

# Structural basis for the inhibition of a bacterial NLR and insights into its activation



Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Yuang Wu

aus Huzhou, China

Köln, Juli 2022



Die in diesem Beitrag beschriebenen Arbeiten wurden unter der Leitung von Prof. Dr. Jijie Chai am Max-Planck-Institut für Züchtungsforschung (Selbständige Forschungsgruppe Strukturbiologie) durchgeführt.

Berichtersteller: Prof. Dr. Jijie Chai  
Prof. Dr. Jane Parker  
Prüfungsvorsitz: Prof. Dr. Alga Zuccaro

Tag der Disputation:



MAX-PLANCK-GESELLSCHAFT



# Abstract

To cope with a constantly changing environment, living organisms acquire different signaling pathways in response to internal and external stimuli. The STAND (signal transduction ATPases with numerous domains) is a family of intracellular signaling proteins widely exist across kingdoms. In eukaryotes, STAND proteins have a key role in innate immunity and act as sensors for the detection of microbes, defects in STAND signaling are often found as the underlying cause of cancer and other human diseases. Bacterial STAND proteins mainly function as transcription regulators, which monitor the availability of nutrients and facilitate their efficient utilization. MalT is a transcription activator of the maltose operon in *E. coli* (*Escherichia coli*) and has all the hallmarks of STAND family proteins. MalT transcriptionally controls the activity of genes involved in maltose/maltodextrin uptake and catabolism, while its own activity is governed by inducer maltotriose and repressor proteins such as MalK, MalY, and Aes. MalT harbors a NOD (nucleotide-binding and oligomerization) module at the protein N-terminal (amino-terminal) end, followed by a SUPR-type sensor domain and a DNA-binding domain at the C-terminal (carboxyl terminal) end. In this study, the structure of an inhibitory complex containing MalT and MalY was solved by cryo-EM (cryo-electron microscopy), revealing the mechanisms of MalT inhibition and activation. Inhibition of MalT was shown to be mediated by inter- and intramolecular interactions. The very N-terminal region of MalT makes contacts with both NBD (nucleotide-binding domain) and WHD (wing-helix domain), keeping the NOD module closed. This resting state is further maintained by MalY, which interacts with MalT via one oligomerization surface on the NBD. Upon maltotriose binding, domain remodeling of MalT occurs and allosterically disengages the inhibitory contacts, exposing another oligomerization surface composed of structural elements from the N-terminal region, leading to nucleotide exchange and MalT activation. Ligand-binding together with oligomerization stabilizes the C-terminal domains of MalT for DNA-binding and transcription initiation at MalT-dependent promoters. The presented results suggest a common strategy for protein inhibition adopted by STAND members during the course of evolution and provide insights into transcription regulation by STAND proteins, demonstrating the functional versatility of this protein family with a relatively conserved domain architecture.



# Zusammenfassung

Lebende Organismen haben verschiedene Signalwege als Reaktion auf interne und externe Stimuli entwickelt, um in einer sich stetig ändernden Umwelt zurechtzukommen. Die STAND-Proteine (signal transduction ATPases with numerous domains) sind eine Familie von intrazellulären Signalproteinen, die in allen Reichen weit verbreitet sind. In Eukaryonten spielen STAND-Proteine eine Schlüsselrolle bei der angeborenen Immunität und fungieren als Sensoren für die Erkennung von Mikroben. Defekte in der STAND-Signalübertragung sind häufig eine Ursache von Krebs und anderen menschliche Krankheiten. Bakterielle STAND-Proteine fungieren hauptsächlich als Transkriptionsregulatoren, die die Verfügbarkeit von Nährstoffen überwachen und deren effiziente Nutzung regulieren. MalT ist ein Transkriptionsaktivator des Maltose-Operons in *E. coli* (*Escherichia coli*) und weist alle Merkmale von Proteinen der STAND-Familie auf. MalT steuert die Transkription von Genen, die an der Aufnahme und Stoffwechsel von Maltose/Maltodextrin beteiligt sind, während seine eigene Aktivität durch den Induktor Maltotriose und Repressorproteine wie MalK, MalY und Aes gesteuert wird. MalT verfügt über ein NOD-Modul (Nukleotidbindungs- und Oligomerisierungsmodul) am N-terminalen (aminoterminalen) Ende des Proteins, das von einer Sensordomäne vom SUPR-Typ und einer DNA-Bindungsdomäne am C-terminalen (carboxylterminalen) Ende gefolgt wird. In dieser Arbeit wurde die Struktur eines inhibitorischen Komplexes aus MalT und MalY durch Kryo-EM (Kryo-Elektronenmikroskopie) gelöst und die Mechanismen der MalT-Hemmung bzw. -Aktivierung entschlüsselt. Die Hemmung von MalT wird durch inter- und intramolekulare Wechselwirkungen vermittelt. Die ersten 20 N-terminalen Aminosäuren stellen den Kontakt sowohl mit der NBD (Nukleotidbindungsdomäne) als auch mit der WHD (Flügelhelixdomäne) her und hält das NOD-Modul geschlossen. Dieser Ruhezustand wird außerdem durch MalY aufrechterhalten, das mit MalT über eine Oligomerisierungsfläche an der NBD interagiert. Durch die Bindung von Maltotriose kommt es zu einem Domänenumbau von MalT, der die allosterisch inhibierenden Kontakte aufhebt und eine weitere Oligomerisierungsfläche freilegt, die aus Strukturelementen der N-terminalen Region besteht. Dieser Domänenumbau führt zum Nukleotidaustausch und zur Aktivierung von MalT. Die Ligandenbindung (Maltotriose) zusammen mit der Oligomerisierung stabilisiert die C-terminalen Domänen von MalT für die DNA-Bindung und Transkriptionsinitiation an MalT-abhängigen Promotoren. Die

vorgestellten Ergebnisse zeigen eine allgemeine Strategie der Proteininhibition, die von STAND-Mitgliedern im Laufe der Evolution etabliert wurde. Diese Einblicke in die Transkriptionsregulierung durch STAND-Proteine zeigt, welche funktionelle Vielseitigkeit mit einer relativ konservierten Domänenarchitektur dieser Proteinfamilie erreicht wird.



# Contents

Abstract .....	I
Zusammenfassung.....	III
1. Introduction.....	1
1.1 Preface.....	1
1.2 STAND proteins across kingdoms.....	1
1.2.1 Structural studies of STAND proteins in animals.....	1
1.2.2 Structural studies of STAND proteins in plants .....	5
1.2.3 Structural studies of STAND proteins in prokaryotes.....	8
1.3 The maltose system in <i>E. coli</i> ( <i>Escherichia coli</i> ).....	9
1.3.1 Transcription regulation in <i>E. coli</i> and the maltose regulon .....	9
1.3.2 Current model for the regulation of the maltose regulon .....	10
1.3.3 MalK, the ABC subunit of the maltose transporter regulates MalT activity .....	11
1.3.4 MalY as a repressor for the maltose regulon .....	12
1.3.5 Aes as a repressor for the maltose regulon .....	13
1.3.6 MalT, the central regulator of the maltose regulon.....	14
2. Aims of this study .....	16
3. Materials and methods .....	17
4. Results .....	35
4.1 Reconstitution of the MalT-MalY complex.....	35
4.1.1 Protein expression and purification .....	35
4.1.2 MalT interacts with MalY <i>in vitro</i> .....	36
4.1.3 Cryo-EM analysis of the MalT-MalY complex.....	37
4.1.4 The overall structure of the MalT-MalY complex.....	38
4.2 The molecular mechanism of MalT repression by MalY .....	40
4.2.1 The MalY-MalT interaction interface in detail .....	40
4.2.2 MalY inhibits MalT oligomerization <i>in vitro</i> .....	41
4.2.3 Structural alignment of the MalT-MalY complex with known NLRs .....	42
4.2.4 MalT interface mutants are defective in oligomerization.....	43
4.3 Autoinhibition of MalT and insights into its activation .....	45
4.3.1 Autoinhibition of MalT is maintained via interdomain interactions.....	45
4.3.2 MalT N-terminal region contributes to protein oligomerization .....	48
4.3.3 Low-order oligomerization of MalT is sufficient for DNA-binding .....	49
4.3.4 MalT <sup>NBD-HD</sup> dimer modeled in its active form.....	50
4.4 Reconstitution of MalT in its active form .....	51

4.4.1 Preliminary cryo-EM study of oligomeric MalT .....	51
4.4.2 Activation of MalT stabilizes its C-terminal domain for binding DNA.....	52
4.4.3 An active MalT dimer .....	53
4.4.4 Mapping of the maltotriose-binding site based on a predicted MalT structure.....	54
5. Discussion .....	57
5.1 Inhibition of an NLR by blocking its oligomerization.....	57
5.2 A dual role for the very N-terminal region of MalT in its activation .....	59
5.3 Mechanism of ligand-induced MalT DNA-binding .....	60
5.4 MalT oligomeric states vs its transcriptional activity .....	62
5.5 MalT stability vs its protein conformation .....	64
5.6 A model for MalT regulation and comparison with other STAND proteins .....	65
6. References .....	68
List of abbreviations .....	75
Acknowledgement.....	78
Erklärung zur Dissertation .....	79
Curriculum vitae .....	80



# 1. Introduction

## 1.1 Preface

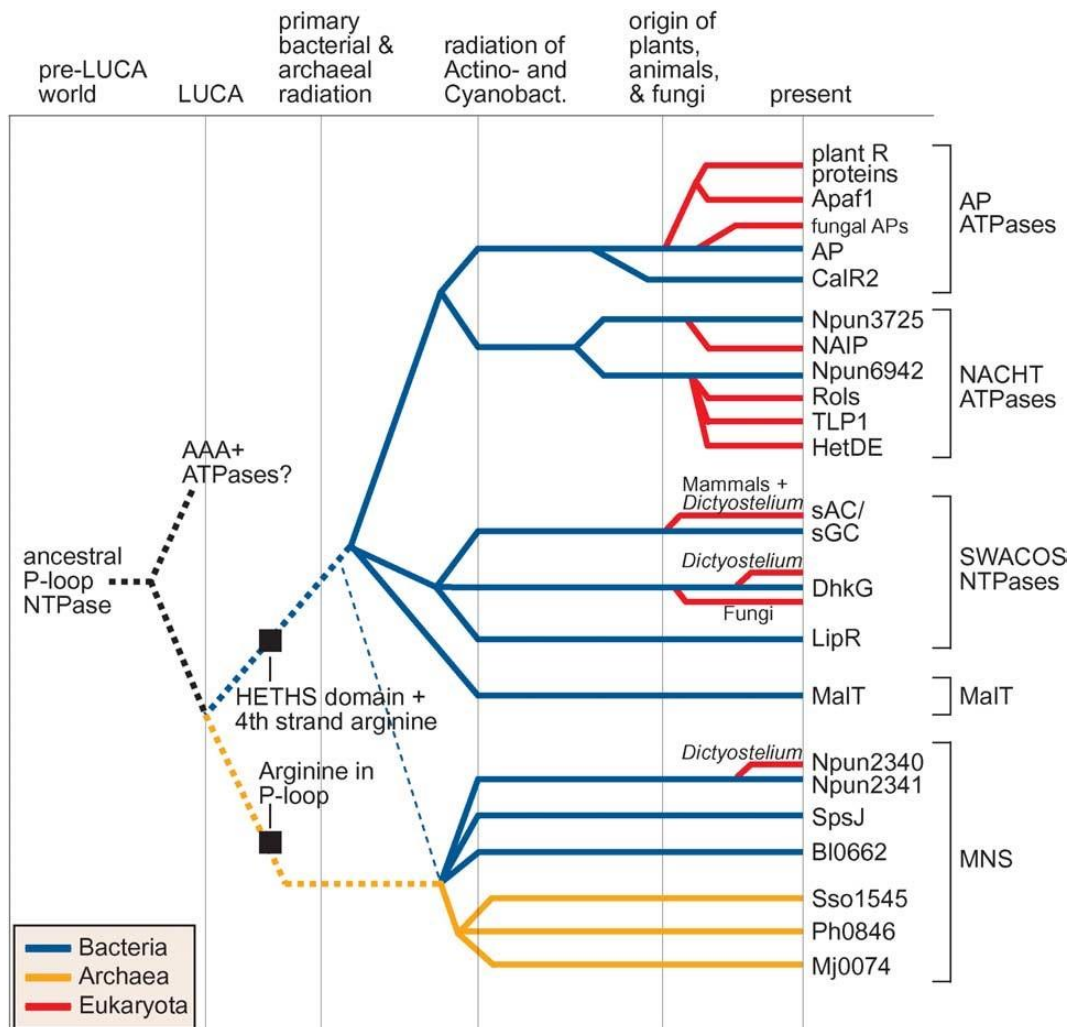
Just as sailors use barometers to predict the weather and navigate, so life on earth monitors its surroundings in order to survive. To cope with a constantly changing environment, it is necessary for both eukaryotes and prokaryotes to translate extracellular signals precisely into specific responses. Cells integrate these signals through different machineries, and receptor proteins or secondary messengers participate in this process have been identified over the years (Hunter, 2000). The STAND (signal transduction ATPases with numerous domains) is a protein family involved in a vast array of biological processes, including transcriptional regulation, programmed cell death, and innate immune response (Leipe et al., 2004). As an important paradigm in intracellular signaling, members of the STAND class share conserved features in their domain architecture, whereby an effector domain mediating the interaction with downstream signaling components, a central module for nucleotide-binding and oligomerization, and a sensor domain responsible for ligand-recognition are integrated as a single unit. Binding of ligand to the C-terminal of protein induces conformational changes that trigger protein activation, and via the effector domain signal can be transmitted towards downstream cascades (Jones et al., 2016; Kaparakis et al., 2007; Meunier and Broz, 2017; Yin et al., 2015).

## 1.2 STAND proteins across kingdoms

### 1.2.1 Structural studies of STAND proteins in animals

Based on the evolutionary history and sequence conservation, STAND proteins are classified into five different clades (the AP-ATPase clade, the NACHT clade, the SWACOS clade, the MalT clade, and the MNS clade) (Figure 1.1) (Leipe et al., 2004). Among them, the AP-ATPase clade and the NACHT clade containing many key players controlling cell death and host defense have

been most extensively studied, from which major mechanistic insights into the regulation and function of STAND proteins were gained.



**Figure 1.1 Inferred evolutionary history of STAND NTPases.** This figure shows relative temporal epochs and marks major evolutionary events by vertical lines. The evolution of the protein-coding gene is traced with horizontal colored lines. A broken line indicates uncertainty with respect to the exact point of origin. Taken from Leipe et al., 2004.

Amongst the early discovered STAND members, Apaf-1 (APOPTOTIC PROTEASE ACTIVATING FACTOR 1) and its nematode homolog CED-4 (CELL DEATH PROTEIN 4) are two essential regulators of developmental and p53-dependent programmed cell death, both of them belong to the AP-ATPase clade (Li et al., 1997; Yuan and Horvitz, 1992). Apaf-1 is the main component of the mammalian apoptosome. When an apoptotic stimulus activates a cell, Apaf-1 forms a heptameric complex in the presence of *cytochrome c* and ATP (adenosine triphosphate)/dATP (deoxyadenosine triphosphate), which allows proximity of procaspase-9 and its enzymatic self-activation. Mature caspase-9 then activates caspase-3 and caspase-7, committing the cell

to apoptosis (Chai and Shi, 2014). Structural analysis of Apaf-1 reveals that the protein exists in two different forms. Prior to *cytochrome c* binding, the N-terminal CARD (caspase recruitment domain) packs against NBD (nucleotide-binding domain), HD1 (helical domain 1), and WHD (winged-helix domain). The interactions between these four adjoining domains are strengthened by an ADP (adenosine diphosphate) molecule as the organizing center, locking Apaf-1 in an inactive conformation. A WD40 repeat domain at the C-terminal end further stabilizes this autoinhibited conformation by interacting with NBD and HD2 (helical domain 2) via hydrogen bonds (Riedl et al., 2005). Assembly of the apoptosome is primed by *cytochrome c* binding, followed by nucleotide exchange that drives conformational changes and leads to relief of autoinhibition (Zhou et al., 2015). The apoptosome has a central hub consisting of NBD, HD1, and WHD, while WD40 repeat forms spokes connected to the central hub through HD2, exhibiting a wheel-shaped appearance of 7-fold symmetry. The CARD at the N-terminal end is flexible and would be stabilized upon homotypic CARD-CARD interaction with procaspase-9. Different from its mammalian homolog, CED-4 does not have a ligand-binding domain and maintains a constitutively active conformation even in the resting form, of which an ATP-bound CED-4 dimer is sequestered to the mitochondria-bound CED-9 (CELL DEATH ABNORMALITY GENE 9) until its release by EGL-1 (EGG-LAYING DEFECTIVE) (Yan et al., 2005). CED-9 has a higher binding affinity for EGL-1 compared to CED-4, therefore EGL-1 binding would induce allosteric changes of CED-9 and result in the dissociation of the CED4-CED9 complex. After the displacement of CED-9, free CED-4 forms a funnel-shaped apoptosome arranged as a tetramer of four asymmetric dimers, and associates with CED-3 (CELL DEATH PROTEIN 3) to form a stable holoenzyme (Qi et al., 2010; Yan et al., 2005).

In addition to Apaf-1, several mammalian STAND proteins from the NACHT clade that function as NLRs (nod-like receptors) have been characterized biochemically, exemplified by NLRC4 (NLR FAMILY CARD DOMAIN-CONTAINING PROTEIN 4), NLRP1 (NLR FAMILY PYRIN DOMAIN-CONTAINING 1), and NLRP3 (NLR FAMILY PYRIN DOMAIN CONTAINING 3) (Andreeva et al., 2021; Hollingsworth et al., 2021; Hu et al., 2013; Hu et al., 2015; Huang et al., 2021; Sharif et al., 2019). These cytosolic sensors reside in an inactive conformation under homeostatic conditions, and can be activated upon microbial invasion by a PAMP (pathogen-associated molecular pattern) or DAMP (host-derived danger-associated molecular pattern), leading to the formation of multiprotein complexes termed inflammasome. Once the inflammasome is formed, it recruits procaspase-1 either directly or through ASC (APOPTOSIS-ASSOCIATED

SPECK-LIKE PROTEIN CONTAINING A CARD)-dependent machinery. Procaspase-1 would then undergo proximity-induced autoproteolysis, generating an active enzyme that cleaves and activates inflammatory cytokines and GSDMD (GASDERMIN D) which induces pyroptotic cell death (Broz and Dixit, 2016). NLRC4 contains an N-terminal CARD domain, a central NOD (nucleotide-binding and oligomerization domain) module, and an LRR (leucine-rich repeat) domain at the C-terminal end. The inactive NLRC4 is monomeric and harbors an ADP molecule, with an overall shape mimicking an inverted question mark. This conformation is maintained synergistically by interdomain interactions within the NOD and sequestration of the NOD by LRR (Hu et al., 2013). Unlike Apaf-1, NLRC4 does not directly recognize the ligand but via NAIP (NLR FAMILY APOPTOSIS INHIBITORY PROTEIN) family proteins, which also belong to the NACHT clade. In the NAIP-NLRC4 inflammasome, one NAIP molecule is incorporated in a substoichiometric manner and confers ligand specificity. In macrophages infected by bacterial pathogens carrying flagellin or components of T3SS (type III secretion system), flagellin-bound NAIP5/6 or PrgJ-bound NAIP2 would interact with and activate an NLRC4 through one oligomerization surface formed by structural elements from NBD and HD1 (hereafter called receptor surface), triggering structural rearrangement and exposing the other oligomerization surface responsible for initiating NLRC4 autoactivation (hereafter called catalytic surface). The activated NLRC4 then interacts with another inactive NLRC4, leading to its activation and progressive assembly of a wheel-like inflammasome structure (Hu et al., 2015; Yang et al., 2018).

NLRP1 was the first NLR discovered to form an inflammasome (Chavarría-Smith and Vance, 2015). Human NLRP1 has an N-terminal PYD (pyrin domain) followed by NOD and LRR, with a C-terminal extension containing a FIIND (function-to-find domain) and a CARD. FIIND consists of ZU5 (as initially found in ZO-1 and UNC5) and UPA (conserved in UNC5, PIDD and Ankirins) subdomains and undergoes post-translational autoproteolysis after ZU5, which is required for NLRP1 inflammasome activation (Finger et al., 2012; Xu et al., 2019). Recently, progress has been made in elucidating the molecular basis of DPP9 (DIPEPTIDYL PEPTIDASES 9)-mediated NLRP1 inhibition, where two distinct NLRP1 molecules are bound by one DPP9 (Hollingsworth et al., 2021; Huang et al., 2021). Three interaction surfaces are identified in the complex structure, including two mediating the interactions between DPP9 and different domains of each NLRP1 molecule (ZU5 and UPA, respectively), and one homodimerization surface between two UPA domains from NLRP1. In resting cells, two NLRP1 polypeptide chains

generated by autoproteolysis are noncovalently associated with one another. A certain fraction of NLRP1 molecules is additionally subjected to N-terminal degradation, generating a pool of C-terminal UPA-CARD fragments. Full-length NLRP1 and the UPA-CARD fragment from a second NLRP1 cooperatively bind to DPP9, resulting a ternary complex that keeps the inflammasome-forming UPA-CARD fragments sequestered (Bauernfried and Hornung, 2021).

NLRP3 is another inflammasome-forming intracellular sensor that can detect a wide range of microbial patterns, endogenous danger signals, and environmental stimuli. Numerous researches have been carried out to investigate the regulation of NLRP3, and recent studies have provided new knowledge about the structural mechanisms of NLRP3 autoinhibition and activation (Andreeva et al., 2021; Bauernfeind et al., 2009; Broz and Dixit, 2016; Elliott and Sutterwala, 2015; Lamkanfi and Dixit, 2014; Schroder and Tschopp, 2010; Sharif et al., 2019). Like other STAND proteins, NLRP3 has a typical structural domain arrangement consisting of an N-terminal PYD, a central NOD, and a C-terminal LRR. Interestingly, the inactive NLRP3 appears to be a double ring cage structure containing 12-16 monomers, formed mainly by interactions between the curved LRRs, while the PYDs are shielded within the cavity inside two rings to avoid premature activation (Andreeva et al., 2021). The oligomeric inactive complex is predominantly localized to the membrane which serves as a scaffolding platform. Upon stimulation, the complex is transported to the MTOC (microtubule-organizing center) via dTGN (trans-Golgi network dispersion), where it gets disrupted following NEK7 (NIMA-RELATED KINASE 7) binding, leading to the oligomerization of NLRP3 into an active form of the inflammasome complex, which is essential for ASC speck formation and caspase 1 activation (Sharif et al., 2019).

### 1.2.2 Structural studies of STAND proteins in plants

Unlike animals that have an adaptive immune system, plants solely depend on innate immunity to detect and restrain the growth of pathogenic invaders, and have developed an elaborate repertoire of innate immune receptors during evolution, including plasma membrane-localized PRRs (pattern recognition receptors) and cytosolic NLRs, of which the latter belongs to the AP-ATPase clade of STAND class proteins (Leipe et al., 2004; Saur et al., 2021). Besides a central NOD module and C-terminal LRR, the N-terminal domain of plant NLRs



mainly has two different types, which are CC (coiled-coil) and TIR (toll/interleukin-1 receptor). Recent research in structural studies of CC-type and TIR-type NLRs have greatly advanced our understanding of ligand-sensing and post-activation mechanisms of NLRs in plants.

ZAR1 (HOPZ-ACTIVATED RESISTANCE 1) is a CC-domain containing NLR from the model plant *Arabidopsis thaliana*, and is the first plant NLR reported to form an oligomeric structure. In response to *Xanthomonas campestris* effector AvrAC, a uridylyl transferase, the RLCK (RECEPTOR-LIKE CYTOPLASMIC KINASES) PBL2 (PBS1-LIKE PROTEIN 2) acts as a decoy which can be uridylylated by the effector, resulting in a modified protein (PBL2<sup>UMP</sup>) that binds to a preformed complex of ZAR1 and RKS1 (RESISTANCE RELATED KINASE 1), forming an active complex termed resistosome (Wang et al., 2019a). In the ZAR1-RKS1 complex structure, the autoinhibition of ZAR1 is maintained primarily through intramolecular interactions, with the CC domain remaining in the inactive state through contacts with LRR, HD1, and WHD. The LRR of ZAR1 is not only involved in the sequestration of the N-terminal domains, a function observed for the LRR from both Apaf-1 and NLRC4, but also mediates the interaction between ZAR1 and RKS1. Binding of PBL2<sup>UMP</sup> to RKS1 triggers steric clash between RKS1 and the NBD of ZAR1, allosterically inducing the release of ADP and stimulating the activation of ZAR1 (Wang et al., 2019b). In the presence of dATP or ATP, the primed ZAR1-RKS1-PBL2<sup>UMP</sup> complex further oligomerizes and assembles into a pentameric resistosome, with all the structural domains of ZAR1 being involved in resistosome formation. Of note, the CC domain of ZAR1 not only contributes to oligomerization, but also confers a functional role. Once activated, the very N-terminal  $\alpha$  helix of ZAR1 protrudes from the CC domain and forms a funnel-shaped structure, which is essential for plasma membrane association and triggering cell death. In a follow-up study, the cell death activity of the ZAR1 resistosome was further delineated in plant cells by showing Ca<sup>2+</sup> influx, perturbation of subcellular structures, and production of reactive oxygen species after ZAR1 activation, suggesting that the ZAR1 resistosome acts as a calcium-permeable cation channel in the effector-triggered immune response (Bi et al., 2021).

The structural and functional conservation of CC-type NLRs has been nicely demonstrated with the Sr35-AvrSr35 complex, which is the first model of direct recognition of a pathogen effector by CC domain-containing NLR, providing potential guidance towards NLR engineering for crop improvement (Förderer et al., 2022). Sr35 is a CC-domain containing NLR identified in a *Triticum monococcum* landrace. When transferred to bread wheat, Sr35 also confers

immunity to *Pgt* Ug99 via recognition of its effector AvrSr35. Like the ZAR1 resistosome, the Sr35-AvrSr35 complex appears to be pentameric and forms a nonselective calcium channel, with both CC and NBD contributing to the packing of protomers. The active conformation is further stabilized by an ATP molecule that forms direct interactions with NBD and HD1. The conserved 'EDVID' motif (Rairdan et al., 2008; Wróblewski et al., 2018) in CC domain interacts with a positively charged surface patch formed by arginine residues on LRR, which is required for the cell death activity.

Other than CC-type NLRs that induce cell death by forming a calcium channel on plasma membrane, a different function has been discovered for resistosome-forming TIR-NLRs (Ma et al., 2020; Martin et al., 2020). RPP1 (RECOGNITION OF *PERONOSPORA PARASITICA* 1) is a TIR-domain containing NLR in *Arabidopsis*, it has a C-JID (C-TERMINAL JELLY ROLL AND IG-LIKE DOMAIN) in addition to a canonical STAND class protein structure. Together with LRR domain, C-JID mediates direct recognition of the effector protein ATR1 (*ARABIDOPSIS THALIANA* RECOGNIZED 1). The NBD of RPP1 undergoes conformational change and tetramerizes following effector binding, bringing the four TIR domains close together, which is a prerequisite for the Mg<sup>2+</sup>/Ca<sup>2+</sup>-dependent NADase holoenzyme activity. RPP1 resistosome appears to be a two-fold dimer of dimers, while the symmetric dimers stabilize the complex, the asymmetric dimers create two NAD<sup>+</sup> binding sites with two centrally located BB-loops, which is essential for RPP1-mediated immunity. Activation of RPP1 results in the hydrolysis of NAD<sup>+</sup>, leading to the induction of host cell death (Ma et al., 2020). A similar activity is also reported for ROQ1 (RECOGNITION OF XOPQ 1), a TIR protein in *Nicotiana benthamiana*. ROQ1 monitors bacterial infections through direct recognition of the *Xanthomonas euvesicatoria* type III effector XopQ (*XANTHOMONAS* OUTER PROTEIN Q) (Martin et al., 2020).

Recently, a cyclic nucleotide monophosphate synthetase activity has been discovered for TIR proteins consisting of a TIR domain only, complementing the NADase activity in TIR-mediated immune responses (Yu et al., 2022). The TIR domain of flax NLR L7 along with other TIR-only proteins have a strong nuclease activity *in vitro*. Using RNA/DNA as substrates, L7 TIR generates 2',3'-cAMP and 2',3'-cGMP in a non-sequence-specific manner, exhibiting cNMPs synthetase activity shared among TIR proteins. Cryo-EM (cryo-electron microscopy) analysis of L7 TIR shows it can form a filamentous structure, with two L7 TIR protofilaments periodically winding around two dsDNA (double stranded DNA) in between as the scaffold. Each L7 TIR

protofilament is tetrameric and is composed of two types of symmetric homodimers mediated by two different interfaces, which together with nucleic acid binding, are important for the synthetase activity and TIR-mediated cell death. Further supporting the crucial role of 2',3'-cAMP/cGMP in immune response, *AtNUDT7* (*ARABIDOPSIS* NUDIX HYDROLASE 7), a negative regulator of TIR-NLR signaling, is found to function as a phosphodiesterase that specifically targets 2',3'-cAMP/cGMP for hydrolysis and attenuates TIR-mediated cell death *in vivo* (Yu et al., 2022).

### 1.2.3 Structural studies of STAND proteins in prokaryotes

While eukaryotic STAND proteins are mainly responsible for the regulation of programmed cell death and self/non-self-discrimination, their counterpart in prokaryotes has various roles in different signaling contexts, including nutrient utilization and biosynthesis of secondary metabolites (Leipe et al., 2004). Given the consensus in domain organization of STAND members across different kingdoms, bacterial STAND proteins often act as regulatory nexuses involved in the integration of multiple signals that are transmitted by various fused signaling domains. Despite decades of genetic and biochemical research, a clear picture of the molecular mechanism regarding the function of bacterial STAND is still lacking. In a recent study, the first crystal structure of a STAND with a TPR (tetratricopeptide repeat)-type sensor domain was reported, shedding light on the understanding towards inhibitor-enhanced autoinhibition of STAND regulation in gene transcription (Lisa et al., 2019).

PH0952 is a predicted transcription factor from *Pyrococcus horikoshii* with an N-terminal ArsR-like DNA-binding domain, and is a close homolog of MalT, one of the best studied STAND proteins. Phylogenetic and clustering analysis shows that PH0952 homologs are positioned between the Apaf-1 and MalT clades, suggesting that PH0952-like proteins represent an evolutionary missing link between Apaf-1 and MalT (Lisa et al., 2019). The inactive structure of PH0952 is devoid of its N-terminal effector domain, and is comprised of a canonical NOD module followed by an arm domain and a TPR-type sensor domain, with a conformation in the shape of the letter G. An ADP molecule is found to be accommodated similarly as in Apaf-1, where it establishes interactions with all the subdomains of the NOD. The autoinhibition is essentially maintained by interactions between NBD with arm and sensor, which is similar to

that of Apaf-1, regardless of a clear distinction between the types of sensor domains in both proteins. Based on the structure of PH0952 and a previous reported MalT sensor domain, a model of MalT lacking its DNA-binding domain was built using homology modeling, suggesting the NBD-sensor contact is a recurrent feature of resting STAND that also exists in MalT, and this seemingly weak interaction might facilitate MalT repression by MalK, a membrane-localized inhibitory protein (Lisa et al., 2019).

## 1.3 The maltose system in *E. coli* (*Escherichia coli*)

### 1.3.1 Transcription regulation in *E. coli* and the maltose regulon

As in most areas of molecular biology, studies of *E. coli* have provided the model for following investigations of transcription in eukaryotic cells. Subsequent to the discovery of mRNA, the first RNAP (RNA polymerase) was also purified from *E. coli*. The basic mechanisms by which transcription is regulated were likewise elucidated in *E. coli* by François Jacob and Jacques Monod, of which regulated gene expression enables the cell to respond to variations in the environment, such as fluctuations in nutrients availability (Brenner et al., 1961; Gros et al., 1961; Jacob and Monod, 1961). An understanding of transcription in *E. coli* has thus provided the foundation for studies of the far more complex mechanisms that regulate gene expression in eukaryotic cells. The first model for the control of gene transcription was illustrated by the lactose operon, of which a repressor protein binds the operator and blocks transcription until its release triggered by lactose binding, leading to the transcription of genes involved in lactose metabolism. In contrast to the lactose model, where the control of gene expression is negatively exerted by repressor proteins, the maltose regulon is represented by a group of genes positively regulated under the control of MalT, a transcriptional activator at all *mal* promoters. This system consists of 10 genes that together orchestrate the uptake and efficient catabolism of polysaccharide in *E. coli*, including *lamB*, *malE*, *malF*, *malG*, and *malK* that mediate the translocation of maltose and maltodextrins, as well as *malP*, *malQ*, *malS*, and *malZ* which catalyze the degradation of the glucose polymers to glucose/glucose-1-phosphate (Boos and Shuman, 1998).

### 1.3.2 Current model for the regulation of the maltose regulon

The activity of the maltose regulon is governed by two branches of regulatory circuits, both of which converge toward MalT. At transcriptional level, the expression of *malT* itself is controlled by CAP (CATABOLITE ACTIVATOR PROTEIN) and Mlc (MAKES LARGE COLONIES). CAP is a cAMP receptor that exists as a homodimer, it functions as a trans-acting transcriptional activator by binding to a DNA region upstream of the RNAP binding site. Two cAMP molecules bind to dimeric CAP and allosterically increase its affinity for DNA, CAP in turn forms a ternary complex together with RNAP holoenzyme and *malT* promoter to activate transcription (Chapon, 1982; Chapon and Kolb, 1983; Debarbouille and Schwartz, 1979). Mlc is a global regulator for sugar-metabolizing systems. In the absence of glucose transport, Mlc binds the Mlc-box, a regulatory region of *malT* with a sequence of palindromic structure, and represses *malT* transcription. When glucose is transported via the membrane-localized EIIC (ENZYME IIC COMPONENT) of the PTS (PHOSPHOTRANSFERASE SYSTEM), Mlc gets sequestered to the unphosphorylated EIIC from promoter, leading to derepression of *malT* and allows transcription to occur (Decker et al., 1998). Meanwhile, the formation of cAMP/CAP complex is glucose-dependent. When the amount of glucose transported into the cell is low, a cascade of events results in the increase of cytosolic cAMP level. This increase in cAMP level is sensed by CAP, which further goes on to activate *malT* transcription (Lee et al., 2000; Plumbridge, 1999). Therefore, the expression of *malT* is influenced both positively and negatively by glucose transportation.

Another layer of regulation affects the maltose system by modulating MalT activity. MalT is activated by maltotriose, an inducer which is not only formed when the cells are growing on maltodextrin but also, in low amounts, endogenously when the cells grow on non-maltodextrin carbon sources, due to degradation of intracellular galactose or trehalose (Boos and Shuman, 1998). On the other hand, MalT can be sequestered at an inactive state by inhibitory proteins. Repression of MalT is maintained by MalY along with MalK, the ATP-hydrolyzing subunit of the maltodextrin transporter, and Aes, a cytoplasmic protein with esterase activity (Panagiotidis et al., 1998; Peist et al., 1997; Reidl and Boos, 1991; Reyes and Shuman, 1988; Zdych et al., 1995). Previous studies suggest that MalK sequesters MalT at the cytoplasmic membrane, while MalY and Aes prevent MalT activation in cytoplasm. Interactions between the inhibitory proteins and MalT are impeded by a high concentration

of maltotriose and, reciprocally, the presence of repressor proteins reduces inducer binding by MalT (Joly et al., 2004; Joly et al., 2002; Panagiotidis et al., 1998; Richet et al., 2012; Schiefner et al., 2014; Schlegel et al., 2002; Schreiber et al., 2000).

The promoters of *mal* operons are delineated by MalT-binding box with a consensus sequence 5'-GGGGA (T/G) GAGG-3', centered at bp -37.5 bp or at -38.5 bp upstream of the transcriptional start point (Bedouelle, 1983; Bedouelle et al., 1982; Gutierrez and Raibaud, 1984; Raibaud et al., 1985). In addition to the MalT box proximal to the transcription initiation site, a structural motif formed by two MalT sites in a direct repeat has been found to be essential for *mal* gene expression. The two repeats are located upstream of the -38 bp region, and are involved in promoter activation either alone or in conjunction with the proximal MalT site irrespective of their orientation (Vidal-Ingigliardi et al., 1991). Expression of *malk lamB malM* and *malEFG* operons are also subject to the control of cAMP/CAP, which form a nucleoprotein complex structure at the promoter region together with MalT (Chapon, 1982; Chapon and Kolb, 1983; Debarbouille and Schwartz, 1979).

### 1.3.3 MalK, the ABC subunit of the maltose transporter regulates MalT activity

It has long been known that mutations in *malk*, a gene encoding the ATP-binding subunit of *E. coli* maltose transporter, lead to an elevated gene expression in the maltose regulon, indicating MalK acts as a repressor that negatively regulates the activity of the maltose regulon (Panagiotidis et al., 1998). Mutational analysis of MalK shows that its regulatory function is essentially conferred by the C-terminal domain of protein, and is independent of the role of MalK in substrate transport, which is mainly contributed by the N-terminal nucleotide-binding domain (Kühnau et al., 1991). The first evidence showing MalK participates in MalT inhibition comes from pulldown assays. A biotinylation site was introduced at the C-terminal of MalT to generate a MalT-BCCP (biotinylated MalT). Avidin-coated beads were then used to attach MalT-BCCP, and these beads were able to bind MalK, confirming a physical contact between MalK and MalT (Panagiotidis et al., 1998). In addition, using purified proteins, the transcriptional activity of MalT was shown to be strongly reduced by MalK in an *in vitro* system. This reduction was not observed for MalT-T38R or MalK-G346S, two mutants known to decrease the inhibitory effect of MalK. Both T38 of MalT and G346 of MalK are involved in the

recognition of MalT by MalK, as mutating either of them is sufficient to disrupt the interaction between MalT and MalK. Furthermore, the repression of MalT by MalK is alleviated in the presence of maltotriose, and MalK is found to have an antagonistic effect on maltotriose binding by MalT, suggesting MalK may inhibit MalT activation by competing with inducer binding (Joly et al., 2004).

Structural studies of the maltose transporter MalFGK2 show that depending on the ATP hydrolysis state, dimeric MalK resides in different configurations during maltose translocation (Chen et al., 2003; Lu et al., 2005; Oldham and Chen, 2011). The transporter consists of two transmembrane subunits, MalF and MalG, and two copies of the ATP-binding subunit, MalK. When in resting state, the MalK dimer is maintained solely through contacts between its C-terminal regulatory domains, and interacts with the transmembrane components via NBDs which are separated from each other. Upon diffusion of maltose into the periplasm, MBP (MALTOSE BINDING PROTEIN) would be loaded and is able to interact with the resting state MalFGK2. Binding of MBP to MalFGK2 triggers conformational changes of MalK, leading to a concerted motion of NBD closure facilitated by the bound ATP molecules, allowing ATP hydrolysis to occur. ATP hydrolysis and release of Pi then open the MalK dimer at the NBD interface, bringing the transporter back to a resting state. In a more recent study, MalK, together with other components of the maltose transporter, have been shown to sequester MalT to the membrane when maltose is absent from the growth media. This sequestration can be relieved by substrate internalization, or substituting a glutamine for the catalytic glutamate in MalK (E159), both of which render the transporter in a conformation distinct from the resting state and result in the dissociation of MalT from MalK (Richet et al., 2012).

### 1.3.4 MalY as a repressor for the maltose regulon

*malY* was discovered in search of genes controlling the synthesis of an endogenous inducer of the maltose system (Ehrmann and Boos, 1987; Reidl and Boos, 1991). A chromosomal T10 insertion that disrupts a gene later named *mall* was isolated. The insertion was outside of any known mal operons and strongly reduced the activity of the maltose system. However, *mall* has an amino acid sequence that highly resembles other classical repressor proteins, and negatively regulates its own transcription as well as the *malX malY* operon located next to it.

Genetic analysis confirms it was not MalI, but MalY that was directly involved in repression of the maltose regulon when overexpressed. Although MalY is found to confer  $\beta$ C-S lyase ( $\beta$ -cystathionase) activity and is able to complement *metC* mutations for growth in the absence of methionine, its enzymatic activity is not essential for the regulation of mal genes (Zdych et al., 1995).

Similar to MalK, MalY inhibits transcription activation by MalT in a purified transcription system. This inhibition can be relieved by maltotriose and is influenced by the identity of nucleotide bound to MalT, as the maltotriose concentration required for derepression is 1,000 times higher in the presence of ADP compared to ATP (Schreiber et al., 2000). Using purified MalT protein and MalY-agarose column, a direct interaction between MalT and MalY was demonstrated with pulldown assay, and the N-terminal 250 amino acids of MalT is found to bear the interaction site for MalY (Schlegel et al., 2002). The MalT/MalY interaction was further biochemically characterized by gel filtration, indicating the formation of a complex with 1:1 stoichiometry (Schreiber et al., 2000). The crystal structure of MalY was solved, revealing a dimeric configuration whose subunits are held together by numerous salt bridges and hydrogen bonds, as well as extensive hydrophobic interactions (Clausen et al., 2000). Each MalY protomer consists of two domains, a large pyridoxal 5-phosphate binding domain and a small domain comprising the C- and N-terminal portions of the protein. MalY mutants defective in *mal* gene repression were identified by genetic screen, the region of interaction with MalT was consequently deduced and mapped on the MalY structure, which is situated on the solvent-accessible side of the dimer and has a circular, convex shape with a hydrophobic core at its center, surrounded by highly polar residues. The cofactor pyridoxal 5-phosphate is bound far away from this interaction patch, making it unlikely to be involved directly in MalT inhibition (Clausen et al., 2000).

### 1.3.5 Aes as a repressor for the maltose regulon

Aes is an enzyme with acetyl esterase activity and has similar repressive functions on *mal* gene expression as MalK and MalY (Joly et al., 2002; Peist et al., 1997). The physiological role of Aes as well as its regulation in *E. coli* remain to be elucidated yet. Aes represses MalT activity both *in vitro* and *in vivo* through direct interaction, which can be reversed by maltotriose binding.



Conversely, maltotriose binding by MalT is also reduced by Aes. MalT bears the same surface determinant for association with MalK and Aes, indicating the interaction surfaces of these two repressor proteins on MalT may partially overlap (Joly et al., 2002). The crystal structure of Aes shows a globular architecture which can be subdivided into two structural parts: a core exhibiting a canonical  $\alpha/\beta$ -hydrolase fold and a funnel-like cap structure that forms the substrate binding site (Schiefner et al., 2014). Aes appears to be a homodimer in solution, with two helices mediating the dimerization. Residues that affect interaction with MalT are mapped on the surface of Aes, which are located at the opposite site of the dimerization interface, suggesting that each site can bind one MalT monomer, resulting in a heterotetrameric complex (Schiefner et al., 2014).

### 1.3.6 MalT, the central regulator of the maltose regulon

*malT* was initially identified as the positive regulator gene of the maltose regulon. Since its identification in the late sixties, numerous studies have been carried out to characterize it both genetically and biochemically (Hatfield et al., 1969a, b; Hofnung and Schwartz, 1971). The purified MalT protein appears monomeric in solution with a molecular weight of 103,000. In the presence of inducers, it forms multimers and triggers transcriptional activation of *mal* genes (Richet and Raibaud, 1987; Schreiber and Richet, 1999). Oligomerization of MalT followed by DNA-binding is supported by several lines of evidence. Firstly, the C-terminal part of MalT encompassing the last 95 amino acid residues confers DNA-binding specificity, and it was found the truncated forms of MalT lacking its C-terminal domain are negatively dominant over the function of the WT (wild type) protein (Cole and Raibaud, 1986). Analysis of the DNA-protein interactions during transcription activation shows MalT binds cooperatively to all of the activating sites present in each promoter, indicating the protomers occupying MalT sites interact with each other (Danot and Raibaud, 1994). Moreover, using gel filtration and sedimentation assay, MalT was found to self-associate into oligomers in the presence of maltotriose and adenine nucleotide. This oligomeric assembly was confirmed by cryo-EM analysis, albeit at a low resolution, which reveals a polydisperse homopolymers structure (Larquet et al., 2004; Schreiber and Richet, 1999).

MalT contains a NOD module followed by an arm domain and a sensor domain primarily composed of SUPR motifs, which corresponds to two-helix antiparallel bundles (Steggborn et al., 2001). Notably, while most STAND class proteins carry their effector domain on the N-terminal end, the DNA-binding domain of MalT is located on its C-terminal end. As suggested by Apaf-1, MalT might undergo an ATP-dependent, inducer-promoted domain rearrangement during activation (Riedl et al., 2005; Zhou et al., 2015). Binding of maltotriose destabilizes the resting form of MalT, presumably by removing an inhibitory contact between the sensor and the N-terminal, thereby nucleotide exchange can take place and self-association ensues, allowing the effector domain to bind MalT boxes present in target promoters. In the meantime, RNAP can be recruited to the promoter through direct interaction with the effector domain (Danot et al., 1996). It has been proposed maltotriose binding requires an intermediate configuration of MalT that does not pre-exist in the autoinhibited form, and both the arm and the sensor domains are needed to form the binding site, as any of them alone displays a low affinity for the inducer (Danot, 2001, 2015). Apart from its role in initiating MalT conformational change, binding of the inducer is also an indispensable step for transcription activation. In an *in vitro* transcription system, maltotriose is still required for the maximal activity of MalT-W317R, an autoactive mutant that can oligomerize regardless of the presence of inducer (Liu et al., 2013). Transcription activation by MalT is largely unaffected when ATP is substituted by its nonhydrolyzable analog, or when the ATPase activity of MalT is abolished by mutating the catalytic aspartate to alanine (MalT-D129A). On the other hand, MalT variant that is specifically defective in ATP hydrolysis tends to escape negative controls *in vivo* and *in vitro*, indicating that ATP hydrolysis of MalT is crucial for resetting the protein to inactive state, but not necessary in transcription initiation (Marquenet and Richet, 2007).

## 2. Aims of this study

Despite major breakthroughs in the understanding of STAND signaling in plant and animal systems, little is known about their bacterial counterparts and, more importantly, the role of STAND proteins in transcriptional regulation. MalT is a bacterial STAND protein that has been well-characterized genetically, and MalY is one of the repressors of the *E. coli* maltose system which has been reported to interact with MalT *in vitro*. In this study, I use MalT and MalY for structural analysis, and I aim to understand the biochemical basis underlying bacterial STAND regulation from the following two specific aspects:

- 1) The mechanism of MalT inhibition;
- 2) The mechanism of MalT activation and inducer-dependent transcription initiation.

## 3. Materials and methods

### 3.1 Buffer and media

**Table 3.1 Buffer and Components**

Buffer name	Components
50× TAE buffer	Tris base (2 M), acetic acid (1 M), EDTA (50 mM)
10× TBE buffer	Tris base (1 M), boric acid (1 M), EDTA (20 mM)
1× PBS buffer	NaCl (137 mM), KCl (2.7 mM), Na <sub>2</sub> HPO <sub>4</sub> (10 mM), and KH <sub>2</sub> PO <sub>4</sub> (1.8 mM)
1×TBS-T buffer	Tris-HCl pH 7.6 (50 mM), NaCl (150 mM), Tween <sup>®</sup> 20 (0.1%)
TSI buffer	Tris-HCl pH 8.0 (50 mM), sucrose (10%), KI (300 mM)
TSCI buffer	Tris-HCl pH 8.0 (50 mM), sucrose (10%), KCl (300 mM)
TKCl buffer	Tris-HCl pH 8.0 (10 mM), KCl (100 mM), PLP (10 μM)
10× SDS running buffer	Tris base (0.25 M), Glycine (1.9 M), SDS (35 mM)
buffer A	Tris-HCl pH 8.0 (50 mM), sucrose (10%), KI (200 mM), magnesium acetate (10 mM), EDTA (0.1 mM), Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub> (0.2 mM) with 0.1 mM ATP or 0.4 mM ADP
buffer I	Tris-HCl pH 8.0 (55 mM), tripotassium citrate (33 mM), DTT (1 mM), maltotriose (1 mM), and ATP (0.1 mM)
buffer E	Tris-HCl pH 8.0 (50 mM), NaCl (50 mM), tripotassium citrate (33 mM), DTT (1 mM)
pyruvate kinase reaction buffer	Tris-HCl pH 7.4 (125 mM), MgSO <sub>4</sub> (2.5 mM) and phosphoenolpyruvate (5 mM)

**Table 3.2 Media and Components**

Media name	Components
M9	Na <sub>2</sub> HPO <sub>4</sub> (42 mM), KH <sub>2</sub> PO <sub>4</sub> (22 mM), NaCl (8.5 mM) and NH <sub>4</sub> Cl (18.7 mM)
P1 salts solution	CaCl <sub>2</sub> (10 mM), MgSO <sub>4</sub> (5 mM)
top-agar	LB media with 0.7% agar

2YT	tryptone (16 g/l), yeast extract (10 g/l), NaCl (5 g/l)
M63B1	KH <sub>2</sub> PO <sub>4</sub> (0.1 M), (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> (15 mM), MgSO <sub>4</sub> (0.8 mM)

## 3.2 Primer, strain, and vector

**Table 3.3 Primer**

Name	Sequence 5'-3'
MalT_F_BamHI	CGCGGATCCATGCTGATTCCGTCAAACTAAG
MalT_R_Sall	ACGCGTCTGACTTACACGCCGTACCCATC
MalT_seq_1	CTTTCGGATTATCTGGTCGATGAGG
MalT_seq_2	AAAGGCGAATTGACCCGCTCACTGGC
MalYF_NdeI	GGAATCCATATGATGTTGATTTTTCAAAGG
MalYR_XhoI	CCGCTCGAGTTAACGAACAGCGCGGATGGCG
pGEX6P-1-seqF	TTGAAACTCTCAAAGTTGATTTTCTTAGCA
pGEX6P-1-seqR	ACCGTCTCCGGGAGCTGCATGTGTCAGAGG
malP800F_108CY3	CY3_TGCACTAGCTCCTCGACGCG
malP800R_108CY3	CY3_CCAGGCTCGATGTTGACCTT
MalTfseq	TCGAGATCTCGATCCC CGGAAATTAATACG
MalTrseq	GGGCTTTGTTAGCAGCCGGATCTCAGTGGT
N_del7F	CGTCCGGTTCGACTCGACCATACCG
N_del14F	ACCGTGGTTCGTGAGCGCCTGCTGGC
N_del21R	GTGATGGTGATGATGATGCATATGTATATC
S5L_F	CCTGATTCCGTTAAACTAAGTC
S5L_R	TGATGGTGATGATGATGCATATG
K6A_F	CCTGATTCCGTCAGCACTAAGTC
K6A_R	TGATGGTGATGATGATGCATATG
S8F-1	CGTCAAACTAGCTCGTCCGGTTC
S8R-2	GAATCAGGTGATGGTGATGATGATGCATATG
R9S_F	GATTCGTCAAACTAAGTAGTCCGGTTC
R9S_R	AGGTGATGGTGATGATGATGCATATG
R12F-1	GTCGTCCGGTTGCACTCGACCATAC
R12R-2	TTAGTTTTGACGGAATCAG
T16A_F	CTCGACCATGCCGTGGTTCG
T16A_R	TCGAACCGGACGACTTAG
R19F-1	CATACCGTGGTTGCTGAGCGCCTGC
R19R-2	GTCGAGTCGAACCGGACGAC
E140F	CCAGTGATCCACGCGTCAATGCGCTTC

---

E140R	ATTAGTGATCAGATGATAGTCATC
R143F-1	CACGAGTCAATGGCCTTCTTTATTC
R143R-2	GATCACTGGATTAGTGATCAGATG
N169F-1	CTGGGCATTGCCGCTCTGCGTGTTCCG
N169R-2	TTGCGGAAGGTTGCGTGACAAC
R173F-1	CAATCTGCGTGTTGCTGATCAACTGCTG
R173R-2	GCAATGCCCAGTTGCGGAAG
WHD_S1_F	CATTTTGCGCGCAATGAACGATGCC
WHD_S2_R	GCGCTTTTCAACAGAAAATGGCGCG
WHD_MD1_F	GTTTTTACAGCGGGCGGCTGATACCGGC
WHD_MD2_R	AGCCCCTGACGCTCAATCTCTTCGAGG
WHD_M1_F	GTTTTTACAGCGGGCGGATGATACCGGC
SEQ_WHD	CCATCAGGAAGCGAAGCAGTTTTTTGATTG
WHD_D312A_F	GTTTTTACAGCGGATGGCTGATACCGGC
S121-2HAIF	GCGGAATGGCATAGCGGATCCTACCCATACGATG
S121-2HAIR	ATAAAGTGGGGATCCTGCATAGTCCGGG
S121-2HABF	GGATCCCCACTTTATCTGGTCATCG
S121-2HABR	GCTATGCCATTCCGCCAGCTCAATG
383/386F	CACGCCTGGAGGCTGTTCCAGGCATAGCG
383/386R	ATTAAGCAGAATATCGCGCAG
410F	GGAAAATCCGCGGTTGGTGTTATTG
410R	AGCAAGCTGTCCCACGGCAGGGC
421F	CTGATGCAATACCAACATCGCTAC
421R	CCACGCCTGCAATAACACC
524/532F	GATGTCCGGCACTACGCTTTGTGGAGTTTAAGGCAGC
524/532R	GTGCTGGCGTGCCATCTGTTCCGG
555F	GAAAAAGCACGCCAGCTGATC
555R	CTGCGTTTTCCACGCGGTTTG
129F	CTGGTCATCGATGCCTATCATCTGATC
129R	ATAAAGTGGGCTATGCCATTCCG
258_K6A_F	ATGCTGATTCCGTCAGCACTCTCGAGAC
258_S5L_R	AGTTAATCACTTCACTGTGGAAAATGAG
258_S8A_F	CCGTCAAACTCGCGAGACCGGTTCCG
258_S8A_R	AATCAGCATAGTTAATCACTTCACTG
258_R9S_F	GTCAAACTCTCGAGTCCGGTTCGACTC
258_R9S_R	GGAATCAGCATAGTTAATCACTTC
258_S5L_F	ATGCTGATTCCGTTAAAATCTCGAGAC
258_S5L_R	AGTTAATCACTTCACTGTGGAAAATGAG
215_NdeI7F	CGTCCGGTTCGACTCGACCATAACCG
215_NdeI7R	CATAGTTAATCACTTCACTGTGG

---

pJM241F-Fseq	CGCACATTTCCCGAAAAGTGCCAC
pJM241R-Rseq	GTTTTCTTAATTCTCTGCTGGCTG
Y82F	CACCCAGCATGCCACCGCCATCG
Y82R	GAAAACCAGTGGGCAATAGCCGCGAG
D182F	TGTGGACGTGCGCTGAGCTGGAGATCATGGCTG
D182R	CTTCCCGGTAGGATTCTGTGGGC
E185F	TGTGGACGTGCGATGAGCTGGCGATCATGGCTG
D182E185F	TGTGGACGTGCGCTGAGCTGGCGATCATGGCTG
malT220_1	CCCAAGCTTGGGCTCATTCTTAATAGATTTATTAAG
malT220_2	GGAAGATCTTCTATTACATAACCTGCCGCGCG
Aes_1	CCCAAGCTTGGGCCATAACGCTCAATAAATTCATC
Aes_2	CCGCTCGAGCGGCATACAAAACTCCTTTCAAATTAC
Aes_3	CCGCTCGAGCGGTTCTTTACCGCTCAGCTTTAGCG
Aes_4	CGCGGATCCGCGACTGGCTTCCGCATTGGGGATG
M13seq1 F	GTGAGTTAGCTCACTCATTAGGC
M13seq2 R	GATAATCAGAAAAGCCCCAAAAACAG

**Table 3.4 Strain**

Name	Genotype
DH5 $\alpha$	F <sup>-</sup> <i>endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 <math>\phi</math>80dlacZ<math>\Delta</math>M15<math>\Delta</math>(lacZYA-argF)U169, hsdR17(<i>r<sub>K</sub><sup>-</sup>m<sub>K</sub><sup>+</sup></i>), <math>\lambda</math><sup>-</sup></i>
BL21 (DE3) $\Delta$ malT220	<i>E. coli</i> str. B F <sup>-</sup> <i>ompT gal dcm lon hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup>)</i> $\lambda$ (DE3 [ <i>lacI lacUV5-T7p07 ind1 sam7 nin5</i> ]) [ <i>malB<sup>-</sup></i> ] <sub>K-12</sub> ( $\lambda$ <sup>S</sup> ) $\Delta$ malT220
MC4100	<i>araD139 <math>\Delta</math>(argF-lac) U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR</i>
ED169	MC4100 $\Delta$ malB107
pop4062	MC4100 <i>malT::Tn10</i>
pop4146	MC4100 <i>aroB glpD <math>\Delta</math>malEpKp99::[P<sub>tac</sub>-P<sub>con</sub>] trp::[Kan<sup>r</sup>-malEp<math>\Delta</math>92-lac]<sub>op</sub></i> (F <sup>+</sup> )
pop7128	MC4100 <i>malTp10</i> (F <sup>+</sup> )
Strain G	MC4100 $\Delta$ malB107 <i>trp::[Kan<sup>r</sup>-malEp<math>\Delta</math>92-lac]<sub>op</sub> <math>\Delta</math>malT220 <math>\Delta</math>csgA::aadA7 <math>\Delta</math>aes</i> (F <sup>+</sup> )
Strain H	MC4100 $\Delta$ malB107 <i>trp::[Kan<sup>r</sup>-malEp<math>\Delta</math>92-lac]<sub>op</sub> <math>\Delta</math>malT220 <math>\Delta</math>csgA::aadA7 <math>\Delta</math>aes <math>\Delta</math>malY::Zeo<sup>r</sup></i> (F <sup>+</sup> )
TG1	K-12 <i>supE thi-1 <math>\Delta</math>(lac-proAB) <math>\Delta</math>(mcrB-hsdSM)5 (r<sub>K</sub>m<sub>K</sub>) F<sup>+</sup> [traD36 proAB<sup>+</sup> lac<sup>l<sub>o</sub></sup> lacZ<math>\Delta</math>M15]</i>
MG1655	K-12 F <sup>-</sup> $\lambda$ <sup>-</sup> <i>ilvG<sup>-</sup> rfb-50 rph-1</i>
UGB935	MG1655 $\Delta$ csgA::aadA7

**Table 3.5 Vector**

Name	Notes
pJM241	single-copy R1 run-away plasmid
pOM215	pJM241 containing P <sub>KAB-TTCT</sub> - <i>malT</i> , P <sub>KAB-TTCT</sub> is a constitutive promoter
pOM258	pJM241 containing P <sub>KAB-TTGG</sub> - <i>malT</i>
pOM206	pET24a (+) containing <i>malT</i> between NdeI and HindIII sites
pGEX-6P-1	Bacterial vector for expressing GST fusion proteins with a PreScission protease site
M13zeo2:: $\Delta$ malT220	zeocin-resistant M13 variant carrying <i>malT</i> deletion between the HindIII and BamHI sites
M13zeo2:: $\Delta$ aes	zeocin-resistant M13 variant carrying <i>aes</i> deletion between the HindIII and BamHI sites

M13mp11:: $\Delta$ malY::Zeo<sup>r</sup>M13mp11 containing zeocin-resistant gene harbored by the flanking region of *malY*

### 3.3 Molecular cloning

#### 3.3.1 PCR (polymerase chain reaction)

All the PCR for cloning were performed using Q5 High Fidelity DNA Polymerase (NEB) according to manufacturers' instructions. A typical reaction set-up and the thermocycling conditions are described as follows:

**Table 3.6 PCR Reaction Mix**

Components	Volume
5× Q5 Reaction Buffer	10 $\mu$ l
10 mM dNTPs	1 $\mu$ l
10 $\mu$ M Forward Primer	2.5 $\mu$ l
10 $\mu$ M Reverse Primer	2.5 $\mu$ l
Templete DNA	10 ng
Q5 High-Fidelity DNA Polymerase	0.5 $\mu$ l
ddH2O	to 50 $\mu$ l

**Table 3.7 Thermocycling Conditions**

Step	Temperature	Time
Initial Denaturation	98 °C	30 seconds
	98 °C	10 seconds
25-35 Cycles	50-72 °C	30 seconds
	72 °C	20-30 seconds/kb
Final Extension	72 °C	2 minutes
Hold	4-10 °C	



### 3.3.2 Agarose gel electrophoresis

Gels were prepared at a concentration of 1-1.5% (w/v) agarose in 1×TAE buffer, supplemented with 10 µl SYBR Green for a total volume of 100 ml. Samples were run at 120 V for 15-30 minutes and visualized by Gel Doc XR+Imager (Bio-Rad).

### 3.3.3 DNA extraction from agarose gels

DNA fragments separated by electrophoresis were cut out from gel and purified with the PCR clean-up and gel extraction kit (Macherey Nagel) according to manufacturer's instructions.

### 3.3.4 Restriction digestion of DNA

Purified DNA fragments or circular plasmids were digested by restriction enzymes from NEB according to manufacturer's instructions. A typical digestion reaction is set up as follows:

**Table 3.8 Restriction Digestion**

Components	Volume
DNA	1 µl
10× CutSmart Buffer	5 µl
Restriction Enzyme A	1 µl
Restriction Enzyme B	1 µl
ddH <sub>2</sub> O	to 50 µl

### 3.3.5 T4 ligation

Ligation of DNA fragments with linearized vectors containing 5' and 3' overhangs were done using T4 ligase (NEB) according to manufacturer's instructions, the ligation product was then used for transformation. A typical ligation reaction is set up as follows:

**Table 3.9 T4 Ligation**

Components	Volume
T4 DNA Ligase Buffer (10×)	2 $\mu$ l
Vector DNA (4 kb)	50 ng (0.02 pmol)
Insert DNA (1 kb)	37.5 ng (0.06 pmol)
T4 DNA ligase	1 $\mu$ l
ddH <sub>2</sub> O	to 20 $\mu$ l

### 3.3.6 Site-directed mutagenesis

Desired point mutations or deletions were introduced into the corresponding constructs with Q5 Site-Directed Mutagenesis Kit (NEB) according to manufacturer's instructions. Constructs containing WT gene sequences were used as template and PCR amplified by mutagenic primers. After being treated with KLD enzyme mix, 4  $\mu$ l of PCR product devoid of template plasmid was transformed into DH5 $\alpha$  and plated on selective LB plates for colony isolation and sequence validation.

### 3.3.7 Gibson assembly

A 2×HA tag was inserted into the pOM215 and pOM258 series vectors via Gibson assembly (Gibson Assembly Cloning Kit, NEB). The insertion and the backbone were individually amplified by PCR using primers with 15-20 bp overlapping ends, then assembled together according to manufacturer's instructions. 5  $\mu$ l of the reaction mixture was used for transformation and further cloning. A typical reaction is set up as follows:

**Table 3.10 Gibson assembly**

Total Amount of Fragments	0.02-0.5 pmol (X $\mu$ l)
Gibson Assembly Master Mix (2×)	10 $\mu$ l
ddH <sub>2</sub> O	10-X $\mu$ l

### 3.3.8 Plasmid DNA isolation from bacteria

For all the plasmid or phage replicative form DNA preparation, NucleoSpin® Plasmid Kit (Macherey-Nagel) was used according to manufacturer's instructions.

### 3.3.9 Preparation of constructs for recombinant protein expression

pOM206 is a pET24a (+) (Merck-Millipore) derivative encoding an N-terminally His-tagged MalT, in which the full-length *malT* gene was cloned between NdeI and HindIII sites. Sequence of full-length *malY* was amplified from *E. coli* genomic DNA with the introduction of a NdeI site and a XhoI site on 5' and 3' end of the gene, then digested by restriction enzymes and ligated to a linearized pGEX-6P-1 vector between the corresponding sites. All the variants of *malT* or *malY* for recombinant protein expression in *E. coli* were derived from the constructs containing WT sequence by Q5 site-directed mutagenesis.

### 3.3.10 Preparation of constructs for allelic exchange via M13 bacteriophage

Using a strain that carries *malT* deletion (BL21 (DE3)  $\Delta malT220$ ) as template, the fragment extending from position -14 bp to +3100 bp relative to *malT* transcription start site was amplified by PCR with the introduction of a HindIII site and a BglII site upstream and downstream of the homologous region, respectively. After restriction enzyme digestion, the fragment was ligated between the HindIII and BamHI sites of M13zeo2, a zeocin-resistant M13 variant, to generate M13zeo2:: $\Delta malT220$ . M13zeo2:: $\Delta aes$  was obtained in a similar way, except that the flanking regions upstream and downstream of *aes* (-327 bp to +3 bp and +939 bp to +1269 bp, respectively) were amplified separately and ligated via an XhoI site before being inserted into M13zeo2 between the HindIII and BamHI sites.

### 3.3.11 Preparation of constructs for *in vivo* assays

Plasmids used for  $\beta$ -galactosidase assay were variants of pJM241. pJM241 is a single-copy R1 run-away plasmid whose replication is under the control of a thermosensitive allele of the  $\lambda$  CI repressor, incubation of a transformed culture at 37°C leads to plasmid amplification. Two series of pJM241 derivatives encoding MalT were constructed, namely pOM258 and pOM215. pOM258 contains a PKAB-TTCT-*malT* fragment inserted between the EcoRI and BamHI sites of pJM241 vector, and pOM215 contains a PKAB-TTGG-*malT* insertion. Due to the difference in promoter strength, the expression level of *malT* from pOM215 was about ten times higher than that from pOM258. Similar as described previously, all the *malT* mutations were introduced into pOM258 and pOM215 by Q5 site-directed mutagenesis. Since a high concentration of MalT protein has been reported to be inhibitory for transcription *in vivo*, pOM258 series constructs were used for the quantification of MalT activity, while pOM215 series constructs were used in western blotting for a better detection of MalT variants whose protein level was low.

MalT-specific polyclonal antibodies were initially obtained from company for western blotting, but pilot experiments showed the antibodies raised against purified MalT had a low specificity, so tagged proteins were used instead to facilitate antibody detection. Based on our study and previous genetic data, both the N-terminal and the C-terminal of MalT are critical for its activity *in vivo*, thus it was unfavorable to attach tags to either ends. To minimize the effect of tag on protein functionality, a surface exposed loop region was chosen and a 2×HA tag was inserted after position S121. pOM215 and pOM258 series constructs were both modified similarly.

## 3.4 Protein and *in vitro* assay methods

### 3.4.1 Protein expression and purification

For the purification of MalT, *E. coli* strain BL21 (DE3)  $\Delta malT220$  was transformed with plasmid pOM206 which encodes N-terminally hexahistidine-tagged *E. coli malT*, or derivatives thereof, plated on LB agar supplemented with 25  $\mu$ g/ml kanamycin and grown at 37°C for overnight. An isolated colony was picked up for preparing a starter culture in LB containing 50  $\mu$ g/ml

kanamycin, then 1-liter LB containing 50 ug/ml kanamycin was inoculated with the starter culture and grown at 37°C until OD<sub>600</sub>=0.8, then induced by 1 mM IPTG (isopropyl β-d-1-thiogalactopyranoside) and grown at 20°C for overnight. The culture was harvested and resuspended in 100 ml TSI buffer with 0.4 mM ATP or ADP. Cells were disrupted by sonication for 10 minutes on ice. After centrifuging for 90 minutes at 14,000g, 4°C, protein was purified from the supernatant by affinity chromatography with Ni Sepharose 6 Fast flow resin (GE Healthcare), then further purified by gel filtration.

*malY* or variants thereof were cloned into pGEX-6P-1 vector and expressed similarly as described for MalT, except that the antibiotic was replaced by 100 µg/ml ampicillin. Cell cultures were harvested and resuspended in TSCI buffer with 10 µM PLP (pyridoxal 5'-phosphate). Cells were lysed by sonication, and GST (Glutathione S-transferases)-MalY was purified from the supernatant by affinity chromatography using Glutathione Sepharose 4B resin (GE Healthcare), then further purified with gel filtration upon removal of the GST tag by enzymatic cleavage.

#### 3.4.2 SDS-PAGE (SDS-polyacrylamide gel electrophoresis)

Proteins were separated by SDS-PAGE based on their electrophoretic mobility. Samples were mixed with 5×SDS-PAGE loading buffer (Bio-Rad) and boiled for 10 minutes at 95°C, then loaded on 10% TGX Stain-Free polyacrylamide gels (Bio-Rad) and run in SDS running buffer at 280 V for 30 minutes. Lower voltage and longer running time were used for western-blotting.

#### 3.4.3 Gel filtration assay

Affinity purified proteins were subjected to gel filtration with different columns and buffers. Monomeric MalT and dimeric MalY were purified using a Superdex 200 Increase 10/300 GL column (GE Healthcare) with buffer A and buffer TKCl, respectively. The MalT-MalY inhibitory complex was reconstituted by incubating purified MalT and MalY at a molar ratio of 1:1 on ice for 30 minutes, then purified by gel filtration using a Superdex 200 Increase 10/300 GL column (GE Healthcare) with buffer TKCl containing 0.4 mM ADP. To reconstitute the MalT oligomer, purified MalT monomer was first incubated in buffer I at 20°C for 20 minutes, then filtered

through a Superose 6 Increase 10/300 GL column with buffer E containing 1 mM maltotriose and 0.1 mM ATP. Samples from relevant fractions were applied to an SDS-PAGE gel and visualized by Coomassie blue staining.

#### 3.4.4 Pull-down assay

To analyze the interaction between MalT and MalY, N-terminally GST-tagged MalY was expressed and affinity purified as described, incubated with an excess amount of purified WT or mutant MalT proteins on ice for 30 min, then loaded on 100  $\mu$ l Glutathione Sepharose 4B resin (GE Healthcare). During incubation, 0.4 mM ADP and 10  $\mu$ M P5P were supplemented to the mixture. The resin was washed three times with 500  $\mu$ l buffer TKCl containing 0.4 mM ADP, proteins were then eluted with the same buffer containing 20 mM GSH (glutathione). The eluted samples were analyzed by SDS-PAGE and visualized by Coomassie blue staining.

#### 3.4.5 Quantification of protein-bound adenine nucleotide

For nucleotide quantification, bound ADP or ATP were released from protein samples by 10 minutes boiling at 92°C and centrifugation at 14,000 g for 15 minutes at 4°C. The resulting supernatant was divided equally and diluted 2.5 times in pyruvate kinase reaction buffer. To convert ADP to ATP for luciferase assay, 30 U pyruvate kinase (Sigma-Aldrich) or a same volume of 50% glycerol as control were added to the samples and incubated at room temperature for 50 min. 100  $\mu$ l of the reaction mixture was added to an equal volume of ATP Assay Mix solution (Sigma Aldrich) and photon emissions were read immediately using a Synergy H1 microplate reader.

#### 3.4.6 Electrophoretic mobility shift assay

The DNA used for this assay was a 108 bp promoter fragment from *malPp800*, which is a derivative of *malPp500* (Danot and Raibaud, 1994) that mimics the natural *pulAp* promoter and has a high binding affinity for MalT. Using M13malP3malPp800 phage as template. The fragment was amplified by PCR with Cy3 labeled primers, followed by purification using a

NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). To reconstitute the nucleoprotein complex, promoter DNA was mixed with purified WT or mutant MalT proteins at stoichiometric ratio, then incubated with buffer I in a total volume of 20  $\mu$ l at 20°C for 20 min. After addition of native gel loading buffer, samples were loaded onto a 6% native PAGE gel (Thermo Fisher, pre-run for 30 minutes) and electrophoresed at room temperature in a 0.5 $\times$ TBE buffer. The electrophoresis was done at 120 V for 60 minutes. Cy3 labeled DNA was visualized by Gel Doc XR+Imager (Bio-Rad).

#### 3.4.7 Immunoblots

3 ml of M9 cell culture was centrifuged and resuspended in a 160  $\mu$ l 1 $\times$ PBS buffer. After adding 40  $\mu$ l of 5 $\times$ SDS PAGE loading buffer, samples of which OD600 was about 15 were boiled at 100°C for 10 min, and electrophoresed through a 10% SDS PAGE gel and transferred to nitrocellulose membrane. The membrane was blocked with TBS-T buffer containing 3% non-fat dry milk at room temperature for 2 hours, probed with peroxidase-conjugated anti-HA antibody overnight at 4°C, then detected using after incubating with 2 ml SuperSignal West Dura and SuperSignal West Femto solution.

## 3.5 Strain construction and *in vivo* assays

### 3.5.1 P1 transduction

P1 is a bacteriophage with a rather sloppy packaging mechanism. During the infection of a donor strain, P1 occasionally packages the DNA of its bacterial host instead of its own into the protein capsid, generating a lysate containing either packaged phage DNA or packaged bacterial DNA, which could be used for infecting a recipient strain and transferring the DNA pieces from the donor via homologous recombination. The transductants harboring desired genetic information can then be selected by specific genetic markers closely linked to the transferred DNA fragment. The protocol for P1 transduction used in this study was modified based on a published method (Thomason et al., 2007) and is described as follows:

1. Inoculate 5 ml LB media with the *E. coli* donor strain (from which the P1 lysate is to be made), incubate overnight at 37°C. Grow the recipient strain at the same time.
2. Dilute the overnight donor culture 1:100 into 5 ml LB media containing 0.2% glucose and 5 mM CaCl<sub>2</sub>, incubate 30 to 45 minutes at 37°C with shaking in a 50 ml Falcon tube.
3. Add 100 µl of a freshly prepared P1 stock and continue to incubate with shaking until the culture lyses (about 3 hours). Then add a few drops of CHCl<sub>3</sub> and continue shaking a few minutes more.
4. Transfer the supernatant (the P1 lysate) to eppendorf tubes and centrifuge 10 minutes at 9200 g, 4°C.
5. Pour off the P1 lysate and purify it by passing through a sterile 0.45 µm filter to remove any residual bacteria.
6. Centrifuge 1.5 ml of the recipient strain cell culture in an eppendorf tube for 2 minutes at maximum speed, room temperature, to pellet the cells.
7. Discard the media and resuspend the cells in 750 µl sterile P1 salts solution.
8. Mix 100 µl of the cells/P1 salts mixture with varying amounts of lysate (1, 10, and 100 µl) in sterile tubes. As a control, include a tube containing 100 µl of cells without P1 lysate.
9. Allow phage to adsorb to the cells for 30 minutes at room temperature. Treat the controls similarly.
10. Add 1 ml LB media plus 200 µl of 1 M sodium citrate, then incubate for about 1 hour at 37°C with aeration.
11. Centrifuge each culture in a 1.5 ml microcentrifuge tube for 2 minutes at maximum speed, room temperature, to pellet the cells. Remove and discard the supernatant.
12. Resuspend the cell pellets in 50 to 100 µl of LB media and spread the entire suspension onto the appropriate selective plate supplemented with 5 mM sodium citrate. Also spread 100 µl of the P1 stock on another plate. Incubate the plates overnight at 37°C.
13. From the plate that shows growth with the least amount of P1 phage, pick one colony to isolate single clones, and incubate plates overnight at 37°C. Repeat this step a second time.



### 3.5.2 Phage preparation

A fresh phage stock of M13zeo2 derivative was prepared as follows:

1. Grow a 5 ml overnight culture of TG1 at 37°C.
2. Transform TG1 competent cells with ligation mixture containing M13zeo2 vector derivatives.
3. Add the transformation mixture and 50 µl of overnight TG1 culture into 3 ml melted top-agar (48°C), vortex vigorously and pour immediately on a LB plate, incubate at 37°C overnight.
4. Take 12 core plaques from the plate and add each of them in a 3 ml 2YT media inoculated with 30 µl overnight TG1 culture, incubate for 6 hours with shaking at 37°C.
5. Centrifuge in an eppendorf tube to pellet the bacteria and extract the replicative form DNA for sequencing.
6. Repeat step 2-4 by transforming TG1 competent cells using sequence-verified replicative form DNA.
7. Pellet the bacteria by centrifugation in an eppendorf tube and incubate the supernatant for 15 minutes at 65 °C to eliminate remaining bacteria. The phage supernatant can then be used for transferring deletions.

### 3.5.3 Allelic exchange via M13 bacteriophage

Since donor strains with *aes* and *malY* deletions were not readily available, allelic exchange was used instead of P1 transduction to make gene knockouts during strain construction. Mutant alleles, which are flanked by regions of homology to the recipient chromosome, were first cloned into allelic exchange vectors with a zeocin cassette, then introduced into recipient cells by conjugation. Single-crossover mutants in which the plasmid has integrated site-specifically into the chromosome can be selected based on zeocin resistance. The protocol for allelic exchange is described as follows:

1. Grow a 5 ml LB overnight culture of the recipient strain.
2. Dilute the overnight culture 1:100 in LB media and incubate at 37°C till OD<sub>600</sub>=1-2.
3. Add 20 µl of the culture and 20 µl of a M13zeo2 derivative phage stock to 600 µl of M63B1 media.

4. Incubate the mixture for 1 hour at 37°C without shaking, then spread 150 µl of the mixture on an LB agar plate supplemented with 25 µg/ml zeocin and incubate overnight.
5. Streak 8 clones for colony isolation on the same media and incubate the plate overnight at 37°C.
6. For each clone, streak one isolated colony and check the sensitivity to M13mp18 phage. Incubate the plate at 37°C for 6 hours. Isolate M13mp18-resistant colonies on LB agar plate supplemented with 25 µg/ml zeocin, incubate the plate overnight at 37°C.
7. Inoculate 0.5 ml of LB containing 0.2% sodium deoxycholate with one isolated colony. Prepare 4 lysogenic candidates.
8. Incubate the LB cultures for 5-6 hours at 37°C, then plate the cultures on MacConkey plates with lactose and incubate overnight at 37°C.
9. For each lysogenic candidate, assess the M13mp18 sensitivity of three healthy red clones with incubation for 6-7 hours at 37°C. Isolate M13mp18-sensitive clones on MacConkey plates with lactose and incubate overnight at 37°C.
10. Streak one isolated colony on a LB agar plate for each of the two independent M13mp18-sensitive candidates, and check their corresponding phenotypes.
11. Further verify the deletion of genes by PCR and retain one candidate.

### 3.5.4 Conjugation

M13 infection requires the F pilus from the host strain. During strain construction, the F plasmid was acquired by conjugating the starter strain with pop7128. Steps for conjugation are described as follows:

1. Grow a 5 ml overnight LB culture for both the donor and recipient strains.
2. Inoculate a 5 ml LB culture with the overnight culture for both strains (1:100 dilution). Incubate them with shaking at 37°C to till OD<sub>600</sub>=0.5.
3. Mix 0.4 ml of the recipient strain culture with 4 ml of the donor strain culture, then incubate at 37°C for 2 hours with mild shaking.
4. Streak the mixture on the selective plate, and test the M13mp18 sensitivity of a single colony to M13 the second day. About 90 % of the clones should correspond to an F<sup>+</sup> recipient.

### 3.5.5 Construction of reporter strains

To construct the reporter strains for  $\beta$ -galactosidase assay, ED169 was used as the starting strain.  $\Delta malB107$  deletes most, if not all, of the *malK* gene. The transcriptional *malEp $\Delta$ 92-lac* fusion was brought into ED169 by P1 transduction using a P1 lysate of pop4146, and the resulting kanamycin resistant clones were selected based on lactose and tryptophan utilization (Lac<sup>+</sup>, Tryp<sup>-</sup>), generating Strain A. To facilitate M13 phage infection, a WT F factor was introduced into Strain A by conjugating Strain A with pop7128. Deletion of *malT*, *aes*, and *malY* was carried out by allelic exchange via M13 bacteriophage in a sequential order, and verified by PCR with primers flanking the deleted gene. In addition, the *csgA* gene, which encodes the structural subunit of the curli fimbriae, was deleted by P1 transduction with lysate from UGB935. Curli fimbriae are synthesized at temperatures below 37°C, which may lead to bacterial aggregation and thus prevent a correct measurement of OD of the culture.

### 3.5.6 $\beta$ -Galactosidase enzyme assay

Strains harboring pJM241, pOM215 or pOM258 derivatives were grown in 10 ml of M9 media supplemented with 0.4% glycerol, 0.01% tryptophan, 1  $\mu$ g/ml thiamine and 30  $\mu$ g/ml ticarcillin. The media was inoculated with 500  $\mu$ l of overnight culture (grown at 25°C in LB media with 30  $\mu$ g/ml ticarcillin) and grown at 25°C till OD<sub>600</sub>=1.  $\beta$ -Galactosidase activity was assayed at 28°C as described (Miller, 1972), using chloroform and 0.02% SDS to permeabilize the cells. Each culture was assayed as triplicate and the enzyme activity value obtained was corrected for the background as measured with pJM241 (about 11 Miller units). Fold change in the  $\beta$ -galactosidase activity was then determined for each variant as the ratio of the variant activity to the activity of WT MalT. Means of the fold change values were then calculated from at least two independent series of cultures.

## 3.6 Cryo-EM analysis

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

Cryo-EM sample was prepared with purified MalT-MalY protein complex at a concentration of about 0.5 mg/ml, 3  $\mu$ l of protein was applied to holeycarbon grids (Quantifoil Cu 1.2/1.3, 300 mesh) glow-discharged for 30 seconds at medium level in HarrickPlasma after 2 minutes evacuation. Grids were then blotted on filter paper (TED PELLA, INC.) for 2.5 seconds in 8 °C with 100% humidity and flash-frozen in liquid ethane using FEI Vitrobot Marked IV.

Titan Krios microscope operated at 300 kV, equipped with Gatan K2 Summit direct electron detector and a Gatan Quantum energy filter, was used to collect micrographs. Stacks were automatically recorded using AutoEMation in super-resolution mode (Lei and Frank, 2005). A nominal magnification of 64,000 $\times$  was used for imaging the samples, corresponding to a final pixel size of 1.061 Å on image. Defocus values varied from -1.0  $\mu$ m to -2.0  $\mu$ m. Exposure rate of data collection was 23 electron per pixel per second. The exposure time for both data sets was 2.56 seconds dose-fractionated into 32 sub-frames, leading to a total electron exposure of approximately 50 electrons per Å<sup>2</sup> for each stack.

### 3.6.2 Image processing and 3D reconstruction

The flowchart of the data processing of MalT-MalY complex is presented in Figure 4.3. All micrographs were 2  $\times$  2 binned, generating a pixel size of 1.061 Å. Motion correction was performed using the MotionCorr2 program (Zheng et al., 2017). CTF (contrast transfer function) models of dose-weighted micrographs were determined using CTFFIND4 (Mindell and Grigorieff, 2003). Around 1,854,962 particles were picked for several rounds of 2D classification using GPU-accelerated Relion 3.1 (Scheres, 2012a, b, 2016). A total of 1,145,822 good particles were selected and subjected to further 3D classification, from which 176,969 particles were selected and subject to 3D auto-refinement, CTF refinement and post-processing were performed, resulting in a final resolution of 2.94 Å for MalT-MalY complex. The resolution was determined according to the gold-standard Fourier shell correlation 0.143 criteria with a high-resolution noise substitution method (Rosenthal and Henderson, 2003). Local resolution distribution was evaluated using RELION (Kucukelbir et al., 2014).

### 3.6.3 Model building and refinement

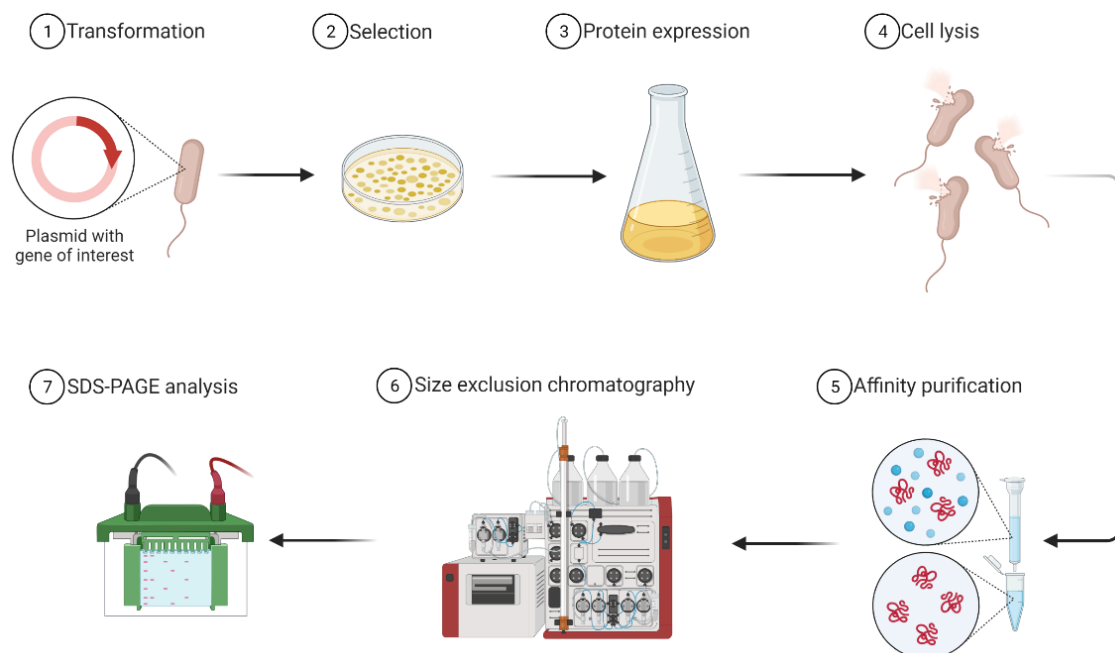
The crystal structure of MalY (1D2F) was docked into the EM map in Chimera (Pettersen et al., 2004). Model building of MalT-MalY was manually adjusted in COOT based on the 2.94 Å reconstruction map (Emsley et al., 2010). Structural refinement was performed in PHENIX in real space with secondary structure and geometry restraints (Adams et al., 2010). The final model after refinement was validated using molprobity in the PHENIX package (Adams et al., 2010).

## 4. Results

### 4.1 Reconstitution of the MalT-MalY complex

#### 4.1.1 Protein expression and purification

Since both MalT and MalY are bacterially-derived proteins, the *E. coli* system was chosen for recombinant protein expression (Figure 4.1). The pOM206 construct used for WT MalT protein expression was provided by Evelyne Richet (Institut Pasteur, Paris), which contains a N-terminally hexahistidine-tagged *malT* gene cloned between the NdeI and the HindIII sites of a pET-24b (+) vector. WT *malY* gene was cloned into a pGEX-6P-1 vector between the BamHI and the XhoI sites, resulting in a construct bearing N-terminally GST-tagged *malY*. All the constructs carrying *malT* or *malY* variants were derived from the WT plasmids via site-directed mutagenesis. Both MalT and MalY were expressed using conventional methods as described in the Materials and methods section.



**Figure 4.1 Schematic workflow of recombinant protein expression and purification.** Transformed bacteria were grown on LB plates with antibiotics, then a single colony was picked to prepare an expression culture. After cell disruption, proteins were purified with affinity purification and size exclusion chromatography, followed by SDS-PAGE analysis.

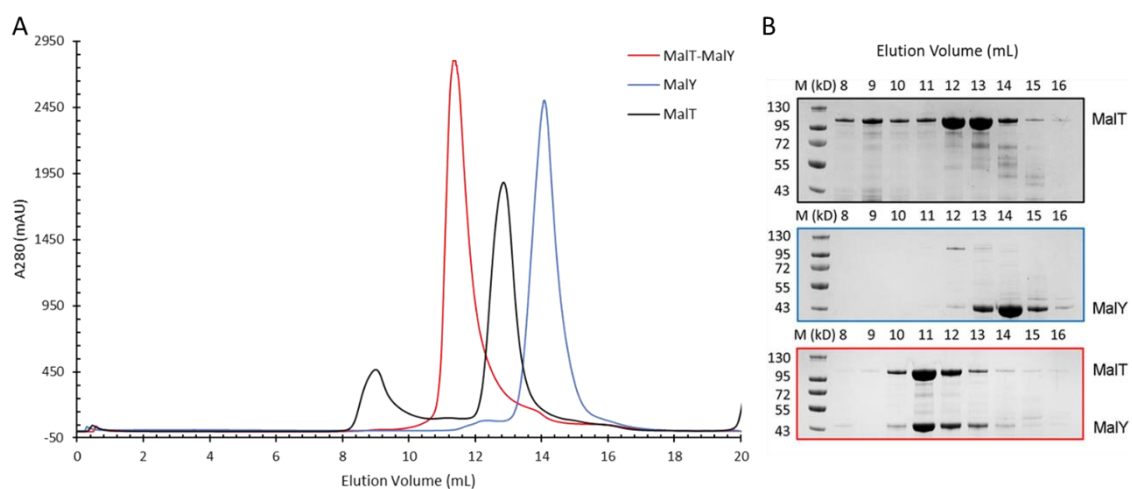
Proteins of MalT and MalY were purified with affinity purification followed by size exclusion chromatography. For MalY, the GST tag was enzymatically removed after affinity purification.

To ensure both MalT and MalY purified were in autoinhibition-competent form, 0.1 mM ADP and 10  $\mu$ M PLP were supplemented during all the purification steps. Both proteins were relatively sturdy and their yield was high (>3 mg protein from a 400 mL expression culture). In gel filtration, the elution volume of MalT corresponds to a monomer (103 kDa), and MalY eluted as a dimer (87 kDa) (Figure 4.2).

#### 4.1.2 MalT interacts with MalY *in vitro*

To test if the recombinantly expressed proteins were able to interact, pull-down assays were performed. One tagged protein was used as the bait immobilized on affinity resin, and the other protein partially purified was captured as the prey. Pull-down assays followed by SDS-PAGE analysis showed a clear interaction between MalT and MalY with either of them as the bait, although a definite stoichiometry of the protein complex could not be assessed.

To further purify the MalT-MalY complex, size exclusion chromatography was used (Figure 4.2 A). Based on previous studies, MalT and MalY enter the binding reaction in a 1:1 molar ratio (Schreiber et al., 2000), therefore MalT was incubated with a same amount of MalY, then the formation of inhibitory complex was evaluated by gel filtration. Gel filtration profiles showed the two proteins can form a stable complex with a smaller elution volume compared to any of them filtered alone, while no additional peak was observed, indicating a heterotetrameric assembly. The presence of MalT and MalY in fractionated samples was confirmed by SDS-PAGE analysis (Figure 4.2 B).



**Figure 4.2 MalT interacts with MalY *in vitro*.** Affinity-purified MalT and MalY proteins were further analyzed by size exclusion chromatography with a Superdex 200 Increase 10/300 GL column. MalT eluted as a monomer (103 kDa) and MalY eluted a dimer (87 kDa) in gel

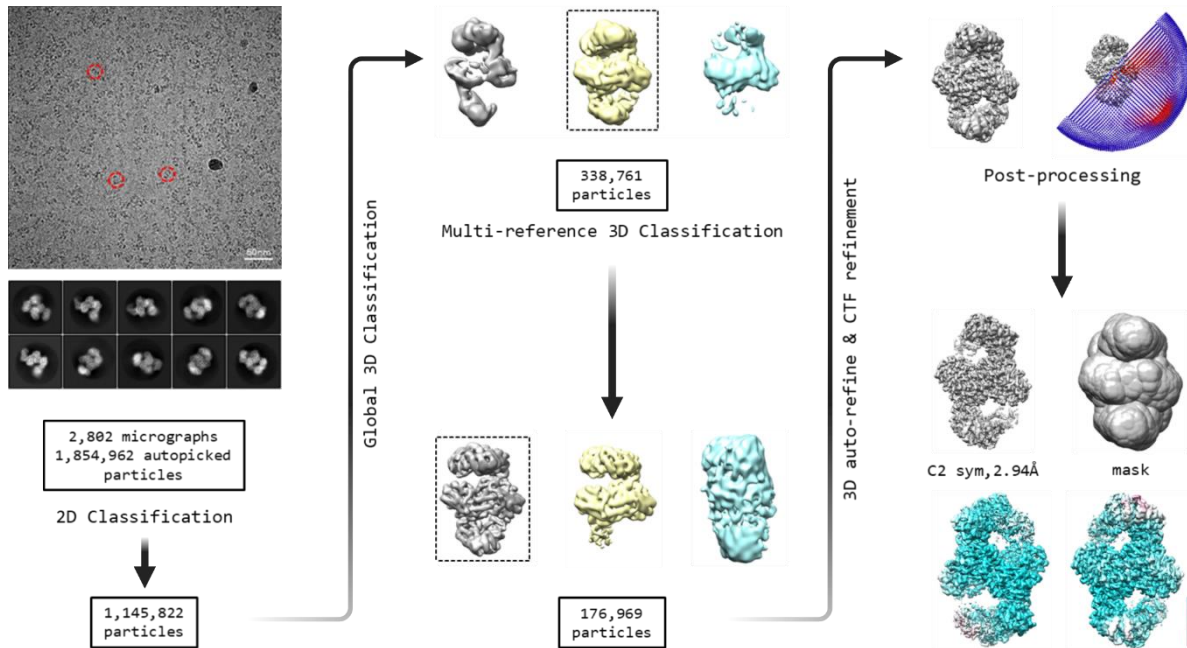
filtration. When incubated together at a 1:1 molar ratio, MalT and MalY formed a complex and eluted at earlier fractions. The first peak for MalT at 9 mL when filtered alone corresponds to protein aggregates.

#### 4.1.3 Cryo-EM analysis of the MalT-MalY complex

Based on the experience of our group and others (Lisa et al., 2019), crystallization of MalT protein has proved difficult. In hope that the binding of another protein could stabilize MalT and facilitate its crystallization, I tried to crystalize the MalT-MalY complex. However, despite several attempts, I was unable to obtain crystals containing both proteins, but only MalY, probably because some regions of MalT remain flexible in the complexed form, whereas the MalY dimer is more stable. I, therefore, attempted to solve the inhibitory complex structure using cryo-EM.

Purified MalT-MalY complex proteins were first examined by negative staining EM. Although it was difficult to estimate the features of individual particles without image processing, the sample was generally homogeneous and cryogenic samples were prepared. For sample preparation, different grids and blotting conditions were tested, and the proteins were found to be the most resistant to denaturation and to form the fewest aggregates on the gold grid. After an initial quality check, the best grid was recovered for further data collection. A total of 2,802 micrographs were taken and further processed using RELION 3.1. For the 2D classification, 1,854,962 particles were automatically selected. The average of the 2D classification showed that the dimeric MalY was bound by two MalT protomers in a centrosymmetric manner on each side opposite to the dimer interface. The best categories containing 1,145,822 particles were used for 3D reconstruction. After 3D reconstruction and 3D classification, a subset of 176,969 particles was used for the final refinement, yielding a map with a global resolution of 2.94 Å (Figure 4.3).





**Figure 4.3 Cryo-EM analysis of the MalT-MalY complex.** Proteins containing the MalT-MalY complex were used for cryo-EM sample preparation. After data collection, 2,805 micrographs were processed with RELION 3.1. 1,854,962 particles were extracted by autopicking and used for 2D classification, the resulting best categories containing 1,145,822 particles were then used for 3D reconstruction followed by 3D classification. The best 3D model with 176,969 particles was further refined, generating a density map with a global resolution of 2.94 Å. The analysis was done by Yue Sun from Tsinghua University, Beijing.

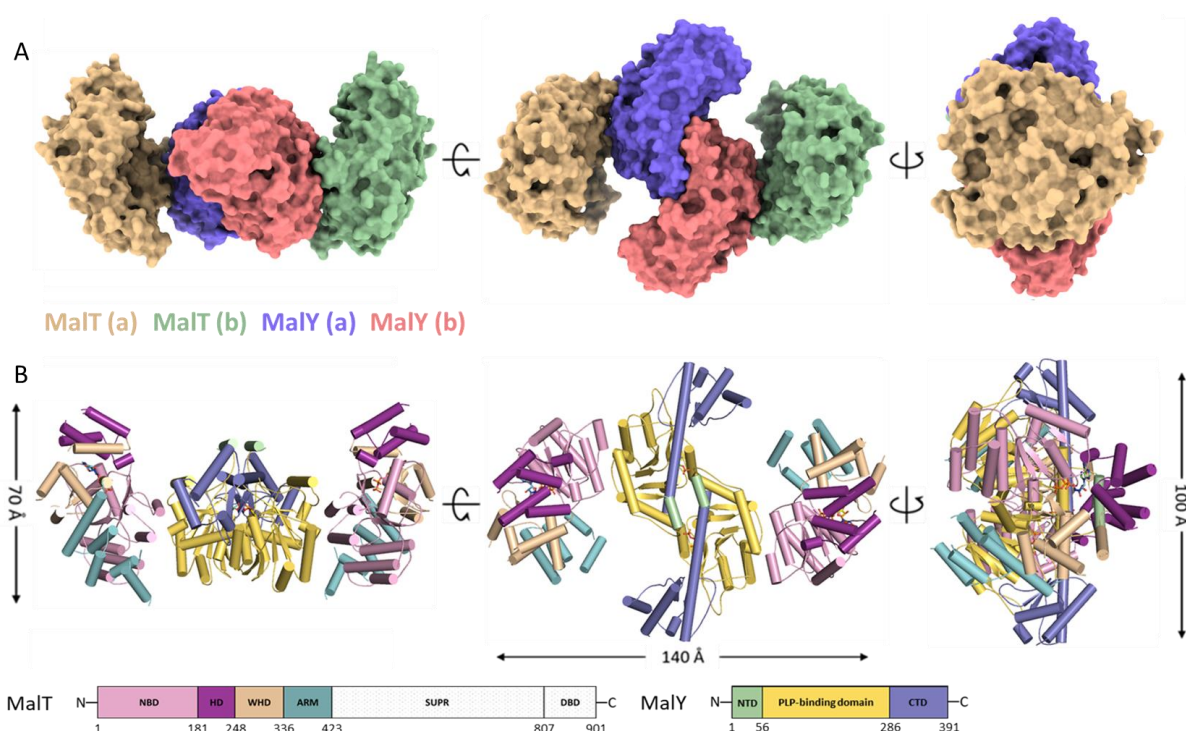
#### 4.1.4 The overall structure of the MalT-MalY complex

Consistent with the previously reported crystal structure (Clausen et al., 2000), MalY forms a homodimer with each of its subunits bound by one cofactor PLP. The N-terminal helices of two subunits together form a zipper-like interaction interface, comprising many polar interactions and hydrophobic interactions. The interaction between MalT and MalY is mainly manifested by hydrogen bonds, which are formed between NBD of MalT and the centrally located PLP-binding domain of MalY (Figure 4.4 A). The sensor domain of MalT, which is composed of superhelical repeats (SUPR), as well as the C-terminal DNA-binding domain, are not well defined in the complex structure, though it was proposed that a NOD-sensor contact observed in other inactive STAND proteins also contributes to MalT autoinhibition (Lisa et al., 2019). The structure of MalT consists of four domains: NBD with its Walker A and B motifs (residues 1 to 181), HD (residues 182 to 248), WHD (residues 249 to 336), and arm domain (residues 337 to 423) (Figure 4.4 B). Extensive interdomain interactions and hydrophobic packing give rise to a compact structure. The density of an ADP molecule is found between

NBD and HD, of which the identity was biochemically confirmed by luciferase assay (Table 4.1). Notably, intramolecular interactions between NBD and WHD of MalT are directly mediated by the very N-terminal segment from NBD, indicating a critical role of this loop region in maintaining the autoinhibition of MalT.

**Table 4.1 Adenine nucleotide content in purified MalT or MalT-MalY complex.** The amount of protein-bound nucleotide was measured by luciferase assays in the presence or absence of pyruvate kinase. The values represent the molar ratio of nucleotide to MalT protomer, which are the means  $\pm$  SD from results of two independent experiments performed in duplicate.

Sample	ADP+ATP (+Pyruvate)	ATP (-Pyruvate)
MalT monomer	1.08 $\pm$ 0.10	0.07 $\pm$ <0.01
MalT-MalY	1.03 $\pm$ 0.09	0.04 $\pm$ <0.01



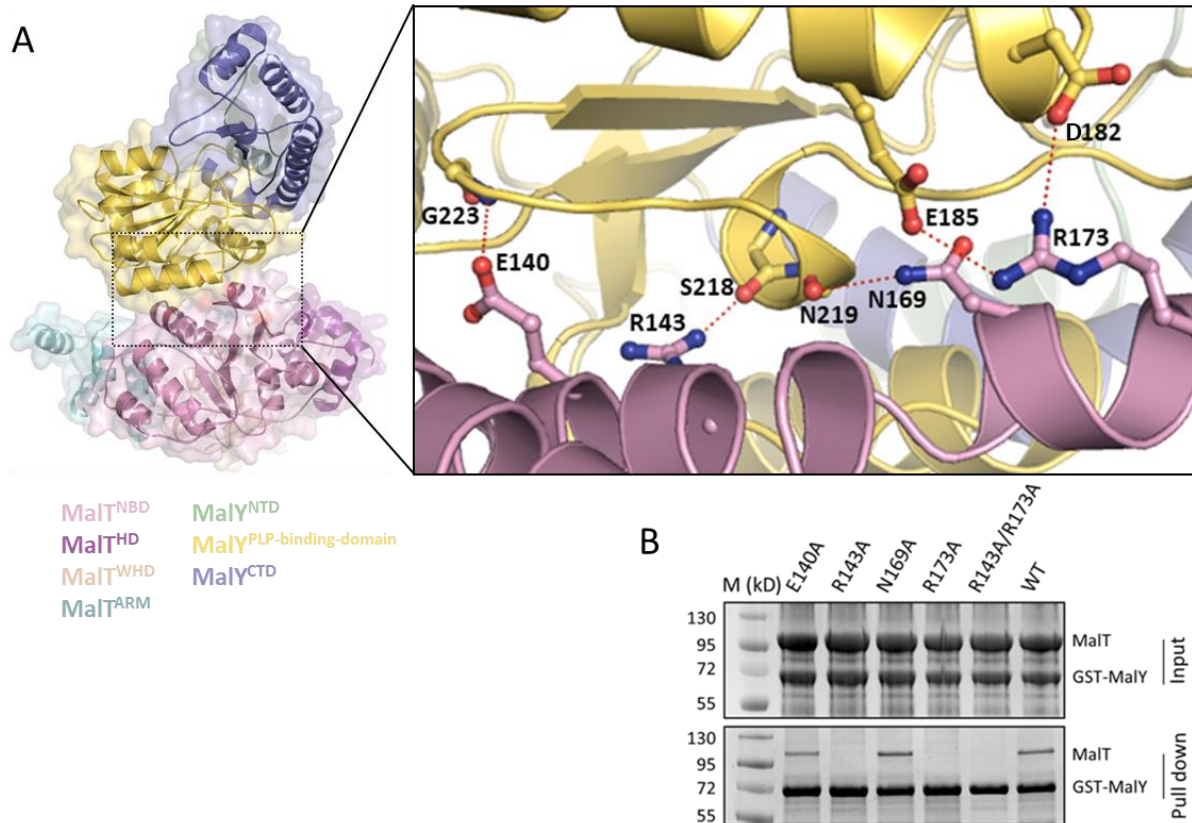
**Figure 4.4 The overall structure of the MalT-MalY complex. (A)** Surface representation of the MalT-MalY complex. Two subunits of the MalY dimer (MalY (a) and MalY (b)) are colored in blue-purple and pink, respectively. Two MalT molecules (MalT (a) and MalT (b)) are colored in wheat and light green, with each of them binds to MalY in a centrosymmetric manner. **(B)** Cartoon representation of the MalT-MalY complex. NBD, HD, WHD, and ARM are colored in pink, purple, wheat, and cyan, respectively. MalT autoinhibition is maintained primarily by interdomain interactions, and MalT interacts with the PLP-binding domain of MalY (colored in yellow) via NBD. The C-terminal domains of MalT are not well resolved in the structure.

## 4.2 The molecular mechanism of MalT repression by MalY

### 4.2.1 The MalY-MalT interaction interface in detail

The interaction interface between MalY and MalT involves helices  $\alpha 7$  and  $\alpha 8$  from MalY PLP-binding domain and helices  $\alpha 7$  and  $\alpha 8$  from the NBD of MalT, in total comprising five hydrogen bonds (Figure 4.5 A). Consistent with previous genetic data from which the MalT interaction patch on MalY was deduced by mutational screening (Clausen et al., 2000), key residues and residues in their vicinity are identified on MalY in our structure (D182, E185, S218, N219, and G223). Residues on the MalT interface that contribute to intermolecular interactions can be grouped into two clusters, E140-R143 and N169-E173, both of which are situated on the solvent-accessible side of NBD. Remarkably, R171, a conserved residue that was predicted to be the arginine finger among MalT clade STAND protein (Marquenet and Richet, 2010), is also located within the interacting surface patch. Substitution of R171 with a glutamic acid reduces both MalT self-association and the ATPase activity observed in the presence of maltotriose (Marquenet and Richet, 2010), suggesting MalT might use the same interface for interacting with MalY and oligomerization.

To verify the structural observation, alanine substitutions were introduced in NBD and their impact on MalY interaction was evaluated using pull-down assays (Figure 4.5 B). As predicted by the structure, interaction between MalT and MalY was greatly impaired with MalT-R143A and MalT-R173A, in agreement with a gain-of-function phenotype observed for the MalT-R173H mutant (O. Danot and E. Richet, unpublished results). Meanwhile, mutating E140 or N169 of MalT had no pronounced effect on protein-protein interaction, suggesting these two residues are less important for MalY binding.



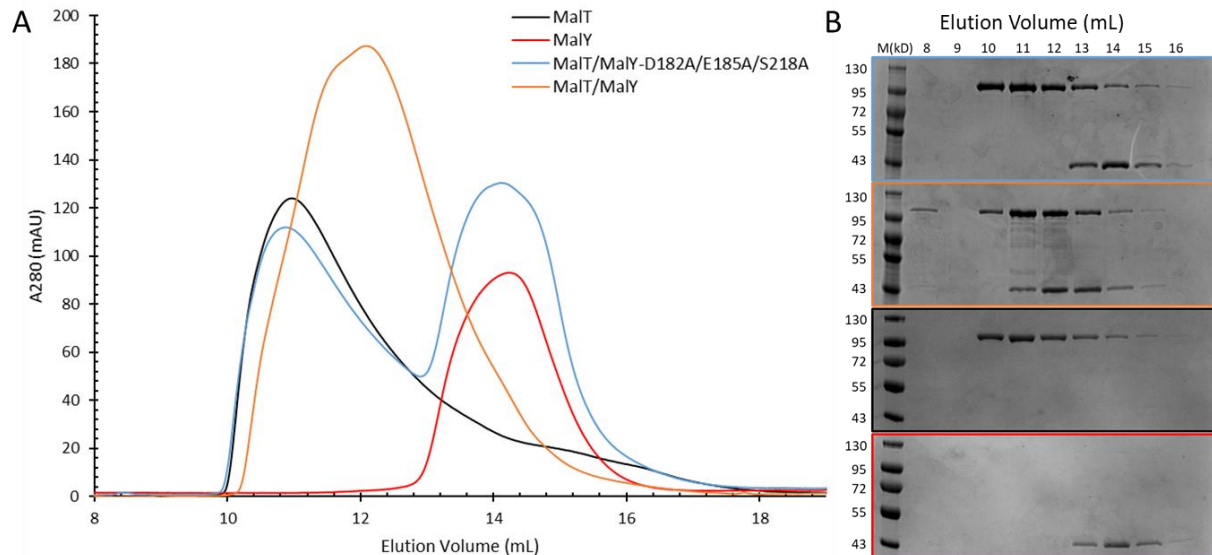
**Figure 4.5 The MaLY-MaIT interaction interface in detail. (A)** The interaction between MaIT and MaLY is mainly facilitated by polar contacts, with residues from both NBD and the PLP-binding domain of MaLY. **(B)** Alanine substitution of residues on MaIT disrupted the interaction to different extents, as revealed by pull-down assays. R143 and R173 from MaIT were identified to be two key residues maintaining the interaction.

#### 4.2.2 MaLY inhibits MaIT oligomerization *in vitro*

When activated, MaIT self-associates into oligomers and elutes in early fractions during gel filtration. To test whether MaLY inhibits MaIT by preventing its oligomerization, the complex formation was compared after incubation of MaIT with WT MaLY or MaLY-D182A/E185A/S218A, a triple mutant carrying the interface mutations.

Since ADP is also able to induce MaIT activation, albeit less effectively than ATP (Schreiber and Richet, 1999), protein incubation and gel filtration were performed in the presence of 1 mM maltotriose and 0.1 mM ADP (Figure 4.6). As expected, when incubated with mutant MaLY, MaIT readily oligomerizes, producing the same profile as MaIT being incubated alone followed by a separated peak of the MaLY-D182A/E185A/S218A dimer. When incubated with WT MaLY, most of the MaIT protein was complexed with MaLY, and the amount of protein eluted in early fractions significantly reduced, in agreement with previous reports that MaIT is much less

sensitive to maltotriose when complexed with MalY (Schreiber et al., 2000). Taken together, these results suggest that MalY represses the activity of MalT by sequestering it from oligomerization.



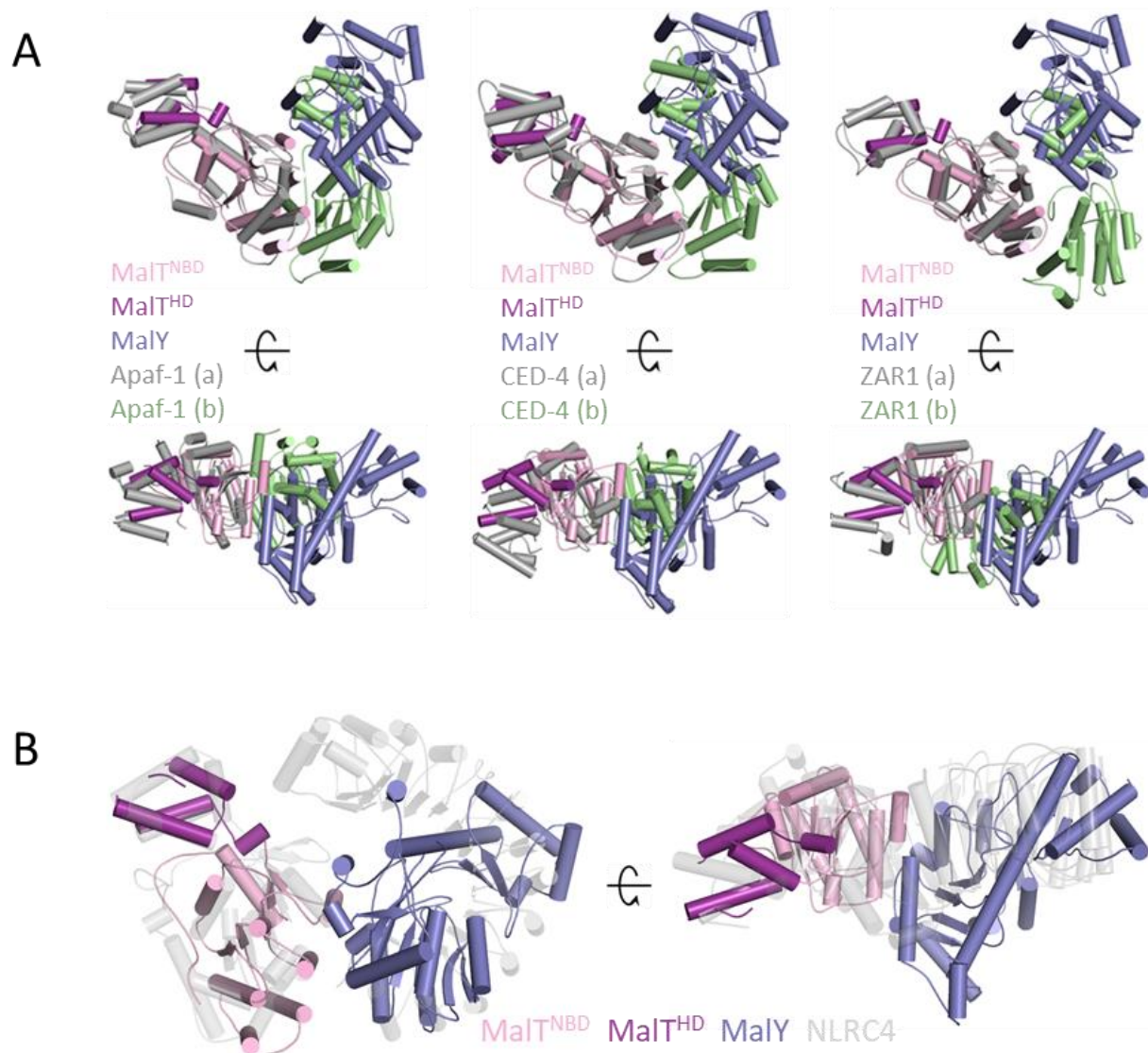
**Figure 4.6 MalY inhibits MalT oligomerization *in vitro*.** **(A)** A same amount of MalT, MalY, or MalY-D182A/E185A/S218A proteins were subjected to gel filtration analysis using a Superdex 200 Increase 10/300 GL column, either separately or together after incubation. In the presence of 1 mM maltotriose and 0.1 mM ADP, WT MalY, but not MalY carrying interface mutations, formed complex with MalT and inhibited its oligomerization. **(B)** SDS-PAGE analysis of respective samples after gel filtration.

#### 4.2.3 Structural alignment of the MalT-MalY complex with known NLRs

To further explore the structural role of MalY in MalT inhibition, structural alignments were done using MalT-MalY and other known protein structures from the STAND family. Using the NOD module of MalT as a query protein structure, the NOD of Apaf-1 was identified as one of the closest homologs by a DALI search. When MalT from the complex structure was superimposed to one protomer in a lateral dimer of active Apaf-1, MalY almost completely overlapped with the other protomer, supporting the notion that the oligomerization interface of MalT is masked upon MalY binding. Similar results were also found when comparing MalT-MalY to a lateral dimer of CED-4 or ZAR1 in their active conformation (Figure 4.7 A). Notably, the alignment of MalT-MalY with the inactive NLRC4 molecule resulted in severe steric clash between MalY and the LRR of NLRC4 (Figure 4.7 B), suggesting a functional similarity between MalY and the ligand-binding domain in sequestering the protein at the resting state (Hu et al.,



2013), which may partially account for the dispensability of MalT C-terminal domains in autoinhibition in our structure.

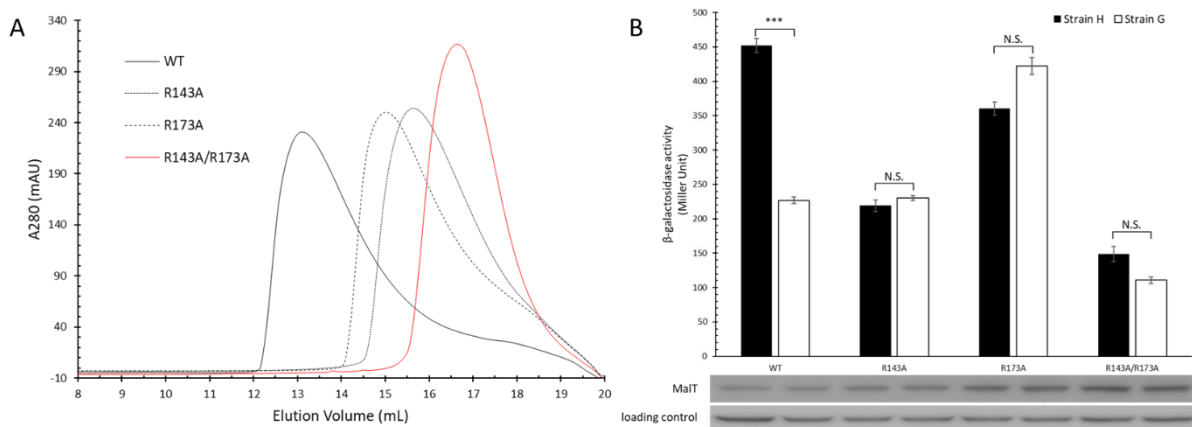


**Figure 4.7 Structural alignment of the MalT-MalY complex with known NLRs.** The MalT-MalY complex structure containing MalT<sup>NBD-HD</sup> and one interacting MalY molecule was aligned to **(A)** active lateral dimers comprising the N-terminal domains of Apaf-1, CED-4, and ZAR1, or **(B)** an inactive NLRC4 molecule consisting of its NBD, HD1, and LRR domains. While the MalT N-terminal domains aligned well with one protomer from the active dimer or the N-terminal domains from NLRC4, the MalY molecule overlapped with the other protomer or the LRR domain of NLRC4.

#### 4.2.4 MalT interface mutants are defective in oligomerization

To examine whether the inhibition of MalT by MalY is indeed executed through its oligomerization interface, the self-association capability of different NBD mutants that disrupted MalY interaction was analyzed using gel filtration. In the presence of 1 mM

maltotriose and 0.1 mM ATP, the size of oligomers formed by MalT-R143A and MalT-R173A was largely reduced compared to WT protein, and a double mutant carrying both substitutions could barely multimerize (Figure 4.8 A). Meanwhile, the oligomer formation of MalT-E140A and MalT-N169A, two mutants that could still interact with MalY, was not compromised, confirming a critical role for residue R143 and R173 in both MalT inhibition and activation. The effect of NBD mutations on MalT transcriptional activity was further investigated *in vivo* using two  $\Delta malT$  reporter strains, namely, strain G and strain H (Figure 4.8 B). In both strains, the chromosomal *lacZ* reporter gene that encodes  $\beta$ -galactosidase was placed under the control of *malEp $\Delta$ 92*, a MalT-dependent promoter deleted for the sequence lying upstream from -92 bp in regard to the transcription start site of *malE* (Raibaud et al., 1989). While strain H is devoid of *malk*, *maly*, and *aes*, strain G still retains *maly*, therefore comparing the  $\beta$ -galactosidase activities from two strains would allow the dissection of MalY inhibition placed on MalT. As a result, all three mutants that abolished MalY interaction (MalT-R143A, MalT-R173A, and MalT-R143A/R173A) showed comparable  $\beta$ -galactosidase activities between the two strains and, not surprisingly, the activity of WT MalT was diminished in the presence of MalY, suggesting the MalY-mediated MalT repression is lifted by mutating MalT residue R143/R173. Notably, the transcriptional activity of the MalT double mutant was strongly reduced in comparison to WT, although its cellular protein level was increased, indicating the competency for transcription initiation is dependent on the size of the MalT oligomer formed. Taken together, the biochemical and *in vivo* data confirm the MalT-MalY interaction observed in our structure, and lead to the conclusion that MalY represses MalT by competing with MalT self-association.



**Figure 4.8** The *in vitro* and *in vivo* effect of MalT interface mutations on MalT activity. (A) A same amount of WT, MalT-R143A, MalT-R173A, and MalT-R143A/R173A proteins were subject to gel filtration analysis with a Superose 6 Increase 10/300 GL column in the presence

of 1 mM maltotriose and 0.1 mM ATP. All three MalT interface mutants were defective in oligomer formation. **(B)** MalY-mediated repression was released by interface mutations on MalT, as shown by MalT transcriptional activities in strains from different genetic backgrounds (*malK*, *aes* and *malY* are deleted in Strain H, while Strain G still contains *malY*).

## 4.3 Autoinhibition of MalT and insights into its activation

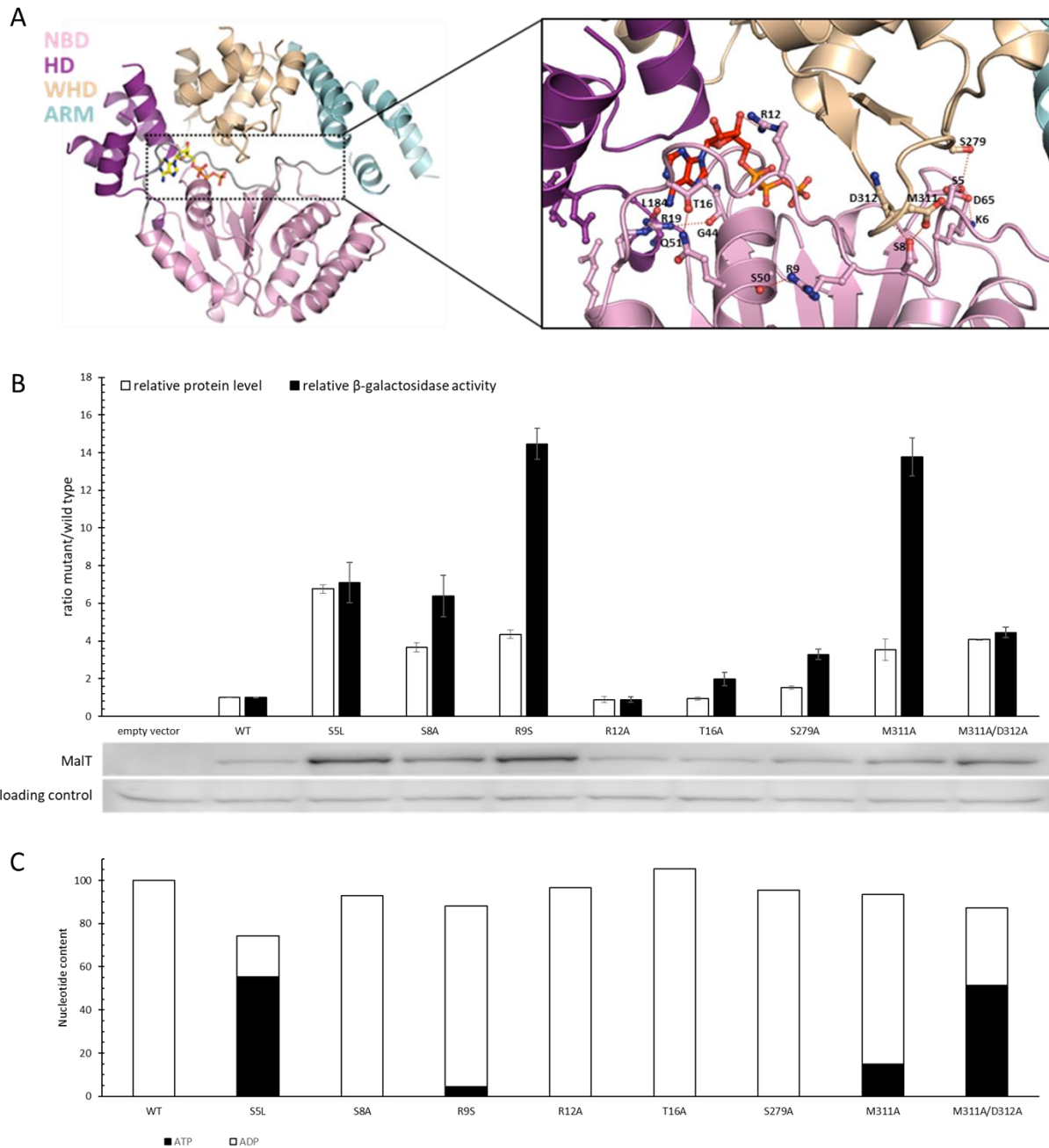
### 4.3.1 Autoinhibition of MalT is maintained via interdomain interactions

In Apaf-1 and its *Drosophila* homolog Dark, the ligand-binding domain WD40 repeats serves as an anti-apoptotic factor in the absence of stimuli (Hu et al., 1998; Rodriguez et al., 1999; Srinivasula et al., 1998). A similar mechanism also exists in NLRC4 and ZAR1, in which the LRR domain sequesters the N-terminal domains to stabilize the resting form (Hu et al., 2013; Wang et al., 2019b). In the case of MalT, a weak contact between NBD and the C-terminal domains was proposed to strengthen autoinhibition by facilitating MalK binding (Lisa et al., 2019). However, such a contact seems to be dispensable in our structure, as the C-terminal domains of MalT were flexible and not well defined.

An important feature of MalT is that the effector domain is not at the N-terminal end, but is connected to the sensor domain at the C-terminal end, while the 20 amino acids at the very beginning of NBD form a loop structure that is wedged between WHD and the rest of NBD (Figure 4.9 A). In both Apaf-1 and NLRC4, ADP-mediated NBD/WHD interaction plays a pivotal role in autoinhibition (Hu et al., 2013; Riedl et al., 2005). The ADP molecule in the inactive MalT structure is deeply buried and coordinated by NBD and HD, and the NOD module is kept in closed form mainly by direct interdomain interactions, thus disruption of these interactions should consequently promote autoactivation. Gain-of-function mutations in the N-terminal region have been isolated by genetic approaches (MalT-S5L, MalT-R9S), which result in higher transcriptional activity compared to the WT (Schlegel et al., 2002). Using structure-guided mutagenesis, residues in the first 20 amino acids involved in polar interactions were systematically mutated to alanine and their *in vivo* effect was evaluated by  $\beta$ -galactosidase assays (Figure 4.9 B). In addition to S5 and R9, mutations of S8 and T16 caused 8.5-fold and 2.5-fold enhancement of MalT activity, respectively. S5 and S8 form three hydrogen bonds with S279, M311, and D312 from WHD. Not surprisingly, mutating the three residues on WHD also leads to an increase in transcriptional activity (3-12 fold). As shown for many STAND



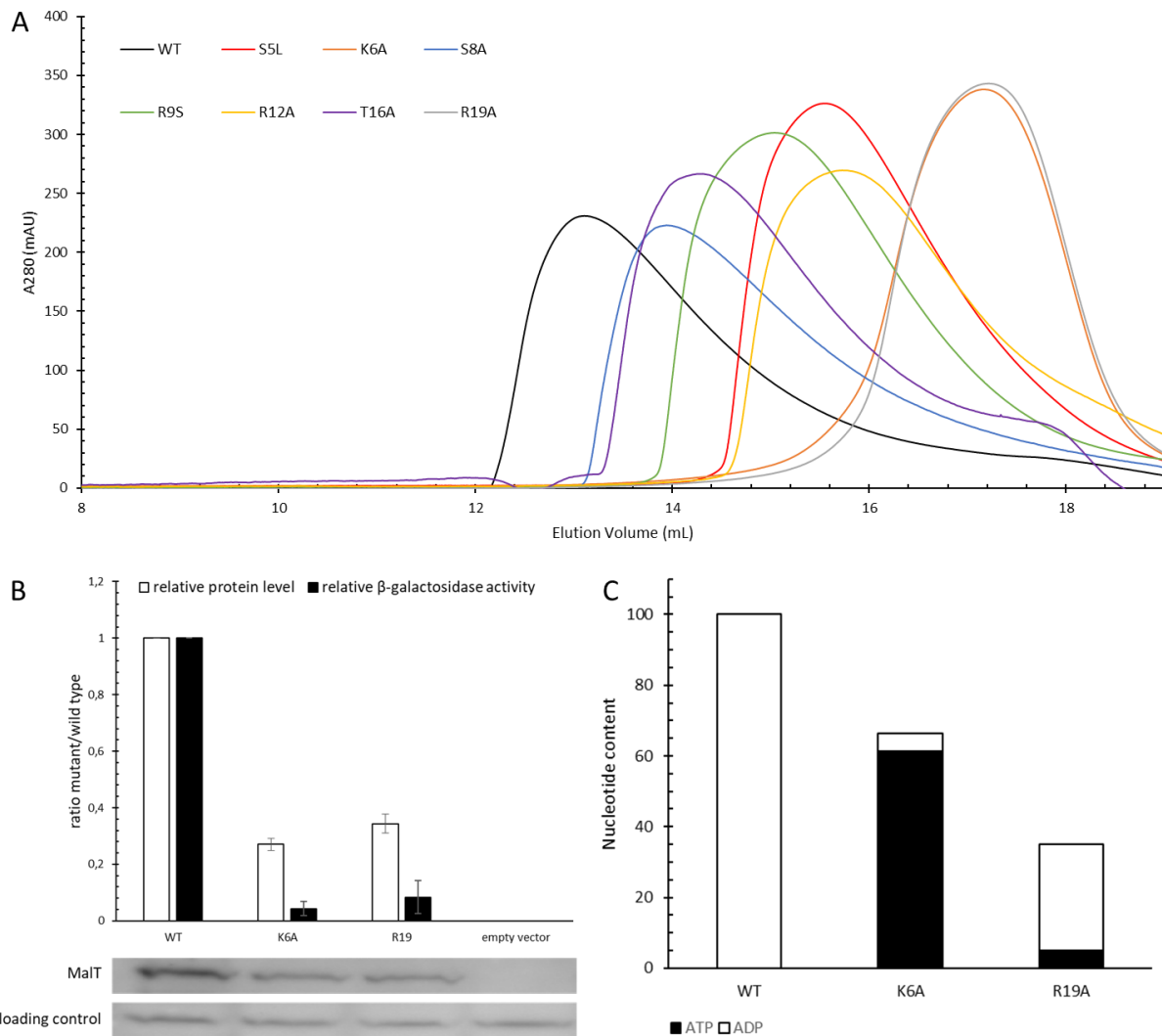
proteins, nucleotide exchange is an essential step coupled with protein activation (Hu et al., 2015; Wang et al., 2019a; Zhou et al., 2015). If N-terminal mutations facilitate MalT activation by removing the NBD/WHD contact, it is likely that the autoactive mutants may be more prone to ATP/ADP exchange than WT. To investigate this *in vitro*, WT or mutant proteins were first purified in the presence of 0.1 mM ATP. After removing free nucleotides from the sample by gel filtration, the protein-bound nucleotides were released and their amounts quantified by luciferase assay (Figure 4.9 C). While it was clear that WT MalT protein contained only ADP, both ATP and ADP were detected for MalT-S5L, MalT-R9S, MalT-M311A, and the MalT-M311A/D312A double mutant. Unexpectedly, nucleotide exchange was not observed for MalT-S8A, MalT-T16A, and MalT-S279A. Mutation of R12 results in a similar protein activity to that of WT MalT and does not affect nucleotide binding. When the total nucleotide occupancy was calculated, all the up-regulated mutants except MalT-T16A showed decreased total nucleotide occupancies compared to WT (74% for MalT-S5L, 93% for MalT-S8A, 88% for MalT-R9S, 95% for MalT-S279A, 93% for MalT-M311A, and 87% for MalT-M311A/D312A), though the reduction was marginal for MalT-S279A, suggesting MalT N-terminal mutations destabilize the NOD module to different extents. Interestingly, the western-blot results showed that all the up-regulated mutants had higher levels of cellular proteins compared to WT, which was also observed in previous studies (Lisa et al., 2019; Liu et al., 2013). Among them, MalT-R9S and MalT-M311A displayed 4.5-fold and 3.5-fold higher protein levels, respectively, while the *in vivo* activity was approximately 14-fold and 12-fold higher than that of the WT, thus confirming these two are *bona fide* gain-of-function mutants.



**Figure 4.9 Mutations in Malt N-terminal region facilitate nucleotide exchange and lead to a higher protein activity. (A)** Malt autoinhibition is maintained by both inter- and intradomain interactions among NBD, HD, and WHD, primarily through polar contacts formed within the N-terminal loop region. **(B)**  $\beta$ -galactosidase assay showed mutations in Malt N-terminal region or WHD increased protein activity to various degrees, with R9S (14-fold) and M311A (12-fold) being the most prominent ones. **(C)** Luciferase assay showed mutations in Malt N-terminal region or WHD decreased the nucleotide-binding affinity of Malt to different levels, and lead to nucleotide exchange in the presence of ATP.

### 4.3.2 MalT N-terminal region contributes to protein oligomerization

Given the repressive function of the N-terminal segment of MalT, it is intuitive to assume that truncating it would lead to autoactivation. Surprisingly, the transcriptional activity of MalT was completely abolished merely by deleting the first 8 amino acids. Also, when assessing its repressive effect, two loss-of-function mutations were identified in the N-terminal region (K6A and R19A) (Figure 4.10 B). Taken together, this evidence suggests that the MalT N-terminal region may also have a role in regulating protein activation. To test this hypothesis, the oligomer formation of the N-terminal mutants was examined by gel filtration in the presence of 1 mM maltotriose and 0.1 mM ATP (Figure 4.10 A). The results showed that all mutants exhibited a lower degree of oligomerization compared to WT protein, revealing a function for the MalT N-terminal peptide in mediating inter-protomer interaction. Moreover, mutations of K6 and R19 that eliminated MalT activity *in vivo* render the protein inducer-insensitive and freeze it in the monomeric form. Luciferase assay suggested that K6 and R19, two important residues that are involved in oligomerization, may also be critical for the autoinhibition of MalT, as mutating either of them dramatically decreased the total nucleotide occupancy of the protein (66% for MalT-K6A and 35% for MalT-R19A) (Figure 4.10 C). Notably, K6 forms a hydrogen bond with D65 on the NBD, and mutations of D65 or its nearby residues have been shown to strongly enhance the transcriptional activity of the protein (Liu et al., 2013), likely by disrupting polar interactions and releasing the autoinhibitory effect. Furthermore, whilst mutations of D65 led to higher cellular protein levels, the amount of MalT-K6A and MalT-R19A proteins in the cell was much lower than that of the WT protein, as shown by western-blot, indicating MalT variants with a destabilized NOD module and cannot self-associate be most susceptible to proteolysis. In short, the MalT N-terminal region is not only important for autoinhibition but is also an essential structural element for protein oligomerization.

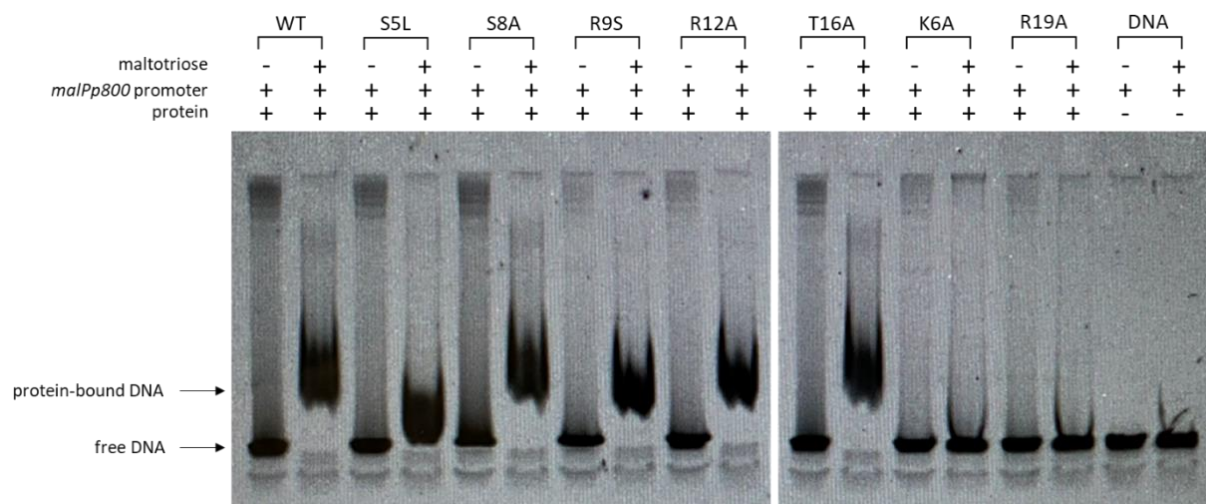


**Figure 4.10 MalT N-terminal mutants are defective in oligomerization. (A)** A same amount of WT or mutant proteins of MalT were subjected to gel filtration analysis with a Superose 6 Increase 10/300 GL column in the presence of 1 mM maltotriose and 0.1 mM ATP, all the mutants exhibited a lower degree of oligomerization compared to WT. **(B)**  $\beta$ -galactosidase assay showed the *in vivo* activity of MalT-K6A and MalT-R19A were strongly reduced compared to that of WT. **(C)** Luciferase assay showed MalT-K6A and MalT-R19A had a lower total nucleotide occupancy and were more prone to nucleotide exchange in the presence of ATP compared to WT.

#### 4.3.3 Low-order oligomerization of MalT is sufficient for DNA-binding

It has long been known that self-association is required for MalT to trigger transcriptional activation (Schreiber and Richet, 1999). Despite the compromised self-association ability of MalT N-terminal mutants, some of these mutants still displayed a high *in vivo* activity. To explain this *in vitro*, EMSA was performed to compare the DNA-binding capability of mutant and WT MalT proteins (Figure 4.11). A 108 bp DNA fragment derived from the semi-synthetic

MalT-dependent promoter *malPp800* (Danot and Raibaud, 1994) was used, and the assay was carried out as described in the Materials and methods section. In the absence of maltotriose, no nucleoprotein complex was formed and the promoter fragment migrated as fast as free DNA. When maltotriose was added during incubation, promoter bound by WT MalT protein migrated more slowly than free DNA, which was also observed for other up-regulated mutants as well as MalT-R12A, but not for MalT-K6A and MalT-R19A, two loss-of-function mutants that cannot oligomerize. This is consistent with the *in vivo* results that only MalT N-terminal mutants retaining DNA-binding capability are transcriptionally competent. Considering the number of MalT boxes present in promoters used for EMSA experiment and  $\beta$ -galactosidase assay (four and two, respectively), a low-order oligomerization of MalT involving two protomers or more may be sufficient to bind DNA and form a stable transcription initiation complex.

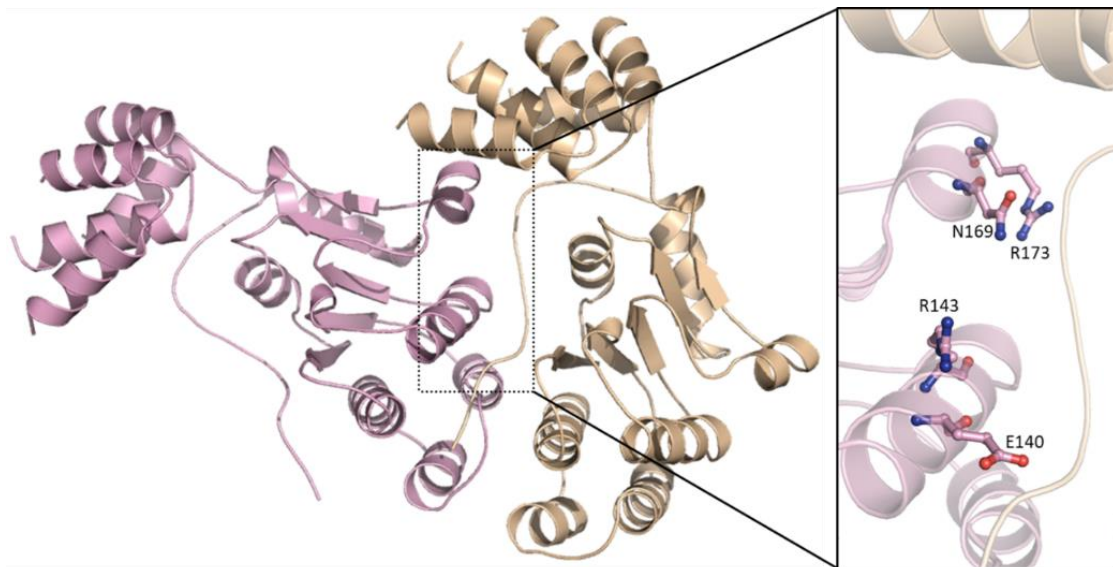


**Figure 4.11 Effect of MalT N-terminal mutations on promoter-binding capability of MalT.** 5  $\mu$ M WT or mutant proteins of MalT were incubated with 1  $\mu$ M Cy3-labelled promoter fragment of *malPp800* in the presence or absence of maltotriose, then run on native PAGE gels and detected with Cy3-specific filter. Results show that MalT N-terminal mutants with a compromised oligomerization capability could still bind *malPp800*, but MalT-K6A and MalT-R19A that lost self-association ability could not.

#### 4.3.4 MalT<sup>NBD-HD</sup> dimer modeled in its active form

During the activation of Apaf-1 or NLRC4, NBD and HD1 on the N-terminal move as a rigid body and undergo little change toward each other (Hu et al., 2015; Zhou et al., 2015), rotation of the NBD-HD1 subunit relative to WHD exposes the oligomerization interface in NOD and renders the protein competent for self-association. Assuming a similar domain

rearrangement, movement of WHD would then open up the NOD of MalT and generates an oligomerization surface, which should contain the N-terminal region. To examine this assumption, the NOD module of MalT was aligned to a lateral dimer of Apaf-1, and a dimeric MalT containing NBD and HD domains was modeled in its active form (Figure 4.12). As anticipated, the N-terminal peptide was found to be located right in between two protomers, in support of the concept that it mediates interprotomer interactions and is required for MalT oligomerization.



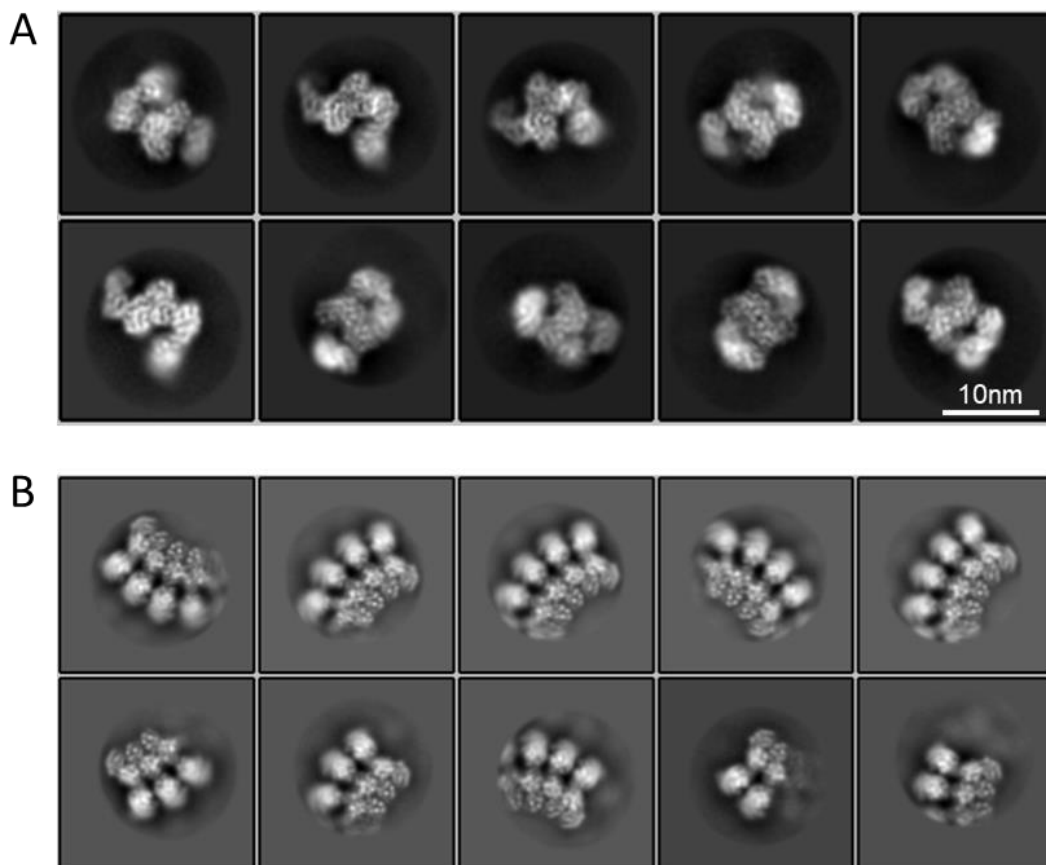
**Figure 4.12 A modeled MalT dimer containing NBD and HD domains.** In the modeled dimer structure, both the N-terminal region and MalY-interacting interface of MalT are located in the dimerization interface.

## 4.4 Reconstitution of MalT in its active form

### 4.4.1 Preliminary cryo-EM study of oligomeric MalT

To gain more insights into the activation mechanism of MalT, oligomeric protein samples were prepared for cryo-EM analysis following gel filtration. Copper grids were used for data collection after the initial sample quality check. In the presence of maltotriose and ATP, MalT forms ring-like structures of different lengths via its N-terminal (Figure 4.13 B), in agreement with a previous study (Larquet et al., 2004). However, all the particles adopt unidirectional orientation, with two neighboring protomers from one oligomer both facing the air-water interface. This preferred specimen orientation results in non-uniform distribution of angular projection views of the individual particles, making it difficult for 3D reconstruction of the

oligomer structure. I tried to address this problem by tilting the grid, but still could not get much improvement. Nevertheless, the average of the 2D classification obtained showed a higher resolution in comparison to the previously reported one (Larquet et al., 2004), from which more structural information could be gleaned (Figure 4.13 B). Based on the observation that promoter DNA interacts with the outer layer of MalT oligomer (Larquet et al., 2004), the DNA-binding domain of MalT should be located in the peripheral of the ring-like structure, which corresponds to the C-terminal region.



**Figure 4.13 2D class averages of different MalT samples. (A)** Cryo-EM analysis of the MalT-MalY complex after 2D classification. **(B)** Cryo-EM analysis of the MalT in its active form after 2D classification. This analysis was done by Yue Sun from Tsinghua University, Beijing.

#### 4.4.2 Activation of MalT stabilizes its C-terminal domain for binding DNA

When the 2D class averages of inactive and active MalT samples were compared, density of the C-terminal region was only found in oligomeric samples, indicating the sensor domain together with the DNA-binding domain are flexible prior to activation when MalT is complexed with MalY, and are stabilized upon oligomerization, a prerequisite for promoter binding and transcription initiation (Figure 4.13). The binding of maltotriose may also have a structural role

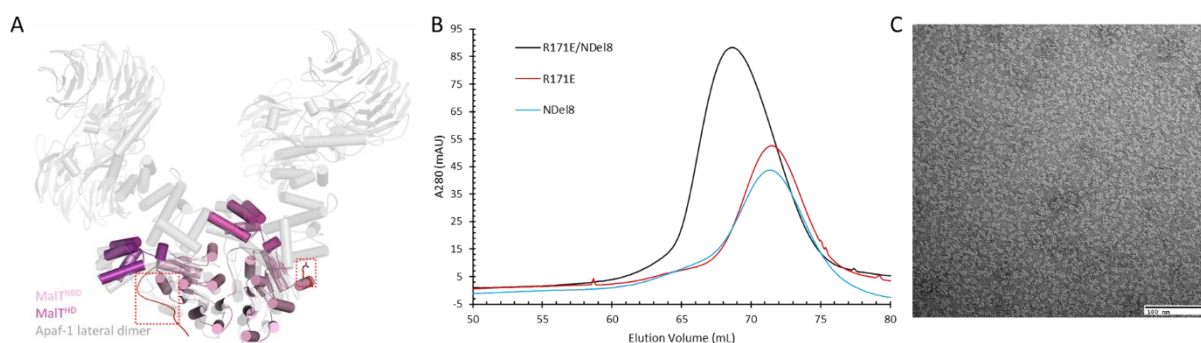


in stabilizing the C-terminal domains of MalT, since MalT-W317R, a MalT mutant that can oligomerize in the absence of inducers, is still responsive to maltotriose in an *in vitro* transcription system (Liu et al., 2013).

#### 4.4.3 An active MalT dimer

In the flagellin-bound NAIP5-NLRC4 complex structure, an NLRC4 mutant carrying R288A and L435D was used instead of WT NLRC4 (Yang et al., 2018). This mutant can interact with NAIP via its receptor surface, but has a defective catalytic surface, preventing it from self-propagating when activated (Hu et al., 2015). In an attempt to circumvent the preferred orientation problem that occurred in the cryo-EM samples of active MalT, I tried to restrict the self-association of MalT using a similar strategy, and reconstituted a MalT dimer consisting of MalT-R171E (plasmid provided by Evelyne Richet from Institut Pasteur, Paris) and MalT-NDe18, a MalT variant with its N-terminal eight amino acids deleted. Residue R171 and the N-terminal region of MalT are predicted to be located at two oligomerization interfaces on opposite sides of NBD (Marquenet and Richet, 2010; this study), thus two mutant proteins carrying either of these two mutations can interact with each other but not with themselves, generating a dimeric complex upon activation (Figure 4.14 A).

To purify the dimeric complex, MalT-R171E and MalT-NDe18 were purified independently, incubated at a molar ratio of 1:1 in the presence of 1 mM maltotriose and 0.1 mM ATP, then filtered through gel filtration. Formation of the dimer was confirmed by comparing the resulting profile to one of the two mutant proteins being filtered alone (Figure 4.14 B). Dimeric protein samples were then examined by negative staining EM, which appeared to be homogeneous and were suitable for further cryo-EM analysis (Figure 4.14 C).



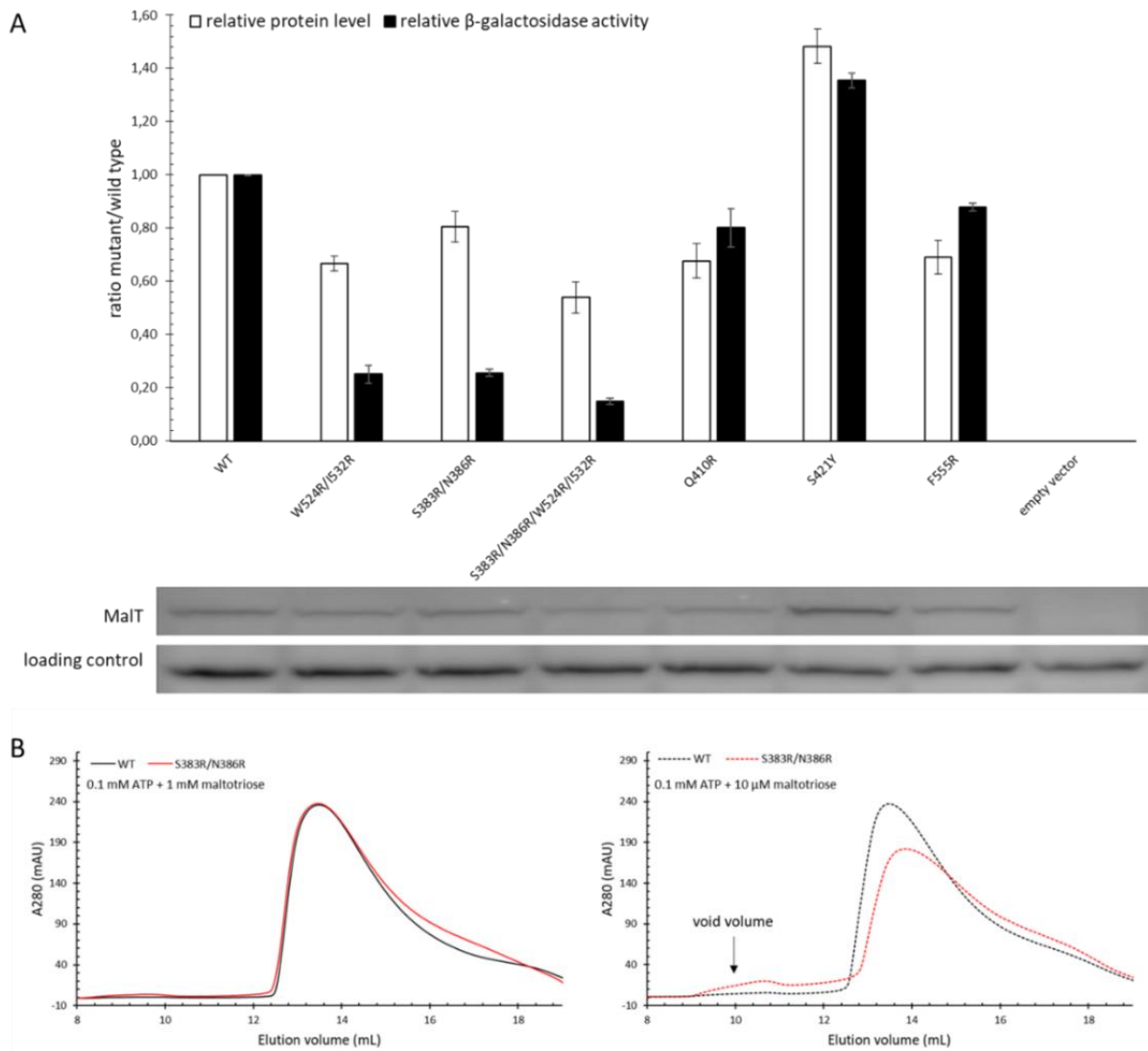
**Figure 4.14 Reconstitution of a MalT dimer. (A)** A dimeric MalT structure modeled based on an active lateral dimer of Apaf-1. Residue R171 and the N-terminal region of MalT are highlighted. **(B)** A same amount of MalT-R171E and MalT-NDe18 proteins were analyzed by gel



filtration with a Hiloal 16/600 Superdex 200 column in the presence of 1 mM maltotriose and 0.1 mM ATP, either separately or together after incubation. **(C)** A negative staining EM image for the MalT-R171E/NDel8 dimer sample.

#### 4.4.4 Mapping of the maltotriose-binding site based on a predicted MalT structure

To probe the mechanism by which MalT binds maltotriose, a MalT structure including its N-terminal domains and part of the sensor domain was predicted using AlphaFold2, and the putative maltotriose-binding site was mapped on this structure based on the relatively conserved ligand recognition pattern between ZAR1 and Sr35 (Förderer et al., 2022; Wang et al., 2019a). Consistent with previous studies, the arm and sensor domain together facilitate the formation of this binding site (Danot, 2015), and residues surrounding the binding site which span over both domains were identified (S383, N386, Q410, S421, W524, I532, F555). In the hope that substituting these residues with bulky ones would impede ligand binding, I measured the *in vivo* activity of the maltotriose-binding site mutants and compared it with that of WT protein (Figure 4.15 A). Two double mutants, S383R/N386R and W524R/I532R exhibited lower transcriptional activity, while the *in vivo* activity of the quadruple mutant carrying four of these substitutions was reduced to about 15% of that of the WT control. Although a decrease in cellular protein levels was observed in these mutants, the decrease was much smaller compared to the reduction of activity *in vivo*.



**Figure 4.15 The *in vitro* and *in vivo* effect of putative maltotriose-binding site mutations on MalT activity.** (A) Residues involved in maltotriose binding of MalT were mapped based on comparisons of a modelled MalT structure with ZAR1 and Sr35. *In vivo* assay showed MalT activity was strongly reduced by mutations of S383/N386/W524/I532. (B) 5  $\mu$ M proteins of WT or MalT-S383R/N386R were incubated and filtered through a Superose 6 Increase 10/300 GL column in the presence of 0.1 mM ATP and different concentrations of maltotriose. When 1 mM maltotriose was used, the oligomerization of WT and MalT-S383R/N386R proteins was almost the same. When maltotriose concentration was reduced to 10  $\mu$ M, MalT-S383R/N386R was less competent in terms of oligomerization and formed more aggregates compared to WT protein, indicating double mutant MalT-S383R/N386R is less sensitive to maltotriose than WT.

Alongside these observations, gel filtration assays were performed using the mutant proteins, showing a reduced binding affinity for maltotriose (Figure 4.15 B). Unfortunately, only MalT-S383R/N386R could be purified in the presence of maltotriose, while the mutant proteins carrying W524R/I532R were unstable under the same conditions. When 1 mM maltotriose was used for incubation and gel filtration, MalT-S383R/N386R could oligomerize similarly to

WT and gave essentially the same profile. However, when the concentration of maltotriose was reduced to 10  $\mu\text{M}$ , MalT-S383R/N386R formed smaller oligomers compared to WT, and a peak appeared at the void volume. When maltotriose concentration was further reduced to 1  $\mu\text{M}$ , i.e. below the protein concentration used for the assay (5  $\mu\text{M}$ ), both MalT-S383R/N386R and the WT protein gave rise to a peak at the void volume in gel filtration, indicating the unliganded protein tended to aggregate under the assay conditions. Taken together, the results of *in vitro* experiments support that residues S383 and N386 contribute to maltotriose binding and their mutation decreases the responsiveness of MalT to maltotriose.

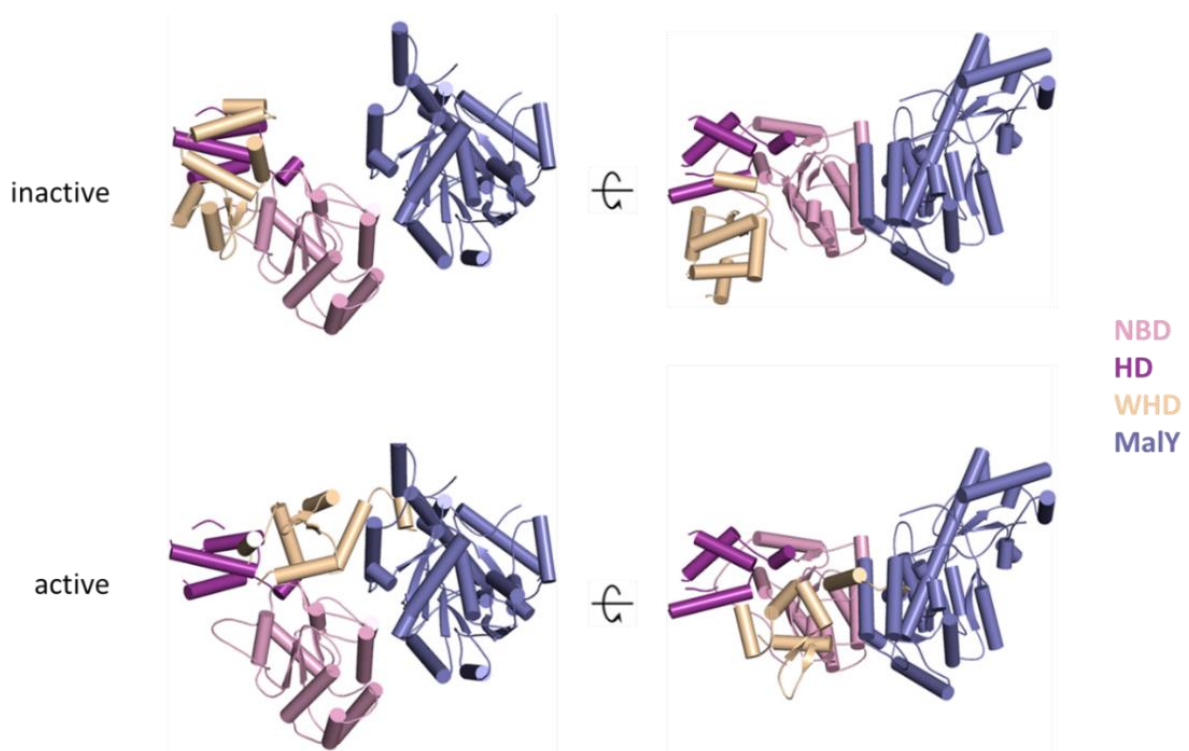
## 5. Discussion

### 5.1 Inhibition of an NLR by blocking its oligomerization

Preventing activation of an NLR protein by masking its oligomerization interface is a recurrent theme in NLR inhibition. In the inactive NLRC4 structure, while the receptor surface composed of NBD and HD1 is preexistent and largely exposed, HD2 and LRR fold back and contact the N-terminal domains, resulting in the burial of the catalytic surface which is required for initiating NLRC4 autoactivation (Hu et al., 2013; Hu et al., 2015). Conformational changes during NLRC4 activation generate this catalytic surface consisting of structural elements from both NBD and WHD, enabling interprotomer interactions between two neighboring NBDs, and WHD and HD1 (Hu et al., 2013; Hu et al., 2015). NLRC4 autoinhibition is reminiscent of CED-9 mediated inhibition of CED-4, though CED-4 lacks the C-terminal LRR domain. Unlike NLRC4 or Apaf-1, CED-4 exists in a constitutively active conformation and forms an asymmetrical dimer prior to activation. The dimeric CED-4 is blocked from further oligomerization by CED-9 (Yan et al., 2005). Structural study showed that CED-9 interacts with one CED-4 monomer through the CARD and NBD of CED-4, sequestering the CED-4 dimer from further oligomerization. Binding of EGL-1 to CED-9 allosterically disassociates CED-9 from the CED-4 dimer, rendering the CED-9-blocked oligomerization interface of CED-4 accessible for another CED-4 molecule to form CED-4 apoptosome (Qi et al., 2010; Yan et al., 2005).

In the MalT-MalY inhibitory complex, the recognition interface of MalY is formed by helices  $\alpha 7$  and  $\alpha 8$  from the NBD of MalT. Structural comparison revealed that this MalY-recognizing interface overlaps with the one MalT oligomerization surface (Figure 4.7). Thus, MalY must be displaced in order for MalT to self-associate. This result indicates functional similarity between MalY, CED-9, and the C-terminal domains of NLRC4 in inhibiting NLR activation. CED-4 activation is dependent on EGL-1 binding and the NLRC4 activation is induced by PrgJ-bound NAIP2 or flagellin-bound NAIP5. However, how MalT is released from MalY sequestration for MalT activation remained enigmatic. The available structural information showed that conserved domain reorganization occurs to NLR during activation (Hu et al., 2015; Wang et al., 2019a; Zhou et al., 2015). It is reasonable to assume that similar structural remodeling is also required for MalT activation. However, modeling study using Apaf-1 as the template indicated

that the recognition interface of MalY on MalT remains intact in the modeled structure of active MalT. Furthermore, the WHD of inactive MalT is distant from MalY in the inactive complex structure. These results suggest that the positioning of the WHD could sterically hinder interaction of the activated MalT with MalY. Supporting this hypothesis, structural comparison between the cryo-EM structure of inactive MalT and the modeled active MalT shows that WHD would clash with MalY upon MalT oligomerization. This structural observation further strengthens the idea that MalY preferentially binds MalT in its resting state (Figure 5.1).



**Figure 5.1 MalY preferentially binds MalT in its resting state.** The upper panel: the NOD module of an inactive MalT in complex with an MalY protomer at different angles. The lower panel: same as the upper panel, except that the WHD is modeled in its active state based on the structure of active Apaf-1, which sterically clashes with MalY.

Since MalT activation is known to be triggered by maltotriose, one model for the maltotriose-induced release of MalY would be that increases in the intracellular maltotriose levels lead to release MalY, consequently resulting in MalT activation. However, MalT alone remains inactive in the absence of maltotriose, why is MalY needed to inhibit MalT? *malX* from the *malX malY* operon encodes the EIIC of the PTS, a sugar-specific permease that catalyzes

the transport of glucose and maltose. This might cause accidental activation of the *malX malY* operon (Reidl and Boos, 1991) if the activity of MalT were not held in check by other components. Thus, MalY might evolve to prevent accidental activation of MalT by the substrate transported from the PTS. Consistent with this hypothesis, structural comparison showed that MalY is positioned similarly to the C-terminal LRR domain of inactive NLRC4. This result further strengthens our conclusion that MalY functions to sequester MalT in a monomeric state.

## 5.2 A dual role for the very N-terminal region of MalT in its activation

The MalT N-terminal domain is not involved in interaction with downstream signaling components like that of other inflammasome-forming proteins. The very first 20 amino acids of MalT forms a loop that is sandwiched between NBD and WHD, tightly stabilizing the NOD module via five interdomain hydrogen bonds and seven intradomain hydrogen bonds like a double-sided tape (Figure 4.9 A). This is a unique feature of MalT, since in many NLRs the interaction between NBD and WHD is only mediated by an ADP molecule without the involvement of other structural elements as revealed by the inactive structures of NLRC4, NOD2, Zar1, and the MalT homologue protein PH0952 (Hu et al., 2013; Lisa et al., 2019; Maekawa et al., 2016; Wang et al., 2019b). As discussed above, upon MalT activation, WHD would undergo structural re-organization in relation to NBD. Thus, the NBD/WHD interaction mediated by the N-terminal peptide is predicted to act to stabilize the inactive conformation of MalT. Thus, disrupting either inter- or intradomain interactions formed along the N-terminal loop is able to destabilize the NOD module. Indeed, all the mutations at the NBD/WHD interface of MalT (MalT-S5L, MalT-M311A, and MalT-M312A) as well as those within NBD (MalT-K6A, MalT-R9S, MalT-R19A) could promote MalT activation, confirming the role of the N-terminal region in MalT autoinhibition. In the autoinhibited Apaf-1 structure, the N-terminal domain is also important for maintaining the resting form. The CARD of Apaf-1 stacks closely against NBD and WHD, leading to the burial of the procaspase-9-binding interface. Meanwhile, exchange of the bound ADP molecule, the organizing center for inactive

Apaf-1, is restricted due to the packing of the CARD of Apaf-1 against other domains (Riedl et al., 2005). An inhibitory role of the N-terminal domain has also been proposed for human NLRP1 following genetic evidence. Gain-of-function mutations related to skin disorders are located within the N-terminal PYD domain, and overexpression of these mutant proteins can significantly enhance inflammasome formation *in vitro* (Zhong et al., 2016).

In several structures of activated NLRs, the loop region N-terminal to the NBD is involved in mediating oligomerization of these NLRs. For example, the N-terminal loop region N-terminal to the NBD of ZAR1 mediates protein oligomerization and positively regulates the cell-death activity, and mutation in this region disrupts ZAR1-mediated disease resistance (Wang et al., 2019a). Based on the modeled structure of the MalT<sup>NBD-HD</sup> dimer, K6 is situated on the dimeric interface and likely to mediate MalT oligomerization through direct interaction, while R19 is located further away towards HD and probably does not contact the neighboring protomer when oligomerized. In support of these structural observations, substitution of K6 or R19 with alanine both abolished MalT self-association, in contrast with mutations at the other positions of the N-terminal loop region. In the inactive MalT structure, R19 hydrogen bonds with G44, a residue located on the P-loop and interacts with the  $\beta$ -phosphate group of the MalT-bound ADP. Substitution of R19 could affect the conformation of the P-loop, thereby reducing the nucleotide binding affinity of MalT and leading to an altered oligomerization interface. Therefore, the role of R19 in MalT oligomerization should be indirect. This idea is supported by the observation that the yield of recombinantly expressed MalT-R19A is much lower than MalT-K6A during protein purification. One reason of this may be because of a destabilized NOD module of MalT-R19A.

### 5.3 Mechanism of ligand-induced MalT DNA-binding

In the cryo-EM structure of the MalY-MalT complex, the C-terminal domains of MalT are largely flexible. This is evident from the 2D averages of the complex particles (Figure 4.13). Interestingly, the 2D averages of the particles from the activated MalT clearly showed the C-terminal domains, indicating that these domains become stabilized upon activation. This is consistent with a previous study, which discovered that besides its role in relieving

autoinhibition, maltotriose is also required for MalT DNA-binding (Liu et al., 2013). In the same study, an autoactive mutant of MalT carrying W317R mutation was isolated by genetic screen. W317 is located at the NBD/WHD interface in the inactive MalT structure, and is not directly involved in interdomain interactions. However, substitution of W317 by an arginine residue could lead to severe steric clash with residues from the NBD side, thereby promoting NOD opening and MalT activation. The mutant protein is competent for oligomerization in the presence of ATP alone, bypassing the requirement of maltotriose-binding for conformational changes, but is strongly defective in the activation of transcription initiation in a purified *in vitro* system. Using DNase I protection assay, it was found that the binding affinity of MalT-W317R to promoter DNA was very low when only ATP was present, and the affinity increased significantly upon addition of maltotriose. These results suggest that that an inducer-bound sensor is essential for the dsDNA-binding domain to adopt a correct configuration, and to effectively bind the MalT boxes at the target promoter following oligomer formation (Liu et al., 2013). It may be that maltotriose binding acts to stabilize the C-terminal domains for dsDNA binding, but the underlying mechanism remains unknown.

The ligand-binding domain of NLRs has been reported to positively regulate protein activity in several instances. In the NLRC4 inflammasome, LRRs from two adjacent protomers contact each other through complementary charged surfaces, contributing to the assembly of the wheel-like structure (Hu et al., 2015). Similar contact exists in the RPP1 resistosome as well, in which two neighboring LRRs stabilize the RPP1 tetramer by direct interaction (Ma et al., 2020). In addition, for the ZAR1 resistosome, LRR is also involved in ZAR1 oligomerization by interacting with HD from a neighboring protomer (Wang et al., 2019a). More recently, the structure of the Sr35 wheat resistosome was solved. Though the LRR domain of Sr35 does not directly participate in the pentamerization of the resistosome, it forms strong contacts with the 'EDVID motif' in the CC domain (Rairdan et al., 2008; Wróblewski et al., 2018). The intramolecular interactions between these two domains are required for Sr35-mediated cell death, and may reflect an evolutionarily conserved mechanism for stabilization of resistosomes formed by CC-domain containing NLRs (Förderer et al., 2022). Combining these knowledges with our structural observations of MalT, it is reasonable to propose that the sensor domain of MalT loaded with maltotriose may promote MalT oligomerization by forming intermolecular interactions between two adjacent protomers, which somehow



maintains the C-terminal DNA-binding domain of MalT in register with promoter dsDNA for transcription initiation.

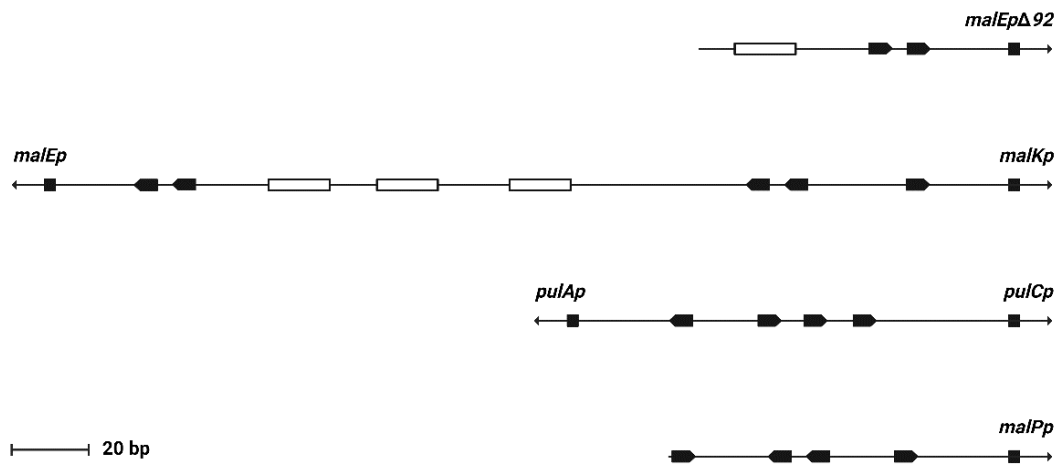
The positioning of the effector domain makes MalT unique among STAND family members, in the sense that its function in signal transduction is directly influenced by the conformation of the sensor domain. For other NLR proteins that can form inflammasomes, the downstream signaling components are recruited via the N-terminal CARD or PYD domain which forms the central hub upon inflammasome assembly, either directly or indirectly (Yin et al., 2015). In the case of plant resistosomes, the nucleated N-terminal CC and TIR domains confer cation channel and NADase activity, respectively, and are less-likely to be directly influenced by the ligand-binding state of the sensor domain compared with MalT (Ma et al., 2020; Wang et al., 2019a). Indeed, constitutively active mutants that bypass the requirement of their cognate ligands for downstream signaling have been identified for both NLRC4 and ZAR1 (Hu et al., 2013; Wang et al., 2019b). Besides MalT, human NLRP1 is another NLR protein that harbors a C-terminal effector domain and has been extensively studied, yet a clear distinction between NLRP1 and MalT is that the oligomerization of NLRP1 is mediated by its C-terminal domains. The UPA and CARD of NLRP1 are released from the N-terminal polypeptide chain after autoproteolysis, forming a two-layered filament with the CARD of NLRP1 being clustered in the center and drives the formation of ASC speck, a process which does not directly involve the LRR domain of NLRP1 (Gong et al., 2021; Hollingsworth et al., 2021). Taken together, ligand-induced intermolecular interactions stabilize MalT for dsDNA-binding, and as pointed out by Liu et al. (Liu et al., 2013), this could effectively keep the protein activity in check after oligomerization, which may represent a general mechanism for STAND proteins that possess a C-terminal effector domain.

## 5.4 MalT oligomeric states vs its transcriptional activity

When analyzing oligomerization activity of MalT mutants and their *in vivo* activity, an intriguing observation was that higher-order oligomerization does not necessarily result in higher transcription activity of MalT. In the strains used for  $\beta$ -galactosidase assay, the promoter of *lacZ* gene was replaced by *malEp $\Delta$ 92*, a MalT-dependent promoter harboring two

MalT-binding motifs proximal to the transcription start site of *malE* (Raibaud et al., 1989). Since MalT N-terminal mutants with a reduced oligomeric size still display a high activity *in vivo*, it is likely that an active MalT trimer or even dimer could be sufficient for transcription initiation, as long as a stable nucleoprotein complex comprising *malEp* $\Delta$ 92/MalT/CAP/RNAP can be formed. This hypothesis can be tested by comparing the binding affinity of a reconstituted MalT-R171E/NDel8 dimer with WT MalT oligomer for *malEp* $\Delta$ 92. As neither MalT-R171E nor MalT-NDel8 alone is transcriptionally competent, it would also be interesting to test if these two proteins can rescue the transcription activity in a purified *in vitro* system.

The dependence of MalT transcriptional activity on its oligomerization states could be varied with the promoters it binds (Figure 5.2). For instance, when a WT *malEp* promoter is used for the *in vivo* assay, a large oligomer formed by WT MalT is expected to demonstrate a higher transcriptional activity than mutant protein with an impaired self-association ability. The reason for this is that the WT *malEp* promoter contains five MalT boxes that span over a 240 bp region upstream of the *malE* transcription start site. Deletion of MalT boxes from the distal one results in a progressive reduction of the promoter activity, and *malEp* $\Delta$ 92 is about 20% as efficient as the WT promoter (Raibaud et al., 1989; Richet, 2000; Richet and Raibaud, 1991). Assuming each MalT box in the promoter region is recognized by one MalT protomer, a WT *malEp* promoter bound by a MalT oligomer containing less than five protomers should phenocopy the truncated *malEp* promoter with a reduced promoter activity. Furthermore, for the full-length *malKp* promoter, higher-order oligomerization of MalT may be a prerequisite for transcription activation. *malKp* shares the same set of MalT boxes with *malEp* but has an opposite direction. Both the proximal and the distal MalT boxes (the first one and the fifth one upstream of the *malK* transcription start site, respectively) are required for *malKp* activation, and the promoter activity is completely abolished by just deleting the distal one (Raibaud et al., 1989). Therefore, in the case of *malKp*, an active MalT trimer or tetramer may not be able to initiate transcription activation, because a simultaneous binding of the proximal and the distal MalT boxes would be more difficult to achieve. In addition, larger MalT oligomers could be favored in other natural MalT-dependent promoters such as *malPp* or *pulAp/pulCp* from *K. pneumoniae*, since cooperative non-specific dsDNA-binding of MalT may facilitate nucleoprotein complex formation (Danot and Raibaud, 1994; Vidal-Ingigliardi et al., 1991).



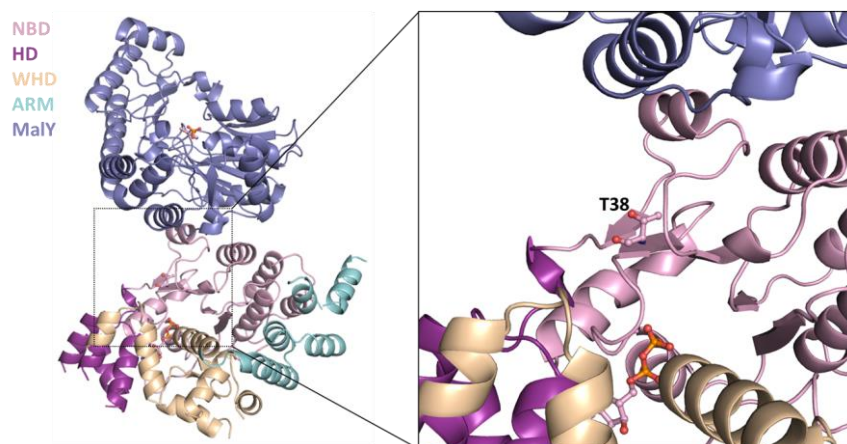
**Figure 5.2 A schematic view of different MalT-dependent promoters.** MalT-binding boxes are represented as filled black arrows, CAP-binding sites are represented as white rectangles, and the Pribnow boxes are represented as black squares. *malEp-malKp* and *pulAp-pulCp* are two pairs of divergent promoters, whose regulatory region within each is largely shared by two genes transcribed reversely. Figure adapted from Vidal-Ingigliardi, Richet et al. (Vidal-Ingigliardi et al., 1991).

## 5.5 MalT stability vs its protein conformation

When assessing the effect of MalT N-terminal mutations on its transcriptional activity, one difficulty was to disentangle the *in vivo* activity from protein stability, since all the MalT N-terminal mutants with upregulated transcriptional activity displayed an enhanced cellular protein amount, though the increases in protein levels were generally less significant than the changes of protein activities (Figure 4.9 B). Compared with previously reported MalT mutations, it seems that all the gain-of-function mutants are more resistant to proteolysis than WT (Lisa et al., 2019; Liu et al., 2013). Meanwhile, MalT variants with a destabilized NOD but cannot oligomerize are most susceptible to proteolysis, such as MalT-K6A and MalT-R19A, whose cellular protein levels are much lower than WT. The reason behind this obvious discrepancy is not clear yet. It has been shown that depending on the ligand binding state, MalT can generate different digestion patterns when subjected to limited proteolysis (Danot, 2001). Therefore, the stability of MalT might also differ due to conformational changes in the protein.

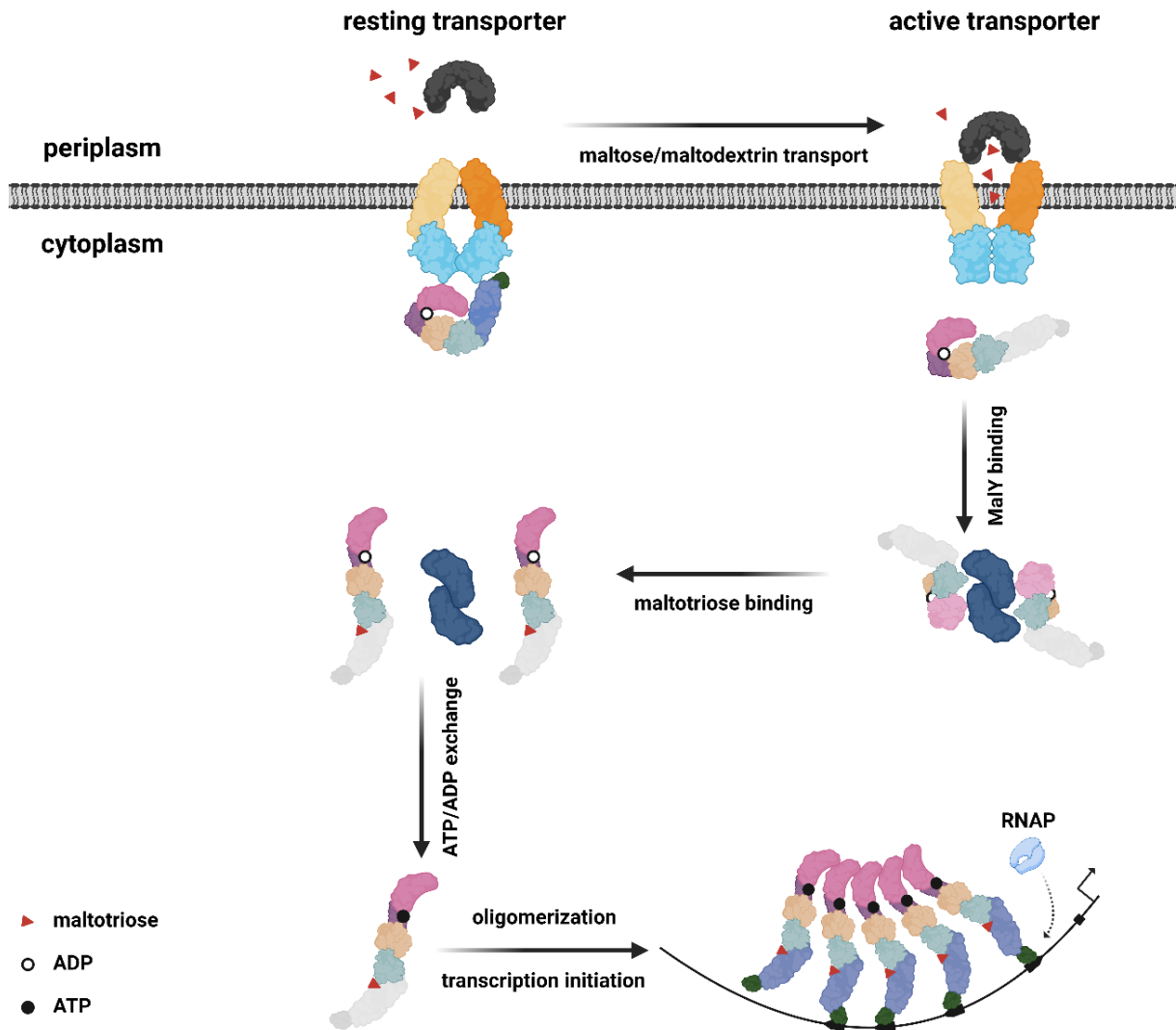
## 5.6 A model for MalT regulation and comparison with other STAND proteins

Combining the results from this and previous studies, a mechanistic model regarding the regulation of MalT can be deduced (Figure 5.4). When no substrate is transported through the maltose-maltodextrin transporter, MalT is anchored to the membrane via its NBD and C-terminal domains. At this stage, the sensor domain of MalT is stabilized by interacting with NBD and MalK (Lisa et al., 2019). Transport of sugar frees MalT from membrane, hence the unliganded sensor domain may become flexible. MalT is then sequestered by other inhibitory proteins, such as MalY and Aes, to prevent its accidental activation. The action of MalK, MalY, and Aes is unlikely to be simultaneous in the same MalT molecule, since the repressions of MalT by MalK and Aes are both mediated by residue T38 from MalT (Joly et al., 2004; Joly et al., 2002). Based on our structure, MalK and Aes would sterically clash with MalY if they bind MalT through this residue on NBD. When MalT is complexed with MalY, two oligomerization surfaces on NBD are both unrevealed due to MalY sequestration and autoinhibition mediated by interdomain interactions. Binding of maltotriose to the sensor of MalT triggers conformational changes that allosterically disengage the inhibitory contacts and renders MalT oligomerization-competent, in the meantime completing an inducer-binding site with high-affinity and stabilizing the effector domain for downstream signaling (Danot, 2015). The MalT-MalY complex structure here provides new insights into the inhibition mechanism of a canonical STAND by a repressor protein, and sheds light on the mode of activation and transcriptional regulation for other STAND family proteins.



**Figure 5.3 Position of residue T38 in the MalT-MalY inhibitory complex.**

Over the past years, oligomeric structures of different STAND proteins have been reported, providing knowledge regarding STAND activation and downstream signaling. Following Apaf-1 activation, the apoptosome recruits the initiator caspase procaspase-9 through CARD-CARD interaction on the central hub, leading to the formation of a holoenzyme and caspase 3 cleavage (Cheng et al., 2016; Li et al., 2017). Other animal NLRs, such as NLRC4 or NLRP6, assemble into inflammasomes upon activation (Hu et al., 2015; Shen et al., 2019). The CARD or PYD of these inflammasome-forming proteins are thought to be positioned in the middle and interact with ASC to initiate ASC polymerization, which further recruits procaspase-1. Though the NLRC4 inflammasome is able to recruit procaspase-1 directly and induce pyroptosis in *asc*-deficient cells, ASC speck formation is still required for cytokine maturation, suggesting the polymerization of adaptor protein amplifies the signal derived from ligand-binding for a robust immune response (Broz and Dixit, 2016). Similar signal-propagation mechanism also exists in RIP2 signaling, whereby RIP2 can be activated by NOD1 or NOD2 and forms filamentous structure (Gong et al., 2018). Plant NLRs are known to form resistosomes when activated, with different signaling output depending on the type of the N-terminal domain. Pentamerization of CC domain-containing NLRs mediates cell death by forming a calcium-permeable cation channel, as exemplified by ZAR1 and Sr35-AvrSr35 resistosomes, while activation of TIR-domain containing NLRs such as ROQ1 and RPP1 generates NADase activity required for TIR-mediated cell death (Förderer et al., 2022; Ma et al., 2020; Martin et al., 2020; Wang et al., 2019a). Different from eukaryotic STANDs whose effector domains are often found to be centrally clustered after oligomerization, the DNA-binding domains of oligomeric MalT spread out in a radial pattern, facilitating synergetic binding of downstream components including the promoter and RNAP (Danot et al., 1996; Larquet et al., 2004). These distinct roles exhibited by STAND proteins that form different macromolecular signaling complexes demonstrate the functional versatility of this protein family with a conserved domain architecture.



**Figure 5.4 A model for MalT regulation.** In the absence of substrate transportation, the *E. coli* maltose transporter MalFGK2 anchors monomeric MalT to plasma membrane via MalK-mediated interactions. Binding of MBP to MalFGK2 and sugar uptake triggers a conformational change in MalK, resulting in the release of MalT into the cytoplasm (Chen et al., 2003; Oldham and Chen, 2011; Richet et al., 2012). MalT then binds to inhibitory proteins such as MalY in the cytoplasm and remains in a resting state. The inhibitory effect of MalY can be released by binding of maltotriose, leading to structural remodeling of MalT and nucleotide exchange, exposing two surfaces on MalT N-terminal domains for oligomerization. Maltotriose binding and oligomerization together stabilize the C-terminal domains of MalT, allowing binding to dsDNA and recruitment of RNAP and subsequent transcription initiation. Components of the MalFGK2 transporter MalF, MalG, MalK are colored in light yellow, orange, and light blue, respectively. NBD, HD, WHD, ARM, the sensor domain, and the DNA-binding domain of MalT are colored in pink, purple, wheat, cyan, slate, and green, respectively. Domains that are flexible are colored in light gray. MalY dimer is colored in dark blue, MBP is colored in black.

## 6. References

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D, Biological crystallography* *66*, 213-221.
- Andreeva, L., David, L., Rawson, S., Shen, C., Pasricha, T., Pelegrin, P., and Wu, H. (2021). NLRP3 cages revealed by full-length mouse NLRP3 structure control pathway activation. *Cell* *184*, 6299-6312.e6222.
- Bauernfeind, F.G., Horvath, G., Stutz, A., Alnemri, E.S., MacDonald, K., Speert, D., Fernandes-Alnemri, T., Wu, J., Monks, B.G., Fitzgerald, K.A., *et al.* (2009). Cutting edge: NF-kappaB activating pattern recognition and cytokine receptors license NLRP3 inflammasome activation by regulating NLRP3 expression. *Journal of immunology (Baltimore, Md : 1950)* *183*, 787-791.
- Bauernfried, S., and Hornung, V. (2021). DPP9 restrains NLRP1 activation. *Nature structural & molecular biology* *28*, 333-336.
- Bedouelle, H. (1983). Mutations in the promoter regions of the malEFG and malK-lamB operons of Escherichia coli K12. *Journal of molecular biology* *170*, 861-882.
- Bedouelle, H., Schmeissner, U., Hofnung, M., and Rosenberg, M. (1982). Promoters of the malEFG and malK-lamB operons in Escherichia coli K12. *Journal of molecular biology* *161*, 519-531.
- Bi, G., Su, M., Li, N., Liang, Y., Dang, S., Xu, J., Hu, M., Wang, J., Zou, M., Deng, Y., *et al.* (2021). The ZAR1 resistosome is a calcium-permeable channel triggering plant immune signaling. *Cell* *184*, 3528-3541.e3512.
- Boos, W., and Shuman, H. (1998). Maltose/maltodextrin system of Escherichia coli: transport, metabolism, and regulation. *Microbiology and molecular biology reviews : MMBR* *62*, 204-229.
- Brenner, S., Jacob, F., and Meselson, M. (1961). An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* *190*, 576-581.
- Broz, P., and Dixit, V.M. (2016). Inflammasomes: mechanism of assembly, regulation and signalling. *Nature reviews Immunology* *16*, 407-420.
- Chai, J., and Shi, Y. (2014). Apoptosome and inflammasome: conserved machineries for caspase activation. *National Science Review* *1*, 101-118.
- Chapon, C. (1982). Role of the catabolite activator protein in the maltose regulon of Escherichia coli. *Journal of bacteriology* *150*, 722-729.
- Chapon, C., and Kolb, A. (1983). Action of CAP on the malT promoter in vitro. *Journal of bacteriology* *156*, 1135-1143.
- Chavarría-Smith, J., and Vance, R.E. (2015). The NLRP1 inflammasomes. *Immunological reviews* *265*, 22-34.
- Chen, J., Lu, G., Lin, J., Davidson, A.L., and Quijcho, F.A. (2003). A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. *Molecular cell* *12*, 651-661.
- Cheng, T.C., Hong, C., Akey, I.V., Yuan, S., and Akey, C.W. (2016). A near atomic structure of the active human apoptosome. *eLife* *5*.
- Clausen, T., Schlegel, A., Peist, R., Schneider, E., Steegborn, C., Chang, Y.S., Haase, A., Bourenkov, G.P., Bartunik, H.D., and Boos, W. (2000). X-ray structure of MalY from Escherichia coli: a pyridoxal 5'-phosphate-dependent enzyme acting as a modulator in mal gene expression. *The EMBO journal* *19*, 831-842.

- Cole, S.T., and Raibaud, O. (1986). The nucleotide sequence of the malT gene encoding the positive regulator of the Escherichia coli maltose regulon. *Gene* *42*, 201-208.
- Danot, O. (2001). A complex signaling module governs the activity of MalT, the prototype of an emerging transactivator family. *Proceedings of the National Academy of Sciences of the United States of America* *98*, 435-440.
- Danot, O. (2015). How 'arm-twisting' by the inducer triggers activation of the MalT transcription factor, a typical signal transduction ATPase with numerous domains (STAND). *Nucleic acids research* *43*, 3089-3099.
- Danot, O., and Raibaud, O. (1994). Multiple protein-DNA and protein-protein interactions are involved in transcriptional activation by MalT. *Molecular microbiology* *14*, 335-346.
- Danot, O., Vidal-Ingigliardi, D., and Raibaud, O. (1996). Two amino acid residues from the DNA-binding domain of MalT play a crucial role in transcriptional activation. *Journal of molecular biology* *262*, 1-11.
- Debarbouille, M., and Schwartz, M. (1979). The use of gene fusions to study the expression of malT the positive regulator gene of the maltose regulon. *Journal of molecular biology* *132*, 521-534.
- Decker, K., Plumbridge, J., and Boos, W. (1998). Negative transcriptional regulation of a positive regulator: the expression of malT, encoding the transcriptional activator of the maltose regulon of Escherichia coli, is negatively controlled by Mlc. *Molecular microbiology* *27*, 381-390.
- Ehrmann, M., and Boos, W. (1987). Identification of endogenous inducers of the mal regulon in Escherichia coli. *Journal of bacteriology* *169*, 3539-3545.
- Elliott, E.I., and Sutterwala, F.S. (2015). Initiation and perpetuation of NLRP3 inflammasome activation and assembly. *Immunological reviews* *265*, 35-52.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta crystallographica Section D, Biological crystallography* *66*, 486-501.
- Finger, J.N., Lich, J.D., Dare, L.C., Cook, M.N., Brown, K.K., Duraiswami, C., Bertin, J., and Gough, P.J. (2012). Autolytic proteolysis within the function to find domain (FIIND) is required for NLRP1 inflammasome activity. *The Journal of biological chemistry* *287*, 25030-25037.
- Förderer, A., Li, E., Lawson, A., Deng, Y.-n., Sun, Y., Logemann, E., Zhang, X., Wen, J., Han, Z., Chang, J., *et al.* (2022). A wheat resistosome defines common principles of immune receptor channels. [2022.2003.2023.485489](https://doi.org/10.26434/chemrxiv-2022-2003-2023-485489).
- Gong, Q., Long, Z., Zhong, F.L., Teo, D.E.T., Jin, Y., Yin, Z., Boo, Z.Z., Zhang, Y., Zhang, J., Yang, R., *et al.* (2018). Structural basis of RIP2 activation and signaling. *Nature communications* *9*, 4993.
- Gros, F., Hiatt, H., Gilbert, W., Kurland, C.G., Risebrough, R.W., and Watson, J.D. (1961). Unstable ribonucleic acid revealed by pulse labelling of Escherichia coli. *Nature* *190*, 581-585.
- Gutierrez, C., and Raibaud, O. (1984). Point mutations that reduce the expression of malPQ, a positively controlled operon of Escherichia coli. *Journal of molecular biology* *177*, 69-86.
- Hatfield, D., Hofnung, M., and Schwartz, M. (1969a). Genetic analysis of the maltose A region in Escherichia coli. *Journal of bacteriology* *98*, 559-567.
- Hatfield, D., Hofnung, M., and Schwartz, M. (1969b). Nonsense mutations in the maltose A region of the genetic map of Escherichia coli. *Journal of bacteriology* *100*, 1311-1315.
- Hofnung, M., and Schwartz, M. (1971). Mutations allowing growth on maltose of Escherichia coli K 12 strains with a deleted malT gene. *Molecular & general genetics : MGG* *112*, 117-132.
- Hollingsworth, L.R., Sharif, H., Griswold, A.R., Fontana, P., Mintseris, J., Dagbay, K.B., Paulo, J.A., Gygi, S.P., Bachovchin, D.A., and Wu, H. (2021). DPP9 sequesters the C terminus of NLRP1 to repress inflammasome activation. *Nature* *592*, 778-783.



- Hu, Y., Ding, L., Spencer, D.M., and Núñez, G. (1998). WD-40 repeat region regulates Apaf-1 self-association and procaspase-9 activation. *The Journal of biological chemistry* *273*, 33489-33494.
- Hu, Z., Yan, C., Liu, P., Huang, Z., Ma, R., Zhang, C., Wang, R., Zhang, Y., Martinon, F., Miao, D., *et al.* (2013). Crystal structure of NLRC4 reveals its autoinhibition mechanism. *Science (New York, NY)* *341*, 172-175.
- Hu, Z., Zhou, Q., Zhang, C., Fan, S., Cheng, W., Zhao, Y., Shao, F., Wang, H.W., Sui, S.F., and Chai, J. (2015). Structural and biochemical basis for induced self-propagation of NLRC4. *Science (New York, NY)* *350*, 399-404.
- Huang, M., Zhang, X., Toh, G.A., Gong, Q., Wang, J., Han, Z., Wu, B., Zhong, F., and Chai, J. (2021). Structural and biochemical mechanisms of NLRP1 inhibition by DPP9. *Nature* *592*, 773-777.
- Hunter, T. (2000). Signaling--2000 and beyond. *Cell* *100*, 113-127.
- Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology* *3*, 318-356.
- Joly, N., Böhm, A., Boos, W., and Richet, E. (2004). MalK, the ATP-binding cassette component of the *Escherichia coli* maltodextrin transporter, inhibits the transcriptional activator malt by antagonizing inducer binding. *The Journal of biological chemistry* *279*, 33123-33130.
- Joly, N., Danot, O., Schlegel, A., Boos, W., and Richet, E. (2002). The Aes protein directly controls the activity of MalT, the central transcriptional activator of the *Escherichia coli* maltose regulon. *The Journal of biological chemistry* *277*, 16606-16613.
- Jones, J.D., Vance, R.E., and Dangl, J.L. (2016). Intracellular innate immune surveillance devices in plants and animals. *Science (New York, NY)* *354*.
- Kaparakis, M., Philpott, D.J., and Ferrero, R.L. (2007). Mammalian NLR proteins; discriminating foe from friend. *Immunology and cell biology* *85*, 495-502.
- Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. *Nature methods* *11*, 63-65.
- Kühnau, S., Reyes, M., Sievertsen, A., Shuman, H.A., and Boos, W. (1991). The activities of the *Escherichia coli* MalK protein in maltose transport, regulation, and inducer exclusion can be separated by mutations. *Journal of bacteriology* *173*, 2180-2186.
- Lamkanfi, M., and Dixit, V.M. (2014). Mechanisms and functions of inflammasomes. *Cell* *157*, 1013-1022.
- Larquet, E., Schreiber, V., Boisset, N., and Richet, E. (2004). Oligomeric assemblies of the *Escherichia coli* MalT transcriptional activator revealed by cryo-electron microscopy and image processing. *Journal of molecular biology* *343*, 1159-1169.
- Lee, S.J., Boos, W., Bouché, J.P., and Plumbridge, J. (2000). Signal transduction between a membrane-bound transporter, PtsG, and a soluble transcription factor, Mlc, of *Escherichia coli*. *The EMBO journal* *19*, 5353-5361.
- Lei, J., and Frank, J. (2005). Automated acquisition of cryo-electron micrographs for single particle reconstruction on an FEI Tecnai electron microscope. *Journal of structural biology* *150*, 69-80.
- Leipe, D.D., Koonin, E.V., and Aravind, L. (2004). STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. *Journal of molecular biology* *343*, 1-28.
- Li, P., Nijhawan, D., Budihardjo, I., Srinivasula, S.M., Ahmad, M., Alnemri, E.S., and Wang, X. (1997). Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* *91*, 479-489.

- Li, Y., Zhou, M., Hu, Q., Bai, X.C., Huang, W., Scheres, S.H., and Shi, Y. (2017). Mechanistic insights into caspase-9 activation by the structure of the apoptosome holoenzyme. *Proceedings of the National Academy of Sciences of the United States of America* *114*, 1542-1547.
- Lisa, M.N., Cvirkaite-Krupovic, V., Richet, E., André-Leroux, G., Alzari, P.M., Haouz, A., and Danot, O. (2019). Double autoinhibition mechanism of signal transduction ATPases with numerous domains (STAND) with a tetratricopeptide repeat sensor. *Nucleic acids research* *47*, 3795-3810.
- Liu, P., Danot, O., and Richet, E. (2013). A dual role for the inducer in signalling by MalT, a signal transduction ATPase with numerous domains (STAND). *Molecular microbiology* *90*, 1309-1323.
- Lu, G., Westbrook, J.M., Davidson, A.L., and Chen, J. (2005). ATP hydrolysis is required to reset the ATP-binding cassette dimer into the resting-state conformation. *Proceedings of the National Academy of Sciences of the United States of America* *102*, 17969-17974.
- Ma, S., Lapin, D., Liu, L., Sun, Y., Song, W., Zhang, X., Logemann, E., Yu, D., Wang, J., Jirschwitz, J., *et al.* (2020). Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. *Science (New York, NY)* *370*.
- Maekawa, S., Ohto, U., Shibata, T., Miyake, K., and Shimizu, T. (2016). Crystal structure of NOD2 and its implications in human disease. *Nature communications* *7*, 11813.
- Marquenet, E., and Richet, E. (2007). How integration of positive and negative regulatory signals by a STAND signaling protein depends on ATP hydrolysis. *Molecular cell* *28*, 187-199.
- Marquenet, E., and Richet, E. (2010). Conserved motifs involved in ATP hydrolysis by MalT, a signal transduction ATPase with numerous domains from *Escherichia coli*. *Journal of bacteriology* *192*, 5181-5191.
- Martin, R., Qi, T., Zhang, H., Liu, F., King, M., Toth, C., Nogales, E., and Staskawicz, B.J. (2020). Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. *Science (New York, NY)* *370*.
- Meunier, E., and Broz, P. (2017). Evolutionary Convergence and Divergence in NLR Function and Structure. *Trends in immunology* *38*, 744-757.
- Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of structural biology* *142*, 334-347.
- Oldham, M.L., and Chen, J. (2011). Crystal structure of the maltose transporter in a pretranslocation intermediate state. *Science (New York, NY)* *332*, 1202-1205.
- Panagiotidis, C.H., Boos, W., and Shuman, H.A. (1998). The ATP-binding cassette subunit of the maltose transporter MalK antagonizes MalT, the activator of the *Escherichia coli* mal regulon. *Molecular microbiology* *30*, 535-546.
- Peist, R., Koch, A., Bolek, P., Sewitz, S., Kolbus, T., and Boos, W. (1997). Characterization of the *aes* gene of *Escherichia coli* encoding an enzyme with esterase activity. *Journal of bacteriology* *179*, 7679-7686.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* *25*, 1605-1612.
- Plumbridge, J. (1999). Expression of the phosphotransferase system both mediates and is mediated by Mlc regulation in *Escherichia coli*. *Molecular microbiology* *33*, 260-273.
- Qi, S., Pang, Y., Hu, Q., Liu, Q., Li, H., Zhou, Y., He, T., Liang, Q., Liu, Y., Yuan, X., *et al.* (2010). Crystal structure of the *Caenorhabditis elegans* apoptosome reveals an octameric assembly of CED-4. *Cell* *141*, 446-457.

- Raibaud, O., Gutierrez, C., and Schwartz, M. (1985). Essential and nonessential sequences in malPp, a positively controlled promoter in Escherichia coli. *Journal of bacteriology* *161*, 1201-1208.
- Raibaud, O., Vidal-Ingigliardi, D., and Richet, E. (1989). A complex nucleoprotein structure involved in activation of transcription of two divergent Escherichia coli promoters. *Journal of molecular biology* *205*, 471-485.
- Rairdan, G.J., Collier, S.M., Sacco, M.A., Baldwin, T.T., Boettrich, T., and Moffett, P. (2008). The coiled-coil and nucleotide binding domains of the Potato Rx disease resistance protein function in pathogen recognition and signaling. *The Plant cell* *20*, 739-751.
- Reidl, J., and Boos, W. (1991). The malX malY operon of Escherichia coli encodes a novel enzyme II of the phosphotransferase system recognizing glucose and maltose and an enzyme abolishing the endogenous induction of the maltose system. *Journal of bacteriology* *173*, 4862-4876.
- Reyes, M., and Shuman, H.A. (1988). Overproduction of MalK protein prevents expression of the Escherichia coli mal regulon. *Journal of bacteriology* *170*, 4598-4602.
- Richet, E. (2000). Synergistic transcription activation: a dual role for CRP in the activation of an Escherichia coli promoter depending on MalT and CRP. *The EMBO journal* *19*, 5222-5232.
- Richet, E., Davidson, A.L., and Joly, N. (2012). The ABC transporter MalFGK(2) sequesters the MalT transcription factor at the membrane in the absence of cognate substrate. *Molecular microbiology* *85*, 632-647.
- Richet, E., and Raibaud, O. (1987). Purification and properties of the MalT protein, the transcription activator of the Escherichia coli maltose regulon. *The Journal of biological chemistry* *262*, 12647-12653.
- Richet, E., and Raibaud, O. (1991). Supercoiling is essential for the formation and stability of the initiation complex at the divergent malEp and malKp promoters. *Journal of molecular biology* *218*, 529-542.
- Riedl, S.J., Li, W., Chao, Y., Schwarzenbacher, R., and Shi, Y. (2005). Structure of the apoptotic protease-activating factor 1 bound to ADP. *Nature* *434*, 926-933.
- Rodriguez, A., Oliver, H., Zou, H., Chen, P., Wang, X., and Abrams, J.M. (1999). Dark is a Drosophila homologue of Apaf-1/CED-4 and functions in an evolutionarily conserved death pathway. *Nature cell biology* *1*, 272-279.
- Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of molecular biology* *333*, 721-745.
- Saur, I.M.L., Panstruga, R., and Schulze-Lefert, P. (2021). NOD-like receptor-mediated plant immunity: from structure to cell death. *Nature reviews Immunology* *21*, 305-318.
- Scheres, S.H. (2012a). A Bayesian view on cryo-EM structure determination. *Journal of molecular biology* *415*, 406-418.
- Scheres, S.H. (2012b). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology* *180*, 519-530.
- Scheres, S.H. (2016). Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods in enzymology* *579*, 125-157.
- Schiefner, A., Gerber, K., Brosig, A., and Boos, W. (2014). Structural and mutational analyses of Aes, an inhibitor of MalT in Escherichia coli. *Proteins* *82*, 268-277.
- Schlegel, A., Danot, O., Richet, E., Ferenci, T., and Boos, W. (2002). The N terminus of the Escherichia coli transcription activator MalT is the domain of interaction with MalY. *Journal of bacteriology* *184*, 3069-3077.

- Schreiber, V., and Richet, E. (1999). Self-association of the Escherichia coli transcription activator MalT in the presence of maltotriose and ATP. *The Journal of biological chemistry* *274*, 33220-33226.
- Schreiber, V., Steegborn, C., Clausen, T., Boos, W., and Richet, E. (2000). A new mechanism for the control of a prokaryotic transcriptional regulator: antagonistic binding of positive and negative effectors. *Molecular microbiology* *35*, 765-776.
- Schroder, K., and Tschopp, J. (2010). The inflammasomes. *Cell* *140*, 821-832.
- Sharif, H., Wang, L., Wang, W.L., Magupalli, V.G., Andreeva, L., Qiao, Q., Hauenstein, A.V., Wu, Z., Núñez, G., Mao, Y., *et al.* (2019). Structural mechanism for NEK7-licensed activation of NLRP3 inflammasome. *Nature* *570*, 338-343.
- Shen, C., Lu, A., Xie, W.J., Ruan, J., Negro, R., Egelman, E.H., Fu, T.M., and Wu, H. (2019). Molecular mechanism for NLRP6 inflammasome assembly and activation. *Proceedings of the National Academy of Sciences of the United States of America* *116*, 2052-2057.
- Srinivasula, S.M., Ahmad, M., Fernandes-Alnemri, T., and Alnemri, E.S. (1998). Autoactivation of procaspase-9 by Apaf-1-mediated oligomerization. *Molecular cell* *1*, 949-957.
- Steedborn, C., Danot, O., Huber, R., and Clausen, T. (2001). Crystal structure of transcription factor MalT domain III: a novel helix repeat fold implicated in regulated oligomerization. *Structure (London, England : 1993)* *9*, 1051-1060.
- Thomason, L.C., Costantino, N., and Court, D.L. (2007). E. coli genome manipulation by P1 transduction. *Current protocols in molecular biology Chapter 1*, 1.17.11-11.17.18.
- Vidal-Ingigliardi, D., Richet, E., and Raibaud, O. (1991). Two MalT binding sites in direct repeat. A structural motif involved in the activation of all the promoters of the maltose regulons in Escherichia coli and Klebsiella pneumoniae. *Journal of molecular biology* *218*, 323-334.
- Wang, J., Hu, M., Wang, J., Qi, J., Han, Z., Wang, G., Qi, Y., Wang, H.W., Zhou, J.M., and Chai, J. (2019a). Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science (New York, NY)* *364*.
- Wang, J., Wang, J., Hu, M., Wu, S., Qi, J., Wang, G., Han, Z., Qi, Y., Gao, N., Wang, H.W., *et al.* (2019b). Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science (New York, NY)* *364*.
- Wróblewski, T., Spiridon, L., Martin, E.C., Petrescu, A.J., Cavanaugh, K., Truco, M.J., Xu, H., Gozdowski, D., Pawłowski, K., Michelmore, R.W., *et al.* (2018). Genome-wide functional analyses of plant coiled-coil NLR-type pathogen receptors reveal essential roles of their N-terminal domain in oligomerization, networking, and immunity. *PLoS biology* *16*, e2005821.
- Xu, H., Shi, J., Gao, H., Liu, Y., Yang, Z., Shao, F., and Dong, N. (2019). The N-end rule ubiquitin ligase UBR2 mediates NLRP1B inflammasome activation by anthrax lethal toxin. *The EMBO journal* *38*, e101996.
- Yan, N., Chai, J., Lee, E.S., Gu, L., Liu, Q., He, J., Wu, J.W., Kokel, D., Li, H., Hao, Q., *et al.* (2005). Structure of the CED-4-CED-9 complex provides insights into programmed cell death in Caenorhabditis elegans. *Nature* *437*, 831-837.
- Yang, X., Yang, F., Wang, W., Lin, G., Hu, Z., Han, Z., Qi, Y., Zhang, L., Wang, J., Sui, S.F., *et al.* (2018). Structural basis for specific flagellin recognition by the NLR protein NAIP5. *Cell research* *28*, 35-47.
- Yin, Q., Fu, T.M., Li, J., and Wu, H. (2015). Structural biology of innate immunity. *Annual review of immunology* *33*, 393-416.
- Yu, D., Song, W., Tan, E.Y.J., Liu, L., Cao, Y., Jirschtzka, J., Li, E., Logemann, E., Xu, C., Huang, S., *et al.* (2022). TIR domains of plant immune receptors are 2',3'-cAMP/cGMP synthetases mediating cell death. *Cell* *185*, 2370-2386.e2318.

Yuan, J., and Horvitz, H.R. (1992). The *Caenorhabditis elegans* cell death gene *ced-4* encodes a novel protein and is expressed during the period of extensive programmed cell death. *Development (Cambridge, England)* *116*, 309-320.

Zdych, E., Peist, R., Reidl, J., and Boos, W. (1995). MalY of *Escherichia coli* is an enzyme with the activity of a beta C-S lyase (cystathionase). *Journal of bacteriology* *177*, 5035-5039.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods* *14*, 331-332.

Zhong, F.L., Mamaï, O., Sborgi, L., Boussofara, L., Hopkins, R., Robinson, K., Szeverényi, I., Takeichi, T., Balaji, R., Lau, A., *et al.* (2016). Germline NLRP1 Mutations Cause Skin Inflammatory and Cancer Susceptibility Syndromes via Inflammasome Activation. *Cell* *167*, 187-202.e117.

Zhou, M., Li, Y., Hu, Q., Bai, X.C., Huang, W., Yan, C., Scheres, S.H., and Shi, Y. (2015). Atomic structure of the apoptosome: mechanism of cytochrome c- and dATP-mediated activation of Apaf-1. *Genes & development* *29*, 2349-2361.

# List of abbreviations

2',3'-cAMP	2',3'-cyclic adenosine monophosphate
2',3'-cGMP	2',3'-cyclic guanosine monophosphate
ADP	adenosine diphosphate
Apaf-1	APOPTOTIC PROTEASE ACTIVATING FACTOR 1
ASC	APOPTOSIS-ASSOCIATED SPECK-LIKE PROTEIN CONTAINING A CARD
AtNUDT7	<i>ARABIDOPSIS</i> NUDIX HYDROLASE 7
ATP	adenosine triphosphate
ATR1	<i>ARABIDOPSIS THALIANA</i> RECOGNIZED 1
bp	base pair
CAP	CATABOLITE ACTIVATOR PROTEIN
CARD	caspase recruitment domain
CC	coiled-coil
CED-3	CELL DEATH PROTEIN 3
CED-4	CELL DEATH PROTEIN 4
CED-9	CELL DEATH ABNORMALITY GENE 9
C-JID	C-TERMINAL JELLY ROLL AND IG-LIKE DOMAIN
Cryo-EM	cryo-electron microscopy
C-terminal	carboxy-terminal
CTF	contrast transfer function
DAMP	danger-associated molecular pattern
dATP	deoxyadenosine triphosphate
DNA	deoxyribonucleic acid
DPP9	DIPEPTIDYL PEPTIDASES 9
dTGN	trans-Golgi network dispersion
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
EGL-1	EGG-LAYING DEFECTIVE
EIIC	ENZYME IIC COMPONENT
FIIND	function-to-find domain

GSDMD	GASDERMIN D
GSH	glutathione
GST	glutathione S-transferases
HD1	helical domain 1
HD2	helical domain 2
IPTG	isopropyl $\beta$ -d-1-thiogalactopyranoside
LRR	leucine-rich repeat
MBP	MALTOSE BINDING PROTEIN
Mlc	MAKES LARGE COLONIES
mRNA	messenger ribonucleic acid
MTOC	microtubule-organizing center
NAIP	NLR FAMILY APOPTOSIS INHIBITORY PROTEIN
NBD	nucleotide-binding domain
NEK7	NIMA-RELATED KINASE 7
NLR	nod-like receptor
NLRC4	NLR FAMILY CARD DOMAIN-CONTAINING PROTEIN 4
NLRP1	NLR FAMILY PYRIN DOMAIN-CONTAINING 1
NLRP3	NLR FAMILY PYRIN DOMAIN CONTAINING 3
NOD	nucleotide-binding and oligomerization domain
N-terminal	amino-terminal
PAMP	pathogen-associated molecular pattern
PBL2	PBS1-LIKE PROTEIN 2
PLP	pyridoxal phosphate
PRR	pattern recognition receptor
PTS	phosphotransferase system
PYD	pyrin domain
RKS1	RESISTANCE RELATED KINASE 1
RLCK	RECEPTOR-LIKE CYTOPLASMIC KINASES
RNA	ribonucleic acid
RNAP	RNA polymerase
ROQ1	RECOGNITION OF XOPQ 1
RPP1	RECOGNITION OF PERONOSPORA PARASITICA 1

SDS-PAGE	SDS-polyacrylamide gel electrophoresis
STAND	signal transduction ATPases with numerous domains
SUPR	superhelical repeats
T3SS	type III secretion system
TIR	toll/interleukin-1 receptor
TPR	tetratricopeptide repeat
UPA	CONSERVED IN UNC5, PIDD AND ANKIRINS
WHD	winged-helix domain
WT	wild type
XopQ	<i>XANTHOMONAS</i> OUTER PROTEIN Q
ZAR1	HOPZ-ACTIVATED RESISTANCE 1
ZU5	FOUND IN ZO-1 AND UNC5
$\beta$ C-S lyase	$\beta$ -cystathionase



# Acknowledgement

I would like to thank Professor Dr. **Jijie Chai**, Professor Dr. **Jane Parker**, Professor Dr. **Alga Zuccaro**, and Dr. **Jan Jirschitzka** for being part of my thesis defense committee.

I am grateful to many people who supported me at any stage of my project. First and foremost, Professor **Jijie Chai** for introducing me to the field of structural biology and for giving me guidance as well as freedom to develop this project. His dedication and enthusiasm for science will continue to inspire me in different aspects of life. Dr. **Evelyne Richet** (Institut Pasteur, Paris, France) sent me the first *malT* construct and *E. coli* expression strain, which provided the starting point of this project. Evelyne also generously shared her strains and protocols for *in vivo* assay, and gave me lots of insightful advice throughout the whole time. **Yue Sun** (Tsinghua University, Beijing, China) and Dr. **Xiaoxiao Zhang** (ShanghaiTech University, Shanghai, China) offered me tremendous help during cryo-EM analysis. The days we stayed up late in the lab for data collection is a memorable memory. Dr. **Wen Song** trained me on all the essential techniques in a structural biology lab and helped me to make a smooth transition into areas I was previously unfamiliar with. I enjoyed every football match we went to together. I owe special thanks to **Dongli Yu**, Dr. **Ertong Li**, and **Keiichi Hasegawa** for all the stimulating discussions and sleepless nights. Last but not least, I would like to thank all the people, colleagues and friends whom I met in the past years for inspiration and coffee breaks.

I am also grateful to the Alexander von Humboldt Foundation and the Max Planck Society for funding.

感谢家人在我求学过程中提供的陪伴和帮助。感谢乔思雯在读博期间，尤其是最后阶段对我的关心和支持。这都将成为我未来生活里的宝贵财富。

# Erklärung zur Dissertation

gemäß der Promotionsordnung vom 12. März 2020

***Diese Erklärung muss in der Dissertation enthalten sein.***

***(This version must be included in the doctoral thesis)***

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.“

Teilpublikationen:

Teilpublikationen liegen nicht vor.

Datum, Name und Unterschrift

Juli 2022, Yuang Wu

# Curriculum vitae

## Personal data

Name	Yuang Wu
Place of birth	Huzhou, China
Date of birth	04.09.1992

## Education

### Doctoral studies

since 10/2017	<p><b>Max Planck Institute for Plant Breeding Research, independent research group (Structural Biology, Prof. Jijie Chai), Köln, Germany</b></p> <p>Thesis project: Structural basis for the inhibition of a bacterial NLR and insights into its activation</p>
---------------	---

### Master's studies

2015-2017	<p><b>Wageningen University &amp; Research, Wageningen, the Netherlands</b></p> <p>Master of Sciences, focus on plant genetics</p> <p>Master thesis: Slow dance in the dark: protein interactions of bZIP23 during Zink deficiency in <i>Arabidopsis thaliana</i></p> <p><b>Max Planck Institute for Plant Breeding Research, Department of Plant Developmental Biology (Prof. George Coupland), Köln, Germany</b></p> <p>Internship project: Mapping QTL for flowering time that segregate in Alpine populations of perennial <i>Arabis alpina</i></p>
-----------	---

### Undergraduate studies

2011-2015	<p><b>Huazhong Agricultural University, Wuhan, China</b></p> <p>Bachelor of Sciences, focus on plant breeding and genetics</p> <p>Bachelor thesis: Genetic analysis of rice grain chalkiness</p>
-----------	--