Structural studies of macromolecules of therapeutic interest



Mattia Bertinelli

Medical Sciences Doctoral Training Centre

Wolfson College

University of Oxford

A thesis submitted in fulfilment of the requirements for the degree of *Doctor of Philosophy* at the University of Oxford

Trinity Term 2020

To my family and to Holly,
for their continuous support

Abstract

Human beings are constantly challenged by nature on multiple fronts. Bacteria and viruses directly threaten lives and our inability to treat certain chronic diseases, as well as the rise of antimicrobial resistance are complicating the problem further. Indirectly, agricultural pests put human lives in danger, by causing billions of dollars worth of damages and reduced access to food, especially in poorer areas of the world that are already more at risk of infections. Here, I present three projects where I attempt to tackle some of these issues.

Despite the availability of a vaccine, the Hepatitis B virus is still a major burden for healthcare systems. Worldwide, deaths associated with a Hepatitis B virus infection are higher than those associated with HIV. In this project, I present attempts at expressing and purifying the virus polymerase, as the basis to obtain its long sought-after structure.

The rise of antimicrobial resistance among bacterial species is one of the main challenges in public health. The New Delhi metallo β-lactamase is a bacterial enzyme capable of hydrolysing a broad variety of β-lactam antibiotics and it is resistant to inhibitors. In this project I present two structures of this enzyme bound to two Fabs derived from monoclonal antibodies. I describe how these antibodies, despite binding a region of the enzyme distant from its active site, exert an inhibitory or enhancing effect on its activity. Furthermore, I show how the presence of palmitic acid appears to induce the formation of clusters of the enzyme.

Finally, a growing threat to agriculture is that of Spodoptera frugiperda, a moth whose larvae feed on a broad variety of cultivated plants. In this project I present the serendipitous discovery of the structure of its only TNF ligand, Eiger. I also model a possible arrangement of the ligand with its receptor and discuss how this arrangement could support an avidity-enhancing clustering at the membrane. I also identify a large number of histidine residues on the surface of the ligand, some of which are conserved and could play a role in the mechanism of the ligand.

Declaration of work

All the work described within this thesis is my own, except:

Chapter 3 (Hepatitis B polymerase)

- Assistance was provided in the collection of fluorescence microscopy images by Dr. Jonathan Elegheert
- Some assistance was provided in the generation of the recombinant baculoviruses with the biGBac system by Dr. Jeremy Keown

Chapter 4 (New Delhi metallo β-lactamase)

- Generation of the panel of monoclonal antibodies was done by an external company (ENVIGO), under the direction by our collaborators at the Oxford Vaccine Group (Dr. Andrew Pollard group). Preliminary biochemical assays were performed by the Oxford Vaccine Group. Sequencing of hybridomas of monoclonal antibodies by Dr. Juthathip Mongkolsapaya.
- The assay that originally pointed towards the interaction between NDM-1 and Dengue virus stems from a project of Dr. Max Renner.
- The assay that originally pointed towards the interaction between NDM-1 and palmitic acid stems from a project of Luca Vecchia and FSEC was jointly collected
- Mounting and cryocooling of crystals was done by Dr. Karl Harlos
- Some assistance was provided in the structure refinement by Dr. Max Renner
- Some assistance was provided in setting up and running the follow-up biochemical assays for NDM-1 by Dr. Jurgen Brem
- Attempts at generating nanobodies were performed by Dr. Jiandong Huo

- Mounting of crystals on EM grids and subsequent SEM imaging was done by Dr. Adam Crawshaw.
- Data collection for Cryo-EM was performed by Dr. Jeremy Keown and Dr. Loic Carrique
- Some assistance was provided for processing of EM data by Dr. Jeremy Keown

Chapter 5 (Eiger)

- The project stems from a collaboration between me and Dr. Max Renner and is based on a publication co-authored by him, Dr. Max Renner, Prof. Jonathan Grimes and myself. Parts of it are quoted verbatim.
- Dr. Guido Paesen mounted and cryocooled the crystal
- Molecular Dynamics simulations were carried out by Dr. Max Renner
- Some assistance was provided for the Molecular Replacement by Dr. Max Renner

Some figures are reproduced with permission from relevant literature. Citations for each figure are shown in the respective captions. Copyright information and license numbers are listed in **Chapter 7**.

Acknowledgements

How can you describe your DPhil? This has been the first thing that crossed my mind when I was staring at the first, inevitably blank, page of this thesis. How do you summarise four year of your life in a few paragraphs? Looking back at when I started, I knew it would be a big challenge. Little did I know back then that the final year of my DPhil would coincide with the COVID-19 pandemic.

The good thing is that these past few years I met some amazing people, that helped me grow in a professional and in a personal way. The first people I want to thank are obviously my supervisors, Jon and Dave. Jon in particular helped me a lot by "steering the boat" in the right direction whenever I needed it, but otherwise providing me with enough space to pursue any project (crazy ones included) I found interesting.

All of this would not have been possible without the support of the Wellcome Trust and of Diamond Light Source, for providing me the with the financial support needed for my studies. I also wish to thank some of my collaborators for providing support with our shared project, Jurgen in particular has been amazing at helping me set up some assays that I desperately needed. I also want to thank Karl, for helping me with really precious samples

Everyone in Jon's group has been amazing and I have to thank in particular Itziar, for showing me the ropes and Jeremy and Loic for the much needed help with electron microscopy. Big thanks again to Jeremy for the help with some lab stuff and a lot of patience. Guido is one of the people I have to thank the most, always reliable and happy to help and he has been doing so for a very long time. His

advice has always been useful, as well as his ability to listen to my long chats while working.

I was very lucky in having Luca as a colleague and a friend: being able to have another Italian voice in the lab has been really comforting throughout this journey. We helped and supported each other on so many occasions I can't even count them. Best of luck for everything!

Finally, huge thanks to Max for all his help throughout my DPhil. We have collaborated (we still are) on a lot of projects, including the ones we weren't even looking for! He's a great friend and one of the few people that appreciates my peculiar brand of optimism. Somehow he also likes pretty much all the things I like. Again, best of luck for everything! (And at some point we should definitely finish our Divinity co-op game)

Lab life, would have been not as great without all the people that shared the office with me. We had many of constructive (and less constructive) discussions, so thank Vicky, Helen, Rhys and Weng! Strubi has been my home for these past four years and unfortunately it would be impossible to adequately thank all of the people I have met.

All of this would not have been possible without the continuous support of my family, and their example has shown me how hard work and dedication always pays off in the end.

Last, but not least (in fact she should be first on this list), I want to thank Holly, for dealing with me every day and for simply being just amazing. At the end of these four years, I look forward to starting the next chapter of my life with her.

List of abbreviations

atm: atmosphere

BAFF: B-cell activating factor

BLAST: basic local alignment search tool

BLI: biolayer interferometry

cccDNA: covalently closed circular DNA

CDR : complementarity-determining regions

CRD: cysteine rich domain

CRISPR: clustered regularly interspaced short palindromic repeats

CTF: contrast transfer function

Dlp: dally-like protein

DMEM: Dulbecco's Modified Eagle's medium

DMSO: Dimethyl sulfoxide

DNA: deoxyribonucleic acid

E.coli: escherichia coli

EDTA: ethylenediaminetetraacetic acid

ELISA: enzyme-linked immunosorbent assay

ESCRT: endosomal sorting complexes required for transport

Fab: fragment antigen-binding

FACS: fluorescence-activated cell sorting

FBS: fetal bovine serum

FPLC: fast protein liquid chromatography

GCTL: galactose-specific binding C-type lectin

GST: glutathione s-transferase

HBcAg : hepatitis B core antigen

HBeAg: hepatitis B virus e antigen

HBsAg: hepatitis B surface antigen

HBV: hepatitis B virus

HCl: hydrochloric acid

HEK: human embryonic kidney

HIV: human immunodeficiency virus

HPLC: high-performance liquid chromatography

HRP: horseradish peroxidase

Hsp: heat-shock protein

IgG: immunoglobulin G

IMGT: immunogenetics information system

IPTG: Isopropyl β- d-1-thiogalactopyranoside

IRES: internal ribosome entry site

K.pneumoniae: Klebsiella pneumoniae

MALDI: matrix-assisted laser desorption/ionization

MBP: maltose binding protein

NCS: non-crystallographic symmetry

NDM-1: New Delhi metallo-β-lactamase 1

NLS: nuclear localisation signal

NTP: nucleotide tri-phosphate

OD: optical density

OMV: outer membrane vesicle

PBS: phosphate buffer saline

PCR: polymerase chain reaction

PDB: protein databank

pgRNA: pre-genomic RNA

RIPA: radioimmunoprecipitation assay

RNA: ribonucleic acid

Rpm: revolutions per minute

RT domain: retrotranscriptase domain

SDS-PAGE: sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEC: size-exclusion chromatography

SEM: scanning electron microscope

SF9: *Spodoptera frugiperda* 9 cells

SUMO: Small Ubiquitin-like Modifier

TACE: TNF alpha converting enzyme

TEM: transmission electron microscope

TLS: translation-liberation-screw-rotation

TNF: tumour necrosis factor

TP domain: terminal protein domain

TRAF: TNF receptor associated factor

Table of contents

1 Outline of the thesis 1

2 Materials and Methods 6

2.1 General Methods 6

2.1.1 *PCR* 6

2.1.2 *RT-PCR* 8

2.1.3 *Cloning* 8

2.1.4 *E.coli transformation* 9

2.1.5 *Plasmid extraction* 9

2.1.6 *Plasmids* 9

2.1.7 *SDS-PAGE and Western blotting* 12

2.1.8 *Insect cells culturing* 13

2.1.9 *Insect cells transfection and baculovirus generation* 13

2.1.10 *HEK cells culturing* 14

2.1.11 *Mammalian cells transfections* 14

2.1.12 *Lentivirus generation* 15

2.1.13 *Fluorescence microscopy* 16

2.1.14 *E.coli cells culturing and protein expression* 16

2.1.15 *Differential scanning fluorescence (Thermofluor)* 16

2.1.16 *Miscellaneous* 16

2.2 Structural Biology Methods 17

2.2.1 *Crystallization* 17

2.2.2 *Data collection and refinement* 18

2.2.3 *Crystals SEM* 18

2.2.4 *Cryo-EM sample preparation*.....17

2.2.5 *Cryo-EM data collection and single-particle analysis of the NDM-1-Fab 5 complex* 19

2.3 Hepatitis B virus 20

2.3.1 *Bioinformatics* 20

2.3.2 *Cloning* 21

2.3.3 *Expression and purification tests* 21

2.4 NDM-1 22

2.4.1 *Expression and purification of NDM-1* 21

2.4.2 *Antibodies generation*.....23

2.4.3 *Preliminary assays* 23

2.4.4 *Generation of Fab fragments from monoclonal antibodies* 23

2.4.5 *Expression and purification of recombinant fab fragments* 24

2.4.6 *NDM-1-Fab complex formation* 24

2.4.7 *Crystallization, data collection and refinement* 24

2.4.8 *NDM-1-Palmitic acid complex formation* 26

2.4.9 *NDM-1 inhibition assays*.....26

2.5 SfEiger.....27

2.5.1 *Crystallization, data collection and refinement* 27

2.5.2 *Bioinformatic identification of SfEiger* 28

2.5.3 *Structural analysis*.....28

2.5.4 *Molecular dynamic simulations*.....28

3 Attempts at expression and purification of the Hepatitis B virus polymerase.....29

3.1 Introduction 29

3.1.1 *A brief overview of the Hepatitis B Virus*.....29

3.1.2 *Structure of the Hepatitis B Virus virion*.....33

3.1.3 *The Hepatitis B Virus genome and proteins*.....34

3.1.4 *Life cycle and replication mechanism of the Hepatitis B Virus* 37

3.1.5 *The structure of a viral RNA polymerase* 41

3.1.6 *The Hepatitis B Virus Polymerase* 43

3.1.7 *Aim of the project* 45

3.2 Results 47

3.2.1 *Construct design and bioinformatics*.....47

3.2.2 *Cloning of the Hepatitis B polymerase*.....51

3.2.3 *Expression tests with lentiviral vectors* 53

3.2.4 *Expression tests in E.coli*.....54

3.2.5 *Transient expression tests in HEK293T* 56

3.2.6 *Expression tests in insect cells* 58

3.2.7 *Purification tests of the HBV polymerase* 61

3.3 Discussion.....67

4 Structural studies on antibodies inhibiting the activity of the New Delhi Metallo β-lactamase 1 (NDM-1) 72

4.1 Introduction 72

4.1.1 *A brief overview of the New Delhi Metallo β-lactamase 1* 72

4.1.2 *New Delhi Metallo β-lactamase 1 bioinformatic analysis* 75

4.1.3 *New Delhi Metallo β-lactamase 1 structural analysis* 80

4.1.4 *Aim of the project* 83

4.2 Results 85

4.2.1 *A panel of monoclonal antibodies exhibits inhibitory activity against NDM-1* 85

4.2.2 *Cloning, Expression and purification of NDM-1* 87

4.2.3 *NDM-1 is bound by Fabs 5 and 17* 88

4.2.4 *NDM-1 forms a complex with palmitic acid*.....91

4.2.5 *Crystallization, data collection, structure solution and refinement of Fab 5 and 17, complexes with NDM-1* 95

4.2.6 *SEM imaging of NDM-1-Fab 5 crystals*.....99

4.2.7 *Nanobodies* 100

4.2.8 *Sample preparation, data collection and processing of the NDM-1-Fab 5 complex by Cryo-EM* 101

4.2.9 *Overview of the Fab fragments* 107

4.2.10 *The structure of the NDM-1-Fab 17 shows an unexpected binding site* 107

4.2.11 *Fab 5 binds a similar epitope to that of Fab 17 on NDM-1* 111

4.2.12 *Comparison between the two Fabs* 112

4.2.13 *Cryo-EM reveals the formation of higher order assemblies of NDM-palmitic acid* 113

4.2.14 *NDM-1 inhibition assays* 115

4.3 Discussion 117

4.3.1 *Future work* 123

5 High-resolution crystal structure of arthropod Eiger TNF suggests a mode of receptor engagement and altered surface charge within endosomes 128

5.1 Introduction 128

5.1.1 *A brief overview of the Eiger protein* 128

5.1.2 *Aim of the project* 130

5.2 Results 131

5.2.1 *Identification of SfEiger* 131

5.2.2 *The structure of SfEiger TNF domain* 138

5.2.3 *SfEiger forms a canonical trimer* 141

5.2.4 *Model of SfEiger – TNF receptor interactions* 146

5.3 Discussion 149

5.3.1 *Future work* 153

6 Conclusions and final remarks.....155

7 Copyright info.....158

8 References.....160

9 Appendix: List of relevant sequences 190

9.1 Table of sequences used in the generation of the phylogenetic tree in **Figure 7**.....190

9.2 Table of PDB codes of structures used in the structural phylogenetic tree in **Figure 91**.....191

9.3 Sequences of Fab 5 and Fab 17 191

1 Outline of the thesis

This work originated from the desire of obtaining long sought-after structural knowledge on the hepatitis B virus polymerase (**Chapter 3**), which is used by the virus to replicate its genome. From the first identification of its transcript in 1988 [1], there have been multiple attempts at expressing the enzyme in different expression systems, both prokaryotic and eukaryotic (more specifically in insect cells). Despite significant efforts and some reports of high level of expression [2], [3], suitable for structural studies, no structure of the polymerase has been reported to date. With recent advances in the field of structural biology, in particular Cryo-Electron Microscopy, I believe there could be a path to obtaining important information to use towards the development of new treatment options against the virus.

While initial work on the project proceeded smoothly, it became immediately clear that expression of the viral enzyme would be a significant challenge. A large number of different constructs were thus generated and tested in different expression systems, following a two-pronged approach. One approach was to truncate the polymerase in several ways, based on both bioinformatic analysis and available cellular expression data, to increase the chances of obtaining stable, crystallizable fragments. The other approach relied on the available literature showing the need for the presence of host cell chaperones for the stability of the apo-enzyme. Thus, the full length polymerase was co-expressed with said enzymes. Unfortunately, given the lack of expression of soluble protein and the diversion caused by a contaminant during the purification, the project was deemed unsuccessful and it was decided to stop any further work.

As we were fully aware of the high risk of the hepatitis B polymerase project, a separate line of research was also pursued, involving the New Delhi Metallo β-Lactamase 1 (NDM-1), a bacterial enzyme capable of hydrolysing β-lactam antibiotics. Due to its resistance to inhibitors [4], broad spectrum activity [5] and potential for horizontal gene transfer [6], the enzyme is a direct threat to global public health. This second project (**Chapter 4**) was done in collaboration with the Oxford Vaccine Group, which identified a panel of monoclonal antibodies targeting the enzyme. Two of these antibodies (5 and 17) showed an inhibitory effect against NDM-1.

The initial focus for the project was to obtain structures for the complexes between NDM-1 and the antibodies, processed into Fab fragments for structural analysis. Over time, it became increasingly clear that our collaborators would not be able to supply us the necessary amounts of antibodies required for crystallization and we decided to recombinantly express and purify the Fab fragments ourselves. NDM-1 proved to be harder to crystallise than anticipated and a structure for the NDM-1/Fab17 complex was obtained from a crystal which grew over the course of ~280 days.

As for the complex between NDM-1 and Fab5, several crystals grew in multiple screened conditions within ~60 days. All of them, however, were in the form of thin needles and needle clusters. Despite extensive optimisation work, complicated by the relatively long time required for crystallization, even optimised crystals retained the same morphology and did not diffract or did so very poorly. A crystal of Fab5 alone eventually grew after ~360 days. In order to

obtain the structure of the missing complex, we employed Cryo-EM. Through several steps of sample optimisation it was possible to obtain a good level of contrast and collect a reasonably sized dataset. While 2D classes look remarkably good, it was not possible to achieve the desired high resolution. While the reason for this could not be determined, the small size of the particles may have posed a challenge during alignment. Furthermore, it is worth noting that data was not collected using an energy filter. Nonetheless, fitting of the crystal structures of NDM-1 and Fab5 was unambiguous and deemed sufficient to identify its epitope. Perhaps puzzling, both antibodies bind very far from the active site raising questions on their mechanism of action.

Alongside the structural analysis, I also carried out some biochemical assays to ascertain the activity of the Fab fragments against NDM-1. These assays painted a somewhat more complex picture than the preliminary data suggested: while the inhibition is consistent (albeit low-level) for all Fabs with a specific substrate, when the substrate changes one of the Fabs does not show any activity, while the other appears to enhance the activity. It has been previously reported than mutations of NDM-1 far from the active site are capable of altering its activity [7] and the similar epitope bound by the two Fabs suggests a region of NDM-1 worth investigating in the development of new therapies.

One final, unexpected, finding about NDM-1 is its apparent ability to bind palmitic acid, which was discovered through its unexpected behaviour as a negative control in two different unrelated studies. This behaviour was experimentally confirmed and preliminary data by Cryo-EM show apparent higher-order assembly of NDM-1 upon addition of palmitic acid.

The NDM-1 line of research, however, has been hindered by the COVID-19 outbreak, as laboratory activity was suspended. For this reason it was not possible to carry out further experiments aimed at elucidating the role of palmitic acid within NDM-1, as well as obtaining better structural and biochemical data on the NDM-1-Fab complexes, especially when taking into account the slow crystallization of NDM-1. Further information on planned experiments is given in (**4.3.1**)

A third line of research (**Chapter 5**) presented in this thesis is the fruit of a serendipitous discovery stemming from a project in collaboration with Dr. Max Renner. While trying to obtain the structure of the lammi virus glycoprotein, a high-resolution dataset was collected from a single crystal. Initial attempts at phasing using a model related to the intended target failed. For this reason, hypothesizing a contaminant from the expression system, the Wide Search Molecular Replacement server [8] was used to identify a suitable molecular replacement model.

As no further information about this protein was available, a somewhat unconventional strategy was used where a partially completed sequence was obtained by manually building unambiguously identifiable residues into the electron density. By reverse-searching the genome of *S.frugiperda* using tBLASTn, it was possible to identify the unknown protein as Eiger. Eiger is the only TNF homologue in arthropods [9] and our structure is the first reported for any of the Eiger proteins and is involved in a variety of cellular processes, such as tissue development and remodelling, immunity and pain sensitivity. Eiger exerts its activity through the activation of the Jun kinase cascade[10], upon binding its receptors (termed Wengen and Grindelwald in arthropods) [11] and

its subsequent internalisation. By a careful examination of the structure and the crystal packing it was possible to identify a striking number of histidine residues decorating the surface of the protein, likely explaining how we came to co-purify this protein along with our target of interest. Histidine residues act as a pH sensor, which could point towards a possible role for at least some of these residues (including one in the putative receptor-binding site) in its mechanism of action. Analysis of the packing in the crystal suggests a higher order assembly of tightly packed Eiger trimers, with sufficient space to accommodate its receptors. Furthermore, this arrangement could provide an enhancement of the avidity towards Eiger. Finally, the structure of Eiger suggests that the trimeric arrangement of TNF-like ligands has been conserved since the split between Protostomes and Deuterostomes, which occurred more than 500 millions of years ago[12].

Further studies will be necessary to validate and expand upon these findings. However, due to the serendipitous nature of this work, no further experiments were planned.

2 Materials and Methods

2.1 General Methods

2.1.1 PCR

Polymerase chain reactions were performed to amplify the genes of interest by using Phusion polymerase master mix (NEB) to a final concentration of 1X. The general protocol for gene amplification uses the following parameters (***Table 1***):

Step	Temperature	Length	
Initial denaturation	95C	3 minutes	
Denaturation	95C	30 s	35 cycles
Annealing	Determined by primers	30 s	
Extension	72C	30 s/kb	
Final extension	72C	5 minutes	
Storage	4C	∞	

Table 1: Standard PCR protocol followed for the amplification of DNA fragments.

Annealing temperature was determined using the online tool from NEB (http://tmcalculator.neb.com/#!/main) to account for the specific needs of the enzyme.

The protocol used to introduce mutations and/or other DNA sequences is based on the overlap-extension PCR [13]. Briefly, two DNA fragments are generated using the standard protocol outlined in ***Table 1***. These fragments contain an overlapping region with the other fragment spanning about 24 nucleotides. In the case of a mutation, the new sequence for the mutation is contained in the middle of the overlapping region. The fragments are then purified by agarose gel electrophoresis, followed by gel

	Step	Temperature	Length	
No primers added	**Initial denaturation**	95C	1 minute	
	Denaturation	95C	30s	3 cycles
	Annealing	Determined by overlap	30s	
	Extension	72C	30s/kb	
	Final extension	72C	5 minutes	
With primers added	**Initial denaturation**	95C	1 minute	
	Denaturation	95C	30s	35 cycles
	Annealing	Determined by overlap	30s	
	Extension	72C	30s/kb	
	Final extension	72C	5 minutes	
	Storage	4C	∞	

Table 2: Standard protocol followed during overlap-extension PCR

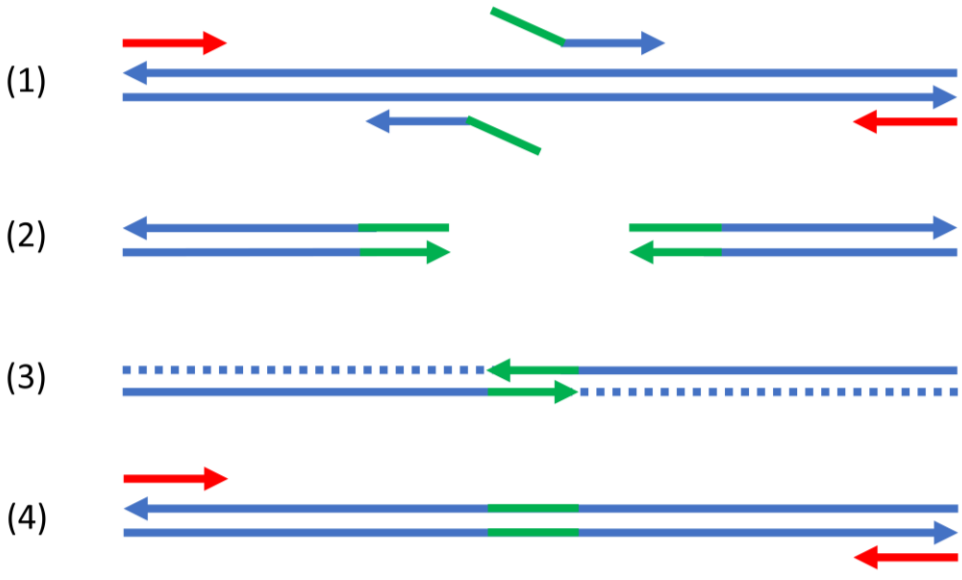


Figure 1: Generation of mutants using PCR. (**1**) Two separate PCRs are setup, with two primers (green) carrying the mutation/insertion/deletion, to obtain 2 complementary fragments (**2**). (**3**) The two fragments are joined for 3 PCR cycles, without external primers (red). (**4**) External primers are added for amplification of the resulting product.

extraction using a kit (QIAGEN), following manufacturer instructions. The fragments were then combined and the protocol in **Table 2** was used (**Figure 1**). The resulting PCR product was again purified by agarose gel electrophoresis and gel extraction. If necessary, higher DNA yields were obtained by doing a standard PCR on the purified product.

2.1.2 RT-PCR

RNA from $1*10^6$ cells was extracted using the RNeasy mini kit (QIAGEN) and incubated with RQ1 RNAse-free DNAse (Promega) to remove eventual DNA contamination. Reverse transcription polymerase chain reaction was then performed using the Superscript IV enzyme (ThermoFisher) following the manufacturer instructions.

2.1.3 Cloning

Cloning of the desired gene of interest into a suitable vector was performed either by ligation-based cloning with restriction enzymes or with In-fusion Eco-dry (Takara) ligation independent cloning.

For ligation-based cloning, the transfer plasmid (see **2.1.6**) and the inserts were digested using EcoRI and KpnI to bypass the secretion signal and allow intracellular expression. The digested DNA was loaded onto an 1% agarose gel and the band excised and purified using a Gel-extraction kit (QIAGEN) following manufacturer instructions. 100ng of cut plasmid and the appropriate amount of insert were then mixed in a 3:1 molar ratio and the ligation reaction was performed using a T4 DNA Ligase (Promega) with an overnight incubation at 4C. 2µL of the resulting DNA were used to transform *E.coli* DH5α cells, as described in **2.1.4**.

In the second case, the reaction was performed according to manufacturer instructions. Briefly, the vector was cut with the appropriate restriction enzymes to linearize it, followed by purification by agarose gel electrophoresis and gel extraction. A mixture of 100ng of linearized vector and 100ng of insert for a final volume of 10µL was used to resuspend the lyophilized enzyme. The reaction was then incubated at 37C for 15 minutes, then at 50C for 15 minutes. The mixture was then diluted with 40µL of TE buffer (QIAGEN) and used for the subsequent transformation step (see **2.1.4**).

2.1.4 *E.coli* transformation

E.coli DH5α cells (Invitrogen) were transformed by heat-shock. Briefly, 50µL of thawed cells were incubated on ice with 50ng or more of plasmid DNA for 15 minutes. The cells were then heat-shocked at 42C for 45 seconds and immediately transferred to ice for 5 minutes. The cells were then resuspended in 400µL of SOC media (Invitrogen) and cultured at 37C for hour. A 100µL aliquot was then plated on LB-Agar plates supplemented with the appropriate antibiotic. The plates were then incubated at 37C overnight.

2.1.5 Plasmid extraction

Selected colonies from the cloning step were grown overnight at 37C as indicated, then centrifuged (4000 x g, 4C, 10 minutes) discarding the supernatant. Using the Miniprep kit (QIAGEN) following the manufacturer instructions the plasmid DNA was obtained and sequenced (GATC).

2.1.6 Plasmids

A system[14] composed of 3 plasmids was used to generate recombinant lentiviruses: the transfer plasmid (***Figure 2***) is a custom plasmid that allows the

insertion of the gene of interest between 5' and 3' LTR sequences and will be integrated in the genome upon infection. The plasmid also contains the gene for a fluorescent protein under an IRES, which is used to rapidly assess transfection efficiency and, if needed, enrich a specific subpopulation of cells via FACS. Envelope and packaging plasmids instead contain the remaining genes that are necessary for the virus production and are commercially available (Addgene plasmids no.12259 and 12260, respectively).

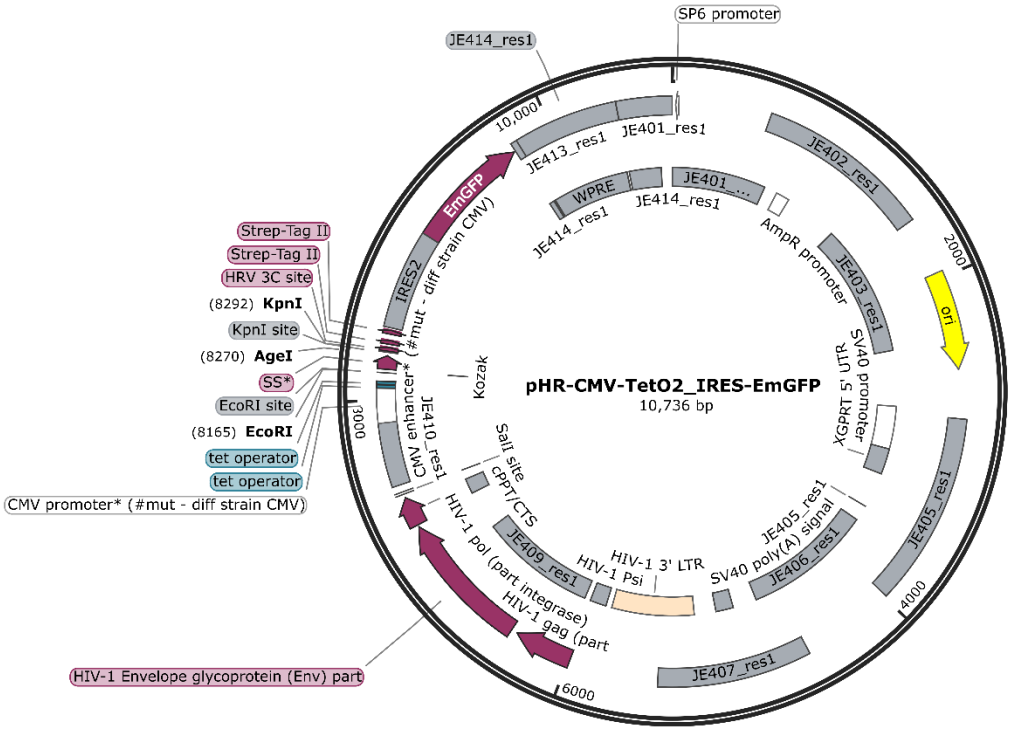


Figure 2: Vector map for the transfer plasmid used for generation of the recombinant lentivirus. Some relevant sites are highlighted. The plasmid map was generated using the SnapGene software.

Plasmids from the pOPIN suite of vectors were used For *E.coli* expression, baculovirus generation in insect cells and transient transfection in mammalian cells [15]. Most of the vectors used support expression in all the aforementioned organisms (***Figure 3*** and ***Table 3***).

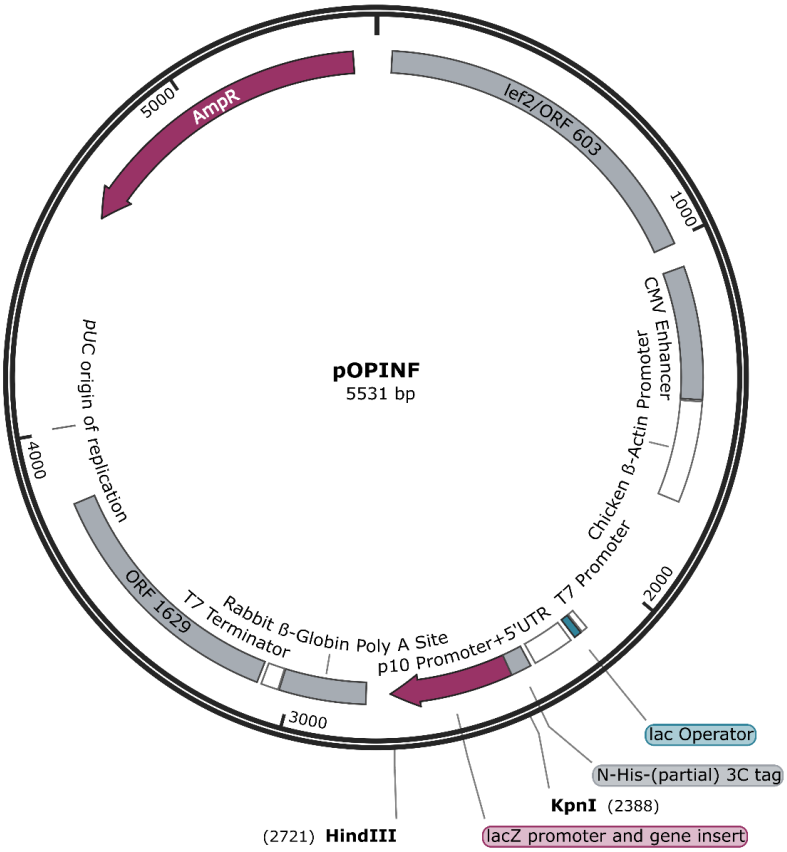


Figure 3: Plasmid map for pOPIN-F. Restriction sites used for plasmid linearization are highlighted. Some relevant sites are highlighted. The plasmid map was generated using the SnapGene software.

Vector	Product	Notes
pOPINF	HISx6 - 3C - POI	
pOPINS3C	HISx6 - SUMO - 3C - POI	
pOPINM	HISx6 - MBP - 3C - POI	
pOPINJ	HISx6 - GST - 3C - POI	
pOPINFB	HISx6 - 3C - POI - BAP	
pOPINVH	SS[RPTPmu] - POI - VH - KHISx6	HEK293/Baculo only
pOPINVL	SS[RPTPmu] - POI - VL - KHISx6	HEK293/Baculo only

Table 3: List of pOPIN vectors used. POI: protein of interest, BAP: biotin acceptor peptide, SS[RPTPmu]: secretion signal of µ-phosphatase.

Table 4 contains a list of constructs that were subcloned in each vector.

In particular, pOPINVH/VL were used to express recombinant Fab fragments and pOPINFB was used to express biotinylated NDM-1. In this case, a biotin acceptor peptide (GLNDIFEAQKIEWHE) is genetically fused at the C-terminus of the protein. Expression is carried out by co-transforming *E.coli* using both the pOPINFB plasmid with NDM-1 and a plasmid containing BirA (biotin ligase) [16].

2.1.7 SDS-PAGE and Western blotting

Samples were loaded onto a NuPAGE 4-12% SDS-PAGE gel using a NuPAGE MES SDS Running Buffer (Invitrogen) and the electrophoresis was ran at 180V. Proteins separated by SDS-PAGE were transferred to a nitrocellulose membrane (GE Healthcare) by western blotting using a Mini Blot Module (Invitrogen). For his-tag detection, the membrane was blocked by incubating with 5% milk-PBS. After blocking, the membrane was incubated for 3 hours at room temperature with an anti-his tag antibody (QIAGEN) diluted 1:5000 in milk-PBS supplemented with Tween 20 (0.05% - PBS-T) . After extensive washing with PBS-T a secondary goat anti-mouse antibody coupled with HRP (Sigma) diluted 1:5000 in PBS-T was added and incubated for 1 hour at room temperature. After more extensive washes with PBS-T, detection was carried out using Clarity™ Western ECL substrate (Bio-Rad) following manufacturer instructions.

For anti-twin-strep tag, blocking was performed with 1% BSA-PBS, followed by incubation with biotin-HRP (IBA) diluted 1:1000 in PBS-T as per manufacturer instructions. After extensive washes with PBS-T detection was carried out using Clarity™ Western ECL substrate (Bio-Rad) following manufacturer instructions.

2.1.8 Insect cells culturing

Spodoptera frugiperda SF9 cells were cultured in SF-900 II media (Gibco) at 27C with the addition of penicillin-streptomycin solution (Sigma) to a final concentration of 1X (100 units of penicillin/100µg streptomycin per L of media). Cultures were shaken at 120 rpm and cells were maintained between $0.6*10^6$ and $6*10^6$ cells/mL.

2.1.9 Insect cells transfection and baculovirus generation

Approximately $1*10^6$ SF9 cells were seeded in a 6-well plate, leaving them for 15 minutes to allow adhesion to the well. The cells were washed 2 times with SF-900 II media, then covered with 2mL of the same media supplemented with 1.5% of FBS (Invitrogen). In a vial, 100ng of bacmid (Oxford Expression Systems) and 2-4µg of plasmid DNA were resuspended in 100µL of SF-900 II media. In a separate vial, 4µL of Cellfectin II (Invitrogen) transfection reagent were resuspended in 100µL of SF-900 II media. The two vials were mixed and the resulting mixture was pipetted dropwise on the well. The cell culture was incubated at 27C for 5 hours, then the media was replaced with SF900 II media without serum, supplemented with 1X (final concentration) of penicillin-streptomycin (Sigma). The cells were then incubated for 72 hours.

After the incubation period, cells were resuspended and added, together with the media, to a 30mL culture of SF9 cells at ~$1*10^6$ cells/mL and incubated for at least 72 hours to generate the first viral stock. Empirical determination of a satisfactory infection was assessed using by staining a small aliquot of cells at t=72 hours with Trypan blue. Dead cells under a light microscope appear blue. A satisfactory infection was characterised by the presence of swollen, non stained

cells and ~20% of dead cells. This step was repeated one more time to generate the second viral stock using 250µL of the first to infect a 30mL culture.

For protein production, cultures of 250mL of cells were infected with 4mL of the second viral stock containing the recombinant baculovirus with the gene of interest. Cells were incubated for at least 72 hours assessing the infection stage using Trypan blue as detailed above. The cultures were centrifuged (4000 x g, 4C, 20 minutes), the supernatant discarded and the pellet was frozen at -20C until needed.

2.1.10 HEK cells culturing

Human embryonic kidney 293 cells were cultured in DMEM media (Sigma) supplemented with L-glutamine (Gibco, final concentration 2mM), Non-Essential Amino Acid (Gibco, final concentration 1X) and 10% FBS. Cells were cultured at 37C with 5% CO_2.

2.1.11 Mammalian cells transfections

For small-scale transient transfection, HEK293T cells cultured in DMEM in a 6-well plate were transfected with 4µg of DNA using Lipofectamine 2000 (Thermo-Fisher) following manufacturer instructions. After 72 hours of incubation (37C, 5%CO_2), cells were harvested by centrifugation (1000xg, 4C, 5 minutes) and cell pellets lysed at room temperature using RIPA buffer (150 mM NaCl, 1.0% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0). Lysate was then centrifuged (13000xg, 10 minutes) and supernatant was used for western blotting.

For large-scale transient transfection, HEK293T cells from 6 confluent T175 flasks were trypsinized and used to seed 6 roller bottles. Cells were grown in 10%

FBS DMEM for 72 hours, before replacing the media with 2%FBS DMEM. A total of 3mg of plasmids, generated using a Gigaprep kit (QIAGEN), containing the genes for the heavy and the light chain of Fabs 5 and 17 were mixed in a 1:1 ratio. To precipitate the salts and further purify the DNA for the transfection, 80µL of chloroform was mixed in and the result centrifuged (10minutes, 16000xg, 25C). The aqueous supernatant, containing the nucleic acids was collected and added to 300mL of serum-free DMEM together with 6mg of PEI (polyethylenimine) to create the transfection solution. The mixture was incubated for 10 minutes, then 2mg of kifunesine, an inhibitor of mannosidase I [17], was added during the expression of Fab 5 and 17 to ensure that an eventual glycosylation of the recombinantly expressed Fab fragments would be homogeneous. The addition of kifunesine is likely not necessary as no glycosylation site is predicted within the Fab sequences. The mixture was then equally divided between the 6 roller bottles and transfected cells were grown for 5 days.

2.1.12 Lentivirus generation

Recombinant lentiviruses were generated as described in [14]. Briefly, HEK293T Lenti-X cells were co-transfected with the 3 plasmids (see **2.1.6**, 30 µg of total DNA; 1:1:1 ratio) using polyethylenimine. After 72 hours, supernatant containing the lentivirus was collected, filtered through a 0.45µm filter and supplemented with 1X Polybrene (Sigma) to aid the subsequent infection. Efficiency of transfection was assessed by fluorescence microscopy. The cell culture supernatant, containing the recombinant lentiviral particles was then collected and was added dropwise to HEK293S GnTI^{-} TetR cells to establish a stable cell line. Addition of blasticidin (2µg/mL) to the growth medium adds selective

pressure to the cells to maintain the plasmid with the TetR inducible system. These cells allow induction of the synthesis of the protein of interest only with the addition of doxycycline to the growth medium. Lysis of cells for western blotting was performed as described in **2.1.11**.

2.1.13 Fluorescence microscopy

To assess the success of the lentiviral infection, HEK293S GnTI⁻ TetR were imaged by Dr. Jonathan Elegheert. Image processing and pseudocoloring was done using ImageJ (https://imagej.nih.gov/ij/download.html).

2.1.14 *E.coli* cell culturing and protein expression

E.coli cells were cultured in LB media at 37C with the addition of the appropriate antibiotic for the vector and strain used. Cultures were shaken at 180 rpm until an OD of 0.6 was reached. The culture was then cooled to 20C and IPTG was added to a final concentration of 100µM, then incubated overnight. The culture was then centrifuged (4000 x g, 4C, 20 minutes), the supernatant discarded and the pellet was frozen at -20C until needed.

2.1.15 Differential scanning fluorescence (Thermofluor)

Differential scanning fluorescence [18] was employed as a method to assess the thermal stability of proteins. A non-specific hydrophobic protein binding dye is mixed with the protein of interest. As the dye-protein mixture is heated, denaturation occurs and the dye can bind to the newly exposed hydrophobic cores of the protein. As the dye is sensitive to its environment, its quantum yield increases in hydrophobic environment, resulting in a higher fluorescence signal.

Experiments were performed using the SYPRO orange dye (Invitrogen - λ_{ex} 470 nm /λ_{em} 570 nm) to a final concentration of 3X using a Stratagene Mx3005P qPCR System (Agilent).

2.1.16 Miscellaneous

Figures were prepared with PyMOL (https://pymol.org) and Chimera (https://www.cgl.ucsf.edu/chimera/).

2.2 Structural Biology Methods

2.2.1 Crystallization

Crystallization screening was carried out using a Cartesian dispenser system[19] with the sitting drop method. Briefly, 100nL of concentrated protein solution are deposited onto a crystallization plate, then 100nL of a precipitant solution is added, taken from a larger reservoir. The plate is then sealed and over time, more water vapour flows towards the reservoir, increasing the concentration of protein and precipitant in the drop.

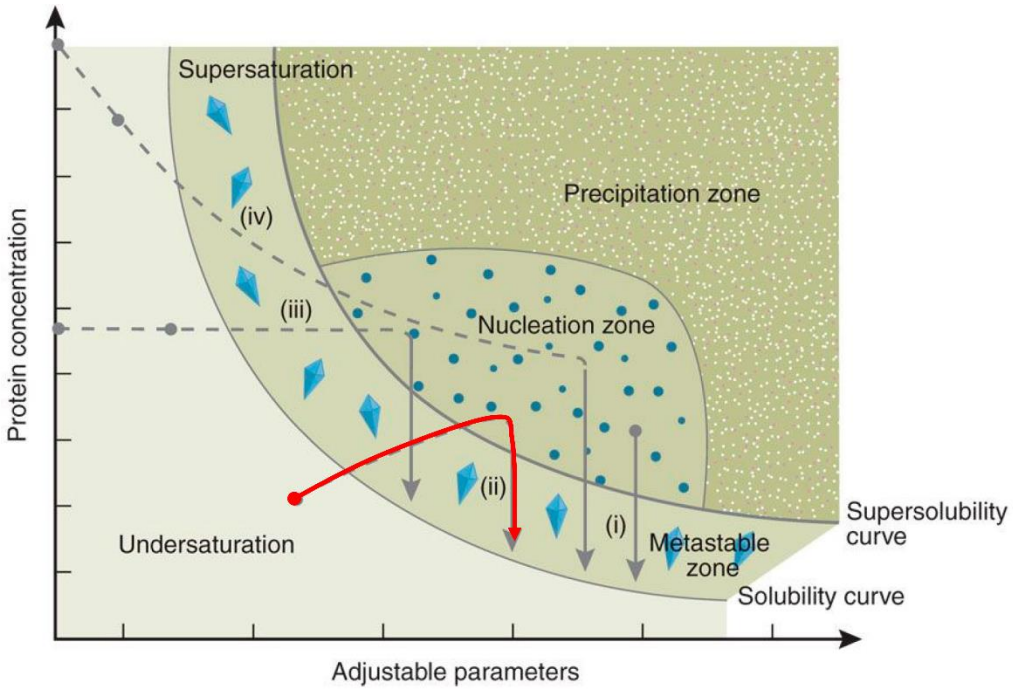


Figure 4: Protein crystallization phase diagram. Adapted from [20]. Arrows indicate the change in protein concentration throughout an experiment. Different techniques such as (**i**)Microbatch (**ii**)Vapour diffusion (**iii**) Microdialysis and (**iv**) free interface diffusion follow different patterns. The red arrow refers to the method employed in this work (vapour diffusion). In a Microbatch experiment, the protein solution and the precipitant are pre-mixed, then covered in oil to prevent evaporation. In a vapour diffusion experiment, a drop containing the protein/precipitant mixture is separated from a reservoir containing pure precipitant: evaporation of water from the protein drop continues until an equilibrium is reached. In a microdialysis experiment the drop with the protein solution is separated from the precipitant by a dialysis membrane. Finally, in a free interface diffusion experiment the protein and precipitant solutions are not premixed, but allowed to equilibrate via diffusion once in contact with each other.

The aim is to bring the concentration of protein and precipitant to a level where spontaneous nucleation and crystal growth can occur (**Figure 4**, red line). When that happens, the concentration of protein in solution slowly decreases until the crystal stops growing.

Following crystals growth, they were collected using microloops, briefly soaked in the mother liquor supplemented with 25% glycerol as a cryoprotectant and flash-cooled in liquid nitrogen.

2.2.2 Data collection and refinement

Refer to sections **2.4.7** and **2.5.1** for details on specific projects.

2.2.3 Crystals SEM

All procedures were done at Diamond Light Source (Didcot, UK) beamline I02-1 (VMXm). Crystals were collected from drops, resuspended in a small volume of mother liquor and applied to Quantifoil holey carbon copper 2/1 grids, then plunge frozen using an automatic Plunge Freezer EM GP2 (Leica). Grids were then mounted onto a sample holder and brought to vacuum for loading and imaging using an InTouchScope SEM (JEOL).

2.2.4 Cryo-EM sample preparation

All NDM-1-Fab 5 samples for electron microscopy were concentrated in (Tris 10mM pH 7.5, NaCl 50mM) and centrifuged to remove any debris (16000xg, 10 minutes, 20C). Glow-discharged gold Quantifoil 200 mesh, R2/1 grids were loaded onto a Vitrobot mark IV (FEI) system for plunge freezing at a 100% humidity and 3.5µL of the complex applied to them (Blot time: 3s, Blot force: -25). Upon plunge freezing in liquid ethane, the grids were clipped and stored under liquid nitrogen for data collection.

The gel-filtered complex of NDM-1-Fab 5 (see section **4.2.3**) was brought to a final concentration of 0.6 mg/mL (measured by nanodrop, assuming 1 AU=1 mg/mL) before plunge freezing. After several buffer optimisation steps the final buffer was (Tris 10mM pH 7.5, NaCl 50mM).

In the case of NDM-1-palmitic acid, NDM-1 (Tris 10mM pH 7.5, NaCl 50mM) at 0.3mg/mL was incubated with a 20-fold excess of palmitic acid dissolved in DMSO for an hour and directly used for plunging or loaded onto a Superdex S200 for size exclusion chromatography, collecting only the desired peak.

2.2.5 Cryo-EM data collection and single-particle analysis of the NDM-1-Fab 5 complex

Quality control and optimisation checks of plunge-frozen samples were done by imaging the grids on a 200kV Glacios Cryo-TEM (Thermo), equipped with a Falcon III detector (Thermo).

Full data collection was done on a 300kV Titan Krios Cryo-TEM with a Falcon III detector (Thermo). Parameters for the data collection are detailed in **Table 10**.

Single-particle data analysis of the NDM-1-Fab 5 complex was carried out with a combined approach of cryoSPARC[21] and RELION 3.0 [22]. A total of 2891 collected movies were motion corrected with motioncor2 v1.1.0 [23] and contrast transfer function (CTF) estimation was done with GCTF v1.06 [24]. Upon manual inspection of the results, all micrographs were used to pick particles.

Particles were picked using the "blob picker" function of cryoSPARC by estimating the size of the complex (40 x 150 Å), yielding a total of 3912328 particles, extracted in a 352 pixel box. After several round of 2D classification, followed by manual curation of the obtained classes, 443886 particles were left. Three ab-initio classes were generated and separately refined. Upon manual inspection of the resulting maps, the best subset (118874 particles) was selected and used for non-uniform refinement.

For the template picking, the crystal structure of the NDM-1-Fab 17 complex (where a full-length NDM-1 was modelled onto the degradation fragment) was used as a template. The same strategy (starting with 1519419 particles) as the blob picking was used. For the Topaz picking, particles from the final subset of blob picking were used to train the model, followed by particle extraction,

yielding a total of 544698 particles and the same strategy was followed. Both the template and the topaz picking datasets gave worse resolution than the blob picking and were not carried forward.

The final particle dataset from the blob picking was exported to RELION 3.0, along with the automatically generated mask (corrected in Chimera to exclude a region of noise). A first round of 3D classification was performed and manually inspected. As all generated 3D classes did not show any significant difference, all particles were used for a final gold-standard 3D refinement. A final step of post-processing was then performed to sharpen the map .

2.3 Hepatitis B virus

2.3.1 Bioinformatics

Sequences of the HBV genomes used to build the phylogenetic tree in **Figure 7** were downloaded from NCBI (see **9.1** for a full list of the sequence identifiers). The multiple sequence alignment was built using Clustal Omega [25] and the output from the program (in PHYLIP format [26]) was used to display the tree using the program SplitsTree[27].

Disorder predictions were carried out using the DisMeta prediction server [28] using sequences NC_003977.2 (Human HBV, subtype ayw) and AY392760.1 (Duck HBV) after translation using Expasy translate (https://web.expasy.org/translate/).

2.3.2 Cloning

Sequences NC_003977.2 and AY392760.1 from **2.3.1** were used to obtain synthetic genes from GeneArt (Thermo) and cloning was performed according to **2.1.3**. Baculoviruses harbouring the gene for a single protein were generated with

flashBAC (Oxford Expression Systems) and for multiple proteins with biGBac [29].

2.3.3 Expression and purification tests

Protocols for expression tests are outlined in sections **2.1.13**, **2.1.7** and **2.1.10**. Purification tests were carried out by lysing the cells in (20mM Tris pH 8, NaCl 500 mM), supplemented with cOmplete EDTA-free protease inhibitor (Roche). Lysis of the cells was performed via sonication or with the addition of 0.5% Triton X-100. To aid with nucleic acids degradation, Turbonuclease (Sigma) was added to the lysate. Following lysis, the solution was centrifuged (30minutes, 20000xg, 4C) then the supernatant was either loaded onto a HisTrap FF column (GE Healthcare) or incubated on a rotating plate with the appropriate resin (Ni-NTA Agarose – Qiagen, Excel – GE Healthcare, TALON – Sigma) for 3 hours to overnight.

After extensive washes with (20mM Tris pH 8, NaCl 500 mM, 5-20mM imidazole, depending on the type of resin), the protein was eluted with (Tris pH 8, NaCl 500 mM, 500mM imidazole). NaCl content was lowered to 50mM NaCl and imidazole was removed with a HiTrap Desalting column (GE Healthcare) or by dilution with (20mM Tris pH 8). The protein was then loaded onto a HiTrap Heparin HP column (GE Healthcare) and eluted with increasing amounts of NaCl (up to 1M). Size exclusion chromatography was performed using a Superdex 200 increase 10/300 GL or with a superose 6 10/300 GL using (20mM Tris pH 8, NaCl 200 mM).

For Twin-strep purification, the clarified lysate from SF9 cells was loaded onto a Strep-Tactin XT column (IBA). After extensive washes with the loading buffer, the protein was eluted with the buffer supplemented with 50mM biotin.

Samples for mass spectrometry were loaded onto a SDS-PAGE gel and bands of interest were cut from the gel with a scalpel. Mass-spectrometry analysis was carried out at York University using MALDI-MS/MS.

2.4 NDM-1

2.4.1 Expression and purification of NDM-1

E.coli BL21*(DE3) cells (NEB) were transformed with the relevant plasmid and grown at 37C in a shaker incubator at 180rpm, until an OD_{600nm} of 0.6-0.8 was reached. The culture was then cooled to a temperature of 20C and IPTG was added to a final concentration of 1mM. Following overnight incubation, cells were harvested by centrifugation (6000xg, 20minutes, 4C) and the supernatant discarded.

The cell pellet was then resuspended in the lysis buffer (Tris 20mM pH 7.5, NaCl 500mM) and a small amount of lysozyme (Sigma) was added to aid with the lysis. The resuspended cells were lysed by sonication and the lysate was cleared by centrifugation (18000xg, 45 minutes, 4C) prior to loading onto a Ni Sepharose HisTrap FF column (GE Healthcare) using an ÄKTA FPLC system (GE Healthcare). The column was washed extensively using wash buffer (Tris 20mM pH 7.5, NaCl 500mM, imidazole 20mM) and the protein was then eluted with elution buffer (Tris 20mM pH 7.5, NaCl 500mM, imidazole 500mM).

The purified protein was then concentrated and loaded onto a Superdex S75 column (GE Healthcare) pre-equilibrated in Tris 20mM pH 7.5, NaCl 150mM for

size exclusion chromatography. Fractions containing the purified protein were incubated, if necessary, with 3C protease overnight (20C) and re-purified by SEC with the same buffer.

2.4.2 Antibodies generation

Monoclonal antibodies against NDM-1 were generated by ENVIGO (https://www.envigo.com/) with protein supplied by our collaborators.

2.4.3 Preliminary assays

Preliminary assays were carried out by our collaborators

2.4.4 Generation of Fab fragments from monoclonal antibodies

The ascites fluid containing the monoclonal antibodies was centrifuged to removed cellular debris (10 minutes, 25C, 13000xg), then loaded onto a HiTrap Protein G HP column (GE Healthcare). After extensive washes with PBS, the protein was eluted with 0.1M glycine-HCl (pH 2.7) and immediately neutralised with an appropriate amount of 1M Tris-HCl (pH 9.0). The recovered antibody was then processed to Fab fragments using the Pierce Fab Preparation Kit (ThermoFisher) following manufacturer instructions. Briefly, the antibody was incubated with papain immobilised on agarose beads for 5 hours, then loaded onto a protein G column to capture uncleaved antibodies and Fc fragments. The column flow-through contained the Fab fragments.

2.4.5 Expression and purification of recombinant Fab fragments

Cloning of the Fab fragment was carried out as detailed in **2.1.3** from sequences obtained by our collaborators (see **9.3** for the full sequences). Large scale transfection and cell cultured was performed as described in **2.1.11**. The growth media was then recovered, centrifuged to remove cellular debris (20 minutes,

4000xg, 4C) and filtered through a 0.25µm filter before being loaded onto a QuixStand benchtop system with a 10kDa membrane to reduce the volume. The resulting media was then loaded onto a HisTrap FF column using an Akta FPLC system. After extensive washes with (20mM Tris pH 7.5, NaCl 200mM, imidazole 20mM), the Fabs were eluted using (20mM Tris pH 7.5, NaCl 200mM, imidazole 500mM). The recovered protein was then loaded onto a Superdex S75 column for size exclusion chromatography using (10mM Tris pH 7.5, NaCl 50mM) as buffer. Purity was then assessed by SDS-PAGE.

2.4.6 NDM-1-Fab complex formation

NDM-1 and Fab fragments 5 and 17 were mixed with an excess of NDM-1 and incubated on ice for 30 minutes, then loaded onto a Superdex 75 (Ge Healthcare) for size exclusion chromatography. The buffers used for SEC were Tris 20mM pH 7.5, NaCl 150mM (crystallography) or Tris 10mM pH 7.5, NaCl 50mM (Cryo-EM).

2.4.7 Crystallization, data collection and refinement

Crystallization screening was performed as outlined in section **2.2.1** with trays set up at 20C and 4C, using the following commercial screens: Index, SaltRx, PEGRx, Morpheus, Morpheus II, PACT, ProPlex, PentaErythrol, NucPro. Furthermore, five in-house screens (I to V) were set up.

Optimisation steps for the NDM-1-Fab 5 complex were carried out by:

- Performing an additive screen (Hampton) on conditions yielding needle crystals.
- Re-screening using seeds prepared from needle crystals using a Seed Bead kit (Hampton)

- Testing *in-situ* proteolysis with trypsin (1:100 w/w) [30]

Diffraction experiments were performed at Diamond Light Source (I03 and I04 beamlines - 0.9795 Å, 100 K) and processed with the Xia2 [31] or autoPROC softwares [32]. Quality of the data was assessed by phenix.Xtriage [33] and Molecular replacement was carried out using Phaser[34], followed by multiple rounds of manual building using COOT [35] and refinement with PHENIX and REFMAC [36]. The resulting structures were then validated using Molprobity [37].

For both Fab 5 and the NDM-1-Fab 17 complex, the search model was a Fab derived from a mouse monoclonal antibody (PDB id: 2VXT), separated into a constant and a variable region and where CDR loops were deleted to eliminate sources of potential bias. For the complex, an additional fragment of NDM-1 (PDB id: 3ZR9), residues 50-140 was added.

For Fab 5, parameters that were refined were: cartesian, individual B-factor, TLS and torsion-angle NCS. For Fab 17 parameters that were refined were: cartesian, individual B-factor and TLS.

2.4.8 NDM-1-palmitic acid complex formation

Biolayer Interferometry experiments were conducted using an Octet RED96 (FortéBio) and standard streptavidin biosensors (Product code 18-5019). The sensors were first hydrated in Tris 20mM pH 7.4 NaCl 200mM for 10 minutes, then loaded with 0.05mg/mL of a chemically synthesized, palmitoylated and biotinylated peptide containing part of the protein sequence of Wnt7a (CH3CO-KC(S-)HGMS(-C=0-$(CH_2)_7$CH=CH$(CH_2)_5CH_3$)GSC(S-)TVLK-(EDA)-Bio) - see also methods in [38]. Excess peptide was removed by washing with the same

buffer. The loaded biosensors were made to interact with analytes (NDM-1 at 12.5 µM and Dlp at 9 µM). Analyte excess was removed by washing with the same buffer. Association and dissociation times were 180 and 300 seconds, respectively. The protocol followed is based on a published study involving Wnt7a [38].

NDM-1 was incubated with palmitic acid-BODIPY (Sigma cat no. D3821) in Tris 10mM NaCl 50mM 7.5% DMSO for 1 hour at 25C at a 1:20 molar ratio. The protein was then briefly centrifuged (2 minutes, 16000xg, 25C) and loaded onto a Superdex 200 Increase 3.2/200 using a FSEC Prominence HPLC system (Shimadzu). The excitation wavelength was 505nm and the emission 512nm.

Alternatively, NDM-1 was incubated with palmitic acid (Sigma) at the same conditions as above and loaded onto a Superdex 200 Increase 10/300 using an Akta FPLC system (GE Healthcare).

2.4.9 NDM-1 inhibition assays

Data were collected using a PHERAstar FS microplate reader (BMG Labtech), all experiments were conducted in 50mM HEPES pH 7.2, 10µM $ZnSO_4$, 1µg/mL BSA, 0.01% (v/v) Triton X-100. Meropenem (Sigma) was used to a final concentration of 100µM and absorbance was detected at a wavelength of 305nm. For Nitrocefin (Sigma), the final concentration was 25µM and absorbance was detected at a wavelength of 492nm. The appropriate amounts of NDM-1, Fab fragments were incubated in a 96-well plate in buffer for 10 minutes, prior to addition of the substrate, then immediately brought to the microplate reader to monitor the reaction. Analysis of the data was carried out with Microsoft Excel and GraphPad Prism. Relative activity of the enzyme was calculated by averaging

the values of three independent replicates at 5 minutes and 15 minutes. Relative activity was calculated with the formula:

$$RA: \frac{A_{+Fab\ (15)} - A_{+Fab\ (5)}}{A_{-Fab\ (15)} - A_{-Fab\ (5)}}$$

2.5 SfEiger

2.5.1 Crystallization, data collection and refinement

Crystallization screening was performed as outlined in section **2.2.1**. Diffraction experiments were performed at Diamond light source (I04 beamline - 0.9795 Å, 100 K) and processed with Xia2[31]. The model used for molecular replacement using Phaser[34] was identified by the Wide Search Molecular Replacement server[8] (PDB ID: 1U5Y). The tracing of the main chain was performed using SHELXE[39]. This was followed by multiple rounds of building side chains using COOT[35] and refining with PHENIX[33] and REFMAC[36]. The resulting structure was then validated using Molprobity[37].

2.5.2 Bioinformatic identification of SfEiger

The complete genome of *Spodoptera frugiperda*[40] was downloaded from NCBI (GenBank ID:NJHR00000000.1) was downloaded from NCBI. The genome was first converted using the makeblastdb command of BLAST[41]. The partial sequence of SfEiger built in the density was used as query for tblastn to identify regions in the genome containing the sequence. BLAST searches were performed using Genome Workbench (https://www.ncbi.nlm.nih.gov/tools/gbench/). The command tblastn[42] compares the query sequence with the translation in all six frames of the database. The results of the search were then manually assembled defining putative exon-intron junctions, supporting the work

with a multiple sequence alignment of different Eiger proteins available on the NCBI database of non-redundant protein sequences. The multiple sequence alignment was generated using Clustal omega[25] and plotted using Jalview[43].

2.5.3 Structural analysis

The Structural Homology Program (SHP)[44] was used to calculate the phylogenetic distances between a set of structures of TNF ligands that were pre-aligned using Pymol. The PHYLIP program [26] was then used to generate the corresponding phylogenetic tree. To visualize the tree, Splitstree[27] was used. The PISA server[45] was used to analyse protein interfaces.

2.5.4 Molecular dynamic simulations

Molecular dynamic simulations were carried out by Dr. Max Renner.

3 Attempts at expression and purification of the hepatitis B virus polymerase

3.1 Introduction

3.1.1 A brief overview of the hepatitis B virus

According to the World Health Organization (WHO), it is estimated that 257 million people are infected with the hepatitis B virus (HBV), with 887000 deaths in 2015, mainly due to complications from the infection [46](**Figure 5**). As with other types of viral hepatitis (A,C,D and E), the HBV has a marked tropism for hepatocytes. Unlike other hepatitis viruses, the largest burden of hepatitis B is that in some cases (5-10% in adults and 90% in infants) the infection becomes chronic and can further develop into cirrhosis and hepatocellular carcinoma (HCC) which is one of the leading causes of cancer-related deaths [47].

The HBV can be transmitted horizontally via bodily fluids such as blood, saliva, semen and vaginal fluids [48]. Transmission of HBV via corneal transplantation has also been reported [49]. Vertical transmission of the virus is also possible perinatally from mother to child, further adding to the burden of the virus on healthcare [50].

In HIV-positive patients, chronic hepatitis B is comorbid in about 10% of the cases and, as treatment with antiretroviral drugs has increased their lifespan, the need for combined therapies can increase hepatotoxicity and complicate the management of both pathologies [51]. Unlike HIV-1, where thanks to public health efforts to contain the number of new infections is decreasing, the number of HBV infections is increasing [52].

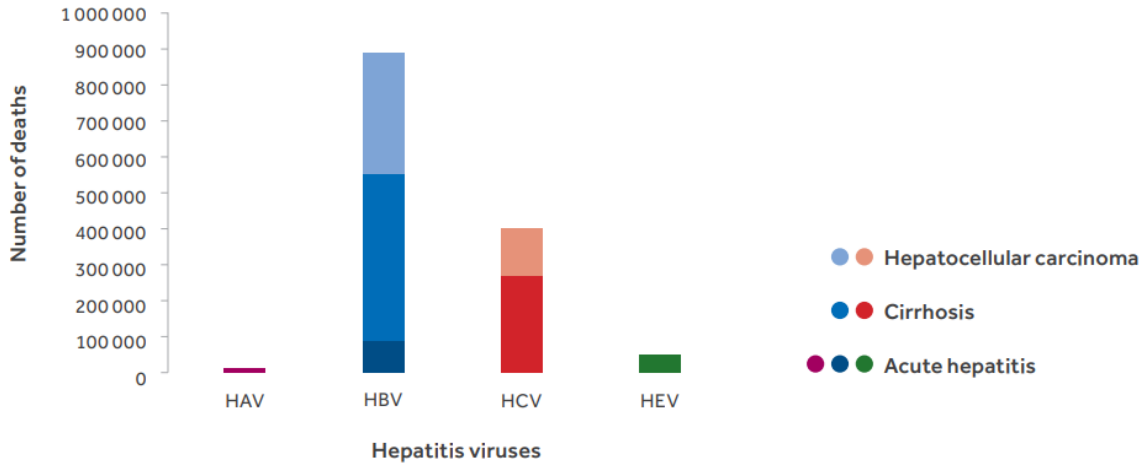


Figure 5: Deaths by viral hepatitis in 2015. HAV: hepatitis A virus, HBV: hepatitis B virus, HCV: hepatitis C virus, HEV: hepatitis E virus. From [52].

Ten different genotypes of the HBV (plus sub-genotypes) have been identified, called "A" to "J". Their geographical distribution is markedly different, with some countries reporting almost exclusively a single genotype[53](**Figure 6**).

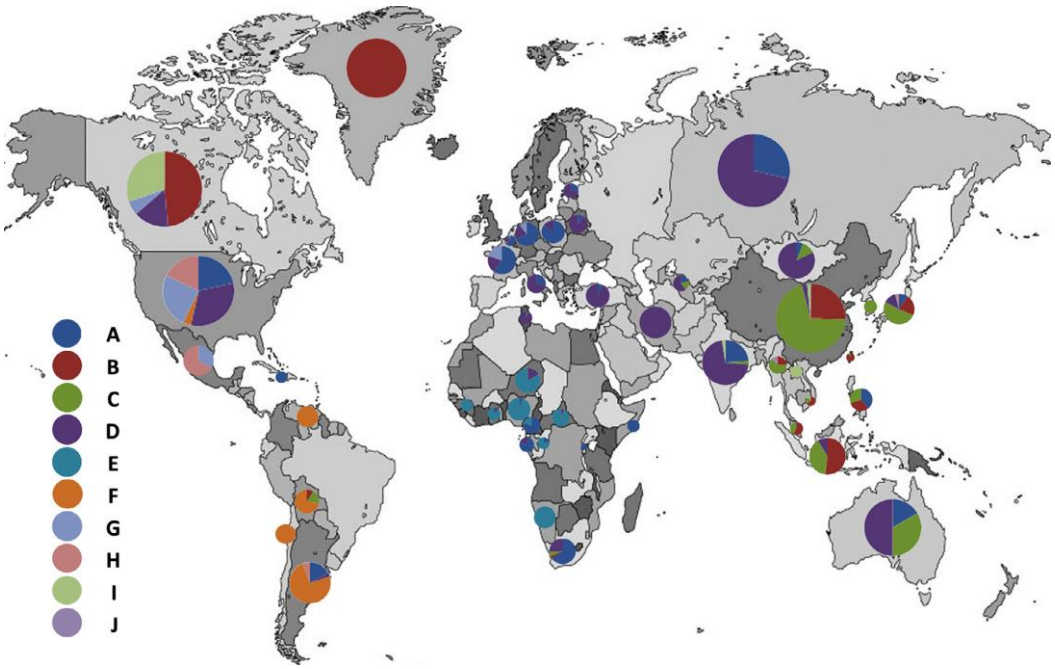


Figure 6: Geographical distribution of different hepatitis B virus genotypes. The pie charts overlaying different nations show the distribution of the different genotypes in each nation. From [53].

The virus genotypes correlates with the clinical prognosis of the HBV infection, where genotypes C, D and F are more likely to result in cirrhosis and hepatocellular carcinoma[54].

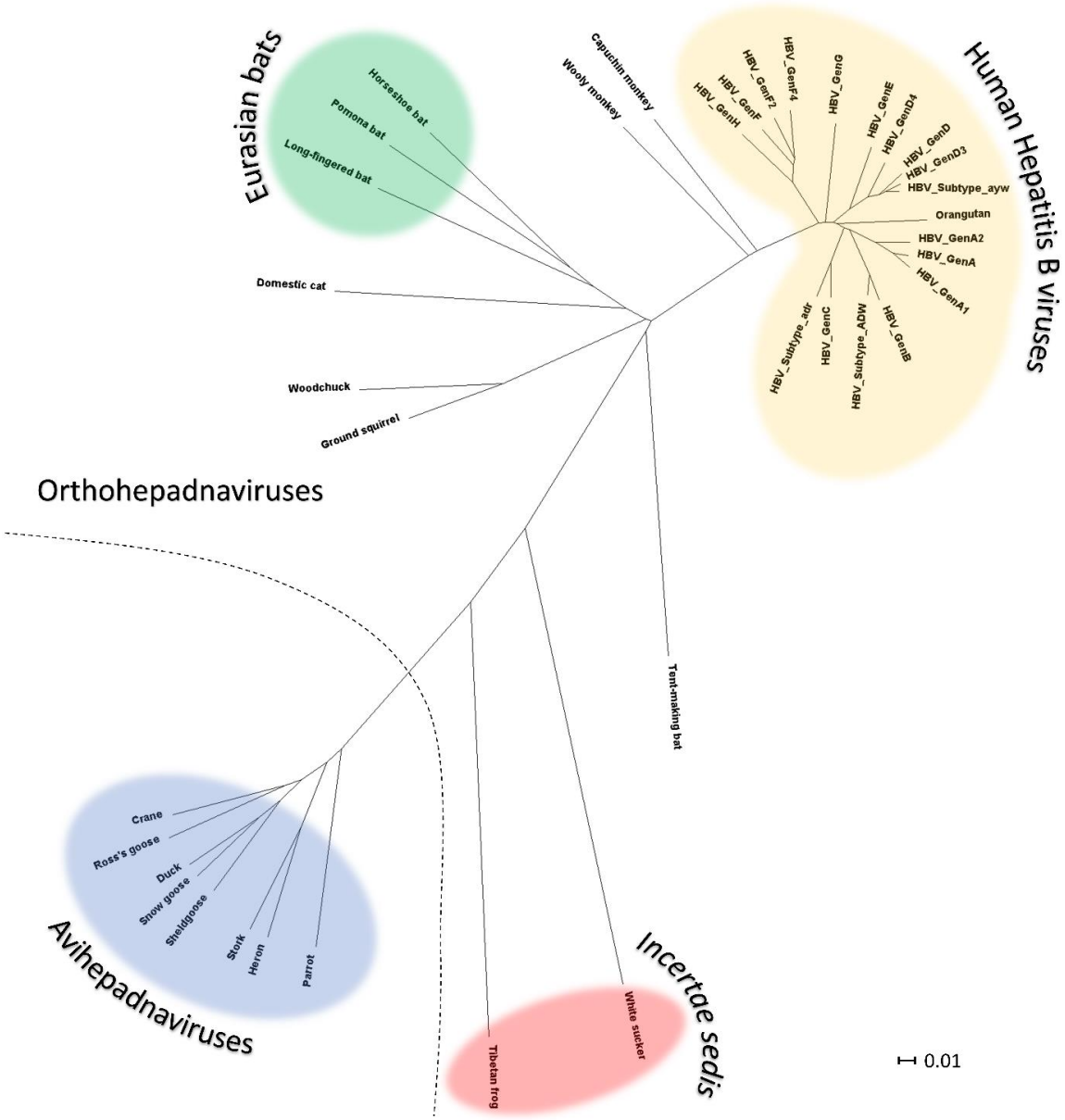


Figure 7*:* Phylogenetic tree of the genomes of the *Hepadnaviridae* family. Human HBVs are highlighted in a yellow bubble, specifying genotypes (as "HBV_Gen_X") and subtypes (as HBV_subtype_xxx). The dashed line separates the *genera Orthohepdnavirus* from the *Avihepadnavirus* (also highlighted by a blue bubble). Other highlighted groups are Eurasian bats (green) and viruses yet without a *genus* (red, *Incertae sedis*). Note that the Tent-making bat is geographically separated from the Eurasian bats, as it is only found in South America. The full list of identifiers of genomes used to generate this figure is given in **9.1**.

The *Hepadnaviridae* family of viruses, is divided into two *genera*: the *Orthohepadnavirus* and the *Avihepadnavirus*. The first comprises mammalian

viruses (including the human HBV), while the latter birds[55]. Notable exceptions to this classifications are the HBVs infecting the white sucker (a fish) and the Tibetan frog (amphibian). In these cases, it is not clear what the *genus* is and the viruses have been designed as *Incertae Sedis* (***Figure 7***).

As no currently available treatment are capable of eradicating the virus once the infection is chronic, the best weapon currently available against the hepatitis B virus is a vaccine containing a recombinant, yeast-expressed, surface antigen [56]. However, there have been reported cases where mutated viruses are able to infect patients previously vaccinated against the virus. Gene mutations in the surface antigen are correlated with these vaccine-escape mutants [57].

It is worth mentioning that hepatitis caused by the hepatitis Delta virus (HDV) is dependent on a pre-existing HBV infection, as HDV is a defective pathogen whose genome does not encode surface proteins. Instead, it relies on HBV surface protein synthesized in HBV-infected hepatocytes. The two viruses, however, belong to different families and have a completely different genome organisation (reviewed in [58]).

Following acute infection from HBV, many patients spontaneously clear the virus from the body, but in some cases the infection becomes chronic and requires lifetime care. While the discriminant between viral clearance and chronic HBV infection is not currently known, there is an inverse relationship with age. Approximately 90% of infections within the first year of age become chronic, while this number falls to about 5% in adults [59]. Chronic HBV infection is managed with nucleoside/nucleotide analogues, such as Tenofovir and Entecavir [60], but none of the available treatments is capable of fully clearing the body

from the virus and over time, prolonged drug treatment can lead to drug resistance [61].

3.1.2 Structure of the hepatitis B virus virion

Virions of the Human HBV are enveloped with an internal capsid with icosahedral symmetry and measure ~40 nm, while the Duck HBV virions range from ~35 to 60 nm [62]. *E.coli* expressed core antigen (HBcAg) self-assembles into icosahedral particles with a triangulation number T=4. Here the icosahedral asymmetric unit is composed of four copies of the HBcAg which, upon application of the symmetry operators, generate the fully closed capsid (240 copies total) [63]. About 10% of the capsids, however, have a T number of 3, with 180 copies in total, and this is also seen in samples collected from infected cells and imaged by Cryo-EM [64]. Icosahedral particles, however, are not necessarily predominant in the context of a viral infection: HBV infected cells also synthesize and secrete via multivesicular bodies smaller, spherical 22nm particles containing only small and medium HBsAg, as well as rod-shaped, filamentous particles with the same diameter. Both types of particles, however, are defective and non-infectious. Secretion of these particles vastly outnumbers that of infectious particles, with ratios between 1000 and 100000 to 1 [65]. This has been shown to enhance the infectivity of the virus, perhaps by acting as a decoy and sequestering host-derived antibodies and preventing their binding to mature particles [66].

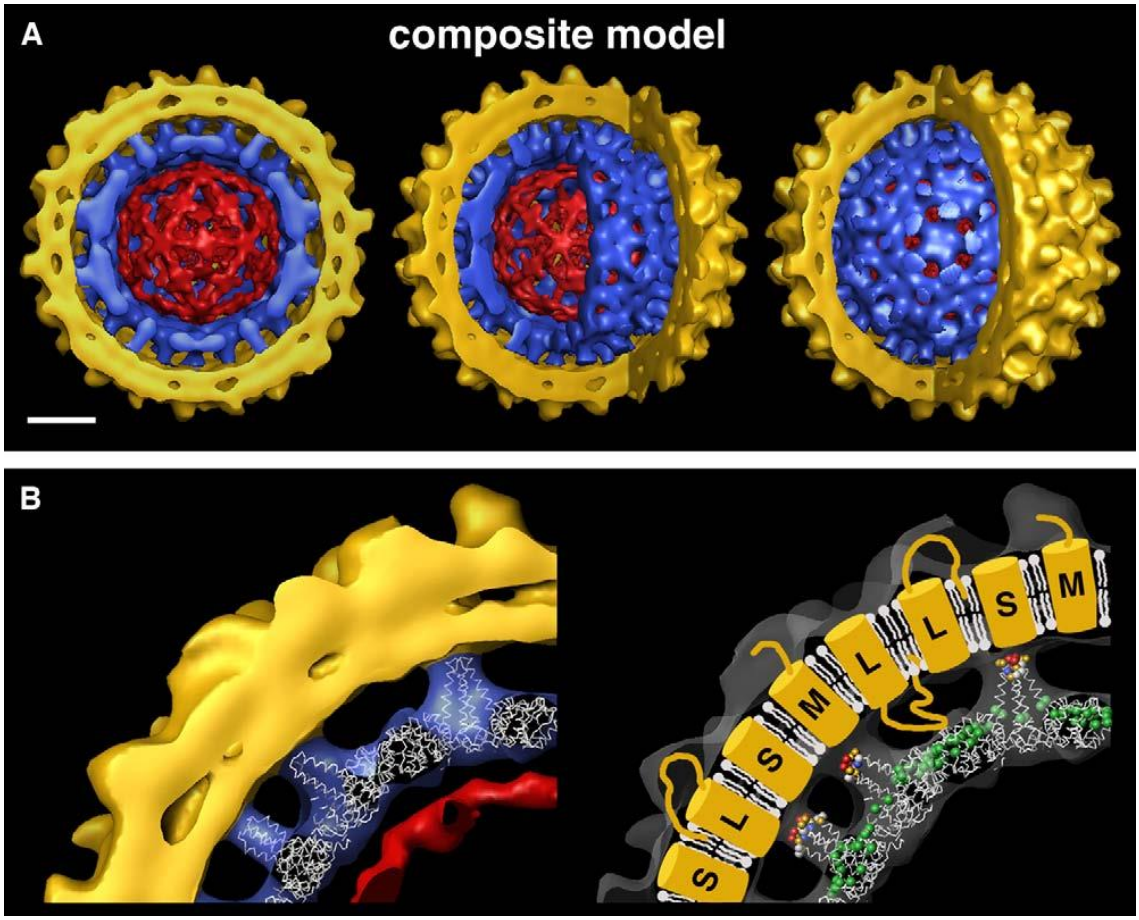


Figure 8: (**A**) Views of native, T=4 virions of HBV. In yellow, density pertaining to the envelope and HBsAg, in blue that of the capsid (with HBcAg) and in red lower-resolution density for the virus genome. (**B**) Closeup and modelling of the three layers. A crystal structure of the capsid is modelled in the EM density (PDB id:1QGT). Scale bar: 100Å. S, M, L indicate Small, Medium and Large surface antigens. From [67].

Fully mature viruses imaged with Cryo-EM show three layers of density (***Figure 8***), starting from the outer, hepatocytes-derived lipid bilayer where multiple copies of S,M and L-HBsAg are inserted. The L-HBsAg interact, via their inner loops with the capsid proteins, forming a middle layer. The innermost layer harbours the viral genome [67].

3.1.3 The hepatitis B virus genome and proteins

HBV is a small DNA viruses with a genome of ~3.2kb. The HBV genome is a partially double stranded DNA where the (-) strand is the complete coding strand and the (+) strand is non-coding and incomplete. While the 5' end of the (+) strand

is fixed, anchored by two direct repeats (DRI and DRII), its 3' end is random and depends on the replication process (see **3.1.4**) [68][69]. From this genome, 4 different transcripts are synthesized: the pre-genomic RNA (pgRNA) is the longest and contains the whole sequence of the virus. The pgRNA, at 3.5kb is longer than the genome of the virus, this happens because it contains a duplicated sequence (see **3.1.4**). Other transcripts are shorter: the 2.4kb pre-S mRNA, the 2.1kb S mRNA and a smaller 0.7kb mRNA. The last of these transcripts, however, is not consistently found in infected cells and is only present in orthohepadnaviruses [70].

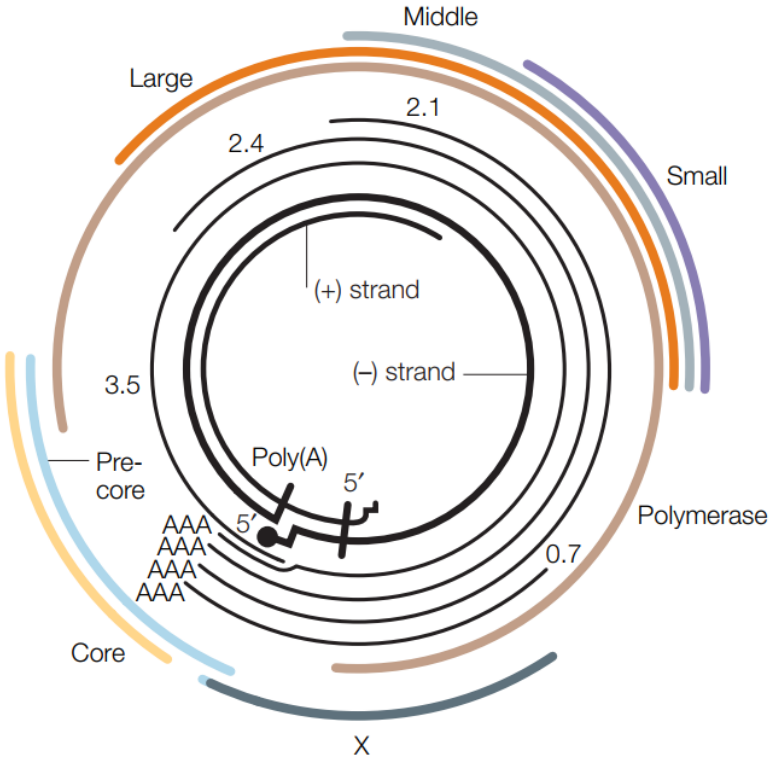


Figure 9: The hepatitis B virus genome. From [71]. Thick black lines at the center of the circle represent the two strands of the genome. Going towards the outside, thin black lines indicate the four mRNA transcripts and their length in kb. Coloured lines represent the region of the transcript giving rise to the complete protein.

The HBV genome is covalently linked to the polymerase (see **3.1.5**): in the context of a viral infection, the bond between the two is cleaved. Host cellular enzymes then complement the (+) strand forming the covalently closed circular

DNA genome (cccDNA), which binds histone proteins, ultimately forming a minichromosome (see **3.1.4**).

From these 4 transcripts, four different open reading frames are identifiable (**Figure 9**). The pre-Core protein (or HBeAg, e antigen), the Core protein and the Polymerase are transcribed from the pgRNA; the Large envelope protein (also called L-HBsAg) from the pre-S mRNA. The medium and small envelope protein (M/S-HBsAg) are transcribed from the S mRNA while the X-protein from the 0.7kDa transcript.

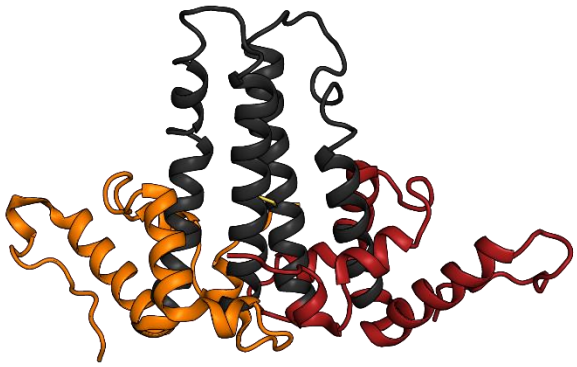


Figure 10: hepatitis B Virus Core protein dimer. PDB id : 4BMG. In black the dimerization interface, in red and orange the domains contacting the genome.

The Core protein (HBcAg) is a 21kDa phosphoprotein, that can be divided in two regions: the N-terminal domain, involved in the formation of the viral capsid and the basic C-terminal domain that can bind nucleic acids (**Figure 10**) [72]. While its most studied function is in the formation of the viral capsid, the core protein is also involved in encapsidation and reverse transcription of the genome as well as immune response regulation[73]. In a clinical setting, antibodies directed against HBcAg can be used to detect previous exposure to the virus [74].

The E antigen (HBeAg) is synthesized from the longer transcript of the Core protein and results in a small protein of 15kDa, due to RNA processing. The HBeAg is not involved in the replication process, but is present in the virion and also secreted from the cell where it has immune modulation properties [75]. The HBeAg is also used as a biomarker used to monitor the course of an HBV infection [74].

The three envelope proteins (HBsAg) are respectively 25, 31 and 42kDa with the smallest being the most abundant. Envelope proteins are post-translationally modified by the addition of sugars and lipids [76]. The pre-S1 region of the L-HBsAg binds to the cellular receptor to allow virus entry.[77]

The X protein (HBx) is a multifunctional, cytoplasmic, 16kDa protein that has also been found in the nucleus where it acts as an activator of transcription factors [78]. It has been shown to be used by the HBV to alter a variety of cellular processes and is required *in vivo* for an efficient replication [79]. HBx has been shown to have oncogenic potential, especially in the case of C-terminal truncated versions [80], [81]. HBx is also responsible for what is known in literature as "hepatitis B virus X Protein-Induced Autophagy". This is believed to happen through differential localization of HBx during the replication cycle of the virus[82]. HBx is directly involved in the permeabilization of mitochondrial membranes [83], leading to apoptosis through the release of apoptotic factors, ROS production [84], or by interaction with death-associated kinases [85].

For a detailed description of the polymerase, see **3.1.5**.

3.1.4 Life cycle and replication mechanism of the hepatitis B virus

A schematic representation of the life cycle of the hepatitis B virus is shown in **Figure 11**. The cellular receptor on hepatocytes that allows entry of the HBV (and in turn for HDV) into the cell has been recently determined to be the Sodium Taurocholate Co-transporting Polypeptide (NTCP) [86], but the entry also appears to be dependent on heparan sulphate proteoglycans [87]. At least in the case of avian HBV, it appears that the Carboxypeptidase D (initially erroneously identified as a receptor) is also involved [88][89].

Once the HBV binds the receptor, internalization of the particle happens via clathrin-mediated endocytosis and translocation into endosomes. As the pH in the endosome gradually decreases, the fusion of surface proteins with the endosomal membrane is triggered [90][91]. There are, however, reports that for the Duck HBV a pH-independent mechanism may be predominant [92].

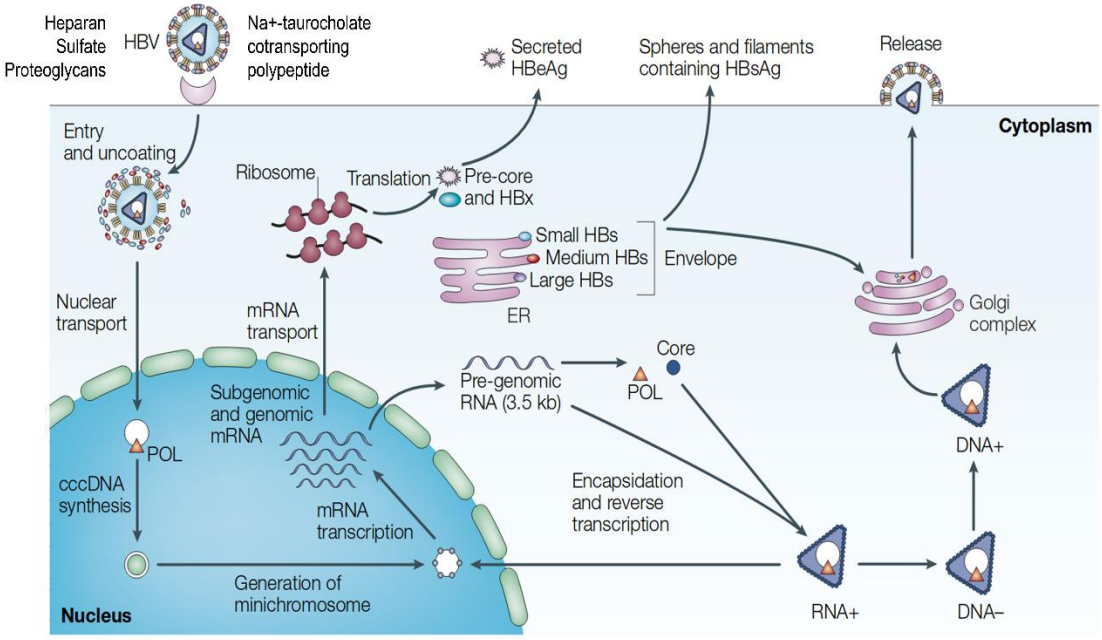


Figure 11: Overview of the hepatitis B virus life cycle. POL: polymerase, cccDNA: covalently closed circular DNA, HBsAg: hepatitis B surface antigen, HBx: hepatitis B x protein, HBeAg: hepatitis B e-antigen. Adapted from [71].

At present it is unclear how the genome is transported into the nucleus, but it is thought to involve the nuclear pore complex. A nuclear localisation signal (NLS) has been identified on the core protein, which would support a model where the capsid contacts the nuclear pore complex and the genome is released [93]. An alternative model involves the polymerase, covalently bound to the genome, contacting nuclear import proteins. This would happen upon phosphorylation of an inducible NLS sequence at the N-terminal of the polymerase [94].

In the nucleus, the polymerase is first cleaved from the genome, seemingly by the host tyrosyl-DNA-phosphodiesterase 1 or 2 [95]. The host polymerases κ, λ and

H, as well as DNA ligases, participate in complementing the (+) strand of the viral genome, giving rise to the Covalently Closed Circular DNA (cccDNA) [96][97]. The cccDNA is then bound by histones, giving rise to a stable minichromosome [98]. The minichromosome acts as the template for the transcription and the synthesis of viral RNAs that will be later used for the translation of the viral protein. The pgRNA will also be encapsidated to form new virions.

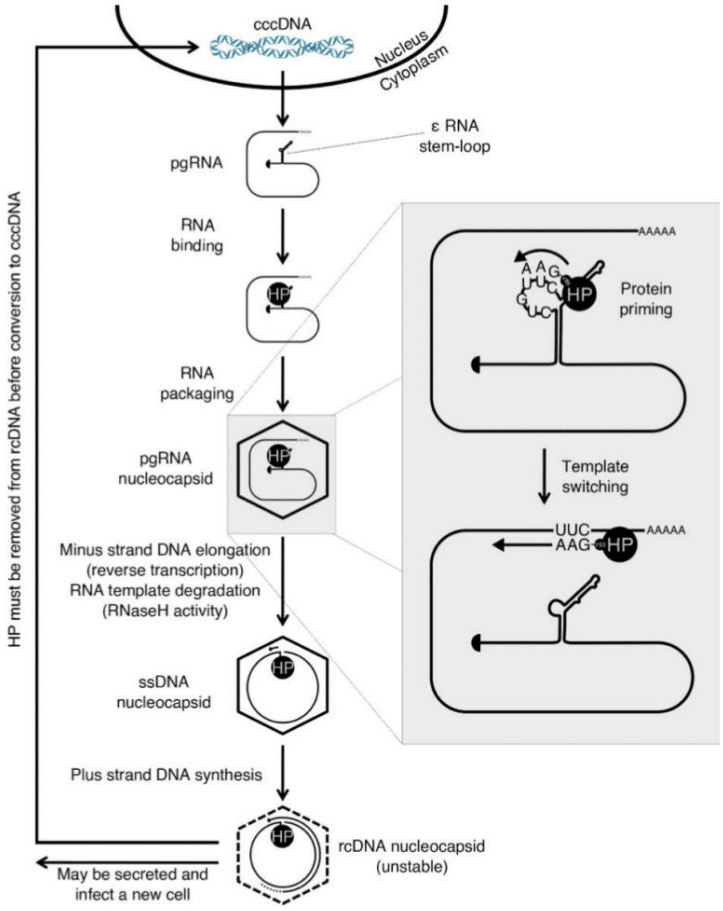


Figure 12: Schematic representation of the hepatitis B virus encapsidation and replication process. pgRNA: pre-genomic RNA, HP; hepatitis B virus polymerase. From [99].

The pgRNA contains two copies of a specific sequence, called ε, with a stem-loop structure[100]. After the pgRNA is exported from the nucleus (**Figure 12**), it is bound by the polymerase, which recognises and binds the 5' ε. In order to be able to bind ε the polymerase requires other host factors, Hsp90 in particular. Hsp

proteins are thought to stabilise the polymerase in an energetically unfavourable conformation, ready to recognise the ε stem-loop [101].

Following this event, encapsidation of the viral genome is initiated [102] and, within the capsid, retrotranscription of the pgRNA. As the capsid matures, some viral particles are transported back to the nucleus to maintain a pool of cccDNA and thus a perduring infection [103]. Interaction between the other capsids and envelope proteins (post-translationally modified in the Golgi) takes place in the cell multivesicular bodies [104]. Secretion of the viral particles, as well as the defective filaments is dependent on the ESCRT machinery [65][105].

The replication process of the HBV, following synthesis of the pgRNA template, takes place entirely within the capsid of the forming virion. Here, the polymerase itself acts as a primer through tyrosine 63, which remains covalently bound to the (-) strand of the HBV genome [106][107][108]. Simultaneously with the retrotranscription, the RNAseH domain of the polymerase degrades the pgRNA template [109]. The polymerase "jumps" between sequences on the genome in a process called "template switching" (**Figure 13**). Upon binding of the ε stem-loop and priming, the 5' DR1 is retrotranscribed. At this point the first "switch" takes place and the newly formed DNA, together with the polymerase, pairs with the 3' DR1 region [110]. Retrotranscription of the (-) strand occurs, coupled with degradation of the RNA, until the polymerase reaches the 5' of the RNA. What remains of the pgRNA is used as primer for the synthesis of the (+) strand, following a second event of template switching [111]. The (+) strand is not fully synthesized and while the reason for this hasn't been conclusively proven, it is currently thought that this is due to steric hindrance as the space inside the capsid

is limited. Another possibility is that, as the capsid matures, the replication process itself depletes the nucleotides available for the synthesis of the (+) filament [112].

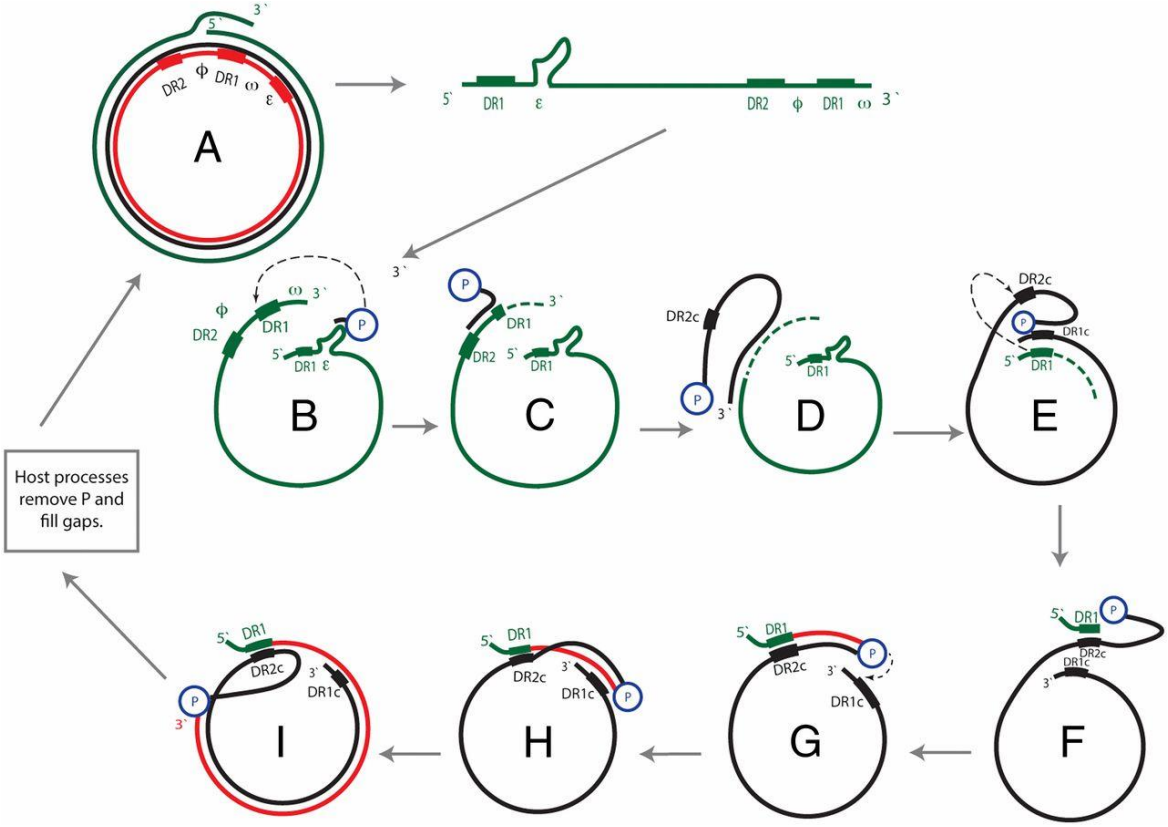


Figure 13: Template switching during the HBV replication process. (**A**)cccDNA. (**B**)The polymerase binds to ε and starts retrotranscribing. (**C**) The newly formed DNA binds to the 3'DR1 and (**D**) retrotranscription of the (-) strand occurs. (**E,F**) Second switching with a small piece of the template RNA acting as primer for the synthesis of the (+) strand. (**G,H,I**) Circularization and partial synthesis of the (+) strand. From [113].

3.1.5 The structure of a viral RNA polymerase

During its life cycle, a virus needs to replicate its genome, in order to make more copies of itself, as well as transcribing it to synthesize a broad range of proteins, required throughout. Different viruses use different strategies to do so but many viruses, such as those that replicate in the cytoplasm, carry a gene that encodes for a viral polymerase.

Like cellular polymerases, a viral polymerase is an enzyme that catalyses the reaction that adds a nucleotide to the 3' end of a nascent chain of RNA or DNA, based on complementarity with a pre-existing nucleic acid template. The catalytic mechanism requires a divalent ion cofactor (typically Mg^{2+} or Mn^{2+}) [114]. In order for the enzyme to add a nucleotides, it needs a free 3'-OH and different mechanisms are employed by different polymerases to provide this [115].

Viral RNA polymerases share a common right hand fold, with "palm", "fingers" and "thumb" domains (**Figure 14**).

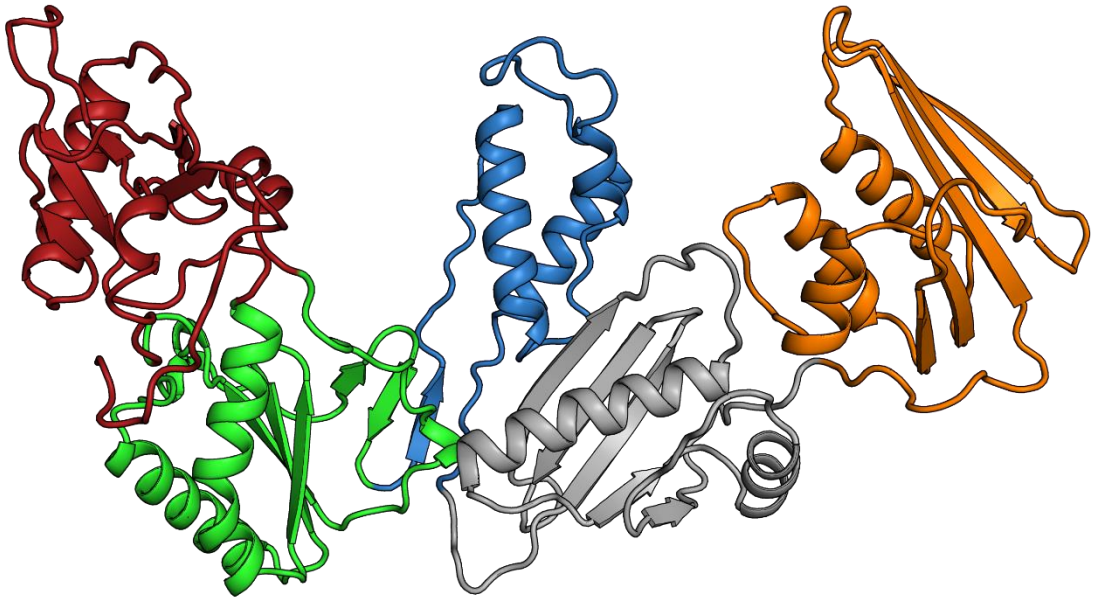


Figure 14: The structure of the HIV-1 RNA polymerase. (**PDB ID**: 1RTH). The palm domain is shown in green, the fingers in red and the thumb in blue. The RNAse H domain is shown in orange and is connected to the main polymerase by a linker domain (in grey).

Within these domains, several sequence motifs are conserved among all polymerases: A, B, C and D [116]. These motifs contain several key residues involved in the activity of the polymerase. In particular, two aspartate residues in motif C are conserved residues that coordinate the two metal ions in the active site involved in the catalytic mechanism.

3.1.6 The hepatitis B virus Polymerase

The hepatitis B virus polymerase (***Figure 15***) is the only enzyme encoded by the virus genome, and it is translated from the pgRNA into a ~90kDa enzyme. The polymerase can be divided into 4 domains: the Terminal Protein (TP) domain, a spacer, the retrotranscriptase (RT) domain and an RNAse H domain.

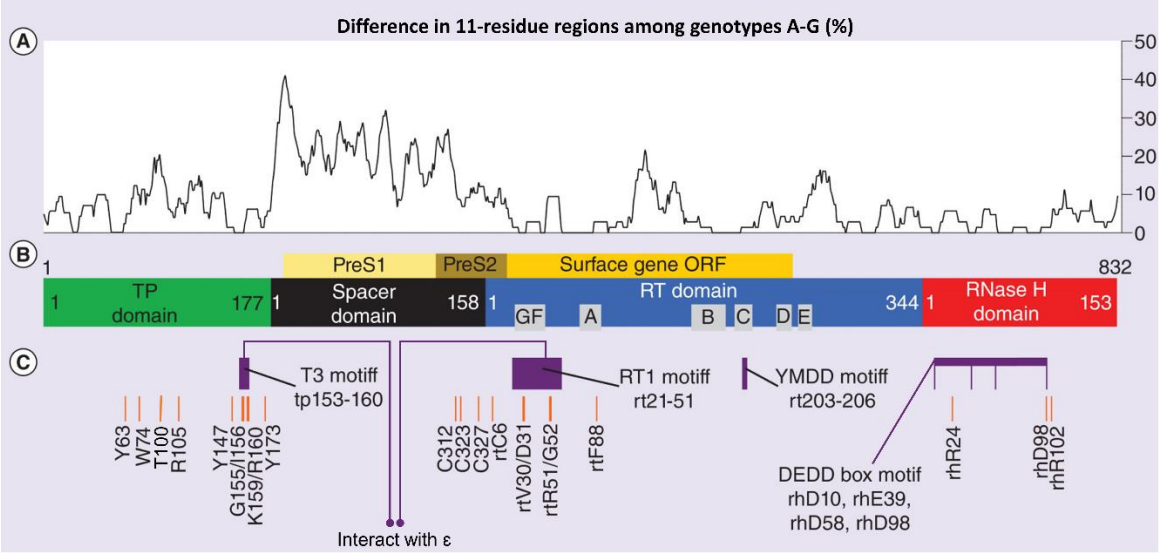


Figure 15: The hepatitis B virus polymerase. (**A**) Graphical representation of sequence variability between HBV genotypes. (**B**) Domain subdivision of the polymerase. In different shades of yellow are highlighted regions coding (upon frame-shift) for surface antigens. (**C**) Some residues and motifs that have been shown to be involved in the function of the enzyme. Adapted from [99].

While no structural data is available for the polymerase, multiple different residues have been identified being involved through mutational studies[117][118]. The TP domain participates in the binding of the genome by interacting with the ε stem-loop and act as a primase during replication, through the tyrosine 63 [119][120]. This process, however, requires also cellular enzymes [106]. While not the predominant way for DNA and proteins to interact, covalent linkage between them has been known for many years for cellular enzymes, as well as bacteria and viruses. This appears to be associated with novel structures

with a specific function [121]. This is the case for the TP domain as well, as no homologues have been identified for this domain.

The TP domain is followed by a spacer region of ~150 amino acids that tethers it to the rest of the polymerase. The spacer has the highest sequence divergence between the HBV genotypes [119]. It is worth noting that the spacer domain is generally considered non relevant for the catalytic activity of the enzyme [99].

The HIV-1 polymerase is the closest homolog to the HBV polymerase: taking together the RT and the RNAse H domains, the sequence identity with the HIV-1 polymerase is ~22%. Given the lack of a structure, several homology models have been generated over the years to investigate the polymerase as a druggable target [122][123][124]. The RT domain incorporates new nucleotides in the nascent strand. The three characteristic features of a polymerase (finger, palm and thumb) are present (***Figure 16***).



Figure 16: Model of the hepatitis B virus polymerase RT domain. A double stranded DNA model is shown in stick representation and a molecule of Lamivudine (nucleoside inhibitor) in space-filling representation. Adapted from [122].

Within the sequence of the RT domain, several conserved boxes ("A" to "G) have been identified. Box A contains a phenylalanine residue that can discriminate dNTPs from NTPs [125]. Box C contains the YMDD motif that canonically binds magnesium ions in the active site that is also conserved in the HIV-1 polymerase. The RT1 motif interacts with the T3 motif (on the TP domain) to induce the binding of ε [120]. Finally, the RNAse H domain sits at the C terminus of the polymerase and is capable of hydrolysing the template pgRNA molecule. It contains a DEDD box motif, involved in the binding of metallic ions [126]. While some compounds that are active against the HIV-1 RNAse H are already active against the HBV one [127], this domain is rapidly becoming a target for the development of new antivirals against the HBV [128].

3.1.7 Aim of the project

With almost 900,000 yearly deaths resulting from chronic HBV-derived hepatitis, the virus is a significant burden on society as a whole. Currently available antiviral therapies for chronic HBV infection have to be taken for a lifetime, as they are unable to completely clear the body of the virus. With the exception of interferon, all the drugs currently in use to treat HBV infection target the polymerase. Yet, the lack of any structural data on the polymerase itself makes the design of better compounds much harder. New approaches in development are instead based on RNA interference using short oligonucleotides (Phase II trials) [129], but CRISPR/Cas9 systems are being considered as well, targeting the cccDNA for degradation [130][131].

The interest in obtaining the elusive structure of the HBV polymerase also lies on the multiple challenging biological questions that it would answer: what is the

structure of the TP domain of the HBV polymerase? The domain has no homologues, thus modelling is unreliable. The structure would also shed light on the mechanistic details of the formation of the DNA-protein covalent bond.

Where do the existing drugs used to treat the HBV bind and how can we improve them? There is a wealth of biochemical data currently available, but no structure to relate it to, which would greatly benefit the rational design of new inhibitors.

Finally, how does the unusual replication mechanism of the HBV relate with the structural features of the polymerase?

For all these reasons, the interest and the challenges associated with this project are great and warranted pursuing this project. Unlike in the past, where significant amounts of pure protein were needed for crystallographic studies, small but meaningful amounts of polymerase could be enough to obtain a structure via single particle Cryo-EM analysis.

3.2 Results

3.2.1 Construct design and bioinformatics

As of now, the polymerase activity has been investigated only through functional assays in HEK cells [118][132][133] by achieving low level of expression of the whole polymerase or truncated constructs. Alternative approaches that can be found in literature are based on expression in insect cells [134][3][135]. More reductionist is the expression of single domains in *E.coli* and purification from inclusion bodies [2]. None of these approaches, however, are suited for structural biology and they have not resulted in any structures. A critical analysis of the literature shows that expression level of the HBV polymerase in the overwhelming majority of cases is only suitable for in vitro assays using radioactive nucleotides to detect small amounts of products. Other studies with higher level of expression, potentially suitable for structural studies, were never replicated. As hinted in some papers, refolding approaches seem unlikely to be an alternative, as the presence of 19 cysteines in the polymerase makes finding a reliable protocol unlikely. Furthermore, in the case of *E.coli* expression [2], the necessity for molar excesses of added amphipoles in order to keep a single domain soluble makes the strategy non viable at least for crystallization. Other approaches, based on mutagenesis of the HIV-1 polymerase in order to simulate the HBV behaviour [136], are clearly non applicable.

Bioinformatic analysis of the sequence of the HBV polymerase, using the disorder prediction server DisMeta [28], reveal agreement with already available information (**Figure 17**). With the exception of the linker region, the rest of the polymerase appears well-ordered. Similarly, the same predicted disorder profile

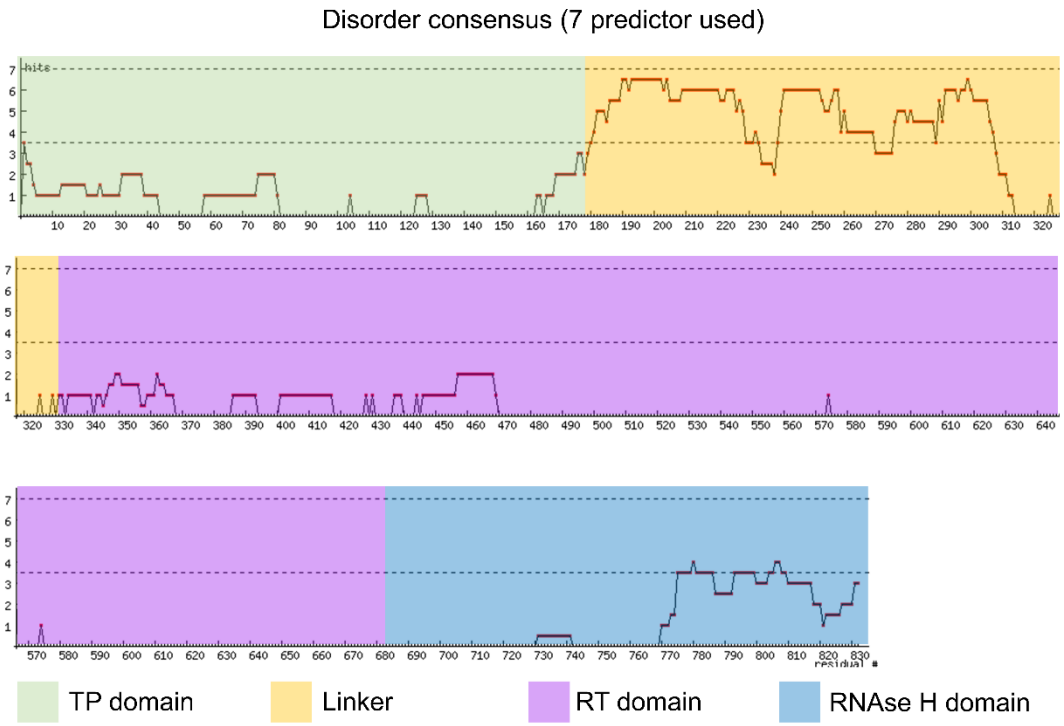


Figure 17: Summary of the disorder prediction of the Human HBV polymerase (NC_003977.2) obtained with the DisMeta server. On the X axis the residue number, on the Y axis the number of predictors that agree with a disordered sequence prediction.

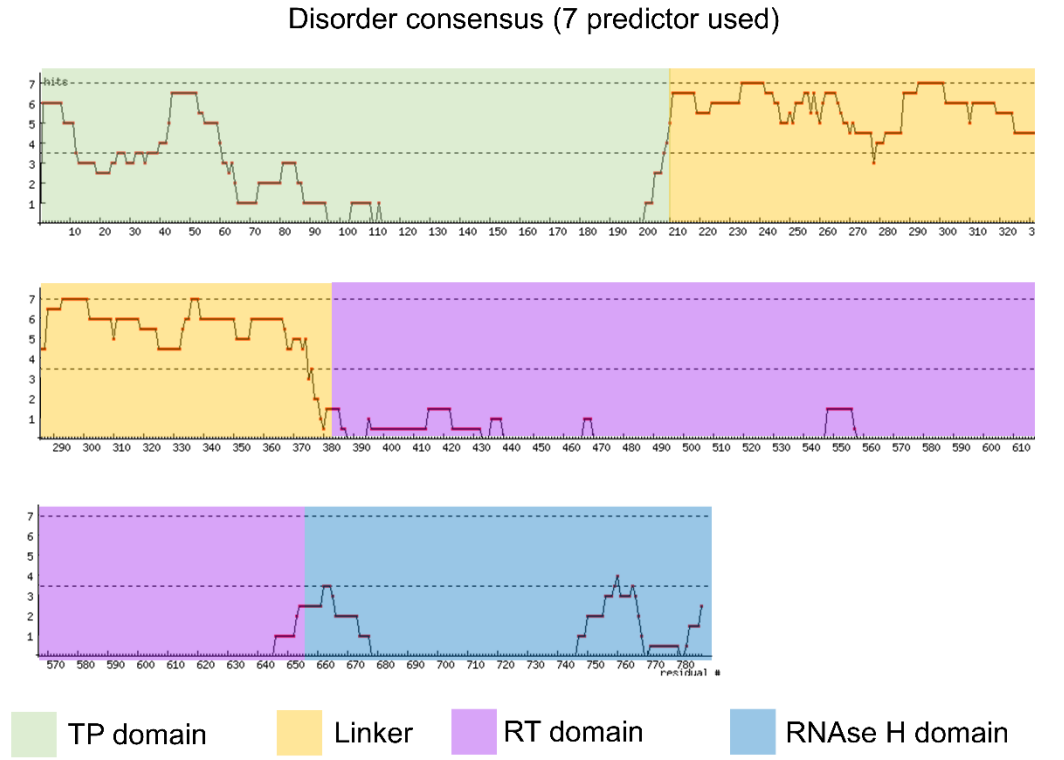


Figure 18: Summary of the disorder prediction of the Duck HBV polymerase (AY392760.1) obtained with the DisMeta server. On the X axis the residue number, on the Y axis the number of predictors that agree with a disordered sequence prediction.

can be seen for the Duck HBV polymerase (**Figure 18**). The HBV sequences used for expression were based on sequences used in other studies (e.g. [94][2]).

Constructs were designed with the intent of reproducing already published successfully expressing HBV polymerase (or single domains). In particular, by using vectors of the pOPIN system, expression tests in *E.coli*, SF9 and HEK cells could be performed without the need for more cloning. In parallel, expression of full-length HBV and Duck HBV polymerases was attempted using a lentiviral system[14]. Expression of single domains was tested using different purification tags: His, Maltose Binding Protein (MBP), Glutathione S-transferase(GST), SUMO for affinity purification and to enhance solubility [137][138]. Expression of a truncated versions of the polymerase, with the linker domain, as well as N- and C-terminal shortened (Mini-RT) [139], was also attempted. More novel constructs were based on findings in the existing literature: the presence of a NLS on the TP domain of the polymerase [94] was the base for a fusion between the TP domain and Importin α-5, one of the main parts of the cellular trafficking system that shuttles protein in the nucleus [140]. Based on the finding that Hsp proteins, in particular Hsp90, are required to keep the polymerase in the right conformation [101], co-expression of these element was also another strategy. The biGBac system [29], was used to co-express a variety of combinations between the HBV polymerase, Hsp proteins and the ε stem-loop. Co-expression of the HBV polymerase with the X-protein was also attempted, as it is required for efficient replication of the virus *in vivo* [79].

An overview of the expression and purification attempts is shown in **Figure 19**.

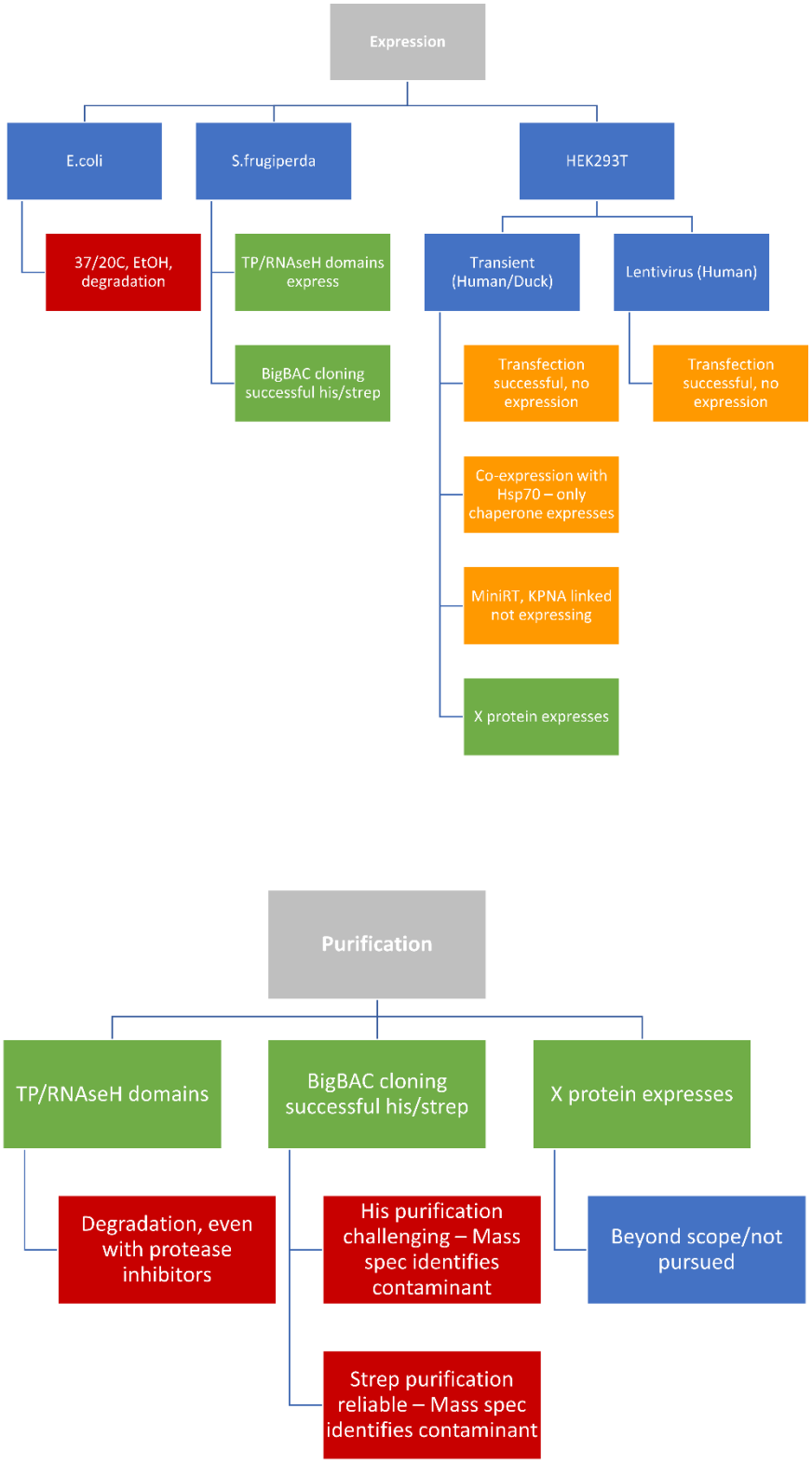


Figure 19: Overview of the different Expression (top) and purification (bottom) attempts. Colour coding of the blocks represent different outcomes. **Red**: not successful, **Orange**: partially successful, **Green**: successful. Blue and grey blocks are informational.

3.2.2 Cloning of the hepatitis B polymerase

Synthetic genes were amplified from their plasmids with a PCR. In general, most reactions were successful using pre-calculated parameters. In some cases, minor adjustments of the reaction (annealing temperature, extension time or addition of DMSO) were necessary to obtain a sufficient amount of the correct product. A summary of tested, successfully cloned constructs is shown in **Table 4**.

NO.	Description	Comments
1	KPNA1 (residues 69-538)	In pOPINF (his x6), M (MBP), S (SUMO)
2	KPNA1 WT (residues 1-538)	
3	HBV P WT (residues 1-831)	
4	HBV P TP domain (residues 1-182)	
5	HBV P TP domain (residues 1-174)	
6	HBV P RT domain (residues 329-682)	
7	HBV P RNAseH domain (residues 682-831)	
8	HBV P RNAseH domain (residues 682-771)	
9	DHBV P WT (Residues 1-788)	
10	DHBV P TP domain (Residues 1-205)	
11	DHBV P RT domain (Residues 383-655)	
12	DHBV P RNAseH domain (Residues 658-788)	
13	HBx WT (Residues 1-188)	
14	HBx (Δ154-188)	
15	HBV TP (1-182) + Linker + KPNA1 (residues 69-538)	
16	DHBV TP +linker + KPNA1 (residues 69-538)	
17	HBV MiniRT-like Hε binding (residues 42-196/291-520)	
18	DHBV MiniRT-like Hε binding (residues 75-244/353-575)	
19	HBV P RT domain + RNAse H domain (residues 329-831)	
20	DHBV P RT domain + RNAse H domain (residues 329-831)	
21	HBV P (Δ 203-287)	
22	HBV P (Δ 203-287 - Δ RNAseH)	
23	RT_A: HBV_RT (residues 286-676)	
24	RT_B: HBV_RT (residues 305-676)	

25	RT_C: HBV_RT_H (residues 286-676)	
26	RT_D: HBV_RT_H (residues 305-676)	
27	RT_E: HBV_RT_H (residues 286-832)	See above
28	RT_F: HBV_RT_H (residues 305-832)	
29	RNASEH_1: (residues 677-832)	
30	RNASEH_2: (residues 677-771)	
31	HBV Polymerase WT (residues 1-831)	Lentivirus
32	DHBV Polymerase WT (residues 1-788)	
33	HBx WT (Residues 1-188)	
34	HBV Pol + Hsp70 + ε	biGBac
35	HBV Pol + Hsp40	
36	HBV Pol + Hsp70	
37	HBV Pol + Hsp90 (expression was misleading due to purification of a contaminant)	
38	HBV Pol + Hsp40 + Hsp70	
39	HBV Pol + Hsp40 + Hsp90	
40	HBV Pol + Hsp70 + Hsp90	

Table 4: Summary of successfully cloned constructs. In the rightmost column the plasmid system used for the expression of each construct. Constructs in the pOPIN vectors have been tested for expression in *E.coli*, SF9 and HEK293T cells. Constructs detected as a single band on a western blot are highlighted in green, while in orange and red constructs that showed degradation and no expression, respectively.

A sample agarose gel with some correct PCR products is shown in **Figure 20**.

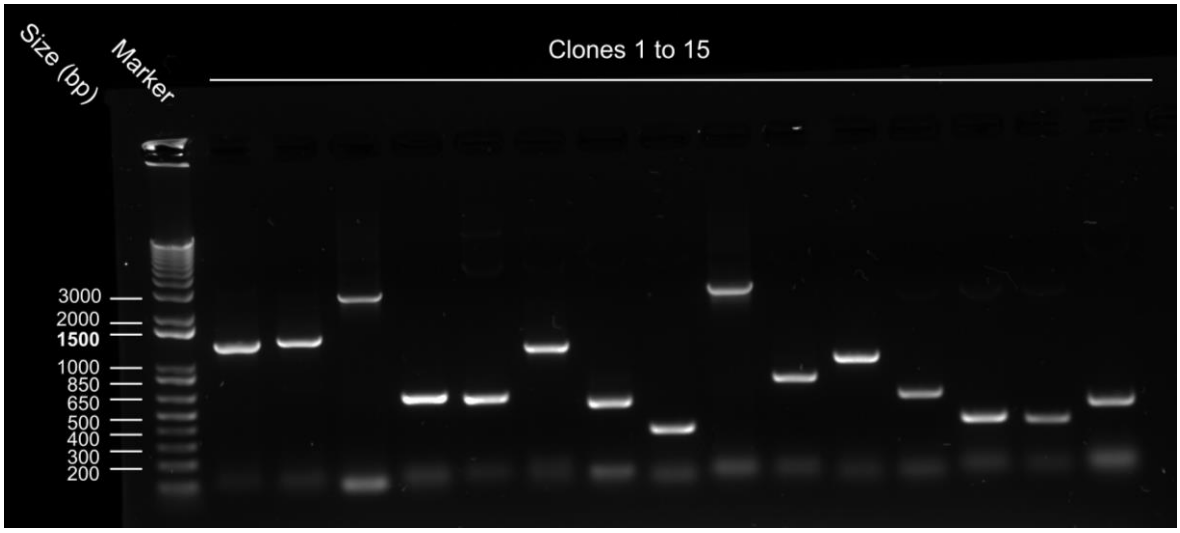


Figure 20: Sample agarose gel with PCR products of different clones of the HBV polymerase

With the exception of the lentivirus plasmids, which rely on ligation-based cloning, all other genes were inserted into their respective plasmids via ligation-independent cloning.

3.2.3 Expression tests with lentiviral vectors

Transfer plasmids harbouring the gene for the HBV polymerase were mixed with the appropriate quantity of packaging and envelope plasmids, then transfected into Lenti-X cells [14]. In the transfer plasmid, an IRES sequence, followed by the gene for a fluorescent protein, lies after the inserted gene. As some RNA is transcribed and translated, success of the transfection and infection steps can be assessed by fluorescence microscopy. For all the tested constructs, infection was successful (**Figure 21**). After harvesting the virus-containing supernatant, HEK293T cells were infected.

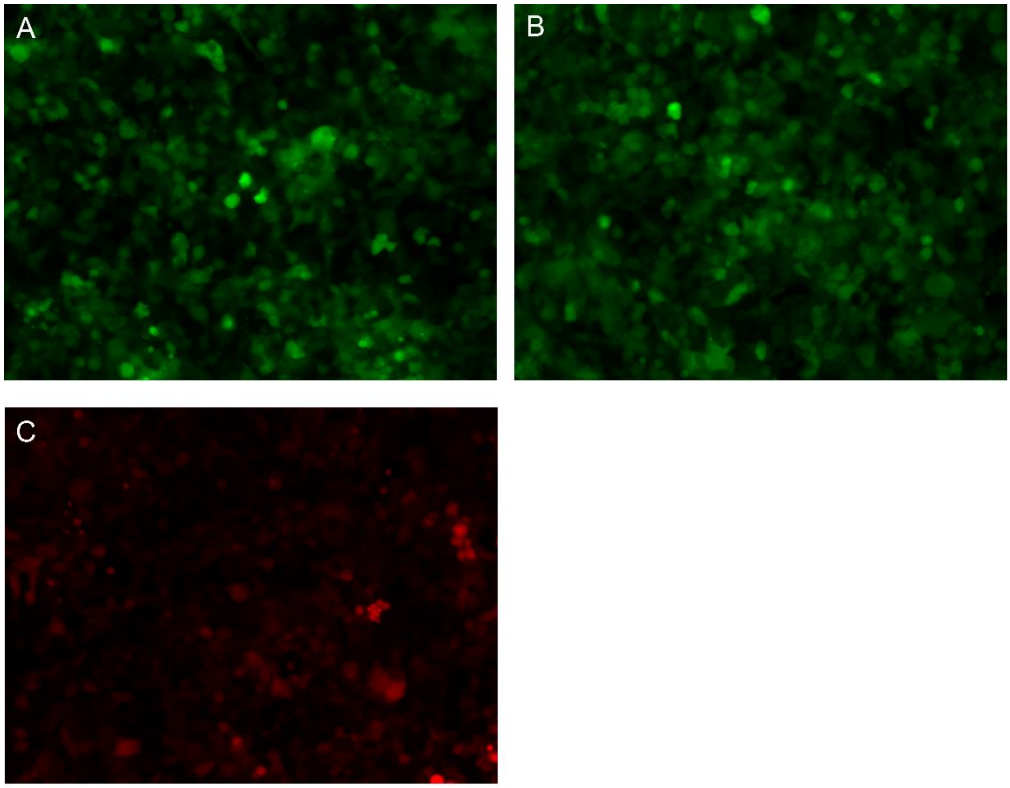


Figure 21: Fluorescence microscopy images of HEK293T cells infected with (**A**) Human HBV polymerase – IRES – EGFP, (**B**) Duck HBV polymerase – IRES – EGFP and (**C**) HBx – IRES – mCherry. Images were pseudo coloured in ImageJ. No scale bar was provided by the collaborator collecting these images.

Expression of the HBV polymerase was induced by supplementing the culture media with doxycycline and incubating for 72 hours at 37 or 30C. Cells were then harvested, lysed and the supernatant checked for expression by western blotting. In this specific case, as the polymerase has a twin-strep tag, the membrane was incubated for a single step with biotinylated horseradish peroxidase. While extremely specific, the signal-to-noise ratio of the resulting image is low (**Figure 22**).

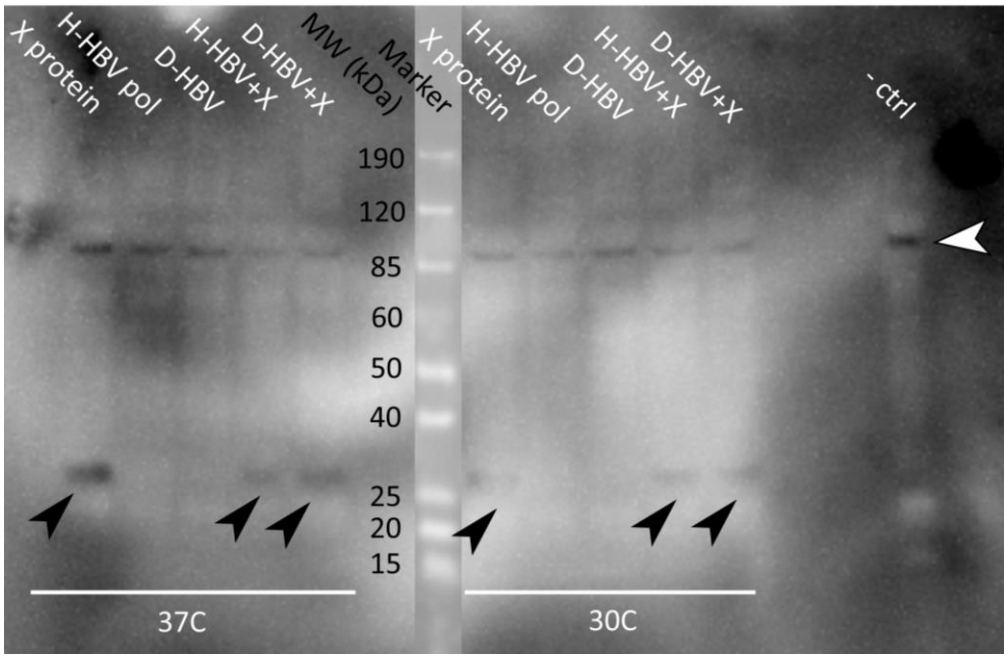


Figure 22: Anti-strep tag western blot of expression tests for the HBV polymerase in HEK293T cells. The white arrow points to a non-specific band (present also in the negative control). Black arrows indicate bands for the HBV X-protein.

Unfortunately, while expression of the X-protein was successful, no bands for the polymerases were detected, whether expressed alone or in combination with HBx. Therefore, it doesn't appear that in HEK293T cells, coexpression of HBx alone can support expression of the HBV polymerase.

3.2.4 Expression tests in *E.coli*

Expression of different domains of the HBV polymerase was also tested in *E.coli*, by transforming BL21* cells with the appropriated plasmids, then growing the

culture at 37C before inducing overnight expression with IPTG at 37C, 20C and 20C supplementing the media with 3% EtOH. This has been shown to enhance expression and solubility for some recombinant proteins [141].

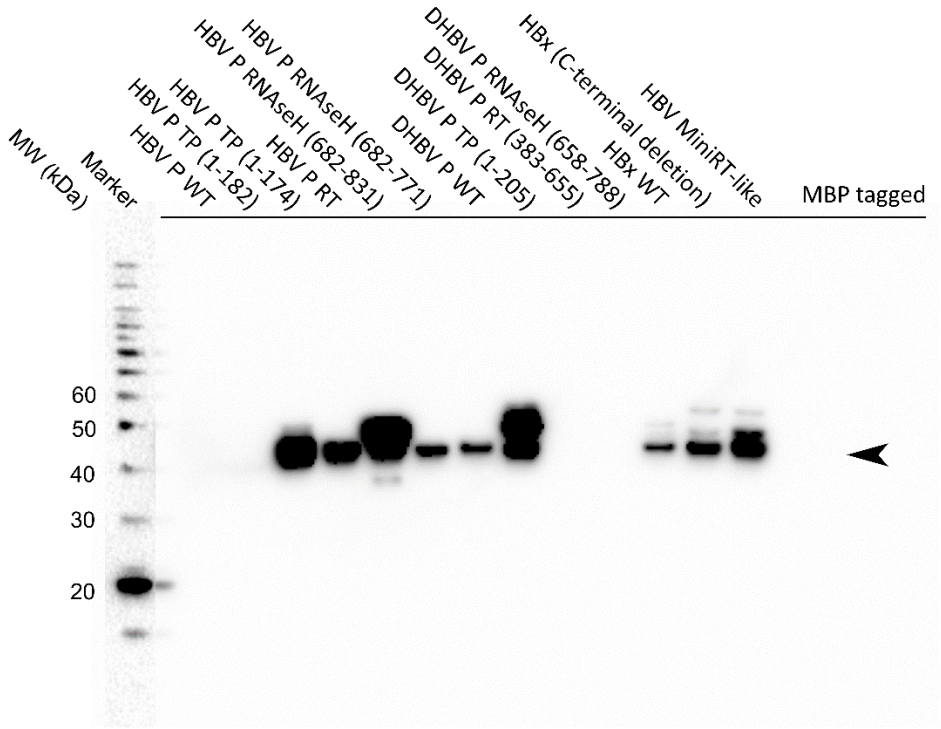


Figure 23: Anti-His tag western blot of *E.coli* expression test at 20C. Black arrowhead indicates the expected molecular weight for MBP tags. Several degradation products can be observed for multiple clones, with only the MBP band visible in some cases. Principal molecular weight bands are highlighted.

Expression at 20C is shown in **Figure 23**. While expression can be detected for some of the clones, in almost all the cases it appears that only the MBP tag is expressing.

Some higher bands can be detected (between 50 and 60 kDa) but they constitute the minority of the product. Similar results were detected at 37C, and while the addition of EtOH in the media changed the expression profile (**Figure 24**), it did not result in non-degraded protein. Despite previous reports of successful

expression of single domains of the HBV polymerase in *E.coli* [2], it was not possible to reproduce the published results and the strategy was abandoned.

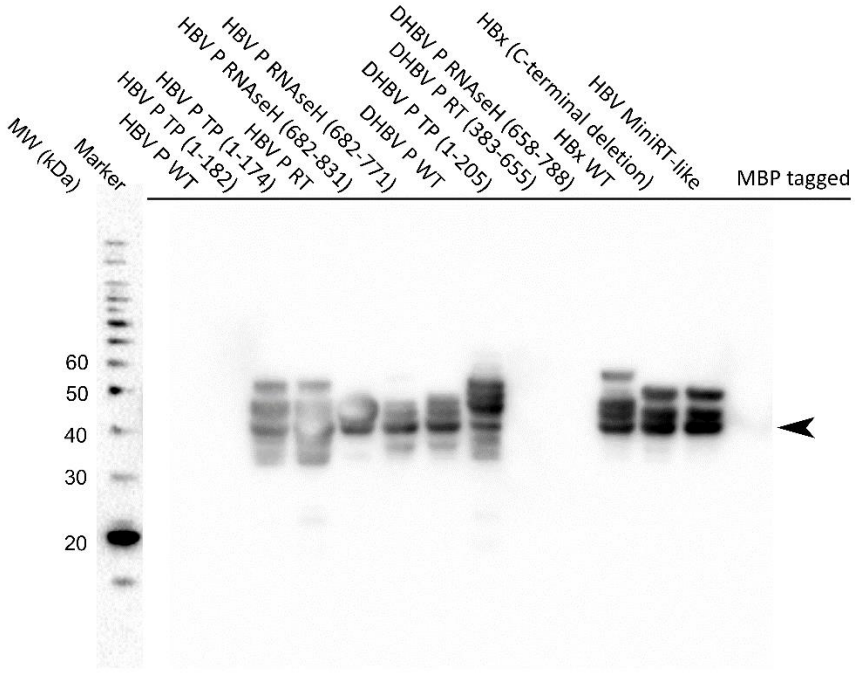


Figure 24: Anti-His tag western blot of an *E.coli* expression test at 20C, with 3% EtOH supplemented media. Black arrowhead indicates the expected molecular weight for MBP tags. Several degradation products can be observed for multiple clones, with a markedly different pattern compared to expression without EtOH. Principal molecular weight bands are highlighted.

3.2.5 Transient expression tests in HEK293T

Mammalian HEK293T cells are a system that is closer to the physiological environment of the virus and where a low level of expression of HBV polymerase has previously been reported [118][132][133]. Multiple rounds of transient expression tests were performed, screening different clones.

In **Figure 25,** it can be seen that while most of the tested clones were negative, there appears to be a low level of expression of Hsp70 when co-expressed with

the HBV polymerase. Unfortunately, it does not appear that there is any expression for the polymerase or any of the his-tagged clones.

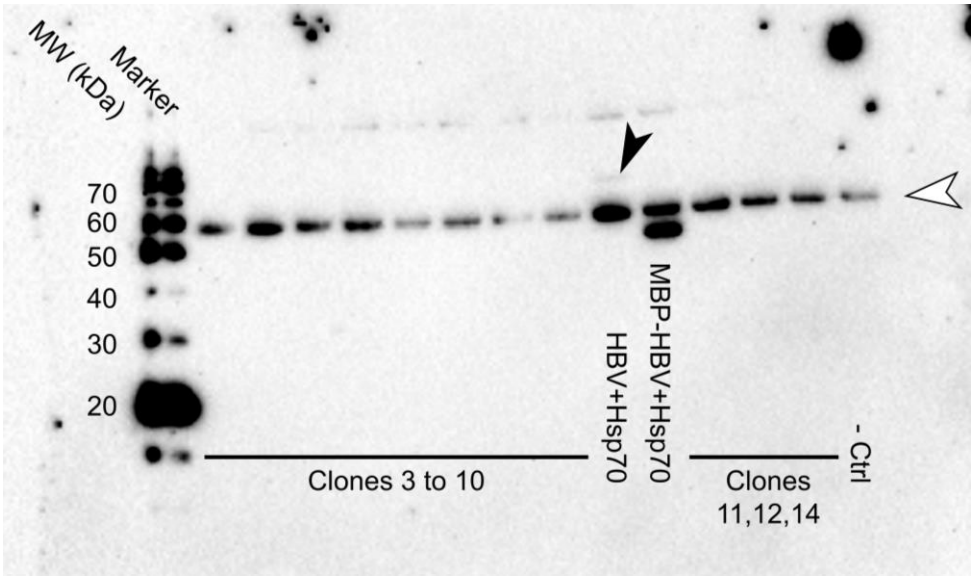


Figure 25: Anti-His tag western blot of transient expression tests in HEK293T cells. The white arrow indicates a protein that is non-specifically recognised by the anti-his antibody. The band is strongly visible likely due to overexposure of the western blot. The black arrow indicates likely low-level expression of Hsp70. Principal molecular weight bands are highlighted.

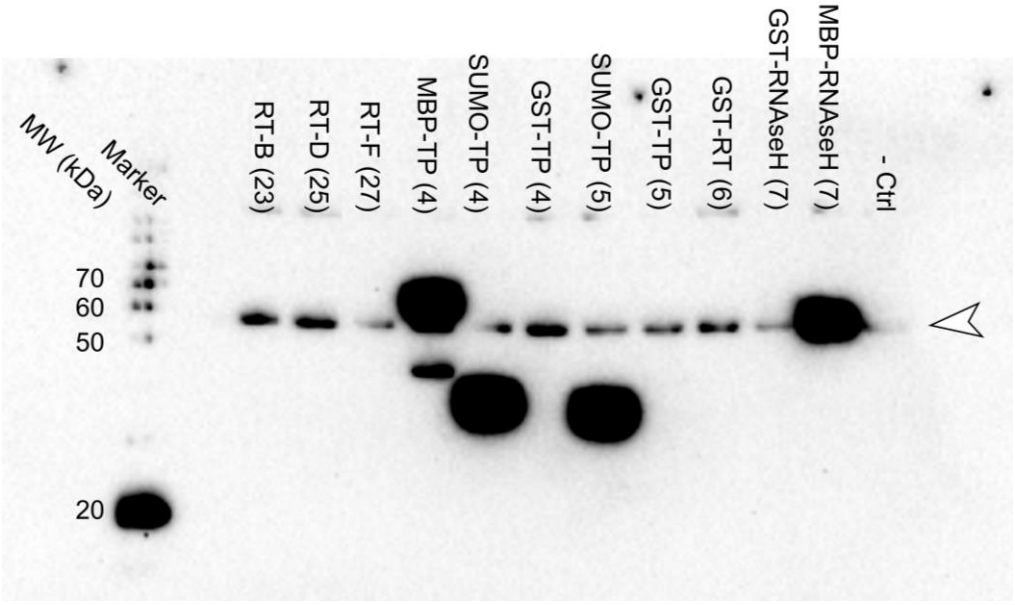


Figure 26: Anti-His tag western blot of transient expression tests in HEK293T cells. The white arrow indicates a faint band likely related to an endogenous protein that is non-specifically recognised by the anti-his antibody. The band is strongly visible likely due to overexposure of the western blot. While some of the constructs appear to express, their bands correspond to the expected molecular weight of the tag, plus some degradation fragments (i.e. MBP – 42 kDa and SUMO – 15 kDa). Principal molecular weight bands are highlighted.

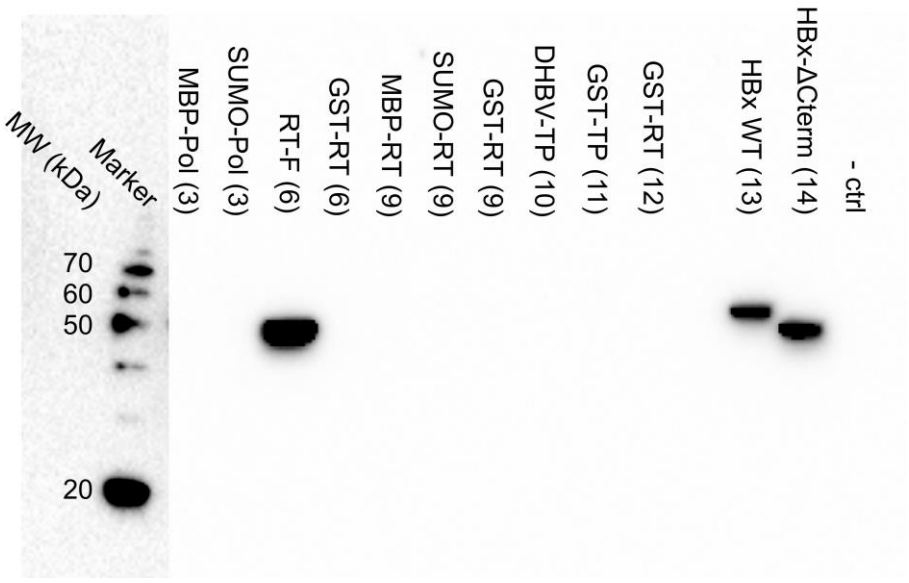


Figure 27: Anti-His tag western blot of transient expression tests in HEK293T cells. A degradation product of the HBV polymerase RT domain and full-length/ΔC-terminal HBV X protein can be observed. Principal molecular weight bands are highlighted.

Other clones tested in **Figure 26** and **Figure 27** had varying degrees of success: while most of them do not express (or only the tag seems to be present), single bands for what appears to be a degradation product of the polymerase, as well as HBx (roughly at twice the expected apparent molecular weight) are present. Overall, the rate of success with transient expression in HEK293T cells was very low, but these results are at least in agreement with the results of the co-infection with HBx using lentiviral vectors. Despite the fact that it was possible to express HBx, no further purifications were pursued as it was beyond the scope of this work.

3.2.6 Expression tests in insect cells

Expression of clones of the HBV polymerase in insect cells was also attempted. Again, while most tests were unsuccessful there were some outliers. It is worth noting that, unlike in HEK293T cells, it was not possible to successfully obtain a single non degraded band of HBx (**Figure 28**).

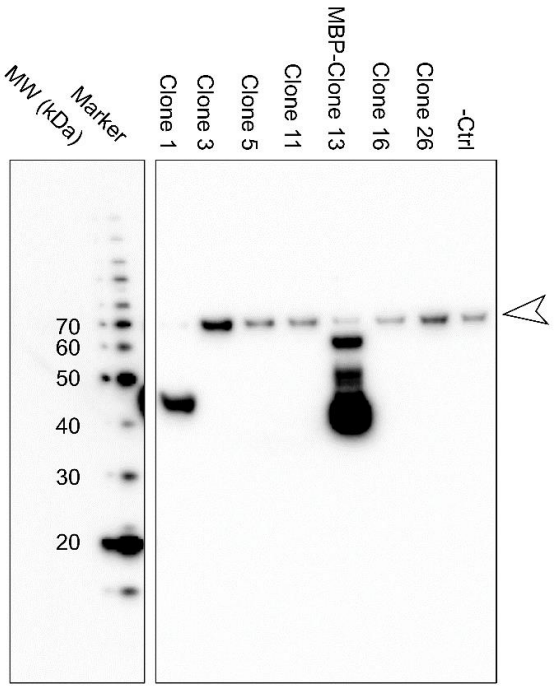


Figure 28: Anti-His tag western blot of HBV polymerase expressed in insect cells. The white arrow indicates a protein that is non-specifically recognised by the anti-his antibody. The band is strongly visible likely due to overexposure of the western blot. Principal molecular weight bands are highlighted.

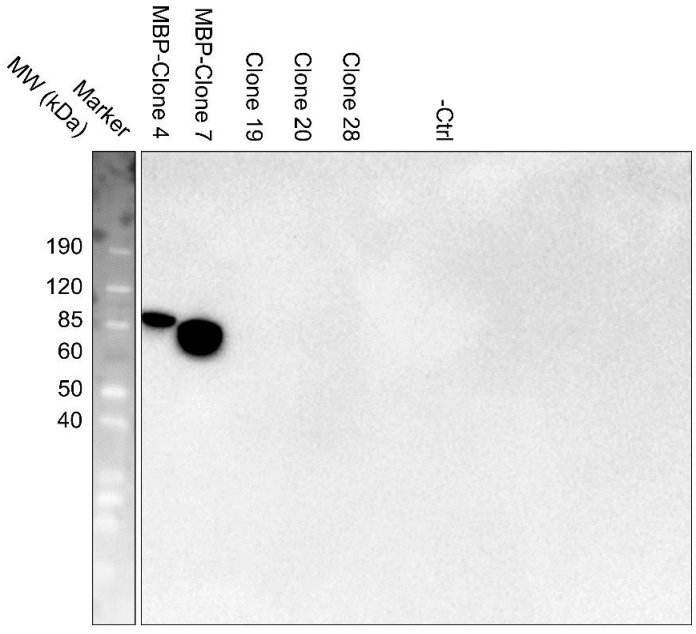


Figure 29: Anti-His tag western blot of HBV polymerase expressed in insect cells. Constructs 4 and 7, linked to MBP can be seen expressed as a single band. Principal molecular weight bands are highlighted.

Some constructs, however, were expressing successfully: MBP-tagged TP and RNAse H domains were showing at the expected apparent molecular weight (**Figure 29**). This two successful clones were used in subsequent purification tests.

Concerning biGBac-expressed polymerase, two of the tested combinations appeared to be successful: in one case, where the HBV polymerase was expressed together with Hsp90 a main band at ~70kDa, with a more faint band over 80kDa could reasonably be interpreted as a degradation product of the majority of the polymerase (***Figure 30***).

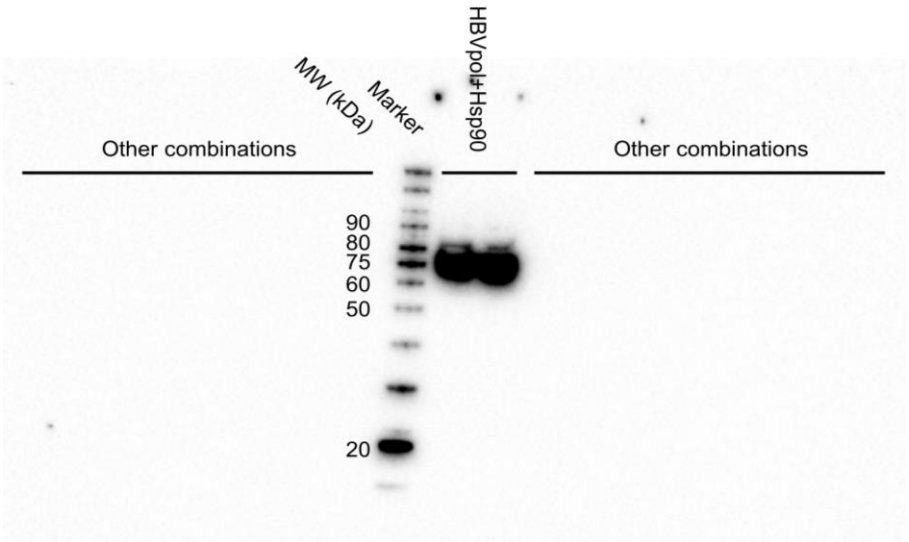


Figure 30: Anti-His tag western blot of HBV polymerase expressed in insect cells with the biGBac system. Bands corresponding to the molecular weight of the full-length HBV polymerase can be observed as well as more intense degradation products below. Other combination refer to tested clones 34 to 40. Principal molecular weight bands are highlighted.

The same combination, with a Twin-strep purification tag, yielded a clean band at ~90kDa, seemingly a full-length HBV polymerase (***Figure 31***).

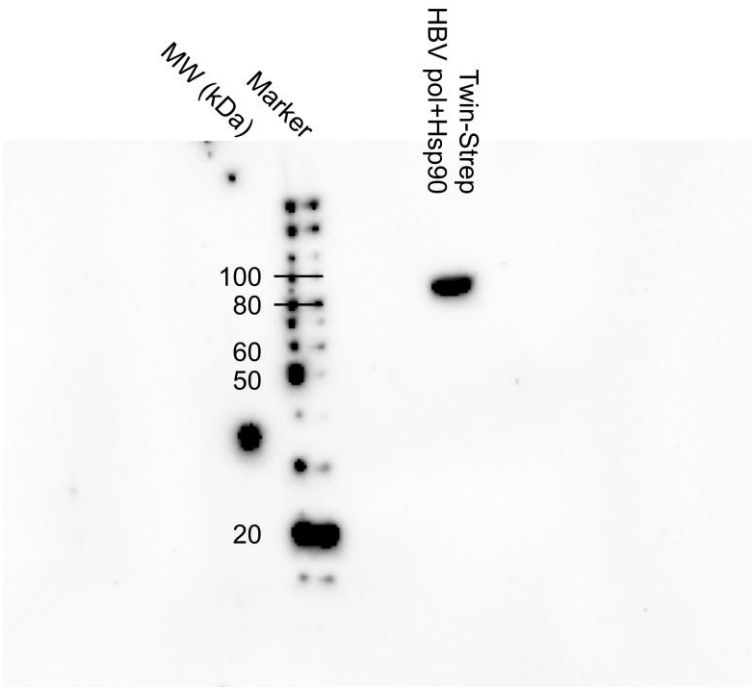


Figure 31: Anti-His tag western blot of HBV polymerase expressed in insect cells using the biGBac system. A his-tag was added for detection to the twin-strep tagged polymerase. Principal molecular weight bands are highlighted.

In retrospect, the addition of a positive control, on top of negative controls to many of the western blot would have been beneficial. It is worth noting, however, that in some cases (e.g. **Figure 27**) some of the bands are internally consistent, providing some confirmation of successful expression following transfections.

To summarise, the most successful expressing clones insect cell-expressed, MBP-tagged TP and RNAse H domains, and (near) full-length polymerase co-expressed with Hsp90. All of these clones were then used for further purification tests.

3.2.7 Purification tests of the HBV polymerase

Single domains were purified by affinity chromatography, followed by size exclusion chromatography. Several buffers were tested, starting from buffers reported in [2]. Overall, all of the single domain constructs suffered from massive

degradation issues, despite the buffer being supplemented with protease inhibitors (**Figure 32**). Multiple attempts at mitigating this issue by changing buffer composition and temperature proved to be unsuccessful. Western blot of the sample shows laddering of protein bands, with the predominant product sitting at ~40kDa, the molecular weight of the MBP tag (**Figure 33**).

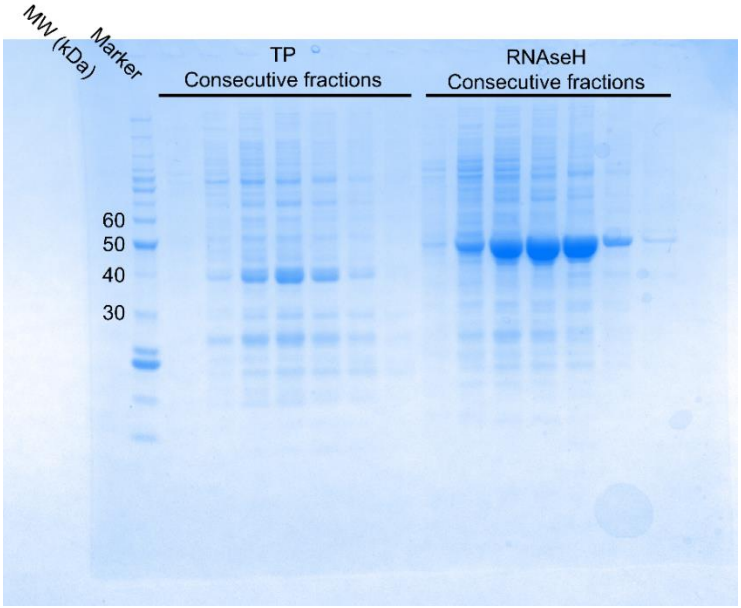


Figure 32: Post-SEC SDS-PAGE of MBP-tagged HBV polymerase TP and RNAseH domains. Principal molecular weight bands are highlighted. No corresponding SEC traces are shown, as no clear peak could be identified.

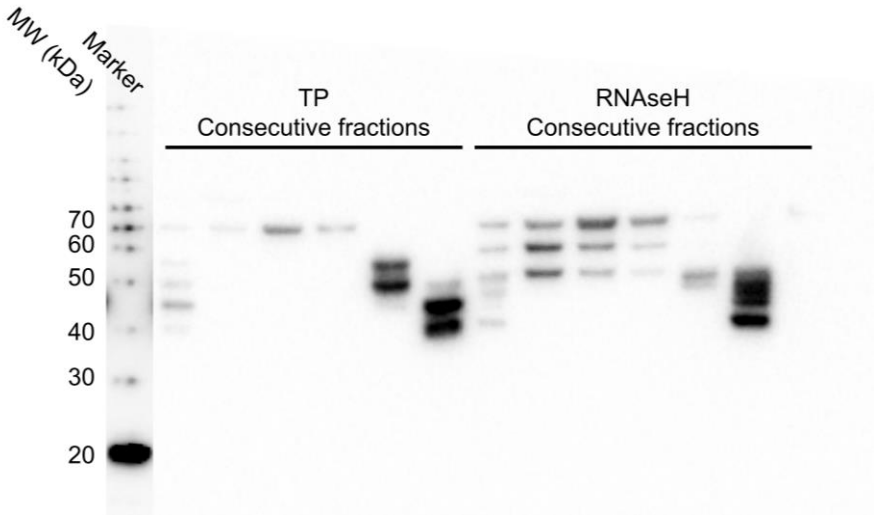


Figure 33: Post-SEC anti-His tag Western blot of MBP-tagged HBV polymerase TP and RNAseH domains. Principal molecular weight bands are highlighted.

For the his-tagged HBV polymerase, binding of the protein to the Nickel column for affinity purification was the main issue encountered: only a small amount of the desired protein could be recovered. Increasing the incubation time of the lysate (up to overnight) with the resin did not increase the yield. Under the assumption that this may have been caused by a less accessible histidine tag, some testing using different types of resins and buffers was done in a systematic way (**Table 5, Figure 34**).

Other resins tested were cobalt-based TALON resin which has been shown to increase specificity for histidine tags [142] and NiSepharose excel resin. Addition of 1M urea [143] or 5mM MgCl2 was also tested.

Resin / Buffer	TALON	Ni-NTA	Excel
Tris/NaCl Buffer	++	+	-
+1M Urea	+++	+	-
+$MgCl_2$	+	++	-

Table 5: Matrix of resins and buffers tested in the purification of the his-tagged HBV polymerase. The (+) signs indicate the intensity of the band relative to the faintest that could be identified, with more + indicating a higher intensity. A (-) sign indicates the lack of a band.

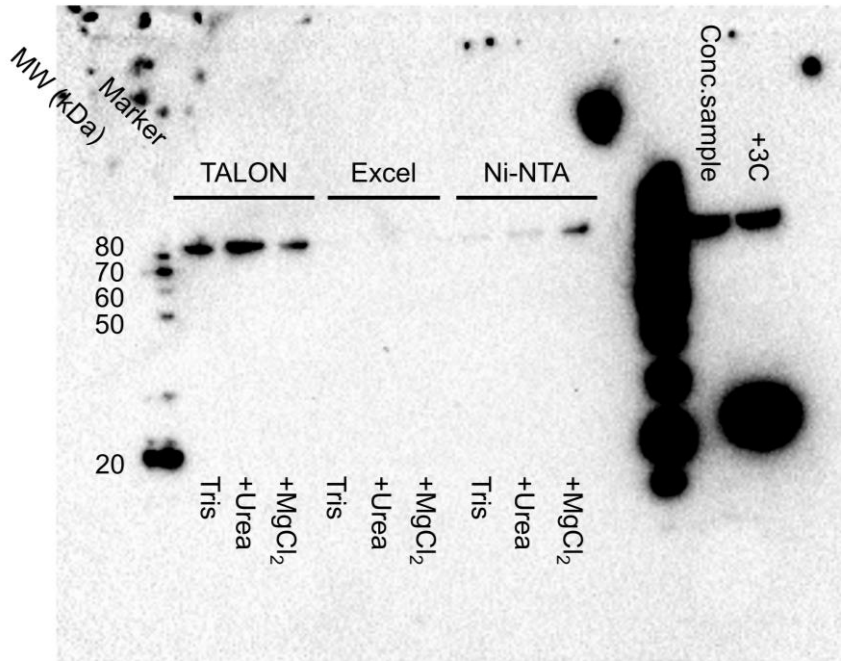


Figure 34: Anti-His tag western blot of the resins/buffer test. The same amount of lysate/beads was used all the lane. The two lanes at the end show the a concentrated sample of the polymerase and the same sample incubated with 3C. Principal molecular weight bands are highlighted.

The results of this screening clearly show a better binding of the protein to TALON beads. For this reason, all further purification attempts of his-tagged polymerase were carried out with this resin. Another approach that was tested exploits the well-known binding of polymerases to heparin [144]. A subsequent purification step using a heparin resin was also tested (**Figure 36**).

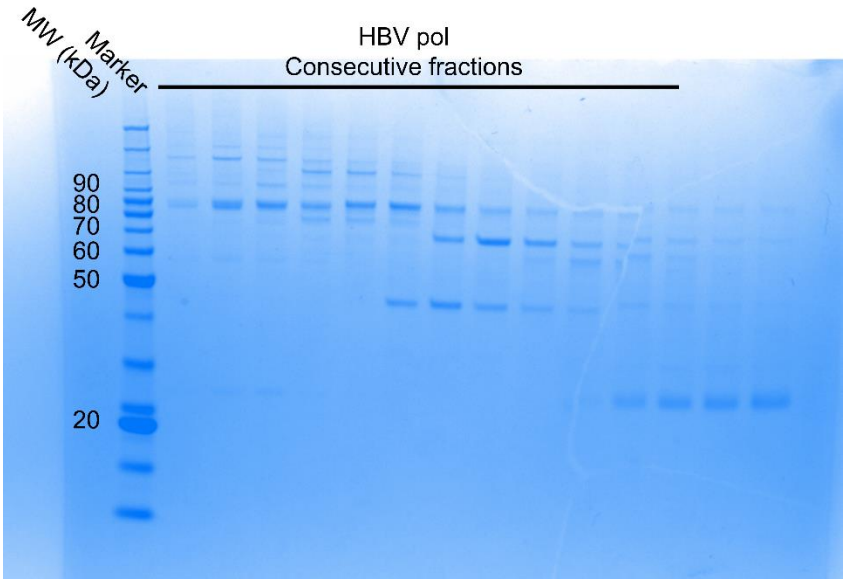


Figure 35: Post-Heparin and SEC SDS-PAGE of the his-tagged HBV polymerase. No corresponding SEC traces are shown, as no clear peak could be identified.

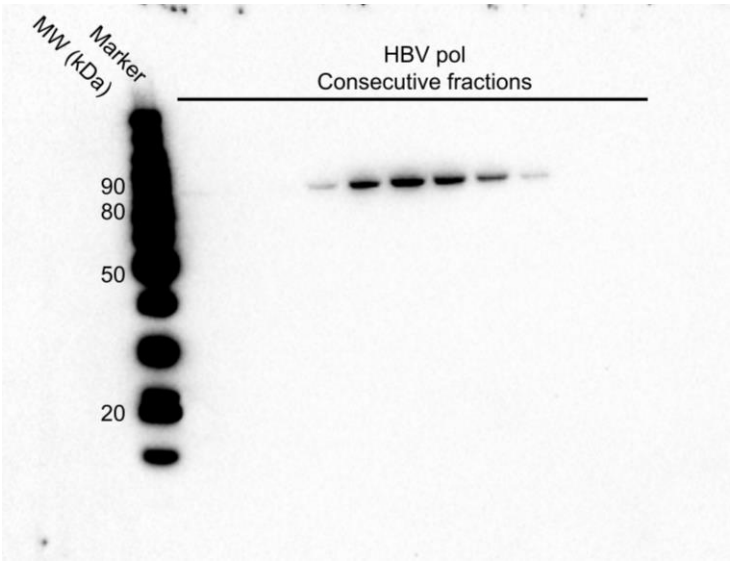


Figure 36: Anti-His tag post-Heparin and SEC western blot of the his-tagged HBV polymerase. Principal molecular weight bands are highlighted.

Despite clearly showing presence of the desired protein, this multi-step purification method was deemed too unreliable and wasteful and focus was shifted on the newly developed Twin-strep tagged polymerase constructs.

In this case, with a single purification step a clean, visible band could be seen. Western blot of the purified fraction with biotinylated HRP indicates the presence of a band at the expected molecular weight (**Figure 37**).

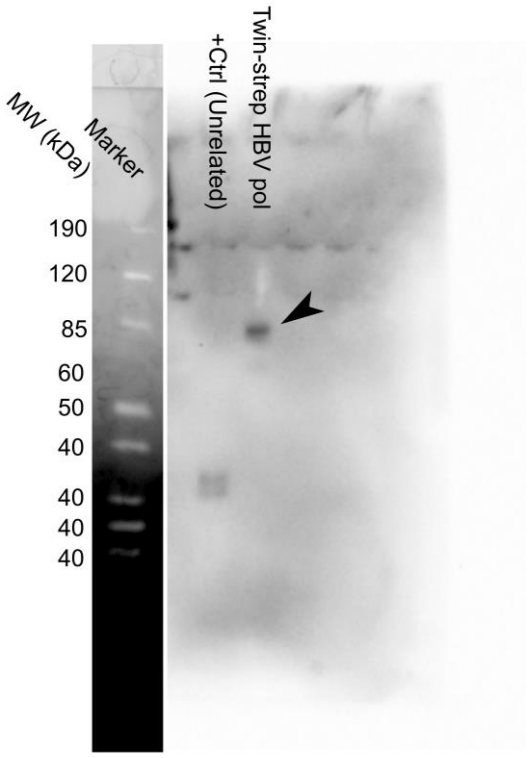


Figure 37: Biotin-HRP western blot of HBV polymerase. The black arrow points to the HBV polymerase band.

Puzzlingly, the output of downstream purification steps (Heparin/SEC) did not show the presence of any protein. Despite multiple attempts with different buffers, no improvement was observed. The observed bands from SDS-PAGE gels of the his- and twin-strep tagged HBV polymerases were then cut and analysed by mass-spectrometry to correctly identify the protein as the polymerase. Unfortunately, in both cases the protein was identified as a native *Spodoptera* protein, aspecifically purified during the process.

In the case of the his-tagged polymerase clone, the protein was unfortunately identified as the putative RNA-binding protein 15 (XP_022837292.1). Analogously, in the case of the Twin-strep tagged clone, identification of the protein by mass-spectrometry revealed the *Spodoptera* pyruvate carboxylase (XP_022834355.1). Due to the unfortunate results of this analysis and in

consideration of the time already invested in the project, all work related to the purification of the hepatitis B virus polymerase was stopped.

3.3 Discussion

Since its identification in 1983 [145], determination of a structure for the hepatitis B virus polymerase has eluded researchers for almost 40 years. Public health efforts to eradicate the virus from the planet are only supported by the existing vaccine. While the number of cases of chronic HBV infections in western countries is relatively small, the virus constitutes a major burden in poorer areas of the world, especially because of mother-child transmission [50]. Development of new, more potent and targeted drugs is also being hampered by the lack of structural information to aid drug design.

Aside from interest in the polymerase as a target for structure based drug-design, the polymerase itself is an attractive target in the study of the biology of the virus. Lacking any related structural homologue, the TP domain is of particular interest and it is possible that this domain represents a new protein fold.

The RNA-dependent DNA polymerase of the HBV has to support many functions, from DNA replication, to RNA retrotranscription, with protein-dependent priming activity in replication, and potentially shuttling of the genome into the nucleus of infected cells [94]. The peculiar replication mechanism adopted by the virus [68] implies that all these functions have to be carried out in the confined space of a very small virion. I believe this constitutes a fascinating biological puzzle that warrants more attention than is currently being given.

While the outcome of this project cannot be considered a success, a retrospective analysis of the results can give some clues as to how the two contaminants were accidentally purified. In the case of the putative RNA-binding protein 15, a brief analysis of its sequence highlights the presence of a large number of histidine

residues (12) within a region that, according to the Uniprot notation (entry A0A2H1VB60_SPOFR), is predicted to be disordered (**Figure 38**).

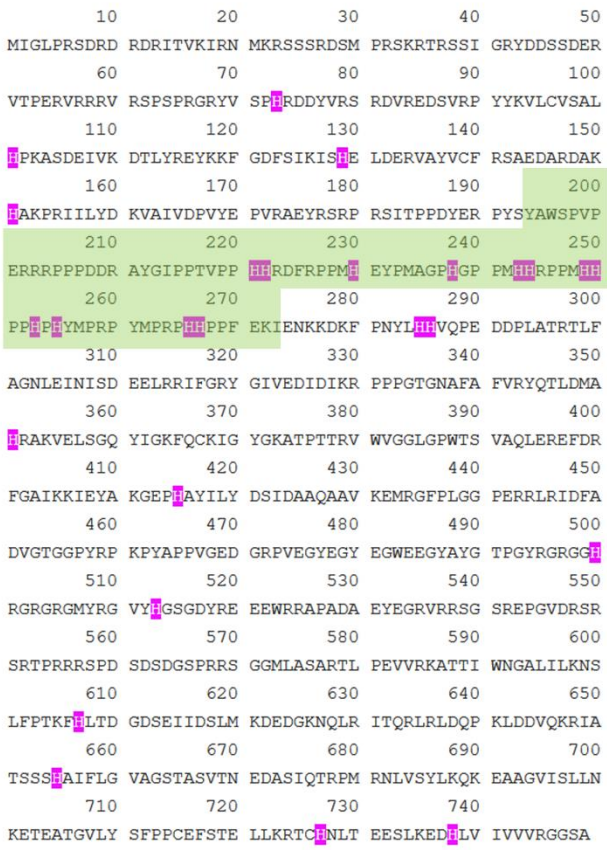


Figure 38: Sequence of the *Spodoptera* putative RNA-binding protein 15. All the histidine residues in the sequence are highlighted in pink, while a predicted disordered region within the protein is highlighted in green.

While it cannot be conclusively proven, the presence of this region with potentially exposed histidine residues could explain a weak binding to the Nickel resin, as well as binding by the anti-His antibody. One thing worth mentioning is that throughout the experimental work, in routine quantitation of the sample using a nanodrop spectrophotometer, the 260/280 ratio was consistently approximately 1. This persisted despite the addition of Turbonuclease to degrade nucleic acids and was wrongly attributed to a small amount of residual RNA bound to the polymerase. Assuming that this putative protein is RNA-binding, this would

constitute evidence that the function is indeed real. Homologues of this protein for which a structure is available are the hypothetical RNA binding protein BC052180 from *Mus musculus* (~70% sequence identity, 20% coverage), and various transcription factors/repressors (~20-30% sequence identity, 14-20% coverage).

When trying to understand the inadvertent purification of the pyruvate carboxylase, it is well known that the enzyme has biotin as a covalently bound cofactor that intervenes during the catalysis reaction [146]. In the pyruvate carboxylase, the biotin is in the active site (**Figure 39**). It is therefore reasonable to speculate that it is accessible enough for binding of the Twin-strep resin to take place. This has also been observed before [147].

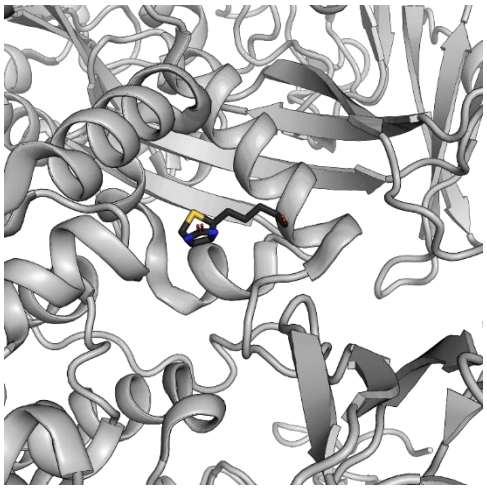


Figure 39: The biotin cofactor in the active site of *S.aureus* pyruvate carboxylase (PDB id: 3BG5).

Accidental purification and sometimes crystallization of contaminant proteins is fairly common [148] (see also **5.2.1**). This case has been particularly unfortunate, as two different proteins, with two different purification methods were misleadingly purified.

While the project did not reach the expected conclusion, cloning of essentially all the planned constructs, as well as their testing worked well. In retrospect, however, no significant expression of the HBV polymerase could be detected.

I consider the requirement for cellular proteins of the Hsp family [101] a key element for the successful expression of significant amounts of the polymerase. While these proteins are normally available in the cell, they will likely be unable to support overexpression of this protein. To solve this, one possibility is to use a weaker promoter and increase the volume of cultured cells instead.

Alternatively (or in combination with), a sequence for the ε stem-loop could be inserted after an IRES to couple transcription of the polymerase gene with that of the RNA sequence that is first bound by the enzyme. With this system, following binding of the polymerase to ε, cellular Hsps would be free to bind another copy of the overexpressed polymerase and so on.

The lentiviral expression system [14] would also be useful in testing the possibility that other cell lines that support HBV replication may be better equipped for expressing this protein (e.g. HepG2, Huh-6, Huh7, HepCHLine-7 [149]).

Ideally this approach could be coupled with a directed evolution approach [150], aimed at improving and expression levels of the enzyme. By subcloning the mutated constructs into a vector with a fluorescent protein in-frame, selection of the best performing clones would be easily carried out by detecting the resulting fluorescent signal. In this case, generation of the recombinant lentivirus and infection of the cells would be a significant bottleneck, likely benefitting from an automated approach.

4 Structural studies on antibodies inhibiting the activity of the New Delhi Metallo β-lactamase 1 (NDM-1)

4.1 Introduction

4.1.1 A brief overview of the New Delhi Metallo β-lactamase 1

One of the biggest threats to public health is bacterial antibiotic resistance, with an estimated 700,000 worldwide deaths each year associated with infections caused by drug-resistant microorganisms [151]. Antibiotic resistance can arise in multiple ways including: mutation of the antibiotic target, decrease of the import or increase of the export of the antibiotic, alteration of cellular pathways and, or by modification or degradation of the antibiotic itself [152]. Bacteria that carry multiple antibiotic-resistant genes, commonly known as "superbugs", represent an increasingly growing problem worldwide.

Following Alexander Fleming's discovery of penicillin from *Penicillum chrysogenum* in 1928, and the identification of its structure by Dorothy Hodgkin in 1945 [153], many β-lactam antibiotics such as Amoxicillin and Ceftriaxone have been discovered and introduced in clinical practice. Over the years, the usage of β-lactam drugs has increased massively: in the USA, from 2004 to 2014 this class of antibiotics constituted ~65% of all prescribed injectable antibiotics [154].

β-lactam antibiotics are bactericidal molecules characterised by the presence of a β-lactam ring (highlighted in red in **Figure 40**). β-lactam antibiotics bind to

penicillin binding proteins, bacterial enzymes involved in the formation of the peptidoglycan, thus blocking the synthesis of new cell wall.

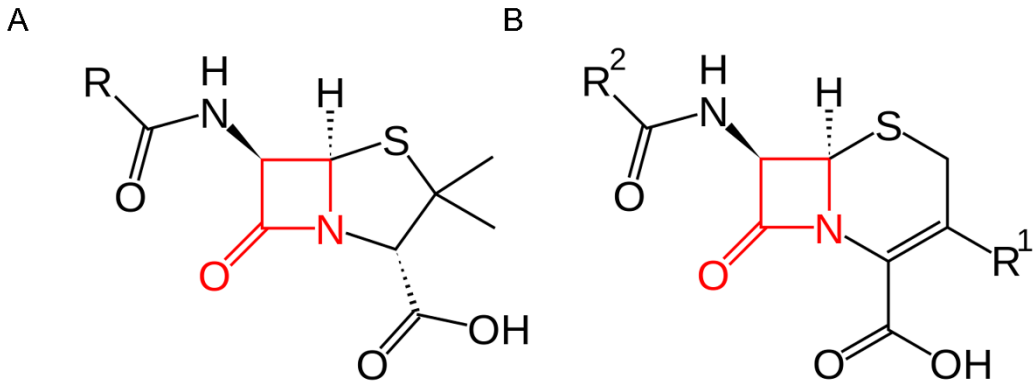


Figure 40: Core chemical structures of (A)Penicillins and (B)Cephalosporins. The β-lactam ring is highlighted in red in both cases. R/R^2 indicate points where different functional groups give rise to the different β-lactam antibiotics

Some bacterial species produce β-lactamases, enzymes capable of hydrolysing these β-lactam drugs. Since the identification of penicillinase, the first β-lactamase to be characterised in bacteria [155], the number of these enzymes has grown to a large family. In order to counteract the activity of these enzymes, several inhibitors have been developed [156] as well as new antibiotics and combinations of the two[157], beginning an "arms race" with these bacterial species.

To categorise these enzymes, the most widely accepted system is the Ambler classification[158] (and subsequent revision [159]). This classification method, based on the sequence of these proteins, divides β-lactamases into 4 different classes A, B, C and D (**Figure 41**). Classes A, C and D are serine β-lactamases, while class B rely on metallic zinc ions for the catalysis mechanism. The cluster formed by subclasses B1 and B2 are only structurally related with subclass B3.

According to this classification, the New Delhi Metallo β-lactamase 1 is part of the B1 subclass of β-lactamases.

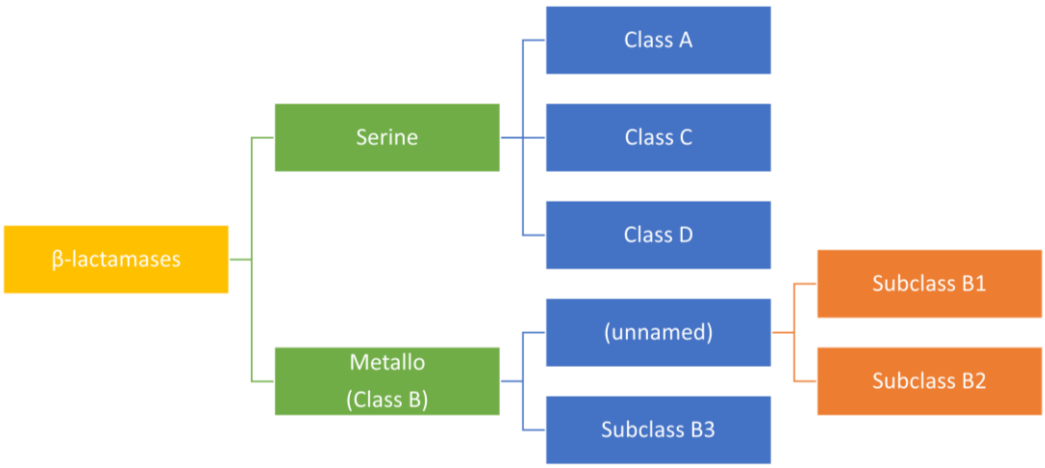


Figure 41: Ambler classification of β-lactamases

A functional classification of β-lactamases is also currently in use, where the behaviour of these enzymes in the presence of different substrates and inhibitors is used to sort these molecules into 17 different functional groups [160].

Of these groups, metallo β-lactamases (MBLs) in particular are the main cause of concern for antibiotic resistance, as available inhibitors for serine β-lactamases are ineffective against MBLs and the relatively low prevalence of MBLs in western countries has resulted in less resources being invested in the development of new inhibitors[4]. Testing for β-lactamases in a clinical setting is also non standardised and can give rise to conflicting results, further complicating diagnosis and treatment[161].

The New Delhi metallo β-lactamase 1 (NDM-1) was discovered in 2009 in a Swedish patient of Indian origin that acquired a urinary tract infection, following a trip to New Delhi [162]. The etiological agent for this infection was identified

as *Klebsiella pneumoniae* harbouring the gene for a previously unknown, inhibitor resistant β-lactamase. The most common bacteria harbouring a gene for NDM-1 are gram-negative bacteria from the *Enterobacteriaceae* family such as *E.coli* and *K.pneumoniae*, but the gene has spread through horizontal gene transfer into other species and has the potential to spread even further [163]. The NDM-1 gene has indeed been found in several plasmids: it is located within transposon Tn125, which is being implicated as the main driver for the dissemination of antibiotic resistance within the *Acinetobacter* species[164]. It has also been found in *Acinetobacter lwoffii* in a plasmid containing the gene for a type IV secretion system as a virulence factor [6]. The gene for NDM-1 has also been found within other big plasmids together with other antibiotic resistance related genes [165].

4.1.2 New Delhi Metallo β-lactamase 1 bioinformatic analysis

NDM-1 is a 270 residue long polypeptide, with the 28 N-terminal amino acids encoding a signal peptide, cleaved following the export of the protein to the periplasmic space, where it is anchored to the outer membrane [166],[167]. This signal peptide is predicted to contain a lipidation signal [167].

Several β-lactamases are lipidated in gram-positive bacteria, such as Penicillinase 1 in *Staphylococcus aureus*, or β-lactamase III in *Bacillus cereus* [168][169]. In gram-negative bacteria, however, few of them are known to be lipidated, with the most well studied being BRanhamella-mOraxella-1 (BRO-1) from *Moraxella catarrhalis* [170]. While the first report that NDM-1 was lipidated was almost overlooked [167], this property is now being investigated as a potential druggable target. Compared to a lipidation-deficient mutant, lipidated and membrane-bound NDM-1 has been shown to be resistant against Zn depletion via calprotectin (a

host defence protein that binds metallic ions) (**Figure 42**). The presence of a lipid moiety allows NDM-1 to be inserted into outer membrane vesicles (OMV) that are secreted by the bacteria during infection. It has been speculated that these OMV can contain plasmids harbouring the gene for NDM-1 [166].

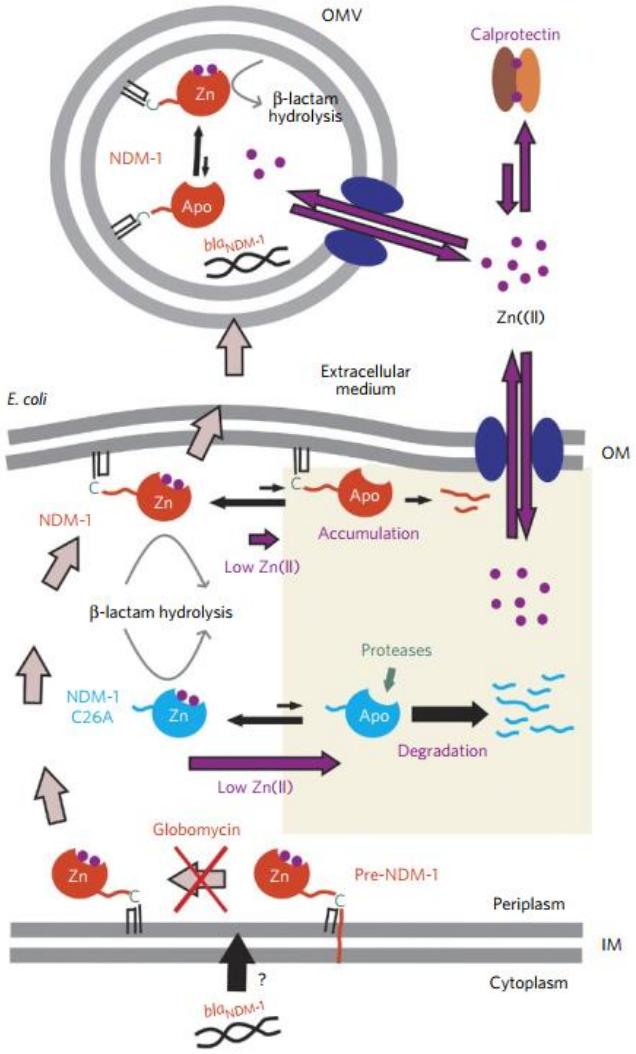


Figure 42: Model of the stabilization of NDM-1 by membrane anchoring and protection from calprotectin-mediated Zn depletion. NDM-1 can also be found bound to the membrane of secreted outer membrane vesicles. IM: Inner Membrane, OM: Outer membrane, OMV: Outer membrane vesicle. C: cysteine. From [166].

Bioinformatic analysis of the NDM-1 sequence using the GPS-lipid server[171] identifies with high confidence a lipidation signal for S-palmitoylation on Cys

26. No structural information for that region is currently available. This is consistent with the result obtained from the analysis of the protein sequence using the SignalP 5.0 server [172], which identifies a lipoprotein signal peptide, with the cleavage just before Cys26 with a probability of 0.9970.

Several variants of NDM-1 that differ by 1 to 5 amino acids have been identified (**Figure 43**) and (**Table 6**). These variants have likely appeared due to increased selective pressure from antibiotic treatment, especially in countries where antibiotic usage is widespread. The vast majority of NDM-1 and of its variants have been identified in Asia with ~58% of the isolates (mainly in China and India), followed by Europe. Americas, Africa and Australia combined comprise less than 25% of the total isolates (2013-2017 data) [173].

While the mutations giving rise to these natural variants do not appear to alter the secondary structure of the enzyme, they have been reported to alter both the thermal stability and the kinetics of the enzyme [7]. Overall, most of the mutations are located on the loops of the enzyme (see **Figure 44**) and, out of all the NDM variants, only NDM-10 and NDM-12 have mutations in the active site. While NDM-10 has an increased activity, the opposite is true for NDM-12 [174],[175] Comparing NDM-7 and NDM-8, aspartate 130 is mutated into an asparagine and a glycine, respectively. Relative to NDM-1, the catalytic activity of NDM-7 is increased, whereas a similar level was observed for NDM-8 [176],[177]. Other residues not in the enzyme active site, such as Trp93 have been shown to alter the enzyme activity [178]. To summarise, while for many single mutants it has been possible to ascertain the effect of the mutation, for variants with multiple mutations, discerning the effect of each in combination is much more difficult [179].

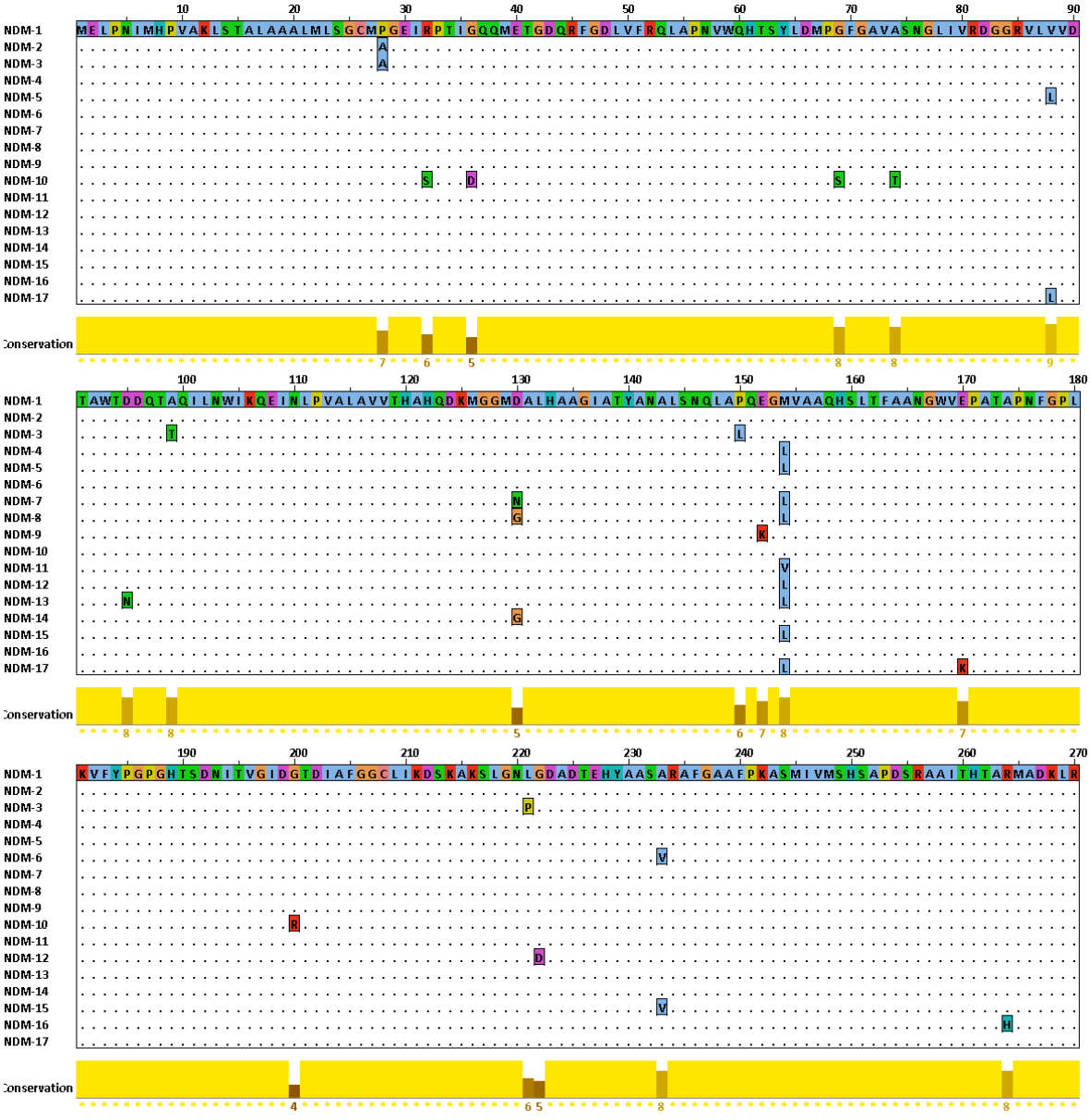


Figure 43: Multiple sequence alignment of different NDM variants. Residue conservation is calculated based on the physico-chemical properties of the mutated residues, based on [180].

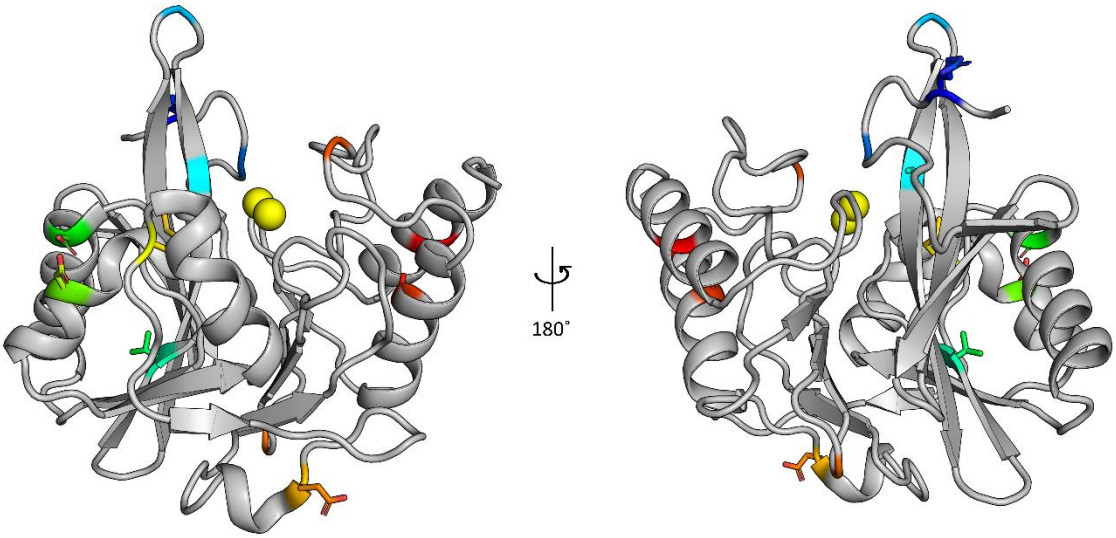


Figure 44: Sites of mutation of natural variants of NDM-1. Sites are highlighted in different colours and the side chains of the original residues are shown in stick representation.

NDM-1 Variant	Mutation(s)	No. of mutations
NDM2	P28A	1
NDM3	D95N	1
NDM4	M154L	1
NDM5	V88L, M154L	2
NDM6	A233V	1
NDM7	D130N, M154L	2
NDM8	D130G, M154L	2
NDM9	E152K	1
NDM-10	R32S, G36D, G69S, A74T, G200R	5
NDM-11	M154V	1
NDM-12	G222D, M154L	2
NDM-13	D95N, M154L	2
NDM-14	D130G	1
NDM-15	A233V, M154L	2
NDM-16	R264H	1
NDM-17	V88L, M154L, E170K	3

Table 6: Natural variants of NDM-1 and their mutations relative to NDM-1.

4.1.3 New Delhi Metallo β-lactamase 1 structural analysis

NDM-1 is a small globular protein, with an αβ-βα sandwich fold, comprising two β-sheets surrounded by four α-helices. The enzyme is capable of accommodating a wide variety of β-lactam compounds thanks to the wide groove that harbours the active site [5].

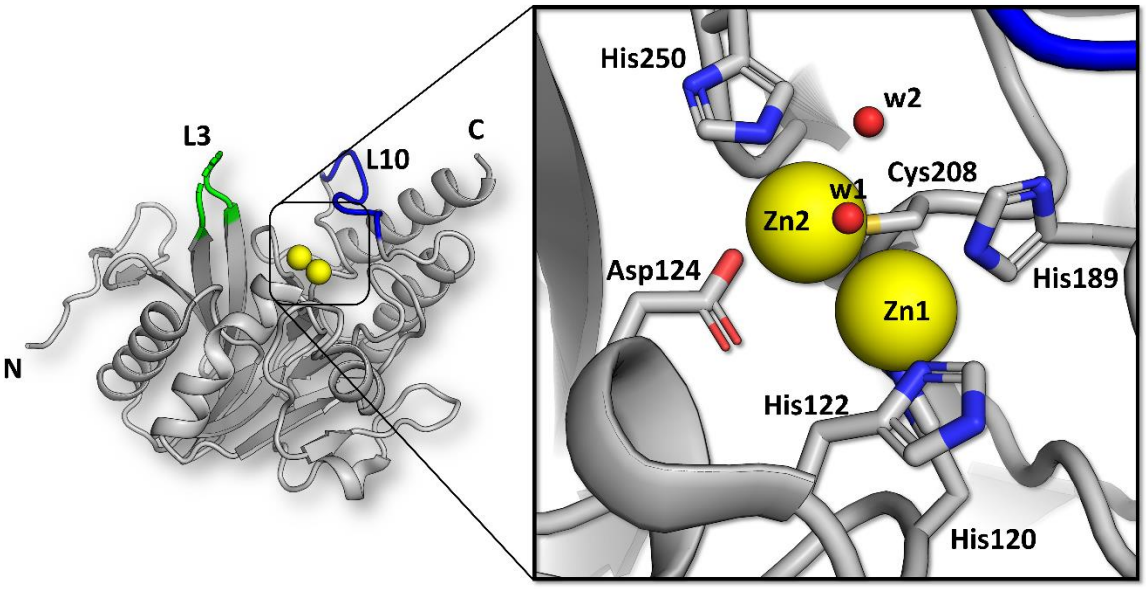


Figure 45: The structure of NDM. Highlighted in yellow are the two zinc atoms, in green loop L3 and in blue loop L10. In the insert, close up view of the active site, where the amino acids coordinating the zinc atoms are highlighted. PDB ID: 5NBK.

In the active site (**Figure 45**), two zinc atoms are coordinated by several residues: zinc 1 is coordinated by 3 histidine residues (His120, His122, His189) and a water molecule, while zinc 2 is by Asp124, Cys208, His250 and a water molecule.

Two loops surrounding the active site are involved in substrate recognition, loop L3 and L10 Loop L3 in particular, assumes multiple different conformations (**Figure 46** and **Figure 47**). to accommodate different compounds by tailoring the size of the groove to the substrate: for penicillins and cephalosporins, L3 opens up the active site, the opposite is true for carbapenem substrates [181].

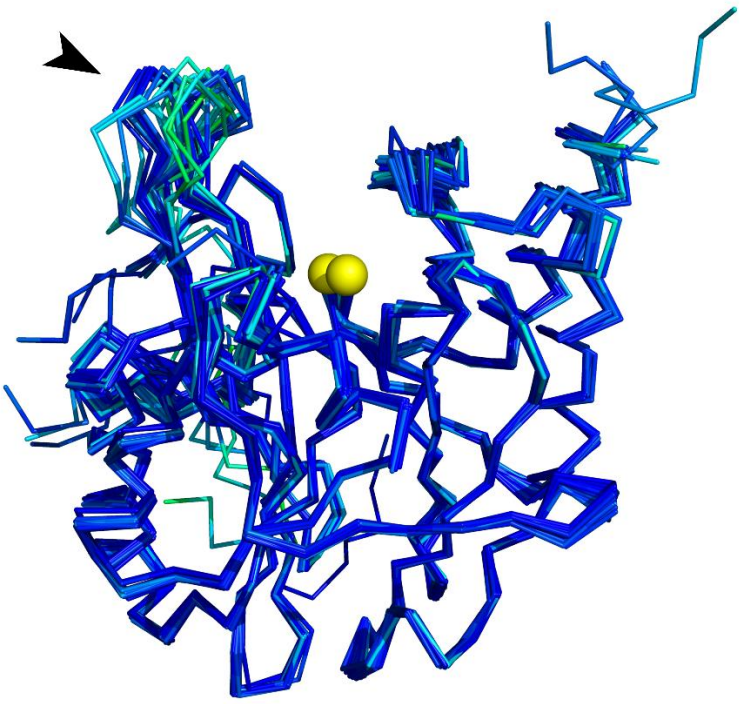


Figure 46: Superimposition of NDM-1 structures. The ribbon traces indicate the main chain paths and the structures are coloured according to their B-factors. Highlighted in yellow the two zinc atoms in the active site and the black arrowhead indicates the L3 loop.

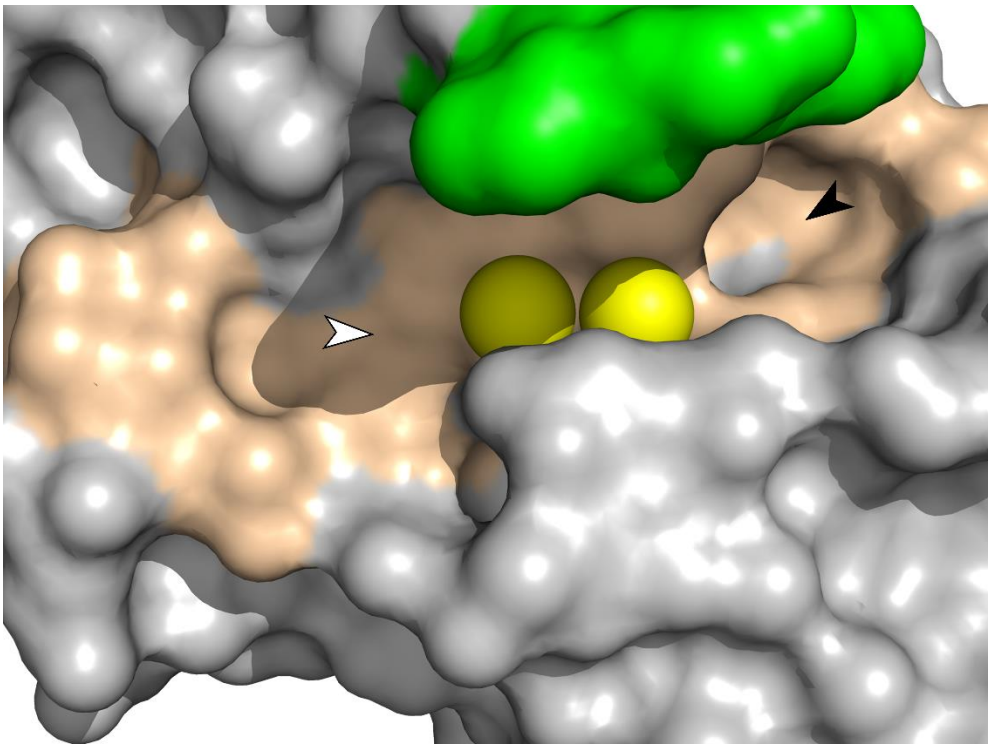


Figure 47: Surface representation of the groove that harbours the active site of NDM-1. Highlighted in green the L3 loop and in yellow the two zinc atoms coordinated in the active site. The brown patch shows a large groove, capable of accommodating a variety of molecules. The black and white arrows indicate where R groups (see **Figure 40**) would lie for carbapenems and penicillins, respectively.

Different reaction mechanisms have been proposed for different substrates, but in all of them the reaction starts with the hydroxyl from the water molecule coordinated to zinc 1 acting as a nucleophile, attacking the carbonyl carbon on the β-lactam ring [182]. For all the proposed mechanisms, the end result of the reaction is the breakage of the ring and the inactivation of the substrate, followed by the release of the deactivated product [183]. **Figure 48** shows the proposed mechanism for the hydrolysis of a cephalosporin substrate.

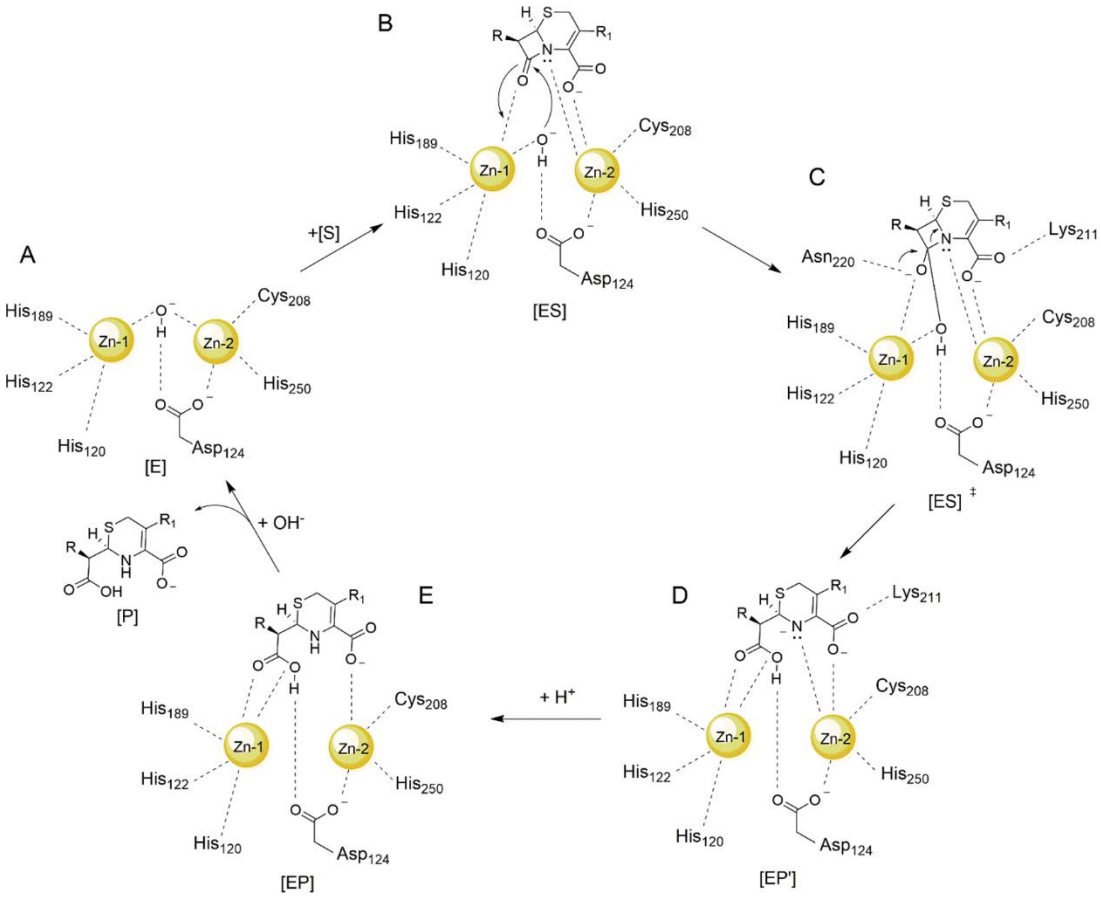


Figure 48: Proposed mechanism for the hydrolysis of a cephalosporin. (**A**) Initial state of the active site. (**B**) Substrate is bound and nucleophilic to the β-lactam ring. (**C**) Transition state. (**D**) The hydrolysed substrate is bound to the active site. (**E**) The product is first neutralised, then released from the active site. From [183]

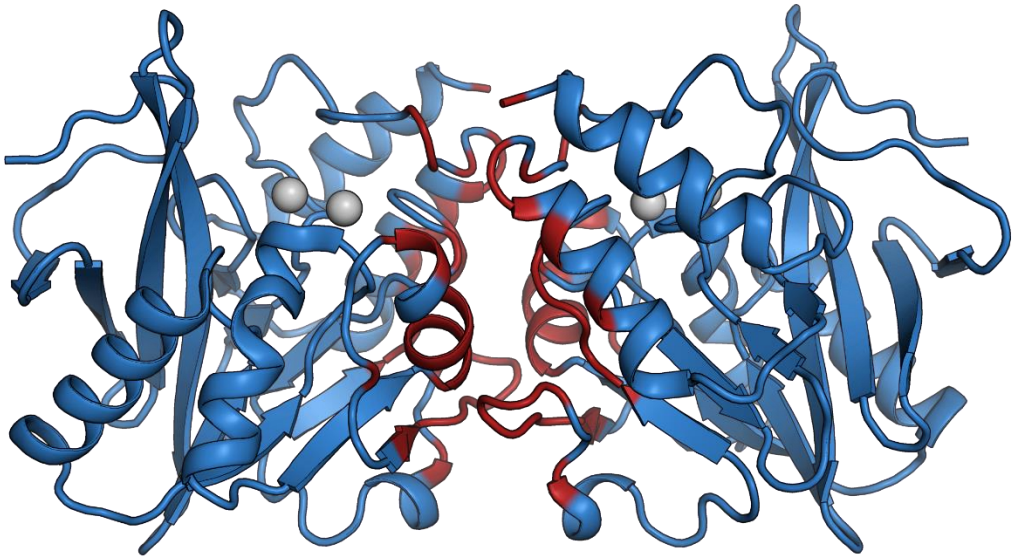


Figure 49: Crystallographic dimer of NDM1. The binding interface is highlighted in red. PDB id: 5ZGE

Dimerization of full-length NDM-1 has also been reported [184] and can be disrupted by mutating the interface. The dimer has been observed in crystal structures (**Figure 49**), as well as in solution, as verified by size exclusion chromatography [167]. The dimerization interface, as calculated by the PISA server [45] extends for ~1600 $Å^2$ and include several hydrogen bonds and salt bridges. However, it has been shown that this interface alone is not sufficient to induce a stable dimer *in vivo* and likely requires other factors, such as membrane anchoring [185].

4.1.4 Aim of the project

Antibiotic usage in clinical and agricultural settings is pervasive and while there is now an increased focus on reducing unnecessary antibiotic prescriptions[186], antibiotic resistance is still increasing [187]. Furthermore, development of new antibiotics has slowed down massively due to a combination of increasing cost, lack of projected revenue and intrinsic development difficulty. Most of the new

antibiotics being developed at the moment have only small benefits compared to existing ones[188]. With the widespread diffusion of antibiotic resistant bacteria causing a large number of deaths every year, the need to find alternative or complementary ways to treat these infection is greater than ever. Parallel to traditional antibiotic therapy, researchers are now exploring the possibility of employing bacteriophages or the CRISPR/Cas9 system to either kill the bacteria or re-sensitize them to traditional antibiotics [189][190]. Both these approaches, however, are likely to need prior identification of the pathogen to be tailored to the specific case.

Therapeutic antibodies have been used for several years against a variety of conditions [191], including cancer, autoimmune diseases and bacterial infections such as those from *S.aureus* [192]. In the case of bacterial infections, it has also been shown that combining a traditional antibiotic treatment with antibodies can increase the efficacy of both in a synergic way [193].

IgGs have been shown to interact with the bacterial outer membrane, inducing permeabilization and resulting in leaking of periplasmic β-lactamases [194][195]. Indeed, antibodies against β-lactamases have been previously found in the serum and the sputum of infected patients [196], which points towards accessibility of β-lactamase to the host immune system.

Under these premises, a panel of antibodies targeting NDM-1 was generated by collaborators via immunization of mice. Of these antibodies, some appear to exhibit inhibitory properties towards NDM-1. We sought to elucidate the molecular details of this interaction to better understand their mechanism of action and their potential to be employed as therapeutics or as diagnostic markers.

4.2 Results

4.2.1 A panel of monoclonal antibodies exhibits inhibitory activity against NDM-1

E.coli expressed NDM-1 with an N-terminal deletion of the first 28 residues (NDM-1 Δ28) was used to immunise mice and, following generation of hybridomas, monoclonal antibodies (mAb) were expressed by a private company (ENVIGO, Indianapolis, USA). Selected mAb clones were used to carry out assays to determine the effect of these antibodies against NDM-1 and some of its variants. An *in vitro* assay using Nitrocefin was first performed to screen the available antibodies. Nitrocefin is a chromogenic cephalosporin: upon hydrolysis it develops a red colour that can be readily detected by measuring the absorbance at 500 nm [197].

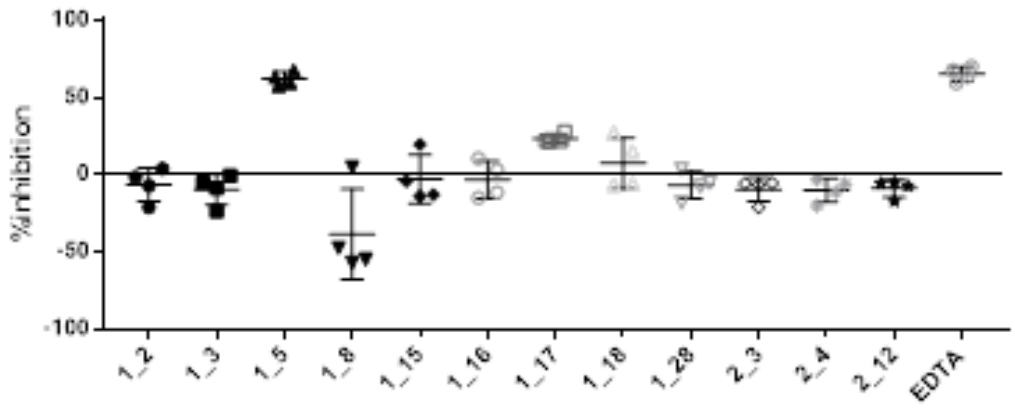


Figure 50: *In vitro* Nitrocefin hydrolysis assay. Each individual replicate (n=4) is plotted together with the median and interquartile range. % inhibition has been calculated relative to the enzyme alone and to the no enzyme control. This figure was provided by our collaborators.

Of the whole panel, antibodies 1_5, 1_8 and 1_17 (from here antibodies 5, 8 and 17) exhibited a certain level of activity, with antibodies 5 and 17 inhibiting the activity of NDM-1 and antibody 8 seemingly enhancing it (**Figure 50**). Focus was then shifted on the most promising antibodies (5 and 17), testing their activity

against NDM-1 and NDM-6. In both cases, both antibodies appear to show an inhibitory effect, more evident in the case of NDM-1 (**Figure 51**).

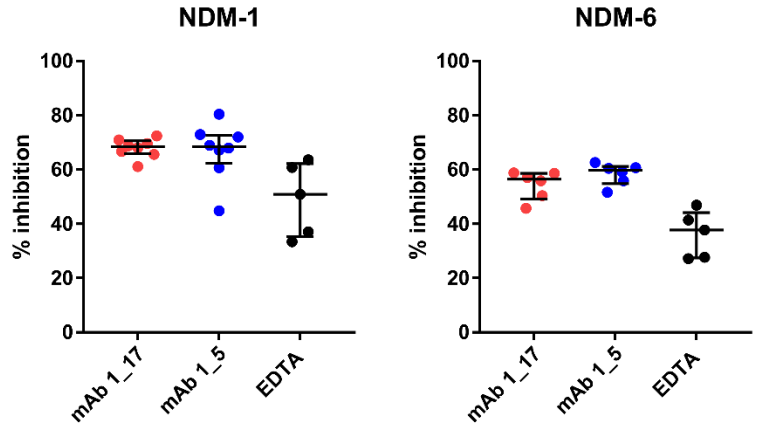


Figure 51: In vitro Nitrocefin hydrolysis assay. Each individual replicate is plotted together with the median and interquartile range. % inhibition has been calculated relative to the enzyme alone and to the no enzyme control. Left panel: activity against NDM-1. Right panel: activity against NDM-6. This figure was provided by our collaborators. At least 5 replicates were performed for each sample.

A final set of *in vitro* bactericidal assays was performed by growing *K.pneumoniae* cultures in the presence (or absence) of serum containing the mAbs, then streaking the cultures on agar plates and counting the colonies following overnight incubation.

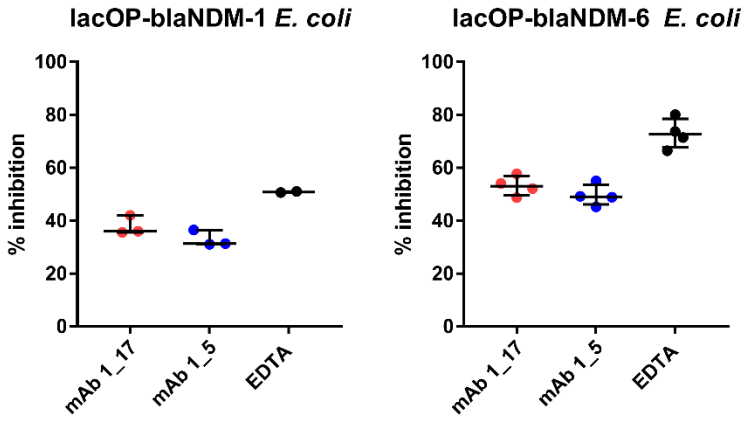


Figure 52: : *In vitro* bactericidal assay. Each individual replicate is plotted together with the median and interquartile range. % inhibition has been calculated relative to the bacteria alone. Left panel: activity against *E.coli* expressing NDM-1. Right panel: activity against *E.coli* expressing NDM-6. This figure was provided by our collaborators. At least two replicates were performed for each sample.

In both assays, an inhibitory effect on the growth of NDM-1 and NDM-6 expressing cells can be seen. In conclusion, it appears that mAb 5 and 17 have an inhibitory activity on NDM-1 and NDM-6.

4.2.2 Cloning, Expression and purification of NDM-1

Full-length NDM-1 and N-terminal truncations of the first 28 and 42 residues (NDM-1Δ28 and NDM-1Δ42, respectively) were subcloned into pOPINF and expressed in *E.coli*. Additionally, NDM-1 Δ28 was subcloned into pOPINFB to allow biotinylation of the final product to use in subsequent experiments. His-tagged NDM-1 was expressed in *E.coli* at 20C for 16-18 hours following addition of IPTG. The overnight culture was harvested by centrifugation, lysed by sonication and centrifuged to remove cell debris. The supernatant was loaded onto a HisTrap FF column (GE Healthcare) and eluted after extensive washes with low concentrations of imidazole. NDM-1 was further purified and the imidazole removed by size exclusion chromatography using a Superdex s75 column.

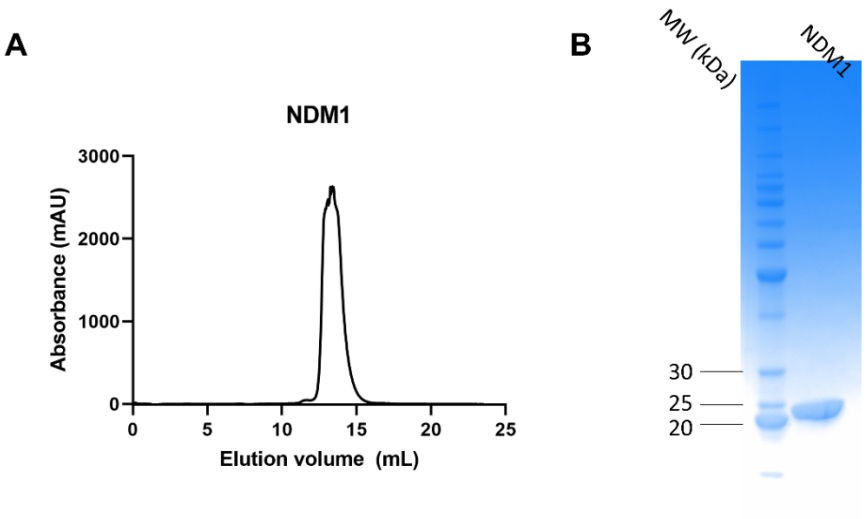


Figure 53: Recombinantly expressed NDM-1. (**A**) SEC profile of NDM-1 on a Superdex 75 10/300 GL column. (**B**) Reducing SDS-PAGE of NDM-1. Principal molecular weight bands are highlighted.

Removal of the N-terminal his-tag was done by addition of 3C protease in a 1:100 protease:target ratio (25C, overnight). Uncleaved protein and remaining tags were removed via reverse affinity chromatography, prior to a second SEC, yielding approximately 60mg of pure protein per litre of bacterial culture, assessed by SDS-PAGE (***Figure 53***).

4.2.3 NDM-1 is bound by Fabs 5 and 17

For preliminary experiments, antibodies 5 and 17 were purified from ascites fluid, then processed to Fabs using the Pierce Fab preparation kit (Thermo Fisher). As obtaining sufficient amounts of both antibodies from the ascites fluid expression was unreliable it was decided to switch to recombinant Fab expression using sequences from hybridoma clones. When the Fab fragments were recombinantly expressed, HEK293T cells were transfected according to the manufacturer instructions. Expression was confirmed by western blotting (***Figure 54***), then the Fab fragments were purified by affinity chromatography and subsequently by size exclusion chromatography. Purity was assessed by SDS-PAGE (***Figure 55***).

The recovered Fabs were incubated with NDM-1 in a 1:2 (Fab:NDM-1) molar ratio, then loaded onto a Superdex 75 column for size exclusion chromatography. The protein complex migrates at about 11mL of elution volume, followed by a smaller peak containing the excess of NDM-1 (***Figure 56*** and ***Figure 57***). Attempts to co-incubate NDM-1 with multiple Fabs (i.e. Fab 5 and Fab 17) were deemed unsuccessful as the size exclusion chromatography traces were analogous to that of NDM-1 and a single Fab fragment (***Figure 58***).

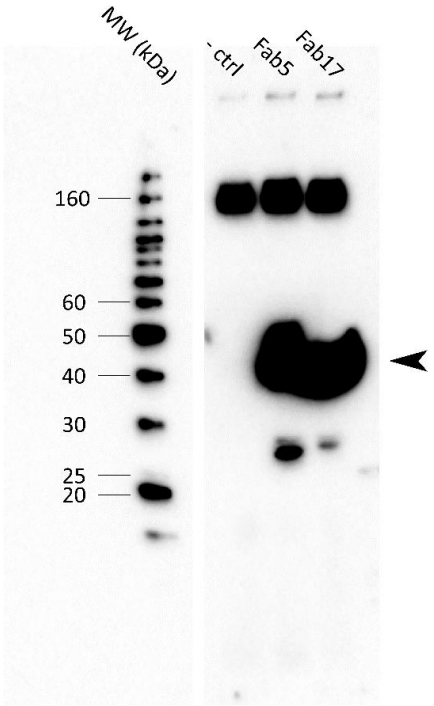


Figure 54: Anti-His western blot of HEK293T supernatant (non-reducing conditions). High molecular weight bands are aspecific interactions, while faint bands at low molecular weight are likely an excess of his-tagged heavy chain. Bands corresponding to Fab 5 are highlighted by a black arrowhead.

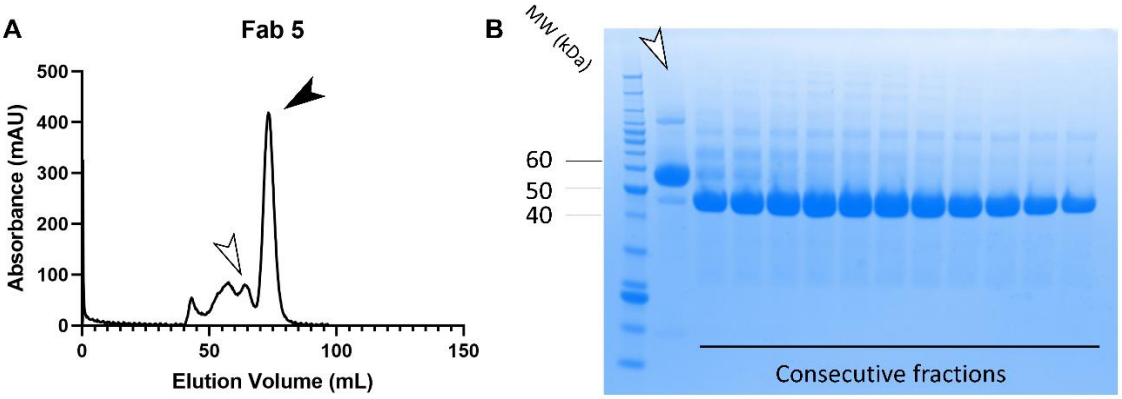


Figure 55: Recombinantly expressed Fab 5. (Left) SEC profile of Fab 5 on a Superdex 75 16/60 column. (Right) Non-reducing SDS-PAGE of recombinant Fab 5. The first fraction (indicated by a white arrowhead) represents some contaminant proteins, followed by consecutive fractions from the main peak (black arrowhead).

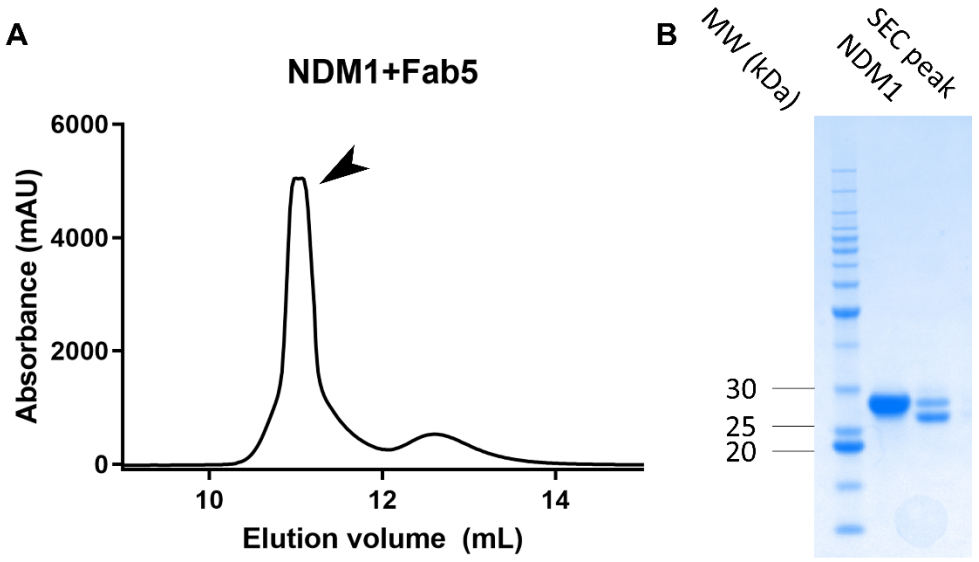


Figure 56: (Left): SEC profile of the NDM-1+Fab5 complex. (Right): SDS-PAGE of the run. The "SEC peak" lane refers to the peak highlighted in panel A by a black arrow. The column used was a Superdex S75 10/300 GL.

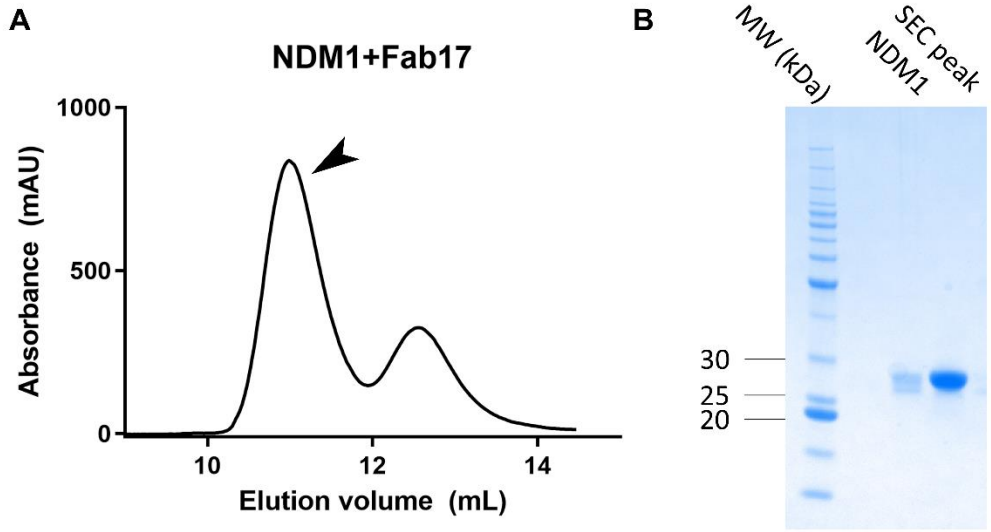


Figure 57: SEC profile of the NDM-1+Fab17 complex (Right): SDS-PAGE of the run. The "SEC peak" lane refers to the peak highlighted in panel A by a black arrow. The column used was a Superdex S75 10/300 GL.

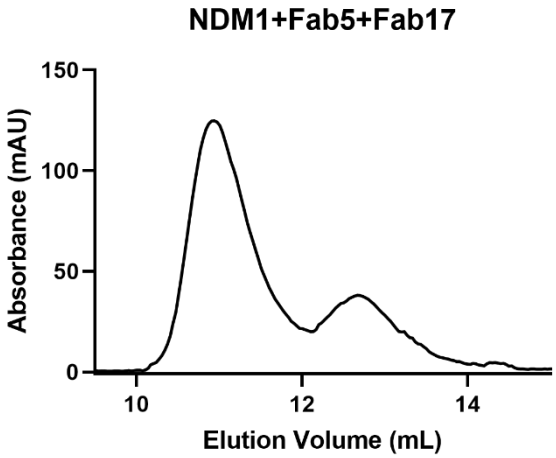


Figure 58: SEC profile of NDM-1 incubated with both Fab 5 and Fab 17complex. The column used was a Superdex S75 10/300 GL.

Taken together, these data show that both Fabs are capable of binding NDM-1 and forming a stable complex, but as binding of multiple Fabs at the same time was not possible, data suggests that the Fabs target the same (or a slightly different) epitope.

4.2.4 NDM-1 forms a complex with palmitic acid

Purified NDM-1 was also used as a putative negative control in other (unrelated) experiments. Oddly, NDM-1 seemed to behave in a vastly different way than expected. NDM-1 was used as negative control in an ELISA assay aimed at showing binding between Dengue virus and GCTL-3, a mosquito C-type lectin that binds the virus, promoting its entry in the cell [198]. Unexpectedly, it appears (**Figure 59**) that NDM-1 binds consistently better than GCTL-3, sometimes at levels comparable to those of 4G2, a control antibody known to bind and neutralise Dengue virus [199].

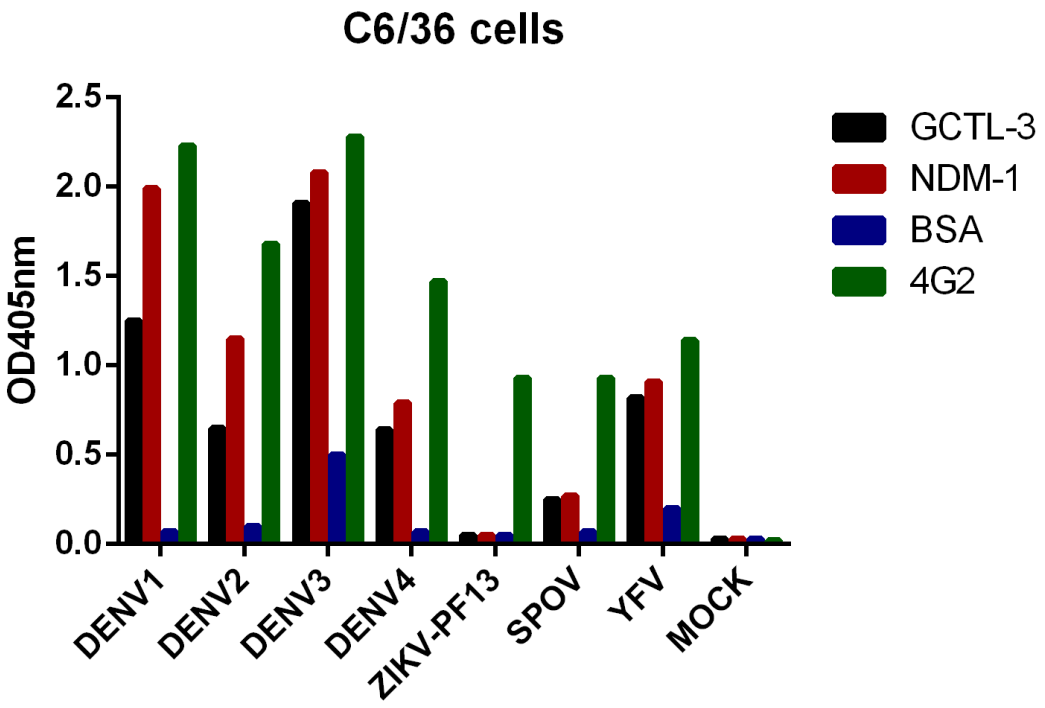


Figure 59: ELISA binding assay to different types of flaviviruses grown in C6/36 cells. This was a test experiment carried out by our collaborators, no replicates have been carried out. DENV1-4: Dengue virus, ZIKV: Zika virus, SPOV: Spondweni virus, YFV: yellow fever virus, MOCK: mock assay

While this data looks interesting, it is hard to speculate what a possible mechanism for this binding may be. However, NDM-1 was also used as a negative control in an unrelated BioLayer Interferometry (BLI) experiment.

Briefly, the experiment was originally meant to investigate the binding of Dally-like protein (Dlp) to a palmitic acid immobilized on a BLI tip via a peptide, as detailed in [38]. The experiment showed that NDM-1 is capable of binding the palmitic acid at a comparable level of that of Dlp (**Figure 60**). This case presented a simpler system, where the specific NDM-1 interactor is known (palmitic acid). While full-length NDM-1 is predicted to be s-palmitoylated [167], nothing has been reported on NDM-1 binding any lipid molecule. Furthermore, the truncated version used in these experiments lacks any sort of lipidation signal.

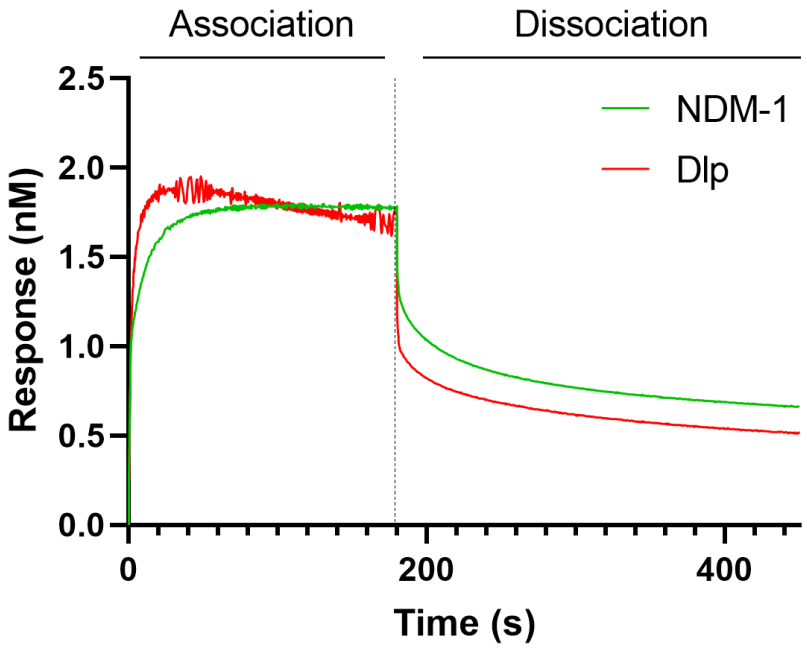


Figure 60: BLI curves of NDM-1 (12.5 µM) and Dlp (9 µM) binding an palmitoyl moiety immobilised on a streptavidin BLI tip via biotinylation. Following subtraction of the baseline values, only the curves pertaining to the association and the dissociation of each analyte are shown. As this was a serendipitous finding, a full dilution series of NDM-1 was not performed and the result should be taken qualitatively.

To further confirm the veracity of this interaction, NDM-1 (6 µM) was incubated with a 20-fold excess of palmitic acid conjugated with BODIPY (a fluorescent dye), then loaded onto a Superdex S200 using a FSEC HPLC system.

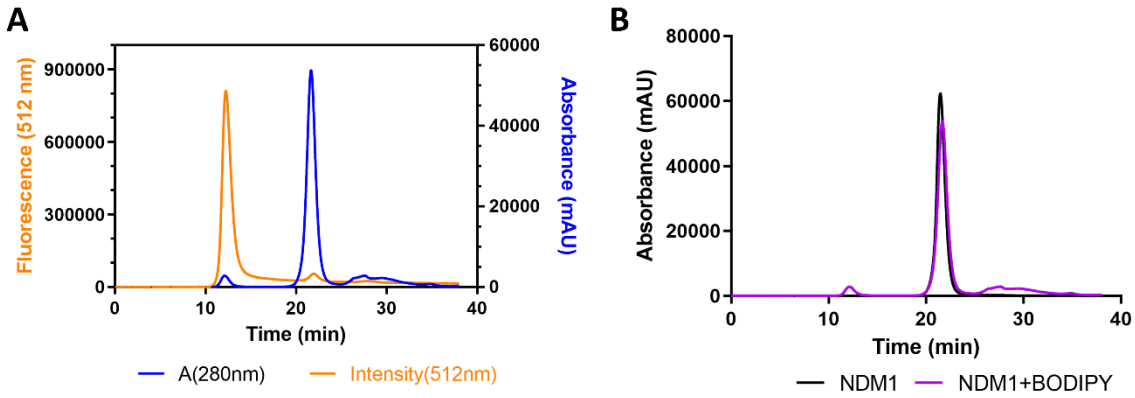


Figure 61: (**A**): FSEC elution profile of NDM-1+palmitic acid-BODIPY. (**B**) SEC profile of NDM-1 with and without palmitic-BODIPY added. The column used was a Superose 6 Increase 3.2/200 and the experiments were performed at a flow rate of 0.08 ml/minute.

In the presence of palmitic-BODIPY, a considerable fluorescence signal is detected at a higher molecular weight than that of a monomer, but not in the column void (**Figure 61A**, orange trace), followed by a smaller peak. This second peak is also detected with NDM-1 only, potentially being a low level of autofluorescence. At 280nm, a very small peak is detected in correspondence of the fluorescent one (**Figure 61B**), indicating that NDM-1 forms complexes with palmitic-BODIPY, potentially including multiple NDM-1 molecules.

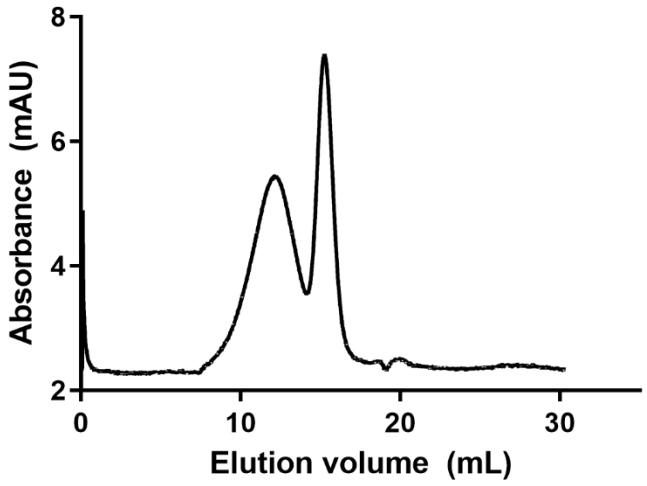


Figure 62: SEC elution profile of purified, crystal-grade NDM-1 incubated with palmitic acid on a Superdex S200 Increase 10/300 GL column.

To exclude that this effect may be due to (or influenced by) the BODIPY moiety of the lipid used, purified, crystal-grade NDM-1 was incubated with palmitic acid diluted in DMSO, then loaded onto a Superdex S200 column for size exclusion chromatography. A second peak at an earlier elution volume appears (**Figure 62**). This peak would correspond to a molecular weight of ~100-150kDa, consistent with the results of the BODIPY experiment.

4.2.5 Crystallization, data collection, structure solution and refinement of Fab 5 and 17, complexes with NDM-1

Crystallization of both single Fabs and their complexes with NDM-1 proved to be significantly more challenging than anticipated. The complexes were concentrated to different concentrations depending on the buffer composition, and assessed by a pre-crystallization test. Final concentration of the protein solution ranged from 10 to 60mg/mL. This is consistent with reported concentrations for structures available in the PDB of NDM proteins. Crystals selected for diffraction experiments were cryoprotected by a short incubation in their respective mother liquor, supplemented with 25% glycerol, then cryocooled by submerging them in liquid nitrogen. **Table 7** recapitulates the conditions for the various crystals.

Crystallization of the NDM-1-Fab 5 complex was possible and small needles or needle clusters were grown within 30 to 60 days in a broad variety of conditions. Optimisation of these crystals was performed by diluting conditions with hits and varying the concentrations of reagents in the mother liquor. Seeding was also attempted using seeds from successful initial hits, as well as screens with additives. Overall, after the challenging optimisation, only some slightly larger needles could be obtained. While most of these crystals did not diffract, a few of them did so, albeit only to ~9Å resolution.

From trays with NDM-1-Fab 5, some crystals of Fab5 alone grew after ~344 days and several datasets were collected. Diffraction experiments were carried out at Diamond Light Source, Didcot (UK) at the I03 beamline. Automatic processing of the diffraction images was done using the autoPROC software [32], to a final resolution of 1.9Å.

Construct	Condition	Days to growth	Image
Fab 5	20 %w/v PEG 3350 0.200 M KCl	344	
Fab 5	2.000 M Ammonium Sulphate 0.100 M Sodium Acetate pH 4.6	344	
NDM-1+Fab 17	0.1M Na-HEPES pH 7.0 20% w/vPEG 8000	238	
NDM-1+Fab 5	Multiple conditions in the ProPlex screen	28-132	
NDM-1+Fab 5	Optimisation of the ProPlex B4 condition	28-132	
NDM-1+Fab 5	Multiple conditions in the PEGRx screen	28-132	

Table 7: Examples of crystals of NDM-1 in complex with Fab5 or Fab17 and crystals of Fab 5.

Initial quality check steps using the Xtriage software (part of the Phenix suite [33]), detected significant translational non-crystallographic symmetry (tNCS). Structure solution was obtained via molecular replacement using PHASER [34]. The search model used in the process was that of a mouse monoclonal antibody (PDB id: 2VXT). The structure was divided into constant and variable regions, which were searched independently to account for possible flexibility. Within the variable regions, CDR loops were deleted to eliminate sources of potential bias. Model building and refinement were carried out through several cycles of manual

building using COOT [35] and refinement with Phenix.refine [33]. Restraints used during refinement were: cartesian, individual B-factor, TLS and torsion-angle NCS. Data collection and refinement statistics are given in **Table 8**. Unfortunately, R-factors were relatively high for a structure of this resolution (22/26%), but this is not entirely unexpected, as the crystal shows a high degree of translational non-crystallographic symmetry tNCS, as detected by phenix.xtriage.

In the case of the complex between NDM-1 and Fab 17, plate crystals could be obtained after ~238 days. Diffraction experiments were performed at Diamond Light Source, Didcot (UK) at the I04 beamline. Automatic processing of the diffraction images was done using the Xia2 [31] and DIALS [200] softwares, to a final resolution of 3.2Å. Molecular replacement was carried out as described for Fab 5, except for the addition to the search model of NDM-1 (PDB id: 3ZR9). The addition of full-length NDM-1 did not yield a suitable solution, whereas a truncation of it (from residue 50 to 140, determined by trial-and-error) was successful. Model building and refinement were carried out through several cycles of manual building using COOT [35] and refinement with Phenix.refine [33]. Restraints used during refinement were: cartesian, individual B-factor and TLS. Data collection and refinement statistics are given in **Table 9**.

As of writing, it has been impossible to grow crystals of Fab 17 alone.

For NDM-1 in complex with palmitic acid, some crystals appeared after 397 days. None of these crystals, however, diffracted.

	Fab 5
Data Collection	
Wavelength (Å)	0.9795
Resolution range (Å)	80.04 – 1.91 (1.94 – 1.91)
Space group	*C 2*
Cell dimensions	
a, b, c (Å)	161.57, 65.93, 106.5
α, β, γ (°)	90, 130.95, 90
Total reflections	460280 (23449)
Unique reflections	65764 (3264)
Multiplicity	7.0 (7.2)
Completeness (%)	99.9 (100.0)
Mean I/σ(I)	6.4 (0.6)
R-merge	0.192 (3.366)
CC ½	0.998 (0.314)
Refinement	
R_{work} / R_{free}	22.1/26.4
Number of atoms	
Protein	6616
Ligand	12
Solvent	478
RMSD(bonds)	0.004
RMSD(angles)	0.7
Ramachandran plot	
Favoured (%)	98.35
Allowed (%)	1.65
Outliers (%)	0.00
Average B-factor	
Protein	42.12
Ligand	52.58
Solvent	40.69

Table 8: Data collection and refinement statistics for Fab 5. Statistics for the highest-resolution shell are shown in parentheses. The ligand atoms refer to the two glycerol molecules that were built into the electron density map.

NDM-1-Fab 17	
Data Collection	
Wavelength (Å)	0.9795
Resolution range (Å)	50.03 – 3.21 (3.26 – 3.21)
Space group	*C 1 2 1*
Cell dimensions	
a, b, c (Å)	70.44, 71.09, 122.78
α, β, γ (°)	90, 91.31, 90
Total reflections	265351 (9474)
Unique reflections	10034 (455)
Multiplicity	26.4 (20.8)
Completeness (%)	99.6 (92.3)
Mean I/σ(I)	5.2 (0.4)
R-merge	1.440 (34.224)
CC ½	0.993 (0.288)
Refinement	
R_{work} / R_{free}	23.1/27.7
Number of atoms	
Protein	4448
Solvent	583
RMSD(bonds)	0.004
RMSD(angles)	0.96
Ramachandran plot	
Favoured (%)	96.88
Allowed (%)	2.95
Outliers (%)	0.17
Average B-factor	
Protein	97.03

Table 9: Data collection and refinement statistics for Fab 17. Statistics for the highest-resolution shell are shown in parentheses

4.2.6 SEM imaging of NDM-1-Fab 5 crystals

With crystals of the NDM-1-Fab 5 complex diffracting very poorly, it was decided to investigate the possibility to use Diamond Light Source VMXm beamline to attempt data collection under vacuum. Crystals were mounted and plunge-frozen on EM grids to determine sample suitability to diffraction experiments under vacuum. Using a SEM, EM grids were imaged, revealing small, rod-like crystals.

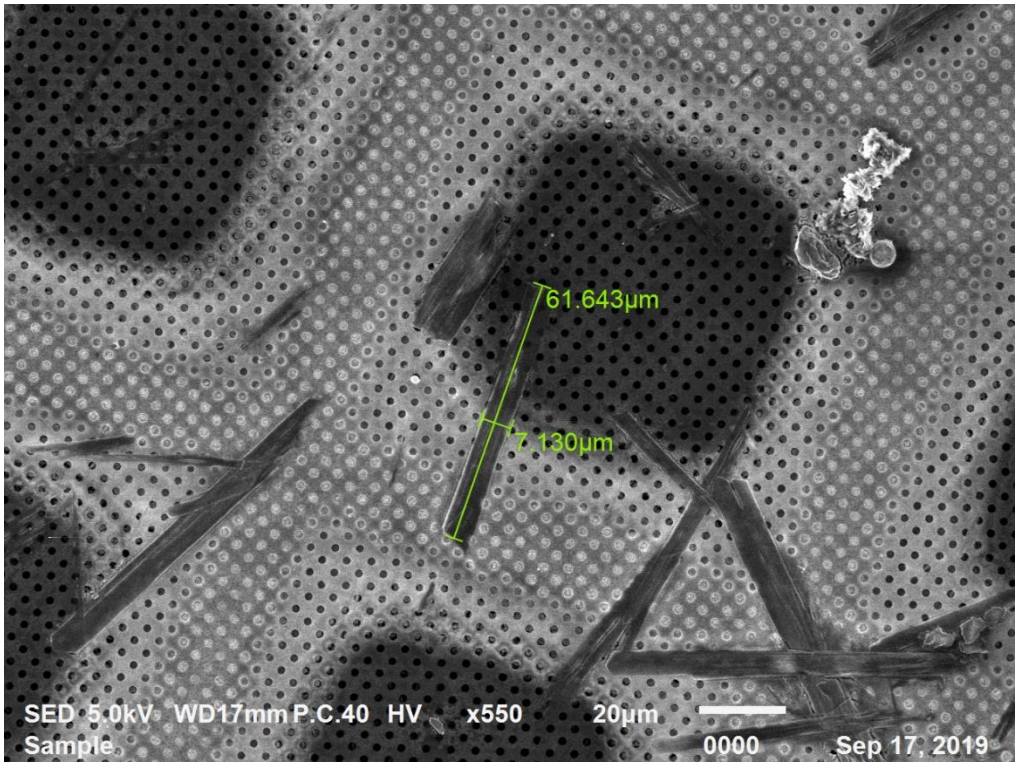


Figure 63: SEM micrograph of NDM-1-Fab 5 crystals plunge-frozen on a cryo-EM grid. Scale bar indicates 20µm.

Interestingly the crystals appeared very thin, and the holes in the carbon layer are visible through the crystal (**Figure 63**). It appears that the sample may be suited for experiments of Microcrystal Electron Diffraction (MicroED) [201]. As of writing, however, no MicroED experiments have been performed.

4.2.7 Nanobodies

In order to explore other possible therapeutic venues, as well as aiding in the crystallization of NDM-1, an attempt was made at raising nanobodies against NDM-1. Purified NDM-1 Δ28 was used to immunise a llama, and screening for nanobodies was done with biotinylated NDM-1. However, despite eliciting an immune response in the llama, no positive clones could be identified following two rounds of panning against NDM-1 using two different libraries. These experiments were carried out Dr. Jiandong Huo at the Rosalind Franklin institute (Harwell campus, Didcot, UK).

4.2.8 Sample preparation, data collection and processing of the NDM-1-Fab 5 complex by Cryo-EM

As obtaining well diffracting crystals for the complex between NDM-1 and Fab 5 appeared unlikely, the possibility of obtaining the structure via single particle Cryo-EM was explored. The gel-filtered complex of NDM-1-Fab 5 (see section **4.2.3**) was plunge frozen onto glow-discharged grids using a Vitrobot mark IV (FEI) system.

Initial screening for quality control was done on a 200kV Glacios Cryo-TEM, whereas full data collection on a 300kV Titan Krios Cryo-TEM.

A workflow of the strategies employed for the processing of the data is shown in **Figure 66**. Data processing was initially carried out using cryoSPARC. The collected movies were motion-corrected and a Contrast Transfer Function (CTF) estimated for all the micrographs.

Several strategies were attempted to pick particles: for blob picker, minimum and maximum diameter were estimated based on the crystal structure of the NDM-1 Fab 17 complex. For template picker, the structure of the same complex was used as template and for the Topaz picker, particles from the final classes of the blob picker pipeline were used to train the model. Picked particles from all three strategies were separately extracted, downsampled and subjected to subsequent rounds of 2D classification. 2D classes were compared with reference 2D classes generated from the crystal structure of the NDM-1 Fab 17 complex (**Figure 64B** and **C**) Once the classes were deemed satisfactory, per-particle motion correction was carried out on particles from the selected classes.

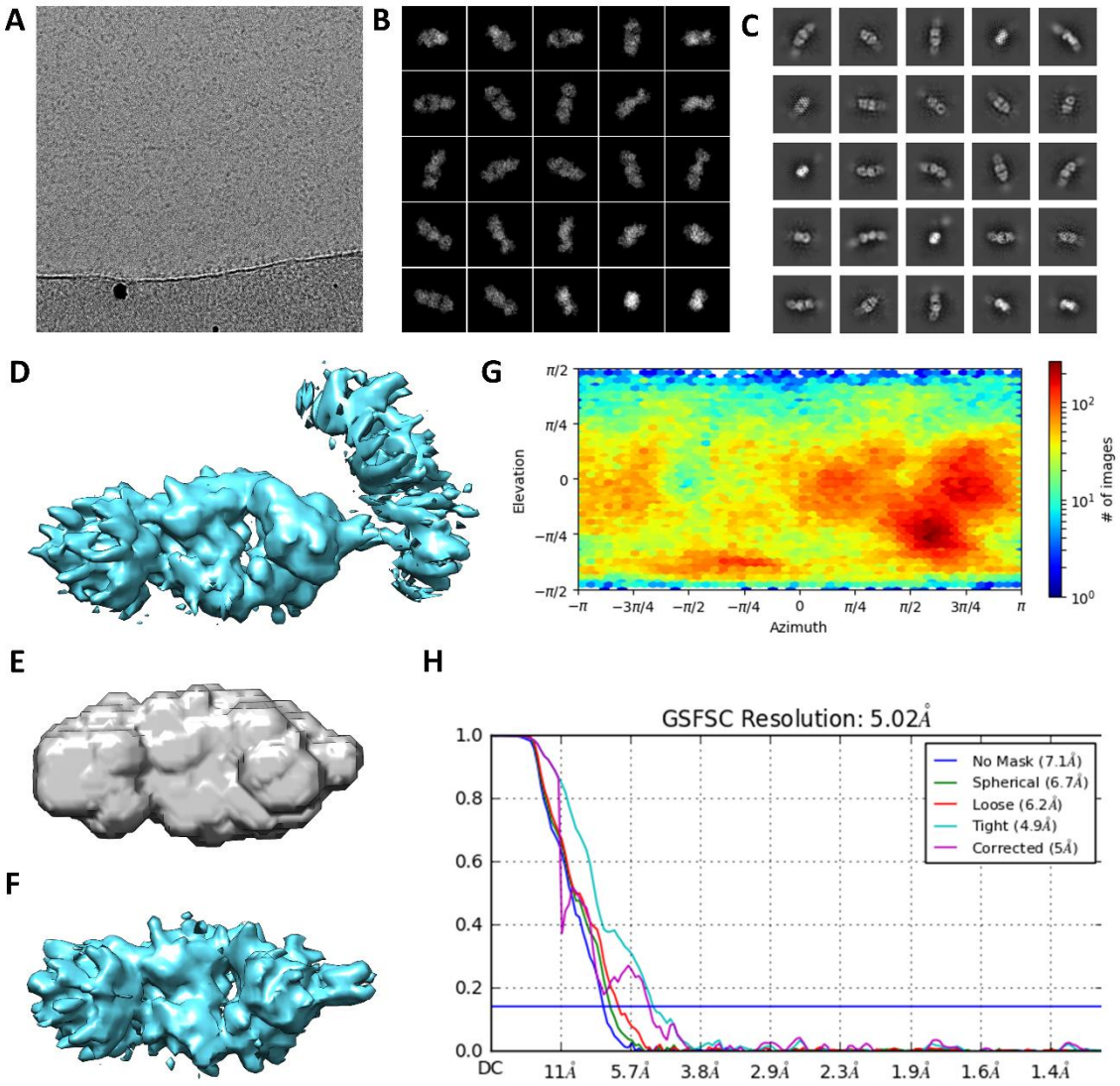


Figure 64: Data processing of the NDM-1-Fab 5 complex in cryoSPARC [21]. (**A**) Sample micrograph of the complex. (**B**) Reference 2D classes generated in cryoSPARC from the NDM-1-Fab 17 complex. (**C**) Final selected 2D classes. (**D**) Density including edge noise (**E**) Mask applied to eliminate noise (**F**) Final sharpened cryoSPARC map (**G**) Angular distribution of the particle projections following non-uniform refinement. (**H**) FSC curves following non-uniform refinement

Several ab-initio models were generated and a separate heterogeneous refinement carried out. Particles from the model with the highest reported resolution (7.11Å, ~118000 particles) were used as input for a final non-uniform refinement to improve the resolution. Noise at the edge of the of the density was removed by modifying the automatically generated mask (**Figure 64D**, **E** and **F**) in Chimera. The final reported resolution was 5.02Å (**Figure 64H**), however this is most likely an overestimation, as the map appears to have significantly less features

than expected for this resolution. This may be due to incorrect processing of noisy high resolution components in the data, but a definitive answer cannot be provided with absolute certainty.

The particles from this dataset were exported to RELION and used to perform a 3D classification. Manual inspection of the obtained classes did not reveal significant differences and thus all the particles were used for a final round of gold-standard 3D refinement. Following post-processing, the reported resolution was 7.6 Å (**Figure 65**). Final parameters for the data collection are reported in **Table 10**.

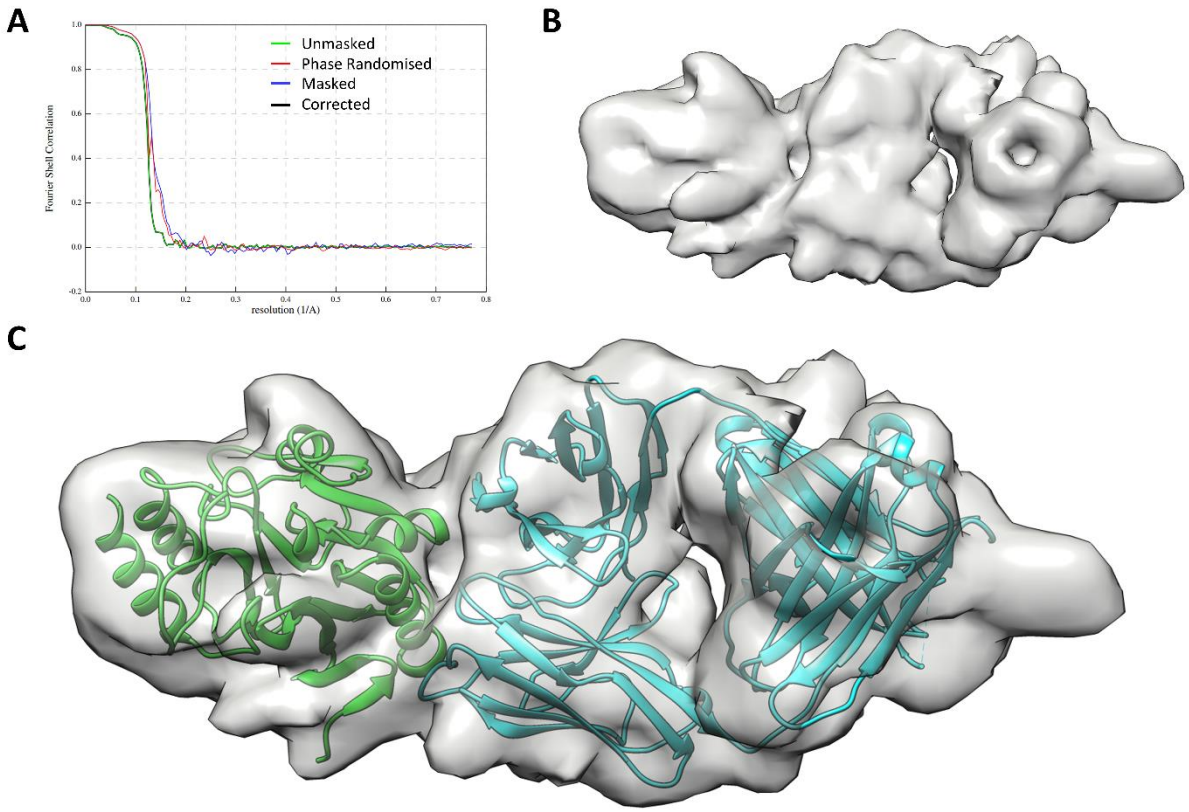


Figure 65: Data processing in RELION. (A) FSC curves after gold-standard refinement in RELION. (B) Final map after gold-standard refinement in RELION. (C) Crystal structures of NDM-1 (PDB id: 3ZR9) and Fab 5 fitted into the final map

NDM-1-Fab 5	
Data Collection	
Magnification	240000x
Voltage (kV)	300
Electron exposure ($e^-/Å^2$)	60
Defocus range (μm)	-1.2 to -4.2
Pixel size (Å)	0.648
Symmetry imposed	C1
Initial particles images (no.)	3912328
Final particles images (no.)	118874
Map resolution (Å)	7.6
FSC threshold 0.143	
Map resolution range (Å)	6.07-9.28

Table 10: Data collection statistics for the NDM-1-Fab 5

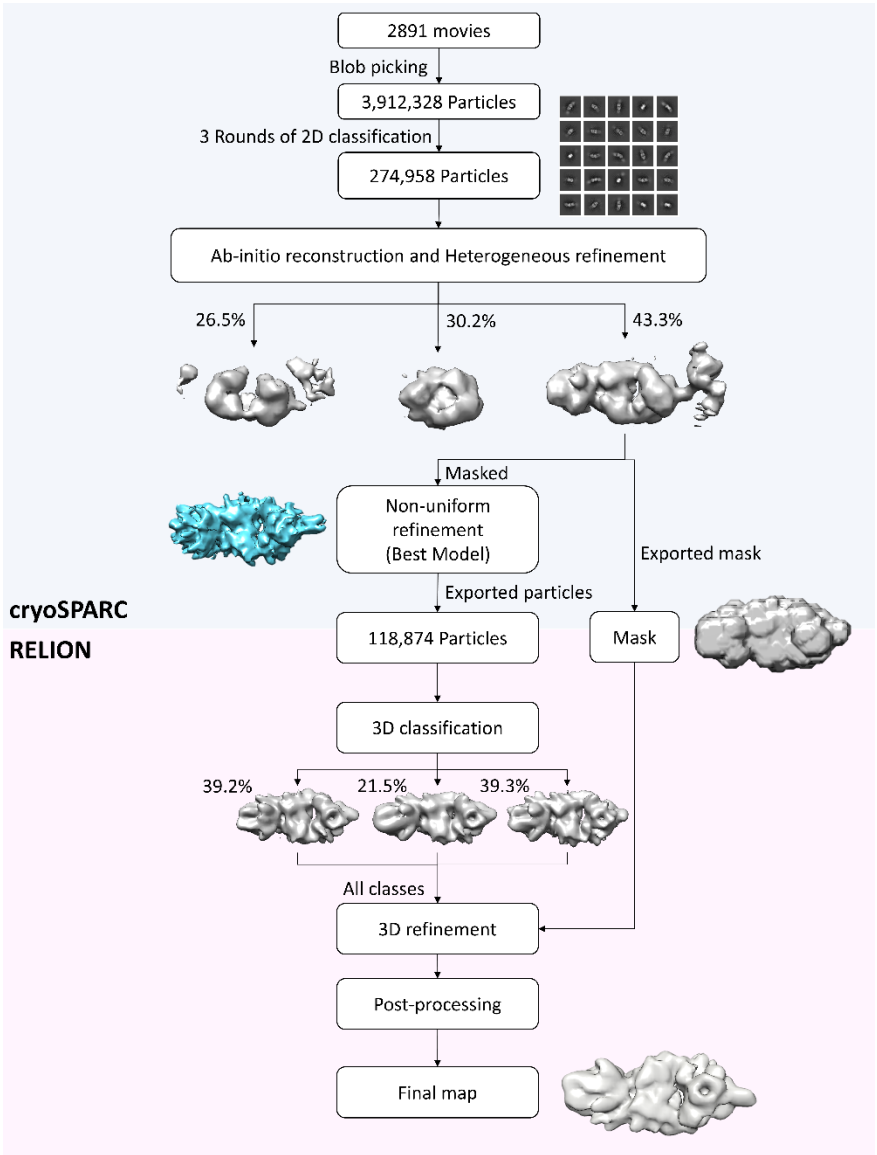


Figure 66: Single particle Cryo-EM workflow. Portions of the workflow carried out in cryoSPARC and RELION are highlighted in blue and pink, respectively.

4.2.9 Overview of the Fab fragments

Identification of the CDR sequences for both Fab 5 and 17 was done using the international ImMunoGeneTics information system® (IMGT) with the DomainGapAlign program and used to generate the "collier de perles" representation. Results of these analyses are shown in **Figure 67**.

It is worth noting that within the crystal structure of Fab 5, in the proximity of His 202 on the constant region of the light chain, it was possible to identify a region with strong electron density that is not part of the Fab itself (**Figure 68**). In the tNCS related molecule, a similar (albeit less defined) density can be found in the same region. None of the components of the buffers used during the purification or the mother liquor, however, can explain said density. Another hypothesis to explain the density may be related to a sugar. Fab fragments are known to be glycosylated and the roles of glycosylation are diverse, such as diversification, elimination of auto-reactivity, modification of the activity and more ([202], also reviewed in [203]).

In the proximity of this site, however, no asparagine residues are available for N-linked glycosylation. In fact, there appears to be a continuous density between His 202 and the unknown region. Attempts at automatic identification of the ligand using phenix.ligands_identification were also not successful. As a satisfactory interpretation of this region of the electron density could not be reached, it was decided to leave the region empty when building the model.

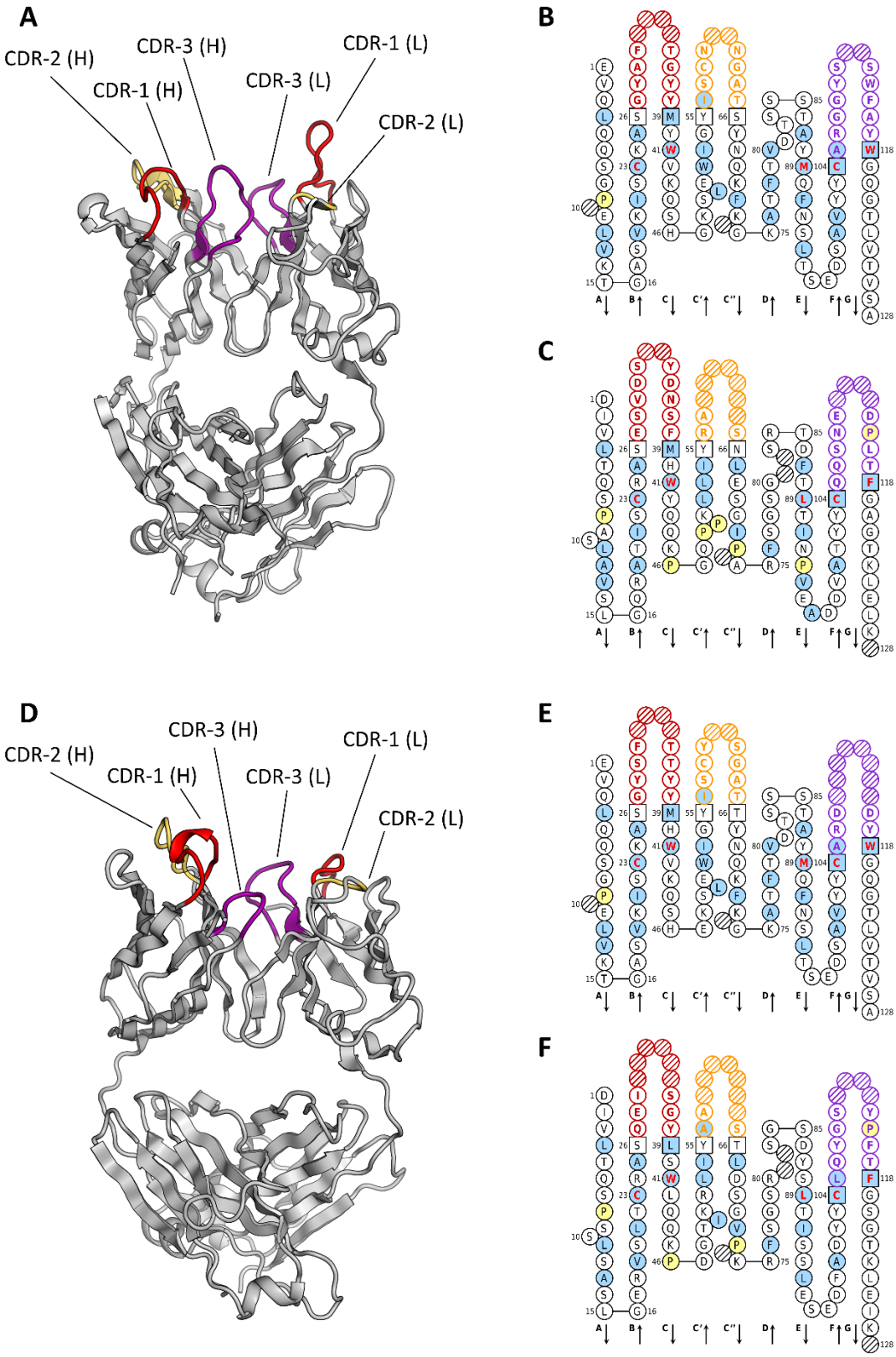


Figure 67: Overview of Fab 5 and 17. (**A**) Structure of Fab 5 (**B**) Organisation of the variable domain of Fab 5 heavy chain. (**C**) Organisation of the variable domain of Fab 5 light chain. (**D**) Structure of Fab 17 (**E**) Organisation of the variable domain of Fab 17 heavy chain. (**F**) Organisation of the variable domain of Fab 17 light chain. In the "collier de perles" representation, generated by the IGMT DomainGapAlign program, blue circles represent hydrophobic residues (in ≥ 50% of variable domains in the database). Red letters are conserved residues and empty circles are gaps when compared to the standard IMGT numbering scheme.

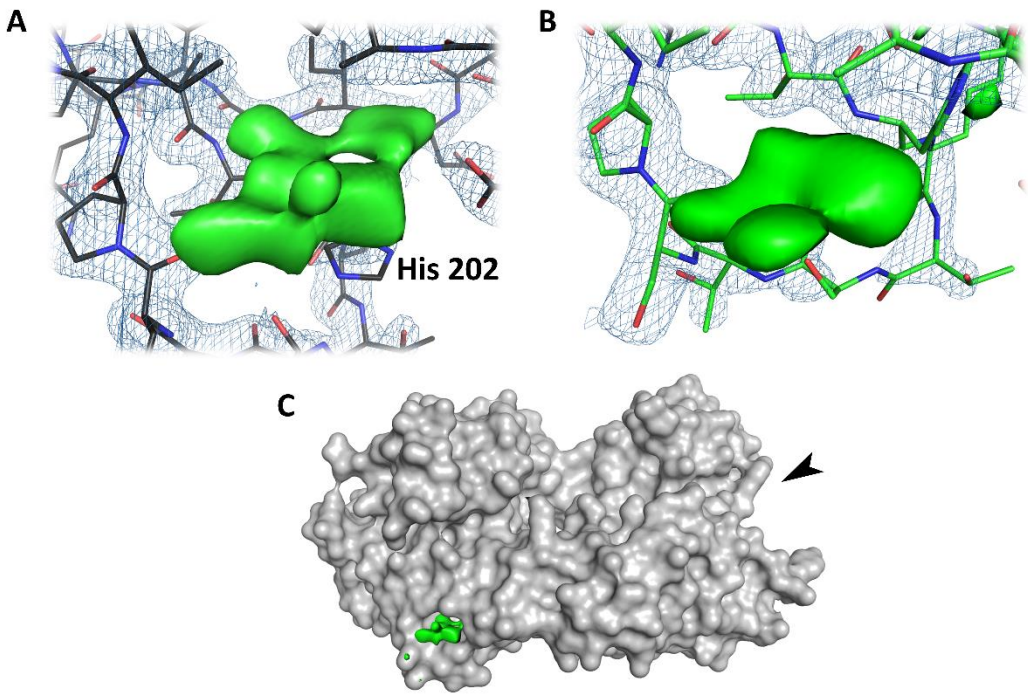


Figure 68: Unidentified density in the crystal structure of Fab 5. (**A**) Overview of the most defined density. Electron density for the neighbouring residues is in wireframe representation from a 2Fo-Fc map (σ 1.2). The unidentified density is in surface representation from a Fo-Fc map (σ 3.5). (**B**) Overview of the least defined density. Electron density for the neighbouring residues is in wireframe representation from a 2Fo-Fc map (σ 1.2). The unidentified density is in surface representation from a Fo-Fc map (σ 2.0). (**C**) The unidentified density sits in the constant region of the light chain, highlighted in green. The variable region of the Fab is indicated by a black arrowhead

4.2.10 The structure of the NDM-1-Fab 17 shows an unexpected binding site

The feature of the crystal structure of the complex between NDM-1 and Fab 17 that is immediately evident from analysing the structure is that only a portion of full-length NDM-1 is present. Only NDM-1 residues from Pro 41 to Asp 199 can be identified and this results in the absence of a complete active site for the enzyme. A superimposition of the structure of the NDM-1 degraded fragment with full-length NDM-1, however, shows a very similar structure, with a RMSD of 0.678Å (**Figure 69**). It is therefore highly likely that this interaction can recapitulate the interaction between Fab 17 and full-length NDM-1.

The NDM-1 epitope bound by Fab 17 is located on the opposite side of the molecule to the active site in the full-length protein (identified in **Figure 70** by two yellow spheres, modelling the coordinated zinc atoms).

Analysis of the structure with the PISA server [45] identifies a total surface for the epitope of 925.6Å^2 (401.8Å^2 and 523.8Å^2 from the heavy and light chain, respectively. A list of the interactions that contribute to the Fab17-NDM-1 interface is given in **Table 11**. On NDM-1 one residue in particular, Asp 130 appears to form a network of hydrogen bonds with Tyr 33 (CDR1) and Ser 52 (CDR2) on the heavy chain of Fab 17 (**Figure 71**).

Finally, Fab 17 appears to also interact with the so-called Ω-loop (Residues 149-159). The Ω-loop is known in the literature to modulate the activity of NDM-1 towards different substrates and while the specific details of how this happens are unknown, its flexibility is deemed to be important [204].

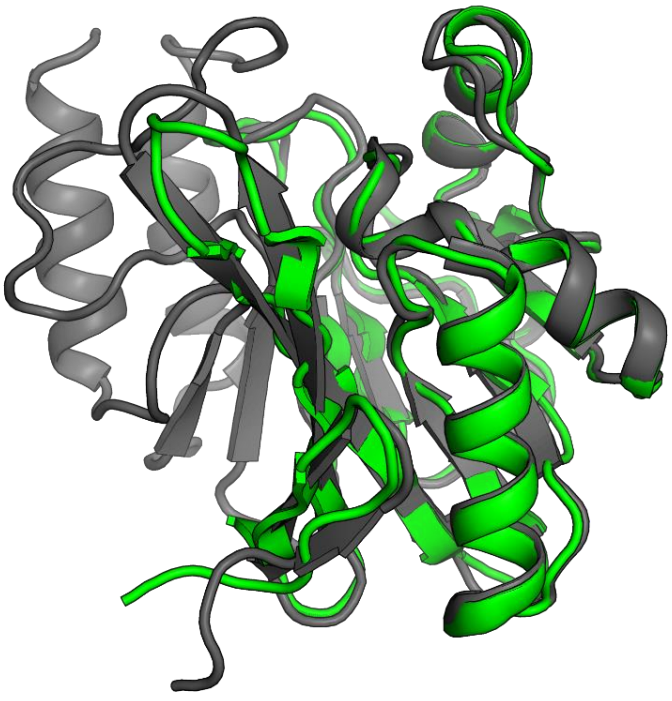


Figure 69: Superimposition of the structures of full-length NDM-1 (grey, PDB id: 3ZR9) and the degraded fragment in the structure of the NDM-1-Fab 17 complex (green).

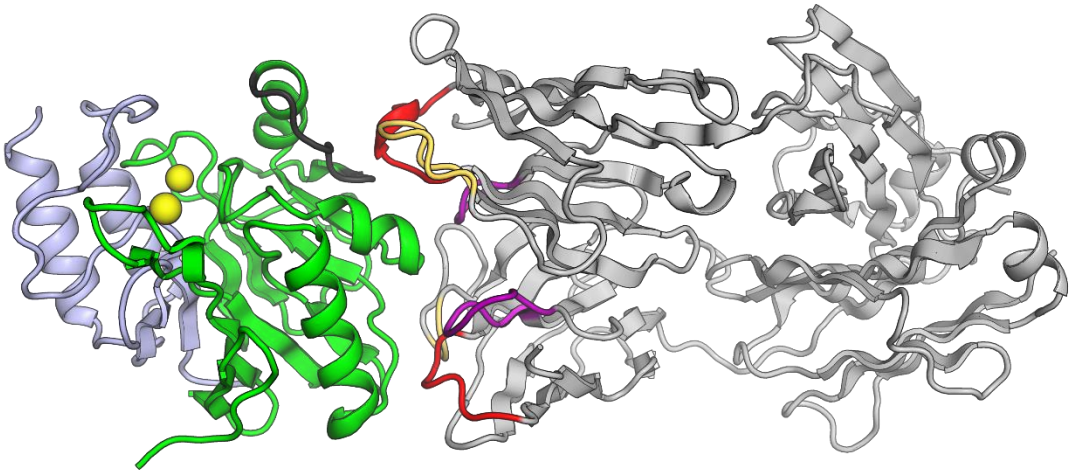


Figure 70: Complex between NDM-1 and Fab 17. Fab 17 is coloured in grey and the CDR loop follow the IGMT colouring scheme (see **Figure 67**). The NDM-1 fragment is coloured in green and the missing portion to complete the full-length chain in blue. The zinc atoms in the active site are modelled based on the structure of full-length NDM-1 (PDB id: 3ZR9). The Ω-loop is shown in black.

NDM-1		**Fab 17**			
No.	**Residue**	**No.**	**Residue**	**Distance (Å)**	**Bond type**
159	HIS [NE2]	32 (H)	TYR [OH]	3.32	Hydrogen bond
130	ASP [O]	33 (H)	TYR [OH]	3.36	Hydrogen bond
130	ASP [OD2]	33 (H)	TYR [OH]	2.22	Hydrogen bond
130	ASP [OD2]	52 (H)	SER 52[OG]	3.00	Hydrogen bond
153	GLY [O]	54 (H)	TYR [OH]	3.44	Hydrogen bond
159	HIS [NE2]	100 (H)	ASP [OD1]	3.87	Salt bridge
106	LYS [NZ]	92 (L)	GLY [O]	3.37	Hydrogen bond
112	PRO [N]	30 (L)	SER [OG]	3.49	Hydrogen bond
138	ALA [N]	49 (L)	TYR [OH]	2.65	Hydrogen bond
83	GLY [O]	52 (L)	SER [OG]	3.85	Hydrogen bond
136	GLY [O]	49 (L)	TYR [OH]	2.92	Hydrogen bond

Table 11: List of polar/ionic interactions between NDM-1 and Fab 17. Residues from Fab 17 are identified by (H) or (L) if they belong to the Heavy or Light chain, respectively. Distance values were calculated by the PISA server.

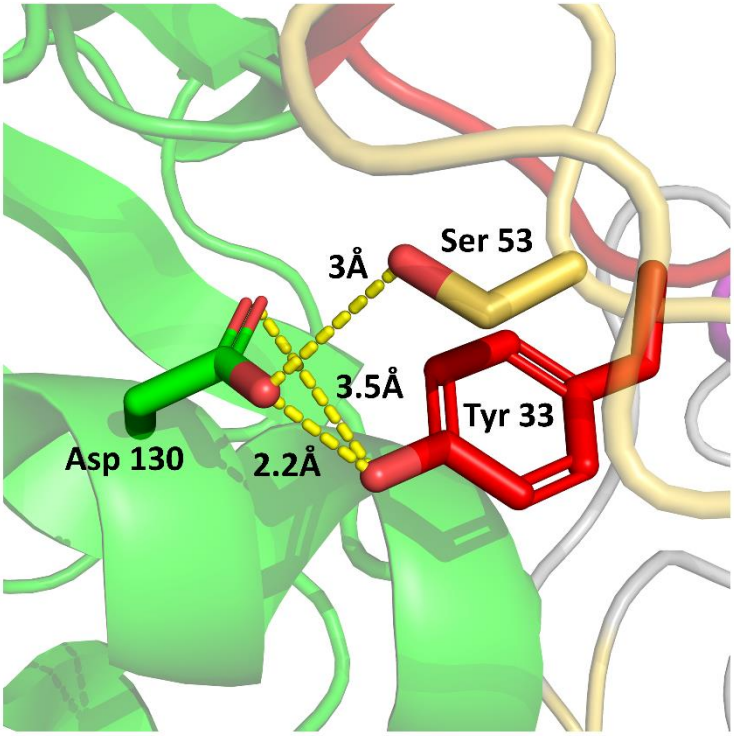


Figure 71: Hydrogen bonds involving NDM-1 Asp 130 as detected by the PISA server. Colouring follows the IGMT scheme.

Within the epitope bound by Fab 17, Asp 130 is the only residue for which natural variants of NDM are known, specifically Asp 130 is mutated in an Asn (NDM-7) or in a Gly (NDM-8). For both mutants Met 154 is also mutated in a Leucine, but that is not part of the epitope.

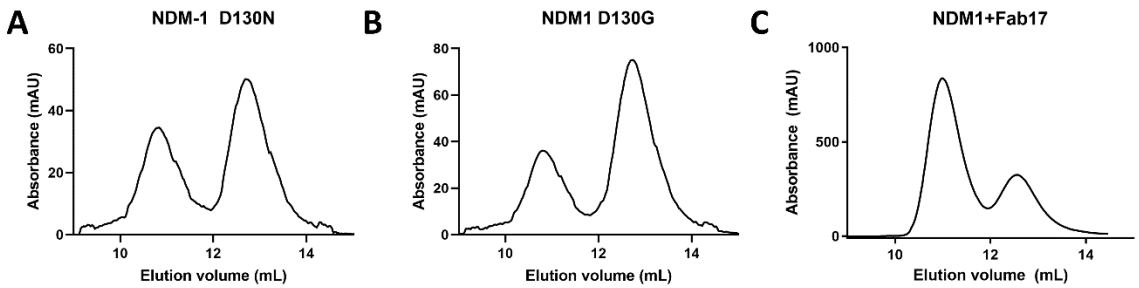


Figure 72: SEC data of NDM-1 mutants bound to Fab 17. A Superdex S200 increase 10/300 GL was used. (**A**) NDM-1 D130N mutant. (**B**) NDM-1 D130G mutant. (**C**) NDM-1 wild-type

Two mutants (NDM-1 D130N and NDM-1 D130G) were generated to verify if Fab 17 can still bind this region if the interactions with Asp 130 are lost. Both mutants were purified as the wild-type and incubated with Fab 17, before loading

the putative complex on a size exclusion chromatography column. As shown in **Figure 72**, it appears that binding of Fab 17 is retained for both mutants.

4.2.11 Fab 5 binds a similar epitope to that of Fab 17 on NDM-1

With the low resolution map obtained from cryo-EM, it was possible to unambiguously fit both a full-length NDM-1 and the crystal structure of Fab 5 within the density (**Figure 73**). This was done in Chimera by using a map simulated from the atoms to 8Å resolution (to approximate the resolution reported by Relion for the experimental map). Correlation values with the map for NDM-1 and Fab 5 were 0.8985 and 0.8589, respectively.

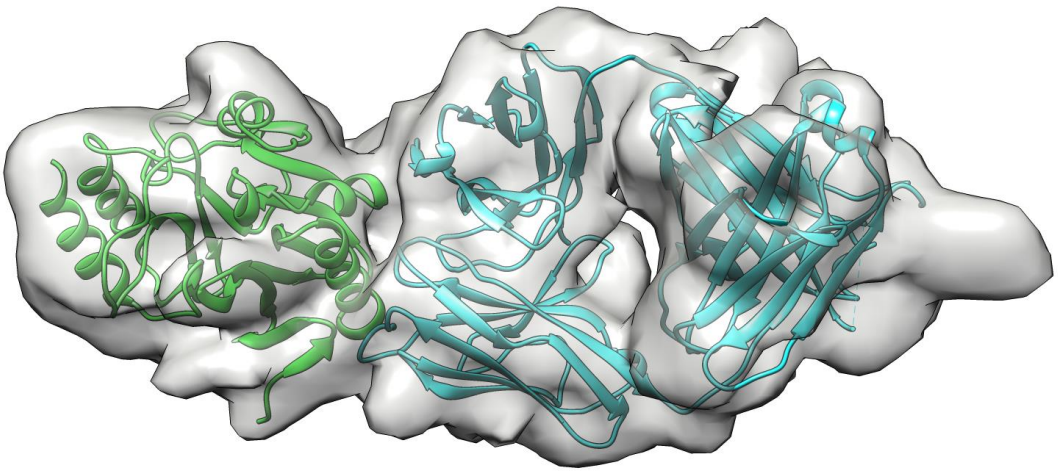


Figure 73: Fitted crystal structures of NDM-1 (PDB id: 3ZR9) and Fab 5. Fitting was performed with the Chimera software.

Similarly to Fab 17, Fab 5 also appears to bind NDM-1 on the opposite side from the active site. It is worth noting, however, that in this fitting there are clashes between the CDR-1 loop of the Fab light chain and one helix (residues 96-108) of NDM-1. This is likely due to the fact that Fab 5 crystallized alone, therefore the conformation of the protein itself that is captured within the crystal is likely to be slightly different to that of the Fab bound to NDM-1. Analysis of this

interface with the PISA server is unlikely to be meaningful, as the coarse fitting of the two crystal structures within the density does not allow correct placement of the residues involved. Despite the coarse fitting, the putative interface between NDM-1 and Fab 5 does not appear to involve either the helix with Asp 130 or the Ω-loop on NDM-1.

4.2.12 Comparison between the two Fabs

By taking the structures of the two complexes and superimposing the two NDM-1 molecules we can observe how the two Fabs relate to each other. The angle formed by the two Fabs is ~60˚ and they are rotated by ~90˚ respective to each other. Looking at the structures, it becomes immediately clear that both Fabs bind a similar epitope on NDM-1 (**Figure 74**).

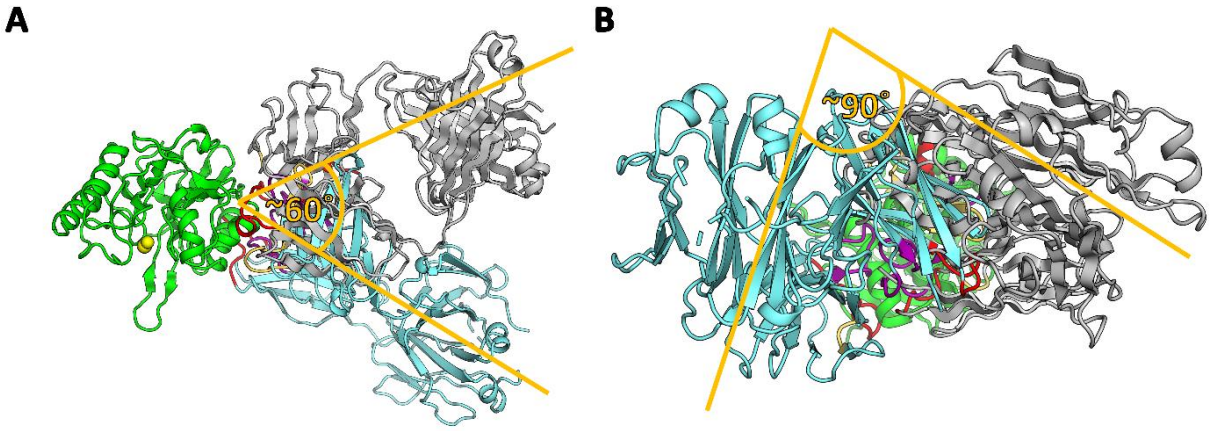


Figure 74: Superimposition between NDM-1-Fabs complexes. NDM-1 is coloured green (yellow spheres represent the zinc atoms coordinated in the active site). Fab 5 and 17 are coloured in grey and light blue, respectively. CDR loops on both Fabs are coloured according to the IGMT colouring scheme. (**A**) Top view. (**B**) Side view. Angles are coloured in yellow and approximate value indicated.

Furthermore, binding of both Fabs to NDM-1 would be impossible due to major clashes that would occur between overlapping regions. This is consistent with size exclusion chromatography data (see **4.2.3**) where only one of the Fabs can be bound at the same time. As previously mentioned, Fab 5, unlike Fab 17, does

not appear to bind the helix with Asp 130 (henceforth referred to as λ-helix) or the Ω-loop (**Figure 75**).

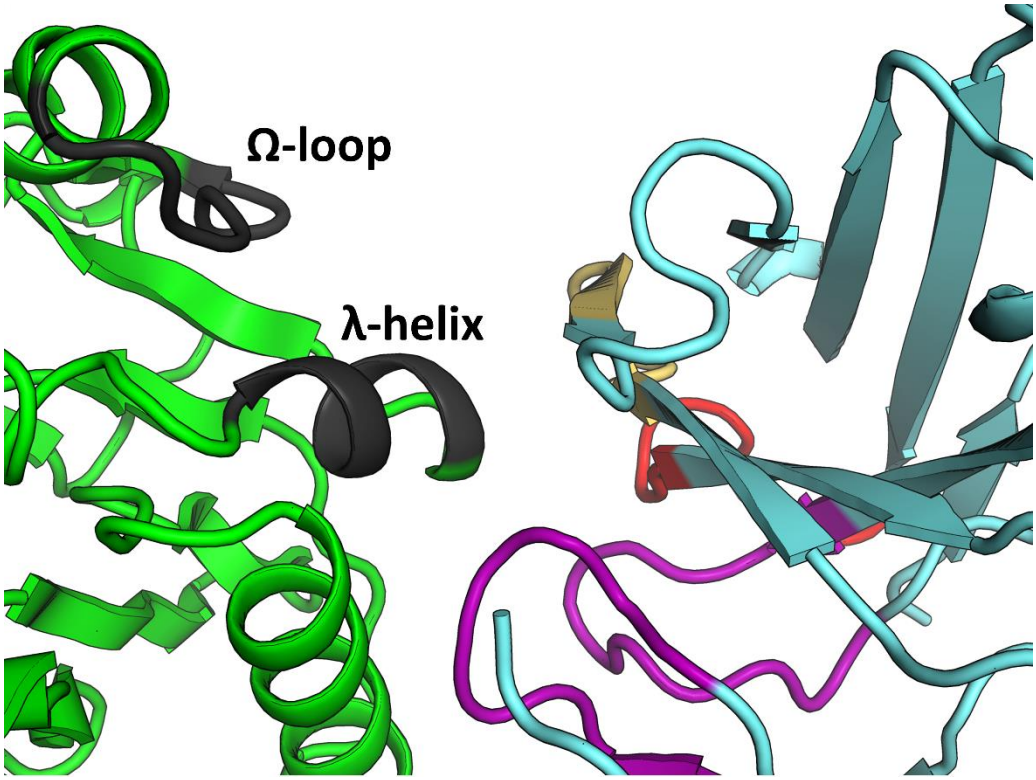


Figure 75: Closeup of the Ω-loop (black, top) and the helix with Asp130 (black, bottom) in complex with Fab5 (light blue). CDR regions are coloured according to the IGMT scheme.

4.2.13 Cryo-EM reveals the formation of higher order assemblies of NDM-palmitic acid

As crystallization of NDM-1 has proven to be mostly unreliable, Cryo-EM was used to investigate the potential formation of higher order assemblies of NDM-1 in the presence of palmitic acid. Upon incubation with palmitic acid, the sample was plunge-frozen and imaged on a Glacios Cryo-TEM.

NDM-1 appears as a well-dispersed monomeric protein and there does not appear to be any significant aggregation or oligomerisation (**Figure 76**). In contrast, when imaging NDM-1 incubated with palmitic acid, a change in morphology is immediately evident, with the formation of ring-like clusters (**Figure 77**). Due

to the impact of COVID-19, however, it was not possible to pursue this any further.

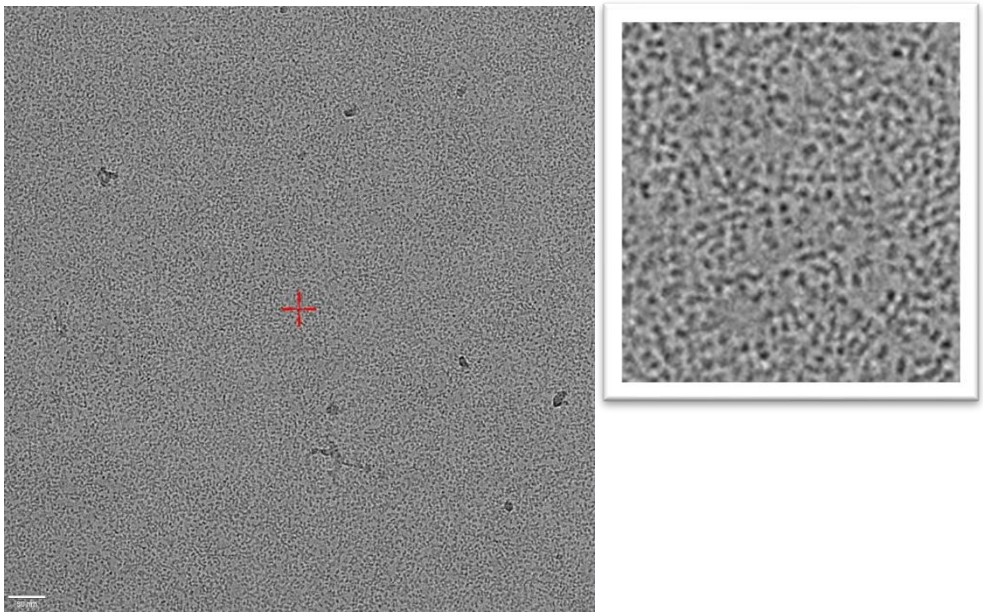


Figure 76: Monomeric NDM-1. Collected on a Glacios Cryo-TEM. Magnification: 73kx, Defocus: -3um. A closeup of the particles is shown in the white insert.

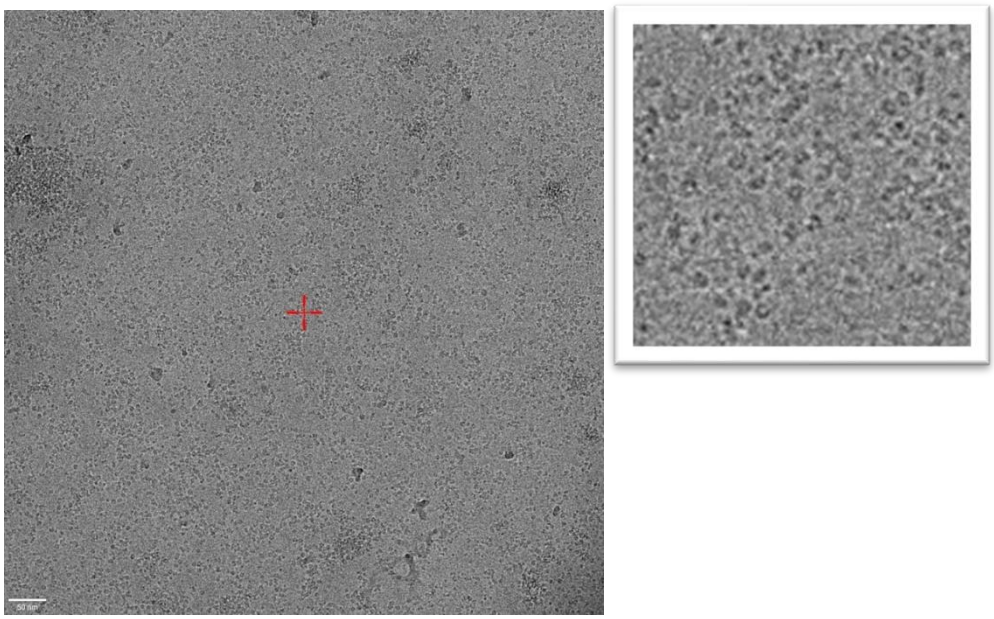


Figure 77: Ring-like clusters of NDM-1. Collected on a Glacios Cryo-TEM. Magnification: 73kx, Defocus: -3um. A closeup of the particles is shown in the white insert.

4.2.14 NDM-1 inhibition assays

To confirm the preliminary data, additional experiments were performed to monitor the inhibition of the hydrolysis of β-lactams by NDM-1. NDM-1 was incubated with Fab 5 and Fab 17 for 10 minutes before addition of the appropriate substrate, then immediately brought to a plate reader to monitor the progress of the reaction. In the tested experimental conditions, both Fabs exhibited different behaviours (**Figure 78**).

When Nitrocefin (a cephalosporin substrate) is used as the reaction substrate, both Fab 5 and 17 show an inhibitory effect, consistent with the results of the preliminary assays (see **4.2.1**).

When the substrate is changed to Meropenem (a carbapenem substrate), Fab 5 does not appear to have any inhibitory effect on the enzyme but surprisingly, Fab 17 shows a marked enhancing effect, almost doubling the relative activity of NDM-1.

Values for the relative activity of NDM-1 with the highest concentration of Fab applied are given in **Table 12**.

	Concentration	**Relative activity (%)**
Fab 5 – Nitrocefin	12µM	32.38 ± 1.25
Fab 17 - Nitrocefin	24µM	54.00 ± 0.93
Fab 5 – Meropenem	12µM	101.72 ± 0.82
Fab 17 – Meropenem	24µM	187.84 ± 0.96

Table 12: Relative Activity values for NDM-1 in the presence of different Fabs and substrates. Relative activity is calculated as the ratio between the activity of the enzyme in the presence of Fabs and without and expressed as percentage. Relative activity was calculated in the linear range, between the 5 and 15 minutes timepoints using the maximum inhibition/enhancement curves versus the control. Due to the impact of COVID-19 it was not possible to carry out a more thorough analysis with multiple concentrations.

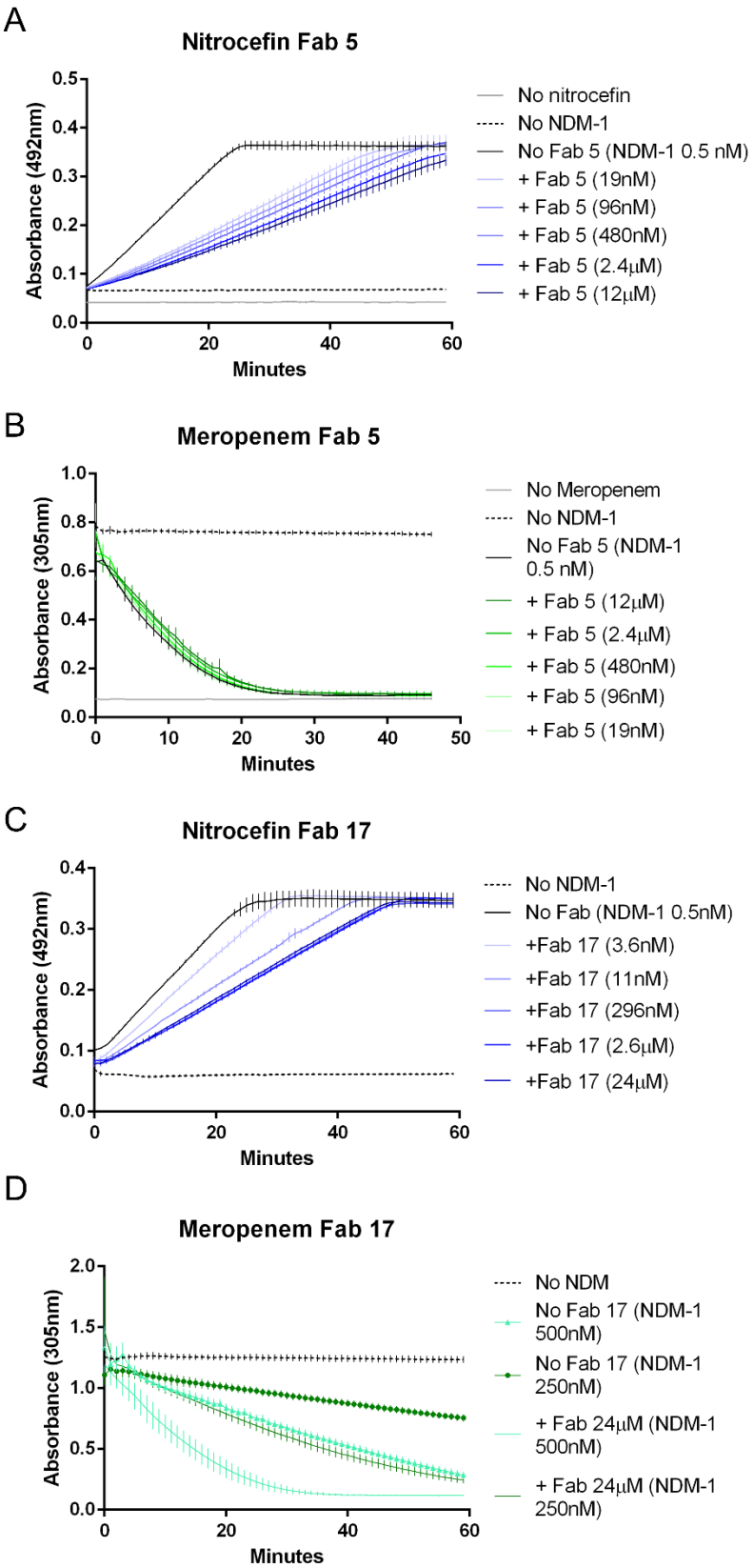


Figure 78: Activity assay of Fab 5 on NDM-1 in the presence of Nitrocefin (A) or Meropenem (B) substrates. Non-hydrolysed Nitrocefin has a low absorbance at 492nm and the signal increases as the reaction progresses. Non hydrolysed Meropenem has high absorbance at 305nm and the signal decreases as the reaction progresses. Reactions were run in triplicates.

Due to the impact of COVID-19, it was unfortunately impossible to perform further assays. Given an appropriate amount of time, it would have been possible to determine Michaelis-Menten kinetic parameters such as K_M and V_{max} to better understand how Fab 5 and 17 alter the activity of the enzyme and explore the possibility of allosteric inhibition of NDM-1.

4.3 Discussion

The need for new effective treatments to counter the growing threat of antibiotic resistance is now more important than ever. One of the main causes of concern for antibiotic resistance is the New Delhi metallo β-lactamase (NDM-1). This is due to multiple factors: the broad spectrum of antibiotics hydrolysed [5], the resistance to commonly used inhibitors[4], the lack of standardised and reliable diagnostic tools [161] and the potential for horizontal gene transfer [165]. Furthermore, the lack of satisfactory progress in the discovery of new antibiotics is likely to impair our ability to counter these pathogens. [188]

In this context, our collaborators obtained a panel of murine monoclonal antibodies directed against this important target. Two of these antibodies, according to preliminary analyses showed an inhibitory effect on the enzyme activity (see **4.2.1**).

Cloning, expression and purification of both NDM-1 and the two Fabs have been successful, with high yields and purity. In terms of crystallography, the complexes, somewhat surprisingly, have proven to be a challenging target, with slow-growing crystals (up to a year) and in many cases not well-diffracting. Despite these difficulties it was possible to obtain crystallographic structural details on Fab 5 (see **4.2.9**) and the complex between NDM-1 and Fab 17 (see

4.2.10). To obtain structural details on the remaining NDM-1-Fab 5, however, it was necessary to employ single-particle Cryo-EM (see **4.2.11**).

Finally, it was possible to identify palmitic acid as an interactor of NDM-1. The interaction was confirmed with multiple techniques and visualized by Cryo-EM see **4.2.12**).

Both these Fabs bind NDM-1 at a similar epitope, situated far from the active site of the enzyme. As the binding site does not involve the active site of the enzyme, it was not immediately clear what the mechanism of inhibition of these antibodies may be.

As discovered in subsequent biochemical assays **4.2.14**, the inhibitory activity of the two Fabs is not consistent between substrates. While for Nitrocefin (part of the class of cephalosporins) the inhibition is consistent between the two Fabs, the situation for Meropenem (part of the class of carbapenems) is different. In this case, upon binding of Fab 5, no appreciable change in activity for the enzyme can be observed. When NDM-1 is bound to Fab 17 the activity of the enzyme is increased by almost 90%. While this may appear surprising at first, recall that the enzyme reaction mechanism for the hydrolysis of various substrate classes is different [181].

Two interesting known natural variants of NDM-1 are NDM-7 and NDM-8: both variants have mutations on two specific residues (Asp 130 and Met 154), however while for both of them Met 154 is mutated into a Leucine, Asp 130 has different mutations. In NDM-7 Asp 130 is mutated into an Asparagine, whereas in NDM-8 into a Glycine. When comparing the carbapenemase activity of these two mutants, in NDM-7 it has been shown to be increased [176], while in NDM-8 it

is similar to that of NDM-1 [177]. Given that the two variants only differ by one amino acid change, it is likely that Asp 130 plays a significant role in the modification of the enzymatic activity, despite not being part of the active site.

Analysis of the structures of the two complexes shows that Fab 17 forms a network of hydrogen bonds between Asp130 (on NDM-1) and three residues on the Fab part of the CDR1 and CDR2 regions (Tyr 33, Ser 53, Ser 55). The same does not hold true for Fab 5, where this interaction, does not occur. Furthermore, Fab 5 does not interact with the Ω-loop, which is known to alter the activity of the enzyme [204].

Sitting before the λ-helix are 3 conserved residues (His 120, His 122, Asp 124) that coordinate the zinc atoms in the active site, followed by two highly flexible glycine residues (Gly 127 and 128). All of these residues are also conserved in several other metallo β-lactamases. These glycine residues induce a kink, structurally conserved among other metallo β-lactamases (**Figure 79**). As glycine residues have greater conformational flexibility, they may be able to support flexibility for the active site [205]. While in the case of Nitrocefin both antibodies appear to exert a mild but consistent inhibitory activity, it is necessary to attempt to explain the difference in activity in the case of a Meropenem substrate.

If, as the sequence conservation among metallo β-lactamases suggests, a certain degree of flexibility is important for the λ-helix, the structures of the two Fab complexes could explain the discrepancy. As Fab 5 does not bind close to this region, it would not interfere with the activity of the enzyme, thus resulting in no appreciable difference in activity, as evidenced in **4.2.14**. On the contrary, the epitope bound by Fab 17 involves both the λ-helix and the Ω-loop, likely resulting

in a stabilization of that region. A reduced flexibility could, therefore, result in an increased carbapenemase activity, perhaps by stabilizing the active site or the [ES] complex.

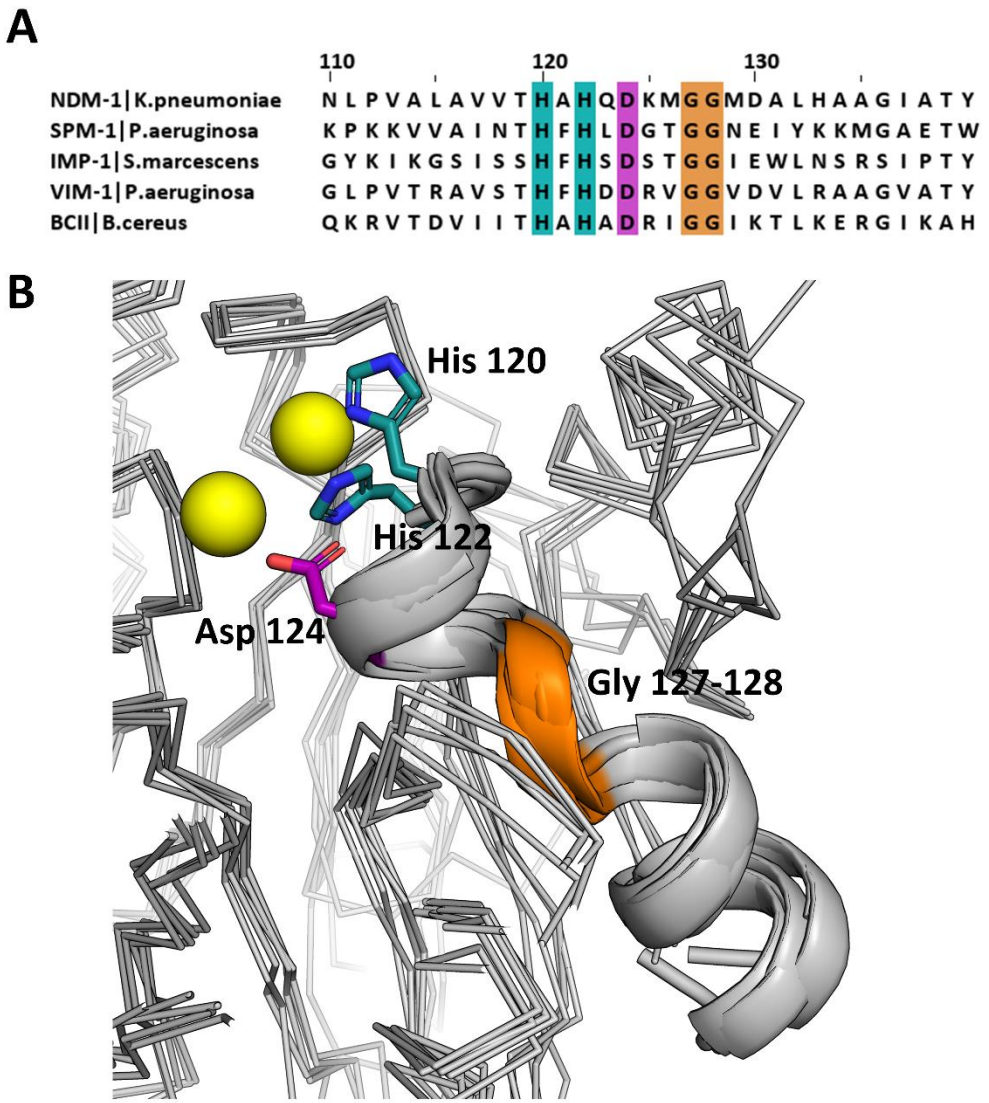


Figure 79: (**A**) Multiple sequence alignment of metallo β-lactamases. Conserved residues are coloured according to the CLUSTALX scheme. (**B**) Superimposed structures of metallo β-lactamases, conserved residues are coloured as above.

From an evolutionary point of view, however, for a reduction in the carbapenemase activity of the enzyme to be selected, the increased flexibility

must either provide some other advantage or have less disadvantages than a more stable structure. Further work, however, will be needed to validate this hypothesis.

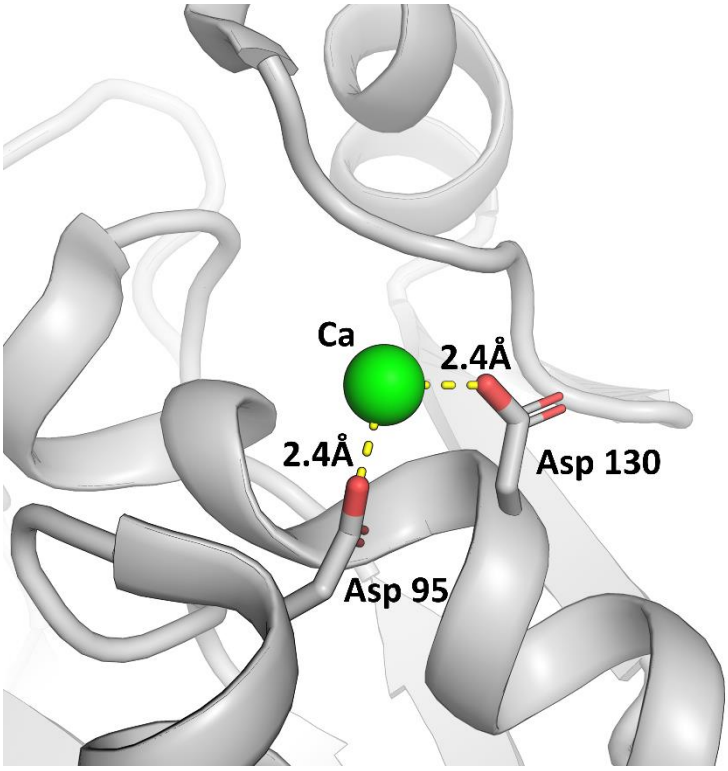


Figure 80: Calcium coordinated by Asp 95 and Asp 130 (PDB id: 6TWT)

It must also be noted that in some reported structures of NDM-1, a metal ion (identified as either Calcium or Zinc in other published structures) can be found close to the λ-helix, near the distortion caused by the aforementioned Glycine residues. In particular, in the case of NDM-1 it appears to be coordinated by Asp 95 and Asp 130 (***Figure 80***). At present it is not known if it plays a role in the mechanism of the enzyme or if it's an experimental artifact.

In the context of viral infections, it has been known for years that the immune system can, at times, promote the viral infection (e.g. antibody-dependent enhancement in Dengue virus [206]). In this case, some of the antibodies that were produced by the organism exposed to the NDM-1 antigen (i.e. *M.musculus*) appear to be capable of enhancing the hydrolytic activity of the enzyme, at least for one of its substrates. By analogy, it is not unreasonable to speculate that a similar process can also happen in the context of a bacterial infection, with the host organism indirectly reducing the efficacy of β-lactam antibiotics. At present, however, no further data is available to support this hypothesis and, given the implication of such a statement, exceptional evidence will be needed to verify this possibility.

Finally, it was possible to serendipitously identify palmitic acid as an interactor of N-terminally truncated NDM-1 (see **4.2.4**). It is also known that full-length NDM-1 is lipidated [167] and inserted into the outer membrane in gram-negative bacteria [166]. Cryo-EM micrographs of NDM-1 incubated with palmitic acid show a marked change in morphology with the protein transitioning from a well-dispersed monomeric distribution to a series of circular clusters (**4.2.13**).

Binding of palmitic acid may be a way for NDM-1 to shield the lipid moiety from the surrounding hydrophilic environment during transport of the protein across the periplasmic space. This has indeed been shown to be the case for the HIV-1 matrix protein [207].

Taking into account the binding of lipid molecules by NDM-1, it is possible to speculate that ring-like assemblies, such as the ones that can be observed via cryo-

EM upon incubation of NDM-1 with palmitic acid may form in a cellular context, with each protomer daisy-chained together with the next via lipid moieties.

Furthermore, outer membranes in gram-negative bacteria contain significant amounts of palmitic acid [208][209] and membrane lipids are known to be transiently exposed to the solvent, which is thought to be important for membrane fusion [210]. It is therefore possible for NDM-1 to be able to bind these exposed lipids, which would explain the puzzling data obtained in the Dengue virus ELISA assay (see **4.2.4**), as lipids with a palmitoyl- fatty acid are known constituents of the virus envelope [211]. More work, however, will be needed to investigate this possibility.

4.3.1 Future work

While the available data allows us to speculate on the mechanism of action of these antibodies, further experiments will be needed to confirm the proposed mechanism. Due to the extremely long crystallization times of the Fab complexes and the impact of COVID-19 many experiments could not be carried out.

In principle, structure-guided mutagenesis of NDM-1 could be employed to establish disulphide bridges in order to artificially restrict the hypothesized movement of NDM-1 within the Fab 17 epitope. Two proposed mutations are shown in **Figure 81**. The first proposed disulphide would involve mutating two Alanine residues (Ala 99 and Ala 131) into Cysteine residues, which would be at an ideal distance of ~2Å. The second pair would be Asp 130 and Val 155. The resulting Cysteines would be ~3.5Å apart. While not ideal, Val 155 resides within the Ω-loop and the inherent flexibility of the loop region could compensate for

the excessive distance. As NDM-1 is a periplasmic protein, oxidising conditions in the compartment should allow the formation of the disulphide bridges.

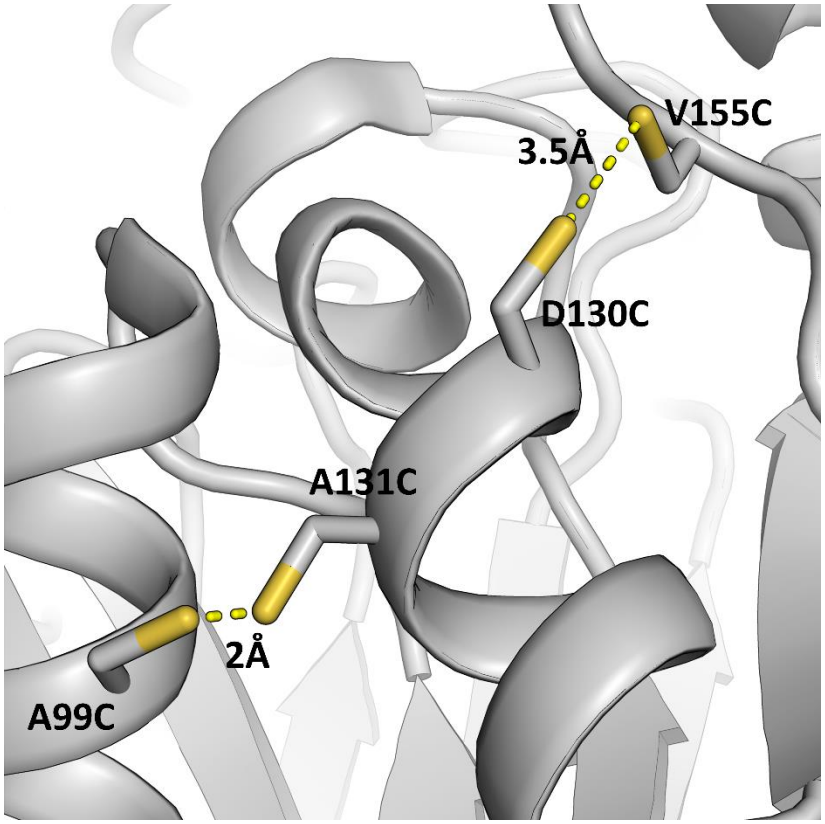


Figure 81: Proposed mutations for the formation of disulphide bridges in NDM-1.

Once the mutants are obtained it would be possible to measure kinetic parameters for the hydrolysis reactions of both Nitrocefin and Meropenem, as detailed in **4.2.14**.

The original aim of the project was to employ these antibodies as a potential therapeutic against antibiotic resistant bacteria. Subsequent analyses have shown the inhibitory effect to be mild and, depending on the substrate, unreliable. This pilot project screened a relatively small number of antibodies, but a more extensive screen (ideally of human patients-derived antibodies) could identify more promising candidates. Even though there is evidence for the membrane

penetrating power of IgGs [194][195][196], such an approach would benefit from a smaller molecule, both for tolerability, efficacy and cost. Synthetic peptides derived from CDR sequences have found some success against *Candida albicans* [212] and it could be a venue of research worth investigating.

One immediate application, however, could be the use of these antibodies to study the biology of NDM-1, perhaps in combination with mutants, such as the ones suggested for the disulphides. The antibodies could be used as a tool to rapidly detect an approximate binding region for small-molecule inhibitors of NDM-1, without the need to obtain the structure of the complexes. This could accelerate the identification of potential allosteric modulators of its activity. As mentioned in **4.2.1**, Antibody 8 has also been identified in the preliminary screen as an enhancer of the enzyme activity against Nitrocefin. Preliminary data collected shows that it is not possible to bind both Fab 8 and either Fab 5 or 17 at the same time. This suggests that the epitope for Fab 8 likely overlaps with that of both other Fabs. Obtaining structural details of this third complex could prove to be very valuable in accurately pinpointing regions of NDM-1 involved in the hydrolysis of its substrates.

One more potential application for Antibody 17 is diagnostics. Currently there are no reliable tests to detect metallo β-lactamases expressing bacteria and the test are slow, laborious, error-prone and expensive, due to the need to culture specimens, extract and amplify DNA or use mass-spectrometry [161]. As detailed in **4.2.10**, Fab 17 is capable of binding mutants of NDM-1 engineered to mimic two of its natural variants, where Asp 130 is mutated into Asparagine or Glycine. It should be possible to develop a proof-of-concept for an indirect ELISA assay capable of detecting NDMs in a simple, cheap and fast way. To do so, the first

step is to ensure that Antibody 17 can bind *E.coli* expressed and purified untagged NDM-1: after running a native PAGE gel and transfer to a membrane for western blotting, the membrane can be incubated with antibody 17. As antibody 17 is derived from mouse, a simple anti-mouse antibody conjugated with a molecule for detection (e.g. horseradish peroxidase) could be used to detect the presence of these proteins. Once this is established, a protocol will need to be generated for assaying patient-derived samples. Given that NDMs are periplasmic (as well as present in extracellular vesicles) [166], it should be possible to free the protein and make it detectable by the antibody via chemical methods or osmotic shock. Further work, however, will be necessary to establish the feasibility of this assay, the appropriate parameters to reliably identify these proteins and validating it against samples from patients.

Finally, the serendipitous identification of palmitic acid as an interactor of NDM-1, warrants further investigation. Preliminary data collected via Cryo-EM, shows that the addition of this lipid induces the formation of seemingly ordered clusters of NDM-1. However, the small dataset collected before the impact of COVID-19 did not yield a satisfactory reconstruction. The first obvious experiment on this topic would include the collection of a large dataset of NDM-1 with palmitic acid. Building on that, structural analysis of NDM-1 in the context of a membrane is the next logical step. For eukaryotic proteins, palmitoylation has been shown to be one important driver of their association to lipid rafts [213] and such rafts have been also recently identified in bacteria [214]. If palmitoyl-mediated targeting of protein holds true in *E.coli,* purification of bacterial lipid rafts and their imaging could be within technical capabilities. Alternatively, a simpler system, using

NDM-1 inserted in the lipid bilayer of liposomes could provide meaningful structural information, albeit in a more artificial system.

5 High-resolution crystal structure of arthropod Eiger TNF suggests a mode of receptor engagement and altered surface charge within endosomes

5.1 Introduction

5.1.1 A brief overview of the Eiger protein

The regulation of a wide variety of cellular processes such as apoptosis, immune system regulation, tissue development and proliferation is affected by members of the Tumour Necrosis Factor (TNF) alpha superfamily, which play a critical role in these mechanisms[215]–[217]. Members of this superfamily are synthesized as type II transmembrane proteins, i.e. proteins with a cytosolic N-terminus, a transmembrane portion and an extracellular C-terminus. In these proteins, the TNFα domain is part of the extracellular domain which can be cleaved and released as a soluble signalling molecule[218].

TNFα ligands become active after oligomerization, forming a trimer that is in turn recognized by TNF receptors and activates downstream signalling. This happens through the action of TNF receptor-associated factors (TRAFs) that relay the signal and trigger the activation of the Jun-kinase cascade[10], [219]–[223].

In *Drosophila melanogaster* (a widely used model organism) only one gene for a TNF ligand and two for TNF receptors are present. The ligand is called Eiger, while the two receptors are called Wengen[9] and Grindelwald[11]. Which of these receptors are bound by Eiger can be tissue and context specific.

In mammalians, however, the repertoire is much wider, with at least 19 and 29 genes, respectively[218]. This is thought to be the result of several whole-genome duplication events that happened throughout the evolutionary history[12].

Furthermore, while invertebrates rely only on the innate immune system, vertebrates have also an adaptive immune system. It is worth noting that several TNFs in vertebrates are implicated in processes like B and T cells regulation. The current view is that these events of genome duplication have increased the number of TNFs, allowing them to be repurposed for the regulation of the ancestral adaptive immune system[224].

For these reasons, Eiger can be considered an evolutionary homologue of TNF in vertebrates. Drosophila can then be used to study the function of this class of proteins irrespective of the adaptive immunity[225].

In the cell, Eiger is involved in a number of different processes during development as well as in the adult organism. As a tumour suppressor, Eiger is involved in the maintenance of the epithelium, more specifically in the removal of polarity-deficient cells[226]. However, in certain contexts, such as when oncogenic Ras is present, Eiger can instead promote tumour growth[227].

Eiger is also involved in tissue development and remodelling: when overexpressed in the wing and in the eye of the insect, it results in cell death[228], [229]. Eiger can not only exert its signalling activity on directly adjacent cells but also at a much bigger distance. This is possible due to the TNFα converting enzyme (TACE)[230], that cleaves the extracellular domain of Eiger making it soluble. This results in larger groups of cells dying[231]. This process, while still not completely understood, is believed to be important in the remodelling of

existing tissues[232], [233]. Eiger is also involved in autophagy: by acting on the JNK pathway it induces the expression of sigmar (salivary glands marred) triggering autophagy and remodelling of the cytoskeleton[234].

Eiger is also involved in a variety of other processes: it intervenes in the immune response where it activates the signalling cascade that results in the release of the prophenoloxidase stored in crystal cells, a key immunity mechanism in arthropods[235]. Eiger also helps to clear extracellular pathogens by activating the cellular immune response[236], as well as facilitating the migration of macrophages between tissue barriers[237].

Finally, Eiger is involved (along with Wengen) in Drosophila larvae in allodynia (i.e. heightened sensitivity to pain) following UV damage, by acting on sensory neurons[238].

5.1.2 Aim of the project

This project started with the serendipitous crystallization of *Spodoptera frugiperda* Eiger. The challenge in this project lies in adopting a different approach to structure determination. As diffraction data was the only data available, careful assumptions had to be made in order to determine the identity of the target. Whilst at the beginning of this project structural details were available for several members of the TNFα superfamily, no structure was available for any of the Eiger proteins. Structural information on Eiger could shed light on the insect homologue of the TNF system, as well as provide a possible target for the development of targeted pesticides against *S.frugiperda*, an important agricultural pest [239].

5.2 Results

5.2.1 Identification of SfEiger

While trying to obtain structural information on the lammi virus glycoprotein[240], a crystal (***Figure 82***) grew after about 80 days at 298K in the C2 condition of the Morpheus HT-96 screen[32] (0.09M Sodium nitrate, Sodium phosphate dibasic, Ammonium sulphate; 0.1M imidazole, MES monohydrate (acid); 16.530% v/v Ethylene glycol, PEG 8000).

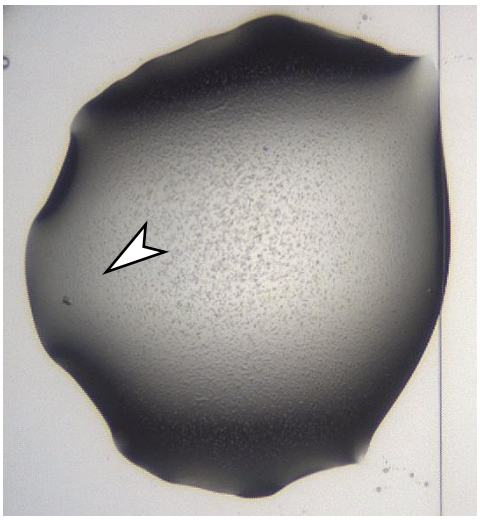


Figure 82: A crystal of SfEiger grew in condition C2 of the Morpheus screen after 80 days

The crystal was cryocooled after the addition of 20% (v/v) of glycerol to the mother liquor and subjected to a diffraction experiment at Diamond Light Source (Didcot, UK) at Beamline I04. Diffraction data collected to 1.7Å was then processed with Xia2[31].

The protein crystallised in the R 3 space group, with a trigonal unit cell. Full data collection statistics are given in **Table 13**. A careful consideration of the Matthews coefficient, assuming a 50% solvent content, suggests ~17kDa in the asymmetric unit, which appeared unlikely to be compatible with the original target. This suggested that the crystal was not of our intended target, but of a contaminant protein. Alternatively, it could be a degradation product of the original target.

Data Collection	SfEiger TNF domain
Wavelength (Å)	0.9795
Resolution range (Å)	47.4 – 1.7 (1.72 – 1.69)
Space group	R 3 :H
Cell dimensions	
a, b, c (Å)	55.2, 55.2, 142.1
α, β, γ (°)	90, 90, 120
Total reflections	152050 (2675)
Unique reflections	17657 (757)
Multiplicity	8.6 (3.5)
Completeness (%)	97.5 (81.2)
Mean I/σ(I)	10.1 (1.2)
R-merge	0.131 (0.974)
CC ½	0.997 (0.401)

Table 13: Data collection statistics for SfEiger. Statistics for the highest-resolution shell are shown in parentheses

A possibility to consider was that the crystallised protein is endogenous to our expression system, *Spodoptera frugiperda*. Given the serendipitous nature of the crystal, reproducing it seemed unlikely and the resolution of the data was not high enough to allow for direct phasing. It was necessary to adopt a different approach to solve the structure, using the Wide Search Molecular Replacement (WSMR) server[8] to obtain a suitable model for molecular replacement. The server attempts molecular replacement using Phaser[34] with up to 100000 SCOP domains as search models. The best solution (TFZ: 9.1, LLG: 65.00), corresponded to the crystal structure of murine APRIL at pH 8.0 (PDB ID: 1U5Y). Chain "D" of the APRIL trimer was used for the molecular replacement, as suggested by the WSMR.

SHELXE[39] was then used to trace the main chain in the density, along with iterative density modification to improve the phases, giving a ~88% complete main chain (150 amino acids over 3 chains of 109, 29 and 12 residues) . Features in the map were clearly visible and allowed for unambiguous identification of a

majority of the side chains (***Figure 83***). The initial model that could be built was ~60% complete.

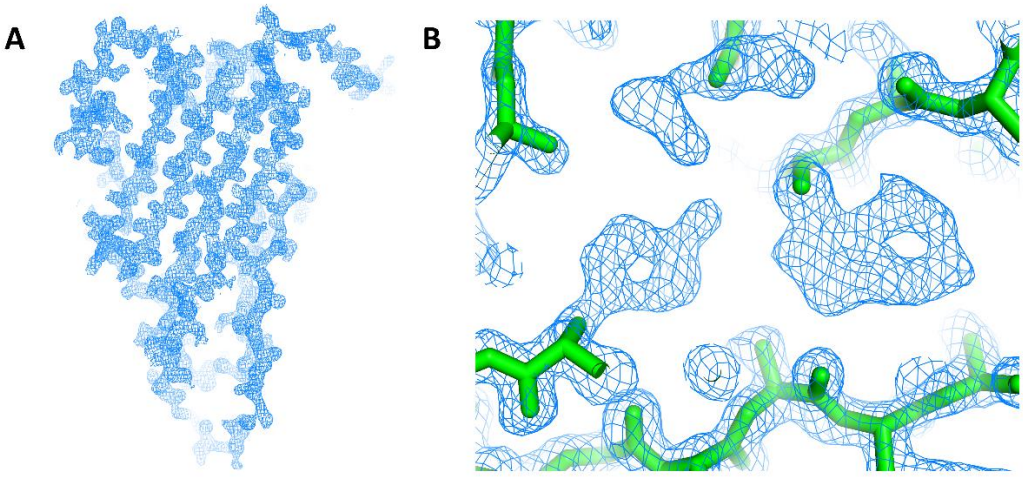


Figure 83: 2Fo-Fc electron density map (contour level 1.5σ) of SfEiger after SHELXE and one round of refinement. (**A**) Overview of the structure. (**B**) Example of some unambiguously identifiable side chains (i.e. a tyrosine on the left, a tryptophan on the right. Adapted from [241].

The sequence inferred from this initial model was then used as query for tBLASTn[42] using the genome of *Spodoptera frugiperda* as nucleotide database. At the time of the study, however, this genome was sequenced[40] but not annotated and directly searchable on NCBI. In order to run BLAST on the genome it was first necessary to download the genome and convert it using the makeblastdb command of BLAST.

Based on the available structural information, it was possible to identify 3 nucleotide sequences that, when translated, covered the entire structure. Genomes, however, contain both coding and non-coding sequences (i.e. exon and introns, respectively) and it was necessary to identify the exact splicing sites on these sequences.

To do this, the canonical splicing pathway (**Figure 84**), which accounts for the majority of splicing events[242], was considered when choosing the possible splicing sites.

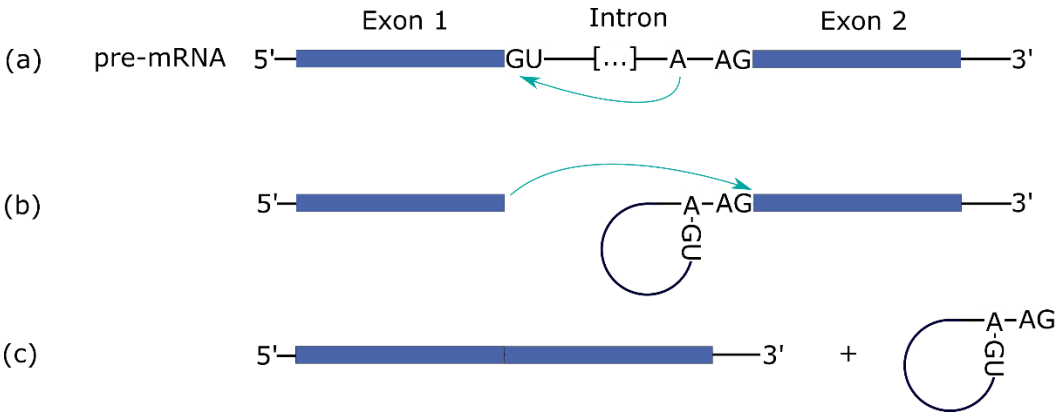


***Figure 84**:* Canonical splicing mechanism. The splicing reactions are catalysed by the spliceosome. In the first reaction (**a**), the 2'-OH of the branchpoint nucleotide executes a nucleophilic attack to the first Guanine in the intron. In the second reaction (**b**) the 3'-OH of the first exon executes another nucleophilic attack, this time to the first nucleotide of the second exon. Finally (**c**), the exons are joined and the *lariat* intron is released.

Annotation of the exons for SfEiger is shown in **Figure 85**. It is worth noting that the numbering of the exons starts from the first Methionine residue before the first residue visible in the density. More exons will code for the full-length SfEiger, but the lack of further information limits the scope of the analysis. For the same reason it was decided to number the residues starting from the first amino acid clearly visible in the electron density.

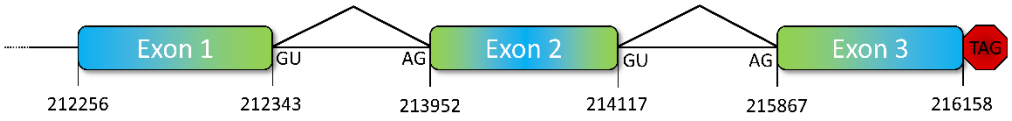


Figure 85: Exons coding for the TNF domain of SfEiger. Numbering refers to the nucleotide positions of the boundaries of each exon in the *Spodoptera frugiperda* genome. Adapted from [241].

Recognition of the splice site between Exon 2 and 3 of SfEiger was immediate, thanks to the supporting structural data, but for the splice site between Exon 1 and 2 the lack of a clear electron density made necessary the use of supporting information.

A multiple sequence alignment (***Figure 86***) was performed using BLAST with the partially complete sequence of SfEiger as query against the database of non-

redundant protein sequences, choosing homologues in closely related organisms as well as more distant human homologue. The closest homologue is from *Spodoptera litura* (94% sequence identity), followed by *Helicoverpa armigera* (88%) and *Vanessa tameamea* (86%).

These findings were further validated by RT-PCR, by extracting the total RNA from SF9 cells and retrotranscribing the putative sequence of the TNF domain of SfEiger. The cDNA was further amplified by PCR (***Figure 87*A**) and sequenced.

Analysis of the sequences (***Figure 87*B and C**) shows a limited numbers of mismatches that, when translated, give rise to the exact sequence determined from the published genome.

Taking all these findings together, the identity of the crystallized SfEiger was confirmed and the final protein sequence was built in the model. While no quantitation of the levels of the messenger RNA in the context of live SF9 cells was performed, its presence could be confirmed.

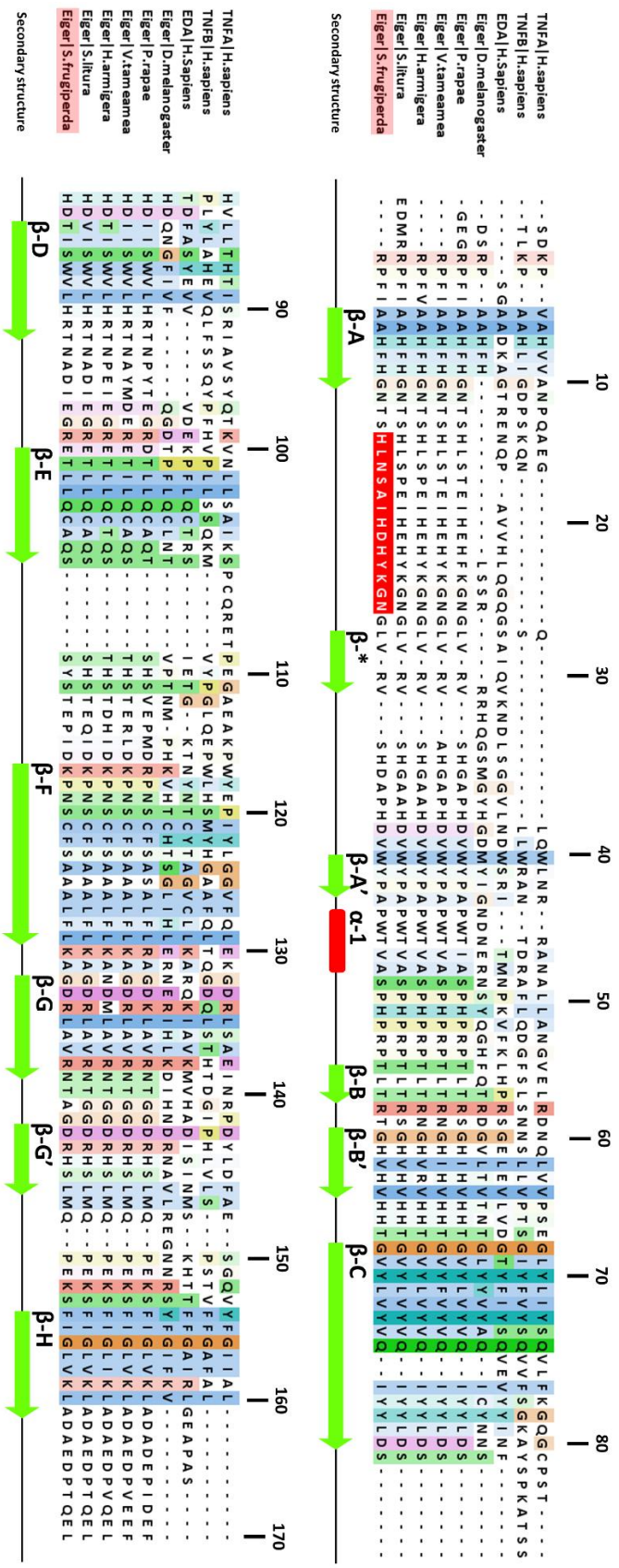


Figure 86: Multiple sequence alignment of SfEiger. Numbering refers to the sequence of SfEiger ad coloring is according to the ClustalX color scheme. Residues 13-26 (in red) are disordered in the structure. Secondary structures elements are highlighted below the alignment. Adapted from [241].

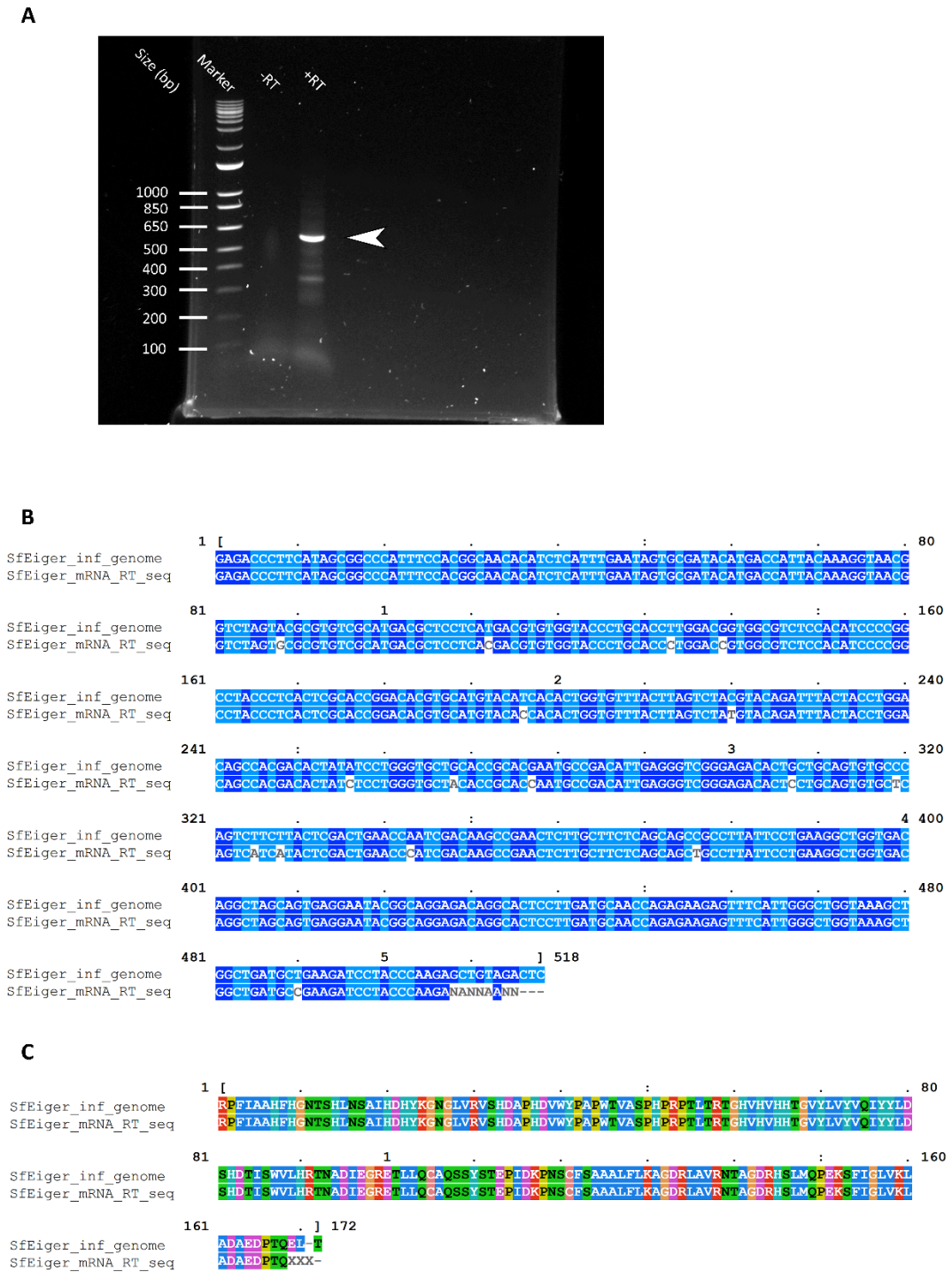


Figure 87:RT-PCR of SfEiger. (A) Agarose gel of the product both in the absence and presence of the retrotranscription enzyme. (B) Alignment between the putative coding sequence of SfEiger and the experimentally determined sequence. (C) Alignment between the translated putative and experimental sequences. From [241].

5.2.2 The structure of SfEiger TNF domain

After building and refining the final model (***Figure 88*** and ***Table 14***), a closer look at the structure of the TNF domain of SfEiger shows that the protein adopts a jelly-roll fold like other members of the TNF superfamily[219].

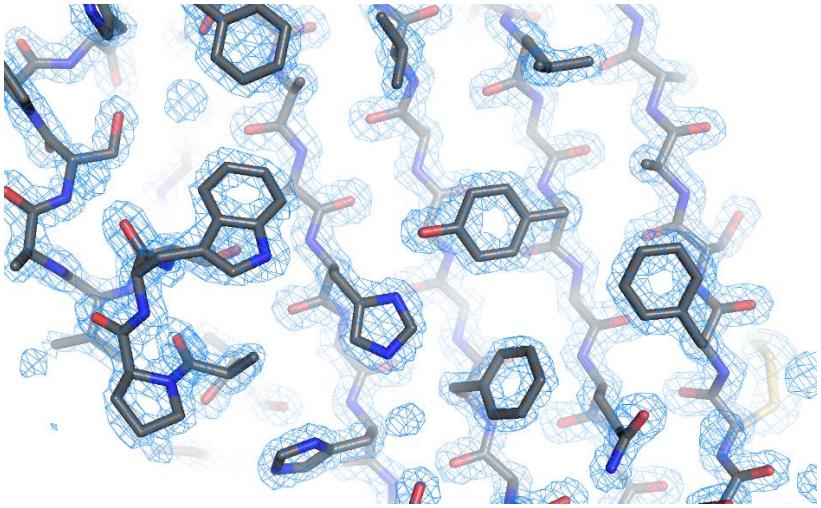


Figure 88: Representative 2Fo-Fc electron density map (contour level 1.5σ) after the final round of refinement. From [241].

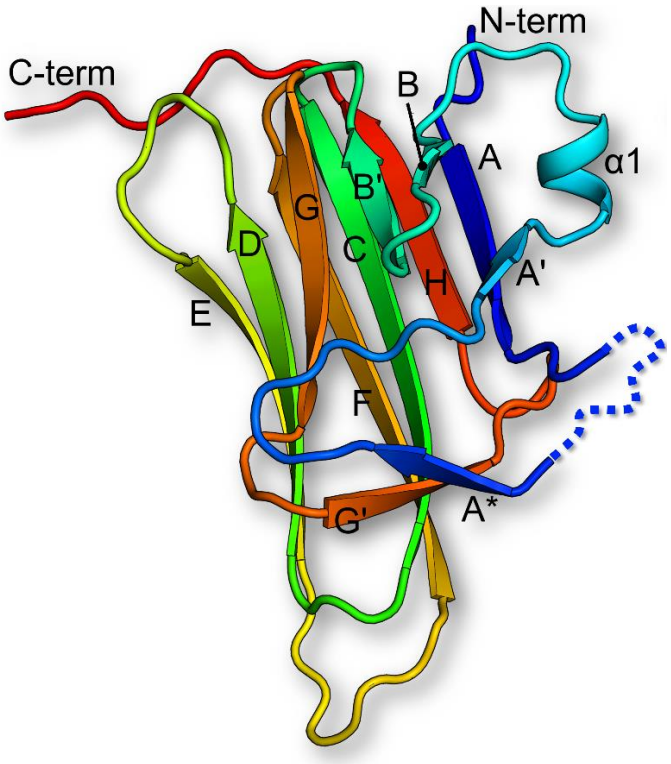


Figure 89: Monomer of SfEiger. Secondary structure elements are labelled. A dotted line shows the stretch of amino acids that is not visible in the electron density. Adapted from [241].

Refinement	SfEiger TNF domain
R_{work} / R_{free}	15.5/18.2
Number of atoms	
Protein	1240
Solvent	175
RMSD(bonds)	0.005
RMSD(angles)	1.15
Ramachandran plot	
Favoured (%)	97.39
Allowed (%)	2.61
Outliers (%)	0.00
Average B-factor	22.47
Protein	20.37
Solvent	37.33

Table 14: Refinement statistics for SfEiger.

The SfEiger TNF domain (***Figure 89***) is composed of 170 amino acids (13 of which are not clearly visible in the density) mostly composed of β-strands organised in two β-sheets. A disulphide bridge is also present between Cys 105 and Cys 121, linking strands E and F.

The DALI server[243] was used to compare the structure of SfEiger to others present in the PDB and identify its closest structural homologue. The structure of Ectodysplasin A1 (PDB ID: 1RJ7) was the hit with the highest score (22% sequence identity, 1.5 Å RMSD). This is also corroborated in the literature[12].

Comparing the two structures (***Figure 90***) shows a few differences: in SfEiger there is an α helix in the region between strands A' and B and the loops between strands D and E (DE flap), as well as between E and F. This could have functional implications as α1 helix contacts neighbouring protomer in the canonical trimer (**section 5.2.3**). The DE flap has also been shown to be necessary in the formation of viral-like clusters at neutral or basic pH in another TNF superfamily member, BAFF[244], [245].

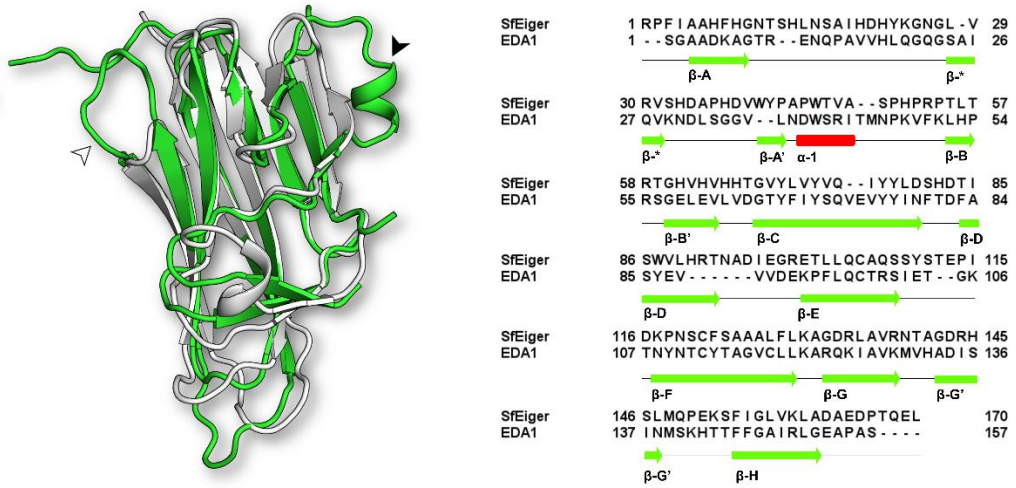


Figure 90: (**Left**) Comparison between the structure of SfEiger (PDB ID: 6I50) and of Ectodysplasin A1 (PDB ID: 1RJ7). The black arrowhead points to the helix α1, the white arrowhead to the extended DE loop. (**Right**) Sequence alignment between SfEiger and Ectodysplasin A1. Adapted from [241].

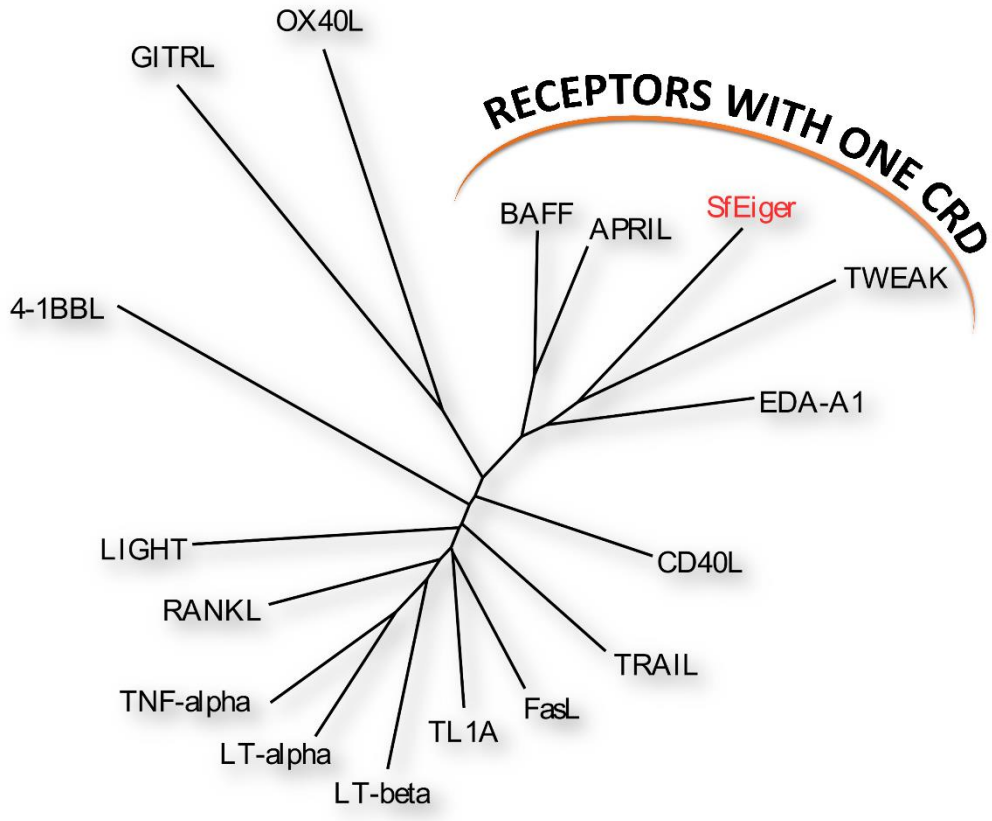


Figure 91: Structural phylogeny of the TNF superfamily. The structural phylogenetic tree was calculated using the structural homology program (SHP) [44]. Figure adapted from [241].

To compare how the structure of SfEiger relates with other members of the TNF superfamily the Structural Homology Program (SHP)[44] was used to generate a structure-based phylogenetic tree using available structures for each member of the superfamily (***Figure 91***). Supporting the previous findings, SfEiger clusters close to Ectodysplasin A1, BAFF, APRIL and TWEAK, all members of the superfamily that bind receptors with a single cysteine rich domain (CRD). More distant members bind to receptors with multiple CRDs. As both the receptors for Eiger, Wengen and Grindelwald have a single CRD[12], [233] it can be argued that receptors with single CRDs have appeared earlier in the evolution of the TNF system. As no structures have been reported for either receptor in arthropods, **Figure 92** shows the B-cell Activating Factor Receptor (BAFF-R) as an example of a receptor with a single CRD.



Figure 92: The CRD of the BAFF receptor. PDB id: 4V46

It is also worth noting that the disulphide bridge in SfEiger is also conserved in Ectodysplasin A, BAFF, APRIL and TWEAK[246]–[249], but not in more distant structures (e.g. TNFα).

5.2.3 SfEiger forms a canonical trimer

While the asymmetric unit of the crystal of SfEiger contains a single monomer, it is known that biologically active TNF ligands are trimeric[218]. The trimer of

SfEiger sits on the three-fold axis generated by the packing in the space group (***Figure 93***).

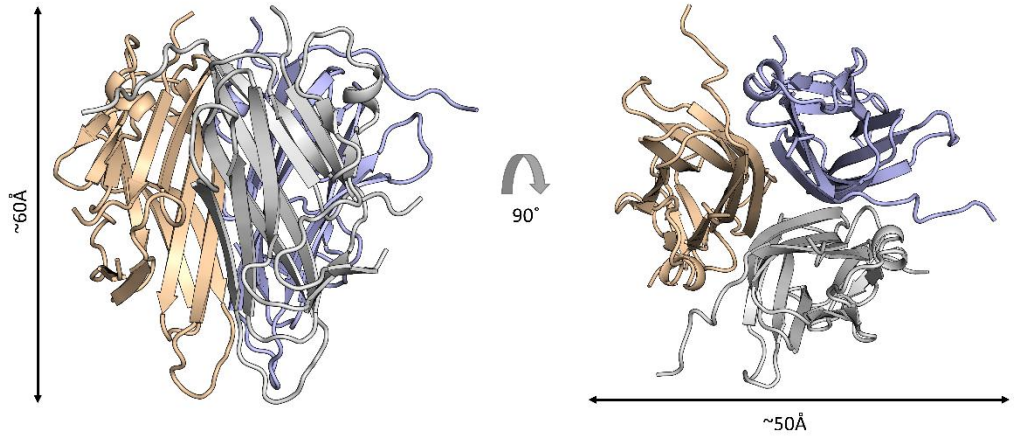


Figure 93: Structure of the SfEiger trimer: front and top views (left and right, respectively). Adapted from [241].

The trimer has the shape of an inverted pyramid, with a height of ~60 Å, and a width of ~50Å. The β-strands A, H, C and F are part of the β-sheet that forms the majority of the oligomerization interface. At the opposite side, β-strand B,B',D,E are part of the β-sheet exposed the solvent. Some additional contacts between monomers involve strands E and F.

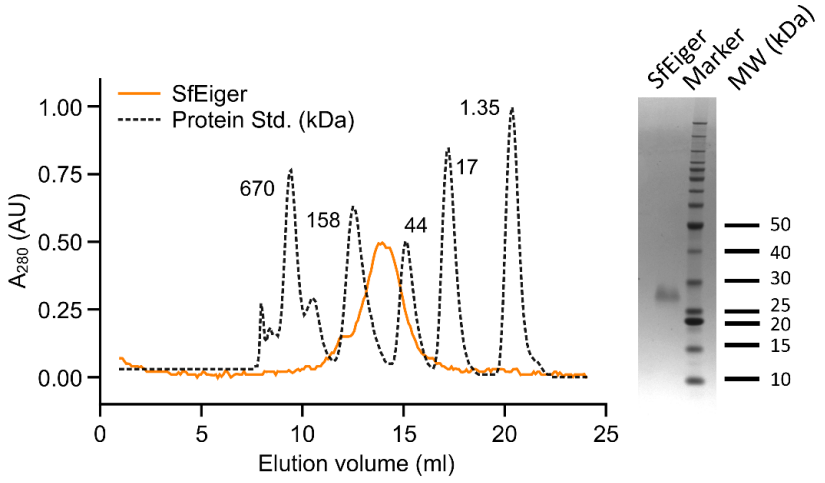


Figure 94: Size exclusion chromatography of the TNF domain of SfEiger. A Superdex S200 increase 10/300 GL was used.The orange line indicates the elution profile of the protein, the dashed line that of a protein standard. On the right, reducing SDS-PAGE of the purified SfEiger. Adapted from [241].

Experimental confirmation of the trimer required cloning and recombinant expression of the His-tagged TNF domain of SfEiger. In size exclusion chromatography, the trimer can be seen migrating at an apparent molecular weight consistent with that of a trimer (**Figure 94**).

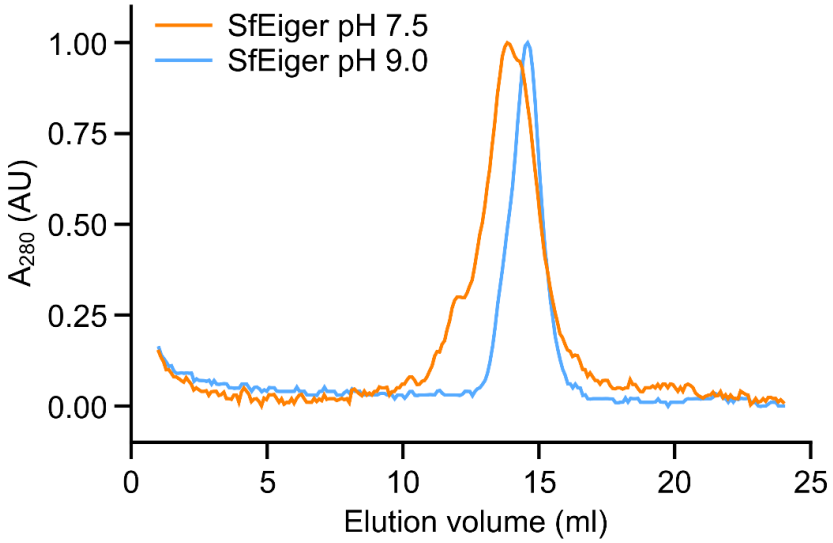


Figure 95: SEC profile of SfEiger at pH 7.5 (orange line) and 9 (blue line). Adapted from [241].

The same holds true at pH 9: unlike BAFF, SfEiger does not seem to form higher order assemblies at neutral or basic pH (**Figure 95**).

In the interface between the monomers, residues His7, Leu71, Tyr73, Phe154 and Val158 are conserved residues at the core of the oligomerization region (**Figure 96**). Indeed, the whole interface is dominated by hydrophobic interaction, shielding hydrophobic residues from interacting with the solvent. As mentioned previously, the helix α1 packs against the neighbouring monomer at the C-terminal tail.

Finally, a striking number of histidine residues (30) decorate the surface of the assembled trimer (***Figure 97***). This is not only the likely reason why SfEiger was

co-purified with the lammi virus glycoprotein but, given that histidine residues often function as pH sensors in protein[250], they may also confer a change in the surface charge of the protein, during the endocytic pathway of Eiger[251].

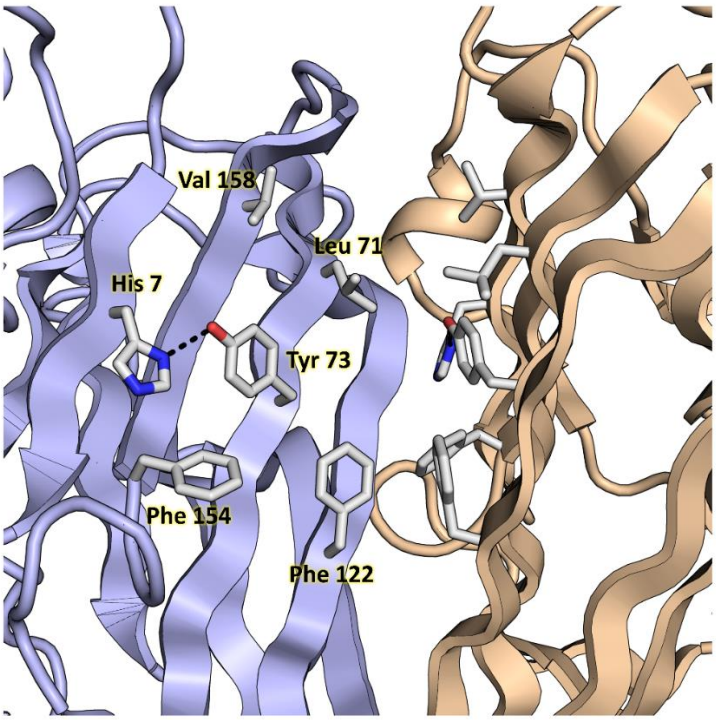


Figure 96: Oligomerization interface of SfEiger. Adapted from [241].

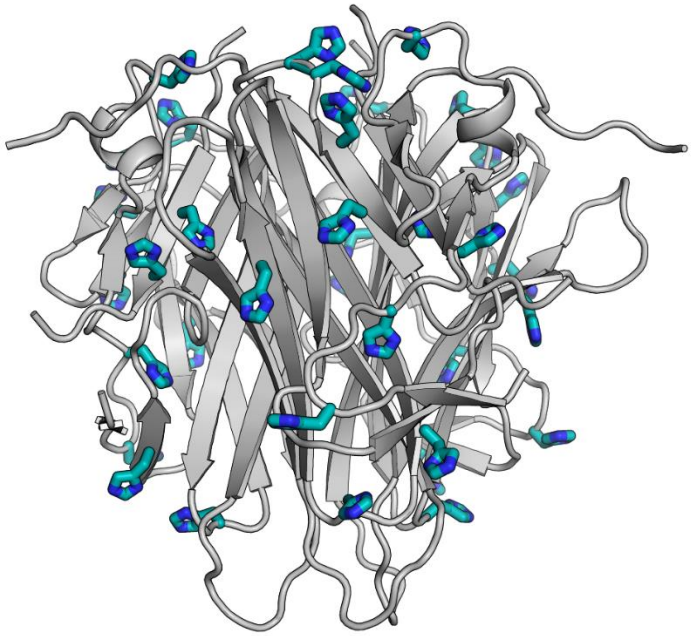


Figure 97: Surface histidine residues of SfEiger. Adapted from [241].

To assess the stability of the trimer, the PISA server[45] was used to calculate the buried surface area upon oligomerization which, for SfEiger, is ~1930 $Å^2$. This value is comparable to that of other members of the TNF superfamily such as TL1A (1977 $Å^2$) or TRAIL (2261 $Å^2$). Experimental validation of the stability of the trimer was determined by differential scanning fluorescence: the thermal denaturation curve shows a melting temperature of >80C, indicating a highly stable protein.

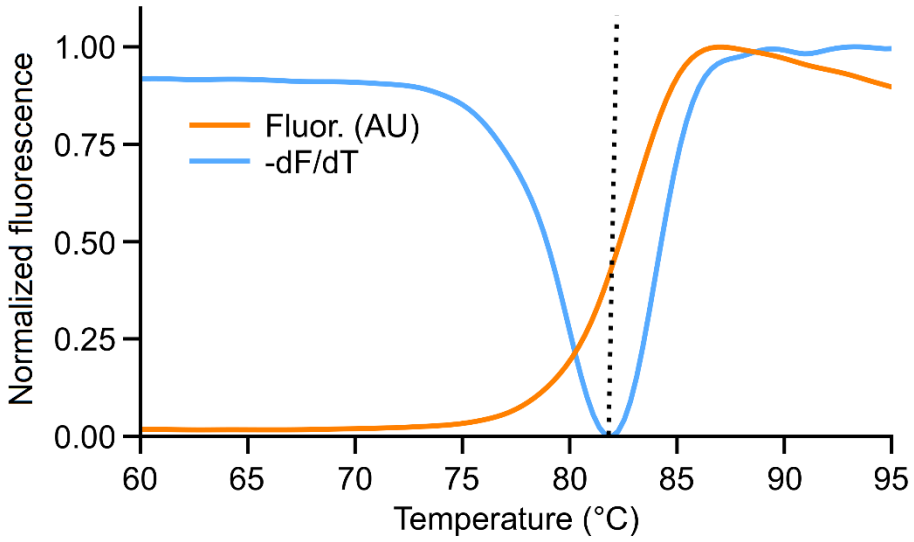


Figure 98: Differential scanning fluorescence of SfEiger. The orange line indicates the fluorescent signal, while the blue its derivative as function of the temperature. Adapted from [241].

To further assess the stability of the SfEiger trimer, an explicit-solvent molecular dynamic simulations was performed (**Figure 99**). The average RMSF over 100ns were very low and for most residues below 0.1nm with the exception of some conformational flexibility in the DE loop, as well as in the A'A* loop. This is consistent with the lack of a clear electron density in the crystal in this region.

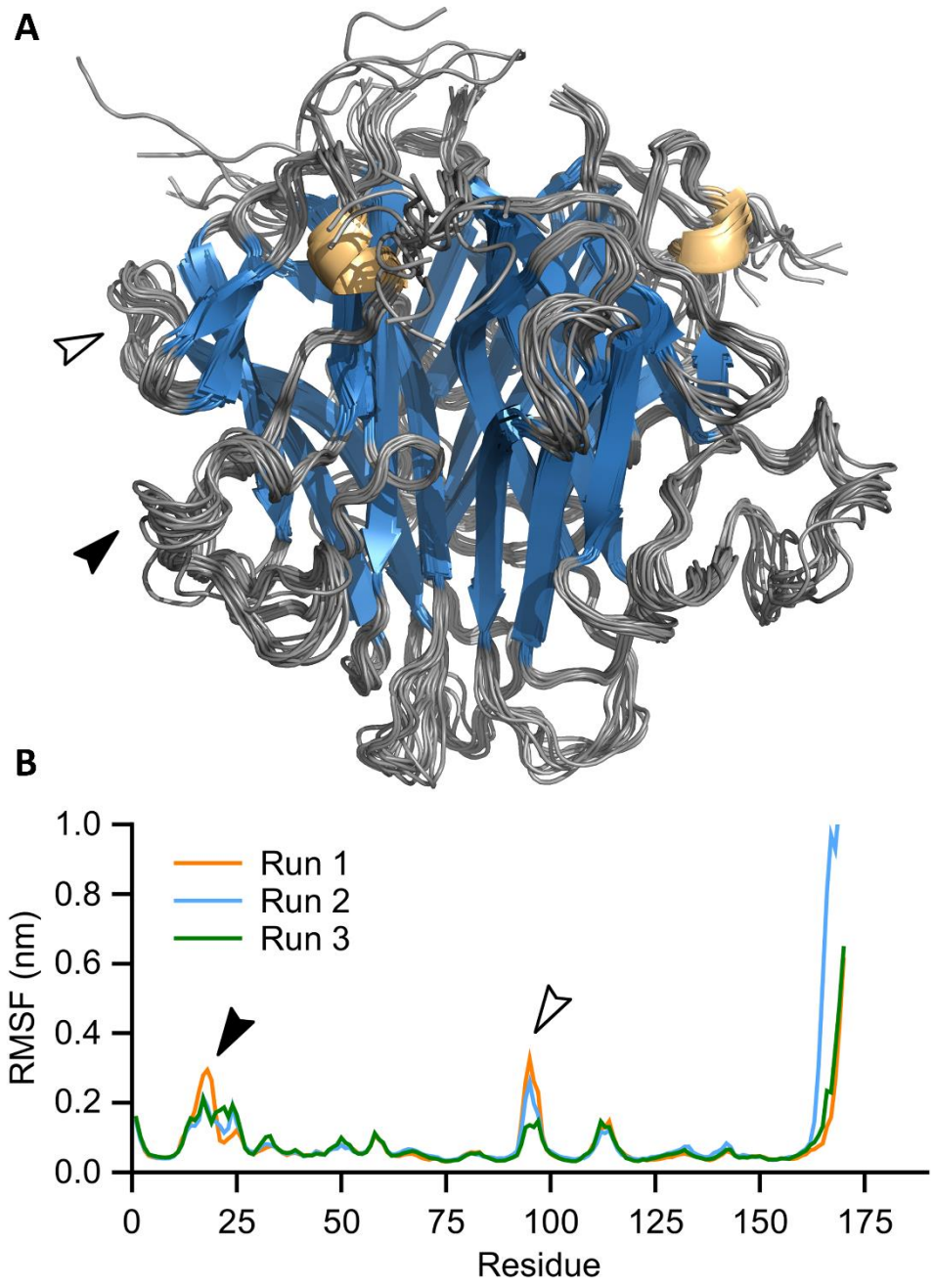


Figure 99: Molecular dynamic simulations. (A) Superposition of 10 conformers extracted at 10ns intervals during the simulation. (B) Per-residue RMSF of Cα atoms over 3 replicate simulations. The white arrow points to the DE loop, the black arrow to the A'A* loop. Molecular dynamics simulations were performed by Dr. Max Renner. Adapted from [241].

5.2.4 Model of SfEiger – TNF receptor interactions

TNF receptors with a single CRD such as Wengen and Grindelwald[12], [233], possess a conserved β-hairpin preceding variable helix-loop-helix motifs[252].

Taking into account the existence of structures of TNF ligands bound to their receptors with one CRD (such as BAFF and APRIL) and the available information on structural conservation (***Figure 86***) it was possible to generate a coarse model of this interaction in the case of SfEiger. This was done by docking the structure of the complex between BAFF and its receptor (PDB ID: 4V46) onto that of SfEiger.

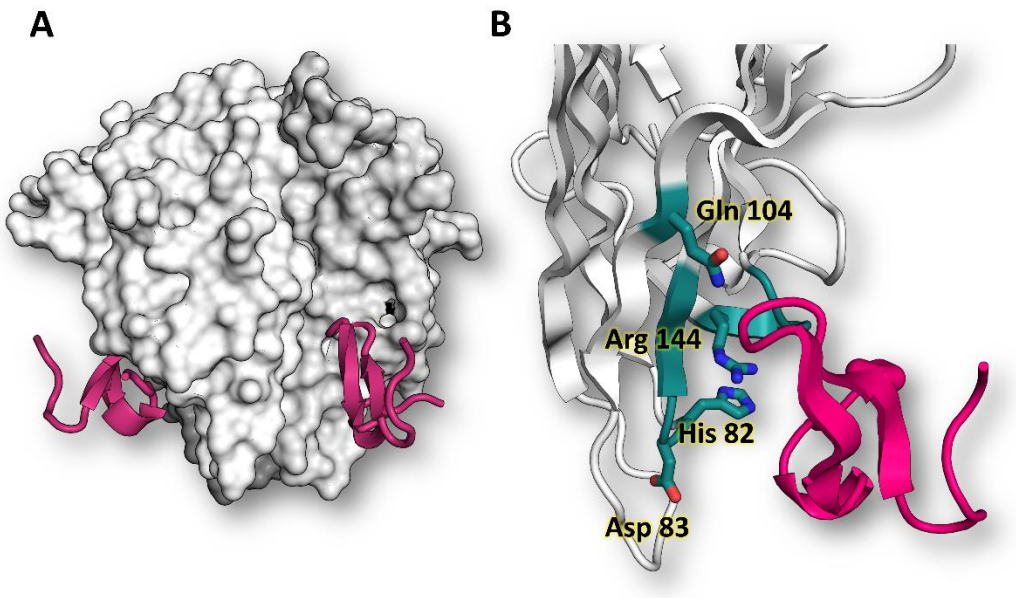


Figure 100: Model of the interaction of SfEiger with its receptor. (**A**) Overview of the interaction. The receptor is coloured in pink, while the SfEiger trimer in white. (**B**) Closeup of the putative interface region. Colouring follows the previous panel. In teal are conserved residues of SfEiger. Adapted from [241].

In the resulting model (***Figure 100***), three copies of the prototype receptor bind SfEiger at the bottom of the inverted pyramid formed by the SfEiger trimer. In particular, engagement of the receptor is primarily via the CD and GH loops, as well as the D strand of SfEiger. Several residues (i.e. His82, Asp83, Gln104, Arg144) are fully conserved in Eiger proteins (***Figure 86***). This suggests they are important for receptor engagement. Furthermore, as mentioned before, given its pH sensing ability, His82 could have a role in pH-dependent modulation of binding.

A closer look at the crystal packing over multiple unit cells shows sheets of trimeric SfEiger all oriented in the same way, forming a densely packed trigonal lattice where the DE loop and the C-terminal tail mediate interaction between trimers (***Figure 101***).

In this arrangement, all the predicted binding interfaces on SfEiger are available and can be contacted by a modelled receptor without major clashes. This tight packing is possible only with a single-CRD receptor (as is the case for Eiger) and would not be compatible with receptors with multiple CRDs, such as the majority of mammalian TNF receptors. Indeed the existence of such lattices has been suggested as a mechanism for clustering and enhancement of signalling in TNFs[218].

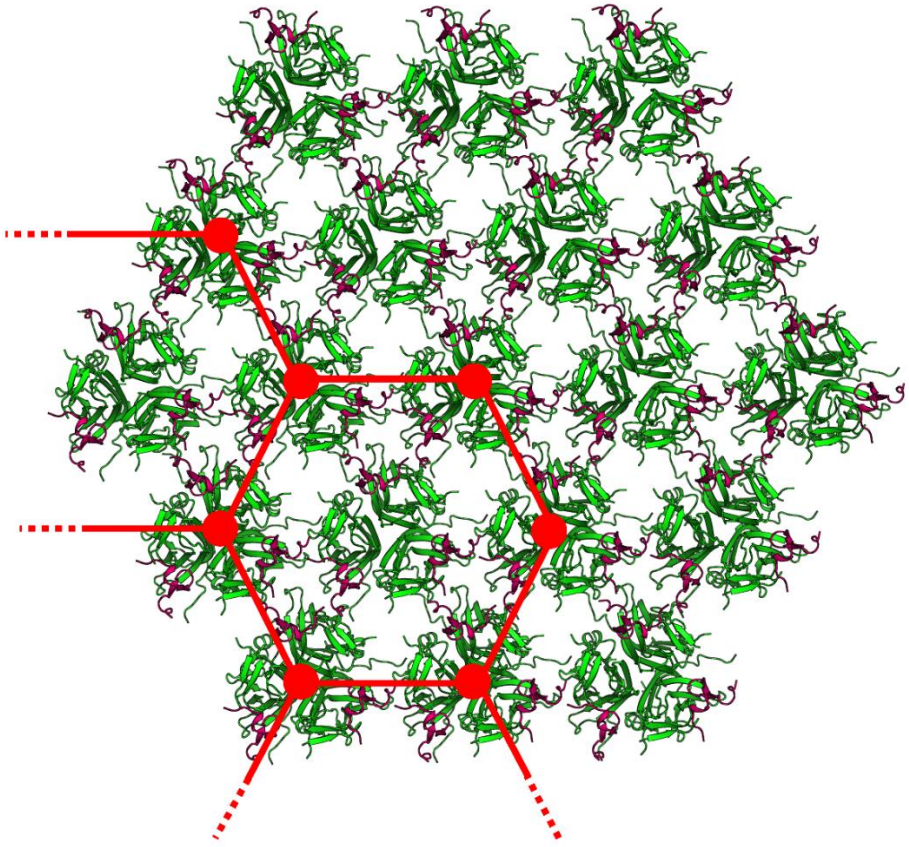


Figure 101: A sheet of SfEiger (green). The modelled receptor is coloured in pink to show the lack of significant clashes. A potential sparser arrangement is indicated with a red line and dots (see **5.3**). Adapted from [241].

5.3 Discussion

Members of the TNF superfamily have crucial roles in the regulation of a variety of cellular processes such as development, cellular proliferation and immunity. Through the action of ligand molecules, either anchored to the membrane or soluble, the signal is relayed to neighbouring cells upon binding of their receptor(s). Due to their involvement in programmed cell death and in immunity, these elements have been investigated as target for therapeutics in cancer and autoimmune diseases. One example is the use of antibodies that target cancer cells, coupled with apoptosis-inducing TNFs[253].

The basic pathway that ends with the activation of the JNK cascade is conserved throughout evolution down to the arthropod *phylum*. In arthropods a single TNF ligand, Eiger, has been identified and is a well characterized system to study the aforementioned processes in *Drosophila melanogaster*[254].

The fall armyworm (*Spodoptera frugiperda*) is major pest in agriculture, currently impacting food supply in Africa and Asia[239], [255]. During the course of an unrelated study, the TNF domain of *S. frugiperda* Eiger was serendipitously crystallized. By adopting a somewhat unorthodox approach, it was possible do determine its structure starting from the electron density alone. The crystal structure of SfEiger shows the canonical trimeric arrangement, shared even with distantly related members of the superfamily. This indicates that the trimerization of TNF ligands has been conserved since before the phylogenetic split between protostomes (from which arthropods descend) and deuterostomes 500 million years ago.

In mammals, it has been reported that the clustering of multiple receptors at the cell surface, in a loose hexagonal lattice, enhances the transduction of signal carried by TNF ligands[218]. In said model, a trimer of TNF receptor sits at each of the vertices of a hexagon, accommodating the apparent symmetry mismatch between TNF receptors (trimeric) and TRAFs (dimers of trimers)[256].

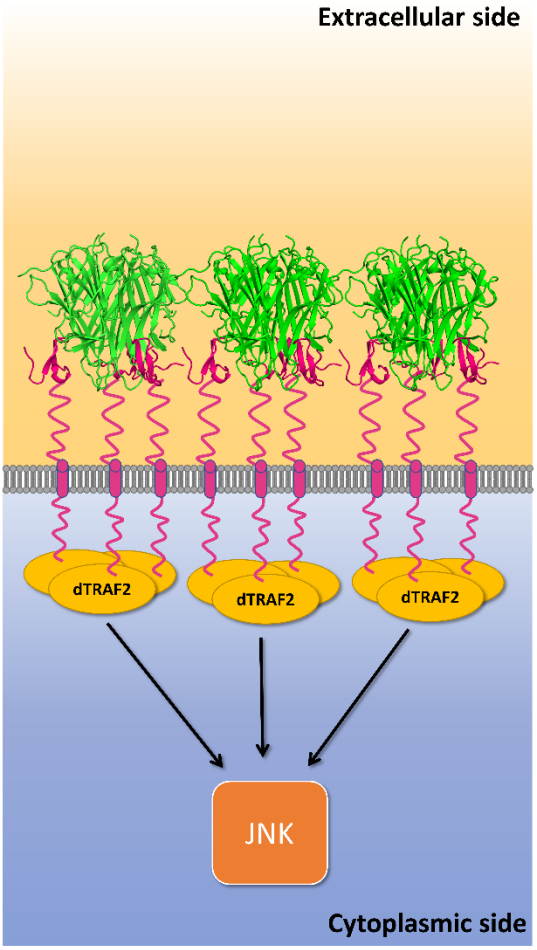


Figure 102: Model of laterally interacting TNF ligand trimers and downstream signalling. Adapted from [241].

In this model, the clustering would be mediated mainly by TRAFs, as the TNF receptors would be too far apart to interact. However, it has also been reported that TNF ligands can interact with each other, bridging this gap. This is the case of BAFF, which can form virus-like clusters (i.e. 60-mers of BAFF in an icosahedral arrangement) at neutral or basic pH. This clustering is dependent on the extended DE loop of BAFF[244], [245].

In this crystal structure of SfEiger, sheets of tightly packed trimers can be observed (***Figure 101***) and the DE loop is again involved in the formation of the contacts between the trimers. It is tempting to speculate, bearing in mind the limitations of the approach, that this packing may point towards a mechanism of signal enhancement at the cell membrane as the arrangement would be able to easily accommodate single CRD receptors. Alternatively, a sparser arrangement

could occur: this would also allow the formation of networks of TRAF dimers on the intracellular side. While these mechanisms would allow for enhancement of binding strength through avidity, more experimental work is required to verify their veracity in a biological context.

During tissue remodelling and development, events of coordinated cell death are necessary to correctly shape tissues. This can be accomplished via apoptosis-induced-apoptosis, a process where group of cells undergo apoptosis following signalling from other apoptotic cells. In Drosophila, the expression of Eiger in apoptotic cells in the posterior compartment of wing disc can induce apoptosis in the neighbouring anterior compartment[232]. It has also been demonstrated that various other organisms employ similar mechanisms of collective, coordinated cell death[231].

As it was possible to initially crystallize SfEiger without overexpressing the protein, a sufficient amount of the protein must have been secreted in the growth medium, hinting at a propagated death-inducing signal. This could be the result of cell stress from overexpression of the original target protein, but also a defence response to the infection of the cells by the baculovirus. Indeed, baculoviruses have evolved a way of counteracting said defence mechanism in the form of the protein p35, which can prevent apoptosis of the host cells by inhibiting pro-apoptotic caspases[257].

Receptors in the plasma membrane that are internalized into endosomes can have an array of functions within a signalling pathway. Other signalling components can also assemble onto endosomes[258]. Earlier experiments using endocytosis inhibitors have shown selective inhibition of response to TNF signalling

molecules[259]. The internalization of TNF receptor 1 is also required during signalling, leading to the formation of the death-inducing signalling complex (DISC)[260], [261]. It has also been shown that *Drosophila* Eiger is internalized and transported to endosomes in a process that involves Deltex[262], and that leads to the initiation of the JNK cascade[226], [251].

As shown previously (***Figure 97***), Histidine residues on the surface of the assembled trimer of SfEiger are abundant (it is, in fact, the second most abundant residue in the protein after Alanine). In SfEiger, His82 is located at the putative receptor site. In *Drosophila* Eiger His342 (the corresponding residue to His82 in SfEiger) would also be located in the centre of the receptor binding surface, together with another histidine, His394 which is part of the G-G' loop (see ***Figure 103***).

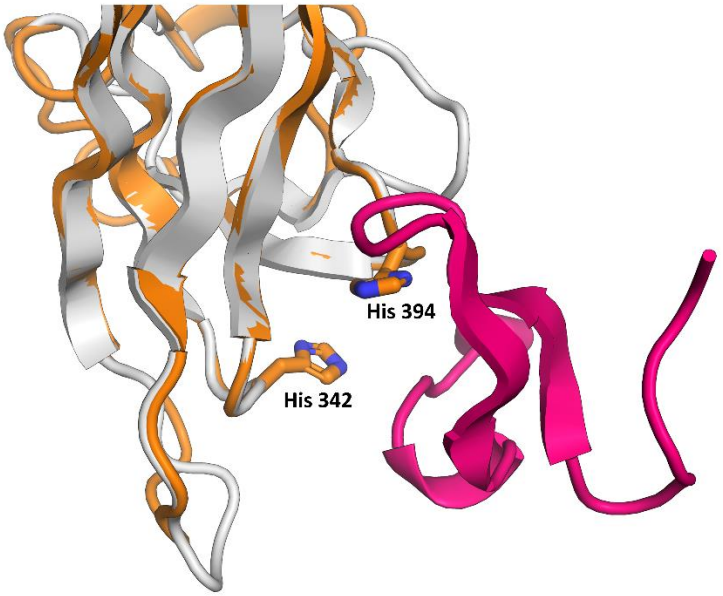


Figure 103: Superimposition of SfEiger and *D.melanogaster* Eiger (model). The *D.melanogaster* Eiger (orange) was modelled with Swiss-model and superimposed onto the structure of SfEiger (white). Histidine residues within the putative receptor binding site are highlighted and the modelled BAFF receptor is shown in magenta.

In the cellular environment, histidine residues can be either neutral or positively charged. Thus, they are known to be key determinants in mediating pH-dependent

conformational changes in proteins (e.g. [263]) upon alteration of the residue charge. The pH-dependent icosahedral viral-like clusters formed by BAFF for instance, have also been shown require His218 (located within the DE flap) [245]. Unlike in BAFF, in SfEiger there are no histidine residues residues within the DE flap and we experimentally demonstrate that formation of these viral-like clusters does not happen (***Figure 95***).

Taking into account these considerations, and the fact that upon endocytosis, Eiger is routed to the low-pH endosomes, it is possible that the some of the histidine residues decorating the surface of SfEiger play a key role in its mechanism. More work, however, will be necessary to determine to what extent.

5.3.1 Future work

Given the serendipitous nature of the work, the scope of the experiments had to be limited. It was eventually possible to clone, express and purify the protein

In order to validate our coarse model, the first thing that could be attempted is to obtain a structure of the complex between SfEiger and one (or both) of its receptors (Wengen and Grindelwald). Design of the constructs would have to be done by relying on published structures of single-CRD receptors (e.g. the BAFF receptor [249]) as well as our data.

Irrespective of whether or not such a complex structure can be obtained, a mutational analysis can be carried out if expression of the receptor can be achieved. The role of histidine residues in the interaction between the two proteins (especially the histidine in the putative binding site) can be assessed via an *in vitro* mutational study, by altering key residues and varying the pH. This could be done by employing Biolayer Interferometry or, more simply, by

immobilizing one of the two proteins (e.g. via strep-tag) and loading the column with the other interaction partner. Upon washing the eluted fraction can be assessed via SDS-PAGE.

Furthermore, the clustering of SfEiger can be investigated by means of Electron Microscopy. If such clustering happens at the cell membrane, it should be possible to image via EM. There are many potential issues that can arise from this approach. Aside from obvious parameters such as concentrations, timing and pH, one other factor to investigate is the experimental system to use: a cell would be the ideal system, but thickness of the sample can prevent imaging if the clustering does not occur in favourable positions. Alternatively, a simplified system where the receptor is embedded in a lipid bilayer could provide better clarity, but the lack of other cellular protein in the system may make the approach unsuccessful if such factors are required.

6 Conclusions and final remarks

In this work I presented three different lines of research: attempts at expressing the hepatitis B virus polymerase, structural studies on complexes between NDM-1 and Fab fragments and the structure of the *Spodoptera frugiperda* Eiger protein.

The first part unfortunately did not progress to an advanced stage, as expression of the protein itself has proven to be more difficult than anticipated. This further emphasizes the challenge that structural characterisation of the hepatitis B virus polymerase poses. In order to investigate the molecular details of this protein, I suggested different approaches that can be potentially undertaken to test for expression of the protein.

The peculiar replication mechanism of the virus, with only a fraction of the virions produced by the infected cell fully assembled, makes purification from a native source highly challenging. Thus, recombinant expression of the protein will have to be pursued. Based on the available information about the replication mechanism, I suggested approaches aimed at supplementing the polymerase with cofactors that may be required to keep the protein in a correctly folded state. This would need to be coupled with extensive, possibly automatised screening in different cell types and, ideally, with a directed evolution approach, screening expression via a reporter system.

All these approaches, however, require significant amounts of manpower, time and money and for this reason all the work on this project had to be stopped.

In the second part of this work, I employed X-ray crystallography and single-particle Cryo-EM to show how two Fab fragments, derived from mice monoclonal antibodies, bind to the New Delhi metallo β-lactamase (NDM-1).

Both of these Fabs bind NDM-1 at a similar epitope, located far from the enzyme active site. Surprisingly, despite the similarities they exhibit a different effect on the enzyme. While they both are capable of inhibiting the β-lactamase activity when the substrate is Nitrocefin (a cephalosporin), with Meropenem (a carbapenem) Fab 5 does not show any activity and Fab 17 appear to enhance it. This is further evidence that the hydrolysis mechanisms used by the enzyme to inactivate the substrates are different.

Furthermore, I show that NDM-1 interacts with palmitoleic acid and preliminary Cryo-EM and size exclusion data suggests the formation of higher order assemblies upon incubation with this lipid.

Unfortunately, due to the impact of COVID-19, it was not possible to further investigate either of these findings. Therefore I suggested some follow-up experiments, which entail the generation of mutants of NDM-1. Analysis of the structures of both Fab complexes suggests that the flexibility of the NDM-1 Ω-loop, as well as the λ-helix may play a role in the enzyme catalytic mechanism.

The original aim of the project was to use these antibodies as therapeutics. By looking at the result I present in this work, however, this appears to be unlikely. Therefore, I suggested new potential ways these antibodies could be repurposed: firstly as a toolkit to study the biology of this enzyme and advance our knowledge of its mechanism. Secondly, as currently all tests to detect metallo β-lactamases expressing bacteria are expensive, unreliable and time consuming, these antibodies could be used as an improved diagnostic tool.

Finally, the biological relevance of the formation of clusters of NDM-1 in the presence of palmitoleic acid could be investigated by means of Cryo-EM. This

could prove to be an attractive target in order to fight the threat of antibiotic resistant bacteria.

The third part of the work I presented stems from the serendipitous discovery of a well-diffracting crystal of a contaminant protein accidentally co-purified with the intended target. In this work I showed an unorthodox method to obtain the structure of *S.frugiperda* Eiger, starting from the diffraction data only. This is the first published structure of a TNF ligand in arthropods.

By analysing the structure of this protein I am able to propose a model for the interaction of Eiger with its receptors. The model suggests the possibility of clustering of the receptor-ligand complex at the membrane. This clustering would allow enhancement of the avidity of the receptor for its ligand. Finally, I suggest that the large number of histidine residues on the surface of the protein is not only the likely factor responsible for the co-purification, but could also play a role as a pH sensor in the mechanism of this ligand.

Further experiments, aimed at obtaining molecular details on the complex with the receptor, could employ X-ray crystallography, as well as Cryo-EM to gather evidence to verify the model outlined here.

7 Copyright info

All figures except the ones listed below are of my creation. Licenses to reuse the other figures for the purpose of this thesis are listed below.

Figure 5: © World Health Organization 2017. CC BY-NC-SA 3.0 IGO (https://creativecommons.org/licenses/by-nc-sa/3.0/igo)

Figure 6: Copyright © 2013 Elsevier B.V. All rights reserved. License Number 4803070431296

Figure 8: Copyright © 2006 Elsevier Inc. All rights reserved. License Number 4803070983785

Figure 9: Copyright © 2005, Springer Nature. License Number 4803080041279

Figure 11: Copyright © 2005, Springer Nature. License Number 4803080688903

Figure 12, **Figure 15**: Copyright © 2015 FUTURE MEDICINE LTD. License number 1026915-1

Figure 13: Permission obtained from publisher for J. C.-Y. Wang, D. G. Nickens, T. B. Lentz, D. D. Loeb, and A. Zlotnick, "Encapsidated hepatitis B virus reverse transcriptase is poised on an ordered RNA lattice," *Proc. Natl. Acad. Sci.*, vol. 111, no. 31, pp. 11329–11334, Aug. 2014.

Figure 16: Copyright © 2010 The Protein Society. License Number 4803091223185

Figure 40: Public domain. From https://en.wikipedia.org/wiki/File:Beta-lactam_antibiotics_example_1.svg

Figure 42: Copyright © 2016, Springer Nature. License Number 4807611075085

Figure 48: Reprinted with permission from P. Linciano, *et al*. “Ten Years with New Delhi Metallo-β-lactamase-1 (NDM-1): From Structural Insights to Inhibitor Design,” ACS Infect. Dis., vol. 5, no. 1, pp. 9–34, Jan. 2019. Copyright © 2019, American Chemical Society

Figure 83, **Figure 85**, **Figure 87**, **Figure 86**, **Figure 88**, **Figure 89**, **Figure 90**, **Figure 91**, **Figure 93**, **Figure 94**, **Figure 95**, **Figure 96**, **Figure 97**, **Figure 98**, **Figure 99**, **Figure 100**, **Figure 101** and **Figure 102**: CC BY 4.0 license http://creativecommons.org/licenses/by/4.0/.

Figure 4: CC BY 4.0 license http://creativecommons.org/licenses/by/4.0/.

8 References

[1] D. H. Mack, W. Bloch, N. Nath, and J. J. Sninsky, “Hepatitis B virus particles contain a polypeptide encoded by the largest open reading frame: a putative reverse transcriptase.,” *J. Virol.*, vol. 62, no. 12, pp. 4786–90, Dec. 1988.

[2] J. Voros *et al.*, “Large-Scale Production and Structural and Biophysical Characterizations of the Human Hepatitis B Virus Polymerase,” *J. Virol.*, vol. 88, no. 5, pp. 2584–2599, 2014.

[3] B. Ayola, P. Kanda, and R. E. Lanford, “High level expression and phosphorylation of hepatitis B virus polymerase in insect cells with recombinant baculoviruses.,” *Virology*, vol. 194, no. 1, pp. 370–3, May 1993.

[4] C. M. Rotondo and G. D. Wright, “Inhibitors of metallo-β-lactamases.,” *Curr. Opin. Microbiol.*, vol. 39, pp. 96–105, Oct. 2017.

[5] V. L. Green, A. Verma, R. J. Owens, S. E. V Phillips, and S. B. Carr, “Structure of New Delhi metallo-β-lactamase 1 (NDM-1).,” *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.*, vol. 67, no. Pt 10, pp. 1160–4, Oct. 2011.

[6] H. Hu *et al.*, “Novel Plasmid and Its Variant Harboring both a bla NDM-1 Gene and Type IV Secretion System in Clinical Isolates of Acinetobacter lwoffii,” *Antimicrob. Agents Chemother.*, vol. 56, no. 4, pp. 1698–1702, Apr. 2012.

[7] A. Makena *et al.*, “Biochemical characterization of New Delhi metallo-β-

lactamase variants reveals differences in protein stability.," *J. Antimicrob. Chemother.*, vol. 70, no. 2, pp. 463–9, Feb. 2015.

[8] I. Stokes-Rees and P. Sliz, "Protein structure determination by exhaustive search of Protein Data Bank derived databases.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 50, pp. 21476–81, Dec. 2010.

[9] S. Kauppila *et al.*, "Eiger and its receptor, Wengen, comprise a TNF-like system in Drosophila.," *Oncogene*, vol. 22, no. 31, pp. 4860–7, Jul. 2003.

[10] A. Corti, G. Fassina, F. Marcucci, E. Barbanti, and G. Cassani, "Oligomeric tumour necrosis factor alpha slowly converts into inactive forms at bioactive levels.," *Biochem. J.*, vol. 284 (Pt 3, pp. 905–10, Jun. 1992.

[11] D. S. Andersen *et al.*, "The Drosophila TNF receptor Grindelwald couples loss of cell polarity and neoplastic growth.," *Nature*, vol. 522, no. 7557, pp. 482–6, Jun. 2015.

[12] G. D. Wiens and G. W. Glenney, "Origin and evolution of TNF and TNF receptor superfamilies.," *Dev. Comp. Immunol.*, vol. 35, no. 12, pp. 1324–35, Dec. 2011.

[13] R. Higuchi, B. Krummel, and R. Saiki, "A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions," *Nucleic Acids Res.*, vol. 16, no. 15, pp. 7351–7367, 1988.

[14] J. Elegheert *et al.*, "Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins.,"

Nat. Protoc., vol. 13, no. 12, pp. 2991–3017, 2018.

[15] N. S. Berrow, D. Alderton, and R. J. Owens, "The precise engineering of expression vectors using high-throughput In-Fusion PCR cloning.," *Methods Mol. Biol.*, vol. 498, pp. 75–90, 2009.

[16] M. Fairhead and M. Howarth, "Site-Specific Biotinylation of Purified Proteins Using BirA," 2015, pp. 171–184.

[17] A. D. Elbein, J. E. Tropea, M. Mitchell, and G. P. Kaushal, "Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I," *J Biol Chem*, vol. 265, no. 26, pp. 15599–605, 1990.

[18] F. H. Niesen, H. Berglund, and M. Vedadi, "The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability," *Nat. Protoc.*, vol. 2, no. 9, pp. 2212–2221, Sep. 2007.

[19] T. S. Walter *et al.*, "A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 61, no. 6, pp. 651–657, Jun. 2005.

[20] L. Govada and N. Chayen, "Choosing the Method of Crystallization to Obtain Optimal Results," *Crystals*, vol. 9, no. 2, p. 106, Feb. 2019.

[21] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker, "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination," *Nat. Methods*, vol. 14, no. 3, pp. 290–296, Mar. 2017.

[22] S. H. W. Scheres, "A Bayesian View on Cryo-EM Structure Determination," *J. Mol. Biol.*, vol. 415, no. 2, pp. 406–418, Jan. 2012.

[23] S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, and D. A. Agard, “MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy,” *Nat. Methods*, vol. 14, no. 4, pp. 331–332, Apr. 2017.

[24] K. Zhang, “Gctf: Real-time CTF determination and correction,” *J. Struct. Biol.*, vol. 193, no. 1, pp. 1–12, Jan. 2016.

[25] F. Sievers *et al.*, “Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.,” *Mol. Syst. Biol.*, vol. 7, p. 539, Oct. 2011.

[26] J. Felsenstein, “PHYLIP - Phylogeny Inference Package (Version 3.2),” *Cladistics*, vol. 5, pp. 164–166, 1989.

[27] D. H. Huson and D. Bryant, “Application of phylogenetic networks in evolutionary studies.,” *Mol. Biol. Evol.*, vol. 23, no. 2, pp. 254–267, Feb. 2006.

[28] Y. J. Huang, T. B. Acton, and G. T. Montelione, “DisMeta: a meta server for construct design and optimization.,” *Methods Mol. Biol.*, vol. 1091, pp. 3–16, 2014.

[29] F. Weissmann *et al.*, “biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes,” *Proc. Natl. Acad. Sci.*, vol. 113, no. 19, pp. E2564–E2569, May 2016.

[30] A. Dong *et al.*, “In situ proteolysis for protein crystallization and structure determination.,” *Nat. Methods*, vol. 4, no. 12, pp. 1019–21, Dec. 2007.

[31] G. Winter, “xia2 : an expert system for macromolecular crystallography

data reduction," *J. Appl. Crystallogr.*, vol. 43, no. 1, pp. 186–190, Feb. 2010.

[32] C. Vonrhein *et al.*, "Data processing and analysis with the autoPROC toolbox.," *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 67, no. Pt 4, pp. 293–302, Apr. 2011.

[33] P. D. Adams *et al.*, "PHENIX : a comprehensive Python-based system for macromolecular structure solution," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 2, pp. 213–221, Feb. 2010.

[34] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read, "Phaser crystallographic software," *J. Appl. Crystallogr.*, vol. 40, no. 4, pp. 658–674, Aug. 2007.

[35] P. Emsley and K. Cowtan, "Coot : model-building tools for molecular graphics," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 60, no. 12, pp. 2126–2132, Dec. 2004.

[36] G. N. Murshudov *et al.*, "REFMAC5 for the refinement of macromolecular crystal structures.," *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 67, no. Pt 4, pp. 355–67, Apr. 2011.

[37] V. B. Chen *et al.*, "MolProbity : all-atom structure validation for macromolecular crystallography," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 1, pp. 12–21, Jan. 2010.

[38] I. J. McGough *et al.*, "Glypicans shield the Wnt lipid moiety to enable signalling at a distance.," *Nature*, Jul. 2020.

[39] G. M. Sheldrick, "Experimental phasing with SHELXC / D / E : combining

chain tracing with density modification," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, no. 4, pp. 479–485, Apr. 2010.

[40] S. Nandakumar, H. Ma, and A. S. Khan, "Whole-Genome Sequence of the Spodoptera frugiperda Sf9 Insect Cell Line.," *Genome Announc.*, vol. 5, no. 34, Aug. 2017.

[41] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool.," *J. Mol. Biol.*, vol. 215, no. 3, pp. 403–10, Oct. 1990.

[42] E. M. Gertz, Y.-K. Yu, R. Agarwala, A. A. Schäffer, and S. F. Altschul, "Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST.," *BMC Biol.*, vol. 4, p. 41, Dec. 2006.

[43] A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, and G. J. Barton, "Jalview Version 2--a multiple sequence alignment editor and analysis workbench," *Bioinformatics*, vol. 25, no. 9, pp. 1189–1191, May 2009.

[44] N. G. A. Abrescia, D. H. Bamford, J. M. Grimes, and D. I. Stuart, "Structure Unifies the Viral Universe," *Annu. Rev. Biochem.*, vol. 81, no. 1, pp. 795–822, Jul. 2012.

[45] E. Krissinel and K. Henrick, "Inference of macromolecular assemblies from crystalline state.," *J. Mol. Biol.*, vol. 372, no. 3, pp. 774–97, Sep. 2007.

[46] "Hepatitis B Fact sheet." [Online]. Available:

www.who.int/mediacentre/factsheets/fs204/en/. [Accessed: 25-Jul-2017].

[47] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA. Cancer J. Clin.*, vol. 61, no. 2, pp. 69–90, Mar. 2011.

[48] D. M. Knipe and P. M. Howley, "Fields Virology." Philadelphia, 2013.

[49] R. H. Hoft, S. C. Pflugfelder, R. K. Forster, S. Ullman, F. M. Polack, and E. R. Schiff, "Clinical evidence for hepatitis B transmission resulting from corneal transplantation.," *Cornea*, vol. 16, no. 2, pp. 132–7, Mar. 1997.

[50] A. M. Cryer and J. C. Imperial, "Hepatitis B in Pregnant Women and their Infants.," *Clin. Liver Dis.*, vol. 23, no. 3, pp. 451–462, 2019.

[51] V. Soriano *et al.*, "Emerging challenges in managing hepatitis B in HIV patients.," *Curr. HIV/AIDS Rep.*, vol. 12, no. 3, pp. 344–52, Sep. 2015.

[52] World Health Organization, *Global hepatitis report, 2017*. 2017.

[53] W. Shi *et al.*, "Hepatitis B virus subgenotyping: history, effects of recombination, misclassifications, and corrections.," *Infect. Genet. Evol.*, vol. 16, pp. 355–61, Jun. 2013.

[54] C.-L. Lin and J.-H. Kao, "Natural history of acute and chronic hepatitis B: The role of HBV genotypes and mutants.," *Best Pract. Res. Clin. Gastroenterol.*, vol. 31, no. 3, pp. 249–255, Jun. 2017.

[55] B. LA Fauquet CM, Mayo MA, Maniloff J, Desselberger U, *Virus Taxonomy, VIIIth Report of the ICTV*. London: Elsevier/Academic Press, 2005.

[56] A. Hodgens and R. Marathi, *Hepatitis B Vaccine*. 2020.

[57] W. F. Carman *et al.*, "Vaccine-induced escape mutant of hepatitis B virus.," *Lancet (London, England)*, vol. 336, no. 8711, pp. 325–9, Aug. 1990.

[58] N. Mentha, S. Clément, F. Negro, and D. Alfaiate, "A review on hepatitis D: From virology to new therapies.," *J. Adv. Res.*, vol. 17, pp. 3–15, May 2019.

[59] C. Seeger and W. S. Mason, "Molecular biology of hepatitis B virus infection," *Virology*, vol. 479–480, pp. 672–686, May 2015.

[60] M. Buti, L. Roade, M. Riveiro-Barciela, and R. Esteban, "Optimal management of chronic hepatitis B patients receiving nucleos(t)ide analogues.," *Liver Int.*, vol. 40 Suppl 1, pp. 15–21, Feb. 2020.

[61] E. De Clercq, G. Férir, S. Kaptein, and J. Neyts, "Antiviral treatment of chronic hepatitis B virus (HBV) infections.," *Viruses*, vol. 2, no. 6, pp. 1279–305, Jun. 2010.

[62] W. S. Mason, G. Seal, and J. Summers, "Virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus.," *J. Virol.*, vol. 36, no. 3, pp. 829–36, Dec. 1980.

[63] S. A. Wynne, R. A. Crowther, and A. G. Leslie, "The crystal structure of the human hepatitis B virus capsid.," *Mol. Cell*, vol. 3, no. 6, pp. 771–80, Jun. 1999.

[64] S. Seitz, S. Urban, C. Antoni, and B. Böttcher, "Cryo-electron microscopy of hepatitis B virions reveals variability in envelope capsid interactions.,"

EMBO J., vol. 26, no. 18, pp. 4160–7, Sep. 2007.

[65] B. Jiang, K. Himmelsbach, H. Ren, K. Boller, and E. Hildt, "Subviral Hepatitis B Virus Filaments, like Infectious Viral Particles, Are Released via Multivesicular Bodies.," *J. Virol.*, vol. 90, no. 7, pp. 3330–3341, Apr. 2015.

[66] M. Bruns, S. Miska, S. Chassot, and H. Will, "Enhancement of Hepatitis B Virus Infection by Noninfectious Subviral Particles," *J. Virol.*, vol. 72, no. 2, pp. 1462 LP – 1468, Feb. 1998.

[67] K. A. Dryden, S. F. Wieland, C. Whitten-Bauer, J. L. Gerin, F. V Chisari, and M. Yeager, "Native hepatitis B virions and capsids visualized by electron cryomicroscopy.," *Mol. Cell*, vol. 22, no. 6, pp. 843–50, Jun. 2006.

[68] J. Beck and M. Nassal, "Hepatitis B virus replication.," *World J. Gastroenterol.*, vol. 13, no. 1, pp. 48–64, 2007.

[69] M. Nassal, "Hepatitis B virus replication: novel roles for virus-host interactions.," *Intervirology*, vol. 42, no. 2–3, pp. 100–16, 1999.

[70] S. Liu, S. S. Y. Koh, and C. G. L. Lee, "Hepatitis B Virus X Protein and Hepatocarcinogenesis," no. Figure 1, pp. 1–14.

[71] B. Rehermann and M. Nascimbeni, "Immunology of hepatitis B virus and hepatitis C virus infection," *Nat. Rev. Immunol.*, vol. 5, no. 3, pp. 215–229, 2005.

[72] C. G. Alexander, M. C. Jürgens, D. A. Shepherd, S. M. V Freund, A. E. Ashcroft, and N. Ferguson, "Thermodynamic origins of protein folding,

allostery, and capsid formation in the human hepatitis B virus core protein.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 30, pp. E2782-91, Jul. 2013.

[73] L.-Y. Mak, D. K.-H. Wong, W.-K. Seto, C.-L. Lai, and M. F. Yuen, "Hepatitis B core protein as a therapeutic target.," *Expert Opin. Ther. Targets*, vol. 21, no. 12, pp. 1153–1159, 2017.

[74] C. S. Coffin, K. Zhou, and N. A. Terrault, "New and Old Biomarkers for Diagnosis and Management of Chronic Hepatitis B Virus Infection.," *Gastroenterology*, vol. 156, no. 2, pp. 355-368.e3, 2019.

[75] M. T. Chen *et al.*, "A function of the hepatitis B virus precore protein is to regulate the immune response to the core antigen.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 41, pp. 14913–8, Oct. 2004.

[76] V. Bruss and D. Ganem, "The role of envelope proteins in hepatitis B virus assembly.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. 3, pp. 1059–63, Feb. 1991.

[77] D. Glebe *et al.*, "Mapping of the hepatitis B virus attachment site by use of infection-inhibiting preS1 lipopeptides and tupaia hepatocytes.," *Gastroenterology*, vol. 129, no. 1, pp. 234–45, Jul. 2005.

[78] M. Doria, N. Klein, R. Lucito, and R. J. Schneider, "The hepatitis B virus HBx protein is a dual specificity cytoplasmic activator of Ras and nuclear activator of transcription factors.," *EMBO J.*, vol. 14, no. 19, pp. 4747–57, Oct. 1995.

[79] S. Fu, R. Zhou, N. Li, Y. Huang, and X. Fan, "Hepatitis B virus X protein

in liver tumor microenvironment," *Tumor Biol.*, vol. 37, no. 12, pp. 15371–15381, Dec. 2016.

[80] K. M. F. Sze, G. K. Y. Chu, J. M. F. Lee, and I. O. L. Ng, "C-terminal truncated hepatitis B virus x protein is associated with metastasis and enhances invasiveness by C-Jun/matrix metalloproteinase protein 10 activation in hepatocellular carcinoma.," *Hepatology*, vol. 57, no. 1, pp. 131–9, Jan. 2013.

[81] K.-Y. Ng *et al.*, "C-terminal truncated hepatitis B virus X protein promotes hepatocellular carcinogenesis through induction of cancer and stem cell-like properties.," *Oncotarget*, vol. 7, no. 17, pp. 24005–17, Apr. 2016.

[82] F. Henkler *et al.*, "Intracellular localization of the hepatitis B virus HBx protein.," *J. Gen. Virol.*, vol. 82, no. Pt 4, pp. 871–82, Apr. 2001.

[83] H.-R. Lee, Y. Y. Cho, G. Y. Lee, D.-G. You, Y. Do Yoo, and Y. J. Kim, "A direct role for hepatitis B virus X protein in inducing mitochondrial membrane permeabilization.," *J. Viral Hepat.*, Nov. 2017.

[84] L. Zhong, W. Shu, W. Dai, B. Gao, and S. Xiong, "Reactive Oxygen Species-Mediated c-Jun NH2-Terminal Kinase Activation Contributes to Hepatitis B Virus X Protein-Induced Autophagy via Regulation of the Beclin-1/Bcl-2 Interaction.," *J. Virol.*, vol. 91, no. 15, Aug. 2017.

[85] H.-T. Zhang *et al.*, "Hepatitis B virus x protein induces autophagy via activating death-associated protein kinase.," *J. Viral Hepat.*, vol. 21, no. 9, pp. 642–9, 2014.

[86] H. Yan *et al.*, "Sodium taurocholate cotransporting polypeptide is a

functional receptor for human hepatitis B and D virus.," *Elife*, vol. 3, pp. 1–28, Nov. 2012.

[87] A. Schulze, P. Gripon, and S. Urban, "Hepatitis B virus infection initiates with a large surface protein-dependent binding to heparan sulfate proteoglycans.," *Hepatology*, vol. 46, no. 6, pp. 1759–68, Dec. 2007.

[88] K. Kuroki, R. Cheung, P. L. Marion, and D. Ganem, "A cell surface protein that binds avian hepatitis B virus particles.," *J. Virol.*, vol. 68, no. 4, pp. 2091 LP – 2096, Apr. 1994.

[89] S. Tong, J. Li, and J. R. Wands, "Carboxypeptidase D is an avian hepatitis B virus receptor.," *J. Virol.*, vol. 73, no. 10, pp. 8696–702, Oct. 1999.

[90] A. Cooper and Y. Shaul, "Clathrin-mediated endocytosis and lysosomal cleavage of hepatitis B virus capsid-like core particles.," *J. Biol. Chem.*, vol. 281, no. 24, pp. 16563–9, Jun. 2006.

[91] C. N. Hayes, Y. Zhang, G. N. Makokha, M. Z. Hasan, M. D. Omokoko, and K. Chayama, "Early events in hepatitis B virus infection: From the cell surface to the nucleus.," *J. Gastroenterol. Hepatol.*, vol. 31, no. 2, pp. 302–9, Feb. 2016.

[92] A. Funk, M. Mhamdi, H. Hohenberg, H. Will, and H. Sirma, "pH-independent entry and sequential endosomal sorting are major determinants of hepadnaviral infection in primary hepatocytes.," *Hepatology*, vol. 44, no. 3, pp. 685–93, Sep. 2006.

[93] H. Li *et al.*, "Nuclear export and import of human hepatitis B virus capsid protein and particles.," *PLoS Pathog.*, vol. 6, no. 10, p. e1001162, Oct.

2010.

[94] J. Lupberger, S. Schaedler, A. Peiran, and E. Hildt, “Identification and characterization of a novel bipartite nuclear localization signal in the hepatitis B virus polymerase.,” *World J. Gastroenterol.*, vol. 19, no. 44, pp. 8000–10, Nov. 2013.

[95] C. Königer, I. Wingert, M. Marsmann, C. Rösler, J. Beck, and M. Nassal, “Involvement of the host DNA-repair enzyme TDP2 in formation of the covalently closed circular DNA persistence reservoir of hepatitis B viruses.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 40, pp. E4244-53, Oct. 2014.

[96] Y. Qi *et al.*, “DNA Polymerase κ Is a Key Cellular Factor for the Formation of Covalently Closed Circular DNA of Hepatitis B Virus.,” *PLoS Pathog.*, vol. 12, no. 10, p. e1005893, Oct. 2016.

[97] Q. Long *et al.*, “The role of host DNA ligases in hepadnavirus covalently closed circular DNA formation.,” *PLoS Pathog.*, vol. 13, no. 12, p. e1006784, 2017.

[98] C. T. Bock *et al.*, “Structural organization of the hepatitis B virus minichromosome.,” *J. Mol. Biol.*, vol. 307, no. 1, pp. 183–96, Mar. 2001.

[99] D. N. Clark and J. Hu, “Unveiling the roles of HBV polymerase for new antiviral strategies.,” *Future Virol.*, vol. 10, no. 3, pp. 283–295, 2015.

[100] A. Kramvis and M. C. Kew, “Structure and function of the encapsidation signal of hepadnaviridae.,” *J. Viral Hepat.*, vol. 5, no. 6, pp. 357–67, Nov. 1998.

[101] J. Hu, D. O. Toft, and C. Seeger, “Hepadnavirus assembly and reverse transcription require a multi-component chaperone complex which is incorporated into nucleocapsids.,” *EMBO J.*, vol. 16, no. 1, pp. 59–68, Jan. 1997.

[102] R. Bartenschlager and H. Schaller, “Hepadnaviral assembly is initiated by polymerase binding to the encapsidation signal in the viral RNA genome.,” *EMBO J.*, vol. 11, no. 9, pp. 3413–20, Sep. 1992.

[103] J. S. Tuttleman, C. Pourcel, and J. Summers, “Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells.,” *Cell*, vol. 47, no. 3, pp. 451–60, Nov. 1986.

[104] T. Watanabe, E. M. Sorensen, A. Naito, M. Schott, S. Kim, and P. Ahlquist, “Involvement of host cellular multivesicular body functions in hepatitis B virus budding.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 24, pp. 10205–10, Jun. 2007.

[105] K. Takahashi *et al.*, “Immunochemical structure of hepatitis B e antigen in the serum.,” *J. Immunol.*, vol. 130, no. 6, pp. 2903–7, Jun. 1983.

[106] G. H. Wang and C. Seeger, “The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis.,” *Cell*, vol. 71, no. 4, pp. 663–70, Nov. 1992.

[107] M. Weber, V. Bronsema, H. Bartos, A. Bosserhoff, R. Bartenschlager, and H. Schaller, “Hepadnavirus P protein utilizes a tyrosine residue in the TP domain to prime reverse transcription.,” *J. Virol.*, vol. 68, no. 5, pp. 2994–9, May 1994.

[108] F. Zoulim and C. Seeger, “Reverse transcription in hepatitis B viruses is primed by a tyrosine residue of the polymerase.,” *J. Virol.*, vol. 68, no. 1, pp. 6–13, Jan. 1994.

[109] G. Radziwill, W. Tucker, and H. Schaller, “Mutational analysis of the hepatitis B virus P gene product: domain structure and RNase H activity.,” *J. Virol.*, vol. 64, no. 2, pp. 613–20, Feb. 1990.

[110] G. H. Wang and C. Seeger, “Novel mechanism for reverse transcription in hepatitis B viruses.,” *J. Virol.*, vol. 67, no. 11, pp. 6507–12, Nov. 1993.

[111] D. D. Loeb, R. C. Hirsch, and D. Ganem, “Sequence-independent RNA cleavages generate the primers for plus strand DNA synthesis in hepatitis B viruses: implications for other reverse transcribing elements.,” *EMBO J.*, vol. 10, no. 11, pp. 3533–40, Nov. 1991.

[112] J. Summers and W. S. Mason, “Replication of the genome of a hepatitis B--like virus by reverse transcription of an RNA intermediate.,” *Cell*, vol. 29, no. 2, pp. 403–15, Jun. 1982.

[113] J. C.-Y. Wang, D. G. Nickens, T. B. Lentz, D. D. Loeb, and A. Zlotnick, “Encapsidated hepatitis B virus reverse transcriptase is poised on an ordered RNA lattice,” *Proc. Natl. Acad. Sci.*, vol. 111, no. 31, pp. 11329–11334, Aug. 2014.

[114] T. A. Steitz, “A mechanism for all polymerases.,” *Nature*, vol. 391, no. 6664, pp. 231–2, Jan. 1998.

[115] A. A. van Dijk, E. V Makeyev, and D. H. Bamford, “Initiation of viral RNA-dependent RNA polymerization.,” *J. Gen. Virol.*, vol. 85, no. Pt 5,

pp. 1077–1093, May 2004.

[116] O. Poch, I. Sauvaget, M. Delarue, and N. Tordo, “Identification of four conserved motifs among the RNA-dependent polymerase encoding elements.,” *EMBO J.*, vol. 8, no. 12, pp. 3867–74, Dec. 1989.

[117] S. A. Jones, D. N. Clark, F. Cao, J. E. Tavis, and J. Hu, “Comparative Analysis of Hepatitis B Virus Polymerase Sequences Required for Viral RNA Binding, RNA Packaging, and Protein Priming,” *J. Virol.*, vol. 88, no. 3, pp. 1564–1572, 2014.

[118] D. N. et al. Clark, J. M. Flanagan, and J. Hu, “Mapping of Functional Subdomains in the Terminal Protein Domain of Hepatitis B Virus Polymerase.,” *J. Virol.*, vol. 91, no. 3, pp. 1–15, Feb. 2017.

[119] R. Bartenschlager and H. Schaller, “The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription.,” *EMBO J.*, vol. 7, no. 13, pp. 4185–92, Dec. 1988.

[120] M. P. Badtke, I. Khan, F. Cao, J. Hu, and J. E. Tavis, “An interdomain RNA binding site on the hepadnaviral polymerase that is essential for reverse transcription,” *Virology*, vol. 390, no. 1, pp. 130–138, Jul. 2009.

[121] Y. Drygin, “Natural covalent complexes of nucleic acids and proteins: some comments on practice and theoryon the path from well-known complexes to new ones,” *Nucleic Acids Res.*, vol. 26, no. 21, pp. 4791–4796, Nov. 1998.

[122] P. R. Daga, J. Duan, and R. J. Doerksen, “Computational model of

hepatitis B virus DNA polymerase: molecular dynamics and docking to understand resistant mutations.," *Protein Sci.*, vol. 19, no. 4, pp. 796–807, Apr. 2010.

[123] A. Bartholomeusz, B. G. Tehan, and D. K. Chalmers, "Comparisons of the HBV and HIV polymerase, and antiviral resistance mutations.," *Antivir. Ther.*, vol. 9, no. 2, pp. 149–60, Apr. 2004.

[124] D. R. Langley *et al.*, "Inhibition of hepatitis B virus polymerase by entecavir.," *J. Virol.*, vol. 81, no. 8, pp. 3992–4001, Apr. 2007.

[125] J. Beck, M. Vogel, and M. Nassal, "dNTP versus NTP discrimination by phenylalanine 451 in duck hepatitis B virus P protein indicates a common structure of the dNTP-binding pocket with other reverse transcriptases.," *Nucleic Acids Res.*, vol. 30, no. 7, pp. 1679–87, Apr. 2002.

[126] Y. Zuo and M. P. Deutscher, "Exoribonuclease superfamilies: structural analysis and phylogenetic distribution.," *Nucleic Acids Res.*, vol. 29, no. 5, pp. 1017–26, Mar. 2001.

[127] J. E. Tavis *et al.*, "The hepatitis B virus ribonuclease H is sensitive to inhibitors of the human immunodeficiency virus ribonuclease H and integrase enzymes.," *PLoS Pathog.*, vol. 9, no. 1, p. e1003125, Jan. 2013.

[128] E. Tramontano, A. Corona, and L. Menéndez-Arias, "Ribonuclease H, an unexploited target for antiviral intervention against HIV and hepatitis B virus.," *Antiviral Res.*, vol. 171, p. 104613, 2019.

[129] V. Soriano, "Hot News: Hepatitis B Gene Therapy Coming to Age.," *AIDS Rev.*, vol. 20, no. 2, pp. 125–127.

[130] G. Lin, K. Zhang, and J. Li, “Application of CRISPR/Cas9 Technology to HBV.,” *Int. J. Mol. Sci.*, vol. 16, no. 11, pp. 26077–86, Nov. 2015.

[131] C. Lee, “CRISPR/Cas9-Based Antiviral Strategy: Current Status and the Potential Challenge.,” *Molecules*, vol. 24, no. 7, Apr. 2019.

[132] S. A. Jones, R. Boregowda, T. E. Spratt, and J. Hu, “In vitro epsilon RNA-dependent protein priming activity of human hepatitis B virus polymerase.,” *J. Virol.*, vol. 86, no. 9, pp. 5134–50, May 2012.

[133] J. Hu and M. Boyer, “Hepatitis B virus reverse transcriptase and epsilon RNA sequences required for specific interaction in vitro.,” *J. Virol.*, vol. 80, no. 5, pp. 2141–50, Mar. 2006.

[134] R. E. Lanford, L. Notvall, and B. Beames, “Nucleotide priming and reverse transcriptase activity of hepatitis B virus polymerase expressed in insect cells.,” *J. Virol.*, vol. 69, no. 7, pp. 4431–9, Jul. 1995.

[135] X. Wang, L. Gao, F. Deng, Y. Zhang, Y. Li, and J. Lin, “High-level production of a functional recombinant hepatitis B virus polymerase in insect cells with a baculovirus expression system.,” *J. Huazhong Univ. Sci. Technolog. Med. Sci.*, vol. 27, no. 3, pp. 269–73, Jun. 2007.

[136] A. Nakamura, N. Tamura, and Y. Yasutake, “Structure of the HIV-1 reverse transcriptase Q151M mutant: insights into the inhibitor resistance of HIV-1 reverse transcriptase and the structure of the nucleotide-binding pocket of Hepatitis B virus polymerase,” *Acta Crystallogr. Sect. F Struct. Biol. Commun.*, vol. 71, no. 11, pp. 1384–1390, Nov. 2015.

[137] J. G. Marblestone, S. C. Edavettal, Y. Lim, P. Lim, X. Zuo, and T. R. Butt,

"Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO.," *Protein Sci.*, vol. 15, no. 1, pp. 182–9, Jan. 2006.

[138] D. R. Smyth, M. K. Mrozkiewicz, W. J. McGrath, P. Listwan, and B. Kobe, "Crystal structures of fusion proteins with large-affinity tags.," *Protein Sci.*, vol. 12, no. 7, pp. 1313–22, Jul. 2003.

[139] X. Wang, X. Qian, H.-C. Guo, and J. Hu, "Heat Shock Protein 90-Independent Activation of Truncated Hepadnavirus Reverse Transcriptase," *J. Virol.*, vol. 77, no. 8, pp. 4471–4480, Apr. 2003.

[140] J. Moroianu, M. Hijikata, G. Blobel, and A. Radu, "Mammalian karyopherin alpha 1 beta and alpha 2 beta heterodimers: alpha 1 or alpha 2 subunit binds nuclear localization signal and beta subunit interacts with peptide repeat-containing nucleoporins.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 92, no. 14, pp. 6532–6, Jul. 1995.

[141] G. Chhetri, P. Kalita, and T. Tripathi, "An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli.," *MethodsX*, vol. 2, pp. 385–91, 2015.

[142] G. Chaga, J. Hopp, and P. Nelson, "Immobilized metal ion affinity chromatography on Co2+-carboxymethylaspartate-agarose Superflow, as demonstrated by one-step purification of lactate dehydrogenase from chicken breast muscle.," *Biotechnol. Appl. Biochem.*, vol. 29 (Pt 1), pp. 19–24, Feb. 1999.

[143] V. N. Danilevich, L. E. Petrovskaya, and E. V. Grishin, "A Highly Efficient Procedure for the Extraction of Soluble Proteins from Bacterial

Cells with Mild Chaotropic Solutions," *Chem. Eng. Technol.*, vol. 31, no. 6, pp. 904–910, Jun. 2008.

[144] C. I. Hammond and M. J. Holland, "Purification of yeast RNA polymerases using heparin agarose affinity chromatography. Transcriptional properties of the purified enzymes on defined templates.," *J. Biol. Chem.*, vol. 258, no. 5, pp. 3230–41, Mar. 1983.

[145] H. Toh, H. Hayashida, and T. Miyata, "Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus," *Nature*, vol. 305, no. 5937, pp. 827–829, Oct. 1983.

[146] M. St Maurice *et al.*, "Domain architecture of pyruvate carboxylase, a biotin-dependent multifunctional enzyme.," *Science*, vol. 317, no. 5841, pp. 1076–9, Aug. 2007.

[147] C. A. Praul, K. D. Brubaker, R. M. Leach, and C. V Gay, "Detection of endogenous biotin-containing proteins in bone and cartilage cells with streptavidin systems.," *Biochem. Biophys. Res. Commun.*, vol. 247, no. 2, pp. 312–4, Jun. 1998.

[148] E. Niedzialkowska *et al.*, "Protein purification and crystallization artifacts: The tale usually not told.," *Protein Sci.*, vol. 25, no. 3, pp. 720–33, Mar. 2016.

[149] L.-T. Sai, Y.-Y. Yao, Y.-Y. Guan, L.-H. Shao, R.-P. Ma, and L.-X. Ma, "Hepatitis B virus infection and replication in a new cell culture system established by fusing HepG2 cells with primary human hepatocytes," *J. Microbiol. Immunol. Infect.*, vol. 49, no. 4, pp. 471–476, Aug. 2016.

[150] M. S. Packer and D. R. Liu, “Methods for the directed evolution of proteins,” *Nat. Rev. Genet.*, vol. 16, no. 7, pp. 379–394, Jul. 2015.

[151] J. O’Neill, *Antimicrobial resistance : tackling a crisis for the health and wealth of nations*. 2014.

[152] J. M. Munita and C. A. Arias, “Mechanisms of Antibiotic Resistance.,” *Microbiol. Spectr.*, vol. 4, no. 2, 2016.

[153] D. C. HODGKIN, “The X-ray analysis of the structure of penicillin.,” *Adv. Sci.*, vol. 6, no. 22, pp. 85–9, Jul. 1949.

[154] K. Bush and P. A. Bradford, “β-Lactams and β-Lactamase Inhibitors: An Overview.,” *Cold Spring Harb. Perspect. Med.*, vol. 6, no. 8, 2016.

[155] E. P. ABRAHAM and E. CHAIN, “An Enzyme from Bacteria able to Destroy Penicillin,” *Nature*, vol. 146, no. 3713, pp. 837–837, Dec. 1940.

[156] S. M. Drawz and R. A. Bonomo, “Three decades of beta-lactamase inhibitors.,” *Clin. Microbiol. Rev.*, vol. 23, no. 1, pp. 160–201, Jan. 2010.

[157] M. S. Butler and D. L. Paterson, “Antibiotics in the clinical pipeline in October 2019.,” *J. Antibiot. (Tokyo).*, Mar. 2020.

[158] R. P. Ambler, “The structure of beta-lactamases.,” *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 289, no. 1036, pp. 321–31, May 1980.

[159] B. G. Hall and M. Barlow, “Revised Ambler classification of β-lactamases,” *J. Antimicrob. Chemother.*, vol. 55, no. 6, pp. 1050–1051, Jun. 2005.

[160] K. Bush and G. A. Jacoby, “Updated functional classification of beta-

lactamases.," *Antimicrob. Agents Chemother.*, vol. 54, no. 3, pp. 969–76, Mar. 2010.

[161] F. Del Bianco *et al.*, "Comparison of Four Commercial Screening Assays for the Detection of blaKPC, blaNDM, blaIMP, blaVIM, and blaOXA48 in Rectal Secretion Collected by Swabs.," *Microorganisms*, vol. 7, no. 12, Dec. 2019.

[162] D. Yong *et al.*, "Characterization of a new metallo-beta-lactamase gene, bla(NDM-1), and a novel erythromycin esterase gene carried on a unique genetic structure in Klebsiella pneumoniae sequence type 14 from India.," *Antimicrob. Agents Chemother.*, vol. 53, no. 12, pp. 5046–54, Dec. 2009.

[163] C. M. Hudson, Z. W. Bent, R. J. Meagher, and K. P. Williams, "Resistance determinants and mobile genetic elements of an NDM-1-encoding Klebsiella pneumoniae strain.," *PLoS One*, vol. 9, no. 6, p. e99209, 2014.

[164] S. Bontron, P. Nordmann, and L. Poirel, "Transposition of Tn125 Encoding the NDM-1 Carbapenemase in Acinetobacter baumannii.," *Antimicrob. Agents Chemother.*, vol. 60, no. 12, pp. 7245–7251, 2016.

[165] P. Nordmann, L. Poirel, T. R. Walsh, and D. M. Livermore, "The emerging NDM carbapenemases.," *Trends Microbiol.*, vol. 19, no. 12, pp. 588–95, Dec. 2011.

[166] L. J. González, G. Bahr, T. G. Nakashige, E. M. Nolan, R. A. Bonomo, and A. J. Vila, "Membrane anchoring stabilizes and favors secretion of New Delhi metallo-β-lactamase.," *Nat. Chem. Biol.*, vol. 12, no. 7, pp. 516–22, 2016.

[167] D. King and N. Strynadka, “Crystal structure of New Delhi metallo-β-lactamase reveals molecular basis for antibiotic resistance.,” *Protein Sci.*, vol. 20, no. 9, pp. 1484–91, Sep. 2011.

[168] J. B. Nielsen and J. O. Lampen, “Membrane-bound penicillinases in Gram-positive bacteria.,” *J. Biol. Chem.*, vol. 257, no. 8, pp. 4490–5, Apr. 1982.

[169] J. B. Nielsen and J. O. Lampen, “Beta-lactamase III of Bacillus cereus 569: membrane lipoprotein and secreted protein.,” *Biochemistry*, vol. 22, no. 20, pp. 4652–6, Sep. 1983.

[170] H. J. Bootsma *et al.*, “Moraxella (Branhamella) catarrhalis BRO beta-lactamase: a lipoprotein of gram-positive origin?,” *J. Bacteriol.*, vol. 181, no. 16, pp. 5090–3, Aug. 1999.

[171] Y. Xie *et al.*, “GPS-Lipid: a robust tool for the prediction of multiple lipid modification sites.,” *Sci. Rep.*, vol. 6, p. 28249, 2016.

[172] J. J. Almagro Armenteros *et al.*, “SignalP 5.0 improves signal peptide predictions using deep neural networks.,” *Nat. Biotechnol.*, vol. 37, no. 4, pp. 420–423, 2019.

[173] A. U. Khan, L. Maryam, and R. Zarrilli, “Structure, Genetics and Worldwide Spread of New Delhi Metallo-β-lactamase (NDM): a threat to public health.,” *BMC Microbiol.*, vol. 17, no. 1, p. 101, 2017.

[174] A. Khajuria, A. K. Praharaj, M. Kumar, and N. Grover, “Presence of a novel variant NDM-10, of the New Delhi metallo-beta-lactamase in a Klebsiella pneumoniae isolate.,” *Indian J. Med. Microbiol.*, vol. 34, no. 1, pp. 121–3.

[175] T. Tada *et al.*, "NDM-12, a novel New Delhi metallo-β-lactamase variant from a carbapenem-resistant Escherichia coli clinical isolate in Nepal.," *Antimicrob. Agents Chemother.*, vol. 58, no. 10, pp. 6302–5, Oct. 2014.

[176] S. Göttig, A. G. Hamprecht, S. Christ, V. A. J. Kempf, and T. A. Wichelhaus, "Detection of NDM-7 in Germany, a new variant of the New Delhi metallo-β-lactamase with increased carbapenemase activity.," *J. Antimicrob. Chemother.*, vol. 68, no. 8, pp. 1737–40, Aug. 2013.

[177] T. Tada *et al.*, "NDM-8 metallo-β-lactamase in a multidrug-resistant Escherichia coli strain isolated in Nepal.," *Antimicrob. Agents Chemother.*, vol. 57, no. 5, pp. 2394–6, May 2013.

[178] A. U. Khan and M. T. Rehman, "Role of Non-Active-Site Residue Trp-93 in the Function and Stability of New Delhi Metallo-β-Lactamase 1," *Antimicrob. Agents Chemother.*, vol. 60, no. 1, pp. 356–360, Jan. 2016.

[179] S. Khan, A. Ali, and A. U. Khan, "Structural and functional insight of New Delhi Metallo β-lactamase-1 variants.," *Future Med. Chem.*, vol. 10, no. 2, pp. 221–229, 2018.

[180] C. D. Livingstone and G. J. Barton, "Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation.," *Comput. Appl. Biosci.*, vol. 9, no. 6, pp. 745–56, Dec. 1993.

[181] H. Feng, X. Liu, S. Wang, J. Fleming, D.-C. Wang, and W. Liu, "The mechanism of NDM-1-catalyzed carbapenem hydrolysis is distinct from that of penicillin or cephalosporin hydrolysis.," *Nat. Commun.*, vol. 8, no. 1, p. 2242, 2017.

[182] M.-N. Lisa *et al.*, “A general reaction mechanism for carbapenem hydrolysis by mononuclear and binuclear metallo-β-lactamases.,” *Nat. Commun.*, vol. 8, no. 1, p. 538, 2017.

[183] P. Linciano, L. Cendron, E. Gianquinto, F. Spyrakis, and D. Tondi, “Ten Years with New Delhi Metallo-β-lactamase-1 (NDM-1): From Structural Insights to Inhibitor Design,” *ACS Infect. Dis.*, vol. 5, no. 1, pp. 9–34, Jan. 2019.

[184] L. Dortet, P. Nordmann, and L. Poirel, “Association of the Emerging Carbapenemase NDM-1 with a Bleomycin Resistance Protein in Enterobacteriaceae and Acinetobacter baumannii,” *Antimicrob. Agents Chemother.*, vol. 56, no. 4, pp. 1693–1697, Apr. 2012.

[185] P. W. Thomas *et al.*, “Characterization of Purified New Delhi Metallo-β-lactamase-1,” *Biochemistry*, vol. 50, no. 46, pp. 10102–10113, Nov. 2011.

[186] F. D’Atri *et al.*, “Targets for the reduction of antibiotic use in humans in the Transatlantic Taskforce on Antimicrobial Resistance (TATFAR) partner countries.,” *Euro Surveill.*, vol. 24, no. 28, 2019.

[187] CDC, “Antibiotic Resistance Threats in the United States,” 2019.

[188] WHO, “Lack of new antibiotics threatens global efforts to contain drug-resistant infections.” [Online]. Available: https://www.who.int/news-room/detail/17-01-2020-lack-of-new-antibiotics-threatens-global-efforts-to-contain-drug-resistant-infections. [Accessed: 19-Mar-2020].

[189] F. L. Gordillo Altamirano and J. J. Barr, “Phage Therapy in the Postantibiotic Era.,” *Clin. Microbiol. Rev.*, vol. 32, no. 2, 2019.

[190] J.-S. Kim *et al.*, “CRISPR/Cas9-Mediated Re-Sensitization of Antibiotic-Resistant Escherichia coli Harboring Extended-Spectrum β-Lactamases.,” *J. Microbiol. Biotechnol.*, vol. 26, no. 2, pp. 394–401, Feb. 2016.

[191] Z. Elgundi, M. Reslan, E. Cruz, V. Sifniotis, and V. Kayser, “The state-of-play and future of antibody therapeutics,” *Adv. Drug Deliv. Rev.*, vol. 122, pp. 2–19, Dec. 2017.

[192] B. François, O. Barraud, and H. S. Jafri, “Antibody-based therapy to combat Staphylococcus aureus infections,” *Clin. Microbiol. Infect.*, vol. 23, no. 4, pp. 219–221, Apr. 2017.

[193] L. Hua *et al.*, “MEDI4893* Promotes Survival and Extends the Antibiotic Treatment Window in a Staphylococcus aureus Immunocompromised Pneumonia Model.,” *Antimicrob. Agents Chemother.*, vol. 59, no. 8, pp. 4526–32, Aug. 2015.

[194] A. Dalhoff, “In vitro and in vivo effect of immunoglobulin G on the integrity of bacterial membranes.,” *Infection*, vol. 12, no. 3, pp. 214–22.

[195] A. Dalhoff, “Synergy between acylureidopenicillins and immunoglobulin G in experimental animals.,” *Am. J. Med.*, vol. 76, no. 3A, pp. 91–100, Mar. 1984.

[196] B. Giwercman, J. W. Rasmussen, O. Cioufu, I. Clemmentsen, H. Schumacher, and N. Høiby, “Antibodies against chromosomal beta-lactamase.,” *Antimicrob. Agents Chemother.*, vol. 38, no. 10, pp. 2306–10, Oct. 1994.

[197] C. H. O’Callaghan, A. Morris, S. M. Kirby, and A. H. Shingler, “Novel

method for detection of beta-lactamases by using a chromogenic cephalosporin substrate.," *Antimicrob. Agents Chemother.*, vol. 1, no. 4, pp. 283–8, Apr. 1972.

[198] Z. N. Adelman and K. M. Myles, "The C-Type Lectin Domain Gene Family in Aedes aegypti and Their Role in Arbovirus Infection.," *Viruses*, vol. 10, no. 7, 2018.

[199] R. Rajamanonmani *et al.*, "On a mouse monoclonal antibody that neutralizes all four dengue virus serotypes.," *J. Gen. Virol.*, vol. 90, no. Pt 4, pp. 799–809, Apr. 2009.

[200] G. Winter *et al.*, "DIALS: implementation and evaluation of a new integration package.," *Acta Crystallogr. Sect. D, Struct. Biol.*, vol. 74, no. Pt 2, pp. 85–97, 2018.

[201] B. L. Nannenga and T. Gonen, "The cryo-EM method microcrystal electron diffraction (MicroED).," *Nat. Methods*, vol. 16, no. 5, pp. 369–379, 2019.

[202] F. S. van de Bovenkamp *et al.*, "Adaptive antibody diversification through N -linked glycosylation of the immunoglobulin variable region," *Proc. Natl. Acad. Sci.*, vol. 115, no. 8, pp. 1901–1906, Feb. 2018.

[203] F. S. van de Bovenkamp, L. Hafkenscheid, T. Rispens, and Y. Rombouts, "The Emerging Importance of IgG Fab Glycosylation in Immunity," *J. Immunol.*, vol. 196, no. 4, pp. 1435–1441, Feb. 2016.

[204] J. E. Raczynska *et al.*, "Flexible loops of New Delhi metallo-β-lactamase modulate its activity towards different substrates," *Int. J. Biol. Macromol.*,

vol. 158, pp. 104–115, Sep. 2020.

[205] B. X. Yan and Y. Q. Sun, “Glycine Residues Provide Flexibility for Enzyme Active Sites,” *J. Biol. Chem.*, vol. 272, no. 6, pp. 3190–3194, Feb. 1997.

[206] S. B. Halstead, “Dengue Antibody-Dependent Enhancement: Knowns and Unknowns.,” *Microbiol. Spectr.*, vol. 2, no. 6, Dec. 2014.

[207] V. Monje-Galvan and G. A. Voth, “Binding mechanism of the matrix domain of HIV-1 gag on lipid membranes,” *Elife*, vol. 9, Aug. 2020.

[208] E. L. Wu *et al.*, “E. coli outer membrane and interactions with OmpLA.,” *Biophys. J.*, vol. 106, no. 11, pp. 2493–502, Jun. 2014.

[209] C. P. Schultz *et al.*, “Evidence for a new type of outer membrane lipid in oral spirochete Treponema denticola. Functioning permeation barrier without lipopolysaccharides.,” *J. Biol. Chem.*, vol. 273, no. 25, pp. 15661–6, Jun. 1998.

[210] M. A. Tahir, R. C. Van Lehn, S. H. Choi, and A. Alexander-Katz, “Solvent-exposed lipid tail protrusions depend on lipid membrane composition and curvature,” *Biochim. Biophys. Acta - Biomembr.*, vol. 1858, no. 6, pp. 1207–1215, Jun. 2016.

[211] T. Reddy and M. S. P. Sansom, “The Role of the Membrane in the Structure and Biophysical Robustness of the Dengue Virion Envelope.,” *Structure*, vol. 24, no. 3, pp. 375–82, Mar. 2016.

[212] L. Polonelli *et al.*, “Antibody complementarity-determining regions (CDRs) can display differential antimicrobial, antiviral and antitumor

activities.," *PLoS One*, vol. 3, no. 6, p. e2371, Jun. 2008.

[213] J. H. Lorent *et al.*, "Structural determinants and functional consequences of protein affinity for membrane rafts," *Nat. Commun.*, vol. 8, no. 1, p. 1219, Dec. 2017.

[214] M. Bramkamp and D. Lopez, "Exploring the existence of lipid rafts in bacteria.," *Microbiol. Mol. Biol. Rev.*, vol. 79, no. 1, pp. 81–100, Mar. 2015.

[215] B. B. Aggarwal, "Signalling pathways of the TNF superfamily: a double-edged sword.," *Nat. Rev. Immunol.*, vol. 3, no. 9, pp. 745–56, Sep. 2003.

[216] C. F. Ware, "The TNF Superfamily-2008.," *Cytokine Growth Factor Rev.*, vol. 19, no. 3–4, pp. 183–6.

[217] R. M. Locksley, N. Killeen, and M. J. Lenardo, "The TNF and TNF receptor superfamilies: integrating mammalian biology.," *Cell*, vol. 104, no. 4, pp. 487–501, Feb. 2001.

[218] É. S. Vanamee and D. L. Faustman, "Structural principles of tumor necrosis factor superfamily signaling.," *Sci. Signal.*, vol. 11, no. 511, Jan. 2018.

[219] M. J. Eck and S. R. Sprang, "The structure of tumor necrosis factor-alpha at 2.6 A resolution. Implications for receptor binding.," *J. Biol. Chem.*, vol. 264, no. 29, pp. 17595–605, Oct. 1989.

[220] C. Zhan *et al.*, "Biochemical and structural characterization of the human TL1A ectodomain.," *Biochemistry*, vol. 48, no. 32, pp. 7636–45, Aug. 2009.

[221] H. J. A. Wallweber, D. M. Compaan, M. A. Starovasnik, and S. G. Hymowitz, “The crystal structure of a proliferation-inducing ligand, APRIL.,” *J. Mol. Biol.*, vol. 343, no. 2, pp. 283–90, Oct. 2004.

[222] S. G. Hymowitz *et al.*, “The crystal structures of EDA-A1 and EDA-A2: splice variants with distinct receptor specificity.,” *Structure*, vol. 11, no. 12, pp. 1513–20, Dec. 2003.

[223] J.-L. Bodmer, P. Schneider, and J. Tschopp, “The molecular architecture of the TNF superfamily.,” *Trends Biochem. Sci.*, vol. 27, no. 1, pp. 19–26, Jan. 2002.

[224] M. F. Flajnik and M. Kasahara, “Origin and evolution of the adaptive immune system: genetic events and selective pressures.,” *Nat. Rev. Genet.*, vol. 11, no. 1, pp. 47–59, Jan. 2010.

[225] J. Shklover, F. Levy-Adam, and E. Kurant, “The role of Drosophila TNF Eiger in developmental and damage-induced neuronal apoptosis.,” *FEBS Lett.*, vol. 589, no. 8, pp. 871–9, Apr. 2015.

[226] T. Igaki, J. C. Pastor-Pareja, H. Aonuma, M. Miura, and T. Xu, “Intrinsic tumor suppression and epithelial maintenance by endocytic activation of Eiger/TNF signaling in Drosophila.,” *Dev. Cell*, vol. 16, no. 3, pp. 458–65, Mar. 2009.

[227] J. B. Cordero, J. P. Macagno, R. K. Stefanatos, K. E. Strathdee, R. L. Cagan, and M. Vidal, “Oncogenic Ras diverts a host TNF tumor suppressor activity into tumor promoter.,” *Dev. Cell*, vol. 18, no. 6, pp. 999–1011, Jun. 2010.

[228] E. Moreno, M. Yan, and K. Basler, “Evolution of TNF signaling mechanisms: JNK-dependent apoptosis triggered by Eiger, the Drosophila homolog of the TNF superfamily.,” *Curr. Biol.*, vol. 12, no. 14, pp. 1263–8, Jul. 2002.

[229] T. Igaki *et al.*, “Eiger, a TNF superfamily ligand that triggers the Drosophila JNK pathway.,” *EMBO J.*, vol. 21, no. 12, pp. 3009–18, Jun. 2002.

[230] N. Agrawal *et al.*, “The Drosophila TNF Eiger Is an Adipokine that Acts on Insulin-Producing Cells to Mediate Nutrient Response.,” *Cell Metab.*, vol. 23, no. 4, pp. 675–84, Apr. 2016.

[231] M. Eroglu and W. B. Derry, “Your neighbours matter - non-autonomous control of apoptosis in development and disease.,” *Cell Death Differ.*, vol. 23, no. 7, pp. 1110–8, 2016.

[232] A. Pérez-Garijo, Y. Fuchs, and H. Steller, “Apoptotic cells can induce non-autonomous apoptosis through the TNF pathway.,” *Elife*, vol. 2, p. e01004, Sep. 2013.

[233] G. de Vreede *et al.*, “A Drosophila Tumor Suppressor Gene Prevents Tonic TNF Signaling through Receptor N-Glycosylation.,” *Dev. Cell*, vol. 45, no. 5, pp. 595-605.e4, 2018.

[234] S. Chittaranjan *et al.*, “The Drosophila TIPE family member Sigmar interacts with the Ste20-like kinase Misshapen and modulates JNK signaling, cytoskeletal remodeling and autophagy.,” *Biol. Open*, vol. 4, no. 5, pp. 672–84, Apr. 2015.

[235] G. Bidla, M. S. Dushay, and U. Theopold, “Crystal cell rupture after injury in Drosophila requires the JNK pathway, small GTPases and the TNF homolog Eiger.,” *J. Cell Sci.*, vol. 120, no. Pt 7, pp. 1209–15, Apr. 2007.

[236] D. S. Schneider *et al.*, “Drosophila eiger mutants are sensitive to extracellular pathogens.,” *PLoS Pathog.*, vol. 3, no. 3, p. e41, Mar. 2007.

[237] A. Ratheesh *et al.*, “Drosophila TNF Modulates Tissue Tension in the Embryo to Facilitate Macrophage Invasive Migration.,” *Dev. Cell*, vol. 45, no. 3, pp. 331-346.e7, 2018.

[238] D. T. Babcock, C. Landry, and M. J. Galko, “Cytokine signaling mediates UV-induced nociceptive sensitization in Drosophila larvae.,” *Curr. Biol.*, vol. 19, no. 10, pp. 799–806, May 2009.

[239] R. D. Harrison *et al.*, “Agro-ecological options for fall armyworm (Spodoptera frugiperda JE Smith) management: Providing low-cost, smallholder friendly solutions to an invasive pest.,” *J. Environ. Manage.*, vol. 243, pp. 318–330, Aug. 2019.

[240] E. Huhtamo *et al.*, “Characterization of a novel flavivirus from mosquitoes in northern europe that is related to mosquito-borne flaviviruses of the tropics.,” *J. Virol.*, vol. 83, no. 18, pp. 9532–40, Sep. 2009.

[241] M. Bertinelli, G. C. Paesen, J. M. Grimes, and M. Renner, “High-resolution crystal structure of arthropod Eiger TNF suggests a mode of receptor engagement and altered surface charge within endosomes,” *Commun. Biol.*, vol. 2, no. 1, p. 293, Dec. 2019.

[242] B. Ng *et al.*, “Increased noncanonical splicing of autoantigen transcripts

provides the structural basis for expression of untolerized epitopes.," *J. Allergy Clin. Immunol.*, vol. 114, no. 6, pp. 1463–70, Dec. 2004.

[243] L. Holm and L. M. Laakso, "Dali server update.," *Nucleic Acids Res.*, vol. 44, no. W1, pp. W351-5, 2016.

[244] Y. Liu, L. Xu, N. Opalka, J. Kappler, H. B. Shu, and G. Zhang, "Crystal structure of sTALL-1 reveals a virus-like assembly of TNF family ligands.," *Cell*, vol. 108, no. 3, pp. 383–94, Feb. 2002.

[245] T. G. Cachero *et al.*, "Formation of virus-like clusters is an intrinsic property of the tumor necrosis factor family member BAFF (B cell activating factor).," *Biochemistry*, vol. 45, no. 7, pp. 2006–13, Feb. 2006.

[246] S. G. Hymowitz *et al.*, "The Crystal Structures of EDA-A1 and EDA-A2," *Structure*, vol. 11, no. 12, pp. 1513–1520, Dec. 2003.

[247] A. Lammens *et al.*, "Crystal structure of human TWEAK in complex with the Fab fragment of a neutralizing antibody reveals insights into receptor binding.," *PLoS One*, vol. 8, no. 5, p. e62697, 2013.

[248] S. G. Hymowitz *et al.*, "Structures of APRIL-receptor complexes: like BCMA, TACI employs only a single cysteine-rich domain for high affinity ligand binding.," *J. Biol. Chem.*, vol. 280, no. 8, pp. 7218–27, Feb. 2005.

[249] H. M. Kim *et al.*, "Crystal structure of the BAFF-BAFF-R complex and its implications for receptor activation.," *Nat. Struct. Biol.*, vol. 10, no. 5, pp. 342–8, May 2003.

[250] J. Srivastava, D. L. Barber, and M. P. Jacobson, "Intracellular pH sensors: design principles and functional significance.," *Physiology (Bethesda).*,

vol. 22, pp. 30–9, Feb. 2007.

[251] W. Ruan, A. Srinivasan, S. Lin, K.-I. Kara, and P. A. Barker, “Eiger-induced cell death relies on Rac1-dependent endocytosis.,” *Cell Death Dis.*, vol. 7, p. e2181, Apr. 2016.

[252] C. Bossen and P. Schneider, “BAFF, APRIL and their receptors: structure, function and signaling.,” *Semin. Immunol.*, vol. 18, no. 5, pp. 263–75, Oct. 2006.

[253] E. Bremer, “Targeting of the tumor necrosis factor receptor superfamily for cancer immunotherapy.,” *ISRN Oncol.*, vol. 2013, p. 371854, 2013.

[254] T. Igaki and M. Miura, “The Drosophila TNF ortholog Eiger: emerging physiological roles and evolution of the TNF system.,” *Semin. Immunol.*, vol. 26, no. 3, pp. 267–74, Jun. 2014.

[255] A. Silver, “Caterpillar’s devastating march across China spurs hunt for native predator,” *Nature*, vol. 570, no. 7761, pp. 286–287, Jun. 2019.

[256] J. Napetschnig and H. Wu, “Molecular basis of NF-κB signaling.,” *Annu. Rev. Biophys.*, vol. 42, pp. 443–68, 2013.

[257] R. J. Clem, “Baculoviruses and apoptosis: a diversity of genes and responses.,” *Curr. Drug Targets*, vol. 8, no. 10, pp. 1069–74, Oct. 2007.

[258] J. Cendrowski, A. Mamińska, and M. Miaczynska, “Endocytic regulation of cytokine receptor signaling.,” *Cytokine Growth Factor Rev.*, vol. 32, pp. 63–73, 2016.

[259] J. R. Bradley, D. R. Johnson, and J. S. Pober, “Four different classes of inhibitors of receptor-mediated endocytosis decrease tumor necrosis

factor-induced gene expression in human endothelial cells.," *J. Immunol.*, vol. 150, no. 12, pp. 5544–55, Jun. 1993.

[260] W. Schneider-Brachert *et al.*, "Compartmentalization of TNF receptor 1 signaling: internalized TNF receptosomes as death signaling vesicles.," *Immunity*, vol. 21, no. 3, pp. 415–28, Sep. 2004.

[261] S. Schütze, V. Tchikov, and W. Schneider-Brachert, "Regulation of TNFR1 and CD95 signalling by receptor compartmentalization.," *Nat. Rev. Mol. Cell Biol.*, vol. 9, no. 8, pp. 655–62, Aug. 2008.

[262] D. Dutta, A. Singh, M. S. Paul, V. Sharma, M. Mutsuddi, and A. Mukherjee, "Deltex interacts with Eiger and consequently influences the cell death in Drosophila melanogaster.," *Cell. Signal.*, vol. 49, pp. 17–29, Sep. 2018.

[263] Y. Vercoulen *et al.*, "A Histidine pH sensor regulates activation of the Ras-specific guanine nucleotide exchange factor RasGRP1," *Elife*, vol. 6, Sep. 2017.

9 Appendix: List of relevant sequences

9.1 Table of sequences used in the generation of the phylogenetic tree in ***Figure 7***

Accession number	Organism
LC519823.1	Human genotype A
HE974381.1	Human genotype A1
HE974383.1	Human genotype A2
LC500247.1	Human genotype B
LC519822.1	Human genotype C
LC519824.1	Human genotype D
HE974379.1	Human genotype D3
HE974382.1	Human genotype D4
LC513657.1	Human genotype E
HM585193.1	Human genotype F
HE974369.1	Human genotype F2
HE974368.1	Human genotype F4
AB625343.1	Human genotype G
LC491577.1	Human genotype H
D00630.1	Human subtype adr
D00329.1	Human subtype adw
NC_003977.2	Human subtype ayw
AF193864.1	Orangutan
NC_020881.1	Long-fingered bat
NC_038503.1	Pomona bat
NC_024444.1	Horseshoe bat
NC_024445.1	Tent-making bat
NC_004107.1	Woodchuck
NC_028129.1	Woolly monkey
MH700450.1	Tibetan frog
NC_027922.1	White sucker
NC_040719.1	Domestic cat
NC_001484.1	Ground squirrel
NC_043528.1	Capuchin monkey
NC_016561.1	Parrot
NC_001486.1	Heron
AJ251937.1	Stork
NC_005950.1	Snow goose
NC_005890.1	Sheldgoose
NC_005888.1	Ross's goose
AJ441113.1	Crane
AY392760.1	Duck

9.2 Table of PDB codes of structures used in the structural phylogenetic tree in ***Figure 91***

PDB identifier	Name
4mxv	Human Lymphotoxin alpha
1tnf	Human Tumour necrosis factor
4mxw	Human Lymphotoxin beta
2hev	Human OX40 ligand
3lkj	Human CD40 ligand
5l19	Human Fas ligand
6bwv	Human CD137 ligand
1dg6	TNF-related apoptosis-inducing ligand
3urf	Human Receptor activator of nuclear factor kappa-B ligand
4ht1	Human TNF-related weak inducer of apoptosis
4zch	Human A proliferation-inducing ligand
5y9j	Human B-cell activating factor
4en0	Human LIGHT
2rjl	Human Vascular endothelial growth inhibitor
2r32	Human TNF superfamily member 18
1rj7	Human Ectodysplasin A

9.3 Sequences of Fab 5 and Fab 17

Fab 5 light chain :

DIVLTQSPASLAVSLGQRATISCRASESVDSYDNSFMHWYQQKPGQPPKL

LIYRASNLESGIPARFSGSGSRTDFTLTINPVEADDVATYYCQQSNEDPL

TFGAGTKLELKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINV

KWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEA

THKTSTSPIVKSFNRNEC

Fab 5 heavy chain:

EVQLQQSGPELVKTGASVKISCKASGYAFTGYYMYWVKQSHGKSLEWIGY

ISCNNGATSYNQKFKGKATFTVDTSSSTAYMQFNSLTSEDSAVYYCARGG

YSSWFAYWGQGTLVTVSAAKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGY

FPEPVTVTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVPSSTWPSETVTC

NVAHPASSTKVDKKIVPRDCGKHHHHHH

Fab 17 light chain :

DIVLTQSPSSLSASLGERVSLTCRASQEISGYLSWLQQKPDGTIKRLIYA

ASTLDSGVPKRFSGSRSGSDYSLTISSLESEDFADYYCLQYGSYPFTFGS

GTKLEIKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKI

DGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKT

STSPIVKSFNRNEC

Fab 17 heavy chain :

EVQLQQSGPELVKTGASVKISCKASGYSFTTYYMHWVKQSHEKSLEWIGY

ISCYSGATTYNQKFKGKATFTVDTSSSTAYMQFNSLTSEDSAVYYCARDD

YWGQGTLVTVSAAKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVT

VTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVPSSTWPSETVTCNVAHPA

SSTKVDKKIVPRDCGKHHHHHH