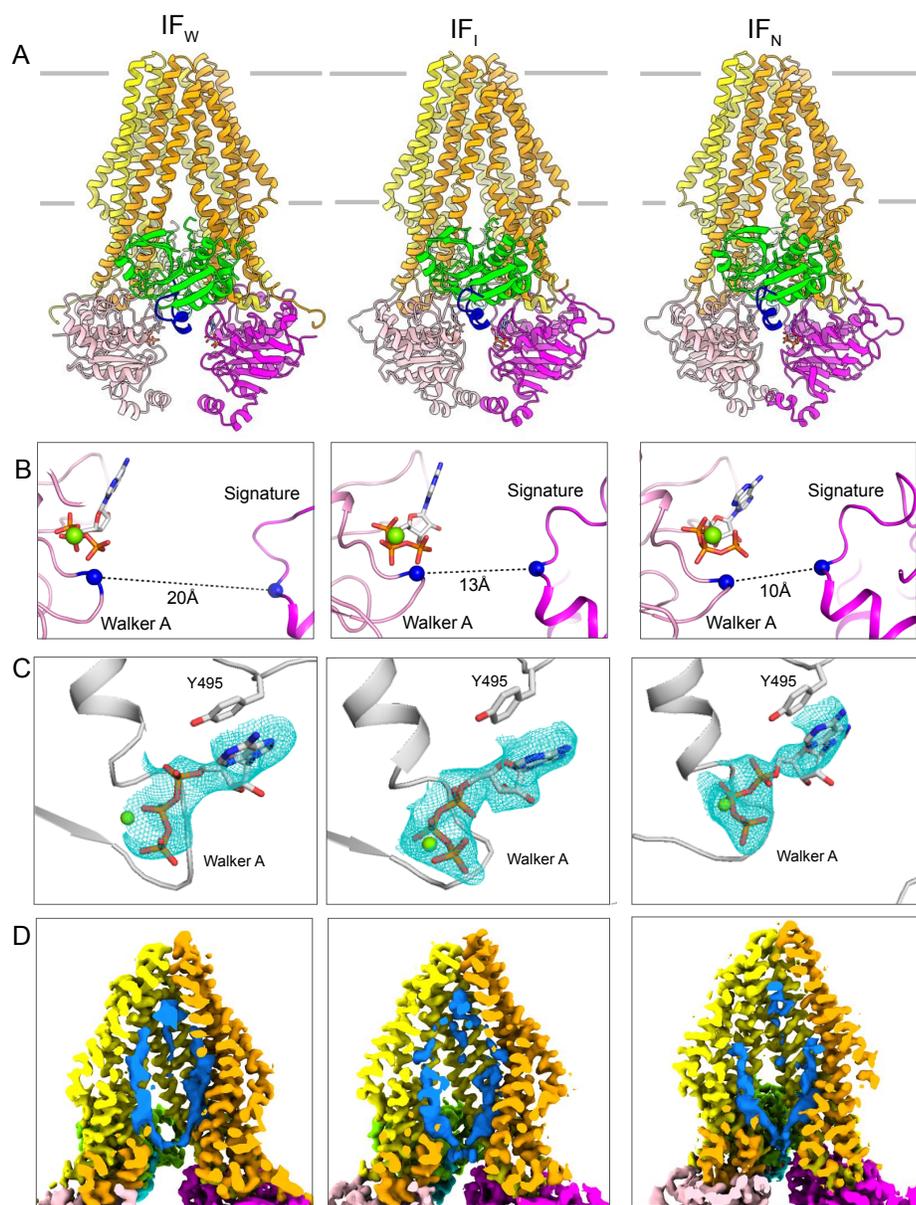


Fig 3.4 Summary of cryo-EM image processing of PCAT1 under active turnover condition



**Fig 3.5 Quality of the cryo-EM density map of *w*PCAT1 in the IF conformations**

- (A) Fourier shell correlation (FSC) curve for the two half maps used in the final reconstruction
- (B) Model validation of the PCAT1 in the inward-facing conformations under active turnover condition
- (C) Local resolution estimation of the PCAT1 in the ATP-bound outward-facing conformation



### Fig 3.6 The cryo-EM structure of PCAT1 IF conformations

The left panels: IF<sub>W</sub>; The middle panels: IF<sub>I</sub> ; The right panels: IF<sub>N</sub>.

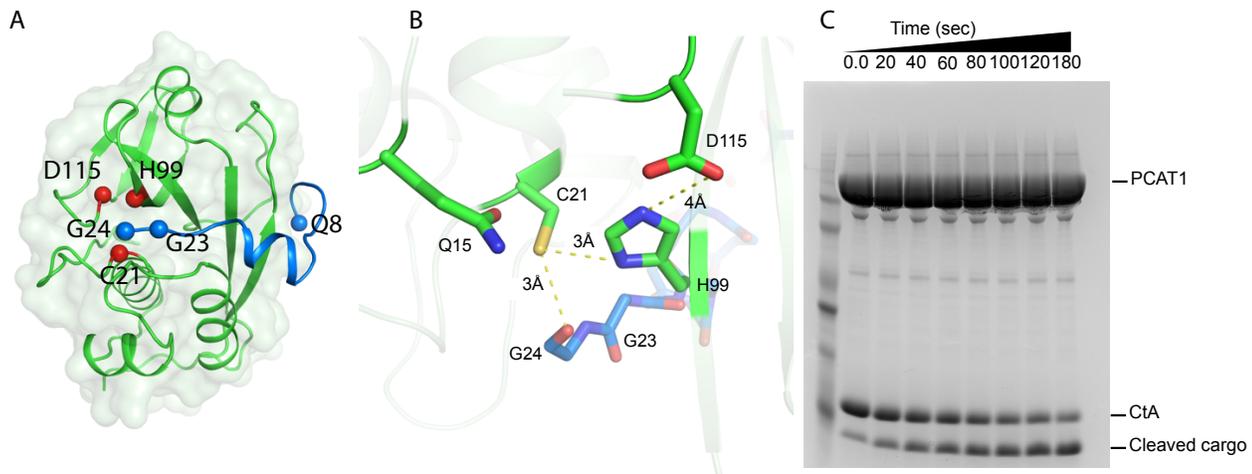
(A) The frontal views of cryo-EM structures of PCAT1 in IF conformations.

The density is color-coded by domains. Blue, CtA; magenta, NBD; orange, TMD; green, PEP.

(B) The degrees of NBDs separation in the IF conformations. The C $\alpha$  distances between G522 in the Walker A motif and S624 in the ABC signature motif of the other NBD are indicated.

(C) The presence of ATP-Mg in the IF conformations. The cryo-EM density of ATP-Mg is shown as a blue mesh. The conserved aromatic residue Y495 is shown forming  $\pi$ - $\pi$  stacking interaction with the adenine ring stabilizing the ATP.

(D) The cargo densities inside the transmembrane cavity. Cross sectional views through the cryo-EM densities maps of PCAT1. The cargo density is shown in blue.



**Fig. 3.7 Substrate cleavage under active turnover condition**

(A) Docking of the leader peptide (blue) on the PEP surface (green).

The PEP catalytic triad residues are indicated as red spheres and the CtA double glycine motif are indicated as blue spheres.

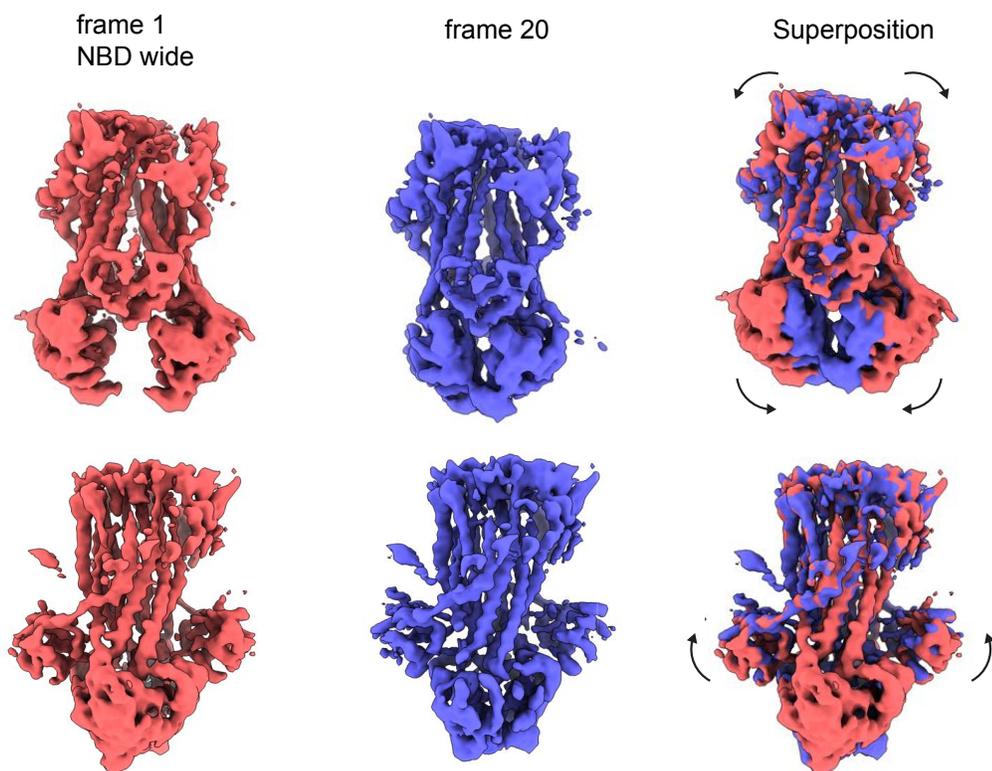
(B) A zoomed-in view of the catalytic site.

(C) Substrate cleavage assay recapitulating the cryo-EM sample preparation

of PCAT1 under active turnover condition: 10 mM ATP-Mg, 2 molar ratio substrate, and creatine phosphokinase ATP regenerating system.

### 3.2.3 Conformational dynamic in the active turnover condition

Next, we used the 3D variability analysis (Punjani and Fleet 2021) to visualize the conformational dynamics of PCAT1 in the active transport cycle. The result illustrated that the conformational transition is a continuous movement of NBD, coupled with rotational movements of the TMDs and dissociation of the PEP domains (Fig. 3.8). Using one NBD subunit as a frame of reference, dimerization of NBD is accompanied by concerted inward movement of the cytosolic/inner leaflet regions of the TMDs and outward motion of the outer leaflet region. The movements of the TMDs leads to conformation changes of the PEP-docking site, resulting in dissociation of the PEP domains. The 3D variability analysis indicates that the association of the NBD dimer is directly coupled with the formation of the OF conformation and the dissociation of the PEP domains. We suggest that these structures represent conformations *en route* to the NBD-dimerized, OF state.



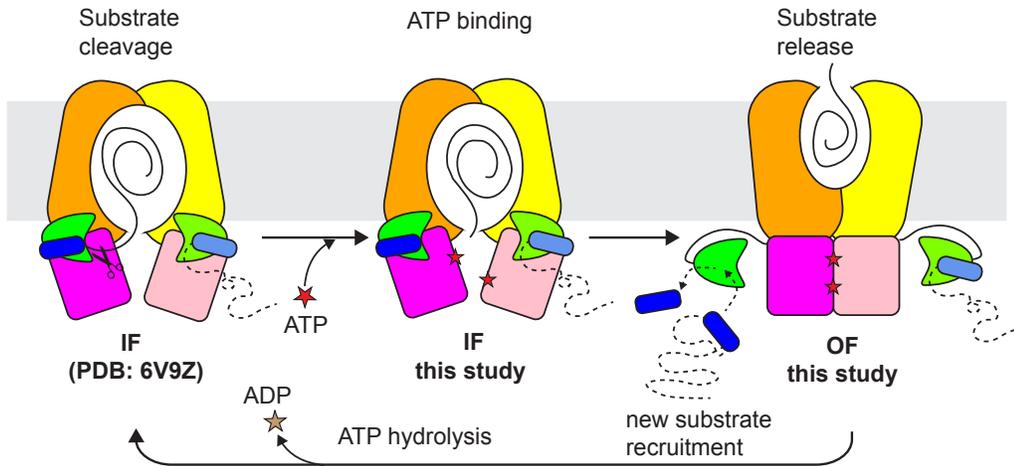
**Fig. 3.8 Conformational variability in the active turnover dataset**

The variability component resolves a motion along the NBD separation-dimerization axis which correspond to the inward-outward motion in the transport cycle. The first and last frame of 3D volume series from the 3D variability analysis is displayed in magenta and blue respectively. Two perpendicular orientations are shown. The front views and the side views are shown with arrows to indicate the direction of domain motion along the peptide transport cycle.

### 3.3 Discussion

#### 3.3.1 Functional coupling between substrate cleavage and substrate transport is achieved by domain motion coupling

In this study, we determined the structures of *wt*PCAT1 in the presence of substrate and ATP, with and without  $Mg^{2+}$  ion. In the absence of  $Mg^{2+}$ , PCAT1 predominately exhibits an OF conformation, in which the PEP domains are detached from the translocation pathway, the NBDs are dimerized with two ATP poised for hydrolysis, and the cargo substrate has been released. Inclusion of  $Mg^{2+}$  to enable ATP hydrolysis resulted in completely different structures. The majority of PCAT1 adopts IF conformation in which the NBDs are separated, the substrate is enclosed inside the TM cavity and the PEP domains remain associated with the transporter core. These results, together with previous biochemical and structural information (Lin, Huang, and Chen 2015; Kieuvongngam et al. 2020), support a mechanism to explain how substrate processing, translocation, and ATP hydrolysis are coupled in the PCAT1 transport cycle (Fig. 3.9). Substrate is recruited in the IF states and proteolytic cleavage takes place near the intracellular opening. After proteolytic cleavage, the cargo is released into the TM cavity, whereas the leader peptide remains attached to the PEP domain. ATP binding stabilizes the NBD-dimerized, OF conformation, to release the cargo to the extracellular space. In this conformation, the PEP domains are disengaged from the transporter core to recruit new substrates. Dissociation of the PEP domains also diminishes proteolytic activity of PCAT1, thereby prevents cleavage of substrates before positioning them inside the translocation pathway. Finally, ATP hydrolysis leads to NBD dissociation, thus resetting the transporter to the IF state.



**Fig. 3.9 Peptide transport cycle model of PCAT1**

PCAT1 is dynamically transitioned between multiple IF conformations. Substrate can bind and is cleaved in the IF conformations. The rate limiting step is NBD dimerization. After NBD dimerization, the substrate is released extracellularly. ATP hydrolysis occurs bringing PCAT1 to IF conformations.

### 3.3.2 Behaviors of PCAT1 in active turnover condition compared to other ATP transporters

The marked difference in the conformational distributions observed with and without  $Mg^{2+}$  reveals energetic and kinetic properties of PCAT1. To explain the cryo-EM observations, we show a simplified transport cycle with two-states (Fig. 3.10). This simplification subsumes multiple NBD-separated IF conformations into a single state (Fig. 3.10). PCAT1 catalyzes the hydrolysis of ATP with a Michaelis constant ( $K_m$ ) of 0.23 mM (Lin, Huang, and Chen 2015). Under our experimental conditions with 10 mM ATP, we assume that all IF conformations contain ATP. The transition from NBD-separated to NBD-dimerized states involves an isomerization step ( $k_1$ ). The NBD-dimerized OF PCAT1 has two routes to return to the NBD-separated state: through isomerization ( $k_{-1}$ ) or through ATP hydrolysis followed by NBD-dissociation and nucleotide exchange. A composite rate constant  $k_2$  is used to describe the latter route.

Assuming the cryo-EM samples were vitrified after the system reaches steady state, then the probability of observing the NBD-separated IF state ( $P_{IF}$ ) versus that of the NBD-dimerized OF state ( $P_{OF}$ ) is determined by equation 1:

$$\frac{P_{IF}}{P_{OF}} = \frac{k_2 + k_{-1}}{k_1} \quad (1)$$

In the first experiment,  $Mg^{2+}$  is omitted to prevent ATP hydrolysis, i.e.  $k_2=0$  (Williams 2000). Under this condition, the IF and OF states are in equilibrium and the Boltzmann distribution law states that the conformational distribution depends on the energy difference between the two states ( $E_{OF}-E_{IF}$ ):

$$\frac{P_{IF}}{P_{OF}} = \frac{k_{-1}}{k_1} = e^{(E_{OF}-E_{IF})/k_B T} \quad (2)$$

where  $k_B$  is the Boltzmann constant and  $T$  is the absolute temperature of the system.

Cryo-EM analysis only resolved the NBD-dimerized OF conformation, indicating that  $P_{OF} \gg P_{IF}$ . Using equation 2, we conclude that the energy of the OF conformation is lower than that of the IF conformation. The fact that  $P_{OF} \gg P_{IF}$  also means  $k_1 \gg k_{-1}$ , i.e, in the presence of ATP, the NBD-dimerized OF conformation is stable, the reverse isomerization ( $k_{-1}$ ) is much slower than that of NBD-dimerization.

In the second experiment, addition of  $Mg^{2+}$  enables energy dissipation from ATP hydrolysis, shifting the system out of equilibrium. Under this condition, the Boltzmann distribution no longer applies. As we learned from the first experiment, that  $k_1 \gg k_{-1}$ , equation 1 can be approximated as

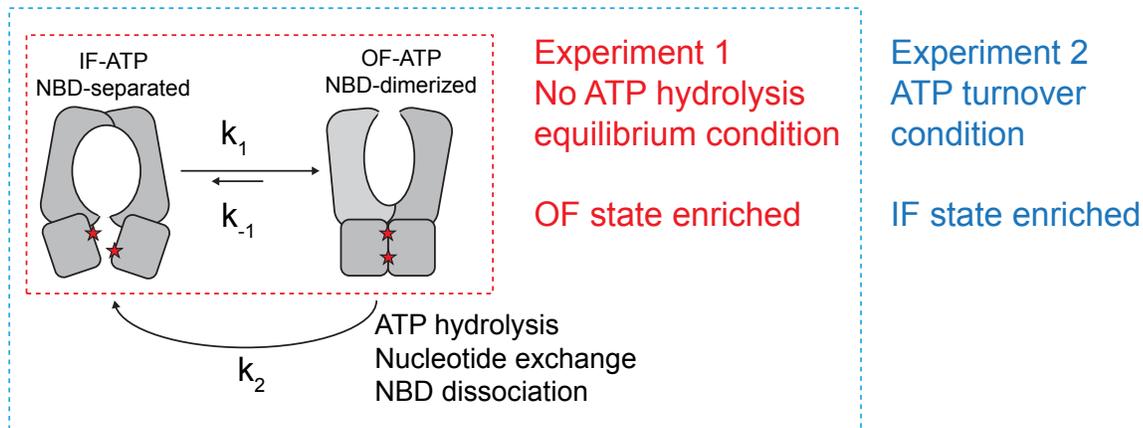
$$\frac{P_{IF}}{P_{OF}} = \frac{k_2 + k_{-1}}{k_1} = \frac{k_2}{k_1} + \frac{k_{-1}}{k_1} \cong \frac{k_2}{k_1} \quad (3)$$

The 3D classification analysis in the presence of  $Mg^{2+}$  shows that approximately 83% of the particles exhibit the NBD-separated IF conformations and about 14% are in the NBD-dimerized

OF conformations (Fig. 3.4). If we were to make inference from this result, then the ratio of  $k_2/k_1$  is about 6. In other words, resetting the NBD-dimerized OF conformation to the NBD-dissociation IF conformation ( $k_2$ ) takes place six times faster than formation of the NBD-dimerized OF conformation ( $k_1$ ).

Recent advances in cryo-EM enabled us to study the conformational landscape of dynamic molecules such as ABC transporters. It is important to keep in mind that which conformation dominates under ATP turnover conditions is a kinetic property specified by the relative rates of transitions in the transport cycle. In contrast, in equilibrium conditions where ATP hydrolysis is abolished, the relative abundance of conformations is a thermodynamic property of the system determined by the relative energies of states. Our analyses of PCAT1 give rise to two conclusions: 1. In the presence of ATP, the NBD-dimerized conformation is the lowest energy state. 2. In the PCAT1 transport cycle, the rate limiting step is NBD dimerization.

So far, the structures of many ABC transporters determined under equilibrium conditions using non-hydrolysable ATP analogs or catalytic-deficient mutants resulted in an NBD-dimerized conformation (Dawson and Locher 2006; Ward et al. 2007; Choudhury et al. 2014; Kim and Chen 2018; Johnson and Chen 2018; Manolaridis et al. 2018; Perez et al. 2019; Hutter et al. 2019; Fan, Kaiser, and Rees 2020). Thus, it is likely a general property of ABC transporters that in the presence of ATP, the NBD-dimerized conformation is the lowest energy state. In contrast, under ATP turnover conditions, different conformational distributions were observed for different ABC transporters. The multidrug transporter ABCG2 exhibits only IF, NBD-separated conformations (Q. Yu et al. 2021), indicating that similar to PCAT1, NBD dimerization is rate limiting. The opposite result was obtained with MRP1, which adopts the NBD-dimerized, OF conformation in the presence of ATP and  $Mg^{2+}$  (Wang et al. 2020). On the other hand, the *Thermus thermophilus* multidrug-resistance proteins A and B (TmrAB) exhibits both NBD-separated and NBD-dimerized conformations (Hofmann et al. 2019). These different results suggest that the kinetic bottleneck of the transport cycle varies from transporter to transporter.



**Fig. 3.10 Kinetic model diagram summarizing the cryo-EM experiments**

Under physiological condition, ATP rapidly bind to the IF state. The process of IF-OF transition is governed by the forward isomerization rate of NBD dimerization ( $k_1$ ) and the backward isomerization rate of NBD dissociation ( $k_{-1}$ ). The rate of ATP hydrolysis, nucleotide exchange, and NBD dissociation back to IF state is represented as  $k_2$ .

The cryo-EM experiment 1 under  $Mg^{2+}$  free condition investigated the NBD dimerization and dissociation under equilibrium condition. It is found that OF state is enriched under this condition, revealing that OF state is the lower energy state.

The cryo-EM experiment 2 under active ATP turnover condition investigated the rate limiting step of the overall transport cycle. It is found that IF state is enriched under this condition, revealing that  $k_2$  is faster than  $k_1$  and NBD dimerization step is the rate limiting step of the transport cycle.

### **3.4 Materials and Methods**

#### **3.4.1 Protein expression and purification**

The wild-type and E648Q PCAT1 in pMCSG20 vector with an N-terminal glutathione-S-transferase (GST) tag and a Tobacco Etch Virus (TEV) protease cleavage site were expressed and purified using a protocol previously described in chapter 2. Briefly, the construct was expressed in *E. coli* strain BL21(DE3) codon plus (RIL) cells expressing PCAT1. The cells were lysed and solubilized in buffer containing 1% n-dodecyl- $\beta$ -D-maltoside (DDM; Anatrace), 50 mM Tris pH 7.0, 500 mM NaCl, 10% glycerol, and 5 mM DTT. The supernatant was applied to the Glutathione Sepharose 4B affinity resin followed by buffer exchange to 50 mM Tris pH 7.0, 500 mM NaCl, 10% Glycerol, 5 mM DTT, and 2 mM n-undecyl- $\beta$ -D-maltopyranoside (UDM; Anatrace). The GST tag was removed by TEV protease and the protein was further purified by gel filtration chromatography using a Superdex 200 increase column (GE Healthcare) in a buffer containing 50 mM Tris pH 7.0, 150 mM NaCl, 2 mM UDM.

The wild-type PCAT1 substrate CtA in MCSG7 vector with an N-terminal TEV cleavable 6x His tag were expressed and purified using a protocol previously described in chapter 2. Briefly, the cell pellet containing inclusion bodies were solubilized in 8 M urea, 50 mM Tris pH 7.0, 150 mM NaCl, and 10% glycerol. The denatured protein was purified on cobalt affinity resin (Clontech Laboratories) and refolded via dialysis in lysis buffer plus 5 mM DTT. The His-tag was removed by cleavage with TEV protease and the protein were further purified by gel-filtration chromatography (Superdex 75 HiLoad 16/60, GE Healthcare).

#### **3.4.2 Cryo-EM sample preparation and data collection of the active turnover dataset**

Purified wild-type PCAT1 in the ATP regenerating buffer condition (0.1 mg/ml creatine phosphokinase, 10 mM creatine phosphate) was mixed with twofold molar excess of refolded CtA. Immediately 10 mM ATP and 10 mM Mg<sup>2+</sup> was added to the mixture solution. The solution was applied onto glow-discharged holey carbon grids (Quantifoil gold R1.2–1.3), incubated for 20 s at 100% humidity, and blotted with filter paper for 3 s before being plunge-frozen into liquid ethane using a Vitrobot Mark IV (FEI). A dataset of 11,966 movies was collected on the Titan Krios Transmission Electron Microscope (FEI) outfitted with a K2 Summit direct electron detector (Gatan) with a super-resolution pixel size of 0.515 Å. The electron dose rate was eight electrons/pixel/s with a total exposure time of 10 s resulting in a total electron dose of 80 electrons/Å<sup>2</sup> over 50 frames.

#### **3.4.3 Single particle reconstruction of the active turnover dataset**

Movie frames were corrected for gain reference and binned by two to yield a pixel size of 1.03 Å/pixel. Sub-frame alignment was carried out using MotionCorr2, and the contrast transfer function (CTF) was estimated using Gctf software. From 11,966 micrographs, 2.8 million particles were selected using RELION particle picking with 3D template. To remove false positive and broken particles, two rounds of batched 3D classification was carried out in

RELION using a cryo-EM map of PCAT1-substrate complex low-passed-filtered to 60Å as a reference model. Four distinct classes, IF<sub>W</sub>, IF<sub>I</sub>, IF<sub>N</sub>, and OF emerged. The particle numbers for these classes are 134 K, 88 K, 80K, and 94K particles respectively. The OF class is too poorly resolved and a separated dataset in the absence of Mg<sup>2+</sup> was collected instead.

For each inward-facing class, masked 3D classification without alignment was performed to ensure the homogeneity of each class. Each class was then subjected to one round of 3D refinement followed by masked 3D refinement. Iterative cycles of CTF refinement, Bayesian polishing, and masked 3D refinement in RELION were then performed to improve the map quality. The final round of refinement was performed without symmetry imposed in CryoSPARC-2 software using non-uniform refinement. The resolution of each cryo-EM map are 4.2 Å, 3.7 Å, and 3.7 Å for IF<sub>W</sub>, IF<sub>I</sub>, and IF<sub>N</sub> respectively.

For the 3D variability analysis, all particles belonging to each conformation were pooled together. One round of a consensus refinement and another round of non-uniform refinement were performed in CryoSPARC-2 on this set of particles. The consensus map was then used as starting volume for 3D variability analysis. The maps were filtered to 5 Å and three eigenvector components were solved. One eigenvector component displays variability corresponding to the IF-OF conformation change. The variability is visualized as movies in a series of 20 volumes.

#### **3.4.4 Cryo-EM sample preparation, data collection, and single particle reconstruction of the ATP-bound outward-facing conformation**

Procedure in the previous section was followed except 10 mM ATP in the absence of Mg<sup>2+</sup> was added immediately before the freezing. A dataset of 11,236 movies was collected on the Titan Krios Transmission Electron Microscope (FEI) outfitted with a K2 Summit direct electron detector (Gatan) with a super-resolution pixel size of 0.515 Å. The electron dose rate was eight electrons/pixel/s with a total exposure time of 10 s resulting in a total electron dose of 80 electrons/Å<sup>2</sup> over 50 frames.

The reconstruction strategy as in the previous section was followed. The 3D classification yielded a single class revealing outward-facing features. A round of 3D refinement followed by a masked 3D refinement was performed. Iterative cycles of CTF refinement, Bayesian polishing, and masked 3D refinement in RELION were then performed to improve the map quality. The final round of refinement was performed without symmetry imposed in CryoSPARC-2 software using non-uniform refinement yielding a map of 4.5 Å resolution.

#### **3.4.5 Improving the cryo-EM map qualities using density modification procedure**

The fundamental concept of the density modification is that information about the expected values of density in one part of a cryo-EM density map can be used to partially correct errors in the Fourier terms representing that map and the corrected Fourier terms lead to an improved map everywhere (Terwilliger, Sobolev, et al. 2020; Terwilliger, Ludtke, et al. 2020). The density modification procedure requires two assumptions. First, the errors in the Fourier space of the two

independent half maps are uncorrelated. Second, prior knowledge about the macromolecules is known in advance and can provide knowledge about the features of a cryo-EM map. For example, the molecular weights can give an estimate of a volume size. Consequently, the ratio of solvent region and macromolecule region can be estimated. It is also known in advance that the macromolecular region has distribution of density like those found among other macromolecules and the solvent region has nearly constant (flattened) density.

The importance of having two independent half maps with uncorrelated error in Fourier space can be reasoned as followed. A single error in Fourier term causes global error throughout the cryo-EM map in real space, including the solvent region. Thus, adjusting the Fourier coefficient to yield flattened solvent region will improve the map quality in real space globally including the map quality of the molecule region. The density modification procedure works by adjusting Fourier coefficients of a density modified map to agree with both the original independent half maps, and the expected features.

the maximum likelihood-based density-modification procedure is implemented in Phenix (Terwilliger, Ludtke, et al. 2020). Two independent unsharpened half maps, a polyalanine model without nucleotide ligands, and a generous mask are used as input. The output density-modified cryo-EM map was used to improve model building and figure preparation.

### **3.4.6 Model building and refinement**

The cryo-EM structure of PCAT1 substrate complex (PDB:6V9Z) was docked into sharpened inward-facing conformation cryo-EM maps. The docking was performed in Chimera using rigid body fitting, followed by manual adjustments in Coot. The final models in all of the inward-facing structures consist of residues 9-722 on PCAT1, residues 14-24 on both CtAs, 2 ATPs and 2 Mg<sup>2+</sup>.

The crystal structure of PCAT1 complexed with ATP $\gamma$ S in an occluded conformation (PDB: 4S0F) was rigid body fit into the outward-facing conformation cryo-EM map similar to the procedure previously described. The final model consists of 9-722 on PCAT1, as well as two molecules of ATP.

In all cases, the models were subjected to real-space refined in PHENIX followed by iterative cycles of refinement in Refmac and manual rebuilding in COOT. MolProbity was used to assess the quality of the final model. To assess the degree of overfitting, we calculated the FSC curves between the model and working half map, the non-working half map, and the full map using SPIDER.

### **3.4.6 Figure preparation**

Figures were prepared with PyMOL, UCSF Chimera, and UCSF ChimeraX.

## CHAPTER 4: Additional native mass spectrometric study of PCAT1

### 4.1 Introduction

PCATs are highly dynamic ABC peptide exporters that function both as a peptide substrate maturing processor and an ATP-driven peptide exporter. PCAT1 contains four enzymatic active sites: two peptidase sites and ATPase sites. Previous structural studies have yet to characterize whether one or two substrates are processed or whether one or two ATP molecules are hydrolyzed in a cycle, and whether ATP hydrolysis happens in a parallel or sequential manner. In addition, capturing short-lived intermediates with structural studies can be challenging.

The standard analytical methods for assaying enzyme-catalyzed reactions involve absorbance, fluorescence, or radioactivity readouts that require reaction components to be derivatized, labeled, or indirectly coupled to chromophores and tags (Bloch 2006). These strategies can work well for monitoring single reactant and product species but do not necessarily capture reactions with multiple intermediates. Furthermore, it is challenging to design spectroscopic experiments that can discriminate the activity of specific active sites in a multi-subunit enzyme complex, determine the stoichiometry of the intermediates or monitoring the sequence of multiple catalytic events.

The challenges of monitoring the PCAT1 transport cycle presents an opportunity to further develop a novel reaction monitoring strategy. Since chemical transformations mainly involve changes in mass from substrate to product, techniques that profile the masses of species present in the reaction and monitor mass changes are highly suitable (reviewed in (Ben-Nissan and Sharon, 2011; Fabris, 2005; Greis, 2007; Liesener and Karst, 2005; Ray et al., 2018). Native mass spectrometry (nMS) enables direct mass measurement of intact noncovalent macromolecular assemblies upon gas-phase transfer from nondenaturing solution conditions (Hernandez and Robinson, 2007; Leney and Heck, 2017; Loo, 1997; Tamara et al., 2021). By maintaining the quaternary structure of protein assemblies, nMS has provided key complementary structural information for hybrid and integrative structural studies, including complex composition, subunit stoichiometry and inter-subunit connectivity. The utility of nMS has also been extended to capture not only snapshots of these structural information under one particular condition and timepoint but also the associated changes in time when protein assemblies undergo dynamic processes such as subunit assembly/disassembly, conformational transitions, protein-ligand and protein-protein interactions and chemical reactions (reviewed in (Ben-Nissan and Sharon, 2011). With nMS being able to characterize noncovalent assemblies, it can detect substrate-bound complexes and any mass changes associated with the chemical reaction throughout the entire process and thus directly provide information on substrate occupancy, enzymatic turnover, intermediates, and order of catalytic steps in the reaction.

Several continuous flow methods hyphenated with MS detection for kinetic studies have been performed using various online setups and instrument configurations that monitored glycosidase, peptidase, nuclease and other enzymatic reactions (Dennhart et al. 2009; 2008; Dennhart and Letzel 2006; Kaufmann et al. 2014; Lee et al. 1989; van den Heuvel et al. 2005; Z. Yu et al. 2012). Here we perform nMS-based continuous monitoring of PCAT1 enzymatic reactions. this application by nMS-based continuous monitoring of PCAT1 enzymatic reactions. With high

resolution and accurate nMS readouts, the proteolytic cleavage and ATP hydrolysis reactions can be tracked unambiguously from characteristic mass shifts due to specific enzymatic activity (*e.g.*, release of cargo upon protease cleavage or loss of phosphate upon ATP hydrolysis). Furthermore, whether one or two substrates are bound and then catalytically processed can be directly measured from the corresponding mass losses. Overall, from the continuous recording, the reaction trajectories of all the relevant species including short-lived intermediates can be extracted and analyzed to uncover the stoichiometry, sequence of events, dynamics, and the mechanism of the reaction.

## 4.2 Summary

In this chapter, we performed nMS-based reaction monitoring of PCAT1 substrate cleavages and ATP hydrolysis. Four reactions are monitored: (1) PCAT1 peptidase activity in the absence of ATP, (2) PCAT1 ATPase activity in the absence of the substrate CtA, (3) PCAT1 activity with CtA and ATP but without Mg<sup>2+</sup> thereby preventing ATP hydrolysis, and (4) PCAT1 activity with CtA and Mg-ATP.

This chapter is adapted from the work co-authored with Dr. Paul Dominic Olinares from the laboratory of Brian Chait at the Rockefeller university. I performed the sample preparation and provided help with analyses. Dr. Dominic Olinares performed all the nMS experiments and time course data analyses.

## 4.3 Results

### 4.3.1 General workflow of the time-resolved nMS-based reaction monitoring

We performed real-time continuous reaction monitoring of PCAT1 activities with nMS readout (Fig 4.1). In this online MS setup, the reaction mixture is prepared and then loaded in an electrospray emitter that is directly coupled to the ion source of the MS instrument for continuous sampling and analyte detection. Because of this direct coupling, the solution conditions for the reaction must be nondenaturing, amenable for electrospray ionization and are free from non-volatile salts, buffers, and additives that could otherwise interfere with nMS characterization. First, PCAT1 and the substrate CtA are buffer exchanged into the nMS-compatible solution 150 mM ammonium acetate, 0.5% C, which is a nonvolatile solution maintained at physiological pH. The choice of MS-compatible detergent was based on previous detergent screening experiments for optimal native MS analysis of PCAT1 membrane protein wherein C was shown to require the lowest activation energy for detergent removal in the gas phase with minimal destabilization of membrane protein complexes (Kieuvongngam et al. 2020; Laganowsky et al. 2013).

For our nMS experiments, we use a commercial Exactive Plus EMR, an Orbitrap-based mass analyzer that is configured for measuring the masses of macromolecular assemblies at high resolving power and exceptional sensitivity (Rose et al. 2012). The samples are manually infused into the instrument at nanoflow rates (typically 20 – 50 nL/min) with a modified static nanospray source that enables easy access and on-the-fly troubleshooting of the electrospray (Olinares and

Chait 2020). At these nanoflow rates, a 5- $\mu$ L reaction mixture is generally sufficient for at least an hour or more of continuous reaction monitoring.

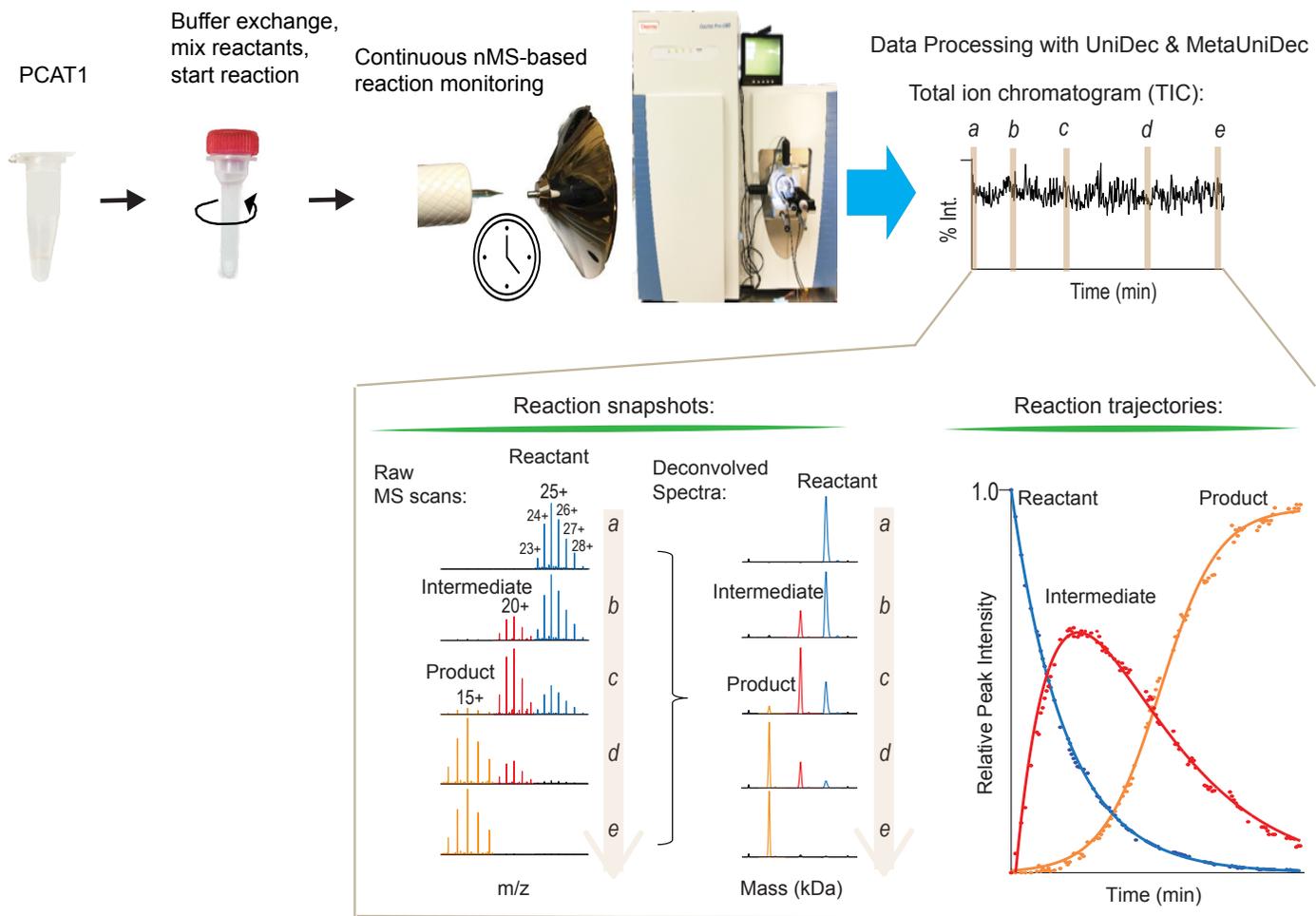
The reaction start time is recorded upon mixing of the buffer-exchanged reactants at the desired molar ratios. The reaction mixture is then quickly loaded into a conductive emitter and electrosprayed at room temperature into the ion source of the mass spectrometer to transfer the ionized analytes from solution to gas-phase for nMS detection. Upon mixing the reactants, several steps including manual sample loading, source setup and electrospray initiation contribute to an overall delay time of 1 – 2 min overall between the start of the reaction and MS data acquisition.

Each nMS scan generates a mass-to-charge ( $m/z$ ) spectrum. Mass spectra are acquired continuously with each scan taking about 600 – 1,200 ms depending on the ion injection time, automatic gain control (AGC) value, number of microscans, and instrument resolution settings, which impact the overall data acquisition cycle. In all the time course experiments performed here, this time window is sufficient to capture all the relevant species given their observed lifetimes and reaction rates.

The complete recording of all the nMS scans across the entire duration of the reaction yields a total ion chromatogram (TIC). Each species are detected as a population of multiply charged ions and thus register multiple  $m/z$  peaks on the spectrum. To obtain the actual mass values for each species present, the charge-state series are then processed and converted into zero-charge state spectra (also called deconvolved spectra) showing a single mass peak for each observed species. To obtain reaction snapshots at specific timepoints within the chromatogram, we use UniDec, a robust, rapid and freely available Bayesian-based deconvolution software (Marty et al. 2015). Each nMS spectrum takes  $\sim$ 1 min to deconvolve using UniDec enabling us to quickly process raw nMS spectra during or immediately after spectra acquisition.

To obtain higher signal-to-noise ratios and accurate mass measurements, the spectra across every 10 – 12 scans are averaged, which covers about 0.2 – 0.25 min time window in the chromatogram. As such, every 10 minutes of acquired data yields about 40 – 50 averaged spectra that need to be deconvolved. For efficient data processing, we use MetaUniDec, an expansion of the UniDec software that enables automated, streamlined and high-throughput processing of large nMS datasets (Reid et al. 2019b). The software extracts deconvolved mass and intensity information for each species detected across the entire chromatogram and provides additional post-processing features. To obtain the relative distribution of all relevant species at each timepoint, each extracted peak intensity is normalized against the sum of the peak intensities for all the relevant species that were observed within the spectrum. Note that as with other MS techniques, nMS is not inherently quantitative as peak signals may vary for each species depending on ionization efficiency, ion transmission and instrument response. Nevertheless, changes in relative abundances along the progress of the reaction can still be captured as reaction trajectories for each analyte. Based on the shapes of these plots, the type of reaction component (reactants, intermediates and products) and the sequence of events can be readily identified. The disappearance of the reactants during the reaction follows an exponential decay curve whereas the appearance of the products follows a sigmoidal or logistic curve characterized by an increase in abundance then subsequent plateau when all the starting material have been consumed. Specific intermediates appear and subsequently disappear delineating a rise-and-fall or a hill-shaped profile. In a multi-step reaction, the sequence of steps can be delineated by the reaction

profiles of the intermediates. Note however that the exact mathematical model that fits the intensity profile for each intermediate depends on the reaction mechanism, which can be challenging to determine. As such, the reaction trajectories for the intermediates observed in this study are derived from best-fit curves (typically a gaussian function or a sum of multiple exponential functions) without minimal assumptions regarding the reaction mechanism.



**Fig. 4.1 Workflow for nMS-based continuous reaction monitoring.**

### 4.3.2 Time course of substrate cleavage revealed by nMS reaction monitoring

In Chapter 2 of this dissertation, the cryo-EM structure of the ATP-free, peptidase-inactive PCAT1 mutant (C21A) incubated with CtA was elucidated. The structure shows that two substrates can bind PCAT simultaneously, but only one substrate is poised for cleavage and translocation while the other bound substrate has its cargo still oriented outside of the PEP domain. The asymmetric positioning suggests that PCAT1 might process one substrate at a time.

To characterize how PCAT1 cleaves its substrate and specifically whether two substrates are cleaved simultaneously or separately, we use native mass spectrometry to monitor the cleavage reaction over time. (Fig 4.2) This would require the ability to discriminate what happens to each of the two substrates. Because the binding and processing of substrates involves mass changes both in the substrate alone and with substrate-bound transporter, nMS is well-suited to continuously record all the events within the reaction and capture intermediates that would reveal the stoichiometry of substrate processing.

Each PCAT1 dimer has a mass of 162 kDa. The CtA substrate is a 10 kDa protein consisting of an N-terminal, 3-kDa leader peptide (LP) and a C-terminal, 7-kDa cargo (Fig 4.2 A). The substrate binds to the PEP domain on PCAT1 and gets cleaved thereby releasing the 7-kDa cargo and retaining the 3-kDa LP. Thus, a characteristic mass loss of 7 kDa indicates CtA proteolysis and cargo release. A range of possible substrate-transporter and mass combinations is possible that correspond to binding and cleavage of substrate

*wt*PCAT1 were incubated with two molar-excess of the full-length substrate. The reaction samples were continuously infused into the MS instrument at a constant rate. To capture the proteolytic cleavage reaction, we monitored the PCAT1-CtA assemblies at the high mass range (162-183 kDa) and the free substrate or cleaved cargo protein at the low mass range (7-10 kDa) within the same time-course experiment. We alternately switched the MS parameters for high mass range measurement and low mass range measurement (Fig 4.2 B).

In the beginning of the experiment, predominantly 183 kDa assemblies corresponding to PCAT1 bound to two full-length substrates were detected (Fig 4.2 C-D). This indicated that the PCAT1 can bind mainly two substrates. Over the course of the experiment, we detected the presence of 175 kDa assemblies. The 8-kDa mass difference corresponded to 66 amino acid residues of the cargo leaving the complex. Indicating that one substrate was cleaved while the other substrate remained bound to the complex. Later, we observed 172-, 168-, and 165-kDa assemblies corresponding to the PCAT1 with one full-length substrate, PCAT1 bind to two N-terminal binding domains of the substrates, and PCAT1 bound to only one N-terminal binding domain.

On the low mass range, the predominant peak at the earliest timepoint matched the mass of the full length CtA (10 kDa) which was initially present in excess. The 7-kDa cleaved cargo is observed coinciding with the presence of the 175-kDa in the high mass range that indicated that substrates were being processed. As the reaction progressed, the peak intensities for full-length CtA decreased as it was being consumed with concomitant increase in the peak intensities for the 7-kDa cargo.

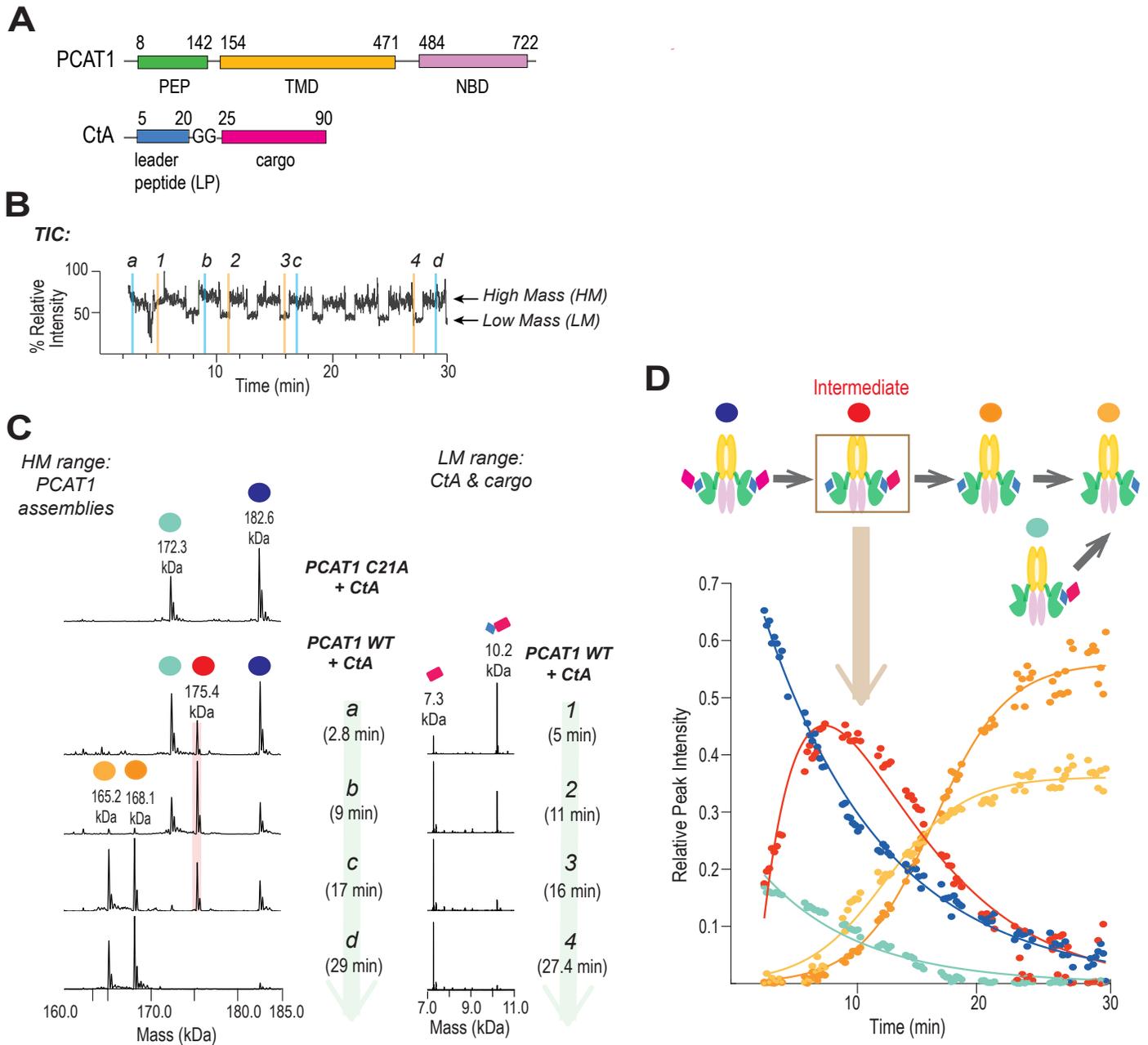
Overall, the results of the time course study suggest that PCAT1 binds to two substrates but only one substrate is cleaved at a time. Moreover, the C-terminal cargo domain is dissociated from the

core transporter while the N-terminal binding domain remains tethered to the PEP. This finding strengthens the structural evidence presented in Chapter 2 that only one substrate is cleaved and inserted into the TM cavity while the other substrate remains bound outside the TM cavity. The result implies that the binding interaction between PCAT1 and the substrate is primarily composed of the interaction at the PEP. The cleaved cargo is weakly associated with PCAT1 inside the TM cavity and can dissociate from the complex. The presence of a cleaved cargo domain is detected as well indicating that ATP is not necessary for cleavage or dissociation of cargo from the TM cavity consistent with previous studies (Lin, Huang, and Chen 2015).

**Table 4.1 Mass measurements from the nMS-based reaction monitoring of PCAT1 activities.**

Sample	Measured Mass $\pm$ SD (Da) <sup>a</sup>		Protein or Protein Complex	Expected Mass (Da)	$\Delta$ Mass (Da)	% Mass Error
<i>PCAT1 peptidase activity, ATP-free</i>						
	182,649	$\pm$ 4	PCAT1 + 2 CtA	182,592	57	0.03
	175,399	$\pm$ 4	PCAT1 + 1 CtA + 1 LP	175,339	60	0.03
	172,426	$\pm$ 13	PCAT1 + 1 CtA	172,384	42	0.02
	168,149	$\pm$ 5	PCAT1 + 2 LP	168,087	63	0.04
	165,192	$\pm$ 8	PCAT1 + 1 LP	165,131	61	0.04
	10,203.9	$\pm$ 0.4	full-length CtA	10,207.9	-4	-0.04
	7,267.2	$\pm$ 0.3	CtA cargo (C-terminal)	7,270.6	-3	-0.05
<i>PCAT1 ATPase activity, substrate-free</i>						
	163,203	$\pm$ 3	PCAT1 + 2 ATP	163,190	12	0.01
	163,114	$\pm$ 2	PCAT1 + 1 ATP + 1 ADP	163,110	4	0.002
	163,031	$\pm$ 1	PCAT1 + 2 ADP	163,030	1	0.001
	162,606	$\pm$ 1	PCAT1 + 1 ADP	162,603	3	0.002
	162,175	$\pm$ 2	PCAT1	162,176	-1	0.000
<i>PCAT1 with substrate and ATP, Mg<sup>2+</sup> absent</i>						
	183,636	$\pm$ 5	PCAT1 + 2 CtA + 2 ATP	183,606	30	0.02
	173,426	$\pm$ 4	PCAT1 + 1 CtA + 1 ATP	173,398	28	0.02
	163,216	$\pm$ 3	PCAT1 + 2 ATP	163,190	26	0.02
<i>PCAT1 peptidase and ATPase activity</i>						
	183,649	$\pm$ 11	PCAT1 + 2 CtA + 2ATP	183,606	43	0.02
	183,563	$\pm$ 4	PCAT1 + 2 CtA + 1 ATP + 1 ADP	183,526	37	0.02
	183,500	$\pm$ 9	PCAT1 + 2 CtA + 2 ADP	183,446	54	0.03
	183,046	$\pm$ 4	PCAT1 + 2 CtA + 1 ADP	183,018	28	0.02
	176,251	$\pm$ 12	PCAT1 + 1 CtA + 1 LP + 2 ADP	176,194	57	0.03
	175,799	$\pm$ 9	PCAT1 + 1 CtA + 1 LP + 1 ADP	175,766	33	0.02
	173,433	$\pm$ 5	PCAT1 + 1 CtA + 2 ATP	173,398	35	0.02
	173,351	$\pm$ 5	PCAT1 + 1 CtA + 1 ATP + 1 ADP	173,318	33	0.02
	173,287	$\pm$ 4	PCAT1 + 1 CtA + 2 ADP	173,238	49	0.03
	172,840	$\pm$ 5	PCAT1 + 1 CtA + 1 ADP	172,811	29	0.02
	169,003	$\pm$ 12	PCAT1 + 2 LP + 2 ADP	168,941	62	0.04
	168,546	$\pm$ 9	PCAT1 + 2 LP + 1 ADP	168,514	32	0.02
	166,042	$\pm$ 11	PCAT1 + 1 LP + 2 ADP	165,986	56	0.03
	165,587	$\pm$ 4	PCAT1 + 1 LP + 1 ADP	165,558	29	0.02
	163,074	$\pm$ 6	PCAT1 + 2 ADP	163,030	44	0.03

<sup>a</sup> Masses were determined from averaged spectra within a specific timepoint range. Measured masses are reported as the average and S.D. of all the calculated masses across the charge-state distribution ( $n \geq 4$ ) for each species.



**Fig. 4.2 nMS-based monitoring of the PCAT1 peptidase activity.**

(A) Domain composition of PCAT1 and the substrate CtA.

(B) The TIC for the complete peptidase reaction. The nMS parameters were set to alternate in monitoring the PCAT1 assemblies at the high-mass range and the substrate at the low-mass range.

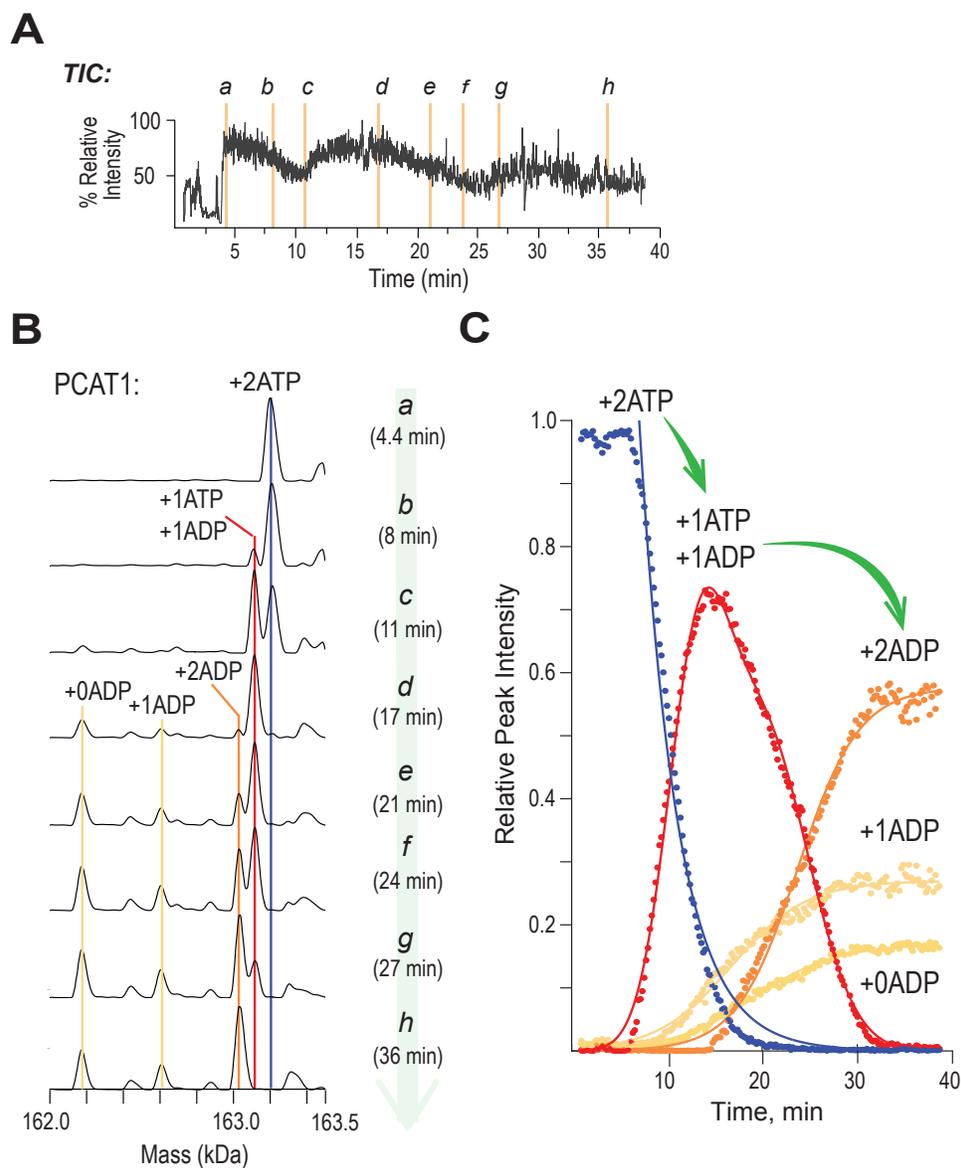
(C) Representative deconvolved spectra at high-mass and low-mass ranges from specific timepoints.

(E) Reaction trajectories for the substrate-bound PCAT1 assemblies revealing the main reaction intermediate.

### 4.3.3 Transient ATP hydrolysis intermediates captured by time-resolved native mass spectrometry

We monitored the ATP hydrolysis activity of PCAT1 in the absence of the substrate (Fig. 4.3). To setup the experiment, PCAT1 was incubated with 0.5 mM ATP, and 1 mM EDTA and then buffer-exchanged into an ATP-free nMS solution to remove unbound ATP and other non-volatile components. An aliquot of magnesium acetate stock was then mixed with the buffer-exchanged sample to a final concentration of 0.5 mM  $Mg^{2+}$  and the reaction mixture was immediately loaded on the emitter for continuous nMS recording (Fig. 4.3 A).

At the earliest timepoint, the two nucleotide binding sites were occupied by ATP (Fig. 4.3 B-C). The first mass change observed was a loss of 89 Da corresponding to the release of an inorganic phosphate group from one round of ATP hydrolysis. At later time points, the remaining bound ATP was also hydrolyzed resulting in another 83 Da loss forming the final product PCAT1–2ADP. The nucleotide binding site has lower affinity for ADP. The dissociation of ADP is observed generating PCAT1-1ADP and *apo* PCAT1. The overall reaction trajectories and the presence of the intermediate indicates that ATP hydrolysis processed sequentially one at a time.



**Fig. 4.3 nMS-based monitoring of the PCAT1 ATP hydrolysis activity.**

(A) The TIC for the complete ATPase reaction.

(B) Representative deconvoluted spectra at specific timepoints showing successive hydrolysis of the bound ATP. (C) Reaction trajectories for the assemblies observed in the ATPase reaction revealing the main reaction intermediate (in red).

#### 4.3.4 Determining the coupling of proteolytic cleavage function and ATP hydrolysis of PCAT1

In the previous section, we have separately characterized how the peptidase domain processes the substrate, and how NBD domains hydrolyses ATP. In the Chapter 3 of this dissertation, we present structural evidence that the two functions are tightly coupled. Here we recapitulate the cryo-EM finding using a different approach and answer the remaining challenge as to how many ATP hydrolysis is required for a single substrate transport.

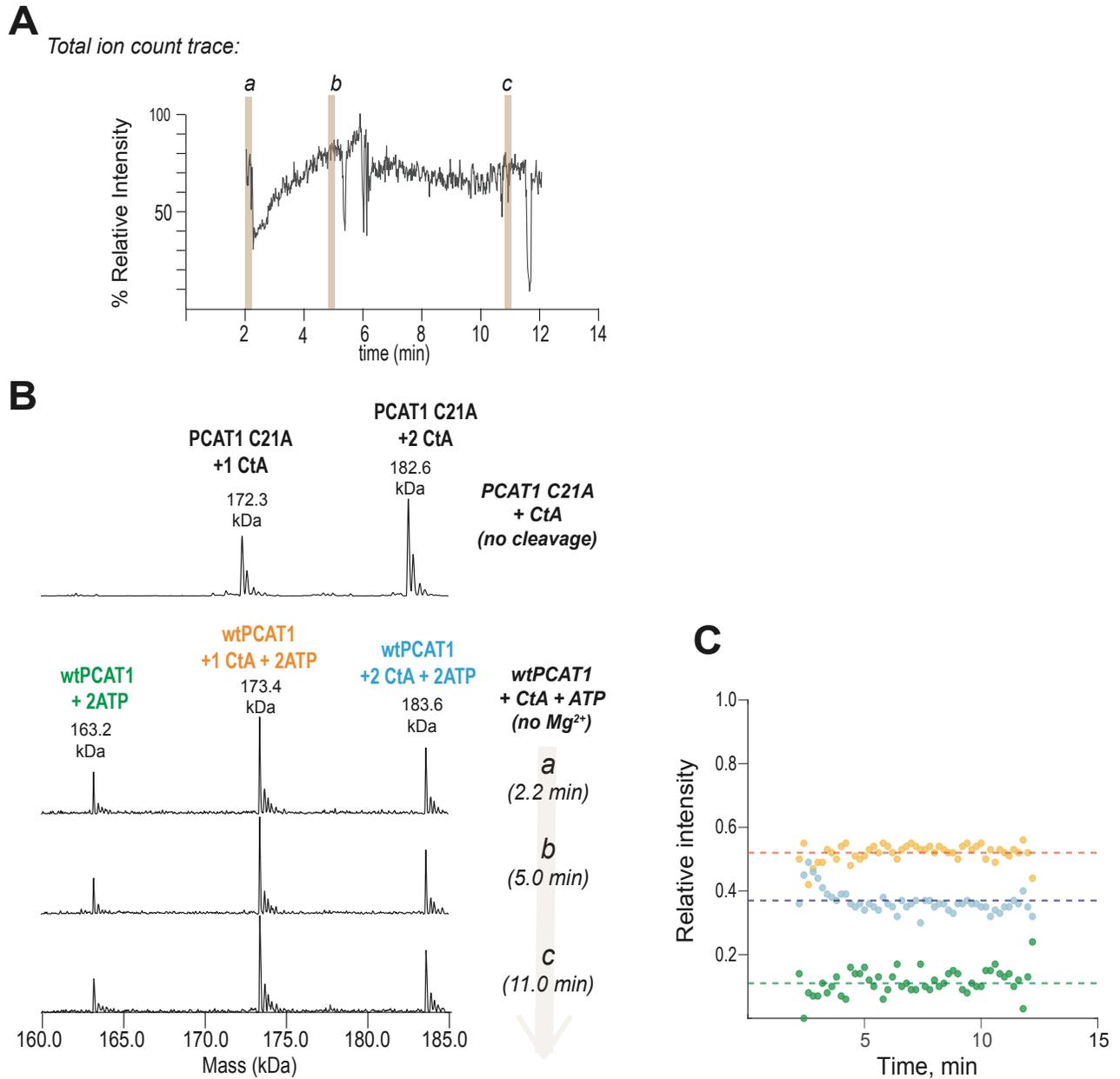
First, we performed a time course with PCAT1 in the presence of CtA and ATP but in the absence of  $Mg^{2+}$  to prevent ATP hydrolysis (Fig. 4.4). A previous cryo-EM experiment in chapter 3 showed that ATP binding without  $Mg^{2+}$  inhibited PCAT1 peptidase activity. We recapitulated this by incubating PCAT1 with 0.5 mM ATP and 1 mM EDTA first then buffer-exchanged into the nMS solution. The substrate was buffer-exchanged separately into the same nMS solution and then mixed with the PCAT1-ATP mixture. Despite having fully loaded active sites (PCAT1 + 2CtA + 2ATP), no substrate cleavage occurred (i.e., no 7-kDa mass loss observed) at any time within the 12 min reaction (Fig. 4.4 B-C), consistent with previous structural conclusion in the chapter 3 that ATP binding without hydrolysis traps the transporter in the proteolytic inactive state.

Next, to monitor what happens when ATP hydrolysis is allowed to proceed, the same mixture is prepared as before with PCAT1-ATP but with a slight excess of ATP (10  $\mu$ M) followed by substrate addition and then flushed with 0.5 mM  $Mg^{2+}$  (Fig. 4.5). From the time course analysis, fifteen assemblies were observed and their reaction trajectories were monitored (Fig. 4.5B-C). At the beginning (Fig. 4.5B, point a), the two starting assemblies (PCAT1 + 2 ATP + 2CtA and PCAT1 + 2ATP + 1 CtA) are present. Then, the first reaction that took place was the hydrolysis of one bound ATP on both starting assemblies as indicated by the mass loss of 80 Da (Fig. 4.5B, point b-c). The dominant assemblies contained one ADP and one ATP bound to either one or two full-length uncleaved substrate. This is consistent with the nMS-based monitoring for the ATPase activity of substrate-free PCAT1 whereby two ATP molecules were loaded but ATP hydrolysis proceeded one at a time. The remaining bound ATP was then hydrolyzed in the two intermediates (Fig. 4.5B, point d-g) forming the +2 ADP assemblies and ADP diffuses out forming +1 ADP assemblies.

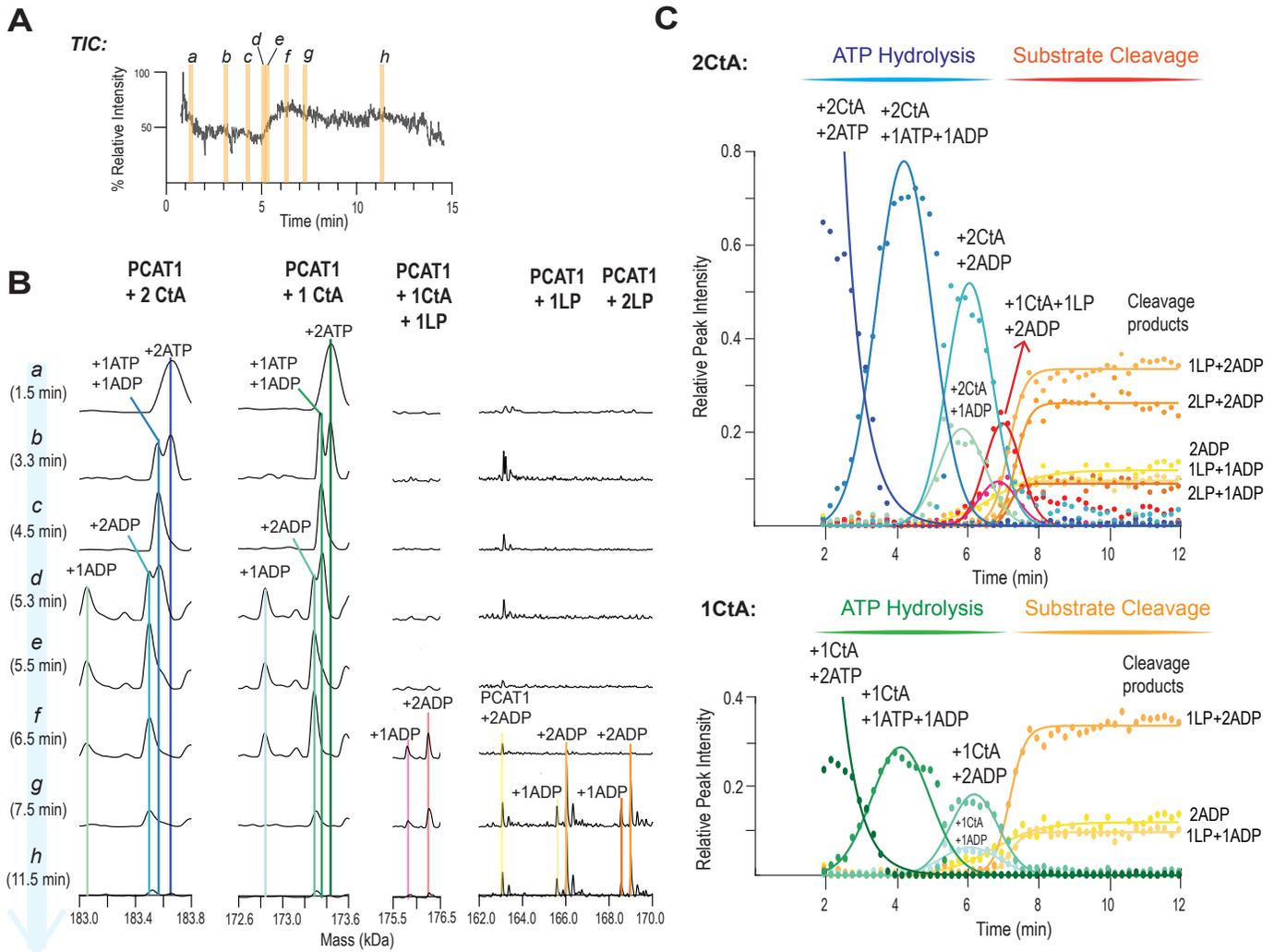
Since the earliest time point to up to point d where +2 ATP were bound and then the two subsequent hydrolysis occurred where at least one ATP remained bound, no substrate cleavage was observed. However, between points d and e, where the predominance of +2ADP and +1ADP assemblies indicating that all the bound ATP and the excess ATP in solution had been hydrolyzed, cleavage of the bound substrate(s) began. At this point, new assemblies were observed where one or two 7-kDa mass losses were detected due to cargo release (Fig. 4.5B, point f-h). However, the cycle of one round of ATP hydrolysis and one round of substrate processing was not observed. This point will be pointed out in the Discussion section.

Consistent with what was found in the time course analysis of peptidase activity with ATP-free PCAT1, the bound substrate was cleaved one substrate at a time. Overall, substrate cleavage only proceeded when there was no ATP bound or all the bound ATP had been hydrolyzed in PCAT1

(Fig. 4.5 C). Together with the structural findings in chapter 3, we conclude that (1) ATP binding does not preclude the substrate binding, (2) one ATP binding is sufficient to inhibit proteolytic cleavage activity, and (3) hydrolysis of both ATPs is required to reset the PCAT1 back to inward-facing cleavage competent state.



**Fig. 4.4 nMS-based monitoring of PCAT1 peptidase and ATP hydrolysis activity in the absence of  $Mg^{2+}$**   
 (A) Total ion count trace for the complete reaction. (B) MS snapshots at different timepoints along the reaction. (C) Reaction trajectory plots for the three species observed. No major changes in species occurred throughout the entire duration indicating that no substrate cleavage reaction had occurred.



**Fig. 4.5. nMS-based monitoring of the peptidase and ATP hydrolysis activity of PCAT1.**

(A) The TIC for the complete PCAT1 reaction.

(B) Representative deconvolved spectra at specific timepoints for all the assemblies observed.

(C) Reaction trajectories for the observed PCAT1 assemblies that started out with 2 CtA (top panel) or 1 CtA (bottom panel). In both cases, peptidase activity only occurred when the two bound ATP have been hydrolyzed.

## 4.4 Discussion

In this work, a novel native mass spectrometry-based continuous reaction monitoring strategy was developed for characterizing the dynamics of large multimeric enzyme assemblies. The method was tested on PCAT1 and had yielded valuable insights into the function of PCAT1.

First, PCAT1 binds to two peptide substrates and only one substrate is cleaved at a time. This finding supports our structural evidence in chapter 2 that even though two substrates are found at the PEP, only one is oriented for translocation.

Second, PCAT1 under ATP non-hydrolysable condition permits substrate binding but inhibits substrate cleavage. This finding further supports biochemical evidence in David Lin paper and strengthens the structural evidence in chapter 3. The cryo-EM structure of wtPCAT1 in ATP-bound conformation shows blob-like low resolution densities of disengaged PEP, making it difficult to discern whether a substrate is bound or not. The nMS result here conclusively demonstrates that indeed the substrate can still bind to a disengaged PEP.

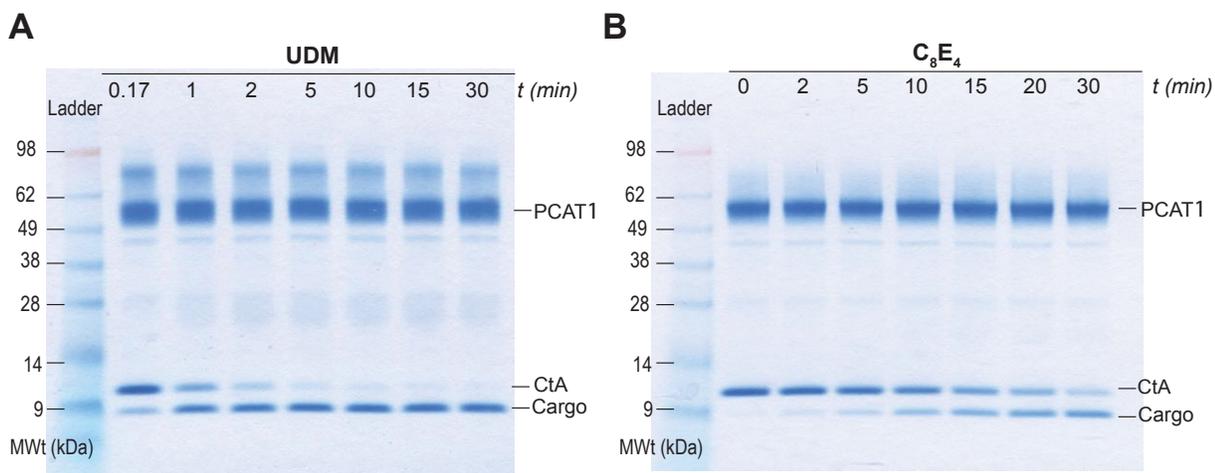
Lastly, the time course experiment of PCAT1 in the presence of substrates and ATP-Mg reveals that ATP hydrolysis occurs sequentially. Both ATP are hydrolyzed before PCAT1 can resume substrate cleavage processing. This implies that the intermediate species PCAT1 + ATP adopts a PEP disengaged conformation. It is curious to see if this prediction holds true if one engineers heterodimeric PCAT1 having one inactive nucleotide binding site.

Despite the advances made by this approach, limitations should be noted as well. The nMS-compatible condition for membrane protein complexes requires screening for an optimal detergent. On the one hand the detergent stabilizes the membrane protein for nMS analysis. On the other hand, the detergent should have a low activation energy for detergent removal in the gas phase without destabilizing the protein complex. C<sub>8</sub>E<sub>4</sub> is shown to be a promising nMS-compatible detergent with minimal gas-phase dissociation. However, the choice of detergents affects the structure and subsequently the functional properties of membrane proteins. For example, PCAT1 in C<sub>8</sub>E<sub>4</sub> has much slower substrate cleavage activity compared to PCAT1 in UDM, the detergent used in chapter 2 and 3 for structural studies (Fig. 4.6).

We observe that in the time course experiment of PCAT1 in the presence of substrates and ATP-Mg, the active turnover cycle was not detected. Rather, we see that all the available ATP (even the excess ATP present in solution) was hydrolyzed first before substrate cleavage occurred. One possible reason is that substrate cleavage activity is much slower in C<sub>8</sub>E<sub>4</sub> while ATP hydrolysis and nucleotide exchange can occur at relatively faster rate. The faster cycle of ATP exchange and hydrolysis blocks the slower substrate cleavage process. After ATP is exhausted then the substrate cleavage can proceed.

In conclusion, this chapter has shown the utility of nMS for monitoring the multi-catalytic activity and dynamics of wtPCAT1. The setup is simple and robust. The focus of the nMS experiment is to capture stoichiometric information. In addition, the time-resolved experiment has answered fundamental questions of PCAT1 dynamics, namely sequences of events in ATP hydrolysis and substrate cleavage, which would have been much more difficult to determine

from other currently available analytical and biophysical methods. Although the nMS-compatible conditions used (e.g., choice of C<sub>8</sub>E<sub>4</sub> detergent and ammonium acetate solution) may affect the reaction rate and other kinetic parameters, the stoichiometry and sequence of events are likely to be preserved and can yield important insights into the function of PCAT1.



**Fig. 4.6. Reaction monitoring of PCAT1 peptidase activity in nMS-compatible detergent C8E4 and detergent used for structural study UDM**

(A) 0.058% w/v UDM and (B) 0.5% v/v C8E4 detergent with SDS-PAGE as readout.

For these time course-experiments, 5  $\mu\text{M}$  PCAT1 (monomer) was mixed with 10  $\mu\text{M}$  CtA

(1:2 protease:substrate ratio) in 200 mM ammonium acetate with detergent at 2X CMC.

Aliquots were collected at specific timepoints and quenched with gel loading buffer prior to SDS-PAGE analysis.

## 4.4 Materials and Methods

### 4.5.1 Sample preparation for native mass spectrometric study

Sample preparation for native mass spectrometry is modified from the standard sample preparation described in previous chapters. The modification aims at minimizing detergents and replacing standard Tris buffer with a more volatile ammonium acetate buffer. Purifying PCAT1 in 1.2 mM UDM (2x critical micelle concentration) produces a functional PCAT1 that exhibit cleavage activity and ATP hydrolysis activity. It has been shown that detergent C8E4 produces a well-behaved sample for the purpose of native-MS for membrane proteins (Reading et al. 2015). We buffer-exchanged the PCAT1 from 1.2 mM UDM to 0.5% C8E4 just prior to native MS analysis.

### 4.5.2 Native mass spectrometry analysis

All PCAT1 and CtA samples were buffer exchanged into nMS solution (200 mM ammonium acetate pH 7.5, 0.5% C8E4) using Zeba microspin desalting columns (Thermo Scientific) with a 7 kDa or 40 kDa molecular weight cut-off (MWCO).

The typical concentrations used for native MS analysis were 4  $\mu$ M PCAT1 monomer for the PCAT1-only sample and 4  $\mu$ M PCAT1 monomer + 8  $\mu$ M CtA (a two-fold excess of CtA). An aliquot (2–3  $\mu$ L) of the sample was loaded into a gold-coated quartz capillary that was fabricated in-house. The sample was then electrosprayed into an Exactive Plus EMR instrument (Thermo Fisher Scientific) using a static nanospray source. The MS parameters used included: spray voltage, 1.0–1.4 kV; capillary temperature, 125°C; S-lens RF level, 200; resolving power, 8750 or 17,500 at m/z of 200; AGC target,  $1 \times 10^6$ ; number of microscans, 5; maximum injection time, 200 ms; injection flatapole, 8 V; interflatapole, 4 V; bent flatapole, 4 V; ultrahigh vacuum pressure,  $7\text{--}10 \times 10^{-10}$  mbar; total number of scans, 100. The in-source dissociation (ISD) and higher-energy collisional dissociation (HCD) parameters were varied and optimized accordingly. Mass calibration was performed using cesium iodide.

For PCAT1 peptidase activity monitoring, the buffer-exchanged PCAT1 and CtA were mixed in a tube at 1:2 mole ratio and 5  $\mu$ L of the mixture was analyzed. It took 1.5 – 2 min to load the sample, insert the tip to the instrument source and initiate nanoelectrospray; this lag time between mixing the reactants and the start of spectra acquisition was recorded for every experiment. Two alternating MS settings were used to capture changes in both the PCAT1 assemblies (162 – 182 kDa, designated here as high mass or HM) and the substrate (7 – 10 kDa, designated here as low mass or LM). The differences in MS parameters when switching from HM to LM included: HCD, from 200 V to 50 V; interflatapole voltage, from 4 V to 7 V; bent flatapole voltage, from 4 V to 6 V. At least 100 scans at the HM range and 50 scans at the LM range were acquired for every switching cycle. The rest of the MS parameters used were as stated above. Mass spectra were acquired continuously in real-time mode without averaging until no further changes in the spectral pattern were observed. Spectra were processed manually.

For the PCAT1 ATPase monitoring, PCAT1 was incubated with 0.5 mM ATP and 1 mM EDTA on ice for 5 min and then buffer exchanged into nMS solution. Magnesium acetate was added to a final concentration of 0.5 mM and immediately loaded onto spray emitter for nMS analysis.

For the PCAT1 full activity profiling, PCAT1 was incubated with 0.5 mM ATP and 1 mM EDTA on ice for 5 min and then buffer exchanged into nMS solution. An additional aliquot of ATP to a final concentration of 10  $\mu$ M was then added. Twofold molar excess of buffer-exchanged CtA was then mixed to the reaction mixture. Magnesium acetate was added to a final concentration of 0.5 mM and immediately loaded onto spray emitter for nMS analysis.

### 4.5.3 Data Processing

The nMS RAW file for each time course was processed using MetaUniDec software (v 4.4.1). First, the RAW file was converted into a hierarchical data format 5 file (HDF5) using the Automatic Chromatographic Parsing function with auto import by time (either 0.2 or 0.25 min). To obtain higher signal-to-noise ratios and accurate mass measurements and to reduce the number of data points, the spectra across every 10 – 12 scans are averaged, which covers about 0.2 – 0.25 min time window in the chromatogram. The resulting HDF5 was then used for deconvolution to extract mass and intensity information for each species detected.

For the PCAT1 peptidase activity monitoring wherein the nMS settings were collected at high mass (HM) and low mass (LM) ranges, the corresponding time points were manually inputted using the HDF5 Import Wizard to create separate HDF5s for the two mass analysis ranges.

For PCAT1 peptidase activity, High Mass: m/z range: 5,000 – 10,000 Th; no background subtraction; mass range: 160,000 – 185,000; sampling mass: 1 Da; smooth charge distribution, on; suppress artifacts, other with beta set to 20; picking range: 500 Da; picking threshold: 0.1.

For PCAT1 peptidase activity, Low Mass: m/z range: 500 – 5,000 Th; no background subtraction; mass range: 5,000 – 15,000; sampling mass: 1 Da; smooth charge distribution, on; suppress artifacts, none; picking range: 500 Da; picking threshold: 0.1.

For PCAT1 ATPase monitoring: m/z range: 5,000 – 10,000 Th; no background subtraction; mass range: 160,000 – 185,000; sampling mass: 1 Da; smooth charge distribution, on; suppress artifacts, other with beta set to 20; picking range: 50 Da; picking threshold: 0.2.

For PCAT1 peptidase-ATP binding ( $Mg^{2+}$ -free) monitoring: m/z range: 5,000 – 15,000 Th; no background subtraction; mass range: 160,000 – 185,000; sampling mass: 1 Da; smooth charge distribution, on; suppress artifacts, other with beta set to 20; picking range: 500 Da; picking threshold: 0.1.

For PCAT1 peptidase-ATPase monitoring: m/z range: 4,000 – 12,000 Th; no background subtraction; mass range: 160,000 – 185,000; sampling mass: 1 Da; smooth charge distribution, on; suppress artifacts, other with beta set to 20; picking range: 10 Da; picking threshold: 0.05.

For all the processed data, the mass and peak intensities were extracted using the following settings: Peak normalization, Max; Peak extraction, height; Extract normalization, Sum. With these parameters, the peak heights for each deconvolved peak were normalized to the sum of the peak heights for all species detected at each timepoint.

From the peak intensity and time course information, the reaction trajectories for the relevant species were plotted and determined using Matlab R2021a (version update 3,

Mathworks) with scripts prepared in-house. The time course data for the reactants and products were either fit to an exponential decay or a sigmoidal/logistic function. The reaction trajectories for the intermediates were plotted with the best-fit curves (typically a gaussian function or a sum of two exponential functions) without minimal assumptions regarding the reaction mechanism.

## Chapter 5: Conclusion

### 5.1 Concluding remarks

Previous investigations of peptide processing and secretion ABC transporters have shown that PCATs are proteolytically active and have identified the catalytic mechanism responsible for substrate maturation cleavages (K. H. Wu and Tai 2004; K. Wu, Hsieh, and Tai 2012; van Belkum, Worobo, and Stiles 1997). Prior structural studies have been focused on using X-ray crystallography to study the peptidase domain and the full-length transporters (Ishii et al. 2010; Bobeica et al. 2019; Lin, Huang, and Chen 2015). These investigations hint that the active transport mechanism follows the alternating access model, and the substrate processing and transport are coupled. The previous endeavors have been limited by the challenge of obtaining crystals of PCATs in complex with the substrates. This limitation was circumvented with the advent of cryo-EM techniques.

In the preceding chapters, I have determined the structure of PCAT1 in complex with its peptide substrate using the cryo-EM technique. The structure establishes the determinants for substrate specificity. Biochemical characterization has confirmed the conserved and unique residues that participate in substrate binding. In addition, the work establishes a working model of how polypeptides are processed and transported. I have found that PCAT1 lacks a substrate binding site in the transmembrane cavity, a feature widely found among ABC exporters. Instead, PEP functions as an intracellular substrate recruiting site. The absence of a substrate binding site in the cavity prevents mature cargo from re-entering the cavity in the outward-facing conformation. This feature is consistent with the alternating access model that a high affinity substrate binding site is found in one side of the membrane and is absent in the other side.

To study the mechanism responsible for substrate processing and transport coupling, I have determined the structure of PCAT1 in the  $Mg^{2+}$  free condition. This structure provides the first outward-facing structures of PCATs and have established that the transport mechanism is indeed consistent with the alternative access model. The coupling mechanism is mediated through the concerted disengaging motion of PEP away from the core transporters as NBD becomes dimerized.

In addition, to study the dynamic of the peptide transport cycle, I have determined the structures of PCAT1 under the active turnover condition. The study has shown the existence of multiple inward-facing conformations, suggesting that PCAT1 is a highly dynamic transporter. The existence of inward-facing ATP bound states suggests that the NBD dimerization represents the rate limiting step in the transport cycle. The role of ATP is to stabilize the NBD dimerized state and provides directionality needed for active transport.

Finally, I have collaborated with the laboratory of mass spectrometry and gaseous ion chemistry at The Rockefeller University to investigate the short-lived intermediates of PCAT1 in the transport cycle. A novel method for continuous, real-time reaction monitoring by native mass spectrometry has been applied to the study of PCAT1 peptidase and ATP hydrolysis activity. The study has uncovered the stoichiometric information of the intermediates of PCAT1 in the act of substrate cleavage. It has also captured the sequence of events in substrate cleavage and ATP hydrolysis.

In conclusion, these structural, biochemical, and biophysical studies further expand our knowledge of peptide processing and secretion by ABC transporters. The work not only provides a more in-depth understanding of gram-positive peptide processing and secretion, but it also provides new insight into structural dynamics of ABC transporters.

## **5.2 Future directions**

In the following sections, I discuss the relevance of the work in the context of the advancements in ABC transporter research. I further discuss promising research avenues that the work can lead to and what limitations and challenge may lie ahead.

### **5.2.1 PCATs in Gram-negative bacteria**

While my present work has focuses entirely on PCAT1 from Gram positive bacteria, PCATs that exist in Gram-negative bacteria forms a large membrane transporter complex, called Type-1 secretion systems (S. Thomas, Holland, and Schmitt 2014). The T1SS are large macromolecular machinery composed of a PCAT, a periplasmic membrane-fusion protein, and an outer-membrane porin. This assembly is thought to create a continuous channel spanning inner and outer membranes, allowing for transport of a large protein substrate to be extruded extracellularly (Lenders et al. 2013). This transport mechanism is different from PCATs in Gram-positive bacteria which operates via the alternating access mechanism. Moreover, the peptidase domain in T1SS lacks the catalytic residues for cleavage and is proteolytically inactive (Lecher et al. 2012). The role of peptidase domain in binding the substrate and orienting the substrate for translocation remains poorly understood.

It is curious to compare structures of PCATs in T1SS to PCAT1 to understand how evolution has adapted PCATs for different functions. Other research goals may include investigations into what role the inactive peptidase domain plays in the protein substrate binding and translocation, and what transport mechanism is required for transport of a large protein substrate.

The current challenge in studying the structure and function of T1SS include the difficulty in preserving T1SS assembly during sample preparation and purification. This is perhaps due to the assembly spanning two membranes with different lipid compositions, making it difficult to stabilize the whole protein complex. However, with the advent of high-resolution cryo-electron tomography (cryo-ET) it may be possible to visualize the T1SS machinery *in situ* (Kaplan et al. 2021; Doerr 2017).

### **5.2.2 energetic landscapes of ABC transporters**

Recent advances in ABC transporter research focuses on understanding the conformational spaces in the transport cycle (Hofmann et al. 2019; Q. Yu et al. 2021; Wang et al. 2020). These advances are enabled by single-particle cryo-EM multiple reconstructions from heterogenous samples. My present work in chapter 3 was possible thanks to this technical breakthrough. I conclude in chapter 3 that PCAT1 conformations obtained in the active turnover condition reflects the kinetic bottleneck in the transport cycle, while conformations obtained under

thermodynamic equilibrium reflect the energetic landscape of the system according to the Boltzmann's distribution.

Synthesizing from previous active turnover experiments that has been carried out thus far, I conclude that the rate limiting steps differ among ABC transporters (Wang et al. 2020; Q. Yu et al. 2021; Hofmann et al. 2019). The rate limiting steps of PCAT1 and ABCG2 is NBD dimerization. Both PCAT1 and ABCG2 are homodimer having two "consensus" nucleotide binding sites (Q. Yu et al. 2021). On the contrary, the rate limiting steps for MRP1 is ATP dissociation from the degenerate site (Wang et al. 2020). It is tempted to think that the difference is due to the presence of the degenerate site. However, future work on is needed to settle whether this hypothesis holds for other ABC transporters, and what the structural determinant of the rate limiting step difference is.

One should be cautious not to draw quantitative conclusions of conformation distribution from the single particle 3D classification procedure alone. This is because the particles can be misclassified to a wrong conformation at low resolution. In addition, conformational distribution is influence by temperature according to Boltzmann distribution. The current plunge freezing method for cryo-EM sample preparation is estimated to yield a cooling rate of  $\sim 10^5$  K per second (Dobro et al. 2010). Therefore, it is theoretically possible that a conformational change faster than 1 microsecond to 1 microsecond could be influenced by sample freezing. Furthermore, in the current 3D classification algorithms, users must specify the parameters for class of particles ( $k$ ). Users may not know *a priori* how many conformations exist in each condition. Therefore, a class that contains a small number of particles may not be reconstructed from the procedure. Moreover, the class distribution cannot inform us of how the states evolve temporally. On the other hand, techniques such as single-molecule fluorescence resonance transfer (smFRET) is more suitable for studying conformation distribution and kinetic mechanism of ABC transporters (Husada et al. 2018; Liu et al. 2018; Wang et al. 2020). The synergy between two approaches may enable the fields to both understand the mechanism at the atomic level and the physiological temporal scale.

## References:

- Adams, Paul D., Pavel v. Afonine, Gábor Bunkóczi, Vincent B. Chen, Ian W. Davis, Nathaniel Echols, Jeffrey J. Headd, et al. 2010. "PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution." *Acta Crystallographica Section D: Biological Crystallography* 66 (2): 213–21. <https://doi.org/10.1107/S0907444909052925>.
- Beek, Josy ter, Albert Guskov, and Dirk Jan Slotboom. 2014. "Structural Diversity of ABC Transporters." *The Journal of General Physiology* 143 (4): 419–35. <https://doi.org/10.1085/jgp.201411164>.
- Belkum, Marco J. van, Randy W. Worobo, and Michael E. Stiles. 1997. "Double-Glycine-Type Leader Peptides Direct Secretion of Bacteriocins by ABC Transporters: Colicin V Secretion in *Lactococcus Lactis*." *Molecular Microbiology* 23 (6): 1293–1301. <https://doi.org/10.1046/j.1365-2958.1997.3111677.x>.
- Binet, Rachel, Sylvie Létoffé, Jean Marc Ghigo, Philippe Delepelaire, and Cécile Wandersman. 1997. "Protein Secretion by Gram-Negative Bacterial ABC Exporters - A Review." *Gene* 192 (1): 7–11. [https://doi.org/10.1016/S0378-1119\(96\)00829-3](https://doi.org/10.1016/S0378-1119(96)00829-3).
- Bloch, Will. 2006. "J.-L. Reymond (Ed.): Enzyme Assays. High-Throughput Screening, Genetic Selection and Fingerprinting." *Analytical and Bioanalytical Chemistry* 386 (6): 1583–84. <https://doi.org/10.1007/s00216-006-0813-7>.
- Bobeica, Silvia C, Shi-hui Dong, Liujie Huo, Nuria Mazo, Satish K Nair, Martin I McLaughlin, and Gonzalo Jime. 2019. "Insights into AMS / PCAT Transporters from Biochemical and Structural Characterization of a Double Glycine Motif Protease," 1–27.
- Brown, Alan, Fei Long, Robert A. Nicholls, Jaan Toots, Paul Emsley, and Garib Murshudov. 2015. "Tools for Macromolecular Model Building and Refinement into Electron Cryo-Microscopy Reconstructions." *Acta Crystallographica Section D: Biological Crystallography* 71: 136–53. <https://doi.org/10.1107/S1399004714021683>.
- Chen, Jue, Gang Lu, Jeffrey Lin, Amy L Davidson, and Florante A Quiocho. 2003. "A Tweezers-like Motion of the ATP-Binding Cassette Dimer in an ABC Transport Cycle." *Molecular Cell*. Vol. 12.
- Chen, Min, Rupert Abele, and Robert Tampé. 2003. "Peptides Induce ATP Hydrolysis at Both Subunits of the Transporter Associated with Antigen Processing\*." *Journal of Biological Chemistry* 278 (32): 29686–92. <https://doi.org/https://doi.org/10.1074/jbc.M302757200>.
- Chen, Vincent B., W. Bryan Arendall, Jeffrey J. Headd, Daniel A. Keedy, Robert M. Immormino, Gary J. Kapral, Laura W. Murray, Jane S. Richardson, and David C. Richardson. 2010. "MolProbity: All-Atom Structure Validation for Macromolecular Crystallography." *Acta Crystallographica Section D: Biological Crystallography* 66 (1): 12–21. <https://doi.org/10.1107/S0907444909042073>.
- Davidson, Amy L., and Jue Chen. 2004. "ATP-Binding Cassette Transporters in Bacteria." *Annual Review of Biochemistry* 73 (1): 241–68. <https://doi.org/10.1146/annurev.biochem.73.011303.073626>.
- Dennhart, Nicole, Tamo Fukamizo, Ryszard Brzezinski, Marie-Ève Lacombe-Harvey, and Thomas Letzel. 2008. "Oligosaccharide Hydrolysis by Chitosanase Enzymes Monitored by Real-Time Electrospray Ionization–Mass Spectrometry." *Journal of Biotechnology* 134 (3): 253–60. <https://doi.org/https://doi.org/10.1016/j.jbiotec.2008.02.004>.
- Dennhart, Nicole, and Thomas Letzel. 2006. "Mass Spectrometric Real-Time Monitoring of Enzymatic Glycosidic Hydrolysis, Enzymatic Inhibition and Enzyme Complexes."

- Analytical and Bioanalytical Chemistry* 386 (3): 689–98. <https://doi.org/10.1007/s00216-006-0604-1>.
- Dennhart, Nicole, Linda M M Weigang, Maho Fujiwara, Tamo Fukamizo, Karen Skriver, and Thomas Letzel. 2009. “26kDa Endochitinase from Barley Seeds: Real-Time Monitoring of the Enzymatic Reaction and Substrate Binding Experiments Using Electrospray Ionization Mass Spectrometry.” *Journal of Biotechnology* 143 (4): 274–83. <https://doi.org/https://doi.org/10.1016/j.jbiotec.2009.08.003>.
- Dobro, Megan J, Linda A Melanson, Grant J Jensen, and Alasdair W McDowall. 2010. “Chapter Three - Plunge Freezing for Electron Cryomicroscopy.” In *Methods in Enzymology*, edited by Grant J Jensen, 481:63–82. Academic Press. [https://doi.org/https://doi.org/10.1016/S0076-6879\(10\)81003-1](https://doi.org/https://doi.org/10.1016/S0076-6879(10)81003-1).
- Doerr, Allison. 2017. “Cryo-Electron Tomography.” *Nature Methods* 14 (1): 34. <https://doi.org/10.1038/nmeth.4115>.
- Drenth, J., J. N. Jansonius, R. Koekoek, H. M. Swen, and B. G. Wolthers. 1968. “Structure of Papain.” *Nature* 218 (5145): 929–32. <https://doi.org/10.1038/218929a0>.
- Emsley, Paul, and Kevin Cowtan. 2004. “Coot: Model-Building Tools for Molecular Graphics.” *Acta Crystallographica Section D: Biological Crystallography* 60 (12 I): 2126–32. <https://doi.org/10.1107/S0907444904019158>.
- Fath, M J, and R Kolter. 1993. “ABC Transporters: Bacterial Exporters.” *Microbiol Rev* 57 (4): 995–1017. [https://doi.org/10.1016/S0005-2736\(99\)00158-3](https://doi.org/10.1016/S0005-2736(99)00158-3).
- Fischer, Hannes, Igor Polikarpov, and Aldo F. Craievich. 2009. “Average Protein Density Is a Molecular-Weight-Dependent Function.” *Protein Science* 13 (10): 2825–28. <https://doi.org/10.1110/ps.04688204>.
- Ford, Robert C., and Konstantinos Beis. 2019. “Learning the ABCs One at a Time: Structure and Mechanism of ABC Transporters.” *Biochemical Society Transactions* 47 (1): 23–36. <https://doi.org/10.1042/BST20180147>.
- Frank, Joachim, Michael Radermacher, Pawel Penczek, Jun Zhu, Yanhong Li, Mahieddine Ladjadj, and Ardean Leith. 1996. “SPIDER and WEB: Processing and Visualization of Images in 3D Electron Microscopy and Related Fields.” *Journal of Structural Biology* 116 (1): 190–99. <https://doi.org/10.1006/jsbi.1996.0030>.
- Gebhard, Susanne. 2012. “ABC Transporters of Antimicrobial Peptides in Firmicutes Bacteria - Phylogeny, Function and Regulation.” *Molecular Microbiology* 86 (6): 1295–1317. <https://doi.org/10.1111/mmi.12078>.
- Grant, Timothy, and Nikolaus Grigorieff. 2015. “Measuring the Optimal Exposure for Single Particle Cryo-EM Using a 2.6 Å Reconstruction of Rotavirus VP6.” *ELife* 4 (MAY): 1–19. <https://doi.org/10.7554/eLife.06980>.
- Havarstein, Leiv Sigve, Dzung Bao Diep, and Ingolf F. Nes. 1995. “A Family of Bacteriocin ABC Transporters Carry out Proteolytic Processing of Their Substrates Concomitant with Export.” *Molecular Microbiology* 16 (2): 229–40. <https://doi.org/10.1111/j.1365-2958.1995.tb02295.x>.
- Heuvel, Robert H H van den, Sara Gato, Cees Versluis, Pascal Gerbaux, Colin Kleanthous, and Albert J R Heck. 2005. “Real-Time Monitoring of Enzymatic DNA Hydrolysis by Electrospray Ionization Mass Spectrometry.” *Nucleic Acids Research* 33 (10): e96–e96. <https://doi.org/10.1093/nar/gni099>.

- Heymann, J. Bernard, and David M. Belnap. 2007. “Bsoft: Image Processing and Molecular Modeling for Electron Microscopy.” *Journal of Structural Biology* 157 (1): 3–18. <https://doi.org/10.1016/j.jsb.2006.06.006>.
- Hofmann, Susanne, Dovile Janulienė, Ahmad R Mehdipour, Christoph Thomas, Erich Stefan, Stefan Brüchert, Benedikt T Kuhn, et al. 2019. “Conformation Space of a Heterodimeric ABC Exporter under Turnover Conditions.” *Nature* 571 (7766): 580–83. <https://doi.org/10.1038/s41586-019-1391-0>.
- Husada, Florence, Kiran Bountra, Konstantinos Tassis, Marijn de Boer, Maria Romano, Sylvie Rebuffat, Konstantinos Beis, and Thorben Cordes. 2018. “Conformational Dynamics of the ABC Transporter McjD Seen by Single-Molecule FRET.” *The EMBO Journal* 37 (21): e100056. <https://doi.org/10.15252/embj.2018100056>.
- Husain, S. S., and G. Lowe. 1968. “Evidence for Histidine in the Active Site of Papain.” *Biochemical Journal* 108 (5): 855–59. <https://doi.org/10.1042/bj1080855>.
- Ishii, Seiji, Takato Yano, Akio Ebihara, Akihiro Okamoto, Miho Manzoku, and Hideyuki Hayashi. 2010. “Crystal Structure of the Peptidase Domain of Streptococcus ComA, a Bifunctional ATP-Binding Cassette Transporter Involved in the Quorum-Sensing Pathway.” *Journal of Biological Chemistry* 285 (14): 10777–85. <https://doi.org/10.1074/jbc.M109.093781>.
- Ishii, Seiji, Takato Yano, and Hideyuki Hayashi. 2006. “Expression and Characterization of the Peptidase Domain of Streptococcus Pneumoniae ComA, a Bifunctional ATP-Binding Cassette Transporter Involved in Quorum Sensing Pathway.” *Journal of Biological Chemistry* 281 (8): 4726–31. <https://doi.org/10.1074/jbc.M512516200>.
- Jardetzky, Oleg. 1966. “Simple Allosteric Model for Membrane Pumps.” *Nature* 211 (5052): 969–70. <https://doi.org/10.1038/211969a0>.
- Jin, Mi Sun, Michael L. Oldham, Qiuju Zhang, and Jue Chen. 2012. “Crystal Structure of the Multidrug Transporter P-Glycoprotein from *Caenorhabditis Elegans*.” *Nature* 490 (7421): 566–69. <https://doi.org/10.1038/nature11448>.
- Johnson, Zachary Lee, and Jue Chen. 2017. “Structural Basis of Substrate Recognition by the Multidrug Resistance Protein MRP1.” *Cell* 168 (6): 1075–1085.e9. <https://doi.org/10.1016/j.cell.2017.01.041>.
- . 2018. “ATP Binding Enables Substrate Release from Multidrug Resistance Protein 1.” *Cell* 172 (1–2): 81–89.e10. <https://doi.org/10.1016/j.cell.2017.12.005>.
- Kamphuis, I. G., K. H. Kalk, M. B.A. Swarte, and J Drenth. 1984. “Structure of Papain Refined at 1.65 Å Resolution.” *Journal of Molecular Biology* 179 (2): 233–56. [https://doi.org/10.1016/0022-2836\(84\)90467-4](https://doi.org/10.1016/0022-2836(84)90467-4).
- Kaplan, Mohammed, Georges Chreifi, Lauren Ann Metskas, Janine Liedtke, Cecily R Wood, Catherine M Oikonomou, William J Nicolas, et al. 2021. “In Situ Imaging of Bacterial Outer Membrane Projections and Associated Protein Complexes Using Electron Cryo-Tomography.” Edited by Bavesh D Kana. *ELife* 10: e73099. <https://doi.org/10.7554/eLife.73099>.
- Kaufmann, Christine, Johanna Grassmann, Dieter Treutter, and Thomas Letzel. 2014. “Utilization of Real-Time Electrospray Ionization Mass Spectrometry to Gain Further Insight into the Course of Nucleotide Degradation by Intestinal Alkaline Phosphatase.” *Rapid Communications in Mass Spectrometry* 28 (April). <https://doi.org/10.1002/rcm.6855>.

- Khare, Dheeraj, Michael L. Oldham, Cedric Orelle, Amy L. Davidson, and Jue Chen. 2009. “Alternating Access in Maltose Transporter Mediated by Rigid-Body Rotations.” *Molecular Cell* 33 (4): 528–36. <https://doi.org/10.1016/j.molcel.2009.01.035>.
- Kieuvongngam, Virapat, Paul Dominic B. Olinares, Anthony Palillo, Michael L. Oldham, Brian T. Chait, and Jue Chen. 2020. “Structural Basis of Substrate Recognition by a Polypeptide Processing and Secretion Transporter.” *ELife* 9 (January). <https://doi.org/10.7554/eLife.51492>.
- Kim, Youngjin, and Jue Chen. 2018. “Molecular Structure of Human P-Glycoprotein in the ATP-Bound, Outward-Facing Conformation.” *Science* 7389 (January): 1–10. <https://doi.org/10.1126/science.aar7389>.
- Kimanius, Dari, Bjorn O Forsberg, Sjors Scheres, and Erik Lindahl. 2016. “Accelerated Cryo-EM Structure Determination with Parallelisation Using GPUs in RELION-2.” *BioRxiv*, 059717. <https://doi.org/10.1101/059717>.
- Korkhov, Vladimir M., Samantha A. Mireku, and Kaspar P. Locher. 2012. “Structure of AMP-PNP-Bound Vitamin B 12 Transporter BtuCD-F.” *Nature* 490 (7420): 367–72. <https://doi.org/10.1038/nature11442>.
- Korkhov, Vladimir M., Samantha A. Mireku, Dmitry B. Veprintsev, and Kaspar P. Locher. 2014. “Structure of AMP-PNP-Bound BtuCD and Mechanism of ATP-Powered Vitamin B12 Transport by BtuCD-F.” *Nature Structural and Molecular Biology* 21 (12): 1097–99. <https://doi.org/10.1038/nsmb.2918>.
- Kotake, Yatsugu, Seiji Ishii, Takato Yano, Yoji Katsuoka, and Hideyuki Hayashi. 2008. “Substrate Recognition Mechanism of the Peptidase Domain of the Quorum-Sensing-Signal - Producing ABC Transporter ComA from *Streptococcus*.” *Biochemistry* 47 (8): 2531–38. <https://doi.org/10.1021/bi702253n>.
- Laganowsky, Arthur, Eamonn Reading, Jonathan T S Hopper, and Carol v Robinson. 2013. “Mass Spectrometry of Intact Membrane Protein Complexes.” *Nature Protocols* 8 (4): 639–51. <https://doi.org/10.1038/nprot.2013.024>.
- Lecher, Justin, Christian K W Schwarz, Matthias Stoldt, Sander H J Smits, Dieter Willbold, and Lutz Schmitt. 2012. “An RTX Transporter Tethers Its Unfolded Substrate during Secretion via a Unique N-Terminal Domain.” *Structure* 20 (10): 1778–87. <https://doi.org/10.1016/j.str.2012.08.005>.
- Lee, Edgar D, Wolfgang Mueck, Jack D Henion, and Thomas R Covey. 1989. “Real-Time Reaction Monitoring by Continuous-Introduction Ion-Spray Tandem Mass Spectrometry.” *Journal of the American Chemical Society* 111 (13): 4600–4604. <https://doi.org/10.1021/ja00195a012>.
- Lenders, Michael H H, Sven Reimann, Sander H J Smits, and Lutz Schmitt. 2013. “Molecular Insights into Type I Secretion Systems.” *Biological Chemistry* 394 (11): 1371–84. <https://doi.org/doi:10.1515/hsz-2013-0171>.
- Lin, David Yin-wei, Shuo Huang, and Jue Chen. 2015. “Crystal Structures of a Polypeptide Processing and Secretion Transporter.” *Nature* 523 (7561): 425–30. <https://doi.org/10.1038/nature14623>.
- Liu, Yanqing, Yue Liu, Lingli He, Yongfang Zhao, and Xuejun C Zhang. 2018. “Single-Molecule Fluorescence Studies on the Conformational Change of the ABC Transporter MsbA.” *Biophysics Reports* 4 (3): 153–65. <https://doi.org/10.1007/s41048-018-0057-z>.

- Locher Kaspar, Lee Allen, and Rees Douglas. 2002. “The E. Coli BtuCD Structure: A Framework for ABC Transporter Architecture and Mechanism.” *Science* 296 (5570): 1091–98. <https://doi.org/10.1126/science.1071142>.
- Locher, Kaspar P. 2016. “Mechanistic Diversity in ATP-Binding Cassette (ABC) Transporters.” *Nature Structural & Molecular Biology* 23 (6): 487–93. <https://doi.org/10.1038/nsmb.3216>.
- Manolaridis, Ioannis, Scott M. Jackson, Nicholas M.I. Taylor, Julia Kowal, Henning Stahlberg, and Kaspar P. Locher. 2018. “Cryo-EM Structures of a Human ABCG2 Mutant Trapped in ATP-Bound and Substrate-Bound States.” *Nature* 563 (7731): 426–30. <https://doi.org/10.1038/s41586-018-0680-3>.
- Marty, Michael T., Andrew J. Baldwin, Erik G. Marklund, Georg K.A. Hochberg, Justin L.P. Benesch, and Carol v. Robinson. 2015. “Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles.” *Analytical Chemistry* 87 (8): 4370–76. <https://doi.org/10.1021/acs.analchem.5b00140>.
- Ménard, Robert, Julie Carrière, Pierre Laflamme, Céline Plouffe, Henri E. Khouri, Andrew C. Storer, Julie Carrière, et al. 1991. “Contribution of the Glutamine 19 Side Chain to Transition-State Stabilization in the Oxyanion Hole of Papain.” *Biochemistry* 30 (37): 8924–28. <https://doi.org/10.1021/bi00101a002>.
- Mi, Wei, Yanyan Li, Sung Hwan Yoon, Robert K Ernst, Thomas Walz, and Maofu Liao. 2017. “Structural Basis of MsbA-Mediated Lipopolysaccharide Transport.” *Nature* 549 (7671): 233–37. <https://doi.org/10.1038/nature23649>.
- Michiels, Jan, Gunter Dirix, Jos Vanderleyden, and Chuanwu Xi. 2001. “Processing and Export of Peptide Pheromones and Bacteriocins in Gram-Negative Bacteria.” *Trends in Microbiology* 9 (4): 164–68. [https://doi.org/10.1016/S0966-842X\(01\)01979-5](https://doi.org/10.1016/S0966-842X(01)01979-5).
- Nicholls, Robert A., Marcus Fischer, Stuart Mcnicholas, and Garib N. Murshudov. 2014. “Conformation-Independent Structural Comparison of Macromolecules with ProSMART.” *Acta Crystallographica Section D: Biological Crystallography* 70 (9): 2487–99. <https://doi.org/10.1107/S1399004714016241>.
- Nishie, Mami, Makoto Sasaki, Jun Ichi Nagao, Takeshi Zendo, Jiro Nakayama, and Kenji Sonomoto. 2011. “Lantibiotic Transporter Requires Cooperative Functioning of the Peptidase Domain and the ATP Binding Domain.” *Journal of Biological Chemistry* 286 (13): 11163–69. <https://doi.org/10.1074/jbc.M110.212704>.
- Nöll, Anne, Christoph Thomas, Valentina Herbring, Tina Zollmann, Katja Barth, Ahmad Reza Mehdipour, Thomas M Tomasiak, et al. 2017. “Crystal Structure and Mechanistic Basis of a Functional Homolog of the Antigen Transporter TAP.” *Proceedings of the National Academy of Sciences* 114 (4): E438. <https://doi.org/10.1073/pnas.1620009114>.
- Oldham, M. L., and J. Chen. 2011. “Snapshots of the Maltose Transporter during ATP Hydrolysis.” *Proceedings of the National Academy of Sciences* 108 (37): 15152–56. <https://doi.org/10.1073/pnas.1108858108>.
- Olinares, Paul Dominic B, and Brian T Chait. 2020. “Native Mass Spectrometry Analysis of Affinity-Captured Endogenous Yeast RNA Exosome Complexes.” In *The Eukaryotic RNA Exosome: Methods and Protocols*, edited by John LaCava and Štěpánka Vaňáčková, 357–82. New York, NY: Springer New York. [https://doi.org/10.1007/978-1-4939-9822-7\\_17](https://doi.org/10.1007/978-1-4939-9822-7_17).
- Pettersen, Eric F., Thomas D. Goddard, Conrad C. Huang, Gregory S. Couch, Daniel M. Greenblatt, Elaine C. Meng, and Thomas E. Ferrin. 2004. “UCSF Chimera - A Visualization System for Exploratory Research and Analysis.” *Journal of Computational Chemistry* 25 (13): 1605–12. <https://doi.org/10.1002/jcc.20084>.

- Punjani, Ali, and David J Fleet. 2021. “3D Variability Analysis: Resolving Continuous Flexibility and Discrete Heterogeneity from Single Particle Cryo-EM.” *Journal of Structural Biology* 213 (2): 107702.  
<https://doi.org/https://doi.org/10.1016/j.jsb.2021.107702>.
- Punjani, Ali, John L. Rubinstein, David J. Fleet, and Marcus A. Brubaker. 2017. “CryoSPARC: Algorithms for Rapid Unsupervised Cryo-EM Structure Determination.” *Nature Methods* 14 (3): 290–96. <https://doi.org/10.1038/nmeth.4169>.
- Rahman, Suhaila, and Hassane S Mchaourab. 2020. “ATP-Dependent Interactions of a Cargo Protein with the Transmembrane Domain of a Polypeptide Processing and Secretion ABC Transporter.” *Journal of Biological Chemistry* 295 (43): 14678–85.  
<https://doi.org/10.1074/jbc.RA120.014934>.
- Reading, Eamonn, Ildir Liko, Timothy M. Allison, Justin L.P. Benesch, Arthur Laganowsky, and Carol v. Robinson. 2015. “The Role of the Detergent Micelle in Preserving the Structure of Membrane Proteins in the Gas Phase.” *Angewandte Chemie - International Edition* 54 (15): 4577–81. <https://doi.org/10.1002/anie.201411622>.
- Rees, Douglas C, Eric Johnson, and Oded Lewinson. 2009. “ABC Transporters: The Power to Change.” *Nature Reviews Molecular Cell Biology* 10 (3): 218–27.  
<https://doi.org/10.1038/nrm2646>.
- Reid, Deseree J., Jessica M. Diesing, Matthew A. Miller, Scott M. Perry, Jessica A. Wales, William R. Montfort, and Michael T. Marty. 2019. “MetaUniDec: High-Throughput Deconvolution of Native Mass Spectra.” *Journal of the American Society for Mass Spectrometry* 30 (1): 118–27. <https://doi.org/10.1007/s13361-018-1951-9>.
- Rohou, Alexis, and Nikolaus Grigorieff. 2015. “CTFFIND4: Fast and Accurate Defocus Estimation from Electron Micrographs.” *Journal of Structural Biology* 192 (2): 216–21.  
<https://doi.org/10.1016/j.jsb.2015.08.008>.
- Rose, Rebecca J, Eugen Damoc, Eduard Denisov, Alexander Makarov, and Albert J R Heck. 2012. “High-Sensitivity Orbitrap Mass Analysis of Intact Macromolecular Assemblies.” *Nature Methods* 9 (11): 1084–86. <https://doi.org/10.1038/nmeth.2208>.
- Rubinstein, John L., and Marcus A. Brubaker. 2015. “Alignment of Cryo-EM Movies of Individual Particles by Optimization of Image Translations.” *Journal of Structural Biology* 192 (2): 188–95. <https://doi.org/10.1016/j.jsb.2015.08.007>.
- Schechter, Israel, and Arieh Berger. 1967. “On the Size of the Active Site in Proteases. I. Papain.” *Biochemical and Biophysical Research Communications* 27 (2): 157–62.  
[https://doi.org/10.1016/S0006-291X\(67\)80055-X](https://doi.org/10.1016/S0006-291X(67)80055-X).
- Shaikh, Tanvir R., Haixiao Gao, William T. Baxter, Francisco J. Asturias, Nicolas Boisset, Ardean Leith, and Joachim Frank. 2008. “Spider Image Processing for Single-Particle Reconstruction of Biological Macromolecules from Electron Micrographs.” *Nature Protocols* 3 (12): 1941–74. <https://doi.org/10.1038/nprot.2008.156>.
- Srikant, Sriram, and Rachele Gaudet. 2019. “Mechanics and Pharmacology of Substrate Selection and Transport by Eukaryotic ABC Exporters.” *Nature Structural & Molecular Biology*, 1–6. <https://doi.org/10.1038/s41594-019-0280-4>.
- Suloway, Christian, Jian Shi, Anchi Cheng, James Pulokas, Bridget Carragher, Clinton S. Potter, Shawn Q. Zheng, David A. Agard, and Grant J. Jensen. 2009. “Fully Automated, Sequential Tilt-Series Acquisition with Leginon.” *Journal of Structural Biology* 167 (1): 11–18.  
<https://doi.org/10.1016/j.jsb.2009.03.019>.

- Tang, Xiaodi, Shenghai Chang, Qinghua Luo, Zhengyu Zhang, Wen Qiao, Caihuang Xu, Changbin Zhang, et al. 2019. "Cryo-EM Structures of Lipopolysaccharide Transporter LptB2FGC in Lipopolysaccharide or AMP-PNP-Bound States Reveal Its Transport Mechanism." *Nature Communications* 10 (1): 4175. <https://doi.org/10.1038/s41467-019-11977-1>.
- Taylor, Nicholas M I, Ioannis Manolaridis, Scott M Jackson, Julia Kowal, Henning Stahlberg, and Kaspar P Locher. 2017. "Structure of the Human Multidrug Transporter ABCG2." *Nature* 546 (7659): 504–9. <https://doi.org/10.1038/nature22345>.
- Terwilliger, Thomas C, Steven J Ludtke, Randy J Read, Paul D Adams, and Pavel v Afonine. 2020. "Improvement of Cryo-EM Maps by Density Modification." *Nature Methods* 17 (9): 923–27. <https://doi.org/10.1038/s41592-020-0914-9>.
- Terwilliger, Thomas C, Oleg v Sobolev, Pavel v Afonine, Paul D Adams, and Randy J Read. 2020. "Density Modification of Cryo-EM Maps." *Acta Crystallographica Section D* 76 (10): 912–25. <https://doi.org/10.1107/S205979832001061X>.
- Thomas, Christoph, and Robert Tampé. 2020. "Structural and Mechanistic Principles of ABC Transporters." *Annual Review of Biochemistry* 89 (1): 605–36. <https://doi.org/10.1146/annurev-biochem-011520-105201>.
- Thomas, Sabrina, I. Barry Holland, and Lutz Schmitt. 2014. "The Type 1 Secretion Pathway - The Hemolysin System and Beyond." *Biochimica et Biophysica Acta - Molecular Cell Research* 1843 (8): 1629–41. <https://doi.org/10.1016/j.bbamcr.2013.09.017>.
- Vernet, Thierry, Daniel C. Tessier, Jean Chatellier, Céline Plouffe, Tak Sing Lee, David Y. Thomas, Andrew C. Storer, and Robert Ménard. 1995. "Structural and Functional Roles of Asparagine 175 in the Cysteine Protease Papain." *Journal of Biological Chemistry* 270 (28): 16645–52. <https://doi.org/10.1074/jbc.270.28.16645>.
- Voss, Neil R., and Mark Gerstein. 2010. "3V: Cavity, Channel and Cleft Volume Calculator and Extractor." *Nucleic Acids Research* 38 (SUPPL. 2): 555–62. <https://doi.org/10.1093/nar/gkq395>.
- Wang, Ling, Zachary Lee Johnson, Michael R Wasserman, Jesper Levring, Jue Chen, and Shixin Liu. 2020. "Characterization of the Kinetic Cycle of an ABC Transporter by Single-Molecule and Cryo-EM Analyses." Edited by László Csanády, Olga Boudker, László Csanády, Markus A Seeger, and Gergely Szakacs. *ELife* 9: e56451. <https://doi.org/10.7554/eLife.56451>.
- Williams, Nicholas H. 2000. "Magnesium Ion Catalyzed ATP Hydrolysis." *Journal of the American Chemical Society* 122 (48): 12023–24. <https://doi.org/10.1021/ja0013374>.
- Wu, Kai Hui, and Phang C. Tai. 2004. "Cys32 and His105 Ate the Critical Residues for the Calcium-Dependent Cysteine Proteolytic Activity of CvaB, an ATP-Binding Cassette Transporter." *Journal of Biological Chemistry* 279 (2): 901–9. <https://doi.org/10.1074/jbc.M308296200>.
- Wu, Kai-hui, Ying-hsin Hsieh, and Phang C Tai. 2012. "Mutational Analysis of CvaB, an ABC Transporter Involved in the Secretion of Active Colicin V" 7 (4). <https://doi.org/10.1371/journal.pone.0035382>.
- Yu, Qin, Dongchun Ni, Julia Kowal, Ioannis Manolaridis, Scott M Jackson, Henning Stahlberg, and Kaspar P Locher. 2021. "Structures of ABCG2 under Turnover Conditions Reveal a Key Step in the Drug Transport Mechanism." *Nature Communications* 12 (1): 4376. <https://doi.org/10.1038/s41467-021-24651-2>.

- Yu, Zhan, Lee Chuin Chen, Mridul Kanti Mandal, Hiroshi Nonami, Rosa Erra-Balsells, and Kenzo Hiraoka. 2012. "Online Electrospray Ionization Mass Spectrometric Monitoring of Protease-Catalyzed Reactions in Real Time." *Journal of the American Society for Mass Spectrometry* 23 (4): 728–35. <https://doi.org/10.1007/s13361-011-0323-5>.
- Zhang, Kai. 2016. "Gctf: Real-Time CTF Determination and Correction." *Journal of Structural Biology* 193 (1): 1–12. <https://doi.org/10.1016/j.jsb.2015.11.003>.
- Zheng, Shawn Q., Eugene Palovcak, Jean Paul Armache, Kliment A. Verba, Yifan Cheng, and David A. Agard. 2017. "MotionCor2: Anisotropic Correction of Beam-Induced Motion for Improved Cryo-Electron Microscopy." *Nature Methods* 14 (4): 331–32. <https://doi.org/10.1038/nmeth.4193>.
- Zi Tan, Yong, Philip R. Baldwin, Joseph H. Davis, James R. Williamson, Clinton S. Potter, Bridget Carragher, and Dmitry Lyumkis. 2017. "Addressing Preferred Specimen Orientation in Single-Particle Cryo-EM through Tilting." *Nature Methods* 14 (8): 793–96. <https://doi.org/10.1038/nmeth.4347>.
- Zivanov, Jasenko, Takanori Nakane, Björn O. Forsberg, Dari Kimanius, Wim J.H. Hagen, Erik Lindahl, and Sjors H.W. Scheres. 2018. "New Tools for Automated High-Resolution Cryo-EM Structure Determination in RELION-3." *ELife* 7: 1–22. <https://doi.org/10.7554/eLife.42166>.
- Zivanov, Jasenko, Takanori Nakane, and Sjors H.W. Scheres. 2019. "A Bayesian Approach to Beam-Induced Motion Correction in Cryo-EM Single-Particle Analysis." *IUCrJ* 6: 5–17. <https://doi.org/10.1107/S205225251801463X>.