A MASS SPECTROMETRIC STUDY OF THE CONFORMATIONAL CHANGE OF *ESCHERICHIA COLI* LIPID II FLIPPASE MURJ

ZHENG YUANYUAN

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University

Department of Applied Biology and Chemical Technology

A mass spectrometric study of the conformational change of *Escherichia coli* lipid II flippase MurJ

ZHENG Yuanyuan

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

August 2024

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

_____

ZHENG Yuanyuan

Abstract

MurJ is a membrane protein that is responsible for the transport of peptidoglycan precursor lipid II to the outside of the membrane during bacterial cell wall synthesis in *E. coli*. Previous studies have characterized the structure of MurJ in different conformations, identified the potential binding site of lipid II, and hypothesized that MurJ transports its substrate lipid II by alternating its structure between the inward and outward conformations. However, the molecular mechanism of lipid II translocation is largely unknown. For instance, the driving force involved in the conformational alternation and the dynamic changes between these conformations are still unclear. In this study, three recombinant MurJ proteins from *E. coli* were expressed and purified. The protein structure was analyzed by circular dichroism (CD) spectroscopy. The results indicated that the helical secondary structure of the MurJ protein was maintained despite variations in pH over the range of 1.0-10.0 and exposure to high concentrations (6 M) of the denaturing agent Guanidine-HCl. This observation highlights the stability of protein secondary structures in response to the environment.

To elucidate the driving force in structural transition of MurJ, a combination of structure-guided cysteine cross-linking and protein native mass spectrometry was employed to detect alternations in protein conformation *in vivo*. Two double mutant protein models were constructed, which could fix the protein in one preferred conformation after cross-linking, namely V43C_T251C (inward-open) and E70C_A296C (outward-open). By adding compounds that interfere with the cell membrane potential, it was found that the disruption of the proton motive force by the compound 3,3',4',5-Tetrachlorosalicylanilide (TCS) significantly hindered the ability of mutant E70C_A296C to transition into the outward-facing configuration required for

successful cross-linking. However, the mutant V43C_T251C was able to achieve *in vivo* cross-linking in the inward configuration under this condition. This indicates that the proton motive force is essential for the conformational transition of MurJ.

The dynamic movement of MurJ protein in two different conformations (inward-open and outward-open) in solution was studied by using hydrogen-deuterium-exchange mass spectrometry (HDX-MS). Comparative analysis of deuterium uptake in the two conformations at different timepoints showed that amino acid residues located in the central cavity of the protein displayed little variation in deuterium uptake differences. This region corresponds to the binding site of the lipid II headgroup. On the other hand, amino acid residues located in transmembrane helices TM1, 2, 7 and 9 showed significant change (> 0.5 Da) in deuterium uptake, indicating these peptides are more flexible in nature. This suggests that MurJ may facilitate the substrate flipping process by maintaining its interaction with the lipid II headgroup while changing the binding with lipid tail. Thus, the protein structure located within the central cavity exhibits relative rigidity, while the region adjacent to the lipid tails exhibits greater flexibility.

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisor, Professor Kwok-Yin Wong, for his generous advice, encouragement, support and guidance throughout my PhD studies. He always giving me new insights into my research from different perspectives. His experience and expertise in science provided a clear direction for my research. His encouragement and support gave me great confidence during the challenging moments of the project. Additionally, I am grateful for his sincere concern during the difficult period of the COVID-19 pandemic.

I would also like to thank all the people who have been supporting me during these few years. Special thanks to Dr. Yu-Wai Chen for his guidance in my study and assistance in thesis writing, Dr. Wai-Hong Chung for his support in plasmid mutagenesis experiments and paper writing, Dr. Lok-Yan So for her assistance in membrane protein expression, Dr. Pui-Kin So for his technical support and useful suggestions in mass spectrometry experiments. Many thanks go to Dr. Sirius Pui-Kam Tse and Dr. Chi-Hang Chow for their instrumental supports. I deeply appreciate Mr. Furong Gong, Dr. Wai-Po Kong, and Dr. Patrick Tsz-Fung Wong for sharing their experience, especially in HDX-MS. Additionally, I would like to thank all my past and present group members, especially Dr. Jia Yan, Mr. Cheung-Hin Hung, Miss Yi-Hui Cao, Mr. Ka-Hin Chan and Miss Jie Yu.

The technical and administrative supports by the staff in Department of Applied Biology and Chemical Technology, the University Research Facility in Chemical and Environmental Analysis, and the University Research Facility in Life Sciences, are gratefully acknowledged. This work was supported by the State Key Laboratory of

Chemical Biology and Drug Discovery of The Hong Kong Polytechnic University, the funding support of the Research Grants Council of Hong Kong and the Innovation and Technology Commission.

Last, but not least, I want to extend my heartfelt thanks to my parents and friends, especially my mother, Mrs. Zhaoying Huang, and my friends, Miss Cuixian Li and Miss Shuyi Liang. Their encouragement, assistance, and kindness comforted me during challenging times.

Table of contents

List of abbreviations

BMH	1,6-Bismaleimidohexane
CD	Circular dichroism
CL	Cardiolipin
cryo-EM	cryo-Electron microscopy
D2O	Deuterium oxide
DDM	n-dodecyl-β-D-maltoside
DMSO	Dimethylsulfoxide
E. coli	*Escherichia coli*
ESI	Electrospray ionization
GuHCl	Guanidine-HCl
HDX	Hydrogen/Deuterium eXchange
HDX-MS	Hydrogen-deuterium-exchange mass spectrometry
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
His	Histidine
HPLC	High-performance liquid chromatography
IPTG	Isopropyl-β-D-thiogalactoside
LB	Luria-Bertani
LC-MS	Liquid chromatographic mass spectrometry
MALDI	Matrix-assisted laser desorption ionization

MATE	Multi-drug and toxin extrusion
MCS	Multiple cloning site
MD	Molecular dynamics
MOP	Multidrug/oligosaccharide-lipid/polysaccharide
MPs	Membrane proteins
MR	Molecular replacement
MVF	Mouse virulence factor family
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
nMS	Native mass spectrometry
OLF	Oligosaccharidyl-lipid flippase
PBPs	Penicillin-binding proteins
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PLGS	ProteinLynx Global Server
PMF	Proton motive force
PST	Polysaccharide transporter
SCAM	Substituted-cysteine accessibility method
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulfate

SDS-PAGE	SDS polyacrylamide gel electrophoresis
T. africanus	*Thermosipho africanus*
TCS	3,3',4',5-Tetrachlorosalicylanilide
TFE	2,2,2-Trifluoroethanol
TM	Transmembrane
T_m	Melting temperature
TP	Transpeptidase
Tris	Trisaminomethane
UDP-GlcNAc	UDP-N-acetylglucosamine
UDP-MurNAc	UDP-N-acetylmuramic
ΔpH	pH difference
$\Delta\psi$	Membrane potential

Chapter 1 Introduction

1.1 Bacterial cell wall

The peptidoglycan cell wall layer exists in nearly all bacterial species [1]. It is essential for bacteria to maintain cell shape against osmotic pressure, and to regulate the division process to ensure successful cell division [2, 3]. The cell wall described here is based on *E. coli* which serves as a model for Gram-negative bacteria. The structure of the bacterial cell wall is characterized as a 3D network structure, and its biosynthesis can be divided into three parts based on reaction location: cytoplasm, membrane, and periplasm (**Figure 1.1**) [4].

In the cytoplasm, the proteins MurA and MurB collaborate to convert UDP-*N*-acetylglucosamine (UDP-GlcNAc) into UDP-*N*-acetylmuramic (UDP-MurNAc). Proteins MurC-F sequentially add five amino acids (L-Ala, D-Glu, m-DAP, and two D-Ala) to UDP-MurNAc.

In the membrane, proteins MraY and MurG link the lipid tail and UDP-GlcNAc to UDP-MurNAc-pentapeptide, respectively, to form lipid II. Subsequently, lipid II is translocated to the periplasmic by the flipping enzyme MurJ in *E. coli*.

In the periplasm, penicillin-binding proteins (PBPs) catalyze the polymerization and cross-linking of lipid II to generate the peptidoglycan cell walls.

It is noteworthy that the synthesis of the cell wall precursor, lipid II, occurs within the cytoplasm, while the final assembly takes place outside the cell membrane. Therefore, lipid II must be translocated across the membrane prior to its polymerization into peptidoglycan. This critical role of MurJ in cell wall biosynthesis makes it a popular target for the development of novel antibiotics [5, 6].

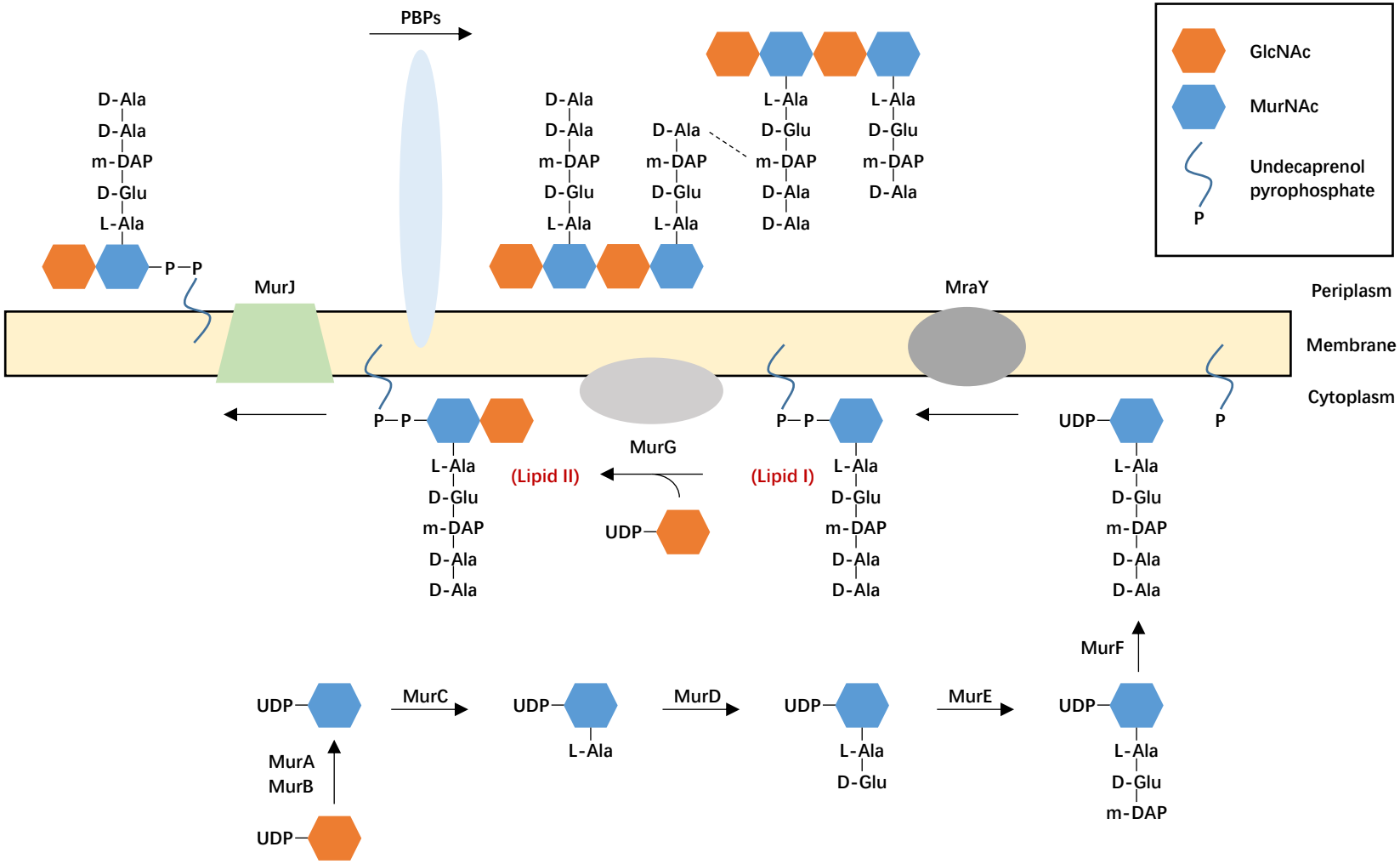


Figure 1.1 Schematic diagram of the steps of peptidoglycan biosynthesis in *E. coli* [4].

1.2 The lipid II flipping enzyme MurJ

The hypothesis that the MurJ protein is the key flipping enzyme in cell wall synthesis in *E. coli* was first proposed by Natividad Ruiz in 2008 [7]. Ruiz's study demonstrated that the depletion of MurJ results in the accumulation of lipid II within the cytoplasm, resulting in bacterial deformation and cell lysis. It is noteworthy that MurJ, formerly known as MviN, belongs to the mouse virulence factor family (MVF) family within the multidrug/oligosaccharide-lipid/polysaccharide (MOP) family [8].

1.2.1 The multidrug/oligosaccharide-lipid/polysaccharide (MOP) family

To escape the host immune responses and maintain the cell integrity and biofilms, a number of surface-exposed polysaccharides will be produced by bacteria [9]. Synthesis of these polymers, including peptidoglycan cell wall, generally initiates inside the cell and must be transported outside the membrane by transferases [10]. These metastatic activities are mostly attributed to the multidrug/oligosaccharide-lipid/polysaccharide (MOP) family of proteins. The MOP flippase superfamily is a group of integral membrane protein families [8]. It consists of four families: (a) the ubiquitous multi-drug and toxin extrusion (MATE) family; (b) the eukaryotic oligosaccharidyl-lipid flippase (OLF) family; (c) the prokaryotic polysaccharide transporter (PST) family and (d) the bacterial mouse virulence factor family (MVF). This superfamily classification was established by Rikki N. Hvorup and colleagues in 2003 [8].

By integrating these families into a superfamily framework, the functions of proteins in other families can be inferred based on proteins whose functions have been well characterized [11]. The PST family is the core of the MOP superfamily, establishing connections with the other three families [8]. The MVF family exhibits the highest sequence similarity to the PST family, suggesting that its function may parallel that of the PST family in the export of complex carbohydrates or analogous substances. Furthermore, certain members of the PST family are capable of translocating lipid-linked oligosaccharides. Comparative homology studies involving the PST and MATE families suggest that MVF proteins function as secondary carriers, facilitating the export of specific biologically significant molecules [8]. Consequently, it is hypothesized that proteins within the MVF family including MurJ are likely to function as transporters [8].

1.2.2 Structure of MurJ

To understand how MurJ transfers lipid II, Alvin CY Kuk et al. obtained MurJ crystals from *Thermosipho africanus* using the lipid cubic phase method in 2017 [12]. The crystal analysis resulted in a structure in an inward-facing conformation with substrate lipid (Protein Data Bank--PDB ID: 5T77) (**Figure 1.2**). This is the first crystal structure overview of MurJ. The structure contains a large central cavity and a hydrophobic groove. The cavity is mainly positively charged and is considered to be interacting with the negatively-charged lipid II head group. In 2018, Sanduo Zheng et al. obtained the crystal structure of MurJ from *E. coli*, at 3.5-Å resolution (PDB ID: 6CC4) (**Figure 1.2**) [13]. The crystal structure of $MurJ_{Ec}$ is similar to the $MurJ_{Ta}$. However, some transmembrane (TM) helices exhibit slight positional differences, which may be attributed to the substrate-free conditions of $MurJ_{Ec}$ crystallization or to inherent differences between the MurJ proteins of the two species. Despite these variations, the overall structural similarities suggest that basic characteristics of MurJ are conserved across species.

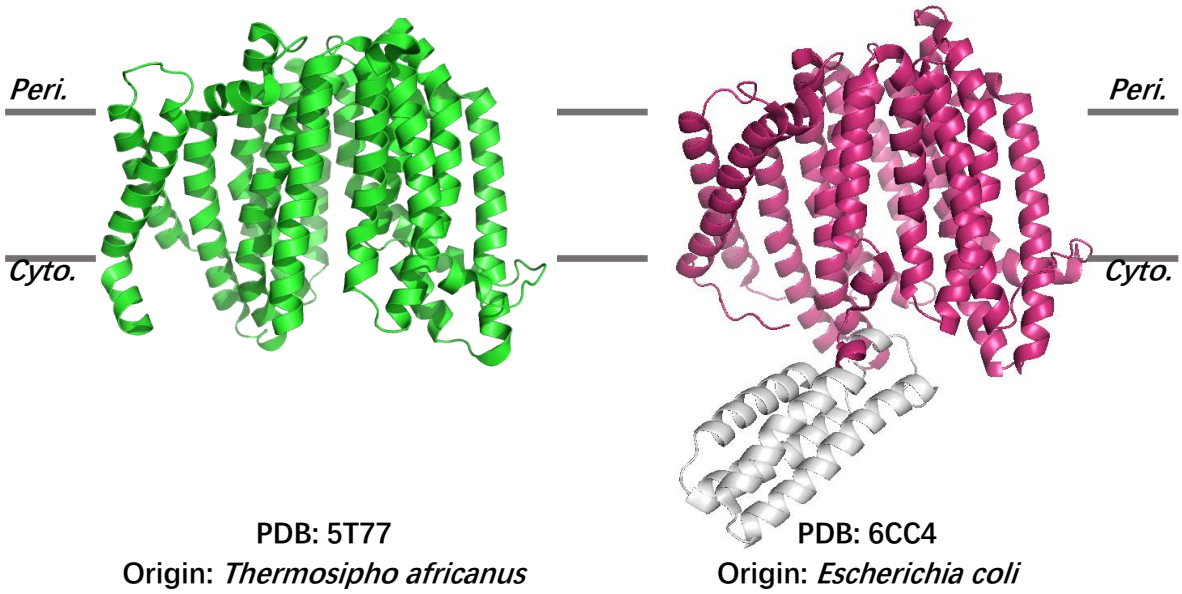


Figure 1.2 Crystal structures of MurJ. Images were downloaded from Protein Data Bank (https://www.rcsb.org/). MurJ protein from *T. africanus* and *E.coli* were shown in green and magenta, respectively. The fused apocytochrome b562RIL (BRIL) protein in $MurJ_{EC}$ was shown in grey.

1.2.3 Flipping mechanism of MurJ

1.2.3.1 Alternating-access model

In 2019, Kuk et al. successfully elucidated four new crystal structures of $MurJ_{Ta}$, which include the inward-closed (PDB ID: 6NC6), inward-open (PDB ID: 6NC7), inward-occluded (PDB ID: 6NC8), and outward-facing conformations (PDB ID: 6NC9) [14]. Based on these structures, a proposed ordered trajectory model for MurJ during its transport cycle has been developed. Initially, in the inward-open conformation, lipid II is introduced into the central cavity. Subsequently, the C-lobe bends inward, facilitating the transition of MurJ to the inward-occluded state, blocking lipid II from the cytoplasm. Following this, the N-lobe closes, leading to the transformation of the protein into an outward-facing structure, thereby opening the periplasmic gate and allowing the release of the lipid II head group into the periplasm. Finally, MurJ reverts to the inward-open conformation via a transition state of inward-closed, thereby preparing for the subsequent lipid turnover process. In 2022, Hidetaka Kohga and colleagues reported a crystal structure of the "squeezed" form of MurJ from *E. coli*, which lacks a cavity capable for substrate binding [15]. The results of molecular dynamics (MD) simulation prediction support the possibility of a conformational transition from the squeezed form to the inward form. This provides an important puzzle piece for the completion of the mechanistic model. The updated flipping model of the MurJ transport mechanism was shown in **Figure 1.3** [15].

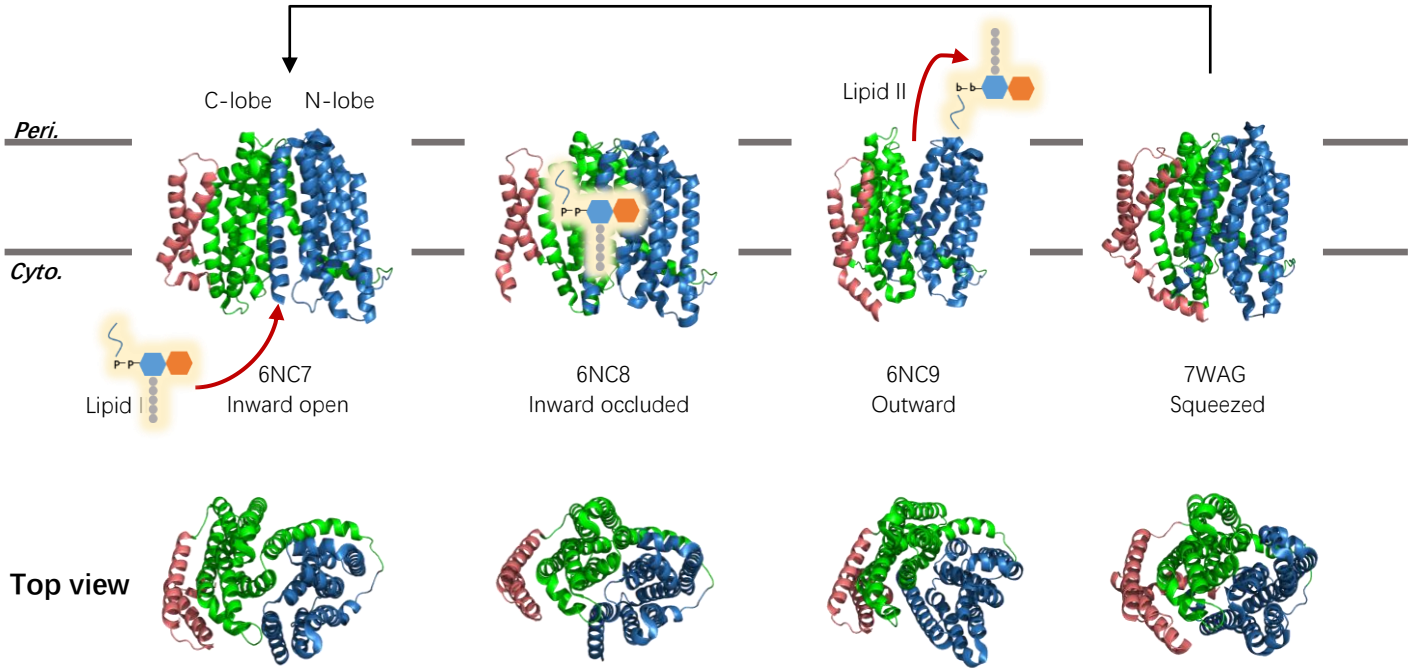


Figure 1.3 Updated flipping model of the MurJ transport mechanism [15]. The N lobe (TM1-TM6) of MurJ was labeled in blue while the C lobe (TM 7-TM12) was labeled in green. Additionally, TM13 and TM14 were labeled in red.

1.2.3.2 Driving force during transition

Currently, the predominant model for protein translocate substrate is based on an alternating access mechanism [16, 17]. However, the specific energy source responsible for facilitating the conformational transitions of proteins within this mechanism remains ambiguous. It is essential to elucidate the energy source that underpins these processes. MurJ protein belongs to the MOP protein family, which serves as a valuable reference for investigating its energy source. Notably, the MATE protein family utilizes the electrochemical gradient of transmembrane counterions (a higher ion concentration in the periplasm compared to the cytoplasm) to facilitate the efflux of drugs [18-20]. This process involves the coupling of ion influx with the uphill transport of amphiphilic drugs. The ion coupling mechanisms employed by MATE transporters are diverse, typically relying on the transmembrane electrochemical gradients of H^+ or Na^+ to mediate the excretion of drugs from the cellular environment. Kuk et al. identified a potential Na^+ binding site within the C lobe of $MurJ_{Ta}$ from crystal structures [14]. By comparing the position change of residue D235 in the TM7 hinge region in the inward-facing and outward-facing conformation crystals, it is speculated that Na^+ binding may induce conformational changes in TM7. However, the relevance of this binding site in transport has not been confirmed because its destruction will lead to protein low expression level or degradation [21]. Bolla et al. conducted an investigation into the influence of varying concentrations of Na^+ and H^+ on the binding of purified MurJ to Lipid II using native mass spectrometry [22]. The results showed that the binding affinity between MurJ and lipid II was affected by changes in H^+ concentration, and the optimal binding condition was pH 7.0, which is consistent with the pH value in the *E. coli* cytoplasm (pH 7.2–7.8). However, this study did not explore the driving force for MurJ to transport lipid II.

In addition to Na^+ and H^+, Cl^- may also be one of the counterion candidates for MurJ. In the crystal structure of $MurJ_{Ta}$ (PDB: 5T77), chloride ions were observed to bind at a proximal site near to the critical arginine residues R24 and R255 (R24 and R270 in $MurJ_{Ec}$) [12]. Given that the $MurJ_{Ta}$ (PDB: 5T77) crystals were crystallized under 150 mM concentrations of Cl^-, the researchers intentionally removed Cl^- in subsequent crystallization experiments to exclude nonspecific ion binding. This approach yielded an outward-open conformation and three distinct inward-facing conformations in the absence of Cl^- [14]. It is hypothesized that during the transport cycle of MurJ, the import of Cl^- ions into the cytoplasm drives the transition from the outward-open conformation to the inward-facing open conformation [21].

Rubino et al. proposed that MurJ requires the membrane potential for lipid II export [23]. The proton motive force (PMF) consists of two parts: membrane potential ($\Delta\psi$) and pH difference (ΔpH) [24]. 3,3',4',5-Tetrachlorosalicylanilide (TCS) has been shown to influence both the membrane potential and the pH gradient of cells. In contrast, valinomycin disrupts the membrane potential without affecting the pH gradient, while nigericin specifically impairs the ΔpH. A comparative analysis of lipid II levels in *E. coli* following treatment with these three agents revealed that only nigericin did not induce an accumulation of lipid II within the cells [23]. Consequently, it is suggested that the energy source utilized by MurJ for the transport of lipid II in *E. coli* may involve ions other than H^+. Taken together, the occasionally contradictory speculations presented above underscore the necessity for conducting further experiments to elucidate the driving force of MurJ flipping activity.

1.3 Native Mass Spectrometry

In recent years, target-based models for new drug development have gained prominence in the pharmaceutical industry [25, 26]. Research in genomics and proteomics plays a crucial role in the identification of drug targets [27, 28]. The elucidation of the function of novel genes or proteins, along with the clarification of their molecular biological mechanisms, can reveal critical components within the signaling pathways of these biological effects, thereby identifying potential drug targets [29]. The accelerated advancement of proteomics is significantly attributed to enhancements in mass spectrometry (MS) technology [30]. Following several decades of development, significant progress in MS was achieved by the 1980s with the invention of two soft ionization techniques: electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) [31]. In recognition of their contributions to these innovations, John Fenn and Koichi Tanaka were awarded the Nobel Prize in Chemistry in 2002. Their development of these ionization methods has fundamentally impacted the field of biomedical mass spectrometry, particularly in the area of proteomics. In proteomics, the mass of a protein provides information not only about its identity but also its chemical modification and structure [32]. The term "mass spectrometry" has been inaccurately applied [33]. It is not a direct measurement of mass; instead, it involves the measurement of the mass-to-charge ratio (m/z) and related properties. The mass spectrum illustrates the correlation between ion abundance and m/z value. Electrospray ionization (ESI), a soft ionization technique pioneered by J. B. Fenn in 1900, is primarily employed for the analysis of polar compounds with low volatility and poor thermal stability [34]. ESI has the advantages of high repeatability, no crystallization process, and flexibility to connect to various types of mass spectrometers.

Native mass spectrometry (nMS) is an electrospray ionization-based method in which the sample is ejected from a non-denaturing solvent and then analyzed in the gas phase of a mass spectrometer to assess intact proteins and protein complexes [35]. The term “native mass spectrometry” has been considerable ambiguity in its definition for more than a decade after it was coined in 2004 [36]. The word “native” may evoke images of the protein’s state in its natural environment (inside a cell). In 2016, Aneika C. Leney and Albert J. R. Heck unified and redefined the concept of “native mass spectrometry” [35]. The designation "native" within the context of native mass spectrometry refers to the biological state of the analyte in solution before the occurrence of the ionization process [35]. Native mass spectrometry preserves the non-covalent interactions between proteins and their ligands, the resulting mass spectra provide insight into the composition and stoichiometry of the complexes formed [37, 38]. Parameters such as pH and ionic strength are regulated to maintain the native folded state of the biological analyte in solution. To date, most proteins studied by native mass spectrometry have been shown to retain at least partially their native structure in the gas phase [35]. Native mass spectrometry plays an important role in mass spectrometric structural analysis of proteins and protein complexes [39, 40].

1.4 Hydrogen-deuterium-exchange mass spectrometry (HDX-MS)

Hydrogen-deuterium exchange is a chemical process in which covalently bonded hydrogen atoms are replaced by deuterium atoms [41]. This technique uses isotopic labeling to measure the deuterium uptake rate of amide hydrogen within the protein backbone [42]. The hydrogen atoms in peptides that can undergo hydrogen-deuterium exchange can be divided into three types according to the structure of the peptide: (1) alkyl hydrogens, which may also be referred to as α-hydrogens; (2) side chain hydrogens, which are associated with functional groups on the side chains; (3) amide hydrogens, which are distributed along the polypeptide backbone (except proline) [43]. Due to the extremely long exchange times of alkyl hydrogens and the notably short exchange times of side chain hydrogens, the observation of their exchange is not feasible [41, 42]. In contrast, the exchange times of amide hydrogens fall within the detectable range of mass spectrometry [42]. Therefore, the focus is exclusively on amide hydrogens (also known as backbone hydrogens) in HDX analysis. Despite membrane proteins comprise approximately 26% of the human proteome, their structural characterization accounts for only approximately 2% of entries in the Protein Data Bank [44, 45]. Structural studies of membrane proteins are hampered by several challenges, primarily due to protein instability during overexpression, purification, and subsequent experimental preparation [46-48].

Hydrogen-deuterium exchange mass spectrometry (HDX-MS) has been used in recent years to study membrane protein dynamics, although it has been used to characterize soluble proteins for decades [42, 49]. HDX-MS monitors the time-dependent exchange of the deuterium oxide (D_2O) with the amide hydrogen of the protein backbone under physiological conditions (**Figure 1.4**) [50]. The rate of deuterium exchange depends on protein solvent accessibility, secondary structure, and structural rigidity [42]. When coupled with proteolytic digestion and peptide separation, HDX-MS can quantitatively analyze the exchange rate in short protein fragments, providing a 'spatial resolution' of about 5 residues on average [43].

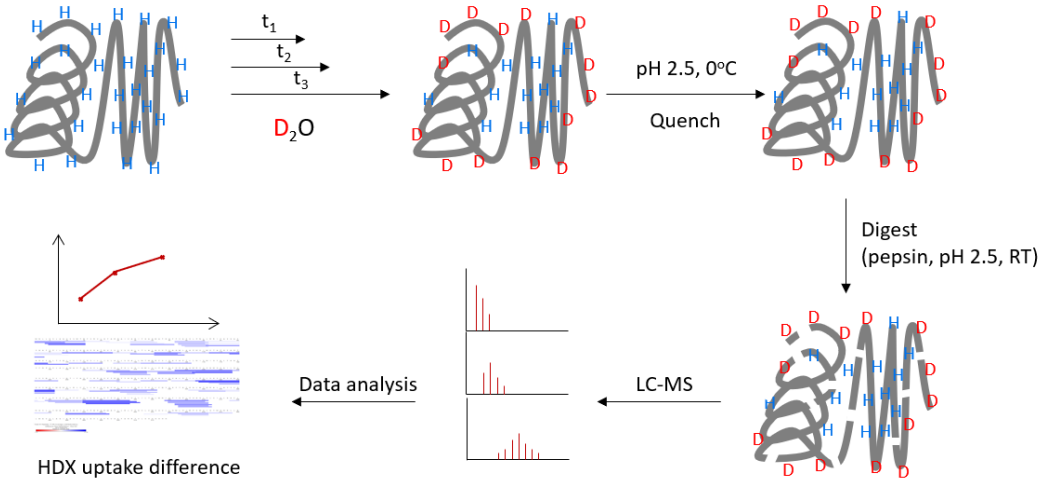


Figure 1.4 General HDX-MS workflow chart [50]. Symbols H and D represent hydrogen atoms and deuterium atoms, respectively. t1, t2, and t3 represents different labeling timepoints.

In the HDX experiment, after diluting the protein in D_2O buffer, deuterium exchanges with unstable protein hydrogen in a time-dependent manner [51]. The HDX rate is mainly influenced by 4 factors: pH, temperature, solvent accessibility, and hydrogen bonding [50]. At a pH of 7, the exchange rate of amide hydrogen that is fully solvent-exposed can reach up to 1,000 s^{-1} [52]. For folded proteins, amide hydrogens show a variety of exchange rates, depending on their position in the protein and whether they are involved in intramolecular hydrogen bonds [53, 54]. Amide hydrogen can be simply divided into "closed state" and "open state" [41]. The closed state refers to hydrogen atoms that are involved in stable hydrogen bonds in the secondary structure or are not exchanged with the solvent deuterium due to the inaccessibility of the solvent [55]. Highly dynamic and solvent-exposed regions will exchange rapidly, while rigid regions or regions involved in hydrogen bonding networks or buried inside proteins will exchange more slowly. The transition between the closed and open states occurs via a partial expansion event. Under physiological conditions where the protein is well folded, the rate of HDX is mainly determined by the rate of transition back to the closed state [56]. HDX-MS can provide insights into conformational changes upon ligand binding, as well as the pathways of protein folding and unfolding [53]. This technique is particularly advantageous for investigating the structures of proteins that are challenging to crystallize or unsuitable for nuclear magnetic resonance (NMR) analysis. The original mass spectra of intact protein or digested peptides can be converted into deuterium uptake graphs, where uptake level is plotted against the labeling time. Both the intact protein and the digested peptide deuterium uptake graph can provide information about the solvent exposure and dynamics of the protein in solution.

1.5 Aims and Objectives

The MurJ protein, which is extensively found across various bacterial species and exhibits a relatively conserved structure, is essential for bacterial survival and reproduction. Understand how MurJ works is important from both fundamental and application viewpoints.

While previous studies have explored the different conformations and substrate binding sites of MurJ, the underlying mechanisms driving its conformational transitions remain poorly understood. This study aims to:

1. Purify MurJ proteins from *E. coli* and improve the expression levels through various expression systems.
2. Investigate the role of proton motive force in the conformational change of MurJ *in vivo*.
3. Study the dynamic behavior of MurJ in solution under different conformations by HDX-MS.

Chapter 2 Expression, purification, mutagenesis, and characterization of MurJ from *Escherichia coli*

2.1 Introduction

E. coli MurJ ($MurJ_{Ec}$) was selected in this study because *E. coli* is often used as Gram-negative bacterial model in scientific research, particularly within the field of genetics [57]. Its genome is relatively small and simple, making it easier to study and manipulate [57]. By understanding gene function and regulatory mechanisms in *E. coli*, researchers can gain insight into more complex organisms.

Purification of membrane proteins like MurJ is challenging due to their poor solubility in aqueous solutions, which necessitates the use of detergents to extract and stabilize the proteins [44, 58]. A commonly used detergent n-Dodecyl β-D-maltoside (DDM) was chosen to solubilize and extract the membrane protein. MurJ protein is usually heterologously expressed in prokaryotic expression systems such as *E. coli* because these systems are easy to manipulate, cost-effective, and provide high protein yields [59]. To enhance the expression level and facilitate subsequent purification, the gene sequence of MurJ protein was cloned into a plasmid vector that encodes an additional affinity histidine (His) tag [60]. After cell lysis, cell debris and membrane components were separated by ultracentrifugation. Then, a suitable detergent (such as DDM) was used to extract and solubilize the membrane protein, to preserve the native conformation and activity of the protein as much as possible [61]. Affinity chromatography was used to enrich the target protein through the specific binding of His tag on the protein to the Ni-NTA (nickel-nitrilotriacetic acid) resin. Subsequently, the target protein was eluted by imidazole in the elution buffer.

The purity of the purified protein was verified by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis, the molecular weight was

determined and confirmed by mass spectrometry, and the secondary structure of the protein was studied by circular dichroism (CD). CD spectroscopy is a powerful technique for studying protein secondary structure, folding, and conformational changes [62-66]. Membrane proteins are essential components of cells and play crucial roles in various biological processes such as signaling and transport [67]. CD spectroscopy provides a non-destructive and relatively simple method to gain insights into the structure and dynamics of membrane proteins [68].

2.2 Methods

2.2.1 Materials

Primers and plasmids were synthesized by Thermo Fisher Scientific (HK), GenScript (Zhejiang, China). PCR purification kits and QIAprep Spin Miniprep Kit are purchased from QIAGEN. Enzymes used in plasmid reconstruction and site-direct mutagenesis were purchased from New England Biolabs (NEB). Media used in cultures and chemicals used in purification are purchased from Sigma-Aldrich. OverExpress™ chemically competent cells were purchased from Lucigen Corporation (US). The cOmplete™ protease inhibitor cocktail was purchased from Sigma-Aldrich. Isopropyl β-D-thiogalactopyranoside (IPTG) was purchased from Thermo Fisher Scientific (HK). n-Dodecyl-β-D-maltoside (DDM) was purchased from Sigma-Aldrich. Amicon centrifugal filters were purchased from Merck Millipore. Gravity (Poly-Prep®) and spin chromatography (Bio-Spin® and Micro Bio-Spin™) columns were purchased from Bio-Rad. Ni-NTA resins were purchased from Thermo Fisher Scientific (HK). Precision Plus Protein Dual Color Standards used in SDS-PAGE was purchased from Bio-Rad. Organic solvents used in LC-MS were HPLC grades purchased from Merck Millipore. Formic acids, 8 M Guanidine-HCl (GuHCl) Solution, and 2,2,2-Trifluoroethanol (TFE) were purchased from Sigma-Aldrich. Cardiolipin assay kits (MAK362) were purchased from Sigma-Aldrich.

2.2.2 Plasmid constructions

The plasmids N-His-MurJ and C-His-MurJ used in this study were constructed as follows. The gene fragment of full-length MurJ from *E. coli* was obtained by polymerase chain reaction (PCR) amplification using *E. coli* K12 genomic DNA as a template. The primers used in this study are listed in **Table 2.1**. Primers were designed using the SnapGene software (www.snapgene.com), and the T_m difference of primer pairs was controlled within 5 °C. These primers were then optimized using the IDT codon optimization tool (www.idtdna.com).

The PCR was performed in a Swift MaxPro thermal cycler (ESCO) using iProof High-Fidelity DNA Polymerase (Bio-Rad). The PCR reaction mixture was prepared as shown in **Table 2.2**. The reaction includes three steps: denaturation, annealing, and extension. The PCR protocol starts with denature at 95 °C for 3 minutes, followed by 16 cycles of 95 °C for 20 seconds, 58 °C for 30 seconds, and 72 °C for 6 minutes. The final stage at 72 °C lasted 6 minutes to complete the reaction. The PCR product was purified using MinElute PCR Purification Kit (Qiagen) and digested with the corresponding restriction enzyme (1 µL, 50,000 units/ml) at 37 °C for 1 hour. The plasmid pET-28a (+) was digested and purified similarly. T4 DNA ligase (1 µL, 2,000,000 units/mL) was added to the purified plasmid and PCR product, and the mixture was incubated at 4 °C for 16 hours to allow DNA fusion. The fused recombinant plasmid was transformed into Top10 competent cells for plasmid propagation. The plasmid was then extracted using the Invitrogen Plasmid Miniprep Kit and sent for sequencing (BGI Hong Kong Co. Limited). The plasmid with the correct sequence was transformed into the *E. coli* expression system for protein expression.

The recombinant plasmids constructed in this study were listed in **Table 2.3**. For the plasmid N-His-MurJ, the gene fragment encoding the MurJ protein from *E. coli* was amplified by PCR using *E. coli* K12 genomic DNA as a template. The primers EcMurJ-F1 and EcMurJ-R1 contained BamHI and HindIII restriction sites, respectively. The vector and PCR products were digested with enzymes BamHI and HindIII, and the PCR product was inserted between the BamHI and HindIII sites of pET-28a (+) to obtain the reconstructed N-His-MurJ plasmid. For the plasmid C-His-MurJ, the MurJ DNA was inserted between the NcoI and XhoI sites of pET-28a (+). Plasmid N-His-TEV-MurJ was generated through PCR and ligation reaction using N-His-MurJ as a template.

Table. 2.1 Primers used in plasmid reconstruction.

Plasmid	**Primers**	**Sequence (5'-3')**
N-His-MurJ	EcMurJ-F1	ATATGGATCCATGAATTTATTAA AATCGCTGGCCGCCGTC
	EcMurJ-R1	AATTAAGCTTTTACACCGTCCGG CGGGCAAATTCTTTAAC
N-His-TEV-MurJ	His-TEV-EcMurJ-F1	ATGAATTTATTAAAATCGCTGGC CGCCGTC
	His-TEV-EcMurJ-R1	GCTGCCCTGAAAATACAGGTTTT CGTGATGATGATGATGATGGCTG CT
C-His-MurJ	C_EcMurJ-F2	ATATCCATGGATGAATTTATTAA AATCGCTGGCCGCCGTC
	C_EcMurJ-R2	AATTCTCGAGCACCGTCCGGCGG GCAAATTCTTTAAC
	C_EcMurJ-F1	AGAAGGAGATATACCATGAATT TATTAAAA
	C_EcMurJ-R1	TTTTAATAAATTCATGGTATATC TCCTTCT

Table. 2.2 Setup for PCR reaction.

Components	**Volume (uL)**
H_2O	35.00
5 × buffer	10.00
dNTP (10 mM)	1.00
Forward and reverse primer mixture (10 uM)	2.50
Template	1.00
DNA polymerase (2 U/μL)	0.50
Total volume	50.00

Table. 2.3 Summary of plasmids used in this study.

Protein Origin	**Vector**	**Reconstructed Plasmid**
Escherichia coli	pET-28a (+)	N-His-MurJ
		N-His-TEV-MurJ
		C-His-MurJ

2.2.3 Plasmid transformation

The competent *E. coli* cells were thawed on ice for 10 minutes. Then, 1-5 µL of plasmid DNA (1 pg-100 ng) or water (negative control) was added to the cell mixture and mixed by gently flicking the tube. After incubating for 30 minutes on ice, the mixture was heat shocked at 42 °C for 30 to 60 seconds and then incubated on ice for 10 minutes. 1 mL of room temperature LB medium was added and incubated at 37 °C, 250 rpm for 1 to 3 hours. 100 µL of cells were spread onto Luria-Bertani (LB) plates containing kanamycin (50 µg/mL) or ampicillin (100 µg/mL) and incubated overnight at 37 °C.

2.2.4 Cysteine mutation

The plasmid of mutant E70C_A296C was purchased from Genescript, while the plasmid for mutant V43C_T251C was generated using site-directed mutagenesis (SDM). The cloning vector was pET-28a (+) and the restriction sites chosen were NcoI and XhoI. The complete mutant plasmid constructs were shown in **Figures S1** and **S2** in the **Appendix**. Wild-type MurJ contains two intrinsic cystines which needs to be first mutated to alanine. Then, a plasmid without Cys codon was used as a template to introduce new double Cys substitutions. The templates and primers used in the experiments are listed in **Table 2.4**. Primers were designed using the SnapGene software (www.snapgene.com), and the T_m difference of primer pairs was controlled within 5 °C. These primers were then optimized using the IDT codon optimization tool (www.idtdna.com).

PCR was conducted using ESCO® Swift MaxPro following the outlined procedure. The template plasmid was initially denatured at 95 °C for 3 minutes, followed by amplification of the target sequence in 16 cycles. Each cycle involved denaturation of plasmids at 95 °C for 40 seconds, annealing at 58 °C for 30 seconds, and extension at 72 °C for 6 minutes. Following the completion of 16 amplification cycles, an additional extension at 72 °C for 6 minutes was performed. Verification of the mutant plasmid size was carried out through 0.9% agarose gel electrophoresis.

Upon confirmation of the product, DpnI (1 µL, 20,000 units/mL) was added to 50 µL PCR product and incubated for 2 hours at 37 °C to cleave the template plasmid. Prior to ligation, the PCR product was purified using the QIAquick PCR purification kit. T4 DNA ligase (1 µL, 2,000,000 units/mL) was added to 20 µL of sample mixture and incubated at room temperature for 2 hours for ligation.

The mutated plasmid was subsequently introduced into competent Top 10 cells for plasmid amplification. Mutant plasmids were isolated using the QIAprep miniprep kit and send to BGI for sequencing. Sequence alignment was conducted using the SnapGene software. The validated plasmid was then introduced into the *E. coli* C43 (DE3) expression system, and transformants were selected on LB agar plates supplemented with 50 µg/mL kanamycin.

Table 2.4 Primers used in SDM.

Primers	**Sequence (5'-3')**
EcMurJ_C314S_F1	GGGTTGCGTCTT**TCA**TTCCTGTTGGCG
EcMurJ_C314S_R1	CGCCAACAGGAA**TGA**AAGACGCAACCC
EcMurJ_C419S_F2	TCTATTGGTCTGGCGGCG**TCA**CTGAATGCTTCG
EcMurJ_C419S_R2	CGAAGCATTCAG**TGA**CGCCGCCAGACCAATAGA
EcMurJ_V43C_F1	GACGCCTTTTTC**TGT**GCTTTTAAACTTC
EcMurJ_V43C_R1	GAAGTTTAAAAGC**ACA**GAAAAAGGCGTC
EcMurJ_T251C_F1	CTTAATCATCAAC**TGT**ATTTTTGCCTCGT
EcMurJ_T251C_R1	ACGAGGCAAAAAT**ACA**GTTGATGATTAAG

2.2.5 Protein expression and purification

Full-length MurJ was overexpressed and purified using a modified method of Alvin C Y Kuk *et al* [12]. A single colony of the successfully transformed recombinant plasmid was selected, inoculated into LB medium containing the corresponding antibiotics, and incubated at 37 °C with shaking overnight. The overnight culture was mixed with 1 L 2 × YT medium at a ratio of 1:100 and then cultured with shaking at 37 °C. When the bacteria were cultured until the optical density (OD) at 600 nm reached 0.6 to 0.8, a final concentration of 1 mM IPTG was added and then incubated at 37 °C for 3 hours to induce the overexpression of the target protein.

The cell pellet was collected by centrifugation at 4000 × g for 15 minutes and then resuspended in 20 mL of resuspension buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl) containing 1/2 piece of protease inhibitor cocktail tablet. The cells were disrupted three times using a cell disruptor under a French press at 20 × kpi, and the resulting cell fragments were harvested by centrifugation at 10,000 × g for 10 minutes. The supernatant was ultracentrifuged (HITACHI CP70ME) at 200,000 × g for 1 hour to collect the membrane fraction. Then, 20 mL of solubilization buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 30 mM DDM) was added, and the mixture was stirred at 4 °C overnight to extract MurJ from the membrane.

After protein extraction, the mixture was centrifuged to collect the supernatant. The supernatant was filtered with a 0.45 μm filter, and 1 mL of Ni-NTA agarose was added and incubated at 4 °C for 1 hour. The resin mixture was loaded onto a gravity column, and 20 mL of buffer A (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 15 mM imidazole, 1

mM DDM) was added to elute unbound or low-bound impurities. Protein MurJ was eluted with 5 mL of buffer B (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 200 mM imidazole, 1 mM DDM). The eluted fractions were analyzed by 12% SDS-PAGE and LC-MS.

2.2.6 Protein identification and characterization

2.2.6.1 Protein identification by SDS-PAGE and LC-MS

The purified protein elution fraction was characterized by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and liquid chromatography-mass spectrometry (LC-MS). The SDS-PAGE gel contained 5% stacking gel and 12% separation gel. Before sample loading, an equivalent volume of loading buffer was added to the samples, the mixtures were incubated at room temperature for 10 minutes. The protein bands were separated at a voltage of 160 V for 1 hour.

LC-MS experiments were performed using a SYNAPT G2-Si mass spectrometer (Waters). The ACQUITY UPLC BEH300 C4 analytical column (1.7 µm, 2.1 × 50 mm) and VanGuard Pre-column (1.7 µm, 2.1 mm × 5 mm) were used at a flow rate of 60 µL/min. The mobile phase A and B were H_2O and ACN. Both mobile phase A and B contained 1% formic acid. The LC elution gradient method is shown in **Table 2.5**.

Table 2.5 LC mobile phase gradient.

Time (min)	Mobile phase A (%)	Moblie phase B (%)
0.50	92.00	8.00
10.00	40.00	60.00
19.50	5.00	95.00
20.50	5.00	95.00
22.00	92.00	8.00

2.2.6.2 Circular Dichroism (CD) measurements

CD spectra were performed using a J-1500 circular dichroism spectrophotometer (JASCO). A cuvette with a pathlength of 1 mm was used and the sample volume was 200 µL. The cuvette was rinsed thoroughly with H_2O and 70% ethanol between samples. Purified MurJ was concentrated and exchanged to the buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM DDM) using an Amicon device with a 30 kDa molecular weight cutoff (Merck Millipore). The aggregated protein particles were removed by centrifugation at 10,000 × g for 10 min. Protein secondary structure CD spectra were recorded in the range of 190-280 nm, with a bandwidth of 1 nm, a step slit of 1 nm, and a time constant of 4 s. Each curve was scanned 3 times at a 50 nm/min scan rate. CD spectra of MurJ at different pH values were obtained by diluting the concentrated protein stock into buffers with varying pH values. In the protein chemical unfolding experiment, 8 M GuHCl solution was used to prepare buffer solutions containing different concentrations of GuHCl. The protein melting temperature (T_m) was determined in the temperature range of 5 to 95 °C with a gradient temperature of 2 °C/min, a bandwidth of 1 nm, a step slit of 1 nm, and a time constant of 4 s. CD melting spectra of MurJ at different pH (1, 3, 8, and 10) were collected.

2.2.7 Protein delipidation and buffer exchange

Protein samples must undergo delipidation procedures prior to conducting native mass spectrometry (nMS) experiments. During the protein purification process, membrane proteins will co-purify with numerous lipids, which can affect the protein signal in nMS, making signal peaks broadened and difficult to analyze [69]. The delipidation process is achieved by adding empty detergent micelles. After the bacteria was lysed, the membrane fragments were collected by ultracentrifugation, and then extraction buffer (20 mM HEPES, 150 mM NaCl, 30 mM DDM, pH 7.0) was added to extract the protein overnight at 4 °C. The supernatant was then collected through ultracentrifugation and incubated with Ni-NTA resin for 30 minutes at 4 °C, allowing histidine-tagged proteins to bind to the resin. The resin was collected via low-speed centrifugation, fresh detergent-containing buffer (20 mM HEPES, 150 mM NaCl, 1 mM DDM, pH 7.0) was added, and the mixture was incubated at 4°C for 1 hour, with this process repeated twice. Then the protein purification steps described in **section 2.2.5** was followed, the protein was eluted with 200 mM imidazole-containing buffer. The purified proteins were analyzed by 12% SDS-PAGE.

Protein delipidation was assessed through LC-MS analysis or by utilization of a cardiolipin kit. LC-MS experiments were performed using Agilent 6540 UHD Q-TOF and Agilent1200 Series Binary LC System instruments. The ACQUITY UPLC HSS T3 analytical column (1.8 µm, 2.1 mm × 100 mm) and VanGuard Pre-column (1.8 µm, 2.1 mm × 5 mm) were used at a flow rate of 0.3 mL/min. The mobile phase used were ACN/H_2O 6:4 (mobile phase A) and IPA/ACN 9:1 (mobile phase B). Both mobile phase A and B contained 10 mM ammonium formate and 0.1% formic acid. The LC elution gradient method is shown in **Table 2.6**. Protein samples were digested with

trypsin (sequencing grade in 50 mM acetic acid) overnight at 37 °C, and the supernatants were collected by centrifugation at 10,000 × g for 5 min. The protease:protein ratio is 1:50 (w/w). Then the supernatant was vacuum dried using CentriVap Benchtop Vacuum Concentrators (Labconco) at 4 °C and reconstituted in $CHCl_3$/MeOH 1:1 before sample loading.

The cardiolipin assay kit uses a probe that specifically binds to cardiolipin and produces fluorescence. The cardiolipin content in the sample is calculated by measuring the fluorescence intensity of the sample at 480 nm (λ_{ex} = 340 nm). The standard curve is determined using the purified cardiolipin included in the kit. The experimental protein concentration was 5 μM.

To better ionize protein samples, it is required to exchange the sample into a volatile ammonium salt buffer before the experiment [70]. The presence of detergents can also affect the efficiency of protein ionization in nMS [71, 72]. In the study by Bolla, it was observed that LDAO performed better than DDM in nMS of the MurJ protein [22]. The protein was exchanged into the buffer (200 mM ammonium acetate, 0.05% LDAO, pH 7.0) using a P-30 Bio-spin column (Bio-Rad). Then, the samples were concentrated using a 30 kDa Amicon (Merck) Ultra-0.5 mL centrifuge device. The resulting concentrate was centrifuged at 10000 × g for 10 min to remove precipitates prior to MS analysis.

Table 2.6 LC mobile phase gradient used in lipid analysis.

Time (min)	Mobile phase A (%)	Moblie phase B (%)
0.00	60	40
2.00	60	40
2.50	42.00	58.00
18.00	1.00	99.00
20.00	1.00	99.00

2.3 Results and discussion

2.3.1 Recombinant Plasmids

The recombinant plasmid fused the His-tag to the N-terminus or C-terminus of the protein to facilitate subsequent purification of the protein by affinity chromatography. Studies have found that the expression level of proteins with a C-terminal His_6 tag is higher than that of proteins with an N-terminal His_6 tag [22]. The positions of the fused N-terminal and C-terminal His_6 tags relative to the binding cavity of MurJ are shown in **Figure 2.1**. These two positions are very close to the entrance of the cavity and may affect subsequent substrate-binding experiments. Taking all the above into consideration, proteins with N-terminal fusion or C-terminal fusion tags were constructed here. Additionally, a TEV restriction site was added to N-His-MurJ to prepare for the His-tag removal.

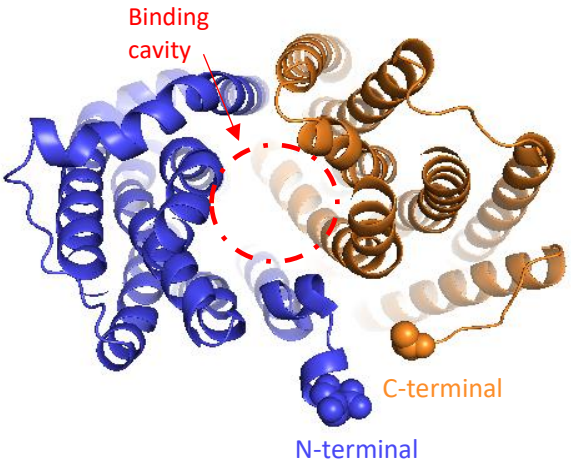


Figure 2.1. Schematic diagram of the location of the His_6-tag fusion in MurJ. The protein structure schematic used here is 6CC4. The fused apocytochrome b562RIL (BRIL) protein has been removed. The N-terminal bundle, C-terminal bundle, and central binding cavity of the protein are labeled in blue, orange, and red, respectively.

2.3.2 Protein over-expression

Bacteria are widely used as a fast and easy recombinant protein expression system with advantages such as low cost and easy reproduction. The most widely used bacterial expression system is *E. coli* because its genetics, genome sequence and physiology are relatively clear [73]. Its simple cultivation process and ability to achieve high cell densities make it suitable for large-scale fermentation.

Strains C41 (DE3), C43 (DE3), and C41 (DE3) pLysS all contain genetic mutations and have a certain tolerance to toxic proteins which can express toxic proteins more effectively [74]. Strain C41 (DE3) is derived from BL21 (DE3) [75]. This strain has a mutation that reduces the level of T7 RNA polymerase activity which can reduce many cell deaths associated with the overexpression of toxic recombinant proteins. Strain C43 (DE3) was derived from C41 (DE3) by selecting resistance to different toxic proteins. It carries at least one additional mutation and has a higher tolerance level to toxic proteins. The C41 (DE3) pLysS strain has a chloramphenicol-resistant plasmid that expresses a small amount of T7 lysozyme, which is a natural inhibitor of T7 RNA polymerase [76]. This strain can inhibit the basic expression of T7 RNA polymerase before the target protein expression, thereby stabilizing the encoding of particularly toxic recombinant proteins.

The recombinant plasmids were transformed into these expression systems for the IPTG induction test to screen out a more suitable expression system (**Table 2.7**). The induction results were checked by SDS-PAGE gel, and the gel images are shown in **Figures 2.2**. The theoretical molecular weight of these MurJ proteins is around 55-56 kDa. The MurJ protein band has a ‘gel shift’, which may be due to the protein

characteristics and the detergent used in the purification process. Gel shift is commonly exists in many membrane protein SDS-PAGE experiments, but the mechanism remains unclear [77]. The samples in **Figure 2.2** are the membrane debris obtained by centrifugation after bacterial cell lysis. The expression of MurJ protein in these cells was induced by adding IPTG (1mM) at 37 °C for 3 hours. Referring to the reported band positions of MurJ protein on SDS-PAGE, there is a obvious band (black arrow) in the range of 37~50 kDa, which is considered to correspond to the MurJ protein [6, 22]. The intensity of the bands was measured using the software ImageJ [78]. The results (**Figure 2.3**) showed that the expression levels of C43 (DE3) of the 3 recombinant plasmids mentioned above were similar and satisfactory. Therefore, the C43 (DE3) expression system was selected for subsequent protein preparation experiments.

Table 2.7. Summary of *E. coli* expression systems used in MurJ expression.

Origin	**Reconstructed Plasmid**		**Expression system**		
			C41 (DE3)	C43 (DE3)	C41 (DE3) pLysS
E. coli	pET-28a (+)	N-His-MurJ	√	√	√
		N-His-TEV-MurJ	√	√	√
		C-His-MurJ	√	√	√

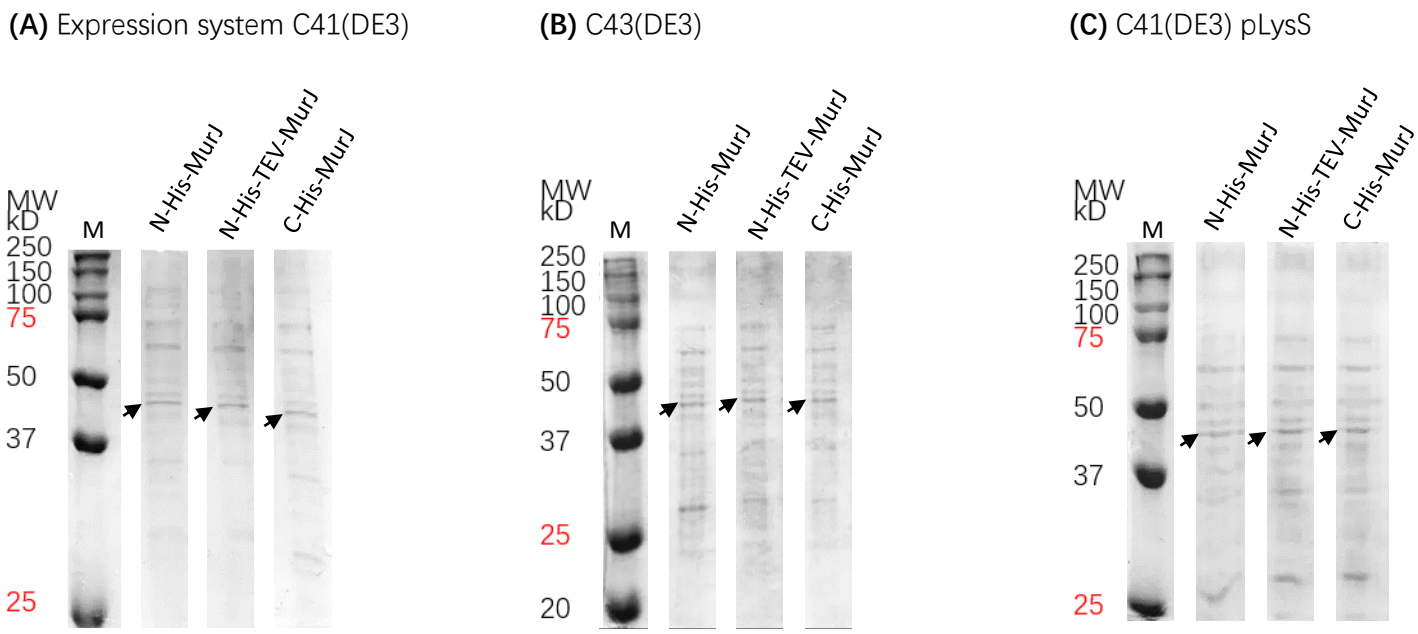


Figure 2.2. SDS-PAGE images of IPTG test of 3 recombinant plasmids in different *E. coli* expression systems. The expression level of protein MurJ induced by 1 mM IPTG in expression system (A) C41 (DE3), (B) C43 (DE3), (C) C43 (DE3) pLysS. The MurJ protein bands were marked by the black arrows.

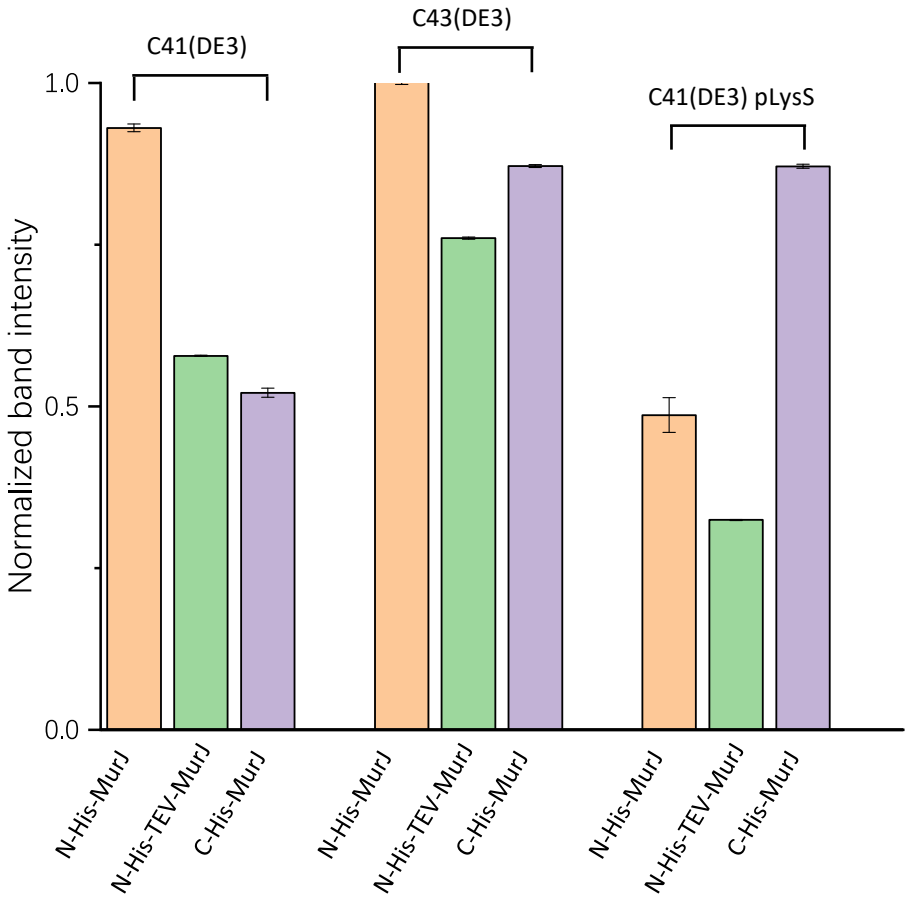


Figure 2.3 Protein band intensity of different expression systems. The orange columns correspond to plasmid N-His-MurJ, the green columns correspond to plasmid N-His-TEV-MurJ, the purple columns correspond to plasmid C-His-MurJ. The first group is C41 (DE3), the second group is C43 (DE3) and the third group is C41 (DE3) pLysS.

2.3.3 Protein purification

The proteins expressed by the system C43 (DE3) mentioned above were purified and preliminarily confirmed by SDS-PAGE (**Figure 2.4**). Compared with the standard ladder, the protein band position was about 40 kDa, which was consistent with the reported value in the literature [6, 22]. The intensity of individual bands in each lane was quantified using ImageJ software. It was calculated that the band intensity of MurJ protein in each lane accounted for 85%-95% of the total bands.

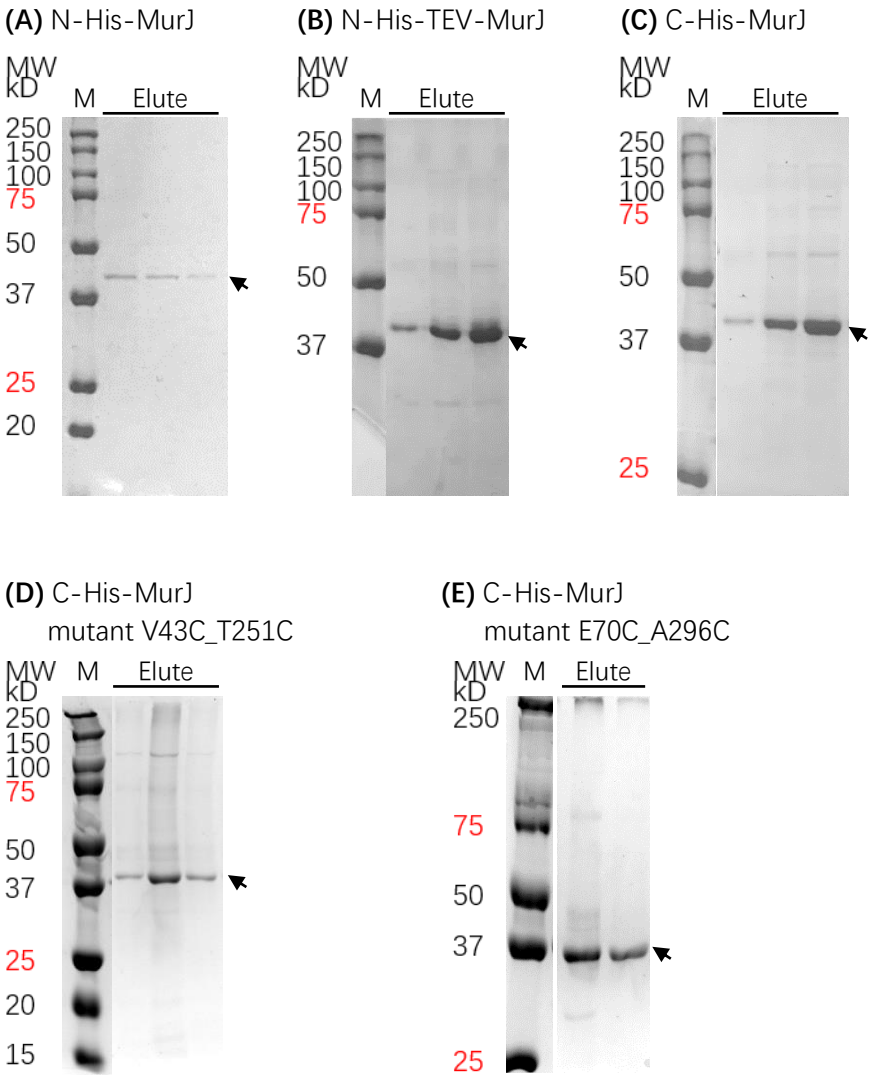


Figure 2.4 SDS-PAGE images of purified MurJ protein fractions. (A) N-His-MurJ (B) His-TEV-MurJ (C) C-His-MurJ (D) C-His-MurJ mutant V43C_T251C (E) C-His-MurJ mutant E70C_A296C.

The desired protein fractions were collected and concentrated, and then subjected to mass spectrometry analysis for further protein identification. The protein mass spectrometry data are shown in **Table 2.8** and **Figure S3-S8** in the **Appendix**. The experimental protein molecular weights were consistent with the theoretical values. Except that N-His-MurJ is 131 kDa smaller, which is -1 methionine difference, corresponds to the protein N-terminal methionine excision (NME). This is a common phenomenon in proteins post-translational modification. The role of NME is still unclear, but it is believed that up to 80% of the proteins in each proteome will undergo this modification [79]. In summary, the protein MurJ was successfully expressed and purified for further experiments.

Table 2.8. Summary of purified MurJ molecular weight data.

Origin	Plasmid	Molecular Weight (Da)	
		Theoretical	Experimental
E. coli	N-His-MurJ	58811.36	58680.37±1.3
	N-His-TEV-MurJ	57391.71	57261.34±1.6
	C-His-MurJ (wild-type)	56332.59	56332.54±1.9
	V43C_T251C	56306.51	56306.64±0.7
	E70C_A296C	56363.60	56233.47±0.8

2.3.4 Secondary structure of MurJ measured by CD spectroscopy

Circular dichroism (CD) spectroscopy is a technique used to study the chiral properties of molecules, which depend on the chirality of the molecule [80]. CD spectroscopy is used to determine the secondary structure of proteins, such as the presence of α-helices, β-sheets, and random coils. When a protein contains α-helical regions, the peptide bonds will arrange in a consistent spiral manner. α-Helices exhibit a characteristic CD spectrum with two negative bands around 222 nm and 208 nm and a positive band around 190 nm [64]. The secondary structure of MurJ protein is mainly composed of α-helices. These α-helices pass through the cell membrane to form a channel or pore that allows specific molecules to pass through. Between the transmembrane helices, MurJ protein has loop structures, which are located on both the intracellular and extracellular sides.

The proteins used in the CD experiment were C-His-MurJ. The MurJ CD spectrum (**Figure 2.5**) shows peaks at 190 nm, 208 nm, and 222 nm, consistent with typical spectral features associated with helical structures. Previous structural studies of MurJ, including site-directed mutagenesis combined with the substituted-cysteine accessibility method (SCAM) and protein crystal structure analysis, all suggested that the structure of MurJ is mainly composed of 14 helices [12-14, 81].

It was also found that the protein retained its helical conformation under acidic conditions (pH 1.0) and alkaline conditions (pH 8.0). Compared to β-barrel proteins and water-soluble proteins, α-helical membrane proteins generally retain a considerable amount of structure in their experimentally unfolded state [82, 83]. This partially structured unfolded state may be the scaffold required for the successful folding of the protein [83]. To investigate whether the helical structure maintained by MurJ under pH changes is the helical core left after unfolding, chemical denaturation and thermal unfolding studies were performed.

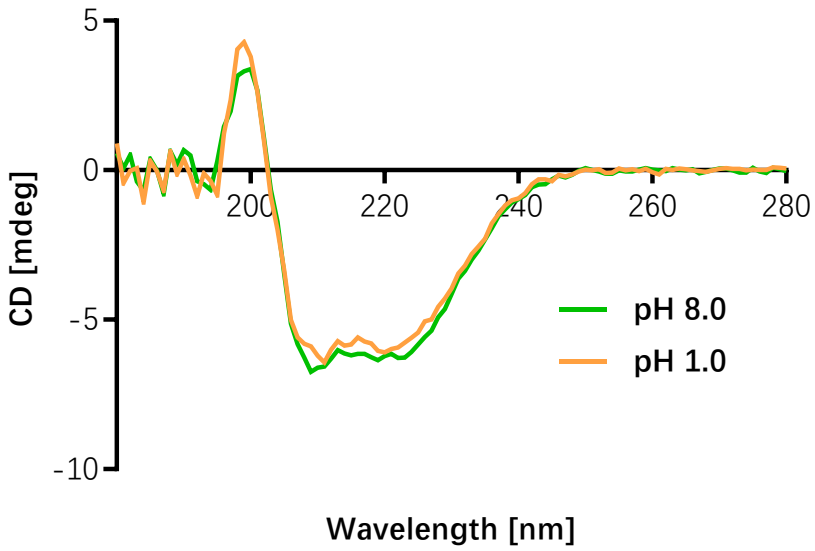


Figure 2.5 CD spectrum of MurJ under different pH. The concentration of MurJ was 0.5 μM. The experimental baseline was subtracted using buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM DDM).

2.3.4.1 MurJ chemical denaturation by Guanidine-HCl (GuHCl)

GuHCl is a commonly used protein denaturant in protein unfolding studies [63, 84]. Hydrogen bonding is one of the main stabilizing forces of protein secondary structures such as α-helices and β-sheets. GuHCl disrupts the secondary structure of proteins by breaking these hydrogen bonds. Additionally, GuHCl acts as a potent electrolyte that affects the protein folding processes by changing the ionic strength of the solution and interfering with the electrostatic interactions between protein molecules.

The CD spectra of MurJ under various concentrations of GuHCl are shown in **Figure 2.6**. The green curve is the CD spectrum of the protein without the addition of GuHCl denaturant. Two distinct minimum values were observed at 208 nm and 222 nm, indicating the existence of α-helix structure within the protein. As the denaturant concentration increases, the negative peak at 222 nm gradually decreases. It should be noted that the increase in the negative value of CD here represents an increase in helicity.

The peak (222nm) signal exhibited minimal variation under 1 M GuHCl. However, a notable decrease in this signal was observed when the concentration exceeded 3 M, suggesting a significant reduction in the helical structure within the sample. The CD values at 222 nm are plotted in **Figure 2.7**. It was calculated that in the presence of 6 M GuHCl, approximately 35% of the helical structure retains in the sample compared to its initial helicity content. Therefore, the helical structures that can withstand pH fluctuations mentioned before should not be considered as helical cores preserved after the unfolding process.

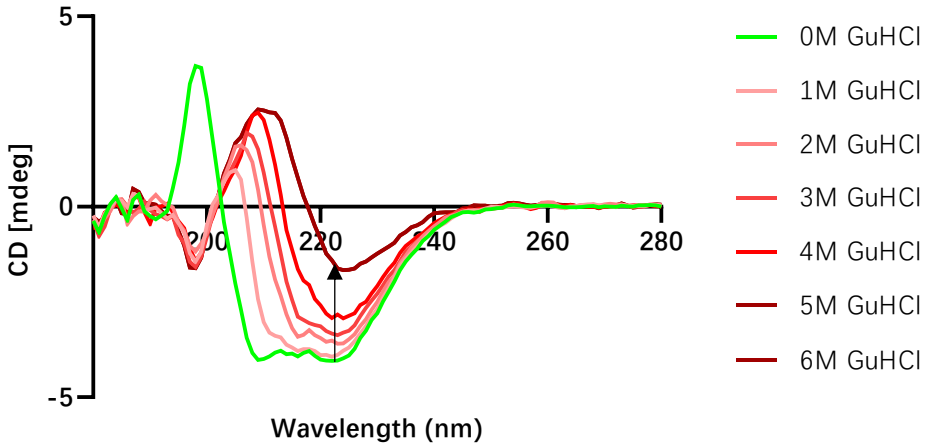


Figure 2.6 CD spectrum of MurJ under different concentrations of denaturing reagent GuHCl. The concentration of MurJ was 0.5 µM. The experimental baseline was subtracted using buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM DDM). The green line represents the protein sample without the addition of GuHCl. The red lines represent protein samples with the addition of 1-6 M GuHCl.

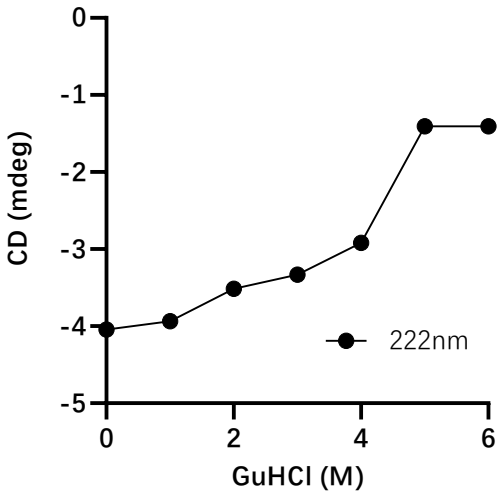


Figure 2.7 The plot of CD value of MurJ at 222 nm against GuHCl concentration during the unfolding processes.

2.3.4.2 Thermal unfolding of MurJ

Thermal unfolding studies of membrane proteins are significant in biology and drug development [63]. Studying the thermal denaturation of membrane proteins can help to understand the stability of these proteins at different temperatures. The melting temperature (T_m) is the temperature at which, during thermal denaturation, half of the protein is denatured while the other half remains in its native state. It is an important indicator for measuring the thermal stability of a protein. Understanding the T_m value of a membrane protein can help to optimize production and storage conditions and ensure that the protein remains stable during production and storage. In general, membrane proteins are designed to remain stable at their natural ambient temperatures.

Figure 2.8 shows the temperature-induced unfolding curves of MurJ protein at different pH values monitored by CD spectroscopy. The protein was exposed to temperatures ranging from 5 °C to 95 °C. The T_m values (**Table 2.9**) were calculated from these curves. The corresponding fitting curves for each T_m value can be found in the **Figure S9-S12** in **Appendix**. The observed T_m fell within the range of 58 °C to 62 °C. Therefore, it can be considered that the MurJ protein appears to be relatively stable under pH changes, with a melting temperature of around 58 °C to 62 °C.

Table 2.9 T_m value of MurJ at different pH.

pH	**1**	**3**	**8**	**10**
T_m (°C)	59.84 ± 2.5	58.50 ± 1.3	61.02 ± 0.5	60.60 ± 0.8

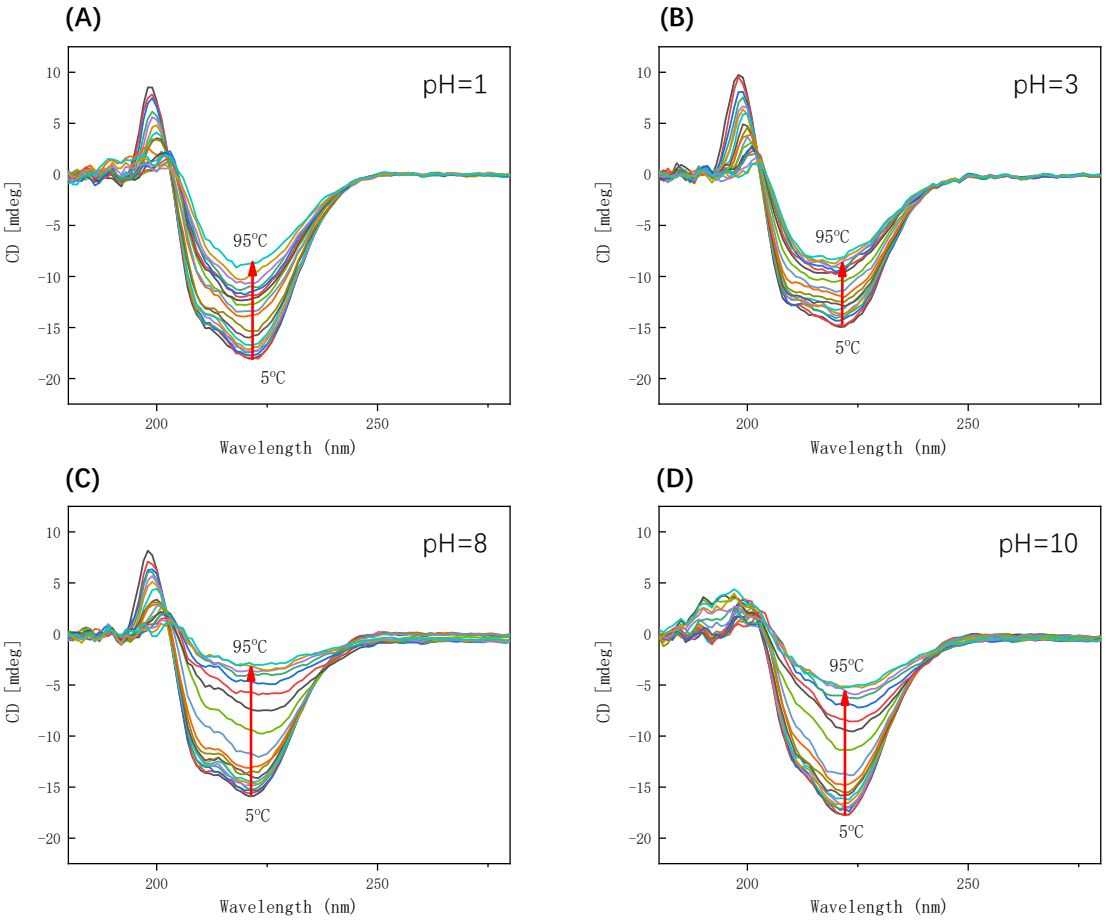


Figure 2.8 CD spectra of temperature-induced unfolding of MurJ. The thermal unfolding of MurJ was monitored using CD in 50 mM Tris buffer at different pH. (A) to (D) correspond to pH values of 1, 3, 8, and 10 respectively. The red arrow shows the change of CD value at 222 nm with increasing temperature.

2.3.5 Determination of cardiolipin content in purified MurJ proteins

The sample delipidation process before mass spectrometry reported in the literature is to dilute the sample in a buffer containing fresh empty micelles, followed by incubation at 37°C overnight [22, 69]. During the incubation, the micelles with lipids were removed by ultrafiltration repeatedly and fresh empty micelles were added. It takes more than ten hours to complete this degreasing process. Here this process was shortened to about 3 hours. The delipidation process was conducted after the protein was bound to the resin. The collection of the protein is relatively easy and does not require several rounds of centrifugal filtration and concentration. This method can maintain a consistent protein concentration throughout the delipidation process, thereby preventing protein aggregation caused by localized concentration disparities during centrifugal concentration. To evaluate the effectiveness of this time-shorten delipidation method, experiments were performed on proteins for 0 h, 3 h, and 16 h. The lipid content of the samples was quantified using a cardiolipin assay kit and liquid chromatography-mass spectrometry (LC-MS). The results are shown in **Figure 2.9** and **Table 2.10**.

From the data calculated from the kit results, it was found that approximately 4 cardiolipins were co-purified with each protein molecule. The amount of lipids contained in the proteins after 16 hours and 3 hours of delipidation was greatly reduced compared to the protein without delipidation. The cardiolipin removal effect of the protein after 3 hours of delipidation reached 87% of that of the 16-hour delipidation process.

From the spectra of the LC-MS, lipids was found to be eluted at about 14 minutes. Based on the change of peak intensity, it was observed that a large amount of lipids has been removed in the flowthrough step of protein purification, but the purified protein without delipidation still retains some lipids. In **Figure 2.9**, D1-1 and D2-2 are the eluate samples collected during the first and second delipidation incubations during the 3-hour delipidation process. Delipidated protein-1 and -2 are proteins purified by imidazole elution after 3-hours delipidation and 16-hours delipidation, respectively. As shown in **Figure 2.9**, the lipid peaks in both the 3h and 16h delipidated protein samples have disappeared.

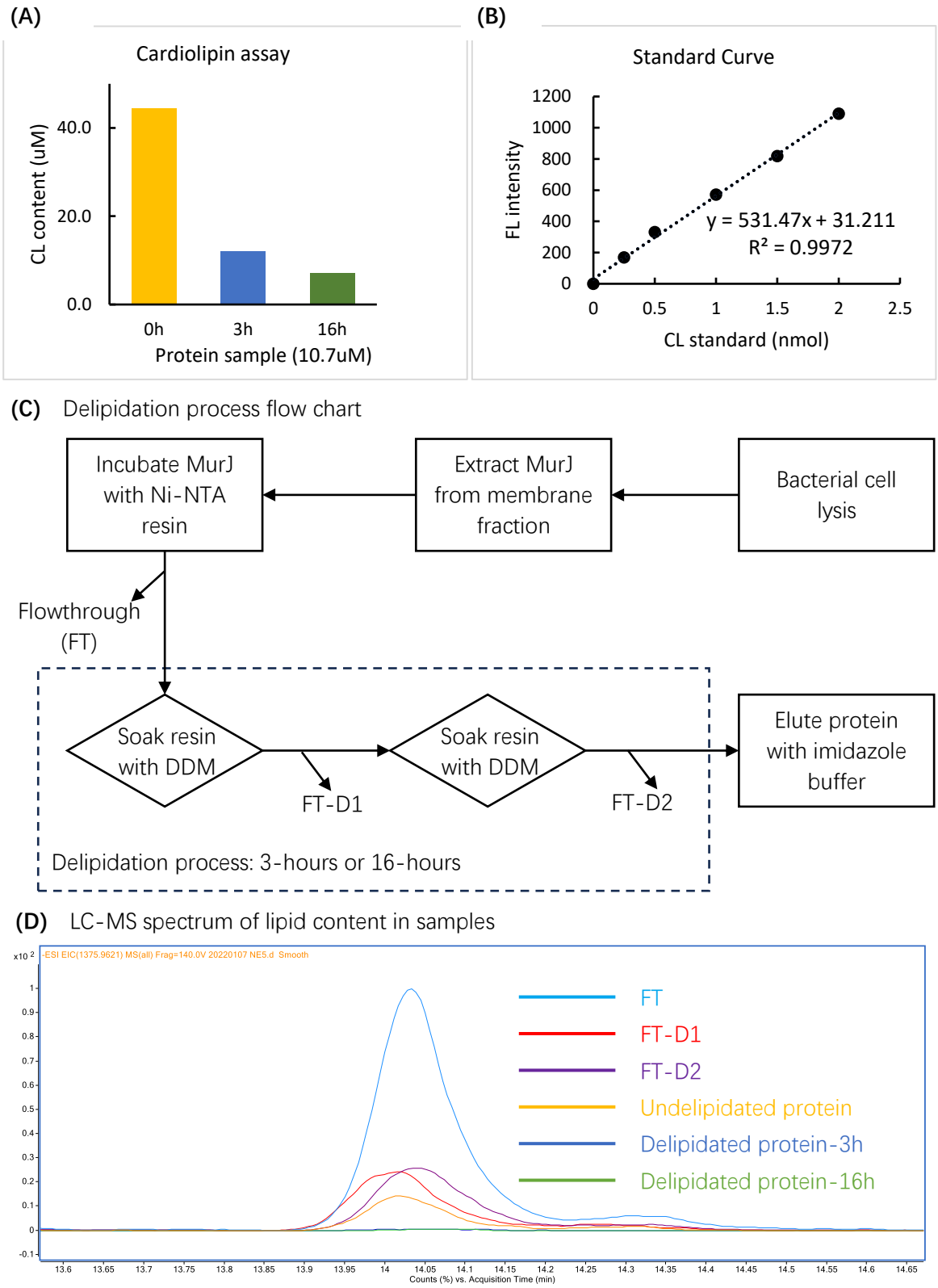


Figure 2.9 Results of lipid content in purified proteins after delipidation process. (A) Cardiolipin content in purified MurJ with or without delipidation. (B) Standard curve used for cardiolipin content calculation. Standard samples were provided with the kit. The standard curve formula is y = 531.47x + 31.211, R^2=0.9972. (C) Flow chart of delipidation process. (D) Spectrum was normalized using software MassHunter.

Table 2.10 Comparison of delipidation process results measured by cardiolipin assay kit.

	CL content (uM)	**CL:protein ratio**	**Delipidated process duration (h)**
Undelipidated protein	44.4	4.2	0
Delipidated protein-1	12.0	1.1	3
Delipidated protein-2	7.1	0.7	16

2.4 Summary

In this chapter, three recombinant plasmids N-His-MurJ, N-His-TEV-MurJ and C-His-MurJ, were designed. These MurJ proteins from *E. coli* were expressed and purified, and the purified proteins were identified by SDS-PAGE and mass spectrometry. To obtain a relatively stable protein expression effect, these three plasmids were used to screen three types of *E. coli* expression systems. The expression of MurJ in C43 (DE3) is relatively stable. Compared with C41 (DE3) and C41 (DE3) pLysS, C43 (DE3) has the highest expression ability of the 3 plasmids with little difference. For subsequent *in vivo* cross-linking experiments, two double-cysteine MurJ mutant were constructed based on previous studies [17], namely E70C_A296C and V43C_T251C.

The CD spectroscopy analysis revealed that the C-His-MurJ protein, which was produced and isolated using the described methodology, exhibits a characteristic helical secondary structure. In particular, MurJ can retain this helical structure when the pH changes drastically from 8.0 to 1.0. The GuHCl unfolding experiment showed that even at concentrations of up to 6 M GuHCl, approximately 35% of the helical structure was retained in the sample compared to its initial helicity level. It is speculated that this retained helical structure may be important for the folding of the MurJ protein. The temperature-induced unfolding study showed that the T_m of MurJ remains relatively consistent under different pH conditions, hovering around 60 °C. In summary, it is hypothesized that the secondary structure of the MurJ protein has resilience to significant fluctuations in pH conditions.

During sample preparation for protein nMS experiments, a rapid delipidation process was optimized. After a 3-hour delipidation process, the lipid signal in the protein sample

disappeared and the protein signal resolution was improved. The overnight delipidation method, which originally required more than 16 hours, was shortened to only 3 hours.

Chapter 3 Driving force of MurJ in conformation transition

3.1 Introduction

Studying activity of flippase proteins is difficult as their substrates maintain their chemical structure both before and after flipping. Enzyme activity is usually assessed by measuring the concentration of product formed or substrate consumed during the enzymatic reaction [85]. It is challenging to distinguish between the structure of a protein's initial and final state. Structural analysis is critical to the study of protein function. X-ray crystallography and nuclear magnetic resonance (NMR) are widely used techniques for studying protein structures in biology [86-90]. However, these methods necessitate the use of highly concentrated purified proteins as samples for analysis. The Substituted Cysteine Accessibility Method (SCAM) is a technique employed to investigate the accessibility and surrounding environment of specific amino acid residues within proteins. In recent years, the SCAM has emerged as a biochemical method for studying protein topology [91, 92].

Butler E. K. et al utilized SCAM and discovered that MurJ contains 14 transmembrane domains (TM) [81]. These domains are arranged into a V-shaped structure, which is consistent with the MOP export proteins. The MATE transporters within the superfamily exhibit great similarity in their structures [93]. Computer modeling predict that TMs 1, 2, 7, and 8 of MurJ will form a central hydrophilic cavity, several charged residues in this cavity were shown to be required for MurJ function, suggesting that MurJ may function in a similar way to the MATE exporters [81, 94]. MATE exporters utilize an antiport mechanism that relies on cation or proton gradients to transport toxic drugs from the cytosol to the periplasm [95].

Recent structural information suggests that the central cavity of MurJ alternates between inward- and outward-open conformations to translocate lipid II molecules across the cytoplasmic membrane, but how these conformational changes occur remains unclear [96]. MurJ is a member of the MOP export transporter superfamily, some of which are driven by sodium or proton electrochemical gradients [97-100]. Recent reports indicated that MurJ couples the membrane potential across the cytoplasmic membrane to lipid II export [23]. In this study, structure-guided cysteine cross-linking and protein native mass spectrometry were used to study the effect of membrane potential on the alternating conformation states of MurJ.

By making reference to previous research [17], two pairs of sites were selected on the protein that can be cross-linked with homobifunctional cysteine cross-linkers and mutated them to cysteine residues to construct two double-cysteine mutants of MurJ, namely E70C_A296C and V43C_T251C. The mutant E70C_A296C will allow 1,6-bismaleimidohexane (BMH) crosslinking when the protein is in an outward-open conformation and then the protein will be structurally locked in this conformation. Another mutant, V43C_T251C, will be locked while in an inward-open configuration. Successfully cross-linked species can be detected by peak shifts in protein mass spectrometry spectra. Details of the location of the cysteine residue pairs are shown in **Figure 3.3** of the results and discussion section.

To study the effect of membrane potential on protein movement *in vivo*, the proteins will be preincubated with three compounds that affect proton motive force before cross-linking, and then cross-links will be monitored using mass spectrometry. The proton motive force (PMF) is composed of two parts: membrane potential ($\Delta\psi$) and pH

difference (ΔpH) [24]. Membrane potential is the electrical potential difference that occurs when the interior of the cytoplasm is negative relative to the periplasm and is a measure of the difference in charge across the membrane. The pH difference refers to the alkalinity inside the cytoplasm relative to the periplasm and represents the difference in proton concentration across the membrane. 3,3',4',5-Tetrachlorosalicylanilide (TCS) can affect both the membrane potential and ΔpH, valinomycin can destroy the membrane potential but has no effect on the pH difference, and nigericin solely affects the pH difference [23].

Native mass spectrometry (nMS) with nano-electrospray ionization (nano-ESI), allows the use of volatile aqueous buffers, milder temperatures and voltages to retain non-covalent interactions and quaternary protein structures [37, 38]. It presents several advantages in studying membrane proteins compared to traditional biophysical techniques used for structural analysis [70, 101]. Conventional techniques such as nuclear magnetic resonance (NMR) and cryo-electron microscopy (cryo-EM) are usually limited by the size of the samples being studied [102-104]. For example, NMR is suitable for small proteins, while cryo-EM is more appropriate for larger protein complexes. nMS is compatible with a broad mass range, ranging from small transmembrane peptides to large multiprotein complexes [35, 105, 106]. Furthermore, native MS is fast and sensitive, capable of detecting samples at concentrations as low as nanomolar levels [70]. In contrast, traditional structural biology methods such as X-ray crystallography usually require relatively high concentrations and quantities.

Recently, in addition to its application in studying higher-order structures of membrane proteins (MPs), nMS has also been used in investigating the effects of lipids on proteins [107, 108]. Bolla and colleagues utilized native MS analysis to discover that a co-

purified endogenous lipid, cardiolipin (CL), competes with the substrate of the protein MurJ for binding [22]. The observation of the protein-ligand complex, along with the determination of the binding constant, serves as crucial evidence supporting the identification of MurJ as the lipid II flippase. Here, the nMS technique was used to study the proton motive force (PMF) influence of MurJ transition.

3.2 Methods

3.2.1 Materials

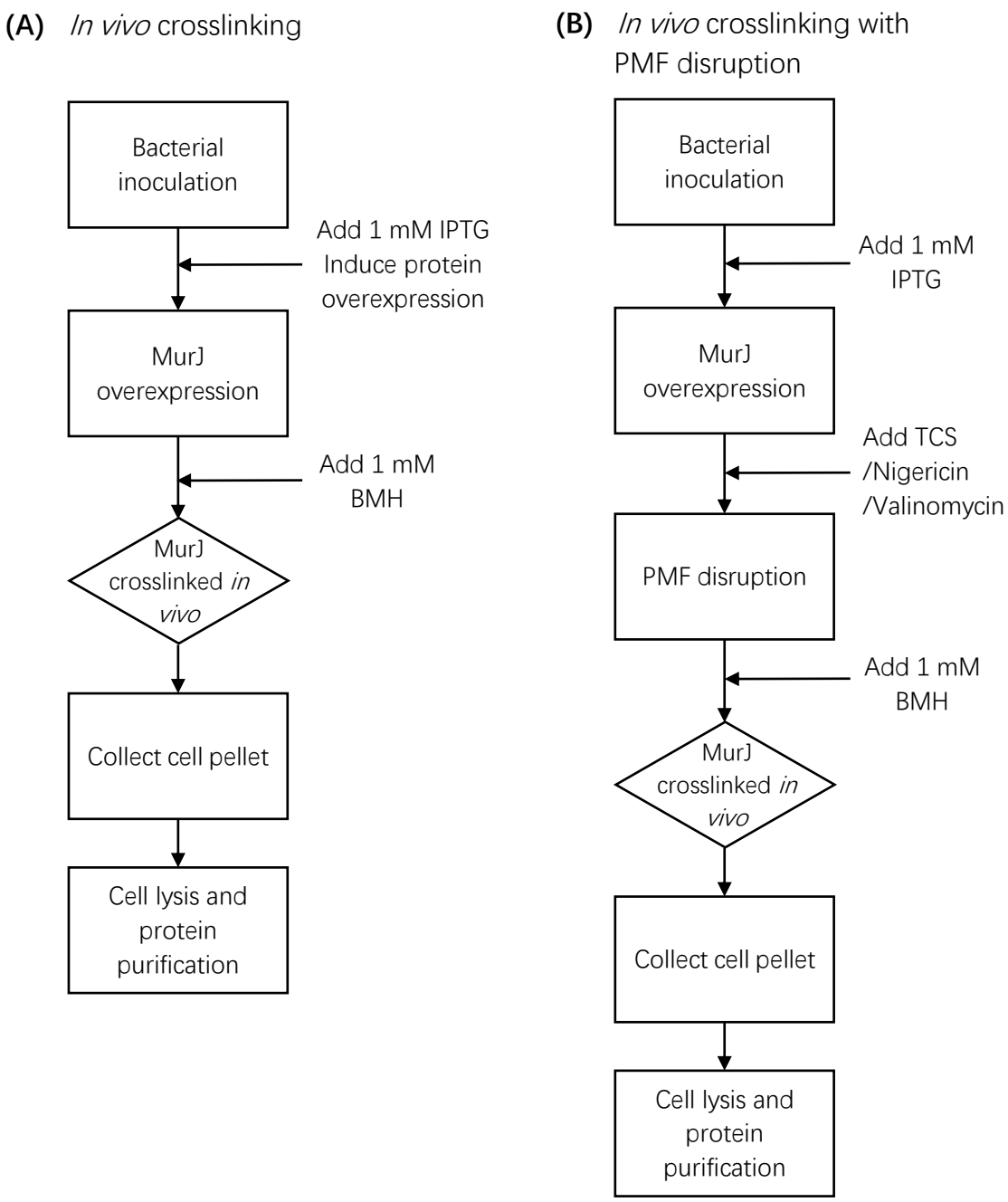
Dimethylsulfoxide (DMSO) and BMH (bismaleimidohexane) were purchased from Thermo Fisher Scientific (HK). Amicon centrifugal filters were purchased from Merck Millipore. Gravity and spin chromatography columns were purchased from Bio-Rad. Formic acids were purchased from Sigma-Aldrich.

3.2.2 *In vivo* cysteine cross-linking

Plasmid transformation and protein expression were completed as described in Chapter 2. After bacteria were diluted 1:100 in LB medium and grown to OD_{600} = 0.6 - 0.8, IPTG was added to 1mM to initiate the overexpression of target proteins at 37 °C for 3 hours. The pellet collected from each 350 mL culture medium was resuspended in 5 mL of LB medium. To this, 250 µL of 20 mM homobifunctional cysteine cross-linker BMH was added and incubated at 37 ºC in the dark for 30 minutes. For control, DMSO was added instead of BMH.

For membrane potential disruption experiments, 50 µL of TCS (10 mM), nigericin (2.5 mg/mL), valinomycin (2.5 mg/mL), or DMSO as a control was added prior to the addition of the cross-linker. The mixture was then incubated at 37 ºC for 10 min. In the case of the valinomycin treatment group, bacteria cultured in LB Miller medium containing 170 mM NaCl necessitate the addition of 200 mM KCl before compound incubation. And for the TCS, Nigericin, and DMSO control group, bacteria need to be supplemented with 200 mM NaCl.

Following the cross-linking process, the mixtures were centrifuged at 10,000 × g for 3 minutes to collect the cell pellet, and washed with lysis buffer (20 mM HEPES, 150 mM NaCl, pH 7.0). Then the cell pellet was resuspended at 20 mL lysis buffer and disrupted through sonication. The ultrasonic energy level is set at 40%, with each treatment cycle lasting 5 seconds followed by a cooling period of 15 seconds. The overall duration of ultrasonic treatment amounts to 20 minutes. The protein purification was performed as described in Chapter 2. The workflows of *in vivo* cross-linking experiments are shown in **Scheme 3.1**.



Scheme 3.1 The workflows of *in vivo* cross-linking experiments.

3.2.3 Native Mass Spectrometry

Native mass spectrometry (nMS) was performed using a SYNAPT G2-Si (Waters) instrument. Referring to the protein native mass spectrum reported by previous articles, due to the large molecular weight of the protein, the mass-to-charge peak will appear in the relatively high area [22]. The instrument was calibrated using the compound sodium iodide (NaI) at the range of m/z 1000 to 7000 (error < 5 ppm), and data acquisition parameters was adjusted to enhance the signal between the m/z 5000 to 7000 region (refer to **Table 3.1**). The sample was loaded into a homemade gold-coated nano-electrospray capillary. The capillary voltage used in the native MS experiments was 1.5 kV, the sampling cone was 80 L/h, the ionization source temperature was 40°C, and gas flow of the nano flow was 0.2 Bar. Other parameters for nano-electrospray ionization mass spectrometry (nano-ESI-MS) application were set as follows, the trap collision energy is 80~110 V, the transfer collision energy is 60~80 V, and the trap gas is controlled at 2.0~6.0 mL/min. Subsequent analysis of the acquired mass spectrometry data was conducted using the MassLynx platform (version 4.1, Waters, UK), which mainly performed noise subtraction and peak smoothing of the combined signals. The corresponding molecular weights were calculated on the ESIprot Online website [109].

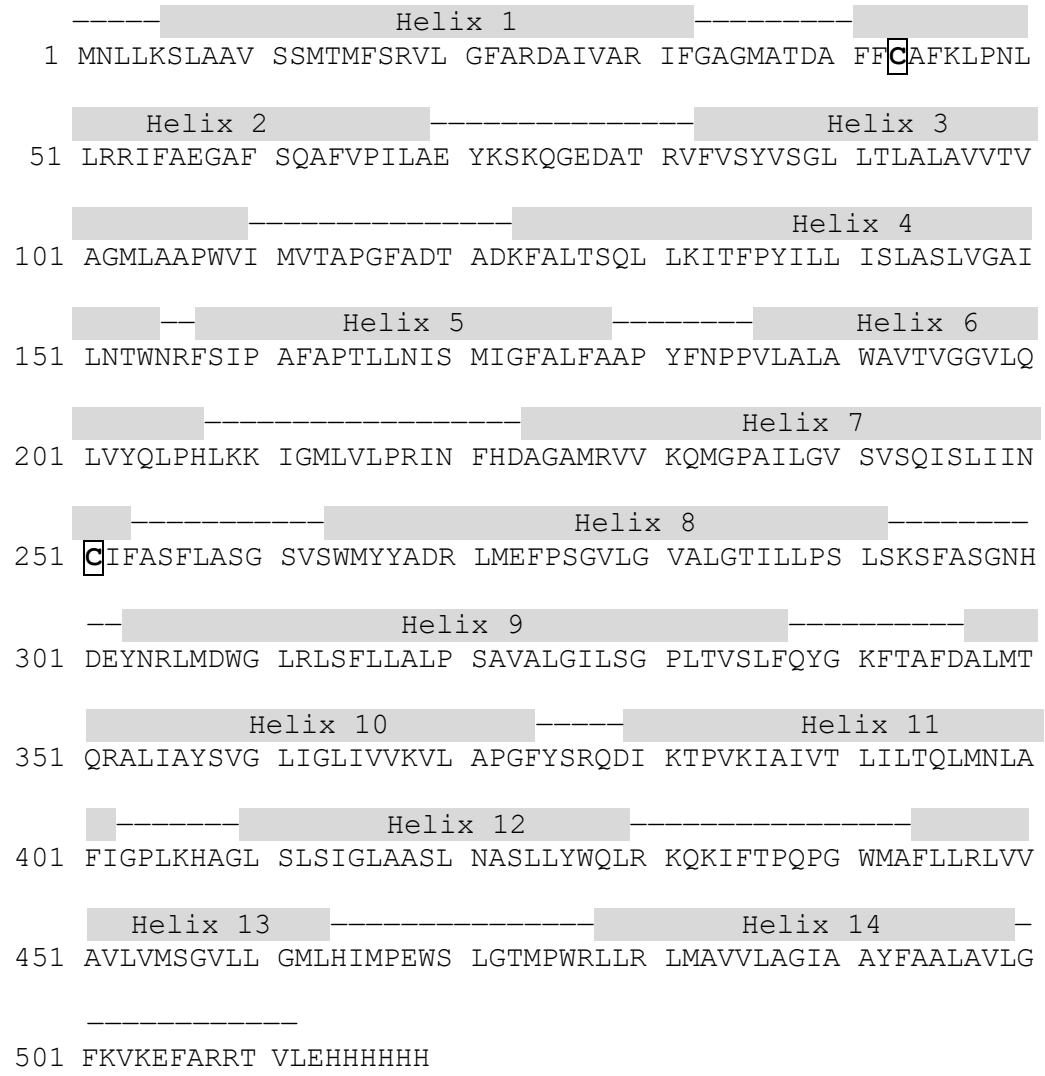
Table 3.1 Acquisition parameters used in nMS.

	Mass	**Dwell Time [% Scan Time]**	**Ramo Time [% Scan Time]**
1	1000	5	5
2	4000	10	80
3	6000	-	-

3.3 Results and discussion

3.3.1 *In vivo* cross-linking

The protein amino acid sequences of mutants V43C_T251C and E70C_A296C are shown below. Theoretical molecular weights of proteins were calculated using an online Compute pI/Mw tool of Expasy (SIB Swiss Institute of Bioinformatics, Switzerland). The full amino acid sequences of V43C_T251C and E70C_A296C are shown in **Figure 3.1** and **Figure 3.2**, respectively.



Theoretical Mw: 56306.51

Figure 3.1 Full amino acid sequence of mutant V43C_T251C.

Theoretical Mw: 56363.60

Figure 3.2 Full amino acid sequence of mutant E70C_A296C.

The location of the mutated cysteine pairs in the protein is shown in **Figure 3.3**. The inward-open conformation model used the PDB file of *E. coli* MurJ protein (PDB code: 6CC4). The outward-open conformation model used the PDB file of the *T. africanus* MurJ protein (PDB code: 6NC9). The cysteine positions marked in this model are E70 and K281, which were determined by the alignment result of *E. coli* MurJ and *T. africanus* MurJ protein sequences [14]. The sequence alignment map of MurJ from various strains reported was shown in **Figure S13** in **Appendix** [60]. The cysteine pair in V43C_T251C will be close to each other when the protein changes to the inward-open configuration, and the cysteine pair in E70C_A296C will be close to each other when the protein shifts to the outward-open configuration. The shortened distance between cysteine residues allows bifunctional cross-linkers with fixed lengths to achieve intra-protein cross-linking, thereby fixing the protein the respective conformation.

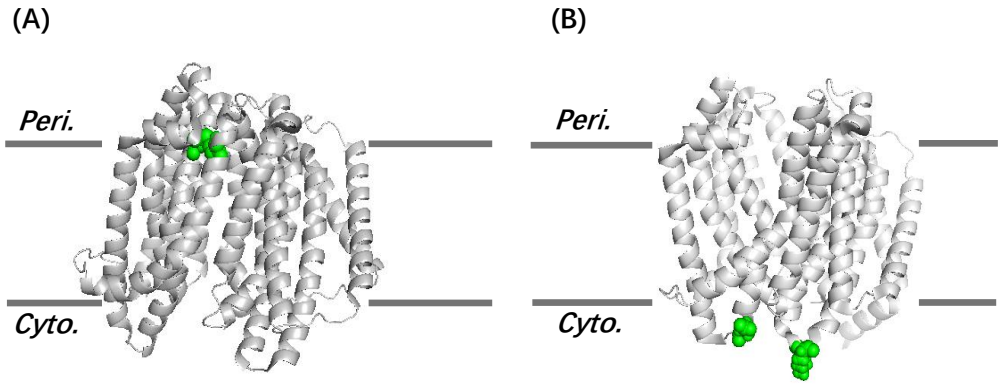


Figure 3.3 Locations of the residue pairs that were mutated into double cysteines in MurJ. The green spheres represent the mutated residues. (A) shows MurJ in the inward-open conformation (PDBID 6CC4) with V43 and T251. The fused apocytochrome b562RIL (BRIL) protein has been removed. (B) shows an outward-open conformation (PDBID 6NC9) with E70 and A296.

Purified proteins were analyzed using 12% SDS-PAGE and the dual color standards protein (Bio-Rad) was used as marker (**Figure 3.4**). The MurJ protein band was observed at approximately 37 kDa in size. The mass change of the cross-linked protein was undetectable in SDS-PAGE because the molecular weight of the cross-linker is low compared to the protein itself.

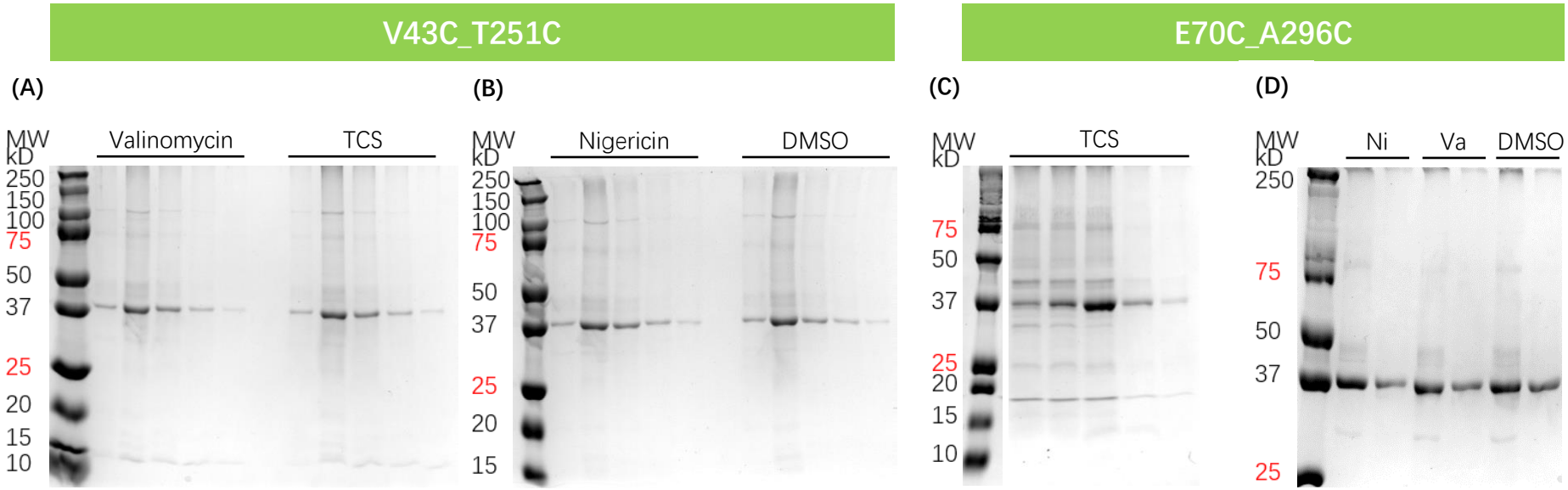


Figure 3.4 SDS-PAGE results of purified cross-linked protein.

(A) (B) are SDS-PAGE data of mutant V43C_T251C, and (C) (D) are data of mutant E70C_A296C. (A) Samples are purified fractions of BMH cross-linked MurJ after treatment with valinomycin or TCS. (B) The figure shows the nigericin treatment group and the DMSO control group data. (C) The figure shows the TCS treatment group. (D) The figure shows the nigericin group, valinomycin group and DMSO control group of mutant E70C_A296C. Ni represents nigericin, Va represents valinomycin.

The flipping cycle of MurJ *in vivo* can be delineated into two distinct processes. Initially, in the inward-open conformation, MurJ transitions to an outward-open state upon substrate binding. Subsequently, in the second process, MurJ reverts from the outward-open state back to the inward-open conformation following substrate release.

In this chapter, the effect of the proton motive force (PMF) on the flipping cycle was investigated by *in vivo* cross-linking experiments using the two mutants constructed above (**Figure 3.5**). *In vivo*, MurJ undergoes cyclic conformational alterations. For the inward-open mutant, cross-linking will occur when MurJ is in the inward-open state. Conversely, for the outward-open mutant, MurJ will be cross-linked in the outward-open conformation.

The *in vivo* cross-linking results of the protein were confirmed by the native mass spectrum. Successfully cross-linked proteins can be detected by peak shifts in the protein mass spectrum. The influence of PMF disruption can be assessed by observing variations in the peak intensity of cross-linked and uncross-linked proteins. If PMF disruption does not affect the flipping cycle, MurJ is able to flip to specific conformation, resulting in the emergence of the cross-linked protein peak. Conversely, if PMF disruption does impede the flipping cycle, MurJ is unable to achieve specific conformation, leading to the absence of the cross-linked protein peak.

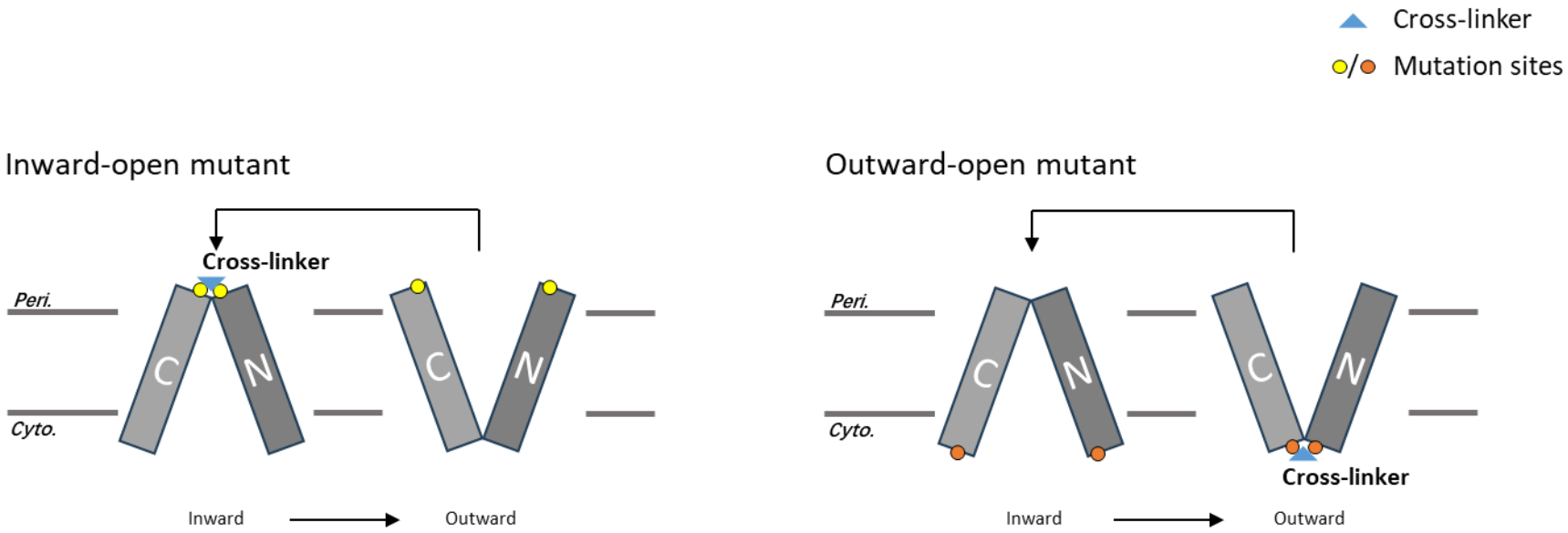


Figure 3.5 Designed cross-linking model of MurJ mutants *in vivo*. The blue triangles represent cross-linkers, and the yellow and orange dots represent the selected residue pairs in the inward-open mutant and outward-open mutant, respectively.

There are four possible situations for the influence of PMF on MurJ conformational cycling (**Table 3.2**). The first situation is that PMF disruption does not affect the flipping of MurJ. In this case, as the conformation of MurJ changes *in vivo*, the protein in the inward mutant will be cross-linked upon transitioning to the inward state, while the protein in the outward mutant will be cross-linked when it flips to the outward state. Consequently, cross-linked protein peaks will be observable in the spectra of both mutants.

The second situation is that PMF disruption affects the transition of MurJ from the inward to the outward state. In this instance, the MurJ protein *in vivo* would be retained in the inward state. Therefore, cross-linking is feasible in the inward-facing mutant but not in the outward-facing mutant. As a result, the cross-linked protein peaks will exclusively be present in the spectra of the inward-facing mutant. The third situation is that PMF disruption influences MurJ flipping on the outward to inward process. The cross-linking situation will be the opposite of the second case.

The last situation is that PMF disruption affects the flipping activity of MurJ in both processes. Therefore, MurJ protein *in vivo* will be trapped in both inward and outward states, so in both mutants, partial MurJ protein will be cross-linked. Therefore, the cross-linked protein peak will be observable in the spectra of both mutants.

Table 3.2 Possible situations for the influence of PMF on MurJ conformational cycling.

	Influence of PMF disruption	Crosslinking effect (cross-linked protein peaks)	
		Inward-open	Outward-open
1	No influence	YES	YES
2	Influence inward to outward process	YES	NO
3	Influence outward to inward process	NO	YES
4	Influence the cycle on both processes	Partial	Partial

3.3.2 Native mass spectrometry results

3.3.2.1 Native mass spectrum of purified MurJ

The native mass spectra of purified MurJ V43C_T251C with (A) or without (B) delipidation treatment are shown in **Figure 3.6**. The signals in the untreated sample were too noisy for peak annotation and the subsequent mass calculation. The data of the delipidated samples were calculated using the online platform ESIprot [109]. The calculation details are shown in **Table 3.3**. The molecular weight of this series of peaks was 56306.64 ± 0.7 Da, which was consistent with the theoretical value 56306.51 Da of MurJ protein. The results indicate that delipidation prior to native mass spectrometry analysis of membrane proteins can significantly enhance protein ionization.

Table 3.3 Protein's charge states and molecular weight of proteins calculated from ESI mass spectrometry data. The theoretical molecular weight of V43C_T251C is 56306.51 Da. The standard error was calculated using the following formula, where $\bar{x}$ is the sample's mean and n is the sample size.

Formular: $\sigma_{\bar{x}} = {}^{\sigma}/_{\sqrt{n}}$

peak	m/z	charge	MW [Da]	Error [Da]
1	3313.1162	17+	56305.84	-0.8
2	3520.1804	16+	56306.76	0.1
3	3754.9141	15+	56308.60	2.0
4	4022.8181	14+	56305.34	-1.3
Deconvoluted MW [Da]: 56306.64				
Standard deviation [Da]: 1.4				

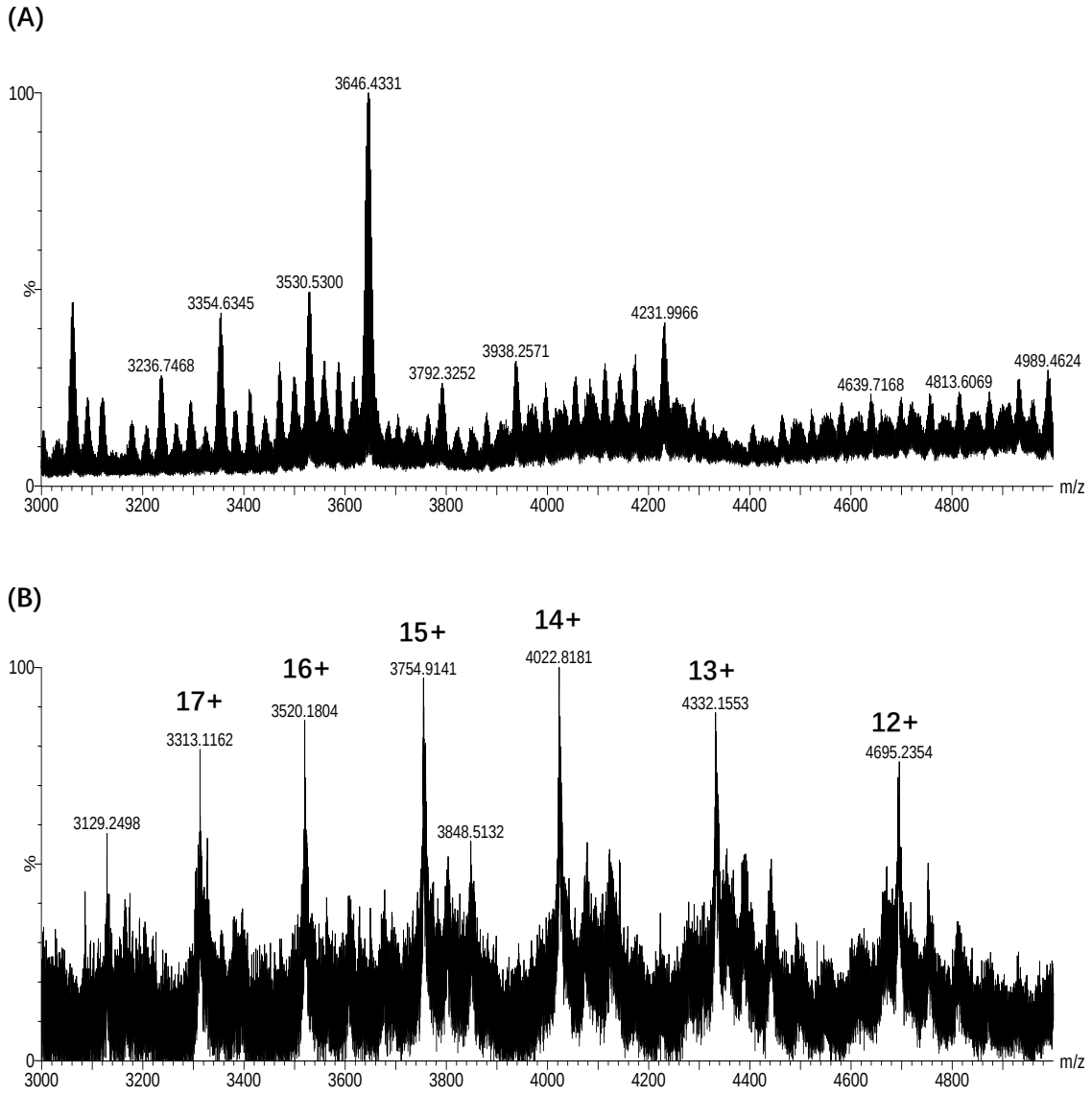


Figure 3.6 Native mass spectrum of purified V43C_T251C MurJ protein with or without delipidation. (A) nMS spectrum of purified MurJ without delipidation. (B) nMS spectrum of purified MurJ with delipidation.

3.3.2.2 Native mass spectrum of cross-linked MurJ

Before the *in vivo* cross-linking experiment was conducted, the effect of incubation time on the cross-linking of the MurJ protein was first studied *in vitro*. The MurJ samples were incubated for 0, 10, 20, 30, and 60 minutes to find out the optimal cross-linking time. In this experiment, the purified MurJ protein was added with the cross-linker for different incubation time. The protein used in this experiment is the purified mutant E70C_A296C. When the protein is successfully cross-linked, the peak will shift toward higher m/z values. The effect of protein cross-linking is determined by comparing the intensity of the protein peaks between the free and cross-linked species.

From the results shown in **Figure 3.7**, even after 60 minutes of incubation, some proteins still cannot be cross-linked with BMH. In the 20-minute incubation, a peak indicating non-specific binding (unsuccessful cross-linking) between MurJ and BMH appeared. Comparison between the data at 20-minute and 30-minute incubation reveals that the formation of the MurJ:BMH complex continued to increase over the latter 10 minutes. Considering all these results, a cross-linking incubation time of 30 minutes was selected.

The software XlinkX 2.5 (Thermo Fisher) was used to examine the intramolecular cross-linking of MurJ by BMH. Analysis revealed the presence of a peptide sequence RIFAEGAFSQAFVPILA(**C**)YK-SF(**C**)SGNHDEYNR, indicating that the peak of MurJ:BMH in nMS represents successful cross-linking. The cross-linked samples were digested following the same method described above for LC-MS sample preparation. SpeedVac dried samples were resuspended with 0.1% formic acid (FA).

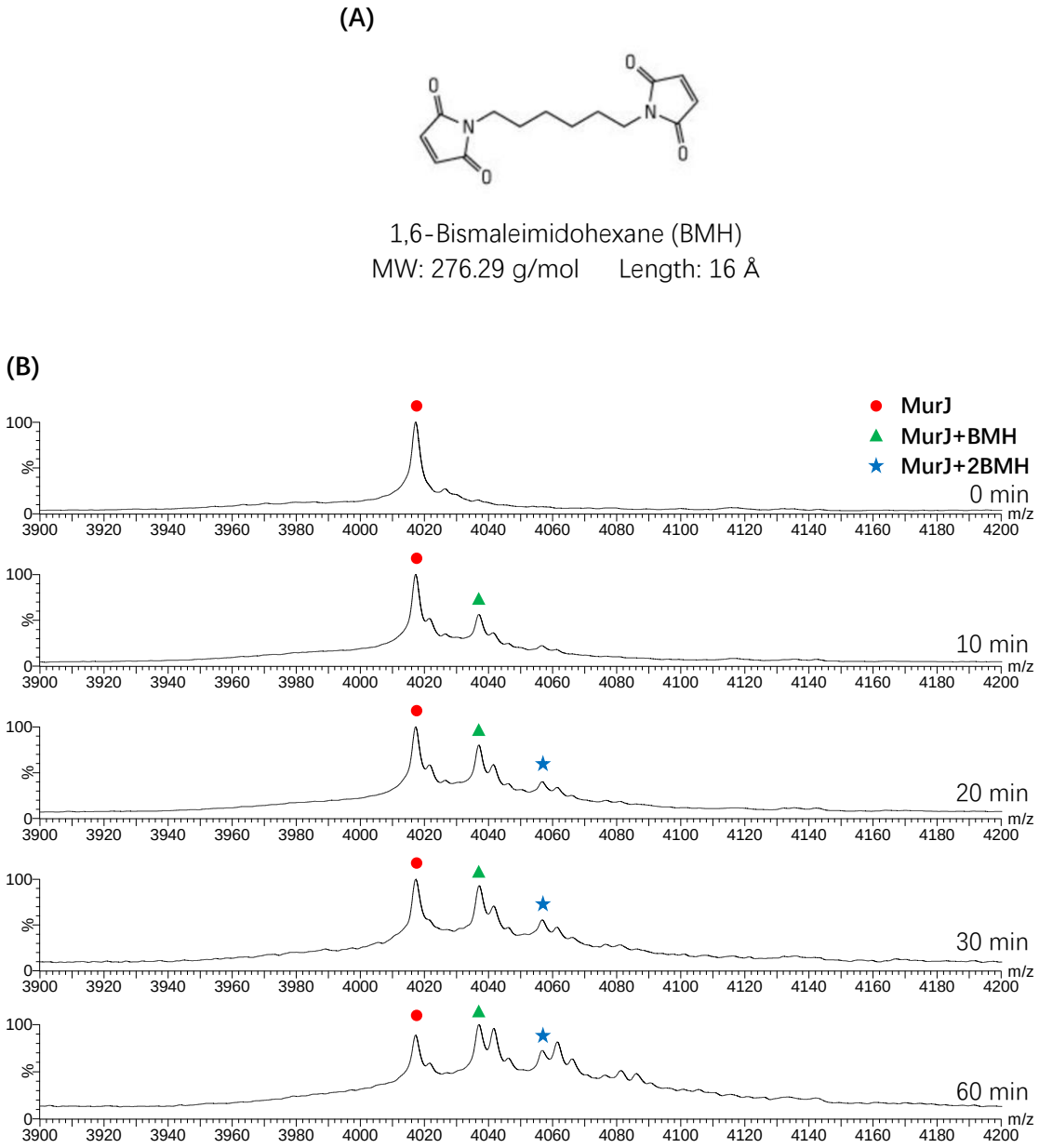


Figure 3.7 *In vitro* protein intramolecular cross-linking efficiency at different incubation times. (A) The chemical structure of cross-linker 1,6-Bismaleimidohexane (BMH). (B) nMS spectrum of purified MurJ incubated with BMH at different times. The red dots mark the peaks of purified MurJ E70C_A296C, green triangles are the peaks of the MurJ:BMH complexes (possibly successful cross-linking), and blue stars are peaks of the complexes of MurJ adducted with 2 BMH (unsuccessful cross-linking).

Figures 3.8-3.11 are the native mass spectrum of the two purified mutants V43C_T251C and E70C_A296C with or without cross-linking with BMH *in vivo*. The theoretical mass and experimental mass details of each sample are listed in **Table 3.4**. The mass of mutant E70C_A296C is 131 Da less than the theoretical mass, which corresponds to the loss of a methionine. N-terminal methionine excision (NME) is one of the most common post-translational modifications of proteins [79]. Although methionine is used to initiate protein synthesis of almost all proteins, in most cases it is subsequently removed.

Table 3.4 The theoretical mass and experimental mass of purified mutants V43C_T251C and E70C_A296C with or without cross-linking with BMH *in vivo*.

sample	Theoretical MW [Da]	Experimental MW[Da]
V43C_T251C	56306.51	56306.64±0.7
V43C_T251C_BMH	56582.80	56583.34±2.7
E70C_A296C	56363.60	56233.47±0.8
E70C_A296C_BMH	56639.89	56523.73±9.5

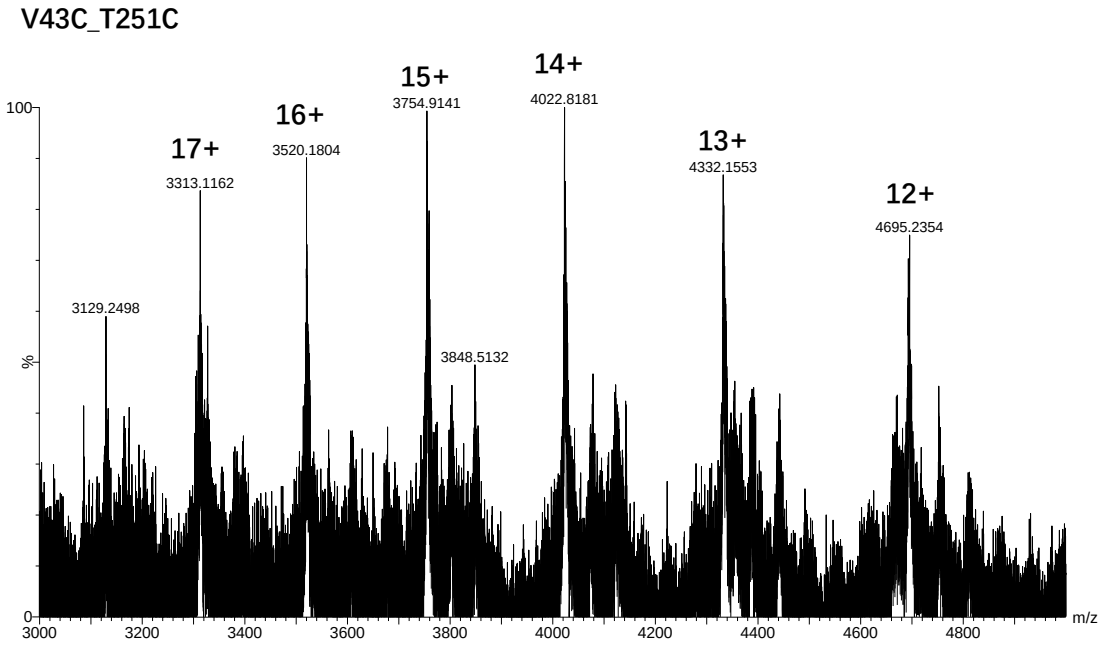


Figure 3.8 nMS spectra of purified MurJ V43C_T251C.

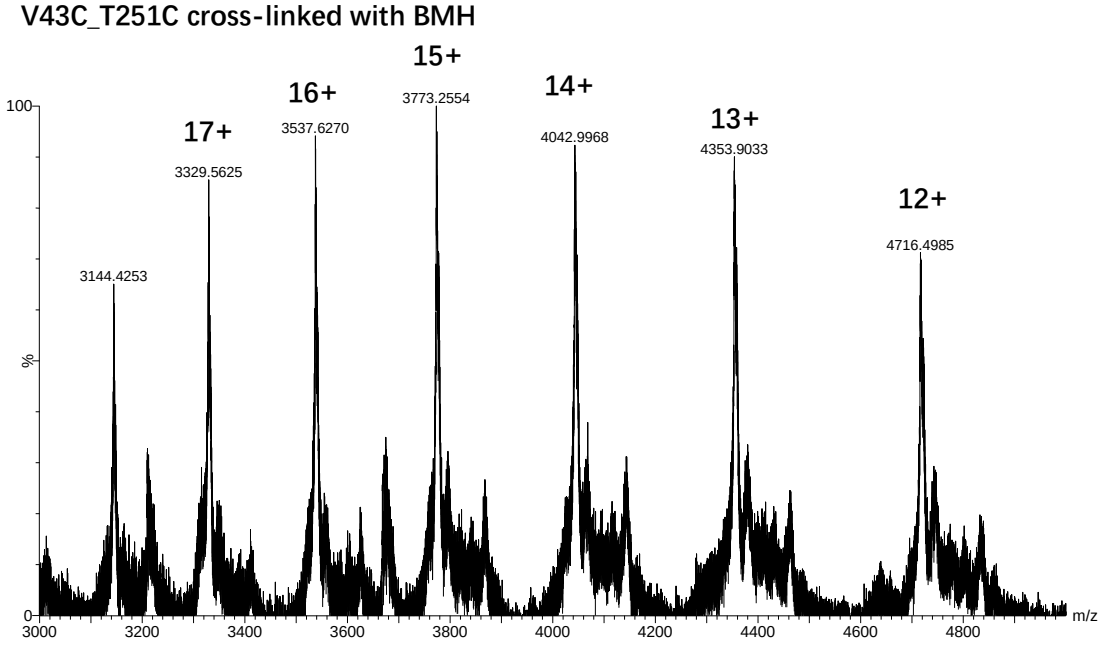


Figure 3.9 nMS spectra of purified MurJ V43C_T251C cross-linked with BMH *in vivo*.

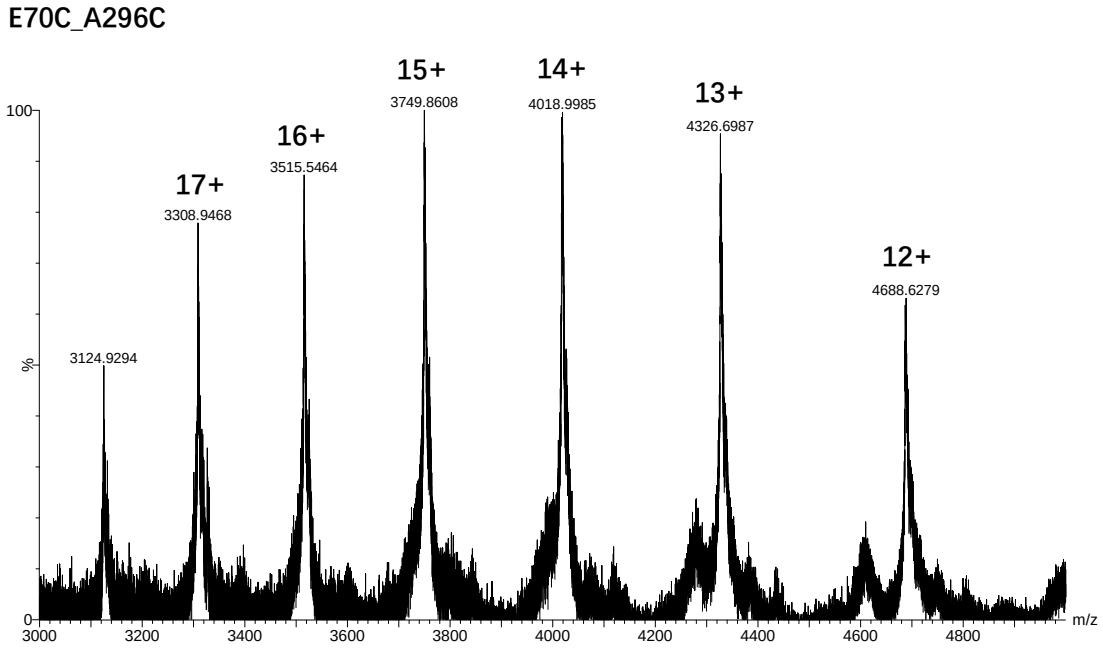


Figure 3.10 nMS spectra of purified MurJ E70C_A296C.

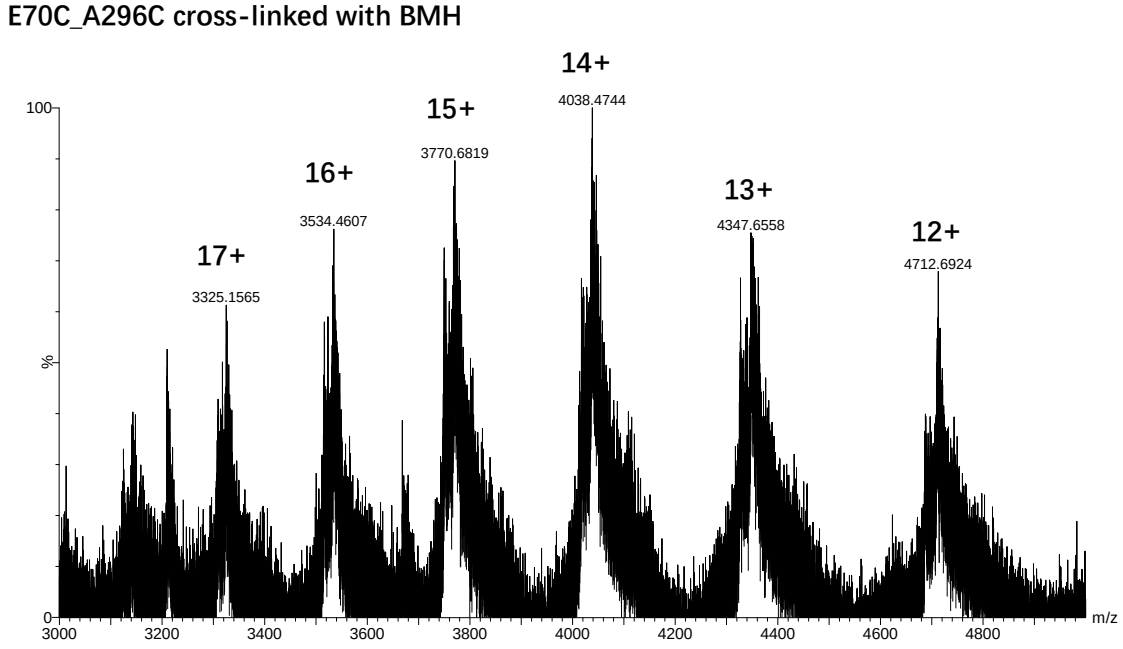
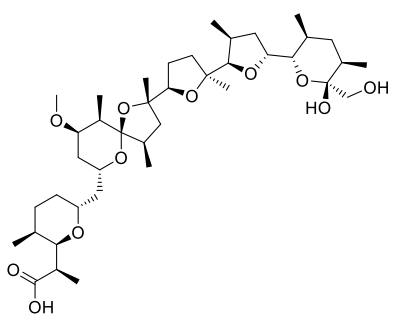


Figure 3.11 nMS spectra of purified MurJ E70C_A296C cross-linked with BMH *in vivo*.

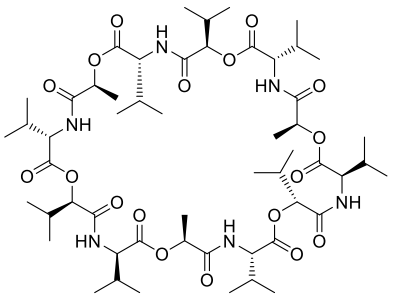
3.3.2.3 Native mass spectrum of MurJ V43C_T251C cross-linking with/without PMF disruption

Previous studies suggest that MurJ acts as a flippase that uses an alternating-access mechanism to transport lipid II across the membrane, cycling between the inward-facing and outward-facing conformations [13-15, 17, 96]. After incubation with the cross-linker BMH, the pair of mutated cysteine residues in the protein will bind to the homobifunctional maleimide cross-linkers to form covalent bonds, thereby fixing the protein in one conformation. Here, two double–cysteine mutants of MurJ, mutant V43C_T251C and E70C_A296C would be fixed in either the inward-facing or outward-facing conformation, respectively, when bound to the linker. The lipid II flippase activity of MurJ has been reported to rely on the membrane potential, but it is still unclear at which stage of the conformational cycle does the membrane potential take its effect [23]. In this study, three compounds were used to disrupt the proton motive force and the changes in the conformations of MurJ were studied *in vivo*.

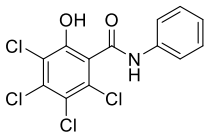
The proton motive force (PMF) consists of two components: the membrane potential ($\Delta\psi$) and the pH difference (ΔpH). The compound 3,3',4',5-Tetrachlorosalicylic acid aniline (TCS) can affect both the membrane potential and ΔpH. Valinomycin can disturb the membrane potential without impacting the pH gradient. Conversely, nigericin solely influences the pH gradient. The molecular structures of these three compounds are shown in **Figure 3.12**.



Nigericin



Valinomycin



Tetrachlorosalicylanilide
(TCS)

Figure 3.12 Structures of nigericin, valinomycin and tetrachlorosalicylanilide.

The *in vivo* cross-linking results of proteins after interference by different chemical compounds are shown in **Figures 3.13-3.18**. **Figures 3.13-3.15** are the results of mutant V43C_T251C, which can be cross-linked in an inward-open conformation. After cross-linking, the protein mass peak shifted to the right, a +1BMH peak appeared. We added compounds that have different effects on the membrane before cross-linking to observe the effects of these compounds on protein conformational change.

In **Figure 3.13**, the cross-linked protein spectra without treatment, after pretreatments, and the overlaid are shown from top to bottom, respectively. As shown in the overlaid spectrum, there is no new uncross-linked protein peak appears, indicating that the compound treatment does not affect protein cross-linking. Taken together, the results in **Figure 3.13-3.15** showed that MurJ can be cross-linked and fixed in the inward-open conformation in the presence of the three compounds. This indicates that the disruption of the proton motive force does not affect the progression from other conformations to the inward-open conformation during conformational cycling. This stage probably occurs after MurJ has translocated lipid II outside the membrane and is reset in preparation for interaction with a new substrate [14]. In other words, after the protein completes the substrate flipping, it returns to the inward-open state to prepare for the next substrate binding and transport. In summary, this process does not rely on the proton motive force.

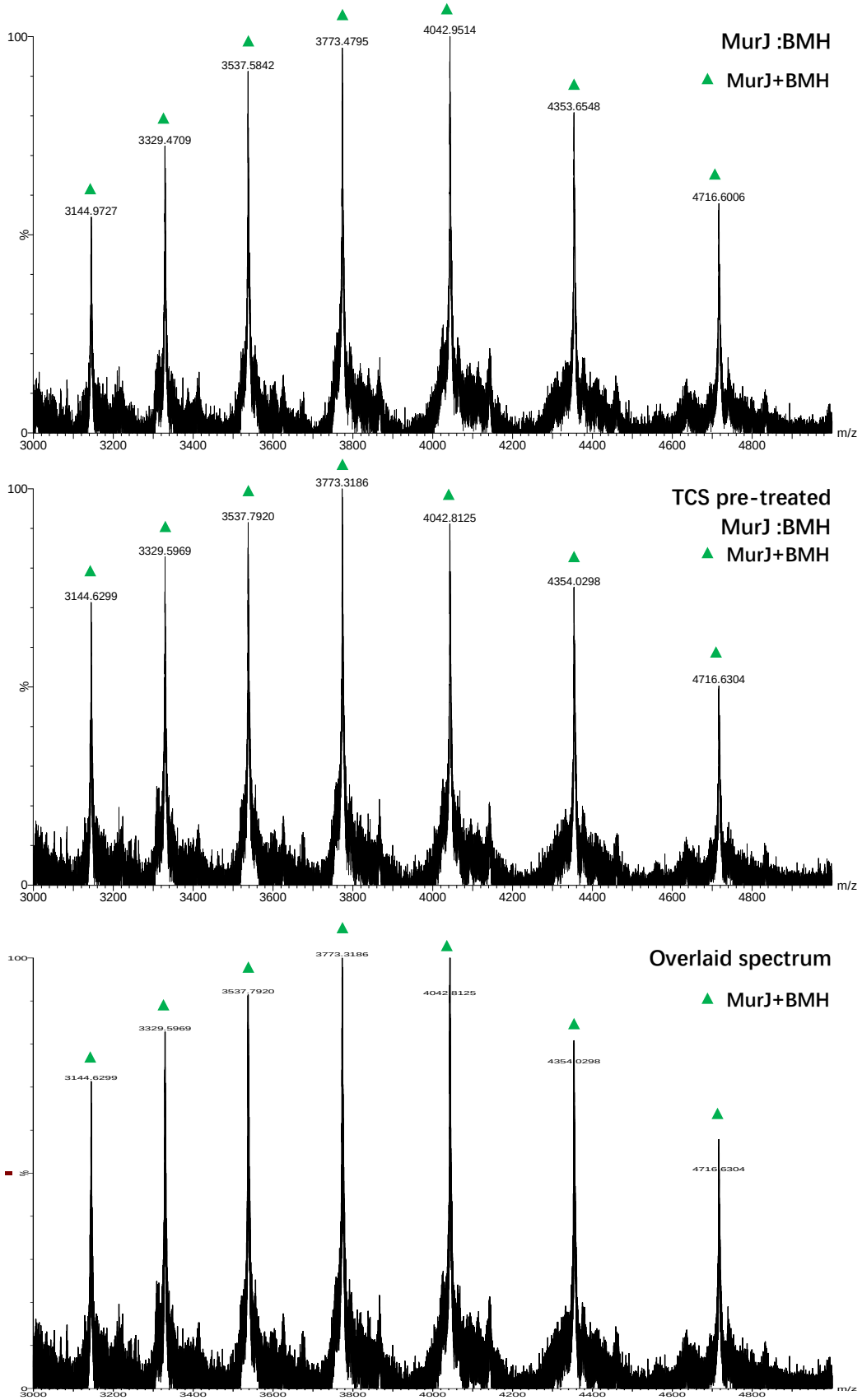


Figure 3.13 nMS spectrum of BMH cross-linked V43C_T251C with/without PMF disruption (TCS-pretreated).

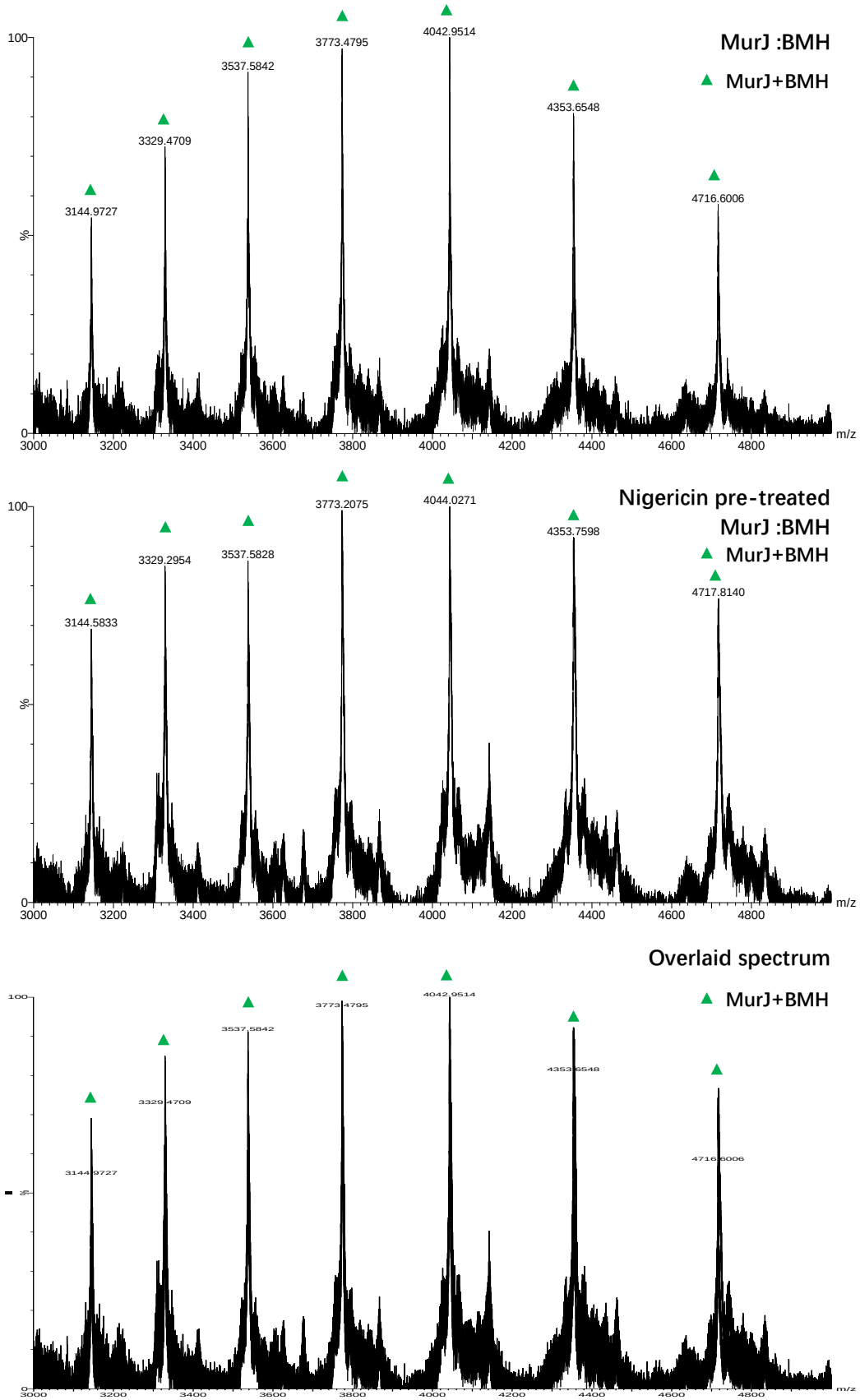


Figure 3.14 nMS spectrum of BMH cross-linked V43C_T251C with/without PMF disruption (Nigericin-pretreated).

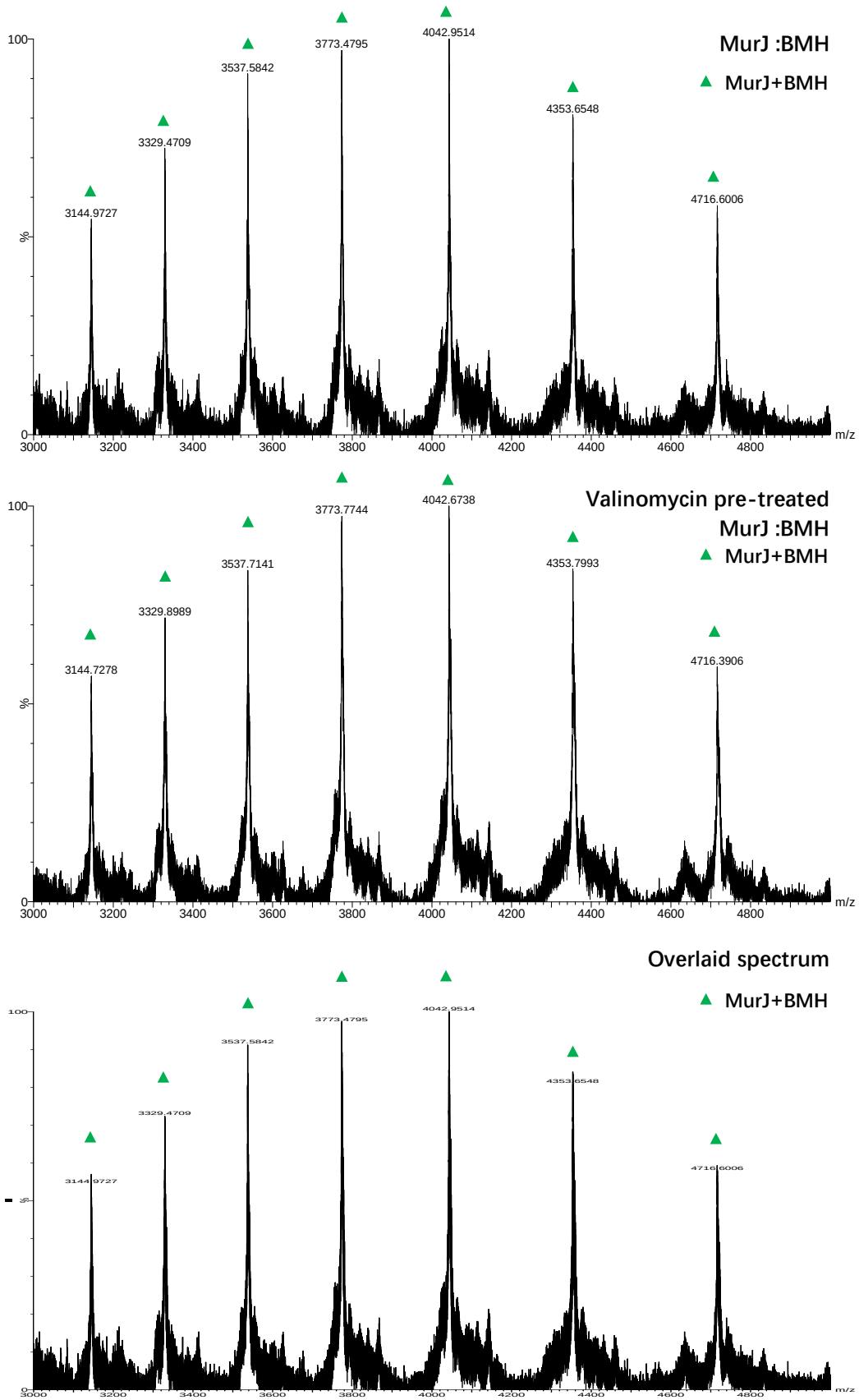


Figure 3.15 nMS spectrum of BMH cross-linked V43C_T251C with/without PMF disruption (Valinomycin-pretreated).

3.3.2.4 Native mass spectrum of MurJ E70C_A296C cross-linking with/without proton motive force (PMF) disruption

Figures 3.16-3.18 show the native mass spectrum of mutant E70C_A296C, which can be cross-linked in an outward-facing conformation. In **Figure 3.16**, an increase in the intensity of the non-cross-linked MurJ peak (red dots) was observed after incubation with TCS resulting in a shift in the relative ratio between the MurJ and cross-linked MurJ:BMH peaks. A portion of MurJ proteins failed to cross-link in the outward-facing conformation, suggesting that movement of these proteins toward the outward-open conformation is impeded. The outward-facing conformation occurs when MurJ transports the substrate out of the membrane. In other words, the disturbance in PMF will cause MurJ to stagnate in the inward-open conformation and fail to complete the transmembrane transport.

Figures 3.17 and **3.18** are the cross-linking results of E70C_A296C under the PMF disruption caused by nigericin or valinomycin, respectively. The cross-linking effects of the proteins after treatment with these two compounds are similar. Nigericin showed a slight inhibitory effect, while valinomycin treatment had no obvious impact on the cross-linking effect of proteins.

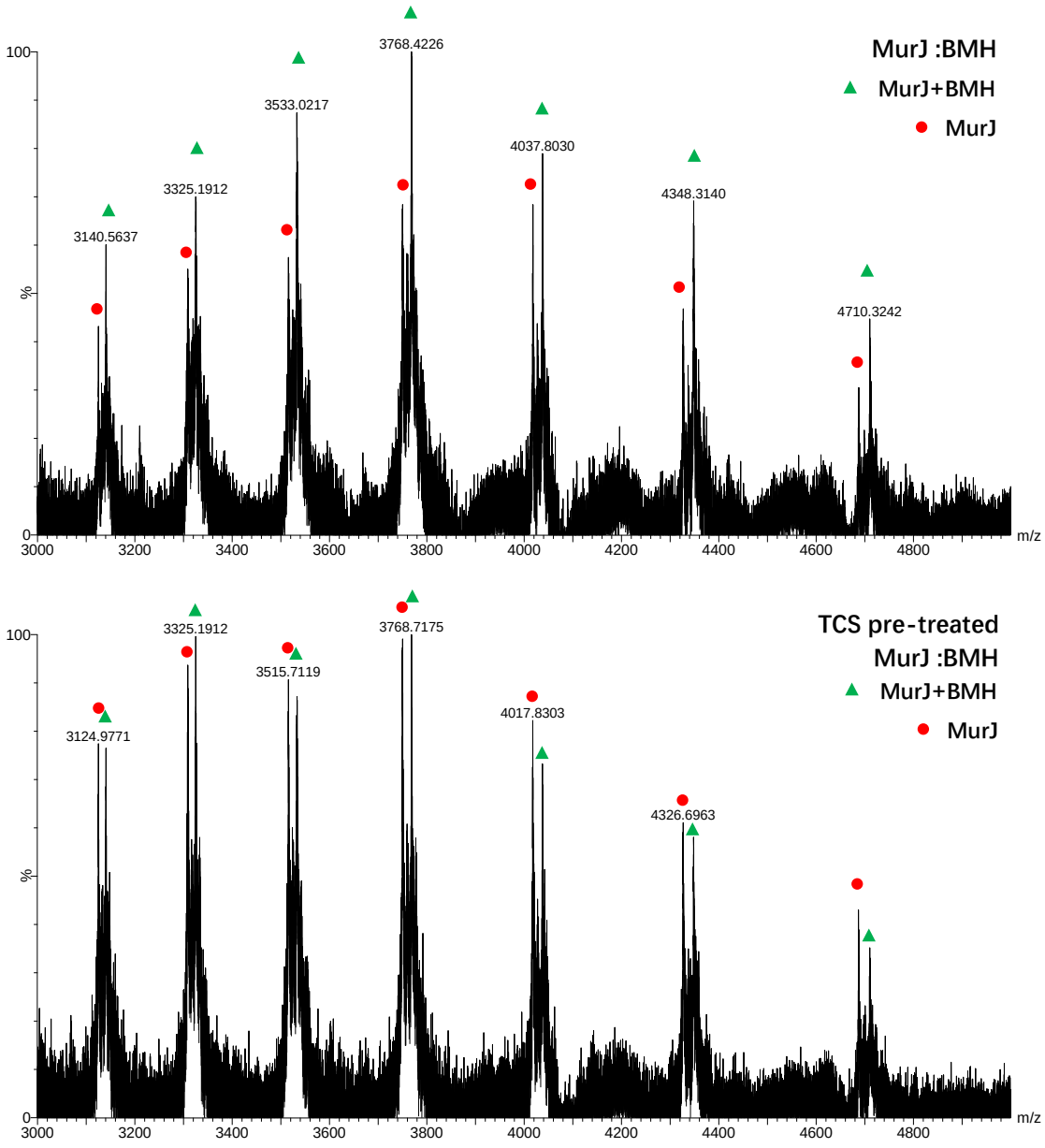


Figure 3.16 nMS spectrum of BMH cross-linked E70C_A296C with/without PMF disruption (TCS-pretreated). The m/z range is 3000 to 5000. The green triangles mark the peaks of the MurJ:BMH complexes (successful cross-linking), and the red dots are the peaks of MurJ E70C_A296C (unsuccessful cross-linking).

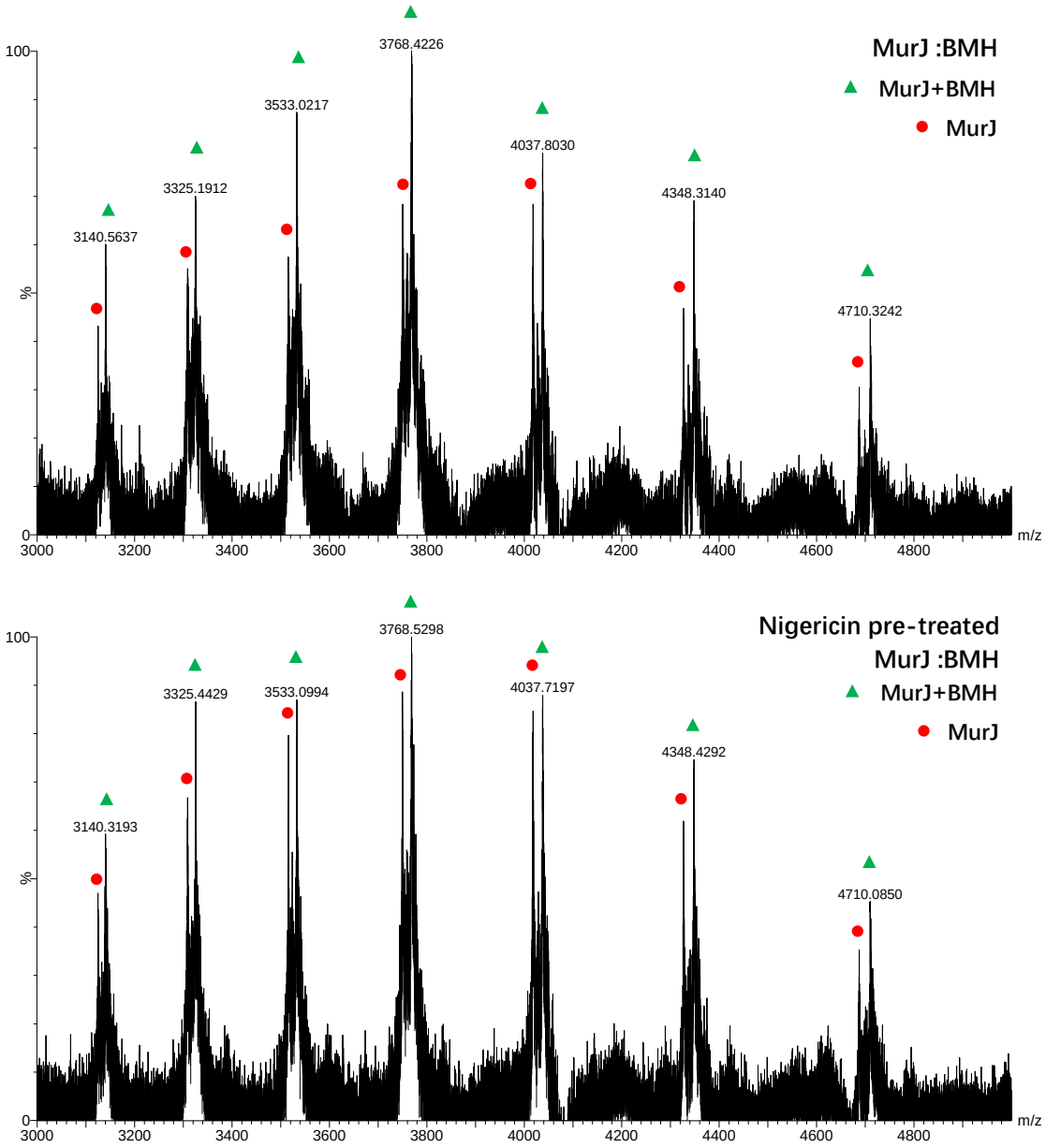


Figure 3.17 nMS spectrum of BMH cross-linked E70C_A296C with/without PMF disruption (Nigericin-pretreated). The m/z range is 3000 to 5000. The green triangles mark the peaks of the MurJ:BMH complexes (successful cross-linking), and the red dots are the peaks of MurJ E70C_A296C (unsuccessful cross-linking).

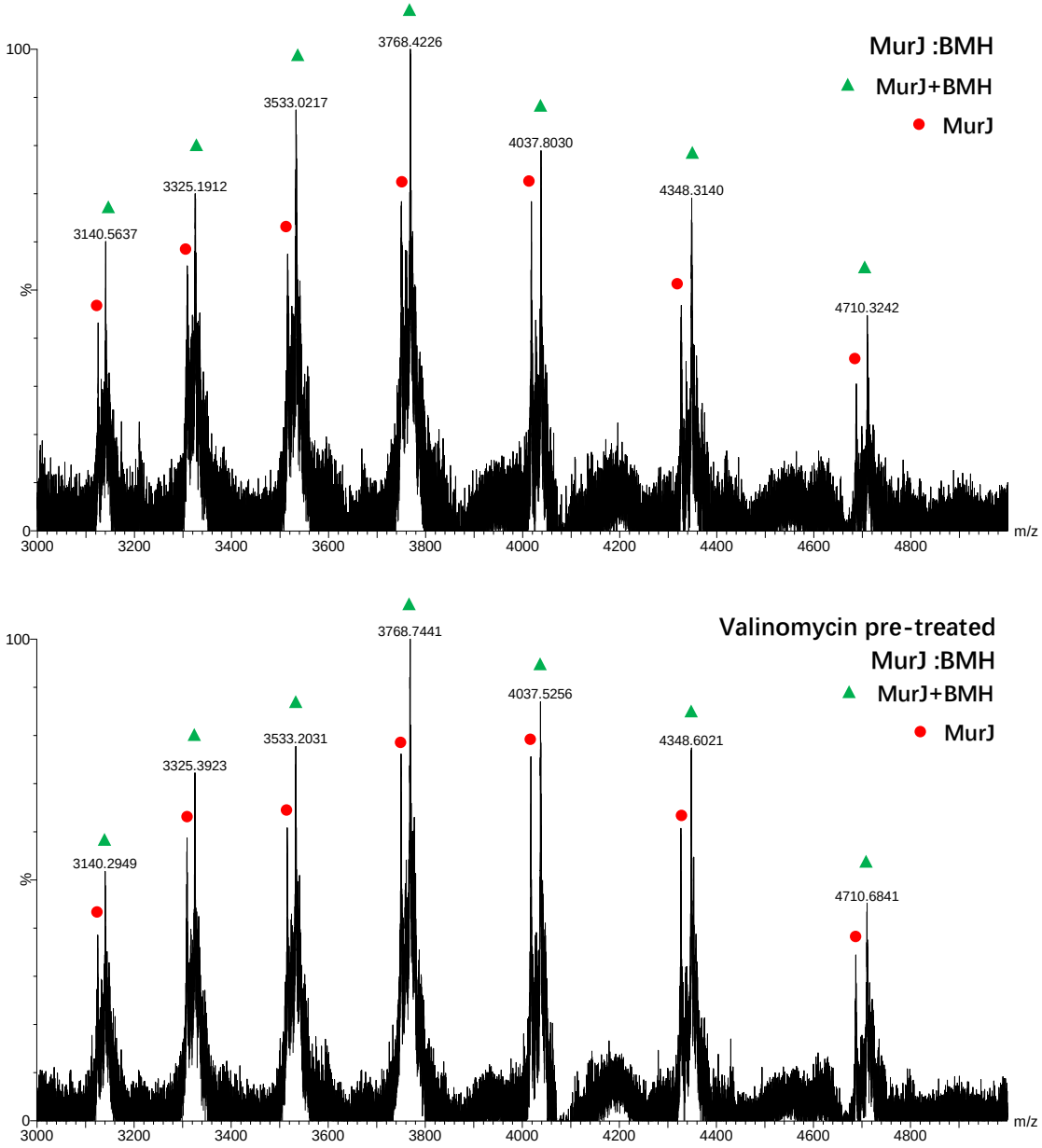


Figure 3.18 nMS spectrum of BMH cross-linked E70C_A296C with/without PMF disruption (Valinomycin-pretreated). The m/z range is 3000 to 5000. The green triangles mark the peaks of the MurJ:BMH complexes (successful cross-linking), and the red dots are the peaks of MurJ E70C_A296C (unsuccessful cross-linking).

3.4 Summary

Current structural data indicates that MurJ translocates lipid II across the membrane by alternating between the inward- and outward-open conformations, but the driving force for the alternation remains unclear. In this study, structure-guided cysteine cross-linking and protein native mass spectrometry were used to investigate the role of the proton motive force in the conformational alternation cycle of MurJ.

Based on previous studies [17], two pairs of sites on the protein that can be cross-linked with homobifunctional cysteine cross-linkers were selected and mutated into cysteine residues. Two double-cysteine MurJ models were therefore built, namely E70C_A296C and V43C_T251C. When the protein is in an outward-open conformation, the mutant E70C_A296C will allow cross-linking with BMH and the reacted MurJ will be structurally fixed in this conformation. On the other hand, the mutant V43C_T251C allows fixation to an inward-open configuration. Species that are successfully cross-linked can be detected with peak shifts in protein mass spectrum. Under normal growth circumstances, the MurJ proteins undergo conformational changes *in vivo*. When the bacteria were incubated with BMH, it is expected that a portion of the proteins would be cross-linked. To investigate the driving force of protein conformational changes, three compounds that can affect membrane proton motive force through different mechanisms were explored for their effects on protein conformational changes.

In the protein cross-linking experiments, it was found that both E70C_A296C and V43C_T251C can be cross-linked with BMH efficiently *in vivo*. The cross-linking efficiency of V43C_T251C (probes the "inward-open" state) did not change significantly after the disruption of the proton motive force. However, for E70C_A296C, when the PMF is disturbed, the relative ratio of cross-linked and uncross-linked species reversed. The addition of TCS affected the movement of MurJ and trapped it in the inward-open conformation.

MurJ protein functions as a flippase involved in lipid II transmembrane transport through conformational changes. The conformational flipping process is a cyclic mechanism, the action of TCS influencing the transition of protein conformation to an outward-open state. Therefore, PMF plays a crucial role in facilitating the flipping process, especially when the substrate is ready to flip out of the membrane after binding to the protein. TCS impacts the PMF of the membrane by modulating hydrogen ion transmembrane transport. Nigericin is associated with H^+ and K^+ transport, while valinomycin affects the K^+ transport. These findings are consistent with previous studies that changes in pH impact the binding of MurJ to lipid II [22]. Therefore, it is speculated that the key ion for flipping is H^+ rather than Na^+.

With these experiments, a sensitive method to probe MurJ conformational changes *in vivo* has been demonstrated. It can be further developed by acquiring data from additional double mutants. This mass spectrometry-based method has great potential to be exploited to identify MurJ inhibitors and study their mechanisms of actions. For instance, it is reported and speculated that the Lys^M peptide inhibits MurJ by binding to the protein and trapping it in one configuration [110]. But it is not yet clear which

configuration it was trapped or at which stage of the alternating flipping process it was. It is hopeful that this newly developed strategy can be applied to answer these questions, for Lys^{M} and other possible inhibitors. The structural knowledge will reveal mechanistic details of MurJ that are indispensable for inhibitor optimization.

Chapter 4 Investigation of MurJ dynamics in solution at different conformations using hydrogen/deuterium exchange mass spectrometry (HDX-MS)

4.1 Introduction

Solvent accessibility and protein dynamics are two key factors that affect the rate and extent of hydrogen-deuterium exchange in HDX-MS [52, 111]. Solvent accessibility refers to the extent to which a region of a protein is exposed to the solvent (H_2O or D_2O in HDX-MS experiments) [112]. Regions that are more exposed to the solvent are more likely to undergo hydrogen-deuterium exchange, which can be determined by comparing the differences in deuterium uptake between different peptides over time [113]. Protein dynamics refers to the motion and conformational changes that occur within proteins over time, which include local fluctuations, domain motions, and larger-scale conformational changes [51]. Dynamic regions of proteins often have higher exchange rates because they may be undergoing conformational changes or fluctuations that lead to temporarily exposing amide hydrogens to the solvent [114-116].

According to the reported crystal data, MurJ can adopt various configurations that will have different solvent accessibility in solution which will complicate the data analysis [12-15]. In this chapter, the protein MurJ fixed in inward-facing or outward-facing conformations obtained in **Chapter 3** were used. The solvent accessibility of peptides in these protein samples is simplified during the initial stages of hydrogen-deuterium exchange, which is conducive to the analysis of regions that undergo movement and conformational changes in the protein. Here, HDX-MS will be employed to investigate the structural characteristics of MurJ in solution.

4.1.1 Bottom-up HDX-MS

Deuterium uptake can be measured throughout the entire protein (global HDX), but is more frequently performed using a "bottom-up" local HDX methodology [43]. This approach involves quenching of the exchange reaction before conducting liquid chromatography-mass spectrometry (LC/MS) analysis, followed by proteolytic digestion [51]. Global HDX enables high-throughput screening, but only provides an overview of protein conformational dynamics. Local HDX analysis method allow for a more detailed examination of local exchange data, thereby providing a level of "spatial resolution" to the findings [43, 117]. Amide hydrogens located in flexible, solvent-exposed regions exhibit a higher exchange rate compared to those situated in rigid, buried regions or regions involved in hydrogen bonding. This hydrogen-deuterium exchange rate depends on the local environment of the amide hydrogens that provides insights into the structural and dynamic properties of proteins.

The rate of hydrogen exchange is significantly influenced by the pH of the surrounding environment and temperature (**Figure 4.1**) [55, 118]. The slowest exchange rates fall within the pH range of 2.5 to 3, which is approximately 105 times slower than the rates observed at pH 7 [55, 119, 120]. A reduction in temperature from 25°C to 0°C leads to an approximate 14-fold decrease in the hydrogen exchange rate [121-123]. Both temperature and pH establish the "quenching conditions" that allow accurate deuterium uptake curves to be measured at different time points with minimized back exchange [124].

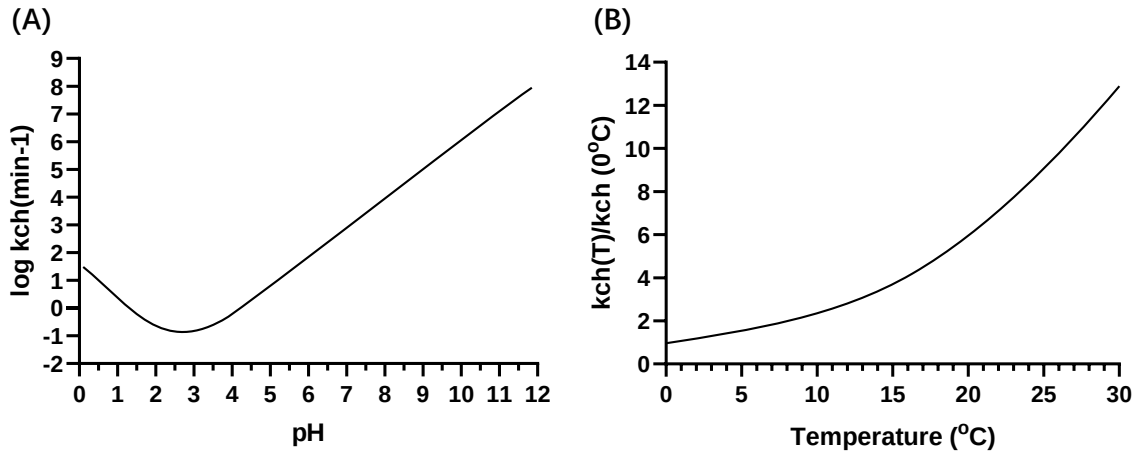


Figure 4.1 The effect of pH and temperature on HDX rates [55]. (A) shows the relationship between the HDX rate of amide hydrogen and pH. (B) shows the relationship between the HDX rate of amide hydrogen and temperature in base-catalyzed exchange.

The fundamental procedures involved in a bottom-up HDX-MS following the labeling process include: (1) Quenching, (2) Digestion, (3) Separation, and (4) Mass analysis [51]. When protein is introduced into a D_2O solution, the amide hydrogens exchange with deuterium is initiated. To obtain the deuterium exchange rate of the protein at specific time points, the exchange reaction is quenched by lowering the pH and temperature (usually pH 2.5 and 0 °C). Subsequently, the quenched protein is digested into smaller peptide fragments using an acid-stable protease like pepsin. The resulting peptides are separated using liquid chromatography (LC) and then analyzed by mass spectrometry to determine the extent of deuterium incorporation. Mass spectrometry is used to map deuterium incorporation in specific regions of a protein [41]. Deuterium incorporation is calculated for each peptide by comparing the mass spectra of a deuterium-labeled sample to an unlabeled control sample. The resulting data is used to plot deuterium uptake curves for each peptide as a function of time [124]. These curves provide information on exchange rates, enabling inferences about structural dynamics [43]. Mapping deuterium incorporation data onto the protein's three-dimensional framework aids in the recognition of protein segments that exhibit flexibility, exposure to solvent, or contribute to stabilizing secondary structures [125-128]. The conformational change, interaction sites, and regions of dynamic behavior within the protein structure can be studied by comparing deuterium incorporation patterns under various conditions, such as ligand presence [129].

4.2 Methods

4.2.1 Materials

Detergent n-dodecyl-β-D-maltoside (DDM) was purchased from Sigma-Aldrich. Amicon centrifugal filters were purchased from Merck Millipore. Organic solvents used in LC-MS were HPLC grades purchased from Merck Millipore. Formic acid was purchased from Sigma-Aldrich. D_2O was purchased from Cambridge Isotope Laboratory. The C18 analytical column and VanGuard Pre-column used in LC-MS were purchased from Waters Corporation. Nepenthesin-2/pepsin column was purchased from company AffiPro.

4.2.2 Hydrogen/Deuterium eXchange (HDX) reaction

Cross-linked protein samples V43C_T251C (inward-open) and E70C_A296C (outward-open) were prepared as described in **Chapter 3**. D_2O buffer (50 mM Tris-HCl, pD 8.0, 150 mM NaCl, 1 mM DDM) was prepared based on the protein purification buffer used in **Chapter 2**. The HDX reaction was initiated by diluting the purified protein in D_2O buffer at a ratio of 1:10 and was carried out at room temperature. Different labeling time points (1 min, 10 min, 60 min) were set according to the experimental purpose. The exchange reaction was terminated by adding an equal volume of quenching buffer (50 mM Tris-HCl, pH 2.5, 150 mM NaCl, 1 mM DDM) precooled to 0 °C. The sample was then injected into the LC-MS for analysis.

4.2.3 Local HDX-MS analysis

The quenched samples were manually injected into the Waters nanoACQUITY UPLC system for enzymatic digestion and peptide separation. Enzymatic digestion was performed using an online Nepenthesin-2/pepsin Column (2.1 mm × 20 mm) at 25 ℃. The digested peptides were desalted and separated at 0 °C using a Waters ACQUITY UPLC BEH C18 VanGuard pre-column (1.7 μm, 2.1 mm × 5 mm) and an analytical column (1.7 μm, 1 mm × 100 mm) at a flow rate of 0.06 mL/min. The mobile phase used were H_2O (mobile phase A) and ACN (mobile phase B). Both mobile phase A and B contained 1% formic acid. The LC elution gradient method is shown in **Table 4.1**. The eluted components were detected by a Waters Synapt G2-Si mass spectrometer with the capillary voltage set to 3.0 kV, the source temperature to 150 °C, the desolvation temperature to 350 °C, and the mass spectra with an m/z range of 100 to 2000 were collected.

Table 4.1 LC mobile phase gradient used in local HDX-MS analysis.

Time (min)	Mobile phase A (%)	Moblie phase B (%)
0.00	92.0	8.0
10.00	40.0	60.0
11.50	5.0	95.0
12.50	5.0	95.0
13.00	92.0	8.0

4.2.4 Data analysis

The mass data from unlabeled control sample were first imported to the software ProteinLynx Global Server (PLGS) version 3.0.2 (Waters) to generate a peptide list using the target protein sequence as a template. The list was then imported into another software DynamX 3.0 (Waters) to map a peptide coverage of the target protein for subsequent hydrogen-deuterium exchange data analysis. Filters were applied to exclude unreliable and noisy weak measurements, peptides which do not meet any of these thresholds will be removed. The protein sequence coverage using the screened peptides was automatically calculated by the software. The filtering parameters in DynamX were set as listed in **Table 4.2**.

Table 4.2 Parameter settings for peptide coverage map generation in DynamX.

Filters	Thresholds
minimum intensity	3000
minimum product per amino acid	0.33
minimum score	6
maximum MH+ error	5 ppm

4.3 Results and discussion

4.3.1 Peptide coverage map

The first part of the data analysis of a bottom-up HDX-MS experiment is to identify the numerous peptides generated after the enzymatic digestion step [130]. The MS data of these peptides will initially be processed and searched by the software PLGS (Waters) using the amino acid sequence of the target protein as a template to identify the corresponding peptide sequences. The obtained peptide list was then imported into the software DynamX (Waters), peptides with poor mass spectrometry signals were excluded to generate a peptide coverage map of the target protein. This map can also serve as supporting evidence for protein identification.

The peptide maps were generated using reference samples that had not expose to deuterium. The protein sample was diluted 1:10 with dilution buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM DDM), followed by addition of pre-chilled quenching buffer and mixing. The resulting peptide coverage maps were presented in **Figure 4.2** and **Figure 4.3**. The coverage rates of BMH-cross-linked MurJ mutants V43C_T251C and E70C_A296C are 65.5% and 57.5%, respectively. Both peptide maps cover several important helices, including transmembrane helices (TM) 1, 2, 7, and 8 that constitute the hydrophilic central cavity, and TM 9 located near the lipid tail binding site [14].

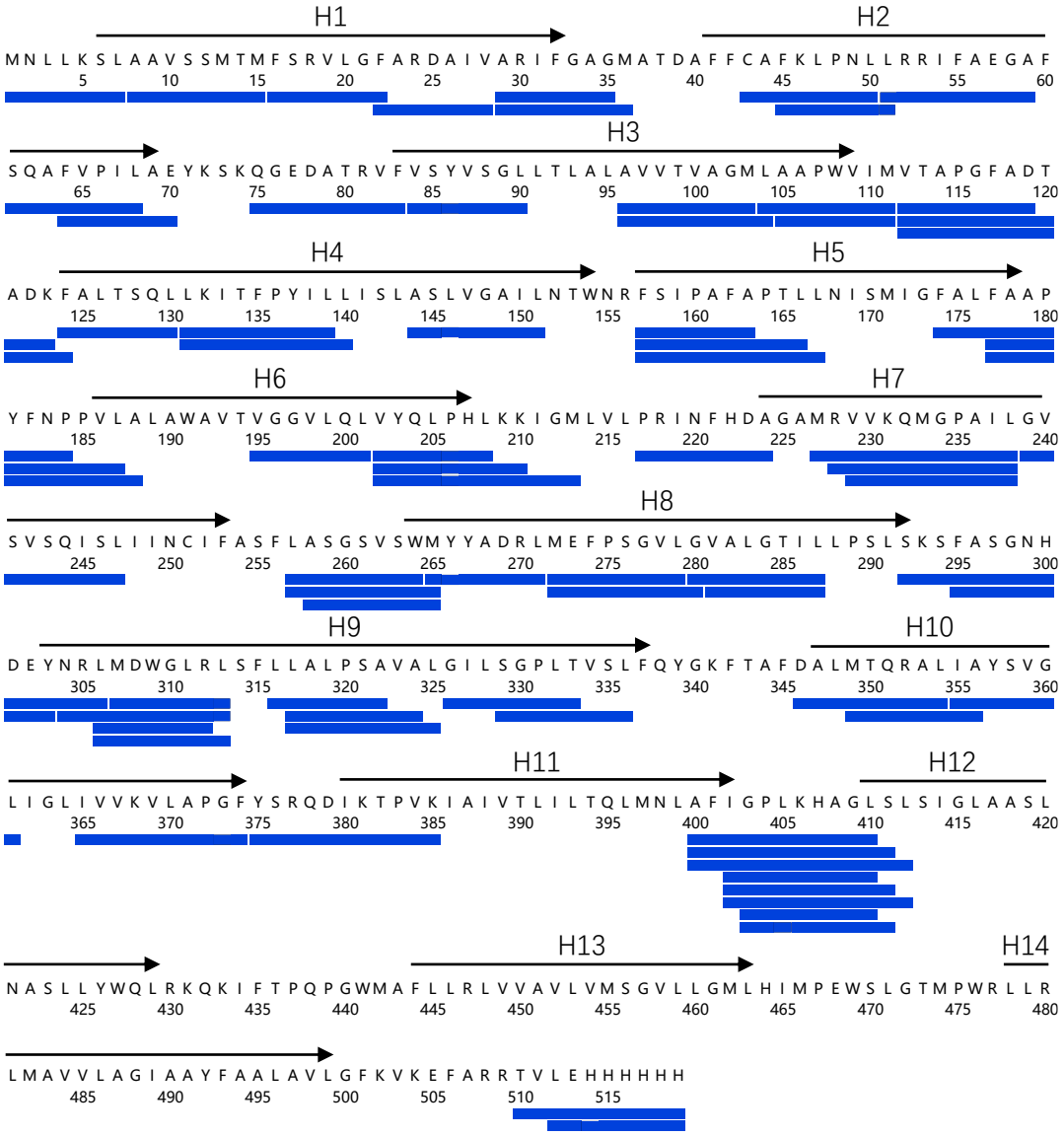


Figure 4.2 Peptide coverage map of BMH cross-linked MurJ V43C_T251C. The coverage rate is 65.5%. The black arrows indicate the helices structures in the protein. The blue bars represent the peptides generated by enzymatic digestion. H stands for transmembrane helices.

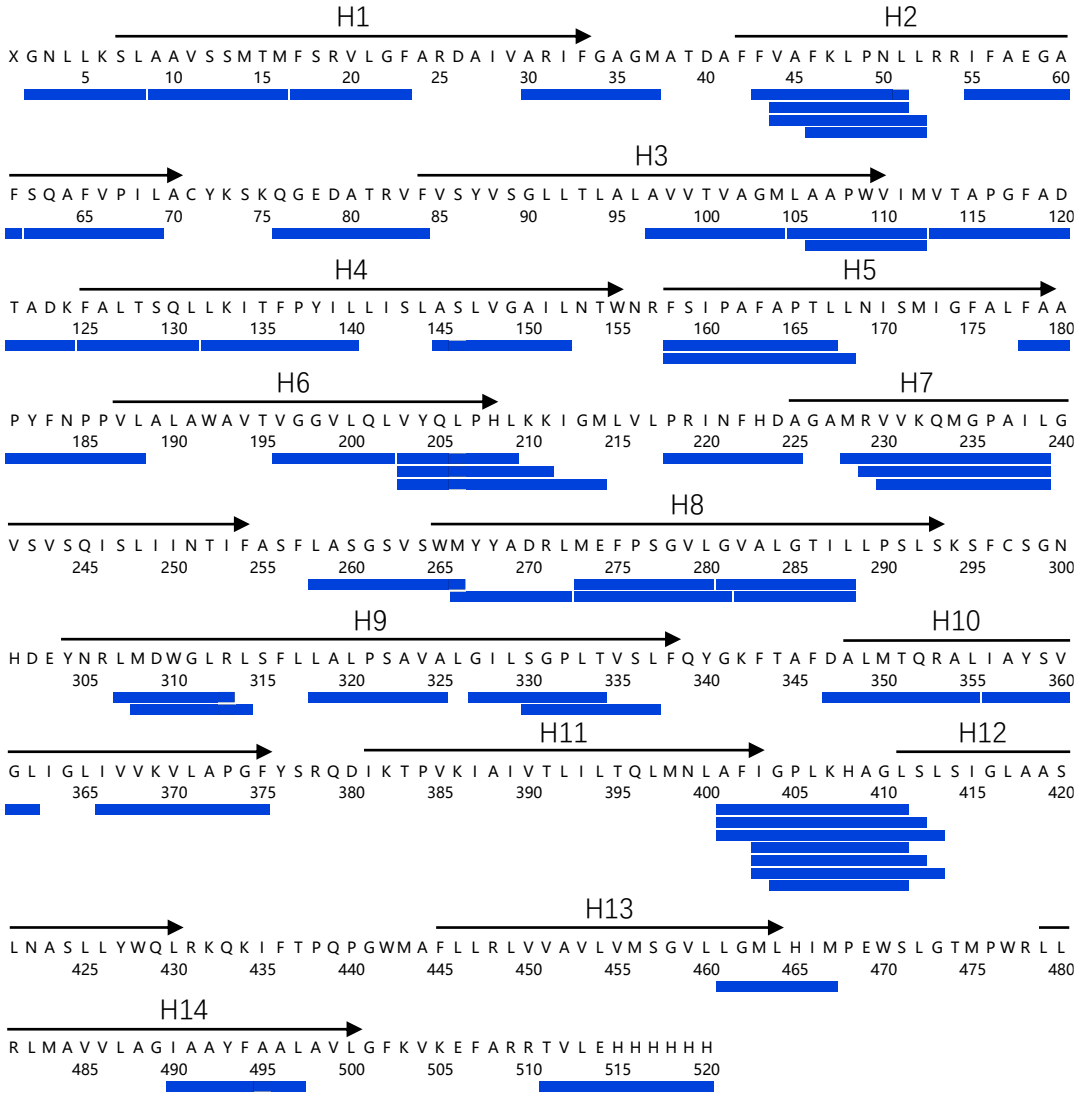


Figure 4.3 Peptide coverage map of BMH cross-linked MurJ E70C_A296C. The coverage rate is 57.5%. The black arrows indicate the helices structures in the protein. The blue bars represent the peptides generated by enzymatic digestion. H stands for transmembrane helices.

4.3.2 Local HDX-MS results of cross-linked MurJ

Site-directed mutagenesis and cross-linking techniques were used to fix MurJ into one of two configurations, inward-open (BMH-cross-linked V43C_T251) and outward-open (BMH-cross-linked E70C_A296C). The dynamic changes of the protein under different configurations were studied by hydrogen-deuterium exchange (**Figure 4.4**). HDX experiments were conducted on purified MurJ at various time points (1 min, 10 min, 60 min), and the experimental data were analyzed by DynamX software to obtain deuterium uptake difference maps. To ensure the correct allocation of mass peaks to peptides, the spectral plot assigned by software was checked manually. The results of the difference map plots are the deuterium uptake values of the BMH-cross-linked E70C_A296C (outward-open) minus the deuterium uptake values of BMH-cross-linked V43C_T251C (inward-open). The scale represents the difference in deuterium uptake in Daltons (Da), with values ranging from -1.0 to + 1.0. In the graphical representation, regions where the exchange difference is less than 0.5 were represented by blue, while regions where the difference exceeds 0.5 were represented by green bands. The number of peptides presented in this figure is less than that depicted in the peptide coverage maps shown above. Certain peptides were excluded by the peptide filters, because the deuterium exchange reaction could result in a significant decrease in the signal intensity of the peak due to a shift in the isotopic peak. The results were partially mapped onto the corresponding positions of the protein structure model according to the peptide sequence number. The relative uptake curves and detailed data of these peptides were given in the **Figure S14-S15** and **Table S1-S2** in **Appendix**.

The regions (TM2, TM5, TM8) with less change (< 0.5) in deuterium uptake in MurJ protein under different conformations is mainly located in the hydrophilic central cavity of the protein. This indicates that the solvent accessibility and dynamic changes of these peptide chains are similar in both the inward and outward structural configurations. In reported crystal structure data, this position is speculated to be the binding site for the head group of substrate lipid II [13-15]. The regions exhibiting significant deuterium uptake difference (> 0.5) are predominantly located within transmembrane segments TM1, TM2, TM7, and TM9. Notably, TM7 is considered to act as a lever that promotes overall conformational changes of MurJ between different configurations [14]. Furthermore, the TM9 region located near the binding site for the elongated lipid tail of lipid II, and variations in deuterium uptake in this area may be closely related to the inherent flexibility of the long lipid chain [13, 14].

It can be found that the peptide (45-FKLPNLL-51) located at the N-terminal region of TM2 exhibit less change, whereas the peptide (61-SQAFVPIL-68) located at C-terminal region demonstrate significant difference between the two configurations [14, 96]. Crystal structure data suggest that a flexible G/A-E-G-A motif located in the middle of TM2 induces the bending of the helix.

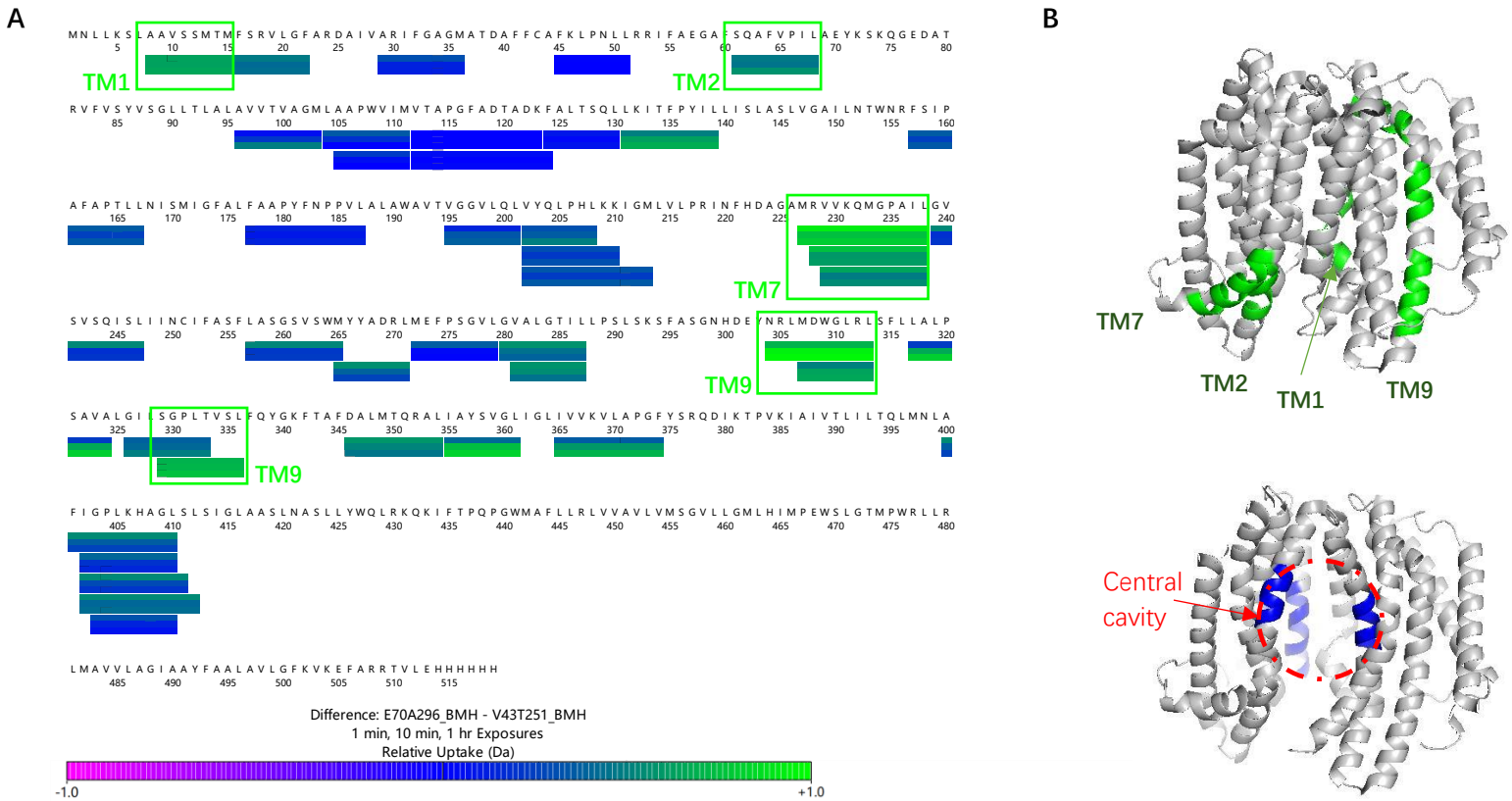


Figure 4.4 Hydrogen deuterium exchange difference between inward-open (V43C_T251C) and outward-open (E70C_A296C) conformation. The mutated proteins were cross-linked with a cross-linker (BMH) and fixed in one conformation. (A) Relative uptake difference map of protein V43C_T251C. Blue color shows less or no exchange difference. The green color shows a higher uptake difference. Each band consists of data from three different time points, from top to bottom, 1 minute, 10 minutes, and 60 minutes. (B) Crystal structure of MurJ in inward-open conformation. PDB code: 6CC4.

4.4 Summary

In this chapter, HDX-MS was used to investigate the dynamic behavior of the MurJ protein in two different conformations (inward-open and outward-open) in solution. The BMH cross-linked MurJ mutants V43C_T251C (inward-open) and E70C_A296C (outward-open) were digested by online nepenthesin-2/pepsin column to obtain peptide coverage maps, with coverage rates of 65.5% and 57.5% respectively. The resulting peptide maps cover the central cavity and lipid tail binding area where MurJ binds to the substrate lipid II.

By comparing the differences in the deuterium uptake results of MurJ in two different configurations at different time points, it was observed that certain regions exhibited little uptake difference, while some areas showed significant differences. The peptides with little deuterium uptake difference were mainly located in the central cavity, where protein MurJ binds to the head group of substrate lipid II, indicating that the protein structure at this position is relatively rigid [13, 14, 131]. The peptides with large deuterium uptake differences were distributed in TM1, 2, 7 and 9, suggesting that these areas of MurJ may have experienced some dynamic alterations in solution. Among them, TM7 is identified as a crucial site for conformational change, whereas TM9 located near the binding region of the long lipid tail of the substrate lipid II.

The uptake difference plots were generated by subtracting the deuterium uptake of BMH cross-linked V43C_T251C from the E70C_A296C, the resulting values predominantly fell within the range of 0 to + 1 Da. MurJ in the inward-open conformation predominantly has a greater deuterium uptake compared to the outward-open conformation. The current hypothesis is that the MurJ protein undergoes a cyclical

transition between different configurations, flipping the substrate to the outside of the membrane [14, 15]. When it is in an outward configuration, it releases the substrate out of the membrane, and then changes back to an inward configuration through an squeezed form [15]. The driving force for the recovery of the outward-facing configuration to the inward-facing configuration remains unclear. One possibility of this phenomenon may be attributed to the tendency of the protein to revert to the inward-facing state after substrate release, resulting in relatively dynamic behavior in solution. This hypothesis requires further experimental validation through additional protein samples with various configurations. Currently, a minimum of three distinct inward-facing configurations have been identified in the crystal structure [12-15]. Through the mutation and cross-linking of sites at various locations within the protein, it is theoretically possible to obtain proteins fixed in different inward-facing configurations. The dynamics of other conformations of the MurJ protein during the flipping cycle in solution can be studied by HDX-MS.

Chapter 5 Conclusion

MurJ is a bacterial membrane protein that plays a key role in the biosynthesis of bacterial cell walls. It is responsible for flipping lipid II, a precursor for cell wall peptidoglycan synthesis, from the inside to the outside of the cell membrane. It is a potential target for antibiotic action. Some important results have been achieved in the study of bacterial membrane protein MurJ, especially the three-dimensional structure information of MurJ obtained through SCAM and X-ray crystallography. These structural analyses have revealed the transmembrane helical configuration and potential lipid II binding site of MurJ, providing important clues for understanding its functional mechanism. It is hypothesized that the process by which MurJ facilitates the flipping and transport of lipid II is accomplished through the protein transitioning between inward and outward conformations. However, the flipping mechanism of these conformations is still not fully understood [13]. A comprehensive understanding of the function and mechanism of MurJ is crucial to revealing the detailed process of bacterial cell wall synthesis and is of great significance to basic biological research.

Prior research has predominantly focused on MurJ proteins derived from *E. coli* and *T. africanus*. In this study, three recombinant plasmids (N-His-MurJ, N-His-TEV-MurJ, C-His-MurJ) were designed, and MurJ proteins from *E. coli* were expressed and purified. Three *E. coli* expression systems were evaluated using the above three different plasmids, and the C43 (DE3) system was screened to show the highest stability for MurJ expression. CD spectral analysis showed that the C-His-MurJ protein could maintain a helical secondary structure when the pH value changed from 8.0 to 1.0. Unfolding experiments utilizing GuHCl demonstrated that even at concentrations reaching 6 M, approximately 35% of the helical structure was retained within the sample. Additionally, temperature-induced unfolding studies showed that the T_m of

MurJ remained around 60 °C under different pH conditions. In conclusion, the secondary structure of the MurJ protein exhibits considerable stability in response to substantial environmental fluctuations.

This study aims to elucidate the mechanisms underlying the conformational alterations of the MurJ protein during the transmembrane transport of lipid II. Two dual-cysteine MurJ models, specifically E70C_A296C and V43C_T251C, were developed through structure-guided cysteine cross-linking in conjunction with native mass spectrometry techniques. The study found that the relative ratio of cross-linked and uncross-linked species in mutant E70C_A296C was altered when the proton motive force was disrupted, whereas the cross-linking efficiency of mutant V43C_T251C remained largely unchanged. The addition of TCS influenced the dynamics movement of MurJ, trapping it in an inward-open conformation. The findings suggest that the proton motive force plays a crucial role in the conformational transition of MurJ, especially when it is ready to flip the substrates out of the membrane after substrate binding. It was speculated that protons (H^+) are the key ions that promote this flipping process. A sensitive method to detect conformational changes of MurJ *in vivo* was subsequently proposed.

The dynamic behavior of the MurJ protein in two different conformations (inward-open and outward-open) in solution was investigated using HDX-MS. Peptide coverage maps with coverage rate of 65.5% and 57.5%, respectively, were obtained by online nepenthesin-2/pepsin column digestion. By comparing the deuterium uptake results of the two configurations at different time points, it was found that the peptides with smaller differences in deuterium uptake were mainly located in the central cavity where

MurJ binds to the head group of substrate lipid II, while the peptides with larger differences (>0.5 Da) were distributed in TM1, 2, 7, and 9. Studies have shown that MurJ undergoes cyclical transitions between different configurations, flipping the substrate outside the membrane. The dynamics of the outward-facing conformation reverting to the inward-facing conformation may originate from the tendency of the protein to return to the inward-facing state upon substrate release. HDX-MS provides a valuable method to further investigate the dynamics of alternative conformations of MurJ protein during the flipping cycle in a solution environment.

Future research of MurJ conformational changes *in vivo* could utilize mutational cross-linking techniques combined with the mass spectrometry detection methods outlined in this study. More double mutant proteins with different configurations in the flipping cycle can be constructed, and the dynamic motion of these double mutants in solution *in vitro* can also be studied by HDX-MS. The mass spectrometry-based method proposed in **Chapter 3** has potential for the identification of MurJ inhibitors and the exploration of their mechanisms of action. The development of MurJ inhibitors could provide new antibiotic treatment strategies, especially in the case of multidrug-resistant bacterial strains. The methodologies and techniques employed in research of MurJ may serve as valuable reference points for the investigation of other membrane proteins. Considering the significant functions that membrane proteins fulfill in numerous biological processes and disease states, studying MurJ is crucial for the discovery and design of new antibiotics, which is imperative in tackling the urgent challenge of antibiotic resistance.

References:

1. Pazos, M. and K. Peters, *Peptidoglycan.* Bacterial cell walls and membranes, 2019: p. 127-168.
2. Scheffers, D.-J. and M.G. Pinho, *Bacterial cell wall synthesis: new insights from localization studies.* Microbiology and molecular biology reviews, 2005. **69**(4): p. 585-607.
3. Seltmann, G. and O. Holst, *The bacterial cell wall*. 2002: Springer Science & Business Media.
4. Laddomada, F., et al., *The MurG glycosyltransferase provides an oligomeric scaffold for the cytoplasmic steps of peptidoglycan biosynthesis in the human pathogen Bordetella pertussis.* Scientific reports, 2019. **9**(1): p. 4656.
5. Boes, A., et al., *Fluorescence anisotropy assays for high throughput screening of compounds binding to lipid II, PBP1b, FtsW and MurJ.* Scientific Reports, 2020. **10**(1): p. 6280.
6. Rubino, F.A., et al., *Detection of transport intermediates in the peptidoglycan flippase MurJ identifies residues essential for conformational cycling.* Journal of the American Chemical Society, 2020. **142**(12): p. 5482-5486.
7. Ruiz, N., *Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flippase in Escherichia coli.* Proceedings of the National Academy of Sciences, 2008. **105**(40): p. 15553-15557.
8. Hvorup, R.N., et al., *The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily.* European journal of biochemistry, 2003. **270**(5): p. 799-813.
9. Rehm, B.H., *Bacterial polymers: biosynthesis, modifications and applications.* Nature Reviews Microbiology, 2010. **8**(8): p. 578-592.
10. Young, K.D., *A flipping cell wall ferry.* Science, 2014. **345**(6193): p. 139-140.
11. Wu, C.H., et al., *Protein family classification and functional annotation.* Computational Biology and Chemistry, 2003. **27**(1): p. 37-47.

12. Kuk, A.C., E.H. Mashalidis, and S.-Y. Lee, *Crystal structure of the MOP flippase MurJ in an inward-facing conformation.* Nature structural & molecular biology, 2017. **24**(2): p. 171-176.

13. Zheng, S., et al., *Structure and mutagenic analysis of the lipid II flippase MurJ from Escherichia coli*. Proceedings of the National Academy of Sciences, 2018. **115**(26): p. 6709-6714.

14. Kuk, A.C., et al., *Visualizing conformation transitions of the Lipid II flippase MurJ.* Nature communications, 2019. **10**(1): p. 1736.

15. Kohga, H., et al., *Crystal structure of the lipid flippase MurJ in a "squeezed" form distinct from its inward-and outward-facing forms.* Structure, 2022. **30**(8): p. 1088-1097. e3.

16. Kuk, A.C., A. Hao, and S.-Y. Lee, *Structure and mechanism of the lipid flippase MurJ.* Annual review of biochemistry, 2022. **91**(1): p. 705-729.

17. Kumar, S., et al., *The bacterial lipid II flippase MurJ functions by an alternating-access mechanism.* Journal of Biological Chemistry, 2019. **294**(3): p. 981-990.

18. Claxton, D.P., et al., *Sodium and proton coupling in the conformational cycle of a MATE antiporter from Vibrio cholerae.* Proceedings of the National Academy of Sciences, 2018. **115**(27): p. E6182-E6190.

19. Steed, P.R., et al., *Na+–substrate coupling in the multidrug antiporter NorM probed with a spin-labeled substrate.* Biochemistry, 2013. **52**(34): p. 5790-5799.

20. Islam, S.T., et al., *Proton-dependent gating and proton uptake by Wzx support O-antigen-subunit antiport across the bacterial inner membrane.* MBio, 2013. **4**(5): p. 10.1128/mbio. 00678-13.

21. Kumar, S., et al., *The bacterial cell wall: from lipid II flipping to polymerization.* Chemical reviews, 2022. **122**(9): p. 8884-8910.

22. Bolla, J.R., et al., *Direct observation of the influence of cardiolipin and antibiotics on lipid II binding to MurJ.* Nature chemistry, 2018. **10**(3): p. 363-371.

23. Rubino, F.A., et al., *Membrane potential is required for MurJ function.* Journal of the American Chemical Society, 2018. **140**(13): p. 4481-4484.

24. Krulwich, T.A., G. Sachs, and E. Padan, *Molecular aspects of bacterial pH sensing and homeostasis.* Nature Reviews Microbiology, 2011. **9**(5): p. 330-343.

25. Campaniço, A., R. Moreira, and F. Lopes, *Drug discovery in tuberculosis. New drug targets and antimycobacterial agents.* European journal of medicinal chemistry, 2018. **150**: p. 525-545.

26. Byrne, D.P., D.M. Foulkes, and P.A. Eyers, *Pseudokinases: update on their functions and evaluation as new drug targets.* Future medicinal chemistry, 2017. **9**(2): p. 245-265.

27. Sleno, L. and A. Emili, *Proteomic methods for drug target discovery.* Current opinion in chemical biology, 2008. **12**(1): p. 46-54.

28. Kramer, R. and D. Cohen, *Functional genomics to new drug targets.* Nature Reviews Drug Discovery, 2004. **3**(11): p. 965-972.

29. Colland, F., et al., *Functional proteomics mapping of a human signaling pathway.* Genome research, 2004. **14**(7): p. 1324-1332.

30. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.

31. Griffiths, J., *A brief history of mass spectrometry.* Anal. Chem, 2008. **80**(15): p. 5678-5683.

32. Phizicky, E., et al., *Protein analysis on a proteomic scale.* Nature, 2003. **422**(6928): p. 208-215.

33. Glish, G.L. and R.W. Vachet, *The basics of mass spectrometry in the twenty-first century.* Nature reviews drug discovery, 2003. **2**(2): p. 140-150.

34. Fenn, J.B., et al., *Electrospray ionization–principles and practice.* Mass spectrometry reviews, 1990. **9**(1): p. 37-70.

35. Leney, A.C. and A.J. Heck, *Native mass spectrometry: what is in the name?* Journal of the American Society for Mass Spectrometry, 2016. **28**(1): p. 5-13.

36. van den Heuvel, R.H. and A.J. Heck, *Native protein mass spectrometry: from intact oligomers to functional machineries.* Current opinion in chemical biology, 2004. **8**(5): p. 519-526.

37. Benesch, J.L., et al., *Protein complexes in the gas phase: technology for structural genomics and proteomics.* Chemical reviews, 2007. **107**(8): p. 3544-3567.

38. Fiorentino, F., D. Rotili, and A. Mai, *Native mass spectrometry-directed drug discovery: Recent advances in investigating protein function and modulation.* Drug Discovery Today, 2023. **28**(5): p. 103548.

39. Barth, M. and C. Schmidt, *Native mass spectrometry—A valuable tool in structural biology.* Journal of Mass Spectrometry, 2020. **55**(10): p. e4578.

40. Heck, A.J., *Native mass spectrometry: a bridge between interactomics and structural biology.* Nature methods, 2008. **5**(11): p. 927-933.

41. Vinciauskaite, V. and G.R. Masson, *Fundamentals of hdx-ms.* Essays in Biochemistry, 2023. **67**(2): p. 301-314.

42. Konermann, L., J. Pan, and Y.-H. Liu, *Hydrogen exchange mass spectrometry for studying protein structure and dynamics.* Chemical Society Reviews, 2011. **40**(3): p. 1224-1234.

43. Narang, D., C. Lento, and D. J. Wilson, *HDX-MS: an analytical tool to capture protein motion in action.* Biomedicines, 2020. **8**(7): p. 224.

44. Pandey, A., et al., *Current strategies for protein production and purification enabling membrane protein structural biology.* Biochemistry and Cell Biology, 2016. **94**(6): p. 507-527.

45. Fagerberg, L., et al., *Prediction of the human membrane proteome.* Proteomics, 2010. **10**(6): p. 1141-1149.

46. Marciano, D.P., V. Dharmarajan, and P.R. Griffin, *HDX-MS guided drug discovery: small molecules and biopharmaceuticals.* Current opinion in structural biology, 2014. **28**: p. 105-111.

47. Lounnas, V., et al., *Current progress in structure-based rational drug design marks a new mindset in drug discovery.* Computational and structural biotechnology journal, 2013. **5**(6): p. e201302011.

48. Vinothkumar, K.R. and R. Henderson, *Structures of membrane proteins.* Quarterly reviews of Biophysics, 2010. **43**(1): p. 65-158.

49. Vadas, O., et al., *Using hydrogen–deuterium exchange mass spectrometry to examine protein–membrane interactions.* Methods in enzymology, 2017. **583**: p. 143-172.

50. Trabjerg, E., Z.E. Nazari, and K.D. Rand, *Conformational analysis of complex protein states by hydrogen/deuterium exchange mass spectrometry (HDX-MS): Challenges and emerging solutions.* TrAC Trends in Analytical Chemistry, 2018. **106**: p. 125-138.

51. Masson, G.R., M.L. Jenkins, and J.E. Burke, *An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery.* Expert opinion on drug discovery, 2017. **12**(10): p. 981-994.

52. Hodge, E.A., M.A. Benhaim, and K.K. Lee, *Bridging protein structure, dynamics, and function using hydrogen/deuterium-exchange mass spectrometry.* Protein Science, 2020. **29**(4): p. 843-855.

53. Marcsisin, S.R. and J.R. Engen, *Hydrogen exchange mass spectrometry: what is it and what can it tell us?* Analytical and bioanalytical chemistry, 2010. **397**: p. 967-972.

54. Bai, Y., et al., *Primary structure effects on peptide group hydrogen exchange.* Proteins: Structure, Function, and Bioinformatics, 1993. **17**(1): p. 75-86.

55. Oganesyan, I., C. Lento, and D.J. Wilson, *Contemporary hydrogen deuterium exchange mass spectrometry.* Methods, 2018. **144**: p. 27-42.

56. Weis, D.D., et al., *Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis.* Journal of the American Society for Mass Spectrometry, 2006. **17**: p. 1498-1509.

57. Blount, Z.D., *The unexhausted potential of E. coli.* elife, 2015. **4**: p. e05826.

58. Lin, S.-H. and G. Guidotti, *Purification of membrane proteins.* Methods in enzymology, 2009. **463**: p. 619-629.

59. Gordon, E., et al., *Effective high-throughput overproduction of membrane proteins in Escherichia coli.* Protein expression and purification, 2008. **62**(1): p. 1-8.

60. Gopal, G.J. and A. Kumar, *Strategies for the production of recombinant protein in Escherichia coli.* The protein journal, 2013. **32**: p. 419-425.

61. Seddon, A.M., P. Curnow, and P.J. Booth, *Membrane proteins, lipids and detergents: not just a soap opera.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 2004. **1666**(1-2): p. 105-117.

62. Kelly, S.M. and N.C. Price, *The use of circular dichroism in the investigation of protein structure and function.* Current protein and peptide science, 2000. **1**(4): p. 349-384.

63. Harris, N.J. and P.J. Booth, *Folding and stability of membrane transport proteins in vitro.* Biochimica Et Biophysica Acta (BBA)-Biomembranes, 2012. **1818**(4): p. 1055-1066.

64. Rodger, A., *Far UV Protein Circular Dichroism*, in *Encyclopedia of Biophysics*, G.C.K. Roberts, Editor. 2013, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 726-730.

65. Martin, S.R. and M.J. Schilstra, *Circular dichroism and its application to the study of biomolecules.* Methods in cell biology, 2008. **84**: p. 263-293.

66. Kelly, S.M., T.J. Jess, and N.C. Price, *How to study proteins by circular dichroism.* Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2005. **1751**(2): p. 119-139.

67. Cho, W. and R.V. Stahelin, *Membrane-protein interactions in cell signaling and membrane trafficking.* Annu. Rev. Biophys. Biomol. Struct., 2005. **34**(1): p. 119-151.

68. Miles, A.J. and B.A. Wallace, *Circular dichroism spectroscopy of membrane proteins.* Chemical society reviews, 2016. **45**(18): p. 4859-4872.

69. Bechara, C., et al., *A subset of annular lipids is linked to the flippase activity of an ABC transporter.* Biophysical Journal, 2015. **108**(2): p. 202a.

70. Keener, J.E., G. Zhang, and M.T. Marty, *Native mass spectrometry of membrane proteins.* Analytical chemistry, 2020. **93**(1): p. 583-597.

71. Reading, E., et al., *The role of the detergent micelle in preserving the structure of membrane proteins in the gas phase.* Angewandte Chemie International Edition, 2015. **54**(15): p. 4577-4581.

72. Laganowsky, A., et al., *Mass spectrometry of intact membrane protein complexes.* Nature protocols, 2013. **8**(4): p. 639-651.

73. Baneyx, F., *Recombinant protein expression in Escherichia coli.* Current opinion in biotechnology, 1999. **10**(5): p. 411-421.

74. Hattab, G., et al., *Escherichia coli as host for membrane protein structure determination: a global analysis.* Scientific reports, 2015. **5**(1): p. 12097.

75. Miroux, B. and J.E. Walker, *Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels.* Journal of molecular biology, 1996. **260**(3): p. 289-298.

76. Studier, F.W., *Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system.* Journal of molecular biology, 1991. **219**(1): p. 37-44.

77. Rath, A., et al., *Detergent binding explains anomalous SDS-PAGE migration of membrane proteins.* Proceedings of the National Academy of Sciences, 2009. **106**(6): p. 1760-1765.

78. Abràmoff, M.D., P.J. Magalhães, and S.J. Ram, *Image processing with ImageJ.* Biophotonics international, 2004. **11**(7): p. 36-42.

79. Bonissone, S., et al., *N-terminal protein processing: a comparative proteogenomic analysis.* Molecular & Cellular Proteomics, 2013. **12**(1): p. 14-28.

80. Miles, A., R.W. Janes, and B.A. Wallace, *Tools and methods for circular dichroism spectroscopy of proteins: a tutorial review.* Chemical Society Reviews, 2021. **50**(15): p. 8400-8413.

81. Butler, E.K., et al., *Structure-function analysis of MurJ reveals a solvent-exposed cavity containing residues essential for peptidoglycan biogenesis in Escherichia coli.* Journal of bacteriology, 2013. **195**(20): p. 4639-4649.

82. Booth, P.J. and P. Curnow, *Folding scene investigation: membrane proteins.* Current opinion in structural biology, 2009. **19**(1): p. 8-13.

83. Booth, P.J. and P. Curnow, *Membrane proteins shape up: understanding in vitro folding.* Current opinion in structural biology, 2006. **16**(4): p. 480-488.

84. Findlay, H.E., et al., *Unfolding free energy of a two-domain transmembrane sugar transport protein.* Proceedings of the National Academy of Sciences, 2010. **107**(43): p. 18451-18456.

85. Punekar, N. and N. Punekar, *Principles of enzyme assays.* ENZYMES: Catalysis, Kinetics and Mechanisms, 2018: p. 115-129.

86. Becker, W., et al., *Investigating protein–ligand interactions by solution nuclear magnetic resonance spectroscopy.* ChemPhysChem, 2018. **19**(8): p. 895-906.

87. Boutet, S., et al., *High-resolution protein structure determination by serial femtosecond crystallography.* Science, 2012. **337**(6092): p. 362-364.

88. Caffrey, M., D. Li, and A. Dukkipati, *Membrane protein structure determination using crystallography and lipidic mesophases: recent advances and successes.* Biochemistry, 2012. **51**(32): p. 6266-6288.

89. Hong, M., Y. Zhang, and F. Hu, *Membrane protein structure and dynamics from NMR spectroscopy.* Annual review of physical chemistry, 2012. **63**: p. 1-24.

90. Turnbull, A.P. and P. Emsley, *Studying protein–ligand interactions using x-ray crystallography.* Protein-Ligand Interactions: Methods and Applications, 2013: p. 457-477.

91. Bogdanov, M., et al., *Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAMTM): application to lipid-specific membrane protein topogenesis.* Methods, 2005. **36**(2): p. 148-171.

92. Bogdanov, M., *Mapping of membrane protein topology by Substituted Cysteine Accessibility Method (SCAM™).* Bacterial Protein Secretion Systems: Methods and Protocols, 2017: p. 105-128.

93. Kusakizako, T., et al., *Structural biology of the multidrug and toxic compound extrusion superfamily transporters.* Biochimica et biophysica Acta (BBA)-biomembranes, 2020. **1862**(12): p. 183154.

94. Butler, E.K., et al., *Charge requirements of lipid II flippase activity in Escherichia coli.* Journal of bacteriology, 2014. **196**(23): p. 4111-4119.

95. Claxton, D.P., K.L. Jagessar, and H.S. Mchaourab, *Principles of alternating access in multidrug and toxin extrusion (MATE) transporters.* Journal of molecular biology, 2021. **433**(16): p. 166959.

96. Kuk, A.C., A. Hao, and S.-Y. Lee, *Structure and mechanism of the lipid flippase MurJ.* Annual review of biochemistry, 2022. **91**: p. 705-729.

97. He, X., et al., *Structure of a cation-bound multidrug and toxic compound extrusion transporter.* Nature, 2010. **467**(7318): p. 991-994.

98. Jin, Y., A. Nair, and H.W. van Veen, *Multidrug transport protein NorM from Vibrio cholerae simultaneously couples to sodium-and proton-motive force.* Journal of Biological Chemistry, 2014. **289**(21): p. 14624-14632.

99. Mousa, J.J., et al., *ClbM is a versatile, cation-promiscuous MATE transporter found in the colibactin biosynthetic gene cluster.* Biochemical and biophysical research communications, 2017. **482**(4): p. 1233-1239.

100. Tanaka, Y., et al., *Structural basis for the drug extrusion mechanism by a MATE multidrug transporter.* Nature, 2013. **496**(7444): p. 247-251.

101. Liko, I., et al., *Mass spectrometry guided structural biology.* Current opinion in structural biology, 2016. **40**: p. 136-144.

102. Olinares, P.D.B., et al., *Native mass spectrometry-based screening for optimal sample preparation in single-particle Cryo-EM.* Structure, 2021. **29**(2): p. 186-195. e6.

103. Ziarek, J.J., D. Baptista, and G. Wagner, *Recent developments in solution nuclear magnetic resonance (NMR)-based molecular biology.* Journal of Molecular Medicine, 2018. **96**: p. 1-8.

104. Cheng, Y., *Single-particle cryo-EM at crystallographic resolution.* Cell, 2015. **161**(3): p. 450-457.

105. Walker, L.R., et al., *Native mass spectrometry of antimicrobial peptides in lipid nanodiscs elucidates complex assembly.* Analytical chemistry, 2019. **91**(14): p. 9284-9291.

106. Kim, S.J., et al., *Integrative structure and functional anatomy of a nuclear pore complex.* Nature, 2018. **555**(7697): p. 475-482.

107. Kaur, H., et al., *The antibiotic darobactin mimics a β-strand to inhibit outer membrane insertase.* Nature, 2021. **593**(7857): p. 125-129.

108. Bolla, J.R., et al., *Membrane protein–lipid interactions probed using mass spectrometry.* Annual Review of Biochemistry, 2019. **88**: p. 85-111.

109. Winkler, R., *ESIprot: a universal tool for charge state determination and molecular weight calculation of proteins from electrospray ionization mass spectrometry data.* Rapid Communications in Mass Spectrometry, 2010. **24**(3): p. 285-294.

110. Chamakura, K.R., et al., *A viral protein antibiotic inhibits lipid II flippase activity.* Nature microbiology, 2017. **2**(11): p. 1480-1484.

111. McAllister, R.G. and L. Konermann, *Challenges in the interpretation of protein H/D exchange data: a molecular dynamics simulation perspective.* Biochemistry, 2015. **54**(16): p. 2683-2692.

112. Balasubramaniam, D. and E.A. Komives, *Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins.* Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2013. **1834**(6): p. 1202-1209.

113. Yan, X., et al., *Deuterium exchange and mass spectrometry reveal the interaction differences of two synthetic modulators of RXRα LBD.* Protein Science, 2007. **16**(11): p. 2491-2501.

114. Engen, J.R., et al., *Identification and localization of slow, natural, cooperative unfolding in the hematopoietic cell kinase SH3 domain by amide hydrogen exchange and mass spectrometry.* Biochemistry, 1997. **36**(47): p. 14384-14391.

115. Giladi, M. and D. Khananshvili, *Hydrogen-deuterium exchange mass-spectrometry of secondary active transporters: from structural dynamics to molecular mechanisms.* Frontiers in Pharmacology, 2020. **11**: p. 70.

116. Zheng, J., et al., *Protein dynamics and conformational changes explored by hydrogen/deuterium exchange mass spectrometry.* Current opinion in structural biology, 2019. **58**: p. 305-313.

117. Edgeworth, M.J., et al., *Global and local conformation of human IgG antibody variants rationalizes loss of thermodynamic stability.* Angewandte Chemie International Edition, 2015. **54**(50): p. 15156-15159.

118. Morgan, C.R. and J.R. Engen, *Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry.* Current protocols in protein science, 2009. **58**(1): p. 17.6. 1-17.6. 17.

119. Dempsey, C.E., *Hydrogen exchange in peptides and proteins using NMR spectroscopy.* Progress in Nuclear Magnetic Resonance Spectroscopy, 2001. **39**(2): p. 135-170.

120. Eigen, M., *Proton transfer, acid-base catalysis, and enzymatic hydrolysis. Part I: elementary processes.* Angewandte Chemie International Edition in English, 1964. **3**(1): p. 1-19.

121. Englander, S.W., *Hydrogen exchange and mass spectrometry: A historical perspective.* Journal of the American Society for Mass Spectrometry, 2006. **17**: p. 1481-1489.

122. Englander, S.W. and N.R. Kallenbach, *Hydrogen exchange and structural dynamics of proteins and nucleic acids.* Quarterly reviews of biophysics, 1983. **16**(4): p. 521-655.

123. Englander, S.W., N.W. Downer, and H. Teitelbaum, *Hydrogen exchange.* Annual review of biochemistry, 1972. **41**(1): p. 903-924.

124. Brown, K.A. and D.J. Wilson, *Bottom-up hydrogen deuterium exchange mass spectrometry: data analysis and interpretation.* Analyst, 2017. **142**(16): p. 2874-2886.

125. Baerga-Ortiz, A., et al., *Epitope mapping of a monoclonal antibody against human thrombin by H/D-exchange mass spectrometry reveals selection of a diverse sequence in a highly conserved protein.* Protein Science, 2002. **11**(6): p. 1300-1308.

126. Puchades, C., et al., *Epitope mapping of diverse influenza Hemagglutinin drug candidates using HDX-MS.* Scientific reports, 2019. **9**(1): p. 4735.

127. Malito, E., et al., *Defining a protective epitope on factor H binding protein, a key meningococcal virulence factor and vaccine antigen.* Proceedings of the National Academy of Sciences, 2013. **110**(9): p. 3304-3309.

128. Wei, H., et al., *Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications.* Drug discovery today, 2014. **19**(1): p. 95-102.

129. Abzalimov, R.R., et al., *Protein conformations can be probed in top-down HDX MS experiments utilizing electron transfer dissociation of protein ions without hydrogen scrambling.* Journal of the American Society for Mass Spectrometry, 2009. **20**(8): p. 1514-1517.

130. Sørensen, L. and R. Salbo, *Optimized workflow for selecting peptides for HDX-MS data analyses.* Journal of The American Society for Mass Spectrometry, 2018. **29**(11): p. 2278-2281.

131. Chu, J., et al., *Human microbiome inspired antibiotics with improved β-lactam synergy against MDR Staphylococcus aureus.* ACS infectious diseases, 2018. **4**(1): p. 33-38.

Appendix

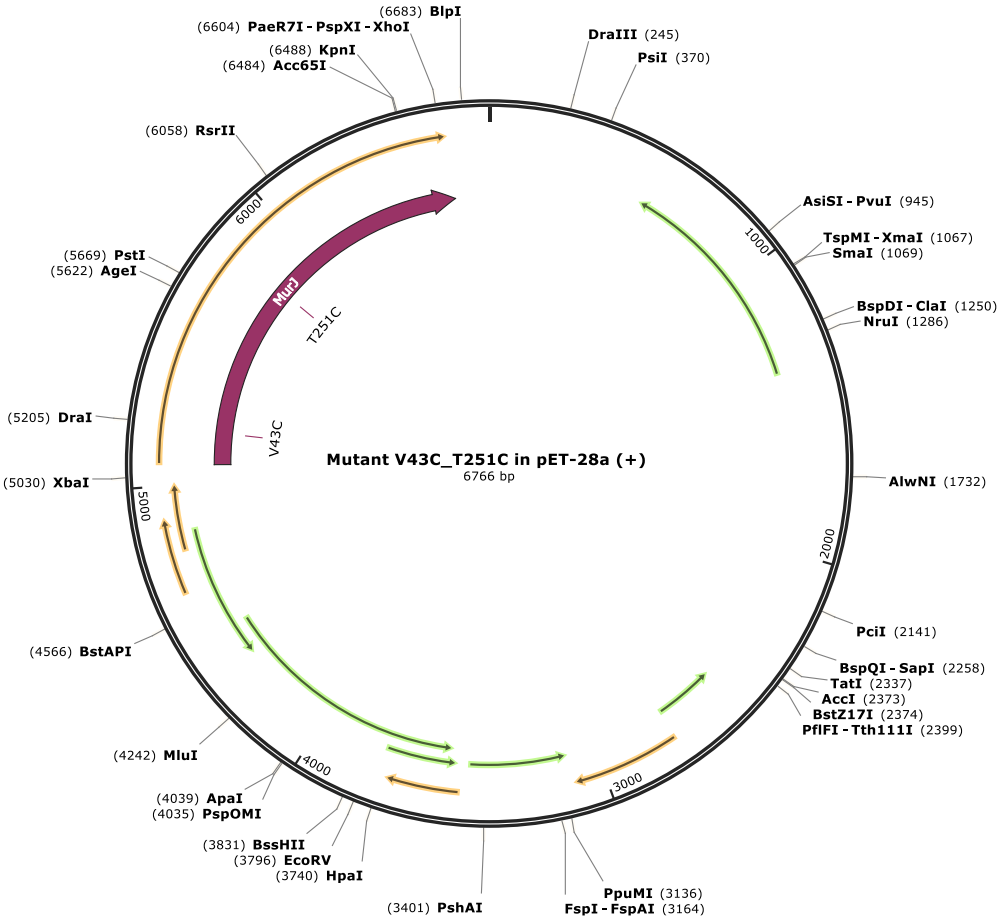


Figure S1 The reconstructed mutant V43C_T251C in pET-28a (+) plasmid.

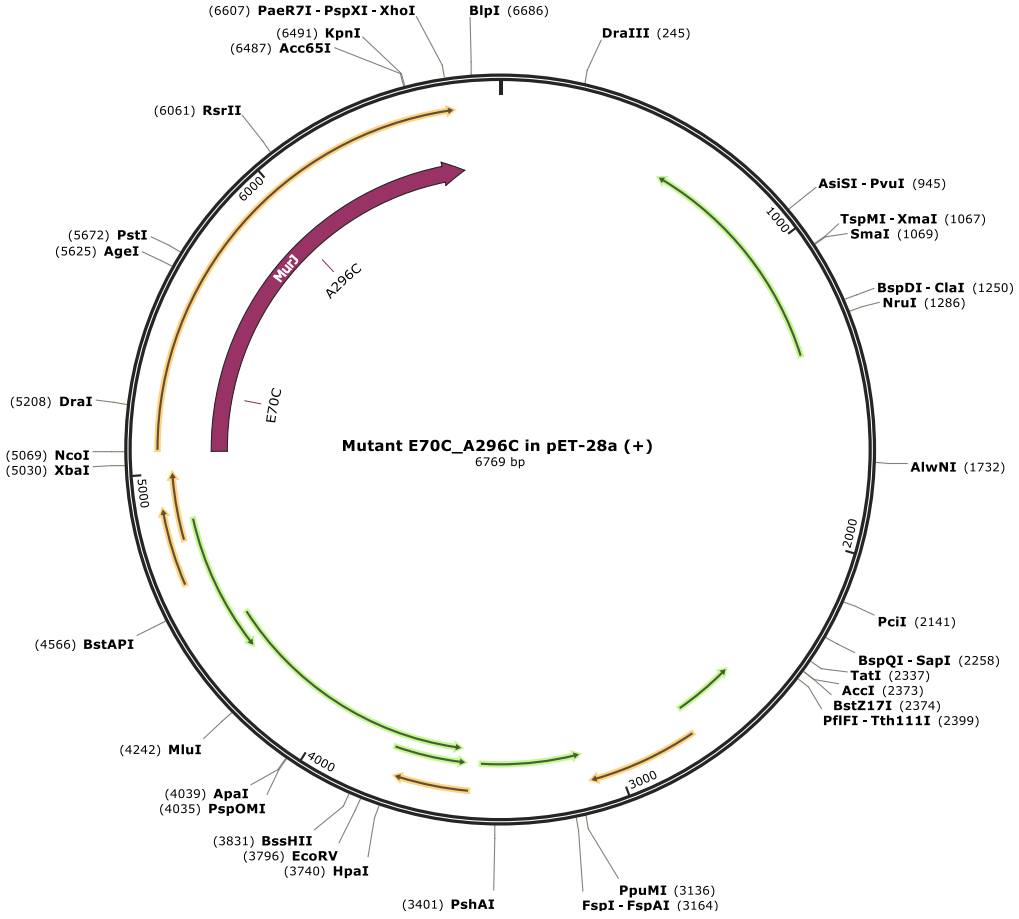


Figure S2 The reconstructed mutant E70C_A296C in pET-28a (+) plasmid.

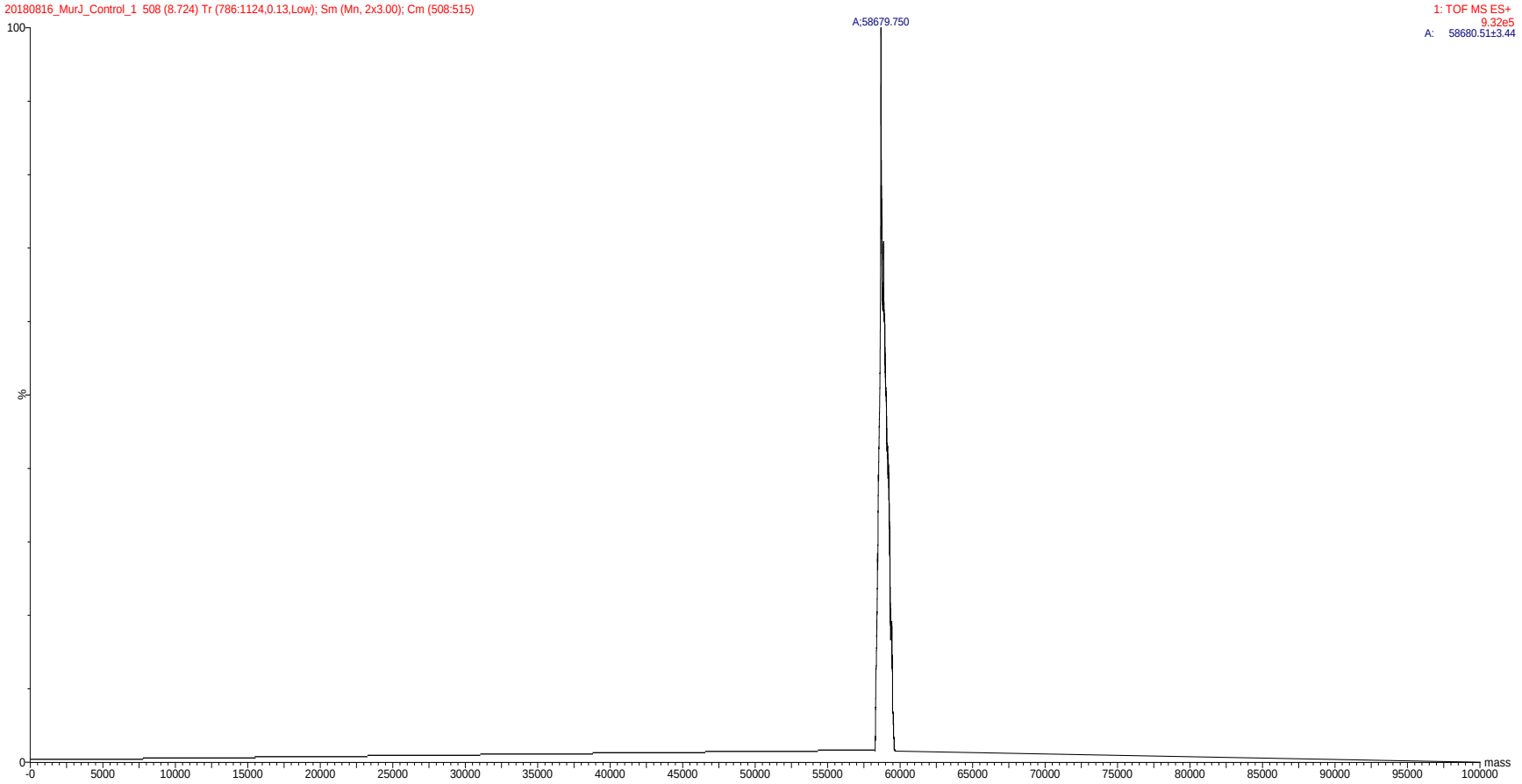


Figure S3 Deconvoluted mass spectrum of N-His-MurJ.

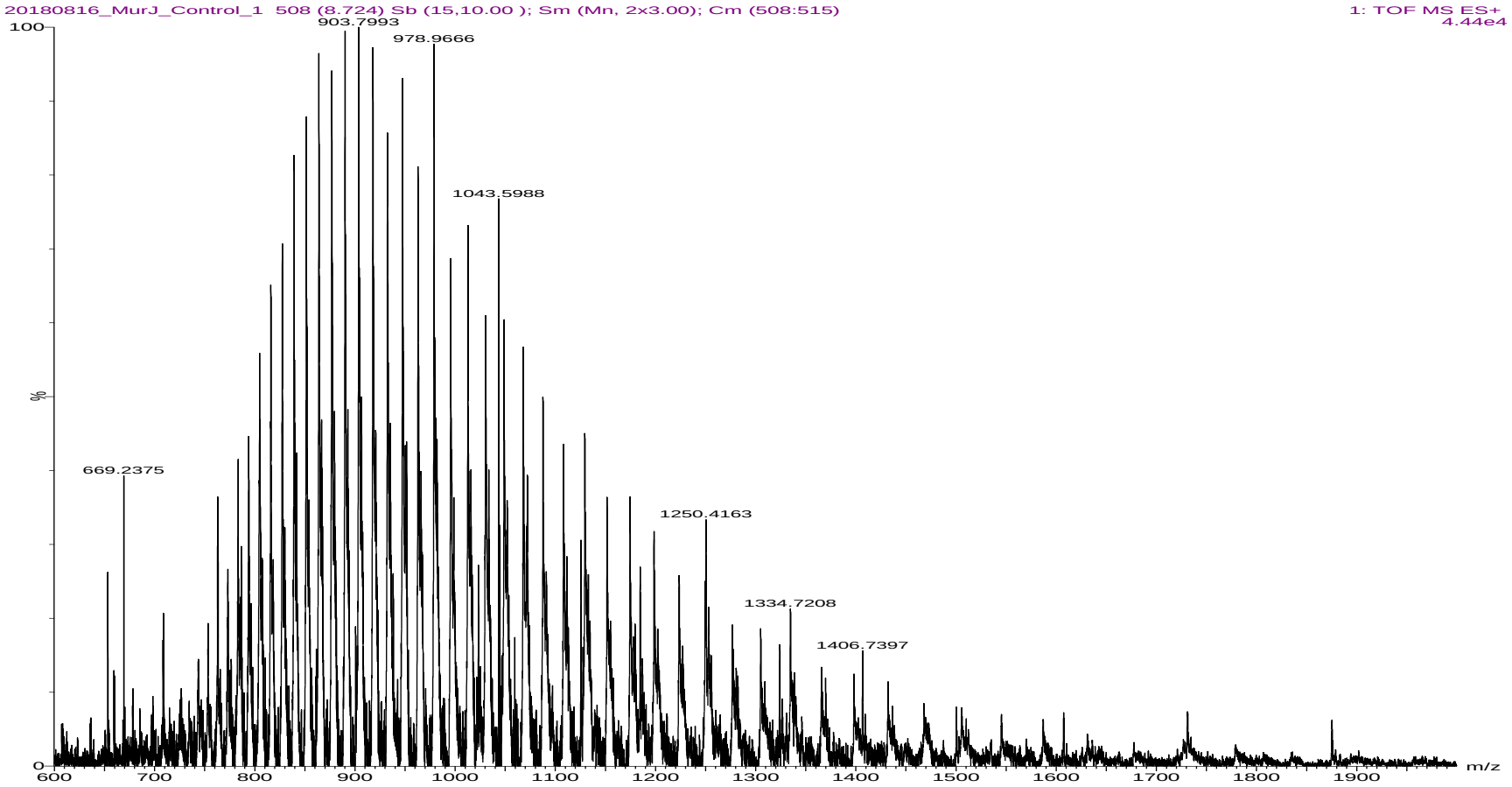


Figure S4 Mass spectrum of N-His-MurJ.

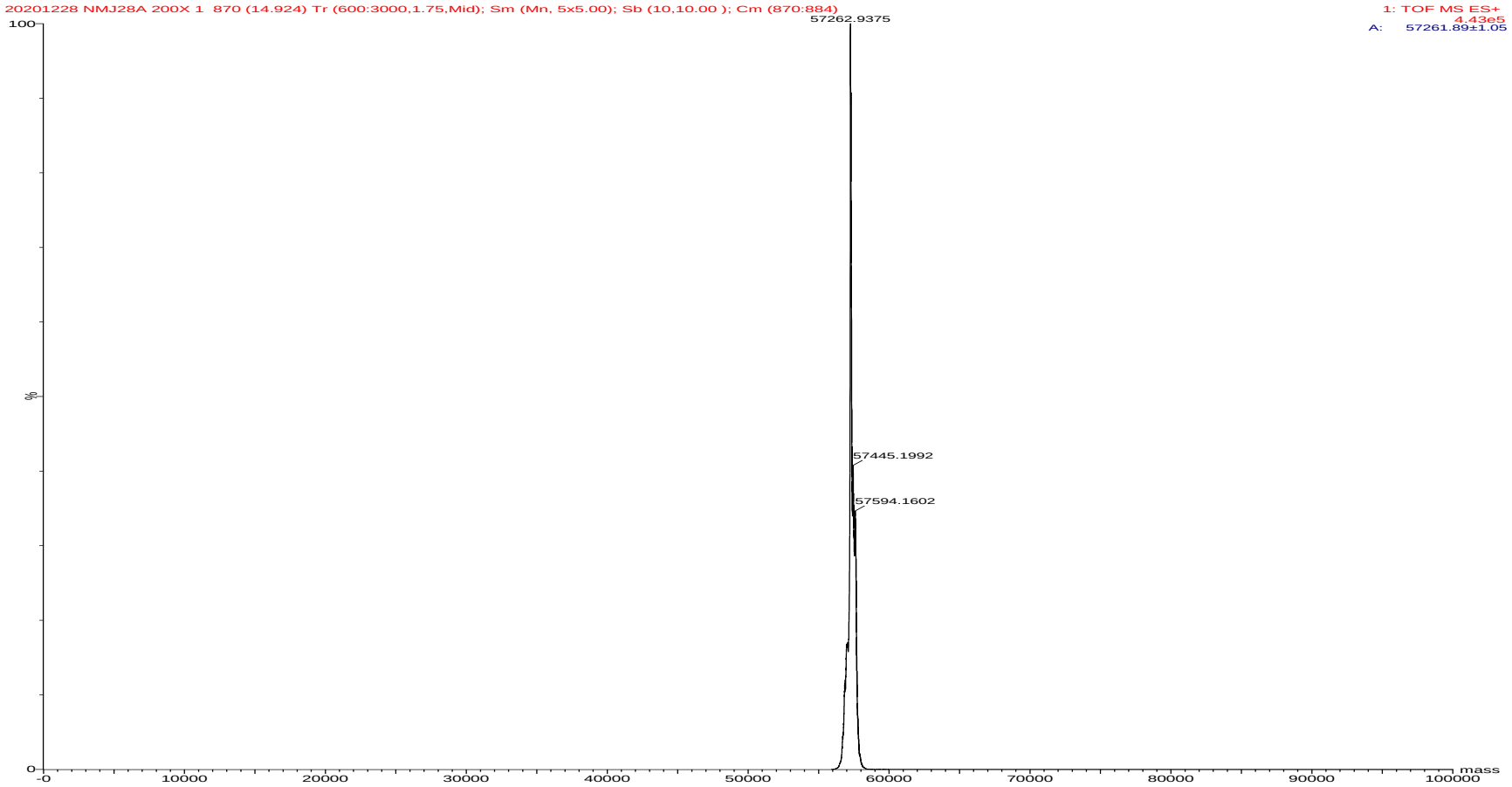


Figure S5 Deconvoluted mass spectrum of N-His-TEV-MurJ.

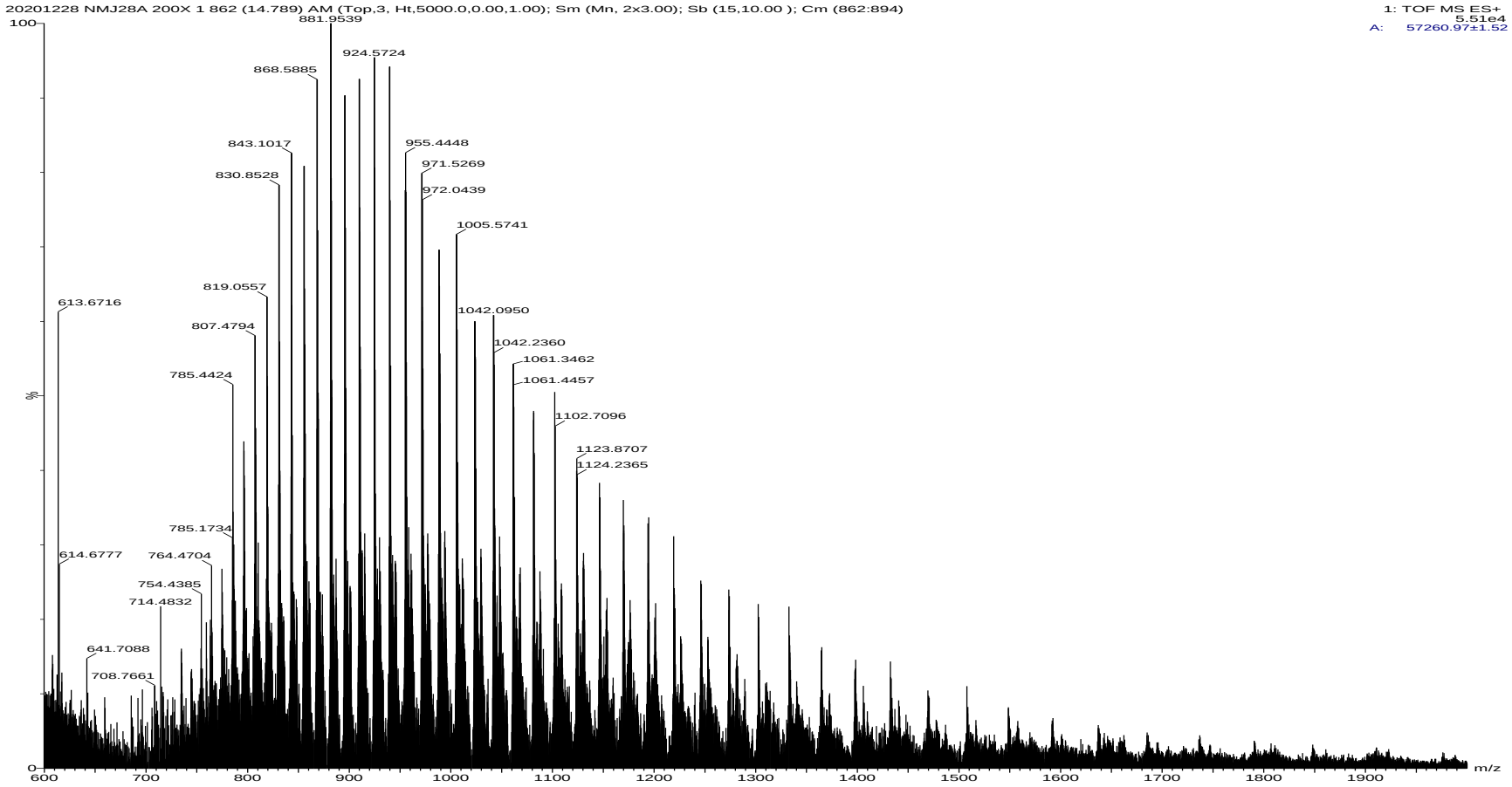


Figure S6 Mass spectrum of N-His-TEV-MurJ.

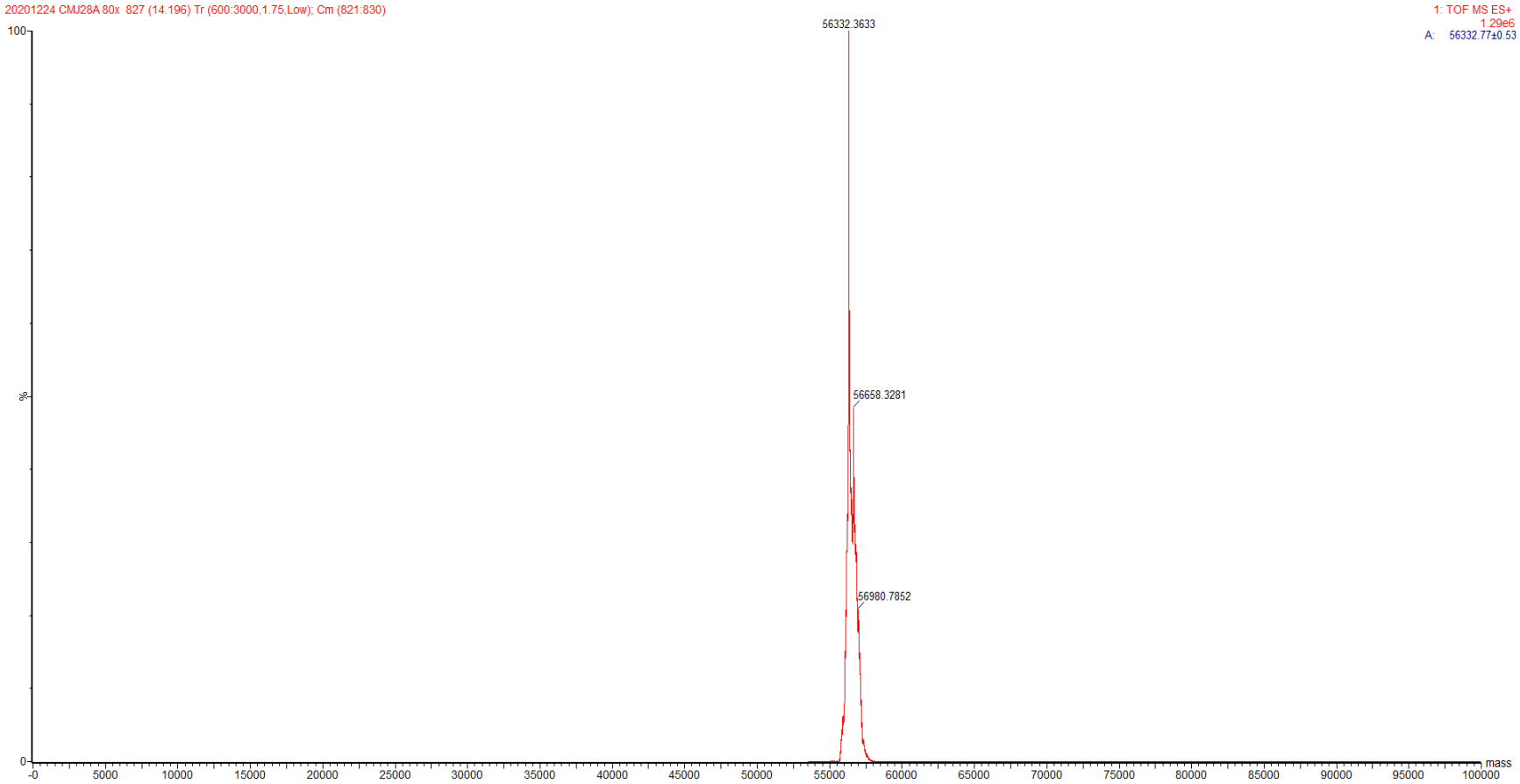


Figure S7 Deconvoluted mass spectrum of C-His-MurJ.

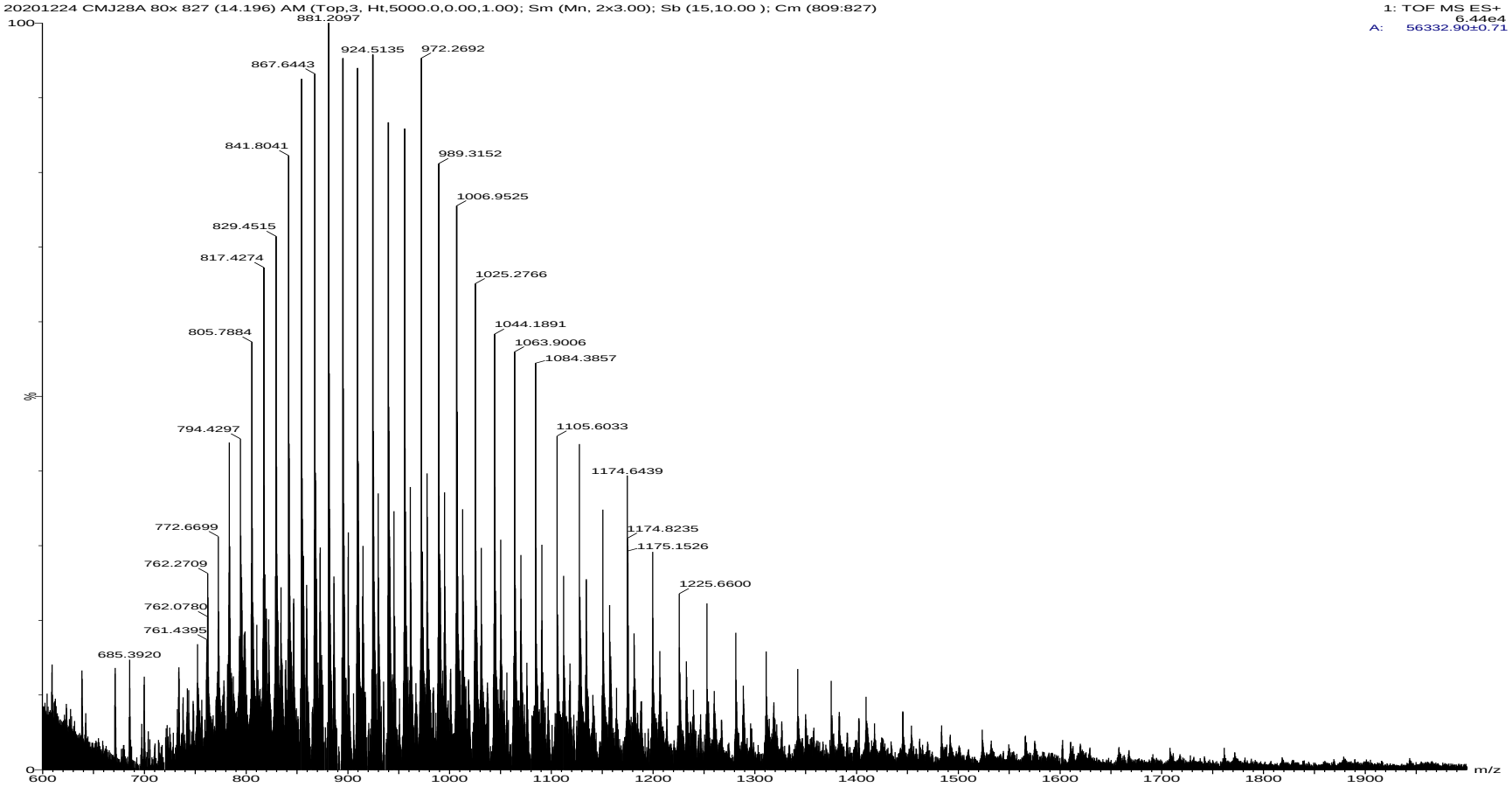


Figure S8 Mass spectrum of C-His-MurJ.

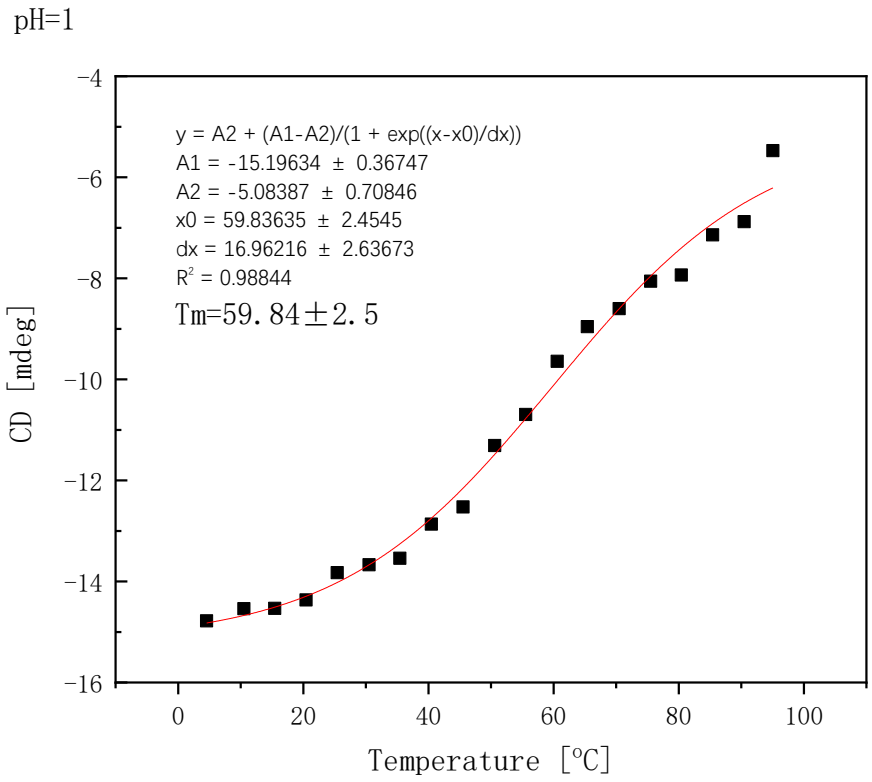


Figure S9 Melting curve (210nm) of MurJ at different pH 1.

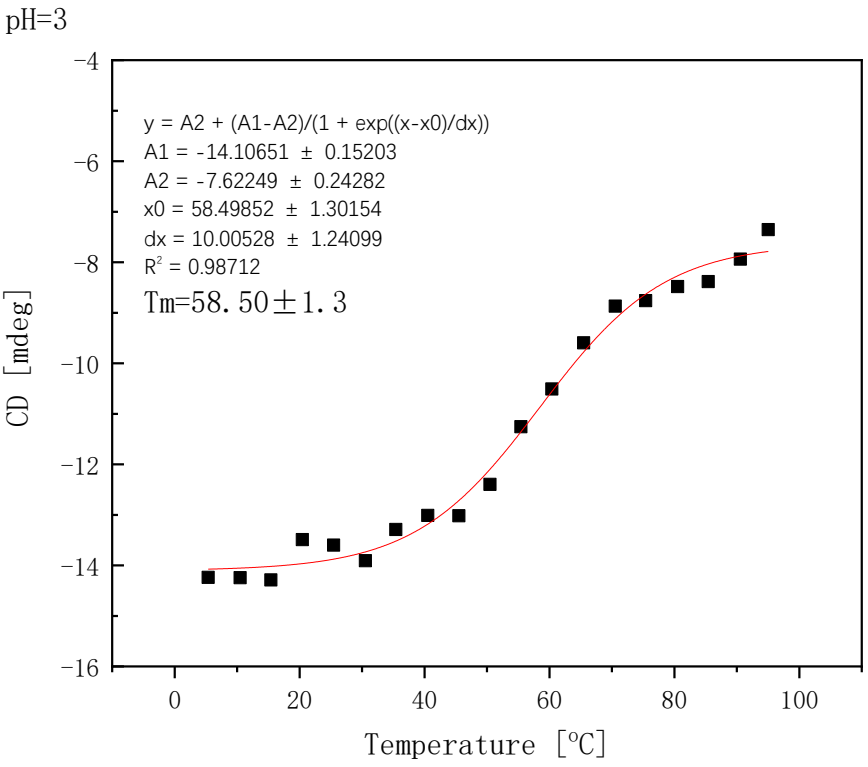


Figure S10 Melting curve (210nm) of MurJ at different pH 3.

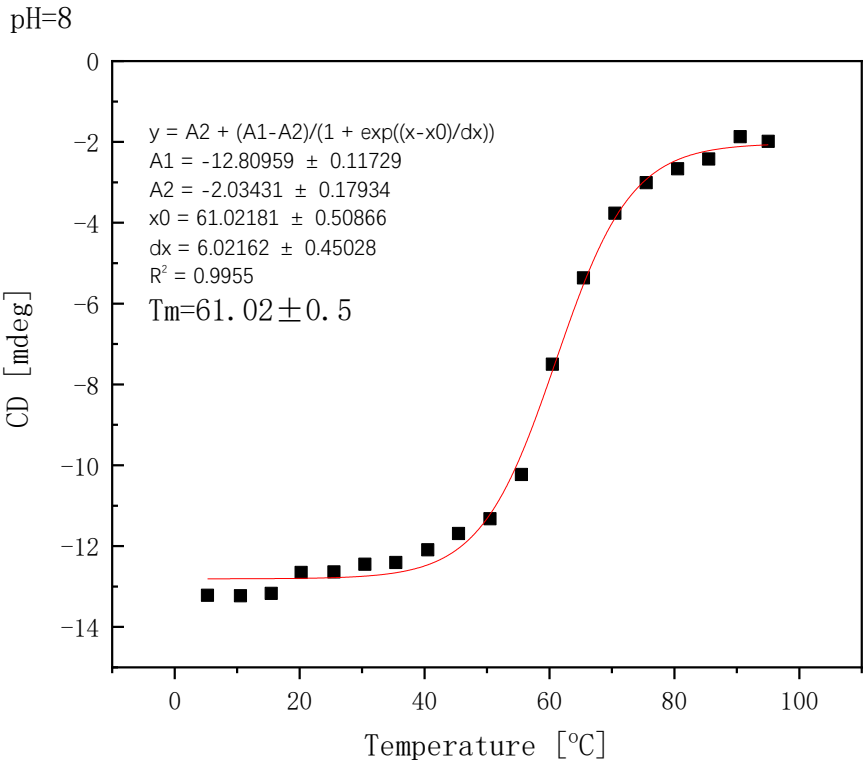


Figure S11 Melting curve (210nm) of MurJ at different pH 8.

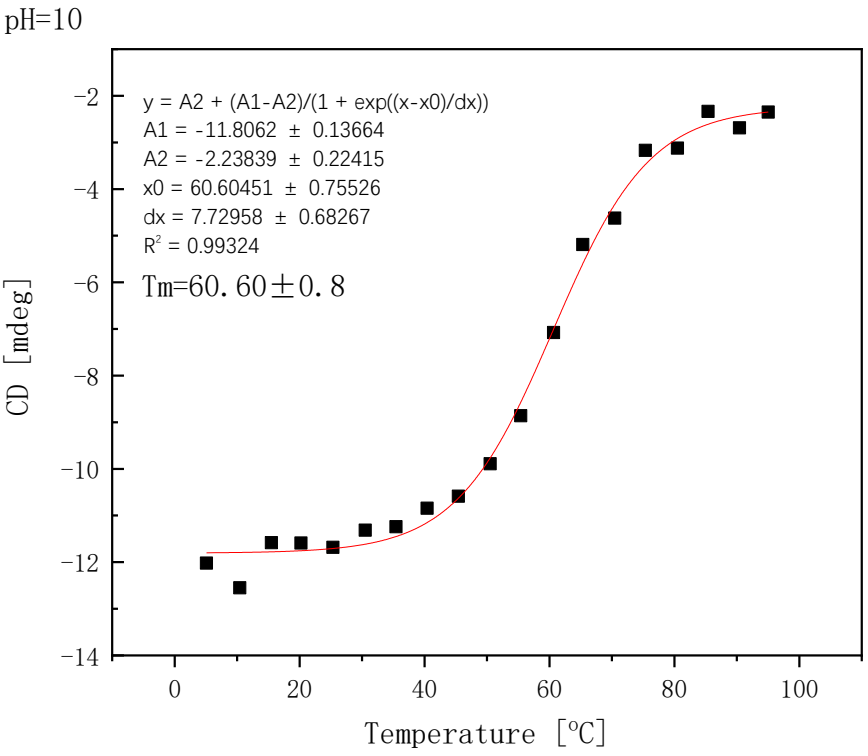


Figure S12 Melting curve (210nm) of MurJ at different pH 10.

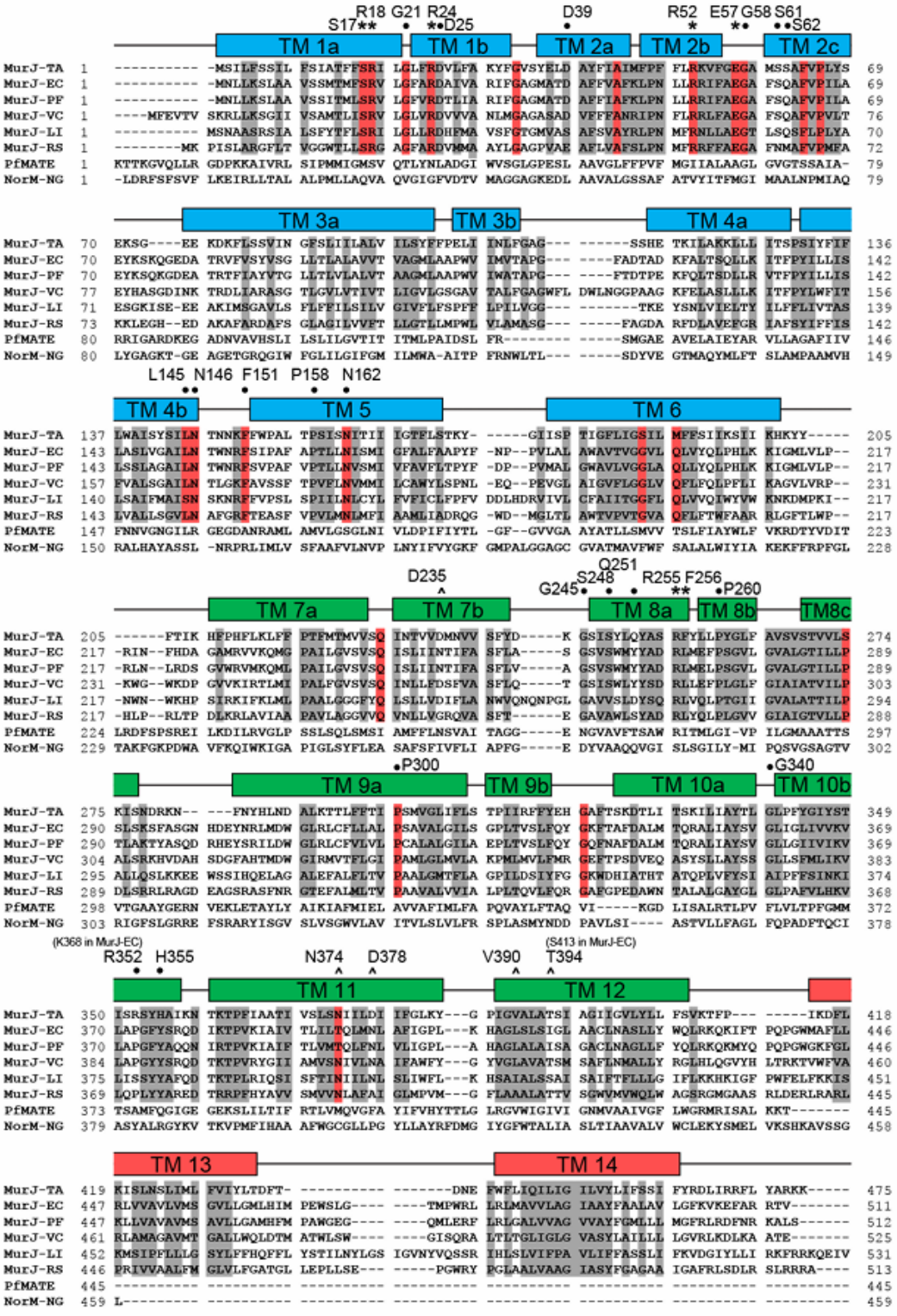


Figure S13 Alignment of MurJ sequences. This figure is from supplementary data reported by Alvin C. Y. Kuk, Aili Hao, Ziqiang Guan, and Seok-Yong Lee in 2019 [14]. It is an open access article with unrestricted use permission.

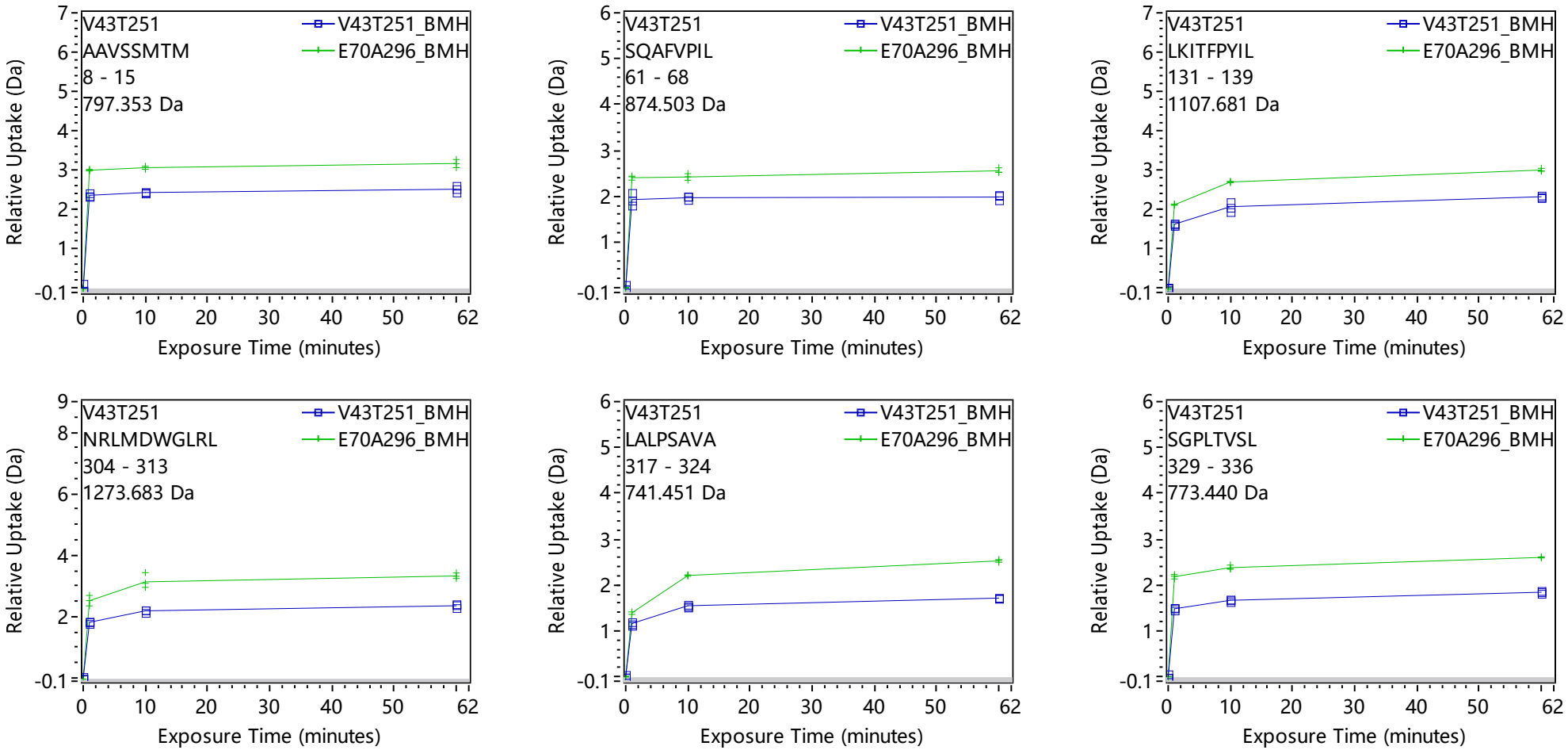


Figure S14 Deuterium uptake difference curves of peptides with significant change between V43C_T251C (inward-open) and E70C_A296C (outward-open) conformations.

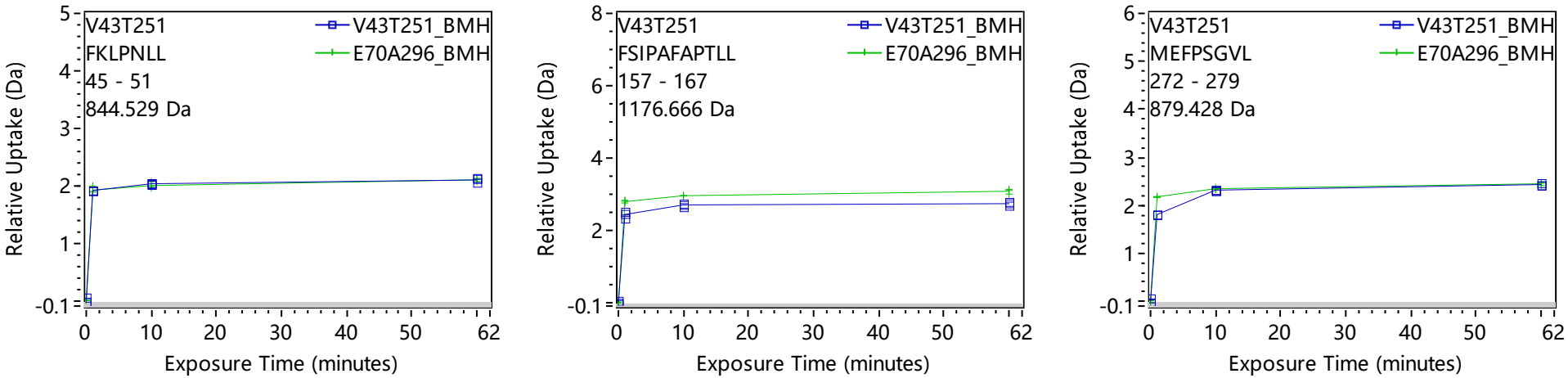


Figure S15 Deuterium uptake difference curves of peptides with less change between V43C_T251C (inward-open) and E70C_A296C (outward-open) conformations.

Table S1. Detailed data for local HDX-MS of MurJ in inward-open (BMH cross-linked V43C_T251C) conformation.

Start	End	Sequence	Maximum Uptake	Peptide monoisotopic mass (Da)	Retention Time (min)	Deuterium Uptake (Da)			Standard deviation (Da) (n=3)		
						1min	10min	60min	1min	10min	60min
8	15	AAVSSMTM	7	797.3532	4.53	2.366	2.436	2.522	0.091	0.083	0.108
16	22	FSRVLGF	6	825.4618	5.66	1.753	1.938	2.046	0.079	0.073	0.088
29	36	ARIFGAGM	7	822.4291	4.96	2.401	2.653	2.756	0.072	0.075	0.095
45	51	FKLPNLL	5	844.5291	6.26	1.914	2.033	2.098	0.057	0.071	0.076
61	68	SQAFVPIL	6	874.5033	6.31	1.927	1.967	1.981	0.118	0.057	0.068
96	103	AVVTVAGM	7	747.4069	4.89	0.338	0.904	1.618	0.040	0.029	0.036
104	111	LAAPWVIM	6	900.5012	6.92	1.320	2.228	2.755	0.112	0.206	0.248
105	111	AAPWVIM	5	787.4171	6.54	1.179	1.886	2.309	0.027	0.070	0.037
112	123	VTAPGFADTADK	10	1192.5844	4.05	3.533	3.609	3.584	0.091	0.118	0.141
112	124	VTAPGFADTADKF	11	1339.6529	5.14	3.944	4.019	3.973	0.099	0.134	0.166
124	130	FALTSQL	6	779.4298	5.59	2.399	2.556	2.599	0.133	0.160	0.164
131	139	LKITFPYIL	7	1107.6812	6.81	1.630	2.079	2.329	0.068	0.115	0.054
157	167	FSIPAFAPTLL	8	1176.6663	7.14	2.433	2.700	2.734	0.096	0.164	0.107
177	187	FAAPYFNPPVL	7	1235.6459	6.51	1.790	2.347	2.576	0.120	0.150	0.174
195	201	VGGVLQL	6	685.4243	5.71	0.930	1.487	1.705	0.422	0.311	0.257
202	208	VYQLPHL	5	869.488	5.51	1.211	1.269	1.282	0.131	0.179	0.239
202	210	VYQLPHLKK	7	1125.6779	4.08	2.235	2.473	2.477	0.067	0.063	0.070
202	213	VYQLPHLKKIGM	10	1426.8239	4.85	2.844	3.206	3.278	0.115	0.062	0.074
227	238	MRVVKQMGPAIL	10	1342.7698	4.93	3.301	3.481	3.649	0.380	0.378	0.418
228	238	RVVKQMGPAIL	9	1211.7293	4.72	2.641	2.818	2.998	0.222	0.215	0.214
229	238	VVKQMGPAIL	8	1055.6282	5.06	2.336	2.487	2.584	0.057	0.059	0.068
239	247	GVSVSQISL	8	889.4989	5.38	2.830	3.351	3.614	0.071	0.112	0.103
257	265	LASGSVSWM	8	937.4448	5.83	2.395	2.737	2.803	0.255	0.269	0.277
265	271	MYYADRL	6	931.4342	4.85	1.256	1.587	1.921	0.090	0.097	0.049
272	279	MEFPSGVL	6	879.4281	6.08	1.808	2.315	2.434	0.120	0.177	0.185

280	287	GVALGTIL	7	743.4662	5.96	1.956	2.479	2.804	0.050	0.049	0.049
281	287	VALGTIL	6	686.4447	5.88	1.546	1.916	2.225	0.031	0.031	0.036
304	313	NRLMDWGLRL	9	1273.6834	6.07	1.816	2.189	2.354	0.172	0.201	0.214
307	313	MDWGLRL	6	890.4553	6.21	1.030	1.286	1.403	0.052	0.064	0.064
317	324	LALPSAVA	6	741.4505	4.97	1.157	1.544	1.713	0.042	0.036	0.027
326	333	GILSGPLT	6	757.4454	5.51	1.207	1.430	1.478	0.022	0.040	0.089
329	336	SGPLTVSL	6	773.4403	5.55	1.481	1.664	1.841	0.042	0.041	0.043
346	354	DALMTQRAL	8	1018.535	4.63	1.520	1.811	2.259	0.264	0.243	0.125
355	361	IAYSVGL	6	722.4083	5.36	0.558	1.193	1.557	0.021	0.035	0.021
365	374	IVVKVLAPGF	8	1042.6659	5.58	0.683	1.287	1.902	0.036	0.055	0.039
400	410	AFIGPLKHAGL	9	1123.6622	5.00	2.289	2.576	2.880	0.064	0.061	0.085
402	410	IGPLKHAGL	7	905.5567	4.16	1.596	1.827	2.030	0.073	0.045	0.050
402	411	IGPLKHAGLS	8	992.5887	4.67	2.066	2.362	2.649	0.072	0.102	0.072
402	412	IGPLKHAGLSL	9	1105.6728	4.85	2.301	2.547	2.850	0.114	0.098	0.081
403	410	GPLKHAGL	6	792.4727	5.00	2.387	2.530	2.707	0.393	0.469	0.273

Table S2. Detailed data for local HDX-MS of MurJ in outward-open (BMH cross-linked E70C_A296C) conformation.

Start	End	Sequence	Maximum Uptake	Peptide monoisotopic mass (Da)	Retention Time (min)	Deuterium Uptake (Da)			Standard deviation (Da) (n=3)		
						1min	10min	60min	1min	10min	60min
8	15	AAVSSMTM	7	797.3532	4.52	3.003	3.067	3.174	0.016	0.031	0.084
16	22	FSRVLGF	6	825.4618	5.66	2.232	2.355	2.480	0.073	0.074	0.061
29	36	ARIFGAGM	7	822.4291	4.95	2.711	2.779	2.879	0.059	0.041	0.062
45	51	FKLPNLL	5	844.5291	6.26	1.920	2.001	2.099	0.069	0.070	0.093
61	68	SQAFVPIL	6	874.5033	6.31	2.400	2.420	2.552	0.038	0.060	0.045
96	103	AVVTVAGM	7	747.4069	4.89	0.393	1.120	2.107	0.048	0.061	0.017
104	111	LAAPWVIM	6	900.5012	6.92	1.668	2.418	2.758	0.183	0.228	0.266
105	111	AAPWVIM	5	787.4171	6.53	1.405	1.942	2.228	0.055	0.037	0.038
112	123	VTAPGFADTADK	10	1192.5844	4.04	3.512	3.518	3.568	0.054	0.054	0.118
112	124	VTAPGFADTADKF	11	1339.6529	5.13	3.989	3.940	3.996	0.042	0.024	0.123
124	130	FALTSQL	6	779.4298	5.59	2.486	2.584	2.670	0.158	0.145	0.171
131	139	LKITFPYIL	7	1107.6812	6.81	2.127	2.705	3.007	0.049	0.042	0.076
157	167	FSIPAFAPTLL	8	1176.6663	7.13	2.791	2.952	3.078	0.113	0.140	0.161
177	187	FAAPYFNPPVL	7	1235.6459	6.50	1.924	2.436	2.708	0.105	0.159	0.159
195	201	VGGVLQL	6	685.4243	5.71	1.073	1.826	2.081	0.419	0.255	0.235
202	208	VYQLPHL	5	869.488	5.50	1.526	1.630	1.778	0.047	0.094	0.222
202	210	VYQLPHLKK	7	1125.6779	4.08	2.542	2.694	2.791	0.046	0.038	0.074
202	213	VYQLPHLKKIGM	10	1426.8239	4.84	3.179	3.439	3.627	0.138	0.152	0.161
227	238	MRVVKQMGPAIL	10	1342.7698	4.92	4.234	4.236	4.453	0.356	0.407	0.352
228	238	RVVKQMGPAIL	9	1211.7293	4.71	3.392	3.415	3.611	0.197	0.199	0.196
229	238	VVKQMGPAIL	8	1055.6282	5.05	2.979	2.972	3.130	0.047	0.046	0.031
239	247	GVSVSQISL	8	889.4989	5.37	3.363	3.537	3.823	0.090	0.115	0.105
257	265	LASGSVSWM	8	937.4448	5.83	2.921	2.956	3.033	0.324	0.341	0.312
265	271	MYYADRL	6	931.4342	4.85	1.816	1.971	2.186	0.175	0.130	0.087
272	279	MEFPSGVL	6	879.4281	6.07	2.174	2.346	2.451	0.219	0.204	0.198

280	287	GVALGTIL	7	743.4662	5.96	2.529	2.921	3.191	0.025	0.031	0.010
281	287	VALGTIL	6	686.4447	5.88	2.121	2.463	2.672	0.040	0.038	0.042
304	313	NRLMDWGLRL	9	1273.6834	6.07	2.516	3.132	3.328	0.260	0.351	0.279
307	313	MDWGLRL	6	890.4553	6.21	1.536	1.923	2.058	0.040	0.062	0.061
317	324	LALPSAVA	6	741.4505	4.96	1.390	2.206	2.526	0.026	0.015	0.025
326	333	GILSGPLT	6	757.4454	5.50	1.585	1.840	1.970	0.016	0.047	0.016
329	336	SGPLTVSL	6	773.4403	5.55	2.179	2.376	2.599	0.040	0.040	0.011
346	354	DALMTQRAL	8	1018.535	4.62	2.107	2.329	2.716	0.131	0.123	0.126
355	361	IAYSVGL	6	722.4083	5.35	0.969	1.937	2.382	0.029	0.028	0.037
365	374	IVVKVLAPGF	8	1042.6659	5.57	1.043	1.882	2.605	0.043	0.025	0.052
400	410	AFIGPLKHAGL	9	1123.6622	4.99	2.835	2.939	3.145	0.067	0.081	0.086
402	410	IGPLKHAGL	7	905.5567	4.15	1.997	2.025	2.179	0.084	0.056	0.085
402	411	IGPLKHAGLS	8	992.5887	4.66	2.602	2.653	2.858	0.036	0.053	0.126
402	412	IGPLKHAGLSL	9	1105.6728	4.84	2.841	2.952	3.277	0.081	0.146	0.135
403	410	GPLKHAGL	6	792.4727	4.99	2.710	2.663	2.778	0.366	0.423	0.358