Characterisation of the Prokaryotic Elongation Factor-Tu with Structural Mass Spectrometry

by

Cameron Baines, MRes

Thesis submitted to the University of Nottingham for the degree
of Doctor of Philosophy

September 2024

Abstract

The elongation factor EF-Tu is an essential and ubiquitous component of prokaryotic mRNA translation. It is directly responsible for the transport of aminoacyl-tRNA to the actively translating ribosome. EF-Tu is a translational GTPase, classically thought to have two conformational extremes induced by the binding of either GDP or GTP, the *open* and *closed* conformations respectively. It is now hypothesised that, in solution, EF-Tu does not adopt these binary conformations and instead samples a wide range of conformations, with the average becoming more *open* or *closed* depending on the bound nucleotide. As an essential prokaryotic protein, EF-Tu presents itself as a desirable antimicrobial target. The elfamycins are a group of bacteriostatic antimicrobials specifically targeting EF-Tu, acting predominantly in one of two ways. Kirromycin-like elfamycins lock EF-Tu in the *closed* confirmation irrespective of GTP hydrolysis, preventing EF-Tu dissociating from the ribosome. Pulvomycin-like elfamycins simply prevent the binding of tRNA to EF-Tu, acting as a competitive inhibitor.

In this thesis the behaviour of the clinically relevant *Acinetobacter baumannii* EF-Tu in complex with two contrasting elfamycins, alongside a recently identified prolyl hydroxylase found to target EF-Tu is described. In Chapter 3, the binding of GE2270A, a pulvomycin-like elfamycin, to EF-Tu was explored using native mass spectrometry and carbene footprinting, a covalent protein labelling technique. The binding site of GE2270A was identified, alongside some tentative GE2270A-induced subtle conformational changes. In Chapter 4, the binding of, and global conformational change induced by, the kirromycin-like elfamycin enacyloxin IIa was examined using carbene footprinting. Alongside this, three enacyloxin IIa resistant EF-Tu mutants were produced and characterised using carbene footprinting and native mass spectrometry. In chapter 5, the interactions between an EF-Tu-specific prolyl hydroxylase and EF-Tu were mapped with carbene footprinting. Finally, chapter 6 demonstrated the use of native mass spectrometry to investigate novel enacyloxin IIa analogues and their binding to EF-Tu. Ion mobility spectrometry-mass spectrometry was used to characterise the stability of *Escherichia coli* EF-Tu in the gas phase. Here, the tertiary structure of EF-Tu was found to be stabilised by ligand binding through collisionally induced unfolding experiments. These results were corroborated with molecular dynamics simulations exploring the thermal stability of EF-Tu in *apo* and *holo* states.

Acknowledgements

Jumping into the world of biomolecular mass spectrometry from a background in micro and molecular biology was always going to present a challenge. A challenge that would have been insurmountable without the guidance and support of Professor Neil Oldham over the past four years. I would like to thank him for providing the opportunity to work on this project, both initially as a DTP rotation, and as my full PhD.

The members of the Oldham Group, past and present, namely Dr James Lloyd, Dr Charles Ducker, Dr Harry Taylor and Arppana Varughese, fostered a friendly and supportive environment, and for that I am truly grateful. A special thanks goes out to Dr James Lloyd for his training and support in carbene footprinting, and to Dr Charles Ducker for his training and support in protein expression. In that regard I would also like to thank Dr Daniel Scott and Professor Robert Layfield for allowing me to use laboratory space and equipment for molecular biology work.

I would like to thank Professor Greg Challis and his research group at the University of Warwick, namely Dr Dan Van, Dr Lona Alkhalaf, Matt Pledge and Mabilly Cox for the purification and supply of enacyloxin IIa and its analogues. I also thank Professor Chris Schofield and his group for providing *Pseudomonas* EF-Tu and PPHD samples. The members of the technical team within the School of Chemistry provided excellent instrument and equipment maintenance, without which some of this work may not have been possible. I thank Dr Jeddidiah Bellamy-Carter for his work creating PepFoot, various other Python scripts, and useful discussion over the years.

I thank my friends for their support, especially Dr Sam Dawson for his discussions about our projects, PhD life and everything in between. Finally, to my parents and family for their lifetime of love and support, I offer my heartfelt thanks. I wish to dedicate this thesis to them, and to those who are no longer with us.

“Looks convincing!” – Professor Neil Oldham

Contents

3. Mapping Interactions of EF-Tu and the Elfamycin GE2270A using Carbene Footprinting. 87

List of Figures

List of Tables

Abbreviations

AC	Alternating Current
AKHD	α-Ketoglutarate-Dependent Hydroxylase
ANOVA	Analysis of Variance
APCI	Atmospheric Pressure Chemical Ionisation
APPI	Atmospheric Pressure Photoionisation
BCA	Bicinchoninic Acid
BCC	*Burkholderia cepacia* Complex
BSA	Bovine Serum Albumin
CCS	Collisional Cross Section
CEM	Chain Ejection Model
CHARMM	Chemistry at Harvard Macromolecular Mechanics
CI	Chemical Ionisation
CID	Collision Induced Dissociation
CIU	Collision Induced Unfolding
CPU	Central Processing Unit
CRM	Charge Residue Model
DC	Direct Current
DDA	Data-Dependent Acquisition
DEC	Decarbamoylated
DESI	Desorption Electrospray Ionisation
DHCCA	Dihydroxycyclohexanecarboxylic Acid
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DTM	Diffuse Trajectory Method
DTT	Dithiothreitol
EB	Elution Buffer
ECD	Electron Capture Dissociation

EDTA	Ethylenediaminetetraacetic Acid
EF	Elongation Factor
EF-Tu	Elongation Factor Thermal Unstable
EI	Electron Ionisation
ENX	Enacyloxin IIa
ESI	Electrospray Ionisation
ETD	Electron Transfer Dissociation
FASTA	Fast-A
FFT	Fast Fourier Transformation
FMet	N-Formylmethionine
FPOP	Fast Photochemical Oxidation of Proteins
FRET	Fluorescence Resonance Energy Transfer
FT	Fourier Transform
FT-ICR	Fourier Transform-Ion Cyclotron Resonance
GAC	GTPase-Associated Centre
GAP	GTPase Activator Protein
GC	Gas Chromatography
GDP	Guanosine Diphosphate
GDPNP	5'-Guanylyl Imidodiphosphate
GEF	Guanosine Exchange Factor
GTP	Guanosine Triphosphate
HDX	Hydrogen-Deuterium Exchange
HEPES	4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid
HIF	Hypoxia-Inducible Factor
HPLC	High-Performance Liquid Chromatography
H-RPF	Hydroxyl Radical Protein Footprinting
ICR	Ion Cyclotron Resonance
IEM	Ion Evaporation Model
IF	Initiation Factor
IMS	Ion Mobility Spectrometry

IMS-MS	Ion Mobility Spectrometry-Mass Spectrometry
IPTG	Isopropyl ß-D-1-Thiogalactopyranoside
IRMPD	Infrared Multiphoton Dissociation
KIR	Kirromycin
KrF	Krypton Fluoride
LB	Lysogeny Broth
LC	Liquid Chromatography
LIT	Linear Ion Trap
MALDI	Matrix-Assisted Laser Desorption Ionisation
MD	Molecular Dynamics
mRNA	Messenger Ribonucleic Acid
MS	Mass Spectrometry
MWCO	Molecular Weight Cut Off
Nd:YAG	Neodymium-Doped Yttrium Aluminium Garnet
Nd:YLF	Neodymium-Doped Yttrium Lithium Fluoride
NEB	New England Biolabs
nESI	Nano-Electrospray Ionisation
Ni-NTA	Nickel-Nitrotriacetic acid
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NRP	Nonribosomal Peptide
PA	Projection Approximation
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PHD	Prolyl Hydroxylase Domain
PKS	Polyketide Synthase
PPHD	P. putida Prolyl Hydroxylase
PPHD	Putative Prolyl Hydroxylase Domain
Q-TOF	Quadrupole Time-Of-Flight
RF	Release Factor

RF	Radio Frequency
RMSD	Root Mean Square Deviation
RNA	Ribonucleic Acid
RP-HPLC	Reverse-Phase High Performance Liquid Chromatography
RPM	Revolutions Per Minute
rRNA	Ribosomal Ribonucleic Acid
SAM	*S*-Adenosyl Methionine
SARS-CoV-2	Severe Acute Respiratory Syndrome *Coronavirus* 2
SASA	Solvent Accessible Surface Area
SDS-PAGE	Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SIMS	Secondary Ion Mass Spectrometry
SRL	Sarcin-Ricin Loop
TBD	Tris-Buffered Saline
TBST	Tris-Buffered Saline with Tween
TDBA	4-[3-(Trifluoromethyl)-3H-Diazirin-3-yl]Benzoic Acid
TEM	Transmission Electron Microscopy
TEV	Tobacco Etch Virus
TFA	Trifluoroacetic Acid
TF-RPF	Trifluoromethyl Radical Protein Footprinting
TJM	Trajectory Method
TOF	Time-Of-Flight
tRNA	Transfer Ribonucleic Acid
TWIMS	Travelling Wave Ion Mobility Spectrometry
UV	Ultraviolet
VHL	von Hippel-Lindau
XIC	Extracted Ion Chromatogram

Amino Acids

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
Hydroxyproline	Hyp	O
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Selenocysteine	Sec	U
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	Y
Tyrosine	Tyr	W
Valine	Val	V

Publications

Baines, C., Sargeant, J., Fage, C.D., Pugh, H., Alkhalaf, L.M., Challis, G.L. and Oldham, N.J., 2024. Native ESI-MS and Collision-Induced Unfolding (CIU) of the Complex Between Bacterial Elongation Factor-Tu and the Antibiotic Enacyloxin IIa. *Journal of the American Society for Mass Spectrometry*, 37, 1490-1496.

Ducker, C., Baines, C., Guy, J., Euzébio Goulart Santana, A., Pickett, J.A. and Oldham, N.J., 2024. A Diterpene Synthase from the Sandfly *Lutzomyia longipalpis* Produces the Pheromone Sobralene. *Proceedings of the National Academy of Sciences*, 121, e2322453121.[*]

---

[*] *Work not described in this thesis.*

1

Introduction

1.1. Proteins

Almost the entirety of processes, across all forms of life, are governed by biomacromolecular assemblies such as proteins and nucleic acids. In the prokaryotic workhorse *Escherichia coli*, proteins constitute approximately 55% of the cellular dry mass, which equates to a conservative estimate of 3 – 4 million proteins per individual cell.[1] The proteomic differences between organisms, or between individual cell types in multicellular organisms, are largely responsible for their observed phenotypic differences.[2] Protein interactions maintain cellular homeostasis through gene expression, signalling, metabolic pathways, defence, and many other processes, mostly through interactions with small molecules, or other proteins.[3,4] Proteins can have a defined three-dimensional structure, maintained by intramolecular interactions between the amino acid building blocks, or instead lack a defined structure and be categorised as intrinsically disordered. In reality, proteins exist on a spectrum of order, with most containing disordered, or highly flexible, regions and are highly dynamic. Often, disorder is a function of activity and may only be present under certain conditions.[5,6] As protein structure is closely tied to function, elucidation of structure plays a vital role in understanding the roles these macromolecules play.

1.1.1. Protein Structure

At the basic level, proteins are composed of linear sequences of amino acids, covalently linked together through the peptide bond. Their structures can be categorised into four groups. The *primary structure* refers to the sequence of individual amino acids within the polypeptide chain. The overall sequence of the polypeptide is encoded by its corresponding gene within the genome, with the identity of the individual amino acid residues specified by the codons within the gene. Intramolecular interactions between amino acids lead to the formation of defined *secondary structure* features. The most common of these are the α-helix and β-sheet. The arrangement of these features in space corresponds to the *tertiary structure*. Folding into the tertiary structure is largely governed by the burial of hydrophobic residues which is then stabilised by non-covalent interactions, such as hydrogen bonding, salt bridges, and covalent disulfide bonds. The tertiary structure may include multiple domains, typically linked together with disordered loops, that may have defined sub-structures including β-barrels and coiled-coils. If two or more folded polypeptide chains assemble together, this forms the *quaternary structure*.

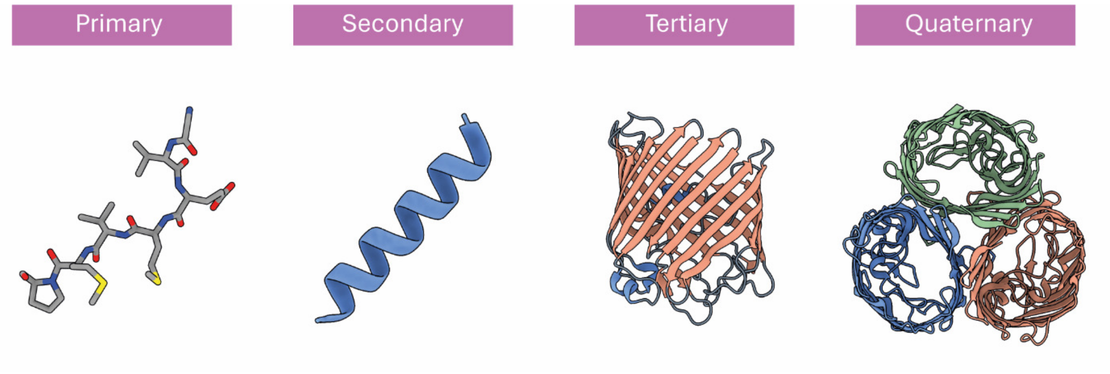


Figure 1.1 | The four levels of protein structure. Left to right; *primary structure* polypeptide; *secondary structure* α-helix; *tertiary structure* folded protein (OmpF, PDB 3POU); *quaternary structure* OmpF homo-tetramer, 3POU.

The individual chains (subunits) may have the same peptide sequence, forming *homo*-multimers or they may have difference sequences, forming *hetero*-multimers. Assembly of quaternary structures largely occurs at highly hydrophobic regions, whereby binding of subunits buries hydrophobic residues within the interface, shielding them from the solvated cellular environment. An overview of protein structures can be seen in figure 1.1. For the majority of proteins, folding is a spontaneous process, and therefore must be thermodynamically favourable. A minority of proteins require chaperones to aid folding, acting as either *foldases* to aid folding in an ATP-dependent manner, or *holdases* to stabilise aggregation prone intermediate structures. As the protein folds, its Gibbs free energy must become more negative, until it reaches a stable state at a free energy minimum.[7] This process can be imagined as a funnel-like free energy landscape (figure 1.2). The landscape is rugged and filled with traps bordered by energy barriers, as a protein folds it travels down the landscape, existing in potentially many meta-stable states prior to reaching the stable minimum. Native protein structures may not be the lowest energy state of the system and may exist in deep traps within the landscape.

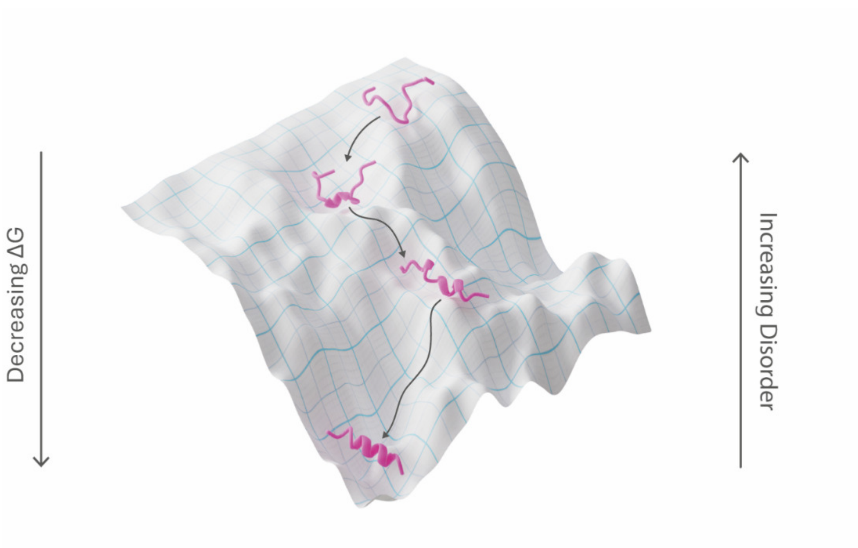


Figure 1.2 | Graphical representation of protein folding navigating the Gibbs free energy landscape. The disordered polypeptide becomes more folded in a discrete manner, with each sequential folding bringing a further decrease in free energy.

In neurodegenerative diseases induced by protein misfolding and aggregation, the protein structure responsible for amyloid fibril aggregation is of lower free energy than that of the correctly folded native state.[8]

1.1.2. Protein Interactions

As stated previously, proteins are dynamic structures containing regions differing in flexibility. Functional domains are typically linked through disordered loop regions which are often critical to their function. In an extreme example, disordered regions in the well-known molecular motor kinesin are vital to impart the flexibility required for the protein to “walk” along cellular microtubes, without which intracellular transport would largely shut down.[9] Protein dynamics are crucial for their ability to bind small molecules and other proteins. Conformational changes may be small, effecting only proximal residues or they may be large, resulting in entire domain rearrangements.[10] Small-molecule ligands binding to proteins are often either cofactors required for protein activity or substrates for an enzymatic reaction.

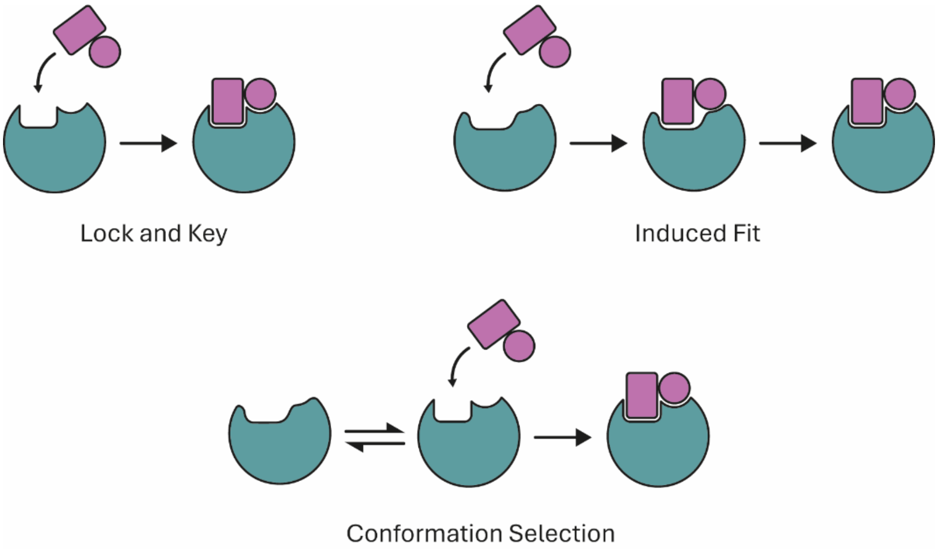


Figure 1.3 | Graphical representation of the three protein-ligand binding interactions. Protein is represented in blue and ligand in orange.

The specificity of a protein towards a ligand was originally proposed to occur through the *lock and key* model by Emil Fischer, where the ligand and protein have rigid, specific geometries that perfectly complement each other, explaining enzyme-substrate specificity.[11] Improvements to this hypothesis by Daniel Koshland led to the *induced fit* model, incorporating structural flexibility of both protein and ligand.

Here, the binding site is flexible and allows for binding of less complimentary ligands through ligand-dependent conformational changes, whilst still allowing for conformational proofreading.[12,13] Finally, the *conformational selection* model proposed by Monod, Wyman, and Changeux hypothesises that the protein binding site undergoes ligand-independent conformational changes, and that the ligand can then bind to the most suitable conformer.[14] An overview of these hypotheses can be found in figure 1.3.

1.1.3. GTPases

GTPases are a broad family of hydrolytic enzymes, all of which bind to guanosine nucleotides. Acting as molecular switches, GTPases are considered to be in their *active* conformation when bound to guanosine triphosphate (GTP) and transition to their *inactive* conformation by hydrolysing GTP to guanosine diphosphate (GDP).[15] Magnesium, present as the Mg^{2+} ion, is an essential cofactor. GTPase activity is typically modulated by interaction with GTPase activator proteins (GAPs). Following hydrolysis, the exchange of GDP to GTP is facilitated by guanine nucleotide exchange factors (GEFs). Binding of which increases the rate of GDP dissociation from the GTPase, allowing GTP to passively bind due to its higher cellular concentration.[16] GTPases have a wide range of cellular activities including signalling, translation and membrane translocation.[17] Although posing a broad range of activities, GTPases (or GTPase domains within larger proteins) are structurally and sequentially conserved. Five motifs (G1 – G5) are essential for their function. G1 contains the highly conserved *P-loop* (Walker Motif, GxxxxGK[T/S]) and contacts the α- and β-phosphates of GDP/GTP.[18] G2 and G3, or *switch I* and *switch II*, are semi-disordered regions that adopt distinct conformations in the GTP or GDP bound state. They contact the γ-phosphate of GTP, aid coordination of the Mg^{2+} cofactor, and are both bound by GAPs. G3 also contains a basic residue, typically glutamine or histidine which coordinates the water molecule used in nucleophilic attack on the γ-phosphate during GTP hydrolysis.[19] G4 and G5 both contact the guanine nucleobase and impart guanine specificity.[20] A representative GTPase (Human Ras) can be seen in figure 1.4.

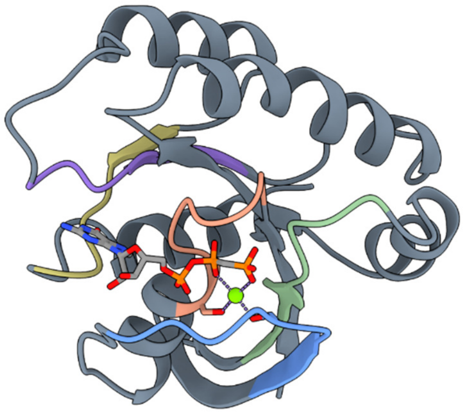


Figure 1.4 | Human Ras GTPase (PDB 121P). Bound GTP is shown in stick representation, coloured by atom. The five key motifs are coloured; G1, orange; G2, blue; G3, green; G4, purple; G5, yellow. Mg^{2+} is represented by a green sphere, coordinating the β- and γ-phosphates to Ser^{17} and Thr^{35} in G1 and G2 respectively. Hydrogens and solvent water omitted for clarity.

1.1.3.1. Translational GTPases

Translational GTPases (tr-GTPase) are a subset of the GTPase family, all responsible for the facilitation and stringent control of mRNA translation at the ribosome. All stages of translation at the ribosome are controlled by tr-GTPases, demonstrating their vital role. In prokaryotes, translational rates are under the control of four major translational GTPases: initiation factor (IF) 2, elongation factor (EF) Tu, EF-G and release factor (RF) 3. A unique elongation factor, SelB, is responsible solely for the transport of selenocysteine charged transfer RNA (Sec-tRNA^{Sec}) to the ribosome and adopts a hybrid IF-2/EF-Tu-like structure.[21] All of these tr-GTPases bind the ribosome in the GTPase-associated centre (GAC), found within the 23S ribosomal RNA (rRNA). The GAC is associated to a stalk-like, typically pentameric protein complex (subunits L10, 2x L12, 2x L7) that protrudes from the large (50S) ribosomal subunit, L11, and binds directly to the 23S rRNA.[22] A unique feature of these tr-GTPases, is the use of the ribosome itself as the GAP. Except for EF-Tu, tr-GTPases in general exhibit a greater affinity towards GTP than GDP, rendering the need for a GEF null. EF-Tu, however, requires binding of a GEF, EF-Ts, to facilitate GDP/GTP exchange.[23] Eukaryotic homologues of EF-Tu, SelB, EF-G and RF-3 (eEF1A, EF-SEC, eEF2, and eRF3) perform a similar role, whereas translation initiation is controlled by a structurally unique tr-GTPase, eEIF2.[24] With the focus of this work being on the prokaryotic EF-Tu, further discussion will focus on prokaryotic translation.

1.2. EF-Tu

1.2.1. General Structure

EF-Tu (figure 1.5, A) is a three-domain protein, consisting of *domain I* (residues 1 – 200), a canonical Ras-like GTPase domain that is mostly α-helical and *domains II* (residues 209 – 292) and *III* (residues 300 – 393), which are β-barrel based, tRNA binding domains.[25] Domains I and II are linked by a flexible 9-residue typically unstructured, loop and domains II and III by a shorter, 7-residue linker. Similar to other tr-GTPases, EF-Tu adopts a range of conformations, irrespective of the nucleotide bound. This contrasts with the classical view of the distinct binary *active* and *inactive* conformations adopted by GTPases, in complex with GTP and GDP respectively.[20,26] In contrast to other tr-GTPases however, the conformational extremes sampled by EF-Tu involve a dramatic rotation (figure 1.5, B), almost 90°, of domains II and III with respect to domain I as opposed to more localised, domain interface changes seen in SelB, and EF-G.[21,27,28] To avoid confusion, for EF-Tu, these extremes will be referred to as *open* and *closed* instead of *active* and *inactive*. Again, similar to other tr-GTPases, the closed confirmation is reached only upon binding of EF-Tu to the ribosome, free EF-Tu in solution appears to be sample a broad range of open to closed conformations.[27,29]

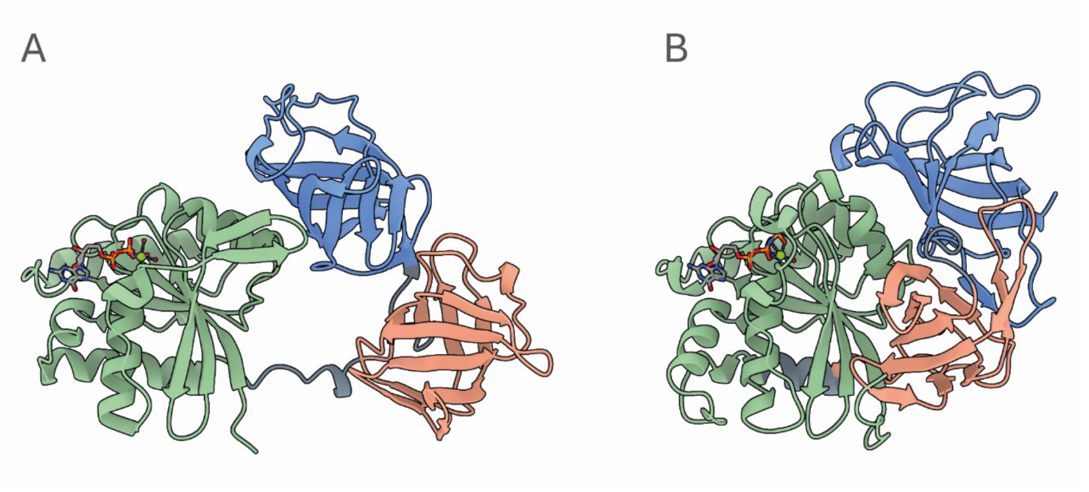


Figure 1.5 | *E. coli* EF-Tu represented as cartoon structures. Domain I is coloured green, domain II, orange, and domain III, blue. Structures were aligned by domain I. A | EF-Tu in the open conformation, bound to GDP (stick representation), PDB 1EFC. B | EF-Tu in the closed conformation, bound to GDNP (stick representation), PDB 1EFT.

1.1.2.1. Domain I

Domain I (figure 1.6, A) exhibits a classical nucleotide-binding α/β structure, consisting of a central β-sheet core (five parallel β-strands, *a, d, e, f, g*, and one antiparallel β-strand, *c*,) surrounded by five α-helices, *B, C, D, E* and *F*.[30] The P-loop is located directly after the C-terminus of β-strand *a*, and extends into N-terminus of α-helix *A*. Importantly, alongside the nucleotide binding pocket, domain I contains the highly flexible switch I and II regions. Referring to the GDP-bound crystal structure obtained by Song *et al.*, switch I forms a single α-helical turn, *A''*, followed by a short loop which subsequently folds into a small anti-parallel β-sheet, *b'* and *b*. The connecting hairpin turn between them forms a hydrogen bonding network between Arg^{58} and Arg^{373} in domain III through a bridging water molecule.[25] Switch II is composed entirely of helix *B* (His^{84} to Ile^{92}), with the catalytic histidine, His^{84}, oriented away from the nucleotide binding pocket. The Mg^{2+} cofactor is coordinated by the side-chain hydroxyl group of Thr^{25}, three water molecules and a β-phosphate oxygen in GDP.

In the 5'-guanylyl imidodiphosphate (GDPNP)-bound crystal structure solved by Nissen *et al.* (figure 1.6, B), switch I undergoes a subtle rearrangement, where the *b'* *β*-strand folds into helix *A''* and rotates inwards, over the nucleotide binding pocket.[31] Switch II undergoes a more dramatic conformational shift, where helix *B* unwinds a turn at the N-terminus and folds several new turns at the C-terminus (Ala^{85} to Gln^{98}). Here, GDPNP acts as a non-hydrolysable GTP analogue where the bonding oxygen between the β and γ-phosphates is replaced with a nitrogen, forming an imidophosphate. The use of GDPNP allows for the study of GTP-bound EF-Tu, without intrinsic GTP hydrolysis by EF-Tu. This, paired with an approximately 40° rotation perpendicular to its axis, moves the catalytic histidine, and positions it facing closer to the nucleotide binding site.[31] His^{84} is still positioned in such a way that autohydrolysis of GTP is exceedingly rare, a rate of about $5\times10^{-5}\ s^{-1}$.[32] The Mg^{2+} cofactor loses a bound water molecule and instead coordinates a γ-phosphate oxygen of GTP.

1.1.2.2. Domains II and III

Domain II consists of 10 anti-parallel *β*-stands (*a2* – *j2*), coalesced into a *β*-barrel structure. Two looping regions connecting strands *b2* – *c2* and *g2* – *h2* protrude outwards into the bulk solvent and produce part of the binding site for the cyclic thiazolyl antibiotic, GE2270A.[31,33] Domain III consists of 6 anti-parallel *β*-stands (*a3* – *f3*) also forming a *β*-barrel structure. During the large conformational change of domains II and III relative to I, the interactions between domains II and III remain relatively stable.[31] New interactions between domains I and III are formed predominantly by helices *B, C* and *D*. Ala^{86} in helix *B* and Glu^{118} in helix *C* form hydrogen bonds with Arg[389] in strand *f3* in domain III, the former through a bridging water molecule. Tyr[161] in helix *D* forms a new hydrogen bond with Glu[327] in domain III.

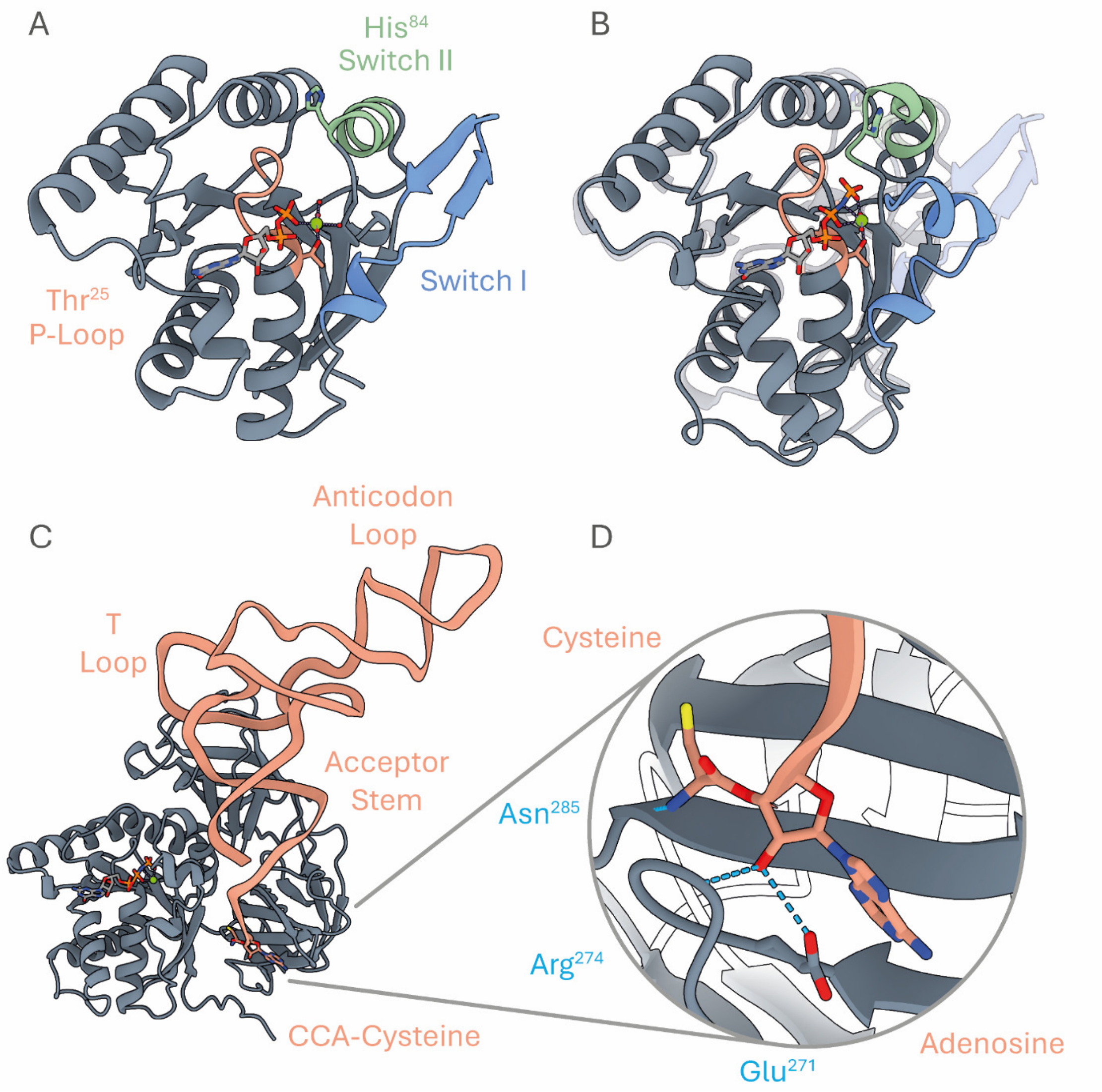


Figure 1.6 | *E. coli* EF-Tu domain I, with key features highlighted, represented as cartoon structures. Mg^{2+} represented as a green sphere, crystal water molecules as a red sphere and nucleotides in stick representation A | EF-Tu in the open conformation, bound to GDP, PDB 1EFC. B | EF-Tu in the closed conformation, bound to GDNP overlain over the open configuration, PDB 1EFT. C | EF-Tu•GTP•aa-tRNA ternary complex, Cys-$tRNA^{Cys}$ shown in orange, PDB 1B23. D | Zoom of C showing domain II residues binding the 3' adenosine alongside the ester conjugated cysteine.

Domains II and III are predominantly responsible for tRNA binding, however some contacts are made with domain I. All residue numbering here is with respect to the *Thermus aquaticus* EF-Tu. The acceptor stem of an aminoacyl-tRNA (aa-tRNA), containing the 3' terminal adenosine, covalently linked to an amino acid through an ester bond, slots into the cleft between formed between domains I and III and places the amino acid into the binding pocket formed by the loops connecting strands a*2* – *b2* and *d2* – *e2*. Depending on the amino acid bound, further contacts can be made to Glu^{271}, His^{273}, Arg^{274} and Asn^{285}.[34] The tRNA 5' end protrudes into the pocket formed at the interface of all three domains, interacting with Arg^{300} (domain II); Lys^{90}, Asn^{91} and Tyr^{88} (helix *B,* domain I); and Glu^{55} (helix *A''*). Finally, the T-loop binds with domain III exclusively through residues Arg^{330}, His^{331}, Gln^{341}, Thr^{350}, Lys^{376} and Gly^{391} (figure 1.6, C and D).[34]

1.2.2. Prokaryotic Translation

1.2.2.1. Translation Initiation

In prokaryotes, DNA transcription to mRNA, and then translation to a polypeptide chain, are physically coupled. The 30S ribosomal subunit interacts directly with actively transcribing RNA-polymerase, and the emerging mRNA fed directly into the ribosome to be translated.[35] Mature ribosomal formation is a highly complex process and will only be touched upon through the lens of the tr-GTPase IF-2. The 30S pre-initiation complex is formed of the 30S ribosomal subunit, bound to IF-1, IF-2•GTP, IF-3 and N-formylmethionine-$tRNA^{FMet}$ (positioned in the ribosomal *A* site). Recruitment of mRNA, followed by start codon (AUG) recognition by FMet-$tRNA^{Fmet}$ forms the 30S initiation complex.[36] Upon binding of the 50S large ribosomal subunit and dissociation of IF-3, the 70S pre-initiation complex is formed. An IF-2 conformational change induces GTP hydrolysis, releasing inorganic phosphate (PO_4^{3-}, P_i) and shifts FMet-$tRNA^{FMet}$ to the ribosomal *P* site. This sequentially leads to dissociation of IF-1 and IF-2•GDP, forming the complete 70S initiation complex. [37] The ribosome is now primed to receive coding tRNA and begin elongation.

1.2.2.2. Translation Elongation – Decoding

Polypeptide elongation consists of repeating cycles of decoding, peptide bond formation and ribosomal translocation (figure 1.7). Elongation initiates as soon as the ribosomal *A* site is free and continues until a stop codon is reached. During the decoding stage, the current mRNA codon in the *A* site is translated. This is achieved by the delivery of aa-tRNA to the ribosome by EF-Tu, as the ternary complex EF-Tu•GTP•aa-tRNA. The ternary complex initially interacts with ribosomal stalk, which shuttles the complex towards the GAC and *A* site.[38] Codon:anticodon recognition is a vital process in translation, incorrect pairing leads to incorporation of the wrong amino acid into the growing polypeptide. If there are successful codon:anticodon interactions, conformational movements in 16S rRNA cause closure of the 30S domain, which in turn, pulls the EF-Tu complex down towards the sarcin-ricin loop (SRL) of 23S rRNA.[39] Interaction with the SRL, which acts as EF-Tu's GTPase activator, induces GTP hydrolysis. In brief, interaction with the RNA backbone raises the local pK_a of the imidazole histidine side chain, rendering it protonated at native pH. What likely follows is a proton abstraction of the catalytic water molecule, bound to His^{84}, by the γ-phosphate in GTP. The resulting nucleophilic hydroxide ion, stabilised by the basic histidine, attacks the γ-phosphate resulting in hydrolysis.[40–42] This interaction with the SRL and the resulting electrostatic environment is responsible for increasing EF-Tu's rate of hydrolysis by five orders of magnitude.[32]

Following hydrolysis, liberated P_i remains bound to EF-Tu, held in place by the conformation of switch I and II. To be released, local conformational changes must take place to facilitate dissociation.[41,43] Delayed release of P_i is seen in EF-G and IF-2 and is linked to their function on the ribosome.[44,45] Current understanding has not linked delayed P_i release to a biological function of EF-Tu.

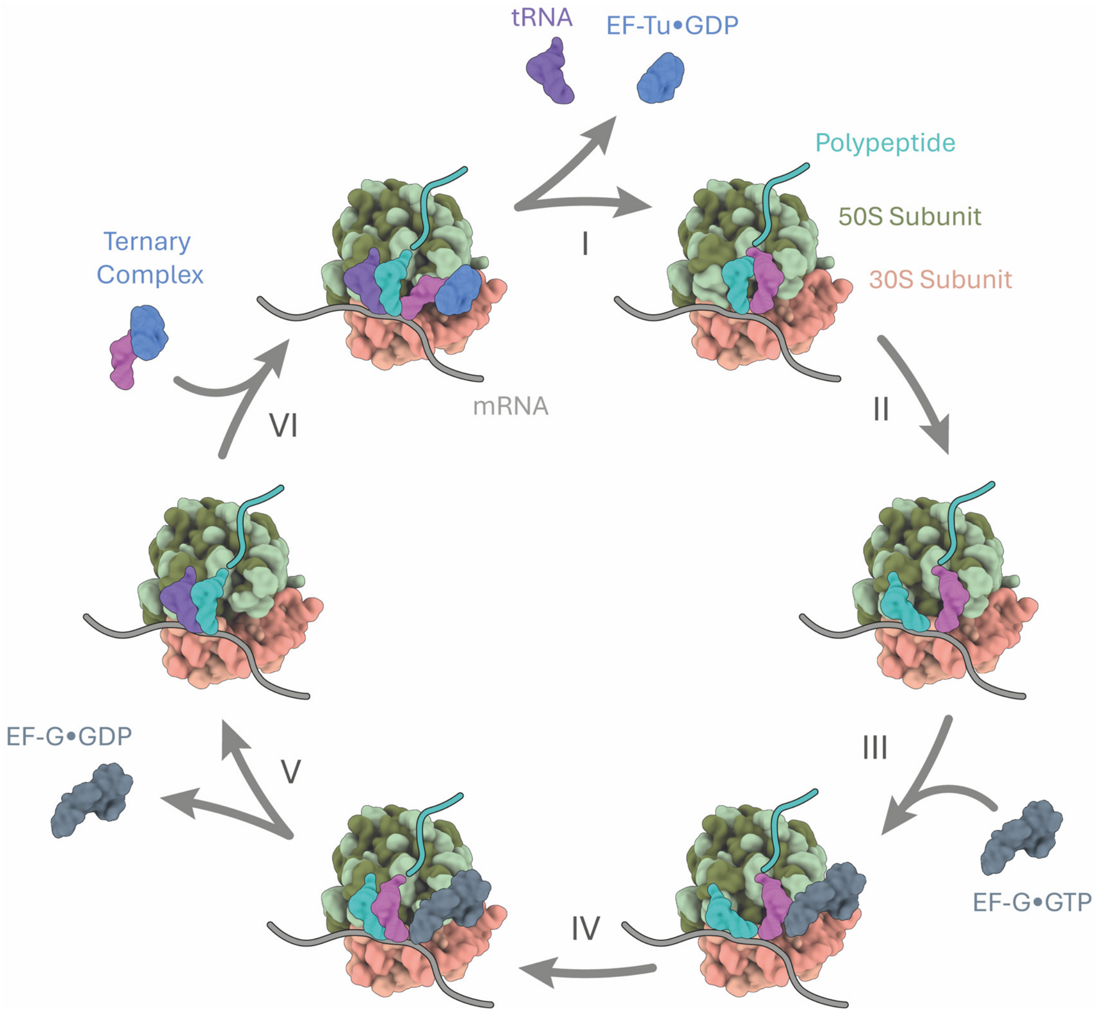


Figure 1.7 | A graphical overview of the major steps in elongation. 50S ribosomal subunit, green; 30S subunit, orange; polypeptide, blue line; mRNA, grey line. I | EF-Tu GTP hydrolysis, release, and release of empty tRNA from the *E* site. II | Peptidyl transfer to the new aa-tRNA (pink) in the *A* site from old peptidyl-tRNA (cyan) in the *P* site. III | EF-G binding with tRNA in a hybrid state between sites. IV | EF-G hydrolysis, ribosomal translocation, new peptidyl-tRNA fully occupying the *P* site and the old tRNA now in the *E* site. V | EF-G release. VI | EF-Tu ternary complex free to bind to the vacant *A* site. Figure adapted from Voorhees & Ramakrishnan. Ribosome•tRNA•EF-Tu, PDB 4V5G; EF-G, PDB 7SSL.[316]

It is known to not be related to proofreading, as cognate and near-cognate codon:anticodon complexes demonstrate the same delay. It is likely due to the tight association between EF-Tu and the ribosome ($K_D < 1$ nM) preventing efficient conformational shifts in the switch I and II region, and keeping the binding pocket closed.[41,43] Once P_i has been released EF-Tu rapidly rearranges into the open conformation, this switch is hypothesised to force the aa-tRNA fully into the ribosomal *A* site whilst dissociating EF-Tu from the ribosome.[46] Total accommodation into the *A* site is further dependent on the aa-tRNA adopting a specific conformation. The rate of accommodation is reduced for near-cognate aa-tRNA, which provides a further layer of proofreading, where near- or non-cognate aa-tRNA can dissociate from the ribosome before fully entering the *A* site.[47,48]

1.2.2.3. Translation Elongation – Peptide Bond Formation and Translocation

The mechanisms involved in formation of the peptide bond between adjacent aa-tRNAs and the subsequent translocation of the ribosome along the mRNA chain will be briefly explored. The peptide bond is formed between the new aa-tRNA in the *A* site and the current peptidyl-tRNA occupying the *P* site. The ribosome, specifically rRNA, acts as an entropic ribozymal catalyst for the bond formation.[49,50] It does not provide catalytic functional groups but instead places the tRNA reactants, along with catalytic water molecules, into their ideal orientations for reaction.[51] The ribozymal activity accelerates this reaction by a factor of 10^7 when compared to a non-catalysed solution-based reaction.[49] The reaction proceeds with a nucleophilic attack by the amino group of aa-tRNA on the ester carbonyl carbon of peptidyl-tRNA. Two separate pathways are hypothesized, but both involve a catalytic water and the movement of three protons, resulting in covalent linking of the peptide chain to the aa-tRNA.[50,51]

Following bond formation, the ribosome moves along the mRNA chain by one codon. This is a multi-step process requiring several dramatic conformational changes. Focussing on the role of EF-G•GTP, ribosomal binding of which induces the forward movement of the 30S subunit.[52] GTP hydrolysis by EF-G allows the tRNA codon:anticodon complexes to dissociate from the 30S subunit, and through a series of further steps, movement of tRNA to the *E* site and the peptidyl-tRNA to the *P* site.[53,54] Dissociation of EF-G•GDP and *E* site tRNA restore the ribosome to a position to accept new EF-Tu ternary complexes into the *A* site, completing the elongation cycle.[55]

1.2.3. The EF-Tu Cycle

As mentioned previously EF-Tu, unlike the other tr-GTPases, exhibits greater affinity towards GDP than GTP, approximately 60-fold higher.[23] This results in much slower rates of passive nucleotide exchange, and the requirement of a GEP, EF-Ts, to facilitate this exchange. EF-Ts interacts with EF-Tu•GDP domain I, through contacting helix *D*, the P-loop, and switch II whilst securing itself to domain III (figure 1.8, A).[56] Binding causes domain I to separate slightly from domains II and III, allowing switch II to be pushed away from the bound Mg^{2+} by domain N of EF-Ts. This pulls away the stabilising Asp^{80} and allows for Mg^{2+} to dissociate from EF-Tu.[56] Helix *D* becomes slightly withdrawn, pulling the guanine-binding Asp^{138} away from the base and allowing it to leave the binding pocket. Alongside this, the P-loop Lys^{24} rotates outwards, following the translocated Asp^{80}, freeing the β-phosphate from the P-loop. Finally, disruption of the P-loop Asp^{21} liberates the α-phosphate, allowing total dissociation of the nucleotide (figure 1.8, B).[56] The approximate 10-fold higher concentration of cellular GTP over GDP allows probabilistic binding of GTP to EF-Tu•EF-Ts.[16,23] Binding of GTP effectively follows these steps in reverse, by first accommodating the α-phosphate, then sequentially the guanine ring, β- and γ-phosphates, and finally the Mg^{2+} cofactor.[56]

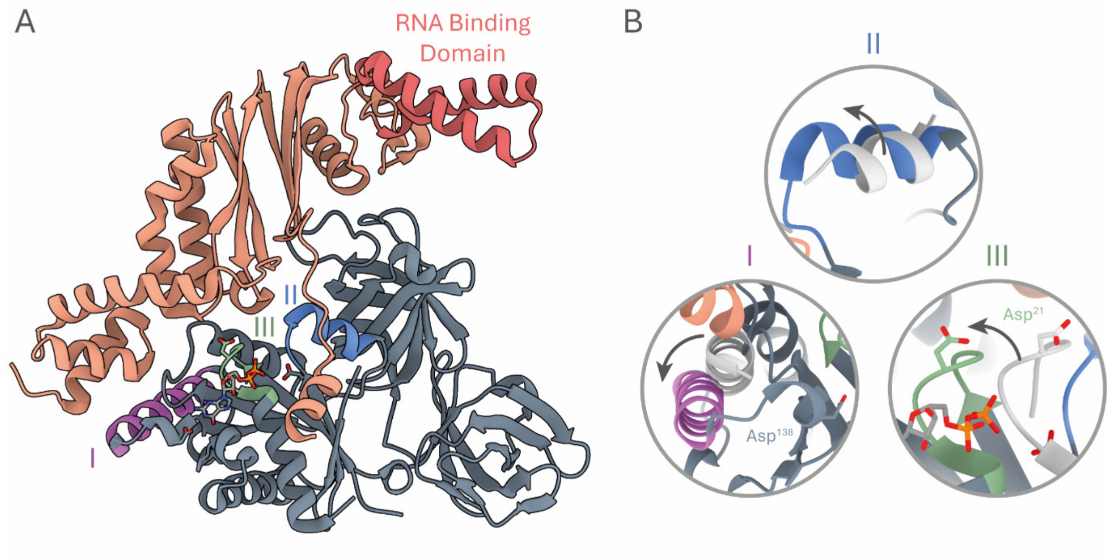


Figure 1.8 | A cartoon representation of the EF-Tu•GDP•EF-Ts (PDB 4PC2) complex and select induced conformational changes leading to GDP release. A | EF-Tu, grey; EF-Ts, orange; putative RNA binding domain, red; helix *D*, purple; P-loop, green; switch II, blue. B | Zoom of key regions showing deviation from the standard EF-Tu•GDP, white; I, backwards movement of helix *D,* pulling Asp^{138} away from the guanine; II movement of switch II away from the binding site, allowing release of Mg^{2+}; III, movement of the P-loop backwards from the binding site, importantly pulling Asp^{21} away from the α-phosphate.

Classically, cofactor binding was thought to then trigger the conformation of EF-Tu to adopt the closed conformation, in the process dissociating from EF-Ts. Instead, Thirup *et al.,* postulate that EF-Ts binding is partially disrupted by the structural changes in domain I (namely switch I and II) upon GTP binding, and not due to the global conformational change. In fact, EF-Ts was shown to remain bound to EF-Tu whilst the latter was in the closed conformation.[56] Thirup *et al.* also captured EF-Tu•GDPNP•EF-Ts in the open conformation, corroborating Kavaliauskas and colleagues' demonstration of GTP/GDPNP bound EF-Tu existing in the open conformation.[26,56] The current accepted theory is that EF-Ts aids the binding of tRNA to EF-Tu through an aminoacyl transferase-like domain (figure 1.8, A), and a quaternary EF-Tu•GTP•EF-Ts•tRNA complex is formed, although the structure has yet to be elucidated.[56–58] Binding of an aa-tRNA causes a steric clash between the aa-tRNA T-loop and EF-Ts, leading to EF-Ts dissociation, and formation of the ternary complex. If non-aa-tRNA is sequestered in this way, there is no steric interference due to the different conformation of uncharged tRNA.[57] Uncharged tRNA does not make the sufficient bonding network to domain II of EF-Tu, and as a result is released from the quaternary complex without disrupting EF-Ts binding. This interaction allows for efficient rates of translation by reducing the likelihood of uncharged tRNA being transported to the ribosome. The overall EF-Tu/EF-Ts cycle is summarised in figure 1.9.

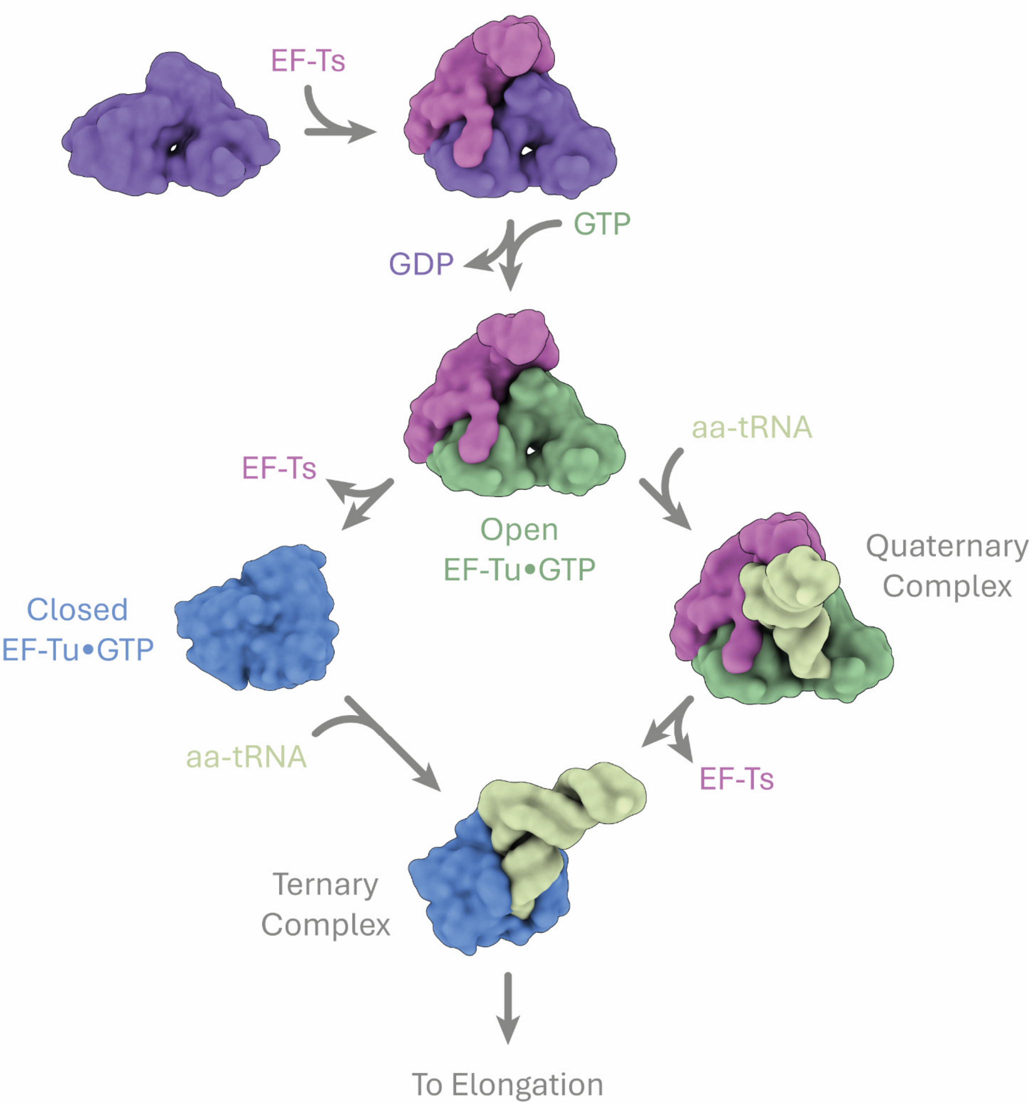


Figure 1.9 | A graphical overview of the major steps of the EF-Tu/EF-Ts recharging cycle, starting from EF-Tu•GDP (purple). Highlighting the two alternate ways EF-Tu•GTP•EF-Ts can progress, either through dissociation, or formation of the quaternary complex prior to the ternary. PDB files used; EF-Tu•GDP, 1EFC; EF-Tu•GDP/GTP•EF-Ts, 4PC2; Quaternary Complex, 4PC2 with tRNA from 1TTT; closed EF-Tu•GTP, 2BVN; ternary complex, 1TTT.

1.2.4. Elfamycins

EF-Tu, due to its ubiquitous nature and essential role in protein biosynthesis, makes for a desirable antibiotic target. Antibiotics that target protein synthesis *as a whole* have been in clinical use for several decades, but these tend to focus on direct ribosome binding and inhibition.[59] Antibiotics including tetracycline and streptomycin target the *A* site, whereas kasagumycin and dityromycin target the *P*/*E* sites and EF-G binding respectively.[60–62] There are myriad compounds that exhibit anti-ribosomal activity that bind across the ribosome, composed of polyketides, macrolides and antimicrobial-peptides to name a few.[62] The elfamycins are a group of structurally diverse natural products, all specifically targeting EF-Tu.[63] One common factor amongst members of the group, aside from their shared low aqueous solubility, is their poor pharmacokinetic profiles. To date only two semi-synthetic elfamycins, NAI003 and LFF571, are currently in clinical trials, with LFF571 having completed phase II.[64,65] Currently, elfamycins see substantial laboratory use due to the ability of some to trap EF-Tu on the ribosome, allowing crystal structures to be collected of that normally transient complex.[63] The elfamycins can be subcategorised into two groups based on their major mode of action; kirromycin-like and pulvomycin-like, named after the prototypical compounds, summarised in figure 1.10, A.

Kirromycin (KIR), the first discovered elfamycin, binds to the interface between domains I and III of EF-Tu, trapping it on the ribosome through inhibition of the conformational change induced by GTP hydrolysis, stalling elongation by both EF-Tu and ribosomal inhibition (figure 1.10, B).[66,67] Kirromycin also has a destabilising effect on switch I and II, causing disordering of helix *A”*, and a rotation of His^{84} closer to the nucleotide phosphoryl groups. As a consequence of this, GTP autohydrolysis rates are increased four-fold, but still lag greatly behind that of the ribosomal-bound complex.[67,68] Globally, kirromycin binding induces a greater compaction of the closed

EF-Tu structure, forcing domains II and III closer to domain I by acting as a wedge to widen the interface between I and III.[69]

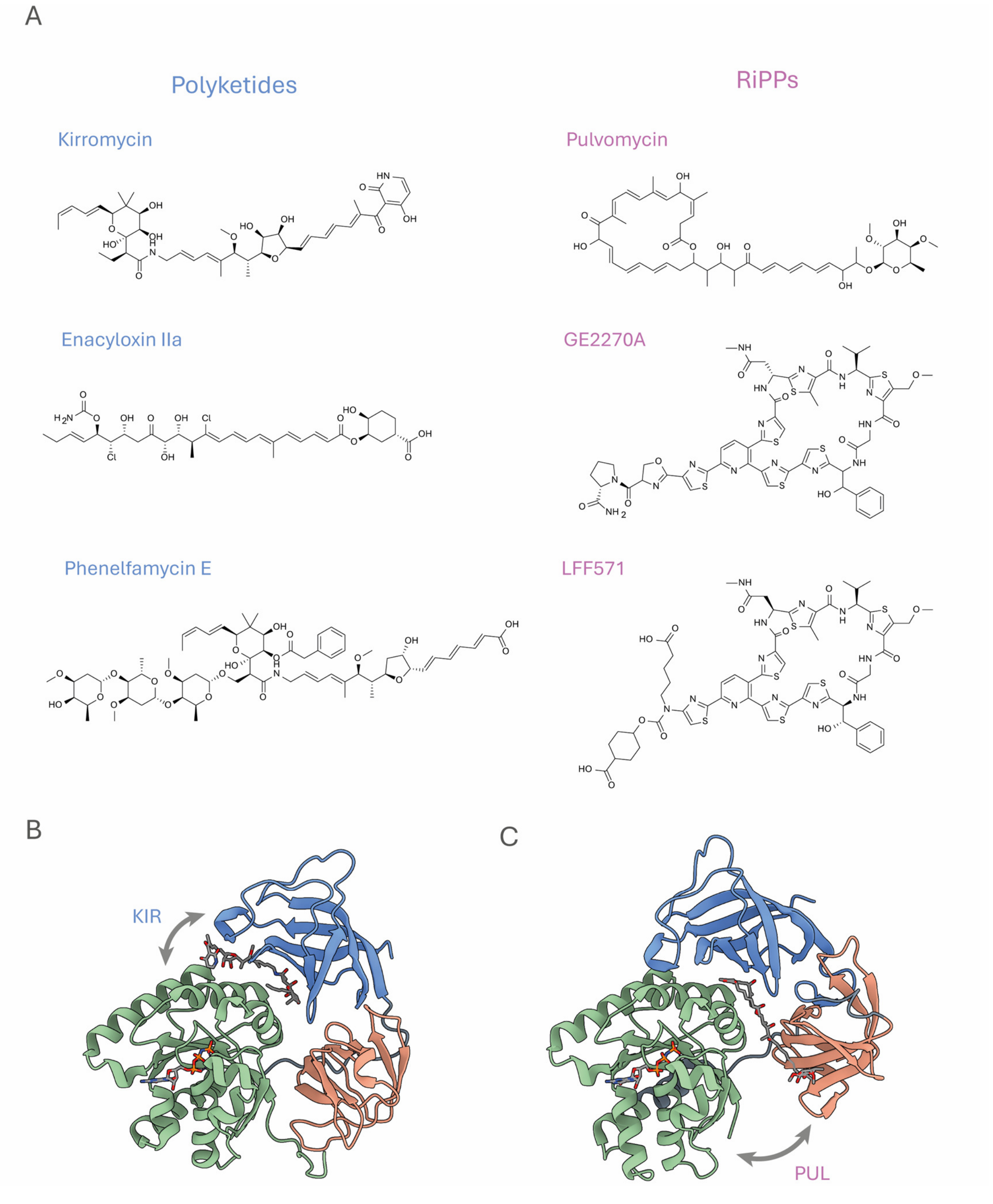


Figure 1.10 | Representative overview of the elfamycins. A | Select members of the two elfamycin groups, showing extensive structural diversity. RiPPS - Ribosomally synthesized and post-translationally modified peptides. B | EF-Tu, (domain I, green; domain III, orange; domain III, blue), bound to kirromycin (KIR) showing widening of the domain I/II interface. PDB 1OB2. C | EF-Tu bound to pulvomycin (PUL) showing extensive widening of the domain I/II interface. PDB 2C78. Structures were aligned by domain I, both bound to GDPNP.

In contrast, pulvomycin binds in the junction of all three closed EF-Tu domains, whilst importantly extending down into the aa-tRNA anchoring site in domain II (figure 1.10, C).[68] Mainly, pulvomycin acts as a competitive inhibitor of aa-tRNA binding, preventing formation of the ternary complex and therefore stalling the delivery of aa-tRNA to the ribosome.[69,70] Similar to kirromycin, pulvomycin induces a modest 5-fold increase in GTP autohydrolysis. However, this is due to movement of Ile^{61}, a residue shielding the γ-phosphate from His^{84}, instead of a direct induced movement of His^{84}.[69,70] In contrast to kirromycin, however, pulvomycin induces a more expanded EF-Tu conformation, through forcing the domain I-II interface apart, as measured through crystallography and by native sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE).[69] Modification of the resulting *holo*-EF-Tu structure, compared to the *apo*-, through both kirromycin and pulvomycin binding indicates their accommodation via the induced fit mechanism.[68] The elfamycins used in this work, enacyloxin IIa and GE2270A (kirromycin-like and pulvomycin-like respectively) will be discussed in detail within their relevant chapters.

1.3. Studying Protein Structure

As alluded to previously, the diversity of protein function is intimately tied to their structure, determination of one may give clues to the other. Elucidation of protein structure is therefore a key prerequisite of the complete understanding of protein interactions at a molecular level. Since the first protein structure of myoglobin was solved in 1958 by John Kendrew and colleagues (leading to his joint Nobel prize shared with Max Perutz, for solving of the haemagglutinin structure) by X-ray crystallography, huge developments have been made in accuracy and resolution of structural elucidation. In the decades following, the number of solved structures increased exponentially (figure 1.11).[71,72] In recent years, the application of machine learning and artificial intelligence for *ab initio* structure prediction has skyrocketed, with DeepMind's AlphaFold and its derivatives contributing 214 million individual predicted structures, compared to the initial deposition of over 360,000 structures in 2022.[73–75] Despite this, predicted structures are just that – a prediction, and experimental verification will always be required.

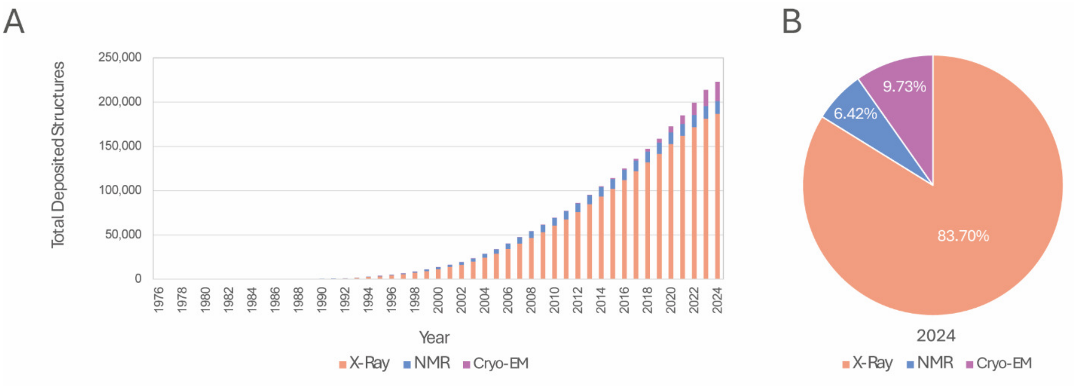


Figure 1.11 | Structure deposition data collected from the RCSB PDB. A | All deposited structures since the PDB formation in 1976. B | Proportion of deposited structures solved by the three major methods to date in 2024.

1.1.3. X-Ray Crystallography

Since its initial introduction, X-ray crystallography has produced a treasure trove of protein structures, representing approximately 85% of structures deposited in the Protein Data Bank (PDB, figure 1.11). At its heart, protein X-ray crystallography relies on the ability of a crystal to act as a diffraction grating for X-ray electromagnetic waves. To grow a protein crystal for diffraction, a large concentration of highly pure protein (typically obtained through heterologous expression, and purification) is required. A general process for production of protein crystals is as followed; first the highly pure, concentrated protein solution is slowly reduced in solubility to promote supersaturation, this may be achieved by modifying solvent conditions (pH, temperature *etc*) or by the addition of additives (ions or polymers).[76] Protein crystal nucleation is a spontaneous phenomenon in solution but may be induced by seeding, the introduction of artificial nucleation sites. However, care must be taken to not induce the growth of too many individual smaller crystals, as a single larger one is preferred. Ideally, solvent conditions should *disfavour* nucleation, but *favour* crystal growth. Through screening of multiple crystallisation conditions, a resulting high-grade regular crystal is hopefully obtained.

Crystals are then mounted and typically flash-frozen to reduce radiation damage, this process and the use of cryoprotectants may also require multiple trial-and-error modifications to produce a suitable crystal.[77] The crystal is then exposed to synchrotron-generated, monochromatic X-rays and the resulting diffraction pattern recorded. Each pattern represents only a slice of the total structure, and the crystal must be rotated, typically through 180°, and the diffraction patterns recorded in an incremental process.[76] Reconstruction of the protein structure from diffraction patterns is computationally difficult, using fast Fourier transformations (FFT) to generate 3D electron density maps, to which individual atoms can be modelled. Iterative refinement of the density with the model is continued until the model best matches the original diffraction data.

Despite its high accuracy, X-ray crystallography does come with drawbacks. Not all proteins are amenable to the crystallisation process, membrane proteins specifically pose specific problems.[78] The process requires large amounts of protein sample, and potentially multiple rounds of screening for ideal crystallisation conditions. Most importantly, the protein structure obtained is that of a low-energy, crystal-packed protein, and likely misrepresents the solution phase fluidity of the protein. Highly flexible regions tend to be poorly resolved in X-ray structures, if at all.[79]

1.1.4. Cryogenic Electron Microscopy

Cryogenic electron microscopy (most commonly, cryogenic transmission electron microscopy (cryo-TEM) is a relatively recent advancement of classical electron microscopy, where a high-energy beam of electrons is fired at an ultra-thin specimen section towards a detector (charge-coupled or direct electron detectors). The interactions of the electrons with the specimen generates an image, with resolution down to the atomic level possible due to the much smaller de Broglie wavelength of electrons compared to visible light microscopes. Exposure of the sample to cryogenic temperatures offers several benefits over traditional TEM, imparting a protective effect against the incoming electrons and the high-vacuum environment of the TEM stage. Importantly, plunge freezing of thin, aqueous biological samples generates macromolecules encased in a vitrified (amorphous, non-crystalline) ice, better preserving their solution-phase structures.[80] The technique proved successful, with the publishing of cryo-TEM micrographs of adenovirus and T4 bacteriophages, amongst others viral capsids in 1984 by Adrian *et al.*[81] In recent times, cryo-TEM has greatly increased in resolution with both technical and software improvements, culminating in an atomic, 1.15 Å resolution structure of human apoferritin by Yip *et al.* in 2020 and a yet unpublished 1.09 Å mouse apoferritin by Nazarov *et al.*[82]

To produce a structural map of a protein, a pure protein sample is vitrified and imaged with low electron doses to minimise structural damage. Due to this, the images typically have a poor signal to noise ratio and consequently, multiple image sets must be averaged to boost signal quality, through a process called *single particle analysis*. Proteins adopt random orientations in vitrified ice, and therefore must be segregated into orientation classes where they can be averaged, with minimal loss of resolution. For 1.25 Å apoferritin this was approximately 430,000 individual particle images.[83] This, paired with physical rotation of the specimen sample stage to acquire multiple angle images, allows for the reconstruction of a 3D electron density volume, to which an atomistic structure can be fit. Similar to X-ray crystallography, the volumes and atom models go through iterative refinement processes, to generate an optimal reconstruction.[83]

Although cryo-TEM offers advantages over X-ray crystallography, namely reduced protein concentrations/volumes and near-native conformations, it too comes with drawbacks. Highly flexible proteins prove troublesome to average, resulting in low resolution maps with high error; small proteins (<50 kDa) are difficult to image and require exogenous additives to resolve; resolution still lags behind cutting-edge X-ray crystallography, where sub-Ångström level resolutions are routinely possible; and prohibitive cost of equipment.[84–86]

1.1.5. Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a powerful analytical technique exploiting the propensity of magnetically active atomic nuclei to respond to changing radio frequency fields.[87,88] Sub-atomic particles exhibit an intrinsic angular momentum, which is typically represented as a dimensionless value, *spin*, obtained by the division of their angular momentum by the reduced Planck constant, $\hbar$.[89] Baryons (protons, neutrons *etc*), and all other fermions, have a half odd-integer spin ($\frac{1}{2}$, $\frac{3}{2}$ *etc*). Atomic nuclei, being fermionic in nature, have a net spin, I, and in nuclei with an even mass number I is 0. In nuclei with odd mass numbers, their net spin is a half odd-integer. This property leads them to be magnetically active, and therefore detectable by NMR.[90] In an NMR measurement, sample nuclei are held in a stationary magnetic field, B_0, application of a radiofrequency (RF) field, specifically its magnetic component B_1, causes active nucleic to flip their spin state. Removal of B_1 allows these nuclei to relax via emission of their own RF waves. Detection of the released energy as a function of time produces Free Induction Decays (FID), which can undergo a Fourier transformation into the frequency domain.[87] A collected NMR spectrum contains a series of peaks at specific frequencies (chemical shift, δ), corresponding to individual nuclei and their surrounding chemical space. As the hydrogen atom is NMR active, acquisition of spectra from an aqueous sample would lead to masking of a samples innate ^1H response due to the hydrogen in water molecules. To combat this, samples are prepared in deuterated solvent (D_2O, CCl_3D *etc*), typically deuterated NMR solvents contain an internal control, 3-trimethylsilyl-deuterosodium propionate, to provide a known deuterium chemical shift, which can be used to calibrate other observed frequences.[87]

NMR excels at analyses of small compounds, where modest numbers of nuclei are present. Proteins contain very high numbers of active nuclei in a multitude of chemical environments, and therefore generate highly complex spectra. Typically, multiple NMR measurements are taken of a multiply radioisotope labelled protein generating chemical shifts for ^1H, ^{13}C and ^{15}N, which allows specifically for analysis of the peptide backbone.[91] Additional techniques, including the Nuclear Overhauser effect (NOE) can aid the 3D geometric positioning of nuclei by imparting distance constraints between coupled nuclei.[92]

NMR analyses of proteins was historically limited to modest (<35 kDa) protein sizes. However, through improvements in high-field magnets, simultaneous multi-dimensional acquisition processes and cryogenic probes have pushed this limit to >180 kDa.[92–94] NMR is highly penetrating and non-disruptive, allowing for so-called *in cellulo* analyses of over-expressed protein structures and importantly, reports solution-phase structures allowing for true conformational dynamics to be observed.[95–97] However, measurements are hampered by difficult data analysis, proportional to protein size, and high concentrations of radioisotopically labelled protein.[97]

1.4. Mass Spectrometry

Mass spectrometry (MS) is an invaluable analytical technique in chemistry, biology, and their interwoven fields of study. Developed by Joseph Thompson alongside Francis Aston in the early 20th century, mass spectrometry employs the separation of gas-phase ions according to their mass-to-charge ratio (*m/z*), through interactions with electric and magnetic fields.[98,99] In the following century, mass spectrometry has progressed through analysis of elemental isotopes, amino acids, individual proteins, to megadalton protein complexes.[100–103] In its most basic form, a mass spectrometer is composed of an ionisation source to produce charged, gas-phase ions, a mass analyser to perform *m/z* separation, and a detector to quantify the ions. Mass spectrometry has proven invaluable to the study of protein interactions, probing ligand binding sites, conformational changes and stoichiometries.[104–106] It is highly sensitive, requiring low concentrations and volumes of sample, and when coupled with chromatography (liquid, LC, or gas, GC) allows for highly accurate analyses of complex protein digests.[107]

1.4.1. Ionisation

Ionisation sources generate gas-phase ions from the analyte sample, which may be solid, liquid or itself gaseous. Electrospray ionisation (ESI), and atmospheric pressure ionisation (chemical ionisation, APCI, photoionisation, APPI) are reserved for liquid samples. Here the analyte solution is nebulised at atmospheric pressure and then introduced into the mass spectrometer. Matrix-assisted laser desorption ionisation (MALDI), desorption electrospray ionization (DESI) and secondary ion mass spectrometry (SIMS) are used to ionise solid-state samples, however DESI can also be applied to liquid surface samples. Electron ionisation (EI) and chemical ionisation (CI) are typically used to ionise gas-phase samples. Ionisation sources can be broadly classified into *hard* or *soft* techniques, based on how much energy they impart into the sample to generate ions. Hard ionisation typically results in analyte ion fragmentation, which may be beneficial to some applications, whereas soft ionisation imparts much lower energy into the analyte, resulting in minimal fragmentation.[107] Both mass spectrometers used here were equipped with electrospray ionisation sources, which is the focus herein.

1.4.1.1. Electrospray Ionisation

Electrospray ionisation, as a soft ionisation source, is an excellent technique for producing ions from biological samples. Soft ionisation can preserve fragile, non-covalent interactions that may otherwise be lost via hard ionisation, which facilitates the study of protein-protein/ligand complexes in the gas-phase.[108] Fundamental ESI was demonstrated in the early 20th century where solvent held in capillary tubes was subject to a strong electric potential of up to 10 kV by John Zeleny.[109] Observation showed droplet emission occurred from the tip of cone, formed by a distortion of the liquid meniscus. This was further characterised in the mid-century by Geoffrey Taylor, leading to the widely accepted *Taylor cone* nomenclature (figure 1.12, A).[110] For ionisation by ESI, a non-volatile analyte is dissolved into a volatile solvent (ammonium acetate, water, acetonitrile *etc*) and passed through an emitter capillary. Here, a high electric potential is applied (1 – 4 kV) which causes accumulation of positively charged ions towards the meniscus of the emerging droplet. Eventually, the coulombic repulsion between the accumulated charges overcome the droplet surface tension and form the Taylor cone. Continuous extension of the cone tip results in the expulsion of positively charged droplets, which are accelerated by the electric potential, towards the mass spectrometer inlet. As the droplets are released, they begin to desolvate, often aided by a stream of dry gas and/or gentle heating. During desolvation, the charge density of the droplets increases causing a rise in coulombic repulsion. Once this repulsion reaches the *Rayleigh limit*, the surface tension of the droplet cannot contain the charge, and the droplets explode into several smaller daughter droplets (figure 1.12, A).[111]

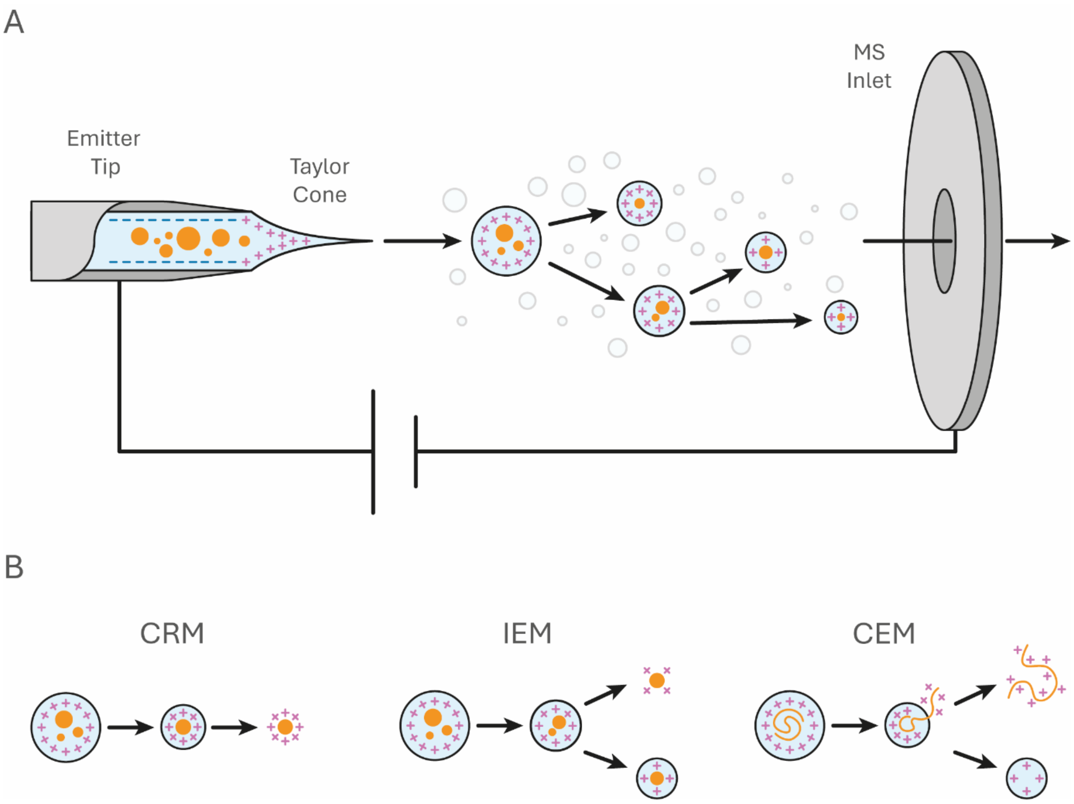


Figure 1.12 | Graphical representation of the principles behind ESI droplet formation in positive ion mode. A | Taylor cone formation, droplet formation, and coulombic explosion. B | Representation of analyte ion ejection mechanisms.

Typically, each daughter droplet has approximately a 7-fold greater charge density compared to the mother droplet, which aids rapid further coulombic fission.[112] This is governed by equation 1.1, where z_R is the total number of charges, e is the elementary charge (1.60×10^{-19} C), ε_0 is the relative permittivity, γ is the surface tension and D the diameter of the droplet.

$$z_R e = \pi\sqrt{8\varepsilon_0 \gamma D^3} \qquad (1.1)$$

The process of droplet desolvation and fission continue, producing smaller and smaller droplets, eventually containing a single charged analyte ion, until the ion is ejected. This ejection takes place through one of three hypothesized methods: ion evaporation model (IEM), charge residue model (CRM) or chain ejection model (CEM) (figure 1.12, B).[113] IEM postulates that the analyte ion is ejected from the droplet prior to complete evaporation due to coulombic repulsion. This method favours low molecular weight ions that eject carrying a low number of charges. In contrast, higher molecular weight ions, such as natively structured proteins, are more favoured to eject via CRM. Here, the droplet undergoes complete evaporation and deposits all of its carried charges onto the analyte ion, resulting in multiply charged analyte ions.[114] These ions have total charges similar to that of the Rayleigh limit of an equivalently sized water droplet.[114–116] CEM explains how a denatured protein, or similar large polymeric compound, ejects from a droplet in a higher charged state than the equivalent native (compact) structure. A denatured protein is able to hold greater numbers of discrete charges, due to its extended conformation. As the Rayleigh limit is reached, droplet charges and analyte charges migrate to opposite sides of the droplet, which causes a perturbance of the spherical shape. A terminus of the protein is typically ejected from the droplet which pulls away from the solvent, stretching the polypeptide into a tail-like structure and eventually pulling the entire polypeptide out.[116] This behaviour has been observed experimentally, and through atomistic molecular dynamics simulations of denatured myoglobin within Rayleigh limit charged solvent droplets.[115]

1.4.1.2. *Nano-Electrospray Ionisation*

Nano-electrospray ionisation (nESI) is a low-flow (nL min^{-1}), low-volume form of ESI (μL min^{-1} – mL min^{-1}), made possible by the development of emitters with sub-micrometre orifice diameters. This results in much smaller initial solvent droplets, which in turn reduce adduction from salt ions due to lower numbers of salt ions present in the initial droplets.[117] nESI sources typically operate at lower electrical potentials (0.5 – 2 kV), which further aid the stability of delicate non-covalent interactions in the gas-phase.[118] The low-flow nature of nESI, however, does lead to a propensity for unstable flow, and the much smaller emitter size are liable for blockage.[119] As stated, ESI and more-so nESI are preferred methods for the study of intact protein in the gas-phase. Native-like protein ions are generated with Rayleigh limit-approximate charge densities, resolved in a mass spectrum as $[M + zH]^{z+}$ peaks with total charges, and charge state distributions, influenced predominantly by the conformation of the protein ion. Protein samples for ESI and nESI should be prepared in such a way to minimise potential adducting cationic ions, namely sodium and potassium, though desalting or dialysis. Adduction of protein ions by non-proton cations leads to reduced signal intensity and resolution, due to the now heterogenous species of ions.[120]

1.4.2. Mass Analysers

After ionisation, gas-phase ions enter a mass analyser, where they are subsequently separated according to their *m/z*. There are many types of mass analysers which use differing methods to perform *m/z* separation. For this work, four mass analysers will be discussed. Quadrupole analysers exploit an ions trajectory stability at a given applied RF and direct current (DC) voltage to selectively transmit ions of a given *m/z*. Time-of-flight (TOF) analysers use an ions velocity to separate them, as for a given accelerating voltage, ion velocity is a function of their *m/z*. Linear ion trap (LIT) analysers are, as the name suggests, trapping analysers. Here, the continuous beam of ions is interrupted, and instead ions are trapped in discrete quanta. Trapped ions are subject to sweeping RF and DC voltages, which causes ions of a given *m/z* to be ejected from the trap towards a detector, not dissimilar to quadrupole stability. Finally, ion cyclotron resonance (ICR) analysers exploit the intrinsic cyclotron motion of charged ions in a magnetic field, which is a function of *m/z* to perform separation.

1.4.2.1. Quadrupole

The quadrupole mass analyser was developed by Wolfgang Paul and Helmut Steinwedel, hence the alternate name of quadrupole ion traps as *Paul traps*, in the 1950's and since then have become a staple in mass spectrometers.[121] Quadrupoles consist of four parallel metal rods, between which ions are transmitted (figure 1.13). Ideally, each rod would have a hyperbolic cross section, but highly precise circular cross sections function adequately. Each rod is pairwise connected to the opposite and a variable superimposed RF/DC potential is applied, such that each pair experiences the opposite polarity. This generates an oscillating electric field, through which the ions propagate.

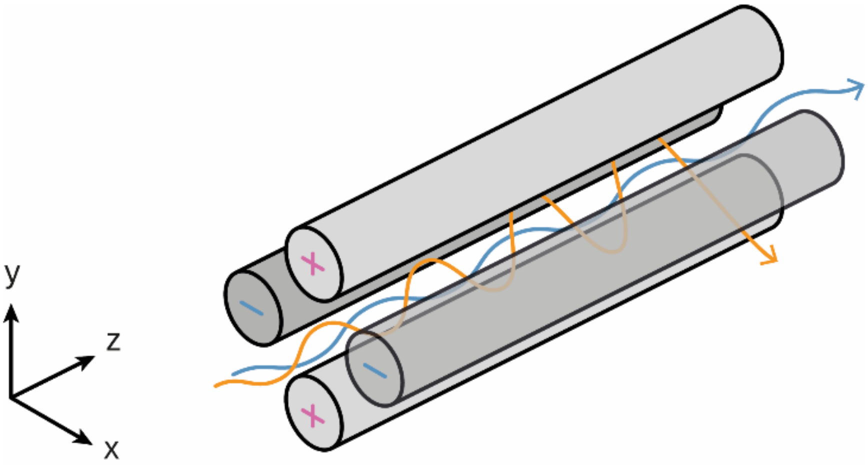


Figure 1.13 | Graphical quadrupole schematic. A stable (blue) and unstable (orange) ion trajectories are shown.

The motion of the ions can be described by equations 1.2 – 1.4, where $V\cos(\omega t)$ is RF (ω is angular frequency, t is time), z is the ion charge, e the elementary charge, r_0 the radius between the quadrupoles and U, the DC voltage.

$$\frac{d^2x}{dt^2} = -\frac{ze(U - V\cos(\omega t))}{mr_0}x \tag{1.2}$$

$$\frac{d^2y}{dt^2} = +\frac{ze(U - V\cos(\omega t))}{mr_0}y \tag{1.3}$$

$$\frac{d^2z}{dt^2} = 0 \tag{1.4}$$

Ion trajectory is only influenced by the quadrupole in the x and y planes, and not in the axial z plane. Ions oscillate back and forth within the x,y plane and remain stable so long as their position does not exceed r_0. The ion positions, therefore, trajectories, demonstrate a relationship between ion *m/z* and the RF/DC potentials. Substituting the above equations with the Mathieu function (equation 1.5) allows for calculation of *stability regions* for ions of a given *m/z*. Where ζ, a, and u are dimensionless parameters, which can be derived from real-world parameters (expressions 1.6), within equations 1.2 and 1.3:

$$\frac{d^2u}{d\zeta^2} + (a - 2q\cos(2\zeta))u = 0 \tag{1.5}$$

$$\zeta = \frac{\omega t}{2} \qquad a = \frac{4zeU}{mr_0\omega^2} \qquad q = \frac{2zeV}{mr_0\omega^2} \tag{1.6}$$

First, rearrangement of equation 1.2 and 1.3 to equations 1.7 and 1.8 respectively:

$$\frac{d^2x}{dt^2} + \frac{ze(U - V\cos(\omega t))}{mr_0}x = 0 \tag{1.7}$$

$$\frac{d^2y}{dt^2} - \frac{ze(U - V\cos(\omega t))}{mr_0}y = 0 \tag{1.8}$$

Which, when substituted with expressions 1.6 become:

$$\frac{d^2x}{d\zeta^2} + (a - 2q\cos(2\zeta))x = 0 \tag{1.9}$$

$$\frac{d^2y}{d\zeta^2} - (a - 2q\cos(2\zeta))y = 0 \tag{1.10}$$

Solutions to the Mathieu function determine bounds of stability in the *a,q* plane and as can be seen in expressions 1.6, a relates to DC and q to RF potentials respectively. From this the region of stability for an ion of given *m/z* can be found in the DC,RF plane (figure 1.14, A). By scanning the quadrupole through a RF and DC range, stable ions are sequentially transmitted through the quadrupole. Ions whose regions of stability intersect this scan line are transmitted, and if there is no overlap of ion stability regions, the ions will be fully resolved (figure 1.14, B). Resolving power can be increased by using a steeper $\frac{U}{V}$ scan, where the regions of intersection are smaller, thus reducing overlap at the cost of sensitivity. Note, that if the DC potential is held at 0 V and only RF potential is engaged, all ions stable below the given RF potential will be transmitted (effectively a resolving power of zero). This is sometimes used in *quadrupole ion guides* to transmit ions within instruments.

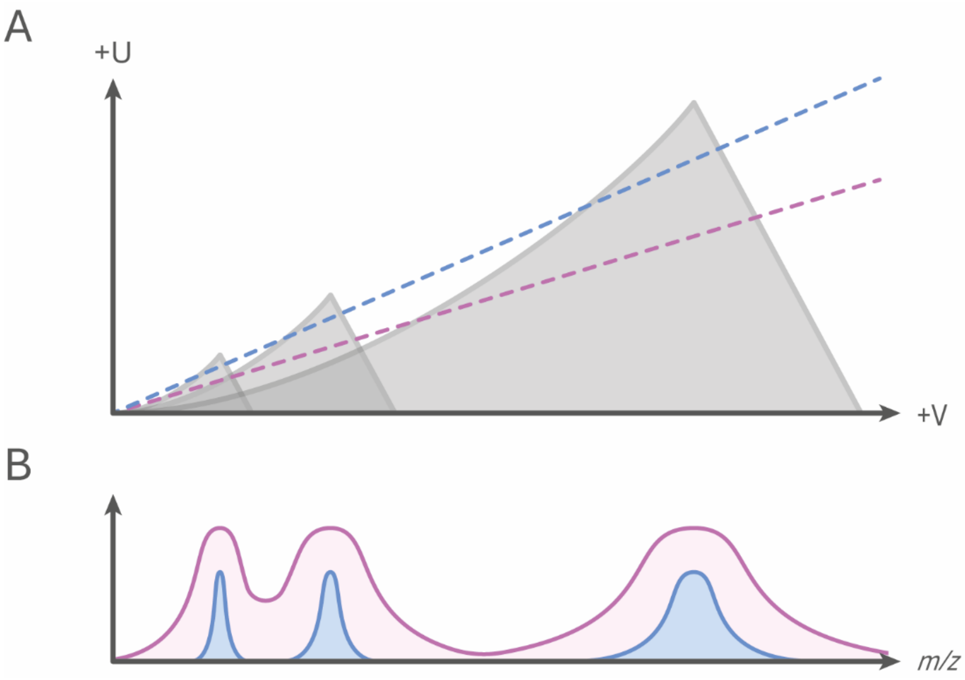


Figure 1.14 | Quadrupole induced regions of stability. A | Areas of stability for three individual ions with increasing *m/z*. Two different scan lines are shown (blue and pink). B | Resulting mass spectrum of the three ions, ions with overlapping regions for a given *U/V* scan are not fully resolve (pink). Increasing *U/V* increases resolution, at the cost of sensitivity (blue).

1.4.2.2. *Time-of-Flight*

TOF analysers, developed in the late 1940's by William Stephens, separate ions based upon their velocities after acceleration by a given potential, external to any further electric or magnetic field influences.[122] Ions of the same m/z, accelerated under the same potential, V, have the same nominal kinetic energy E_k, and therefore the same velocities, v, given by equation 1.11. Acceleration of the ion imparts a potential energy (E) to it, equal to its kinetic energy.

$$E = zeV = E_k = \frac{1}{2}mv^2 \tag{1.11}$$

Where z is the ion charge, e the elementary charge and m the ion mass. Velocity is equal to distance as a function of time, with distance in this instance being the length of the flight tube, l. Substitution into equation 1.11 yields equation 1.12 which can be rearranged to solve for time, t, given in equation 1.13.

$$zeV = \frac{1}{2}m\left(\frac{l}{t}\right)^2 \tag{1.12}$$

$$t = l\sqrt{\frac{m}{2zeV}} \tag{1.13}$$

In a TOF analyser, a continuous ion beam accelerated into the flight tube, held under high vacuum is first converted into discrete packets of ions, using an orthogonal pusher. Ions accumulate in the source region of the TOF, before they are accelerated and released by the pusher into the flight tube towards the detector. In reality, ions of equal *m/z* enter the flight tube with variant kinetic energies due to unequal acceleration imparted by the pusher. Because of this, ion arrival times become distributed, leading to limited resolution. To combat this, ion mirrors, or *reflectrons,* can be introduced (figure 1.15). Here, the ions are directed into a stacked set of ring electrodes, held at increasing potentials. Ions with the same *m/z* but differing velocities, penetrate the reflectron to different depths before their velocities become zero, and they are accelerated out of the reflectron towards the detector. As the ion with the higher velocity penetrates deeper into the mirror, it spends more time there prior to acceleration out. The end result is a focussing of arrival times by correction of the variant kinetic energies which therefore increases resolution.

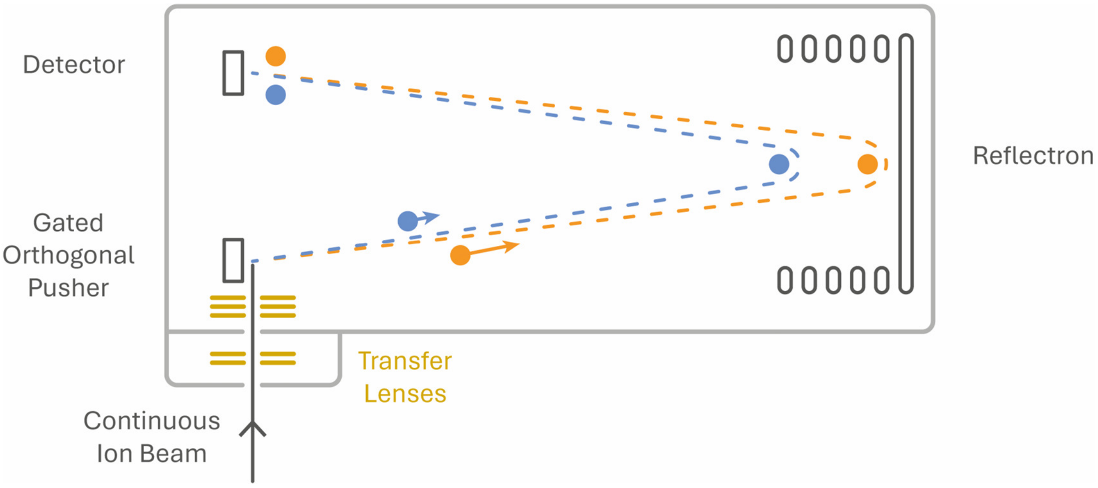


Figure 1.15 | Graphical overview of a single reflectron TOF. Two ions of equal *m/z* have accelerated to different velocities. The faster (orange) penetrates deeper into the reflectron than the slower (blue) such that they arrive at the detector with a much smaller deviance,

1.4.2.3. Linear Ion Trap

LIT analysers act in principle similar to quadrupole analysers, where ions enter into a hyperbolic quadrupole, with opposing RF/DC potentials applied to the opposite pairs, constraining the ions in the *x,y* plane. Unlike quadrupole analysers, LIT's have additional endcap rods, to which a DC potential is applied to trap the ions in the *z* axis (figure 1.16). Typically, ions injected into the LIT from a continuous beam, by reducing the potential of the "front" endcap. Ions enter the trap and are reflected by the maintained potential of the "back" endcap, at this point ions will have lost kinetic energy due to collisions and therefore lack the energy to overcome the potential of the front endcap, trapping them inside. Ions process in the *x,y* plane within the trap with the motion described by equations 1.9 and 1.10, and along the *z* axis in a forward and reverse motion. By altering the applied RF potential, ions of given *m/z* become unstable in the *x,y* plane. This causes them to be ejected from the trap towards a detector, though slits in the *x* axis rods. Individual ions of interest can be isolated in a LIT through selective broadband RF sweeping. By applying RF potentials to specifically destabilise all but the ion of interest, unwanted ions can be ejected from the trap. This is of particular importance for performing tandem mass spectrometry, where an ion can be isolated, fragmented and the resulting fragment ions ejected from the trap towards a detector.

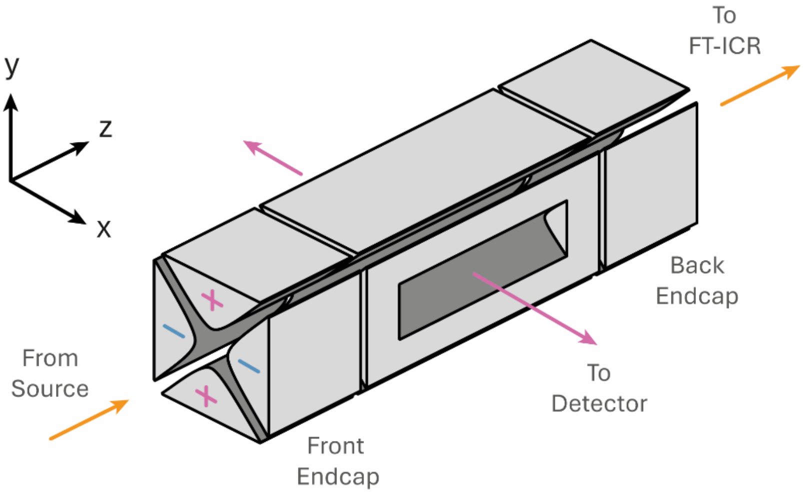


Figure 1.16 | Graphical representation of the LIT mass analyser inside a LTQ FT Ultra Mass Spectrometer, where the LIT is coupled with a FT-ICR mass analyser.

1.4.2.4. Ion Cyclotron Resonance

ICR mass analysers, or more commonly Fourier transform ICR (FT-ICR) analysers, were developed in the 1970's by Melvin Comisarow and Alan Marshall.[123,124] The analyser consists of a cylindrical cell composed of four separate plates, two RF excitation plates and two detection plates, capped by an additional two trapping plates (figure 1.17). Unlike the other discussed mass analysers, an FT-ICR cell is both the *analyser* and the *detector,* removing the need for a dedicated detector in a solely FT-ICR-based instrument. The FT-ICR cell is held under high-vacuum and seated within a superconducting magnet. Ions, with a velocity vector, v, and charge, ze, entering the cell perpendicular to the magnetic field vector, B, experience the Lorentz force, F_B, (equation 1.14), causing them to adopt deflected trajectories around B.

$$F_B = zevB \tag{1.14}$$

If the kinetic energy of the ions is sufficiently small, they adopt circular orbits perpendicular to B with a cyclotron frequency, ω_C, in rads^{-1} as per equation 1.15. Which can be re-expressed to give frequency, f, in hertz (equation 1.16), where m is the ion mass and r is the orbit radius.

$$\omega_C = \frac{zeB}{m} = \frac{v}{r} \tag{1.15}$$

$$f = \frac{zeB}{2\pi m} \tag{1.16}$$

These show that in a constant magnetic field, the frequency of the ion's orbit is inversely proportional to its m/z. These frequencies, typically in the high kHz – mHz range, correspond to the RF range and can therefore be modulated by application of an RF field. In operation, ions are guided into the FT-ICR cell by hexa- or octapole trapping ion guides, which allow for accumulation of discrete ion packets prior to entry into FT-ICR cell.[107,119] Entry is permitted by reducing the DC potential of the front trapping plate and releasing the ion packet from the trapping guides. Once in the FT-ICR cell, the initial orbit of each ion has too small a radius for them to pass by the detector plates. By application of a broadband RF pulse, all ions are excited to a higher radius orbit, permitting detection. As ions pass by the detection plate, the induced change in current is recorded. Removal of the RF pulse allows the ions to relax to their ideal orbit radii, as they move further from the detection plates, a transient (or free induction decay) is produced. This is a complex sine wave composed of the sum of each individual ions ω_C over time, which can undergo a

Fourier transformation into the frequency domain and thus the individual ions *m/z* determined through equation 1.16, (figure 1.17).

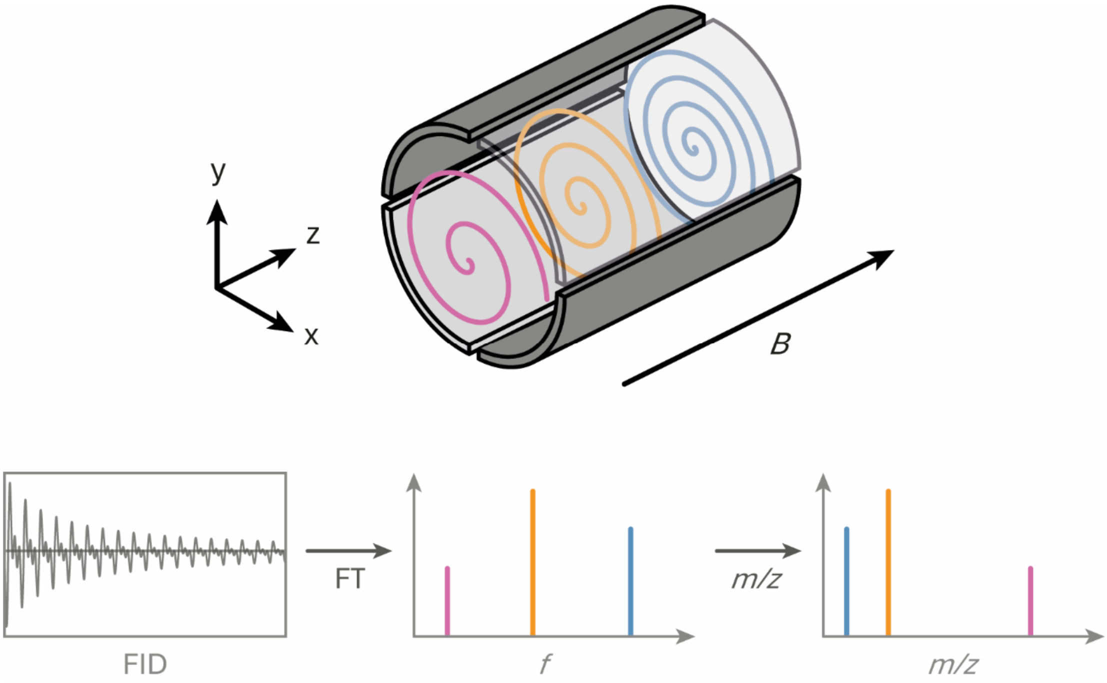


Figure 1.17 | Graphical representation of an FT-ICR mass analyser (dark grey plates are detector, light grey are excitation). Front and back trapping plates are omitted for clarity. Three ions of different *m/z* (blue, pink and orange) are relaxing to their ideal obits. Below are their summed FIDs which are Fourier transformed into the frequency domain and then solved for the *m/z* domain through equation 1.16.

1.4.3. Tandem Mass Spectrometry

Mass spectrometry-based analyses often include steps to isolate and fragment precursor ions to garner further information. In proteomic studies, specifically top-down analysis and *de novo* sequencing, this is typically the isolation of a protein or peptide ion followed by fragmentation, yielding information of their amino acid make up.[125] Tandem mass spectrometry (MS/MS) refers to the coupling of multiple mass analyses together in series, where each n^{th} sequential analysis is referred to as MS^n. Typically, ion isolation is performed in one mass analyser, and a separate one is used to analyse the resulting fragment ions. The most common of these being the tandem-quadrupole (typically triple-quadrupole) instruments. Here, the first and third (Q_1 and Q_3) quadrupoles are used as mass analysers, and the second (q_2) as a collision cell for ion activation. However, as discussed previously, mass analysers such as LITs or FT-ICRs can perform both roles.

1.4.3.1. Hybrid Mass Spectrometers

Hybrid mass spectrometers contain multiple *different* mass analysers within the same instrument, with the most well-known being the quadrupole-time-of-flight analysers (Q-TOF). Two hybrid instruments were used in this work, a Waters Synapt G1 (Q-TOF) and a Thermo Scientific LTQ FT Ultra (LIT-FT-ICR). In a Q-TOF (typically a Q_1q_2-TOF) instrument the resolving Q_1 quadrupole can transmit the full range of ions, or a select *m/z* window through to the q_2 collision cell, where ions are activated and fragmented. The fragment ions are then analysed by a TOF. In the Synapt G1, q_2 is replaced by three individual travelling wave, stacked-ring ion guides (as opposed to hexapole ion guides). The second of which functions as a mobility cell for ion-mobility spectrometry measurements. This is sandwiched between two trapping collision cells, which can be used to fragment ions pre- or post-mobility separation. The LTQ FT Ultra is a combination of a LIT and a FT-ICR mass analysers, allowing for isolation and fragmentation by multiple methods. The LIT provides the benefit of scan speed and ion accumulation, whilst being able to isolate and activate ions. The resulting fragments can either be analysed by the LIT or passed into the FT-ICR cell to take advantage of the ultra-high resolving power conferred by ICR. Ions of interest can be activated in multiple ways; collision induced dissociation (CID), electron capture dissociation (ECD), electron transfer dissociation (ETD) and infrared multiphoton dissociation (IRMPD). ECD and IRMPD are typically used in FT-ICR cells, where the introduction of collision gas (CID) or radical ion precursors (ETD) into the high vacuum of the FT-ICR would be deleterious to its operation. CID was the method of ion activation used in all work herein.

1.4.3.2. Collision Induced Dissociation

During CID ion activation, analyte ions of interest are accelerated to a high kinetic energy before colliding with a neutral gas, typically molecular nitrogen, helium or argon. CID can be broadly split into either high or low energy depending on the kinetic energy (E_K) of the ion (high energy, E_K >100 eV and low energy <100 eV). The energy from the collision imparted on the ion, E_{com}, is given by equation 1.17, with kinetic energy derived from equation 1.11. Where m_N is the mass of the neutral collision gas, and m_P the mass of the ion.

$$E_{\mathrm{com}} = \left(\frac{m_N}{m_N + m_P}\right) E_K \qquad (1.17)$$

The collisional energy, E_{com}, is distributed throughout the internal bonds of the ion as vibrational energy. In low-energy CID, multiple collision events allow this internally distributed energy to build over time until the weakest bond is broken, and fragment ions dissociate. The rate of dissociation is much slower than this internal distribution of energy, and the fragmentation is said to be *ergodic*.[126] This is opposed to high-energy CID, where sufficient energy is imparted into the precursor ion to induce bond breaking quicker than it can dissipate through the ion, leading to *non-ergodic* fragmentation.[127]

In protein or peptide ions, low-energy CID of protonated ions typically induces breaking of weak covalent modifications first, followed by the kinetically weak peptide bond leading to the formation of *b* and *y* ions (figure 1.18). In peptides containing basic residues, the proton affinity of the sidechain facilitates location of the charge to fragments containing those residues and causing them to dominate in the resulting MS^2 spectrum.[128] In peptides resulting from a tryptic digestion, with C-terminal lysine and arginine, MS^2 spectra are typically dominated by *Y* ions, as they originate from the basic residue containing C-terminus. Gentle CID can also be used to induce dissociation of non-covalent complexes from native protein ions, these can be small molecules, or protein subunits from multimeric complexes.[129]

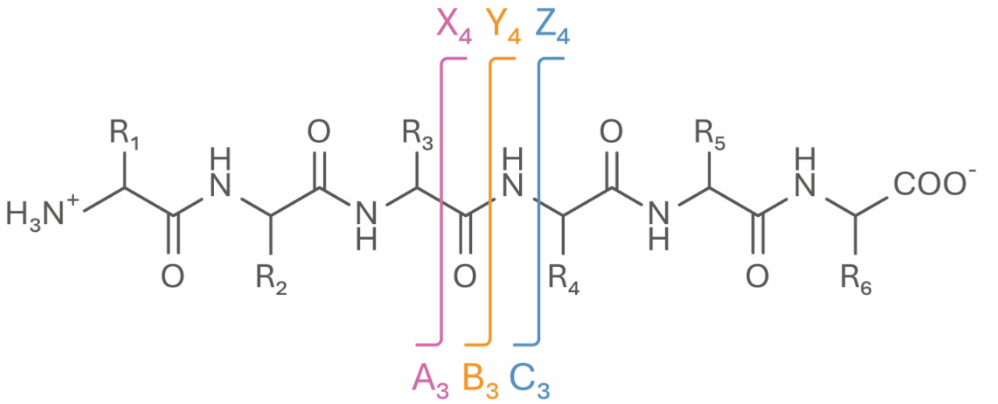


Figure 1.18 | Representative hexapeptide showing fragmentation sites for generation of A + X, B + Y and C + X ions in a MS^2 analysis.

1.4.4. Native Mass Spectrometry

Native mass spectrometry (nMS) is broadly described as any mass spectrometric technique that permits the transmission, and study of, native-like protein ions in the gas-phase. Importantly this maintains non-covalent interactions, be them between protein subunits in a quaternary structure, or complexes of protein and small molecule. As mentioned previously, ESI and nESI sources are preferred due to their gentle ionisation of protein macromolecules into the gas phase.

Through the use of non-denaturing conditions; including neutral pH; non-denaturing solvents, such as ammonium acetate; and modest source temperatures and energies, protein ions can be produced maintaining their native, solution-like conformations.[130–132] Graphical representations of native-like and denatured protein mass spectra can be found in figure 1.19, A and B respectively). Retention of a truly native state in the gas-phase is impossible, the loss of a solvent environment causes extremely quick (picosecond level) collapse of charged side-chains, causing an inevitable collapse of the protein to a more compacted structure.[133] However, it is known that gaseous protein ions do maintain non-covalent interactions through IR spectroscopic analysis of gas phase ions, and that overall conformations are similar to "native" crystal structures through ion mobility spectrometry-mass spectrometry experiments.[134–136] Conventional native mass spectrometry using Q-TOF instruments has recently been used to study large protein complexes, including the approximately 700 kDa, eukaryotic 20S proteasome and 450 kDa assemblies of immunoglobulins.[137,138] Pioneering work by the Carol Robinson group has allowed for the study of membrane proteins and complexes from their native membrane vesicles, and has opened up a whole field of "nativeomics" where endogenous ligands and lipids of membrane-bound proteins can be analysed through nMS.[139,140] Through complexing of nESI sources to Orbitrap mass analysers, the intact 2 MDa 70S ribosomal complex has been detected, along with its subsequent CID induced dissociation into the 30S and 50S subunits.[141]

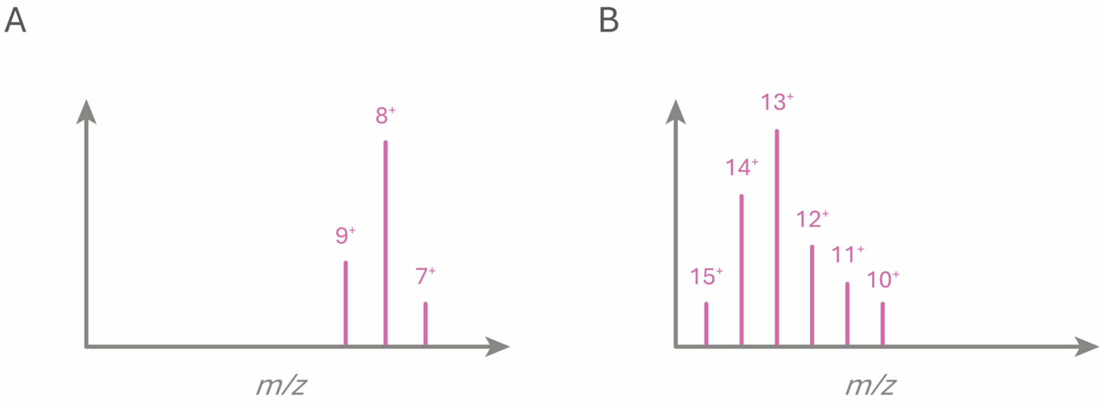


Figure 1.19 | Graphical representations of denatured (A) and native-like (B) protein mass spectra. The extended denatured conformation facilitates greater charging during ionisation, and therefore a reduced *m/z* distribution compared to a globular, native-like conformation.

1.4.4.1. Ion Mobility Spectrometry-Mass Spectrometry

Ion mobility spectrometry (IMS) is a technique where gas-phase ions are separated based on their charge, size and shape. The latter two characteristics combine to give a *collisional cross-section* (CCS) of an ion, which is the rotationally averaged combination of its cross sections. IMS is frequently used as part of a hybrid instrument, where analysis of ions separated by IMS is performed by a mass spectrometer (IMS-MS). During IMS analysis ions are accelerated into a pressurised IMS cell, typically either a hexapole or stacked-ring ion guides, where they interact with an inert collision gas (typically molecular nitrogen or helium). The resistance each ion experiences travelling through the cell is a function of their CCS, where larger ions interact more and therefore travel slower through the cell, this is said to be an increase in *drift time*.[142] Because of this, two native protein ions with the same *m/z*, but different conformations (say differing levels of denaturation) can be separated by their differences in drift time.[143] In a standard drift time spectrometer, ions move through a homogenous electric field, E, with a velocity, v, against a counter flow of collision gas. Their motility, K, through the cell is defined by equation 1.17.

$$K = \frac{v}{E} \tag{1.17}$$

Motility is inversely proportional to CCS (Ω) as given by the Mason-Schamp equation, where z is the ion charge state, e, the elementary charge, N, the collision gas number density, μ, the reduced mass of collision gas and ion ($\frac{Mm}{M+m}$), K_B the Boltzmann constant (1.38×10^{-23} JK^{-1}) and T, the temperature of the collision gas (equation 1.17).[144]

$$\Omega = \frac{ze}{K}\frac{3}{16N}\sqrt{\frac{2\pi}{\mu K_B T}} \tag{1.18}$$

This holds true for standard drift time spectrometer, however the Synapt G1 contains a travelling wave ion mobility separation (TWIMS) cell, who's basis of operation is a constantly changing electric field. Because of this Ω is no longer inversely proportional to K, and instead the instrument must be calibrated to known literature to perform CCS calculations.[145]

1.4.4.2. Travelling Wave Ion Mobility Separation

In a TWIMS cell, a travelling DC potential is applied along the stacked-ring ion guides to propel ions through the cell (figure 1.20, A). A RF potential is applied to each ring, with adjacent rings have opposite RF phases to radially focus ions. To move the ions, a DC potential offset is superimposed over the RF potential of the first pair of two (or more) rings, this then travels along the rings as a group, moving in one (or more) ring increments (figure 1.20, B).[142,144]

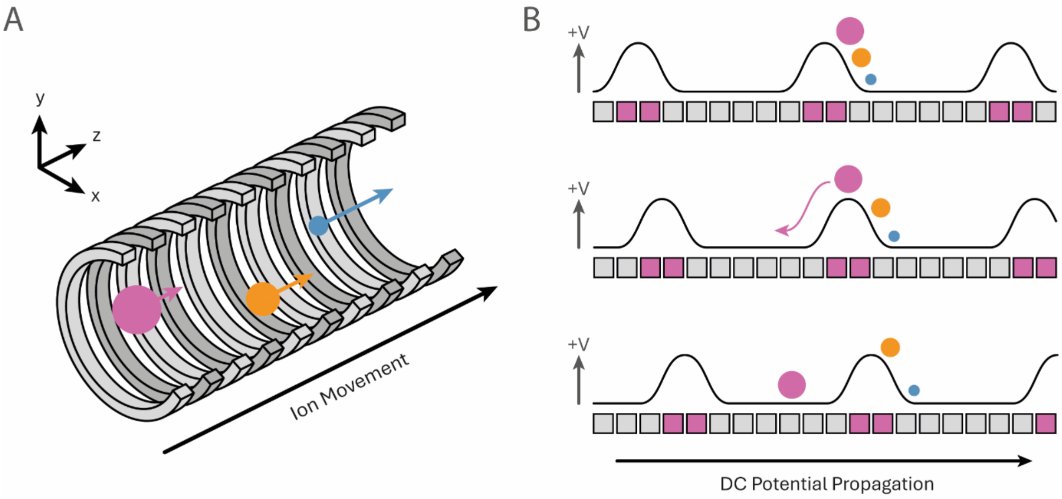


Figure 1.20 | Graphical overview of TWIMS. A | Ion movement through the stacked-ring ion guide. Adjacent rings have opposite RF potentials. Larger ions (pink and orange) move slower than smaller ions (blue). Collision gas not shown for clarity. B| Principle of travelling wave movement. DC potentials applied to groups of two rings (pink) propagate one ring at a time. Larger ions move slowly and can gradually roll over the waves.

Analyte ions are pushed through the cell by the movement of the potential wave (effectively surfing), interacting with a stationary collision gas. Ions of a greater CCS are slowed down and may roll over the top of the wave, impeding their travel and allowing for seperation.[146,147]

1.4.4.3. Collision Induced Unfolding

Native protein ions, collisional activated, may adopt transient, partially unfolded conformations.[143] As drift times are within millisecond time scales, conformations short lifespans can be separated and resolved by IMS-MS. Collision induced unfolding (CIU) analysis involve subjecting a native protein ion to a ramped activation energy, prior to IMS and monitoring the changes in drift time as a result. Here, the stabilising influences of ligands or, conversely, the destabilising influences of residue mutations can be inferred by the propensity of a native protein ion to unfold at higher or lower activation energies respectively, compared to an *apo*-protein. [148–151] CIU has also been applied to separate heterologous mixtures of immunoglobulins, due to differing disulfide bonding and glycosylation patterns imparting differing stability on ions with unresolvable differences in CCS.[152]

1.5. MS-Based Protein Labelling

Protein labelling, or *footprinting*, paired with mass spectrometry analysis have been used to directly study protein interactions since the late 1990's, where acetylation of calmodulin lysine residues showed which regions were masked upon binding with melittin.[153] This was followed by development of hydrogen-deuterium exchange (HDX) to monitor the exposure of amide hydrogens to bulk solvent, initially through nMS but soon followed by peptide analysis.[154,155] Typically, protein footprinting experiments follow a *bottom-up* proteomic workflow which, in brief, consists of an enzymatic digest of a protein (which may be first resolved by SDS-PAGE) into constituent peptides. The peptides are then subject to chromatographic separation, often through reverse-phase high-performance liquid chromatography (RP-HPLC) coupled with ESI-MS for analysis. Peptides can be fragmented (section 1.4.3.2) and, if the protein sequence is unknown, searched against large databases for identification (*de novo* sequencing).[156]

Protein footprinting is usually performed on an *apo*- and *holo*-protein complexed with a binding partner. By binding, a ligand impacts a shielding effect on the protein, preventing labelling in the binding site (figure 1.21). When comparing *apo*- and *holo*-proteins, binding sites can be identified through this reduction in labelling. Direct ligand masking is not the only cause of differential labelling, ligand modulated protein conformational dynamics can be reported through the increase or decrease in labelling of specific regions as they become either more or less exposed. MS^2 level analysis can be performed to produce sub-peptide level footprinting resolution and elucidate labelling of specific residues. Broadly, protein footprinting techniques can be split into either *reversible* or *irreversible* depending on the nature of the labelling probe. Carbene, hydroxyl radical and trifluoromethyl radical footprinting are all irreversible methods, whereas HDX is reversible.

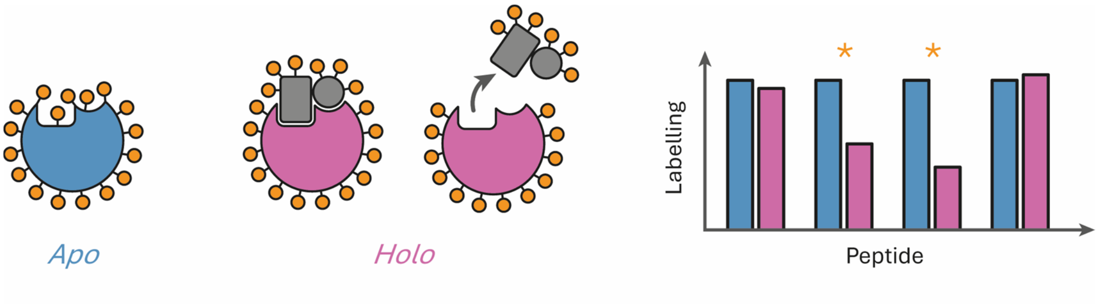


Figure 1.21 | Graphical representation of protein footprinting. The presence of a bound ligand in the *holo* protein imparts a masking effect (left), reducing peptide labelling in those regions involved in ligand binding (orange asterisk, right).

1.5.1. Hydrogen-Deuterium Exchange

The principle behind HDX is the isotopic exchange of amide hydrogens to deuterium through exposure of a protein to a deuterated solvent.[157] Although most hydrogen atoms not involved in hydrogen bonding are amenable for exchange, the amide hydrogen is ubiquitous through all residues (except proline) and exchanges at a rate to facilitate MS analysis.[158,159] Exchange is rate limited by both the solvent accessibility of the amide, and by its hydrogen bonding network. This yields spatial information about specific amide, and its structural characteristics. Solvent exposed secondary structure elements tend to exchange slowly due to the amide hydrogen bonds responsible for their structure.[159] Initially NMR was used to interrogate the deuterated protein, with deuterium having an active NMR nucleus, this has been slowly phased out in favour of MS-based workflows but does still see some use, providing powerful residue-level deuterium localisation.[160,161] The exchange of each hydrogen to deuterium imparts a +1 Da mass shift, which is easily detectable in MS through a change in the ion isotopic distribution.

A typical HDX experiment involves the incubation of a protein in D_2O held at a neutral pH to induce exchange. To quench this reaction, the pH is dropped to below 2.5 and the sample held at 0°C, under these conditions the rate of exchange is approximately 10^5 orders of magnitude slower.[159] These quenching conditions must be maintained as much as possible to minimise the unwanted back-exchange to H_2O. Refrigerated HPLC separation of the deuterated samples can be performed immediately, or flash-freezing in liquid nitrogen and cryogenic storage is possible for un-coupled HPLC-MS analysis.[162] Bottom-up proteomics for HDX is challenging, the quenching conditions restrict enzyme selection to acid-stable proteases, classically pepsin however more commercially available acid-stable proteases are becoming available.[163,164]

Quenching is performed at multiple timescales, with modern automated, in-line systems reactions can be quenched at 1 millisecond intervals, allowing protein dynamics to be captured in unprecedented temporal resolution.[165] This paired with highly complex, automated instrumentation and data analysis has greatly minimised many downsides of HDX.[165,166] One area in which HDX falls short, however, is the limited MS^2 analysis that can be performed on deuterated peptides. Traditional CID of peptides is not feasible due to proton motility under collisional activation.[166,167] This leads to *scrambling* of the deuterium throughout the peptide and therefore a complete loss of residue-level resolution.[168] This can be avoided using alternate fragmentation methods, namely ECD and ETD.[169] This is also mitigated slightly by pepsin digestion, which produces a multitude of small peptides. High coverage from small peptides yields pseudo-MS^2 resolution due to the high numbers of overlapping peptides.

1.5.2. Radical Protein Footprinting

Several protein footprinting techniques utilise *in situ* generated radical species to covalently label the accessible protein surface. Hydroxyl radical protein footprinting (H-RPF) and its successor, fast photochemical oxidation of proteins (FPOP) utilises hydroxide radicals to modify protein structures. Trifluoromethyl radical protein footprinting (TF-RPF) is a modification of FPOP, where trifluoromethyl (CF_3) radicals are produced and label the protein structure. These are examples of irreversible protein footprinting workflows, where stable covalent modifications allow for extensive downstream processing, whilst maintaining the bound label.[170]

1.5.2.1. Hydroxyl Radical Protein Footprinting

Similar to initial HDX studies, H-RPF was first developed to study the interactions between nucleic acids and proteins, where hydroxyl radical-induced DNA cleavage was prevented in protein-bound DNA regions, before direct protein labelling was investigated by MS.[171,172] Reaction of hydroxyl radicals imparts easily discernible common mass shifts of +16 and +14 Da (from hydroxyl and carbonyl modification of CH_2 groups respectively). Classically, hydroxyl radicals were formed in one of two ways; by Fenton's reagent or through synchrotron radiolysis of water.[173] In the former, ethylenediaminetetraacetic acid (EDTA)-chelated Fe^{2+} is oxidized by hydrogen peroxide (H_2O_2), to Fe^{3+} leading to generation of a hydroxide anion and a hydroxyl radical.[174] This is a cheap and readily available method for generation of the radical, however it suffers from long exposure times required for sufficient oxidation, leading to poor temporal resolution of protein dynamics.[175] The extended protein exposure to H_2O_2 may also lead to oxidative damage of the protein itself, questioning whether the protein remains in a native conformation. In contrast, X-ray radiolysis of water generates hydroxyl radicals, amongst other radical species, quickly (milliseconds) with no need for H_2O_2, leading to a theoretically higher temporal resolution. However, synchrotrons are not typical laboratory equipment, so instrument availability is a concern, along with exposure to high-energy (GeV), high-flux photons which are known to induce protein aggregation at exposures greater than 50 ms.[176]

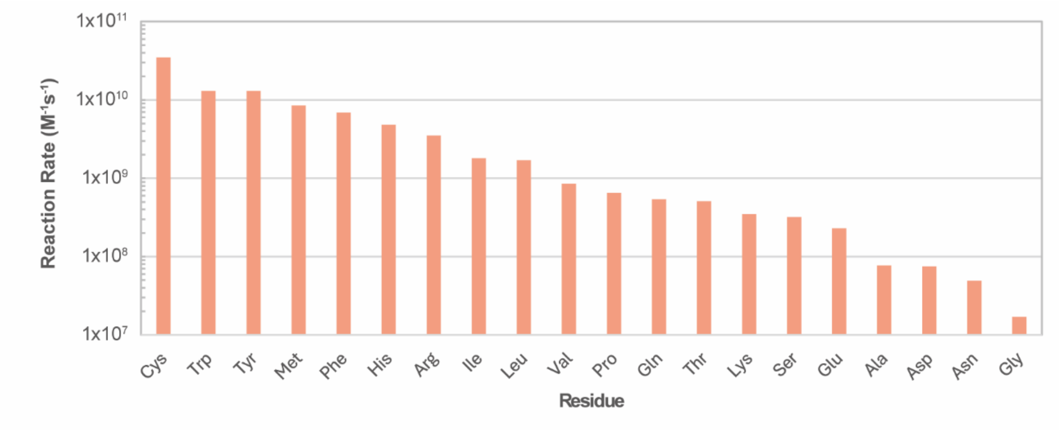


Figure 1.22 | Reaction rate constants for reaction of hydroxyl radicals with each amino acid, showing the approximate 1000-fold difference between glycine and cysteine.

To combat some of these downsides Michael Gross and David Hambly, and independently Aye *et al.* developed the fast photochemical oxidation of proteins workflow (FPOP).[176,177] Here, photolysis of H_2O_2 by a 17 ns-pulsed 248 nm KrF or 266 nm frequency-quadrupled Nd:YAG laser generates hydroxyl radicals. The experimental setup allows for flow-based irradiation, where samples were passed down fused-silica tubing and irradiated through a UV-transparent section. Irradiation and oxidation times are such that protein dynamics can be captured. The inclusion of antioxidants allowed for radical quenching, and pulsed irradiation results in internal controls to measure background oxidation.[175] However, the use of a far-UV laser may have consequences on protein stability, being similar in wavelength to peak absorbance of aromatic residues. Hydroxyl radicals in general suffer from a 1000-fold disparity of residue reaction rates (figure 1.22).[178] High potential sites of reaction across a protein/peptide leads to complex data, but this is remedied somewhat by advancements in automated data analysis.[179]

1.5.2.2. *Trifluoromethyl Radical Protein Footprinting*

The FPOP workflow has recently been applied to generate alternate radical species. One of those being the CF_3 radical, by Zhang *et al*. Here, 300-fold molar excess sodium triflate ($NaSO_2CF_3$) was included alongside H_2O_2 and standard FPOP performed. The hydroxyl radical is rapidly quenched by the triflate anion, producing a CF_3 radical as a result. The radical then reacts with protein, forming trifluoromethylated products with a +68 Da mass shift.[180] The radical was found to react with 18 residues, including glycine, alanine, aspartic acid, glutamic acid, serine, and threonine. All of which are relatively unreactive towards the hydroxyl radical.[178] The increased hydrophobicity of modified peptides allows for easy RP-HPLC separation, alongside labelling of hydrophobic residues within detergent-bound transmembrane helices.[180] More recently, the same group applied CF_3 footprinting using synchrotron generated hydroxyl radicals.[181] This process suffers from similar issues as FPOP, namely far-UV photolysis alongside even greater difficulty in data analysis, due to the often multiplexed labelling of hydroxyl and trifluoromethyl groups in the same peptides.[182]

1.5.3. Carbene Footprinting

1.5.3.1. Carbenes

As discussed above, irreversible protein footprinting relies on the generation of highly reactive radical species. Both hydroxyl and trifluoromethyl radicals possess one unpaired valence electron and are therefore *monovalent* radicals. Carbenes are molecules containing a *divalent* carbon radical that possesses two unpaired electrons. The simplest carbene, methylene, is a divalent carbon bound to two hydrogens (H_2C:). The configuration of the unpaired electrons restricts the classification of a carbene to either *singlet* or *triplet*. (figure 1.23). Singlet carbenes adopt a trigonal planar sp^2 orbital structure and their unpaired electrons have opposite spin (total spin of 0), due to the electrons occupying the free sp^2 hybridised orbital. In triplet carbenes, the unpaired electrons have parallel spin (total spin of 1) and therefore cannot occupy the same orbital. Typically, triplet carbenes adopt a sp^2 geometry, but linear *sp* geometries are possible. The ground state of most carbenes is the sp^2 hybridised triplet state, however nucleophilic groups can stabilise singlet carbenes by donation of an electron pair into the available *p* orbital.[183] Carbene reactivity allows for insertion in X-H bonds (X being C, O, N, S), C=C, C=O, C-C and aromatic rings.[184,185] As a result of their electron configuration, singlet and triplet carbenes differ in their reaction mechanisms. Singlet carbene insertion is a stereospecific concerted reaction, whereas triplet insertion is non-stereospecific. Regardless, carbenes demonstrate a nanosecond-level lifetimes making them a potent protein labelling agent, with reaction rates of $10^7 – 10^9$ M^{-1} s^{-1}.[186]

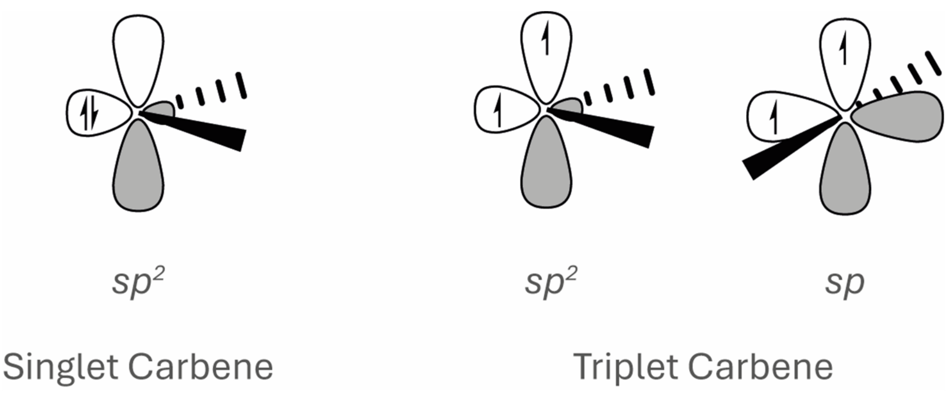


Figure 1.23 | Carbene orbital geometries for both singlet and triplet species.

1.5.3.2. Carbene Protein Footprinting

Carbene protein footprinting employs the use of diazirine-containing precursor probes, which generate carbenes by photolysis. The diazirine group is a highly strained, three-membered ring composed of a carbon and two nitrogen atoms. Irradiation at approximately 350 nm causes expulsion of molecular nitrogen and formation of the carbene (figure 1.24). However, irradiation also induces the rearrangement to the linear diazo group which must undergo a second photolysis to form the carbene, slowing the rate of carbene insertion.[187] Rational design of trifluoromethyl- and aryl-substituted diazirines, which stabilise singlet carbenes due to *P* orbital interactions, principally form the carbene upon irradiation and stabilise the diazo isomer, leading to reduced unwanted diazo-reacted products have led to their increased use as footprinting agents.[188–190] Carbenes show indiscriminate amino acid labelling and the propensity of residue labelling is instead determined by the precursor probe.[191]

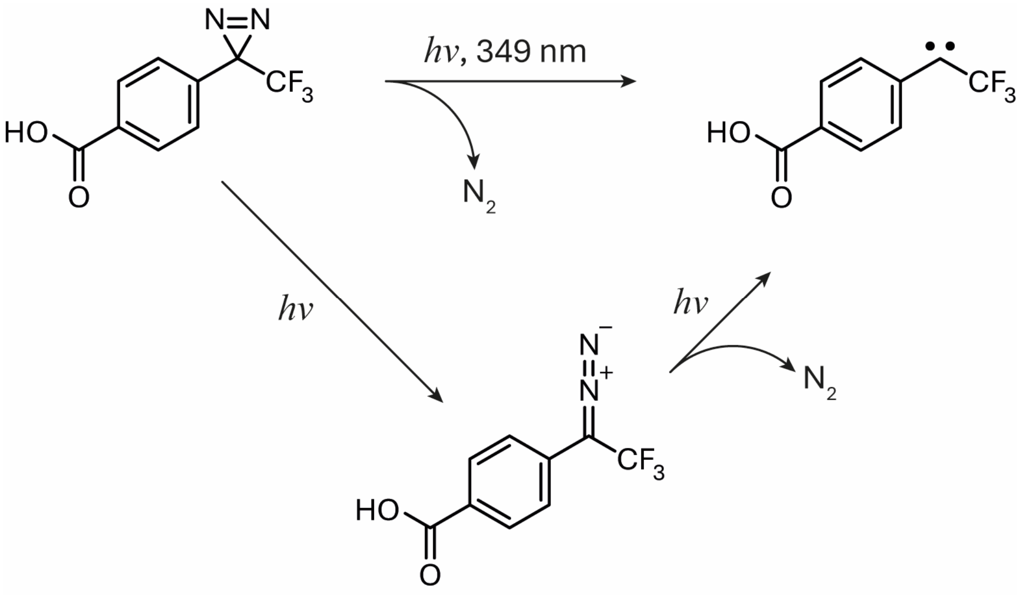


Figure 1.24 | TDBA photolysis under 349 nm laser light. The diazirine group can leave, directly producing the carbene or rearrange to form a diazo group, which can undergo a separate photolysis step.

Initial use of carbenes as footprinting agents by Richards *et al.* utilised diazirine gas to label native hen egg-white lysosome and an intrinsically disordered small peptide, subsequently analysed through ESI-MS.[192] Labelling was performed by saturation of an aqueous sample with diazirine, before photolysis by illumination with a 500 W mercury arc-lamp. This procedure was marred by low solubility, inefficient photolysis (and therefore labelling), and the quote "annoying" flash-burn risk from handling diazirine gas. Chanelle Jumper and David Schriemer improved upon this by using a commercially available diazirine-containing modified leucine, photoleucine, as the carbene precursor. Nd:YAG pulsed laser (frequency-tripled 355 nm, 1000 Hz) irradiation of *apo*- and *holo*-calmodulin with photoleucine and ESI-nMS analysis was successful at differentiating the two, with *holo*-calmodulin accruing fewer labels due to ligand-conferred protein masking.[193] Photoleucine required long irradiation times (>2 minutes) to reach maximum labelling with a 10,000-fold concentration of photoleucine over the protein, and because of such was marred by significant diffusion of the probe. The group improved on this with new probes, 2-amino-4,4-azipentanoic acid and 4,4-azipentanoic acid, which highlighted the differing labelling specificities conferred by the probe. The protein samples were also flash-frozen prior to irradiation, which greatly reduced probe diffusion. This study was the first to include a bottom-up proteomic workflow post labelling, with residue-level resolution obtained by MS^2 analysis of tryptic digests.[194]

More recent advancements showed the development of a benzoic acid-substituted diazirine 4-[3-(trifluoromethyl)-3H-diazirin-3-yl]benzoic acid (TDBA, figure 1.24) by Manzi *et al.* allowed for increased water solubility and greater interactions with both polar and hydrophobic protein residues prior to carbene generation.[195] The end result being much lower concentrations of probe were required for efficient labelling. Differential TDBA labelling was demonstrated on lysozyme and ubiquitin specific peptidase 5 in *apo* and *holo* states, using a Nd:YLF laser (frequency-tripled 349 nm, 1000 Hz). The same group has then gone on to apply TDBA-based footprinting to the eukaryotic tr-GTPase eIF4A, and membrane proteins OmpF and Gasdermin D, where TDBA was able to penetrate the solvating detergent, and label the hydrophobic transmembrane regions.[195–197] Sub-residue carbene insertion was investigated by Ye *et al.* through IMS-MS separation of TDBA labelled peptides, where isobaric peptides containing a single phenylalanine labelled at different positions were separated by ion mobility.[198]

Compared to alternate protein labelling workflows, carbene footprinting offers several advantages. The covalent, irreversible nature of the labelling allows easy downstream processing of samples, including HPLC-MS and MS^2 analysis. The indiscriminate nature of carbene insertion allows labelling of all 20 amino acids, and specificity conferred by the probe can be accounted for by using multiple carbene precursors. Insertion of carbene imparts large mass shifts, +202.02 Da for TDBA, which facilitate easy identification compared to those seen in H-RPF and TF-RPF. Photolytic irradiation at 349 nm removes protein exposure to perturbing far-UV radiation. However, carbene footprinting technologies lack far behind the mature, automated procedures available for other methods, especially HDX. Precursor probe sizes make quantitative comparisons to solvent accessible surface areas difficult, with masking effects seen to be exaggerated by rigid aryldiazirines over more flexible alkyldiazirines.[199] Advancements have been made by development of semi-automated software and trials of FPOP-style continuous flow photoleucine carbene footprinting proved successful, closing the gap between carbene footprinting and more traditional techniques.[200,201]

1.5.4. Overall Aims and Objectives

The work presented in this thesis aims to use multiple structural mass spectrometry techniques, namely carbene footprinting, native mass spectrometry and ion-mobility spectrometry-mass spectrometry, to study the conformational dynamics of EF-Tu. As discussed prior, the global conformation of EF-Tu can be influenced by its bound ligands, as evident by multiple published crystal structures. However, these static structures do not convey how EF-Tu behaves in solution, specifically how binding of these ligands effects its conformational dynamics.

2

Materials and Methods

2.1. EF-Tu Expression and Purification

2.1.1. Transformation and quality assessment of plasmid DNA

Table 2.1 | Buffers used during heterologous EF-Tu protein expression and their respective pH (if given). Constituents denoted with * indicate they were added fresh directly before use.

Buffer	Constituents	pH
Elution	50 mM Tris-HCl, 100 mM NaCl, 400 mM imidazole	7.4
Laemmli (4x)	250mM Tris-HCl, 8% SDS, 40% glycerol, 0.05% bromophenol blue, 6% 2-mercaptoethan	6.8
Lysis	50 mM Tris-HCl, 100 mM NaCl, 20 mM imidazole, 10% glycerol 1 mM DTT *, 1% Igepal CA-60 *, cOmplete protease inhibitor *	7.4
N3	4.2 M guanidinium hydrochloride, 0.9 M potassium acetate	4.8
P1	50 mM Tris-HCl, 10 mM EDTA	8.0
P2	200 mM NaOH, 1% SDS	-
SEC	50 mM Tris-HCl, 150 mM NaCl	8.0
Storage	50 mM Tris-HCl, 150 mM NaCl, 1 mM DTT *	8.0
TBS	200 mM Tris-HCl, 1.5 M NaCl	7.6
TBST	200 mM Tris-HCl, 1.5 M NaCl, 0.05% Tween-20	7.6
TEV	50 mM Tris-HCl, 1 mM DTT *	7.5
Wash	50 mM Tris-HCl, 100 mM NaCl, 20 mM imidazole	7.4

Table 2.1 | *E. coli* expression strains used during heterologous EF-Tu protein expression and their respective genotypic features.

E. coli Strain	Genotypic Markers					
BL21(DE3)	F^-	*fhuA2*	*lon*	*ompT*	*gal*	*dcm*
	No F plasmid	T1 bacteriophage resistance	Lon protease deficient	OmpT protease deficient	Cannot grow in galactose	Inhibition of certain restriction enzymes
	ΔhsdS	*(λ DE3)*				
	Methylation deficient	T7 RNA polymerase expression				
NEB® 5-alpha	F^-	*fhuA2*	Δ(argF-lacZ)205(U169)	*phoA*	*glnV44*	*Φ80Δ(lacZ)M15*
	No F plasmid	T1 bacteriophage resistance	Enhanced resistance to H_2O_2	Deficient in periplasmic alkaline phosphatase	Supression of amber nonsense mutation	Blue/White Screening
	gyrA96	*recA1*	*relA1*	*endA1*	*thi-1*	*hsdR17*
	Nalidixic acid resistance	Reduced plasmid recombination	Inhibition of starvation response	Facilitates plasmid insertion	Thiamine synthesis inhibition	Restriction deficient

2.1.1.1. Transformation

Two purified pET151 plasmids (pET151_tufA1 and pET151_tufA2, figure 2.1) containing wild-type *A. baumannii tufA* gene were kindly provided by Dr Daniel Van (University of Warwick). The *tufA* gene is N-terminally tagged with a 6xHis-tag, followed by a TEV protease cleavage site. Expression is controlled by a standard T7-promoter with a *lac* operon suppression system. DNA was purified from *E. coli* TOP10 (Thermo Fisher Scientific) grown in liquid culture supplemented with 100 µg mL^{-1} carbenicillin. For long-term storage of each plasmid, transformation into chemically competent *E. coli* NEB 5-alpha (New England Biolabs, table 2.2) was performed. Plasmid DNA (1 µL) was mixed with *E. coli* NEB 5-alpha (50 µL) in a 1.5 mL microcentrifuge tube, mixed by gentle inversion and incubated on ice for 20 minutes. The cells were then subject to heat shock (45 °C, 45 seconds) in a water bath before being placed back onto ice. LB broth (900 µL) was added and the transformants recovered *via* incubation at 37 °C for 30 minutes. Post-recovery, 100 µL of each transformant was plated onto LB agar supplemented with 100 µg mL^{-1} ampicillin and incubated at 37°C overnight (18 hours). Single colonies of successful transformants were used to inoculate liquid cultures (LB broth, 10 mL, 100 µg mL^{-1} ampicillin) which were incubated at 37°C overnight with shaking (200 RPM).

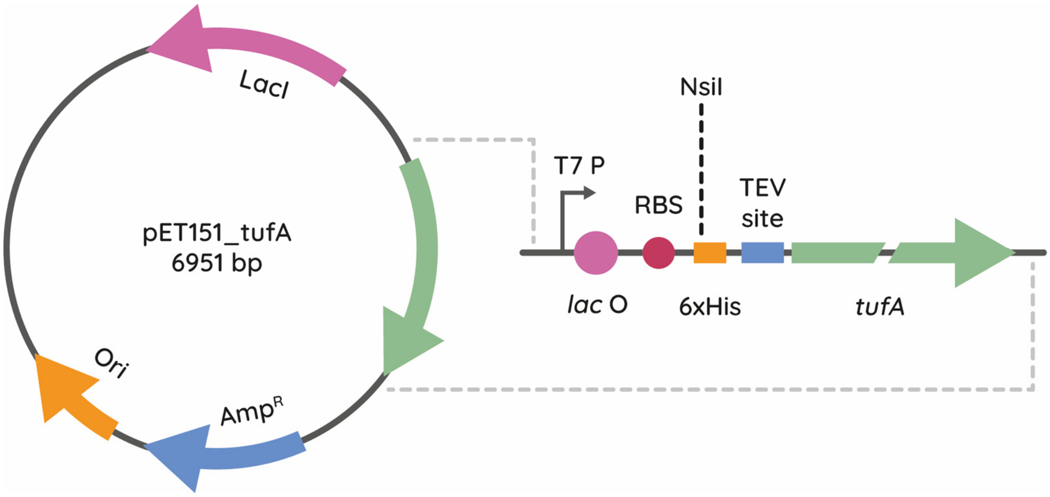


Figure 2.1 | Organisation of the pET151_tufA1/2 plasmids (left) with key features annotated. *lacI*, lac repressor; Ori, origin of replication; Amp^R, ampicillin resistance conferring β-lactamase. The *tufA* gene and associated features are highlighted (right, not to scale). T7 P, T7 bacteriophage promoter; *lac* O, lac operator sequence; RBS, ribosome binding site; NsiI, NsiI restriction site; 6xHis, 6xHis tag; TEV site, TEV protease cleavage site; *tufA*, gene encoding wild-type EF-Tu.

To isolate plasmid DNA, QIAprep Spin Miniprep Kits (Qiagen) were used as per the manufacturers protocol. In brief overnight cultures were centrifuged at 3220x g for 5 minutes and the pellet resuspended in buffer P1 (250 µL, table 2.2). The solutions were transferred to 1.5 mL microcentrifuge tubes, lysis buffer P2 (250 µL) added and incubated on ice for 2 minutes. Neutralising buffer N3 (350 µL) was added followed by a further 2-minute incubation. Solutions were centrifuged (16100x g, 10 minutes), 800 µL of supernatant transferred to a Miniprep column and centrifuged again (16100x g, 1 minute), discarding the flow-through. Samples were then washed three times *via* addition of ethanol (750 µL), incubation for 1 minute and further centrifugation (16100x g, 1 minute), discarding the flow-through each time. The column was then transferred to a clean 1.5 mL microcentrifuge tube, elution buffer EB (50 µL) added, incubated for one minute and centrifuged (16100x g, 1 minute), retaining the flow-through. Purified plasmid was quantified using a NanoDrop Microvolume Spectrophotometer (Thermo Fisher Scientific) blanked with buffer EB, returning DNA concentrations of 68.0 and 140.4 ng µL^{-1} for pET151_tufA1 and pET151_tufA2 respectively.

2.1.1.2. Plasmid DNA Assessment

Plasmid DNA was assessed by a diagnostic restriction digest alongside targeted DNA Sanger sequencing of the tufA genes. Restriction enzyme digestions were performed with NsiI (New England Biolabs) with a targeted cutting site of ATGCA//T. Digestion reactions were performed in 25 µL volumes (3.7 µL plasmid DNA, 1.5 µL restriction enzyme, 2.5 µL of 10x NEBuffer r3.1 buffer (New England Biolabs), 17.3 µL nuclease-free water). The digestion reaction was incubated at 37 °C for 15 minutes, the products subject to agarose gel electrophoresis and visualised using SYBR Safe DNA stain (Thermo Fisher Scientific). Sequencing was performed by the DNA Sequencing Facility (School of Life Sciences, University of Nottingham) using generic T7-promoter forward primers and T7-terminator reverse primers (table 2.3) on an Applied Biosystems® 3130xl Genetic Analyser (Life Scientific).

Table 2.3 | Primers used for introduction of enacyloxin IIa resistant conferring mutations, with the base responsible for the missense mutation underlined within the codon (bracketed). The *tufA* flanking T7 promoter and terminator primer pair were used as sequencing primers. Melting temperatures (T_M) are provided for each primer.

Oligo	Direction	Sequence	T_M (°C)
Q125K	FWD	AAGGTACACCTAC(CTT)ACGAGAAAGAAGGATGTGC	73.4
	REV	GCACATCCTTCTTTCTCGT(AAG)GTAGGTGTACCTT	
G319D	FWD	AATGGAGTGTGACGACC(ATC)TTCTTCTTTAGAAAGTACG	75.3
	REV	CGTACTTTCTAAAGAAGAA(GAT)GGTCGTCACACTCCATT	
A378T	FWD	CTTCACGGAT(CGT)AAAACGTAGACCTGGGTCCATTG	79.5
	REV	CAATGGACCCAGGTCTACGTTTT(ACG)ATCCGTGAAG	
T7 Promoter	FWD	TAATACGACTCACTATAGGG	55.9
T7 Terminator	REV	TATGCTAGTTATTGCTCAGCGG	63.3

2.1.2. Test Expression of Wild-type *A. baumannii* EF-TU

2.1.2.1. Transformation and Over Expression

Purified pET151_tufA1 and pET151_tufA2 was transformed into chemically competent *E. coli* BL21(DE3) (New England Biolabs, table 2.2) for heterologous expression. Plasmid DNA (1 µL) was mixed with *E. coli* BL21(DE3) (50 µL) in a 1.5 mL microcentrifuge tube, mixed by gentle inversion and incubated on ice for 20 minutes. The cells were then subject to heat shock (45 °C, 45 seconds) in a water bath before being placed back onto ice. LB broth (900 µL) was added and the transformants recovered *via* incubation at 37 °C for 30 minutes. Post-recovery, the transformants were centrifuged (16100x g, 1 minute), 850 µL supernatant removed and the pellet resuspended in the remaining supernatant (100 µL). The transformants were then plated onto LB agar supplemented with 100 µg mL^{-1} ampicillin alongside 34 µg mL^{-1} chloramphenicol and incubated at 37 °C overnight.

Single colonies of successful transformants were used to inoculate liquid cultures (LB broth, 10 mL, 100 µg mL^{-1} ampicillin, 34 µg mL^{-1} chloramphenicol) which were incubated at 37°C overnight with shaking (200 RPM). Overnight cultures were used to produce long-term glycerol stocks (500 µL culture, 500 µL sterile 50% glycerol) kept at -80°C and to inoculate test expression cultures (LB broth, 10 mL, 100 µg mL^{-1} ampicillin, 34 µg mL^{-1} chloramphenicol, 200 µL overnight culture). Expression cultures were incubated at 37 °C with shaking (200 RPM) until a turbidity of approximately 0.6 (600 nm) was reached. At this point, expression was induced by the addition of isopropyl ß-D-1-thiogalactopyranoside (IPTG, 0.75 mM) and the cultures incubated at 37 °C with shaking (200 RPM) for a further 3 hours. After which, the cultures were centrifuged (4 °C 3220x g, 15 minutes) and the pellet stored at -20 °C overnight. Cells were lysed by the addition of lysis buffer (5 mL, table 2.2), incubated at 4 °C on a rotary spinner for 15 minutes, and subject to sonication on ice (15 × 10 seconds, 25 Hz tuning, 60% amplitude). Lysed cells were then centrifuged (3220x g, 15 minutes).

2.1.2.2. SDS-PAGE and Immunoblotting

Soluble (30 µL supernatant, 10 µL 4x Laemmli buffer, table 2.2) and insoluble (pellet resuspended in 50 µL 4x Laemmli buffer) fractions were incubated at 95 °C for 5 minutes followed by centrifugation (16100x g, 1 minute). Fractions were resolved by SDS-PAGE *via* a 5-20% gradient gel run at 100 mA current with uncapped voltage. Gels were stained with Coomassie Blue R-250 for 1 hour followed by overnight destaining (50:40:10 methanol, water, acetic acid). For immunoblotting, protein resolved on a separate gel was transferred to an Amersham Potran nitrocellulose membrane (Cytiva) overnight (40 mA at 8-9 V). Following this, the membrane was reversibly stained with Ponceau S (Thermo Fisher Scientific) to confirm successful protein transfer. The stain was removed by several washes with TBST buffer (table 2.2), and the membrane blocked by incubation in 5% Marvel milk powder in TBST for 1 hour. Following this, the membrane was incubated with a mouse anti-6xHis tag primary IgG antibody (4 mL, 1/2000 dilution , AD1.1.10, Bio-Rad) for a further hour. After 3 washes in TBST, the membrane was incubated for 1 hour with a rabbit anti-mouse secondary IgG antibody conjugated to horseradish peroxidase (5 mL 1/5000 dilution, P0260, Agilent). The immunoblot was visualised *via* chemiluminescent exposure of blue/green optimized Hyperfilm photography film (Cytiva) performed in the dark and developed using Ilford multigrade developer.

2.1.3. Over Expression and Purification of Wild-type *A. baumannii* EF-TU

2.1.3.1. Over Expression

An overnight *E. coli* pET151_tufA1 culture (LB broth 20 mL, 100 µg mL^{-1} ampicillin, 34 µg mL^{-1} chloramphenicol) was produced from a glycerol stock of a BL21(DE3) transformant. The following day, full-scale expression cultures (LB broth, 1 L, 100 µg mL^{-1} ampicillin, 34 µg mL^{-1} chloramphenicol) were initiated by the addition of the overnight culture. The culture was grown and induced as in section 2.1.2.1 before centrifugation in two 500 mL flasks (4 °C 4658x g, 15 minutes). The pellet was resuspended in 30 mL TBS buffer (table 2.2), transferred to 50 mL centrifuge tubes and centrifuged (3220x g, 15 minutes) before storage overnight at -20 °C. To lyse the cells, 30 mL of lysis buffer was added to a single pellet prior to incubation and sonication as in section 2.1.2.1. Post sonication, the lysate was transferred to thick-walled centrifuge tubes and centrifuged (4 °C, 23000x g, 15 minutes) and the supernatant transferred to a 50 mL centrifuge tube, alongside a soluble fraction aliquot for SDS-PAGE analysis.

2.1.3.2. Ni-NTA Purification

Nickel-nitrilotriacetic acid (Ni-NTA) agarose matrix solution (1 mL, Qiagen) was prepared by mixing, centrifugation (540x g, 1 minute) and resuspension in lysis buffer (500 µL). This washing step was repeated twice more, and the final resuspension (lysis buffer, 1 mL) kept on ice. The prepped Ni-NTA solution was added to the 30 mL lysate supernatant, and the solution incubated at 4 °C on a rotary spinner for 2 hours. Following this, the lysate was added in batches to a 10 mL Poly-Prep chromatography column (Bio-Rad), an aliquot of flow-through set aside and the remainder discarded. The Ni-NTA agarose matrix was washed 3 times with wash buffer (10 mL, table 2.2) and the flow-through discarded. Protein was eluted by the addition of elution buffer (1 mL, table 2.2), incubation for 5 minutes on a rotary spinner and the flow-through collected. Elution was repeated twice more yielding 3 eluents. SDS-PAGE was performed on the various fractions to confirm protein purity.

2.1.3.3. TEV Cleavage and Size-exclusion Chromatography

A PD10 desalting column (Cytiva) was prepared by the addition of 5 mL TEV protease buffer (table 2.2), discarding the flow-through and repeating a further 4 times. To the equilibrated column, 2.5 mL of pooled eluent was added followed by 2.5 mL of TEV buffer, the flow-through collected and kept on ice. A TEV protease aliquot (New England Biolabs) was added to the flow-through and incubated at 4 °C on a rotary spinner overnight. The following day, sodium chloride (100 mM) and imidazole (20 mM) were added, the solution transferred to a Poly-Prep column containing 1 mL of pre-prepared Ni-NTA agarose matrix (as in section 2.1.3.2) and incubated at 4 °C on a rotary spinner for 1 hour. The flow-through was collected and the column washed twice with SEC buffer (1 mL, table 2.2), collecting the flow-through both times. To assess the post-TEV cleavage purity of wild-type EF-Tu, the flow-through aliquots were pooled and subject to size-exclusion chromatography (SEC) using an ÄKTAprime Plus (Cytiva) equipped with an HiLoad Superdex 200 pg column, equilibrated with SEC buffer and operating at a flowrate of 2.6 mL per minute. In total, 34x 10 mL fractions were collected, SDS-PAGE was performed to assess efficacy of TEV digestion, as well as to confirm protein content of SEC fractions 18 – 23. Following this, fraction 21 (10 mL) was added to a Vivaspin 20 centrifugal concentration (VIVAproducts) and centrifuged (3220x g, 10 minutes). The protein was then washed three times with storage buffer (10 mL) and the flow-through discarded each time. Finally, the protein was concentrated *via* further centrifugation (3220x g, 20 minutes) to a volume of approximately 500 µL.

2.1.3.4. *Quantification of EF-Tu Concentration*

To quantify the final concentration of wild-type *A. baumannii* EF-Tu the bicinchoninic acid (BCA) assay was used. A range of bovine serum albumin (BSA) concentrations were produced in storage buffer as per the manufacturer's guidelines (see below). Each concentration was added to a 96-well microplate (25 µL), in triplicate, alongside a storage buffer blank, a four-fold and ten-fold dilution of concentrated wild-type *A. baumannii* EF-Tu. To each well, 200 µL of prepared BCA reagent (Thermo Fisher Scientific) was added and the plate incubated at 37 °C with shaking for 30 minutes. Following incubation, the absorbance of each well at 562 nm was measured using a SPECTROstar Nano microplate spectrometer (BMG LabTech), resulting in protein concentration of 2.095 mg mL^{-1}. The concentrated wild-type *A. baumannii* EF-Tu was then diluted to a concentration of 50 µM, aliquoted into individual 20 µL volumes, flash-frozen in liquid nitrogen and stored at -80 °C.

2.1.4. Production and Expression of Enacyloxin IIa Resistant *A. baumannii* EF-Tu Mutants

2.1.4.1. *Site-Directed Mutagenesis*

Three individual enacyloxin IIa (ENX) resistant *A. baumannii* EF-Tu mutants (EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T}) were produced, each conferring a different level of resistance. Each resistance conferring mutation was a single amino acid substitution, generated *via* introduction of a single base missense mutation using site-directed mutagenesis. To introduce the missense mutations into pET151_tufA1, primers were designed containing the base substitutions required (table 2.3), and a linear amplification, non-PCR-based QuikChange Lightning site-directed mutagenesis

(Agilent) was used, as per the manufacturer's instructions. In brief, reactions (5 µL reaction buffer, 0.5 µL purified pET151_tufA1, 0.8 µL forward/reverse mutagenic primers, 1 µL dNTP mix, 1.5 µL QuikSolution, 1 µL QuikChange Lightning enzyme, 40.4 µL nucleotide-free water) were performed in a thermocycler, with the cycling parameters given in table 2.4. Successful mutagenesis was confirmed with Sanger DNA sequencing as in section 2.1.1.2, again using standard T7 promoter and terminator primers.

Table 2.4 | Thermocycler cycling parameters for QuikChange Lightning site-directed mutagenesis kit. The length of 68 °C extension in segment 2 is dependent on plasmid length (30 seconds for each Kbp of DNA, pET151_tufA1 ≈ 7 Kbp).

Segment	Cycles	Temperature (°C)	Time
1	1	95	2 minutes
2	18	95	20 seconds
		60	10 seconds
		68	3 minutes, 30 seconds
3	1	68	5 minutes

2.1.4.2. Expression and Purification

Mutant *A. baumannii* EF-Tu encoding plasmids (pET151_tufA1_K, pET151_tufA1_D, pET151_tufA1_T) were transformed into chemically competent *E. coli* BL21(DE3) as in section 2.1.1.1 and single colonies of successful transformants used to inoculate an overnight culture (LB broth 20 mL, 100 µg mL^{-1} ampicillin, 34 µg mL^{-1} chloramphenicol). Protein expression was performed as described in section 2.1.3.1, followed by Ni-NTA purification (section 2.1.3.2) and TEV cleavage (section 2.1.3.3). Cleaved *A. baumannii* EF-Tu mutants were not purified with SEC, as the additional purity gained from SEC was not required over that reached by Ni-NTA purification alone. Each protein concentration was quantified using a BCA assay (section 2.1.3.4) resulting in concentrations of 9.395, 3.935 and 14.125 mg mL^{-1} for EF-Tu^{Q125K}, EF-Tu^{G319D} and EF-Tu^{A378T} respectively. Each protein was then concentrated, aliquoted and stored as described in section 2.1.3.4. Further confirmation of successful mutagenesis was attained by collision-induced dissociation tandem mass spectrometry (CID-MS/MS, (sections 2.3.1.2 to 2.3.1.5) and, for EF-Tu^{Q125K}, a diagnostic tryptic digest performed (sections 2.3.1.2 to 2.3.1.5) due to the introduction of a novel tryptic cleavage site.

2.2. Native Mass Spectrometry

All native mass spectrometry and ion-mobility spectrometry-mass spectrometry (IMS-MS) measurements were performed on a Synapt G1 HDMS High-Definition mass spectrometer (Waters). The mass spectrometer was operated in positive ion mode. All aqueous solutions were prepared with Milli-Q ultrapure water (>18 MΩ resistivity) and all compounds and reagents were purchased from Merck (Sigma-Aldrich) or Thermo Fisher Scientific, unless otherwise stated.

2.2.1. Native Mass Spectrometry of EF-Tu Complexes

2.2.1.1. Preparation of Nano-Electrospray Ionisation Emitter Tips

Emitter tips used for nano-electrospray ionisation (nESI) experiments were pulled, in-house, from borosilicate 3.3 glass capillary tubes (0.8 mm internal diameter, 1.0 mm external diameter, 1000 mm length, Hirschmann) using a Flaming/Brown P-97 micropipette puller (Sutter Instruments). This resulted in tips with an approximately 1 cm taper leading to a 0.3 – 0.6 µm internal diameter orifice. The prepped emitter tips were then loaded into the housing assembly, and 5 µL samples loaded through

GELoader micropipette tips (Eppendorf). The sample was forced to the orifice of the emitter tip through flicking the assembly, prior to the insertion of a back-fitted, 99.9% platinum wire filament (0.127 mm diameter, Sigma-Aldrich), such that the bottom of the filament was submerged within the sample (figure 2.2) The whole emitter housing was then assembled into the three-axis interlocked stage of the nESI source.

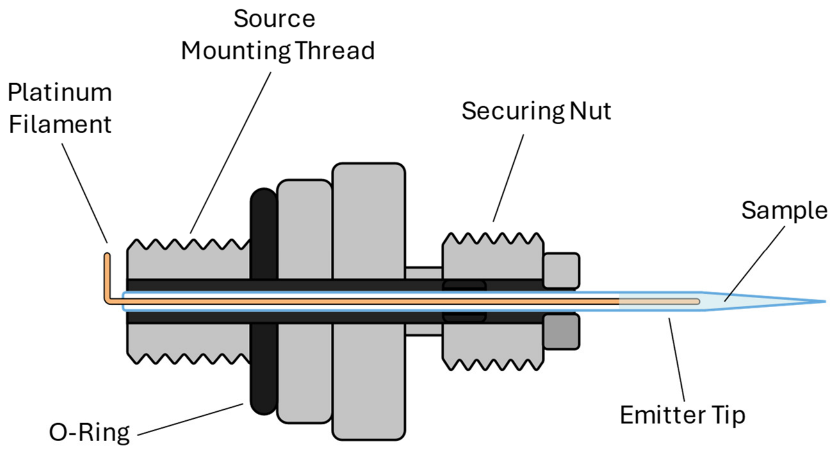


Figure 2.2 | Cut-away schematic representation of the assembled nESI emitter tip housing assembly.

2.2.1.2. Sample Preparation for Native Mass Spectrometry

Protein samples require exchange into a buffer compatible with native mass spectrometry prior to submission to the spectrometer. The high salt presence in common biological buffers presents a challenge to native mass spectrometry, leading to signal suppression and peak broadening. All protein was buffer exchanged into ammonium acetate (50 mM, pH 6.9) using Zeba Spin desalting columns (Thermo Fisher Scientific) with a 7 KDa MWCO and either a 75 or 500 µL bed volume). Buffer exchange was performed following the manufacturer's instructions. In brief, columns were centrifuged (4 °C, 1250x g, 1 minute) to remove storage buffer and equilibrated with 5x washes in ammonium acetate (50 or 400 µL depending on bed volume) by addition and centrifugation (4 °C, 1250x g, 1 minute). The column was placed into a clean 1.5 mL microcentrifuge tube, the protein samples loaded (2-12 or 30-130 µL depending on bed volume) and the column centrifuged again (4°C, 1500x g, 2 minutes), collecting the eluent containing the buffer exchanged protein.

2.2.1.3. Native Mass Spectra Acquisition

Table 2.5 | Select representative Synapt operational parameters. Some parameters, specifically voltages, were often changed slightly per measurement to optimise signal.

Condition	Native (TOF)	IMS-MS (Mobility)
Capillary Voltage (kV)	1.4 - 1.8	1.4 - 2.2
Sampling Cone Voltage (V)	80	80
Extraction Cone Voltage (V)	5	50
Collision Energy (V)	10	10 - 110
Trap Cell Bias (V)	5	12
Trap T-Wave Velocity (ms^{-1})	-	300
IMS T-Wave Velocity (ms^{-1})	-	280
Transfer T-Wave Velocity (ms^{-1})	-	200
Backing Pressure (mbar)	4.06	4.07
Source Pressure (mbar)	2.66×10^{-3}	2.68×10^{-3}
Trap Pressure (mbar)	1.84×10^{-2}	2.98×10^{-2}
IMS Cell Pressure (mbar)	5.17×10^{-4}	4.31×10^{-1}
TOF Pressure (mbar)	1.60×10^{-6}	1.96×10^{-6}

After buffer exchange, protein samples were diluted further in ammonium acetate to working concentrations of 1 – 15 µM prior to loading into the emitter tip assembly (section 2.2.1.1). Dry gaseous nitrogen was used for desolvation and nanoflow, with a nanospray backing pressure of 100 mbar. A gentle capillary voltage was applied, typically between 1.0 – 2.5 kV to ionise the sample. Capillary voltage and the cartesian position of the emitter stage were altered with each measurement in order to give the greatest signal strength. Cone voltages and collision voltages were kept constant between measurements. Representative voltages, pressures *etc* can be seen in table 2.5. Dry gaseous argon was used as the collision gas. The Synapt was operating in TOF acquisition mode and spectra were produced from 1 – 2 minutes of acquisition to give the greatest signal to noise ratio. Mass spectra were analysed in MassLynx 4.0 (Waters).

2.2.1.4. Native Mass Spectra of, and Collision-induced Dissociation of Enacyloxin IIa bound EF-Tu

Initial studies of EF-Tu and enacyloxin IIa were performed with *E. coli* wild-type EF-Tu, which was kindly provided by Professor Greg Challis (University of Warwick). All enacyloxin IIa samples were also provided by the Challis group. Enacyloxin IIa was extracted from *B. ambifaria* by bacterial growth on solid media, followed by ethyl acetate extraction of diced agar pieces. The crude enacyloxin IIa was then purified by HPLC. Work conducted with enacyloxin IIa was performed in the dark, and concentrated pure stocks stored in foil at -20 °C. Enacyloxin IIa working stocks of 20 µM were produced as needed in ammonium acetate.

E. coli apo-EF-Tu aliquots (1 µL) were thawed on ice before dilution in ammonium acetate (4 µL) to approximately 15 µM. To produce ligand bound EF-Tu, the *apo*-protein was diluted in magnesium acetate tetrahydrate (5 mM, 2 µL) and aqueous GDPNP (5 mM, 2 µL). Aliquots were incubated on ice for 10 minutes prior to buffer exchange (section 2.2.1.2). Buffer exchanged *apo*-EF-Tu was further diluted two-fold in ammonium acetate, whilst EF-Tu•GDPNP was diluted two-fold in ammonium acetate working stocks of enacyloxin IIa, resulting in a final enacyloxin IIa concentration of 5 µM. Proteins for use in enacyloxin IIa bound experiments were purposefully incubated with sub-stoichiometric concentrations of enacyloxin IIa, resulting in individual samples containing *apo*-EF-Tu, EF-Tu•GDPNP, and EF-Tu•GDPNP•ENX complexes. Protein samples were incubated on ice for 2 minutes before loading into prepared emitter tips (section 2.2.1.1) and spectra acquired (section 2.2.1.3). For collisional activation of EF-Tu•GDPNP•ENX, the dominant charge state $[M + 12H]^{12+}$ was isolated in the quadrupole, and collision voltage increased from 20 to 60 V to induce dissociation of enacyloxin IIa from the complex.

2.2.1.5. Native Mass Spectra of, and Collision-induced Dissociation of GE2270A bound EF-Tu

Native mass spectrometry measurements on EF-Tu bound to GE2207A were performed on wild-type *A. baumannii* EF-Tu. GE2270A (AdipoGen Life Sciences) was dissolved in DMSO to give a stock solution of 10 mM and stored at -20 °C. Working stocks of 100 µM produced as needed in ammonium acetate. *A. baumannii* wild-type EF-Tu aliquots were thawed on ice before dilution in ammonium acetate, giving 110 µM protein. This was then diluted further by the addition of GDP (5 mM, 2 µL), magnesium acetate tetrahydrate (5 mM, 1 µL) and GE2270A (100 µM, 4 µL) yielding 36.6 µM EF-Tu and 40 µM GE2270A. Aliquots were incubated on ice for 10 minutes prior to buffer exchange (section 2.2.1.2). Buffer exchanged EF-Tu•GDP•GE2270A was then diluted two-fold in ammonium acetate. Protein samples were incubated on ice for 2 minutes before loading into prepared emitter tips (section 2.2.1.1) and spectra acquired (section 2.2.1.3). For collisional activation of EF-Tu•GDP•GE2270A, the dominant charge state $[M + 12H]^{12+}$ was isolated in the quadrupole and collision voltage increased to from 10 - 80 V to induce dissociation of GE2270A from the complex.

2.2.1.6. Investigating binding of Enacyloxin IIa Analogues to EF-Tu

Structural analogues of wild-type enacyloxin IIa were produced by either mutagenesis of the biosynthetic gene cluster responsible for enacyloxin IIa synthesis, or by introduction of alternative precursor DHCCA analogues. All enacyloxin IIa analogues were provided by Professor Greg Challis and produced by members of his research group. Analogues were provided as crude (without HPLC purification) solid residue produced by evaporation of ethyl acetate extracts (section 2.2.1.4). Unless otherwise specified, crude analogue extracts were dissolved in methanol (100 µL), in the dark, and stored at -20°C. Working stocks were produced as needed by ten-fold dilution in ammonium acetate. *A. baumannii* wild-type EF-Tu aliquots (50 µM, 10 µL) were thawed on ice, prior to a 2 µL sample being diluted in magnesium acetate tetrahydrate (5 mM, 4 µL), crude extract working stock (1 µL) and aqueous GDPNP (5 mM, 2 µL). Aliquots were incubated on ice for 10 minutes prior to buffer exchange (section 2.2.1.2), loaded into prepared emitter tips (section 2.2.1.1) and spectra acquired (section 2.2.1.3) with no further dilutions. If bound EF-Tu was present, the $[M + 12H]^{12+}$ charge state was isolated in the quadrupole and collisionally activated (40 V), liberating the bound analogue for detection.

2.2.1.7. Native mass spectra of EF-Tu bound Prolyl Hydroxylase

P. putida EF-Tu and prolyl hydroxylase (PPHD, Putative Prolyl Hydroxylase) were provided by Professor Chris Schofield (University of Oxford). Protein was thawed on ice and diluted to working concentrations (20 µM, 10 µL) in ammonium acetate and stored on ice prior to buffer exchange (section 2.2.1.2). Buffer exchanged EF-Tu•GDP and *apo*-PPHD were diluted two-fold in ammonium acetate, whereas the *holo* complex was formed by mixing EF-Tu•GDP and *apo*-PPHD at an equimolar concentration, resulting in 10 µM concentrations of *apo* and *holo*-protein. Samples were incubated on ice for 15 minutes before loading into prepared emitter tips (section 2.2.1.1) and native spectra acquired (section 2.2.1.3).

2.2.1.8. K_D Characteristics of Mutant EF-Tu and Enacyloxin IIa

A. Baumannii enacyloxin IIa resistant mutant EF-Tu aliquots (50 µM, 20 µL) were thawed on ice prior to dilution with ammonium acetate to 10 µM and incubated on ice for 10 minutes prior to buffer exchange using 500 µL bed volume Zeba columns (section 2.2.1.2). Enacyloxin IIa stocks were produced covering a wide range of concentrations (320, 240, 200, 120, 80, 40, 20, 10, 5 and 2.5 µM) by dilution in ammonium acetate and stored in the dark, on ice until needed. Buffer exchanged protein was diluted two-fold with an enacyloxin IIa sample and incubated on ice for 2 minutes before loading into prepared emitter tips (section 2.2.1.1) and spectra acquired for 2 minutes (section 2.2.1.3). In these experiments, an open vial of acetonitrile was placed in the Synapt nESI source to induce a charge state reduction, which can stabilise weak protein-ligand interactions.[202] The proportion of EF-Tu•GDP•ENX respective to total protein signal was calculated, and a non-linear regression model fit, allowing calculation of K_D by way of the enacyloxin IIa concentration required to reach 50 % of the maximum equilibrium binding.

2.2.2. Ion-mobility Spectrometry-mass Spectrometry

2.2.2.1. IMS-MS Calibration

Travelling-wave IMS-MS was employed to study *E. coli* EF-Tu conformational stability in the gas phase as an *apo*-protein, alongside EF-Tu•GDPNP, and EF-Tu•GDPNP•ENX complexes. IMS-MS experiments were performed using the standard nESI source, with the Synapt G1 HDMS instrument operating in the mobility acquisition mode. Representative operating parameters can be seen in (table 2.5). The TWIMS cell of the instrument was calibrated for collisional cross section (CCS) measurements by acquiring arrival time distributions of denatured myoglobin, native cytochrome C and native ubiquitin and comparing to those reported in literature databases.[203,204]

Arrival time distributions were converted to calibrated CCS values (${}^{TW}CCS_{N2\rightarrow He}$) using the method developed by Ruotolo *et al.* with the required enhanced duty cycle coefficient (EDC) being 1.41.[145] In brief, calibrant drift times are first corrected for mass-dependent total flight time by equation 2.1 where t'_D is the corrected drift time (ms), t_D the measured drift time (ms), C the EDC constant and m/z being the mass to charge ratio of the ion.

$$t'_D = t_D - \left(\frac{C\sqrt{m/z}}{1000}\right) \tag{2.1}$$

Calibrant CCS values were also corrected for charge state and reduced mass by equation 2.2 where Ω' is corrected CCS (nm^2), Ω is literature CCS (nm^2), z is the ion charge state and μ is the reduced mass.

$$\Omega' = \frac{\Omega}{\left(z \times \sqrt{1/\mu}\right)} \tag{2.2}$$

By plotting the natural logarithms of t'_D against Ω' and performing a linear regression (equation 2.3) the factors X and A, referred to as exponent factor and the fit-constant respectively, can be derived from the gradient and y-axis intercept respectively.

$$\ln(\Omega') = X\ln(t'_D) + \ln(A) \tag{2.3}$$

The exponent factor can then be used to further correct the drift time (t''_D) using equation 2.4. This can then be re-plotted against Ω, which allows for the conversion of experimentally derived drift times to a corrected CCS ($^{TW}CCS_{N2\rightarrow He}$).

$$t''_D = t'_D{}^X \times z \times \sqrt{1/\mu} \qquad (2.4)$$

2.2.2.2. IMS-MS of EF-Tu in Complex with Enacyloxin IIa

E. coli EF-Tu samples were prepared as described in section 2.2.1.4 with sub-stoichiometric enacyloxin IIa, resulting in samples containing *apo*-EF-Tu, EF-Tu•GDPNP, and EF-Tu•GDPNP•ENX. Protein samples were incubated on ice for 2 minutes before loading into prepared emitter tips (section 2.2.1.1) and spectra acquired (section 2.2.1.3) in mobility acquisition mode, with 2-minute acquisition times per sample. The $[M + 12H]^{12+}$ peaks corresponding to the *apo*-EF-Tu, EF-Tu•GDPNP, and EF-Tu•GDPNP•ENX species were isolated in the quadrupole and subject to incrementally increased collision voltages (20 – 60 V in 1 V steps, alongside 42 – 44 V in 0.5 V steps). Arrival time distributions for each species were extracted at each collision energy and corrected based on the calibration (section 2.2.2.1). The corrected arrival time distributions were then used to calculate $^{TW}CCS_{N2\rightarrow He}$ values. Arrival time distributions fingerprint and waterfall plots were produced using custom Python scripts, generated by Dr Jeddidiah Bellamy-Carter (University of Loughborough).

CIU_{50} (collision energy at which 50 % of the protein species has undergone an unfolding event) values were calculated using CIUSuite 2.[205] Fitting of a four-parameter sigmoidal logistic regression to EF-Tu $^{TW}CCS_{N2\rightarrow He}$ as a function of collision energy allowed for continuous extraction of CIU_x (collision energy at which x % of the protein species has undergone an unfolding event). Analyses were performed in

Prism 10.1.2 (GraphPad), T-tests and analysis of variance (ANOVA) statistical tests were employed with significance denoted as $P < 0.05$.

2.3. Carbene Footprinting

All carbene footprinting experiments analysed using an LTQ FT Ultra Hybrid Mass Spectrometer (Thermo Fisher Scientific) operated in positive ion mode, coupled to a Dionex U3000 nano-HPLC system. All aqueous solutions were prepared with Milli-Q ultrapure water (>18 MΩ resistivity) and liquid chromatography mobile-phase prepared with LC-MS grade acetonitrile. All compounds and reagents were purchased from Merck (Sigma-Aldrich) or Thermo Fisher Scientific, unless otherwise stated. All proteases were purchased from New England Biolabs.

2.3.1. Sample Preparation

2.3.1.1. TDBA Preparation

The photochemical probe TDBA was solubilised as the sodium salt by the addition of a sub-stoichiometric volume of sodium hydroxide solution (100 mM, 977 µL, 0.9 molar equivalent) to solid TDBA (25 mg). The solution was stirred at room temperature for 1 hour, prior to filtration through cotton wool to remove excess undissolved TDBA, yielding a 100 µM solution of the sodium salt. Solubilised TDBA was stored in foil at 4 °C until use.

2.3.1.2. SDS-PAGE of Protein for Proteolytic Digestion

Protein samples were mixed with 6x Laemmli reducing buffer (5:1 µL protein:buffer) and incubated at 95 °C for 5 minutes prior to loading into either in-house made 12% gel or into 12% TGX SDS-PAGE gel (BioRad). Gels were run at 150 V with an uncapped current for 45 minutes to 1 hour to ensure band separation. Gels were washed with ultrapure water, heated for 2 minutes by microwave, and the water discarded. Gels were washed twice more prior to the addition of GelCode Blue coomassie G-250 stain (20 mL) and further heating by microwave for 90 seconds. The gel was stained for a further 15 minutes at room temperature with rocking. Following this, the stain was discarded, and the gel rapidly destained by adding ultrapure water and microwave heating for 2 minutes.

2.3.1.3. In-Gel Proteolytic Digestion

Protein bands of interest was excised from the gel using a scalpel, diced into approximately 1 mm^3 pieces and destained for 15 minutes in 500 µL microcentrifuge tubes by the addition of 50:50 acetonitrile:ammonium bicarbonate (100 mM, 50 µL). Following this, gel pieces were dehydrated with acetonitrile (450 µL) and vortex mixing for 3 minutes. The acetonitrile was removed, the gel pieces reduced with DTT (10 mM, 50 µL, ammonium bicarbonate) and incubated at 55 °C for 30 minutes. After this, the gel pieces were dehydrated again, alkylated by the addition of iodoacetamide (55 mM, 50 µL) and incubated for 30 minutes at room temperature, in the dark. The gel pieces were dehydrated a final time, protease solution added (10 ng μL^{-1}, 50 µL, 50 mM ammonium bicarbonate) and incubated overnight at 37°C (AspN, GluC and trypsin) or 25°C (chymotrypsin). The digestion reaction was quenched the following day by the addition of formic acid (1 µL), the supernatant removed and centrifuged (5000x g, 2 minutes) and the top half transferred to plastic autosampler vials for submission to nano-HPLC-MS analysis.

2.3.1.4. In-Solution Proteolytic Digestion

Protein samples (5 µL) were reduced by the addition of ammonium bicarbonate (100 mM, 8.5 µL) and DTT (100mM, 1.5 µL, ammonium bicarbonate) and incubation at 55 °C for 30 minutes. Following this, the samples were alkylated with iodoacetamide (120 mM, 3 µL, ammonium bicarbonate) and incubated for 30 minutes at room temperature, in the dark. Excess iodoacetamide was then reduced with DTT (100 mM, 2 µL) and incubated for 30 minutes at room temperature. Finally, protease solution was added (50 ng µL^{-1}, 5 µL, 50 mM ammonium bicarbonate) and the samples incubated overnight, quenched and centrifuged (section 2.3.1.3).

2.3.1.5. Nano-HPLC-MS analysis

Protein digests were analysed using a Dionex U3000 nano-HPLC coupled with an LTQ FT Ultra Hybrid Mass Spectrometer. Digests (typically 2 – 5 µL) were loaded onto a C18 Pepmap300 trapping column (10 mm length, 300 Å pore size, 5 µm particle size, Thermo Fisher Scientific) with mobile phase A (5 % acetonitrile, 0.1 % formic acid). Peptides were eluted onto a C18 Pepmap300 resolving column (150 mm length, 75 µm diameter, 300 Å pore size, 5 µm particle size, Thermo Fisher Scientific) and separated with a 30-minute linear gradient up to 55 % mobile phase B (95 % acetonitrile, 0.1 % formic acid), followed by a rapid 5-minute hold at 90% mobile phase B and finally a 20-minute equilibration back to mobile phase A. The resolving column was coupled directly to a CoAnn pulled fused silica emitter tip (20 µm internal diameter, 365 µm external diameter, 10 µm orifice internal diameter, MSWil) operated at 1.8 – 1.9 kV with the ion transfer tube held at 275 °C and the tube lens operating at 145 V. For standard peptide analysis, the spectrometer was operated in full scan mode utilising the ultra-high mass accuracy and resolving power of the FT-ICR mass analyser. For peptide identification, the spectrometer was operated in data-dependent analysis

(DDA) mode, whereby the three most intense ions in the full scan are isolated in a 2 *m/z* window and subject to collision-induced dissociation fragmentation and detection in the LTQ XL linear ion trap with 35% nominal energy. Parent ions observed for 3 consecutive scans were rejected within a 6-minute window of the last detection. Targeted MS/MS was performed on selected precursor ions within an 8 *m/z* window of a specific *m/z* also in the LTQ XL linear ion trap with the same operating parameters.

2.3.2. Carbene Footprinting of EF-Tu

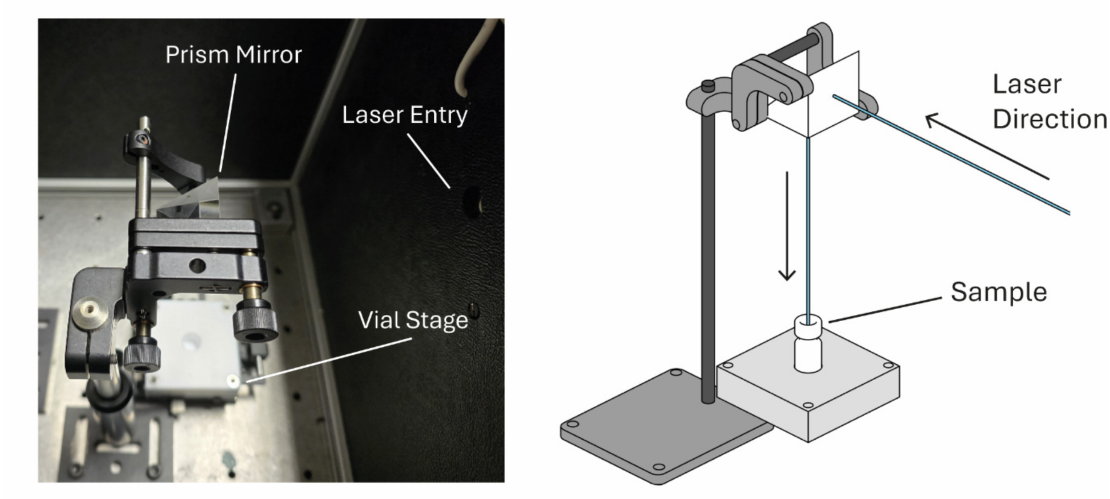


Figure 2.3 | Carbene footprinting irradiation assembly. Left is a photograph of the right-angle prism mirror and two-axis stage mount for sample vials, the assembly is housed within a laser safety enclosure. A schematic representation is shown on the right.

2.3.2.1. *EF-Tu in Complex with Enacyloxin IIa*

Wild-type *A. baumannii* EF-Tu aliquots (50 µM, 20 µL) were thawed on ice prior to dilution. *Apo*-EF-Tu master mixes were produced composed of EF-Tu (50 µM, 2.5 µL), TDBA (100 mM, 2.5 µL) and ultrapure water (20 µL). Nucleotide bound EF-Tu master mixes contained EF-Tu (50 µM, 2.5 µL), TDBA (100 mM, 2.5 µL), GDP/GDPNP (5 mM, 5 µL) and ultrapure water (15 µL). EF-Tu•GDPNP•ENX master mixes contained EF-Tu (50 µM, 2.5 µL), TDBA (100 mM, 2.5 µL), GDPNP (5 mM, 5 µL), enacyloxin IIa (30 µM, 5 µL) and ultrapure water (10 µL). Each master mix was incubated on ice for 30 minutes before being separated into 5x 5 µL aliquots in plastic autosampler vials, the samples flash-frozen in liquid nitrogen and transported on ice to the laser assembly. Protein samples were irradiated for 15 seconds by a computer-controlled Explorer One 349 laser (Nd:YLF at the tripled frequency harmonic of 349 nm, 1000 Hz repetition frequency, 125 µJ Q-switched pulsed energy, Spectra Physics). The laser beam enters perpendicular to the sample and is reflected down into the vial by a right-angle prism mirror (figure 2.3). After irradiation, samples were allowed to thaw on ice before separation *via* SDS-PAGE (section 2.3.1.2), the labelled protein excised and digested in-gel with Trypsin (section 2.3.1.3) and the resulting peptides analysed (section 2.3.1.5). Optimisation enzymatic digests, and footprinting with different TDBA concentrations and irradiation times were also performed as described above.

2.3.2.2. EF-Tu Mutants in Complex with Enacyloxin IIa

Mutant and wild-type *A. baumannii* EF-Tu (EF-Tu^{Q125K}, EF-Tu^{G319D} and EF-Tu$^{A378T)}$ aliquots (50 µM, 20 µL) were thawed on ice prior to dilution. Master mixes were produced for each EF-Tu species consisting of EF-Tu•GDPNP (EF-Tu, 50 µM, 4 µL; TDBA, 100 mM, 2 µL; GDPNP, 4 mM, 4 µL; ultrapure water 10 µL), EF-Tu•GDPNP•ENX (EF-Tu, 50 µM, 4 µL; TDBA, 100 mM, 2 µL; GDPNP, 4 mM, 4 µL; enacyloxin IIa, 55 µM, 4 µL; ultrapure water 6 µL). A master mix was also produced to investigate mutant-EF-Tu•GDP•ENX binding to higher enacyloxin IIa concentrations (EF-Tu, 50 µM, 3 µL; TDBA, 100 mM, 1.5 µL; GDP, 4 mM, 3 µL; enacyloxin IIa, 200 µM, 3 µL; ultrapure water 4.5 µL). Samples were aliquoted, frozen and irradiated as in section 2.3.2.2 and then subject to in-gel enzymatic digestion with Trypsin (section 2.3.1.3) prior to analysis (section 2.3.1.5).

2.3.2.3. EF-Tu in Complex with GE2270A

Wild-type *A. baumannii* EF-Tu aliquots (50 µM, 20 µL) were thawed on ice prior to dilution. Master mixes were produced consisting of EF-Tu•GDP (EF-Tu, 50 µM, 5 µL; TDBA, 100 mM, 2.5 µL; GDP/GDPNP, 5 mM, 5 µL; ultrapure water 12.5 µL) and EF-Tu•GDP/GDPNP•GE2270A (EF-Tu, 50 µM, 5 µL; TDBA, 100 mM, 2.5 µL; GDP/GDPNP, 5 mM, 5 µL; GE2270A, 60 µM, 5 µL; ultrapure water 7.5 µL). Samples were aliquoted, frozen and irradiated as in section 2.3.2.2 and then subject to in-solution enzymatic digestion with Trypsin (section 2.3.1.4) prior to analysis (section 2.3.1.5).

2.3.2.4. EF-Tu in Complex with PPHD

P. putida EF-Tu and PPHD aliquots (100 µM, 2 µL) were thawed on ice prior to dilution. Separate footprinting experiments were performed focusing on either EF-Tu or PPHD with respect to the other protein as a binding partner *e.g.* footprinting of EF-Tu with respect to PPHD binding was performed with 8 µM EF-Tu and 10 µM PPHD to cause saturation of EF-Tu and *vice versa* for PPHD with respect to EF-Tu binding. Master mixes were produced consisting of *apo*-EF-Tu/PPHD (EF-Tu/PPHD, 50 µM, 9 µL; TDBA, 100 mM, 4.5 µL; ultrapure water 31.5 µL), *holo*-EF-Tu•PPHD (EF-Tu, 50 µM, 7.2 µL; PPHD, 50 µM, 9 µL; TDBA, 100 mM, 4.5 µL; ultrapure water 24.3 µL) and *holo*-PPHD•EF-Tu (EF-Tu, 50 µM, 9 µL; PPHD, 50 µM, 7.2 µL; TDBA, 100 mM, 4.5 µL; ultrapure water 24.3 µL). Samples were aliquoted, frozen and irradiated as in section 2.3.2.2 and then subject to in-gel enzymatic digestion with Trypsin and Chymotrypsin (section 2.3.1.3) prior to analysis (section 2.3.1.5).

2.3.3. Data Analyses

2.3.3.1. Peptide Identification

`RAW` data files acquired from DDA CID MS/MS of precursor peptides were subject to database searching for peptide identification. Files were processed in SearchGUI 4.1.11 and searched against a customised protein database `FASTA` file with the X!Tandem search algorithm.[206,207] Protease digestion was allowed with up to three missed cleavages. Peptide precursor ions between 4 – 40 residues in length and charge states between 1 –4^+ were selected within an error tolerance of 10 ppm, and B and Y fragment ions with a mass error of 1 Da. Carbamidomethylation was selected as a fixed modification on cysteine residues (+58.0215 Da) and deamidation (asparagine, +0.9840 Da) plus oxidation (methionine, +15.9949 Da) selected as variable modifications. The search results were visualised in PeptideShaker 3.0.8, and data analysed in Excel (Microsoft).[208]

2.3.3.2. Peptide-Level Carbene Footprinting Quantification

Analyses of carbene footprinting experiments and quantification of labelling at the peptide level were performed in PepFoot 2.1.1. `RAW` data files were converted to the required `mz5` format in MSConvert (ProteoWizard) and uploaded to PepFoot.[201,209] An *in silico* protease digest was performed (three missed cleavages, peptide length of 4 – 40) and peak matching with a tolerance of 0.02 mmu. Mass spectra were combined from extracted ion chromatograms for each detected peptide (labelled and unlabelled) and their peak areas integrated. Fractional modifications (F_{Mod}) of the peptides were calculated by the ratio of labelled peptide peak area (A) to the sum of labelled and unlabelled peak areas (equation 2.5). F_{Mod} analyses and plotting were performed in Prism 10.1.2, statistical analysis consisted of unpaired Multiple T-tests and corrected for multiple comparisons with the Holm-Šídák method (significance defined as $p < 0.05$).[210,211]

$$F_{Mod} = \frac{A_{Labelled}}{A_{Labelled} + A_{Unlabelled}} \qquad (2.5)$$

2.3.3.3. Residue-Level Carbene Footprinting Quantification

For sub-peptide-level carbene footprinting quantification, `RAW` data files of precursor peptides targeted CID MS/MS were analysed in the Qual Browser suite in Xcalibur 2.0 (Thermo Fisher Scientific). MS/MS spectra were combined from the extracted ion chromatogram of the selected precursor ion. MS/MS spectra were processed manually to identify B and Y fragment ions (both labelled and unlabelled). F_{Mod} for each fragment ion ($F_{Mod}(i)$) was calculated by the ratio of labelled peptide peak intensity (I) to the sum of labelled and unlabelled peak intensities (equation 2.6).

$$F_{Mod}(i) = \frac{I_{Labelled}}{I_{Labelled} + I_{Unlabelled}} \qquad (2.6)$$

$$F_{Mod-Abs} = \left(F_{Mod}(i) - F_{Mod}(i-1)\right) \times F_{Mod} \qquad (2.7)$$

This can be converted further to absolute ($F_{Mod\text{-}Abs}$) by subtraction of the previous sequential fragment ion $F_{Mod}(i\text{-}1)$ and multiplied by the total $\mathrm{F_{Mod}}$ across the whole peptide (equation 2.7). In the cases where $F_{Mod\text{-}Abs}$ returned anomalous, *e.g.* negative, values the residue was discarded and grouped with the previous adjacent residue. If the fragment ion corresponding to a specific residue was not present, that residue was also grouped with the last detected previous residue. $F_{Mod\text{-}Abs}$ analyses and plotting were performed in Prism 10.1.2 (GraphPad), statistical analysis consisted of unpaired Multiple T-tests and corrected for multiple comparisons with the Holm-Šídák method (significance defined as $p < 0.05$).

2.4. Computational Analyses

2.4.1. Protein Homology modelling

Homology models of *E. coli* wild-type, *A. baumannii* wild-type, and *A. baumannii* mutant EF-Tu alongside *P. putida* EF-Tu and PPHD were produced using I-TASSER, AlphaFold2 (DeepMind) and AlphaFold2-Multimer servers.[73,212–215] In all cases `PDB` template files were provided, consisting of codes; 1EFC for EF-Tu in the open configuration, 2BVN for EF-Tu in the closed configuration, 2D8T for EF-Tu in the GE2270A bound state, and 4IW3 for the EF-Tu•PPHD complex. The AlphaFill server was used to transplant missing ligands and co-factors into AlphaFold models.[216]

2.4.2. *In-silico* Protein Digests.

To generate a theoretical list of peptides from a given protein sequence, the webserver based PeptideCutter (Expasy) was used.[217] To produce peptide fragmentation ions, the Peptide Fragment calculator (Institute for Systems Biology) server was used.

2.4.3. Molecular Dynamics Simulations

Molecular dynamics simulations were performed on EF-Tu in the gas-phase to investigate the degree of thermal unfolding, with the aim to corroborate unfolding seen through CIU experiments. All simulations were performed with GROMACS 5.1.2, equipped with the CHARMM36 forcefield and running through the Windows Subsystem for Linux (Ubuntu 18.04.6 LTS), supplementary scripts were run using Python 3.12.[218,219] Simulations were performed CPU-bound on a 16-core CPU operating on 15 cores (AMD Ryzen 9 5950X, 4.9 GHz boost).

2.4.3.1. *Protein File Preparation for Simulation*

EF-Tu homology models, free of bound ligands, were processed using ChargePlacer (designed by Dr Jeddidiah Bellamy-Carter) which assigns protons to the protein to generate the net 12^+ charge state, as observed experimentally. In brief, ChargePlacer randomly protonates chargeable sites (aspartic acid, glutamic acid, lysine, arginine, histidine, and the N- or C-terminal residues) to give the desired charge state, after which the minimum energy of the protonated sequence is calculated. If this minimum energy is lower than the current minimum, the protonated sequence is used to reseed the charge placement algorithm, and the process repeated until a stable energy minimum is reached. Charged protein files were converted into GROMACS compatible file formats (Gromos87, `gro`) with CHARMM36 parameters, using CHARMM-modified TIP3P as a water model, NH_3^+ as the N-termini and COO^- as the C-termini.

To prepare ligand topology, each of the ligand files `pdb` (GDP, GDPNP, enacyloxin IIa) had missing hydrogen atoms added, and were converted to the `mol2` format, and manually altered for compatibility. Ligand topology files (`str`) were produced using the CHARMM General Force Field server before being converted to GROMACS compatible syntax through Python scripting and converted to the `gro` format. *Holo-*

EF-Tu complexes were formed by combining protein and ligand topology files in plain-text, simply by inserting ligand atom coordinates below those of the protein. Simulation topology files were also updated in a similar fashion, to reflect the ligand addition.

2.4.3.2. Molecular Simulations

The prepared proteins were then centred in a large, cubic cell (900 nm^3) and energy minimised *via* steepest descent for 10,000 steps. Long-range interactions were given a large threshold of 300 nm, as recommended by Konermann.[220] The energy minimised proteins were then temperature equilibrated at 298 K over 50 ps, constraining hydrogen bonds (LINCS, iterations = 1, order = 4). Full-length production runs were initiated from the final thermally equilibrated state. Control simulations consisted of 5 ns runs, with the temperature held at 298 K and a simulation step size of 2 fs. Thermal unfolding simulations consisted of a 5 ns run with a ramped linear gradient of 298 – 950 K over 1 ns, followed by a further 4 ns held at 950 K. To aid stability due to the high energy imparted onto the simulation, the step size was reduced to 1 fs for the thermal unfolding simulations. Simulation analyses were performed using in-built GROMACS tools. Output trajectory files were converted to `pdb` format and further analysed in Houdini 20.5.332 (SideFX Software). For comparison with experimentally determined CIU data, `pdb` output trajectories were binned to 2 fs steps and projection approximation (PA) CCS calculations performed by the CCSCalc function within DriftScope 2.0 (Waters). CCS_{PA} was calculated with a gas collision radius of 1.4, and then resulting data multiplied by an empirical value of 1.14 to give a corrected value, CCS_{CALC}, which can be compared to the experimentally derived $^{TW}CCS_{N2\rightarrow He}$ values.[221] A modified atom list was used in CCSCalc to accept non-proteinaceous atom types, such as those found in bound ligands.

Mapping Interactions of EF-Tu and the Elfamycin GE2270A using Carbene Footprinting

3

3.1. Introduction

3.1.1. GE2270A

The thiazolyl peptide GE2270A belongs to the pulvomycin-like class of elfamycins, acting mostly as a competitive inhibitor of aa-tRNA binding.[69,222] GE2270A was first isolated and characterised by Selva *et al.* in the early 1990's, alongside structural elucidation by Kettenring *et al.*[222,223] The compound was isolated from a strain (ATCC 53773) of the relatively rare, soil-dwelling Actinomycetota, *Planobispora rosea.* At the time of discovery, and to this day, the *Planobispora* genus remains relatively uncharacterised, it is a genetically intractable genus with only four formally defined species and a single sequenced genome.[224] GE2270A was isolated though fermentation growth and found to be present in both the supernatant, and mycelium of the bacteria. A combined 50 L ethyl acetate extraction of the supernatant and 20 L methanol extraction of the mycelium yielded 6.06 g of crude GE2270A. This was purified via column chromatography, yielding approximately 1.2 g of pure compound from an initial 50 L fermentation culture. Since then, GE2270A has since been found to be produced by several other unclassified *Nonomuaea* species, another rare Actinomycetota, alongside other thiopeptide-based elfamycins, thiomuracin A – H.[225]

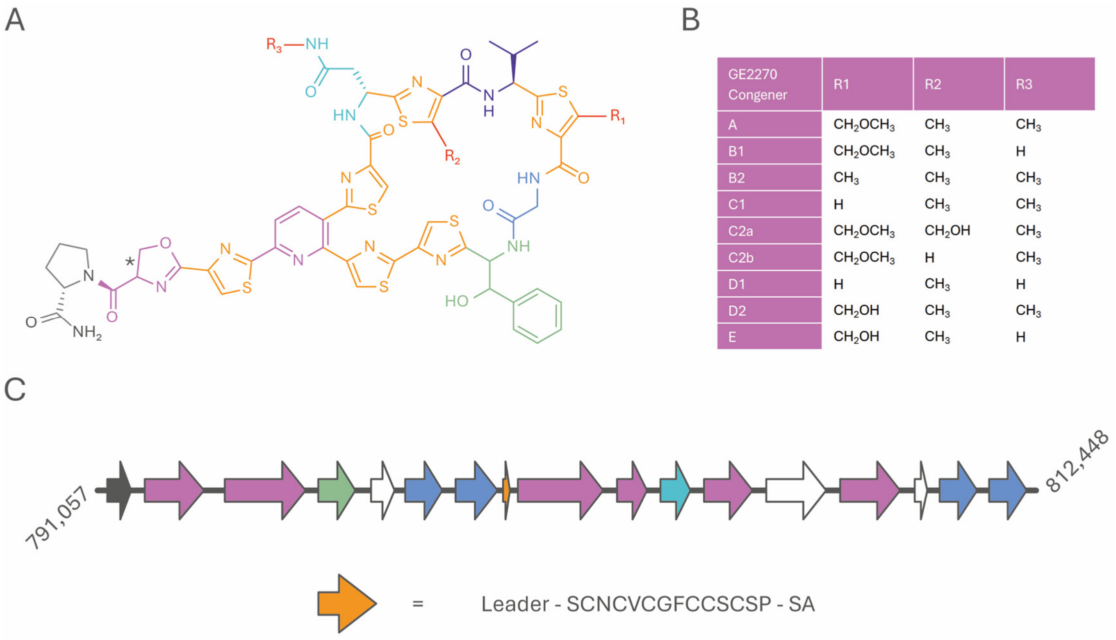
3.1.1.1. Structure and Biosynthesis of GE2270A

GE2270A is a highly modified ribosomal peptide, consisting of six thiazole, one oxazoline and one pyridine rings, formed into a macrocyclic structure (figure 3.1, A). It is the best studied member of the GE2270 class of compounds produced by *P. rosea* (E, D1, D2, C1, C2a, C2b, B1, A and T), which differ in the constituents of five functional groups (figure 3.1, B).[226] The total biosynthetic pathway has recently been characterised in *P. rosea,* and heterologous expression of the entire 21.4 kb (figure 3.1, C) gene cluster, *pbt*, in a more genetically tractable *Nonomuaea* strain (ATCC 39727) through a cosmid[†] system yielded successful biosynthesis. The gene cluster encodes for 16 individual proteins (excluding that which encodes the precursor peptide) composed of various dehydratases, monooxygenases, methyltransferases and radial *S*-adenosyl methionine (SAM) enzymes.[227] The cluster is located in a highly conserved region of the genome, flanked by upstream DNA-directed RNA polymerase subunits and downstream by EF-G and EF-Tu coding sequences. The cysteine-rich precursor peptide, SCNCVCGFCCSCSPSA, is thought to first undergo thiazole synthesis followed by oxazoline synthesis, cleavage of C-terminal serine and alanine, and macrocyclization by pyridine ring formation from Ser^1 and Ser^{11}.[227] Congener defining methylation and hydroxylation of specific sites is thought to take place on the post-thiazole linear structure, prior to macrocyclization. Interestingly, *pbtA*, the gene encoding the precursor peptide, is one of the three most expressed genes during *P. rosea* growth. In fact, after 48 hours of growth, it became the most expressed gene as measured by transcriptomic analysis.[224] Expression of the necessary enzymes is

---

[†] A cosmid refers to a large plasmid containing a λ phage *cos* sequence to allow for bacteriophage packing and transduction.[315]

much slower, and *ptbA* peak expression occurs when sufficient concentrations of the enzymes have accumulated, in a so-called *just-in-time* transcription strategy.[224]



GE2270 Congener	R1	R2	R3
A	CH_2OCH_3	CH_3	CH_3
B1	CH_2OCH_3	CH_3	H
B2	CH_3	CH_3	CH_3
C1	H	CH_3	CH_3
C2a	CH_2OCH_3	CH_2OH	CH_3
C2b	CH_2OCH_3	H	CH_3
D1	H	CH_3	H
D2	CH_2OH	CH_3	CH_3
E	CH_2OH	CH_3	H

Figure 3.1 | GE2270A overview. A | General GE2270 structure coloured by precursor residue and variable groups labelled R_1 – R_3 (grey, proline; pink, serine; orange, cysteine; green, phenylalanine; blue, glycine; purple, valine; cyan asparagine). Congener GE2270T differs from GE2270A solely by the presence of a double bond in the oxazoline ring (asterisk). B | Variable groups for each GE2270 congener (T not included). C | *ptb* biosynthetic gene cluster coloured by general function alongside the precursor peptide sequence (grey, regulator; pink, dehydratase; green, P450; white, hypothetical; blue, methyltransferase; orange precursor peptide). Start and end base numbers highlighted.

3.1.1.2. Mode of Action

As a pulvomycin-like elfamycin, GE2270A inhibits formation of the EF-Tu•GTP•aa-tRNA ternary complex, acting mostly upon Gram-positive bacteria.[222] GE2270A differs substantially from pulvomycin in structure and EF-Tu binding. GE2270A binds with moderately affinity ($K_D < 3$ μM) to EF-Tu in domain II, of both the GDP and GTP-bound protein through four or five hydrogen bonds respectively (figure 3.2, A). In the GDP-bound state of *Thermus thermophilus* EF-Tu, GE2270A forms hydrogen bonds with Asp^{227}, Thr^{239}, Arg^{234} and His^{273} all within domain II. Asp^{227} is found at the N-terminus of the β-strand *b2,* Thr^{239} at the N-terminus of strand c*2*, and Arg^{234} and His^{273} within connecting loops of *b2* – *c2* and *g2* – *h2* respectively.[33,68] In the GTP-bound state of EF-Tu, an additional hydrogen bond is formed with Gln^{98} in domain I, proximal to the C-terminus of helix *B*.[68]

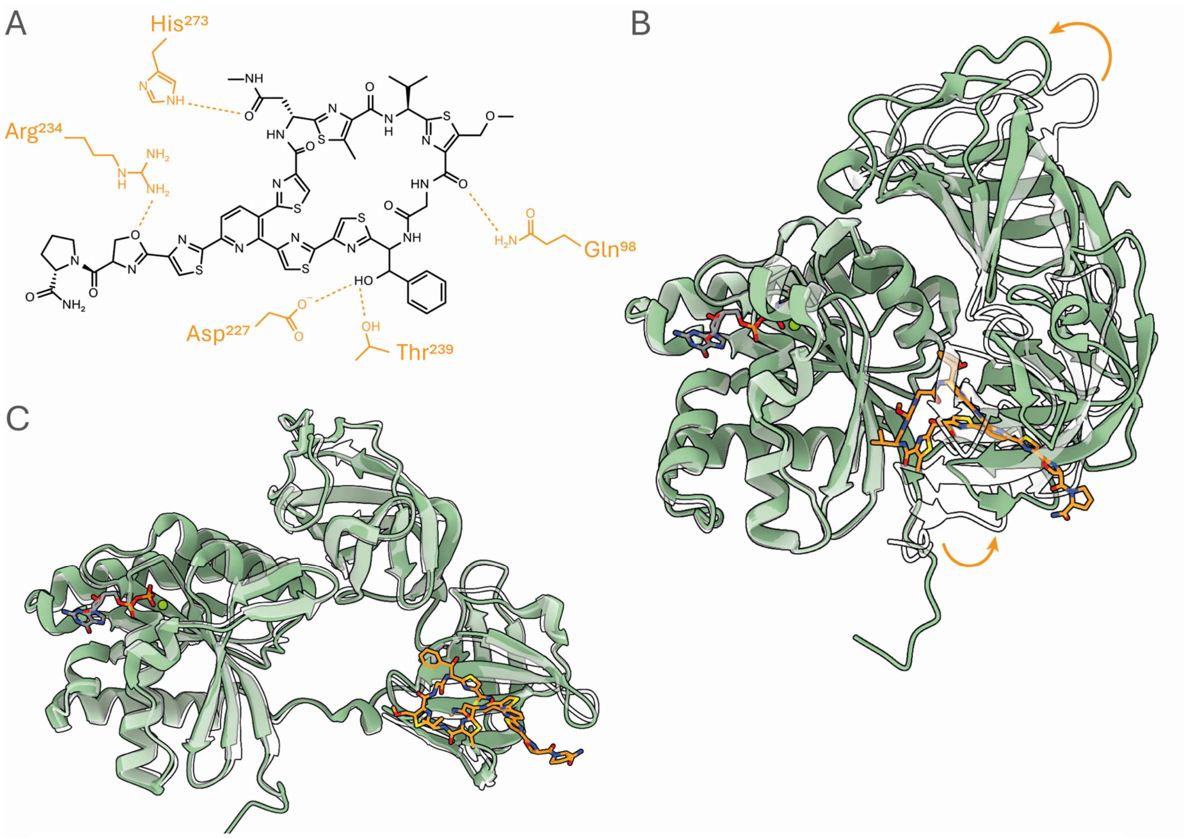


Figure 3.2 | Structural GE2270A overview. A | GE2270A with the five possible hydrogen bonding residues, dotted lines represent hydrogen bonds (not to scale) from *T. thermophilus* EF-Tu. B | Closed EF-Tu•GDPNP•GE2270A (green, PDB 2C77) overlayed on EF-Tu•GDPNP (white, PBD 1EXM) showing widening of the domains I – II interface and the associated movements of domains II and III caused by GE2270A (orange). C | Open EF-Tu•GDP•GE2270A (green, PDB 1D8T) overlayed on EF-Tu•GDP (white, PBD 1EFC) showing minimal structural perturbations caused by GE2270A (orange).

Binding of GE2270A to EF-Tu•GDP causes minimal structural perturbations outside of domain II (figure 3.2, C); however, its large size does prevent full closure of EF-Tu (figure 3.2, B). The inability for EF-Tu to fully close, alongside GE2270A bonding to His^{273} prevent aa-tRNA from forming the ternary complex with EF-Tu•GTP. His^{273} is involved in direct hydrogen bonding with the amino acid carried by aa-tRNA.[34] Unlike pulvomycin, GE2270A does not induce any significant auto-hydrolytic GTPase activity, the contacts made with the switch regions in domain I are minor compared to pulvomycin. GE2270A alongside pulvomycin, can form a stable complex with both EF-Tu•GDP•EF-Ts and EF-Tu•GTP•EF-Ts, and in fact stimulates the release of GDP from the complex, whilst allowing exchange to GTP.[70,228]

3.1.1.3. EF-Tu Resistance to GE2270A

P. rosea natively encodes a GE2270A-resistant form of EF-Tu, EF-Tu1, which is coincidently also resistant to kirromycin. This is alongside an inessential copy, EF-Tu3, which is thought to act as a decoy, limiting binding to the, albeit resistant, EF-Tu1.[229] *Streptomyces ramocissimus,* the native producer of kirromycin, encodes three separate EF-Tus, one of which (EF-Tu3) demonstrates GE2270A resistance.[230,231] Both exhibit multiple residue substitutions across their sequences but two specific mutations, namely G257S and G275A (strands *g2* and *j2* respectively), were found to induce ten-fold GE2270A resistance in *E. coli*, with Ala^{275} being principally responsible.[231] Interestingly, this resistance prevents GE2270A binding only to open EF-Tu, binding to the closed conformation was still seen, likely stabilised by residues homologous to Gln^{98} in domain I. Resistant EF-Tu•GTP•GE2270A was even able to successfully bind aa-tRNA, albeit at two-fold lower affinity.[231] A similar resistance conferring mutation (G257E) has been found in *Clostridium difficile* exposed to the semi-synthetic GE2270A analogue, LFF571, showing spontaneous resistance is possible, albeit at low rates ($\leq 1.2 \times 10^{-9}$ resistant:susceptible).[232]

3.2. Aims

The aim of this work was to characterise GE2270A binding to wild-type EF-Tu from a clinically relevant pathogen, *A. baumannii*, through native mass spectrometry, and to explore the ability of carbene footprinting to map the binding site of GE2270A to domain II of EF-Tu. Although the crystal structure of GE2270A-bound EF-Tu has been solved, only several deposited structures exist. This, paired with the intrinsic inability of X-ray crystallography to capture protein dynamics discussed in chapter 1, has limited investigation into GE2270A-mediated EF-Tu dynamics. Carbene footprinting has been previously shown to capture protein conformational dynamics and as such, presents itself as a promising methodology to investigate the conformational changes induced in EF-Tu by binding of GE2270A.

3.3. Results and Discussion

3.3.1. Native Mass Spectrometry of EF-Tu with GE2270A

Wild-type *A. baumannii* EF-Tu was heterologously expressed in *E. coli* as a 6xHis-tagged protein, before cleavage with TEV protease as described in section 2.1.3. GE2270A was purchased as a solid and prepared in DMSO. For native mass spectrometry, EF-Tu•GDP•GE2270A samples were prepared as described in section 2.2.1.5, resulting in final concentrations of 18.3 and 20 µM EF-Tu•GDP and GE2270A respectively. Spectra were acquired using nESI emitter tips, as prepared in section 2.2.1.1, on a Synapt G1 HDMS High-Definition mass spectrometer. A typical native spectrum is shown in figure 3.3, A. Typically, nESI EF-Tu spectra exhibited three principal charge states: $[M + 11H]^{11+}$, $[M + 12H]^{12+}$, and $[M + 13H]^{13+}$, with the $[M + 12H]^{12+}$ charge state being dominant under these conditions.

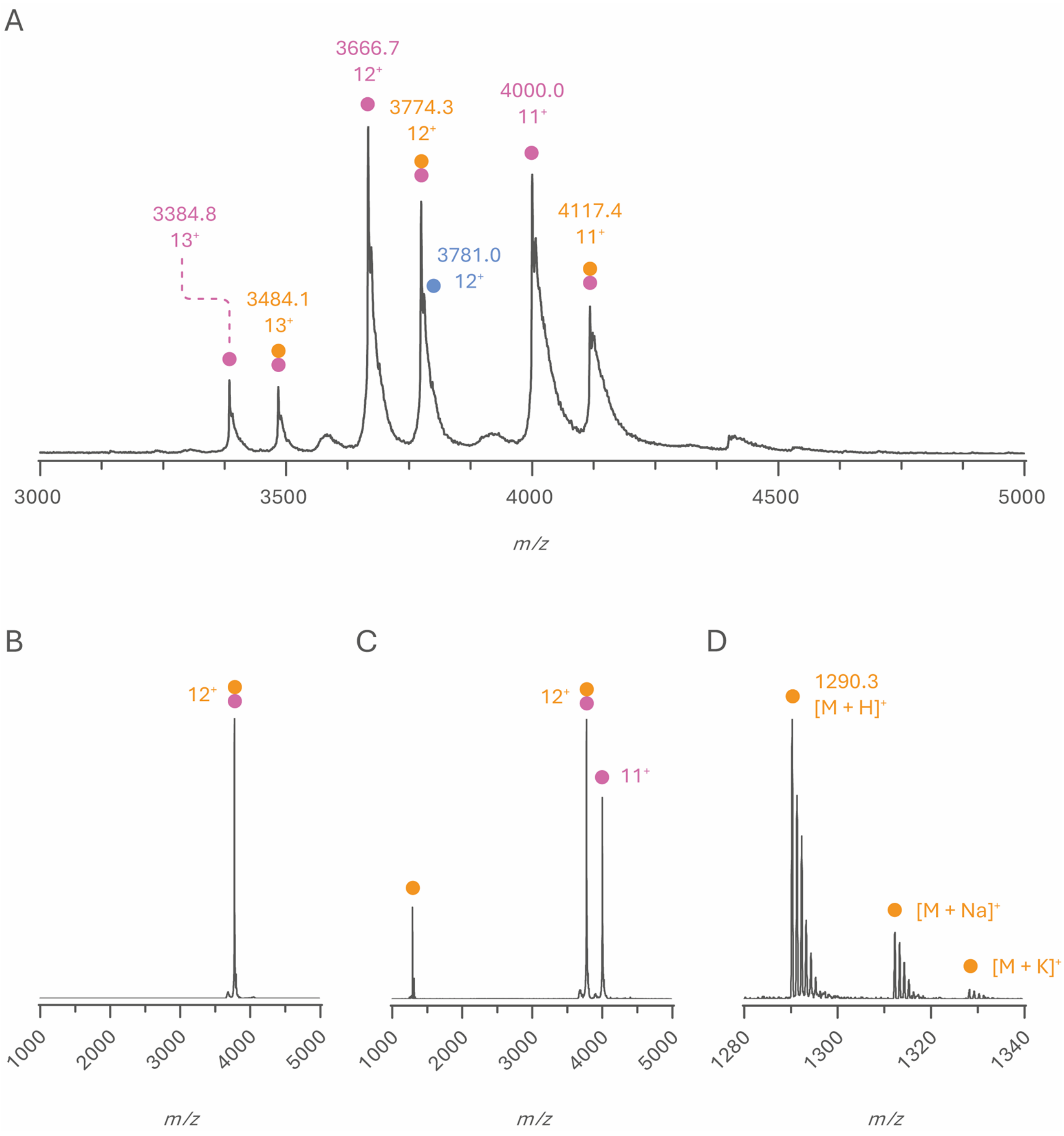


Figure 3.3 | Normalised native mass spectra of EF-Tu in complex with GDP and GE2270A. A | nESI produced three charge states: $[M + 11H]^{11+}$, $[M + 12H]^{12+}$, and $[M + 13H]^{13+}$. Each charge state contains EF-Tu•GDP (pink circle) and EF-Tu•GDP•GE2270A (pink and orange circles). A representative DMSO adduct peak is indicated by the blue circle. B | Quadrupole isolated EF-Tu•GDP•GE2270A $[M + 12H]^{12+}$ ion, with no collisional activation. C | Collisional activation (50 V) of EF-Tu•GDP•GE2270A $[M + 12H]^{12+}$ ejects free GE2270A (orange circle) and charge-reduced EF-Tu•GDP $[M + 11H]^{11+}$ (pink circle). D | Zoom of ejected GE2270A showing protonated, sodiated and potassiated ions.

Within each charge state two species are observed: EF-Tu•GDP and EF-Tu•GDP•GE2270A, with the latter accounting for approximately 42% of the total protein ions. Deconvolution of the charge states results in masses of 43,988 ± 0.50 Da and 45,280 ± 0.43 Da for EF-Tu•GDP and EF-Tu•GDP•GE2270A respectively. The calculated mass based upon the *A. baumannii* wild-type EF-Tu sequence of 43,522 Da, compared to the observed EF-Tu•GDP deconvoluted mass, gives a deviance of 466 Da. This can be explained by the presence of bound GDP (443.20 Da) and the magnesium cofactor (24.31 Da) with some small error (1.5 Da). The binding of GE2270A to EF-Tu•GDP imparts a theoretical mass shift of 1290 Da, with the calculated molecular weight of GE2270A being 1290.5 Da. Each EF-Tu•GDP and EF-Tu•GDP•GE2270A peak is accompanied by a secondary shoulder peak with a deconvoluted mass shift of +78 Da, corresponding to an adduction by DMSO.

Quadrupole isolation of the EF-Tu•GDP•GE2270A $[M + 12H]^{12+}$ charge state (figure 3.3, B) followed by collision-induced dissociation liberated ionised GE2270A, evidenced by detection of $[M + 11H]^{11+}$ EF-Tu•GDP ions (figure 3.3, C). In some instances, small amounts of GE2270A were ejected as a neutral ion, resulting in the emergence of $[M + 12H]^{12+}$ EF-Tu•GDP ions, but this behaviour was inconsistent. Examination of liberated GE2270A (figure 3.3, D) revealed a dominant ion with a *m/z* of 1290.3, corresponding to protonated GE2270A. Alongside this, minor sodiated and potassiated ions were seen at 1312.3 and 1328.2 *m/z* respectively.

The quadrupole isolated EF-Tu•GDP•GE2270A $[M + 12H]^{12+}$ ion was subject to a range of collision voltages (10 – 80 V) to probe the stability of the complex (figure 3.4). At each collision voltage, the proportion of EF-Tu•GDP•GE2270A to total EF-Tu ions was calculated. At a collision voltage of 10 V, no GE2270A ejection was seen, with noticeable ejection beginning at approximately 30 V. Above 65 V there was little further GE2270A ejection, and between 70 – 80 V, approximately 20% of EF-Tu•GDP remains bound to GE2270A. At collision voltages greater than 80 V, spectra quality degraded significantly, and it became impossible to accurately quantify the proportion of EF-Tu•GDP bound to GE2270A.

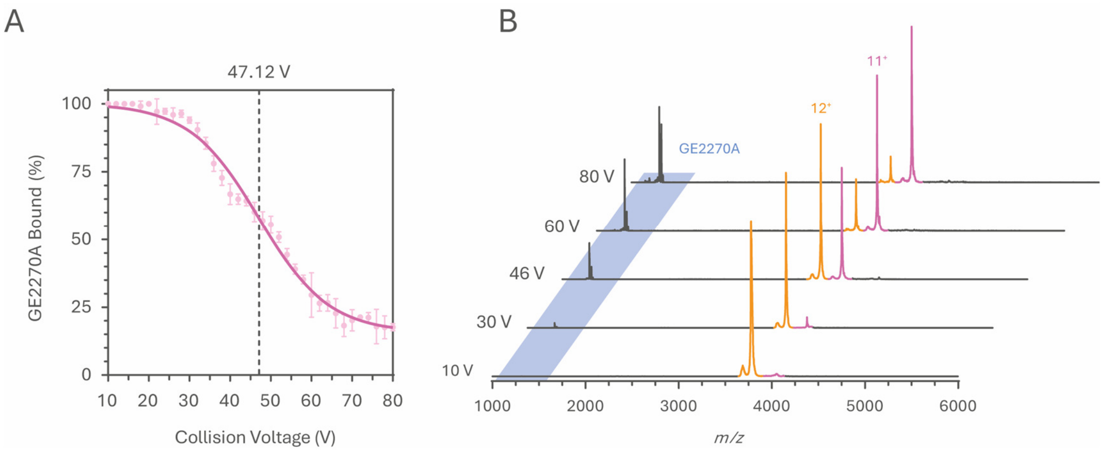


Figure 3.4 | CID of EF-Tu•GDP•GE2270A. A | CID of the quadrupole isolated EF-Tu•GDP•GE2270A $[M + 12H]^{12+}$ ion, with voltages of 10 – 80 V. Data is an average of at least 3 repeats. Error bars represent one standard deviation. B | Representative normalised spectra from A, showing liberation of GE2270A (blue shading), emergence of charge-reduced EF-Tu•GDP $[M + 11H]^{11+}$ (pink) and disappearance of EF-Tu•GDP•GE2270A $[M + 12H]^{12+}$ (orange) as the collision voltage increases.

In CIU IMS-MS experiments performed on *E. coli* EF-Tu•GDPNP•ENX (see chapter 6), EF-Tu was found to undergo a significant unfolding event when subject to collisional voltages between 40 – 50 V. Complete GE2270A ejection may be hindered by the collapse of EF-Tu tertiary structure around the ligand.

Symmetrical sinusoidal non-linear regression analysis of the data (figure 3.4, A) allows for the extraction of CE_{50}, which refers to the collision voltage at which the survival yield of the original ion is 50%.[233] For EF-Tu•GDP•GE2270A, this was found to be 47.12 ± 0.9 V, comparable to that calculated for enacyloxin IIa (42.22 ± 0.2 V, chapter 4). Both are large compounds, with similar numbers of non-bonding contacts and hydrogen bonds formed upon binding to EF-Tu (21 contacts, 4 hydrogen bonds and 16 contacts, 6 hydrogen bonds for GE2270A and enacyloxin IIa respectively). Representative spectra from the ramped collision energies (figure 3.4, B) show the simultaneous emergence of ejected GE2270A (protonated, sodiated and potassiated, figure 3.3, D) and EF-Tu•GDP $[M + 11H]^{11+}$ ions, alongside the disappearance of EF-Tu•GDP•GE2270A $[M + 12H]^{12+}$ ions.

3.3.2. Carbene Footprinting of EF-Tu with GE2270A

Carbene footprinting of *A. baumannii* EF-Tu was performed in the presence and absence of GE2270A. The choice of proteolytic enzyme, TDBA concentration and irradiation time were based upon the initial EF-Tu carbene footprinting optimisations, which can be found in chapter 4. Carbene footprinting was performed as described in section 2.3.2.3, where 10 µM of EF-Tu•GDP was incubated with 12 µM of GE2270A, followed by irradiation in the presence of 10 mM TDBA for 15 seconds.

Irradiated protein was digested in-solution with trypsin, and the peptides analysed by HPLC-MS. In total, 48 individual tryptic peptides were identified, representing 87.5% of the complete *A. baumannii* EF-Tu sequence, including multiple large peptides with three missed cleavages (figure 3.5). Of those identified peptides, 3 showed a significant reduction in F_{Mod} in EF-Tu•GDPNP•GE2270A and 10 showed a significant reduction in F_{Mod} in EF-Tu•GDP•GE2270A, when compared to the EF-Tu•GDP control. Significance was defined as $p < 0.05$ as calculated by a Multiple T-test and corrected for multiple comparisons with the Holm-Šídák method. EF-Tu•GDP peptides with a $F_{Mod} < 0.05$ were disregarded from statistical analyses (section 2.3.3.2).

Interestingly, in both EF-Tu•GDPNP•GE2270A and EF-Tu•GDP•GE2270A, no significant reductions in F_{Mod} were seen in the three peptides containing residues involved in hydrogen bonding to GE2270A in domain II (figure 3.5 and figure 3.6, B). The peptides Ala^{214} – Arg^{232} and Ala^{218} – Arg^{232} showed very low labelling in the control and ligand-treated samples in contrast to peptide Ala^{218} – Arg^{239}, which was instead close to 100% labelled in all samples. These large peptides span the linker between domains I and II whilst incorporating strands *a2* and *b2*, where hydrogen bonding occurs.

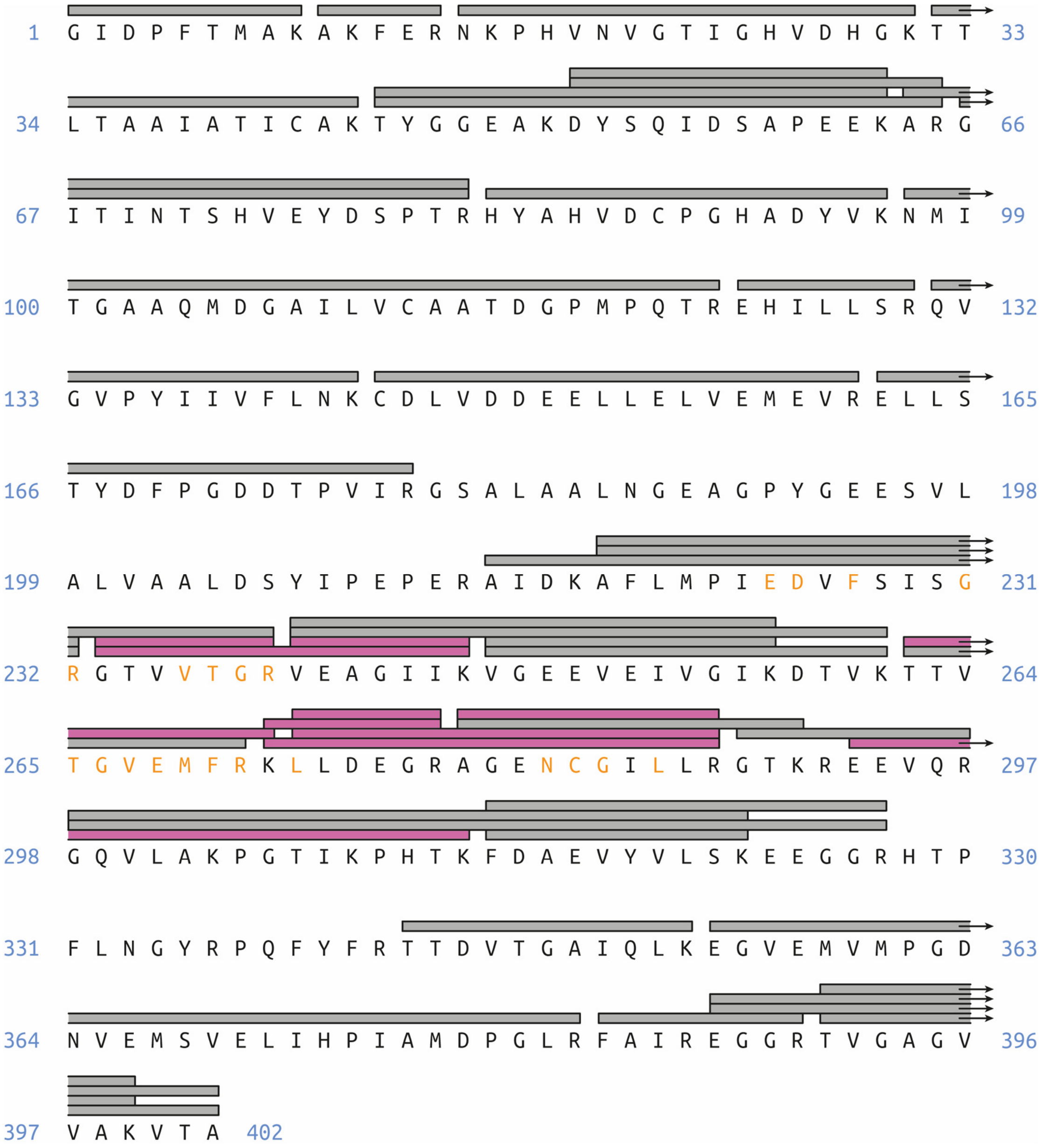


Figure 3.5 | Tryptic peptide coverage of *A. baumannii* EF-Tu post-carbene footprinting. Peptides coloured grey showed no significant reductions in F_{Mod} in EF-Tu•GDP•GE2270A compared to EF-Tu•GDP, and those coloured pink did show significant reductions. Residues coloured orange are those involved in non-bonding contacts and hydrogen bonding with GE2270A.

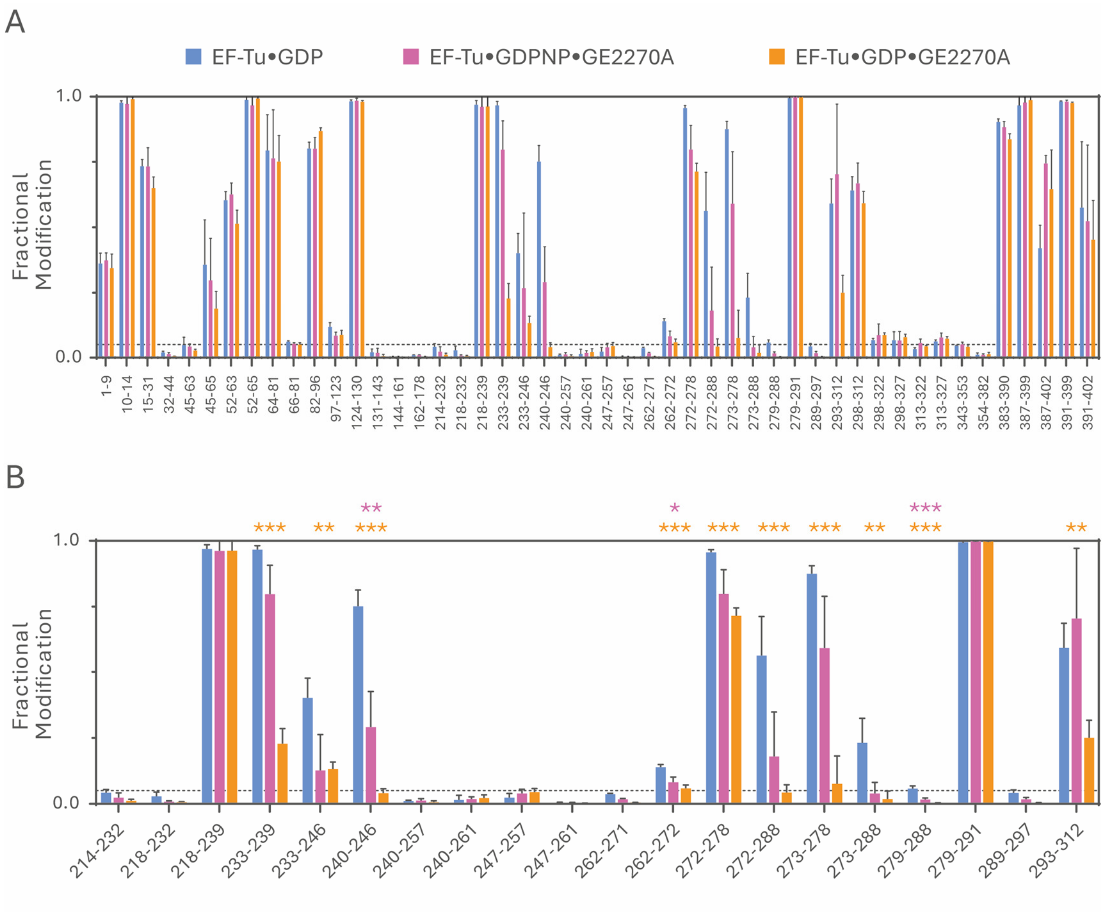


Figure 3.6 | F_{Mod} histograms for EF-Tu•GDP (blue), EF-Tu•GDPNP•GE2270A (pink) and EF-Tu•GDP•GE2270A (orange). A | All 48 tryptic peptides reported. B | Zoom of domain II tryptic peptides. Data is an average of 5 replicates, error bars represent standard deviation. Stars indicate significant reductions in F_{Mod}, compared to EF-Tu•GDP and coloured as above. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Dotted line indicates an F_{Mod} of 0.05.

On balance, there were fewer significant reductions in F_{Mod} in EF-Tu•GDPNP•GE2270A compared to EF-Tu•GDP•GE2270A (figure 3.6, B). As EF-Tu•GDPNP on average adopts a more closed conformation then EF-Tu•GDP, consequently there will be less binding of GE2270A. Although EF-Tu•GDPNP•GE2270A has been crystalised, the GE2270A binding region in domain II becomes slightly occluded by domain I due to the closing of domains II and III. Therefore, it is likely that GE2270A binding occurs in a more open conformation, followed by closure.[69] Although there are fewer significant reductions in F_{Mod}, it is worth noting that in all but one peptide, Glu^{293} – Lys^{313}, F_{Mod}

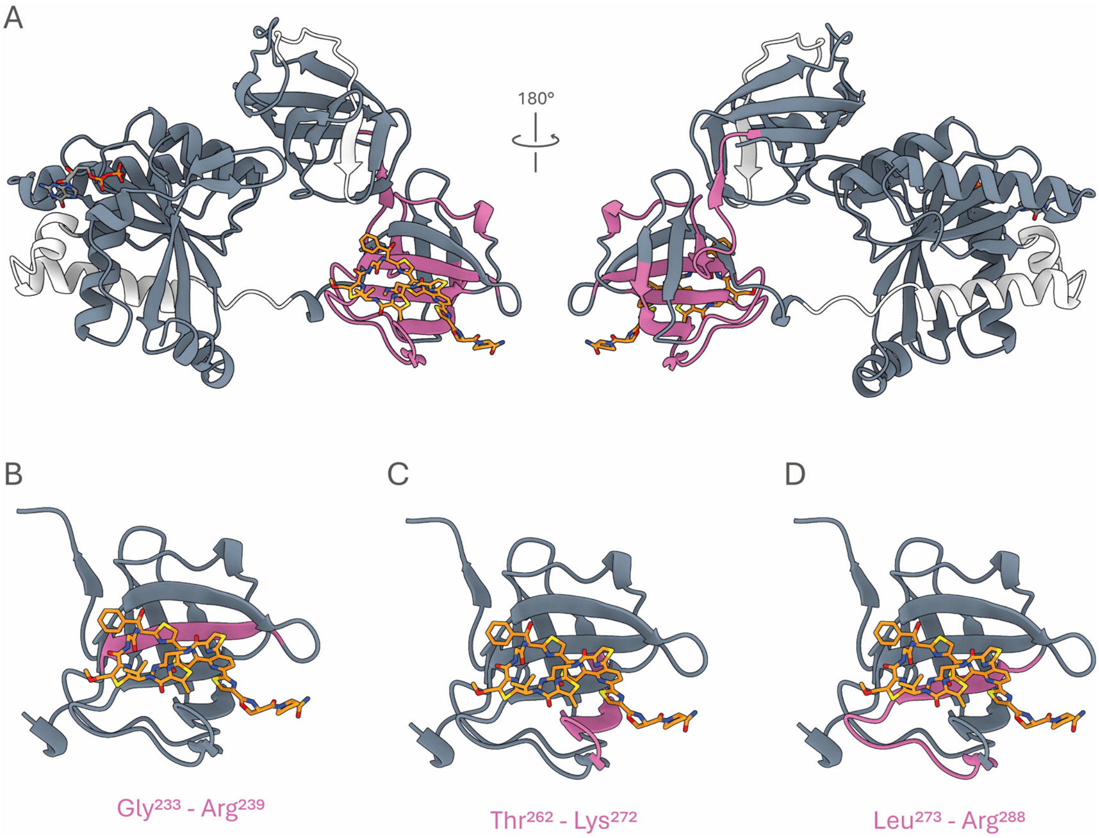


Figure 3.7 | EF-Tu•GDP•GE2270A F_{Mod} data mapped to an open *A. baumannii* EF-Tu AlphaFold 2 homology model. A | Peptides with significantly reduced F_{Mod} (pink) localise to GE2270A (orange) binding site in domain II. Peptides in grey showed no significant differences when compared to EF-Tu•GDP. Peptides in white were not detected by HPLC-MS. B | Zoom of domain II highlighting Gly^{233}–Arg^{239} (pink). C | Zoom of domain II highlighting Thr^{262}–Lys^{272} (pink). D | Zoom of domain II highlighting Leu^{273}–Arg^{288} (pink). These three peptides contain all EF-Tu residues contacting GE2270A.

showed qualitative, non-significant reductions compared to EF-Tu•GDP implying GE2270A binding, albeit at lower efficacy. Peptide Thr^{262} – Lys^{272} showed a small, but significant reduction in F_{Mod} (-0.058, $p = 0.04$). This peptide, encompassing the C-terminus of strand *f2*, strand *g2* and the N-terminus of strand *h2* contains seven residues (Thr^{265}, Gly^{266}, Val^{267}, Glu^{268}, Met^{269}, Phe^{270} and Arg^{271}) homologous to those contacting (< 3.8 Å) GE2270A in *E. coli* and *T. thermophilus*. These residues are all found within strand *f2* and the loop connecting to *h2*.[69] Peptide Ala^{279} – Arg^{288} shows a small, but highly significant reduction in F_{Mod} (-0.041, $p < 0.001$). The peptide solely encompasses strand *j2* and contains a further four residues contacting GE2270A (Asn^{282}, Cys^{283}, Gly^{284}, Lys^{286}).

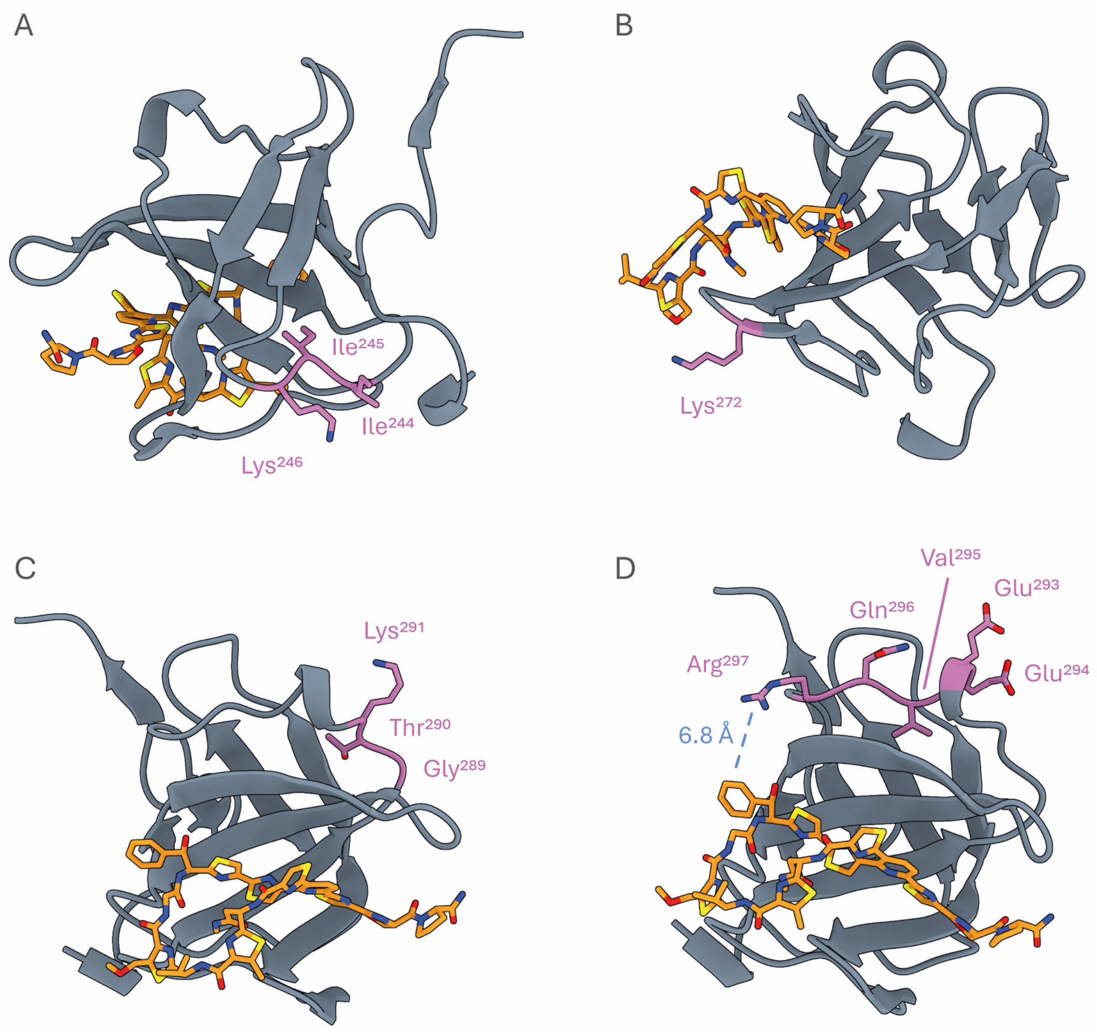


Figure 3.8 | Residues of interest in domain II A | Ile^{244}, Ile^{245} and Lys^{246} which show significant masking in EF-Tu•GDP•GE2270A despite being present on the opposite side of domain II to GE2270A. B | Lys^{272} remains heavily labelled in EF-Tu•GDP•GE2270A due to its accessible position. C | Gly^{289}, Thr^{290} and Lys^{291} also remain highly labelled in EF-Tu•GDP•GE2270A. D | Glu^{293} – Arg^{297} which show masking in EF-Tu•GDP•GE2270A despite their position distal to the GE2270A binding site.

EF-Tu•GDP•GE2270A showed significant F_{Mod} reductions in 10 peptides, split coarsely into three regions (figure 3.6, B and figure 3.7, A, B). The first spanning strand *c2* and the connecting loop to *d2*. Peptide Gly^{233} – Arg^{239} contains four GE2270A contacting residues (Val^{236}, Thr^{237}, Gly^{238} and Arg^{239}) and showed a highly significant reduction in F_{Mod} compared to EF-Tu•GDP (-0.739, p < 0.001, figure 3.7, B). Overall, the missed cleavage peptide Gly^{233} – Lys^{246} showed reduced labelling compared to Gly^{238} – Arg^{239}, but still showed significantly reduced F_{Mod} (-0.269, $p = 0.01$).

Interestingly, the correctly cleaved Val^{240} – Lys^{246} showed highly significant F_{Mod} reduction in both EF-Tu•GDP•GE2270A and EF-Tu•GDPNP•GE2270A (-0.711, $p < 0.001$ and -0.461, $p = 0.002$ respectively). This peptide is spatially far removed from GE2270A, on the opposite side of domain II. Sub-peptide MS^2 analysis places TDBA insertion on Ile^{244}, Ile^{245} and Lys^{246}, which are within the loop connecting strands *f2* and *i2* (figure 3.8, A and 3.9, C). Strands *j2* and *c2* form the bulk of the GE2270A binding pocket, so it is possible that GE2270A-induced perturbations in those strands propagates to the connecting loops in such a way as to shield Ile^{244}, Ile^{245} and Lys^{246} from TDBA insertion. Unfortunately, as a low intensity, minimally labelled peptide in GE2270A-bound EF-Tu, targeted MS^2 was not feasible to elucidate TDBA insertion in the *holo* protein.

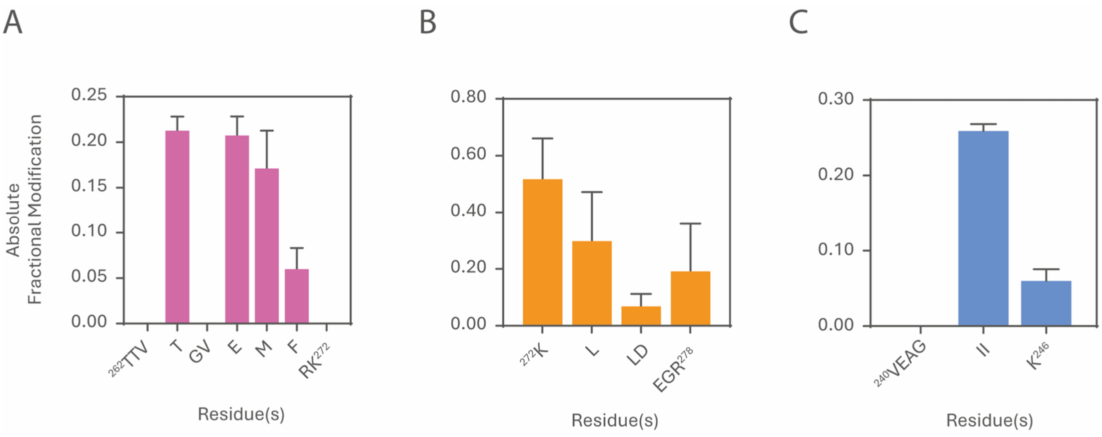


Figure 3.9 | MS^2 of EF-Tu•GDP peptides showing labelling distribution within each peptide. A| Peptide Thr^{262} – Lys^{272}. B | Peptide Lys^{272} – Arg^{278}. C | Peptide Val^{240} – Lys^{246}. Data is an average of 3 replicates, error bars represent standard deviation.

Peptide Thr^{262} – Lys^{272} (figure 3.7, C) contains the majority of contacting residues and, as in EF-Tu•GDPNP•GE2270A, showed a relatively small, yet highly significant reduction in F_{Mod} (-0.080, $p < 0.001$). This peptide is a missed cleavage of the previous peptide, Thr^{262} – Arg^{272}, which demonstrated even lower F_{Mod}, below the threshold for statistical analysis.

The presence of the additional basic lysine in the missed cleavage peptide may induce greater labelling due to the carboxylic acid group in TDBA conferring a higher affinity towards basic residues.[195] MS^2 analysis narrows the probe insertion to Thr^{265}, Glu^{268}, Met^{269} and Phe^{270}, all residues contacted by GE2270A (figure 3.9, A).

Peptides Lys^{272} – Arg^{278} and Leu^{273} – Arg^{278} both contain the GE2270A contacting Leu^{273} and show significant reductions in F_{Mod}, the latter having a much higher reduction (-0.243 $p < 0.001$ and -0.799, $p < 0.001$ respectively). The additional lysine in Lys^{272} – Arg^{278} protrudes out of strand *h2* beyond the GE2270A binding pocket, providing easy access to TDBA, which may explain its higher overall labelling (figure 3.8, B). MS^2 analysis placed TBDA labelling predominantly on Lys^{272} and Leu^{273} (figure 3.9, B). The next series of peptides; Lys^{272} – Arg^{288}, Leu^{273} – Arg^{288} (figure 3.7, D) and Ala^{279} – Arg^{288} all show significant reductions in F_{Mod}, (-0.520, $p < 0.001$; -213, $p = 0.02$; -0.055, $p < 0.001$) with all three-encompassing strand *j2*, containing the GE2270A contacting residues Asn^{282}, Cys^{283}, Gly^{284} and Leu^{286}. Interestingly, peptide Ala^{279} – Lys^{291}, which also covers those four residues, does not show significant masking. The additional three C-terminal residues are basic and highly accessible, on the top of domain II and are likely to be labelled irrespective of GE2270A being bound (figure 3.8, C). Finally, the large double missed cleavage peptide Glu^{293} – Lys^{312} demonstrates a curious significant reduction in F_{Mod} (-0.037, $p = 0.001$) in EF-Tu•GDP•GE2270A but not EF-Tu•GDPNP•GE2270A, where there was a high standard deviation in F_{Mod}. This peptide spans the mostly disordered region connecting domains II and III and is located directly after the C-terminus of strand *j2*. It is possible that structural perturbations in *j2* due to GE2270A binding propagate through this region, resulting in masking. Unfortunately, the peptide is of very low intensity, meaning direct MS^2 analysis was not possible. However, when comparing to the adjacent peptides Gly^{298} – Lys^{322} and Gly^{298} – Lys^{327} (figure 3.6, A) which show insignificant changes in F_{Mod} we can tentatively assume the masking occurs within the first five residues, which are proximal to the binding site of GE2270A (figure 3.8, D).

3.4. Conclusions

Here, I have explored the binding, and interactions, of GE2270A to *A. baumannii* EF-Tu through native mass spectrometry and carbene footprinting. The susceptibility of *A. baumannii* to GE2270A has not been studied, but as resistance conferring mutations prevent GE2270A binding in the GDP conformation of EF-Tu it can be assumed that this EF-Tu is sensitive. Native mass spectrometry demonstrated clear binding of GE2270A to EF-Tu•GDP, which could be dissociated under moderate collisional energy. Through carbene footprinting, the binding site of GE2270A could be pinned down to domain II, in agreement with published structures of *E. coli* and *T. thermophilus* GE2270A-bound EF-Tu.[69,228] Tryptic peptides containing residues homologous to those known to form non-bonding contacts with GE2270A in *E. coli* and *T. thermophilus* were found to be significantly masked in the presence of GE2270A in *A. baumannii* EF-Tu. Masking was also seen in peptides proximal to those directly contacting GE2270A. This potentially hints at subtle conformational dynamics induced by GE2270A across domain II, which would be difficult to elucidate from the two published crystal structures of EF-Tu•GE2270A. Overall, carbene footprinting was successful at identifying the GE2270A binding site, and may give clues to more subtle induced conformational changes in EF-Tu.

Investigating Interactions of EF-Tu and the Elfamycin Enacyloxin IIa using Native Mass Spectrometry and Carbene Footprinting

4

4.1. Introduction

4.1.1. Enacyloxin IIa

The kirromycin-like elfamycin, enacyloxin IIa, was first isolated as a congener of an assortment of polyenic compounds produced by *Frateuria* sp. W-315 (reported at the time as *Gluconobacter* sp. W-315) by Watanabe *et al.* in 1982.[234] The genus *Frateuria,* consists of a six characterised and a multitude of uncharacterised species. A Gram-negative Xanthomonadale, an order consisting of several phytopathogens, W-315 was first isolated from a soil sample and initially characterised based upon its growth inhibition of the fungi *Neurospora crassa* and *Aspergillus oryzae*.[235] The antibiotic mix, designated AB-315, was extracted from W-315 fermentation broth cultured in spent *N. crassa* media by diethyl ether extraction, yielding approximately 100 mg L^{-1} of crude powder.[234] HPLC analysis identified eight individually eluting peaks, with the second being the most intense (by 365 nm UV detection). The powder demonstrated inhibition of Gram-negative and Gram-positive bacterial growth and had no effect on *N. crassa* or *Saccharomyces cerevisiae*. The same group elucidated the NMR structure of the major component, named enacyloxin II and subsequently enacyloxin IIa, in 1990 as a 3,4-dihydroxycyclohexanecarboxylic acid conjugated to a chlorinated polyene sidechain.[236] During the course of their work, the group elucidated the structure of enacyloxin IVa, and in 2007 the structures of enacyloxin Ia and IIIa (figure 4.1, F, A and C respectively) were published. The structures of a further two constituents separated by HPLC were found to be decarbamoylated (DEC) enacyloxin IIa and IV (figure 4.1, E and F respectively), however structures of the final two remain unknown.[237–240]

4.1.1.1. Structure and Biosynthesis of Enacyloxin IIa

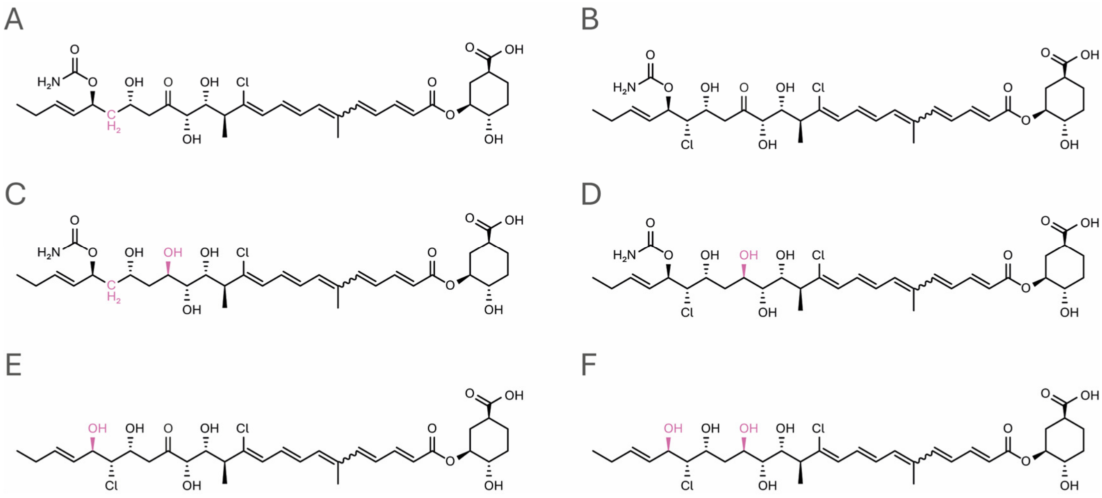
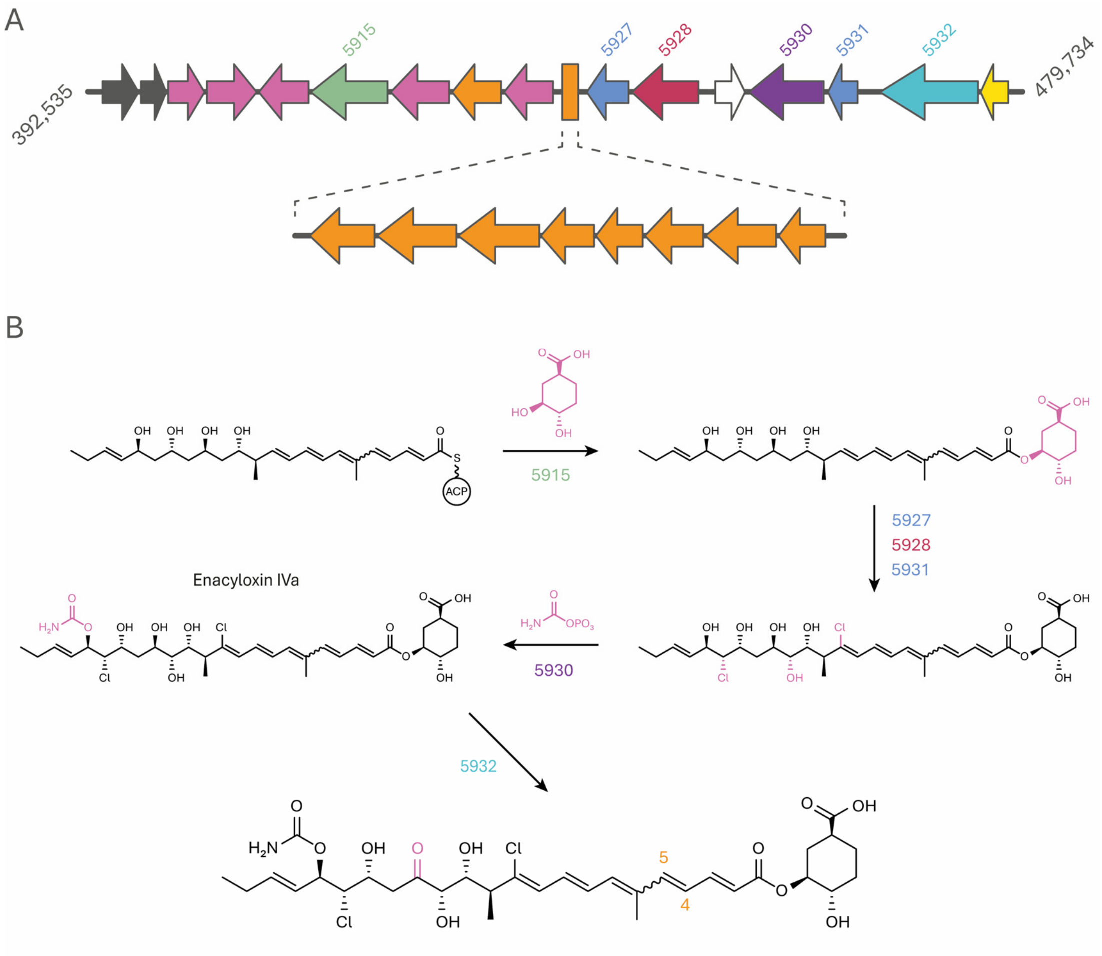


Figure 4.1 | Known enacyloxin structures. A | Enacyloxin Ia. B | Enacyloxin IIa. C | Enacyloxin IIIa. D | Enacyloxin IVa. E | DEC-Enacyloxin IIa. F | DEC-Enacyloxin IVa.

Enacyloxin IIa, as stated prior, consists of a dihydroxycyclohexanecarboxylic acid (DHCCA) *head* group, linked to a pentaene *tail* through an ester bond at the hydroxyl at the 3rd position (figure 4.1, B). The mature compound contains two chlorines, bonded to *tail* carbons 11 and 18, alongside a carbamoyl group at carbon 19.[241]



Enacyloxin IIa (Δ-4,5 = *E*)
iso-Enacyloxin IIa (Δ-4,5 = *Z*)

Figure 4.2 | Select stages of enacyloxin IIa biosynthesis. A | Enacyloxin IIa biosynthetic gene cluster, coloured by gene function. Zoom shows the major PKS codons. Grey, regulators; Pink, DHCCA synthases; Green, NRP synthetase; Orange, PKS; Blue, non-haeme iron-dependent oxygenase/chlorinase; Red, $FADH_2$-dependent chlorinase; White, hypothetical protein; Purple, carbamoyl transferase; Cyan, enacyloxin oxidase; Yellow, MATE-like transporter. Arrows drawn to scale and chromosomal location highlighted. Numbered codons refer to those involved in B. B | Post-PKS synthetic steps. Numbers refer to enzymes encoded by the codons highlighted in A. Results of each step are highlighted in pink.

Very little is known about the enacyloxin IIa biosynthetic pathway in *Frateuria.* Watanabe *et al.,* hypothesised that enacyloxin IVa (figure 4.1, D) is a metabolic precursor to enacyloxin IIa, which was confirmed by Izaki *et al.* who demonstrated the extracellular quinoprotein, enacyloxin oxidase, catalyses the oxidation of secreted enacyloxin IVa to enacyloxin IIa.[242,243]

In 2011, Mahenthiralingam *et al*. published their finding of a novel biosynthetic gene cluster in *Burkholderia ambifaria* responsible for the synthesis of enacyloxin IIa (figure 4.2).[241] Members of the *Burkholderia* genus demonstrate an array of genotypic and phenotypic diversity, with many producing biological active secondary metabolites.[244] They are often highly genetically tractable and contain genomes with up to 20% resulting from lateral gene transfer.[245] The *Burkholderia cepacia* complex (BCC) is a group of highly genetically and phenotypically similar bacteria, surrounding *B. cepacia*, with *B. ambifaria* a member. Antimicrobial screening highlighted *B. ambifaria* as a potent inhibitor of other Gram-positive Gram-negative bacteria. Random transposon mutants with reduced antimicrobial activity showed predominant transposon insertion into a large, cryptic 87.4 kb genomic island within the *B. ambifaria* genome. Insertion into secretion-related loci and, interestingly, quorum sensing-related loci also resulted in reduced antimicrobial activity. The genomic island (figure 4.2, A) has an overall GC content of 71.5% compared to the genome average of 66.5% suggesting a xenobiotic origin, which is highly likely in a BCC member. The island contains 24 coding sequences, with over a third having homology to polyketide synthase domains. The whole island is under the control of two orphaned LuxR-like regulators responsive towards N-acyl homoserine lactones, a common Gram-negative quorum sensing molecule.

The biosynthesis is complex, the polyketide-based pentaene *tail* is hypothesised to be assembled by a hybrid, type I, *cis*- and *trans*-acyltransferase containing polyketide synthase (PKS) before being released conjugated to a stand-alone acyl carrier protein. At this stage, the *tail* is not chlorinated, carbamoylated and each ketone reduced to their corresponding hydroxyl groups (figure 4.2, B). Post-PKS, the *tail* is first conjugated to DHCCA by a nonribosomal peptide synthetase (NRP, DHCCA synthesised from 3-dehydroshikimate by five enzymes encoded in the gene cluster). Addition of two chlorines and an additional hydroxyl group yields decarbamoyl-enacyloxin IVa, carbamoylation yields enacyloxin IVa, and finally extracellular oxidation yields enacyloxin IIa. Washes of *B. ambifaria* grown on solid agar yielded enacyloxin IIa as detected by ESI-MS, confirming production. Semipreparative-HPLC eluted two enacyloxin IIa stereoisomers, Δ-4,5-*E* and Δ-4,5-*Z* corresponding to enacyloxin IIa and the novel *iso*-enacyloxin IIa respectively (figure 4.2, B).[241]

In *B. ambifaria* transposon mutants with no antimicrobial activity also showed insertion into loci associated with secretion, specifically to recombination hotspot and valine-glycine-rich protein encoding genes. These have been shown to associate with the complex type-VI secretion systems found in Gram-negative bacteria, suggesting these as the method of enacyloxin IVa efflux, prior to oxidation by enacyloxin oxidase.[246]

4.1.1.2. Mode of Action

Enacyloxin IIa is a kirromycin-like elfamycin, binding to closed EF-Tu•GTP and preventing the conformational switch back to the open structure upon GTP hydrolysis. This behaviour prevents EF-Tu•GDP from dissociating from the ribosome, and stalling protein synthesis, and is the predominant enacyloxin IIa mode of action. In 1991 Watanabe *et al.* demonstrated enacyloxin IIa to be a bacteriostatic agent, acting to reduce the rate of radio-labelled amino acid incorporation into insoluble protein in a concentration-dependent manner.[247] The same group went on to hypothesise that enacyloxin IIa inhibits aa-tRNA binding to the ribosome by showing an enacyloxin IIa concentration-dependent reduction of radio-labelled Phe-tRNA^{Phe} bound to ribosomes.[248] Cetin *et al.* expanded on this, demonstrating kirromycin-like tangential effects of enacyloxin IIa on EF-Tu, namely: increased affinity towards GTP,

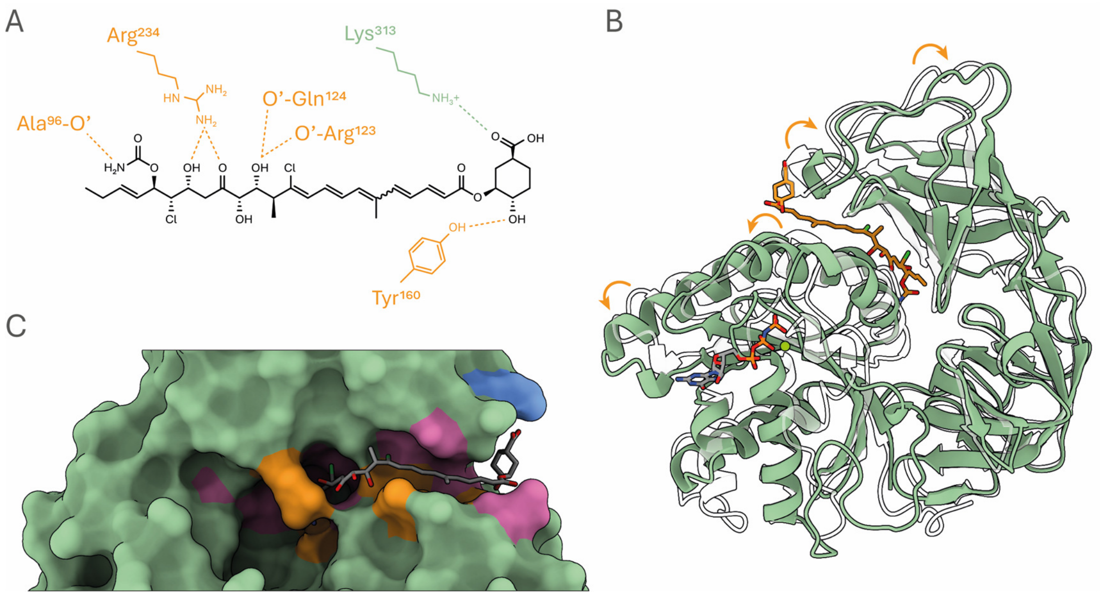


Figure 4.3 | Enacyloxin IIa binding to EF-Tu. A | EF-Tu•GTP•ENX hydrogen bonding network. Hydrogen bond forming residues coloured in orange, salt bridge forming residue in green. O' refers to backbone carbonyl carbon. Residues numbered as in *E. coli* EF-Tu. B | Cartoon crystal structure of *E. coli* EF-Tu•GDPNP•ENX (EF-TU green, ENX orange, PDB 2BVN). Note the shifting of structural elements in domains I and III compared to EF-Tu•GDPNP (white, PDB 1EFT). C | Residues contacting ENX (orange), hydrogen bond-forming (pink) and salt bridge forming (blue). ENX shown as sticks.

inhibition of EF-Ts mediated GDP release, and reduced EF-Tu-based protection of aa-tRNA from spontaneous diacylation.[249] The enacyloxin IIa binding site on EF-Tu was inferred to be similar to that of kirromycin by Zuurmond *et al.*, who showed that kirromycin resistance-conferring mutations also conferred enacyloxin IIa resistance.[250] The first crystal structure of EF-Tu•GDPNP•ENX was reported by Parmeggiani *et al.* in 2006, showing enacyloxin IIa binding to the domain I and III interface (figure 4.3, B), as predicted by Zuurmond *et al.*[251] In *E. coli* EF-Tu, enacyloxin IIa forms six hydrogen bonds (figure 4.3, A); the backbone carbonyl in Ala^{96} (helix *B*) binds to the carbamoyl amine; the carbonyls in Arg^{123} and Gly^{124} (helix *C*) bond with the enacyloxin IIa hydroxyl group at carbon 13; a terminal side chain nitrogen in Arg^{373} (strand *e3*) binds to both the carbonyl and hydroxyl groups at carbon 15 and 17 respectively; and the side chain hydroxyl group in Tyr^{160} (loop *D* – *g*) bonds to the hydroxyl group *ortho* to the carboxylic acid in DHCCA. An additional salt-bridge is formed between Lys^{313} (loop *a3* – *b3*) and the DHCCA carbonyl.[251] Alongside these, a further 16 non-bonding contacts (< 3.8 Å) are formed between enacyloxin IIa and residues in domains II and III (figure 4.3, C). Binding of enacyloxin IIa causes a widening of the domain I and III interface, and subtle shifts of helices *B* and *C*. As a result of this, the switch I region becomes totally disordered and absent from the crystal structure. The hydrogen bonding network formed between enacyloxin IIa and domains I and III stabilise the closed structure irrespective of the nucleotide bound, leading to closed EF-Tu•GDP•ENX remaining stuck on the ribosome. An EF-Tu•GDPNP•ENX•aa-tRNA quaternary structure is still able to form (albeit with slightly deformed aa-tRNA 3'-end geometry), presence of aa-tRNA induces structural stability within switch I, with helix *A''* reforming bound to the tRNA. The catalytic His^{84} in switch II remains oriented away from the γ-phosphate, in contrast to EF-Tu•GDPNP•KIR where His^{84} is pushed much closer to the γ-phosphate, explaining the reduced autohydrolysis of GTP in EF-Tu•GDPNP•ENX compared to EF-Tu•GDPNP•KIR. Both kirromycin and enacyloxin IIa inhibit the formation of EF-Tu•EF-Ts, competing for similar binding sites.

4.1.1.3. EF-Tu Resistance to Enacyloxin IIa

As mentioned earlier, resistance towards kirromycin also confers resistance towards enacyloxin IIa.[250,252] In screening *E. coli* kirromycin-resistant mutants against enacyloxin IIa, four were found to confer enacyloxin IIa resistance. Three of which (Q131K, G325D and A384T) localise to the enacyloxin IIa binding site (figure 4.4, numbered as in *A. baumannii*), and the fourth (Q329H) is directly involved in aa-tRNA binding.[34] Q131K, G325D and A384T all reduce the binding affinities of enacyloxin IIa to EF-Tu in both the GDP (40 ± 20 µM, 20 ± 4 µM and 6 ± 1 µM respectively *vs* < 1 µM for the wild-type) and GTP-bound EF-Tu (3 ± 1 µM, 4 ± 1 µM and < 1 µM respectively *vs* < 1 µM for the wild-type) as measured by nucleotide-dissociation. Q329H instead reduces the affinity of aa-tRNA towards EF-Tu•GDP•ENX (80 ± 30 µM *vs* 2 ± 1 µM for the wild-type), allowing dissociation of EF-Tu from the tRNA after ribosomal delivery and hydrolysis. For the purposes of this work only Q131K, G325D and A384T are of interest, being directly related to enacyloxin IIa binding. Interestingly *Frateuria* W-315 does not produce an enacyloxin IIa resistant EF-Tu but does show phenotypic resistance. This implies an alternate resistance mechanism, hypothesised to be active efflux, but the deterministic cause remains unknown.[253]

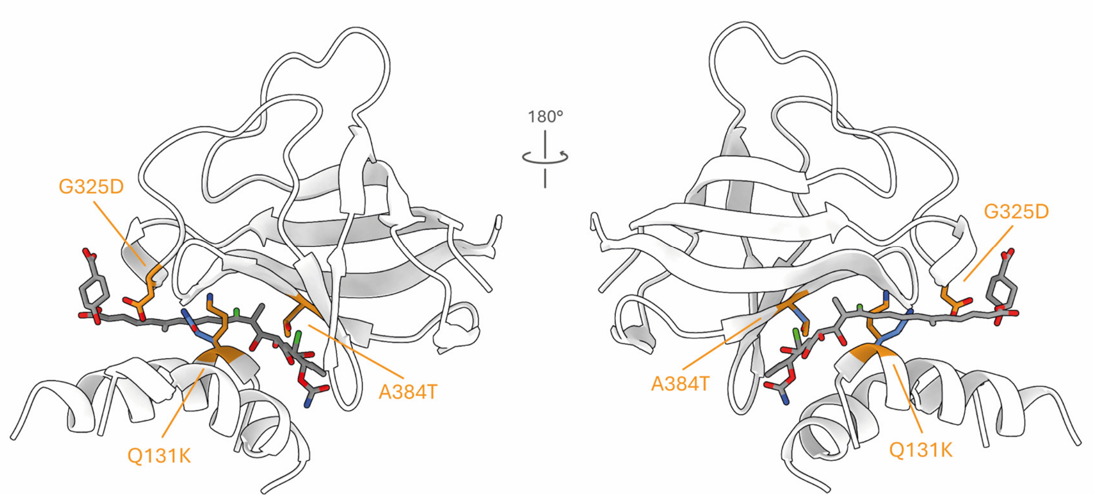


Figure 4.4 | Enacyloxin IIa conferring resistance mutations. Wild-type residues are coloured in blue, and mutant residues in orange. Only domain III and helices *B* and *C* of domain I are shown for clarity. Enacyloxin IIa represented as a stick structure. Residue numbers correspond to the heterologously expressed *A. baumannii* EF-Tu.

4.2. Aims

The aim of the work presented in this chapter was to apply carbene footprinting to *A. baumannii* EF-Tu to study the highly dynamic nature of its conformational change, especially when bound to enacyloxin IIa. Similar to the work described in chapter 3, several enacyloxin IIa-bound EF-Tu crystal structures have been solved, yet they are unable to capture subtle conformational dynamics. Due to its solution-phase nature, carbene footprinting is able to elucidate these dynamics. Several enacyloxin IIa resistant mutants were produced to confirm enacyloxin IIa binding being responsible for the transition of EF-Tu from its *open* to *closed* conformation, described by their carbene footprinting profiles.

4.3. Results and Discussion

4.3.1. Heterologous Expression of *A. baumannii* EF-Tus

Wild-type *A. baumannii* EF-Tu (referred to from this point as just *EF-Tu*), was heterologously expressed and purified from *E. coli* as described in section 2.1.1 – 2.1.3. Enacyloxin IIa resistant EF-Tu mutants (EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T}) were produced by targeted mutagenesis as described in section 2.1.4.1, and heterologously expressed and purified as above. Wild-type EF-Tu purified from a 1 L *E. coli* culture yielded 2.10 mg mL^{-1} (equivalent to 48.14 μM), found predominantly in the soluble fraction of lysed *E. coli* (figure 4.5, A, orange arrow), and presence was confirmed by anti-His western blotting. Ni-NTA affinity purification yielded predominately 6xHis-tagged EF-Tu (figure 4.5, B), with a theoretical molecular weight of 46,114 Da. Cleavage of the N-terminal 6xHis tag with TEV protease (figure 4.5, C) yielded EF-Tu (pink arrow, theoretical 43,522 Da), visualised alongside free TEV (green arrow). Analytical SEC (appendix II) indicated EF-Tu elution within collected fractions 19, 20 and 21, which was confirmed by SDS-PAGE (figure 4.5, C). Fraction 21 was concentrated (section 2.1.3.3), diluted, aliquoted and stored at -80°C until needed.

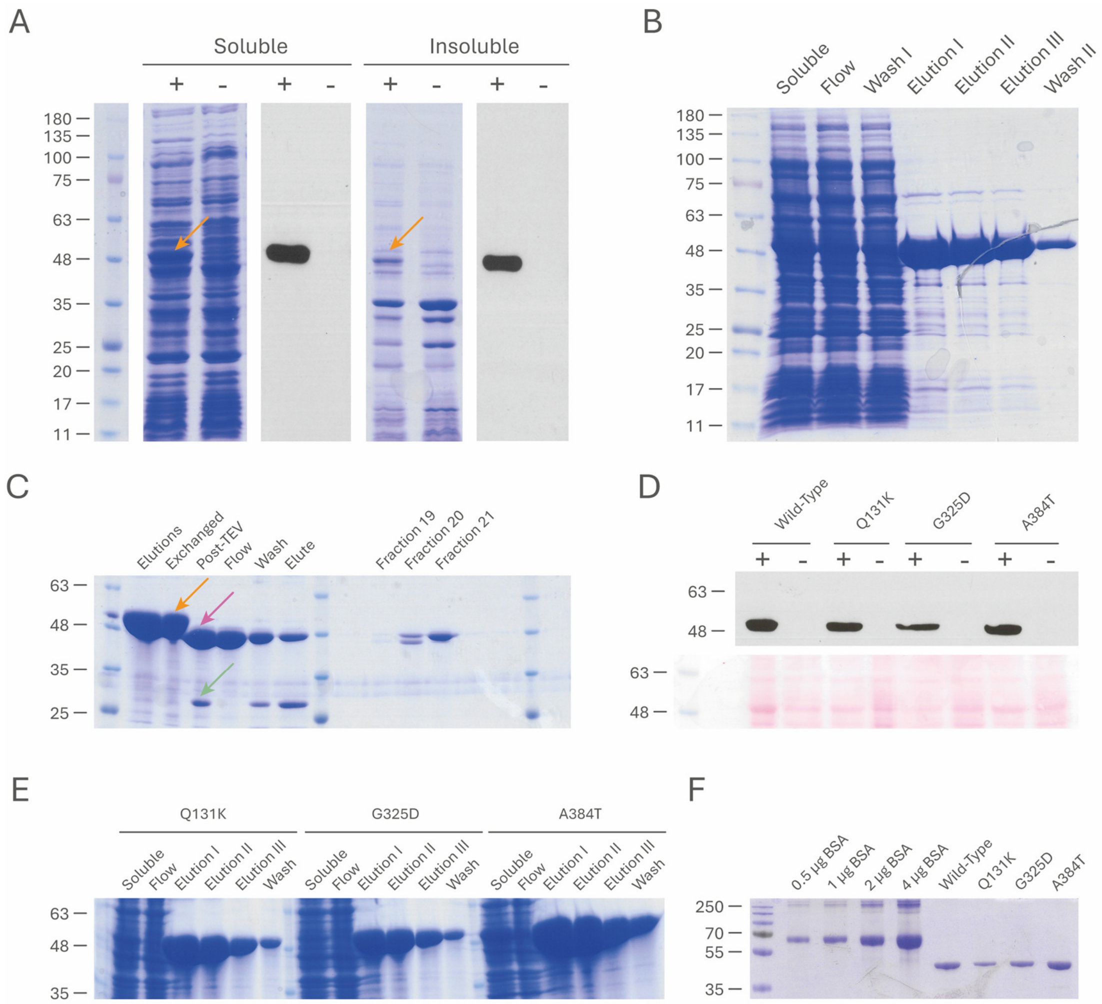


Figure 4.5 | Heterologous *A. baumannii* EF-Tu expression and purification. All molecular weight markers show kDa. A | Wild-type EF-Tu (orange arrow) localised predominantly in the soluble fraction and identity confirmed by western blotting. B | Ni-NTA affinity purification yielded highly pure EF-Tu, with some losses remaining bound to the resin. C | TEV cleavage of the N-terminal 6xHis•EF-Tu (orange arrow) followed by Ni-NTA to remove TEV (green arrow) yielded clean EF-Tu (pink arrow), confirmed by SEC (fractions 19 – 21 containing EF-Tu). D | Expression of EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} mutants. Soluble fraction visualised with Ponceau S (bottom) and identify confirmed by anti-His western blot (top). E | Ni-NTA purification of EF-Tu mutants yielded very high protein concentrations. F | Mini-SDS-PAGE of each EF-Tu variant post dilution and aliquoting compared to BSA standards.

EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} (figure 4.5, D), were purified (figure 4.5, E), cleaved and concentrated yielding 9.40, 3.94 and 14.22 mg mL^{-1} (215.86, 90.29 and 326.39 µM respectively) as calculated by a BCA protein assay. Samples of each EF-Tu were diluted to 1 µM and subject to SDS-PAGE alongside prepared BSA standards to confirm correct concentrations obtained by BCA (figure 4.5, F). Confirmation of successful mutagenesis was performed at the DNA level by sequencing the *tufA* gene of each mutant EF-Tu (section 2.1.1.2). To corroborate this, mass spectrometric confirmation was performed at the protein level (figure 4.6).

Substitution of glutamine for lysine, as in EF-Tu^{Q131K}, made direct mass spectrometric analysis difficult due to the nominally isobaric nature of the residues. Conveniently, this substitution introduced a novel trypsin cleavage site, yielding the peptide Val^{132} – Lys^{143} unique to EF-Tu^{Q131K}. Extracted ion chromatograms corresponding to this peptide (figure 4.6, A, m/z of 681.41 for $[M+2H]^{2+}$) confirmed its presence only in EF-Tu^{Q131K}. The identity of this peptide was confirmed by targeted MS^2 (figure 4.6, B) showing excellent coverage in both b and y ions. For EF-Tu^{G325D}, the substitution of glycine for aspartic acid imparts an easily detectable mass shift of +58.04 Da. At the MS^1 level, the doubly-charged Gly^{316} containing peptide, Phe^{313} – Arg^{327}, had a m/z of 849.92 compared to 878.93 for the Asp^{316} substituted peptide ($\Delta = +29.01$). Targeted MS^2 of each peptide (figure 4.6, C) confirmed their identities, with excellent coverage from both b and y ions. Examination of EF-Tu^{G325D} Phe^{313} – Arg^{327} MS^2 spectrum (figure 4.6, C, left) showed comparable detected y ions ($y_3 - y_{12}$) having an average m/z shift of +58.01 ± 0.04 when compared to the wild-type (figure 4.6, C, right). The $b_2 - b_{12}$ ions are identical between the two spectra, except for b_{13} being present in the EF-Tu^{G325D} peptide. The presence of aspartic acid in an arginine-containing peptide demonstrates increased cleavage at its C-terminus, which may explain the abundance of the b_{13} ion in EF-Tu^{G325D}.[254–256] Combined, these indicated the presence of Asp^{325} in EF-Tu^{G325D}, confirming successful mutagenesis. Unfortunately, despite multiple digestion attempts, peptides containing Ala^{384} were often not detected. Chymotryptic digestion of EF-Tu^{A384T} yielded the C-terminal peptide

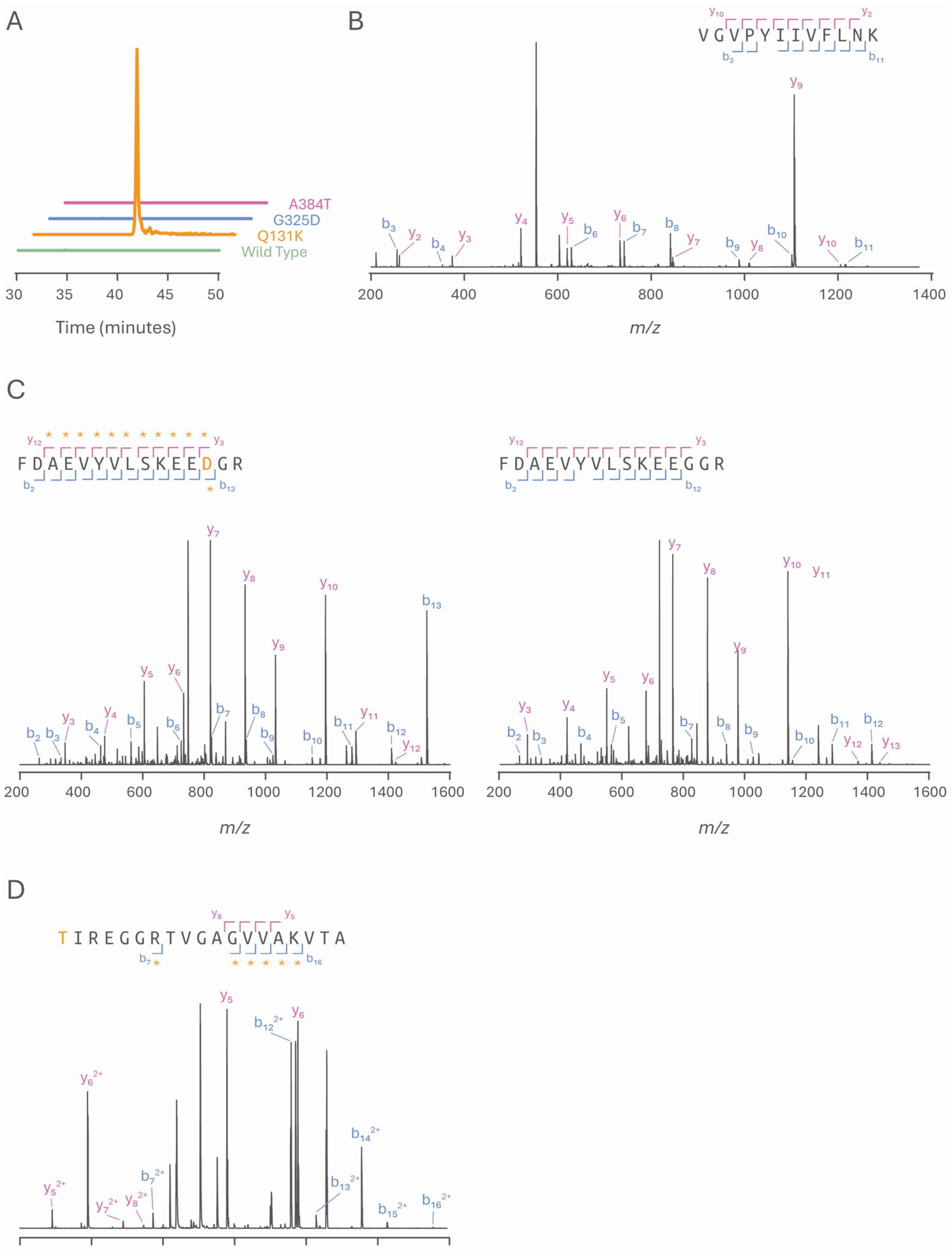


Figure 4.6 | Confirmation of mutagenesis. A | XIC of novel tryptic peptide Val^{132} – Lys^{143} in EF-Tu^{Q131K}. B | MS^2 spectrum of Val^{132} – Lys^{143}. C | MS^2 spectra of Phe^{313} – Arg^{327} in EF-Tu^{G325D} (left) and wild-type EF-Tu (right). D | MS^2 spectra of Thr^{384} – Ala^{402} in EF-Tu^{A384T}. Fragment ions asterisked carry a *m/z* attributed to the amino acid substitution. Observed and expected *m/z* values of each fragment ion can be found in appendix III, i – iv.

Thr^{84} – Ala^{402}, however digestion of wild-type EF-Tu did not produce the corresponding Ala^{384} – Ala^{402}. Although direct comparison of these peptides was therefore not possible, MS^2 of $[M+4H]^{4+}$ Thr^{84} – Ala^{402} (figure 4.6, D) identified all $[M+2H]^{2+}$ *b* ions carrying an average *m/z* shift of + 15.03 ± 0.19, when compared to *in silico* theoretical Ala^{384} – Ala^{402} *b* ions (indicated with the orange asterisk). This is in agreement to the calculated mass difference between threonine and alanine (15.01 Da). Presence of b_7 restricts this *m/z* shift to Ala^{384}, which is corroborated by all detected $[M+2H]^{2+}$ *y* ions matching their corresponding wild-type theoretical *m/z* (average *m/z* Δ of 0.05 ± 0.03). This paired with DNA sequencing indicated successful substitution of Ala^{384} with Thr^{384}.

4.3.2. Native Mass Spectrometry of EF-Tu with Enacyloxin IIa

Samples of each EF-Tu were prepared for native-MS as described in section 2.2.1.4. For acquisition of EF-Tu•GDP, protein aliquots were diluted post-buffer exchange (section 2.2.1.2) to final concentrations of 10 µM. Spectra were acquired using nESI emitter tips, as prepared in section 2.2.1.1, on a Synapt G1 HDMS High-Definition mass spectrometer. Native spectra of each EF-Tu can be seen in figure 4.7, A. As shown in chapter 3, nESI-MS native EF-Tu spectra typically show three charge states: $[M + 11H]^{11+}$, $[M + 12H]^{12+}$, and $[M + 13H]^{13+}$, with $[M + 12H]^{12+}$ being the principal charge state. Without the addition of supplementary nucleotide, purified EF-Tu is almost entirely observed as EF-Tu•GDP (figure 4.7, A, filled circles). Very minor peaks corresponding to *apo*-EF-Tu can be seen within each charge state (figure 4.7, A, empty circles). Deconvolution of *apo*-EF-Tu charge states yields molecular weights of 43,561 ± 5.71, 43,616 ± 11.45, 43,681 ± 6.97 and 43,564 ± 6.62 Da for wild-type, $EF\text{-}Tu^{Q131K}$, $EF\text{-}Tu^{G325D}$ and $EF\text{-}Tu^{A384T}$ respectively.

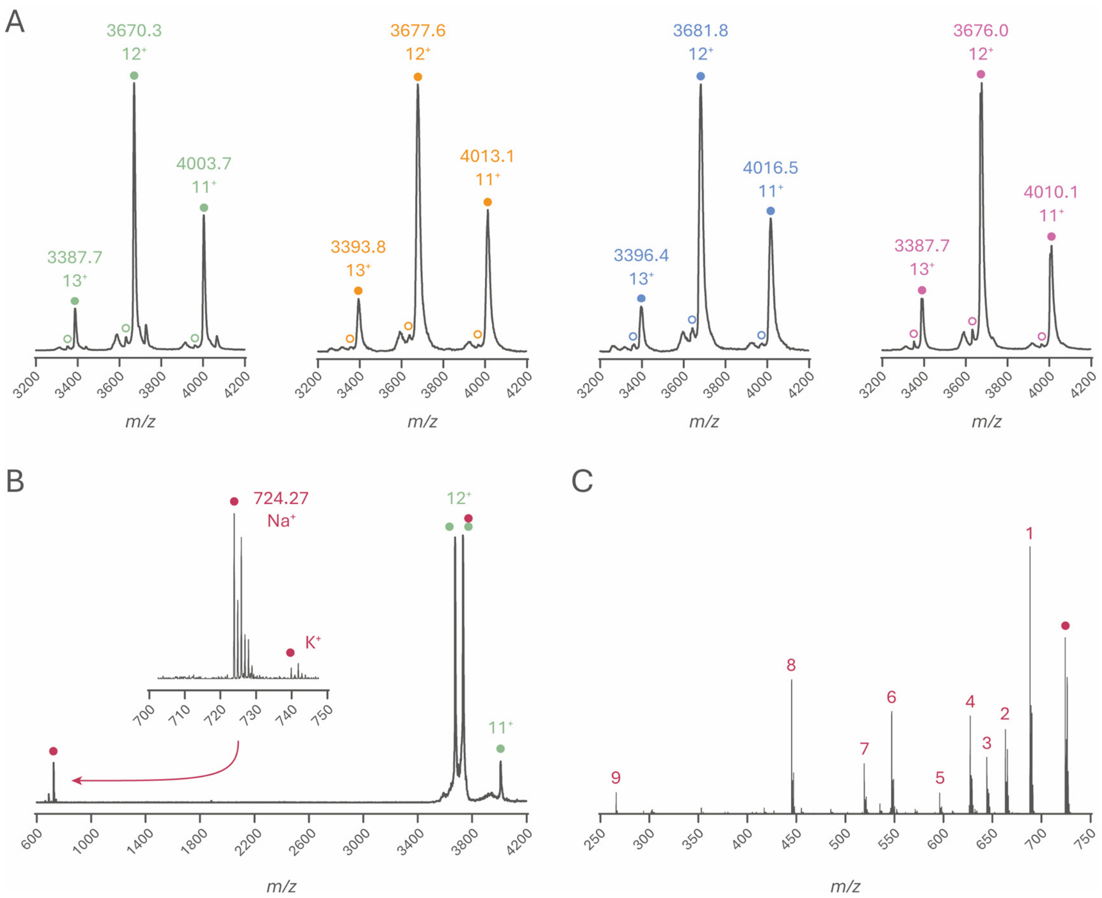


Figure 4.7 | A | Native mass spectra of wild-type, EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} (green, orange, blue and pink respectively). Charge states and *m/z* highlighted. Closed circles indicate EF-Tu•GDP and open *apo*-EF-Tu (Mg^{2+} bound). B | Quadrupole isolated and collisionally activated [M + 12H]$^{12+}$ EF-Tu•GDPNP•ENX (red circle) showing predominant ejection of enacyloxin IIa as a neutral ion, with minor liberated sodiated and potassiated enacyloxin IIa ions. C | CID fragmentation pattern of sodiated enacyloxin IIa (red circle), fragment numbers correspond to those in table 4.1.

Accounting for bound Mg^{2+}, the remaining discrepancy between observed and theoretical masses can be explained by cationic adduction, with adducting peaks visible upon close inspection of the individual peaks. Calculated mass shifts between the *apo*-protein and the dominant EF-Tu•GDP peaks indicated an average deviance of 453 ± 8.32 Da, corresponding to addition of GDP (443.2 Da) to the Mg^{2+} containing *apo*-protein. Incubation of wild-type EF-Tu with GDPNP and enacyloxin IIa prior to nESI-MS yields exclusively EF-Tu•GDPNP•ENX. Quadrupole isolation of [M + 12H]$^{12+}$ followed by collisional activation induces dissociation of enacyloxin IIa (figure 4.7, B, 42 V). Enacyloxin IIa is ejected predominantly as a neutral ion, indicated by the emergence of EF-Tu•GDPNP [M + 12H]$^{12+}$. Some cationic ejection is observed, indicated by EF-Tu•GDPNP [M + 11H]$^{11+}$emergence alongside detection of sodiated and potassiated enacyloxin IIa ions, however no protonated enacyloxin IIa ions are observed (figure 4.7, B, inset). Quadrupole isolation and collisional activation of the dominant sodiated enacyloxin IIa ion yields a reproducible fragmentation pattern (figure 4.7, C), with tentative fragment assignments given in table 4.1.

Table 4.1 | Tentative fragment ions produced from CID activation of enacyloxin IIa, corresponding to the spectrum in figure 4.7, C. All fragments, bar 9, are present as sodiated ions. Fragments all originate from sequential losses from the enacyloxin IIa *tail*.

Fragment	*m/z*	Loss from Parent Ion (*m/z*)	Loss
1	688.30	-35.99	HCl
2	663.27	-61.02	CH_3NO_2
3	644.32	-79.97	HCl + C_3H_8
4	627.30	-96.99	HCl + CH_3NO_2
5	596.22	-128.07	?
6	547.21	-177.08	HCl + $C_7H_{11}NO_2$
7	519.22	-205.08	$C_8H_{12}ClNO_3$
8	445.18	-279.12	$C_{11}H_{18}ClNO_5$
9	266.13	-435.11	$C_{19}H_{27}Cl_2NO_6$

Quadrupole isolated [M + 12H]$^{12+}$ EF-Tu•GDPNP•ENX ion was collisionally activated with a range (20 – 60 V) of collision voltages (figure 4.8). Measurable enacyloxin IIa ejection was observed starting at 30 V (figure 4.8, A) and the minima of approximately 5 % remaining EF-Tu•GDPNP•ENX reached at 55 V. Fitting of a symmetrical sinusoidal non-linear regression again allowed for the extraction of CE_{50}, calculated to be 42.2 ± 0.2 V. Representative spectra for the given collision voltages are shown in figure 4.8, B, demonstrating the preferential ejection of neutral enacyloxin IIa. Interestingly, charged enacyloxin IIa ejection appeared to occur preferentially at the lower collision voltages (< 40 V) before neutral ejection becomes dominant (figure 4.8, C). As precursor EF-Tu•GDPNP•ENX ions are typically adducted by sodium ions, gentle collisional activation would remove these salts in a process deemed *collisional cleaning*.[257] It is possible that enacyloxin IIa ejected under these conditions could sequester these salt ions, leading to observation of sodiated enacyloxin IIa.

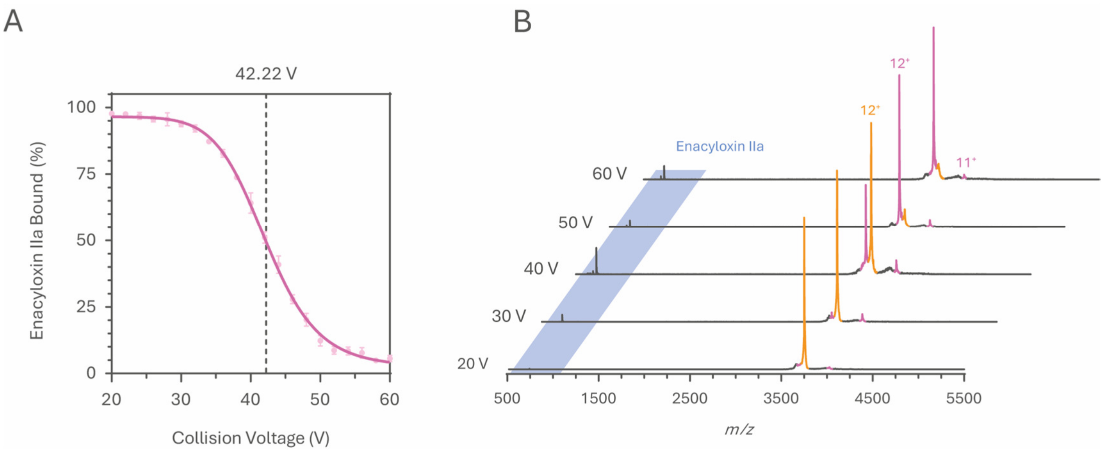


Figure 4.8 | CID of EF-Tu•GDP•ENX. A | CID of the quadrupole isolated EF-Tu•GDP•ENX [M + 12H]$^{12+}$ ion, with voltages of 20 – 60 V. Data is an average of 3 repeats. Error bars represent one standard deviation. B | Representative normalised spectra from A, showing liberation of enacyloxin IIa (blue shading), emergence of EF-Tu•GDPNP [M + 12H]$^{12+}$ (pink), minor emergence of charge reduced EF-Tu•GDPNP [M + 11H]$^{11+}$, and disappearance of EF-Tu•GDP•ENX [M + 12H]$^{12+}$ (orange) as the collision voltage increases.

4.3.2.1. Binding Affinities of Enacyloxin IIa to EF-Tu•GDP

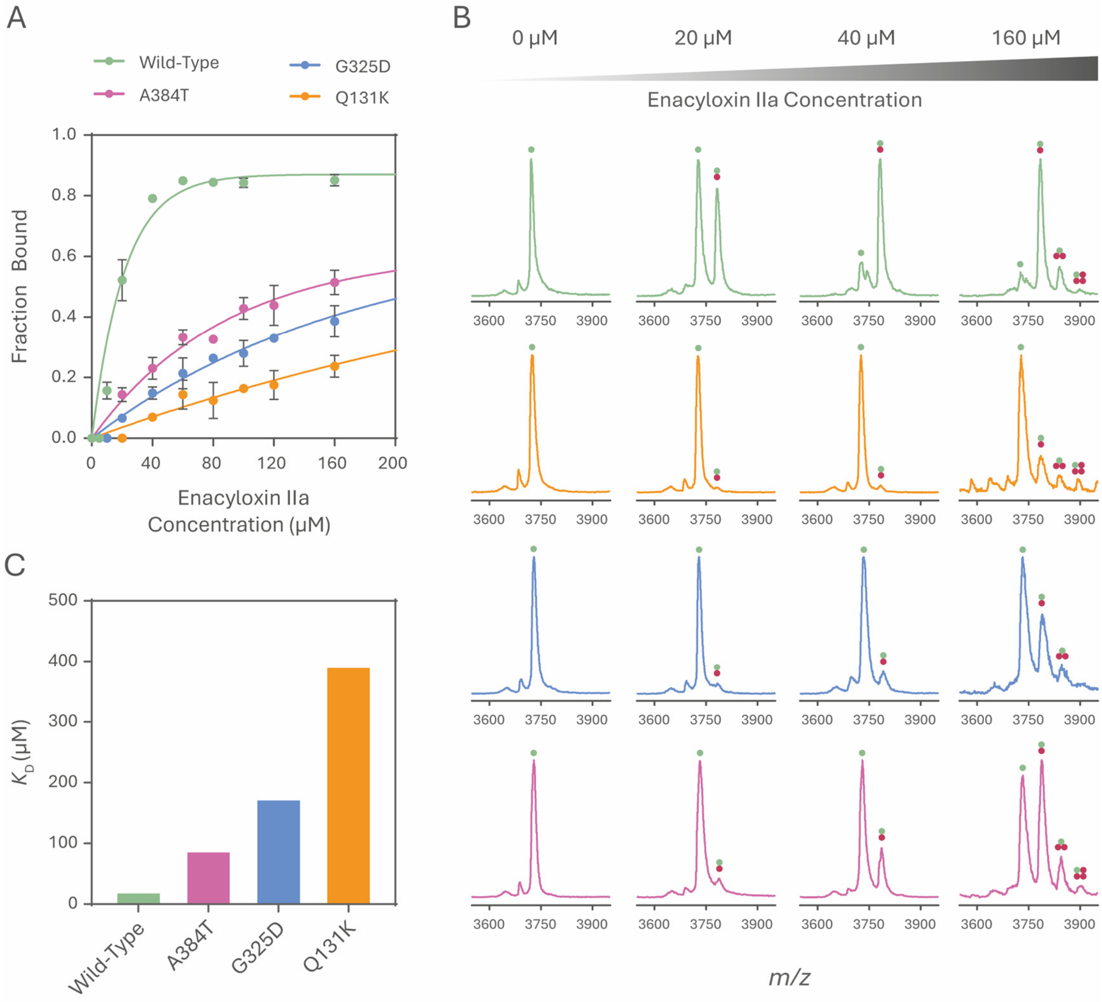


Figure 4.9 | EF-Tu•GDP•ENX binding affinities. A | Binding curves for EF-Tu•GDP•ENX, coloured as indicated. Data points show average values of three replicates and error bars, one standard deviation. B | Representative normalised spectra from A, coloured the same. EF-Tu•GDP is represented by the green circle, and enacyloxin IIa the red circle. C | Extracted K_D values from A. Note, due to the robust regression method used in A, confidence intervals for each K_D are not calculated.

As the resistance mechanism of EF-Tu towards enacyloxin IIa is induced by reductions in binding, to confirm the resistant nature of EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} their binding affinities towards enacyloxin IIa were investigated. This was achieved through native-MS, as described in section 2.2.1.8. Initial attempts to characterise the affinity of EF-Tu•GDPNP towards enacyloxin IIa were unsuccessful as a result of the nanomolar binding affinity of enacyloxin IIa to EF-Tu.

At concentrations of EF-Tu (5 µM) required to produce spectra of acceptable quality, incubation with equimolar enacyloxin IIa led to observation solely of the EF-Tu•GDPNP•ENX complex even in EF-Tu^{Q131K}, the mutant reported to have the highest binding affinity.[252] In light of this, characterisation of EF-Tu•GDP•ENX binding affinities were investigated (figure 4.9, A). Although wild-type EF-Tu•GDP binds enacyloxin IIa strongly (K_D reportedly below detection in nucleotide dissociation experiments), EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} exhibit a broader range of binding affinities (up to 60-fold *vs* wild-type compared to 4-fold for EF-Tu•GDPNP).[250]

By calculating the proportion of detected EF-Tu•GDP•ENX with respect to total EF-Tu and fitting non-linear regression (one-phase association) curves, K_D values could be obtained (figure 4.9, A). Here K_D was calculated as the enacyloxin IIa concentration where half-maximal binding is observed. K_D values of 21.58, 393.50, 174.3 and 88.9 µM were observed for wild-type EF-Tu, EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} respectively (figure 4.9, C). These are at odds with values previously reported by solution phase nucleotide dissociation (< 1, 40 ± 20, 20 ± 4, and 6 ± 1 for wild-type EF-Tu, EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} respectively) by approximately a factor of 10.[250] However, the qualitative pattern of resistance remains constant, with EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} showing decreasing levels of resistance. With weak binding ligands that do not reach binding saturation, curve fitting is inaccurate often leading to an over-estimation of K_D values.[258] Representative native-MS spectra for each EF-Tu can be seen in figure 4.9, D, for 0, 20, 40 and 60 µM of enacyloxin IIa. In each EF-Tu, non-specific binding of enacyloxin IIa was seen at the highest enacyloxin IIa concentration, with binding of up to an additional three enacyloxin IIa observed. In EF-Tu^{Q131K}, the most resistant mutant, comparable levels of specific and non-specific binding were observed, demonstrating the much-reduced affinity towards enacyloxin IIa within the binding site.

4.3.3. Carbene Footprinting of EF-Tu with Enacyloxin IIa

4.3.3.1. *Optimisation of Proteolytic Digestion.*

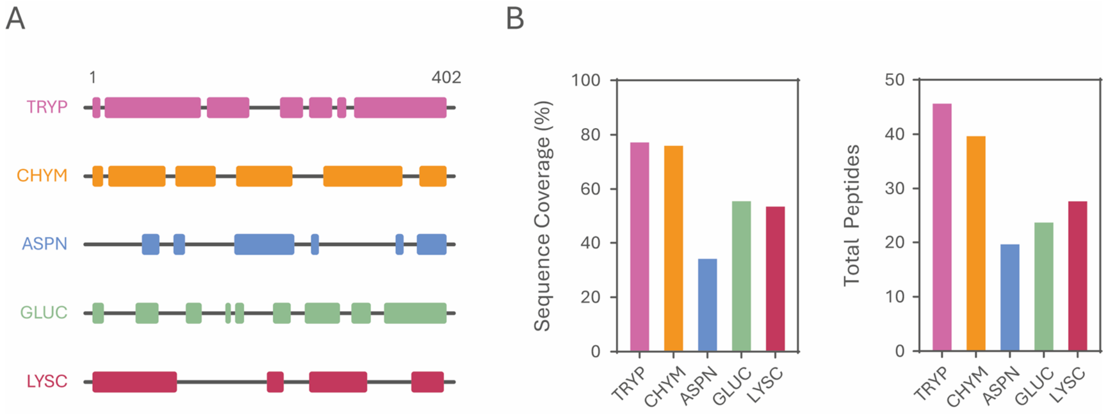


Figure 4.10 | HPLC-MS^2 measured, proteolytic EF-Tu sequence coverage. A | Graphical representation of protease sequence coverage, bars indicate regions containing detected peptides. B | Sequence coverage and total peptide number measured for each protease, coloured as in A.

In silico proteolytic digestions of *A. baumannii* EF-Tu with five proteases (trypsin, chymotrypsin, AspN, GluC and LysC) were performed using SearchGUI and PeptideShaker with search settings described in section 2.3.3.1. Theoretical sequence coverage of 90.6, 100, 85.3, 79.4 and 54.7 % was reported for trypsin, chymotrypsin, AspN, GluC and LysC respectively, suggesting chymotrypsin produces the greatest possible sequence coverage. *In vitro* proteolytic digests (in-gel, section 2.3.1.3) were performed on EF-Tu with each enzyme, and the obtained sequence coverage compared (figure 4.10). Samples were analysed using DDA HPLC-MS^2 (section 2.3.1.5) and analysed as described section 2.3.3.1. As expected, *in vitro* digestion sequence coverage did not equate to the maximum possible coverage obtained by *in silico* theoretical digestions. Interestingly, these results showed trypsin yielded the greatest observed sequence coverage (figure 4.10, A, B), covering 77.9 % of the total sequence, across 46 individual peptides (these peptides being reported as confident by PeptideShaker). Chymotrypsin performance was disappointing when compared to the theoretical potential, producing 76.6 % coverage over 40 peptides. Chymotrypsin did provide additional coverage between Gly^{179} – Arg^{213} (helices *E* and *F*), but this region, linking domains I and II, is of little interest with regards to the enacyloxin IIa binding site. Neither AspN (34.8 % coverage, 20 peptides), GluC (56.2 % coverage, 24 peptides) or LysC (54.2 % coverage, 28 peptides) provided any additional benefit over trypsin. For this reason, single protease digestion using trypsin was selected for use in carbene footprinting experiments. Note, when tryptic digestions are performed in-solution (section 2.3.1.4), sequence coverage tends to be slightly higher (*e.g.* observed 87.5 % coverage of EF-Tu•GDP•GE2270A in chapter 3).

4.3.3.2. Optimisation of Carbene Footprinting Parameters

Prior to performing differential carbene footprinting on EF-Tu (sections 2.3.2.1 – 2.3.2.2) method optimisation was performed. *Apo*-EF-Tu was subject to carbene footprinting, as described in section 2.3.2.1, whilst altering several parameters. Both the concentration of TDBA (present as the sodium salt), alongside laser irradiation time were varied to obtain the greatest labelling, whilst minimising the likelihood of double TDBA insertion within individual peptides. Footprinted EF-Tu was purified through SDS-PAGE (section 2.3.1.2), digested in-gel (section 2.3.1.3) and subject to HPLC-MS (section 2.3.1.4). F_{Mod} was quantified as described in section 2.3.3.2.

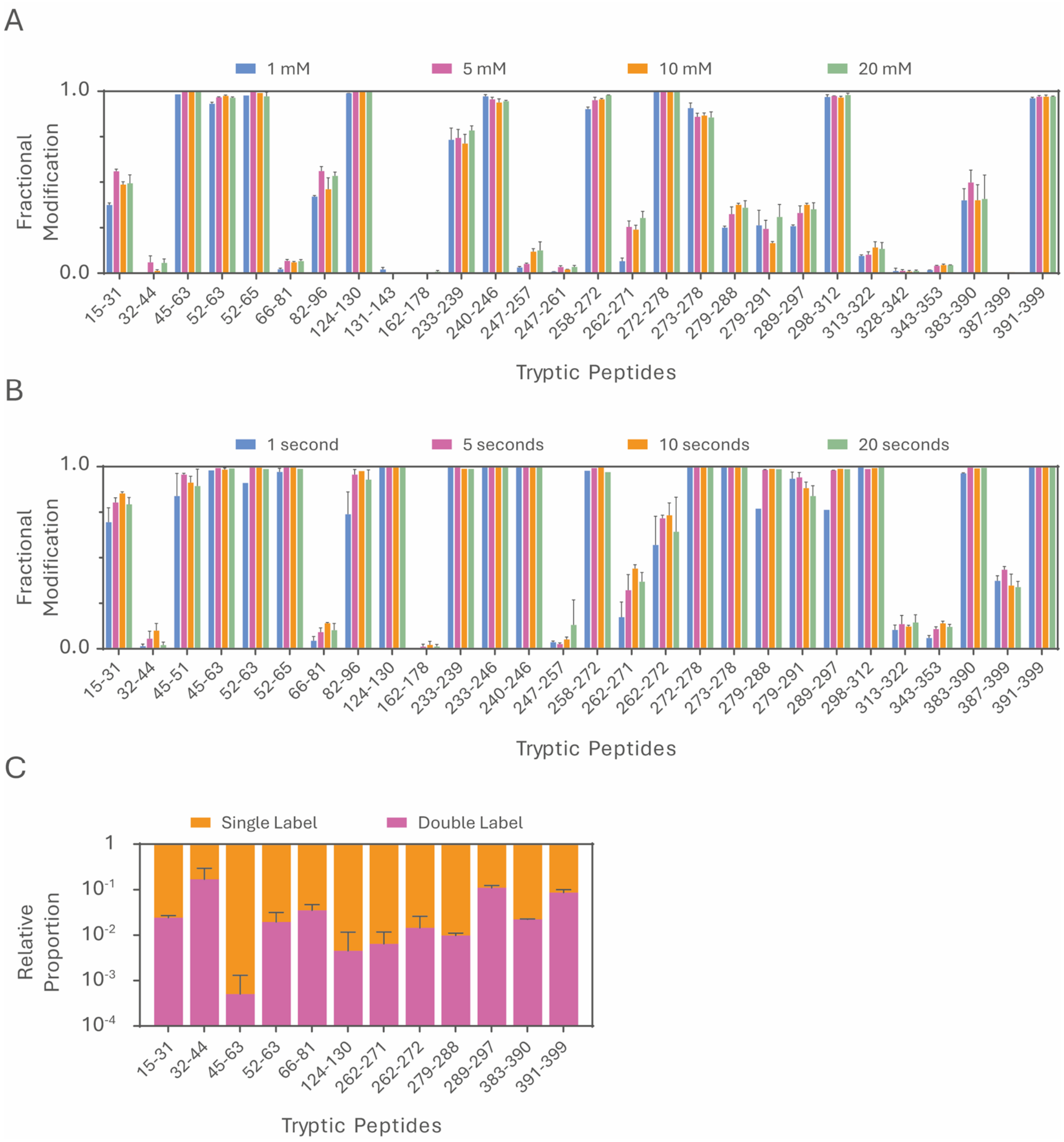


Figure 4.11 | Optimising footprinting conditions. A | Varying TDBA concentration shows a mild concentration-dependent effect on F_{Mod}. B | Varying laser irradiation time shows a more pronounced concentration-dependent effect on F_{Mod}, especially on peptides with lower overall F_{Mod}. C | Peptides from B irradiated for 20 seconds demonstrating detectable double labelling. Proportion of double labelled XIC area with respect to total labelled XIC. Note, base-10 logarithmic *y*-axis. All data is an average of 3 repeats, error bars represent one standard deviation.

A range of TDBA concentrations (1 – 20 mM) were tested (figure 4.11, A) which demonstrated a mild concentration-dependent effect on F_{Mod}, particularly on peptides with low overall F_{Mod}. Eight individual peptides were found to be heavily labelled, irrespective of TDBA concentration. On balance, no peptides demonstrated significantly higher F_{Mod} between 10 and 20 mM TDBA, so a concentration of 10 mM was used going forward. This is in line with the concentrations used by Manzi *et al.* to successfully label both lysozyme and ubiquitin specific protease 5.[195] Moving forward, the effect of total irradiation time (1 – 20 s) on F_{Mod} was investigated, using 10 mM TDBA (figure 4.11, B).

Altering irradiation time produced a more noticeable, mostly proportional effect on F_{Mod}, again on those peptides reporting an $F_{Mod} < 1.0$. Interestingly, almost all peptides (with $F_{Mod} < 1.0$) showed a reduction in F_{Mod} when subject to 20 seconds of irradiation compared to 10 seconds. It was hypothesised that this may be due to an increase of double labelling. Greater double labelling reduces the proportion of singly labelled peptides, and therefore leads to a reduction in F_{Mod}. To investigate this, effected peptides were analysed to determine the amount of double labelling (figure 4.11, C). Overall, double labelled peptides accounted for between 0.005 and 16.9 % of their respective total (single and double) labelling XIC peak area, no triply labelled peptides were identified. Thr^{32} – Lys^{44} showed the greatest relative drop in F_{Mod}, and the corresponding greatest extend of double labelling (16.9 %). This peptide showed labelling (section 4.3.3.4) across three clearly defined residues, which may explain its propensity towards double labelling, despite an overall low F_{Mod}. Two further peptides, Gly^{289} – Arg^{291} and Thr^{391} – Lys^{399} showed close to 10 % double labelling (10.9 and 8.5 % respectively), but these are heavily labelled peptides showing F_{Mod} values of > 0.95 irrespective of irradiation time. To minimise the effects of double labelling, a middling irradiation time of 15 seconds was chosen going forward. Note, extreme discrepancies between certain equivalent peptide F_{Mod} in figures 4.11, A and B were found to be caused by prior use of trifluoroacetic acid (TFA) in the HPLC mobile phase. Leading to a curious suppression of labelled peptides, resulting in overall reduced F_{Mod}. Thus, TFA was not used as a mobile phase modifier for any labelling experiments. Instead, 0.1 % formic acid was added (chapter 2, section 2.3.1.5).

4.3.3.3. Carbene Footprinting of wild-type EF-Tu with Enacyloxin IIa

EF-Tu samples were produced and footprinted as described in section 2.3.2.1 prior to SDS-PAGE purification (section 2.3.1.2), in-gel digestion (section 2.3.1.3) and subject to HPLC-MS analysis (section 2.3.1.4). F_{Mod} was quantified as described in section 2.3.3.2. Initial footprinting experiments were performed with a final EF-Tu concentration of 5 µM, limited by available enacyloxin IIa (and therefore methanol) concentration. All enacyloxin IIa samples were kind gifts from Professor Greg Challis (University of Warwick). Tryptic digestion yielded 28 peptides, representing 56.5% of the total EF-Tu sequence (figure 4.12 and 4.13). Differential footprinting of *apo*-EF-Tu and EF-Tu•GDP (figure 4.13, A) showed no significant differences in F_{Mod} between any peptides. Interestingly, there was no significant masking imparted by GDPNP, despite relatively good coverage of the binding site (figure 4.14, B). A non-significant reduction in F_{Mod} (-0.073) was seen in Asn^{15} – Lys^{31}, a large peptide encompassing the G1 *P-loop* motif, associated with nucleotide binding. Thr^{32} – Lys^{44} showed a minor increase in F_{Mod} (+0.029), but this peptide (helix *A*) is proximal to the highly flexible switch I region (specifically helix *A''*). No structures of *apo*-EF-Tu have been solved, and as discussed earlier, crystal structures can show an artificial structure associated with crystal packing, making interpretation of these highly flexible regions difficult. Molecular dynamics simulations of human K-RAS, which shares high structural similarity to domain I of EF-Tu demonstrate the flexibility of switch I, and the dampening effect of GTP on this flexibility, which may lead to these (insignificant) unmasking effects proximal to switch I.[259]

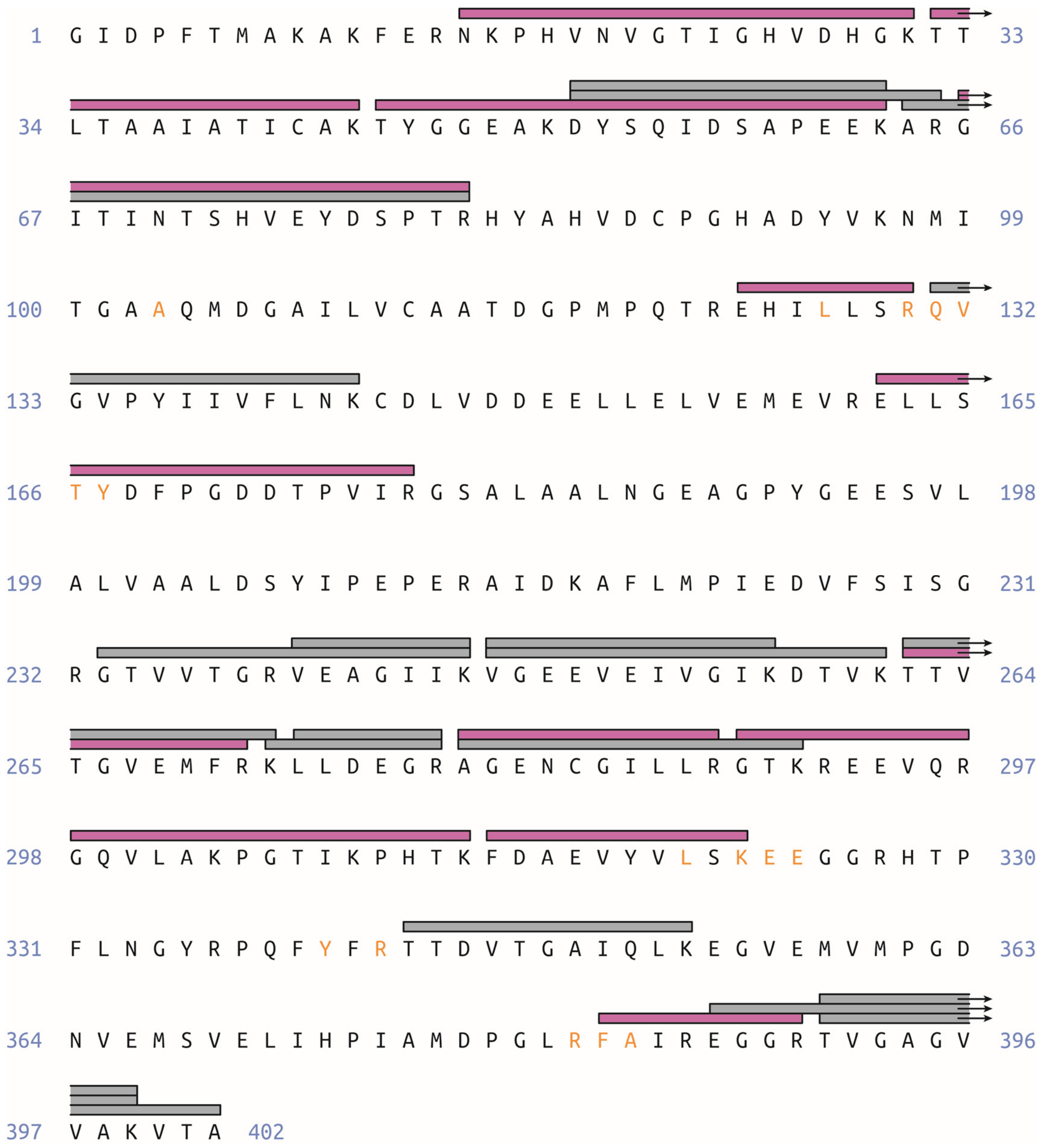


Figure 4.12 | Tryptic peptide coverage of *A. baumannii* EF-Tu post-carbene footprinting. Peptides coloured grey showed no significant reductions in F_{Mod} in EF-Tu•GDPNP•ENX compared to EF-Tu•GDPNP, and those coloured pink did show significant reductions. Residues coloured orange are those involved in non-bonding contacts and hydrogen bonding with enacyloxin IIa.

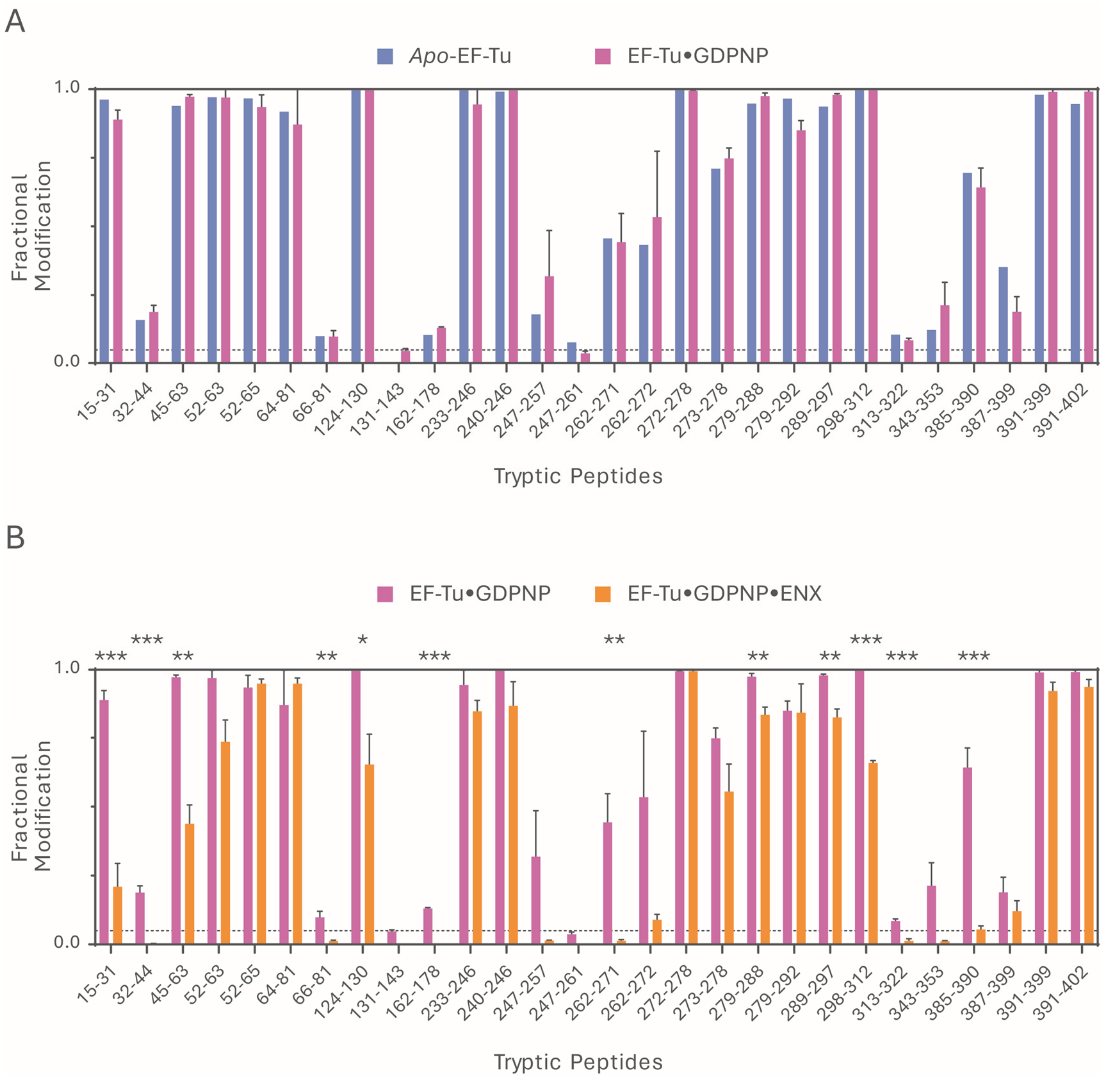


Figure 4.13 | F_{Mod} histograms for *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX. A | *apo*-EF-Tu (blue) *vs*. EF-Tu•GDPNP (pink) B | EF-Tu•GDPNP (pink) *vs* EF-Tu•GDPNP•ENX (orange). Data is an average of four replicates, error bars represent one standard deviation. Stars indicate significant reductions in F_{Mod}, in EF-Tu•GDPNP•ENX compared to EF-Tu•GDPNP. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Dotted line indicates an F_{Mod} of 0.05.

This suggests that GDPNP binding to EF-Tu may influence switch I dynamics, which could in turn lead to greater TDBA accessibility to helix *A* underneath. However, if this is the case, the effect must be minimal to only induce such a small, insignificant increase in F_{Mod}. The dynamics of EF-Tu in solution were recently elucidated in detail by Johansen *et al.* who, using fluorescence resonance energy transfer (FRET) and x-ray crystallography, demonstrated the striking conformational flexibility of EF-Tu. *E. coli* EF-Tu•GDPNP to adopt, and crystalise in, the classical GDP *open* conformation.[26] FRET showed EF-Tu•GDPNP to adopt an average conformation in solution between the *open* and *closed* crystal structures, compared to that of EF-Tu•GDP which adopted a conformation in agreement with the *open* crystal structures.

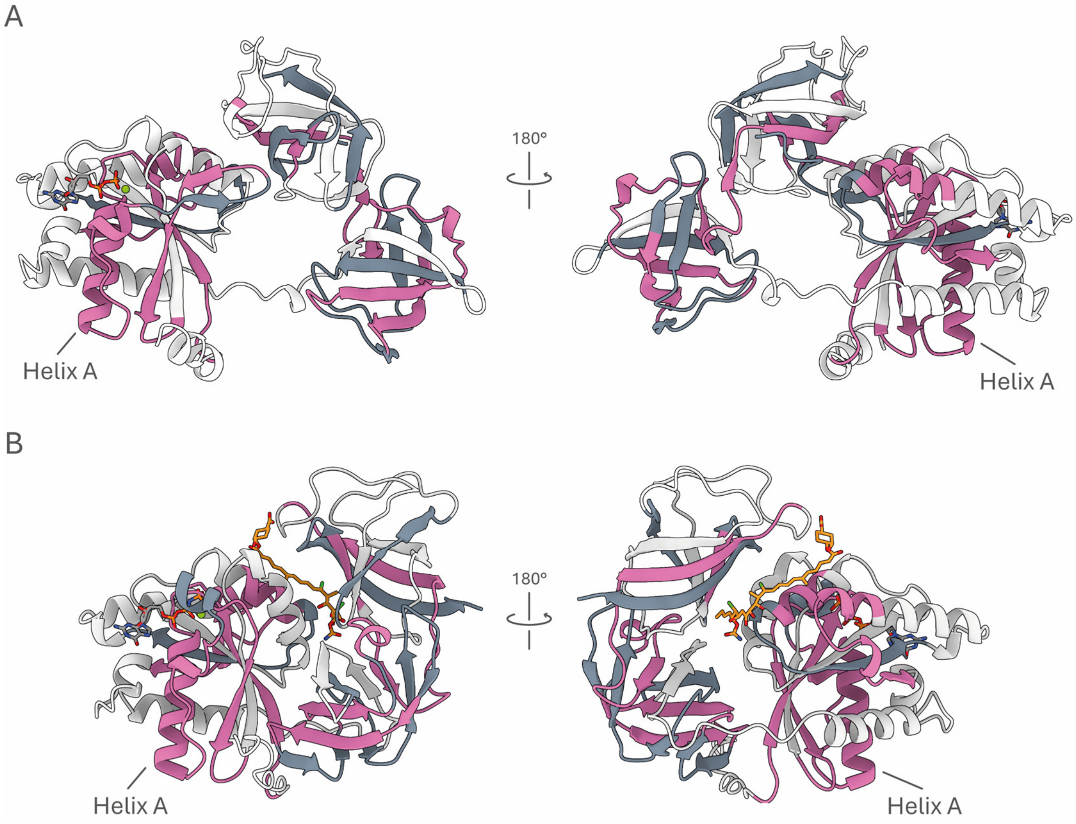


Figure 4.14 | EF-Tu•GDPNP•ENX F_{Mod} data mapped to an *A. baumannii* EF-Tu homology model. A | F_{Mod} data mapped to *open* EF-Tu. Peptides with significantly reduced F_{Mod} (pink) are found in all three EF-Tu domains. Peptides in grey showed no significant differences when compared to EF-Tu•GDP. Peptides in white were not detected by HPLC-MS. B | F_{Mod} data mapped to *closed* EF-Tu, enacyloxin IIa represented as sticks (orange). Masking can be seen in peptides covering the enacyloxin IIa binding site. GDP (top) and GDPNP (bottom) represented as sticks.

In contrast, comparison of EF-Tu•GDPNP to EF-Tu•GDPNP•ENX revealed striking differences in F_{Mod}. Overall, 12 EF-Tu•GDPNP•ENX peptides showed significant reductions in F_{Mod} compared to EF-Tu•GDPNP (figure 4.12; 4.13, B and 4.14). These peptides not only covered the enacyloxin IIa binding site but were also distributed across the accessible surface of EF-Tu, suggesting global structural perturbations associated with enacyloxin IIa binding. Pleasingly, peptides containing 62.5% of residues either bonding to or contacting EF-Tu were detected, and all showed significant reductions in F_{Mod} (figure 4.12).

The first three peptides significantly masked, Asn^{15} – Lys^{31}, Thr^{32} – Lys^{44} and Thr^{45} – Lys^{63} (F_{Mod} reductions of -0.680, -0.186 and -0.533 respectively, all $p < 0.001$) sequentially cover strand *a*, helices *A* and *A"* and strand *b'*. As discussed previously, these encompass the flexible switch I region which lays in front of helix *A*. Although switch I is flexible, in the *open* conformation the motif is structured, consisting of the small helix *A"* and antiparallel sheet *b'* and *b*. In the *closed* conformation, this region is often found to be wholly disordered and often absent in crystal structures due to its high variability.[56] It is likely that in the more *closed* EF-Tu•GDPNP•ENX structure, the resulting flexibility in this region induced a masking effect on Asn^{15} – Lys^{31} and Thr^{32} – Lys^{44}. Notably, although Thr^{45} – Lys^{63} showed a significant reduction in F_{Mod}, two peptides, Asp^{52} – Lys^{63} and Asp^{52} – Arg^{65}, showed no significant reduction in F_{Mod}. MS^2 analysis of labelled Thr^{45} – Lys^{63} placed TDBA insertion within Lys^{51} – Asp^{57} (figure 4.16, A). It is worth noting that Asp^{52} – Lys^{63} showed a relatively high reduction in F_{Mod} (-0.233), albeit insignificant ($p = 0.08$). Structural flexibility is likely to be the cause of this lack of significant masking. The additional Ala^{64} and Arg^{65} in Asp^{52} – Arg^{65}, which are found C-terminal to strand *b'* (*open*) and in the highly exposed C-terminus of refolded helix *A"* (closed) are likely to be heavily labelled irrespective of EF-Tu conformation, leading to no significant difference in F_{Mod} (figure 4.16, A). Solvent accessible surface area (SASA) analysis showed *closed* Asp^{52} – Lys^{63} to have a larger average SASA than Thr^{45} – Lys^{63} which may explain its higher labelling (1185.97 and 499.31 Å^2 respectively, SASA probe here being TDBA, not water, table 4.2) The peptide Gly^{66} – Arg^{81} did show a minor, but significant reduction in F_{Mod} (-0.086, $p < 0.001$). In *open* EF-Tu this forms the accessible strand *b,* and more buried strand *c.* MS^2 analysis placed TDBA labelling on Ile^{67} – Thr^{71}, which exclusively forms strand *b* and the connecting loop to strand *c* (figure 4.16, B). In *closed* EF-Tu, strand *b* is disordered and falls behind the newly structured helix *A''* greatly reducing its accessibility towards TDBA (figure 4.15, B). The missed cleavage peptide Ala^{64} – Arg^{81} contains the N-terminal aforementioned Ala^{64} and Arg^{65} and again, did not show a significant reduction in F_{Mod}.

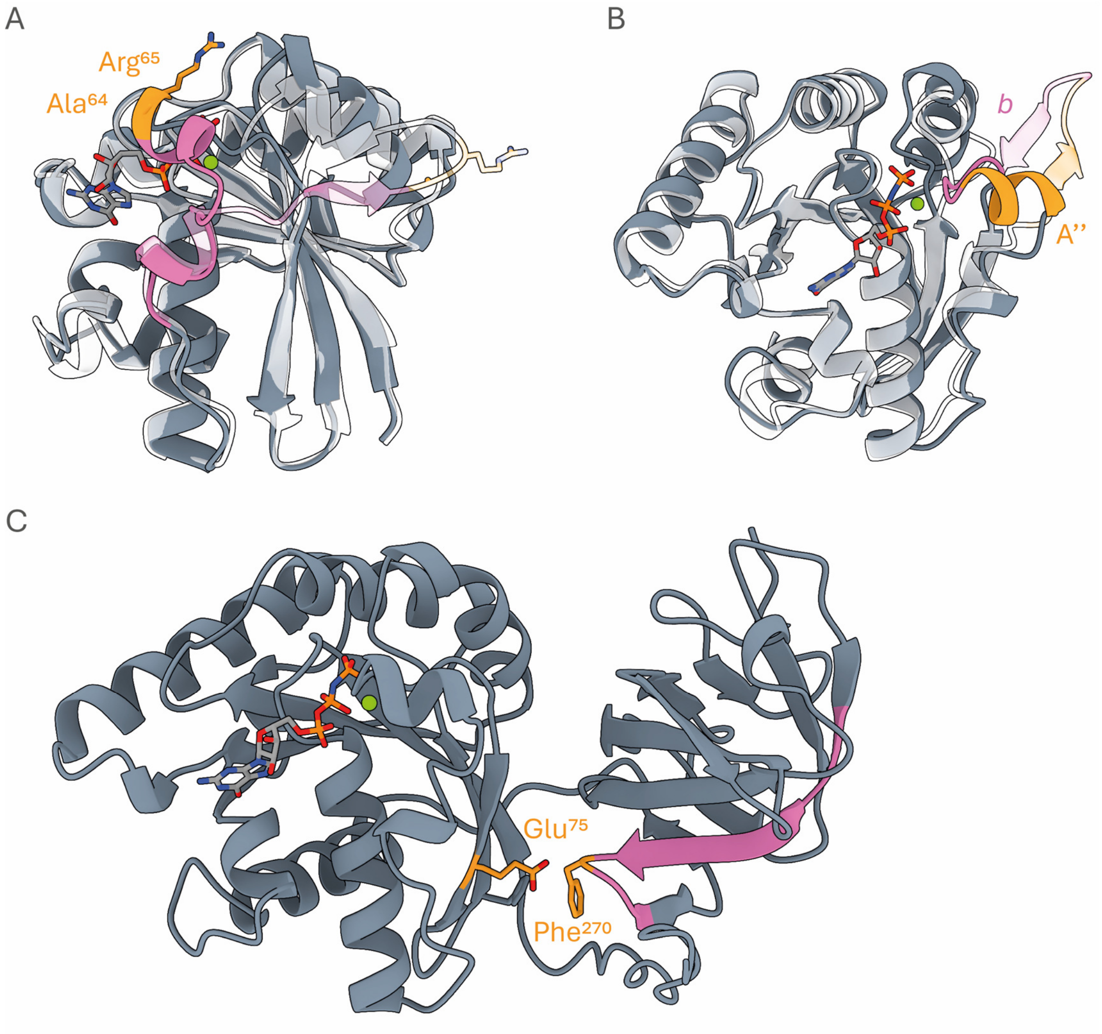


Figure 4.15 | Zoom of some key regions. A | Peptide Asp^{52} – Arg^{65} (pink) with prominent Ala^{64} and Arg^{65} (orange). B | Refolding of strand *b'* to helix *A''* (orange) which results in it laying in front of the now disordered strand *b* (pink). In A and B only domain I is shown, and the N-terminus removed. *Closed* EF-Tu is shown in grey and *open* EF-Tu in white. C | Peptide Thr^{262} – Arg^{271}(pink) containing Phe^{270} which bonds to Glu^{71} in domain I. Domain III omitted for clarity. GDPNP shown as a stick representation in all structures, and Mg^{2+} cofactor as a green sphere.

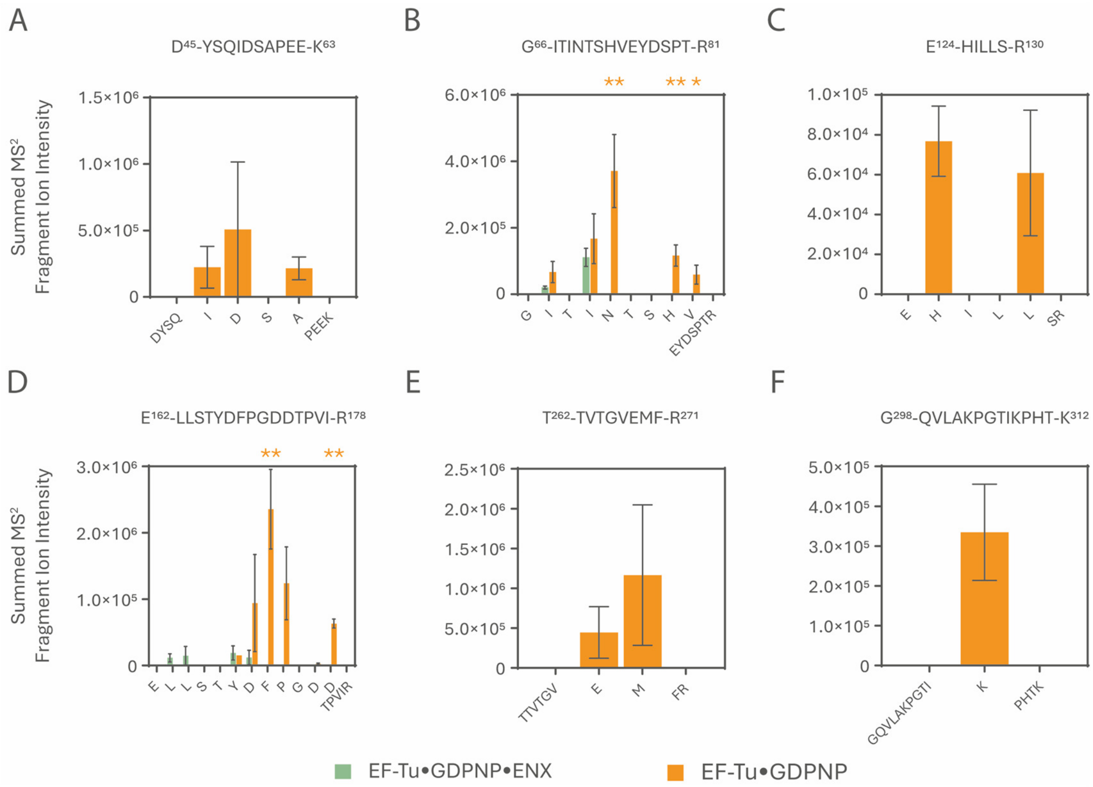


Figure 4.16 | MS^2 analyses of peptides giving successful fragmentation. A| Asp^{45} – Lys^{63}. B | Gly^{66} – Arg^{81}. C | Glu^{124} – Lys^{130}. D | Glu^{162} – Arg^{178}. E| Thr^{262} – Arg^{271}. E| Gly^{298} – Lys^{312}. Data is an average at least 3 replicates, error bars represent standard deviation. Stars indicate significant reductions in EF-Tu•GDPNP•ENX F_{Mod} (if detected) compared to EF-Tu•GDPNP. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

The peptide Glu^{124} – Arg^{130} showed a significant reduction in F_{Mod} (-0.345, $p = 0.003$) in EF-Tu•GDPNP•ENX, containing two contacting residues (figure 4.12). In *E. coli* EF-Tu, the backbone carbonyl in Arg^{123} forms a hydrogen bond with enacyloxin IIa, the equivalent residue in *A. baumannii* is Arg^{130} (figure 4.3, A). Glu^{124} – Arg^{130} also encompasses Leu^{127}, homologous to the contacting residue Leu^{120} in *E. coli*. MS^2 analysis showed TDBA insertion on His^{125} and Leu^{127} (figure 4.16, C). The following peptide, Gln^{131} – Lys^{143} was both very weakly labelled (EF-Tu•GDPNP F_{Mod} of 0.043), and of very low intensity, and consequently was below the threshold for statistical comparison. The significantly masked Glu^{162} – Arg^{178} (Δ F_{Mod} of -0.129, $p < 0.001$) contains the residues Thr^{166} and Tyr^{167}, homologous to the *E. coli* contacting and hydrogen bonding residues Gln^{159} and Tyr^{160} respectively. MS^2 analysis showed labelling predominantly on the adjacent Asp^{168} and Phe^{169}, which were significantly masked when bound to enacyloxin IIa (figure 4.16, D).

The next significantly masked peptide, Thr^{262} – Arg^{271} (ΔF_{Mod} of -0.429, $p < 0.001$), is located in domain II and is brought into close contact with domain I in the *closed* conformation, such that the Phe^{261} amide nitrogen forms a hydrogen bond with Glu^{68} in *T. thermophilus* EF-Tu (homologous to Phe^{270} and Glu^{75} in *A. baumannii.*[251] MS^2 analysis placed TDBA insertion between Glu^{268} – Met^{269}, a region that in *closed* EF-Tu becomes not only buried next to domain I, but may also be inadvertently masked by the free, disordered N-terminus (figure 4.15, C, figure 4.16, E, figure 4.17, table 4.2). The missed cleavage Thr^{262} – Lys^{272} showed highly variable F_{Mod} in EF-Tu•GDPNP, and despite a relatively high reduction in F_{Mod} (-0.445) was not found to be significantly masked in EF-Tu•GDPNP•ENX. In EF-Tu•GDPNP•ENX Thr^{262} – Lys^{272} did demonstrate a higher F_{Mod} compared to Thr^{262} – Arg^{271} (0.090 and 0.014 respectively), likely due to the additional basic C-terminal lysine. Within this these experiments the missed cleavage peptide and the sequential Lys^{272} – Arg^{278} were of too low intensity to perform MS^2, however successful fragmentation of the sequential Lys^{272} – Arg^{278} in chapter 3 (figure 3.9, B) demonstrated the majority of labelling to be on the N-terminal Lys^{272}, Leu^{273}, and Leu^{274}.

Table 4.2 | Calculated average SASA from crystal structures of *open* and *closed* EF-Tu (PDB 1EFC and 1EFT respectively) and its respective change from *open* to *closed*.

Tryptic Peptide	Open Sasa (Å^2)	Closed SASA (Å^2)	ΔSASA (Å^2)	% Change in SASA
15-31	437.00	315.67	-121.33	-27.76
32-44	235.50	214.61	-20.89	-8.87
45-63	1379.74	1200.53	-179.21	-12.99
56-63	632.19	499.31	-132.88	-21.02
52-65	1121.76	1185.97	+64.21	+5.72
64-81	499.89	680.48	+180.58	+36.12
66-81	380.04	310.23	-69.81	-18.37
124-130	106.50	98.03	-8.47	-7.95
131-143	147.63	149.97	+2.33	+1.58
162-178	851.84	842.77	-9.06	-1.06
233-246	317.08	185.07	-132.02	-41.63
240-246	149.93	179.68	+29.75	+19.84
247-257	378.87	392.51	+13.64	+3.60
247-261	720.29	865.76	+145.47	+20.20
262-269	147.76	27.74	-120.03	-81.23
262-272	913.29	93.68	-819.61	-89.74
272-278	549.32	208.11	-341.21	-62.11
273-278	403.39	203.11	-200.28	-49.65
279-288	331.58	239.74	-91.84	-27.70
279-292	674.94	718.51	+43.57	+6.45
289-297	724.45	776.72	+52.27	+7.22
298-312	470.01	588.60	+118.59	+25.23
313-322	189.38	355.92	+166.54	+87.94
343-353	287.86	255.73	-32.13	-11.16
385-390	475.85	268.77	-207.07	-43.52
387-399	509.73	527.71	+17.98	+3.53
391-399	118.24	258.93	+140.69	+118.98
391-402	363.67	507.87	+144.20	+39.65

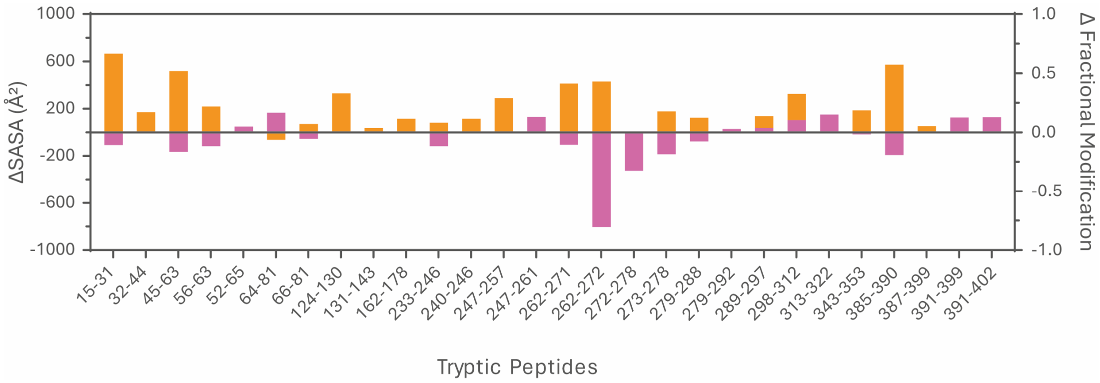


Figure 4.17 | Change in SASA (pink) and change in F_{Mod} (orange) for each tryptic peptide. The heavily reduced peptides between Thr^{261} and Arg^{278} corresponds to the region in domain II which rotates and become buried next to domain I. Peptides with reduced SASA in *closed* EF-Tu typically exhibit a greater reduction in F_{Mod}.

Ala^{279} – Arg^{288} was significantly masked in EF-Tu•GDPNP•ENX (Δ F_{Mod} of 0.140, $p < 0.001$). The N-terminal region of this peptide forms part of the interface between domains I and II in the *closed* conformation, and as a result showed a corresponding drop in solvent-accessible surface area (SASA, Δ SASA of -91.84 Å^2) when compared to *open* EF-Tu (table 4.2, figure 4.17). The missed cleavage Ala^{279} – Lys^{292} showed no significant masking in EF-Tu•GDPNP•ENX, the additional C-terminal residues occupy an exposed position in domain II irrespective of EF-Tu conformation and are likely to remain heavily labelled.

Interestingly, Gly^{289} – Arg^{297} showed significant reductions in F_{Mod} (-0.153, $p < 0.001$) in EF-Tu•GDPNP•ENX. Whilst this peptide does undergo a significant translocation to the *closed* conformation, overall, the SASA increases by 52.27 Å^2 (albeit only a 7.22 % increase over *open* EF-Tu, table 4.2, figure 4.17). Unfortunately, this peptide, despite showing heavy labelling, was of too low intensity for successful MS^2 fragmentation. Similarly, Gly^{298} – Lys^{312} also showed significant reductions in F_{Mod} (-0.341, $p < 0.001$) whilst becoming more exposed in EF-Tu•GDPNP•ENX (Δ SASA of +118.59 Å^2, or +25.23 %). Although of low intensity, tentative MS^2 analysis placed TDBA labelling to Lys^{303}, which did individually show a reduction in SASA within the overall peptide indicating a partial burial of the peptide in the *closed* conformation (Lys^{303} Δ SASA of -15.20 Å^2, or -44.06 %, figure 4.16, F).

Phe^{313} – Lys^{322} showed a minor, but significant reduction in F_{Mod} (-0.072, $p < 0.001$) in EF-Tu•GDPNP•ENX. The C-terminal Lys^{322} is homologous to *E. coli* Lys^{313}, which forms a salt bridge to the enacyloxin IIa DHHCA carboxylic acid group, and Leu^{320} is homologous to *E. coli* Leu^{311}, which forms a non-bonding contact with enacyloxin IIa. Unfortunately, this peptide was too weak to perform MS^2 analysis, however due to the presence of bonding and contacting residues, it is highly likely that masking is due to binding of enacyloxin IIa, as this peptide forms a major part of the domain III binding site. Finally, the small peptide Phe^{385} – Arg^{390} showed a large significant reduction in F_{Mod} (-0.588, $p < 0.001$) and contains two N-terminal residues (Phe^{385} and Ala^{386}), homologous to those in *E. coli* contacting enacyloxin IIa (Phe^{374} and Ala^{384}). MS^2 analysis placed labelling predominantly on Ile^{387} and Arg^{388}, proximal to those contacting residues and well within the enacyloxin IIa binding site. Overall, these data show that carbene footprinting has captured both the localised binding site of enacyloxin IIa on EF-Tu and, importantly, the large conformational change to the *closed* conformation associated with enacyloxin IIa binding. This is the first time carbene footprinting has been used to investigate a conformational change of this magnitude.

4.3.3.4. Analysis of an Isomerically Labelled Peptide

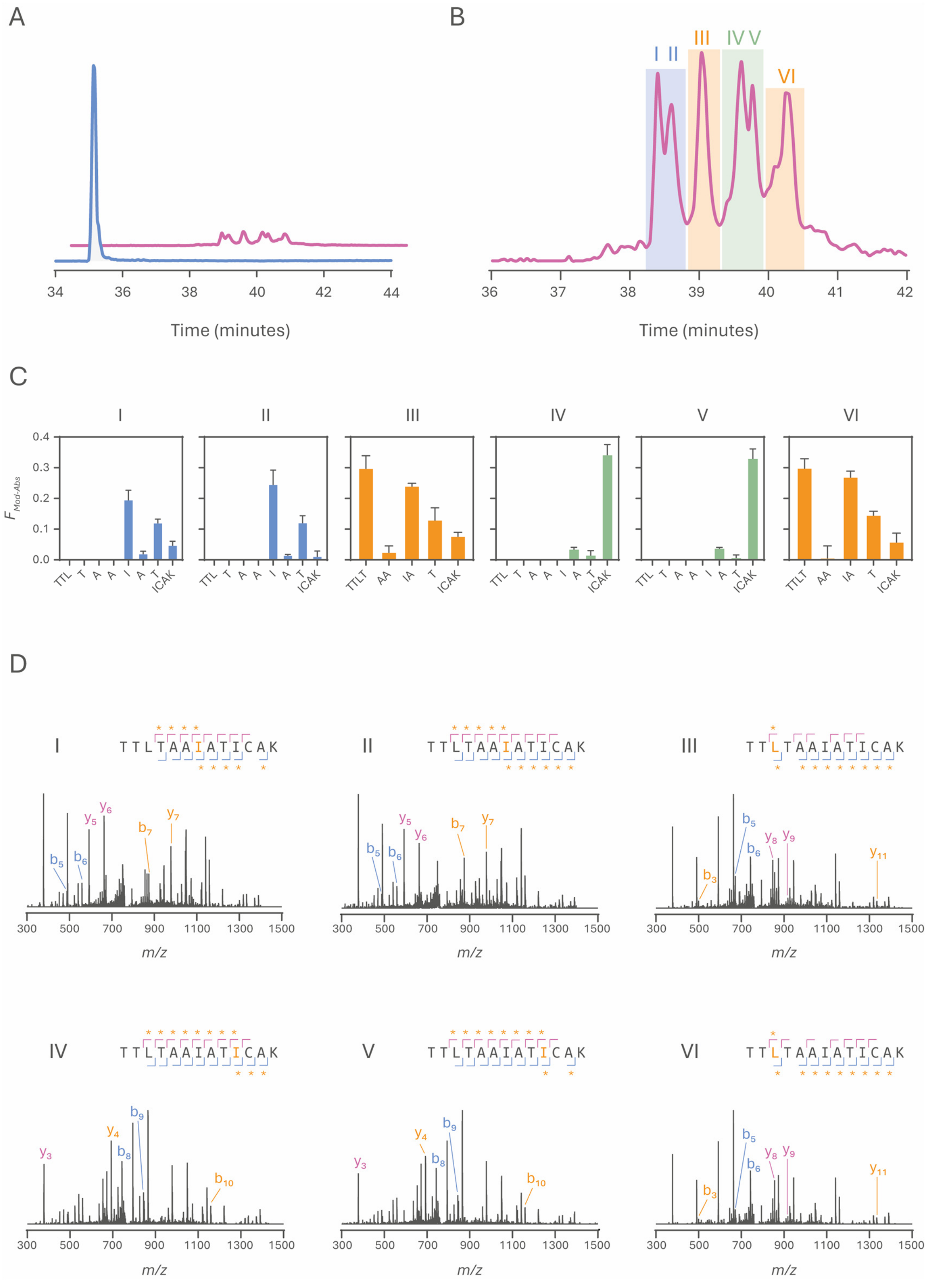


Figure 4.18 | MS^2 analyses of the isomerically labelled peptide Thr^{32} – Lys^{44}. Main figure legend overleaf.

Figure 4.18 | Cont. A | XICs of labelled (pink) and unlabelled (blue) Thr^{32} – Lys^{44}, normalised to the unlabelled. B | Normalised labelled XIC showing six distinct chromatographic peaks, demonstrating three individual labelling patterns (shaded regions). C | $F_{Mod\text{-}Abs}$ histograms compiled from summed MS^2 scans extracted from each peak (coloured as in B). Data are an average of four repeats, error bars show one standard deviation. D | ProteinMetrics derived representative MS^2 scans from each labelled chromatographic peak, showing TDBA insertion localised to three specific residues (orange). For clarity, only *b* and *y* ions corresponding to those carrying the additional *m/z* from TDBA insertion (orange) and immediately adjacent ions are shown. All detected ions are indicated in the label above each spectrum, orange asterisks indicate ions carrying the additional *m/z* from TDBA insertion. Observed and expected *m/z* values of each fragment ion can be found in appendix III, v – x for isomers I – VI respectively.

The TDBA labelled tryptic peptide, Thr^{32} – Lys^{44} (helix *A*) exhibits a complex chromatographic profile (figure 4.18, A. B). Six distinct chromatographic peaks were observed eluting over a combined two minutes which, as expected, was approximately three minutes later than the unlabelled peptide. Absolute F_{Mod} MS^2 analysis was performed by manual combining of MS^2 scans under each chromatographic peak (figure 4.18, B), alongside a separate DDA search using ProteinMetrics (section 2.3.3.3; figure 4.18, D). Interestingly, the six individual chromatographic peaks demonstrated only three pairs of distinct absolute F_{Mod} distributions. The first pair, chromatographically separated by approximately 15 seconds showed preferential

TDBA insertion on Ile^{38} ($F_{Mod\text{-}Abs}$ of 0.197, peaks I and II figure 4.18, B, C). Followed by Thr^{40} ($F_{Mod\text{-}Abs}$ of 0.122), and minor insertion on Ala^{39} and Ile^{41} – Lys^{44}. Here, TDBA insertion into two different positions on the isoleucine sidechain is likely responsible for this chromatographic difference. This localisation was corroborated by repeated analysis in ProteinMetrics (figure 4.18, D), where TDBA insertion was found to localise to Ile^{38} in MS^2 scans within those chromatographic peaks. A similar effect was seen for peaks IV and V (chromatographic separation of approximately 15 seconds, figure 4.18, B, C). Here, absolute F_{Mod} narrowed TDBA insertion, with limited resolution, to Ile^{41} – Lys^{44} ($F_{Mod\text{-}Abs}$ of 0.343), alongside minor insertion within Ile^{38} and Thr^{40}. Helpfully, analysis in ProteinMetrics localised the TDBA insertion specifically to Ile^{41} (figure 4.18, D), again likely showing TDBA insertion into different positions of the isoleucine sidechain.

More interestingly, peaks III and VI (figure 4.18, B, C) showed extensive chromatographic separation (approximately 80 seconds). Absolute F_{Mod} showed TDBA insertion throughout the peptide, but preferentially into Thr^{32} – Thr^{35} ($F_{Mod\text{-}Abs}$ of 0.299). These residues were not found to be labelled in any of the other chromatographic peaks. ProteinMetrics showed TDBA insertion to be localised to Leu^{34} (figure 4.18, D) in both chromatographic peaks. Interestingly, the site of insertion in peak VI must impart a significant reduction in polarity compared to III. It is possible that TBDA is inserting into the leucine sidechain in peak III, and into the backbone in peak IV, possibly either into the amine or carbonyl. Unfortunately, sub-residue (MS^3) analysis would be difficult to perform in this example. The labelled peptide is of significantly lower intensity than the unlabelled (figure 4.18, A), and even more so in EF-Tu•GDPNP•ENX. The combined use of alternative fragmentation methods to generate $a - x$ and $c - z$ ions could elucidate backbone TDBA localisation. The sub-residue level carbene footprinting performed by Lu *et al,* focused on *ortho meta* and *para* TDBA insertion into the phenylalanine phenyl group and not on backbone/sidechain variations, so no comparisons can be drawn there.[198]

4.3.3.5. Repeat EF-Tu EF-Tu•GDPNP•ENX Carbene Footprinting with Four-fold Excess Enacyloxin IIa.

Carbene footprinting of EF-Tu•GDPNP•ENX was repeated with greater concentrations of EF-Tu (10 µM) and enacyloxin IIa (40 µM) with the aim to a) increase EF-Tu tryptic coverage and b) ensure full occupancy of EF-Tu with enacyloxin IIa. Samples were irradiated as before (section 2.3.2.1) but were then subject to in-solution tryptic digestion (section 2.3.1.4). Data were analysed as previously described (section 2.3.3.2). Rather disappointingly, the increased EF-Tu concentration paired with in-solution digestion yielded a modest increase in protein coverage (71.1 % *vs* 56.5 %) across 29 peptides. This is lower than that seen in Chapter 3 under similar conditions (enacyloxin IIa *vs* GE2270A). Perhaps minor contaminating compounds in the enacyloxin IIa samples, or enacyloxin IIa itself, induces a less efficient tryptic digestion. Nevertheless, the global masking effect induced by enacyloxin IIa binding to EF-Tu•GDPNP was found to be more pronounced (figure 4.19) under these conditions. Of the 29 peptides detected, 22 showed a significant reduction in $F_{Mod,}$ as identified by T-test analysis.

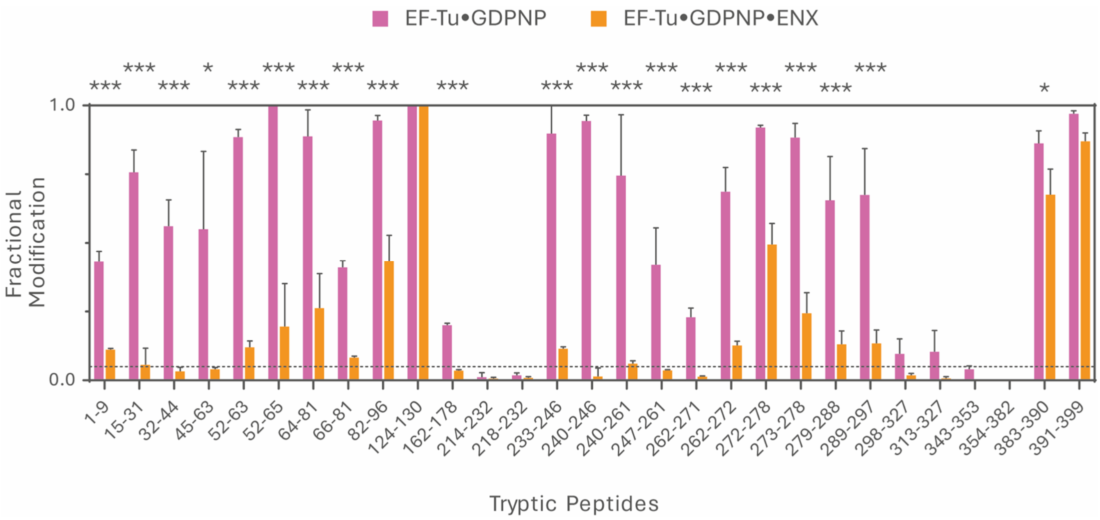


Figure 4.19 | EF-Tu•GDPNP (pink) *vs* EF-Tu•GDPNP•ENX (orange) with 40 µM enacyloxin IIa. Data is an average of five replicates, error bars represent one standard deviation. Stars indicate significant reductions in F_{Mod}, in EF-Tu•GDPNP•ENX compared to EF-Tu•GDPNP. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Dotted line indicates an F_{Mod} of 0.05.

At first glance, this would appear to be a huge increase in significant masking compared to data shown in figure 4.13, A. However, a large amount of significantly masked peptides in figure 4.19 are missed cleavages of peptides shown in figure 4.13, A, or novel peptides covering structural regions shown to be masked prior. On balance, peptides shown to be masked at 1.1 molar equivalent enacyloxin IIa exhibited greater reductions in F_{Mod} under higher enacyloxin IIa conditions, showing that perhaps there was non-saturated binding of EF-Tu by enacyloxin IIa under the previous conditions. There were peptides found to have significantly reduced F_{Mod} under these conditions that did not exhibit the same response in section 4.3.3.3. Peptides Asp^{52} – Lys^{63} and the missed cleavage Asp^{52} – Arg^{65} both showed significant reductions in F_{Mod} with 40 µM enacyloxin IIa compared to the EF-Tu•GDPNP control (-0.766 and -0.802 respectively, both $p < 0.001$). Incubated with 1.1-fold enacyloxin IIa (figure 4.13, B), Asp^{52} – Lys^{63} showed a modest, but insignificant reduction in F_{Mod} (-0.234) and Asp^{52} – Lys^{65} showed no real difference between EF-Tu•GDPNP and EF-Tu•GDPNP•ENX (F_{Mod} of 0.935 and 0.949 respectively). These peptides, along with the prior Thr^{45} – Lys^{63} cover the switch I region of EF-Tu, which as discussed previously, undergoes a large conformational change between *open* and *closed* EF-Tu.

His^{82} – Lys^{96}, a peptide not seen in section 4.3.3.3, showed significant F_{Mod} reduction in EF-Tu•GDPNP•ENX (-0.511, $p < 0.001$). This peptide encompasses strand *d* and the N-terminus of helix *B* (switch II), MS^2 analysis placed TDBA insertion to Pro^{89} and His^{91} localised in the connecting strand, proximal to the N-terminus of helix *B*. Pro^{89} falls under the disordered looping region of switch I in *closed* EF-Tu, resulting in complete burial (SASA reduced from 24.24 to 0.00 Å^2 in *open* and *closed* EF-Tu respectively). Interestingly, the detected peptides covering Gly^{233} to Arg^{297}, a large proportion of domain II, all showed significant reductions in F_{Mod}.

The two peptides, Gly^{233} – Lys^{246} and Val^{240} – Lys^{246}, showed minor and insignificant masking in figure 4.13, B but showed large, significant reductions in F_{Mod} under these conditions (-0.383 and -0.930 respectively, both $p < 0.001$). Gly^{233} – Lys^{246} does become highly buried in *closed* EF-Tu (SASA reduced from 317.08 to 185.07 Å^2 compared to *open* EF-Tu, table 4.2), forming the interface between domains I and II. This masking is indicative of a more *closed* conformation being achieved under this much greater excess of enacyloxin IIa. Peptides Val^{247} – Lys^{261}, Val^{247} – Lys^{261}, and Thr^{262} – Lys^{272}, which all showed qualitative non-significant reductions in F_{Mod} in figure 4.13, B, were found to all be significantly masked in figure 4.19 (F_{Mod} reductions of -0.384, -0.216 and -0.560 respectively, all $p < 0.001$). Again, these form the interface between domains I and II, indicating a more *closed* conformation. Peptides spanning Lys^{272} –

Arg^{297} showed a similar effect, with all showing large, significant reductions in F_{Mod}. Two, Ala^{279} – Arg^{288} and Gly^{289} – Arg^{297} were significantly masked under both conditions.

Overall, these data suggest that EF-Tu has likely adopted a more *closed* average conformation with this greater excess of enacyloxin IIa. It is possible that under 1.1-fold excess enacyloxin IIa, as seen in section 4.3.3.3, EF-Tu was not fully bound by enacyloxin IIa, and therefore the population on average was not in a *closed* conformation. Enacyloxin IIa is an extremely fragile compound, and even under storage at -20 °C in the dark, had likely degraded to some degree. Because of this, it is difficult to know an accurate concentration of active enacyloxin IIa. The carbene footprinting in this section was performed with fresh enacyloxin IIa, at a much higher molar excess than EF-Tu, likely leading to full saturation and a more detectable conformational change.

4.3.3.6. *Carbene Footprinting of Enacyloxin IIa Resistant EF-Tu*

Section 4.3.2.1 shows the successful characterisation of enacyloxin IIa resistant EF-Tu mutants, each demonstrating a differing level of resistance. As discussed previously, each of the three mutants induces enacyloxin IIa resistance by preventing binding of the antibiotic to EF-Tu. If the dramatic reductions in F_{Mod} shown in section 4.3.3.3 are a direct result of enacyloxin IIa binding, we hypothesised that these resistant mutants would show fewer extreme changes in F_{Mod}. To investigate this initially EF-Tu•GDPNP•ENX and the three resistant mutants (EF-Tu^{Q131K}•GDPNP•ENX, EF-Tu^{G325D}•GDPNP•ENX and EF-Tu^{A384T}•GDPNP•ENX) were subject to carbene footprinting performed as described in section 2.3.2.1 prior to SDS-PAGE purification (section 2.3.1.2), in-gel digestion (section 2.3.1.3) and HPLC-MS analysis (section 2.3.1.4). F_{Mod} was quantified as described in section 2.3.3.2.

As stated in section 4.3.2.1, the sub-micromolar binding affinity of enacyloxin IIa to the wild-type and mutant EF-Tu•GDPNP complexes led to their discrimination with carbene footprinting impossible.

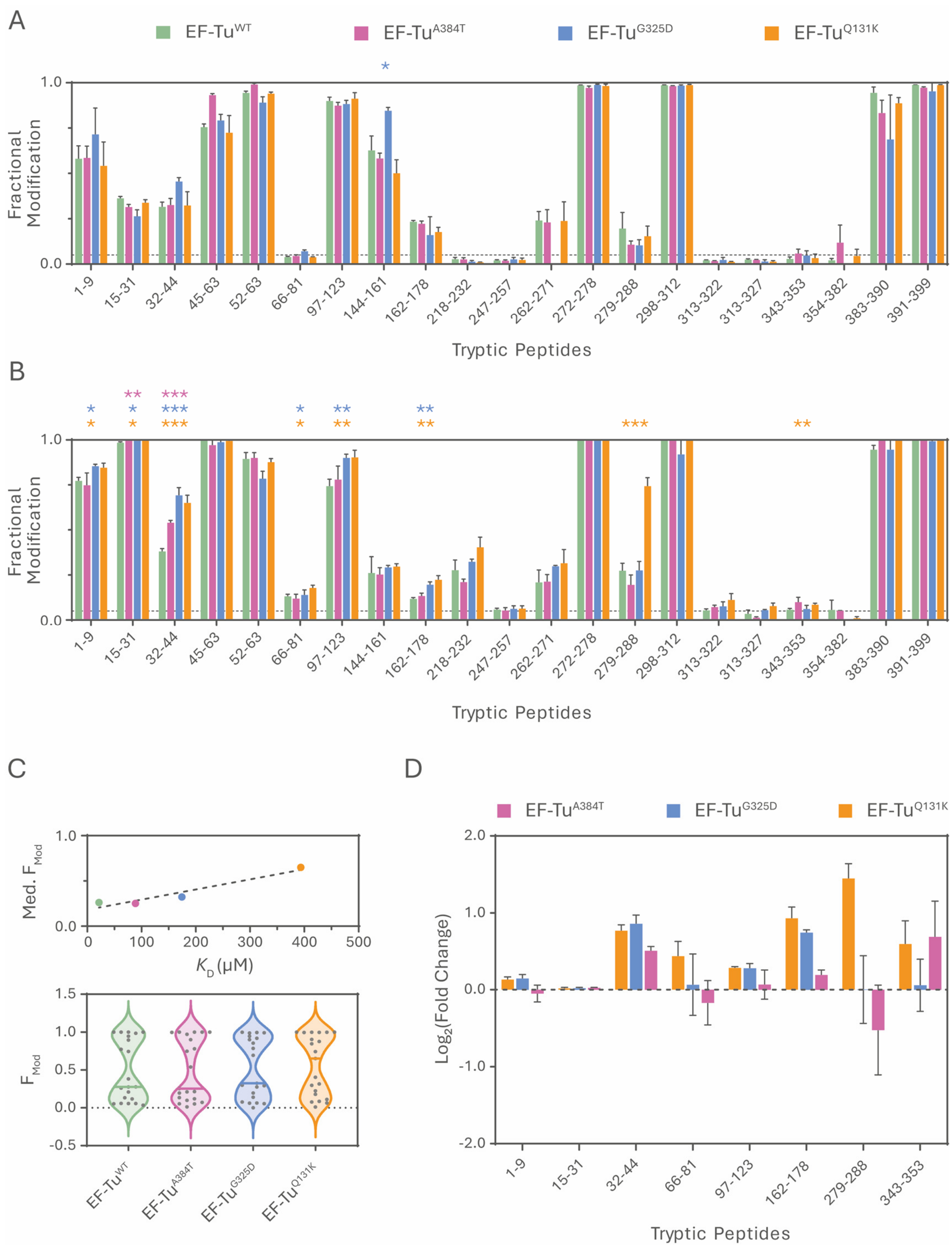


Figure 4.20 | Carbene footprinting of enacyloxin IIa resistant EF-Tu Main figure legend overleaf.

Figure 4.20 | Cont. A | Carbene footprinting of wild-type (EF-Tu^{WT}) and enacyloxin IIa resistant mutants; EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} in complex with GDPNP and enacyloxin IIa. B | Carbene footprinting of wild-type (EF-Tu^{WT}) and enacyloxin IIa resistant mutants; EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} in complex with GDP and enacyloxin IIa. C | Correlation of K_D values to median F_{Mod} (top) for EF-Tu^{WT}, EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T}, $R^2 = 0.93$, coloured as below. Violin plots (bottom) of footprinting data in B showing protein-level labelling data for each species. Median values represented by the solid horizontal lines. Coloured as above. D | Fold change in F_{Mod} for peptides from B that have a significant increase in F_{Mod} in EF-Tu^{Q131K} compared to EF-Tu^{WT}. Data is an average of four replicates, error bars represent one standard deviation. Stars indicate significant changes in F_{Mod}, compared to EF-Tu^{WT}, coloured by mutant species. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Dotted line in A and B indicates an F_{Mod} of 0.05.

Figure 4.20, A shows that no significant differences in F_{Mod} were seen between wild-type EF-Tu•GDPNP or the three mutants, when footprinted in the presence of a 1.1-fold excess of enacyloxin IIa. As in section 4.3.2.1, we instead turned to investigation of GDP-bound EF-Tu which, as shown in figure 4.9, A, demonstrated a large range of binding affinities towards enacyloxin IIa.

Wild-type EF-Tu•GDP•ENX, alongside EF-Tu^{Q131K}•GDP•ENX, EF-Tu^{G325D}•GDP•ENX and EF-Tu^{A384T}•GDP•ENX (all 5 μM), were footprinted in the presence of 40 μM enacyloxin IIa. From native mass spectrometry measurements (figure 4.9, B) at this enacyloxin IIa concentration, wild-type EF-Tu•GDP, EF-Tu^{Q131K}•GDP, EF-Tu^{G325D}•GDP and EF-Tu^{A384T}•GDP should show discriminatory levels of enacyloxin IIa binding, in a K_D-dependent manner. Pleasingly, a K_D-dependent increase in F_{Mod} was seen under these conditions (figure 4.20, B). As enacyloxin IIa binding induces a shift to a more *closed* EF-Tu structure, with an associated reduced overall F_{Mod}, inhibition of enacyloxin IIa binding should lead to a less *closed* conformation and therefore a greater overall F_{Mod} when compared to the wild-type. Overall, all EF-Tu•GDP•ENX samples showed an average increase in F_{Mod} over EF-Tu•GDPNP•ENX (figure 4.20, C, averages of 0.435, 0.488, 0.560, 0.516 and 0.493 for wild-type EF-Tu•GDPNP•ENX, wild-type EF-Tu•GDP•ENX, EF-Tu^{Q131K}•GDP•ENX, EF-Tu^{G325D}•GDP•ENX and EF-Tu^{A384T}•GDP•ENX respectively). This result is as expected, if EF-Tu adopts, on average, a less *closed* structure when bound to GDP *vs* GDPNP. When comparing individual peptides in mutant EF-Tu•GDP•ENX to wild-type EF-Tu•GDP•ENX, 8 show a significant increase in F_{Mod} in EF-Tu^{Q131K}•GDP•ENX ($p \leq 0.04$), 6 in EF-Tu^{G325D}•GDP•ENX ($p \leq 0.03$) and 2 in EF-Tu^{A384T}•GDP•ENX ($p \leq 0.005$).

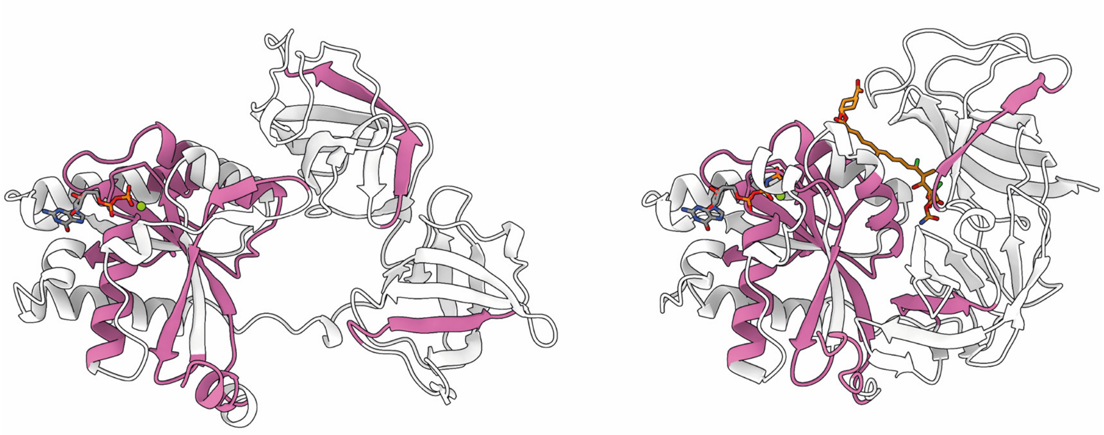


Figure 4.21 | Peptides (pink) in EF-Tu^{Q131K}•GDP•ENX demonstrating a significantly greater F_{Mod} than in wild-type EF-Tu•GDP•ENX mapped to *open* and *closed* EF-Tu homology models. GDP and GDPNP (both coloured by element) and enacyloxin IIa (orange) are represented as sticks.

This again follows the measured trend in K_D for these species, showing an increase in the number of peptides showing significantly greater F_{Mod} as the K_D, and therefore resistance, increases (Figure 4.21). Figure 4.20, D shows the fold-F_{Mod} change between mutant EF-Tu•GDP•ENX and wild-type EF-Tu•GDP•ENX for those 8 peptides with significantly higher F_{Mod} in EF-Tu^{Q131K}•GDP•ENX compared to wild-type EF-Tu•GDP•ENX. This is plotted as the Log_2(fold-F_{Mod} change) as to equate the visual differences of equivalent positive and negative fold changes.[260] Upon examination, the N-terminal peptide Gly^1 – Lys^9 had a significantly F_{Mod} in EF-Tu^{Q131K}•GDP•ENX and EF-Tu^{G325D}•GDP•ENX (0.845 and 0.854, $p = 0.02$ and 0.003 respectively) compared to wild-type EF-Tu•GDP•ENX (0.772). The free N-terminus is positioned where domain II contacts domain I in the *closed* EF-Tu conformation. Although it is structurally disordered, it is possible that N-terminal interactions with domain II in *closed* EF-Tu limit TDBA accessibility, and this protection is lost in the *open* conformation.

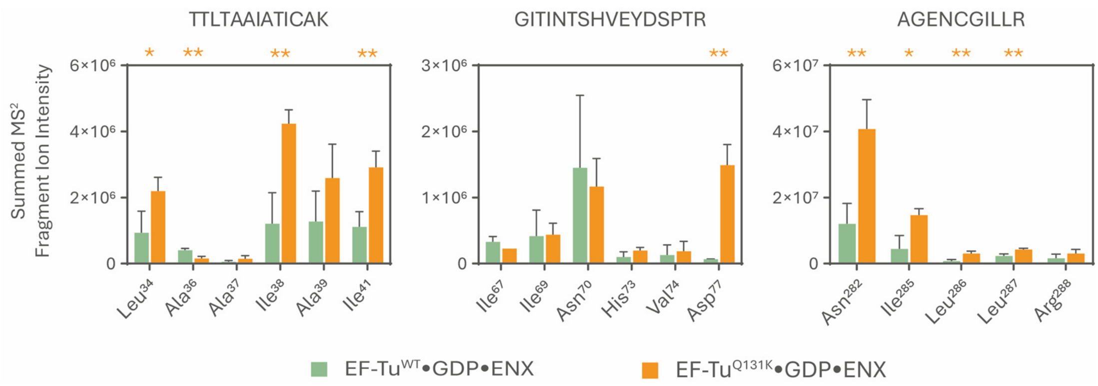


Figure 4.22 | MS^2 residue level footprinting data. Comparison of select peptides Thr^{32} – Lys^{44} (A), Gly^{66} – Arg^{81} (B) and Ala^{279} – Arg^{288} (C)showing significantly higher F_{Mod} in EF-Tu^{Q131K} (orange) over EF-Tu^{WT} (green) and having sufficient ion intensities. The summed intensities of each fragment ion are reported. Data is an average of three replicates, error bars represent one standard deviation. Stars indicate significant changes in F_{Mod}, compared to EF-Tu^{WT}. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

Asn^{15} – Lys^{31} showed a very minor, yet significantly ($p \leq 0.04$) greater F_{Mod} in all mutants compared to the wild-type (F_{Mod} of 0.985, 1.000, 0.997 and 0.996 for wild-type, EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} respectively). As discussed in section 4.3.3.3, this peptide covers strand *a*, and the nucleotide-binding P-loop. Importantly, major TDBA accessibility of this peptide is dictated by the central pore in *open* EF-Tu, demonstrated by the significant reductions in F_{Mod} seen in enacyloxin IIa-induced *closed* EF-Tu (figure 4.13). Here, significant increases in F_{Mod} were seen as the EF-Tu conformations are expected to become more *open*.

The same is true for Thr^{32} – Lys^{44}, as this peptide is expected to be masked in *closed* EF-Tu, due to the conformational changes induced in the proximal switch I region compared to *open* EF-Tu. Again, a significant (all $p < 0.001$) increase in F_{Mod} was seen for all mutant EF-Tu over the wild-type (0.381, 0.540, 0.692 and 0.650 for wild-type, EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T} respectively), indicating a resistance-dependent *open* conformation. MS^2 analysis of wild-type EF-Tu and EF-Tu^{Q131K} (figure 4.22, A) indicated significantly differential increases in labelling localise predominantly to Ile^{38}, Ile^{41} and Leu^{34} and to a lesser extent, Ala^{36}, indicative of the whole of helix *A* becoming more accessible in EF-Tu^{Q131K} (therefore a more *open* structure).

Gly^{66} – Arg^{81}, another switch I associated peptide masked in *closed* EF-Tu, showed significant increases in F_{Mod} in EF-Tu^{Q131K} over wild-type EF-Tu (0.187 and 0.132 respectively, $p = 0.01$), and a qualitative but non-significant K_D-dependent change in F_{Mod}. The significant increase in F_{Mod} was found to be localised entirely to Asp^{77} by MS^2 analysis, a residue occluded by domain II in *closed* EF-Tu and accessible in *open* EF-Tu (figure 4.22, B). The large peptide Asn^{97} – Arg^{123}, proximal to the enacyloxin IIa binding cleft, is significantly ($p \leq 0.008$) more labelled in EF-Tu^{Q131K} and EF-Tu^{G325D} over the wild-type (F_{Mod} of 0.903, 0.900 and 0.743 respectively), again showed a K_D-dependent increase in F_{Mod}. Glu^{162} – Arg^{178} contains the enacyloxin IIa binding residue, Tyr^{167}, and showed a significantly ($p = 0.002$ in both) higher F_{Mod} in both EF-Tu^{Q131K} and EF-Tu^{G325D} over wild-type EF-Tu (0.223, 0.196 and 0.117 respectively), a perfect K_D-dependent increase in F_{Mod}. The final two peptides, Ala^{279} – Arg^{288} and Thr^{343} – Lys^{322}, found in domains II and III respectively, were both significantly more labelled in EF-Tu^{Q131K} compared to wild-type EF-Tu (0.743 *vs* 0.274 and 0.085 *vs* 0.057, $p = 0.01$ and 0.02 respectively). Both peptides become part of the interface of the three domains in *closed* EF-Tu, and because of that, less accessible to TDBA (figure 4.17). Greater labelling would infer EF-Tu adopting a more *open* conformation.

Interestingly, Ala^{279} – Arg^{288} had a reduced F_{Mod} in EF-Tu^{A384T} compared to wild-type EF-Tu, although this was not a significant difference. MS^2 analysis of the same peptide in EF-Tu^{Q131K} (figure 4.22, C) narrowed down this significant increase in F_{Mod} to Asn^{282} and Ile^{285} – Ile^{287}, with the predominant difference being at Asn^{282}. This residue is buried entirely next to domain I in *closed* EF-Tu (SASA of 0.00 compared to 84.74 Å^2 in *open* EF-Tu), again, indicative of EF-Tu^{Q131K} adopting a more *open* conformation. Together, the protein, peptide and residue level F_{Mod} values provide strong evidence that there is a K_D-, or resistance-dependent effect on F_{Mod}, and that binding of enacyloxin IIa is responsible for the global conformational change seen in EF-Tu.

4.4. Conclusions

EF-Tu proved to be an interesting protein to study using carbene footprinting. The high conformational flexibility of EF-Tu made it an excellent subject to assess the capabilities of carbene footprinting to elucidate such large conformational changes. Through native mass spectrometry, enacyloxin IIa was shown to bind to *A. baumannii* EF-Tu. Although not unexpected, *A. baumannii* currently has no published structures of EF-Tu in complex with enacyloxin IIa or other kirromycin-like elfamycins. Enacyloxin IIa resistant mutants were produced, with single residue substitutions based upon those conferring resistance to kirromycin in *E. coli* EF-Tu. These mutations were found to induce enacyloxin IIa resistance at differing rates, as quantified by native mass spectrometry K_D measurements (in descending level of binding affinities; EF-Tu^{Q131K}, EF-Tu^{G325D} and EF-Tu^{A384T}). Both *apo*-EF-Tu and EF-Tu•GDPNP were found to be conformationally similar by carbene footprinting experiments. This is in agreement with published literature, demonstrating the fluidity of EF-Tu to move from *open* and *closed* conformations. GDPNP (GTP) was classically thought to induce EF-Tu's adoption of the *closed* conformation, however recent solution-phase, and crystal, structural analyses have demonstrated EF-Tu can adopt an *open* conformation bound to GDPNP.

Through addition of enacyloxin IIa to EF-Tu, a markedly different carbene footprinting profile as adopted. EF-Tu•GDPNP•ENX has been shown to adopt a more *closed* conformation than EF-Tu•GDPNP, markedly one with a much-reduced SASA. Carbene footprinting measured global reductions in F_{Mod} across EF-Tu when in complex with enacyloxin IIa. These reductions were seen directly in the enacyloxin IIa binding site, corroborating the native mass spectrometric data showing enacyloxin IIa binding to *A. baumannii* EF-Tu. More interestingly, F_{Mod} reductions were seen globally, residues and peptides known to become less accessible in *closed* EF-Tu were shown to less labelled than in the more *open* EF-Tu•GDPNP. These suggest carbene footprinting is capturing not only the direct binding of enacyloxin IIa to EF-Tu, but also the transition to the more compact structure.

When the three enacyloxin IIa resistant EF-Tu species were subject to carbene footprinting, they showed an overall K_D-dependent increase in F_{Mod} when compared to wild-type EF-Tu. As resistance is conferred by inhibition of enacyloxin IIa binding, it is plausible to assume the adoption of the *closed* conformation is also inhibited. If EF-Tu is more *open* an increase in F_{Mod} is expected, and that is what was seen with each of the mutants. This further strengthens the conclusion that carbene footprinting is capturing this conformational change and would be a valuable tool to investigate other highly dynamic protein systems.

Mapping interactions of EF-Tu and a Putative Prolyl Hydroxylase with Carbene Footprinting

5

5.1. Introduction

5.1.1. Hydroxyproline

Hydroxylation of proline is one of the most common post-translational modifications found in the human proteome. In fact, (2*S*,4*R*)-4-hydroxyproline is found in higher abundance than seven standard amino acids (cysteine, glutamine, histidine, methionine, phenylalanine, tryptophan and tyrosine).[261,262] Enzymatic hydroxylation occurs on one of two carbons, forming (2*S*,3*S*)-3- or (2*S*,4*R*)-4-hydroxyproline, for the purpose of this work I will focus on the latter (henceforth referred to as *Hyp*). In mammals Hyp (alongside 3-hydroxyproline) is found in high concentrations in collagen, representing approximately 14% of the total residues.[263] Collagen aggregates are a quaternary complex formed of three individual tropocollagen polypeptides, each having a repeating pattern of Gly-*X*-*Y*, where X and Y represent any amino acid. If *Y* is occupied by a proline, it is hydroxylated to form Gly-*X*-Hyp. Tropocollagen monomers form a *left-handed* helical structure, which in turn coil around each other to form a *right-handed* coiled coil structure. The presence of Hyp within collagen significantly increases its thermal stability, proportional to its abundance, by approximately 15.5 °C.[264] The full molecular basis behind Hyp-induced collagen stabilisation has not been elucidated, but it is hypothesised to be a combination of two factors: ring puckering and hydrogen bonding networks. In Gly-Pro-Pro collagen, the *X* position proline typically adopts an *endo* configuration (figure 5.1, A). If instead the *Y* proline is hydroxylated, the *X* position proline is structurally frustrated, and can adopt either an *endo* or *exo* conformation (figure 5.1, A). This induces an inherent flexibility in the collagen tetramer and allows it to adopt its ideal 7/2 helical structure (figure 5.1, B). Alongside this, it is hypothesised that Hyp allows for formation of more inter-tropocollagen hydrogen bonds through bridging water molecules, inducing greater stability of the quaternary structure.[262,265,266] Although lack of prolyl hydroxylation in collagen is not a lethal condition, it does lead to several connective tissue disorders including skeletal dysplasia and hypermobility.[267] Major symptoms of scurvy, namely fragile capillaries, bone and teeth weakening are caused by reduced Hyp abundance in collagen, due to reduced availability of ascorbic acid for use as an essential prolyl-hydroxylase co-factor.[268]

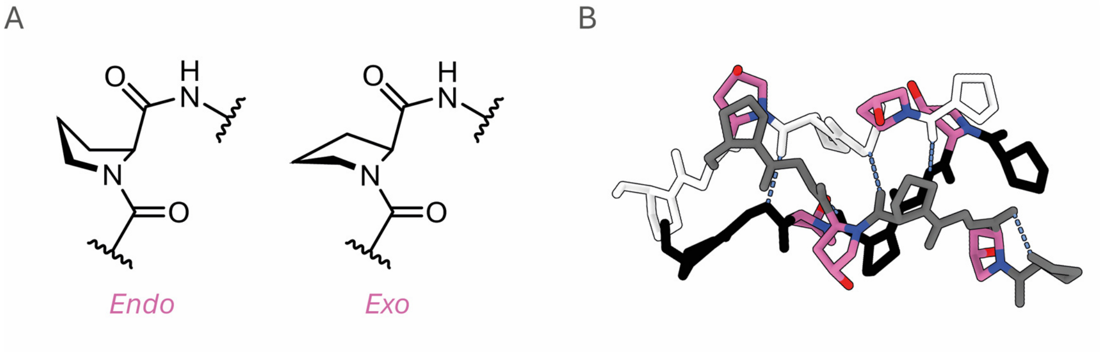


Figure 5.1 | Collagen and proline overview. A | Skeletal representations of *endo* and *exo* configurations of proline. B | One helical turn of collagen with the repetitive sequence Pro-Hyp-Gly. Tropocollagen peptides are coloured white, grey and black. Hydroxyprolines coloured pink and hydrogen bonds represented as dashed blue lines.

Outside of collagen, Hyp has been localised to relatively few proteins or peptides. Typically, collagen-like proteins (*e.g.* elastins) contain Hyp at concentrations of up to 25 %.[269] Hyp has been found to play some roles in cell signalling, specifically acting upon kinases. The dual-specificity tyrosine-phosphorylation-regulated kinases 1A and 1B (DYRK1A, 1B) and protein kinase B are two of the seven proposed prolyl hydroxylated kinases.[270–272] In higher eukaryotes, prolyl hydroxylation plays a part in the hypoxia-inducible factor (HIF) oxygen sensing pathway.[273]

5.1.2. HIF Prolyl Hydroxylation

HIFs are highly conserved, heterodimeric eukaryotic transcription factors formed of an α and β subunit. They respond exclusively to hypoxic or pseudohypoxic cellular conditions and induce upregulation of glycolytic enzymes, erythropoiesis, and production of vascular endothelial growth factors.[274] To understand the process of this transcriptional regulation, the molecular mechanism behind prolyl hydroxylation must be first introduced.

HIFs (specifically their alpha-subunits) are hydroxylated by HIF prolyl-hydroxylases domain (PHD) proteins (PHD 1, 2 and 3), which are members of the α-ketoglutarate-dependent hydroxylases (AKDH) class of non-haem iron proteins. AKHDs have diverse roles across prokaryotic and eukaryotic life, from DNA damage repair, flavonoid biosynthesis, epigenetic regulation, collagen biosynthesis and protein hydroxylation.[275–278] In brief, AKHD binds Fe(II) as a cofactor, coordinated by two histidine and one acidic residue, alongside three water molecules. The Fe(II) centre coordinates α-ketoglutarate and is subsequently bound by molecular oxygen, forming an Fe(III) centre. The superoxide attacks α-ketoglutarate, forming succinate and a new Fe(IV)-oxo intermediate, whilst releasing carbon dioxide. This then first abstracts a proton from the locally coordinated substrate proline, forming a hydroxyl-bound Fe(III) centre, which is then attacked by the proline radical forming a transient hydroxylated proline bound to the Fe(III) centre. Subsequent release of the hydroxylated proline and succinate allows the cycle to repeat.[279]

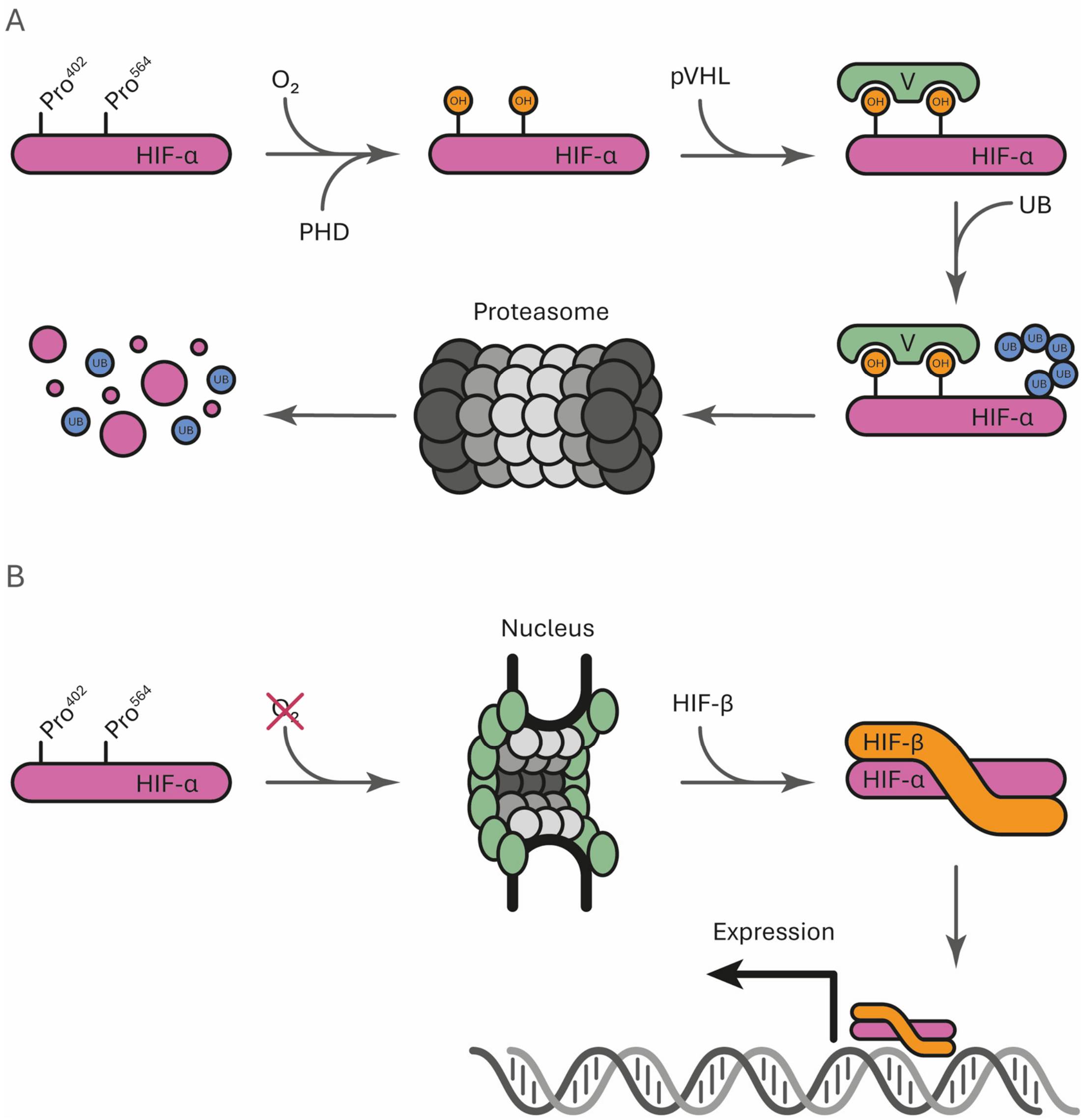


Figure 5.2 | A simplified graphical overview of the HIF oxygen sensing pathway under normoxic (A) or hypoxic (B) conditions. V, von Hippel–Lindau tumour suppressor; UB, ubiquitin.

As described, prolyl hydroxylation is an oxygen-dependent reaction and therefore not possible, or possible with reduced efficiency, in hypoxic or partially hypoxic conditions. Under normoxic conditions, the cytosolic HIF-α is hydroxylated at Pro^{402} and Pro^{564} (numbered as found in human HIF-1α).[280] This leads to its recognition by an E3 ubiquitin ligase complex, specifically by a von Hippel–Lindau (VHL) tumour suppressor.[281,282] Ubiquitination then marks HIF-α for degradation through the ubiquitin–proteasome system, therefore preventing formation of the full HIF transcription factor complex.[281] Under hypoxic conditions, insufficient oxygen is available to perform HIF-α hydroxylation, and therefore it is not ubiquitinated for degradation. HIF-α is then able to move to the nucleus, dimerise with HIF-β, and upregulate hypoxia-associated genes.[280] Interestingly, the hydroxylation of the two prolines differ in a hypoxia-dependent relationship. Hyp^{402} was found to be more sensitive to hypoxic conditions than Hyp^{564}, showing approximately 50 and 75 % hydroxylation respectively at physiologically severe hypoxia (0.5 % O_2).[283] Hydroxylation of both prolines was found to be required for fast degradation.[284] Iron-chelation, introduction of non-ferrous transition metal ions or influx of succinate lead to *pseudohypoxic* conditions, where HIF-α is not hydroxylated, despite the presence of molecular oxygen, by inhibition of PHDs.[285] A summary of this process can be found in figure 5.2.

5.1.3. EF-Tu as a Prolyl-Hydroxylase Target

Schofield and colleagues characterised a *Pseudomonas putida* strain (PA0310) containing a transposon insertion into a putative PHD (PPHD) gene, finding it to encode a fully functioning PHD.[286] Incubation of the strain in the presence of iron chelators (therefore pseudohypoxic) showed a reduced growth rate compared to wild-type *P. putida*. The PA0310 strain also demonstrated increased pyocyanin production under these conditions, indicative of environmentally stressed cells.[287] These suggest a vital role of PPHD in *P. putida,* tentatively linked to iron uptake. Mass spectrometric proteomic data from *Shewanella oneidensis*, reported EF-Tu possessing a +16 Da mass shift, indicative of hydroxylation. Further bioinformatic analyses showed the presence of a tentative PHD in *S. oneidensis.* Incubation of heterologously expressed PPHD with 14 proline-containing *P. putida* EF-Tu (henceforth, EF-Tu$^{\text{Putida}}$) peptides led to specific hydroxylation of Pro^{54} (figure 5.3, C), contained within a peptide spanning Ile^{44} – Ile^{63} (These residues, and those from here are numbered as according to the crystal structure of PPHD•EF-Tu$^{\text{Putida}}$). The authors then went on to

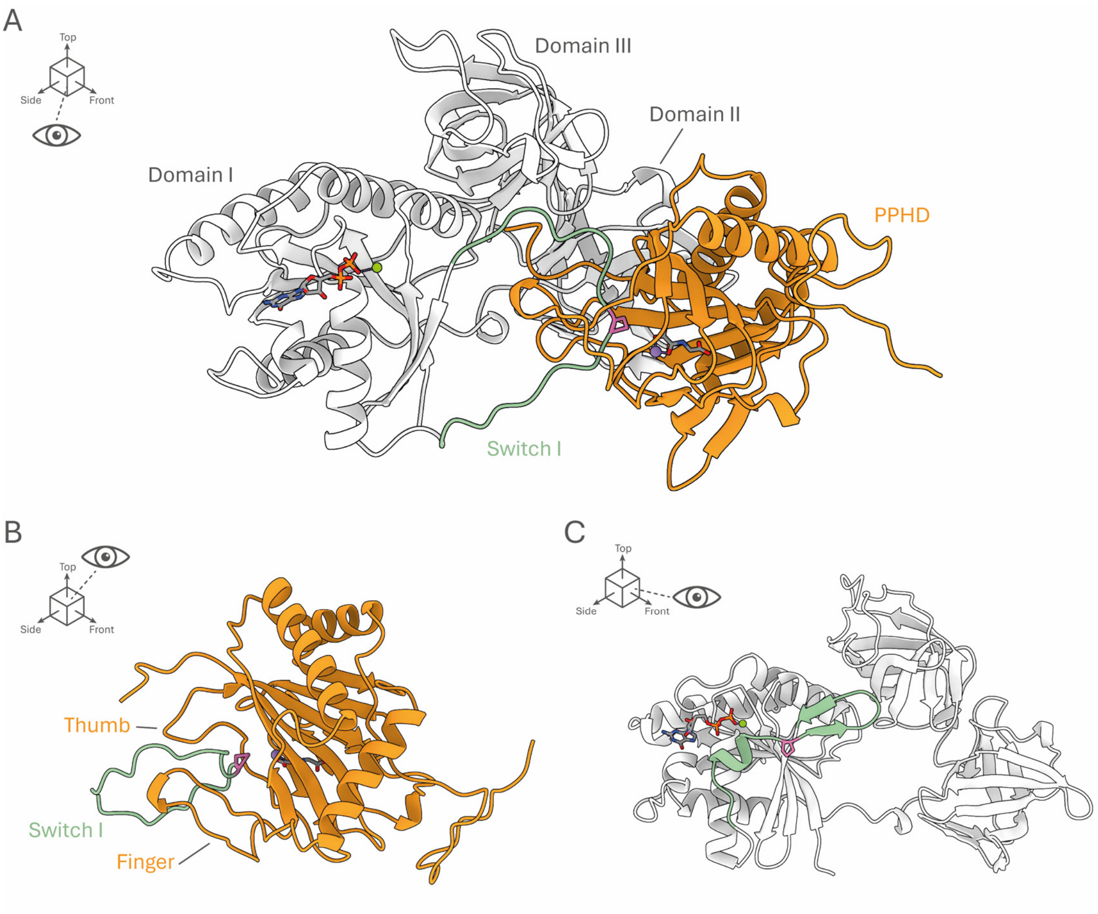


Figure 5.3 | Cartoon homology structures of *P. putida* PPHD and EF-Tu. A | PPHD (orange) and EF-Tu^{Putida} (orange) complex. The unfolded switch I region (green) extends into the PPHD binding site, bringing Pro^{54} (pink) in close contact with the catalytic centre. B | Zoom of PPHD in complex with swich I, the rest of EF-Tu^{Putida} is not shown. The two protruding *finger* and *thumb* loops sandwich switch I. C | Front view of EF-Tu^{Putida}, coloured as in A. In all structures, homology modelling was used to fill missing disordered regions in the crystal structure, namely switch I. GDP and α-ketoglutarate are shown as sticks in EF-Tu^{Putida} and PPHD respectively. Mg^{2+} and Mn^{2+} are shown as green and purple spheres respectively.

successfully crystalise EF-Tu^{Putida} in complex with PPHD (figure 5.3, A), showing binding of PPHD between domains I and II/III. The switch I region of EF-Tu was found to be completely disordered and projected away from domain I, positioning Pro^{54} into the active site of PPHD. This was secured by the two projecting *finger* and *thumb* loops in PPHD, which grasp the looping switch I (figure 5.3, B). Four residues in PPHD, Leu^{101}, Phe^{102}, Leu^{103} and Leu^{114} interact hydrophobically with Phe^{222}, Ile^{224},

Val^{230} and Phe^{265} in $EF\text{-}Tu^{Putida}$ domain II. The rate of hydroxylation was found to be low on full-length $EF\text{-}Tu^{Putida}$, as shown by MS^2 analysis of switch I.

The function of EF-Tu switch I hydroxylation is unknown but evidently not essential, the switch I region is vital to translate GTP hydrolysis into the associated conformational change. Pro^{54} (Pro^{53} in *E. coli* and Pro^{60} in *A. baumannii* EF-Tu) is highly conserved but not involved in any molecular interactions with either bound nucleotide or switch II. Interrogation of published EF-Tu crystal structures does not implicate switch I proline in binding to aa-tRNA or the ribosome, although it does form close contacts (approximately 3.6 Å from the carbon at the 4th position) with the 3' terminus of aa-tRNA. Hydroxylation at this residue must not be a common modification, as no EF-Tu structures are deposited in the PDB containing any hydroxyproline residues. Either PPHD-like enzymes are rare amongst prokaryotes, or $EF\text{-}Tu^{Hyp}$ is a relatively minor modification, and forms a tiny subset of cellular EF-Tu. Hydroxylation of proline, arginine and histidine are common in eukaryotic translation machinery, including in ribosomal proteins, release factors and elongation factors.[288] In some cases, these are essential to their function, yet research into prokaryotic translational hydroxylation is an unexplored field.

5.2. Aims

The work presented within this chapter aims to apply carbene protein footprinting to capture the interactions between $EF\text{-}Tu^{Putida}$ and PPHD. A crystal structure of $EF\text{-}Tu^{Putida}$• PPHD has been solved, however a portion of unfolded switch I is not resolved. Carbene footprinting may allow for the study of PPHD interactions with $EF\text{-}Tu^{Putida}$ in this highly dynamic region. Ideally, production of hydroxylated $EF\text{-}Tu^{Putida}$ could lead to investigation into potential structural ramifications using carbene footprinting. However, the exceedingly low rate of hydroxylation and subsequent inability to selectively isolate and purify hydroxylated $EF\text{-}Tu^{Putida}$, at least within the bounds of this work, makes this very challenging.

5.3. Results and Discussion

5.3.1. Native Mass Spectra of PPHD and EF-Tu^{Putida}.

Samples of *P. putida* and EF-Tu^{Putida} were kind gifts from Professor Chris Schofield (University of Oxford). Samples of PPHD, EF-Tu^{Putida} and PPHD•EF-Tu^{Putida} were prepared as described in section 2.2.1.7, resulting in final concentrations of 10 µM of each protein individually. For PPHD•EF-Tu^{Putida}, 20 µM of each protein was prepared as described in section 2.2.1.7 prior to mixing, resulting in final concentrations of 10 µM. Samples were incubated on ice before loading into prepared emitter tips (section 2.2.1.1) and native spectra acquired on a Synapt G1 HDMS High-Definition mass spectrometer. (section 2.2.1.3).

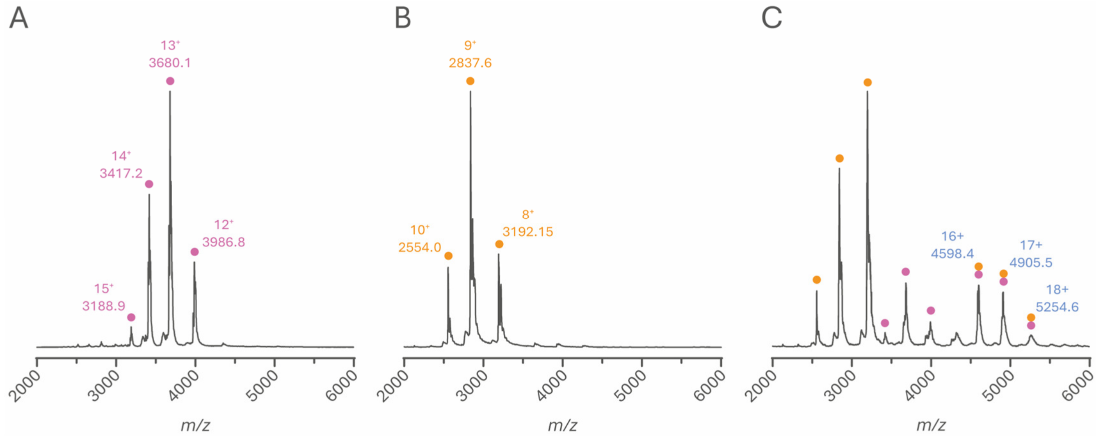


Figure 5.4 | Individually normalised native mass spectra of EF-Tu^{Putida} (A), PPHD (B) and the PPHD•EF-Tu^{Putida} complex (C). Charge states and *m/z* of major peaks are annotated.

Typical native mass spectra can be seen in figure 5.4. EF-Tu^{Putida} was provided with an intact N-terminal 6xHis-tag, with a theoretical expected mass of 47,184.75 Da. Figure 5.4, A shows a typical EF-Tu^{Putida} mass spectrum exhibiting four principal charge states: $[M + 12H]^{12+}$, $[M + 13H]^{13+}$, $[M + 14H]^{14+}$ and $[M + 15H]^{15+}$, with the $[M + 13H]^{13+}$

charge state being dominant under these conditions. Within each charge state three individual peaks can be seen, corresponding to EF-Tu$^{\text{Putida}}$•GDP and adduction by one and two unknown compounds, with an average deconvoluted mass of 177.63 Da. Interestingly, the peak corresponding to EF-Tu$^{\text{Putida}}$•GDP•Adduct$^{177.63}$ was the most dominant species. Deconvolution of peaks corresponding to EF-Tu$^{\text{Putida}}$•GDP gave a mass of 47,649.38 ± 0.15 Da, a difference of +464.63 Da over the theoretical. Bound GDP (443.20 Da) and Mg^{2+} (24.31 Da) accounted for this deviance, with some small error (-2.88 Da).

Figure 5.4, B shows a typical PPHD mass spectrum exhibiting three principal charge states: $[M + 8H]^{8+}$, $[M + 9H]^{9+}$ and $[M + 10H]^{10+}$, with the $[M + 9H]^{9+}$ charge state being dominant. Again, each charge state contained three major peaks with the first corresponding to of N-terminally 6xHis-tagged *apo*-PPHD, deconvolution of which gave a mass of 25,528.17 ± 0.27 Da. This is inexplicably less than the theoretical mass of 25,603.70 Da by -75.53 Da. The two additional peaks show deconvoluted masses of 25,706.99 ± 0.17 and 25,763.12 ± 3.80 Da, corresponding to mass shifts of +178.82 (unknown adduct) and +234.95 Da (addition of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, HEPES, with a molecular weight of +238.30 Da and a small error of -3.35 Da).

Figure 5.4, C shows a typical mass spectrum of co-incubated *apo*-PPHD and EF-Tu$^{\text{Putida}}$, predominantly showing each protein individually, and minor formation of the heterodimeric complex (representing only 16 % of total detected protein). Deconvolution of the $[M + 14H]^{14+}$, $[M + 15H]^{15+}$ and $[M + 16H]^{16+}$ PPHD•EF-Tu$^{\text{Putida}}$ charge states yielded a mass of 73,558.95 ± 8.54 Da. Summation of the measured EF-Tu$^{\text{Putida}}$•GDP and *apo*-PPHD deconvoluted masses yields 73,177.55 Da, giving a deviance of +381.40 Da in the measured PPHD•EF-Tu$^{\text{Putida}}$ mass. The peaks corresponding to PPHD• EF-Tu$^{\text{Putida}}$ were of low intensity and are broad, covering an approximately 100 *m/z* range, so this was not unexpected.

5.3.2. Carbene Footprinting of PPHD and EF-Tu^{Putida}

Prior to carbene footprinting, *in vitro* proteolytic digests (in-gel, section 2.3.1.3) were performed on PPHD and EF-Tu^{Putida} (10 µM each) with trypsin and chymotrypsin, and the obtained sequence coverage compared (figure 5.5). Samples were analysed using DDA HPLC-MS^2 (section 2.3.1.5) and analysed as described in section 2.3.3.1.

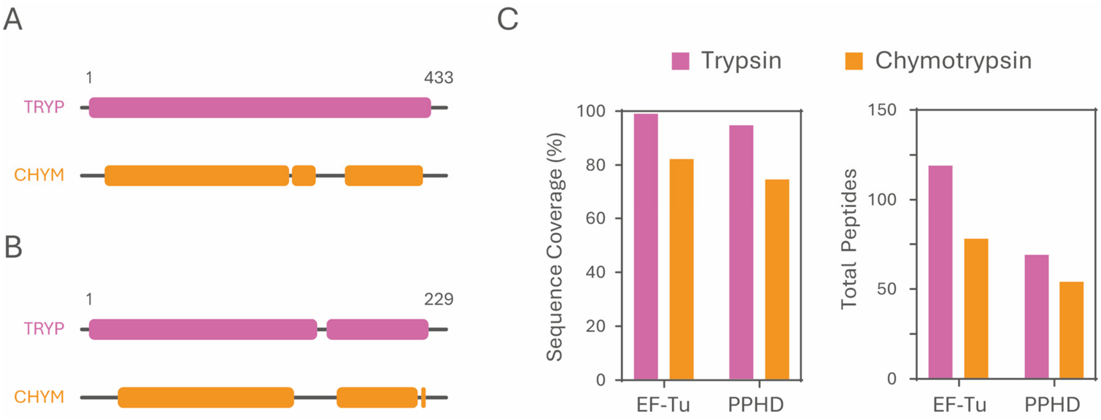


Figure 5.5 | HPLC-MS^2 measured, proteolytic EF-Tu^{Putida} and PPHD sequence coverage. Graphical representation of EF-Tu^{Putida} (A) and PPHD (B) protease sequence coverage, bars indicate regions containing detected peptides. C | Sequence coverage and total peptide number measured for each protein and protease, coloured as in A.

Under these digestion conditions, excellent sequence coverage was obtained by tryptic digestion of both EF-Tu^{Putida} (figure 5.5, A) and PPHD (figure 5.5, B), returning 99.08 and 94.32 % coverage across 119 and 78 peptides respectively (figure 5.5, C). Chymotryptic digestion also produced good sequence coverage (94.32 and 74.68 % across 69 and 54 peptides for EF-Tu^{Putida} and PPHD respectively (figure 5.5, C). Both proteases were used for digestion post-carbene footprinting.

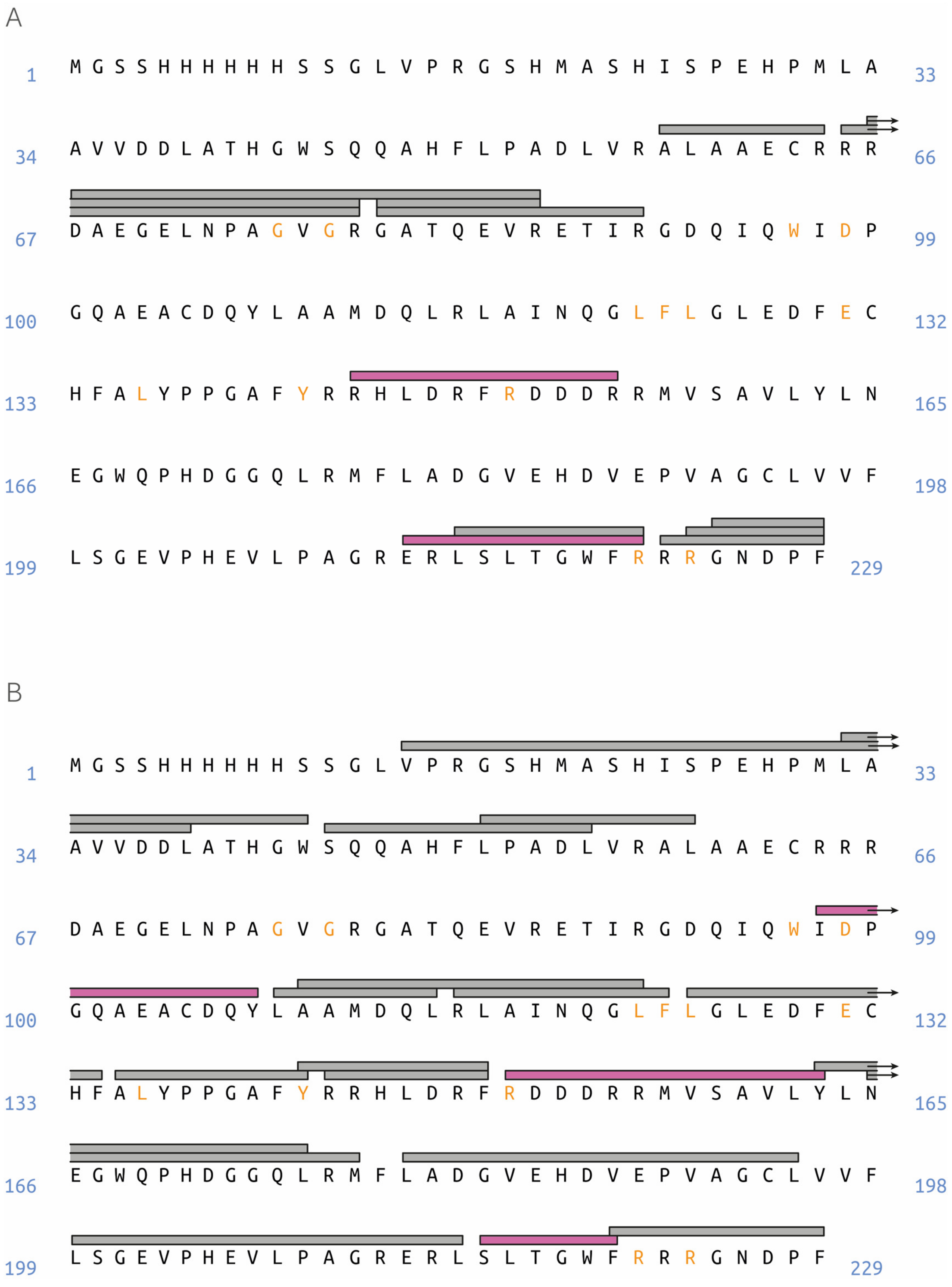


Figure 5.6 | Labelled peptide sequence coverage of PPHD post-carbene footprinting for tryptic (A) and chymotryptic (B) digests. Peptides coloured grey showed no significant reductions in F_{Mod} in PPHD•EF-Tu^{Putida} *vs apo*-PPHD, and those coloured pink did show significant reductions. Residues coloured orange are those involved in non-bonding contacts and hydrogen bonding with EF-Tu^{Putida}.

Carbene footprinting of PPHD and EF-Tu^{Putida} was performed on their respective *apo* and *holo* states (*apo*-EF-Tu^{Putida} in this instance being EF-Tu^{Putida}•GDP). Tryptic and chymotryptic digestions were used for each protein. Carbene footprinting was performed as described in section 2.3.2.4, where 8 µM of the protein of interest was incubated with 10 µM of its binding partner, in order to provide a molar excess. Samples were irradiated in the presence of 10 mM TDBA for 15 seconds.

5.3.2.1. Carbene Footprinting of PPHD

Disappointingly, tryptic digestion of PPHD post-carbene footprinting yielded poor sequence coverage of labelled peptides (figure 5.6, A). Only 13 peptides were identified, representing 26.64 % of the protein sequence. Of these, two peptides showed significant reductions in F_{Mod} in PPHD•EF-Tu^{Putida} compared to *apo*-PPHD (figure 5.7, A and figure 5.8, A). Significance was defined as $p < 0.05$ as calculated by a Multiple T-test and corrected for multiple comparisons with the Holm-Šídák method. *Apo*-PPHD peptides with a $F_{Mod} < 0.05$ were disregarded from statistical analyses. Residues from here are numbered according to the 6xHis-tagged sequences of PPHD and EF-Tu^{Putida}, not those reported in the crystal structure of PPHD•EF-Tu^{Putida}. These two peptides, Arg^{145} – Arg^{155} and Glu^{213} – Arg^{222} span two residues (Arg^{151} and Arg^{222} respectively) shown to be involved in hydrogen bonding to EF-Tu^{Putida}. Arg^{145} – Arg^{155} showed a modest, yet significant reduction in F_{Mod} (-0.320, p = 0.03, figure 5.7, A) in PPHD•EF-Tu^{Putida} compared to *apo*-PPHD. This peptide forms a major part of the *thumb* loop (figure 5.3, B and 5.8, B), responsible for grasping the unfolded EF-Tu^{Putida} switch I and coordinating it into the PPHD binding groove. The backbone amide nitrogen of Arg^{151} forms a hydrogen bond with the Asp^{87} side-chain carboxylic acid in EF-Tu^{Putida}, located on the N-terminal side of the hydroxylation target, Pro^{90}. The second peptide, Glu^{213} – Arg^{222}, showed a minor but significant drop in F_{Mod} (-0.053, p = 0.005) in PPHD•EF-Tu^{Putida}. This peptide was labelled at a very low level in *apo*-PPHD, likely due to its position within the binding groove (figure 5.8, C). The accessible C-terminal Arg^{222} is likely to be the location of TDBA insertion (figure 5.8, C). The same residue forms a hydrogen bond to the Glu^{92} carboxylic acid sidechain in EF-Tu^{Putida} switch I from its guanidino group.

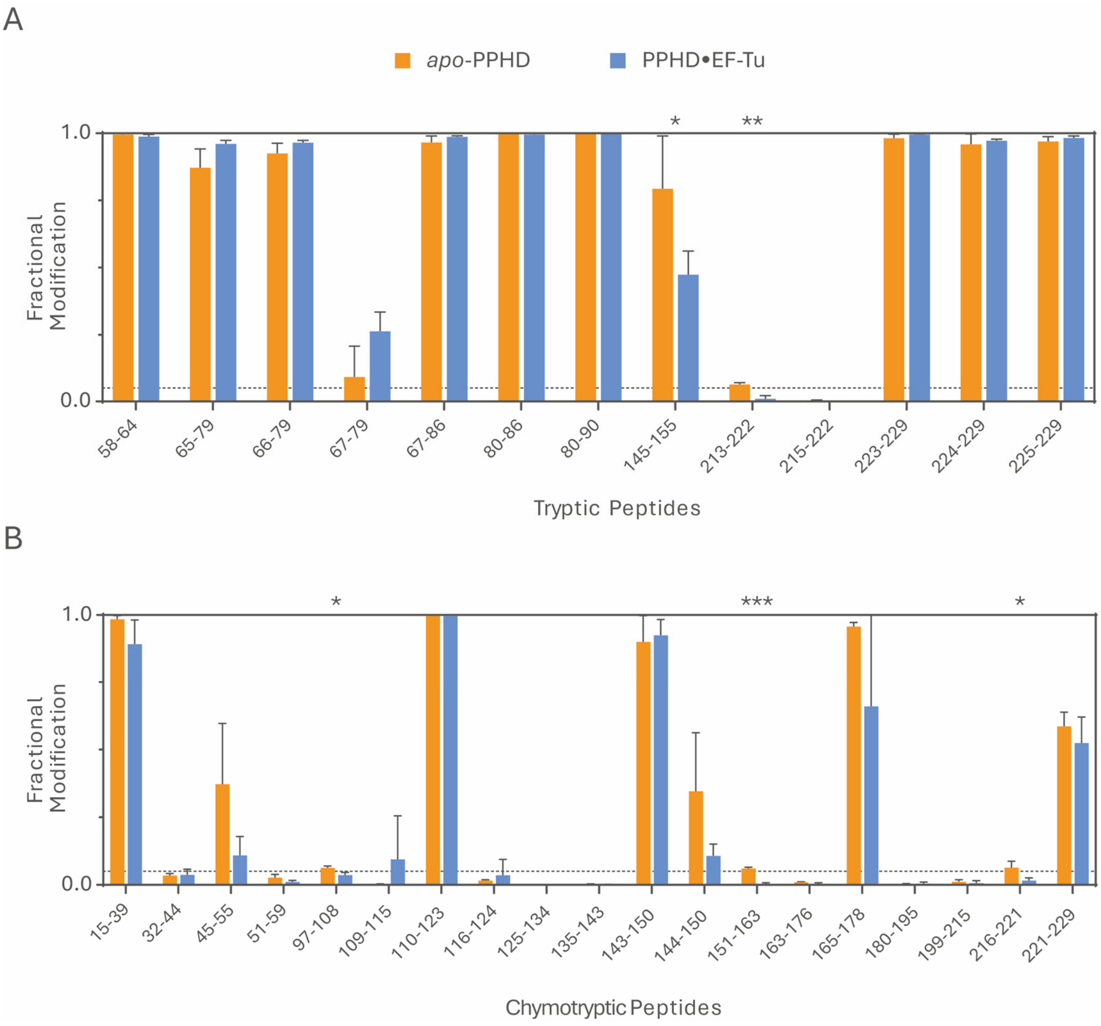


Figure 5.7 | F_{Mod} histograms for *apo*-PPHD (orange) and PPHD•EF-Tu^{Putida} (blue). A | Tryptic PPHD peptides. B | Chymotryptic PPHD peptides. Data is an average of 4 replicates, error bars represent standard deviation. Stars indicate significant reductions in F_{Mod}, compared to *apo*-PPHD. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Dotted line indicates an F_{Mod} of 0.05.

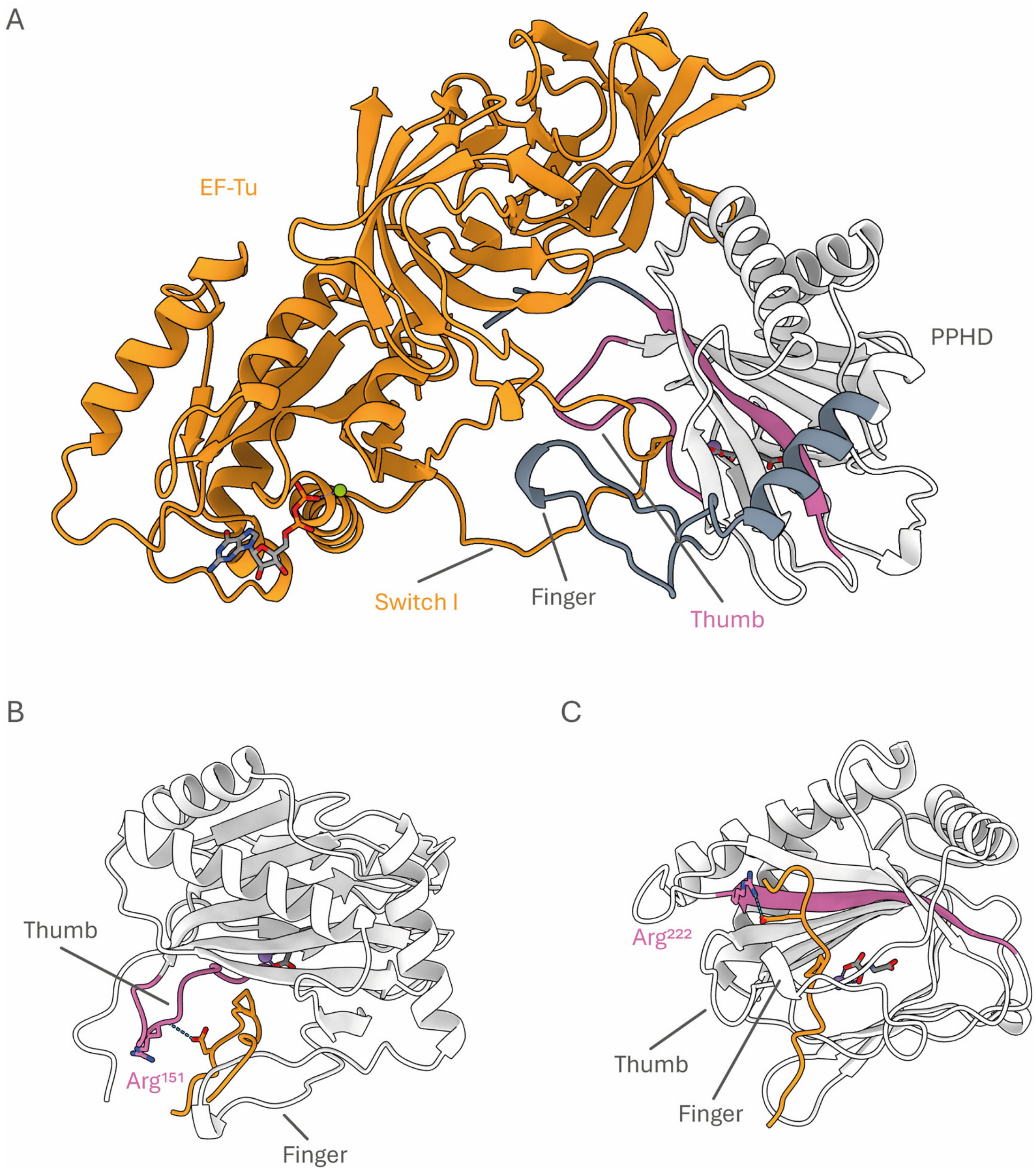


Figure 5.8 | Tryptic PPHD•EF-Tu^{Putida} F_{Mod} data mapped to a homology model of the complex. A | PPHD peptides detected by HPLC-MS are shown in grey, those showing significant reductions in F_{Mod} in PPHD•EF-Tu^{Putida} are coloured pink. EF-Tu is coloured orange. B | Zoom of PPHD with EF-Tu removed, bar a small part of switch I. Arg^{145} – Arg^{155} is coloured pink, bonding Arg^{151} highlighted. C | Glu^{213} – Arg^{222} is coloured pink, bonding Arg^{222} highlighted. GDP and α-ketoglutarate are shown as sticks, Mg^{2+} and Mn^{2+} are shown as green and purple spheres respectively.

Chymotryptic digestion post-carbene footprinting yielded better, but still disappointing sequence coverage of labelled peptides. Overall, 19 peptides were identified, representing 75.98 % of the PPHD sequence (figure 5.6, B). Three peptides (Ile^{97} – Tyr^{108}, Arg^{151} – Tyr^{163} and Ser^{216} – Phe^{221}) showed significant F_{Mod} reductions in PPHD•EF-Tu^{Putida} compared to *apo*-PPHD (figure 5.7, B and figure 5.9, A). The first peptide, Ile^{97} – Tyr^{108} was poorly labelled in *apo*-PPHD (F_{Mod} of 0.061) but did show a minor, yet significant reduction in F_{Mod} (-0.027, $p = 0.02$) in PPHD•EF-Tu^{Putida}. This peptide is located proximal to the switch I binding groove (figure 5.9, B), and the carboxylic acid side-chain of Asp^{98} forms a hydrogen bond with the side-chain amino group of EF-Tu^{Putida} Lys^{93}. Unfortunately, this peptide is also proximal to the large, unstructured N-terminal tag, so it is not possible to rule out that masking effects may be induced by the tag. From interrogation of the crystal structure, this peptide is likely not masked by direct occlusion from switch I, even though it contains a bonding residue.

The second peptide, Arg^{151} – Try^{163}, is complementary to the tryptic peptide Arg^{145} – Arg^{155}, covering the same *thumb* loop and the bonding residue, Arg^{151}. Here, Arg^{151} – Try^{163} shows a minor, but significant reduction in F_{Mod} (-0.060, $p = 0.03$) in PPHD•EF-Tu^{Putida}. The additional C-terminal residues of Arg^{151} – Try^{163} over Arg^{145} – Arg^{155} are found buried in the binding groove, and as a result, much less accessible to TDBA (figure 5.9, C). The loss of the more exposed N-terminal Arg^{145} – Phe^{150} would explain the rather large differences in labelling between these two peptides.

The final peptide, Ser^{216} – Phe^{221}, is complementary to the tryptic peptide Glu^{213} – Arg^{222}, but does not cover the hydrogen bonding C-terminal Arg^{222}. However, Ser^{216} – Phe^{221} forms a portion of the switch I binding groove (figure 5.9, D), and shows a significant reduction in F_{Mod} (-0.048, $p = 0.03$) in PPHD•EF-Tu^{Putida}. Carbene footprinting and tryptic/chymotryptic digestion captured masking associated with binding of switch I to the *finger-thumb* binding groove but did not identify masking associated with PPHD contacting domain II of EF-Tu^{Putida}, despite peptide coverage of the contacting residues.

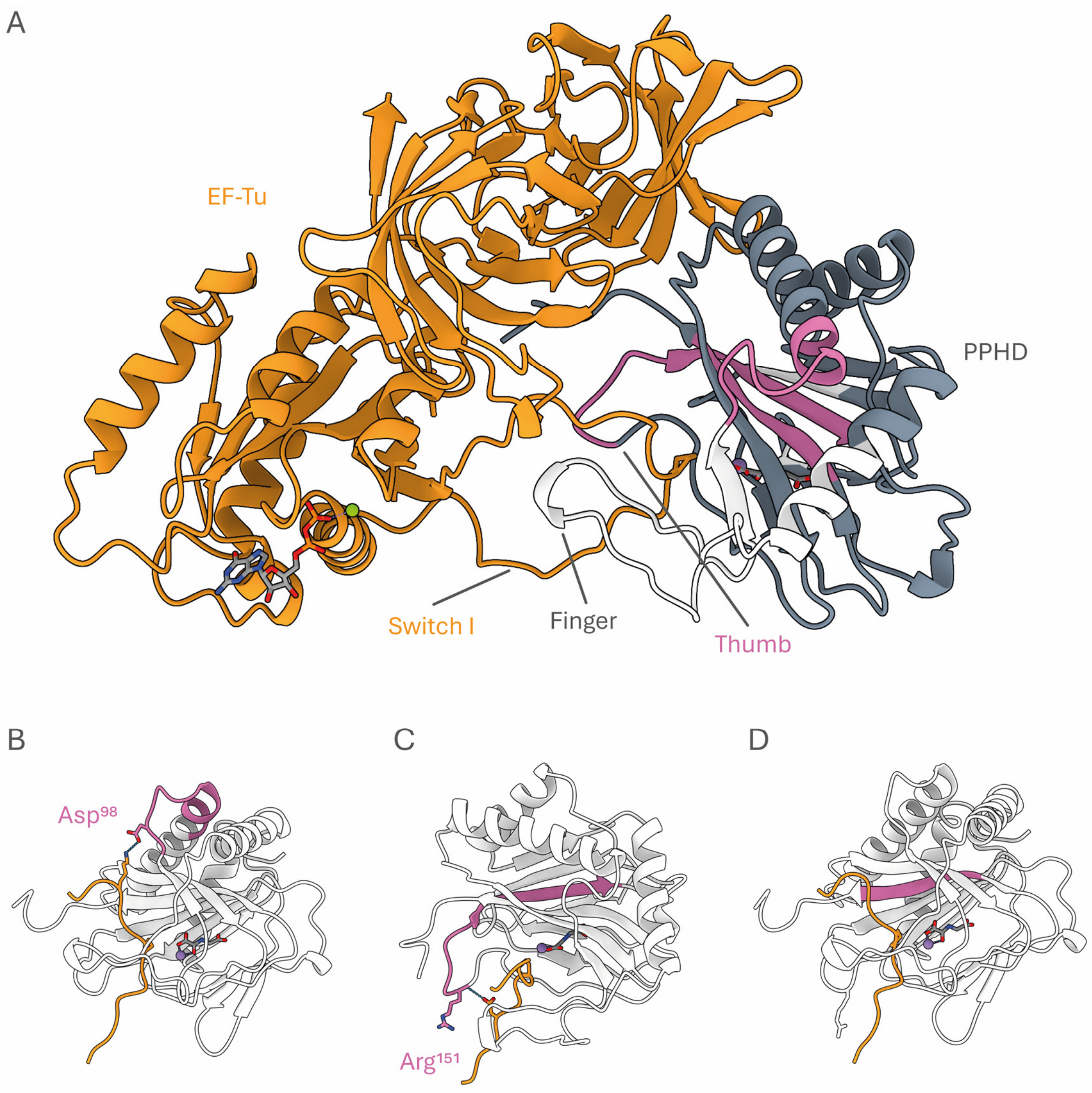


Figure 5.9 | Chymotryptic PPHD•EF-Tu$^{\text{Putida}}$ F_{Mod} data mapped to a homology model of the complex. A | PPHD peptides detected by HPLC-MS are shown in grey, those showing significant reductions in F_{Mod} in PPHD•EF-Tu$^{\text{Putida}}$ are coloured pink. EF-Tu is coloured orange. B | Zoom of PPHD with EF-Tu removed, bar a small part of switch I. Ile^{97} – Tyr^{108} is coloured pink, bonding Asp^{98} highlighted. C | Arg^{151} – Try^{163} is coloured pink, bonding Arg^{151} highlighted. D | Ser^{216} – Phe^{221} is coloured pink. GDP and α-ketoglutarate are shown as sticks, Mg^{2+} and Mn^{2+} are shown as green and purple spheres respectively.

5.3.2.2. Carbene Footprinting of EF-Tu^{Putida}

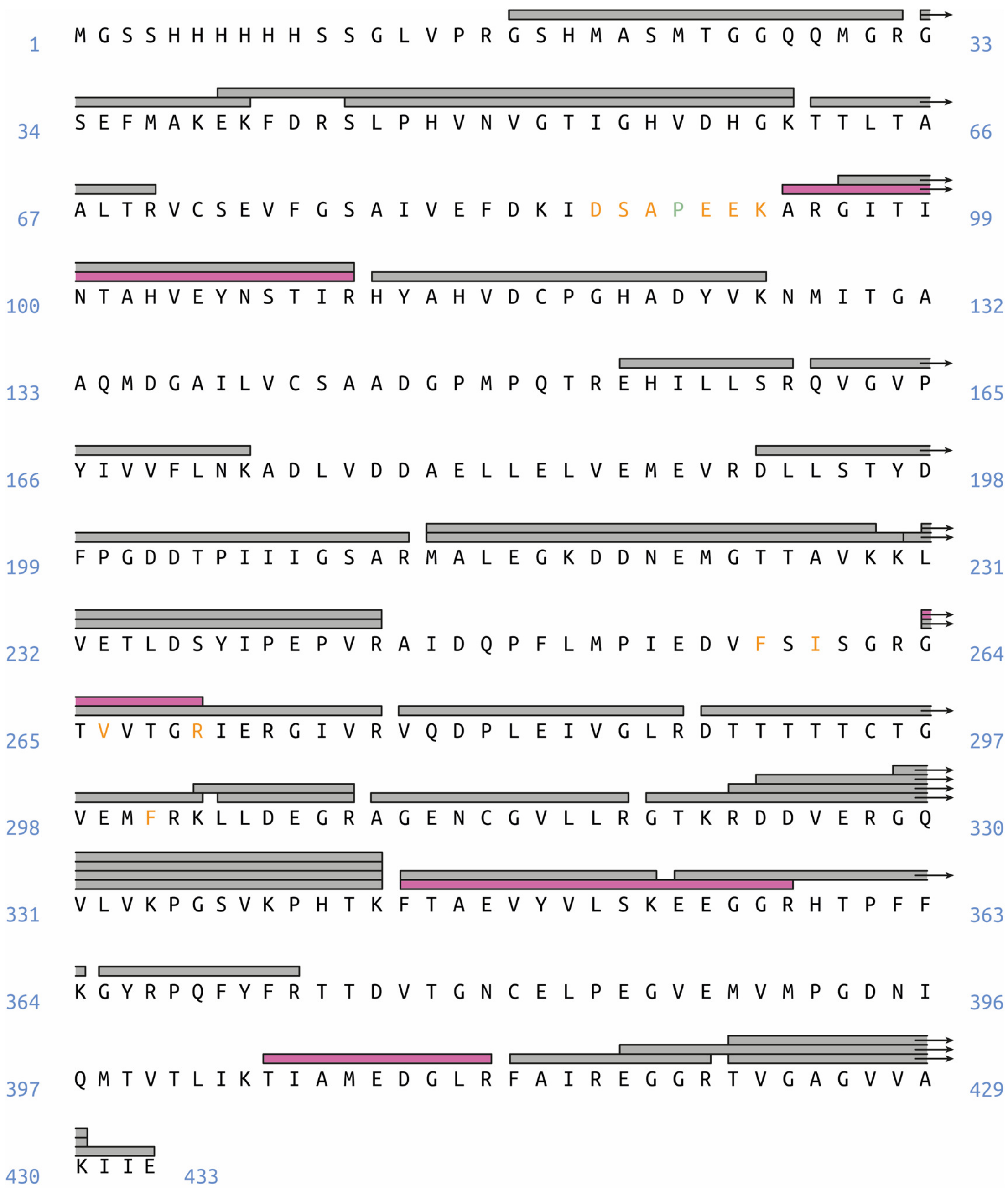


Figure 5.10 | Tryptic labelled peptide sequence coverage of EF-Tu^{Putida} post-carbene footprinting. Peptides coloured grey showed no significant reductions in F_{Mod} in EF-Tu^{Putida}•PPHD *vs* *apo*-EF-Tu^{Putida}, and those coloured pink did show significant reductions. Residues coloured orange are those involved in non-bonding contacts and hydrogen bonding with EF-Tu^{Putida}, the target Pro^{90} is coloured green.

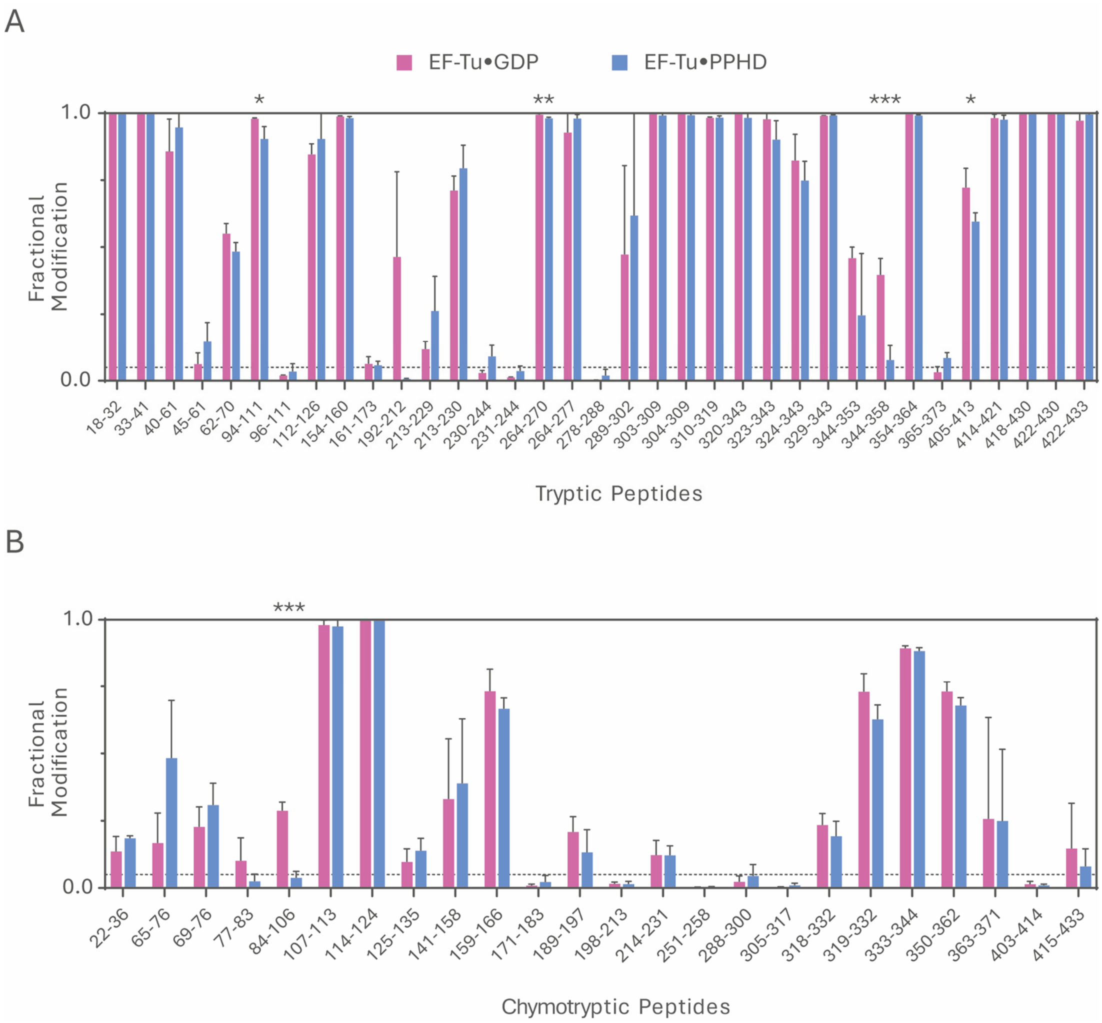


Figure 5.11 | F_{Mod} histograms for EF-Tu$^{\text{Putida}}$•GDP (pink) and EF-Tu$^{\text{Putida}}$•PPHD (blue). A | Tryptic PPHD peptides. B | Chymotryptic PPHD peptides. Data is an average of four replicates, error bars represent standard deviation. Stars indicate significant reductions in F_{Mod}, compared to *apo*-PPHD. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Dotted line indicates an F_{Mod} of 0.05.

EF-Tu$^{\text{Putida}}$, as typical amongst EF-Tu in general, is highly conserved between species. EF-Tu$^{\text{Putida}}$ shares a 78 % identity with *A. baumannii* EF-Tu (as computed by the Expasy SIM and LANVIEW tools).[289] Because of this, tryptic digestion of EF-Tu$^{\text{Putida}}$ produced acceptable sequence coverage of labelled peptides, similar to that seen for *A. baumannii* EF-Tu (62.82 % coverage from 35 peptides, figure 5.10). Although this is significantly less coverage than that seen in figure 5.5, C, the detected labelled peptides are found across all three EF-Tu domains, providing good structural coverage. Four peptides (Ala^{94} – Arg^{111}, Gly^{264} – Arg^{270}, Phe^{344} – Arg^{358} and Thr^{405} – Arg^{413})

displayed significant reductions in F_{Mod} in EF-Tu^{Putida}•PPHD compared to *apo*-EF-Tu^{Putida} (figure 5.11, A and 5.12, A and B).

The first peptide, Ala^{94} – Arg^{111}, covers the C-terminal region of switch I alongside β-strand *c* (figure 5.12, C). Although this peptide does not encompass residues forming hydrogen bonds with PPHD, it does show a small but significant reduction in F_{Mod} (-0.075, $p = 0.04$) in EF-Tu^{Putida}•PPHD. This region of switch I and strand *c* are not found directly within the PPHD binding groove, however they do fall between the *finger* loop of PPHD and domain I of EF-Tu^{Putida}. Due to the complete loss of structure within switch I and its subsequent sequestering by PPHD, it is probable that his masking is imparted by PPDH binding to EF-Tu^{Putida}.

The second masked peptide, Gly^{264} – Arg^{270} shows an incredibly minor yet significant reduction in F_{Mod} (-0.014, $p = 0.005$). This peptide spans strand *c2* in domain II of EF-Tu^{Putida} and encompasses both the hydrophobically interacting Val^{266} and hydrogen bonding Arg^{270} (figure 5.12, D). These two residues form a part of the bonding network between PPHD and domain II of EF-Tu^{Putida}, and as a result, are highly occluded when EF-Tu^{Putida} is bound to PPHD. Interestingly both Thr^{405} – Arg^{413} and Phe^{344} – Arg^{358}, peptides found in domain III of EF-Tu^{Putida} show pronounced reductions in F_{Mod} in EF-Tu^{Putida}•PPHD (-0.319 and -0.127, $p = 0.02$ and $p < 0.001$, figures 5.12, E and F respectively). Thr^{405} – Arg^{413} may be occluded by both the C-terminus of PPHD or the C-terminal region of switch I. As the only comparable structural information comes from the crystal structure of EF-Tu^{Putida}•PPHD, the flexibility and conformational movements of the proteins *in situ* are not known. It is not unreasonable to hypothesise that the inherent flexibility of the disordered switch I is responsible for the masking of Thr^{405} – Arg^{413} due to its proximity. In contrast, Phe^{344} – Arg^{358} is buried within the central pore of EF-Tu^{Putida}, with the C-terminal Arg^{358} only 4.82 Å in distance from the C-terminal phenylalanine of PPHD (figure 5.12, F). Although no MS^2 sub-peptide analyses were performed, due to the inherent acidity of TDBA and its propensity to label basic residues, it is possible that Arg^{358} would be labelled in *apo*-EF-Tu^{Putida}.

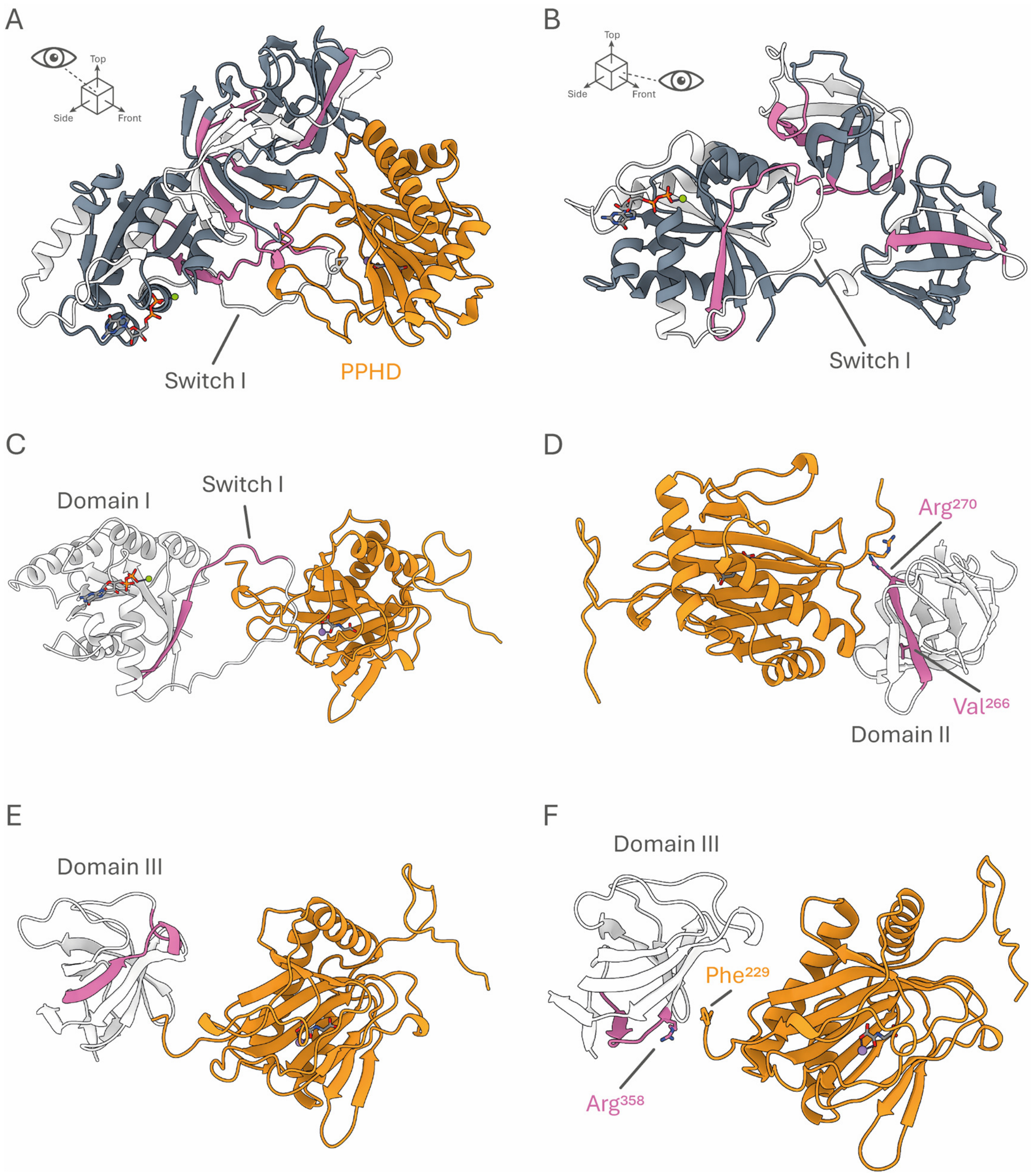
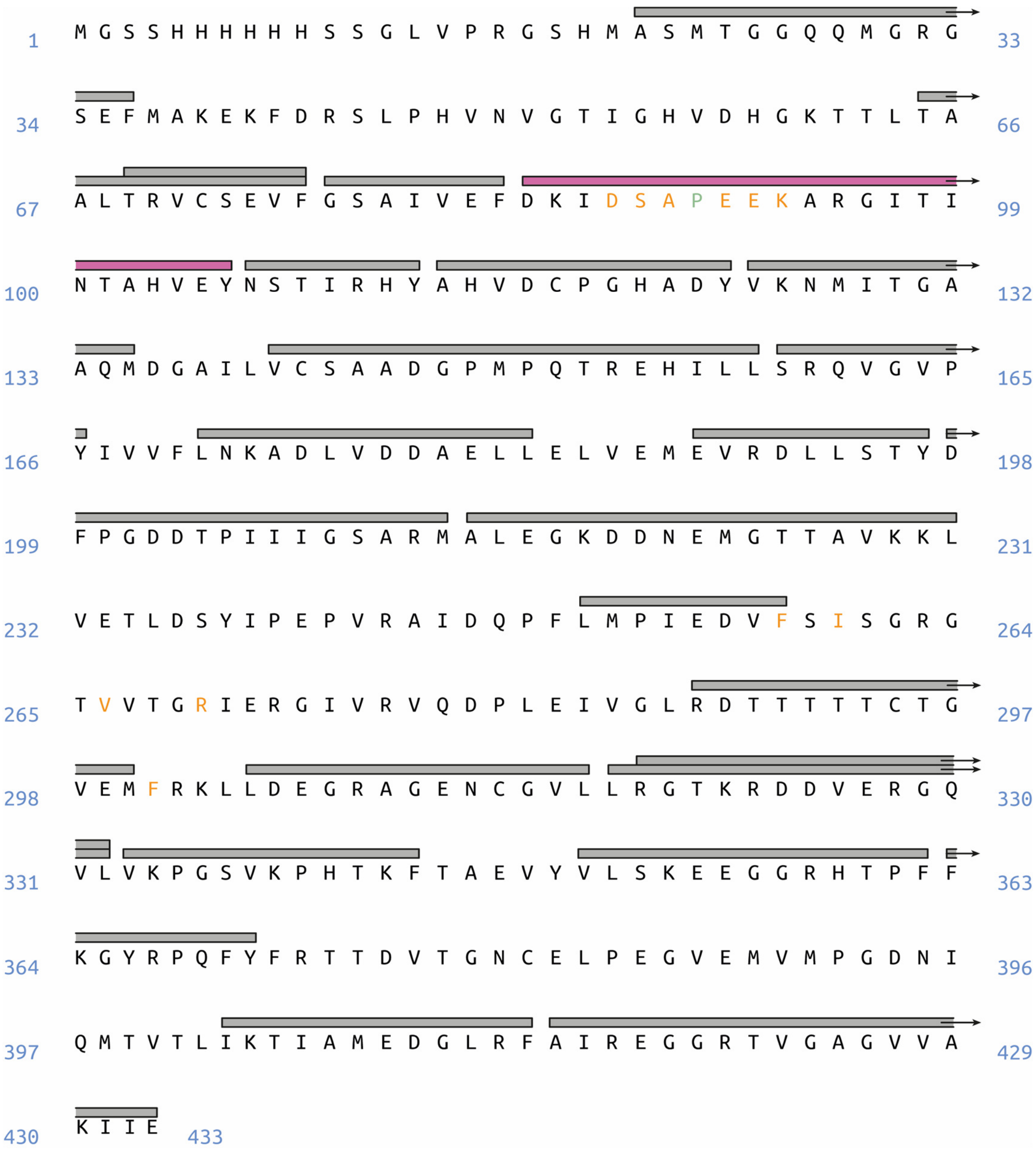


Figure 5.12 | Tryptic EF-Tu^{Putida}•PPHD F_{Mod} data mapped to a homology model of the complex. A | EF-Tu^{Putida} peptides detected by HPLC-MS are shown in grey, those showing significant reductions in F_{Mod} in PPHD•EF-Tu^{Putida} are coloured pink. PPHD is coloured orange. B | Front view of A with PPHD removed for clarity. C | Zoom of EF-Tu^{Putida}•PPHD showing only domain I of EF-Tu^{Putida}. Ala^{94} – Arg^{111} is coloured pink, D | Zoom of EF-Tu^{Putida}•PPHD showing only domain II of EF-Tu^{Putida}. Gly^{264} – Arg^{270} is coloured pink, bonding Arg^{270} and contacting Val^{266} highlighted. E | Zoom of EF-Tu^{Putida}•PPHD showing only domain III of EF-Tu^{Putida}. Thr^{405} – Arg^{413} is coloured pink. F | Zoom of EF-Tu^{Putida}•PPHD showing only domain III of EF-Tu^{Putida}. Phe^{344} – Arg^{358} is coloured pink. GDP and α-ketoglutarate are shown as sticks, Mg^{2+} and Mn^{2+} are shown as green and purple spheres respectively.

The binding of PPHD, specifically the C-terminus, would likely impart a masking effect on that residue. It is worth noting that the correctly cleaved peptide, Phe^{344} – Lys^{353} is not significantly masked in EF-Tu^{Putida}•PPHD, despite being encompassed by Phe^{344} – Arg^{358}. However, Phe^{344} – Lys^{353} does show a moderate non-significant reduction in F_{Mod} (-0.213). Labelling of this peptide in EF-Tu^{Putida}•PPHD was highly variable, with two replicates showing much greater F_{Mod}. This may be indicative of low levels of complex formation, or at least low levels of PPHD binding to EF-Tu^{Putida} outside of switch I.



5.13 | Labelled peptide sequence coverage of EF-Tu^{Putida} post-carbene footprinting. Main figure legend overleaf.

Figure 5.13 | Cont. Chymotryptic labelled peptide sequence coverage of EF-Tu^{Putida} post-carbene footprinting. Peptides coloured grey showed no significant reductions in F_{Mod} in EF-Tu^{Putida}•PPHD *vs* *apo*-EF-Tu^{Putida}, and those coloured pink did show significant reductions. Residues coloured orange are those involved in non-bonding contacts and hydrogen bonding with EF-Tu^{Putida}, the target Pro^{90} is coloured green.

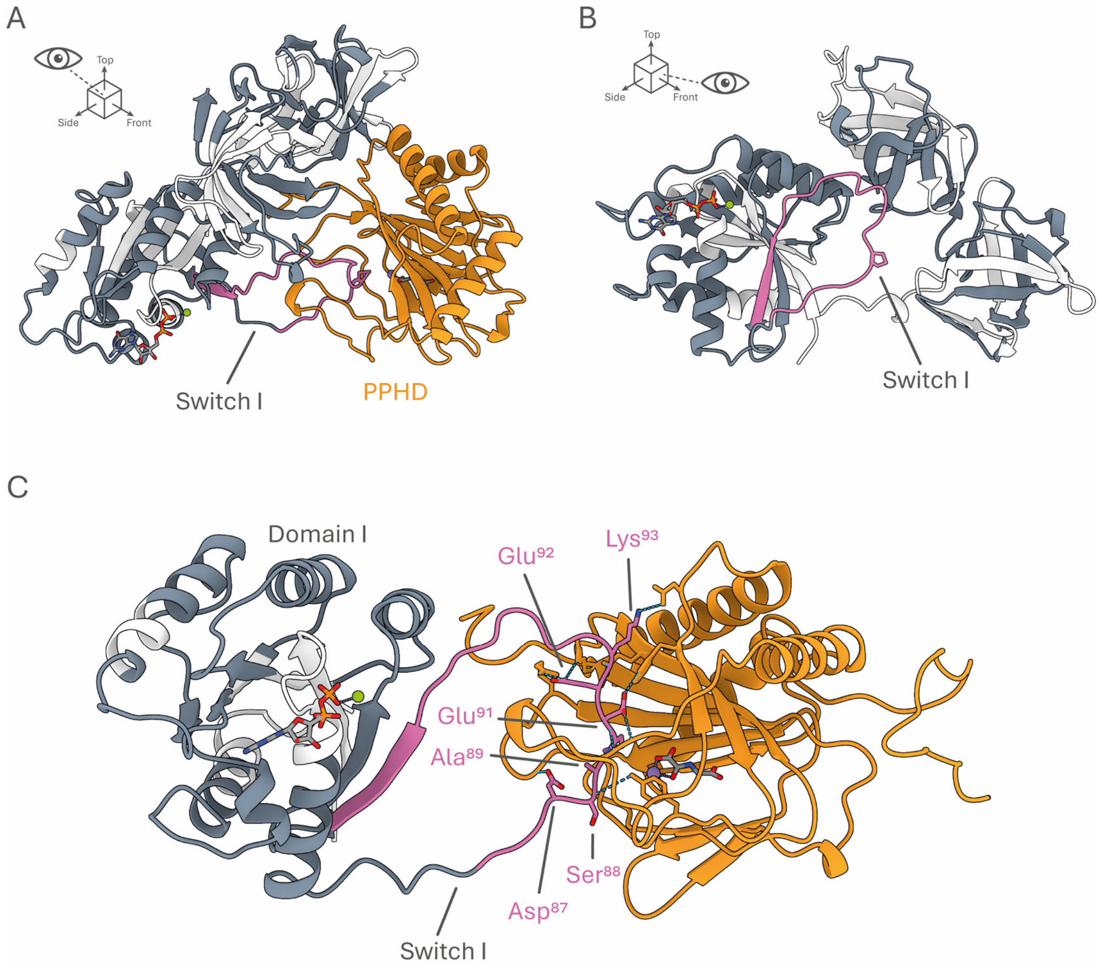


Figure 5.14 | Chymotryptic EF-Tu^{Putida}•PPHD F_{Mod} data mapped to a homology model of the complex. A | EF-Tu^{Putida} peptides detected by HPLC-MS are shown in grey, those showing significant reductions in F_{Mod} in PPHD•EF-Tu^{Putida} are coloured pink. PPHD is coloured orange. B | Front view of A with PPHD removed for clarity. C | Zoom of EF-Tu^{Putida}•PPHD showing only domain I of EF-Tu^{Putida}. Asp^{84} – Tyr^{106} is coloured pink, bonding Asp^{87}, Ser^{88}, Ala^{89}, Glu^{91}, Glu^{92} Lys^{93} are highlighted. GDP and α-ketoglutarate are shown as sticks, Mg^{2+} and Mn^{2+} are shown as green and purple spheres respectively.

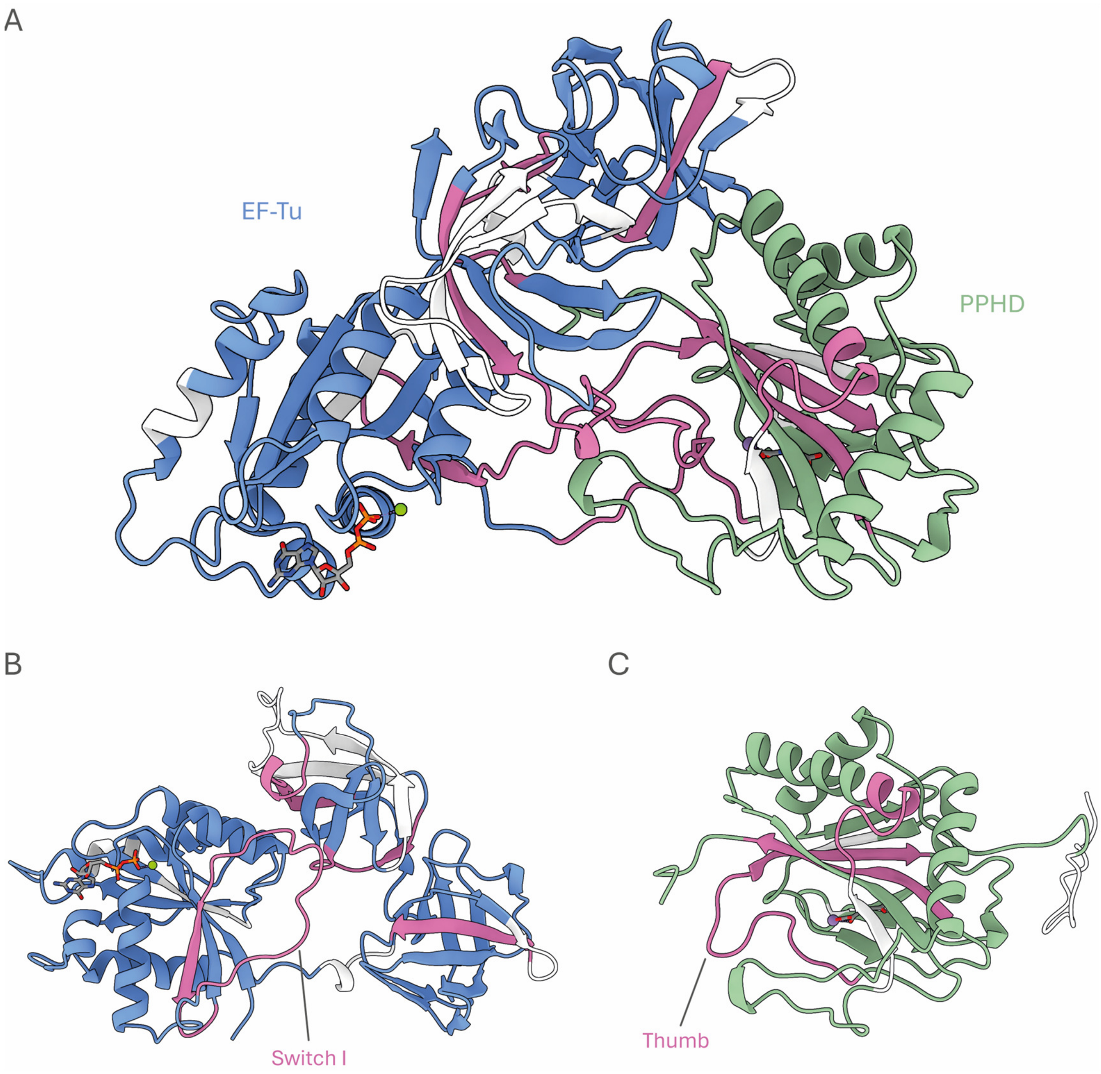


Figure 5.15 | Combined EF-Tu^{Putida}•PPHD F_{Mod} data from both proteases mapped to a homology model of the complex. A | All peptides detected for EF-Tu^{Putida} (blue) and PPHD (green). Peptides showing significant reductions in F_{Mod} in *holo vs apo* proteins are coloured pink. Peptides not detected are coloured white. B | Isolated EF-Tu^{Putida} coloured as in A. C | Isolated PPHD coloured as in A. GDP and α-ketoglutarate are shown as sticks, Mg^{2+} and Mn^{2+} are shown as green and purple spheres respectively.

Chymotryptic digestion of labelled EF-Tu^{Putida} yielded comparable sequence coverage to trypsin (65. 13 % sequence coverage from 24 individual peptides, figure 5.13). Surprisingly, only one peptide, Asp^{84} – Tyr^{106} showed a significant reduction in F_{Mod} in EF-Tu^{Putida}•PPHD compared to *apo*-EF-Tu^{Putida} (-0.288, $p < 0.001$, figure 5.11, B and figure 5.13). Asp^{84} – Tyr^{106} is complementary to the tryptic peptide Ala^{94} – Arg^{111}, encompassing the majority of the tryptic residues, alongside greater N-terminal coverage of switch I (figure 5.14, A and B). Asp^{84} – Tyr^{106} covers all the residues involved in the hydrogen bonding of switch I to PPHD, alongside the target Pro^{90}. As mentioned prior, EF-Tu^{Putida} Asp^{87} binds to PPHD Arg^{151}, the backbone carbonyl of EF-Tu^{Putida} Ser^{88} binds the hydroxyl group in PPHD Tyr^{143} and the backbone carbonyl of EF-Tu^{Putida} Ala^{89} binds the backbone amide nitrogen of PPHD Gly^{78}. C-terminal of Pro^{90}, Glu^{91} forms three hydrogen bonds with Gly^{76} and Trp^{96} in PPHD, as mentioned Glu^{92} forms two hydrogen bonds from its carboxylic acid group to Arg^{155} and Arg^{122} in PPHD. Finally, as mentioned above, EF-Tu^{Putida} Lys^{93} binds to PPHD Asp^{98}. Masking of this peptide is highly indicative of switch I being bound by PPHD. Conversely, the chymotryptic peptide Ile^{403} – Phe^{414} does not show a similar reduction in F_{Mod} to its complementary tryptic peptide, Thr^{405} – Arg^{413}. The additional N-terminal Ile^{403} and Lys^{404} are highly exposed on the back of EF-Tu^{Putida} domain III. As TDBA is liable to preferentially interact with basic and hydrophobic residues, it is probable that these residues would be labelled, and as they are not directly masked by PPHD, would remain labelled in EF-Tu^{Putida}•PPHD. The combined tryptic and chymotryptic footprinting data is shown in figure 5.15.

5.4. Conclusions

Native mass spectrometry showed successful formation of the EF-Tu^{Putida}•PPHD complex, albeit at relatively low abundance compared to the individual constituents. The complex must be stable in solution to enable the production of a crystal structure, but it is possible that this stability does not translate into the gas-phase. It is also possible that complex formation was unfavoured under the conditions used for native mass spectrometry. On the other hand, carbene footprinting of EF-Tu^{Putida}•PPHD showed masking in each protein correlating mostly to regions expecting to be masked by their binding. Masking was seen on EF-Tu^{Putida} switch I and its binding groove on PPHD. Interestingly, minor masking was seen on EF-Tu^{Putida} peptides located in domain II contacted by PPHD, but no masking was seen on corresponding PPHD peptides. This may suggest that PPHD binds strongly to EF-Tu^{Putida} switch I and more transiently to domain II. Production and isolation of hydroxylated EF-Tu^{Putida} would be a potential future avenue of research. Overall, this shows the capability of carbene footprinting to capture protein-protein interactions of EF-Tu.

Characterisation of EF-Tu Interactions with Enacyloxins by Native Mass Spectrometry, Ion Mobility Spectrometry and Molecular Dynamics

6

6.1. Introduction

6.1.1. Novel Analogues of Enacyloxin IIa

Enacyloxin IIa, as introduced in chapter 4, is the major constituent of the enacyloxin group of elfamycins produced by both *Frateuria* sp. W-315 and *B. ambifaria.*[235,241] Elucidation of the enacyloxin IIa biosynthetic gene cluster in *B. ambifaria* by Mahenthiralingam *et al.* showed enacyloxin IIa to be the product of a hybrid polyketide synthase complex, which was then highly decorated by oxidation, chlorination, carbamoylation and conjugation to DHCCA[241]. This process is facilitated by 22 individual proteins, encoded on an 87.5 kb genomic island. Unfortunately, the conjugated polyenic backbone of enacyloxin IIa conveys an inherent instability, particularly towards ultraviolet light. Polyenes display typically poor stability towards heat, ultraviolet radiation and extremes in pH, an issue that has plagued polyenic-based antifungal compounds since the isolation of amphotericin B in the mid 1950's.[290–292] Conjugated carbon-carbon double bonds strongly absorb visible and ultraviolet light in the 200 – 400 nm range, often resulting in polyenic compounds taking on a yellow colouration. The initial enacyloxin complex (enacyloxins 1 – 8, chapter 4) isolated by Watanabe *et al.* demonstrated this property, with a peak ultraviolet absorbance reported at 370 nm and a reported yellow-green colouration, which is explained by the presence of five carbon-carbon bonds conjugated to the carbonyl group in the ester functionality (table 6.1).[235] Enacyloxin IIa extracts used in the work throughout this thesis all showed a bright yellow colouration (figure 6.1), and unique visible light fluorescence when irradiated during carbene footprinting (chapter 4). The hydrophobic polyene chain is also responsible for the poor aqueous solubility of enacyloxin IIa, a feature shared with clinically used polyenic antifungals. Amphotericin B requires complexing with bile acids or liposomes for intravenous use, and nanoparticulate emulsions for oral administration.[293,294] The contrasting elfamycin GE2270A, introduced in chapter 3, displays a similarly poor aqueous solubility despite not being polyenic. A semi-synthetic derivative, NAI003, is completed phase I clinical trials for use as a topical acne treatment, showing good efficacy at treating *Propionibacterium acnes* as a gel emulsion.[295]

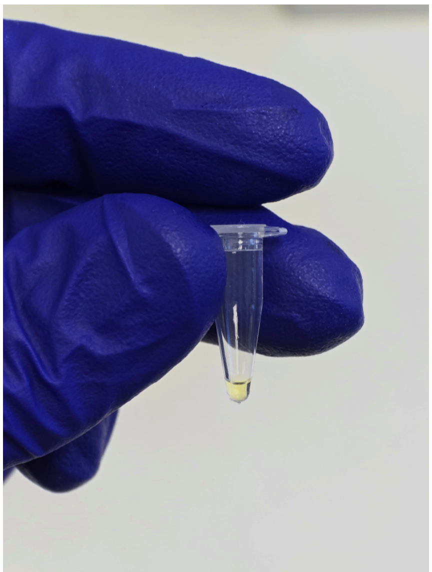
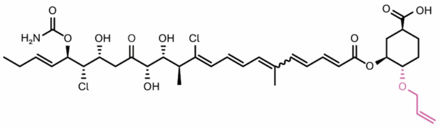
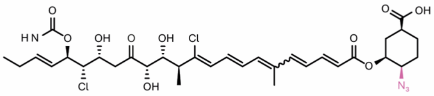
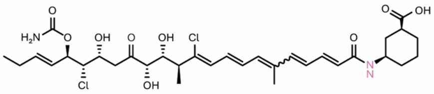
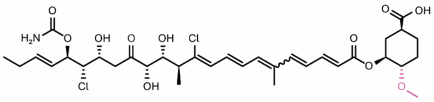
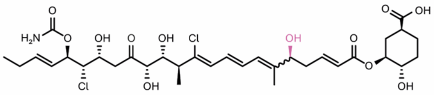
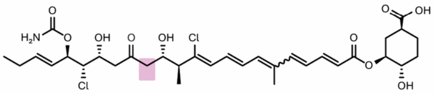
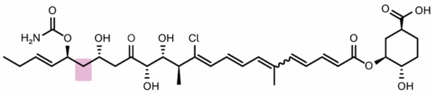
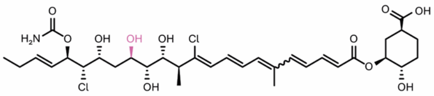
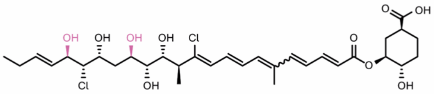
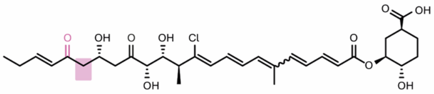
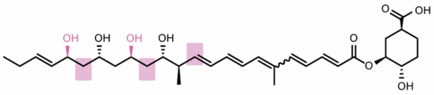


Figure 6.1 | Enacyloxin IIa (10 mM) dissolved in methanol showing yellow colouration.

The group of Professor Greg Challis are actively producing novel analogues of enacyloxin IIa to try to both enhance its solubility and stability, and potentially increase its potency. Two major methods have been employed to produce these analogues, both through targeted editing of the *B. ambifaria* biosynthetic gene cluster. The first is through deletion of bamb_5912 – bamb_5914, the genes responsible for facilitating DHCCA synthesis (chapter 4, figure 4.2). This deletion, paired with supplementation of four different DHCCA analogues, yields enacyloxin IIa analogues with altered *head* groups (enacyloxin IIa1 – IIa4, table 6.1).

Table 6.1 | Enacyloxin IIa analogues, associated genotypes, structures and masses. Atoms coloured pink differ from wild-type enacyloxin IIa.

Enacyloxin	Genotype	Structure	Mass (Da)
IIa1	ΔBamb_5912-14 Allyl DHHCA		741.27
IIa2	ΔBamb_5912-14 Azide DHHCA		726.34
IIa3	ΔBamb_5912-14 Amino DHHCA		684.26
IIa4	ΔBamb_5912-14 O-Methyl DHHCA		716.26
IIa5	Bamb_5920 M9DH		719.25
IIa6	ΔBamb_5927		685.24
IIa7	ΔBamb_5931		667.28
IIa8	ΔBamb_5932		703.25
IIa9	ΔBamb_5930 ΔBamb_5932		660.25
IIa10	ΔBamb_5930 ΔBamb_5931		622.25
IIa11	ΔBamb_5927-32		576.33

The second major method involves the deletion of non-haem iron-dependent oxygenases/chlorinases, enacyloxin oxidase, carbamoyl transferases and $FADH_2$-dependant chlorinases. These have been performed individually, and in multiple combinations, yielding a total of six analogues (enacyloxin IIa6 – IIa11, table 6.1). Finally, a single separate analogue was produced by a single point mutation in in the polyketide subunit-encoding bamb_5920, specifically in the dehydratase domain. This analogue possesses an additional hydroxyl group at the 5th *tail* carbon (enacyloxin IIa5, table 6.1), leading to an interruption of the conjugated chain into a triene and a single alkene. An overview of these analogues can be found in table 6.1.

6.1.2. Collision Induced Unfolding

The concepts behind ion mobility spectrometry (specifically TWIMS-MS) were introduced in detail in chapter 1 and as such, will not be discussed at length here. In brief, the mobility of a gas-phase ion through a pressurised cell is not only dependent on its charge, but also its size and shape. Therefore, the collisional cross section (CCS) of the ion can be measured from IMS arrival distributions.[221,296] Experimentally derived CCS values are often compared to those calculated for atomic coordinates, either from crystal/cryo-EM structures, or resulting from molecular dynamics (MD) simulations. Several methods exist to calculate CCS from atomic coordinates, falling under two broad categories: Monte-Carlo simulations and projection approximation (PA). Monte-Carlo simulations are computationally expensive but produce highly accurate CCS estimations.[297,298] Trajectory method (TJM) and diffuse trajectory method (DTM) are both Monte-Carlo models, simulating gas molecules fired towards a target from a random orientation. Both methods utilise Lennard-Jones forcefields to model long-range interactions between the target and gas molecules, and to also account for momentum and velocity of the gas.[299] They differ by way of DTM also accounting for inelastic collisions between the gas and target.[300] A representative *A. baumannii* EF-Tu TJM projection can be seen in figure 6.2, A, B, as calculated by the software Collidoscope, and visualised using SideFX Houdini (version 20.5.332). In contrast, projection approximation (PA) estimation of CCS is the simplest method, developed in the mid 1920's by Edward Mack. It equates the target CCS to its average orthographically projected "shadow", from random target orientations, which can then be corrected for the inclusion of mobility gas radii.[301] This method of calculation is much quicker, making it ideal for interrogating MD trajectories, which may be composed of many thousands of individual structures. A basic PA-like CCS calculation can be seen in figure 6.2, C, D, produced in SideFX Houdini. In brief, the atomic coordinates of EF-Tu were projected onto an arbitrary 2D plane, and the spherical bounding region calculated. From the surface area of the sphere, the area of the circular hemispherical face can be calculated.

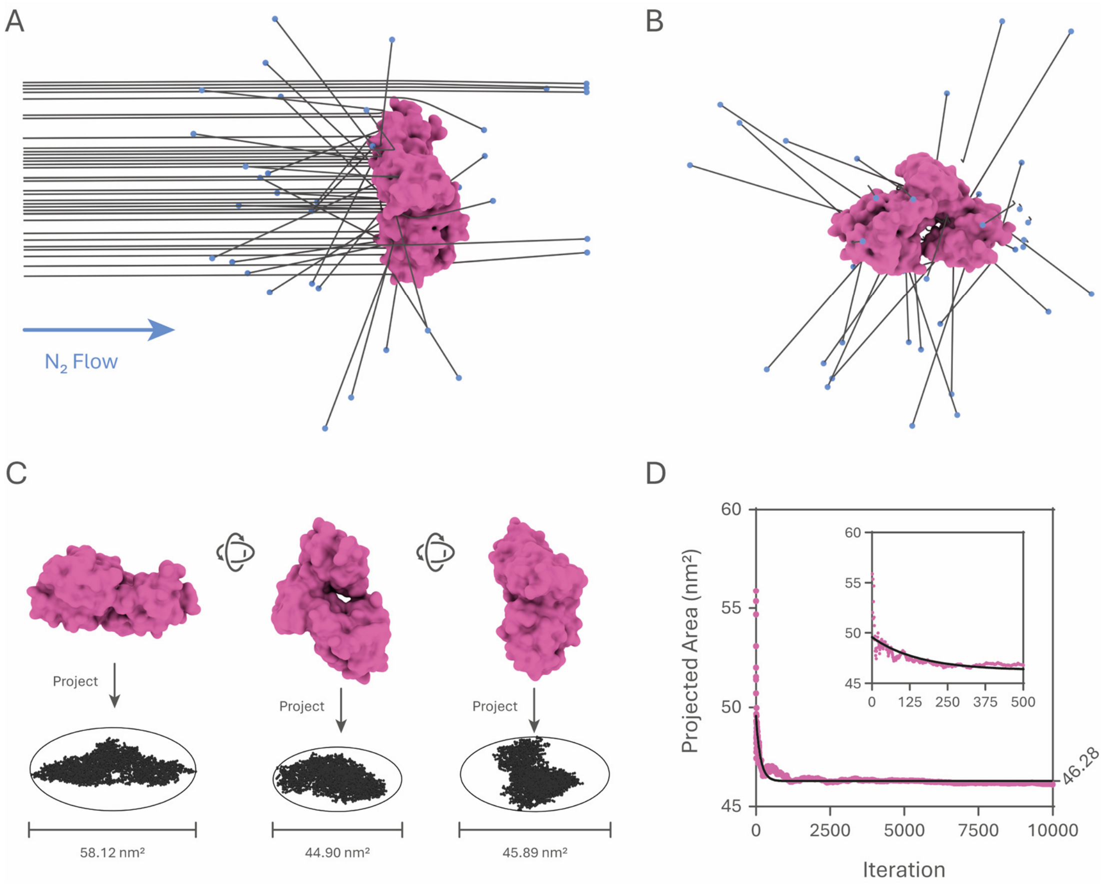


Figure 6.2 | Simulations of TJM and PA CCS calculation methods. Main figure legend overleaf.

Figure 6.2 | Cont. A | EF-Tu CCS calculated using TJM. The trajectories of 50 individual N_2 molecules (blue) out of a total 1000 are shown, which can be seen interacting with EF-Tu (pink). Gas molecules are deflected by electrostatic interactions with the protein. B | Front view of A showing deflection. C | Three representative EF-Tu orientations and their PA-like areas as calculated in Houdini. D | Convergence of an average PA-like area occurs after sampling approximately 500 random EF-Tu orientations (inset) during this 10,000-orientation benchmark.

This process was iterated 10,000 times, taking approximately 3 minutes, randomizing the 3D rotation of EF-Tu each iteration and the running average of the PA area calculated (figure 6.2, D). A stable average PA was reached after approximately 500 iterations. This is an extremely basic implementation of a PA-like calculation, and does not consider collision radii, and in fact overestimates EF-Tu PA CCS by approximately 10 nm^2. But overall is a good demonstration of the concept, allowing rapid calculation and conversion to more accurate representation of CCS for comparison to experimental values.

Collision induced unfolding (CIU) exploits the nature of collisionally activated proteins to adopt transient, partially unfolded conformations. In this case, collisional activation occurs prior to IMS separation. Unfolded or partially unfolded conformations would have a larger CCS than the native structure, and therefore demonstrate an increased drift time through an IMS cell. By subjecting a protein ion to a ramped collision voltage and monitoring the change in CCS, its gas phase stability can be inferred. A more stable ion, such as a *holo*-protein compared to an *apo*-protein, should require higher collision energies to unfold, reflected in a lower CCS at a given collision voltage. The groups of Clemmer and Jarrold were the first to implement this technique to monitor the unfolding of lysozyme, bovine pancreatic trypsin inhibitor and cytochrome C in the gas-phase, through ion activation in the IMS cell.[302,303] The CIU nomenclature was then coined by Hopper and Oldham, whilst showing FK-binding protein, lysozyme, and horse heart myoglobin to be stabilised in the gas-phase through ligand binding.[150] Further CIU research undertaken in the Oldham group has investigated the destabilising effects of mutations in residues implicated in maintaining native conformations. The first, through systematic alanine scanning of an acyl carrier protein, highlighted the importance of polar residues in maintaining a native tertiary structure.[151] A complimentary method was used to probe key ubiquitin residues, through acetylation, ethoxycarbonylation and succinylation of nucleophilic residues, histidine and the free N-terminus respectively, again highlighting the importance of polar residues (specifically lysine) in maintaining the native conformation.[148] In each case, experimental CIU data was complemented with gas-phase MD simulations, which can aid elucidation of the structural perturbations leading to experimentally observed changes in CCS.

6.1.3. Molecular Dynamics

The experimental CIU work presented in this chapter is supplemented by gas-phase MD simulations of *E. coli* EF-Tu. MD simulations are vastly complex, and as such only the basic concepts behind them will be touched upon here, as the underlying mechanics are beyond the scope of this chapter. Specifically, their use for proteinaceous simulations will be explored. MD simulations are, at their core, a computational method to simulate the spatial movements of atoms. During a simulation, the trajectories of each atom are calculated by solving Newton's equations of motion and are governed by the interatomic potentials of interacting atoms. These rules are packaged into *forcefields*, which describe the allowed forces experienced by each atom.[304] The simulations are quantised, and atom coordinates are updated after given timesteps (typically at the femtosecond scale).[305] By running a simulation for a set length of time (typically nano to microseconds), atom coordinates can be saved each timestep and the resulting structures can be interpreted as a *trajectory*, showing the evolution of the atom positions over time. For the purposes of this work, only all-atom MD simulations are considered but the existence of coarse-grain simulations (of various scales) is acknowledged. The application of MD simulations to proteins was pioneered by Levitt and Marshall whose 1975 Nobel Prize winning work used coarse-grain models to predict the folding pathway of pancreatic trypsin inhibitor.[306] Since then MD simulations have become increasingly complex, with all-atom simulations of a solvated influenza virion (approximately 160 million atoms, simulated for 120 ns) and a membrane-bound SARS-CoV-2 spike glycoprotein (approximately 1.7 million atoms, simulated for just short of 5 µs) being two particularly extreme examples.[307,308]

In contrast to solvated simulations, where bulk solvent must be modelled and simulated (either implicitly or explicitly), gas-phase, or *in vacuo*, simulations are much simpler. Coincidently, the field of proteinaceous *in vacuo* simulations blossomed in the late 1990s alongside the pioneering work in CIU by Clemmer and Jarrold, highlighting the intimate relationship between IMS-MS and MD for structural interrogation of gas-phase protein structure.[309–311] Typically these studies involve "denaturation" of protein through increased internal temperatures to mimic CIU which, although not an exact representation, provides helpful insight into tertiary structural stability.[309] Gas-phase simulations have correctly modelled the conformational reduction of protein ions in the gas phase, through both loss of internal hydrophobic interactions and, importantly, the collapse of polar sidechains in the absence of solvent.[312] Recent interesting work by Sever and Konnerman showed the formation of "inside-out" ubiquitin during gas-phase folding. Here, the absence of bulk solvent induced the formation of a compact folded complex with a polar core and exposed hydrophobic surface, in complete contrast to the solvated structure.[313]

6.2. Aims

The work presented here aims to demonstrate the ability to detect binding of novel enacyloxin IIa analogues to EF-Tu directly from crude extracts, thereby removing the need for time-consuming purification steps. I will also explore how binding of enacyloxin IIa to EF-Tu impacts its gas-phase stability through IMS-MS, complemented with all-atom MD simulations. CIU is a well-documented method to study protein, and protein/ligand complexes in the gas phase. EF-Tu, with its large-scale conformation dynamics, and multiple ligand binding sites, represents a complex protein system to interrogate with IMS-MS and CIU.

6.3. Results and Discussion

6.3.1. Detecting Binding of Enacyloxin IIa Analogues to EF-Tu from Crude Extracts

Crude extracts of each enacyloxin IIa analogue were kindly provided by Professor Greg Challis, with the mutagenesis of the enacyloxin IIa PKS and extraction performed by members of his research group. Samples consisted of solid residue resulting from ethyl acetate extracts of *B. ambifaria* culture. Visually, extracts ranged from bright yellow to dark brown and showed various levels of solubility in methanol. For native mass spectrometry analysis, samples were prepared as described in section 2.2.1.6. Briefly, wild-type *A. baumannii* EF-Tu (10 µM) was incubated with 100-fold dilutions of each crude extract (resuspended in methanol), GDPNP and magnesium acetate before buffer exchange (section 2.2.1.2), loading into prepared emitter tips (section 2.2.1.1) and spectra acquired (section 2.2.1.3) on a Synapt G1 HDMS High-Definition mass spectrometer. In the case of analogue enacyloxin IIa5, the crude extract was instead resuspended in 10 % methanol (50 mM ammonium acetate), and EF-Tu incubated with a 10-fold dilution of that. Of the 11 analogues tested, 8 showed successful binding to EF-Tu (mass spectra can be seen in appendix IV). Where binding was observed, the dominant $[M + 12H]^{12+}$ charge state peak corresponding to EF-Tu•GDPNP•*Analogue* was quadrupole isolated and collisionally activated. Liberation of bound analogue allowed for confirmation of its identity. Although this workflow allowed us to quickly identify successful EF-Tu binding analogues, it was not possible to perform quantitative analysis. The unknown concentration of each analogue makes quantitative analyses (*e.g.* K_D determination) impossible.

Table 6.2 | Summary of enacyloxin IIa analogues binding to EF-Tu. The deconvoluted masses shifts and associated errors of EF-Tu EF-Tu•GDPNP•*Analogue* compared to EF-Tu•GDPNP are given alongside the deconvoluted masses and errors of the quadrupole ejected ligands. Error represents deviation from the theoretical masses.

Analogue	Binding	Deconvoluted Mass Shift of Bound EF-Tu (Da)	Error Compared to Theoretical Mass (Da)	Deconvoluted Mass of Ejected Analogue (Da)	Error Compared to Theoretical Mass (Da)
IIa1	✓	741.37	+0.09	741.16	-0.11
IIa2	✓	721.24	-5.00	726.15	-0.09
IIa3	✓	683.98	-0.28	685.15	+0.90
IIa4	✓	706.73	-8.52	738.13	-0.11
IIa5	✗	-	-	-	-
IIa6	✓	685.86	+0.62	685.11	-0.13
IIa7	✓	666.98	-0.30	667.18	-0.09
IIa8	✓	701.84	-1.41	703.15	-0.10
IIa9	✓	657.68	-2.57	660.15	-0.10
IIa10	✗	-	-	-	-
IIa11	✗	-	-	-	-

A summary of the binding of each analogue can be seen in table 6.2. Enacyloxin IIa5, IIa10 and IIa11 did not show any binding to EF-Tu, despite being detectable in their dissolved extracts. In enacyloxin IIa5 hydroxylation at the C5 carbon present in this analogue may cause steric clashes with Glu^{324} (numbered as in *A. baumannii* EF-Tu) located in the enacyloxin IIa binding site (specifically in a partially formed helix localised between *β*-strands *a3* and *b3*), leading to unfavourable binding.

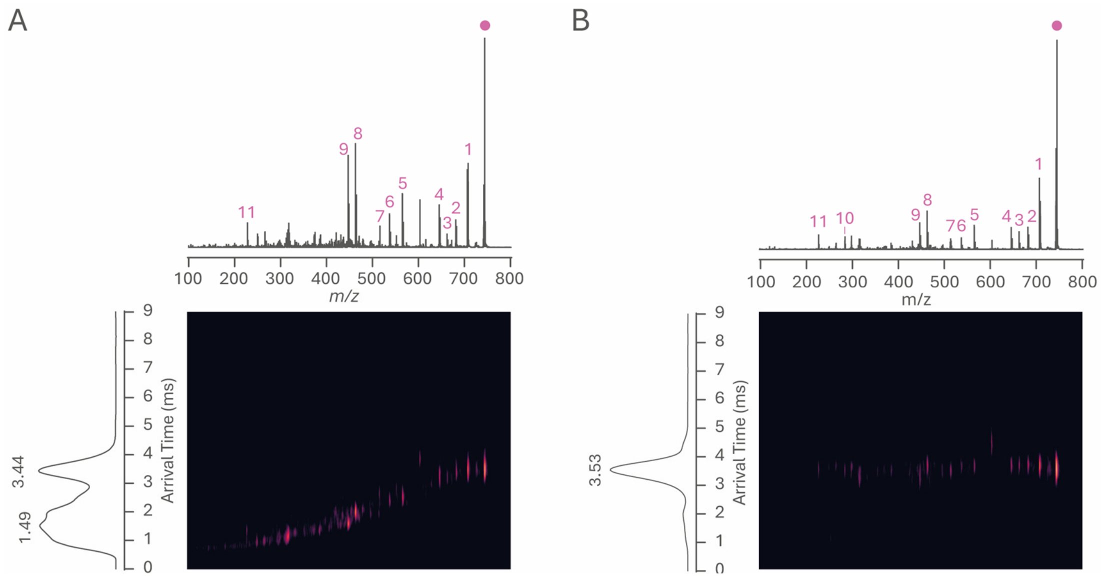


Figure 6.3 | Enacyloxin IIa5 IMS-MS^2 fragmentation. A | CID activation in the *trap* cell prior to IMS separation. B | CID activation in the *transfer* cell post IMS separation. Enacyloxin IIa5 parent ion is represented by the pink circle, fragments numbered as in table 6.3.

Table 6.3 | Tentative fragment ions produced from CID activation of enacyloxin IIa5, corresponding to the spectrum in figure 6.3, A. All fragments originate from sequential losses from the enacyloxin IIa5 *tail*.

Fragment	*m/z*	Loss from Parent Ion (*m/z*)	Loss	Comparison to Equivalent Wild-Type ENX Fragment (*m/z*)
1	706.16	-35.98	HCl	+17.70
2	681.13	-61.08	CH_3NO_2	+17.86
3	662.17	-79.97	HCl + C_3H_8	+17.85
4	645.16	-96.98	HCl + CH_3NO_2	+17.86
5	565.10	-177.04	HCl + $C_7H_{11}NO_2$	+17.89
6	537.11	-205.03	$C_8H_{12}ClNO_3$	+17.89
7	513.07	-228.07	?	-
8	463.08	-279.06	$C_{11}H_{18}ClNO_5$	+17.90
9	447.00	-295.14	$C_{11}H_{18}ClNO_6$	-
10	285.14	-435.11	$C_{19}H_{27}Cl_2NO_6$	+19.01
11	228.13	-492.12	$C_{22}H_{32}Cl_2NO_7$	-1.05 *

** This is a comparison to a theoretical wild-type enacyloxin IIa fragment as this was not detected.*

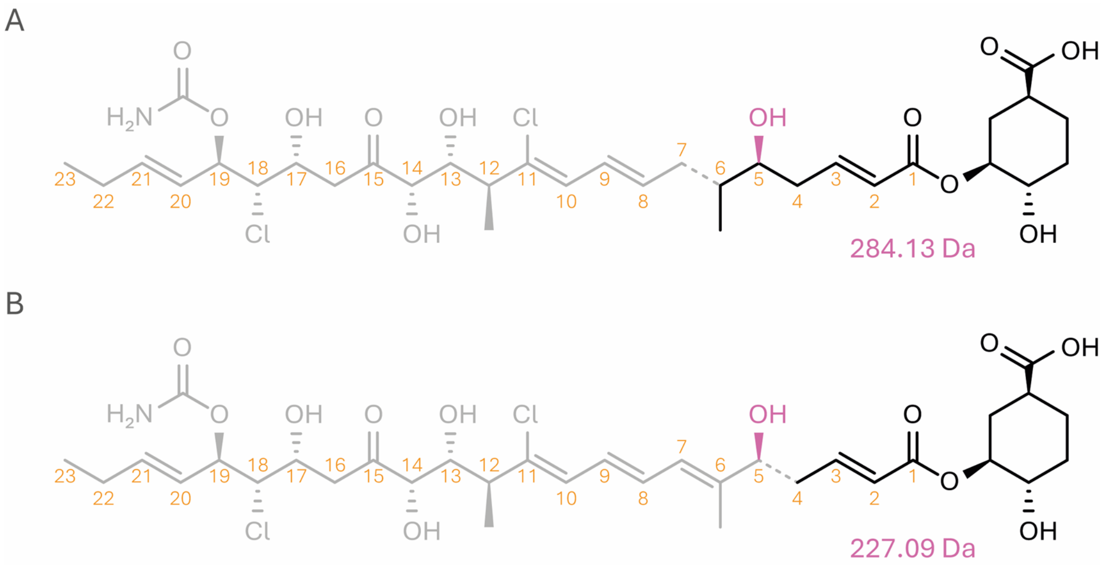


Figure 6.4 | Enacyloxin IIa5 fragments localising hydroxylation to the C5 carbon. A | Fragmentation leading to formation of the $[M + H]^{1+}$ ion at 285.14 *m/z* (fragment 10, table 6.3). B | Fragmentation leading to formation of the $[M + H]^{1+}$ ion at 228.13 *m/z* (fragment 11, table 6.3).

Interestingly, the other two non-binding analogues (enacyloxin IIa10 and IIa11) are both missing the terminal carbamoyl and chlorine (table 6.1). These two mutations are not deleterious to binding individually, as enacyloxin IIa7 and IIa9 bind successfully (table 6.2, appendix III). It is known that enacyloxin IIa shows weaker binding to EF-Tu compared to kirromycin predominantly as a result of its shorter tail.[251] The carbamoyl amine binds to Ala[102] in EF-Tu, and the bulky chlorine forces the binding cleft to accommodate enacyloxin IIa. It is possible that the combined loss of bonding in the already weakly binding tail, alongside the potentially less accommodating blinding cleft between domains I and III results in these analogues being unable to bind EF-Tu.

As enacyloxin IIa5 is the only analogue with a reduced polyene, it was subject to MS^2 analysis to confirm localisation of the hydroxyl group to the C5 carbon (figure 6.3, A and B, 2D plots generated using ORIGAMI).[314] Targeted CID fragmentation of the sodiated ion (*m/z* of 742.24) took place in the *transfer* cell of the Synapt G1 HDMS

High-Definition mass spectrometer (figure 6.3, B) after IMS separation. This leads to fragment ions possessing the same arrival time distributions as the precursor ion, allowing easy separation of desired fragment ions from other contaminating ions. This contrasts with CID fragmentation in the *trap* cell prior to IMS, where the resulting fragments are then separated in the IMS cell (figure 6.3, A). Tentative assignments of the fragment ions can be seen in table 6.3. All comparable fragments carried, on average, an additional +17.99 ± 0.38 Da over those seen for wild-type enacyloxin IIa (chapter 4, table 4.1) corresponding to the additional hydroxyl group at the C5 carbon and an additional hydrogen at the C4 carbon. The smallest detected fragment in both enacyloxin IIa and IIa5 (*m/z* 285.14 in enacyloxin IIa5 and 267.13 in enacyloxin IIa) still carried this mass shift. This ion corresponds to the protonated DHCCA *head* of enacyloxin IIa/IIa5 conjugated to the first 6 carbons of the polyene *tail* (fragmentation of the double bond between carbons 6 and 7, figure 6.4, A). A smaller fragment ion (*m/z* 228.13) was identified from enacyloxin IIa5 corresponding to the protonated *head* conjugated to the first 4 carbons in the polyene *tail* (fragmentation of the double bond between carbons 4 and 5, figure 6.4, B). Importantly the difference between these two ions equates to 57.01 Da, a mass corresponding to $C_3H_5O_3$, indicating hydroxylation of the predicted carbon.

6.3.2. Native Mass Spectrometry and Ion Mobility Spectrometry Analysis of *E. coli* EF-Tu Complexed with Enacyloxin IIa

Samples of *E. coli* EF-Tu were prepared for native-MS as described in section 2.2.1.4. Briefly, *apo*-EF-Tu was incubated with sub-stoichiometric molar concentrations of GDPNP and enacyloxin IIa, yielding samples containing *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX. Spectra were acquired using nESI emitter tips, as prepared in section 2.2.1.1, on a Synapt G1 HDMS High-Definition mass spectrometer. A representative spectrum can be seen in figure 6.5, A. Similarly to *A. baumannii* EF-

Tu, (chapters 3 and 4), native nESI-MS EF-Tu spectra typically show three charge states: $[M + 11H]^{11+}$, $[M + 12H]^{12+}$, and $[M + 13H]^{13+}$, with $[M + 12H]^{12+}$ being the principal charge state. Within each charge state three individual peaks were seen, corresponding to *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX.

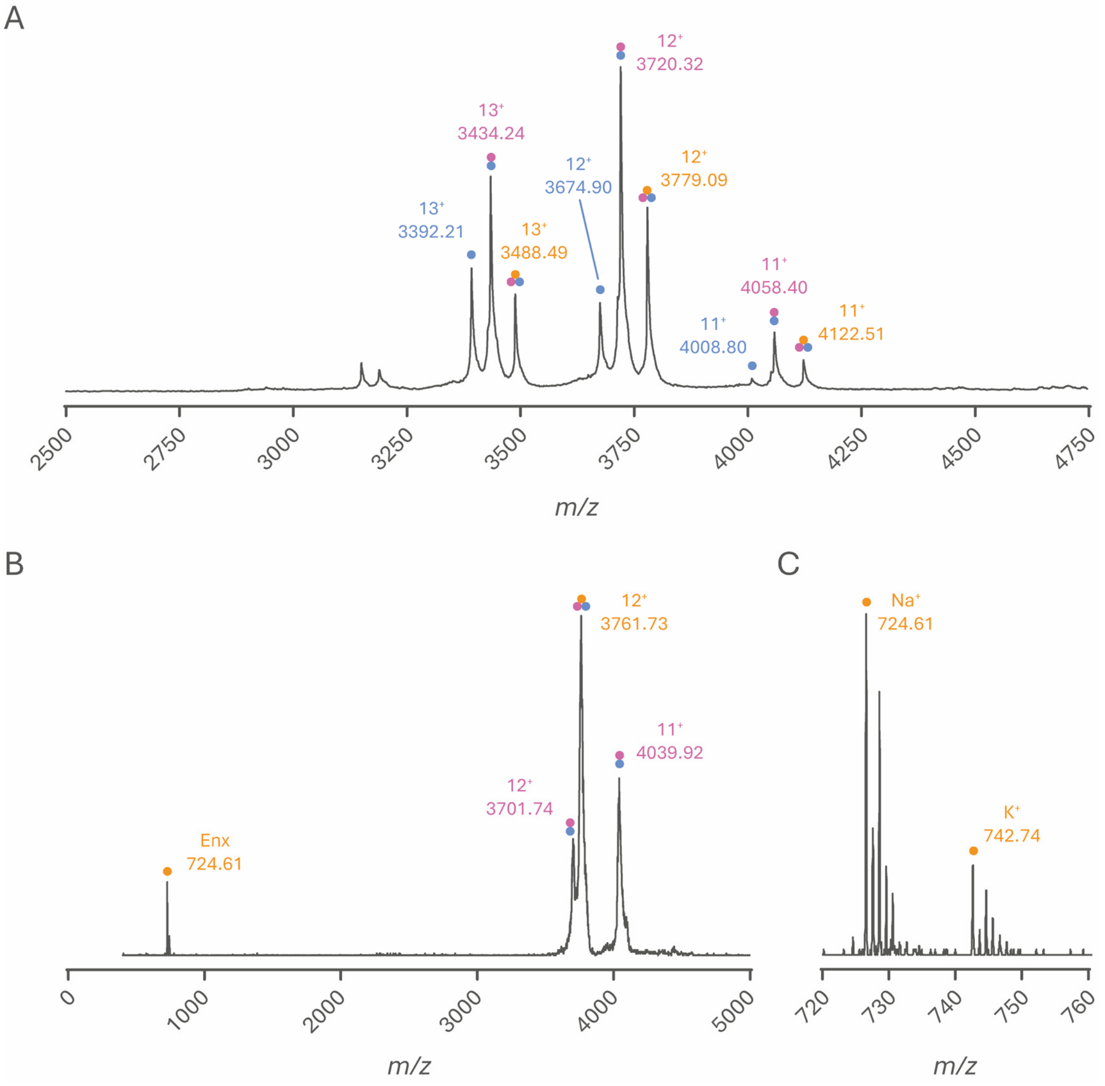


Figure 6.5 | Representative native mass spectra of *E. coli* EF-Tu. *Apo*-EF-Tu (blue) bound to GDPNP (pink) and enacyloxin IIa (orange). Charge states and *m/z* highlighted. B | Quadrupole isolated and collisionally activated $[M + 12H]^{12+}$ EF-Tu•GDPNP•ENX showing ejection of enacyloxin IIa (orange) and emergence of EF-Tu•GDPNP (blue/pink). C | Zoom of enacyloxin IIa in B showing sodiated and potassiated ions.

Deconvolution of each peak yielded masses of 44,085 ± 0.64, 44,631 ± 0.81 and 45,336 ± 0.51 Da for *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX respectively. The theoretical mass of *E. coli* EF-Tu based upon its amino acid sequence is 43,944.28 Da resulting in the observed *apo*-EF-Tu mass being 116 Da larger than expected. Compared to the observed *apo*-EF-Tu mass, EF-Tu•GDPNP is 546 Da larger, corresponding to the addition of GDPNP, with some error (-42 Da). The further addition of enacyloxin IIa confers an observed mass increase of 705 Da, comparable to the calculated enacyloxin IIa molecular weight of 702.6 Da. Quadrupole isolation of $[M + 12H]^{12+}$ followed by collisional activation induces dissociation of enacyloxin IIa, as seen with *A. baumannii* EF-Tu•GDPNP•ENX (chapter 4), which ejected as both a neutral and charged ion, indicated by the emergence of EF-Tu•GDPNP $[M + 12H]^{12+}$ and $[M + 11H]^{11+}$ respectively (figure 6.5, B). Liberated enacyloxin IIa was observed as both sodiated and potassiated cations (figure 6.5, C).

6.3.2.1. *CIU of E. coli EF-Tu Complexed with Enacyloxin IIa*

To investigate the gas-phase stability of EF-Tu, and how it was influenced by bound ligands, the $[M + 12H]^{12+}$ ions of *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX were quadrupole isolated and subject to CIU with collision voltages ranging from 30

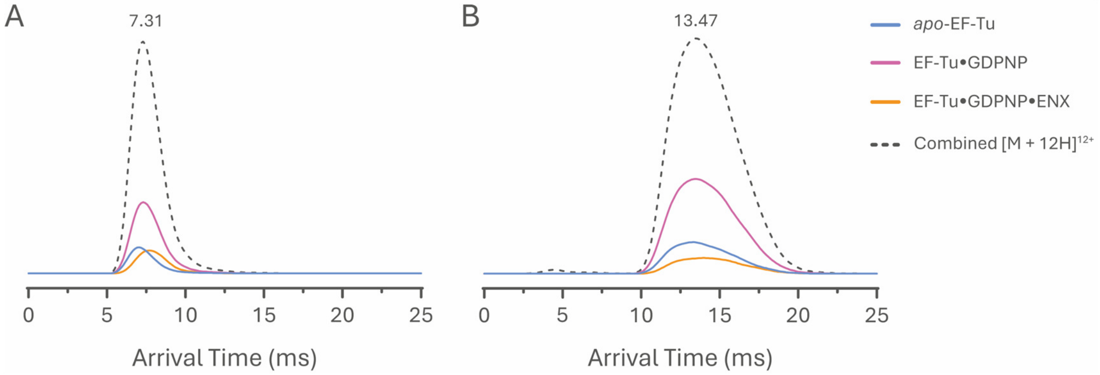


Figure 6.6 | Normalised arrive time distributions of $[M + 12H]^{12+}$ *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX ions (blue, pink and orange respectively) at collision voltages of 30 V (A) and 60 V (B). The dashed line shows the arrive time distribution of the combined EF-Tu species, with arrival time maxima given.

– 60 V (section 2.2.2.2). With low collisional activation (30 V collision energy), average arrival times of 7.05, 7.31, 7.70 ms were observed for *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX ions respectively (figure 6.6, A). At the maximum tested collision voltage of 60 V, average arrival times increased to 13.34, 13.56 and 14.24 ms for *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX ions respectively, indicating a significant degree of unfolding has occurred (figure 6.6, B). Interrogation of the full range of collision voltages was performed in CIUSuite 2 (section 2.2.2.2). CIU fingerprint plots (figures 6.7, A – C for *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX respectively) alongside their corresponding arrival time distribution waterfall plots (figures 6.7, D – F) were produced using Python code.

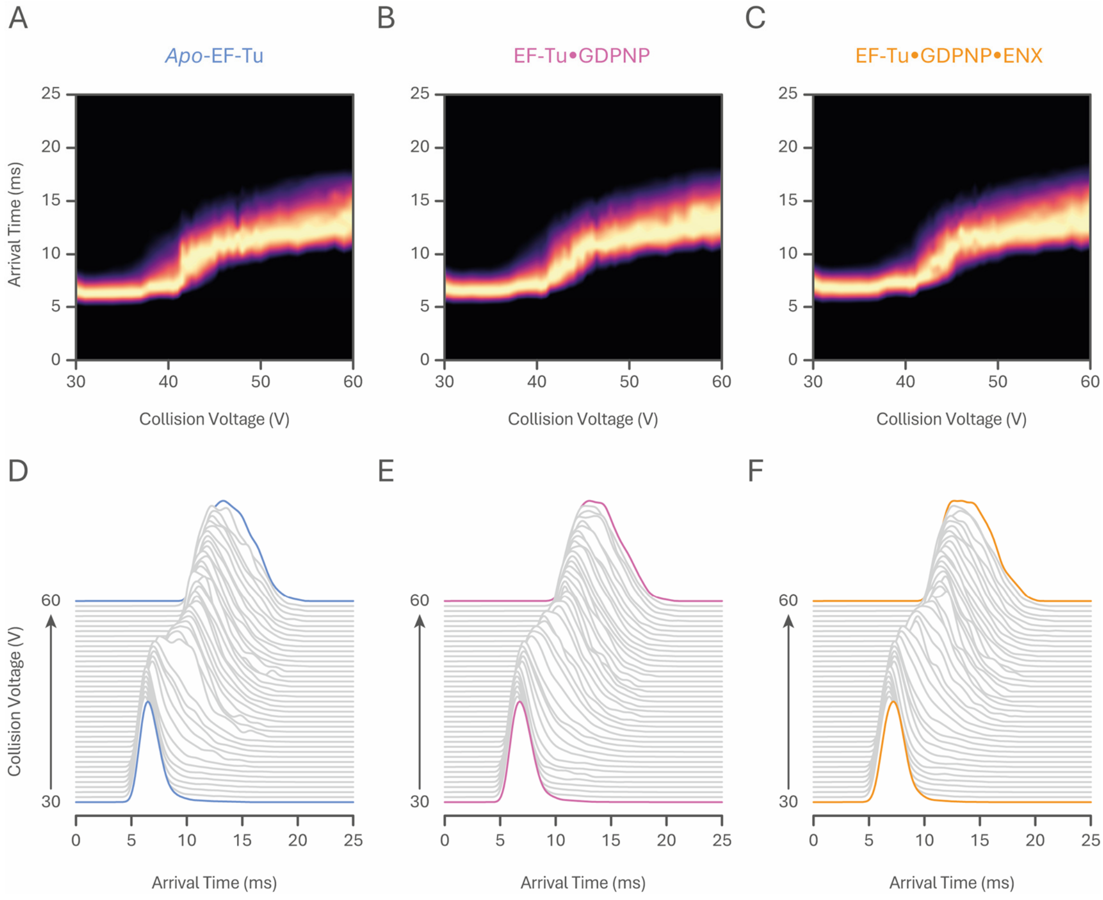


Figure 6.7 | CIU 2D fingerprint plots for $[M + 12H]^{12+}$ *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX ions (A, B and C respectively). Each arrival time distribution at a given collision voltage is normalised and coloured by intensity. The corresponding waterfall plots (D, E and F) show each normalised arrival time distribution.

These data indicated EF-Tu underwent a single unfolding event at a collision voltage of around 40 V, leading to a large increase in average arrival time. Interestingly, it appears that *apo*-EF-Tu underwent a more binary unfolding, were the arrival time distribution (figure 6.8, D) corresponding to a collision voltage of 41.5 V clearly showed two distinct arrival time maxima. This behaviour was not observed in EF-Tu•GDPNP and EF-Tu•GDPNP•ENX, where a more gradual unfolding occurs.

Calibration of TWIMS data using denatured myoglobin, native cytochrome C and native ubiquitin was performed (section 2.2.2.1), which permits accurate conversion of average arrival time to corrected $^{TW}CCS_{N2\rightarrow He}$. CIUSuite 2 allows the export of 2D IMS-MS data in an accessible `csv` data format for easy upstream processing, from which, the $^{TW}CCS_{N2\rightarrow He}$ of *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX can be calculated for each collision voltage (figure 6.8, A). Surprisingly, *apo*-EF-Tu was found to have a smaller $^{TW}CCS_{N2\rightarrow He}$ than the ligand-bound protein (33.8 ± 0.3, 34.7 ± 0.6 and 35.6 ± 0.4 nm^2 for *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX respectively). This relationship contrasted with the expected *open apo*-EF-Tu possessing a greater $^{TW}CCS_{N2\rightarrow He}$ than the *closed* EF-Tu•GDPNP and EF-Tu•GDPNP•ENX. Calculation of $^{TW}CCS_{N2\rightarrow He}$ for crystal structures of *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX yielded values of 45.2, 31.5 and 40.6 nm^2 respectively. These theoretical $^{TW}CCS_{N2\rightarrow He}$ values covered the two extremes of *open* and *closed* EF-Tu, and as disused previously (chapter 4), EF-Tu in solution does not solely occupy these extremes. Instead it samples a range of conformations in between, leading to an average CCS between them. This, paired with the expected initial gas-phase collapse of the tertiary structure, explains the observed $^{TW}CCS_{N2\rightarrow He}$ values being within the theoretical extremes. It is possible that *apo*-EF-Tu was affected by this collapse to a greater extent, without stabilising ligands, leading to an initially more compact structure.

Comparison between each EF-Tu species was best drawn from their respective increases in $^{TW}CCS_{N2\rightarrow He}$ ($\Delta\ ^{TW}CCS_{N2\rightarrow He}$) from their initial, most compact conformation (figure 6.8, B). It was clear that the collision voltage required to induce unfolding of *apo*-EF-Tu was less than that of EF-Tu•GDPNP, which in turn was less than that of EF-Tu•GDPNP•ENX, indicating conformational stability induced by ligand binding. To quantify these differences in stability, the CIU$_{50}$ value was used. This is the collision voltage required to convert 50 % of the compact protein into the extended, unfolded state. By fitting four-parameter sigmoidal logistic curves to $^{TW}CCS_{N2\rightarrow He}$ data (figures 6.8, A and B), CIU$_{50}$ values can be extracted (reported in Prism as IC$_{50}$). *Apo*-EF-Tu reported a CIU$_{50}$ of 40.3 ± 0.8 V, significantly lower than that of EF-Tu•GDPNP (43.2 ± 0.5 V, $p < 0.001$) and EF-Tu•GDPNP•ENX (44.6 ± 1.0 V, $p < 0.001$). Interestingly the additional binding of enacyloxin IIa over EF-Tu•GDPNP did not induce a significant increase in CIU$_{50}$ ($p = 0.42$, figure 6.9, A), although there were clear increases in stability induced by enacyloxin IIa binding over EF-Tu•GDPNP at collision voltages lower than this, as reported by the reduced $^{TW}CCS_{N2\rightarrow He}$.

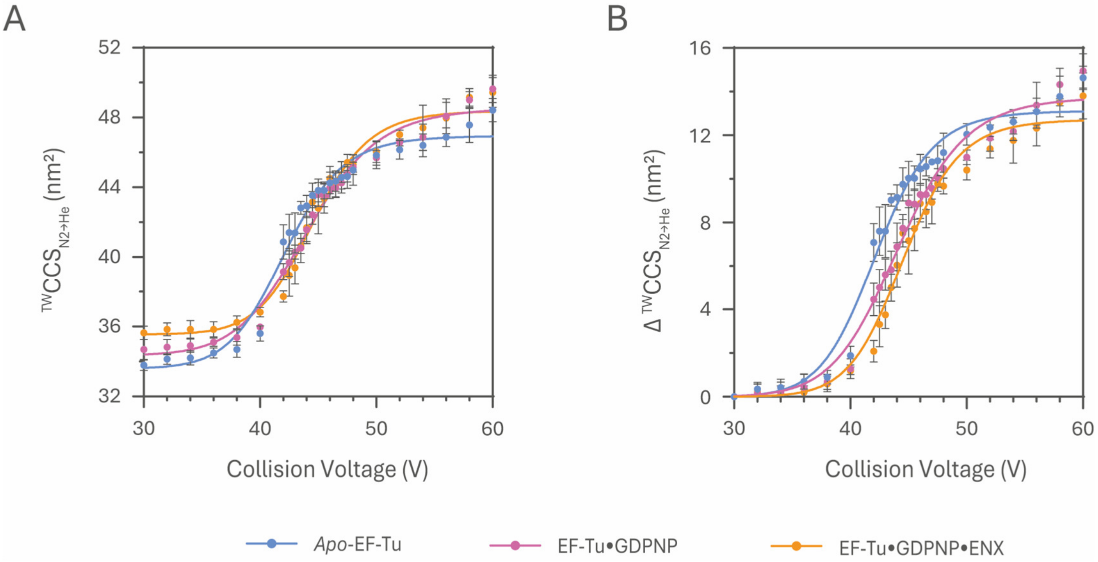


Figure 6.8 | A | Calculated $^{TW}CCS_{N2\rightarrow He}$ for *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX across the collision voltage range. B | Δ $^{TW}CCS_{N2\rightarrow He}$ for each species showing the increase in $^{TW}CCS_{N2\rightarrow He}$ from the starting compact structure.

To investigate this, collision voltages required to convert x % of the compact protein into the unfolded state, where x ranges between 10 – 90 % (figure 6.9, B) were calculated. Using this metric, EF-Tu•GDPNP•ENX was found to have significantly higher CIU$_x$ ($x \leq 40$ %, equating to a collision voltage ≤ 43.3 V) than EF-Tu•GDPNP ($p \leq 0.03$). Interestingly, this showed that significant increases in conformational stability due to bound enacyloxin IIa are lost between collision voltages of 43.2 - 44.6 V. This correlated with enacyloxin IIa ejection studies (chapter 4, figure 4.8) which showed 50 % of bound enacyloxin IIa is ejected *via* CID at a collision voltage of 42.2 ± 0.2 V. It would appear that significantly increased stability induced by enacyloxin IIa binding was lost due to ejection of the antibiotic at collision voltages similar to those calculated for EF-Tu•GDPNP•ENX CIU_{50}. This gives confidence that ligand binding to EF-Tu did increase the gas-phase stability of its compact fold, and that this increased stability was modulated by ligand ejection from the EF-Tu complex.

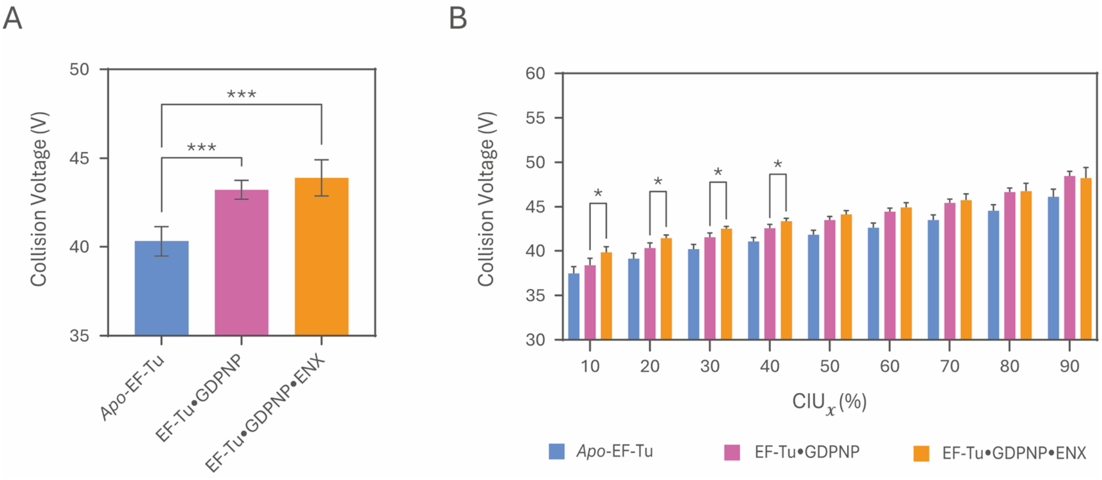


Figure 6.9 | A | Calculated CIU_{50} for *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX. B | Calculated CIU_x for *apo*-EF-Tu, EF-Tu•GDPNP and EF-Tu•GDPNP•ENX. Significance is shown for EF-Tu•GDPNP *vs* EF-Tu•GDPNP•ENX. Data are average values of three replicates and error bars show one standard deviation. Stars indicate significance. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

6.3.3. Molecular Dynamics Simulations of EF-Tu

MD simulations of *E. coli* EF-Tu *in vacuo* were performed in *apo* and *holo* states as described in section 2.4.3 using GROMACS 5.1.2. Briefly, *apo*-EF-Tu homology models in *open* and *closed* conformations were produced (section 2.4.1) and protonated using ChargePlacer to give a net 12^+ charge state. Charged protein was combined with GDP, GDPNP or GDPNP•ENX to produce *holo* complexes. All protein systems were energy minimised, and thermally equilibrated (298 K, 50 ps).

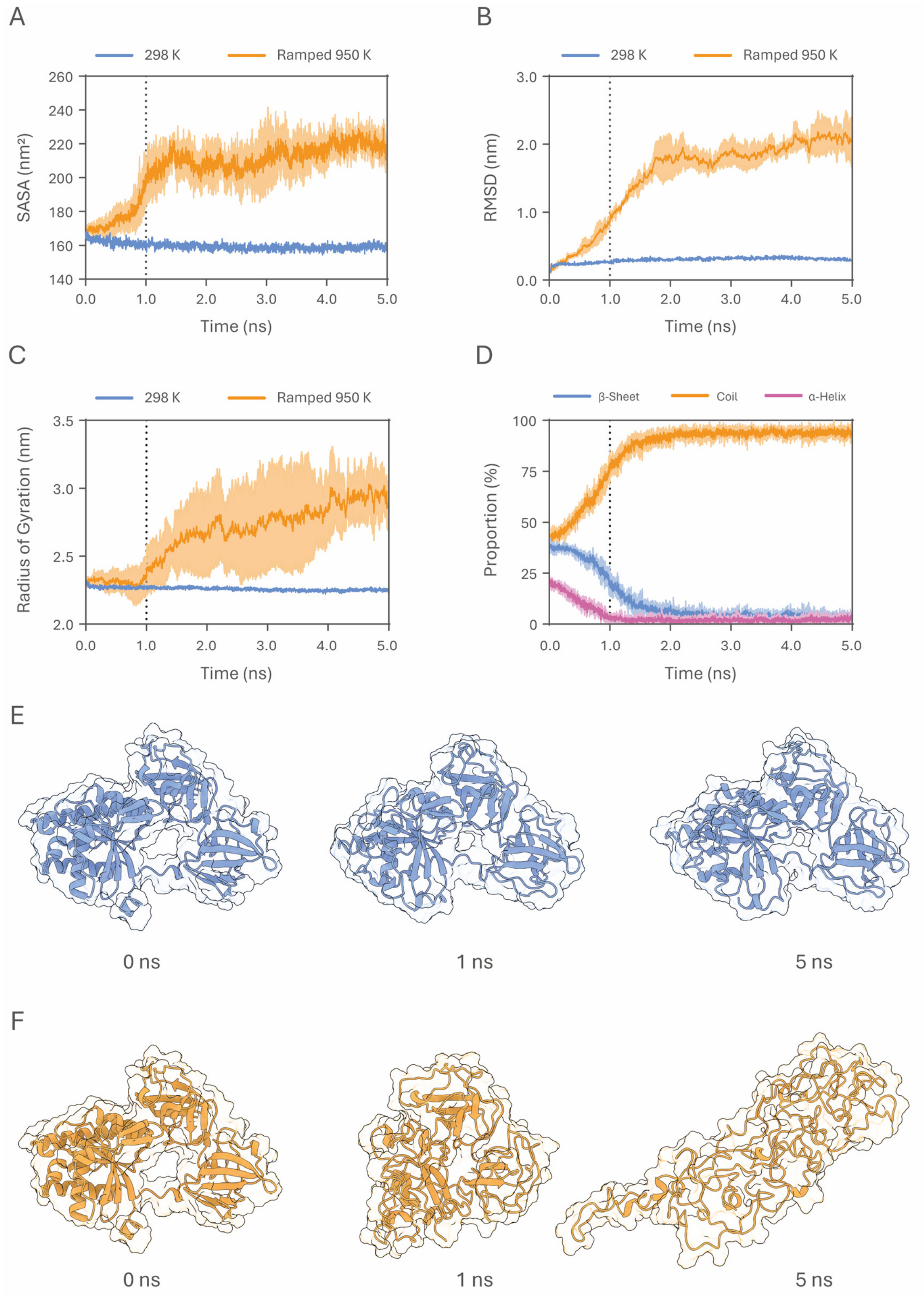


Figure 6.10 | *Apo*-EF-Tu$^{\mathrm{OPEN}}$ MD Simulation data. Main figure legend overleaf.

Figure 6.10 | Cont. A | Calculated SASA for *apo*-EF-Tu^{OPEN} simulated at 298 K (blue) and ramped 298 – 950 K (orange). B | Calculated RMSD for *apo*-EF-Tu^{OPEN} simulated at 298 K and a ramped 298 – 950 K, coloured as in A. C | Calculated radii of gyration for *apo*-EF-Tu^{OPEN} simulated at 298 K and a ramped 298 – 950 K, coloured as in A. D | Secondary structure analysis of *apo*-EF-Tu^{OPEN} simulated at a ramped 298 – 950 K. E | Representative structures from *apo*-EF-Tu^{OPEN} 298 K simulations at 0, 1 and 5 ns. F | Representative structures from *apo*-EF-Tu^{OPEN} 298 - 905 K simulations at 0, 1 and 5 ns. EF-Tu is shown as a cartoon and transparent surface. Data for ramped 298 – 950 K simulations are average values of three replicates and the shaded regions show one standard deviation.

To confirm protein stability *in vacuo, apo*-EF-Tu in the *open* conformation was simulated for 5 ns held at a temperature of 298 K. To interrogate the thermal stability of each complex, 5 ns simulations were performed consisting of a 1 ns linear thermal ramp (298 – 950 K) and a further 4 ns held at 950 K. The resulting MD trajectories were analysed using inbuilt GROMACS functions, alongside analysis in SideFX Houdini.

Initial simulations were performed consisting of *apo*-EF-TU in the *open* configuration held at 298 K and subject to ramped temperature. *Apo*-EF-Tu simulated in the *open* (EF-Tu^{OPEN}) conformation at 298 K showed a relatively minor drop in SASA (167.28 to 158.14 nm^2, 5.47 % decrease) over the 5 ns simulation indicating a small compaction of the structure (figure 6.10, A). When the protein was subjected to ramped heating a large degree of unfolding was observed, as monitored by SASA (figure 6.10, A), root mean square deviation (RMSD, figure 6.10, B), radii of gyration (figure 6.10, C) and secondary structure analysis (figure 6.10, D). SASA measurements show unfolding into an extended conformation consistent with ramped heating. Continued hold at 950 K produced no further conformation expansion but rather an increasingly higher disordered and flexible conformation, as shown by continued increases in RMSD and radii of gyration alongside an almost complete degradation of ordered secondary structure. Representative structures extracted during the simulation can be seen in figure 6.10, E and F. To investigate the stability of EF-Tu upon addition of bound ligands, two *holo*-EF-Tu complexes were made; EF-Tu^{OPEN}•GDP (*open* conformation) and EF-Tu^{CLOSED}•GDPNP•ENX (*closed* conformation) alongside an *apo*-EF-Tu^{CLOSED} structure in the *closed* conformation as a further control. All of these were subjected to a ramped heating simulation for 5 ns. As with *apo*-EF-Tu^{OPEN}, all protein experienced significant unfolding as a result of heating (represented by an increase in SASA (Δ SASA) from their respective starting structures, figures 6.11, A and B).

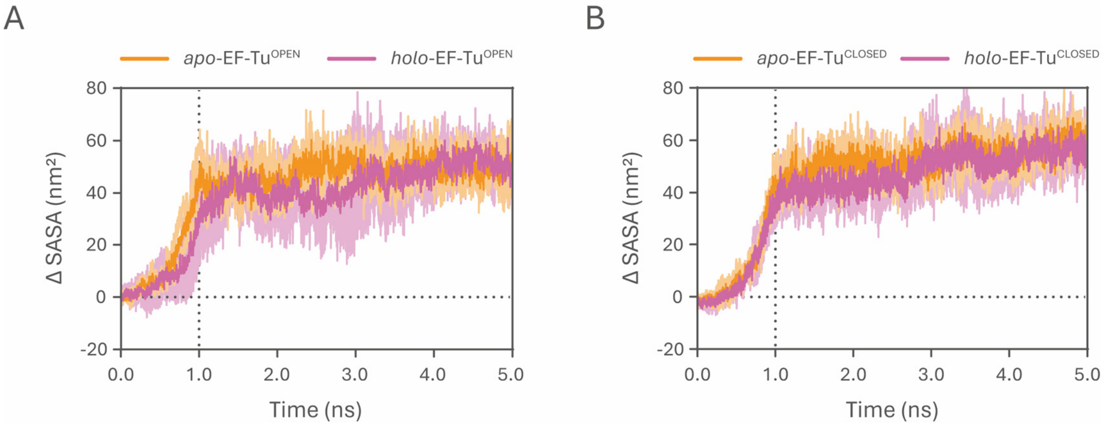


Figure 6.11 | A | Calculated increases in SASA (Δ SASA) for *apo* (orange)and *holo*-EF-Tu^{OPEN} (EF-Tu^{OPEN}•GDP, pink). B | Calculated increases in SASA (Δ SASA) for *apo* (orange) and holo-EF-Tu^{CLOSED} (EF-Tu^{CLOSED}•GDPNP•ENX, pink). Data are average values of three replicates and the shaded regions show one standard deviation.

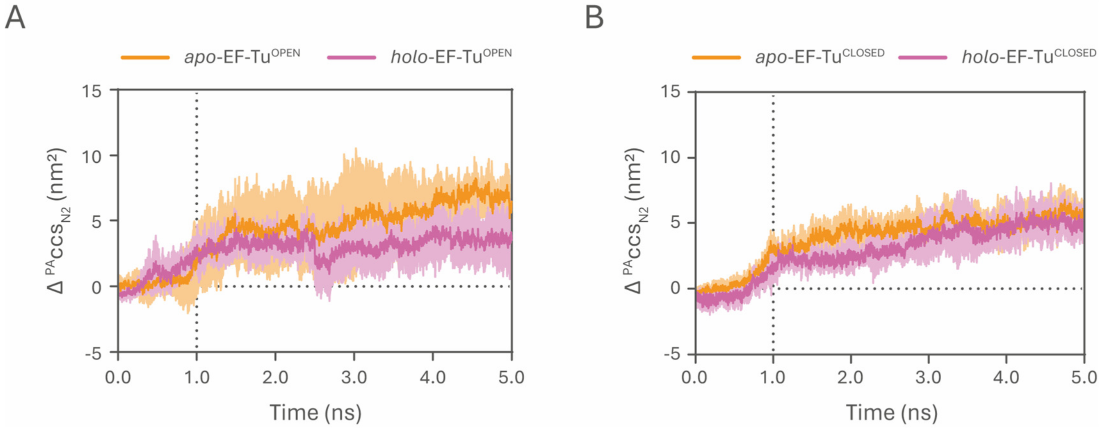


Figure 6.12 | A | Calculated increases in $^{PA}CCS_{N2}$ (Δ $^{PA}CCS_{N2}$) for *apo* (orange) and *holo*-EF-Tu^{OPEN} (EF-Tu^{OPEN}•GDP, pink). B | Calculated increases in $^{PA}CCS_{N2}$ (Δ $^{PA}CCS_{N2}$) for *apo* (orange) and holo-EF-Tu^{CLOSED} (EF-Tu^{CLOSED}•GDPNP•ENX, pink). Data are average values of three replicates and the shaded regions show one standard deviation.

To interrogate this further, corrected PA CCS ($^{PA}CCS_{N2}$) values were calculated for proteinaceous atoms throughout the simulations (shown as the increase in $^{PA}CCS_{N2}$, Δ $^{PA}CCS_{N2}$, figures 6.12, A and B). *Note, $^{PA}CCS_{N2}$ values were calculated at half temporal resolution (every 2 ps) as opposed to the base simulation resolution of 1 ps due to software limitations with CCSCalc*. From these, there were clear qualitative differences between systems. The presence of GDP in EF-Tu^{OPEN}•GDP induced a smaller $^{PA}CCS_{N2}$ over *apo*-EF-Tu^{OPEN} (final values of 36.33 and 39.29 nm^2 respectively), indicating a more compact structure.

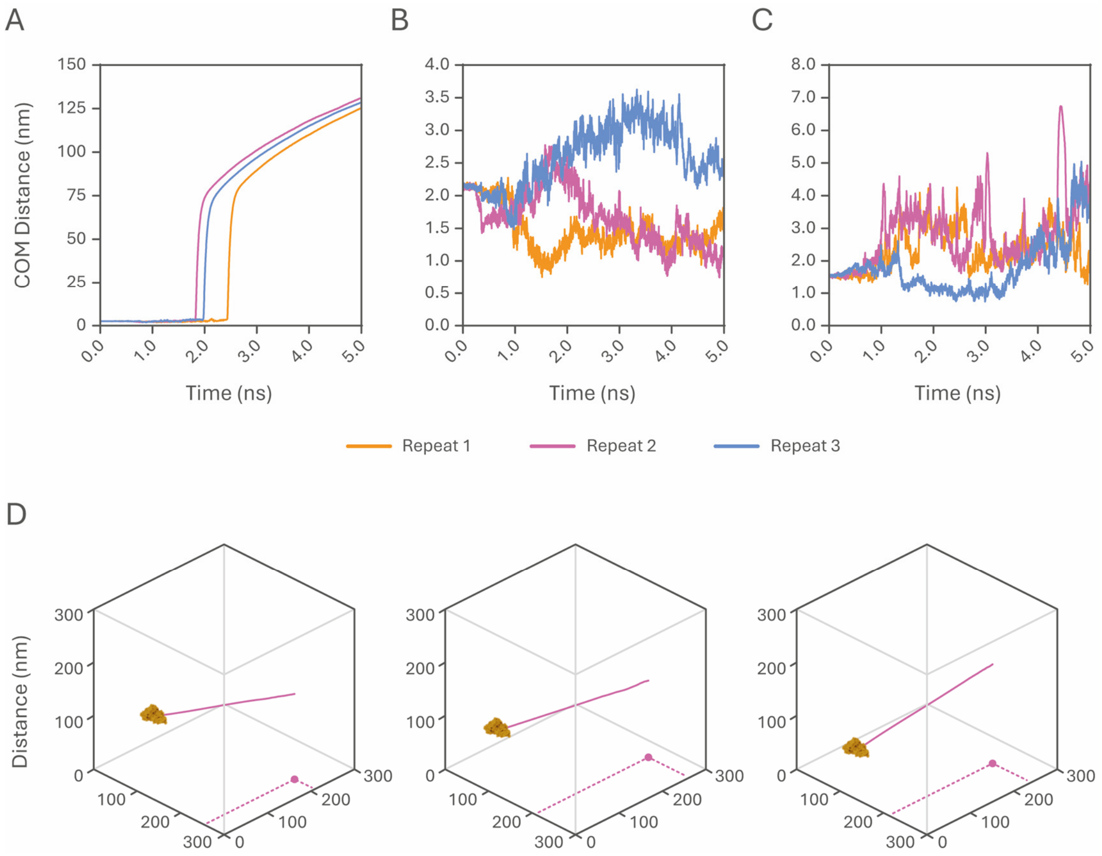


Figure 6.13 | A | Calculated centre of mass (COM) differences between EF-Tu and GDP. B | Calculated COM differences between EF-Tu and GDPNP. C | Calculated COM differences between EF-Tu and enacyloxin IIa. D | 3D (*x,y,z*) representations of GDP ejection trajectories. EF-Tu positions (orange) as of the start of the simulation (EF-Tu structure not to scale). GDP trajectory is represented as a pink line. Final (*x,z*) GDP position highlighted by the pink circle and dashed lines.

Interestingly, during all three replicates of the EF-Tu^{OPEN}•GDP simulations GDP was ejected from the protein after 1.75 – 2.5 ns, corresponding to a slight drop in $^{PA}CCS_{N2}$ (figure 6.12, A, figure 6.13, A and figure 6.14, A). This behaviour was not displayed by GDPNP or enacyloxin IIa which, although move locally through the system, remain associated with EF-Tu (figure 6.13, B and C and figure 6.14, B). GDPNP remains tightly associated to the Mg^{2+} cofactor and buried within the protein structure. Enacyloxin IIa, on the other hand, is fully exposed on the surface of EF-Tu. The 3D ejection trajectories of GDP for each repeat simulation are shown in figure 6.13, D. Initially, enacyloxin IIa did provide a stabilising effect on EF-Tu^{CLOSED}•GDPNP•ENX for the first 3 ns of the simulations (figure 6.12, B).

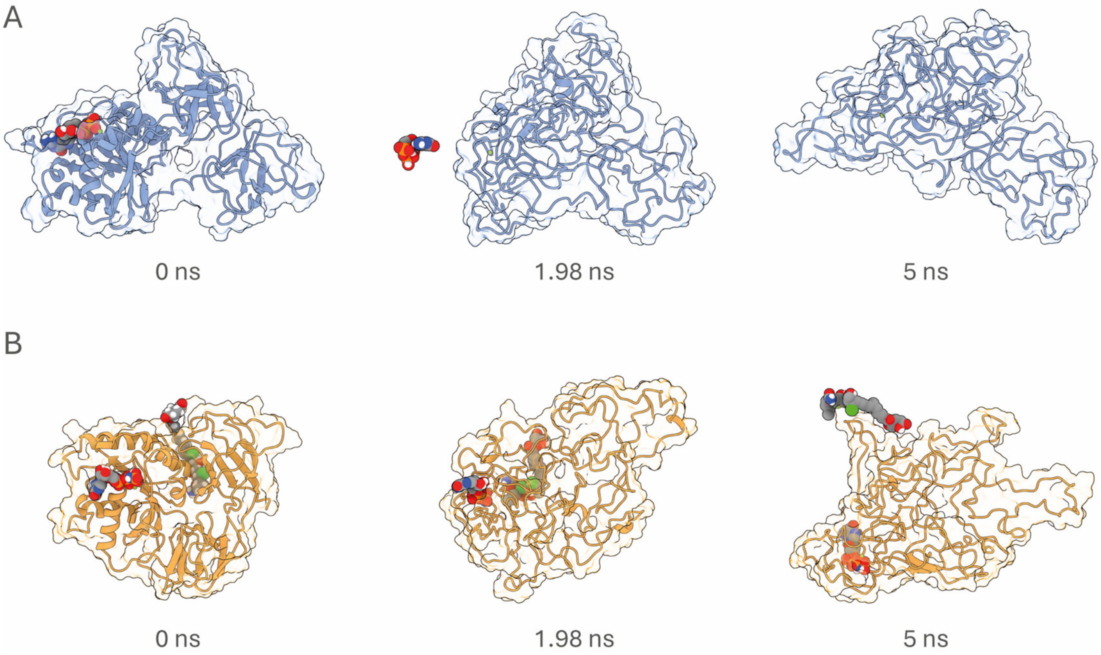


Figure 6.14 | A | Representative structures from EF-Tu^{OPEN}•GDP 298 - 905 K simulations at 0, 1.98 and 5 ns. Note, 1.98 ns was selected to capture ejection of GDP from EF-Tu. F | Representative structures from EF-Tu^{CLOSED}•GDPNP•ENX 298 - 905 K simulations at 0, 1.98 and 5 ns. Note, buried GDPNP and surface associated enacyloxin IIa at 5 ns. EF-Tu is shown as a cartoon and transparent surface. GDP(NP) and enacyloxin IIa are shown as VDW spheres, coloured by element and Mg^{2+} shown as a green sphere.

Loss of stabilisation correlates with movement of enacyloxin IIa from within what remains of its binding pocket to the surface of the protein (figure 6.13, C and figure 6.14, B). To investigate ligand stabilisation effects further, domain-level proteinaceous atom RMSD was calculated for *apo*-EF-Tu^{CLOSED} and EF-Tu^{CLOSED}•GDPNP•ENX (represented in figure 6.15 as the Δ RMSD between *apo*-EF-Tu^{CLOSED} and EF-Tu^{CLOSED}•GDPNP•ENX). Here, domain I was found to have an overall negative Δ RMSD (therefore greater RMSD in *apo*-EF-Tu^{CLOSED} then EF-Tu^{CLOSED}•GDPNP•ENX), with domains II and III remaining at approximately a Δ RMSD of 0. This would imply an overall stabilising effect of GDPNP binding within domain I during the EF-Tu^{CLOSED}•GDPNP•ENX simulation. As can be seen in figure 6.14, B, GDPNP remains within the collapsed protein structure during the simulations, tightly associated with the Mg^{2+} cofactor.

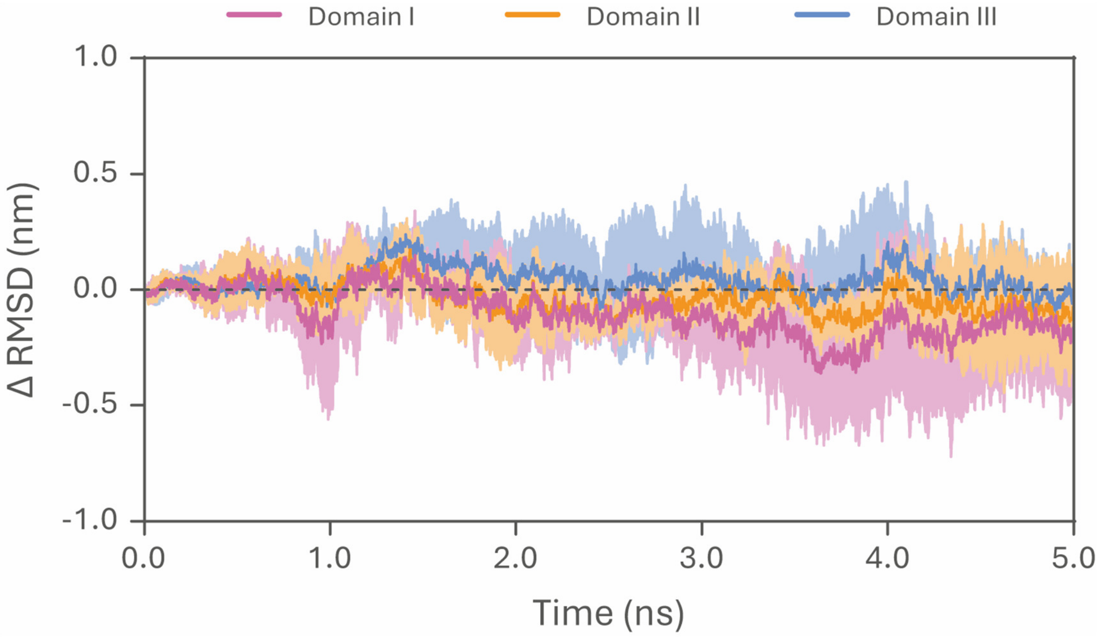


Figure 6.15 | Change in RMSD (Δ RMSD) between *apo*-EF-Tu^{CLOSED} and EF-Tu^{CLOSED}•GDPNP•ENX and per domain. A more negative Δ RMSD would infer greater stability in EF-Tu^{CLOSED}•GDPNP•ENX compared to *apo*-EF-Tu^{CLOSED}. Data are average values of three replicates and the shaded regions show one standard deviation.

In contrast, enacyloxin IIa did not remain associated with its binding groove, as this was largely destroyed during unfolding, and hence did not impart a direct stabilisation effect on EF-Tu. It is possible that additional stability is conferred to domains I and III due to enacyloxin IIa binding prior to collapse of the binding site.

In all MD simulations, EF-Tu did not undergo unfolding to the extent that was seen experimentally (greatest MD $\Delta\ {}^{PA}CCS_{N2}$ of 8.42 nm^2 compared to a greatest experimental $\Delta\ {}^{TW}CCS_{N2\rightarrow He}$ of 14.95 nm^2). This is unsurprising for several reasons; CIU occurs over millisecond timescales compared to the nanosecond timescale MD simulations and CID-mediated unfolding is an intrinsically different process than imparting a linear, protein-wide thermal gradient. Although truly direct comparisons cannot be made between the two methods, the qualitative outcomes are similar: the presence of bound ligands increases EF-Tu stability over the *apo*-protein.

6.4. Conclusions

Here, I have shown the use of native mass spectrometry as a tool to quickly and accurately assess the binding of novel enacyloxin analogues to EF-Tu. Importantly, this was achieved using simple crude extracts of each analogue, removing the need for time consuming preparatory HPLC. The specificity of mass spectrometry allowed "fishing" of bound EF-Tu out of complex sample mixtures, alongside ejection and identification of each bound analogue through CID.

I also investigated the stability of EF-Tu in the gas-phase through CIU and MD simulations. The results show that the EF-Tu compact, gas-phase structure is significantly stabilised by GDPNP binding, and that the EF-Tu•GDPNP complex is significantly further stabilised by binding of enacyloxin IIa. This additional stability is modulated by CID ejection of enacyloxin IIa and is lost at collision voltages higher than that shown to eject 50 % of enacyloxin IIa. Although not a direct comparison, thermal unfolding MD simulations showed a similar phenomenon. Binding of GDP and GDPNP•ENX induced greater thermal stability over their respective *apo* proteins.

Due to hardware limitations, it was not feasible to run all-atom MD simulations for longer than 5 ns. However, to further explore EF-Tu unfolding microsecond simulations could be performed. This would allow for conformational changes from *open* to *closed* EF-Tu to be observed, and although still an order of magnitude smaller in timescale than CIU, would theoretically provide a more accurate comparison to experimental data.

7

Overall Conclusions

This thesis has explored the use of several structural mass spectrometry techniques to study the prokaryotic elongation factor EF-Tu. Specifically, it focussed on the interactions between EF-Tu and several cognate binding partners, both small molecule and protein in nature. Key concepts surrounding protein structure and interactions were introduced, the intimate role of EF-Tu in prokaryotic translation through transport of aminoacyl-tRNA was outlined and the relevant broad field of mass spectrometry discussed (chapter 1). The key experimental methods (chapter 2) used throughout this work were described, along with data analyses and computational methods.

The experimental focus of this thesis was predominately the application of carbene protein footprinting to EF-Tu. As discussed in detail previously, carbene footprinting allows protein-ligand interactions to be elucidated. Accessible residues on the protein surface can be rapidly and irreversibly labelled by a probe, which can then be quantified by mass spectrometry. Binding of a partner, small molecule or protein, should impart a masking effect, protecting the underlying protein from labelling. Conformational changes in a target protein can also be highlighted by carbene footprinting, as regions of the protein become more or less exposed as a result of conformation change, a corresponding increase or decrease in labelling would be expected.

This method was first used to interrogate binding of the thiazolyl elfamycin GE2270A to *A. baumannii* EF-Tu (chapter 3). The elfamycins are a diverse range of compounds that all specifically target EF-Tu, inhibiting its role in prokaryotic translation. GE2270A is known to bind to domain II of EF-Tu, specifically in the *open* conformation of the protein, and prevents binding of aminoacyl-tRNA, acting as a competitive inhibitor. Although there was no reason to question binding of GE2270A to *A. baumannii* EF-Tu, no published research has explored this specific interaction. Carbene footprinting of EF-Tu in complex with GDP was performed as a control, which was compared against EF-Tu in complex with both GDP/GDPNP and GE2270A. These data showed highly localised, significant reductions in labelling within domain II of EF-Tu when complexed with GE2270A, an effect which was found to be GDP/GDPNP dependent. EF-Tu is more likely to adopt an *open* conformation bound to GDP, which is required for GE2270A binding. When bound to GDPNP, on average, EF-Tu adopts a more *closed* conformation, which is detrimental to GE2270A binding. Interrogation of the footprinting data showed GE2270A-dependent masking effects on peptides containing residues predicted to contact the antibiotic. These were paired with several unexpected significant masking effects in peptides proximal, but not directly within, the binding site. These effects are tentatively inferred to be due to structural perturbations within the protein structure induced by GE2270A binding, specifically as these peptides cover disordered loops within domain II which are inherently flexible and my become more ordered upon ligand binding.

Overall, this chapter demonstrates the effectiveness of carbene footprinting to capture binding of GE2270A to *A. baumannii* EF-Tu. Considering other work discussed in this thesis, future work relating to GE2270A could take several avenues, the first being production of GE2270A resistant EF-Tu mutants. Several domain II residue substitutions highlighted in chapter 3 conferred resistance to the antibiotic. If these mutations induce resistance by inhibition of GE2270A binding, this should be reflected in a reversal of the masking effects seen in the wild-type (GE2270A sensitive) protein, an effect shown for the contrasting elfamycin, enacyloxin IIa, in chapter 4. The second potential direction would be a molecular dynamics investigation into the conformational fluidity of the EF-Tu•GE2270A complex. As discussed in chapter 1, crystal structures do not represent the true solution phase structure (or ensemble of structures) of a protein, and do not directly provide information on conformational fluidity. Through microsecond-scale molecular dynamics simulations, the full dynamics of EF-Tu can be captured, including the conformational switching between *open* and *closed* conformations. Simulations of this nature of the EF-Tu•GE2270A complex would elucidate any conformational perturbations induced by GE2270A binding, which are otherwise not able to be captured in the single published crystal structure of the complex. This would allow for greater understanding of the carbene footprinting data, especially in those peptides thought to be masked by changes in conformational flexibility.

Carbene footprinting was also used to probe the interactions between EF-Tu and the polyketide-based elfamycin enacyloxin IIa (chapter 4). In contrast to GE2270A, enacyloxin IIa binds to EF-Tu in the *closed* conformation, at the domain I – II interface, which does not exist in the *open* conformation. Enacyloxin IIa exerts its antibiotic effects through locking EF-Tu in the closed conformation, irrespective of GTP hydrolysis. This prevents its dissociation from the ribosome, therefore causing a cessation of protein translation. Similarly to GE2270A, no structural information regarding binding of enacyloxin IIa to *A. baumannii* EF-Tu has been published. This, paired with the expected large conformational change associated with enacyloxin IIa binding (*open* to *closed*) made EF-Tu•ENX an interesting target for carbene footprinting analysis.

Carbene footprinting was performed initially on wild-type EF-Tu, *apo* protein and EF-Tu•GDPNP were controls. No significant differences in labelling were seen between *apo*-EF-Tu and EF-Tu•GDPNP, highlighting their conformational similarity in solution. In contrast, when the EF-Tu•GDPNP•ENX complex was subject to carbene footprinting, significant reductions in labelling were seen both globally across the protein and locally within the predicted enacyloxin IIa binding site. Global reductions in labelling were found in regions of EF-Tu that become either inaccessible in the *closed* conformation or instead become more disordered *e.g.*

switch I. Together, these indicated that carbene footprinting has captured the large conformational change associated with enacyloxin IIa binding. The presence of localised masking within the enacyloxin IIa binding site gave confidence that these effects were due to enacyloxin IIa binding.

To further investigate this, several enacyloxin IIa resistant EF-Tu mutants were produced through targeted mutagenesis of wild-type EF-Tu. Each mutation conferred differing levels of enacyloxin IIa resistance as measured by their binding affinities to the antibiotic. These measurements were acquired through native mass spectrometry titrations. Native mass spectrometry was also employed to study the CID mediated ejection of enacyloxin IIa from EF-Tu. Carbene footprinting of each mutant bound to GDP and enacyloxin IIa showed a resistance-dependent effect on labelling. As resistance towards enacyloxin IIa increased, the overall amount of labelling increased, which was interpreted as EF-Tu adopting a more *open* conformation. This effect would be expected with reduced enacyloxin IIa binding, as it is this binding that induces adoption of the *closed* conformation.

The work in this chapter highlights the capabilities of carbene footprinting to capture these large-scale conformational changes, an area of research that had not been investigated through carbene footprinting before. Importantly, these effects were modulated by enacyloxin IIa resistance, indicating that they were directly due to its binding to EF-Tu. However, several areas of research remain unexplored. Carbene footprinting was performed on mutant EF-Tu bound to GDP and enacyloxin IIa, not the preferential GDPNP. The much lower K_D values of EF-Tu•GDPNP towards enacyloxin IIa made differentiation of the resistant mutants impossible under these conditions. Methods to reduce the overall concentrations of EF-Tu and enacyloxin IIa in the footprinting experiments were tentatively explored *e.g.* performing footprinting at low concentration of protein before combing aliquots for upstream processing at acceptable protein concentrations. However, time constraints meant this could not be explored to its full potential. Similar to the EF-Tu•GE2270A complex, investigation into conformational fluidity through microsecond-scale molecular dynamics simulations would be a great asset to combine with carbene footprinting data. Especially for regions of EF-Tu showing high flexibility or disorder, which are often not represented in the published crystal structures.

Protein-protein interactions between *P. putida* EF-Tu and a putative prolyl hydroxylase (PPHD) were investigated through carbene footprinting in chapter 5. As discussed, EF-Tu was found to be the target of a putative prolyl hydroxylase in *P. putida*. PPHD was found to bind to domain I and II of EF-Tu, whilst unravelling EF-Tu switch I region and positioning it within its binding site. Native mass spectrometry showed minor formation of the EF-Tu•PPHD complex when the substituent proteins were incubated at equimolar concentrations. Dual chymotryptic and tryptic carbene footprinting experiments on both EF-Tu and PPHD showed masking effects on each protein associated with their binding. Specifically, masking was seen on EF-Tu switch I and its PPHD binding site alongside minor masking in EF-Tu domain II. These results suggested that PPHD binds strongly to switch I/domain I of EF-Tu and perhaps more transiently to domain II, as indicated by the reduced masking effects there. This complex lends itself to a raft of potential further research. Hydroxylation of EF-Tu is possible and, despite being difficult and out of the scope of this project, would be possible to purify out from native EF-Tu. This modified EF-Tu could then be subject to a range of experiments, including repeats of work described in this thesis to investigate the effects, if any, that hydroxylation has on these EF-Tu interactions.

A method for rapidly characterising binding of semi-synthetic enacyloxin IIa analogues was explored in chapter 6. Resuspended crude extracts of enacyloxin IIa analogues were incubated with EF-Tu prior to buffer exchange and desalting. Using native mass spectrometry, the binding of each analogue to EF-Tu could be studied. These data were purely qualitative, as the analogues were of unknown concentrations, but the method allowed for quick screening of novel compounds without the need for preparative HLPC purification.

Finally, the gas-phase stability of EF-Tu was investigated through IMS-MS and CIU also in chapter 6. IMS allows separation of ions based upon the size and shape, not just their mass and charge. From arrival time distributions, CCS of ions can be calculated. CIU exploits the collisional activation of a protein ion prior to IMS separation, if sufficient energy is imparted into a protein ion it may undergo partial unfolding. This can be monitored using IMS, as an unfolded protein tends to have a larger CCS. By subjecting protein ions to a range of collision voltages, the energy at which unfolding occurs can be calculated. *E. coli* EF-Tu as an *apo* protein and bound to GDPNP and enacyloxin IIa was subject to CIU to investigate stabilising effects on the protein imparted by ligand binding. Here, an increase in stability is defined as a reduction in CCS at a given collision voltage. These experiments showed that EF-Tu•GDPNP was significantly more stable than the *apo* protein, and that further significant increases in stability were conferred by enacyloxin IIa binding. However, this additional stability was modulated by the CID ejection of enacyloxin IIa, where at collisional voltages greater than, or sufficient to, eject 50 % of bound enacyloxin IIa, the additional significant stability was lost.

These data were corroborated through short *in vacuo* molecular dynamics EF-Tu simulations. *Apo* and *holo* EF-Tu in their *open* and *closed* initial conformations were subject to 5 ns MD simulations. Control simulations were performed where the protein was held at 298 K, and thermally ramped 298 – 950 K simulations were performed to induce denaturation. Although not a direct comparison to CID-mediated unfolding, these thermal simulations do allow the exploration of ligand stabilisation. Through integration of MD trajectories CCS values can be calculated for each protein throughout the simulation. These showed that the presence of bound ligands did in fact confer greater stability (*i.e.* reduced unfolding) over their respective *apo* protein conformations.

These methods lend themselves to investigation of EF-Tu bound to other compounds. GE2270A and enacyloxin IIa analogues are examples taken directly from this work. Their stabilising effects could be studied through both CIU and MD simulations. For greater confidence in MD data, extension of these simulations to the microsecond scale could be performed. CIU occurs on the millisecond timescale, and although still an order of magnitude greater than microsecond MD simulations, longer simulations will allow for better representations of EF-Tu unfolding.

Overall, this thesis has explored the use of several structural mass spectrometric methods to characterise EF-Tu. Both from an interaction perspective, and a conformational stability perspective. The work presented within demonstrated the ability of carbene footprinting to accurately map protein-ligand interactions and, importantly, large scale conformational changes associated with binding. Finally, ligand binding to EF-Tu was shown to increase its gas-phase stability, which in turn was modulated by ligand ejection at sufficient collision energies.

References

1. Milo, R. What Is the Total Number of Protein Molecules per Cell Volume? A Call to Rethink Some Published Values. *BioEssays* 2013, 35, 1050–1055.
2. Bludau, I.; Aebersold, R. Proteomic and Interactomic Insights into the Molecular Basis of Cell Functional Diversity. *Nat Rev Mol Cell Biol* 2020, 21, 327-340.
3. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Protein Function. *Molecular Biology of the Cell. 4th edition* 2002.
4. Du, X.; Li, Y.; Xia, Y. L.; Ai, S. M.; Liang, J.; Sang, P.; Ji, X. L.; Liu, S. Q. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int J Mol Sci* 2016, 17, 144
5. Pavlovic-Lazetic, G. M.; Mitić, N. S.; Kovačević, J. J.; Obradović, Z.; Malkov, S. N.; Beljanski, M. V. Bioinformatics Analysis of Disordered Proteins in Prokaryotes. *BMC Bioinformatics* 2011, 12, 1-22.
6. Deforte, S.; Uversky, V. N. Molecules Order, Disorder, and Everything in Between. *Molecules* 2016, 21, 1090.
7. Yang, L. Q.; Ji, X. L.; Liu, S. Q. The Free Energy Landscape of Protein Folding and Dynamics: A Global View. *J Biomol Struct Dyn* 2013, 31, 982–992.
8. Agbas, A.; Agbas, A. Trends of Protein Aggregation in Neurodegenerative Diseases. *Neurochemical Basis of Brain Function and Dysfunction* 2018.
9. Seeger, M. A.; Rice, S. E. Intrinsic Disorder in the Kinesin Superfamily. *Biophys Rev* 2013, 5, 233.
10. Plattner, N.; Noé, F. Protein Conformational Plasticity and Complex Ligand-Binding Kinetics Explored by Atomistic Simulations and Markov Models. *Nat Comun* 2015, 6, 1–10.
11. Fischer, E. Einfluss Der Configuration Auf Die Wirkung Der Enzyme. *Der Dtsch Chem Ges* 1894, 27, 2985–2993.
12. Savir, Y.; Tiusty, T. Conformational Proofreading: The Impact of Conformational Changes on the Specificity of Molecular Recognition. *PLoS One* 2007, 2, 468.
13. Koshland, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc* Natl *Acad Sci* 1958, 44, 98–104.

14. Monod, J.; Wyman, J.; Changeux, J.-P. On the Nature of Allosteric Transitions: A Plausible Model. *J Mol Biol* 1965, 12, 88–118.

15. Stouten, P. F. W.; Sander, C.; Wittinghofer, A.; Valencia, A. How Does the Switch II Region of G-Domains Work? *FEBS Lett* 1993, 320, 1–6.

16. Cherfils, J.; Zeghouf, M. Regulation of Small GTPases by GEFs, GAPs, and GDIs. *Physiol Rev* 2013, 93, 269–309.

17. Gilman, A. G. G Proteins: Transducers of Receptor-Generated Signals. *Annu Rev Biochem* 1987, 56, 615–649.

18. Saraste, M.; Sibbald, P. R.; Wittinghofer, A. The P-Loop — a Common Motif in ATP- and GTP-Binding Proteins. *Trends Biochem Sci* 1990, 15, 430–434.

19. Hewitt, N.; Ma, N.; Arang, N.; Martin, S. A.; Prakash, A.; DiBerto, J. F.; Knight, K. M.; Ghosh, S.; Olsen, R. H. J.; Roth, B. L.; Gutkind, J. S.; Vaidehi, N.; Campbell, S. L.; Dohlman, H. G. Catalytic-Site Mutations Confer Multiple States of G Protein Activation. *Sci Signal* 2023, 16 , 7842.

20. Maracci, C.; Rodnina, M. V. Translational GTPases. *Biopolymers* 2016, 105, 463.

21. Leibundgut, M.; Frick, C.; Thanbichler, M.; Böck, A.; Ban, N. Selenocysteine tRNA-Specific Elongation Factor SelB Is a Structural Chimaera of Elongation and Initiation Factors. *EMBO J 2005*, 24, 11.

22. Nomura, T.; Nakano, K.; Maki, Y.; Naganuma, T.; Nakashima, T.; Tanaka, I.; Kimura, M.; Hachimori, A.; Uchiumi, T. In Vitro Reconstitution of the GTPase-Associated Centre of the Archaebacterial Ribosome: The Functional Features Observed in a Hybrid Form with *Escherichia Coli* 50S Subunits. *Biochemical Journal* 2006, 396, 565.

23. Gromadski, K. B.; Wieden, H. J.; Rodnina, M. V. Kinetic Mechanism of Elongation Factor Ts-Catalyzed Nucleotide Exchange in Elongation Factor Tu. *Biochemistry* 2002, 41, 162–169.

24. Paulin, F. E. M.; Campbell, L. E.; O⌧Brien, K.; Loughlin, J.; Proud, C. G. Eukaryotic Translation Initiation Factor 5 (ElF5) Acts as a Classical GTPase-Activator Protein. *Curr Biol* 2001, 11, 55–59.

25. Song, H.; Parsons, M. R.; Ã N Rowsell, S.; Leonard, G.; Phillips, S. E. V. Crystal Structure of Intact Elongation Factor EF-Tu from *Escherichia Coli* in GDP Conformation at 2.05 Å Resolution. *J Mol Bio* 1999, 285, 1245-1256.

26. Johansen, J. S.; Kavaliauskas, D.; Pfeil, S. H.; Blaise, M.; Cooperman, B. S.; Goldman, Y. E.; Thirup, S. S.; Knudsen, C. R. *E. Coli* Elongation Factor Tu Bound to a GTP Analogue Displays an Open Conformation Equivalent to the GDP-Bound Form. *Nucleic Acids Res* 2018, 46, 8641.
27. Al-Karadaghi, S.; Ævarsson, A.; Garber, M.; Zheltonosova, J.; Liljas, A. The Structure of Elongation Factor G in Complex with GDP: Conformational Flexibility and Nucleotide Exchange. *Structure* 1996, 4, 555–565.
28. Polekhina, G.; Thirup, S.; Kjeldgaard, M.; Nissen, P.; Lippmann, C.; Nyborg, J. Helix Unwinding in the Effector Region of Elongation Factor EF-Tu-GDP. *Structure* 1996, 4, 1141–1151.
29. Lin, J.; Gagnon, M. G.; Bulkley, D.; Steitz, T. A. Conformational Changes of Elongation Factor g on the Ribosome during tRNA Translocation. *Cell* 2015, 160, 219–227.
30. Kjeldgaard, M.; Nyborg, J. Refined Structure of Elongation Factor EF-Tu from *Escherichia Coli*. *J Mol Biol* 1992, 223, 721–742.
31. Kjeldgaard, M.; Nissen, P.; Thirup, S.; Nyborg, J. The Crystal Structure of Elongation Factor EF-Tu from *Thermus Aquaticus* in the GTP Conformation. *Structure* 1993, 1, 35–50.
32. Rodnina, M. V.; Pape, T.; Fricke, R.; Wintermeyer, W. Elongation Factor Tu, a GTPase Triggered by Codon Recognition on the Ribosome: Mechanism and GTP Consumption. *Biochem Cell Biol* 2011, 73, 1221–1227.
33. Heffron, S. E.; Jurnak, F. Structure of an EF-Tu Complex with a Thiazolyl Peptide Antibiotic Determined at 2.35 Å Resolution: Atomic Basis for GE2270A Inhibition of EF-Tu. Biochemistry 2000, 39, 37–45.
34. Nissen, P.; Kjeldgaard, M.; Thirup, S.; Polekhina, G.; Reshetnikova, L.; Clark, B. F. C.; Nyborg, J. Crystal Structure of the Ternary Complex of Phe-tRNA$^{\mathrm{Phe}}$, EF-Tu, and a GTP Analog. *Science* 1995, 270, 1464–1472.
35. Kohler, R.; Mooney, R. A.; Mills, D. J.; Landick, R.; Cramer, P. Architecture of a Transcribing-Translating Expressome. *Science* 2017, 356, 194–197.
36. Gualerzi, C. O.; Pon, C. L. Initiation of mRNA Translation in Bacteria: Structural and Dynamic Aspects. *Cell Mol Life Sci* 2015, 72, 4341–4367.
37. Chen, B.; Kaledhonkar, S.; Sun, M.; Shen, B.; Lu, Z.; Barnard, D.; Lu, T. M.; Gonzalez, R.; Frank, J. Structural Dynamics of Ribosome Subunit Association Studied by Mixing-Spraying Time-Resolved Cryogenic Electron Microscopy. *Structure* 2015, 23, 1097–1105.

38. Diaconu, M.; Kothe, U.; Schlünzen, F.; Fischer, N.; Harms, J. M.; Tonevitsky, A. G.; Stark, H.; Rodnina, M. V.; Wahl, M. C. Structural Basis for the Function of the Ribosomal L7/12 Stalk in Factor Binding and GTpase Activation. *Cell* 2005, 121, 991–1004.
39. Fischer, N.; Neumann, P.; Bock, L. V.; Maracci, C.; Wang, Z.; Paleskava, A.; Konevega, A. L.; Schröder, G. F.; Grubmüller, H.; Ficner, R.; Rodnina, M. V.; Stark, H. The Pathway to GTPase Activation of Elongation Factor SelB on the Ribosome. *Nature* 2016, 540, 80–85.
40. Liljas, A.; Ehrenberg, M.; Åqvist, J. Comment on "The Mechanism for Activation of GTP Hydrolysis on the Ribosome." *Science* 2011, 333.
41. Daviter, T.; Wieden, H. J.; Rodnina, M. V. Essential Role of Histidine 84 in Elongation Factor Tu for the Chemical Step of GTP Hydrolysis on the Ribosome. *J Mol Biol* 2003, 332, 689–699.
42. Schweins, T.; Geyer, M.; Scheffzek, K.; Warshel, A.; Kalbitzer, H. R.; Wittinghofer, A. Substrate-Assisted Catalysis as a Mechanism for GTP Hydrolysis of P21ras and Other GTP-Binding Proteins. *Nat Struct Biol* 1995, 2, 36–44.
43. Kothe, U.; Rodnina, M. V. Delayed Release of Inorganic Phosphate from Elongation Factor Tu Following GTP Hydrolysis on the Ribosome. *Biochemistry* 2006, 45, 12767–12774.
44. Algire, M. A.; Maag, D.; Lorsch, J. R. Pi Release from EIF2, Not GTP Hydrolysis, Is the Step Controlled by Start-Site Selection during Eukaryotic Translation Initiation. *Mol Cell* 2005, 20, 251–262.
45. Rodnina, M. V.; Savelsbergh, A.; Matassova, N. B.; Katunin, V. I.; Semenkov, Y. P.; Wintermeyer, W. Thiostrepton Inhibits the Turnover but Not the GTPase of Elongation Factor G on the Ribosome. *Proc Natl Acad Sci* 1999, 96, 9586–9590.
46. Noel, J. K.; Whitford, P. C. How EF-Tu Can Contribute to Efficient Proofreading of Aa-TRNA by the Ribosome. *Nat Commun* 2016, 7, 1–10.
47. Ranjan, N.; Rodnina, M. V. Thio-Modification of TRNA at the Wobble Position as Regulator of the Kinetics of Decoding and Translocation on the Ribosome. *J Am Chem Soc* 2017, 139, 5857–5864.
48. Rodnina, M. V.; Fischer, N.; Maracci, C.; Stark, H. Ribosome Dynamics during Decoding. *Philos T Roy Soc B* 2017, 372.
49. Sievers, A.; Beringer, M.; Rodnina, M. V.; Wolfenden, R. The Ribosome as an Entropy Trap. *Proc Natl Acad Sci* 2004, 101, 7897–7901.

50. Polikanov, Y. S.; Steitz, T. A.; Innis, C. A. A Proton Wire to Couple Aminoacyl-TRNA Accommodation and Peptide-Bond Formation on the Ribosome. *Nat Struct Mol Biol* 2014, 21, 787–793.

51. Wallin, G.; Åqvist, J. The Transition State for Peptide Bond Formation Reveals the Ribosome as a Water Trap. *Proc Natl Acad Sci* 2010, 107, 1888–1893.

52. Wasserman, M. R.; Alejo, J. L.; Altman, R. B.; Blanchard, S. C. Multiperspective SmFRET Reveals Rate-Determining Late Intermediates of Ribosomal Translocation. *Nat Struct Mol Biol* 2016, 23, 333–341.

53. Belardinelli, R.; Sharma, H.; Caliskan, N.; Cunha, C. E.; Peske, F.; Wintermeyer, W.; Rodnina, M. V. Choreography of Molecular Movements during Ribosome Progression along mRNA. *Nat Struct Mol Biol* 2016, 23, 342–348.

54. Savelsbergh, A.; Mohr, D.; Kothe, U.; Wintermeyer, W.; Rodnina, M. V. Control of Phosphate Release from Elongation Factor G by Ribosomal Protein L7/12. *EMBO J* 2005, 24, 4316–4323.

55. Belardinelli, R.; Sharma, H.; Peske, F.; Wintermeyer, W.; Rodnina, M. V. Translocation as Continuous Movement through the Ribosome. *RNA Biol* 2016, 13, 1197–1203.

56. Thirup, S. S.; Van, L. B.; Nielsen, T. K.; Knudsen, C. R. Structural Outline of the Detailed Mechanism for Elongation Factor Ts-Mediated Guanine Nucleotide Exchange on Elongation Factor Tu. *J Struct Biol* 2015, 191, 10–21.

57. Burnett, B. J.; Altman, R. B.; Ferguson, A.; Wasserman, M. R.; Zhou, Z.; Blanchard, S. C. Direct Evidence of an Elongation Factor-Tu/Ts·GTP·Aminoacyl-tRNA Quaternary Complex. *J Biol Chem* 2014, 289, 23917-23927.

58. Burnett, B. J.; Altman, R. B.; Ferrao, R.; Alejo, J. L.; Kaur, N.; Kanji, J.; Blanchard, S. C. Elongation Factor Ts Directly Facilitates the Formation and Disassembly of the *Escherichia Coli* Elongation Factor Tu·GTP·Aminoacyl-tRNA Ternary Complex. *J Biol Chem* 2013, 288, 13917–13928.

59. Mccoy, L. S.; Xie, Y.; Tor, Y. Antibiotics That Target Protein Synthesis. *Interdiscip Rev RNA* 2011, 2, 209–232.

60. Schluenzen, F.; Takemoto, C.; Wilson, D. N.; Kaminishi, T.; Harms, J. M.; Hanawa-Suetsugu, K.; Szaflarski, W.; Kawazoe, M.; Shirouzo, M.; Nierhaus, K. H.; Yokoyama, S.; Fucini, P. The Antibiotic Kasugamycin Mimics mRNA Nucleotides to Destabilize tRNA Binding and Inhibit Canonical Translation Initiation. *Nat Struct Mol Biol* 2006, 13, 871–878.

61. Jenner, L.; Starosta, A. L.; Terry, D. S.; Mikolajka, A.; Filonava, L.; Yusupov, M.; Blanchard, S. C.; Wilson, D. N.; Yusupova, G. Structural Basis for Potent Inhibitory Activity of the Antibiotic Tigecycline during Protein Synthesis. *Proc Natl Acad Sci* 2013, 110, 3812–3816.

62. Lin, J.; Zhou, D.; Steitz, T. A.; Polikanov, Y. S.; Gagnon, M. G. Ribosome-Targeting Antibiotics: Modes of Action, Mechanisms of Resistance, and Implications for Drug Design. *Annu Rev Biochem* 2018, 87, 451-478.

63. Prezioso, S. M.; Brown, N. E.; Goldberg, J. B. Elfamycins: Inhibitors of Elongation Factor-Tu. *Mol Microbiol* 2017, 106, 22.

64. Mullane, K.; Lee, C.; Bressler, A.; Buitrago, M.; Weiss, K.; Dabovic, K.; Praestgaard, J.; Leeds, J. A.; Blais, J.; Pertel, P. Multicenter, Randomized Clinical Trial to Compare the Safety and Efficacy of Lff571 and Vancomycin for Clostridium Difficile Infections. *Antimicrob Agents Chemother* 2015, 59, 1435–1440.

65. Fabbretti, A.; He, C. G.; Gaspari, E.; Maffioli, S.; Brandi, L.; Spurio, R.; Sosio, M.; Jabes, D.; Donadio, S. A Derivative of the Thiopeptide GE2270A Highly Selective against *Propionibacterium Acnes*. *Antimicrob Agents Chemother* 2015, 59, 4560–4568.

66. Wolf, H.; Chinali, G.; Parmeggiani, A. Mechanism of the Inhibition of Protein Synthesis by Kirromycin. *Eur J Biochem* 1977, 75, 67–75.

67. Wolf, H.; Chinali, G.; Parmeggiani, A. Kirromycin, an Inhibitor of Protein Biosynthesis That Acts on Elongation Factor Tu. *Proc Natl Acad Sci* 1974, 71, 4910–4914.

68. Parmeggiani, A.; Nissen, P. Elongation Factor Tu-Targeted Antibiotics: Four Different Structures, Two Mechanisms of Action. *FEBS Lett* 2006, 580, 4576–4581.

69. Parmeggiani, A.; Krab, I. M.; Okamura, S.; Nielsen, R. C.; Nyborg, J.; Nissen, P. Structural Basis of the Action of Pulvomycin and GE2270 A on Elongation Factor Tu. *Biochemistry* 2006, 45, 6846–6857.

70. Wolf, H.; Assmann, D.; Fischer, E. Pulvomycin, an Inhibitor of Protein Biosynthesis Preventing Ternary Complex Formation between Elongation Factor Tu, GTP, and Aminoacyl-tRNA. *Proc Natl Acad Sci* 1978, 75, 5324–5328.

71. Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. A Three-Dimensional Model of the Myoglobin Molecule Obtained by x-Ray Analysis. *Nature* 1958, 18, 662–666.

72. Perutz, M. F.; Rossmann, M. G.; Cullis, A. F.; Muirhead, H.; Will, G.; North, A. C. T. Structure of Haemoglobin: A Three-Dimensional Fourier Synthesis at 5.5-A. Resolution, Obtained by X-Ray Analysis. *Nature* 1960, 185, 416–422.

73. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* 2021, 596, 583-589.

74. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Zídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res* 2022, 50, 439–444.

75. Varadi, M.; Bertoni, D.; Magana, P.; Paramval, U.; Pidruchna, I.; Radhakrishnan, M.; Tsenkov, M.; Nair, S.; Mirdita, M.; Yeo, J.; Kovalevskiy, O.; Tunyasuvunakool, K.; Laydon, A.; Žídek, A.; Tomlinson, H.; Hariharan, D.; Abrahamson, J.; Green, T.; Jumper, J.; Birney, E.; Steinegger, M.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database in 2024: Providing Structure Coverage for over 214 Million Protein Sequences. *Nucleic Acids Res* 2024, 52, 368–375.

76. McPherson, A.; Gavira, J. A. Introduction to Protein Crystallization. *Acta Crystallogr F Struct Biol Commun* 2014, 70, 2-20.

77. Helliwell, J. R. Protein Crystal Perfection and Its Application. *Acta Crystallogr D Struct Biol Commun* 2005, 61, 793–798.

78. Kermani, A. A.; Kermani, A. A. A Guide to Membrane Protein X-Ray Crystallography. *FEBS J* 2021, 288, 5788–5804.

79. Hinsen, K. Structural Flexibility in Proteins: Impact of the Crystal Environment. *Bioinformatics* 2008, 24 521–528.

80. Bhella, D. Cryo-Electron Microscopy: An Introduction to the Technique, and Considerations When Working to Establish a National Facility. *Biophys Rev* 2019, 11, 515.

81. Adrian, M.; Dubochet, J.; Lepault, J.; McDowall, A. W. Cryo-Electron Microscopy of Viruses. *Nature* 1984, 308, 32–36.

82. Yip, K. M.; Fischer, N.; Paknia, E.; Chari, A.; Stark, H. Atomic-Resolution Protein Structure Determination by Cryo-EM. *Nature* 2020, 587, 157-161.

83. Lyumkis, D. Challenges and Opportunities in Cryo-EM Single-Particle Analysis. *J Biol Chem* 2019, 294, 5181.

84. Benjin, X.; Ling, L. Developments, Applications, and Prospects of Cryo-Electron Microscopy. *Protein Sci* 2020, 29, 872–882.

85. Castells-Graells, R.; Meador, K.; Arbing, M. A.; Sawaya, M. R.; Gee, M.; Cascio, D.; Gleave, E.; Debreczeni, J.; Breed, J.; Leopold, K.; Patel, A.; Jahagirdar, D.; Lyons, B.; Subramaniam, S.; Phillips, C.; Yeates, T. O. Cryo-EM Structure Determination of Small Therapeutic Protein Targets at 3 Å-Resolution Using a Rigid Imaging Scaffold. *Proc Natl Acad Sci* 2023, 120.

86. Wang, H. W.; Wang, J. W. How Cryo-electron Microscopy and X-ray Crystallography Complement Each Other. *Protein Sci* 2017, 26, 32-39.

87. Bothwell, J. H. F.; Griffin, J. L. An Introduction to Biological Nuclear Magnetic Resonance Spectroscopy. *Biol Rev* 2011, 86, 493–510.

88. Cala, O.; Guillière, F.; Krimm, I. NMR-Based Analysis of Protein-Ligand Interactions. *Anal Bioanal Chem* 2014, 406, 943–956.

89. Griffiths, D. J. Introduction to Quantum Mechanics. *Cambridge University Press* 2017.

90. Balci, M. Dynamic NMR Spectroscopy. Basic 1H- and 13C-NMR. *Spectroscopy* 2005, 213–231.

91. Poulsen, F. M. A Brief Introduction to NMR Spectroscopy of Proteins. 2002.

92. Ernst, R. R.; Bodenhausen, G.; Wokaun, A. Principles of Nuclear Magnetic Resonance in One and Two Dimensions. *Oxford University Press* 1990.

93. Tugarinov, V.; Muhandiram, R.; Ayed, A.; Kay, L. E. Four-Dimensional NMR Spectroscopy of a 723-Residue Protein: Chemical Shift Assignments and Secondary Structure of Malate Synthase G. *J Am Chem Soc* 2002, 124, 10025–10035.

94. Religa, T. L.; Sprangers, R.; Kay, L. E. Dynamic Regulation of Archaeal Proteasome Gate Opening as Studied by Trosy NMR. *Science* 2010, 328, 98–102.

95. Serber, Z.; Corsini, L.; Durst, F.; Dötsch, V. In-Cell NMR Spectroscopy. *Methods Enzymol* 2005, 394, 17–41.

96. Reinknecht, C.; Riga, A.; Rivera, J.; Snyder, D. A. Patterns in Protein Flexibility: A Comparison of NMR "Ensembles", MD Trajectories, and Crystallographic B-Factors. *Molecules* 2021, 26.

97. Hu, Y.; Cheng, K.; He, L.; Zhang, X.; Jiang, B.; Jiang, L.; Li, C.; Wang, G.; Yang, Y.; Liu, M. NMR-Based Methods for Protein Analysis. *Anal Chem* 2021, 93, 1866–1879.

98. Davis, E. A.; Falconer, I. J.J. Thompson And The Discovery Of The Electron. *CRC Press* 2002.

99. Aston, F. W. A Positive Ray Spectrograph. *Lond Edinb Dubl Phil Mag* 1919, 38, 707–714.

100. Schultz, J. C.; Hack, C. A.; Benner, W. H. Mass Determination of Megadalton-DNA Electrospray Ions Using Charge Detection Mass Spectrometry. *J Am Soc Mass Spectrom* 1998, 9, 305–313.

101. Dempster, A. J. Isotopic Constitution of Uranium. *Nature* 1935, 136, 180–180.

102. Karas, M.; Bachmann, D.; Hillenkamp, F. Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Anal Chem* 1985, 57, 2935–2939..

103. Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. Matrix-Assisted Ultraviolet Laser Desorption of Non-Volatile Compounds. *Int J Mass Spectrom Ion Process* 1987, 78, 53–68.

104. Møller, C.; Sprenger, R. R.; Stürup, S.; Højrup, P. Determination of the Binding Sites for Oxaliplatin on Insulin Using Mass Spectrometry-Based Approaches. *Anal Bioanal Chem* 2011, 401, 1623–1633.

105. Berezovskaya, Y.; Armstrong, C. T.; Boyle, A. L.; Porrini, M.; Woolfson, D. N.; Barran, P. E. Metal Binding to a Zinc-Finger Peptide: A Comparison between Solution and the Gas Phase. *Chem Commun* 2010, 47, 412–414.

106. Esswein, S. T.; Florance, H. V.; Baillie, L.; Lippens, J.; Barran, P. E. A Comparison of Mass Spectrometry Based Hydrogen Deuterium Exchange Methods for Probing the Cyclophilin A Cyclosporin Complex. *J Chromatogr A* 2010, 1217, 6709–6717.

107. Hoffmann Edmond de, S. V. Mass Spectrometry: Principles and Applications, 3rd Edition. *John Wiley & Sons* 2007.

108. Loo, J. A. Studying Noncovalent Protein Complexes By Electrospray Ionization Mass Spectrometry. *Mass Spectrom Rev* 1996, 16, 1-23.

109. Zeleny, J. The Electrical Discharge from Liquid Points, and a Hydrostatic Method of Measuring the Electric Intensity at Their Surfaces. *Phys Rev* 1914, 3, 69.

110. Taylor, G. Disintegration of Water Drops in an Electric Field. *Proc R Soc Lond A Math Phys Sci* 1964, 280, 383–397.

111. Rayleigh, L. On the Equilibrium of Liquid Conducting Masses Charged with Electricity. *Lond Edinb Dubl Phil Mag* 1882, 14, 184–186.

112. Gomez, A.; Tang, K.; Gomeza, A.; Tangb, K. Charge and Fission of Droplets in Electrostatic Sprays. *Phys Fluids* 1994, 6, 404–414.

113. Banerjee, S.; Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Int J Anal Chem* 2012, 2012, 282574.

114. Fernandez De La Mora, J. Electrospray Ionization of Large Multiply Charged Species Proceeds via Dole⊠ Charged Residue Mechanism. *Anal Chim Acta* 2000, 406, 93–104.

115. Metwally, H.; Duez, Q.; Konermann, L. Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry. *Anal Chem* 2018, 90, 10069–10077.

116. Kebarle, P.; Verkcerk, U. H. Electrospray: From Ions in Solution to Ions in the Gas Phase, What We Know Now. *Mass Spectrom Rev* 2009, 28, 898–917.

117. Zampronio, C. G.; Giannakopulos, A. E.; Zeller, M.; Bitziou, E.; Macpherson, J. V.; Derrick, P. J. Production and Properties of Nanoelectrospray Emitters Used in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Implications for Determination of Association Constants for Noncovalent Complexes. *Anal Chem* 2004, 76, 5172–5179.

118. Benesch, J. L. P.; Ruotolo, B. T.; Simmons, D. A.; Robinsons, C. V. Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. *Chem Rev* 2007, 107, 3544–3567.

119. McCullagh, James.; Oldham, Neil. Mass Spectrometry. *Oxford University Press* 2019.

120. Metwally, H.; McAllister, R. G.; Konermann, L. Exploring the Mechanism of Salt-Induced Signal Suppression in Protein Electrospray Mass Spectrometry Using Experiments and Molecular Dynamics Simulations. *Anal Chem* 2015, 87, 2434–2442.

121. Paul, W. Electromagnetic Traps for Charged and Neutral Particles. *Rev Mod Phys* 1990, 62, 531.

122. Wolff, M. M.; Stephens, W. E. A Pulsed Mass Spectrometer with Time Dispersion. *RScI* 1953, 24, 616–617.

123. Comisarow, M. B.; Marshall, A. G. Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chem Phys Lett* 1974, 25, 282–283.

124. Marshall, A. G. Convolution Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chem Phys Lett* 1979, 63, 515–518.

125. Liu, X.; Dekker, L. J. M.; Wu, S.; Vanduijn, M. M.; Luider, T. M.; Tolić, N.; Kou, Q.; Dvorkin, M.; Alexandrova, S.; Vyatkina, K.; Paša-Tolić, L.; Pevzner, P. A. De Novo Protein Sequencing by Combining Top-down and Bottom-up Tandem Mass Spectra. *J Proteome Res* 2014, 13, 3241–3248.

126. Haler, J. R. N.; Massonnet, P.; Far, J.; de la Rosa, V. R.; Lecomte, P.; Hoogenboom, R.; Jérôme, C.; De Pauw, E. Gas-Phase Dynamics of Collision Induced Unfolding, Collision Induced Dissociation, and Electron Transfer Dissociation-Activated Polymer Ions. *J Am Soc Mass Spectrom* 2019,30 , 563–572.

127. Sleno, L.; Volmer, D. A. Ion Activation Methods for Tandem Mass Spectrometry. *J Mass Spectrom* 2004, 39, 1091–1112.

128. Khatun, J.; Ramkissoon, K.; Giddings, M. C. Fragmentation Characteristics of Collision-Induced Dissociation in MALDI TOF/TOF Mass Spectrometry. *Anal Chem* 2007, 79, 3032–3040.

129. Light-Wahl, K. J.; Schwartz, B. L.; Smith, R. D. Observation of the Noncovalent Quaternary Associations of Proteins by Electrospray Ionization Mass Spectrometry. *J. Am. Chem. Soc* 1994, 116, 5271–5278.

130. Karch, K. R.; Snyder, D. T.; Harvey, S. R.; Wysocki, V. H. Native Mass Spectrometry: Recent Progress and Remaining Challenges. *Annu Rev Biophys* 2022, 51, 157–179.

131. Van Den Heuvel, R. H. H.; Heck, A. J. R. Native Protein Mass Spectrometry: From Intact Oligomers to Functional Machineries. *Curr Opin Chem Biol* 2004, 8, 519–526.

132. Chernushevich, I. V.; Thomson, B. A. Collisional Cooling of Large Ions in Electrospray Mass Spectrometry. *Anal Chem* 2004, 76, 1754–1760.

133. Breuker, K.; McLafferty, F. W. Stepwise Evolution of Protein Native Structure with Electrospray into the Gas Phase, 10−12 to 102 s. *Proc Natl Acad Sci* 2008, 105,

134. Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. Biochemistry: Evidence for Macromolecular Protein Rings in the Absence of Bulk Water. *Science* 2005, 310, 1658–1661.

135. Oomens, J.; Polfer, N.; Moore, D. T.; Van Der Meer, L.; Marshall, A. G.; Eyler, J. R.; Meijer, G.; Von Helden, G. Charge-State Resolved Mid-Infrared Spectroscopy of a Gas-Phase Protein. *Phys Chem Chem Phys* 2005, 7, 1345–1348.

136. Ruotolo, B. T.; Robinson, C. V. Aspects of Native Proteins Are Retained in Vacuum. *Curr Opin Chem Biol* 2006, 10, 402–408.

137. Chen, Z.; Kellie, J. F.; Hottenstein, C. S.; Szapacs, M. E. Native High-Resolution Mass Spectrometry Analysis of Noncovalent Protein Complexes up to 450 KDa. *Bioanalysis* 2020, 12, 1353–1362.

138. Vimer, S.; Ben-Nissan, G.; Morgenstern, D.; Kumar-Deshmukh, F.; Polkinghorn, C.; Quintyn, R. S.; Vasil⊠ev, Y. V.; Beckman, J. S.; Elad, N.; Wysocki, V. H.; Sharon, M. Comparative Structural Analysis of 20S Proteasome Ortholog Protein Complexes by Native Mass Spectrometry. *ACS Cent Sci* 2020, 6, 573–588.

139. Gupta, K.; Li, J.; Liko, I.; Gault, J.; Bechara, C.; Wu, D.; Hopper, J. T. S.; Giles, K.; Benesch, J. L. P.; Robinson, C. V. Identifying Key Membrane Protein Lipid Interactions Using Mass Spectrometry. *Nat Prot* 2018, 13, 1106–1120.

140. Chorev, D. S.; Baker, L. A.; Wu, D.; Beilsten-Edmands, V.; Rouse, S. L.; Zeev-Ben-Mordehai, T.; Jiko, C.; Samsudin, F.; Gerle, C.; Khalid, S.; Stewart, A. G.; Matthews, S. J.; Grünewald, K.; Robinson, C. V. Protein Assemblies Ejected Directly from Native Membranes Yield Complexes for Mass Spectrometry. *Science* 2018, 362, 829–834.

141. Van De Waterbeemd, M.; Fort, K. L.; Boll, D.; Reinhardt-Szyba, M.; Routh, A.; Makarov, A.; Heck, A. J. R. High-Fidelity Mass Analysis Unveils Heterogeneity in Intact Ribosomal Particles. *Nat Methods* 2017, 14, 283–286.

142. Cumeras, R.; Figueras, E.; Davis, C. E.; Baumbach, J. I.; Gràcia, I. Review on Ion Mobility Spectrometry. Part 1: Current Instrumentation. *Analyst* 2015, 140, 1376–1390.

143. Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. Naked Protein Conformations: Cytochrome c in the Gas Phase. *J Am Chem Soc* 1995, 117, 10141–10142.

144. Mason, E. A.; Schamp, H. W. Mobility of Gaseous Lons in Weak Electric Fields. *Ann Phys* 1958, 4, 233–270.

145. Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. Ion Mobility-Mass Spectrometry Analysis of Large Protein Complexes. *Nat Prot* 2008, 3, 1139-1152.

146. Shvartsburg, A. A.; Smith, R. D. Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal Chem* 2008, 80, 9689–9699.

147. May, J. C.; McLean, J. A. The Influence of Drift Gas Composition on the Separation Mechanism in Traveling Wave Ion Mobility Spectrometry: Insight from Electrodynamic Simulations. *Int J Ion Mobil Spec* 2013, 16, 85–94.

148. Al-jabiry, A.; Palmer, M.; Langridge, J.; Bellamy-Carter, J.; Robinson, D.; Oldham, N. J. Combined Chemical Modification and Collision Induced Unfolding Using Native Ion Mobility-Mass Spectrometry Provides Insights into Protein Gas-Phase Structure. *Chem Eur J* 2021, 27, 13783–13792.

149. Dixit, S. M.; Polasky, D. A.; Ruotolo, B. T. Collision Induced Unfolding of Isolated Proteins in the Gas Phase: Past, Present, and Future. *Curr Opin Chem Biol* 2018, 42, 93–100.

150. Hopper, J. T. S.; Oldham, N. J. 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, 1851–1858.

151. Bellamy-Carter, J.; O'Grady, L.; Passmore, M.; Jenner, M.; Oldham, N. J. Decoding Protein Gas-Phase Stability with Alanine Scanning and Collision-Induced Unfolding Ion Mobility Mass Spectrometry. *Anal Sens* 2021, 1, 6–6.

152. Tian, Y.; Han, L.; Buckner, A. C.; Ruotolo, B. T. Collision Induced Unfolding of Intact Antibodies: Rapid Characterization of Disulfide Bonding Patterns, Glycosylation, and Structures. *Anal Chem* 2015, 87, 11509–11515.

153. Steiner, R. F.; Albaugh, S.; Fenselau, C.; Murphy, C.; Vestling, M. A Mass Spectrometry Method for Mapping the Interface Topography of Interacting Proteins, Illustrated by the Melittin-Calmodulin System. *Anal Biochem* 1991, 196, 120–125.

154. Katta, V.; Chait, B. T.; Carr, S. Conformational Changes in Proteins Probed by Hydrogen-Exchange Electrospray-Ionization Mass Spectrometry. *Rapid Commun Mass Spectrom* 1991, 5, 214–217.

155. Zhang, Z.; Smith, D. L. Determination of Amide Hydrogen Exchange by Mass Spectrometry: A New Tool for Protein Structure Elucidation. *Protein Sci* 1993, 2, 522–531.

156. Dupree, E. J.; Jayathirtha, M.; Yorkey, H.; Mihasan, M.; Petre, B. A.; Darie, C. C. A Critical Review of Bottom-Up Proteomics: The Good, the Bad, and the Future of This Field. *Proteomes* 2020, 8, 14.

157. Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jørgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D. Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat Methods* 2019 16:7 2019, 16, 595–602.

158. Konermann, L.; Pan, J.; Liu, Y. H. Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem Soc Rev* 2011, 40, 1224–1234.

159. Engen, J. R. Analysis of Protein Conformation and Dynamics by Hydrogen/Deuterium Exchange MS. *Anal Chem* 2009, 81, 7870–7875.

160. Chu, I. Te; Pielak, G. J. Using NMR-Detected Hydrogen-Deuterium Exchange to Quantify Protein Stability in Cosolutes, under Crowded Conditions *in Vitro* and in Cells. *Magnetic Resonance Letters* 2023, 3, 319–326.

161. Peacock, R. B.; Komives, E. A. Hydrogen/Deuterium Exchange and Nuclear Magnetic Resonance Spectroscopy Reveal Dynamic Allostery on Multiple Time Scales in the Serine Protease Thrombin. *Biochemistry* 2021, 60, 3441–3448.

162. Watson, M. J.; Harkewicz, R.; Hodge, E. A.; Vorauer, C.; Palmer, J.; Lee, K. K.; Guttman, M. Simple Platform for Automating Decoupled LC-MS Analysis of Hydrogen/Deuterium Exchange Samples. *J Am Soc Mass Spectrom* 2021, 32, 597–600.

163. Van Schaick, G.; Domínguez-Vega, E.; Gstöttner, C.; Van Den Berg-Verleg, J. H.; Schouten, O.; Akeroyd, M.; Olsthoorn, M. M. A.; Wuhrer, M.; Heck, A. J. R.; Abello, N.; Franc, V. Native Structural and Functional Proteoform Characterization of the Prolyl-Alanyl-Specific Endoprotease EndoPro from *Aspergillus Niger*. *J Proteome Res* 2021, 20, 4875–4885.

164. Tsiatsiani, L.; Akeroyd, M.; Olsthoorn, M.; Heck, A. J. R. *Aspergillus Niger* Prolyl Endoprotease for Hydrogen-Deuterium Exchange Mass Spectrometry and Protein Structural Studies. *Anal Chem* 2017, 89, 7966–7973.

165. Kish, M.; Smith, V.; Lethbridge, N.; Cole, L.; Bond, N. J.; Phillips, J. J. Online Fully Automated System for Hydrogen/Deuterium-Exchange Mass Spectrometry with Millisecond Time Resolution. *Anal Chem* 2023, 95, 5000–5008.

166. Pascal, B. D.; Willis, S.; Lauer, J. L.; Landgraf, R. R.; West, G. M.; Marciano, D.; Novick, S.; Goswami, D.; Chalmers, M. J.; Griffin, P. R. HDXWorkbench: Software for the Analysis of H/D Exchange MS Data. *J Am Soc Mass Spectrom* 2012, 23, 1512–1521.

167. Johnson, R. S.; Krylov, D.; Walsh, K. A. Proton Mobility within Electrosprayed Peptide Ions. *J Mass Spectrom* 1995, 30, 386–387.

168. Rand, K. D.; Zehl, M.; Jørgensen, T. J. D. Measuring the Hydrogen/Deuterium Exchange of Proteins at High Spatial Resolution by Mass Spectrometry: Overcoming Gas-Phase Hydrogen/Deuterium Scrambling. *Acc Chem Res* 2014, 47, 3018–3027.

169. Zehl, M.; Rand, K. D.; Jensen, O. N.; Jorgensen, T. J. D. Electron Transfer Dissociation Facilitates the Measurement of Deuterium Incorporation into Selectively Labelled Peptides with Single Residue Resolution. *J Am Chem Soc* 2008, 130, 17453–17459.

170. Wang, L.; Chance, M. R. Protein Footprinting Comes of Age: Mass Spectrometry for Biophysical Structure Assessment. *Mol Cell Proteom* 2017, 16, 706–716.

171. Tullius, T. D.; Dombroski, B. A. Hydroxyl Radical "Footprinting": High-Resolution Information about DNA-Protein Contacts and Application to λ Repressor and Cro Protein. *Proc Natl Acad Sci* 1986, 83, 5469–5473.

172. Maleknia, S. D.; Brenowitz, M.; Chance, M. R. Millisecond Radiolytic Modification of Peptides by Synchrotron X-Rays Identified by Mass Spectrometry. *Anal Chem* 1999, 71, 3965–3973.

173. McKenzie-Coe, A.; Montes, N. S.; Jones, L. M. Hydroxyl Radical Protein Footprinting: A Mass Spectrometry-Based Structural Method for Studying the Higher Order Structure of Proteins. *Chem Rev* 2022, 122, 7532–7561.

174. Pogozelski, W. K.; Mcneese, T. J.; Tullius, T. D. What Species Is Responsible for Strand Scission in the Reaction of [Fe(II)EDTA]^{2} and H_2O_2 with DNA? *J. Am. Chem. Soc* 1995, 117, 6428–6433.

175. Niu, B.; Gross, M. L. MS-Based Hydroxyl Radical Footprinting: Methodology and Application of Fast Photochemical Oxidation of Proteins (FPOP). *Mass Spectrometry-Based Chemical Proteomics* 2019, 363–416.

176. Hambly, D.; Gross, M. L. Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *J Am Soc Mass Spectrom* 2005, 16, 2057–2063.

177. Aye, T. T.; Low, T. Y.; Sze, S. K. Nanosecond Laser-Induced Photochemical Oxidation Method for Protein Surface Mapping with Mass Spectrometry. *Anal Chem 2005*, 77, 5814–5822.

178. Xu, G.; Chance, M. R. Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem Rev* 2007, 107, 3514–3543.

179. Rojas Ramírez, C.; Espino, J. A.; Jones, L. M.; Polasky, D. A.; Nesvizhskii, A. I. Efficient Analysis of Proteome-Wide FPOP Data by FragPipe. *Anal Chem* 2023, 95, 16131–16137.

180. Cheng, M.; Zhang, B.; Cui, W.; Gross, M. L. Laser-Initiated Radical Trifluoromethylation of Peptides and Proteins and Its Application to Mass Spectrometry-Based Protein Footprinting. *Angew Chem Int Ed Engl* 2017, 56, 14007.

181. Cheng, M.; Asuru, A.; Kiselar, J.; Mathai, G.; Chance, M. R.; Gross, M. L. Fast Protein Footprinting by X-Ray Mediated Radical Trifluoromethylation. *J Am Soc Mass Spectrom* 2020, 31, 1019–1024.

182. Jain, R.; Farquhar, E. R.; Dhillon, N. S.; Jeon, N.; Chance, M. R.; Kiselar, J. Optimized Trifluoromethylation and OH Radical Labeling with Radiolysis Provide Two-Fold Enhancement in the Resolution of Hydroxyl Radical Protein Footprinting. *ChemRxiv,* 2024.

183. Sauers, R. R. Singlet Carbene Stability: Linear Free-Energy Analysis of Substituent Effects. *MI:M Publishing* 2014, 376–383.

184. Park, S.; Kim, C. E.; Jeong, J.; Ryu, H.; Maeng, C.; Kim, D.; Baik, M. H.; Lee, P. H. Selective Ring Expansion and C–H Functionalization of Azulenes. *Nat Commun* 2023, 14, 1–9.

185. Huang, M. Y.; Zhu, S. F. Uncommon Carbene Insertion Reactions. *Chem Sci* 2021, 12, 15790–15801.

186. Das, J. Aliphatic Diazirines as Photoaffinity Probes for Proteins: Recent Developments. *Chem Rev* 2011, 111, 4405–4417.

187. Ge, S. S.; Chen, B.; Wu, Y. Y.; Long, Q. S.; Zhao, Y. L.; Wang, P. Y.; Yang, S. Current Advances of Carbene-Mediated Photoaffinity Labelling in Medicinal Chemistry. *RSC Adv* 2018, 8, 29428–29454.

188. Djordjevic, I.; Wicaksono, G.; Solic, I.; Steele, T. W. J. Diazoalkane Decay Kinetics from UVA-Active Protein Labelling Molecules: Trifluoromethyl Phenyl Diazirines. *Results Chem* 2020, 2, 100066.

189. Preston, G. W.; Wilson, A. J. Photo-Induced Covalent Cross-Linking for the Analysis of Biomolecular Interactions. *Chem Soc Rev* 2013, 42, 3289–3301.

190. Brunner, J.; Senn, H.; Richards, F. M. 3-Trifluoromethyl-3-Phenyldiazirine. A New Carbene Generating Group for Photolabeling Reagents. *J Biol Chem* 1980, 255,

191. West, A. V.; Muncipinto, G.; Wu, H. Y.; Huang, A. C.; Labenski, M. T.; Jones, L. H.; Woo, C. M. Labelling Preferences of Diazirines with Protein Biomolecules. *J Am Chem Soc* 2021, 143, 6691–6700.

192. Richards, F. M.; Lamed, R.; Wynn, R.; Patel, D.; Olack, G. Methylene as a Possible Universal Footprinting Reagent That Will Include Hydrophobic Surface Areas: Overview and Feasibility: Properties of Diazirine as a Precursor. *Protein Sci* 2000, 9, 2506–2517.

193. Jumper, C. C.; Schriemer, D. C. Mass Spectrometry of Laser-Initiated Carbene Reactions for Protein Topographic Analysis. *Anal Chem* 2011, 83, 2913–2920.

194. Jumper, C. C.; Bomgarden, R.; Rogers, J.; Etienne, C.; Schriemer, D. C. High-Resolution Mapping of Carbene-Based Protein Footprints. *Anal Chem* 2012, 84, 4411–4418.

195. Manzi, L.; Barrow, A. S.; Scott, D.; Layfield, R.; Wright, T. G.; Moses, J. E.; Oldham, N. J. Carbene Footprinting Accurately Maps Binding Sites in Protein–Ligand and Protein–Protein Interactions. *Nat Commun* 2016, 7, 1–9.

196. Lloyd, J. R.; Hogan, A.; Paschalis, V.; Bellamy-Carter, J.; Bottley, A.; Seymour, G. B.; Hayes, C. J.; Oldham, N. J. Mapping the Interaction between Eukaryotic Initiation Factor 4A (EIF4A) and the Inhibitor Hippuristanol Using Carbene Footprinting and Mass Spectrometry. *Proteomics* 2021, 21, 2000288.

197. Lloyd, J. R.; Biasutto, A.; Dürr, K. L.; Jazayeri, A.; Hopper, J. T. S.; Oldham, N. J. Mapping the Binding Interactions between Human Gasdermin D and Human Caspase-1 Using Carbene Footprinting. *JACS Au* 2023, 3, 2025–2035.

198. Lu, G.; Xu, X.; Li, G.; Sun, H.; Wang, N.; Zhu, Y.; Wan, N.; Shi, Y.; Wang, G.; Li, L.; Hao, H.; Ye, H. Subresidue-Resolution Footprinting of Ligand-Protein Interactions by Carbene Chemistry and Ion Mobility-Mass Spectrometry. *Anal Chem* 2020, 92, 947–956.

199. Hogan, J. M.; Lee, P. S.; Wong, S. C.; West, S. M.; Morishige, W. H.; Bee, C.; Tapia, G. C.; Rajpal, A.; Strop, P.; Dollinger, G. Residue-Level Characterization of Antibody Binding Epitopes Using Carbene Chemical Footprinting. *Anal Chem* 2023, 95, 3922–3931.

200. Zhang, B.; Rempel, D. L.; Gross, M. L. Protein Footprinting by Carbenes on a Fast Photochemical Oxidation of Proteins (FPOP) Platform. *J Am Soc Mass Spectrom* 2016, 27, 552–555.

201. Bellamy-Carter, J.; Oldham, N. J. PepFoot: A Software Package for Semiautomated Processing of Protein Footprinting Data. *J Proteome Res* 2019, 18, 2925-2930.

202. Hopper, J. T. S.; Sokratous, K.; Oldham, N. J. Charge State and Adduct Reduction in Electrospray Ionization-Mass Spectrometry Using Solvent Vapor Exposure. *Anal biochem* 2011, 421, 788-790.

203. Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V; Ruotolo, B. T.; Duijn, van; R Chem, A. J. Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Proc. Natl. Acad. Sci.* 2010, 39, 9557–9565.

204. Salbo, R.; Bush, M. F.; Naver, H.; Campuzano, I.; Robinson, C. V; Pettersson, I.; D Jørgensen, T. J.; Haselmann, K. F. Traveling-Wave Ion Mobility Mass Spectrometry of

Protein Complexes: Accurate Calibrated Collision Cross-Sections of Human Insulin Oligomers. *Rapid Commun Mass Spectrom* 2012, 26, 1181-1193.

205. Eschweiler, J. D.; Rabuck-Gibbons, J. N.; Tian, Y.; Ruotolo, B. T. CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions. *Anal Chem* 2015, 87, 11516–11522.
206. Muth, T.; Vaudel, M.; Barsnes, H.; Martens, L.; Sickmann, A. XTandem Parser: An Open-Source Library to Parse and Analyse X!Tandem MS/MS Search Results. *Proteomics* 2010, 10, 1522–1524.
207. Vaudel, M.; Barsnes, H.; Berven, F. S.; Sickmann, A.; Martens, L. SearchGUI: An Open-Source Graphical User Interface for Simultaneous OMSSA and X!Tandem Searches. *Proteomics* 2011, 11, 996–999.
208. Vaudel, M.; Burkhart, J. M.; Zahedi, R. P.; Oveland, E.; Berven, F. S.; Sickmann, A.; Martens, L.; Barsnes, H. PeptideShaker Enables Reanalysis of MS-Derived Proteomics Data Sets. *Nat Biotech* 2015, 33, 22-24.
209. Holman, J. D.; Tabb, D. L.; Mallick, P. Employing ProteoWizard to Convert Raw Mass Spectrometry Data. *Curr Protoc Bioinformatics* 2014, 46, 13-24.
210. Holm, S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand J Stat* 1979, 46, 13-24.
211. Šidák, Z. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *J Am Stat Assoc* 1967, 62, 626–633.
212. Yang, J.; Zhang, Y. I-TASSER Server: New Development for Protein Structure and Function Predictions. *Nuc Acid Res* 2015, 43, 174-181.
213. Zheng, W.; Zhang, C.; Li, Y.; Pearce, R.; Bell, E. W.; Zhang, Y. Folding Non-Homologous Proteins by Coupling Deep-Learning Contact Maps with I-TASSER Assembly Simulations. *Cell Reports* 2021, 1.
214. Zhou, X.; Zheng, W.; Li, Y.; Pearce, R.; Zhang, C.; Bell, E. W.; Zhang, G.; Zhang, Y. I-TASSER-MTD: A Deep-Learning-Based Platform for Multi-Domain Protein Structure and Function Prediction. *Nat Methods* 2022, 17, 2326-2353.
215. Mirdita, M.; Ovchinnikov, S.; Steinegger, M.; Schütze, K.; Moriwaki, Y.; Heo, L. ColabFold: Making Protein Folding Accessible to All. *Nat Methods* 2022, 19, 679-682.
216. Hekkelman, M. L.; De Vries, I.; Joosten, R. P.; Perrakis, A. AlphaFill: Enriching AlphaFold Models with Ligands and Cofactors. *Nat Methods* 2023, 20, 205–213.

217. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. Humana Press, 2005, 571–607.

218. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J Comput Chem* 2005, 26, 1701–1718.

219. Huang, J.; Mackerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J Comput Chem* 2013, 32, 2135-2145.

220. Konermann, L.; Metwally, H.; McAllister, R. G.; Popa, V. How to Run Molecular Dynamics Simulations on Electrospray Droplets and Gas Phase Proteins: Basic Guidelines and Selected Applications. *Methods* 2018, 144, 104–112.

221. Benesch, J. L. P.; Ruotolo, B. T. Mass Spectrometry: Come of Age for Structural and Dynamical Biology. *Curr Opin Struct Biol* 2011, 21, 641–649.

222. Selva, E.; Beretta, G.; Montanini, N.; Saddler, G. S.; Gastaldo, L.; Ferrari, P.; Lorenzetti, R.; Landini, P.; Ripamonti, F.; Goldstein, B. P.; Berti, M.; Montanaro, L.; Denaro, M.; Montanaro, L. Antibiotic GE2270 A: A Novel Inhibitor Of Bacterial Protein Synthesis I. Isolation And Characterization. *J Antibiot* 1991, 44, 693–701.

223. Vékey, K.; Colombo, L.; Ferrari, P.; Tavecchia, P.; Nebuloni, M.; Vékey, K.; Gallo, G. G.; Enrico, S. Antibiotic GE2270 A: A Novel Inhibitor Of Bacterial Protein Synthesis II. Structural Elucidation. *J Antibiot* 1991, 44, 702–715.

224. Carratore, F. Del; Iorio, M.; Pérez-Bonilla, M.; Schmidt, K.; Pérez-Redondo, R.; Sosio, M.; Macdonald, S. J.; Gyulev, I. S.; Tsigkinopoulou, A.; Thomas, G. H.; Genilloud, O.; Rodríguez-García, A.; Donadio, S.; Breitling, R.; Takano, E. Multi-Omics Study of *Planobispora Rosea*, Producer of the Thiopeptide Antibiotic GE2270A. *mSystems* 2021, 6, 1110-1128.

225. Morris, R. P.; Leeds, J. A.; Naegeli, H. U.; Oberer, L.; Memmert, K.; Weber, E.; LaMarche, M. J.; Parker, C. N.; Burrer, N.; Esterow, S.; Hein, A. E.; Schmitt, E. K.; Krastel, P. Ribosomally Synthesized Thiopeptide Antibiotics Targeting Elongation Factor Tu. *J Am Chem Soc* 2009, 131, 5946–5955.

226. Selva, E.; Ferrari, P.; Kurz, M.; Tavecchia, P.; Colombo, L.; Stella, S.; Restelli, E.; Goldstein, B. P.; Ripamonti, F.; Denaro, M. Components of the GE2270 Complex Produced by *Planobispora Rosea* ATCC53773. *J Antibiot* 48, 1039-1042.

227. Tocchetti, A.; Maffioli, S.; Iorio, M.; Alt, S.; Mazzei, E.; Brunati, C.; Sosio, M.; Donadio, S. Capturing Linear Intermediates and C-Terminal Variants during Maturation of the Thiopeptide GE2270. *Chem Biol* 2013, 20, 1067–1077.

228. Anborgh, P. H.; Parmeggiani, A. Probing the Reactivity of the GTP- and GDP-Bound Conformations of Elongation Factor Tu in Complex with the Antibiotic GE2270 A. *J Biol Chem* 1993, 268, 24622–24628.

229. Möhrle, V. G.; Tieleman, L. N.; Kraal, B. Elongation Factor Tu1 of the Antibiotic GE2270A Producer *Planobispora Rosea* Has an Unexpected Resistance Profile against EF-Tu Targeted Antibiotics. *Biochem Biophys Res Commun* 1997, 230, 320–326.

230. Vijgenboom, E.; Woudt, L. P.; Heinstra, P. W. H.; Rietveld, K.; Van Haarlem, J.; Van Wezel, G. P.; Shochat, S.; Bosch, L. Three Tuf-like Genes in the Kirromycin Producer *Streptomyces Ramocissimus.* *Microbiology* 1994, 140, 983–998.

231. Zuurmond, A. M.; De Graaf, J. M.; Olsthoorn-Tieleman, L. N.; Van Duyl, B. Y.; Mörhle, V. G.; Jurnak, F.; Mesters, J. R.; Hilgenfeld, R.; Kraal, B. GE2270A-Resistant Mutations in Elongation Factor Tu Allow Productive Aminoacyl-tRNA Binding to EF-Tu·GTP·GE2270A Complexes. *J Mol Biol* 2000, 304, 995–1005.

232. Leeds, J. A.; Sachdeva, M.; Mullin, S.; Dzink-Fox, J.; LaMarche, M. J. Mechanism of Action of and Mechanism of Reduced Susceptibility to the Novel Anti-*Clostridium Difficile* Compound LFF571. *Antimicrob Agents Chemother* 2012, 56, 4463–4465.

233. Kertesz, T. M.; Hall, L. H.; Hill, D. W.; Grant, D. F. CE50: Quantifying Collision Induced Dissociation Energy for Small Molecule Characterization and Identification. *J Am Soc Mass Spectrom* 2009, 20, 1759–1767.

234. Watanabe, T.; Izaki, K.; Takahashi, H. New Polyenic Antibiotics Active Against Gram-positive And -negative Bacteria I. Isolation And Purification Of Antibiotics Produced By *Gluconobacter* Sp. W-315. *J Antibiot* 1982, 35, 1141–1147.

235. Watanabe, T.; Izaki, K.; Takahashi, H. New Polyenic Antibiotics Active Against Gram-positive And-negative Bacteria II. Screening Of Antibiotic Producers And Taxonomical Properties Of *Gluconobacter* Sp. W-315. *J Antibiot* 1982, 35, 1148–1155.

236. Watanabe, T.; Sugiyama, T.; Takahashi, M.; Shima, J.; Yamashita, K.; Izaki, K.; Furihata, K.; Seto, H. The Structure of Enacyloxin II, a Novel Linear Polyenic Antibiotic Produced by *Gluconobacter* Sp. W-315. *Agric Biol Chem* 1990, 54, 259–261.

237. Watanabe, T.; Shima, J.; Izaki, K.; Sugiyama, T. New Polyenic Antibiotics Active Against Gram-positive And Gram-negative Bacteria VII. Isolation And Structure Of Enacyloxin Iva, A Possible Biosynthetic Intermediate Of Enacyloxin IIa. *J Antibiot* 1992, 45, 575-576.

238. Furukawa, H.; Kiyota, H.; Yamada, T.; Yaosaka, M.; Takeuchi, R.; Watanabe, T.; Kuwahara, S. Stereochemistry of Enacyloxins. Part 4: Complete Structural and

Configurational Assignment of the Enacyloxin Family, a Series of Antibiotics from *Frateuria* Sp. W-315. *Chem Biodivers* 2007, 4, 1601-1604.

239. Watanabe, T.; Sugiyama, T.; Takahashi, M.; Shima, J.; Yamashita, K.; Izaki, K.; Furihatat, K.; Seto, H. New Polyenic Antibiotics Active Against Gram-positive And Gram-negative Bacteria IV. Structural Elucidation Of Enacyloxin IIa. *J Antibiot* 1992, 45, 470-475.

240. Watanabe, T.; Kiyota, H.; Takeuchi, R.; Enari, K.; Oritani, T. Stereochemistry of Enacyloxin. Part 2: Structure Elucidation of Decarbamoyl Enacyloxin IIa and IVa, New Members of Enacyloxin Antibiotics from *Frateuria* Sp. W-315. *Heterocycl Comm* 2001, 7, 313–316.

241. Mahenthiralingam, E.; Song, L.; Sass, A.; White, J.; Wilmot, C.; Marchbank, A.; Boaisha, O.; Paine, J.; Knight, D.; Challis, G. L. Enacyloxins Are Products of an Unusual Hybrid Modular Polyketide Synthase Encoded by a Cryptic *Burkholderia Ambifaria* Genomic Island. *Chem Biol* 2011, 18, 665-677.

242. Oyama, R.; Hanzawa, H.; Sano, T.; Sugiyama, T.; Izaki, K.; Watanabe, T. An Extracellular Quinoprotein Oxidase That Catalyzes Conversion of Enacyloxin IVa to Enacyloxin IIa. *Biosci Biotechnol Biochem* 1994, 58, 1914–1917.

243. Watanabe, T.; Oyama, R.; Hanzawa, H.; Sugiyama, T.; Izaki, K. Enzymatic Properties of an Extracellular Quinoprotein, Enacyloxin Oxidase. *Biosci Biotechnol Biochem* 1995, 59, 123–125.

244. Coenye, T.; Vandamme, P. Diversity and Significance of *Burkholderia* Species Occupying Diverse Ecological Niches. *Environ Microbiol* 2003, 5, 719–729.

245. Chain, P. S. G.; Denef, V. J.; Konstantinidis, K. T.; Vergez, L. M.; Agulló, L.; Reyes, V. L.; Hauser, L.; Córdova, M.; Gómeza, L.; González, M.; Land, M.; Lao, V.; Larimer, F.; LiPuma, J. J.; Mahenthiralingam, E.; Malfatti, S. A.; Marx, C. J.; Parnell, J. J.; Ramette, A.; Richardson, P.; Seeger, M.; Smith, D.; Spilker, T.; Sul, W. J.; Tsoi, T. V.; Ulrich, L. E.; Zhulin, I. B.; Tiedje, J. M. Burkholderia *Xenovorans* LB400 Harbors a Multi-Replicon, 9.73-Mbp Genome Shaped for Versatility. *Proc Natl Acad Sci* 2006, 103, 15280–15287.

246. Pukatzki, S.; Ma, A. T.; Sturtevant, D.; Krastins, B.; Sarracino, D.; Nelson, W. C.; Heidelberg, J. F.; Mekalanos, J. J. Identification of a Conserved Bacterial Protein Secretion System in *Vibrio Cholerae* Using the Dictyostelium Host Model System. *Proc Natl Acad Sci* 2006, 103, 1528–1533.

247. Watanabe, T.; Suzuki, T.; Izaki, K. New Polyenic Antibiotics Against Gram-positive And Gram-negative Bacteria V. Mode Of Action Of Enacyloxin IIa. *J Antibiot* 1991, 44, 1457-1459.

248. Watanabe, T.; Okubo, N.; Suzuki, T.; Izaki, K. New Polyenic Antibiotics Active Against Gram-positive And Gram-negative Bacteria VI. Non-lactonic Polyene Antibiotic, Enacyloxin IIa, Inhibits Binding Of Aminoacyl-tRNA To A Site Of Ribosomes. *J Antibiot* 1992, 45, 572-574

249. Cetin, R.; Rab, I. M. K.; Anborgh, P. H.; Cool, R. H.; Watanabe, T.; Sugiyama, T.; Izaki, K.; Parmeggiani, A. Enacyloxin IIa, an Inhibitor of Protein Biosynthesis That Acts on Elongation Factor Tu and the Ribosome. *EMBO J* 1996, 15, 2604–2611.

250. Zuurmond, A. M.; Olsthoorn-Tieleman, L. N.; Martien De Graaf, J.; Parmeggiani, A.; Kraal, B. Mutant EF-Tu Species Reveal Novel Features of the Enacyloxin IIa Inhibition Mechanism on the Ribosome. *J Mol Biol* 1999, 294, 627–637.

251. Parmeggiani, A.; Krab, I. M.; Watanabe, T.; Nielsen, R. C.; Dahlberg, C.; Nyborg, J.; Nissen, P. Enacyloxin IIa Pinpoints a Binding Pocket of Elongation Factor Tu for Development of Novel Antibiotics. *J Biol Chem* 2006, 281, 2893–2900.

252. Mesters, J. R.; Zeef, L. A. H.; Hilgenfeld, R.; Martien De Graaf, J.; Kraal, B.; Bosch, L. The Structural and Functional Basis for the Kirromycin Resistance of Mutant EF-Tu Species in *Escherichia Coli*. *EMBO J* 1994, 13, 4877–4885.

253. Créchet, J. B.; Malosse, C.; Hountondji, C. EF-Tu from the Enacyloxin Producing *Frateuria* W-315 Strain: Structure/Activity Relationship and Antibiotic Resistance. *Biochimie* 2016, 127, 59–69.

254. Foreman, D. J.; Dziekonski, E. T.; McLuckey, S. A. Maximizing Selective Cleavages at Aspartic Acid and Proline Residues for the Identification of Intact Proteins. *J Am Soc Mass Spectrom* 2019, 30, 34.

255. Breci, L. A.; Tabb, D. L.; Yates, J. R.; Wysocki, V. H. Cleavage N-Terminal to Proline: Analysis of a Database of Peptide Tandem Mass Spectra. *Anal Chem* 2003, 75, 1963-1971.

256. Gu, C.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. Selective Gas-Phase Cleavage at the Peptide Bond C-Terminal to Aspartic Acid in Fixed-Charge Derivatives of Asp-Containing Peptides. *Biomed Environ Mass Spectrom* 1997, 8, 5804–5813.

257. Benesch, J. L. P. Collisional Activation of Protein Complexes: Picking Up the Pieces. *J Am Soc Mass Spectrom* 2009, 20, 341–348.

258. M Gavriilidou, A. F.; Holding, F. P.; Coyle, J. E.; Zenobi, R. Application of Native ESI-MS to Characterize Interactions between Compounds Derived from Fragment-Based

Discovery Campaigns and Two Pharmaceutically Relevant Proteins. *SLAS Discov* 2018, 23, 951-959.

259. Vatansever, S.; Gümüş, Z. H.; Erman, B. Intrinsic K-Ras Dynamics: A Novel Molecular Dynamics Data Analysis Method Shows Causality between Residue Pair Motions. *Sci Rep* 2016, 6, 1–12.

260. Aguilan, J. T.; Kulej, K.; Sidoli, S. Guide for Protein Fold Change and P-Value Calculation for Non-Experts in Proteomics. *Mol Omics* 2020, 16, 573–582.

261. McCaldon, P.; Argos, P. Oligopeptide Biases in Protein Sequences and Their Use in Predicting Protein Coding Regions in Nucleotide Sequences. *Proteins* 1988, 4, 99–122.

262. Gorres, K. L.; Raines, R. T. Prolyl 4-Hydroxylase. *Crit Rev Biochem Mol Biol* 2010, 45, 106–124.

263. Rappu, P.; Salo, A. M.; Myllyharju, J.; Heino, J. Role of Prolyl Hydroxylation in the Molecular Interactions of Collagens. *Essays Biochem* 2019, 63, 325–335.

264. Rosenbloom, J.; Harsch, M.; Jimenez, S. Hydroxyproline Content Determines the Denaturation Temperature of Chick Tendon Collagen. *Arch Biochem Biophys* 1973, 158, 478–484.

265. Bella, J. Collagen Structure: New Tricks from a Very Old Dog. *Biochem J* 2016, 473, 1001–1025.

266. Chow, W. Y.; Bihan, D.; Forman, C. J.; Slatter, D. A.; Reid, D. G.; Wales, D. J.; Farndale, R. W.; Duer, M. J. Hydroxyproline Ring Pucker Causes Frustration of Helix Parameters in the Collagen Triple Helix. *Sci Rep* 2015, 5.

267. Zou, Y.; Donkervoort, S.; Salo, A. M.; Foley, A. R.; Barnes, A. M.; Hu, Y.; Makareeva, E.; Leach, M. E.; Mohassel, P.; Dastgir, J.; Deardorff, M. A.; Cohn, R. D.; Dinonno, W. O.; Malfait, F.; Lek, M.; Leikin, S.; Marini, J. C.; Myllyharju, J.; Bönnemann, C. G. P4HA1 Mutations Cause a Unique Congenital Disorder of Connective Tissue Involving Tendon, Bone, Muscle and the Eye. *Hum Mol Genet* 2017, 26, 2207–2217.

268. Gandhi, M.; Elfeky, O.; Ertugrul, H.; Chela, H. K.; Daglilar, E. Scurvy: Rediscovering a Forgotten Disease. *Diseases* 2023, 11.

269. Schmelzer, C. E. H.; Nagel, M. B. M.; Dziomba, S.; Merkher, Y.; Sivan, S. S.; Heinz, A. Prolyl Hydroxylation in Elastin Is Not Random. *Biochim Biophys Acta* 2016, 1860, 2169–2177.

270. Guo, J.; Chakraborty, A. A.; Liu, P.; Gan, W.; Zheng, X.; Inuzuka, H.; Wang, B.; Zhang, J.; Zhang, L.; Yuan, M.; Novak, J.; Cheng, J. Q.; Toker, A.; Signoretti, S.; Zhang, Q.; Asara,

J. M.; Kaelin, W. G.; Wei, W. PVHL Suppresses Kinase Activity of Akt in a Proline-Hydroxylation-Dependent Manner. *Science* 2016, 353, 929–932.

271. Lee, S. B.; Ko, A.; Oh, Y. T.; Shi, P.; D'Angelo, F.; Frangaj, B.; Koller, A.; Chen, E. I.; Cardozo, T.; Iavarone, A.; Lasorella, A. Proline Hydroxylation Primes Protein Kinases for Autophosphorylation and Activation. *Mol Cell* 2020, 79, 376-389.

272. Hu, S.; He, W.; Wu, G. Hydroxyproline in Animal Metabolism, Nutrition, and Cell Signalling. *Amino Acids* 2021, 54, 513–528.

273. Jaakkola, P.; Mole, D. R.; Tian, Y. M.; Wilson, M. I.; Gielbert, J.; Gaskell, S. J.; Von Kriegsheim, A.; Hebestreit, H. F.; Mukherji, M.; Schofield, C. J.; Maxwell, P. H.; Pugh, C. W.; Ratcliffe, P. J. Targeting of HIF-α to the von Hippel-Lindau Ubiquitylation Complex by O2-Regulated Prolyl Hydroxylation. *Science* 2001, 292, 468–472.

274. Gallipoli, P.; Huntly, B. J. P. Histone Modifiers Are Oxygen Sensors. *Science* 2019, 363, 1148–1149.

275. Yu, B.; Hunt, J. F. Enzymological and Structural Studies of the Mechanism of Promiscuous Substrate Recognition by the Oxidative DNA Repair Enzyme AlkB. *Proc Natl Acad Sci* 2009, 106, 14315–14320.

276. Walport, L. J.; Hopkinson, R. J.; Schofield, C. J. Mechanisms of Human Histone and Nucleic Acid Demethylases. *Curr Opin Chem Biol* 2012, 16, 525–534.

277. Markolovic, S.; Wilkins, S. E.; Schofield, C. J. Protein Hydroxylation Catalyzed by 2-Oxoglutarate-Dependent Oxygenases. *J Biol Chem* 2015, 290, 20712–20722.

278. Cheng, A. X.; Han, X. J.; Wu, Y. F.; Lou, H. X. The Function and Catalysis of 2-Oxoglutarate-Dependent Oxygenases Involved in Plant Flavonoid Biosynthesis. *Int J Mol Sci* 2014, 15, 1080-1095.

279. Flashman, E.; Schofield, C. J. The Most Versatile of All Reactive Intermediates? *Nat Chem Bio* 2007, 3, 86–87.

280. Strowitzki, M. J.; Cummins, E. P.; Taylor, C. T. Protein Hydroxylation by Hypoxia-Inducible Factor (HIF) Hydroxylases: Unique or Ubiquitous? *Cells* 2019, 8,

281. Iwai, K.; Yamanaka, K.; Kamura, T.; Minato, N.; Conaway, R. C.; Conaway, J. W.; Klausner, R. D.; Pause, A. Identification of the von Hippel–Lindau Tumor-Suppressor Protein as Part of an Active E3 Ubiquitin Ligase Complex. *Proc Natl Acad Sci* 1999, 96, 12436–12441.

282. Mole, D. R.; Maxwell, P. H.; Pugh, C. W.; Ratcliffe, P. J. Regulation of HIF by the von Hippel-Lindau Tumour Suppressor: Implications for Cellular Oxygen Sensing. *IUBMB Life* 2001, 52, 43–47.

283. Tian, Y. M.; Yeoh, K. K.; Lee, M. K.; Eriksson, T.; Kessler, B. M.; Kramer, H. B.; Edelmann, M. J.; Willam, C.; Pugh, C. W.; Schofield, C. J.; Ratcliffe, P. J. Differential Sensitivity of Hypoxia Inducible Factor Hydroxylation Sites to Hypoxia and Hydroxylase Inhibitors. *J Biol Chem* 2011, 286, 13041-13051.

284. Chan, D. A.; Sutphin, P. D.; Yen, S.-E.; Giaccia, A. J. Coordinate Regulation of the Oxygen-Dependent Degradation Domains of Hypoxia-Inducible Factor 1α. *Mol Cell Biol* 2005, 25, 6415–6426.

285. Mao, H.; Yang, A.; Zhao, Y.; Lei, L.; Li, H. Succinate Supplement Elicited "Pseudohypoxia" Condition to Promote Proliferation, Migration, and Osteogenesis of Periodontal Ligament Cells. *Stem Cells Int* 2020, 1.

286. Scotti, J. S.; Leung, I. K. H.; Ge, W.; Bentley, M. A.; Paps, J.; Kramer, H. B.; Lee, J.; Aik, W.; Choi, H.; Paulsen, S. M.; Bowman, L. A. H.; Loik, N. D.; Horita, S.; Ho, C. H.; Kershaw, N. J.; Tang, C. M.; Claridge, T. D. W.; Preston, G. M.; McDonough, M. A.; Schofield, C. J. Human Oxygen Sensing May Have Origins in Prokaryotic Elongation Factor Tu Prolyl-Hydroxylation. *Proc Natl Acad Sci* 2014, 111, 13331–13336.

287. Lau, G. W.; Hassett, D. J.; Ran, H.; Kong, F. The Role of Pyocyanin in Pseudomonas Aeruginosa Infection. *Trends Mol Med* 2004, 10, 599–606.

288. Graifer, D.; Malygin, A.; Karpova, G. Hydroxylation of Protein Constituents of the Human Translation System: Structural Aspects and Functional Assignments. *Future Med Chem* 2019, 11, 357–369.

289. Duret, L.; Gasteiger, E.; Perriere, G. LALNVIEW: A Graphical Viewer for Pairwise Sequence Alignments. *Comput Appl Biosci* 1996, 12, 507–510.

290. Ghannoum, M. A.; Rice, L. B. Antifungal Agents: Mode of Action, Mechanisms of Resistance, and Correlation of These Mechanisms with Bacterial Resistance. *Clin Microbiol Rev* 1999, 12, 501-517.

291. Groeschke, J.; Solassol, I.; Bressolle, F.; Pinguet, F. Stability of Amphotericin B and Nystatin in Antifungal Mouthrinses Containing Sodium Hydrogen Carbonate. *J Pharm Biomed Anal* 2006, 42, 362–366.

292. Denning, D. W.; Hope, W. W. Therapy for Fungal Diseases: Opportunities and Priorities. *Trends Microbiol* 2010, 18, 195–204.

293. Wasan, K. M.; Wasan, E. K.; Gershkovich, P.; Zhu, X.; Tidwell, R. R.; Werbovetz, K. A.; Clement, J. G.; Thornton, S. J. Highly Effective Oral Amphotericin B Formulation against Murine Visceral Leishmaniasis. *J Infect Dis* 2009, 200, 357–360.

294. Hamill, R. J. Amphotericin B Formulations: A Comparative Review of Efficacy and Toxicity. *Drugs* 2013, 73, 919–934.

295. Fabbretti, A.; He, C. G.; Gaspari, E.; Maffioli, S.; Brandi, L.; Spurio, R.; Sosio, M.; Jabes, D.; Donadio, S. A Derivative of the Thiopeptide GE2270A Highly Selective against *Propionibacterium Acnes.* *Antimicrob Agents Chemother* 2015, 59, 4560-4568.

296. Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V.; Baldwin, A. J.; Benesch, J. L. P. Collision Cross Sections for Structural Proteomics. *Structure* 2015, 23, 791–799.

297. Jurneczko, E.; Barran, P. E. 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* 2010, 136, 20–28.

298. Bleiholder, C.; Wyttenbach, T.; Bowers, M. T. A Novel Projection Approximation Algorithm for the Fast and Accurate Computation of Molecular Collision Cross Sections (I). *Method. Int J Mass Spectrom* 2011, 308, 1–10.

299. Hewitt, D.; Marklund, E.; Scott, D. J.; Robinson, C. V.; Borysik, A. J. A Hydrodynamic Comparison of Solution and Gas Phase Proteins and Their Complexes. *J Phys Chem B* 2014, 118, 8489–8495.

300. Ewing, S. A.; Donor, M. T.; Wilson, J. W.; Prell, J. S. Collidoscope: An Improved Tool for Computing Collisional Cross-Sections with the Trajectory Method. *J Am Soc Mass Spectrom* 2017, 28, 587–596.

301. Mack, E. Average Cross-Sectional Areas of Molecules by Gaseous Diffusion Methods. *J Am Chem Soc* 1925, 47, 2468–2482.

302. Valentine, S. J.; Anderson, J. G.; Ellington, A. D.; Clemmer, D. E. Disulfide-Intact and -Reduced Lysozyme in the Gas Phase: Conformations and Pathways of Folding and Unfolding. *J Phys Chem B* 1997, 101, 3891–3900.

303. Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. Protein Structure *in Vacuo*: Gas-Phase Conformations of BPTI and Cytochrome c. *J Am Chem Soc* 1997, 119, 2240–2248.

304. González, M. A. Force Fields and Molecular Dynamics Simulations. *Collection SFN* 2011, 12, 169–200.

305. Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids: Second Edition. *Oxford University Press* 2017.

306. Levitt, M.; Warshel, A. Computer Simulation of Protein Folding. *Nature* 1975, 253, 694-698.

307. Durrant, J. D.; Kochanek, S. E.; Casalino, L.; Ieong, P. U.; Dommer, A. C.; Amaro, R. E. Mesoscale All-Atom Influenza Virus Simulations Suggest New Substrate Binding Mechanism. *ACS Cent Sci* 2020, 6, 189–196.

308. Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; Mclellan, J. S.; Fadda, E.; Amaro, R. E. Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Cent Sci* 2020, 6, 1722–1734.

309. Arteca, G. A.; Reimann, C. T.; Tapia, O. Proteins *in Vacuo*: Denaturing and Folding Mechanisms Studied with Computer-Simulated Molecular Dynamics. *Mass Spectrom Rev* 2001, 20, 402–422.

310. Shelimov, K. B.; Jarrold, M. F. "Denaturation" and Refolding of Cytochrome c *in Vacuo*. *J Am Chem Soc* 1996, 118, 10313–10314.

311. Gross, D. S.; Schnier, P. D.; Rodriguez-Cruz, S. E.; Fagerquist, C. K.; Williams, E. R. Conformations and Folding of Lysozyme Ions *in Vacuo*. *Proc Natl Acad Sci* 1996, 93, 3143–3148.

312. Rolland, A. D.; Prell, J. S. Computational Insights into Compaction of Gas-Phase Protein and Protein Complex Ions in Native Ion Mobility-Mass Spectrometry. *Trends Anal Chem* 2019, 116, 282–291.

313. Sever, A. I. M.; Konermann, L. Gas Phase Protein Folding Triggered by Proton Stripping Generates Inside-Out Structures: A Molecular Dynamics Simulation Study. *J Phys Chem B* 2020, 124, 3667–3677.

314. Migas, L. G.; France, A. P.; Bellina, B.; Barran, P. E. ORIGAMI: A Software Suite for Activated Ion Mobility Mass Spectrometry (AIM-MS) Applied to Multimeric Protein Assemblies. *Int J Mass Spectrom* 2018, 427, 20–28.

315. Collins, J.; Hohn, B. Cosmids: A Type of Plasmid Gene-Cloning Vector That Is Packageable in Vitro in Bacteriophage Lambda Heads. *Proc Natl Acad* 1978, 75, 4242.

316. Voorhees, R.; Ramakrishnan, V. Structural Basis of the Translational Elongation Cycle. *Annu Rev Biochem* 2013, 82, 203-236.

Appendices

I. Protein Sequences

i. *E. coli* EF-Tu$^{\text{Wild-Type}}$

(GIDPFT)MSKEKFERTKPHVNVGTIGHVDHGKTTLTAAITTVLAKTYGGAARAFDQIDNAPEEKARGITINTS

HVEYDTPTRHYAHVDCPGHADYVKNMITGAAQMDGAILVVAATDGPMPQTREHILLGRQVGVPYIIVFLNKCDM

VDDEELLELVEMEVRELLSQYDFPGDDTPIVRGSALKALEGDAEWEAKILELAGFLDSYIPEPERAIDKPFLLP

IEDVFSISGRGTVVTGRVERGIIKVGEEVEIVGIKETQKSTCTGVEMFRKLLDEGRAGENVGVLLRGIKREEIE

RGQVLAKPGTIKPHTKFESEVYILSKDEGGRHTPFFKGYRPQFYFRTTDVTGTIELPEGVEMVMPGDNIKMVVT

LIHPIAMDDGLRFAIREGGRTVGAGVVAKVLS

ii. *A. baumannii* EF-Tu$^{\text{Wild-Type}}$

(GIDPFT)MAKAKFERNKPHVNVGTIGHVDHGKTTLTAAIATICAKTYGGEAKDYSQIDSAPEEKARGITINTS

HVEYDSPTRHYAHVDCPGHADYVKNMITGAAQMDGAILVCAATDGPMPQTREHILLSRQVGVPYIIVFLNKCDL

VDDEELLELVEMEVRELLSTYDFPGDDTPVIRGSALAALNGEAGPYGEESVLALVAALDSYIPEPERAIDKAFL

MPIEDVFSISGRGTVVTGRVEAGIIKVGEEVEIVGIKDTVKTTVTGVEMFRKLLDEGRAGENCGILLRGTKREE

VQRGQVLAKPGTIKPHTKFDAEVYVLSKEEGGRHTPFLNGYRPQFYFRTTDVTGAIQLKEGVEMVMPGDNVEMS

VELIHPIAMDPGLRFAIREGGRTVGAGVVAKVTA

iii. *A. baumannii* EF-Tu^{Q131K}

(GIDPFT)MAKAKFERNKPHVNVGTIGHVDHGKTTLTAAIATICAKTYGGEAKDYSQIDSAPEEKARGITINTS
HVEYDSPTRHYAHVDCPGHADYVKNMITGAAQMDGAILVCAATDGPMPQTREHILLSRKVGVPYIIVFLNKCDL
VDDEELLELVEMEVRELLSTYDFPGDDTPVIRGSALAALNGEAGPYGEESVLALVAALDSYIPEPERAIDKAFL
MPIEDVFSISGRGTVVTGRVEAGIIKVGEEVEIVGIKDTVKTTVTGVEMFRKLLDEGRAGENCGILLRGTKREE
VQRGQVLAKPGTIKPHTKFDAEVYVLSKEEGGRHTPFLNGYRPQFYFRTTDVTGAIQLKEGVEMVMPGDNVEMS
VELIHPIAMDPGLRFAIREGGRTVGAGVVAKVTA

iv. *A. baumannii* EF-Tu^{G325D}

(GIDPFT)MAKAKFERNKPHVNVGTIGHVDHGKTTLTAAIATICAKTYGGEAKDYSQIDSAPEEKARGITINTS
HVEYDSPTRHYAHVDCPGHADYVKNMITGAAQMDGAILVCAATDGPMPQTREHILLSRQVGVPYIIVFLNKCDL
VDDEELLELVEMEVRELLSTYDFPGDDTPVIRGSALAALNGEAGPYGEESVLALVAALDSYIPEPERAIDKAFL
MPIEDVFSISGRGTVVTGRVEAGIIKVGEEVEIVGIKDTVKTTVTGVEMFRKLLDEGRAGENCGILLRGTKREE
VQRGQVLAKPGTIKPHTKFDAEVYVLSKEEDGRHTPFLNGYRPQFYFRTTDVTGAIQLKEGVEMVMPGDNVEMS
VELIHPIAMDPGLRFAIREGGRTVGAGVVAKVTA

v. *A. baumannii* EF-Tu^{A384T}

(GIDPFT)MAKAKFERNKPHVNVGTIGHVDHGKTTLTAAIATICAKTYGGEAKDYSQIDSAPEEKARGITINTS

HVEYDSPTRHYAHVDCPGHADYVKNMITGAAQMDGAILVCAATDGPMPQTREHILLSRQVGVPYIIVFLNKCDL

VDDEELLELVEMEVRELLSTYDFPGDDTPVIRGSALAALNGEAGPYGEESVLALVAALDSYIPEPERAIDKAFL

MPIEDVFSISGRGTVVTGRVEAGIIKVGEEVEIVGIKDTVKTTVTGVEMFRKLLDEGRAGENCGILLRGTKREE

VQRGQVLAKPGTIKPHTKFDAEVYVLSKEEGGRHTPFLNGYRPQFYFRTTDVTGAIQLKEGVEMVMPGDNVEMS

VELIHPIAMDPGLRFTIREGGRTVGAGVVAKVTA

vi. *P. putida* EF-Tu$^{Wild\text{-}Type}$

(MGSSHHHHHHSSGLVPRGSH)MASMTGGQQMGRGSEFMAKEKFDRSLPHVNVGTIGHVDHGKTTLTAALTRVC

SEVFGSAIVEFDKIDSAPEEKARGITINTAHVEYNSTIRHYAHVDCPGHADYVKNMITGAAQMDGAILVCSAAD

GPMPQTREHILLSRQVGVPYIVVFLNKADLVDDAELLELVEMEVRDLLSTYDFPGDDTPIIIGSARMALEGKDD

NEMGTTAVKKLVETLDSYIPEPVRAIDQPFLMPIEDVFSISGRGTVVTGRIERGIVRVQDPLEIVGLRDTTTTT

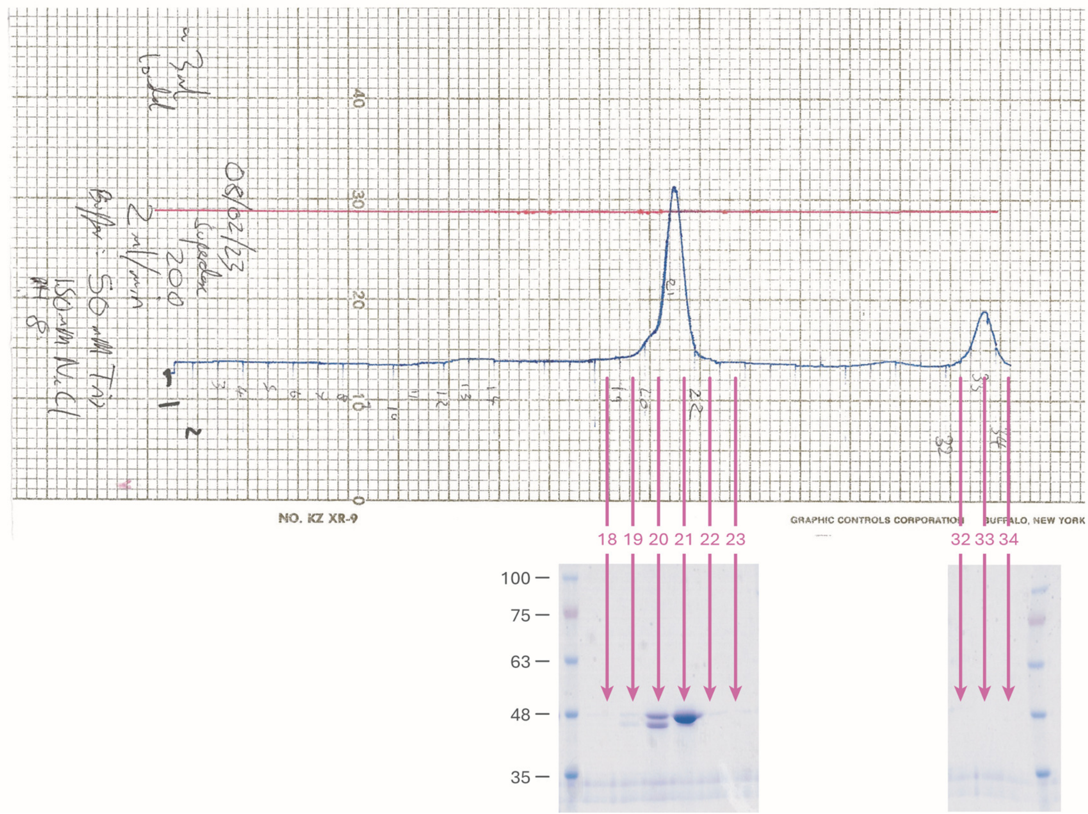
CTGVEMFRKLLDEGRAGENCGVLLRGTKRDDVERGQVLVKPGSVKPHTKFTAEVYVLSKEEGGRHTPFFKGYRP

QFYFRTTDVTGNCELPEGVEMVMPGDNIQMTVTLIKTIAMEDGLRFAIREGGRTVGAGVVAKIIE

vii. *P. putida* PPHD$^{\text{Wild-Type}}$

(MGSSHHHHHHSSGLVPRGSH)MASHISPEHPMLAAVVDDLATHGWSQQAHFLPADLVRALAAECRRRDAEGELNPAGVGRGATQEVRETIRGDQIQWIDPGQAEACDQYLAAMDQLRLAINQGLFLGLEDFECHFALYPPGAFYRRHLDRFRDDDRRM*VS*AVLYLNEGWQPHDGGQLRMFLADGVEHDVEPVAGCLVVFLSGEVPHEVLPAGRERLSLTGWFRRRGNDPF

II. Analytical SEC Chromatogram



Appendix II | Physical SEC chromatogram (top) for wild-type *A. baumannii* EF-Tu. Labelled fractions were subject to SDS-PAGE (bottom), with fraction 21 containing the highest concentration of EF-Tu.

III. MS^2 Fragment Ions of Select Peptides

i. EF-Tu^{Q131K} Val^{132} – Lys^{143}

Appendix III(i) | Observed and calculated *b* and *y* fragment ions for Val^{132} – Lys^{143} in EF-Tu^{Q131K} along with their associated *m/z* errors.

Residue	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
V	b$_1$	100.08	–	–	y$_{12}$	–	–	–
G	b$_2$	157.10	–	–	y$_{11}$	1262.75	1262.50	-0.25
V	b$_3$	256.17	256.01	-0.16	y$_{10}$	1205.73	1205.66	-0.07
P	b$_4$	353.22	353.21	-0.01	y$_9$	1106.66	1106.42	-0.24
Y	b$_5$	516.28	–	–	y$_8$	1009.61	1009.49	-0.12
I	b$_6$	629.37	629.23	-0.14	y$_7$	846.54	846.41	-0.14
I	b$_7$	742.45	742.22	-0.23	y$_6$	733.46	733.33	-0.13
V	b$_8$	841.52	841.24	-0.27	y$_5$	620.38	620.31	-0.07
F	b$_9$	988.59	988.39	-0.20	y$_4$	521.31	521.27	-0.04
L	b$_{10}$	1101.67	1101.39	-0.28	y$_3$	374.24	374.21	-0.03
N	b$_{11}$	1215.71	1215.48	-0.23	y$_2$	261.16	261.05	-0.10
K	b$_{12}$	–	–	–	y$_1$	147.11	–	–
			Average	-0.19			Average	-0.12

ii. EF-Tu$^{\text{Wild-Type}}$ Phe^{313} – Arg^{327}

Appendix III(ii) | Observed and calculated *b* and *y* fragment ions for Phe^{313} – Arg^{327} in EF-Tu$^{\text{Wild-Type}}$ along with their associated *m/z* errors.

Residue	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
F	b_1	148.08	–	–	y_{15}	–	–	–
D	b_2	263.10	263.01	-0.09	y_{14}	1551.76	–	–
A	b_3	334.14	334.00	-0.14	y_{13}	1436.74	–	–
E	b_4	463.18	463.07	-0.11	y_{12}	1365.70	1365.62	-0.08
V	b_5	562.25	562.14	-0.11	y_{11}	1236.66	1236.47	-0.19
Y	b_6	725.31	–	–	y_{10}	1137.59	1137.41	-0.18
V	b_7	824.38	824.23	-0.15	y_9	974.53	974.36	-0.17
L	b_8	937.47	937.32	-0.15	y_8	875.46	875.30	-0.16
S	b_9	1024.50	1024.34	-0.16	y_7	762.37	762.26	-0.11
K	b_{10}	1152.59	1152.37	-0.22	y_6	675.34	675.23	-0.11
E	b_{11}	1281.64	1281.49	-0.15	y_5	547.25	547.12	-0.12
E	b_{12}	1410.68	1410.48	–	y_4	418.20	418.19	–
G	b_{13}	1467.70	–	–	y_3	289.16	289.10	-0.06
G	b_{14}	1524.72	–	–	y_2	232.14	–	–
R	b_{15}	–	–	–	y_1	175.12	–	–
			Average	-0.15			Average	-0.13

iii. EF-Tu^{G325D} Phe^{313} – Arg^{327}

Appendix III(iii) | Observed and calculated *b* and *y* fragment ions for Phe^{313} – Arg^{327} in EF-Tu^{G325D} along with their associated *m/z* errors.

Residue	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
F	b_1	148.08	–	–	y_{15}	–	–	–
D	b_2	263.10	263.11	0.01	y_{14}	1609.77	–	–
A	b_3	334.14	334.08	-0.06	y_{13}	1494.74	–	–
E	b_4	463.18	463.13	-0.05	y_{12}	1423.71	1423.58	-0.13
V	b_5	562.25	562.16	-0.10	y_{11}	1294.66	1294.50	-0.17
Y	b_6	725.31	725.15	-0.16	y_{10}	1195.60	1195.48	-0.12
V	b_7	824.38	824.24	-0.14	y_9	1032.53	1032.40	-0.14
L	b_8	937.47	937.27	-0.20	y_8	933.46	933.36	-0.10
S	b_9	1024.50	1024.23	-0.27	y_7	820.38	820.31	-0.07
K	b_{10}	1152.59	1152.30	-0.29	y_6	733.35	733.30	-0.04
E	b_{11}	1281.64	1281.38	-0.26	y_5	605.25	605.24	-0.01
E	b_{12}	1410.68	1410.37	-0.31	y_4	476.21	476.20	–
D	b_{13}	1525.71	1525.47	-0.24	y_3	347.17	347.12	-0.05
G	b_{14}	1582.73	–	–	y_2	232.14	–	–
R	b_{15}	–	–	–	y_1	175.12	–	–
			Average	-0.20			Average	-0.09

iv. EF-Tu^{A384T} Thr^{384} – Ala^{402}

Appendix III(iv) | Observed and calculated *b* and *y* fragment ions for Thr^{384} – Ala^{402} in EF-Tu^{A384T} along with their associated *m/z* errors.

Residue	Fragment Ion (2+)	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion (2+)	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
T	b_1	51.53	–	–	y_{19}	–	–	–
I	b_2	108.07	–	–	y_{18}	871.01	–	–
R	b_3	186.12	–	–	y_{17}	814.47	–	–
E	b_4	250.65	–	–	y_{16}	736.42	–	–
G	b_5	279.16	–	–	y_{15}	671.89	–	–
G	b_6	307.67	–	–	y_{14}	643.38	–	–
R	b_7	385.72	386.07	0.35	y_{13}	614.87	–	–
T	b_8	436.24	–	–	y_{12}	536.82	–	–
V	b_9	485.78	–	–	y_{11}	486.30	–	–
G	b_{10}	514.29	–	–	y_{10}	436.76	–	–
A	b_{11}	549.80	–	–	y_9	408.25	–	–
G	b_{12}	578.32	–	–	y_8	372.73	372.84	0.10
V	b_{13}	627.85	627.64	-0.21	y_7	344.22	344.26	0.03
V	b_{14}	677.38	677.43	0.04	y_6	294.69	295.15	0.46
A	b_{15}	712.90	712.83	-0.08	y_5	245.16	245.14	-0.02
K	b_{16}	776.95	776.92	-0.03	y_4	209.64	–	–
V	b_{17}	826.48	–	–	y_3	145.59	–	–
T	b_{18}	877.01	–	–	y_2	96.06	–	–
A	b_{19}	–			y_1	45.5311		
			Average	0.02			Average	0.14

v. EF-Tu$^{\text{Wild-Type}}$ Thr^{32} – Lys^{44} Isomer I

Appendix III(v) | Observed and calculated *b* and *y* fragment ions for Thr^{32} – Ala^{44} isomer I in EF-Tu$^{\text{Wild-Type}}$ along with their associated *m/z* errors.

Residue	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
T	b_1	102.06	-	-	y_{19}	-	-	-
T	b_2	203.10	-	-	y_{18}	1435.71	-	-
L	b_3	316.19	-	-	y_{17}	1334.66	-	-
T	b_4	417.23	-	-	y_{16}	1221.58	1221.54	-0.04
A	b_5	488.27	-	-	y_{15}	1120.53	1120.51	-0.02
A	b_6	559.31	559.17	-0.14	y_{14}	1049.49	1049.32	-0.17
I	b_7	874.42	874.25	-0.17	y_{13}	978.46	978.35	-0.11
A	b_8	945.45	945.31	-0.15	y_{12}	663.35	663.35	0.00
T	b_9	1046.50	1046.36	-0.14	y_{11}	592.31	592.29	-0.02
I	b_{10}	1159.59	1159.29	-0.30	y_{10}	491.26	491.22	-0.04
C	b_{11}	1319.62	-	-	y_9	378.18	378.18	-0.01
A	b_{12}	1390.65	1390.26	-0.40	y_2	218.15	-	-
K	b_{13}	-	-	-	y_1	147.11	-	-
			Average	-0.23			Average	-0.04

vi. EF-Tu$^{\text{Wild-Type}}$ Thr^{32} – Lys^{44} Isomer II

Appendix III(vi) | Observed and calculated *b* and *y* fragment ions for Thr^{32} – Ala^{44} isomer II in EF-Tu$^{\text{Wild-Type}}$ along with their associated *m/z* errors.

Residue	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
T	b_1	102.06	-	-	y_{19}	-	-	-
T	b_2	203.10	-	-	y_{18}	1435.71	-	-
L	b_3	316.19	-	-	y_{17}	1334.66	1334.51	-0.15
T	b_4	417.23	-	-	y_{16}	1221.58	1221.44	-0.14
A	b_5	488.27	488.10	-0.18	y_{15}	1120.53	1120.31	-0.22
A	b_6	559.31	559.23	-0.08	y_{14}	1049.49	1049.24	-0.26
I	b_7	874.42	874.29	-0.13	y_{13}	978.46	978.32	-0.14
A	b_8	945.45	945.22	-0.24	y_{12}	663.35	663.31	-0.04
T	b_9	1046.50	1046.28	-0.22	y_{11}	592.31	592.19	-0.12
I	b_{10}	1159.59	1159.22	-0.36	y_{10}	491.26	491.23	-0.04
C	b_{11}	1319.62	1319.37	-0.25	y_9	378.18	378.08	-0.10
A	b_{12}	1390.65	1390.33	-0.32	y_2	218.15	-	-
K	b_{13}	-	-	-	y_1	147.11	-	-
			Average	-0.25			Average	-0.09

vii. EF-Tu$^{\text{Wild-Type}}$ Thr^{32} – Lys^{44} Isomer III

Appendix III(vii) | Observed and calculated *b* and *y* fragment ions for Thr^{32} – Ala^{44} isomer III in EF-Tu$^{\text{Wild-Type}}$ along with their associated *m/z* errors.

Residue	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
T	b_1	102.06	–	–	y_{19}	–	–	–
T	b_2	203.10	–	–	y_{18}	1435.71	–	–
L	b_3	316.19	–	–	y_{17}	1334.66	1334.35	-0.31
T	b_4	417.23	–	–	y_{16}	1221.58	–	–
A	b_5	488.27	–	–	y_{15}	1120.53	918.47	-202.06
A	b_6	559.31	–	–	y_{14}	1049.49	847.30	-202.20
I	b_7	874.42	874.19	-0.23	y_{13}	978.46	–	–
A	b_8	945.45	945.20	-0.25	y_{12}	663.35	663.29	-0.06
T	b_9	1046.50	1046.33	-0.17	y_{11}	592.31	592.28	-0.03
I	b_{10}	1159.59	1159.38	-0.21	y_{10}	491.26	491.17	-0.10
C	b_{11}	1319.62	1319.49	-0.13	y_9	378.18	378.19	0.01
A	b_{12}	1390.65	1390.51	-0.14	y_2	218.15	–	–
K	b_{13}	–	–	–	y_1	147.11	–	–
			Average	-0.19			Average	-0.04

viii. EF-Tu$^{\text{Wild-Type}}$ Thr^{32} – Lys^{44} Isomer IV

Appendix III(viii) | Observed and calculated *b* and *y* fragment ions for Thr^{32} – Ala^{44} isomer IV in EF-Tu$^{\text{Wild-Type}}$ along with their associated *m/z* errors.

Residue	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
T	b_1	102.06	-	-	y_{19}	-	-	-
T	b_2	203.10	-	-	y_{18}	1435.71	-	-
L	b_3	316.19	-	-	y_{17}	1334.66	1334.45	-0.21
T	b_4	417.23	-	-	y_{16}	1221.58	1221.46	-0.12
A	b_5	488.27	488.21	-0.07	y_{15}	1120.53	1120.43	-0.10
A	b_6	559.31	559.21	-0.10	y_{14}	1049.49	1049.34	-0.15
I	b_7	672.39	672.26	-0.13	y_{13}	978.46	978.33	-0.13
A	b_8	743.43	743.19	-0.24	y_{12}	865.37	865.21	-0.16
T	b_9	844.48	844.29	-0.19	y_{11}	794.34	794.20	-0.13
I	b_{10}	1159.59	1159.36	-0.22	y_{10}	693.29	693.17	-0.12
C	b_{11}	1319.62	1319.41	-0.21	y_9	378.18	378.16	-0.02
A	b_{12}	1390.65	1390.37	-0.28	y_2	218.15	-	-
K	b_{13}	-	-	-	y_1	147.11	-	-
			Average	-0.21			Average	-0.11

ix. EF-Tu$^{\text{Wild-Type}}$ Thr^{32} – Lys^{44} Isomer V

Appendix III(ix) | Observed and calculated *b* and *y* fragment ions for Thr^{32} – Ala^{44} isomer V in EF-Tu$^{\text{Wild-Type}}$ along with their associated *m/z* errors.

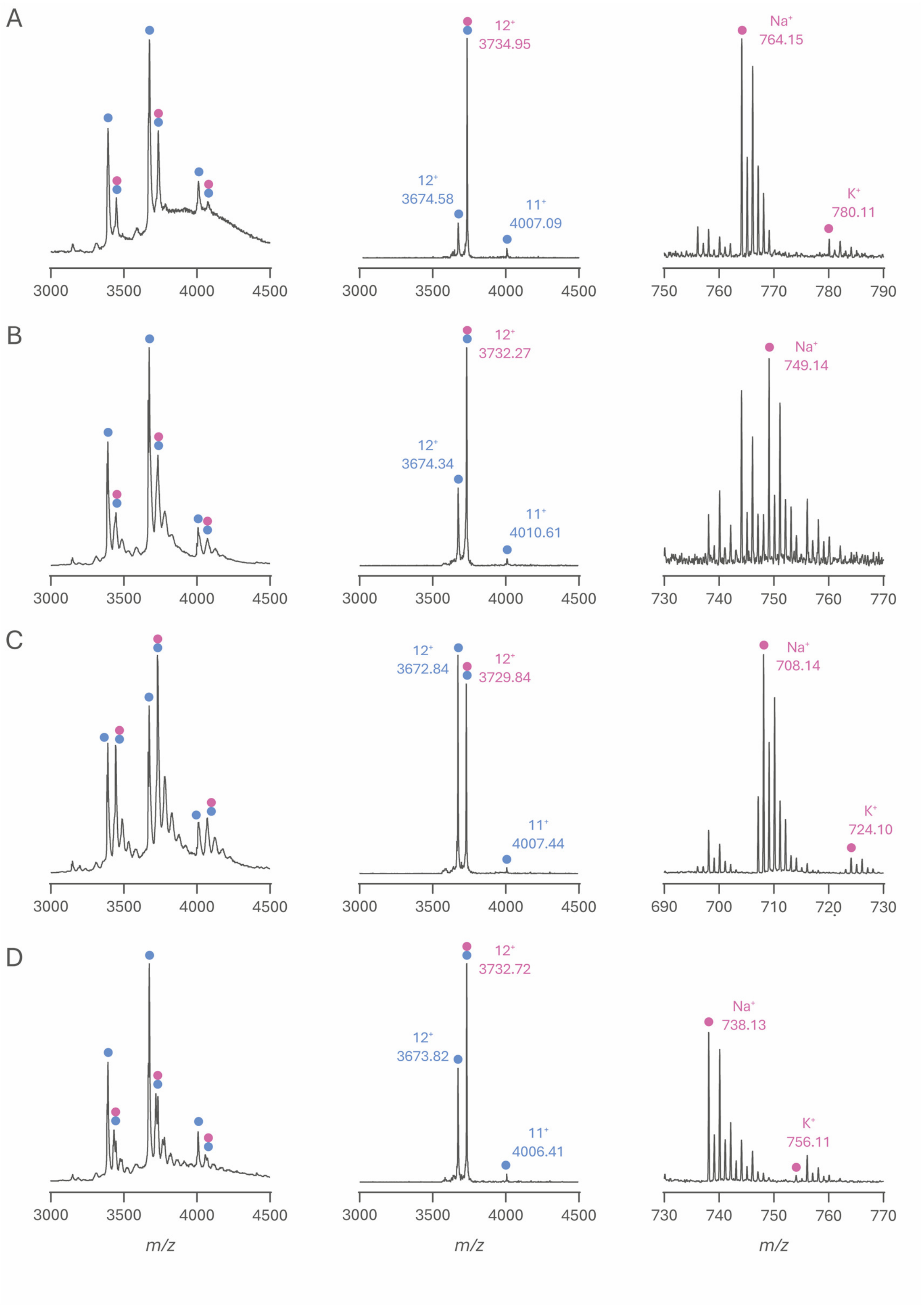
Residue	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
T	b_1	102.06	-	-	y_{19}	-	-	-
T	b_2	203.10	-	-	y_{18}	1435.71	-	-
L	b_3	316.19	-	-	y_{17}	1334.66	1334.48	-0.18
T	b_4	417.23	-	-	y_{16}	1221.58	1221.46	-0.12
A	b_5	488.27	488.14	-0.13	y_{15}	1120.53	1120.50	-0.03
A	b_6	559.31	559.13	-0.18	y_{14}	1049.49	1049.28	-0.22
I	b_7	672.39	672.19	-0.20	y_{13}	978.46	978.31	-0.15
A	b_8	743.43	743.17	-0.25	y_{12}	865.37	865.22	-0.15
T	b_9	844.48	844.33	-0.15	y_{11}	794.34	794.24	-0.10
I	b_{10}	1159.59	1159.38	-0.20	y_{10}	693.29	693.20	-0.09
C	b_{11}	1319.62	-	-	y_9	378.18	378.16	-0.02
A	b_{12}	1390.65	1390.47	-0.18	y_2	218.15	-	-
K	b_{13}	-	-	-	y_1	147.11	-	-
			Average	-0.20			Average	-0.10

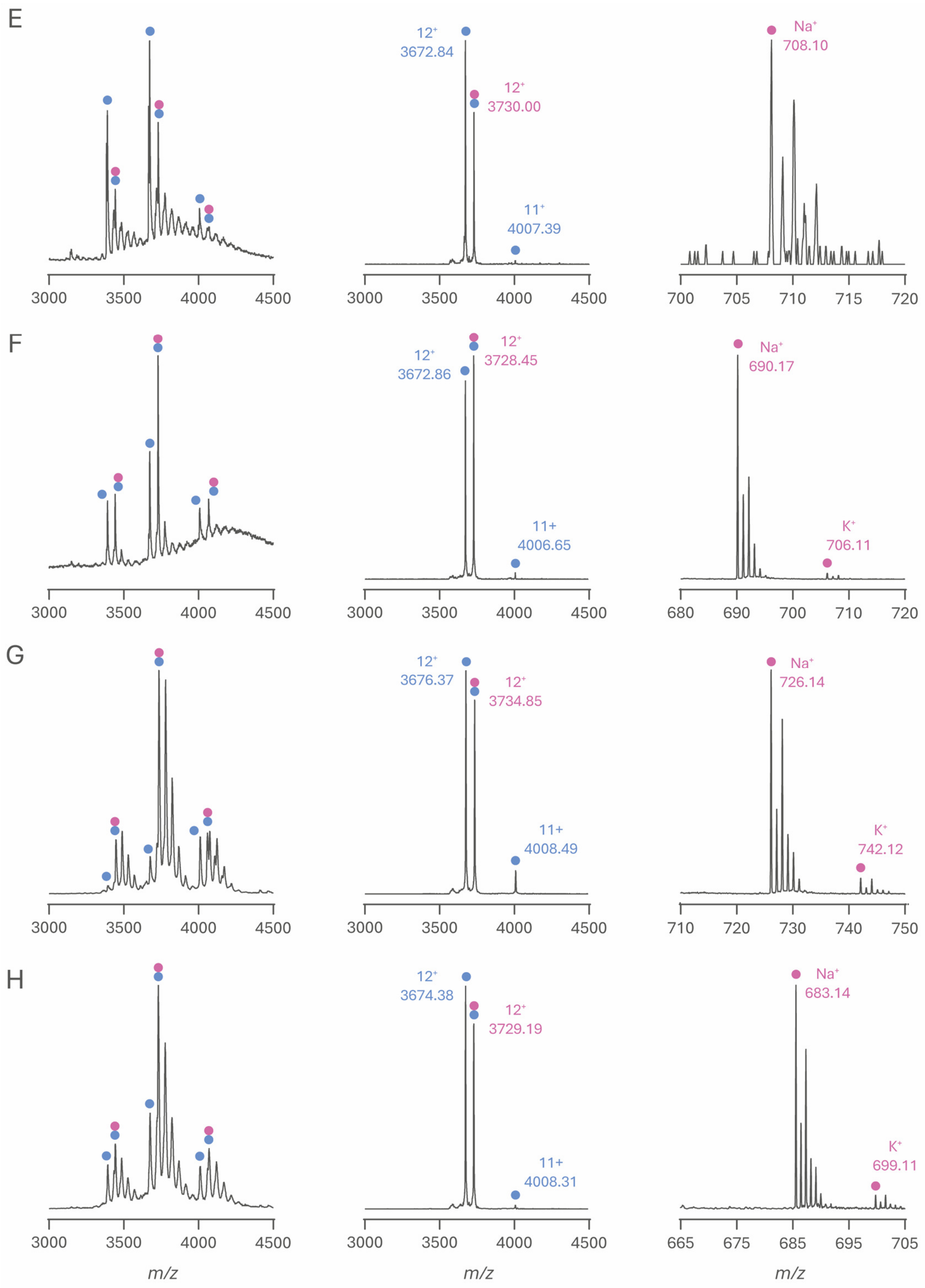
x. EF-Tu$^{\text{Wild-Type}}$ Thr^{32} – Lys^{44} Isomer VI

Appendix III(x) | Observed and calculated *b* and *y* fragment ions for Thr^{32} – Ala^{44} isomer VI in EF-Tu$^{\text{Wild-Type}}$ along with their associated *m/z* errors.

Residue	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)	Fragment Ion	Calculated (*m/z*)	Observed (*m/z*)	Error (*m/z*)
T	b_1	102.06	-	-	y_{19}	-	-	-
T	b_2	203.10	-	-	y_{18}	1435.71	-	-
L	b_3	518.21	-	-	y_{17}	1334.66	1334.56	-0.10
T	b_4	619.26	-	-	y_{16}	1019.56	-	-
A	b_5	690.30	-	-	y_{15}	918.51	-	-
A	b_6	761.33	-	-	y_{14}	847.47	847.31	-0.16
I	b_7	874.42	874.20	-0.22	y_{13}	776.43	-	-
A	b_8	945.45	945.23	-0.22	y_{12}	663.35	663.25	-0.09
T	b_9	1046.50	1046.35	-0.15	y_{11}	592.31	592.21	-0.10
I	b_{10}	1159.59	1159.29	-0.30	y_{10}	491.26	491.16	-0.11
C	b_{11}	1319.62	1319.54	-0.08	y_9	378.18	378.07	-0.11
A	b_{12}	1390.65	1390.32	-0.33	y_2	218.15	-	-
K	b_{13}	-	-	-	y_1	147.11	-	-
			Average	-0.22			Average	-0.10

IV. Native Mass Spectra of Enacyloxin IIa Analogues Bound to EF-Tu





Appendix III | Native mass spectra of each enacyloxin IIa analogue that bound to EF-Tu. Left to right; whole native mass spectrum; quadrupole isolated and collisionally activated $[M + 12H]^{12+}$ EF-Tu•GDPNP•Analogue peak; ejected enacyloxin IIa analogue A – D correspond to enacyloxin analogue IIa1 – IIa4 and E – H to IIa6 – IIa9. Blue circles indicate EF-Tu•GDPNP; pink circles, enacyloxin IIa analogues; blue and pink circles, EF-Tu•GDPNP•Analogue.

V. PIPS Reflective Statement

Note to examiners:

This statement is included as an appendix to the thesis in order that the thesis accurately captures the PhD training experienced by the candidate as a BBSRC Doctoral Training Partnership student.

The Professional Internship for PhD Students is a compulsory 3-month placement which must be undertaken by DTP students. It is usually centred on a specific project and must not be related to the PhD project. This reflective statement is designed to capture the skills development which has taken place during the student's placement and the impact on their career plans it has had.

I had the pleasure to work at Phospho Biomedical Animation (for ease, referred to as just Phospho) for my PIPS, a small, award-winning scientific animation and illustration company based in Tottenham, London. Phospho specialise in visually stunning, data-driven artwork used for a variety of purposes, both academic and industrial. Their imagery has featured on the covers of several prominent journals, and animations commissioned by leading industrial companies. I have a five-plus year background in scientific illustration, initially as a hobby and more recently as freelance commercial work. However, due to hardware and time constraints, I had little experience in animation. I produce and render all my own illustrations using the open-source software Blender, whereas Phospho, uniquely, work exclusively with the commercial software Houdini (SideFX) and Redshift (Maxon). At Phospho I worked under the founder Dr Jeroen Claus, and alongside two biomedical animators.

During my time at Phospho, I produced several commissioned pieces. The first being a journal cover submission for Professor Elton Zeqiraj, for research published in Cell. The second was a product animation for Oxford Nanopore, and the third being a proof-of-concept animation for presentation at a conference for Dr Chris Tape. The first challenge to overcome during the PIPS was learning how to competently Houdini, a vastly more complex and powerful software than Blender, featuring completely different workflows and code-based input. I believe I completed the PIPS with much greater understanding and knowledge of the software, thanks to the help from all members of Phospho. The artwork I produced for the Cell cover submission consisted of a fully 3D rendered image (figure 1, A), alongside photography of a hexagonally tessellated origami model (figure 1, B), complete with drawn representations of the protein subject matter of the publication. Both concepts I produced, refined through continual input from both Jeroen and Elton and ultimately, although not selected for that issue of Cell, artwork I am proud of.

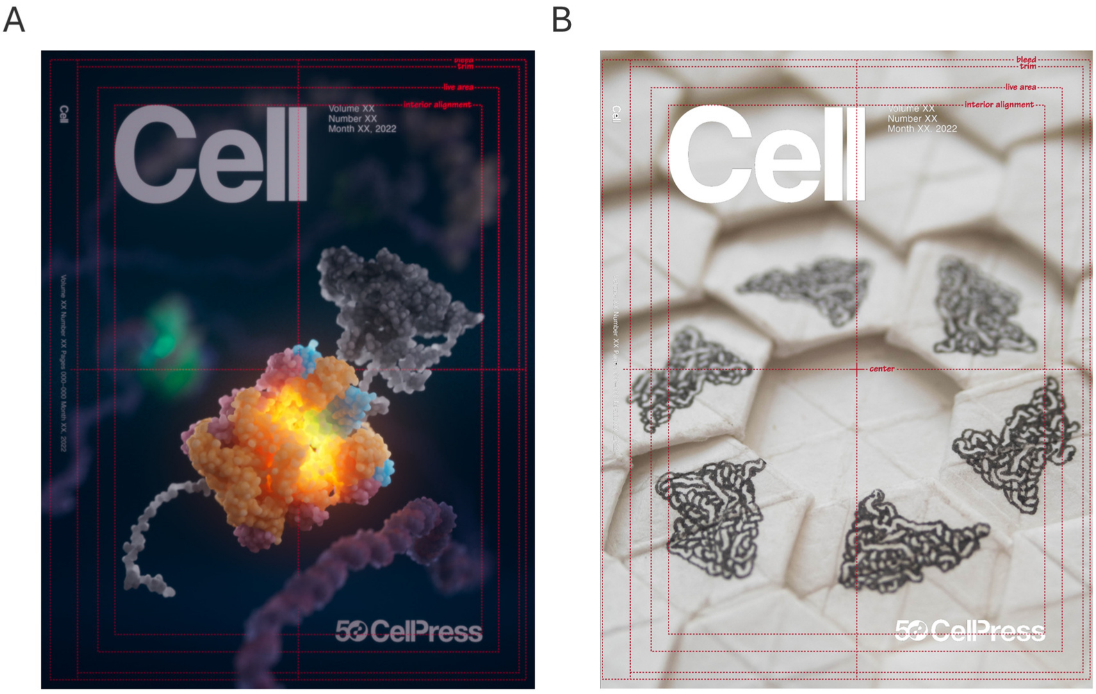


Figure 1 | Example artwork produced on my PIPS for Cell cover image submission.

The work I produced for Oxford Nanopore was a one-minute animation, showcasing their long-read RNA sequencing capabilities, since used in multiple promotional materials (figure 2). Here I learned not only how to produce high-quality animations, but also the process of working with a client. I gained video and audio editing skills, knowledge of client-lead workflows and gained confidence in showing my work for critique and improvement, from both Jeroen and members of the Nanopore team we communicated with.

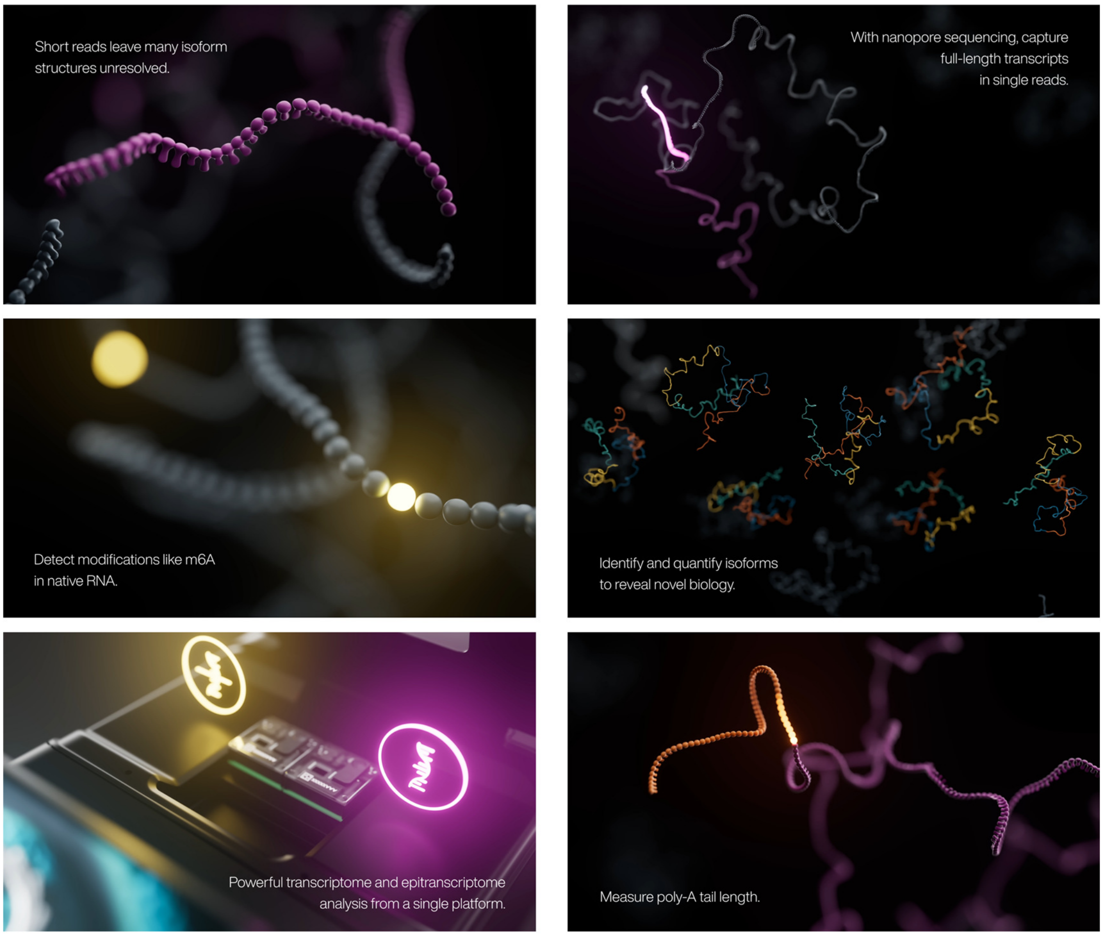


Figure 2 | Representative stills from the animation I produced for Oxford Nanopore.

Finally, the animation I produced for Dr Chris Tape was under a very tight schedule and sought to highlight his work into the phenotypic landscape of differentiating cell types (figure 3). The animation was to be shown as a part of his presentation at the American Association for Cancer Research annual meeting. During this I deepened my skillset in animation, whilst working under tight time constraints.

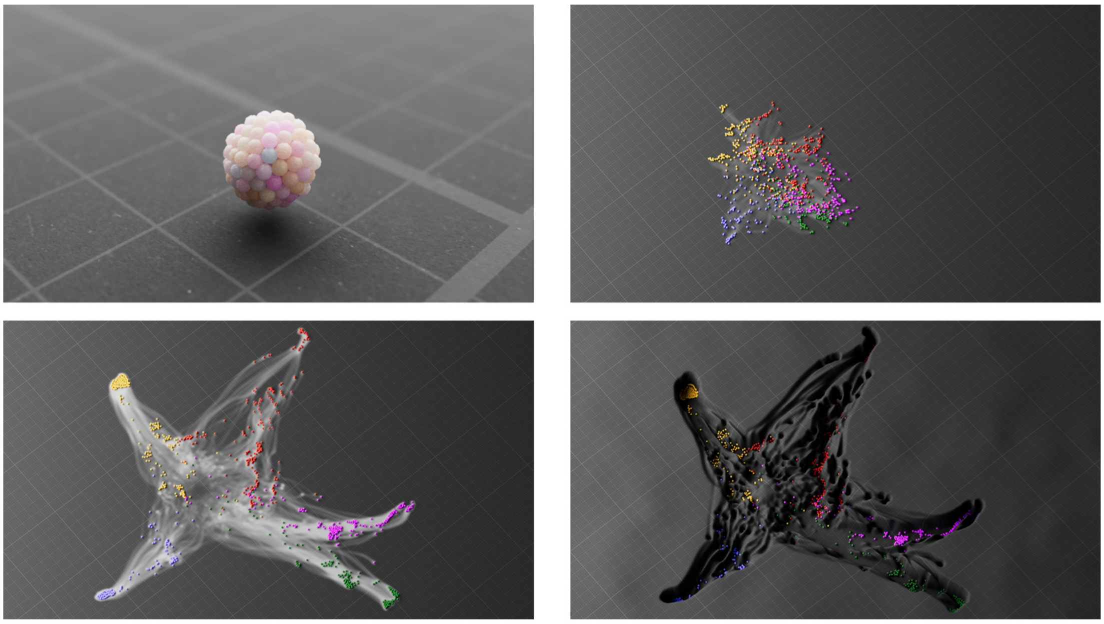


Figure 3 | Representative stills from the animation I produced for Dr Chris Tape.

Overall, I further developed my skills in scientific illustration and communication alongside gaining novel skills in animation, client relationships plus communication, and the use of novel software. These skills translate into academic research, presentation of data is an essential component to our work, and visually appealing representations of research aid its appeal to a wider audience. Working at Phospho allowed me to realistically consider the pursuit of scientific animation and illustration as a career path, and if not, the skills I gained are transferable to the presentation of my own data. I wish to offer my thanks to all those at Phospho who helped me during my placement, especially Jeroen. I would also like to acknowledge their financial support for my accommodation in London for three months, without which the placement would not have been possible.