

Contents

Abstract	I
Acknowledgement	V
Contents	VII
List of Abbreviations	XI

Chapter 1: Introduction of amoebapore A, an antimicrobial protein produced by *Entamoeba histolytica*

1.1 Introduction of <i>Entamoeba histolytica</i>	2
1.1.1 Amoebiasis and <i>Entamoeba histolytica</i> infection	2
1.1.2 Drugs for amoebiasis treatment	4
1.2 Introduction of amoebapore A	5
1.2.1 Discovery of amoebapore	5
1.2.2 The amoebapore family	6
1.2.3 The importance of positively charged residues for amoebapore A-membrane interaction..	8
1.2.4 pH-dependent self-association and pore formation	13
1.2.5 Interactions of amoebapore A with membranes	17
1.2.6 Other pathogenic factors of <i>Entamoeba histolytica</i>	19
1.3 Structural and biophysical studies of amoebapore A	21
1.3.1. Amoebapore A: a member of the saposin-like protein family	21
1.3.2. pH-dependent dimerisation of amoebapore A.....	27
1.3.3 Interaction of amoebapore A with membranes.....	30
1.4 Research Aims	33

Chapter 2: Cell-free production and structural and functional characterisation of recombinant amoebapore A

2.1 Introduction	36
2.1.1 Cell-free protein synthesis	36
2.1.2 Advantages of using cell-free protein synthesis	37
2.1.3 NMR spectroscopy	38
2.1.3.1 One-dimensional ¹ H NMR experiments	38

2.1.3.2 Two-dimensional NMR	39
2.1.3.3 Three-dimensional HNCA to obtain sequence specific assignments	42
2.1.4 Aims	44
2.2 Methods and Materials	45
2.2.1 Amoebapore A plasmid construct	45
2.2.2 Transformation of plasmids.....	46
2.2.3 Preparation of the <i>E.coli</i> S30 extract.....	46
2.2.4 Preparation of the T7 RNA polymerase	46
2.2.4.1 pKO1166 plasmid construct	46
2.2.4.2 Culture preparation	47
2.2.4.3 Salt precipitation	47
2.2.4.4 DEAE chromatography.....	48
2.2.4.5 Cellulose phosphate affinity chromatography	48
2.2.5 Preparation of the Master Mix.....	49
2.2.6 Preparation of amino acids	49
2.2.7 Cell-free protein synthesis	50
2.2.8 Analytical scale cell-free protein synthesis	52
2.2.8.1 Testing the Master Mix and amino acid stock solutions	52
2.2.8.2 Optimisation of the Rosetta (λ DE3)/pRARE S30 extract volume in the cell-free reaction.....	53
2.2.8.3 Optimisation of the RNasin concentration in the cell-free reaction	53
2.2.8.4 Optimisation of the T7 RNAP concentration in the cell-free reaction	53
2.2.8.5 Optimisation of the amino acid concentration in the cell-free reaction.....	53
2.2.8.6 Unlabelled versus isotope labelled amino acids	54
2.2.9 Preparative cell-free protein synthesis.....	54
2.2.10 Synthesis of isotope labelled protein samples	55
2.2.11 Purification of amoebapore-A	56
2.2.11.1 Immobilised metal ion liquid chromatography.....	56
2.2.11.2 Factor Xa Protease digestion to remove the (His) ₆ -tag.....	57
2.2.11.3 Reverse-phased high performance liquid chromatography	57
2.2.12 NMR spectroscopy	58
2.2.12.1 T_2 measurement.....	58
2.2.12.2 Two dimensional ^1H - ^{15}N HSQC experiment.....	59
2.2.12.3 Three-dimensional HNCA experiment.....	59

2.2.13 Pore-forming assay of amoebapore A	59
2.3 Results	60
2.3.1 Preparation and testing of the components of the cell-free synthesis.....	60
2.3.1.1 Testing of the Master Mix and amino acid stocks in the cell-free approach	60
2.3.1.2 Optimisation of the Rosetta (λ DE3)/pRARE S30 extract volume in the cell-free reaction.....	61
2.3.1.3 Optimisation of the RNase inhibitor concentration in the cell-free reaction.....	61
2.3.1.4 Preparation and optimisation of the T7 RNAP concentration in the cell-free reaction.....	62
2.3.1.5 Crude unlabelled amino acid titration.....	66
2.3.1.6 Unlabelled versus ^{15}N -labelled protein expression.....	67
2.3.2 Preparative scale cell-free protein synthesis.....	68
2.3.3 Purification of amoebapore A	69
2.3.3.1 Immobilised metal affinity chromatography	69
2.3.3.2 Factor Xa digestion.....	71
2.3.3.3 Reverse phased-high performance liquid chromatography	75
2.3.4 NMR spectrometry	76
2.3.4.1 One-dimensional ^1H NMR experiments	76
2.3.4.2 Two-dimensional ^1H - ^{15}N HSQC assignment.....	77
2.3.4.3 Three-dimensional HNCA sequence specific assignments	82
2.3.5 Pore-forming activity of amoebapore A.....	85
2.4 Discussion.....	86
2.4.1 Optimisation and troubleshooting during cell-free synthesis of amoebapore A	86
2.4.2 Factor Xa digestion.....	88
2.4.3 Transamination activities.....	89
2.4.4 Assignment of the NMR resonances representing backbone nuclei	90

Chapter 3: Investigation of amoebapore A pH-dependent activity

3.1 Introduction	92
3.1.1 Amoebapore A dimerisation.....	92
3.1.2 NMR spectroscopy to study amoebapore A dimerisation	92
3.1.2.1 Identification of residues at the dimer interface	92
3.1.2.2 Determination of the pKa values of ionisable groups.....	92
3.1.3 Other biophysical methods	93

3.1.3.1 Size exclusion chromatography	94
3.1.3.2 Small angle X-ray scattering.....	94
3.1.4 Design of amoebapore A mutants	97
3.1.5 Aims	98
3.2 Methods and Materials	99
3.2.1 Sample preparation	99
3.2.2 NMR spectroscopy	99
3.2.2.1 pH- and concentration-dependent chemical shift mapping	99
3.2.2.2 Side-chain assignments of Glu, Asp and His residues.....	101
3.2.2.3 pH titration and the determination of the pKa values	102
3.2.2.4 Estimation of the K_d using NMR spectroscopy	104
3.2.3 Size exclusion chromatography.....	106
3.2.4 Small angle X-ray scattering	106
3.2.5 Site directed mutagenesis	107
3.3 Results	109
3.3.1 Variation in the molecular weight as a function of pH.....	109
3.3.2 The influence of pH and concentration on the chemical shifts of resonances – mapping the dimer interface	110
3.3.3 Side-chain resonance assignments	115
3.3.4 pKa determination of acidic amino acids and the histidine-75	121
3.3.5 pH titration monitored by 2D ^1H - ^{15}N HSQC spectroscopy.....	129
3.3.6 Estimation of K_d for amoebapore A dimerisation	131
3.3.7 Small angle X-ray scattering	134
3.3.8 Cell-free synthesis of amoebapore A mutants.....	138
3.4 Discussion.....	146
3.4.1 pH- and concentration-dependent amoebapore A dimerisation	146
3.4.2 Mechanism of amoebapore A activity.....	149
3.4.2.1 Lys residues	149
3.4.2.2 Asp residues	149
3.4.2.3 Glu16 and Glu62 residues.....	150
3.4.2.4 His75 residue	152

Chapter 4: Interaction of amoebapore A with model membranes

4.1 Introduction	156
4.1.1 Amoebapore A against membranes	156
4.1.2 Quartz crystal microbalance with energy dissipation.....	156
4.1.3 Neutron reflectometry.....	158
4.1.4 Aims	160
4.2 Methods and Materials	161
4.2.1 Liposome preparation	161
4.2.2 Protein preparation	161
4.2.3 Quartz crystal microbalance with dissipation.....	162
4.2.4 Sample preparation for neutron reflectometry.....	165
4.2.5 Neutron reflectometry.....	166
4.3 Results	167
4.3.1 Quartz crystal microbalance with energy dissipation.....	167
4.3.2 Neutron reflectometry.....	174
4.4 Discussion	179
4.4.1 Interactions of amoebapore A with different model membranes	179
4.4.2 Membrane binding mechanism of amoebapore A.....	180

Chapter 5: Conclusion

5.1 Mechanism of amoebapore A dimerisation	184
5.2 Amoebapore A interaction with model membranes	188
Appendices:	191
Appendix 1: SDS-PAGE gel electrophoresis	191
Appendix 2: Method of packing cellulose phosphate affinity column.....	191
Appendix 3: Assignments of amoebapore A resonances acquired in different NMR experiments at pH 3.0.....	192
References	205

List of Abbreviations

Abbreviation	Definition
AMP	Adenosine 3'5' cyclic monophosphate
AMPs	Antimicrobial peptides/proteins
ANSTO	Australian Nuclear Science and Technology Organisation
APA-1	Amoebapore A
APD	Antimicrobial Peptide Database
ATP	Adenosine triphosphate
BHK	Baby hamster kidney
CAEphosphate	N-acyl-sphingosylphosphorylethanolamine
CHO	Chinese hamster ovary
CTHSQC	Constant time heteronuclear single quantum coherence
CTP	Cytosine triphosphate
DEAE	Diethylaminoethyl
DEPC	Diethylpyrocarbonate
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DMPG	1,2-dimyristoyl-sn-glycero-3-phospho-rac-(1-glycerol)
DPhyPC	1,2-diphytanoyl-sn-glycero-3-phosphocholine
DPPG	1,2-dipalmitoyl-sn-glycero-3-(1'-rac-glycerol)
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FID	Free induction decay
Gal/GalNAc	Galactose N-acetyl galactosamine carbohydrate
GPI	Glycosylphosphatidylinositol
GTP	Guanosine triphosphate
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
HSQC	Heteronuclear single quantum coherence
IMAC	Immobilised metal ion affinity chromatography
LB	Luria-Bertani
MES	2-[N-morpholino]ethanesulfonic acid
MW	Molecular weight

MWCO	Molecular weight cut off
NMP	Nucleoside monophosphates
NMR	Nuclear magnetic resonance
NTP	Nucleoside triphosphates
PC	Phosphateidylcholine
PE	Phosphatedylethanolamine
PG	Phosphatideylglycerol
PI	Phosphatidylinositol
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPG	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-rac-(1-glycerol)
ppm	Parts per million
PS	Phosphatidylserine
QCM-D	Quartz crystal microbalance with energy dissipation
RBC	Red blood cells
RF	Radio frequency
RNasin	Ribonuclease inhibitor
RP-HPLC	Reverse-phase high performance liquid chromatography
rpm	Revolutions per minute
S30	Sedimentation at 30,000 g
SAPLIP	Saposin-like protein
SAXS	Small angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SLD	Scattering length density
SM	Sphingomyelin
T7 RNAP	T7 RNA polymerase
TFA	Trifluoroacetic acid
T ₀	Time zero
UTP	Uridine triphosphate

1

Introduction of amoebapore A, an antimicrobial protein produced by *Entamoeba histolytica*

1.1 Introduction of *Entamoeba histolytica*

1.1.1 Amoebiasis and *Entamoeba histolytica* infection

Amoebiasis, also termed amoebic dysentery, is a life-threatening parasitic infection caused by the unicellular protozoan parasite *Entamoeba histolytica* [1]. Amoebiasis is the third leading life-threatening parasitic disease after malaria and schistosomiasis [2]. Humans are infected with amoebiasis throughout the world; however, infection predominantly occurs in developing countries in tropical regions with poor sanitary conditions and contaminated drinking water [3]. According to the World Health Organisation, more than 100,000 deaths per year are caused by infection of this parasite [2]. This parasitic disease is also responsible for severely affecting the growth of infected children [4].

Infection by *E. histolytica* occurs by ingestion of mature cysts in contaminated food or water, and the mature cysts can be spread from person to person. After ingestion of the cysts, the parasite inhabits the human colonic lumen, where it feeds primarily on bacteria and cellular debris. Each cyst undergoes excystation and releases eight trophozoites. The trophozoites migrate to the large intestine where they grow and multiply by binary fission to increase their numbers, feeding on bacteria and cell debris [5].

In ~90% of the infected individuals, the trophozoites attach to the mucus and epithelial cells without penetrating the intestinal mucosa. They undergo encystation to become mature cysts that are passed to stools, resulting in asymptomatic carriers. Asymptomatic carriers are often infected by the non-pathogenic *Entamoeba dispar*, which is visually indistinguishable from the pathogenic *E. histolytica*, and is often termed the non-pathogenic *E. histolytica*. In ~10% of the pathogenic *E. histolytica* infections, the trophozoites lyse colonic epithelial cells, penetrate the intestinal mucosa and begin an invasive process. Severe inflammation of the gut causes bloody diarrhoea, abdominal pain, weight loss, fatigue and dehydration, which results in blood loss and depletion of body protein stores (*i.e.*, dysentery). This can be fatal, especially in infants and children. The destruction of the mucosa can lead to the spread of the parasite via the blood stream to other sites such as the lungs, brain, genitals and liver, which are commonly infected, and lead to massive disruption, inflammation and abscesses of host tissue [6-9] (Figures 1.1 & 1.2).

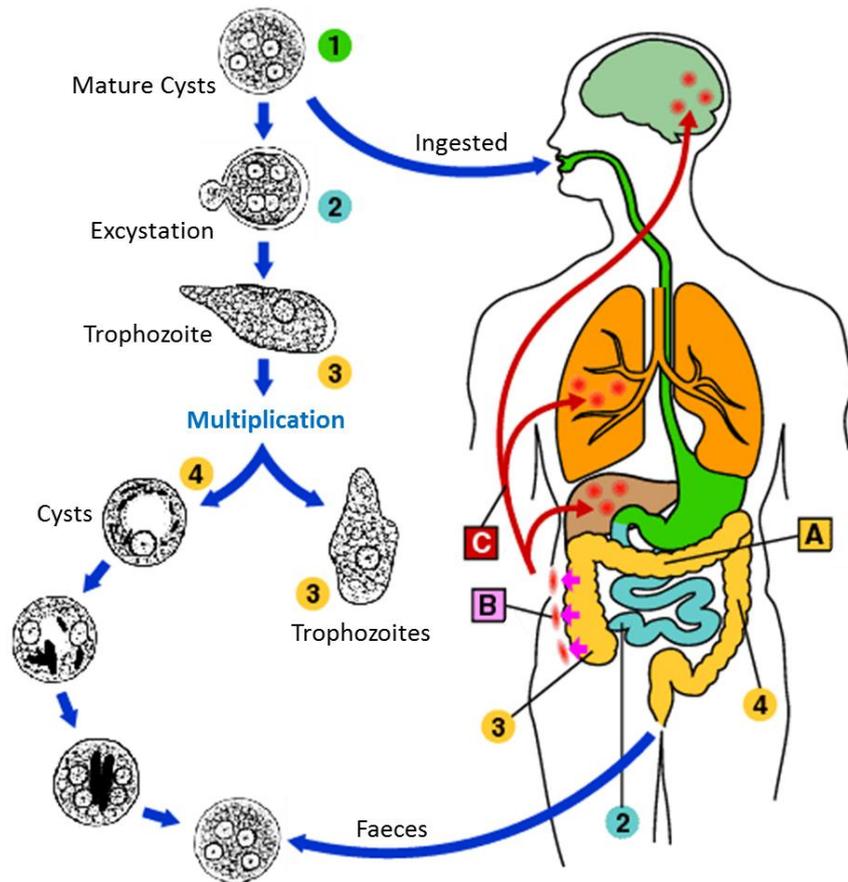


Figure 1.1: Infection cycle of *E. histolytica*. (A) Non-invasive colonisation: the trophozoites remain in the lumen and the cysts are passed to faeces. (B) Intestinal disease: the trophozoites invade intestinal mucosa and results in intestinal dysentery. (C) Extra-intestinal disease: the trophozoites penetrate into the blood stream and migrate to other sites such as the liver, lungs and brain, causing severe infection. (1) Mature cysts; (2) mature cysts undergoing excystation to form trophozoites; (3) trophozoites; (4) trophozoites undergoing encystation to form cysts. The picture was modified from [10].

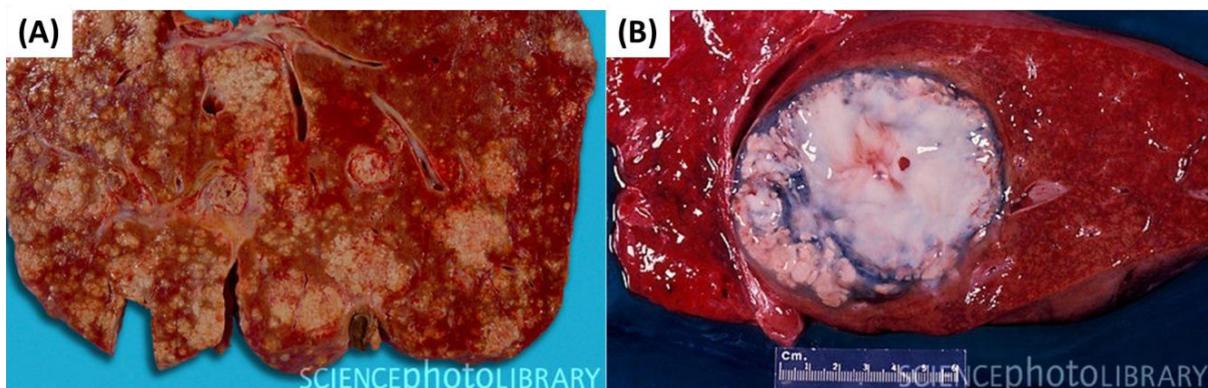


Figure 1.2: Specimens of liver infection and abscess caused by *E. histolytica*. (A) The white patches are regions of damaged tissue caused by the infection (*i.e.*, ulcers). (B) The flask-shaped white ulcer is pus-filled and circumscribed by circular linings formed by the body, preventing the infection from spreading. The pictures were taken from [11].

1.1.2 Drugs for amoebiasis treatment

There are anti-amoebic drugs that target *E. histolytica*. They are generally classified into two groups: tissue amoebicides and luminal amoebicides. The tissue amoebicides kill amoeba in the intestine and host organs, whereas the poorly absorbed luminal amoebicides are active only in the intestinal lumen [3].

Three commonly used tissue amoebicides against invasive amoebiasis are metronidazole, emetine and chloroquine. Metronidazole works against intestinal infection by forming highly reactive free radical species that damage the DNA of the parasite, whereas emetine functions by inhibiting protein synthesis of the trophozoites leading to parasite death [12]. However, the efficiencies of these drugs are limited. For example, metronidazole has no effect on cysts, thus treatment with metronidazole has a high re-infection rate and should be followed by a luminal amoebicide to eliminate luminal parasites. In addition, both metronidazole and emetine have severe side-effects, such as nausea, vomiting, a metallic taste, abdominal pain and headaches, which limit their clinical use [3, 13, 14]. Chloroquine is used to treat hepatic amoebiasis by killing the trophozoites in the liver through inhibition of parasite DNA synthesis. However, chloroquine is less effective in the treatment of intestinal amoebiasis. Luminal amoebicides such as paromomycin, diloxanide furoate and 8-hydroxyquinolones are normally used in conjunction with metronidazole for the treatment of intestinal infections. These drugs are used for treating asymptomatic cyst carriers by killing trophozoites in the lumen, but are less effective against the invasive disease. The luminal amoebicides are especially useful for the control of the spread of infection; however, indiscriminate use of drugs for asymptomatic carriers may lead to drug resistance [6, 15].

To prevent the spread of the parasitic disease, it is important to improve sanitation and hygiene standards. For example, boiling or filtering drinking water to prevent waterborne transmission caused by faecal-contaminated water supplies, frequent hand washing and thorough cooking of food, and proper disposal and disinfection of human waste. However, in third-world countries, sanitation improvement remains an issue, making this disease a lasting international public health problem. Moreover, it is important to extend the access of amoebicides usage in high disease incidence areas to control the infection during the early stages of infection and reduce the risk of the invasive process. Thus there is a challenge to develop new effective drugs against the parasite.

1.2 Introduction of amoebapore A

1.2.1 Discovery of amoebapore

A study performed by Ravdin *et al.* in 1980 showed that invasion of *E. histolytica* to a monolayer of Chinese hamster ovary (CHO) cells involved a contact dependent cytolytic reaction [16]. Initially the parasite establishes an intimate adherence to the target cells which were originally attached to a cover slip, causing the cell membranes to bulge outwardly (*i.e.*, blebbing) and extensive rounding of the cells. The cells were detached from the glass surface and phagocytised by the parasite (Figure 1.3). The invasion process was rapid: approximately 50% of the CHO cells were destroyed within 2 h, and only occurred to cells that were in direct contact with the parasite [16].

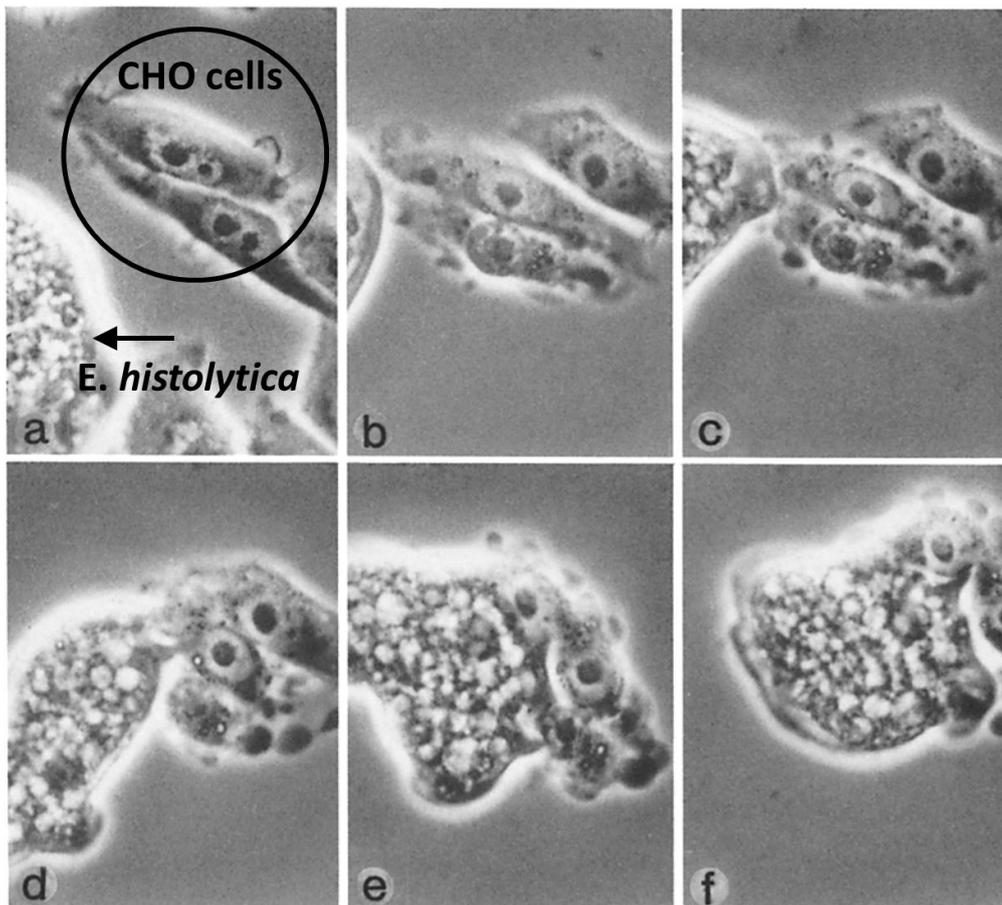


Figure 1.3: Destruction of CHO cells by *E. histolytica*. (a) The CHO cells (indicated by the black circle) were not in direct contact with the parasite and were viable. (b-d) Upon contact with *E. histolytica* (indicated by the arrow), three CHO cells underwent blebbing and rounding (e-f) and were phagocytosed by the parasite. The figure was modified from [16].

Based on the results, it was hypothesised that following the contact between *E. histolytica* and the target cell, a membrane active factor was released from the parasite and transferred to the target cell. The factor is able to spontaneously interact with the cell membrane and affect the ionic permeability of the target cell, leading to extensive cell blebbing followed by phagocytosis and cell death [17]. Studies carried out in 1980s have revealed that the “pathogenic factor”, named amoebapore, appeared to be a protein produced by *E. histolytica* [17, 18]. Amoebapore was found to exist as protein aggregates in target membranes under slightly acidic environments and was able to rapidly increase target membrane conductance upon binding. Based on the observations, it was proposed that amoebapores were able to form pores in membranes. However, these studies provided no detailed description of how the pore formed, and whether amoebapore oligomerisation occurred during pore formation [18-20] or is present in aqueous solution prior to membrane binding [21]. Under conditions of freezing, thawing, or addition of high salt, the aggregates can be dissociated, suggesting reversible oligomerisation [18-20]. More information on the pore-forming properties of amoebapores will be discussed in Sections 1.2.4 and 1.2.5.

1.2.2 The amoebapore family

In 1994, three amoebapore isoforms were isolated from the cytoplasmic granules of *E. histolytica*, and named amoebapore A, B and C [22]. Among the three isoforms, amoebapore A was the most abundant protein and present at a ratio of approximately 35:10:1 (amoebapore A:B:C). All three isoforms showed pore-forming activity. Amoebapore A, B and C all have 77 residues, with molecular weights (MWs) of 8244, 8295.7 and 8025.6 Da, and similar protein isoelectric points calculated to be in the range of 5–6. Amoebapore B and C show 57 and 47% sequence identity to amoebapore A and were predicted to have the same structural organisation. Sequence alignment of the three isoforms (Figure 1.4) shows that all three proteins have six conserved cysteine residues at identical positions. The six cysteines form three disulfide bonds (Figure 1.4). A number of hydrophobic residues are conserved among the isoforms which are dispersed throughout the sequence (Figure 1.4). Since amphipathic helices are the structural elements preferentially used by membrane-penetrating peptides, the hydrophobic regions have been proposed to play important roles in interacting with the membrane surface [23]. Six charged residues Glu16, Lys37, Lys48, Asp55, Glu62 and His75 (numbering for amoebapore A; indicated by red arrows in Figure 1.4) and three uncharged residues Tyr30, Gly40 and Gly53 (Figure 1.4), are conserved across all three isoforms, suggesting these residues may be functionally important.

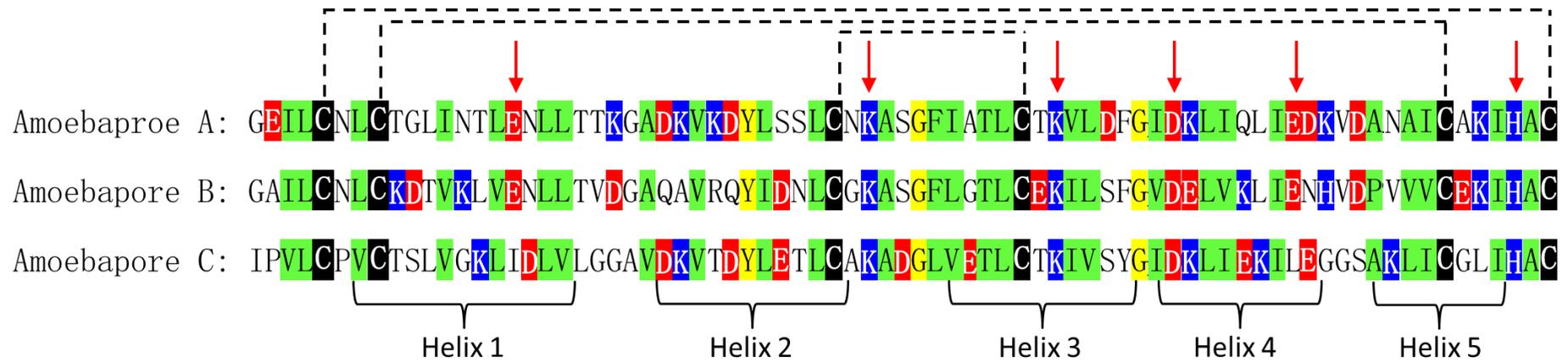


Figure 1.4: Sequence alignment of amoebapore A, B and C. Cysteine residues are highlighted in black with dashed lines representing the three disulfide bonds formed. Configuration of the disulfide bonds were determined based on the solution structure of APA-1 solved in 2004 [24]. The conservative hydrophobic residues across all three isoforms are highlighted in green. The conserved charged residues are indicated by arrows (*i.e.*, E16, K37, K48, D55, E62 and H75) and the non-charged conserved residues (*i.e.*, Y30, G40 and G53) are highlighted in yellow. Positively charged residues are highlighted in blue and negatively charged residues are in red [25, 26].

Since amoebapore A (APA-1) is the most abundant amoebapore among the three isoforms, its pathogenic activity has been the most extensively investigated [27-29]. In both *in vitro* and *in vivo* studies with human red blood cells (RBC), baby hamster kidney cells (BHK) and bacteria cells, it was observed that down-regulation or silencing of APA-1 expression significantly diminished the virulence of *E. histolytica* trophozoites (Figure 1.5). Surprisingly, overexpression of APA-1 did not increase the parasitic virulence but dramatically reduced it. Excess amounts of APA-1 were found in the cytosol of the parasite and surrounding media, suggesting extra internal control mechanisms may be required to prevent gene overexpression and to maintain the protein in the cytoplasmic granules. The results clearly indicated that APA-1 was a key virulent factor of *E. histolytica*, and changes in APA-1 levels or cellular localisation can lead to dramatic reductions of the pathogenicity of *E. histolytica*.

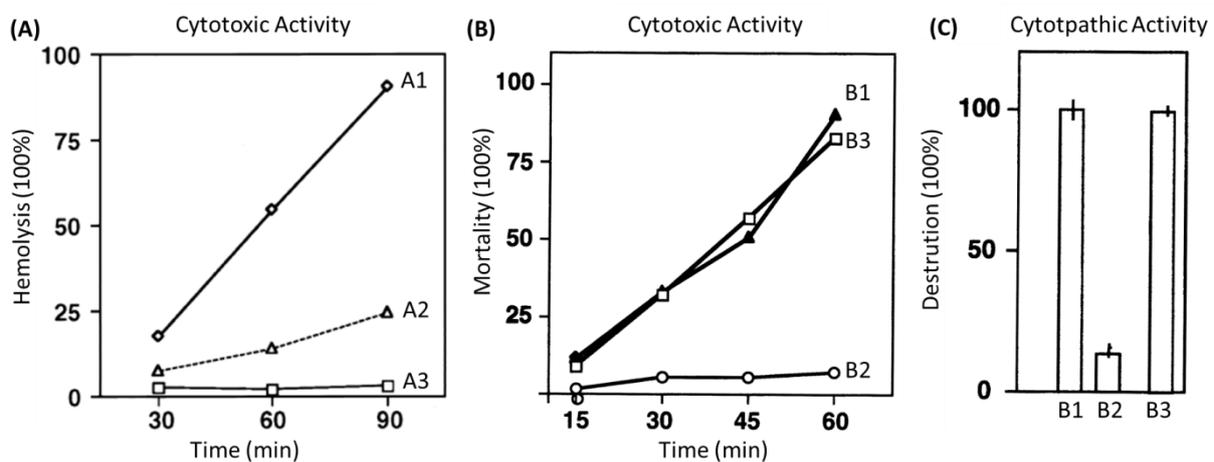


Figure 1.5: Cytotoxic and cytopathic activities of *E. histolytica* trophozoites. (A) Lysis of human RBC by intact trophozoites: (A1) wild-type trophozoites; (A2) trophozoites containing down-regulated APA-1 and (A3) human RBC not exposed to trophozoites. (B) and (C) represent the results of lysis of BHK cells (*in vitro* study) and destruction of a tissue cultured monolayer of BHK cells by intact trophozoites (*in vivo* study): (B1) wild-type trophozoites, (B2) trophozoites containing down-regulated APA-1 and (B3) a non-related transfectant as a control. The trophozoites containing down-regulated APA-1 had drastically lower pathogenic activity against the human RBC and BHK cells. The graphs were modified from [27, 29].

1.2.3 The importance of positively charged residues for amoebapore A-membrane interaction

To further investigate the importance of individual APA-1 helices in the membrane-penetrating mechanism and antibacterial activity, APA-1 helices were separately synthesised and their pore-forming activities were examined using a liposome depolarisation assay (Figure 1.6) [30, 31]. Helices 3 and 4 were synthesised as a single peptide, because in initial studies, the

secondary structure of amoebapore was predicted to consist of four α -helices [8, 25]; where helix 3 was found to be a helical stretch of two continuous helices (helix 3 and 4) connected by a small kink, as revealed by the APA-1 solution structure [24]. The liposome depolarisation results indicated that among the four synthetic peptides, the peptide which consisted of helices 3 and 4 had the most pronounced pore-forming activity; although it was not as potent as full-length APA-1. Less activity was detected for helix 1 and no activity for helices 2 and 5. However, by adding extra amino acids with a positive net charge to the C-terminus of helix 1, its pore-forming activity was remarkably enhanced. A similar result was observed for the helix (3 + 4) construct.

Moreover, interactions between vesicles composed of various phospholipids with APA-1 and its synthetic peptides were examined. The results showed that the absorption of the peptides was influenced by the membrane charge [8, 30, 31]. APA-1 and its synthetic peptides interacted preferably to negatively charged phospholipids such as phosphatidylinositol (PI), phosphatidylglycerol (PG) and phosphatidylserine (PS), and showed no pronounced binding to neutral phospholipids, such as sphingomyelin (SM), phosphatidylethanolamine (PE) and phosphatidylcholine (PC). The helix 1 homologue with the most positive overall charge showed the greatest binding to the anionic phospholipids compared with the other peptides (Figure 1.7). The results show that positively charged residues are critical for APA-1-membrane interaction. APA-1-membrane interaction may be initiated by electrostatic forces without interacting with any specific membrane receptor. Consequently, antibodies against APA-1 are unable to inhibit its toxic effect [32]. In addition, among the five helices, helices 3 and 4 have the most number of conserved hydrophobic residues and charged residues (*i.e.*, K37, D55 and E62), and the helices showed the highest pore-forming activity. This suggests that following initial binding of APA-1 to membranes, the amphipathicity of the helices may be important in membrane penetration, and the charged residues may also play important roles during membrane binding and/or the penetration process. Figure 1.8 shows a helical wheel projection of each APA-1 helix. Based on the APA-1 solution structure, the hydrophobic residues face toward each other [24], thereby forming a hydrophobic core that is likely to contribute to the stability of the protein.

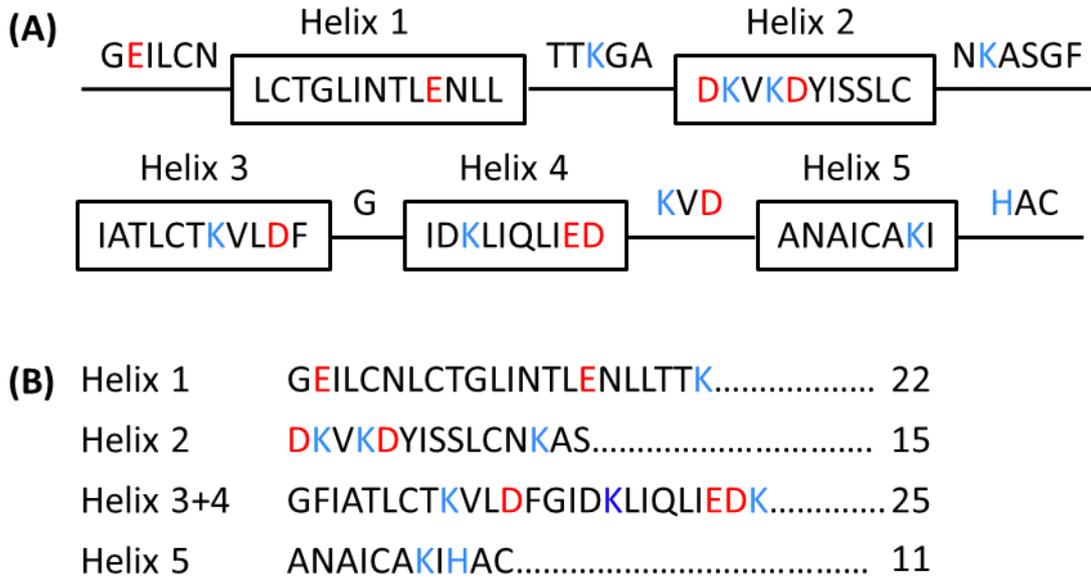


Figure 1.6: (A) Primary structure of APA-1, and (B) the synthetic peptides of APA-1 with the number of amino acids [25, 31]. Positively charged amino acids are coloured in blue and negatively charged amino acids are in red. The boxes represent the five helices defined in the solution structure of APA-1 [24].

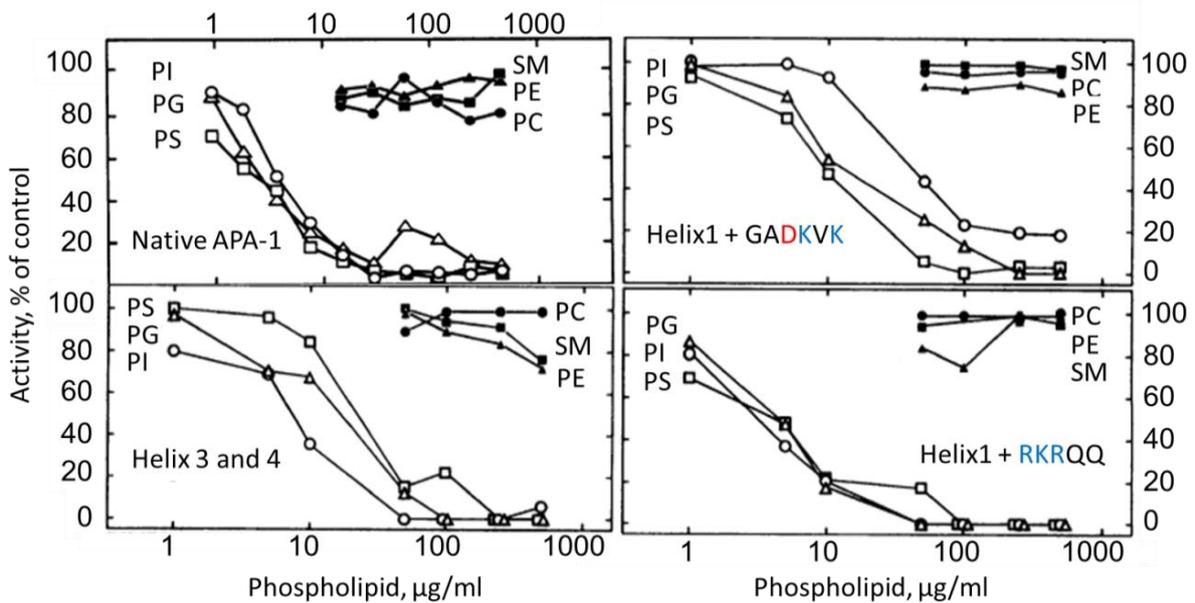


Figure 1.7: Pore-forming activity of a trophozoite extract (native APA-1) and synthetic peptides (helices 3 and 4, and two helix 1 homologues) measured by a liposome depolarisation assay. APA-1 and the synthetic peptides were incubated with vesicles composed of various phospholipids. The remaining APA-1 pore-forming activities were then examined by measuring the depolarisation of the liposome membrane potential. APA-1 and the synthetic peptides interacted more favourably with negatively charged membranes. The positively charged residues on helix 1 homologues are depicted in blue and negatively charged residues are in red. The graphs were modified from [8, 31].

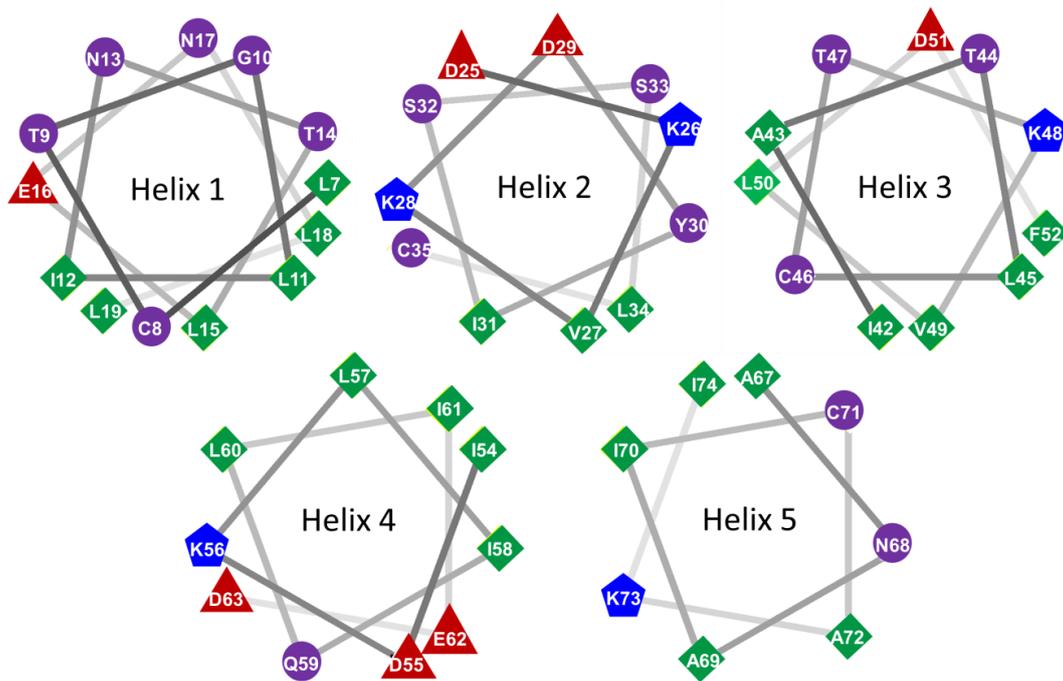


Figure 1.8: Helical wheel projections of the APA-1 sequence. The five helices are respectively viewed from one end with the perimeter of a wheel corresponding to the backbone of the helix. The hydrophilic residues are in purple, hydrophobic residues in green, potentially negatively charged residues are in red, and potentially positively charged residues are in blue. The helical wheel projection is generated by [33] and modified according to [34]. Based on the APA-1 solution structure [24], the angles of each helix were adjusted so that the hydrophobic residues are facing toward the centre, forming a hydrophobic core which stabilises the APA-1 fold.

The three disulfide bonds are crucial for maintaining the rigidity and tertiary fold of APA-1. Reduction of the disulfide bonds has been shown to result in complete loss of APA-1 activity [8]. This leads to the questions of: (1) how the single peptides (helix 1 and helices 3 and 4) orientate for membrane insertion in the absence of disulfide bonds to give pore-forming activity; and (2) how the single peptides oligomerise to higher molecular aggregates, which appears to be a prerequisite for pore formation and membrane disruption [17, 18]. In addition, synthetic helices 2 and 5 showed no pore-forming activity, yet these helices have positive overall charges of +1 and +2 (at neutral pH), and contain 26 and 55% of the hydrophobic residues, respectively. This may be explained by simple model systems for small lytic peptides. A peptide with at least 20 amino acids is required to ensure that the peptide spans the lipid bilayer and aggregates into pore structures [35, 36]. The hydrophobic region of the α -helix peptides interacts with the hydrophobic core of the phospholipids (*i.e.*, acyl chains) and the hydrophilic region forms the inside face of the pore. Helices that are shorter than 20 amino acids may perturb membrane integrity by interacting with phospholipids in the plane of the

lipid bilayer, but cannot easily form pores unless the bilayer narrows [35, 36]. The synthetic peptides of helices 2 and 5 have fifteen and eleven amino acids, respectively. Consequently, their lack of membrane permeabilising activity may be because of the relatively short α -helix lengths formed. Moreover, it was noticed that although the peptide composed of helices 3 and 4 showed the highest pore-forming activity, both these peptides were less potent and had weaker binding affinity to anionic phospholipids when compared with full-length APA-1. This indicates that the presence of certain areas in other helices may be crucial to achieve the full potency of APA-1, such as Lys37 located on helix 2, His75 on helix 5 and the hydrophobic residues, which are conserved across all the amoebapore isoforms. Site-directed mutagenesis of these residues and pore-forming assays of the mutants will provide a better understanding of which residues are important for APA-1 function.

Positively charged residues are considered to be crucial for APA-1 membrane lytic action, reflected by the selective binding to negatively charged phospholipids and by an increase in APA-1 activity under slightly acidic conditions. The pH range where APA-1 is most active coupled with the eight Lys residues interspersed along the entire sequence and a C-terminally placed His residue have been suggested to play crucial roles for APA-1 interaction with anionic phospholipids and thus pore-forming activity. To verify this, Lys and His residues of APA-1 were chemically modified under different conditions by diethylpyrocarbonate (DEPC), citraconic anhydride and manoalide. DEPC at pH 8.5 was used to modify both His and Lys residues and this compound only modifies His at pH 6.0. Citraconic anhydride or manoalide were used to modify Lys only. All treatments resulted in nearly complete loss of APA-1 pore-forming activity (Figure 1.9) [32]. Such binding preference to anionic membranes is also known for a variety of antimicrobial proteins and peptides. According to the Antimicrobial Peptide Database (APD) [37], among the 525 antimicrobial proteins/peptides that have been characterised, 96% of them prefer interacting with an anionic bacterial membrane. For example, positively charged residues of the magainin 2 are crucial for enhancing the initial interaction with negatively charged membranes and facilitating pore-forming activity [38].

Interestingly, both the liposome depolarisation assay and membrane association study showed that modification of the sole His residue (H75) resulted in no pore-forming activity but 70% of the protein remained in contact with liposomes, whereas modification of Lys residues resulted in complete protein dissociation from the membrane and loss of activity. The results suggest that the Lys residues play a major role in membrane interaction, and the His residue appears to be important for pore formation [32]. However, it was unclear as to whether the chemical

modifications caused any alteration to the tertiary structure of APA-1, and thus abolished APA-1 activity. In addition, apart from His and Lys, DEPC is known to react with the side-chains of other amino acids such as Cys and Tyr [39]. Since all the Cys residues are involved in disulfide bonds, the only Tyr residue (*i.e.*, Tyr30), which is conserved among the amoebapore isoforms, may have been modified and this modification may have influenced APA-1 activity. Moreover, it was observed that chemical treatment by manoalide only modified ~60% of the Lys residues. Such modification gave rise to an APA-1 containing solution with 30% activity [32]. This suggests that some Lys residues were not readily accessible for modification. These Lys residues may be buried in the APA-1 structure and contribute to APA-1 activity.

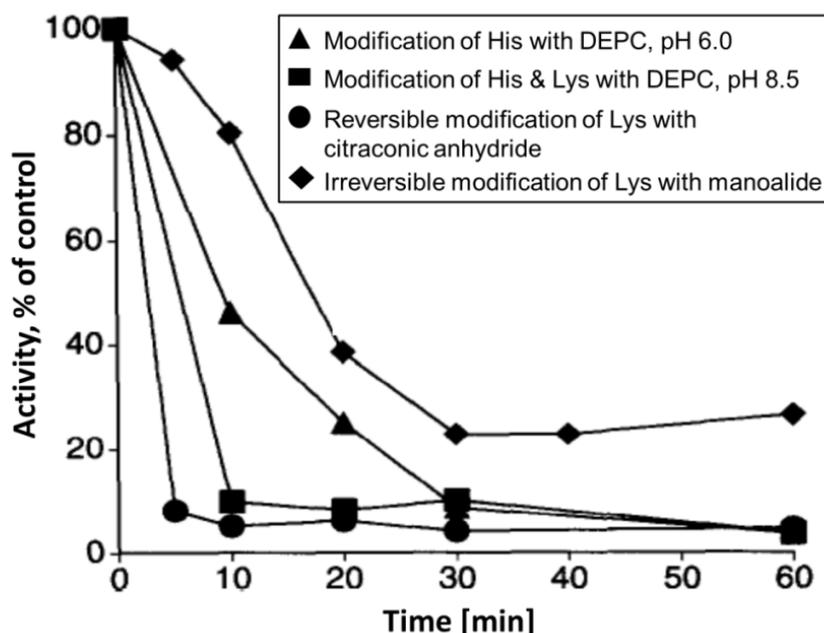


Figure 1.9: The effect of chemical modifications to His and Lys residues on pore-forming activity of APA-1. The pore-forming activity was examined by measuring the dissipation of the liposome membrane potential. Native APA-1 without any chemical modification was used as the positive control. Loss of pore-forming activity during incubation with a chemical modifying reagent is shown as a function of time. All the treatments showed loss of activity. The figure was modified from [32].

1.2.4 pH-dependent self-association and pore formation

Several studies have shown that the pore-forming activity of APA-1 is pH-dependent [8, 18, 32, 40]. A decrease in APA-1-membrane association did not coincide with a decrease of APA-1 pore-forming activity [32]. Figure 1.10 shows that pore-forming activity of APA-1 declined dramatically to ~30% when the pH was raised from 5.2 to ~5.6, and was essentially lost by pH 7. Intriguingly, the figure also shows that no pronounced reduction of APA-1-liposome

association was observed over the same pH range (*i.e.*, pH 5.2 to ~5.6). Nonetheless, APA-1 lipid association was disrupted at pH values > 6.5. These results support the chemical modification study, which suggests that in addition to interaction with the lipids, another event, which was later proposed to be APA-1 dimerisation [24], is involved in APA-1 pore-forming activity [32]. The pH range where APA-1 abruptly loses its association with membranes (*i.e.*, pH 6.5–7) is close to the intrinsic pKa value of His residues, which ranges between 6.3 and 6.8 [41, 42]. Based on these observations together with the chemical modification studies, it has been proposed that the protonation state of H75 of APA-1 is an important determinant for APA-1 pore formation and membrane association.

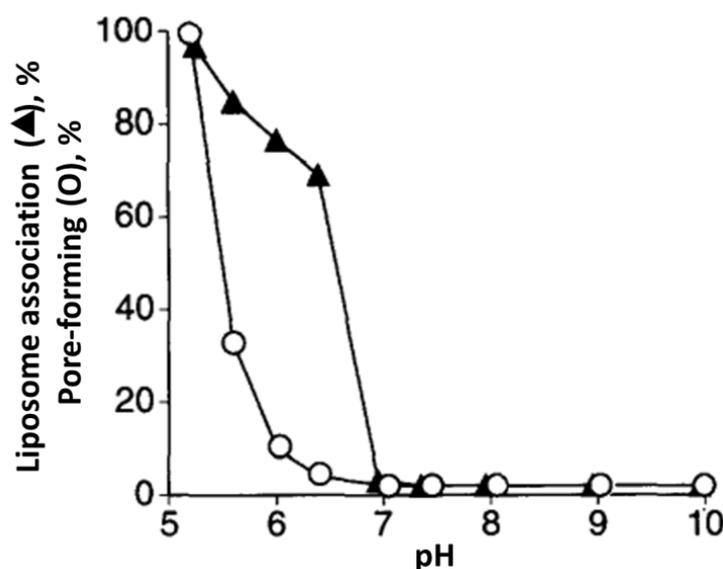


Figure 1.10: The effect of pH on APA-1-liposome association and pore formation. Liposome association was determined by measuring the amount of free APA-1 in solution after incubation with liposomes. Pore-forming activity was studied by measuring liposome depolarisation. Both measured properties at pH 5.2 were normalised as 100% [32].

Using planar lipid bilayers, APA-1 was observed to form undefined structures that resulted in an increase of membrane conductance [20]. By crosslinking APA-1 using glutaraldehyde, homo-oligomers with MWs of dimers up to hexamers were observed in the presence of liposomes. Crosslinking of APA-1 in solution also showed similar behaviour, but oligomerisation was less extensive, suggesting that localisation of the protein in the membrane facilitates the formation of oligomeric species. Self-association of APA-1 is reversible [8]. Based on these results together with the previously described chemical modification and liposome depolarisation assays, it appears that APA-1 self-association occurs in solution prior to membrane binding, whereas formation of the proposed pore structure, which causes

membrane permeabilisation, only occurs once APA-1 is membrane bound. To characterise the role of H75 in pore formation, crosslinking experiments were performed with DEPC-modified His APA-1 [32]. The results showed that following treatment the majority of APA-1 was present as monomers, although a small number of oligomers with lower MWs were also detected. Pore-forming activity of the H75-modified protein was lost. This indicates that H75 modification blocks the formation of oligomers and thus assembly of pores within the membrane. Nevertheless, the influence of the chemical treatments to other residues or to the APA-1 tertiary structure was not characterised in these studies.

Based on the above observations, the protonation state of the sole His residue appears to be a crucial determinant for APA-1 oligomerisation and may play a role in membrane binding. Since Lys residues have pKa values of ~ 10.5 [41-43] and are protonated over the APA-1 pH-active range, initial interactions between APA-1 and anionic membranes are likely to be driven by surface accessible Lys residues. It has been proposed that once APA-1 is membrane-bound and the concentration of bound APA-1 reaches a sufficiently high value, APA-1 undergoes higher aggregation to form pores with the hydrophobic regions exposed to the lipids and the lipid-unfavourable hydrophilic regions located in the centre of the pore structure [24]. It is postulated that under slightly acidic conditions where APA-1 is most active, the protonated H75 is involved in electrostatic interactions with the side-chain of either an Asp or Glu residue of APA-1, and this interaction drives homotypic dimer formation. Deprotonation of H75 or protonation of acidic residues is hypothesised to disrupt this homotypic association. Moreover, as previously described, APA-1 abruptly loses its interaction with membranes at pH 6.5–7, close to the intrinsic pKa of His, indicating the protonation state of H75 may be functionally important for membrane binding.

The intrinsic pKa values of Asp and Glu residues are 3.9–4 and 4.2–4.4, respectively [41, 42, 44]. However, acidic residues can have elevated or suppressed pKa values in proteins. Residues with unusual pKa values are often associated with functionally important events such as salt-bridge formation, hydrogen bonding and protein self-association [45-48]. It is possible that acidic residues that have elevated pKa values close to pH 5.2 may play important roles in regulating APA-1 activity. For example, protonation of acidic residues should lead to an APA-1 surface that has a higher overall positive charge, and this should enhance APA-1 interaction with anionic membranes. Conversely, deprotonation of the acidic residues reduces the overall positive charge of APA-1 and this likely destabilises the pore structure by introducing repulsive forces that do not favour APA-1 oligomerisation or membrane interaction.

Six charged residues are observed to be conserved across all the amoebapore isoforms: E16, K37, K48, D55, E62 and H75, suggesting these residues may play important roles in APA-1 activity. To better understand the roles of these ionisable residues in APA-1 functionality, it is crucial to measure their side-chain pKa values in an effort to identify specific residues that have charged states that are congruent with the active pH values. This can be achieved using NMR experiments and will be reported in Chapter 3.

Conductance measurements of different planar lipid membranes has provided data that indicates APA-1 forms pores with different lifetimes and diameters, with estimated values varying between 6.3–22.1 Å [49]. However, the method does not directly visualise APA-1 pores to provide unambiguous evidence that membrane permeabilisation was solely because of pore formation, particularly under conditions similar to those encountered *in vivo*. Consequently, no evidence describing the detailed structure and mechanism of APA-1 pore formation exists. Electron microscopy has been used to directly visualise pores in membranes for various pore-forming proteins, providing more detailed information of pore structures. For example, melittin, a highly cytolytic peptide produced from the honeybee venom (*Apis mellifera*), was identified to exist as tetramers to octamers and form pores in membranes with a 70–80 Å ring diameter and a 35–45 Å internal diameter [50]. Fragaceatoxin C, an actinoporin produced from the sea anemone *Actinia fragacea*, was found to form pore structures consisting of a protomer with a ring diameter of 112 Å and an internal diameter of 52 Å [51]. Magainin 2, a broad-spectrum antimicrobial peptide isolated from the skin of the African clawed frog *Xenopus laevis*, was recognised to form pores in target membranes with diameters of ~80 Å [52]. And α -toxin produced by *Staphylococcus aureus* forms pores with a diameter of 75 Å. However, the crystallisation of the APA-1 pore structure within membranes and visualising the pore structures with relatively small pore sizes (*i.e.*, 6.3–22.1 Å) by X-ray crystallography and electron microscopy is challenging [53] Detailed structural information at the atomic level that describes the pore structure of APA-1 in a membrane can also be achieved using solid-state NMR spectroscopy. However, the particular sample conditions that are required for collecting suitable high-resolution data using solid-state NMR spectroscopy can be difficult to achieve.

Other methods for investigating protein pore formation include neutron reflectometry and quartz crystal microbalance which indirectly probe protein pore formation in membranes. Neutron reflectometry measures the thickness and scattering length density change of the membrane when a protein is bound. Quartz crystal microbalance monitors the frequency and energy dissipation change of membranes upon protein interactions so that membrane rigidity

and surface mass changes can be determined. However, these techniques do not provide detailed structural information of APA-1-membrane association; for example, the residues that are crucial for pore formation and membrane binding cannot be identified using these methods.

1.2.5 Interactions of amoebapore A with membranes

The mode of action of APA-1 has been hypothesised to involve interactions with anionic phospholipids without binding to specific lipid structures or membrane receptors [54]. It was found that APA-1 was more potent against Gram-positive than against Gram-negative bacteria [30-32], and significantly higher APA-1 concentrations were required to lyse the latter type of bacteria [55]. Since the membrane surfaces of both bacteria are negatively charged due to the presence of a lipopolysaccharide layer in Gram-negative bacteria and teichoic acid bound to the peptidoglycan in Gram-positive bacteria, the external wall structures of the bacteria may play a role in the observed different APA-1 activities. The additional membrane layer of Gram-negative bacteria may act as a barrier with higher resistance against APA-1 invasion (Figure 1.11) [31]. To further investigate this, bacterial cytoplasts of Gram-positive and Gram-negative bacteria were prepared and APA-1 was able to lyse both of them, indicating that the bacterial cytoplasmic membrane is a target for APA-1 invasion [49]. The initial interaction of APA-1 with Gram-negative bacteria involves permeabilisation of its outer membrane. However, pore formation on the outer membrane without reaching the inner membrane is not sufficient to kill the bacteria, either because APA-1 molecules bound to the outer membrane are no longer available for inner membrane invasion, or the pores formed on the outer membrane may be too small to enable APA-1 molecules to enter into the periplasmic space. In contrast to APA-1, NK-lysin is able to kill both Gram-positive and Gram-negative bacteria at similar concentrations [49]. This is most likely because of its overall surface charge being highly positive at functional pH values, thus giving rise to high binding affinity to the lipopolysaccharide moiety on the outer membrane and the anionic groups on the inner membrane such as phosphatidylglycerol and cardiolipin of Gram-negative bacteria. Moreover, the remarkable positive surface of NK-lysin may be capable of displacing Ca^{2+} and Mg^{2+} cations which are important for the stabilisation of the lipopolysaccharide layer of Gram-negative bacteria, leading to membrane destabilisation [49]. APA-1 does not have any pronounced electropositive region (see Session 1.3.1) and thus not easily penetrate the lipopolysaccharide layer. Similar to APA-1, other pore-forming proteins, such as psoriasin expressed from human differentiated keratinocytes, were found to be less toxic against Gram-negative bacteria. Much longer incubation periods were required for cell lysis [56].

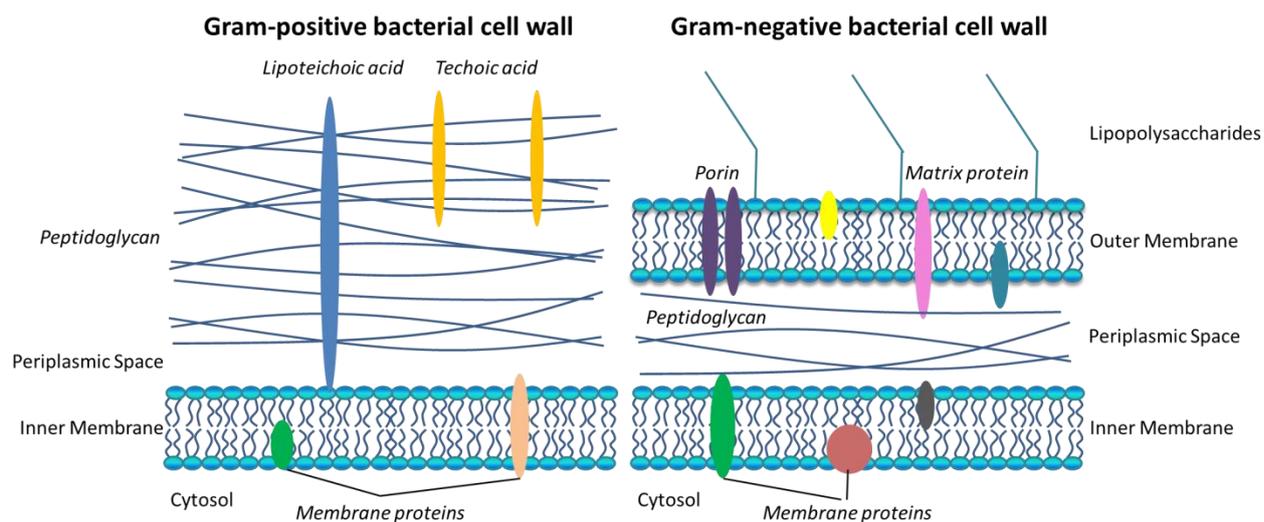


Figure 1.11: Schematic representations of the Gram-positive and Gram-negative bacterial cell walls. The Gram-negative bacteria have an additional membrane layer that acts as an additional barrier against APA-1 cell homeostasis disruption.

Despite being effective against prokaryotic cells, APA-1 is also highly potent against eukaryotic cells such as human intestinal epithelium cells, erythrocytes, neutrophils, lymphocytes and macrophages [18, 20, 57]. The surface of mammalian cell membranes consists predominantly of cholesterol and zwitterionic lipids, such as PC and PE, indicating APA-1 lytic activity toward eukaryotic cells is driven primarily by hydrophobic interactions. However, it has been previously shown that APA-1 at active pH prefers binding to negatively charged membranes and showed no permeabilising activity against neutral liposomes [8]. In contrast, NK-lysin which is highly positively charged and does not show any surface region suitable for hydrophobic interactions also has cytotoxic activity against mammalian cells [55]. These results indicate that in a complicated cellular membrane system such as human membranes, complex events may occur which cannot be easily mimicked by simple model lipid membranes.

The amoeba is not affected by the amoebapores even though the plasma membrane of *E. histolytica* also contains a significant amount (10%) of anionic phospholipids [58]. This implies that the parasite is resistant to its own toxic proteins. The negatively charged lipids of *E. histolytica* are likely to be located on the inner membrane leaflet, as found with other eukaryotic cells, which are not accessible on the membrane surface for APA-1 recognition [59]. In addition, the plasma membrane of *E. histolytica* has a relatively high cholesterol content (46.5% mol/mol of phospholipids) which contributes to membrane rigidity [58]. The presence

of cholesterol on liposomes greatly reduces membrane binding affinity of APA-1 [59]. Therefore, bacteria cells lack cholesterol are more accessible to APA-1 attack than eukaryotic cells. Moreover, large amounts of N-acyl-sphingosylphosphorylethanolamine (CAEphosphate) (39%), which is a sphingolipid known for its stability against hydrolysis, is present in the *E. histolytica* plasma membrane. The CAEphosphate is almost twice as abundant in *E. histolytica* membranes as the sphingolipid content present in other eukaryotic cell membranes [58]. The high content of sphingolipid and cholesterol may form a high density membrane shield that protects the parasite from its own membranolytic proteins, as well as toxic components produced by other microorganisms that may also dwell in the human colon, such as the Shiga toxin produced by the Shigatoxigenic group of *Escherichia coli* that act to inhibit protein synthesis in target cells [60]. Lastly, glycosylphosphatidylinositol (GPI)-anchored lipophosphopeptidoglycan molecules located on the membrane surface can form a densely packed matrix on the entire surface of the trophozoites, which may further contribute to protect the parasite against lytic damage [59, 61].

1.2.6 Other pathogenic factors of *Entamoeba histolytica*

Apart from APA-1, which is an essential virulent factor of *E. histolytica*, other pathogenic factors have been identified. The major surface receptor of the parasite, lectin with galactose N-acetyl galactosamine carbohydrate (Gal/GalNAc lectin), has been shown to participate in the initial recognition and adherence of the parasite to host tissue [62]. Interaction of the Gal/GalNAc lectin with the brush border of intestinal epithelial cells may trigger cell apoptosis signalling and lesions to the host tissue that permits spread of the parasite through the blood stream. Target cell death occurs within 5–15 min after cellular adherence and this is followed by phagocytosis of the target cell. In addition to the Gal/GalNAc lectin, two proteins enriched in Lys and Glu residues, named KERP1 and KERP2, were also identified to be involved in the establishment of amoeba-cell interactions and the progression of parasitic invasion [63]. The GPI-anchored lipophosphopeptidoglycan molecules located on the *E. histolytica* trophozoites surface are also able to interact with host cells [64]. One remarkable feature of *E. histolytica* during tissue invasion is that it is able to lyse human colon cells and destroy the extracellular matrix. Cysteine proteases produced by *E. histolytica* were identified to be directly involved in target tissue invasion and gut inflammation by degradation of extracellular matrix proteins, colon mucus and host antibodies [65]. Similar to APA-1, *E. histolytica* trophozoites with reduced amount of Gal/GalNAc lectin and cysteine proteases showed a dramatic decrease of invasive activity [66-68]. This suggests that cooperative action is present among different

virulent factors in the parasite to maintain its pathogenicity. Decreasing the amount of one factor can lower the overall parasitic virulence, whereas increasing one of these factors can cause other factors to become limited, thereby also reducing the pathogenicity of *E. histolytica*, as shown in Section 1.2.2 [29].

1.3 Structural and biophysical studies of amoebapore A

1.3.1. Amoebapore A: a member of the saposin-like protein family

In 2004, the solution structure of monomeric native APA-1 at pH 3.0 was solved by homonuclear two-dimensional (2D) ^1H nuclear magnetic resonance (NMR) spectroscopy [24]. The tertiary structure of APA-1 revealed that it belongs to the saposin-like protein (SAPLIP) family that is characterised by a bundle of five amphipathic α -helices and three disulfide bonds that provide protein stability (Figure 1.12).

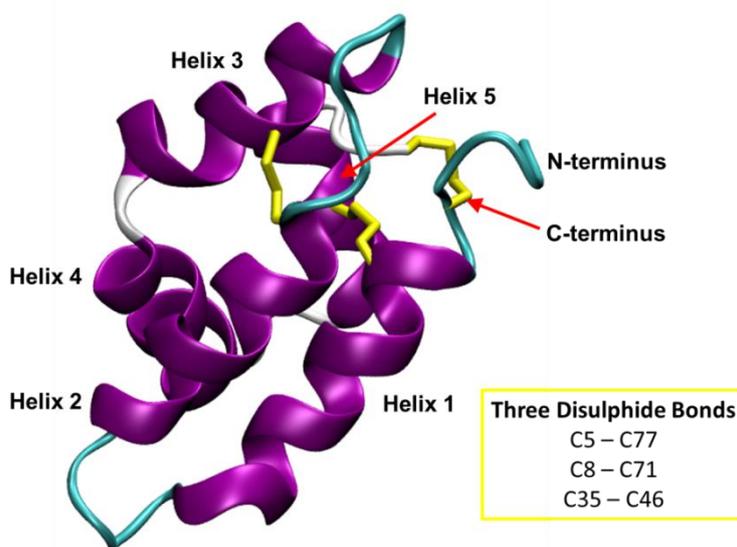


Figure 1.12: Ribbon representation of the structure of monomeric APA-1 solved by NMR spectroscopy at pH 3.0 [24]. The three disulfide bonds are depicted in yellow. The six cysteines that form the disulfide bonds are highlighted in the yellow box.

The SAPLIP family has over 200 members and it is termed SAPLIP because saposin A–D were the first characterised proteins of this family. Although the SAPLIP members share the characteristic saposin-like fold, they perform a highly diverse range of biological functions that are associated with interactions with lipids [69]. Examples of SAPLIP members which are functionally close to APA-1 include caenopore-5, saposin A–D, NK-lysin and granulysin. Caenopore-5, a pore-forming protein found in *Caenorhabditis elegans*, protects the host from pathogens that are ingested by the worm by disrupting the integrity of bacterial cell membranes [70]. Saposin A–D are four small glycoproteins derived from prosaposin and function in sphingolipid degradation and membrane digestion. They work as cofactors during specific sphingolipid hydrolysis in lysosomes by interacting with membranes [71]. NK-lysin [72] and

human granulysin [73] are located in the intracellular granules of porcine and human cytotoxic lymphocytes, respectively. Both SAPLIPs participate in the internal immune mechanisms by exhibiting antibacterial activity against pathogens.

By comparing the primary sequences and tertiary structures of APA-1 and particular SAPLIP members (*i.e.*, caenopore 5, saposin A–C, NK-lysin and granulysin), information about functional properties of each member can be obtained. The protein sequences of these SAPLIP members are diverse, yet there are a number of features that are conserved and these features may be functionally significant across the SAPLIP family (Figures 1.13 and 1.14). First of all, SAPLIP members share the characteristic SAPLIP fold which consists of five amphipathic α -helices. The hydrophobic amino acids are distributed at similar positions in the structure and fourteen hydrophobic amino acids are conserved among all the SAPLIP proteins (presented in Figure 1.13), which is ~52% of the number of conserved hydrophobic residues across the amoebapore isoforms (14 out of 27 hydrophobic residues are conserved across the SAPLIP proteins) (Figure 1.5). This indicates that the hydrophobic core is important for maintaining the SAPLIP fold. Additionally, almost all the proteins share a similar pattern of six cysteine residues which contribute to the stabilisation of the tertiary structures through the formation of disulfide bonds. Granulysin is an exception with only four cysteine residues forming two disulfide bonds. This is due to loss of the C-terminal cysteine residue during a post-translational modification process and the loss of the N-terminal cysteine possibly during genetic evolution [74]. Furthermore, seven charged residues are conserved between APA-1 and at least another three members of the SALIP family, namely E16, K22, D25, D29, K37, K48 and H75. This indicates that the charges at these positions in the structures may play similar functional roles in membrane binding and permeabilisation properties in different proteins. As previously discussed, the Lys residues are important for initial binding of APA-1 with membranes, H75 is crucial for both APA-1 oligomerisation and interaction with membranes, and the acidic residues may be involved in facilitating self-association, membrane binding or regulating the formation of APA-1 pores. Among these residues, K37 is highly conserved among four SAPLIP members, suggesting this residue is likely to be crucial for common functionality among these proteins. In addition, the positively charged residue H75 of APA-1 is also conserved among four SAPLIP proteins, except the His residue is replaced by a Lys and Arg residue in NK-lysin and granulysin, respectively (Figure 1.13 and 1.15). The replacement may be responsible for the differences in their pH-dependent activities, as previously described in Section 1.2.4. Site-directed mutagenesis of K37 and/or H75 to remove the positive charges and replacement of H75 with a Lys or Arg residue combined with pore-forming assays should

provide a better understanding of the role of these residues in the activity of APA-1 and other SAPLIP members.

Despite these similarities, the electrostatic surfaces of the SAPLIP proteins are remarkably different, with net charges ranging from -9 (saposin C) to $+10.8$ (granulysin) at neutral pH, assuming all residues have pKa values that are equivalent to the isolated residues (calculated using on-line Protein Calculator developed by The Scripps Research Institute, San Diego, USA). The diversity of the electrostatic surfaces of the proteins is considered to reflect their distinct mechanisms of protein-lipid association and various functions [75]. APA-1 has a neutral surface charge at pH 7, with a slightly acidic pI value (calculated) of ~ 5.9 , and the charged residues are evenly distributed across the protein surface (Figure 1.14). Similar to APA-1, caenopore-5 has an overall charge of -1.5 , a calculated pI value of ~ 6.3 , and has been postulated to perform its antibacterial action by pore formation. Its activity is pH-dependent with the highest activity at pH 5.4, which is very close to that of APA-1 (*i.e.*, pH 5.2) [70]. In contrast to APA-1, saposin isoforms have low pI values around 4 and perform higher activity under acidic conditions. Although the saposin isoforms share a high degree of structural similarity, they have distinct modes of actions and diverse ligand-binding properties [76]. NK-lysin [72] and granulysin [73] are dominated by positively charged residues with overall surface charges of $+5.7$ and $+10.8$ at neutral pH, and high pI values of ~ 8.9 and ~ 10.8 , respectively. The positive surface of these two proteins is proposed to be responsible for initial interaction with membranes that ultimately lead to membrane disruption. It is proposed that NK-lysin permeabilises membranes in the monomeric state primarily via electrostatic interactions. The high density of positively charged amino acids of NK-lysin, especially on helix 3 may induce an electric field strong enough to perturb the lipid packing order, leading to membrane destabilisation [77]. For granulysin, in order to interact with the hydrophobic core of the target membrane, it undergoes a rotation that creates hydrophobic patches on its surface for membrane interaction [73]. Based on these outcomes, the unique charge distribution and spatial arrangement of APA-1 indicates that the interaction of APA-1 with target membranes involves a unique mechanism that is not understood.

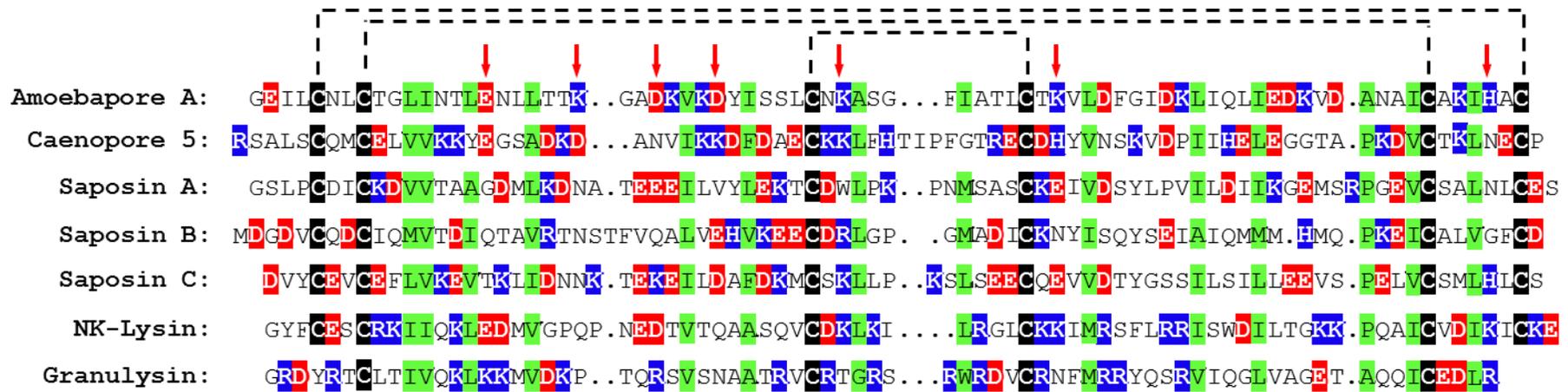


Figure 1.13: Sequence alignment of amoebapore A, caenopore 5, saposin A, B and C, NK-lysin and granulysin. Cysteine residues are highlighted in black with dashed lines linking cysteines that form the three disulfide bonds. The hydrophobic residues which are conserved in more than six SAPLIP proteins are highlighted in green. Positively charged residues are highlighted in blue and negatively charged residues are highlighted in red. The charged residues conserved between APA-1 with at least another three SAPLIP members are indicated by red arrows. These conserved charged residues which are not clearly identified based on the sequence alignment have similar spatial locations in the tertiary structure (see Figure 1.15). The sequence alignment was generated by CLUSTALW [78].

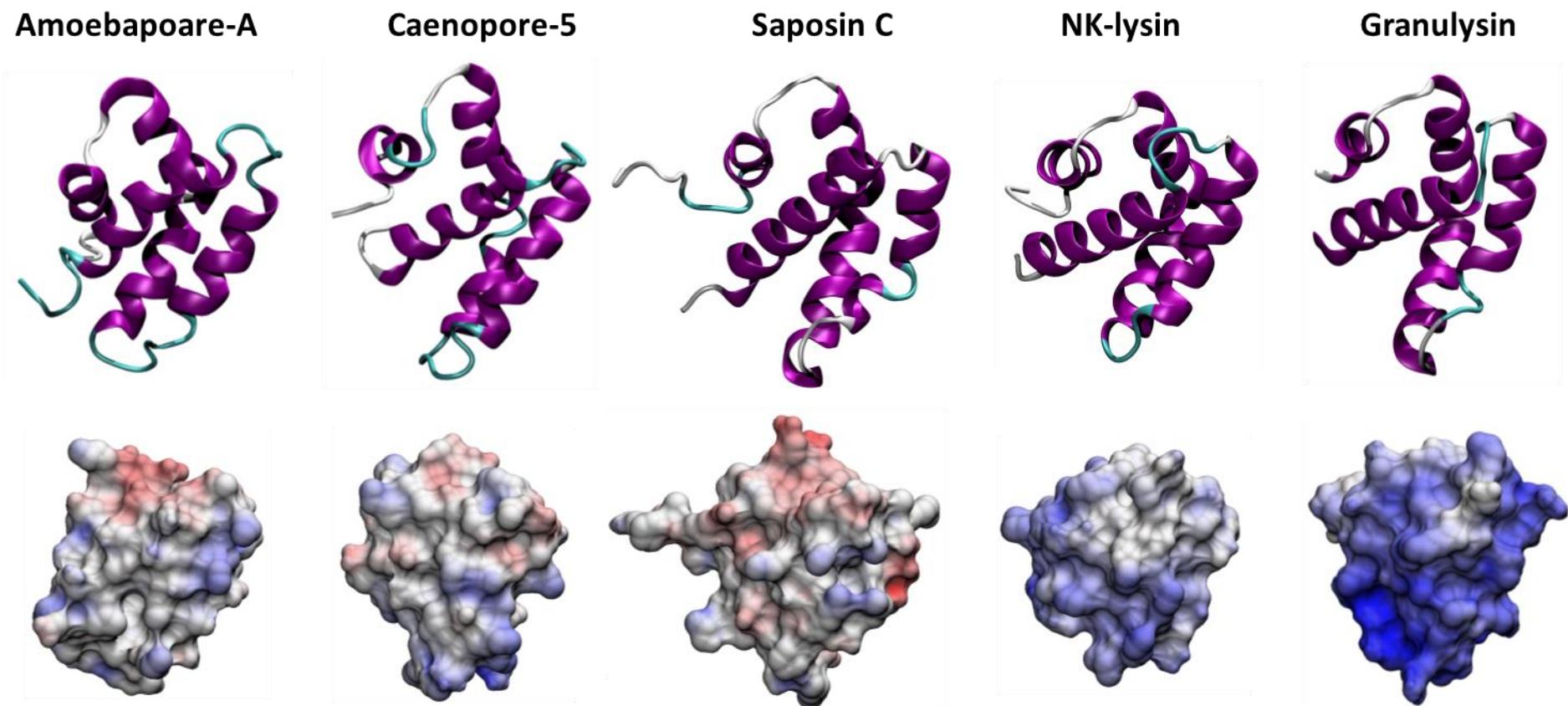


Figure 1.14: Tertiary structures and surface electrostatic potentials of SAPLIP proteins. (A) amoebapore A; (B) caenopore 5; (C) saposin C; (D) NK-lysin, and (E) granulysin. The overall charges for the proteins (at neutral pH) are calculated to be 0, +1, -7, +6 and +11, respectively, assuming all residues have pKa values that are equivalent to the isolated residues. Positively charged surface regions are depicted in blue and negatively charged surface regions are depicted in red. The conserved disulfide bonds are not shown.

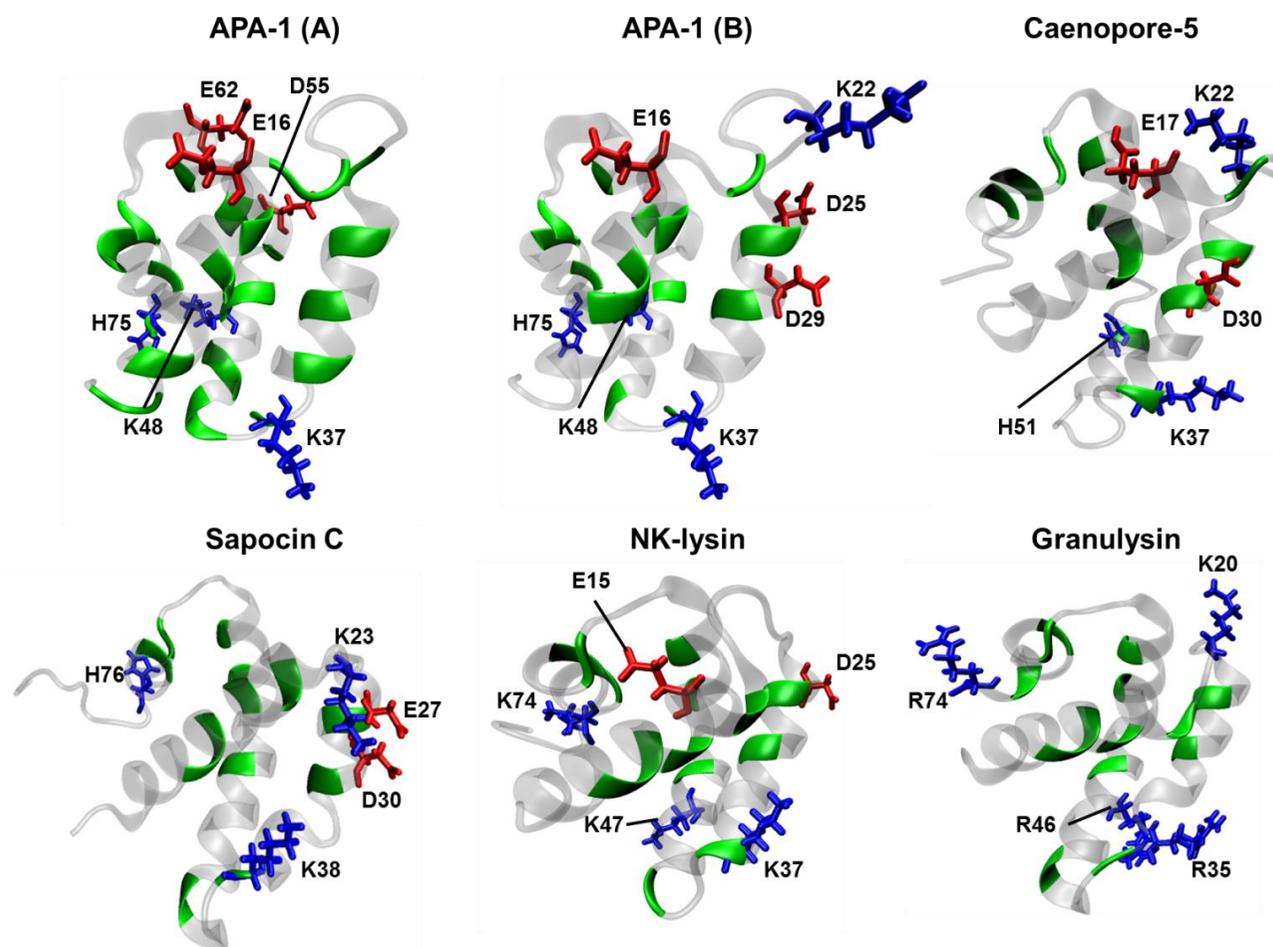


Figure 1.15: The conserved residues present across the three amoebapore isoforms (APA-1 (A)) and between APA-1 and other SAPLIP members (APA-1 (B)). The conserved residues with respect to the sequence of APA-1 in other SAPLIP members are also shown. The conserved hydrophobic residues are depicted in green, the conserved positively charged residues are presented in blue and the negatively charged residues are in red. The conserved hydrophobic residues located in the core of the structure are considered to be important for maintaining the SAPLIP fold. The charged residues which are conserved between APA-1 and other SAPLIP members may be involved in the common functionalities among the SAPLIP members whereas the residues which are conserved across the amoebapore isoforms but not among the SAPLIP members may represent the unique mode of action of APA-1. The side-chain protons for the highlighted residues of APA-1 were rebuilt and energy minimised using programme GROMACS-PDBGMX [79].

1.3.2. pH-dependent dimerisation of amoebapore A

Chemical crosslinking of APA-1 in the liposome-bound state has shown that at active pH values APA-1 undergoes self-association and forms oligomers in membranes [32]. Based on these findings, it was further hypothesised that the membrane permeabilising activity of APA-1 is reversibly regulated by a pH-dependent dimerisation event. Based on the size exclusion chromatography results performed by Hecht *et al.*, it is proposed that APA-1 exists as monomer at pH values ≤ 3.5 and exists as a dimer at pH 5.2. At higher pH values (*i.e.*, > 8) the protein is believed to exist predominantly as monomers [24] (Figure 1.16). The MW of the later eluted peak at pH 8 was estimated to be 2 kDa smaller than the monomeric APA-1 eluted at pH 3.5, and the early eluted peak was approximately 5 kDa larger than the dimeric APA-1 eluted at pH 5.2. The variation of the elution volumes may be explained by different solvent conditions and the protein conformation.

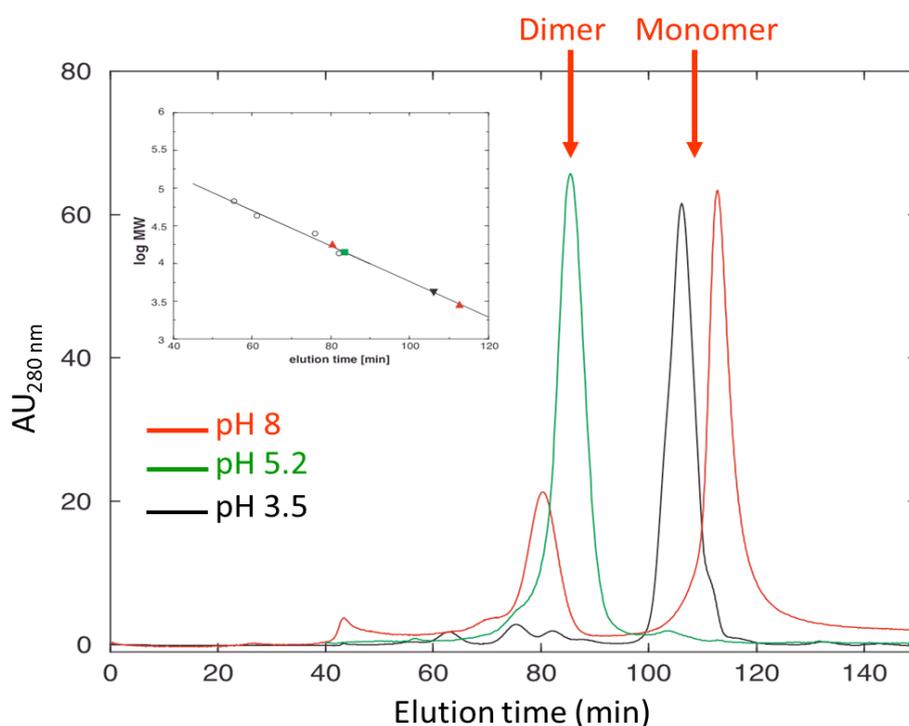


Figure 1.16: Elution profiles of size exclusion chromatography showing the pH-dependent dimerisation of APA-1. APA-1 exists as monomer at pH 3.5 (black), dimer at pH 5.2 (green) and predominantly monomer at pH 8.0 (red). The picture is adapted from [24].

It was proposed that the pH-dependent dimerisation of APA-1 is modulated by the ionisation state of the surface-exposed His residue located at the dimer interface [24]. Without altering the overall APA-1 secondary structure as confirmed by circular dichroism spectroscopy

(190–250 nm wavelength), chemical modification of the His appeared to block APA-1 dimerisation and pore formation [24]; although it cannot be ascertained whether the tertiary structure still remained the same and no other residues were modified. Based on this observation, together with structural data, it was suggested that the protonation state of the His side-chain is responsible for APA-1 pH-dependent activity and appears to be a prerequisite for APA-1 pore formation. Although description of the dimerisation interface at the atomic level is not available, this data identifies particular amino acid side-chains that appear to be responsible for dimerisation. It was proposed that at pH values between 4 and 6, the protonated His75 forms an electrostatic interaction with the negatively charged side-chain of Asp63 on the opposing monomer. A second ion pair is formed between Glu2 and Lys64. These two ion pairs trigger an anti-parallel head-to-head dimerisation of APA-1 with a height of ~ 33 Å and a width of ~ 43 Å (Figure 1.17A and B) [24, 80]. The dimerisation is disrupted by protonation of the acidic residues upon lowering the pH and by deprotonation of His75 at higher pH values (Figure 1.17C) [24]. Since, helices 1 and 2 of APA-1 provide an exclusively hydrophobic surface, it is further proposed that when a threshold concentration of APA-1 is reached in the membrane, this hydrophobic surface is extended by adding further dimers and results in a ring-like hexameric pore structure with a hydrophobic exterior and hydrophilic interior (Figure 1.18) [24]. The proposed hexameric pore has an inner diameter of ~ 20 Å, which is close to the diameter derived from biophysical conductance measurements on planar lipid bilayers [49]. However, this model does not explain the observation of different size pore formations with diameters varying between 6.3 and 22.1 Å in the conductance measurements [49]. In addition, the role of helix 3 and 4 which were previously proposed to be crucial for membrane “seeking” and association based on their remarkable membrane permeabilisation activity was not clarified by this model [30, 31].

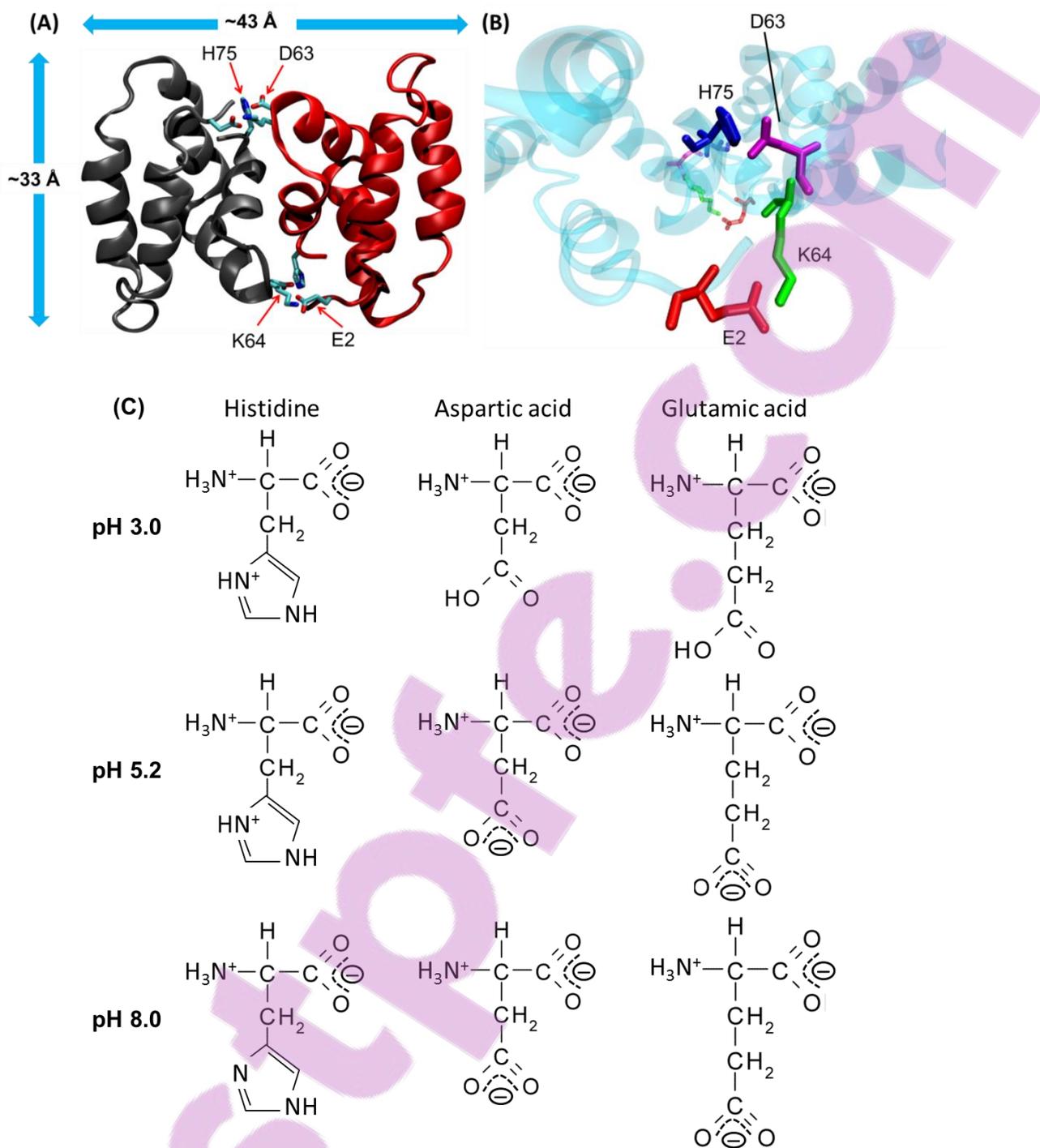


Figure 1.17: Proposed dimer model of APA-1. (A) Two ion pairs are formed at the dimer interface, H75–D63 and K64–E2 [24]. The height and width of the dimer are estimated to be ~ 33 and 43 \AA , respectively. (B) Zoomed in diagram of the dimer interface; the residues proposed to be involved in the electrostatic interactions are coloured. (C) Ionisation states of the dimer interface residues at pH 3.0, where the acidic residues are protonated and do not form salt bridges; pH 5.2, where the His residue is protonated and forms an electrostatic interaction with a deprotonated acidic residue; and pH 8.0, where the His is deprotonated and the salt bridge is disrupted. The ionisation state of Lys residues is not presented because it is very likely to be always protonated (*i.e.*, positively charged) over the pH range of APA-1 activity.

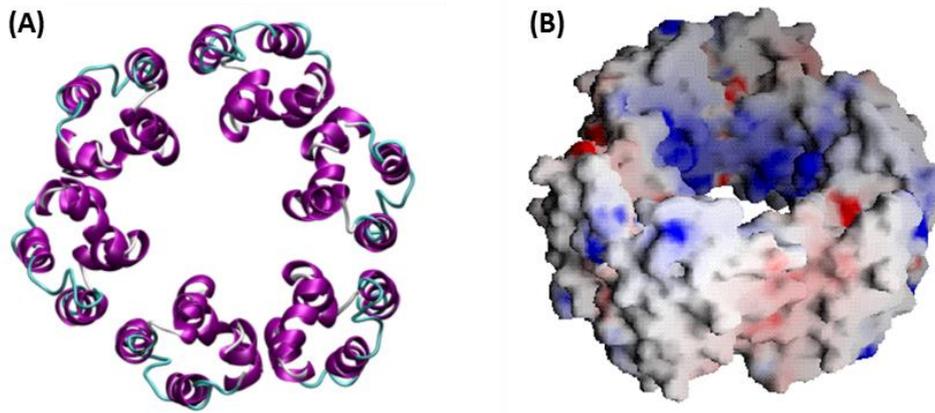


Figure 1.18: Proposed APA-1 hexameric pore model. (A) and (B) Addition of further APA-1 dimers results in a hexameric pore structure in the membrane with a hydrophobic exterior and a hydrophilic interior. In diagram (B), the positive surface potential is indicated in blue, negative surface potential in red and non-polar residues are coloured white [24].

1.3.3 Interaction of amoebapore A with membranes

Antimicrobial peptides/proteins (AMPs) are key components of innate immunity of all life forms ranging from bacteria to insects and mammals, including humans. According to the APD [37], more than 500 mature peptides have been isolated and characterised from a variety of natural sources. Currently, natural AMPs are attracting particular attention because of their potential use against antibiotic-resistant pathogens. Understanding the mode of action of these peptides is essential for novel therapeutic developments. It may also give insight into the mechanism of the more complex protein APA-1.

Most AMPs are able to permeabilise target cell membranes. Two major models have been used to describe the membrane permeabilisation activity induced by AMPs binding: the barrel stave model and the carpet model [81]. In the carpet model, AMPs orient parallel to the surface of the lipid bilayer, accumulate and form an extensive layer over the target membrane surface. The membrane structure is destabilised after a threshold concentration is reached [81] (Figure 1.19A). Examples of proteins which use this mode of action include NK-lysin and granulysin, where their membrane-active behaviour is caused mainly by electrostatic interactions with the membranes. The barrel stave model describes AMPs that form transmembrane pores. The proteins attach to the target membrane and aggregate. Once the threshold concentration is reached, AMPs insert into the membrane as oligomers perpendicular to the plane of the bilayer. The hydrophobic peptide regions align with the lipid core region and the hydrophilic peptide regions form the interior region of the pore [81] (Figure 1.19B). One typical example is

alamethicin, which is isolated from the *Trichoderma viride* fungus. It penetrates the lipid bilayer as transmembrane-oriented helix bundles surrounding the central pore after a certain peptide concentration is reached [82].

Since APA-1 has been proposed to form pores with channel-like characteristics, its interaction with membranes has been postulated according to the “barrel-stave” model [19, 24]. It is hypothesised that the initial association of APA-1 with target membranes is driven by protonated Lys residues with anionic phospholipids. The protonation state of the single His residue is postulated to drive APA-1 dimer formation in solution by forming a salt-bridge with D63 in the proposed dimer model. The dimers aggregate on the membrane and once a critical concentration is reached the protein undergoes oligomerisation via helices 1 and 2 which provide an exclusively hydrophobic surface, leading to hexameric pore formation in the target membrane with a hydrophobic exterior and a hydrophilic interior. However, it is unclear whether dimer formation is a prerequisite for APA-1-membrane interaction and the role the His residue plays in APA-1 activity. The protein may associate with the target membrane in both monomeric and dimeric forms. Accumulation of APA-1 on the surface of the membrane consequently leads to the formation of pore structures within the bilayer which results in cell death (Figure 1.20).

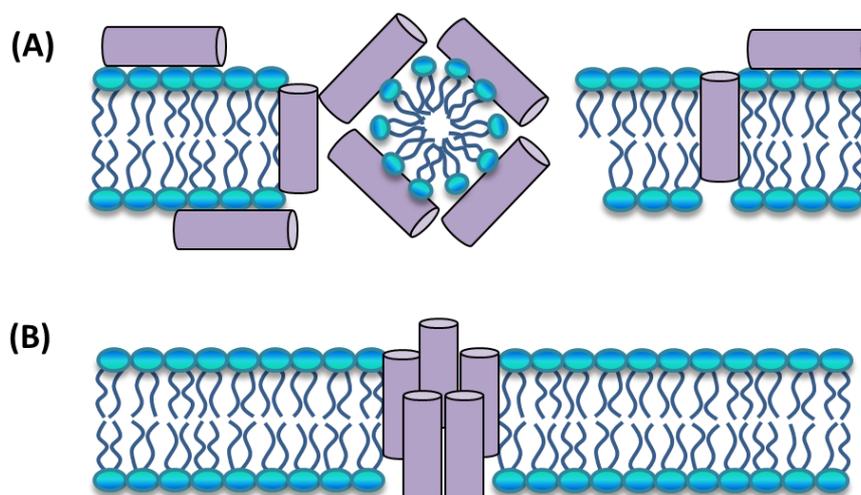


Figure 1.19: Schematic diagrams of the two models of action of AMPs. (A) Carpet model: AMPs adsorb parallel to the membrane via electrostatic interactions. The hydrophilic regions are exposed to the solvent and the hydrophobic regions face the membrane. After reaching sufficient coverage, the AMPs produce a detergent like effect that destabilises the membrane structure. (B) Barrel stave model: The positively charged residues of the AMPs interact with the anionic phospholipids, insert perpendicular to the bilayer via hydrophobic residues and oligomerise to form a pore structure. Both models require a threshold concentration of the AMPs for membrane disruption to occur.

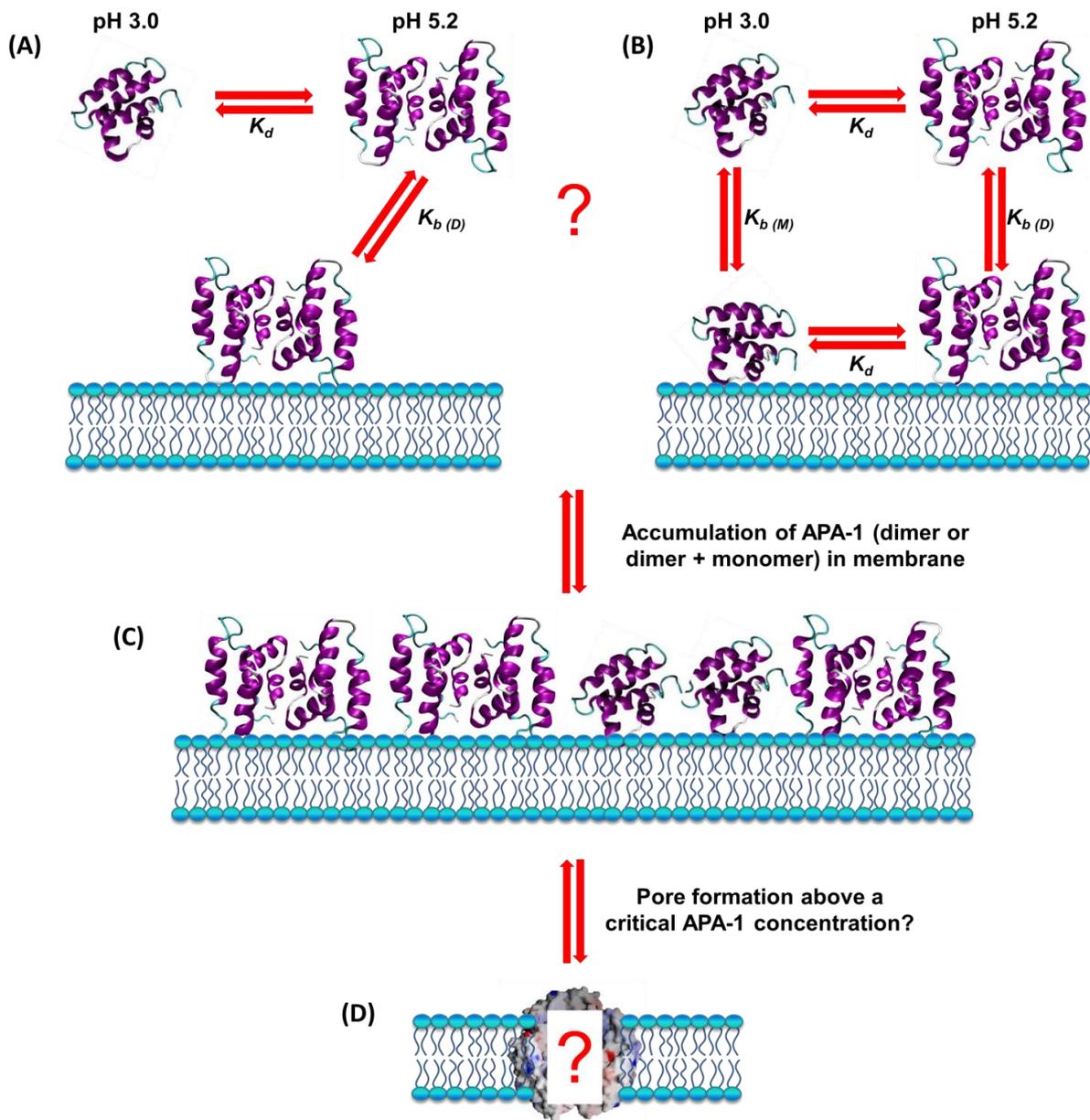


Figure 1.20: Schematic diagrams representing the proposed mechanism of APA-1 assembly into a pore structure in target membranes. (A) APA-1 dimerisation only occurs in solution; dimerisation is required prior to APA-1-membrane association. (B) APA-1 dimerisation happens both in solution and on the membrane surface (or within). Both APA-1 monomer and dimer are able to interact with the membrane. (C) In both models, a threshold concentration of APA-1 may be required for the proposed pore to form in the target membrane (D) The APA-1-membrane interactions presented in (A) and (B) are initially driven by electrostatic interactions between the positively charged residues of APA-1 and the negatively charged membrane headgroups. The K_d , $K_{b(M)}$ and $K_{b(D)}$ labelled in the diagram respectively represent the dimer dissociation constant, monomer to membrane binding constant, and dimer to membrane binding constant. All processes are considered reversible. Information such as the size of the pore structure, whether dimerisation is functionality important for APA-1 pore formation, whether conformational changes to APA-1 occur upon pore formation have not been fully resolved.

1.4 Research Aims

APA-1 is an antimicrobial protein that acts as one of the major pathogenic factors of *E. histolytica*. The activity of APA-1 has a broad killing spectrum, including both prokaryotic and eukaryotic cells. When the parasite is attached to a target cell, it discharges cytoplasmic granules to release APA-1. It has been proposed that the slightly acidic environment in the contact space between the amoeba and the target cell results in the formation of active APA-1 dimers and, ultimately, the creation of oligomeric pores in target cell membranes. The pore structure releases cell contents and this process eventually leads to cell death. However, detailed molecular structural information of the dimer, how dimerisation facilitates pore formation and the membrane binding mechanism of APA-1 remain unanswered. The central aim of this project is to use different biophysical techniques to investigate how APA-1 self-associates in solution and interacts with model membrane systems. The specific aims include:

- (1) Recombinant production of APA-1 using cell-based methods has failed. Therefore, to produce the milligram quantities needed for biophysical studies and to isotope enrich APA-1 with NMR-active nuclei, cell-free synthesis was used to produce APA-1 samples.
- (2) Develop a model that explains the pH-dependent activity of APA-1, by identifying residues that are located at the dimer interface, model the dimer structure based on biophysical studies, and determine the pKa values of acidic and His residues.
- (3) Prepare different APA-1 mutants based on the results derived in (2) and test their biological activities and examine their overall fold.
- (4) Elucidate the pH-dependent and lipid selective membrane binding activity of APA-1 using model membranes.

The results arising from this research project provide more detailed information of the molecular mechanism of APA-1 action. Such information may facilitate the design of more effective drugs against *E. histolytica* infection, and provide a better general understanding of antimicrobial protein-membrane interactions that can potentially aid in the design of APA-1 analogues for therapeutic anti-infective compounds against particular bacteria.

2

Cell-free production and structural and functional characterisation of recombinant amoebapore A

2.1 Introduction

2.1.1 Cell-free protein synthesis

The cytotoxicity of APA-1 has prevented recombinant production of this protein using conventional cell-based approaches. Extracting the protein from the parasite is both time consuming and labour intensive, and only small amounts of unlabelled native protein are retrieved using this approach [25]. To produce APA-1, including isotopically labelled APA-1, cell-free protein synthesis was used because it does not depend on cell viability [83-87].

In cell-free protein synthesis, a biologically active protein can be synthesised using crude cell extracts and the desired DNA template [88-90]. The S30 extract is obtained from homogenisation and centrifugal fractionation of lysed cells, usually bacteria, but other cell types have been used, including yeast, insect and plant [91]. The S30 extract contains all the necessary enzymes and other transcription and translation components for protein synthesis, and together with an exogenous supply of amino acids, nucleotides, salt and energy-generating factors, protein overexpression in a test tube is feasible. The transcription of the target gene in the cell-free system is under the control of the T7 promoter [92-94]. Although cell-free protein synthesis from several organisms has been developed (*e.g.*, *E. coli* cells, wheat germ embryos and rabbit reticulocytes), *E. coli* extracts were used for cell-free APA-1 expression because they generally give higher protein yields and provide more NMR homogeneous samples suitable for protein structural analysis [88, 95, 96]. On the negative side, amino acid metabolism by *E. coli* can lead to isotope scrambling, *e.g.*, transaminase activities [84, 88, 97, 98]. Nonetheless, the cell-free system is far more inert with regard to isotope scrambling because the pool of metabolic enzymes present in the cell extract is not regenerated, and transamination between Asp/Asn and Glu/Gln occurs as a side reaction. This will be discussed in detail in Section 2.3.4.2.

To extend the cell-free reaction time and increase protein productivity, a continuous exchange cell-free system was developed [91, 99]. In this system, a reaction mixture is supplied with a feeding solution through a semi-permeable membrane and agitated in a shaking incubator (Figure 2.1). The reaction mixture consists of high MW compounds necessary for gene expression, such as the cell extract, nucleic acids, tRNAs and the DNA template. The feeding solution contains essential amino acids, salts, ATP and energy generating factors to supply fresh precursors into the reaction mixture while continuously removing inhibitory by-products

from the reaction mixture. The volume:volume ratio of the reaction mixture and feeding mixture is usually 1:10. The semi-permeable membrane with a molecular weight cut-off (MWCO) between 10 and 50 kDa is used to ensure effective exchange between the reaction mixture and the feeding mixture with no loss of essential proteins due to the formation of large complexes through multiple interactions [100].

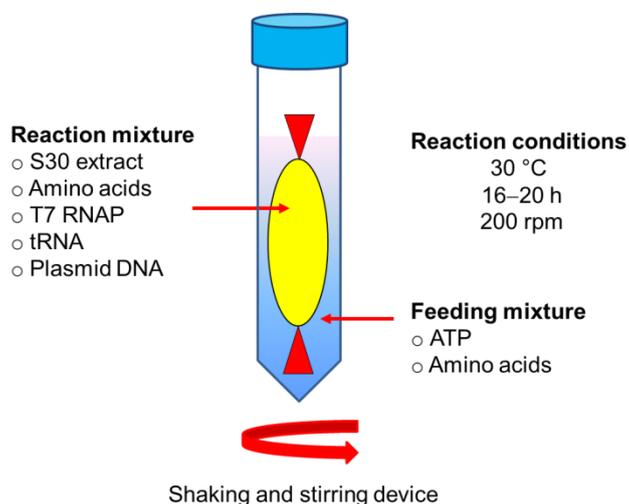


Figure 2.1: Schematic picture of the cell-free protein expression dialysis (continuous flow) system.

2.1.2 Advantages of using cell-free protein synthesis

Cell-free protein synthesis offers several advantages over conventional *in vivo* protein production. The reaction is fast and can be carried out in small volumes, and it is independent of cell viability, thereby allowing the production of cytotoxic and membrane proteins. In addition, the target protein is the only protein synthesised and isotope-labelled during the reaction. Consequently, isotope-labelled amino acids are used very efficiently allowing inexpensive and rapid isotope labelling for studying proteins by NMR spectroscopy. Proteins can be uniformly, selectively, site-specifically or combinatorially isotope-labelled for NMR spectroscopic analysis by simply incorporating particular isotope-labelled amino acids in the cell-free reaction system [84, 101-105]. Moreover, cell-free protein synthesis allows incorporation of non-natural or chemically modified amino acids into the expressed protein at desired positions during translation by including modified tRNAs [106, 107]. Proteins can be labelled with fluorescent or biotinylated amino acids, or photo-reactive cross-linked groups for sensitive detection and functional analysis [108].

2.1.3 NMR spectroscopy

2.1.3.1 One-dimensional ^1H NMR experiments

One-dimensional (1D) ^1H NMR experiments can provide insight into whether APA-1 expressed from the cell-free synthesis is folded and monomeric. The dispersion of signals in the amide proton region (6–10 ppm) and the up-field shift of resonances arising from methyl protons (~ 0 ppm) buried within the protein core provide the main indicators of a folded globular state in a 1D spectrum. In contrast, poorly folded proteins give rise to poor peak dispersion and resonance broadening of backbone amide protons due to chemical exchange phenomena (Figure 2.2). Consequently, these proteins are difficult to study by NMR spectroscopy [109]. Moreover, the approximate size of a protein can be estimated by obtaining the transverse relaxation time from the 1D ^1H spectra [109]. More details will be described in Section 2.2.12.1.

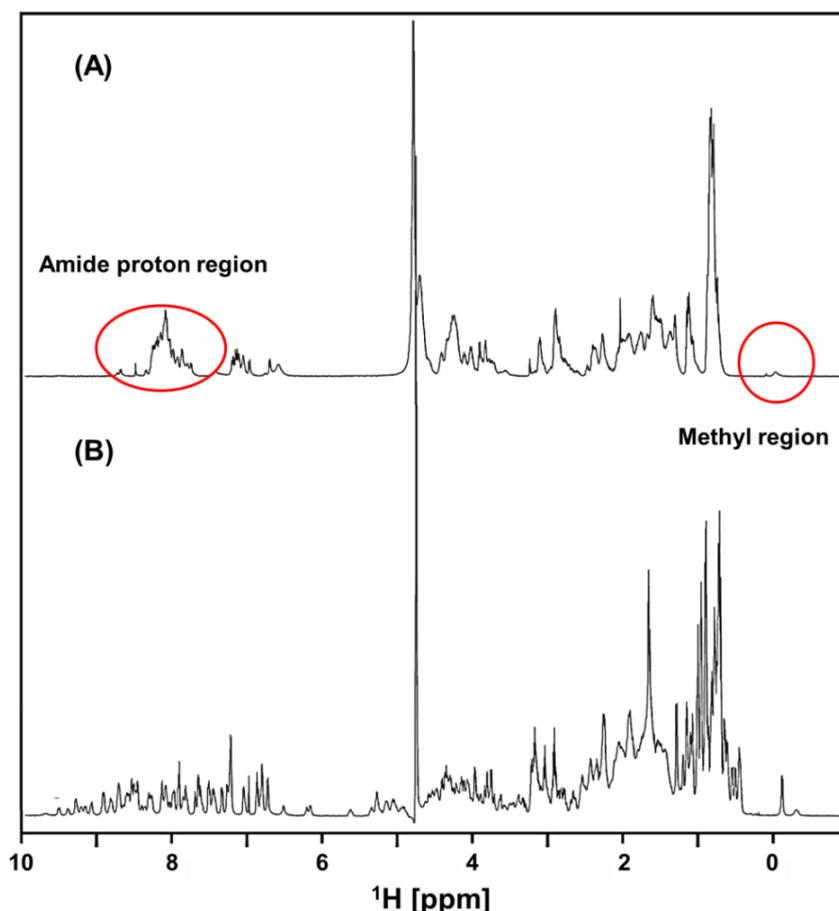


Figure 2.2: One-dimensional spectra of folded and unfolded ubiquitin. (A) 1D spectrum of a 2 mM ubiquitin sample in 8 M urea solution. The reduced chemical shift dispersion (*e.g.*, the amide proton region) is characteristic of unfolded proteins. The methyl group is upfield shifted. (B) 1D spectrum of a 2 mM folded ubiquitin sample in H_2O . Both spectra were acquired at 27 °C. The picture was modified from [110].

2.1.3.2 Two-dimensional NMR

Proteins give rise to complex 1D ^1H NMR data because of the large number of protons within the structure. Such spectra have significant resonance overlap that makes any sensible analysis of the spectra very challenging. By introducing additional spectral dimensions, this resonance overlap is often removed and detailed protein structural analysis or other types of analysis, such as dynamics, is possible. In addition, extra information such as how atoms are bonded and the details of through-space proximity between nuclei can be obtained [111, 112].

A 2D NMR experiment involves four time periods: preparation, evolution, mixing and detection (Figure 2.3) [111, 112]. The preparation period allows the nuclei in the protein sample to reach equilibrium with the external magnetic field environment. After a set of radio frequency (RF) pulses, the nuclear spins precess freely at their own frequencies during the evolution period with delay time t_1 . The frequencies are measured (indirectly) during this evolution period by incrementing this delay period (see below) to give rise to one of the two chemical shift axes. During a mixing period, if two nuclei interact with each other (*e.g.*, either through-bond (electron mediated) or through-space (dipole-dipole coupling) interactions), the free induction decay (FID) signal obtained from one nucleus at the end of t_1 is manipulated by the frequency of the second nucleus, which can be recorded as an observable signal during the detection period (t_2). To obtain 2D spectra, the process is repeated with gradually increased t_1 values (*i.e.*, incremented). A series of FIDs obtained from incremented t_1 values are Fourier transformed along t_2 values to obtain a mixed frequency/time-domain interferogram. The intensities of the signals in the spectra oscillate as a function of t_1 at t_2 in units of Hz. The second dimension is achieved by Fourier transforming along t_1 values (Figure 2.4) [111].

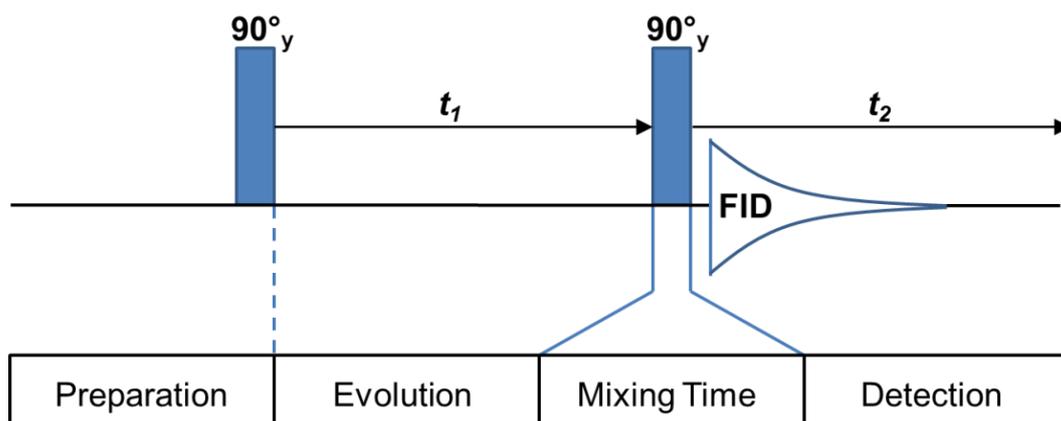


Figure 2.3: Schematic representation for a 2D NMR experiment. The cycle contains four periods: preparation, evolution, mixing time and detection. A series of 1D FIDs are generated with different delay times, t_1 . After the mixing period, the signal is detected during time t_2 . Fourier transformation of the series of FIDs gives a 2D spectrum in which peaks appear that correlate the directly observed chemical shift with the interaction of interest. The preparation and mixing period may contain one or a series of RF pulses [113].

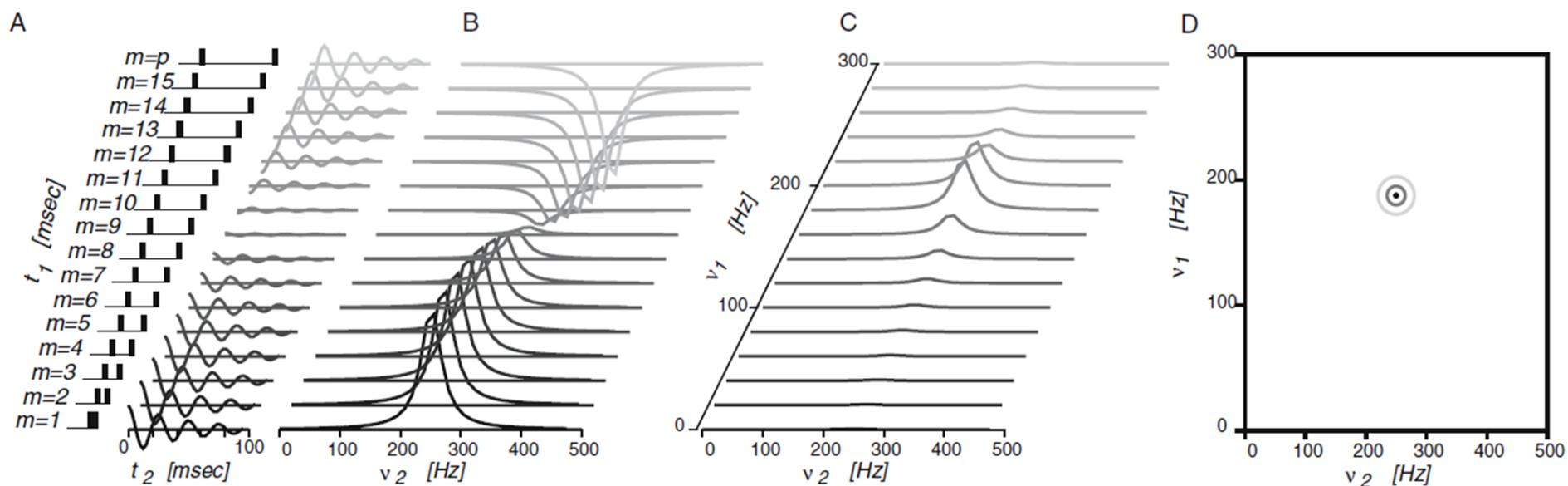


Figure 2.4: Generation of a 2D NMR spectrum. (A) t_1 values in the evolution period are incremented. Each FID is obtained from each t_1 value and a series of FIDs are recorded as a function of t_2 which are recorded in real time. (B) The FIDs are individually Fourier transformed and the x-axis is converted from the time domain to the frequency axis. (C) A second Fourier transformation is carried out along t_1 . (D) A 2D spectrum showing the same spectral information as displayed in (C), the darkest dot represents the strongest height intensity, whereas the feint circle shows regions of less intensity. The picture was taken from [114].

To characterise the conformational change of APA-1 in different solution environments, a fully assigned two-dimensional (2D) ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) spectrum was required. The 2D ^1H - ^{15}N HSQC experiment correlates the nitrogen nucleus of an amide backbone group (*i.e.*, N-H) with the directly attached amide proton nucleus. In the 2D ^1H - ^{15}N HSQC spectrum, each amide backbone group within a uniformly ^{15}N -enriched protein gives rise to one resonance in the spectrum, and each NH_2 group from the side chains of Asn and Gln residues give rise to two resonances in the spectrum. In addition, the indole side chain (NH) group of Trp residues shows one resonance in the spectrum for each Trp residue present, typically around 10–12 ppm in the proton dimension. This spectrum provides a diagnostic fingerprint, and is often used to measure various structural and functional features of a protein [115]. The dispersion of the resonances in the 2D ^1H - ^{15}N HSQC spectra is also indicative of the folded state of the protein. In the spectrum of a folded protein, resonances are usually well-dispersed of equal intensity with the number of resonances corresponding to the number of residues in the protein sequence (excluding prolines, which do not have amide protons). In contrast, a spectrum of an unfolded protein often lacks resonance dispersion and the number of resonances is often inconsistent with the protein sequence [116].

2.1.3.3 Three-dimensional HNCA to obtain sequence specific assignments

In a 2D NMR experiment, the FID signals are collected after the mixing period. In the case of a three-dimensional (3D) experiment, an additional evolution and mixing period are applied before acquiring the FID signal. The t_1 and t_2 (time delays incremented in the evolution periods) are separately Fourier transformed to give a 3D spectrum [117].

The 3D HNCA NMR experiment is a conventional approach to obtain sequence specific assignments of the amide groups and the $\text{C}\alpha$ nuclei, and the assignments can be mapped on to the 2D ^1H - ^{15}N HSQC spectrum. The 3D HNCA spectrum has three orthogonal chemical shift coordinates: the ^1H amide chemical shift, the ^{15}N amide chemical shift and the $^{13}\text{C}\alpha$ chemical shift [118, 119]. The HNCA correlates one ^1H and ^{15}N group (*i.e.*, chemical shift of both nuclei are recorded) to two $^{13}\text{C}\alpha$ nuclei. One correlation is between the amide group with its corresponding $^{13}\text{C}\alpha(i)$ nuclei (*i.e.*, intra-residue), and the other correlation is between the same amide group and the $^{13}\text{C}\alpha(i-1)$ of the previous residue (*i.e.*, inter-residue) (Figure 2.5A) [118]. It is not unambiguous to identity between inter- and intra- correlations in the HNCA experiment alone; although the intra- $\text{C}\alpha$ resonances generally have greater intensity than the inter- $\text{C}\alpha$ resonances because the N to intra- $\text{C}\alpha$ scalar coupling is stronger (~ 11 and ~ 7 Hz for

N-intra- C_{α} and N-inter- C_{α} couplings, respectively) [120]. To achieve unambiguous backbone assignment, a 3D HN(CO)CA experiment is acquired. In the HN(CO)CA experiment, the ^1H and ^{15}N chemical shifts are correlated with only the inter- C_{α} nuclei via the CO group, and therefore only the C_{α} correlation from the previous residue is observed in the spectrum (Figure 2.5B). By overlaying the two spectra, backbone assignments can be achieved (Figure 2.5C). However, in this Chapter, the combinatorial labeling approach coupled with the available ^1H chemical shift assignments (Prof. Joachim Grötzinger, Department of Biochemistry, Christian-Albrechts-Universität zu Kiel, Medical Faculty, Kiel, Germany) were used to obtain the majority of the sequential assignment in the 2D ^1H - ^{15}N HSQC spectrum without the requirement to acquire the 3D HN(CO)CA experiment (Section 2.3.4.2).

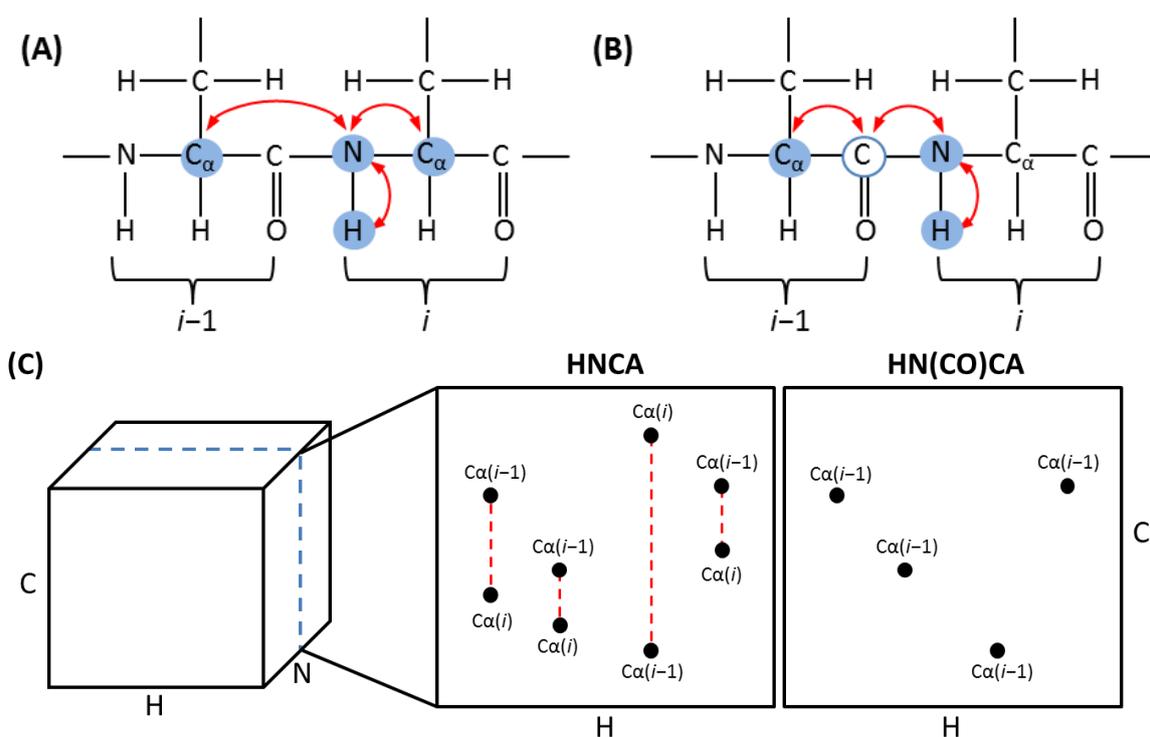


Figure 2.5: Three-dimensional HNCA and HN(CO)CA experiments. (A) and (B) represent the schematic structures illustrating the magnetisation transfer pathway of the HNCA and HN(CO)CA experiments, respectively. (C) Displays a schematic of the 3D HNCA and HN(CO)CA spectra. (A) The amide nitrogen is coupled with the C_{α} nuclei of both inter- ($i-1$) and intra-residues (i). In each NH strip both inter- and intra- C_{α} resonances are visible in the 3D HNCA spectrum as displayed in (C). (B) The amide group is only coupled with the inter- C_{α} nucleus in the HN(CO)CA spectrum at each NH “strip”. By overlaying the HN(CO)CA spectrum with the HNCA, all $C_{\alpha}(i)$ and $C_{\alpha}(i-1)$ resonances can be distinguished. The double headed arrows indicate that the experiment is an “out and back” experiment (*i.e.*, the initially excited proton nucleus and the detected proton nucleus are the same).

2.1.4 Aims

The aims of this chapter were to: (1) use the cell-free system to produce cytotoxic APA-1 in both unlabelled and isotopically labelled forms; (2) develop a purification protocol to obtain pure APA-1 and examine its activity using the liposome depolarisation assay; and (3) use the amino-acid type combinatorially ^{15}N - and uniformly $^{13}\text{C}/^{15}\text{N}$ -labelled APA-1 to facilitate backbone resonance assignments at pH 3.0 using NMR spectroscopy. In Chapter 3, the assignment of the 2D ^1H - ^{15}N HSQC spectrum obtained from this chapter was used to identify amino acids involved at the dimer interface. In addition, the cell-free system was also used to express amino acid type selectively $^{15}\text{N}/^{13}\text{C}$ -labelled APA-1 and various APA-1 mutants to study the APA-1 dimerisation mechanism using NMR spectroscopy and other biophysical methods (*e.g.*, small angle X-ray scattering and size exclusion chromatography).

2.2 Methods and Materials

2.2.1 *Amoebapore A* plasmid construct

The DNA sequence data encoding the APA-1 gene is available from the EMBL/GenBank databases under the accession number M83945 [25]. The gene encoding APA-1 (77 residues) was previously cloned by our collaborators in Germany (Prof. Joachim Grötzinger, Department of Biochemistry, Christian-Albrechts-Universität zu Kiel, Medical Faculty, Kiel, Germany) into the pIVEX 2.4d plasmid (F. Hoffmann-La Roche Ltd., Basel, Switzerland) using *NotI* and *BamHI* cleavage sites (Figure 2.6). The resulting expression vector, pIVEX2.4d-APA-1, consists of the APA-1 gene fused with a N-terminal (His)₆-tag and a Factor Xa cleavage site. The amino acid sequence of the fusion protein, as determined by nucleotide sequencing is: MSGSHHHHHSSG **IEGR** GEILCNLCTG¹⁰ LINTLENLLT²⁰ TKGADKVKDY³⁰ ISSLCNKASG⁴⁰ FIATLCTKVL⁵⁰ DFGIDKLIQL⁶⁰ IEDKVDANAI⁷⁰ CAKIHAC⁷⁷. The bold lettering represents the Factor Xa recognition site.

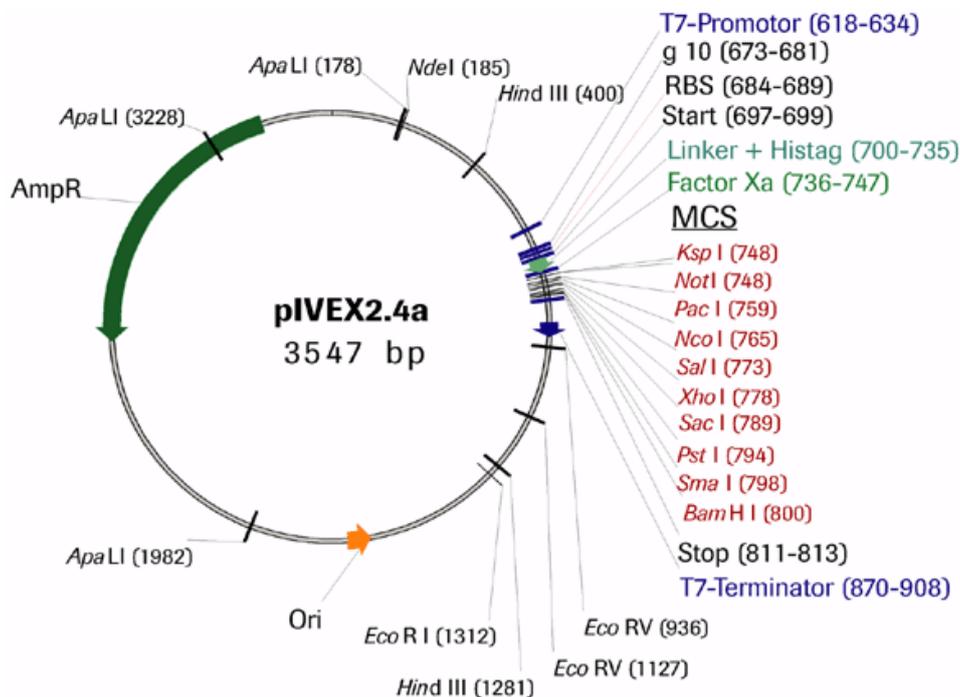


Figure 2.6: pIVEX 2.4a plasmid as a DNA template for the expression of APA-1. The picture was taken from [121].

2.2.2 Transformation of plasmids

To carry out transformations, 25 ng of plasmid was incubated on ice with 50 µl of *E. coli* DH5α or Top 10 competent cells (Invitrogen, Carlsbad, USA) for 30 minutes. Heat shock of the cells was carried out at 42 °C for 2 min to ensure uptake of the plasmid and 500 µl of Luria-Bertani (LB) media was immediately added to the transformed cells, followed by 2 h of incubation at 37 °C with agitation at 250 rpm. One-hundred microliters of the bacterial strain was taken and plated onto LB/agar containing 0.1% (w/v) ampicillin and incubated overnight at 37 °C. A single colony was selected from the plate and inoculated in 10 ml of LB medium/0.1% (w/v) ampicillin for 8 h at 37 °C and 250 rpm. One millilitre of the culture was transferred to 1 L of LB medium/0.1% (w/v) ampicillin and the large culture incubated for a further 12–16 h. Plasmid DNA was extracted from the overnight culture using the Invitrogen Mini/Midi/Maxi Prep kits (Invitrogen, Carlsbad, USA) or the QIAprep[®] Mega Prep kit (QIAGEN, Limburg, Netherlands). The extracted DNA material was analysed on a 0.8% (w/v) agarose gel and the concentrations were measured photometrically at a wavelength of 280 nm using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, USA).

2.2.3 Preparation of the *E. coli* S30 extract

An S30 extract was prepared by MSc student Mikhail Legkodimo (School of Biological Science, The University of Auckland, 2010) as described by Pratt [91] from the *E. coli* strain Rosetta (λDE3)/pRARE, with two-fold concentration by dialysis using polyethylene glycol 2000 [122].

2.2.4 Preparation of the T7 RNA polymerase

2.2.4.1 pKO1166 plasmid construct

The pKO1166 plasmid which contains the phage T7 *gene1* that encodes a T7 RNA polymerase (T7 RNAP) (Figure 2.7) was used to express the T7 RNAP during cell-free reactions. The plasmid was provided by Dr Andrew Kralicek at The New Zealand Institute for Plant and Food Research Limited (Auckland). The pKO1166 uses the transcription factors present in the S30 extract to synthesise the T7 RNAP under the control of the phage ρ_L promoter [92]. With the continuous production of T7 RNAP, APA-1 was expressed under the control of the T7 promoter.

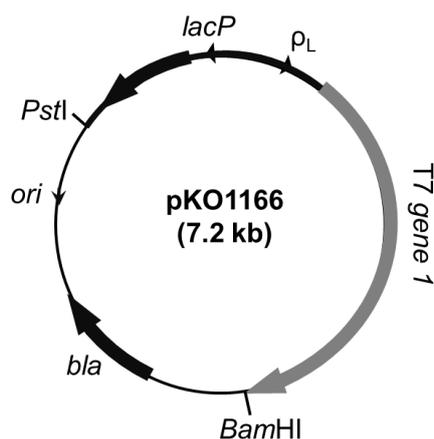


Figure 2.7: The pKO1166 plasmid construct. The pKO1166 plasmid contains the T7 *gene 1* which directs the expression of T7 RNAP under the control of the bacteriophage ρ_L promoter. The figure was made according to [92].

2.2.4.2 Culture preparation

Apart from co-synthesising T7 RNAP during the cell-free reaction, it can also be expressed and purified from cell culture and subsequently added to the reaction mixture. The expression of T7 RNAP was achieved using the pKO1166 vector as described in [92] with minor modifications. The pKO1166 vector was transformed into *E. coli* BL21 competent cells by heat shock, as described in Section 2.2.2, at 30 °C. A 2 L culture of competent cells was grown at 30 °C in LB medium supplemented with 0.1% (w/v) ampicillin and 5% thiamine (w/v). The culture was incubated at 30 °C and 250 rpm. When the A_{600} reached 0.9, the temperature was raised to 42 °C for 30 min and subsequently lowered to 40 °C for a further 2 h incubation. The cells were harvested by centrifugation at 10,816 g for 10 min at 4 °C. Cell pellets were re-suspended in lysis buffer (50 mM Tris-HCl, pH 7.9, 5% (v/v) glycerol, 0.5 M NaCl, 1 mM DTT, 2 tablets of “Complete EDTA-free” protease inhibitors (F. Hoffmann-La Roche Ltd., Basel, Switzerland), 10 mM spermidine and 0.05% Na-deoxycholate) at a ratio of 5 g wet cell pellet per 40 ml buffer.

2.2.4.3 Salt precipitation

The cell solution was passed through an Avestin EmulsiFlex-C5 homogenizer (Avestin Inc, Canada) at ~103,421 kPa (15,000 psi) and the lysed cells were spun at 11,000 g for 45 min. Working at a temperature of 4 °C, proteins were precipitated from the supernatant by slowly adding ground ammonium sulfate to a final concentration of 0.35 g/ml over 20 min while mixing. The solution was incubated for 20 min to maximise the precipitation and centrifuged at

11,000 g for 45 min. The pellet containing T7 RNAP was dissolved into TGED buffer (10 mM Tris-HCl, pH 7.9, 5% (v/v) glycerol, 0.1 mM EDTA and 1 mM DTT) with a volume equal to 20 times of the original weight of the cell pellets and dialysed overnight against the TGED buffer with a volume equal to 20 times the dialysate volume. The dialysate was centrifuged at 11,000 g for 10 min and the supernatant was collected.

2.2.4.4 DEAE chromatography

Twenty millilitre aliquots of the supernatant containing T7 RNAP were loaded onto a High-Trap diethylaminoethyl (DEAE) cellulose-ion exchange column connected to a gradient ÄKTA purifier (GE Healthcare, Little Chalfont, UK) that has been pre-equilibrated with five column volumes of TGE buffer (10 mM Tris-HCl, pH 7.9, 5% (v/v) glycerol and 0.1 mM EDTA). The sample was loaded at 5 ml/min. UV absorption at 280 nm was monitored at room temperature. Bound proteins were eluted with a linear gradient of 0–1 M NaCl in TGE buffer over 30 min at a flow rate of 5 ml/min. Two-and-a-half millilitre fractions were collected and analysed by 4–12% SDS-PAGE (Appendix 1). The T7 RNAP fractions were pooled, concentrated to 30 ml using VivaspinTM centrifugal concentrators (GE Healthcare, Little Chalfont, UK) with 10,000 Da MWCO, then dialysed against 2 L of buffer P (10 mM potassium-phosphate (KPi), pH 7.5 and 0.1 mM EDTA) overnight.

2.2.4.5 Cellulose phosphate affinity chromatography

For each run, 10 ml of the dialysed protein solution was loaded onto a cellulose phosphate affinity column (Whatman P-11, 2.0 cm diameter, 19.5 cm length) connected to a gradient ÄKTA purifier (GE Healthcare, Little Chalfont, UK) that has been pre-equilibrated with five column volumes of buffer P at 1 ml/min. The method for packing the column is described in Appendix 2. UV absorption at 280 nm was monitored at room temperature. The bound proteins were eluted with a linear gradient of 0–1 M of KCl in buffer P over 80 minutes at a flow rate of 1 ml/min. Two millilitre fractions were collected and analysed by 4–12% SDS-PAGE. The T7 RNAP fractions were pooled and concentrated to 10 ml using 10,000 Da MWCO VivaspinTM centrifugal concentrators (GE Healthcare, Little Chalfont, UK). The protein solution was dialysed against 1 L of storage buffer (40 mM KPi, pH 7.5, 0.2 mM EDTA and 2 mM DTT) and further concentrated to 10–20% of the original volume. The protein concentration was determined photometrically at 280 nm using Eq. 2.1:

$$\text{T7 RNAP concentration (mg/ml)} = (Abs_{280}/1.426 \text{ L mol}^{-1}\text{cm}^{-1}) \times 99 \text{ kDa} \quad \text{Eq. 2.1}$$

where $1.426 \text{ L mol}^{-1} \text{ cm}^{-1}$ is the molar extinction coefficient (ϵ_{280}) calculated based on the molar extinction coefficient of tyrosine, tryptophan and cysteine residues in the T7RNAP amino acid sequence. The MW of T7RNAP is 99 kDa. The T7 RNAP solution was mixed with glycerol at a 1:1 ratio (v/v) and stored at $-20 \text{ }^\circ\text{C}$.

2.2.5 Preparation of the Master Mix

Working on ice, the following stock solutions for making the Master Mix were prepared: 2 M HEPES pH 7.5, 0.17 M DTT, 120 mM ATP, 80 mM each of CTP, GTP and UTP, 64 mM cyclic-AMP, 6.8 mM folinic acid, 2.75 M ammonium acetate, 4.16 M KOH-(L)-glutamic acid or 2.08 M KOH-L-glutaric acid, 1.93 M Mg-acetate and 1.6 M creatine phosphate. The stock solutions were mixed together in the above sequence and stored at $-80 \text{ }^\circ\text{C}$.

2.2.6 Preparation of amino acids

Amino acid stock solutions were prepared in four distinct groups, A, B, C and D. The concentration of each amino acid stock solution was 100 mM. The group A stock consisted of strong acid soluble amino acids: Glu, Asp, Gln, Thr, Met and Tyr, dissolved in 0.7 M HCl. The group B stock consisted of base-soluble amino acids: Ile, Trp, Cys and Phe, dissolved in 1.09 M KOH. The group C stock consisted of Ala, Arg, Lys, Pro and Ser in ultra-pure water, and the group D stock consisted of weak acid-soluble amino acids Asn, Gly, His, Leu, and Val dissolved in 0.11 M HCl. The amino acids were dissolved in different solutions so that when the stocks were mixed, the pH of the mixture was ~ 7 . Vigorous vortex mixing and incubation at $50 \text{ }^\circ\text{C}$ for 10 min was required to fully dissolve the amino acids. The stock solutions were mixed together to a final concentration of 15 mM in the order of A, D, C and B to minimise possible precipitation of the amino acids due to sharp changes in pH upon mixing. The final reaction mixture contained 1 mM of each amino acid.

To prepare crude amino acid mixtures (Cambridge Isotope Laboratories, Tewksbury, USA), two amino acid stocks with the same concentration were prepared in 1 M HCl and 1 M KOH. Ten minutes of vigorous mixing was required to dissolve the amino acids. Once the stock solutions were mixed, they were immediately added to the Master Mix solutions to prevent amino acid precipitation. Stock solutions that were not used were stored at $-80 \text{ }^\circ\text{C}$. Prior to the cell-free reactions, the stock solutions were thawed at room temperature and incubated at $50 \text{ }^\circ\text{C}$ for 5 min followed by vigorous mixing to (re)dissolve the amino acids.

For the preparation of ^{15}N - and $^{15}\text{N}/^{13}\text{C}$ -isotope labelled amino acid solutions, unlabelled amino acids were replaced with the required labelled amino acids. Crude ^{15}N -labelled or $^{13}\text{C}/^{15}\text{N}$ -labelled amino acid mixes or individual L- ^{15}N -asparagine: H_2O , L- ^{15}N -cysteine, L- ^{15}N -glutamine, L- ^{15}N -glycine, L- ^{15}N -isoleucine, L- ^{15}N -phenylalanine, L- ^{15}N -serine, L- ^{15}N -tryptophan, L- ^{15}N -tyrosine, L- ^{15}N -valine, L- ^{15}N -alanine, L- ^{15}N -leucine, L- ^{15}N -threonine, L- $^{13}\text{C}/^{15}\text{N}$ - or L- ^{15}N -aspartic acid, L- $^{13}\text{C}/^{15}\text{N}$ - or L- ^{15}N -glutamic acid, L- $^{13}\text{C}/^{15}\text{N}$ - or L- ^{15}N -histidine: $\text{HCl}:\text{H}_2\text{O}$, and L- $^{13}\text{C}/^{15}\text{N}$ - or L- ^{15}N -lysine: 2HCl were purchased from Cambridge Isotope Laboratories (Tewksbury, USA). To isotopically label glutamic acid residues, 100 mM of ^{15}N - or $^{13}\text{C}/^{15}\text{N}$ -glutamic acid in 0.7 M HCl was separately prepared, added to the group A mixture then mixed with D, C and B (in that order) to give a final concentration of 15 mM. When pKO1166 was used to replace pure T7 RNAP for continuous production of T7 RNAP during a cell-free reaction, twice the amount of L-alanine (*i.e.*, 200 mM) was added to the amino acid stock because T7 RNAP is rich in alanine residues (*i.e.*, 11.3% of the T7 RNAP residues are alanine) [123]. This higher amount of alanine ensured sufficient supply of L-alanine for APA-1 production.

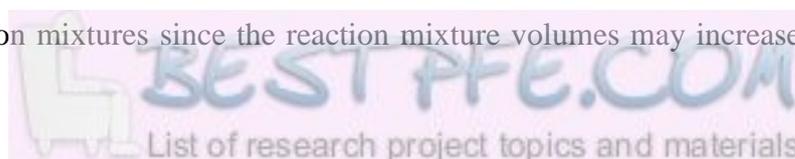
2.2.7 Cell-free protein synthesis

The solutions used in the reaction mixture of the continuous exchange cell-free system are listed (Table 2.1).

Table 2.1: Components of the final reaction mixture of the continuous exchange cell-free system

Reagents	Concentrations
HEPES/KOH, pH 7.5	55 mM
DTT	1.7 mM
ATP	1.2 mM
Each of CTP, GTP and UTP	0.8 mM
Cyclic AMP	0.64 mM
Folinic acid	68 μ M
Ammonium acetate	27.5 mM
KOH-glutamic acid or KOH-glutaric acid for labelling with L-[¹⁵ N]-glutamic acid	208 mM
Mg-acetate	19.3 mM
Creatine phosphate	80 mM
Creatine kinase	0.25 mg/ml
Ribonuclease inhibitor (RNasin) (Promega, Fitchburg, USA)	250 units/ml
<i>E. coli</i> total tRNA	175 μ g/ml
Rosetta (λ DE3)/pRARE S30 extract	24% (v/v)
Purified T7 RNAP or pKO1166	186 μ g/ml T7 RNAP or 32 μ g/ml pKO1166
Individual pure amino acids or crude amino acid mixture	1 mM pure amino acids or 0.03 mg/ml crude amino acid mixture
pIVEX2.4d-APA plasmid	16 μ g/ml

The feeding mixture was identical to the reaction mixture, except that the S30 extract, tRNA, plasmid DNA, T7 RNAP, creatine kinase and RNasin were absent. The pH of the feeding mixture was adjusted to 7.5 using 1 M KOH. The concentrations of the crude amino acids, RNasin, Rosetta (λ DE3)/pRARE S30 extract and purified T7 RNAP were obtained from optimisation of the protein expression level, as described in Section 2.2.8. All the cell-free reactions were performed at 30 °C with agitation at 200 rpm overnight (Figure 2.1). Membranes with MWCO of 12–14 kDa were used. All protein productions were visualised using 4–12% SDS-PAGE by loading 2 μ l of the reaction mixtures onto the gel. For comparative experiments, the SDS-PAGE loading volumes varied according to the final volume of the reaction mixtures since the reaction mixture volumes may increase or decrease



via dialysis during the overnight reaction. In addition, protein production was also assessed by comparing the expression levels of target proteins and the background proteins.

2.2.8 Analytical scale cell-free protein synthesis

Analytical cell-free reactions with 50 μ l of the reaction mixture and 950 μ l of the feeding mixture were initially carried out to optimise the reaction conditions in order to obtain the highest yields of protein production by titration of each component. Control reactions containing 16 μ g/ml of the IL6+ mutant DNA were used to test the viability of the reaction conditions because it generally gives high expression levels and precipitation, giving a visual confirmation of synthesis. Five to ten microliters of the reaction mixture was collected before starting the exchange with the feeding mixture and stored separately at 4 °C as the T₀ control. The reactions were performed in Slide-A-Lyzer mini dialysis units (Thermo Scientific, Waltham, USA) that were pre-soaked in ultra-pure water for 30 min and a 1 ml Nunc Cryotubes which contains the feeding mixture (Thermo Scientific, Waltham, USA) (Figure 2.8).

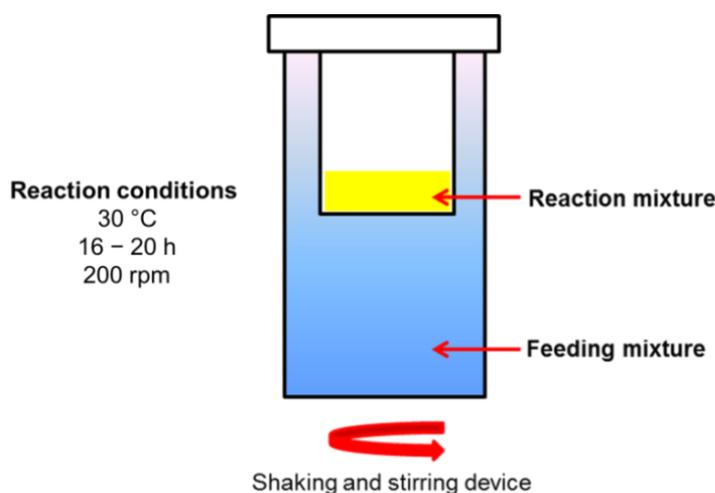


Figure 2.8: Analytical scale cell-free protein synthesis. The analytical scale cell-free reactions were conducted in mini dialysis units. The volumes of the reaction and feeding mixtures were 50 and 950 μ l, respectively.

2.2.8.1 Testing the Master Mix and amino acid stock solutions

Following the preparation of a new Master Mix and/or amino acid mixture, analytical scale cell-free reactions were performed to test the quality of the mixtures. The IL6+ mutant protein was used as the control.

2.2.8.2 Optimisation of the Rosetta (λ DE3)/pRARE S30 extract volume in the cell-free reaction

The concentration of the Rosetta (λ DE3)/pRARE S30 extract prepared (Section 2.2.3) was optimised at the analytical scale to achieve the maximal protein yield while using the lowest amount of the S30 extract. Reaction mixtures containing 12, 24, 30 and 36% (v/v) of the S30 extract were prepared and APA-1 production from the reactions was examined. A cell-free reaction using a previous lab stock of an S30 extract that was optimised to give the highest protein yield at a concentration of 24% (v/v) was used as the control.

2.2.8.3 Optimisation of the RNasin concentration in the cell-free reaction

RNasin (F. Hoffmann-La Roche Ltd., Basel, Switzerland) is one of the most expensive chemicals in the cell-free reaction. To reduce the cost and maximise the level of protein production, the concentration of RNasin was optimised at the analytical scale. Protein productivity from the reaction mixtures containing 50, 150, 250, 300, 350, 400, 450, 500 and 550 units of RNasin per ml of the reaction were tested. Lab made and optimised RNasin by MSc student Mikhail Legkodimo (School of Biological Science, The University of Auckland, 2010) as described in [124] with small modifications was also occasionally used.

2.2.8.4 Optimisation of the T7 RNAP concentration in the cell-free reaction

The activity of the purified T7 RNAP was tested by titrating the T7 RNAP concentration to find the highest yield of APA-1 produced. Protein yields from the reaction mixtures containing 0.7, 1.05, 1.4, 1.75, 2.8 and 4.2 μ g of T7 RNAP per ml of cell-free reaction were compared using 4–12% SDS-PAGE.

2.2.8.5 Optimisation of the amino acid concentration in the cell-free reaction

Crude $^{13}\text{C}/^{15}\text{N}$ -labelled amino acid mixtures (Cambridge Isotope Laboratories, Tewksbury, USA) were used for uniform $^{13}\text{C}/^{15}\text{N}$ -labelling of APA-1. To minimise the requirement of the expensive isotope labelled amino acids and maximise the protein yield, an amino acid concentration titration was performed using crude unlabelled amino acids. The crude amino acid mixtures with concentrations of 0.003, 0.0075, 0.015 and 0.03 g/ml were used for analytical scale cell-free reactions.

2.2.8.6 Unlabelled versus isotope labelled amino acids

Productivity of APA-1 from the cell-free reactions using isotope labelled amino acids (*i.e.*, $^{13}\text{C}/^{15}\text{N}$ - or ^{15}N -labelled) were compared with reactions carried out using the unlabelled amino acids. The reactions were performed at the analytical scale.

2.2.9 Preparative cell-free protein synthesis

For large scale APA-1 expression, the cell-free reactions were scaled-up from the 50/950 μl reaction to 1/10 ml, 2/20 ml and 3/30 ml reactions (reaction/feeding mixture). The yields of APA-1 expression were compared using the three systems. Reaction mixtures containing 1, 2 or 3 ml were put into dialysis bags (MWCO = 12–14 kDa). The dialysis bag was sealed at one end and trimmed to have a minimal overhang. A 100 μl microcentrifuge tube with a cut hinge was inserted into the other end of the dialysis tubing and the cap was closed. The microcentrifuge tube ensures the contents of the dialysis bag are readily accessible. The reaction mixtures were placed in a screw-capped 15 or 50 ml polypropylene tube containing 10, 20 or 30 ml of the feeding mixture (Figure 2.9). The reactions were incubated at 30 °C and 200 rpm. To produce isotope enriched APA-1 the unlabelled amino acids were replaced with labelled amino acids.

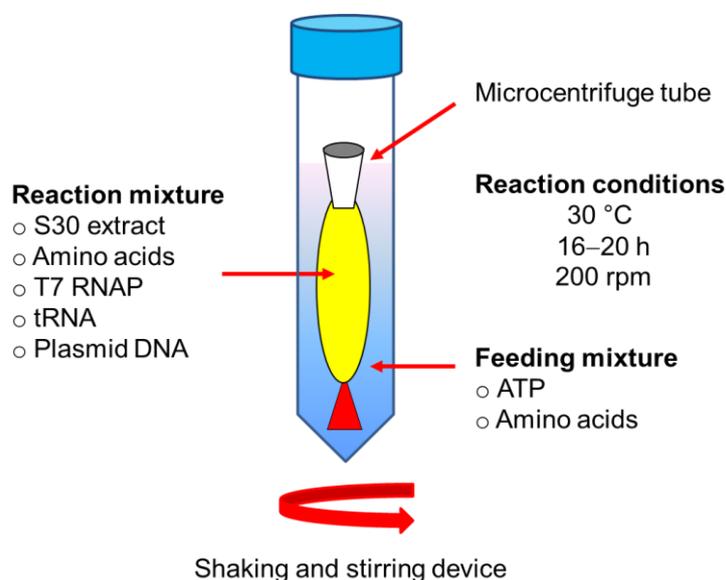


Figure 2.9: Preparative scale cell-free protein synthesis. The preparative scale cell-free reactions were conducted in 15 or 50 ml centrifuge tubes.

2.2.10 Synthesis of isotope labelled protein samples

To make isotopically enriched APA-1, ^{15}N - and $^{13}\text{C}/^{15}\text{N}$ -labelled amino acid mixtures (Cambridge Isotope Laboratories, Tewksbury, USA) were prepared. Combinatorial selectively ^{15}N -, uniformly $^{13}\text{C}/^{15}\text{N}$ - and selectively $^{13}\text{C}/^{15}\text{N}$ -enriched APA-1 samples were synthesised. The combinatorial selectively ^{15}N -enriched APA-1 samples were prepared by Dr Andrew Kralicek at The New Zealand Institute for Plant and Food Research Limited (Auckland). In the selectively $^{13}\text{C}/^{15}\text{N}$ -labelling approach, the charged amino acids (Asp, Gly, His and Lys) of APA-1 were $^{13}\text{C}/^{15}\text{N}$ -labelled for pKa measurements (see Chapter 3).

In the combinatorial labelling approach, five samples were produced each with particular residues ^{15}N -labelled simultaneously according to the combinatorial labelling scheme (Figure 2.10) [84]. The most abundant amino acids generally found in proteins, such as leucine and alanine, were labelled in only one of the five samples, whereas the least abundant amino acids such as histidine and cysteine were labelled in up to three of the samples. This scheme was based on the average amino acid frequencies in proteins reported in the National Centre for Biotechnology Information database. It was designed to minimise the number of resonances in the 2D ^1H - ^{15}N HSQC spectra and avoid any simultaneous ^{15}N -labelling of glutamine and asparagine as these residues increase the chance of signal overlap in the NMR spectra due to the resonances from side chain amide groups. By using the combinatorial labelling scheme, each 2D ^1H - ^{15}N HSQC spectrum recorded from one sample produces only a third of the resonances generated by recording a spectrum of a uniformly ^{15}N -enriched protein. Thus it was possible to minimise the number of resonances observed in each spectrum thereby limiting resonance overlap [84]. The combinatorial labelling method was combined with the available amide proton chemical shifts (Appendix 3) provided by Prof. Joachim Grötzinger (Department of Biochemistry, Christian-Albrechts-Universität zu Kiel, Medical Faculty, Kiel, Germany) as an initial approach to assign the 2D ^1H - ^{15}N HSQC spectrum of APA-1.

Residue type	Sample No.					Freq. (%)
	1	2	3	4	5	
L	●					14.3
R	●	●				0
N	●		●			6.5
Y	●			●		1.3
D	●				●	7.8
C	●	●			●	7.8
H	●		●	●		1.3
A		●				9.1
K		●	●			10.4
Q		●		●		1.3
F		●			●	2.6
W		●	●	●		0
M		●		●	●	0
G			●			6.5
I			●	●		11.7
T			●		●	7.8
S				●		3.9
V				●	●	3.9
E					●	3.9
P						0

Figure 2.10: Combinatorial selective ^{15}N -labelling scheme. Five combinatorial samples depicted by different colours are synthesised. Each sample contains a different combination of ^{15}N -labelled amino acids [88]. The last column represents the frequencies of each amino acid in APA-1.

2.2.11 Purification of amoebapore-A

2.2.11.1 Immobilised metal ion liquid chromatography

After completion of the cell-free reaction, the insoluble components were removed by centrifugation at 4,600 g for 20 min at 4 °C. Chromatography was performed on an ÄKTA Purifier system (GE Healthcare, Little Chalfont, UK). A HiTrap immobilised metal affinity chromatography (IMAC) column (GE Healthcare, Little Chalfont, UK) was used to purify the $(\text{His})_6$ -APA-1 from the cell-free reaction supernatant. The column loaded with 0.1 M $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$ was equilibrated with five column volumes of binding buffer (10 mM Na_2HPO_4 , 500 mM NaCl and 20 mM imidazole, pH 7.4). The cell-free supernatant was diluted by a factor of two with binding buffer and loaded onto the column at a flow rate of 0.5 ml/min. The

column was washed with five column volumes of binding buffer and APA-1 was eluted using the elution buffer containing 500 mM of imidazole (10 mM NaH₂PO₄, 500 mM NaCl and 500 mM imidazole, pH 7.4). Five millilitre fractions were collected and analysed by 4–12% SDS-PAGE. The fractions containing the His-tag APA-1 protein were pooled and extensively dialysed against 10 mM ammonium bicarbonate at pH 8.0 for at least six times (1 h each time) to remove imidazole. This solution was lyophilised.

2.2.11.2 Factor Xa Protease digestion to remove the (His)₆-tag

Lyophilised protein was reconstituted in 500 µl of 10 mM HCl and diluted in TBS digest buffer to give a final protein concentration of approximately 1 mg/ml (50 mM Tris, 6 mM CaCl₂ and 100 mM NaCl, pH 8.0), as determined photometrically. The fusion protein was digested with Factor Xa (F. Hoffmann-La Roche Ltd., Basel, Switzerland) at a protease:protein ratio of 1:50 (v/v) for 120 h at 25 °C. Protease cleavage was terminated by lyophilising the protein solution and the protein was stored at –80 °C. Protein samples from each day of the digestion were collected and analysed by 4–12% SDS-PAGE.

2.2.11.3 Reverse-phased high performance liquid chromatography

Reverse-phased high performance liquid chromatography (RF-HPLC) was used to purify APA-1 after the protease cleavage. After thawing, the protein sample was diluted two fold with 10% acetonitrile/0.1% trifluoroacetic acid (TFA) (v/v) and the pH adjusted to 3.0. The protein solution was loaded onto a semi-preparative C8 column (Grace Discovery Science, Columbia, USA; 10 x 250 mm, 5 µm particle size). A continuous linear acetonitrile/0.1% (v/v) TFA was applied at a flow rate of 5 ml/min with fraction volumes of 2 ml collected. APA-1 purity was confirmed by mass spectroscopy and 4–12% SDS-PAGE analysis. The amount of protein was determined by measuring the A₂₈₀, and using Eq. 2.1 and the calculated molar absorption coefficient (ϵ_{280}) of 1865 L M⁻¹ cm⁻¹. Fractions containing APA-1 were pooled, lyophilised and stored at –80 °C.

The average mass of APA-1 was determined by microTOP-Q MS at the School of Chemical Sciences, The University of Auckland. Samples were mixed with 0.7 µl of the matrix solution (5 mg/ml alpha-cyanohydroxycinnamic acid in 60% acetonitrile/0.1% TFA) and spotted onto a stainless steel sample target. Mass spectra were obtained by a 4700 Proteomics Analyzer (Applied Biosystems, Foster City, USA) in the positive-ion linear mode and calibrated internally using co-spotted standard proteins (Calibration Mixture 3, Applied Biosystems). The

average mass of the protein was compared with the calculated mass of disulfide-bridged APA-1.

2.2.12 NMR spectroscopy

Protein concentrations required for solution-state NMR experiments should be at least 50 μM for 2D experiments and ~ 500 μM for 3D experiments [125, 126]. After RP-HPLC, APA-1 samples were dissolved in 10 mM sodium citrate and disodium hydrogen phosphate (NaPi) 93% $\text{H}_2\text{O}/7\%$ D_2O (v/v). The pH adjusted to 3.0 and 300 μl of the protein solution was placed into a symmetrical matched microtube (Shigemi Inc., Allison Park, USA). The pH values of the samples were adjusted by the addition of 0.1 M NaOH or 0.1 M HCl. The pH values were measured before and after each experiment. The average readings were taken for data analysis.

All NMR spectra were recorded at 25 $^\circ\text{C}$ on a Bruker AV600 spectrometer equipped with a 5-mm z-axis pulsed-field gradient $^1\text{H}/^{15}\text{N}/^{13}\text{C}$ cryoprobe optimised for ^1H detection. All spectra were processed using the program NMRPipe. Analysis of the processed spectra were performed with the program Analysis which is part of the CcpNmr software package [127].

2.2.12.1 T_2 measurement

One-dimensional ^1H spectra of an APA-1 sample were recorded in ~ 5 minutes at pH 3 using spin-echo delay periods of 0.25 and 2.9 ms. The spectra were used to obtain the transverse relaxation time [128]. By measuring the difference in the resonance intensities between the two 1D spectra with the two delay periods, T_2 values were calculated using Eq. 2.2 [109]

$$T_2 \text{ (ms)} = 4(\Delta_A - \Delta_B)/\ln(I_A/I_B)$$

Eq. 2.2

where Δ_A and Δ_B are the delay periods used in the 1D spectra. I_A and I_B are the peak intensities recorded using short and long delay periods in the two 1D spectra, respectively. The T_2 value is predominantly dependent on the size of the protein and can be used to calculate approximate rotational correlation times (τ_c ; Eq. 2.3). The estimated τ_c value is the time it takes the protein to rotate by one radian and it depends on the size of the protein, thus τ_c can be used to provide a rough estimate of the protein MW (assuming the protein is a globular folded structure with no significant asymmetry). The MW of a protein in aqueous solution near room temperature is equal to two times the τ_c value [109]. Consequently, the dimerisation state of APA-1 was roughly estimated by comparing the calculated MWs of APA-1 at different pH values.

$$\tau_c = 1/[5 \times (T_2/1000)]$$

where T_2 is provided in ms and the calculated τ_c is given in ns.

2.2.12.2 Two dimensional ^1H - ^{15}N HSQC experiment

Two-dimensional (2D) ^1H - ^{15}N HSQC spectra were recorded using data matrices consisting of $128^* \times 1024^*$ data points (where n^* refers to complex points) with acquisition times of 62 (t_N) and 136 ms (t_{HN}). Depending on sample concentrations, between 16 to 256 scans per complex t_N increment were collected. A recycle delay of 1.2 s was used. The total measuring time for each experiment ranged between 40 min to 24 h. Data sets were processed using NMRPipe [129] and spectra visualised and resonances assigned using CcpNmr Analysis [130]. The spectra of the five combinatorially labelled samples were overlaid and coupled with the available ^1H chemical shift information used to assign the 2D ^1H - ^{15}N HSQC spectrum of APA-1 at pH 3.0.

2.2.12.3 Three-dimensional HNCA experiment

The three-dimensional HNCA spectrum was recorded using uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched APA-1 to sequentially assign the 2D ^1H - ^{15}N HSQC spectrum at pH 3. The sample concentration was 1.6 mM. The 3D HNCA was recorded as a constant-time water flip-back experiment [118]. The data matrix consisted of $45^* (t_1) \times 34^* (t_2) \times 1024^* (t_3)$ data points with acquisition times of 11.4 (t_1), 21.5 (t_2) and 95 ms (t_3). The total experimental time was 19 h. The ^1H carrier was positioned on the H_2O resonance, the ^{13}C carrier at 53 ppm and the ^{15}N carrier at 115 ppm. ^{15}N decoupling was applied during data acquisition. The data set was processed using NMRPipe and the spectrum visualised using the CcpNmr Analysis software [118]. To assign the resonances, the 3D spectrum was reduced to a series of 2D planes or “strips” which enables an easy approach to complete sequential assignment connectivities.

2.2.13 Pore-forming assay of amoebapore A

The pore-forming assay of APA-1 was performed by our collaborators in Germany (Prof. Joachim Grötzinger, Department of Biochemistry, Christian-Albrechts-Universität zu Kiel, Medical Faculty, Kiel) by monitoring the dissipation of a valinomycin-induced diffusion potential in liposomes at pH 5.2, 25 °C, as described in [131].

2.3 Results

2.3.1 Preparation and testing of the components of the cell-free synthesis

Prior to large scale protein production, the performance of the cell-free reactions were optimised to obtain the highest protein yields. The quantities of the Master Mix and amino acids were tested; the concentration of the S30 extracts, RNasin and T7 RNAP were prepared and optimised; and the expression level of APA-1 using isotope labelled and unlabelled amino acids were compared.

2.3.1.1 Testing of the Master Mix and amino acid stocks in the cell-free approach

When a Master Mix or amino acid stock solution was freshly made, analytical cell-free reactions were carried out to test the performances of the solutions. APA-1 expressions using newly prepared crude unlabelled amino acid mixtures and Master Mixes were investigated. Expression of the IL6+ mutant using the same solutions was performed as the positive control. The reaction mixture that was collected before the dialysis step was used as the T₀ control. The SDS-PAGE result (Figure 2.11) showed that the expression level of APA-1 was similar to the IL6+ control, and APA-1 was synthesised in the soluble form.

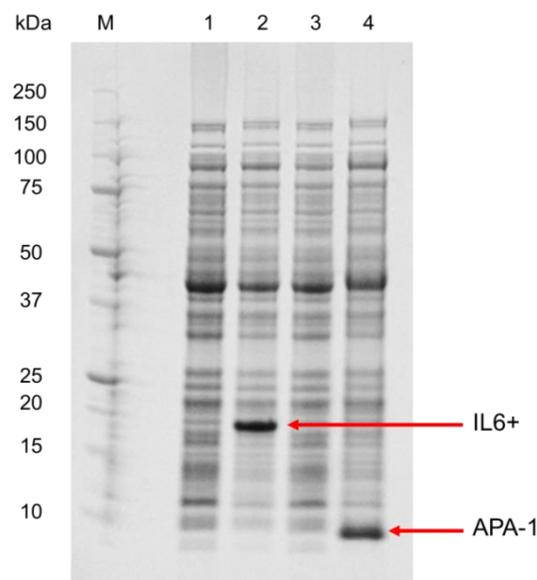


Figure 2.11: 4–12% SDS-PAGE analysis of the expression of IL6+ and APA-1 using a newly prepared Master Mix and a crude unlabelled amino acid mixture. Lane M: Precision plus MW marker (BioRad, Hercules, USA); Lanes 1 and 3: T₀ control; lanes 2 and 4: expressions of IL6+ and APA-1, respectively. The arrows indicate the positions of the protein bands corresponding to IL6+ and APA-1.

2.3.1.2 Optimisation of the Rosetta (λ DE3)/pRARE S30 extract volume in the cell-free reaction

To test the viability of the Rosetta (λ DE3)/pRARE S30 extract and to optimise the required S30 extract concentration for maximal APA-1 production, titration of the extract was carried out at the analytical scale. Reaction mixtures containing 18, 24, 30 and 36% (v/v) of the extract per ml of the cell-free reaction were prepared and used for unlabelled APA-1 expression. Expression of the IL6+ mutant using 24% (v/v) of the extract was performed as the control. The SDS-PAGE result (Figure 2.12) demonstrated that 18% (v/v) of the extract was the minimal amount required to obtain an optimum APA-1 yield. Concentrations of S30 extract below 18% (v/v) gave reduced expression levels (data not shown).

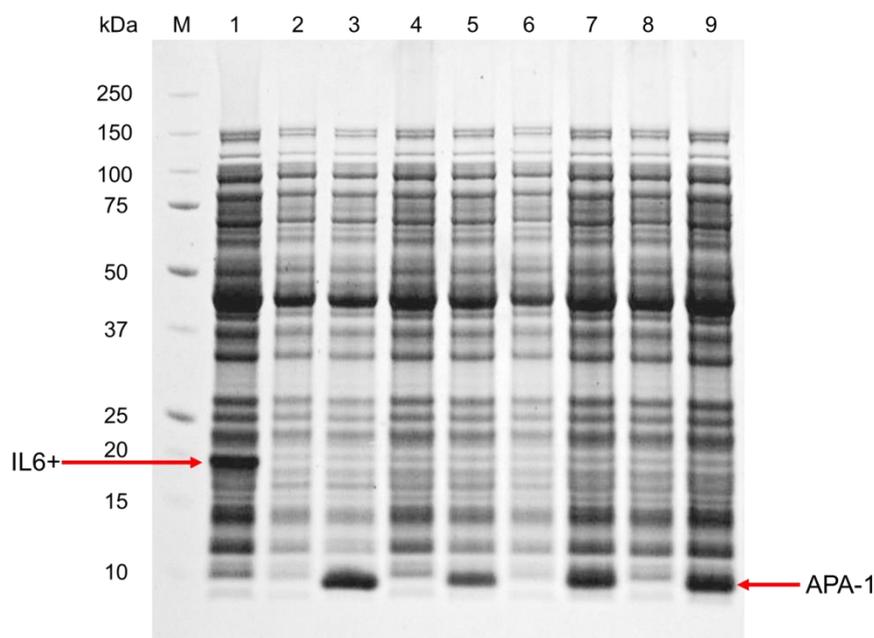
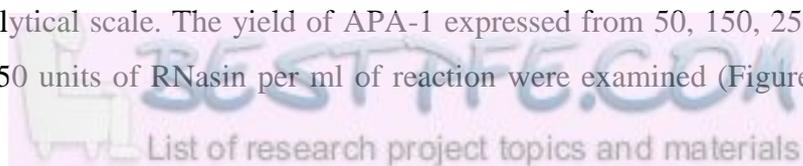


Figure 2.12: 4–12% SDS-PAGE analysis of the Rosetta (λ DE3)/pRARE S30 extract titration. Lane M: Precision plus MW marker (BioRad, Hercules, USA); lane 1: expression of the IL6+ mutant using 24% (v/v) of the S30 extract; lanes 2, 4, 6 and 8: T₀ control; lanes 3, 5, 7 and 9: expressions of APA-1 using 18, 24, 30 and 36% (v/v) of the S30 extract, respectively. The arrows indicate the positions of the protein bands corresponding to IL6+ and APA-1. APA-1 expression levels under the different conditions were assessed by comparing APA-1 and background expression levels.

2.3.1.3 Optimisation of the RNase inhibitor concentration in the cell-free reaction

Titration of the RNasin (F. Hoffmann-La Roche Ltd., Basel, Switzerland) concentration was performed at the analytical scale. The yield of APA-1 expressed from 50, 150, 250, 300, 350, 400, 450, 500 and 550 units of RNasin per ml of reaction were examined (Figure 2.13). The



result showed that the RNasin with a concentration of 250 U/ml or above gave similar APA-1 expression levels, thus 250 U/ml was used for future experiments.

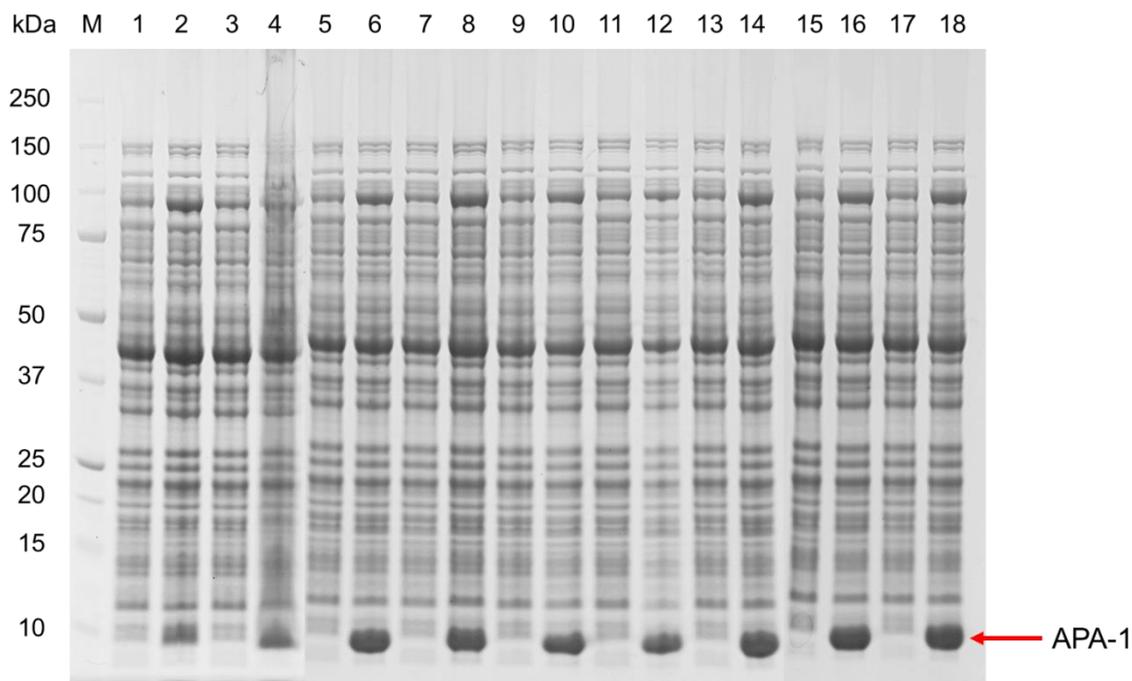


Figure 2.13: 4–12% SDS-PAGE analysis of the effect of RNasin concentration on cell-free expression yield. Lane M: Precision plus MW marker (BioRad, Hercules, USA); lanes with odd numbers are T₀ controls; lanes 2–18 (even numbers) are expressions of APA-1 using 50, 150, 250, 300, 350, 400, 450, 500 and 550 U RNasin/ml of reaction, respectively. The arrow indicates the position of the protein bands corresponding to APA-1. Sample volumes loaded onto the gel were adjusted according to the final reaction volumes following overnight incubation. APA-1 overexpression under the different conditions was further assessed by comparing the levels of APA-1 produced against the background proteins.

2.3.1.4 Preparation and optimisation of the T7 RNAP concentration in the cell-free reaction

Expression of T7 RNAP was induced by heat shock at 42 °C for 30 min followed by 2 h of inoculation at 40 °C (Figure 2.14). A 2 L culture yielded ~12.6 g of wet cell-pellet mass.

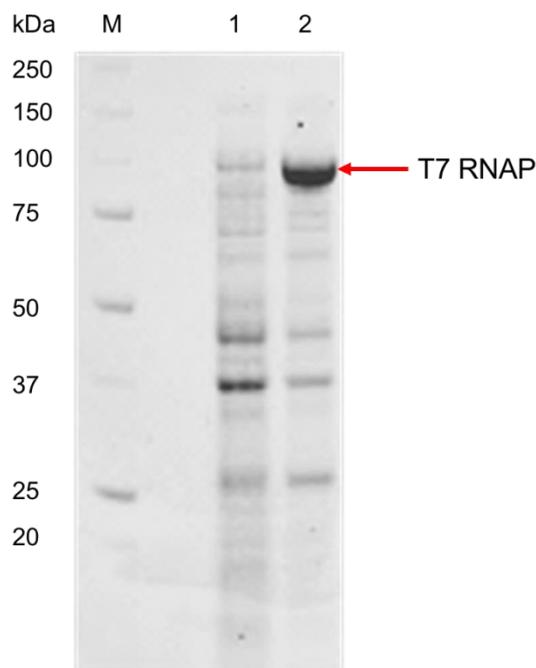


Figure 2.14: 4–12% SDS-PAGE analysis of the overexpression of T7 RNAP. Lane M: Precision plus MW marker (BioRad, Hercules, USA); lane 1: before heat shock induction; lane 2: 2 h after heat shock induction. The arrow indicates the position of the protein band corresponding to T7 RNAP.

A high-trap DEAE-ion exchange column was used to purify T7 RNAP and the protein was eluted at ~14% NaCl in TGED buffer (Figure 2.15). The void volume and collected fractions were analysed using 4–12% SDS-PAGE (Figure 2.16). The gel result showed that the T7 RNAP was also present in the void volume, indicating the High-Trap DEAE-ion exchange column reached its capacity and was over-loaded. The void fractions were collected and re-purified through the column. All the fractions containing T7 RNAP were pooled together (lanes 2–11 in Figure 2.16).

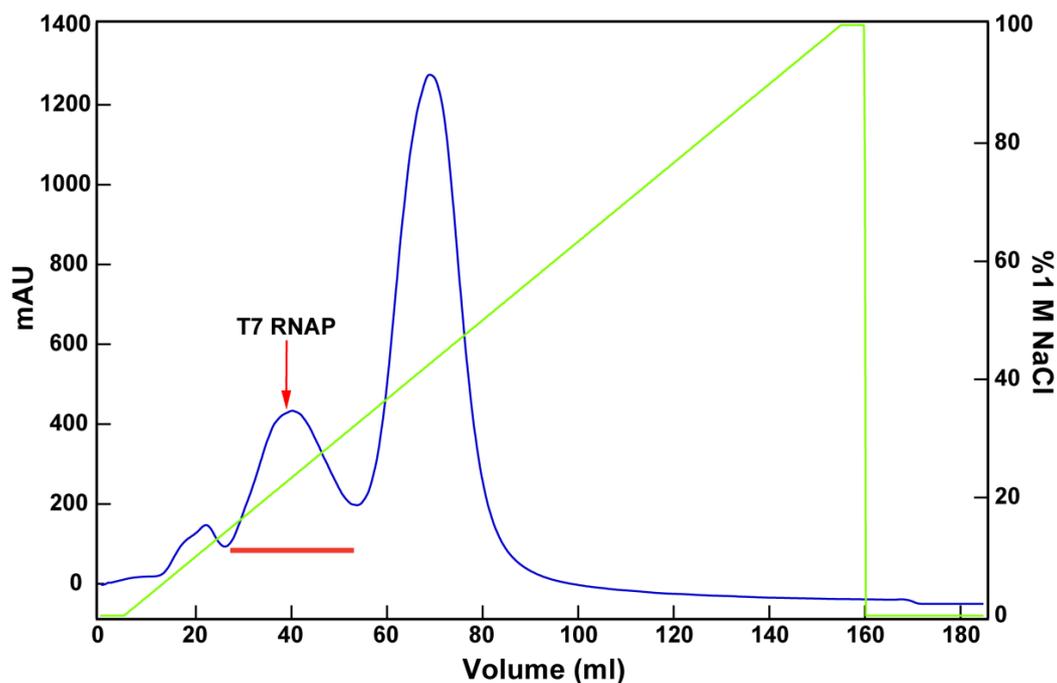


Figure 2.15: Chromatogram of T7 RNAP purification by DEAE-ion exchange chromatography. T7 RNAP eluted over an elution volume range of 25–54 ml, as indicated by the red bar. Blue: UV detection at 280 nm; green: 0–100% gradient of 1 M NaCl in TGED buffer.

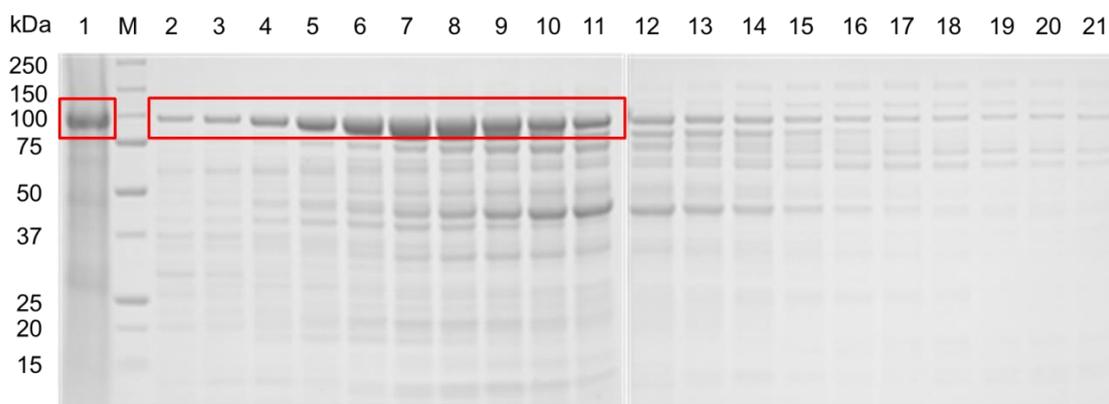


Figure 2.16: 4–12% SDS-PAGE analysis of DEAE-ion exchange chromatography purification of T7 RNAP. Lane 1: void volume; lane M: Precision plus MW marker (BioRad, Hercules, USA); lanes 2–21: fractions collected from the DEAE purification. Fractions corresponding to lanes 2–11 containing T7 RNAP were pooled. The red boxes indicate the positions of the protein bands corresponding to T7 RNAP.

T7 RNAP was then purified through a cellulose phosphate affinity column and the protein was eluted at ~45% (v/v) KCl in buffer P (Figure 2.17). The collected fractions were analysed using 4–12% SDS-PAGE (Figure 2.18). Fractions containing T7 RNAP were pooled and

concentrated. Based on the A_{280} absorbance, it was estimated that ~60 mg of the T7 RNAP was obtained from the 2 L culture.

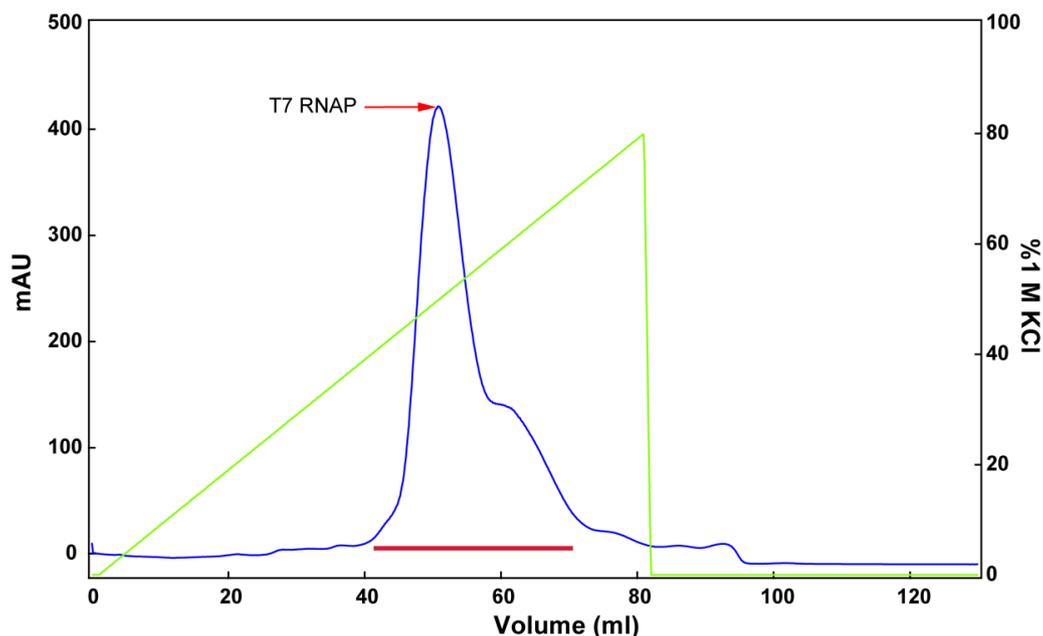


Figure 2.17: Chromatogram of the T7 RNAP purification using the cellulose phosphate affinity chromatography. T7 RNAP was eluted over an elution volume range of 42–72 ml as represented by the red bar. Blue: UV detection at 280 nm; green: 0–80% gradient of 1 M KCl in buffer P.



Figure 2.18: 4–12% SDS-PAGE analysis of the cellulose phosphate affinity chromatography of T7 RNAP. Lanes 1–3 and 4–14: fractions collected from the purification; lane M: Precision plus MW marker (BioRad, Hercules, USA). Fractions corresponding to lanes 2–13 containing T7 RNAP were pooled. The red boxes indicate the positions of the protein bands corresponding to T7 RNAP.

The viability of the purified T7 RNAP was examined using analytical scale cell-free reactions and the T7 RNAP concentration for maximal APA-1 production was optimised. Expression of APA-1 using 0.7, 1.05, 1.4, 1.75, 2.1, 2.8 and 4.2 μg of T7 RNAP per ml of cell-free reaction were performed. The SDS-PAGE result indicated that amounts $\geq 1.75 \mu\text{g/ml}$ of the T7 RNAP gave the highest levels of APA-1 productivity (Figure 2.19). Therefore, 1.75 $\mu\text{g/ml}$ of T7 RNAP was used in future reactions.

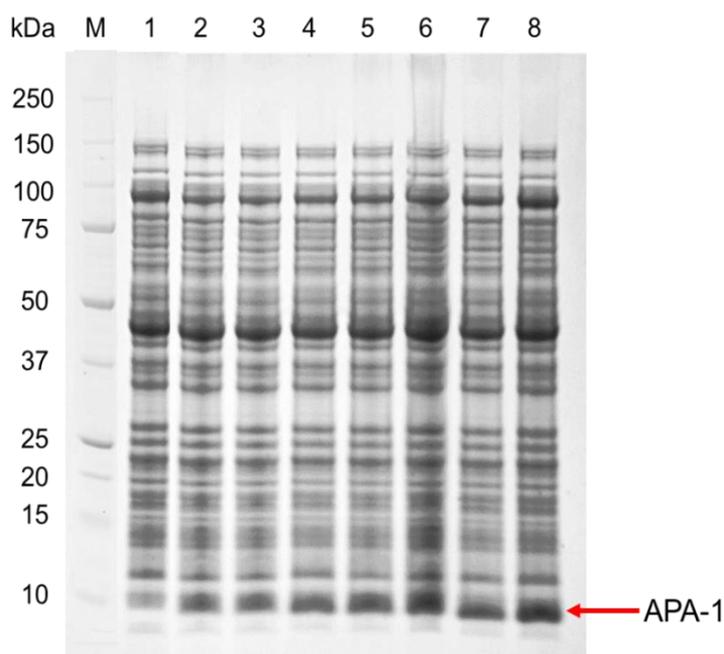


Figure 2.19: 4–12% SDS-PAGE analysis of the effect of T7 RNAP concentration on cell-free expression yield. Lane M: Precision plus MW marker (BioRad, Hercules, USA); lane 1: T_0 control; lanes 2–8: expression of APA-1 using 0.7, 1.05, 1.4, 1.75, 2.1, 2.8 and 4.2 μg of T7 RNAP/ml of reaction. The arrow indicates the position of the protein bands corresponding to APA-1. Sample volumes loaded onto the gel were adjusted according to the final reaction volumes after the overnight incubation and APA-1 expression levels were compared.

2.3.1.5 Crude unlabelled amino acid titration

Crude unlabelled amino acid mixtures with concentrations of 0.03, 0.015, 0.0075 and 0.003 g/ml were prepared to identify the minimal amount required to achieve maximal APA-1 expression. The SDS-PAGE results showed that the highest APA-1 expression was achieved at 0.03 g/ml (Figure 2.20). Expression levels of APA-1 at the amino acid concentration of 0.3 g/ml were further tested and the yield was similar to the yield observed when using 0.03 g/ml (data not shown). Therefore, 0.03 g/ml was used for expression of uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched APA-1 by replacing the unlabelled amino acids with a crude $^{13}\text{C}/^{15}\text{N}$ -labelled amino acid mixture.

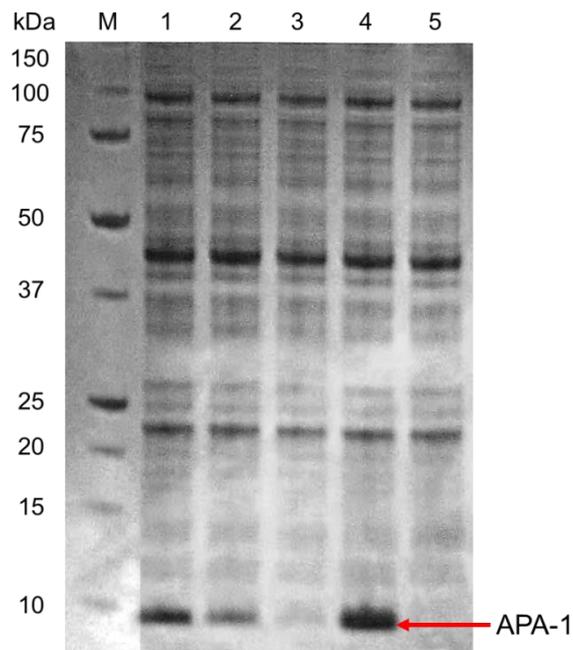


Figure 2.20: 4–12% SDS-PAGE analysis of the effect of crude amino acid concentration for unlabelled APA-1 cell-free expression. Lane M: Precision plus MW marker (BioRad, Hercules, USA); lanes 1–5: expressions of APA-1 using 0.015, 0.0075, 0.003 and 0.03 g/ml of amino acids, respectively; lane 5: T₀ control. The arrow indicates the position of the protein bands corresponding to APA-1. Sample volumes loaded onto the gel were adjusted according to the final reaction volumes after the overnight incubation.

2.3.1.6 Unlabelled versus ¹⁵N-labelled protein expression

Analytical scale cell-free reactions were performed to examine the productivity of APA-1 using unlabelled and isotope labelled amino acids. In the experiments, pure unlabelled and ¹⁵N-labelled amino acids were used. The reactions were performed using the previously optimised conditions (*i.e.*, 18% of the S30 extract, 250 U/ml of RNasin and 1.75 µg/ml of the T7 RNAP). The SDS-PAGE results showed that the expression of APA-1 using the two types of amino acid mixtures was similar (Figure 2.21).

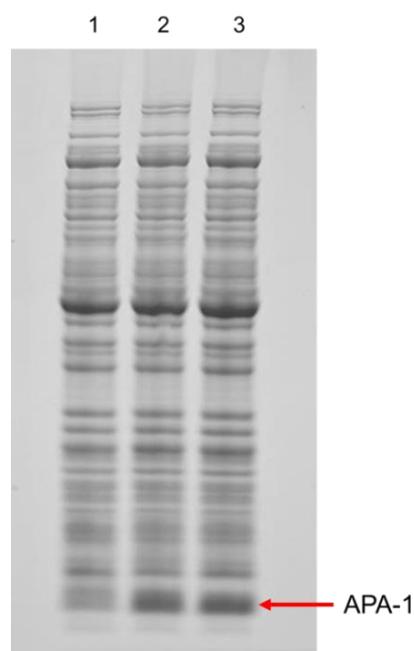


Figure 2.21: 4–12% SDS-PAGE analysis of the expression of APA-1 using pure unlabelled and ¹⁵N-labelled amino acids. Lane 1: T₀ control; lanes 2–3: expression of APA-1 using unlabelled and ¹⁵N-labelled amino acids, respectively. The arrow indicates the protein bands corresponding to APA-1. Sample volumes loaded onto the gel were adjusted according to the final reaction volumes after the overnight incubation.

2.3.2 Preparative scale cell-free protein synthesis

For large scale APA-1 production, the cell-free reactions were modified from the analytical to preparative scales. 1/10, 2/20 and 3/30 ml of the reaction mixture/feeding mixture set-ups were prepared and expression levels of APA-1 were compared with the analytical scale reaction. The results showed that the expression levels of APA-1 using the preparative scales were lower than the analytical scale expressions (Figure 2.22). When using the preparative scale reactions, precipitation/crystallisation was observed on the surface of the dialysis membrane after an overnight reaction. This may be due to the accumulation of by-products produced from the reaction or precipitation of salts on the dialysis membrane, which may block the exchange between the reaction and feeding mixtures thereby hampering target protein expression. The 3/30 ml set-up was observed to have the highest amount of precipitation/crystallisation leading to the lowest APA-1 yield, whereas the 1/10 ml reaction was observed to have the lowest precipitation/crystallisation and resulted in better APA-1 expression. This suggests that protein production is directly proportional to the reaction surface area on the membrane, blocking the membranes decreases the exchange area between the reaction and feeding mixtures, thereby inhibiting protein expression. Therefore, the 1/10 ml set-up was used for the preparative scale

cell-free reactions. Unlabelled, uniformly ^{15}N -, uniformly $^{13}\text{C}/^{15}\text{N}$ -, combinatorial selectively ^{15}N -, and selectively (Glu, Asp, Lys, and His) $^{13}\text{C}/^{15}\text{N}$ -enriched recombinant APA-1 samples were produced using the 1/10 ml preparative scale.

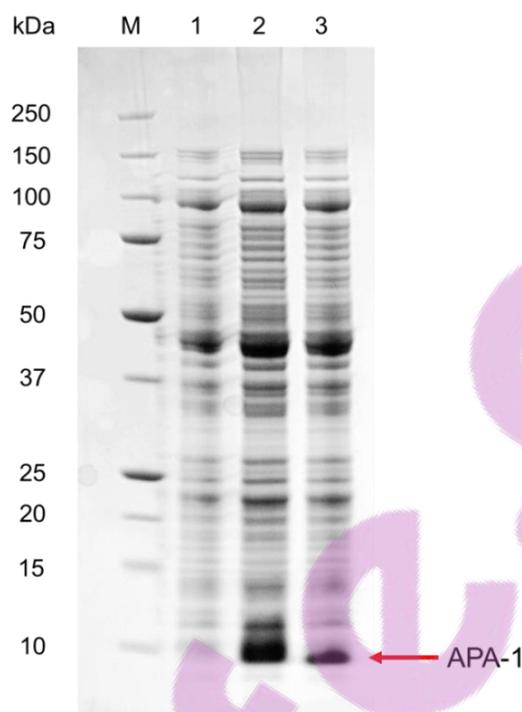


Figure 2.22: 4–12% SDS-PAGE analysis of the expression of APA-1 by the analytical and preparative scale systems. Lane M: Precision plus MW marker (BioRad, Hercules, USA); lane 1: T_0 control; lanes 2 and 3: expression of APA-1 from 50 μl analytical and 1 ml preparative scale reactions, respectively. The arrow indicates the position of the protein bands corresponding to APA-1. Sample volumes loaded onto the gel were adjusted according to the final reaction volumes after overnight incubation.

2.3.3 Purification of *amoebapore A*

2.3.3.1 Immobilised metal affinity chromatography

The supernatant containing APA-1 from the preparative scale cell-free reactions were purified through the IMAC column. Recombinant APA-1 was eluted from the column after about three column volumes using an elution buffer containing 500 mM imidazole (Figure 2.23). The collected fractions were analysed by 4–12% SDS-PAGE (Figure 2.24) and the result showed that IMAC purification did not give pure APA-1 because other protein containments were also present in the eluted solution (not visible in Figure 2.24). The fractions containing APA-1 were pooled and dialysed extensively against 10 mM ammonium bicarbonate at pH 8.0 to remove imidazole which is known to inhibit the activity of Factor Xa [132] and to reduce protein precipitation and enable sample lyophilisation.

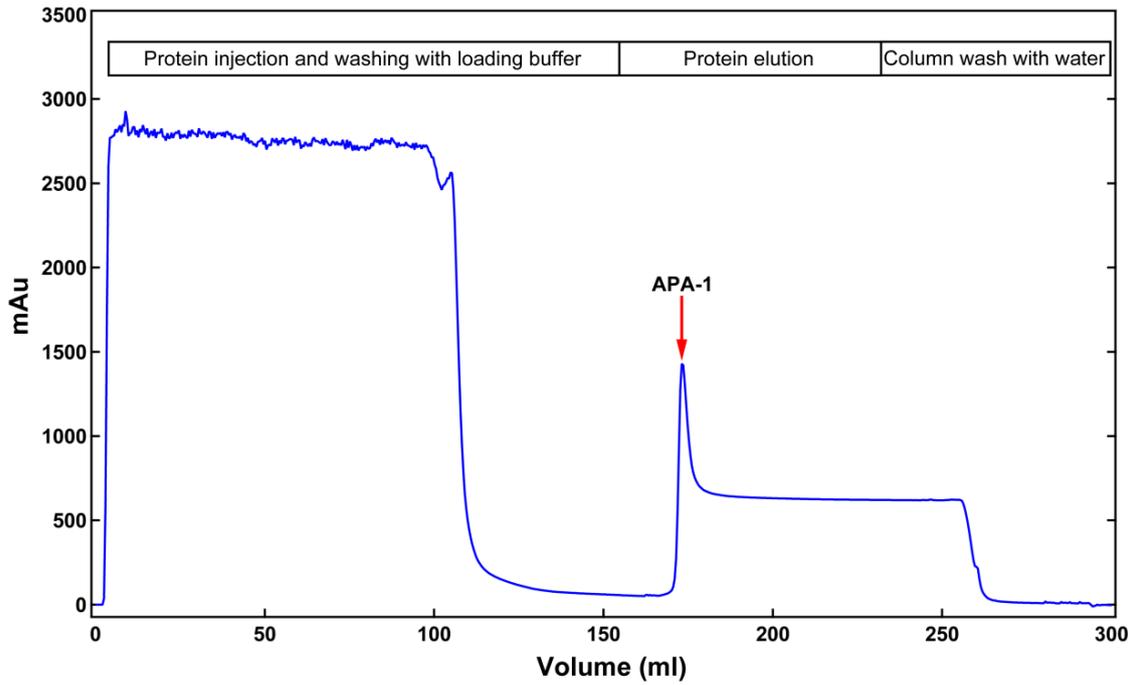


Figure 2.23: Chromatogram of the IMAC purification of APA-1. The red arrow indicates the peak that represents the elution of APA-1.

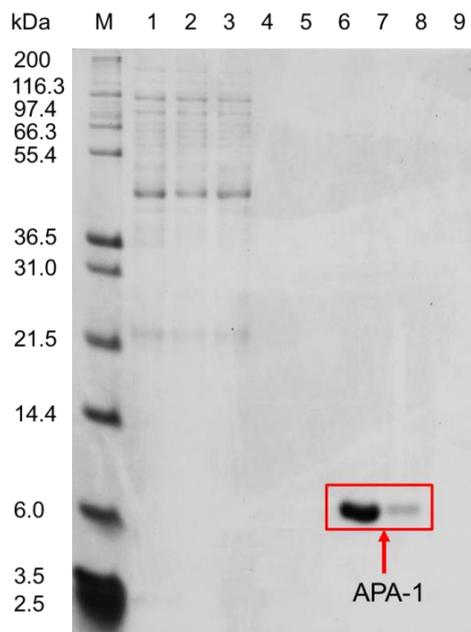


Figure 2.24: 4–12% SDS-PAGE analysis of fractions collected from the IMAC purification. Lane M: Mark 12™ standard protein marker (Invitrogen, Carlsbad, USA); lanes 1–5: void and washing fractions; lanes 6–9: elution fractions containing APA-1. The arrow indicates the position of the protein bands corresponding to APA-1.

2.3.3.2 Factor Xa digestion

Lyophilised APA-1 was diluted into the digestion buffer for (His)₆-tag cleavage. The efficiency of Factor Xa cleavage (F. Hoffmann-La Roche Ltd., Basel, Switzerland) of the fusion protein using different concentrations of CaCl₂ (2–6 mM), pH (6.5 versus 8.0), temperature (15–25°C), and digestion time was examined using 4–12% SDS-PAGE analysis (data not shown). The optimal conditions that gave complete cleavage of the N-terminal (His)₆-tag by Factor Xa were 25 °C, 2 mM CaCl₂ and pH 8.0 for 120 h at a protein:protease ratio of 1:50 (v/v). A lower protease:protein ratio (1:10) did reduce the reaction time; by only one day and was not a cost effective option. The required (His)₆-tag cleavage length by Factor Xa for APA-1 was found to be longer than the digestion periods of other proteins, which usually took between 16–72 h [133-139], although five days of digestion was previously observed [140]. In addition, it was observed that Factor Xa digested APA-1 in a two-step process. In the first 48 h it partially cleaved APA-1 and resulted in two protein bands on the gel that are close to each other. Following this period, full digestion of the protein was completed after 120 h (Figure 2.25). Although Factor Xa has a high degree of specificity, cleavage can occur at other non-recognition sites, usually following a basic amino acid in *E. coli* expression systems [141-143].

It was found that the Factor Xa protease purchased from Qiagen (QIAGEN, Limburg, Netherlands) had lower digestion efficiency than the protease purchased from Roche (F. Hoffmann-La Roche Ltd., Basel, Switzerland). More than two weeks was required for full APA-1 digestion when the protease from Qiagen was used (data now shown). The reason contributing to this observation is not clearly known; it may be due to slight differences of the protease purity or the sources that the protease was extracted from, for example, Factor Xa protease from Roche is extracted from bovine plasma whereas Factor Xa protease from Qiagen is not clarified.

The MW of APA-1 shown on the SDS-PAGE is smaller than its calculated MW. This can be explained by the “gel shift” phenomenon [144]. The phenomenon appears to be a common feature of membrane proteins which are mainly composed of helical structures. The differences in migration are primarily due to the altered protein-SDS binding and protein helical conformation.



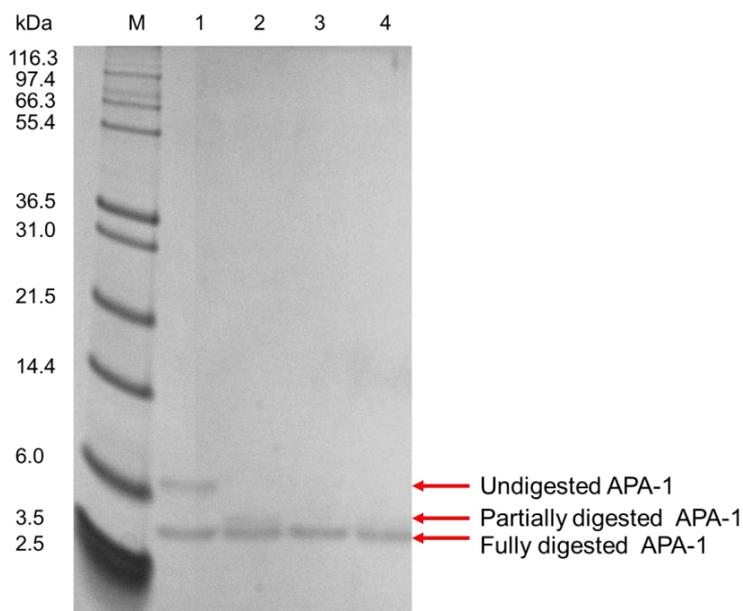


Figure 2.25: 4–12% SDS-PAGE analysis of APA-1 digestion. Lane M: Mark 12™ standard protein marker (Invitrogen, Carlsbad, USA): lanes 1–4: 48, 72, 96 and 120 h digestion, respectively. The arrow indicates the protein band corresponding to the undigested, partially digested and fully digested recombinant APA-1

To verify the location of the secondary cleavage site of Factor Xa on APA-1, the partially digested sample was collected and the two digested products were separated on a C8 RP-HPLC column. One protein species was eluted at 85% acetonitrile/0.1% TFA and the other species was eluted at 100% (Figure 2.26). The proteins were lyophilised and re-dissolved into the matrix solution and microTOP-Q MS analysis was performed. The MS profiles showed that the protein that eluted earlier had a MW of 8833.5 Da (*i.e.*, upper band on the SDS-PAGE), and the other protein had a MW of 8234.2 Da (Figure 2.27). The microTOP-Q MS instrument had a 4 Da constant shift smaller than the real mass. Therefore, the latter species is in agreement with the calculated mass (8238 Da) assuming the formation of the three disulfide bonds. The sequence of the (His)₆-tag is **MSGSHHHHHSSGIEGR**, where IEGR is the recognition site for Factor Xa digestion. APA-1 fused with SGIEGR has a calculated MW of 8838 Da assuming the formation of the three disulfide bonds, in agreement with the MW of the first eluted protein species. The result indicates that during the digestion of APA-1, the Factor Xa cleaves after the third Ser in the above sequence (in bold), leaving SGIEGR fused to APA-1. Subsequent digestion leads to the removal of the recognition site.

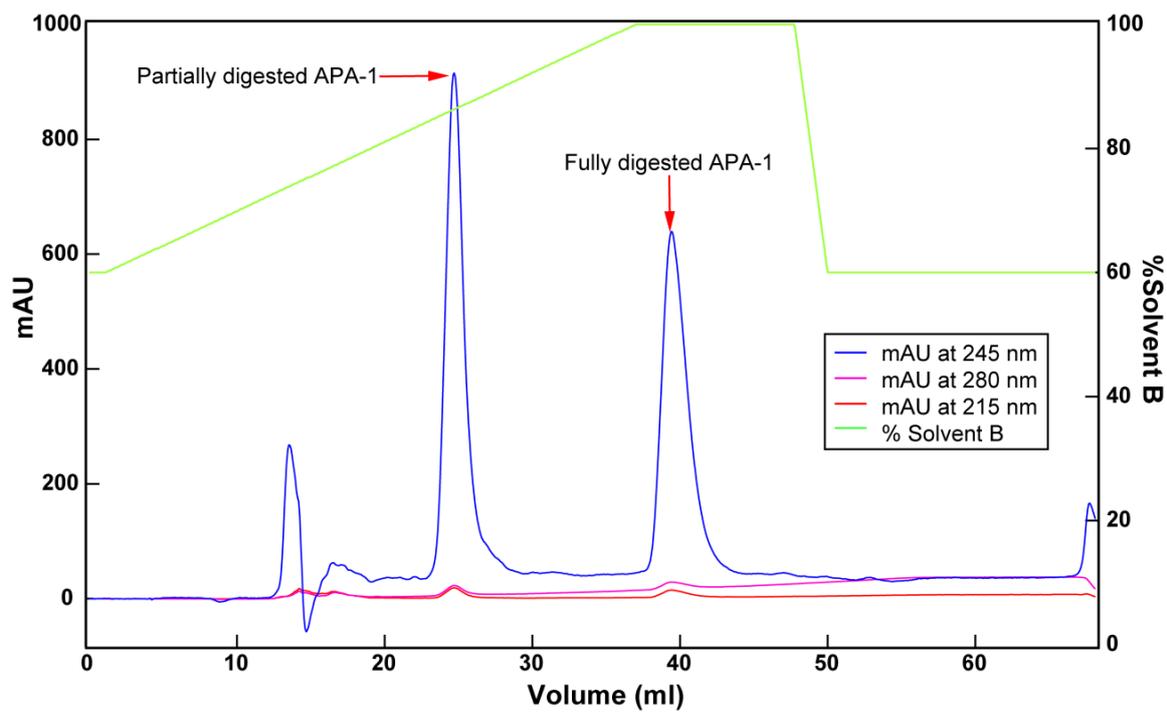


Figure 2.26: Chromatogram showing the purification of APA-1 using the C8 RP-HPLC column. The arrows indicate the peaks eluting at 85% and 100% (v/v) acetonitrile, and correspond to the partially and fully digested APA-1, respectively.

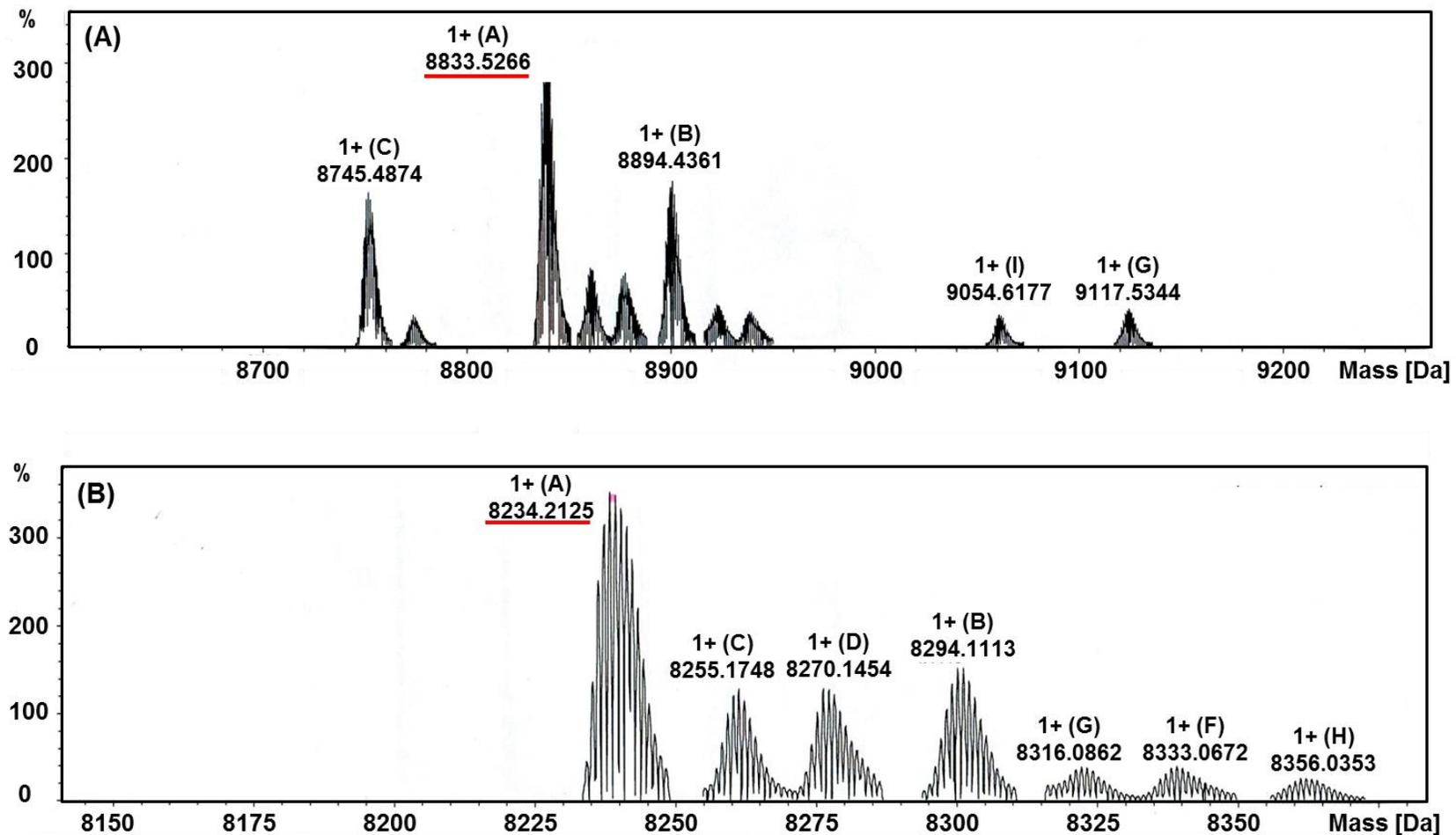


Figure 2.27: Mass spectrometry profiles of partially and fully digested APA-1 purified from RP-HPLC. **(A)** MS profile of the partially digested APA-1 with a MW of 8833.5 Da. **(B)** MS profile of the fully digested APA-1 with a MW of 8234.2 Da. The y-axis represents the relative abundance of protein present in solution, and the x-axis represents the protein mass. Minor species with variable masses are likely to arise from oxidation of residues [145].

2.3.3.3 Reverse phased-high performance liquid chromatography

During the RP-HPLC purification, recombinant APA-1 was eluted from the column in 100% (v/v) acetonitrile (Figure 2.28). The fractions corresponding to the peaks in the RP-HPLC chromatogram were collected and analysed by SDS-PAGE (Figure 2.29). The pooled fractions that contained APA-1 were lyophilised to remove the solvent and re-constituted in 0.01 M HCl. The SDS-PAGE result showed that the RP-HPLC procedure gave pure APA-1. The amount of APA-1 obtained from the cell-free reactions after purification was estimated using the calculated molar absorption coefficient. The uniform $^{13}\text{C}/^{15}\text{N}$ - and combinatorial ^{15}N -reactions 1–4 produced $\sim 150\ \mu\text{g}/\text{ml}$ of cell-free reaction. The combinatorial reaction five produced a significantly lower amount of protein ($\sim 50\ \mu\text{g}/\text{ml}$). This is probably because of the replacement of the major salt, potassium glutamate, with potassium acetate (required for the labelling of ^{15}N -glutamate).

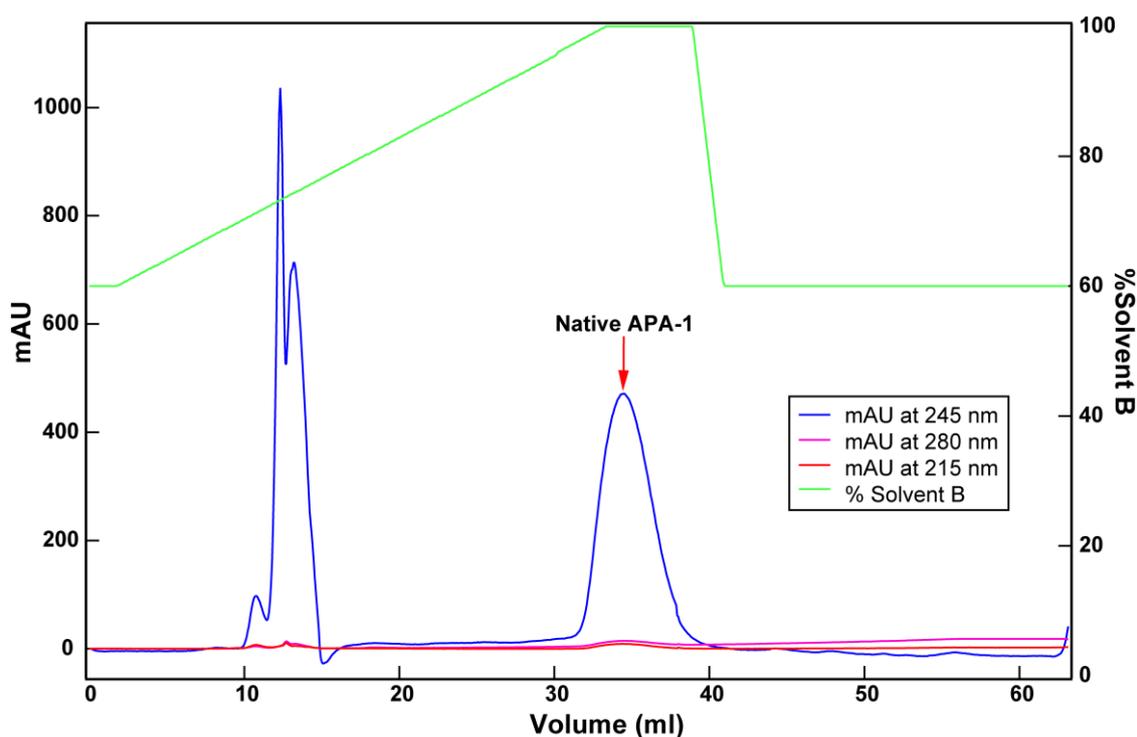


Figure 2.28: Chromatogram showing the C8 RP-HPLC column purification of APA-1. UV 280 nm (Blue), UV 245 nm (red), UV 215 nm (pink) and concentration of solvent B (green). APA-1 was eluted from the column at 100% (v/v) of solvent B. The arrow indicates the peak corresponding to the fully digested APA-1.

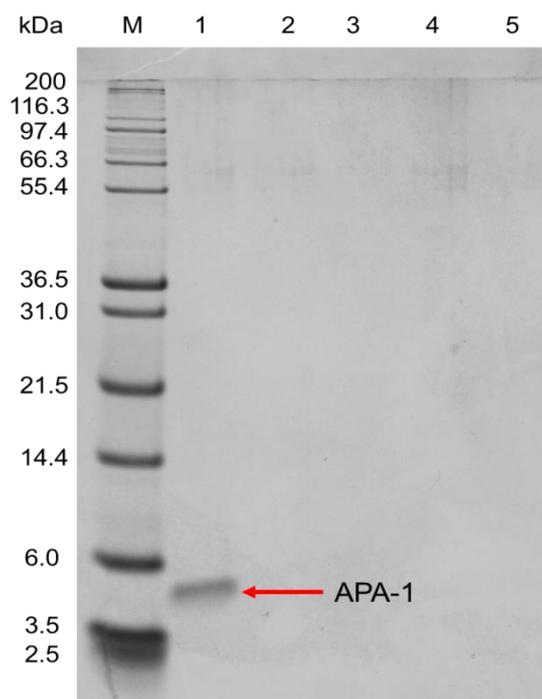


Figure 2.29: 4–12% SDS-PAGE analysis of fractions purified from the RP-HPLC. Lane M: mark 12TM standard protein marker; lane 1: fraction APA-1; lanes 2–3: the two peaks eluted before APA-1. The arrow indicates the position of the protein bands corresponding to APA-1.

2.3.4 NMR spectrometry

2.3.4.1 One-dimensional ¹H NMR experiments

One-dimensional ¹H NMR experiments of the uniformly ¹³C/¹⁵N- and the combinatorial ¹⁵N-enriched APA-1 were recorded at pH 3.0. Methyl proton resonances at ~0 ppm and good dispersion of the amide proton resonances at 7.5–10 ppm were observed, indicating that the recombinant APA-1 samples were folded. Based on the peak intensities between the 1D spectra recorded with different relaxation delay periods (Figure 2.30), the *T*₂ value was estimated to be 45.3 ms and the MW of the uniformly ¹³C/¹⁵N-enriched APA-1 sample at 50 μM was approximated to be 8.84 kDa at pH 3.0. This suggested that APA-1 existed predominantly as a monomer at this pH. The concentrations of the combinatorial ¹⁵N-enriched samples ranged between 50–80 μM, and the MW ranged between 7.95–9.32 kDa, also indicating the presence of primarily monomeric APA-1.

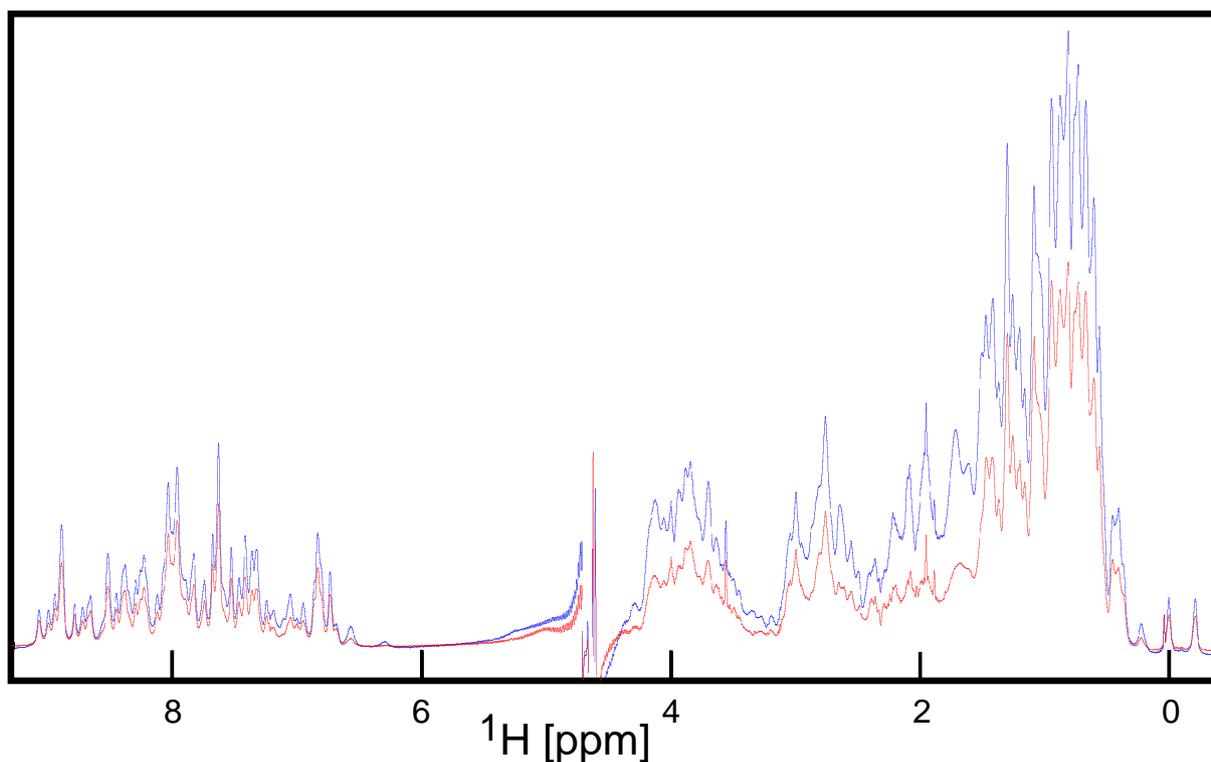


Figure 2.30: 1D ^1H spectrum of 48 μM APA-1 at pH 3.0. Blue: 250 μs delay; red: 2.9 ms delay.

2.3.4.2 Two-dimensional ^1H - ^{15}N HSQC assignment

As an initial approach, the sequence-specific assignment of the 2D ^1H - ^{15}N HSQC spectrum of APA-1 was achieved using the available amide proton chemical shift data in combination with the selective ^{15}N -labeling amino acid approach [84]. 2D ^1H - ^{15}N HSQC spectra were acquired for each of the combinatorial ^1H - ^{15}N -enriched APA-1 samples to identify the 19 different types of amino acids with backbone amide protons (Figure 2.31A). By using the combinatorial labelling approach, the number of observed resonances in each spectrum was minimised thereby limiting resonance overlap. The pattern of occurrence or absence of any particular resonance in the 2D ^1H - ^{15}N HSQC spectra recorded of the five samples enabled the identification of the amino acid type associated with this resonance.

The resonance intensities from the same amino acid type labelled in different samples were closely reproduced between the five 2D ^1H - ^{15}N HSQC spectra (Figure 2.31B). The resonance intensities of glutamate residues were noticeably weaker than resonances arising from other amino acids. This is due to the high concentration of potassium L-glutamate which is unsuitable for selective ^{15}N -labelling of glutamate residues in a protein. The presence of glutamate racemase in the S30 extract converts the L-glutamate to D-glutamate [88]. Reaction

mixtures with potassium glutarate instead of potassium glutamate can be used to give better performance [146]. However, potassium glutarate does not suppress conversions between different types of amino acids as well as the conventional potassium glutamate buffer [146]. Therefore, the glutarate buffer is best suited for ^{15}N -labelling of Glu only, whereas the use of glutamate in the reaction mixture is more advantageous for sample preparations that do not require isotope labelling of Glu.

Amino acid conversions were observed between Asp/Asn and Glu/Gln residues as a consequence of undesired side-reactions caused by the presence of metabolic enzymes in the S30 extract. The reactions decrease the efficiency of ^{15}N -labelling for particular amino acids and scramble ^{15}N - and ^{13}C -labelling [84, 88]. In the 2D ^1H - ^{15}N HSQC spectrum of combinatorial sample 3 where Asn residues were ^{15}N -labelled and Asp residues were not, resonances arising from Asp residues were observed, indicating that Asn was deaminated into Asp. The deamination was due to the hydrolysis of the side chain amide groups of Asn by asparaginase present in the cell extract [88, 105, 147] (Figure 2.32A). No evidence of amination from Asp to Asn was observed in the Asp-labelled samples, which contrasts what has been observed in other studies [92, 97, 146]. The deamination activity can be suppressed by performing a heat treatment of the S30 extract or adding an asparaginase inhibitor [88]. Moreover, amination from Glu to Gln was observed in the 2D ^1H - ^{15}N HSQC spectrum of sample 5. The conversion of Glu to Gln was due to the presence of glutamine synthase in the cell-extract with the help of ammonium ions and ATP (Figure 2.32B), and this enzymatic conversion has been observed in other cell-free reactions [84, 146-148]. The transaminase activity can be suppressed by adding glutamine synthase inhibitors [148]. No deamination of Gln to Glu was observed, which has been observed in cell-free reactions [97]. Evidence has shown that the amino acid conversions occur after the addition of the amino acids to the cell lysate and prior to protein synthesis [105, 146]. In the 2D ^1H - ^{15}N HSQC spectra of Sample 3 and Sample 5, the intensities of the unwanted Asp and Gln cross-peaks were slightly weaker than the resonances arising from other labelled residues, indicating minor transamination [88].

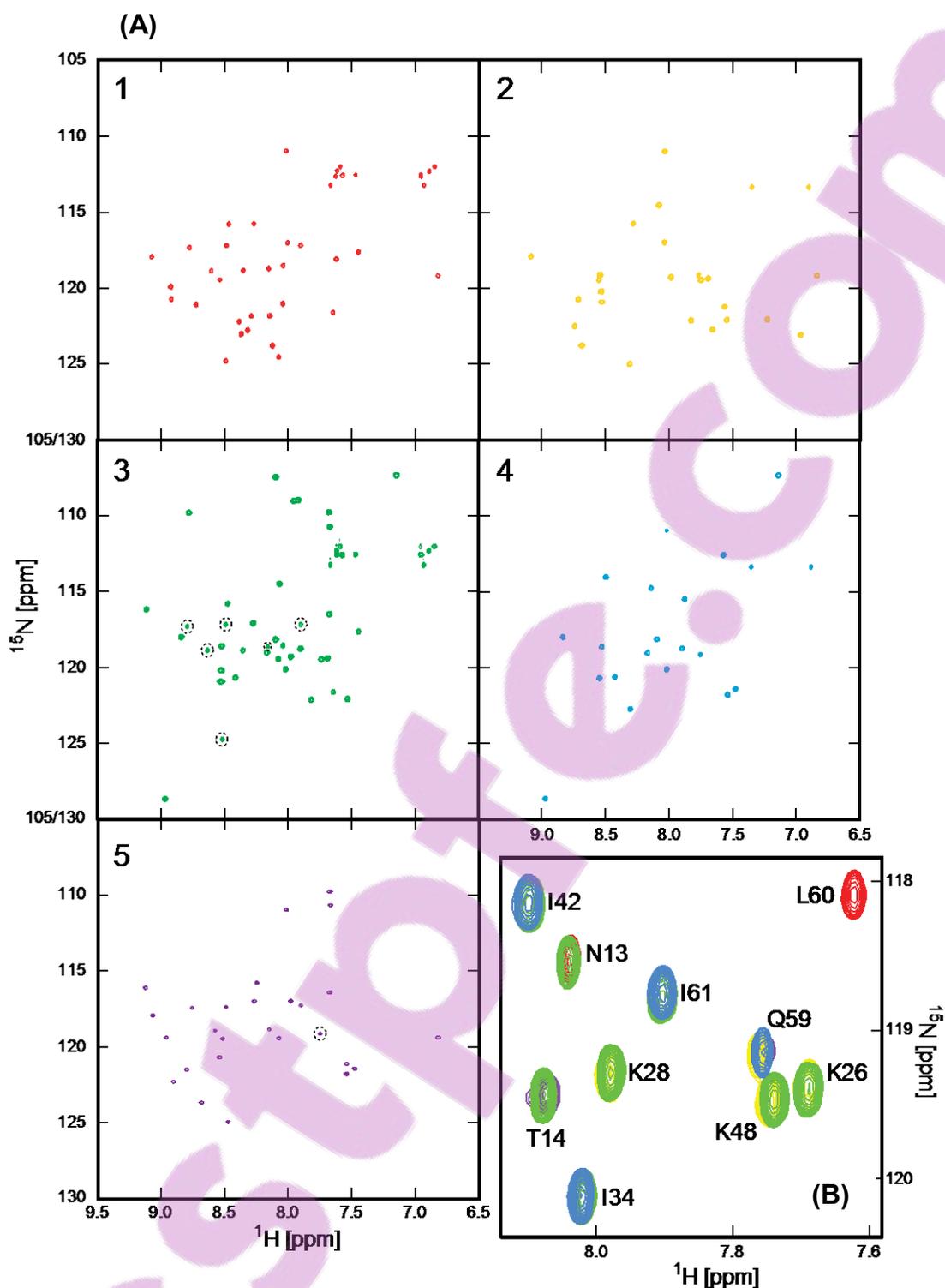


Figure 2.31: 2D ^1H - ^{15}N HSQC spectra of five combinatorially ^{15}N -labelled samples of APA-1. Numbers in the top left corner of (A) refer to the five different labeling patterns used. Sample 1: Leu, Arg, Asp, Asn, Tyr, His, Cys; Sample 2: Ala, Lys, Arg, Phe, Gln, Met, Cys, Trp; Sample 3: Gly, Ile, Lys, Thr, Asn, His, Trp; Sample 4: Ser, Val, Ile, Gln, Tyr, Met, His, Trp; Sample 5: Glu, Val, Thr, Asp, Phe, Met, Cys. The six additional resonances (circled) in Sample 3 are a result of deamination of Asn to Asp. The resonances circled in Sample 5 is the result of transamination of Glu to Gln (Q59) (B) Selected spectral region with all five spectra superimposed. The pattern of resonances occurrence in the different spectra identifies the amino acid type (labelled). The colouring scheme is the same as in (A) and follows the colouring scheme used in Figure 2.10. The appearance of a weak resonance for Q59 in sample 5 is because of transamination of Glu to Gln.

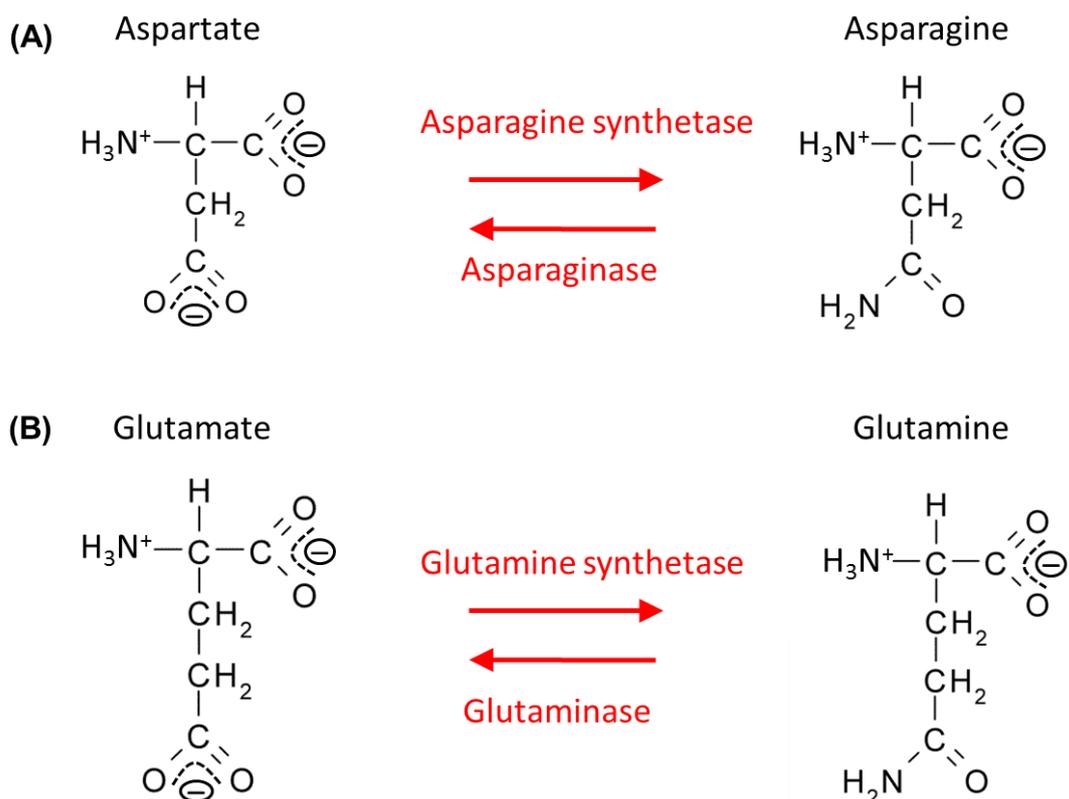


Figure 2.32: Transamination activity occurred in cell-free synthesis. (A) Amination of aspartate to form asparagine by asparagine synthetase, and deamination of asparagine to form aspartate by asparaginase. (B) Amination of glutamate to form glutamine by glutamine synthetase, and deamination of glutamine to form glutamate by glutaminase [149].

For small proteins or protein domains (*i.e.*, ~70 residues) where only ^1H chemical shift data are available, the combination of amide proton assignment information with the ^{15}N -combinatorial labelling approach represents an attractive starting point for rapidly obtaining resonance assignment of heteronuclear NMR data. By overlaying the five 2D ^1H - ^{15}N spectra combined with the ^1H amide proton chemical shift assignment information at pH 3.0, the residue specific assignments of resonances in the 2D spectrum were made. For example in Figure 2.31B, three resonances appeared in the overlaid spectra with only green and blue contours, and therefore can be assigned to the amino acid Ile according to the combinatorial labelling scheme presented in Figure 2.10. The ^1H chemical shift values of the resonances are 8.148, 8.119 and 7.913 ppm, and match the values to the available ^1H value of Ile 34, 42 and 61 (8.148, 8.119 and 7.913 ppm). Therefore, the approach enabled the assignment of the resonances to Ile 34, 42 and 61, respectively, without resorting to 3D NMR spectroscopy.

By using this procedure, 55 out of the 77 residues were unambiguously identified in the 2D ^1H - ^{15}N HSQC spectra without the requirement of conventional heteronuclear 3D experiments. The resonances that could not be unambiguously assigned are due to: (1) The amide proton chemical shift data arising from the structural study were missing for residues G1, L7, G10 and T44. (2) Resonances arising from Glu residues are weaker than other residues due to the presence of potassium L-glutamate which was unsuitable for ^{15}N -labelling of Glu residues, as explained previously. The signal of E2 is too weak that it is not visible in the spectrum obtained of sample 5. (3) Residues with the same amino acid type and very similar $^1\text{H}_\text{N}$ chemical shifts could not be unequivocally assigned, for example, L4 and L11 have $^1\text{H}_\text{N}$ chemical shifts of 8.062 and 8.060 ppm, respectively. The same situation was observed for the following pairs: L15/L50, L18/L34, T20/T21, K22/K37, L45/L57 and D51/D66. (4) As previously described, transaminase and deaminase activity influenced the intensity of resonances arising from aspartate and asparagine ^{15}N -labelled amino acids. As a consequence, particular aspartate and asparagine residues could not be distinguished due to similar ^1H chemical shifts, *i.e.*, N68/N13/D51.



2.3.4.3 Three-dimensional HNCA sequence specific assignments

To confirm the assignments of the resonances in the 2D ^1H - ^{15}N HSQC spectrum and obtain missing assignment information, a 3D HNCA spectrum was recorded using the uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched APA-1 sample at pH 3.0 and sequential assignments were achieved as described in Section 2.2.12.3 (an example is given in Figure 2.33). The intra-residue resonances were observed to be stronger than the inter-residue correlations [109].

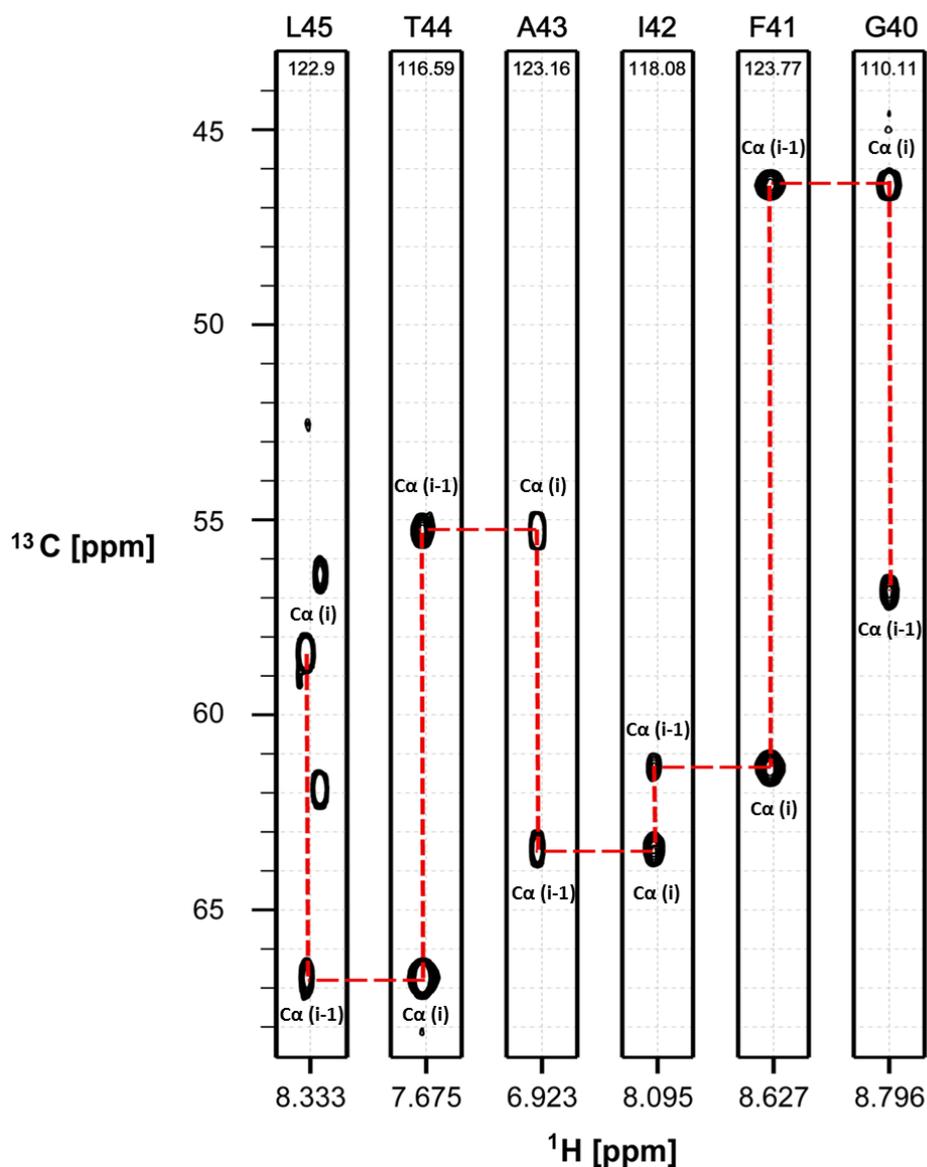


Figure 2.33: The sequential assignment process of assigning G40, F41, I42, A43, T44 and L45 using the 3D HNCA experiment. The x-axis at the bottom is the $^1\text{H}_\text{N}$ chemical shift, y-axis is the $^{13}\text{C}_\alpha$ chemical shift, and the ^{15}N chemical shifts for each residue are located at the top of each 2D strip extracted from the 3D spectrum. Sequential residues are identified by horizontal and vertical lines.

As previously shown [84], by combining the ^{15}N -combinatorial labelling strategy with a 3D HNCA experiment, sequence-specific assignments for all backbone ^1H , ^{15}N and $^{13}\text{C}_\alpha$ nuclei were achieved. The assignment information from the assigned HNCA spectrum was mapped onto the 2D ^1H - ^{15}N HSQC spectrum (Figure 2.34). The 3D HNCA approach confirmed the 2D ^1H - ^{15}N HSQC resonances assignments by matching the colour coding employed in the combinatorial labelling approach and the available ^1H chemical shift data. In addition, missing assignment information was completed using the 3D HNCA experiment, including the assignment of L7, G10 and T44 residues which were missing in the original ^1H chemical shift list (Appendix 3). The ^1H chemical shift values for N13 and T21 were corrected to be 8.022 and 7.665 ppm, rather than 8.506 and 8.332 ppm in the available ^1H chemical shift information. The chemical shifts for G1 are missing in the 3D HNCA spectrum because it is located at the flexible N-terminus of APA-1 and does not give rise to observable correlations. Assignments for E are also missing because of extremely weak resonance intensities, as described in Section 2.3.4.2.

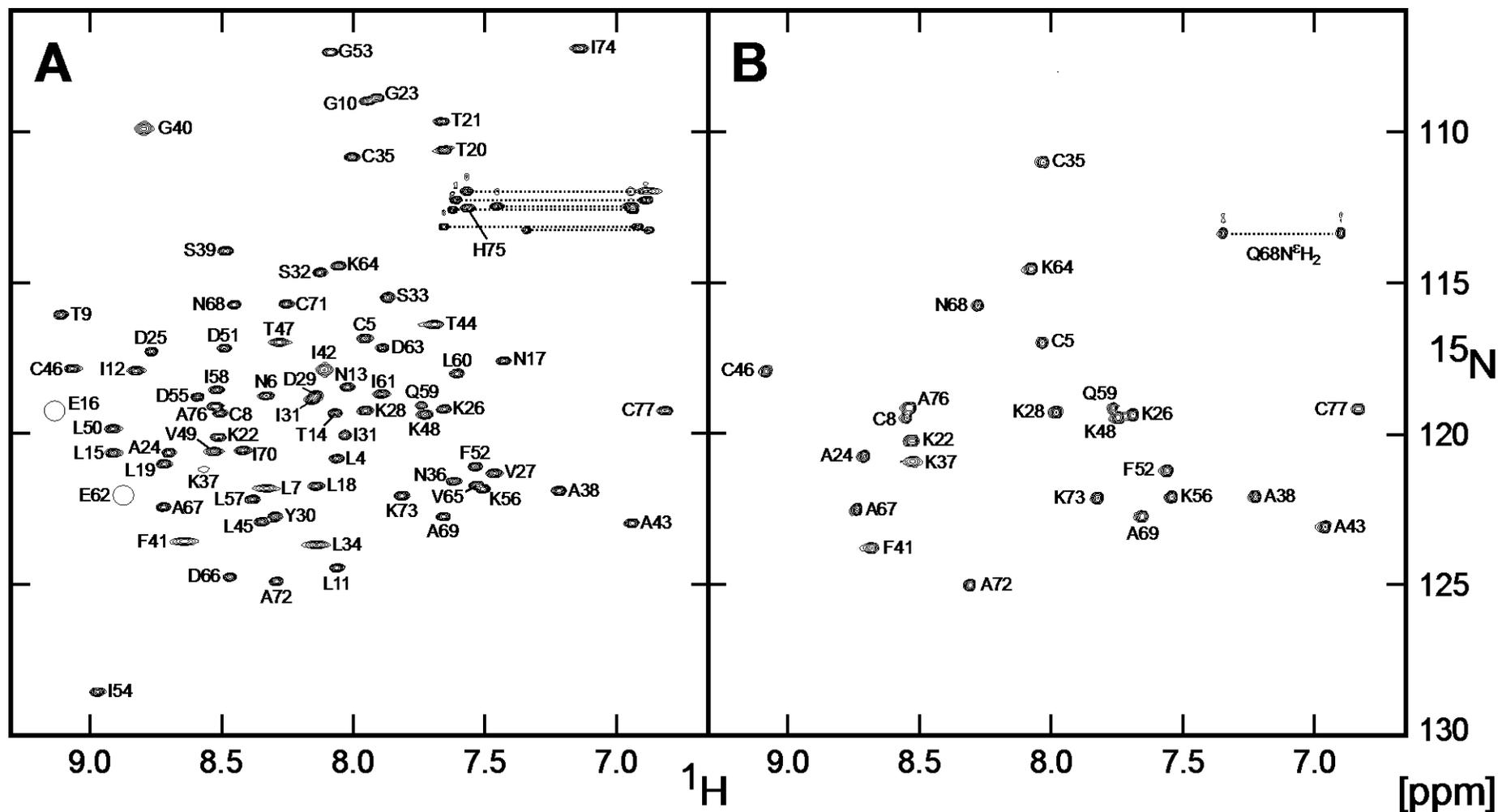


Figure 2.34: (A) Assigned 2D ^1H - ^{15}N HSQC spectrum of uniformly $^{15}\text{N}/^{13}\text{C}$ -enriched APA-1. The dotted lines connect amide protons corresponding to the side chains of glutamine and asparagine residues. Resonances arising from E16 and E62 (circles) were noticeably weak and were only visualised when the noise threshold was significantly reduced. (B) 2D ^1H - ^{15}N HSQC spectra for one of the five combinatorial samples prepared (*i.e.*, sample 2). Amino acids Ala, Lys, Arg, Phe, Gln, Met, Cys and Trp were labelled, thereby giving rise to resonances in the spectrum. Resonances are labelled with assignment information.

2.3.5 Pore-forming activity of amoebapore A

Before protease digestion to remove the (His)₆-tag, the five combinatorial samples were sent to our collaborators in Germany (Prof. Joachim Grötzinger, Department of Biochemistry, Christian-Albrechts-Universität zu Kiel, Medical Faculty, Kiel, Germany) to test the activity of the fusion APA-1. It was found that the fusion APA-1 with the (His)₆-tag showed lower activity when compared to the activity of native APA-1 using the liposome depolarisation assay. This suggested that the (His)₆-tag at N-terminus hampers dimer formation or membrane binding. After removal of the (His)₆-tag, the pore-forming activity of the recombinant APA-1 was comparable with native APA-1 (Figure 2.35).

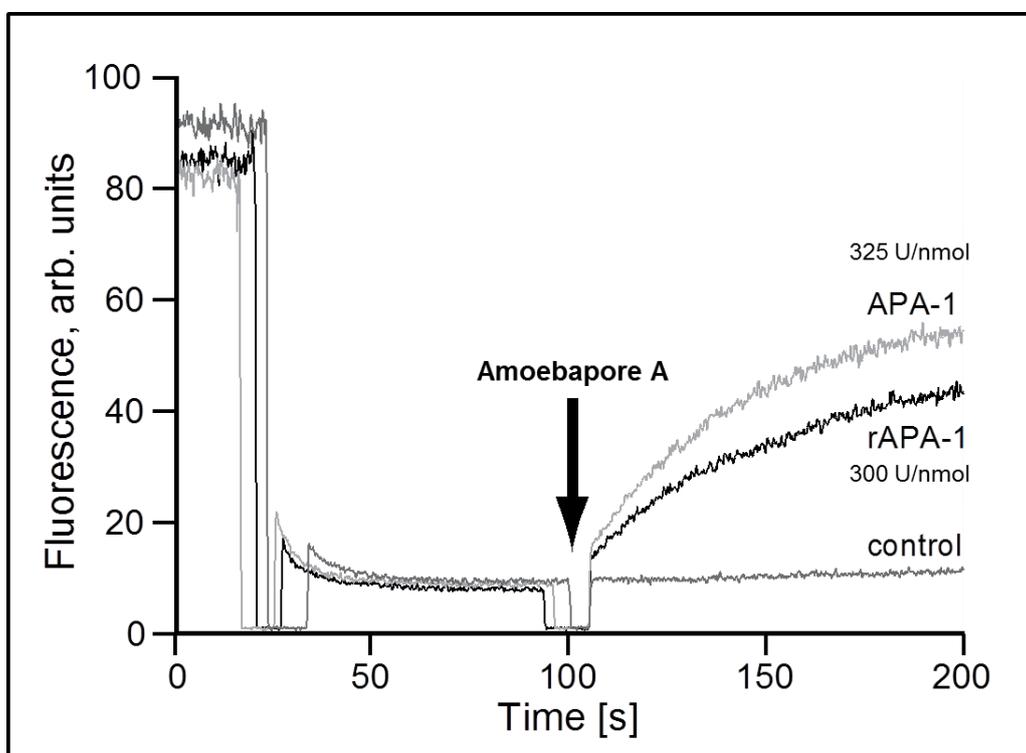


Figure 2.35: Time course of pore formation induced by native/recombinant APA-1. The dissipation of a valinomycin-induced diffusion potential in vesicles after the addition of APA-1 and the peptide solvent as the control were recorded. Pore-forming activity is reflected by the increase in fluorescence as a function of time.

2.4 Discussion

2.4.1 Optimisation and troubleshooting during cell-free synthesis of amoebapore A

Previous attempts to produce recombinant APA-1 using heterologous cell-based expression systems have failed, probably because of the cytotoxicity of APA-1. As cell-free protein synthesis does not depend on cell viability, it represents an obvious alternate route for the production of cytotoxic APA-1.

However, despite the many advantages of using cell-free protein synthesis, the high reagent costs coupled with the time-consuming preparation of the components represent limiting factors of the system. Therefore, it is important to achieve optimal levels of APA-1 expression while maintaining a cost effective and efficient cell-free synthesis system. For example, the RNase inhibitor is one of the most expensive components present in the cell-free reaction. Consequently, to reduce cost, the concentration of the commercially purchased RNase inhibitor was optimised so that minimal amounts were used without compromising APA-1 yields. As a result, the amount required for optimal APA-1 production was effectively reduced by 50% when compared with the standard protocol [123]. In addition, optimisation of the crude amino acid mixture concentration was performed to reduce the cost of using expensive ^{15}N - and $^{13}\text{C}/^{15}\text{N}$ - isotopically labelled amino acids for APA-1 production. Other high cost reagents present in cell-free synthesis include creatine phosphate, which works as an energy source, and nucleoside triphosphates (NTPs), which are essential for the protein translation process. A study has shown that creatine phosphate and NTPs can be replaced by cheaper reagents, glucose and nucleoside monophosphates (NMPs), while retaining a comparable protein expression level of chloramphenicol acetyl transferase [150]. However, it was observed that the reaction pH was not stable in this approach due to the accumulation of acidic side products [150]. Such instability would likely affect protein expression levels.

The concentration of freshly prepared *E. coli* S30 extract and T7 RNAP were also assessed to (1) test their viabilities, and (2) minimise the amount required to achieve comparable APA-1 yields, so as to avoid repeated S30 extract and T7 RNAP preparations. It is important to test occasionally the activity of the S30 extract to check that there is no deterioration in the activity. Optimisation of other cell-free ingredients can be performed, such as the concentrations of Mg^{2+} , K^{+} and DNA template plasmid, which may vary with different batches of S30 extract. In

addition, adjustment of the tRNA concentration and reaction temperature may also help maximise protein yields [92, 151].

Although optimisations tests were carried out, there were still issues with obtaining consistent levels of cell-free production of APA-1, and therefore troubleshooting experiments were performed. For example, it was found that APA-1 expression using HEPES purchased from British Drug House (BDH Merck Ltd., UK) gave reasonable APA-1 expression whereas HEPES purchased from other companies such as Sigma-Aldrich Co. (St. Louis, USA) and Fluka Analytical (Sigma-Aldrich Co., St. Louis, USA) gave noticeably lower or no APA-1 production. This could be due to slight variations in the purity of the HEPES reagent obtained from different companies, and that trace amounts of contaminants may have hampered the expression of APA-1.

It has been suggested that instead of using purified T7 RNAP, the pKO166 plasmid can be added to cell-free reactions for the co-synthesis of T7 RNAP using the transcriptional factors present in the *E. coli* S30 extract (*i.e.*, an auto-induction system) [88, 123]. The pKO166 plasmid is expressed non-competitively because transcription of this plasmid and the desired DNA depend on a different RNA polymerase during the cell-free reaction. However, it was found that using pKO1166 occasionally gave rise to a noticeable reduction of APA-1 expression. A possible reason for this observation is that overexpression of the 883-residue T7 RNAP may have partially exhausted the amino acid pool, thereby hampering optimal levels of APA-1 production. However, the reason why this only happened occasionally remains unresolved.

Furthermore, it was found that expression levels of APA-1 using the preparative scale reactions were generally lower than using the analytical scale reactions. This is likely due to the higher volume ratio of feeding mixture to reaction mixture in analytical scale reactions (*i.e.*, 19:1 for analytical scale and 10:1 for preparative scale). Therefore, the supply of essential precursors into the reaction mixture is more abundant. In addition, precipitation/crystallisation was observed on the surface of the dialysis membrane in preparative scale cell-free reactions as described in Section 2.3.2. This may be due to the accumulation of by-products or salt precipitation on the dialysis membrane that blocks dialysis thereby hampering target protein expression. Expression levels of APA-1 using three different reaction mixture/feeding mixture set-ups were compared (*i.e.*, 1/10, 2/20 and 3/30 ml) and the 1/10 ml set-up was found to give reasonable expression levels. Lastly, it was observed that expression of APA-1 gave

remarkably lower production when the reaction was performed in an Eppendorf thermomixer dry block heating shaker (Eppendorf, Hamburg, Germany) when compared with the use of a waterbath with an agitation device. This may be due to better temperature control and more thorough mixing of the reactions using the waterbath system.

All the issues discussed above were specific to APA-1 production, as other proteins (*e.g.*, IL6+) generally gave similar expression levels using the different approaches. No clear explanation for these observations can be given at this stage. As a result, the cell-free protein synthesis system has been used to produce approximately 150–500 µg of fusion-APA-1 per ml of cell-free reaction. Following protease digestion and purification 50–150 µg of APA-1 per ml of cell-free reaction was achieved. The decrease of the protein yield was mainly due to pH-facilitated aggregation of APA-1 during the protease digestion step (*i.e.*, pH 8.0) and removal of the (His)₆-tag. The final yields fall within the range of previously produced cytotoxic and membrane associated proteins. For example, the antimicrobial proteins human β-defensin-2 and crecopin, and membrane bound *E. coli* transporters have been synthesised at 1.2 mg/ml, 34 µg/ml and 100 µg/ml of reaction mixture, respectively [83, 85, 152]. APA-1 synthesised from the cell-free system was in the soluble form and exhibited similar pore-forming activity to the native protein following the proteolytic removal of the N-terminal (His)₆-tag.

2.4.2 Factor Xa digestion

Factor Xa protease was used for the removal of the (His)₆-tag from APA-1. The optimised digestion conditions showed that five days were required for complete digestion. This period is longer than the usual digestion time found for the digestion of other fusion proteins, which require between 16–72 h [133-139]. In addition, although Factor Xa has a high degree of specificity, cleavage can occur at other non-recognition sites, usually following a basic amino acid in *E. coli* expression systems [141-143]. It was observed that Factor Xa digested APA-1 in a two-step process. The initial cleavage was at a non-specific site that was after the third Ser residue, yet within the linker region between the (His)₆-tag and target protein. This initial digest removed the majority of the (His)₆-tag within 48 h, followed by subsequent digestion at the recognition site IEGR which resulted in complete removal of the N-terminal tag after a further 72 h incubation. Therefore, characterisation of the protein products after cleavage of the fusion protein was essential and carried out by mass spectrometry and SDS-PAGE (Figure 2.25 and 2.27). Inhibition of digestion of particular fusion proteins by Factor X has been shown previously because the protease cannot access the recognition sequence [153, 154]. Therefore,

one possible explanation for the long period required for complete APA-1 digestion may be due to the dimerisation of APA-1 restricting access of the protease to the recognition sequence; note that the dimer interface has been postulated to involve residues at the N-terminus and dimer formation is favoured under the pH used for the digest (see Chapter 3). Reducing the protein concentration and/or dropping the pH to 3.0 creates a predominantly monomeric population of APA-1 thereby exposing the protease recognition site; however, under these conditions Factor Xa is not active. To improve the digestion efficiency and specificity and reduce the high cost associated with purchasing the Factor Xa protease, a tobacco etch virus (TEV) protease site which recognises a cleavage point between Gln and Gly/Ser residues of a Glu-Asn-Leu-Tyr-Phe-Gln-(Gly/Ser) sequence and is a highly stringent protease may be used [155].

2.4.3 Transamination activities

In the cell-free system, transamination activities are essentially suppressed because metabolic enzymes are not regenerated, as observed in cell-based expression systems. However, undesired side-reactions for some amino acids can still occur and lead to the incorporation of unwanted isotope labelled amino acids into the target protein. Such side reactions create undesirable additional resonances in the NMR spectra or significantly reduce the intensities of resonances of residues for which assignment information or observation is desired. Conversions from Asn to Asp and Glu to Gln were observed, resulting in additional cross-peaks in the combinatorial 2D ^1H - ^{15}N HSQC spectra. To suppress the side-reactions, the S30 cell extract can be treated specifically prior to usage and inhibitors of asparaginase and glutamine synthase can be added to the cell-free reactions. However, to completely eliminate metabolic enzymes and avoid side-reactions, the cell-free system needs to be reconstituted with purified components, as in the 'PURE' system [88, 156]. Unfortunately, this method is uneconomical and unnecessary for the purposes of ^{15}N -labelling. Conversions of other amino acids can also occur, such as between Gly and Ser, Cys and Tyr, Ile and Leu, and Thr to Cys, Gly, Ser and Tyr [98]. Moreover, by-products can be formed from isotope scrambling of amino acids such as Arg, His, Lys, Met, Thr, Val, Ala, Cys, Gly Ser and Trp which lead to undesired additional resonances in the NMR spectra. The by-products can be easily removed by dialysis of the sample prior to NMR analysis [92]. However, these side reactions were not observed in APA-1 expression. The transamination activity did not affect the 2D ^1H - ^{15}N HSQC spectrum when a uniformly isotope labelled sample was required.

2.4.4 Assignment of the NMR resonances representing backbone nuclei

For rapid sequence-specific assignment of the resonances in the 2D ^1H - ^{15}N HSQC spectrum, a combinatorial selective ^{15}N -labeling strategy was used in combination with the available amide proton chemical shift information. By using this approach as an initial starting point for resonance assignment of the 2D ^1H - ^{15}N HSQC spectrum of APA-1, 55 out of 77 residues were assigned unambiguously. Although the combinatorial labelling approach does not provide sequential assignment information, it offers great advantages by minimising the number of resonances observed in each spectrum, thus removing spectral overlap. By combining the combinatorial labelling approach, amide proton chemical shifts and a 3D HNCA experiment, full sequential backbone assignments ($^1\text{H}_\text{N}$, ^{15}N and $^{13}\text{C}\alpha$) of APA-1 was obtained at pH 3.0. The assigned 2D ^1H - ^{15}N HSQC spectrum of APA-1 combined with different NMR experiments can be used to characterise dimer formation and potentially membrane binding to model membranes. The cell-free approach was also used to prepare site-specific mutants of APA-1 that may be important for APA-1 activity. The cell-free expression and confirmation of the folding state using NMR spectroscopy of the mutants will be described in Chapter 3.

3

Investigation of amoebapore A pH-dependent activity



3.1 Introduction

3.1.1 *Amoebapore A dimerisation*

The pore-forming activity of APA-1 has been proposed to occur via a pH-dependent dimerisation event, where the protein is inactive in the monomeric form and becomes active when it forms a dimer under slightly acidic conditions [24]. Although no structural description of the dimerisation interface is available, it has been postulated that the ionisation states of four residues are crucial for dimer formation: H75, A63, E2 and K64 [24]. Two salt bridges are formed by these four residues to give an anti-parallel head-to-head dimer. It has been postulated that dimers accumulate on the target membrane and aggregate to form pore structures when a threshold concentration of APA-1 is reached [24]. However, detailed structural information of the dimer and its functionality in APA-1 pore-forming activity remains unsolved. Moreover, it is unclear which ionisable side-chains are functionally important in dimer formation, membrane-binding and pore formation.

3.1.2 *NMR spectroscopy to study amoebapore A dimerisation*

3.1.2.1 Identification of residues at the dimer interface

To model the APA-1 dimer structure, residues located at the dimer interface must be identified. Each NMR-active nucleus has a characteristic chemical shift, reflecting its local chemical environment (Section 2.1.3.2). The transition of APA-1 from monomeric to dimeric states should lead to perturbations of the local chemical environments of nuclei located at the dimer interface, thereby leading to chemical shift changes of the corresponding resonances in the NMR spectra [157]. The 2D ^1H - ^{15}N HSQC spectrum has been frequently used for studying protein-protein interactions [158-162]. By modulating the solution conditions, *e.g.*, pH or protein concentration, to induce APA-1 dimerisation, chemical shift mapping coupled with the backbone assignments can be used to probe the self-association of APA-1, and thus identify the residues located at the dimer interface. The results will be used to guide the design of site-directed mutants that further probe key residues involved in APA-1 activity.

3.1.2.2 Determination of the pKa values of ionisable groups

Given the pH-dependent activity of APA-1, the ionisation states of the Asp, Glu, His and Lys (no Arg in the APA-1 sequence) residues are likely to play important roles in APA-1 function.

The intrinsic pKa values of these ionisable groups are 3.9, 4.3, 6.5 and 10.4 for Asp, Glu, His and Lys, respectively [43]. Nonetheless, in proteins these pKa values vary over a wide range due to the influence of local environments. Based on the pKa values of ionisable residues assessed in 78 different folded proteins, the values have been found to range between 0.5 and 9.2 for Asp, 2.1 and 8.8 for Glu, 2.4 and 9.2 for His, and 5.7 and 12.1 for Lys residues [48]. The pKa values are influenced by factors such as electrostatic interactions, hydrogen bonding, or the extent of solvent exposure in a protein structure [163-166]. Residues with unusual pKa values often reside at ligand binding or protein interaction sites, or are involved in enzymatic catalytic functions [167-169]. Therefore, the determination of pKa values of ionisable residues can provide insights into the function of these residues in proteins. Figure 3.1 presents the charged residues of APA-1.

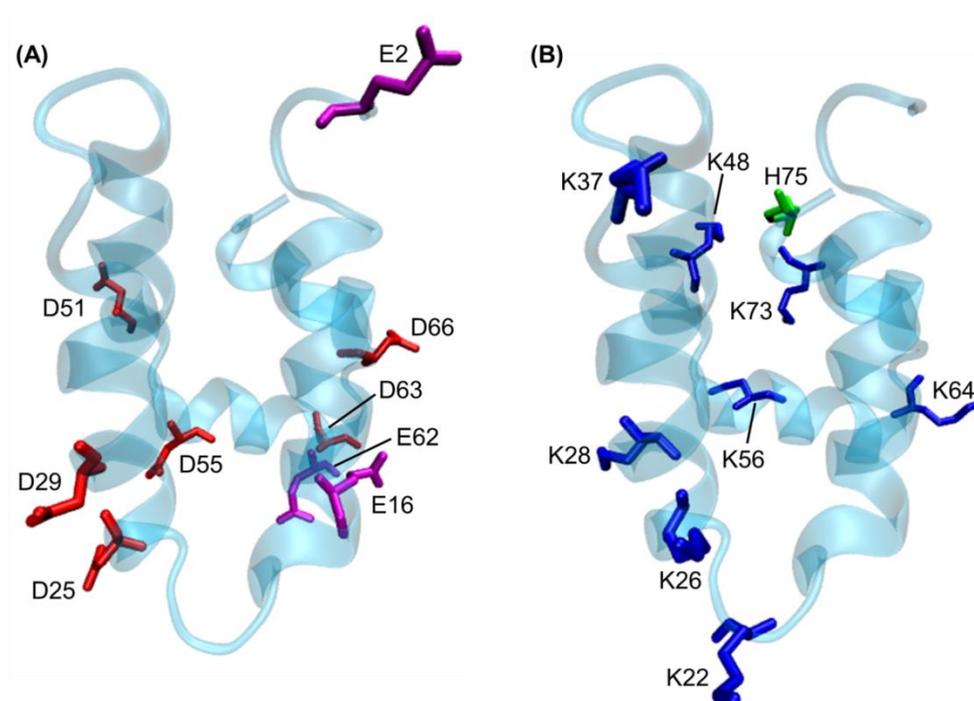


Figure 3.1: Ionisable residues of APA-1. (A) Acidic amino acids: the six Asp residues are presented in red and the three Glu residues are in magenta; (B) basic amino acids: the eight Lys are depicted in blue and the sole His is coloured green. Residue assignment information is also provided.

3.1.3 Other biophysical methods

Besides NMR spectroscopy, other methods such as size exclusion chromatography and small angle X-ray scattering can be used to study the APA-1 monomer-dimer equilibrium.

3.1.3.1 Size exclusion chromatography

Small-zone size exclusion chromatography (SEC) analysis can be used to provide the dimer-monomer equilibrium of APA-1 under different pH values and protein concentrations. However, dilution of the protein concentration (depending on the void volume of the column) and dissociation of the protein complex during passage through the column can influence the population of monomer and dimer species [170] and variable protein retention times [171, 172]. To overcome these problems large-zone SEC can be used. The method requires large protein volumes to saturate an analytical size-exclusion column (usually a few mL volume) and establish a plateau region in which no dilution of the sample occurs. Shifts in the boundary positions of the plateau region as a function of protein concentration can be analysed to examine quantitative information describing the protein self-association equilibrium, such as the fraction of monomer and dimer, and the equilibrium association constant [173, 174]. Unfortunately, because reasonably large sample volumes are required over a range of concentrations (μM to mM), relatively large quantities of protein are needed. Thus, this method was considered to be not suitable for characterising the monomer-dimer equilibrium of APA-1.

3.1.3.2 Small angle X-ray scattering

Small angle X-ray scattering (SAXS) is a low-resolution technique for characterising gross structural information of protein molecules and protein complexes in solution with sizes ranging from a few kDa to several thousand kDa [175]. The sample solution environment can be easily monitored, therefore making SAXS useful for mapping structural changes to a protein upon environmental perturbations [176, 177]. The basic principle of SAXS is that the incident X-ray beam collides with molecules in solution elastically, leading to the scattering of reflected waves in all directions. The reflected waves interfere with each other, resulting in constructive interference along certain angles and a unique scattering pattern. When small angles are applied, the scattering profiles provide information about the global structure and conformation (*e.g.*, size and shape) of the protein molecule [178]. Figure 3.2 shows a schematic diagram of a basic SAXS experiment.

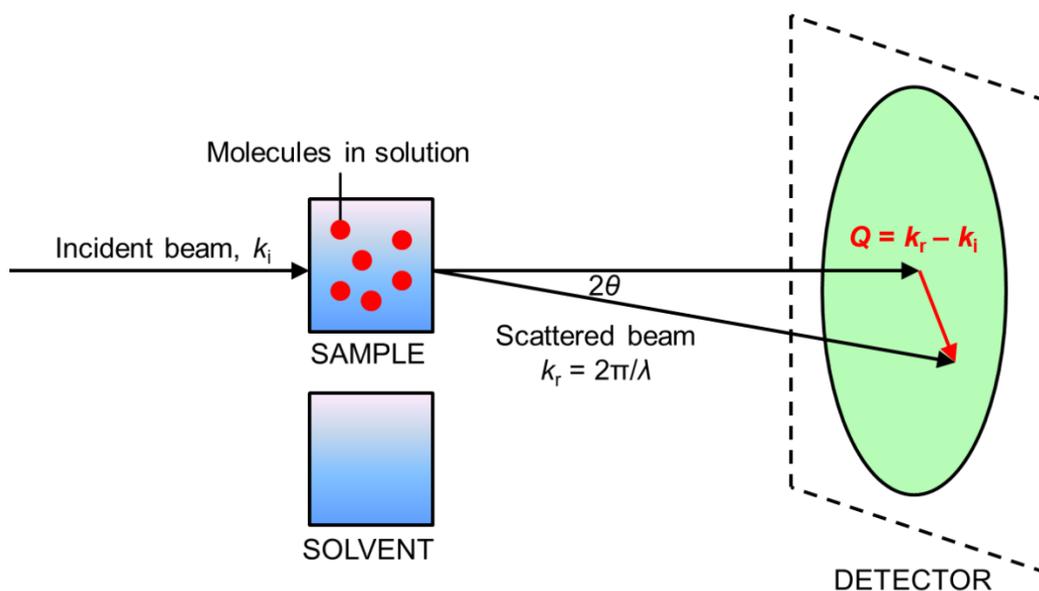


Figure 3.2: Schematic representation of a SAXS experiment. The diagram is made according to [179]. Incident beam (k_i) collides with molecules in solution, scattered with angle 2θ and recorded on the detector. The scattered beam (k_r) is equal to $2\pi/\lambda$. The momentum transfer (Q) is the vector difference between k_i and k_r . The background scattering profile is subtracted from the sample.

When X-rays with a known wavelength (λ) are scattered from a sample, the scattering vector or momentum transfer (Q), can be calculated using Eq. 3.1 [180]:

Eq. 3.1

$$Q = 4 \pi \sin\theta/\lambda$$

where 2θ is the angle between the incident and scattered beam. The scattered intensity, $I(Q)$, is recorded as a function of Q . Since the λ is fixed, $I(Q)$ is essentially dependent on the scattering angle θ . Scattering data from the solvent is subtracted from the scattering profile. SAXS experiments to obtain high-resolution structural information require highly pure, soluble and monodisperse protein samples with sufficient dilution so that no significant attractive or repulsive interactions exist between molecules. All protein molecules in a sample contribute to the SAXS signal, thereby the random positions and orientations of the molecules give rise to an isotropic intensity distribution which reflects the scattering from a single protein molecule averaged over all orientations [181, 182].

The scattering patterns generated from protein molecules in a sample are presented as averaged 1D curves. Structural parameters such as the MW, radius of gyration (R_g) and maximum intramolecular distance (D_{\max}) of the protein can be determined from the 1D curves, providing

a low-resolution overall shape of the protein [183]. The R_g , which provides information of mass distribution within a protein, can be estimated using the Guinier approximation [183]:

$$I(Q) \approx I(0) \exp(-Q^2 R_g^2/3)$$

Eq. 3.2

where $I(0)$ is the scattered intensity at zero angle which is proportional to the MW of the protein. By plotting $\ln(I(Q))$ versus Q^2 and fitting the slope and intercept, R_g can be obtained [183]. Small amounts of aggregation of the sample in solution can lead to an upturn in the Guinier plot and give rise to biased R_g and $I(0)$ values. The software **GNOM** [184] can be used to indirectly Fourier transform the scattering profile to obtain the interatomic distance distribution function, $P(r)$, of the scattering molecules [185]. The $P(r)$ profile can provide the shape and volume occupied by a protein or protein complex and is sensitive to the symmetry and domain structure within a molecule.

In addition, information obtained from a SAXS profile can be used to assess different protein models with defined geometry to estimate the rough shape and size of the protein. Protein models with simple shapes (*e.g.*, sphere, cylinders and ellipsoids) are usually employed as a first step in scattering pattern interpretation. The calculated and the experimental scattering patterns are compared and a close-to-native solution structure from different models of the protein can be achieved [186]. Moreover, the approximate 3D structure of a protein such as size and shape can be reconstructed from the 1D scattering data using *ab initio* analysis, assuming a homogenous solution and constant scattering density within the molecules. For example, programmes **DAMMIN** [187] and **DAMMIF** [188] are able to reconstruct the overall shape of a protein molecule by finite volume elements (*i.e.*, dummy atoms) to fit the experimental data. Both programmes use simulated annealing to construct compact and interconnected models. **GASBOR** [189] can be used to reconstruct proteins structures by a chain-like ensemble of dummy residues. These dummy atom/residue model approaches allow several models to be generated which may fit equally well to the experimental SAXS data. To reduce ambiguities, **DAMAVR** [190] is used to place the models into structurally similar classes and generate an average structure from each class. The generated *ab initio* model can be superimposed with the atomic structure of the target protein using **SUPCOMB** [191]. In addition, programmes **CRYSOL** [192] and **AXES** [193] are modelling tools that directly operate on known protein atomic models (entered into the protein database) to generate theoretical scattering profiles. The theoretical profiles are evaluated against the experimental profiles to validate the atom model, *i.e.*, least-squares type fitting of the experimental data.

Furthermore, **SASREF** [194] can be used to perform quaternary structure modelling of a molecular complex with a known protein structure against the experimental scattering data; multiple non-unique structures may be obtained because one or more molecules are free to move in space.

3.1.4 Design of amoebapore A mutants

Site-directed mutagenesis is a useful tool to verify the proposed APA-1 dimer model and examine the importance of the residues that are proposed to be located at the dimer interface. Based on the previously proposed dimer model [24] (Figure 3.3), five mutants were designed by our collaborator in Germany (Prof. Joachim Grötzinger, Department of Biochemistry, Christian-Albrechts-Universität zu Kiel, Medical Faculty, Kiel, Germany). These mutants are proposed to disrupt or remove the pH-dependency of dimer formation, including E2Q, D63N, D64K, H75Q and H75K, where E2Q, D63N, K64Q and H75Q are considered to abolish the ionic interactions of the dimer, and H75K is proposed to eliminate the pH-dependency of dimer formation (at least over pH values ≤ 8). Based on NMR perturbation analysis which will be described in Section 3.3.2, additional mutants which are postulated to disrupt or stabilise APA-1 activity were designed (*i.e.*, E2A, K37A, K37Q, F41A and E2Q-K37Q). Biological assay tests along with NMR spectroscopy can be used to probe the activity and fold of these mutants.

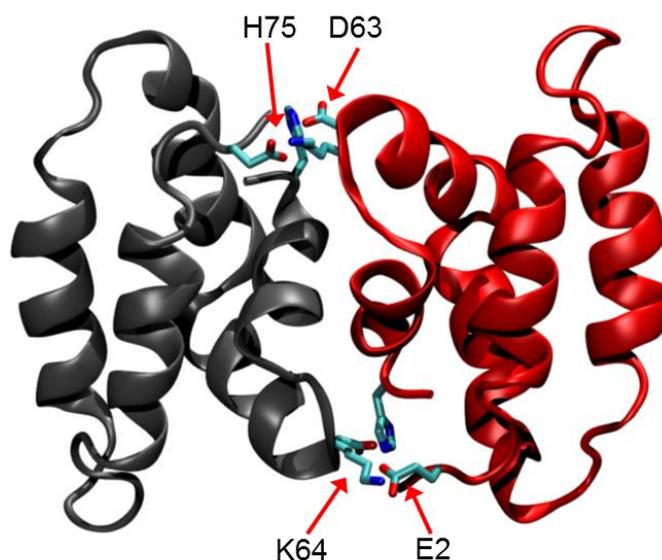


Figure 3.3: Proposed APA-1 dimer model in solution. Two ion pairs are involved in the dimer interface, H75–D63 and K64–E2. Five mutants: E2Q, D63N, D64K, H75Q and H75K, were designed based on this model, which are proposed to either disrupt or stabilise dimer formation.

3.1.5 Aims

The main research aims in this chapter were to:

- (1) Use NMR chemical shift perturbation analysis to identify residues located at the dimer interface;
- (2) Synthesise selectively $^{13}\text{C}/^{15}\text{N}$ -labelled APA-1 (Glu, Asp, Lys and His are labelled) samples to measure side-chain pKa values using NMR spectroscopy;
- (3) Perform concentration-dependent chemical shift mapping combined with small-zone SEC to estimate the dissociation constant of the APA-1 monomer-dimer equilibrium;
- (4) Perform SAXS experiments to investigate the gross shape and size of APA-1 in solution at different protein concentrations and pH values;
- (5) Design and synthesise APA-1 mutants based on the proposed dimer models in order to determine which model is correct, and investigate the importance of dimerisation in APA-1 functionality and identify residues that are crucial for APA-1 activity.

3.2 Methods and Materials

3.2.1 Sample preparation

Uniformly ^{15}N -labelled and selectively $^{13}\text{C}/^{15}\text{N}$ -labelled APA-1 samples were produced from the cell-free protein expression system and purified as described in Chapter 2. To reduce the cost of using double labelled material and minimise peak overlap, only charged amino acids were $^{13}\text{C}/^{15}\text{N}$ -labelled. Potassium glutarate replaced potassium glutamate to ensure sufficient Glu labelling [146].

3.2.2 NMR spectroscopy

APA-1 samples were prepared as described in Chapter 2. NMR experiments were performed at 25 °C. The pH values of the samples were measured before and after each experiment. The average readings were taken for data analysis. The pH showed little variation at most pH values (± 0.05 pH unit). However, the pH was not stable above 8.5, where the solution was not buffered. To ensure stable pH values, a 30 min incubation period was employed after adjusting the pH, and the pH values were measured again before each NMR experiment.

3.2.2.1 pH- and concentration-dependent chemical shift mapping

The 1D ^1H and 2D ^1H - ^{15}N HSQC spectra of uniformly $^{13}\text{C}/^{15}\text{N}$ -labelled APA-1 samples with concentrations of 210 μM were recorded as a function of pH between 3.0 and 8.5 at 0.5 pH unit increments. The MWs of APA-1 at each pH value were estimated by performing a 1D experiment, as described in Section 2.2.12.1 (Eq. 2.2 and 2.3). The errors of the T_2 values were estimated by taking the signal-to-noise ratio:

$$\text{Error}(T_2) = \frac{4(\Delta_A - \Delta_B)}{\ln(S/N_{0.25\text{ ms}} + S/N_{2.9\text{ ms}})} \quad \text{Eq. 3.3}$$

where Δ_A and Δ_B are the delay periods used in the 1D spectra, $S/N_{0.25\text{ ms}}$ and $S/N_{2.9\text{ ms}}$ are the signal to noise peak intensity ratios recorded using spin-echo delay periods of 0.25 and 2.9 ms in the two 1D spectra, respectively.

The 2D ^1H - ^{15}N HSQC backbone assignments obtained from Chapter 2 at pH 3.0 were used as an initial starting point for the titration. To investigate the pH and concentration effects on

dimer formation, 2D ^1H - ^{15}N HSQC spectra of uniformly $^{13}\text{C}/^{15}\text{N}$ -labelled APA-1 samples were acquired at pH 3.0 and 5.2, and at protein concentrations of 25 and 560 μM (*i.e.*, four samples). Chemical shift changes of the spectra recorded at pH 3.0 and 5.2 at each protein concentration were compared using Eq. 3.4:

$$\Delta\delta_{\text{NH}} = \sqrt{\Delta\delta_{\text{H}}^2 + \left(\frac{\Delta\delta_{\text{N}}}{6}\right)^2} \quad \text{Eq. 3.4}$$

where $\Delta\delta_{\text{H}}$ and $\Delta\delta_{\text{N}}$ represent the ^1H and ^{15}N chemical shift differences of a resonance, recorded at two pH values and one of the two concentrations. One standard deviation above the mean of the weighted chemical shift perturbation ($\Delta\delta_{\text{NH}}$) for all the resonances was used as a threshold, in which, any residue with a $\Delta\delta_{\text{NH}}$ value greater than the threshold value was considered to be influenced over the pH range studied. To separate the pH and concentration effects on the chemical shifts, the $\Delta\delta_{\text{NH}}$ obtained at the high and low protein concentrations were then compared to each other. The difference in the absolute weighted chemical shift perturbation values ($|\Delta\Delta\delta|$) was extracted by subtracting the pH-dependent chemical shift differences for the two protein concentrations [157]:

$$|\Delta\Delta\delta| = \Delta\delta_{560\ \mu\text{M}} - \Delta\delta_{25\ \mu\text{M}} \quad \text{Eq. 3.5}$$

Again, one standard deviation above the mean of the absolute chemical shift difference for all the resonances was adopted as a threshold. Any residue which gave differences greater than the threshold value was considered to show a significant pH-dependent chemical shift change at the high protein concentration. The residues identified were considered to correspond to residues that reside at the dimer interface.

It was observed in the 2D ^1H - ^{15}N HSQC spectra recorded at pH 5.2 that a set of resonances were missing, presumably because of slow to intermediate exchange on the NMR chemical shift time scale. This hampered the analysis of the dimerisation study at this pH. Consequently to study the influence of concentration on APA-1 dimerisation, a study was performed at pH 3.0 at two protein concentrations: 25 and 560 μM . Using pH values between 3 and 5.2 showed similar resonance broadening observed at pH 5.2. Averaged chemical shift differences at the two concentrations were calculated according to Eq. 3.5. Resonances that had chemical shift differences greater than one standard deviation above the mean were considered to show a concentration-dependent chemical shift change at pH 3.0.

3.2.2.2 Side-chain assignments of Glu, Asp and His residues

To obtain the side-chain assignments of Asp and Glu, a combination of NMR experiments were performed at pH 3.0 and 1.0 mM using the selectively $^{13}\text{C}/^{15}\text{N}$ -labelled APA-1 sample, including 2D ^1H - ^{15}N HSQC, 3D HCCH TOCSY [195], 2D CBCANH [196], 2D ^1H - ^{13}C constant time (CT) HSQC [197, 198], 3D HBHANH [199], 3D ^{15}N -edited TOCSY [200] and 2D and 3D HCCO [201, 202] experiments (Table 3.1, Figure 3.4). These experiments provided connectivities between the side-chain C_γ (Asp) and C_δ (Glu) (Figure 3.5) with other nuclei in the backbone and side-chain. A ^1H - ^{13}C CTHSQC spectrum was acquired to obtain the assignment of the imidazole CH_{ϵ_1} of the His residue (Figure 3.5) and this resonance was used as the probe nucleus to measure the imidazole ring pKa value.

Table 3.1: The NMR experiments used to achieve side-chain assignments of Asp, Glu and His residues.

NMR experiment type	Time (h)	Acquisition time (ms)	No. of complex points	Residue
2D HSQC	0.2	68.2 (H) x 105.3 (N)	1024 (H) x 128 (N)	All
3D HCCH TOCSY	66	68.2 (H) x 2.9 (C) x 4.3 (H)	1024 (H) x 35 (C) x 64 (H)	All
2D CBCANH	0.7	68.2 (H) x 5.6 (C)	1024 (H) x 54 (C)	All
2D CTHSQC ¹	13	68.2 (H) x 21.2 (C)	1024 (H) x 256 (C)	All
3D HBHANH	67	68.2 (H) x 15.4 (N) x 6.4 (H)	1024 (H) x 15 (N) x 49 (H)	All
3D ^{15}N -edited TOCSY	133	68.2 (H) x 17.5 (N) x 8.5 (H)	1024 (H) x 32 (N) x 128 (H)	All
2D HCCO	2–12	68.2 (H) x 32.6 (C)	1024 (H) x 64 (C)	Glu/Asp
3D HCCO	19	68.2 (H) x 3.7 (C_α) x 26.5 (CO)	1024 (H) x 16 (C_α) x 52 (CO)	Glu/Asp
2D CTHSQC ²	0.3	65.7 (H) x 14.1 (C)	1024 (H) x 64 (C)	His



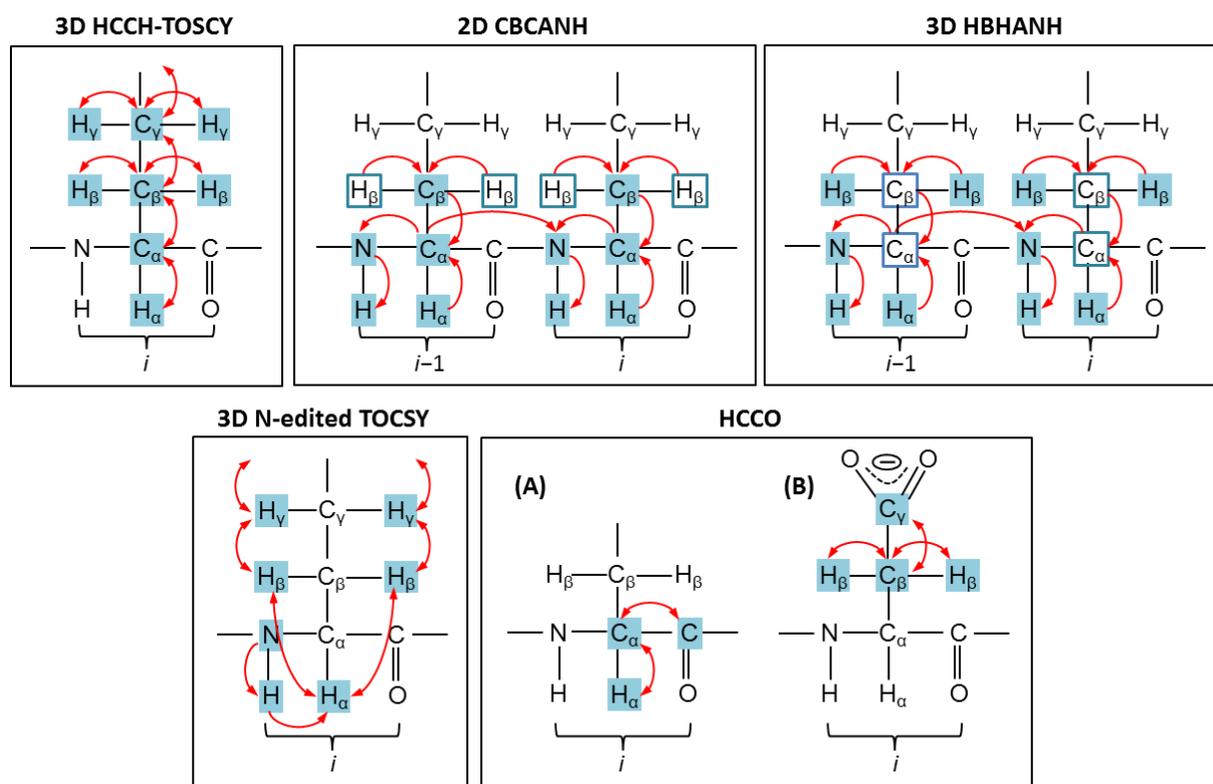


Figure 3.4: Schematic overview of the different NMR experiments used for assigning the side-chain of Glu and Asp residues. Shaded nuclei are frequency labelled, whereas open circled nuclei are only involved in the transfer of magnetisation between nuclei. Double-headed arrows indicate that the experiment is an “out and back” experiment and single headed arrows indicate that the experiment starts at a particular ^1H spin and detects another ^1H nucleus. **(3D HCCH-TOCSY)** The side-chain aliphatic ^1H and ^{13}C resonances are correlated. **(2D CBCANH)** The amide ^1H and ^{15}N resonances are correlated with those of the intra- and inter-residue $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ resonances. Since a selectively labelled APA-1 sample was used, for each amide group only one set of $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ resonances were observed (unless residues were sequentially labelled). **(3D HBHANH)** The amide ^1H and ^{15}N resonances are correlated with the intra- and inter- $^1\text{H}\alpha$ and $^1\text{H}\beta$ resonances. **(3D ^{15}N -edited TOCSY):** All the aliphatic ^1H nuclei are correlated with the amide ^{15}N and amide ^1H nuclei. **(2D HCCO and 3D HCCO)** Usually the intra-residue $^1\text{H}\alpha$, $^{13}\text{C}\alpha$ and ^{13}CO resonances are correlated (A). To measure the side-chain pKa of Asp and Glu residues, the β - and γ -methylene proton are correlated with carboxyl $\text{C}\gamma$ and $\text{C}\delta$ nuclei, respectively. The example for Asp is given in (B).

3.2.2.3 pH titration and the determination of the pKa values

Two selectively $^{13}\text{C}/^{15}\text{N}$ -labelled APA-1 samples with concentrations of 1.0 and 0.1 mM were prepared for side-chain pKa measurements. Each of the two samples was split into equal volumes and their pH values were adjusted to the extremes of the pH range examined (*i.e.*, pH 2 and 8.5 for the 1 mM sample, and pH 2 and 10 for the 0.1 mM sample). The desired pH values of the NMR samples were achieved by mixing the appropriate portions of the two samples.

The 2D HCCO and 2D CTHSQC experiments were performed for estimation of the side-chain pKa values of Asp and Glu residues and His residues, respectively. The pH titration using the 1.0 mM APA-1 sample was performed between pH 2–8.5 for the 2D HCCO experiment (13 titration points) and pH 3–9 for the 2D CTHSQC experiment (13 titration points). The titrations were performed on the 0.1 mM protein sample over the pH range of 2–8 for the 2D HCCO (24 titration points) and 3.5–10 for the 2D CTHSQC (19 titration points). For all the titrations, the ^1H chemical shifts were referenced to TSP at 0 ppm, and the ^{13}C chemical shifts were indirectly referenced [203]. The influence of pH on the TSP signals was corrected using Eq. 3.6 and 3.7 [204].

$$\delta_{\text{corr(C)}} = \delta_{\text{C}} + 0.1 \times (1 + 10^{5.0 - \text{pH}})^{-1} \quad \text{Eq. 3.6}$$

$$\delta_{\text{corr(H)}} = \delta_{\text{H}} + 0.019 \times (1 + 10^{5.0 - \text{pH}})^{-1} \quad \text{Eq. 3.7}$$

where δ_{C} and δ_{H} are the chemical shifts of the ^{13}C and ^1H signals before the pH correction, and $\delta_{\text{corr(C)}}$ and $\delta_{\text{corr(H)}}$ are the chemical shifts after correcting the pH influence on the TSP signal.

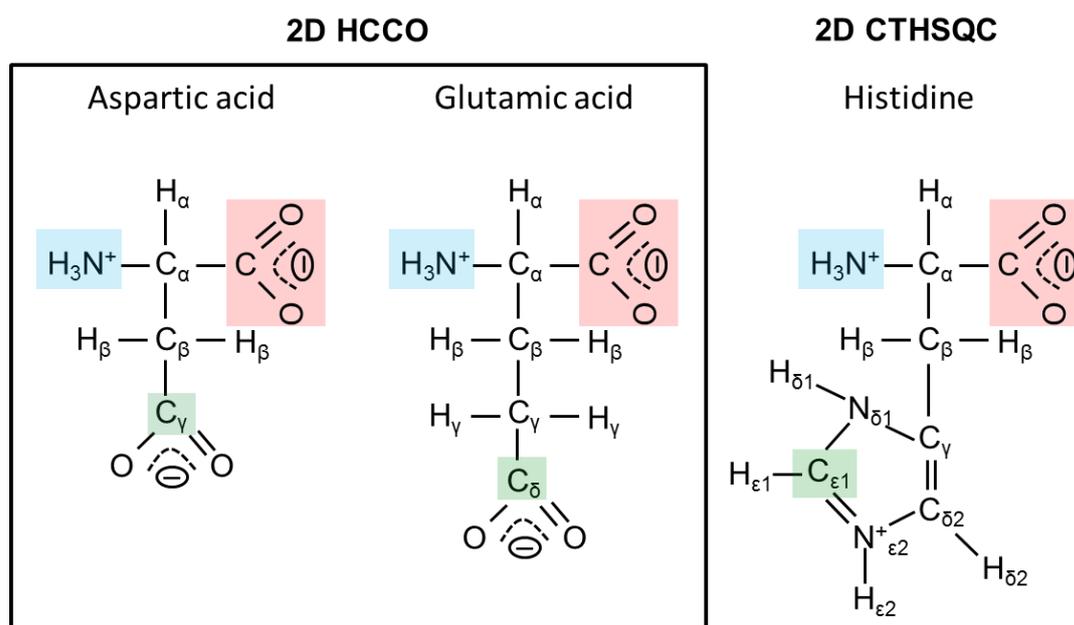


Figure 3.5: Schematic diagrams representing the structures of residues Asp and Glu at neutral pH and His at low pH (*e.g.*, pH 3). Chemical shift changes of the resonances representing C_γ and C_δ nuclei in Asp and Glu (shaded in green), respectively, were followed in the 2D HCCO experiments. Chemical shift changes of the imidazole CH_{ϵ_1} moiety (shaded green) of His were followed in the 2D CTHSQC experiment to determine the pKa of the imidazole ring. At low pH, either of N_{δ_1} and N_{ϵ_2} can release a proton (H^+), and therefore His is in tautomeric state. The shaded areas depicted in blue and red represent the amide and carboxylic groups, respectively.

The pKa values of the side-chain carboxyl groups of Asp (C γ) and Glu (C δ) and the imidazole region (C ϵ_1) (Figure 3.5) of the His residue were determined by following the chemical shift changes of the corresponding resonances as a function of pH (Figure 3.5). The pH titration data for the ^{13}C chemical shifts were analysed using the Python Programming Language (Python Software Foundation, Delaware, USA) and the script was written by Dr Michael Schmitz, Romel Bobby and Wei Li (School of Biological Sciences and School of Chemical Sciences, The University of Auckland). According to previous studies performed to estimate side-chain pKa values [205-208], the curve fitting was carried out using a modified Henderson-Hasselbalch Hill equation [209]:

Eq. 3.8

$$\delta_{\text{obs}} = \frac{(\delta_{\text{HA}} + \delta_{\text{A}}^- 10^{n(\text{pH} - \text{pKa})})}{1 + 10^{n(\text{pH} - \text{pKa})}}$$

where δ_{obs} is the observed chemical shift, δ_{HA} and δ_{A}^- are the chemical shifts in the acidic and basic pH limits, respectively, and n is the Hill coefficient which represents the slope of the titration curve in the transition region. The Hill coefficient measures the cooperativity resulting from the influence of hydrogen ion binding at other proximate sites (*e.g.*, amide group and ionisable groups) to the desired ionisable sites being examined. An n value of 1 indicates independent binding, $n < 1$ and $n > 1$ indicate negative and positive cooperative binding, respectively. The error of the pKa value for the His residue was estimated by adding a small number (*i.e.*, ± 0.1 ppm) to the chemical shifts of the resonances in the 2D CTHSQC spectra; and the errors of the pKa values for Asp and Glu residues were obtained by taking the ^{13}C chemical shift differences between the paired resonances (the H β for Asp and H γ for Glu) in the 2D HCCO spectra. Five hundred simulations (Monte-Carlo approximation of error) were used in the Python Script to fit the curves.

The pH titration was repeated at 0.1 mM APA-1 by performing a series of 2D ^1H - ^{15}N HSQC spectra with 0.25 pH increments between pH 2–9.5 (22 titration points). The pKa values were estimated by following the ^1H chemical shifts as a function of pH. The data which showed a single titration event was fitted using the modified Henderson-Hasselbalch equation with the Hill coefficient (Eq. 3.8) and data which showed complex titration events were not fitted.

3.2.2.4 Estimation of the K_d using NMR spectroscopy

To study the dissociation constant (K_d) of the monomer-dimer equilibrium of APA-1, 2D ^1H - ^{15}N HSQC spectra at eight protein concentrations ranging from 25 to 840 μM at pH 3.0

were recorded. Higher pH values were not used, as mentioned, because of significant line-broadening of particular resonances of interest. The K_d can be defined using Eq. 3.9 and 3.10 [210]:

$$K_d = \frac{[M]^2}{[D]} \quad \text{Eq. 3.9}$$

$$K_d = \frac{2p_M^2 \times [APA-1]}{p_D} \quad \text{Eq. 3.10}$$

where $[M]$ and $[D]$ are the concentrations of monomer and dimer in a protein sample, respectively, p_M and p_D are the populations of APA-1 in monomeric and dimeric states which add to 1, and $[APA-1]$ is the total protein concentration. Based on the assumption that monomers and dimers undergo fast exchange on the NMR chemical shift time scale, the observed ^{15}N chemical shift (ω_{obs}) can be defined by the population weighted average of the monomeric and dimeric APA-1 chemical shift, ω_M and ω_D (Eq. 3.11). By combining Eq. 3.10 and 3.11, the ω_{obs} can also be defined as Eq. 3.12 [210]:

$$\omega_{\text{obs}} = p_M \omega_M + p_D \omega_D = \omega_D + p_M (\omega_M - \omega_D) \quad \text{Eq. 3.11}$$

$$\omega_{\text{obs}} = \omega_D + \frac{\sqrt{K_d (K_d + 8[APA-1])} - K_d}{4[APA-1]} (\omega_M - \omega_D) \quad \text{Eq. 3.12}$$

Since E2 was the only residue whose corresponding resonance showed chemical shift changes upon changing the concentration of APA-1 and visible over the concentration range, the K_d , ω_M and ω_D of E2 were fitted using the Levenberg-Marquardt non-linear least squares optimisation to different APA-1 concentrations and ω_{obs} (^{15}N chemical shifts) to predict the K_d of the APA-1 monomer-dimer equilibrium [211] at pH 3.0. The equation was implemented into the Python Programming Language (Python Software Foundation, Delaware, USA) by Dr Michael Schmitz (School of Chemical Sciences, The University of Auckland) to perform the fitting.

3.2.3 Size exclusion chromatography

The SEC experiments were performed on an ÄKTA Purifier system (GE Healthcare, Little Chalfont, UK) using an analytical Superdex 75 15/150 gel filtration column (3 ml) (GE Healthcare, Little Chalfont, UK). The column was pre-equilibrated with 10 mM sodium citrate/NaPi at pH 5.2. All experiments were performed at 4 °C. Unlabelled APA-1 samples at concentrations ranging between 10 µM and 2.5 mM, pH 5.2, were prepared by a series of dilutions and incubated for at least 1 h at 4 °C prior to loading onto the column. In each run, 20 µl of the APA-1 solution was loaded onto the column at a flow rate of 0.5 ml/min. The column was calibrated using five proteins of known MW (*i.e.*, aprotinin (6.5 kDa), ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa), ovalbumin (44 kDa) and conalbumin (75 kDa)). The elution volume (V_e) of APA-1 was analysed to study the dimerisation behaviour of APA-1.

To investigate the monomer-dimer equilibrium of APA-1, the peak areas corresponding to the monomer and dimer peaks were calculated using the UNICORN software routines that came with the ÄKTA Purifier system. The APA-1 concentration was plotted against the total peak area and fitted to a linear equation $y = 0.00524009x$ where y is the total peak area and x is the protein concentration, $R^2 = 0.9997$. The concentration of monomers ([M]) and dimers ([D]) present in the solution at each protein concentration were estimated using the peak areas and the coefficient obtained from the fitted linear equation:

$$[M] = \frac{A_M}{0.00524009} \quad \text{Eq. 3.13}$$

$$[D] = \frac{A_D}{0.00524009 \times 2} \quad \text{Eq. 3.14}$$

where A_M and A_D represent the peak areas corresponding to the APA-1 monomer and dimer concentrations in a sample. Based on the monomer and dimer concentrations, the K_d was estimated using Eq 3.9. To exclude the dilution factor on K_d , four protein concentrations (*i.e.*, 12, 24, 47, and 95 µM, where monomers and dimers co-exist) were used to estimate K_d and the calculated values were averaged.

3.2.4 Small angle X-ray scattering

One set of SAXS data were collected at the Australian Nuclear Science and Technology Organisation (ANSTO) (Sydney, Australia) on a Bruker Nanostar instrument with a copper

target, three-pinhole collimation and HiStar 2D detector with 100 μm pixel size. SAXS data of APA-1 samples with a concentration range between 550 μM and 2.1 mM at pH values of 3.0 and 5.2 were collected. The sample with a volume of 15 μl and matched buffer blank were inserted sequentially in the same sealed quartz capillary. All experiments were performed at 25 $^{\circ}\text{C}$. The sample to detector distance was 65 cm which gave a Q range of $0.01 \text{ \AA}^{-1} \leq Q \leq 0.43 \text{ \AA}^{-1}$ ($Q = 4\pi \sin(\theta)/\lambda$, 2θ is the angle between the incident and scattered beams and $\lambda = 1.5418 \text{ \AA}$ is the X-ray wavelength). Each SAXS measurement was recorded for 3 h for a protein concentration $> 730 \mu\text{M}$ and 6 h for samples with concentrations $\leq 730 \mu\text{M}$.

A second set of X-ray scattering experiments of APA-1 were performed on the SAXS/WAXS beamline of the Australian Synchrotron (Melbourne, Australia) using a Pilatus 1M image plate detector. The scattering data was collected with a 10 s exposure time (10 \times 1 sec frame exposures) for protein concentrations ranging from 95 μM to 190 μM , at pH 3.0. During the exposure time, protein samples were continuously flowed through the 1.5 mm diameter capillary at a flow rate of 6 $\mu\text{l sec}^{-1}$ using a syringe pump. The SAXS data of a buffer blank was recorded before each experiment and a water rinsing cycle was performed after each sample run. Using a sample-detector distance of 1.6 m, a Q range of $0.006 \text{ \AA}^{-1} \leq Q \leq 0.65 \text{ \AA}^{-1}$ was covered ($\lambda = 1.0322 \text{ \AA}$).

Reduction, background subtraction and analysis of scattering data were performed using the SAXS15ID program [212]. The R_g , D_{max} , $P(r)$ and the forward scattering intensity were calculated from the experimental scattering data collected from ANSTO using the indirect Fourier transform method with the program GNOM [184]. Theoretical scattering profiles were generated from the 20 lowest energy NMR structures of APA-1 monomer and two APA-1 dimer models using AXES [193]. *Ab initio* shape restoration was only performed using the dataset collected from Australian Synchrotron (*i.e.*, with protein concentration range between 95 μM to 190 μM) using GASBOR [189] because data recorded at ANSTO was found to be of insufficient quality to enable GASBOR to derive sensible models. Dummy atom models were averaged, filtered and superimposed with the atomic resolution APA-1 monomer and dimer structures using DAMAVER, DAMFILT and SUPCOMB [191, 213]. Default parameters were used for all calculations.

3.2.5 Site directed mutagenesis

Five mutants were provided by our collaborator (Prof. Joachim Grötzinger, Department of Biochemistry, Christian-Albrechts-Universität zu Kiel, Medical Faculty, Kiel) based on the

previously proposed dimer model [24] (Figure 3.3): E2Q, D63N, D64K, H75Q and H75K. Another set of APA-1 mutants were designed based on a new dimer model proposed in Section 3.3.2: E2A, K37A, K37Q, F41A and E2Q-K37Q. The plasmids of these mutants were prepared by MSc student Mikhail Legkodimov (School of Biological Sciences, The University of Auckland) using *de novo* synthesis by ZyGEM (ZyGEM Corporation Ltd., Hamilton). The correct sequences of the DNA constructs were verified by DNA sequencing carried out at the Allan Wilson Centre Genome Service (Massey University, Palmerston North). Analytical cell-free reactions were performed to test the viability of the APA-1 mutants and the expressions were examined by 4–12% SDS-PAGE. The mutants which gave reasonable levels of cell-free protein expression were scaled up to the 1 ml preparative scale using ¹⁵N-labelled amino acids. The overall fold of the mutants were examined using 2D ¹H–¹⁵N HSQC experiments at pH 3.0.

3.3 Results

3.3.1 Variation in the molecular weight as a function of pH

One-dimensional NMR experiments were performed as a function of pH using a 210 μM sample. The MWs of APA-1 at different pH values were determined from 1D spin-echo experiments (Section 2.2.12.1). The molecular mass of APA-1 increased as the pH increased (Figure 3.6). The MW of APA-1 was estimated to be 8.5 ± 0.4 kDa at pH 3.0 and 12.9 ± 0.8 kDa at pH 5.2, which corresponds approximately to the monomer (*i.e.*, 8.2 kDa) and dimer (*i.e.*, 16.4 kDa) species. The approximate MW determined at pH 5.2 is underestimated because the equation used (Eq. 2.3) assumes that the protein is a symmetrical globular fold. This is not the case for the APA-1 dimer structure and thus leads to an underestimate of the MW. In addition, the NMR experiment measures the weight averaged MW of the dimeric and monomeric APA-1 populations (possibly even higher MW order), thus it is possible that at pH 5.2 the MW determined reflects the presence of a fraction of monomers. Nonetheless, there is a clear increase in MW at pH 5.5, in agreement with the previous result (Figure 1.16) [24]. In addition, the MW at pH 8.5 was estimated to be 12.9 ± 0.9 , similar to what was obtained at pH 5.2, suggesting the presence of a dimeric APA-1 population at this pH value, also consistent with Figure 1.16. Similar MWs were obtained when the pH was dropped back to 5.2 and 3.0, indicating reversible self-association.

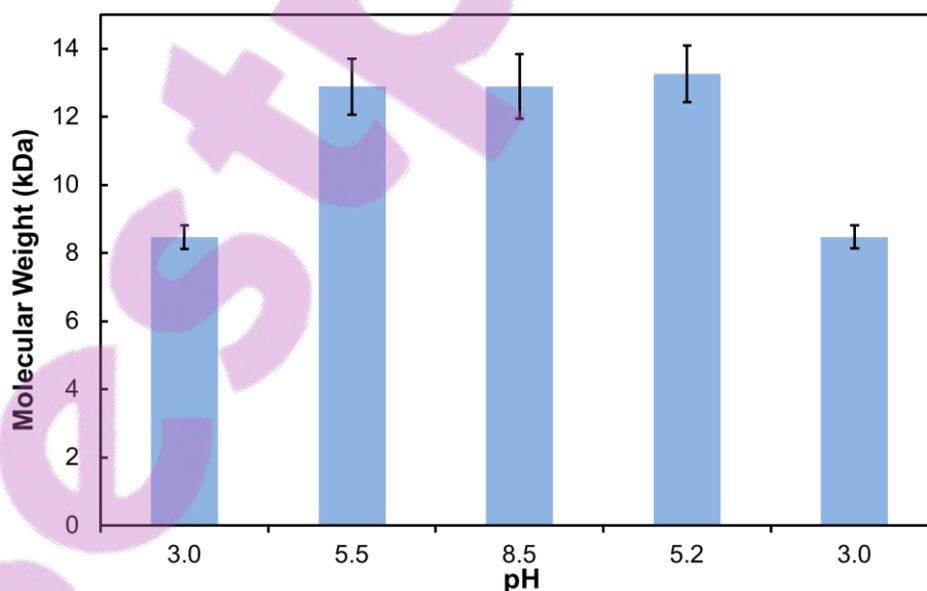


Figure 3.6: The MWs of APA-1 estimated using T_2 and τ_c values obtained from the 1D spin-echo ^1H NMR experiments. Based on the estimation, APA-1 was proposed to exist predominantly as monomers at pH 3, and in a monomer-dimer population at pH 5.5 and 8.5 at 210 μM .

3.3.2 *The influence of pH and concentration on the chemical shifts of resonances – mapping the dimer interface*

To characterise the dimer interface using NMR spectroscopy assignment of resonances in the 2D ^1H - ^{15}N HSQC spectra at different pH values was required. Figure 3.7 shows an overlay of 2D ^1H - ^{15}N HSQC spectra recorded at pH values of 3.0, 5.2 and 8.5. Intriguingly, the intensity of particular resonances decreased dramatically as the pH was increased (circled resonances: L7, L34, K37, G40, F41, I42 and T44), presumably due to chemical exchange caused by pH and/or APA-1 dimerisation. To distinguish changes in chemical shifts of resonances because of pH and protein concentration, 2D ^1H - ^{15}N HSQC spectra were recorded at pH 3.0 and 5.2, and at protein concentrations of 25 and 560 μM (data not shown). At the low concentration and as the pH rose from 3.0 to 5.2, resonances corresponding to residues G40, F41, L7, K37, I42 and T44 showed significant line-broadening, and these resonances were no longer detected above the background noise as the pH increases. Ten resonances showed chemical shift differences above one standard deviation of the mean upon pH changes. Four of the resonances correspond to Asp residues (D29, D51, D63 and D66) in which the change in the protonation state has influenced the nearby chemical environment of the backbone amide group. The other six resonances correspond to residues F52, I61, V65, A67, A69 and I70, are located in close structural proximity to the Asp residues, suggesting the chemical shift changes are due to changes in the ionisation states of the Asp side-chains. The experiment was repeated at a protein concentration of 560 μM and similar results were obtained. The observed line-broadening may also arise from particular motional events that show a pH-dependence, *i.e.*, chemical exchange. Although highly unlikely, such intrinsic motional changes may account for the line-broadening as the pH was raised, and only through further analysis, *e.g.*, mutational analysis and extensive dynamics analysis, can this explanation be eliminated.

To separate the chemical shift differences resulting from pH and dimerisation, the chemical shift differences obtained at high and low protein concentrations were then compared to each other. The absolute differences in the chemical shift changes $|\Delta\Delta\delta|$ were extracted by subtracting the normalised pH-dependent chemical shift differences achieved at the two protein concentrations. No chemical shift difference was calculated for the resonances that were not visible at pH 5.2. A total of five resonances were identified to show significant pH-dependent chemical shift differences that were greater than one standard deviation above the mean at the higher concentration. The corresponding residues are I3, C5, S33, A38 and T44. In addition, resonances which were not visible at pH 5.2 represent residues in close structural proximity to

the five residues identified to show large chemical shift differences. A previous study of rat CD2 domain1, which dimerises in a pH-dependent manner and has weak self-association ($K_d \sim 3\text{--}6 \text{ mM}$), has shown that residues which displayed concentration-dependent chemical shifts and showed line-broadening were located at the dimer interface [168, 214]. Consequently, these residues identified in APA-1 are likely to form the dimer interface and the association is likely to be weak.

Since a number of resonances were broadened at pH 5.2 and thus an unambiguous identification of the dimer interface was not possible, a similar set of NMR experiments were recorded at pH 3.0. As the protein concentration increased from 25 to 560 μM , a subset of residues was observed to show substantial line-broadening and chemical shift changes. The spectra acquired at the two concentrations were overlaid (Figure 3.8). The chemical shift differences at the two concentrations were normalised and plotted as shown in Figure 3.9. One standard deviation from the mean weighted chemical shift perturbation was considered to be a significant chemical shift change. Fourteen resonances were found to display significant changes in chemical shifts and correspond to residues E2, I3, L4, L7, I31, L34, N36, K37, A38, G40, F41, I42, T44 and T47. These residues were also identified in the previous study at pH 5.2. Therefore, these residues are likely to be located at the dimer interface. The residues which showed concentration-dependent chemical shift changes were mapped to the APA-1 dimer model. These residues were clearly clustered to form a contiguous region near the N-terminus and the loop between helices 2 and 3 (Figure 3.9). The residues which were previously proposed to be located at the dimer interaction surface showed no significant chemical shift changes and are located on the opposite site of the clustered region, and thus are not involved in APA-1 dimerisation.

Based on the chemical shift mapping study, a new dimer model was generated by Dr Michael Schmitz (School of Chemical Sciences, The University of Auckland) using the docking programme HADDOCK [215] (Figure 3.10). Based on the dimer model, K37 is in close proximity to E2 on the adjacent monomer. An electrostatic interaction is proposed to be crucial in facilitating APA-1 pH-dependent dimerisation. Two F41 residues from each monomer on the backside of the model are postulated to form a pi-stacking interaction. This interaction is also considered crucial in stabilising the dimer.



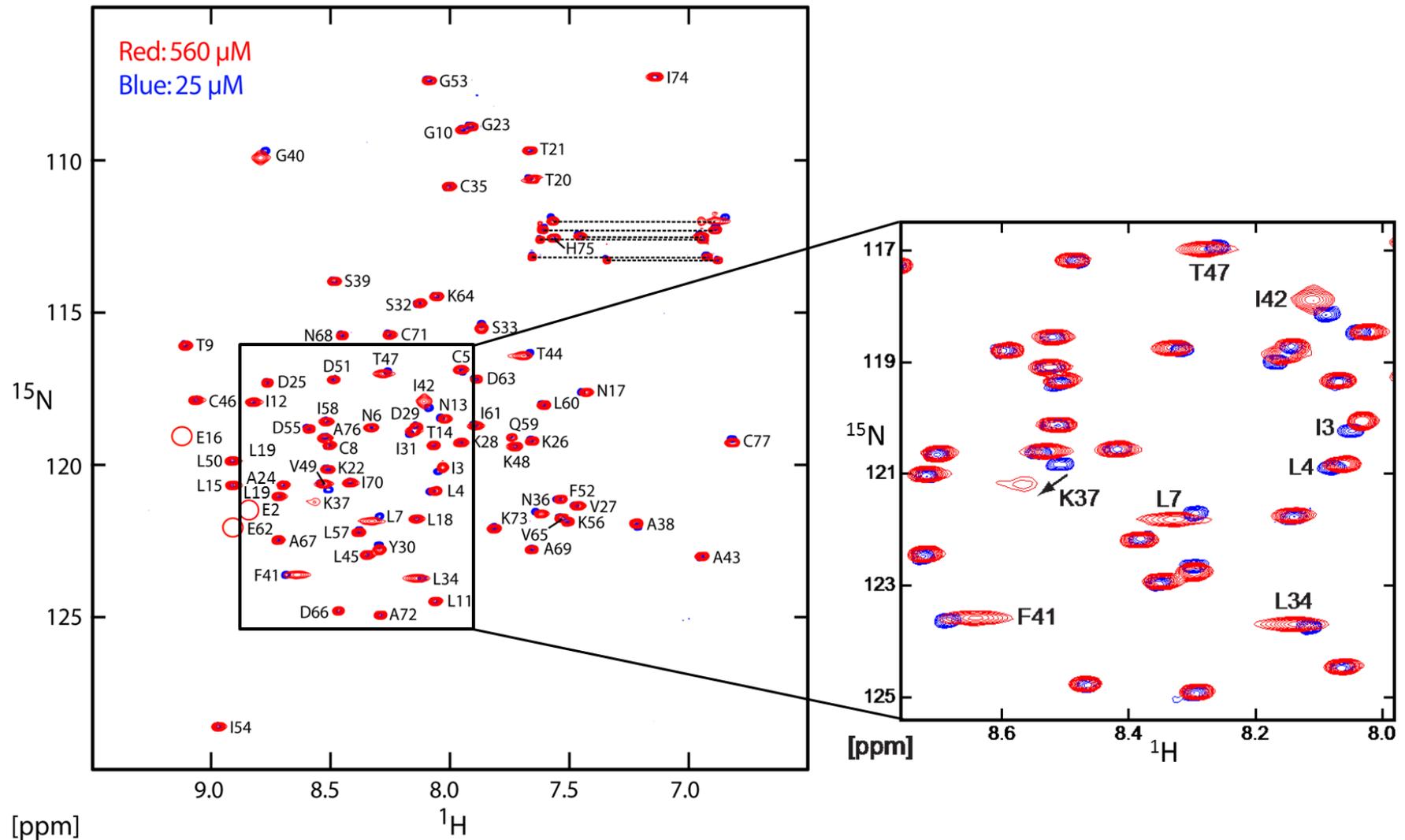


Figure 3.8: 2D ^1H - ^{15}N HSQC spectra of APA-1 at pH 3.0, and concentrations of 25 μM (blue) and 560 μM (red). Labeled resonances on the right panel showed noticeable line-broadenings and chemical shift changes. The assignment of Glu resonances which were not clearly visible in the spectra due to insufficient ^{15}N -labelling are circle in red.

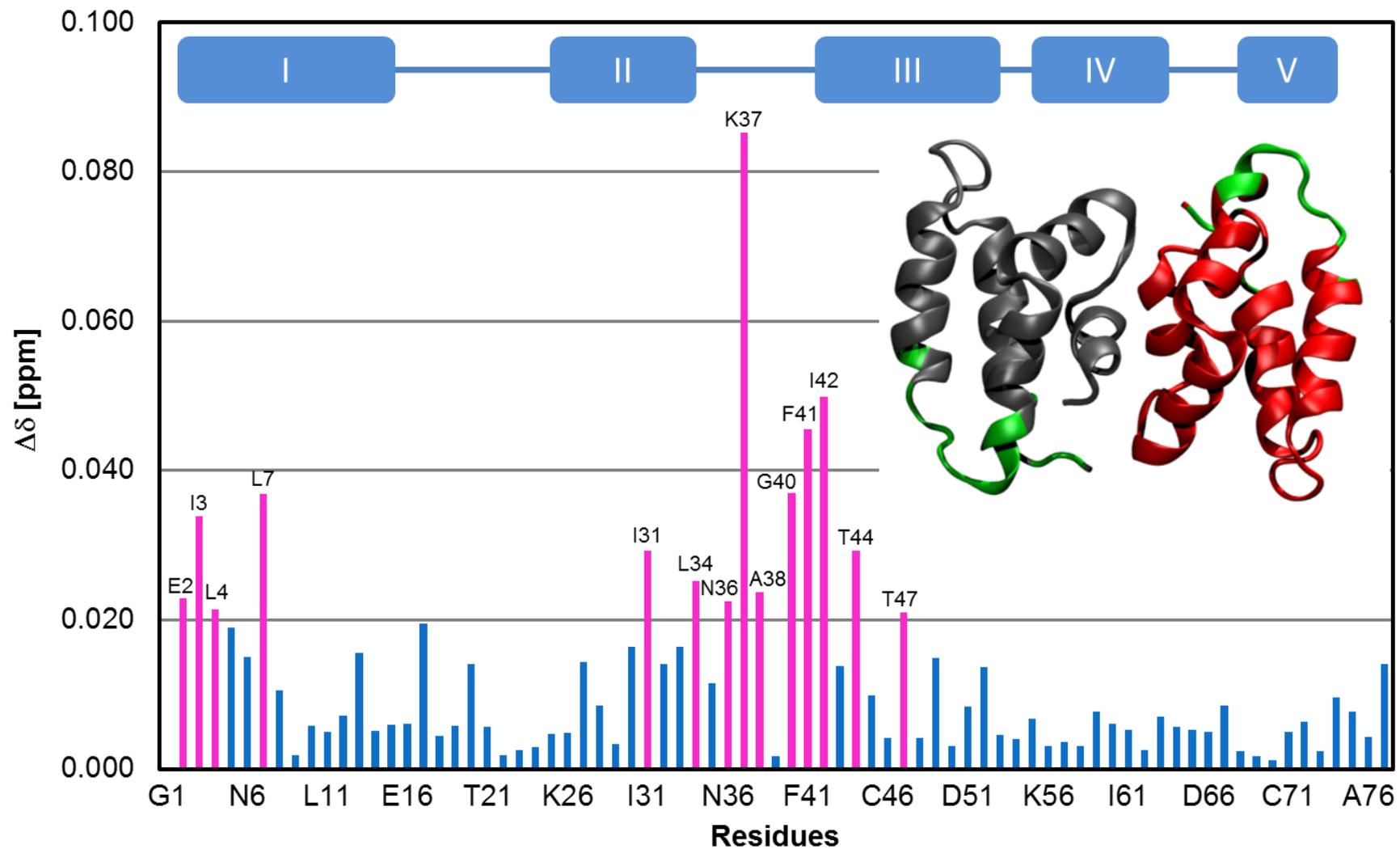


Figure 3.9: Normalised chemical shift differences of residues at protein concentrations of 25 μM and 560 μM , pH 3.0. Changes colored in pink indicate resonances that showed chemical shift changes above one standard deviation of the mean. Inset: these residues (coloured in green) were mapped to the original dimer model.

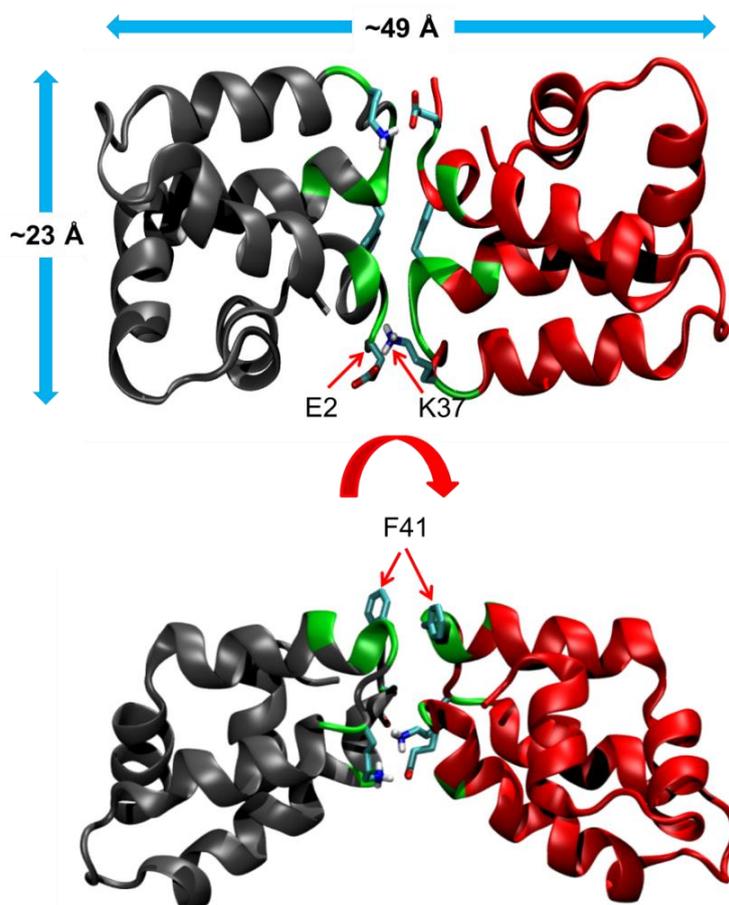


Figure 3.10: Newly proposed dimer model. The dimer model was proposed based on the NMR chemical shift perturbation study in combination with the monomeric structure and the programme HADDOCK [215]. Residue K37 is hypothesised to form an ion pair with E2. Two F41 residues from each monomer on the backside of the model are postulated to form a pi-stacking interaction. The dimensions of the dimer are ~23 and ~49 Å.

3.3.3 Side-chain resonance assignments

APA-1 has nine acidic amino acids: six Asp residues and three Glu residues. Side-chain resonance assignment of the acidic amino acids was carried out using selectively $^{13}\text{C}/^{15}\text{N}$ -enriched APA-1. Although ^1H chemical shifts for nuclei were available (although not complete), the assignment of heteronuclei was required to measure the pKa of Glu and Asp side-chains. A two-dimensional $^1\text{H}-^{15}\text{N}$ HSQC spectrum of the selectively labelled APA-1 was acquired and assigned at pH 3.0 using previously reported assignment information (Figure 3.11, Section 2.3.4.3). Amino acid scrambling from Glu to Gln and from Asp to Asn was observed, resulting in six additional resonances. The resonance corresponding to the Gln residue had a similar peak intensity to the resonances arising from Glu and Asp residues, indicating moderate transamination activity by glutamine synthetase (Section 2.3.4.2); whereas the signal intensities

of the Asn resonances were noticeably weaker indicating minor transamination by asparagine synthetase has taken place (Section 2.3.4.2). Residue E16 and K56 gave rise to more than one resonance in the spectrum, indicating that these residues are present in different conformations. Based on the resonance intensities, the populations corresponding to the two weak resonances of E16 were estimated to be ~15 and 17% of the total population. A 17.1% population was estimated for the weak resonance arising from K56.

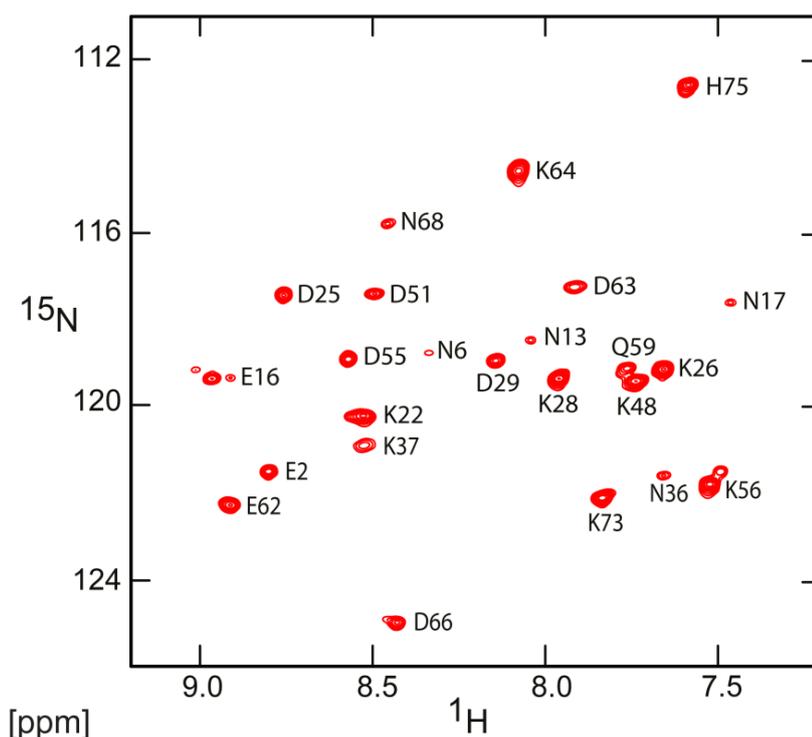


Figure 3.11: Assigned 2D ^1H - ^{15}N HSQC spectrum of selectively $^{13}\text{C}/^{15}\text{N}$ -enriched APA-1 at concentration of 100 μM and pH 3.0.

To determine the residue specific pKa values of acidic residues, assignments of $\text{C}\gamma$ (Asp) and $\text{C}\delta$ (Glu) resonances (Figure 3.5) were required. To achieve this, a series of 2D and 3D NMR experiments were performed. The assignments of the 3D HCCH-TOCSY spectra were achieved by combining the $\text{C}\alpha$ assignments taken from the previously recorded 3D HNCA experiment (Section 2.3.4.3) and the available 1D ^1H chemical shift assignments (Prof. Joachim Grötzinger, Department of Biochemistry, Christian-Albrechts-Universität zu Kiel, Medical Faculty, Kiel). The ^1H and ^{13}C assignments of 21 residues of a possible 24 residues were completed. Resonances arising from D25, D29 and D51 could not be assigned because 1D ^1H side-chain assignment information was not available, as well as weak signal intensities and signal overlap in the $\text{C}\beta/\text{H}\beta$ region of the 3D spectrum.

To obtain the C β chemical shifts of the Asp residues, a 2D ^1H - ^{13}C CBCANH spectrum of APA-1 was acquired. The correlations for the C α and C β nuclei are of opposite sign which aids the assignment process. Eighteen resonances arising from C α and fifteen resonances from C β atoms were assigned, using the H $_N$ and C α chemical shifts obtained from the 2D ^1H - ^{15}N HSQC and 3D HNCA spectra (Sections 2.3.4.2 and 2.3.4.3). Resonances arising from D51 and Q59 were in overlap with resonances arising from K22 and K48, respectively (assignments are shown in brackets in Figure 3.12). Resonances arising from C β nuclei of E2, N6, N13, N17, D25, D29, N36, K37 and D63 were not observed. The amino acid scrambling of Asp to Asn led to a reduction of the signal intensities of resonances arising from Asp residues. Resonances for E2 and K37 were not observed.

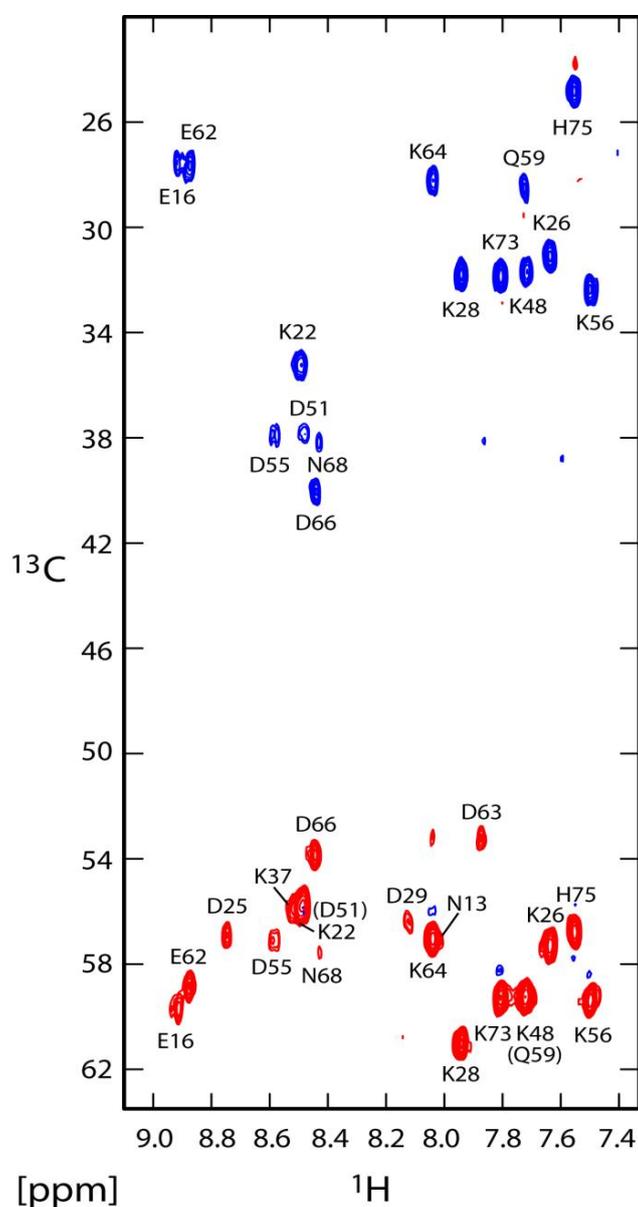


Figure 3.12: 2D CBCANH spectrum of APA-1 recorded at 1 mM and pH 3.0. The C α resonances are coloured red and C β resonances are coloured blue.

The assignment information obtained from the 3D HCCH-TOCSY and 2D CBCANH experiments were used to assign the 2D ^1H - ^{13}C CTHSQC of APA-1 (Figure 3.13). The weak resonances at frequencies between 20–26 ppm (^{13}C) arise from methyl groups. The origin of the methyl groups is probably because of minor amino acid scrambling that had occurred during the cell-free protein synthesis which led to the production of Ala, Val, Leu and Ile residues [216]. The population of the methyl groups was estimated to be 1.6% of the main APA-1 species, as determined by the relative signal intensities against resonances arising from the desired labelled amino acids. Most of the resonances arising from $\text{C}\alpha$ atoms were not observed due to signal overlap with the water signal between ~ 4.5 – 5 ppm. The $\text{C}\beta\text{H}\beta$ resonance correlation of the sole His was easily identified. No signals corresponding to $\text{C}\gamma\text{H}\gamma$ of Glu and $\text{C}\beta\text{H}\beta$ of Asp residues were observed. This was because the refocusing period used in the constant-time experiment was set to refocus the C-C (aliphatic) scalar coupling of ~ 35 Hz and not the larger C-C (side-chain C=O) coupling that is ~ 55 Hz.

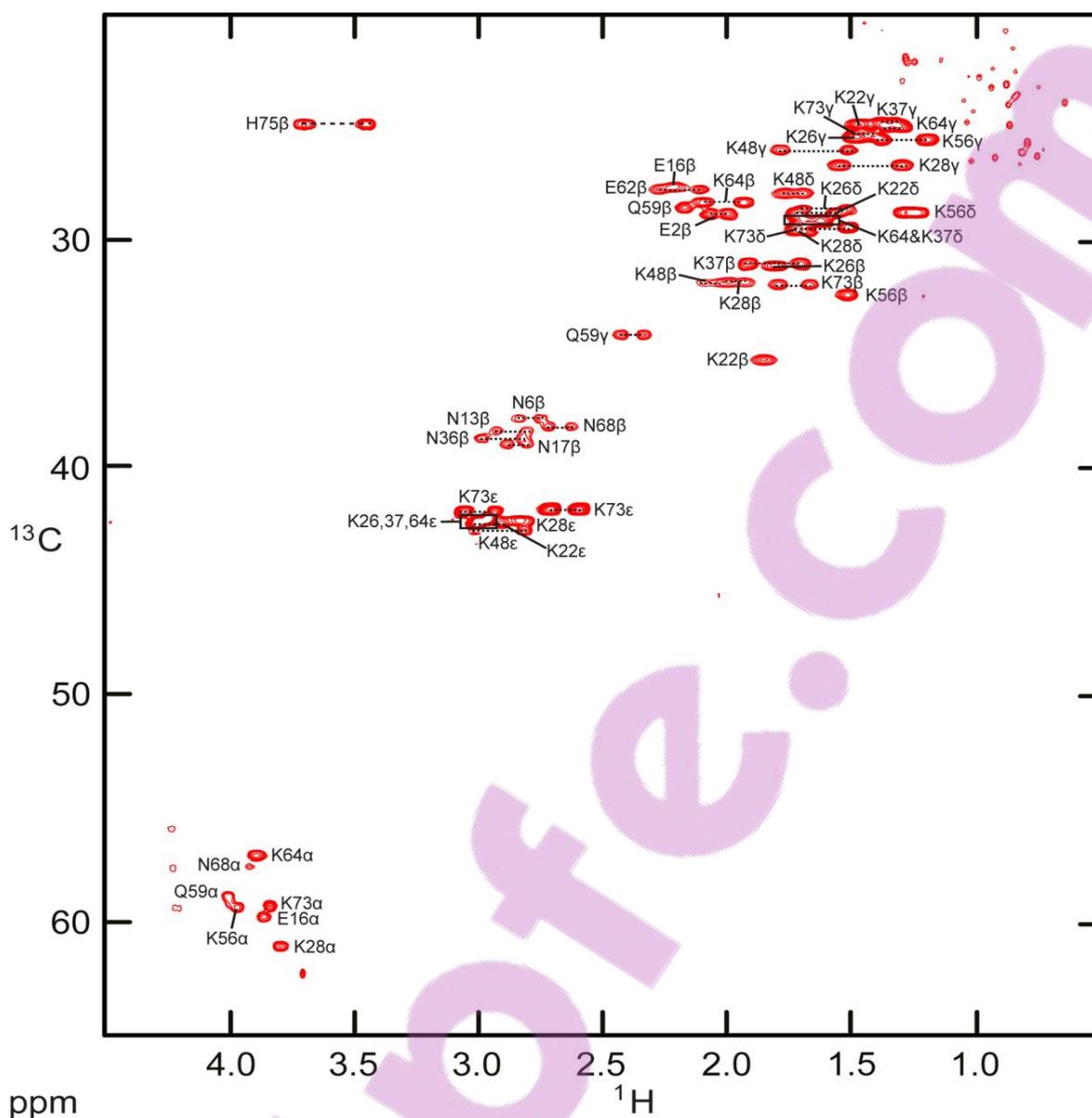


Figure 3.13: 2D ^1H - ^{13}C CTHSQC spectrum of APA-1 recorded at 1mM and pH 3.0. Dashed lines represent methylene groups (CH_2). The unassigned and weak resonances between 20–26 ppm are raised from methyl groups. Not all C_α atoms were assigned because of signal overlap with the water resonances. Signal overlap was observed for resonances arising from C_δ and C_ϵ of Lys residues, especially around 28.8–29.2 and 42.0–42.5 ppm (^{13}C).

Missing side-chain assignments for residues D25, D29 and D51 were obtained using 3D HBHANH and 3D ^{15}N -edited TOCSY experiments. To obtain the C_γ (Asp) and C_δ (Glu) assignments, a 3D HCCO experiment was recorded. Beside D25 and D29, unambiguous assignments of the carboxyl groups were made. Missing assignment information of D25 and D29 was because of severe signal overlap. Figure 3.14 shows an example (residue D66) of how the side-chain resonances were assigned in the 2D HCCO spectrum using a combination of NMR experiments. The assigned 2D HCCO spectrum at pH 3.0 is shown in Figure 3.15.

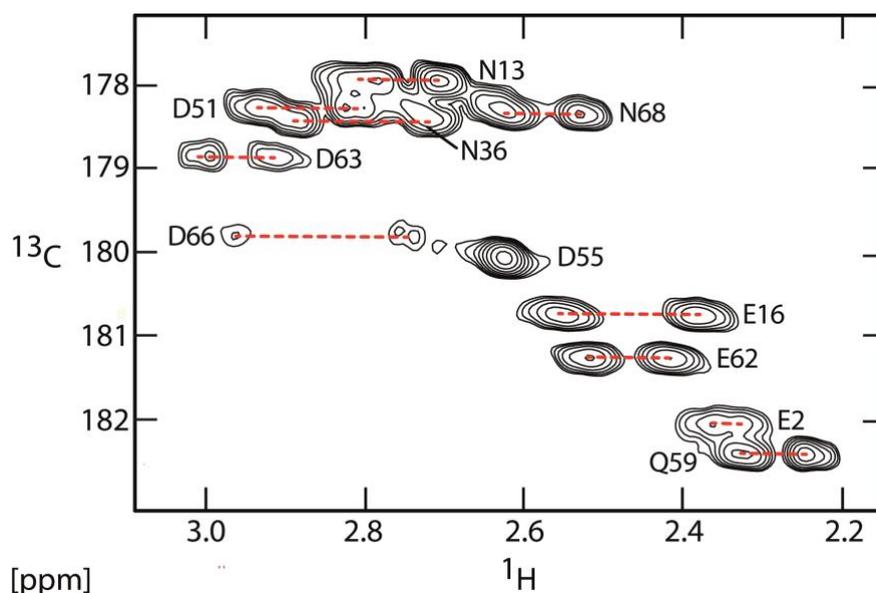


Figure 3.15: Assigned 2D ^1H - ^{13}C HCCO spectrum of APA-1 at pH 3.0 and 1.0 mM.

3.3.4 pKa determination of acidic amino acids and the histidine-75

The pH-dependent chemical shifts of C_γ (Asp) and C_δ (Glu) were monitored by recording a series of 2D HCCO spectra. Unambiguous assignment of the resonances arising from the C_γ nuclei of D25 and D29 over the full pH range examined was not possible. For protonated Asp and Glu residues, the carboxyl ^{13}C resonances are located between 177.8–178.6 ppm and 180.6–181.6 ppm, respectively. Figure 3.16 illustrates that the resonances of the $^{13}\text{C}_\gamma$ and $^{13}\text{C}_\delta$ resonances are downfield-shifted by ~3–5 ppm as the pH was raised, indicating deprotonation of the carboxylic side-chains. No pronounced pH-dependent chemical shift changes were observed for the transaminated Gln and Asn residues, for example Q59 (Figure 3.16). To determine the pKa values, the pH-dependent chemical shifts of the carboxyl ^{13}C resonances were plotted against pH. The titration curves were analysed using a modified Handerson-Hasselbalch equation with the Hill coefficient (Figure 3.17). The estimated pKa values and the corresponding Hill coefficients are presented in Table 3.2 at both APA-1 concentrations examined.

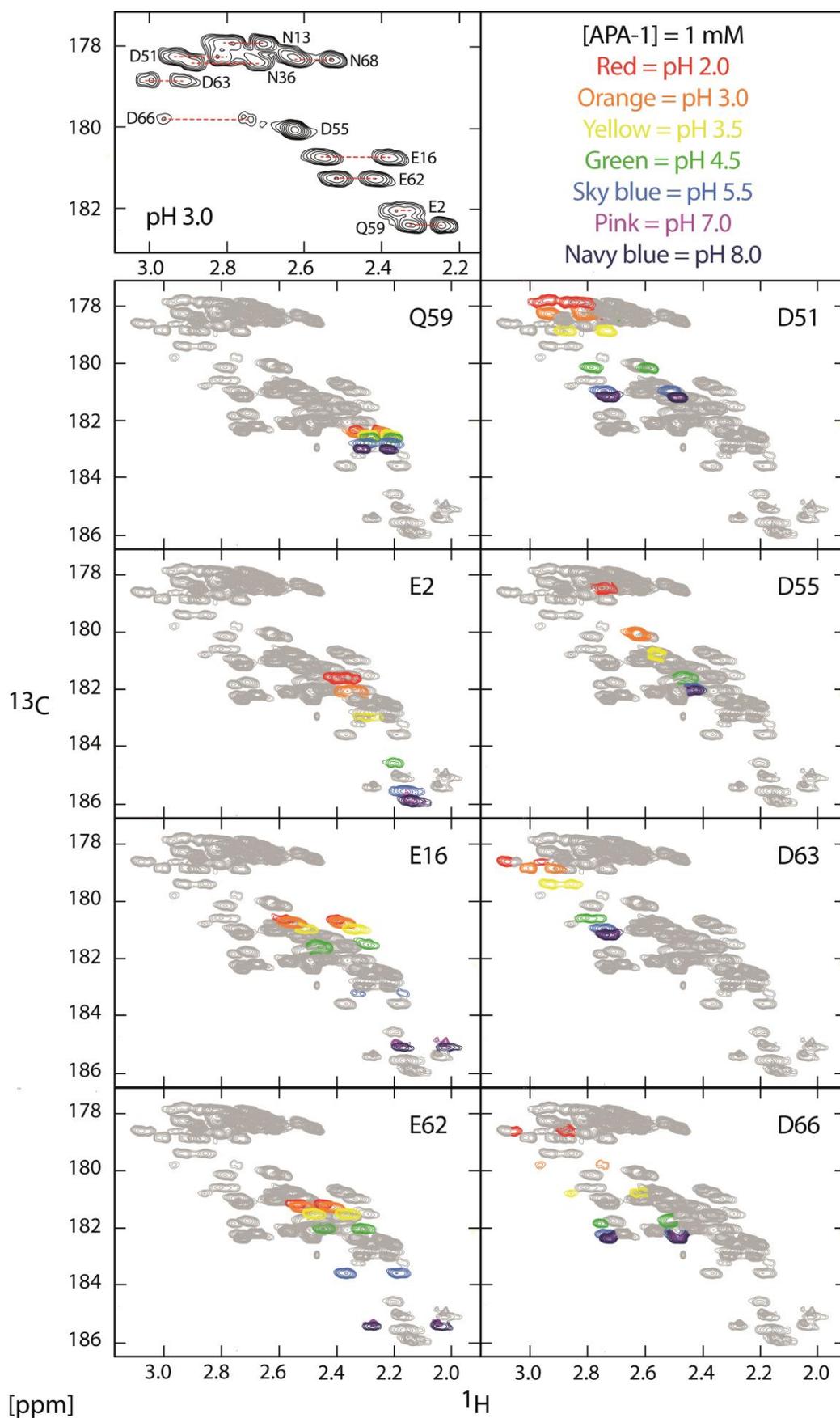


Figure 3.16: The 2D HCCO spectra showing the pH-dependent chemical shift movements of the carboxyl groups of Asp and Glu residues. The first panel represents the assigned 2D HCCO spectrum at pH 3.0. The coloured resonances in each panel represent the carboxyl group chemical shift movements of a particular residue as a function of pH.

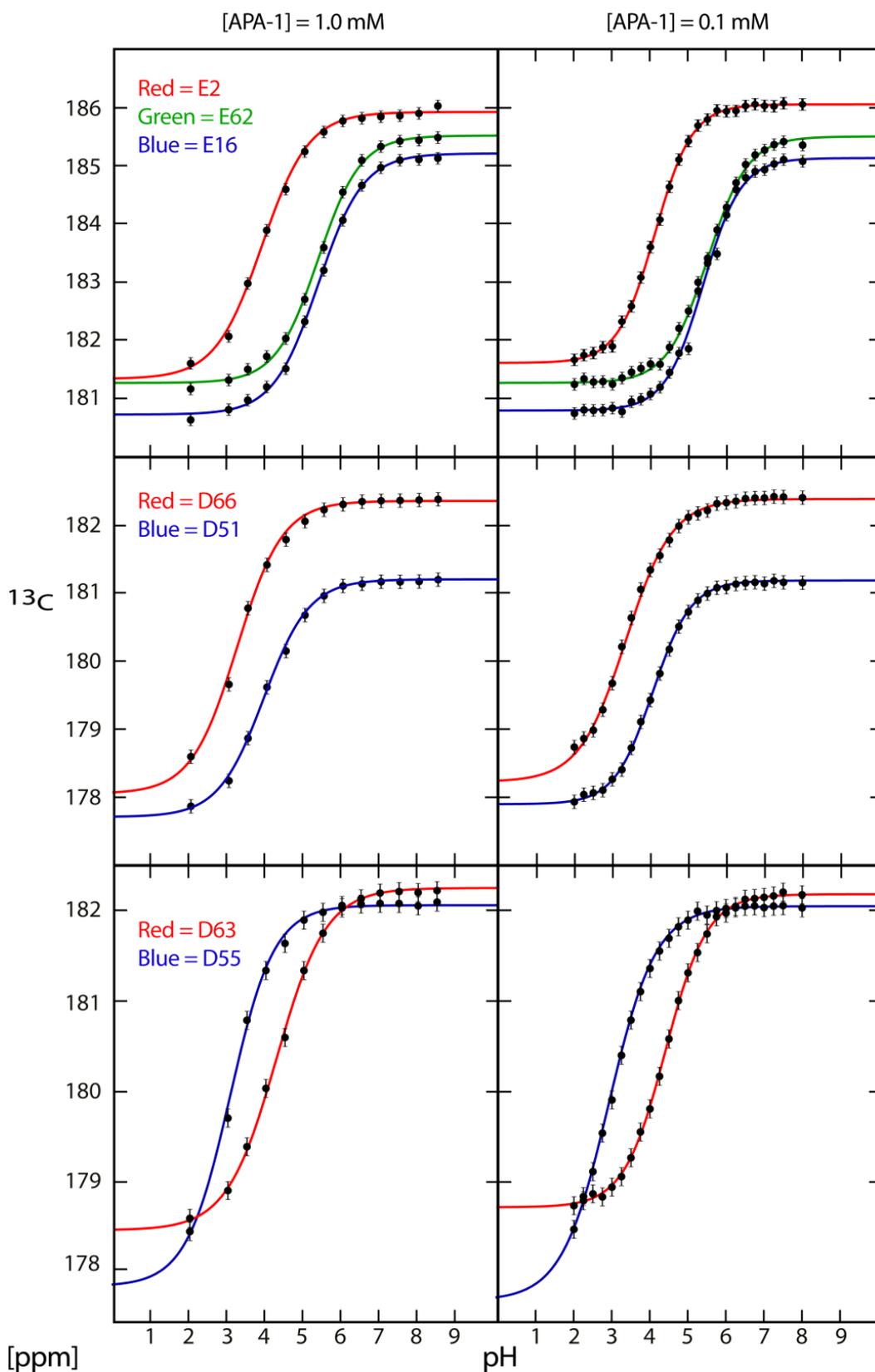


Figure 3.17: The carboxyl ^{13}C chemical shift titration of Glu and Asp residues at APA-1 concentrations of 1.0 (left) and 0.1 mM (right). The data were used for the determination of the pKa values by fitting to the modified Henderson-Hasselbalch equation.

The pKa values of ionised amino acids are affected by the local environment [163-166]. From studies examining the pKa values of ionisable residues in proteins, three general rules have been derived that will be used to make general interpretation of the pKa values derived for APA-1 ionisable residues, as illustrated in Figure 3.18.

(A) Dehydration	
-COOH ↔ -COO ⁻ + H ⁺	↑ nonpolar environment: pKa ↑
-NH ₃ ⁺ ↔ -NH ₂ + H ⁺	↑ nonpolar environment: pKa ↓
(B) Charge-charge interactions	
-COOH ↔ -COO ⁻ + H ⁺	↑ positive charge: pKa ↓
-NH ₃ ⁺ ↔ -NH ₂ + H ⁺	↑ positive charge: pKa ↓
(C) Charge-dipole interactions	
-COOH ↔ -COO ⁻ + H ⁺	↑ hydrogen bonding to protonated form: pKa ↑
-NH ₃ ⁺ ↔ -NH ₂ + H ⁺	↑ hydrogen bonding to protonated form: pKa ↑

Figure 3.18: Three general rules for interpreting pKa values of ionisable groups. (A) The ionisable groups of acidic and basic residues buried in the interior of a protein generally have pKa values above and below the intrinsic values, respectively. This is because the neutral side-chain is favoured in a hydrophobic environment. (B) In a nonpolar environment, the presence of neighbouring charges can affect the pKa of the ionisable groups. For example, the side-chain pKa of acidic residues involved in a salt bridge is usually lowered because the salt bridge interaction favours the amino acid in the negatively charged state. (C) Ionisable groups can form hydrogen bonds with partial charges or dipoles on neighbouring polar groups and form hydrogen bonds. The effect of hydrogen bonding on the pKa value depends on whether a protonated or deprotonated form of the ionisable group is favoured. For example, when an acidic residue is involved in hydrogen bonding and the protonated state is favoured, a higher pKa value is anticipated. The picture is made according to [48].

At an APA-1 concentration of 1.0 mM, the side-chain pKa values of E16 and E62 were calculated to be 5.35 ± 0.04 and 5.35 ± 0.05 , respectively. These values are noticeably higher than the intrinsic pKa value (*i.e.*, 4.2–4.4) [41, 42, 44]. Examination of the APA-1 structure shows that E16 and E62 are in close proximity (Figure 3.19). Moreover, along with D63 and D66, these four acidic amino acids form a potentially negatively charged patch on APA-1. Also present in that region is K64. No major pKa differences for the pKa values of E16, E62, D63 and D66 were observed at a protein concentration of 0.1 mM (Table 3.2).

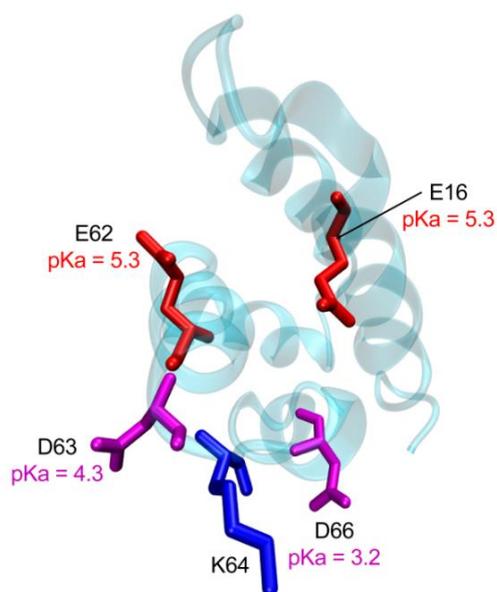


Figure 3.19: Representation of the side-chains of E16, E62, D63, D66 and K64 in the APA-1 structure. The Glu, Asp and Lys residues are coloured in red, magenta and blue, respectively.

The side-chain pKa values of D51 and D55 were calculated to be 3.93 ± 0.07 and 3.09 ± 0.12 , and are in close proximity with K48 and K56, respectively. Based on the rules described in Figure 3.18, these residues may form salt bridges. Unfortunately, the positions of the side-chains are not well defined in the solution structure (Figure 3.20). Nonetheless, the low pKa value of D55 suggests that a salt bridge formed with K56 is plausible. Similar pKa values for D51 and D55 were observed at a protein concentration of 0.1 mM (Table 3.2).

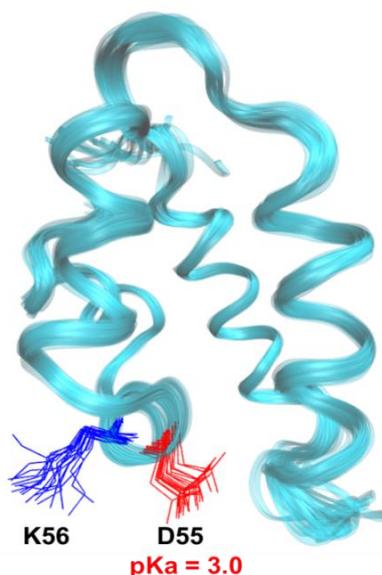


Figure 3.20: Representation of the side-chains of K56 and D55 in the APA-1 structure. The 20 lowest energy minimised structures are presented. The side-chain of K56 is coloured in blue and the side-chain of D55 is coloured in red in each model.

Residue E2, which has been proposed to be involved at the dimer interface in both dimer models (Figure 3.3 and 3.10), has a calculated side-chain pKa value of 3.90 ± 0.06 at an APA-1 concentration of 1.0 mM. Based on the APA-1 solution structure, the side-chain of E2 is solvent exposed, and is located distal from other ionisable residues (Figure 3.21). When the protein concentration was lowered to 0.1 mM, E2 displayed the largest pKa variation among all the ionisable residues, which had increased by 0.22 pH units. The average absolute pKa difference for the rest of the ionisable residues was 0.11. This larger change may be due to weak self-association, and will be discussed in more detail in the discussion.

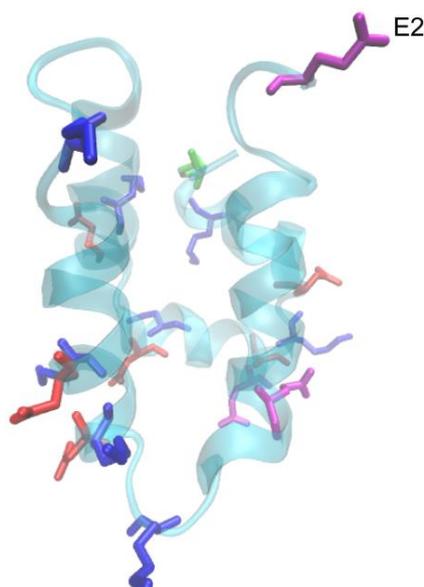


Figure 3.21: Representation of the side-chains of ionisable residues in APA-1 structure. The side-chains of Glu, Asp, Lys and His residues are coloured in magenta, red, blue and green, respectively. Residue E2 is labelled and the position of its side-chain is observed to locate distal from other ionisable residues.

At 1.0 mM, the Hill coefficients for all the ionisable groups were comparable, lying between 0.67–0.76. According to literature [207, 217, 218], Hill coefficients significantly below 1 are considered to indicate negative cooperativity. This could be because of the ionisable state of various side-chains located nearby which titrate over the same range of pH. Such low Hill values were not observed in the APA-1 titrations, indicating minor negative cooperativity due to weak interactions with other titrating groups. At a protein concentration of 0.1 mM, the Hill coefficient for all the carboxyl groups were slightly raised to a range of 0.70–0.89, suggesting less negative cooperativity.

To obtain the C ϵ H assignment of H75, a 2D ^1H - ^{13}C CTHSQC spectrum was recorded at pH 3.0. The C ϵ H resonance was distinguishable and easily assigned. The C δ H resonances were also observed (121.48 (^{13}C) and 7.134 ppm (^1H) at pH 3.0) but were not used for the pH titration. Figure 3.22 shows the chemical shift movements of the $^{13}\text{C}_\epsilon$ resonance as a function of pH. No significant chemical shift movement was observed at pH values less than 5.5 and pronounced changes occurred between pH 6.5 and 8.5. The pK $_a$ of the imidazole ring was determined by plotting the $^{13}\text{C}_\epsilon$ chemical shift changes as a function of pH (Figure 3.23) and was calculated to be 7.29 ± 0.07 at 1.0 mM. A similar pK $_a$ (7.19 ± 0.07) was obtained at a protein concentration of 0.1 mM.

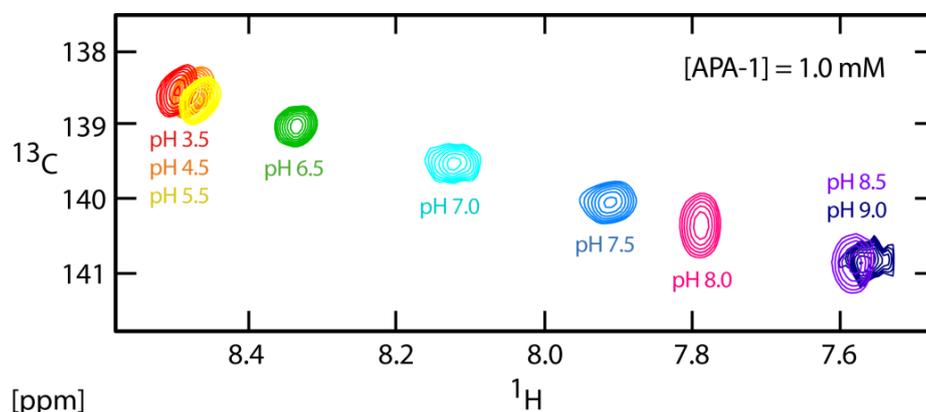


Figure 3. 22: The 2D ^1H - ^{13}C CTHSQC spectrum showing the pH-dependent chemical shift movement of the C ϵ H group of H75 residue at 1.0 mM.

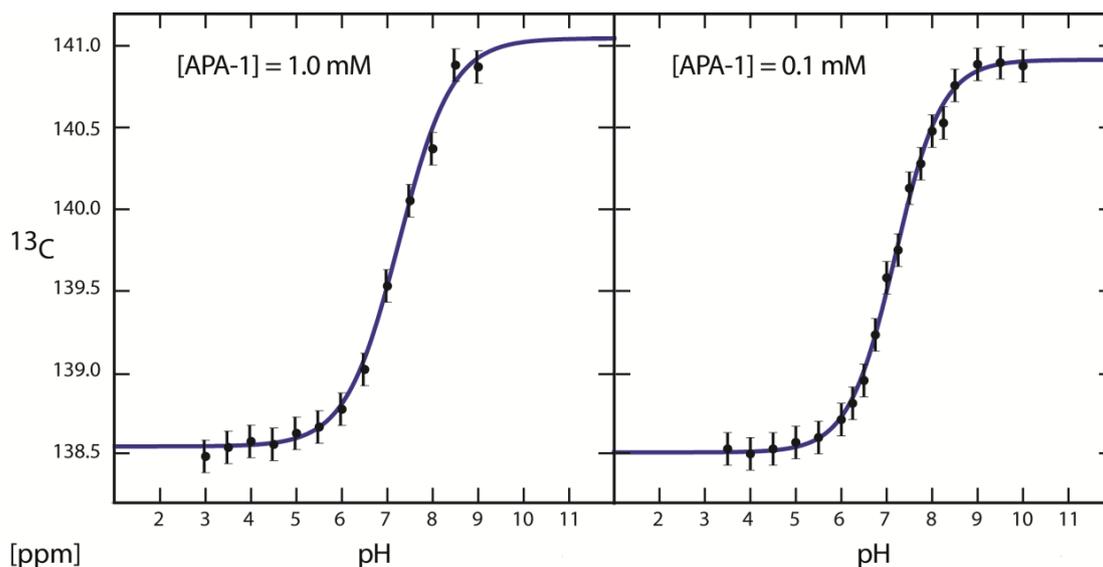


Figure 3.23: pH titration curves of the C ϵ group of H75. The $^{13}\text{C}_\epsilon$ chemical shifts are plotted as a function of pH at APA-1 concentrations of 1.0 (left) and 0.1 mM (right). The data were used for the determination of the side-chain pK $_a$ value by fitting the curves using the modified Henderson-Hasselbalch equation.

Table 3.2: pKa values of acidic groups and the sole His residue of APA-1 determined from 2D ^1H - ^{13}C HCCO and 2D ^1H - ^{13}C CTHSQC experiments, respectively.

Residue	[APA-1] = 1.0 mM			[APA-1] = 0.1 mM		
	pKa ^a	Hill coefficient ^b	$\Delta\delta$ (ppm) ^c	pKa ^d	Hill coefficient ^e	$\Delta\delta$ (ppm) ^f
Glu2	3.90 ± 0.06	0.71 ± 0.05	4.41	4.12 ± 0.03	0.89 ± 0.05	4.38
Glu16	5.35 ± 0.04	0.74 ± 0.06	4.48	5.38 ± 0.03	0.87 ± 0.04	4.35
Asp51	3.93 ± 0.07	0.73 ± 0.07	3.32	4.06 ± 0.04	0.86 ± 0.06	3.23
Asp55	3.09 ± 0.12	0.76 ± 0.10	3.61	2.95 ± 0.12	0.70 ± 0.07	3.58
Glu62	5.35 ± 0.05	0.77 ± 0.06	4.30	5.49 ± 0.03	0.82 ± 0.04	4.12
Asp63	4.27 ± 0.07	0.67 ± 0.06	3.61	4.41 ± 0.04	0.80 ± 0.06	3.45
Asp66	3.25 ± 0.10	0.72 ± 0.09	3.77	3.34 ± 0.07	0.71 ± 0.06	3.70
His75	7.29 ± 0.07	0.74 ± 0.03	2.39	7.19 ± 0.07	0.85 ± 0.10	2.35

pKa^{a,d} and Hill coefficient^{b,e} values were determined from the pH dependent chemical shift of the carboxyl ^{13}C resonances at APA-1 concentrations of 1.0 and 0.1 mM by curve-fitting to the Henderson-Hasselbalch equation with the Hill coefficient (Eq. 3.8). $\Delta\delta^{c,f}$ represents the amplitude of chemical shift differences between protonated and deprotonated forms of individual residues at both concentrations.

3.3.5 pH titration monitored by 2D ^1H - ^{15}N HSQC spectroscopy

The pKa values for the side-chains of residues D25 and D29 were not obtained from the 2D HCCO experiments due to unavailable assignment information. As an alternative, the titrating behaviour of these two residues were obtained using 2D ^1H - ^{15}N HSQC experiments at protein concentration of 0.1 mM (Figure 3.24). The amide ^1H chemical shifts were followed as a function of pH and the titration curves were plotted and fitted using Eq. 3.8 (Figure 3.25). The amide ^1H and ^{15}N chemical shifts are usually sensitive to changes in the local electronic environment due to the high polarisability of the N-H bonds [219]. Nearby titration groups can influence the amide ^1H and ^{15}N chemical shifts and may result in complex titration curves of the amide groups. Therefore, the amide group pKa values do not necessarily reflect (de)protonation of the intra-residue side-chain. Nonetheless, the amide ^1H analysis provides a useful tool for rough estimation of side-chain pKa values

The amide ^1H and side-chain (as determined above) pKa values are listed in Table 3.3. The side-chain pKa of E2 was not obtained due to broadening of the amide ^1H resonance beyond detection above pH 5.0. The pH titration of D51 is not shown because its pH-dependent amide ^1H chemical shift changes were small ($\Delta\delta = 0.05$ ppm) thus leading to large errors when fitted to Eq. 3.8. Most of the amide ^1H titrations behaved as simple titration curves except for D29 and E62, which displayed multiphasic behaviours (Figure 3.25). Titration of other groups located adjacent to D29 and E62 may contribute to this observation, thus the results do not represent the true side-chain titration behaviours of D29 and E62. Consequently, the titration curves of these two residues were not fitted.

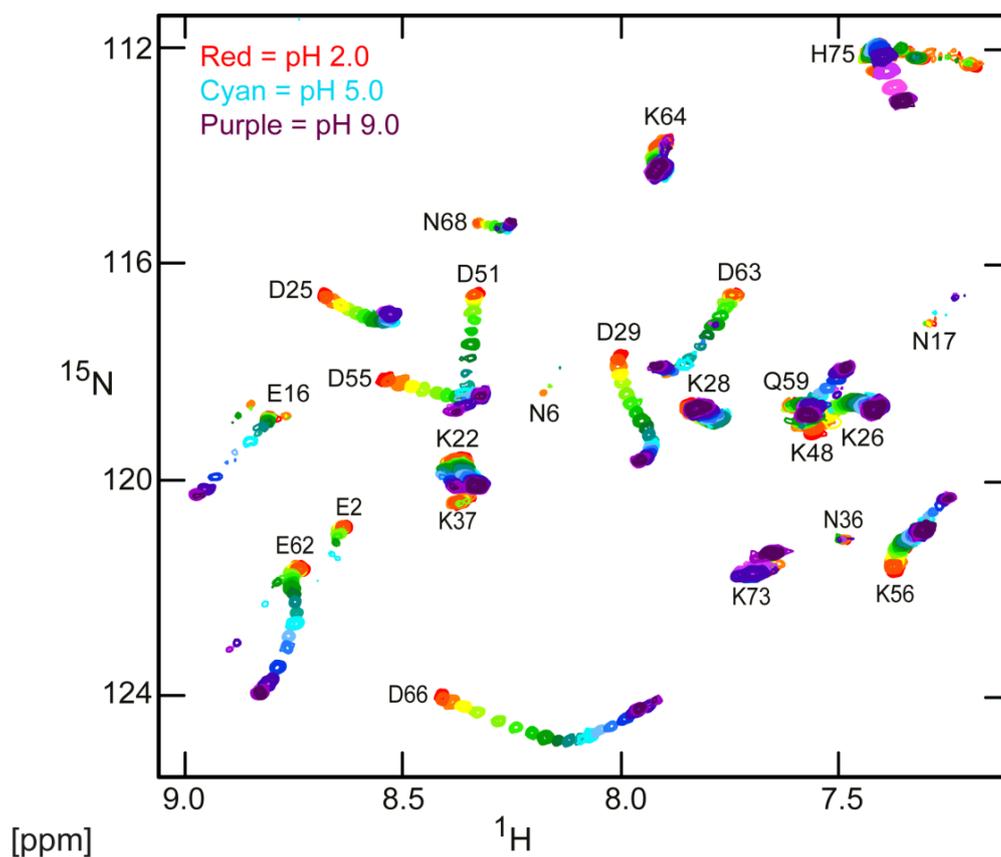


Figure 3.24: Superimposed 2D ^1H - ^{15}N HSQC spectra showing the pH-dependent chemical shift changes of the selectively labelled APA-1 at 0.1 mM. The pH values at which the APA-1 spectra were recorded are represented by the colours of the rainbow ranging from red to purple indicating the lowest pH level (pH 2.0) to the highest pH level (pH 9.0).

Table 3.3: Side-chain pKa values determined from the $^1\text{H}_\text{N}$ and ^{13}C (C_γ and C_δ) nuclei of Glu and Asp residues of APA-1.

Residue	Direct determination of the side-chain pKa using C_γ and C_δ chemical shifts	Indirect determination of the side-chain pKa using the $^1\text{H}_\text{N}$ chemical shifts
Glu16	5.38 ± 0.03	5.34 ± 0.07
Asp25	-	3.12 ± 0.21
Asp29	-	-
Asp55	2.95 ± 0.12	2.93 ± 0.17
Glu62	5.49 ± 0.03	-
Asp63	4.41 ± 0.04	4.59 ± 0.08
Asp66	3.34 ± 0.07	3.62 ± 0.09

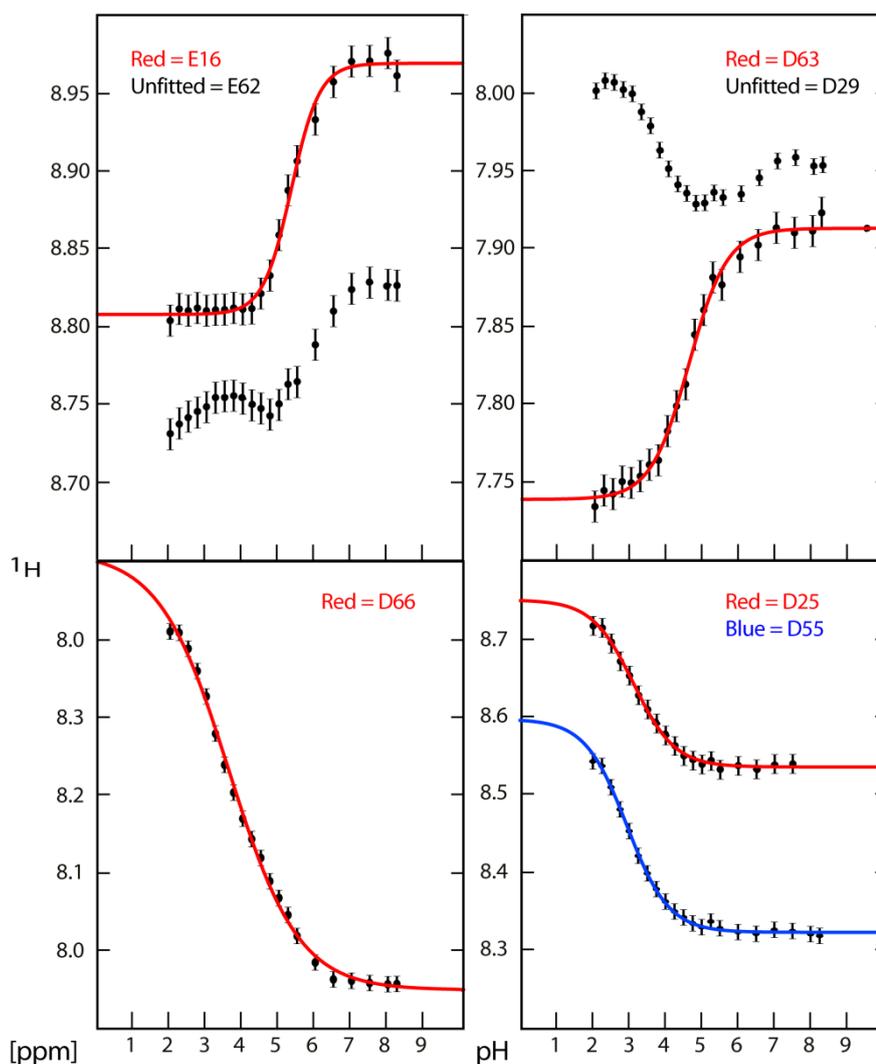


Figure 3.25: The amide ^1H group pH titration curves of Asp and Glu residues. The amide ^1H chemical shifts were plotted as a function of pH at a protein concentration of 0.1 mM. Titrations which showed monophasic behaviours were used for side-chain pKa estimations. No fitting was performed for E62 and D29 which exhibited multiphasic titration behaviours.

3.3.6 Estimation of K_d for amoebapore A dimerisation

To investigate the K_d of APA-1 dimerisation, 2D ^1H - ^{15}N HSQC spectra were acquired at eight protein concentrations ranging from 25–850 μM at pH 3.0. Resonances corresponding to E2 and K37 showed the largest chemical shift change and line-broadening (Figure 3.26). The resonance representing residue K37 was broadened to such an extent that it was no longer detectable above 250 μM . Therefore, only the ^{15}N chemical shifts of E2 were followed as a function of protein concentration for the K_d study (Figure 3.27). Attempts were made to fit the data to the monomer-dimer equilibrium model (Eq. 3.12), assuming the monomer-dimer equilibrium was in the fast-exchange regime on the NMR chemical shift time scale. However,

REST.PE.COM
List of research project topics and materials

correct data fitting was not possible, and therefore, a true K_d value was not obtained (data not shown). The result suggest that the K_d of APA-1 dimerisation is beyond the highest concentration studied (*i.e.*, $K_d > 850 \mu\text{M}$). Higher concentrations were not attempted because of limited sample availability, possible non-specific protein aggregation [210, 220] and severe line-broadening of the resonance arising from E2. The study was not performed at pH 5.2 because resonances corresponding to E2 and K37 are not visible at pH 5.2 (Section 3.3.2).

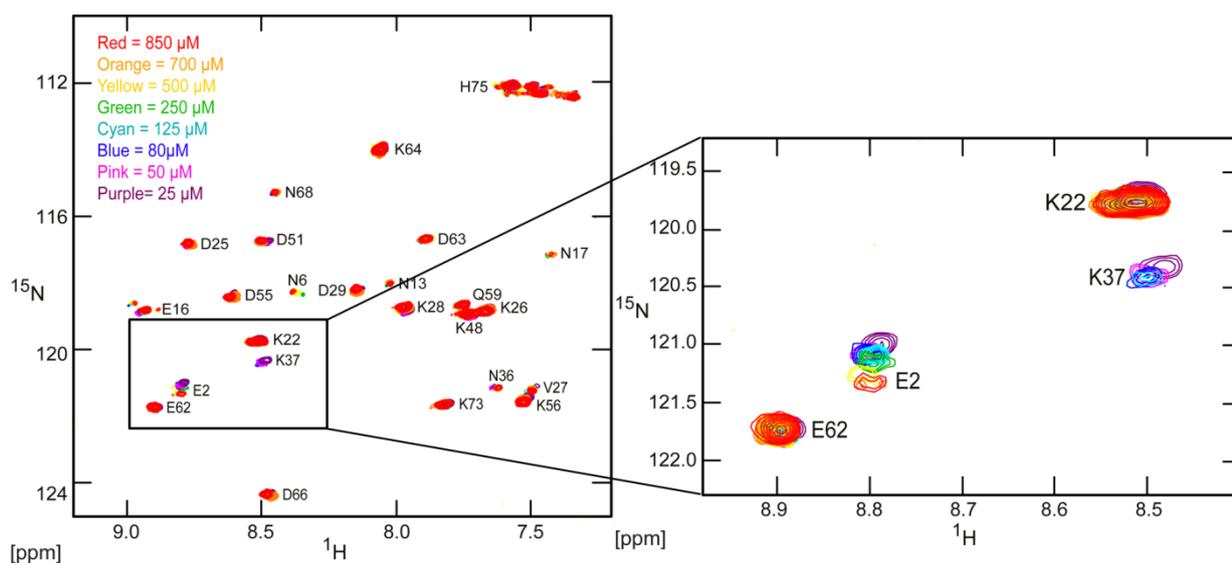


Figure 3.26: Superimposed 2D ^1H - ^{15}N HSQC spectra recorded at various APA-1 concentrations at pH 3.0. Chemical shift changes and line-broadening of resonances corresponding to E2 and K37 were observed, as presented in the selected spectral region (right).

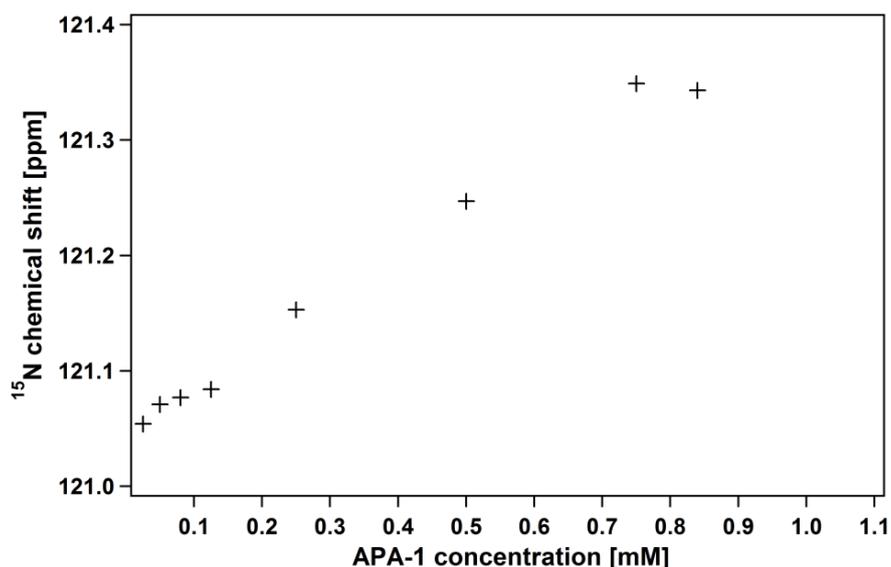


Figure 3.27: Titration of amide ^{15}N resonances of residue E2 plotted against the protein concentration at pH 3.0. The last data point is an outlier because the signal intensity was very weak (at the noise threshold) and correct assignment of this resonance was challenging.

To K_d of APA-1 dimerisation was further studied at pH 5.2 using analytical SEC with APA-1 concentrations ranging between 1.5 μM to 3.0 mM. Four chromatograms are represented in Figure 3.28 as examples. At very low protein concentration (1.2 μM) APA-1 eluted from the column as a single peak at a retention time corresponding to the MW of the monomer species. Upon raising the protein concentration, an additional peak was eluted at a retention time corresponding to the APA-1 dimer and the peak area of this species increased as the concentration increased, indicating an increase in the dimer population. The monomer-dimer equilibrium was reached above 190 μM , as reflected by small variations of the ratios between monomer and dimer peak areas. No pronounced tailing of the peaks and shift of elution volumes was observed upon increasing the protein concentration, suggesting that the exchange rate between monomers and dimers was slow on a timescale of minutes (SEC experiment time). No higher order oligomers were detected. To exclude the dilution factor on the K_d , four protein concentrations (*i.e.*, 12, 24, 47, and 95 μM , where monomers and dimers co-existed) were used to estimate the K_d (Eq. 3.9). The K_d was estimated at each concentration and the average value was $247 \pm 32 \mu\text{M}$, indicating stronger dimer association at pH 5.2 than 3.0 ($K_d > 850 \mu\text{M}$).

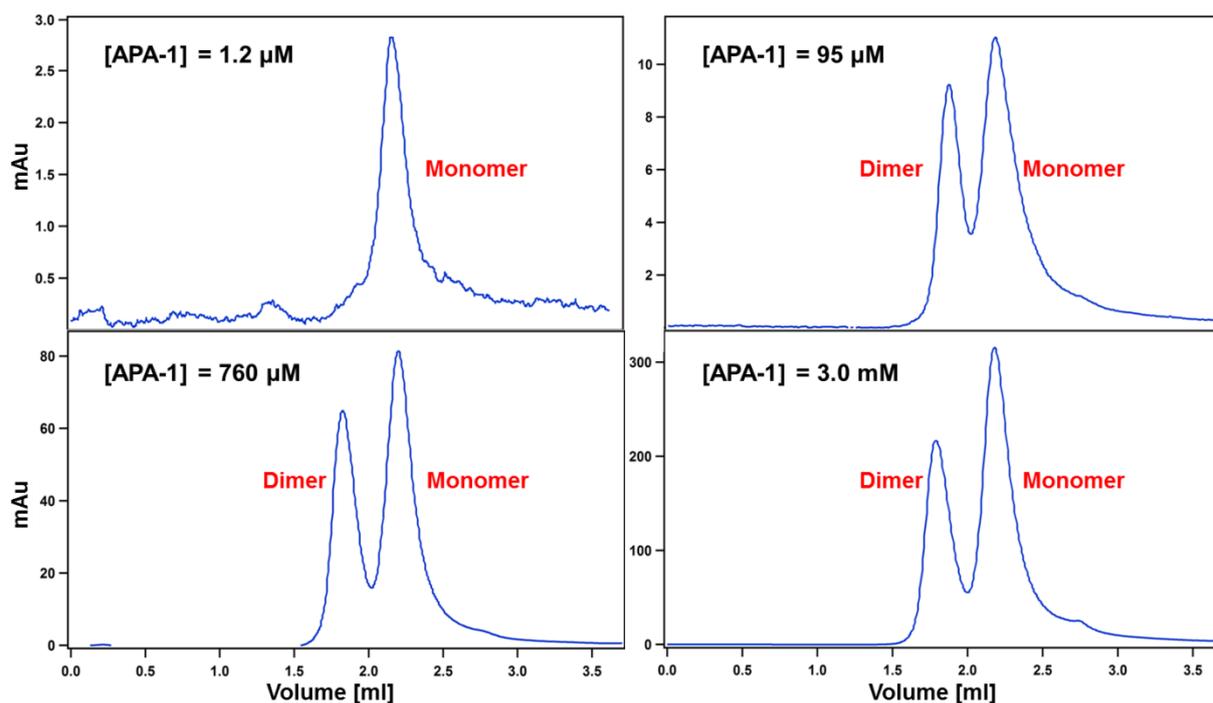


Figure 3.28: The SEC profiles representing the concentration-dependent behaviour of APA-1 dimerisation at pH 5.2. Chromatograms obtained at protein concentrations of 1.2, 95, 760 and 3000 μM are presented. The protein elutes primarily as monomer at low concentrations and as a mixture of monomer and dimer at higher concentrations. The slight variation of the hydrodynamic volumes is responsible for the minor shift of retention volumes of the monomeric and dimeric APA-1 under slightly different buffer conditions. The absorption was measured at 280 nm.

3.3.7 Small angle X-ray scattering

The SAXS experiments were designed to resolve the gross shape of APA-1 in solution as a function of APA-1 concentration. Two sets of SAXS experiments were performed: one using a bench top SAXS instrument at ANSTO and the second one was carried out at the Australian Synchrotron, which records data with significantly higher sensitivity. Due to the limited instrument sensitivity at ANSTO, the SAXS experiments were performed over a relatively high protein concentration, ranging from 550 μM to 2.1 mM at pH values of 3.0 and 5.2. Experimental SAXS curves of APA-1 at the various concentrations were fitted using an elliptical cylinder shape structure factor [221] because this shape most closely matched the dimer models presented earlier. From the SAXS data (Figure 3.29), clear differences were observed between the scattering profiles at pH 3.0 (low dimer population) and 5.2 (high dimer population), even at the lowest concentrations, implying that even lower concentrations were necessary to achieve scattering from a monomer only solution. This is in agreement with the result obtained from the SEC experiments.

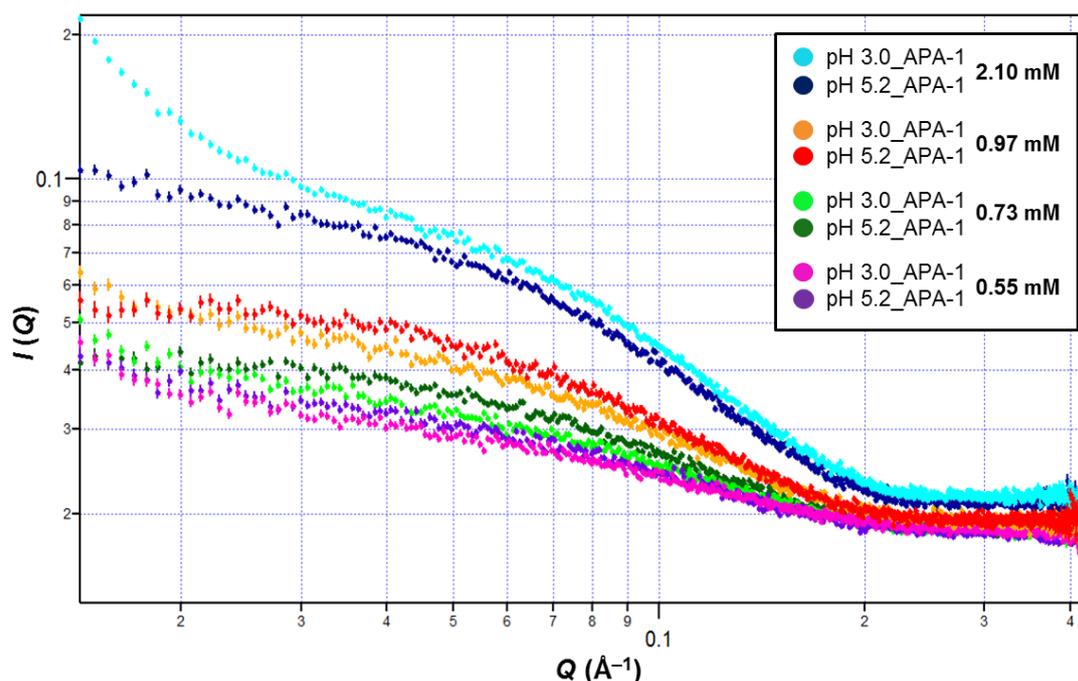


Figure 3.29: Experimental SAXS curves collected using native APA-1 in solution at different concentrations. The protein is expected to exist as a mixed population of dimers and monomers, with a larger dimer ratio at higher concentrations and pH 5.2 where APA-1 is most active.

The experimental SAXS data of APA-1 at pH 3.0 showed an upturn in intensity in the low Q range, indicating the presence of micro-aggregation. Consequently, data at pH 3.0 was not fitted. The pH 5.2 experimental SAXS data at concentrations of 2.1 and 0.55 mM also had

minor aggregation, thus the data was fitted over a narrower Q range (Figure 3.30). Based on the elliptical cylinder model, at the highest protein concentration and pH 5.2, a model with an $R_g = \sim 13 \text{ \AA}$ and $D_{\max} = \sim 57 \text{ \AA}$ (with the hydration shell) was generated from the SAXS data. This SAXS-derived model is similar to the dimer model proposed in Section 3.3.2 which has a radius of 11.5 \AA and a length of 49 \AA (without the hydration shell) (the two dimer models are presented in Figure 3.31).

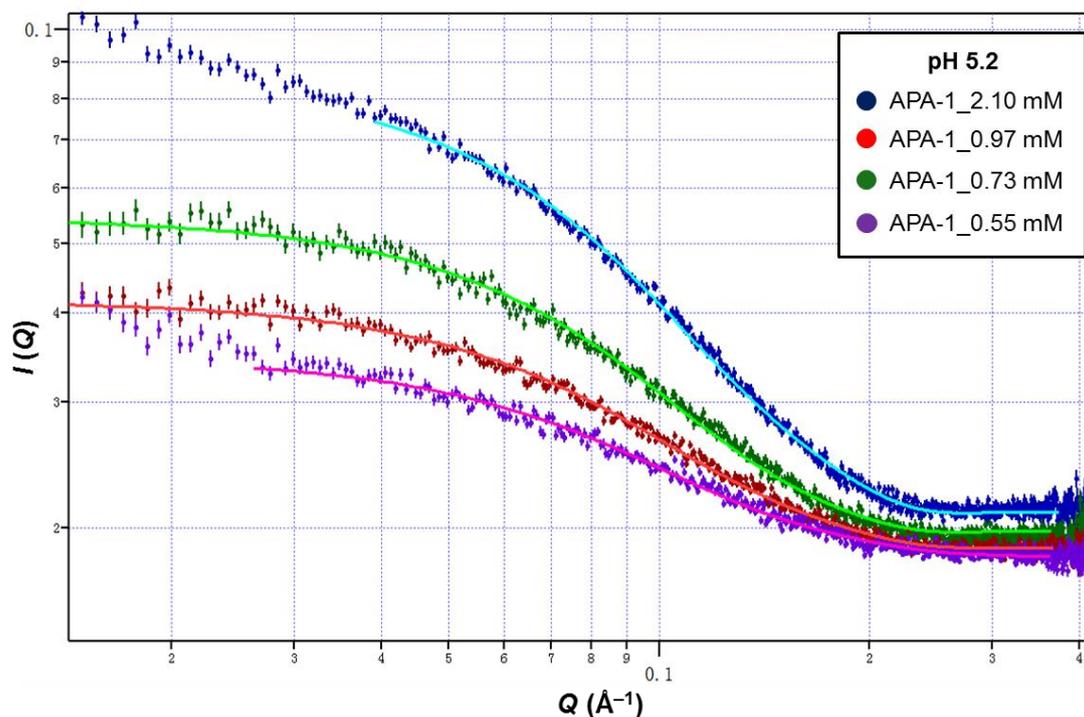


Figure 3.30: Experimental SAXS curves (solid circles) of APA-1 at various concentrations and corresponding fitted curves (solid lines) using the elliptical shape structure factor. The fitting was achieved using IGor Pro (WaveMetrics, Inc, Lake Oswego, USA).

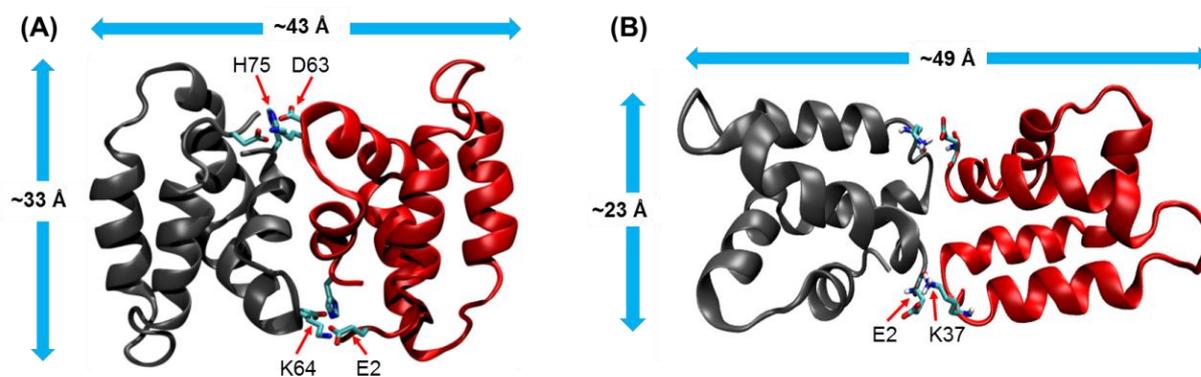


Figure 3.31: Two proposed dimer models of APA-1. (A) Two ion pairs are formed at the dimer interface, H75–D63, K64–E2, forming a dimer model with dimensions of ~ 43 and $\sim 33 \text{ \AA}$, respectively. (B) Residue K37 forms an ion interaction with E2, giving rise to a dimer model with dimensions of ~ 49 and $\sim 23 \text{ \AA}$, respectively.

Experimental SAXS curves were fitted by minimising the χ^2 values to the theoretical scattering profiles calculated from the monomeric and proposed APA-1 dimer structures using AXES [193]. AXES functions by calculating the best fit between the experimental scattering data and a structure model by varying a number of parameters, including the hydration shell. By using the χ^2 values and F -test (Table 3.4), the results showed that the theoretical curves of the new dimer model (Figure 3.31B) fit the experimental data statistically better than the previously proposed model (Figure 3.31A) at the highest concentration of 2.1 mM with $P_F = 1.796 \text{ e}^{-09}$ ($F = 2.68$, $\nu_1 = 579$, $\nu_2 = 579$) (Table 3.4). Although lower P_F values were obtained, the fitting of the data at the lower protein concentrations still supported the result observed at the highest concentration (*i.e.*, at the $P_F = 0.05$ level). The data did not give a good fit to the monomer structure as indicated by the large χ^2 values, and the P_F values at all concentrations was significant (Table 3.4). The results show that APA-1 exists as a dimer-monomer mixture and at these concentrations and the dimer is the dominant species. The SAXS results are in agreement with the K_d value obtained from the SEC experiment (Section 3.3.6). In addition, the R_g and D_{\max} values of the experimental data were determined to be over a range of 12–14.6 and 55–58.7 Å, respectively, using GNOM [184] (Table 3.4). Similar to the results obtained from the fittings using the elliptical cylinder model, the R_g tended to decrease as the protein concentration decreased. This is an indicator of increasing population of monomeric APA-1 which has dimensions of approximately 22 and 30 Å.

Table 3.4: Calculated parameters for SAXS data of APA-1 at different concentrations, pH 5.2

Concentration (mM)	GNOM		AXES $d_f=579; P_F=0.05$				
	R_g (Å)	D_{max} (Å)	Reduced χ^2 (model A)	Reduced χ^2 (model B)	Reduced χ^2 (monomer)	P_F ($F; d_A, d_B$)	P_F ($F; m, d_B$)
2.1	14.6	58.7	3.212	1.961	37.34	$1.796 e^{-09}$	$1.70 e^{-211}$
0.97	13.8	55.4	1.272	1.157	11.23	0.128	$1.82 e^{-138}$
0.73	13.0	55.0	1.894	1.496	9.875	0.002	$8.00 e^{-101}$
0.55	12.0	57.9	3.087	2.473	6.914	0.003	$4.91 e^{-34}$

d_f represents the degrees of freedom. Reduced χ^2 is calculated from (No. of data points / d_f) $\times \chi_{(AXES)}^2$ [222];

$P_F(F; \nu_A, \nu_B)$ represents the probability generated from integration of F distribution ($F = \chi_{\nu_A}^2 / \chi_{\nu_B}^2$) to analyse the differences between two fits [222];

ν_A, ν_B are the d_f of the two fitting profiles;

d_A represents the d_f of dimer model A (Figure 3.32A); d_B represents the d_f of dimer model B (Figure 3.32B); and m represents the d_f of APA-1 monomer.

Attempts to fit the ANSTO data using GASBOR failed, possibly because the quality of the data was not sufficient for obtaining reasonable fits. The SAXS instrument at the Australian Synchrotron with higher sensitivity was suitable for collecting data at lower protein concentrations. DAMMIN and GASBOR were used to generate *ab initio* dummy atom models from the experimental scatterings recorded at a protein concentration of 190 μ M, pH 3.0. The 15 generated dummy atom models fit the SAXS data with χ^2 values ranging between 0.61–0.64. The dummy atom models were averaged using DAMAVER and the *ab initio* molecular shape superimposed with the two proposed dimer models and the monomeric APA-1 structure. It was observed that the *ab initio* model matched the structure of monomeric APA-1 (Figure 3.32), but not the dimer models. This indicates that at this concentration, the majority of APA-1 exists as a monomer. Superimposition of the average molecular envelope with monomeric APA-1 shows that the volume calculated from the SAXS data is larger than the NMR structure and a small lobe section between helices 2 and 3 protrudes from the envelope. This suggest that at this concentration, a small population of APA-1 dimer may exist which caused minor overestimation of the size of monomeric APA-1. Dummy atom models at lower protein concentrations were not examined because SAXS data quality was poor.

Concentrations higher than 190 μM were not examined because the amount of protein material was limited.

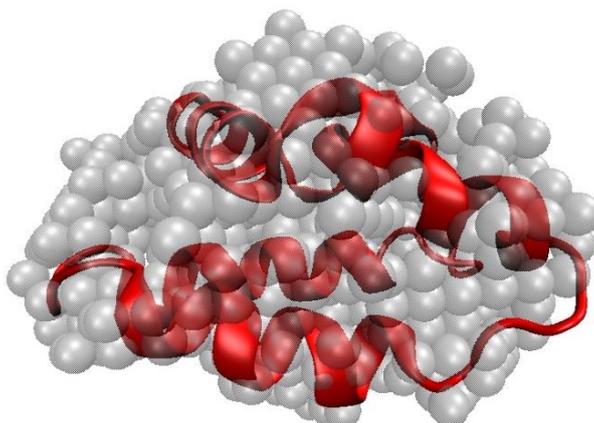


Figure 3.32: Ensemble averaged NMR-derived APA-1 structure superimposed with the molecular envelope generated from the ensemble of 15 GASBOR models using DAMAVER and DAMFILT.

3.3.8 Cell-free synthesis of amoebapore A mutants

APA-1 mutants that are postulated to stabilise/destabilise the dimer models were designed. E2Q, E2A, D63N, K64Q, H75Q, H75K are designed to disrupt the old dimer model (Figure 3.32A) whereas E2Q, E2A, E2Q-K37Q, K37A, K37Q and F41A are designed to disrupt the newly dimer model (Figure 3.32B). Expressions of the mutants were performed on the analytical scale to test protein expression. IL6+ expression was used as the control. The SDS-PAGE analysis showed that APA-1 mutants migrated differently on the SDS-PAGE (Figure 3.33). This gel shift phenomenon occurs commonly for membrane binding proteins [144]. Single mutation may alter the protein-SDS binding by changing the residue size, charge, hydrophathy and conformation of the unfolded protein in an SDS gel, resulting in unusual migration rates [144]. The expressions of the mutants were performed using 1/10 ml preparative scale cell-free reactions. Both uniformly ^{15}N -labelled and unlabelled mutants were produced.

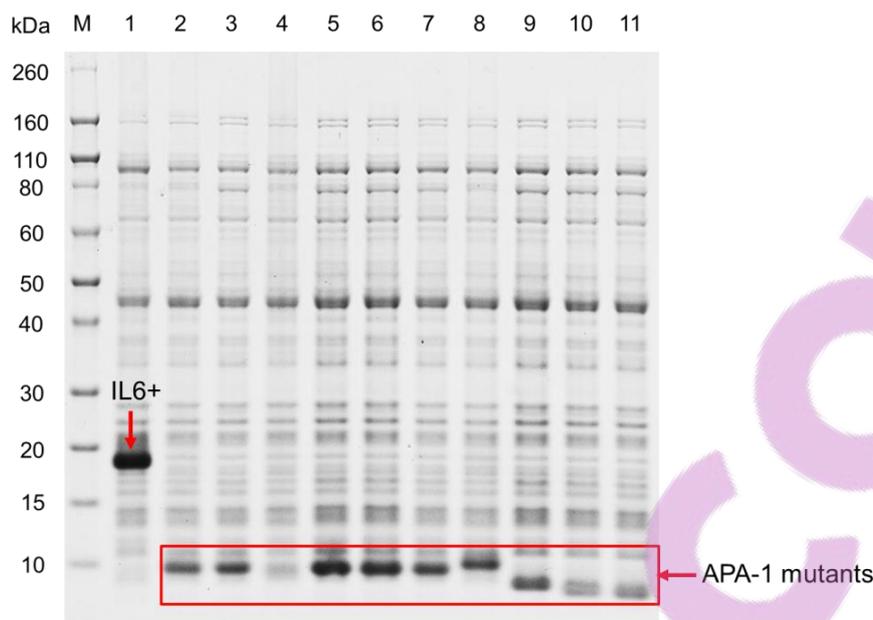
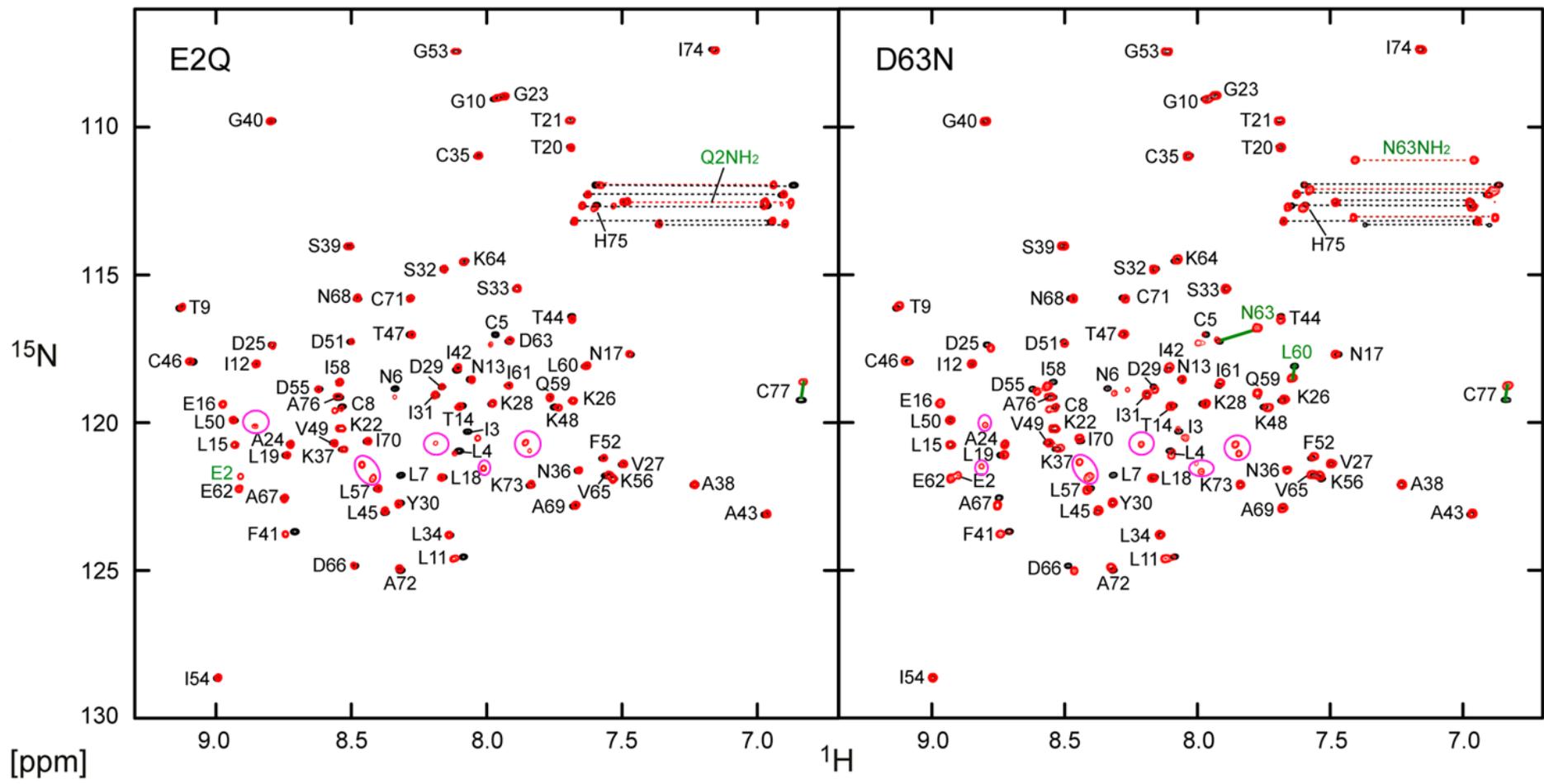
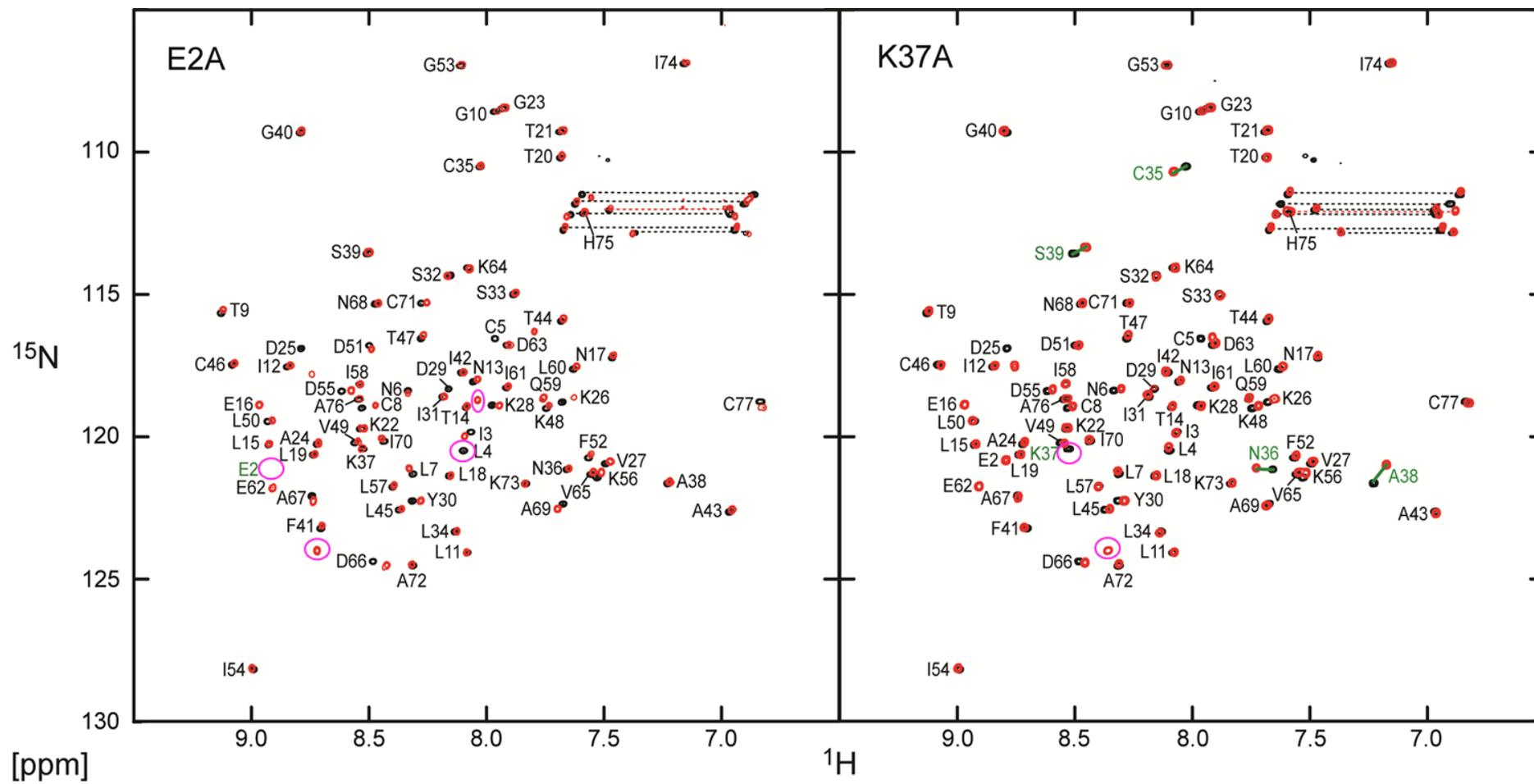


Figure 3.33: 4–12% SDS-PAGE analysis of the expressions of APA-1 mutants. Lane M: Precision plus MW marker (BioRad, Hercules, USA); lane 1: IL6+ control; lanes 2–11: expressions of uniformly ^{15}N -labelled E2Q, D63N, H75K, H75Q, K64Q, E2A, K37A, K37Q, F41A and E2Q-K37Q. The arrows indicate the positions of protein bands corresponding to IL6+ and the bands representing APA-1 mutants are boxed.

The cell-free expressed mutants were $(\text{His})_6$ -tag digested using Factor Xa. It was challenging to obtain full digestion of the mutants because (1) the Factor Xa purchased from Qiagen (QIAGEN, Limburg, Netherlands) was used, which was later discovered to have low digestion efficiency as described in Section 2.3.3.2 and (2) the mutations may have reduced the Factor Xa efficiency when compared with the efficiency of the digestion of recombinant APA-1. As a result, only partially digested mutants were achieved for the majority of the mutants. Consequently, bioassays were not performed using these protein products. Nonetheless, the 2D ^1H - ^{15}N HSQC spectra of the partially digested mutants acquired at pH 3.0 showed a similar overall resonance pattern when compared with the recombinant APA-1. This indicates no major difference in the protein fold between the recombinant APA-1 and the APA-1 mutants produced. Minor chemical shifts were observed for resonances associated with residues adjacent to the mutation site (Figure 3.34). Purification of unlabelled mutants gave rise to four peaks in the HPLC chromatograms, and each peak was collected and examined by mass spectrometry in an effort to understand the digestion pattern by Factor Xa. An example is presented in Figure 3.35.





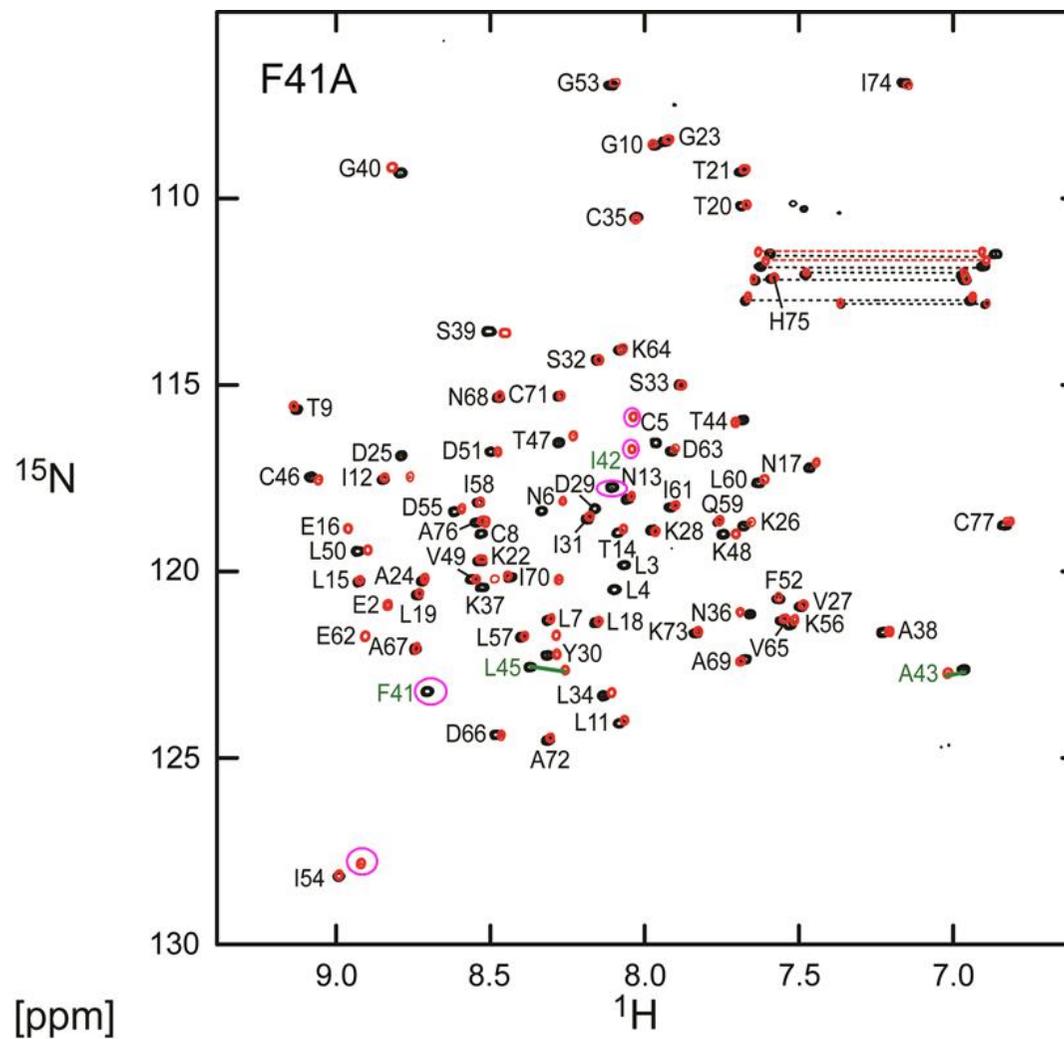


Figure 3.34: Overlaid 2D ^1H - ^{15}N HSQC spectra of APA-1 mutants (red) and the fully digested recombinant APA-1 (black). All spectra were recorded at pH 3.0. The spectra of the mutants were assigned based on the 2D ^1H - ^{15}N HSQC assignments of the recombinant APA-1. Unassigned resonances are likely to arise from the undigested (His)₆-tag or mutation (circled in magenta). Resonances arising from residues that are located adjacent to the mutation sites usually showed chemical shift changes and are assigned in green.

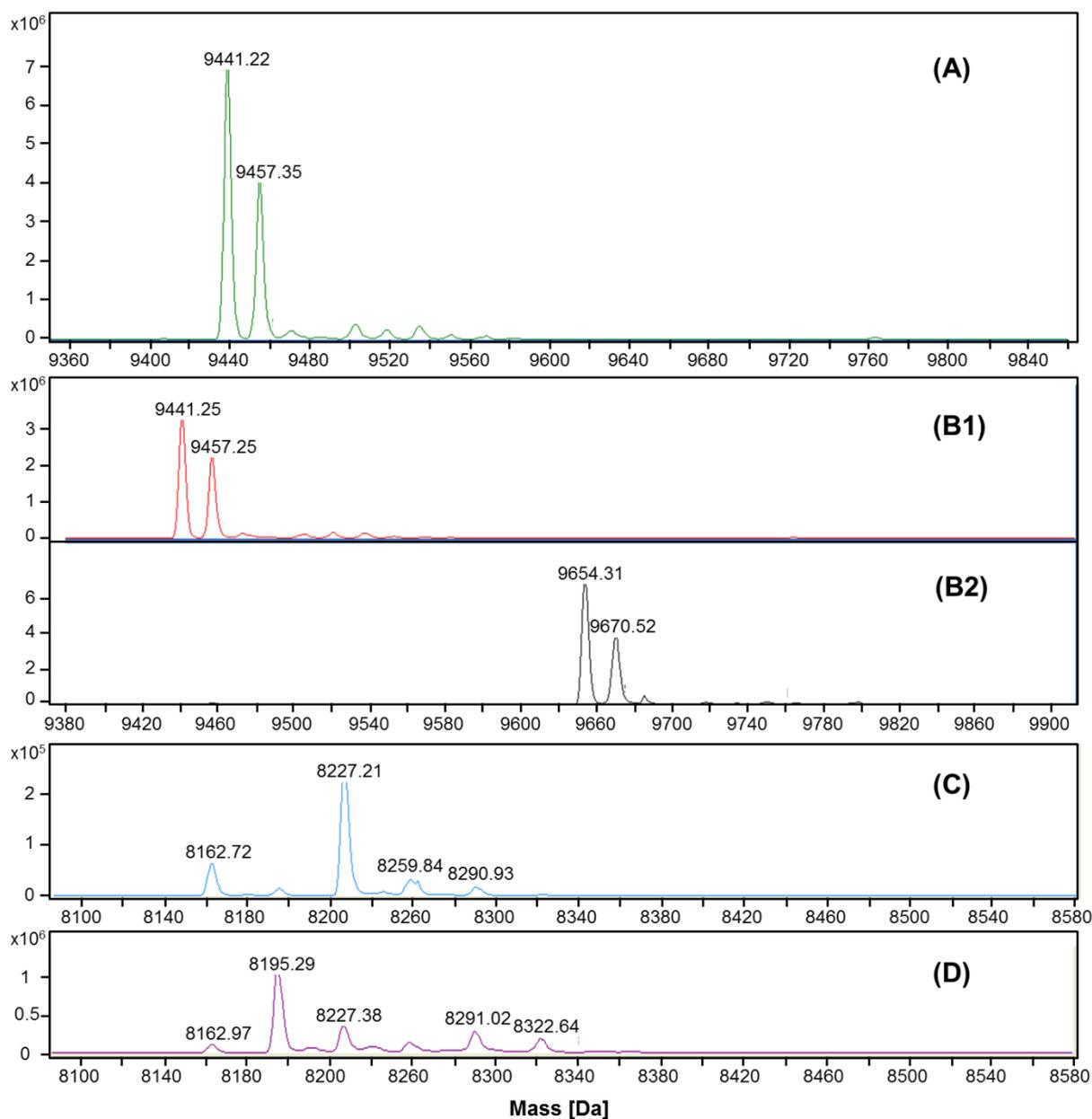


Figure 3.35: Mass spectroscopy profiles presenting the results of Factor Xa digestion of mutant F41A. Panel (A)–(D) represent the mass spectroscopy profiles corresponding to the species contained in the four peaks eluted from the RP-HPLC column. Interpretation of the partially digested F41A species was carried out; however, a full interpretation was not achieved. The species presented in (A), (B1) and (B2) are partially digested F41A with MW higher than the fully digested species which would have a mass of 8244 assuming formation of the three disulfide bonds. The species presented in (C) and (D) have MW closer to the expected MW of the fully digested APA-1. Similar mass spectroscopy complexity was observed for the other mutants (data not shown).

3.4 Discussion

3.4.1 pH- and concentration-dependent amoebapore A dimerisation

Dimerisation of APA-1 was investigated using NMR spectroscopy. Both 1D ^1H and 2D ^1H - ^{15}N HSQC NMR experiments showed that APA-1 dimerisation is a pH- and concentration-dependent event. A number of resonances were observed to show significant line-broadening and chemical shift changes upon pH and concentration changes, suggesting that the dimer-monomer equilibrium is in the slow to intermediate exchange regime on the NMR chemical shift time scale. Studies examining the dimerisation of rat CD2 have also reported similar observations [168], and therefore the resonances that showed line-broadening represent residues located at the dimer interface. Note that while line-broadening is hypothesised to be linked to protein dimerisation, it is also possible that intrinsic slow motion dynamics arise for these residues under particular conditions; albeit, this is unlikely given the pH- and concentration- dependent line-broadening observations coupled with previous SEC research [24]. No major conformational change was observed during APA-1 dimerisation, as confirmed by no significant change in the 2D ^1H - ^{15}N HSQC spectrum. The monomer structure combined with chemical shift changes were used to build a new model of the APA-1 dimer with the docking programme HADDOCK [215]. The new dimer model places juxtaposition K37 and E2 on opposing monomers. Such residue positioning enables the formation of a salt bridge between these two residues, and the dimer interface is also stabilised by pi-stacking interactions between F41 residues on each monomer (Figure 3.10). In addition, based on the side-chain pKa study, residue E2 has a calculated side-chain pKa value of 3.90 ± 0.06 at an APA-1 concentration of 1.0 mM. The APA-1 solution structure indicates that the side-chain of E2 is solvent exposed and located away from other ionisable residues. When the protein concentration was lowered to 0.1 mM, E2 displayed the largest side-chain pKa variation compared with the other ionisable residues studied (*i.e.*, increased by 0.22 pH units). Previous studies on CD2 have shown that anomalous, yet small, pKa variations can be correlated with weak protein dimerisation [157]. A Glu residue located at the dimer interface displayed a pKa variation of 0.38 pH units attributed to a protein concentration effect, suggesting self-association may account for such variation [157, 168]. Consequently, the concentration-dependent shift of the side-chain pKa of E2 suggests a weak dependence of the pKa on protein concentration. This dependence appears to suggest E2 is located at the dimer interface, in agreement with the dimer model, and the interaction of APA-1 monomers is weak with a K_d

value at pH 5.2 determined to be $247 \pm 32 \mu\text{M}$ and a significantly lower pKa at pH 3 of $> 850 \mu\text{M}$. In addition, the resonance in the HSQC arising from K37, showed the largest chemical shift perturbation and the strongest line-broadening as a function of protein concentration, supporting the role of this residue at the dimer interface.

The interface areas of the two dimer models were calculated using a programme “Protein Interfaces, Surfaces and Assemblies” (PDBePISA) [223]. Interface areas obtained for the original and new dimer models were estimated to be 583.2 and 441.4 \AA^2 and represent 5.9% and 4.4% of the total surface areas of the dimer models, respectively. Both interface areas are significantly lower than the standard-size protein interfaces (*i.e.*, 1200–2000 \AA^2) as well as the cut-off interface area (600 \AA^2) for an energetically favourable interaction. This further supports the results presented herein that APA-1 self- association is weak [224].

The SAXS experiments were performed to investigate the gross shape of APA-1 in solution. At the lowest concentration studied (190 μM), DAMMIN and GASBOR were successfully used to generate an *ab initio* dummy atom model of the APA-1 monomer, suggesting at this concentration and at pH 3.0, the protein behaves predominantly as a monomer. This is in agreement with the NMR T_2 measurement (Section 3.3.1) and the K_d estimation using NMR spectroscopy (Section 3.3.6). In addition, the SEC and the SAXS data recorded at 5.2 and concentrations ranging between 1.5 μM to 3.0 mM and 550 μM and 2.1 mM, respectively, clearly indicated that APA-1 exists in a monomer-dimer equilibrium. Using Eq. 3.18 and excluding any dilution effect on the SEC column, the APA-1 dimer population at the highest concentration studied in the SAXS experiment (2.1 mM) was determined to be 79.2%. This number is in agreement with the modelling of the SAXS data using Igor and AXES, and the SAXS data analysis was found to fit statistically better to the new dimer model. A previous SEC study on pH-dependent APA-1 dimerisation [24] has shown that APA-1 existed exclusively as dimers at pH 5.2 and primarily as a monomer at pH 8.5, whereas the NMR T_2 measurements showed that APA-1 (210 μM) appears to have a similar monomer-dimer equilibrium at pH 5.2 and 8.5 because the T_2 value at these pHs were similar. It is likely that this difference in observations at the high pH may arise from the different protein concentrations used in each study, as well as the dilution of the protein concentration on the preparative-scale size exclusion column leading to a shift in the equilibrium to the monomer species.

Based on the previous dimer model (Figure 3.32A), a hexameric pore model has been proposed with a pore thickness of ~ 33 Å and diameter of 22 Å [24]. However, no clear scientific evidence has been used in defining the stoichiometry of the pore. In addition, previous conductance study has shown that APA-1 formed pores with diameters varied from 6.3–22.1 Å. The hexameric pore model does not explain the observation of the smaller pore diameter (*e.g.*, 6.3 Å) [49]. In contrast, to fit through the membrane, the newly proposed dimer structure (Figure 3.27B) is likely to insert perpendicular to the membrane plane and lead to a pore structure with a thickness closer to physiological membranes (~ 49 Å) [225]. Additionally, it is likely that the stoichiometry of the pore will be different to the previously suggested hexamer. To unambiguously determine the pore structure of APA-1, other techniques such as electron microscopy, which enable direct visualisation of protein pores in membranes, and solid-state NMR spectroscopy, which provides detailed pore structural information at the atomic level can be used. Furthermore, it is not clear whether APA-1 dimerisation is crucial for pore formation or arises under the conditions of the *in vitro* studies undertaken (given that the interaction is weak). For example, residue K37, which is proposed to be important for APA-1 dimerisation, is highly conserved in the three amoebapore isoforms as well as SAPLIP members such as caenopore 5, saposin C and NK-lysin (Figure 1.15), indicating this residue may play an important role in APA-1 and SAPLIP protein functions. However, no evidence of dimerisation (or dimerisation via K37) exists in the other amoebapore isoforms and SAPLIP members. In addition, the closely related, non-pathogenic amoeba, termed *E. dispar*, possess orthologous of amoebapores named disparpores [226]. Disarpore-A has significantly less pore-forming activity, yet have near identical sequences when compared with APA-1 [226]. Four residues are different between the two protein sequences, including a key change of E2 to P2. This amino acid difference is significant because it would likely disrupt the proposed dimer models as well as membrane binding activity. Hence, this single amino acid difference may represent the primary reason for the lower activity of disparpore-A when compared with APA-1. Further research into the functional role of E2 is required. As such, ten mutants were designed to stabilise or disrupt the dimerisation models and will be used to characterise the importance of dimer formation and APA-1 activity.

3.4.2 Mechanism of amoebapore A activity

3.4.2.1 Lys residues

The pH-dependent activity of APA-1 clearly dictates that the ionisation states of charged amino acids are responsible for APA-1 function. Positively charged Lys residues have been shown to be essential in initial membrane binding, reflected by the selective binding to negatively charged phospholipids [8, 30, 31], and the disruption of APA-1 activity upon chemical modification of Lys residues [32]. However, the reported pKa range of Lys side-chains is 5.7–12.1 [48], which do not lie in the active pH range of APA-1, suggesting the ionisation state of Lys may play a negligible role in the pH-dependent activity. Nonetheless, to study the residue specific function of Lys residues, their side-chain pKa values can be calculated. This can be done using triple-resonance NMR spectroscopy by following the side chain $^{15}\text{N}\zeta$ (Lys) chemical shift [227]. However, such experiments can be difficult because of possible protein aggregation at extreme pH values, which has been observed for APA-1 (data not presented). As a complementary approach, site-directed mutagenesis of Lys residues can be performed and the activity of the APA-1 mutants studied.

3.4.2.2 Asp residues

In contrast, given the reported side-chain pKa ranges for acidic amino acids and histidine (0.5–9.2 for Asp, 2.1–8.8 for Glu and 2.4–9.2 for His), which do lie in the active pH range of APA-1 where changes in pH shift protein activity from inactive to active states, it is likely that the ionisation states of these amino acids are crucial in APA-1 function. Consequently, in this chapter the pKa values of the acidic amino acids and His residue were determined using NMR spectroscopy (Section 3.3.4). Using the calculated pKa values and structural data, further insights describing how the ionisation states influence APA-1 activity and which residues are linked to APA-1 pH-dependent activity can be identified.

The results showed that three side-chain pKa values were suppressed below the intrinsic value, D25, D55 and D66. Taking D55 as an example, the side-chain pKa was calculated to be 3.09 ± 0.12 at 1.0 mM with no significant variation at 0.1 mM (2.95 ± 0.12). D55 is located adjacent to K56 in the APA-1 structure (Figure 3.20). According to the general rules for interpreting side-chain pKa behaviours (Figure 3.18C), D55 may form a salt bridge with the side-chain of K56. Such a non-covalent interaction may contribute to the stabilisation of the protein fold [228]. Similar proximity features were observed for D25/K28 and D66/K64. Unfortunately, the

position of the side-chain is not well defined (Figure 3.20), and the structure which was resolved at pH 3.0 may not represent reliable side-chain positions at other pH values. Thus, clear evidence that salt bridges were forming between these residues requires further studies. This can be achieved by using a mutagenesis approach where the Lys residues are mutated and the pKa values of the aspartic acid residues measured to determine any change in the pKa values. A rise in the pKa value would indicate that the salt bridge was disrupted.

3.4.2.3 Glu16 and Glu62 residues

Residues E16 and E62 were observed to have elevated side-chain pKa values of 5.35 ± 0.04 and 5.35 ± 0.05 at 1.0 mM, respectively. As previously reported in various proteins, residues with unusual side-chain pKa values are normally located within active sites or important for protein function [167-169]. No pronounced pKa shift was observed when the protein concentration was lowered (5.38 ± 0.03 for E16 and 5.49 ± 0.03 for E62). This concentration independence, suggests that the molecular origin for the elevated side-chain pKa is not linked to APA-1 dimerisation. This is in agreement with the newly proposed dimer model (Figure 3.36) where the residues are located distal from the dimer interface. Therefore, the elevated pKa must values are a likely consequence of the chemical environment of these two residues.

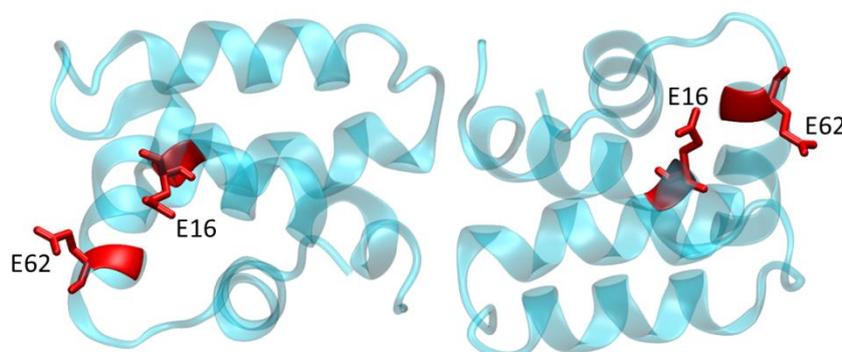


Figure 3.36: Representation of residue E16 and E62 on the proposed new dimer model. Residues E16 and E62 are labelled in red and located distal from the dimer interface.

Based on the APA-1 structure, E16 and E62 are located in close proximity along with D63, which also has a slightly elevated side-chain pKa (4.27 ± 0.07) (Figure 3.37). According to the general rules for interpreting side-chain pKa behaviours (Figure 3.18B), deprotonation of these residues would give rise to repulsive interactions, resulting in a negatively charged environment that contributes to the elevated pKa values. APA-1 has the highest pore-forming activity at pH 5.2 and the activity was dramatically reduced above 5.2, in particular between

pH 5.2–5.5. The side-chain pKa values of E16 and E62 reside in this pH range, making these two residues potential candidates responsible for stabilising/abolishing the pH-dependent activity of APA-1. The region becomes more negatively charged upon deprotonation of the residues, which may contribute to the loss of APA-1 activity.

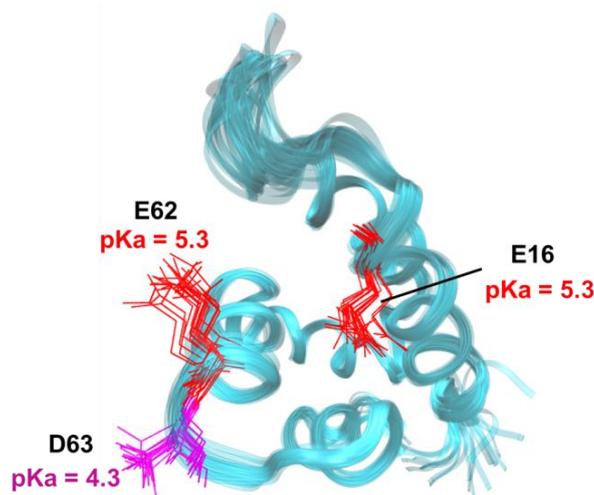


Figure 3.37: Representation of the side-chains of E16, E62 and D63 in the APA-1 structure. The 20 lowest energy minimised structures are presented. The Glu (E16 and E62) and D63 residues are respectively coloured in red and magenta. All three residues have elevated side-chain pKa values.

Interestingly, E16 and E62 are among the six charged residues which are conserved across the amoebapore isoforms (Figure 1.4), suggesting they may be functionally important. Moreover, similar to APA-1, saposin C and caenopore 5 which show pH-dependent activity and interact favourably with negatively charged membranes, have the highest activity observed at pH 5.3 and 5.6, respectively [70, 229]. Intriguingly, a number of Glu residues in saposin C were found to have elevated pKa values close to 5.5 [75], including E64 which is conserved with E62 in APA-1 (Figure 3.38). Unusual side-chain pKa values (pKa = ~5) were also obtained for residues E17 and E63 of caenopore 5, which are structurally conserved with E16 and E62 of APA-1 (performed by Wei Li, School of Biological Sciences, The University of Auckland) (unpublished data) (Figure 3.38). All the elevated side-chain pKa values are close to the highest activity of these SAPLIP members, suggesting the ionisation state of the Glu residues is a key determinant of protein activity. By mutating two glutamate residues in saposin C, hence decreasing the overall negative charge of the protein electrostatic surface, stronger interaction with membranes was observed. The result indicates that side-chain neutralisation of Glu residues is essential for triggering saposin C-anionic lipid interaction [75]. In contrast, NK-lysin [72] and granulysin [73], whose activities are pH-independent, display remarkable positive surfaces which are postulated to drive membrane destabilisation via carpet-like

mechanisms [77]. Interestingly, E16 and/or E62 of APA-1 are not conserved in these proteins, indicating ionisation of acidic residues may be a unique feature for particular SAPLIP proteins that present pH-dependent activity. Bioassays with mutated E16 and/or E62 should provide insights about their roles in APA-1 activity.

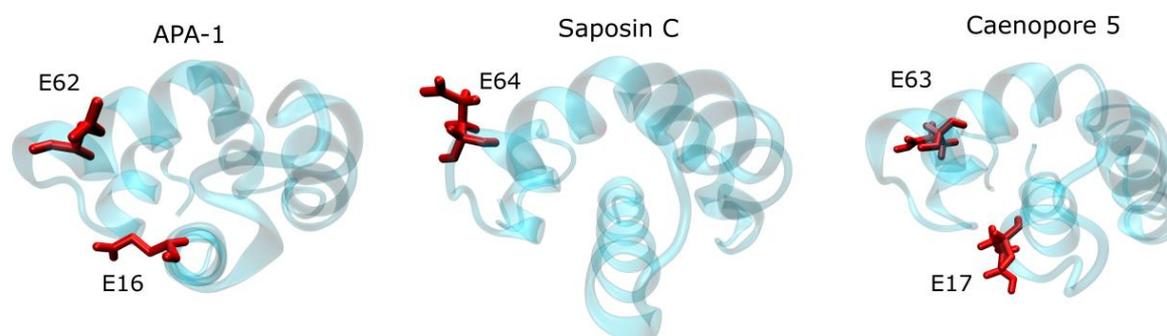


Figure 3.38: Superimposed structures of APA-1, saposin C and caenopore 5. E16 and E62 are structurally conserved in saposin C and caenopore 5.

3.4.2.4 His75 residue

The sole His (H75) has been hypothesised to be crucial in driving APA-1 dimerisation (Figure 3.37A) [24]. However, the chemical shift perturbation study showed that the resonance of H75 displayed no noticeable chemical shift change over the concentration range examined (Figure 3.8). In the new dimer model derived in this thesis, H75 is not located at the dimer interface (Figure 3.39). The side-chain pKa of the H75 was calculated to be 7.29 ± 0.07 at 1.0 mM, elevated above the intrinsic value of 6.5 [43], and no noticeable variation of the pKa was observed at a 10-fold lower concentration (7.19 ± 0.07). In conclusion, the results showed that the elevated pKa of H75 is not related to APA-1 dimerisation, but may play a role in the membrane binding activity of this protein.

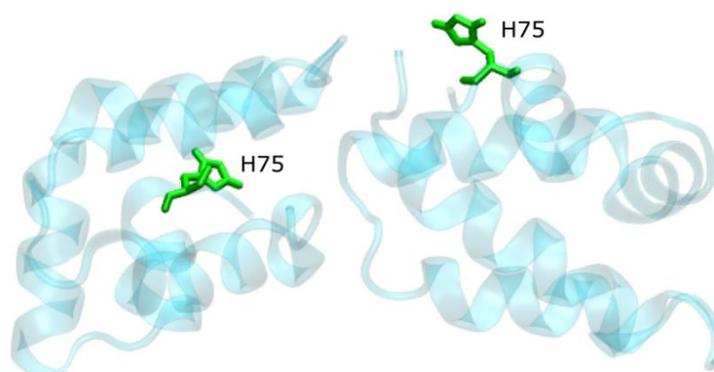


Figure 3.39: Proposed APA-1 dimer structure. Residue H75 in each monomer is coloured in green.

Chemical modification of H75 removes pore forming activity [32] and the residue is highly conserved across the amoebapore isoforms, suggesting the residue is functionally important. In addition, the elevated side-chain pKa of H75 suggests that charged form of the residue is favoured in APA-1 and the pKa is close to the pH range of 6.5–7, where APA-1 abruptly loses its association with membranes [32] (Section 1.2.4). This suggests that deprotonation of H75 may be a key determinant for APA-1-membrane interaction. Based on the APA-1 structure, H75 is located close to K48 and K73 (Figure 3.40). At pH values well below 7, the protonated H75 together with K48 and K73 may form a positively charged cluster, which facilitates the interaction between APA-1 with an anionic membrane. Upon deprotonation of H75, the overall electrostatic surface charge of APA-1 is reduced (coupled with the deprotonation of Asp and Glu residues), which may be partly responsible for the loss of APA-1-membrane interactions. Based on the sequence alignment of SAPLIP members, the positive charge at position 75 in APA-1 is conserved in saposin C, NK-lysin and granulysin (Figure 1.13). Intriguingly, the H75 in APA-1 is replaced with Lys and Arg residues in NK-lysin and granulysin, respectively. The much higher intrinsic pKa values of Lys and Arg (10.4 and 12.0)[43] dictates that these residues are always protonated over a wide pH range, which may be responsible for the pH-independent activity of these two proteins. To further examine the role of H75, site-directed mutagenesis studies are required.

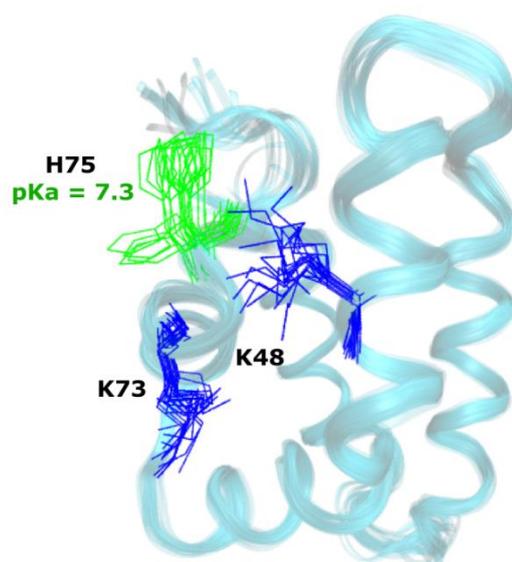


Figure 3.40: Representation of the side-chains of H75, K48 and K73 in APA-1 structure. The 20 lowest energy minimised structures are presented. The side-chains of His and Lys are respectively coloured in green and blue. The side-chains of these three residues are located close to each other.

4

Interaction of amoebapore A with model membranes

4.1 Introduction

4.1.1 Amoebapore A against membranes

Chapter 1 and 3 showed that APA-1 is an AMP that destroys cells via a pH-dependent mechanism. It is postulated that APA-1 dimers aggregate on the target membrane, and once a threshold protein concentration is reached the dimers assemble to a pore-like structure in the membrane [24]. There is, however, no evidence of APA-1 pore-formation and molecular information describing the membrane binding mechanism of APA-1 is missing.

A range of biophysical techniques such as quartz crystal microbalance and neutron reflectometry have been used to determine what mechanisms are implemented by membrane binding proteins. Since biological cellular membranes are complex, difficult to study, and challenging to prepare *in vitro*, simpler biomimetic membrane models, which still capture important biological characteristics, such as lipid composition, charge and fluidity, are commonly used. Changes in the property of different membrane models upon interaction with APA-1 can be readily assessed using model membrane systems.

4.1.2 Quartz crystal microbalance with energy dissipation

Quartz crystal microbalance (QCM) is a highly sensitive instrument to investigate adsorption and binding events to surfaces. It is well suited for the investigation of the interactions between protein and phospholipid *in situ* [230-234]. Information about the formation of membranes and the mass and structural changes occurring to the membrane upon protein binding can be obtained in real-time [230]. The method relies on the piezoelectric effect of a quartz crystal (*e.g.*, silicon) sandwiched between electrodes (*e.g.*, gold). When a voltage is applied across the quartz crystals, a mechanical stress is induced. Alternating the applied voltage causes the crystal to oscillate at its fundamental resonance frequency and harmonics of the fundamental frequency. The measured frequency is dependent on the amount of mass adhered to the crystal surface. Therefore, mass change induced by phospholipid membrane deposition onto the crystal and protein adsorbed can be probed by measuring the change in frequency based on the Sauerbrey relation (Eq. 4.1) [235]:

$$\Delta F = -C_f \cdot \Delta m \quad \text{Eq. 4.1}$$

where ΔF is the change of frequency (Hz), C_f is the sensitivity factor of a crystal, and Δm is the change of mass per area (ng cm^{-2}). The C_f for a 5 MHz quartz crystal at room temperature is $56.6 \text{ Hz cm}^2 \mu\text{g}^{-1}$. A decrease of the resonance frequency is an indicator that the adhered mass has increased [235]. However, the Sauerbrey relation is only valid for rigid and homogeneous films, and hence is not appropriate for loosely structured mass that is adhered to the surface, for example, a disrupted membrane surface. To overcome this problem, a QCM with energy dissipation (QCM-D) is introduced. The instrument monitors the dissipation factor (D) to measure the energy loss of an adsorbed film. A driving alternating voltage is removed periodically (approximately once per second) to measure the energy damping of the crystal into the surrounding environment. The dissipation factor can be calculated using Eq. 4.2 [236],

Eq. 4.2

$$D = E_{\text{dissipated}} / (2\pi E_{\text{stored}})$$

where $E_{\text{dissipated}}$ is the energy lost from the crystal during a single resonance after the voltage is removed, and E_{stored} is the initial energy of the crystal. The D is an indicator of how a crystal oscillation decays after the driving voltage is removed, hence, giving information about the viscoelastic properties of a film. A high D value indicates that energy is lost from the surface quickly, suggesting a thick, soft or loose film; whereas a low D implies the formation of a rigid and compact film [237]. The penetration depth of the harmonic waves is inversely proportional to the resonance frequency of the wave [238]. Therefore, higher harmonics probe closer to the surface of the quartz crystal, whereas lower harmonics probe closer to the bulk solvent. By combining the measured F and D at different harmonics, 3D information of the mass and viscoelasticity change of a film before and after protein adsorption can be obtained [230]. This information can be used to model the protein mechanisms.

In addition, a change of resonant frequency versus a change of energy dissipation plot (ΔF - ΔD plot) at all harmonics provides useful information about the complexity of the adsorption process [237]. The ΔF values are plotted reversely on the x-axis to reflect mass increase, and the ΔD values are plotted on the y-axis (Figure 4.1). The ΔF - ΔD plot consists of a number of discrete points, and each data point refers to the ΔF and ΔD values at a particular time point, and the point (0,0) corresponds to time 0 [237]. Once the QCM-D measurement starts, the data points shift. The plot is able to show how the structure of the membrane changes upon mass addition. A change in the direction of the trace suggests a different process is occurring. A linear relationship indicates a single mechanism whereas a more complex relationship indicates that different processes are involved. In addition, kinetic information can be extracted from the

plots. The data points are acquired about every second, therefore dispersed data is an indicator of a rapid process whereas closely spaced data suggests a slow process [237]. The traces can be used to give a more comprehensive fingerprint for understanding the protein-membrane interactions.

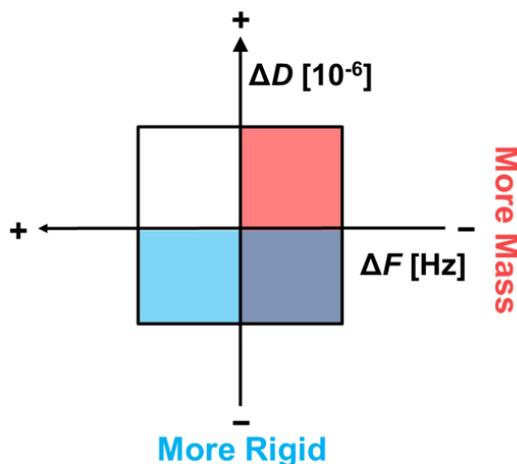


Figure 4.1: A schematic diagram interpreting the ΔF - ΔD plot. Data points on the plot pointing south suggest that the process is causing the surface to become more rigid; conversely, data points pointing north suggest that the surface is loose. Data points moving east indicate adsorption of mass onto the surface whereas data points moving west indicate mass lost from the surface. The picture was made according to [239].

4.1.3 Neutron reflectometry

Neutron reflectometry is used to structurally characterise protein adsorption onto a flat membrane under a wide range of environments [240-243]. Neutron reflectometry probes variation in the neutron scattering length density (SLD) of a film perpendicular to a surface with molecular resolution. The neutron source produces beam of neutrons that is highly collimated and strike on a flat film at a known angle (θ) beyond the critical angle for total external reflection with incident neutron beam intensity (I_i) and wavelength λ (Å) (Figure 4.2). Materials which are highly transparent to neutrons, flat, and smooth (*e.g.*, polished silicon crystals) are usually used as solid substrates for membrane deposition so that the membrane-lipid interface can be approached through the solid surface to avoid scattering in the bulk solution. Low-energy neutrons are non-destructive and thus impact no physical damage to a sample in comparison to X-ray scattering. Some incident neutrons are transmitted through or absorbed by the sample, whereas other neutrons are reflected. The intensity of the reflected neutrons (I_r) is detected and reflectivity (R) is measured as the reflected intensity relative to the incident intensity ($R = I_r / I_i$). The wave-vectors k_i and k_r define the incident and reflected

beams, respectively. The momentum transfer vector (Q) that is perpendicular to the surface is the difference in the wave-vectors and can be calculated using Eq. 4.3:

$$Q = 4\pi \sin \theta / \lambda \quad \text{Eq. 4.3}$$

Thus the reflectivity can be measured in two ways: either by a single λ and different θ , or using a fixed θ and different λ to cover a desired Q range. High Q data is essential in neutron reflectometry experiments because the data can provide molecular resolution. The reflectivity R is measured as a function of Q . From $Q = 0$ to Q_c (Q at critical angle), there is absolute reflectivity followed by a rapid decline beyond Q_c , yielding a reflectivity profile $R(Q)$ [240-243] that can be defined using Eq. 4.4:

$$R(Q) = (16 \pi^2 / Q^4) / |\rho(z)| \quad \text{Eq. 4.4}$$

where $|\rho(z)|$ is the Fourier transform of the SLD. The SLD of a film with known composition and density can be defined as:

$$\text{SLD} = N_A b_i \rho_i / M_i \quad \text{Eq. 4.5}$$

where N_A is Avagadro's constant, b_i is the neutron scattering length of component i with mass density ρ_i and atomic mass M_i . Consequently, the $R(Q)$ profile contains information of the SLD gradient within the interface and the SLD gradient contains information of the thickness, density and roughness of the interfaces [244].

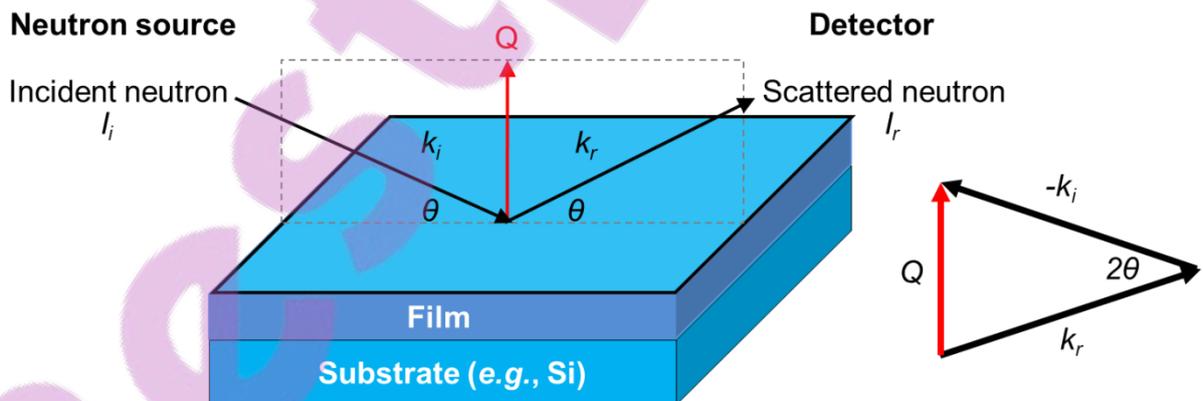


Figure 4.2: A schematic diagram representing reflectivity experiments performed on a thin film. Substrate is used for film formation. An incident neutron beam strikes the surface at reflected angle θ . The momentum transfer Q is the difference between vectors k_i and k_r .

A unique set of structural parameters of an interface cannot be directly determined from a single reflectivity measurement because phase information in the reflected waves is lost (known as the ‘phase problem’). Consequently, the $R(Q)$ data cannot be directly inverted – there is a number of equally valid $\rho(z)$ profiles that will give the same reflectivity. Therefore, it is customary to model the SLD profile as a set of homogeneous layers on a surface. Each layer is defined with parameters of film thickness (d), interfacial roughness (σ) and the SLD. The parameters are then adjusted until the calculated reflectivity agrees well with the measured $R(Q)$ [244]. Naturally, this gives rise to some uncertainty about the uniqueness of the model fit - neutron reflectometry is able to reduce this ambiguity using isotopic hydrogen and deuterium (H/D) substitution in the solvents to create enhanced contrast variation because the nuclei of H and D scatter neutrons with opposite phase [240-243]. The SLD values of a few common nuclei present in proteins and lipids are listed in Table 5.1. Moreover, by isotopically deuterating different components in a complex planar surface, different contrast variation can be achieved to obtain structural information of the interaction interface such as the relative location of lipids and proteins in a protein-membrane complex system. However, some protein-lipid membranes may be too delicate for the physical exchange from H to D [240-242, 245].

Table 5.1: The scattering length, b , of common nuclei [246].

Nucleus	Scattering length / fm
^1H	-3.74
^2D	6.67
^{12}C	6.65
^{16}O	5.80
^{15}N	6.44
^{32}S	2.80

4.1.4 Aims

The research aim of this chapter is to prepare various membrane models with different properties, such as charge, composition and fluidity, and characterise the interactions of APA-1 with the membrane models using QCM-D and neutron reflectometry under different solvent conditions. The information enables a critical assessment of the proposed mechanisms of APA-1 activity, and lays the framework for further studies of APA-1-membrane interactions.

4.2 Methods and Materials

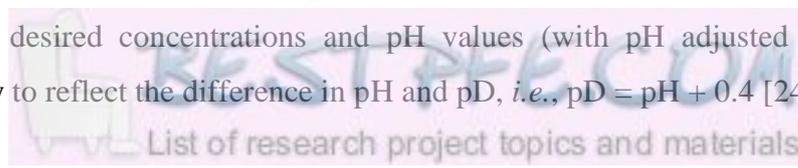
4.2.1 Liposome preparation

1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 1,2-dimyristoyl-sn-glycero-3-phospho-rac-(1-glycerol) (DMPG), 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhyPC), 1,2-dipalmitoyl-sn-glycero-3-(1'-rac-glycerol) (DPPG), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-rac-(1-glycerol) (POPG) and *E. coli* polar lipid extract were purchased from Avanti Polar Lipids (Alabaster, USA). Chloroform, methanol, sodium citrate and disodium hydrogen phosphate (NaPi) were purchased from Sigma-Aldrich Co. (St.Louis, USA). All solutions were filtered through a 0.22 μm nitrocellulose membrane.

Ten milligrams of each lipid was dissolved in chloroform/methanol (75:25 v/v) to create 10 mg/ml individual stock solutions. The stock solutions were mixed in glass tubes to the desired molar ratios. The solvent was evaporated under a gentle stream of N_2 gas and fully dried overnight under vacuum. The dried lipids were re-suspended in 10 mM sodium citrate/NaPi (pH 5.2 or 8.5) and 150 mM NaCl to give a concentration of 0.5 mg/ml. This was followed by vigorous mixing until homogenous solutions were achieved. The lipid solutions were incubated in a water bath at 50 $^\circ\text{C}$ for 30 min then sonicated using a probe sonicator (ThermoFischer Scientific, Waltham, USA) at 3 W for cycles of 10 sec sonication with 10 sec delays in between. This avoided sample heating and was continued until a clear solution was observed. The lipid solutions were further diluted in the same buffer to 0.1 mg/ml. Brief sonication was required before use. Seven lipid solutions were prepared: DMPC, DMPC:DMPG (2:1), DPhyPC, DPhyPC:DPPG (2:1), POPC, POPC:POPG (2:1) and *E. coli* polar lipid extract.

4.2.2 Protein preparation

APA-1 was expressed from a cell-free protein expression system and purified as described in Chapter 2. The purified protein was lyophilised and re-dissolved into small volume of ultra-pure water to ~ 1.0 mM. The protein concentration was determined by UV absorption with $\mathcal{E}_{280} = 1865 \text{ M}^{-1} \text{ cm}^{-1}$. The protein solution was aliquoted and diluted with 10 mM sodium citrate/NaPi in either H_2O (for QCM-D experiments) or in D_2O (for neutron reflectometry experiments) to the desired concentrations and pH values (with pH adjusted in the D_2O solutions accordingly to reflect the difference in pH and pD, *i.e.*, $\text{pD} = \text{pH} + 0.4$ [247]).



4.2.3 Quartz crystal microbalance with dissipation

Silicon/silicon oxide (Si/SiO₂) crystals with deposition area of 1 cm² were purchased from ATA Scientific (Sydney, Australia). Before use, they were thoroughly cleaned using a Bioforce UV tip-cleaner (BioForce Nanoscience Inc., Ames, USA) for 15 min, incubated in 2% (v/v) Hellmanex for 10 min at room temperature, washed extensively in ultrapure water and dried under a gentle stream of N₂ gas followed by cleaning in the Bioforce UV tip-cleaner for another 15 min. The cleaned crystals were assembled in the QCM-D system which can hold four crystals and rinsed with 2.5 ml of ultra-pure water at 500 µl/min, 2 ml of Hellmanex detergent solution, 2 ml of HPLC grade ethanol followed by another 2 ml of ultra-pure water at 200 µl/min. Finally the system was equilibrated with 2 ml of 10 mM sodium citrate/NaPi buffer and 150 mM NaCl at 200 µl/min.

QCM-D measurements were carried out using a Q-SENSE E4 system (Gothernburg, Sweden) which has four chambers for sensor crystals and the deposition of materials and flow rates controlled by a peristaltic pump. The changes of resonance frequency and energy dissipation upon mass deposition were measured simultaneously at seven different overtones of the natural frequency (5 MHz) to provide more structural information on different layers attached to the crystal surface. Measurements at the first overtone were not considered due to their high sensitivity to bulk solution changes and thus high fluctuations. All experiments were performed at 25 °C.

To determine the structural changes of the membranes upon APA-1 binding, the formation of lipid bilayers on the solid supported surface were initially assessed using QCM-D. Four different bilayers mimicking eukaryotic and prokaryotic bilayers on surfaces were chosen, DMPC, DMPC:DMPG (2:1), DPhyPC and DPhyPC:DPPG (2:1). The bilayer formations were prepared by *in situ* liposome deposition onto a Si/SiO₂ surface. QCM-D cells were pre-equilibrated with buffer (10 mM sodium citrate/NaPi, 150 mM NaCl, pH 5.2 or 8.5). Eight hundred and fifty microliters of a 0.1 mg/ml vesicle solution was flowed over the surface at 50 µl/min for deposition and allowed to incubate for 1 h to achieve a stable baseline. Stable base line at $\Delta F = -20$ to -25 Hz and $\Delta D < 1$ should be achieved after the incubation period, indicating the formation of a stable bilayer with negligible viscoelasticity and low defects [237]. Such bilayer formations were observed for DMPC and DPhyPC:DPPG (2:1) lipids. Figure 4.3A demonstrates a QCM-D profile for the formation of a DPhyPC:DPPG (2:1) bilayer. Firstly the frequency and dissipation reached extreme values (-58 Hz and 2.8×10^{-6} , respectively),

reflecting fusion of the vesicles onto the surface. The frequency and dissipation profiles then started to stabilise to values of -22 Hz and zero, respectively, corresponding to the spontaneous rupture of the vesicles and lipid bilayer formation [248]. Rinsing with $850\ \mu\text{L}$ of buffer at $50\ \mu\text{l}/\text{min}$ was applied to remove excess lipid material and a stable bilayer on the surface was achieved.

If after the incubation period the values are greater than those quoted above ($\Delta F = -20$ to -25 Hz and $\Delta D < 1$), then intact vesicles are bound to the surface. To rupture the vesicles a $2.0\ \text{M}$ NaCl wash was required ($850\ \mu\text{l}$ at $50\ \mu\text{l}/\text{min}$) followed by the buffer wash to leave a stable bilayer membrane on the surface. The extra vesicle rupture step was required for bilayer formation using DPhyPC and DMPC:DMPG (2:1) lipids. Figure 4.3B illustrates DMPC:DMPG (2:1) bilayer formation as an example. A thick layer of intact vesicles were deposited onto the Si/SiO₂ surface indicated by the low frequency value of -60 Hz. A solution of $2\ \text{M}$ NaCl was applied to help rupturing the vesicles. After rinsing the surface, the frequency slowly equilibrated to -22 Hz, suggesting the formation of a stable bilayer. Consistent frequency and energy dissipation at different overtones shows that a rigid and compact film has formed (Figure 4.4).

After the formation of the bilayers, $180\ \mu\text{L}$ of APA-1 at the desired pH value (5.2 or 8.5) was injected into the chambers at $50\ \mu\text{l}/\text{min}$. Peptide that remained in the dead volume of the flow system was introduced into the chamber by injecting a small amount of buffer ($180\ \mu\text{l}$) immediately after APA-1 addition. The protein was incubated with the membranes for at least 20 minutes until a stable base line was reached and then excess APA-1 was removed by a buffer rinse. All frequency and energy dissipation were followed in real-time. The original data were plotted and analysed using Igor Pro 6.22A (WaveMetrics Inc., Lake Oswego, USA). QCM-D experiments were repeated between 4 to 6 times and reproducible data were obtained. All QCM-D experiments were performed at ANSTO (Sydney, Australia).

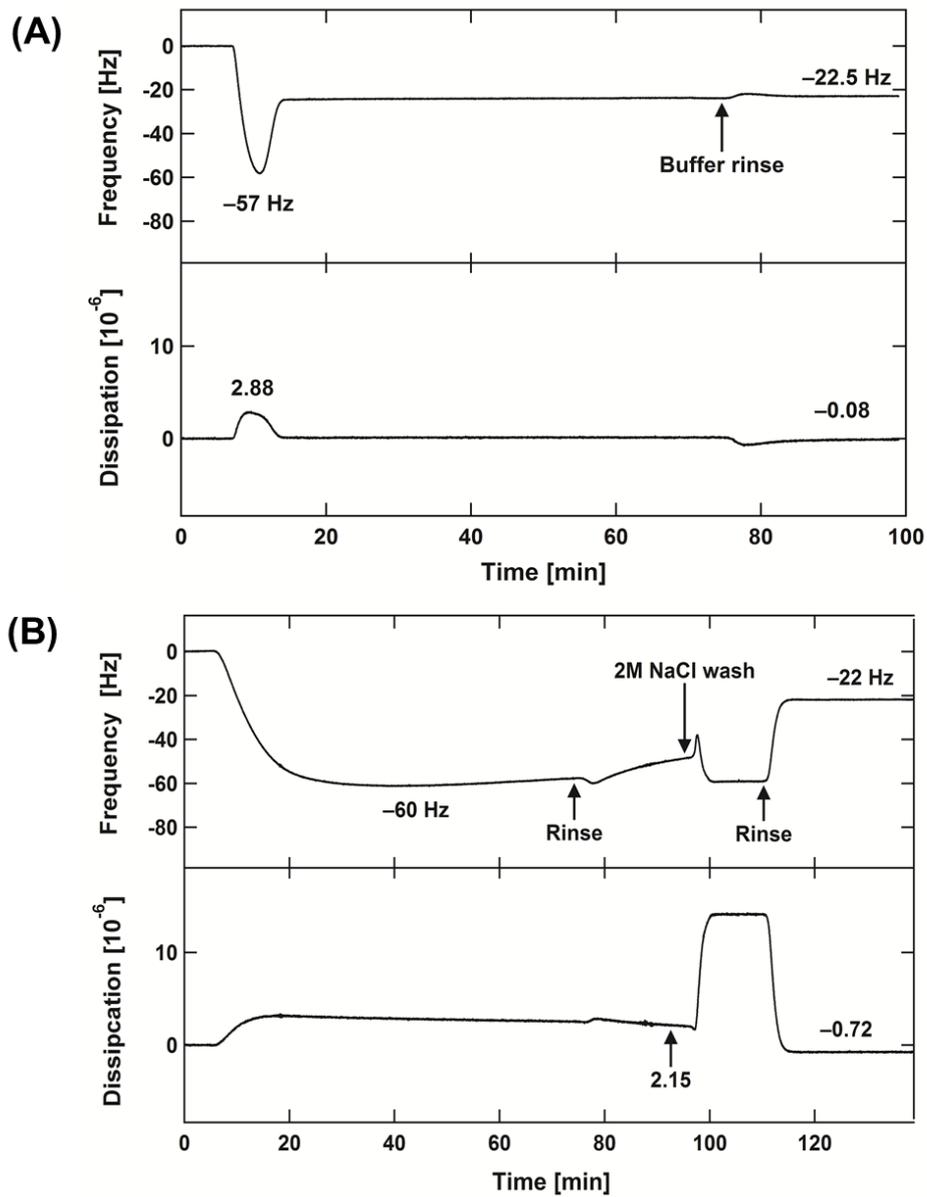


Figure 4.3: The changes of the resonant frequency and energy dissipation versus time for vesicle deposition onto the Si/SiO₂ surfaces. Panel (A) represents DPhyPC/DPPG bilayer formation and panel (B) represents DMPC/DMPG bilayer formation.

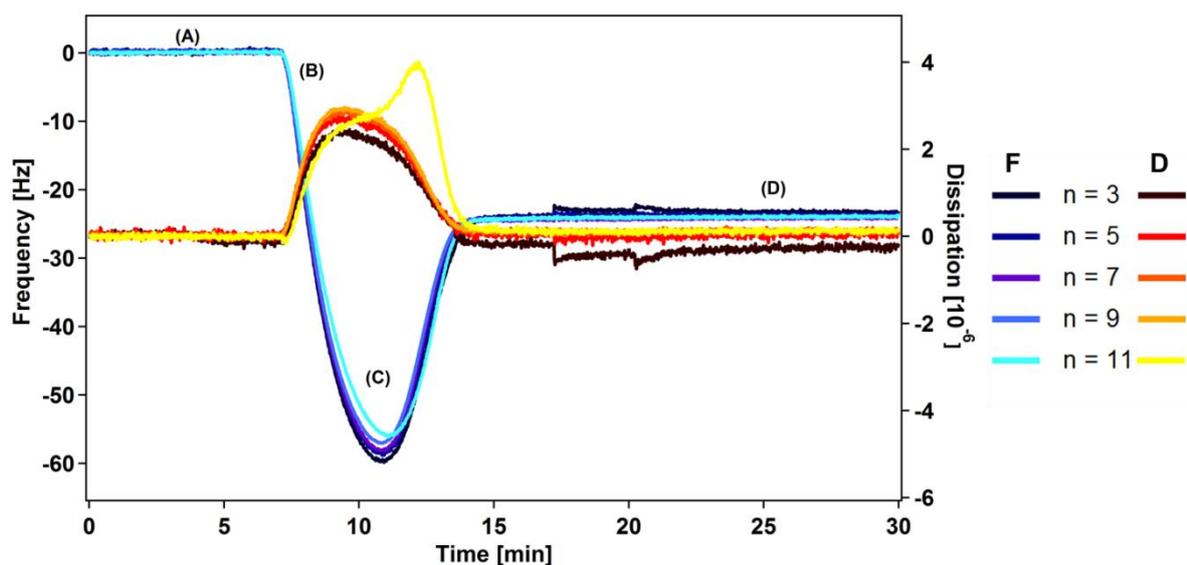


Figure 4.4: The changes of resonant frequencies (F) and energy dissipations (D) versus time for DPhyPC:DPPG (2:1) bilayer formation on the Si/SiO₂ surface at different overtones. Consistent frequency and energy dissipation at different overtones indicate formation of a rigid membrane. Processes (A)–(D) respectively correspond to initial equilibrium of the Si/SiO₂ surface, vesicle deposition, vesicle rupture and bilayer formation.

4.2.4 Sample preparation for neutron reflectometry.

The silicon oxide wafers for neutron reflectometry experiments were thoroughly rinsed with ultra-pure water and ethanol, dried with air and cleaned in a UV-ozone chamber (Jelight Company Inc., Irvine, USA) for 15 min. Vesicle solutions were prepared in the same way as stated in the QCM-D experiments and the same buffer was used. The silicon oxide wafers were heated to ~50 °C then rinsed with 1.5 ml of 10 mM sodium citrate/NaPi buffer with 150 mM NaCl (pH 5.2 or 8.5). One millilitre of each of the vesicle solutions were slowly deposited onto the wafers and incubated for 1 h. Two millilitres of 2 M NaCl was then applied to DMPC:DMPG (2:1) and DPhyPC membranes to facilitate the rupture of the vesicles, and therefore the formation of a complete and stable lipid bilayer. All membranes were equilibrated with 5 ml of 10 mM sodium citrate/NaPi buffer in D₂O (pH 5.2 or 8.5). When the wafers were cooled to room temperature, 800 µl of APA-1 in the D₂O buffer was slowly applied to the wafers immediately followed by injecting another 300 µl of D₂O buffer to push the protein solution from the tube onto the silica surface. Measurements were performed after incubating APA-1 with membranes for 30 min and rinsed with buffer.

4.2.5 Neutron reflectometry

Neutron reflectometry data were collected from supported lipid membranes using the Platypus time-of-flight neutron reflectometer at the OPAL 20 MW research reactor (ANSTO, Sydney, Australia) [249, 250]. Cold neutron pulses of bandwidth $2.8 \text{ \AA} \leq \lambda \leq 18.0 \text{ \AA}$ and 23 Hz frequency were generated using a boron-coated disc chopper system in the low resolution mode ($\Delta\lambda/\lambda = 7\%$). Direct and reflected beam spectra were recorded using a two-dimensional helium-3 neutron detector. All measurements were carried out at room temperature (25 °C). Reflected beam spectra from the supported lipid bilayers were collected at $\theta_1 = 0.45^\circ$ for 10 minutes, $\theta_2 = 1.6^\circ$ for 20 min and at $\theta_3 = 5.0^\circ$ for 1 h. Constant illumination of the sample was achieved using collimating slits set at 37 mm wide \times 0.25 mm high for θ_1 , 37 mm wide \times 0.88 mm high for θ_2 and 37 mm wide \times 2.74 mm high for θ_3 . Direct beam measurements were collected in transmission mode through the Si blocks under the same collimation conditions as for θ_1 , θ_2 and θ_3 for 5, 10 and 4 minutes, respectively.

Direct and reflected beam time-of-flight spectra were reduced using the SLIM software [251] that runs within the Igor Pro data processing environment (IGor Pro, WaveMetrics Inc., Lake Oswego, USA). Analysis of the neutron reflectivity profiles were performed using the MOTOFIT analysis program, with data presented as a function of the momentum transfer normal to the surface ($Q = 4\pi \sin \theta / \lambda$). A $\sim 22 \text{ \AA}$ thick oxide layer on the surface of the Si substrates was included in each of the structural models. The supported bilayer membranes are described by a three-layer model consisting of a headgroup near the Si/SiOx surface, a tail layer that encompasses both acyl leaflets of the bilayer and a second headgroup at the solution interface. Structures were refined by varying the scattering length density, thickness and roughness between adjacent layers, and a Levenberg-Marquardt/Genetic algorithm used to minimize values of χ^2 .

4.3 Results

4.3.1 Quartz crystal microbalance with energy dissipation

The frequency and energy dissipation changes in the QCM-D profiles in response to the binding of 5 μM APA-1 to the DMPC, DMPC:DMPG (2:1), DPhyPC and DPhyPC:DPPG (2:1) bilayers were assessed. The QCM-D penetration signal is approximately 0.25 μm in water [252], allowing the detection of both bilayer and APA-1 bound to the surface of the bilayer. Two pH values were used to investigate the interactions: pH 5.2, where APA-1 is most active; and pH 8.5 where the protein has been shown to be inactive [32]. APA-1-membrane interactions at pH 3.0 were not performed because preliminary result showed that formation of a stable bilayer could not be achieved under acidic conditions. Only the frequencies and energy dissipations at the 7th overtone are presented in figures, unless otherwise stated.

Changes in the lipid bilayer architecture upon addition of APA-1 were investigated at 25 °C. Figure 4.5 represents the changes of membrane properties upon binding of 5 μM APA-1. APA-1 was observed to have the most pronounced effect on the DPhyPC:DPPG (2:1) bilayer. In comparison with other membranes, binding of APA-1 led to significant mass removal from the membrane, reflected by the increase of resonant frequency. The interaction of APA-1 with the DPhyPC:DPPG (2:1) bilayer was repeated at pH 8.5. No binding was observed, in agreement with the previous study showing that APA-1 completely lost its association with liposomes at pH above 7 [32]. Minor changes were detected for the DMPC, DMPC:DMPG (2:1) and DPhyPC bilayers upon addition of 5 μM APA-1. A slow increase of the resonant frequencies was observed for the DPhyPC and DMPC:DMPG (2:1) bilayers and a decrease of resonant frequency was observed for the DMPC bilayer. These subtle changes are comparable to the lower end of the sensitivity of QCM-D (± 1 Hz), and are thus not considered to be significant. More discussion will be given in the ΔF - ΔD plots analysis.

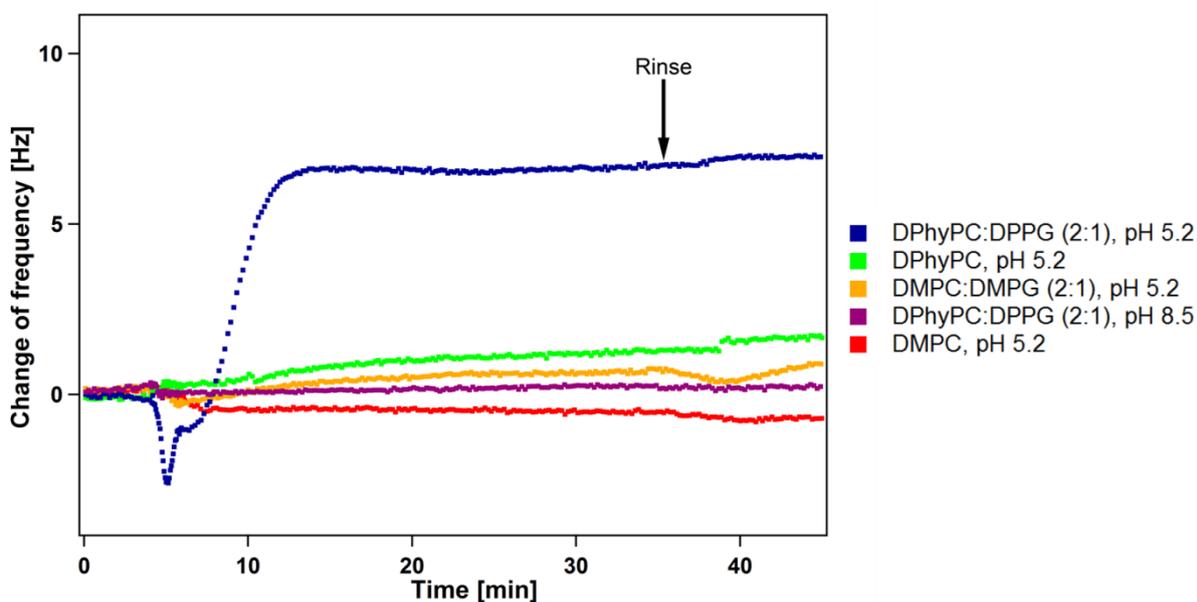


Figure 4.5: Time-dependent changes in the resonant frequency for different membranes upon binding of 5 μM APA-1. Binding of APA-1 caused the most significant mass removal on the DPhyPC:DPPG (2:1) bilayer at pH 5.2. Frequency changes at the 7th overtone are presented. Time 0 is the point when APA-1 was injected to the QCM-D system.

To further analyse the structural changes to the DPhyPC:DPPG (2:1) bilayer upon binding of APA-1, frequencies and energy dissipations at different overtones of the natural frequency of the QCM-D sensor crystal are presented in Figure 4.6. Such an examination allows a more complete characterisation of the mass and viscoelasticity properties of the membrane as a function of distance from the surface. A representative set of the third, fifth, seventh, ninth and eleventh overtones of the frequency and dissipation for the DPhyPC:DPPG (2:1) bilayer with APA-1 bound is shown in Figure 4.6. The first overtone was not used because it probes mostly the bulk solution and thus gives fluctuating data. It was observed that upon APA-1 binding, the frequencies measured by all overtones decreased to the same level, suggesting the structural characteristics were the same across the entire thickness of the membrane. This may be an indicator of a vertical transmembrane insertion [230]. Immediately after the insertion process, mass gain and removal was observed at different overtones followed by removal of considerable amount of lipid materials from the bilayer, as indicated by the increased frequencies at all overtones. The frequencies at different overtones became widely dispersed after APA-1 binding, suggesting different amounts of mass were removed from different layers of the membranes hence an asymmetrical disruption mechanism. Dispersed energy dissipation was also observed, reflecting different viscoelasticity properties at different layers on the surface. The third overtone showed the most pronounced change of both resonant frequency

and energy dissipation, suggesting the most significant change occurred closer to the surface of the membrane and the layer is loose and soft (disordered). No pronounced change was observed after rinsing the surface with buffer.

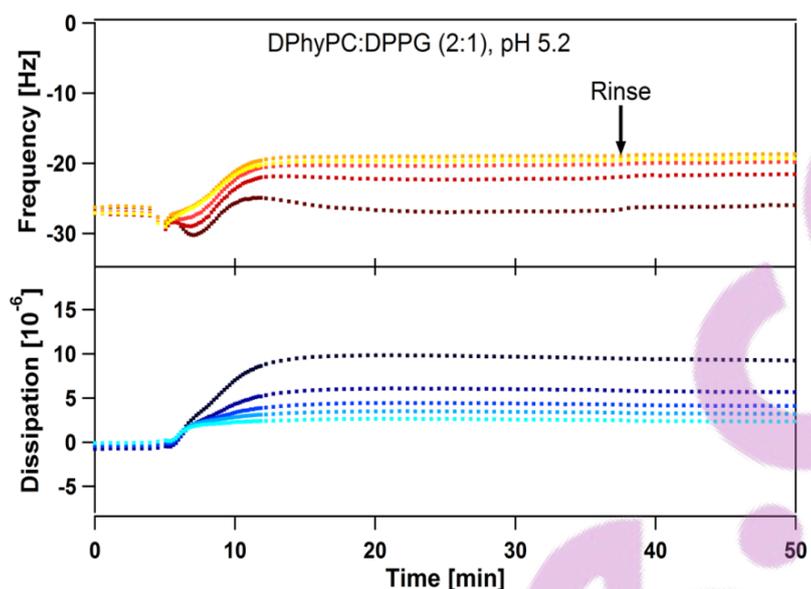


Figure 4.6: Time-dependent changes in the frequencies (F) and energy dissipations (D) on the DPhyPC:DPPG (2:1) bilayer upon binding of 5 μ M APA-1. Consistent changes in frequency and energy dissipation at different overtones followed by a divergent change suggests a transmembrane insertion followed by an asymmetrical membrane disruption and loss of membrane rigidity. Time 0 is when APA-1 was added.

To further understand the binding mechanism of APA-1 with membrane models, $\Delta F-\Delta D$ plots were performed for all the QCM-D profiles. It was observed that APA-1 interactions with both DMPC and DMPC:DMPG (2:1) bilayers resulted in a one phase adsorption mechanism, indicated by the single black arrows (Figure 4.7). The interaction of APA-1 with a DMPC bilayer can be represented by a north-east arrow (Figure 4.7A), which, according to Figure 4.2, is interpreted as an increase in mass of the membrane with a corresponding loss of rigidity. The $\Delta F-\Delta D$ plot of the 3rd overtone displayed the greatest change to the membrane upon APA-1 interaction, whereas no pronounced change was observed at the 11th overtone. This indicates that binding of APA-1 has caused minor changes to the surface of the DMPC bilayer: a subtle amount of APA-1 was bound to the surface of the membrane and slightly reduced the surface membrane rigidity. However, since both the frequency and dissipation variations were small, the process could also be due to an apparent effect caused by a difference in density or viscosity of the protein solution being introduced in the QCM-D chamber. Overall, the observations clearly showed that APA-1 interactions with a PC-type membrane were weak and

transient, and did not lead to dramatic changes in the overall architecture of the DMPC bilayer. In contrast, the interaction of APA-1 with a DMPC:DMPG (2:1) bilayer clearly showed that APA-1 bound to the membrane and a mass increase was observed without any major change in energy dissipation (Figure 4.7B).

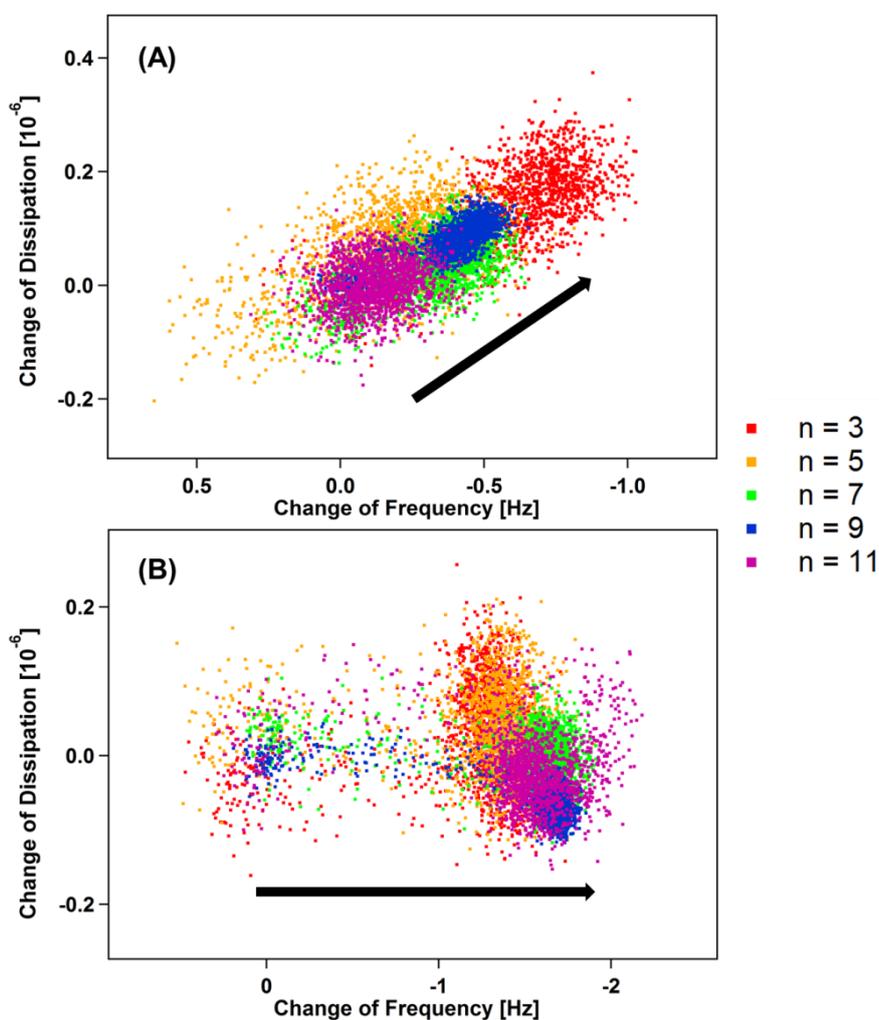


Figure 4.7: ΔF - ΔD plots representing the interaction of 5 μM APA-1 with (A) DMPC, and (B) DMPC:DMPG (2:1) membranes at pH 5.2. To assist with the interpretation, arrows representing the mechanism process are shown.

Figure 4.8 shows the ΔF - ΔD plots of APA-1 interacting with DPhyPC and DPhyPC:DPPG (2:1) membranes. Similar to Figure 4.7A, binding of APA-1 did not cause pronounced changes in the frequency or energy dissipation to the DPhyPC membrane (Figure 4.8A). The processes at the 3rd and 5th overtone and the 7th, 9th and 11th overtones present opposite results. The processes indicated by the 3rd and 5th overtones can be interpreted by a north-east arrow suggesting the addition of mass and a reduction in membrane rigidity, whereas the processes for the remaining overtones indicate a reduction in mass and an increase in membrane rigidity.

However, the frequency changes of the 7th, 9th and 11th overtones were at the lower end of the sensitivity of QCM-D and are thus not considered to be significant.

From the ΔF - ΔD plot of the DPhyPC:DPPG (2:1)-APA-1 interaction, three distinct turning points were observed suggesting that four mechanistic processes were involved (labelled as 1, 2, 3 and 4 in Figure 4.8B). The first process involved a mass gain through the entire thickness of the membrane without affecting the membrane rigidity, similar to what was observed for the interaction of APA-1 with the DMPC:DMPG (2:1) membrane. Immediately after the initial insertion, the observed direction of change indicates more mass is present on the membrane with a corresponding reduction in membrane rigidity. The processes of the five overtones do not overlap but rather they are divergent. The 3rd and the 5th overtones, which probe further from the crystal surface, show the greatest addition of mass, suggesting that more mass is associated towards the top of the membrane [230]. During the third process, represented by a north-west arrow, mass is rapidly lost from the membrane and the membrane became even less rigid. Similar frequency changes were observed for all overtones suggesting that APA-1 has disrupted the membrane across the whole bilayer. The lower overtones also showed a greater increase of dissipation, suggesting more pronounced loss of membrane rigidity. Finally, the fourth process involves very slow mass gain, more pronounced on the lower overtones. No noticeable change of energy dissipation was observed, indicating stabilised membrane rigidity. This process is consistent with a rearrangement of the protein and lipid remaining on the surface [239]. Since DPPG has a gel-liquid phase transition temperature of 41 °C [253], performing the interaction at room temperature would result in the coexistence of gel-phase DPPG and liquid-phase DPhyPC (gel-fluid phase transition temperature = -140°C [254]) lipids. This may lead to the formation of phase segregation between the lipid domains, thereby creating membrane packing defects. The defects may facilitate interaction of APA-1 with the DPhyPC/DPPG membrane. In addition, the interaction may also be facilitated by membrane defects formed by the mixture of PG and the steric hindrance covered by the methyl groups on DPhyPC. In contrast, the ΔF - ΔD plot for the APA-1-DPhyPC:DPPG (2:1) interaction at pH 8.5 shows no pronounced ΔF and ΔD changes (Figure 4.8C).



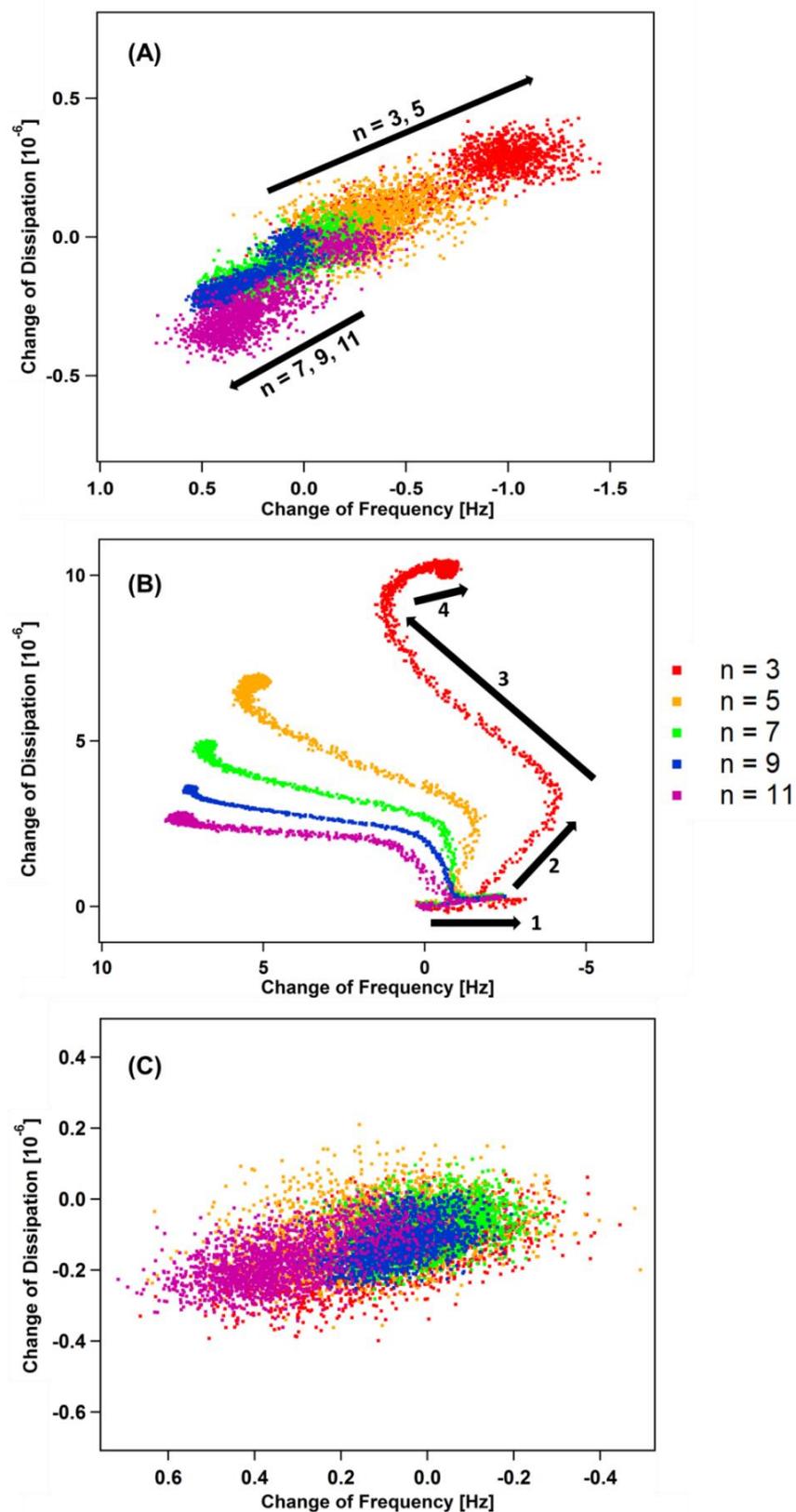


Figure 4.8: The ΔF - ΔD plots representing the interaction of APA-1 to (A) DPhyPC, (B) DPhyPC:DPPG (2:1) bilayers at pH 5.2, and (C) DPhyPC:DPPG (2:1) bilayer at pH 8.5. No pronounced changes in the signals at pH 8.5 were observed. To assist with the interpretation, arrows representing the number of processes are shown.

Moreover, it was found that APA-1 interacted strongly to loosely deposited phospholipids. Two DPhyPC membranes were prepared with and without using the 2 M NaCl rinsing step, which is critical for the formation of a complete DPhyPC bilayer. Consequently, two DPhyPC membranes were obtained with resonant frequencies at levels of -21.5 Hz and -42.5 Hz, suggesting the formation of a complete bilayer and a surface with loosely deposited phospholipids, respectively. Interactions of $5 \mu\text{M}$ of APA-1 with both membranes were investigated at pH 8.5 (Figure 4.9). No change in the resonant frequency of the complete DPhyPC membrane was observed upon addition of APA-1, consistent with the previous results showing that APA-1 does not associate with phospholipid at pH 8.5. In contrast, addition of APA-1 to the thick and loosely attached DPhyPC membrane resulted in the distinct removal of lipid from the surface. Interestingly, such disruption has led to the formation of a more complete DPhyPC bilayer reflected by a resonant frequency value equilibrated at -22.5 Hz. The corresponding $\Delta F-\Delta D$ plot of APA-1 interacting with the loosely packed membrane is shown in Figure 4.10. A single process was observed and represented by a south-west arrow. The remarkable increase of resonant frequency and decrease of energy dissipation suggest rapid loss of lipid material and increased membrane rigidity. The slightly separated traces indicate that more mass is removed towards the surface of the membrane. Clearly, the result has shown that the introduction of APA-1 has caused an instantaneous removal of the loosely attached DPhyPC from the surface and the membrane structure becomes more rigid and compact presumably because of rearrangement of the lipid remaining on the surface. Based on these results, it is proposed that APA-1 is capable of effectively binding to loosely attached phospholipids thus facilitating the removal of lipid material. The experiment was repeated using a DMPC:DMPG (2:1) membrane at pH 8.5 and the same result was obtained.

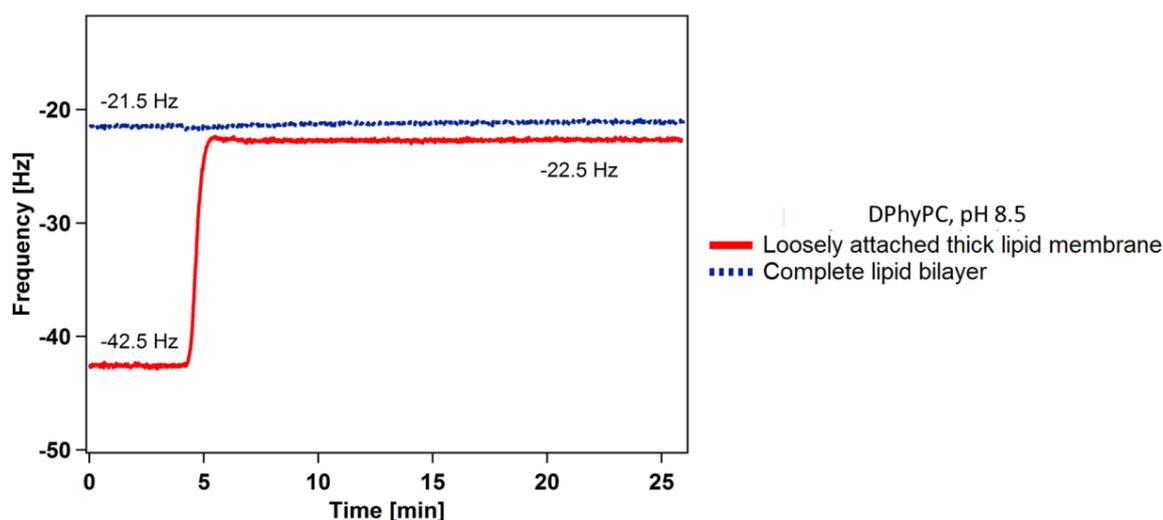


Figure 4.9: Time-dependent changes in the QCM frequency value of the DPhyPC membranes upon binding of 5 μM APA-1 at pH 8.5. The addition of APA-1 to loosely deposited phospholipids caused significant mass removal, leading to bilayer that is compact and stable. Time 0 is when APA-1 was introduced.

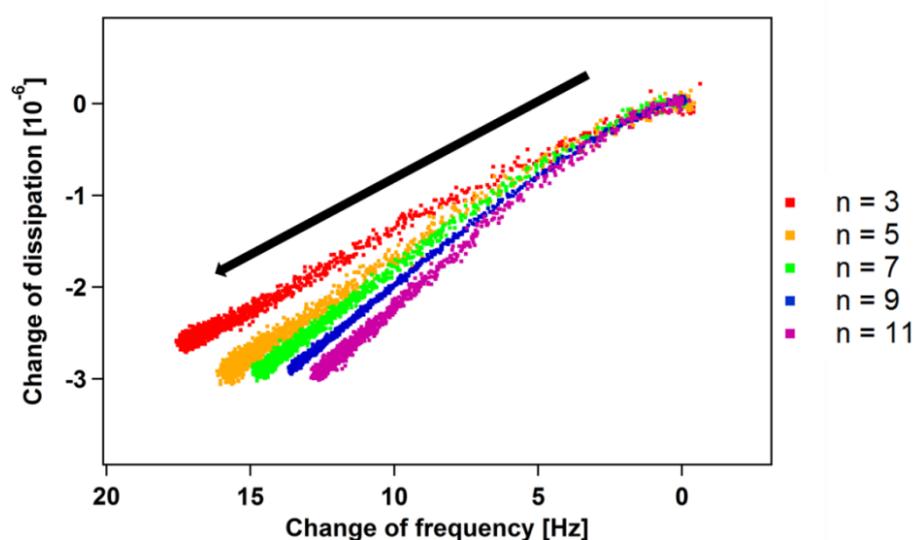


Figure 4.10: The ΔF – ΔD plots representing the interaction of APA-1 to loosely attached DPhyPC phospholipids at pH 8.5. To assist with the interpretation, arrows representing the mechanism process are shown. The binding caused significant mass removal and increased the rigidity of the membrane.

4.3.2 Neutron reflectometry

Based on the QCM-D results, neutron reflectometry experiments were performed to study the changes to model membrane properties caused by APA-1 association. By varying lipid compositions APA-1-membrane interactions were modulated. Seven different membrane models were used: DMPC, DMPC:DMPG (2:1), DPhyPC, DPhyPC:DPPG (2:1), POPC,

POPC:POPG (2:1) and an *E. coli* polar lipid extract on Si/SiO₂ supported surfaces. The POPC/PG membrane models and the membrane prepared from the *E. coli* polar lipid extract were chosen to study the impact of unsaturated membranes and natural lipids on APA-1-membrane interaction. APA-1 solutions were introduced to the membrane at pH 5.2. Figure 4.11 shows the measured and fitted reflectivity profiles and the corresponding SLD profiles before (blue) and after (red) incorporation of APA-1 to (A) DMPC and (B) DMPC:DMPG (2:1) membranes. Fitting the reflectivity curves of the membrane gave a membrane thickness of ~30 and ~35 Å, respectively, slightly thinner than the previously stated thickness [255]. Comparing the SLD profiles before and after binding of APA-1, it was observed that subtle mass adsorption was obtained for both membranes. This is consistent with the previous results obtained using QCM-D, which is a technique that is more sensitive to small mass changes when compared with neutron reflectometry.

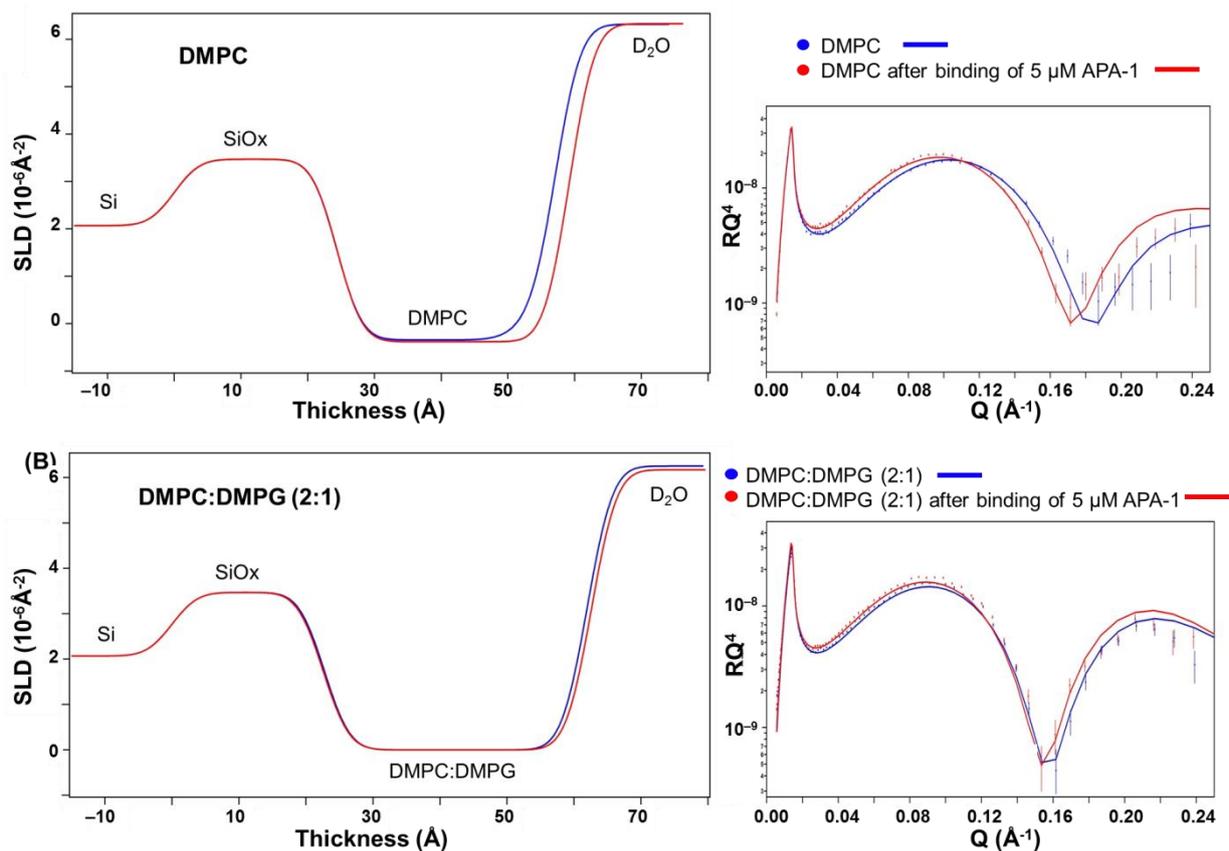


Figure 4.11: Reflectivity modelling of APA-1 binding to DMPC and DMPC:DMPG (2:1) membranes. (A) & (B) are the SLD and reflectivity profiles of the APA-1 interaction with DMPC and DMPC:DMPG (2:1) membranes, respectively. Data recorded in D₂O is presented. Subtle mass adsorptions were observed on both membranes.

Figure 4.12 shows the reflectivity modelling of the interactions between APA-1 and (A) DPhyPC and (B) DPhyPC:DPPG (2:1) membranes. The membrane thickness were modelled to be ~ 31 and ~ 34 Å. Noticeably, the DPhyPC:DPPG (2:1) membranes were found to have defects, reflected by the membrane SLD value that is slightly above the average SLD value of lipid membranes (*i.e.*, $\sim -0.4 \cdot 10^{-6} \text{ \AA}^{-2}$ [256]). This indicates the presence of solvent in the membrane. The defects are likely due to lipid domains formed by mixing the gel-phase DDPG and liquid-phase DPhyPC lipids, as described in Section 4.3.1. Near identical SLD profiles were obtained for the DPhyPC membrane in the absence and presence of APA-1, suggesting that negligible changes had occurred to the membrane when APA-1 was present. This observation showed that APA-1 does not interact with the DPhyPC membrane. In contrast, injection of 5 μM APA-1 has caused the most pronounced membrane disruption to the DPhyPC:DPPG (2:1) membrane. Based on the SLD profile, the thickness of the membrane was modelled to have decreased by ~ 6 Å. Intriguingly, binding of APA-1 has reduced the solvent content which was originally present in the bilayer, suggesting a more compact packing of the membrane.

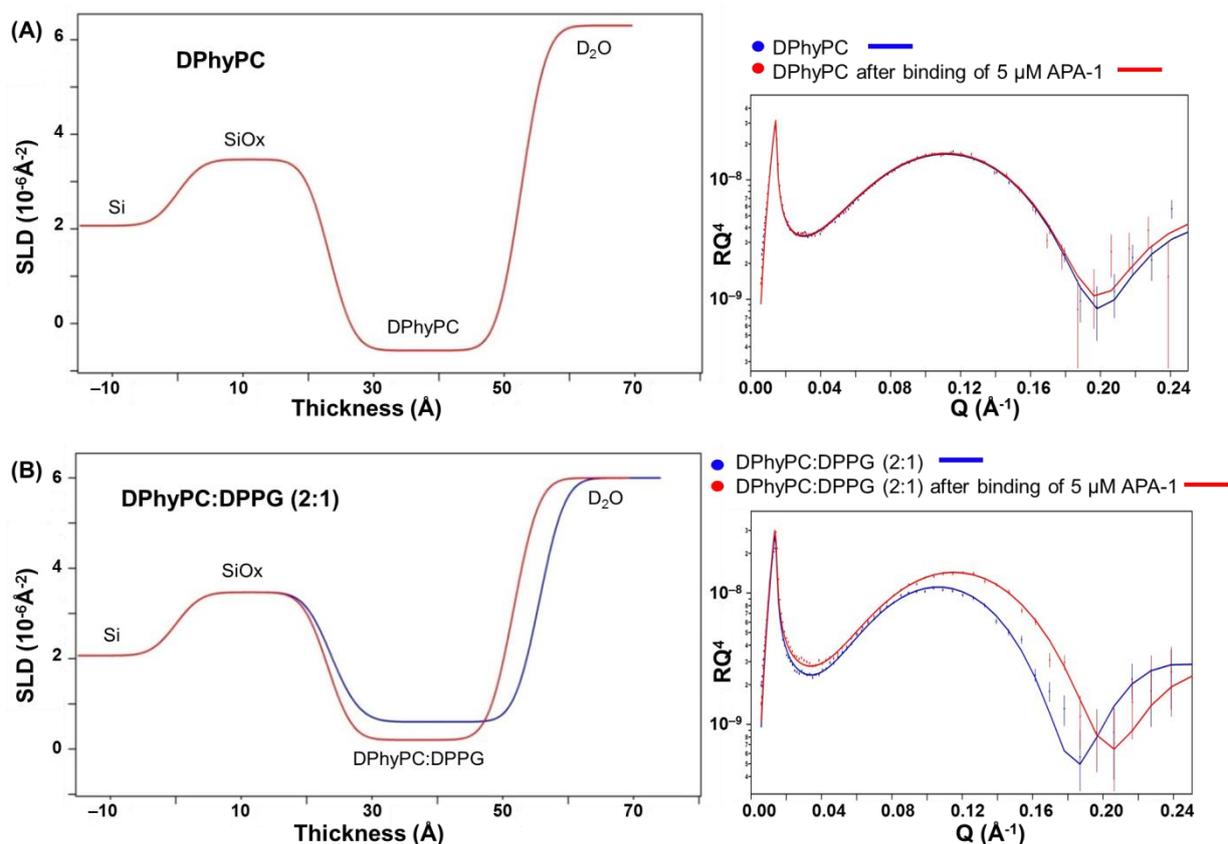


Figure 4.12: Reflectivity modelling of APA-1 binding to DPhyPC and DPhyPC:DPPG (2:1) membranes. (A) & (B) are the SLD and reflectivity profiles of the APA-1 interaction with DPhyPC and DPhyPC:DPPG (2:1) membranes, respectively. Data recorded in D₂O is presented. Disruption and material removal from the DPhyPC:DPPG (2:1) membrane was observed upon injection of APA-1.

The interaction of APA-1 with POPC and POPC:POPG (2:1) membranes were examined. Thickness of ~ 36 Å and ~ 38 Å were respectively obtained for the POPC and POPC:POPG (2:1) membranes. Figure 4.13 shows that APA-1 did not cause any change to the POPC membrane and subtle mass adsorption was observed for the POPC:POPG (2:1) membrane, similar to the results obtained for the DMPC:DMPG (2:1) membrane.

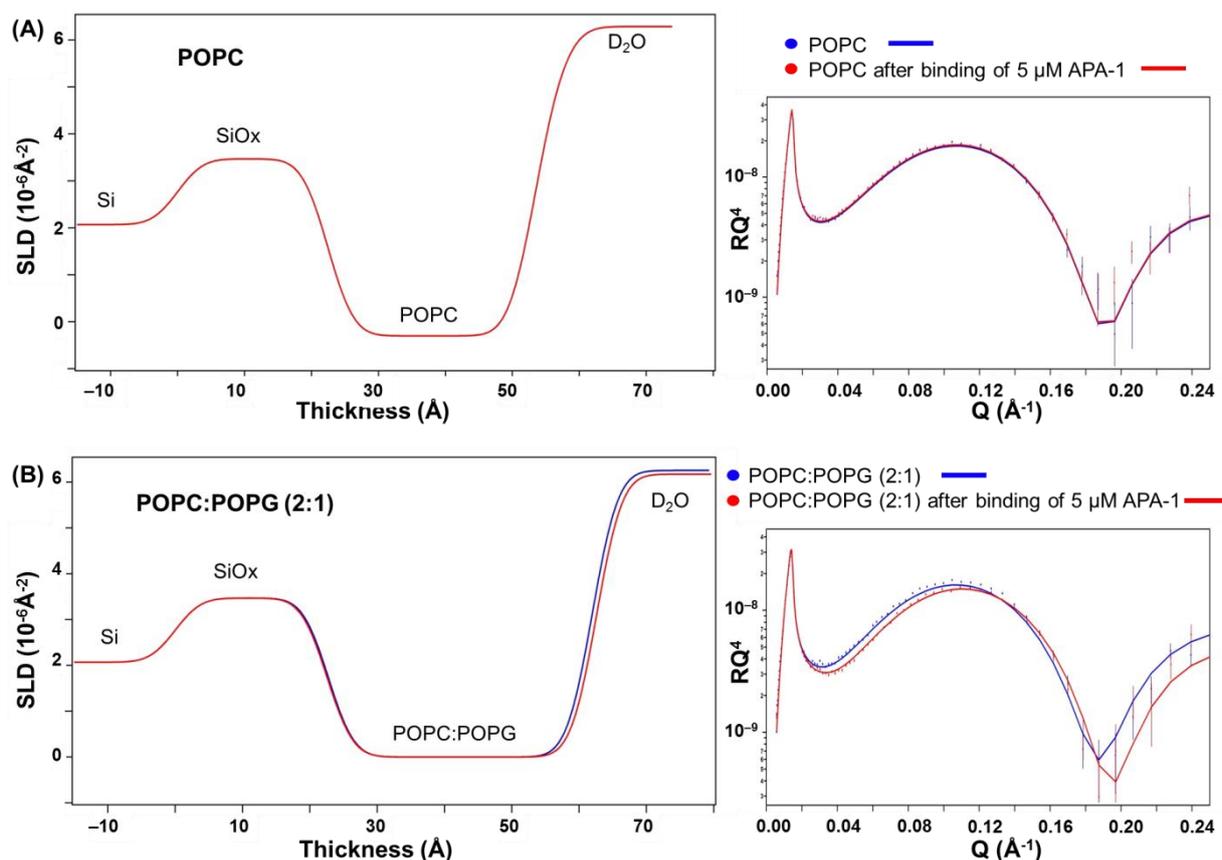


Figure 4.13: Reflectivity modelling of APA-1 binding to POPC and POPC:POPG (2:1) membranes. (A) & (B) respectively represent the SLD and reflectivity profiles of APA-1 interaction with POPC and POPC:POPG (2:1) recorded in D₂O. Subtle mass adsorption was estimated on the POPC:POPG (2:1) membrane upon introduction of APA-1.

An *E. coli* polar lipid extract which contains $\sim 33\%$ negatively charged lipids (67% PE, 23.2% PG and 9.8% CA (w/w)) was used to study the incorporation of APA-1. Interestingly, it was observed that the *E. coli* polar lipid extract had spontaneously formed a Gram-negative-like bacterial membrane, in which, two layers were formed and separated by a thick layer of solvent. The outer layer was modelled to be a loosely attached and thin layer with poor surface coverage, whereas the inner layer appeared to be more complete and has a membrane thickness modelled to be ~ 41 Å. The distance between the two layers was modelled to be ~ 180 Å, close to the suggested value in the literature (~ 160 Å) [257]. Upon addition of 5 μM APA-1, the

outer lipid layer was completely removed and another $\sim 8 \text{ \AA}$ of lipids were modelled to be removed from the inner membrane. The SLD of the inner membrane was increased upon APA-1 addition, indicating the membrane had become more solvent accessible. Introduction of another 10 \mu M APA-1 further disrupted the membrane and increased the membrane solvent content.

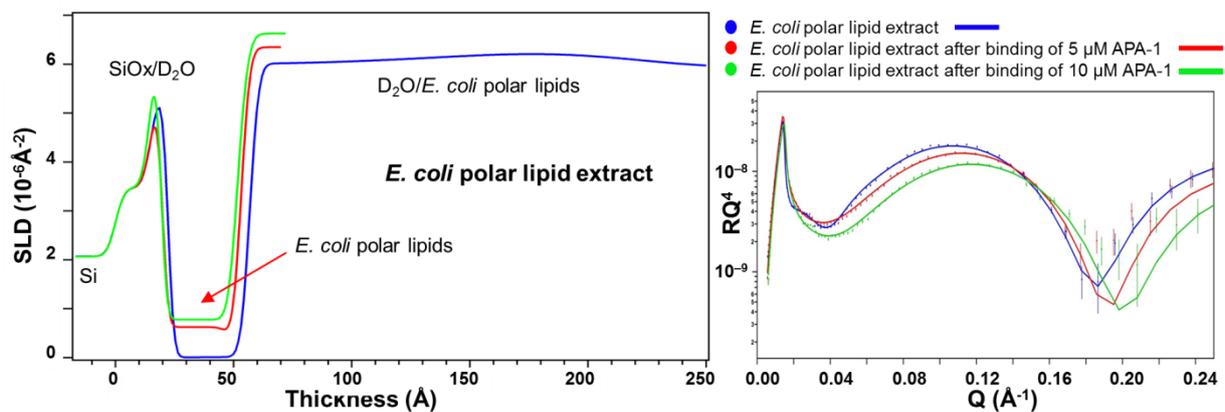


Figure 4.14: Reflectivity modelling of APA-1 in D_2O binding to membrane made from *E. coli* polar lipid extract.

4.4 Discussion

4.4.1 Interactions of amoebapore A with different model membranes

Generally, zwitterionic PC and anionic PG lipids are used as models for eukaryotic and bacterial membranes, respectively [258]. It was shown that APA-1 bound more strongly to membranes containing PG lipids at pH 5.2, such as DMPC/DMPG, DPhyPC/DPPG and POPC/POPG bilayers, supporting the hypothesis that positively charged regions on APA-1 are important for lytic activity by facilitating the interaction with negatively charged membranes [30-32]. Such a binding preference is known for a variety of AMPs, such as magainins [38] and sarcotoxin IA [259].

Binding of APA-1 with the neutral PC membranes were also detected, although to a much lower extent when compared with the PC/PG membranes. This interaction may be primarily driven by hydrophobic interactions between the hydrocarbon acyl chains of the membranes and the amphiphilic α -helical structure of APA-1 [24], which is a common feature for many membrane binding proteins. Although APA-1 interactions with the neutral lipids were observed to be much weaker than with anionic PG lipids [8], APA-1 was found to be highly potent against mammalian cells which are predominantly composed of zwitterionic lipids. This indicates that other complex events may be involved during the interaction in the complicated cellular membrane system which cannot be easily mimicked by simple model membranes.

APA-1 interactions with DMPC/DMPG, DPhyPC/DPPG and POPC/POPG membranes were compared. Although all three membranes are negatively charged, APA-1 was observed to have the largest impact on the architecture of the DPhyPC/DPPG membrane. Based on the ΔF - ΔD plots, it was clearly shown that APA-1 does not function in the same manner with different membranes. Upon addition of APA-1, a one-phase mechanism (*i.e.*, mass adsorption) was observed for the DMPC/DMPG membrane (Figure 4.8B) whereas a four-phase process (*i.e.*, symmetrical mass adsorption, asymmetrical mass adsorption/removal, membrane disruption and membrane structure rearrangement) was detected for the DPhyPC/DPPG membrane (Figure 4.9B). The more significant effect of APA-1 on the interaction with DPhyPC/DPPG may be explained by the coexistence of gel-phase DPPG and liquid-phase DPhyPC lipids, which may exhibit phase segregation between the gel- and liquid-domains. Such lipid domain boundaries present in the bilayer are known to cause membrane packing defects, which may facilitate the interaction of APA-1 with the membrane [260-262]. Similar observations have

been obtained in other studies on proteins binding to membranes that had packing defects. For example, DMPC and DMPC/PG membranes with defects enhanced binding of aurein 1.2 and caerin 1.1 [260], and POPC membranes that had a phase boundary in lipid domains were found to facilitate the pore forming activity of cobra CTX A3 [261]. In addition, in the DPhyPC/DPPG membrane, the mixture of PG and the steric hindrance caused by the methyl groups in DPhyPC may also cause defects within the bilayer that aided access for APA-1 binding. On the other hand, the DMPC/PG and POPC/PG lipids would give rise to more ordered membranes that are not easily accessible for interaction by APA-1.

4.4.2 Membrane binding mechanism of amoebapore A

It has been proposed that APA-1 permeabilises membranes via a pH-dependent dimerisation event and accumulation of the dimers on the cell membrane leads to trans-membrane pore formation [24]. The SEC results (Section 3.3.6) showed that APA-1 has a weak monomer-dimer equilibrium dissociation constant ($K_d = 247 \pm 32 \mu\text{M}$ at pH 5.2). This suggests that at 5 μM , ~96% of APA-1 molecules are present in the monomeric state (Eq.3.18). Both QCM-D and neutron reflectometry results showed that APA-1 at a concentration of 5 μM was able to interact with membranes, suggesting: (1) monomeric APA-1 is able to interact with membranes prior to dimerisation, or (2) the protein builds up a local high concentration close to the membrane surface that leads to dimer formation and then membrane binding.

The ΔF - ΔD plots showed that the introduction of APA-1 caused subtle mass adsorption and minor loss of rigidity on the DMPC, DMPC/DMPG and DPhyPC membranes. For DMPC and DPhyPC membranes, the mass adsorption primarily occurred on the membrane surface. In comparison, mass adsorption on the DMPC/DMPG membrane was observed across the entire membrane thickness without disturbing membrane rigidity, as indicated by the consistent frequency and dissipation changes at different overtones. Similar behaviours were observed in a QCM-D study of oncocin, an AMP which is isolated from *Oncopeltus fasciatus*. The protein performs a non-lytic antimicrobial activity by freely penetrating bacterial membranes [263], suggesting APA-1 is able to traverse the DMPC/DMPG membrane with little requirement of energy, presumably because there is no conformational change. In addition, it is likely that no higher MW pore structure was formed in the membrane as this would otherwise lead to disruption of membrane integrity, which can be reflected by separated dissipation changes at different overtones. To exert more invasive antimicrobial activity, a threshold APA-1 concentration may be required. The threshold concentration is required to perform lytic activity

for many AMPs [264]. For example, alamethicin, a pore forming peptide produced by fungus *Trichoderma viride*, initially associates parallel to the membrane surface, once a threshold concentration is reached, the protein reorientates and inserts into the bilayer as a pore [264].

In contrast, APA-1 interaction with the DPhyPC:DPPG (2:1) membrane resulted in a more pronounced membrane disruption. The $\Delta F-\Delta D$ plot showed that mass adsorption was initially observed at all overtones without affecting the membrane rigidity, suggesting a vertical transmembrane insertion. The interaction was found to be reversible. Following the first step, an asymmetrical mass gain and removal occurred. More mass was adsorbed towards the membrane surface and material was removed from deeper into the membrane, both events caused minor reduction of the membrane rigidity. Immediately after the process, a remarkable mass removal was observed at all overtones with more pronounced loss of membrane rigidity toward the membrane surface. The traces at different overtones were well separated in the $\Delta F-\Delta D$ plot, consistent with an asymmetrical membrane disruption. Similar behaviour was observed for maculatin 1.1, which is an AMP secreted from the skin of Australian tree frogs. Maculatin 1.1 was found to switch from a transmembrane insertion to a lysis mechanism once a threshold peptide concentration on the membrane is reached [230]. However, no evidence of pore-formation could be provided. At the end, subtle mass adsorption was observed on the membrane surface and the membrane became more stabilised with no more change in the rigidity. This process is consistent with rearrangement of the protein and lipid remaining on the surface [239]. Consistent with the QCM-D results, the neutron reflectometry profile showed membrane disruption of the DPhyPC:DPPG (2:1) membrane upon APA-1 interaction.

As previously described (Section 1.2.5), APA-1 is more potent against Gram-positive than against Gram-negative bacteria [30-32]. Based on the neutron reflectometry profile, it was observed that similar to the natural Gram-negative bacteria membrane, deposition of the *E. coli* polar lipid extract lead to the formation of a bilayer membrane with an additional thin film of loosely attached phospholipids. This layer was separated from the membrane by a thick layer of solvent. Upon addition of APA-1, the outer layer was completely removed and the inner membrane was also disrupted, thereby reflecting lipid removal. This is in agreement with the previous hypothesis, that the presence of external wall structure in the Gram-negative bacteria membrane acts as a barrier with higher resistance against APA-1 invasion [31]. Initial lytic activity of APA-1 against Gram-negative bacteria involves permeabilisation of its outer membrane. Without the presence of complete outer membrane APA-1 is able to attack the cytoplasmic membrane which is fatal for bacteria. Binding of APA-1 to the membrane caused

increase of solvent content of the membrane, reflected by increased SLD value. However, it is unclear whether the increase of solvent was because of APA-1 aqueous pore formation or disruption of membrane packing which allowed passage of D₂O into the membrane. To further characterise APA-1 pore formation and investigate the relative location and amount of APA-1 bound to the membrane using neutron reflectometry, selective labelling of APA-1 and/or the membrane by deuterium can be performed [240-242, 245].

No interaction of APA-1 with membranes (*i.e.*, DMPC:DMPG and DPhyPC:DPPG) were observed at pH 8.5, in agreement with the liposome depolarisation assay (Section 1.2.4), which showed that APA-1 membrane association was lost at high pH values (*i.e.*, ≥ 7) [32]. This may be because of deprotonation of acidic and His residues which lead to reduction of the overall electrostatic surface of APA-1 and hampered the interaction of APA-1 with anionic lipids, as discussed in Section 3.4.2. In contrast, when loosely deposited phospholipids were present, interaction of 5 μ M of APA-1 resulted in a remarkable loss of material from the membrane. Surprisingly, this process was found to facilitate the formation of more complete membrane bilayers, as shown in the QCM-D results. The corresponding ΔF - ΔD plot showed a one-phase mechanism of this process which involves an instantaneous mass removal from the membrane and a dramatic increase in membrane rigidity. Based on the observations, it is proposed that APA-1 is capable of effectively binding to loosely attached phospholipids thus facilitating removal of lipid material. This process appeared to be largely independent of pH, suggesting it is likely to be primarily driven by hydrophobic interaction.

From the study it is shown that the activity of APA-1 is pH- and membrane property-dependent. Different lipid headgroup charges, composition, and membrane packing can have impact on APA-1 binding. The results indicate that APA-1 is capable of inserting into membranes and loss of lipid material. However, no clear evidence of pore formation can be provided. To further study the pore-forming activity of APA-1, much higher protein concentrations are required. In addition, a tethered bilayer system can be adopted that is designed to create an aqueous reservoir between the solid substrate and the membrane [265], so that the transmembrane APA-1 aggregation (*i.e.*, pore) can be extended beyond the membrane, thereby aiding the study of the interaction between membranes and APA-1. Moreover, APA-1 mutants which are hypothesised to inhibit APA-1-membrane interaction can be designed for further understand the APA-1-membrane interaction mechanism.

5

Conclusions

The effectiveness of conventional antibiotics is being reduced because numerous bacterial strains are gaining resistance and this number of bacterial species will continue to grow. Thus, the development of new antimicrobial agents is urgently needed. AMPs are key components of innate immunity that protect organisms from microbial infection [81, 266]. They have less susceptibility to the development of bacterial resistance [81]. Understanding the mode of action of AMPs and their interaction with membranes is useful for the therapeutic development of new antimicrobial analogues. In this project, the activity mechanism of APA-1, an AMP whose activity is regulated by a pH-dependent process, was investigated. The main aims were to investigate the APA-1 dimerisation and the APA-1-membrane binding mechanism. The information arising from the research has provided a better understanding of the molecular mechanism of APA-1 activity and aids our general understanding of AMP-membrane interactions.

5.1 Mechanism of amoebapore A dimerisation

Cytotoxic APA-1 was produced using a cell-free protein synthesis approach that does not depend on cell viability [83-87]. Isotope labelling was adapted to the system for the structural analysis of APA-1 dimerisation using NMR spectroscopy. The NMR experiments showed that APA-1 dimerisation is both pH- and concentration-dependent and a new APA-1 dimer model was proposed based on the NMR chemical shift perturbation analysis. The side-chain pKa of E2 showed the strongest dependence on protein concentration compared with other ionisable residues studied, indicating that E2 is located at the dimer interface, in agreement with the dimer model. Nonetheless, the NMR data points to weak self-association. The NMR results were supported by the K_d estimation using NMR spectroscopy and SEC experiments. An accurate K_d value at pH 3.0 was not measured because of possible protein aggregation and severe NMR resonance line-broadening at the higher protein concentrations. Such line-broadening did not permit a full titration curve analysis and only a lower-limit K_d value was obtained at pH 3.0. Modelling of the SAXS data was consistent with an APA-1 monomer at low concentration (190 μ M) and pH 3.0, whereas a model with dimensions close to the proposed new dimer model (based on the NMR analysis) was achieved at the highest concentration (2.1 mM) and pH 5.2. The result is consistent with the K_d estimation. Nonetheless, although the proposed dimer model allows the identification of certain amino acid side-chains responsible for dimerisation, the role of the dimer in APA-1 activity has not been determined. Ten mutants which were designed to either stabilise or disrupt the APA-1

dimerisation, and thus activity, were designed. Liposome depolarisation assays using the mutants should pinpoint which residues are important for APA-1 activity, and coupled with biophysical studies, the potential role of dimerisation to membrane activity be resolved.

The side-chain pKa values of acidic and His residues were determined using NMR experiments at two APA-1 concentrations. Based on the results, it is proposed that the ionisation states of particular charged amino acids are likely to modulate the pH-dependent activity of APA-1, as presented in Figure 5.1 below.

Figure 5.1 (A) and (B): APA-1 exists in monomer-dimer equilibrium with weak association. At low pH, protonation of the ionisable residues results in APA-1 with a net positive electrostatic surface charge. Deprotonation of residue E2 which is located at the dimer interface facilitates APA-1 dimerisation via the formation of a salt-bridge with K37.

Figure 5.1 (C): Both monomeric and dimeric APA-1 species may interact with the anionic membrane via the positively charged surface of the protein. This interaction has been shown to involve lysine residues, as chemical modification leads to a loss of membrane binding activity of APA-1 [32]. A few Asp residues which have suppressed side-chain pKa values are proposed to form salt bridge with their neighbouring Lys residues, contributing to extra stabilisation of the protein fold [228]. Mutagenesis of Lys residues would provide further insights into the role of these residues in APA-1 function. Deprotonation of the Asp residues reduced the overall positive surface charge of APA-1.

Figure 5.1 (D): When a critical APA-1 concentration is reached, APA-1 forms a pore structure in the membrane. However, information of the pore structure and whether APA-1 dimerisation is a prerequisite for the formation of the pore remains unresolved. APA-1 has the highest activity at pH 5.2.

Figure 5.1 (E): It has been previously observed that APA-1 abruptly loses pore-forming activity at pH values between 5.2–5.5 [32], close to the side-chain pKa of E16 and E62 (pKa = 5.3), suggesting these residues are the primary determinant for APA-1 activity. Deprotonation of E16 and E62 further reduces the APA-1 electrostatic charge. This may lead to disruption of the pore structure and reduction of APA-1-membrane binding affinity. Consequently, neutralisation of the Glu side-chains may be essential to trigger APA-1 activity. The same Glu residues are structurally conserved in other SAPLIP members that have pH-dependent activity, such as saposin C and caenopore 5 [75].

Figure 5.1 (F): A further increase of the pH leads to the deprotonation of H75, which has an elevated side-chain pKa of 7.3, close to the pH where APA-1-membrane association is completely lost. It has been previously shown that chemical modification of H75 resulted in loss of APA-1 activity [32], suggesting that H75 is important for APA-1 activity. This charge-type amino acid is conserved in the sequence of other SAPLIP members such as NK-lysin and granulysin, whose membrane binding activities are pH-independent. However, the His residue is replaced with a Lys and Arg in NK-lysin and granulysin, respectively, which has much higher intrinsic pKa values. This suggests that the ionisation state of H75 is important in modulating APA-1 activity. Overall, the requirement on the (de)protonation states of the ionisable residues limits APA-1 activity to a relatively narrow pH range, and presenting a unique mode of action of APA-1.

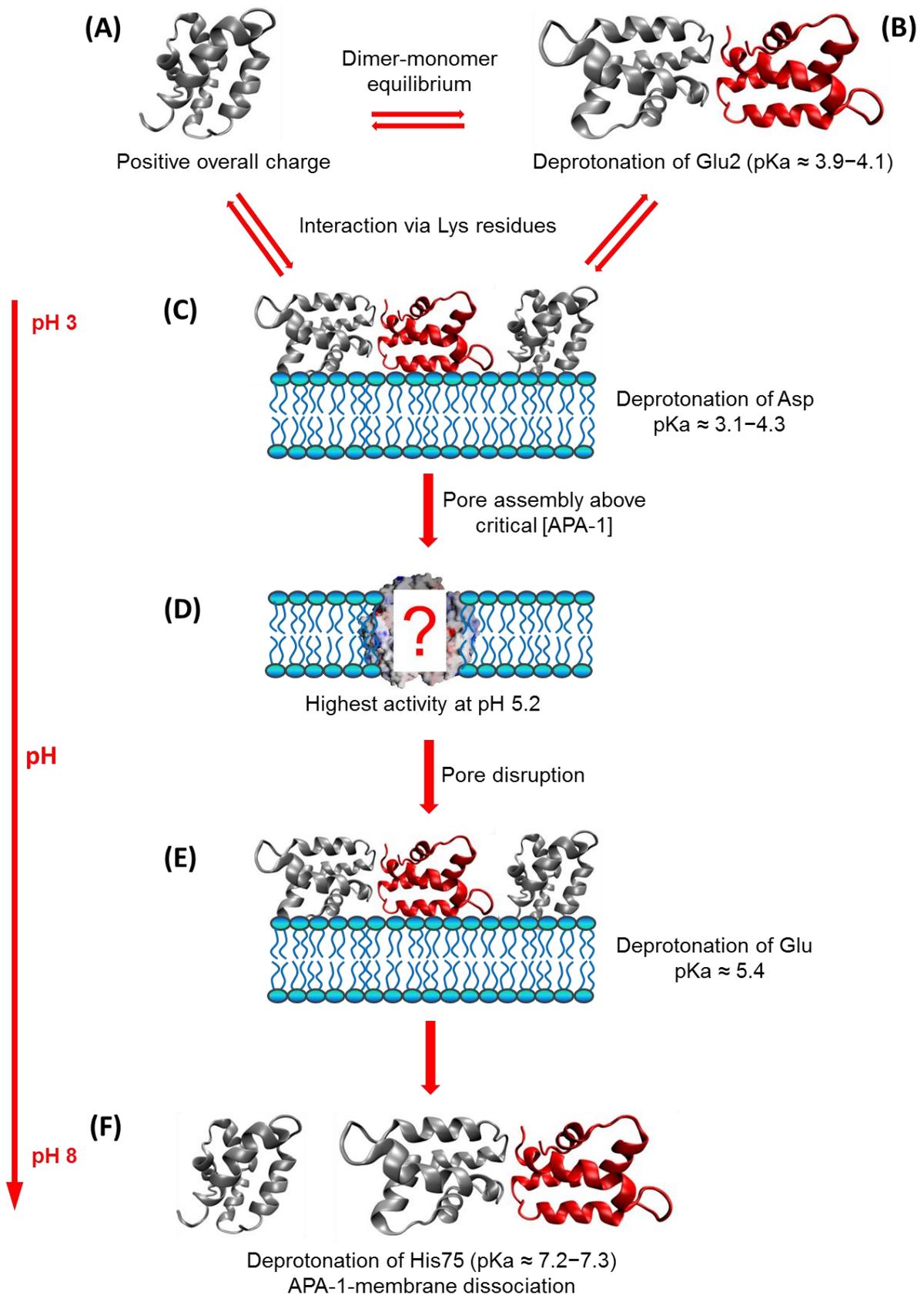


Figure 5.1: Schematic diagram representing the proposed mechanism of the pH-dependent APA-1 activity.

5.2 Amoebapore A interaction with model membranes

Since biological membranes are complex systems to study and prepare *in vitro*, the interaction of APA-1 with membranes was assessed using the model membranes. Model membranes were prepared using single and mixed lipids and cellular lipid extracts. Using such an approach enabled the identification of what properties of lipids were important in defining the mode of action of APA-1. The mechanism of APA-1 adsorption at low concentration (5 μM) and pH 5.2 onto membranes was elucidated. The QCM-D and neutron reflectometry results showed that the membranes responded differently to the introduction of APA-1 and these signatures are characteristic of the membrane property differences.

- (1) APA-1 bound to membranes at 5 μM at pH 5.2. Based on the K_d estimation, APA-1 exists primarily as a monomer at 5 μM , suggesting: (1) monomeric APA-1 is able to bind to the membranes without dimerisation, or (2) APA-1 is able to accumulate to a local high concentration proximate to the membrane and lead to dimerisation and membrane binding. The study was not performed at pH 3.0 where APA-1 was proposed to be inactive. This is because a QCM-D study showed that a stable bilayer was not formed at pH 3.0. In addition, no scientific evidence is available describing the activity of APA-1 at lower pH values (*e.g.*, pH values < 5.2). Consequently, it remains unclear whether APA-1 is at active pH values below 5, even though there is negligible dimer formation.
- (2) Similar to the majority of AMPs, APA-1 bound more strongly to negatively charged membranes than neutral membranes at pH 5.2, supporting the hypothesis that positively charged regions on APA-1 enhance the binding with anionic membranes [30-32].
- (3) APA-1 was also found to interact with the neutral membranes and such interactions led to subtle mass adsorption onto the membrane surface. This may be primarily driven by hydrophobic interactions, suggesting the amphipathicity of APA-1 helices are important in function. This may contribute to APA-1 lytic activity against mammalian cells.
- (4) Intriguingly, APA-1 appeared to have transmembrane insertion into the DPPC:DPPG membrane without altering the membrane rigidity, suggesting that at low protein concentration APA-1 is able to insert into membranes in a low, non-oligomeric state (*e.g.*, monomer). Similar effects have been observed in non-lytic AMPs such as apidaecin 1a and 1b [267] and oncocin [263]. A critical concentration may be required for APA-1 to perform membrane disruption.

- (5) APA-1 showed the most pronounced influence on an incomplete DPhyPC:DPPG membrane, which likely consisted of a mixture of gel- and lipid-phase lipids, suggesting that the presence of membrane defects enhanced the activity of APA-1.
- (6) No binding of APA-1 to anionic membranes was observed at pH 8.5, in agreement with our hypothesis that the ionisation states of charged residues are able to modulate the pH-dependent activity of APA-1. Deprotonation of acidic residues and H75 reduce the overall electrostatic surface of APA-1, thereby may lead to abolishing of the interaction of APA-1 with anionic membranes.
- (7) APA-1 was able to effectively bind to loosely attached phospholipids and facilitate removal of lipid materials, even at pH 8.5 where it is inactive against intact membranes. Based on the QCM-D data, such lipid removal may have led to rearrangement of lipids that resulted in the formation of more complete membranes. This pH-independent manner suggests that the interaction is primarily driven by hydrophobic interactions.

Several structures of SAPLIP proteins have been reported. Although the proteins share the characteristic saposin-like fold, they perform a highly diverse range of biological functions that are associated with interactions with lipids [69]. Besides the minor differences in the arrangement of the helices, the major attribution for various functions of these proteins may be due to their electrostatic surfaces. In contrast to other SAPLIP members, the (de)protonation states of the ionisable residues appears to limit the activity of APA-1 to a relatively narrow pH range, and APA-1 displays a more complicated activity mechanism, in which different events may take place (*i.e.*, dimerisation and pore formation). However, it is unclear whether dimerisation is crucial for APA-1 activity or is an event which is unique for APA-1 but not functionally important. The results herein provided no clear evidence that APA-1 forms a pore, even though conductance measurements have indicated that APA-1 forms pores with various diameters [49]. Nonetheless, these conductance results do not provide information of the pore structure and it is unclear whether APA-1 activity is solely relying on pore formation, particularly under conditions similar to those encountered *in vivo*. To further characterise the APA-1 pore structure, electron microscopy and solid-state NMR spectroscopy can be used. Electron microscopy allows direct visualisation of the pores in membranes. However, crystallising APA-1 pores within membranes and visualising the pore structure by electron microscopy can be difficult because of the relatively small APA-1 pore size and weak interaction [53]. Solid-state NMR spectroscopy can be used to characterise the APA-1 pore

structure at the atomic level. However, particular sample conditions, such as sufficient protein materials, ultrahigh magnetic fields, sensitive techniques (*e.g.*, magic angle spinning), are required for collecting suitable high-resolution data [268] .

In addition, site-directed mutagenesis (*e.g.*, E16, E62 and H75) can be performed which should aid identification of residues that are crucial for APA-1-membrane interaction and pore formation. Finally, human and bacterial cells can be used to validate the results obtained with simple model membranes. However, application of biophysical techniques with cells can be a challenge; for instance, other complex events may be involved during the interaction in the complicated cellular membrane system that cannot be easily mimicked.

Appendices:

Appendix 1: SDS-PAGE gel electrophoresis

Protein samples were analysed by 4–12% SDS-PAGE. Ten microliters were taken from each sample and mixed with 4 x SDS-PAGE loading buffer (50 mM Tris-HCl (pH 6.8), 2% SDS (w/v), 10% glycerol (v/v), 1% β -mercaptoethanol (w/v), 12.5 mM EDTA (w/v), 0.02% bromophenol blue (w/v)). Samples were heated at 99 °C for 2 min, mixed well, and loaded onto the gel. All the SDS-PAGE electrophoresis experiments were performed at 200 V for 30 min using 1x MES SDS running buffer (Invitrogen, Carlsbad, USA). To visualise proteins, gels were stained in 0.1% coomassie blue solution and detained in H₂O using microwave oven for 2 min at 800 W followed by 6 min at 190 W. Mark 12™ standard protein marker (Invitrogen, Carlsbad, USA) was used for estimating the protein MWs.

Appendix 2: Method of packing cellulose phosphate affinity column

To prepare for column media, 3 g of cellulose phosphate was weighed, stirred into 75 ml of 0.5 M NaOH and left for 5 min to settle. The supernatant was discarded and the cellulose phosphate media was transferred into a funnel/filter and washed with ultra-pure water until the pH of the flow through was below 11. The media was then stirred into 75 ml of 0.5 M HCl and left for 5 min to settle. The supernatant was discarded and the media was washed in a funnel/filter with ultra-pure water until the pH of the flow through reached above 3.0. The media was transferred into 60 ml of 0.5 M potassium phosphate (pH 7.5), stirred, and allowed to settle (the pH should be 7.5). The supernatant was decanted and the media was stirred into 60 ml of buffer P (10 mM potassium-phosphate (KPi), pH 7.5 and 0.1 mM EDTA) and left for 4 min to settle. The supernatant was discarded. The last two processes (*i.e.*, washed with potassium phosphate and buffer P) were repeated for a further four times.

To prepare for the cellulose phosphate affinity column, 60 ml of buffer P was added to the media, mixed and poured into the C10/10 column (GE Healthcare, Little Chalfont, UK) which has been stoppered. The column was left overnight to settle at room temperature. The column was connected to a gradient ÄKTA system (GE Healthcare, Little Chalfont, UK). The column was ready for purification after equilibrated with five column volumes of buffer P.



Appendix 3: Assignments of amoebapore A resonances acquired in different NMR experiments at pH 3.0.

APA-1 concentration			200 μ M	200 μ M	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM
Residue	Atom	Published assignment	2D HSQC	3D HNCA	3D HCCH-TOCSY	HCCTHSQC	2D HBHANH	2D CBCANH	3D HCACO	3D HHN-TOCSY	2D HCCO	2D HCCO TSP correction
G1												
E2	H _N	8.806	8.786				8.775			8.776		
	N		121.33				121.29			121.36		
	C α				57.60	57.52						
	H α	4.202			4.231	4.231	4.231			4.248		
	C β				29.30	28.69						
	H β *				2.011	2.008	2.039			2.040		
	H β *				2.082	2.080						
	C γ				33.77				34.20			
	H γ *				2.431				2.448	2.452	2.446	2.326
H γ *				2.471				2.476		2.481	2.364	
C=O								181.27		181.01	183.07	
I3	H _N	8.040	8.031									
	N		120.06									
	C α											
L4	H _N	8.085	8.063									
	N		120.82									
	C α											
C5	H _N	7.989	8.009	7.995								
	N		116.96	117.05								
	C α			57.83								
N6	H _N	8.354	8.332	8.330			8.300			8.314		
	N		118.88	118.96			118.74			118.74		
	C α			55.73	55.87							
L7	H _N		8.330	8.331								
	N		121.82	122.03								
	C α			57.67								
C8	H _N	8.537	8.506	8.507								
	N		119.47	119.34								
	C α			61.59								

T9	H _N	9.129	9.115	9.112								
	N		116.18	116.27								
	C α			66.91								
G10	H _N		7.945	7.944								
	N		108.99	108.99								
	C α			47.35								
L11	H _N	8.082	8.068	8.062								
	N		124.43	124.64								
	C α			58.03								
I12	H _N	8.827	8.824	8.825								
	N		117.92	118.14								
	C α			62.78								
N13	H _N	8.506	8.022	8.023			8.025	8.023		8.025		
	N		118.49	118.66			118.43			118.45		
	C α			57.31	56.97			57.08				
	H α				4.450		4.477			4.451		
	C β				38.60	38.23			38.71			
	H β^*				2.822	2.821	2.805		2.822	2.802	2.826	2.719
	H β^*				2.940	2.936	2.942		2.938	2.945	2.936	2.812
	C γ								178.35		178.33	180.37
T14	H _N	8.082	8.078	8.068								
	N		119.44	119.55								
	C α			67.21								
L15	H _N	8.928	8.919	8.912								
	N		120.74	120.88								
	C α			57.95								
E16	H _N	8.970	8.944	8.949			8.944	8.928		8.940		
	N		119.36	119.00			119.33			119.25		
	C α			59.59	59.59	59.51		59.72				
	H α	3.882			3.866	3.867	3.886			3.868		
	C β				27.52	27.46		27.54				
	H β 1	2.234								2.237		
	H β 2	2.228			2.224	2.222	2.217					
	C γ				34.39				34.14			
	H γ 1	2.666			2.660				2.671		2.670	2.552
	H γ 2	2.494			2.492				2.492		2.502	2.380
C=O								180.20		180.16	182.22	

N17	H _N	7.456	7.430	7.429			7.434			7.430		
	N		117.65	117.80			117.54			117.54		
	C α			56.61	56.50							
L18	H _N	8.159	8.140	8.142								
	N		121.84	121.94								
	C α			57.60								
L19	H _N	8.736	8.717	8.718								
	N		121.09	121.21								
	C α			58.65								
T20	H _N	7.683	7.654	7.654								
	N		110.61	110.81								
	C α			65.29								
T21	H _N	8.332	7.665	7.664								
	N		109.67	109.87								
	C α			63.51								
K22	H _N	8.534	8.510	8.512			8.502	8.495		8.502		
	N		120.14	120.33			120.17			120.08		
	C α			55.91	56.35			55.81				
	H α				4.709		4.707					
	C β				35.64	35.23		35.24				
	H β 1	3.292										
	H β 2				1.865	1.865	1.881					
	C γ				24.84	24.85						
	H γ *				1.481	1.480				1.461		
	H γ *											
	C δ				28.59	28.67						
	H δ *				1.574	1.576						
	H δ *					1.725						
	C ϵ				42.97	42.34						
He1				2.912	2.913							
He2				2.984	2.983							
G23	H _N	7.932	7.911	7.911								
	N		108.88	109.08								
	C α			44.75								
A24	H _N	8.719	8.701	8.702								
	N		120.75	120.85								
	C α			55.81								

D25	H _N	8.795	8.766	8.767			8.752	8.754		8.758		
	N		117.35	117.48			117.31			117.38		
	C α			56.96				56.93				
	H α	4.377					4.375			4.347		
	C β											
	H β 1						2.815					
	H β 2	2.842					2.823			2.825		
	C=O											
K26	H _N	7.687	7.654	7.654			7.646	7.640		7.649		
	N		119.38	119.39			119.07			119.15		
	C α			57.30	57.60	57.69		57.32				
	H α	4.239			4.221	4.218	4.225			4.217		
	C β				31.29	31.00		31.10				
	H β 1	1.834			1.816	1.821	1.829			1.807		
	H β 2											
	C γ				25.66	25.40						
	H γ 1	1.501			1.491	1.493				1.484		
	H γ 2											
	C δ				28.16	28.53						
	H δ *				1.709	1.707						
	H δ *											
	C ϵ				42.54	42.28						
H ϵ 1	3.008			3.003	3.003							
H ϵ 2												
V27	H _N	7.490	7.463	7.464								
	N		121.41	121.51								
	C α			67.19								
K28	H _N	7.984	7.954	7.954			7.944	7.945		7.945		
	N		119.30	119.44			119.35			119.21		
	C α			61.06	61.07	60.95		61.03				
	H α	3.813			3.807	3.803	3.807			3.799		
	C β				31.89	31.69		31.79				
	H β 1	1.945			1.937	1.937	1.973			1.962		
	H β 2				1.997	1.998						
	C γ				26.89	26.55						
	H γ *				1.311	1.309						
H γ *				1.562	1.561				1.532			

K28	C δ				29.39	29.43						
	H δ^*				1.701	1.715						
	H δ^*											
	C ϵ				42.52	42.21						
	H ϵ^*				2.836	2.837						
	H ϵ^*											
D29	H _N	8.164	8.141	8.141			8.129	8.127		8.129		
	N		118.73	118.95			118.81			118.78		
	C α			56.41				56.38				
	H α	4.408					4.400					
	C β											
	H β 1	2.899					2.884			2.865		
	H β 2	2.960										
	C=O											
Y30	H _N	8.327	8.296	8.297								
	N		122.75	122.95								
	C α			61.90								
I31	H _N	8.177	8.152	8.154								
	N		119.05	119.08								
	C α			64.49								
S32	H _N	8.147	8.127	8.127								
	N		114.81	114.89								
	C α			61.98								
S33	H _N	7.888	7.869	7.869								
	N		115.51	115.71								
	C α			61.43								
L34	H _N	8.149	8.139	8.139								
	N		123.80	123.88								
	C α			57.61								
C35	H _N	8.030	8.004	8.004								
	N		111.00	111.05								
	C α			54.22								
N36	H _N	7.650	7.618	7.618			7.622			7.621		
	N		121.63	121.79			121.58			121.65		
	C α			55.53	55.25							
	H α	4.514			4.512		4.508					
	C β				38.73	38.58						
	H β 1	2.839			2.830	2.828	2.819		2.843	2.820	2.829	2.733

N36	H β 2	3.008			2.994	2.994	3.002		3.014	2.959	3.003	2.911
	C γ								178.63		178.64	180.66
K37	H $_N$	8.556	8.569				8.518	8.523				
	N		121.07				120.99					
	C α				55.59	55.81		55.90				
	H α	4.223			4.227	4.233	4.230					
	C β				30.67	30.89						
	H β^*					1.712	1.725					
	H β^*				1.927	1.925	1.944					
	C γ				24.41	24.76						
	H γ^*				1.334	1.334						
	H γ^*				1.394	1.394						
	C δ				28.77	29.01						
	H δ^*				1.630	1.630						
	H δ^*				1.682	1.680						
	C ϵ				42.54	42.14						
H ϵ^*				2.971	2.971							
H ϵ^*												
A38	H $_N$	7.234	7.218	7.218								
	N		122.08	122.10								
	C α			52.54								
S39	H $_N$	8.505	8.485	8.486								
	N		114.06	114.16								
	C α			56.84								
G40	H $_N$	8.805	8.795	8.796								
	N		109.79	110.10								
	C α			46.43								
F41	H $_N$	8.679	8.640	8.644								
	N		123.79	123.77								
	C α			61.35								
I42	H $_N$	8.119	8.110	8.112								
	N		118.15	118.08								
	C α			63.47								
A43	H $_N$	6.965	6.940	6.940								
	N		123.10	123.16								
	C α			55.30								

T44	H _N		7.691	7.692							
	N		116.41	116.59							
	C α			66.74							
L45	H _N	8.371	8.347	8.350							
	N		123.05	122.93							
	C α			58.42							
C46	H _N	9.085	9.067	9.068							
	N		117.94	118.05							
	C α			62.83							
T47	H _N	8.295	8.281	8.282							
	N		116.95	117.19							
	C α			67.31							
K48	H _N	7.750	7.728	7.728			7.712	7.724		7.711	
	N		119.49	119.55			119.35			119.34	
	C α			59.33	59.34	59.31		59.27			
	H α	4.214			4.207	4.206	4.212			4.188	
	C β				32.09	31.72		31.69			
	H β 1	2.060			2.054	2.045	2.072			2.073	
	H β 2	2.117			2.108	2.103					
	C γ				25.89	25.88					
	H γ *				1.534	1.527					
	H γ *				1.797	1.798					
	C δ				28.39	27.78					
	H δ *				1.708	1.707					
	H δ *										
	C ϵ				42.77	42.64					
He*				2.825	2.825						
He*				3.207	3.022						
V49	H _N	8.549	8.527	8.528							
	N		120.73	120.80							
	C α			67.37							
L50	H _N	8.935	8.913	8.914							
	N		119.85	120.03							
	C α			57.97							
D51	H _N	8.504	8.489	8.490			8.471	8.474		8.473	
	N		117.20	117.40			117.23			117.26	
	C α			55.39				55.75			
	H α	4.402					4.484				

D51	C β							37.86	37.98			
	H β^*								2.930	2.903	2.926	2.815
	H β^*						3.038		3.054	3.036	3.043	2.943
	C=O								178.537		178.54	180.58
F52	H _N	7.563	7.536	7.536								
	N		121.23	121.31								
	C α			61.08								
G53	H _N	8.107	8.086	8.086								
	N		107.43	107.58								
	C α			44.32								
I54	H _N	8.985	8.971	8.972								
	N		128.65	128.78								
	C α			65.29								
D55	H _N	8.625	8.591	8.591			8.573	8.586			8.580	
	N		118.88	119.01			118.75				118.81	
	C α			57.21	57.23			57.11				
	H α	4.381			4.369		4.376				4.365	
	C β				37.77			37.91	37.97			
	H β 1				2.727		2.726		2.766	2.732	2.742	2.658
	H β 2	2.776			2.636							
	C=O								179.56		179.71	181.59
K56	H _N	7.534	7.508	7.509			7.491	7.497			7.500	
	N		122.08	122.04			121.74				121.63	
	C α			59.39	59.94	59.26		59.39				
	H α	3.985			3.980	3.973	3.985				3.976	
	C β				32.31	32.26		32.36				
	H β 1											
	H β 2	1.538			1.535	1.527	1.520				1.526	
	C γ				25.99	25.42						
	H γ^*				1.221	1.208						
	H γ^*				1.397	1.392					1.389	
	C δ				29.01	28.64						
	H δ^*				1.287	1.235						
	H δ^*					1.299						
	C ϵ				41.52	41.71						
	H ϵ^*				2.602	2.601						
	H ϵ^*				2.718	2.717						

L57	H _N	8.399	8.382	8.383								
	N		122.19	122.39								
	C α			59.20								
I58	H _N	8.533	8.517	8.516								
	N		118.64	118.56								
	C α			66.26								
Q59	H _N	7.766	7.740	7.740			7.732	7.732		7.733		
	N		119.16	119.27			119.06			119.05		
	C α			59.22	58.85	58.77						
	H α	4.015			4.004	4.005	4.013			4.004		
	C β				28.60	28.42		28.51				
	H β *				2.185	2.180	2.171			2.175		
	H β *											
	C γ				34.22	34.03						
	H γ *				2.351	2.352			2.365		2.364	2.255
	H γ *				2.434	2.434			2.444		2.446	2.326
C δ								181.28		181.25	183.29	
L60	H _N	7.636	7.606	7.607								
	N		118.09	118.22								
	C α			58.45								
I61	H _N	7.913	7.891	7.892								
	N		118.78	118.88								
	C α			65.40								
E62	H _N	8.914	8.891	8.908			8.882	8.882		8.887		
	N		122.25	121.89			122.16			122.21		
	C α			57.95	58.99	58.88		58.83				
	H α	4.019			4.015	4.020	4.014			4.001		
	C β				28.16	27.57		27.66				
	H β 1	2.130			2.129	2.117	2.119			2.127		
	H β 2	2.289			2.289	2.289	2.286			2.283		
	C γ				34.07				34.09			
	H γ *	2.539			2.531				2.500	2.499	2.536	2.423
	H γ *	2.636			2.623				2.648	2.645	2.636	2.520
C=O								180.53		180.50	182.54	
D63	H _N	7.911	7.888	7.889			7.886	7.882		7.874		
	N		117.20	117.38			117.08			117.20		
	C α			53.22	53.66			53.26				

D63	H α	4.691			4.685		4.686					
	C β				38.39				38.28			
	H β 1	3.035			3.047				3.035	3.044	3.029	2.923
	H β 2	3.129			3.108		3.091		3.119		3.114	3.020
	C=O								178.96		178.93	180.99
K64	H _N	8.079	8.055	8.057			8.047	8.041		8.057		
	N		114.55	114.65			114.44			114.35		
	C α			57.03	56.98	57.11		57.11				
	H α	3.912			3.906	3.895	3.900			3.896		
	C β				28.34	28.18		28.21				
	H β * ₁				1.945	1.948	1.953			1.940		
	H β * ₂				2.114	2.109	2.095			2.093		
	C γ				24.78	24.79						
	H γ * ₁				1.312	1.319				1.320		
	H γ * ₂				1.376	1.375				1.380		
	C δ				29.02	28.98						
	H δ * ₁				1.640	1.639						
	H δ * ₂				1.713	1.710						
	C ϵ				42.29	42.29						
He* ₁				3.015	3.017							
He* ₂												
V65	H _N	7.555	7.531	7.532								
	N		121.80	121.92								
	C α			63.56								
D66	H _N	8.506	8.468	8.468			8.438	8.443		8.443		
	N		124.77	124.96			124.79			124.80		
	C α			53.80	54.11			53.87				
	H α	4.433			4.429		4.425			4.404		
	C β				40.27			40.07	39.98			
	H β 1	2.896			2.866				2.868	2.899	2.867	2.762
D66	H β 2	3.109			3.069		3.084		3.088	3.072	3.082	2.990
	C=O								179.48		179.58	181.49
A67	H _N	8.742	8.720	8.721								
	N		122.52	122.66								
	C α			56.25								
N68	H _N	8.476	8.452	8.453			8.437	8.434		8.442		
	N		115.80	115.95			115.79			115.71		

N68	C α			57.51	56.97	57.56		57.54				
	H α	3.931			3.928	3.920	3.940			3.929		
	C β				38.59	38.07		38.15	38.23			
	H β 1	2.644			2.644	2.640	2.672		2.648		2.650	2.542
	H β 2	2.728			2.730	2.728			2.743		2.742	2.631
C γ								178.60		178.55	180.61	
A69	H _N	7.661	7.656	7.657								
	N		122.75	122.96								
	C α			54.83								
I70	H _N	8.431	8.418	8.419								
	N		120.63	120.77								
	C α			66.88								
C71	H _N	8.284	8.254	8.254								
	N		115.76	115.92								
	C α			60.50								
A72	H _N	8.311	8.291	8.291								
	N		125.02	125.11								
	C α			54.93								
K73	H _N	7.833	7.814	7.815			7.812	7.810		7.816		
	N		122.14	122.27			122.04			122.10		
	C α			59.26	59.35	59.33		59.28				
	H α	3.861			3.837	3.844	3.850			3.831		
	C β				31.92	31.81		31.86				
	H β 1	1.814			1.683	1.680	1.668			1.652		
	H β 2				1.809	1.804	1.818			1.787		
	C γ				25.02	25.23						
	H γ *				1.435	1.434						
	H γ *				1.499	1.496						
	C δ				29.41	29.29						
	H δ 1				1.731	1.723						
	H δ 2	1.691			1.524	1.524						
	C ϵ				41.92	41.82						
H ϵ *				2.952	2.951							
H ϵ *				3.059	3.059							
I74	H _N	7.160	7.139	7.139								
	N		107.35	107.45								
	C α			60.65								

H75	H _N	7.586	7.564	7.565			7.556	7.560		7.565		
	N		112.60	112.74			122.55			112.48		
	C α			56.77	57.23	57.23		56.77				
	H α	4.362			4.364	4.364	4.356			4.353		
	C β				25.04	24.75		24.82				
	H β 1	3.473			3.461	3.459	3.457			3.453		
	H β 2	3.725			3.706	3.706	3.709			3.694		
A76	H _N	8.536	8.523	8.524								
	N		119.14	119.30								
	C α			52.66								
C77	H _N	6.836	6.815	6.814								
	N		119.19	119.45								
	C α			57.31								

The nuclei which cannot be detected in the corresponding NMR experiments are shaded in grey and the nuclei corresponding resonances that were too weak to be visible are shaded in green.

Co-Authorship Form

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Abstract.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

In Chapter 2: "Cell-free synthesis and combinatorial selective 15N-labeling of the cytotoxic protein amoebapore A from *Entamoeba histolytica*", Protein Expression and Purification, vol 68, pp. 22-27, 2009.

Nature of contribution by PhD candidate	Performed most of experiments including cell-free protein expression, protein purification, acquiring NMR experiments and assigning the NMR spectra.
Extent of contribution by PhD candidate (%)	70%

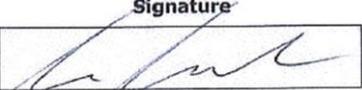
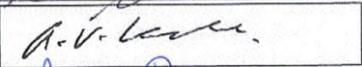
CO-AUTHORS

Name	Nature of Contribution
Colm Carraher	Provided reagents for the cell-free reactions.
Andrew V Kralicek	Provided training on the cell-free expression and produced combinatorial labelled sample, and edited the manuscript.
Andrew J Dingley	Provided main guidance on the NMR experiments and assignment procedure, and edited the manuscript.

Certification by Co-Authors

The undersigned hereby certify that:

- ❖ the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- ❖ in cases where the PhD candidate was the lead author of the work that the candidate wrote the text.

Name	Signature	Date
Colm Carraher		26/02/2013
Andrew V Kralicek		26/02/2013
Andrew J Dingley		26/02/2013
		Click here
		Click here
		Click here

References

1. Sargeant, P.G., Williams, J. E., & Grene, J. D., *The differentiation of invasive and non-invasive Entamoeba histolytica by isoenzyme electrophoresis*. Transactions of the Royal Society of Tropical Medicine and Hygiene, 1978. **72**(5): p. 519-521.
2. Organisation, W.H. *Amoebiasis: Parasitic Diseases* [cited 2012 27/09/2012].
3. Stanley, S.L., Jr., *Amoebiasis*. Lancet, 2003. **361**(9362): p. 1025-1034.
4. Mondal, D., Petri, W. A., Jr., Sack, R. B., Kirkpatrick, B. D., & Haque, R., *Entamoeba histolytica-associated diarrheal illness is negatively associated with the growth of preschool children: evidence from a prospective study*. Transactions of the Royal Society of Tropical Medicine and Hygiene, 2006. **100**(11): p. 1032-1038.
5. Dobell, C., *Research on the intestinal protozoa of monkeys and man*. Parasitology, 1928. **20**: p. 357-412.
6. Carrero, J.C., Cervantes-Rebolledo, C., Aguilar-Diaz, H., Diaz-Gallardo, M. Y., Laclette, J. P., & Morales-Montor, J., *The role of the secretory immune response in the infection by Entamoeba histolytica*. Parasite Immunology, 2007. **29**(7): p. 331-338.
7. Ravdin, J.I., *Immunobiology of human infection by Entamoeba histolytica*. Pathology and Immunopathology Research, 1989. **8**(3-4): p. 179-205.
8. Leippe, M., Ebel, S., Schoenberger, O. L., Horstmann, R. D., & Muller-Eberhard, H. J., *Pore-forming peptide of pathogenic Entamoeba histolytica*. Proceedings of the National Academy of Sciences of the United States of America, 1991. **88**(17): p. 7659-7663.
9. Ravdin, J.I., *Amebiasis*. Clinical Infectious Diseases, 1995. **20**(6): p. 1453-1464; quiz 1465-1466.
10. *Amoebiasis, Life Cycle The Centers for Disease Control and Prevention*. 2012 [cited 2012 18/04]; Available from: <http://www.dpd.cdc.gov/dpdx/HTML/Amebiasis.htm>.
11. Ansary, M.A. *Liver abscess*. [cited 2012 27/09/2012]; Available from: <http://www.sciencephoto.com>.
12. Muller, M., *Mode of action of metronidazole on anaerobic bacteria and protozoa* Surgery 1983. **93**: p. 165-171.
13. Wright, C.W., & Phillipson, J. D., *Natural products and the development of selective antiprotozoal drugs*. Phytotherapy Research Journal, 1990. **4**: p. 127-139.
14. Haque, R., Huston, C. D., Hughes, M., Houpt, E., & Petri, W. A., *Amoebiasis*. The New England Journal of Medicine, 2003. **348**: p. 1565-1573.
15. Bansal, D., Malla, N. & Mahajan, R.C., *Drug resistance in amoebiasis* Indian Journal of Medical Research, 2006. **123**: p. 115-118.
16. Ravdin, J.I., Croft, B. Y., & Guerrant, R. L., *Cytopathogenic mechanisms of Entamoeba histolytica*. The Journal of Experimental Medicine, 1980. **152**(2): p. 377-390.

17. Lynch, E.C., Rosenberg, I. M., & Gitler, C., *An ion-channel forming protein produced by Entamoeba histolytica*. The EMBO Journal, 1982. **1**(7): p. 801-804.
18. Young, J.D., Young, T. M., Lu, L. P., Unkeless, J. C., & Cohn, Z. A., *Characterization of a membrane pore-forming protein from Entamoeba histolytica*. The Journal of Experimental Medicine, 1982. **156**(6): p. 1677-1690.
19. Keller, F., Hanke, W., Trissl, D., & Bakker-Grunwald, T., *Pore-forming protein from Entamoeba histolytica forms voltage- and pH-controlled multi-state channels with properties similar to those of the barrel-stave aggregates*. Biochimica et Biophysica Acta, 1989. **982**(1): p. 89-93.
20. Rosenberg, I., Bach, D., Loew, L. M. & Gitler, C., *Isolation, characterization and partial purification of a transferable membrane channel (amoebapore) produced by Entamoeba histolytica*. Mol. Biochem. Parasitol., 1989. **33**: p. 237-248.
21. Young, J.D., & Cohn, Z. A., *Molecular mechanisms of cytotoxicity mediated by Entamoeba histolytica: characterization of a pore-forming protein (PFP)*. Journal of Cellular Biochemistry 1985. **29**(4): p. 299-308.
22. Leippe, M., Andra, J., Nickel, R., Tannich, E., & Muller-Eberhard, H. J., *Amoebapores, a family of membranolytic peptides from cytoplasmic granules of Entamoeba histolytica: isolation, primary structure, and pore formation in bacterial cytoplasmic membranes*. Molecular Microbiology, 1994. **14**(5): p. 895-904.
23. Leippe, M., *Membrane perforation by Entamoeba histolytica: structural implications derived from the sequence of the pore-forming peptide*. Archives of Medical Research, 1992. **23**(2): p. 35-37.
24. Hecht, O., Van Nuland, N. A., Schleinkofer, K., Dingley, A. J., Bruhn, H., Leippe, M., & Grotzinger, J., *Solution structure of the pore-forming protein of Entamoeba histolytica*. The Journal of Biological Chemistry, 2004. **279**(17): p. 17834-17841.
25. Leippe, M., Tannich, E., Nickel, R., van der Goot, G., Pattus, F., Horstmann, R. D., & Muller-Eberhard, H. J., *Primary and secondary structure of the pore-forming peptide of pathogenic Entamoeba histolytica*. The EMBO Journal, 1992. **11**(10): p. 3501-3506.
26. Herbst, R., Marciano-Cabral, F., & Leippe, M., *Antimicrobial and pore-forming peptides of free-living and potentially highly pathogenic Naegleria fowleri are released from the same precursor molecule*. The Journal of Biological Chemistry, 2004. **279**(25): p. 25955-25958.
27. Bracha, R., Nuchamowitz, Y., Leippe, M., & Mirelman, D., *Antisense inhibition of amoebapore expression in Entamoeba histolytica causes a decrease in amoebic virulence*. Molecular Microbiology 1999. **34**(3): p. 463-472.
28. Bracha, R., Nuchamowitz, Y., & Mirelman, D., *Transcriptional silencing of an amoebapore gene in Entamoeba histolytica: molecular analysis and effect on pathogenicity*. Eukaryotic Cell, 2003. **2**(2): p. 295-305.
29. Bracha, R., Nuchamowitz, Y., & Mirelman, D., *Amoebapore is an important virulent factor of Entamoeba histolytica*. Journal of Biosciences, 2002. **27**: p. 925-933.
30. Andra, J., Berninghausen, O., Wulfken, J., & Leippe, M., *Shortened amoebapore analogs with enhanced antibacterial and cytolytic activity*. FEBS Letters, 1996. **385**(1-2): p. 96-100.

31. Leippe, M., Andra, J., & Muller-Eberhard, H. J., *Cytolytic and antibacterial activity of synthetic peptides derived from amoebapore, the pore-forming peptide of Entamoeba histolytica*. Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(7): p. 2602-2606.
32. Andra, J., & Leippe, M., *Pore-forming peptide of Entamoeba histolytica. Significance of positively charged amino acid residues for its mode of action*. FEBS Letters, 1994. **354**(1): p. 97-102.
33. Armstrong, D., Zidovetzki, R., & Hu, J. *Helical Wheel Projections*. 2001 [cited 2012 9/12/2012]; Available from: <http://trimer.tamu.edu/cgi-bin/wheel/wheel.pl>.
34. Campbell, M.K., & Farrell, S.O., *Biochemistry* 4th edition ed. 2003.
35. Samson, M.S.P., *The biophysics of peptide models of ion channels*. Progress in Biophysics and Molecular Biology, 1991. **55**: p. 139-235.
36. Bechinger, B., Kim, Y., Chirlian, L. E., Gesell, J., Neumann, J.-M., Montal, M., Tomich, J., Zasloff, M., & Opefla, S. J., *Orientations of amphipathic helical peptides in membrane bilayers determined by solid-state NMR spectroscopy*. . Journal of Biomolecular NMR 1991. **1**(2): p. 167-173.
37. Wang, Z., & Wang, G., *APD: the Antimicrobial Peptide Database*. Nucleic Acids Research, 2004. **32**(Database issue): p. D590-592.
38. Matsuzaki, K., Sugishita, K., Harada, M., Fujii, N., & Miyajima, K., *Interactions of an antimicrobial peptide, magainin 2, with outer and inner membranes of Gram-negative bacteria*. Biochimica et Biophysica Acta, 1997. **1327**(1): p. 119-130.
39. Miles, E.W., *Modification of histidyl residues in proteins by diethylpyrocarbonate*. Methods in Enzymology 1977. **47**: p. 431-442.
40. Ravdin, J.I., Schlesinger, P. H., Murphy, C. F., Gluzman, I. Y., & Krogstad, D. J., *Acid intracellular vesicles and the cytolysis of mammalian target cells by Entamoeba histolytica trophozoites*. Journal of Protozoology, 1986. **33**(4): p. 478-486.
41. Nozaki, Y., & Tanford, C., *Examination of titration behavior*. Methods in Enzymology, 1967. **11**: p. 715-734.
42. Gurd, F.R.N., Keim, P., Glushko, V.G., Lawson, P.J., Marshall, R.C., Nigen, A.M., & Vigna, R.A. , *Carbon-13 nuclear magnetic resonance of some pentapeptides*. In *Chemistry and biology of peptides*, in *Chemistry and biology of peptides*, J. Meienhofer, Editor. 1972, Ann Arbor Science Publishers: Ann Arbor, MI. p. 45–49.
43. Thurlkill, R.L., Grimsley, G. R., Scholtz, J. M., & Pace, C. N., *pK values of the ionizable groups of proteins*. Protein Science 2006. **15**(5): p. 1214-1218.
44. Wuthrich, K., *NMR of proteins and Nucleic Acids*. 1986: John Wiley & Sons, New York.
45. Honig, B., & Nicholls, A., *Classical electrostatics in biology and chemistry*. Science, 1995. **268**(5214): p. 1144-1149.
46. Perez-Canadillas, J.M., Campos-Olivas, R., Lacadena, J., Martinez del Pozo, A., Gavilanes, J. G., Santoro, J., Rico, M., & Bruix, M., *Characterization of pKa values and titration shifts in the*

- cytotoxic ribonuclease alpha-sarcin by NMR. Relationship between electrostatic interactions, structure, and catalytic function.* Biochemistry, 1998. **37**(45): p. 15865-15876.
47. Alba, E., Weiler, S., & Tjandra, N., *Solution Structure of Human Saposin C: pH-Dependent Interaction with Phospholipid Vesicles.* Biochemistry 2003. **42**: p. 14729-14740.
 48. Pace, C.N., Grimsley, G. R., & Scholtz, J. M., *Protein ionizable groups: pK values and their contribution to protein stability and solubility.* The Journal of Biological Chemistry, 2009. **284**(20): p. 13285-13289.
 49. Gutschmann, T., Riekens, B., Bruhn, H., Wiese, A., Seydel, U., & Leippe, M., *Interaction of amoebapores and NK-lysin with symmetric phospholipid and asymmetric lipopolysaccharide/phospholipid bilayers.* Biochemistry, 2003. **42**(32): p. 9804-9812.
 50. Park, S.C., Kim, J. Y., Shin, S. O., Jeong, C. Y., Kim, M. H., Shin, S. Y., Cheong, G. W., Park, Y., & Hahm, K. S., *Investigation of toroidal pore and oligomerization by melittin using transmission electron microscopy.* Biochemical and Biophysical Research Communications 2006. **343**(1): p. 222-228.
 51. Mechaly, A.E., Bellomio, A., Gil-Carton, D., Morante, K., Valle, M., Gonzalez-Manas, J.M., & Guerin, D.M.A., *Structural Insights into the Oligomerization and Architecture of Eukaryotic Membrane Pore-Forming Toxins.* Structure, 2011. **19**: p. 181-191.
 52. Han, M., Mei, Y., Khant, H., & Ludtke, S.J, *Pore Formation by Antimicrobial Peptide, Magainin 2, in Phospholipid Vesicles Detected and Visualized by Cryo-Electron Microscopy.* Microscopy and Microanalysis, 2008. **14**: p. 1610-1611.
 53. Baumeister, W., & Steven, A. C., *Macromolecular electron microscopy in the era of structural genomics.* Trends in Biochemical Sciences, 2000. **25**(12): p. 624-631.
 54. Leippe, M., & Herbst, R., *Ancient weapons for attack and defense: the pore-forming polypeptides of pathogenic enteric and free-living amoeboid protozoa.* Journal of Eukaryotic Microbiology 2004. **51**(5): p. 516-521.
 55. Bruhn, H., Riekens, B., Berninghausen, O., & Leippe, M., *Amoebapores and NK-lysin, members of a class of structurally distinct antimicrobial and cytolytic peptides from protozoa and mammals: a comparative functional analysis.* Biochemical Journal, 2003. **375**(Pt 3): p. 737-744.
 56. Michalek, M., Gelhaus, C., Hecht, O., Podschun, R., Schroder, J. M., Leippe, M., & Grotzinger, J., *The human antimicrobial protein psoriasin acts by permeabilization of bacterial membranes.* Developmental & Comparative Immunology, 2009. **33**(6): p. 740-746.
 57. Leippe, M., Bahr, E., Tannich, E., & Horstmann, R. D., *Comparison of pore-forming peptides from pathogenic and nonpathogenic Entamoeba histolytica.* Molecular and Biochemical Parasitology, 1993. **59**(1): p. 101-109.
 58. Aley, S.B., Scott, W. A., & Cohn, Z. A., *Plasma membrane of Entamoeba histolytica.* The Journal of Experimental Medicine, 1980. **152**(2): p. 391-404.
 59. Andra, J., Berninghausen, O., & Leippe, M., *Membrane lipid composition protects Entamoeba histolytica from self-destruction by its pore-forming toxins.* FEBS Letters, 2004. **564**: p. 109-115.

60. Sandhu, K.S., & Gyles, C. L., *Pathogenic Shiga toxin-producing Escherichia coli in the intestine of calves*. Canadian Journal of Veterinary Research, 2002. **66**(2): p. 65-72.
61. Moody-Haupt, S., Patterson, J.H., Mirelman, D. & McConville, M.J. , *The Major Surface Antigens of Entamoeba histolytica Trophozoites are GPI-anchored Proteophosphoglycans*. Journal of Molecular Biology 2000. **297**: p. 409-420.
62. Petri, W.A., Jr., Haque, R., & Mann, B. J., *The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite Entamoeba histolytica*. Annual Review of Microbiology, 2002. **56**: p. 39-64.
63. Seigneur, M., Mounier, J., Prevost, M. C., & Guillen, N., *A lysine- and glutamic acid-rich protein, KERP1, from Entamoeba histolytica binds to human enterocytes*. Cellular Microbiology, 2005. **7**(4): p. 569-579.
64. Marinets, A., Zhang, T., Guillen, N., Gounon, P., Bohle, B., Vollmann, U., Scheiner, O., Wiedermann, G., Stanley, S. L., & Duchene, M., *Protection against invasive amebiasis by a single monoclonal antibody directed against a lipophosphoglycan antigen localized on the surface of Entamoeba histolytica*. The Journal of Experimental Medicine, 1997. **186**(9): p. 1557-1565.
65. Que, X., & Reed, S. L., *Cysteine proteinases and the pathogenesis of amebiasis*. Clinical Microbiology Reviews, 2000. **13**(2): p. 196-206.
66. Stanley, S.L., Jr., Zhang, T., Rubin, D., & Li, E. , *Role of the Entamoeba histolytica cysteine proteinase in amebic liver abscess formation in severe combined immunodeficient mice*. Infection and Immunity, 1995. **63**: p. 1587-1590.
67. Ankri, S., Padilla-Vaca, F., Stolarsky, T., Koole, L., Katz, U., & Mirelman, D., *Antisense inhibition of expression of the light subunit (35 kDa) of the Gal/GalNac lectin complex inhibits Entamoeba histolytica virulence*. Molecular Microbiology 1999. **33**(2): p. 327-337.
68. Ankri, S., Stolarsky, T., & Mirelman, D., *Antisense inhibition of expression of cysteine proteinases does not affect Entamoeba histolytica cytopathic or haemolytic activity but inhibits phagocytosis*. Molecular Microbiology 1998. **28**(4): p. 777-785.
69. Munford, R.S., Sheppard, P. O., & O'Hara, P. J., *Saposin-like proteins (SAPLIP) carry out diverse functions on a common backbone structure*. The Journal of Lipid Research, 1995. **36**(8): p. 1653-1663.
70. Mysliwy, J., Dingley, A. J., Stanisak, M., Jung, S., Lorenzen, I., Roeder, T., Leippe, M., & Grotzinger, J., *Caenopore-5: the three-dimensional structure of an antimicrobial protein from Caenorhabditis elegans*. Developmental & Comparative Immunology, 2010. **34**(3): p. 323-330.
71. Kishimoto, Y., Hiraiwa, M., & O'Brien, J. S., *Saposins: structure, function, distribution, and molecular genetics*. The Journal of Lipid Research, 1992. **33**(9): p. 1255-1267.
72. Liepinsh, E., Andersson, M., Ruyschaert, J. M., & Otting, G., *Saposin fold revealed by the NMR structure of NK-lysin*. Nature Structural & Molecular Biology, 1997. **4**(10): p. 793-795.
73. Anderson, D.H., Sawaya, M. R., Cascio, D., Ernst, W., Modlin, R., Krensky, A., & Eisenberg, D., *Granulysin crystal structure and a structure-derived lytic mechanism*. Journal of Molecular Biology, 2003. **325**(2): p. 355-365.

74. Bruhn, H., & Leippe, M., *Comparative modeling of amoebapores and granulysin based on the NK-lysin structure-structural and functional implications*. The Journal of Biological Chemistry, 1999. **380**(7-8): p. 1001-1007.
75. de Alba, E., Weiler, S., & Tjandra, N., *Solution structure of human saposin C: pH-dependent interaction with phospholipid vesicles*. Biochemistry, 2003. **42**(50): p. 14729-14740.
76. Qi, X., & Grabowski, G. A., *Differential membrane interactions of saposins A and C: implications for the functional specificity*. The Journal of Biological Chemistry, 2001. **276**(29): p. 27010–27017.
77. Miteva, M., Andersson, M., Karshikoff, A. & Otting, G. , *Molecular electroporation: a unifying concept for the description of membrane pore formation by antibacterial peptides, exemplified with NK-lysin*. FEBS Letters, 1999. **462**: p. 155-158.
78. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. & Higgins, D.G., *ClustalW and ClustalX version 2*. Bioinformatics, 2007. **23**(21): p. 2947-2948.
79. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J., *GROMACS: fast, flexible, and free*. Journal of Computational Chemistry, 2005. **26**(16): p. 1701-1718.
80. Leippe, M., Bruhn, H., Hecht, O., & Grotzinger, J., *Ancient weapons: the three-dimensional structure of amoebapore A*. Trends in Parasitology 2005. **21**(1): p. 5-7.
81. Brogden, K.A., *Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?* Nature Reviews Microbiology, 2005. **3**(3): p. 238-250.
82. Boheim, G., *Statistical analysis of alamethicin channels in black lipid membranes*. The Journal of Membrane Biology, 1974. **19**(3): p. 277-303.
83. Xu, Z., Chen, H., Yin, X., Xu, N., & Cen, P., *High-level expression of soluble human beta-defensin-2 fused with green fluorescent protein in Escherichia coli cell-free system*. Applied Biochemistry and Biotechnology, 2005. **127**(1): p. 53-62.
84. Wu, P.S., Ozawa, K., Jergic, S., Su, X. C., Dixon, N. E., & Otting, G., *Amino-acid type identification in 15N-HSQC spectra by combinatorial selective 15N-labelling*. Journal of Biomolecular NMR 2006. **34**(1): p. 13-21.
85. Martemyanov, K.A., Shirokov, V. A., Kurnasov, O. V., Gudkov, A. T., & Spirin, A. S., *Cell-free production of biologically active polypeptides: application to the synthesis of antibacterial peptide cecropin*. Protein Expression and Purification, 2001. **21**(3): p. 456-4561.
86. Klammt, C., Schwarz, D., Lohr, F., Schneider, B., Dotsch, V., & Bernhard, F., *Cell-free expression as an emerging technique for the large scale production of integral membrane protein*. FEBS Journal, 2006. **273**(18): p. 4141-4153.
87. Katzen, F., Chang, G., & Kudlicki, W., *The past, present and future of cell-free protein synthesis*. Trends in Biotechnology 2005. **23**(3): p. 150-156.
88. Ozawa, K., Dixon, N. E., & Otting, G., *Cell-free synthesis of 15N-labeled proteins for NMR studies*. IUBMB Life, 2005. **57**(9): p. 615-622.

89. Quash, G., Fournet, G., & Reichert, U., *Anaplerotic reactions in tumour proliferation and apoptosis*. *Biochemical Pharmacology*, 2003. **66**(3): p. 365-370.
90. Engleking, L., *Physiological Chemistry*. 2005: Teton NewMedia.
91. Pratt, J.M., *Coupled transcription-translation in prokaryotic cell-free systems*, in *Transcription and Translation: a Practical Approach*, S.J.H. B.D. Hames, Editor. 1984, IRL Press: Oxford, UK. p. 179-209.
92. Ozawa, K., Headlam, M. J., Schaeffer, P. M., Henderson, B. R., Dixon, N. E., & Otting, G., *Optimization of an Escherichia coli system for cell-free synthesis of selectively N-labelled proteins for rapid analysis by NMR spectroscopy*. *European Journal of Biochemistry* 2004. **271**(20): p. 4084-4093.
93. Liu, D.V., Zawada, J. F., & Swartz, J. R., *Streamlining Escherichia coli S30 extract preparation for economical cell-free protein synthesis*. *Biotechnology Progress*, 2005. **21**(2): p. 460-465.
94. Zawada, J., & Swartz, J., *Effects of growth rate on cell extract performance in cell-free protein synthesis*. *Biotechnology and Bioengineering*, 2006. **94**(4): p. 618-624.
95. Swartz, J., *Developing cell-free biology for industrial applications*. *Journal of Industrial Microbiology and Biotechnology*, 2006. **33**(7): p. 476-485.
96. Zawada, J.F., Yin, G., Steiner, A. R., Yang, J., Naresh, A., Roy, S. M., Gold, D. S., Heinsohn, H. G., & Murray, C. J., *Microscale to manufacturing scale-up of cell-free cytokine production--a new approach for shortening protein production development timelines*. *Biotechnology and Bioengineering*, 2011. **108**(7): p. 1570-1578.
97. Krishnarjuna, B., Jaipuria, G., Thakur, A., D'Silva, P., & Atreya, H. S., *Amino acid selective unlabeled for sequence specific resonance assignments in proteins*. *Journal of Biomolecular NMR* 2011. **49**(1): p. 39-51.
98. Tanio, M., Tanaka, R., Tanaka, T., & Kohno, T., *Amino acid-selective isotope labeling of proteins for nuclear magnetic resonance study: proteins secreted by Brevibacillus choshinensis*. *Analytical Biochemistry* 2009. **386**(2): p. 156-160.
99. Davanloo, P., Rosenberg, A. H., Dunn, J. J., & Studier, F. W., *Cloning and expression of the gene for bacteriophage T7 RNA polymerase*. *Proceedings of the National Academy of Sciences of the United States of America*, 1984. **81**: p. 2035-2039.
100. Buck, M.A., Olah, T. A., Weitzmann, C. J., & Cooperman, B. S., *Protein estimation by the product of integrated peak area and flow rate*. *Analytical Biochemistry*, 1989. **182**(2): p. 295-299.
101. Vinarov, D.A., Lytle, B. L., Peterson, F. C., Tyler, E. M., Volkman, B. F., & Markley, J. L., *Cell-free protein production and labeling protocol for NMR-based structural proteomics*. *Nature Methods*, 2004. **1**(2): p. 149-153.
102. Tyler, R.C., Aceti, D. J., Bingman, C. A., Cornilescu, C. C., Fox, B. G., Frederick, R. O., Jeon, W. B., Lee, M. S., Newman, C. S., Peterson, F. C., Phillips, G. N., Jr., Shahan, M. N., Singh, S., Song, J., Sreenath, H. K., Tyler, E. M., Ulrich, E. L., Vinarov, D. A., Vojtik, F. C., Volkman, B. F., Wrobel, R. L., Zhao, Q., & Markley, J. L., *Comparison of cell-based and cell-free protocols for producing target proteins from the Arabidopsis thaliana genome for structural studies*. *Proteins*, 2005. **59**(3): p. 633-643.

103. Staunton, D., Schlinkert, R., Zanetti, G., Colebrook, S. A., & Campbell, I. D., *Cell-free expression and selective isotope labelling in protein NMR*. Magnetic Resonance in Chemistry, 2006. **44 Spec No**: p. S2-9.
104. Ozawa, K., Wu, P. S., Dixon, N. E., & Otting, G., *N-Labelled proteins by cell-free protein synthesis. Strategies for high-throughput NMR studies of proteins and protein-ligand complexes*. FEBS Journal, 2006. **273**(18): p. 4154-4159.
105. Craven, C.J., Al-Owais, M., & Parker, M. J., *A systematic analysis of backbone amide assignments achieved via combinatorial selective labelling of amino acids*. Journal of Biomolecular NMR, 2007. **38**(2): p. 151-159.
106. Ozawa, K., Loscha, K. V., Kuppan, K. V., Loh, C. T., Dixon, N. E., & Otting, G., *High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites*. Biochemical and Biophysical Research Communications 2012. **418**(4): p. 652-656.
107. Noren, C.J., Anthony-Cahill, S. J., Griffith, M. C., & Schultz, P. G., *A general method for site-specific incorporation of unnatural amino acids into proteins*. Science, 1989. **244**(4901): p. 182-188.
108. Jackson, A.M., Boutell, J., Cooley, N., & He, M., *Cell-free protein synthesis for proteomics*. Brief Funct Genomic Proteomic, 2004. **2**(4): p. 308-319.
109. Dingley, A.J., I. Lorenzen, and J. Grotzinger, *NMR analysis of viral protein structures*. Methods Mol Biol, 2008. **451**: p. 441-62.
110. Cavanagh, J., *Chapter 6-Experimental ¹H NMR methods in Protein NMR Spectroscopy: Principles And Practice*. 2007, Academic Press.
111. Aue, W.P., Bartholdi, E., & Ernst R.R., *Two-dimensional spectroscopy application to nuclear magnetic resonance*. Journal of Chemical Physics, 1976. **64**(5): p. 2229-2246.
112. Jeener, J., M., B. H., Bachmann, P., & Ernst, R. R. , *Investigation of exchange processes by two-dimensional NMR spectroscopy*. Journal of Chemical Physics, 1979. **71**: p. 4546-4553.
113. Keeler, J., *Understanding NMR Spectroscopy*. 2010: Wiley
114. Rule, G.S., & Hitchens, T. K., in *Fundamentals of Protein NMR Spectroscopy*. 2006. p. 170-175.
115. Guignard, L., Ozawa, K., Pursglove, S. E., Otting, G., & Dixon, N. E., *NMR analysis of in vitro-synthesized proteins without purification: a high-throughput approach*. FEBS Letters, 2002. **524**(1-3): p. 159-162.
116. Hill, J.M., *NMR Screening for Rapid Protein Characterization in Structural Proteomics*, in *Structural Proteomics, High-Throughput Methods*, M.G. B. Kobe., T. Huber, Editor. 2008, Humana Press. p. 437-446.
117. Oschkinat, H., Griesinger, C., Kraulis, P. J., Sorensen, O. W., Ernst, R. R., Gronenborn, A. M., & Clore, G. M., *Three-dimensional NMR spectroscopy of a protein in solution*. Nature, 1988. **332**(6162): p. 374-376.
118. Ikura, M., Kay, L. E., & Bax, A., *A novel approach for sequential assignment of ¹H, ¹³C, and ¹⁵N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin*. Biochemistry, 1990. **29**(19): p. 4659-4667.

119. Wishart, D.S., B.D. Sykes, and F.M. Richards, *The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy*. *Biochemistry*, 1992. **31**(6): p. 1647-51.
120. Bax, A., & Ikura, M., *An efficient 3D NMR technique for correlating the proton and ¹⁵N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly ¹⁵N/¹³C enriched proteins*. *Journal of Biomolecular NMR*, 1991. **1**(1): p. 99-104.
121. *Vector map and linker sequence: pIVEX2.4a*. 2008 [cited 2008 10/09/2008]; Available from: <https://www.roche-applied-science.com/servlet/RCConfigureUser?URL=StoreFramesetView&storeId=10202&catalogId=10202&langId=-1&countryId=us>.
122. Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T., & Yokoyama, S., *Cell-free production and stable-isotope labeling of milligram quantities of proteins*. *FEBS Letters*, 1999. **442**(1): p. 15-19.
123. Ozawa, K., Jergic, S., Crowther, J. A., Thompson, P. R., Wijffels, G., Otting, G., Dixon, N. A., *Cell-free protein synthesis in an autoinduction system for NMR studies of protein-protein interactions*. *Journal of Biomolecular NMR*, 2005. **32**(3): p. 235-241.
124. Klink, T.A., Vicentini, A. M., Hofsteenge, J., & Raines, R. T., *High-level soluble production and characterization of porcine ribonuclease inhibitor*. *Protein Expression and Purification* 2001. **22**(2): p. 174-179.
125. Moore, J., Abdul-Manan, N., Fejzo, J., Jacobs, M., Lepre, C., Peng, J., & Xie, X., *Leveraging structural approaches: applications of NMR-based screening and X-ray crystallography for inhibitor design*. *Journal of Synchrotron Radiation*, 2004. **11**(Pt 1): p. 97-100.
126. Horst, R., Bertelsen, E. B., Fiaux, J., Wider, G., Horwich, A. L., & Wuthrich, K., *Direct NMR observation of a substrate protein bound to the chaperonin GroEL*. *Proceedings of the National Academy of Sciences of the United States of America*, 2005. **102**(36): p. 12748-12753.
127. Fogh, R.H., Vranken, W. F., Boucher, W., Stevens, T. J., & Laue, E. D., *A nomenclature and data model to describe NMR experiments*. *Journal of Biomolecular NMR* 2006. **36**(3): p. 147-155.
128. Sklenar, G., & Bax, A., *Spin-echo water suppression for the generation of pure-phase two-dimensional NMR Spectra*. *Journal of Magnetic Resonance*, 1987. **74**: p. 469-479.
129. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. *Journal of Biomolecular NMR*, 1995. **6**(3): p. 277-293.
130. Vranken, W.F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., & Laue, E. D., *The CCPN data model for NMR spectroscopy: development of a software pipeline*. *Proteins*, 2005. **59**(4): p. 687-696.
131. Xun, Y., Tremouilhac, P., Carraher, C., Gelhaus, C., Ozawa, K., Otting, G., Dixon, N. E., Leippe, M., Grotzinger, J., Dingley, A. J., & Kralicek, A. V., *Cell-free synthesis and combinatorial selective ¹⁵N-labeling of the cytotoxic protein amoebapore A from Entamoeba histolytica*. *Protein Expression and Purification* 2009. **68**(1): p. 22-27.

132. Abdullah, N., & Chase, H. A., *Removal of poly-histidine fusion tags from recombinant proteins purified by expanded bed adsorption*. Biotechnology and Bioengineering, 2005. **92**(4): p. 501-513.
133. Moss, J., Haun, R. S., Tsai, S. C., Welsh, C. F., Lee, F. J., Price, S. R., & Vaughan, M., *Activation of cholera toxin by ADP-ribosylation factors: 20-kDa guanine nucleotide-binding proteins*. Methods in Enzymology, 1994. **237**: p. 44-63.
134. Hanck, T., Stricker, R., Krishna, U. M., Falck, J. R., Chang, Y. T., Chung, S. K., & Reiser, G., *Recombinant p42IP4, a brain-specific 42-kDa high-affinity Ins(1,3,4,5)P4 receptor protein, specifically interacts with lipid membranes containing Ptd-Ins(3,4,5)P3*. European Journal of Biochemistry 1999. **261**(2): p. 577-584.
135. Guan, M., Su, B., Ye, C., & Lu, Y., *Production of extracellular domain of human tissue factor using maltose-binding protein fusion system*. Protein Expression and Purification, 2002. **26**(2): p. 229-234.
136. Ogino, T., Kaji, T., Kawabata, M., Satoh, K., Tomoo, K., Ishida, T., Yamazaki, H., Ishidoh, K., & Kominami, E., *Function of the propeptide region in recombinant expression of active procathepsin L in Escherichia coli*. The Journal of Biochemistry, 1999. **126**(1): p. 78-83.
137. Rihs, H.P., Chen, Z., Rueff, F., Petersen, A., Rozynek, P., Heimann, H., & Baur, X., *IgE binding of the recombinant allergen soybean profilin (rGly m 3) is mediated by conformational epitopes*. Journal of Allergy and Clinical Immunology 1999. **104**(6): p. 1293-1301.
138. Kushwaha, A., Rao, P. P., Duttu, V. S., Malhotra, P., & Chauhan, V. S., *Expression and characterisation of Plasmodium falciparum acidic basic repeat antigen expressed in Escherichia coli*. Molecular and Biochemical Parasitology, 2000. **106**(2): p. 213-224.
139. Huys, I., Dyason, K., Waelkens, E., Verdonck, F., van Zyl, J., du Plessis, J., Muller, G. J., van der Walt, J., Clynen, E., Schoofs, L., & Tytgat, J., *Purification, characterization and biosynthesis of parabutoxin 3, a component of Parabuthus transvaalicus venom*. European Journal of Biochemistry, 2002. **269**(7): p. 1854-1865.
140. Reed, D.W., & Hartzell, P. L., *The Archaeoglobus fulgidus D-lactate dehydrogenase is a Zn(2+) flavoprotein*. Journal of Bacteriology, 1999. **181**(24): p. 7580-7587.
141. Eaton, D., H. Rodriguez, and G.A. Vehar, *Proteolytic processing of human factor VIII. Correlation of specific cleavages by thrombin, factor Xa, and activated protein C with activation and inactivation of factor VIII coagulant activity*. Biochemistry, 1986. **25**(2): p. 505-12.
142. Quinlan, R.A., Moir, R. D., & Stewart, M., *Expression in Escherichia coli of fragments of glial fibrillary acidic protein: characterization, assembly properties and paracrystal formation*. Journal of Cell Science, 1989. **93** (Pt 1): p. 71-83.
143. Zaitseva, J., Zhang, H., Binnie, R. A., & Hermodson, M., *The proteins encoded by the rbs operon of Escherichia coli: II. Use of chimeric protein constructs to isolate and characterize RbsC*. Protein Science, 1996. **5**(6): p. 1100-1107.
144. Rath, A., Gilbowicka, M., Nadeau, V. G., & Deber, C. M., *Detergent binding explains anomalous SDS-PAGE migration of membrane proteins*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(6): p. 1760-1765.

145. Houde, D., Kauppinen, P., Mhatre, R., & Lyubarskaya, Y., *Determination of protein oxidation by mass spectrometry and method transfer to quality control*. Journal of Chromatography A 2006. **1123**(2): p. 189-198.
146. Jia, X., Ozawa, K., Loscha, K., & Otting, G., *Glutarate and N-acetyl-L-glutamate buffers for cell-free synthesis of selectively ¹⁵N-labelled proteins*. Journal of Biomolecular NMR 2009. **44**(2): p. 59-67.
147. Huang, H., Yu, Y., Yi, X., & Zhang, Y., *Nitrogen metabolism of asparagine and glutamate in Vero cells studied by (1)H/ (15)N NMR spectroscopy*. Applied Microbiology and Biotechnology, 2007. **77**(2): p. 427-436.
148. Morita, E.H., Shimizu, M., Ogasawara, T., Endo, Y., Tanaka, R., & Kohno, T., *A novel way of amino acid-specific assignment in (1)H-(15)N HSQC spectra with a wheat germ cell-free protein synthesis system*. Journal of Biomolecular NMR, 2004. **30**(1): p. 37-45.
149. Ireland, R.J., & Lea, P. J. , *The enzymes of glutamine, glutamate, asparagine, and aspartate metabolism*, in *Plant amino acids: biochemistry and biotechnology*, B.K. Singh, Editor., Marcel Dekker, Inc: New York. p. 49-99.
150. Calhoun, K.A., & Swartz, J. R. , *An Economical Method for Cell-Free Protein Synthesis using Glucose and Nucleoside Monophosphates*. Biotechnology Progress, 2005. **21**: p. 1146-1153.
151. Shin, J., & Noireaux, V., *Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70*. Journal of Biological Engineering, 2010. **4**: p. 8.
152. Wu, J.J., & Swartz, J. R., *High yield cell-free production of integral membrane proteins without refolding or detergents*. Biochimica et Biophysica Acta, 2008. **1778**(5): p. 1237-1250.
153. Belmouden, A., Le, K. H., Lederer, F., & Garchon, H. J., *Molecular cloning and nucleotide sequence of cDNA encoding rat kidney long-chain L-2-hydroxy acid oxidase. Expression of the catalytically active recombinant protein as a chimera*. European Journal of Biochemistry, 1993. **214**(1): p. 17-25.
154. Ko, Y.H., Thomas, P. J., Delannoy, M. R., & Pedersen, P. L., *The cystic fibrosis transmembrane conductance regulator. Overexpression, purification, and characterization of wild type and delta F508 mutant forms of the first nucleotide binding fold in fusion with the maltose-binding protein*. The Journal of Biological Chemistry, 1993. **268**(32): p. 24330-24338.
155. Kapust, R.B., Tozser, J., Copeland, T. D., & Waugh, D. S., *The P1' specificity of tobacco etch virus protease*. Biochemical and Biophysical Research Communications, 2002. **294**(5): p. 949-955.
156. Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., & Ueda, T., *Cell-free translation reconstituted with purified components*. Nature Biotechnology, 2001. **19**(8): p. 751-755.
157. Chen, H.A., Pfuhl, M., & Driscoll, P. C., *The pH dependence of CD2 domain 1 self-association and ¹⁵N chemical exchange broadening is correlated with the anomalous pKa of Glu41*. Biochemistry, 2002. **41**(50): p. 14680-14688.
158. Shuker, S.B., Hajduk, P. J., Meadows, R. P., & Fesik, S. W., *Discovering high-affinity ligands for proteins: SAR by NMR*. Science, 1996. **274**(5292): p. 1531-1534.

159. Chen, Y., Reizer, J., Saier, M. H., Jr., Fairbrother, W. J., & Wright, P. E., *Mapping of the binding interfaces of the proteins of the bacterial phosphotransferase system, HPr and IIAGlc*. *Biochemistry*, 1993. **32**(1): p. 32-37.
160. Gronenborn, A.M., & Clore, G. M., *Identification of the contact surface of a streptococcal protein G domain complexed with a human Fc fragment*. *Journal of Molecular Biology*, 1993. **233**(3): p. 331-335.
161. Foster, M.P., Wuttke, D. S., Clemens, K. R., Jahnke, W., Radhakrishnan, I., Tennant, L., Reymond, M., Chung, J., & Wright, P. E., *Chemical shift as a probe of molecular interfaces: NMR studies of DNA binding by the three amino-terminal zinc finger domains from transcription factor IIIA*. *Journal of Biomolecular NMR* 1998. **12**(1): p. 51-71.
162. Higashi, K., Terui, Y., Inomata, E., Katagiri, D., Nomura, Y., Someya, T., Nishimura, K., Kashiwagi, K., Kawai, G., & Igarashi, K., *Selective structural change of bulged-out region of double-stranded RNA containing bulged nucleotides by spermidine*. *Biochemical and Biophysical Research Communications*, 2008. **370**(4): p. 572-577.
163. Qin, J., Clore, G.M., & Gronenborn, A.M., *Ionization equilibria for side-chain carboxyl groups in oxidized and reduced human thioredoxin and in the complex with its target peptide from the transcription factor NFB*. *Biochemistry*, 1996. **35**: p. 7-13.
164. Sundd, M., Iverson, N., Ibarra-Molero, B., Sánchez-Ruiz, J.M., & Robertson, A.D., *Electrostatic interactions in ubiquitin: stabilization of carboxylates by lysine amino groups*. *Biochemistry*, 2002. **41**: p. 7586-7596.
165. Song, J., Leskowsky, M., Qasim, M., & Markley, J.L., *NMR determination of pKa values for Asp, Glu, His, and Lys mutants at each variable contiguous enzyme-inhibitor contact position of the turkey ovomucoid third domain*. *Biochemistry* 2003. **42** p. 2847-2856.
166. Laurents, D.V., Huyghues-Despointes, B.M.P., Bruix, M., Thurlkill, R.L., Schell, D., Newson, S., Grimsley, R.G., Shaw, K.L., Treviño, S., Rico, M., Briggs, J.M., Antosiewicz, J.M., Scholtz, M., & Pace, C.M., *Charge-charge interactions are key determinants of the pKa values of ionizable groups in ribonuclease Sa [pI=3.5] and a basic variant [pI=10.2]*. *Journal of Molecular Biology* 2003 **325**: p. 1077-1092.
167. Tudor, J.E., Pennington, M. W., & Norton, R. S., *Ionisation behaviour and solution properties of the potassium-channel blocker ShK toxin*. *European Journal of Biochemistry*, 1998. **251**(1-2): p. 133-141.
168. Chen, H.A., Pfuhl, M., McAlister, M. S., & Driscoll, P. C., *Determination of pK(a) values of carboxyl groups in the N-terminal domain of rat CD2: anomalous pK(a) of a glutamate on the ligand-binding surface*. *Biochemistry*, 2000. **39**(23): p. 6814-6824.
169. Nyarko, A., Cochrun, L., Norwood, S., Pursifull, N., Voth, A., & Barbar, E., *Ionization of His 55 at the dimer interface of dynein light-chain LC8 is coupled to dimer dissociation*. *Biochemistry*, 2005. **44**(43): p. 14248-14255.
170. Sebillé, B., *The measurement of interactions involving proteins by size exclusion chromatography*. , in *HPLC of biological macromolecules* K.M. Gooding, Regnier, F.E., Editor. 1990, Marcel Dekker: New York. p. 585-662.

171. Walsh, D.M., Lomakin, A., Benedek, G. B., Condrón, M. M., & Teplow, D. B., *Amyloid beta-protein fibrillogenesis. Detection of a protofibrillar intermediate*. The Journal of Biological Chemistry, 1997. **272**(35): p. 22364-22372.
172. Nichol, L.W., Siezen, R. J., & Winzor, D. J., *The study of multiple polymerization equilibria by glass bead exclusion chromatography with allowance for thermodynamic non-ideality effects*. Biophysical Chemistry 1978. **9**(1): p. 47-55.
173. Winzor, D.J., *Analytical exclusion chromatography*. Journal of Biochemical and Biophysical Methods, 2003. **56**(1-3): p. 15-52.
174. Darling, P.J., Holt, J. M., & Ackers, G. K., *Coupled energetics of lambda cro repressor self-assembly and site-specific DNA operator binding I: analysis of cro dimerization from nanomolar to micromolar concentrations*. Biochemistry, 2000. **39**(37): p. 11500-11507.
175. Feigin, L.A., & Svergun, D. I., *Structure Analysis By Small-angle X-ray and Neutron Scattering*. 1987, New York: Plenum Press.
176. Yamagata, A., & Tainer, J. A., *Hexameric structures of the archaeal secretion ATPase GspE and implications for a universal secretion mechanism*. The EMBO Journal, 2007. **26**(3): p. 878-890.
177. Krukenberg, K.A., Förster, F., Rice, L., Sali, A., & Agard, D. A., *A novel conformation of E.coliHsp90 in solution: insights into the conformational dynamics of Hsp90*. Structure, 2008. **16**: p. 755-756.
178. Glatter, O., & Kratky, O. I., *Small Angle X-ray Scattering*. 1982, 111 Fifth Avenue, New York: Academic Press.
179. Haydyn, D.T.M., Dmitri, I. S., *Structural characterization of proteins and complexes using small-angle X-ray solution scattering*. Journal of Structural Biology, 2010. **172**: p. 128-141.
180. Dainese, E., Sabatucci, A., & Cozzani, I., *Small Angle X-Ray Scattering: A Powerful Tool to Analyze Protein Conformation in Solution*. Current Organic Chemistry, 2005. **9**: p. 1781-1800.
181. Neylon, C., *Small angle neutron and X-ray scattering in structural biology: recent examples from the literature*. European Biophysics Journal 2008. **37**(5): p. 531-541.
182. Jacques, D.A., & Trewella, J., *Small-angle scattering for structural biology--expanding the frontier while avoiding the pitfalls*. Protein Science, 2010. **19**(4): p. 642-657.
183. Guinier, A., *La diffraction des rayons X aux tres petits angles: application a l'etude de phenomenes ultramicroscopiques*. Annals of Physics, 1939. **12**: p. 161-237.
184. Svergun, D.I., *Determination of the regularization parameter in indirect-transform methods using perceptual criteria*. Journal of Applied Crystallography, 1992. **25**: p. 495-503.
185. Glatter, O., *A New Method for the Evaluation of Small-Angle Scattering Data*. Journal of Applied Crystallography, 1977. **10**: p. 415-421.
186. Zheng, W., & Doniach, S., *Fold recognition aided by constraints from small angle X-ray scattering data*. Protein Engineering, Design and Selection, 2005. **18**(5): p. 209-219.
187. Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing*. Biophysical Journal 1999. **76**: p. 2879-2886.

188. Franke, D., & Svergun, D. I., *DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering*. Journal of Applied Crystallography, 2009. **42**: p. 342-346.
189. Svergun, D.I., Petoukhov, M. V., & Koch, M. H., *Determination of domain structure of proteins from X-ray solution scattering*. Biophysical Journal 2001. **80**(6): p. 2946-2953.
190. Volkov, V.V., & Svergun, D. I., *Uniqueness of ab initio shape determination in small-angle scattering*. Journal of Applied Crystallography, 2003. **36**: p. 860-864.
191. Kozin, M.B., & Svergun, D. I., *Automated matching of high- and low-resolution structural models*. Journal of Applied Crystallography, 2001. **34**: p. 33-41.
192. Svergun, D.I., Barberato, C., & Koch, M. H. J., *CRY SOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates*. Journal of Applied Crystallography, 1995. **28**: p. 768-773.
193. Grishaev, A., Guo, L., Irving, T., & Bax, A., *Improved fitting of solution X-ray scattering data to macromolecular structures and structural ensembles by explicit water modeling*. Journal of the American Chemical Society, 2010. **132**(44): p. 15484-15486.
194. Petoukhov, M.V., & Svergun, D. I., *Global rigid body modeling of macromolecular complexes against smallangle scattering data*. Biophysical Journal 2005. **89**: p. 1237-1250.
195. Bax, A., Clore, G. M., & Gronenborn, A. M., *1H-1H correlation via isotropic mixing of 13C magnetization, a new three-dimensional approach for assigning 1H and 13C spectra of 13C-enriched proteins*. Journal of Magnetic Resonance, 1990. **88**(2): p. 425-431.
196. Wittekind, M., & Mueller, L. , *HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the α -carbon and β -carbon resonances in proteins*. . Journal of Magnetic Resonance, Series B, 1993. **101**: p. 201-205.
197. Vuister, G.W., & Bax, A *Measurement of Two- and Three-Bond Proton to Methyl-Carbon J Couplings in Proteins Uniformly Enriched with ^{13}C* . J. Magn. Reson., 1993. **102**(2): p. 228-231.
198. Hallenga, K., & Lippens, G. M, *A constant-time ^{13}C - ^1H HSQC with uniform excitation over the complete ^{13}C chemical shift range*. Journal of Magnetic Resonance, 1994. **5**(1): p. 59-66.
199. Grzesiek, S., & Bax, A., *Amino acid type determination in the sequential assignment procedure of uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched proteins*. Journal of Biomolecular NMR 1993. **3**(2): p. 185-204.
200. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M., *Overcoming the overlap problem in the assignment of 1H NMR spectra of larger proteins by use of three-dimensional heteronuclear 1H-15N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta*. Biochemistry, 1989. **28**(15): p. 6150-6156.
201. Powers, R., Gronenborn, A. M., Clore, G. M., & Bax, A., *Three-Dimensional Triple-Resonance NMR of $^{13}\text{C}/^{15}\text{N}$ -Enriched Proteins Using Constant-Time Evolution* Journal of Magnetic Resonance, 1991. **94**: p. 209-213.
202. Grzesiek, S., & Bax, A. , *The origin and removal of artifacts in 3D HCACO spectra of proteins uniformly enriched with ^{13}C* . Journal of Magnetic Resonance, Series B, 1993. **102**: p. 103-106.

203. Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., & Sykes, B. D., *1H, 13C and 15N chemical shift referencing in biomolecular NMR*. Journal of Biomolecular NMR, 1995. **6**(2): p. 135-140.
204. Bundi, A., & Wuthrich, K., *Use of Amide 1H-NMR Titration Shifts for Studies of Polypeptide Conformation*. Biopolymers, 1979. **18**: p. 299-311.
205. Onufriev, A., Case, D. A., & Ullmann, G. M., *A novel view of pH titration in biomolecules*. Biochemistry, 2001. **40**(12): p. 3413-3419.
206. Chang, D.K., Cheng, S. F., Lin, C. H., Kantchev, E. B., & Wu, C. W., *Self-association of glutamic acid-rich fusion peptide analogs of influenza hemagglutinin in the membrane-mimic environments: effects of positional difference of glutamic acids on side chain ionization constant and intra- and inter-peptide interactions deduced from NMR and gel electrophoresis measurements*. Biochimica et Biophysica Acta 2005. **1712**(1): p. 37-51.
207. Quijada, J., Lopez, G., Versace, R., Ramirez, L., & Tasayco, M. L., *On the NMR analysis of pKa values in the unfolded state of proteins by extrapolation to zero denaturant*. Biophysical Chemistry 2007. **129**(2-3): p. 242-250.
208. Xue, R., Wang, S., Qi, H., Song, Y., Wang, C., & Li, F., *Structure analysis of the fourth transmembrane domain of Nramp1 in model membranes*. Biochimica et Biophysica Acta, 2008. **1778**(6): p. 1444-1452.
209. Markley, J.L., *Observation of histidine residues in proteins by nuclear magnetic resonance spectroscopy*. Accounts of Chemical Research, 1975. **8**(2): p. 70-80.
210. Akerud, T., Thulin, E., Van Etten, R. L., & Akke, M., *Intramolecular dynamics of low molecular weight protein tyrosine phosphatase in monomer-dimer equilibrium studied by NMR: a model for changes in dynamics upon target binding*. Journal of Molecular Biology, 2002. **322**(1): p. 137-152.
211. Press, W.H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T., *Numerical Recipes. The Art of Scientific Computing*. 1986, Cambridge Cambridge University Press.
212. Cookson, D.J., *Saxs15id - a software package for SAXS/WAXS data processing and analysis*. 2006.
213. Volkov, V.V., & Svergun, D.I., *Uniqueness of ab initio shape determination in small-angle scattering*. Journal of Applied Crystallography, 2003. **36**: p. 860-864.
214. Chen, H.A., Pfuhl, M., Driscoll, P.C., *The pH Dependence of CD2 Domain 1 Self-Association and 15N Chemical Exchange Broadening Is Correlated with the Anomalous pKa of Glu41*. Biochemistry Pharmacology, 2002. **41**: p. 14680-14688.
215. Dominguez, C., Boelens, R., & Bonvin, A. M., *HADDOCK: a protein-protein docking approach based on biochemical or biophysical information*. Journal of the American Chemical Society, 2003. **125**(7): p. 1731-1737.
216. McIntosh, L.P., & Dahlquist, F. W., *Biosynthetic incorporation of 15N and 13C for assignment and interpretation of nuclear magnetic resonance spectra of proteins*. Quarterly Reviews of Biophysics 1990. **23**(1): p. 1-38.

217. Tomlinson, J.H., Ullah, S., Hansen, P. E., & Williamson, M. P., *Characterization of salt bridges to lysines in the protein G B1 domain*. Journal of the American Chemical Society, 2009. **131**(13): p. 4674-4684.
218. Sundd, M., Iverson, N., Ibarra-Molero, B., Sanchez-Ruiz, J. M., & Robertson, A. D., *Electrostatic interactions in ubiquitin: stabilization of carboxylates by lysine amino groups*. Biochemistry, 2002. **41**(24): p. 7586-7596.
219. de Dios, A.C., Pearson, J. G., & Oldfield, E., *Secondary and tertiary structural effects on protein NMR chemical shifts: an ab initio approach*. Science, 1993. **260**(5113): p. 1491-1496.
220. Pfuhl, M., Chen, H. A., Kristensen, S. M., & Driscoll, P. C., *NMR exchange broadening arising from specific low affinity protein self-association: analysis of nitrogen-15 nuclear relaxation for rat CD2 domain 1*. Journal of Biomolecular NMR, 1999. **14**(4): p. 307-320.
221. Kline, S.R., *Reduction and Analysis of SANS and USANS Data using Igor Pro*. Journal of Applied Crystallography 2006. **39**(6): p. 895-900.
222. Mills, R.D., Trehwella, J., Qiu, T. W., Welte, T., Ryan, T. M., Hanley, T., Knott, R. B., Lithgow, T., & Mulhern, T. D., *Domain organization of the monomeric form of the Tom70 mitochondrial import receptor*. Journal of Molecular Biology 2009. **388**(5): p. 1043-1058.
223. Krissinel, E., & Henrick, K., *Inference of macromolecular assemblies from crystalline state*. Journal of Molecular Biology, 2007. **372**(3): p. 774-797.
224. Moreira, I.S., Fernandes, P. A., & Ramos, M. J., *Hot spots--a review of the protein-protein interface determinant amino-acid residues*. Proteins, 2007. **68**(4): p. 803-812.
225. Fattal, D.R., & Ben-Shaul, A., *A molecular model for lipid-protein interaction in membranes: the role of hydrophobic mismatch*. Biophysical Journal, 1993. **65**(5): p. 1795-1809.
226. Nickel, R., Ott, C., Dandekar, T., & Leippe, M., *Pore-forming peptides of Entamoeba dispar. Similarity and divergence to amoebapores in structure, expression and activity*. European Journal of Biochemistry, 1999. **265**(3): p. 1002-1007.
227. Andre, I., Linse, S., & Mulder, F. A., *Residue-specific pKa determination of lysine and arginine side chains by indirect 15N and 13C NMR spectroscopy: application to apo calmodulin*. Journal of the American Chemical Society 2007. **129**(51): p. 15805-15813.
228. Scholtz, J.M., Qian, H., Robbins, V. H., & Baldwin, R. L., *The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide*. Biochemistry, 1993. **32**(37): p. 9668-9676.
229. Vaccaro, A.M., Tatti, M., Ciaffoni, F., Salvioli, R., Serafino, A., & Barca, A., *Saposin C induces pH-dependent destabilization and fusion of phosphatidylserine-containing vesicles*. FEBS Letters, 1994. **349**(2): p. 181-186.
230. Mechler, A., Praporski, S., Atmuri, K., Boland, M., Separovic, F., & Martin, L. L., *Specific and selective peptide-membrane interactions revealed using quartz crystal microbalance*. Biophysical Journal 2007. **93**(11): p. 3907-3916.
231. Govorukhina, N., Bergsma-Schutter, W., Mazeret-Dubut, C., Mazeret, S., Drakopoulou, E., Bystrykh, L., Oling, F., Mukhopadhyay, A., Reviakine, I., Lai Kee Him, J., & Brisson, A., *Chapter 4: Self-Assembly of annexin A5 on lipid membranes*, in *Annexins: Biological Importance and Annexin-related Pathologies*. 2003. p. 67-78.

232. Janshoff, A., & Steinem, C., *Label-free detection of protein - ligand interactions by the quartz crystal microbalance*. *Methods in molecular biology*, 2005. **305**: p. 47-64.
233. Richter, R.P., & Brisson, A. R., *Following the formation of supported lipid bilayers on mica: a study combining AFM, QCM-D, and ellipsometry*. *Biophysical Journal*, 2005. **88**(5): p. 3422-3433.
234. Richter, R.P., Him, J. L., Tessier, B., Tessier, C., & Brisson, A. R., *On the kinetics of adsorption and two-dimensional self-assembly of annexin A5 on supported lipid bilayers*. *Biophysical Journal*, 2005. **89**(5): p. 3372-3385.
235. Sauerbrey, G., *The use of quartz oscillators for weighing thin layer and for micro weighing*. *Zeitschrift fuer Physik*, 1959. **155**: p. 206-222.
236. Rodahl. M., H.F., Krozer. A., Brzezinski. P., & Kasemo. B *Quartz crystal microbalance setup for frequency and Q-factor measurments in gaseous and liquid environments*. *Review of Scientific Instruments*, 1995. **66**: p. 3924-3930.
237. Rodahl, M., Hook, F., Fredriksson, C., Keller, C. A., Krozer, A., Brzezinski, P., Voinova, M., & Kasemo, B., *Simultaneous frequency and dissipation factor QCM measurements of biomolecular adsorption and cell adhesion*. *Faraday Discuss*, 1997(107): p. 229-246.
238. Rodahl, M., & Kasemo, B., *On the measurement of thin liquid overlayers with the quartz-crystal microbalance*. *Sensors and Actuators A: Physical* 1996. **54**: p. 448-456.
239. McCubbin, G.A., Praporski, S., Piantavigna, S., Knappe, D., Hoffmann, R., Bowie, J. H., Separovic, F., & Martin, L. L., *QCM-D fingerprinting of membrane-active peptides*. *European Biophysics Journal*, 2011. **40**(4): p. 437-446.
240. James, M., *Characterization of Thin-Film Surfaces and Interfaces Using Neutron Reflectometry*. *Australian Journal of Chemistry*, 2001. **54**: p. 487-491.
241. Wacklin, H.P., *Neutron reflection from supported lipid membranes*. *Current Opinion in Colloid & Interface Science*, 2010. **15**: p. 445-454.
242. Zhou, X.L., & Chen, S. H, *Theoretical foundation of X-ray and neutron reflectometry*. *Physics Reports*, 1995. **257**: p. 223-348.
243. Tun, Z., *Detection and characterization of ultra-thin films with neutron reflectometry*. *Acta Crystallographica*, 2009. **A65**: p. 1-4.
244. Lovell, M.R., & Richardson, R.M. , *Analysis methods in neutron and X-ray reflectometry*. *Current Opinion in Colloid & Interface Science*, 1999. **4**: p. 197-204.
245. Choi, E.J., & Foster, M. D. , *Effect of Flow on Human Serum Albumin Adsorption to Self-Assembled Monolayers of Varying Packing Density*. *Langmuir*, 2003. **19**(13): p. 5464-5474.
246. Sears, V.F., *Neutron scattering lengths and cross sectioirn*. *Neutron News* 1992. **3**(3): p. 26-37.
247. Schowen, K.B.J., *Transition States of Biochemical Processes*, ed. R.D. Gandour, and Schowen, R. L. 1978, New York: Plenum Press.
248. Johnson, J.M., et al., *Early steps of supported bilayer formation probed by single vesicle fluorescence assays*. *Biophys J*, 2002. **83**(6): p. 3371-9.

249. James, M., Nelson, A., Brule, A., Schulz, J.C., *Platypus: A time-of-flight neutron reflectometer at Australia's new research reactor*. J. Neutron Research 2006. **14**: p. 91-108.
250. James, M., Nelson, A., Holt, A., Saerbeck, T., Hamilton, W.A., Klose, F., *The multipurpose time-of-flight neutron reflectometer "Platypus" at Australia's OPAL reactor*. Nuclear Instruments and Methods in Physics Research A 2011. **632**: p. 112-123.
251. Nelson, A., *Motofit - integrating neutron reflectometry acquisition, reduction and analysis into one, easy to use, package*. J. Physics: Conference Series 2010. **521**: p. 012094.
252. Nimeri, G., Fredriksson, C., Elwing, H., Liu, L., Rodahl, M., & Kasemo, B., *Neutrophil interaction with protein-coated surfaces studied by an extended quartz crystal microbalance technique*. Colloids and Surfaces B: Biointerfaces, 1998. **22**(5): p. 255-264.
253. Biltonen, R.L., & Lichtenberg, D., *The use of differential scanning calorimetry as a tool to characterize liposome preparation*. Chemistry and Physics of Lipids, 1993. **64**: p. 129-142.
254. Lindsey, H., Petersen, N. O., & Chan, S. I., *Physicochemical characterization of 1,2-diphytanoyl-sn-glycero-3-phosphocholine in model membrane systems*. Biochimica et Biophysica Acta, 1979. **555**(1): p. 147-167.
255. Kucerka, N., Tristram-Nagle, S., & Nagle, J. F., *Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains*. The Journal of Membrane Biology 2005. **208**(3): p. 193-202.
256. Krueger, S., M.C.W., Majkrzak, C. F., Dura, J. A., Berk, N. F., Tarek, M., & Plan, A. L., *Investigation of Hybrid Bilayer Membranes with Neutron Reflectometry: Probing the Interactions of Melittin*. Langmuir, 2001. **17**: p. 511-521.
257. Woolverton, C., Willey, J., & Sherwood, L., *Prescott's Microbiology*. 2005: McGraw-Hill Higher Education.
258. Pozo Navas, B., Lohner, K., Deutsch, G., Sevcsik, E., Riske, K. A., Dimova, R., Garidel, P., & Pabst, G., *Composition dependence of vesicle morphology and mixing properties in a bacterial model membrane system*. Biochimica et Biophysica Acta, 2005. **1716**(1): p. 40-48.
259. Nakajima, Y., Qu, X. M., & Natori, S., *Interaction between liposomes and sarcotoxin IA, a potent antibacterial protein of Sarcophaga peregrina (flesh fly)*. The Journal of Biological Chemistry, 1987. **262**(4): p. 1665-1669.
260. Fernandez, D.I., Sani, M. A., Miles, A. J., Wallace, B. A., & Separovic, F., *Membrane defects enhance the interaction of antimicrobial peptides, aurein 1.2 versus caerin 1.1*. Biochimica et Biophysica Acta, 2013. **1828**(8): p. 1863-1872.
261. Wu, P.L., Chiu, C. R., Huang, W. N., & Wu, W. G., *The role of sulfatide lipid domains in the membrane pore-forming activity of cobra cardiotoxin*. Biochimica et Biophysica Acta 2012. **1818**(5): p. 1378-1385.
262. Cannon, B., Hermansson, M., Gyorke, S., Somerharju, P., Virtanen, J. A., & Cheng, K. H., *Regulation of calcium channel activity by lipid domain formation in planar lipid bilayers*. Biophysical Journal, 2003. **85**(2): p. 933-942.
263. Knappe, D., Piantavigna, S., Hansen, A., Mechler, A., Binas, A., Nolte, O., Martin, L. L., & Hoffmann, R., *Oncocin (VDKPPYLPRPRPPRRYNR-NH₂): a novel antibacterial peptide optimized*

- against gram-negative human pathogens*. Journal of Medicinal Chemistry 2010. **53**(14): p. 5240-5247.
264. Huang, H.W., *Action of antimicrobial peptides: two-state model*. Biochemistry, 2000. **39**(29): p. 8347-8352.
265. McGillivray, D.J., Valincius, G., Vanderah, D. J., Febo-Ayala, W., Woodward, J. T., Heinrich, F., Kasianowicz, J. J., & Lösche, M., *Molecular-scale structural and functional characterization of sparsely tethered bilayer lipid membranes*. Biointerphases, 2007. **2007**(2): p. 21-33.
266. Lehrer, R.I., *Primate defensins*. Nature Reviews Microbiology 2004. **2**(9): p. 727-738.
267. Piantavigna, S., Czihal, P., Mechler, A., Richter, M., Hoffmann, R., & Martin, L. L., *Cell penetrating apidaecin peptide interactions with biomimetic phospholipid membranes*. International Journal of Peptide Research and Therapeutics, 2009. **15**: p. 139-146.
268. Kang, S. J., L.B.J., *Solid-state NMR Study on Membrane Protein Structure in Biological Condition*. Journal of the Korean Magnetic Resonance Society, 2012. **16**: p. 103-110.