Mass Spectrometry-Guided Structural Analysis Of Protein Kinase Inhibitor Complexes

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

By

Lauren Tomlinson

March 2022

Acknowledgements

Firstly, I would like to thank my supervisor, Claire Eyers, for your support and guidance during the last four years. I have enjoyed your look of puzzlements as I ramble around trying to explain something. Thank you also to my second supervisor, Richard Bayliss, and Matthew Batchelor at the University of Leeds, who both helped contribute greatly to the work in this thesis with the molecular modelling aspect of Aurora A. I would also like to thank Pat Eyers and all the members of Lab D for their continued enthusiasm and support throughout this PhD. In particular, thank you to Dom Byrne for all the initial training on protein purification and his continued guidance during the PhD.

Thanks to the past and present members of the Centre for Proteome Research for all your ongoing encouragement. I'm sure you will all miss hearing the word 'conformer' in your weekly meetings. A special mention must go to Philip Brownridge, who can fix any problem you might come across, which was needed when the good old Synapt G2Si didn't fancy playing ball.

I would like to thank my family and friends for their support, who I'm very confident have no idea what I have been doing for the last four years, but you can all read the wonders of this thesis to get to grips on the world of IM-MS (a test will be carried out at the end).

Last but not least, I would like to thank Alice Clubbs Coldron, who I started this PhD alongside and is now one of my very good friends. The PhD experience certainly wouldn't have been the same without Alice by my side. Those tea/coffee trips and tennis battles got us through the last four years. Alice also generated the NME proteins for this thesis (had to think of something to get those extra coffee trip science chats in).

Abstract

Protein kinases are involved in the vast majority of cellular processes including cell cycle regulation and metabolism, with the dysregulation of protein kinases being a significant contributing factor in many diseases. Protein kinase inhibitors can thus be very effective in treating diseases that are driven by abnormal kinase signalling. To aid in the development of small molecule protein kinase inhibitors, either as potential therapeutics, or as chemical tools to explore cellular signalling mechanisms, it is therefore important to understand the effects that such small molecules have on protein kinase structure and activity. Mass spectrometry (MS) and in particular native ion-mobility mass spectrometry (IM-MS) can be used to explore conformational dynamics, folding/unfolding intermediates and ligand-induced conformational changes.

The work presented in this thesis used IM-MS to characterise changes in the conformational landscape and stability of two protein kinases, Aurora A and NME1, as a function of activity. To explore conformational dynamics of the essential mitotic Ser/Thr kinase, I made use of peptide activators, small molecule inhibitors and a single point mutation which rendered the protein kinase catalytically inactive. Aided by molecular modelling, I established three major conformations, the relative abundance of which were dependent on the Aurora A activation status: one highly populated compact conformer similar to that observed in most crystal structures, a second highly populated conformer possessing a more open structure infrequently found in crystal structures, and an additional low-abundance conformer not currently represented in the protein databank. Notably, inhibitor binding induced more compact configurations of Aurora A, as adopted by the unbound enzyme, with both IM-MS and modelling revealing inhibitor-mediated stabilization of active Aurora A.

In addition to the serine/threonine Aurora A kinase, the structural conformation of the His kinase NME1 protein was also investigated, using a number of point mutations to unravel key regulatory sites. NME1 activity and dimerisation status were also explored in the presence of a variety of small molecule compounds. These results reveal that dissociation from the native active hexamer to a dimer could be associated with oxidative stress due to some form of resistance to dissociation observed with a NME1 C109A mutated protein and the oxidising abilities of the compounds tested.

Table of Contents

List of Figures

List of Tables

Abbreviations

ATD	Arrival time distribution
ATP	Adenosine triphosphate
APK	Atypical protein kinase
BCR	Breakpoint cluster region
CCS	Collision cross section
CE	Collision energy
CCSD	CCS half-height width distribution
CIU	Collision-induced unfolding
CML	Chronic myelogenous leukaemia
DC	Direct Current
DDA	Data dependent acquisition
DSF	Differential scanning fluorimetry
DTIMS	Drift-tube ion mobility spectrometry
EHSS	Exact hard-sphere scattering
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
EPK	Eukaryotic protein kinase
FAIMS	Field-asymmetric ion mobility spectrometry
FDA	Food Drug Administration
FDR	False discovery rate
HCD	Higher-energy collisional dissociation
IMPACT	Ion Mobility Projection Approximation Calculation Tool
IM-MS	Ion mobility-mass spectrometry
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
MALDI	Matrix-assisted laser desorption/ionisation
MS	Mass spectrometry
MW	Molecular weight
MYCN	N-myc proto-oncogene protein
m/z	Mass-to-charge ratio
NDPK	Nucleoside diphosphate protein kinases

NME	Non-metastatic cells
NMR	Nuclear magnetic resonance
PA	Projection approximation
PD	Proteome Discoverer
pHis	Phosphohistidine
PKA	Protein kinase A
PTM	Post-translational modification
RF	Radio Frequency
SD	Standard Deviation
TACC	Transforming acidic coiled coil
TCS	Two component system
TM	Trajectory method
TOF	Time-of-flight
TPX2	Targeting protein for Xklp2
TWIM	Travelling Wave Ion Mobility
V	Voltage
WT	Wild-type

Amino Acids

Amino acid	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Met	K
Methionine	Phe	M
Phenylalanine	Pro	F
Proline	Ser	P
Serine	Thr	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Publications

Tomlinson LJ, Batchelor M, Sarsby J, Byrne DP, Brownridge PJ, Bayliss R, Eyers PA, Eyers CE. (2022). Exploring the Conformational Landscape and Stability of Aurora A Using Ion-Mobility Mass Spectrometry and Molecular Modeling. Journal of the American Society for Mass Spectrometry. 33(3):420-435. doi: 10.1021/jasms.1c00271. PMID: 35099954.

Tomlinson LJ, Eyers CE. Ion Mobility-Mass Spectrometry to Evaluate the Effects of Protein Modification or Small Molecule Binding on Protein Dynamics. (2020). Methods Mol Biol. doi: 10.1007/978-1-0716-0030-6_11. PMID: 31729661.

Tomlinson LJ, Clubbs Coldron AKM, Eyers PA, Eyers CE. (2020). Determination of Phosphohistidine Stoichiometry in Histidine Kinases by Intact Mass Spectrometry. Methods Mol Biol. doi: 10.1007/978-1-4939-9884-5_6. PMID: 31707653.

Chapter 1. Introduction

1.1   Protein kinases

Eukaryotic cells are exposed to many extracellular signals that result in a specific cellular response. Signalling within the cell is controlled by post-translational modifications (PTM), with phosphorylation one of the most common PTM's. The role of protein kinases is to transfer phosphoryl groups to target proteins, resulting in protein phosphorylation and can often result in the modification of their activity status. Phosphorylation was first identified in the 1950s by Krebs and Fischer, with this PTM now one of the most studied mechanisms that drives cellular responses (1). Phosphorylation is reversed by phosphatases, which are responsible for the removal of phosphoryl moieties from target proteins (2). Protein phosphorylation results in the activation of signal transduction pathways, which are associated with various biological processes. In particular, protein kinases have important roles in the regulation of cell growth, differentiation, development and death. There are two main classifications of mammalian protein kinases depending on which amino acid they phosphorylate: serine/threonine or tyrosine residues (Figure 1.1). Both of these classifications of protein kinases consist of a glycine-rich N-terminal ATP binding pocket and a conserved aspartic acid residue, which is essential for catalytic activity (3). Protein kinases can act as either intracellular signal mediators or cell membrane receptors. Activation of membrane receptors leads to a chain of events associated with kinase interactions that induces a cellular response often through regulation of gene transcription. The loss of kinase receptor interactions, mutations which alter kinase or phosphatase catalytic activity, or regulatory phosphorylation sites can lead to uncontrolled cell proliferation and cancer (4).
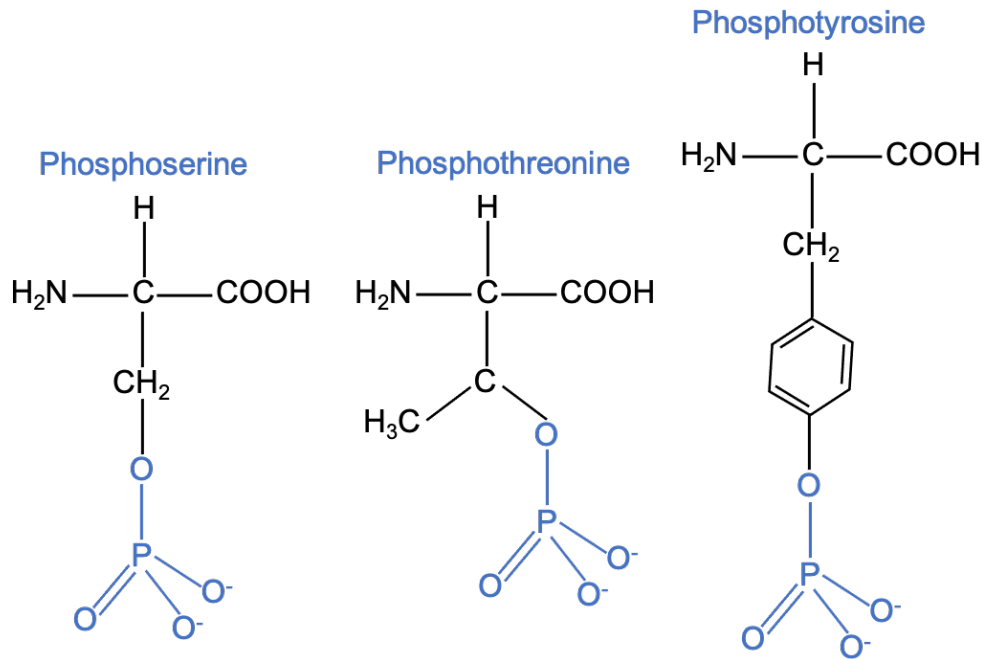
Figure 1.1. Structures of phosphoserine, phosphothreonine and phosphotyrosine.

The human genome encodes more than 500 protein kinases making up 2% of human genes. Protein kinases are separated into groups depending on their structure, specificity, and regulation. Eukaryotic protein kinases (ePK) make up a large proportion of kinases, classified into eight groups, whilst other kinases known as atypical protein kinases (aPKs) are categorised into an additional four groups, which are kinases that do not have sufficient sequence similarity to be classified as an ePK (5). The ePK group was further divided into nine groups (Figure 1.2): AGC consisting of 63 members, CAMK with 74 members, CK1 with 12 members, CMGC with 61 members, STE with 47 members, TK with 90 members, TKL with 43 members, RCG with 5 members, and a final group known as 'Other' with 83 members. In addition to the nine kinase families, a further 40 kinases are classified as atypical (6). All kinases shown in the phylogeny diagram have structural similarities, including a universal conserved core region, which is associated with the ATP binding pocket and is where the residues involved in phosphotransfer reactions are situated. The conserved region is modified based on the spatial placement of the structural elements. The structural elements are distinctive for each kinase, which are joined to the N- and C-terminal ends of the universal core region. Within the universal core region, many kinases have structural insertions that have no spatial similarity between families (7).
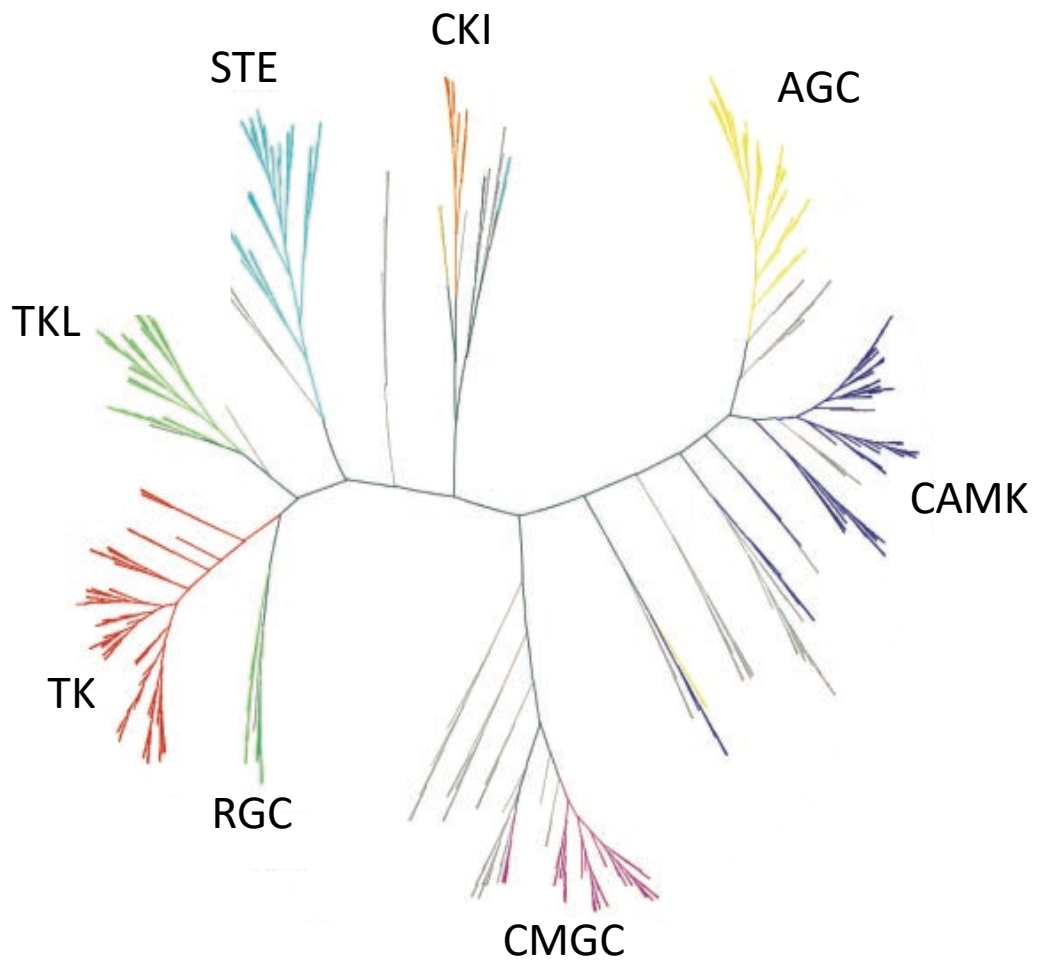
Figure 1.2. Protein kinase classification of the human genome. Eukaryotic protein kinases are classified into nine groups (eight shown here) based on their structure, specificity, and regulation properties. Figure adapted from (5).

1.1.1 Protein kinase structure

Protein kinases consist of a fold of two lobes: an N-terminal lobe and a C-terminal lobe (Figure 1.3). The N-terminal lobe consists of a five-stranded β-sheet with an α-helix, termed the C-helix, whereas the C-terminal consists of six α-helices. Each of the two lobes are connected via a hinge region that forms the ATP-binding site for the protein, which plays a vital role in enzymatic activity. The activation loop begins with a conserved DFG motif and extends up to an APE motif, with this loop forming a cleft that binds substrates when in an active kinase conformation. The bound substrate peptide then subsequently form interactions with the conserved HRD motif that sits within the catalytic loop (8).
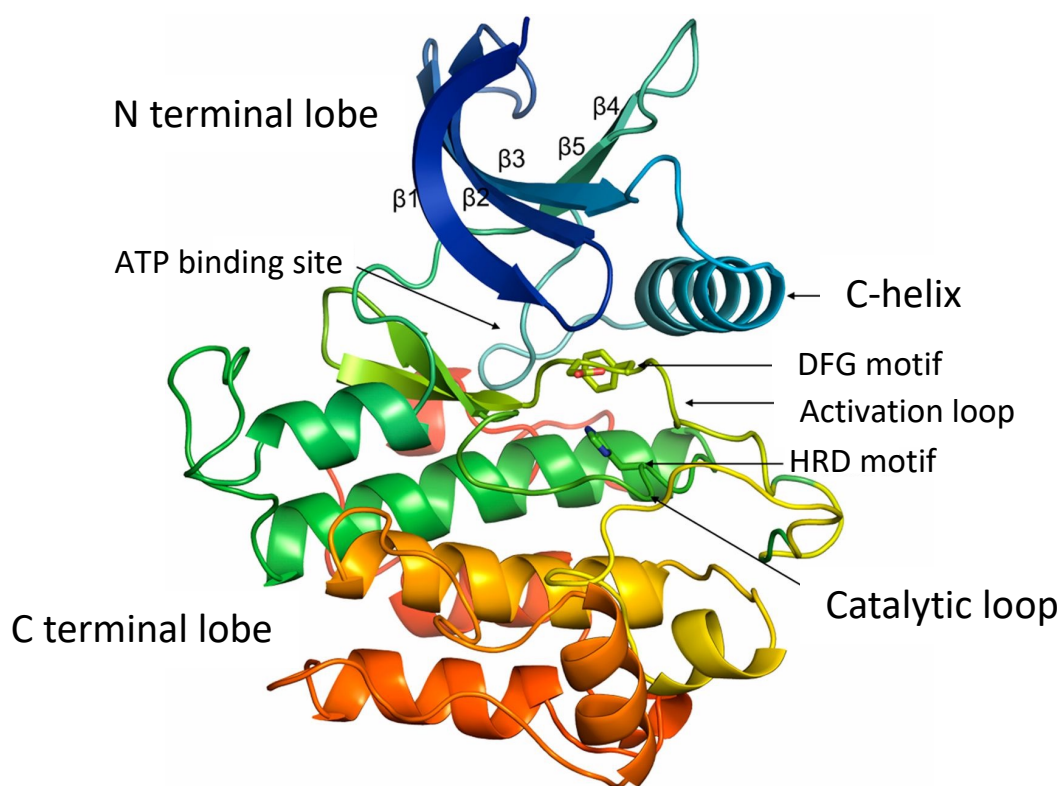
Figure 1.3. Example of a protein kinase structure. Figure adapted from (8).

1.1.2 Magnesium binding DFG motif

The Asp in the DFG motif is positioned to bind magnesium and interact with an oxygen atom of the β phosphate of ATP when in the active conformation. Active conformations result in an inwards positioning of the C-helix that enables a Glu to form a salt bridge with a Lys residue. The positively charged β3 lysine residue is responsible for the coordination of the α and β phosphates of ATP to the C-helix, resulting in stabilisation of the nucleotide due to the clamping down of the glycine rich loop, which leads to the exposure of the γ-phosphate. On

formation of the salt bridge between β3 and a conserved Glu residue, hydrogen bonds with oxygen on the α and β phosphates of ATP take place, which are stabilised by a GxGxxG motif within the N-lobe (9).

1.1.3 The αC-helix

Positioning of the αC-helix plays an important role in conformational changes that occur within the catalytic domain. The αC-helix is the only conserved helix in the β-sheet rich N-lobe, with connections to various aspects of the molecule, and therefore, acts as a signal integration motif (10). The positioning of the αC-helix between the N-lobe and C-helix determines whether the kinase will adopt an open or closed conformation for an active or inactive kinase respectively. This position can be allosterically modulated via the interaction from activating or inhibitory mechanisms (11). Many AGC protein kinases, including protein kinase A (PKA), are associated with a hydrophobic motif (HM), which is a 50-60 amino acid C-terminal tail sequence situated after the conserved catalytic domain (12, 13). The general consensus sequence of the HM is F/Y-X-X-F/Y-S/T/E/D-F/Y (where X is any amino acid), with this motif docking to a hydrophobic surface in the N-lobe that is created by the αC-helix, known as the PDK1-interacting fragment (PIF)-pocket (14). Phosphorylation of AKT at the Ser473 residue of the HM leads to further stabilisation of the kinase in the active conformation due to the enhanced binding of the PIF pocket (15).

1.1.4 The catalytic motif

The Asp residue in the HRD catalytic motif becomes protonated by the acceptance of a hydrogen from the hydroxyl group of the R group of a substrate amino acid, commonly being a Ser, Thr or Tyr (16). The Asn residue within the HRD motif enables the binding of two Mg^{2+} ions, resulting in the orientation of the catalytic Asp residue (17). The histidine situated at the beginning of the HRD motif is highly conserved and acts as a central scaffold for the catalytic spine. This histidine binds to the carbonyl of the Asp in the DFG motif, which also hydrophobically contacts with the Phe residue of the DFG motif (18).

1.1.5 Activation loop phosphorylation

Protein kinases that contain the HRD catalytic motif require phosphorylation of the activation loop for catalysis. The positively charged Arg residue provides a link between the catalytic motif, DFG motif and activation loop via a phosphorylated residue within the activation loop. In PKA, the first kinase crystal structure to be solved, phosphorylation of the Thr197 residue, situated next to the catalytic Asp of HRD, is required for catalytic activity. Phosphorylation of Thr197 results in a conformational change from DFG-out to the catalytically active DFG-in mode, where the coordination of His87 on the αC-helix, Arg165 in the catalytic motif and Lys189 of the β9 sheet all takes place (19).

1.1.6 The αF-helix

The αF-helix has been shown to acts as a central scaffold for all the elements that are required for full catalytic function (Figure 1.4) (9). The mobility of the αC-helix and the activation loop enables the hydrophobic spine to be actively assembled and disassembled, which leads to regulation of protein kinase activity, and therefore, has been termed the regulatory spine (16). The regulatory spine and the C-lobe are both anchored to the F-helix, with the C-spine connected to the hydrophobic C-terminus and the R-spine to the N-terminus. One residue of the regulatory spine in the C-lobe is associated with the HRD motif of the catalytic loop. This residue is commonly a histidine. However, in PKA and the majority of the AGC family, the histidine is substituted for a Tyr residue. The histidine or tyrosine residue is anchored to the F-helix by a conserved aspartate, which then acts as the base of the regulatory spine. The additional hydrophobic spine, known as the catalytic spine, also connects with residues from both the N and C-lobes. However, in contrast to the regulatory spine, the catalytic spine also consists of the adenine ring of ATP. The formation of the catalytic spine occurs from the Ala residue within the VAIK motif of the β3 sheet, with this residue interacting with the adenine ring of ATP (9).
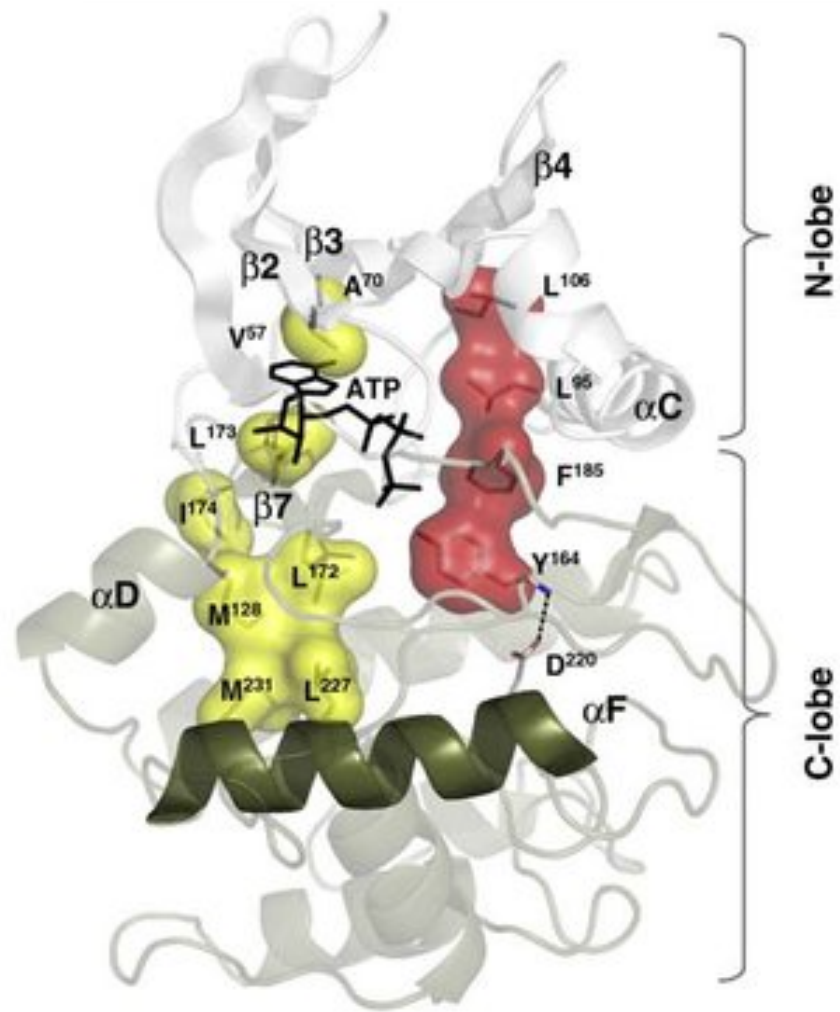
Figure 1.4. The F-helix interaction with the two hydrophobic spines of PKA. Regulatory spine (red) and catalytic spine (yellow) connected to the F-helix (green) at the centre of the C-lobe. ATP provides a link between the N-lobe and C-lobe. Figure obtained from (9).

1.2 Aurora kinases

Aurora kinases belong to the family of serine/threonine protein kinases, which are encoded by the IpI1p gene, also known as the Aurora gene (20). Three Aurora kinases have been identified in mammals: Aurora A, Aurora B and Aurora C. The three Aurora kinases share 60-75% sequence similarity in the kinase domain. Aurora C lacks the N-terminal domain that is present in both Aurora A and B, suggesting that Aurora C has different regulation (21). Within the catalytic domain, the amino acid sequence for each of the Aurora A isoforms is highly conserved between different organisms, which highlights their importance for various protein functions of different species.

The N-terminal domain of each Aurora kinase, which is responsible for substrate specificity, intracellular localisations and other functions, have much less similarity (22). Aurora kinases contain a regulatory domain at the N-terminus and a catalytic domain at the C-terminal (Figure 1.5). Within the N-terminal region, Aurora A contains an A-box, which is involved in kinase degradation. The D-box within the C-terminal region also plays an important role in self-degradation. The regulation of expression levels and activity of Aurora A is associated with phosphorylation/dephosphorylation and protein degradation. Kinase activity of Aurora A is stimulated by phosphorylation, with phosphorylation on Thr288, situated within the activation loop of human Aurora A required for kinase activity (23).
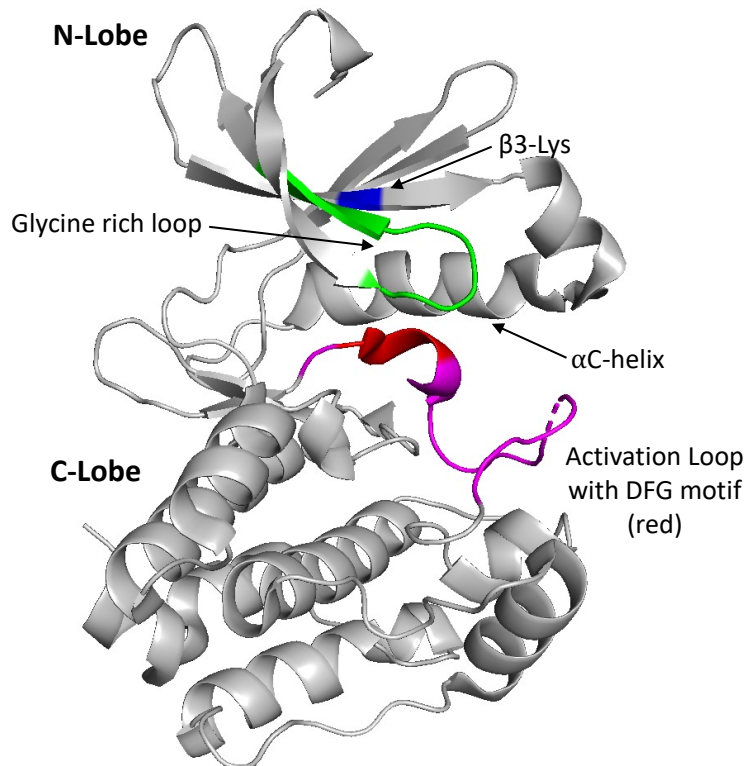
Figure 1.5. Aurora A protein kinase structure. Structural features including the activation loop (magenta), DFG motif (red), glycine rich loop (green), Lys162 (blue). Diagram produced in PyMol from PDB file 1MQ4.

1.2.1 Aurora cell cycle

The cell cycle is divided into different stages, including interphase, prophase, prometaphase, metaphase, anaphase and telophase (24). Human Aurora kinases are cell cycle regulated; low in the G1/S phase and high in the G2/M phase, with each kinase having a distinct role during mitosis (25). In order for a cell to undergo mitosis, replication of the organelles and DNA must be carried out prior to the separation of the chromosomes into the daughter cell (24). Aurora kinases are associated with various functions of mitosis, including mitotic entry, centriole pair separation, bipolar spindle assembly, metaphase chromosome alignment and cytokinesis completion (26). Aurora A has a distinct characteristic with its association to centrosomes and microtubule regions close to the centrosome in comparison to Aurora B and C. Aurora A specifically associates with centrosomes that separate during the late stages of S phase and early G2 phase (27). The centrosome association is directed by the carboxyl-terminal catalytic

domain and the amino-terminal region. However, this does not require kinase activity. Aurora A has also been located at the spindle midzone and midbody in the late stages of mitosis (28).

1.2.2 Aurora and cancer

Aurora kinases are of importance to cancer research, as the genes encoding all three of the Aurora kinases can be mapped to regions of chromosomal abnormalities in various cancer types, with the detection of overexpression in tumour cell lines (28). Aurora A has been the predominant focus for research of the three Aurora kinases. Human Aurora A is mapped to chromosomal region 20q13.2, with this region being shown to be amplified in cancer cell lines and primary tumours. Irregular expression of Aurora A can lead to cancer, by uncoupling centrosome duplication during the cell cycle, which then leads to centrosome amplification (29). Overexpression of Aurora A has shown to result in mitotic abnormalities, which lead to cytokinesis failure and tetraploid cell production (30). The inhibition of Aurora A can lead to abnormalities in mitotic spindle assembly, which results in a spindle checkpoint-dependent mitotic arrest (31). Research has also shown that Aurora A inhibition can lead to MYCN degradation. Under physiological conditions, Aurora A binds to MYCN, sequestering it away from ubiquitin-mediated proteolytic degradation (32).

1.2.3 Aurora A and TPX2 activator

Aurora A is activated through autophosphorylation, enhanced by its interaction with TPX2 (Targeting protein for Xklp2) during mitosis (33). TPX2 is regarded as both a substrate and Aurora A activator, with its role associated with the positioning of the kinase to the spindle microtubules. TPX2 binds at two regions of Aurora A: the N-terminal lobe and a region between the N-terminal and C-terminal lobe (Figure 1.6), inducing a conformational change that enables autophosphorylation of the activating T288 residue (34). Aurora A on its own is deactivated by dephosphorylation of T288 by protein phosphatase 1. However, with TPX2-bound, the phosphorylated T288 residue is protected from phosphatases and activity is maintained (35). In the absence of phosphorylation, TPX2 has also been shown to increase the activity of Aurora A when the T288 is mutated to a non-phosphorylated residue due to a conformational change in the catalytic domain (34).
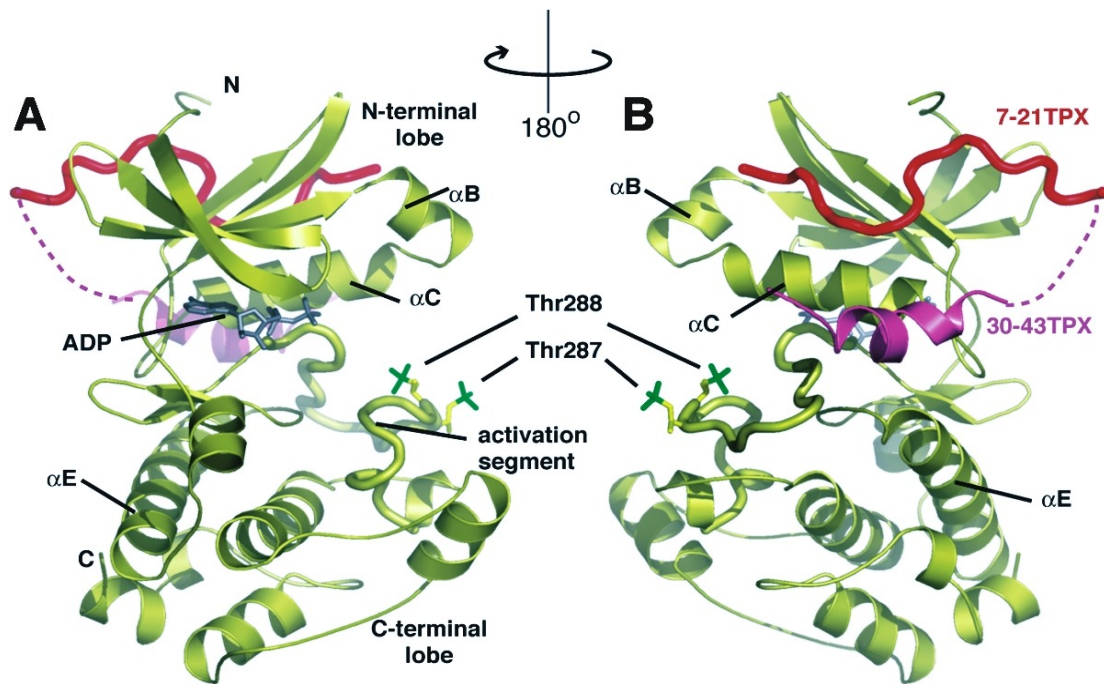
Figure 1.6. Aurora A bound to TPX2. A) Catalytic domain of Aurora A (yellow) and the N-terminal domain of TPX2. TPX2 binds at the N-terminal lobe (red) and between the two lobes (pink) of Aurora A. Dotted pink line connects the two regions where TPX2 binds. **B)** Complex oriented by 180° rotation to show the two regions of TPX2 binding to Aurora A. Figure adapted from (35).

1.2.4 Aurora A and N-MYC

The family of Myc transcription factors are oncogenic, being de-regulated in over 50% of all cancers (36).The inhibition of Myc proteins has been targeted as a validated therapeutic. However, compounds targeting Myc proteins have failed during the clinical stages (37). N-Myc was first discovered as the product of amplified genes in neuroblastoma (38). Conserved sequence motifs known as Myc boxes (MBO-IV) provide a docking site that enables protein-protein interactions (39), with the Myc transactivation domain situated in the N-terminal conserved motifs MB0, MBI and MBII. Myc stability is regulated by phosphorylation of MBI which targets the protein for ubiquitination and proteolysis (40). Aurora A inhibits this process in neuroblastoma cells, which leads to accumulation of N-Myc protein (41). Aurora A binds to the complex of N-Myc/SCF^{FbxW7}, leading to a decrease K48 linkages present in the polyubiquitin chains. The stabilisation of N-Myc is not dependent on catalytically active Aurora A. The Aurora A inhibitor, MLN8237, have been shown to alter the conformation of Aurora A, destabilising the interaction with N-Myc (32). Other inhibitors, such as MK-5108,

Hesperadin and CCT-137690, that compete with ATP and fail to induce any change in conformation, are understood to leave the Aurora A/N-Myc complex intact (42).

1.2.5 Aurora A and TACC3

TACC3, a member of the transforming acidic coiled coil (TACC) family of centrosome proteins is thought to enhance microtubule stability in mitosis (43). TACC3 dysregulation is implicated in neurodegenerative diseases and various cancers, including breast, ovarian and bladder (44, 45). Aurora A in human cells phosphorylates TACC3 on three conserved regions: S34 in the N-terminal domain, S552 in the clathrin-interaction domain, and S558 in the C-terminal domain (45). Phosphorylation of S558 in mammalian cells induces TACC3 localisation on the spindle microtubules and plays an important role in its binding to clathrin (46). TACC3 bound to clathrin and ch-TOG forms interactions between kinetochore fibres, suggesting a potential for cross-linking and the stabilisation of the mitotic microtubules (47).

1.3   Protein kinase inhibitors

Dysregulation and mutations of protein kinases can result in many human diseases; therefore, kinases have become key drug targets. Imatinib, the first FDA-approved small molecule kinase inhibitor to reach the market, is used to treat chronic myelogenous leukaemia (CML) (6). CML is associated with the activation of the Abl kinase following gene translocation to the breakpoint cluster region (BCR). Binding of Imatinib to BCR-Abl via six hydrogen bond interactions stabilises the complex and prevents ATP from reaching the binding site (48). The development of inhibitors used for many other cancers is challenging due to kinase dysregulation of various signalling pathways compared to the inhibition of a single signalling pathway that is effective with Imatinib. Many of the FDA-approved kinase inhibitors target the ATP-binding site. However, this can lead to selectivity issues, as many ATP sites of protein kinases have common features, which can ultimately result in unwanted off-target effects (6).

1.3.1 Kinase inhibitor nomenclature

The conformational state of a kinase on the inhibitor potency is a fundamental issue in the drug discovery world of selective kinase inhibitors. Protein kinase inhibitors target a variety

of active or inactive conformations, with the two major classes termed DFG-in and DFG-out. The DFG motif, situated within the kinase activation loop, plays an important role in catalysis. In active conformations, the Asp of the DFG motif is orientated to bind magnesium and interact with ATP. In contrast, inactive kinase conformations have a collapsed activation loop that inhibits the binding of substrates. The DFG-in mode can be a range of active or inactive conformations, where the Phe residue contacts both the C-helix of the N-terminal lobe. The DFG-out mode is associated with the inactive conformation, where the Phe residue occupies the ATP pocket, resulting in the exposure of the C-helix pocket (Figure 1.7) (8). The gatekeeper residue, situated within the R spine of protein kinases, plays an important role in the effectiveness of inhibitor binding. The amino acid that is situated as the gatekeeper residue determines how accessible the hydrophobic pocket is for inhibitor binding. Small residues such as Ala, Cys, Gly, Ser, Thr or Val, result in a large hydrophobic pocket, whereas large residues (Tyr or Phe) have a small pocket reducing the accessibility for binding (49). This residue is often mutated when the kinase is targeted with an inhibitor to trigger drug resistance (50).

Kinase inhibitors can be classified as Types I-IV. Type I inhibitors typically bind in the DFG-in conformation, contacting only the ATP site and not the conserved allosteric pocket. Type I inhibitors do not require a specific conformation for binding as they do not target the allosteric pocket. In contrast, Type II inhibitors, which are commonly associated with the DFG-out conformation, contact both the ATP site and an allosteric pocket. The allosteric pocket is only available for binding when the kinase is in its catalytically inactive conformation, where the DFG motif in the N-terminus of the activation loop is flipped out (51). In addition to Type I and II, Type I ½ inhibitors have been identified that act as a fusion of Type I and II classifications. The Type I ½ inhibitors can bind to various regions including the adenine-binding pocket, hinge residues, and the hydrophobic pocket II in the DFG-in conformation. Two Type 1 ½ inhibitors have been approved: vemurafenib for the treatment of melanoma, and lapatinib for the treatment of breast cancer. Type III inhibitors contact an allosteric site, which inhibits kinase catalytic activity without having any effect on ATP binding. Trametinib, is an FDA approved Type III MEK inhibitor responsible for treating melanoma (52). Type IV inhibitors form covalent bonds with the target enzyme, which can cause issues with safety and toxicity, making them less desirable as potential drug targets (53).
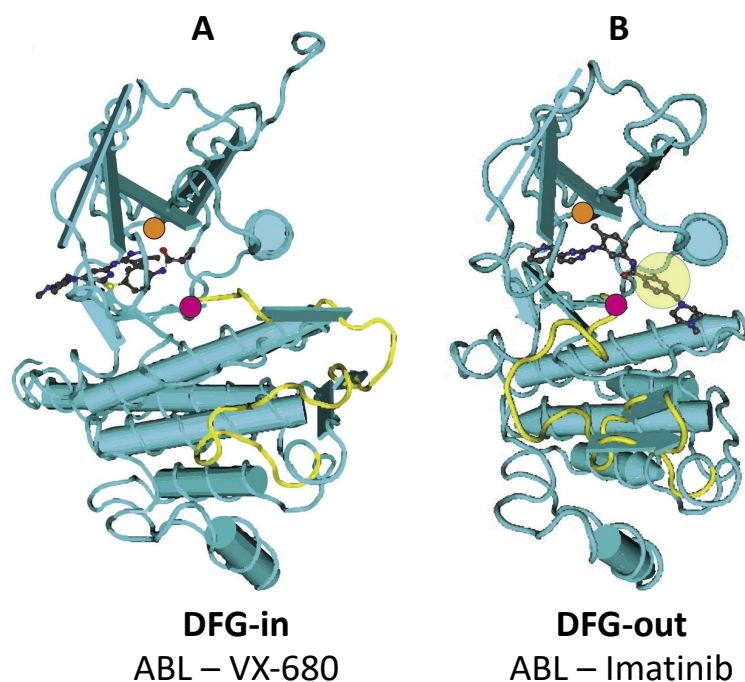
Figure 1.7. DFG-in and DFG-out configurations that protein kinases can adopt. A) Cocrystal structure of ABL (cyan) bound to VX-680 inhibitor (grey). Activation loop (yellow) adopts DFG-in configuration. **B)** ABL (cyan) bound to imatinib (grey) in DFG-out configuration exposing allosteric pocket (yellow circle). Gatekeeper residue shown as orange circle and DFG residue that is responsible for DFG configuration shown as magenta circle for both A and B. Figure obtained from (51).

The conformations of the DFG-in and DFG-out active kinases have been used as a defining tool to group classes of inhibitor compounds (54). However, there are a variety of kinases where the Phe residue of the DFG motif is positioned in a different orientation that is within an intermediate position between a DFG-in and DFG-out conformation. Work carried out by Modi and Dunbrack have classified all the human kinase structures to two levels of structural detail. They firstly clustered the kinase structures into three groups that focused on the position of the Phe side chain of the DFG-motif, which consisted of DFG-in, DFG-out and DFG-inter. The DFG-in group was the largest set with 227 kinases that orientated the DFG-Phe against or underneath the C-helix. The second largest group, the DFG-out conformation, consisted of 60 kinases where the DFG-Phe was orientated into the ATP binding pocket. DFG-inter was the smallest group identified with 27 kinases that positioned the DFG-Phe side chain out of the C-helix but not to a fully extended DFG-out conformation. The majority of the structures showed the DFG-Phe to be orientated upwards and towards the β-sheets that separate the active site into two distinctive sections. This conformation has previously been

identified in Aurora A as 'DFG-up' by Dodson et al (55) and as a group of conformations termed 'ωCD' by Ung et al (56).

The three groups were then divided into subclasses dependent on the dihedral angles that the Phe side chain required for placement, which resulted in eight groups, six for DFG-in, one for DFG-out and one for DFG-inter. A nomenclature was developed based on the regions of the Ramachandran map that were occupied by the X, D, and F residues of the X-DFG motif. 'A' for alpha-helical region, 'B' for beta-sheet region and 'L' for left- handed helical region. The $\chi 1$ rotamer of the Phe side chain was termed 'minus' for the − 60° rotamer, 'plus' for the + 60° rotamer and 'trans' for the 180° rotamer. The active kinase conformation, the most abundant found in the Protein Data Bank (PDB), was termed as 'BLAminus' using this new nomenclature. All structures that were deemed catalytically active containing bound ATP and Mg^{2+}/Mn^{2+} and a phosphorylated activation loop belonged to the BLAminus subclass. Previously, all inactive DFG-in conformations were categorised together. However, this work identified the most prominent inactive DFG-in conformations were split into BLBplus and ABAminus groups. BLBplus was found to be the most abundant inactive conformation of kinases where the DFG-Phe ring is situated underneath the C-helix and is positioned upwards, resulting in the C-helix being orientated outwards. This orientation generates an extra region that could be exploited for design of new inhibitors (8).

Analysis performed by the Levinson group using Time resolved- Förster resonance energy transfer (TR-FRET) has identified conformational changes of Aurora A based on its phosphorylation status and ligand/inhibitor binding. Analysis revealed a variety of conformational preferences based on a DFG-in or DFG-out configuration, with each inhibitor adopting the configuration to a different extent. The DFG-in inhibitors favoured binding to Aurora A when it was constrained in the DFG-in configuration by the binding of the activating TPX2 peptide. DFG-out inhibitors favoured binding to Aurora A when it was phosphorylated on the activation loop. The addition of TPX2 showed a transition from DFG-out to DFG-in state for both forms of non-phosphorylated and phosphorylated Aurora A, with phosphorylated Aurora A being more prominently DFG-in compared to non-phosphorylated. Conformational changes were also evident upon inhibitor binding, with the largest shift seen with SNS-314, where an almost full transition to the DFG-in conformation was apparent for non-

phosphorylated and phosphorylated Aurora A. Transitions were also seen to the DFG-out state, with AMG-900 causing the largest shift in the phosphorylated + TPX2 form. The data concluded that three distinct groups of conformational changes were evident, including type I inhibitors that promote the DFG-in state, type I inhibitors that promote the DFG-out state and type II inhibitors (57).

1.3.2 Aurora inhibitors

Recently, many preclinical studies have been carried out on small molecules selectively targeting Aurora A for anticancer purposes including MK8745 (58), LY3295668 (59), BPR1KO609S1 (60), AKI603 (61), LDD970 (62) and CYC3 (63).

Alisterib (MLN8237) is a selective Aurora A kinase inhibitor which has been characterised using *in vitro* and *in vivo* pre-clinical models (64) demonstrating an IC_{50} value against Aurora A of 1.2 and 396.4 nM for Aurora B (65). The development of MLN8237 was due to its predecessor, MLN8054, showing central nervous system side effects in phase I clinical studies (66). MLN8237 has been shown to demonstrate antiproliferative activity in a number of human tumour cell lines, such as ovarian, prostate, lung and lymphoma cells. Additionally, MLN8237 has been shown to be effective in paediatric type cancers, including neuroblastoma, acute lymphoblastic leukaemia and Ewing sarcoma cell lines (65). MLN8237 has been assessed in phase I and II clinical trials for haematological malignancies, patients with advanced solid tumours and children with refractory/recurrent solid tumours. (67, 68). MLN8237 is the only Aurora A inhibitor that has been tested in phase III trials, showing antitumour responses and tolerance in patients with Peripheral T-Cell Lymphoma (69).
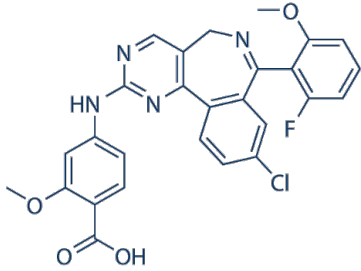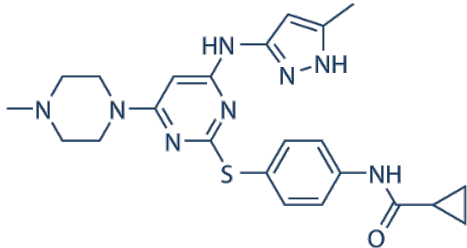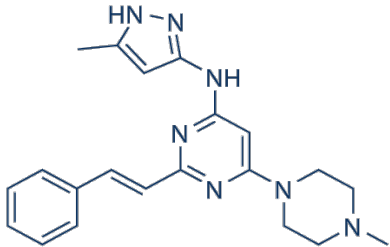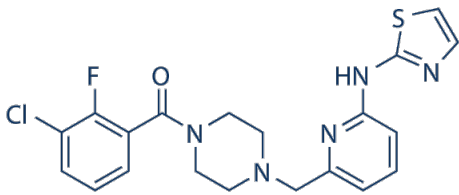
VX-680 has shown to inhibit Aurora kinases as an anti-cancer therapeutic and was the first Aurora inhibitor to enter clinical trials (70). VX-680 in cells inhibits mitotic Histone H3 phosphorylation, inhibits proliferation and induces polyploidy and apoptosis, whilst other cells remain unaffected (71, 72).

ENMD-2076, a DFG-in inhibitor (57), is a multitarget inhibitor with an IC_{50} value of 14 nM for Aurora A (73). It permits growth of a variety of human solid tumour and hematopoietic cancers, colorectal cancer, triple-negative breast cancer and myeloma (73-76). Many phase I

and II clinical trials have taken place with ENMD-2076 in a range of solid tumour and hematologic malignancies studies (77, 78).

MK8745, a complete DFG-out (57) selective Aurora A inhibitor has over a 100-fold selectivity for Aurora A (IC_{50} of 0.6 nM) in comparison to Aurora B (IC_{50} of 280 nM) and other kinases (58).

Table 1.1. Aurora A kinase inhibitors.

Inhibitor	Mass (Da)	Structure
MLN8237 (Alisertib)	518.92	
VX680 (Tozasertib)	464.59	
ENMD2076	375.47	
MK8745	431.91	

1.4   Non-canonical phosphorylation

In addition to phosphorylation of serine, threonine and tyrosine, many other amino acids are now thought to play a role in mammalian protein signalling, namely His, Arg, Asp, Cys, Glu and Lys, which are all termed as non-canonical phosphorylation (79). An estimated total number of 748 histidine kinases have been predicted from searches of the genomes of 67 eukaryotic species from the phylogenetic tree (80). Phosphohistidine (pHis) was initially identified in the 1960s as part of an oxidation reaction of an enzyme intermediate (81). Generic and site-specific antibodies have been identified against pHis and been used to demonstrate direct and indirect roles in T-cell signalling, cell proliferation, differentiation, and migration (82).

1.4.1 Two component signalling system

The most studied histidine phosphorylation has been in bacteria, fungi, and plants due to their involvement in the two-component signalling system (TCS), which were first reported in 1980. To date, there are no known function homologues of TCS in vertebrates. TCS are understood to play a crucial role in cellular responses to many environmental factors (83). A TCS is composed of a sensor histidine kinase and a response regulator, with both aspects associated with connecting a cellular or environmental signal with a suitable response. This is carried out by communication through a phosphoryl-group transfer from a histidine on the histidine kinase to an aspartate on the response regulator. In addition, histidine kinases are able to function as a phosphatase for the response regulator (84).

1.4.2 Histidine phosphorylation

The phosphorus of the phosphoryl group in phosphohistidine binds can bind to either the 1- or 3- position of the imidazole ring of histidine, forming 1- or 3-phosphohistidine (pHis) respectively through a phosphoramidate bond (Figure 1.8) (85). The N1 pHis isomer is thought to be less stable in comparison to N3 pHis due to N1 pHis being more electrophilic (86). The analysis of pHis is very challenging due to the free energy of hydrolysis of the phosphoramidate (N-P) bond (87). This results in the subsequent loss of the phosphate group at low pH and high temperatures (88). Many protein phosphorylation assays require acidic treatments; therefore, the analysis of pHis requires some method adjustments. Studies have

carried out the detection of proteins containing the phosphoramidate bond using high performance liquid chromatography and thin layer electrophoresis (89, 90). Mass spectrometry has previously been exploited for the detection of synthetic and enzymatic derived pHis (79, 91, 92).
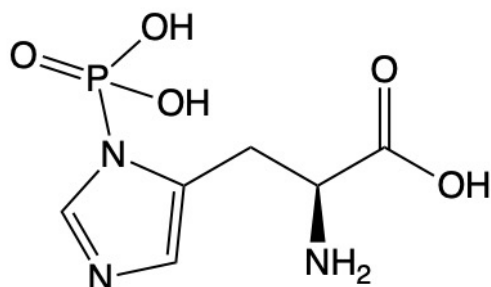
N1-phosphohistidine (N1-pHis)   N3-phosphohistidine (N3-pHis)
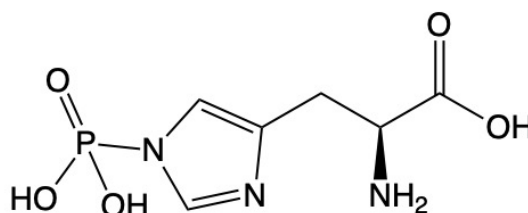
Figure 1.8. Two pHis isomers found in physiological conditions, N1 phosphohistidine (N1-pHis) and N3 phosphohistidine (N3-pHis). Structures were produced in ChemDraw 3.1.

1.4.3 Nucleoside diphosphate kinases

NME1 and NME2, both known mammalian histidine kinases, belong to the family of nucleoside diphosphate kinases (NDPKs), which are a group of proteins encoded by the *nme* (non-metastatic cells) genes. Metastasis suppressor genes inhibit the growth of biological processes during the progression of metastases, whilst not affecting the primary tumour development (93). In addition to their metastasis suppressor functions, NME1 and NME2 have shown to be associated with proliferation and differentiation. NME proteins use pHis as an enzyme intermediate to carry out a catalytic reaction (94). There are ten known human NME proteins, with NME1 the first of the family of proteins to demonstrate the ability to supress the metastatic phenotype of cancer cells (95). NME1 and NME2 share 88% sequence similarity and have both been shown to demonstrate histidine kinase activity auto-phosphorylating at H118 (96). The different isoforms of NME are associated with different subcellular localisations including being found in the nucleus (NME1/2), mitochondria (NME3/4) and cytoplasm (NME1/2) (97).

NDPKs play a role in catalysing the transfer of *y*-phosphate from nucleoside triphosphate to nucleoside diphosphate, which results in the generation of a high energy phosphohistidine

intermediate (Figure 1.9) (85). NDPKs play an important role in maintaining the balance between intracellular NTP and dNTP pools, which are required for essential cellular processes, such as DNA replication, protein translation and RNA synthesis (97). The nucleotide base has been reported to position near the protein surface, which is stabilised by the Phe60 residue on the helix and Val112 on the kpn loop to create a clamp structure. The ribose is located in the nucleotide binding cleft of the protein, with K12 and N115 forming hydrogen bonds between the 2' and 3' hydroxyl groups and the phosphate moiety orientated towards the His118 residue. His118 has previously been shown to be stabilised by Glu129, with a hydrogen bond between N3 of the His118 imidazole group resulting in the position of N1 being available for phosphate transfer (98).
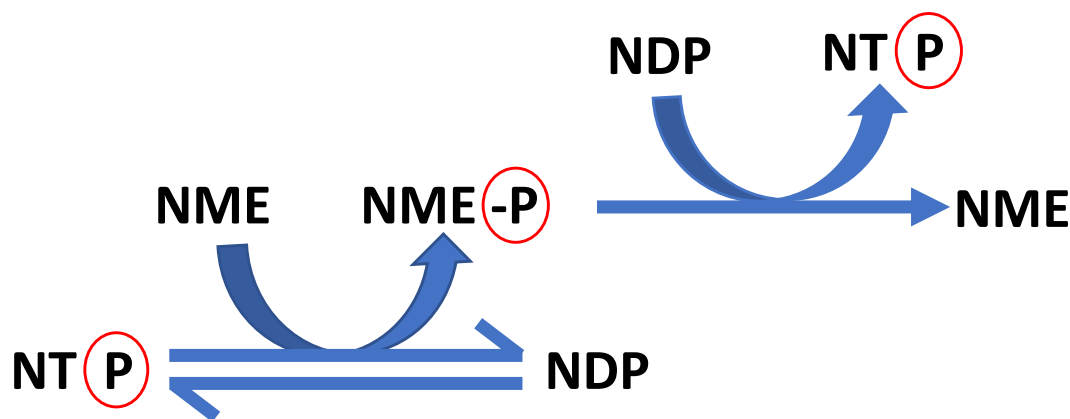
Figure 1.9. Schematic diagram of the catalytic mechanism of NDPKs. Phosphate transfer is carried out via a ping-pong mechanism.

1.4.4 NME structure and multimerisation

NME2 was the first structure identified of the human NME proteins and revealed that the protein adopts a hexameric configuration of a dimer of trimers or a trimer of dimers (Figure 1.10). Each monomeric subunit mass is 17 kDa, folding into an alpha/beta domain around an anti-parallel ß sheet, with a C-terminal extension. The high sequence similarity between the NME protein family suggests that other NME proteins would be structurally similar to NME2, with a hexamer formation. NME oligomerisation is thought to be regulated though redox-dependent mechanisms, with oxidising conditions disrupting the hexameric structure. Upon oxidising conditions, NME1 has been shown to form an intramolecular disulphide bond between Cys4 and Cys145. This disulphide bond formation disrupts the protein interface

leading to a conformational change in the C-terminal region, which destabilises the hexamer. The resultant dimer exhibits reduced enzymatic activity.
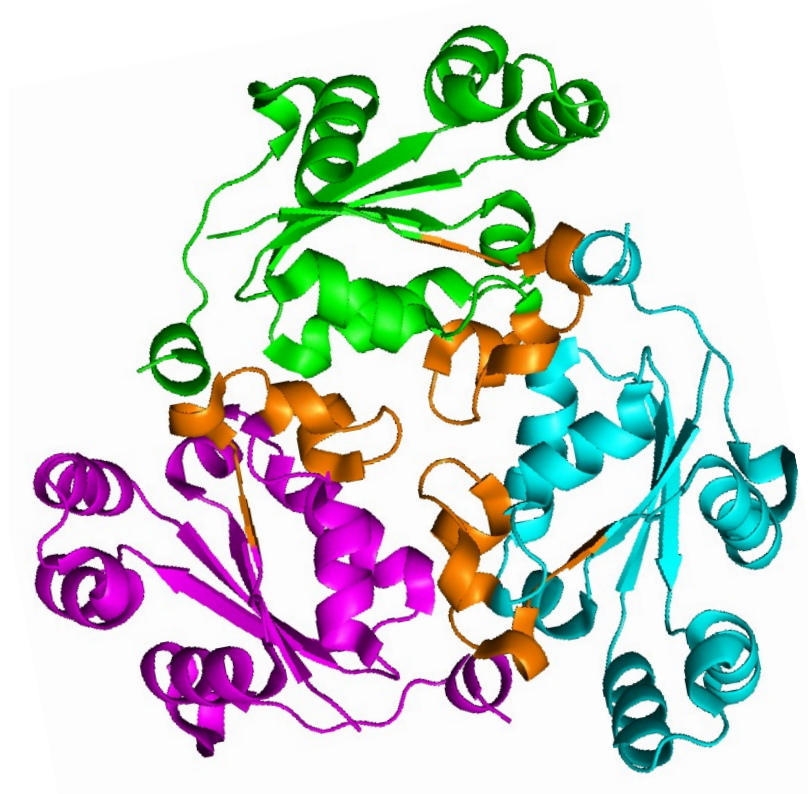
Figure 1.10. Structure of NME1 trimer. Three monomers of NME1 shown in green, magenta and cyan. Kpn loop shown in orange. Ribbon diagram produced in PyMol from PDB file 2HVD.

1.5   Mass spectrometry proteomics

Proteomics is defined as the complete characterisation of all the proteins expressed in a cell, tissue, or organism (99). An important application of proteomics is the study of PTMs, as protein kinases function by signal transduction and phosphorylation, it is crucial that PTMs are analysed to provide interpretation of kinase activity within the cell (100). Mass spectrometry-based proteomics predominately relies on protein enzymatic digestion, peptide ionisation, separation of ions based on their *m/z* ratio and then ion detection (101), an approach termed 'bottom-up' proteomics. Peptides can be analysed by mass spectrometry using a variety of ionisation forms including electrospray ionisation (ESI) and matrix-assisted laser desorption/ionisation (MALDI), with liquid chromatography MS (LC-MS) commonly used for peptide separation in-line with (ESI) (102).

1.5.1 Tandem mass spectrometry

MS-based proteomics often uses reverse-phase liquid chromatography (LC) prior to MS to facilitate separation of peptide analytes from a complex mixture for analysis. Peptides are commonly bound to a C18 column, which is packed with silica particles that contain alkyl chains. Elution of peptides occurs with increased concentration of organic solvent and separation of peptides is based on their hydrophobicity (103). On elution from the LC column, peptides are introduced into the ion source of the MS instrument, where the generation of gaseous ions takes place. In a tandem MS experiment, the *m/z* ratio of all ions entering the ion source are recorded in the MS1 scan. In a data-dependent acquisition (DDA) mode, a TOP*N* method is implemented to analyse the ions with the most intensity from the MS1 scan. The ion of interest is then isolated and fragmented to produce a MS2 (MS/MS) spectrum of the product ions (104) which can be used to determine peptide (and thus protein) sequence.

1.5.2 Electrospray ionisation

Electrospray ionisation (ESI) is a soft ionisation technique that transfers ions from solution to the gas phase prior to mass spectrometry analysis. During ESI, the sample in solution is sprayed from the capillary, which results in the generation of a charged droplet of the same polarity as the capillary voltage. The droplets reach higher temperature on entry to the ESI source, which leads to evaporation of the solvent and the reduction in size of the charged

droplets (Figure 1.11). The electric field strength of the charged droplets reaches an equilibrium where it is kinetically and energetically stable for ions positioned at the droplet surface to be transferred into the gas phase. NanoESI, where ESI is performed at flows rates between 200 and 1000 nl/min with the use of a smaller diameter emitter in the range of 10 to 100 microm id, enables smaller droplets to be produced, which results in more efficient ionisation. The reduced droplet size reduces the number of salt ions in each droplet, which results in less adducts on the protein following solvent evaporation and better spraying capabilities. The advantages of nanoESI include higher sensitivity, decreased competition for ionisation and an enhanced dynamic range (105).
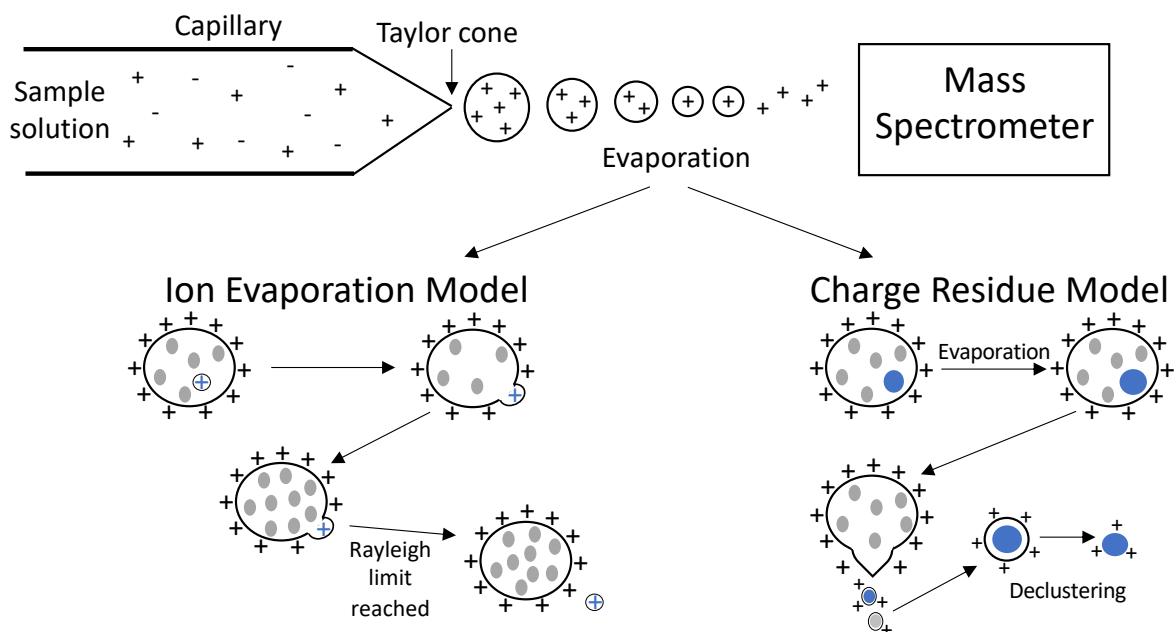
Figure 1.11. Schematic of the electrospray ionisation process. Sample solution exits the capillary tip with the assistance of an electric field in form of a Taylor cone. Charged droplets emerge from the tip and travel through the electric field to the mass spectrometer. Two models described: Ion Evaporation Model shows how smaller ions enter the gas phase and Charge Residue Model shows how larger ions enter the gas phase. Red circles demonstrate proteins and grey circles demonstrate peptides in the droplet. Figure adapted from (106).

1.5.3 Mass analysers

The mass analyser is the component of the mass spectrometer that enables the main goal of the analysis to be achieved, which is the mass determination of the analyte in question. A range of mass analysers are available that can also trap and store ions in addition to resolving ions of different *m/z*. Most common analysers include quadrupole, Time-of-Flight (ToF) and Fourier transform, with many setups employing different combinations or hybrids of these

analyser types. A quadrupole analyser acts as a mass filter, with four hyperbolic rods that run in parallel to each other. The rods that are situated opposite one another are electrically connected, with a radio frequency (RF) potential and direct current (DC) applied the rods. The combination of both the RF and DC potentials enable the ions to oscillate once they travel through in the quadrupole in the z-direction. Ions of a specific *m/z* will have stable trajectories, whereas ions not in the specific *m/z* range with have instable trajectories resulting in their collision with the rods and them being filtered out. The variation in the RF and DC potentials enable ions of different *m/z* regions to be scanned and filtered through the quadrupole. A ToF mass analyser consists of a flight tube and an acceleration grid that pushes a series of ions from the ionisation source to the MS detector. The basic understanding of the ToF analyser is if two ions of different *m/z* values are accelerated from the ion source with the same energy, the arrival times at the detector will be different for each of the ions. Ions with larger mass will travel slower through the flight tube in comparison to smaller mass ions. Once the ion hits the detector, the mass spectrometer will record the time it took for the ion to travel through the flight tube, with the drift time through the tube proportional to the *m/z* of the ion (107).

1.5.4 Higher-energy collisional dissociation (HCD)

Characterisation of peptide sequence, either manually, or using algorithms that permit protein database search, requires peptide fragmentation. Higher-energy collisional dissociation (HCD) in a Q-Exactive mass spectrometer utilises an octopole collision cell, which can function at higher collisional energies compared to collision induced dissociation (CID) (108). The high RF voltage of the collision cell enables high resolution mass analysis and no lower mass cut-off, resulting in increased protein sequence coverage. HCD enables higher energy dissociations compared to ion trap CID resulting in the analysis of further fragmentation pathways. One limitation of using HCD in an Orbitrap instrument is the length of time it takes to acquire spectra due to the increased number of ions that are needed for Fourier transform detection. (109, 110).

1.5.5 Electron transfer dissociation (ETD)

The application of CID and HCD for the identification of modified peptides with labile PTMs can lead to issues regarding site localisation of the modification. In 2004, electron transfer

dissociation (ETD) was reported as a new technique to fragment gaseous ions. ETD implements the use of reagent anions for ion-ion reactions, which then leads to electron transfer to the peptide cation. This reaction results in the peptide cleavage at the N-Cα bond and subsequently a series of c (N-terminal peptide ion fragments) and z (C-terminal peptide ion fragments) ions (111, 112). ETD requires the generation of reagent anions, which is commonly carried out with the addition of fluoranthene into the chemical ionisation source. The source can generate near-thermal electrons which the fluoranthene can capture to produce a radical anion. ETD fragmentation does not result in energy distribution throughout the ion, which enables peptides to retain a labile PTM, such as a phospho moiety, and therefore, leads to the improvement of site localisation of the phosphorylated site. The non-energetic reaction used for ETD results in the limitation to only multiply charged species (3+), which is due to the decrease in electron-induced charge and therefore, its ability to fragment a singly charged ion (113). More recent advances have implemented the use of dual fragmentation strategies that consist of HCD and ETD, referred to as EThcD. Fragmentation during EThcD results in a mixture of b, y, c and z ions, which results in additional information for protein sequence determination, whilst also maintaining the identification of any PTM's (114).

1.5.6 Thermo Orbitrap Fusion

Advances in the field of mass spectrometry have led to the use of tribrid mass spectrometers. The Thermo Orbitrap Fusion is a tribrid instrument that utilises three mass analysers; a quadrupole mass filter, Orbitrap, and Ion Trap (Figure 1.12), which enables MS1, MS2 and MS^n analysis in the Orbitrap and Ion Trap (115).
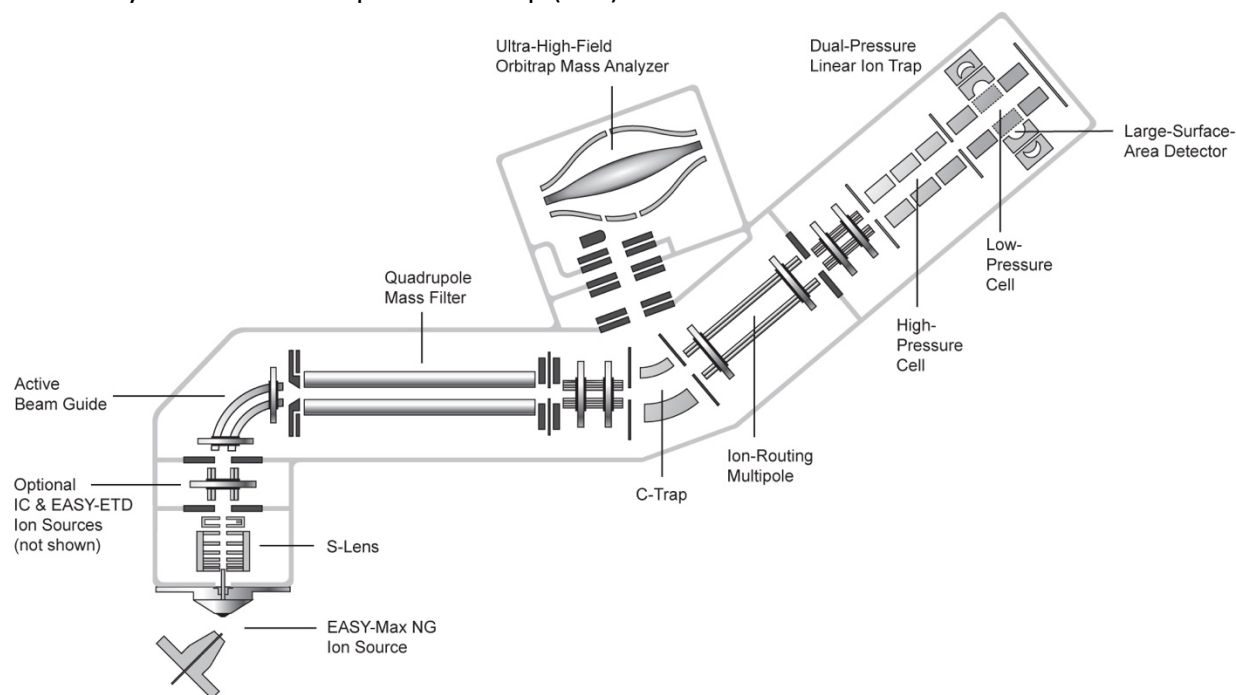
Figure 1.12. Schematic of the Thermo Orbitrap Fusion Tribrid mass spectrometer. The Orbitrap Fusion Tribrid mass spectrometer consists of a quadruple mass filter and both an Ion Trap and Orbitrap mass analyser. HCD can be applied in the Ion-Routing Multipole and ETD in the linear Ion Trap of the instrument.

1.5.7 Orbitrap

Prior to entry in the Orbitrap, ions accumulate in the C-Trap, which is an RF quadrupole consisting of four hyperbolical rods situated perpendicular to the Orbitrap. The Orbitrap mass analyser consists of an outer barrel shaped electrode and an inner spindle-like pole electrode (Figure 1.13). A linear electric field is applied between the two electrodes that results in ions being trapped around the inner spindle and m/z values are measured by the frequency of harmonic ion oscillations (116). The oscillation frequencies are detected via the measurement of an image current from the outer barrel shaped electrodes, which is then followed by Fourier transform for conversion of the recorded time-domain signal to a m/z spectrum (117). The Fourier transform breaks down the original time-based waveform into several sinusoidal forms, which all have an individual magnitude, frequency, and phase. This results in the

conversion of a waveform in the time domain to a series of sinusoidal equations. The spectrum produced represents the frequency on one axis and the magnitude on the other axis (118).
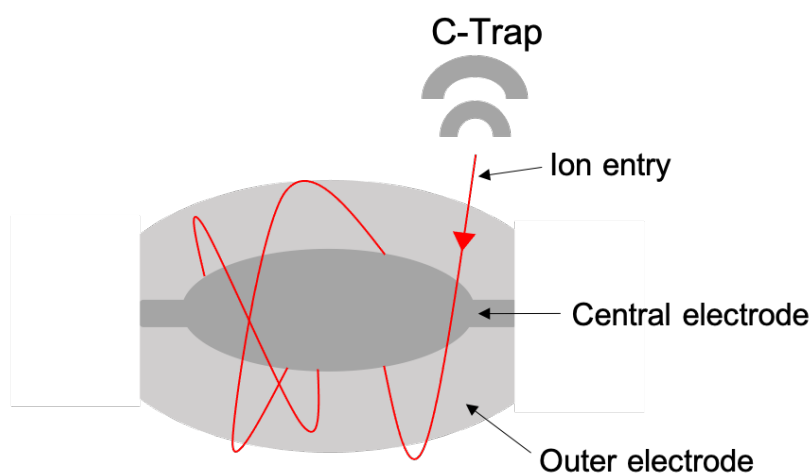
Figure 1.13. Orbitrap mass analyser. Ions are transported from the C-Trap into the Orbitrap. An electric field is applied between the central and outer electrodes to allow ions to oscillate with the frequency of oscillations proportional to the *m/z* output.

1.5.8 Ion Trap

Ion Trap mass analysers accumulate ions in an oscillating field using two different types: a 3D (Paul) trap or a 2D (linear) trap, with the latter used in the Orbitrap Fusion instrument (Figure 1.14). The 3D trap consists of a hyperbolic ring electrode and two hyperboloidal end-cap electrodes, whereas the 2D trap consists of four hyperbolic or cylinder-shaped rod electrodes (119). Ion trap mass analysers use a helium gas environment to fragment peptides by the application of an auxiliary excitation frequency, which increases the motion of the precursor ion. The induced motion results in the increase of further energetic collisions within the gas leading to increased vibrational excitation of the precursor ion. Following the breakage of bonds, the fragment ions can no longer be excited by the auxiliary frequency. For ions to be ejected from the trap, the RF voltage at one end of the electrode is altered to force the ions to be instable and subsequently be ejected from the trap and into the mass detector (120). A 2D trap has a higher capacity for trapping ions and is more efficient at externally injecting ions due to its larger trapping area in comparison to a 3D trap. Ions entering the Ion Trap must be under control to prevent space-charging effects due to ion repulsion when the level of ions present increase too much, which results in a decrease in resolving power performance and mass accuracy (119).
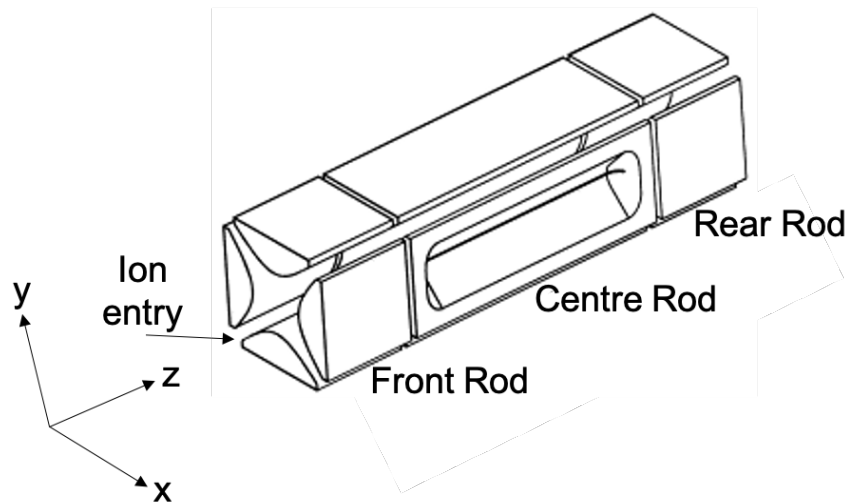
Figure 1.14. 2D Ion Trap mass analyser. 2D linear ion trap consisting of four hyperbolic rod electrodes. Application of varying RF and DC voltages enables the ions to oscillate in the field and travel towards the mass detector. Figure adapted from (121).

1.5.9 Proteomics data analysis

Following the acquisition of MS/MS data, search engines such as MASCOT can be utilised to identify peptides and subsequently, proteins for high throughput, rather that manual, data interrogation. Search engines compare MS1 and MS2 experimental data to theoretical values of peptide or fragment masses obtained from searching a database containing the proteins of interests. Identified proteins are scored to determine the peptides with the highest matches against the inputted database (122). False positive PSMs (peptide spectral matches) assignments can occur during the database search algorithm, therefore, the false discovery rate (FDR) of peptides needs to be determined. A popular way of determining FDR is to search a decoy database, which can be created by reversing the target database, such that the composition (in terms of amino acid content and distribution) of the two are identical. The number of matches to the decoy database allow computation of the number of false positives within the data set (123). Additional tools can be incorporated into the search parameters of MASCOT and Proteome Discoverer, such as phosphoRS, which gives a score based on the prediction of the PTM site localisation by the analysis of specific product ions (124).

1.6   Native mass spectrometry

Native mass spectrometry (MS) is a form of electrospray ionisation where analytes are sprayed from a non-denaturing solvent. The term 'native' describes the biological state of the analyte in solution, prior to the ionisation step during MS (125). This technique enables large biomolecules and complexes to be transferred from a liquid phase to gas phase, whilst experimental conditions are mild enough to enable the preservation of non-covalent interactions and the proteins native state (126).

1.6.1 Ion-Mobility Mass Spectrometry

Ion-Mobility (IM) MS separates ions on the basis of their charge and shape. Therefore, information such as collision cross-sections (CCS) of proteins in the gas phase can be obtained and compared to those obtained from other structural biology techniques, including x-ray crystallography (125). IM-MS has been emerging as a popular technique used to gather information regarding conformational dynamics, ligand-induced conformational changes, and folding/unfolding intermediates. IM-MS of smaller charge states would commonly relate to the more compact protein structure, as a smaller number of exposed protonation sites are available, which is more favourable to represent the native folded protein (127). Proteins that are less than 100 kDa with fewer than five charge states are often associated with a tightly configured structure, whereas more than five charge states are likely representative of an unfolded protein, which can be due from spraying under denaturing conditions or intrinsic disorder. The charge states of the protein can be manipulated to produce lower charge states and reduce the chance of coulombic repulsion or unfolding. This can be achieved by charge stripping with the addition of a base into the desolvation region (128).

In comparison to other structural techniques, IM-MS only requires a small amount of sample in the nanogram quantities (129). Native MS coupled with nESI enables the use of nL/min flow rates, which produces smaller droplets than standard ESI, which allows for lower sprayer voltage and source temperatures, and therefore, maintains the native protein structure in the gas phase (130). Other techniques, such as cryo-electron microscopy, require the protein to be above 50 kDa, whereas IM-MS does not have this limitation with the analysis of smaller proteins. Cryo-EM also has the challenge of needing higher quantities of material and the difficulty in embedding small particles into ice correctly. However, the advantage of cryo-EM

is the ability to gain information regarding the 3D structure, whilst the sample is in solution and preserving the native state, which can enable the visualisation of how molecules move and interact when performing different functions (131, 132). Additional structural techniques including x-ray crystallography and NMR spectroscopy require the protein sample to be in its purest form, however, native MS can be used to explore the protein heterogeneity of the sample (125). X-ray crystallography does have the advantage of higher resolution, where mechanisms of action and complex interactions can be observed. However, the solid structure is not representative of the native form, which does not enable different conformations to be identified. The technique can also result in protein precipitation, and it is a very time-consuming method to carry out (133). NMR has the ability to probe protein dynamics and observe multiple conformations, including the ratios in how the conformers can flip from one to another, whilst the protein is in solution and its native state. In addition, protein interactions can be detected using chemical shift mapping or chemical shift titration. The main drawback of using NMR is its limit to smaller proteins below the range of 30 kDa due to the slower tumbling rate of larger proteins (134).

1.6.2 Protein solution to gas phase considerations

Prior to entry into the mass spectrometer vacuum, proteins need to be desolvated and ionised, which gives rise to the question of whether the ionised species in the gas phase are a representation to those found in solution (135). Studies have shown that it is possible to partially retain the structures and topology found in solution and observe these findings when carrying out gas phase experiments (136, 137). It is of high importance in native MS to control parameters of pH and ionic strength in order to sustain the folded native state of the protein or analyte in solution (125). The sample and spraying conditions that are adopted can be modified with solvents that stabilise structures, such as ammonium acetate, which enables the retention of protein interactions and non-covalent protein complexes (126). When transferring a protein to the gas phase from solution, a series of events can occur over a time period, including the collapse of charged side chains (picosecond time-scale), and a decrease of hydrophobic interaction and increase in hydrogen bonding (millisecond time-scale) (138). Proteins can lose their native structural fold over longer time periods, with factors such as protein charge state playing a role. For example, experiments have shown that cytochrome c can lose its native folded structure at 400 milliseconds, however, IM-MS is capable of

detecting these solution-like structural features within 1-2 milliseconds, prior to the protein beginning to unfold (139). In order to maintain the native structures of proteins, the temperature and energy must be carefully observed, to not induce any unwanted changes on the native protein (140).

1.6.3 IM-MS instrumentation

The aim of IMS instrumentation is to separate a series of analyte ions through a drift tube filled with buffer gas under the influence of an electric field. The relationship between the drift velocity of the ions and the electric field is influenced by the mobility (K) of the ion in the chosen buffer or gas. Smaller ions will travel faster through the electric field in comparison to larger ions. The instrument parameters, including the gas temperature and pressure can impact the rate of K. The separation of analyte ions is also influenced by the physicochemical properties of the ion and gas, such as the ion and gas masses, ion charge state and the CCS value (141).

There are three primary forms of IM-MS: drift-tube ion mobility spectrometry (DTIMS), travelling-wave ion mobility spectrometry (TWIMS), and field-asymmetric ion mobility spectrometry (FAIMS) (142).

In DTIMS, the duration of time an ion takes to pass through the drift tube relates to its mass, charge, and rotationally averaged CCS, with CCS determination possible from measured experimental parameters. Ions are trapped following ionisation and prior to their entry into the drift cell via an ion gate. The ions are then introduced into the mobility device, which is filled with a neutral gas buffer. Ions are transported through the device under weak electric field conditions. The drift time associated with each ion can then be used to calculate Ω based on the Mason-Shamp equation (143):

$$\Omega = \frac{3ze}{16N} \left(\frac{2\pi}{\mu k_B T}\right)^{1/2} \frac{1}{K0}$$

where K0 is the reduced mobility, z is the ion charge state, e is the elementary charge, N is the number density of the drift gas, μ is the reduced mass of the ion-neutral drift gas pair, kB is the Boltzmann constant and T is the temperature of the gas.

The drift cell aspect of the TWIMS device is similar to the DTIMS, with a stacked ring of electrodes that apply a voltage to enable the ions to travel through the drift cell at a low pressure of ca. 2-4 Torr. A difference between the two devices is that DTIMS has a uniform electric field and the TWIMS has an oscillating electric field, which induces waves that enable the ions to be propelled through the device to the mass analyser. The application of alternative sections of positive and zero electric fields enables the creation of the travelling wave for the ions to propel through to the detector. Before ions enter the ion mobility device, they are stored in the trap ion guide, which allows the ions to be transferred to the ion mobility cell at the same time. Ions can travel on the top of the waves or travel through the waves depending on the interactions between the ions and the drift gas, therefore ions with higher mobility will travel quicker and reach the detector sooner (Figure 1.15). (142).
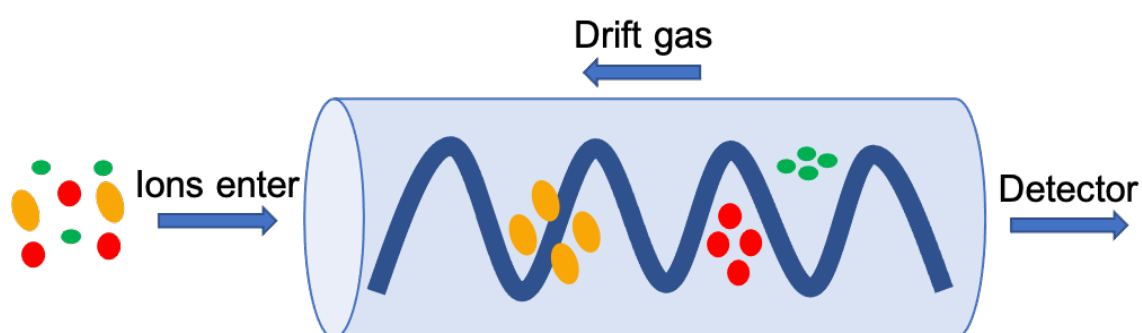
Figure 1.15. Travelling Wave Ion Mobility (TWIMS) device. The drift tube is filled with buffer gas and an electric field is applied to enable ions to travel through the tube in a wave formation. Ions enter the TWIMS device with the duration of time spent in the TWIMS dependant on the size of each ion. Smaller, compact ions, travel over the wave and have smaller CCS, whereas larger ions will travel through the wave, taking longer to exit and have larger CCS.

The use of IM-MS as an analytical technique was commercialised in 2006 with the arrival of the Waters Synapt HDMS instrument and then followed by the introduction of the Synapt G2 in 2011 and the Synapt G2Si (Figure 1.16) in 2013.

The instrumentation of the Synapt G2Si includes a transfer ion guide that focuses the separated ions prior to their analysis in the quadrupole analyser. The quadrupole is used as a mass filter step prior to separation in the ion mobility device. A quadrupole consists of four hyperbolic rods, which are charged with direct current (DC) and radio frequency (RF) voltages. A stable DC is applied to the rods and the rods situated adjacently are applied with voltages of opposite polarity, whilst the fluctuation of RF voltages enables the transition of polarity of the rods. The oscillation of DC potentials and RF frequencies creates an electric field for the ions to be transported through the quadrupole. Ions in a selected *m/z* range have a stable trajectory enabling them to travel to the detector, whereas ions outside the *m/z* range will be unstable and collide with the rods leading to their expulsion from the analyser (144).

In addition to the quadrupole analyser, the Synapt G2Si also consists of a Time-of-Flight (ToF) analyser. A ToF mass analyser is based on the principle that ions with different masses, but the same energy, will travel at different velocities. Ions are accelerated by an electric field and the *m/z* of an ion is determined based on the duration of time it takes to travel from the electric pusher to the detector, with lighter ions arriving prior to the heavier ions (145).
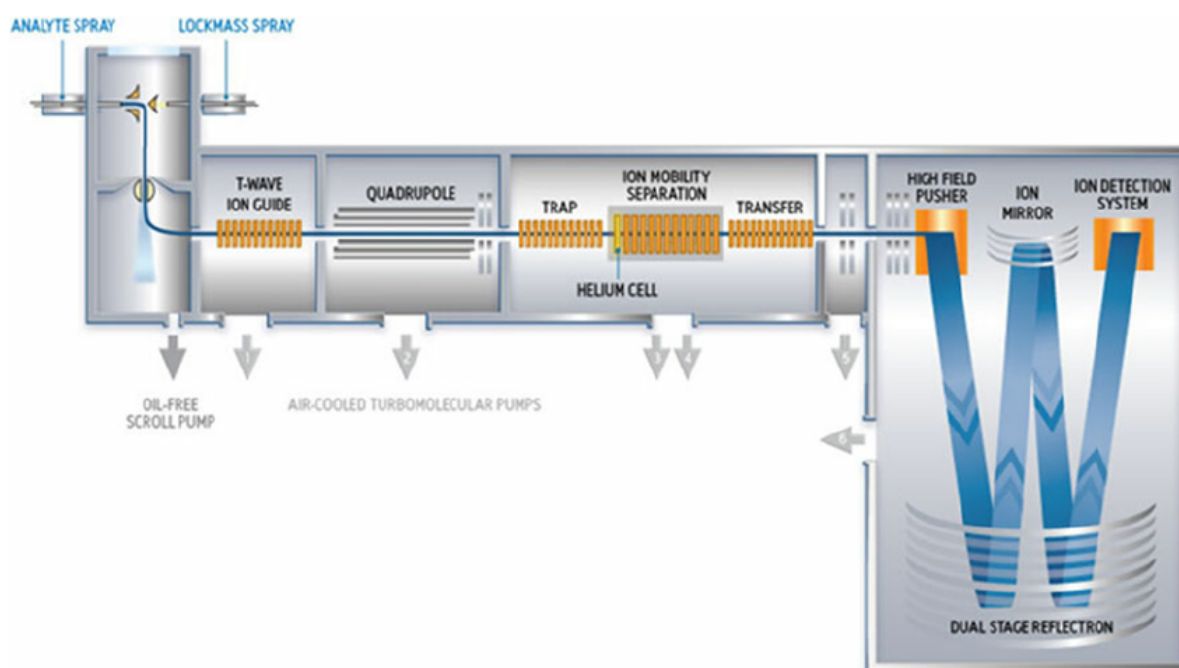
Figure 1.16. Schematic of the Waters Synapt G2Si mass spectrometer. The Synapt G2Si consists of a quadruple mass analyser up to 8000 *m/z* and a Travelling Wave Ion Mobility cell device. Collisional activation voltage can be applied in the Trap region of the instrument.

1.6.4 Collision Cross Section (CCS)

A TWIMS device cannot be used to directly determine CCS in contrast to DTIMS due to the continuous alteration within the electric field and the association between Ω and K0 no longer being related. The drift time through the travelling-wave mobility cell requires calibration of the cell to determine CCS. The TWIMS cell is calibrated for drift time under defined conditions such as gas type/pressure and travelling wave speed/height using ions similar in nature to the analyte, of known CCS value. Drift time calibration is carried out with the analysis of analytes that possess the same chemical or physical properties as the analyte in question (146).

CCS values can provide information on the conformational shape of an ion in the gas phase during defined conditions, which can then be compared if needed with theoretical CCS values obtained from other structural techniques, such as x-ray crystallography. The CCS value represents an average of all the geometric orientations and interaction types within the measured time of the experiment. The value is influenced by the drift gas due to the momentum transfer and gas polarisation abilities, and the influencing factors of the IM-MS experiment including the magnitude of the electric field and the temperature. CCS values provide an additional physicochemical measurement that can help complement accurate mass data for protein identification (147). CCS measurements alongside *m/z* ratio and retention time can enhance identification purposes when in used in combination with databases that contain known CCS values and accurate mass measurements. The analysis of CCS values have additionally been shown to be consistent between instruments with a range of experimental conditions tested and results have shown CCS to be more reliable than retention time values. However, one limitation of using CCS values as an identification tool could be the lack of databases that provide CCS measurements for contaminants (148).

The separation of an ion during an IM-MS experiment is represented by a distribution of arrival time, as opposed to a single peak of the arrival time. The distribution is a cause of ion packet diffusion and the presence of multiple interconnecting conformer species. CCS values are commonly determined from the apex of the main arrival time distribution peak, which results in a single CCS value for each protein. However, the CCS determination of each individual conformational state can provide additional information alongside molecular

modelling at interpreting the structural properties that each conformer is associated with (149).

Molecular modelling algorithms can be implemented to determine CCS values by inputting structural information obtained from NMR or x-ray structures (150). CCS values calculated from high-resolution coordinates and compared to experimentally derived CCS from TWIMS instruments have shown to have an exceptional correlation with an error of <3% for globular proteins. However, a significant challenge in carrying out these calculations is the reliability of determining CCS from atomic coordinates that provide values from $>10^6$ models in the >100 kDa range. This is due to larger proteins and protein complexes often resulting in complications during the calculation due to difficulty in determining the inter and intramolecular interactions (151).

There are a range of methods available to calculate CCS that use Monte Carlo integrations that consist of 'probes' to represent the IM gas, which are then 'fired' upon a randomly orientated 'target', which is the structure of interest (151). The algorithms vary with the assumptions and approximations that they make with regards to the collisions between the probe and the target. There are three commonly used methods to determine CCS values that use helium as their buffer gas; projection approximation (PA), the trajectory method (TM) and exact hard-sphere scattering (EHSS) (138).

The PA method is the simplest method, which calculates the CCS based on the average projected area of the target (protein structure) in relation to the size of the IM gas probes. The molecule is randomly rotated several times during the calculation to enable the rotational space to be assessed and the average projected area is then calculated based on the Monte Carlo integration. As the PA method doesn't take into account the scattering and long-range interactions, calculations are quick, however, the speed in which the calculation is completed can lead to an underestimation of the CCS. This method is predominately used for ions below 2 kDa due to its inability to consider multiple collisions (151).

The EHSS method calculates CCS by averaging the motion of the transfer cross section, which relates to the scattering angles between the arriving and departing gas atom trajectories.

EHSS does factor in the scattering and collisions with the drift gas, however, in similarity to PA, it does not factor in long-range interactions with the buffer gas and molecular ion (138).

The TM, the most complex and reliable, considers both longer interactions and collisions between the ions and buffer gas atoms, which makes it a more suitable algorithm to analyse larger biomolecules including proteins. In order to utilise this method, the effective potential must be determined, which then enables trajectories to be calculated within this specific potential to gain the scattering angles. The disadvantage of performing TM is the longer time required due to the analysis of the longer interactions and the requirement of the integration of forces to calculate the trajectory of the probes makes this model more expensive (138).

IMPACT (Ion Mobility Projection Approximation Calculation Tool) is an algorithm that can analyse models from x-ray crystallography, nuclear magnetic resonance spectroscopy, electron microscopy and x-ray scattering to calculate CCS values 10^6 quicker than previous models without compromising on accuracy. This method has previously been shown to calculate CCS values of all the proteins in the European PDB in a few hours, providing information on how similar a structure is to others of a comparable mass. IMPACT enables the calculation of quick CCSs during molecular dynamic simulations, obtaining information that corresponds to the gyration radius of the molecule, which makes it a useful tool for molecular modelling purposes (151).

1.6.5 Collision-Induced Unfolding

Collision activation of the ions can be performed in the trap and transfer domains of the TWIMS device to gain further structural analysis of protein stability and structure (Figure 1.17). Collision-induced unfolding (CIU) involves the isolation of a precursor ion and subjecting it to energetic collisions with a background gas. The collisions must be high enough to increase the internal energy of the analyte in question to induce a conformational change associated with the protein unfolding in the gas phase. Collision energy too high will induce too much energy and will cause disruption of the significant covalent bond interactions within the protein. Arrival time distributions (ATDs) are acquired during each voltage interval, with changes in ATD corresponding to different structural transitions of the protein. Quantitative tools can then be carried out to determine binding effects and the related structural changes

upon ligand binding (152). The development of CIU has enabled the technique to be used as an analytical fingerprint tool to study protein stabilities and protein-ligand complexes (153). CIU can assess partially unfolded intermediates in the millisecond time frame, providing information regarding the structure and stability of isolated protein complexes (154).
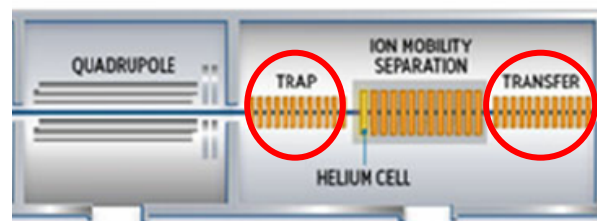
Figure 1.17. Collision-induced unfolding in a TWIMS device. Specific ion is isolated in the quadrupole and then subject to energetic collisions in either the trap or transfer regions of the TWIMS device.

Chapter 2. Materials and Methods

2.1   Reagents

General lab consumables were purchased from Sigma-Aldrich unless otherwise stated. Details of Aurora A inhibitors and TPX2 peptides are shown in Table 2.1.

Table 2.1. Inhibitor and peptide purchasing information.

Inhibitor/Peptide	Catalogue number	Purchased from
MLN8237	S1133	SelleckChem
VX-680	S1048	
ENMD-2076	S1181	
MK-8745	S7075	
Staurosporine	S1421	
TPX2 (MSQVKSSYSYDAPSDFINFSSLDDEGDTQNIDSWFEEKANLEN)	N/A	Pepceuticals
TPX2 modified peptide (MSQVKSSASAAAPSDFINFSSLDDEGDTQNIDSWFEEKANLEN)		

2.2   Plasmid purification

2.2.1 Escherichia coli competent cells

Glycerol stocks of TOP10 and BL21(DE3) pLysS *E. coli* were inoculated on Luria-Bertani (LB) agar plates and incubated for 18 hours at 37°C. A single colony was picked from each plate and grown in 50 mL of LB media for 18 hours at 37°C in a shaking incubator (250 RPM). A total volume of 2 mL was added to 200 mL of new LB media until an OD_{600nm} of 0.6 was reached. Cells were centrifuged to attain a pellet for 10 minutes at 4°C (5000 x *g*). The pellet was lysed on ice in buffer 1 ((100 mM rubidium chloride, 50 mM manganese chloride, 30 mM potassium acetate, 10 mM calcium chloride and 15% (v/v) glycerol). Cells were incubated for 5 minutes on ice and then centrifuged for 10 minutes at 4°C (5000 x g) and suspended in buffer 2 (10 mM MOPS, 10 mM rubidium chloride, 75 mM calcium chloride and 15% (v/v) glycerol) for 1 hour on ice. Bacterial cells were aliquoted on ice into sterile 1.5 mL Eppendorf tubes in volumes of 200 µL prior to flash freezing in liquid nitrogen and stored at -80°C.

2.2.2 TOP10 transformation

pET30 Aurora A WT and D274N plasmids (50 ng) were transformed into 50 μL TOP10 *E. coli* on ice for 30 minutes in a sterile Eppendorf. The Eppendorf was then transferred to a water bath at 42°C for 30 seconds followed by 2 minutes on ice to induce heat shock. 250 μL SOC (Super Optimal broth with Catabolite repression) media (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was added and incubated for 30 minutes at 37°C in a shaking incubator (250 RPM). Cultures were inoculated on agar plates containing 50 μg/mL kanamycin and 37 μg/mL chloramphenicol and incubated for 18 hours at 37°C. Colonies were picked and grown in 5 mL of LB media containing kanamycin and chloramphenicol and incubated for 18 hours at 37 °C in a shaking incubator (250 RPM). Cultures were centrifuged to collect pellets for 10 minutes at 5000 x g. Purification of plasmids was carried out using the QIAGEN QIAprep Spin Miniprep kit. Plasmid DNA quantity was assessed with Nanodrop 2000c (Thermo Fisher Scientific).

2.3   Protein expression and purification

2.3.1 pLysS *E. coli* transformation

pET30 Aurora A WT and D274N plasmids were transformed into BL21 (DE3) pLysS *E. coli* (Novagen) as described in section 2.2.2. A single colony for each expressed protein was transferred in to 100 mL of LB media with addition of antibiotics and incubated for 18 hours at 37 °C in a shaking incubator (250 RPM). Cultures (5 mL per litre of LB media) were transferred to 8 L of LB media flasks and incubated at 37 °C (250 RPM) until an OD_{600nm} between 0.6 – 0.8 was reached. Protein expression was induced with the addition of 0.4 mM isopropyl-1-thio-β-galactosidase (IPTG) and incubated for 18 hours at 18 °C (250 RPM). Cultures were centrifuged for 10 minutes at 5000 *g*, the supernatant discarded, and the bacterial pellet harvested.

2.3.2 *E.coli* pellet lysis

E.coli pellets were lysed in 100 mL of ice cold lysis buffer (50 mM Tris pH 7.4, 10% glycerol, 300 mM NaCl, 10 mM imidazole, 1 mM DTT, 100 mM EDTA, 100 mM EGTA, protease inhibitor tablet (Roche)). The lysed cells were then sonicated on ice using a 3 mm microprobe attached to a MSE Soniprep 150 plus motor unit at an amplitude of 16 microns in 30 second intervals.

Samples were centrifuged for 1 hour at 8°C (43,000 *x g*) to pellet the cellular debris and then filtered through a 0.22 μm filter. Supernatants were collected and snap frozen if needed prior to protein purification.

2.3.3 His-tag cleavage

Removal of the His-Tag from the Aurora A proteins was carried out using a Nickel His-Trap HP column. The column was equilibrated in elution (50 mM Tris - HCl 7.0, 10% glycerol, 300 mM NaCl, 500 mM imidazole, 1 mM $MgCl_2$, 1 mM DTT) and wash (50 mM Tris - HCl 7.0, 10% glycerol, 300 mM NaCl, 20 mM imidazole, 1 mM $MgCl_2$) buffers using a syringe and filter motion prior to the addition of *E. coli* lysate. The column was then washed with 10 mL of wash buffer and 10 mL of elution buffer in 0.5 mL fractions and the eluent fractions collected. The fractions containing the protein as determined by SDS-PAGE and protein Bradford assay were pooled and incubated with 25 μg TEV for 18 hours at 4°C on a rotating wheel. The sample was centrifuged for 20 minutes at 4°C (16,000 RCF) prior to loading for size exclusion chromatography.

2.3.4 Gel Filtration

Following removal of the His-Tag with TEV, proteins were loaded on to a Superdex 200 16 600 column (GE Healthcare) attached to an AKTA FPLC system and a Frac-920 (GE Healthcare), which was equilibrated in filtered and degassed gel filtration buffer (20 mM Tris pH 7.0, 10% glycerol, 200 mM NaCl, 40 mM imidazole, 5 mM $MgCl_2$, 1 mM DTT). The system was set up with a flow rate of 0.5 mL/min and fractions were collected in 1.5 mL volume. 5 μL of collected fractions containing the protein were analysed by SDS-PAGE to determine the purest fractions. Fractions were pulled together and passed through the His-Trap column until the protein contained no additional non-cleaved material.

2.4   SDS-PAGE

SDS-PAGE was carried out using 10% polyacrylamide gels to determine protein purity, molecular weight and for the initial step prior to western blotting. Gels were made using a resolving buffer (1.5 M Tris-HCl pH 8.8) and stacking buffer (0.5 M Tris-HCl pH 6.8). Proteins were loaded through the stacking gel (0.1 M stacking buffer, 4% Bis-Acrylamide (v/v), 0.01%

(v/v) SDS, 0.01 % (v/v) APS and 0.1% (v/v) N,N,N',N'-Tetramethylethylenediamine (TEMED)) and resolved in the resolving gel 0.4 M resolving buffer, 10% or 12% Bis-Acrylamide (v/v), 0.01% (v/v) SDS, 0.01 % (v/v) ammonium persulfate (APS) and 0.1% (v/v) TEMED. Proteins were denatured by heating in 5x SDS sample buffer (0.25 M Tris-HCI pH 6.8, 0.25% bromophenol blue, 500 mM DTT, 5% SDS, 50% glycerol) for 5 minutes at 95°C. Samples were loaded on to the gel and separated by electrophoresis for ~ 50 minutes at 200 V in running buffer (25 mM Tris-HCI, 190 mM glycine, 0.1% SDS). Proteins were visualised using Coomassie stain (0.2% Brilliant blue R-250, 7.5% acetic acid, 50% methanol) and then the gel was washed using destain buffer (50% H_2O, 40% methanol, 10% acetic acid).

2.5   Bradford assay

Protein concentration was determined by a Bradford assay with a standard curve of bovine serum albumin standards ranging from 0 – 1 mg/mL (Figure 2.1). Proteins were diluted as required in 1X Tris buffer pH 7.4 prior to the assay and 5 µL of protein or buffer for the blank was added to 200 µL of Coomassie Plus Protein Assay Reagent in a 96-well plate. The absorbance was recorded using a spectrophotometer at 595 nm and protein concentration was determined relative to the blank and BSA calibration curve readings.
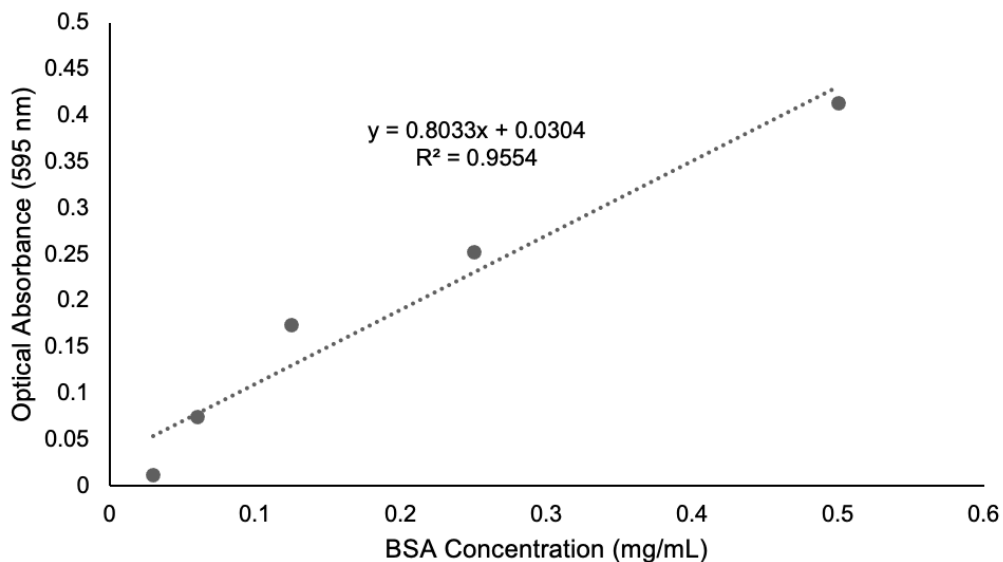
Figure 2.1. BSA standard curve to determine protein concentration.

2.6   Western blotting

Samples were separated by SDS-PAGE prior to transfer onto a nitrocellulose membrane (GE Healthcare) at 100 V for 1 hour in pre-chilled transfer buffer (25 mM Tris, 190 mM glycine, 0.1% SDS, 20% methanol). Membranes were blocked in 5% milk powder (Marvel) in Tris-buffered saline and 0.1% Tween 20 (TBST) (20 mM Tris pH 7.6, 137 mM NaCl, 0.1% Tween-20 (v/v)) for 1 hour at room temperature on a shaking rocker. Primary antibodies were made up in 5% milk TBST at 1:5000 dilution and incubated with the membranes for 18 hours at 4°C prior to washing for 3x 5 minutes in TBST. Washed membranes were incubated in secondary antibody (diluted 1:5000 in 5% milk in TBST) for 1 hour at room temperature on a rocker prior to washing for 3x 10 minutes in TBST. Immunoblion Western Chemiluminescent HRP Substrate (Millipore) developing reagent was applied to the membranes prior to their transfer into an X-ray film cassette and exposure to film for various time periods. The films were development using an ECOMAX X-ray film processor (Protec).

2.7   Differential Scanning Fluorimetry (DSF)

Protein solutions were diluted in 50 mM Tris pH 7.4, 100 mM NaCl to a final concentration of 5 µM with the addition of 1 mM ATP, 10 mM $MgCl_2$ and 40 µM of inhibitors as required. Tris (10 mM Tris, pH 7.4, 20 mM NaCl) buffer or 4% DMSO were used as controls for the assay. SYPRO Orange fluorescent dye (Invitrogen) was diluted 1:1000 in 10 mM Tris, pH 7.4, 20 mM NaCl buffer and added to the MicroAmp Fast Optical 96-well reaction plate (Applied Biosystems) prior to the addition of an optical adhesive cover. The plate was vortexed and centrifuged for 30 seconds at 500 x g prior to analysis using an Applied Biosystems StepOnePlus Real-Time PCR instrument. The instrument was set to measure fluorescence with thermal ramping (0.3°C per minute between 25°C and 95°C). Data was collected using the StepOne Software (version 2.1). Data was processed to generate denaturation profiles, which was then normalised to include the lowest fluorescent reading at 0% and highest at 100%. Linear regression analysis was carried out using the Boltzmann equation to generate sigmoidal denaturation curves, and average Tm/ΔTm values were calculated using GraphPad Prism software.

2.8   Enzymatic assays

Enzymatic assays were performed on a PerkinElmer LabChip EZ Reader using a 384-well plate to assess the mobility shift of a peptide substrate upon WT and D274N variants of Aurora A. Final volume of 80 µL in the presence of 50 mM Hepes pH 7.4, 0.015% (v/v) Brij-35 and 5 mM $MgCl_2$, with the addition of 1 mM DTT, 2 µM peptide substrate (5'-FAM-LRRASLG-CO_{NH2}) and 10 ng protein. Activity of both Aurora A variants was analysed with the addition of 1 mM ATP and quantified over the assay time relative to the control assay with no ATP. Data was plotted as % peptide conversion using GraphPad Prism software.

2.9   Desalting with Amicon spin filters

Prior to all mass spectrometry experiments, Aurora A was buffer exchanged into 150 mM ammonium acetate (for IM-MS) or 50 mM ammonium bicarbonate (for phosphopeptide analysis) using Amicon spin filter 10K cut-off columns at 4°C. Spin columns were pre-washed with buffer prior to addition of protein and centrifuged 3x 10 minutes at 13,000 RPM. Following the final spin, the filter was inverted into a new collection tube and spun for 2 minutes at 3,000 RPM to collect the protein. Final protein concentration was determined using a Nanodrop and adjusted accordingly to experiment. Nanodrop was blanked with the appropriate buffer using the Protein 280 nm setting prior to loading 1 µL of sample to determine the concentration.

2.10   Native ion mobility-mass spectrometry

2.10.1 Waters Synapt G2Si

Analysis was performed on a Waters Synapt G2Si instrument, which was calibrated with sodium caesium for a mass range between 500 – 8000 *m/z*. Following desalting (as in 2.9), proteins were adjusted to 5 µM in 150 mM ammonium acetate and the sample was subject to electrospray ionization at ~ 2 kV with a pulled nanospray tip (Section 2.10.3). Pressure in the travelling wave mobility cell was set to 2.2 mbar (nitrogen), IMS was carried out with a travelling wave height of 23 V, velocity of 496 m/s and a trap bias of 33. For collision-induced unfolding (CIU) experiments, the 11+ charge state of bound or unbound protein was quadrupole isolated and subjected to collisional activation in the trap region of the TriWave. Step-wise activation voltage was applied between 16 and 34 V in two-volt intervals before

IMS measurement. CIU was carried out a travelling wave height of 27 V, velocity of 497 m/s and a trap bias of 35.

2.10.2 Calculation of CCS

Drift time calibration of the TriWave device for determination of rotationally averaged collision cross section (CCS) was carried out using β-lactoglobulin (Sigma-Aldrich L3908), cytochrome c (Sigma-Aldrich C2506) and bovine serum albumin (Sigma-Aldrich A2153). CCSs in nitrogen buffer gas were estimated according to the database provided by Bush et al (150) and then calibrated against helium cross sections (^{TW}CCS $_{N2>He}$). A logarithmic plot of the reduced drift time for multiple charge states for each of the protein standards used for CCS determination versus the charge and reduced CCS value was calculated, prior to the generation of a straight-line graph to establish the slope and intersection. The experimentally derived CCS values were then calculated using the formula shown in Figure 2.3.

$$y = 0.58x - 0.7625$$
$$R^2 = 0.9986$$
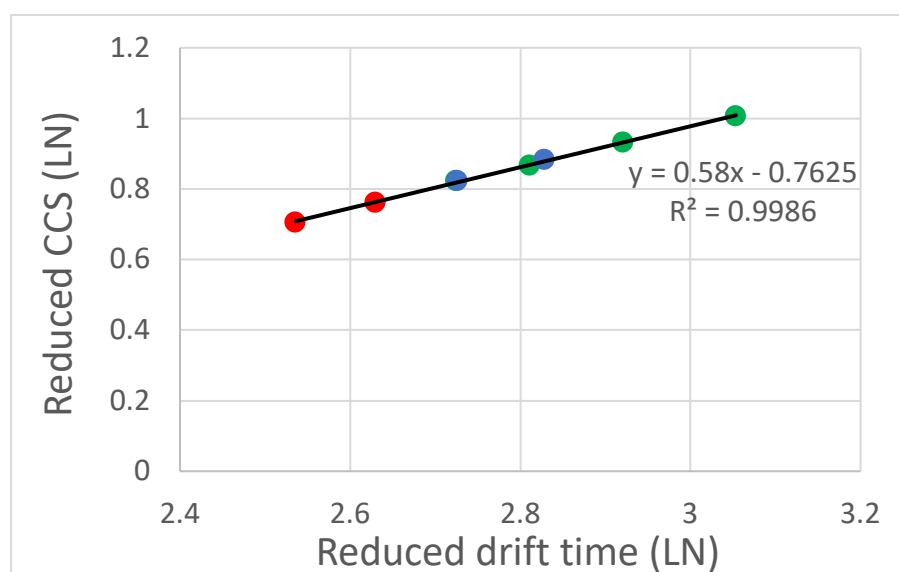
Figure 2.2. CCS calibration plot used for calibration of TriWave. Different charge states for each protein calibrant shown in red for Avidin, blue for β-lactoglobulin and green for BSA.

$$11*((C18-1.5*SQRT(3050)/1000))^0.58*EXP(-0.7625)$$

Charge state for the
protein of interest

m/z for protein
of interest

Inputted from y
intercept shown in
Figure 2.2

Intensity of chromatogram list
copied from MassLynx
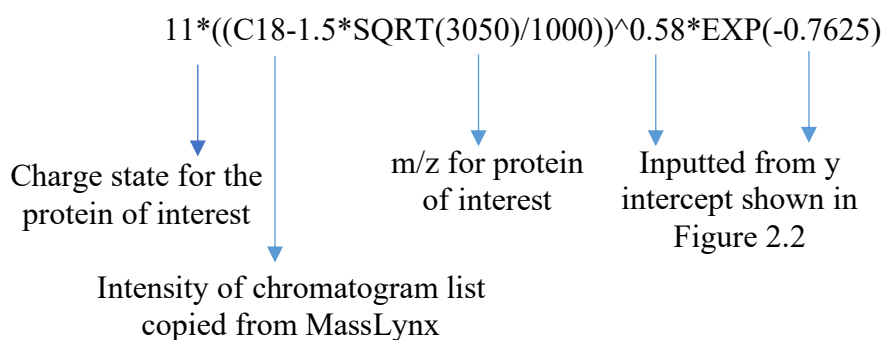
Figure 2.3. Example for calculation of CCS.

2.10.3 Nanospray pulled tips

TW120–4 Thin-Wall Capillaries (4 × 1.2 mm) (World Precision Instruments 1B100-3) were pulled using a Sutter P-1000 puller using settings displayed in Table 2.2. Pulled capillary tips were placed in a secure clasp holder and ~ 3 µL of protein was added using a gel loading tip. The capillary was centrifuged briefly using a clasp holder to ensure all the liquid had reached the tip of the capillary. A platinum wire was placed in the capillary tip following removal from the secure clasp. A metal clasp and protective rubber insert were applied to the outside of the capillary prior to mounting on the G2Si source stage.

Table 2.2. Sutter P-1000 puller settings.

Line	Heat	Pull	Vel.	Delay	Pressure	Ramp
1 x 1	489	0	20	120	235	491
2 x 1	489	0	18	120	Delay mode **X**	
3 x 1	489	0	18	120	Safe heat **X**	
4 x 1	0	0			Jaw temp **42°C**	

2.10.4 Inhibitor/nucleotide binding

Inhibitor compounds and nucleotides were made up in 10 mM DMSO and then diluted to 200 µM in 150 mM ammonium acetate prior to use (stored at -20°C). 10:1 excess

(inhibitor/nucleotide:protein) was added following protein desalting (2.9) and incubated for 10 minutes at R.T. prior to native (IM-)MS analysis.

2.10.5 Gaussian fitting of CCS distributions

Average CCS profiles from three replicate analyses were used for Gaussian fitting of Aurora A individual conformational states using Origin (Version 2021b). The Fit Peaks Pro function was implemented to initially assign the most abundant peak for each data set. Additional peaks were added for either conformer I, II, III, and IV, where applicable. The CCS, CCSD, and area parameters for each assigned conformer were manually adjusted using Fit Control and the iteration feature, in order to obtain the best fit between experimental data (black line) and sum of Gaussians (red line). Black error bars are representative of the SD, and the error between the experimental line and sum of Gaussians was reported with R^2.

2.10.6 RStudio

Scatter plots of $^{TW}CCS_{N2>He}$ (nm^2) values versus CCS distribution (CCSD) (nm^2) were generated using ggplot in RStudio (code included in Appendix 2.1). Aurora A MS/MS fragmentation plots were generated in RStudio (code included in Appendix 2.2)

2.10.7 CIUSuite 2

Raw data files were converted from Waters.raw to a text-based format using the drift time extraction tool in TWIMExtract. The data from the isolated charge state was summed across the *m/z* range and IM drift time was converted to CCS to compile a file of collision-voltage IM sets. TWIMExtract compiles the data in to a _raw.csv file, which is then used as the input file for the Python operated CIUSuite 2 software. The ion intensities at each collision voltage energy were smoothed with a Savitsky-Golay filter using a window length of three and a polynomial order of two. The CIUSuite_plot function was used to process the CIU data set within the working directory to create a contour unfolding plot in the form of a .png file.

2.11   Intact mass analysis

The mass of the proteins was determined using a Waters Nano Acquity Ultra Performance Liquid Chromatography (UPLC) system running a C4 trap column (Waters - MassPREP Micro

Desalting Column) coupled with a Waters Synapt G2Si. The mass spectrometer was calibrated with GluFib for a mass accuracy of 1 ppm or below and performance tested with 250 fmol/mL myoglobin. Chromatography was performed using the gradient Table 2.2, with buffer A (0.1% (v/v) formic acid in HPLC grade water) and buffer B ((0.1% (v/v) formic acid in HPLC grade acetonitrile). MS data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1.

Table 2.3. LC-MS gradient used for intact mass determination.

Time (min)	Flow (µL/min)	Buffer A (%)	Buffer B (%)	Curve
Initial	25	95	5	6
0.10	40	95	5	6
5.10	40	95	5	6
5.20	25	95	5	6
6.00	25	95	5	6
6.10	25	95	5	6
7.60	25	10	90	6
7.90	25	95	5	6
8.60	25	10	90	6
8.90	25	95	5	6
9.60	25	10	90	6
11.60	25	10	90	6
11.70	25	95	5	6

2.11.1 NME autophosphorylation assay

Catalytic activity of the histidine kinase NME1 was evaluated by intact mass analysis following protein incubation (1 µg) with 1 mM nucleotide (or without for control) in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl in a final volume of 100 µL. Reactions were incubated for 5 minutes at R.T. prior to LC-MS analysis (as in 2.11) and the ratio of non-phosphorylated to phosphorylated protein quantified. To determine the effect of inhibitors on NME1 auto-phosphorylation and dimerisation, 200 µM compound (in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl) was included in the assay (final volume of 50 µL, equating to a 1:10 protein: drug ratio). Reactions were incubated for 10 minutes at R.T prior to addition of 1 mM ATP. Samples were centrifuged briefly and analysed immediately.

2.12   Phosphopeptide mapping

2.12.1 Sample preparation and LC/MS/MS analysis

Following desalting of proteins into 50 mM ammonium bicarbonate, disulphide bonds were reduced with 3 mM DTT in 50 mM ammonium bicarbonate for 30 minutes at 60°C. After cooling to R.T., cysteine residues were then alkylated with 14 mM iodoacetamide for 1 hour in the dark at R.T. Proteins were digested with trypsin ((50:1) Promega) for 18 hours at 37°C, prior to loading a 3 pmol injection onto the instrument. Samples were made up in loading buffer (3% (v/v) ACN, 0.1% (v/v) TFA) and loaded onto a trapping column (PepMap100, C18, 300 μm x 5 mm) at a flow rate of 9 μL/min for seven minutes. Peptides were subsequently resolved at a flow rate of 0.3 μL/min on an analytical column (Easy-Spray C18 75 μm x 500 mm, 2 μm bead diameter column) over a 60-minute gradient with 3% buffer A (0.1% (v/v) Formic acid in H_2O):97% buffer B (80% (v/v) ACN, 0.1% (v/v) Formic acid in H_2O) to 20% buffer A:80% buffer B and a final 5 min 100% B wash. MS1 spectra was acquired on a Thermo QExactive mass spectrometer over a range of 350-2000 *m/z* in the Orbitrap (resolution at 200 *m/z*), AGC target was set at $2E^5$ and a maximum injection time of 50 ms. MS2 data was acquired in data-dependent acquisition (DDA) mode using a top speed approach with a cycle time of 3 s. Normalised collision energy was set to 32% HCD. MS2 settings included 15,000 resolution at 200 *m/z*, maximum injection time of 50 ms, fragmentation intensity threshold of $5E^4$ for charge states 2+ to 5+. A dynamic exclusion window of 60 s with a mass tolerance of 10 ppm was implemented.

2.12.2 Data analysis with Proteome Discoverer

LC-MS/MS data was processed using Thermo Proteome Discoverer (2.4) and MASCOT (2.6). Raw mass spectrometry data files were converted to mzML format to enable processing with Proteome Discoverer. Data was searched against a human UniProt Aurora A database modified with the two expressed protein sequence (122-403 residues of the WT and D274N variant). Processing settings were set as follows:MS1 tolerance – 10 ppm and MS2 mass tolerance – 0.01 Da; enzyme specificity was defined as trypsin with up to two missed cleavages allowed; carbamidomethyl Cys was set as a fixed modification; Met oxidation and Ser/Thr/Tyr phosphorylation were defined as variable modifications. Data were filtered to a 1% false discovery rate (FDR) on peptide spectrum matches (PSMs) using automatic decoy

searching with MASCOT. ptmRS node with Proteome Discoverer was used to determine phosphosite localisation confidence.

2.13   Molecular Modelling

Molecular modelling was carried out by Dr Matthew Bachelor, at the University of Leeds. Missing parts of the Aurora A sequence were modelled into the structure using the PYMOD plugin in PYMOL (155) (PyMOL, The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). A homology model of the full 122–403 sequence was built using MODELLER (156) based on one of the Aurora A crystal structures. The MODELLER loop modelling function in PYMOD was then used to build ten new improved, Aurora A models, allowing only the newly added residues of the N- and C-termini (and any new A-loop residues) to change. The model with the lowest 'objective function' and without obvious new contacts made with the rest of the protein was chosen as the starting structure for modelling. This procedure was performed for several Aurora A crystal structures, PDB codes: 1OL7, 1OL5, 2WTV (chain A and B), 3E5A, 4C3P, 5G1X, 5L8K, 5ODT, 6HJK.

All atom simulations were performed with the CHARMM36m force field (157) using NAMD (158). Inputs for NAMD simulations were generated using CHARMM-GUI (159) based on the PYMOD generated models. Where required, phosphorylation of Thr287/Thr288 was maintained using the doubly-deprotonated Thr patch (THPB). N- and C-termini were uncapped. The protein was solvated in a rectangular waterbox with a minimum distance of 10 Å between the protein and the box edge (~20,000 TIP3P water molecules). Cl⁻ ions were added to neutralise the protein. Solvated structures were first subjected to 10,000 conjugate gradient energy-minimization steps. Prior to the collection of trajectory data, a heating protocol that raised the temperature of the system from 0 to 300 K over 60,000 steps and a short pre-equilibration at 300 K for 125,000 steps, were used. The time step used was 2 fs throughout. Trajectory frames were recorded every 5000 steps (10 ps) and simulations ran for >100 ns with temperature controlled at 300 K and pressure at 1 atm using Langevin dynamics.

Gō-like models and potentials were generated from PYMOD generating initial structures using the MMTSB web service (https://mmtsb.org/webservices/gomodel.html) (160, 161). MD simulations of Gō-like models were carried out using Langevin dynamics and the CHARMM package, version 44/45 (162). The timestep was 10 or 15 fs. Simulations across a range of different temperatures were performed to gauge where the unfolding transition occurs then production simulations were performed below this temperature. Simulation trajectories were processed and analysed using Wordom (163). The protein component of the system was isolated and aligned, and individual trajectory frames extracted for CCS measurements. The stability/variability of the protein as a function of time through the simulation was monitored using the root mean square deviation (RMSD) of CA atoms with reference to the initial structure.

Clustering of Gō-like model conformers was performed using a 15-Å RMSD cut-off value between clusters. Native contact fractions in Gō-like models were calculated as described by (164) using a low temperature (250 K) simulation to define native contacts distances at 80% occupancy.

The software IMPACT (165) was used to estimate CCS values for protein structures. The default atomic radii and convergence parameters were used for all-atom and united-atom simulations. For Gō-like models, the atomic radii were estimated to be the average distance between each CA atom (3.8 Å). This provided reasonable comparison with the all-atom simulation results. In all cases the raw IMPACT CCS value based on projection approximation (rather than the recalibrated TJM value) was used, as this provided much better comparison with experimental data for AurA models and for a bovine serum albumin model.

Chapter 3. Evaluating the differences in conformation and dynamics of WT and D274N Aurora A as a function of activity.

3.1   Introduction

As discussed in Chapter 1, Aurora kinases are involved in many cell cycle events, and belong to the family of serine/threonine protein kinases. Aurora A controls progression of mitosis via the phosphorylation of various targets. Autophosphorylation of the Thr288 residue is required for catalytic activity of Aurora A and with the binding of activating partners, a fully active conformation is present, which can regulate the assembly of mitotic spindles. Aurora kinases consist of an N-terminal domain (39-139 aa), a kinase domain (250-300 aa) and a C-terminal domain (15-20 aa) (Figure 3.1) (166). The kinase domain consists of a β-stranded N-terminal lobe connected to an α-helical C-terminal lobe by a hinge region that is associated with the active conformation (167). The C-terminal domain contains the conserved residue Thr288 for Aurora A, with phosphorylation of this residue resulting in a conformation change associated with its activity (34).
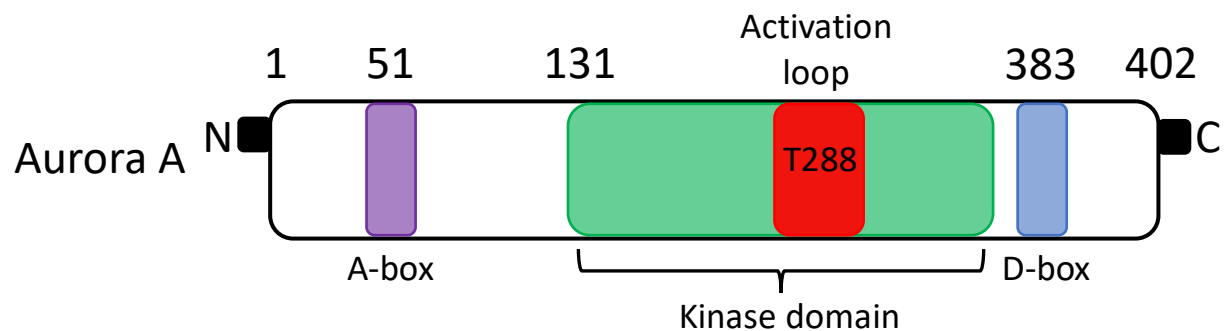
Figure 3.1. Aurora A structural domain schematic. The domain structure of full-length Aurora A, including the A-box at residue 51, kinase domain containing the activation loop with activated residue T288, and the D-box at residue 383.

The N-terminal and C-terminal lobes are connected via a flexible hinge region, which forms the ATP-binding site. The activation loop plays an important role in the enzymatic activity of the protein kinase and is situated within the active site, with the loop commonly between 20-30 residues in length, beginning with a conserved DFG motif (Asp-Phe-Gly) and extending up to an APE motif (Ala-Pro-Glu) (Figure 3.2A). In active kinases, the Asp residue within the DFG motif is positioned and orientated to bind magnesium, which interacts with the oxygen atom of the β phosphate of ATP. Active kinases position the C-helix in an inward position, which places a conserved Glu residue into the helix forming a salt bridge with a Lys residue in the β3

strand. Following formation of the salt bridge, the lysine side chain is able to form hydrogen bonds with ATP via the oxygen atoms of α and β phosphates. The GxGxxG motif situated in the N-terminal lobe stabilises the bound ATP phosphates during the catalysis step, which creates a stable environment for the phospho-transfer reaction (8).
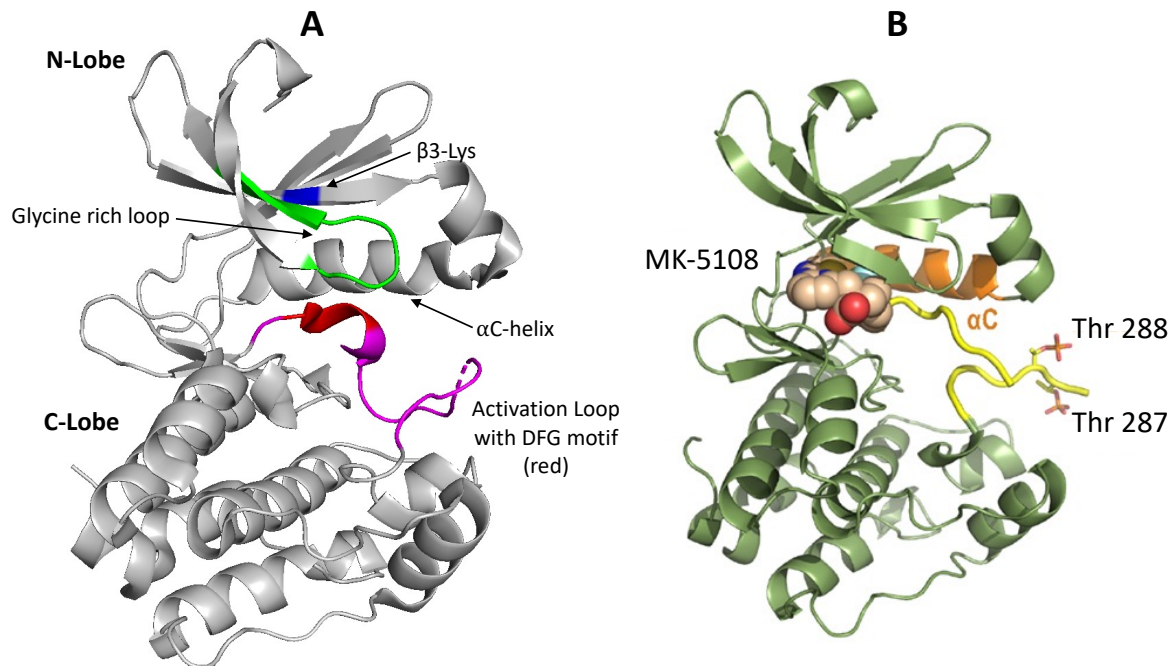
Figure 3.2. Aurora A protein kinase structures with and without inhibitor. A) Aurora A structural features including the activation loop (magenta), DFG motif (red), glycine rich loop (green) and Lys162 (blue). Diagram produced in PyMol from PDB file 1MQ4. **B)** Aurora A (ribbon diagram) in complex with MK-5108 (spacefilling model). Majority of kinase domain shown in green, with activation loop shown in yellow, αC helix shown in orange, and the phosphorylated Thr 287 and Thr 288 residues highlighted. Figure adapted from (168).

In this chapter, a truncated version of Aurora A, with only the residues 122-403 will be used for analysis due to this shorter version being sufficient to bind activating partners. The TPX2 binding partner has previously been shown to bind to this truncated version of Aurora A and is able to stimulate activity of this kinase *in vitro* as effectively as binding to full-length Aurora A (35). In addition, the Aurora A (122-403) protein contains the important conserved Thr288 residue required for catalytic activity. Aurora A (122-304) has been shown to be phosphorylated when expressed in *E. coli,* as detected by immunoblotting with a phosphospecific antibody (33). Crystal structures of phosphorylated Aurora A (122-403) bound to TPX2 with both Mg2$^+$ and ATP have been determined to show that the binding of TPX2 results in the kinase active conformation (35). Various other phosphorylated crystal

structures of this truncated version have also been reported including with phosphorylation on Thr287 and Thr288, and of Aurora A bound to the MK-5108 inhibitor. Binding of MK-5108 has shown to result in Aurora A adopting an inactive conformation, where the activation loop and C helix are positioned improperly for catalysis. The structure was shown to be phosphorylated on the Thr287 and Thr288 residues, with the MK-5108 inhibitor able to interact with the active form of Aurora A and occupy the nucleotide-binding pocket in-between the two lobes of the kinase (Figure 3.2B) (168).

In addition to a phosphorylated form of Aurora A (122-403), a mutated version of this protein will also be analysed in this chapter. The aspartic acid residue situated in the DFG motif will be mutated to an asparagine to result in a D274N mutated variant. This residue in Aurora A is important to the catalytic activity of the protein. By mutation of the aspartic acid, the protein is unable to bind ATP/MgCl$_2$, deeming it catalytic inactive. Studies have shown the mutated D274N Aurora A variant to be catalytically inactive, but still retains TPX2 binding capability (35). When expressed in *E. coli*, the D274N variant unlike wild-type (WT) protein, was not phosphorylated, suggesting that the WT Aurora A autophosphorylates, and the observed phosphorylation does not arise due to active bacterial kinases (169).

Studies have shown that phosphorylation can induce local and global changes to the protein structure, altering catalytic activity and binding specificity (170). Changes in binding capability, including the ability to form oligomeric states (homo- or hetero- oligomers) can also influence protein activity (171, 172). This chapter will exploit the use of ion-mobility mass spectrometry (IM-MS), where non-denaturing conditions are used to mimic the proteins physiological surroundings. The analysis of the two forms of Aurora A, phosphorylated WT and non-phosphorylated D274N, was carried out by IM-MS to determine whether changes are evident in protein conformation upon activation.

3.1.1 Aims

The aims of this chapter were to express, purify and characterise the conformational dynamics and stability of Aurora A (phosphorylated WT and non-phosphorylated D274N), evaluating differences in conformation and dynamics as a function of activity.

3.2 Results & Discussion

3.2.1 Purification of WT and D274N Aurora A (122-403)

Two Aurora A (122-403) forms, a WT, active version which autophosphorylates during bacterial expression, and an inactive, non-phosphorylated D274N mutant, were purified to evaluate the effects of phosphorylation on protein conformation by IM-MS. 6His-N-terminal-tagged Aurora A (122-403) WT and D274N plasmids were provided from the Prof Richard Bayliss lab at the University of Leeds. The truncated version of Aurora A was used as this is sufficient to bind activating partners, such as TPX2, while removing the N-terminal region, which is poorly conserved across species.

His-tagged Aurora A was expressed in *E. coli* and separated from the bacterial cell lysate using a Nickel His-Trap HP column, with the His-Tag cleaved using TEV protease, as described in Chapter 2. Following removal of the His-Tag, proteins were loaded onto a Superdex 200 16 600 column attached to an AKTA FPLC system with a Frac-920 and eluted in 1.5 mL fractions, monitoring absorbance at 595 nm. Peaks were observed between 48 – 63 mL for WT Aurora A (Figure 3.3A) and between 45 – 65 mL for D274N Aurora A (Figure 3.3B). Analysis of the eluted fractions by SDS-PAGE revealed a molecular mass of ~32 kDa (Figure 3.3C/D), which was consistent with the expected molecular mass of Aurora A (122-403). The fractions containing the highest concentration of protein were pooled and passed through the His-Trap column again to remove any remaining non-cleaved (His-tagged) protein (Figure 3.4).
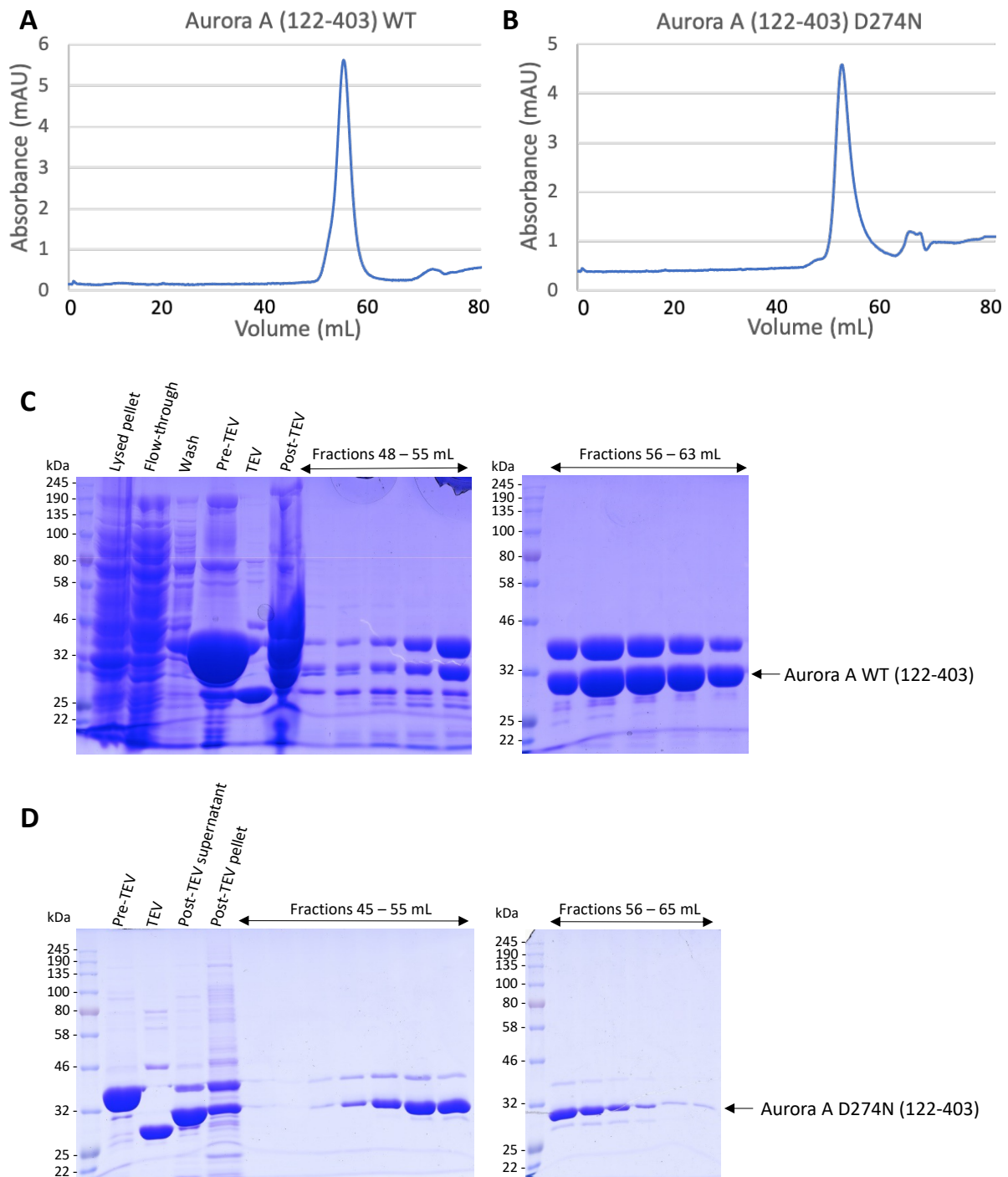
Figure 3.3. Purification of WT and D274N Aurora A (122-403). 6His-N-terminally tagged human wild-type (WT) Aurora A (122-403) and D274N Aurora A (122-403) were expressed from a pET30TEV vector in BL21 (DE3) pLysS *Escherichia coli.* His-tagged Aurora A was separated from bacterial cell lysate using a Nickel His-Trap HP column and eluted with 500 mM imidazole in 0.5 mL fractions. Protein-containing fractions were pooled, and the His-Tag cleaved from Aurora A with addition of 25 μg TEV protease for 18 hours at 4°C. Cleaved His-Tag protein was loaded on to a Superdex 200 16 600 column attached to an AKTA FPLC system and 1.5 mL fractions collected whilst monitoring protein elution at an absorbance of 280 nm (**A, B**). 5 μL of protein-containing fractions were analysed by SDS-PAGE to evaluate purity (**C, D**).
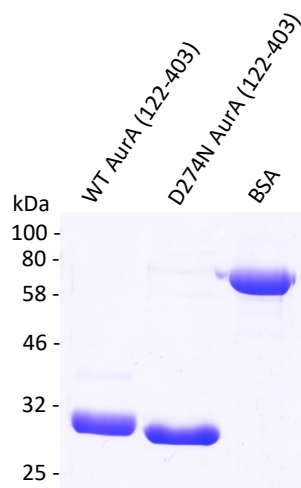
Figure 3.4. Molecular weight determination of WT and D274N Aurora A (122-403). SDS-PAGE gel of BSA standard and purified WT and D274N Aurora A (122-403) proteins (2 μg). His-Aurora A proteins were purified using a Nickel HisTrap HP column, prior to His-Tag cleavage and gel filtration elution. Aurora A (122-403) has a predicted MW of 33 kDa.

3.2.2 Intact mass analysis of Aurora A

Prior to IM-MS analysis, the two Aurora A proteins were analysed by intact MS under denaturing conditions to determine the molecular weight of the two expressed and purified proteins (WT and D274N) (Figure 3.5). Raw MS data was deconvoluted in MassLynx 4.1 using MaxEnt1 to determine experimental masses for WT and D274N Aurora A (122-403). The theoretical mass of WT Aurora A (122-403) with the addition of the mass of the TEV cleavage site (+ serine) is 32,724.49 Da. Therefore, a mass difference of 254.51 Da was determined between the theoretical and observed (32979 Da) masses. WT Aurora A is known to be phosphorylated, which could account for a difference of 240 Da with three phosphorylated sites. The mass shifts to the right of the peak at 32979 Da (81 Da) could also be representative of an additional three phosphorylated sites, which results in a final number of six phosphorylated sites. An oxidated methionine could account for the remaining mass difference between theoretical and observed mass (16 Da) (173). The smaller intensity peaks (shown in red) in-between the more abundant peaks could be indicative of disodium clusters, which could have come from glassware and the protein purification steps.
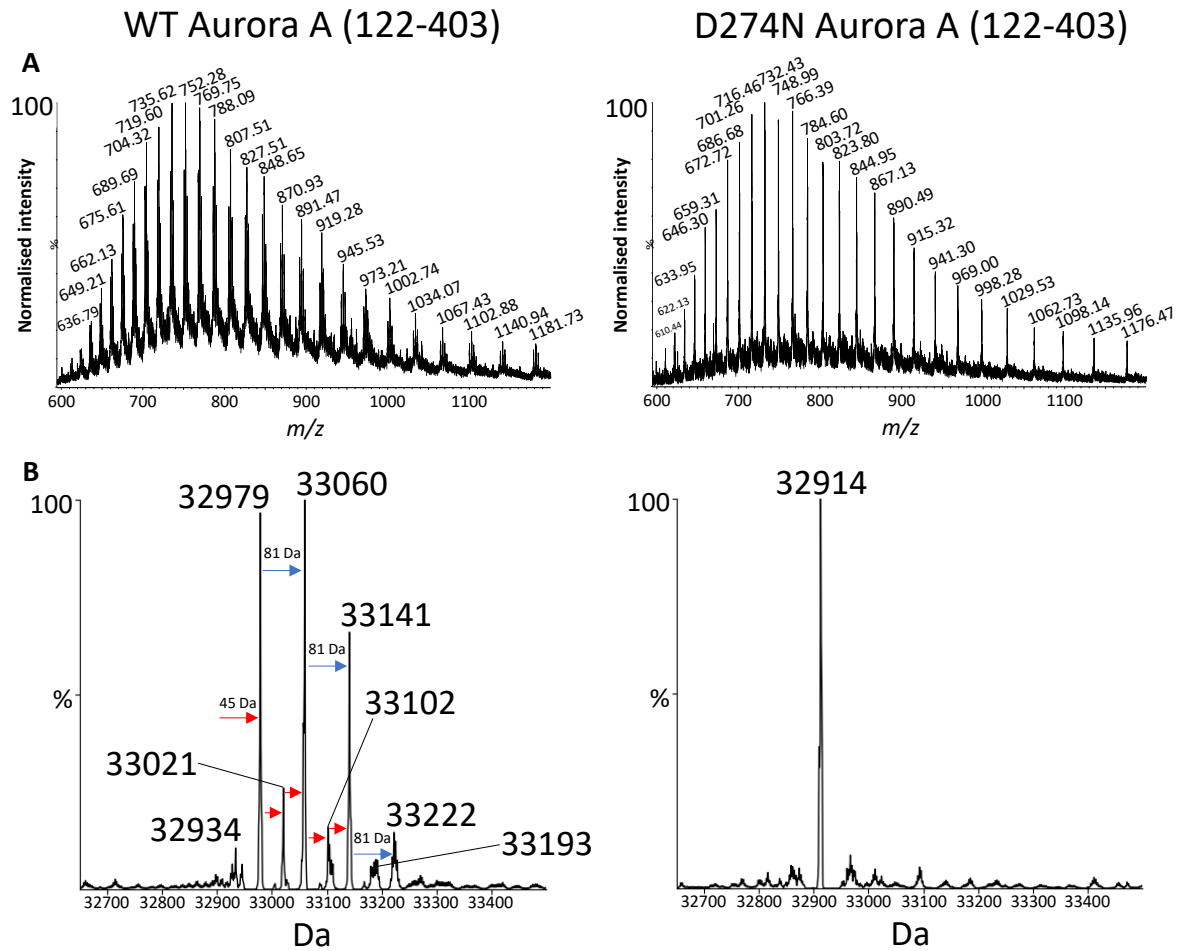
Figure 3.5. Molecular weight determination shows phosphorylation mass shifts associated with active Aurora A WT protein. A) Pre-deconvoluted spectra for WT (left) and D274N (right). **B)** Deconvoluted mass spectra for WT (left) and D274N (right). Data were analysed using MassLynx 4.1 and deconvoluted in MaxEnt1. Following deconvolution in MaxEnt1, the data was then copied into UniDec to produce the figures shown in B. Protein was diluted to 1 µg in 0.1% formic acid prior to analysis.

3.2.3 Aurora A phosphorylation status

A phosphospecific antibody against pThr288, the residue that sits within the activation loop and phosphorylation of which is required for catalytic activity of Aurora A (174), was used as a proxy to determine catalytic activity of these two Aurora A proteins. The immunoblot showed WT Aurora A to be phosphorylated at Thr288, whereas D274N Aurora A showed a negligible amount of phosphorylation at this residue (Figure 3.6), confirming lack of (auto) catalytic activity of the D274N variant.
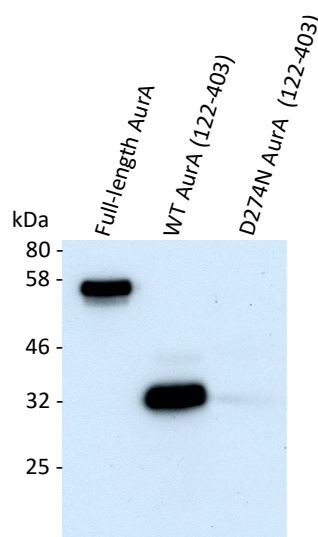
Figure 3.6. WT but not D274N Aurora A is phosphorylated on Thr288 that is required for catalytic kinase activity. Immunoblot of full-length Aurora A (CST), and purified WT and D274N Aurora A (122-403) using an anti-phospho T288 antibody (200 ng of protein).

To confirm these findings, the sites of phosphorylation on these two proteins were characterised by LC-MS/MS following tryptic digestion, which cleaves at arginine and lysine residues. A total of six sites of autophosphorylation were identified for WT Aurora A, including the Thr288 residue required for catalytic activity of this protein kinase (Table 3.1. MS/MS fragmentation shown in Figure 3.7), in agreement with the immunoblot data. Phosphorylation of S226, S266, S278, T287 and T288 have all been previously identified in human recombinant Aurora A following tryptic digestion (175, 176). However, phosphorylation on the S186 residue of Aurora A has not been reported to date. No sites of phosphorylation were identified on purified D274N Aurora A under the same conditions.

Peptide sequence	Modification	Charge	m/z (Da)	DeltaM (ppm)	Ion score	ptmRS
EVEIQsHLR	S186 (Phospho)	2	595.78	1.16	22	100
ELQKLsKFDEQR	S226 (Phospho)	3	534.26	1.89	32	100
DIKPENLLLGsAGELK	S266 (Phospho)	2	888.96	0.99	61	129
IADFGWsVHAPSSR	S278 (Phospho)	2	805.36	1.05	73	100
RttLAGTLDYLPPEmIEGR	T287 (Phospho); T288 (Phospho); M300 (Oxidation)	2	1155.01	-3.84	43	99.75; 100

Table 3.1. Phosphosite mapping of WT Aurora A (122-403).

Aurora A was buffer exchanged into 50 mM ammonium bicarbonate, reduced, alkylated, and digested with trypsin. Tryptic peptides were analysed on a Thermo QExactive with 32% HCD fragmentation. Data was processed using Thermo Proteome Discoverer 2.4 against a human UniProt Aurora A database modified with the 122-403 residues and D274N mutation. Processing settings were as follows: dynamic modifications – Phospho (S/T/Y), maximum missed cleavages – 2, MS1 tolerance – 10 ppm and MS2 mass tolerance – 0.01 Da. Modified phosphorylated residues are shown in lower case in peptide sequences. Mass error for each peptide was calculated using DeltaM (ppm) score. Ion score measures how closely the MS/MS spectrum matches the peptide of interest. ptmRS is an algorithm for PTM assignment, which uses MS/MS spectra alongside the peptide sequence to calculate site probabilities for all the potential modified sites (124).
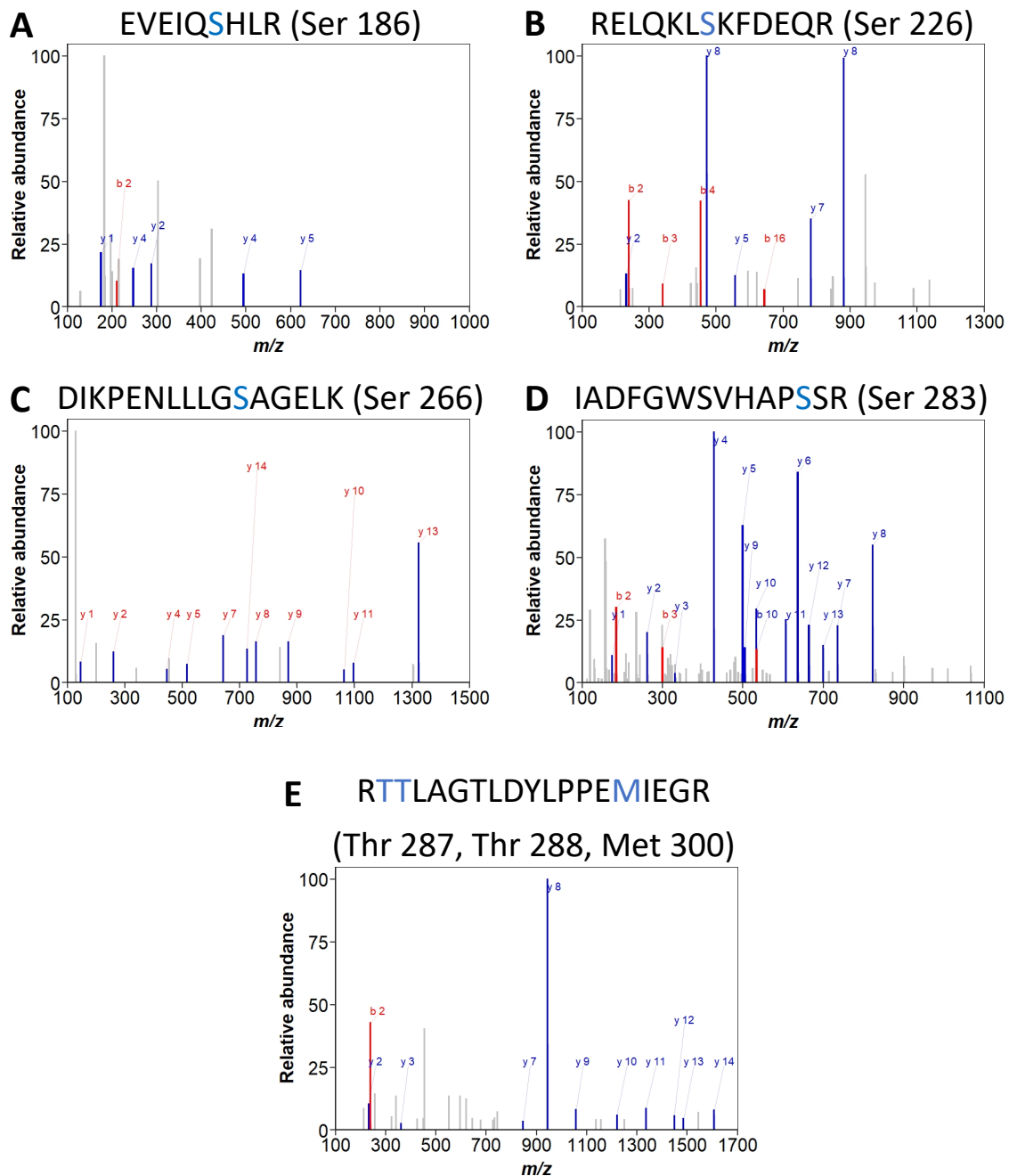
Figure 3.7. Product ion spectra for phosphorylated sites.
Purified WT Aurora A was digested using trypsin and analysed by LC-MS/MS. MS2 spectra generated by HCD shows autophosphorylation site at the following residues **A)** Ser186, **B)** Ser226, **C)** Ser266, **D)** Ser283 and **E)** Thr287, Thr288, Met300. MS2 spectra plots were created using RStudio (Version 1.3.1073).

3.2.4 Aurora A activity

The enzymatic activity of WT and D274N Aurora A was directly evaluated using a peptide mobility shift assay on the EZ Reader platform. WT and D274N Aurora A were incubated with a fluorescent peptide substrate (5'-FAM-LRRASLG-CO$_{NH2}$) and Mg/ATP with the conversion of substrate to phosphorylated product quantified over time to determine the catalytic activity of both forms of Aurora A. These data were in agreement with the observations of catalytic activity (as determined by autophosphorylation) of WT but not D274N Aurora A; WT protein exhibited a high conversion rate, with ~48% of substrate being phosphorylated within 40 min, while the D274N inactive mutant failed to show any activity against this peptide substrate in the presence of Mg/ATP (Figure 3.8).
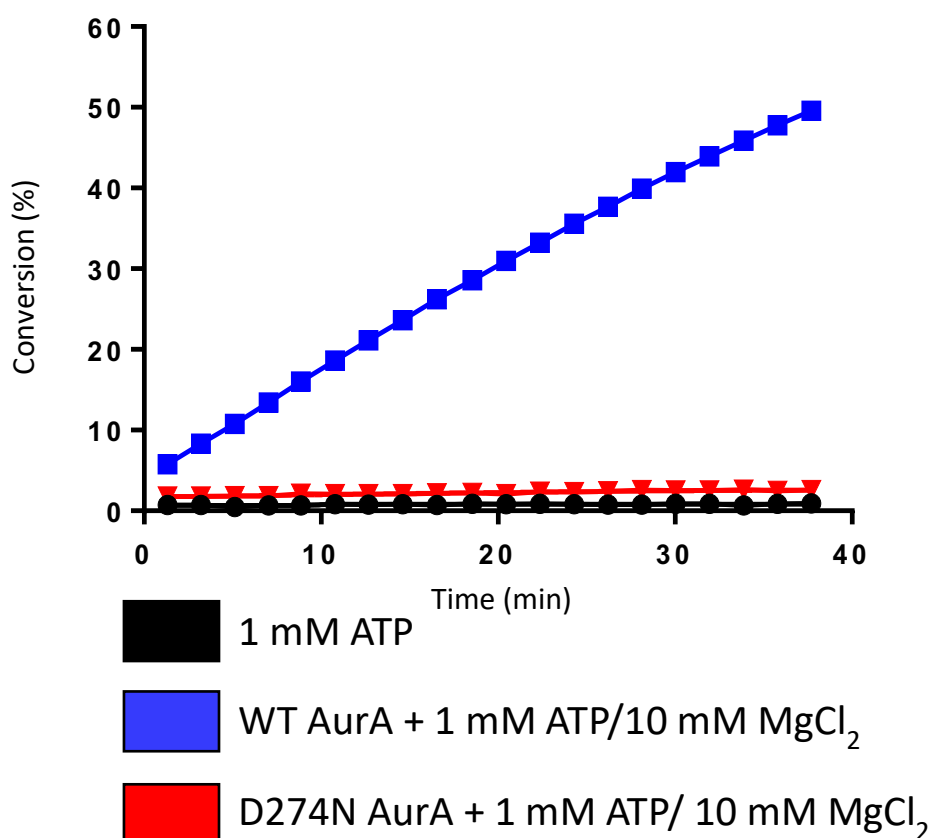
Figure 3.8. WT Aurora A can phosphorylate a substrate peptide in comparison to non-phosphorylated D274N mutant. Enzyme activity assay of WT and D274N Aurora A (122-403) in the presence of 1 mM ATP/ 10 mM MgCl$_2$ (10 ng protein, 2 µM substrate peptide).

3.3.5 Aurora A stability by DSF

The thermal stability of Aurora A was analysed by DSF to determine the effect of phosphorylation and kinase activation status on protein stability for WT Aurora A versus the D274N mutant. WT Aurora A was shown to have a lower melting temperature in comparison to D274N Aurora A, indicating that the inactive variant protein is more stable than the phosphorylated active form (Figure 3.9A/B). Melting temperature of the WT Aurora A increased by 8.3°C in the presence of ATP/MgCl$_2$. The increased ΔT_m observed for WT Aurora A with ATP/MgCl$_2$ is indicative of tight ATP binding, whereas the negligible change with D274N Aurora A is representative of this protein's inability to co-ordinate ATP/MgCl$_2$. The stabilisation of WT Aurora A with ATP alone (+ 2.5 °C) demonstrates that Mg^{2+} is required for high affinity ATP binding to Aurora A.
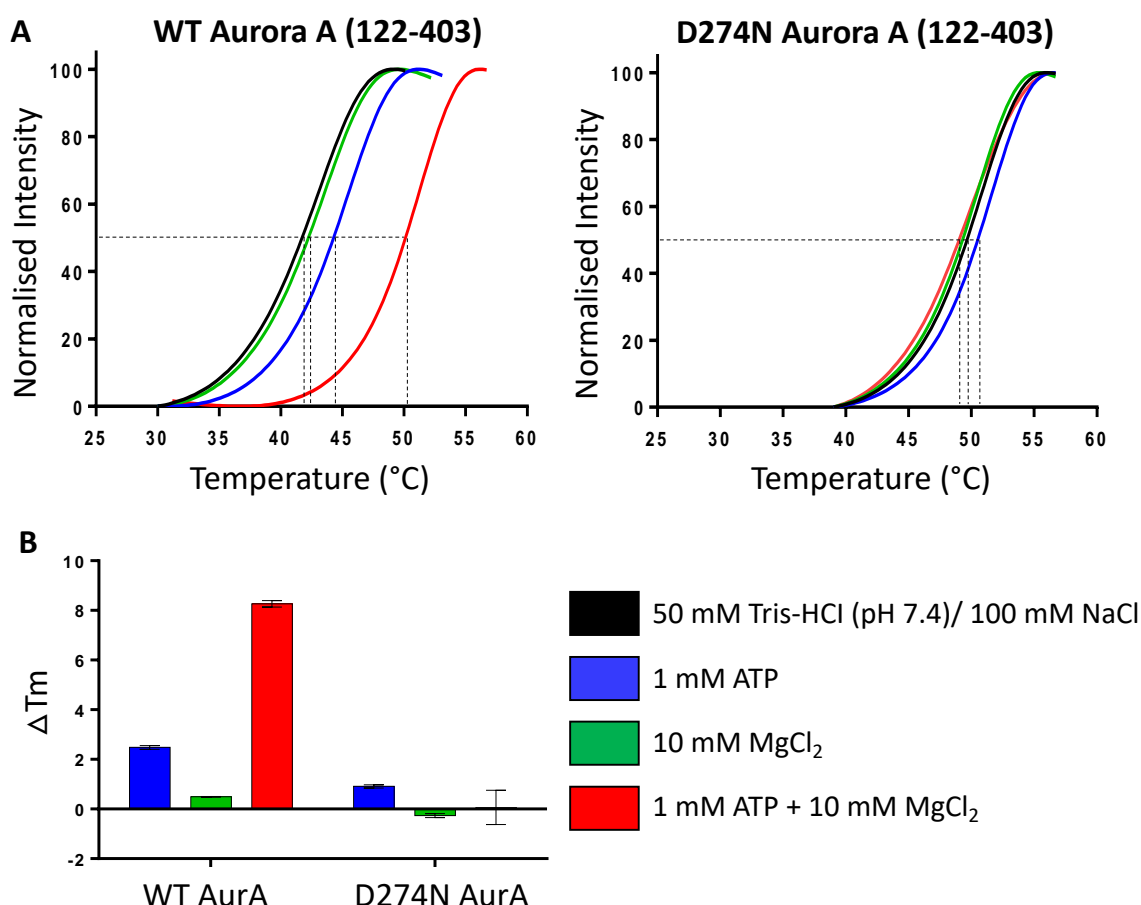
Figure 3.9. Catalytically active WT Aurora A can bind ATP/MgCl$_2$ in comparison to catalytically inactive D274N mutant. A) DSF thermal stability assay with 5 µM Aurora A (black), in the presence of 1 mM ATP (blue), 10 mM MgCl$_2$ (green), or 1 mM ATP + 10 mM MgCl$_2$ (red). **B)** Difference in melting temperature (ΔTm) compared with buffer control is presented for both WT and D274N Aurora A (122-403).

3.2.6 IM-MS suitable buffers

Native IM-MS was implemented in this chapter for the analysis of Aurora A due to this technique being able to evaluate protein structure under (near) physiological conditions, which enables the study of conformational flexibility and differences seen associated with phosphorylated versus non-phosphorylated protein (126). Prior to native IM-MS analysis of Aurora A, desalting of the protein was required to remove any interfering substances obtained from the purification steps, such as detergents used in buffers for protein stabilisation that could lead to signal suppression (177, 178). A suitable buffer must be volatile to be compatible for MS analysis and is required to maintain the noncovalent interactions of the native protein structure or complex in a concentration of approximately 1-20 µM (179). Ammonium acetate at pH 7 is the standard solution used for IM-MS experiments due to its volatility and thus lack of interference during ESI. While it does not function as a buffer per se at neutral pH, due to the NH_4^+ and $CH_3\text{-}COO^-$ not being a conjugate acid/base pair, the reduction in pH to 4.75 (pK_a of acetic acid) at the point of ionisation is not as extensive as might be observed with other solutions such as pure water (180). High concentrations (between 50 and 300 mM) can reduce the signal suppression effects of residual purification buffers, such as Tris (181).

Spin filter molecular weight cut-off columns are commonly used for desalting prior to native MS experiments, serving to concentrate the sample and removing small molecular weight non-volatile components of purification buffers that can cause interference (126). Several alternative solutions for native IM-MS were tested to determine whether they were more suitable in comparison to ammonium acetate. The use of ammonium acetate as a suitable solution can result in stability issues associated with the pH, leading to protein deactivation, unfolding and precipitation, which can reduce protein recovery (180).

To evaluate different solutions for native MS analysis of Aurora A for maximal protein recovery, while retaining 'native' conformation, purified protein was desalted into either Tris/ammonium acetate, Trehalose, or using a mixture of amino acids (Figure 3.11). The amino acid mixture (20 mg/mL solution of arginine, lysine, aspartic acid, and glutamic acid dissolved in 150 mM ammonium acetate) was used as a pre-wash coating step prior to normal desalting to prevent Aurora A from sticking to the spin column and resulting in a high volume

of sample loss. In addition to the use of the spin filter columns, 600 μL (4x the amount used for spin columns) of Aurora A was desalted using a column attached to a AKTA FPLC system to compare to the spin filter column method. The column was pre-equilibrated with ammonium acetate and fractions were collected in volumes of 0.3 mL. A Bradford assay was used to determine the fractions with the highest protein concentration, and these were then selected for IM-MS analysis. The Tris/ammonium acetate solution was tested to determine if the addition of Tris would provide more stability to the protein, as it is thought that the substitution for some ammonium acetate with Tris can stabilise the folded form of the protein (182). Similar to the addition of Tris, Trehalose was also tested as a solution that could provide enhanced stability to the folded form of the protein (183).

Analysis of protein recovery for each desalting method used demonstrated that desalting into ammonium acetate using the Amicon spin-filter columns resulted in the highest recovery at 57% (Figure 3.10). The Tris/ammonium acetate buffer had the next best recovery at 36%. However, the addition of Tris compromised ESI resulting in poor spectral quality (Figure 3.11C). Desalting into ammonium acetate using the AKTA FPLC system rather than the spin columns showed the lowest recovery, likely due to a high level of sample dilution and subsequent sample loss. The use of the amino acid mixture to pre-coat the spin columns prior to ammonium acetate desalting failed to enhance protein recovery, with only 17% recovered. Trehalose had a low recovery of 21% and resulted in poor resolution between the 11+ and 10+ charge states (Figure 3.11D). The poor resolution and quality of the charge distribution data could be due to the protein containing 10% glycerol, in order to keep the protein stable, however, this would result in the spraying conditions not being optimal. Alternative desalting methods could have been employed to remove the glycerol, such as dialysis. The amount of protein recovered for the different methods demonstrated that ammonium acetate resulted in the highest recovery, with the pre-coat amino acid method having the lowest (Table 3.2). The different desalting methods suggested that using 150 mM ammonium acetate in the spin filter columns resulted in the best quality data with the highest resolution between charge states in contrast to the other methods carried out (Figure 3.11).
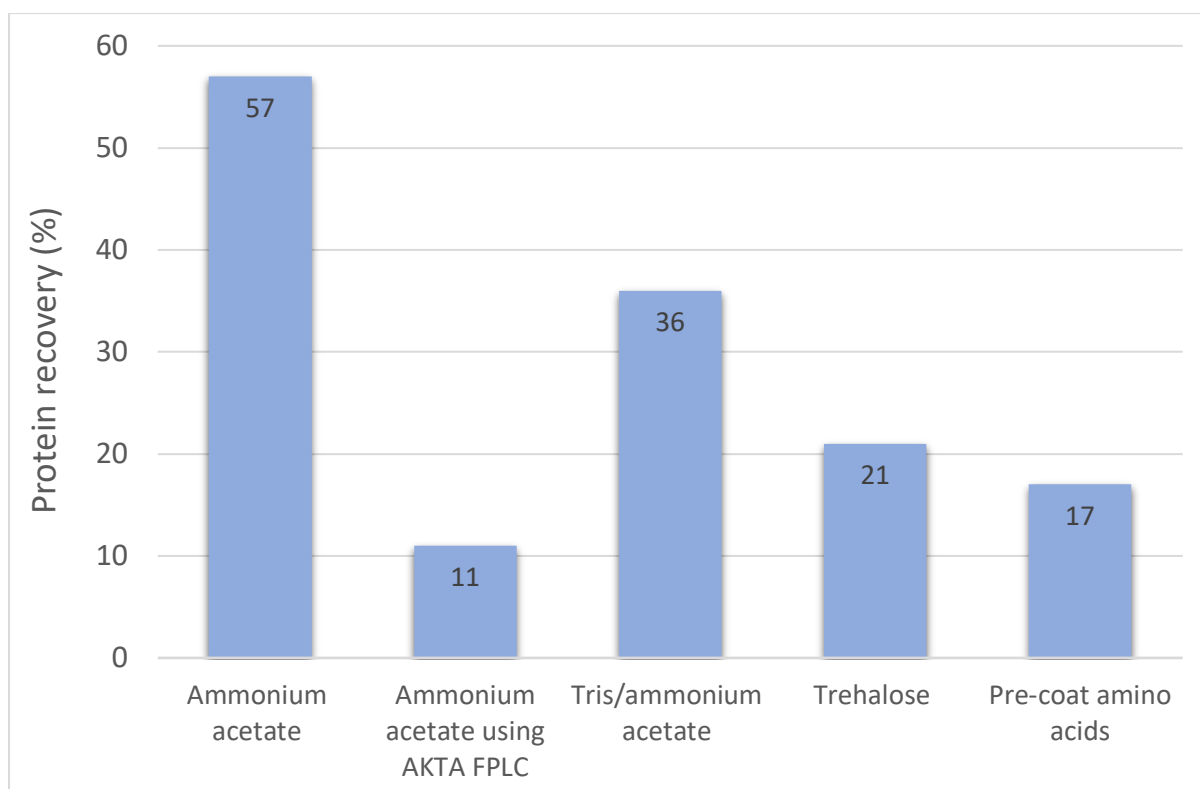
Figure 3.10. Protein recovery of different desalting methods used for IM-MS analysis. Recovery expressed as a percentage for each desalting method used. Protein concentration determined using Nanodrop following the desalting step and recovery calculated based on starting material. All methods (except for AKTA FPLC) used Amicon 10 kDa cut-off spin filter columns for 3x 10 min spins and an initial 10 min spin to pre-coat the filter with the appropriate buffer.

Buffer exchange method	Protein recovery (µg)	Volume (µL)
Ammonium acetate	313	80
AmAc using AKTA FPLC	59	300
Tris/ammonium acetate	194	77
Trehalose	112	75
Pre-coat amino acids	94	75

Table 3.2. Protein recoveries of desalting methods for IM-MS.
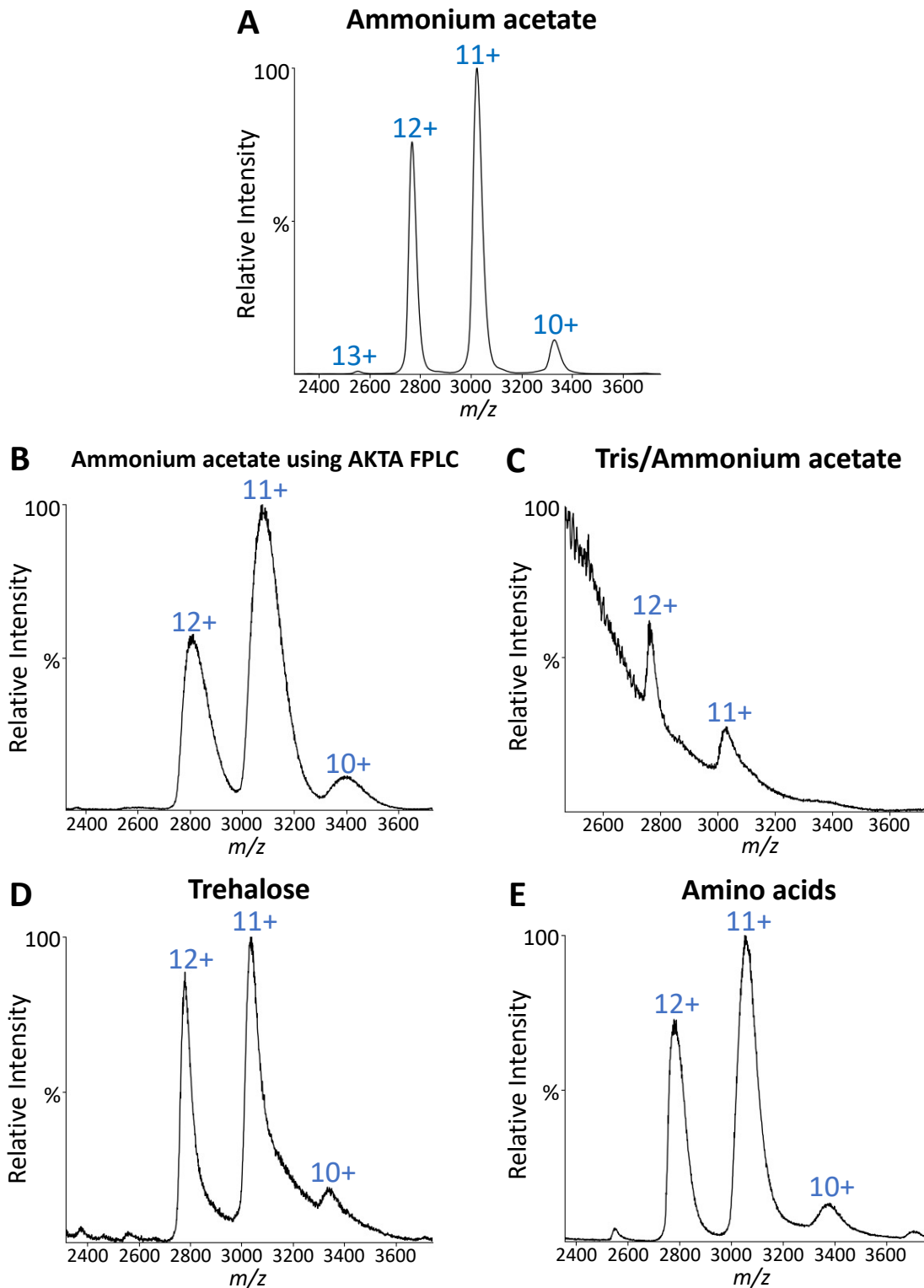
Figure 3.11. Analysis of suitable native MS desalting strategies for WT Aurora A. WT Aurora A was desalted into **A)** 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns. **B**) 150 mM ammonium acetate using the AKTA FPLC system. **C)** 50 mM Tris/150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns. **D)** 20 mM Trehalose using Amicon 10 kDa cut-off spin columns. **E)** 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns that had been pre-equilibrated in 150 mM ammonium acetate containing 20 mg/mL of amino acids (arginine, lysine, aspartic acid, and glutamic acid) prior to application of Aurora A.

3.2.7 Aurora A under native IM-MS conditions

Following the desalting step, WT and D274N Aurora A were subject to electrospray under 'native' conditions. The charge state distribution for both forms of Aurora A show three and four charge states for D274N and WT respectively. Native mass spectra with a limited number of charge states in a Gaussian distribution is indicative of a protein or protein complex with limited intrinsic disorder in contrast to a denatured protein that will have a high distribution of charge states (128). The 11+ and 12+ charge states were predominantly observed for WT protein with the 10+ and 13+ states being observed at much lower levels. Interestingly, for D274N Aurora A, while the relative ratio of the 10+ and 11+ charge states was as observed for WT protein, the relative abundance of the 12+ state was much lower, suggesting that the non-phosphorylated variant exhibited less flexibility than phosphorylated protein (Figure 3.12).
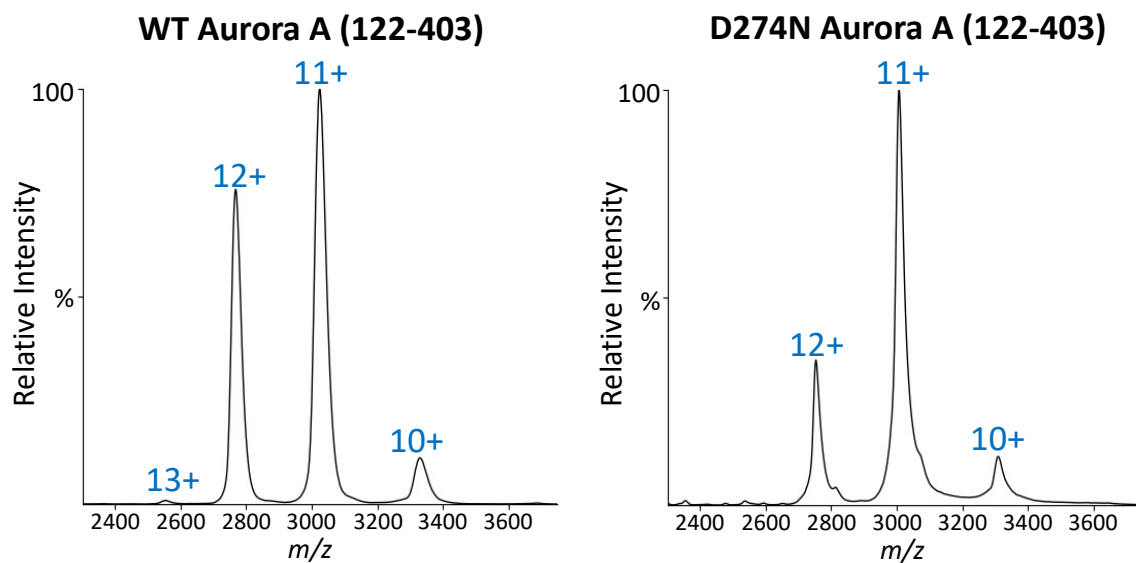
Figure 3.12. Aurora A native MS spectra. Native ESI mass spectrum of WT Aurora A (left) and D274N (right). Data analysis carried out using MassLynx 4.1 and figures created using UniDec.

3.2.8 CCS drift time calibration

In Travelling Wave Ion Mobility (TWIM), the calculation of the rotationally averaged collision cross section ($^{TW}CCS_{N2\rightarrow He}$) for protein analytes of interest is determined by calibration of the drift time using protein calibrants of similar charge state and CCS to the protein being analysed, as discussed in the introduction (1.5). To generate a drift time/CCS calibration curve, arrival time distributions (ATD) were acquired for multiple charge states for each of the three calibrant proteins, avidin, BSA and β-lactoglobulin, at the same travelling wave height and velocity as the protein of interest (184) (Figure 3.13). The ATD of a protein is the time it takes for each ion to travel through the mobility region of the mass spectrometer, with the measurement carried out on multiple ions providing a distribution of arrival times (185).

The reduced drift time was calculated by extracting the drift time at each different charge state alongside the *m/z* for the three calibrant proteins. The drift times as measured in the TWIMS device in the presence of nitrogen for each protein calibrant were converted to helium CCS using the Matthew Bush CCS database (150). Each protein has a known CCS value reported in the Bush database, which can be used to determine the reduced CCS value with the addition of the charge state at each drift time. The reduced drift time and CCS were then converted to their natural logarithm to gain a *y* value from the reduced CCS/drift time calibrant graph to use for determination of the CCS scale. The chromatogram list from the drift time window in MassLynx, the charge state and *m/z* of Aurora A were used alongside the *y* value from the graph to determine a CCS scale to calculate unknown CCS's for WT and D274N Aurora A.
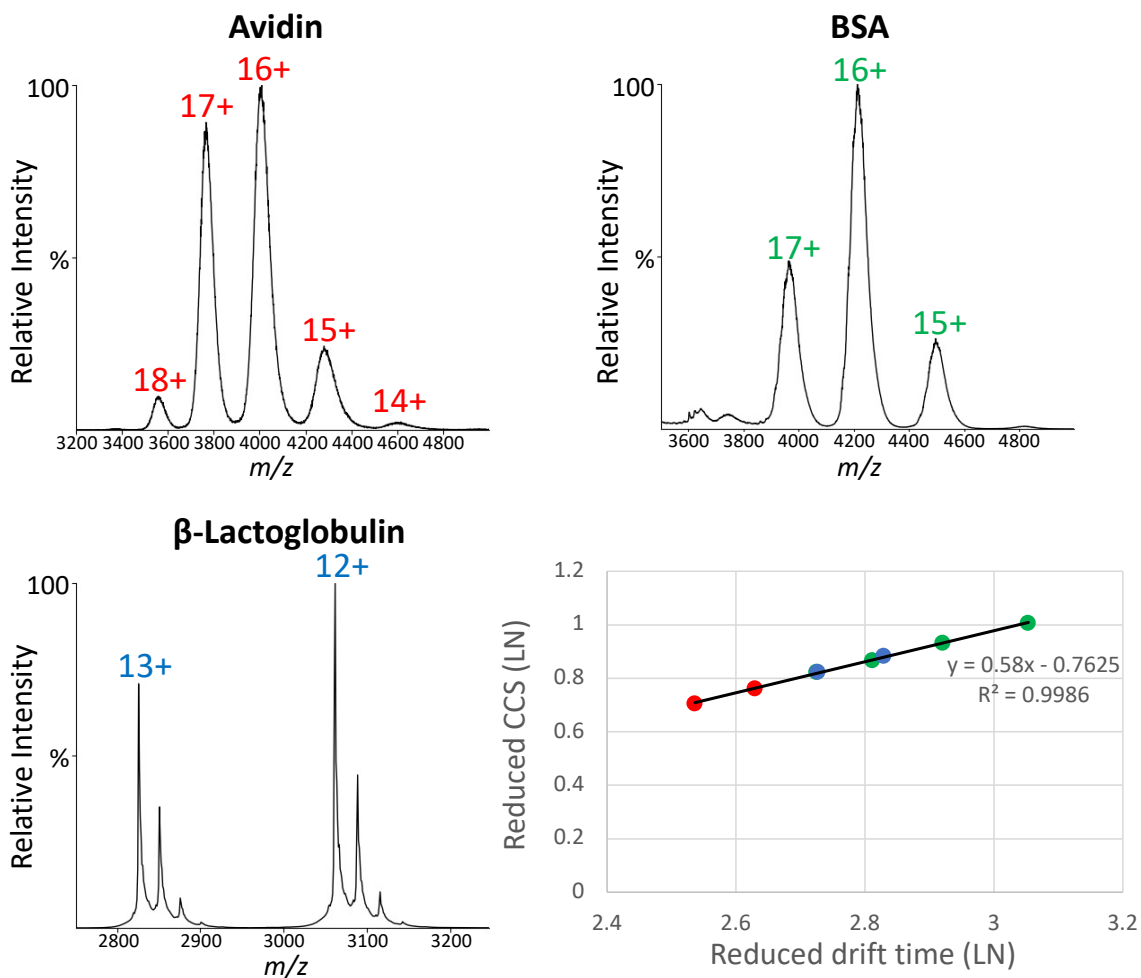
Figure 3.13. Protein calibrants used to determine CCS calculation for Aurora A. Protein calibrants were prepared as described in (186) at a concentration of 10 µM prior to calculating their drift time through the T-wave cell.

3.2.9 Aurora A CCS determination

Following drift time calibration, the CCS of WT and D274N Aurora A was determined using TWIMS for the 10+, 11+ and 12+ charge states (Figure 3.14). Each charge state was subject to quadrupole isolation and resulted in a wide CCS distribution for both WT and D274N Aurora A. The 12+ charge state yielded the widest CCS distribution from ~13 nm^2 to ~34 nm^2 for both forms of Aurora A. The 10+ and 11+ charge states were similar in their distribution of CCS, with the 11+ showing more detail in the lower CCS region for WT Aurora A. The 11+ charge state was chosen to be the charge state for further analysis due to it being the most abundant state of the three (Figure 3.12) and revealing sufficient detail across the full range of CCS distribution for both WT and D274N Aurora A. The smallest charge state (10+) would have been the favoured charge state to choose for further analysis due to the protein being in its most native state with the lowest number of protonated species on, however, the 10+ charge state was not sufficiently abundant in the D274N mutated variant, and therefore, analysis would have been challenging.
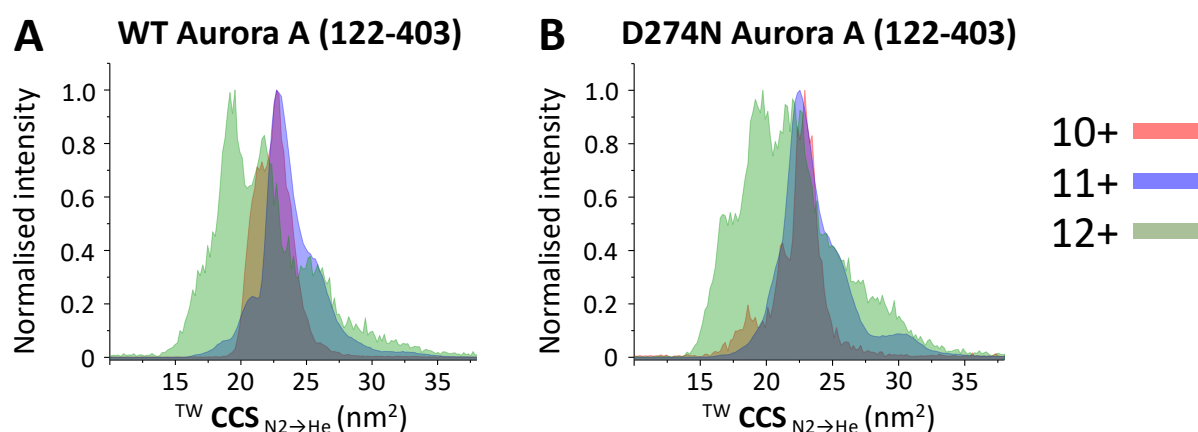
Figure 3.14. CCS determination for the three most abundant charge states of Aurora A. $^{TW}CCS_{N2 \rightarrow He}$ for the [M+10H] $^{10+}$, [M+11H] $^{11+}$ and the [M+12H] $^{12+}$ species of WT **(A)** and D274N **(B)** Aurora A (122-403). 10+ (red), 11+ (blue) and 12+ (green). Line plots were generated in Origin (Version 2021b).

Initial visualisation of the CCS represented in Figure 3.15A suggested some subtle differences at the lower CCS region (~22 nm^2) between WT and D274N Aurora A. To gain a better understanding into the different conformational states that contribute to the broad CCS distributions, Gaussian fitting in Origin (Version 2021b) was carried out to reveal a number of different conformers. The Fit Peaks Pro function was implemented to initially assign the most abundant conformer (II) and then further conformers were added for I, III and IV, where

applicable. The CCS, CCSD, and area parameters for each assigned conformer were manually adjusted using the Fit Control and iteration feature to obtain the best fit between experimental data (black line) and the sum of Gaussians (red line). To effectively fit the Gaussian data, an R^2 value of 0.995 and above was set as a reference point, however, due to the quality of the spectra, an R^2 value of above 0.900 was deemed acceptable for fitting. Black error bars are representative of the SD, and the error between experimental and sum of Gaussians was reported with R^2.

A

B

WT Aurora A (122-403)

$R^2 = 0.991$

Conformer	CCS (nm^2)	CCSD (nm^2)
I	20.4	1.5
II	22.5	1.9
III	24.8	3.3
IV	N/A	N/A

D274N Aurora A (122-403)

$R^2 = 0.996$

Conformer	CCS (nm^2)	CCSD (nm^2)
I	21.0	1.6
II	23.0	2.3
III	25.5	3.0
IV	30.5	4.2
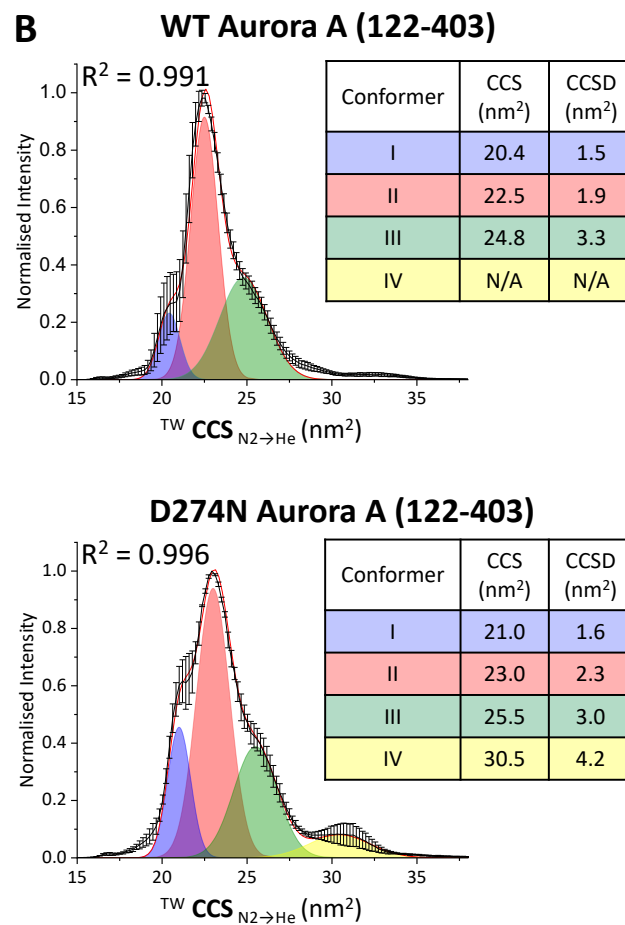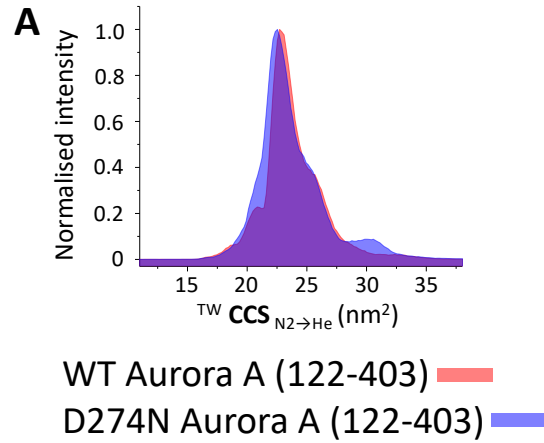
Figure 3.15. Active WT Aurora A (122-403) is more conformationally compact than inactive D274N Aurora A. A) $^{TW}CCS_{N2 \to He}$ for the [M+11H] [11+] species of WT or D274N Aurora A (122-403). Line plots were generated in Origin (Version 2021b). **B)** $^{TW}CCS_{N2 \to He}$ Gaussian fitting [M+11H] [11+] species of (top) WT or (bottom) D274N Aurora A (122-403). Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed in Origin (Version 2021b), with R^2 values listed.

The weighted average CCS for WT Aurora A (23.9 nm^2) was slightly larger than that observed for D274N (22.3 nm^2), which was consistent with the compared CCS calculated theoretical values (Table 3.3). IMPACT (Ion Mobility Projection Approximation Calculation Tool) was used to calculate an estimated CCS value with the use of structures obtained from the protein data bank (PDB). PDB code 1OL7 was used to estimate the WT Aurora A CCS based on this structure representing a phosphorylated version of Aurora A, with the activation loop present and only minor parts of the N-terminus and C-terminus missing. The 1OL6 structure was used due to it being a dephosphorylated structure and a good structural match for D274N Aurora A The modelling CCS values obtained from IMPACT are within the expected 10% range of the experimentally derived CCS's, which has been previously observed with different proteins (187).

	Experimental weighted average CCS (nm^2)	IMPACT CCS (nm^2)
WT Aurora A	23.9	25.3
D274N Aurora A	22.3	23.7

Table 3.3. Comparison of Aurora A CCS derived from IM-MS experimental values and IMPACT software. Experimentally derived CCS values were calculated based on the weighted averages of three independent replicates. UniProt human Aurora A databases 1OL7 and 1OL6 were used to determine IMPACT CCS values based on the trajectory method.

Although the WT form demonstrated a larger CCS value, the overall CCS half-height width distribution (CCSD) was narrower compared to D274N (Figure 3.15B). This would suggest that the D274N Aurora A protein had greater conformational flexibility in comparison to the WT protein. The CCS Gaussian fitting of this data set showed three conformers (I-III) that were observed in both WT and D274N Aurora A, with an additional fourth conformer (IV) present in the D274N protein (Figure 3.16B). The two prominent conformers (II and III) in both proteins were within the 3% variance expected for IM-IM experiments (WT - conformer II: CCS = 22.5 nm^2, CCDS = 1.9 nm^2; conformer III: CCS = 24.8 nm^2, CCDS = 3.3 nm^2, D274N - conformer II: CCS = 23.0 nm^2, CCDS = 2.3 nm^2; conformer III: CCS = 25.5 nm^2, CCDS = 3.0 nm^2,), with WT showing a larger abundance (54%) compared to D274N (48%) for conformer II (Figure 3.16A). In contrast, conformer I was higher in relative abundance in D274N (17%) compared to WT (11%) and there was an additional conformer IV (8%) that was not present for the WT protein. The lack of conformer IV for WT Aurora A suggests that phosphorylation and/or

activation status results in a more constrained conformational profile than that observed for D274N.
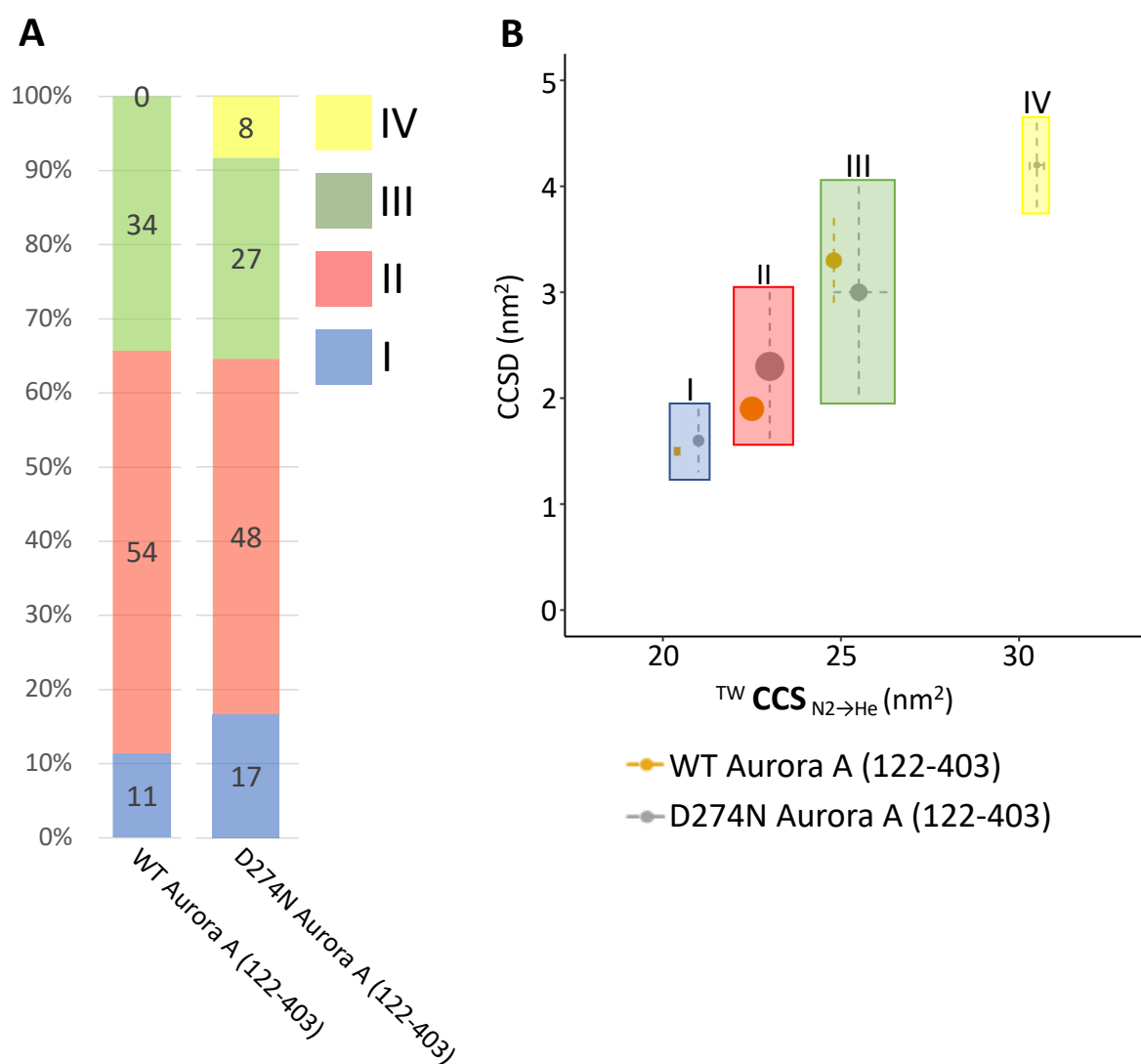
Figure 3.16. Proportional conformational space adopted by the four different conformers of WT and D274N Aurora A (122-403). Determined by Gaussian fitting in Figure 3.15 for three individual experiments. **A)** Percentage (%) area of different conformational states (I (blue), II (red), III (green), IV (yellow)). **B)** Proportional scatter plots (CCS (nm^2) versus CCSD (nm^2)) for the different conformational states. Size of dot representative of area.

3.2.10 Aurora A stability by CIU

To assess the stability of WT and D274N Aurora A, collision induced unfolding (CIU) analysis was performed, with applied collision energy (CE) sufficient to induce unfolding, but not to induce fragmentation (153). The CE was applied in two-volt intervals from 16-34 V within the trap region of the instrument to compare CCS unfolding profiles for both proteins. Data is

represented as CIU fingerprint plots (Figure 3.17A), and a comparison of WT and D274N across each CE step is represented in Figure 3.17B. A total of four unfolding transitions were present for both WT and D274N Aurora A, beginning with the folded state protein, which unfolds at ~24 V and ~26 V for WT and D274N respectively. The lower CE required to initiate unfolding for WT compared to D274N suggests that WT Aurora A is less stable in contrast to the non-phosphorylated form, which correlated with the DSF thermal stability findings presented earlier (Figure 3.9). Two partial unfolding intermediates were observed over a CCS range of ~23-29 nm^2 for both proteins. The final unfolded states were present between ~31-33 nm^2, with the D274N protein showing a larger CCSD at this state, similar to that seen in the previous CCS data set.
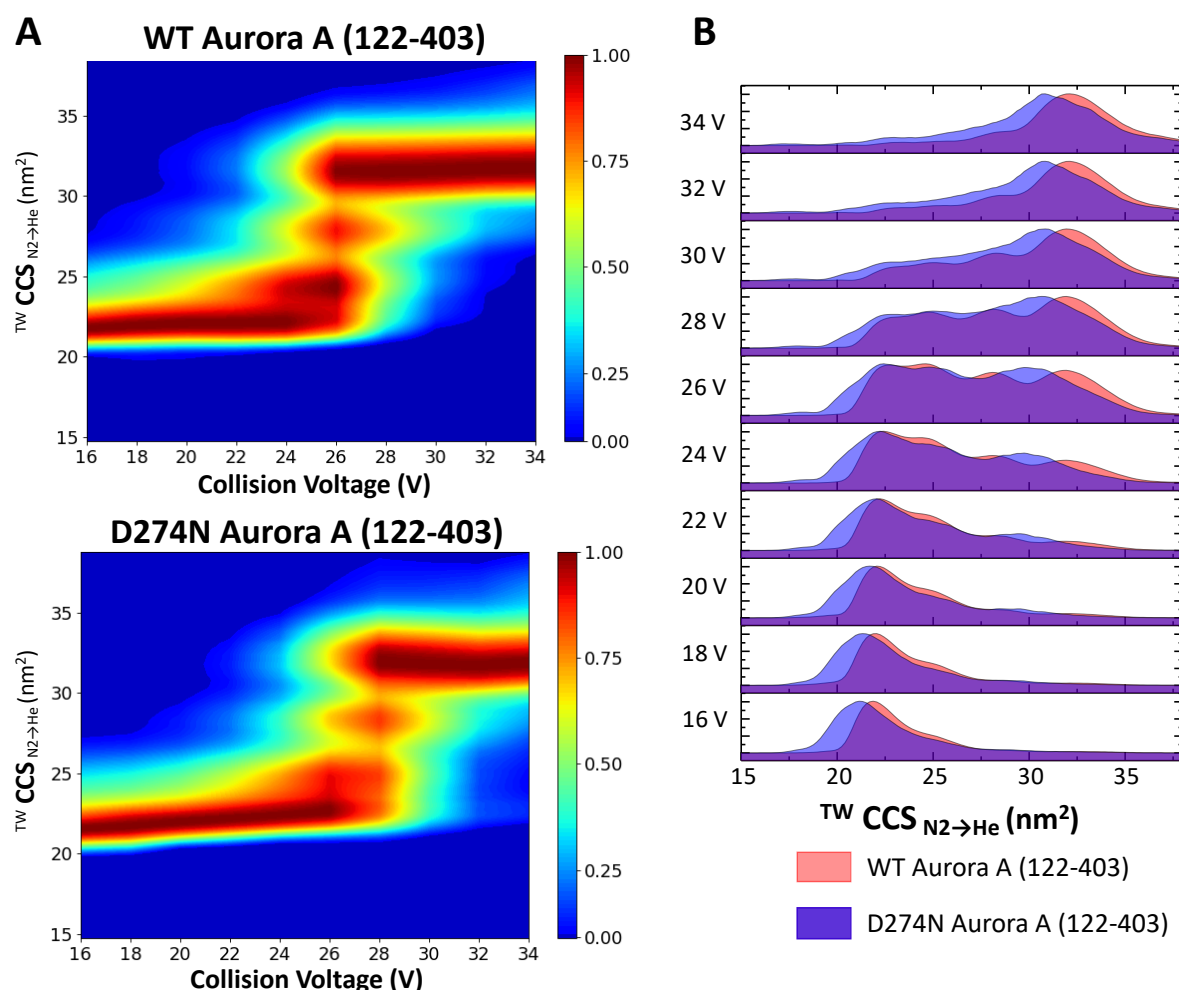
Figure 3.17. WT Aurora A has less kinetic stability than D274N Aurora A. A, B) Collision-induced unfolding profiles for the isolated 11+ charge state of WT (top) and D274N (bottom) Aurora A (122-403) (or overlaid in (**B**)). Stepped collision energy (CE) was applied between 16 and 34 V in two-volt intervals. Data analysis was carried out in MassLynx 4.1, (**A**) generating heat-maps using CIUSuite 2 and (**B**) mountain plots using Origin (Version 2021b). Presented are data from an average of 3 independent experiments.

3.3.   Conclusion

This chapter has explored the conformational dynamic differences of an active, phosphorylated protein in comparison to an inactive non-phosphorylated protein by IM-MS. WT Aurora A (122-403) and D274N Aurora A (122-403) were successfully purified to produce catalytically active and inactive Aurora A. Prior to IM-MS experiments, thermal stability and kinase enzyme assays confirmed the catalytic activity of WT Aurora A, demonstrating that this protein could bind ATP/Mg and phosphorylate a substrate peptide. WT Aurora A also had evidence of phosphorylation on the Thr288 residue required for catalytic activity, as well as five other sites of autophosphorylation, in contrast to no phosphorylation for D274N Aurora A.

IM-MS experiments revealed four primary conformational states for the two Aurora A proteins following Gaussian fitting of the CCS distribution. Differences in the ratio of these conformers and the CCSD values were evident between phosphorylated and non-phosphorylated Aurora A. Phosphorylated WT Aurora A demonstrated reduced conformational dynamics and flexibility in comparison to D274N Aurora A suggesting that phosphorylation partially constrains the protein. Further exploration of the stability of the two proteins by DSF and CIU revealed that the non-phosphorylated form is more stable than phosphorylated Aurora A.

This chapter has successfully utilised IM-MS to determine subtle differences associated with active/phosphorylated proteins. Chapter 4 will investigate the impact of small molecule inhibitors and peptides on the phosphorylated and non-phosphorylated Aurora A proteins.

Chapter 4. Effect of small molecule inhibitors on the conformational dynamics of Aurora A

4.1   Introduction

Chapter 3 discussed the effects of phosphorylation on the conformational landscape and dynamics of Aurora A, comparing the active WT phosphorylated protein with an inactive D274N, non-phosphorylated variant, by IM-MS. This chapter explores the effect of different small molecule Aurora A inhibitors and activating peptides on the structure, stability, and conformation dynamics of these two proteins using IM-MS.

Protein kinase inhibitors can target a range of active or inactive conformations, with the two major classes termed DFG-in and DFG-out. The DFG motif is situated within the kinase activation loop and plays an important role in catalysis. Active conformations enable the Asp of the DFG motif to orientate to a position where it can bind magnesium and interact with ATP. Inactive kinase conformations have a collapsed activation loop, which blocks substrates from binding. A DFG-in configuration can be a variation of active or inactive conformations, with the Phe residue contacting the C-helix of the N-terminal lobe. The DFG-out configuration is associated with the inactive conformation, with the Phe residue occupying the ATP pocket, which results in the exposure of the C-helix pocket (8).

In addition to the DFG-in and DFG-out inhibitor binding mechanism classifications of inhibitors, different groups have classified further structures based on the positioning of the DFG motif. Modi et al have classified structures based on the position of the Phe residue and the backbone dihedral angles of the residues X-D-F (X = residue prior to DFG motif) to recognise a DFG-inter conformation alongside the DFG-in and DFG-out conformations. This clustering system recognised eight distinct conformations, including six that could be classified into different configurations of the DFG-in conformation (8). The Möbitz group classified all the mammalian kinases present in the Protein Data Bank (PDB) using the dihedral angles of four Cα atoms from the residues that make up the DFG motif and their distance from the C-helix, revealing twelve classifications (188). Ung et al identified five groups by classifying 3,708 kinase structures present in the PDB by assessing the αC helix distance to

the DFG motif, which provided structural information on the conformational space of the catalytic domain of the protein kinases (56).

Levinson et al used a Förster resonance energy transfer (FRET) sensor to track structural movements of the activation loop of Aurora A using a range of inhibitors (57). The analysis revealed a variety of conformational preferences based on a DFG-in or DFG-out configuration, with each inhibitor adopting the configuration to a different extent (Figure 4.1). The DFG-in inhibitors favoured binding to Aurora A when it was constrained in the DFG-in configuration by the binding of the activating TPX2 peptide. DFG-out inhibitors favoured binding to Aurora A when it was phosphorylated on the activation loop (57).
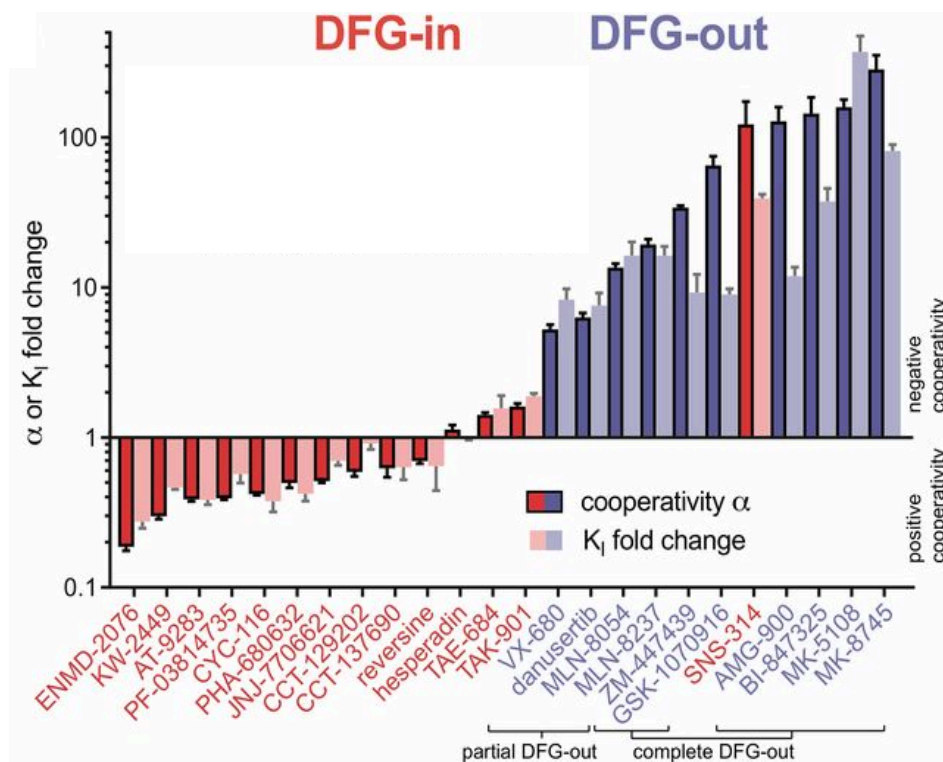
Figure 4.1. FRET analysis tracks structural movements of the activation loop of Aurora A to differentiate between the DFG-in and DFG-out inhibitors. Cooperativity α-values determined by the relative shift towards the DFG-in or DFG-out configuration by global fitting of fluorescence binding data. Figure reproduced from (189).

Based on the Levinson analysis, the two inhibitors that favoured the most extreme DFG-in (ENMD-2076) and DFG-out (MK-8745) configurations were initially analysed. In addition, the two partial DFG-out inhibitors, MLN8237 and VX-680, were also chosen for analysis. These

four inhibitors are all selective to Aurora A, therefore, a generic protein kinase inhibitor, staurosporine (190), was also chosen for comparison.

Aurora A can be activated by either phosphorylation or binding of an activating peptide derived from the microtubule-associated protein TPX2. Crystal structures have previously been published of Aurora A in its phosphorylated form with a 43-length residue of TPX2, which enables activation and protection from dephosphorylation. The binding of TPX2 has been shown to cause no global conformational change, however, the peptide pulls on the activation segment, which orientates the T288 activating phosphorylation site into a buried position that locks the kinase into an active conformation (35).

4.1.1 Aims

The aim of this chapter was to analyse the effects of inhibitor compounds and activating peptides on the structure and dynamics of Aurora A by IM-MS. The conformational landscape of the two proteins bound either to small molecule inhibitors or activating (or control) peptide were assessed to determine any differences in conformational flexibility using CCS determination. The stability of inhibitor-bound complexes by collision-induced unfolding were also explored. Experimental data were compared with computational modelling (performed by Dr Matthew Bachelor at The University of Leeds). A key aim was to determine whether native IM-MS was able to distinguish binding modes of the different classes of inhibitor.

4.2   Results & Discussion

4.2.1 Effect of small molecule inhibitors on CCS distribution

To determine the effects of small molecule inhibitors on the conformational landscape of Aurora A, a series of IM-MS experiments were performed to determine the CCS distribution of Aurora A in the presence of the five small molecule inhibitors described above (MLN8237, VX-680, ENMD-2076, MK-8745 and staurosporine). The inhibitors were selected based on their previously characterised mode of action resulting in either DFG-in or DFG-out configurations. Staurosporine is a generic type I inhibitor, which favours the DFG-in conformation (191). With exception to staurosporine, the remaining four inhibitors are specific Aurora A inhibitors (64). MLN8237 and VX-680 are both understood to favour a partial DFG-out conformation (57). ENMD-2076 has been shown to favour a complete DFG-in position, whereas MK-8745, has shown to adopt the complete DFG-out position (57).

Following desalting of the proteins into 150 mM ammonium acetate as previously described in Chapter 3, 10-fold molar excess of each inhibitor was added to either WT or D274N Aurora A and incubated for ten minutes at room temperature prior to IM-MS analysis. Each Aurora A protein with inhibitor complex was then subject to native electrospray ionisation to determine whether the inhibitor had bound under these conditions. As expected, incubation of Aurora A induced an increase in mass compared with Aurora A alone, demonstrating inhibitor binding (Figure 4.2). The TICs of D274N Aurora A + inhibitor bound observed less resolution between the charge states when compared to WT Aurora A due to the difficulty in obtaining good spraying stability for this variant protein. The inhibitors ranged from 375 Da to 519 Da, therefore a small increase in *m/z* region of the MS profile would be expected to indicate inhibitor binding. The increase in mass shift suggests that each inhibitor had bound and the 10-fold excess with a ten-minute incubation was sufficient for this analysis. Upon inhibitor binding, the charge state distribution altered in comparison to no inhibitor bound, in particular with VX680 and ENMD-2076, where the 10+ charge state was more prevalent. This could be due to the addition of the inhibitor causing a change in the pK values of the ionised groups. Binding of the inhibitor compounds can also cause some burial of parts of the protein, which would result in some charge states not being available for ionisation as well as others, and therefore a different charge state distribution profile would be observed (192).

The most abundant charge state (11+) observed for Aurora A in the absence of inhibitor remained present in the inhibitor-bound complexes. Therefore, the 11+ charge state for each inhibitor-bound complex was quadruple isolated to determine a CCS profile following drift time calibration for comparison with the unbound protein. Average IM-MS profiles from three replicate analyses were used for Gaussian fitting of Aurora A individual conformational states using Origin (Version 2021b) (Figure 4.3). The Fit Peaks Pro function was implemented to initially assign the most abundant peak (conformer II). Additional peaks were added for conformers I, III, and IV, where applicable. The CCS, CCSD, and area parameters for each assigned conformer were manually adjusted using Fit Control and the iteration feature, to enable the best fit between experimental data (black line) and sum of Gaussians (red line).
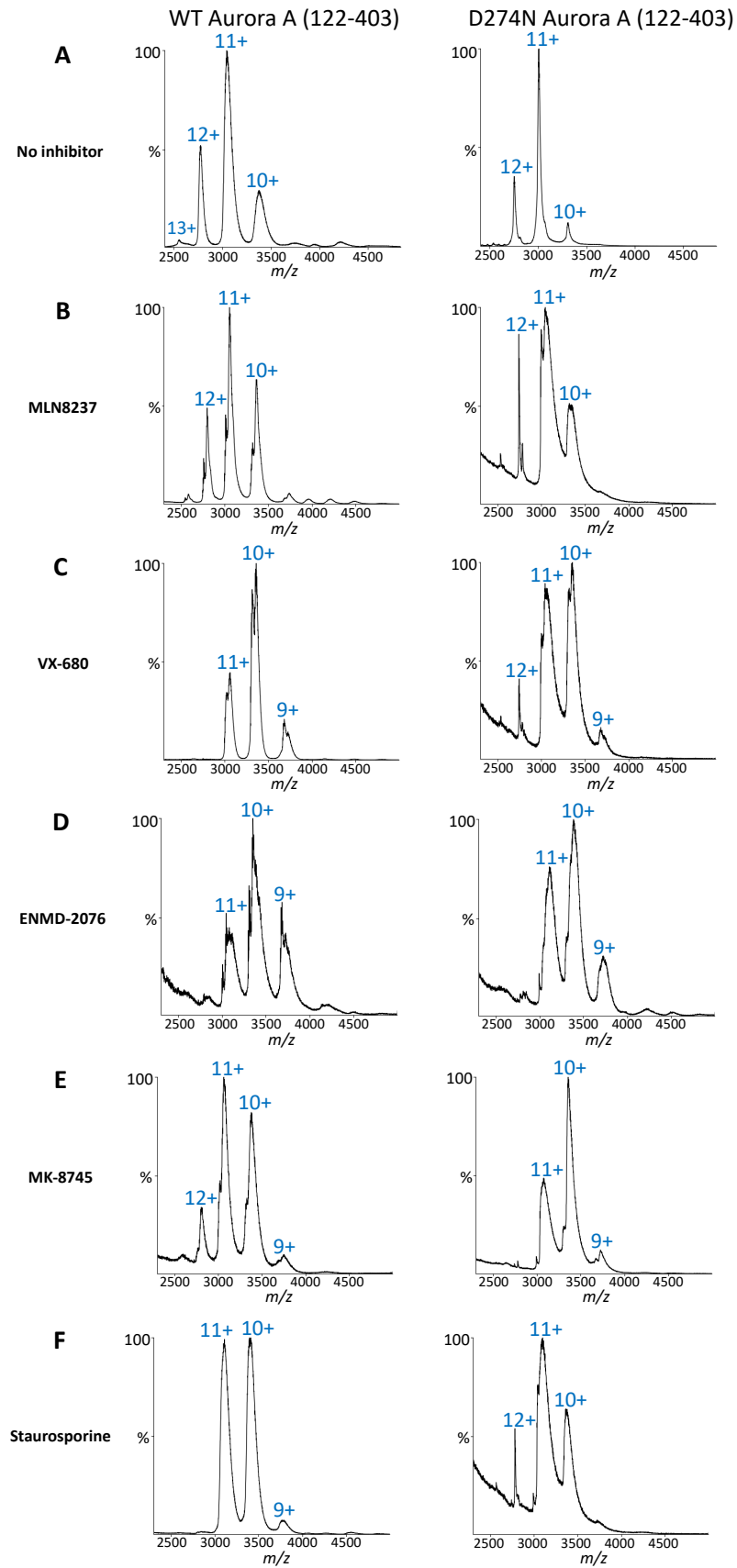
Figure 4.2. Inhibitor compounds bind to WT and D274N Aurora A. Mass spectra chromatograms of **A)** WT (left) or D274N (right) Aurora A (122-403) in the presence of 10-molar excess of **B)** MLN8237, **C)** VX-680, **D)** ENMD-2076, **E)** MK-8745, or **F)** staurosporine.

Staurosporine, the generic inhibitor, and MLN8237 (DFG-out) induced the greatest conformational change of WT Aurora A in the presence of an inhibitor, with only conformers II and III being observed (Figure 4.3). With exception to MLN8237 and staurosporine, all other inhibitor-bound complexes observed the smallest conformer (I), with a CCS of 19.8-20.1 nm^2, which was similar when compared to WT Aurora A alone (20.4 nm^2). The most abundant conformer (II) was observed in all inhibitor-bound complexes in addition to WT Aurora A alone. The CCS for conformer II was consistent between inhibitors with CCS values ranging from 22.7-23.0 nm^2 compared to 22.5 nm^2 in the absence of inhibitor. Conformer III was also present in all conditions ranging from 25.1-25.7 nm^2 for inhibitor complexes and 24.8 nm^2 for WT Aurora A alone. Both conformers II and III showed a marginal increase in CCS for inhibitor-bound complexes compared to WT Aurora A alone. However, overall, the increase was within the 3% variance expected between CCS values (with exception of conformer III for VX-680 and MK-8745), and therefore the increase was not sufficient to suggest that the inhibitor-bound complexes induced a larger size in the conformation of WT Aurora A.

The CCSD for conformer I of WT Aurora A (1.5 nm^2) in the absence of inhibitor was smaller than the equivalent conformational state for all inhibitor-bound complexes, with values ranging from 2.0-2.7 nm^2. These values are all above the 3% variance expected, which suggests that VX-680, ENMD-2076 and MK-8745 increase the conformational flexibility of the protein. ENMD-2076, MK-8745 and staurosporine had little effect on the CCSD value for conformer III (3.4-3.5 nm^2) compared to WT Aurora A alone (3.3 nm^2). However, the two partial DFG-out inhibitors, MLN8237 and VX-680, observed CCSD values for conformer III of 2.9 nm^2 and 2.8 nm^2. The smaller CCSD values for these two inhibitors outside of the 3% variance suggests that these inhibitors constrain the conformational landscape of WT Aurora A.

Upon addition of the inhibitor compounds to D274N Aurora A, the protein adopts a different conformational landscape in comparison to D274N Aurora A alone or WT Aurora A with inhibitors. Binding of inhibitors to D274N Aurora A results in the absence of conformer I in all cases, in contrast to this conformer being present when WT Aurora A was bound to VX-680, ENMD-2076 and MK-8745. Conformer II remained the most abundant conformer observed for all inhibitors bound to D274N Aurora A (23.2-23.4 nm^2), with a similar CCS observed for

D274N Aurora A alone (23.0 nm^2). Little variation was observed for CCS values of conformer III between D274N Aurora A alone (25.5 nm^2) and inhibitor-bound complexes (25.4-25.6 nm^2). Similar to the CCSD values observed for conformer II, conformer III also showed a wide range in values when inhibitors were bound (2.6-3.5 nm^2) compared to 3.0 nm^2 for D274N Aurora A alone. Smaller CCS values were observed for conformer IV in all inhibitor-bound complexes ranging from (29.3-30.2 nm^2 – staurosporine N/A) compared to D274N Aurora A alone (30.5 nm^2). However, with exception to MK-8745, all the values observed for conformer IV were within the 3% variance and therefore, suggests that binding of inhibitors does not increase the size of the protein conformation adopted by D274N Aurora A. The 3% variance between CCS values has become a standard reference point for these types of native IM-MS experiments (193). CCSD values observed for conformer IV were also within a wide range between 3.6-5.0 nm^2 for inhibitor bound complexes compared to 4.2 nm^2 for D274N Aurora A alone. Although subtle differences were observed between binding of inhibitors for WT and D274N Aurora A, no significant observations were apparent between the different modes of inhibitors that were analysed. Thereby, making it challenging to observe any inhibitor structural effects by the determination of CCS Gaussian fitting.
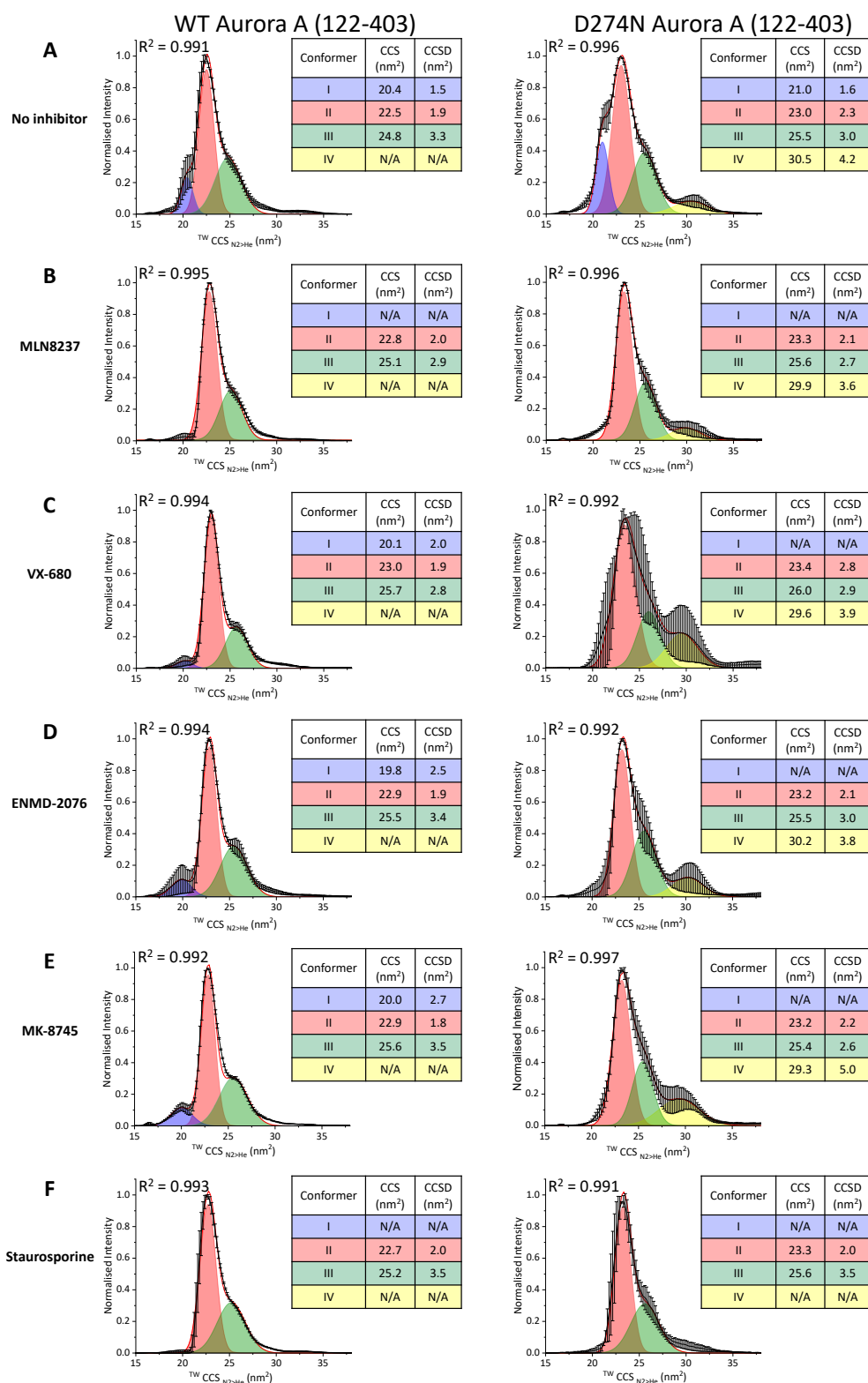
Figure 4.3. CCS distribution of inhibitor bound WT and D274N Aurora A. $^{TW}CCS_{N2\rightarrow He}$ of the [M+11H] $^{11+}$ species of **A)** WT (left) or D274N (right) Aurora A (122-403) in the presence of 10-molar excess of **B)** MLN8237, **C)** VX-680, **D)** ENMD-2076, **E)** MK-8745, or **F)** staurosporine. Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed using the Fit Peaks Pro function in Origin (Version 2021b), with R^2 values listed.

To further investigate the differences between the conformational landscape that WT and D274N Aurora A adopt in the presence of inhibitors, the relative proportion of each conformer was determined, as shown in Figure 4.4. The dot plots in Figure 4.4A/B provide an overview of the conformers adopted by WT and D274N Aurora A in the presence of inhibitors, showing the main difference to be the absence of conformer IV with all WT Aurora A conditions. The proportional plots of relative abundance for each conformer (Figure 4.4C/D) provide a further insight into the % of conformer landscape adopted by Aurora A in the presence of inhibitors. The abundance of conformer I decreases in the presence of VX-680, ENMD-2076 and MK-8745 for WT Aurora A. Additionally, conformer I is fully absent for WT Aurora A in the presence of MLN8237 and staurosporine and was not observed for any inhibitor complexes of D274N. Conformer II was observed to be the most abundant conformer in all conditions, ranging from 48-68%. The relative abundance of conformer II increased in the presence of all inhibitor complexes (exception to WT Aurora A + MK-8745), with the largest increase observed for WT Aurora A in the presence of VX-680 (68%) and D274N Aurora A in the presence of staurosporine (63%). The abundance of conformer III in the presence of inhibitors showed small changes for both WT and D274N Aurora A. However, no common trend was observed where all inhibitors induced either an increase or decrease in the relative abundance of this conformer. Conformer IV abundance increased in the presence of inhibitors (with the exception of staurosporine) for D274N Aurora A, ranging from 9-26% of the conformational landscape, compared to 8% in the presence of no inhibitor. Interestingly, both WT and D274N Aurora A had the same relative abundance for conformers II and III in the presence of the most generic inhibitor, staurosporine. This could be due to staurosporine not being selective for Aurora A and therefore having little influence on overall structure irrespective of activation status. Overall, the subtle changes in conformer relative abundancies that WT and D274N Aurora A adopt in the presence of inhibitors are not significant enough to distinguish between the different modes of inhibitors analysed in this study.
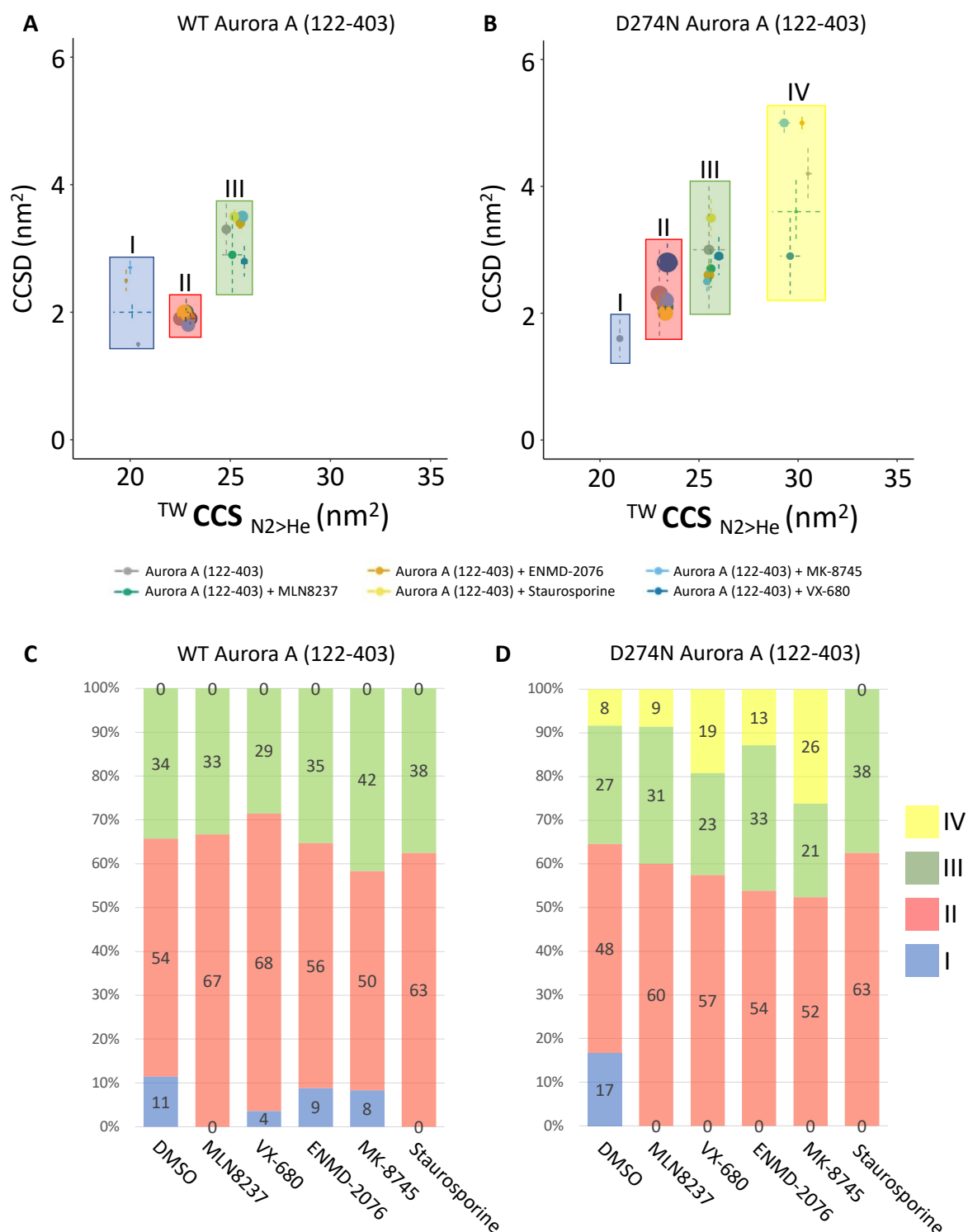
Figure 4.4. Conformational space adopted by Aurora A alone and in the presence of different inhibitors suggests inhibitors stabilise WT Aurora A. A, B) Proportional scatter plots (CCS (nm^2) versus CCSD (nm^2)) for the different conformational states (as determined by Gaussian fitting in Figure 4.3) for WT **(A)** and D274N **(B)** Aurora A (122-403). Size of dot representative of area. **C, D)** % area of four different conformational states (as determined by Gaussian fitting in Figure 4.3): I (blue), II (red), III (green), IV (yellow) for WT **(C)** and D274N **(D)** Aurora A (122-403) alone or in the presence of different inhibitors as indicated. Average % area presented from 3 individual experiments.

4.2.2 Molecular modelling of CCS distributions

To gain a further understanding into the meaning and relevance of the different conformers, mathematical modelling was carried out in collaboration by Dr Matthew Batchelor, a postdoctoral research associate in the group of my co-supervisor, Prof. Richard Bayliss at the University of Leeds. IMPACT, a model used to determine CCS values from known crystal structures, was used to calculate CCS using a combination of Aurora A models. Aurora A crystal structures obtained from the PDB were modified to add in any missing parts of the protein structure for Aurora A (122-403).

The PDB crystal structures included codes 1OL7, 5L8K and 6HJK, which all represented structures as DFG-in, DFG-up, and DFG-out, respectively (Table 4.1). CCS values obtained from these structures were 22.7, 22.9 and 23.3 nm^2 for DFG-in, DFG-up, and DFG-out, respectively. These CCS values were a good match to the values obtained from the experimentally derived conformers II (WT = 22.5 nm^2, D274N = 23.0 nm^2) and III (WT = 24.8 nm^2, D274N = 25.5 nm^2). PDB code 4C3P was also used, which is a dephosphorylated Aurora A structure in a dimer form bound to the activating TPX2 peptide, with the A-loop and αEF helix adopting an 'open' configuration where it extends out from the kinase domain in a DFG-in conformation. The 4C3P model demonstrated a monomeric CCS value of 24.9 nm^2, which was a good match for conformer III of the inhibitor-bound complexes. The CCS distributions from IMPACT for the all-atom simulation model were overall a good fit to the IM-MS experimentally derived values for conformers II and III, with a difference of 5-10% as has previously been observed between experimental and calculated values (187).

PDB	[1]CCS (nm²)	Crystal description
1OL5	22.6–23.0	Phosphorylated, TPX2, DFG-in
1OL6	22.3–22.8	Dephosphorylated, D274N, DFG-in
1OL7	22.2–22.7	Phosphorylated, DFG-in
1OL7N	22.1–22.6	Based on 1OL7, dephosphorylated, D274N, DFG-in
3E5A	22.5–22.9	Phosphorylated, VX-680, TPX2, DFG-in
4C3P	25.0–25.5	Dephosphorylated, TPX2, dimer, DFG-in, A-loop open
4C3P (alternative)	24.3–24.8	Dephosphorylated, TPX2, dimer, DFG-in, A-loop open
4J8M	23.1–23.7	Dephosphorylated, T287D, CD532, DFG-in
4CEG	21.1–21.6	Phosphorylated, C290A, C393A, DFG-in
5G1X	21.8–22.2	Phosphorylated, N-MYC, DFG-in
5ODT	23.2–23.7	Dephosphorylated, D274N, TACC3, DFG-in
5ODT (alternative)	23.9–24.3	Dephosphorylated, D274N, TACC3, DFG-in
1MUO	23.2–23.6	Dephosphorylated, adenosine, DFG-up
2WTV	23.4–23.7	Phosphorylated, MLN8054, chain A, DFG-up
2WTV	22.6–23.0	Phosphorylated, MLN8054, chain B, DFG-up
4JBQ	21.9–22.5	Dephosphorylated, VX-680, DFG-up
5EW9	22.8–23.4	Phosphorylated, MK-5108, DFG-up
5L8K	22.3–22.7	Phosphorylated, vNAR, DFG-up
6HJK	22.9–23.4	Dephosphorylated, L210C, ASDO2, DFG-out
6HJK (alternative)	23.0–23.4	Dephosphorylated, L210C, ASDO2, DFG-out
4F5S	39.3–40.1	BSA, calibrant

Table 4.1 Collision cross section estimates made using IMPACT on a single initial structural model of Aurora A (122–403) for each of the PDB codes. DFG-in models are grouped in blue, DFG-up (or inter) models are grouped in yellow, and two alternatively built starting structures for the only available DFG-out Aurora A structure (6HJK) are green. [1] As IMPACT is a stochastic method, CCS estimates were performed 200 times on each structure to give a range of values.

A simplified Gō model was used alongside the IMPACT all-atom model for comparison of calculated CCS values. In the Gō model, each residue is thought of as a single 'bead', with stabilising interactions made from contacts with the initial structure (164). Simulations using the Gō model were determined on the 1OL7 PDB structure, with CCS values showing a close match to conformers II and III from the IM-MS experiments (Figure 4.5A). The simulations failed to represent the two smaller conformers (I and IV), which could possibly be due to conformer I arising from experimentally-induced compaction and conformer IV representing a low prevalence configuration that is yet to be determined from the static crystal structures deposited in the PDB.

Further analysis of the CCS values determined from IMPACT using the all-atom sim model separated the CCS values into eight different distributions (Figure 4.5B). A D1/D2 plot, used by Modi et al (8), was implemented to distinguish between the different DFG configurations in the CCS distribution (Figure 4.5C). The DFG-in conformation, which is represented as the density of the upper left quadrant in the D1/D2 plot, was seen to be present in all eight regions of the CCS. The DFG-up/inter conformations, represented as the lower left quadrant, were also present in all eight regions, although less distinct at the highest CCS region shown in purple. The uncategorised DFG configuration that is not observed in the crystal structures with the more 'open structure', represented by the upper right quadrant, is more prevalent at the higher CCS regions. These CCS values are a close match to conformer III of Aurora A, suggesting this conformer represents the 'open' structures. The broadness of conformer III would suggest that it is composed of a mixture of 'open' configurations that could not be fully resolved during the IM-MS experiment. In contrast to the 'open' conformer III, conformer II is more of a 'closed' structure, where the A-loop is inward facing. The increase in the relative abundance of conformer II for inhibitor-bound complexes (Figure 4.4), suggests that the inhibitors constrain the complex into a more 'closed' structure for WT Aurora A. These findings would suggest that the experimentally derived conformers II and III are not representing different DFG configuration, rather, they are instead larger scale conformational changes.
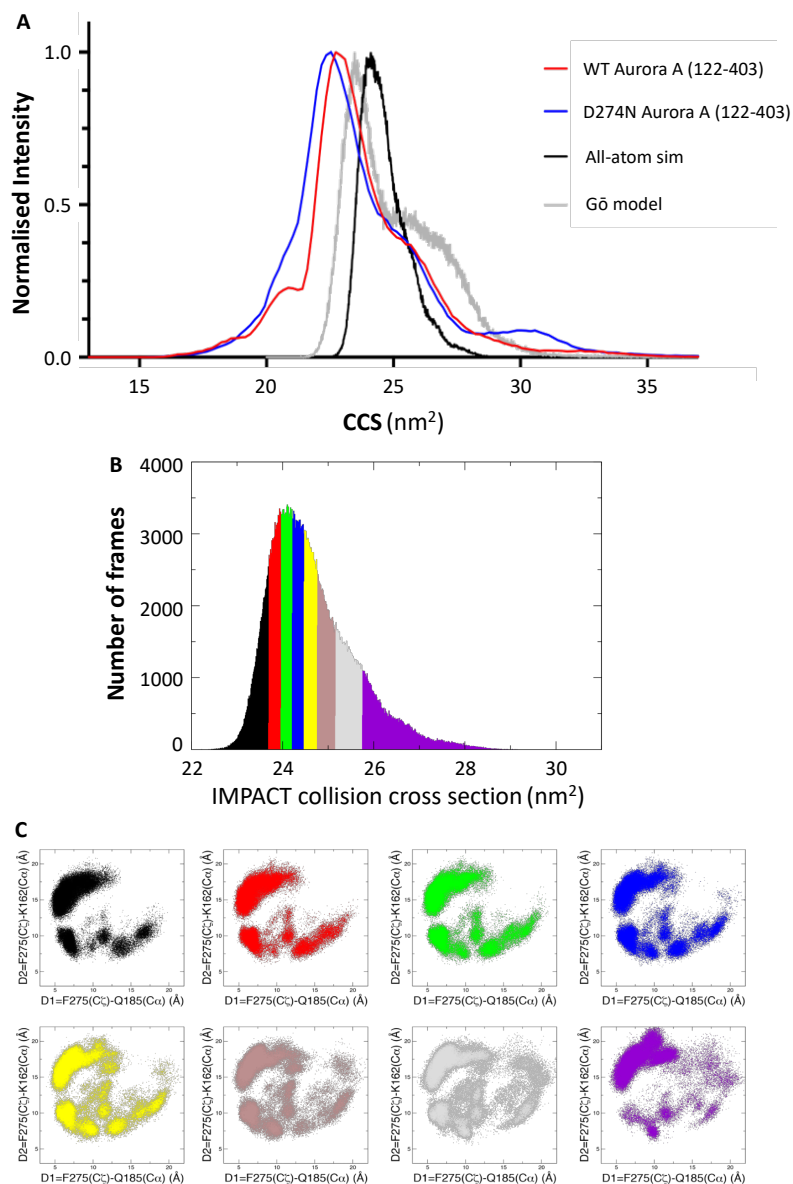
Figure 4.5. Comparison of CCS distribution from experimental IM-MS analysis and molecular modelling. A) Overlaid $^{TW}CCS_{N2\rightarrow He}$ for WT (red), D274N (blue) Aurora A (122-403), and overall distribution from all-atom simulations (black) and Gō model (grey). Four independent Gō model simulations show two components that are similar to conformers II and III. Simulation trajectory frames were separated into two groups: first peak (<24 nm^2, predominately 'conformer II') and second peak (>26 nm^2, predominately 'conformer III'). **B)** CCS distribution from the all-atom sim model is divided into eight regions with ~80,000 structures in each region. **C)** DFG motif conformations in each region are described using a D1/D2 plot. DFG-in conformation (upper left quadrant of D1/D2 plot) is present in all eight regions on the CCS distribution profile. DFG-up/inter and DFG-out conformers (lower left quadrant) are present in all eight regions but less observed at the highest CCS region (purple). The upper right quadrant of the D1/D2 plot is more observed in the higher CCS regions, suggesting that uncategorised DFG motif conformers (not observed in crystal structures) with a more 'open structure' contribute to the high CCS range of the profile.

4.2.3 Aurora A thermal stability

Differential Scanning Fluorimetry (DSF), a thermal stability assay, was used to determine the effects on stability of phosphorylated WT Aurora A and non-phosphorylated D274N mutant Aurora A in the presence of different inhibitors. Data are represented as thermal denaturation curves (Figure 4.6A) and the difference in melting temperature relative to a 4% DMSO control (Figure 4.6B). D274N Aurora A was more stable than the WT Aurora A form (in the absence of inhibitors), consistent with similar experiments (collision-induced infolding) in Chapter 3 (Figure 3.16). All five inhibitors increased the thermal stability of WT Aurora A, with MLN8237, VX-680 and MK-8745 all increasing the Tm by >7.5 °C. Further stabilisation was apparent for ENMD-2076 and staurosporine with an increase in Tm of >9.2 °C. The stabilisation-induced effects were less marked with D274N Aurora A due to its inherently higher relative stability than WT protein, with only a small difference in Tm being observed (<2.5 °C).
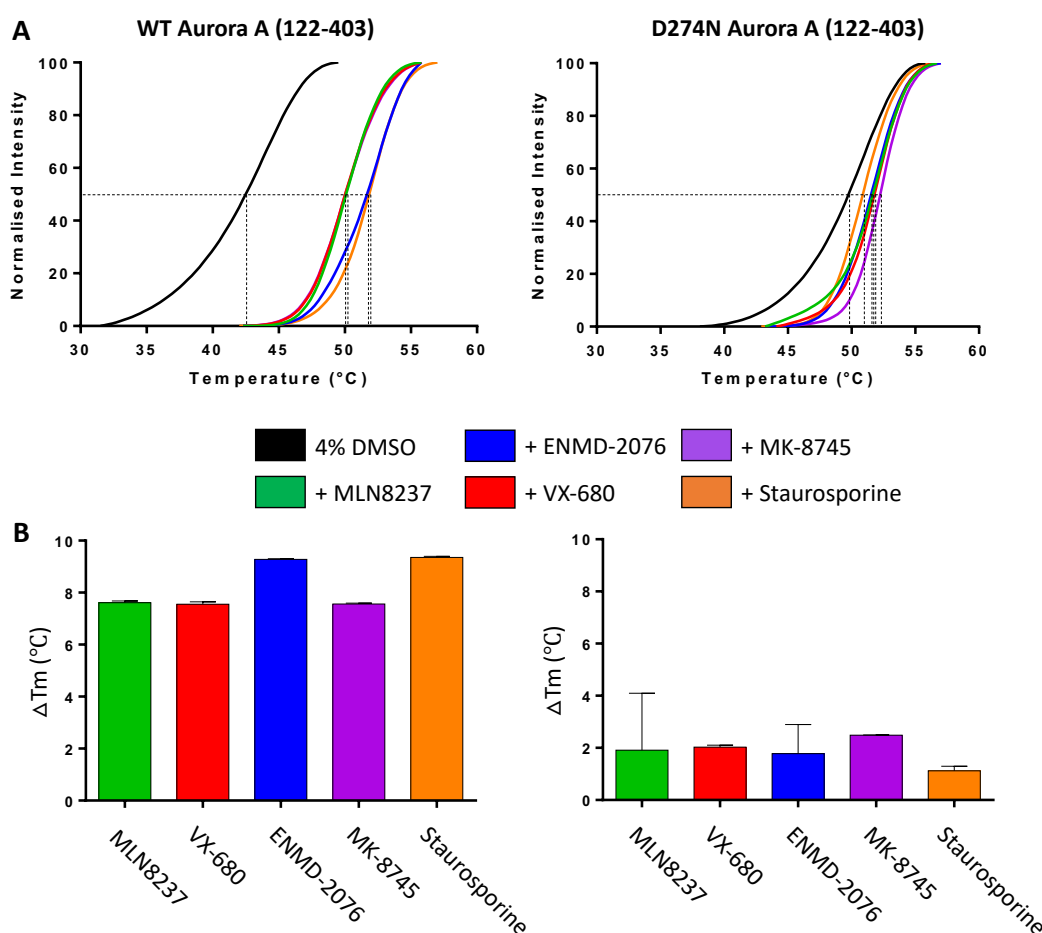
Figure 4.6. Inhibitor bound complexes stabilise WT and D274N Aurora A. A) DSF thermal stability assay with 5 µM Aurora A + 4% DMSO (black), in the presence of 40 µM of each inhibitor. **B)** Difference in melting temperature (ΔTm) relative to 4% DMSO control is presented for both WT and D274N Aurora A.

4.2.4 Aurora A stability determined by CIU

Collision-induced unfolding (CIU), which determines relative kinetic stability, was also performed to evaluate the effect of inhibitor binding, and to compare with the thermal stability changes observed using DSF. Collision energy (CE) was applied in the trap region of the Synapt G2Si instrument in two-volt intervals between 16-34 V, and the CCS values compared under each condition. The application of CE was high enough to induce protein unfolding, but not to induce protein fragmentation, to gain information on the kinetic stability of Aurora A, with each CE able to trap the protein in a specific conformational state (153, 189, 194). Applying collision activation to the inhibitor-bound complexes can enable the observation of many partially unfolded intermediates that are stable within the millisecond time frame, which can gain information relating to the structure of each isolated inhibitor-bound complex. This information can then be related to the stability of the inhibitor complexes and help to understand if any of the inhibitors result in more stabilisation, which could be due to their DFG mode preference (153).

CIU profiles (processed using CIUSuite 2 (195)), revealed that WT Aurora A was less stable (unfolding commencing at ~ 24 nm^2) in comparison to D274N Aurora A (~ 26 nm^2), which was in agreement with the previous findings of the DSF analysis (Figure 4.6). Binding of inhibitors revealed that all the inhibitors had some effect on the kinetic stability of both WT and D274N Aurora A (Figure 4.7). The most prominent difference was the lack (or significant reduction) of intermediate unfolding transition states between the initial conformation and the final unfolded forms. All of the inhibitor complexes resulted in a CCS value of ~ 33-35 nm^2 at the final unfolded form of 34 V (Figure 4.8). However, there were variations evident between the different inhibitors in terms of the energy needed to commence unfolding, and the transitions that were adopted prior to reaching the final unfolded state.
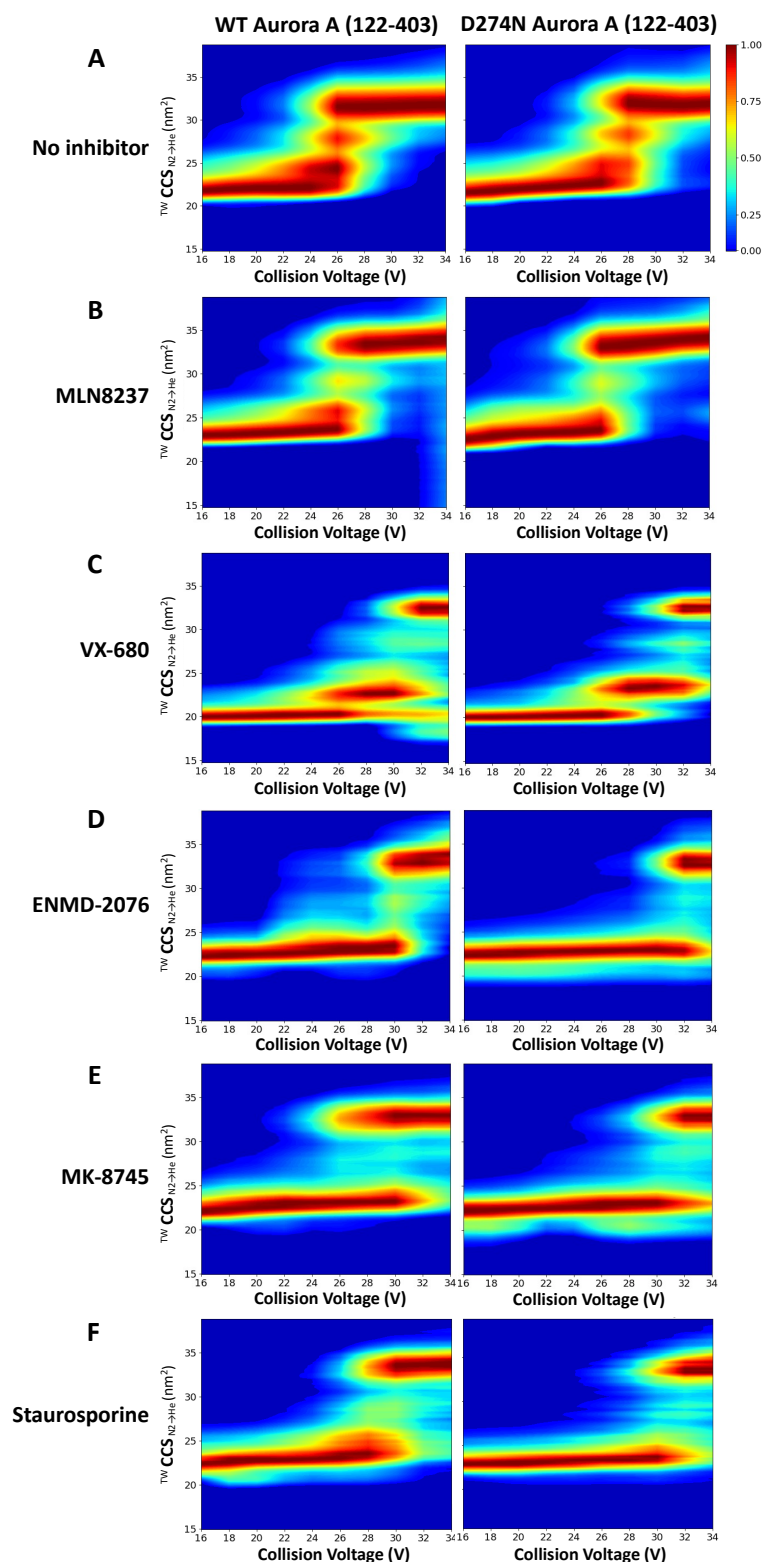
Figure 4.7. Collision-induced unfolding profiles of inhibitor bound Aurora A. The isolated 11+ charge state of **A**) WT (left) and D274N (right) Aurora A (122-403) in the presence of 10-molar excess of **B)** MLN8237, **C)** VX680, **D)** ENMD2076, **E)** MK8745, or **F)** staurosporine were subject to CIU using a stepped collision energy (CE) between 16 and 34 V (two-volt intervals). Data analysis was carried out in MassLynx 4.1, and CIUSuite 2 used to generate the heat maps). Presented are data from a single experiment, representative of the data from independent triplicate analyses.

To further investigate the differences between each inhibitor for the two forms of Aurora A, 'mountain plots' were generated (Figure 4.8). The binding of the MLN8237 inhibitor observed the smallest change to the unfolding profile of Aurora A alone, with similar CE required to initiate unfolding of the two forms. The CCS of the final unfolded conformer for Aurora A in the presence of bound MLN8237 at 34 V was slightly larger than observed for Aurora A alone. This finding would suggest that the intermediate unfolding transitions that Aurora A adopts in the presence of this inhibitor caused a degree of destabilisation. The other partial DFG-out inhibitor, VX-680, showed differences when compared to MLN8237, with the effect of VX-680 resulting in a smaller CCS (<20 nm^2) for all CE, and inducing stabilisation for both forms of Aurora A. ENMD-2076 and MK-8745 revealed the most stabilisation of Aurora A, with the highest stabilisation of all the conditions tested being D274N Aurora A + ENMD-2076 with unfolding beginning at ~ 32 V, compared to ~ 26 V in the presence of no inhibitor. The unfolding profiles of Aurora A in the presence of the complete DFG-in and DFG-out inhibitors, ENMD-2076 and MK-8745, failed to demonstrate any significant differences to be able to distinguish between a DFG-in or DFG-out configuration. Aurora A bound to staurosporine, the generic inhibitor, showed little difference to the effect of the complete DFG-in or DFG-out inhibitors, with this inhibitor stabilising D274N more than WT Aurora A. Although CIU was unable to determine any significant differences between the unfolding profiles of the DFG-in or DFG-out inhibitors, differences were observed between the partial DFG-out inhibitors and DFG-in/out inhibitors. The partially unfolded transition states observed for the DFG-out inhibitors MLN8237 and VX-680 were not present for the other inhibitors, which suggests that these partial DFG-out inhibitors could lock Aurora A into specific configurations.
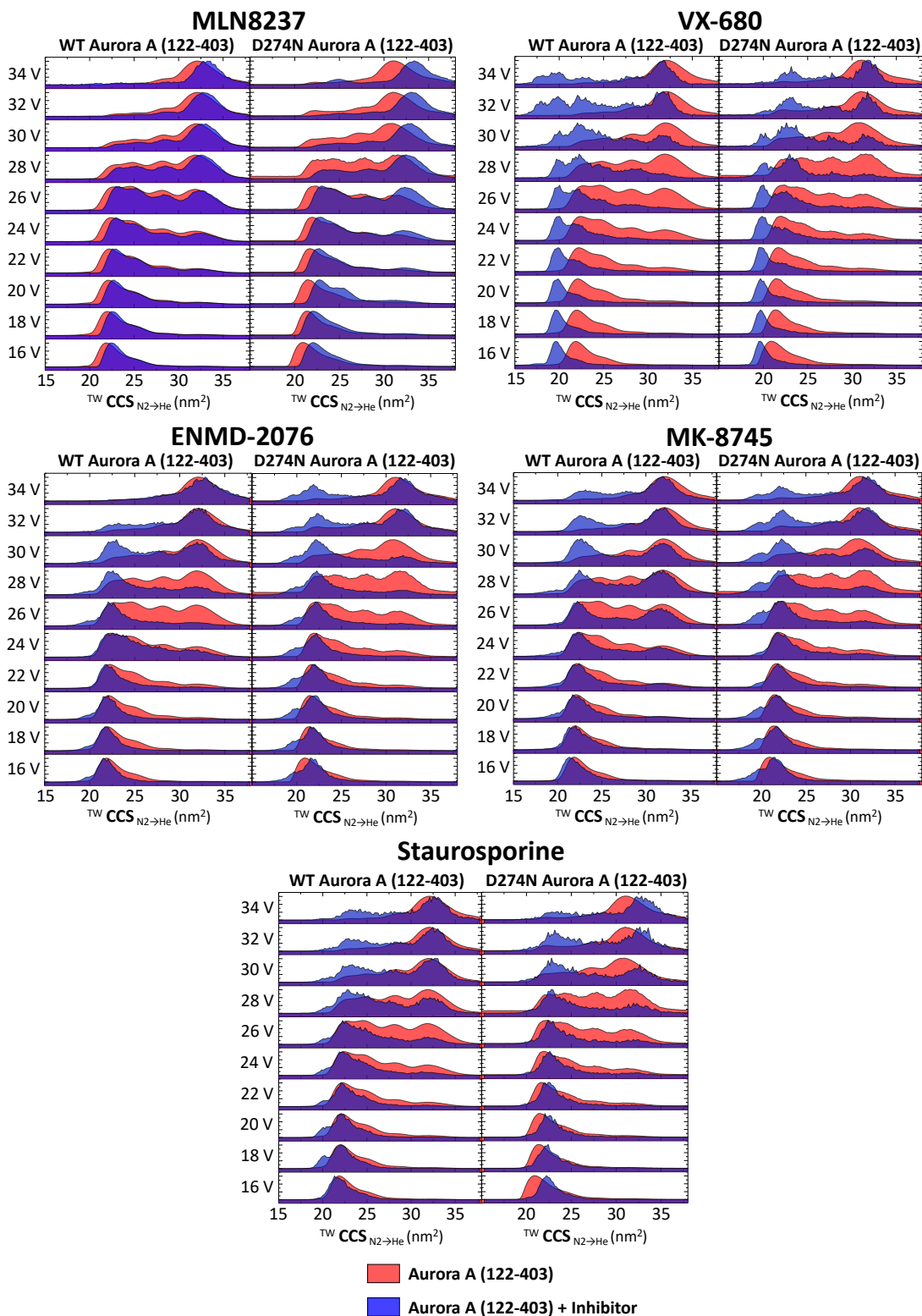
Figure 4.8. Collision-induced unfolding profiles of Aurora A in the absence and presence of inhibitors. The isolated 11+ charge state of WT (left) and D274N (right) Aurora A (122-403) in the presence (blue) or absence (red) of 10-molar excess of inhibitor were subject to CIU using a stepped collision energy (CE) between 16 and 34 V (two-volt intervals). Data analysis was carried out in MassLynx 4.1, generating mountain plots using Origin (Version 2016 64Bit). Presented are data from an average of three replicates.

To directly compare the unfolding profiles for each inhibitor, the data were plotted for a single voltage reading at 26 V, where the highest number of differences were seen across the conditions (Figure 4.9). At 26 V, little difference was observed between both forms of Aurora A alone and with MLN8237 bound, confirming the findings previously discussed and as represented in Figures 4.7 and 4.8. The other four inhibitors for both forms of Aurora A induced stabilisation at 26 V, evident by a more compact and lower CCS value for inhibitor-bound complexes. When directly comparing the inhibitor complexes for WT and D274N Aurora A (Figure 4.8), little difference is observed between the two forms, with exception of MK-8745, where D274N shows far more compaction and stabilisation in comparison with WT Aurora A.
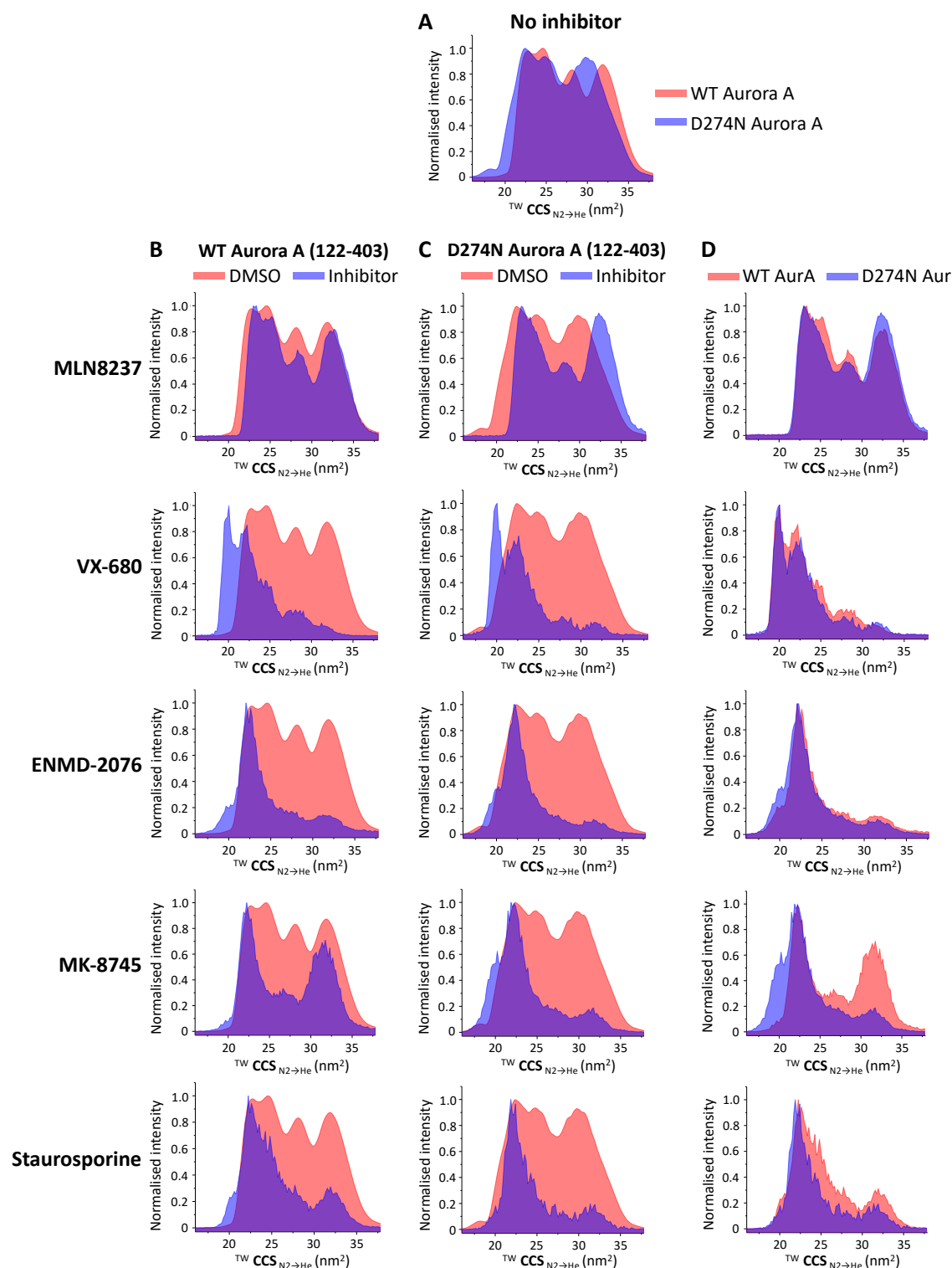
Figure 4.9. Collision-induced unfolding profiles at 26 V of Aurora A in the absence and presence of inhibitors. The isolated 11+ charge state at 26V collision energy of **A)** Aurora A (122-403) with no inhibitor. **B)** WT Aurora A (122-403) and **C)** D274N Aurora A (122-403) in the presence (blue) or absence (red) of 10-molar excess of inhibitor. **D)** Aurora A (122-403) in the presence of inhibitor for WT (red) and D274N (blue). Data analysis was carried out in MassLynx 4.1, and line plots were generated using Origin (Version 2016 64Bit). Presented are data from an average of three replicates.

4.2.5 Effect of TPX2 activating peptide on Aurora A conformation and stability

As previously described, TPX2 plays a role in spindle assembly and has been previously shown to result in activation of Aurora A, whilst inducing a conformational change that enhances autophosphorylation and prevents dephosphorylation. A TPX2 peptide (residues 1-43) has been shown to stabilise the active conformation of Aurora A. This minimally activating peptide was incubated with the two forms of Aurora A at a range of concentrations alongside a mutated TPX2 (Y8A, Y10A, D11A) variant predicted not to bind to assess their effect on WT and D274N Aurora A using an *in vitro* kinase assay (Figure 4.10). Residues [8]YSYDAPS[14] are the conserved region of TPX2, with Y8 and Y10 sitting within the hydrophobic groove between the β sheet, helix αB, and helix αC, when bound to Aurora A. The mutation of these three residues affects the binding ability of TPX2 to Aurora A (35). Activity of WT Aurora A (5 µM) increased upon prior incubation with 0.4 µM (and higher) TPX2. The two higher concentrations of TPX2 peptide (40 and 4 µM) promoted a maximal increase in Aurora A activity with significant reduction in activating ability below 0.4 µM TPX2. In contrast to the activation induced with the TPX2 peptide, the mutated TPX2 peptide (Y8A, Y10A and D11A) failed to increase the activity of WT Aurora A at any of the concentrations tested (Figure 4.10B). The inactivity of D274N Aurora A was unaffected by either the minimally activating, or mutant TPX2 peptides, as expected (Figure 4.10C).
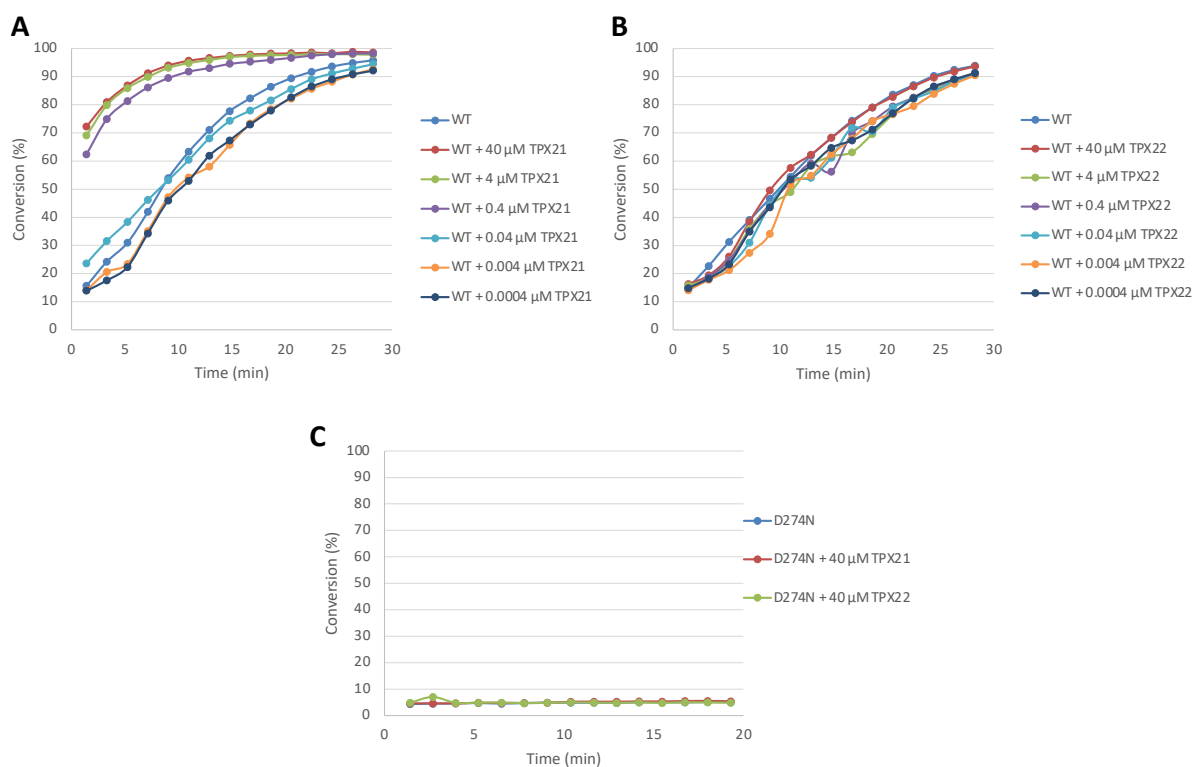
Figure 4.10. WT Aurora A is activated by TPX2 peptide. *In vitro* peptide-based Aurora A kinase assays using 5 µM WT Aurora A in the presence of TPX2 peptide **(A)**, or a non-activating variant TPX2 peptide (Y8A, Y10A, D11A) **(B)**. **C)** 5 µM D274N Aurora A with 40 µM of either TPX2 or the variant TPX2 peptide. The amount of phosphorylation was calculated in real-time using the EZ Reader software by measuring the phosphorylated: non-phosphorylated ratio at each time point in the assay and this was then converted to a % to obtain the peptide conversion (%) plot.

IM-MS analysis was subsequently used to evaluate the effect of peptide binding on the conformational landscape of Aurora A. Following incubation for ten minutes at room temperature with 10x molar excess of the TPX2 and TPX2 variant peptides, the MS profiles of WT and D274N Aurora A (Figure 4.11B/C) were shown to induce a change in the *m/z* distribution compared to Aurora A alone (Figure 4.11A). Additional charge states were present with the binding of TPX2 either side of the 10+, 11+ and 12+ that was observed for Aurora A alone. Binding of TPX2 peptides resulted in an increase in the overall *m/z* region that the peaks adopted, which was expected as the TPX2 peptide has a molecular weight of 4.9 kDa (4.7 kDa for mutated TPX2), and therefore, the overall mass and charge state distribution would increase when bound to Aurora A. The MS profile revealed charge states that were not evenly distributed between each other, which suggests that there are both Aurora A bound to TPX2 (red) and Aurora A unbound to TPX2 (blue) within the complex. For the purpose of

CCS determination, the 11+ charge state of Aurora A + TPX2 or TPX2 variant peptide was quadrupole isolated.
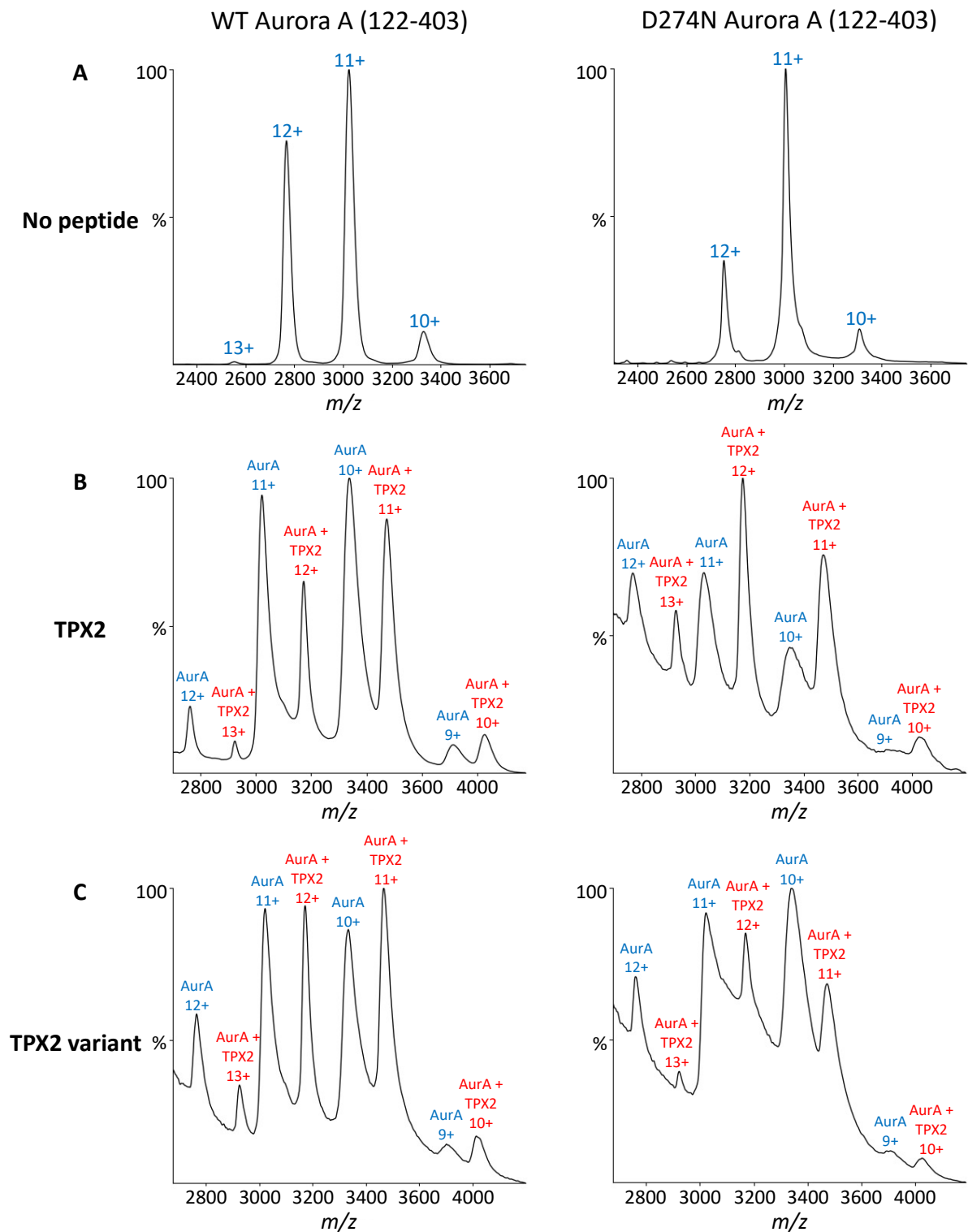
Figure 4.11. TPX2 peptides bind to WT and D274N Aurora A. Mass spectra chromatograms of **A)** WT (left) or D274N (right) Aurora A (122-403) alone, in the presence of 10-molar excess of **B)** TPX2 peptide or **C)** variant TPX2 peptide. Aurora A unbound with TPX2 charge states shown in blue and TPX2-bound shown in red. Aurora A + TPX2 11+ charge isolated for CCS determination.

As previously mentioned, with the increased size of the protein complex when bound to TPX2, the CCS conformer descriptions were changed to I*-IV* (from I-IV) to acknowledge the higher mass of the complex. The addition of TPX2 did not drastically change the conformational landscape of Aurora A compared to the absence of TPX2 peptides (Figure 4.12/4.13A). The CCS values for TPX2-bound conformers I* and II* were similar to WT Aurora A alone (I* = 20.2 nm^2, II* = 22.4 nm^2 compared with I = 20.4 nm^2, II = 22.5 nm^2). However, the CCSD values for TPX2-bound conformers I* and II* increased to 2.0 nm^2 and 2.8 nm^2 respectively compared to 1.5 nm^2 (I) and 1.9 nm^2 (II) for WT Aurora A alone. The increased CCSD values suggest that the binding of TPX2 increases the conformationally flexibility of Aurora A, although could potentially be accounted for by flexibility of the TPX2 peptides themselves. The broad CCSD distribution of conformer II* suggests that this conformer could be a result of various configurations, similar to what was described previously for conformers II and III as 'closed' and 'open' structures. The binding of TPX2 appears to change this conformational equilibrium and, based on the current understanding of activation mechanism, potentially reduces the transition between 'closed' and 'open' configurations. The larger conformer (III*) was notably distinct from the comparable conformer for WT Aurora A alone, where conformer III exhibits a much wider CCSD (3.3 nm^2 for conformer III as opposed to 2.2 nm^2 for III*). In addition, the CCS for conformer III* was larger compared to that of conformer III of Aurora A alone (III* = 26.6 nm^2, III = 24.8 nm^2) and comparable to the IMPACT calculated CCS for the TPX2-bound PDB model (4C3P). The addition of the mutated TPX2 peptide resulted in a similar conformational landscape of Aurora A when bound to the activated TPX2 peptide, with similar CCS and CCSD values observed for all conformers.

The activating TPX2 peptide bound to D274N Aurora A revealed three conformers (II*-IV*), with the absence of conformer I* that was present in the same condition for WT Aurora A and D274N Aurora A alone. The CCS value for D274N Aurora A in the presence of TPX2 peptide for conformer II* (22.3 nm^2) was smaller when compared to D274N Aurora A alone (II = 23.0 nm^2). In contrast, the CCS value of conformer III* (26.4 nm^2) was larger when compared to D274N Aurora A alone (25.5 nm^2), and conformer IV* (30.0 nm^2) had minimal change compared to IV (30.5 nm^2). The addition of TPX2 to WT Aurora A previously showed to result in a larger CCSD value for conformer II*, however, upon addition of TPX2 to D274N Aurora A, this was not the case. The CCSD value for conformer II* had little change to D274N Aurora A

in the presence of TPX2 (2.5 nm^2) in comparison to without TPX2 (2.3 nm^2). This would suggest that the TPX2 peptide fails to induce the same increased conformational flexibility that was observed for WT Aurora A with TPX2 bound. Additionally, no increase in CCSD for conformer III* was observed for D274N Aurora A with TPX2 bound (2.9 nm^2) when compared to without TPX2 (3.0 nm^2). Conformer IV* for D274N Aurora A in the presence of TPX2 showed a large decrease in CCSD value (3.2 nm^2) compared to without TPX2 (4.2 nm^2) suggesting that the peptide significantly reduces the flexibility of this protein. Similar to the effect of the TPX2 variant on WT Aurora A, the variant effect on D274N Aurora A was consistent with the activating TPX2 peptide, where little change was observed between CCS and CCSD values between all conformers.
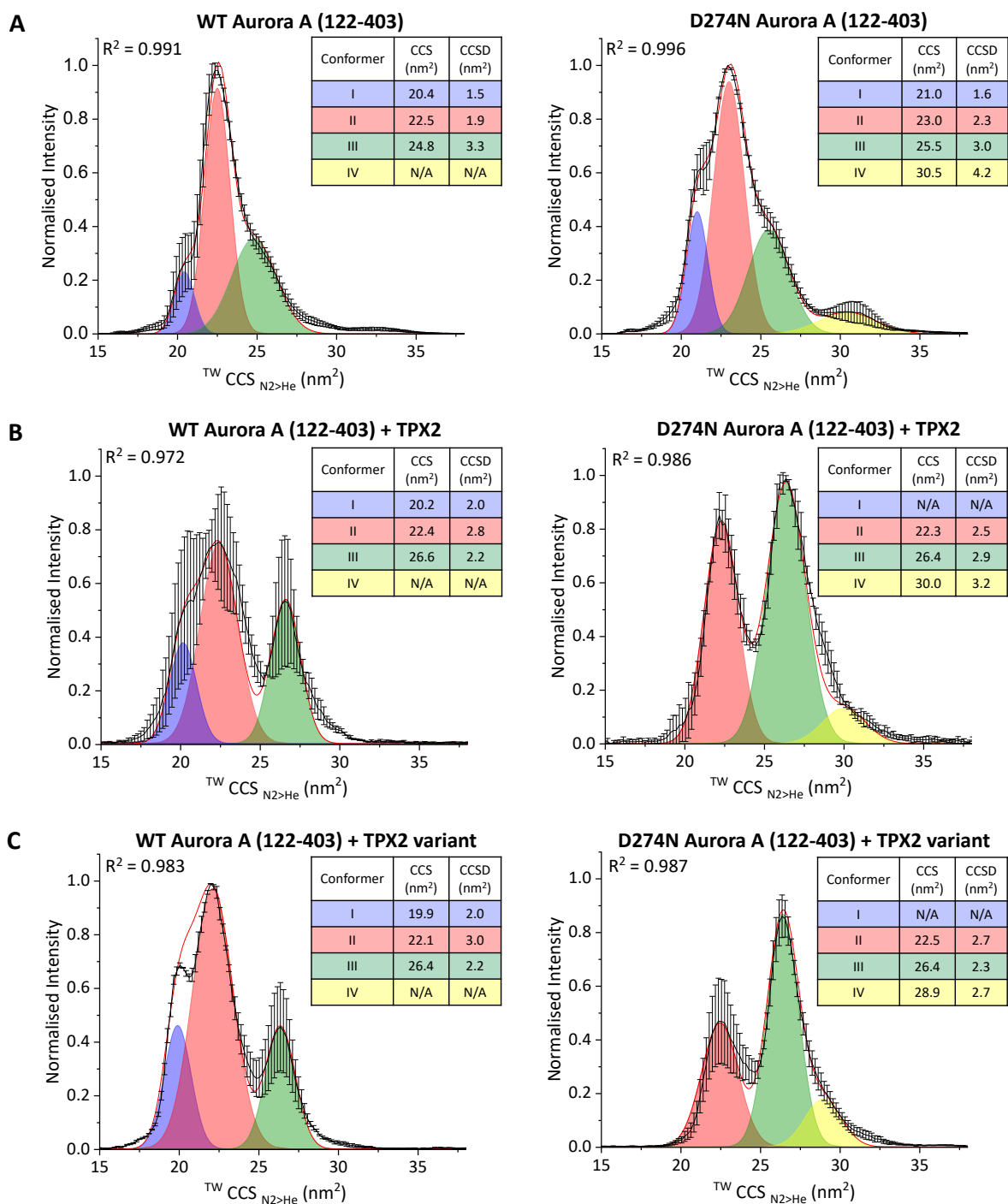
Figure 4.12. TPX2 peptide alters the conformational landscape of Aurora A. A) $^{TW}CCS_{N2 \rightarrow He}$ of the [M+11H] $^{11+}$ species of WT (left) or D274N (right) Aurora A in the presence of 10-molar excess of **B)** TPX2 peptide or **C)** variant TPX2 peptide. Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed using the Fit Peaks Pro function in Origin (Version 2021b), with R^2 values listed.

To further explore the effect of the TPX2 peptides on Aurora A, the relative abundancies of all conformers were determined to see whether subtle changes between conditions could be distinguished. The relative proportion of the most abundant conformer (II*) was comparable to WT Aurora A (conformer II = 54%) when in the presence of the activating TPX2 peptide (49%) or variant TPX2 (60%) (Figure 4.13B). Little change in the relative abundance of the other two conformers (I* and III*) was also observed between WT Aurora A alone and in the presence of both activating and variant TPX2 peptides. Larger differences were observed for conformer II* for D274N Aurora A in presence of activating TPX2 peptide (39%) and variant TPX2 (33%) when compared to without TPX2 (48%). This is in contrast to the same conditions for WT Aurora A, where this conformer remained the most abundant form with the binding of both TPX2 peptides. Upon binding of the TPX2 peptides to D274N Aurora A, the most abundant conformer switched to III*, with 53% abundance for activating TPX2 and 55% for TPX2 variant, compared to 27% in the absence of TPX2, suggesting the more 'open' configuration was adopted.
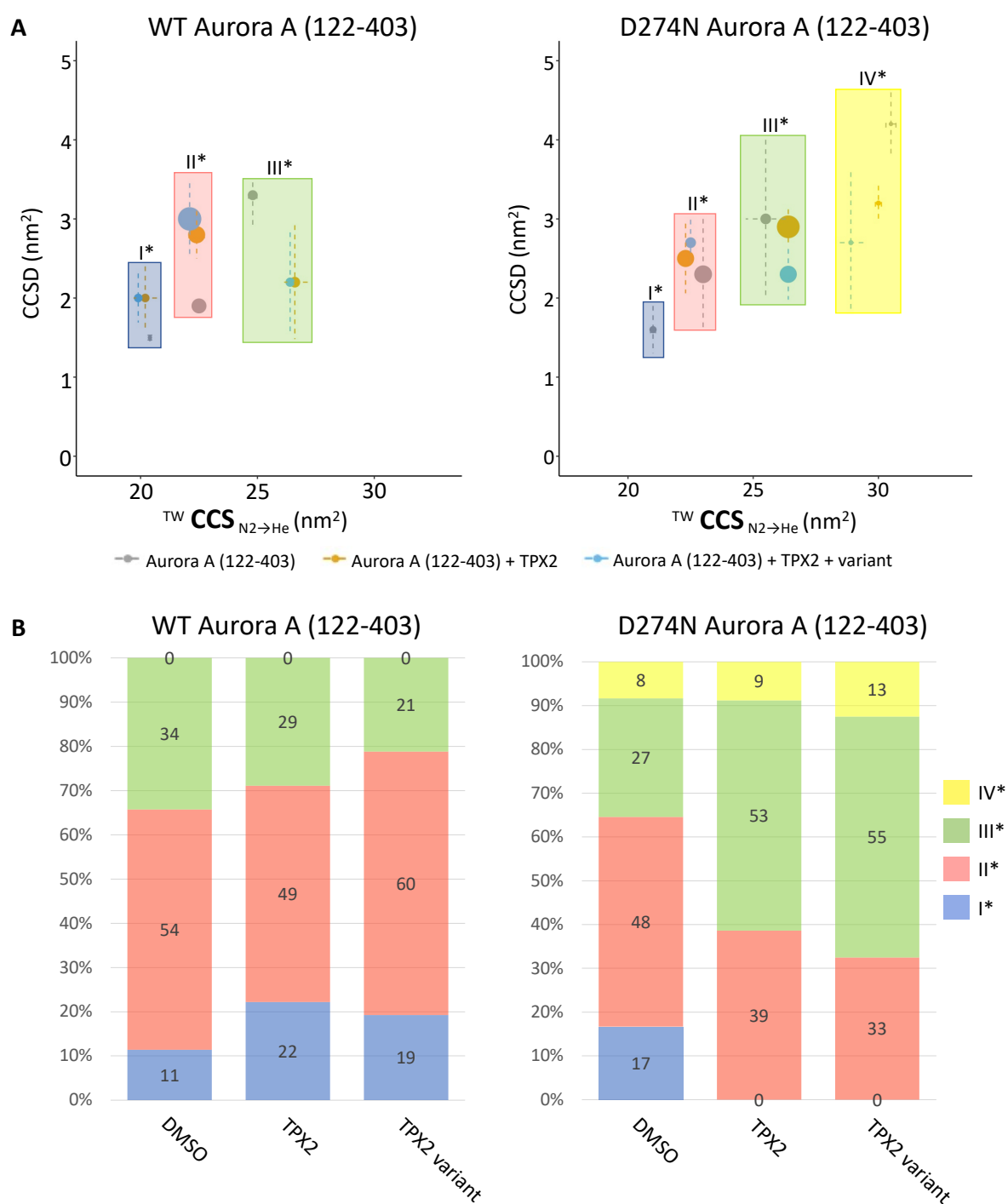
Figure 4.13. Conformational space adopted by Aurora A + TPX2 peptide. A) Proportional scatter plots (CCS (nm²) versus CCSD (nm²)) for the different conformational states (as determined by Gaussian fitting in Figure 4.12) for WT (left) and D274N (right) Aurora A (122-403). Size of dot representative of area. **B)** % area of up to four different conformational states (as determined by Gaussian fitting in Figure 4.12): I* (blue), II* (red), III* (green), IV* (yellow) for WT (left) and D274N (right) Aurora A (122-403) alone or in the presence of TPX2 peptides. Average % area presented from 3 individual experiments.

4.3   Conclusion

This chapter describes the use of IM-MS to assess the conformational dynamics of phosphorylated and non-phosphorylated Aurora A, with the addition of small molecule inhibitors and TPX2 substrate peptides. Gaussian fitting of the IM-MS data resulted in the observation of up to four conformers (I-IV), which varied based on Aurora A activation status and inhibitor binding. The addition of inhibitors revealed a switch in relative abundancies for the two most prominent conformers of Aurora A, II and III, with the conformational landscape of WT Aurora A observing a higher abundance for conformer II (~23 nm^2) when bound with the inhibitors. The CCS findings obtained from the range of inhibitors failed to distinguish between the complete DFG-in (ENMD-2076) and DFG-out (MK-8745) inhibitors. This could be due to a requirement of needing higher IMS resolution to detect the different modes or the intrinsic flexibility that the proteins are adopting in the gas phase. Molecular simulations of Aurora A using different structure configurations obtained from the PDB suggested that conformer II was representative of a 'closed' structure, with the A-loop facing inwards. In contrast, conformer III was more representative of 'open' structures, with the A-loop being in an extended-out orientation. The broad CCSD values obtained from conformer III could also be a combination of 'open' configurations that might not be completely resolved by IM-MS. Inhibitor-bound complexes of Aurora A showed an increase in abundance for conformer II, which suggests that the inhibitors provide a level of constraint to the protein and lock it more into the 'closed' structure configuration.

CIU and DSF were carried out to assess whether the addition of inhibitor compounds resulted in enhanced stabilisation for both forms of Aurora A. The DSF thermal shift assay showed increased stability upon inhibitor binding for all inhibitors bound to WT Aurora A, with a smaller increase in stabilisation observed for D274N Aurora A inhibitor complexes. CIU analysis demonstrated that all inhibitors stabilised both forms of Aurora A, with an increased voltage required to induce unfolding and fewer intermediate unfolding transitions present for inhibitor complexes. The complete DFG-in inhibitor ENMD-2076 and complete DFG-out inhibitor MK-8745 induced the most stabilisation for Aurora A, with MLN8237 showing the least difference to Aurora A alone. The two partial DFG-out inhibitors, MLN8237 and VX-680, had more intermediate unfolding transitions compared to ENMD-2076 and MK-8745. Therefore, this suggests that by carrying out CIU, changes in the relative kinetic stability of

Aurora A bound to a DFG-in/out inhibitor can result in their differentiation from a partial DFG-out inhibitor. The lack of unfolding transition states observed for Aurora A in the presence of ENMD-2076, MK-8745 and staurosporine implies that the partial DFG-out inhibitors (MLN8237 and VX-680) lock Aurora A into specific configurations. This suggest that the effects observed of binding the two partial DFG-out inhibitors are based on both the position of the DFG/P-loop and the noncovalent interactions that these inhibitors associate with.

The addition of the TPX2 peptide was shown to activate WT Aurora A, with an increase in activity by ~4-fold. Minimal changes were observed in the conformational landscape that Aurora A adopts in the presence of TPX2 when compared to Aurora A alone. However, the larger CCSD value for II* in the presence of TPX2 suggests that this conformer could be a result of various configurations that cannot be observed in this type of experiment. Conformer III* of TPX2 bound complexes showed a larger relative abundance and CCS, with a smaller CCSD value, when compared to WT Aurora A alone, which suggests that the activating TPX2 peptide stabilises the III* conformer, termed the 'open' configuration.

Overall, this chapter has utilised a combination of IM-MS and molecular simulations to successfully determine subtle structural differences associated with binding of small molecule inhibitors and TPX2 peptides to phosphorylated and non-phosphorylated Aurora A. Further studies carrying out IM-MS and CIU on the >500 protein kinases members could provide a greater understanding of the conformational landscape adopted by these enzymes, which could be a beneficial insight for the discovery of future small molecule inhibitors for a range of diseases.

Chapter 5. Effect of nucleotide binding and small molecule compounds on the structural conformation of NME1

5.1   Introduction

Nucleoside diphosphate kinase A (NME1) is a known mammalian histidine kinase belonging to the family of nucleoside diphosphate kinases (NDPKs) (95). NME1 has previously been described as a metastasis suppressor, with several pathways shown to associate NDPK activity with suppression of tumour motility and metastasis (196-198). The NDPK family of proteins are located in the nucleus, cytosol and mitochondria, and have shown to be involved in T-cell signalling, cell proliferation, differentiation and migration. The biological functions of NME make it a good candidate as a potential therapeutic target. However, the mechanisms of NME function in tumour metastasis have been challenging to unravel and an inhibitor has yet to be identified. Identifying an inhibitor for NME would help to understand the mechanisms of how NME carries out its functions and therefore, provide a good basis for the discovery and design of future therapeutic targets.

NDPKs play a role in catalysing the transfer of y-phosphate from nucleoside triphosphate to nucleoside diphosphate, which results in the generation of a high energy phosphohistidine (pHis) intermediate via a ping-pong mechanism (85) at His118.

Analysis of pHis is inherently challenging due to the free energy of the phosphoramidate (N-P) bond (87) that can cause loss of the phosphate group at low pH and high temperatures (88). Generic and site-specific antibodies have been identified against the two isomers of pHis, 1-pHis and 3-pHis (82, 199, 200).

This chapter will analyse a range of nucleotide triphosphates to determine whether different nucleotides can act as phosphate donors for NME1 auto-phosphorylation. Previous analysis of H118G NME1, where His118 was mutated to glycine revealed that it was unable to be auto-phosphorylated, identifying H118 as the site of autophosphorylation (201). An NME1 protein with this histidine residue mutated to an alanine (H118A) will be analysed alongside the wild-type form to determine whether the mutation results in the inhibition of phosphorylated NME1.

X-ray structures of human NME1 have revealed that the protein forms a homo-hexameric structure, where the C-terminal domain enables two dimer subunits to contact resulting in their trimerisation to form the hexamer (201). Stabilisation of the hexamer is maintained by a C-terminal 15 residue segment that results in the burial of approximately 30% (300 $Å^2$) of the protein surface. The hexameric form of NME1 is understood to be required for its functional enzymatic activity and the phosphate transfer to histidine (202).

The kpn loop is situated within the active site of NDPKs and plays a key role in the protein surface contacts that define the oligomerisation state of the protein. The kpn loop and the helical hairpin α_A-α_2 form an interaction that enables the stabilisation of any bound nucleotide by the burial of the phosphate and its orientation with respect of the catalytic histidine. The nucleotide phosphate group associates with the side chains of the conserved residues L11, Y51, R87, T93 and R105 of NME1 (203). Consequently, mutation of any of these five residues decreases the enzymatic activity of the protein due to the reduction in hexamer formation (204).

The hexameric form of NME1 that is required for full enzymatic activity has previously been shown to dissociate into a dimer under oxidative conditions (205). Under oxidative conditions, where purified recombinant NME1 was treated with H_2O_2, modifications on redox-active cysteines take place, which recognise a change in the redox environment, leading to dissociation of the hexamer (206). Specifically, Cys109 in NME1 has shown to be reversibly oxidised with glutathione and irreversibly to sulfonic acid when treated with H_2O_2 *in vitro* (207). Residues Cys4 and Cys145 have also been shown to form a disulphide bond under oxidative conditions stabilising the dimer. Oxidised NME1 in its dimer form exhibits lower enzymatic activity, which then leads to a decrease in its ability to act as a suppressor of tumour metastasis (208).

5.1.1 Aims

The aim of this chapter was to explore the autophosphorylation activity and multimerization of NME1 in its wild-type (WT) form and when subject to site-specific mutations that are reported to influence function to gain a more detailed understanding of structure-function relationship.

Activity (as determined by auto-phosphorylation status) was investigated in the presence of different nucleotides using intact MS to quantify the levels of phosphorylation. NME1 WT was initially tested to explore whether a mass shift of 80 Da could be observed in the presence of different nucleotides. A phosphonull variant (H118A) was also analysed as a control. To evaluate the role of Arg 105 on NME1 catalytic activity, I also analysed autophosphorylation and multimerisation of NME1 R105A, in addition to the double mutant (H118A/R105A), which lacks both the site of autophosphorylation and a key catalytic residue.

In addition, to investigate the role of Cys 109 in NME1 complex formation under oxidative stress conditions, I also analysed NME1 C109A under the same conditions. Furthermore, and as previously undertaken with Aurora A (Chapters 3 & 4), I used IM-MS to explore whether the addition of nucleotides influenced protein conformation. In an attempt to identify small molecule inhibitors of NME1 for future use as chemical tools to investigate the cellular roles of NME1, including in metastasis, I also investigated binding, activity and multimer formation of NME1 following screening of an FDA-approved drug library. The library screen was used as a selection process to choose specific compounds to analyse by intact MS and IM-MS to assess any influence the compounds could have on protein conformation, oligomeric state, and activity.

5.2   Results & Discussion

5.2.1 Purification of NME1 proteins

Five forms of purified recombinant NME1 (WT, H118A, R105A, H118A/R105A, C109A as previously described) were provided by Alice Clubbs Coldron, a PhD student with Prof Patrick Eyers (University of Liverpool) allowing me to assess the impact of different nucleotides on NME1 auto-phosphorylation (Figure 5.1). Proteins were expressed in BL21 (DE3) pLysS *E. coli* from pOPINJ plasmids and purified by virtue of an N-terminal 6His tag. The tag was cleaved using 3C protease and NME proteins further purified by size exclusion chromatography.
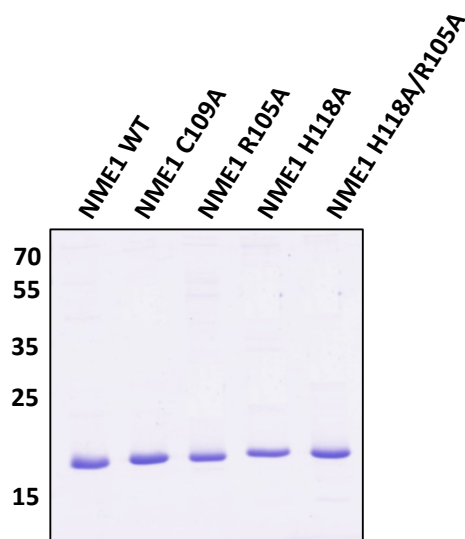
Figure 5.1. SDS-PAGE analysis of NME1 proteins. 1 µg of purified recombinant NME1 protein (WT or variants as indicated) was heated to 95 °C with an equal volume of 2x SDS sample buffer and subjected to electrophoresis on a 12% acrylamide gel, which was stained with Coomassie blue reagent.

5.2.2 Nucleotide effect on NME1 phosphorylation

To determine the nucleotide specificity of NME1 catalytic activity, WT NME1 was incubated with each of the 5 nucleotides (ATP, CTP, GTP, TTP and UTP) and analysed by intact LC/MS to identify a shift in mass associated with auto-phosphorylation. As NME is reported to control the homeostasis of nucleotide pools, the range of five nucleotides were analysed to confirm that they all bind, whether NME had specificity to a specific nucleotide, and whether the level of phosphorylation is consistent between nucleotides. In the absence of nucleotide, the mass of WT NME1 protein was observed at 17,304 Da (Figure 5.2) closely matching the expected mass of 17,303 Da. A smaller peak at 17384 Da (17,304 + 80 Da) was also observed under these conditions, suggesting that a small amount of singly phosphorylated NME1 was also present following expression and purification in *E. coli*. In the presence of ATP, TTP and UTP a single peak was observed at 17,384 Da indicative of stoichiometric phosphorylation of WT NME1 at a single site. While the most abundant mass observed following incubation with either CTP or GTP was also 17,384 Da, there was also a smaller peak at 17,304 Da representing the non-phosphorylated protein, suggesting that CTP and GTP were less efficient substrates than ATP, TTP or UTP.
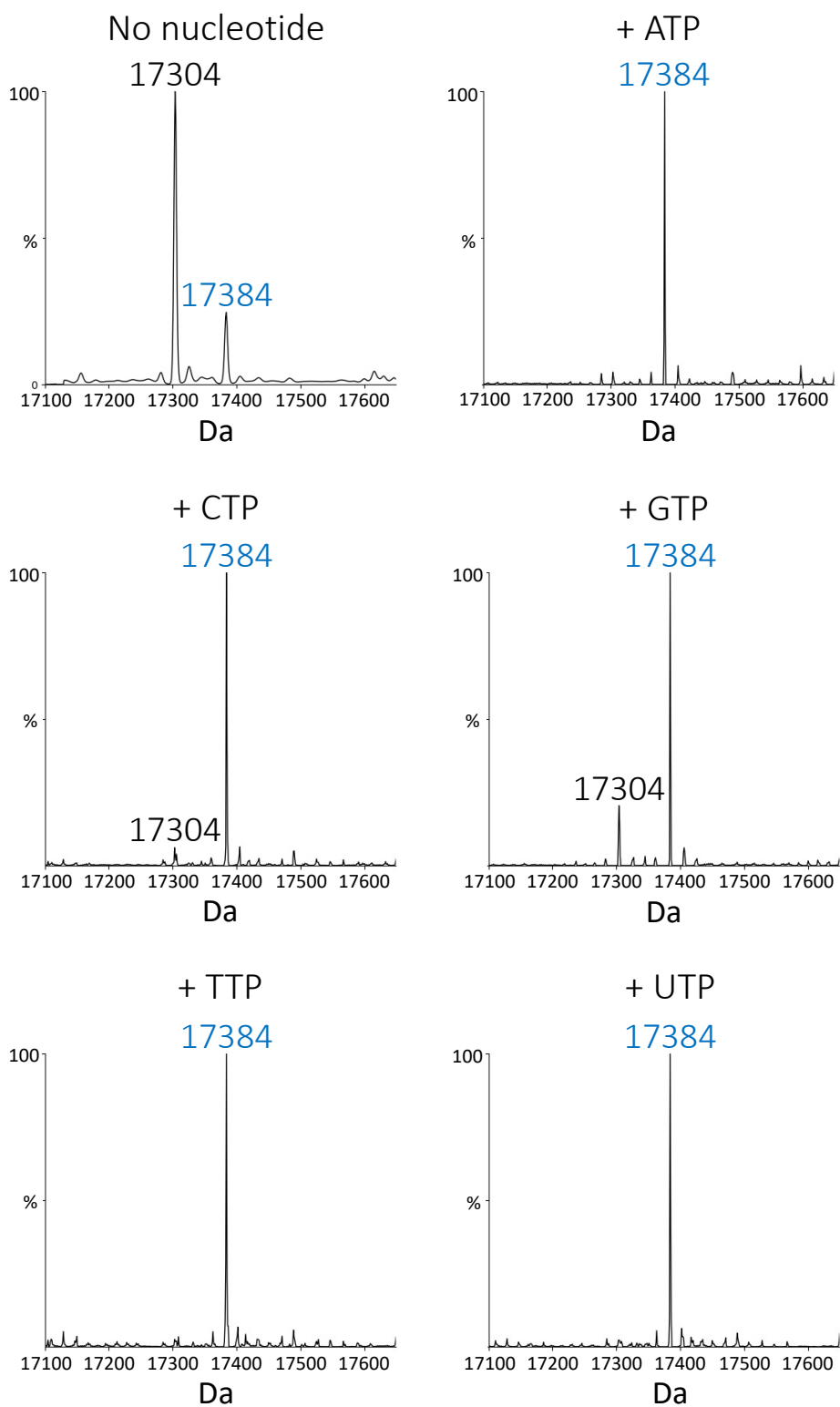
Figure 5.2. NME1 WT uses nucleotides as effective phosphate donors to auto-phosphorylate on a single site. 1 μg of NME1 WT was incubated with 1 mM nucleotide (or without for control) in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 5 minutes at R.T. Sample acquisition was performed on a Synapt G2-Si using a C4 column and 580 nmol injection of material. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1. Deconvolution settings were as follows: 0.5 Da/channel resolution; Uniform Gaussian distribution with a width at half height of 0.500 Da; Minimum intensity ratios for left and right at 33%.

To confirm whether the site of phosphorylation was H118, the H118A variant of NME1 was analysed as previously described for NME1 WT. Irrespective of the nucleotide used, no phosphorylated H118A NME1 was observed, with a mass of 17,238 Da (identical to the theoretical mass of this variant) being recorded under all conditions (Figure 5.3). These data thus confirm the hypothesis that H118 is indeed the site of NME1 auto-phosphorylation under these conditions.
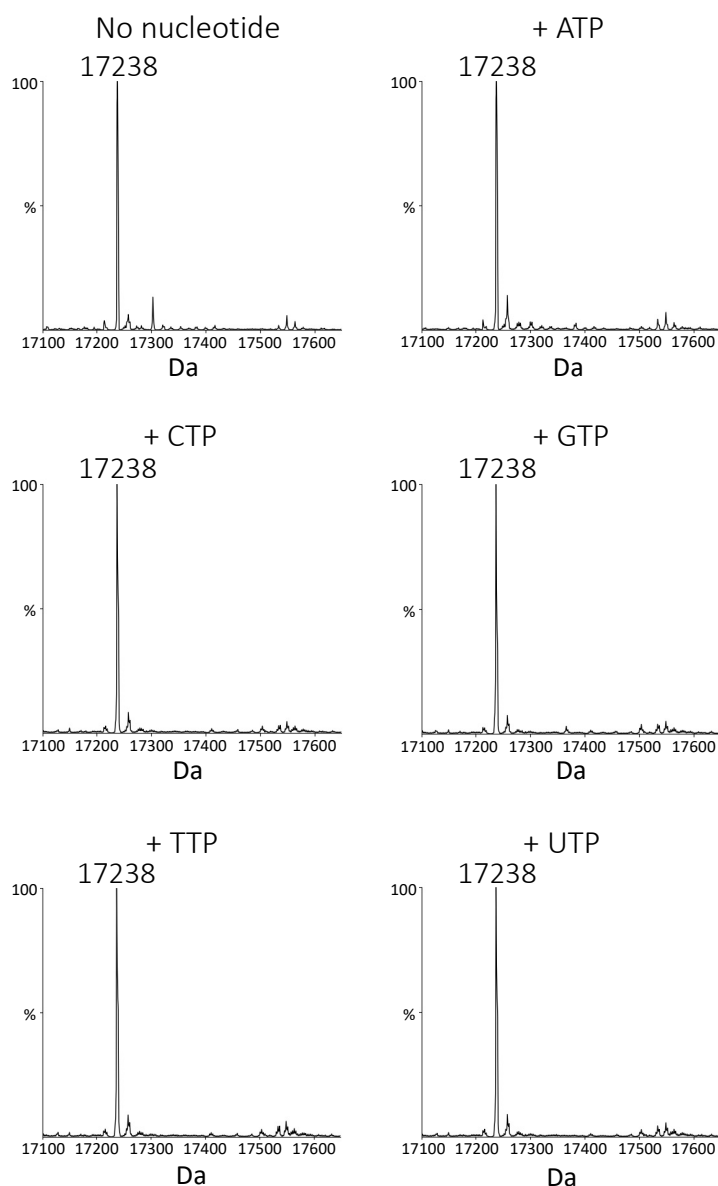
Figure 5.3. Addition of nucleotides to the mutated H118 variant fail to act as an effective phosphate donor for auto-phosphorylation. 1 µg of NME1 H118A was incubated with 1 mM nucleotide (or without for control) in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 5 minutes at R.T. Sample acquisition was performed on a Synapt G2-Si using a C4 column and 580 nmol injection of material. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1. Deconvolution settings were as follows: 0.5 Da/channel resolution; Uniform Gaussian distribution with a width at half height of 0.500 Da; Minimum intensity ratios for left and right at 33%.

NME1 R105A was then analysed under the same conditions to investigate the role of this residue within the active site on its ability to auto-phosphorylate. A mass of 17,219 Da was observed for NME1 R105A in the absence and presence of all nucleotides (compared with a theoretical mass of 17,218 Da) (Figure 5.4), suggesting that mutation of this residue within the active site has an impact on its catalytic function irrespective of the phosphate donor, given the lack of any observed phosphorylated protein.
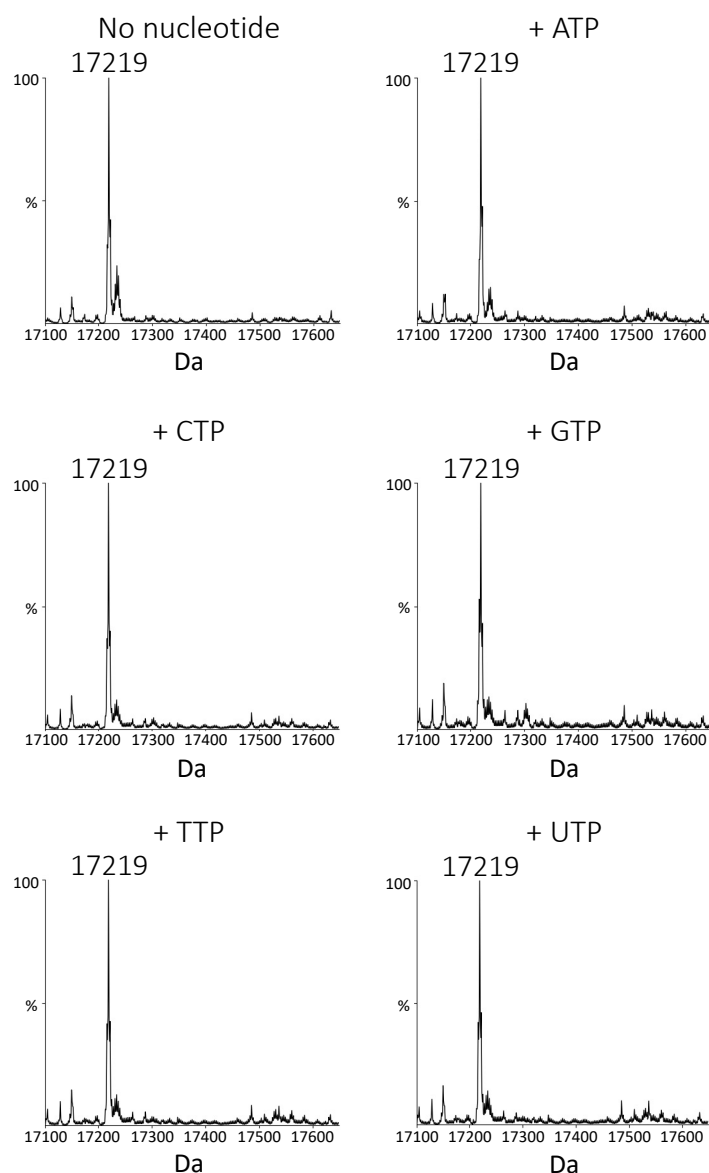
Figure 5.4. Addition of nucleotides to the mutated R105 variant fail to act as an effective phosphate donor for auto-phosphorylation. 1 µg of NME1 R105A was incubated with 1 mM nucleotide (or without for control) in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 5 minutes at R.T. Sample acquisition was performed on a Synapt G2-Si using a C4 column and 580 nmol injection of material. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1. Deconvolution settings were as follows: 0.5 Da/channel resolution; Uniform Gaussian distribution with a width at half height of 0.500 Da; Minimum intensity ratios for left and right at 33%.

Analysis of the H118A/R105A NME1 incorporating both the active site and the acceptor site mutations, supported the previous data. The only major peak observed under any of the conditions was 17,153 Da, compared with the theoretical mass of this protein of 17,152 Da (Figure 5.5).
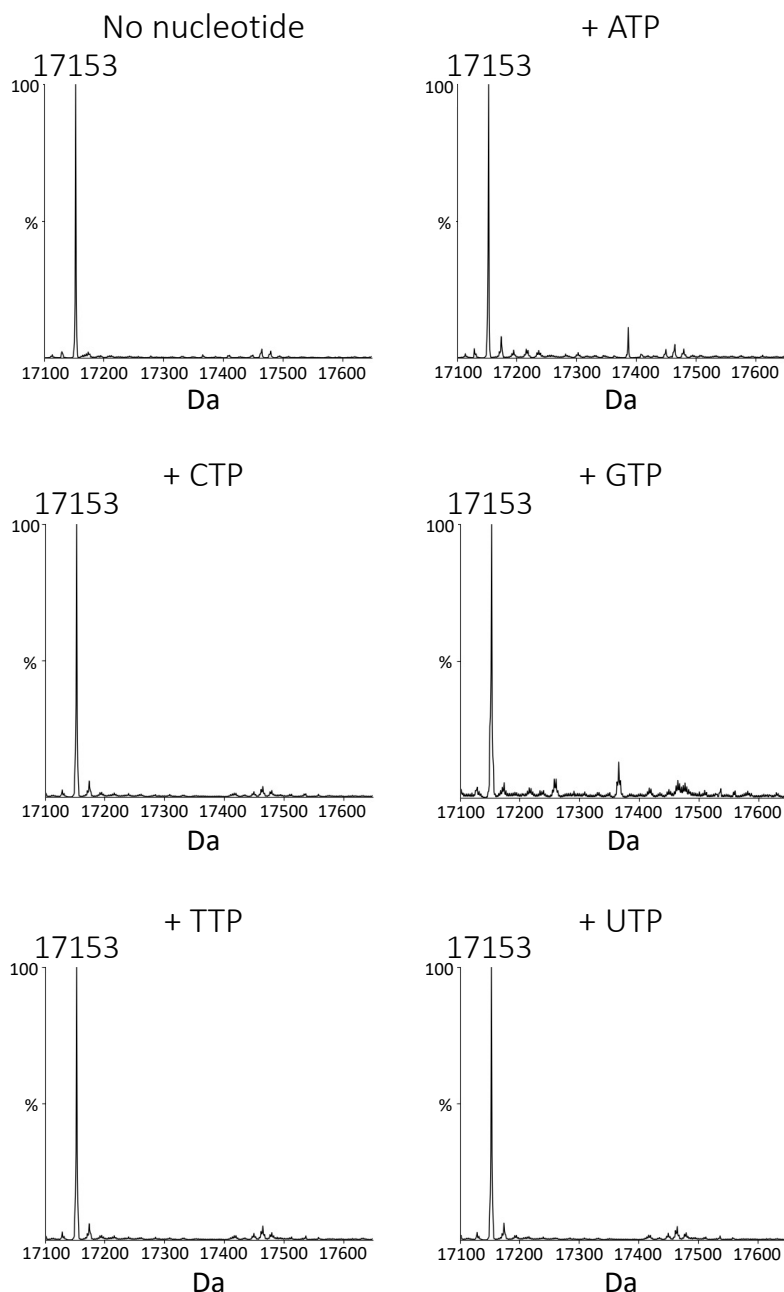
Figure 5.5. Addition of nucleotides to the mutated H108/R105 variant fail to act as an effective phosphate donor for auto-phosphorylation. 1 µg of NME1 H118A/R105A was incubated with 1 mM nucleotide (or without for control) in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 5 minutes at R.T. Sample acquisition was performed on a Synapt G2-Si using a C4 column and 580 nmol injection of material. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1. Deconvolution settings were as follows: 0.5 Da/channel resolution; Uniform Gaussian distribution with a width at half height of 0.500 Da; Minimum intensity ratios for left and right at 33%.

The final NME1 protein variant analysed incorporated a point mutation at C109 (C109A), the site reported to be required for redox-regulated control of NME1 hexamer formation (205). Although this NME1 C109A variant was not expected to have any direct effect on NME1 auto-phosphorylation under these conditions, investigation of this variant is interesting as it could help establish the role of the C109 residue in mediating the effects of specific small molecule (oxidising) inhibitors through regulation of NME1 oligomerisation. Phosphorylation was observed following incubation with all nucleotides, with a major peak apparent at 17,351 Da, compared with a mass of 17,272 Da in the absence of nucleotide (theoretical mass = 17271 Da). However, there was a slight reduction in the levels of autophosphorylation seen compared to NME1 WT (Figure 5.6), with only ATP inducing stoichiometric auto-phosphorylation during the time-course of this assay.
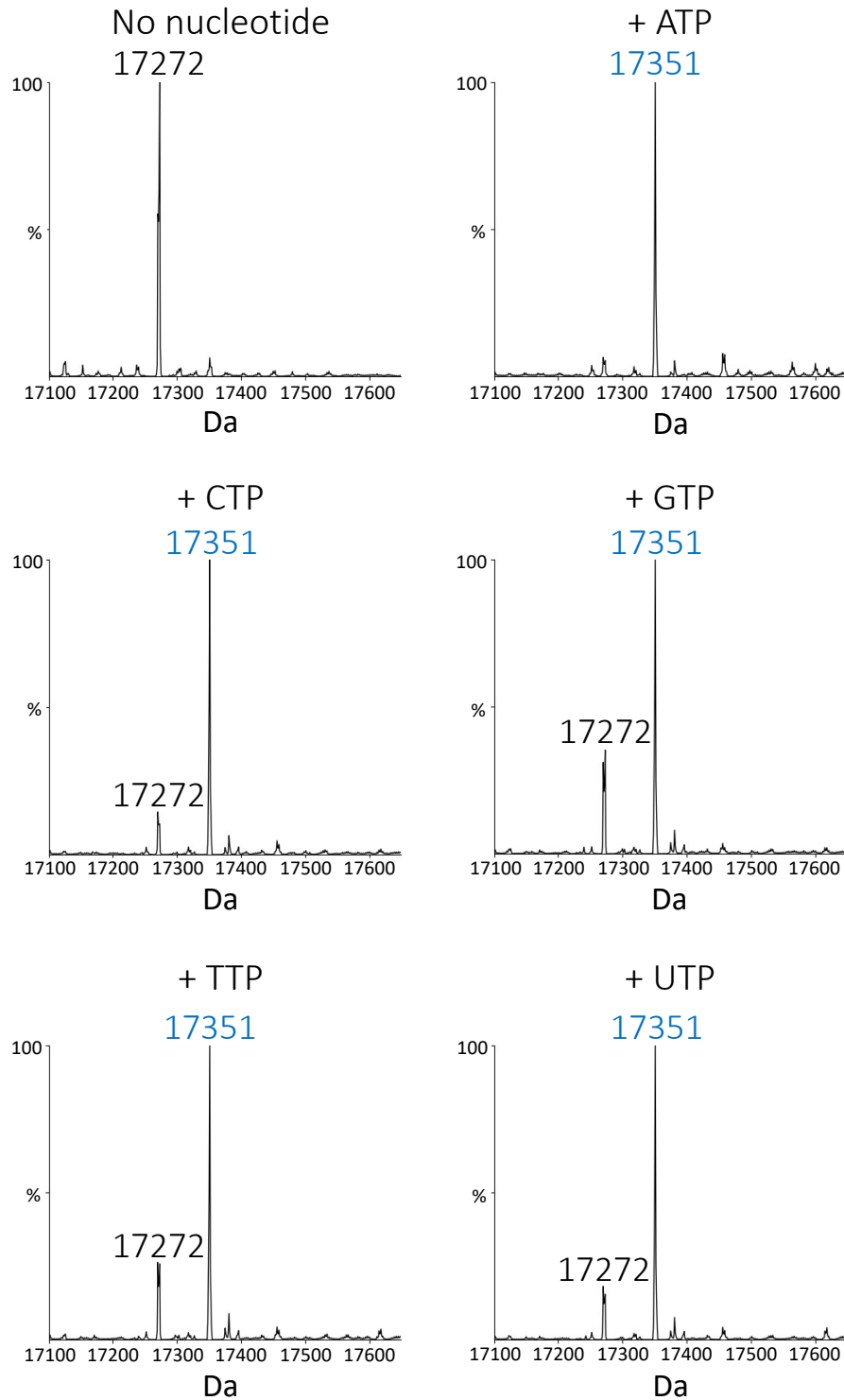
Figure 5.6. NME1 C109A uses nucleotides as effective phosphate donors to auto-phosphorylate on a single site. 1 µg of NME1 C109A was incubated with 1 mM nucleotide (or without for control) in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 5 minutes at R.T. Sample acquisition was performed on a Synapt G2-Si using a C4 column and 580 nmol injection of material. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1. Deconvolution settings were as follows: 0.5 Da/channel resolution; Uniform Gaussian distribution with a width at half height of 0.500 Da; Minimum intensity ratios for left and right at 33%.

A summary of all masses observed for the five NME1 proteins +/- nucleotides is presented in Table 5.1, confirming the findings that only the NME1 WT and C109A proteins had masses associated with the phosphorylated form. As previously stated, mutation of the H118 and R105 residues inhibited NME1 auto-phosphorylation irrespective of nucleotide phosphate donor, confirming H118 as the site of auto-phosphorylation and R105 as essential for catalytic activity, likely as a result of its being required for nucleotide binding.

NME1 protein	Observed Mass (Da)											
	Theoretical mass (Da)	No nucleotide (Da)	+ ATP (Da)		+ CTP (Da)		+ GTP (Da)		+ TTP (Da)		+ UTP (Da)	
WT	17303	17304	17384	95%	17384	93%	17384	78%	17384	96%	17384	97%
H118A	17237	17238	17238		17238		17238		17238		17238	
R105A	17217	17218	17219		17219		17219		17219		17219	
H118A/R105A	17152	17153	17153		17152		17153		17153		17153	
C109A	17271	17272	17351	94%	17351	84%	17351	63%	17351	71%	17351	81%

Table 5.1. NME1 WT and C109A use nucleotides as effective phosphate donors to auto-phosphorylate on a single site. Samples prepared as detailed in Figure 5.2. Stoichiometry of histidine phosphorylation was determined by comparing the relative heights of the phosphorylated and non-phosphorylated protein within the sample mixture containing the phosphate donor (shown in blue as % phosphorylated for each nucleotide for NME1 WT and C109A).

To assess the stoichiometry of NME1 auto-phosphorylation under the different conditions (Figures 5.2-6), the relative abundance (intensity) of the phosphorylated and non-phosphorylated forms of the protein within any given experiment were compared (Table 5.1). The addition of ATP, CTP, TTP and UTP resulted in a high level of phosphorylation at >93% for NME1 WT. GTP showed the lowest level of phosphorylation for all nucleotides for both WT (78%) and C109A (63%). NME1 C109A also showed reduced levels of phosphorylation for all other nucleotides in comparison to NME1 WT. In addition to the reduced levels of phosphorylation with the bound nucleotides for C109A compared to WT, a proportion of WT was phosphorylated in the absence of nucleotide, whereas this was not observed for C109A.

The lower % of phosphorylation observed for all nucleotides bound to C109A compared to WT could be due to the Cys to Ala mutant enforcing a change in conformation that reduces the efficiency of nucleotide binding, which would result in reduced levels of autophosphorylation. The mutation of the Cys to Ala residue could result in a reduction of the hexamer formation that is required for full activity in the WT protein, and therefore, a reduction in the level of autophosphorylation would occur. Small differences associated with the abundance of phosphorylation between the different nucleotides could also be due to the ping-pong mechanism in which the nucleotides bind and differences in the rate of each reaction.

5.2.3 Impact of nucleotides on the conformational dynamics of NME1 proteins

Native IM-MS was carried out to determine the oligomeric states of all of the different NME1 proteins. Following buffer exchange into 150 mM ammonium acetate, NME1 proteins were subject to ESI under 'native' conditions to retain any higher order complexes. Four charge states were present in NME1 WT ranging from 23+ to 20+, with the 21+ charge being predominant (Figure 5.7). The charge distribution data was deconvoluted using UniDec software, which uses Bayesian deconvolution to produce a zero-charge mass distribution and determine the oligomeric form of the protein (209). NME1 WT has a monoisotopic mass of 17.3 kDa and following deconvolution, this protein form presented natively as a hexamer, with a mass of ~ 104 kDa (Figure 5.7). This correlates with previous findings that suggests NME1 exists as a hexamer and the hexameric form results in multifunctional enzymatic activity for this protein (210).
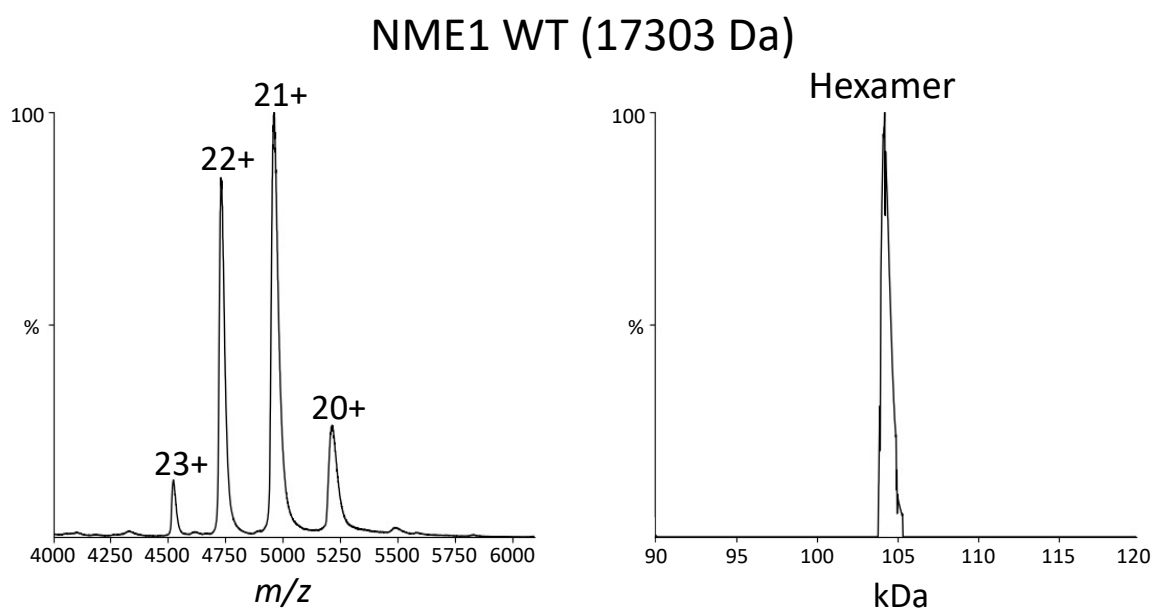
Figure 5.7. NME1 WT natively forms as a hexamer. Native ESI mass spectrum of NME1 WT (left) and deconvolution to determine oligomeric state, revealing a hexamer (right). Spectrum list was copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis.

Following determination of the oligomeric state of NME1 WT alone, the oligomeric state of WT when bound to each nucleotide was also assessed (Figure 5.8). Addition of all nucleotides to NME1 WT resulted in a hexamer, suggesting that NME1 is able to form the active hexameric state irrespective of nucleotide phosphate donor. The slight shift in deconvoluted mass observed for different nucleotides is likely a result of native MS data quality under the different conditions.
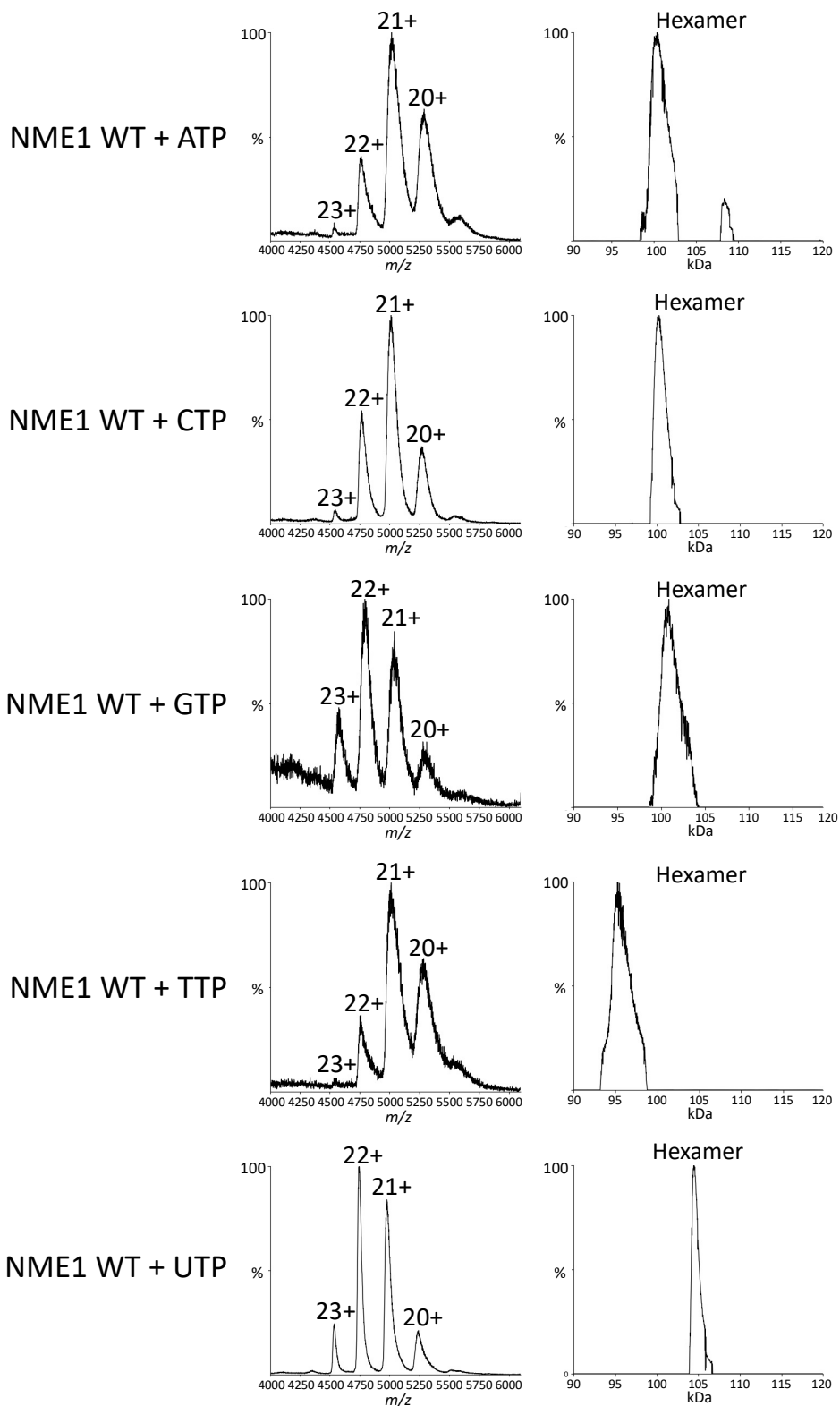
Figure 5.8. Addition of nucleotides does not alter the oligomeric state of NME1 WT. Native ESI mass spectrum of NME1 WT + nucleotide (left) and oligomeric determination showing hexamer form (right) in the presence of each nucleotide. Spectrum lists were copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

Following oligomeric determination of NME1 WT, the most abundant charge state (21+) was subject to quadrupole isolation to determine the rotationally averaged collision cross section (^{TW}CCS $_{N2 \rightarrow He}$) of NME1 WT +/- nucleotides. The drift time of avidin, BSA and β-lactoglobulin were calibrated using the Travelling Wave Ion Mobility (TWIM) device, as described in Chapter 3. Although the range of calibrants used were the same for the Aurora A experiments in Chapters 3 and 4 for the smaller sized Aurora A protein, the selection of these calibrants sufficiently covered the range of mobilities required to account for NME1 forming as the larger hexamer structure. Gaussian fitting was performed in Origin (Version 2021b) using the Fit Peaks Pro function to initially assign the most abundant peak (conformer II). Additional peaks were added for conformers I, III, and IV, where applicable. The CCS, CCSD, and area parameters for each assigned conformer were manually adjusted using the Fit Control and the iteration feature, in order to obtain the best fit between experimental data (black line) and sum of Gaussians (red line).

Gaussian fitting of the CCS distribution for NME1 WT revealed two predominant conformers (II & III), which were observed in all of the nucleotide-bound conditions (Figure 5.9). The CCS values for conformer II of the nucleotide-bound forms were similar to NME1 WT without nucleotide, with conformer II ranging from 51.9 - 52.6 nm^2 for all nucleotide conditions compared to 52.0 nm^2 for WT alone. The CCS values of conformer III were similar for nucleotide conditions (53.8 – 54.4 nm^2), and closely matched NME1 WT alone with 53.9 nm^2. This suggests that the nucleotides have little impact on the protein conformation when compared to NME1 WT alone. Likewise, there was minimal variation in the observed CCSD values for conformers II and III across all these conditions, suggesting that the addition of nucleotide does not increase the flexibility of these conformers. However, nucleotide addition did induce a smaller conformer (conformer I), which was distinct from the non-phosphorylated form of NME1 WT in the absence of any nucleotide. Conformer I represented the smallest CCS value ranging from 50.6 – 51.1 nm^2 for all nucleotides, suggesting that the smaller CCS is evidence of a more compact structure that forms when the protein is in its phosphorylated form. In order to confirm that the complex contained the bound nucleotide, each conformer could have been extracted to determine the mass of the complexes following deconvolution.

Although the CCS and CCSD values were similar between conditions, the relative abundance of the conformers altered with addition of nucleotide. Conformer II, the most abundant in NME1 WT with DMSO control, decreased from 66% to a range of 18 – 51% in the presence of nucleotide, with UTP having the most prominent effect. Conformer III also had a wide range

of relative abundance between nucleotides, with 31% for CTP/GTP and 79% for UTP compared to 34% for NME1 WT + DMSO (Figure 5.10).
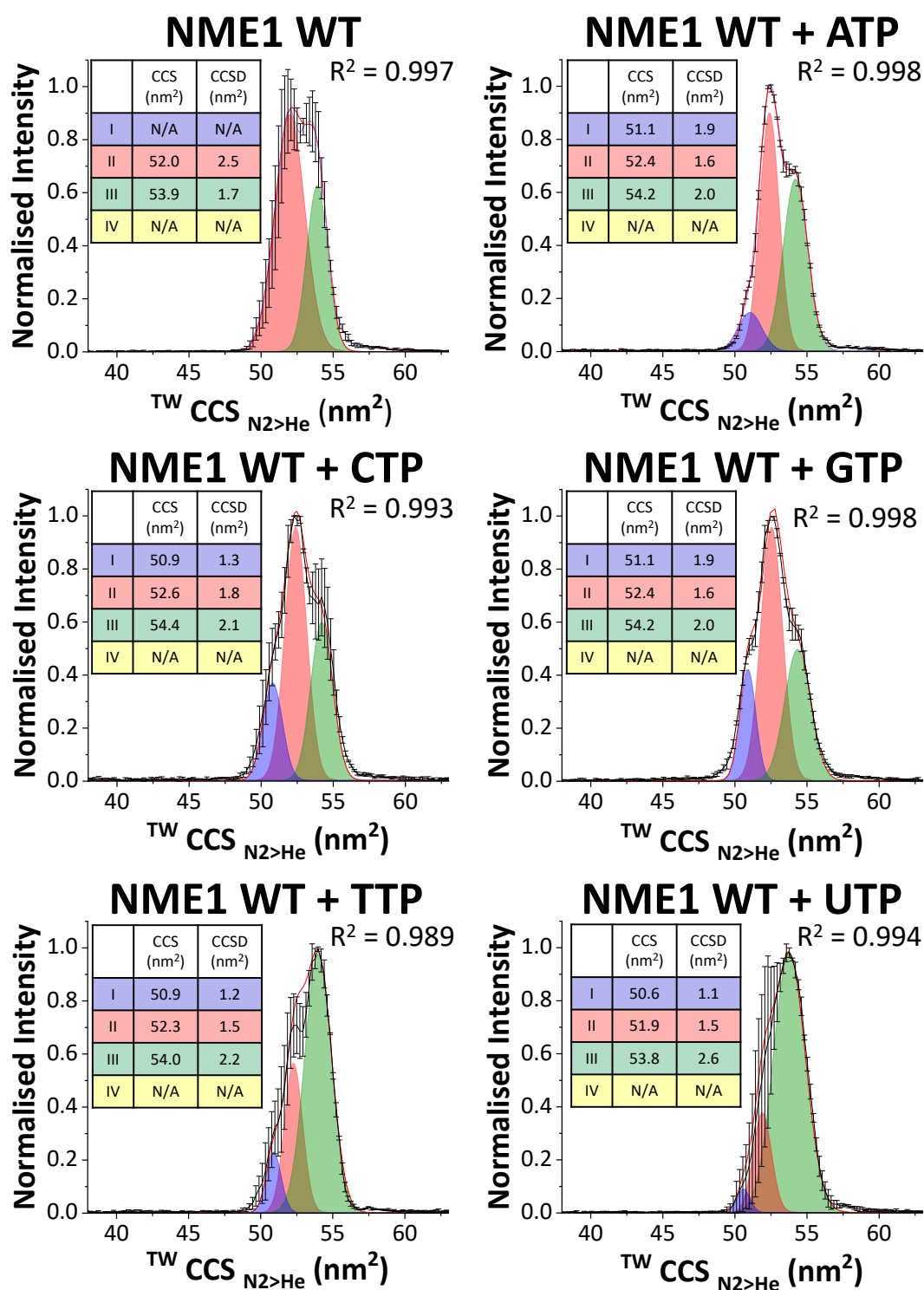
Figure 5.9. CCS distribution of NME1 WT + nucleotides. $^{TW}CCS_{N2 \to He}$ Gaussian fitting [M+21H] $^{21+}$ species of NME1 WT + ATP, CTP, GTP, TTP or UTP. Nucleotide was added in 10-molar excess to protein and incubated for 10-minutes at R.T. Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed using the Fit Peaks Pro function in Origin (Version 2021b), with R^2 values listed.
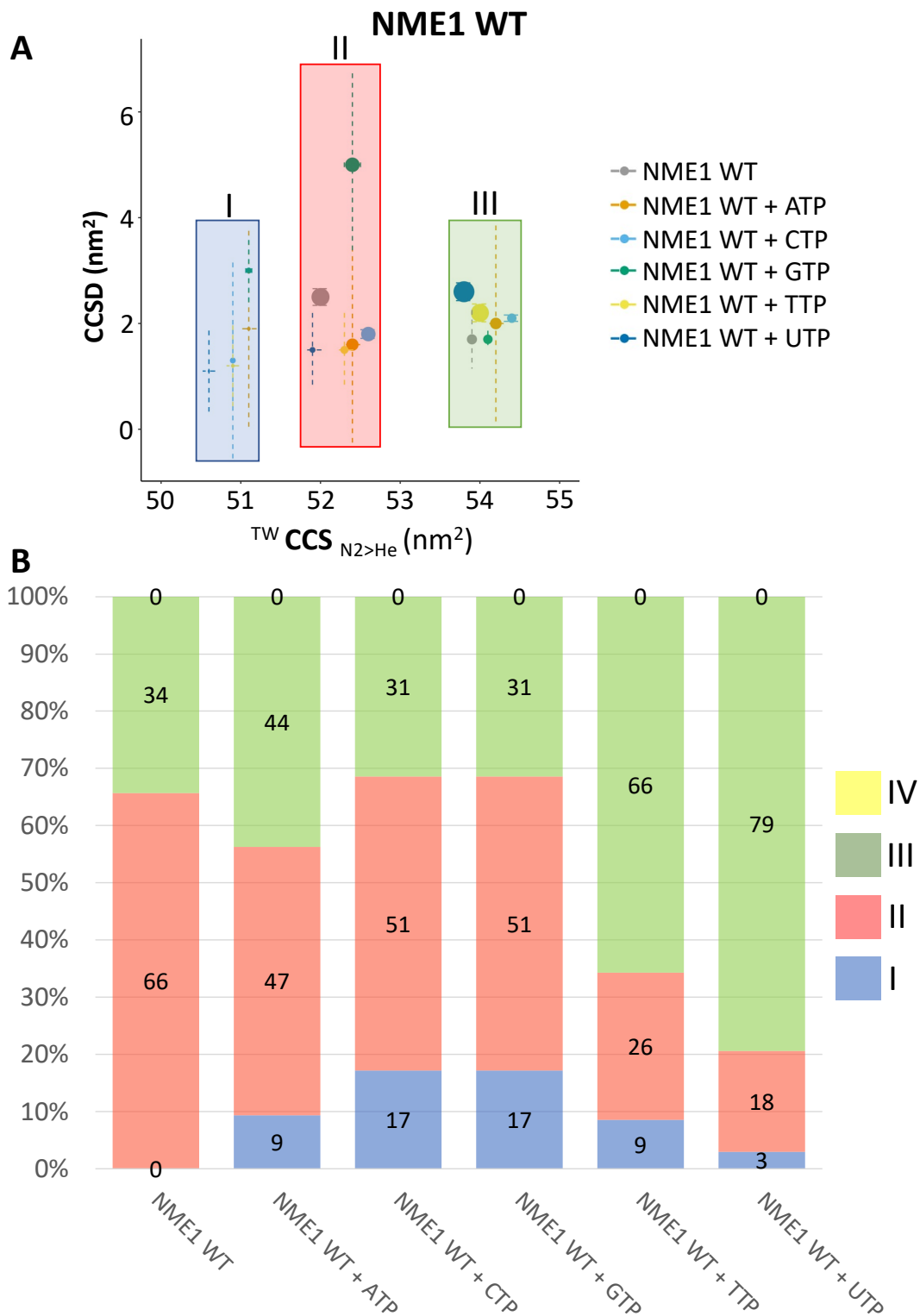
Figure 5.10. Conformational space adopted by NME1 WT alone and in the presence of different nucleotides. A) Proportional scatter plots (CCS (nm^2) versus CCSD (nm^2)) for the different conformational states (as determined by Gaussian fitting in Figure 5.9) for NME1 WT + nucleotides. Size of dot representative of area. **B)** % area of four different conformational states (as determined by Gaussian fitting in Figure 5.9): I (blue), II (red), III (green), IV (yellow) for NME1 WT alone or in the presence of different nucleotides as indicated. Average % area presented from 3 individual experiments.

The NME1 H118A variant had the same charge state distribution as previously seen with NME1 WT, with charges states 23+ to 20+ observed and 21+ the most prominent charge state (Figure 5.11). Following deconvolution of the charge state distribution, a mass was observed at ~ 101 kDa, suggesting that the mutation of the histidine residue had no impact on the hexameric form of NME1. The deconvoluted peak presented itself over a wider range in comparison to NME1 WT, which could be due to a difference in the quality of the charge state distribution data. Addition of all nucleotides to NME1 H118A resulted in a hexamer formation of ~ 100 kDa (Figure 5.12). In similarity to the effect of the nucleotides on the oligomeric state of NME1 WT, no change was observed in oligomeric state when nucleotides were bound to NME1 H118A. These data confirm that hexamer formation is not dependent on NME1 phosphorylation status.

NME1 H118A (17237 Da)
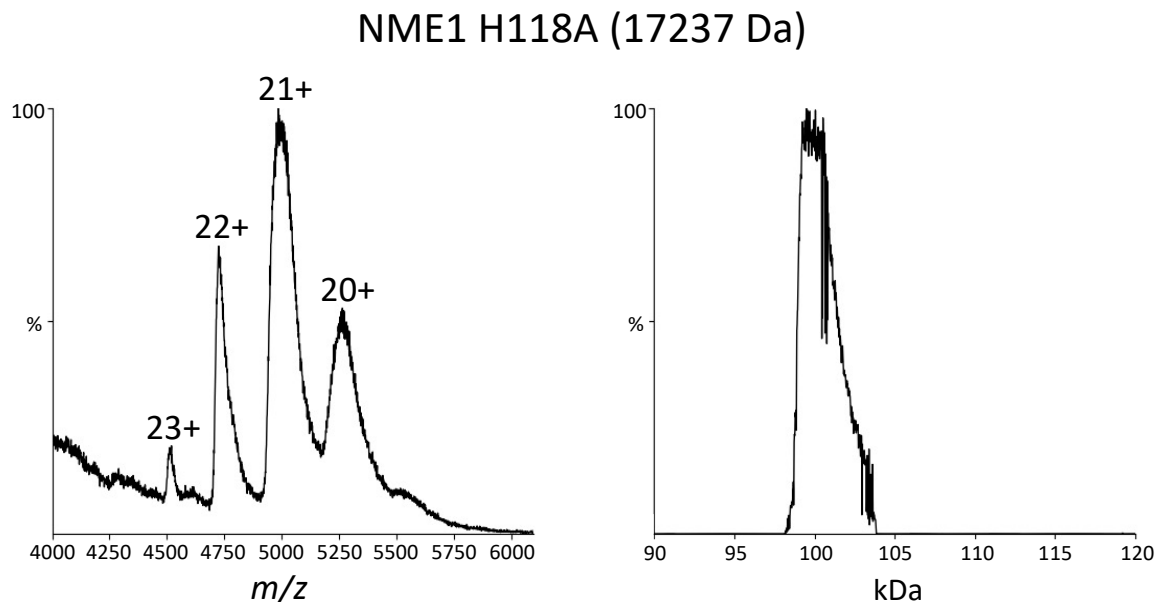
Figure 5.11. NME1 H118A natively forms as a hexamer. Native ESI mass spectrum of NME1 H118A (left) and oligomeric determination showing hexamer form (right). Spectrum list was copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.
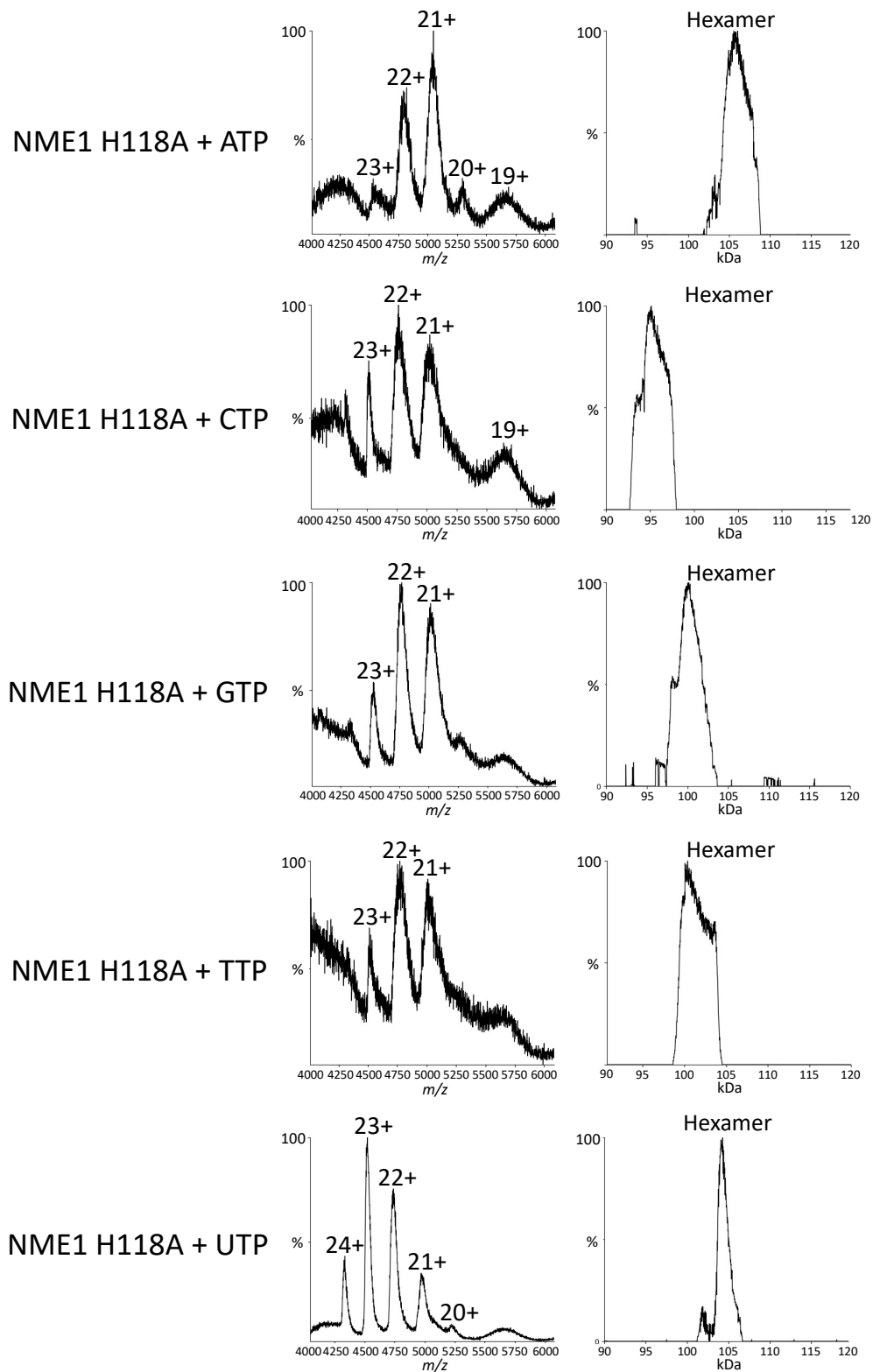
Figure 5.12. Addition of nucleotides does not alter the oligomeric state of NME1 H118A. Native ESI mass spectrum of NME1 H118A + nucleotide (left) and oligomeric determination showing hexamer form (right) for in the presence of each nucleotide. Spectrum lists were copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

143

In contrast to NME1 WT, native IM-MS of the 21+ charge state of NME1 H118A revealed a much narrower CCS profile fitting to a single conformer at 50.7 nm^2 (Figure 5.13), in comparison with the larger conformers for WT of II (52.0 nm^2) and III (53.9 nm^2). The smaller CCS values observed for NME1 H118A suggest that this non-phosphorylatable variant adopts a slighter smaller, more compact structure compared to NME1 WT.

Addition of nucleotides reduced the relative abundance of conformer II to between 58-82% (Figure 5.14), although it remained the most abundant state irrespective of nucleotide. In contrast to the observations of NME1 H118A alone where conformer I was not observed, this state was seen upon inclusion of all the nucleotides. Conformer I had a wide range of smaller CCS values ranging between 47.3 – 49.5 nm^2 for all nucleotide conditions. The CCS values for NME1 H118A + nucleotides are overall smaller than the DMSO control, a similar trend to that previously observed for NME1 WT + nucleotide. This suggests that any change to a smaller conformation observed between NME1 + nucleotides is not due to the effect of the nucleotides acting as a phosphate donor to induce auto-phosphorylation, as H118A lacks the histidine residue required for this to occur, but rather a conformational change induced by binding of these nucleotides.
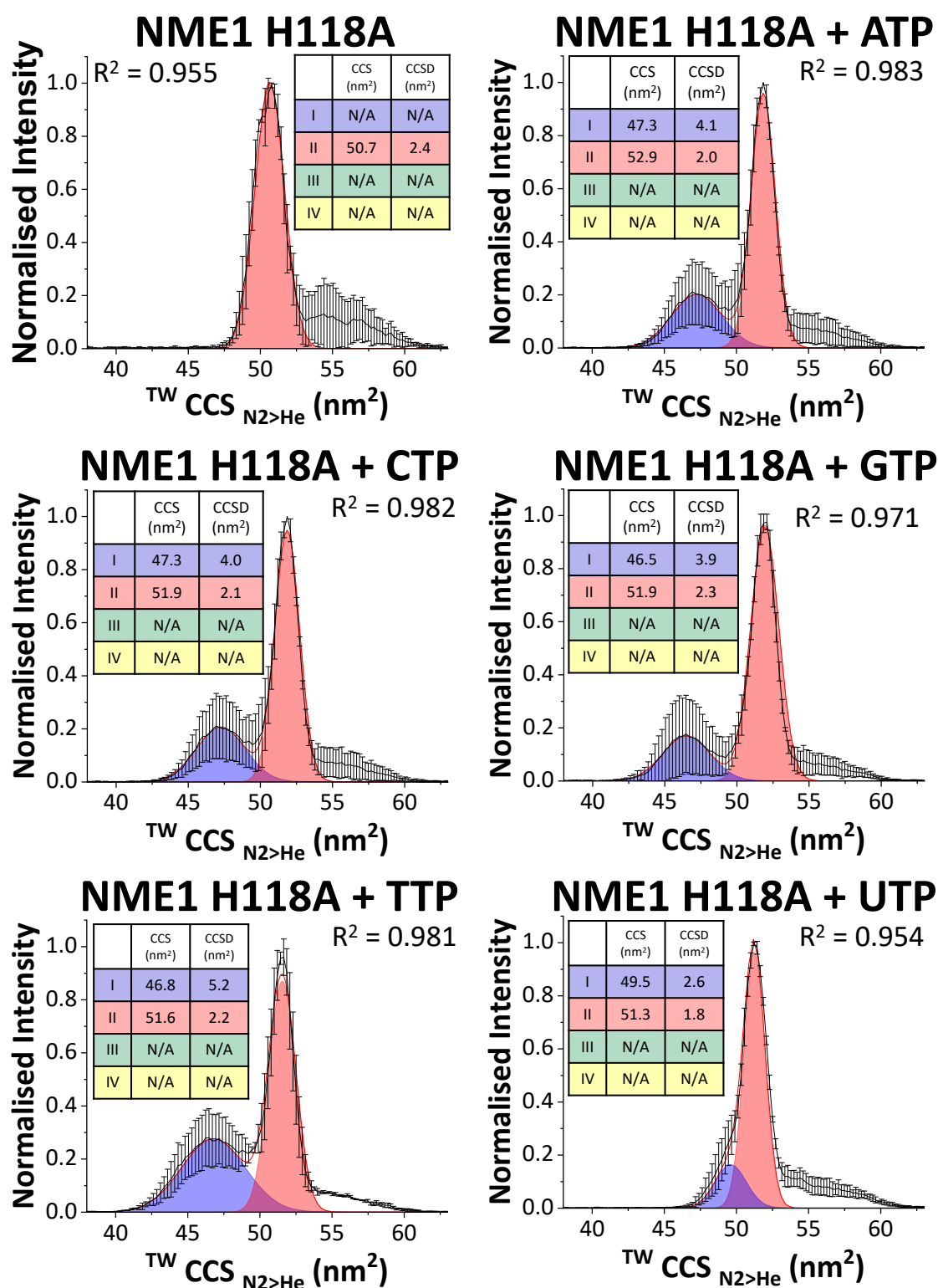
Figure 5.13. CCS distribution of NME1 H118A + nucleotides. $^{TW}CCS_{N2\rightarrow He}$ Gaussian fitting [M+21H] $^{21+}$ species of NME1 WT + ATP, CTP, GTP, TTP or UTP. Nucleotide was added in 10-molar excess to protein and incubated for 10-minutes at R.T. Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed using the Fit Peaks Pro function in Origin (Version 2021b), with R^2 values listed.
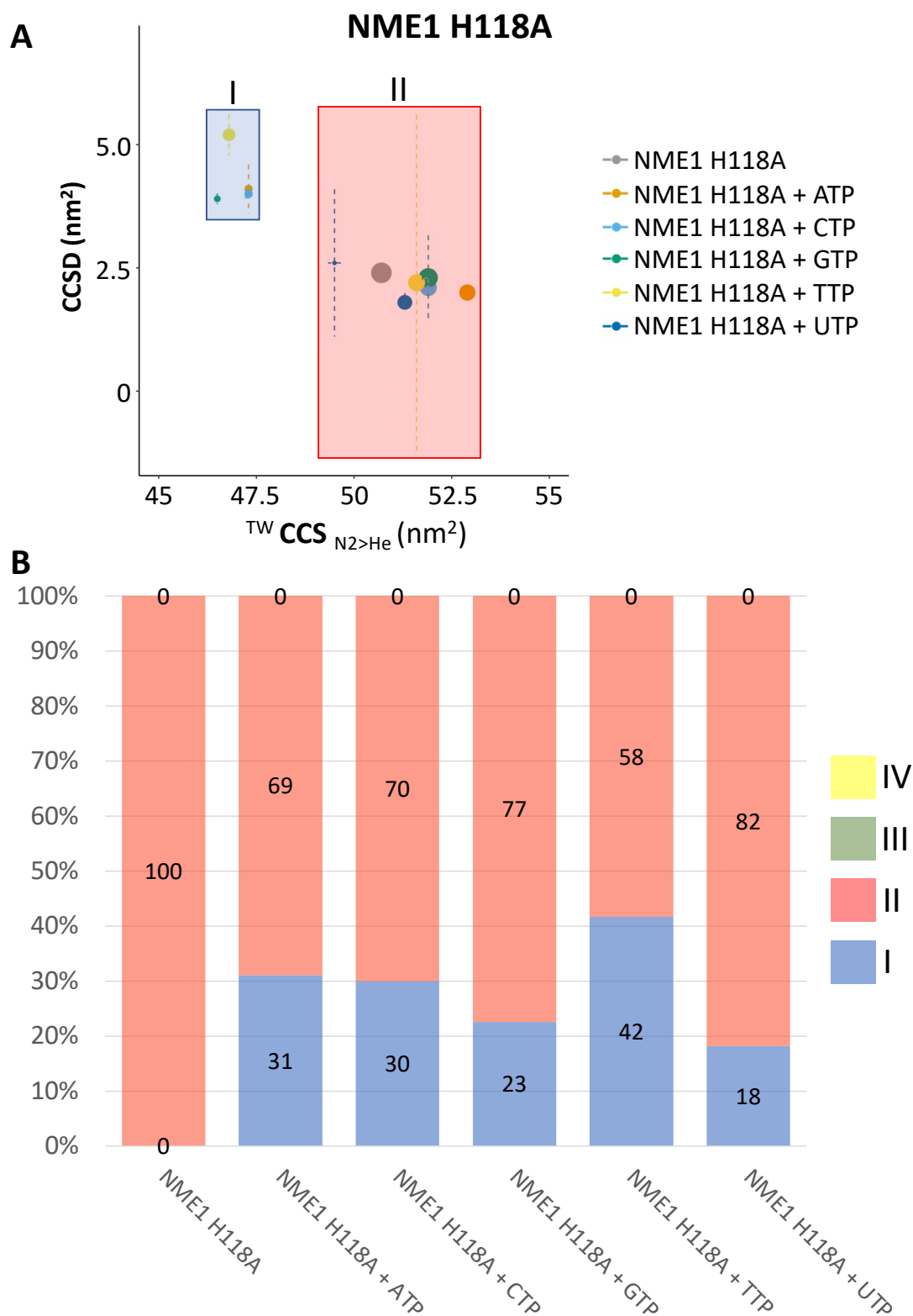
Figure 5.14. Conformational space adopted by NME1 H118A alone and in the presence of different nucleotides. A) Proportional scatter plots (CCS (nm^2) versus CCSD (nm^2)) for the different conformational states (as determined by Gaussian fitting in Figure 5.13) for NME1 H118A + nucleotides. Size of dot representative of area. **B)** % area of four different conformational states (as determined by Gaussian fitting in Figure 5.13): I (blue), II (red), III (green), IV (yellow) for NME1 H118A alone or in the presence of different nucleotides as indicated. Average % area presented from 3 individual experiments.

NME1 R105A resulted in a charge state distribution between 14+ and 11+, with the 12+ charge being predominant (Figure 5.15). The mass spectrum for NME1 R105A was notably different from that for WT protein (Figure 5.7), suggesting that R105A might not form as a hexamer. Deconvolution of this spectrum yielded a mass of ~ 33 kDa, suggesting that this variant forms natively as a dimer. As the R105 residue sits within the kpn loop of NME1, it is thought to play an important role in enzymatic activity and binding of substrates. We demonstrate here that it is also required for formation of the functional NME1 hexamer. Nucleotide addition had no effect on the ability of NME1 R105A to form higher order complexes, with dimer formation occurring in the presence of all nucleotides (Figure 5.16). The is consistent with the two previous NME1 proteins analysed, WT and H118A, where the oligomeric state is not altered when nucleotide is bound.
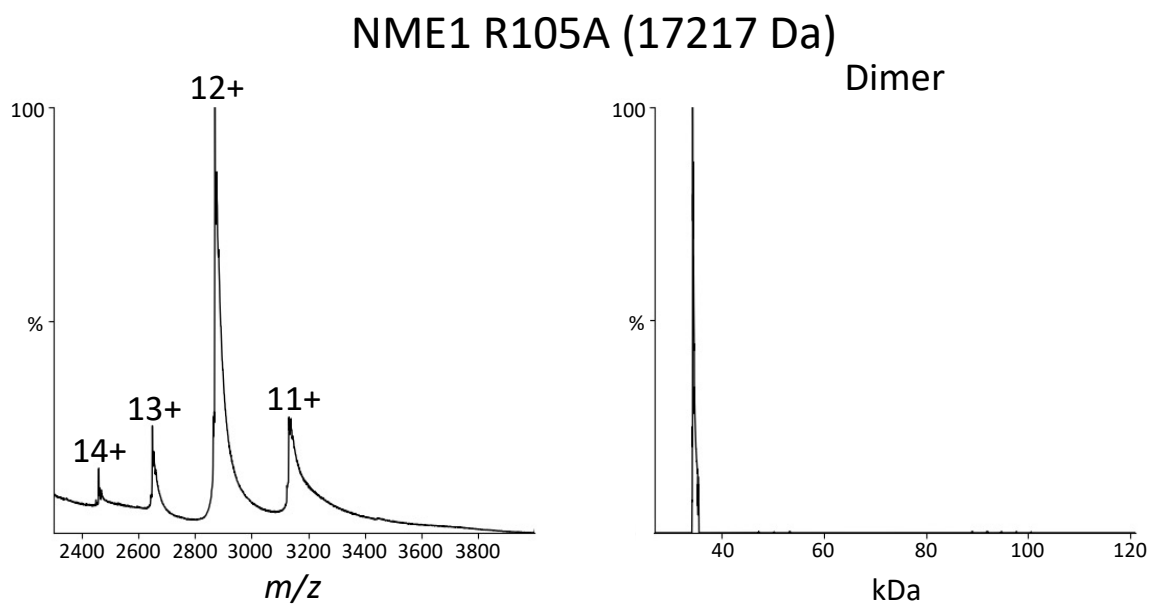
Figure 5.15. NME1 R105A natively forms as a dimer. Native ESI mass spectrum of NME1 R105A (left) and oligomeric determination showing dimer form (right). Spectrum list was copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.
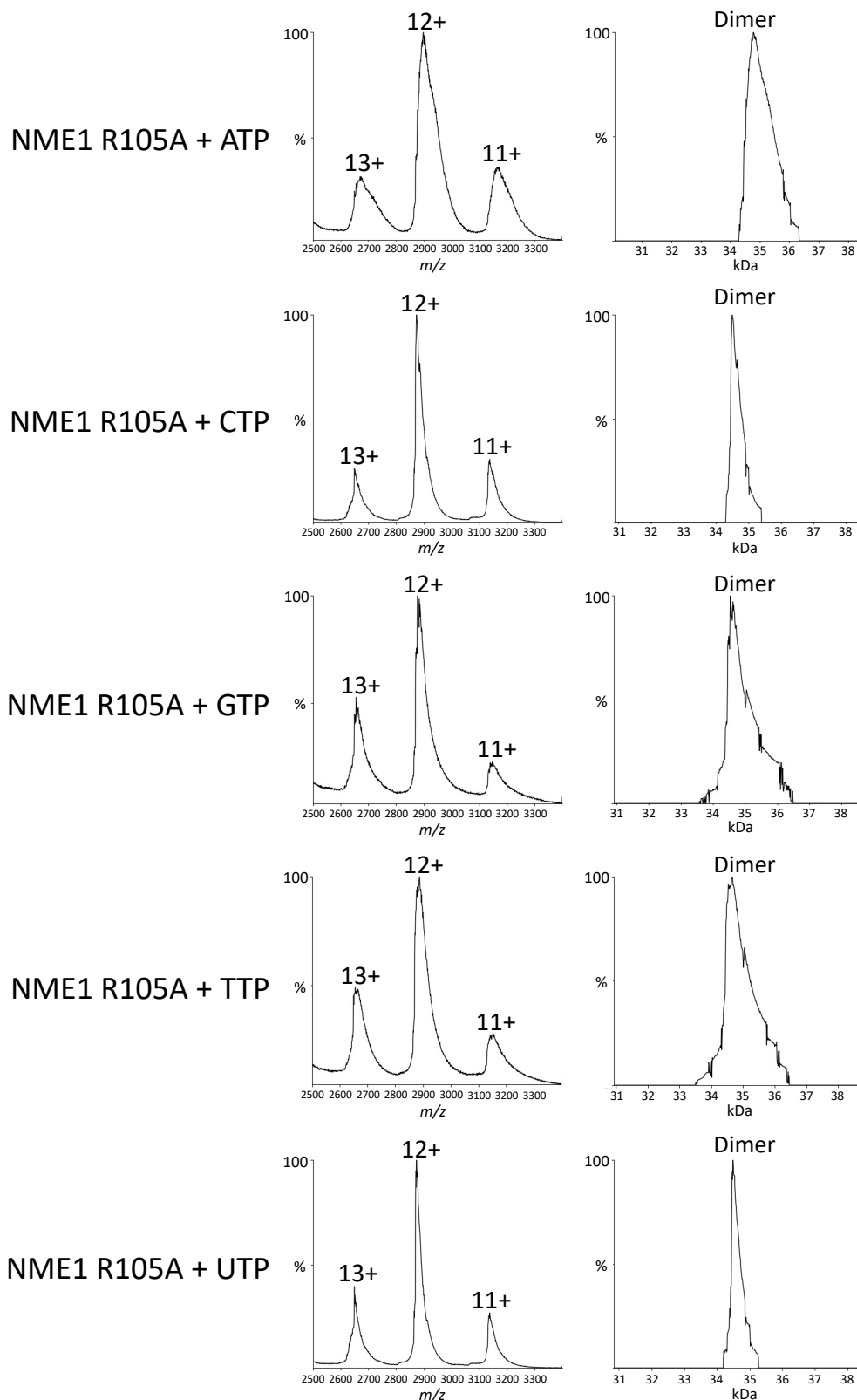
Figure 5.16. Addition of nucleotides does not alter the oligomeric state of NME1 R105A. Native ESI mass spectrum of NME1 R105A + nucleotide (left) and oligomeric determination showing dimer form (right) in the presence of each nucleotide. Spectrum lists were copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

CCS Gaussian fitting of the 12+ charge state for NME1 R105A revealed smaller CCS values when compared to NME1 WT and H118A, which are termed as I*-IV* to reflect the change in oligomeric state. This is expected due to this variant forming as a dimer rather than a hexamer, therefore, the structure is smaller and will result in a smaller CCS value (Figure 5.17). Conformers ranged from 26.3 – 28.6 nm^2 in the absence of nucleotide, whereas the largest conformers for NME1 WT and H118A were 53.9 nm^2 and 50.7 nm^2 respectively.

The CCS values for NME1 R105A in the presence of any nucleotide was not significantly changed compared to the DMSO control condition. Addition of nucleotides resulted in the same conformational landscape, with the presence of conformers I* and II*, ranging from 26.7 – 29.1 nm^2 compared to 26.3 – 28.6 nm^2 for the DMSO control. A slight decrease in CCSD values for conformer II* was observed upon addition of nucleotide, ranging from 2.1 – 2.6 nm^2 compared to the DMSO control of 3.3 nm^2. The small decrease in CCSD values in the presence of nucleotide suggests that these phosphate donors serve to constrain the conformational landscape of NME1 R105A. Upon addition of nucleotides, the overall abundance for each conformer had little change, with the majority of conditions observing conformer II as the most abundant (Figure 5.18). The little effect that the nucleotides had on NME1 R105A could be due to this variant being catalytically inactive as a result of the lack of ability to bind a nucleotide.
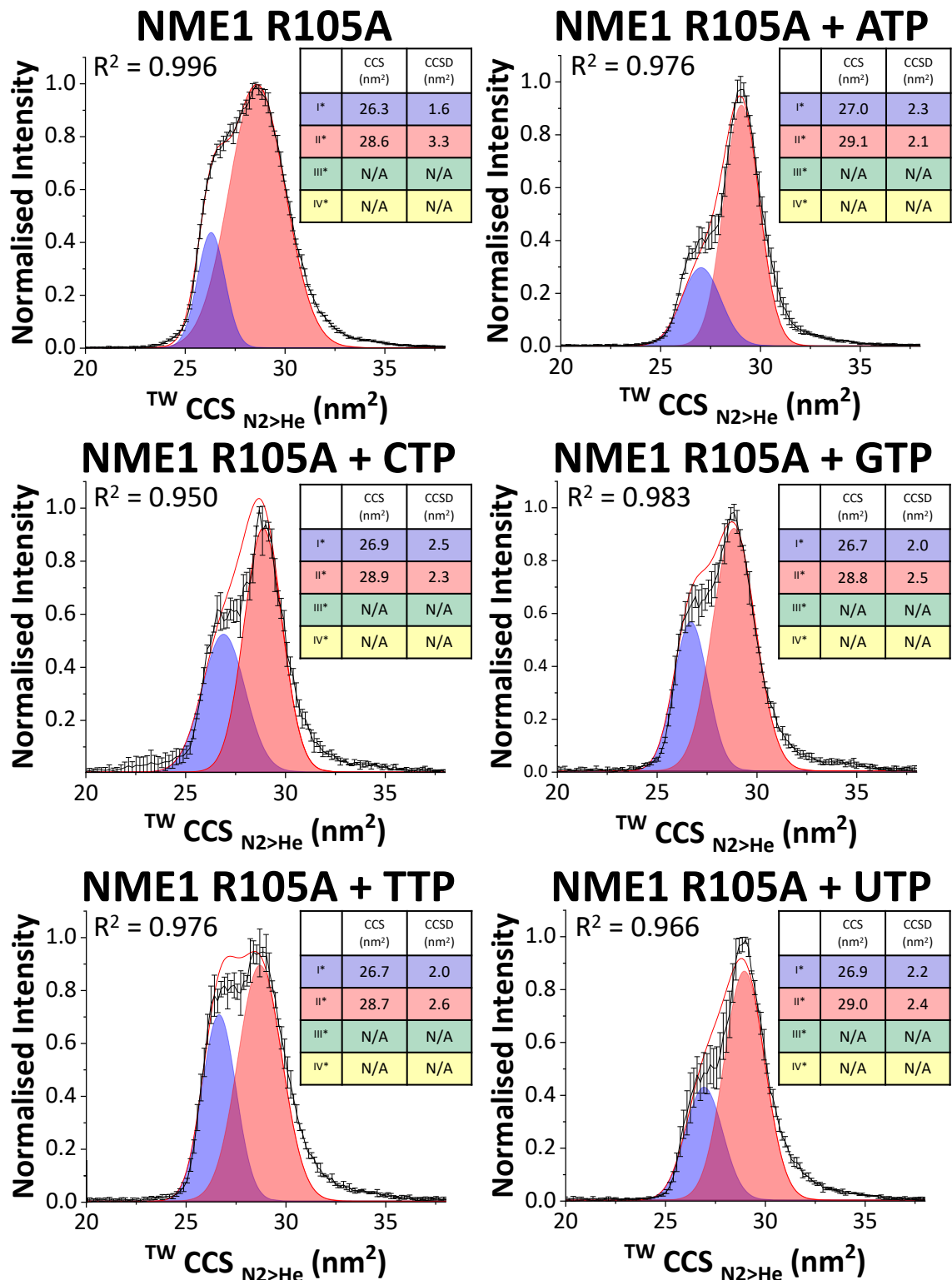
Figure 5.17. CCS distribution of NME1 R105A + nucleotides. $^{TW}CCS_{N2\rightarrow He}$ Gaussian fitting [M+12 H]$^{12+}$ species of NME1 WT + ATP, CTP, GTP, TTP or UTP. Nucleotide was added in 10-molar excess to protein and incubated for 10-minutes at R.T. Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed using the Fit Peaks Pro function in Origin (Version 2021b), with R^2 values listed.
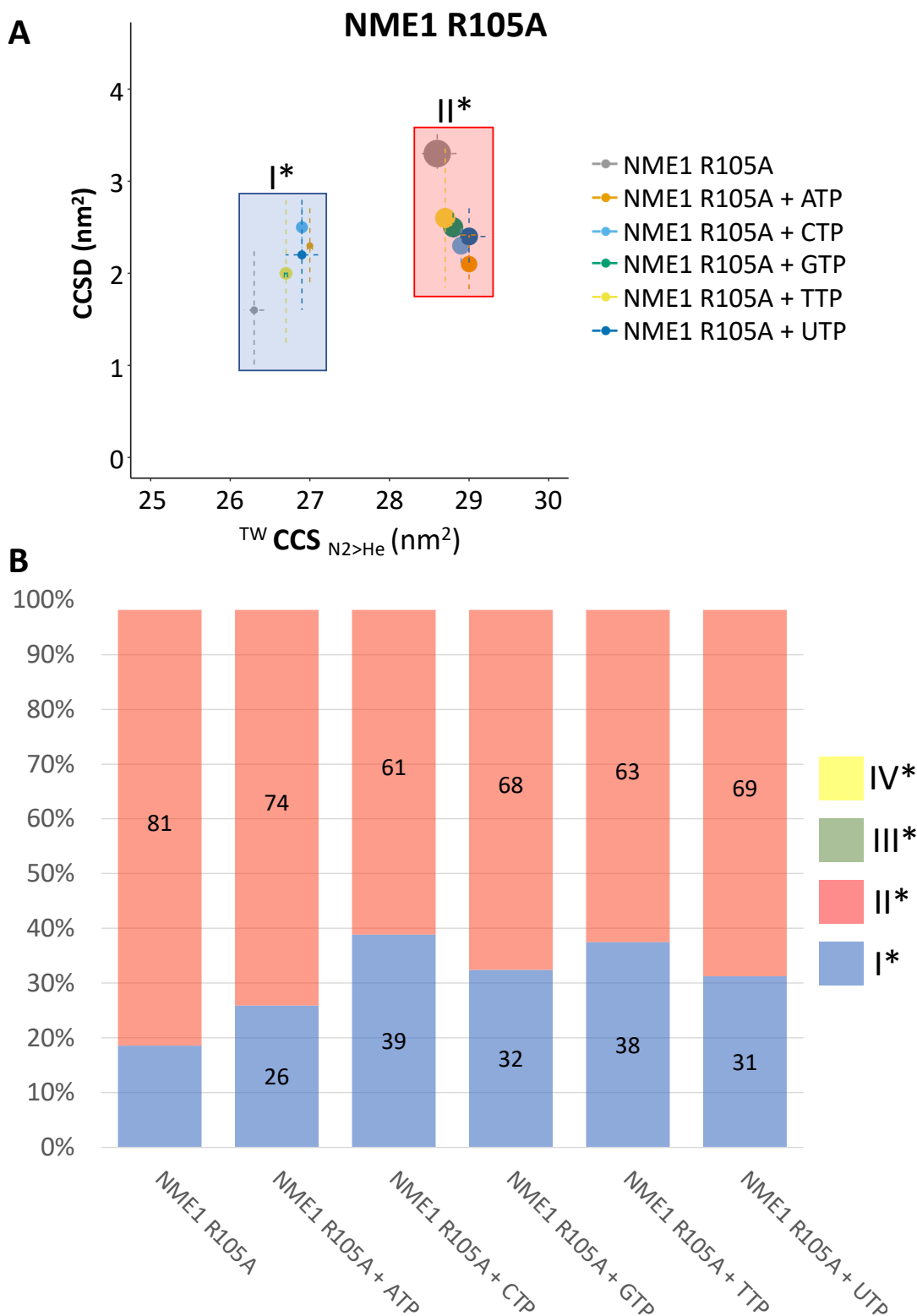
Figure 5.18. Conformational space adopted by NME1 R105A alone and in the presence of different nucleotides. A) Proportional scatter plots (CCS (nm^2) versus CCSD (nm^2)) for the different conformational states (as determined by Gaussian fitting in Figure 5.17) for NME1 R105A + nucleotides. Size of dot representative of area. **B)** % area of four different conformational states (as determined by Gaussian fitting in Figure 5.17): I (blue), II (red), III (green), IV (yellow) for NME1 R105A alone or in the presence of different nucleotides as indicated. Average % area presented from 3 individual experiments.

NME1 H118A/R105A showed a similar charge state distribution to NME1 R105A, with charge states 14+ to 11+ observed and 12+ being predominant (Figure 5.19). Following deconvolution, this variant appeared natively as a dimer, with an observed mass of ~ 33 kDa. Based on the deconvolution findings for the single residue variants, NME1 H118A and NME1 R105A, where H118A formed as a hexamer and R105A as a dimer, it would suggest that the mutation of R105 is the cause of this switch from hexamer to dimer. Nucleotide addition had no effect on the propensity of this NME1 variant to form dimer, with all yielding a complex of ~ 35 kDa (Figure 5.20).

NME1 H118A/R105A (17152 Da)
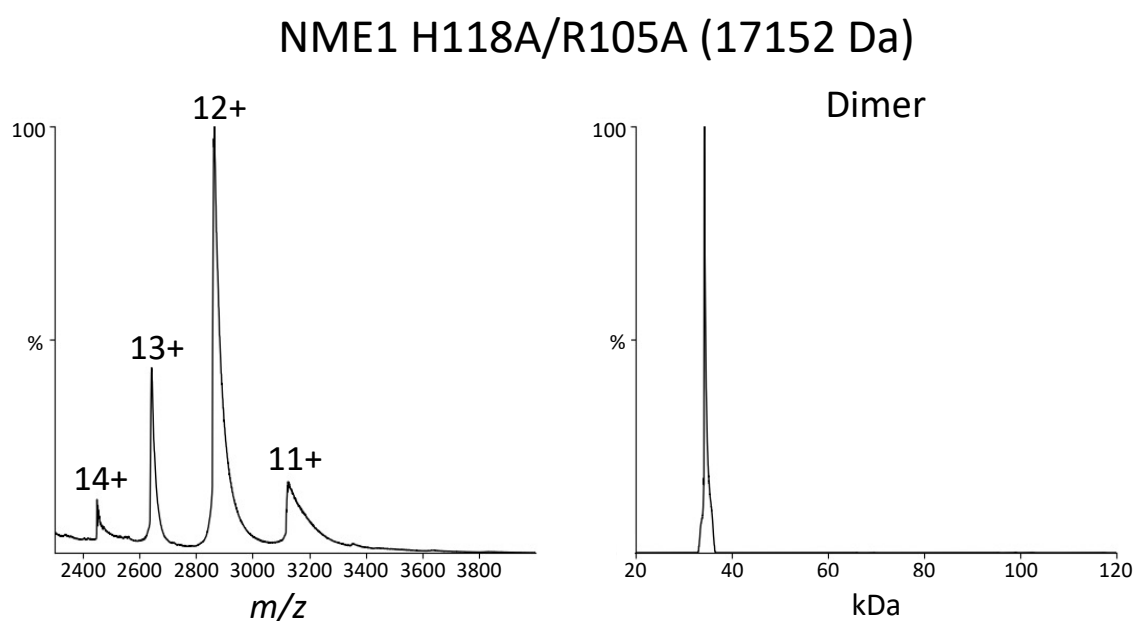
Figure 5.19. NME1 H118A/R105A natively forms as a dimer. Native ESI mass spectrum of NME1 H118A/R105A (left) and oligomeric determination showing dimer form (right). Spectrum list was copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.
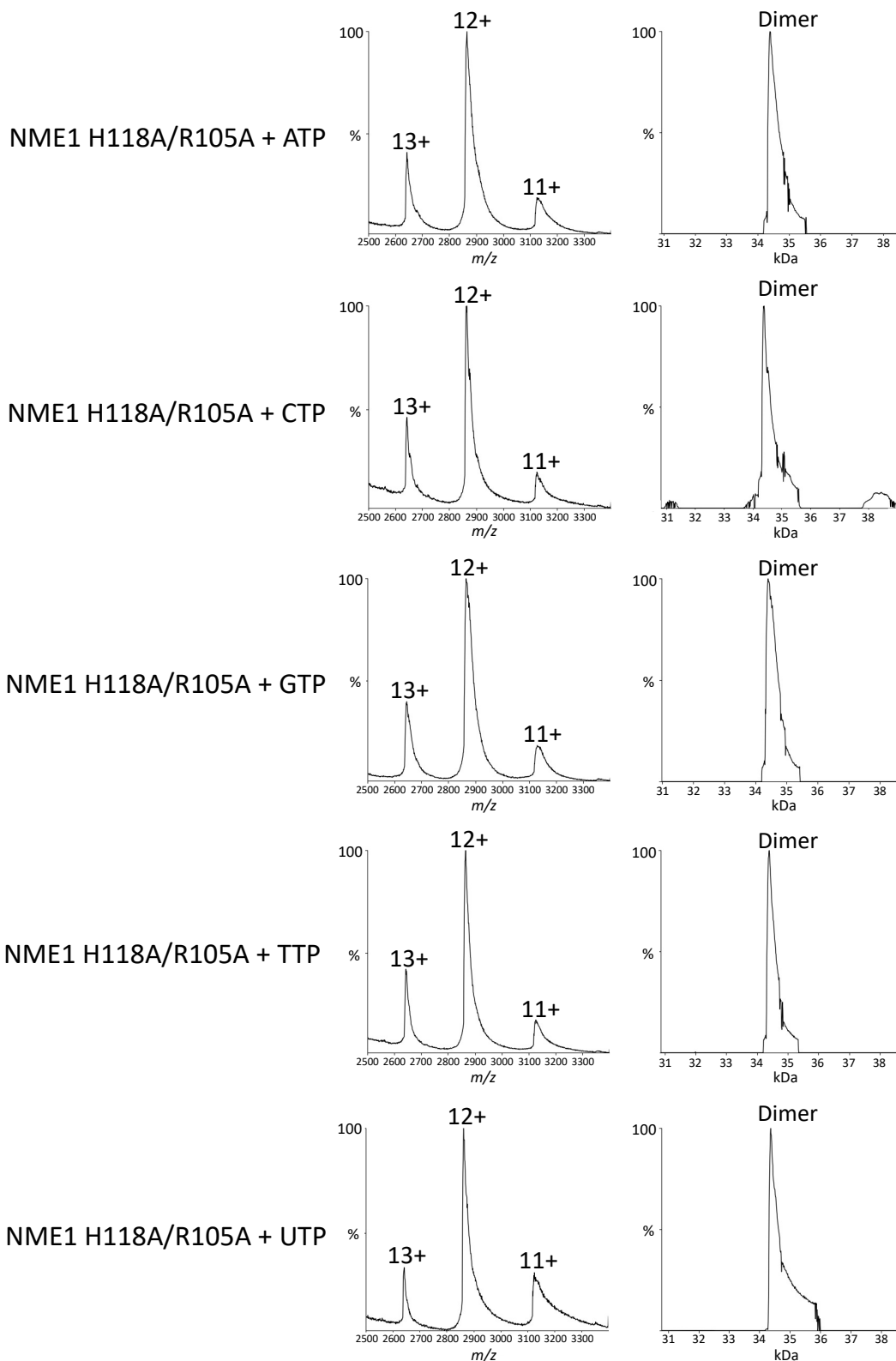
Figure 5.20. Addition of nucleotides does not alter the oligomeric state of NME1 H118A/R105A. Native ESI mass spectrum of NME1 H118A/R105A + nucleotide (left) and oligomeric determination showing dimer form (right) in the presence of each nucleotide. Spectrum lists were copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

Gaussian fitting of the 12+ charge state for NME1 H118A/R105A revealed the presence of two conformers (III* & IV*), with CCS values of 25.1 nm^2 and 27.2 nm^2 (Figure 5.21). These CCS values were similar to the two smallest conformers observed for the previous dimer form of NME1 R105A of 26.3 nm^2 and 28.6 nm^2, albeit slightly smaller, suggesting that the double mutant is more compact than the single R105A mutated variant. Conformer III* was the most abundant for NME1 H118A/R105A without the presence of nucleotide (61%), which was in contrast to NME1 R105A, where conformer III* was not observed (Figure 5.22).

Interestingly, the CCS landscape of NME1 H118A/R105A revealed nucleotide-dependent differences. While conformer II* and III* (at ~25.1-25.7 nm^2 and 27.2-29.0 nm^2 respectively) were observed in the absence and presence of all nucleotides for NME1 H118A/R105A, ATP, CTP/GTP/TTP also induced a smaller conformer at 22.4-22.8 nm^2. Moreover, CTP and UTP additionally invoked much smaller (20.7 nm^2) and larger (30.1 nm^2) conformational states respectively. This double mutated variant should result in little effect when bound to a nucleotide due to the absence of the residues responsible for auto-phosphorylation and catalytic activity. However, this was not observed with this data set, with a large degree of differences seen between the relative conformer abundancies when each nucleotide was bound, which could potentially suggest different binding modes between the nucleotides.

NME1 H118A/R105A

R² = 0.991

	CCS (nm²)	CCSD (nm²)
I*	N/A	N/A
II*	N/A	N/A
III*	25.1	2.0
IV*	27.2	2.9

NME1 H118A/R105A + ATP

R² = 0.984

	CCS (nm²)	CCSD (nm²)
I*	N/A	N/A
II*	22.8	2.1
III*	25.7	3.1
IV*	29.0	1.8

NME1 H118A/R105A + CTP

R² = 0.997

	CCS (nm²)	CCSD (nm²)
I*	20.7	2.2
II*	22.6	1.9
III*	25.6	2.0
IV*	27.9	2.4

NME1 H118A/R105A + GTP

R² = 0.987

	CCS (nm²)	CCSD (nm²)
I*	N/A	N/A
II*	22.5	2.7
III*	25.6	2.7
IV*	28.3	2.3

NME1 H118A/R105A + TTP

R² = 0.994

	CCS (nm²)	CCSD (nm²)
I*	N/A	N/A
II*	22.4	2.8
III*	25.7	1.9
IV*	27.9	2.3

NME1 H118A/R105A + UTP

R² = 0.988

	CCS (nm²)	CCSD (nm²)
I*	N/A	N/A
II*	25.3	2.3
III*	27.5	2.5
IV*	30.1	2.0

Figure 5.21. CCS distribution of NME1 H118A/R105A + nucleotides. $^{TW}CCS_{N2\rightarrow He}$ Gaussian fitting [M+12H] $^{12+}$ species of NME1 WT + ATP, CTP, GTP, TTP or UTP. Nucleotide was added in 10-molar excess to protein and incubated for 10-minutes at R.T. Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed using the Fit Peaks Pro function in Origin (Version 2021b), with R² values listed.
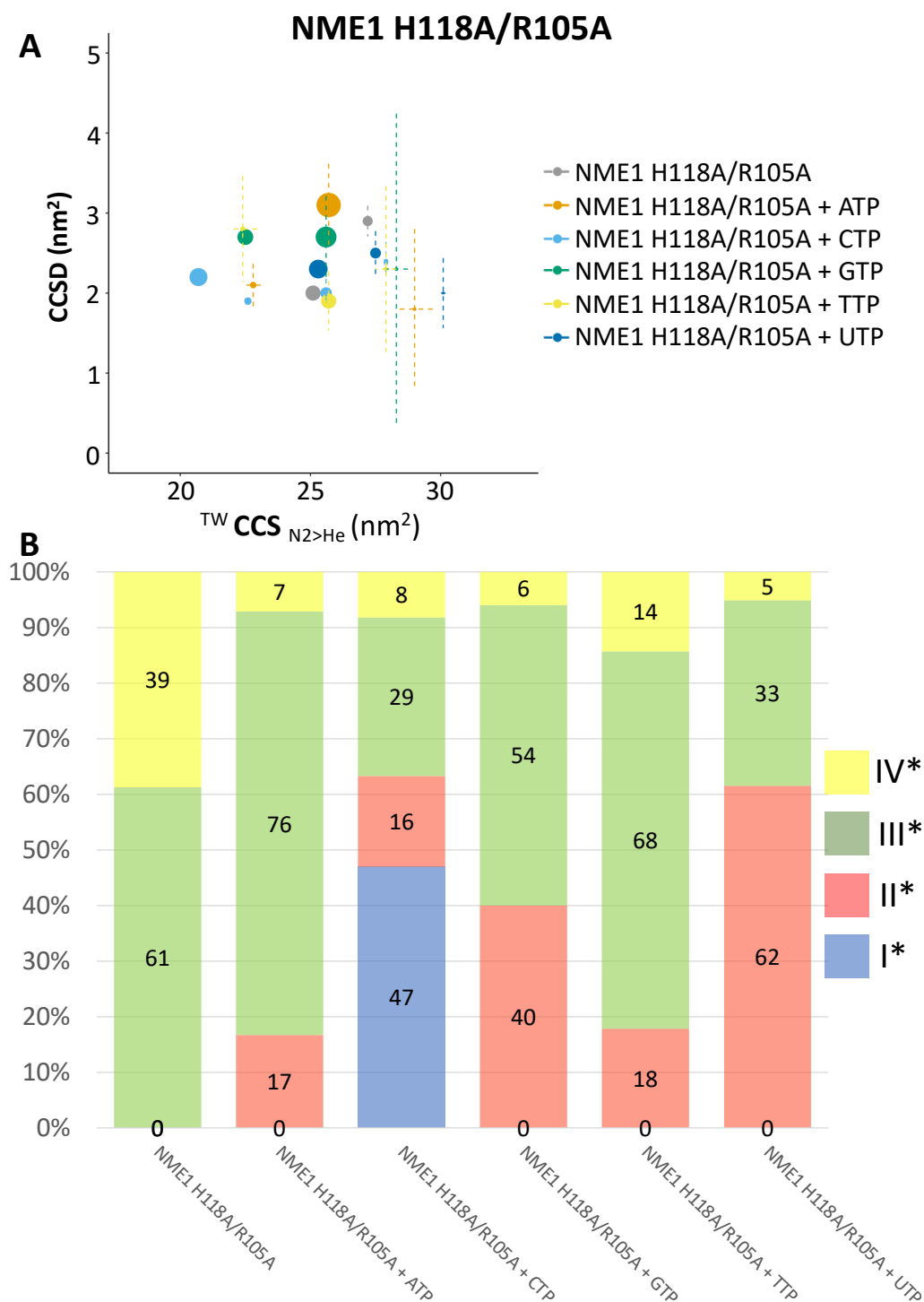
Figure 5.22. Conformational space adopted by NME1 H118A/R105A alone and in the presence of different nucleotides. A) Proportional scatter plots (CCS (nm^2) versus CCSD (nm^2)) for the different conformational states (as determined by Gaussian fitting in Figure 5.21) for NME1 H118A/R105A + nucleotides. Size of dot representative of area. Conformers not represented by different boxes due to level of overlap between conformational states. **B)** % area of four different conformational states (as determined by Gaussian fitting in Figure 5.21): I (blue), II (red), III (green), IV (yellow) for NME1H118A/R105A alone or in the presence of different nucleotides as indicated. Average % area presented from 3 individual experiments.

NME1 C109A had a charge state distribution between 22+ and 20+, with relatively poor signal quality rendering limited peak resolution between the predominant 21+ and 20+ charge states (Figure 5.23). The deconvolution of the mass spectra resulted in a broad peak around ~ 105 kDa, representative of hexamer formation. Similar to NME1 WT, the addition of nucleotides had no effect on the hexameric state of NME1 C109A (Figure 5.24).
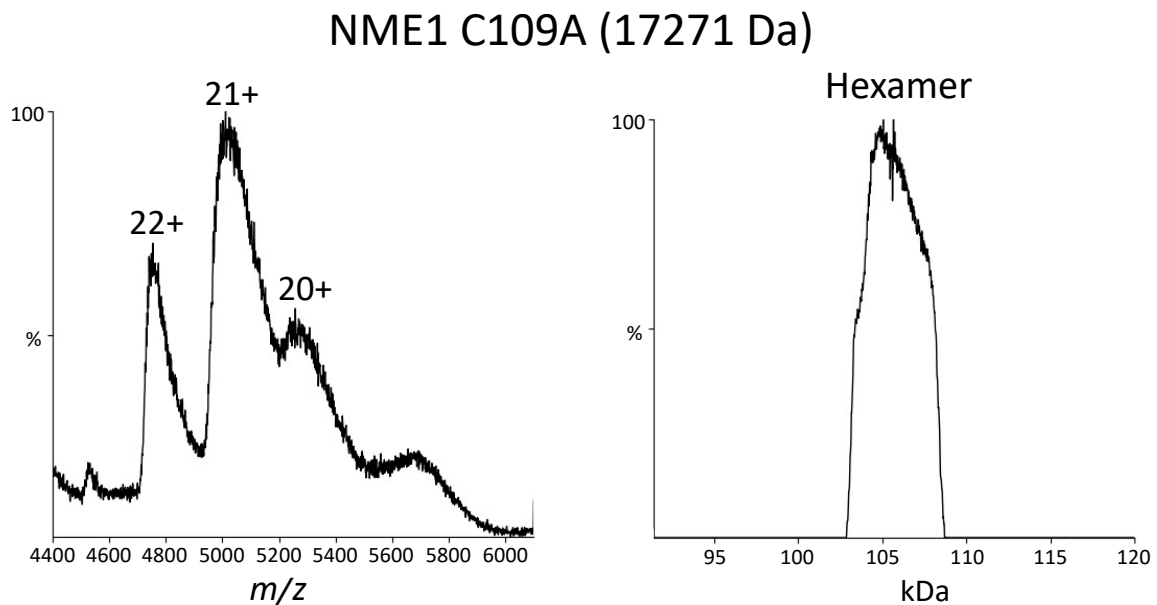
Figure 5.23. NME1 C109A natively forms as a hexamer. Native ESI mass spectrum of NME1 C109A (left) and oligomeric determination showing hexamer form (right). Spectrum list was copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.
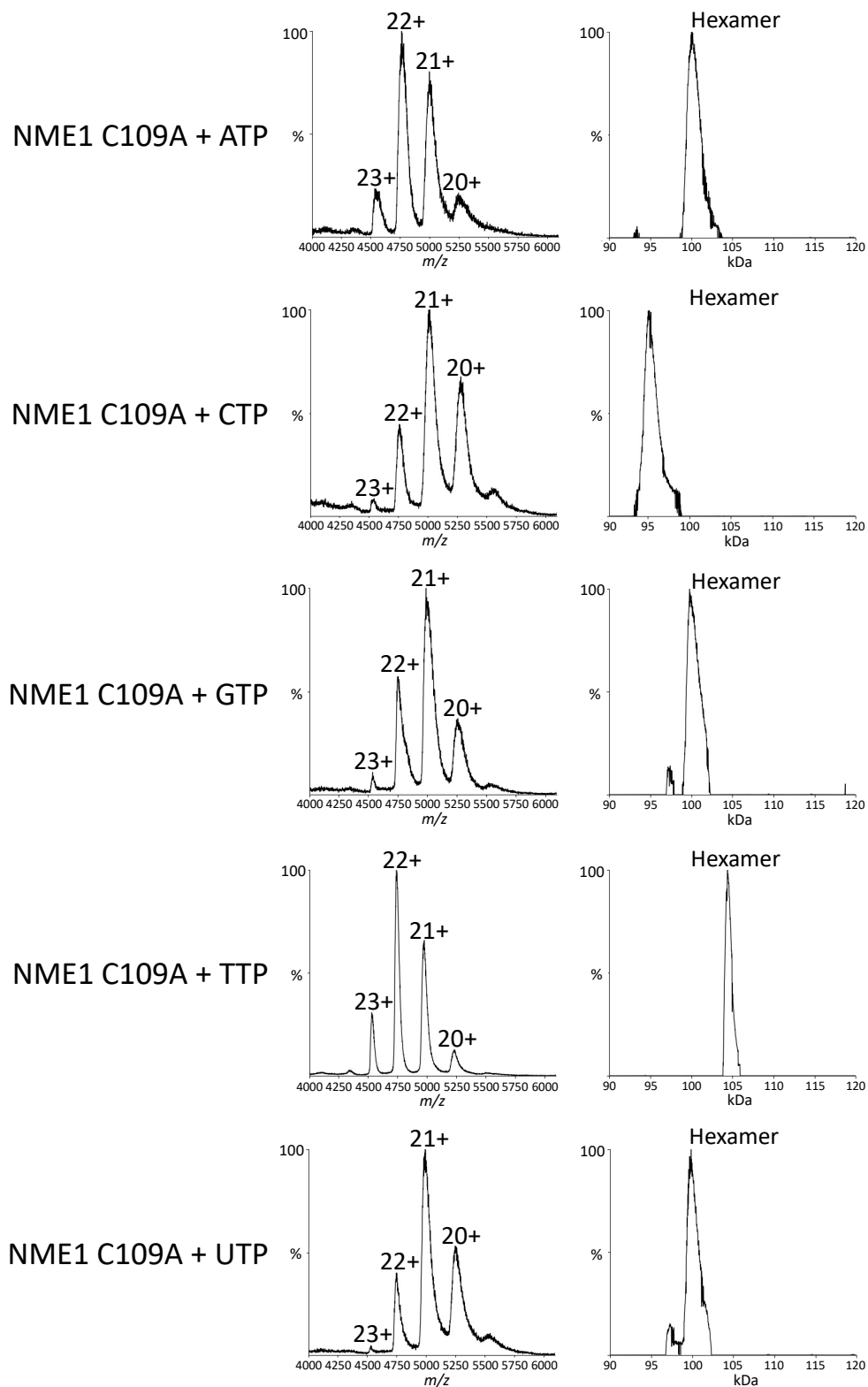
Figure 5.24. Addition of nucleotides does not alter the oligomeric state of NME1 C109A. Native ESI mass spectrum of NME1 C109A + nucleotide (left) and oligomeric determination showing hexamer form (right) in the presence of each nucleotide. Spectrum lists were copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.
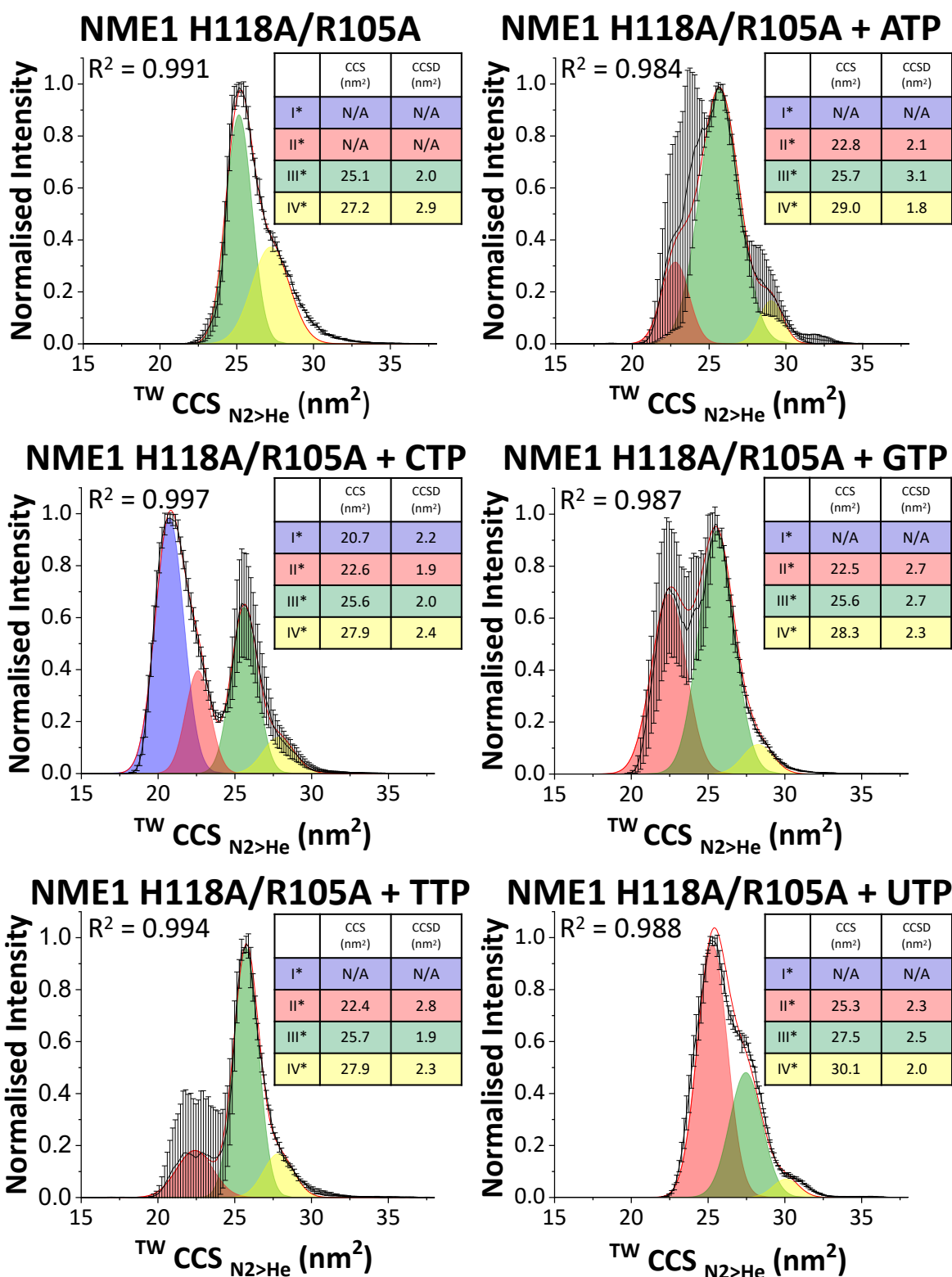
Gaussian fitting of the CCS for the 21+ charge state of NME1 C109A revealed CCS values that were similar to the previous hexameric forms of NME1, WT and H118A. The most predominant conformer observed was conformer II (53.5 nm^2) (Figure 5.25), which was similar to NME1 WT conformer III (53.9 nm^2). In contrast to WT and H118A, C109A revealed the presence of the smaller conformer (I) at 51.8 nm^2, however, the CCS for this conformer closely relates to conformer II for WT (52.0 nm^2) and H118A (50.7 nm^2).

Upon addition of each nucleotide to NME1 C109A, a single conformer (II) distribution was observed (Figure 5.26) ranging from 52.7 – 53.1 nm^2 compared to the slightly larger 53.5 nm^2 with DMSO control. In addition to little effect in the CCS distribution of conformer II between conditions, the CCSD values for this conformer observed minimal variation ranging from 2.2 – 2.6 nm^2 with nucleotides compared to 2.3 nm^2 for the DMSO control. The addition of nucleotides appeared to have less impact on the CCS values for NME1 C109A compared to WT, where a slight decrease was observed. This chapter previously demonstrated that C109A + nucleotide conditions were not as effective as acting as a phosphate donor for auto-phosphorylation with lower levels of phosphorylation observed. Therefore, this could be a factor in the little alteration that is present in the protein conformational landscape with the nucleotide-bound conditions when compared to DMSO control.
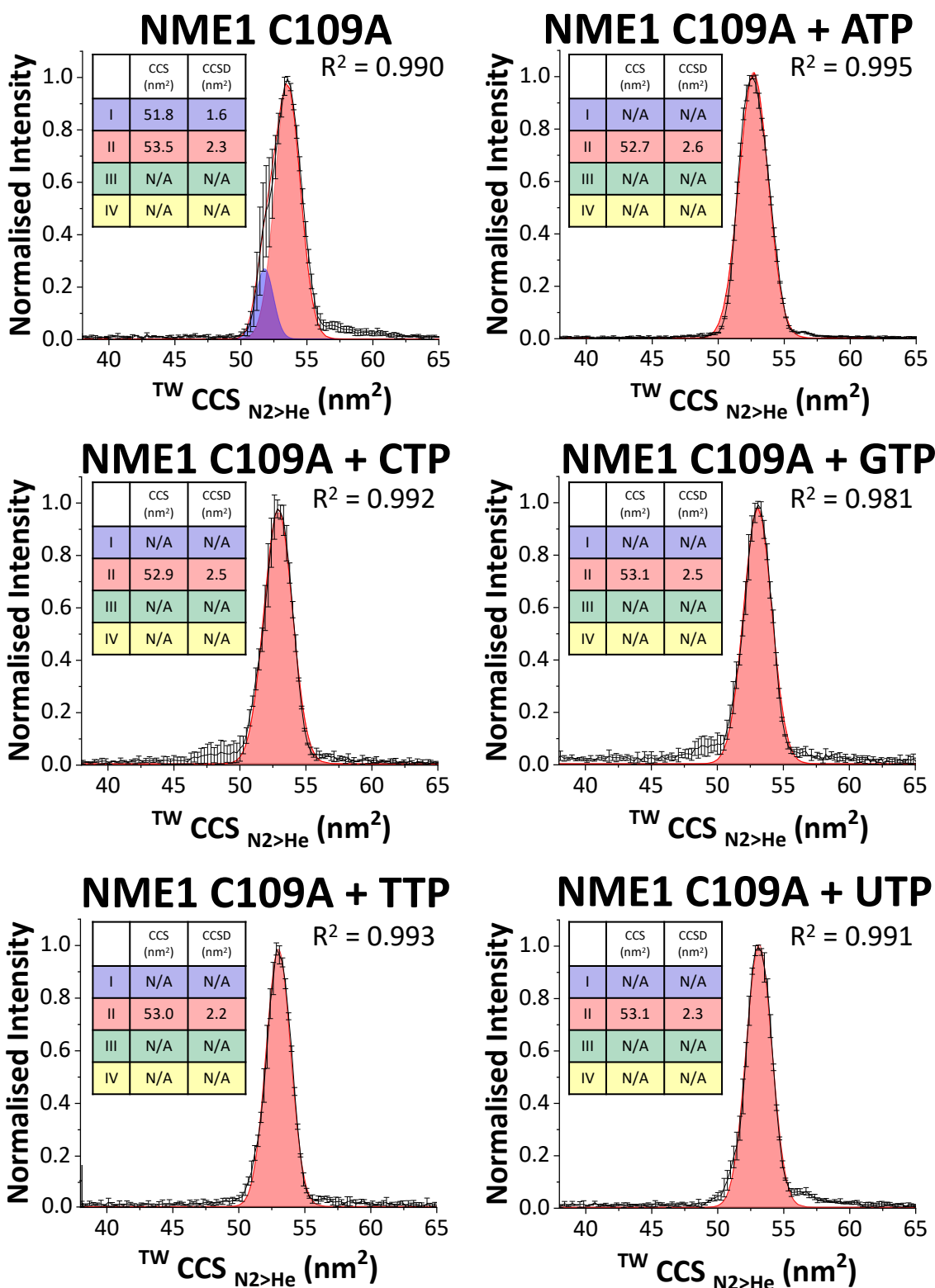
Figure 5.25. CCS distribution of NME1 C109A + nucleotides. $^{TW}CCS_{N2\rightarrow He}$ Gaussian fitting [M+21H] $^{21+}$ species of NME1 WT + ATP, CTP, GTP, TTP or UTP. Nucleotide was added in 10-molar excess to protein and incubated for 10-minutes at R.T. Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed using the Fit Peaks Pro function in Origin (Version 2021b), with R^2 values listed.
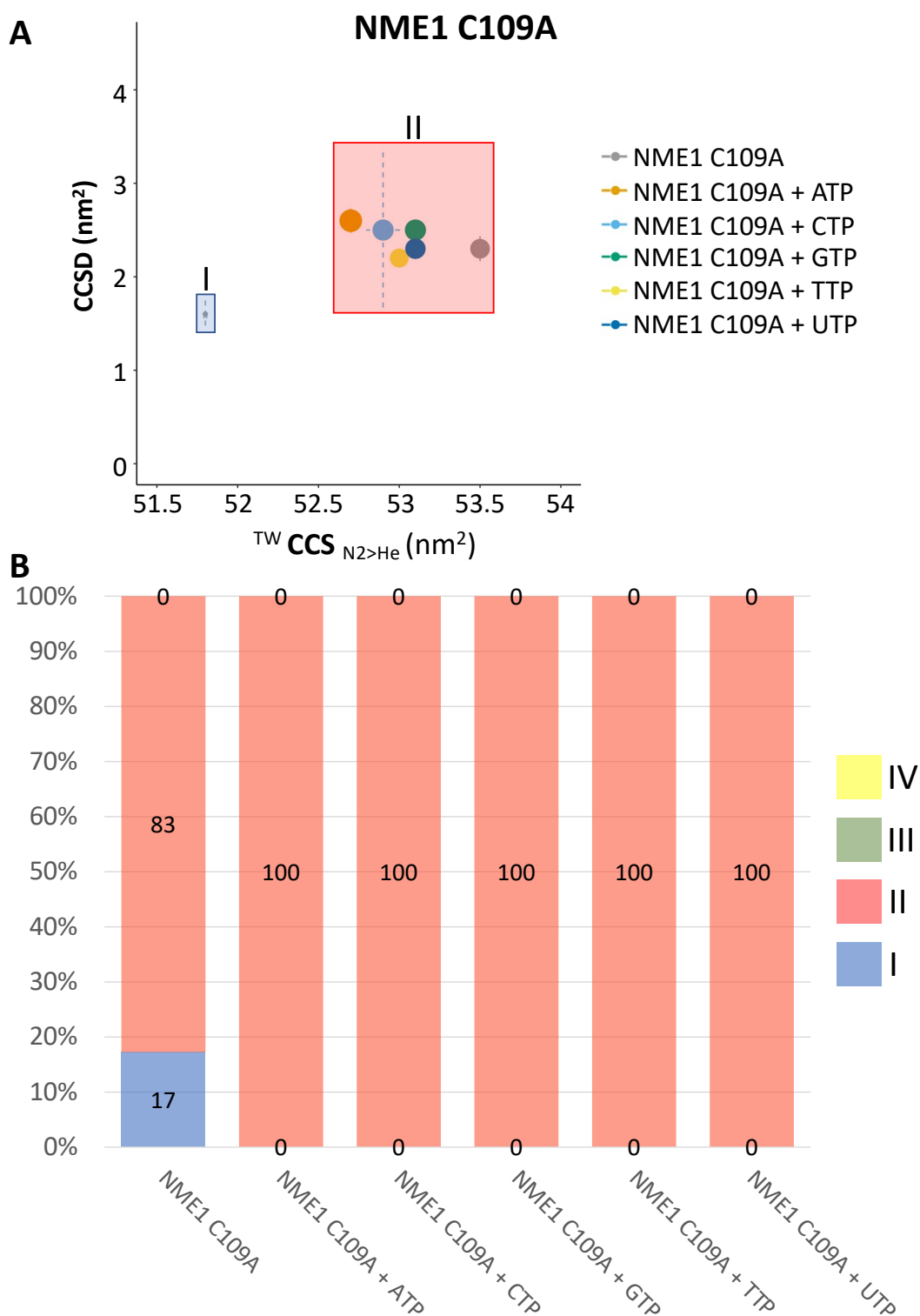
Figure 5.26. Conformational space adopted by NME1 C109A alone and in the presence of different nucleotides. A) Proportional scatter plots (CCS (nm^2) versus CCSD (nm^2)) for the different conformational states (as determined by Gaussian fitting in Figure 5.25) for NME1 C109A + nucleotides. Size of dot representative of area. **B)** % area of four different conformational states (as determined by Gaussian fitting in Figure 5.25): I (blue), II (red), III (green), IV (yellow) for NME1 C109A alone or in the presence of different nucleotides as indicated. Average % area presented from 3 individual experiments.

5.2.3 Effect of compounds on NME1

To identify small molecule compounds that might have an effect on the activity and/or stability of NME1, a screen was carried out against a 1000 compound FDA-approved drug library (SelleckChem (211)) using DSF by Alice Clubbs Coldron. As NME proteins are not standard protein kinases, the FDA-approved drug library rather than a standard kinase inhibitor library was screened as it contains a variety of classes of small molecule drugs. The FDA-approved drug library contains a range of compounds associated with various diseases, including oncology, immunology, and anti-inflammatory, and therefore was deemed appropriate to screen for NME1 inhibitors. The binding of the compounds to NME1 was indicated by a shift in the melting temperature of the protein compared to that of the protein in the absence of any compound (buffer control). In order to assess how the compounds affected the stability and conformation of NME1 WT, NME1 C109A was also analysed by DSF in the presence of FDA library hits (ΔTm +/- 1.0 °C). NME1 C109A is hypothesised to have more resistance to hexamer dissociation as a result of its inability to form dimer stabilising disulphide bonds. Compounds that induced large differences between NME1 WT and NME1 C109A, and that had the potential to oxidise NME1 by virtue of a carbonyl group were thus selected for further analysis and are represented by a red dot (Figure 5.27A). Compounds shown as blue dots with large negative shifts for NME1 WT were not selected based on these compounds demonstrating either no or marginal T_m differences in NME1 C109A, therefore suggesting they do not change the stability of C109A and not applicable to test the hypothesis.

The ΔTm of NME1 WT compared to C109A in the presence of the selected compounds are shown in Figure 5.27B. The addition of several of the compounds to NME1 WT induced a shift of -1 °C or greater, suggesting that the compounds are having an effect on the stability of NME1 WT (Figure 5.27B/Table 5.2). Following the DSF analysis, it was recognised that many of the compounds consisted of functional groups that can accept electrons (Figure 5.28), therefore enabling them to act as oxidising agents. S1949 (Menadione), the compound shown to induce one of the largest negative shifts at -5.8 °C, is a known oxidising agent (212). The screen of the twenty compounds with NME1 C109A overall revealed less of a negative shift (Figure 5.27C) in comparison to those observed for NME1 WT (with exception to S1082 and S1457). S1730 showed a similar temperature shift for both NME1 WT and C109A, therefore this compound was selected to act as a control between the two forms of NME1. As less of a

negative shift was observed with NME1 C109A + compounds, this confirmed the initial hypothesis that this mutated protein could be causing some resistance to dissociation and/or unfolding. To explore this hypothesis further, native MS was carried out to determine the oligomeric state of NME1 WT and C109A when bound to each compound.
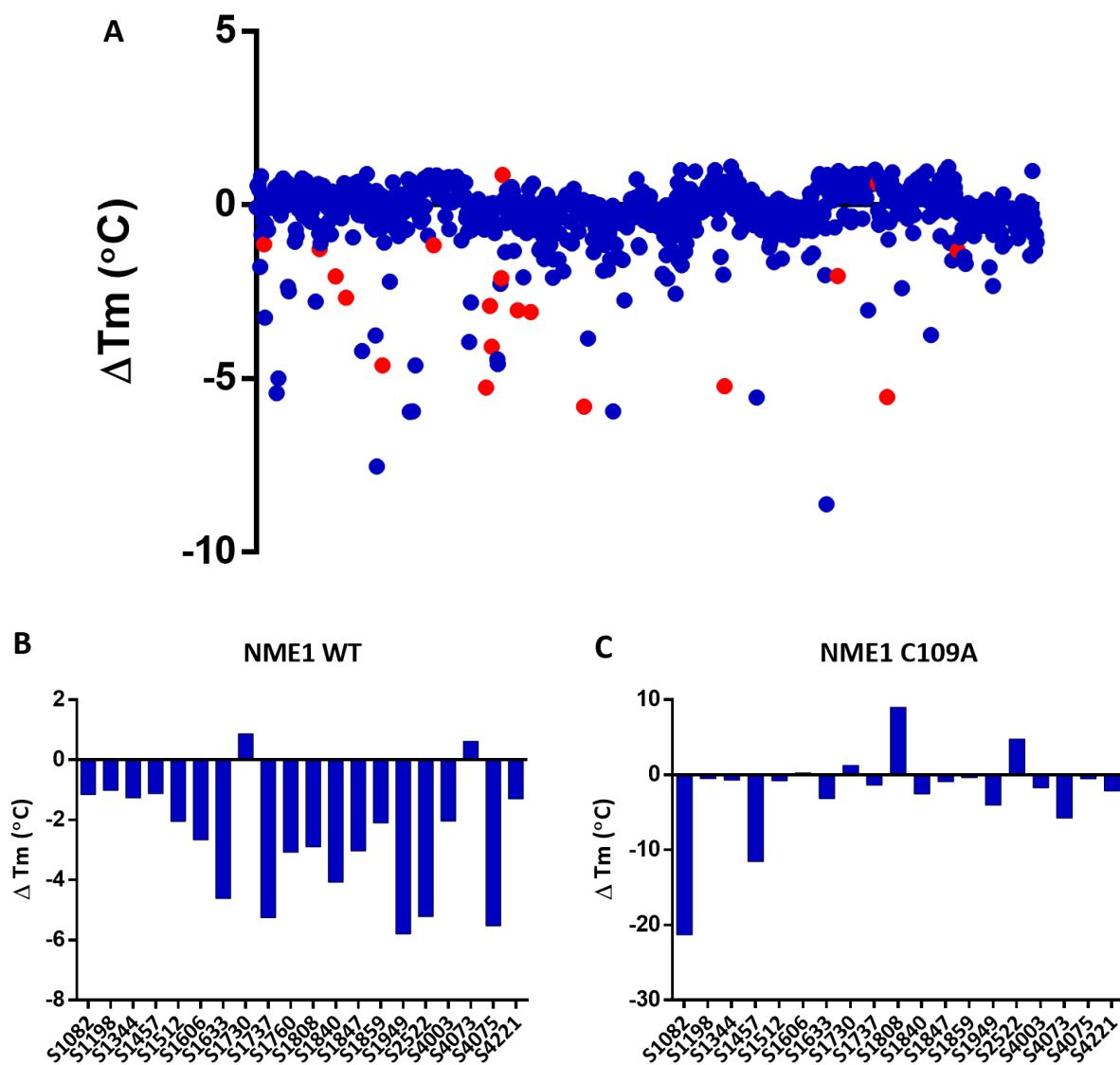
Figure 5.27. NME1 WT and NME1 C109A screening with the food drug approved (FDA) library. (A) WT NME1 was incubated with 200 µM of an FDA drug compound and the thermal stability was determined in the presence of these compounds using differential scanning fluorimetry (DSF). Difference in melting temperature (ΔTm) relative to control is presented for NME1 WT (B) and NME1 C109A (C) in the presence of selected compounds based on their effect on NME1 WT and C109A, alongside their oxidising potential by assessing their chemical structures (red dots). Data provided by Alice Clubbs Coldron, a PhD student with Prof Patrick Eyers (University of Liverpool).

Compound	ΔTm (℃) for NME1 WT	ΔTm (℃) for NME1 C109A
S1082 (Vismodegib)	- 1.60	- 21.33
S1198 (Irinotecan)	- 1.02	- 0.53
S1344 (Glimepiride)	- 1.27	- 0.72
S1457 (Atazanavir sulfate)	- 1.13	- 11.55
S1512 (Tadalafil)	- 2.05	- 0.82
S1606 (Clotrimazole)	- 2.67	+ 0.29
S1633 (Zafirlukast)	- 4.62	- 3.19
S1730 (Indapamide)	0.66	+ 1.27
S1737 (Prednisolone)	- 5.26	- 1.37
S1808 (Nifedipine)	- 1.84	+ 9.03
S1840 (Lomustine)	- 4.08	- 2.55
S1760 (Rifapentine)	- 3.08	- 4.78
S1847 (Clemastine fumarate)	- 3.87	- 0.92
S1859 (Diethylstilbestrol)	- 2.02	- 0.37
S1949 (Menadione)	- 5.80	- 5.74
S2522 (Adrenaline)	- 5.22	4.77
S4003 (Lithocholic acid)	- 2.05	- 1.76
S4221 (Benzbromarone)	- 1.30	- 2.14
S4073 (Aminosalicylate sodium)	0.62	- 1.76
S4075 (Zinc pyrithione)	- 5.53	- 0.55

Table 5.2. Difference in melting temperature (ΔTm) for NME1 WT and C109A + compounds in FDA library screen. DSF screen carried out as described in Figure 5.27 legend.

S1082 (Vismodegib) 421 Da
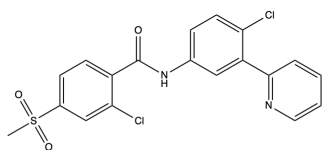
S1198 (Irinotecan) 587 Da
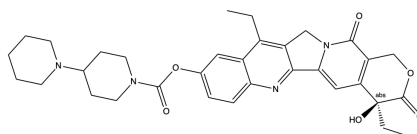
S1344 (Glimepiride) 491 Da
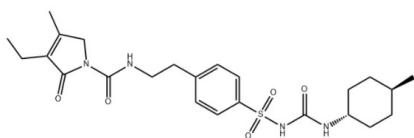
S1457 (Atazanavir sulfate) 705 Da
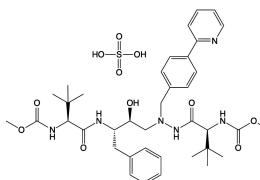
S1512 (Tadalafil) 389 Da
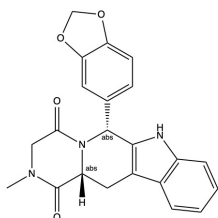
S1606 (Clotrimazole) 345 Da
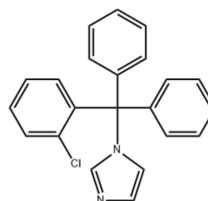
S1633 (Zafirlukast) 576 Da
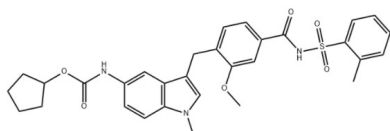
S1730 (Indapamide) 366 Da
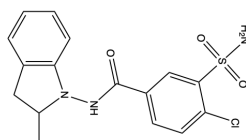
S1737 (Prednisolone) 360 Da
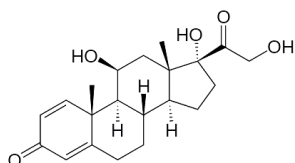
S1760 (Rifapentine) 877 Da
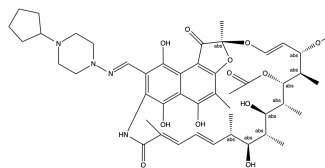
Figure 5.28. Chemical structures for FDA screened compounds (Part 1). Structures produced in ChemDraw (Version 19.1).

S1808 (Nifedipine) 346 Da
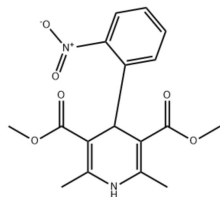
S1840 (Lomustine) 234 Da
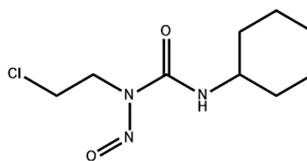
S1847 (Clemastine fumarate) 460 Da
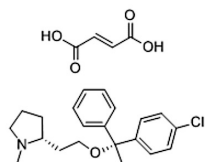
S1859 (Diethylstilbestrol) 268 Da
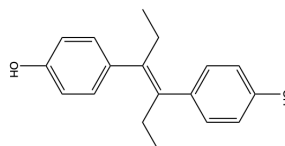
S1949 (Menadione) 172 Da
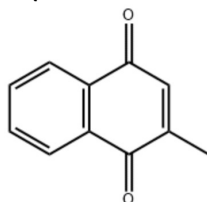
S2522 (Adrenaline) 183 Da
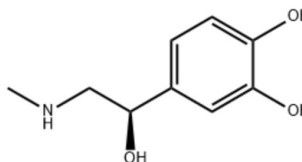
S4003 (Lithocholic acid) 377 Da
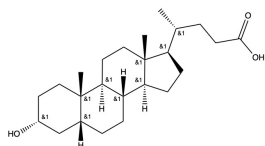
S4073 (Aminosalicylate sodium) 152 Da
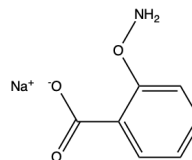
S4075 (Zinc pyrithione) 318 Da
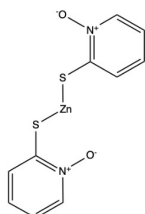
S4221 (Benzbromarone) 424 Da
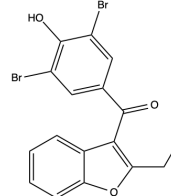
Figure 5.28. Chemical structures for FDA screened compounds (Part 2). Structures produced in ChemDraw (Version 19.1).

5.2.4 Effect of compounds on NME1 oligomeric state

The DSF screen provided a useful understanding to determine which compounds bound to NME1. However, the change in melting temperature observed by DSF does not necessarily relate to structural changes that could be occurring, multimer formation, or activity. Therefore, native MS was carried out determine whether any of the compounds were inducing structural changes in NME1 and whether the hexameric state is compromised. To assess the ratio of compound:protein required to induce any change in oligomeric state, NME1 WT was incubated with either 5x, 8x or 10x molar excess of S1949 for ten minutes prior to native MS (Figure 5.29). The addition of S1949 compound at 5x molar concentration to NME1 revealed a mass of ~ 100 kDa (represented by the number 6) as the most abundant peak, with a smaller peak at ~ 50 kDa (represented as 3) also present. This suggests that the S1949 at 5x molar excess was sufficient to initiate hexameric dissociation, but not sufficient for complete dissociation. More complete dissociation was achieved with 8x molar excess, yielding ions representative of NME1 complexes between ~ 25-40 kDa (represented by 2) and higher peaks at ~ 75, 90 and 100 kDa (represented by 4, 5 and 6). The higher concentration of 8x molar excess induced dissociation to a mass representative of the dimer form. However, the peaks at the larger masses indicated that there was still some hexamer form remaining and this concentration was therefore not sufficient to induce complete dissociation to the dimer. Complete dissociation of the hexamer was achieved with 10x molar excess of S1949 and was thus used to test all the other compounds.
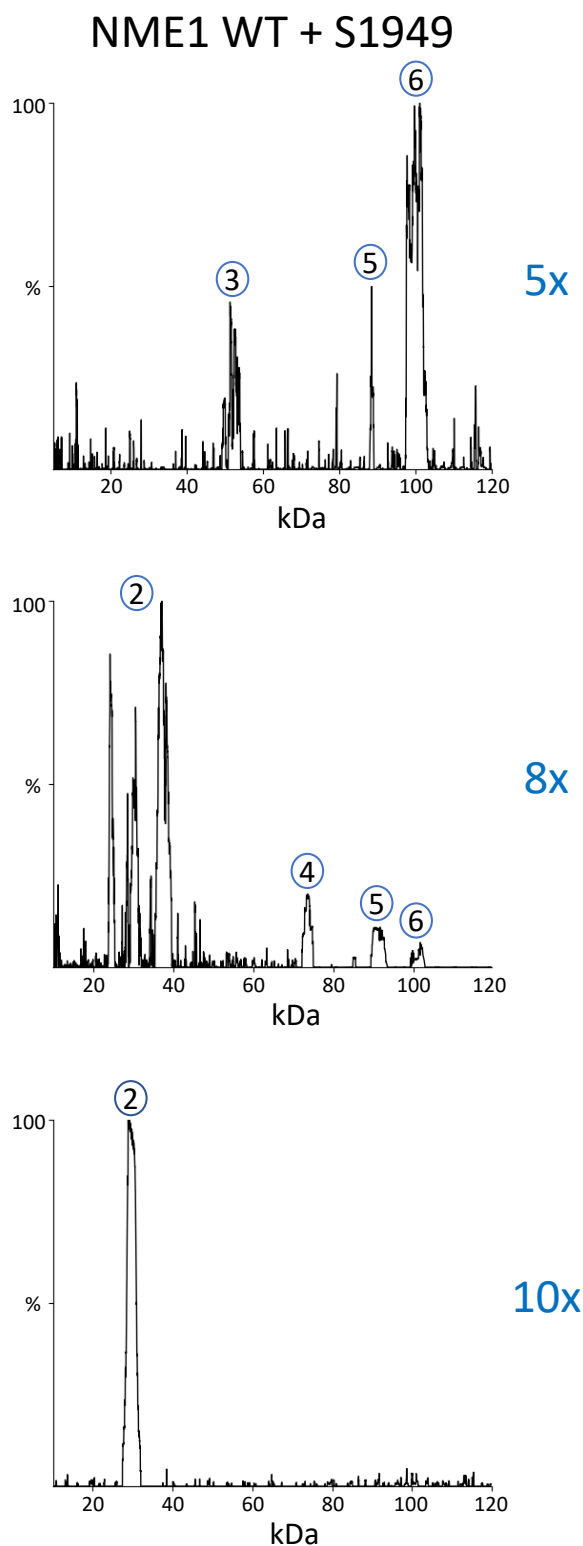
Figure 5.29. NME1 WT + S1949 compound shows dissociation to dimer at 10x excess concentration. Oligomeric determination of 5x, 8x and 10x excess of S1949. Number above peaks representative of multimer state, such as 2 for dimer and 6 for hexamer formation. Spectrum lists were copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

Following determination of the appropriate molar excess concentration of compound to use to induce changes on the oligomeric state, twenty compounds were incubated with NME1 for ten minutes prior to native MS analysis (refer to Appendix 5.1. for all native MS data). Upon addition of the compounds S1949 and S1082 to NME1 WT (Figure 5.30), peaks at the lower m/z region (2000 – 4400 *m/z*) were observed in comparison to the absence of compound, where peaks were observed between 4500 – 5500 *m/z*. In addition to peaks at the smaller *m/z* regions being present, peaks were also evident at the higher mass range of 5500 – 7500 *m/z*. Binding of S4075 to NME1 WT also demonstrated a wider *m/z* region for the observed peaks, however, the region was not as low as compared to S1949 and S1082, suggesting that the S4075 compound did not induce as much dissociation of the higher order NME1 complex. The change in the observed m/z values of the abundant peaks suggests that the compounds are binding to NME1 WT and potentially causing a change in the overall oligomeric state of the protein. Binding of the compounds to NME1 C109A also resulted in a change to the peaks observed in the native MS spectra, where more peaks are observed in a smaller *m/z* region when compared to the absence of compound.
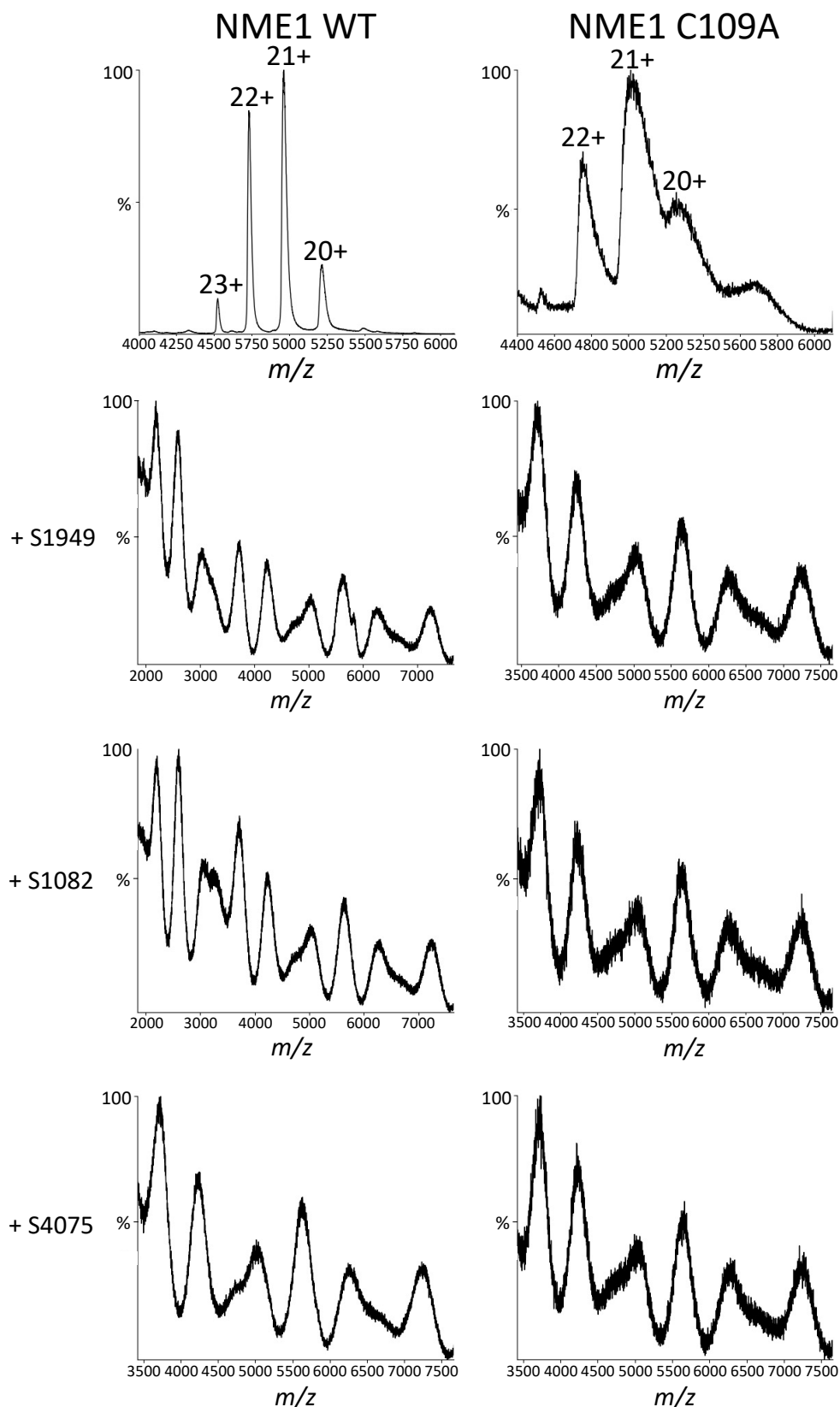
Figure 5.30. NME1 proteins bind compounds. Native ESI spectrum for NME1 WT (left) and C109A (right) + S1949, S1082 or S4075. Figures created by copying the spectrum lists into UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

To determine the oligomeric state of NME1 WT and C109A when bound to the compounds, the native MS spectra (Figure 5.30/Appendix 5.1) were deconvoluted using UniDec software. A mass range between 0 – 120 kDa was used to identify all multimers present up to the hexamer mass of native NME1, which was previously identified in the absence of any compound. Upon addition of the compounds S1949, S1082 and S4075 to NME1 WT, peaks between 30 – 40 kDa were observed (Figure 5.31), which suggests that the compounds induce dissociation to the dimer (theoretical mass of 34.6 kDa). For NME1 WT + S1949, no peaks in the larger mass region were observed, which suggests that this compound induces complete dissociation of NME1 WT.

Upon addition of the compounds to NME1 C109A, it was expected that the hexamer form would be most abundant, with this mutated variant expected to be resistant to oxidative stress and therefore, resistant to dissociation to the dimer. However, the binding of the S1949 and S1082 compounds revealed a peak at ~50 kDa, which would be indicative of a trimer form (theoretical trimer mass of 51.9 kDa). This shows that the NME1 C109A variant is partially resistant to dissociation induced by these compounds, suggesting a role for C109 in stabilisation of the higher order complex. Binding of S1949 and S1082 to NME1 C109A suggests that the kpn loop that is associated with the trimer interaction remains intact in comparison to when these compounds are bound to NME1 WT, where the kpn loop might be disordered, which would promote the dimer form. Binding of the S4075 compound to NME1 C109A yielded an abundant peak at ~ 100 kDa, representative of the hexamer form. Although the most abundant peak in the presence of S4075 was representative of the hexamer, smaller peaks (particularly around ~52 kDa) were evident suggesting some dissociation to the trimer.
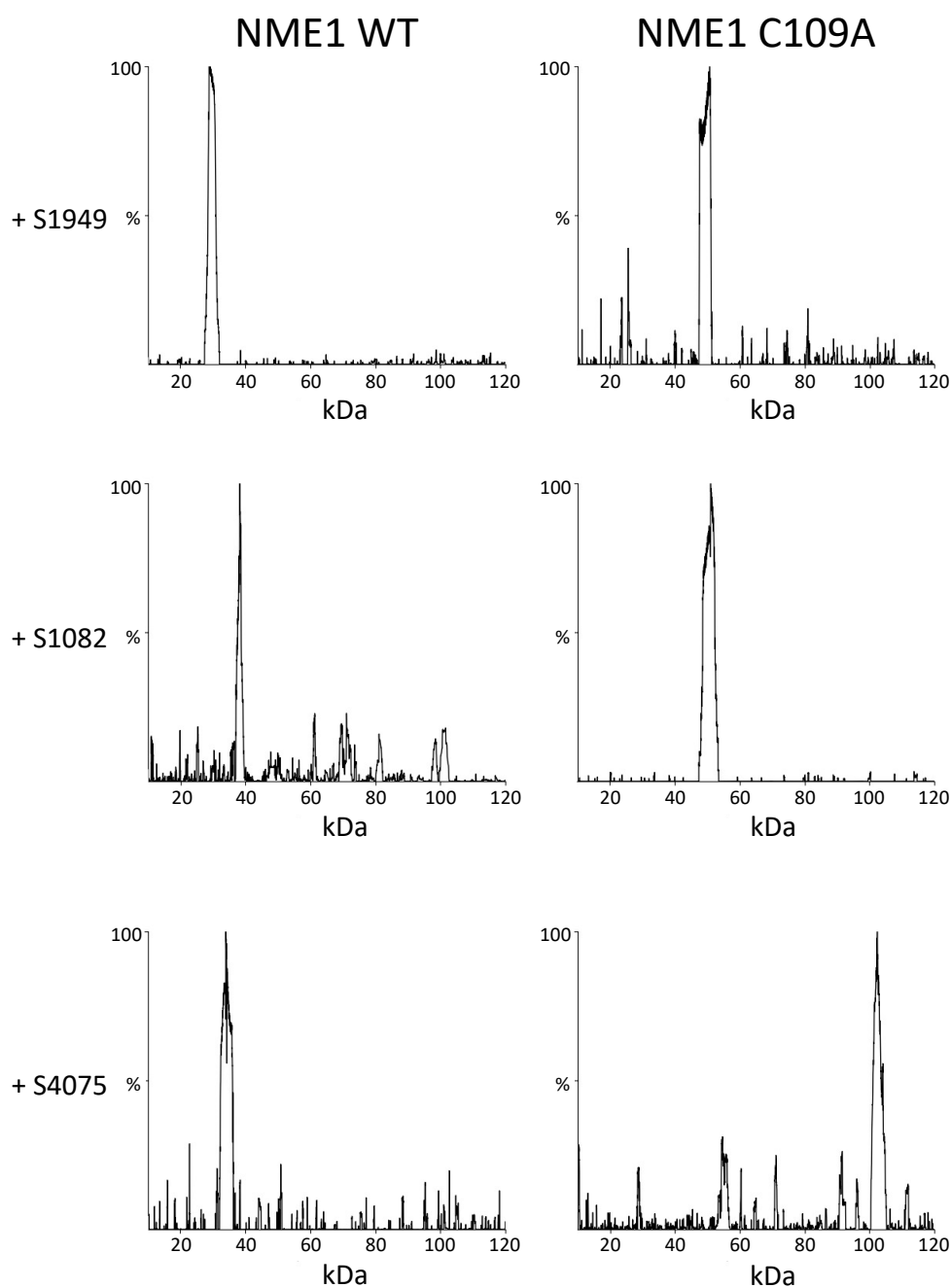
Figure 5.31. Oligomeric state of NME1 proteins with compounds. Oligomeric determination for 10x excess compound with NME1 WT or C109A. Spectrum lists were copied from MassLynx 4.1 and inputted into UniDec to carry out the deconvolution analysis. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

From the full set of compounds analysed (twenty), thirteen induced NME1 dissociation from hexamer to dimer upon addition to NME1 WT (Table 5.3/data shown in Appendix 5.1). Five of the compounds induced trimer formation, while one (S1198) promoted formation of a tetramer and another (S4073) did not induce hexamer dissociation. The addition of the compounds to NME1 C109A showed that only two compounds (S1198 and S4075) failed to induce hexameric dissociation. A high proportion of the compounds (nine) resulted in dissociation to a trimer when bound to NME1 C109A and six of the compounds resulted in the dimer form. The evidence that the majority of the compounds caused some form of dissociation of the hexamer form when added to NME1 WT suggests that the oligomeric state of NME1 could be destabilised under oxidative conditions due to the compounds containing carbonyl groups that may cause oxidative stress. The mutation of C109 was expected to cause some resistance to dissociation due to the cysteine no longer being available for dimerisation. However, full resistance (with exception to S1198 and S4075) was not observed with NME1 C109A, and only partial resistance of dissociation was observed, with many compounds dissociating to a trimer. Although C109 was mutated, two other cysteine residues, C4 and C145, were still present in this protein, which form an intramolecular disulphide bond (205). Therefore, oxidation could occur on these residues and inhibit the full resistance to dissociation of the hexamer form.

Compound	NME1 WT	NME1 C109A
S1082 (Vismodegib)	Dimer	Trimer
S1198 (Irinotecan)	Tetramer	Hexamer
S1344 (Glimepiride)	Trimer	Trimer
S1457 (Atazanavir sulfate)	Dimer	Dimer
S1512 (Tadalafil)	Dimer	Trimer
S1606 (Clotrimazole)	Dimer	Trimer
S1633 (Zafirlukast)	Dimer	Trimer
S1730 (Indapamide)	Dimer	Dimer
S1737 (Prednisolone)	Dimer	Tri/Hex
S1760 (Rifapentine)	Dimer	Dimer
S1808 (Nifedipine)	Trimer	Dimer
S1840 (Lomustine)	Trimer	Tetramer
S1847 (Clemastine fumarate)	Dimer	Dimer
S1859 (Diethylstilbestrol)	Dimer	Trimer
S1949 (Menadione)	Dimer	Trimer
S2522 (Adrenaline)	Dimer	Trimer
S4003 (Lithocholic acid)	Trimer	Trimer
S4073 (Aminosalicylate sodium)	Hexamer	Dimer
S4075 (Zinc pyrithione)	Dimer	Hexamer
S4221 (Benzbromarone)	Trimer	Trimer

■ Some resistance of dissociation by C109A

■ No dissociation with C109A

Table 5.3. Oligomeric state of NME1 proteins with compounds. Oligomeric determination for 10x excess compound with NME1 WT or C109A. Some resistance of dissociation from hexamer by C109A shown in blue and no dissociation shown in red. Data analysis carried out using MassLynx 4.1. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis. Data shown in Appendix 5.1.

5.2.5 Compound effects on the CCS conformational landscape of NME1

After the initial experiments to determine the oligomeric state of NME1 upon addition of the compounds, eight compounds were selected to analyse their effect on the CCS landscape of NME1. The selection of the eight were chosen based on their dissociation to the dimer form for NME1 WT and the resistance of dissociation by NME1 C109A. Addition of all the selected compounds to NME1 WT revealed a decrease in CCS values from 53.8 nm^2 (conformer III) to between 26.1 - 32.7 nm^2 (largest conformer for each compound) (Figure 5.32/34). This would be expected as the compounds had previously been shown to dissociate to the dimer form and would therefore have a smaller CCS as compared to the hexamer protein, as the protein complex is now adopting a much smaller conformation. NME1 WT + S1737 observed an additional conformer not seen in any of the other conditions, which was termed as a * symbol due to it having a much smaller CCS value (22.3 nm^2) in comparison to the smallest conformer I in other conditions. Without molecular modelling data on the structural configurations adopted by each of the conformational states, it is difficult to speculate what is happening structurally to NME1 WT when bound to the different compounds.

NME1 WT

	CCS (nm²)	CCSD (nm²)
I	N/A	N/A
II	52.0	2.5
III	53.9	1.7
IV	N/A	N/A

$R^2 = 0.997$

NME1 WT + S1082

	CCS (nm²)	CCSD (nm²)
I*	23.3	1.8
II*	25.3	2.0
III*	28.2	2.5
IV*	N/A	N/A

$R^2 = 0.992$

NME1 WT + S1512

	CCS (nm²)	CCSD (nm²)
I*	N/A	N/A
II*	25.1	2.2
III*	27.5	2.5
IV*	N/A	N/A

$R^2 = 0.988$

NME1 WT + S1606

	CCS (nm²)	CCSD (nm²)
I*	N/A	N/A
II*	23.9	2.4
III*	26.0	2.6
IV*	N/A	N/A

$R^2 = 0.996$

NME1 WT + S1633

	CCS (nm²)	CCSD (nm²)
I*	24.1	1.8
II*	25.6	1.7
III*	27.9	2.2
IV*	N/A	N/A

$R^2 = 0.991$

NME1 WT + S1737

	CCS (nm²)	CCSD (nm²)
*	22.3	0.6
I*	N/A	N/A
II*	26.4	2.0
III*	28.6	3.3
IV*	32.4	2.8

$R^2 = 0.991$

NME1 WT + S1949

	CCS (nm²)	CCSD (nm²)
I*	24.5	1.8
II*	26.0	1.4
III*	27.4	1.3
IV*	N/A	N/A

$R^2 = 0.996$

NME1 WT + S2522

	CCS (nm²)	CCSD (nm²)
I	N/A	N/A
II	26.0	2.7
III	28.7	2.4
IV	32.0	2.2

$R^2 = 0.995$

NME1 WT + S4075

	CCS (nm²)	CCSD (nm²)
I*	23.5	1.1
II*	25.1	2.2
III*	27.7	2.3
IV*	N/A	N/A

$R^2 = 0.996$

Figure 5.32. CCS distribution for NME1 WT with compounds. $^{TW}CCS_{N2\rightarrow He}$ Gaussian fitting for the [M+21H] $^{21+}$ species of NME1 WT. Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed using the Fit Peaks Pro function in Origin (Version 2021b), with R^2 values listed. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

Addition of all the compounds to NME1 C109A (with exception to S4075) revealed a decrease in CCS values from 53.5 nm^2 (conformer III) to between 28.0 – 33.9 nm^2 (largest conformer for each compound) (Figure 5.33/34). As the compounds had previously been shown to cause some degree of dissociation from the hexamer form, the smaller CCS would be expected to match the new oligomeric state of the protein in the presence of each compound. The S4075 compound, which caused full resistance of dissociation to the hexamer form in NME1 C109A revealed only the presence of conformer II, with a CCS of 59.2 nm^2, which was larger than the CCS values observed for NME1 C109A alone (I = 51.8 nm^2, II = 53.5 nm^2).

When comparing the relative abundance between conformers for NME1 WT/C109A + compounds, the most prominent difference between the two is that the compounds induce NME1 C109A into one main conformational state (III*/IV*), with exception of S1949 and S2522 (Figure 5.34). NME1 WT when bound to each compound results in the presence of an additional conformer (I* or IV*), and also a decrease in the most abundant conformer II* (S4075 exception) from 70% to 57%. The compounds bound to NME1 WT induced a range of relative conformer abundance difference between compounds, which could be simply due to the array of compounds tested and the different mechanisms in which they induce on NME1. The larger differences associated with the conformational landscape between NME1 WT and C109A when bound to different compounds suggests that the compounds have a greater effect on NME1 WT in comparison to NME1 C109A. This is expected as the compounds induce a further level of dissociation on NME1 WT in contrast to some resistance shown with NME1 C109A.
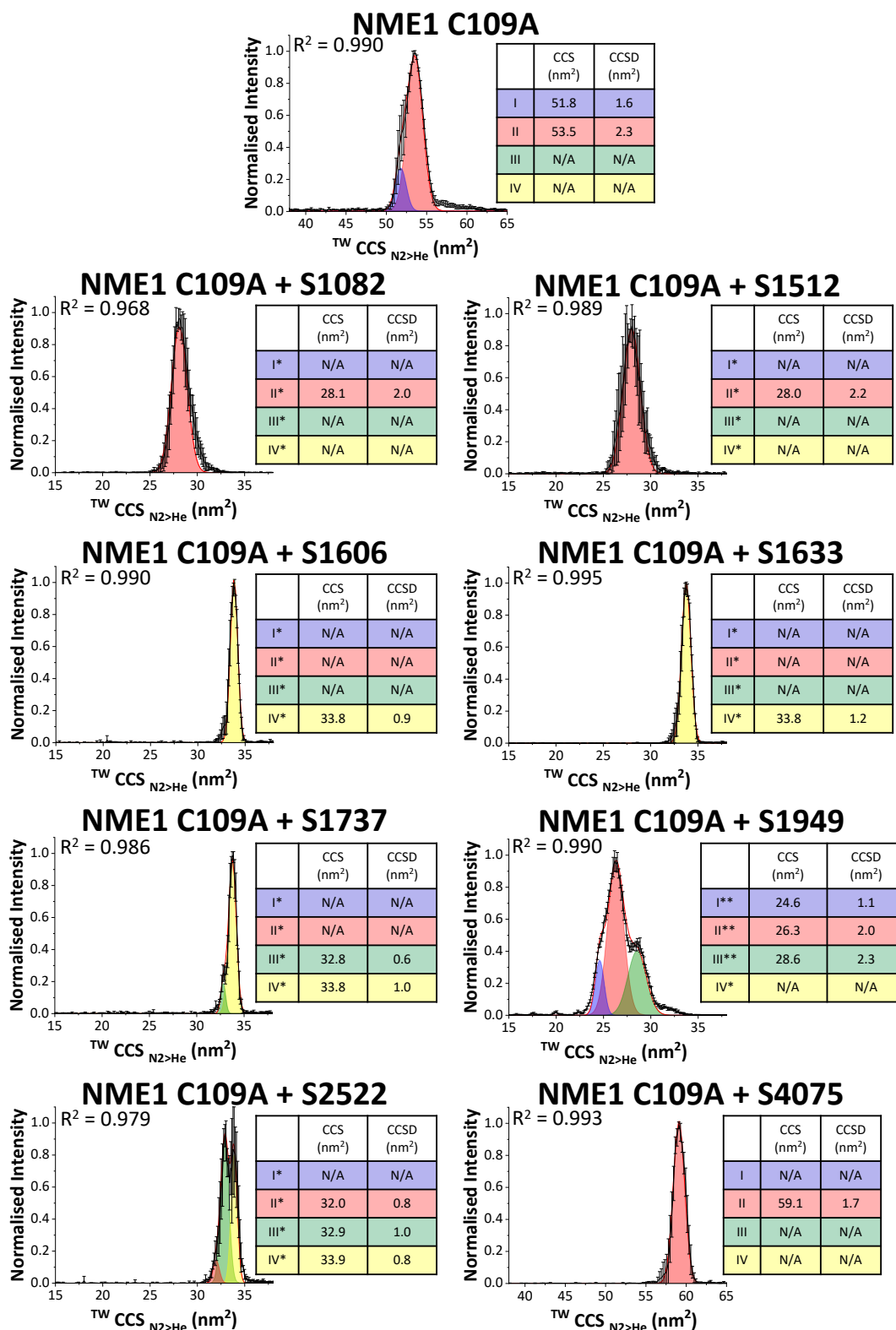
Figure 5.33. CCS distribution for NME1 C109A with compounds. $^{TW}CCS_{N2\rightarrow He}$ Gaussian fitting for the $[M+21H]^{21+}$ species of NME1 C109A. Red line is the average of three independent replicates. Black error bars represent the S.D. Gaussian fitting was performed using the Fit Peaks Pro function in Origin (Version 2021b), with R^2 values listed. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.
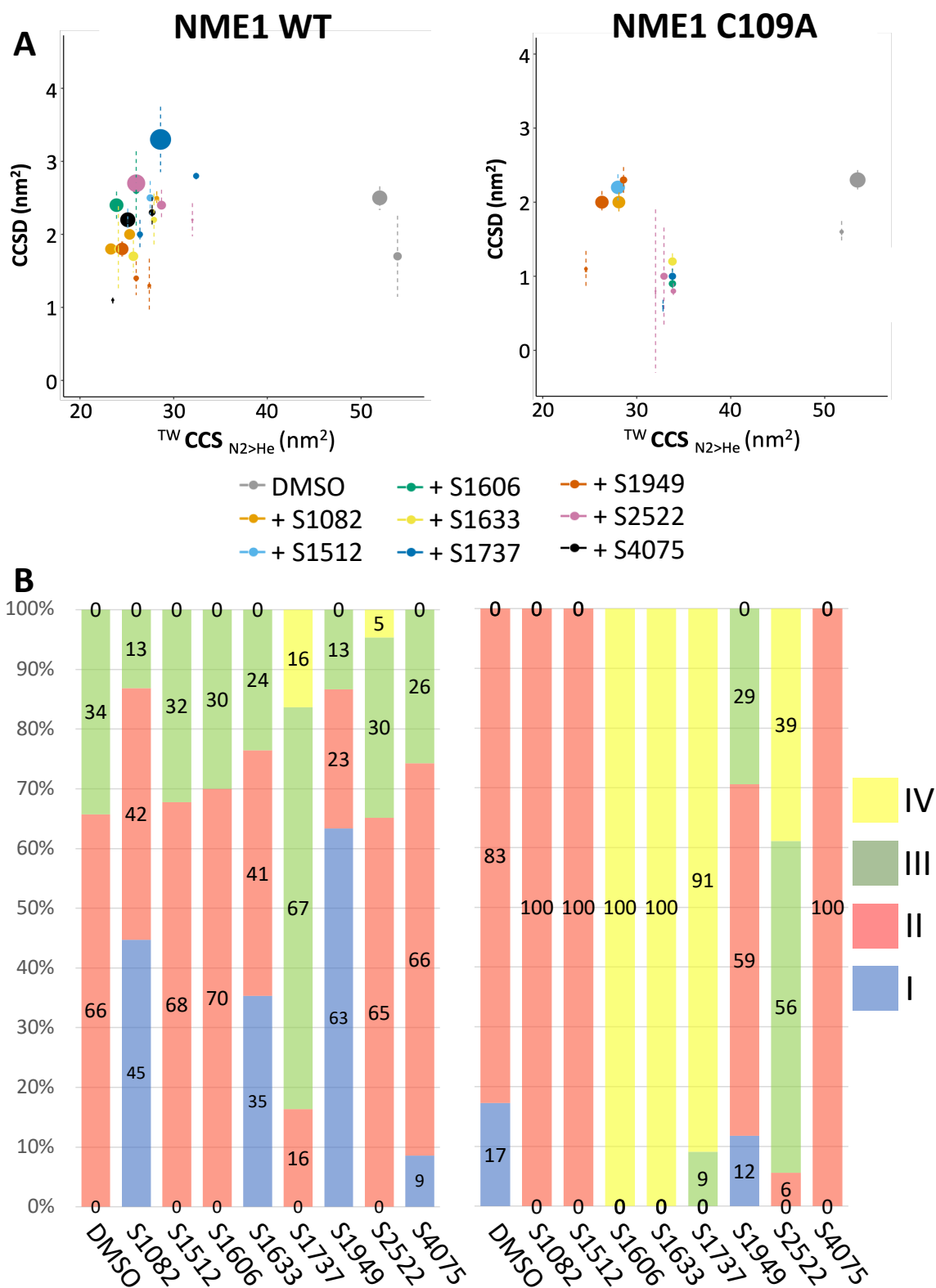
Figure 5.34. Proportional conformational space adopted by the NME1 WT/C109A + compounds. Determined by Gaussian fitting in Figure 5.32/33 for three individual experiments. **A)** % area of different conformational states (I (blue), II (red), III (green), IV (yellow)) for NME1 WT (left) and NME1 C109A (right). **B)** Proportional scatter plots (CCS (nm²) versus CCSD (nm²)) for the different conformational states for NME1 WT (left) and NME1 C109A (right). Size of dot representative of area.

5.2.6 Effect of compounds on the phosphorylation of NME1

Following the observation that the compounds induced dissociation for both NME1 WT and C109A, the next aspect of the analysis was to determine whether the addition of compounds prior to ATP had any effect on phosphorylation. This was carried out as before, using an intact LC/MS experiment to assess the abundance of phosphorylation for NME1 WT and C109A when bound to compounds. Prior to the addition of compounds and ATP, NME1 WT was desalted using the same method for IM-MS experiments to determine whether desalting NME1 into ammonium acetate had any impact on phosphorylation. This additional step prior to the incubation with ATP did not reduce the level of phosphorylation (85%) compared to 75% without the desalting step (Figure 5.35).
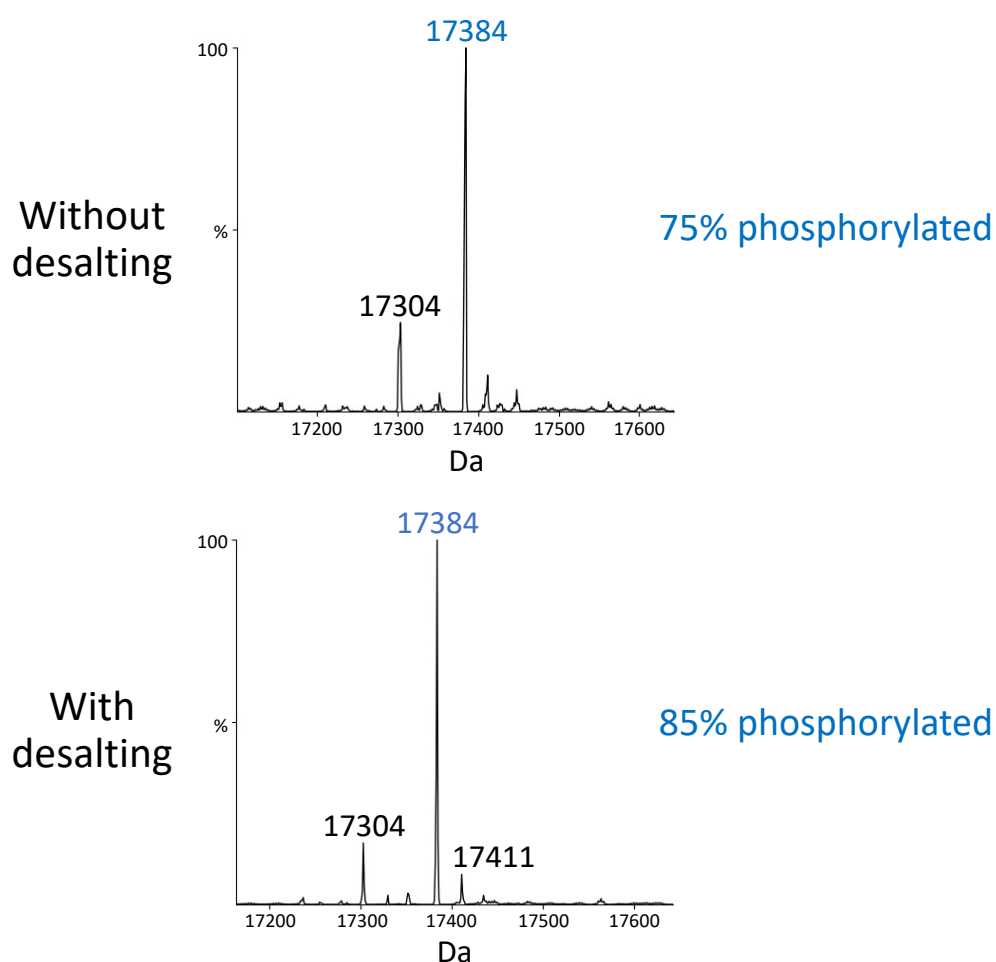
Figure 5.35. Desalting NME1 into ammonium acetate does not reduce level of phosphorylation. NME1 WT was buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to setting up phosphorylation reaction. 1 µg of NME1 WT was incubated with 10 ATP µM in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 5 minutes at R.T. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1.

To establish the optimal incubation time for the addition of compounds, S1949 was incubated at R.T. for 10, 30 and 60 minutes with NME1 WT. The relative abundance of the smaller peaks (ranging from 17560 to 17850 Da) decreased with the increase in time for the 30 and 60 minute incubations (Figure 5.36). The peak intensity of NME1 WT without compound at 17303 Da reduced with the two longer time points and the mass accounting for NME1 WT + S1949 (17474 Da) also had reduced intensity. The 10-minute incubation time was selected for all future experiments with compounds due to the longer time points not gaining any additional information and reducing the intensity of the two important masses.
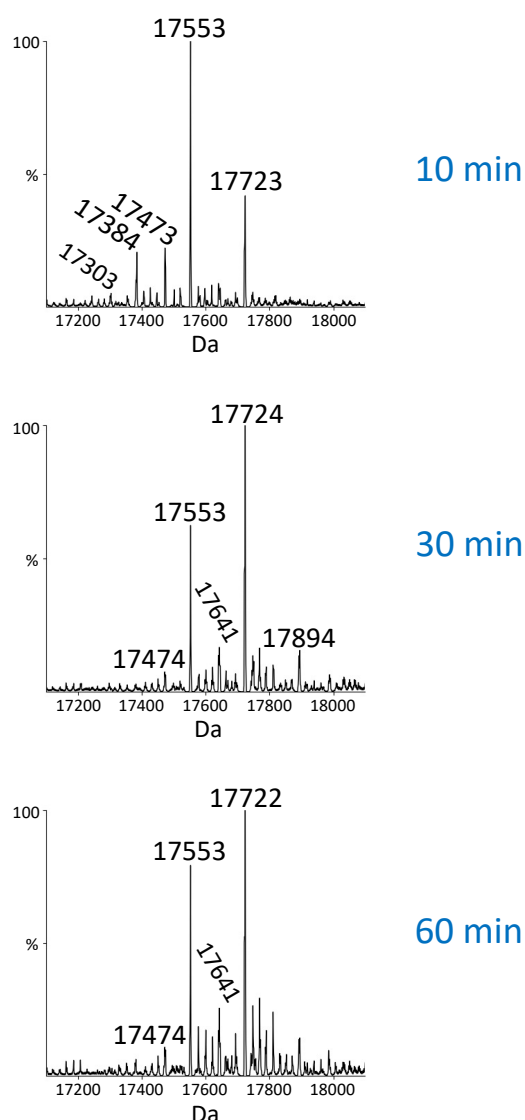
Figure 5.36. NME1 WT + S1949 compound has no major additional effects with longer incubation. 1 µg of NME1 WT was incubated with 200 µM S1949 compound in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 10, 30 and 60 minutes at R.T. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1.

The previous experiments associated with determining phosphorylation mass shifts due to the addition of nucleotides were at a concentration of 1 mM nucleotide. As the compounds tested had a variety of unique structures, it could not be determined whether they are ATP-competitive. However, if the compounds were competing with the ATP binding site, the binding potential of the compounds to NME1 WT would be greater in the presence of minimal ATP due to NME1 having such a high affinity for the binding of ATP. To overcome this possibility, the concentration of ATP was tested at decreased amounts in order to reduce the competition between ATP and compound. The reduced ATP concentrations of 50 µM and 10 µM decreased the level of phosphorylation to 92% and 75% respectively compared to 93% for 1 mM ATP (Figure 5.37). Although the level of phosphorylation decreased, there was still a higher proportion of phosphorylated that non-phosphorylated NME1. Therefore, the 10 µM ATP concentration was chosen for further experiments to provide the least competition to the binding of the compounds.
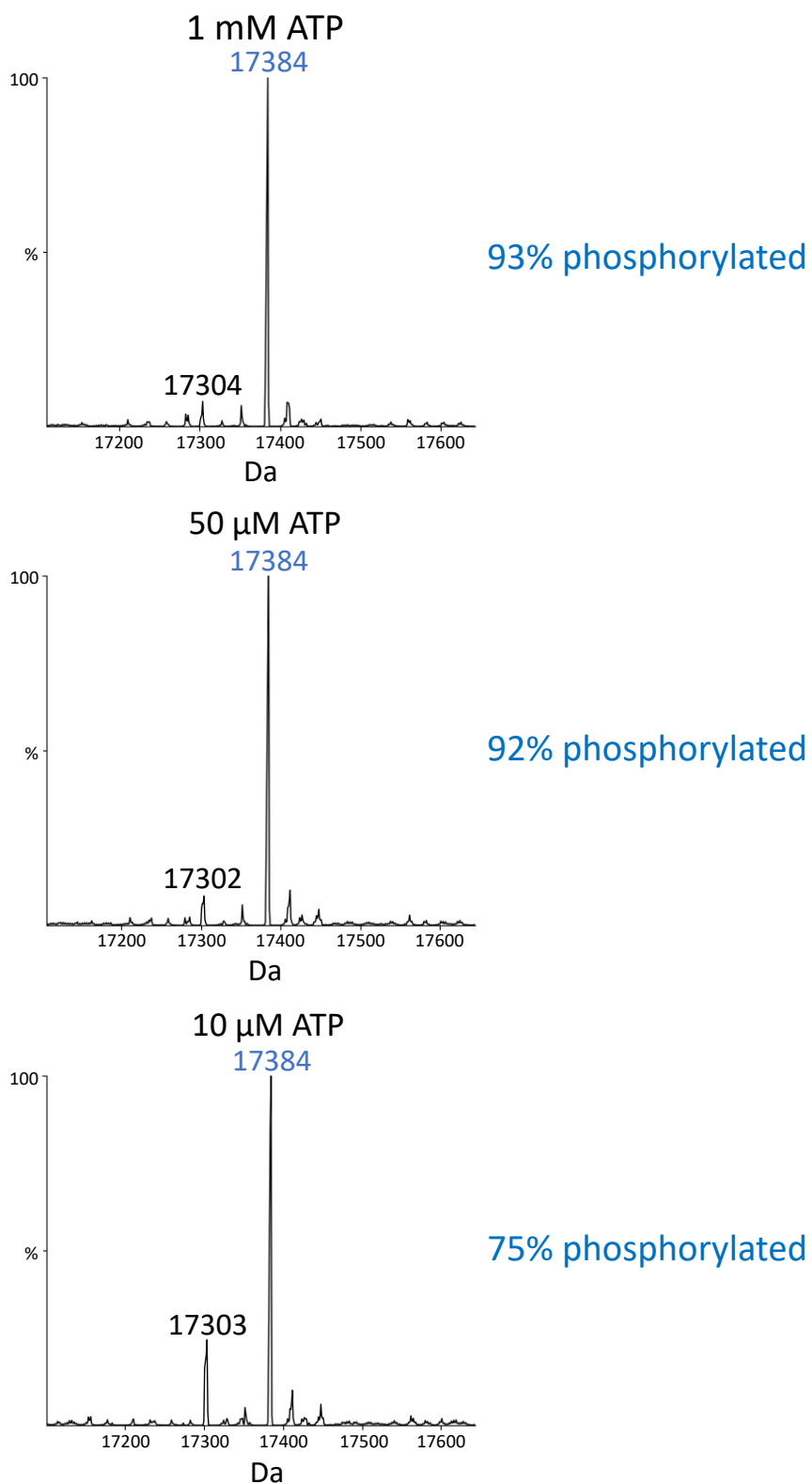
Figure 5.37. NME1 WT is phosphorylated by lower concentrations of ATP. 1 μg of NME1 WT was incubated with ATP in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 5 minutes at R.T. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1.

Based on the FDA library screen and dissociation of the hexameric state for NME1 WT when bound with compounds, eight were selected to analyse for the effects on NME1 activity, and thus auto-phosphorylation. The addition of compounds S1082, S1512, S1606 and S1633 had little impact on the relative abundance of phosphorylated NME1 WT (Figure 5.38) when compared to ATP alone (Figure 5.37), with ~75% appearing in its phosphorylated form. The addition of these four compounds all failed to observe a mass that corresponded with the theoretical mass of NME1 WT + compound. The compounds ranged from 345 to 576 Da, therefore, a peak would be expected in the region of 17728 - 17959 Da for the combined mass of NME1 WT, selected compound, and phosphorylation. The lack of peak observed in this range for any of these compounds suggests that the compounds do not covalently bind to NME1 WT. Another potential explanation for the lack of mass observed with bound compound could be due to the compounds causing dissociation leading to the presence of smaller fragment peaks, in particular with the addition of S1606. One of the mass shifts observed for the S1606 compound between the 17413 and 17435 Da peaks, had a mass of 22 Da, which would be commonly associated with a sodium adduct obtained from the sample preparation methods.
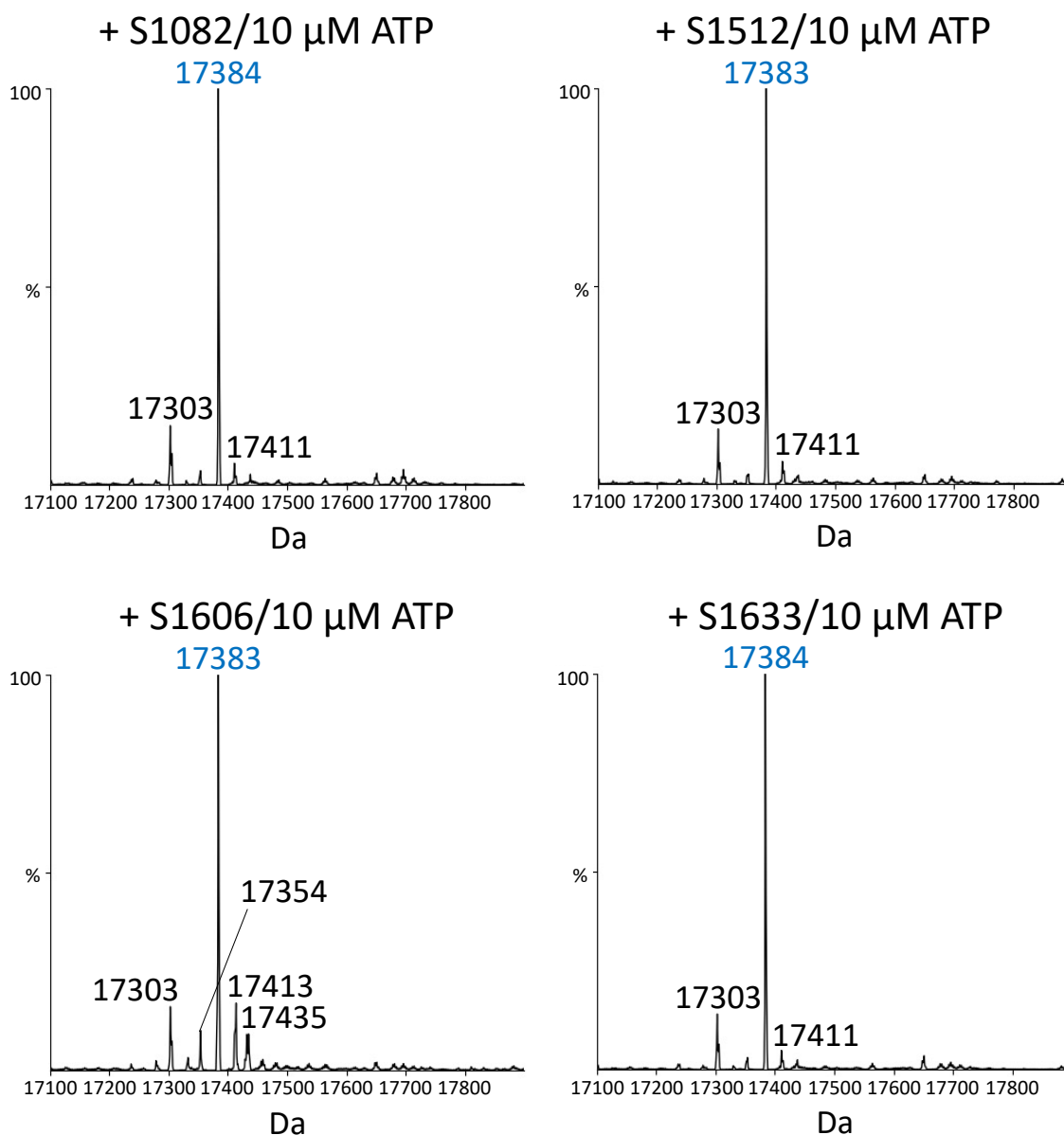
Figure 5.38. Compounds do not alter phosphorylation of NME1 WT. 1 µg of NME1 WT was incubated with 200 µM of each compound in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 10 minutes at R.T. 10 µM ATP was added to the sample mixture and then immediately analysed by LC-MS under non-native conditions. The mass corresponding to phosphorylated (unbound) NME1 is labelled in blue. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1.

The spectra obtained for NME1 following incubation with either S1737 or S4075 (Figure 5.39) revealed peaks corresponding to phosphorylated (17382 Da) and non-phosphorylated unbound NME1 (17303 Da), with S1737 incubation resulting in significant peaks at 17413 Da and 17429 Da that could be representative of compound mass adducts, such as a methyl group for the peak at 17429 Da. Interestingly, analysis of NME1 incubated with either S1949 (172 Da) and S2522 (183 Da) revealed masses consistent with phosphorylated compound bound (17552 Da and 17560 Da respectively) as well as unbound NME1. The difference in mass (Δ4 amu) between that observed for these two compounds and as would be expected for non-covalent binding, suggests that both bind through a covalent mechanism. The mass of 17474 Da for S1949 compound is likely representative of non-covalently bound compound that is not phosphorylated. The mass of 17725 Da could account for a non-covalently bound S1949 molecule in addition to the mass of 17552 Da, therefore, two molecules of compound bound to phosphorylated NME1, one covalent and one non-covalent.
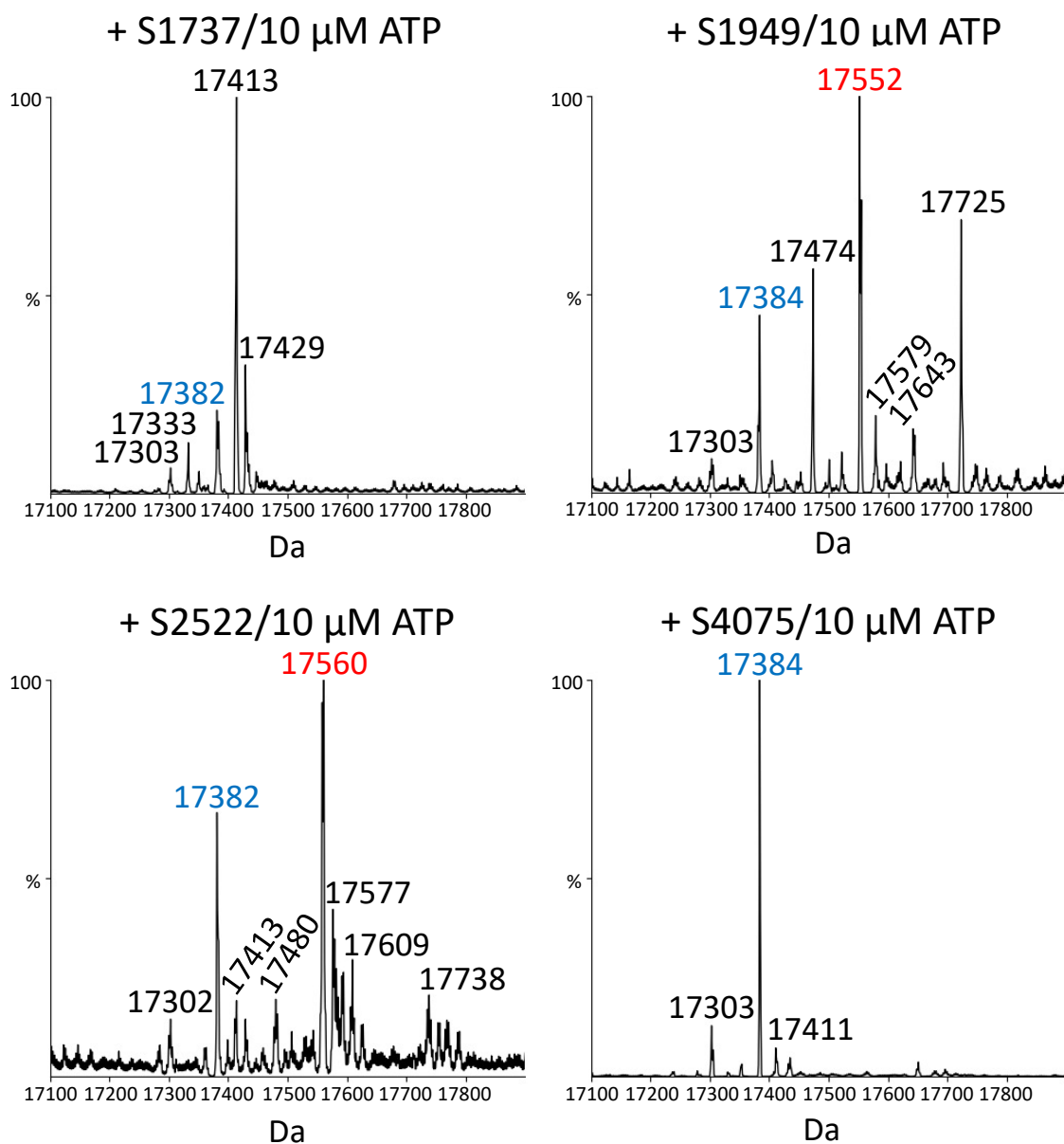
Figure 5.39. Compounds do not alter phosphorylation of NME1 WT. 1 µg of NME1 WT was incubated with 200 µM of each compound in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 10 minutes at R.T. 10 µM ATP was added to the sample mixture and then immediately analysed. Unbound compound phosphorylated mass shown in blue and bound compound phosphorylated mass shown in red. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1.

The binding of the compounds S1082, S1512, S1606 and S1633 to NME1 C109A had no evidence to suggest that any of these compounds covalently bound to the protein, with no peaks observed at the expected mass for compound complex (Figure 5.40). The addition of the compounds also had little effect on the mass associated with the phosphorylated form at ~ 17352 Da. The smaller peak observed at 17273 Da is representative of non-phosphorylated NME1 C109A.
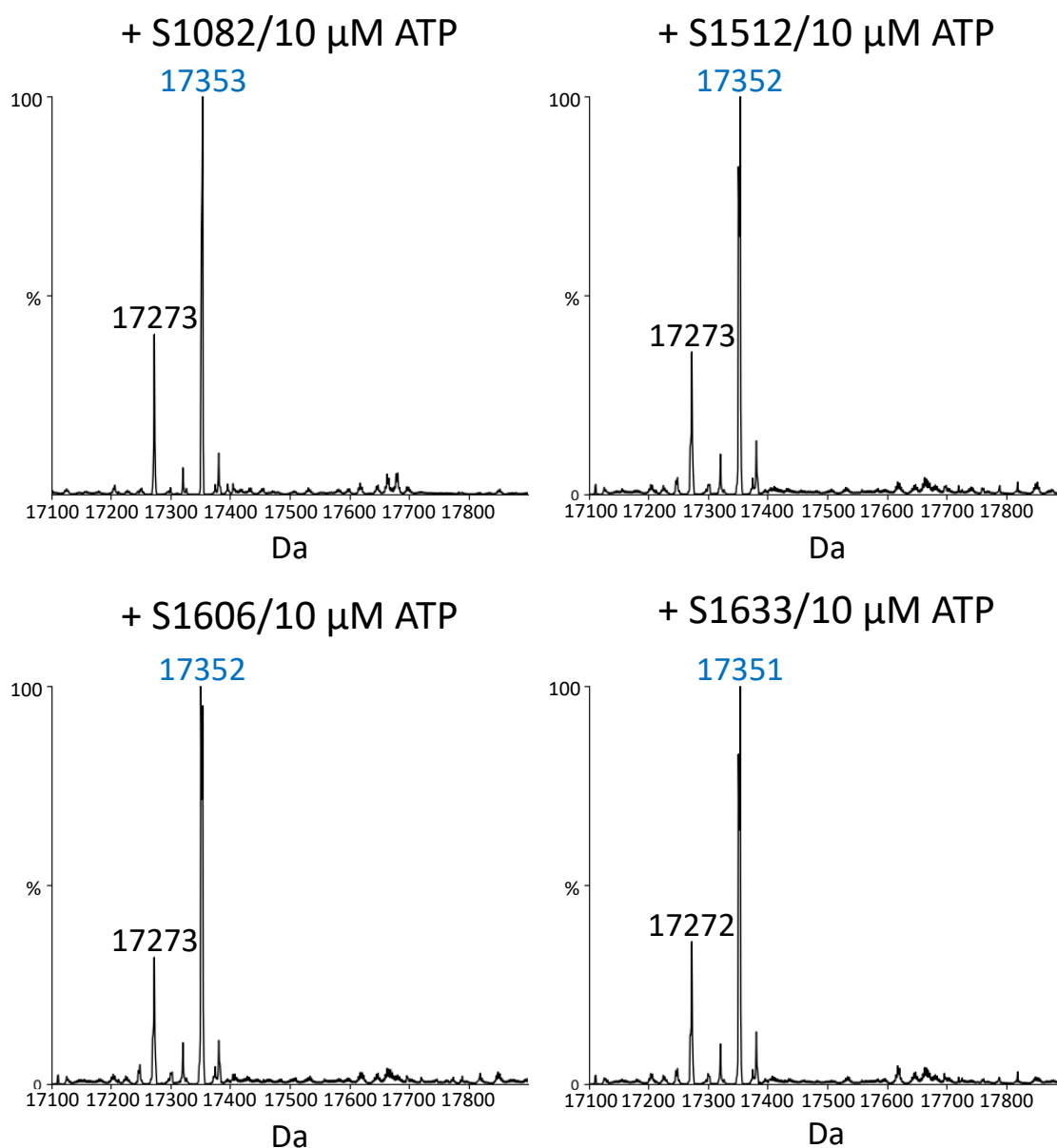
Figure 5.40. Compounds do not alter phosphorylation of NME1 C109A. 1 µg of NME1 C109A was incubated with 200 µM of each compound in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 10 minutes at R.T. 10 µM ATP was added to the sample mixture and then immediately analysed. Unbound compound phosphorylated mass shown in blue. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1.

The addition of the compounds S1737 and S4075 to NME1 C109A had similar spectra profiles when compared to the same compounds for NME1 WT, with no peaks observed that matched the masses for NME1 C109A + compound (+/- the phosphorylated form). Similar to the data for NME1 WT, analysis of NME1 C109A following incubation with S1949 or S2522 (Figure 5.41) revealed masses associated with non-phosphorylated (17351 Da and 17351 Da respectively) and phosphorylated NME1 C109A (17522 Da and 17528 Da respectively). The mass of 17546 Da for S2522 compound could represent the mass of an OH adduct from the compound. The mass of 17705 Da could be representative of a further S2522 molecule in addition to the mass of phosphorylated NME1 C109A with covalently bound S2522.
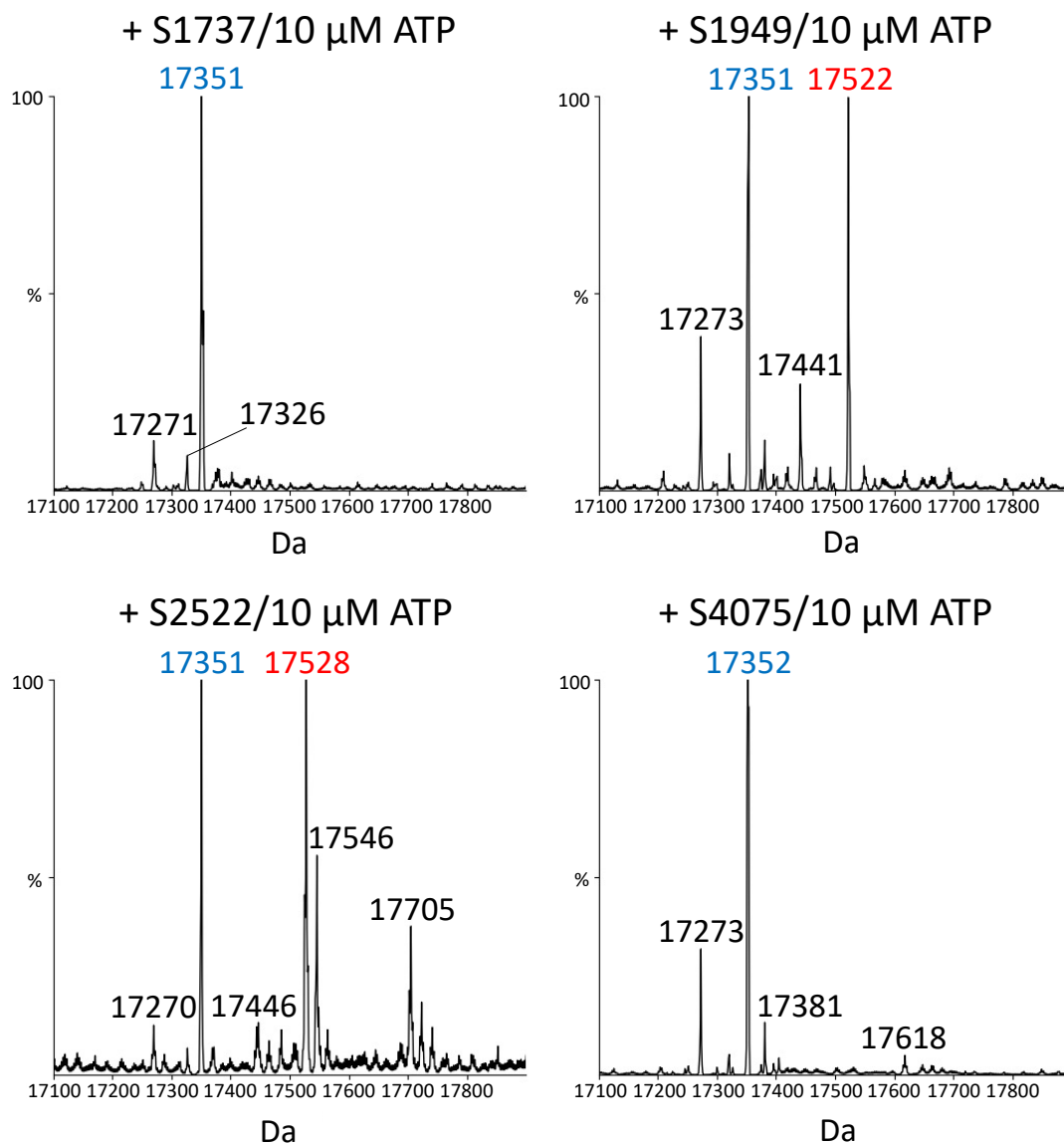
Figure 5.41. Compounds do not alter phosphorylation of NME1 C109A. 1 µg of NME1 C109A was incubated with 200 µM of each compound in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 10 minutes at R.T. 10 µM ATP was added to the sample mixture and then immediately analysed. Unbound compound phosphorylated mass shown in blue and bound compound phosphorylated mass shown in red. Data was processed using MassLynx (4.1) and deconvoluted using MaxEnt1.

As previously mentioned, the binding of the compounds induced an oligomeric switch for NME1 WT from hexamer to dimer, which could result in changes to the enzymatic activity of this protein. The dimer form of this protein is expected to not have the same functional activity as the hexamer form, and therefore, not be as effective in binding ATP. Despite this, the addition of the compounds prior to the addition of ATP did not have a large impact on the level of phosphorylation for either NME1 WT or C109A (Figure 5.42). This suggests that the

dimer form of NME1 when bound to compounds appears to be as active as the hexamer form in the absence of compounds.
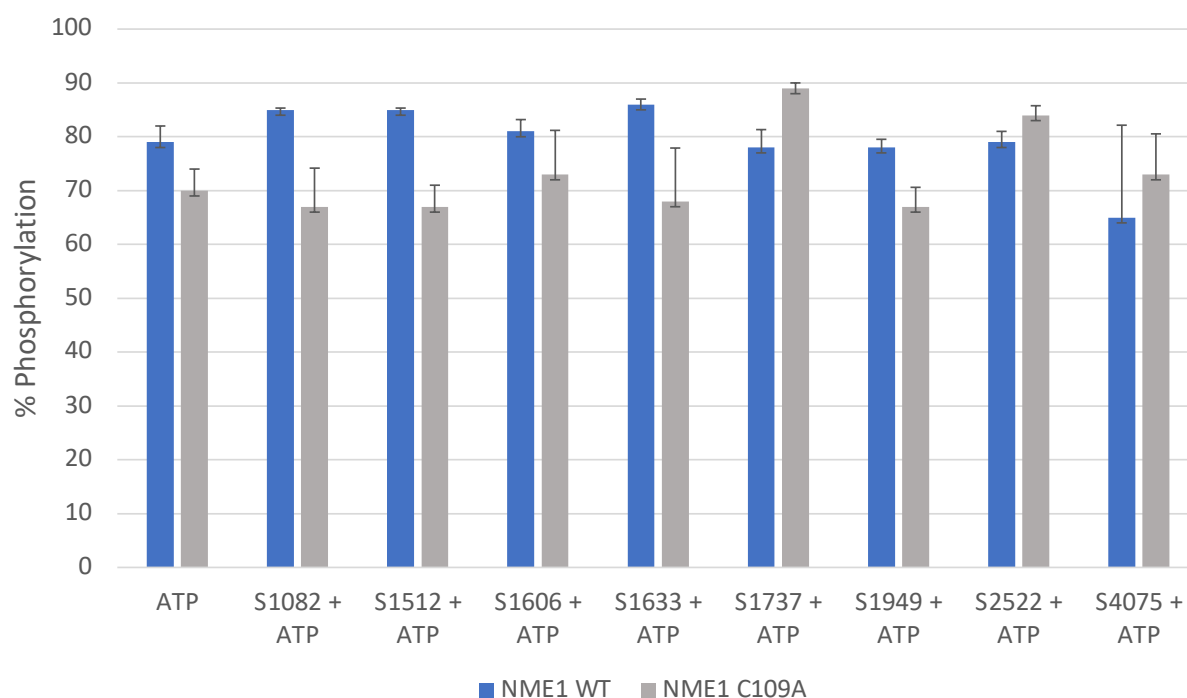
Figure 5.42. Addition of compound does not impact level of phosphorylation for NME1 WT and C109A. A total of 1 µg of NME1 WT or C109A was incubated with 200 µM of each compound in 50 mM Tris-HCl, pH 8.0, 100 mM NaCl for 10 minutes at R.T. 10 µM ATP was added to the sample mixture and then immediately analysed. Stoichiometry of histidine phosphorylation was determined by comparing the relative heights of the phosphorylated and non-phosphorylated protein within the sample mixture containing the phosphate donor. Data presented as average of three independent replicates with error expressed as standard deviation.

5.3   Conclusion

The first aim of this chapter was to determine whether the addition of nucleotides to NME1 would induce a mass shift associated with auto-phosphorylation, thereby establishing an activity assay. Addition of nucleotides to the NME1 WT induced a shift in mass indicating phosphorylation, which was ablated upon mutation of H118 to Ala. This finding correlated with previous studies suggesting that H118 in NME1 is phosphorylated (213). No autophosphorylation was observed for NME1 R105A, suggesting that the mutation of this Arg residue that sits within the kpn loop of NME1 supports previous published data that shows this residue affects the ability of this protein to bind a nucleotide and act as a phosphate donor to induce autophosphorylation. The kpn loop within NME1 is situated within the active

site and the R105 residue forms interactions that neutralise the negatively charged phosphate group and enables phosphorylation to occur (203). It is clear that the R105A mutation decreases the enzymatic activity of this protein, which may be a result of reduced nucleotide binding and/or dissociation of the hexamer to a dimer.

The addition of the nucleotides to NME1 WT caused a large variation in terms of the relative abundance of conformers that the protein adopted in the presence of each nucleotide. The variations observed in protein conformation by the addition of the different nucleotides could be a result of the ping-pong mechanism that occurs when the high-energy phosphohistidine intermediate is formed (96). Therefore, the reaction could have occurred too quickly prior to the IM-MS experiment being carried out, which potentially could have resulted in the acquisition taking place when phosphorylation was not at its most abundant, when the complex is switching between its intermediate and standard state. In addition to the large variation observed for each nucleotide, the CCSD values between conditions had little alteration between nucleotides and DMSO control. This would suggest that upon addition of nucleotide, the protein doesn't induce any enhanced conformational flexibility. Therefore, this IM-MS experiment might not be suitable to distinguish any sufficient differences in protein conformation upon addition of a range of nucleotides due to the mechanism in which they bind effectively.

This chapter confirmed by IM-MS that NME1 WT natively forms as a hexamer, which is required for full enzymatic activity (203). Upon addition of a range of compounds selected from an FDA-drug approved library screen of small molecules, dissociation of this hexameric form to a dimer was observed by IM-MS. Menadione (S1949), a known oxidising agent (212), resulted in the greatest negative shift in Tm when analysed by DSF. This suggested that the compounds inducing a negative shift could be destabilising the hexamer structure under oxidative conditions, which has been confirmed in previous studies (205). Under oxidative *in vitro* conditions, the cysteines in NME1 are modified due to their high nucleophilic properties, resulting in the formation of disulphide bonds that can alter the protein complex and decrease activity. The change in conformation results in dissociation of the natively formed hexamer into a dimer form. The conformation change has been shown to alter the C-terminal domain by the disruption of two hydrogen bonds between the kpn loop and C-terminal

domain (205). The mutated C109A variant was analysed with the selected compounds to determine whether the mutation of the cysteine would result in resistance to dissociation of the hexamer to dimer. The findings in this chapter did not suggest that the NME1 C109A protein resulted in full resistance to dissociation when incubated with the selected compounds. However, partial resistance to dissociation was observed with the majority of the compounds causing a dissociation to the trimer form. This could be due to the remaining cysteines, C4 and C145, forming a disulphide bond that undergoes oxidation and results in some dissociation being observed. In order to examine this hypothesis, additional Cys to Ala NME1 mutated variants would need to be analysed alongside the C109A mutated form.

To determine whether the addition of the compounds could alter the activity of either WT or C109A NME1, auto-phosphorylation of these two proteins was quantified upon incubation with the compound and ATP. As the compounds induced dissociation from a hexamer to dimer for NME1 WT, the dimer form was expected to have reduced activity and thus auto-phosphorylation. NME1 in a dimer form should have reduced enzymatic activity and also not be able to effectivity bind a nucleotide to induce phosphorylation (203). The results in this chapter showed that the binding of a compound prior to the addition of ATP resulted in little change in the level of phosphorylation when compared to ATP only. Previous data shown in Section 5.2.6 demonstrated that not all the compounds analysed bound covalently, therefore, an increase in mass was not observed and the binding efficiency could not be fully determined by IM-MS or intact MS. To gain a further understanding into the effects these compounds have on NME1, additional experiments such as computational modelling could be done to support the analysis by DSF. In addition, further MS experiments with other compounds analysed from the initial FDA library screen could be carried out to confirm whether NME1 is dissociated under oxidative conditions and whether the dimer form does have reduced enzymatic activity in the presence of other compounds.

Chapter 6. Final Conclusions & Future Perspectives

The work carried out in this thesis aimed to explore the use of IM-MS as a technique to evaluate the effects of inhibitor binding and phosphorylation on the conformational landscape and stability of two different protein kinases, the serine/threonine protein kinase Aurora A, and the histidine kinase NME1 (Nucleoside diphosphate kinase A). I also used native MS to evaluate changes in the oligomerisation state of NME1 induced by the binding of small molecule compounds.

Chapter 3 described the purification of the two forms of Aurora A (122-403) (that were then analysed in Chapters 3 and 4), a WT active form, and an inactive D274N variant that lacked the ability to co-ordinate ATP/MgCl$_2$. The truncated version of this protein with only the residues 122-403 had previously been shown to be sufficient at binding the Aurora A binding partner TPX2 (35), confirming its suitability for these structural studies. Results in Chapter 3 showed the WT Aurora A autophosphorylates, with six sites of phosphorylation determined by LC/MS/MS, including the Thr288 residue, which is situated within the activation loop, and is required for Aurora A catalytic activity (214). D274N Aurora A showed a negligible amount of phosphorylation at the Thr288 residue when analysed by a phosphospecific antibody. Additionally, no phosphorylation sites were observed by LC/MS/MS for D274N Aurora A. The generation of the D274N Aurora A variant was important for the IM-MS analysis in order to assess any potential differences between phosphorylated (active) and non-phosphorylated (inactive) protein either alone or when bound to small molecule inhibitors or the variants of the activating TPX2 minimal peptide.

Prior to the analysis of Aurora A by IM-MS, Chapter 3 assessed different solutions for desalting the protein into a volatile buffer suitable for MS analysis. In agreement with many other IM-MS studies, ammonium acetate (150 mM) was the preferred solution and was used for the IM-MS work carried out in this thesis (182). IM-MS analysis revealed both WT and D274N Aurora A to form monomers when sprayed under native conditions, which was consistent with published structures available for the Aurora A (122-403) proteins (215). Following CCS determination of WT and D274N Aurora A, a broad CCS distribution was observed, which

prompted the need to Gaussian fit the CCS landscape to enable determination of the different overlapping conformers that both proteins were adopting. Gaussian fitting revealed three and four conformational states for WT and D274N Aurora A respectively, with D274N demonstrating greater conformational flexibility in comparison to WT Aurora A due to D274N's broader CCS landscape and smaller CCS average weighted value. The experimentally derived average CCS values were within the 10% range expected when compared to theoretical values calculated using IMPACT software based on comparison of phosphorylated and non-phosphorylated structures of Aurora A found in the PDB. Finally, Chapter 3 revealed that D274N Aurora A was more stable than WT when analysed by both DSF and CIU techniques. The findings in Chapter 3 provided a good understanding of the two Aurora A proteins and the differences between phosphorylated and non-phosphorylated forms, which was fundamental prior to the analysis of inhibitor bound complexes in Chapter 4.

Chapter 4 aimed to explore any differences in the conformational landscape adopted by WT and D274N Aurora A in the presence of Aurora A specific inhibitors with different binding mechanisms, as well as the generic proteins kinase inhibitor staurosporine. The inhibitors selected for analysis were chosen based on the different DFG modes that each inhibitor preferred (as previously reported (57)), such as the DFG-in inhibitor, ENMD-2076, and DFG-out inhibitor, MK-8745. Previous studies in the literature using a Förster resonance energy transfer (FRET) sensor had demonstrated structural movements of the activation loop of Aurora A when bound to a range of different DFG mode inhibitors (57). Therefore, IM-MS was carried out to explore the utility of this rapid structural technique for differentiation modes of small molecule inhibition based on enforcing different conformational landscapes on Aurora A. Chapter 4 demonstrated that all inhibitor compounds bound to both WT and D274N Aurora A, evident by the increased m/z region that the peaks adopted. Staurosporine, the most generic inhibitor tested in this chapter (216), induced the largest conformational change on both forms of Aurora A, constraining to only two conformational states (II and III). However, the analysis of the extreme DFG-in and DFG-out inhibitors was difficult to determine the associated differences upon binding of these two inhibitors on the structure of Aurora A.

Using IM-MS as a technique to differentiate the types of inhibitors based on their DFG preference may not have been possible potentially due to the low resolution of the instrument or the protein flexibility being too high to detect in this manner. An MS instrument with higher resolution, such as the Waters Cyclic IMS could highlight further differences between the effects of different inhibitors on the structural landscape of Aurora A that could not be observed using the Waters Synapt G2Si. In addition to the enhanced resolution of the Cyclic IMS, additional experiments could be carried out that enable ions to be selected and reinjected with collision-induced dissociation to gain IMS^n analysis, which would provide a much greater understanding of the Aurora A inhibitor bound complexes (141).

Although, IM-MS was unable to determine the DFG preference of each inhibitor in Chapter 4, subtle differences were observed for the inhibitor-bound forms of phosphorylated and non-phosphorylated Aurora A. The overall conformational landscape of Aurora A changed when bound to an inhibitor, with Aurora A showing a higher relative abundance for conformer II when bound to an inhibitor. Molecular modelling was carried out in Chapter 4 to gain a further understanding of the different conformational states that Aurora A adopted in the absence and presence of inhibitors. Modelling suggested that conformer II was representative of a more 'closed' structure, where the A-loop is facing inwards, whereas conformer III was representative of an 'open' structure, with the A-loop extended out. The increase in conformer II observed for the inhibitor bound Aurora A complexes therefore suggested that the inhibitors constrained the protein and locked it into a more 'closed' configuration. Using the differences in the relative ratios of conformer II, it may be feasible to further exploit this as an assay to screen for potential small molecule inhibitors, with all inhibitor complexes observing an increase in conformer II (except for MK-8745).

The final aspect of Chapter 4 assessed the stability of Aurora A in the presence of inhibitor complexes by using CIU. All inhibitor complexes of Aurora A exhibited greater stability than unbound proteins, with the largest effects being observed with the DFG-in inhibitor ENMD-2076 and partial DFG-out inhibitor VX-680. The CIU analysis was able to determine differences in Aurora A stability when bound to a partial DFG-out inhibitor in comparison to a complete DFG-in or DFG-out inhibitor. The two partial DFG-out inhibitors MLN8237 and VX680 showed partially unfolded intermediates that were not evident for the other inhibitors. The partial

unfolding intermediate transition states suggested that these inhibitors locked Aurora A into a specific configuration, which was dependent on the positioning of the DFG/P-loop and the noncovalent interactions of the inhibitor bound complexes. Overall, Chapter 4 successfully characterised the changes in the Aurora A conformational landscape and stability induced by the binding of small molecular inhibitors. The differences observed in stability shown by CIU for the partial DFG-out inhibitors suggest that it could be possible to utilise this technique to analyse different binding modes of inhibitors. This could potentially differentiate the inhibitors into different groups based on the partial unfolding intermediates or lack of that was observed for the complete DFG-in/out inhibitors.

Following the analysis of the serine/threonine protein kinase Aurora A, Chapter 5 examined the histidine protein kinase, NME1. It has previously been demonstrated that NME1 functions as a pHis-containing nucleotide diphosphate kinase, which controls the levels of nucleotide and triphosphates in the cell (97). Using a combination of intact LC/MS and native IM-MS, Chapter 5 aimed to evaluate the effect of nucleotides on a range of NME1 proteins, including a WT form and several mutated versions (H118A, R105A, H118A/R105A and C109A). The H118A mutated variant was generated to provide a protein that lacked the suspected histidine residue that was thought to be phosphorylated and therefore, this protein acted as a control for the analysis. The R105 residue plays an important role in the catalytic activity of NME1 and was analysed to determine whether a mutation of this residue would have an effect on nucleotide binding and the ability to act as an effective phosphate donor for auto-phosphorylation. The double mutant variant (H118A/R105A) was analysed to determine the nucleotide effects on a protein that had the removal of the autophosphorylation site and a key catalytic residue. NME1 C109A was analysed to investigate the role of C109 in NME1 complex formation under oxidative stress and the impact this mutation would have on the binding of small molecule compounds. Intact MS revealed that NME1 WT and C109A observed a shift in mass of 80 Da, representative of phosphorylation, in the presence of different nucleotides (ATP, CTP, GTP, TTP and UTP), which suggested that these two forms of NME1 were able to use a variety of nucleotides as effective phosphate donors *in vitro* to auto-phosphorylate on a single site. Previous studies have reported the H118 site to be phosphorylated (96), therefore, the NME1 H118A mutated variant was analysed, which

demonstrated no mass shift associated with phosphorylation, suggesting the H118 residue is indeed the single site of phosphorylation.

As a confirmatory step, LC/MS/MS peptide mapping could be performed on digests from purified NME1, or protein immunoprecipitated from cell lines of interest. Traditional phosphopeptide mapping methods are not suitable for analysis of acid-labile phosphohistidine, so we would need to employ methods specifically developed in the Eyers groups for this purpose. Previous studies have detected pHis in human cell lines with the use of UPAX/MS (79), therefore this technique could be further explored to detect histidine phosphorylation of NME1.

IM-MS analysis of the conformational landscape of NME1 proteins in the presence of bound nucleotides showed a level of variation between the conformer relative abundancies for the different nucleotides when bound to NME1 WT and the H118A/R105A mutated variant. This made the interpretation of the data challenging to deduce the impact that the nucleotides had on the conformational dynamics of NME1. To gain a further understanding of the nucleotide binding effects, molecular modelling similar to the analysis carried out for Aurora A would be useful to explore the structural configurations associated with the abundance of different conformers. Nucleotide binding and phosphate transfer from the different nucleotides could potentially have varied functions, therefore, if differences are observed between the nucleotides, this could provide a further insight into the structural conformation that NME1 is adopting in the presence of each nucleotide.

Following the analysis of the effects of nucleotides on NME1 by intact MS, the oligomeric structure of NME1 was analysed by native MS. It is reported that NME1 forms natively as a hexamer, with this formation required for full enzymatic activity (203). Under oxidising conditions, this hexamer dissociates with the cysteines (Cys4, Cys109 and Cys145) in NME1 forming disulphide bonds and stabilising a dimeric structure with reduced activity (205). Chapter 5 utilised native MS to demonstrate that upon incubation of NME1 with a range of compounds selected from the FDA-drug approved library of small molecules, the hexameric form of NME1 WT dissociated to form a dimer. The compounds inducing a negative shift when analysed by DSF suggested that this could be a result of oxidative stress due to many of the

compounds possessing functional groups that can accept electrons, and S1949 (Menadione) being previously reported as an oxidising agent (212). Although we predicted that analysis of the NME1 C109A variant with the selected compounds would reveal resistance of hexamer dissociation with an increase in the dimeric form, the majority of the compounds tested resulted in the formation of trimers, suggesting a different mechanism of engagement. This C109A mutated version still contained the Cys4 and Cys145 residues, which have been reported to form a disulphide bond (205). It is possible that the presence of these conserved Cys residues that remained in the C109A protein were making it prone to oxidative stress due to the modifications of the highly reactive nucleophilic thiol groups and therefore dissociation of the hexamer would occur under these oxidating conditions. Interestingly, compound-induced change in hexamer formation did not noticeably alter the activity of NME1 as determined by auto-phosphorylation.

Future experiments associated with multimeric status by native MS and disulphide mapping by LC/MS to analyse an NME1 protein with the additional Cys to Ala residue modifications would confirm the hypothesis that NME1 does exhibit Cys-redox regulated multimer formation and the effect of this on activity. Oxidative stress could be induced with the addition of diamide or hydrogen peroxide or additionally, the protein could be made in the presence of peroxide, although this could potentially complicate the MS analysis due to additional adducts seen upon treatment of proteins with hydrogen peroxide. In addition to the analysis of further NME1 mutated variants, more compounds analysed by native MS from the 976 that were screened using DSF would be beneficial to gain a further understanding of the structural changes that these compounds impose on NME1. The FDA-approved drug library of small molecules was analysed in this thesis due to NME1 not being a standard protein kinase. However, future experiments associated with screening protein kinase inhibitor libraries could be of additional importance, such as the protein kinase inhibitor set 1 (PKIS1) library, which has a total of 367 ATP-competitive kinase inhibitors (217). Validating the effect of the compounds in cells would be useful to gain a further understanding into the identification of other putative substrates, which could ultimately lead to the generation of small molecule inhibitors for future investigation of the cellular roles of NME1.

Overall, the work carried out in this thesis described the use of IM-MS to evaluate the differences in the conformational landscape of Aurora A and NME1 when bound to small molecule inhibitor compounds. The use of IM-MS was able to detect differences observed between phosphorylated and non-phosphorylated Aurora A and Gaussian fitting of the CCS distribution revealed subtle changes dependent on the activation status or binding of inhibitor. In addition to the differences observed based on Aurora A inhibitor binding, native MS revealed the oligomeric changes upon NME1 when bound to compounds, showing dissociation from the hexamer to the dimer form. The IM-MS work completed in this thesis could be used to gain a further insight into the conformational space of many other protein kinases to aid in the discovery of future atypical protein kinase inhibitors.

Appendix 2.1

Appendix 2.1.1 R code for CCS scatter plots

```
# Import data
data <-read.csv(file.choose())
install.packages("svglite")
install.packages("Cairo")
library(svglite)


library(ggplot2)
cbbPalette <- c("#999999", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2",
"#D55E00", "#CC79A7")


#make scatterplot - size
ggplot(data=data,   aes(x=CCS,   y=CCSD))  +  geom_point(alpha=1,   size=3*(data$Area),
aes(color=ID))+
  scale_color_manual(values=cbbPalette) + theme_classic() +
  geom_errorbar(aes(ymin=LowerCCSD,ymax=UpperCCSD,color=ID),width=0.2,linetype=2) +
  geom_errorbarh(aes(xmax=UpperCCS,xmin=LowerCCS,color=ID),width=0.2,linetype=2) +
  xlim(40,65) + ylim(0,10)
```

Appendix 2.2

Appendix 2.2.1 R code for fragmentation maps

```
y <- read.csv("y.csv",

b <- read.csv("b.csv",

unmatched <- read.csv("unmatched.csv",

all <- read.csv("all.csv",

y_limits <- c(105,25)


p <- ggplot(data = all,

      aes(x = mz,

         y = i,

          colour = ion,

          label = matches))+

  geom_col(data = unmatched,

      aes(x = mz,

         y = i),

      color = "#C0C0C0",

      fill = "#C0C0C0",

      width = 2,

      size = 1) +

  geom_col(data = y,

      aes(x = mz,

         y = i),

      color = "#0000CC",

      fill = "#0000CC",

      width = 2,

      size = 1) +

  geom_col(data = b,

      aes(x = mz,

         y = i),
```

```r
              color = "#FF0000",

              fill = "#FF0000",

              width = 2,

              size = 1) +

geom_text_repel(vjust= -1,

                hjust = 0,

                size = 5,

                ylim = y_limits,

                min.segment.length = 0.5,

                direction = "y",

                force = 50,

                segment.alpha = 0.2) +

theme(axis.text.x = element_text(size = 25,

                      family = "sans",

                      colour = "black"),

      axis.text.y = element_text(size = 25,

                      family = "sans",

                      colour = "black"),

      axis.title.x = element_text(size = 25,

                       face = "bold.italic",

                       colour = "black",

                       vjust = 0),

      axis.title.y = element_text(size = 25,

                       face = "bold",

                       colour = "black", hjust = 0.5),

      plot.margin = margin(0, 1, 0, 0, "cm"),

      legend.position = "none",

      axis.line = element_line(colour = 'black', size = 1),

      axis.ticks = element_line(size = 1),

      axis.ticks.length = unit(.25, "cm"),

      panel.grid.major = element_blank(),

      panel.grid.minor = element_blank(),
```

```
        panel.background = element_rect(fill = 'NA', colour = "black")) +

  scale_x_continuous(limits = c(75,705), expand = c(0, 0

            labels = c(100, 200, 300, 400, 500, 600, 700),

            breaks = c(100, 200, 300, 400, 500, 600, 700)) +

  scale_y_continuous(limits = c(0,105), expand = c(0, 0)) +

  scale_color_manual(values=c("#C0C0C0","#FF0000", "#0000CC", "#008000"))+

  xlab("m/z") +

  ylab("Relative abundance") +

  geom_segment(aes(x = 80, y = -0.5, xend = 705, yend = -0.5), colour = "black")


aspect_ratio <- 1.67

ggsave("test.svg", height = 6 , width = 6 * aspect_ratio)
```

Appendix 5.1.1 Native MS spectra for NME1 WT/C109A + Compounds

+ S1198 (Irinotecan)
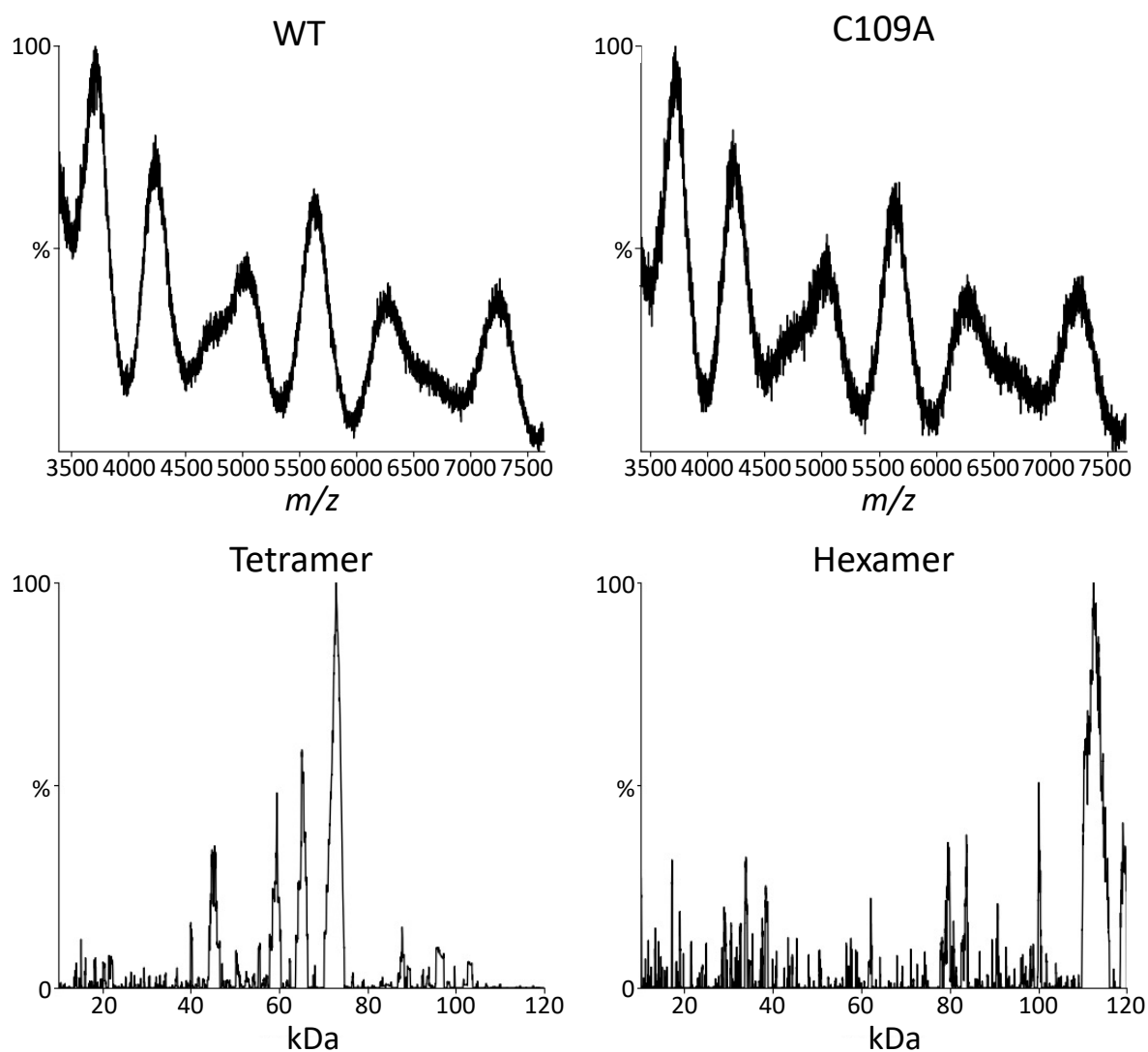
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1344 (Glimepiride)
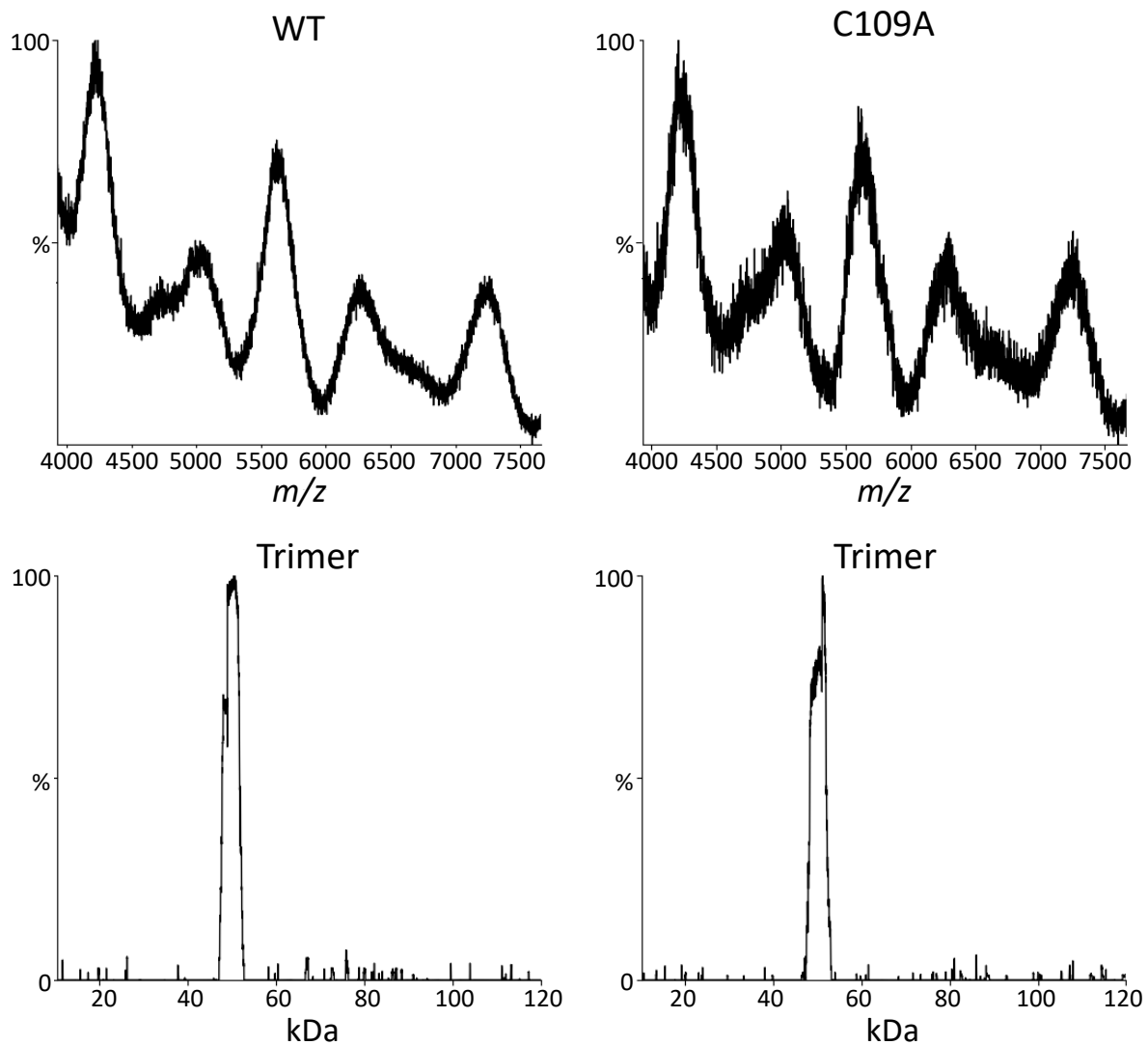
WT

C109A

Trimer

Trimer

Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1457 (Atazanavir sulfate)

Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1512 (Tadalafil)

Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1606 (Clotrimazole)
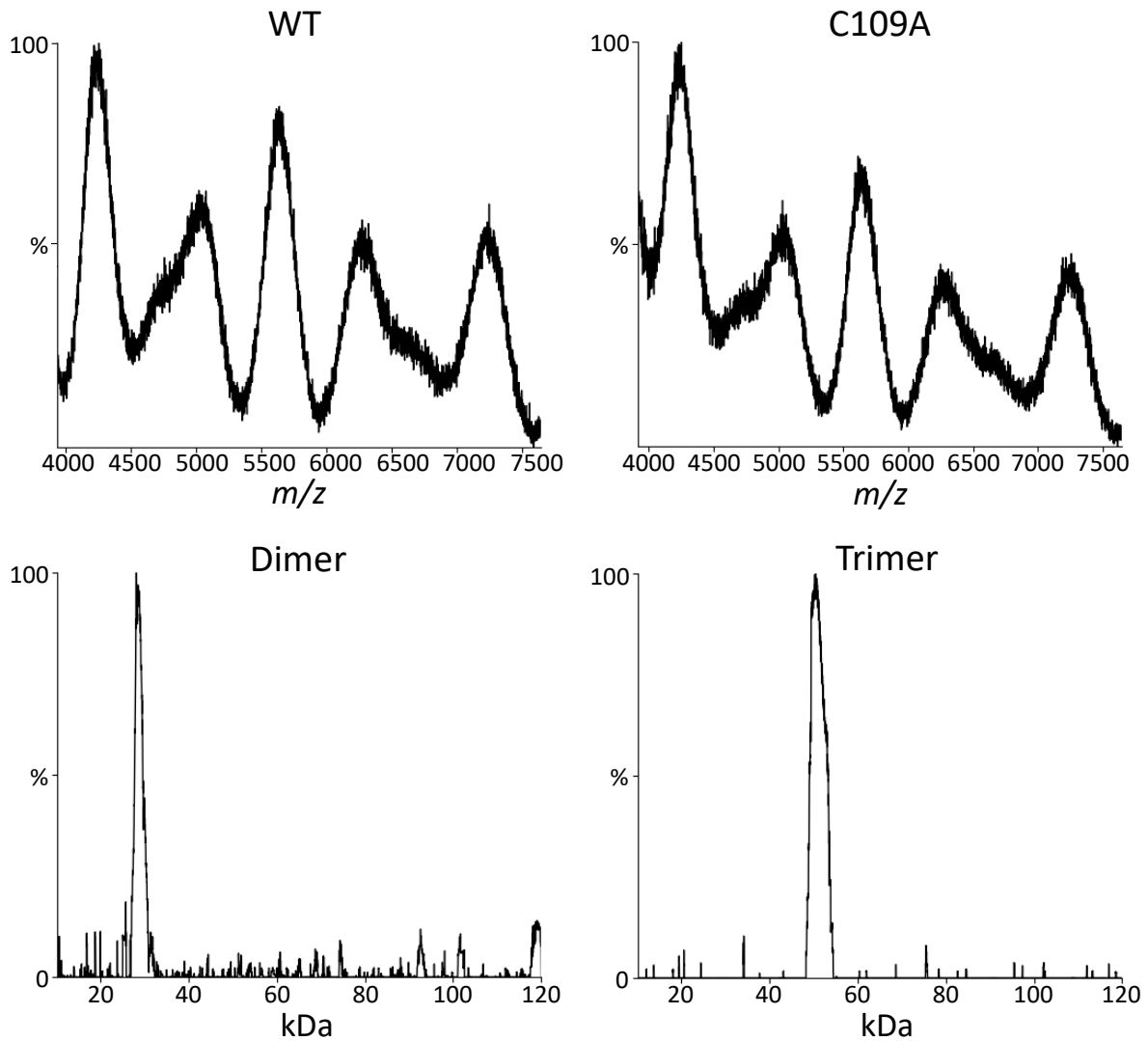
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1633 (Zafirlukast)
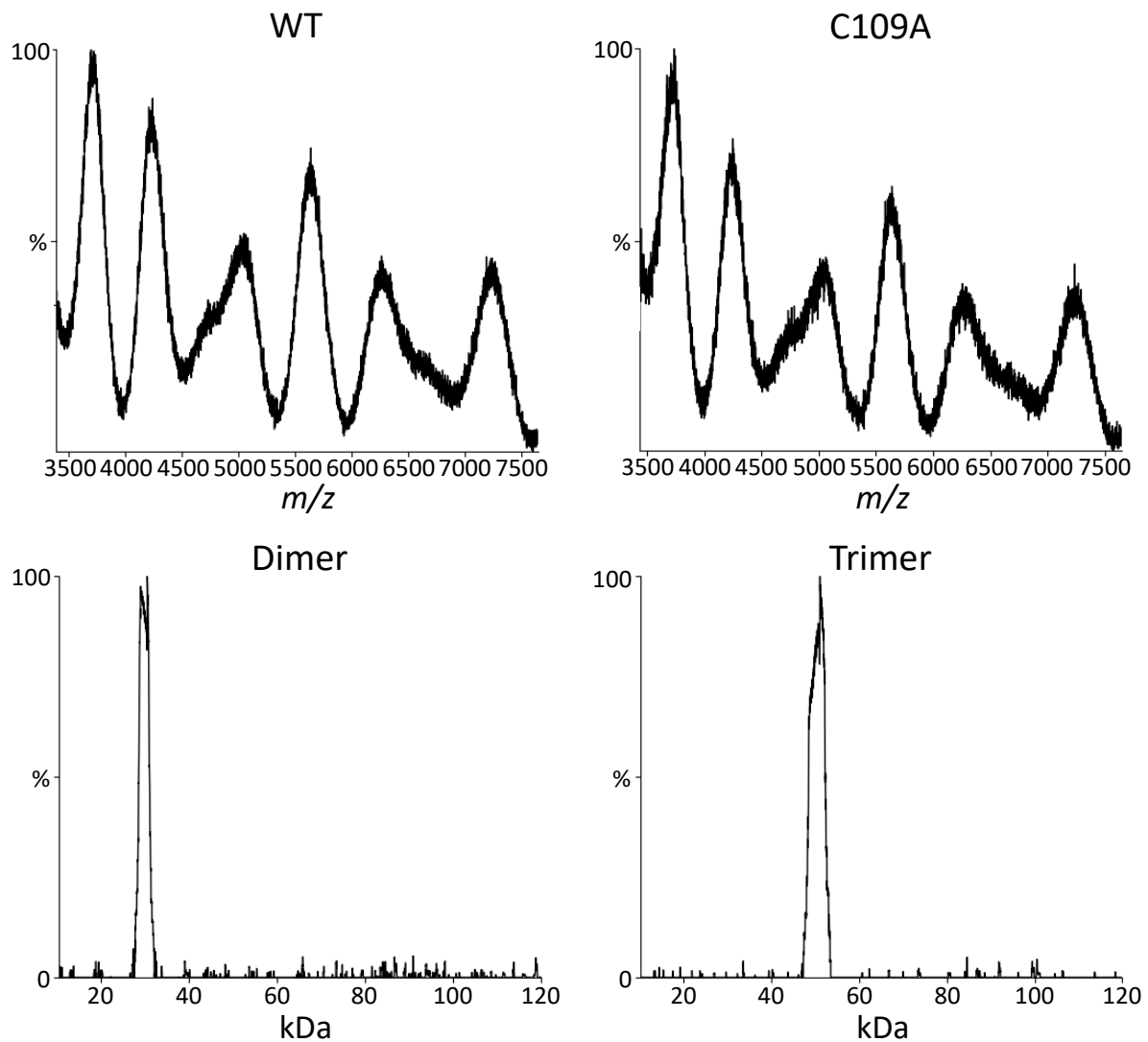
WT

C109A

Dimer

Trimer

Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1730 (Indapamide)
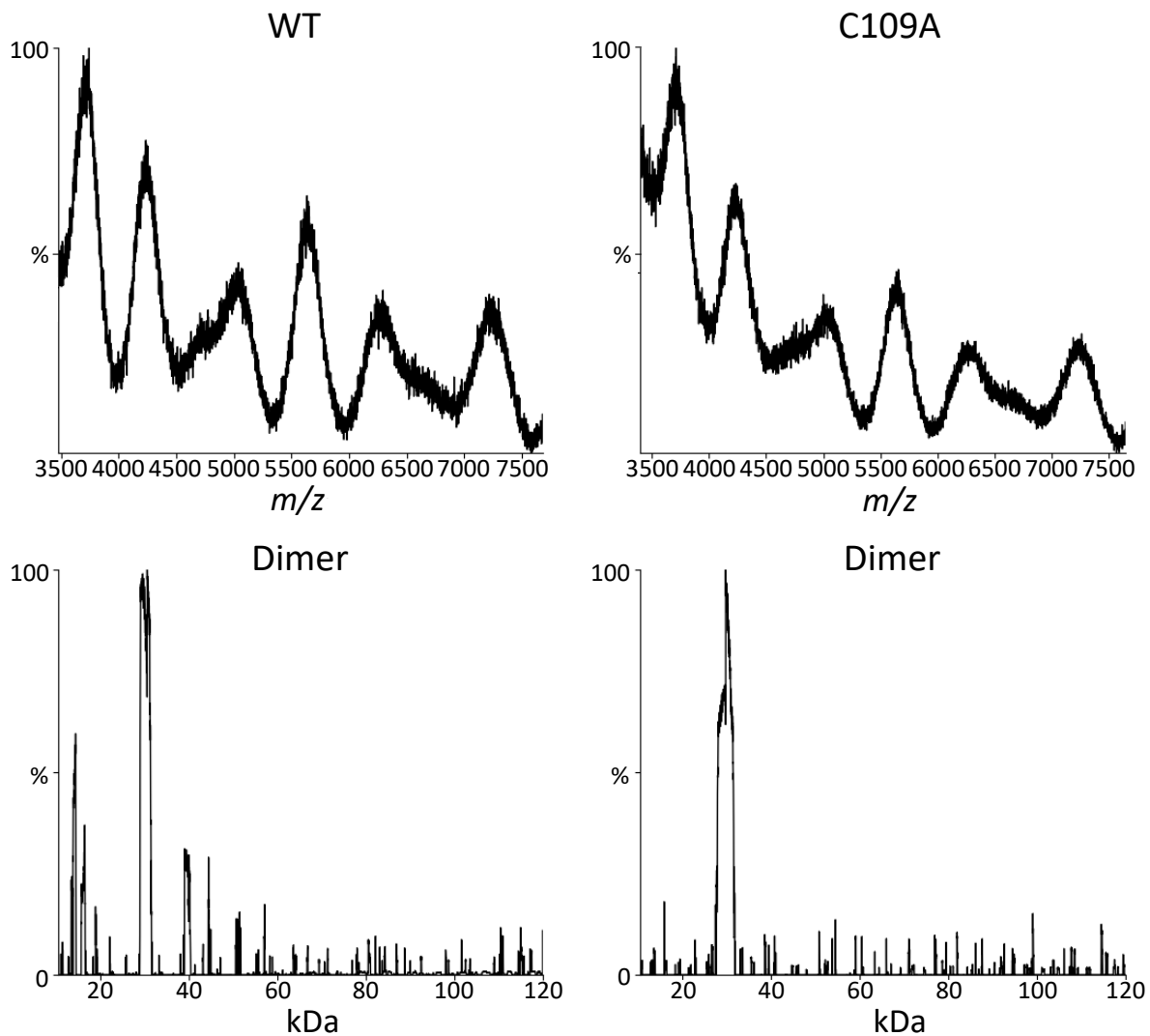
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1737 (Prednisolone)
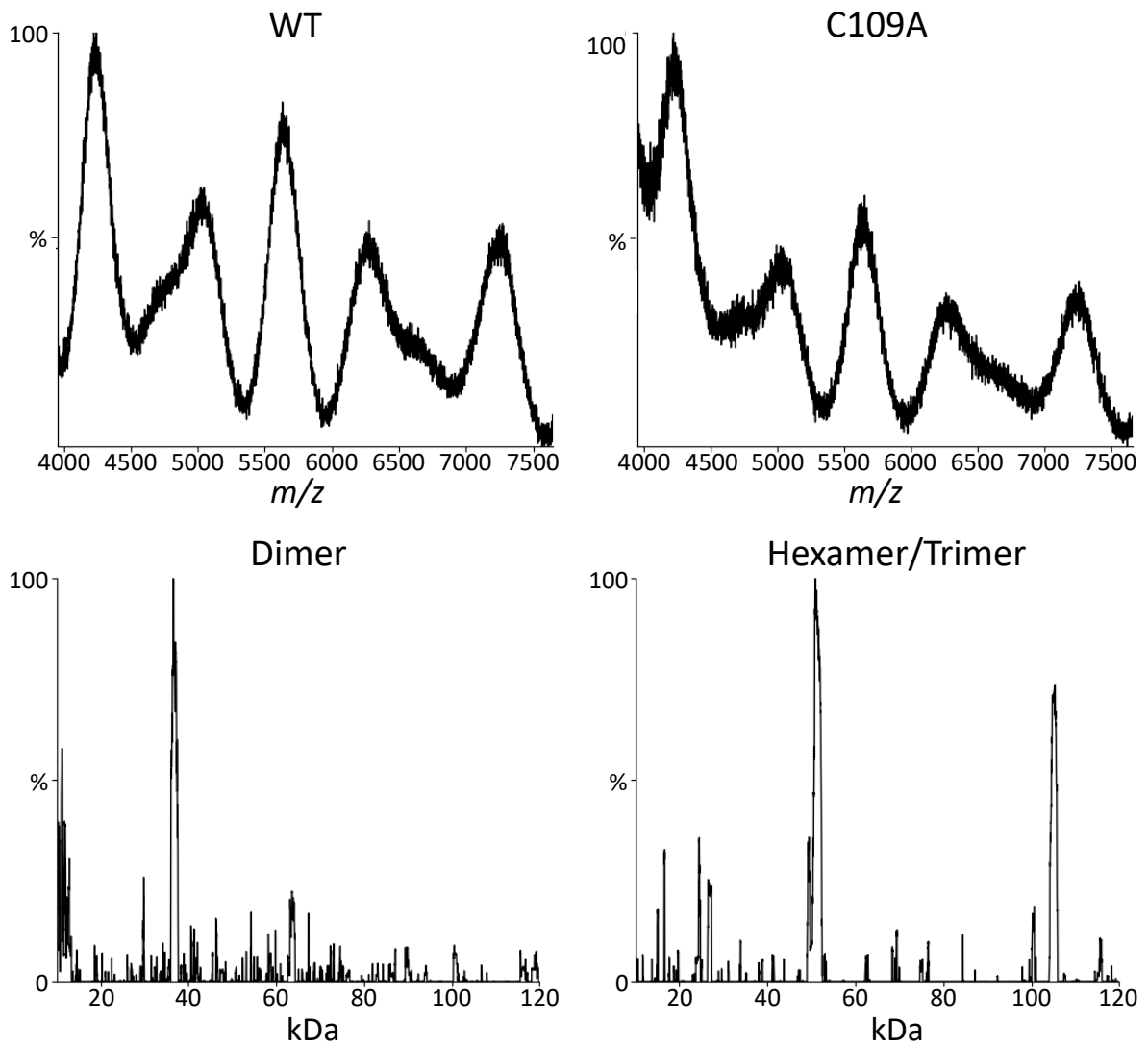
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1760 (Rifapentine)
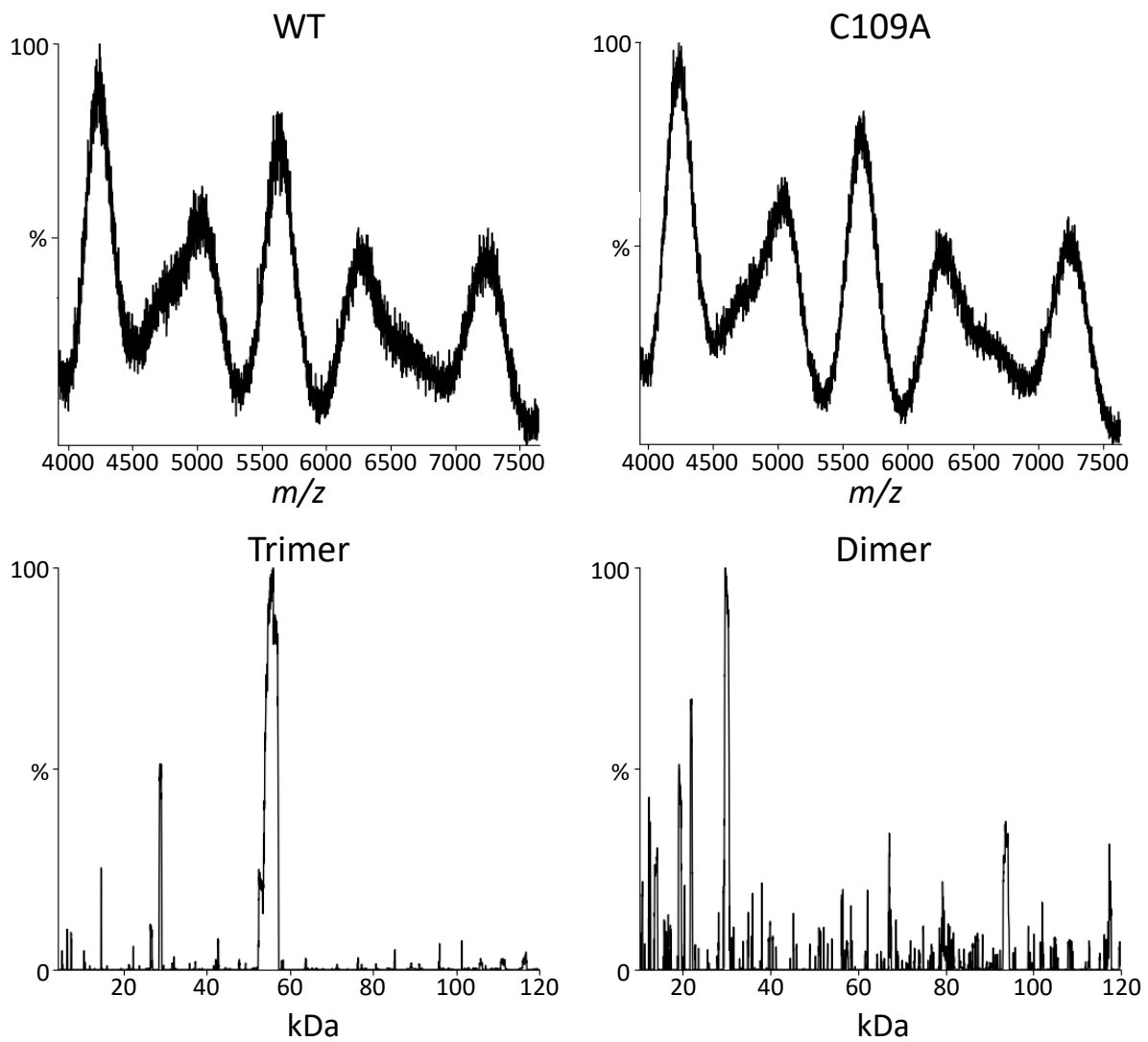
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1808 (Nifedipine)
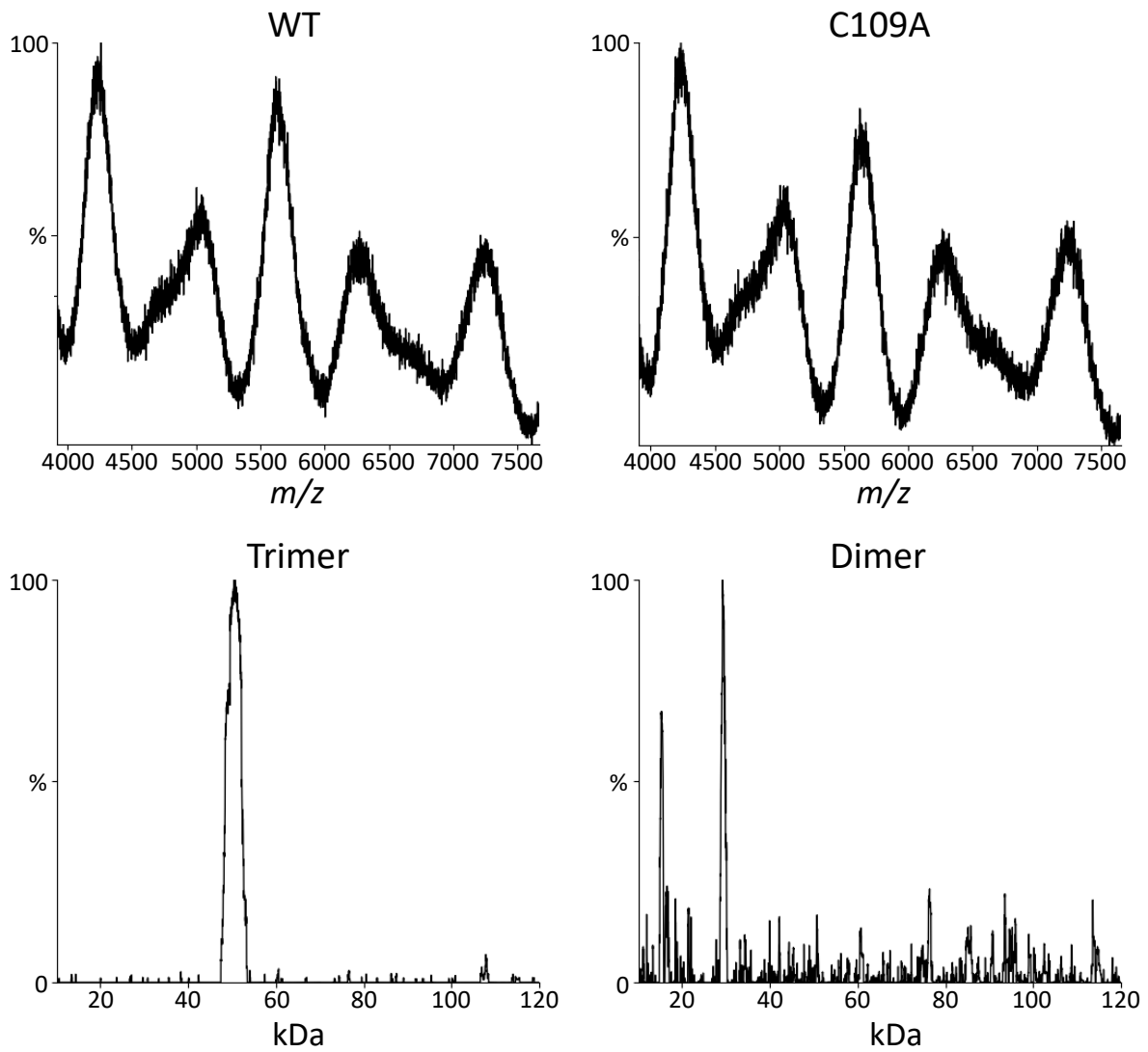
WT

C109A

Trimer

Dimer

Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1840 (Lomustine)
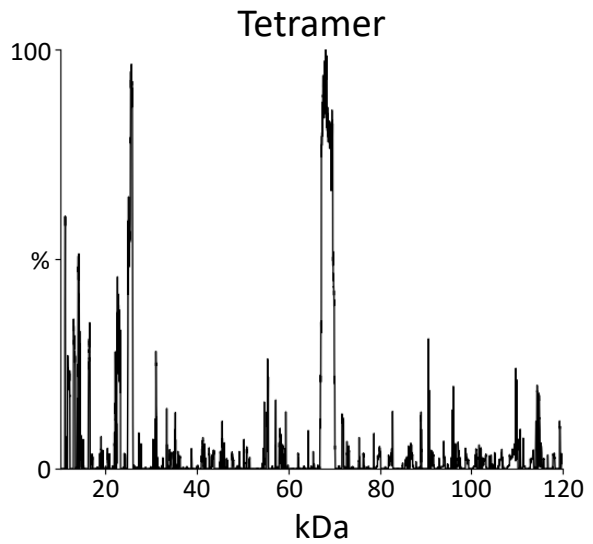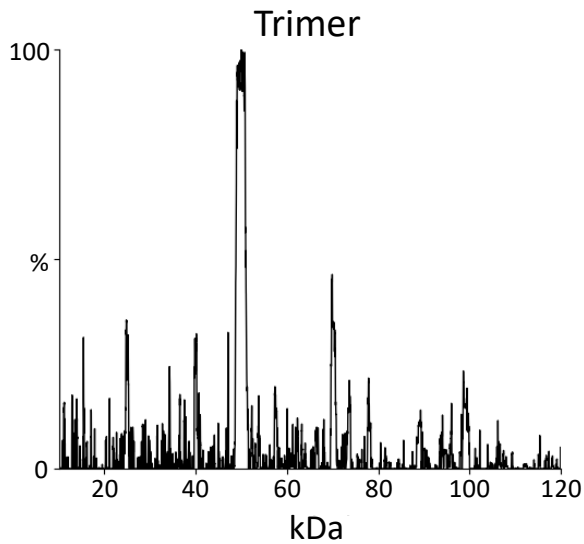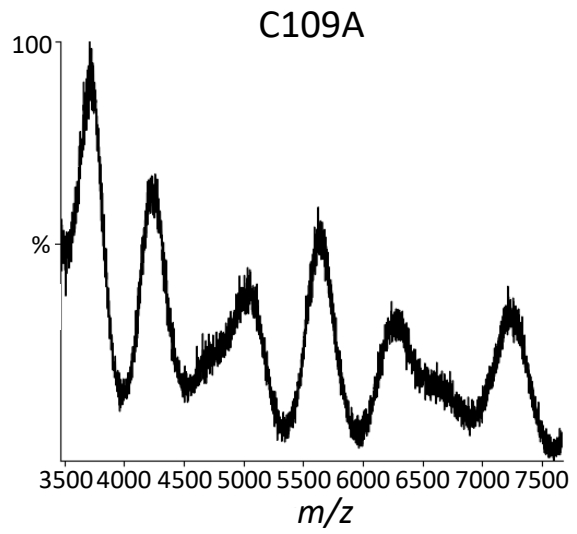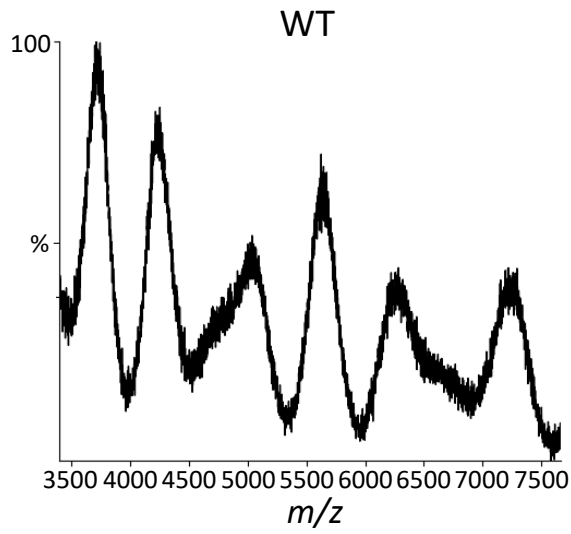
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1847 (Clemastine fumarate)
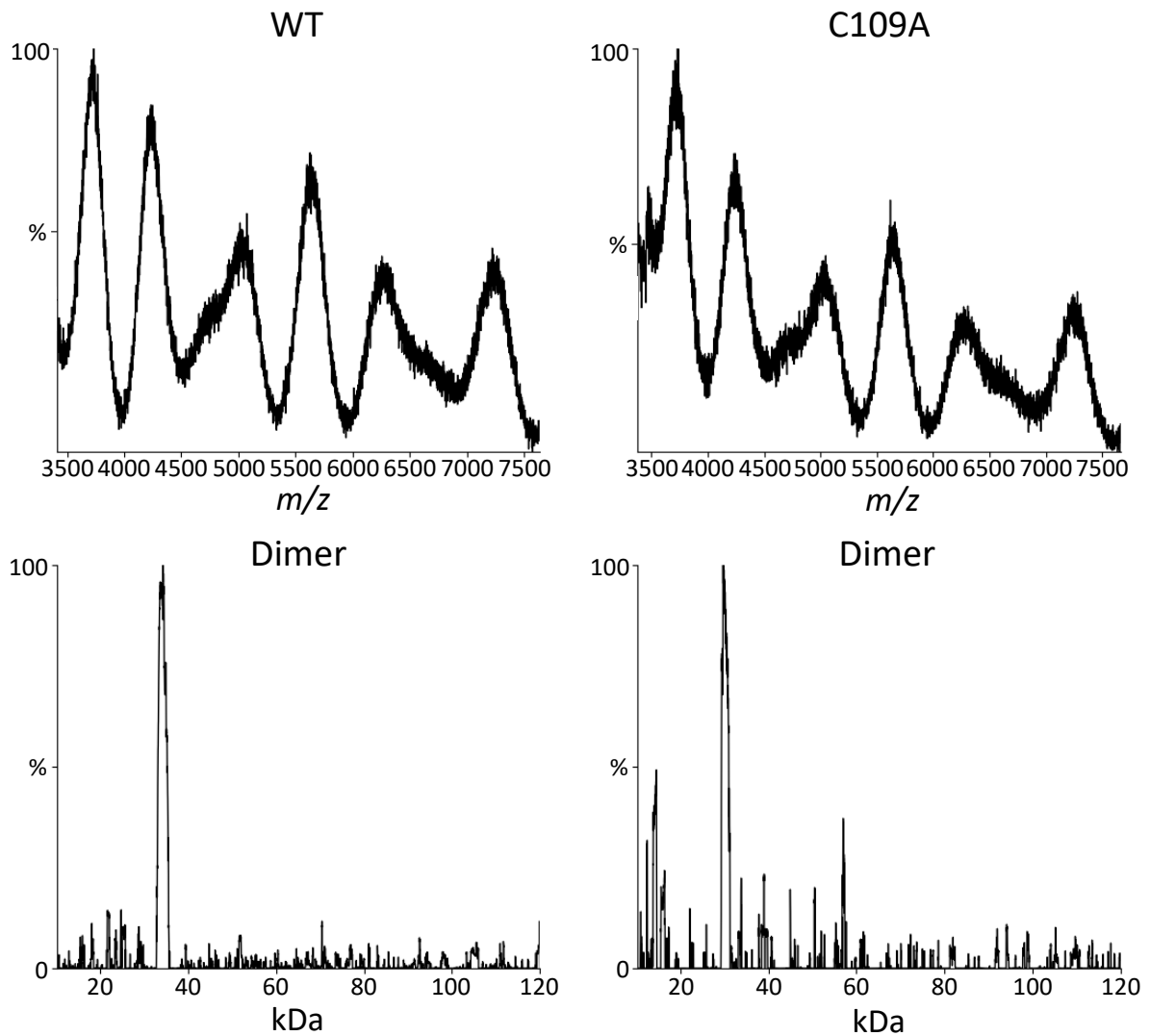
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S1859 (Diethylstilbestrol)
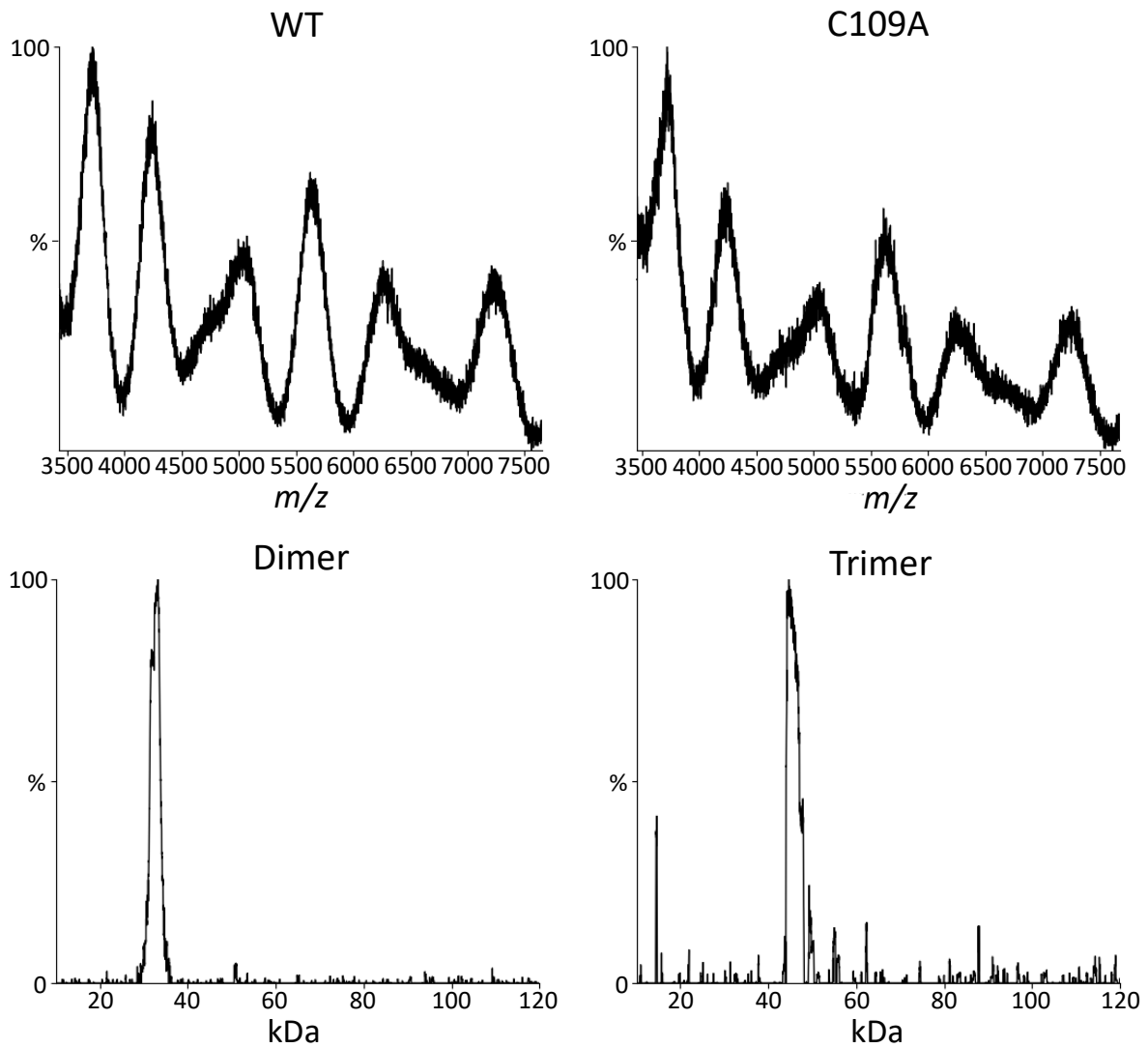
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S2522 (Adrenaline)
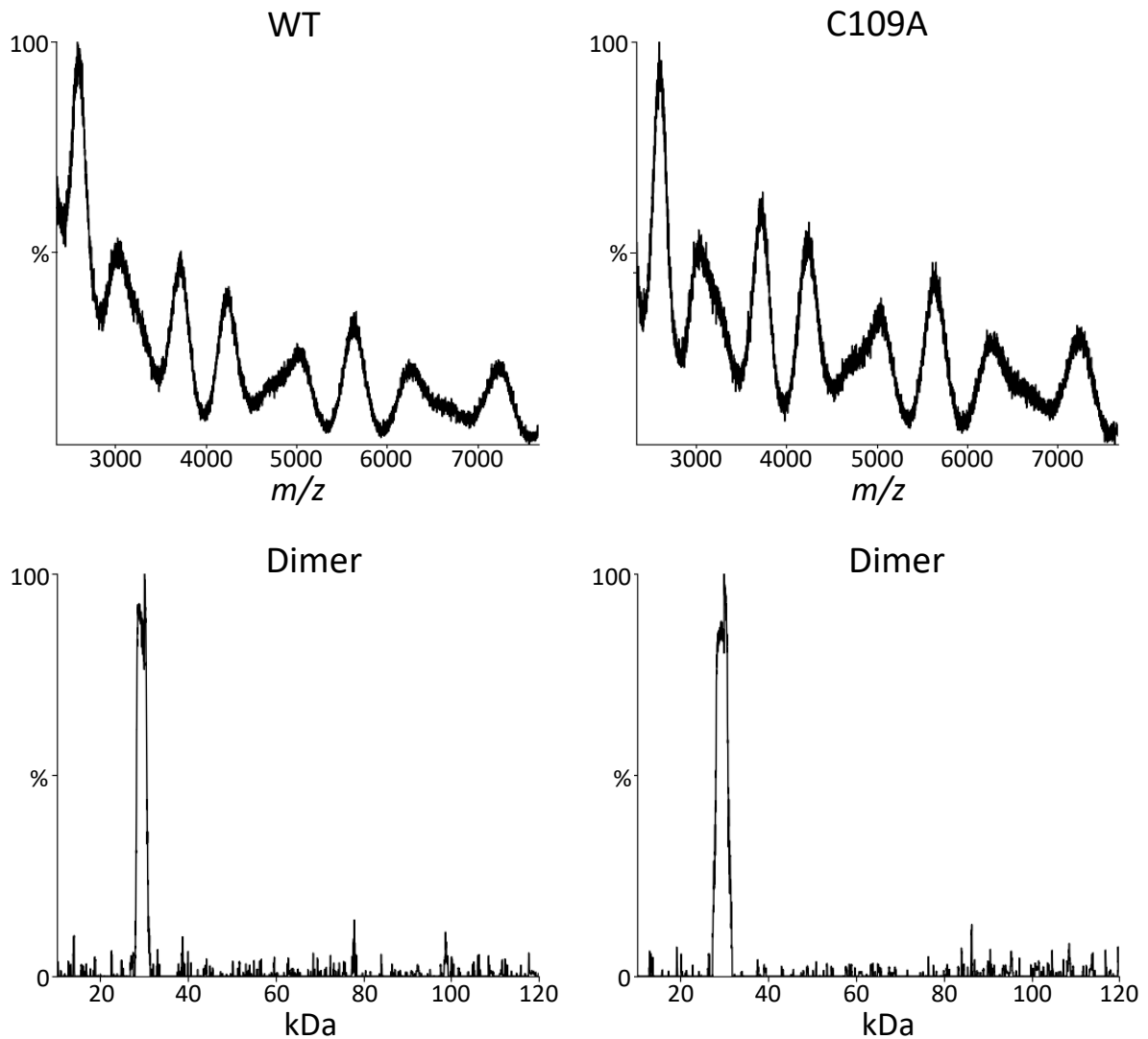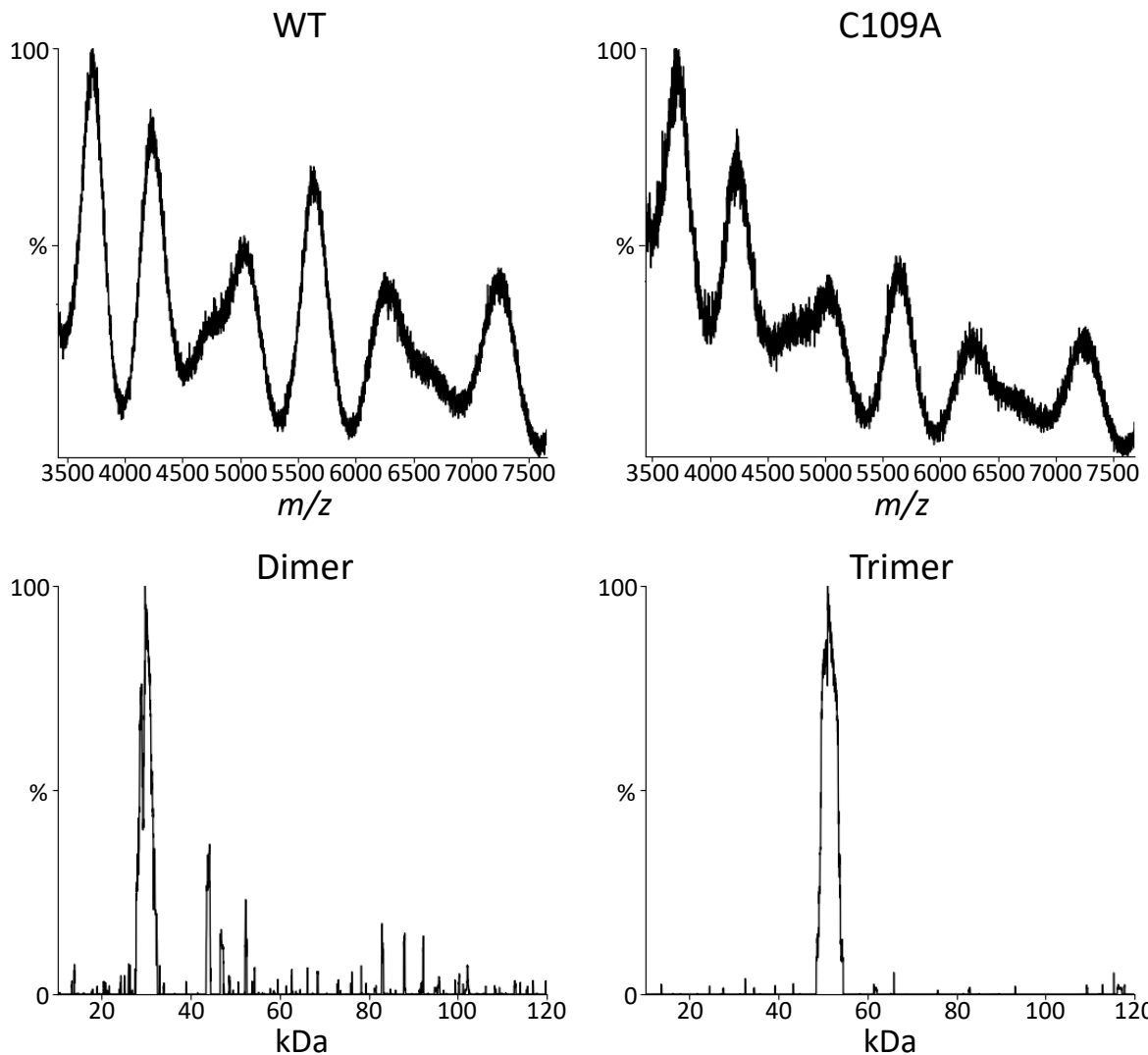
WT

C109A

Dimer

Trimer

Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.
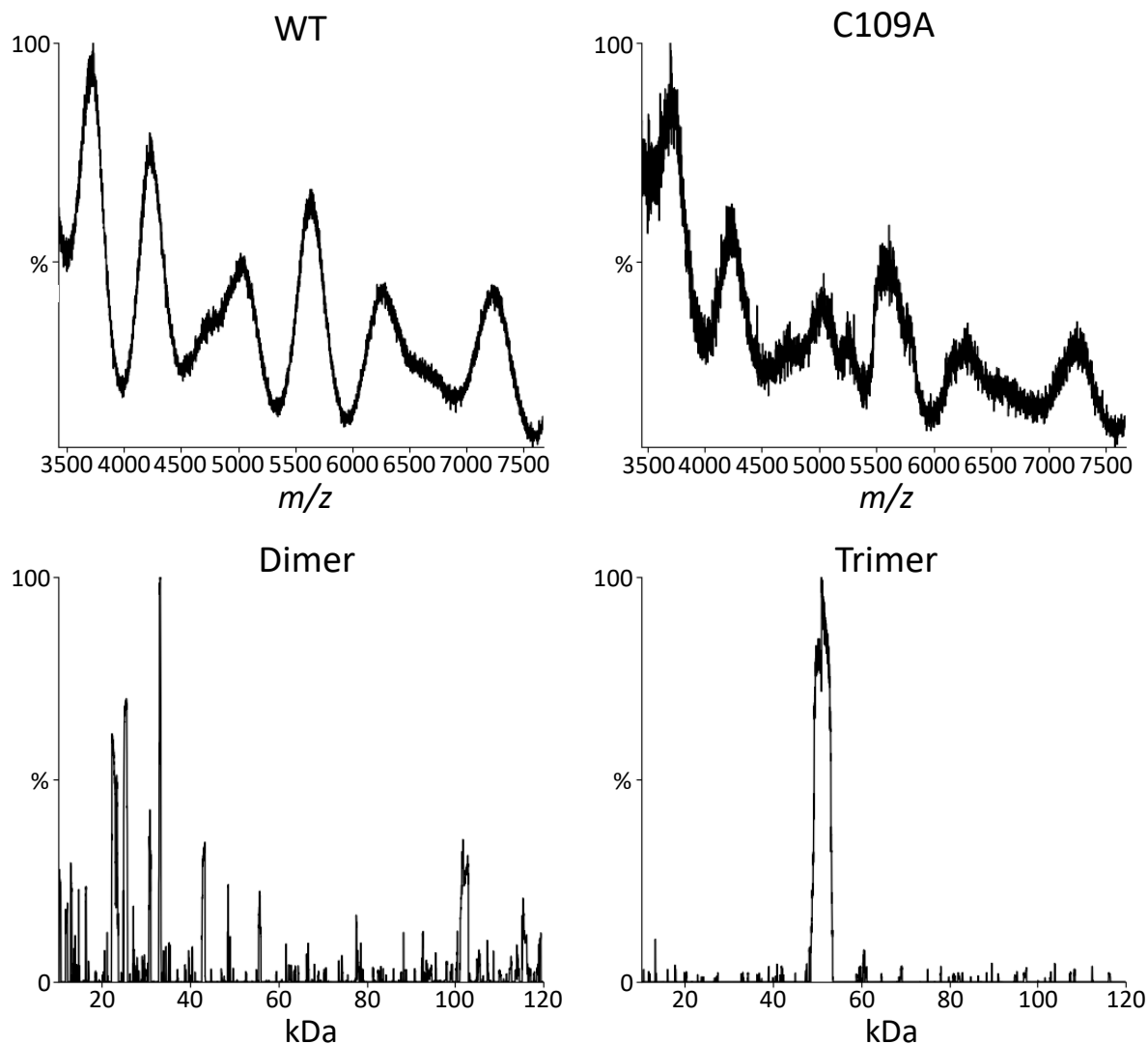
+ S4003 (Lithocholic acid)
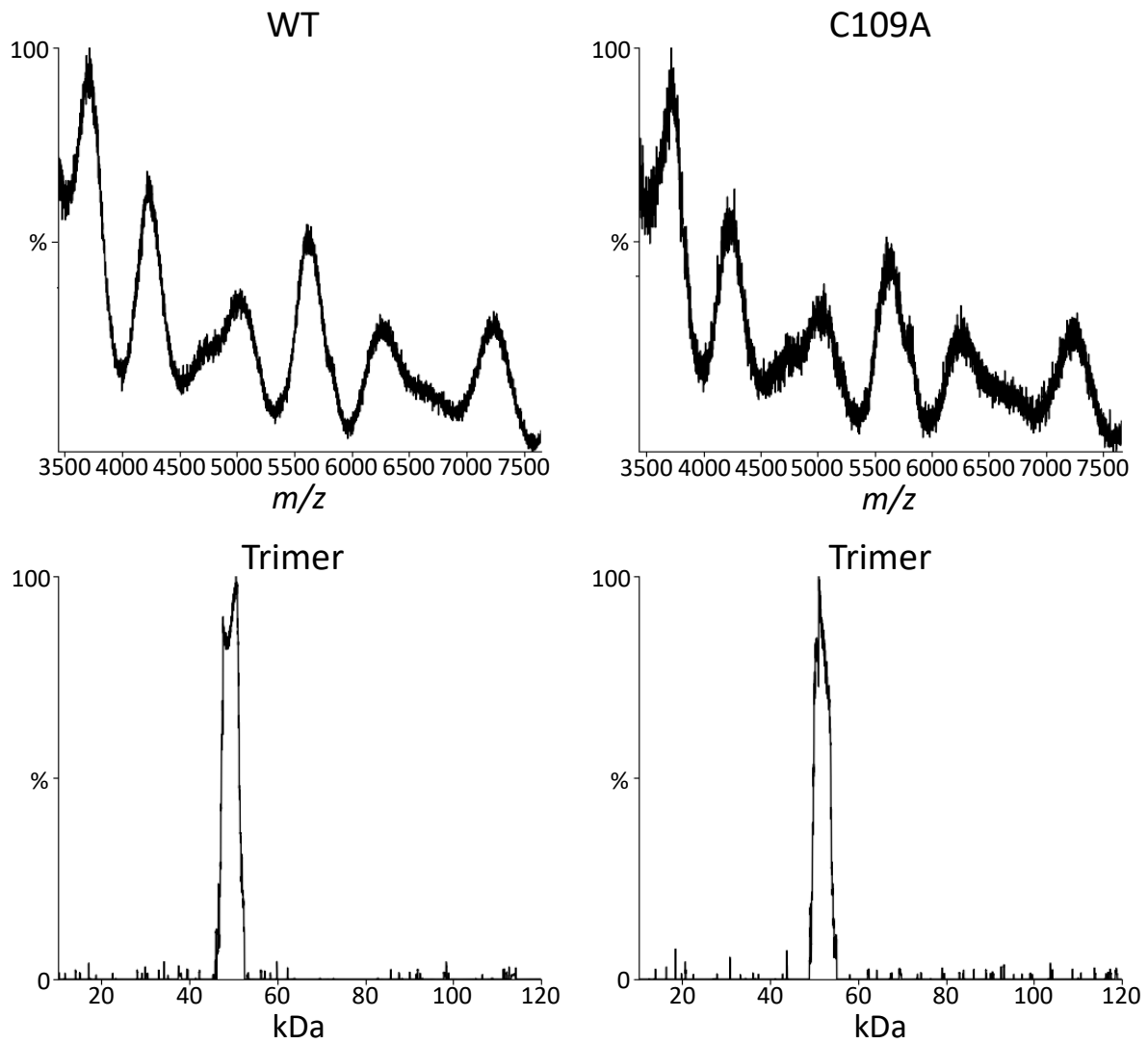
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S4073 (Aminosalicylate sodium)
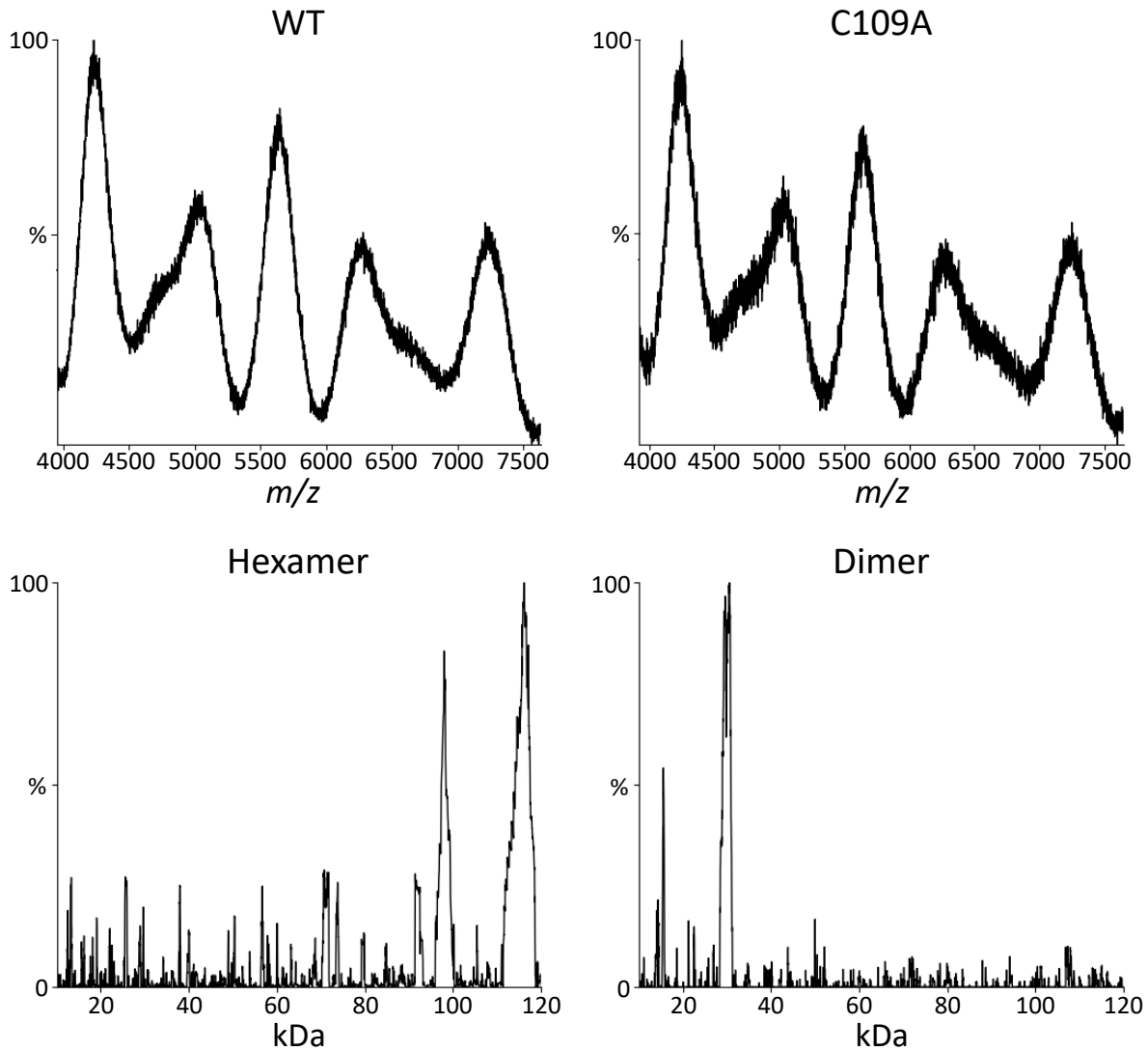
WT

C109A

Hexamer

Dimer

Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

+ S4221 (Benzbromarone)
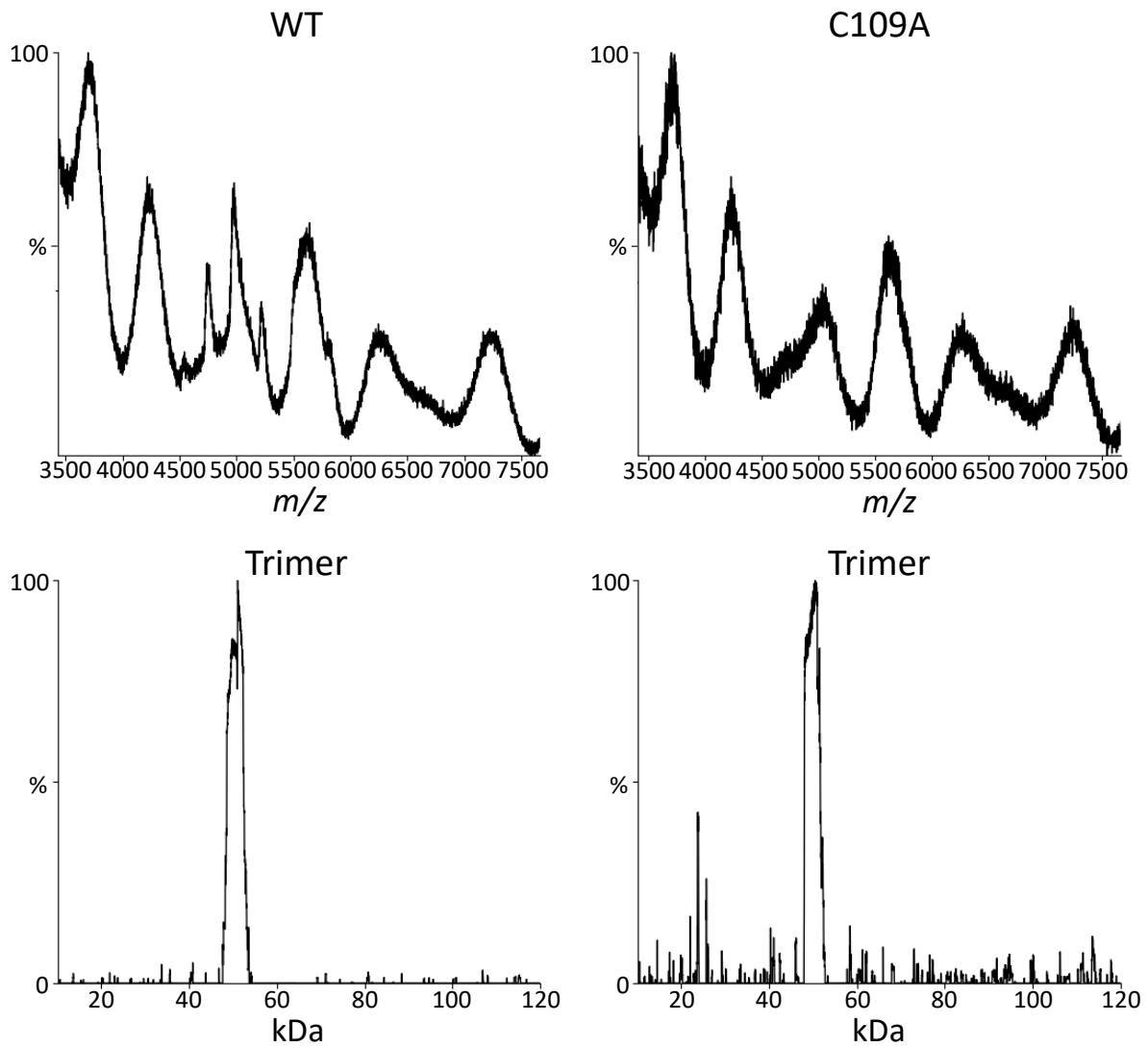
Native ESI spectrum for NME1 WT (left) and C109A (right) + compounds. Data analysis carried out using MassLynx 4.1 and figures created using UniDec. Protein buffer exchanged into 150 mM ammonium acetate using Amicon 10 kDa cut-off spin columns prior to MS analysis.

Bibliography

1.   Krebs EG. An accidental biochemist. Annu Rev Biochem. 1998;67:xii-xxxii.

2.   Ardito F, Giuliani M, Perrone D, Troiano G, Lo Muzio L. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). Int J Mol Med. 2017;40(2):271-80.

3.   Shchemelinin I, Sefc L, Necas E. Protein Kinases, Their Function and Implication in Cancer and Other Diseases. 2006:81.

4.   Kondapalli L, Soltani K, Lacouture ME. The promise of molecular targeted therapies: protein kinase inhibitors in the treatment of cutaneous malignancies. J Am Acad Dermatol. 2005;53(2):291-302.

5.   Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science. 2002;298(5600):1912-34.

6.   Roskoski R, Jr. A historical overview of protein kinases and their targeted small molecule inhibitors. Pharmacol Res. 2015;100:1-23.

7.   Scheeff ED, Bourne PE. Structural evolution of the protein kinase-like superfamily. PLoS Comput Biol. 2005;1(5):e49-e.

8.   Modi V, Dunbrack RL. Defining a new nomenclature for the structures of active and inactive kinases. Proceedings of the National Academy of Sciences. 2019;116(14):6818.

9.   Taylor SS, Kornev AP. Protein kinases: evolution of dynamic regulatory proteins. Trends Biochem Sci. 2011;36(2):65-77.

10.   Johnson DA, Akamine P, Radzio-Andzelm E, Madhusudan M, Taylor SS. Dynamics of cAMP-dependent protein kinase. Chem Rev. 2001;101(8):2243-70.

11.   Narayana N, Cox S, Shaltiel S, Taylor SS, Xuong N. Crystal structure of a polyhistidine-tagged recombinant catalytic subunit of cAMP-dependent protein kinase complexed with the peptide inhibitor PKI(5-24) and adenosine. Biochemistry. 1997;36(15):4438-48.

12.   Kannan N, Haste N, Taylor SS, Neuwald AF. The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module. Proceedings of the National Academy of Sciences of the United States of America. 2007;104(4):1272-7.

13.   Leroux AE, Schulze JO, Biondi RM. AGC kinases, mechanisms of regulation and innovative drug development. Seminars in Cancer Biology. 2018;48:1-17.

14.   Biondi RM, Komander D, Thomas CC, Lizcano JM, Deak M, Alessi DR, et al. High resolution crystal structure of the human PDK1 catalytic domain defines the regulatory phosphopeptide docking site. The EMBO journal. 2002;21(16):4219-28.

15.   Yang J, Cron P, Thompson V, Good VM, Hess D, Hemmings BA, et al. Molecular mechanism for the regulation of protein kinase B/Akt by hydrophobic motif phosphorylation. Mol Cell. 2002;9(6):1227-40.

16.   Huse M, Kuriyan J. The conformational plasticity of protein kinases. Cell. 2002;109(3):275-82.

17.   Kornev AP, Haste NM, Taylor SS, Eyck LFT. Surface Comparison of Active and Inactive Protein Kinases Identifies a Conserved Activation Mechanism. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(47):17783-8.

18.   Bayliss R, Fry A, Haq T, Yeoh S. On the molecular mechanisms of mitotic kinase activation. Open Biol. 2012;2(11):120136.

19.   Steichen JM, Iyer GH, Li S, Saldanha SA, Deal MS, Woods VL, Jr., et al. Global consequences of activation loop phosphorylation on protein kinase A. J Biol Chem. 2010;285(6):3825-32.

20.   Ke YW, Dou Z, Zhang J, Yao XB. Function and regulation of Aurora/Ipl1p kinase family in cell division. Cell Research. 2003;13(2):69-81.

21.   Quartuccio S, Schindler K. Functions of Aurora kinase C in meiosis and cancer. Frontiers in Cell and Developmental Biology. 2015;3.

22.   Bischoff JR, Plowman GD. The Aurora/Ipl1p kinase family: regulators of chromosome segregation and cytokinesis. Trends in Cell Biology. 1999;9(11):454-9.

23.   Walter AO, Seghezzi W, Korver W, Sheung J, Lees E. The mitotic serine/threonine kinase Aurora2/AIK is regulated by phosphorylation and degradation. Oncogene. 2000;19(42):4906-16.

24.   Schafer KA. The cell cycle: a review. Vet Pathol. 1998;35(6):461-78.

25.   Crane R, Gadea B, Littlepage L, Wu H, Ruderman JV. Aurora A, meiosis and mitosis. Biol Cell. 2004;96(3):215-29.

26.   Kollareddy M. Aurora kinases: structure, functions and their association with cancer. Biomedical Papers. 2008:27-33.

27.   Sugimoto K, Urano T, Zushi H, Inoue K, Tasaka H, Tachibana M, et al. Molecular Dynamics of Aurora-A Kinase in Living Mitotic Cells Simultaneously Visualized with Histone H3 and Nuclear Membrane Protein Importin&alpha. Cell Structure and Function. 2002;27(6):457-67.

28.   Carmena M, Earnshaw WC. The cellular geography of Aurora kinases. Nature Reviews Molecular Cell Biology. 2003;4(11):842-54.

29.   Doxsey S. Duplicating Dangerously: Linking Centrosome Duplication and Aneuploidy. Molecular Cell. 2002;10(3):439-40.

30.   Meraldi P, Honda R, A.Nigg E. Aurora-A overexpression reveals tetraploidization as a major route to centrosome amplification in p53−/− cells. The EMBO Journal. 2002;21(4):483-92.

31.   Kaestner P, Stolz A, Bastians H. Determinants for the efficiency of anticancer drugs targeting either Aurora-A or Aurora-B kinases in human colon carcinoma cells. Molecular Cancer Therapeutics. 2009;8(7):2046.

32.   Gustafson William C, Meyerowitz Justin G, Nekritz Erin A, Chen J, Benes C, Charron E, et al. Drugging MYCN through an Allosteric Transition in Aurora Kinase A. Cancer Cell. 2014;26(3):414-27.

33.   Eyers PA, Erikson E, Chen LG, Maller JL. A Novel Mechanism for Activation of the Protein Kinase Aurora A. Current Biology. 2003;13(8):691-7.

34.   Bayliss R, Burgess SG, McIntyre PJ. Switching Aurora-A kinase on and off at an allosteric site. The FEBS Journal. 2017;284(18):2947-54.

35.   Bayliss R, Sardon T, Vernos I, Conti E. Structural Basis of Aurora-A Activation by TPX2 at the Mitotic Spindle. Molecular Cell. 2003;12(4):851-62.

36.   Chen H, Liu H, Qing G. Targeting oncogenic Myc as a strategy for cancer treatment. Signal Transduction and Targeted Therapy. 2018;3(1):5.

37.   Soucek L, Whitfield J, Martins CP, Finch AJ, Murphy DJ, Sodir NM, et al. Modelling Myc inhibition as a cancer therapy. Nature. 2008;455(7213):679-83.

38.   Kohl NE, Kanda N, Schreck RR, Bruns G, Latt SA, Gilbert F, et al. Transposition and amplification of oncogene-related sequences in human neuroblastomas. Cell. 1983;35(2 Pt 1):359-67.

39.   Tu WB, Helander S, Pilstål R, Hickman KA, Lourenco C, Jurisica I, et al. Myc and its interactors take shape. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2015;1849(5):469-83.

40.   Mark W. Richards SGB, Evon Poon, Anne Carstensen, Martin Eilers, Louis Chesler and Richard Bayliss. Structural basis of N-Myc binding by Aurora-A and its destabilization by kinase inhibitors. PNAS. 2016;113.

41.   Otto T, Horn S, Brockmann M, Eilers U, Schüttrumpf L, Popov N, et al. Stabilization of N-Myc is a critical function of Aurora A in human neuroblastoma. Cancer Cell. 2009;15(1):67-78.

42.   Brockmann M, Poon E, Berry T, Carstensen A, Deubzer HE, Rycak L, et al. Small molecule inhibitors of aurora-a induce proteasomal degradation of N-myc in childhood neuroblastoma. Cancer Cell. 2013;24(1):75-89.

43.   Peset I, Vernos I. The TACC proteins: TACC-ling microtubule dynamics and centrosome function. Trends Cell Biol. 2008;18(8):379-88.

44.   Ha G-H, Kim J-L, Breuer E-KY. Transforming acidic coiled-coil proteins (TACCs) in human cancer. Cancer Letters. 2013;336(1):24-33.

45.   Burgess SG, Peset I, Joseph N, Cavazza T, Vernos I, Pfuhl M, et al. Aurora-A-Dependent Control of TACC3 Influences the Rate of Mitotic Spindle Assembly. PLOS Genetics. 2015;11(7):e1005345.

46.   Manfredi MG, Ecsedy JA, Meetze KA, Balani SK, Burenkova O, Chen W, et al. Antitumor activity of MLN8054, an orally active small-molecule inhibitor of Aurora A kinase. Proc Natl Acad Sci U S A. 2007;104(10):4106-11.

47.   Booth DG, Hood FE, Prior IA, Royle SJ. A TACC3/ch-TOG/clathrin complex stabilises kinetochore fibres by inter-microtubule bridging. Embo j. 2011;30(5):906-19.

48.   Roskoski R. STI-571: an anticancer protein-tyrosine kinase inhibitor. Biochemical and Biophysical Research Communications. 2003;309(4):709-17.

49.   Roskoski R, Jr. ERK1/2 MAP kinases: structure, function, and regulation. Pharmacol Res. 2012;66(2):105-43.

50.   Bishop AC, Ubersax JA, Petsch DT, Matheos DP, Gray NS, Blethrow J, et al. A chemical switch for inhibitor-sensitive alleles of any protein kinase. Nature. 2000;407(6802):395-401.

51.   Treiber Daniel K, Shah Neil P. Ins and Outs of Kinase DFG Motifs. Chemistry & Biology. 2013;20(6):745-6.

52.   Monod J, Changeux J-P, Jacob F. Allosteric proteins and cellular control systems. Journal of Molecular Biology. 1963;6(4):306-29.

53.   Eglen R, Reisine T. Drug discovery and the human kinome: Recent trends. Pharmacology & Therapeutics. 2011;130(2):144-56.

54.   Zhang J, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. Nat Rev Cancer. 2009;9(1):28-39.

55.   Dodson CA, Kosmopoulou M, Richards MW, Atrash B, Bavetsias V, Blagg J, et al. Crystal structure of an Aurora-A mutant that mimics Aurora-B bound to MLN8054: insights into selectivity and drug design. Biochem J. 2010;427(1):19-28.

56.   Ung PM, Rahman R, Schlessinger A. Redefining the Protein Kinase Conformational Space with Machine Learning. Cell Chem Biol. 2018;25(7):916-24.e2.

57.   Lake EW, Muretta JM, Thompson AR, Rasmussen DM, Majumdar A, Faber EB, et al. Quantitative conformational profiling of kinase inhibitors reveals origins of selectivity for Aurora kinase activation states. Proceedings of the National Academy of Sciences. 2018;115(51):E11894.

58.   Nair JS, Ho AL, Schwartz GK. The induction of polyploidy or apoptosis by the Aurora A kinase inhibitor MK8745 is p53-dependent. Cell Cycle. 2012;11(4):807-17.

59.   Du J, Yan L, Torres R, Gong X, Bian H, Marugán C, et al. Aurora A-Selective Inhibitor LY3295668 Leads to Dominant Mitotic Arrest, Apoptosis in Cancer Cells, and Shows Potent Preclinical Antitumor Efficacy. Mol Cancer Ther. 2019;18(12):2207-19.

60.   Shionome Y, Lin WH, Shiao HY, Hsieh HP, Hsu JT, Ouchi T. A novel aurora-A inhibitor, BPR1K0609S1, sensitizes colorectal tumor cells to 5-fluorofracil (5-FU) treatment. Int J Biol Sci. 2013;9(4):403-11.

61.   Zheng FM, Long ZJ, Hou ZJ, Luo Y, Xu LZ, Xia JL, et al. A novel small molecule aurora kinase inhibitor attenuates breast tumor-initiating cells and overcomes drug resistance. Mol Cancer Ther. 2014;13(8):1991-2003.

62.   Ndolo KM, Park KR, Lee HJ, Yoon KB, Kim YC, Han SY. Characterization of the Indirubin Derivative LDD970 as a Small Molecule Aurora Kinase A Inhibitor in Human Colorectal Cancer Cells. Immune Netw. 2017;17(2):110-5.

63.   Lin Y, Richards FM, Krippendorff BF, Bramhall JL, Harrington JA, Bapiro TE, et al. Paclitaxel and CYC3, an aurora kinase A inhibitor, synergise in pancreatic cancer cells but not bone marrow precursor cells. Br J Cancer. 2012;107(10):1692-701.

64.   Bavetsias V, Linardopoulos S. Aurora Kinase Inhibitors: Current Status and Outlook. Front Oncol. 2015;5:278-.

65.   Manfredi MG, Ecsedy JA, Chakravarty A, Silverman L, Zhang M, Hoar KM, et al. Characterization of Alisertib (MLN8237), an Investigational Small-Molecule Inhibitor of Aurora A Kinase Using Novel In Vivo Pharmacodynamic Assays. Clinical Cancer Research. 2011;17(24):7614.

66.   Macarulla T, Cervantes A, Elez E, Rodríguez-Braun E, Baselga J, Roselló S, et al. Phase I study of the selective Aurora A kinase inhibitor MLN8054 in patients with advanced solid tumors: safety, pharmacokinetics, and pharmacodynamics. Mol Cancer Ther. 2010;9(10):2844-52.

67.   Dees EC, Cohen RB, von Mehren M, Stinchcombe TE, Liu H, Venkatakrishnan K, et al. Phase I Study of Aurora A Kinase Inhibitor MLN8237 in Advanced Solid Tumors: Safety, Pharmacokinetics, Pharmacodynamics, and Bioavailability of Two Oral Formulations. Clinical Cancer Research. 2012;18(17):4775.

68.   Matulonis UA, Sharma S, Ghamande S, Gordon MS, Del Prete SA, Ray-Coquard I, et al. Phase II study of MLN8237 (alisertib), an investigational Aurora A kinase inhibitor, in patients with platinum-resistant or -refractory epithelial ovarian, fallopian tube, or primary peritoneal carcinoma. Gynecologic Oncology. 2012;127(1):63-9.

69.   O'Connor OA, Özcan M, Jacobsen ED, Roncero JM, Trotman J, Demeter J, et al. Randomized Phase III Study of Alisertib or Investigator's Choice (Selected Single Agent) in Patients With Relapsed or Refractory Peripheral T-Cell Lymphoma. J Clin Oncol. 2019;37(8):613-23.

70.   Bebbington D, Binch H, Charrier J-D, Everitt S, Fraysse D, Golec J, et al. The discovery of the potent aurora inhibitor MK-0457 (VX-680). Bioorganic & Medicinal Chemistry Letters. 2009;19(13):3586-92.

71.   Tyler RK, Shpiro N, Marquez R, Eyers PA. VX-680 Inhibits Aurora A and Aurora B Kinase Activity in Human Cells. Cell Cycle. 2007;6(22):2846-54.

72.   Harrington EA, Bebbington D, Moore J, Rasmussen RK, Ajose-Adeogun AO, Nakayama T, et al. VX-680, a potent and selective small-molecule inhibitor of the Aurora kinases, suppresses tumor growth in vivo. Nat Med. 2004;10(3):262-7.

73.   Tentler JJ, Bradshaw-Pierce EL, Serkova NJ, Hasebroock KM, Pitts TM, Diamond JR, et al. Assessment of the in vivo antitumor effects of ENMD-2076, a novel multitargeted kinase inhibitor, against primary and cell line-derived human colorectal cancer xenograft models. Clin Cancer Res. 2010;16(11):2989-98.

74.   Fletcher GC, Brokx RD, Denny TA, Hembrough TA, Plum SM, Fogler WE, et al. ENMD-2076 is an orally active kinase inhibitor with antiangiogenic and antiproliferative mechanisms of action. Mol Cancer Ther. 2011;10(1):126-37.

75.   Wang X, Sinn AL, Pollok K, Sandusky G, Zhang S, Chen L, et al. Preclinical activity of a novel multiple tyrosine kinase and aurora kinase inhibitor, ENMD-2076, against multiple myeloma. Br J Haematol. 2010;150(3):313-25.

76.   Ionkina AA, Tentler JJ, Kim J, Capasso A, Pitts TM, Ryall KA, et al. Efficacy and Molecular Mechanisms of Differentiated Response to the Aurora and Angiogenic Kinase Inhibitor ENMD-2076 in Preclinical Models of p53-Mutated Triple-Negative Breast Cancer. Front Oncol. 2017;7:94.

77.   Diamond JR, Bastos BR, Hansen RJ, Gustafson DL, Eckhardt SG, Kwak EL, et al. Phase I safety, pharmacokinetic, and pharmacodynamic study of ENMD-2076, a novel angiogenic and Aurora kinase inhibitor, in patients with advanced solid tumors. Clin Cancer Res. 2011;17(4):849-60.

78.   Veitch Z, Zer A, Loong H, Salah S, Masood M, Gupta A, et al. A phase II study of ENMD-2076 in advanced soft tissue sarcoma (STS). Sci Rep. 2019;9(1):7390-.

79.   Hardman G, Perkins S, Brownridge PJ, Clarke CJ, Byrne DP, Campbell AE, et al. Strong anion exchange-mediated phosphoproteomics reveals extensive human non-canonical phosphorylation. The EMBO Journal. 2019;38(21):e100847.

80.   Kabbara S, Hérivaux A, Dugé de Bernonville T, Courdavault V, Clastre M, Gastebois A, et al. Diversity and Evolution of Sensor Histidine Kinases in Eukaryotes. Genome Biology and Evolution. 2018;11(1):86-108.

81.   Boyer PD, Deluca M, Ebner KE, Hultquist DE, Peter JB. Identification of phosphohistidine in digests from a probable intermediate of oxidative phosphorylation. J Biol Chem. 1962;237:Pc3306-pc8.

82.   Panda S, Srivastava S, Li Z, Vaeth M, Fuhs SR, Hunter T, et al. Identification of PGAM5 as a Mammalian Protein Histidine Phosphatase that Plays a Central Role to Negatively Regulate CD4(+) T Cells. Mol Cell. 2016;63(3):457-69.

83.   Tomlinson LJ, Clubbs Coldron AKM, Eyers PA, Eyers CE. Determination of Phosphohistidine Stoichiometry in Histidine Kinases by Intact Mass Spectrometry. In: Eyers CE, editor. Histidine Phosphorylation: Methods and Protocols. New York, NY: Springer US; 2020. p. 83-91.

84.   Igo MM, Ninfa AJ, Stock JB, Silhavy TJ. Phosphorylation and dephosphorylation of a bacterial transcriptional activator by a transmembrane receptor. Genes Dev. 1989;3(11):1725-34.

85.   Attwood PV, Muimo R. The actions of NME1/NDPK-A and NME2/NDPK-B as protein kinases. Laboratory Investigation. 2018;98(3):283-90.

86.   Attwood PV, Piggott MJ, Zu XL, Besant PG. Focus on phosphohistidine. Amino Acids. 2007;32(1):145-56.

87.   Puttick J, Baker EN, Delbaere LT. Histidine phosphorylation in biological systems. Biochim Biophys Acta. 2008;1784(1):100-5.

88.   Attwood PV. P-N bond protein phosphatases. Biochim Biophys Acta. 2013;1834(1):470-8.

89.   Besant PG, Attwood PV. Detection of a mammalian histone H4 kinase that has yeast histidine kinase-like enzymic activity. Int J Biochem Cell Biol. 2000;32(2):243-53.

90.   Fujitaki JM, Smith RA. Techniques in the detection and characterization of phosphoramidate-containing proteins. Methods Enzymol. 1984;107:23-36.

91.   Medzihradszky KF, Phillipps NJ, Senderowicz L, Wang P, Turck CW. Synthesis and characterization of histidine-phosphorylated peptides. Protein Sci. 1997;6(7):1405-11.

92.   Wind M, Wegener A, Kellner R, Lehmann WD. Analysis of CheA Histidine Phosphorylation and Its Influence on Protein Stability by High-Resolution Element and Electrospray Mass Spectrometry. Analytical Chemistry. 2005;77(7):1957-62.

93.   Stafford LJ, Vaidya KS, Welch DR. Metastasis suppressors genes in cancer. Int J Biochem Cell Biol. 2008;40(5):874-91.

94.   Freije JM, Blay P, MacDonald NJ, Manrow RE, Steeg PS. Site-directed mutation of Nm23-H1. Mutations lacking motility suppressive capacity upon transfection are deficient in histidine-dependent protein phosphotransferase pathways in vitro. J Biol Chem. 1997;272(9):5525-32.

95.   Steeg PS, Bevilacqua G, Kopper L, Thorgeirsson UP, Talmadge JE, Liotta LA, et al. Evidence for a novel gene associated with low tumor metastatic potential. J Natl Cancer Inst. 1988;80(3):200-4.

96.   Adam K, Lesperance J, Hunter T, Zage PE. The Potential Functional Roles of NME1 Histidine Kinase Activity in Neuroblastoma Pathogenesis. Int J Mol Sci. 2020;21(9).

97.   Boissan M, Dabernat S, Peuchant E, Schlattner U, Lascu I, Lacombe ML. The mammalian Nm23/NDPK family: from metastasis control to cilia movement. Mol Cell Biochem. 2009;329(1-2):51-62.

98.   Boissan M, Schlattner U, Lacombe M-L. The NDPK/NME superfamily: state of the art. Laboratory Investigation. 2018;98(2):164-74.

99.   Graves PR, Haystead TAJ. Molecular biologist's guide to proteomics. Microbiol Mol Biol Rev. 2002;66(1):39-63.

100.   Huang L-C, Ross KE, Baffi TR, Drabkin H, Kochut KJ, Ruan Z, et al. Integrative annotation and knowledge discovery of kinase post-translational modifications and cancer-associated mutations through federated protein ontologies and resources. Sci Rep. 2018;8(1):6518.

101.   Chait BT. Mass Spectrometry: Bottom-Up or Top-Down? Science. 2006;314(5796):65-6.

102.   Resing KA, Ahn NG. Proteomics strategies for protein identification. FEBS Lett. 2005;579(4):885-9.

103.   Pitt JJ. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. Clin Biochem Rev. 2009;30(1):19-34.

104.   Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem. 2001;70:437-73.

105.   Wickremsinhe ER, Singh G, Ackermann BL, Gillespie TA, Chaudhary AK. A review of nanoelectrospray ionization applications for drug metabolism and pharmacokinetics. Curr Drug Metab. 2006;7(8):913-28.

106.   Manisali I, Chen DDY, Schneider BB. Electrospray ionization source geometry for mass spectrometry: past, present, and future. TrAC Trends in Analytical Chemistry. 2006;25(3):243-56.

107.	Haag AM. Mass Analyzers and Mass Spectrometers. In: Mirzaei H, Carrasco M, editors. Modern Proteomics – Sample Preparation, Analysis and Practical Applications. Cham: Springer International Publishing; 2016. p. 157-69.

108.	Diedrich JK, Pinto AFM, Yates JR, 3rd. Energy dependence of HCD on peptide fragmentation: stepped collisional energy finds the sweet spot. Journal of the American Society for Mass Spectrometry. 2013;24(11):1690-9.

109.	Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M. Higher-energy C-trap dissociation for peptide modification analysis. Nat Methods. 2007;4(9):709-12.

110.	Jedrychowski MP, Huttlin EL, Haas W, Sowa ME, Rad R, Gygi SP. Evaluation of HCD- and CID-type fragmentation within their respective detection platforms for murine phosphoproteomics. Mol Cell Proteomics. 2011;10(12):M111.009910.

111.	Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci U S A. 2004;101(26):9528-33.

112.	Swaney DL, McAlister GC, Wirtala M, Schwartz JC, Syka JE, Coon JJ. Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors. Anal Chem. 2007;79(2):477-85.

113.	Boersema PJ, Mohammed S, Heck AJ. Phosphopeptide fragmentation and analysis by mass spectrometry. J Mass Spectrom. 2009;44(6):861-78.

114.	Yu Q, Wang B, Chen Z, Urabe G, Glover MS, Shi X, et al. Electron-Transfer/Higher-Energy Collision Dissociation (EThcD)-Enabled Intact Glycopeptide/Glycoproteome Characterization. Journal of the American Society for Mass Spectrometry. 2017;28(9):1751-64.

115.	Senko MW, Remes PM, Canterbury JD, Mathur R, Song Q, Eliuk SM, et al. Novel parallelized quadrupole/linear ion trap/Orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. Anal Chem. 2013;85(24):11710-4.

116.	Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R. The Orbitrap: a new mass spectrometer. J Mass Spectrom. 2005;40(4):430-43.

117.	Scigelova M, Hornshaw M, Giannakopulos A, Makarov A. Fourier transform mass spectrometry. Mol Cell Proteomics. 2011;10(7):M111.009431-M111.

118.	Gutmark R. Wave Front Analysis. In: Schmidt-Erfurth U, Kohnen T, editors. Encyclopedia of Ophthalmology. Berlin, Heidelberg: Springer Berlin Heidelberg; 2018. p. 1917-9.

119.	Li X, Jiang G, Luo C, Xu F, Wang Y, Ding L, et al. Ion Trap Array Mass Analyzer: Structure and Performance. Analytical Chemistry. 2009;81(12):4840-6.

120.	Douglas DJ, Frank AJ, Mao D. Linear ion traps in mass spectrometry. Mass Spectrometry Reviews. 2005;24(1):1-29.

121.	Schwartz JC, Senko MW, Syka JEP. A two-dimensional quadrupole ion trap mass spectrometer. Journal of the American Society for Mass Spectrometry. 2002;13(6):659-69.

122.	Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999;20(18):3551-67.

123.	Aggarwal S, Yadav AK. False Discovery Rate Estimation in Proteomics. Methods Mol Biol. 2016;1362:119-28.

124.	Taus T, Köcher T, Pichler P, Paschke C, Schmidt A, Henrich C, et al. Universal and confident phosphorylation site localization using phosphoRS. J Proteome Res. 2011;10(12):5354-62.

125.   Leney AC, Heck AJR. Native Mass Spectrometry: What is in the Name? Journal of the American Society for Mass Spectrometry. 2017;28(1):5-13.

126.   Heck AJR. Native mass spectrometry: a bridge between interactomics and structural biology. Nature Methods. 2008;5(11):927-33.

127.   Lanucara F, Holman SW, Gray CJ, Eyers CE. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nature Chemistry. 2014;6(4):281-94.

128.   Beveridge R, Migas LG, Das RK, Pappu RV, Kriwacki RW, Barran PE. Ion Mobility Mass Spectrometry Uncovers the Impact of the Patterning of Oppositely Charged Residues on the Conformational Distributions of Intrinsically Disordered Proteins. Journal of the American Chemical Society. 2019;141(12):4908-18.

129.   Hilton GR, Benesch JLP. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. J R Soc Interface. 2012;9(70):801-16.

130.   Juraschek R, Dülcks T, Karas M. Nanoelectrospray--more than just a minimized-flow electrospray ionization source. J Am Soc Mass Spectrom. 1999;10(4):300-8.

131.   Liu Y, Huynh DT, Yeates TO. A 3.8 Å resolution cryo-EM structure of a small protein bound to an imaging scaffold. Nat Commun. 2019;10(1):1864.

132.   Benjin X, Ling L. Developments, applications, and prospects of cryo-electron microscopy. Protein science : a publication of the Protein Society. 2020;29(4):872-82.

133.   Acharya KR, Lloyd MD. The advantages and limitations of protein crystal structures. Trends in Pharmacological Sciences. 2005;26(1):10-4.

134.   Silva Elipe MV. Advantages and disadvantages of nuclear magnetic resonance spectroscopy as a hyphenated technique. Analytica Chimica Acta. 2003;497(1):1-25.

135.   Hewitt D, Marklund E, Scott DJ, Robinson CV, Borysik AJ. A Hydrodynamic Comparison of Solution and Gas Phase Proteins and Their Complexes. The Journal of Physical Chemistry B. 2014;118(29):8489-95.

136.   Shvartsburg AA, Smith RD. Fundamentals of Traveling Wave Ion Mobility Spectrometry. Analytical Chemistry. 2008;80(24):9689-99.

137.   Loo JA, He JX, Cody WL. Higher Order Structure in the Gas Phase Reflects Solution Structure. Journal of the American Chemical Society. 1998;120(18):4542-3.

138.   Jurneczko E, Barran PE. How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. Analyst. 2011;136(1):20-8.

139.   Schennach M, Breuker K. Proteins with highly similar native folds can show vastly dissimilar folding behavior when desolvated. Angewandte Chemie (International ed in English). 2014;53(1):164-8.

140.   Wyttenbach T, Bowers MT. Structural stability from solution to the gas phase: native solution structure of ubiquitin survives analysis in a solvent-free ion mobility-mass spectrometry environment. J Phys Chem B. 2011;115(42):12266-75.

141.   Giles K, Ujma J, Wildgoose J, Pringle S, Richardson K, Langridge D, et al. A Cyclic Ion Mobility-Mass Spectrometry System. Analytical Chemistry. 2019;91(13):8564-73.

142.   Dodds JN, Baker ES. Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. Journal of The American Society for Mass Spectrometry. 2019;30(11):2185-95.

143.   Kinetic Theory of Mobility and Diffusion: Sections 5.1 – 5.2. Transport Properties of Ions in Gases1988. p. 137-93.

144.   Thomas SN. Chapter 10 - Mass spectrometry. In: Clarke W, Marzinke MA, editors. Contemporary Practice in Clinical Chemistry (Fourth Edition): Academic Press; 2019. p. 171-85.

145.   Kaklamanos G, Aprea E, Theodoridis G. 11 - Mass spectrometry: principles and instrumentation. In: Pico Y, editor. Chemical Analysis of Food (Second Edition): Academic Press; 2020. p. 525-52.

146.   Eyers CE, Vonderach M, Ferries S, Jeacock K, Eyers PA. Understanding protein–drug interactions using ion mobility–mass spectrometry. Curr Opin Chem Biol. 2018;42:167-76.

147.   Paglia G, Angel P, Williams JP, Richardson K, Olivos HJ, Thompson JW, et al. Ion Mobility-Derived Collision Cross Section As an Additional Measure for Lipid Fingerprinting and Identification. Analytical Chemistry. 2015;87(2):1137-44.

148.   Righetti L, Bergmann A, Galaverna G, Rolfsson O, Paglia G, Dall'Asta C. Ion mobility-derived collision cross section database: Application to mycotoxin analysis. Analytica Chimica Acta. 2018;1014:50-7.

149.   Nielson FF, Colby SM, Thomas DG, Renslow RS, Metz TO. Exploring the Impacts of Conformer Selection Methods on Ion Mobility Collision Cross Section Predictions. Analytical Chemistry. 2021;93(8):3830-8.

150.   Bush MF, Hall Z, Giles K, Hoyes J, Robinson CV, Ruotolo BT. Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. Analytical Chemistry. 2010;82(22):9557-65.

151.   Marklund Erik G, Degiacomi Matteo T, Robinson Carol V, Baldwin Andrew J, Benesch Justin LP. Collision Cross Sections for Structural Proteomics. Structure. 2015;23(4):791-9.

152.   Allison TM, Reading E, Liko I, Baldwin AJ, Laganowsky A, Robinson CV. Quantifying the stabilizing effects of protein-ligand interactions in the gas phase. Nat Commun. 2015;6:8551-.

153.   Dixit SM, Polasky DA, Ruotolo BT. Collision induced unfolding of isolated proteins in the gas phase: past, present, and future. Curr Opin Chem Biol. 2018;42:93-100.

154.   Ruotolo BT, Hyung S-J, Robinson PM, Giles K, Bateman RH, Robinson CV. Ion Mobility–Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes. Angewandte Chemie International Edition. 2007;46(42):8001-4.

155.   Janson G, Paiardini A. PyMod 3: a complete suite for structural bioinformatics in PyMOL. Bioinformatics. 2020.

156.   Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol. 1993;234(3):779-815.

157.   Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods. 2017;14(1):71-3.

158.   Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. Journal of Computational Chemistry. 2005;26(16):1781-802.

159.   Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: A web-based graphical user interface for CHARMM. Journal of Computational Chemistry. 2008;29(11):1859-65.

160.   Karanicolas J, Brooks CL. Improved Gō-like Models Demonstrate the Robustness of Protein Folding Mechanisms Towards Non-native Interactions. Journal of Molecular Biology. 2003;334(2):309-25.

161.   Karanicolas J, Brooks CL. The structural basis for biphasic kinetics in the folding of the WW domain from a formin-binding protein: Lessons for protein design? Proceedings of the National Academy of Sciences. 2003;100(7):3954-9.

162.   Brooks BR, Brooks III CL, Mackerell Jr. AD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: The biomolecular simulation program. Journal of Computational Chemistry. 2009;30(10):1545-614.

163.   Seeber M, Felline A, Raimondi F, Muff S, Friedman R, Rao F, et al. Wordom: A user-friendly program for the analysis of molecular structures, trajectories, and free energy surfaces. Journal of Computational Chemistry. 2011;32(6):1183-94.

164.   Karanicolas J, Brooks CL, 3rd. Improved Gō-like models demonstrate the robustness of protein folding mechanisms towards non-native interactions. J Mol Biol. 2003;334(2):309-25.

165.   Marklund EG, Degiacomi MT, Robinson CV, Baldwin AJ, Benesch JL. Collision cross sections for structural proteomics. Structure. 2015;23(4):791-9.

166.   Bolanos-Garcia VM. Aurora kinases. Int J Biochem Cell Biol. 2005;37(8):1572-7.

167.   Hanks SK, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. Faseb j. 1995;9(8):576-96.

168.   de Groot CO, Hsia JE, Anzola JV, Motamedi A, Yoon M, Wong YL, et al. A Cell Biologist's Field Guide to Aurora Kinase Inhibitors. Front Oncol. 2015;5(285).

169.   Eyers PA, Erikson E, Chen LG, Maller JL. A novel mechanism for activation of the protein kinase Aurora A. Curr Biol. 2003;13(8):691-7.

170.   Dou H, Buetow L, Hock A, Sibbet GJ, Vousden KH, Huang DT. Structural basis for autoinhibition and phosphorylation-dependent activation of c-Cbl. Nature Structural & Molecular Biology. 2012;19(2):184-92.

171.   Ma NJ-L, Stern DF. Regulation of the Rad53 protein kinase in signal amplification by oligomer assembly and disassembly. Cell Cycle. 2008;7(6):808-17.

172.   Hashimoto K, Nishi H, Bryant S, Panchenko AR. Caught in self-interaction: evolutionary and functional mechanisms of protein homooligomerization. Physical Biology. 2011;8(3):035007.

173.   Facilities AoBR. A Database of Post Post Translational Modifications 2021 [

174.   Ohashi S, Sakashita G, Ban R, Nagasawa M, Matsuzaki H, Murata Y, et al. Phospho-regulation of human protein kinase Aurora-A: analysis using anti-phospho-Thr288 monoclonal antibodies. Oncogene. 2006;25(59):7691-702.

175.   Ferrari S, Marin O, Pagano MA, Meggio F, Hess D, El-Shemerly M, et al. Aurora-A site specificity: a study with synthetic peptide substrates. Biochem J. 2005;390(Pt 1):293-302.

176.   Haydon CE, Eyers PA, Aveline-Wolf LD, Resing KA, Maller JL, Ahn NG. Identification of Novel Phosphorylation Sites on Xenopus laevis Aurora A and Analysis of Phosphopeptide Enrichment by Immobilized Metal-affinity Chromatography *. Molecular & Cellular Proteomics. 2003;2(10):1055-67.

177.   Clarke DJ, Campopiano DJ. Desalting large protein complexes during native electrospray mass spectrometry by addition of amino acids to the working solution. Analyst. 2015;140(8):2679-86.

178.   Pan P, McLuckey SA. The effect of small cations on the positive electrospray responses of proteins at low pH. Anal Chem. 2003;75(20):5468-74.

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

180.   Konermann L. Addressing a Common Misconception: Ammonium Acetate as Neutral pH "Buffer" for Native Electrospray Mass Spectrometry. Journal of The American Society for Mass Spectrometry. 2017;28(9):1827-35.

181.   Verkerk UH, Kebarle P. Ion-ion and ion-molecule reactions at the surface of proteins produced by nanospray. Information on the number of acidic residues and control of the number of ionized acidic and basic residues. J Am Soc Mass Spectrom. 2005;16(8):1325-41.

182.   Xia Z, DeGrandchamp JB, Williams ER. Native mass spectrometry beyond ammonium acetate: effects of nonvolatile salts on protein stability and structure. Analyst. 2019;144(8):2565-73.

183.   Kaushik JK, Bhat R. Why Is Trehalose an Exceptional Protein Stabilizer?: AN ANALYSIS OF THE THERMAL STABILITY OF PROTEINS IN THE PRESENCE OF THE COMPATIBLE OSMOLYTE TREHALOSE *. Journal of Biological Chemistry. 2003;278(29):26458-65.

184.   Kune C, Far J, De Pauw E. Accurate Drift Time Determination by Traveling Wave Ion Mobility Spectrometry: The Concept of the Diffusion Calibration. Anal Chem. 2016;88(23):11639-46.

185.   Gabelica V, Shvartsburg AA, Afonso C, Barran P, Benesch JLP, Bleiholder C, et al. Recommendations for reporting ion mobility Mass Spectrometry measurements. Mass Spectrometry Reviews. 2019;38(3):291-320.

186.   Byrne DP, Vonderach M, Ferries S, Brownridge PJ, Eyers CE, Eyers PA. cAMP-dependent protein kinase (PKA) complexes probed by complementary differential scanning fluorimetry and ion mobility–mass spectrometry. Biochemical Journal. 2016;473(19):3159-75.

187.   McCabe JW, Mallis CS, Kocurek KI, Poltash ML, Shirzadeh M, Hebert MJ, et al. First-Principles Collision Cross Section Measurements of Large Proteins and Protein Complexes. Analytical Chemistry. 2020;92(16):11155-63.

188.   Möbitz H. The ABC of protein kinase conformations. Biochim Biophys Acta. 2015;1854(10 Pt B):1555-66.

189.   Deng L, Broom A, Kitova EN, Richards MR, Zheng RB, Shoemaker GK, et al. Kinetic stability of the streptavidin-biotin interaction enhanced in the gas phase. J Am Chem Soc. 2012;134(40):16586-96.

190.   Stepczynska A, Lauber K, Engels IH, Janssen O, Kabelitz D, Wesselborg S, et al. Staurosporine and conventional anticancer drugs induce overlapping, yet distinct pathways of apoptosis and caspase activation. Oncogene. 2001;20(10):1193-202.

191.   Tanramluk D, Schreyer A, Pitt WR, Blundell TL. On the origins of enzyme inhibitor selectivity and promiscuity: a case study of protein kinase binding to staurosporine. Chem Biol Drug Des. 2009;74(1):16-24.

192.   Onufriev AV, Alexov E. Protonation and pK changes in protein-ligand binding. Q Rev Biophys. 2013;46(2):181-209.

193.   May JC, Morris CB, McLean JA. Ion Mobility Collision Cross Section Compendium. Analytical chemistry. 2017;89(2):1032-44.

194.   Hopper JT, Oldham NJ. Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: the effect of ligand binding on conformational stability. J Am Soc Mass Spectrom. 2009;20(10):1851-8.

195.   Eschweiler JD, Rabuck-Gibbons JN, Tian Y, Ruotolo BT. CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions. Analytical Chemistry. 2015;87(22):11516-22.

196.   Boissan M, Montagnac G, Shen Q, Griparic L, Guitton J, Romao M, et al. Membrane trafficking. Nucleoside diphosphate kinases fuel dynamin superfamily proteins with GTP for membrane remodeling. Science. 2014;344(6191):1510-5.

197.   Khan I, Steeg PS. The relationship of NM23 (NME) metastasis suppressor histidine phosphorylation to its nucleoside diphosphate kinase, histidine protein kinase and motility suppression activities. Oncotarget. 2017;9(12):10185-202.

198.   Palacios F, Schweitzer JK, Boshans RL, D'Souza-Schorey C. ARF6-GTP recruits Nm23-H1 to facilitate dynamin-mediated endocytosis during adherens junctions disassembly. Nat Cell Biol. 2002;4(12):929-36.

199.   Fuhs Stephen R, Meisenhelder J, Aslanian A, Ma L, Zagorska A, Stankova M, et al. Monoclonal 1- and 3-Phosphohistidine Antibodies: New Tools to Study Histidine Phosphorylation. Cell. 2015;162(1):198-210.

200.   Kee J-M, Oslund RC, Perlman DH, Muir TW. A pan-specific antibody for direct detection of protein histidine phosphorylation. Nature Chemical Biology. 2013;9(7):416-21.

201.   Chen Y, Gallois-Montbrun S, Schneider B, Véron M, Moréra S, Deville-Bonne D, et al. Nucleotide binding to nucleoside diphosphate kinases: X-ray structure of human NDPK-A in complex with ADP and comparison to protein kinases. J Mol Biol. 2003;332(4):915-26.

202.   Yu BYK, Tossounian M-A, Hristov SD, Lawrence R, Arora P, Tsuchiya Y, et al. Regulation of metastasis suppressor NME1 by a key metabolic cofactor coenzyme A. Redox Biology. 2021;44:101978.

203.   Georgescauld F, Song Y, Dautant A. Structure, Folding and Stability of Nucleoside Diphosphate Kinases. International journal of molecular sciences. 2020;21(18):6779.

204.   Tepper AD, Dammann H, Bominaar AA, Véron M. Investigation of the active site and the conformational stability of nucleoside diphosphate kinase by site-directed mutagenesis. J Biol Chem. 1994;269(51):32175-80.

205.   Kim MS, Jeong J, Jeong J, Shin DH, Lee KJ. Structure of Nm23-H1 under oxidative conditions. Acta Crystallogr D Biol Crystallogr. 2013;69(Pt 4):669-80.

206.   Reddie KG, Carroll KS. Expanding the functional diversity of proteins through cysteine oxidation. Curr Opin Chem Biol. 2008;12(6):746-54.

207.   Lee E, Jeong J, Kim SE, Song EJ, Kang SW, Lee KJ. Multiple functions of Nm23-H1 are regulated by oxido-reduction system. PLoS One. 2009;4(11):e7949.

208.   Kim MS, Jeong J, Lee KJ, Shin DH. A preliminary X-ray study of human nucleoside diphosphate kinase A under oxidative conditions. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2010;66(Pt 11):1490-2.

209.   Marty MT, Baldwin AJ, Marklund EG, Hochberg GKA, Benesch JLP, Robinson CV. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. Analytical chemistry. 2015;87(8):4370-6.

210.   You D-J, Park CR, Mander S, Ahn C, Seong JY, Hwang J-I. Characterization of Functional Domains in NME1L Regulation of NF-κB Signaling. Mol Cells. 2016;39(5):403-9.

211.   SelleckChem. FDA-approved Drug Library 2021 [Available from: https://www.selleckchem.com/screening/fda-approved-drug-library.html.

212.   Santini MT, Morelli G, Fattorossi A, Malorni W, Rainaldi G, Indovina PL. The oxidizing agent menadione induces an increase in the intracellular molecular oxygen concentration in K562 and A431 cells: direct measurement using the new paramagnetic EPR probe fusinite. Free Radic Biol Med. 1996;20(7):915-24.

213.   Adam K, Ning J, Reina J, Hunter T. NME/NM23/NDPK and Histidine Phosphorylation. International Journal of Molecular Sciences. 2020;21(16):5848.

214.   Bayliss R, Conti, E. Structure of Human Aurora-A 122-403 phosphorylated on Thr287, Thr288 2003 [Available from: https://www.rcsb.org/structure/1OL7.

215.   Bayliss R, Conti, E. Structure of Aurora-A 122-403, phosphorylated on Thr287, Thr288 and bound to TPX2 1-43 PDB2003 [Available from: https://www.rcsb.org/structure/1OL5.

216.   Lawrie AM, Noble ME, Tunnah P, Brown NR, Johnson LN, Endicott JA. Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. Nat Struct Biol. 1997;4(10):796-801.

217.   Dranchak P, MacArthur R, Guha R, Zuercher WJ, Drewry DH, Auld DS, et al. Profile of the GSK published protein kinase inhibitor set across ATP-dependent and-independent luciferases: implications for reporter-gene assays. PloS one. 2013;8(3):e57888-e.