Engineered Pores for Stochastic Sensing and Single Molecule Studies



A thesis submitted for the degree of Doctor of Philosophy
in Chemical Biology

Leon E. O. Harrington

New College

2012

For my family

Acknowledgements

I would first like to thank Prof. Hagan Bayley and the BBSRC for financial support. During my time in the Bayley lab I have enjoyed a great deal of intellectual freedom and independence, for which I am very grateful. When I first started in the lab, it was Dr Stephen Cheley who took me under his wing and spent a great deal of time bringing me up to speed in the lab, taught me many useful tricks and has been a great mentor over the course of my DPhil. It took me a long time to become comfortable with electrophysiology and that I am now at all competent is largely thanks to Dr Ruhma Syeda, who not only taught me DIBs but provided a great deal of support during the time I spent collecting and analysing my data. I should also thank Ruhma and my other friends in the Bayley lab, past and present, with whom I have had a great deal of fun and who have made the last four years a most memorable episode of my life. I would like to thank my examiners, for taking the time to read and discuss my thesis. Lastly, I should like to thank my collaborator, Dr Stefan Knapp, for providing the kinases so essential to my work.

Leon Harrington

University of Oxford, 2011

Abstract

Engineered Pores for Stochastic Sensing and Single Molecule Studies

Submitted for examination for the degree of Doctor of Philosophy
Michaelmas Term 2011, Leon Harrington, New College

Stochastic sensing is a powerful approach for the detection of a range of analytes, yielding information at the single molecule level. Protein pores engineered for the binding of desired analytes are inserted into planar lipid bilayers and analyte binding is detected by observing the modulation of current flowing through the pore under an applied potential.

In this work, two complementary approaches for the detection of protein analytes were investigated. Chapter 2 describes the development of genetically encoded hetero-heptameric αHL stochastic sensors that bear a single Pim kinase consensus substrate sequence at the *trans* mouth of the pore. These were achieved by insertional fusion in the *trans* loop of a single subunit of the pore to yield a loop-constrained peptide, or via a novel method wherein site-specific proteolytic cleavage of the *trans* loop fusion yielded a peptide attached by a single terminus. The affinity of Pim-1 for the proteolytically-cleaved sensor was determined to be 99 nM. Additionally, the addition of MgATP to the sensor in the presence of Pim-1 produced an

irreversible change in current behaviour that is consistent with phosphorylation of the sensor.

The sensors and methods outlined here allow for sensitive, label-free detection of Pim-1 kinase. Further, the observation of phosphorylation in the presence of MgATP would suggest that it should be possible to observe catalytic turnover of the sensor substrate if the kinase activity is coupled with an active phosphatase. Further, it should be straightforward to modify the sensor to facilitate the screening of ATP-competitive inhibitors. Hence, the sensors developed here will be useful for the sensitive, label-free detection and assay of kinases by stochastic sensing.

In Chapter 3, a scheme is proposed in which outer membrane protein G (OmpG) pores that bind a desired target are selected from a combinatorial peptide library inserted in an extramembranous loop of OmpG. This would allow the selection of stochastic sensors against targets for which no suitable ligand exists. The suitability of OmpG for bacterial display was first demonstrated, before construction of the combinatorial library. However, OmpG expression had an adverse effect on cell viability that prevented successful screening. Additionally, a rapid method of characterising selected clones using droplet interface bilayers was developed that may be useful for high throughput applications.

Finally, Chapter 4 details efforts to express and engineer the Voltage-Dependent Anion Channel (VDAC) for stochastic sensing. This monomeric β-barrel membrane protein possesses a large lumen of diameter 25 Å. VDAC could therefore have potential for the stochastic sensing of substrates too large for the lumen of αHL or OmpG.

Declaration

I declare that no part of this thesis has been accepted, or is currently being submitted, for any degree of diploma or certificate or any other qualification in this university or elsewhere. Except where explicit reference is made to the work of others, the work contained in this thesis is my own, and is not the outcome of work done in collaboration.

Leon Harrington

University of Oxford, 2011

"It is better to fail in originality

than to succeed in imitation."

–Herman Melville

Contents

1 General introduction **1**

1.1 Stochastic sensing 2

1.1.1 Detection of ions and small molecules 5

1.1.2 Detection and translocation of polymeric species 8

1.1.3 Detection of protein-binding interactions 12

1.2 Single-channel recording 16

1.2.1 Planar Lipid Bilayers 17

1.2.2 Droplet Interface Bilayers 19

1.2.3 Single molecule kinetics 22

1.3 Pore structure and function 28

1.3.1 Alpha-haemolysin 28

1.3.2 Outer membrane protein G 30

1.3.3 The Voltage-Dependent Anion Channel 33

1.4 Protein kinases 35

1.4.1 Structure of the active conformation 39

1.4.2 Assay of kinase function 39

1.4.3 Kinase inhibition 46

1.5 Overview 48

2 Pim kinase detection by design 61

2.1 Introduction 61

2.1.1 The Pim family of kinases 62

2.1.2 Chapter overview 66

2.2 Results 67

2.2.1 Genetically encoded *trans* loop pores 67

2.2.2 Protease-cleavable pores 80

2.2.3 Phosphorylation 95

2.3 Conclusions 99

2.4 Future developments 102

2.5 Materials and methods 103

2.5.1 Plasmids 104

2.5.2 Protein preparation 109

2.5.3 Electrophysiology 113

3 OmpG for bacterial display 122

3.1 Introduction 122

3.1.1 Artificial protein binding peptides 123

3.1.2 Proposed methodology 133

3.1.3 Chapter overview 134

3.2 Results 136

3.2.1 Plasticity of OmpG extracellular loops 136

3.2.2 Expression system 141

3.2.3 OmpG bacterial display library 145

3.2.4 Further expression system optimisation 149
3.2.5 Rapid electrophysiology of OmpG mutants 152
3.3 Conclusions 160
3.4 Future developments 161
3.5 Materials and methods 162
3.5.1 Plasmids 163
3.5.2 Trehalose assay 169
3.5.3 Flow cytometry 170
3.5.4 Library construction 170
3.5.5 Electrophysiology 176

4 VDAC for stochastic sensing 185
4.1 Introduction 185
4.1.1 Chapter overview 186
4.2 Results 187
4.2.1 Recombinant expression in *E. coli* 187
4.2.2 *In vitro* protein expression 192
4.3 Conclusions 197
4.4 Future developments 199
4.5 Materials and methods 199
4.5.1 Plasmids 200
4.5.2 Recombinant protein expression and purification 203
4.5.3 *In vitro* transcription and translation 207

A Selected Plasmid Sequences 213

Abbreviations

αHL	Alpha-haemolysin
λPP	Lambda protein phosphatase
ADP	Adenosine diphosphate
AFM	Atomic force microscopy
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
Bis-tris	2,2-Bis(hydroxymethyl)-2,2',2"-nitrilotriethanol
CB6	Cucurbit[6]uril
CDR	Complementarity determining region
CMC	Critical micelle concentration
CML	Chronic myelogenous leukaemia
DDM	Dodecylmaltoside
DIB	Droplet Interface Bilayer
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
DOPA	1,2-dioleoyl-*sn*-glycero-3-phosphate
DPhPC	1,2-diphytanoyl-*sn*-glycero-3-phosphocholine
dsDNA	Double-stranded DNA
DTT	Dithiothreitol

EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
FACS	Fluorescence activated cell sorting
FP	Fluorescence polarisation
FRET	Förster resonance energy transfer
GdnHCl	Guanidinium hydrochloride
GPCR	G protein-coupled receptor
GUV	Giant unilamellar vesicle
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
IP3	Inositol 1,4,5-triphosphate
IPTG	Isopropylthiogalactoside
ITC	Isothermal titration calorimetry
IVTT	*In vitro* transcription and translation
LB	Luria broth
LDAO	Lauryldimethylamine oxide
LPS	Lipopolysaccharide layer
mAb	Monoclonal antibody
MACS	Magnetic activated cell sorting
MALDI-TOF	Matrix-assisted laser desorption/ionisation-time of flight
MBSA	MOPS-BSA-sodium chloride
MCS	Multiple cloning site
MIL	Maximum interval likelihood
MOPS	3-(*N*-morpholino)propanesulphonic acid
Ni-NTA	Nickel nitriloacetic acid
NMR	Nuclear magnetic resonance

OD	Optical density
OG	Octylglucoside
OmpG	Outer membrane protein G
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethyleneglycol
PIDA	*N*-propyliminodiacetic acid
PKA	Protein kinase A
PKIP	Heat stable protein kinase inhibitor peptide
PLM	Pimtide loop mutant
RNA	Ribonucleic acid
SAPE	Streptavidin-phycoerythrin
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide electrophoresis
SKM	Segmental k-means
SOC	Super-optimal broth with catabolite repression
TAE	Tris-acetate-EDTA
TCEP	Tris(2-carboxyethyl)phosphine
TE	Tris-EDTA
TEV	Tobacco etch virus
THF	Tetrahydrofuran
TRF	Time resolved fluorescence
Tris	Tris(hydroxymethyl)aminomethane

VDAC	Voltage dependent anion channel
WT	Wild type

Chapter 1

General introduction

Biological pores facilitate the passage of ions, molecules and macromolecules across membranes. They range from some that provide non-selective diffusion pathways through to those that exhibit exquisite selectivity. Others are involved in active transport processes and many act as conduits of information as well as chemical entities, interacting with chemical and proteinaceous ligands that trigger cascades along signal transduction networks within the cell [1]. While the study of the basic science of membrane proteins has greatly enhanced our knowledge of biology, these studies have also led to biological pores being engineered for use as nanoscale stochastic sensors [2].

Stochastic sensing has its origins in the development of single-channel electrical recordings of channel-forming biomolecules. The subsequent discovery that molecules (channel blockers) can be observed to bind reversibly to single channels and modulate the current flowing through them forms the phenomenological basis for stochastic sensing [3].

With the development of facile protein engineering techniques and the

discovery of robust biological pores, stochastic sensing has been applied to a range of analytes from ions to large macromolecules. The single-molecule nature of the technique allows determination of analyte concentration and the single-molecule kinetic parameters of binding, but also information that is often sufficient for analyte identification.

This chapter presents a historical overview of stochastic sensing followed by the biochemistry, principles and techniques underpinning this method and its application in this thesis.

1.1 Stochastic sensing

Stochastic sensing is a powerful approach for the detection of a range of analytes, yielding information at the single-molecule level. Protein pores engineered for the binding of desired analytes are inserted into planar lipid bilayers and analyte binding is detected by observing the modulation of current flowing through the pore under an applied potential. In its simplest manifestation, association and dissociation is observed as binary (on/off) fluctuations in the transmembrane current (Fig 1.1).

Analyte concentration can be determined from the frequency of binding events observed as current fluctuations. Moreover, the amplitude of the current modulation and the mean duration of the binding events can be used to infer analyte identity. Stochastic sensors are sometimes described as nanoscale Coulter counters [4]. However, Coulter counters typically measure the passage of analytes through microchannels, without any interaction between the analyte and sensor. While analyte discrimination of Coulter

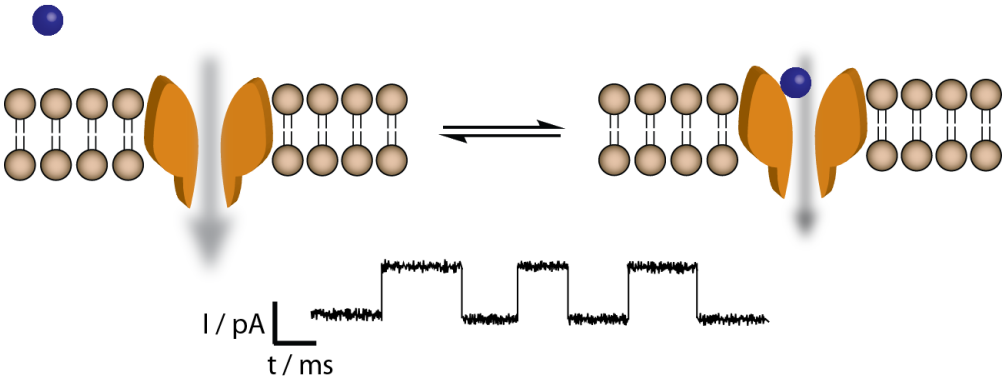


Figure 1.1: Principle of stochastic sensing. A protein pore (orange) with engineered affinity for a desired analyte (blue) is inserted in a planar lipid bilayer. When a potential is applied across the membrane, a current flowing through the pore is observed. Reversible binding of the analyte to the pore modulates this transmembrane current giving an observable signal.

counters is restricted to that based on size, stochastic sensors offer more specific means of analyte identification. Stochastic sensing via reversible current blockades is often referred to as resistive-pulse sensing.

Stochastic sensing is also possible without the production of reversible current blockades. Analytes that bind outside the pore lumen, for example to a ligand on a polyethylene glycol (PEG) linker attached to one entrance to the pore, may not influence the current flow through the pore. However, these analytes may still be detected indirectly, for example through a change in the pore's rectification properties, or through the modification of pore noise on binding [5].

Solid state pores in materials such as organic polymers or inorganic materials such as silicon nitride are also widely used. However, they suffer several limitations compared to biological pores. For example, it is very difficult to fabricate nanoscale solid state pores with consistent dimensions and topology between different pores. Further, biological pores are readily engineered

site-specifically by standard mutagenesis techniques, whereas the modification of solid state pores must be carried out by chemical means that lack specificity. As solid state pores are not relevant to this work they will not be discussed further here. Their use for stochastic sensing has been reviewed elsewhere [6, 7].

Through the development of protein modification and engineering techniques, stochastic sensing has been applied to a broad range of analytes, including ions and small molecules, polymers, polynucleotides, polypeptides, as well as protein-ligand interactions. A range of examples exists for strategies using both engineered and non-engineered pores. Polymeric analytes are commonly detected using non-engineered pores, where their detection arises due to translocation through a pore rather than an interaction with a binding site. However, engineering may be used to alter detection properties, for example by slowing down translocation to increase sensitivity [8]. Additionally, some polymers such as polypeptides may be detected through interactions with engineered binding sites, as with smaller analytes [5, 9–11].

Ions and small molecules have mostly been detected through the use of engineered pores, although some small molecules with pre-existing affinity can be studied without engineering [12]. Another important development has been the use of molecular adapters such as cyclodextrins, which provide portable recognition elements for a variety of small molecules [13]. Further, the targeted chemical modification of the pore-forming toxin alpha-haemolysin (αHL) has allowed ligands to be attached via linker molecules such as PEG, enabling the detection of protein-ligand interactions [5]; crucially, this approach introduces binding sites that lie outside the pore. The

following sections summarise key examples of these different stochastic sensing approaches.

1.1.1 Detection of ions and small molecules

For many non-engineered protein pores, the passage of molecules and analytes occurs faster than can be resolved by the apparatus commonly used for single channel recording [14]. Protein pores must therefore be engineered to introduce or enhance affinity. Because of the single-molecule nature of detection this need *not* necessarily be specific for a particular analyte species, as it is often possible to discriminate and group individual events when different analytes give rise to different event properties such as block amplitude or mean event duration, as is commonly the case. For the detection of ions and small molecules, three strategies have been used: to genetically engineer the pore to introduce amino acids that create enhanced analyte affinity; to use macrocyclic compounds that reside within the transmembrane pore and provide affinity for analyte molecules; or to chemically modify the pore with sensing elements such as ion chelators.

One recent notable example of analyte sensing using non-engineered pores is the single molecule observation of the interaction between antibiotics and the bacterial porin OmpF [12]. The authors found that ampicillin, several other penicillins and cephalosporins interact strongly with the pore at a constricted region. It was hypothesised that this affinity had evolved to increase translocation and permeability rates of antibiotics.

Detection using engineered pores

Alpha-haemolysin (αHL) has been engineered to detect a range of divalent metal ions at nanomolar concentrations [15–18]. This was achieved by introducing a tetra-histidine (4H) binding site. Later implementations used the ability to form mixed heteroheptamers to generate a heptamer containing a single subunit bearing the 4H site [17, 18]. Divalent ions such as Co(II), Ni(II) and Cu(II) each gave characteristic signals that made it possible to distinguish different ions within a mixture. Further, it was shown that the sensor can be linearly calibrated without any knowledge of the kinetic mechanism, provided the metal ions bind independently.

Other genetic modifications have enabled the detection of small molecules. For example, engineering rings of aromatic residues within the αHL lumen to form binding sites for nitroaromatic compounds [19]. Various nitroaromatic compounds could be distinguished from each other on the basis of differing amplitudes of current block or differing mean binding event durations.

A similar engineering strategy but using rings of arginine residues enabled the detection of phosphate anions and the phosphate group-bearing second messenger inositol 1,4,5-trisphosphate (IP3) [20]. A ring of seven arginines allowed phosphate ions to bind, giving rise to an almost complete channel block. In an attempt to generate pores with increased affinity to phosphate esters, a pore with fourteen arginines (i.e. two rings of seven) was created that had nanomolar affinity for IP3 and picomolar affinity for myo-inositol hexaphosphate (IP6).

Detection using molecular adapters

Cyclodextrins are cyclic oligosaccharides that are formed of 5 or more α-1,4-linked α-D-glucopyranoside units. α, β and γ cyclodextrins comprising 6, 7 and 8 sugar subunits respectively, are toroidal in shape. The external surface is hydrophilic whereas the interior, while not hydrophobic, is considerably less hydrophilic. Cyclodextrins are therefore able to form a range of inclusion compounds with small hydrophobic molecules.

Cyclodextrins can be used as molecular adapters for analyte sensing. Gu *et al.* found that β-cyclodextrin bound αHL to give a partial current block. When bound, it was found that this cyclodextrin functioned as a molecular adapter and could be used to detect several adamantane derivatives [13]. In an important paper for the nascent field of nanopore sequencing, use of a heptakis-(6-deoxy-6-amino)-β-cyclodextrin as a molecular adapter in αHL allowed the detection and discrimination of 2'-deoxyribonucleoside 5'-monophosphates as well as methylcytosine [21, 22]. Later work showed that molecular adapters can also be covalently attached [22, 23]. The molecular nature of cyclodextrin interaction with αHL has recently been characterised through structural and functional studies [24].

Other molecular adapters have also been demonstrated, for example cucurbit[6]uril (CB6) was used as a carrier for tetrahydrofuran (THF) [25]. Here, the CB6 was described as a carrier, as association and dissociation of THF from the αHL–CB6 binary complex was not observed. This arises due to the association and dissociation rates of THF with CB6 being much slower than the association and dissociation rates of CB6-THF with αHL.

In another example, several cyclic peptides were shown to act as molecular adapters. Positively charged peptides were able to act as binding sites for several polyanions such as IP3 [26].

Finally, Chen *et al.* used the heptakis-(6-deoxy-6-amino)-β-cyclodextrin molecular adapter on an engineered OmpG pore to detect adenosine diphosphate [27]. This exemplifies one benefit of molecular adapters, in that they provide the ability to transplant a particular binding site between structurally different pores.

Detection using chemically modified pores

Hammerstein *et al.* created a sensor for divalent zinc ions by the chemical attachment of 'half-chelator' ligands comprising the *N*-propyliminodiacetic acid (PIDA) group and an iodoacetamide group for covalent attachment via site-directed cysteine modification [28]. A heteroheptameric αHL pore was formed containing a single modified subunit with two cysteine residues at positions 117 and 143. When PIDA ligands were attached to each cysteine residue, the two ligands were able to chelate zinc ions in a two-step process, with an ion being chelated by either one or both ligands. The single-molecule nature of this approach enabled all major rate constants involved in this process to be determined.

1.1.2 Detection and translocation of polymeric species

The facilitated passage of biopolymers across biological membranes is commonplace and essential for life. In normal cell function, nascent and folded

proteins translocate the membranes of the endoplasmic reticulum [29, 30], mitochondrion [31, 32], chloroplast [33] and the nucleus [34]. Additionally, biopolymer translocation features in some mechanisms of infection, such as for anthrax [35] and phage [36].

In certain *in vivo* settings some polymers are driven through pores electrophoretically [37]. Therefore, model systems based on protein pores in the planar lipid bilayer setup should be useful both for polymer sensing and studying the fundamentals of the translocation process.

Polymers

The first study of uncharged polymer translocation used polyethyleneglycol (PEG) chains to probe the lumen of several protein pores [38]. Bezrukov *et al.* indirectly studied the translocation of PEG through single alamethicin pores and were able to infer the diffusion coefficient of PEG and the average number of molecules within the pore [39]. Subsequent work using αHL by Movileanu *et al.* resolved discrete current blockades due to PEG translocation [40]. In a later study, it was possible to determine the number of ethylene glycol repeat units in PEG, as these molecules caused distinct, mass-dependent conductance states that could be de-convoluted by a statistical method [41]. The relative blockade distribution closely resembled a conventional MALDI-TOF mass spectrum of polydisperse PEG.

Polynucleotides

The translocation of deoxyribonucleic acid (DNA) through pores has been known for some time to be a biologically important process [42, 43]. In

recent years the application of stochastic sensing to polynucleotide strands has generated considerable interest due to the potential for ultrafast nanopore sequencing [44].

Working on the hypothesis that strands of DNA or ribonucleic acid (RNA) could be electrophoretically driven through nanopores, Kasianowicz *et al.* were first to demonstrate this using the αHL pore [45]. Single stranded DNA or RNA homopolymers produced large, reversible current blockades when added to the *cis* side of an αHL-containing membrane under positive applied potential, where *cis* was in this case defined as the chamber connected to the working electrode and to which αHL protein was added. To confirm the translocation hypothesis, competitive polymerase chain reaction (PCR) analysis was performed on samples from the *trans* chamber some time after DNA had been added to the *cis* chamber. This showed that the number of DNA molecules present was consistent with the event rate observed in single channel recordings. Further, the block amplitude and dwell time were found to be dependent on the nature of the homopolymer and its length [45, 46]. The 1.2 nm narrowest diameter of αHL implies that single strands are only able to translocate in a single-file manner.

In steps towards to the resolution of the nucleotide composition of strands, Akeson *et al.* and Meller *et al.* demonstrated that pyrimidine and purine segments in RNA, and DNA homopolymers of different composition could be discriminated [46, 47]. Differing polymers gave different translocation times and different amplitudes of block.

The translocation of double-stranded DNA (dsDNA) through the large pore of the voltage-dependent anion channel (VDAC) as well as membranes

containing *Bacillus subtilis* ion channels has also been investigated [48, 49]. In these studies double-stranded DNA was able to pass across a planar lipid bilayer only if the pore-forming proteins were present. However, these channels possess undesirable characteristics such as voltage-gating that appear to have precluded further exploitation for sequencing purposes. Recently, the MspA porin from *Mycobacterium smegmatis* has been engineered for enhanced polynucleotide detection [50]. MspA possesses structural features that may allow it to be engineered to be more responsive to individual nucleotides [51].

Although the goal of nanopore sequencing is compelling, some challenges remain before nucleotide resolution can be achieved. First, studies using immobilised DNA strands have shown that αHL possesses multiple 'recognition points'—locations within the pore where individual bases can be resolved [52–55]. This could complicate measurements, but equally could be used to provide redundant measurements and hence provide a built-in 'proof-reading' mechanism to improve the reliability of base reads [56]. Second, to achieve adequate temporal resolution, the translocation speed of DNA through αHL must be reduced without reducing the ability to resolve amplitude differences. Currently, single nucleotide resolution of translocating DNA has yet to be achieved.

Polypeptides

The translocation of peptides and proteins is more complicated than for single DNA molecules due to the variety of amino acid side-chains and their wide range of chemical properties. In the first study of this nature, simple cationic

α-helical peptides were translocated through the αHL pore [57]. Translocation was slowed due to electrostatic and steric interactions. However, in a subsequent study it was shown that placement of negatively charged residues at the entrance and exit of the β-barrel created electrostatic traps that lower the energy barrier of translocation for cationic peptides [58].

For larger peptides and proteins, unfolding and unzipping processes become key to translocation if the polypeptides are larger than the pore lumen. Goodrich *et al.* found that peptides with β-hairpin structure gave longer current blockades than peptides lacking the hairpin structure [59]. Translocation time was heavily dependent on the electrical force applied. This suggests that unfolding of the hairpin may be the rate limiting step.

Attempts at investigating protein folding have been made using αHL pores [60]. By using different concentrations of denaturant, populations of partially unfolded protein and completely unfolded protein gave rise to blockades of long and short duration respectively, consistent with the principle that complete unfolding is required for translocation.

1.1.3 Detection of protein-binding interactions

Ligand-receptor interactions are crucial for many cellular functions. Biological pores can provide ideal model platforms for studying such interactions at the single-molecule level.

Early studies used pore-forming peptides such as gramicidin and were not strictly single-molecule detection methods. For example, Cornell *et al.* constructed a sensor comprising a supported lipid bilayer tethered to a gold

electrode [61]. One leaflet contained tethered gramicidin monomers, with the other leaflet containing free biotinylated gramicidin monomers. A biotinylated Fab fragment that bound the desired analyte was attached to the biotinylated gramicidin via streptavidin. Dimerisation of gramicidin monomers between the two leaflets gave a relatively high channel conductance in the absence of analyte. When analyte was added at picomolar concentrations, the Fab fragments were cross-linked, reducing the diffusion rate of gramicidin monomers in the non-tethered leaflet and thus reducing the conductance across the membrane because they were unable to form dimers with channels immobilised in the inner half of the membrane.

The αHL pore has been used to study a variety of protein-ligand interactions. In an early study, αHL was site-specifically chemically modified with a PEG chain that terminated with a biotin molecule [5]. The PEG chain was of such length that it could be reversibly captured from either the *cis* or *trans* side on addition of streptavidin or anti-biotin antibody, yielding distinct current signatures. When captured by streptavidin from the *cis* side, short blockades due to the PEG chain are abolished, but when captured from the *trans* side, this abolition is also accompanied by a decrease in channel conductance due to the presence of the PEG chain within the transmembrane constriction of the pore.

A topic of key importance in biology is that of multivalency or cooperativity, where a molecule possessing several binding sites may bind simultaneously to multiple ligands [62]. The topic is of great interest as it underpins a variety of biological interactions such as antibody binding, cell adhesion and viral attachment. Single-molecule methods have the potential for greater

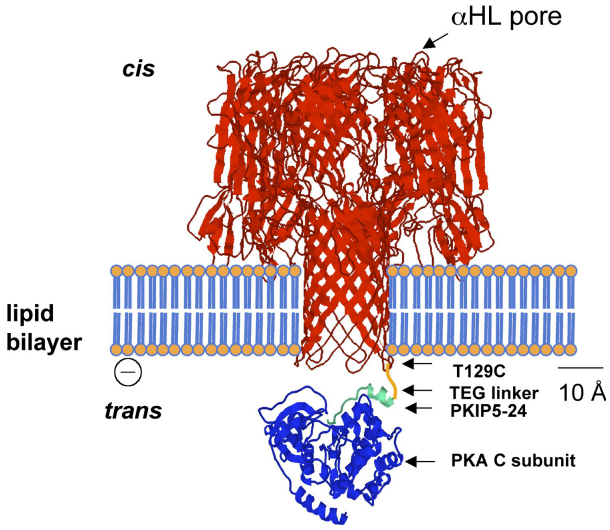


Figure 1.2: Molecular model of the PKIP-pore bound to the catalytic subunit of cAMP-dependent protein kinase. The hetero-heptameric αHL pore contains one subunit modified with residues 5-24 of PKI and six unmodified subunits. The N-terminus of PKIP5-24 is tethered to a Cys residue at position 129 through a disulphide bond and a short flexible tetra(ethylene glycol) (TEG) linker. Reprinted from Xie *et al.* [11] with permission from Elsevier.

insight than bulk methods as individual monovalent and divalent binding events can be observed. Howorka *et al.* attached single or multiple copies of a disaccharide ligand to the αHL pore [9]. This allowed the study of the kinetics of lectin binding as mono- and di-valent binding events were observed.

Other studies using αHL have shown that ligands other than small molecules may be used to study protein-ligand interactions. For example, Xie *et al.* covalently attached a single peptide corresponding to amino-acids 5–24 of the naturally occurring protein kinase A inhibitory peptide (PKIP) to the *trans* mouth of the αHL pore (Fig 1.2) [11]. When the catalytic subunit of protein kinase A was added to the *trans* side, binding events were observed as current blockades under a negative applied potential. As well as allowing the measurement of single-molecule kinetic parameters, the synergistic effect

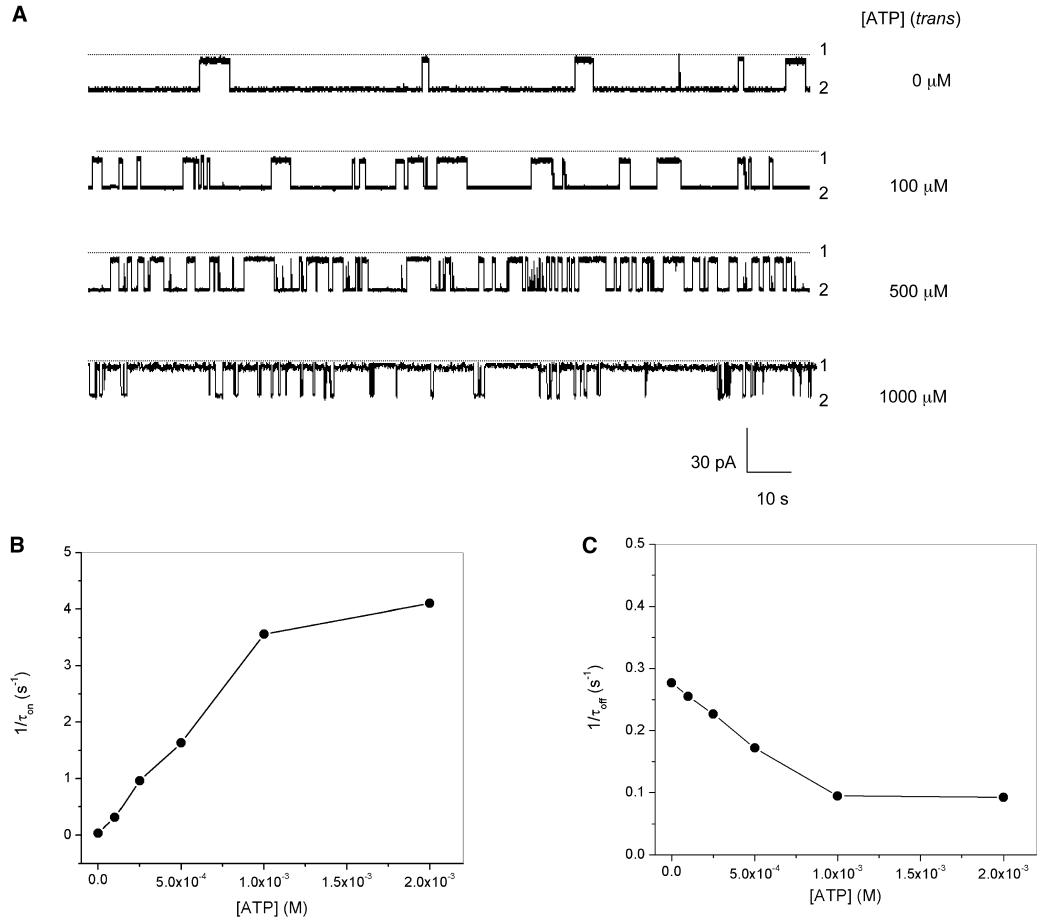


Figure 1.3: Effect of MgATP on the Binding of C Subunit to the PKIP-pore. The PKIP-pore was introduced from the *cis* chamber, and the C subunit and MgATP from the *trans* chamber. Both chambers contained 300 mM KCl, with 15 mM MOPS titrated with KOH (pH 6.8), at 25 ± 1°C. Recordings were made at −80 mV. A) Single-channel currents at various concentrations of ATP and 10 mM $MgCl_2$ in the presence of 100 nM C subunit. Level 1, zero current; level 2, free PKIP-pores. B) Plot of $1/\tau_{\rm on}$ versus the concentration of ATP, where $\tau_{\rm on}$ is the mean inter-event interval. C) Plot of $1/\tau_{\rm off}$ versus the concentration of ATP, where $\tau_{\rm off}$ is the mean blockade time. Reprinted from Xie *et al.* [11] with permission from Elsevier.

of MgATP binding was also observed (Fig 1.3). Subsequent to this study, a genetically-encoded kinase sensor was constructed where PKIP flanked by glycine/serine linkers was inserted into a single *trans*-mouth loop of αHL [10]. This circumvented some of the technical challenges of preparing chemically modified pores, such as coupling efficiencies. Detection occurred in a similar fashion to the chemically modified pore but with additional fast events in the open state due to the motion of the extended loop at the pore *trans* mouth.

1.2 Single-channel recording

Perhaps the most common technique for the electrical measurement of a single ion channel in a membrane is patch clamping, as developed by Neher and Sakmann [63]. In patch clamping, a glass pipette containing an electrode is applied to a cell membrane and a high resistance (>1 GΩ) seal is formed by suction. Different configurations of this basic setup exist that allow single-channel or whole-cell currents to be measured. The patched area of membrane can also be excised from the cell, allowing control of the bath solution composition. Once patched, current can be measured through the channel(s) by holding the membrane potential constant using a patch clamp amplifier.

While patch clamping is an excellent and widely used method for the study of ion channels, in certain applications it may have disadvantages. For example, internal membranes are not readily patchable as they are not directly accessible to the pipette. Further, recordings can be complicated by the presence of other endogenous proteins and cytosolic components.

In many situations, it would be desirable to have an experimental system with fully defined chemical composition. One solution is to reconstitute the protein of study into giant unilamellar vesicles (GUVs), which can then be patch-clamped as if they were cells [64]. Another approach is to use planar lipid bilayers, one of the two techniques used for the work contained in this thesis.

1.2.1 Planar Lipid Bilayers

Planar lipid bilayers, sometimes called Black Lipid Membranes due to their appearance under reflected light, can be formed in two ways. The first was developed by Mueller *et al.* in 1962 [65]. In this method, a solution of lipid in an organic solvent such as decane is painted onto a micron-scale aperture in a Teflon film that separates two chambers containing aqueous phases. The organic solvent wets the hydrophobic Teflon around the aperture and the lipid molecules spontaneously form monolayers at the interface. Due to the wetting effect the film thins in the centre until a bilayer is formed. In practice, however, there is significant residual solvent present in the bilayer with this method.

Ten years later, Montal and Mueller developed an improved method for 'folded' bilayers [67], the method used in this thesis for creating planar lipid bilayers. In this setup, a 20 μm Teflon film separates two 1 ml chambers machined in a Delrin block (Fig 1.4). An aperture of $\sim$100 μm is present in the centre of the Teflon film. The aperture is first pre-treated with hexadecane (10% v/v in pentane). Aqueous buffer is then added to both chambers. A

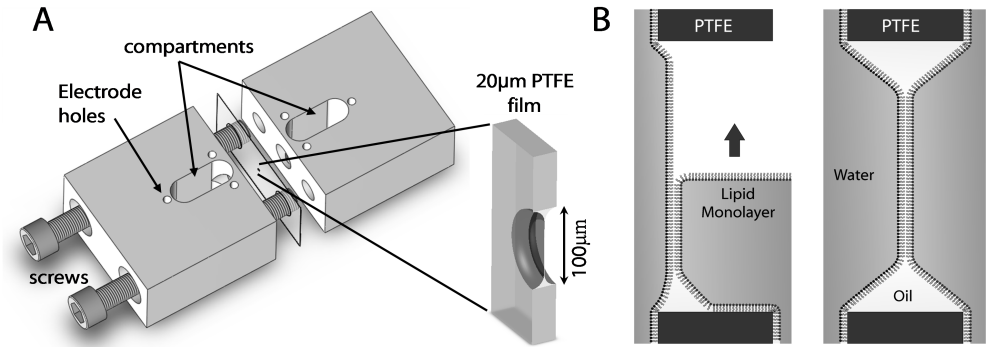


Figure 1.4: Planar lipid bilayer experimental setup. A) A typical chamber for the formation of planar lipid bilayers, shown in exploded form. Two blocks of Delrin are separated by a thin film of Teflon containing a small aperture. The apparatus is held together by two screws and sealed with silicone sealant around the circumference of the Teflon film exposed to the buffer-containing compartments. B) The Montal-Mueller method of bilayer formation. The aperture is pre-treated with hexadecane before lipid monolayers are sequentially folded across both sides to form a bilayer. Reproduced from Maglia *et al.*, [66].

few microlitres of lipid dissolved in pentane are then applied to the aqueous surface of both chambers. By pipetting, the aqueous surface level is lowered below the aperture and then raised back up. Repetition on both sides of the aperture leads to two lipid monolayers being folded together to form a bilayer. This method is a little faster than painting, and the bilayers contain less residual solvent. With care, it is also possible to form asymmetric bilayers with this method [67].

To perform electrical measurements, Ag/AgCl electrodes are most commonly used, with one electrode placed in each chamber. Detergent solubilised protein is added to one chamber and spontaneous insertion is then typically encouraged by application of potential and stirring of the chamber contents, both having the effect of inducing small defects in the lipid bilayer that provide a lower energy barrier for protein insertion. Protein may also be reconstituted into the bilayer through the fusion of proteoliposomes [68].

One disadvantage of planar lipid bilayers compared to patch clamp exper-

iments is that planar lipid bilayers typically have much higher capacitances than a patch. This leads to a much higher background noise than for patch clamping. This noise can of course be reduced by using chambers with smaller apertures, but a compromise has to be found as it will be progressively more difficult to obtain channel reconstitution the smaller the bilayer. Planar lipid bilayers are metastable and inherently fragile. A planar lipid bilayer typically lasts in the order of several hours under ideal conditions.

Conventions

In electrophysiology, the two 'sides' of the membrane are usually defined as *cis* or *trans*. In most patch-clamp experiments, *cis* is defined as the side of the working electrode (i.e. the patch pipette), and *trans* as the side of the ground electrode. In planar bilayer studies *cis* may be defined as the side of the ground electrode or the side to which protein is added. Often in the literature this convention is not clearly defined.

In this work, the convention taken is that the *cis* side is the side to which protein is added *and* that is connected to the ground electrode. For αHL, which inserts vectorially if added to a freshly formed bilayer, the cap domain will be in the *cis* chamber, which therefore corresponds to the physiological extracellular side and hence giving the same polarity as a cell.

1.2.2 Droplet Interface Bilayers

A recent technique for single channel recording that overcomes some of the limitations of planar lipid bilayers is that of Droplet Interface Bilayers (DIBs).

Although developed only recently, the underlying principle had in fact been demonstrated in 1966 by Tsofina *et al.*, [69]. Tsofina *et al.* described the formation of a lipid bilayer between two aqueous phases separated by a solution of lipid in heptane. An oil phase of heptane containing lipids was placed on top of an aqueous phase and a pipette containing an aqueous solution with a small volume at the tip was lowered into the oil phase. When the pipette approached the heptane-aqueous interface, the two monolayers present assembled into a lipid bilayer.

Recently, our laboratory developed a new method based on this principle and named it Droplet Interface Bilayers [70]. Around the same time similar methods, but based on microfluidics, were also developed by Funakoshi *et al.* and Malmstadt *et al.*, [71, 72]. In our DIB system, two fine ($\sim$100 μm diameter) Ag/AgCl electrodes are coated at their ends with a thin layer of agarose to provide a hydrophilic interface to support freely hanging aqueous droplets (Fig 1.5). The electrodes are then submerged in a bath of oil (typically hexadecane or squalene). Small aqueous volumes (typically 200 nl) are hand pipetted onto the agarose-coated electrode tips.

There are several different configurations depending on where the lipid is contained. In the 'lipid out' configuration, lipid is dissolved in the oil phase [70]. When the aqueous droplets are submerged in the oil phase, a lipid monolayer forms at the interface of each droplet. In the 'lipid in' configuration, lipid is present in the aqueous buffer solution in the form of liposomes [73]. When submerged in the oil phase, the liposomes fuse at the interface to form a lipid monolayer. The two configurations may also be combined, i.e. lipids may be present both in the oil and aqueous phases.

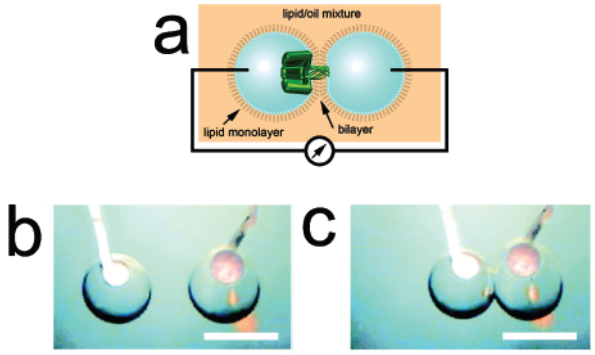


Figure 1.5: Droplet Interface Bilayer (DIB) experimental setup. A) Schematic of a DIB. A bilayer is formed between two aqueous droplets immersed in a lipid/oil mixture. An αHL channel is depicted inserted in the bilayer and current flowing through the pore is measured via Ag/AgCl electrodes inserted into each droplet. B) Bright-field image of two droplets each of approximately 200 nl volume immersed in lipid/oil mixture and suspended on two Ag/AgCl electrodes. Droplets are typically incubated for 5–10 minutes to allow monolayer formation. C) The droplets are then brought together whereupon a bilayer spontaneously forms. Reprinted with permission from Holden *et al.*, [70]. Copyright 2007 American Chemical Society

After a short stabilisation period (typically 5–10 minutes), the droplets can be brought together by means of micromanipulators, which are attached to the electrodes. Oil is displaced from the interface of the two droplets and the two monolayers form a bilayer. *Cis* and *trans* sides are defined in the same way as for planar lipid bilayers, with the grounded droplet containing protein defined as *cis* and the droplet containing the working electrode being defined as *trans*.

The DIB system carries many advantages. The small volume of the aqueous droplets means very little material, such as potentially precious protein or blockers, is consumed. DIBs are considerably more robust than planar lipid bilayers, having lifetimes that can range from days to weeks. DIBs are also easy to physically manipulate, for example the bilayer size is easily regulated by moving the droplets closer or further apart. This is particularly

useful as one can initially form a large bilayer to maximise the likelihood of channel insertion, before reducing the bilayer size to minimise further insertion and reduce the current noise of the bilayer. Droplets can also be joined and rejoined without cross-contamination, which facilitates screening applications [74].

Using the lipid-in method, asymmetric bilayers can be readily formed [73]. Other unique features such as the ability to form networks of many droplets and devices that mimic electrical circuits have been demonstrated [70, 75] and reviewed elsewhere [76].

1.2.3 Single molecule kinetics

A variety of characteristics of single-molecule detection experiments can be analysed, such as the amplitude of block, the inter-event interval, the event dwell time, and rate constants for transitions between different states. These parameters can give insight into the kinetics of the binding process being observed at a level that is often far deeper than ensemble measurements. For example, rate constants for all steps in a chemical reaction can be determined directly at the single-molecule level, whereas an ensemble measurement will only reveal that of the rate determining step [28, 77–80]. While it will be shown that the kinetic equations for stochastic sensing are functionally analogous to those for ensemble measurements, the physical interpretation is different. In this section, the mathematical equations relevant to stochastic sensing are derived and explained in the context of single-molecule kinetics.

Single-molecule kinetics of a unimolecular reaction

First we consider a macroscopic unimolecular elementary reaction:

$$\mathrm{A} \xrightarrow{k_{\mathrm{AB}}} \mathrm{B} \tag{1.1}$$

where the concentration of A is not replenished during the reaction. For a unimolecular reaction the rate is first-order with respect to the concentration of the reactant, A:

$$\frac{\mathrm{d[A]}}{\mathrm{d}t} = -k_{\mathrm{AB}}[\mathrm{A}] \tag{1.2}$$

Integration and rearrangement of Equation 1.2 gives:

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-k_{\mathrm{AB}}t} \tag{1.3}$$

Two useful parameters for describing reaction rate are the half-life, $t_{\frac{1}{2}}$ and the time constant, τ, defined as:

$$t_{\frac{1}{2}} = \frac{\ln 2}{k_{\mathrm{AB}}} \qquad \tau_{\mathrm{AB}} = \frac{1}{k_{\mathrm{AB}}} \tag{1.4}$$

where the half-life of a reaction is the time taken for the concentration of reactant to reach half its initial value, and the time constant is the time taken for the concentration of reactant to reach $1/e$ of its initial value. Substitution of the equation for the time constant into Equation 1.3 gives the equation:

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-\frac{t}{\tau_{\mathrm{AB}}}} \tag{1.5}$$

In a macroscopic reaction, at $t = t_{\frac{1}{2}}$ half of the initial number of reactant molecules have undergone reaction to give product. When considering the reaction at the single-molecule level, it is immediately clear that this interpretation is not valid—the state of the single molecule is binary; it is either unreacted or reacted. At the single-molecule level, the conversion of A to B is stochastic (random, non-deterministic). Hence, rather than describe the reaction in terms of the concentration of reaction, we instead describe it in terms of the probability that A has not yet reacted, i.e. the survival probability.

For a single molecule unimolecular reaction, the probability of a molecule having not yet reacted at a time, t, is given by:

$$p(t) = e^{-k_{\mathrm{AB}}t} \tag{1.6}$$

The probability that the dwell time lies in a particular interval is defined in terms of the probability density function, $f(t)$, as:

$$p(t_1 \leq t \leq t_2) = \int_{t_1}^{t_2} f(t)\mathrm{d}t \tag{1.7}$$

and so, as the probability of a molecule surviving until a time, τ, is given by Equation 1.6, then:

$$p(\tau \leq t < \infty) = \int_{\tau}^{\infty} f(t)\mathrm{d}t = e^{-k_{\mathrm{AB}}\tau} \tag{1.8}$$

The probability density function may therefore be determined by inspection:

$$p(\tau \leq t < \infty) = \int_{\tau}^{\infty} k_{\mathrm{AB}} e^{-k_{\mathrm{AB}} t} \mathrm{d}t \tag{1.9}$$

$$= -e^{-k_{\mathrm{AB}} t} \Big|_{\tau}^{\infty} = e^{-k_{\mathrm{AB}} \tau} \tag{1.10}$$

so explicitly:

$$f(t) = k_{\mathrm{AB}} e^{-k_{\mathrm{AB}} t} \tag{1.11}$$

The mean dwell time can be determined from the probability density function:

$$\overline{\tau} = \int_{0}^{\infty} t f(t) \mathrm{d}t = \int_{0}^{\infty} t k_{\mathrm{AB}} e^{-k_{\mathrm{AB}} t} \mathrm{d}t = \frac{1}{k_{\mathrm{AB}}} \tag{1.12}$$

Hence we see that the mean dwell time is the reciprocal of the macroscopic rate constant. This connects us to the ensemble measurement, in that we see that the ensemble time constant of a unimolecular reaction is the mean survival time of the reactant molecules.

For a large number of measurements of dwell times in the unreacted state $(\tau_1, \tau_3, \tau_5, \ldots)$, a histogram may be constructed by calculating the proportion of the time that the random variable τ falls within certain ranges, determined by the binning width t_{bin}. For single-molecule planar bilayer experiments, the binning width should be such that there are a sufficient number of bins for fitting to be performed (typically > 10). Good fitting results are usually obtained with a binning width near or below the characteristic filtering time ($t_f = 1/\nu$ where ν is the filtering frequency). As the number of measured events increases, this frequency histogram will approach the probability density function. For these experiments, several hundred events are normally

desirable for good statistics.

As the frequency histogram resembles the probability density function as the number of events increases, we can fit the histogram with a function of the form:

$$f(t) = A\exp(-\frac{t}{\overline{\tau}}) + C \tag{1.13}$$

In our model, it is expected that $C \to 0$ for a good fit.

Reversible single-molecule reactions

For a reversible unimolecular reaction of the form:

$$\mathrm{A} \underset{k_{\mathrm{BA}}}{\overset{k_{\mathrm{AB}}}{\rightleftharpoons}} \mathrm{B} \tag{1.14}$$

which is comprised of forward and reverse unimolecular reactions, the result of Equation 1.12 leads to the following rate constants:

$$k_{\mathrm{AB}} = \frac{1}{\overline{\tau}_{\mathrm{AB}}} \tag{1.15}$$

$$k_{\mathrm{BA}} = \frac{1}{\overline{\tau}_{\mathrm{BA}}} \tag{1.16}$$

The equilibrium constant may then be expressed in terms of the forward and reverse rate constants and hence in terms of the mean dwell times:

$$K = \frac{k_{\mathrm{AB}}}{k_{\mathrm{BA}}} = \frac{\overline{\tau}_{\mathrm{BA}}}{\overline{\tau}_{\mathrm{AB}}} = \frac{[B]}{[A]} \tag{1.17}$$

The mean dwell time can be simply defined as the statistical mean of the sample of single-molecule dwell times measured:

$$\overline{\tau}_{\mathrm{BA}} = \frac{\sum \tau_{\mathrm{BA}}}{n_{\mathrm{B}}} \tag{1.18}$$

where n_{B} is the number of dwell times measured that correspond to species B. Under equilibrium conditions $n_{\mathrm{A}} = n_{\mathrm{B}}$ and hence:

$$K = \frac{\sum \tau_{\mathrm{BA}}}{\sum \tau_{\mathrm{AB}}} \tag{1.19}$$

For a bimolecular reaction the ensemble rate equation is:

$$\frac{\mathrm{d[A]}}{\mathrm{d}t} = -k_{\mathrm{AB}}[\mathrm{A}][\mathrm{B}] \tag{1.20}$$

However, in single-molecule experiments, for the following scenario:

$$\mathrm{P} + \mathrm{A} \underset{k_{-\mathrm{PA}}}{\overset{k_{+\mathrm{PA}}}{\rightleftharpoons}} \mathrm{P.A} \tag{1.21}$$

where P is a pore molecule and A is a ligand that reversibly binds to the pore, the change in [A] is negligible over the course of the reaction and so the forward rate equation takes the form of a pseudo-first-order reaction:

$$\frac{\mathrm{d[P]}}{\mathrm{d}t} = -k_{+\mathrm{PA}}[\mathrm{A}]_0[\mathrm{P}] = -k'_{+\mathrm{PA}}[\mathrm{P}] \tag{1.22}$$

and hence the corresponding expressions for the mean dwell times apply:

$$\overline{\tau}_P = \frac{1}{k_{+\mathrm{PA}}[\mathrm{A}]_0} \tag{1.23}$$

$$\overline{\tau}_{PA} = \frac{1}{k_{-\mathrm{PA}}} \tag{1.24}$$

and the corresponding equilibrium constant can be expressed as:

$$K_{\mathrm{PA}} = \frac{k_{+\mathrm{PA}}}{k_{-\mathrm{PA}}} = \frac{\overline{\tau}_{\mathrm{PA}}}{\overline{\tau}_{\mathrm{P}}[\mathrm{A}]_0} = \frac{\sum \tau_{\mathrm{PA}}}{\sum \tau_{\mathrm{P}}} \cdot \frac{1}{[\mathrm{A}]_0} \tag{1.25}$$

1.3 Pore structure and function

Two pores, αHL and outer membrane protein G (OmpG) were used in this work. Attempts were also made to use a further pore, the voltage gated anion channel (VDAC). Here follows a summary of the structural and functional properties of these pores, relevant to their engineering and use as platforms for stochastic sensing.

1.3.1 Alpha-haemolysin

αHL from *Staphylococcus aureus* is a pore-forming toxin that is secreted as a water-soluble 33.2 kDa monomer. This monomer binds to human platelets, erythrocytes, monocytes, lymphocytes, endothelial cells and rabbit erythrocytes [81]. Once membrane-bound, the monomer assembles into a 232.4 kDa heptameric homo-oligomer [82]. This heptameric species forms a transmembrane channel that actuates cytolysis through membrane permeabilisation [81].

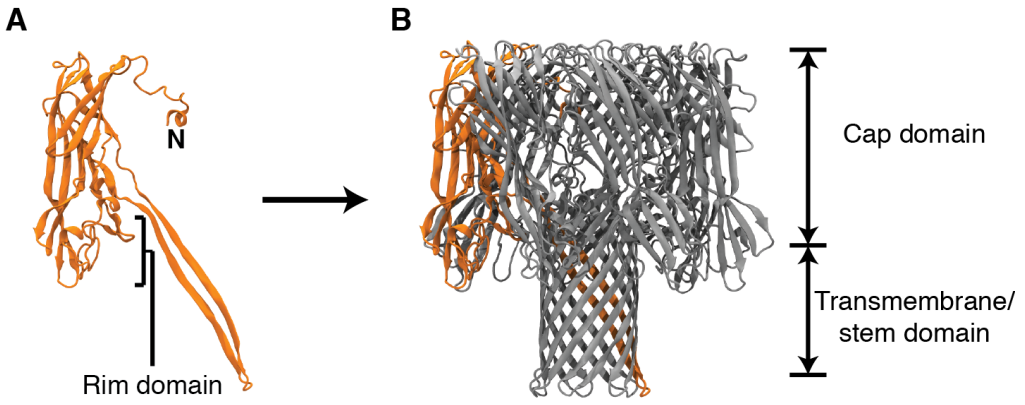


Figure 1.6: Crystal structure of the heptameric pore-forming toxin alpha-haemolysin (αHL). A) Single subunit of the heptameric pore showing the position of the rim domain and N-terminus (N). B) The heptameric pore showing position of the cap and transmembrane domains. Atomic coordinates were taken from PDB: 7AHL and visualised in PyMOL [86].

The crystal structure of heptameric αHL has been solved to 1.9 Å [83]. The heptameric pore is mushroom-shaped, with a wide extracellular 'cap' domain about 100 Å in diameter, and a stem domain that comprises a 14-strand β-barrel with each subunit donating a β-hairpin to the barrel (Fig 1.6). The cap domain comprises seven β-sandwiches and the amino-latch from each subunit. The amino latch was thought to be important for haemolytic activity [84], however, although the ability to form functional pores was diminished when the N-terminus was truncated up to 17 amino acids, it was still substantial. The N-terminus is therefore not critical for the pre-pore to pore transformation, but instead appears to prevent the premature oligomerisation of αHL monomers in solution [85].

The cap domain contains a great many inter-subunit interactions that are

responsible for the robustness of the assembled oligomer. Additionally, there are seven rim domains that do not form significant inter-subunit interactions but appear to interact with the membrane bilayer. A solvent-filled channel runs the length of the pore along the sevenfold axis of symmetry and is at its narrowest in the lumen adjacent to Glu111, Met113, and Lys147 side chains, where it is 14 Å in diameter [83].

A four-step mechanism for heptamer assembly and membrane insertion has been proposed [83]. Monomers first bind the membrane, then heptamerise into a 'pre-pore' complex that does not yet traverse the membrane [87]. The glycine-rich region of the pre-pore then translocates the membrane forming the transmembrane β-barrel [88].

When reconstituted into a planar lipid bilayer, αHL forms pores that have a high unitary conductance of $\sim$1 nS at +100 mV in 1 M KCl solution. The wild type channel exhibits a single open state with no gating events [66]. The heptameric pore is robust and can withstand comparatively extreme conditions such as denaturants [89], extreme pH [90] and high temperature [91].

1.3.2 Outer membrane protein G

Outer membrane protein G (OmpG) is a monomeric non-specific porin whose expression arises due to a rare chromosomal mutation called *cog-192* [92, 93]. This mutation is a large chromosomal deletion extending from the 3' end of *pspA* through to the 5' end of an open reading frame immediately upstream of *ompG*. Hence, as a result of the deletion, the expression of *ompG* is placed under the control of the *pspA* promoter. Although *ompG* DNA is present in

all common *E. coli* strains, its expression under laboratory conditions is only seen due to rare chromosomal mutations [92].

Proteoliposome swelling assays and substrate uptake experiments indicated OmpG to be a general diffusion porin with a larger channel diameter than the other *E. coli* porins such as OmpF and OmpC, and no substrate specificity [92, 93]. OmpG additionally differs from classical porins such as OmpF and OmpC in that it does not appear to exist in stable oligomers, as demonstrated by voltage-induced closure, pH-induced gating, block by Gd^{3+}, and biochemical studies such as chemical cross-linking [92, 94, 95]. Reconstituted into planar lipid bilayers, OmpG forms channels of unitary conductance 0.81 nS in 1 *M* NaCl, pH 7.5 [95]. In common with other porins, OmpG exhibits heat modifiability [95]; native OmpG appears as a 28 kDa band in sodium dodecyl sulphate polyacrylamide electrophoresis (SDS-PAGE) when not heated before running—after heating in SDS it runs as a 34 kDa species [95].

The structure of OmpG has been solved in three independent studies, two by X-ray crystallography [96, 97], one by solution nuclear magnetic resonance (NMR) (Fig 1.7). Crystal packing reinforces the view that OmpG is monomeric [96, 97]. Analysis of the structures confirmed earlier predictions of a 14-strand β-barrel topology [95, 98]. The pore architecture consists of 14 anti-parallel β-sheets arranged in a β-barrel, with short periplasmic turns and longer extracellular loops [96, 97, 99]. However, these loops are not long enough to fold back into the lumen of the pore, as is found for example in loop 3 of the general porins such as OmpF and OmpC. The central channel forms a pore of large diameter $\sim$13 Å.

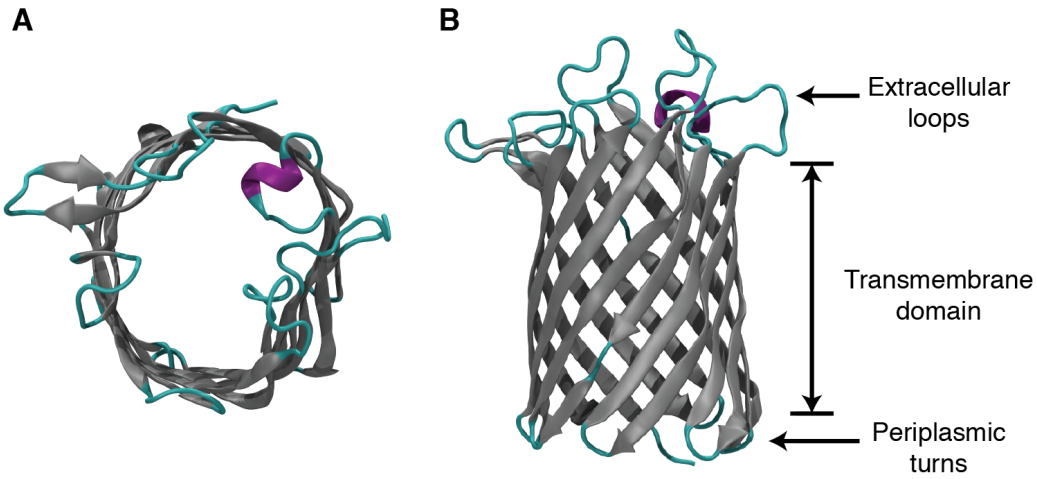


Figure 1.7: Crystal structure of Outer Membrane Porin G (OmpG) in its open conformation at pH 7.5. A) Top view of the pore looking down from the extracellular side. B) Side view showing position of extracellular and periplasmic loops. Atomic coordinates taken from PDB: 2IWV and visualised in PyMOL [86].

Yildiz *et al.* solved the crystal structure of OmpG at pH 7.5 and 5.6, revealing two distinct 'open' and 'closed' conformations [96]. At pH 7.5 the extracellular loops mostly extend into the medium, only minimally constricting the lumen of the pore. However, at pH 6.5, there is a large conformational change centred around loop 6, which folds across and into the barrel due to 'unzipping' of several hydrogen bonds in the adjacent β-sheets. Two histidine residues, His231 and His261 are believed to be involved in this pH-dependent mechanism. These structural observations are consistent with pH-dependent gating observed in single-channel recordings [94]. Recent single-molecule force spectroscopy studies using atomic force microscopy (AFM) provide further support for this mechanism of pH-dependent gating [100–102].

As both αHL and OmpG possess transmembrane domains consisting of 14-strand β-barrels of similar diameter, OmpG has potential as an alternative to αHL in stochastic sensing due to a number of advantageous properties.

One particular advantage lies in the monomeric nature of OmpG, which could allow the introduction of multiple modifications at different sites, or single-point modifications, without the need to form and purify hetero-oligomers with defined stoichiometry. One disadvantage for sensing purposes is that at neutral pH OmpG exhibits fast, spontaneous gating events attributable to the motion of loop 6 [95, 103]. To enhance the pore's potential for sensing applications, Chen *et al.* optimised hydrogen bonding near loop 6 by a point deletion and engineered a disulphide bond by cysteine introduction [103]. These modifications stabilised the region of the pore responsible for spontaneous gating and reduced pore noise as a result. Further, it was shown that the pore could bind the molecular adapter heptakis-(6-deoxy-6-amino)-β-cyclodextrin, allowing the stochastic detection of adenosine diphosphate (ADP).

1.3.3 The Voltage-Dependent Anion Channel

The voltage-dependent anion channel (VDAC) was first discovered in 1976 [104]. Mammalian VDACs exist as three isoforms, VDAC-1, VDAC-2 and VDAC-3, which are predicted to be structurally similar but have different expression levels and tissue distributions [105, 106]. The channel has been extensively studied by biochemical means, although several aspects remain controversial, for example whether VDAC is found in plasma membranes and whether it plays a role in apoptosis. [107] More recently the 3D structure of VDAC-1 has been solved in three independent studies, one by X-ray crystallography [108], one by solution NMR [109], and one by a combined NMR and

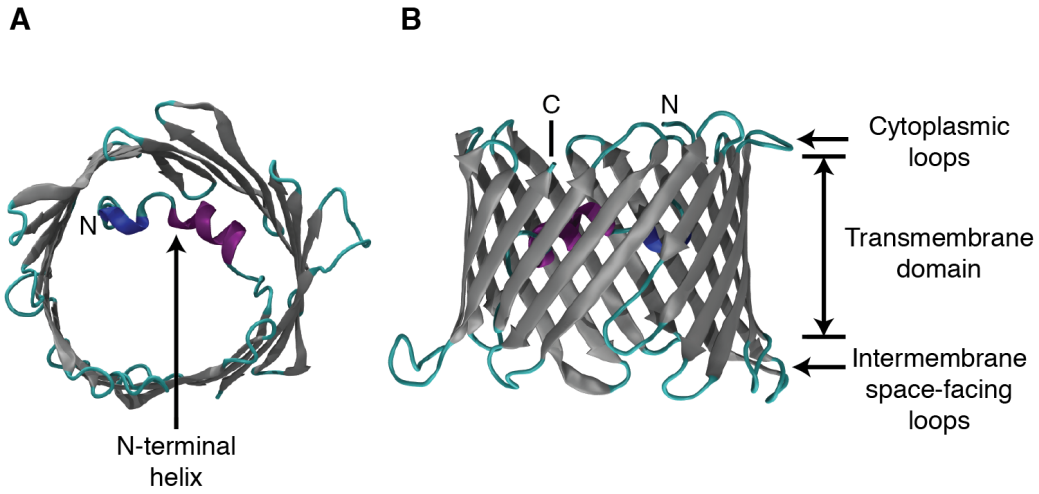


Figure 1.8: Crystal structure of the Voltage Dependent Anion Channel (VDAC). A) Top view of the pore looking down from the intermembrane space-facing side and showing the position of the N-terminal helix. The N-terminus is labeled 'N'. B) Side view of the pore showing the position of the cytoplasmic and intermembrane space-facing loops. The C-terminus is labeled 'C'. Atomic coordinates taken from PDB: 3EMN and visualised in PyMOL [86].

X-ray study [110]. As these structures are of detergent-refolded VDAC, there is some disagreement over whether the structures represent the 'native' conformation that corresponds to the channel-forming species present in lipid bilayers [111] [112].

VDAC is the most abundant protein in the mitochondrial outer membrane and is responsible for the passive transport of metabolites and ions between the mitochondrial inter-membrane space and the cytosol [113, 114]. There have been numerous reports implicating VDAC in apoptosis, with the protein proposed to have a large number of interaction partners such as hexokinase and the Bcl-2 family of apoptosis regulator proteins [115–123]. However, the details of these mechanisms remain disputed [124].

VDAC forms large channels with unitary conductance ∼4 nS at neutral pH in 1 M KCl with applied potential <30 mV, dependent on the source of

the protein [125]. At around ±30 mV and higher, the channel switches to a lower conductance state with altered selectivity [125, 126].

The 3D structures show that VDAC comprises a 19-strand β-barrel formed of anti-parallel β-sheets (Fig 1.8), with the first and last strands parallel [108–110]. Notably, all known prokaryotic β-barrels possess an even number of strands [127, 128]. The N-terminal segment of VDAC is of particular interest as a number of important roles have been suggested for it, such as voltage sensing and protein-protein interactions with anti-apoptotic proteins [115, 129]. The N-terminus was not fully assigned in the NMR structure, whereas it was fully resolved in the high-resolution crystal structure [108]. This shows the N-terminal segment nestling against the interior of the pore, with hydrogen bonds to pore-lining residues. The authors suggest a voltage-sensing mechanism in which the N-terminus undergoes a conformational change, pivoting around a flexible linker sequence and placing the helix in a more occlusive position, restricting metabolite flux [108].

Due to the undesirable voltage-dependent behaviour, VDAC has not been exploited for stochastic sensing, although its large pore diameter (25 Å) and monomeric state could be highly useful. One study has reported the translocation of dsDNA across a planar lipid bilayer containing VDAC, but measured indirectly and no further work has since been reported [49].

1.4 Protein kinases

Protein kinases form an extensive family of enzymes that transfer the γ-phosphate from a nucleoside triphosphate (normally adenosine triphosphate

(ATP)) to an amino acid in a target protein capable of accepting a phosphate group, in a process called phosphorylation. Phosphorylation occurs in all living organisms. Serine/threonine and tyrosine kinases form a large superfamily of enzymes that catalyse the transfer of phosphate to the free hydroxyl groups of serine, threonine or tyrosine [130–132]. Within this superfamily, most act on both serine and threonine, or on tyrosine, while there are some that will act on all three amino acids. There are also other protein kinases that act on other amino acids and are distinct from the serine/threonine and tyrosine kinase superfamily, such as histidine kinases.

Protein kinases are one of the largest gene families in eukaryotes, with the total number of human protein kinases believed to be about 518, which constitutes 1.7% of all human genes [131]. Most signal transduction processes are mediated by a phosphotransfer step or cascade. Further, by substrate modification, kinases regulate the cytoskeleton and cellular processes such as metabolism, transcription, the cell cycle, apoptosis and cellular differentiation. Roles are also played in intercellular communication, homeostasis, and the nervous and immune systems. Reversible phosphorylation for enzyme control was first discovered by Krebs and Fischer in 1956 [133].

The power of phosphorylation as a regulatory mechanism stems from two key features. First, a small signal, i.e. the addition of a phosphate group to a single molecule of enzyme, can produce a large response if this enzyme goes on to catalytically process a great many molecules of substrate, with these substrates also potentially activating further responses in a cascade. Second, phosphorylation is reversible. Phosphorylated proteins may be dephosphorylated by protein phosphatases [134]. These two opposing and reversible

processes hence form a powerful regulatory on/off switching mechanism. As the phosphorylation and dephosphorylation reactions are catalysed by distinct groups of enzymes, further regulation is possible through independent control by distinct stimuli [134].

The effect of phosphorylation on substrate proteins manifests in several ways. First, the phosphate group may alter the activity of the protein. Addition of a phosphate group introduces two negative charges and the capability to form multiple hydrogen bonds. Change in activity may arise from the bulk properties of the phosphate group or by conformational change induced by altered interactions. Second, the phosphate group may produce a new recognition site for another protein, for example the SH2 domain [135].

As phosphorylation is critical to so many crucial cellular processes, it is vital that kinases are tightly regulated. Spatially or temporally inappropriate kinase activity can be extremely dangerous and lead to cell transformation and cancer. For example, the first oncogene to be discovered, *v-Src*, encodes a constitutively active tyrosine kinase [136].

All protein kinases have similar catalytic function, mediated by a high degree of structural similarity in their active conformations. However, in their inactive conformations, there are no catalytic restraints on conformation. A wide range of regulatory mechanisms have therefore evolved that prevent the adoption of a catalytically active conformation [130].

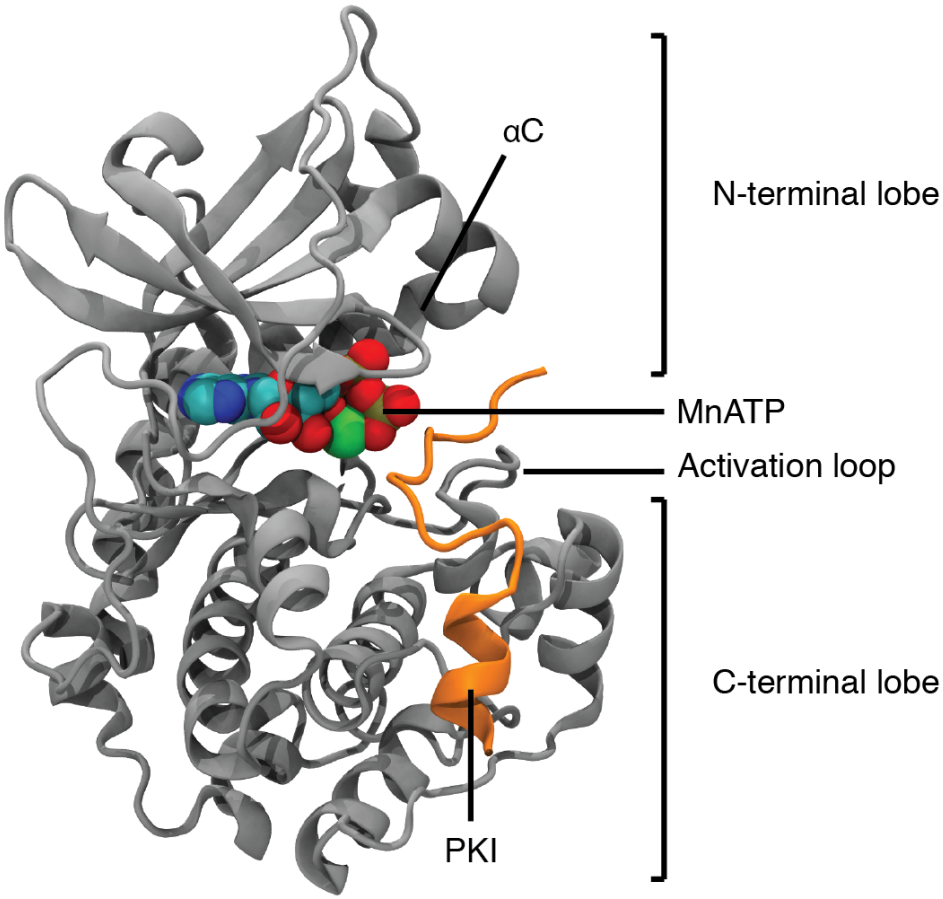


Figure 1.9: Crystal structure of the cyclic AMP-dependent protein kinase/protein kinase A showing the positions of key structural features. The kinase molecule is coloured grey. MnATP is shown with a space-filling representation. The protein kinase A inhibitory peptide (PKI) is coloured orange. Atomic coordinates from PDB: 1ATP and visualised in PyMOL [86].

1.4.1 Structure of the active conformation

The first solved crystal structure of a eukaryotic protein kinase domain was of the catalytic subunit of cyclic AMP-dependent protein kinase (Fig 1.9), also known as protein kinase A (PKA) [137–139]. Protein kinase folds are highly conserved and consist of two lobes or sub-domains. The N-terminal lobe is smaller and comprises a five-stranded β-sheet and a key α-helix termed αC. The large C-terminal lobe is mostly α-helical. A deep cleft exists between the two lobes, in which MgATP binds. The phosphate binding loop (P loop) is a flexible, glycine-rich loop with a conserved motif (GXGXϕG, where ϕ is usually tyrosine or phenylalanine) that sits above MgATP. The phosphates of MgATP are coordinated by backbone interactions with the P loop, with the aromatic residues forming a cap over the site of phosphotransfer.

The peptide substrate is bound across the front of the MgATP-binding pocket in an extended conformation, placing the serine/threonine/tyrosine close to the γ-phosphate of MgATP. The substrate is supported by a 20–30 residue loop called the activation loop. In most kinases this loop is phosphorylated in the catalytically active conformation. Phosphorylation of the activation loop stabilises it in a conformation conducive for substrate binding.

1.4.2 Assay of kinase function

Measurement of kinase activity is hugely important, largely due to the intensity of drug discovery efforts aimed at finding modulators of kinases through high throughput screening. Consequently, a wide range of assays have been developed [140]. Those in widespread use mostly fall into two

groups, with one using radiometric detection and the others using one of several fluorescence-based measurements such as Förster resonance energy transfer (FRET) or fluorescence polarisation (FP).

Radiometric methods

Radiometric kinase assays are still considered the 'gold standard' to which non-radiometric methods are compared. An *in vitro* phosphorylation reaction is spiked with radiolabeled [γ-^{32}P]ATP. The resulting phosphorylated peptides or proteins are captured on phosphocellulose paper, which is immersed in phosphoric acid to convert the acceptor peptides and proteins to positively charged species that are quantitatively immobilised on the negatively charged paper [141, 142]. After stringent washing in phosphoric acid to remove excess negatively charged ATP, the amount of radiolabeled product bound to the paper can be measured, for example by scintillation counting [143].

It is possible to assay the phosphorylation of peptides and proteins at serine, threonine or tyrosine with this method [143]. A certain few kinase substrates that may be chosen for assay will not be adequately positively charged at pH 1.8 and will hence not bind quantitatively to the phosphocellulose paper. This may be rectified by the addition of basic residues to the N- or C-terminus of the peptide/protein.

Radiometric assays have several advantages over fluorescence-based alternatives. Firstly, almost any substrate may be used, including long proteins, whereas many fluorescence-based methods have limitations on the nature and size of the substrate (see below). The radiometric assay provides a direct

measurement of enzyme activity, whereas other methods typically require a secondary detection agent such as an anti-phosphoserine antibody. Finally, the sensitivity of the radiometric assay is much higher than fluorescence-based assays, which measure a typically small change in signal upon phosphorylation.

The major limitations of the radiometric method are its low throughput and use of radioisotopes, which mean it is rarely used for high throughput screening. Techniques are available for high throughput radiometric assays, either through the use of scintillation beads that capture phosphorylated peptides (using the biotin-streptavidin interaction, for example) that can be read with an imaging reader [144], or via Perkin Elmer's FlashPlate product. Nonetheless, there is a strong preference for non-radioactive assays where high throughput screening is used for drug discovery.

Fluorescence-based methods

These methods are favoured for high throughput screening as they are easily automated, homogenous and easy to use. There are a large number of commercial systems available, and so this section will describe only the most common, with a focus on the broad concepts utilised, rather than the specific details of individual products.

Fluorescence intensity assays are commonly used for measuring enzyme activity. However, for the detection of kinase activity, a coupling reaction must be used, for example the detection of ADP by using linked reactions. An example of this is the ADP Hunter product from DiscoveRx [145].

QTL Biosystems market a high throughput assay for kinase and phos-

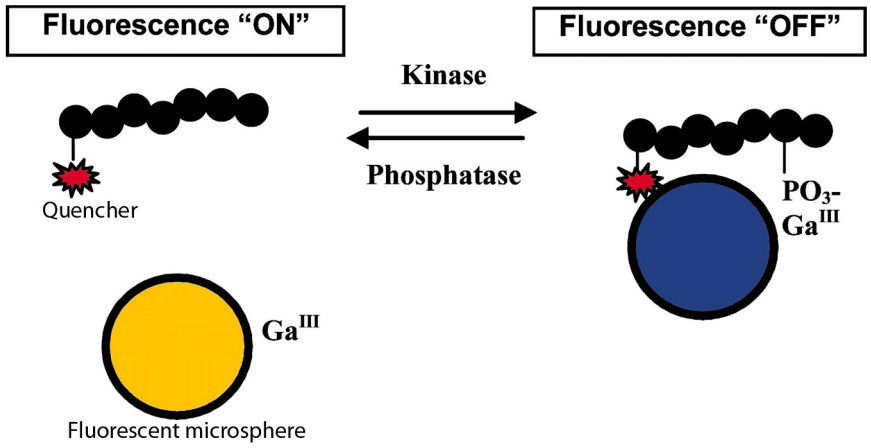


Figure 1.10: General scheme for the detection of kinase and phosphatase activities using the QTL assay. Phosphorylated quencher-conjugated peptide substrate molecules are brought to the surface of fluorescent polyelectrolyte functionalised microspheres via gallium ion-mediated binding to phosphate groups. The resulting quench of fluorescence is proportional to the extent of phosphorylation or dephosphorylation. Figure adapted from Rininsland *et al.*, [146].

phatase activity that utilises metal-ion mediated polymer super-quenching [146]. The assay sensor comprises microspheres that are co-functionalised with highly fluorescent-conjugated polyelectrolytes and gallium ions. The trivalent gallium ions selectively bind phosphate groups on peptide/protein substrates. The peptide/protein substrate is conjugated to a quencher, so that the binding of phosphorylated substrate to the microspheres will quench the fluorescence signal in a quantitative manner.

Phosphatase activity can be measured in this way by the use of phosphorylated substrate, where the activity of the phosphatase will lead to a decrease in quenching, as the dephosphorylated substrate will remain disassociated from the microspheres. Both kinase and phosphatase activity are hence measured via the modulation of the fluorescence signal of the microspheres in a simple, homogenous assay (Figure 1.10). The assay is significantly more

sensitive than many other fluorescent kinase assays as the fluorescence of the polyelectrolyte can be quenched by extremely low concentrations of quencher, but this does require quencher-modified substrates [147].

These methods aside, the three most common fluorescence techniques for kinase assay are FP, FRET and time-resolved fluorescence (TRF), which are reviewed below.

FP methods FP methods measure the proportion of polarisation retained in the emitted light when a molecule is excited by polarised light. A large molecule (>100 kDa) will not rotate significantly on the timescale of the lifetime of the excited state, so the light will be emitted in the same plane of polarisation as the exciting light. In contrast, small molecules (<10 kDa) will have rotated to some extent before emission and so will emit in a different plane to the exciting light; polarisation in the plane of excitation will therefore decrease.

This effect is utilised in kinase assays by the selective capture of fluorescent-labeled phosphopeptides by phospho-antibodies or by immobilised metal affinity using metal-chelated beads. The binding of the antibodies/beads to the phosphorylated substrate peptides slows the rotation of the molecules, increasing the FP signal [148]. The main disadvantages of the FP approach are the limited sensitivity and signal-to-noise ratio, and the susceptibility to fluorescence interference from some reaction components and, if screening compounds, from fluorescent or coloured compounds [149]. This interference can lead to false hits or mask real hits during the screening process [150].

FRET methods FRET is the process where radiationless energy-transfer occurs from an excited donor fluorophore to an acceptor fluorophore. FRET only occurs when the donor and acceptor are within 1–10 nm of each other and in a favourable orientation. This effect has been utilised in several different ways for kinase assays. One example is where a donor fluorophore is conjugated to a phospho-antibody that specifically binds to a phosphorylated substrate that is labelled with an acceptor fluorophore. Upon antibody binding, an increase in FRET will be detected. Another example uses a FRET labelled substrate that is subjected to a phosphorylation reaction. A site specific protease is then added that will only cleave unphosphorylated substrate and so disrupt FRET. The donor/acceptor fluorescence ratio is hence measured to determine the level of enzyme activity [151]. The disadvantages of these assays include the need for carefully designed labeled substrates, which may be expensive. Additionally, the presence of fluorophore labels may alter kinase kinetics. Free fluorophores may also interfere with the energy transfer process.

TRF methods Time-resolved fluorescence methods use fluorophores with long fluorescence decay times, and measure the change in fluorescence over time after excitation by a flash of light. Organic fluorophores typically have short emission lifetimes, so lanthanide ions such as europium, samarium and terbium are commonly used.

An example of a TRF kinase assay is the DELFIA assay from Perkin Elmer. Kinase reactions are performed with a biotinylated substrate, which is subsequently immobilised on a streptavidin-coated capture plate. After

multiple washing steps, a europium-coupled antibody is added to detect phosphorylated substrates. An 'enhancement' solution is then added, which dissociates the europium ions that subsequently form highly fluorescent chelates with components in the solution. This assay is sensitive and avoids the issue of fluorescence interference. However, the immobilisation and multiple washing/labelling steps are less efficient than fully homogeneous 'mix-and-read' methods.

To reduce the cost of reagents and increase the speed of the assay, TRF has also been combined with FRET to create TR-FRET methods. One example is the LanthaScreen assay from Invitrogen. Lanthascreen utilises a fluorescein-labelled substrate. After quenching of the kinase reaction, terbium-labeled anti-phospho antibody is added. Excitation of the terbium brings about FRET to the fluorescein acceptor. The ratio of acceptor emission to donor emission is measured, with the possibility of time-resolved measurement due to the long decay time of the terbium. This method has the advantage of being a simple 'mix-and-read' method. However, due to the short wavelength of fluorescein emission, it is susceptible to interference from coloured compounds that may be screened. Specific antibodies are also required, and may be expensive.

Other methods Kinase activity can also be measured via detection of ATP depletion. For example, the Kinase-Glo assay from Promega uses firefly luciferase to convert luciferin to oxylucferin in an ATP consuming reaction that produces yellow-green light. Phosphorylation by kinases reduces the amount of ATP present, thus decreasing the number of photons produced

by the luciferase reaction. A significant advantage of this assay is that it is universal for any combination of kinase and substrate and does not require antibodies. Fluorescence interference is also avoided. However, when screening inhibitors, some partial inhibitors for the luciferase may complicate interpretation of the observed kinase inhibition.

Enzyme-linked immunosorbent assays (ELISA) were widely used before fluorescent methods became more popular [152–154]. These assays can be very sensitive and are not susceptible to fluorescent interference from screened compounds. However, the requirement of immobilisation and washing steps in these methods limits their use in high throughput applications. This method also requires specific anti-phospho antibodies, which may be difficult to develop.

Finally, mobility shift assays exist that are based on the different electrophoretic shifts of the phosphorylated and unphosphorylated states of a peptide substrate. For high throughput, microfluidic chip-based assays have been developed. As the assay is dependent on the charge difference between the substrate and product, the main limitation is the selection and design of the substrate sequence and size. The use of larger substrates is hence difficult to achieve.

1.4.3 Kinase inhibition

Perhaps the most obvious way to inhibit a kinase is by direct competition with its substrate [155]. Indeed, for PKA, the peptide inhibitor PKI [137] mimics a substrate for the kinase by binding the active conformation di-

rectly and without affecting the binding of MgATP. In many kinases, motifs that mimic substrates are present as C-terminal extensions and so inhibit the kinase in an intramolecular fashion. These sequences are often termed 'pseudosubstrate' sequences [130].

The majority of small molecule inhibitors target the ATP binding site and recognise the active conformation of kinases. This is probably due to many inhibitors having been discovered by enzymatic assays and because many inhibitors have been synthesised to mimic ATP. Along with G protein-coupled receptors (GPCRs), kinases are among the most intensively pursued drug targets, the majority of kinase targets being investigated for cancer treatment [156].

This popularity is due to several factors. First, almost all signal transduction pathways involve a phosphotransfer cascade, so kinase inhibition is highly likely to effect a physiological response. Second, kinase inhibition is usually tolerated in healthy cells, enabling the selective killing of tumour cells. Third, the highly conserved MgATP binding site allows the design of tightly binding small molecules. However, this high level of conservation is also a significant challenge for kinase inhibitor development as achieving selectivity between different kinase families can be very difficult [156]. Nonetheless, there have been high-profile successes such as imatinib (Gleevec) which is used to treat chronic myelogenous leukaemia (CML) [157].

In CML, a chromosomal translocation known as the Philadelphia chromosome leads to the formation of an abnormal fusion gene between part of the breakpoint cluster region (*bcr*) and the *abl* gene [158]. The BCR-ABL1 protein arising from this fusion is a constitutively active tyrosine kinase whose

action is the cause of CML. Imatinib selectively inhibits the BCR-ABL1 protein and exhibits up to 80% response rates in CML cases [159].

Challenges remaining for drug discovery aimed at kinase targets include drug resistance, lack of efficacy and validation difficulties for particular disease settings.

1.5 Overview

This thesis presents two different approaches in applying stochastic sensing to the detection of protein analytes. In one, rationally engineered stochastic sensors for the Pim family of protein kinases were developed using genetically encoded sensor elements. An enhanced approach was also developed that used cleavage by a specific protease to generate a pore with the peptide sensing element attached only by a single terminus. Preliminary experiments using these pores indicated that phosphorylation of the sensor element can be observed, leading to the possibility of further single molecule studies of phosphorylation and dephosphorylation, for example by the introduction of a protein phosphatase to create a catalytic cycle.

In the other, a methodology is developed towards tailored stochastic sensors of protein analytes using OmpG as a scaffold, both for selection of genetically encoded sensor elements by bacterial display, and subsequently for bio-sensing by single channel measurements of the selected pores. The methodology was designed so that prior knowledge of ligands or rationally designed protein engineering steps would not be required, thereby expanding the scope of what can be targeted by stochastic sensing.

Bibliography

[1] Luckey, M. (2008) *Membrane Structural Biology: With Biochemical and Biophysical Foundations*. (Cambridge University Press).

[2] Bayley, H. and Cremer, P. S. (2001) Stochastic sensors inspired by biology. *Nature 413*, 226–30.

[3] Neher, E. and Steinbach, J. H. (1978) Local anaesthetics transiently block currents through single acetylcholine-receptor channels. *J Physiol 277*, 153–76.

[4] Coulter, W. H. (1953) Means for counting particles suspended in a fluid (US Patent 2656508).

[5] Movileanu, L., Howorka, S., Braha, O., and Bayley, H. (2000) Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nat Biotechnol 18*, 1091–1095.

[6] Howorka, S. and Siwy, Z. (2009) Nanopore analytics: sensing of single molecules. *Chem Soc Rev 38*, 2360–84.

[7] Dekker, C. (2007) Solid-state nanopores. *Nat Nanotechnol 2*, 209–15.

[8] Rincon-Restrepo, M., Mikhailova, E., Bayley, H., and Maglia, G. (2011) Controlled Translocation of Individual DNA Molecules through Protein Nanopores with Engineered Molecular Brakes. *Nano Lett.*

[9] Howorka, S., Nam, J., Bayley, H., and Kahne, D. (2004) Stochastic detection of monovalent and bivalent protein-ligand interactions. *Angew Chem Int Ed Engl 43*, 842–846.

[10] Cheley, S., Xie, H., and Bayley, H. (2006) A genetically encoded pore for the stochastic detection of a protein kinase. *Chembiochem 7*, 1923–1927.

[11] Xie, H., Braha, O., Gu, L.-Q., Cheley, S., and Bayley, H. (2005) Single-molecule observation of the catalytic subunit of cAMP-dependent protein kinase binding to an inhibitor peptide. *Chem Biol 12*, 109–120.

[12] Nestorovich, E. M., Danelon, C., Winterhalter, M., and Bezrukov, S. M. (2002) Designed to penetrate: time-resolved interaction of single antibiotic molecules with bacterial pores. *Proc Natl Acad Sci U S A 99*, 9789–9794.

[13] Gu, L. Q., Braha, O., Conlan, S., Cheley, S., and Bayley, H. (1999) Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature 398*, 686–690.

[14] Uram, J. D., Ke, K., and Mayer, M. (2008) Noise and bandwidth of current recordings from submicrometer pores and nanopores. *ACS Nano 2*, 857–872.

[15] Kasianowicz, J., Walker, B., Krishnasastry, M., and Bayley, H. (1994) *Genetically-engineered pores as metal-ion biosensors*, Biomolecular materials by design Book Series: Materials Research Society symposium proceedings. (Materials Research Society) Vol. 330, pp. 217–223.

[16] Walker, B., Kasianowicz, J., Krishnasastry, M., and Bayley, H. (1994) Pore-forming protein with a metal-actuated switch. *Protein Eng 7*, 655–662.

[17] Braha, O., Walker, B., Cheley, S., Kasianowicz, J. J., Song, L., Gouaux, J. E., and Bayley, H. (1997) Designed protein pores as components for biosensors. *Chem Biol 4*, 497–505.

[18] Braha, O., Gu, L. Q., Zhou, L., Lu, X., Cheley, S., and Bayley, H. (2000) Simultaneous stochastic sensing of divalent metal ions. *Nat Biotechnol 18*, 1005–1007.

[19] Guan, X., Gu, L., Cheley, S., Braha, O., and Bayley, H. (2005) Stochastic sensing of TNT with a genetically engineered pore. *Chembiochem 6*, 1875–1881.

[20] Cheley, S., Gu, L., and Bayley, H. (2002) Stochastic sensing of nanomolar inositol 1,4,5-trisphosphate with an engineered pore. *Chem Biol 9*, 829–838.

[21] Astier, Y., Braha, O., and Bayley, H. (2006) Toward single molecule DNA sequencing: direct identification of ribonucleoside and deoxyribonucleoside 5'-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *J Am Chem Soc 128*, 1705–1710.

[22] Clarke, J., Wu, H.-C., Jayasinghe, L., Patel, A., Reid, S., and Bayley, H. (2009) Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol 4*, 265–270.

[23] Wu, H.-C., Astier, Y., Maglia, G., Mikhailova, E., and Bayley, H. (2007) Protein nanopores with covalently attached molecular adapters. *J Am Chem Soc 129*, 16142–16148.

[24] Banerjee, A., Mikhailova, E., Cheley, S., Gu, L.-Q., Montoya, M., Nagaoka, Y., Gouaux, E., and Bayley, H. (2010) Molecular bases of cyclodextrin adapter interactions with engineered protein nanopores. *Proc Natl Acad Sci U S A 107*, 8165–8170.

[25] Braha, O., Webb, J., Gu, L.-Q., Kim, K., and Bayley, H. (2005) Carriers versus adapters in stochastic sensing. *Chemphyschem 6*, 889–892.

[26] Sanchez-Quesada, J., Ghadiri, M., Bayley, H., and Braha, O. (2000) Cyclic peptides as molecular adapters for a pore-forming protein. *J Am Chem Soc 122*, 11757–11766.

[27] Chen, M., Li, Q.-H., and Bayley, H. (2008) Orientation of the monomeric porin OmpG in planar lipid bilayers. *Chembiochem 9*, 3029–36.

[28] Hammerstein, A. F., Shin, S.-H., and Bayley, H. (2010) Single-molecule kinetics of two-step divalent cation chelation. *Angew Chem Int Ed Engl 49*, 5085–90.

[29] Wirth, A., Jung, M., Bies, C., Frien, M., Tyedmers, J., Zimmermann, R., and Wagner, R. (2003) Sec61p complex is a dynamic precursor activated channel. *Mol Cell 12*, 261–268.

[30] Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S. H., and von Heijne, G. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature 433*, 377–381.

[31] Muro, C., Grigoriev, S. M., Pietkiewicz, D., Kinnally, K. W., and Campo, M. L. (2003) Comparison of the TIM and TOM channel activities of the mitochondrial protein import complexes. *Biophys J 84*, 2981–2989.

[32] Rehling, P., Model, K., Brandner, K., Kovermann, P., Sickmann, A., Meyer, H. E., Kühlbrandt, W., Wagner, R., Truscott, K. N., and Pfanner, N. (2003) Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. *Science 299*, 1747–1751.

[33] Hinnah, S. C., Wagner, R., Sveshnikova, N., Harrer, R., and Soll, J. (2002) The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides. *Biophys J 83*, 899–911.

[34] Stewart, M. (2007) Molecular mechanism of the nuclear protein import cycle. *Nat Rev Mol Cell Biol 8*, 195–208.

[35] Krantz, B. A., Melnyk, R. A., Zhang, S., Juris, S. J., Lacy, D. B., Wu, Z., Finkelstein, A., and Collier, R. J. (2005) A phenylalanine clamp catalyzes protein translocation through the anthrax toxin pore. *Science 309*, 777–781.

[36] Dreiseikelmann, B. (1994) Translocation of DNA across bacterial membranes. *Microbiol Rev 58*, 293–316.

[37] Pfanner, N. and Truscott, K. N. (2002) Powering mitochondrial protein import. *Nat Struct Biol 9*, 234–236.

[38] Krasilnikov, O. V., Sabirov, R. Z., Ternovsky, V. I., Merzliak, P. G., and Muratkhodjaev, J. N. (1992) A simple method for the determination of the pore radius of ion channels in planar lipid bilayer membranes. *FEMS Microbiol Immunol 5*, 93–100.

[39] Bezrukov, S. M., Vodyanoy, I., and Parsegian, V. A. (1994) Counting polymers moving through a single ion channel. *Nature 370*, 279–281.

[40] Movileanu, L., Cheley, S., and Bayley, H. (2003) Partitioning of individual flexible polymers into a nanoscopic protein pore. *Biophys J 85*, 897–910.

[41] Robertson, J. W. F., Rodrigues, C. G., Stanford, V. M., Rubinson, K. A., Krasilnikov, O. V., and Kasianowicz, J. J. (2007) Single-molecule mass spectrometry in solution using a solitary nanopore. *Proc Natl Acad Sci U S A 104*, 8207–8211.

[42] Davies, J. (1994) Inactivation of antibiotics and the dissemination of resistance genes. *Science 264*, 375–382.

[43] Koonin, E. V., Makarova, K. S., and Aravind, L. (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol 55*, 709–742.

[44] Branton, D., Deamer, D. W., Marziali, A., Bayley, H., Benner, S. A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X., Jovanovich, S. B., Krstic, P. S., Lindsay, S., Ling, X. S., Mastrangelo, C. H., Meller, A., Oliver, J. S., Pershin, Y. V., Ramsey, J. M., Riehn, R., Soni, G. V., Tabard-Cossa, V., Wanunu, M., Wiggin, M., and Schloss, J. A. (2008) The potential and challenges of nanopore sequencing. *Nat Biotechnol 26*, 1146–1153.

[45] Kasianowicz, J. J., Brandin, E., Branton, D., and Deamer, D. W. (1996) Characterization of individual polynucleotide molecules using a membrane channel. *Proc Natl Acad Sci U S A 93*, 13770–13773.

[46] Akeson, M., Branton, D., Kasianowicz, J. J., Brandin, E., and Deamer, D. W. (1999) Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys J 77*, 3227–3233.

[47] Meller, A., Nivon, L., Brandin, E., Golovchenko, J., and Branton, D. (2000) Rapid nanopore discrimination between single polynucleotide molecules. *Proc Natl Acad Sci U S A 97*, 1079–1084.

[48] Szabò, I., Bàthori, G., Tombola, F., Brini, M., Coppola, A., and Zoratti, M. (1997) DNA translocation across planar bilayers containing Bacillus subtilis ion channels. *J Biol Chem 272*, 25275–25282.

[49] Szabò, I., Bàthori, G., Tombola, F., Coppola, A., Schmehl, I., Brini, M., Ghazi, A., De Pinto, V., and Zoratti, M. (1998) Double-stranded DNA can be translocated across a planar membrane containing purified mitochondrial porin. *FASEB J 12*, 495–502.

[50] Butler, T. Z., Pavlenok, M., Derrington, I. M., Niederweis, M., and Gundlach, J. H. (2008) Single-molecule DNA detection with an engineered MspA protein nanopore. *Proc Natl Acad Sci U S A 105*, 20647–20652.

[51] Derrington, I. M., Butler, T. Z., Collins, M. D., Manrao, E., Pavlenok, M., Niederweis, M., and Gundlach, J. H. (2010) Nanopore DNA sequencing with MspA. *Proc Natl Acad Sci U S A 107*, 16060–16065.

[52] Stoddart, D., Heron, A. J., Mikhailova, E., Maglia, G., and Bayley, H. (2009) Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc Natl Acad Sci U S A 106*, 7702–7707.

[53] Purnell, R. F. and Schmidt, J. J. (2009) Discrimination of single base substitutions in a DNA strand immobilized in a biological nanopore. *ACS Nano 3*, 2533–2538.

[54] Stoddart, D., Maglia, G., Mikhailova, E., Heron, A. J., and Bayley, H. (2010) Multiple base-recognition sites in a biological nanopore: two heads are better than one. *Angew Chem Int Ed Engl 49*, 556–559.

[55] Stoddart, D., Heron, A. J., Klingelhoefer, J., Mikhailova, E., Maglia, G., and Bayley, H. (2010) Nucleobase recognition in ssDNA at the central constriction of the alpha-hemolysin pore. *Nano Lett 10*, 3633–3637.

[56] Stoddart, D., Maglia, G., Mikhailova, E., Heron, A. J., and Bayley, H. (2010) Multiple base-recognition sites in a biological nanopore: two heads are better than one. *Angew Chem Int Ed Engl 49*, 556–9.

[57] Movileanu, L., Schmittschmitt, J. P., Scholtz, J. M., and Bayley, H. (2005) Interactions of peptides with a protein pore. *Biophys J 89*, 1030–1045.

[58] Wolfe, A. J., Mohammad, M. M., Cheley, S., Bayley, H., and Movileanu, L. (2007) Catalyzing the translocation of polypeptides through attractive interactions. *J Am Chem Soc 129*, 14034–14041.

[59] Goodrich, C. P., Kirmizialtin, S., Huyghues-Despointes, B. M., Zhu, A., Scholtz, J. M., Makarov, D. E., and Movileanu, L. (2007) Single-molecule electrophoresis of beta-hairpin peptides by electrical recordings and Langevin dynamics simulations. *J Phys Chem B 111*, 3332–3335.

[60] Oukhaled, G., Mathé, J., Biance, A.-L., Bacri, L., Betton, J.-M., Lairez, D., Pelta, J., and Auvray, L. (2007) Unfolding of proteins and long transient conformations detected by single nanopore recording. *Phys Rev Lett 98*, 158101.

[61] Cornell, B. A., Braach-Maksvytis, V. L., King, L. G., Osman, P. D., Raguse, B., Wieczorek, L., and Pace, R. J. (1997) A biosensor that uses ion-channel switches. *Nature 387*, 580–583.

[62] Hunter, C. A. and Anderson, H. L. (2009) What is cooperativity? *Angew Chem Int Ed Engl 48*, 7488–7499.

[63] Neher, E. and Sakmann, B. (1976) Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature 260*, 799–802.

[64] Tank, D. W., Miller, C., and Webb, W. W. (1982) Isolated-patch recording from liposomes containing functionally reconstituted chloride channels from Torpedo electroplax. *Proc Natl Acad Sci U S A 79*, 7749–53.

[65] MUELLER, P., RUDIN, D. O., TIEN, H. T., and WESCOTT, W. C. (1962) Reconstitution of cell membrane structure in vitro and its transformation into an excitable system. *Nature 194*, 979–80.

[66] Maglia, G., Heron, A. J., Stoddart, D., Japrung, D., and Bayley, H. (2010) Analysis of single nucleic acid molecules with protein nanopores. *Methods Enzymol 475*, 591–623.

[67] Montal, M. and Mueller, P. (1972) Formation of bimolecular membranes from lipid monolayers and a study of their electrical properties. *Proc Natl Acad Sci U S A 69*, 3561–6.

[68] Miller, C. (1986) *Ion Channel Reconstitution*. (Springer).

[69] Tsofina, L., Liberman, E., and Babakov, A. (1966) Production of Bimolecular Protein-Lipid Membranes in Aqueous Solution. *Nature 212*, 681–&.

[70] Holden, M. A., Needham, D., and Bayley, H. (2007) Functional bionetworks from nanoliter water droplets. *J Am Chem Soc 129*, 8650–5.

[71] Funakoshi, K., Suzuki, H., and Takeuchi, S. (2006) Lipid bilayer formation by contacting monolayers in a microfluidic device for membrane protein analysis. *Anal Chem 78*, 8169–74.

[72] Malmstadt, N., Nash, M. A., Purnell, R. F., and Schmidt, J. J. (2006) Automated formation of lipid-bilayer membranes in a microfluidic device. *Nano Lett 6*, 1961–5.

[73] Hwang, W. L., Chen, M., Cronin, B., Holden, M. A., and Bayley, H. (2008) Asymmetric droplet interface bilayers. *J Am Chem Soc 130*, 5878–9.

[74] Syeda, R., Holden, M. A., Hwang, W. L., and Bayley, H. (2008) Screening blockers against a potassium channel with a droplet interface bilayer array. *J Am Chem Soc 130*, 15543–8.

[75] Maglia, G., Heron, A. J., Hwang, W. L., Holden, M. A., Mikhailova, E., Li, Q., Cheley, S., and Bayley, H. (2009) Droplet networks with incorporated protein diodes show collective properties. *Nat Nanotechnol 4*, 437–40.

[76] Bayley, H., Cronin, B., Heron, A., Holden, M. A., Hwang, W. L., Syeda, R., Thompson, J., and Wallace, M. (2008) Droplet interface bilayers. *Mol Biosyst 4*, 1191–208.

[77] Shin, S.-H., Steffensen, M. B., Claridge, T. D. W., and Bayley, H. (2007) Formation of a chiral center and pyrimidal inversion at the single-molecule level. *Angew Chem Int Ed Engl 46*, 7412–6.

[78] Luchian, T., Shin, S.-H., and Bayley, H. (2003) Single-molecule covalent chemistry with spatially separated reactants. *Angew Chem Int Ed Engl 42*, 3766–71.

[79] Luchian, T., Shin, S.-H., and Bayley, H. (2003) Kinetics of a three-step reaction observed at the single-molecule level. *Angew Chem Int Ed Engl 42*, 1926–9.

[80] Shin, S.-H., Luchian, T., Cheley, S., Braha, O., and Bayley, H. (2002) Kinetics of a reversible covalent-bond-forming reaction observed at the single-molecule level. *Angew Chem Int Ed Engl 41*, 3707–9; 3523.

[81] Bhakdi, S. and Tranum-Jensen, J. (1991) Alpha-toxin of Staphylococcus aureus. *Microbiol Rev 55*, 733–51.

[82] Gouaux, J. E., Braha, O., Hobaugh, M. R., Song, L., Cheley, S., Shustak, C., and Bayley, H. (1994) Subunit stoichiometry of staphylococcal alpha-hemolysin in crystals and on membranes: a heptameric transmembrane pore. *Proc Natl Acad Sci U S A 91*, 12828–31.

[83] Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H., and Gouaux, J. E. (1996) Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science 274*, 1859–66.

[84] Walker, B., Krishnasastry, M., Zorn, L., and Bayley, H. (1992) Assembly of the oligomeric membrane pore formed by Staphylococcal alpha-hemolysin examined by truncation mutagenesis. *J Biol Chem 267*, 21782–6.

[85] Jayasinghe, L., Miles, G., and Bayley, H. (2006) Role of the amino latch of staphylococcal alpha-hemolysin in pore formation: a co-operative interaction between the N terminus and position 217. *J Biol Chem 281*, 2195–204.

[86] DeLano, W. (2002) *The PyMOL User's Manual.* (DeLano Scientific, Palo Alto, CA, USA).

[87] Walker, B., Braha, O., Cheley, S., and Bayley, H. (1995) An intermediate in the assembly of a pore-forming protein trapped with a genetically-engineered switch. *Chem Biol 2*, 99–105.

[88] Valeva, A., Palmer, M., Hilgert, K., Kehoe, M., and Bhakdi, S. (1995) Correct oligomerization is a prerequisite for insertion of the central molecular domain of staphylococcal alpha-toxin into the lipid bilayer. *Biochim Biophys Acta 1236*, 213–8.

[89] Japrung, D., Henricus, M., Li, Q., Maglia, G., and Bayley, H. (2010) Urea facilitates the translocation of single-stranded DNA and RNA through the alpha-hemolysin nanopore. *Biophys J 98*, 1856–63.

[90] Maglia, G., Henricus, M., Wyss, R., Li, Q., Cheley, S., and Bayley, H. (2009) DNA strands from denatured duplexes are translocated through engineered protein nanopores at alkaline pH. *Nano Lett 9*, 3831–6.

[91] Kang, X.-F., Gu, L.-Q., Cheley, S., and Bayley, H. (2005) Single protein pores containing molecular adapters at high temperatures. *Angew Chem Int Ed Engl 44*, 1495–9.

[92] Fajardo, D. A., Cheung, J., Ito, C., Sugawara, E., Nikaido, H., and Misra, R. (1998) Biochemistry and regulation of a novel Escherichia coli K-12 porin protein, OmpG, which produces unusually large channels. *J Bacteriol 180*, 4452–9.

[93] Misra, R. and Benson, S. (1989) A novel mutation, cog, which results in production of a new porin protein (OmpG) of Escherichia coli K-12. *J Bacteriol 171*, 4105.

[94] Conlan, S. and Bayley, H. (2003) Folding of a monomeric porin, OmpG, in detergent solution. *Biochemistry 42*, 9453–65.

[95] Conlan, S., Zhang, Y., Cheley, S., and Bayley, H. (2000) Biochemical and biophysical characterization of OmpG: A monomeric porin. *Biochemistry 39*, 11845–54.

[96] Yildiz, O., Vinothkumar, K. R., Goswami, P., and Kühlbrandt, W. (2006) Structure of the monomeric outer-membrane porin OmpG in the open and closed conformation. *EMBO J 25*, 3702–13.

[97] Subbarao, G. V. and van den Berg, B. (2006) Crystal structure of the monomeric porin OmpG. *J Mol Biol 360*, 750–9.

[98] Behlau, M., Mills, D. J., Quader, H., Kühlbrandt, W., and Vonck, J. (2001) Projection structure of the monomeric porin OmpG at 6 A resolution. *J Mol Biol 305*, 71–7.

[99] Liang, B. and Tamm, L. K. (2007) Structure of outer membrane protein G by solution NMR spectroscopy. *Proc Natl Acad Sci U S A 104*, 16140–5.

[100] Korkmaz-Ozkan, F., Köster, S., Kühlbrandt, W., Mäntele, W., and Yildiz, O. (2010) Correlation between the OmpG secondary structure and its pH-dependent alterations monitored by FTIR. *J Mol Biol 401*, 56–67.

[101] Damaghi, M., Bippes, C., Köster, S., Yildiz, O., Mari, S. A., Kühlbrandt, W., and Muller, D. J. (2010) pH-dependent interactions guide the folding and gate the transmembrane pore of the beta-barrel membrane protein OmpG. *J Mol Biol 397*, 878–82.

[102] Mari, S. A., Köster, S., Bippes, C. A., Yildiz, O., Kühlbrandt, W., and Muller, D. J. (2010) pH-induced conformational change of the beta-barrel-forming protein OmpG reconstituted into native E. coli lipids. *J Mol Biol 396*, 610–6.

[103] Chen, M., Khalid, S., Sansom, M. S. P., and Bayley, H. (2008) Outer membrane protein G: Engineering a quiet pore for biosensing. *Proc Natl Acad Sci U S A 105*, 6272–6277.

[104] Schein, S. J., Colombini, M., and Finkelstein, A. (1976) Reconstitution in planar lipid bilayers of a voltage-dependent anion-selective channel obtained from paramecium mitochondria. *J Membr Biol 30*, 99–120.

[105] Hinsch, K.-D., De Pinto, V., Aires, V. A., Schneider, X., Messina, A., and Hinsch, E. (2004) Voltage-dependent anion-selective channels VDAC2 and VDAC3 are abundant proteins in bovine outer dense fibers, a cytoskeletal component of the sperm flagellum. *J Biol Chem 279*, 15281–8.

[106] De Pinto, V., Guarino, F., Guarnera, A., Messina, A., Reina, S., Tomasello, F. M., Palermo, V., and Mazzoni, C. (2010) Characterization of human VDAC isoforms: a peculiar function for VDAC3? *Biochim Biophys Acta 1797*, 1268–75.

[107] Sabirov, R. Z. and Merzlyak, P. G. (2011) Plasmalemmal VDAC controversies and maxi-anion channel puzzle. *Biochim Biophys Acta*.

[108] Ujwal, R., Cascio, D., Colletier, J.-P., Faham, S., Zhang, J., Toro, L., Ping, P., and Abramson, J. (2008) The crystal structure of mouse VDAC1 at 2.3 A resolution reveals mechanistic insights into metabolite gating. *Proc Natl Acad Sci U S A 105*, 17742–17747.

[109] Hiller, S., Garces, R. G., Malia, T. J., Orekhov, V. Y., Colombini, M., and Wagner, G. (2008) Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science 321*, 1206–1210.

[110] Bayrhuber, M., Meins, T., Habeck, M., Becker, S., Giller, K., Villinger, S., Vonrhein, C., Griesinger, C., Zweckstetter, M., and Zeth, K. (2008) Structure of the human voltage-dependent anion channel. *Proc Natl Acad Sci U S A 105*, 15370–15375.

[111] Colombini, M. (2009) The published 3D structure of the VDAC channel: native or not? *Trends Biochem Sci 34*, 382–9.

[112] Hiller, S., Abramson, J., Mannella, C., Wagner, G., and Zeth, K. (2010) The 3D structures of VDAC represent a native conformation. *Trends Biochem Sci 35*, 514–21.

[113] Hodge, T. and Colombini, M. (1997) Regulation of metabolite flux through voltage-gating of VDAC channels. *J Membr Biol 157*, 271–279.

[114] Rostovtseva, T. and Colombini, M. (1997) VDAC channels mediate and gate the flow of ATP: implications for the regulation of mitochondrial function. *Biophys J 72*, 1954–62.

[115] Abu-Hamad, S., Arbel, N., Calo, D., Arzoine, L., Israelson, A., Keinan, N., Ben-Romano, R., Friedman, O., and Shoshan-Barmatz, V. (2009) The VDAC1 N-terminus is essential both for apoptosis and the protective effect of anti-apoptotic proteins. *J Cell Sci 122*, 1906–1916.

[116] Shoshan-Barmatz, V., Keinan, N., and Zaid, H. (2008) Uncovering the role of VDAC in the regulation of cell life and death. *J Bioenerg Biomembr 40*, 183–191.

[117] Abu-Hamad, S., Zaid, H., Israelson, A., Nahon, E., and Shoshan-Barmatz, V. (2008) Hexokinase-I protection against apoptotic cell death is mediated via interaction with the voltage-dependent anion channel-1: mapping the site of binding. *J Biol Chem 283*, 13482–13490.

[118] Abu-Hamad, S., Sivan, S., and Shoshan-Barmatz, V. (2006) The expression level of the voltage-dependent anion channel controls life and death of the cell. *Proc Natl Acad Sci U S A 103*, 5787–5792.

[119] Zaid, H., Abu-Hamad, S., Israelson, A., Nathan, I., and Shoshan-Barmatz, V. (2005) The voltage-dependent anion channel-1 modulates apoptotic cell death. *Cell Death Differ 12*, 751–760.

[120] Zalk, R., Israelson, A., Garty, E. S., Azoulay-Zohar, H., and Shoshan-Barmatz, V. (2005) Oligomeric states of the voltage-dependent anion channel and cytochrome c release from mitochondria. *Biochem J 386*, 73–83.

[121] Shoshan-Barmatz, V. and Gincel, D. (2003) The voltage-dependent anion channel. *Cell biochemistry and biophysics 39*, 279–292.

[122] Tsujimoto, Y. and Shimizu, S. (2000) VDAC regulation by the Bcl-2 family of proteins. *Cell Death Differ 7*, 1174–1181.

[123] Malia, T. J. and Wagner, G. (2007) NMR structural investigation of the mitochondrial outer membrane protein VDAC and its interaction with antiapoptotic Bcl-x(L). *Biochemistry 46*, 514–525.

[124] Rostovtseva, T. K., Tan, W., and Colombini, M. (2005) On the role of VDAC in apoptosis: fact and fiction. *J Bioenerg Biomembr 37*, 129–142.

[125] Colombini, M. (1989) Voltage gating in the mitochondrial channel, VDAC. *J Membr Biol 111*, 103–111.

[126] Colombini, M. (1979) A candidate for the permeability pathway of the outer mitochondrial membrane. *Nature 279*, 643–5.

[127] Zeth, K. and Thein, M. (2010) Porins in prokaryotes and eukaryotes: common themes and variations. *Biochem J 431*, 13–22.

[128] Schulz, G. E. (2000) beta-Barrel membrane proteins. *Curr Opin Struct Biol 10*, 443–7.

[129] Koppel, D. A., Kinnally, K. W., Masters, P., Forte, M., Blachly-Dyson, E., and Mannella, C. A. (1998) Bacterial expression and characterization of the mitochondrial outer membrane channel. Effects of N-terminal modifications. *J Biol Chem 273*, 13794–13800.

[130] Huse, M. and Kuriyan, J. (2002) The conformational plasticity of protein kinases. *Cell 109*, 275–282.

[131] Manning, G., Whyte, D., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of the human genome. *Science 298*, 1912.

[132] Adams, J. A. (2001) Kinetic and catalytic mechanisms of protein kinases. *Chem Rev 101*, 2271–90.

[133] Krebs, E G, E. G. and Fischer, E. H. (1956) The phosphorylase b to a converting enzyme of rabbit skeletal muscle. *Biochim Biophys Acta 20*, 150–7.

[134] Virshup, D. and Shenolikar, S. (2009) From promiscuity to precision: protein phosphatases get a makeover. *Mol Cell 33*, 537–545.

[135] Yaffe, M. (2002) Phosphotyrosine-binding domains in signal transduction. *Nat Rev Mol Cell Biol 3*, 177–186.

[136] Levinson, A., Oppermann, H., Levintow, L., Varmus, H., and Bishop, J. (1978) Evidence that the transforming gene of avian sarcoma virus encodes a protein kinase associated with a phosphoprotein. *Cell 15*, 561–572.

[137] Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S., and Sowadski, J. M. (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science 253*, 407–14.

[138] Bossemeyer, D., Engh, R. A., Kinzel, V., Ponstingl, H., and Huber, R. (1993) Phosphotransferase and substrate binding mechanism of the cAMP-dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0 A structure of the complex with Mn2+ adenylyl imidodiphosphate and inhibitor peptide PKI(5-24). *EMBO J 12*, 849–59.

[139] Zheng, J., Knighton, D. R., ten Eyck, L. F., Karlsson, R., Xuong, N., Taylor, S. S., and Sowadski, J. M. (1993) Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MgATP and peptide inhibitor. *Biochemistry 32*, 2154–61.

[140] Ma, H., Deacon, S., and Horiuchi, K. (2008) The challenge of selecting protein kinase assays for lead discovery optimization. *Expert Opin Drug Discov 3*, 607–621.

[141] Witt, J. J. and Roskoski, Jr, R. (1975) Rapid protein kinase assay using phosphocellulose-paper absorption. *Anal Biochem 66*, 253–8.

[142] Glass, D. B., Masaracchia, R. A., Feramisco, J. R., and Kemp, B. E. (1978) Isolation of phosphorylated peptides and proteins on ion exchange papers. *Anal Biochem 87*, 566–75.

[143] Hastie, C. J., McLauchlan, H. J., and Cohen, P. (2006) Assay of protein kinases using radiolabeled ATP: a protocol. *Nat Protoc 1*, 968–71.

[144] Mallari, R., Swearingen, E., Liu, W., Ow, A., Young, S. W., and Huang, S.-G. (2003) A generic high-throughput screening assay for kinases: protein kinase a as an example. *J Biomol Screen 8*, 198–204.

[145] Charter, N. W., Kauffman, L., Singh, R., and Eglen, R. M. (2006) A generic, homogenous method for measuring kinase and inhibitor activity via adenosine 5'-diphosphate accumulation. *J Biomol Screen 11*, 390–9.

[146] Rininsland, F., Xia, W., Wittenburg, S., Shi, X., Stankewicz, C., Achyuthan, K., McBranch, D., and Whitten, D. (2004) Metal ion-mediated polymer superquenching for highly sensitive detection of kinase and phosphatase activities. *Proc Natl Acad Sci U S A 101*, 15295–300.

[147] Chen, L., McBranch, D. W., Wang, H. L., Helgeson, R., Wudl, F., and Whitten, D. G. (1999) Highly sensitive biological and chemical sensors based on reversible fluorescence quenching in a conjugated polymer. *Proc Natl Acad Sci U S A 96*, 12287–92.

[148] Nikiforov, T. T. and Simeonov, A. M. (2003) Application of fluorescence polarization to enzyme assays and single nucleotide polymorphism genotyping: some recent developments. *Comb Chem High Throughput Screen 6*, 201–12.

[149] Turek-Etienne, T. C., Small, E. C., Soh, S. C., Xin, T. A., Gaitonde, P. V., Barrabee, E. B., Hart, R. F., and Bryant, R. W. (2003) Evaluation of fluorescent compound interference in 4 fluorescence polarization assays: 2 kinases, 1 protease, and 1 phosphatase. *J Biomol Screen 8*, 176–84.

[150] Kashem, M. A., Nelson, R. M., Yingling, J. D., Pullen, S. S., Prokopowicz, 3rd, A. S., Jones, J. W., Wolak, J. P., Rogers, G. R., Morelock, M. M., Snow, R. J., Homon, C. A., and Jakes, S. (2007) Three mechanistically distinct kinase assays compared: Measurement of intrinsic ATPase activity identified the most comprehensive set of ITK inhibitors. *J Biomol Screen 12*, 70–83.

[151] Kupcho, K., Somberg, R., Bulleit, B., and Goueli, S. A. (2003) A homogeneous, nonradioactive high-throughput fluorogenic protein kinase assay. *Anal Biochem 317*, 210–7.

[152] Lehel, C., Daniel-Issakani, S., Brasseur, M., and Strulovici, B. (1997) A chemiluminescent microtiter plate assay for sensitive detection of protein kinase activity. *Anal Biochem 244*, 340–6.

[153] Forrer, P., Tamaskovic, R., and Jaussi, R. (1998) Enzyme-linked immunosorbent assay for measurement of JNK, ERK, and p38 kinase activities. *Biol Chem 379*, 1101–11.

[154] Farley, K., Mett, H., McGlynn, E., Murray, B., and Lydon, N. B. (1992) Development of solid-phase enzyme-linked immunosorbent assays for the determination of epidermal growth factor receptor and pp60c-src tyrosine protein kinase activity. *Anal Biochem 203*, 151–7.

[155] Kobe, B. and Kemp, B. E. (1999) Active site-directed protein regulation. *Nature 402*, 373–6.

[156] Zhang, J., Yang, P. L., and Gray, N. S. (2009) Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer 9*, 28–39.

[157] Druker, B., Guilhot, F., O'Brien, S., Gathmann, I., Kantarjian, H., Gattermann, N., Deininger, M., Silver, R., Goldman, J., Stone, R., et al. (2006) Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *N Engl J Med 355*, 2408–2417.

[158] Daley, G., Van Etten, R., and Baltimore, D. (1990) Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science 247*, 824.

[159] O'Brien, S., Guilhot, F., Larson, R., Gathmann, I., Baccarani, M., Cervantes, F., Cornelissen, J., Fischer, T., Hochhaus, A., Hughes, T., et al. (2003) Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med 348*, 994–1004.

Chapter 2

Pim kinase detection by design

2.1 Introduction

While the stochastic sensing of kinases has been demonstrated using chemically and genetically modified pores, this has only been performed for a single kinase: protein kinase A. With so many kinases encoded in the human genome, and given the critical problem of developing selective inhibitors, it would be desirable to show this method can be applied to other kinases. Further, from an engineering perspective, the creation of sensing elements at the *trans* mouth of αHL has not yet been fully explored, for example the role of electrostatics and what limits may exist to what can be effectively incorporated to yield a functional sensor. Finally, both chemical and genetic modification methods have their individual advantages and disadvantages—further work may make it possible to combine the advantages of both approaches.

Towards these goals, this chapter details the design of stochastic sensors for the Pim family of kinases. These kinases possess a number of character-

istics that make them ideal for study by stochastic sensing as a progression of what has previously been established for protein kinase A. For instance, they are monomeric, constitutively active serine/threonine kinases for which a very high affinity consensus substrate sequence (Pimtide) has been established. They also have significant therapeutic interest due to their involvement in certain types of leukaemia and cancer. The existence of a high affinity substrate sequence is particularly tantalising as, opposed to previous work that used an inhibitory peptide, it should be possible to study enzymatic activity on the substrate.

It was decided to follow the genetic modification approach due to the ease and simplification afforded versus chemical modification, but also because the genetic approach is more amenable to rapid screening and combinatorial studies that may be pursued in the future.

This section begins by introducing the Pim family of kinases in detail. The design process for engineering αHL-based stochastic sensors is then detailed, followed by an account of the subsequent characterisation of the pores.

2.1.1 The Pim family of kinases

The *pim-1* gene was first discovered as a common proviral integration site for the Moloney murine leukaemia virus, which induces T-cell lymphomas in mice [1]. The oncogenic activity of *pim-1* was suggested by experiments where Pim-1 expression was upregulated in mice, resulting in an increased susceptibility to lymphomagenesis [2]. In the resulting T-cell lymphomas, it was found that either *c-myc* or *N-myc* were activated by proviral insertion,

suggesting cooperativity between *pim-1* and *myc*. This was confirmed by double transgenic *Eμ-pim-1 Eμ-myc* mice. Mice co-expressing both genes developed pre-B-cell leukaemia *in utero*, indicating one of the most potent oncogenic cooperative interactions so far described [3].

The *pim-1* gene was subsequently shown to encode a serine/threonine kinase [4–6]. However, Pim-1 deficient mice were found to present a largely normal phenotype, suggesting the existence of similar kinases. In due course two other Pim kinases, Pim-2 and Pim-3, were discovered [7–10]. Mice with all three *pim* genes knocked out develop normally but with reduced body size, presumably due to Pim kinases playing a role in cell proliferation [10].

In keeping with the proposed role in cell proliferation and survival, several substrates have been identified *in vitro* but without strong evidence in vivo. These include transcriptional repressors, activators, co-activators and regulators of the cell cycle [11–17]. Additionally, Pim kinases have been shown to modulate SOCS-1 degradation [18] and are able to phosphorylate the pro-apoptotic protein BAD at a key residue for inactivation [19–21]. Pim kinase substrates are reviewed in more detail elsewhere [22]. Interestingly, many Pim kinase substrates appear to be shared between Akt kinases, suggesting partially overlapping pathways that both regulate survival signalling [23].

Pim-1 and Pim-2 have a strongly basophilic substrate preference, with the strongest preference exhibited at positions -5 and -3 relative to the phosphorylation site. Additionally, a residue with a small side-chain is preferred at position $+1$ [24–26]. The contributions to binding of each of the nine positions surrounding a Pim-2 phosphorylation site were recently measured using a peptide library approach [26]. From this, a consensus substrate sequence

was determined (ARKRRRHPSGPPTA), which is referred to as 'Pimtide'. Pimtide binds to Pim-1 and Pim-3 with very high affinity, with $K_d \sim 40$–60 nM [27]. Attempts to measure Michaelis constants were hampered due to experimental limitations, but they are the same order of magnitude as the dissociation constants [27].

There is currently no evidence for the regulation of Pim kinases by post-translational modification; regulation appears to be at the transcriptional and translational levels. Pim-1 and Pim-2 expression levels are significantly upregulated in leukocytes by the JAK/STAT pathway in response to cytokines [19, 28–31].

Structure

The *pim-1* gene produces two translation products, one a 36 kDa protein and the other a 44 kDa protein through alternative initiation at an upstream CUG codon [32]. Both products have serine/threonine kinase activity [32]. The crystal structure of Pim-1 (36 kDa form) has been solved in the apo-form and bound to a variety of inhibitors and substrates in binary and ternary complexes [27, 33–35]. These structural studies revealed Pim-1 to be in a constitutively active conformation, without the need for phosphorylation of residues in the activation loop. Although Pim-1 possesses the archetypal two-lobe fold it also has several distinguishing features (Fig 2.1).

First, Pim-1 possess an additional β-hairpin in the N-terminal lobe [33]. Two turn residues in this hairpin are subject to diffuse large cell lymphoma (DLCL)-related somatic hypermutation [37]. Some of these mutations have been found to significantly alter kinase activity [35].

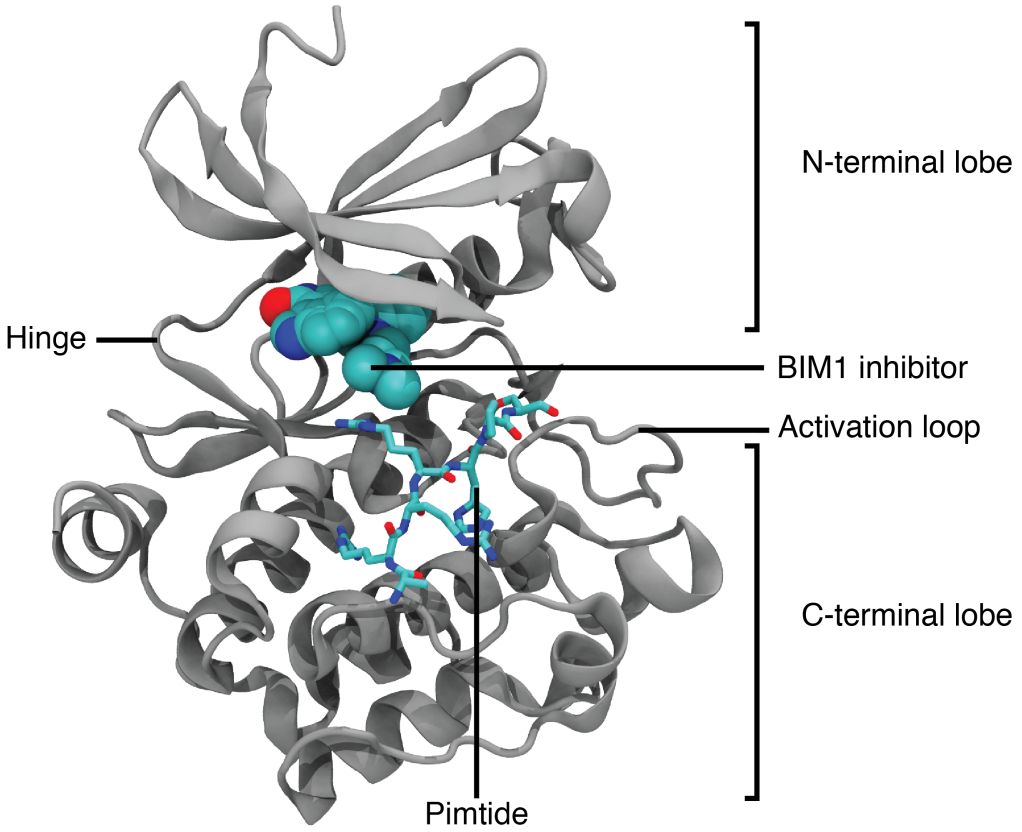


Figure 2.1: Crystal structure of Pim-1 kinase showing the positions of key features. Pim-1 (grey) is shown complexed with the inhibitor BIM1 (space-filling) and the Pim-1 consensus substrate peptide 'Pimtide' (sticks). Atomic coordinates are from PDB: 2BIL and were visualised in PyMOL [36].

Second, Pim-1 possesses a unique hinge region. Most kinases possess a highly conserved residue in the hinge region that forms a backbone NH hydrogen bond with ATP in the active conformation. However, in Pim-1 this residue is Pro123, which is unable to participate in hydrogen bonding [33]. Further, a number of inserted residues relative to other kinases with high homology, as well as a *cis*-proline (Pro125), result in an expansion of the ATP-binding pocket. Hence, ATP is found to bind in a distinct fashion from other kinases as it lacks one of two highly conserved hydrogen bond interactions with the hinge. Further, a typical salt-bridge between the conserved Lys169 residue and the nucleotide phosphate is also absent. This unique binding mode suggests the possibility of developing Pim-selective inhibitors [38].

2.1.2 Chapter overview

This chapter details the engineering of stochastic sensors for the Pim family of kinases. It begins with the engineering of a genetically encoded sensor analogous to that previously engineered to sense protein kinase A [39]. Further optimisation and engineering was required to obtain a detectable signal and this is also detailed. A method for the production of αHL pores bearing a peptide attached by a single terminus at the *trans* mouth was also developed, and is described, whereby such pores can be generated in a simple protocol from a genetically encoded pore by proteolytic cleavage of a mutant *trans* loop. The characterisation of these pores is detailed, followed by preliminary experiments that suggest that phosphorylation of the attached Pimtide peptide can be observed, which paves the way for further studies of enzymatic

activity at the single molecule level.

2.2 Results

2.2.1 Genetically encoded *trans* loop pores

It was decided to design a genetically encoded Pim kinase sensor based on αHL in an analogous manner to the previously described genetically encoded protein kinase A stochastic sensor [39]. This approach offers several benefits over targeted chemical modification (cf. Xie *et al.* [40]). First, the procedures for synthesising the modified pore are simplified, as it is not necessary to first chemically synthesise the Pimtide peptide and then couple it to heteromeric αHL. Second, chemical coupling is not 100% efficient, so a proportion of pores will be unmodified, whereas all genetically encoded pores will possess the Pimtide peptide.

The first design to be tested comprised the Pimtide sequence plus eight-residue glycine/serine linkers at both the N- and C-terminal ends of the Pimtide peptide, all inserted into the *trans* loop of αHL, replacing the residues 129–130 in the wild type αHL sequence (Figure 2.2). This modification is denoted 'PLM' (Pimtide Loop Mutant). The gene constructed also possessed an octa-aspartate 'tail' at the C-terminus to facilitate the separation of heteroheptameric pores by SDS-PAGE [41].

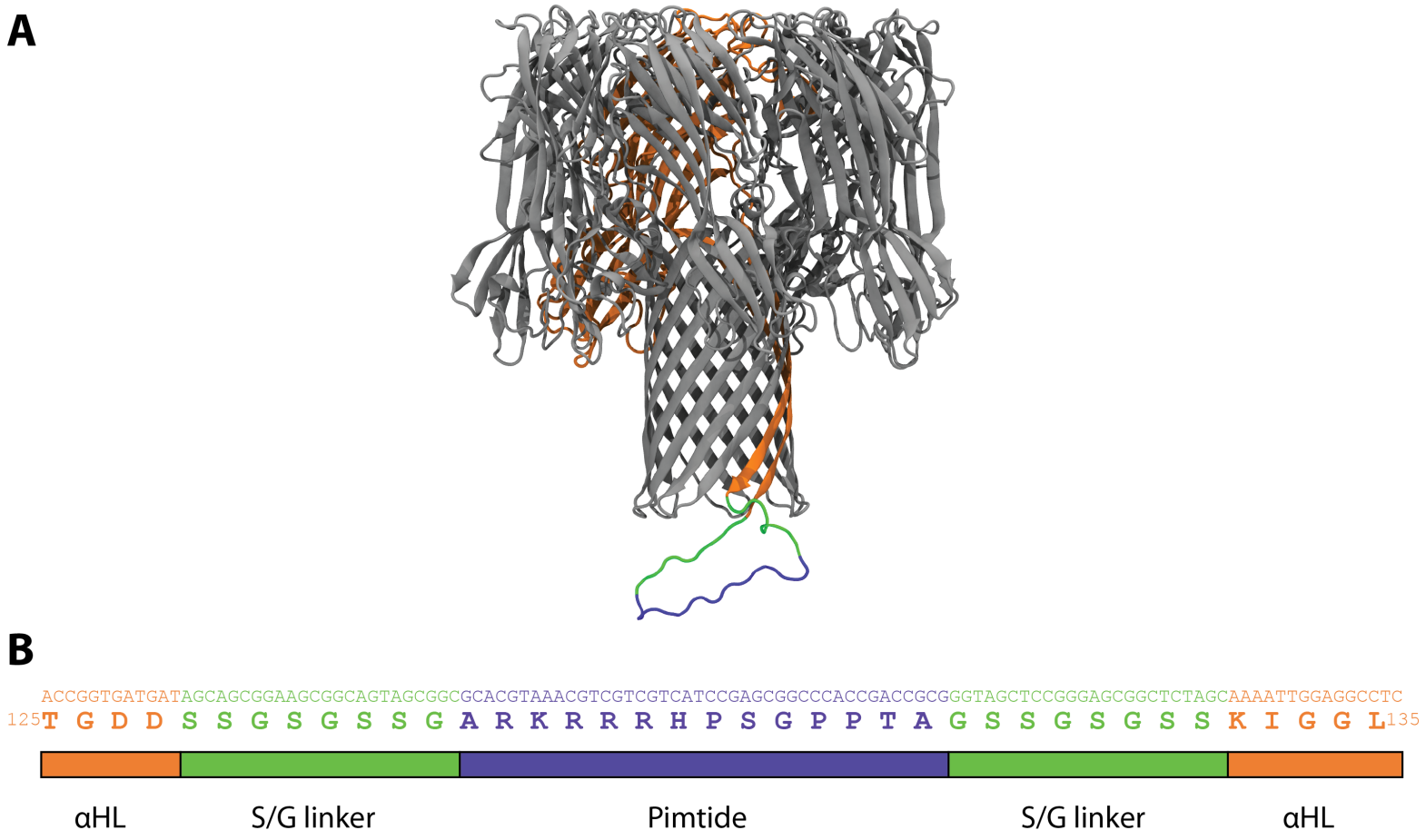


Figure 2.2: Design of αHL-PLM. (A) Model of $(\alpha\text{HL-WT})_6(\alpha\text{HL-PLM})_1$ in cartoon representation. The PLM-containing subunit is coloured orange, the serine/glycine linkers are coloured green and the Pimtide peptide is coloured blue. (B) Sequences of the PLM-modified *trans* loop region showing the DNA coding sequence and amino acid sequence. Colours are as in (A).

Heteromeric pores containing a single Pimtide sequence were prepared by *in vitro* transcription and translation (IVTT) of the wild type αHL and αHL-PLM-D8 genes together in the presence of rabbit erythrocyte membranes, which mediate the oligomerisation of αHL. The resulting mixture of homo- and heteroheptameric species was separated by SDS-PAGE [41]. When the wild type αHL (αHL-WT) and αHL-PLM-D8 plasmids were present in a ratio of 6:2, $(\alpha\text{HL-PLM-D8})_1(\alpha\text{HL-WT})_6$ (αHL-PLM hereafter) was the predominant species. Species containing more than two PLM subunits were not observed, even when αHL-PLM-D8 DNA was present in higher quantities than αHL-WT or was the only plasmid present (Figure 2.3). This may be due to excessive steric crowding and perhaps electrostatic repulsion making the formation of heptamers containing more than two modified subunits energetically unfavourable.

It is important to note that this assignment of stoichiometry relies on the presumption that each step downwards in the ladder of bands in Figure 2.3 corresponds to the presence of one further αHL-PLM-D8 subunit. However, as the full ladder of eight bands is not observed in this case, this cannot be proven unambiguously.

Single αHL-PLM heteromer pores were characterised by single channel recording. In initial experiments it was observed that insertion of these pores was much less frequent than for WT homoheptamers or αHL-PKIP. This could be due to the higher net charge possessed by Pimtide at the experimental pH versus the PKIP5–24 sequence used in the previous genetically encoded kinase sensor (+5 vs +1) [39]. When single αHL-PLM heteromer pores were successfully reconstituted into planar bilayers, the addition of

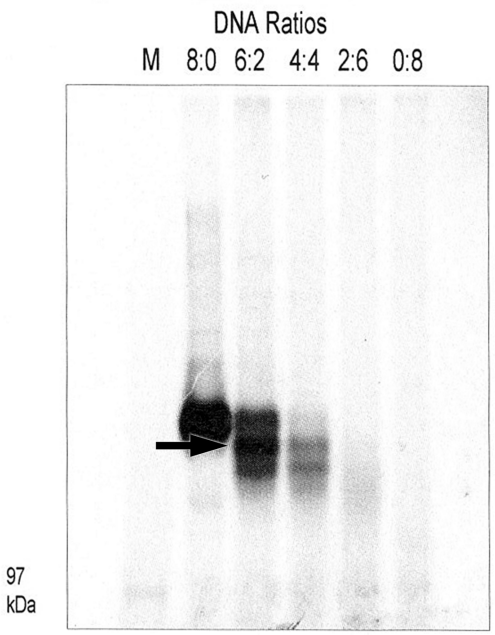


Figure 2.3: Autoradiograph of preparative SDS-PAGE of αHL-PLM. Various ratios of αHL-PLM-D8 and αHL-WT encoding plasmids were co-translated in IVTT reactions with rabbit erythrocyte membranes, as described in the main text (ratio shown in the form WT:PLM). The prepared heptamers were electrophoresed on a 5% SDS-PAGE gel. The arrow marks the location of the $(\alpha\text{HL-PLM-D8})_1(\alpha\text{HL-WT})_6$ heteroheptamer.

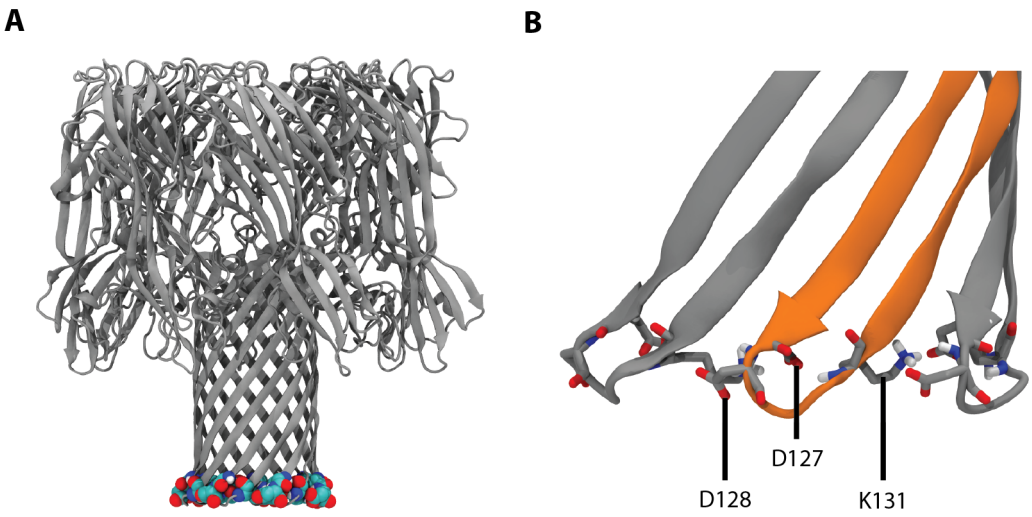


Figure 2.4: Location of charged residues at the αHL *trans* mouth. (A) View of the full heptameric pore; the *trans* mouth charges D127, D128 and K131 are shown in space-filling form. (B) Zoomed view of the *trans* loops of three subunits in the heptameric pore; the three charged residues on a single subunit (highlighted in orange) are labelled; charged residues from all three subunits are shown in stick form.

200 nM Pim-1 to the *trans* chamber did not produce a measurable change in the pore's electrical behaviour (Figure 2.7 A).

Pore insertion was observed to occur more frequently when a lower pH buffer was used (pH 4.0 vs pH 7.4, data not shown). The *trans* mouth of αHL possesses a ring of charged residues at its entrance, formed by D127, D128 and K131 (Figure 2.4). It was hypothesised that the increased pore insertion observed at low pH is due to protonation of the acidic aspartate residues that leads to a net neutral or positive pore *trans*-mouth, either of which could present a lower energetic barrier to insertion of a pre-assembled αHL heptamer. It was desired to perform single channel experiments in a near-neutral buffer to preserve kinase activity, hence the manipulation of the charges at the *trans* mouth of the pore was attempted by mutagenesis.

Trans mouth charge mutations

A set of (αHL-PLM)$_1$(αHL)$_6$ heteroheptamers was created, with each member containing one of the following mutations in all subunits: D127N, D128N, D127N-D128N, D127N-D128N-K131Q, and D127N-D128N-K131A. The corresponding homoheptamers formed of αHL without the PLM modification were also produced. Initial characterisation was performed in 20 mM HEPES pH 7.4, 250 mM KCl. Representative traces and I-V relationships for the homoheptamers are shown in Figure 2.5.

The most notable feature of the majority of the mutants was the presence of strong, slow gating at negative applied potentials (data not shown). This was independent of the PLM modification, as the corresponding homoheptamers without the PLM modification also exhibited this gating. This was considered undesirable as it was expected that kinase-binding events would be observed under negative applied potential (based on previous studies [39, 40]), and gating could obscure or otherwise interfere with binding events.

A trend was observed where the propensity of pores to insert in the bilayer increased with the number of neutralised *trans* mouth residues, i.e. the D127N/D128N/K131A mutant inserted most readily and the single mutants the least (data not shown). However, it is difficult to perform properly controlled experiments to measure insertion propensity as there are many variables that are difficult or impossible control, for example insertion is promoted by destabilisation of the bilayer through application of voltage and/or stirring, and bilayer stability may vary from one preparation to another.

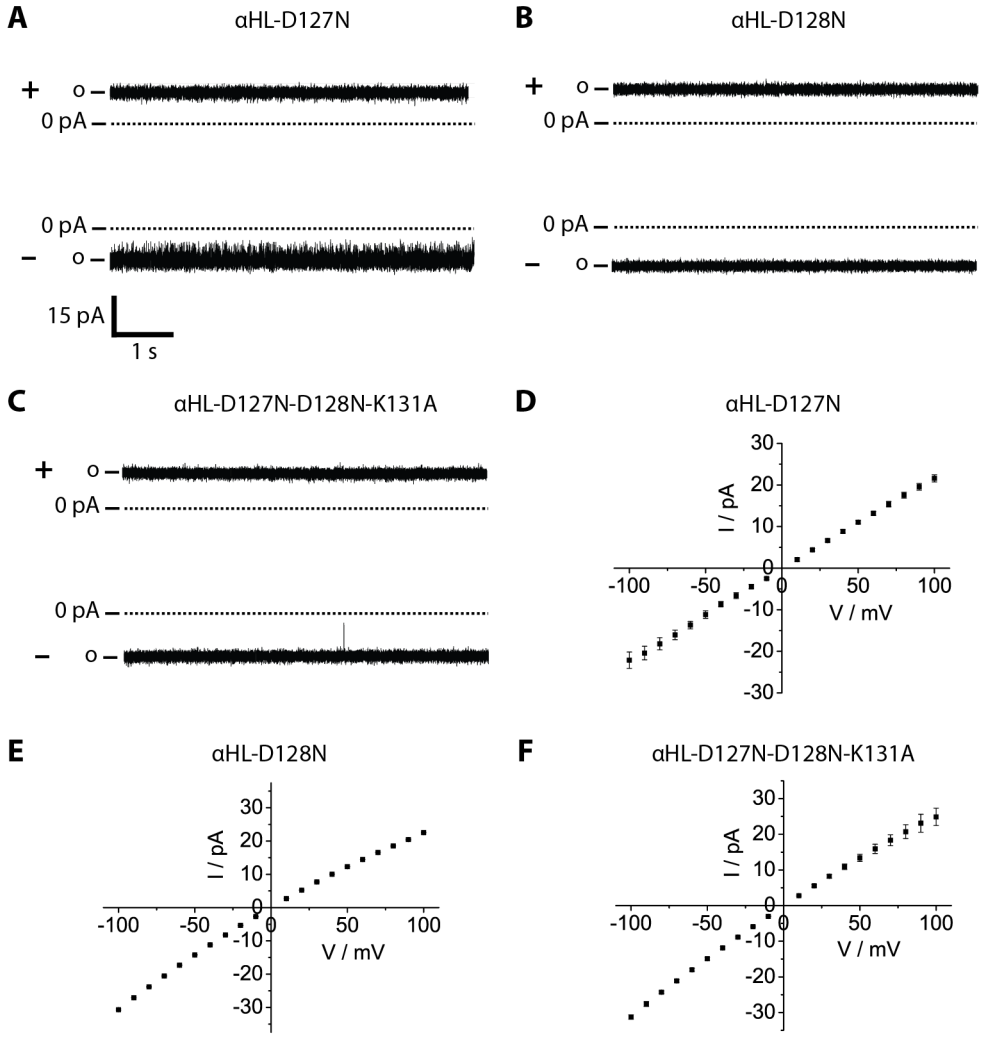


Figure 2.5: Representative traces of several homoheptameric αHL *trans* mouth charge mutants. (A) αHL-D127N at +50 mV and −50 mV. (B) αHL-D128N at +50 mV and −50 mV. (C) αHL-D127N-D128N-K131A at +50 mV and −50 mV. (D) I-V relationship for αHL-D127N. (E) I-V relationship for αHL-D128N. (F) I-V relationship for αHL-D127N-D128N-K131A. All data were recorded in 20 mM HEPES pH 7.4, 250 mM KCl. I-V data points are all $n = 3$. Error bars represent the standard deviation of the mean. The open state in each case is marked 'o'. The filter cut-off frequency was 1 kHz.

Hence these observations must be treated with caution. In any case, the insertion rate of the singly-neutralised mutants was sufficient for the efficient performance of experiments.

One mutant, D127N, did not exhibit gating at negative applied potential. Further, the PLM-containing heteroheptamer exhibited fast events at positive applied potential that may be attributable to conformational motions of the PLM-containing *trans* loop (Fig 2.6 A (i)). The (αHL-D127N-PLM-D8)$_1$(αHL-D127N)$_6$ heteroheptamer was identified as the most promising mutant and was therefore examined in the presence of Pim-1 to see if kinase binding could be detected.

Before addition of Pim-1, (αHL-D127N-PLM-D8)$_1$(αHL-D127N)$_6$ ('αHL-PLM-D127N' hereafter) in 15 mM MOPS pH 6.8, 300 mM KCl at +50 mV exhibited short-lived transitions to a partially closed state with a mean dwell time of 0.41$\pm$0.03 ms ($n = 3$), and an inter-event interval of 6.6$\pm$1.4 ms ($n = 3$) (Figure 2.6 A, (ii) and (iii)). These partially closed events are presumably due to the Pimtide loop interacting with the mouth of the pore and are hence referred to as blockades. The open state current was 12.7 ± 0.4 pA.

Under these conditions, introduction of Pim-1 to the *trans* chamber gave rise to the intermittent abolition of the previously observed short current blockades (Figure 2.6 B (i)). After addition of Pim-1, dwell time analysis indicated that the open state exhibited two distinct dwell time populations (Figure 2.6 B (iii)). In the presence of 100 nM Pim-1, mean dwell times of 4.72 ± 0.16 ms ($n = 3$) and 100 ± 11 ms ($n = 3$) were observed. The mean lifetime of the short blockade events was 0.46$\pm$0.03 ms ($n = 3$) (Figure 2.6 B (ii)). Notably, no alteration in channel behaviour was observed at negative

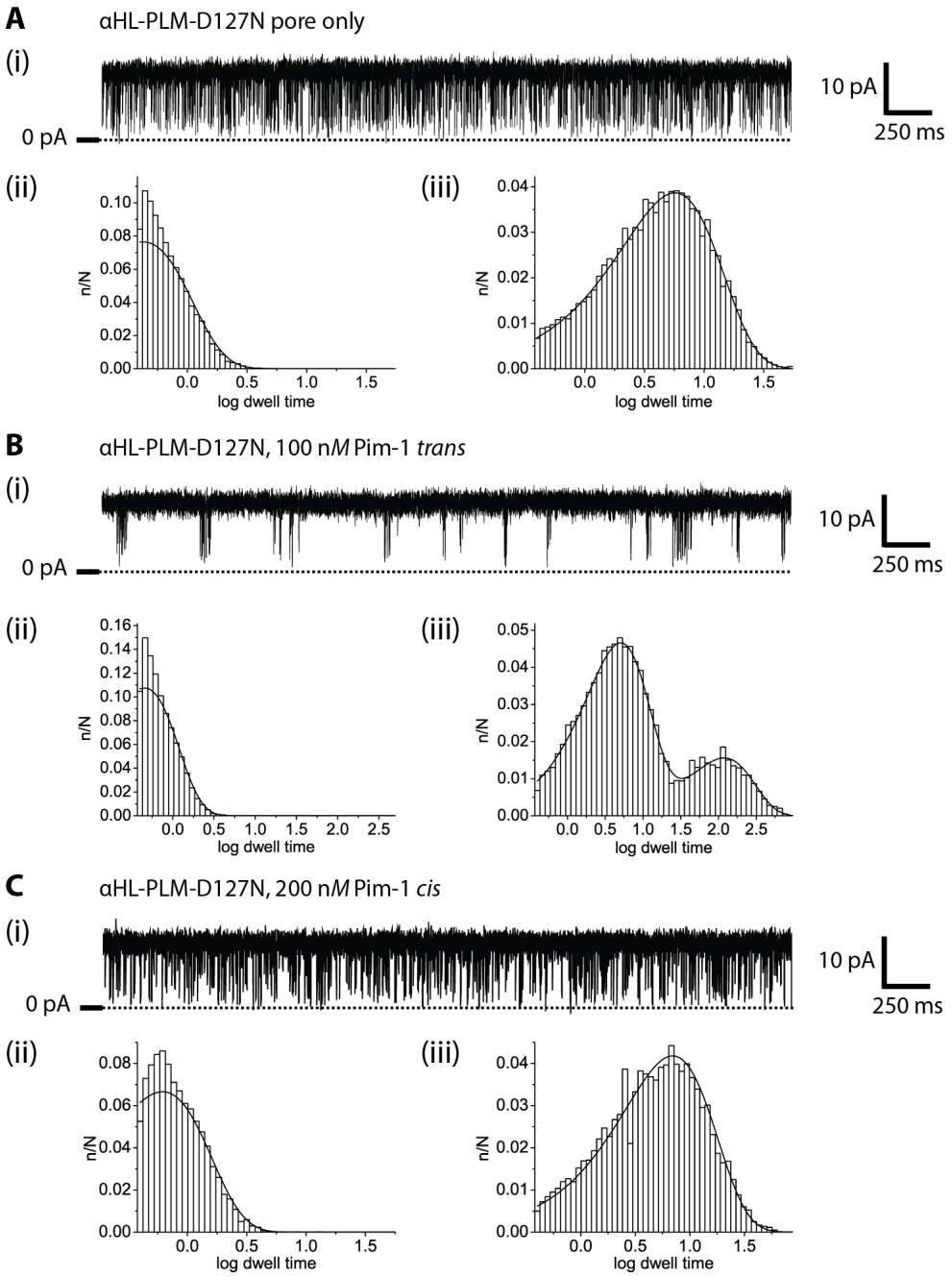


Figure 2.6: Comparison of channel behaviour of αHL-PLM-D127N heteroheptamer before (A), and after the addition of (B) 100 nM Pim-1 to the *trans* chamber, or (C) 200 nM Pim-1 to the *cis* chamber. For each representative current trace (i), representative dwell time histograms and fitted probability density functions are shown for (ii) the blocked level and (iii) the open level. The traces were recorded in 15 mM MOPS pH 6.8, 300 mM KCl buffer with an applied potential of +50 mV. The filter cut-off frequency was 2 kHz, except for (C), where it was 1 kHz. Traces were idealised and dwell time histograms fitted in QuB.

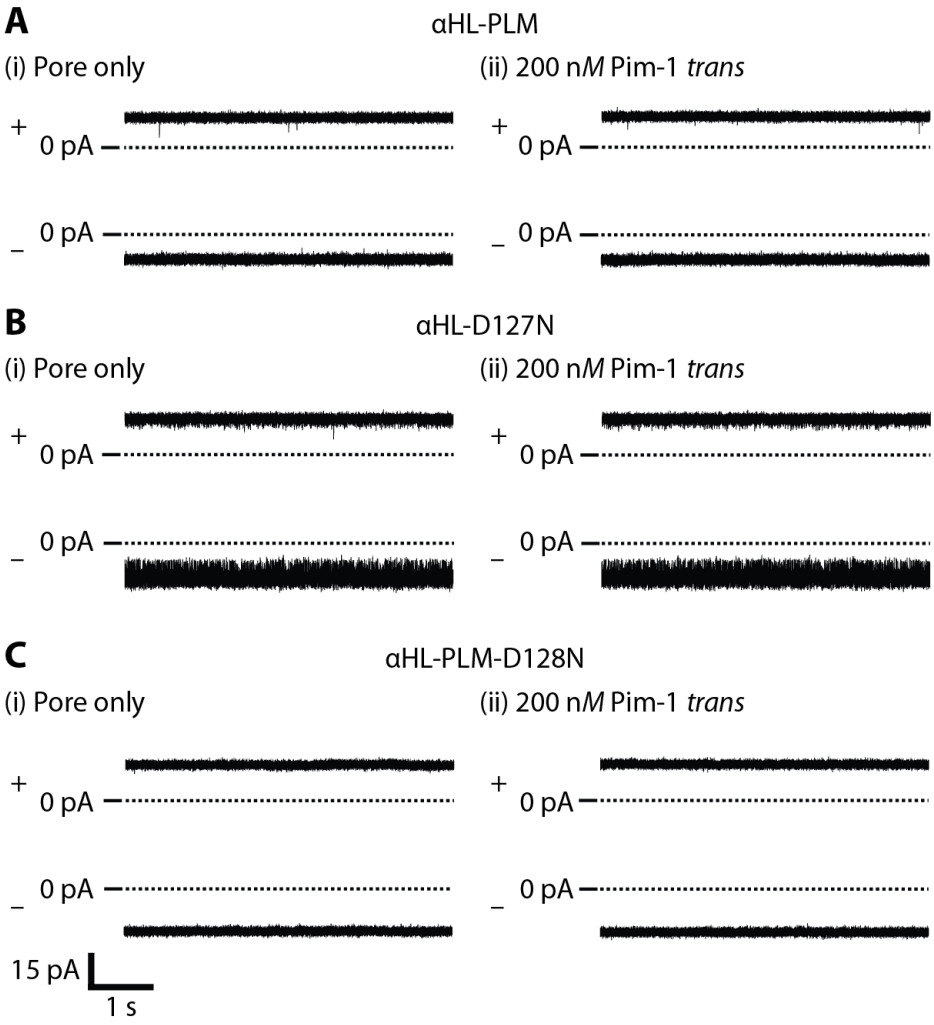


Figure 2.7: Negative binding results. Representative current traces for (A) αHL-PLM heteromer, (B) αHL-D127N homo-heptamer, and (C) αHL-PLM-D128N heteromer, in the (i) absence or (ii) presence of 200 nM Pim-1 in the *trans* chamber. Measurements were made at $\pm$50 mV as indicated, in 15 mM MOPS pH 6.8, 300 mM KCl and 5 mM DTT. The filter cut-off frequency was 2 kHz.

applied potential. To confirm that this change in current behaviour was consistent with binding of Pim-1 to the *trans* loop of αHL-PLM-D127N, Pim-1 was also added to the *cis* chamber in a separate experiment. As expected, no alteration in current behaviour was observed (Figure 2.6 C).

To confirm that the effect of Pim-1 when present in the *trans* chamber is the result of a specific interaction with the Pimtide-containing *trans* loop, the effect of Pim-1 on the αHL-D127N homo-heptamer was studied. With 200 nM Pim-1 present in the *trans* chamber, no alteration in current behaviour was observed (Figure 2.7 B). Additionally, no change in behaviour was seen for the αHL-PLM-D128N heteromer when Pim-1 was added to the *trans* chamber (Figure 2.7 C). Hence, it may be suggested that the importance of the D127N mutation for the detection of Pim-1 binding is due to it allowing current blockades arising from the *trans* loop to be observed, enabling kinase binding to be detected indirectly via its effect on these blockades.

Although Pim-1 binding was apparent, it occured without creating a current blockade, in contrast with the previously studied αHL-PKIP pore [39]. It was hypothesised that this may be because bound kinase molecules were too far from the pore *trans* mouth to measurably alter the current flowing through the channel. Hence, variants of the αHL-D127N-PLM pore were constructed where the serine/glycine linker sections in the *trans* loop were reduced in length to draw the Pimtide section closer to the *trans* mouth.

Linker length variation

The original PLM modification features eight serine/glycine 'linker' residues on either side of the Pimtide peptide. Two variants of αHL-D127N-PLM were created where the linker length was reduced from eight to four (SG4) and two (SG2). The electrical properties of these variants were then examined (Figure 2.8).

The SG4 variant behaved similarly to the original αHL-D217N-PLM pore, with short-lived current blockades observed at positive potential (Figure 2.8 SG4 (i)). However, the open state exhibited a higher level of noise, which prevented reliable idealisation of traces for analysis.

Under the conditions used, the current flowing through the pore is comparatively low, and hence a low frequency cut-off filter is required to observe the signal. However, the blockade events observed with this and the original αHL-D217N-PLM pore are short—somewhat less than $1/f_c$ (where f_c is the filter cut-off frequency, which in this case was 2 kHz). Under such conditions, a significant proportion of blockade events are attenuated in amplitude due to insufficient sampling. Those attenuated $> 50\%$ will be missed during idealisation by the half-maximum threshold method. Difficulty was also experienced using the segmental k-means (SKM) method. However, using a higher cut-off frequency is not possible here as the signal would be lost within the noise. Additionally, channel gating (mostly at negative potential) appears at potentials above ± 50 mV. Hence, these factors present a limitation to the reliable analysis of these recordings and a limitation of these mutants.

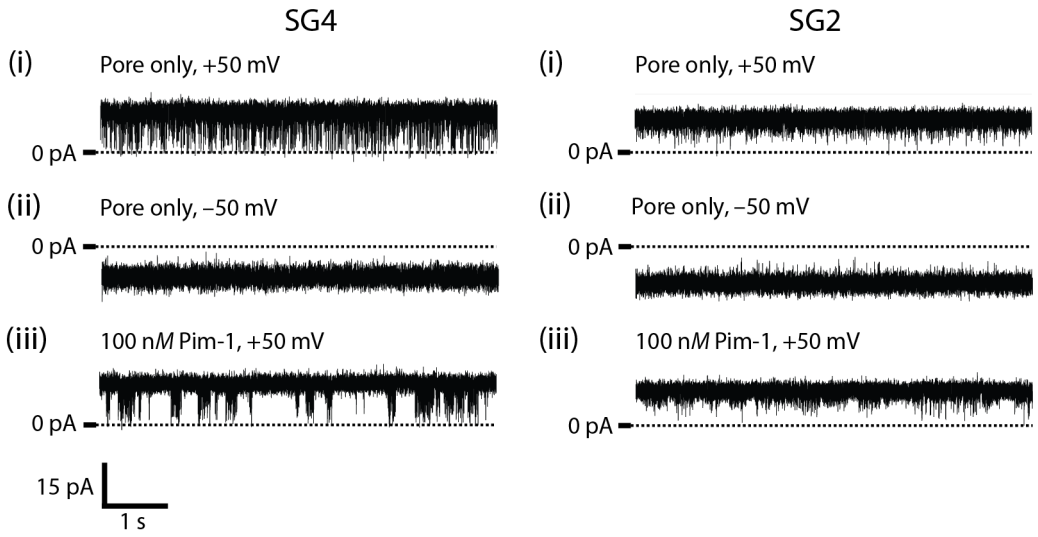


Figure 2.8: Representative traces of the αHL-PLM linker length variants SG4 and SG2: (i) at +50 mV, (ii) at −50 mV, (iii) at +50 mV in the presence of 100 nM Pim-1. All measurements were performed in 15 mM MOPS pH 6.8, 300 mM KCl. The filter frequency cut-off was 2 kHz.

Upon addition of Pim-1 to the *trans* chamber, intermittent abolition of the current blockades was observed in a qualitatively similar fashion as for the αHL-D217N-PLM pore, with no accompanying current blockade (Figure 2.8 SG4 (iii)). The SG2 variant exhibited behaviour similar to SG4 but with further increased noise in the open state such that short-lived blockades if present were obscured (Figure 2.8 SG2 (i)). The addition of Pim-1 did not induce a detectable change in channel behaviour (Figure 2.8 SG2 (iii)).

Based on these observations, it appears that shortening the linker length does not result in the bound kinase being held in such a way as to produce a current blockade. It is clear that the linkers can be shortened by as much as half while still enabling binding, however at one quarter the original length binding is no longer observed. These variants did not present an improvement over the original αHL-PLM-D127N pore so were not studied further.

2.2.2 Protease-cleavable pores

The previously developed chemically modified kinase-sensing αHL pore αHL-PKIP presented a single PKI peptide attached by a single terminus to the *trans* mouth of the pore [40]. Additionally, the pore exhibited a 'clean' signal consisting of a single conductance state, with current blockades upon the binding of protein kinase A molecules. However, the method is laborious and does not result in a quantitative yield of modified protein. The genetic encoding method is more expedient and all pores possess a single copy of the encoded *trans* loop peptide. However, in this case the peptide is conformationally constrained within the loop with each peptide terminus attached to the pore, albeit via flexible linkers.

It was posited that it may be possible to engineer an αHL mutant that contains a genetically encoded *trans* loop peptide insertion that also contains a protease cleavage site, such that cleavage of the *trans* loop at one terminus of the insert peptide would result in a peptide attached to the *trans* mouth by a single terminus (Figure 2.9). This would present an expedient route to producing pores analogous to those chemically modified to attach peptides, eliminating the requirement for peptide synthesis and chemical modification steps. Further, having all elements genetically encoded could be advantageous, for example for the construction of libraries of attached peptides. Finally, it was hoped that this topology would lead to the observation of current blockades when Pim kinase molecules bind, in an analogous fashion to that seen for the chemically modified αHL-PKIP pore [40].

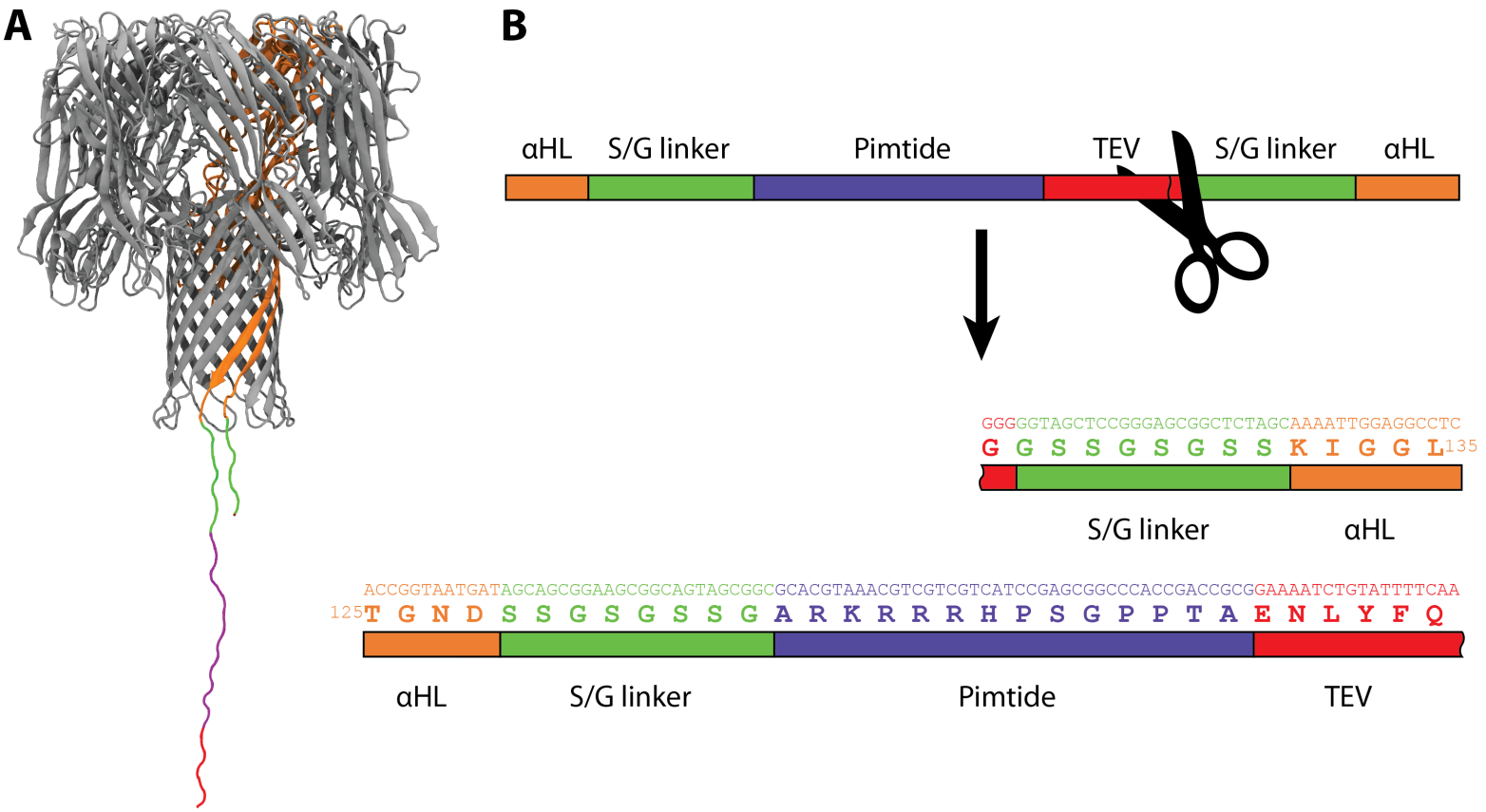


Figure 2.9: Design of the TEV protease cleavable αHL-PLM-TEV heteromer pore. (A) Illustrative model of (αHL-PLM-TEV-D127N-D8)$_1$(αHL-D127N)$_6$ heteroheptamer in cartoon representation, indicating the expected pore topology after oligomerisation of αHL-D127N monomer and TEV protease-cleaved αHL-PLM-TEV-D127N-D8 monomer and purification of the desired heteroheptamer by SDS-PAGE. (B) Sequence view of the *trans* loop region indicating the placement of the TEV protease cleavage site for the αHL-PLM-TEV variant, and indicating the free termini (depicted by wavy lines) produced after TEV protease cleavage.

Construction

A simple protocol was devised whereby a tobacco etch virus (TEV) protease site (ENLYFQG) was inserted either immediately preceding or succeeding the Pimtide sequence of αHL-PLM-D127N-D8 (Figure 2.9 shows the scenario for placement of the cleavage site after the Pimtide sequence). The intention was that these monomers would be cleaved with TEV protease before being oligomerised with unmodified monomers to form hetero-heteromers containing a single, cleaved Pimtide-bearing monomer.

Two constructs were made, αHL-PLM-TEV-D127N-D8 and αHL-TEV-PLM-D127N-D8, where the TEV protease cleavage site was inserted preceding and succeeding the Pimtide sequence, respectively. It was then necessary to confirm that the site was cleaved efficiently by TEV protease and to optimise cleavage conditions. Both mutants were synthesised in 25 μl IVTT reactions. Protein synthesis was then halted by the addition of chloramphenicol, and aliquots were taken before the addition of one unit of TEV protease and at one hour intervals afterwards until four hours had elapsed. The cleavage was performed at 4°C to avoid adventitious proteolysis from endogenous proteases that may have been present in the S30 extract used in the IVTT reaction. The aliquots were then analysed by SDS-PAGE and autoradiography (Figure 2.10).

It was clear that the protease site was accessible and cleavage occurred readily—after two hours the cleaved monomer is the major species. A small amount of uncleaved monomer did remain after four hours. However, the protease is also present during a subsequent oligomerisation step at room

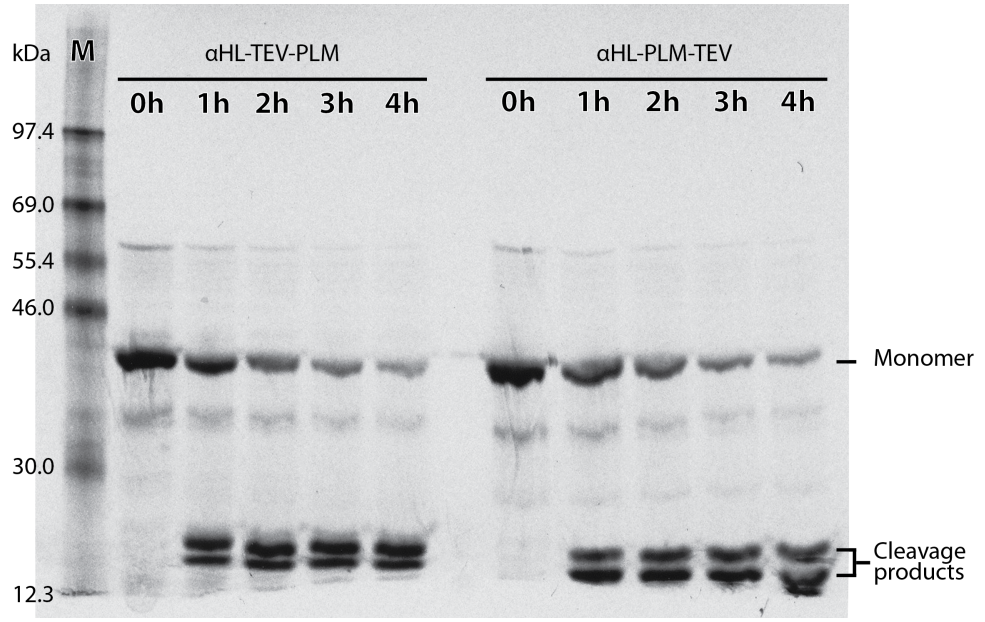


Figure 2.10: TEV protease digestion time courses for αHL-TEV-PLM-D8 monomer and αHL-PLM-TEV-D8 monomer. Samples were electrophoresed on a 10% bis-tris SDS-PAGE gel (Bio-Rad) in XT-MOPS running buffer. The markers (M) are ^{14}C-methylated protein standards (Perkin Elmer).

temperature for one hour when performing the preparative procedure for hetero-oligomers. It was therefore expected that any traces of remaining uncleaved monomer, present as monomer or incorporated into a heteroheptamer, would be cleaved to completion. It was therefore decided to incubate at 4°C for four hours to ensure maximal cleavage but minimise adventitious proteolysis.

To produce heteroheptamers, TEV protease-cleaved monomer was prepared under the optimised conditions. It was then mixed with αHL-D127N monomer, also prepared by IVTT, and rabbit erythrocyte membranes were subsequently added. The mixture was incubated for one hour at room temperature to allow oligomerisation to occur. The mixture was then purified by SDS-PAGE as previously (Figure 2.11 A). An uncleaved heteroheptamer

was also prepared so that the electrical properties of the constrained-loop pore could be compared with those of the TEV protease-cleaved, unconstrained pore and hence demonstrate the incorporation of proteolytically cleaved monomer.

Gel-purified (αHL-PLM-TEV-D127N-D8)$_1$(αHL-D127N)$_6$ was analysed by SDS-PAGE before and after boiling, to confirm the incorporation of TEV-cleaved monomer and determine the stoichiometry of the heteromer. As expected, after boiling, the major band moved from the molecular weight expected for heptamer (heptameric αHL migrates faster than predicted, as it remains folded in SDS) to that expected for monomer (Figure 2.11 B). A second, weaker band appeared at approximately half the molecular weight of the monomer, which is interpreted to be the two cleavage products from the TEV-cleaved monomer at a coincident position, unresolved.

Gel densitometry was performed to determine the relative intensity of monomer to cleavage products. As the cleavage product bands are coincident, a ratio of 6:1 would be expected. The ratio calculated by densitometry was 8:1. The reason for this difference is not known but may be due to the limited dynamic range of the X-ray film resulting in the exposure being outside the linear range. As such, while it is presumed that each band in a mixture of heteromers, such as lane 2 in Figure 2.11 A, differs by an interval of one αHL-PLM-TEV-D127N-D8 subunit, this cannot be proven unambiguously with the current data.

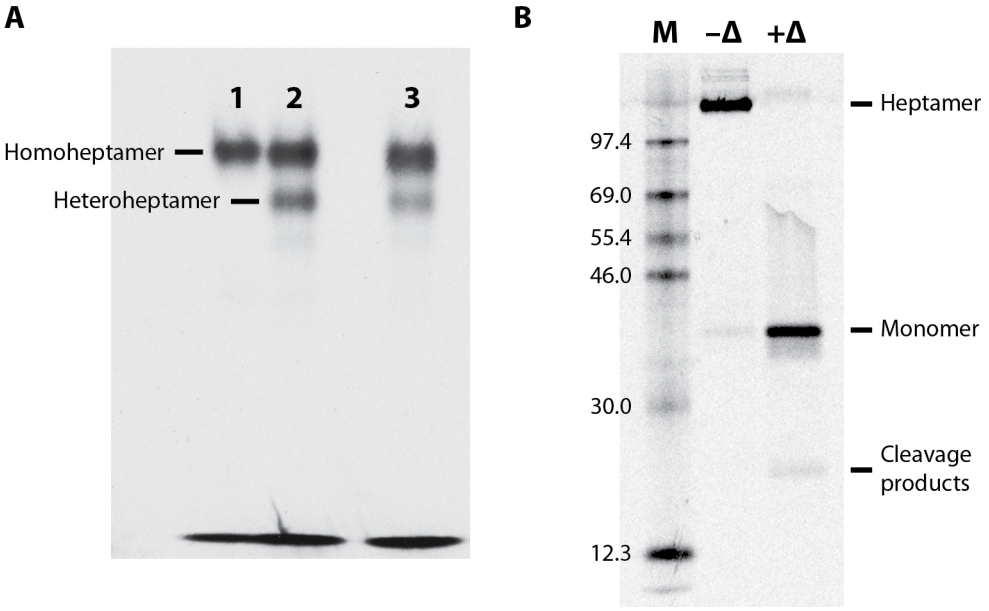


Figure 2.11: (A) Autoradiograph of preparative SDS-PAGE of αHL-PLM-TEV mutants. 1) αHL-D127N homoheptamer. 2) αHL-PLM-TEV-D127N-D8 (uncut) heteroheptamer-containing mixture. 3) αHL-PLM-TEV-D127N-D8 (cut) heteroheptamer-containing mixture. The positions of homoheptamers and heteroheptamers containing a single αHL-PLM-TEV-D127N-D8 subunit are marked. The intense bands at the bottom of the gel correspond to the dye front. (B) Autoradiograph of gel-purified (αHL-TEV-PLM-D127N-D8)$_1$(αHL-D127N)$_6$ before and after boiling, analysed by SDS-PAGE (12% bis-tris gel, Bio-Rad).

Electrical properties

The electrical properties of the hetero-heptameric pores containing single TEV protease-cleaved subunits were then studied by single channel recording. Representative traces are compared in Figure 2.12. (αHL-TEV-PLM-D127N-D8)$_1$(αHL-D127N)$_6$, hereafter referred to as αHL-TEV-PLM, formed noisy channels that exhibited pronounced rectification, with the conductance at negative potentials being much less than at positive potentials (Figure 2.12 A). Further, no alteration in channel behaviour was apparent upon the addition of Pim-1 kinase (data not shown).

In contrast, (αHL-PLM-TEV-D127N-D8)$_1$(αHL-D127N)$_6$, hereafter referred to as αHL-PLM-TEV, formed channels that exhibited defined gating events with notably different kinetics depending on the polarity of the applied potential (Figure 2.12 B (i) and (ii)). At a positive applied potential of +50 mV, transitions to a partially closed current level were observed with several distinct dwell times (Figure 2.12 B (iv)). The dwell time histogram for the partially closed current level was best fitted with three component functions, giving mean durations of 0.22 ± 0.05, 2.24 ± 0.89, and 18.0 ± 6.9 pA (all $n = 3$) (Figure 2.12 B (iv)). The mean inter-event interval was 52 ± 11 ms ($n = 3$) (Figure 2.12 B (iii)). The open state amplitude was 12.7 ± 0.8 pA ($n = 7$) and the partially closed current level was 1.36 ± 0.02 pA ($n = 7$), or approximately 89% of the open level (data not shown). At negative potential, the predominantly occupied state was a partially closed state of mean amplitude -1.30 ± 0.07 pA ($n = 6$). Short-lived transitions to an open state of mean amplitude -10.8 ± 0.7 pA ($n = 6$) were observed. The mean par-

tially closed lifetime was 7.6 ± 2.2 ms ($n = 6$) (Figure 2.12 B (vi)) and the mean inter-event interval (lifetime of the open state) was 1.36 ± 0.22 ms ($n = 6$) Figure 2.12 B (v)). The innate, partially closed states at both positive and negative applied potential presumably arise due to the presence of the Pimtide peptide that protrudes from the *trans* mouth. Hence, these events are hereon described in the terminology of blockade events and blocked current levels, as it is hypothesised that the Pimtide peptide is transiently blocking the *trans* mouth of the pore.

Upon the addition of 200 nM Pim-1 kinase to the *trans* chamber at an applied potential of -50 mV, transitions of αHL-PLM-TEV to the open state were observed with a mean lifetime of 86.8 ± 3.2 ms ($n = 3$), in addition to the much shorter events innate to the pore, which had a mean dwell time in the open state of 1.3 ± 0.4 ms ($n = 3$) (Figure 2.16 A and C). In the presence of 200 nM Pim-1, the mean lifetime of the innate blockades (mean dwell time in the blocked state) was 7.8 ± 2.3 ms ($n = 3$) (Figure 2.16 B). The dwell times corresponding to the innate pore events were hence both unchanged, within error, from the values measured in the absence of Pim-1. All values reported in this section were obtained from measurements performed in 15 mM MOPS pH 6.8, 300 mM KCl and 5 mM dithiothreitol (DTT). The reason for the inclusion of DTT is given below.

Next, the effect of proteolytic cleavage of αHL-PLM-TEV on the electrical properties of the pore was confirmed by comparing its properties with those of un-cleaved αHL-PLM-TEV. Un-cleaved αHL-PLM-TEV formed channels that were predominantly in the open state under both polarities of applied potential, with short transitions to a blocked state (Figure 2.13 (i) and (ii)).

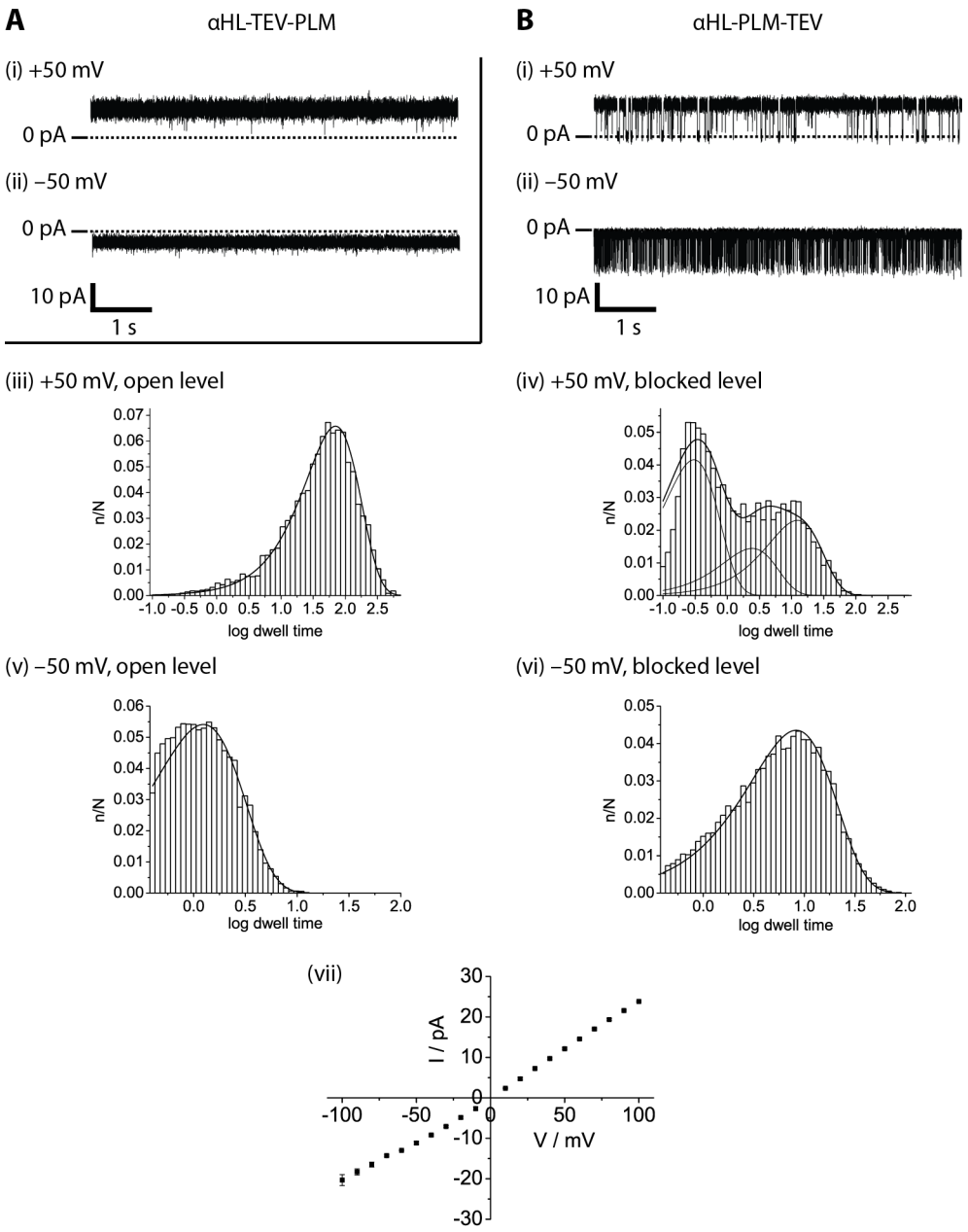


Figure 2.12: Behaviour of TEV protease-cleaved αHL-TEV-PLM and αHL-PLM-TEV heteroheptamers. (A) Representative traces of αHL-TEV-PLM at (i) +50 mV and (ii) −50 mV. (B)Representative traces of αHL-PLM-TEV at (i) +50 mV and (ii) −50 mV. (iii)–(vi) Representative dwell time histograms and fitted probability density functions for the open current levels and partially closed (blocked) current levels of αHL-PLM-TEV at ±50 mV as indicated. (vii) I-V relationship for αHL-PLM-TEV; all data points $n = 3$. Error bars represent the standard deviation of the mean. Traces were acquired in 15 mM MOPS pH 6.8, 300 mM KCl (A), and also 5 mM DTT (B). The filter cut-off frequency was 2 kHz.

As such the pore is reminiscent of αHL-PLM-D127N, but differing at negative potential, where the presence of blockade events is unique to the un-cleaved αHL-PLM-TEV pore. Consistent with this comparison, when 100 nM Pim-1 was added to the *trans* chamber, intermittent abolition of the short current blockades was observed in a qualitatively similar manner to the αHL-PLM-D127N pore (Figure 2.13 (iii), (v) and (vi)). The markedly different electrical properties of cleaved and un-cleaved αHL-PLM-TEV support the conclusion that proteolytic cleavage of the *trans* loop of αHL-PLM-TEV was successful.

A notable advantage of cleaved αHL-PLM-TEV over the loop-constrained αHL-PLM-D127N pore is that its current traces are much more amenable to reliable idealisation and analysis. First, the short-lived transitions to the open state at negative applied potential are of sufficient duration for adequate sampling under the acquisition conditions used. Second, although kinase binding does not create a current blockade, kinase binding events are clearly delimited as a transition from one conductance level (blocked) to a distinct current level (open). As such, this pore presents a significant improvement over the pores described earlier.

Kinetic analysis of Pim-1 interaction

The kinetics of the Pim-1 interaction with the cleaved αHL-PLM-TEV heteromer pore were next examined, to confirm that the longer lifetime events observed upon addition of Pim-1 arise from a bimolecular association and unimolecular dissociation of the complex, and to derive the association and dissociation rate constants and the dissociation equilibrium constant. References hereafter to αHL-PLM-TEV refer to the proteolytically cleaved form

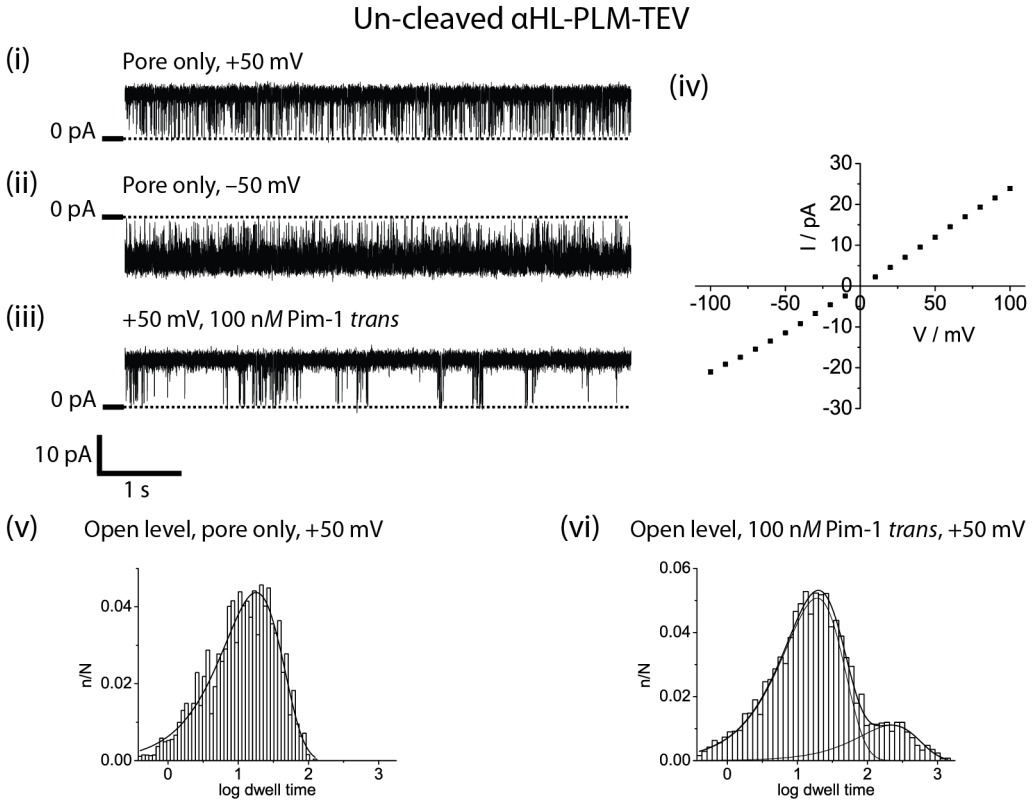


Figure 2.13: Representative traces of the un-cleaved αHL-PLM-TEV heteroheptamer pore. (i) Pore under applied potential of +50 mV. (ii) Pore under applied potential of −50 mV. (iii) Pore after addition of 100 nM Pim-1 to the *trans* chamber, under an applied potential of +50 mV. (iv) I-V relationship for the pore. All data were recorded in 15 mM MOPS pH 6.8, 300 mM KCl and 5 mM DTT. I-V data points are all $n = 3$. Error bars represent the standard deviation of the mean. (v) and (vi) Representative dwell time histograms and fitted probability density functions for the open pore level at +50 mV before and after the addition of 100 nM Pim-1. The filter cut-off frequency was 2 kHz.

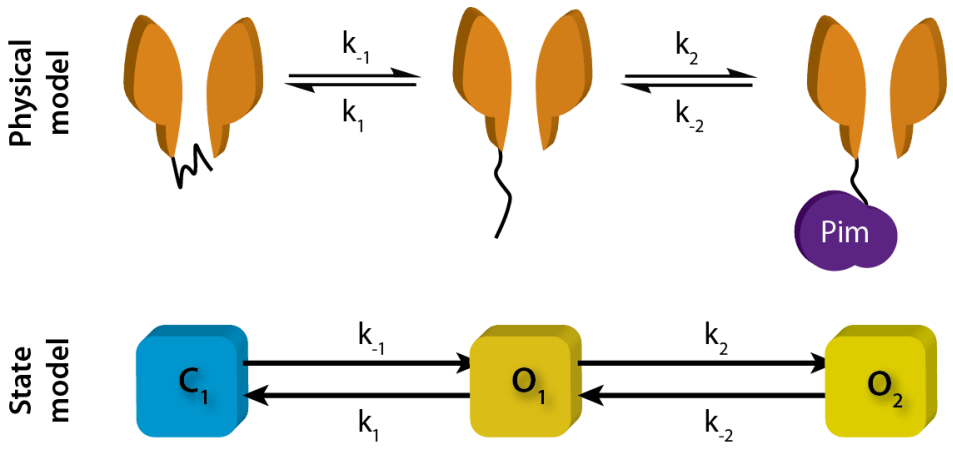


Figure 2.14: State model used for hidden Markov analysis in QuB of αHL-PLM-TEV traces recorded in the presence of Pim-1, and the corresponding physical interpretation. C_1 is a blocked state of mean amplitude -1.30 ± 0.07 pA ($n = 6$), and O_1 and O_2 are two open states with mean amplitude -10.8 ± 0.7 pA ($n = 6$). Standard deviations used are the standard deviation of Gaussian noise around the mean current. Arrows denote connections between states and their associated rate constants are also denoted.

of the pore. Pim-1 was titrated against the pore over the concentration range 50–400 nM and the recorded current traces were analysed using the program QuB (www.qub.buffalo.edu).

The current traces were best described by a three state model with two distinct amplitudes, as shown in Figure 2.14. Consistent with the interpretation that the partially blocked conductance level under negative applied potential corresponds to blockade of the pore by the Pimtide peptide (leading to the blocked state C_1), it is presumed that kinase may only bind when the pore is in the open state (O_1), leading to the kinase-bound open state O_2. Figure 2.14 illustrates the correspondence between these states in the model and the physical interpretation. Rate constants were estimated by the maximum interval likelihood (MIL) method within QuB, and the concentration dependences of the apparent association and dissociation rate constants

for Pim-1 binding (k_2' and k_{-2} respectively, where $k_2' = k_2$[Pim-1]) were examined.

Initial attempts were frustrated by a lack of consistency in the apparent Pim-1 on-rate (k_2') from pore to pore. Further, for a single pore, the apparent on-rate did not increase linearly with concentration of Pim-1 but fell off at higher concentrations. Further investigation found that when short sections of a long recording (e.g. 30 minutes) were analysed at regular intervals the apparent on-rate decreased rapidly with time (Figure 2.15 A). The probable cause was thought to be loss of Pim-1 activity through oxidation of cysteine residues. To counter this, measurements were performed in 15 mM MOPS pH 6.8, 300 mM KCl with the addition of 5 mM DTT. Although the reducing power of DTT is diminished significantly below neutral pH, there was still an excess of DTT in the active thiolate form relative to Pim-1 at the concentrations used.

It was found that the presence of DTT preserved the activity of Pim-1 in the chamber over the course of approximately twenty minutes, as shown by a constant apparent on-rate within error (Figure 2.15 B). Another possible explanation for the role of DTT is that it may chelate small amounts of inhibitory metal ions present as trace contaminants in salts or plasticware. This has been found to be the case for lambda phosphatase, which is strongly inhibited by Zn^{2+} [42]. Pim-1 is also strongly inhibited by Zn^{2+}, which is uncharacteristic of protein kinases [24]. As the goal was simply to preserve Pim-1 activity during single-channel recording, no further evidence for the role of DTT was sought.

Under the new conditions of 15 mM MOPS pH 6.8, 300 mM KCl and

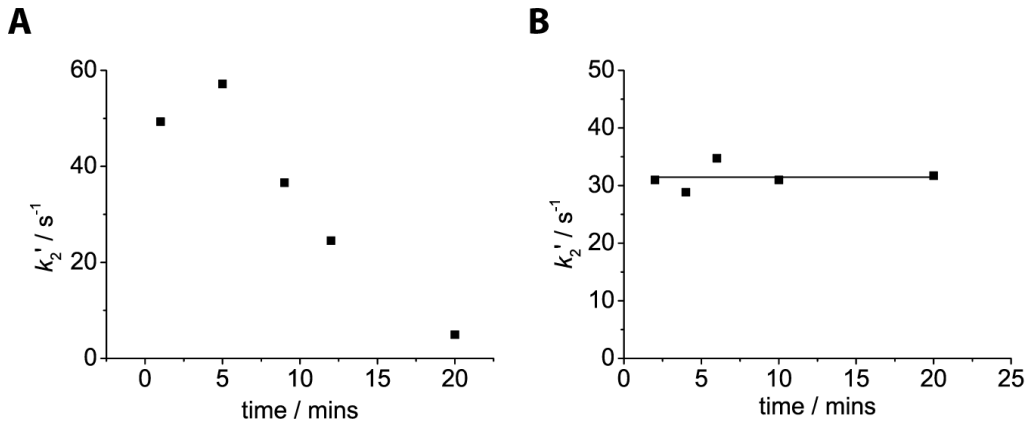


Figure 2.15: Time dependence of the apparent on-rate of Pim-1 binding to αHL-PLM-TEV. (A) Measured in 15 mM MOPS pH 6.8, 300 mM KCl. (B) Measured in 15 mM MOPS pH 6.8, 300 mM KCl, 5 mM DTT. Times were measured from the point of addition of 200 nM Pim-1 to the *trans* chamber. Each data point corresponds to a two minute section of recording. The calculated apparent on-rate (k_2') is plotted with the x-coordinate being the mid-point between the start and end time of the two minute section. Each panel represents a single experiment.

5 mM DTT, titrations were performed of Pim-1 at 50, 100, 200, and 400 nM with an applied potential of -50 mV, and the apparent on- and off-rates were plotted against Pim-1 concentration (Figure 2.16 D). Consistent with a bimolecular process, the apparent on-rate was found to be linearly dependent on Pim-1 concentration, whereas the off-rate was found to be invariant with respect to Pim-1 concentration. The association rate constant, k_2, was determined from the gradient of the fitted linear function ($y = mx$) from the plot of apparent on-rate (k_2') versus concentration and was determined to be $(1.11 \pm 0.07) \times 10^8$ M^{-1} s^{-1}. The off-rate was determined to be 10.9 ± 0.8 s^{-1}. The dissociation constant was hence 99 ± 10 nM.

It is notable that the association rate constant is quite high for a protein-protein interaction, where such rate constants are typically expected to be on the order of $10^6 M^{-1}$ s^{-1} [43]. The association rate constant of pro-

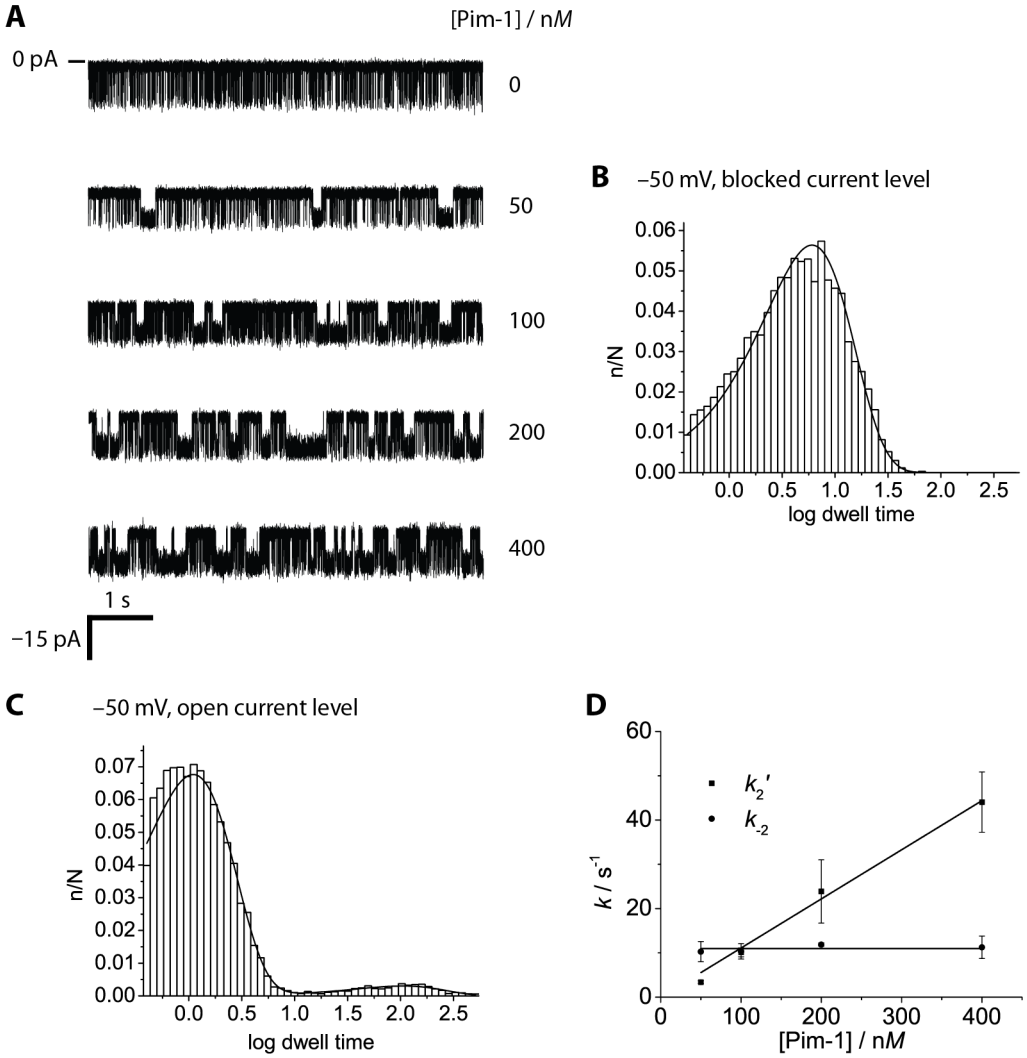


Figure 2.16: Kinetic analysis of Pim-1 interaction with the αHL-PLM-TEV heteromeric pore. (A) Representative traces of channel behaviour in the presence of increasing concentrations of Pim-1. (B) Representative dwell-time histogram and fitted probability density function of the blocked level. (C) Representative dwell-time histogram and fitted probability density function of the open level. (D) Concentration dependence of the kinase off-rate constant, k_{-2}, and the pseudo-first order kinase on-rate constant, k_2', where $k_2' = k_2$[Pim-1]. Error bars represent the standard deviation of the mean ($n = 3$). All measurements were performed in 15 mM MOPS pH 6.8, 300 mM KCl and 5 mM DTT with an applied potential of −50 mV. The filter cut-off frequency was 2 kHz.

tein kinase A for the PKIP-pore developed by Xie *et al.* was found to be $2.5 \times 10^6 M^{-1}$ s^{-1} [40]. However, it is known that long-range electrostatic interactions can enhance association rates by several orders of magnitude [44]. In the case of the Pimtide–Pim-1 interaction, Pimtide has a high positive net charge, whereas Pim-1 is net negative and has a large number of acidic residues in the substrate binding site [27]. Isothermal titration calorimetry (ITC) measurements indicated that binding was strongly driven by enthalpy, presumably due to the large number of polar interactions between the substrate and active site [27]. Hence, it is reasonable to justify the measured association rate constant on the basis of electrostatic 'steering'.

The measured value for the dissociation constant of 99 nM compares well with the literature value of 58 nM [27]. The literature value was determined by ITC in 50 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT. The value presented in this work was measured at higher ionic strength (300 mM KCl) and so is consistent with the weakened affinity expected due to increased electrostatic screening.

2.2.3 Phosphorylation

Previous work on the stochastic sensing of protein kinase A was based on observing the interaction between the kinase and an inhibitory peptide, precluding any study of enzymatic activity. However, the present work has utilised Pimtide, a high-affinity consensus substrate. It was therefore interesting to investigate whether phosphorylation of the Pimtide-bearing αHL pores could be observed by single channel recording.

Initial observations were made using both the αHL-PLM-D127N and αHL-PLM-TEV (cleaved) pores. Pim-1 phosphorylates in the presence of ATP and either Mg^{2+} or Mn^{2+} [24]. To observe phosphorylation, a single pore was reconstituted into a planar bilayer and observed for a short time in buffer consisting of 15 mM MOPS pH 6.8, 300 mM KCl. Pim-1 was then added to the *trans* chamber (100 nM final concentration) and binding and dissociation was observed for a few minutes. To bring about phosphorylation, $MgCl_2$ and ATP were added to the *trans* chamber at final concentrations of 1 mM and 500 μM, respectively. On most occasions an altered channel behaviour was observed immediately after stirring of the *trans* chamber had been completed (typically 30 s in duration). Occasionally, binding of Pim-1 would continue for some time before an alteration in channel behaviour was observed, as would be expected for a stochastic process.

In the case of αHL-PLM-D127N, the altered channel behaviour after the addition of MgATP consisted of short-lived blockades (Figure 2.17 A). Periods of abolished blockades attributed to Pim-1 binding were no longer observed. It was not possible to quantitatively analyse the traces due to the limitations discussed previously (the combination of increased open state noise and blockade event dwell times below $1/f_c$). However, the blockades appear qualitatively altered in kinetics compared to the pore before addition of MgATP. This may arise due to the altered charge of the Pimtide peptide after the addition of a phosphate group.

In the case of αHL-PLM-TEV, the altered channel behaviour consisted of blockade events of mean dwell time 8.0 ± 1.4 ms ($n = 5$) and mean inter-event interval (dwell time in the open state) of 1.31 ± 0.24 ms ($n = 5$) (Figure 2.17

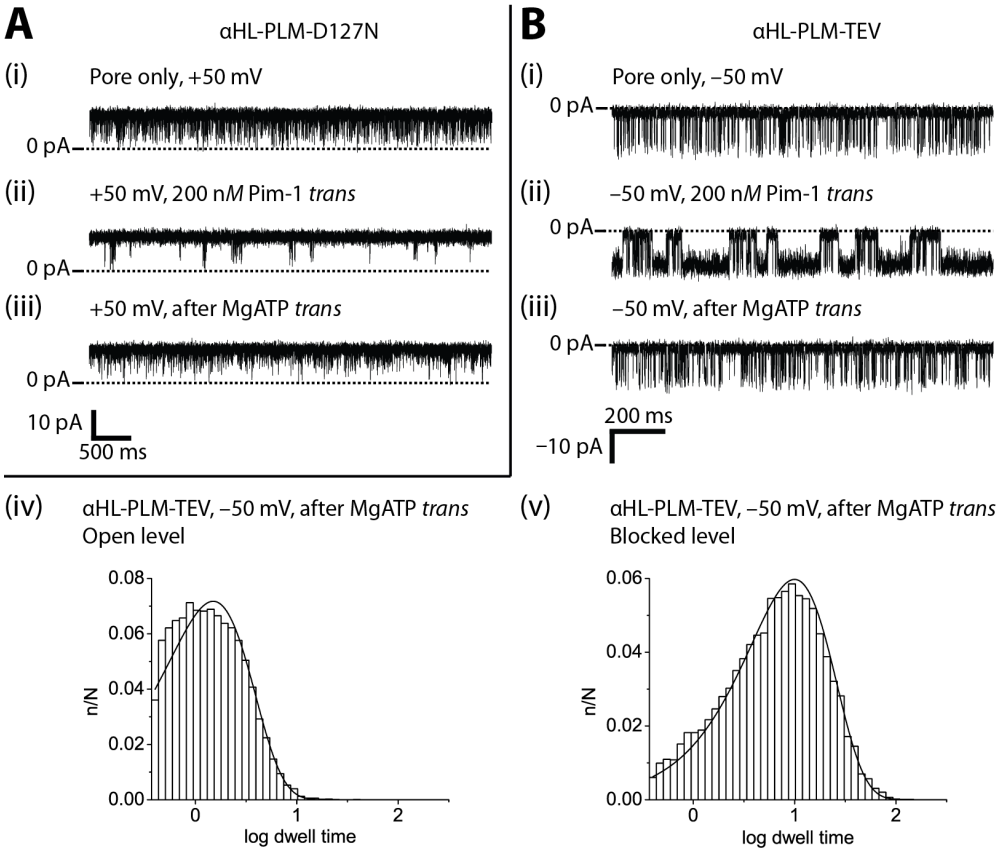


Figure 2.17: Representative traces illustrating the change in channel behaviour following the addition of Pim-1 and MgATP. (A) αHL-PLM-D127N pore: (i) before addition of Pim-1 or MgATP, (ii) after the addition of 200 nM Pim-1, (iii) after the addition of MgATP. (B) αHL-PLM-TEV pore: (i) before addition of Pim-1 or MgATP, (ii) after the addition of 200 nM Pim-1, (iii) after the addition of MgATP. (iv) and (v) Representative dwell time histograms and fitted probability density functions for the open and blocked levels of αHL-PLM-TEV after the addition of Pim-1 and MgATP. Traces were recorded in 15 mM MOPS pH 6.8, 300 mM KCl. The applied potential was (A) +50 mV or (B) −50 mV. The filter cut-off frequency was 2 kHz.

B (iv) and (v)). Transitions to the open state, with longer life-time, were seen very infrequently (i.e. on the order of one event per minute or less). In contrast to the loop-constrained αHL-PLM-D127N pore, it would appear that the blockade kinetics are unaffected by phosphorylation.

These observations appear consistent with phosphorylation of the Pimtide peptide, in that after the transfer of a phosphate group to the Pimtide substrate, kinase binding at the bulk level should be abolished. The single molecule interpretation would be that the presence of the phosphate group would cause the on-rate of the kinase to be dramatically reduced, but still have a finite value, so very infrequent binding events would still be expected to be observed, as was the case in these experiments.

While this observation is consistent with phosphorylation of the pore, several alternative explanations exist. First, ATP could induce a change in Pim-1 that would prevent its binding to the pore. This could arise either due to a conformational change induced by ATP binding to Pim-1, or through autophosphorylation of the kinase. Previous results from experiments using protein kinase A indicate that ATP binding does affect the affinity of protein kinase A for peptide substrates. However, this manifests as a marked increase in affinity for the peptide substrate, rather than a decrease [39, 40, 45, 46].

Autophosphorylation of human Pim-1 has previously been shown to occur [27, 34]. However, comparison of pools of Pim-1 in different phosphorylation states isolated from *E. coli* expression by ion exchange chromatography indicated that each pool had comparable kinetic parameters [34]. Second, it could be possible that ATP is itself inducing the effect by acting directly on the pore. This is unlikely as there are no documented nucleotide binding

sites on αHL, but cannot be ruled out by the present data.

An obvious limitation of performing the experiment as described above is that the observation of phosphorylation is single-shot. In principle, several hundred such experiments could be performed, with the time between addition of MgATP and the abolition of kinase binding events for each experiment plotted as a histogram. It would then be possible to fit the histogram with an exponential probability function to derive the mean lifetime for the unphosphorylated pore, and hence obtain the phosphorylation rate constant. However, this would be extremely laborious. Nonetheless, this observation should pave the way for more sophisticated experiments able to more efficiently measure catalytic activity, as discussed below.

2.3 Conclusions

A series of genetically encoded stochastic sensors of Pim-1 have been developed. In the first iteration, a pore bearing the Pimtide peptide was engineered in an analogous fashion to a previous sensor for protein kinase A [39]. However, this pore did not exhibit any alteration of channel behaviour upon addition of Pim-1 kinase. Further, it was difficult to reconstitute the pore into planar lipid bilayers, presumably due to the high net charge present near the *trans* mouth.

A set of charge neutralisation mutants of the *trans* mouth residues D127, D128 and K131 as homoheptamers, and as heteroheptamers in conjunction with the PLM modification, were generated and characterised with the aim of increasing the propensity to insert into planar lipid bilayers and to facili-

tate the observation of Pim kinase binding. It was found that pore insertion was more favourable the more charges were neutralised at the *trans* mouth. Further, the (αHL-D127N-PLM-D8)$_1$(αHL-D127N)$_6$ heteroheptamer exhibited detectable binding events in the presence of Pim-1, observed as the stochastic abolition of fast blockade events that presumably arise due to the conformational dynamics of the mutant *trans* loop.

Another stage of optimisation was then attempted, whereby the serine/glycine linkers between αHL and the *trans* loop-inserted Pimtide sequence were reduced in length, in an attempt to draw the bound kinase closer to the *trans* mouth entrance and hence induce a current blockade. While the linker lengths could be reduced by half without affecting kinase binding, reduction by three-quarters abolished kinase binding. Further, no current blockade was observed with these shorter linker variants. This strategy was hence not pursued further.

A novel engineering strategy was then developed to produce pores that bear a genetically encoded peptide at the *trans* mouth of the pore, but where the peptide is only attached to the *trans* mouth by a single terminus, rather than being constrained within the *trans* loop. Hence, this form of pore resembles the previously published chemically modified protein kinase A sensor in topology. However, this new method retains the benefits of a genetically encoded sensor strategy, such as being quicker and simpler to produce, being able to easily produce combinatorial libraries, and with 'quantitative' yields of modified pore.

Pores were designed with TEV protease cleavage sites at both the N- and C-terminal ends of the inserted Pimtide sequence. It was shown that αHL

monomers thus modified could be cleaved by TEV protease before being co-oligomerised with αHL-D127N monomers to form heteroheptamers.

The heteroheptamer (αHL-TEV-PLM-D127N-D8)$_1$(αHL-D127N)$_6$ exhibited noisy, featureless channels that did not produce a detectable kinase binding signal. This may be due to obstruction of the Pimtide sequence by the adjacent peptide formed from the glycine/serine linker and TEV protease cleavage site that before cleavage were to the N-terminal side of the Pimtide peptide. In the (αHL-PLM-TEV-D127N-D8)$_1$(αHL-D127N)$_6$ heteroheptamer, the Pimtide-containing peptide is unencumbered by the adjacent peptide arising from the serine/glycine linker as the adjacent peptide is shorter due to the TEV protease recognition site residues remaining fused to the Pimtide sequence.

This heteroheptamer exhibited gating events that differed markedly depending on the polarity of the applied potential. Further, the binding of Pim-1 was clearly apparent at negative potentials. This allowed the characterisation of the kinetics of binding by titration of Pim-1 over the concentration range 50–400 nM.

Finally, it was possible to show that when MgATP was added to the *trans* chamber, phosphorylation of the pore was observed, presenting as the abolishment of Pim-1 binding events. While this is a one-shot experiment, i.e. one occurrence of phosphorylation per independent experiment with a single pore, it should be possible to design more sophisticated experiments to study catalytic turnover of the Pimtide substrate.

2.4 Future developments

There are two main directions of inquiry that build upon the results described here. In the first, phosphorylation and dephosphorylation of a single αHL-PLM-TEV pore might be studied. For example, by adding a protein phosphatase to the *trans* chamber along with Pim-1, it might be possible to observe turnover of the Pimtide substrate molecule between the kinase and phosphatase, forming a catalytic cycle.

One such phosphatase is lambda protein phosphatase (λPP), from lambda phage. Although prokaryotic, it possesses a high degree of sequence similarity with the catalytic subunits of a variety of protein phosphatases from the PP1, PP2A and PP2B groups [47, 48]. Further, and rather unusually, it has an absolute requirement for Mn^{2+} for activity. This could be useful in these experiments as the association of λPP with phosphorylated αHL-PLM-TEV could first be studied in the absence of Mn^{2+}, without completing the catalytic cycle as the phosphatase would not be able to dephosphorylate the pore. Subsequent addition of Mn^{2+} would complete the cycle. Provided the kinase and phosphatase have distinct dwell times (which would seem a reasonable assumption), it will be possible to extract phosphorylation and dephosphorylation rate constants by performing enzyme titrations and using hidden Markov analysis methods.

In the second direction, pores would be developed for the screening and study of kinase inhibitors. In previous studies with protein kinase A, a synergistic effect between the binding of MgATP and the PKIP loop of αHL-PKIP to protein kinase A was demonstrated [39, 40]. A similar effect is expected

for the Pim kinases; it would be necessary to engineer a pseudo-substrate equivalent of Pimtide to prevent phosphorylation, which would be most simply achieved by mutating the serine phosphorylation site to an alanine. It should hence be possible to measure the inhibitory effect of ATP-competitive kinase inhibitors indirectly via their reduction of this synergistic effect. By titrating an inhibitor it should be possible to calculate the inhibitory constant, K_i.

A further progression would then be to integrate these pores into a droplet interface biliayer (DIB) system where an array of droplets could be rapidly screened [49, 50]. These droplets could perhaps contain different kinases to investigate specificity, or may simply contain members of compound libraries for inhibitor discovery.

2.5 Materials and methods

Chemicals and oligonucleotides were from Sigma Aldrich unless otherwise stated. Enzymes and DNA ladders were from New England Biolabs. Cloning was performed in XL-10 Gold cells (Stratagene) that were made chemically competent by the method of Chung *et al.*, [51].

Routine plasmid purification was performed using either Qiaprep spin miniprep or Qiagen Plasmid Maxi kits as directed by the accompanying protocols. DNA quantitation was performed by UV-visible absorption spectroscopy using a NanoDrop 1000 (Thermo Scientific). Water used was purified by reverse osmosis to a final resistivity of 18.2 MΩ cm (Millipore).

2.5.1 Plasmids

pT7-αHL was a gift from Stephen Cheley [52]. All plasmid insert sequences were verified by sequencing (Source BioScience, Oxford).

pT7-αHL-PLM-D8

This plasmid was prepared by Stephen Cheley. A DNA cassette containing the Pimtide sequence, flanking glycine/serine linkers, and AgeI and StuI sites at the 5' and 3' ends respectively, was prepared by assembly polymerase chain reaction (PCR) using the following oligonucleotides: 5'-GCGC GCACCGGTGATGATAG-3', 5'-GCCGCTACTGCCGCTTCCGCTGCT ATCATCACCGGTGCGC-3', 5'-AGCGGCAGTAGCGGCGCACGTAAA CGTCGTCGTCATCCGA-3', 5'-GCTACCCGCGGTCGGTGGGCCGCT CGGATGACGACGACGT-3', 5'-CCGACCGCGGGTAGCTCCGGGAGC GGCTCTAGCAAAATTG-3', and 5'-CCCGGGAGGCCTCCAATTTTGC TAGAGCCGCTC-3'. The cassette was gel purified and then digested with AgeI and StuI, before a further gel purification step. The digested cassette was then ligated into pT7-αHL-RL4-D8 (encoding the wild type αHL gene remodelled to contain silent restriction sites and an octoaspartate tag, as detailed in [39]) that had also been digested with AgeI and StuI. The ligation reaction was transformed into chemically competent XL-10 gold cells, which were subsequently plated on LB carbenicillin plates and incubated at 37°C overnight. Colonies were picked and DNA was isolated and sequenced (Source BioScience, Oxford) to confirm successful mutagenesis. Large scale DNA isolation by plasmid maxi-prep was then performed, and the isolated

DNA was stored at a concentration of 400 ng/μl at -20°C.

pT7-αHL-D127N-D128N-PLM-D8

This plasmid was prepared by homologous recombination by Stephen Cheley using pT7-αHL-PLM-D8 as the template.

pT7-αHL-D127N-D128N

pT7-αHL-D127N-D128N was generated by *in vivo* homologous recombination [53, 54]. Two PCR reactions were performed using pT7-αHL as template. For the first reaction, the template was first linearised by NdeI before PCR using the forward mutagenic primer: 5'-ACTGGTAACAACACAGG AAAAATTGGCGG-3' and the reverse non-mutagenic primer (SC47): 5'-C AGAAGTGGTCCTGCAACTTTAT-3'. The second PCR used template linearised by HindIII; the reverse mutagenic primer was 5'-CCGCCAATTT TTCCTGTGTTGTTACCAGT-3' and the forward non-mutagenic primer (SC46) was 5'-ATAAAGTTGCAGGACCACTTCTG-3'. PCR was performed using Phusion Flash HF Master Mix (New England Biolabs) using cycling conditions as directed by the manufacturer. Aliquots of the completed reactions were analysed by electrophoresis on a 0.8% TAE agarose gel to confirm successful amplification. Equal volumes (typically 5 μl) were then mixed and transformed into chemically competent XL-10 Gold cells, which were subsequently plated on LB carbenicillin plates and incubated at 37°C overnight. Colonies were picked and DNA was isolated and sequenced (Source BioScience, Oxford) to confirm successful mutagenesis. Large scale DNA isolation by plasmid maxi-prep was then performed, and the isolated

DNA was stored at a concentration of 400 ng/μl at -20°C.

pT7-αHL-D127N-D128N-K131(Q/A)-PLM-D8

pT7-αHL-D127N-D128N-K131Q-PLM-D8 was generated by *in vivo* homologous recombination (as above) using pT7-αHL-D127N-D128N as the template. The forward mutagenic primer was 5'-GCTCTAGCCAAATTGGA GGCCTTATTG-3' and the reverse mutagenic primer was 5'-CAATAAGG CCTCCAATTTGGCTAGAGC-3'. The non-mutagenic primers were SC47 and SC46 as above.

pT7-αHL-D127N-D128N-K131A-PLM-D8 was generated by *in vivo* homologous recombination (as above) using pT7-αHL-D127N-D128N-PLM-D8 as the template. The forward mutagenic primer was 5'-GGGAGCGGCTC TAGCGCGATTGGAGGC-3' and the reverse mutagenic primer was 5'-G CCTCCAATCGCGCTAGAGCCGCTCCC-3'. The non-mutagenic primers were SC47 and SC46 as above.

pT7-αHL-D127N-D128N-K131(Q/A)

pT7-αHL-D127N-D128N-K131Q was a gift from Ellina Mikhailova. pT7-αHL-D127N-D128N-K131A was generated by *in vivo* homologous recombination (as above) using pT7-αHL-D127N-D128N-K131Q as the template. The forward mutagenic primer was 5'-GGTAATAATACAGGAGCGATTG GCGGC-3' and the reverse mutagenic primer was 5'-GCCGCCAATCGCTC CTGTATTATTACC-3'. The non-mutagenic primers were SC47 and SC46 as above.

pT7-αHL-D127N, pT7-αHL-D128N

These plasmids were both generated from pT7-αHL by *in vivo* homologous recombination (as above). For pT7-αHL-D127N the forward mutagenic primer was 5'-GTTACTGGTAATGATACAGGAAAAATTGGC-3' and the reverse mutagenic primer was 5'-GCCAATTTTTCCTGTATCATTACCA GTAAC-3'.

For pT7-αHL-D128N, the forward mutagenic primer was 5'-GTTACTG GTGATAATACAGGAAAAATTGGC-3' and the reverse mutagenic primer was 5'-GCCAATTTTTCCTGTATTATCACCAGTAAC-3'. In both cases, the non-mutagenic primers were SC47 and SC46 as above.

pT7-αHL-D127N-PLM-D8, pT7-αHL-D128N-PLM-D8

These plasmids were both generated from pT7-αHL-PLM-D8 by *in vivo* homologous recombination (as above). For pT7-αHL-D127N-PLM-D8 the forward mutagenic primer was 5'-GTCACCGGTAATGATAGCAGCGGAAG C-3' and the mutagenic primer was 5'-GCTTCCGCTGCTATCATTACCG GTGAC-3'.

For pT7-αHL-D128N-PLM-D8 the forward mutagenic primer was 5'-G TCACCGGTGATAATAGCAGCGGAAGC-3' and the reverse mutagenic primer was 5'-GCTTCCGCTGCTATTATCACCGGTGAC. In both cases the non-mutagenic primers were SC47 and SC46 as above.

pT7-αHL-D127N-PLM-D8 linker length variants

These were generated by *in vivo* homologous recombination. The template for the 'insert' PCR reactions was pT7-αHL-D127N-D128N-PLM, linearised by NdeI. The template for the 'vector' PCR reactions was uncut pT7-αHL-D127N-D128N-PLM. For each mutant, one 'insert' PCR and one 'vector' PCR reaction was performed and the two products recombined.

For SG4, the insert primers were 5'-AATGATAGCAGCGGAAGCGCACGTAAACGTCGTCGTC-3' and 5'-AATTTTGCTAGAGCCGCTCGCGGTCGGTGGGCCGCT-3'. The vector primers were 5'-AGCGGCTCTAGCAAAATTGGAGGCCTT-3' and 5'-GCTTCCGCTGCTATCATTACCGGTGAC-3'.

For SG2, the insert primers were 5'-ACCGGTAATGATAGCAGCGCACGTAAACGTCGTCGTC-3' and 5'-GCCTCCAATTTTGCTAGACGCGGTCGGTGGGCCGCT-3'. The vector primers were 5'-TCTAGCAAAATTGGAGGCCTTATTGGT-3' and 5'-GCTGCTATCATTACCGGTGACATTACC-3'.

pT7-αHL-D127N-TEV-PLM-D8, pT7-αHL-D127N-PLM-TEV-D8

These were generated from pT7-αHL-D127N-PLM-D8 by *in vivo* homologous recombination. For pT7-αHL-D127N-TEV-PLM-D8, the forward mutagenic primer was 5'-GAAAATCTGTATTTTCAAGGGGCACGTAAACGTCGTCGT-3' and the reverse mutagenic primer was 5'-CCCTTGAAAATACAGATTTTCGCCGCTACTGCCGCTTCC-3'.

For pT7-αHL-D127N-PLM-TEV-D8, the forward mutagenic primer was

5'-GAAAATCTGTATTTTCAAGGGGGTAGCTCCGGGAGCGGC-3' and the reverse mutagenic primer was 5'-CCCTTGAAAATACAGATTTTCCG CGGTCGGTGGGCCGCT-3'. In both cases the non-mutagenic primers were SC47 and SC46.

2.5.2 Protein preparation

Pim kinases were a gift from Stefan Knapp and colleagues at the Structural Genomics Consortium, Oxford. They were expressed and purified recombinantly from *E. coli* as described [27]. The full-length human Pim-1 (36 kDa isoform, gi 33304198) was supplied in 50 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) and 5% (v/v) glycerol. Pim-1 was expressed as a fusion with a hexa-histidine tag, which was subsequently removed by TEV protease cleavage during purification. The concentration was determined by absorbance measurement at 280 nm to be 3.4 mg/ml, using a molecular weight of 35,545 Da and an extinction coefficient of 48,930 M^{-1} cm^{-1}. The protein identity was verified in our laboratory by mass spectrometry (M^+ calculated: 35545, found: 35540).

αHL proteins were produced using the S30 T7 High-Yield Protein Expression System (Promega), which is a coupled *in vitro* transcription-translation (IVTT) system. A standard reaction comprises: 500 ng plasmid DNA, 10 μl S30 Premix Plus (supplied with kit), 9 μl S30 extract (supplied with kit), 1 μl of L-[^{35}S]methionine (1175 Ci/mmol, 10 mCi/ml, MP Biomedicals), and sufficient nuclease-free water to make a total reaction volume of 25 μl. The S30 extract was pre-treated with rifampicin (1 μg/ml final) before addition to the

IVTT reaction, to suppress expression from endogenous non-T7 promoters in the extract. IVTT reaction mixtures were incubated at 37°C with shaking at 1200 rpm in a Thermomixer Comfort (Eppendorf) for one hour. Specific variations of this standard protocol are described in the following sections. $(\alpha$HL-PLM-D8$)_1(\alpha$HL$)_6$ and $(\alpha$HL-PLM-D127N-D128N-D8$)_1(\alpha$HL-D127N-D128N$)_6$ were made by Stephen Cheley.

Homoheptamers

Alpha-haemolysin mutant homoheptamers were produced in 25 μl reactions as above. These 25 μl reactions were supplemented with 3 μl of rabbit erythrocyte membranes before incubation. After incubation, each reaction was centrifuged for ten minutes at 25,000$\times g$. The supernatant was removed and the pellet was resuspended in 200 μl MBSA buffer (10 mM MOPS, 150 mM NaCl, pH 7.4, 1 mg/ml bovine serum albumin). The suspension was then centrifuged and washed once more before being centrifuged a final time. The resulting pellet was resuspended in 50 μl of 1X Laemmli sample buffer (62.5 mM Tris.HCl pH 6.8, 2.3% w/v SDS, 5% v/v β-mercaptoethanol, 10% w/v glycerol) and the entire sample was electrophoresed in a 5% SDS-PAGE gel at 80 V overnight (approximately 18–20 hours).

Heteroheptamers

Alpha-haemolysin mutant heteroheptamers were produced in 100 μl reactions, scaled up from the quantities described above. Two different plasmids are added in an 6:2 ratio; the 'background' plasmid that does not contain the mutations used for stochastic sensing, is added at the higher concentration;

the sensing mutant, e.g. pT7-αHL-PLM, of which only one subunit is desired in the resulting heteroheptamer, is added at the lower concentration. The ratio was found empirically to give the optimum yield of heteroheptamers containing a single 'modified' subunit and six 'unmodified' subunits. The total quantity of both DNAs together was 500 ng in a 25 μl reaction. The reactions were then processed as described at the start of this section.

Purification by SDS-PAGE

After electrophoresis, the gels were dried on Whatman 3M filter paper under vacuum for 3–4 hours at room temperature. The dried gel was autoradiographed using Kodak BioMax MR film. By aligning the dried gel with the developed autoradiograph, the desired bands were excised from the gel with a scalpel. Homoheptamers appear as a single band. Heteroheptamer reactions produce a ladder of bands because the 'modified' subunit carries a D8 tag that generates a downwards (i.e. faster migrating) gel-shift. This gel-shift is proportional to the number of modified subunits present. A homoheptamer lane is usually run alongside which allows confirmation of the top band in the heteroheptamer lane as the 7:0 heptamer. The desired band, with 6:1 stoichiometry, is the next fastest migrating band (e.g. Figure 2.11).

Each excised gel slice was then rehydrated in 500 μl of TE buffer (10 mM Tris.HCl pH 8.0, 1 mM EDTA) for one hour at room temperature. The backing filter paper was then removed with forceps and the rehydrated gel macerated with a pestle. The resulting mixture was filtered through a 0.2 μm cellulose acetate filter (Rainin Microfilterfuge tube). The filtrate was aliquoted (typically 10 μl portions) and stored at -80°C.

TEV protease-cleaved pores

For optimisation of TEV cleavage, a 25 μl IVTT reaction, as above, containing pT7-αHL-D127N-TEV-PLM-D8 DNA was incubated for one hour at 37°C. Translation was then stopped by the addition of chloramphenicol to a final concentration of 34 μg/ml. 2 μl of the stopped reaction was saved as the un-cleaved control sample. Then 10 units of AcTEV protease (Invitrogen) were added to the stopped reaction and incubated at 4°C for four hours. After 1, 2, 3 and 4 hours, 2 μl samples were retained. These samples plus the control were then electrophoresed in a pre-cast 10% bis-tris acrylamide gel (Bio-Rad) in XT-MOPS buffer for one hour at 200 V. The gel was dried under vacuum and autoradiographed.

For the preparation of TEV-cleaved heteroheptamers, two IVTT reactions were first prepared. First, a 75 μl reaction containing pT7-αHL-D127N DNA was prepared, scaled up from the quantities described at the start of this section. Second, a 25 μl reaction containing pT7-αHL-D127N-TEV-PLM-D8 or pT7-αHL-D127N-PLM-TEV-D8 was prepared. After incubation as described above, translation was stopped by the addition of chloramphenicol to a final concentration of 34 μg/ml. The αHL-D127N reaction was stored on ice until needed.

To the αHL-D127N-TEV-PLM-D8 or αHL-D127N-PLM-TEV-D8 reaction, 10 units of AcTEV protease were added. The reaction was then incubated at 4°C for 4 hours. The TEV protease-digested reaction was then mixed with the αHL-D127N reaction, followed by the addition of 5 μl of rabbit erythrocyte membranes. The resulting mixture was incubated for one

hour at room temperature. It was then centrifuged, washed with MBSA and purified by SDS-PAGE as previously described.

Analysis of heteromer composition

To confirm the composition and stoichiometry of (αHL-PLM-TEV-D127N-D8)$_1$(αHL-D127N)$_6$, 200 μl of gel-purified heteromer was concentrated to a volume of 40 μl in a Microcon YM-50 (50 kDa MWCO, Millipore) centrifugal concentrator. The concentrated sample was then divided into two 20 μl aliquots, and 5 μl of 5X Laemmli sample buffer was added to each. One of the two samples was then boiled at 95°C for 20 minutes. Both samples were then analysed by SDS-PAGE on a 12% bis-tris gel (Bio-Rad) in XT-MOPS buffer. The gel was then dried and analysed by autoradiography. Gel densitometry was performed using Quantity One software (Bio-Rad).

2.5.3 Electrophysiology

Single-channel recording was performed using the folded planar lipid bilayer method of Montal and Mueller [55]. A 25 μm thick Teflon film containing an aperture of 70 μm diameter was fastened between two machined Delrin chambers (see Figure 1.4 in Chapter 1). Prior to bilayer formation, the aperture was pre-treated with 1 μl of 1% (v/v) hexadecane in pentane on each side. Each chamber was then filled with 1 ml of filter-sterile buffer (composition as described in Results). Where DTT was used, this was added to the chamber at the start of the experiment from a 1 M stock solution that had been titrated to pH 6.8 with KOH. 5 μl of a 10 mg/ml solution of

1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC) in pentane was then dropped onto the buffer surface of each chamber. After five minutes, lipid bilayers were folded across the teflon aperture by the repeated raising and lowering of the buffer level below and above the aperture in each chamber by pipetting.

Electrical currents were measured using two Ag/AgCl electrodes, prepared by treating silver wire (1.5 mm diameter) with 30% sodium hypochlorite solution [56]. Where DTT was present in the chamber solution, 3 M KCl agarose bridges were used between the electrodes and chamber solution. The electrodes were connected to the headstage of a patch-clamp amplifier (Axopatch 200B, Molecular Devices) operating in voltage-clamp mode. The *cis* chamber was defined as the chamber connected to the grounded electrode and to which αHL protein was added. The *trans* chamber was connected to the working electrode. Current signals were filtered with a hardware (within the Axopatch 200B) low-pass Bessel filter (80 dB/decade) with a corner frequency of 2 kHz. Signals from the amplifier were digitised using a Digidata 1320A digitiser (Molecular Devices), which was connected to a computer running the pCLAMP software suite (Molecular Devices). Current signals were sampled with a frequency of 10 kHz. The assembly of chamber, electrodes, and headstage was contained within a Faraday cage.

Bilayer size was checked by the application of a triangular voltage wave ($\sim$20 Hz) across the bilayer which produces a square wave current output (as for a capacitor: $I = C \times \mathrm{d}V/\mathrm{d}t$). The input waveform was calibrated using a 100 pF model bilayer cell (MCB-1U, Molecular Devices) so that a current of 100 pA corresponds to a bilayer capacitance of 100 pF. Bilayers were checked

before experiments were begun to ensure the bilayer was of the expected size and did not leak current.

Once a satisfactory bilayer had been formed, αHL protein (typically $\sim$1 ng/μl) was added to the *cis* chamber. Typically, a volume of 0.5 μl was added. Voltage was then applied (typically +150 mV) and the *cis* chamber stirred with a magnetic stirrer flea until the insertion of a pore into the bilayer was observed as a stepwise increase in current.

Analytes and cofactors were introduced by pipetting into the *trans* chamber as required. After introduction, the *trans* chamber was stirred for 30 s with a magnetic stirrer flea to ensure thorough mixing before resumption of measurements. In the case of Pim kinases, additional pre-mixing was performed by removing 200 μl of buffer from the *trans* chamber and mixing it with the desired quantity of Pim kinase by pipetting in a microcentrifuge tube. The entire volume was then returned to the *trans* chamber, whereupon the contents were mixed for 30 s as above before resumption of measurements.

Analysis

Single-channel data were analysed using QuB (www.qub.buffalo.edu). Current traces were idealised using the segmental k-means (SKM) hidden-Markov algorithm of QuB [57] using models as described in the Results section. No additional filtering or processing of the current traces was performed before idealisation. Dwell time analysis and rate constant estimation was performed using the maximum interval likelihood (MIL) algorithm of QuB, with retrospective application of a dead time of 300 μs for approximate correction of missed events due to low-pass filtering [58]. Histogram data was typically

extracted from QuB and plotted in OriginPro (OriginLab) for presentation.

Bibliography

[1] Cuypers, H. T., Selten, G., Quint, W., Zijlstra, M., Maandag, E. R., Boelens, W., van Wezenbeek, P., Melief, C., and Berns, A. (1984) Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. *Cell 37*, 141–50.

[2] van Lohuizen, M., Verbeek, S., Krimpenfort, P., Domen, J., Saris, C., Radaszkiewicz, T., and Berns, A. (1989) Predisposition to lymphomagenesis in pim-1 transgenic mice: cooperation with c-myc and N-myc in murine leukemia virus-induced tumors. *Cell 56*, 673–82.

[3] Verbeek, S., van Lohuizen, M., van der Valk, M., Domen, J., Kraal, G., and Berns, A. (1991) Mice bearing the E mu-myc and E mu-pim-1 transgenes develop pre-B-cell leukemia prenatally. *Mol Cell Biol 11*, 1176–9.

[4] Selten, G., Cuypers, H. T., Boelens, W., Robanus-Maandag, E., Verbeek, J., Domen, J., van Beveren, C., and Berns, A. (1986) The primary structure of the putative oncogene pim-1 shows extensive homology with protein kinases. *Cell 46*, 603–11.

[5] Hoover, D., Friedmann, M., Reeves, R., and Magnuson, N. S. (1991) Recombinant human pim-1 protein exhibits serine/threonine kinase activity. *J Biol Chem 266*, 14018–23.

[6] Padma, R. and Nagarajan, L. (1991) The human PIM-1 gene product is a protein serine kinase. *Cancer Res 51*, 2486–9.

[7] Breuer, M. L., Cuypers, H. T., and Berns, A. (1989) Evidence for the involvement of pim-2, a new common proviral insertion site, in progression of lymphomas. *EMBO J 8*, 743–8.

[8] van der Lugt, N. M., Domen, J., Verhoeven, E., Linders, K., van der Gulden, H., Allen, J., and Berns, A. (1995) Proviral tagging in E mu-myc transgenic mice lacking the Pim-1 proto-oncogene leads to compensatory activation of Pim-2. *EMBO J 14*, 2536–44.

[9] Feldman, J. D., Vician, L., Crispino, M., Tocco, G., Marcheselli, V. L., Bazan, N. G., Baudry, M., and Herschman, H. R. (1998) KID-1, a protein kinase induced by depolarization in brain. *J Biol Chem 273*, 16535–43.

[10] Mikkers, H., Nawijn, M., Allen, J., Brouwers, C., Verhoeven, E., Jonkers, J., and Berns, A. (2004) Mice deficient for all PIM kinases display reduced body size and impaired responses to hematopoietic growth factors. *Mol Cell Biol 24*, 6104–15.

[11] Koike, N., Maita, H., Taira, T., Ariga, H., and Iguchi-Ariga, S. M. (2000) Identification of heterochromatin protein 1 (HP1) as a phosphorylation target by Pim-1 kinase and the effect of phosphorylation on the transcriptional repression function of HP1(1). *FEBS Lett 467*, 17–21.

[12] Leverson, J. D., Koskinen, P. J., Orrico, F. C., Rainio, E. M., Jalkanen, K. J., Dash, A. B., Eisenman, R. N., and Ness, S. A. (1998) Pim-1 kinase and p100 cooperate to enhance c-Myb activity. *Mol Cell 2*, 417–25.

[13] Rainio, E.-M., Sandholm, J., and Koskinen, P. J. (2002) Cutting edge: Transcriptional activity of NFATc1 is enhanced by the Pim-1 kinase. *J Immunol 168*, 1524–7.

[14] Winn, L. M., Lei, W., and Ness, S. A. (2003) Pim-1 phosphorylates the DNA binding domain of c-Myb. *Cell Cycle 2*, 258–62.

[15] Bachmann, M., Hennemann, H., Xing, P. X., Hoffmann, I., and Möröy, T. (2004) The oncogenic serine/threonine kinase Pim-1 phosphorylates and inhibits the activity of Cdc25C-associated kinase 1 (C-TAK1): a novel role for Pim-1 at the G2/M cell cycle checkpoint. *J Biol Chem 279*, 48319–28.

[16] Mochizuki, T., Kitanaka, C., Noguchi, K., Muramatsu, T., Asai, A., and Kuchino, Y. (1999) Physical and functional interactions between Pim-1 kinase and Cdc25A phosphatase. Implications for the Pim-1-mediated activation of the c-Myc signaling pathway. *J Biol Chem 274*, 18659–66.

[17] Wang, Z., Bhattacharya, N., Mixter, P. F., Wei, W., Sedivy, J., and Magnuson, N. S. (2002) Phosphorylation of the cell cycle inhibitor p21Cip1/WAF1 by Pim-1 kinase. *Biochim Biophys Acta 1593*, 45–55.

[18] Chen, X. P., Losman, J. A., Cowan, S., Donahue, E., Fay, S., Vuong, B. Q., Nawijn, M. C., Capece, D., Cohan, V. L., and Rothman, P. (2002) Pim serine/threonine kinases regulate the stability of Socs-1 protein. *Proc Natl Acad Sci U S A 99*, 2175–80.

[19] Fox, C. J., Hammerman, P. S., Cinalli, R. M., Master, S. R., Chodosh, L. A., and Thompson, C. B. (2003) The serine/threonine kinase Pim-2 is a transcriptionally regulated apoptotic inhibitor. *Genes Dev 17*, 1841–54.

[20] Aho, T. L. T., Sandholm, J., Peltola, K. J., Mankonen, H. P., Lilly, M., and Koskinen, P. J. (2004) Pim-1 kinase promotes inactivation of the pro-apoptotic Bad protein by phosphorylating it on the Ser112 gatekeeper site. *FEBS Lett 571*, 43–9.

[21] Yan, B., Zemskova, M., Holder, S., Chin, V., Kraft, A., Koskinen, P. J., and Lilly, M. (2003) The PIM-2 kinase phosphorylates BAD on serine 112 and reverses BAD-induced cell death. *J Biol Chem 278*, 45358–67.

[22] Brault, L., Gasser, C., Bracher, F., Huber, K., Knapp, S., and Schwaller, J. (2010) PIM serine/threonine kinases in the pathogenesis and therapy of hematologic malignancies and solid cancers. *Haematologica 95*, 1004–15.

[23] Amaravadi, R. and Thompson, C. B. (2005) The survival kinases Akt and Pim as potential pharmacological targets. *J Clin Invest 115*, 2618–24.

[24] Friedmann, M., Nissen, M. S., Hoover, D. S., Reeves, R., and Magnuson, N. S. (1992) Characterization of the proto-oncogene pim-1: kinase activity and substrate recognition sequence. *Arch Biochem Biophys 298*, 594–601.

[25] Palaty, C. K., Clark-Lewis, I., Leung, D., and Pelech, S. L. (1997) Phosphorylation site substrate specificity determinants for the Pim-1 protooncogene-encoded protein kinase. *Biochem Cell Biol 75*, 153–62.

[26] Hutti, J. E., Jarrell, E. T., Chang, J. D., Abbott, D. W., Storz, P., Toker, A., Cantley, L. C., and Turk, B. E. (2004) A rapid method for determining protein kinase phosphorylation specificity. *Nat Methods 1*, 27–9.

[27] Bullock, A. N., Debreczeni, J., Amos, A. L., Knapp, S., and Turk, B. E. (2005) Structure and substrate specificity of the Pim-1 kinase. *J Biol Chem 280*, 41675–82.

[28] Allen, J. D., Verhoeven, E., Domen, J., van der Valk, M., and Berns, A. (1997) Pim-2 transgene induces lymphoid tumors, exhibiting potent synergy with c-myc. *Oncogene 15*, 1133–41.

[29] Bachmann, M. and Möröy, T. (2005) The serine/threonine kinase Pim-1. *Int J Biochem Cell Biol 37*, 726–30.

[30] Wang, Z., Bhattacharya, N., Weaver, M., Petersen, K., Meyer, M., Gapter, L., and Magnuson, N. S. (2001) Pim-1: a serine/threonine kinase with a role in cell survival, proliferation, differentiation and tumorigenesis. *J Vet Sci 2*, 167–79.

[31] Zhu, N., Ramirez, L. M., Lee, R. L., Magnuson, N. S., Bishop, G. A., and Gold, M. R. (2002) CD40 signaling in B cells regulates the expression of the Pim-1 kinase via the NF-kappa B pathway. *J Immunol 168*, 744–54.

[32] Saris, C. J., Domen, J., and Berns, A. (1991) The pim-1 oncogene encodes two related protein-serine/threonine kinases by alternative initiation at AUG and CUG. *EMBO J 10*, 655–64.

[33] Qian, K. C., Wang, L., Hickey, E. R., Studts, J., Barringer, K., Peng, C., Kronkaitis, A., Li, J., White, A., Mische, S., and Farmer, B. (2005) Structural basis of constitutive activity and a unique nucleotide binding mode of human Pim-1 kinase. *J Biol Chem 280*, 6130–7.

[34] Jacobs, M. D., Black, J., Futer, O., Swenson, L., Hare, B., Fleming, M., and Saxena, K. (2005) Pim-1 ligand-bound structures reveal the mechanism of serine/threonine kinase inhibition by LY294002. *J Biol Chem 280*, 13728–34.

[35] Kumar, A., Mandiyan, V., Suzuki, Y., Zhang, C., Rice, J., Tsai, J., Artis, D. R., Ibrahim, P., and Bremer, R. (2005) Crystal structures of proto-oncogene kinase Pim1: a target of aberrant somatic hypermutations in diffuse large cell lymphoma. *J Mol Biol 348*, 183–93.

[36] DeLano, W. (2002) *The PyMOL User's Manual.* (DeLano Scientific, Palo Alto, CA, USA).

[37] Pasqualucci, L., Neumeister, P., Goossens, T., Nanjangud, G., Chaganti, R. S., Küppers, R., and Dalla-Favera, R. (2001) Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature 412*, 341–6.

[38] Morwick, T. (2010) Pim kinase inhibitors: a survey of the patent literature. *Expert Opin Ther Pat 20*, 193–212.

[39] Cheley, S., Xie, H., and Bayley, H. (2006) A genetically encoded pore for the stochastic detection of a protein kinase. *Chembiochem 7*, 1923–1927.

[40] Xie, H., Braha, O., Gu, L.-Q., Cheley, S., and Bayley, H. (2005) Single-molecule observation of the catalytic subunit of cAMP-dependent protein kinase binding to an inhibitor peptide. *Chem Biol 12*, 109–120.

[41] Howorka, S., Cheley, S., and Bayley, H. (2001) Sequence-specific detection of individual DNA strands using engineered nanopores. *Nat Biotechnol 19*, 636–9.

[42] Zhuo, S. and Dixon, J. E. (1997) Effects of sulfhydryl regents on the activity of lambda Ser/Thr phosphoprotein phosphatase and inhibition of the enzyme by zinc ion. *Protein Eng 10*, 1445–52.

[43] Northrup, S. H. and Erickson, H. P. (1992) Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc Natl Acad Sci U S A 89*, 3338–42.

[44] Schreiber, G. and Fersht, A. R. (1996) Rapid, electrostatically assisted association of proteins. *Nat Struct Biol 3*, 427–31.

[45] Lew, J., Coruh, N., Tsigelny, I., Garrod, S., and Taylor, S. S. (1997) Synergistic binding of nucleotides and inhibitors to cAMP-dependent protein kinase examined by acrylodan fluorescence spectroscopy. *J Biol Chem 272*, 1507–13.

[46] Whitehouse, S. and Walsh, D. A. (1983) Mg X ATP2-dependent interaction of the inhibitor protein of the cAMP-dependent protein kinase with the catalytic subunit. *J Biol Chem 258*, 3682–92.

[47] Barik, S. (1993) Expression and biochemical properties of a protein serine/threonine phosphatase encoded by bacteriophage lambda. *Proc Natl Acad Sci U S A 90*, 10633–7.

[48] Zhuo, S., Clemens, J. C., Hakes, D. J., Barford, D., and Dixon, J. E. (1993) Expression, purification, crystallization, and biochemical characterization of a recombinant protein phosphatase. *J Biol Chem 268*, 17754–61.

[49] Bayley, H., Cronin, B., Heron, A., Holden, M. A., Hwang, W. L., Syeda, R., Thompson, J., and Wallace, M. (2008) Droplet interface bilayers. *Mol Biosyst 4*, 1191–208.

[50] Syeda, R., Holden, M. A., Hwang, W. L., and Bayley, H. (2008) Screening blockers against a potassium channel with a droplet interface bilayer array. *J Am Chem Soc 130*, 15543–8.

[51] Chung, C. T., Niemela, S. L., and Miller, R. H. (1989) One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A 86*, 2172–5.

[52] Walker, B., Krishnasastry, M., Zorn, L., Kasianowicz, J., and Bayley, H. (1992) Functional expression of the alpha-hemolysin of Staphylococcus aureus in intact Escherichia coli and in cell lysates. Deletion of five C-terminal amino acids selectively impairs hemolytic activity. *J Biol Chem 267*, 10902–10909.

[53] Bubeck, P., Winkler, M., and Bautsch, W. (1993) Rapid cloning by homologous recombination in vivo. *Nucleic Acids Res 21*, 3601–2.

[54] Jones, D. H. and Howard, B. H. (1991) A rapid method for recombination and site-specific mutagenesis by placing homologous ends on DNA using polymerase chain reaction. *Biotechniques 10*, 62–66.

[55] Montal, M. and Mueller, P. (1972) Formation of bimolecular membranes from lipid monolayers and a study of their electrical properties. *Proc Natl Acad Sci U S A 69*, 3561–6.

[56] Maglia, G., Heron, A. J., Stoddart, D., Japrung, D., and Bayley, H. (2010) Analysis of single nucleic acid molecules with protein nanopores. *Methods Enzymol 475*, 591–623.

[57] Qin, F. (2004) Restoration of single-channel currents using the segmental k-means method based on hidden Markov modeling. *Biophys J 86*, 1488–501.

[58] Qin, F., Auerbach, A., and Sachs, F. (1996) Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events. *Biophys J 70*, 264–80.

Chapter 3

Method development: OmpG for bacterial display

3.1 Introduction

The stochastic detection of protein analytes by single channel recording has previously been achieved using pores that are engineered with an attached small molecule [1, 2] or peptide ligand [3, 4]. Such an approach requires the prior knowledge of an existing high affinity ligand. It would therefore be advantageous if a method existed that allowed the rapid selection of a high affinity ligand that is already 'in place' on the pore that will subsequently be used for stochastic sensing, without the need for additional engineering steps. This would allow the selection of stochastic sensors against targets for which no suitable ligand exists, but which also have optimised affinity in the structural context of the pore 'scaffold'.

In this chapter, a scheme for such a method is proposed, in which outer

membrane protein G (OmpG) pores that bind a desired target are selected from a combinatorial peptide library inserted in an extramembranous loop of OmpG. The combinatorial OmpG library is displayed on the cell surface of *E. coli* and hence screened by the method of bacterial cell surface display. Further, the scheme (see Figure 3.1) includes a rapid method of characterising selected clones by single channel recording using droplet interface bilayers (DIBs).

In this section, the various possible methods used for peptide display are reviewed before the proposed methodology is described in detail. Then, there follows a description of the progress made towards implementing this scheme for the selection of stochastic sensors from a combinatorial library of OmpG porins.

3.1.1 Artificial protein binding peptides

Protein-protein interactions are crucial for most biological processes. The domains involved in protein recognition have evolved to give rise to hugely complex networks of interacting proteins that govern essentially all processes required for life. Sometimes novel molecular recognition is required very quickly, for example in the immune system when challenged by a foreign entity.

For this scenario there exist antibodies: Y-shaped glycoproteins comprising heavy and light chains that possess hypervariable regions that form paratopes (binding sites for specific epitopes). The human body maintains a huge diversity of antibodies arising from the hypervariable regions, ensuring

that when challenged by a foreign entity, there should be at least one motif that possesses a high affinity for an epitope on the foreign entity. Further, on recognition of an antigen, B cells begin to proliferate rapidly and undergo a process called somatic hypermutation, whereby the variable domains of antibodies undergo a high mutation rate. This further increases the diversity of potential recognition motifs. Those new hypermutated antibodies that possess higher affinity for the antigen are selected through competition by survival of the antibody-producing B cells. In this way, proteins with high affinity for a new target are rapidly produced.

This process has been hugely important for biotechnology and human health ever since Edward Jenner first tested an immunisation for cowpox as protection against smallpox [5]. The advent of monoclonal antibodies saw the ability to isolate identical antibodies specific to a given antigen [6] and the use of monoclonal antibodies for therapy.

Subsequent research has sought to produce 'artificial' antibodies that can be raised and matured using molecular display techniques. For example, the development of phage display methods for antibodies by Greg Winter, John McCafferty, Richard Lerner and Carlos Barbas [7, 8] paved the way for blockbuster biologic drugs such as adalimumab (Humira), which was the first fully human monoclonal antibody drug approved by the FDA [9, 10]. Adalimumab was derived from the candidate antibody D2E7, discovered during a phage display screen by Cambridge Antibody Technology and BASF Pharma [10]. Humanised antibodies are typically only 95% human as they are typically formed by splicing murine complementarity determining regions (CDRs) into human monoclonal antibodies. Hence, the use of phage display has enabled

the development of fully human therapeutic antibodies that should not have deleterious immunogenicity in patients.

In recent years, researchers have developed a number of binding proteins much smaller than antibodies or antibody fragments, based upon small protein domains that serve as scaffolds for combinatorial peptide libraries. Examples include 'Nanobodies', which are single-domain antibodies that are typically produced from camelid heavy-chain antibodies [11]; 'Monobodies', which are based on the fibronectin type III domain (FN3) [12, 13]; and 'Anticalins', based on lipocalin scaffolds [14].

Possible uses of these affinity agents are not limited to therapeutic applications; others include mechanistic and cellular investigations through the selective disruption of protein interaction domains [15]; as crystallisation chaperones [16]; and for understanding the fundamental principles of protein interactions [17].

Regardless of the scaffold used, whether it be a full antibody, Fab fragment or monobody, synthetic antibody methods all rely on molecular display techniques [18]. While the most common of these is phage display, several other methods have been developed, with varying levels of adoption. Crucially, they all maintain a physical link between phenotype and genotype. This ensures that once positively selected proteins are isolated from a library, their identity can be determined by sequencing the linked genotype. The key features of the main display technologies are summarised below.

Phage display

Phage display was originally developed by George P Smith [19]. In the original embodiment of the method, the peptide(s) of interest are displayed on the surface of filamentous phage through the fusion of the peptide sequence(s) to the N-terminus of the pIII or pVIII phage coat protein genes. All five capsid proteins have subsequently been utilised for phage display to varying extents. Copies of a fusion coat protein are incorporated into phage virions as they are released by the cell. The fused peptides are displayed in an accessible manner. The genotype-phenotype link is between the displayed protein and its encoding gene sequence present within the viral genome.

A phagemid system is commonly used for the construction and screening of phage display libraries [20] [21]. A phagemid is a plasmid that contains an origin of replication for double stranded replication as well as an f1 origin for single stranded replication and packaging into phage particles. The combinatorial DNA library is cloned into the phagemid such that it is expressed as a fusion with the desired coat protein.

This phagemid library is then transformed and maintained in *E. coli*. To generate the phage display library, *E. coli* harbouring library phagemids are infected with a helper phage. Typically, a packaging-deficient helper phage is used. Infection triggers production of wild type and fusion phage proteins. The library fusion coat protein will form a small proportion of the phage proteins contained within the cell ready for phage assembly. As phagemid DNA contains a functional packaging signal, it is preferentially used for virion assembly, with the coat proteins containing a proportion of the corresponding

fusion coat protein. The use of phagemids rather than phage vectors can be beneficial as they are smaller and easier to clone, usually allowing higher library sizes. Further, they are essential for certain experiments such as monovalent display [20].

Phage libraries are typically selected *in vitro* using target binding and elution methods. This selection is often called 'panning'. Due to the robust nature of the viral particles, many different selections have been developed [20]. In a typical protocol, the selection target is immobilised to the bottom of a microtiter plate or magnetic beads. The phage particles are then incubated with the immobilised target. Phage displaying proteins that bind the immobilised target remain bound, while non-binding phage are washed away. The positively selected phage are then eluted and amplified in bacteria. This cycle may be repeated multiple times to increase enrichment of binding motifs. After selection, the DNA contained within the positive clones can be analysed by sequencing.

Many variations and improvements have been proposed since the inception of phage display, including the use of alternative phage such as T7 [22]. The applications of phage display are myriad and expansive in scope [21, 23, 24]. Phage display has been shown to be a highly successful method for a diverse range of displayed proteins. In common with other cell-based methods, the library size is limited by the transformation efficiency of the host cells used for propagation. Certain proteins may be recalcitrant to display due to toxicity to *E. coli* or through interference with phage assembly. However, a wide range of protein types, sizes and even multi-subunit species (e.g. Fab fragments) have been successfully displayed.

Yeast display

Yeast display was first developed in the lab of K Dane Wittrup [25] [26]. In common with bacterial display, yeast display is a cell surface display method. Proteins are displayed as a C-terminal fusion with the Aga2p mating agglutinin protein. This fusion is encoded by a plasmid that places the fusion coding sequence under control of the GAL1,10 galactose inducible promoter. The Aga2p fusion protein is processed and secreted, whereupon it is tethered to the Aga1p protein by disulphide bonds. Aga1p is itself attached to the cell wall covalently by a phosphatidylinositol glycan tail. For the purpose of yeast display, a modified strain EBY100 is used where the Aga1p encoding gene has been placed in a single, galactose inducible open reading frame [25].

Once a combinatorial library is cloned into the host yeast cells, cell surface display is induced using standard procedures. Optimisation of induction conditions is usually required to optimise display levels. Displaying cells are then selected by fluorescence activated cell sorting (FACS) [27]. Epitope tags are often included in the fusion construct to allow detection of display independent of ligand binding. This can be used to ensure the full fusion protein has been expressed and to normalise ligand binding-related fluorescence to the quantity of protein displayed on each cell surface.

A significant limitation of yeast surface display is that library size is limited by the transformation efficiency of yeast cells, which is much lower than bacteria ($\sim 10^6$ transformants/μg versus $\sim 10^9$ transformants/μg for *E. coli*). However, yeast display also possesses several advantages. First, it is a eukaryotic system, so is ideal for displaying proteins that require eukary-

otic post-translational modification and/or efficient disulphide isomerisation. However, yeast glycosylation patterns do differ compared to other eukaryotes. Second, quantitative screening of the library is possible due to the use of FACS.

Ribosome display

Ribosome display was developed in the lab of Andreas Plückthun [28], although the underlying technique was first demonstrated by Mattheakis *et al.*, [29]. In common with mRNA display, ribosome display differs from other display techniques in that it does not depend on living cells for the genotype-phenotype link or production of phage, and is therefore a true *in vitro* technique.

In ribosome display, a DNA library is amplified by PCR, introducing a T7-promoter, ribosome binding site, stabilising stem loops and a spacer sequence, before being transcribed into mRNA. The mRNA is then translated in a modified *E. coli* S30 system under conditions which stabilise mRNA-ribosome complexes/polysomes, such as a high concentration of magnesium ions. The mRNA used does not contain stop codons, and hence a proportion of mRNA molecules will remain attached to the translating ribosome. The spacer encoded in the mRNA allows the displayed protein to fold unhindered, and hence the folded, displayed protein is available for affinity selection [28].

Selection is typically by way of incubating the stabilised ribosome complexes with immobilised target antigen. Non-binding complexes are washed away and positively selected complexes dissociated with EDTA or eluted with antigen. The positively selected mRNA is then recovered and reverse

transcribed. The resultant cDNA can be amplified by PCR and subsequently subjected to further cycles of enrichment [28].

As ribosome display is a fully *in vitro* technique, the library size is not constrained by the transformation efficiency of host organisms and is instead limited by the number of mRNA molecules. For example, 10 μg of mRNA provides $\sim 1.5 \times 10^{13}$ molecules, allowing extremely large libraries to be screened. Further, the technique is unaffected by other host cell constraints such as toxicity of the displayed protein. However, ribosome display is limited in that the displayed protein must be capable of folding correctly in the prokaryotic *in vitro* translation system.

mRNA display

mRNA display was first developed in the lab of Jack Szostak [30]. mRNA is analogous to ribosome display in that the genotype-phenotype link is between mRNA and the translated protein. However, in mRNA display, the species that is screened is mRNA covalently attached directly to translated protein, whereas in ribosome display the screened species is mRNA and translated protein held loosely and non-covalently in complex with the ribosome [30].

In the mRNA display procedure, a DNA library is transcribed into mRNA. This is then ligated to a 3'-puromycin oligonucleotide by splinted ligation. Translation of the resulting mRNA-DNA-puromycin fusion is then performed in a rabbit reticulocyte system. The puromycin moiety stalls translation. Under optimised conditions, the puromycin slowly finds its way to the peptidyl transferase site of the ribosome, whereupon it is covalently joined to the nascent polypeptide, covalently linking the mRNA and protein together.

The fusions are then isolated, for example using poly-dT resin if a poly-dA linker is used in the mRNA-protein fusion [31].

Before selection, the cDNA/mRNA-protein hybrid fusion is usually generated by reverse transcription. This is to remove secondary and tertiary structure from the mRNA that could interfere or compete for binding. The final fusion protein can then be screened in a similar fashion to ribosome display, against an immobilised target, and the encoding DNA/RNA is amplified in a similar manner by PCR [31].

mRNA display possesses many similar advantages to ribosome display. Additionally, the covalent linkage in mRNA display enables more stringent selection conditions. The absence of a very large ribosome complex during selection may also avoid possible steric effects caused by ribosomal presence.

Bacterial display

Bacterial display was first developed in the lab of George Georgiou [32]. In common with yeast display, it is a cell surface display method. As typical transformation efficiencies for bacteria such as *E. coli* are several orders of magnitude higher than yeast, it is possible to screen much larger libraries by bacterial display than yeast display. The largest published bacterial display library contained 5×10^{10} clones [33].

A variety of methods for bacterial surface display have been demonstrated in the literature [34, 35]. A general bacterial display system will consist of a peptide/protein fused to a 'carrier' protein that targets the outer membrane (gram negative bacteria) and thus display the fused 'passenger' protein. The choice of fusion will largely be dependent on the carrier protein's membrane

topology, i.e. whether either or both of the termini are extracellular facing. Alternatively, the passenger protein may be an insertional fusion, for example as an insertion in the extracellular loop of a porin.

Although bacterial display methods display the passenger protein on the cell surface, it is also possible to 'display' proteins periplasmically [36]. These are accessible for screening purposes under conditions where the outer membrane is permeable to larger molecules (up to 10 kDa) [36].

As bacterial display libraries are much larger than yeast libraries, they cannot be directly screened by FACS [27], as even the fastest cell sorters cannot screen so many cells in a practical time. Instead, one or more rounds of pre-enrichment are performed, typically by magnetic activated cell sorting (MACS), whereby the library displaying cells are incubated with magnetic bead-immobilised target. The beads are then washed and placed in culture medium whereupon the bound cells divide. Thus, selected clones are amplified for further selection steps [33, 37].

Several different bacterial species have been investigated for surface display [34, 38]; however *E. coli* remains the most investigated [39]. Many *E. coli* display methods have focused on the display and selection of peptides [33, 38], for example for epitope mapping, but it is possible to display larger proteins such as antibody fragments [32]. However, bacterial display is limited in that it is difficult to display complex eukaryotic proteins correctly due to the lack of eukaryotic post-translation modification.

It is possible to use a number of outer membrane proteins in *E. coli* for cell surface display. One system, the FliTrx system [40], is available commercially from Invitrogen. In the FliTrx system, the *E. coli* thioredoxin gene is inserted

into a redundant section of the flagellin gene. This fusion is assembled and displayed as partially functional flagella on the cell surface. Peptide libraries are cloned into the thioredoxin active-site loop in the fusion protein. The resulting display library is screened by panning.

In another system, a peptide library is displayed as an insertional fusion in an extracellular loop of outer membrane protein A (OmpA) [33]. This fusion is presented on the outer membrane and the inserted peptides are thus displayed in constrained loops on the cell surface. The library is selected by MACS and FACS over several rounds of selection.

Analogous systems using circularly permuted OmpX allow the extracellular display of peptides as N- or C-terminal fusions, reducing bias imposed by the constrained loop conformation [41, 42]. This bi-terminal display scaffold also enables more sophisticated selection schemes such as epitope tagging for display normalisation during FACS. Other carrier proteins for bacterial display in *E. coli* include the general porins, ice nucleation protein, autotransporters and S-layer proteins [34]. Some carrier proteins may have undesirable physiological effects on the host cells under certain conditions.

3.1.2 Proposed methodology

Although OmpG is not normally expressed in lab strains of *E. coli*, when induced it is targeted to the outer membrane where it forms a non-specific porin channel. The extracellular loops of the common porins are poorly conserved and previous studies have shown that they can accommodate insertional or 'sandwich' fusions (the largest documented was 162 amino acids

in an extracellular loop of OmpC) without loss of membrane targeting and insertion [43]. It was hypothesised that OmpG would share this property, and could therefore be used as a bacterial display scaffold.

Further, as the utility of OmpG for stochastic sensing has been demonstrated [44], it was proposed that it would be possible to select high-affinity binding OmpG mutants that could subsequently be rapidly reconstituted into an artificial bilayer system for single channel recording and stochastic sensing. Ideally, clones would be taken from bacterial plates and rapidly reconstituted with minimal purification/processing steps. As such, the technique of droplet interface bilayers (DIBs) seemed well suited. The goal was a single continuous protocol that could be followed in a time-frame typical for display technologies (several weeks) and that would result in the selection of novel high-affinity stochastic sensors for an arbitrarily chosen protein analyte. These sensors would be directly characterised and used in DIBs (Figure 3.1).

3.1.3 Chapter overview

This chapter details the progress made towards realising the proposed methodology (Figure 3.1). Preliminary experiments to determine the suitability of OmpG for bacterial display are described, followed by the production of a high diversity library in an extracellular loop of OmpG. In addition, rapid methods for the electrophysiological characterisation of selected OmpG loop mutants were developed using DIBs. Some technical challenges remain before the methodology can be fully tested and these are discussed at the end.

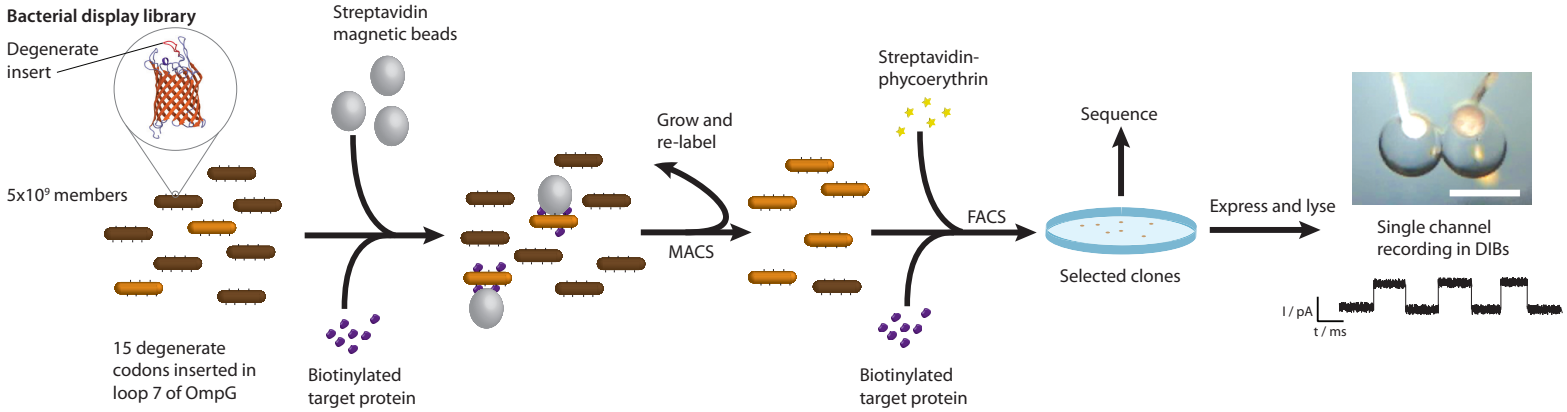


Figure 3.1: Overall scheme for the selection and characterisation of OmpG loop mutants from a combinatorial library that bind a target protein. Binding clones (shown as orange cells among a population of black, non-binding cells) are enriched through several rounds of magnetic activated cell sorting (MACS) before final selection by fluorescence activated cell sorting (FACS). The selected clones are then characterised rapidly by single channel recording using droplet interface bilayers (DIBs). The scheme describes selection against a biotinylated target protein but other labelling schemes are possible and are described in the main text.

3.2 Results

3.2.1 Plasticity of OmpG extracellular loops

The extracellular loops of non-specific diffusion porins are poorly conserved, suggesting they may tolerate extensive mutagenesis [45]. To demonstrate the feasibility of using OmpG as a bacterial display carrier protein, it was necessary to show that: (i) the extracellular loops can accept the insertion of additional peptide sequences, with the mutant protein continuing to be expressed and translocated into the outer membrane as a functional porin, (ii) the inserted peptide sequences are displayed on the cell surface such that they are accessible to binding proteins.

T7-tag loop insertion mutants

To demonstrate the suitability of OmpG for bacterial display, the T7-tag epitope was chosen to test the plasticity of the OmpG extracellular loops. This eleven amino-acid epitope, MASMTGGQQMQ, is of similar length to previously constructed loop insertion peptide libraries in the literature and has also been used to test the plasticity of the extracellular loops of OmpA [33]. Further, commercial antibodies for this epitope are available, allowing display levels to be assessed by FACS. All seven extracellular loops of OmpG were tested for their acceptance of this peptide insertion.

The T7-tag insertion mutants were generated using homologous recombination with the primers listed in Table 3.1 (Section 3.5.1). To facilitate subsequent library construction, flanking NcoI and SpeI sites, at the 5' and 3' ends respectively, were simultaneously introduced. Therefore, the final in-

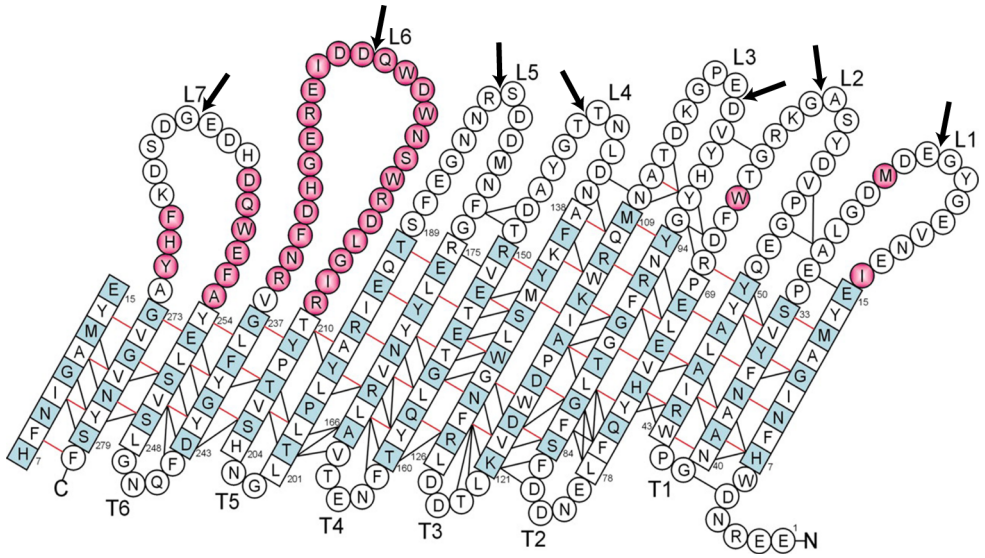
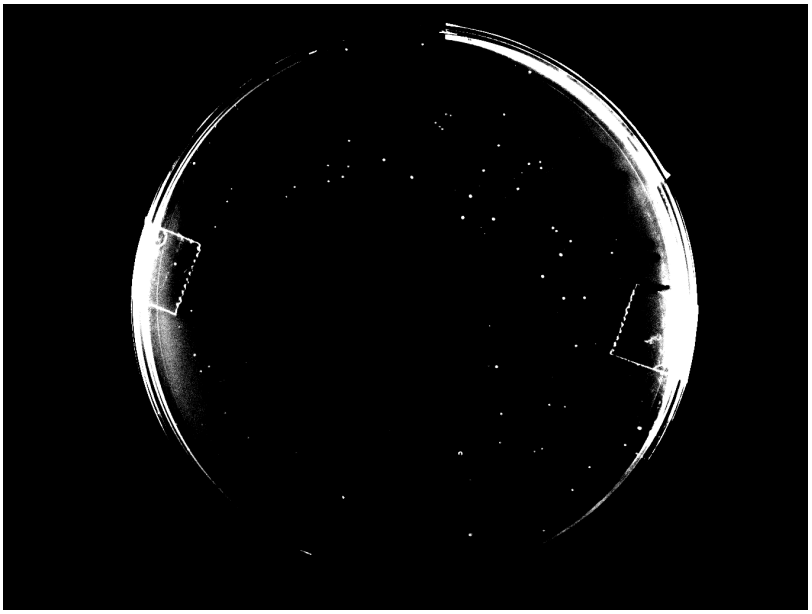


Figure 3.2: T7-tag insertion sites indicated on a topology map of OmpG. Arrows indicate the sites where the T7-tag peptide sequence was inserted between two flanking amino acids. The topology map is reproduced from Liang *et al.*, where it was derived from solution NMR data. Residues depicted by squares are in β-sheets. Other residues are shown as circles. Red lines indicate hydrogen-bond pairs identified by nuclear overhauser effects (NOEs). Black lines indicate long-range backbone NOEs. Unassigned residues are coloured pink. The extracellular loops are labelled L1 to L7; intracellular turns T1 to T6. Reproduced and adapted from Liang *et al.*, [46]. Copyright by the National Academy of Sciences.

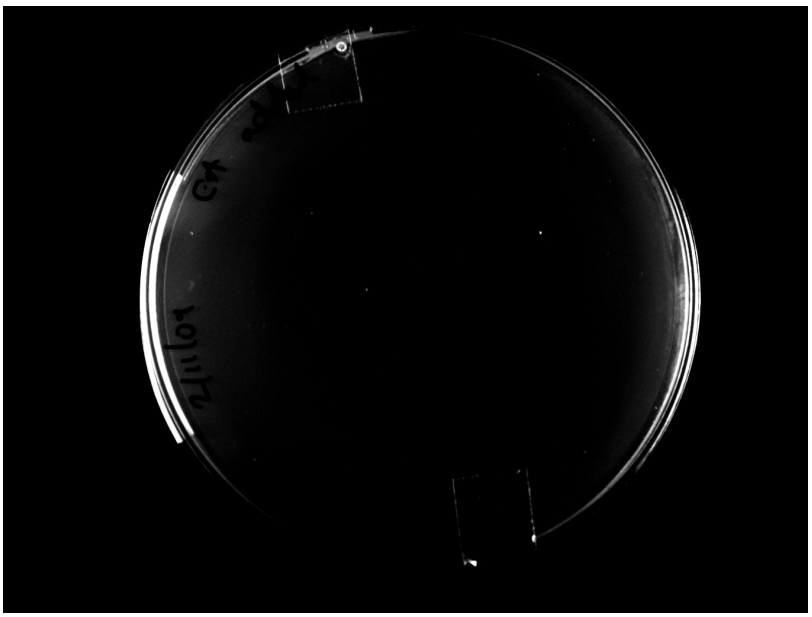
serted sequence was: *S*MASMTGGQQMQ*TS*, where the italicised residues arise from the introduced restriction sites. The insertion sites were chosen to be approximately in the middle of the native extracellular loops as informed by the NMR structure (Figure 3.2) [46]. The template plasmid for this mutagenesis was pT7-OmpG-ΔD215, which was generated from pT7-SMC-OmpG, which encodes the wild type OmpG gene including signal sequence. The ΔD215 mutation was used as it has been previously shown to 'quieten' the pore in single channel recording [44].

Trehalose assay

All seven T7-tag loop insert mutants were first tested using a trehalose uptake assay. The porin-deficient *E. coli* strain PC2889 lacks the large porins OmpF, OmpC and LamB. For this assay PC2889E (PC2889 transformed with the pLysE plasmid) was used to maintain low levels of basal expression from plasmids bearing T7 promoters, and to limit toxicity. Both strains are unable to grow on minimal media supplemented with larger sugars such as trehalose or maltotriose as the sole carbon source, as no suitable uptake pathway is available. However, when PC2889E is transformed with a plasmid containing the wild type OmpG gene under control of a T7 promoter, cell growth is rescued, as OmpG is of sufficient size to allow uptake of larger sugars (Figure 3.3(a)). When instead a plasmid containing a gene encoding mature OmpG is transformed, growth is not rescued as the protein is not targeted to the outer membrane (Figure 3.3(b)). Thus, it is possible to screen the T7-tag loop insert mutants for correct membrane insertion and pore formation.



(a) pT7-OmpG



(b) pT7-OmpGm

Figure 3.3: Trehalose growth rescue assay. The porin-deficient *E. coli* strain PC2889E is unable to grow on M9 minimal media supplemented with trehalose as the sole carbon source, as it lacks uptake pathways for large sugars. (a) When transformed with pT7-OmpG, basal expression of OmpG rescues growth by allowing trehalose uptake. (b) In contrast, mature OmpG does not rescue growth as the porin is not targeted to the outer membrane. The contrast of these photographs has been enhanced to improve the visibility of colonies.

It is important to note that porin-deficient *E. coli* strains are often able to adapt to selective conditions through mutation or altered regulation of other pathways. In PC2889, while the lamB gene has been deleted, OmpF and OmpC are suppressed by an ompR mutation. Thus, reversion is possible if care is not taken. As a result, in this trehalose assay a few colonies can often be seen on plates where growth should not be rescued (Figure 3.3(b)). Colony size variation is also apparent. However, there is at least an order of magnitude more colonies on a plate where growth has been rescued than one where it has not. Therefore, success must be determined in a manner relative to the positive and negative controls based on total colony count.

The seven OmpG loop insertion mutants were transformed into chemically competent PC2889E cells, which were then plated onto M9 minimal agar supplemented with trehalose (0.4% w/v final concentration). The bacterial plates were incubated at 37°C for 72 hours to obtain suitably sized colonies. All bacterial plates were compared to a wt-OmpG control plate as the reference phenotype. OmpG-ΔD215 was also tested to ensure that the residue deletion did not affect the positive assay result. While all seven loop insert mutants gave visible colonies, the loop 1 and loop 2 mutants gave much smaller colonies that did not resemble the control plate. These were deemed to represent a partially defective phenotype and so were considered to have failed the assay. All other loop mutants reproduced the control phenotype and so were deemed to have passed the assay. Those that passed were carried forward to the next evaluation.

Flow cytometry assay

To complement the trehalose assay, loop insert mutants 3–7 were also tested for effective cell surface display by flow cytometry. Small cultures of PC2889E cells harbouring the loop insert mutant plasmids were grown to OD 0.7–0.8 and induced for 1 hour at 37°C. A small quantity of cells for each culture was then harvested and labelled in phosphate buffered saline (PBS): first with biotinylated mouse anti-T7-tag monoclonal antibody (mAb) at 3.33 nM, and then with streptavidin-phycoerythrin conjugate (SAPE) at 10 nM. The cells were then resuspended in PBS to a concentration suitable for analysis in the flow cytometer. Negative controls were also prepared, e.g. cells expressing OmpG-ΔD215 that were: (a) left unlabelled, (b) labelled only with SAPE, and (c) labelled with both mAb and SAPE. These controls allow calibration of the flow cytometer to the non-fluorescent population and confirm the absence of non-specific binding from primary and secondary labelling agents.

All loop insert mutants (3–7) gave rise to highly fluorescent cell populations and so were capable of displaying the T7-tag epitope (Figure 3.4). Loop 7 was ultimately chosen for construction of a loop insert library. This choice was arbitrary as loops 3–7 were essentially indistinguishable in the trehalose and FACS tests.

3.2.2 Expression system

Although the preliminary experiments used the T7-promoter system, there are constraints on the choice of expression system when performing bacterial display. Firstly, the choice of host *E. coli* strain is restricted to those

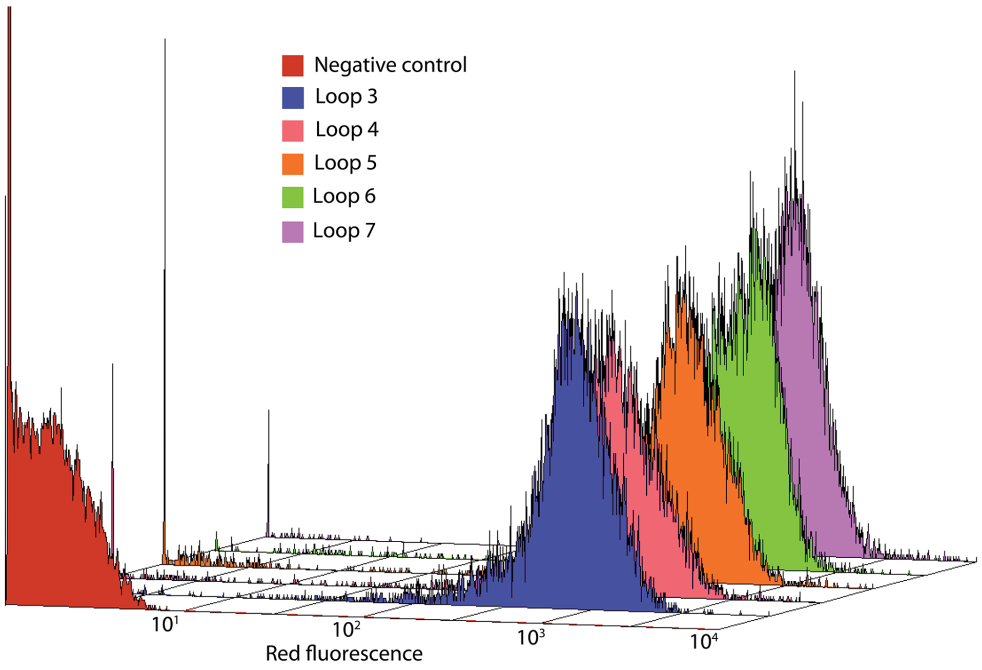


Figure 3.4: Flow cytometry histograms for PC2889E cells harbouring OmpG T7-tag loop insert mutants 3–7 and pT7-OmpG-ΔD215 as the negative control. Cells were labelled with 3.33 nM biotinylated mouse anti-T7-tag mAb and 10 nM SAPE after they had been induced for 1h at 37°C.

that can be made to achieve the highest transformation efficiencies so that a high diversity library may be constructed. Currently, this rules out B strain derivatives as these are typically one to two orders of magnitude less competent than equivalent K12 derivatives. For this project, the K12 derivative MC1061 [47] was chosen as it can achieve efficiencies approaching 1×10^{10} cfu/μg under ideal conditions.

Secondly, a critical factor is that the expression system must allow sufficient protein to be expressed so that it can be detected on the cell surface, but without arresting cell growth or causing cell death. Additionally, it is preferable to have a tightly controlled promoter to prevent the loss of mildly toxic sequences in a library.

The choice of MC1061 cells prohibited the use of T7 promoter-based plasmids as these require DE3 lysogen strains. It was therefore decided to use a T5 promoter-based plasmid based on pQE60. pQE60L is a pQE60 plasmid that has been modified to allow cis repression by a lacI^q gene contained on the plasmid. The multiple cloning site (MCS) was further modified. First, an NdeI site in a non-coding region of the plasmid was destroyed. Then the NcoI site in the MCS was mutated by homologous recombination to an NdeI site, to create the final plasmid pQELH1. This allowed subsequent sub-cloning of T7-tag OmpG loop insert mutants into pQELH1. As the loop 7 T7-tag insert mutant was chosen for library construction, this was subcloned into pQELH1 to give the plasmid pQELH1-OmpG-ΔD215-L7I.

Before beginning library construction and screening, it was necessary to optimise expression conditions, specifically the time after induction at which cells are screened, and the incubation temperature after induction. First, the

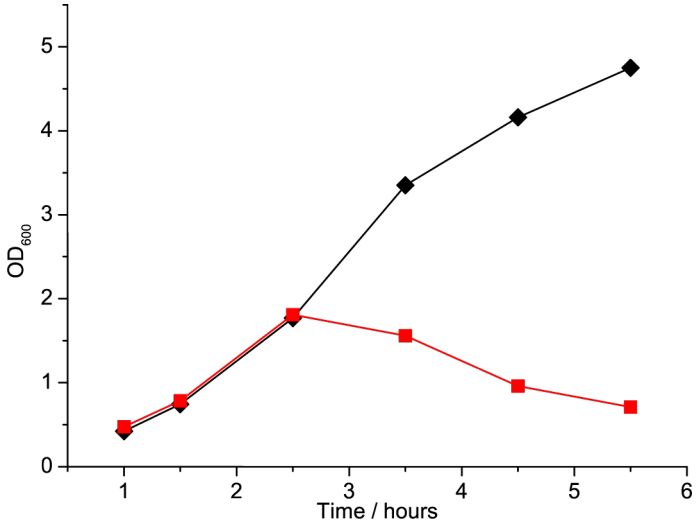


Figure 3.5: Growth curves for uninduced and induced cultures of MC1061 harbouring pQELH1-OmpG-ΔD215-L7I. Two cultures were grown to an OD600 0.7–0.8, at which time one was induced with 0.5 mM IPTG and the other was left uninduced. The uninduced curve is marked with diamonds (◊); induced with squares (□).

effect of OmpG expression on culture growth was examined by simply measuring the growth curves of uninduced and induced cultures of MC1061 harbouring pQELH1-OmpG-ΔD215-L7I. Growth was monitored for four hours after induction at 37°C. As shown in Fig 3.5, the uninduced culture continued to grow to stationary phase. In contrast, the induced culture grew at the same pace up until one hour after induction, after which growth was arrested and cell death occurred.

This result seemed to suggest that these first conditions were suitable. Further, flow cytometry data also showed a high level of surface display of the expressed OmpG mutant. Therefore, a combinatorial library was constructed.

3.2.3 OmpG bacterial display library

Once the optimum promoter system and expression conditions had been found, it was possible to begin construction of a high diversity peptide library inserted in loop 7 of OmpG.

Library design

There are several key parameters for the design of a naive peptide library. First, the number of degenerate codons must be sufficiently high to enable discovery of high affinity motifs, but if the sequence is too long the sampling of sequence space is so sparse that the probability of sampling a rare binding sequence becomes too small to be feasible. Second, the nature of the degenerate codon used can have profound consequences on the fitness of the library [48].

The choice of codon set—exactly which amino acids to include—is an important consideration. The number of possible amino acid combinations quickly becomes impossibly large to sample completely as the number of positions increases (five positions with twenty amino acids gives $20^5 = 3.2 \times 10^6$ possible sequences). Because codons for most amino acids are degenerate, the situation is even worse in terms of the number of codon combinations that can be completely sampled. If one wishes to construct a random library that still contains all twenty common amino acids, NNK or NNS codons are typically used, where N = A, C, T or G, K = G or T, and S = G or C. These codons scale as 32^n, a considerable improvement on using all 64 possible codons. An additional complication is that NNK or NNS codon libraries still

have unequal representations of amino acids.

In the human immune system, the CDRs of antibodies have strong amino acid biases. For example, 39% of the amino acids in the third region of human heavy chains are comprised only of tyrosine, serine or glycine [49]. This has posed the question of whether using reduced codon sets that contain only a subset of the full amino acid repertoire are capable of equal or better fitness than 'complete' codon sets. In fact, recent studies have shown that high-affinity binding peptides can be selected from libraries comprising a minimal subset of amino acids [50, 51]. In the most extreme example, libraries comprised only of tyrosine and serine produced high affinity binding in antibody or 'monobody' scaffolds [52, 53].

For this library, a length of 15 degenerate codons was chosen. This was informed by previous work that had used a highly diverse library of 15 NNS codons [33]. With a reported library size of 5×10^{10}, sampling of 15-mer sequences was very sparse. Nonetheless, it would be expected to contain all possible 7-mer sequences with 99% confidence. The library was successfully screened against five distinct protein targets, suggesting the library design is effective [33].

While previous work has typically used the NNK or NNS codon, for the present project the NDT codon (where D = G, A, or T) was chosen. While this is a more reduced codon set, the various chemical property groups of side-chains are still represented, e.g. basic, acidic, hydrophobic, and aromatic residues [48]. However, no stop codons are present which eliminates truncated mutants. Further, the smaller number of codons increases the degree of sequence space sampling relative to larger codon sets such as NNK for a

library of a fixed size. The library size of Bessette *et al.* is by far the largest reported in the literature [33]. For the present project it was expected that the maximum achievable diversity would most likely be an order of magnitude lower, in common with most other published studies. Thus, the NDT codon was chosen to maximise sampling of the sequence diversity. Further, this specific codon set has not been used for the selection of binding peptides; its use has been restricted to the directed evolution of enzymes. Therefore, it merits further investigation.

Library construction

Using the OmpG loop 7 T7-tag insert mutant generated previously (pQELH1-OmpG-ΔD215-L7I), a loop insert peptide library can be generated by cassette mutagenesis between NcoI and SpeI sites. However, prior to this, a StuI site was introduced approximately equidistant between these two sites. This was introduced to allow the further reduction of background clones by either triple-digestion of the vector or digestion after ligation. As detailed later, this was found to typically reduce background by a factor of 7. In this case, although background should not cause the false positives that can occur with the directed evolution of enzymes, background clones reduce the proportion of 'useful' clones in the library and therefore reduce fitness and sampling.

The degenerate cassette was generated by Klenow fragment extension of an oligonucleotide containing fifteen NDT degenerate codons and primed with a shorter oligonucleotide (Figure 3.6). The double stranded DNA was then digested with NcoI and SpeI to generate compatible sticky ends for the vector. The digested insert was purified by neutral polyacrylamide gel

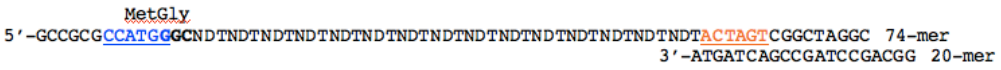


Figure 3.6: Klenow extension scheme showing the two oligonucleotides used to generate a degenerate cassette for library construction. The NcoI restriction site is coloured blue and the SpeI site is coloured orange.

electrophoresis (PAGE) before ligation.

Cut vector was prepared by double digestion of the plasmid. The digested vector was purified using QIAquick spin columns (silica membrane) followed by further purification on PCR Kleen spin columns (gel filtration).

Small-scale test ligations were then performed to (i) determine the optimum insert to vector ratio, (ii) determine the proportion of background clones, (iii) determine the transformation efficiency, (iv) evaluate the effect of background reduction measures.

Within error, insert to vector ratios 3:1. 5:1 or 10:1 did not result in significant changes in ligation efficiency. Therefore, to reduce the likelihood of double insertions, a ratio of 3:1 was used for the large scale ligation. Without background reduction, the proportion of background clones was approximately 25%. To reduce background, after ligation the reaction is subjected to StuI digestion, which linearises any undigested supercoiled vector present, as well as any (re-circularised) singly cut plasmid. With this method, background was reduced to approximately 3%. Alternatively, triple digestion of the plasmid with NcoI, SpeI, and StuI before ligation of the insert gave a similar final level of background.

Before transformation, ligation reactions were treated with QIAquick spin columns. After treatment, the ligation reaction was transformed by electroporation into electrocompetent MC1061 cells. Viable cells were recovered

by incubation in SOC media and the outgrowth was subsequently used to inoculate a larger culture. This culture was grown overnight and then flash frozen as glycerolised aliquots. The final library was estimated to contain 4×10^9 transformants, determined by plating dilutions from the outgrowth and colony counting, consistent with expectations. Colonies were randomly selected from the dilution plates and the insertion sequences encoded by the plasmids they harboured were sequenced. The sequences were statistically analysed to verify the amino acid composition was consistent with a random library (Materials and methods, Section 3.5.4).

3.2.4 Further expression system optimisation

Although from simple growth curve experiments it had appeared that induction for one hour at 37°C would be suitable, it was soon apparent that this is not the case. When screening a cell surface display library, several rounds of selection and then amplification of the selected clones are performed. It is therefore essential that cells remain viable so they may be amplified further or form colonies on agar plates at the final stage. A growth curve alone does not provide sufficient information on cell viability. This was quickly apparent when a trial MACS selection was performed on the newly constructed combinatorial library. It was found that the induced cells did not survive when plated on LB agar and incubated overnight, suggesting a loss of cell viability.

A simple experiment was designed to measure cell survivability. A small culture was induced under the required conditions and then dilutions were

plated onto LB agar. The number of colonies observed the next day were then compared to a control plate that had colonies from an uninduced culture grown under the same conditions. The number of cells initially plated was standardised for variation in culture density. Although the variance of results is high, it is able to provide semi-quantitative results that can serve as a rough guide to optimising conditions. Optimisation was performed in tandem with flow cytometry to confirm the conditions resulted in sufficient surface display levels.

To minimise toxicity, induction of pQELH1-OmpG-ΔD215-L7I was first tested for 30, 60, 90 and 120 minutes at 20°C. First, display levels were compared by flow cytometry by labelling with biotinylated anti-T7-tag mAb and SAPE as previously described. An increase in mean fluorescence intensity was observed as the induction time was increased. Display appeared to be rapid even at reduced temperatures—after only thirty minutes mean fluorescence was greater than 10-fold higher than background. The effect of these conditions on cell viability was then measured using the plate-based assay previously described.

The number of surviving colonies was significantly lower for all induction conditions surveyed, versus the uninduced control, with the number of survivors decreasing further as the induction time was increased (e.g. Figure 3.7). When combined with the flow cytometry results, it is clear that even at low temperature and for short induction periods, expression proceeds rapidly at a high level, which presumably gives rise to toxicity. Due to this result, it was decided that the T5 promoter system of the pQE plasmid was too powerful and could not be easily modulated in a fashion suitable for this

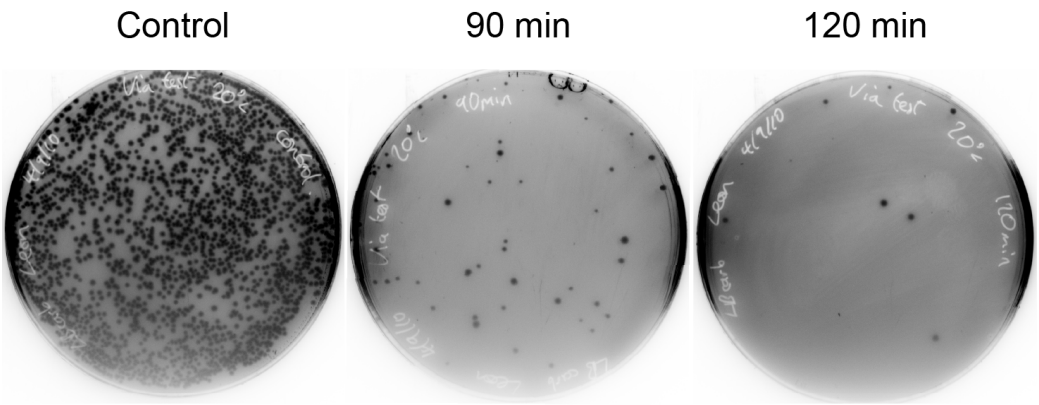


Figure 3.7: Example photographs of bacterial plates illustrating the relationship between cell survivability and induction time. The plasmid used for these tests was pQELH1-OmpG-ΔD215-L7I. Plates shown are the control plate (uninduced culture), 90 minute induction at 20°C, and 120 minute induction at 20°C. The same number of cells were spread on each plate and the number of surviving colonies were counted the next day after incubation at 37°C.

project. As an alternative, it was decided to test pBAD24, which utilises the weaker, tightly repressed araBAD promoter. This has previously been shown to be useful for the controlled expression of toxic genes [54].

The gene OmpG-ΔD215-L7I(S) was sub-cloned into pBAD24 to give pB24-OmpG-ΔD215-L7I(S). The L7I(S) variant has a StuI site introduced between the NcoI and SpeI sites that flank the T7-tag encoding region and was introduced for the purpose of background reduction when constructing libraries. The StuI modified mutant was found to behave the same as the original mutant in the assays used here and so was used to minimise subsequent steps leading to library construction.

In the araBAD system, it was found that induction at 20°C did not produce a significant amount of fluorescence in the FACS assay, even after long induction periods, such as overnight. However, induction at 25°C showed a measurable increase in fluorescence with time. For induction at 30°C,

fluorescence levels increased more rapidly. However, for induction at 25°C only approximately 1 in 4 cells survived, even after just one hour induction. Survivability was markedly lower for induction at 30°C.

3.2.5 Rapid electrophysiology of OmpG mutants

Previous electrophysiological studies of OmpG have typically used protein over-expressed in *E.coli*, purified from inclusion bodies and refolded in detergent [44, 55, 56]. While this method produces a high yield of expressed protein, the procedure is laborious, refolding yields may be low, and there is the possibility that the detergent refolded conformation may not be the native conformation. However, the 3D structures of OmpG obtained from native and refolded protein preparations do not differ significantly (root mean squared deviation 0.66 Å between PDB: 2F1C and 2IWV).

For the workflow described in Section 3.1, it is expected that library sorting by MACS and FACS will result in a number of positively selected clones that will require characterisation by single channel recording. Over-expression and purification of each clone from inclusion bodies would be slow and laborious and so it would be desirable to have a more rapid method for single channel characterisation. In anticipation of this requirement, a rapid method for the single channel characterisation of positively selected OmpG clones was developed. Specifically, desired characteristics were that: (i) the method would avoid laborious purification steps, (ii) the method would be compatible with screening mutants in quick succession, (iii) the method should involve the least number of steps possible between isolation of posi-

tively selected OmpG clone colonies and single channel recording.

It was decided to use the Droplet Interface Bilayers (DIBs) technique for single channel recording as it is compatible with the rapid screening of different clones [57]. Further, as the volumes involved in the droplets are very small, tiny quantities of protein are required for channel reconstitution into the bilayer.

For protein preparation, it was decided to prepare whole cell lysates from small cultures of porin-deficient PC2889E cells over-expressing OmpG mutants with intact signal sequences. A porin-deficient *E. coli* strain was chosen to avoid the observation of the ubiquitous porins OmpF, OmpC and LamB. As OmpG would be over-expressed, it was supposed that this would be the predominant membrane protein in the preparation and so the observation of other endogenous porins would be minimal. Gentle lysis was performed in buffer containing EDTA to inhibit metalloproteases and weaken the lipopolysaccharide layer (LPS), lysozyme for cell lysis, and 1% octylglucoside (OG) to solubilise OmpG from the cell membranes. In addition, Benzonase nuclease (Merck) was added to hydrolyse DNA and RNA and reduce the viscosity of the sample. After lysis and incubation with nuclease, the lysates were clarified by centrifugation and then used directly in the DIB setup.

Wild type OmpG

Wild type OmpG was first used as a test case for the rapid characterisation method described above. The clarified lysate was diluted (typically 100-fold) into a solution of liposomes (10 mol % 1,2-dioleoyl-*sn*-glycero-3-phosphate

(DOPA) in DPhPC) in buffer (10 mM HEPES pH 7.4, 500 mM KCl). The choice of lipid composition was based on anecdotal observations of increased stability rather than any requirement on the part of the properties of OmpG. Pure DPhPC bilayers are also effective but were observed to be more prone to fusion between the droplets and were susceptible to membrane defects that generate noise (data not shown). Once the lysate was diluted, the resulting solution was used to form the *cis* droplet as described in Section 1.2.2. The *trans* droplet was formed of liposome solution without lysate.

After a period of incubation, a droplet interface bilayer was formed in the normal manner. Bilayer formation occurred more slowly than when purified protein samples were used. Subsequently, multiple channel insertions were observed on the application of a potential across the bilayer (Figure 3.8). Single channels could be obtained by using higher dilutions of lysate in liposome solution, to limit pore insertion frequency. The mean open channel conductance at +50 mV was 1.09 ± 0.03 nS ($n = 15$).

OmpG-ΔD215-Pimtide

As a further test case, an OmpG mutant in the 'quiet' ΔD215 background was generated that had the Pimtide sequence inserted at the approximate midpoint of the seventh extracellular loop. This was envisaged as a test case for potential clones from library selection. Additionally, it would allow a further test of the rapid characterisation method by monitoring Pim kinase binding to the pore, which would serve to verify the identity of the channel being observed.

Channel insertion appeared less favourable than for wild type OmpG.

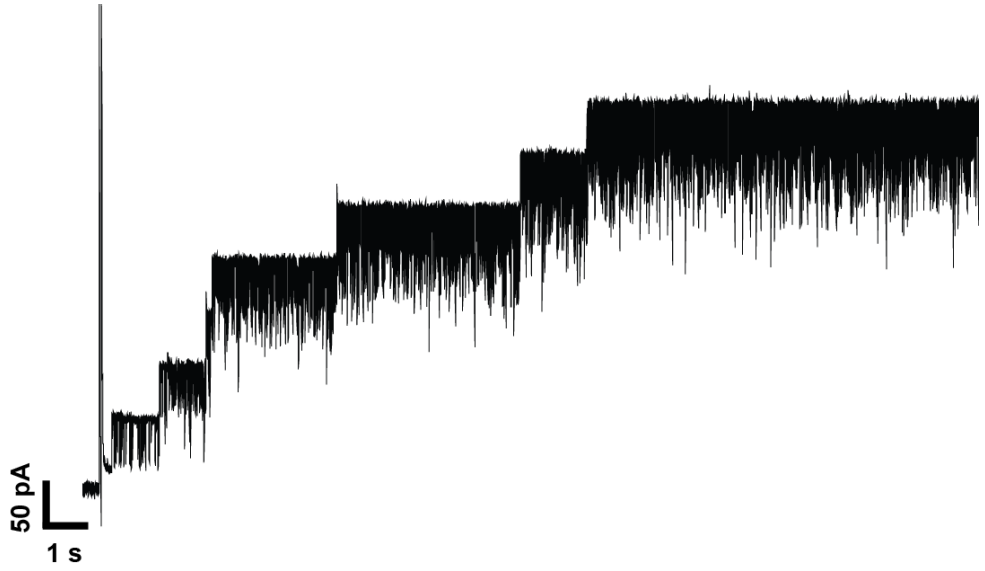


Figure 3.8: Step-wise insertions of wt-OmpG in a Droplet Interface Bilayer. Clarified lysate containing wt-OmpG, prepared as described in the text, was diluted 100-fold in a vesicle solution (1:10 DOPA to DPhPC molar ratio) in 10 mM HEPES pH 7.4, 500 mM KCl. The large amplitude spike visible to the left of the recording was due to the application of a potential of +50 mV, upon which channels began to insert. The filter cut-off frequency was 1 kHz.

This may be due to a lower level of correctly folded protein being present in the *E. coli* lysate as a consequence of the Pimtide mutation. Nonetheless, channels were observed with mean conductance 0.71 ± 0.04 nS ($n = 3$). The current-voltage relationship was also measured between -40 and $+40$ mV (Figure 3.9). These channels exhibited notable asymmetry in gating behaviour (e.g. Figure 3.10).

In repeat measurements both orientations of protein were observed, as inferred from inversion of the asymmetric gating behaviour relative to the polarity of the applied potential. However, the orientation depicted in Figure 3.10 was observed more often. The introduced Pimtide sequence has a net charge of +5 at pH 7.4. This highly charged loop presumably presents a higher energetic barrier to membrane insertion from the extracellular side

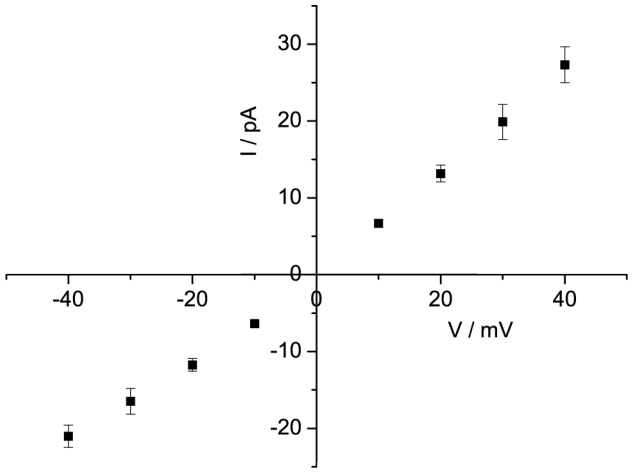


Figure 3.9: Current-voltage relationship of OmpG-ΔD215-Pimtide. Measurements were performed in 10 mM HEPES pH 7.4, 500 mM KCl. The bilayer composition was DPhPC with 10 mol % DOPA. Error bars represent the standard deviation of the mean ($n = 3$).

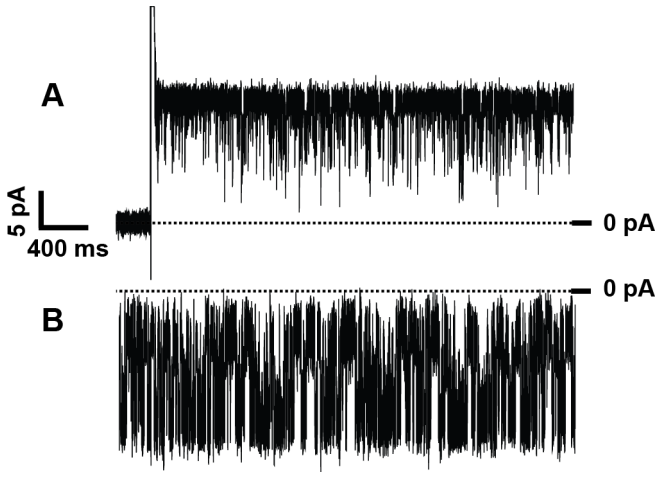


Figure 3.10: Representative current traces of OmpG-ΔD215-Pimtide. The measurement was performed in 10 mM HEPES pH 7.4, 500 mM KCl. The bilayer composition was DPhPC with 10 mol % DOPA. The applied voltage was +30 mV in the top trace (A) and −30 mV in the bottom trace (B). In the top trace, the potential was applied shortly after the beginning of the trace, which gave rise to the capacitive spike observed. The filter cut-off frequency was 1 kHz.

of OmpG than insertion from the intracellular side. Therefore, the working hypothesis is that the orientation corresponding to the trace shown in Figure 3.9 was where the extracellular side of OmpG was exposed to the *cis* chamber and the intracellular side of OmpG was exposed to the *trans* chamber. This may explain the distinct gating behaviour (Figure 3.10)—at negative potential the positively charged loop may experience an electric field that conformationally biases it towards the pore mouth (i.e. towards the negatively charged electrode), giving rise to the longer occupancy of the blocked state observed relative to that at positive potential.

Detection of Pim-2 kinase

As a preliminary test, binding of Pim-2 kinase to the OmpG-ΔD215-Pimtide pore was explored. Pim-2 was diluted in vesicle/lysate mixture (prepared as above) to a final concentration of 1.1 μM before the *cis* droplet was formed as previously. As before, the *trans* droplet contained vesicles and buffer only. The choice to place Pim-2 in the *cis* droplet was guided by the observations on orientation described above.

After bilayer formation and channel insertion, the pore properties were altered at positive applied potential, where an additional current level was observed (Figure 3.11; new level marked as 'b'). The observation was made independently three times, but the nature of the interaction and its kinetics were not further studied as the inherent noise of the pore prohibited reliable idealisation and analysis of the current recordings.

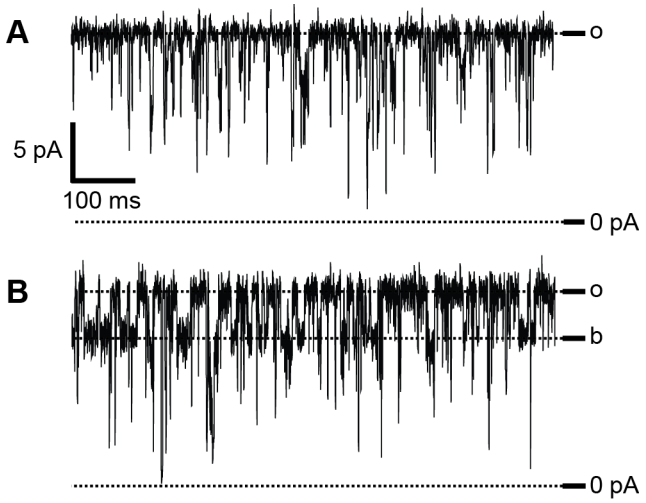


Figure 3.11: Representative current traces of: A) OmpG-ΔD215-Pimtide, B) OmpG-ΔD215-Pimtide in the presence of 1.1 μM Pim-2. The open state is marked 'o' and the 'blocked' state marked 'b'. The applied voltage was +30 mV. The measurement was performed in 10 mM HEPES pH 7.4, 500 mM KCl. The bilayer composition was DPhPC with 10 mol % DOPA. The filter cut-off frequency was 1 kHz.

Electrophysiology conclusions

It is possible to reconstitute natively expressed OmpG pores directly from a clarified lysate into DIBs without any intermediate purification steps. A similar approach was recently reported where eukaryotic ion channels were reconstituted into Droplet Hydrogel Bilayers (DHBs) from the membrane fraction of mechanically disrupted cells over-expressing the desired protein [58]. Channel insertion occurs readily for wild type OmpG and less readily for loop mutants. This may be due to the highly charged nature of the Pimtide insertional mutant, either by presenting an energetic barrier to membrane reconstitution, or because it may cause lower amounts the pore to be successfully exported to the *E. coli* outer membrane during over-expression, resulting in a lower concentration of active protein.

Preliminary observations were also made of OmpG-ΔD215-Pimtide in

the presence of Pim-2 kinase. Under these conditions the pores displayed a new intermediate conductance level. However, further experiments would be necessary to directly attribute this to kinase binding events, for example concentration dependence studies on the associated rate constants. This would be complicated by the inherent noise of the pore, which appeared to obfuscate binding events and prevented satisfactory analysis of the current recordings presented above.

Nonetheless, these experiments show that the approach should be applicable to the screening of positively selected OmpG mutants from a combinatorial library in a rapid manner. In the envisaged scheme, colonies would be picked from a plate, and used to inoculate small over-expression cultures that would then be prepared as described above. The prepared lysates would then be used to form droplets in a chip similar to that previously used for rapid screening [57], whereby each droplet containing a different clone could be rapidly accessed and measured with another droplet attached to a moveable arm.

One disadvantage of the DIB setup is that the aqueous solution contained within a droplet is more difficult to access once formed than in the conventional planar lipid bilayer setup. This can make it slower to perform titrations of substrate. One solution would be to create a series of droplets containing the expressed material and differing concentrations of substrate. However, for the purpose of initial screening, a simple phenomenological result may be sufficient.

3.3 Conclusions

Towards the goal of realising the scheme outlined in Figure 3.1, several important proofs of concept have been demonstrated. First, it was necessary to demonstrate that the loops of OmpG would accept the insertion of peptides within them, and that these insertion mutants would continue to be correctly transported to and inserted into the *E. coli* outer membrane in such a way that the inserted peptide was accessible to extracellular protein binding partners. By engineering OmpG mutants containing the T7 tag epitope within each extracellular loop, correct transport and pore function was demonstrated by a trehalose uptake assay that confirmed that insertion mutants in loops 3–7 (as defined in Figure 3.2) gave normal uptake phenotypes. Further, cells expressing these mutants that were labelled with anti-T7-tag mAb and a fluorescent secondary labelling agent were shown to be highly fluorescent when examined by FACS, indicating that the inserted peptide is displayed on the cell surface in an accessible fashion.

Following this, a combinatorial library of approximately 4×10^9 clones was produced using the NDT degenerate codon in the seventh extracellular loop of OmpG. However, further testing of the expression system and the effects of expression of OmpG mutants in the host *E. coli* cells indicated that even expression under 'mild' conditions was toxic towards the cells, something that is highly undesirable for bacterial cell surface display screening. FACS studies revealed the level of expression from the T5 promoter to be very strong with rapid onset, and it was only weakly decreased at lower temperature.

One possible reason for the toxicity observed is that high levels of over-

expression may produce protein more rapidly than can be transported and folded into the outer membrane. This protein may misfold in the cytoplasm and/or the periplasm and induce cellular stress responses. As an alternative expression system, the araBAD promoter was tested, using the expression plasmid pBAD24. However, while weaker than the original T5 promoter system trialled, it still presented problems in terms of cell viability and toxicity.

Finally, as described in the preceding section, proof of concept experiments have been performed for the rapid characterisation of OmpG mutants in DIBs without the need for purification.

3.4 Future developments

Although most of the foundation experiments have been demonstrated, the most significant challenge remaining is to find a suitable expression system and conditions that allow sufficient expression of OmpG so that cells may be selected by MACS and FACS, but with minimal toxic effects such that selected cells can survive to the next round of selection.

The araBAD system was found to be much more suited to application in bacterial surface display, however the plasmid tested, pBAD24, still resulted in toxic levels of expression. An immediate possibility would be to test a lower copy number plasmid. pBAD24 is often claimed to have a low copy number (i.e. $\sim$10 cell^{-1}), however it does not contain the complete origin of replication of pBR322 as depicted in the original publication [54]. Instead, the origin lacks the *rop* gene that negatively regulates copy number, and hence pBAD24 has a markedly higher copy number than expected [59]. As

it is likely that in this case toxicity arises from over-expression producing translation products faster than they can be secreted and folded into the outer membrane, a lower copy number may prove beneficial.

If toxicity can be overcome, the next stage would be to demonstrate effective selection for clones that bind a suitable test target (for example an easily obtainable enzyme) and characterise them via the rapid DIB method described above.

3.5 Materials and methods

Chemicals and oligonucleotides were from Sigma Aldrich unless otherwise stated. Enzymes and DNA ladders were from New England Biolabs. The PC2889 strain (BL21(DE3)ΔlamB ompR) was from the Netherlands Culture Collection; PC2889E was generated by transformation of PC2889 with the pLysE plasmid (Novagen); MC1061 was from the Yale Coli Genetic Stock Center (CGSC). Cells were made chemically competent by the method of Chung *et al.* [60] Cells were made electrically competent by the method of Dower *et al.* [20, 61] Cloning was performed in XL-10 Gold cells (Stratagene).

The pT7 vector has been previously described [62] and was a gift from Stephen Cheley. pQE60L was a gift from Stephen Tucker. pWQ126 (*E. coli* Wza cloned in pBAD24) was used as a source of pBAD24 vector backbone and was a gift from Chris Whitfield [63]. pT7-SMC-OmpG has been previously described and was a gift from Sean Conlan [56, 64]. Biotinylated anti-T7-tag mouse mAb was from Novagen. SAPE was from Molecular Probes. QIAquick (Qiagen) and PCR Kleen (Bio-Rad) spin columns were

used as directed by standard protocols. Flow cytometry was performed on a FACSort flow cytometer (BD Biosciences).

Routine plasmid purification was performed using either Qiaprep spin miniprep or Qiagen Plasmid Maxi kits as directed by the accompanying standard protocols. DNA quantitation was performed by UV-visible absorption spectroscopy using a NanoDrop 1000 (Thermo Scientific). Water used was purified by reverse osmosis to a final resistivity of 18.2 MΩ cm (Millipore).

3.5.1 Plasmids

All plasmid insert sequences were verified by sequencing (Source BioScience, Oxford).

pT7-OmpG-ΔD215

pT7-OmpG-ΔD215 was generated by *in vivo* homologous recombination [65, 66]. Two polymerase chain reaction (PCR) reactions were performed using pT7-SMC-OmpG as template, which encodes the wild type OmpG sequence including the N-terminal signal sequence. For the first reaction, the template was first linearised by NdeI before PCR using the forward mutagenic primer: 5'-CGCTGGAGTAACTGGGACTGGCAGGAT-3' and the reverse non-mutagenic primer (SC47): 5'-CAGAAGTGGTCCTGCAACTTTAT-3'. The second PCR used template linearised by HindIII; the reverse mutagenic primer was 5'-GTCCCAGTTACTCCAGCGCAGCCCAATGCGCGTATA-3' and the forward non-mutagenic primer (SC46) was 5'-ATAAAGTTGCAG GACCACTTCTG-3'. PCR was performed using Phusion Flash HF Master

Mix (New England Biolabs) using cycling conditions as directed by the manufacturer. Aliquots of the completed reaction were analysed by electrophoresis on a 0.8% TAE agarose gel to confirm successful amplification. Equal volumes (typically 5 μl) of each PCR product were then mixed and transformed into chemically competent XL-10 Gold cells, which were subsequently plated on LB carbenicillin plates and incubated at 37°C overnight. Colonies were picked and DNA was isolated and sequenced (Source BioScience, Oxford) to confirm successful mutagenesis. Large scale DNA isolation by plasmid maxi-prep was performed when required.

T7-tag loop insertion mutants

The T7-tag sequence, flanked by NcoI and SpeI sites at the N and C termini respectively, was inserted into each of the extracellular loop regions of pT7-OmpG-ΔD215 by *in vivo* homolous recombination. This produced seven constructs each containing one T7-tag sequence in a different extracellular loop, denoted as pT7-OmpG-ΔD215-L1I, pT7-OmpG-ΔD215-L2I, ..., pT7-OmpG-ΔD215-L7I. (The loop numbering is as shown in Figure 3.2). The full inserted sequence is *S*MASMTGGQQMQ*TS*, where the italicised residues result from the introduced restriction sites. In all cases, the reverse non-mutagenic primer was SC47, and the forward non-mutagenic primer was SC46. The mutagenic primers used are listed in Table 3.1. Insertion sites were chosen to be in the middle of the extracellular loop and hence as far as possible from the transmembrane pore. The insertion sites for all seven extracellular loops are listed in Table 3.2. Plasmid DNA was isolated and sequenced as described previously.

Table 3.1: Forward and reverse mutagenic primers for the insertion of T7-tag encoding sequences into each of the seven extracellular loops of OmpG

Loop	Direction	Primer
1	Forward	5'-ATGACAGGCGGCCAGCAGATGGGCAC TAGTGAAGATATGGATGGGCTG-3'
	Reverse	5'-CTGCTGGCCGCCTGTCATGCTCGCCA TGGAGCCATAACCCTCGACGTT-3'
2	Forward	5'-ATGACAGGCGGCCAGCAGATGGGCAC TAGTGGTAAACGTGGAACGTGG-3'
	Reverse	5'-CTGCTGGCCGCCTGTCATGCTCGCCA TGGACGCGCTATAATCTACCGC-3'
3	Forward	5'-ATGACAGGCGGCCAGCAGATGGGCAC TAGTGAACCGGGTAAAGACACG-3'
	Reverse	5'-CTGCTGGCCGCCTGTCATGCTCGCCA TGGAATCAACGTAGTGATAACC-3'
4	Forward	5'-ATGACAGGCGGCCAGCAGATGGGCAC TAGTACCGGTTACGCTGATACC-3'
	Reverse	5'-CTGCTGGCCGCCTGTCATGCTCGCCA TGGAAGTGTTCAGATCGTTGGC-3'
5	Forward	5'-ATGACAGGCGGCCAGCAGATGGGCAC TAGTCGCAATAACGGTGAGTTT-3'
	Reverse	5'-CTGCTGGCCGCCTGTCATGCTCGCCA TGGAGCTGTCGTCCATATTGAA-3'
6	Forward	5'-ATGACAGGCGGCCAGCAGATGGGCAC TAGTGATGATATTGAACGTGAA-3'
	Reverse	5'-CTGCTGGCCGCCTGTCATGCTCGCCA TGGACTGCCAGTCCCAGTTACT-3'
7	Forward	5'-ATGACAGGCGGCCAGCAGATGGGCAC TAGTGGCGACAGTGATAAATTC-3'
	Reverse	5'CTGCTGGCCGCCTGTCATGCTCGCCAT GGATTCGTCGTGATCCTGCCA-3'

pQELH1, pQELH1-OmpG-ΔD215-L7I, and pQELH1-OmpG-ΔD215-L7I(S)

pQE60L contains a NdeI site in a non-coding region of the plasmid. This site was destroyed to allow a further (unique) NdeI site to be introduced into the MCS. pQE60L DNA was linearised with NdeI and the resulting 5' overhangs were filled in using Phusion Flash polymerase. The plasmid was then circularised by ligating the blunt ends together with T4 DNA ligase. The ligation reaction was transformed into chemically competent XL-10 Gold cells, plated on LB carbenicillin plates and incubated at 37°C overnight. Colonies were picked and DNA isolated by plasmid mini-prep. Test digestions with NdeI were performed to confirm the restriction site had been destroyed.

Mutagenesis by *in vivo* homologous recombination was then performed on pQE60L lacking the NdeI restriction site, to change the existing NcoI site in the MCS to an NdeI site. For the first PCR reaction, the forward mutagenic primer was 5'-GGAGAAATTACATATGGGAGGATCCAGAT C-3', and the reverse non-mutagenic primer was 5'-GCTCTAGATCATGC GCACCCGTGGC-3'. For the second PCR reaction, the reverse mutagenic primer was 5'-GCCACGGGTGCGCATGATCTAGAGC-3' and the forward non-mutagenic primer was 5'-GATCTGGATCCTCCCATATGTAATTTC TCC-3'. Colonies were picked and DNA isolated as previously. Clones were tested for NdeI insertion by test restriction digests and further checked by DNA sequencing. The resulting plasmid was designated pQELH1.

To prepare pQELH1-OmpG-ΔD215-L7I, the OmpG-ΔD215-L7I gene was subcloned from pT7-OmpG-ΔD215-L7I. pT7-OmpG-ΔD215-L7I and pQELH1

were both double-digested by NdeI and HindIII. The digests were electrophoresed in a 1% TAE agarose gel and the pQELH1 vector fragment and OmpG-ΔD215-L7I gene fragment were extracted using a QIAquick gel extraction kit (Qiagen). The purified insert and vector fragments were then ligated by T4 DNA ligase and the ligation reaction was transformed into chemically competent XL-10 Gold cells. Colonies were picked and DNA isolated as described previously. Successful subcloning was confirmed by restriction digests and further confirmed by DNA sequencing.

Using pQELH1-OmpG-ΔD215-L7I as the template, a StuI site was introduced between the NcoI and SpeI sites flanking the T7-tag encoding region. This changed the encoded T7-tag sequence from MASMTGGQQMQ to MASMTGLQQMQ. Flow cytometry experiments confirmed this sequence was recognised by the biotinylated anti-T7-tag mAb comparably to the true T7-tag sequence. For the first PCR reaction, the forward mutagenic primer was 5'-CATGACAGGCCTCCAGCAGATGGGCACTAGTG-3' and the reverse non-mutagenic primer was 5'-GCTCTAGATCATGCGCACCCGTGG C-3'. For the second PCR reaction, the reverse mutagenic primer was 5'-C ACTAGTGCCCATCTGCTGGAGGCCTGTCATG and the forward non-mutagenic primer was 5'-GATCTGGATCCTCCCATATGTAATTTCTC C-3'.

pB24-OmpG-ΔD215-L7I(S)

pWQ126 was used as a source of the pBAD24 vector backbone. pWQ126 contains *E. coli* Wza cloned between EcoRI and PstI. The gene OmpG-ΔD215-L7I(S) was cloned into pWQ126 using *in vivo* homologous recombi-

nation such that the gene is contained within NdeI (introduced) and HindIII (already present). The new NdeI site begins 5 bp downstream from the Shine-Dalgarno sequence present in pBAD24. In the first PCR reaction, to generate the OmpG-ΔD215-L7I(S) gene plus homologous overlap regions, pQELH1-OmpG-ΔD215-L7I(S) was used as the template, with the forward primer 5'-CTAGCAGGAGGAATTCCATATGAAAAAGTTATTACCC-3' and the reverse primer 5'-ATCCGCCAAAACAGCCAAGCTTTTATCAG AACGAGTA-3'. In the second reaction that amplified the pBAD24 vector backbone, pWQ126 was used as the template with the forward primer 5'-AAGCTTGGCTGTTTTGGCGGATGAGAGAAGATT-3' and the reverse primer 5'-CATATGGAATTCCTCCTGCTAGCCCAAAAAAACG-3'. The resulting plasmid was designated pB24-OmpG-ΔD215-L7I(S).

pT7-OmpG-ΔD215-Pimtide

The Pimtide sequence ARKRRRHPSGPPTA was inserted between E262 and G263 in the seventh extracellular loop of OmpG. This was achieved by *in vivo* homologous recombination. The non-mutagenic primers were SC47 and SC46, as above. The forward mutagenic primer was 5'-CGTCGTCA TCCGAGCGGACCGCCTACCGCGGGCGACAGTGATAAATTC-3', and the reverse mutagenic primer was 5'-TCCGCTCGGATGACGACGACGT TTACGCGCTTCGTCGTGATCCTGCCA-3'. The template plasmid was pT7-OmpG-ΔD215.

3.5.2 Trehalose assay

M9 trehalose minimal media

To make M9 minimal media, a 5X stock solution of 'M9 salts' was first prepared. This contained of 33.9 g/l disodium phosphate, 15 g/l monopotassium phosphate, 2.5 g/l sodium chloride, 5 g/l ammonium chloride, and was autoclaved after preparation.

To make M9 trehalose minimal agar, 1.5% w/v bacto agar (Oxoid) was added to a volume of water equal to 75% of the desired final volume of medium, and this was autoclaved. The autoclaved agar was then cooled to 50–60°C in a water bath. M9 salts were added to give 1X final concentration. The agar was then supplemented with the following compounds from filter sterilised stocks (final concentrations shown in brackets): $MgSO_4$ (2 mM), $CaCl_2$ (100 μM), thiamine hydrochloride (1 μg/ml), carbenicillin (50 μg/ml), chloramphenicol (34 μg/ml) and trehalose (0.4% w/v). Sterile milli-Q water was added to bring the medium up to the final desired volume. The complete medium was then poured into petri dishes, and allowed to cool. The plates were then stored at 4°C until needed.

Growth rescue assays

For the positive and negative controls (Fig 3.3), pT7-SMC-OmpG (wild type OmpG including signal sequence) and pT7-OmpGm (mature OmpG sequence) were each transformed into chemically competent PC2889E cells and then plated on M9 minimal trehalose plates. The plates were incubated at 37°C for 72 hours at which point colonies were visible to the naked eye.

The plates were then compared to the control plates and judged as to whether successful growth rescue had occurred.

3.5.3 Flow cytometry

Cells were prepared by growing 5 ml cultures in LB at 37°C with antibiotic selection. Cultures were induced when the OD_{600} was approximately 0.7. After induction, aliquots of culture containing approximately 1×10^7 cells were centrifuged at 3,000×g for 5 minutes and the supernatant removed. Each cell pellet was resuspended in 100 μl ice-cold PBS containing 3.33 nM biotinylated mouse anti-T7-tag monoclonal antibody (mAb). The resuspended cells were incubated for 30 minutes on ice before being centrifuged at 3,000×g and the supernatant removed. Each cell pellet was resuspended in 100 μl ice-cold PBS containing 10 nM streptavidin-phycoerythrin. The resuspended cells were incubated on ice for 30 minutes before being centrifuged at 3,000×g for 5 minutes and the supernatant removed. The cell pellets were resuspended in 1 ml ice-cold PBS, transferred to 12 × 75 mm, 5 ml polystyrene round-bottom test tubes (BD Biosciences) and immediately analysed on a FACSort flow cytometer (BD Biosciences) equipped with a 488 nM laser. Typically, 1×10^4 cells were analysed. Fluorescence was measured in the FL2 channel.

3.5.4 Library construction

Digested vector preparation

100 μg of pQELH1-OmpG-ΔD215-L7I(S) was digested overnight (16 hours) with 500 units each of NcoI-HF and SpeI at 37°C in a total volume of 1.5 ml.

A small aliquot (1 μl) was analysed by electrophoresis on a 1% agarose TAE gel to confirm successful completion of digestion. Reaction clean-up was then performed using a QIAquick gel extraction kit. 4.5 ml of buffer 'QG' as supplied were added to the digestion reaction followed by 1.5 ml of isopropanol. The standard QIAquick protocol was followed, with the sample divided equally between twelve spin columns. The eluted DNA was concentrated in a Savant SpeedVac (Thermo Scientific) to a volume of 200 μl. The concentrated DNA was then divided in two and further purified on two PCR Kleen spin columns (Bio-Rad) according to the standard instructions. The final yield was 64%.

Degenerate cassette preparation

The long, degenerate oligo was 5'-GCCGCGCCATGGGC$(NDT)_{15}$ACTAGT CGGCTAGGC-3', where N=A, C, T or G and D=G, A, or T. The extension primer was 5'-GGCAGCCTAGCCGACTAGTA-3'. The extension scheme is shown in Figure 3.6. The two oligonucleotides were resuspended in TE buffer (10 mM Tris.HCl pH 8, 1 mM EDTA) at a concentration of 100 μM each. 1.0 nmol of long degenerate oligo was mixed with 3.0 nmol of extension primer in a total volume of 50 μl TE buffer + 100 mM NaCl. The mixture was heated to 98°C and then allowed to cool to room temperature to allow the oligonucleotides to anneal.

The annealed duplex was added to a final reaction volume of 200 μl containing 0.4 mM dNTPs and 10 units of DNA polymerase I Klenow fragment. This extension reaction was incubated at 37°C for one hour. 0.5 M EDTA was added to give a final concentration of 10 mM. The reaction was then heat

inactivated at 65°C for 15 minutes. DNA was extracted by phenol-chloroform extraction and ethanol precipitation according to standard protocols [67]. The precipitated DNA was resuspended in 200 μl TE buffer.

The extracted extension reaction was digested overnight at 37°C with 100 units each of NcoI and SpeI in a total volume of 400 μl. The digestion reaction was then extracted by phenol-chloroform extraction and ethanol precipitation according to standard protocols [67]. The precipitated DNA was resuspended in 40 μl of TE buffer.

The digested, extracted cassette DNA was purified by neutral PAGE. The DNA was mixed with nucleic acid sample loading buffer (Bio-Rad) and loaded onto a 15% TBE polyacrylamide gel. The gel was run in 1X TBE buffer overnight with 100 V applied. Staining of the gel was performed in a 0.002% (w/v) methylene blue solution in 1X TAE buffer for 3–4 hours. The gel was subsequently de-stained in water for 2 hours.

The major stained band was excised from the gel with a scalpel. The gel slice was weighed and two volume equivalents of acrylamide extraction buffer (100 mM sodium acetate pH 4.5, 1 mM EDTA, 0.1% SDS) were added to the gel slice in a 1.5 ml microcentrifuge tube. The gel slice was crushed in the acrylamide extraction buffer with a plastic pestle. After crushing, the mixture was incubated at 37°C overnight with shaking (400 rpm) in a Thermomixer Comfort (Eppendorf). After incubation the mixture was divided in two and the extraction buffer was separated from residual gel fragments by filtration through 0.2 μm Microfilterfuge (Rainin) spin filters. The filtrate was extracted by phenol-chloroform extraction and DNA precipitated by ethanol precipitation according to standard protocols [67]. The precipi-

tated DNA was resuspended in 50 μl TE buffer and quantitated. 4.1 μg of purified cassette DNA was finally recovered.

Ligation and transformation

To determine optimal insert-to-vector ratios, 50 ng of purified digested vector was ligated with purified degenerate insert at 3:1, 5:1 and 10:1 molar ratios using 200 units of T4 DNA ligase in a reaction volume of 20 μl incubated at room temperature for 1 hour. Ligation reactions were then heat inactivated at 65°C for 15 minutes. The inactivated reactions were digested with 5 units of StuI for 30 minutes at 37°C. The digestions were then heat inactivated at 65°C for 15 minutes. A background control was also prepared. 50 ng of purified digested vector was prepared in a ligation reaction of 20 μl where the degenerate insert was omitted and replaced with water. The control reaction was heat inactivated and digested by StuI under identical conditions to the other test ligations. An identical reaction was prepared but not digested by StuI, to examine the effect of StuI digestion on background levels.

From each test reaction, 0.5 μl was electroporated into 50 μl aliquots of electrocompetent MC1061 cells. Cells were electroporated in 0.1 cm gap electroporation cuvettes (Bio-Rad) with an applied voltage of 1.8 kV. Electroporated cells were immediately transferred to 950 μl SOC medium for each electroporation and incubated at 37°C for 1 hour. These outgrowths were diluted as required in SOC and then plated on LB carbenicillin plates. For each reaction, dilution factors of 1,000 and 10,000 were plated. Plates were incubated overnight at 37°C. The next day, colonies were counted and

transformation efficiencies calculated according to the following equation:

$$\text{efficiency (cfu}/\mu g) = \frac{\text{no. of colonies on plate}}{\text{mass of DNA plated } (ng)} \times 1000\ ng/\mu g \qquad (3.1)$$

For the scaled-up ligation, 24.0 μg purified digested vector was ligated with 0.753 μg purified degenerate insert with 2,000 units of T4 DNA ligase in a total reaction volume of 2 ml. The ligation reaction was incubated at 16°C overnight. The next day, the ligation reaction was heat inactivated at 65°C for 20 minutes. 60 units of StuI were then added to the inactivated ligation and the digestion was incubated at 37°C for 4 hours. The digestion was then heat inactivated at 65°C for 20 minutes. The inactivated reaction was divided equally in three. To each aliquot, 3 volumes of buffer 'QG' and 1 volume of isopropanol were added before these were purified across three QIAquick spin columns according to standard protocols. Pure ligated DNA was eluted in a volume of 35 μl from each column. The three eluates were pooled and then quantitated. The yield was 57%.

The purified ligation was divided into five equal aliquots. Each aliquot was electroporated in 350 μl of electrocompetent MC1061 cells in 0.2 cm gap electroporation cuvettes (Bio-Rad) with an applied voltage of 2.5 kV. Each electroporation was rescued with 1 ml of SOC and the rescued aliquots were pooled. Additional SOC was added to the pooled, rescued cells to give a final volume of 50 ml. This was incubated at 37°C for 1 hour. Dilutions were taken from the outgrowth and plated on LB carbenicillin plates that were subsequently incubated overnight at 37°C. Colonies were counted and the efficiency was calculated as described above. The diversity of the library

was calculated as: efficiency (cfu/μg) $\times$ amount of DNA transformed (μg). The remaining outgrowth was added to 500 ml LB carbenicillin and grown overnight with shaking at 20°C. Aliquots containing cell numbers corresponding to five times the calculated diversity, i.e. 2×10^{10} cells, were centrifuged at 3,000$\times g$ for 20 minutes. The cell pellets were then resuspended in 1 ml of LB+15% w/v glycerol. These 1 ml aliquots were then flash frozen in dry ice-ethanol and stored at -80°C.

To examine the quality of the library, ten colonies from the plates used to calculate efficiency were selected and the plasmid DNA isolated by mini-prep. The isolated plasmids were then sequenced. Of the ten plasmids sequenced, seven had successfully inserted random cassettes. Two had base deletion mutations and one had mixed sequences due to double transformation. Codon incidence was statistically analysed using the following equation [68]:

$$P' = [N!/k!(N-k)!][f^k(t-f)^{N-k}]/t^N \tag{3.2}$$

where N is the total number of residues, k is the number of observations of a specific residue, and P' is the probability of observing k residues given that f is the number of codons encoding that residue and t is the total number of codons in the peptide set. This P' value is then corrected for the number of independent samplings by the following formula:

$$P = 1 - (1 - P')^{12j} \tag{3.3}$$

where $j = 1$ for compositional analysis (i.e. not position-by-position).

Table 3.2: Statistical analysis of a sample of combinatorial library insert sequences. Plasmid DNA was purified from ten randomly selected colonies as described in the main text and the inserted genes sequenced. Sequences from the seven clones that did not present anomalies such as double transformants or frame shifts were then analysed using Equations 3.2 and 3.3. P is the corrected probability of observing the number of each residue found in the sample.

Amino acid	Codon	Count	% observed	% expected	observed/expected	P
R	CGT	7	6.7	8.3	0.8	0.8
N	AAT	9	8.6	8.3	1.0	0.8
D	GAT	15	14.3	8.3	1.7	0.2
C	TGT	11	10.5	8.3	1.3	0.7
G	GGT	8	7.6	8.3	0.9	0.8
H	CAT	10	9.5	8.3	1.1	0.8
I	ATT	5	4.8	8.3	0.6	0.6
L	CTT	8	7.6	8.3	0.9	0.8
F	TTT	9	8.6	8.3	1.0	0.8
S	AGT	10	9.5	8.3	1.1	0.8
Y	TAT	6	5.7	8.3	0.7	0.7
V	GTT	7	6.7	8.3	0.8	0.8
	Total	105				

3.5.5 Electrophysiology

Preparation of crude lysates

Plasmid DNA encoding the desired OmpG mutant under the control of a T7 promoter, e.g. pT7-OmpG, was transformed into PC2889E cells, which were plated onto appropriate media and incubated overnight at 37°C. A single colony was picked and used to inoculate 10 ml of LB medium, which was incubated overnight at 37°C with shaking at 200 rpm. 5 ml of the overnight inoculum was used to inoculate 100 ml of LB medium, which was incubated at 37°C with shaking at 200 rpm until the culture reached an OD_{600} of 0.8. The culture was then induced by the addition of isopropylthiogalactoside (IPTG) to a final concentration of 0.5 mM, after which the culture was incubated for a further hour at 37°C with shaking at 200 rpm. The culture

was then aliquoted into 10 ml portions that were centrifuged for 20 minutes at 3,000×g. The supernatant was discarded and the pellets were either frozen at −80°C for later use or lysed for immediate use as described in the next paragraph.

For the preparation of crude lysate, one pellet was resuspended in 200 μl lysis buffer (50 mM Tris.HCl pH 8, 5 mM EDTA, 200 μg/ml lysozyme, 1% octyl glucoside) and incubated on ice for 30 minutes. Magnesium chloride was then added to give a final concentration of 5 mM and 50 units of Benzonase nuclease (Merck) were added. The lysate was incubated on ice for a further 30 minutes. The lysate was then clarified by centrifugation for 20 minutes at 4,000×g. The clarified lysate was used for single channel recording, as described below.

Vesicle preparation

Lipids were purchased from Avanti Polar Lipids as lyophilised powders. Lipid stock solutions were prepared in chloroform with a final concentration of 10 mg/ml. Stock solutions of 1,2-dioleoyl-*sn*-glycero-3-phosphate (sodium salt) (DOPA) and 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC) were mixed to achieve 10 mol % of DOPA in DPhPC. This mixture was dried under a gentle stream of nitrogen in a glass vial, followed by further drying in a vacuum desiccator for two hours. The dried lipid film was rehydrated in a volume of buffer (10 mM HEPES pH 7.4, 500 mM KCl) to achieve a total lipid concentration of 0.5 mM. The rehydrated lipid suspension was mixed by vortexing. The vortexed suspension was subjected to freeze-thaw five times. The suspension was then passed twice through a LIPEX extruder (Northern

Lipids) with 0.1 μm Isopore Membrane Filters (Millipore). The resulting vesicle solution was stored at 4°C.

Droplet Interface Bilayers

A 10×4×4 mm Perspex chamber was filled with hexadecane. This formed the oil phase bath within which the droplets were suspended. Ag/AgCl electrodes were prepared from 0.1 mm diameter silver wire by incubation overnight in aqueous sodium hypochlorite solution. The electrode tips were coated with 3% w/v low-melt agarose in water before two electrodes were each suspended from a micromanipulator (Narishige) to allow controlled translational motion of the electrodes and their attached droplets. One electrode was connected to the ground terminal of the amplifier headstage, and the other connected to the working terminal. Following the convention for planar lipid bilayers described in Section 1.2.1, the *cis* droplet is defined as the droplet connected to ground, and contains the pore-forming protein. The *trans* droplet is defined as the droplet connected to the working side of the headstage.

The assembly of chamber, electrodes, micromanipulators and headstage was contained within a Faraday cage. The headstage was connected to an Axopatch 200B patch clamp amplifier (Molecular Devices), operating in voltage-clamp mode. Current signals were filtered with a hardware (within the Axopatch 200B) low-pass Bessel filter (80 dB/decade) with a corner frequency of 1 kHz. Signals from the amplifier were digitised using a Digidata 1320A digitiser (Molecular Devices), which was connected to a computer running the pCLAMP software suite (Molecular Devices). Current signals were sampled with a frequency of 5 kHz.

Prior to droplet formation, clarified lysate prepared as described above was diluted 100-fold into vesicle solution. The dilution factor was typically varied to optimise the rate of channel insertion for the desired experiment. The agarose-treated electrodes were submerged in the hexadecane phase within the chamber. 150 nl of the lysate/vesicle mixture was applied to the tip of the *cis* electrode using a micro-pipette while submerged in the oil phase. The same volume of vesicle solution only was then applied in the same manner to the *trans* electrode. The droplets were incubated for 5–10 minutes to allow for monolayer formation. The two droplets were brought together by means of the two micromanipulators and spontaneous bilayer formation was observed after a short period of contact. Bilayer capacitance was monitored by the application of a triangular voltage function by means of a function generator (Iso-Tech GFG-8216A).

Bibliography

[1] Movileanu, L., Howorka, S., Braha, O., and Bayley, H. (2000) Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nat Biotechnol 18*, 1091–1095.

[2] Howorka, S., Nam, J., Bayley, H., and Kahne, D. (2004) Stochastic detection of monovalent and bivalent protein-ligand interactions. *Angew Chem Int Ed Engl 43*, 842–846.

[3] Cheley, S., Xie, H., and Bayley, H. (2006) A genetically encoded pore for the stochastic detection of a protein kinase. *Chembiochem 7*, 1923–1927.

[4] Xie, H., Braha, O., Gu, L.-Q., Cheley, S., and Bayley, H. (2005) Single-molecule observation of the catalytic subunit of cAMP-dependent protein kinase binding to an inhibitor peptide. *Chem Biol 12*, 109–120.

[5] Edward Jenner. (1798) *An inquiry into the causes and effects of the Variolae Vaccinae* (Edward Jenner).

[6] Köhler, G. and Milstein, C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature 256*, 495–7.

[7] McCafferty, J., Griffiths, A. D., Winter, G., and Chiswell, D. J. (1990) Phage antibodies: filamentous phage displaying antibody variable domains. *Nature 348*, 552–4.

[8] Barbas, 3rd, C. F., Kang, A. S., Lerner, R. A., and Benkovic, S. J. (1991) Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc Natl Acad Sci U S A 88*, 7978–82.

[9] Brekke, O. H. and Sandlie, I. (2003) Therapeutic antibodies for human diseases at the dawn of the twenty-first century. *Nat Rev Drug Discov 2*, 52–62.

[10] Kempeni, J. (1999) Preliminary results of early clinical trials with the fully human anti-TNFalpha monoclonal antibody D2E7. *Ann Rheum Dis 58 Suppl 1*, I70–2.

[11] Harmsen, M. M. and De Haard, H. J. (2007) Properties, production, and applications of camelid single-domain antibody fragments. *Appl Microbiol Biotechnol 77*, 13–22.

[12] Koide, A., Bailey, C. W., Huang, X., and Koide, S. (1998) The fibronectin type III domain as a scaffold for novel binding proteins. *J Mol Biol 284*, 1141–51.

[13] Koide, A., Abbatiello, S., Rothgery, L., and Koide, S. (2002) Probing protein conformational changes in living cells by using designer binding proteins: application to the estrogen receptor. *Proc Natl Acad Sci U S A 99*, 1253–8.

[14] Skerra, A. (2001) 'Anticalins': a new class of engineered ligand-binding proteins with antibody-like properties. *J Biotechnol 74*, 257–75.

[15] Wojcik, J., Hantschel, O., Grebien, F., Kaupe, I., Bennett, K. L., Barkinge, J., Jones, R. B., Koide, A., Superti-Furga, G., and Koide, S. (2010) A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. *Nat Struct Mol Biol 17*, 519–27.

[16] Koide, S. (2009) Engineering of recombinant crystallization chaperones. *Curr Opin Struct Biol 19*, 449–57.

[17] Kossiakoff, A. A. and Koide, S. (2008) Understanding mechanisms governing protein-protein interactions from synthetic binding interfaces. *Curr Opin Struct Biol 18*, 499–506.

[18] Levin, A. M. and Weiss, G. A. (2006) Optimizing the affinity and specificity of proteins with molecular display. *Mol Biosyst 2*, 49–57.

[19] Smith, G. P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science 228*, 1315–7.

[20] Clackson, T. (2004) *Phage display: A Practical Approach.* (Oxford University Press).

[21] Kehoe, J. W. and Kay, B. K. (2005) Filamentous phage display in the new millennium. *Chem Rev 105*, 4056–72.

[22] Danner, S. and Belasco, J. G. (2001) T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. *Proc Natl Acad Sci U S A 98*, 12954–9.

[23] Pande, J., Szewczyk, M. M., and Grover, A. K. (2010) Phage display: concept, innovations, applications and future. *Biotechnol Adv 28*, 849–58.

[24] Sidhu, S. S. and Koide, S. (2007) Phage display for engineering and analyzing protein interaction interfaces. *Curr Opin Struct Biol 17*, 481–7.

[25] Boder, E. T. and Wittrup, K. D. (1997) Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol 15*, 553–7.

[26] Boder, E. T. and Wittrup, K. D. (2000) Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymol 328*, 430–44.

[27] Daugherty, P. S., Iverson, B. L., and Georgiou, G. (2000) Flow cytometric screening of cell-based libraries. *J Immunol Methods 243*, 211–27.

[28] Hanes, J. and Plückthun, A. (1997) In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A 94*, 4937–42.

[29] Mattheakis, L. C., Bhatt, R. R., and Dower, W. J. (1994) An in vitro polysome display system for identifying ligands from very large peptide libraries. *Proc Natl Acad Sci U S A 91*, 9022–6.

[30] Roberts, R. W. and Szostak, J. W. (1997) RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc Natl Acad Sci U S A 94*, 12297–302.

[31] Liu, R., Barrick, J. E., Szostak, J. W., and Roberts, R. W. (2000) Optimized synthesis of RNA-protein fusions for in vitro protein selection. *Methods Enzymol 318*, 268–93.

[32] Francisco, J. A., Campbell, R., Iverson, B. L., and Georgiou, G. (1993) Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface. *Proc Natl Acad Sci U S A 90*, 10444–8.

[33] Bessette, P. H., Rice, J. J., and Daugherty, P. S. (2004) Rapid isolation of high-affinity protein binding peptides using bacterial display. *Protein Eng Des Sel 17*, 731–9.

[34] Lee, S. Y., Choi, J. H., and Xu, Z. (2003) Microbial cell-surface display. *Trends Biotechnol 21*, 45–52.

[35] Daugherty, P. S. (2007) Protein engineering with bacterial display. *Curr Opin Struct Biol 17*, 474–80.

[36] Chen, G., Hayhurst, A., Thomas, J. G., Harvey, B. R., Iverson, B. L., and Georgiou, G. (2001) Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). *Nat Biotechnol 19*, 537–42.

[37] Kenrick, S., Rice, J., and Daugherty, P. (2007) Flow cytometric sorting of bacterial surface-displayed libraries. *Curr Protoc Cytom Chapter 4*, Unit4.6.

[38] Löfblom, J. (2011) Bacterial display in combinatorial protein engineering. *Biotechnol J.*

[39] van Bloois, E., Winter, R. T., Kolmar, H., and Fraaije, M. W. (2011) Decorating microbes: surface display of proteins on Escherichia coli. *Trends Biotechnol 29*, 79–86.

[40] Lu, Z., Murray, K. S., Van Cleave, V., LaVallie, E. R., Stahl, M. L., and McCoy, J. M. (1995) Expression of thioredoxin random peptide libraries on the Escherichia coli cell surface as functional fusions to flagellin: a system designed for exploring protein-protein interactions. *Biotechnology (N Y) 13*, 366–72.

[41] Rice, J. J. and Daugherty, P. S. (2008) Directed evolution of a biterminal bacterial display scaffold enhances the display of diverse peptides. *Protein Eng Des Sel 21*, 435–42.

[42] Rice, J. J., Schohn, A., Bessette, P. H., Boulware, K. T., and Daugherty, P. S. (2006) Bacterial display using circularly permuted outer membrane protein OmpX yields high affinity peptide ligands. *Protein Sci 15*, 825–36.

[43] Xu, Z. and Lee, S. Y. (1999) Display of polyhistidine peptides on the Escherichia coli cell surface by using outer membrane protein C as an anchoring motif. *Appl Environ Microbiol 65*, 5142–7.

[44] Chen, M., Khalid, S., Sansom, M. S. P., and Bayley, H. (2008) Outer membrane protein G: Engineering a quiet pore for biosensing. *Proc Natl Acad Sci U S A 105*, 6272–6277.

[45] Cruz, N., Le Borgne, S., Hernández-Chávez, G., Gosset, G., Valle, F., and Bolivar, F. (2000) Engineering the Escherichia coli outer membrane protein OmpC for metal bioadsorption. *Biotechnol Lett 22*, 623–629.

[46] Liang, B. and Tamm, L. K. (2007) Structure of outer membrane protein G by solution NMR spectroscopy. *Proc Natl Acad Sci U S A 104*, 16140–5.

[47] Casadaban, M. J. and Cohen, S. N. (1980) Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *J Mol Biol 138*, 179–207.

[48] Reetz, M. T., Kahakeaw, D., and Lohmer, R. (2008) Addressing the numbers problem in directed evolution. *Chembiochem 9*, 1797–804.

[49] Zemlin, M., Klinger, M., Link, J., Zemlin, C., Bauer, K., Engler, J. A., Schroeder, Jr, H. W., and Kirkham, P. M. (2003) Expressed murine and human CDR-H3 intervals of equal length exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures. *J Mol Biol 334*, 733–49.

[50] Fellouse, F. A., Wiesmann, C., and Sidhu, S. S. (2004) Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. *Proc Natl Acad Sci U S A 101*, 12467–72.

[51] Fellouse, F. A., Barthelemy, P. A., Kelley, R. F., and Sidhu, S. S. (2006) Tyrosine plays a dominant functional role in the paratope of a synthetic antibody derived from a four amino acid code. *J Mol Biol 357*, 100–14.

[52] Koide, A., Gilbreth, R. N., Esaki, K., Tereshko, V., and Koide, S. (2007) High-affinity single-domain binding proteins with a binary-code interface. *Proc Natl Acad Sci U S A 104*, 6632–7.

[53] Fellouse, F. A., Li, B., Compaan, D. M., Peden, A. A., Hymowitz, S. G., and Sidhu, S. S. (2005) Molecular recognition by a binary code. *J Mol Biol 348*, 1153–62.

[54] Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol 177*, 4121–30.

[55] Chen, M., Li, Q.-H., and Bayley, H. (2008) Orientation of the monomeric porin OmpG in planar lipid bilayers. *Chembiochem 9*, 3029–36.

[56] Conlan, S., Zhang, Y., Cheley, S., and Bayley, H. (2000) Biochemical and biophysical characterization of OmpG: A monomeric porin. *Biochemistry 39*, 11845–54.

[57] Syeda, R., Holden, M. A., Hwang, W. L., and Bayley, H. (2008) Screening blockers against a potassium channel with a droplet interface bilayer array. *J Am Chem Soc 130*, 15543–8.

[58] Leptihn, S., Thompson, J. R., Ellory, J. C., Tucker, S. J., and Wallace, M. I. (2011) In Vitro Reconstitution of Eukaryotic Ion Channels Using Droplet Interface Bilayers. *J Am Chem Soc 133*, 9370–9375.

[59] Cronan, J. E. (2006) A family of arabinose-inducible Escherichia coli expression vectors having pBR322 copy control. *Plasmid 55*, 152–7.

[60] Chung, C. T., Niemela, S. L., and Miller, R. H. (1989) One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A 86*, 2172–5.

[61] Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res 16*, 6127–45.

[62] Walker, B., Krishnasastry, M., Zorn, L., Kasianowicz, J., and Bayley, H. (1992) Functional expression of the alpha-hemolysin of Staphylococcus aureus in intact Escherichia coli and in cell lysates. Deletion of five C-terminal amino acids selectively impairs hemolytic activity. *J Biol Chem 267*, 10902–10909.

[63] Drummelsmith, J. and Whitfield, C. (2000) Translocation of group 1 capsular polysaccharide to the surface of Escherichia coli requires a multimeric complex in the outer membrane. *EMBO J 19*, 57–66.

[64] Conlan, S. (2003) Ph.D. thesis (Texas A&M University).

[65] Bubeck, P., Winkler, M., and Bautsch, W. (1993) Rapid cloning by homologous recombination in vivo. *Nucleic Acids Res 21*, 3601–2.

[66] Jones, D. H. and Howard, B. H. (1991) A rapid method for recombination and site-specific mutagenesis by placing homologous ends on DNA using polymerase chain reaction. *Biotechniques 10*, 62–66.

[67] Sambrook, J. (2001) *Molecular cloning: a laboratory manual.* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

[68] Abedi, M. R., Caponigro, G., and Kamb, A. (1998) Green fluorescent protein as a scaffold for intracellular presentation of peptides. *Nucleic Acids Res 26*, 623–30.

Chapter 4

VDAC for stochastic sensing

4.1 Introduction

Of the protein pores utilised for stochastic sensing, αHL has been the most extensively studied. However, alternative monomeric pores would be highly desirable. Monomeric pores are advantageous for protein engineering as multiple modifications at different sites, or modifications at a single point on the full protein, are easily achieved without the need to form and purify hetero-oligomers with defined stoichiometry.

To this end, OmpG has been previously studied and engineered to enhance its electrical properties for use in stochastic sensing [1]. However, while this 'quiet' OmpG (qOmpG) was an improvement over the wild type pore, it still has limitations. First, innate gating events, while significantly reduced, are still present. Second, one of the engineered modifications is a disulphide bond that restricts the conformational flexibility of extracellular loop 6. The presence of these engineered cysteine residues could complicate experiments

that use reversible cysteine chemistry for the attachment of sensing elements or for single molecule chemistry studies [2–9].

The voltage-dependent anion channel (VDAC) is one possible alternative. The structures of human and mouse VDAC-1 were recently solved, showing VDAC-1 to be a monomeric 19-strand β-barrel of anti-parallel β-sheets, with the first and last strands parallel [10–12]. Additionally, the pore possesses a large lumen of diameter 25 Å. VDAC-1 could therefore have potential for the stochastic sensing of substrates too large for the lumen of αHL or OmpG. Indeed, a previous study reported the translocation of double stranded DNA across lipid bilayers containing VDAC [13]. However, translocation was not directly observed by electrical recording.

4.1.1 Chapter overview

As a supplementary study, human VDAC-1 (hVDAC1) was investigated as a pore for stochastic sensing. The work detailed here describes the construction of plasmids encoding the hVDAC1 gene and an N-terminally truncated mutant, designed to eliminate voltage gating. Further, attempts at expression and refolding are described. Finally, production of the full-length pore by *in vitro* methods is reported. Although the final goal, to engineer hVDAC1 for stochastic sensing, was ultimately unsuccessful, pertinent observations are recorded here so they may be helpful to future endeavours.

4.2 Results

While the majority of studies of VDAC have used natively purified protein from mitochondria-rich tissue such as rat liver (e.g. Rostovtseva *et al.* [14]), recent structural studies used protein expressed recombinantly in *E. coli* and refolded from inclusion bodies [10–12]. Although there remains controversy as to whether these refolded proteins represent the true native structure of the channels that had previously been studied by biochemical and electrophysiological assays [15, 16], this expression strategy was chosen for its ease of use and for the ability to generate and express mutants for protein engineering.

4.2.1 Recombinant expression in *E. coli*

Plasmids encoding the desired proteins under control of the T7 promoter were prepared using commercial gene synthesis and standard cloning methods (Section 4.5). The hVDAC1 gene used was codon-optimised for *E. coli* and derived from the amino acid sequence deposited in GenBank (gi 4507879).

Expression time course

To confirm expression and optimise conditions, an expression time course was first performed. A 100 ml expression culture of BL21(DE3)pLysE cells harbouring pT7-H6-hVDAC1 was grown to $OD_{600} = 0.7$, at which point a 1 ml aliquot was retained and the culture was induced with IPTG. Further 1 ml aliquots were then retained at hourly intervals until 6 hours after induction. All aliquots were centrifuged immediately after collection and the cell pel-

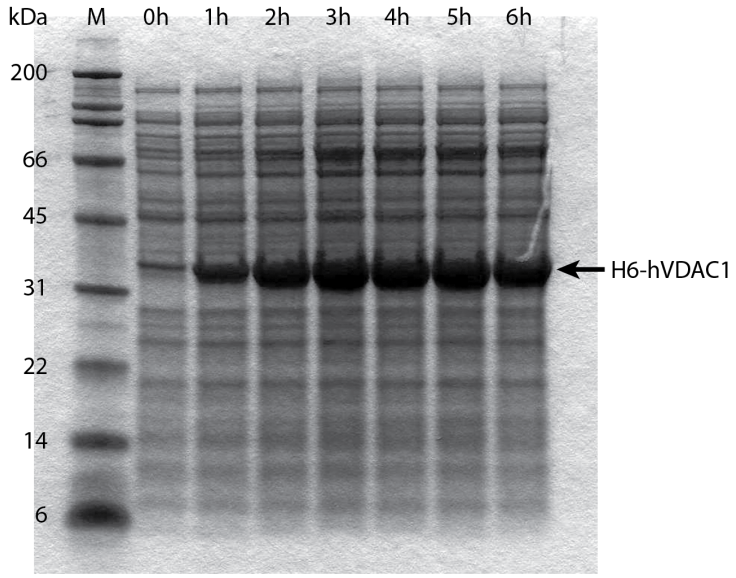


Figure 4.1: Expression time course of H6-hVDAC1 in BL21(DE3)pLysE. Aliquots were taken at hourly intervals immediately before induction and thereafter for six hours. Cells were harvested and lysed under denaturing conditions before being electrophoresed in a 12% bis-tris gel (Bio-Rad) in XT-MES buffer. Markers (M) are broad-range natural unstained protein standards (Bio-Rad). Protein bands were visualised with InstantBlue stain (Expedeon).

lets were frozen at −20°C. The cell pellets were then lysed under denaturing conditions (8 *M* urea) and analysed by SDS-PAGE (Figure 4.1).

H6-hVDAC1 was expressed rapidly and at high levels after induction, with protein levels reaching a plateau after 3–4 hours. Hence, an induction time of four hours was chosen for future expression cultures.

Expression and purification of H6-hVDAC1

Initial expression attempts followed the procedure of Ujwal *et al.* with some modifications (cf. Materials and methods) [10]. However, it was found that refolding of the protein by dialysis as described in their procedure resulted in complete protein precipitation. This is likely due to the reported cut-off of the dialysis tubing being too small to allow the passage of lauryldimethy-

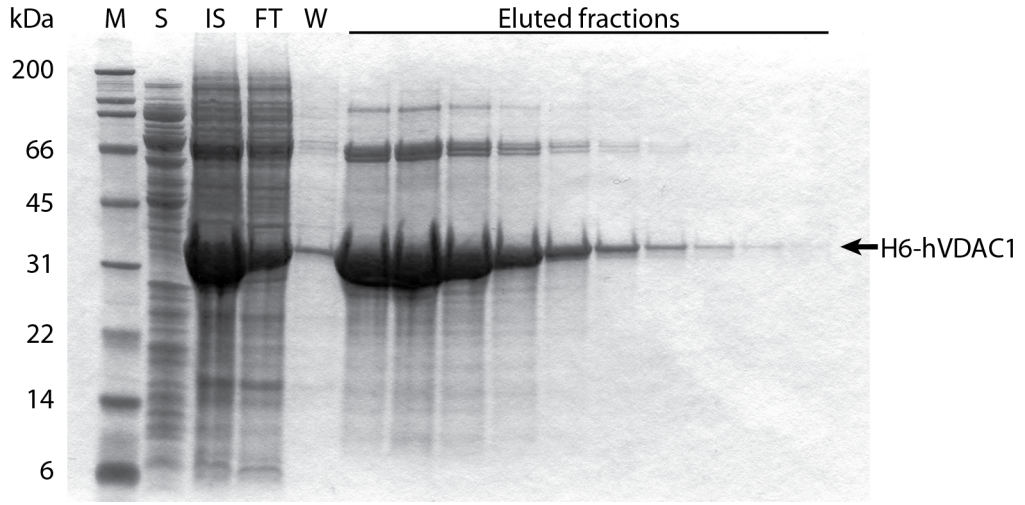


Figure 4.2: Purification of H6-hVDAC1 by Ni-NTA affinity chromatography. (S) Soluble lysate fraction. (IS) Insoluble lysate fraction, solubilised in buffer containing 8 *M* urea. (FT) Column flow-through. (W) Column wash fraction. Markers (M) are broad-range natural unstained protein standards (Bio-Rad). Samples were analysed by SDS-PAGE on a 12% bis-tris gel (Bio-Rad) in XT-MES buffer. Protein bands were visualised with InstantBlue stain (Expedeon).

lamine oxide (LDAO) micelles into the tubing and hence maintaining solubilisation of the protein. Instead, a modified procedure was followed where H6-hVDAC1 was first purified by Ni-NTA affinity chromatography, then deliberately precipitated by dialysis against buffer that did not contain denaturing agent, before being resolubilised in guanidinium hydrochloride (GdnHCl). The resolubilised protein was refolded by drop-wise dilution into a stirred buffer solution containing 2% LDAO, in a manner similar to the procedure of Hiller *et al.*, [12]. Ni-NTA purification yielded approximately 95% pure H6-hVDAC1, although higher bands were observed that are likely to be disulphide-linked oligomers, as the samples were not prepared under reducing conditions (Figure 4.2). The expression level was approximately 100 mg/l. After refolding by drop-wise dilution, no precipitate was visible.

Reconstitution of the refolded protein in planar lipid bilayers was then

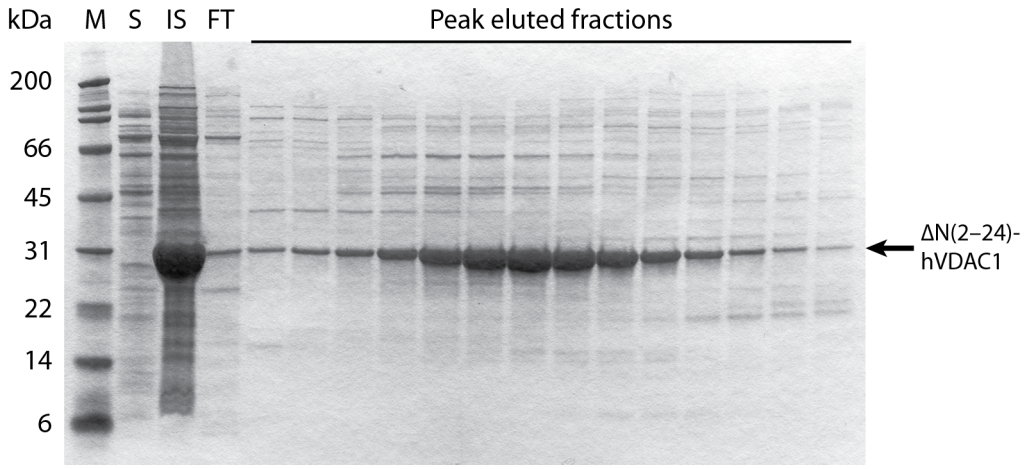
attempted. However, although tested under a range of conditions, for example in the presence of cholesterol and/or Mg^{2+} as in Hiller *et al.* [12], no reproducible channel activity was observed.

Expression of N-terminally truncated hVDAC1

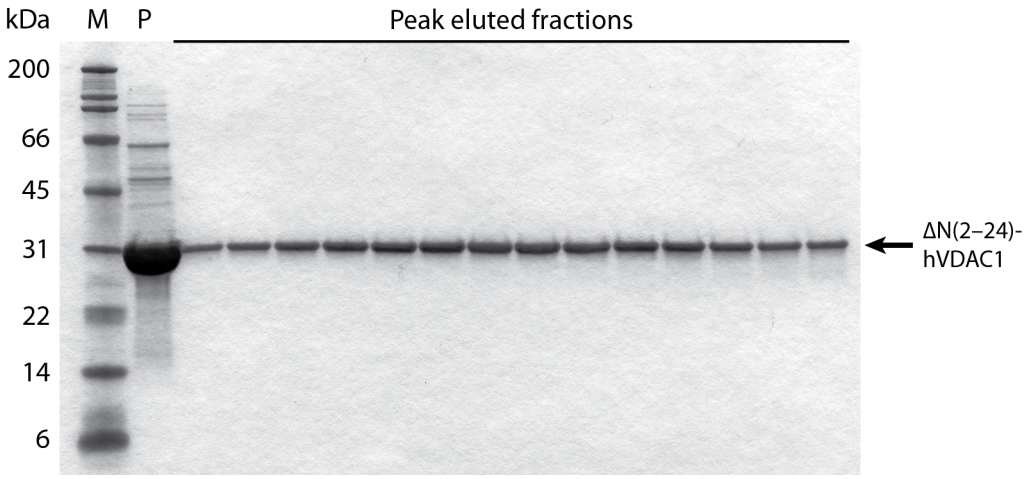
Concurrently with the expression and purification H6-hVDAC1, the expression and purification of ΔN(2–24)-hVDAC1 was attempted. This mutant was designed so that the α-helical N-terminal segment of the wild type protein was removed. This segment is thought to be responsible for the voltage gating of VDAC-1 [17], and hence it was hypothesised that removal of this segment may produce a pore that exhibits a single conductance level without voltage gating. Indeed, after these experiments had been performed, Abu-Hamad *et al.* showed that a similar truncation mutant, ΔN(2–26)-mVDAC1, but produced in a recombinant yeast system, did not exhibit voltage gating [18]. Such a pore would open up the possibility of further engineering for stochastic sensing.

ΔN(2–24)-hVDAC1 was expressed and purified from inclusion bodies in a similar manner to H6-hVDAC1. However, as difficulties obtaining activity from H6-hVDAC1 arose, it was decided not to use a hexahistidine tag for affinity chromatography and instead purify the protein by ion exchange and gel filtration, in case the hexahistidine tag was responsible for the lack of channel-forming activity. After two rounds of cation exchange and one round of gel filtration the protein was purified to homogeneity, with a final yield of 4.73 mg from a 100 ml culture (Figure 4.3).

Purified protein was refolded by drop-wise dilution as for H6-hVDAC1.



(a) Ion exchange



(b) Gel filtration

Figure 4.3: Purification of ΔN(2–24)-hVDAC1. (a) Capture ion exchange using a 5 ml HiTrap SP FF column (GE Healthcare). (S) Soluble lysate fraction. (IS) Insoluble lysate fraction, solubilised in buffer containing 8 *M* urea. (FT) Column flow-through. (b) Final gel purification using a HiLoad 26/600 Superdex 200 pg column (GE Healthcare). (P) Pooled fractions from second ion exchange purification (not shown). Markers (M) are broad-range natural unstained protein standards (Bio-Rad). Samples were analysed by SDS-PAGE on a 12% bis-tris gel (Bio-Rad) in XT-MES buffer. Protein bands were visualised with InstantBlue stain (Expedeon).

Reconstitution of the refolded protein in planar lipid bilayers was then attempted. However, as for H6-hVDAC1, no reproducible channel activity was observed.

4.2.2 *In vitro* protein expression

In vitro protein expression provides several advantages over *in vivo* expression. First, cell-free expression can avoid problems of host toxicity caused by heterologous protein expression, commonly experienced with many integral membrane proteins. Second, production of protein is fast as it does not require cell transfection or cell culturing—protein is usually produced in just a few hours. Third, it is possible to add other components or reagents as required to the cell extract to optimise expression or folding conditions. For example, it is possible to add certain detergents or liposomes to enable correct folding and solubilisation of membrane proteins. While the most commonly used *E. coli*-derived system is not capable of post-translational modifications, other systems are capable of glycosylation (e.g. insect-cell derived systems), and others such as rabbit reticulocyte systems are capable when canine microsomal membranes are added [19–22].

In vitro transcription and translation of hVDAC-1 was investigated to see whether it was possible to produce functional protein. This would present a significantly faster procedure than the existing method of expression into inclusion bodies in *E. coli* followed by purification and refolding.

E. coli S30 extract system

Protein expression was first tested in an *E. coli*-derived S30 extract system (*E. coli* T7 S30 Extract System for Circular DNA from Promega). S30 systems typically produce the highest protein yields, however they are not capable of post-translational modifications that may be required for correct folding, or for producing soluble protein, when expressing mammalian proteins. A C-terminal hexahistidine tagged construct was used for these experiments; both N- and C-terminal tags have been used in the literature and hence the change was simply due to convenience.

An initial test confirmed that while hVDAC1-H6 was produced in the S30 extract system, it was insoluble. It has been shown possible to solubly express some mammalian proteins in S30 systems in the presence of detergent [23–25]. Hence, a screen of several detergents was performed to see if any facilitated the soluble expression of hVDAC1-H6 *in vitro*.

The detergents screened were Triton X-100, dodecylmaltoside (DDM), Brij-58, and Genapol X-080. None of these detergents were found to be inhibitory to protein expression. In addition, LDAO was tested but was found to be inhibitory at all practical concentrations (i.e. above the CMC, data not shown). Detergent concentrations used were guided by prior values in the literature [21], and were: Triton X-100, 0.2%; DDM, 0.1%; Brij-58, 1.5%; Genapol X-080, 1%. Expression in IVTT reactions was analysed by SDS-PAGE and autoradiography. To establish the fraction of soluble protein, samples of supernatant from centrifuged IVTT reactions were also analysed (Figure 4.4).

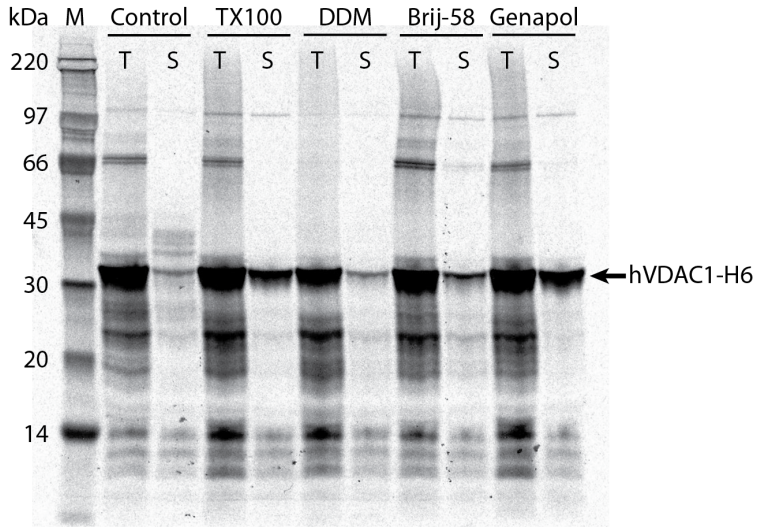


Figure 4.4: Detergent screen for *in vitro* expression of hVDAC1-H6 in S30 extract. IVTT reactions were performed in the presence of detergents as marked (TX100 = Triton X-100, Genapol = Genapol X-080, DDM = Dodecylmaltoside). The control was supplemented with water instead of detergent. For each reaction, aliquots from the total reaction (T), and the supernatant after centrifugation (S) were analysed by SDS-PAGE on a 10% bis-tris gel (Bio-Rad) in XT-MES buffer. The markers (M) are ^{14}C-methylated protein standards (Perkin Elmer).

Of the detergents screened, all but DDM increased the amount of soluble hVDAC1-H6 in the IVTT reaction. Genapol X-080 had the strongest effect, with Triton X-100 only slightly less effective. hVDAC1 is not stable to SDS and so purification by SDS-PAGE in a similar fashion to αHL was not possible. First, reconstitution of channels into planar bilayers and DIBs from the unpurified reaction supernatant was attempted. However, no channel activity was observed. Additionally, purification by Ni-NTA affinity chromatography was attempted prior to reconstitution, but the expressed protein did not bind to the matrix even under mild binding conditions. Hence, although it was shown that detergent maintains protein solubility, it is not clear whether hVDAC1-H6 expressed in S30 extract is functionally active. Further, it is possible that the hexahistidine tag becomes either inaccessi-

ble to matrix binding (perhaps due to misfolding), or may be proteolytically degraded by endogenous proteases that may be present in the S30 extract.

Some time after these experiments were performed, Deniaud *et al.* published a study where hVDAC1 was produced using the Roche RTS 100 and 500 bacterial cell-free expression systems [26]. The authors also found hVDAC1 to be aggregated when detergent was not present, and screened the detergents $C_{12}E_8$, DDM, LDAO, foscholine-12, cymal-5 and C_8E_4 for the maintenance of soluble protein. Of these detergents, they found DDM to be optimal and were able to purify sufficient protein to obtain crystals that diffracted with low resolution (8 Å).

Insect cell extract IVTT

As it is possible that functional protein was not obtained from the S30 system due to a lack of post-translational modification, an insect cell-derived IVTT system was also tested. Expression of hVDAC1-H6 was examined in the TNT T7 Insect Cell Extract Protein Expression System (Promega). Initial attempts to express hVDAC1-H6 from pT7-hVDAC1-H6 produced negligible quantities of protein (data not shown). Hence, the hVDAC1-H6 gene was cloned into pF25A (Promega), a vector optimised for insect cell extract IVTT that bears 5' and 3' untranslated regions (UTRs) for maximal protein expression. Reactions expressing hVDAC1-H6 and luciferase (positive control plasmid supplied with kit) were analysed by SDS-PAGE and autoradiography. To establish the fraction of soluble protein, samples of supernatant from centrifuged IVTT reactions were also analysed (Figure 4.5).

hVDAC1-H6 was well expressed from the pF25A vector. Proteins were

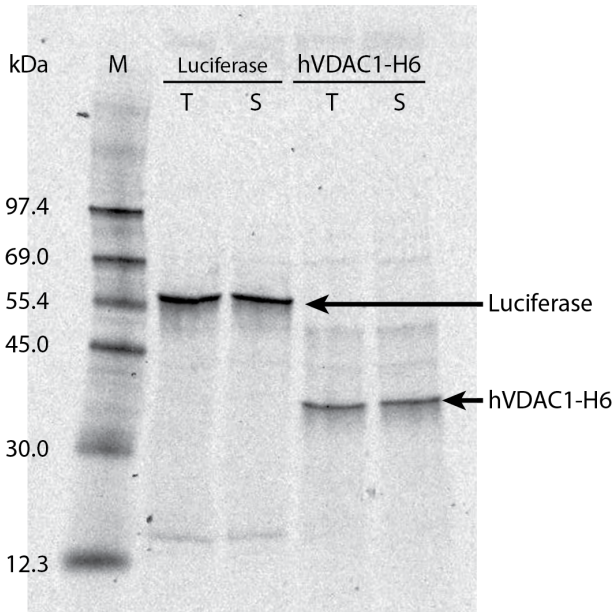


Figure 4.5: Insect cell extract expression of hVDAC1-H6. Expression from the luciferase ICE T7 control vector and pF25A-hVDAC1-H6 was analysed by SDS-PAGE and autoradiography. For each reaction, aliquots from the total reaction (T), and the supernatant after centrifugation (S) were analysed by SDS-PAGE on a 4–12% bis-tris gel (Bio-Rad) in XT-MES buffer. The markers are ^{14}C-methylated protein standards (Perkin Elmer).

radiolabeled with L-[^{35}S]methionine, and hence intensity is dependent on the number of methionine residues present in the protein—in this case, the luciferase contains fourteen methionine residues, versus three for hVDAC1-H6. Notably, the expressed hVDAC1-H6 protein was soluble, with essentially all remaining in the supernatant after centrifugation.

Reconstitution of the expressed hVDAC1-H6 protein was attempted using planar lipid bilayers and DIBs. However, once more it was not possible to observe channel activity. As with *E. coli*-based IVTT in the presence of detergent, it may be that although the protein is soluble, it is not correctly folded for membrane insertion. It may be that hVDAC1 requires chaperones or other accessory proteins for correct functional folding that are not present in the IVTT systems tested.

4.3 Conclusions

Expression plasmids encoding two versions of the human VDAC-1 channel were created, one encoding a hexahistidine-tagged full length protein, the other encoding an N-terminal truncation of the wild type protein. Both proteins were expressed as inclusion bodies in *E. coli* with high yield. These were readily purified under denaturing conditions, and then refolding was performed by drop-wise dilution into buffer containing detergent (LDAO).

Although this appeared to be successful, the reconstitution of purified channels into artificial bilayer systems (Montal-Mueller planar lipid bilayers or DIBs) was unsuccessful. This is particularly unfortunate as the procedures followed here are essentially as described in the literature [10, 12]. Al-

though these papers show some brief electrophysiological data, there remains some controversy over whether the refolded protein represents the channel-forming species observed when natively purified protein is studied [16]. It is possible that there may be some conditions, not adequately documented in the literature, which induce a conformational change in the protein to the channel-forming species but these have not been found in this work.

Attempts to produce hVDAC1 by *in vitro* transcription and translation were also made in *E. coli* and insect-cell derived systems. hVDAC1-H6 was successfully expressed in good yield in an *E. coli* S30 IVTT system. However, the translation product aggregated and was insoluble. Several detergents were screened and Triton X-100, Genapol X-080 and Brij-58 were all shown to increase the fraction of soluble protein produced, with Triton X-100 and Genapol X-080 being the most effective. Unfortunately, as with protein expressed heterologously in *E. coli* cells, attempts to reconstitute protein produced by IVTT were unsuccessful.

IVTT production was attempted in an insect-cell derived system to see if the lack of channel-forming activity was due to the lack of post-translational modification in *E. coli* IVTT systems. H6-hVDAC1 was produced in good yield relative to a control (luciferase). Notably, the translation product showed no sign of aggregation and was fully soluble. However, once more reconstitution into lipid bilayers was unsuccessful.

4.4 Future developments

As it remains unclear whether the production of recombinant VDAC1 in *E. coli* is a viable method for producing *bona fide* channel-forming species, future work would require the determination of conditions that allow the production of protein that can be easily and reproducibly reconstituted into lipid bilayers. This would require extensive optimisation of refolding conditions, if indeed it proves possible to obtain the native conformation via this route. Should this not be the case, alternative expression systems that allow the native expression of hVDAC1 and its mutants may prove useful. For example, a VDAC1-deficient *Saccharomyces cerevisiae* strain was used to express recombinant murine VDAC1, which could then be purified from the mitochondrial fraction of the yeast cells [27]. Once such a system had been established, protein engineering efforts to generate sensing elements within the pore could then begin.

4.5 Materials and methods

Chemicals and oligonucleotides were from Sigma Aldrich unless otherwise stated. Enzymes and DNA ladders were from New England Biolabs. Cloning was performed in XL-10 Gold cells (Stratagene) that were made chemically competent by the method of Chung *et al.*, [28]. Chemically competent BL21-(DE3)pLysE cells (Merck) were prepared by the same method.

Routine plasmid purification was performed using either Qiaprep spin miniprep or Qiagen Plasmid Maxi kits as directed by the accompanying pro-

tocols. DNA quantitation was performed by UV-visible absorption spectroscopy using a NanoDrop 1000 (Thermo Scientific). Water used was purified by reverse osmosis to a final resistivity of 18.2 MΩ cm (Millipore).

4.5.1 Plasmids

All plasmid insert sequences were verified by sequencing (Source BioScience, Oxford). pT7-SC3 and pT7-SC5 were gifts from Stephen Cheley.

pUC57-hVDAC1

This plasmid was supplied as a custom product by GenScript USA Inc. The human VDAC-1 (hVDAC1) amino acid sequence (gi 4507879) was used to generate an encoding nucleotide sequence, codon-optimised for *E. coli* using DNAWorks [29]. Flanking NcoI and HindIII sites were designed at the 5'- and 3'- termini respectively of the hVDAC1-encoding gene to facilitate subsequent subcloning. The resulting gene was synthesised by GenScript and cloned into pUC57 at the EcoRV restriction site.

pT7-H6-hVDAC1

pUC57-hVDAC1 was double-digested with NcoI and HindIII. The excised hVDAC1-encoding fragment was then purified by electrophoresis in a 1% TAE agarose gel. pT7-SC5 was also digested with NcoI and HindIII and the vector fragment was gel purified. pT7-SC5 encodes a hexahistidine tag between NdeI and NcoI sites such that when a gene is cloned between NcoI and HindIII the resulting ORF encodes the cloned protein plus an N-terminal hex-

ahistidine tag. The hVDAC-1 encoding fragment and vector fragment were ligated by T4 DNA ligase. The ligation reaction was transformed into chemically competent XL-10 gold cells, which were then plated on LB carbenicillin plates and incubated overnight at 37°C. Colonies were picked, plasmid DNA was isolated, and insert sequences were verified by sequencing. Large scale DNA isolation by plasmid maxi-prep was then performed using a successful clone, and the isolated DNA was stored at a concentration of 400 ng/μl at -20°C.

pT7-hVDAC1-H6

The hVDAC1 gene from pT7-H6-hVDAC1 was amplified by PCR using primers that changed the 5' NcoI site to an NdeI site and placed a SpeI site immediately before the position of the first stop codon in the template sequence. The forward primer was 5'-CGGCGGCATATGGCGGTTCCACC AACGTATGCCGAC-3' and the reverse primer was 5'-GGGCGCACTAGT CGCTTGAAACTCCAGGCCCAGACC-3'. The PCR product was purified by electrophoresis in a 1% TAE agarose gel. The purified PCR product was double-digested with NdeI and SpeI and the gene fragment was gel purified as before.

pT7-SC3 encodes a hexahistidine tag between SpeI and HindIII sites such that when a gene is cloned between NdeI and SpeI the resulting ORF encodes the cloned protein plus a C-terminal hexahistidine tag. pT7-SC3 was double-digested with NdeI and SpeI and the vector fragment was purified by electrophoresis on a 0.8% TAE agarose gel. The purified vector fragment was ligated with the purified, digested gene fragment with T4 DNA ligase.

The ligation reaction was transformed into chemically competent XL-10 gold cells. Clones were selected and verified by sequencing as above. DNA isolation by plasmid maxi-prep was then performed using a successful clone, and the isolated DNA was stored at a concentration of 400 ng/μl at -20°C.

pT7-ΔN(2–24)hVDAC1

This N-terminal truncation mutant was generated by *in vivo* homologous recombination [30, 31], using pT7-H6-hVDAC1 as the template. The forward mutagenic primer was 5'-GGCTTAATTAAGCTGGACCTGAAGACG-3' and the reverse mutagenic primer was 5'-GTCCAGCTTAATTAAGCCC ATATGTATATCTCCTTC-3'. The forward and reverse non-mutagenic primers were SC47 and SC46, respectively.

pF25A-hVDAC1-H6

The hVDAC1-H6 gene from pT7-hVDAC1-H6 was amplified by PCR. The primers used were designed such that the 5' NdeI site was changed to an NcoI site, and a KpnI site was introduced immediately after the second stop codon at the 3' end. The forward primer was 5'-CGGCGGCCATGGCGG TTCCACCAACGTATGCCGAC-3' and the reverse primer was 5'-GGGCG CGGTACCTCATTAATGATGGTGATGGTGGTG-3'. The PCR product was purified by electrophoresis in a 1% TAE agarose gel. The purified PCR product was double-digested with NcoI and KpnI and the gene fragment was gel purified as before.

pF25A was double-digested with NcoI and KpnI and the vector fragment was purified by electrophoresis on a 0.8% TAE agarose gel. The purified

vector fragment was ligated with the purified, digested gene fragment with T4 DNA ligase. The ligation reaction was transformed into chemically competent XL-10 gold cells. Clones were selected and verified by sequencing as above. DNA isolation by plasmid maxi-prep was then performed using a successful clone, and the isolated DNA was stored at a concentration of 1 μg/μl at -20°C.

4.5.2 Recombinant protein expression and purification

H6-hVDAC1

pT7-H6-hVDAC1 DNA was transformed into chemically competent BL21-(DE3)pLysE cells, which were plated on LB agar supplemented with carbenicillin and chloramphenicol and incubated at 37°C overnight. All media hereafter were supplemented with carbenicillin and chloramphenicol. A single colony was used to inoculate 10 ml of LB, which was incubated overnight at 37°C with shaking at 200 rpm. 5 ml of this overnight inoculum was added to 100 ml of LB in a 500 ml baffled culture flask. The culture was incubated at 37°C with shaking at 200 rpm until it reached $OD_600 = 0.7$. IPTG was added to a final concentration of 1 mM and the culture was incubated at 37°C with shaking at 200 rpm for four hours. The cells were then harvested by centrifugation at 3,000$\times g$ for twenty minutes. The cell pellet was frozen at -80°C.

The frozen pellet was thawed on ice and resuspended in 10 ml of lysis buffer (20 mM Tris.HCl pH 8.0, 50 mM NaCl), before sonication on ice with a probe sonicator for thirty seconds, repeated five times with thirty second

rest intervals between sonication. Magnesium chloride was then added to a final concentration of 5 mM, and 500 units of Benzonase nuclease (Merck) were added, whereupon the mixture was incubated on ice for ten minutes. The lysate was then cleared by centrifugation at 29,000×g for 45 minutes.

The inclusion body pellet was washed with buffer (50 mM Tris.HCl pH 8.0, 500 mM, 10 mM imidazole) and then re-pelleted by centrifugation at 29,000×g for 45 minutes. The pellet was solubilised in binding buffer (50 mM Tris.HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 8 M urea) and then incubated on ice for twenty minutes. Any remaining insoluble matter was then removed by centrifugation at 29,000×g for thirty minutes.

The supernatant was added to a 50% slurry of Ni-NTA fast-flow (Qiagen) in binding buffer. The mixture was incubated for one hour at 4°C with gentle rotation. After incubation, the mixture was poured into a plastic disposable column (Bio-Rad) and the flow-through was collected. The column was then washed with 2×10 ml of binding buffer. Protein was then eluted in 10×1 ml fractions of elution buffer (37.5 mM Tris.HCl pH 8.0, 375 mM NaCl, 250 mM imidazole, 8 M urea). Aliquots were taken at relevant stages of the purification and analysed by SDS-PAGE on a 12% bis-tris gel (Bio-Rad) in XT-MES buffer at 200 V for 45 minutes. The gel was stained with InstantBlue (Expedeon), a Coomassie-based stain, dried under vacuum and imaged using a flatbed scanner.

The first three eluted fractions were pooled and precipitated by dialysis in a FloatALyzer dialysis device (Spectrum) with 10 kDa MW cutoff against 2×500 ml of 20 mM Tris.HCl pH 8.0, 50 mM NaCl, and 1 mM DTT. The precipitate was collected by centrifugation at 29,000×g for one hour and

then dissolved in 20 mM MOPS pH 7.0, 150 mM NaCl, 1 mM DTT and 6 M GdnHCl. The final volume was 1 ml. The solubilised protein was then refolded by drop-wise dilution (∼50 μl/min) into 10 ml of rapidly stirring refolding buffer (20 mM MOPS pH 7.0, 150 mM NaCl, 1 mM DTT and 1% LDAO) at 4°C. After dilution the refolding solution was incubated for a further four hours with stirring, and then centrifuged at 29,000×g for one hour to remove any insoluble material.

ΔN(2–24)hVDAC1

pT7-ΔN(2–24)hVDAC1 was transformed into chemically competent BL21-(DE3)pLysE cells. The expression culture was then prepared and harvested as above. The frozen pellet was thawed on ice and resuspended in 10 ml of lysis buffer (20 mM MES titrated to pH 6.5 with NaOH), before sonication on ice with a probe sonicator for thirty seconds, repeated five times with thirty second intervals between sonication. Magnesium chloride was then added to a final concentration of 5 mM, and 500 units of Benzonase nuclease were added, whereupon the mixture was incubated on ice for ten minutes. The lysate was then cleared by centrifugation at 29,000×g for 45 minutes.

The inclusion body pellet was washed twice with buffer (20 mM MES pH 6.5, 1 M urea) with centrifugation at 29,000×g for 45 minutes after each wash. The pellet was then solubilised in 4 ml of buffer A (20 mM MES pH 6.5, 8 M urea, 1 mM DTT) and incubated on ice for one hour. Any remaining insoluble matter was then removed by centrifugation at 29,000×g for thirty minutes.

Capture cation exchange chromatography was performed on an Akta Pu-

rifier 10 system (GE Healthcare) using a 5 ml HiTrap SP FF column equilibrated with buffer A. ΔN(2–24)hVDAC1 was eluted during a 0–50% gradient over ten column volumes with buffer B (buffer A + 1 *M* NaCl) at about 11% buffer B. Eluted protein was collected in 1 ml fractions. Peak fractions were pooled, yielding 11.6 mg of protein. This protein was precipitated by dialysis in a FloatALyzer dialysis device (Spectrum) with 10 kDa MW cutoff against 2×1l of 20 m*M* MES pH 6.5. The precipitate was collected by centrifugation at 29,000×*g* for one hour and then dissolved in 5 ml of buffer A.

A second cation exchange purification was performed, using a 5 ml HiTrap SP HP column equilibrated with buffer A. ΔN(2–24)-hVDAC1 was eluted during a 0–50% gradient over twenty column volumes with buffer B (buffer A + 1 *M* NaCl) at about 11% buffer B. Eluted protein was collected in 0.5 ml fractions. Peak fractions were pooled and concentrated using an Amicon Ultra-15 centrifugal concentrator with a MW cutoff of 10 kDa, to a final volume of 1 ml.

A final gel filtration step was performed, using a HiLoad 26/600 Superdex 200 prep-grade column (column volume approximately 320 ml) in gel filtration buffer (20 m*M* MES pH 6.5, 150 m*M* NaCl, 1 m*M* DTT, 8 *M* urea). The peak fractions were pooled, yielding 4.7 mg of pure protein, and concentrated using an Amicon Ultra-15 centrifugal concentrator with a MW cutoff of 10 kDa, to a final volume of 1 ml. Buffer exchange was then performed using a PD-10 desalting column (GE Healthcare) equilibrated in 20 m*M* MOPS pH 7.0, 150 m*M* NaCl, 1 m*M* DTT, 6 *M* GdnHCl. 1 ml of the protein solution was then refolded by drop-wise dilution (∼50 μl/min) into 10 ml of rapidly stirring refolding buffer (20 m*M* MOPS pH 7.0, 150 m*M* NaCl, 1 m*M*

DTT and 2% LDAO) at 4°C. After dilution the refolding solution was incubated for a further four hours with stirring, and then centrifuged at 29,000×g for one hour to remove any insoluble material.

Aliquots were taken at all relevant stages of the purifications and analysed by SDS-PAGE on 12% bis-tris gels (Bio-Rad) in XT-MES buffer at 200 V for 45 minutes. The gels were stained with InstantBlue (Expedeon), a Coomassie-based stain, dried under vacuum and imaged using a flatbed scanner.

4.5.3 *In vitro* transcription and translation

Detergent screen using S30 system

E. coli in vitro transcription and translation reactions were performed using the *E. coli* T7 S30 Extract System for Circular DNA (Promega). All reactions had a final volume of 25 μl and contained: 10 μl S30 premix (supplied with kit), 2.5 μl complete amino acid mix (prepared by mixing together 'amino acid mixture minus leucine' and 'amino acid mixture minus methionine' supplied with kit), 1 μl L-[^{35}S]methionine (1175 Ci/mmol, 10 mCi/ml, MP Biomedicals), 2.66 μl pT7-H6-hVDAC1 plasmid DNA (400 ng/μl), and 7.5 μl S30 extract. The S30 extract was pre-treated with rifampicin (1 μg/ml final) before addition to the IVTT reaction, to suppress expression from endogenous non-T7-promoters in the extract. 1.33 μl of stock solutions of the following detergents were added to their respective reactions (final concentrations in reaction in brackets): Triton X-100 (0.2%), dodecylmaltoside (DDM) (0.1%), Brij-58 (1.5%), Genapol X-080 (1%). 1.33 μl of nuclease-free water

was added to the control reaction. Reactions were incubated for one hour at 37°C.

After incubation, 2 μl aliquots of each reaction were retained for analysis by SDS-PAGE. The remaining reactions were centrifuged at 25,000$\times g$ for thirty minutes at 4°C. 2 μl aliquots were taken from each of the resulting supernatants. To each aliquot, 2 μl of 2X and 16 μl of 1X Laemmli sample buffer (1X = 62.5 mM Tris.HCl pH 6.8, 2.3% w/v SDS, 5% v/v β-mercaptoethanol, 10% w/v glycerol) was added. 10 μl from each prepared sample was then analysed by SDS-PAGE on a 12% bis-tris acrylamide gel (Bio-Rad) in XT-MES buffer at 200 V for 45 minutes. The gel was then dried under vacuum and autoradiographed.

Insect cell extract IVTT

Insect cell IVTT reactions were performed using the TNT T7 Insect Cell Extract Protein Expression System (Promega). Each reaction contained: 20 μl TNT T7 ICE master mix (supplied with kit), 2 μl nuclease-free water, 1 μl L-[^{35}S]methionine (1175 Ci/mmol, 10 mCi/ml, MP Biomedicals), and 2 μl plasmid DNA (1 μg/μl). Reactions were incubated at 29 °C for four hours.

After incubation, 2 μl aliquots of each reaction were retained for analysis by SDS-PAGE. The remaining reactions were centrifuged at 25,000$\times g$ for thirty minutes at 4°C. 2 μl aliquots were taken from each of the resulting supernatants. To each aliquot, 2 μl of 2X and 16 μl of 1X Laemmli sample buffer (1X = 62.5 mM Tris.HCl pH 6.8, 2.3% w/v SDS, 5% v/v β-mercaptoethanol, 10% w/v glycerol) was added. 10 μl from each prepared

sample was then analysed by SDS-PAGE on a 4–12% bis-tris acrylamide gel (Bio-Rad) in XT-MES buffer at 200 V for 35 minutes. The gel was then dried under vacuum and autoradiographed.

Bibliography

[1] Chen, M., Khalid, S., Sansom, M. S. P., and Bayley, H. (2008) Outer membrane protein G: Engineering a quiet pore for biosensing. *Proc Natl Acad Sci U S A 105*, 6272–6277.

[2] Shin, S.-H., Steffensen, M. B., Claridge, T. D. W., and Bayley, H. (2007) Formation of a chiral center and pyrimidal inversion at the single-molecule level. *Angew Chem Int Ed Engl 46*, 7412–6.

[3] Shin, S.-H., Luchian, T., Cheley, S., Braha, O., and Bayley, H. (2002) Kinetics of a reversible covalent-bond-forming reaction observed at the single-molecule level. *Angew Chem Int Ed Engl 41*, 3707–9; 3523.

[4] Luchian, T., Shin, S.-H., and Bayley, H. (2003) Kinetics of a three-step reaction observed at the single-molecule level. *Angew Chem Int Ed Engl 42*, 1926–9.

[5] Luchian, T., Shin, S.-H., and Bayley, H. (2003) Single-molecule covalent chemistry with spatially separated reactants. *Angew Chem Int Ed Engl 42*, 3766–71.

[6] Hammerstein, A. F., Shin, S.-H., and Bayley, H. (2010) Single-molecule kinetics of two-step divalent cation chelation. *Angew Chem Int Ed Engl 49*, 5085–90.

[7] Shin, S.-H. and Bayley, H. (2005) Stepwise growth of a single polymer chain. *J Am Chem Soc 127*, 10462–3.

[8] Xie, H., Braha, O., Gu, L.-Q., Cheley, S., and Bayley, H. (2005) Single-molecule observation of the catalytic subunit of cAMP-dependent protein kinase binding to an inhibitor peptide. *Chem Biol 12*, 109–120.

[9] Movileanu, L., Howorka, S., Braha, O., and Bayley, H. (2000) Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nat Biotechnol 18*, 1091–1095.

[10] Ujwal, R., Cascio, D., Colletier, J.-P., Faham, S., Zhang, J., Toro, L., Ping, P., and Abramson, J. (2008) The crystal structure of mouse VDAC1 at 2.3 A resolution reveals mechanistic insights into metabolite gating. *Proc Natl Acad Sci U S A 105*, 17742–17747.

[11] Bayrhuber, M., Meins, T., Habeck, M., Becker, S., Giller, K., Villinger, S., Vonrhein, C., Griesinger, C., Zweckstetter, M., and Zeth, K. (2008) Structure of the human voltage-dependent anion channel. *Proc Natl Acad Sci U S A 105*, 15370–15375.

[12] Hiller, S., Garces, R. G., Malia, T. J., Orekhov, V. Y., Colombini, M., and Wagner, G. (2008) Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science 321*, 1206–1210.

[13] Szabò, I., Bàthori, G., Tombola, F., Coppola, A., Schmehl, I., Brini, M., Ghazi, A., De Pinto, V., and Zoratti, M. (1998) Double-stranded DNA can be translocated across a planar membrane containing purified mitochondrial porin. *FASEB J 12*, 495–502.

[14] Rostovtseva, T. K., Sheldon, K. L., Hassanzadeh, E., Monge, C., Saks, V., Bezrukov, S. M., and Sackett, D. L. (2008 Dec 2) Tubulin binding blocks mitochondrial voltage-dependent anion channel and regulates respiration. *Proc Natl Acad Sci U S A 105*, 18746–18751.

[15] Hiller, S., Abramson, J., Mannella, C., Wagner, G., and Zeth, K. (2010) The 3D structures of VDAC represent a native conformation. *Trends Biochem Sci 35*, 514–21.

[16] Colombini, M. (2009) The published 3D structure of the VDAC channel: native or not? *Trends Biochem Sci 34*, 382–9.

[17] Song, J., Midson, C., Blachly-Dyson, E., Forte, M., and Colombini, M. (1998) The sensor regions of VDAC are translocated from within the membrane to the surface during the gating processes. *Biophys J 74*, 2926–44.

[18] Abu-Hamad, S., Arbel, N., Calo, D., Arzoine, L., Israelson, A., Keinan, N., Ben-Romano, R., Friedman, O., and Shoshan-Barmatz, V. (2009) The VDAC1 N-terminus is essential both for apoptosis and the protective effect of anti-apoptotic proteins. *J Cell Sci 122*, 1906–1916.

[19] He, M. (2008) Cell-free protein synthesis: applications in proteomics and biotechnology. *N Biotechnol 25*, 126–32.

[20] Hovijitra, N. T., Wuu, J. J., Peaker, B., and Swartz, J. R. (2009) Cell-free synthesis of functional aquaporin Z in synthetic liposomes. *Biotechnol Bioeng 104*, 40–9.

[21] Klammt, C., Schwarz, D., Fendler, K., Haase, W., Dötsch, V., and Bernhard, F. (2005) Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *FEBS J 272*, 6024–38.

[22] Katzen, F., Peterson, T. C., and Kudlicki, W. (2009) Membrane protein expression: no cells required. *Trends Biotechnol 27*, 455–60.

[23] Berrier, C., Park, K.-H., Abes, S., Bibonne, A., Betton, J.-M., and Ghazi, A. (2004) Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry 43*, 12585–91.

[24] Elbaz, Y., Steiner-Mordoch, S., Danieli, T., and Schuldiner, S. (2004) In vitro synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. *Proc Natl Acad Sci U S A 101*, 1519–24.

[25] Ishihara, G., Goto, M., Saeki, M., Ito, K., Hori, T., Kigawa, T., Shirouzu, M., and Yokoyama, S. (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr Purif 41*, 27–37.

[26] Deniaud, A., Liguori, L., Blesneac, I., Lenormand, J. L., and Pebay-Peyroula, E. (2010) Crystallization of the membrane protein hVDAC1 produced in cell-free system. *Biochim Biophys Acta 1798*, 1540–6.

[27] Zaid, H., Abu-Hamad, S., Israelson, A., Nathan, I., and Shoshan-Barmatz, V. (2005) The voltage-dependent anion channel-1 modulates apoptotic cell death. *Cell Death Differ 12*, 751–760.

[28] Chung, C. T., Niemela, S. L., and Miller, R. H. (1989) One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A 86*, 2172–5.

[29] Hoover, D. M. and Lubkowski, J. (2002) DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res 30*, e43.

[30] Bubeck, P., Winkler, M., and Bautsch, W. (1993) Rapid cloning by homologous recombination in vivo. *Nucleic Acids Res 21*, 3601–2.

[31] Jones, D. H. and Howard, B. H. (1991) A rapid method for recombination and site-specific mutagenesis by placing homologous ends on DNA using polymerase chain reaction. *Biotechniques 10*, 62–66.

Appendix A

Selected Plasmid Sequences

All sequences are listed in the 5'–3' direction. Cloned/inserted sequences were verified by sequencing.

pT7-αHL

Plasmid was a gift from Stephen Cheley. The cloning sites for the αHL gene, NdeI and HindIII, are highlighted in red.

```
  1 TTCTTGAAGA CGAAAGGGCC TCGTGATACG CCTATTTTTA TAGGTTAATG
 51 TCATGATAAT AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT
101 GTGCGCGGAA CCCCTATTTG TTTATTTTTC TAAATACATT CAAATATGTA
151 TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA
201 GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTT
251 GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT
301 AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG
351 ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT
401 CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG
451 TGTTGACGCC GGGCAAGAGC AACTCGGTCG CCGCATACAC TATTCTCAGA
501 ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC
551 ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC
601 TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG
```

```
 651 CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA
 701 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
 751 TGCAGCAATG GCAACAACGT TGCGCAAACT ATTAACTGGC GAACTACTTA
 801 CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT
 851 GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
 901 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG
 951 GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT
1001 CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC
1051 ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT
1101 AGATTGATTT AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC
1151 CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA
1201 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT
1251 TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA
1301 GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT
1351 AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCCTT CTAGTGTAGC
1401 CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC
1451 GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
1501 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT
1551 CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC
1601 TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT
1651 TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA
1701 CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT
1751 AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
1801 CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTTT
1851 TACGGTTCCT GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCCTGCG
1901 TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA
1951 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG
2001 AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT
2051 ATTTCACACC GCATATATGG TGCACTCTCA GTACAATCTG CTCTGATGCC
2101 GCATAGTTAA GCCAGTATAC ACTCCGCTAT CGCTACGTGA CTGGGTCATG
2151 GCTGCGCCCC GACACCCGCC AACACCCGCT GACGCGCCCT GACGGGCTTG
2201 TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT
2251 GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGGCAGCGC
2301 TCTCCCTTAT GCGACTCCTG CATTAGGAAG CAGCCCAGTA GTAGGTTGAG
2351 GCCGTTGAGC ACCGCCGCCG CAAGGAATGG TGCATGCAAG GAGATGGCGC
2401 CCAACAGTCC CCCGGCCACG GGGCCTGCCA CCATACCCAC GCCGAAACAA
2451 GCGCTCATGA GCCCGAAGTG GCGAGCCCGA TCTTCCCCAT CGGTGATGTC
2501 GGCGATATAG GCGCCAGCAA CCGCACCTGT GGCGCCGGTG ATGCCGGCCA
2551 CGATGCGTCC GGCGTAGAGG ATCGAGATCT AGCCCGCCTA ATGAGCGGGC
2601 TTTTTTTTAG ATCTCGATCC CGCGAAATTA ATACGACTCA CTATAGGGAG
2651 ACCACAACGG TTTCCCTCTA GAAATAATTT TGTTTAACTT TAAGAAGGAG
```

```
2701 ATATACATAT GGCAGATTCT GATATTAATA TTAAAACCGG TACTACAGAT
2751 ATTGGAAGCA ATACTACAGT AAAAACAGGT GATTTAGTCA CTTATGATAA
2801 AGAAAATGGC ATGCACAAAA AAGTATTTTA TAGTTTTATC GATGATAAAA
2851 ATCACAATAA AAAACTGCTA GTTATTAGAA CAAAAGGTAC CATTGCTGGT
2901 CAATATAGAG TTTATAGCGA AGAAGGTGCT AACAAAAGTG GTTTAGCCTG
2951 GCCTTCAGCC TTTAAGGTAC AGTTGCAACT ACCTGATAAT GAAGTAGCTC
3001 AAATATCTGA TTACTATCCA AGAAATTCGA TTGATACAAA AGAGTATATG
3051 AGTACTTTAA CTTATGGATT CAACGGTAAT GTTACTGGTG ATGATACAGG
3101 AAAAATTGGC GGCCTTATTG GTGCAAATGT TTCGATTGGT CATACACTGA
3151 AATATGTTCA ACCTGATTTC AAAACAATTT TAGAGAGCCC AACTGATAAA
3201 AAAGTAGGCT GGAAAGTGAT ATTTAACAAT ATGGTGAATC AAAATTGGGG
3251 ACCATACGAT CGAGATTCTT GGAACCCGGT ATATGGCAAT CAACTTTTCA
3301 TGAAAACTAG AAATGGTTCT ATGAAAGCAG CAGATAACTT CCTTGATCCT
3351 AACAAAGCAA GTTCTCTATT ATCTTCAGGG TTTTCACCAG ACTTCGCTAC
3401 AGTTATTACT ATGGATAGAA AAGCATCCAA ACAACAAACA AATATAGATG
3451 TAATATACGA ACGAGTTCGT GATGATTACC AATTGCATTG GACTTCAACA
3501 AATTGGAAAG GTACCAATAC TAAAGATAAA TGGACAGATC GTTCTTCAGA
3551 AAGATATAAA ATCGATTGGG AAAAAGAAGA AATGACAAAT TAATGTAAAT
3601 TATTTGTACA TGTACAAATA AATATAATTT ATAACTTTAG CCGAAAGCTT
3651 GGATCCGGCT GCTAACAAAG CCCGAAAGGA AGCTGAGTTG GCTGCTGCCA
3701 CCGCTGAGCA ATAACTAGCA TAACCCCTTG GGGCCTCTAA ACGGGTCTTG
3751 AGGGGTTTTT TGCTGAAAGG AGGAACTATA TCCGGATAAT TCGAGCTCGG
3801 TACCCACCCC GGTTGATAAT CAGAAAAGCC CCAAAAACAG GAAGATTGTA
3851 TAAGCAAATA TTTAAATTGT AAACGTTAAT ATTTTGTTAA AATTCGCGTT
3901 AAATTTTTGT TAAATCAGCT CATTTTTTAA CCAATAGGCC GAAATCGGCA
3951 AAATCCCTTA TAAATCAAAA GAATAGACCG AGATAGGGTT GAGTGTTGTT
4001 CCAGTTTGGA ACAAGAGTCC AGTATTAAAG AACGTGGACT CCAACGTCAA
4051 AGGGCGAAAA ACCGTCTATC AGGGCGATGG CCCACTACGT GAACCATCAC
4101 CCTAATCAAG TTTTTTGGGG TCGAGGTGCC GTAAAGCACT AAATCGGAAC
4151 CCTAAAGGGA TGCCCCGATT TAGAGCTTGA CGGGGAAAGC CGGCGAACGT
4201 GGCGAGAAAG GAAGGGAAGA AAGCGAAAGG AGCGGGCGCT AGGGCGCTGG
4251 CAAGTGTAGC GGTCACGCTG CGCGTAACCA CCACACCCGC CGCGCTTAAT
4301 GCGCCGCTAC AGGGCGCGTG GGGATCCTCT AGAGTCGACC TGCAGGCATG
4351 CAAGCTATCC CGCAAGAGGC CCGGCAGTAC CGGCATAACC AAGCCTATGC
4401 CTACAGCATC CAGGGTGACG GTGCCGAGGA TGACGATGAG CGCATTGTTA
4451 GATTTCATAC ACGGTGCCTG ACTGCGTTAG CAATTTAACT GTGATAAACT
4501 ACCGCATTAA AGCTAGCTTA TCGATGATAA GCTGTCAAA
```

pT7-αHL-PLM-D8

Plasmid designed and constructed by Stephen Cheley. The cloning sites for the RL4-D8 gene, NdeI and HindIII, are highlighted in red. The sites for PLM cassette insertion, AgeI and StuI, are highlighted in green.

```
   1 TTGTTGAAGA CGAAAGGGCC TCGTGATACG CCTATTTTTA TAGGTTAATG
  51 TCATGATAAT AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT
 101 GTGCGCGGAA CCCCTATTTG TTTATTTTTC TAAATACATT CAAATATGTA
 151 TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA
 201 GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTT
 251 GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT
 301 AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG
 351 ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT
 401 CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG
 451 TGTTGACGCC GGGCAAGAGC AACTCGGTCG CCGCATACAC TATTCTCAGA
 501 ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC
 551 ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC
 601 TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG
 651 CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA
 701 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
 751 TGCAGCAATG GCAACAACGT TGCGCAAACT ATTAACTGGC GAACTACTTA
 801 CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT
 851 GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
 901 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG
 951 GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT
1001 CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC
1051 ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT
1101 AGATTGATTT AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC
1151 CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA
1201 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT
1251 TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA
1301 GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT
1351 AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCCTT CTAGTGTAGC
1401 CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC
1451 GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
1501 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT
1551 CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC
1601 TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT
```

```
1651 TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA
1701 CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT
1751 AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
1801 CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTTT
1851 TACGGTTCCT GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCCTGCG
1901 TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA
1951 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG
2001 AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT
2051 ATTTCACACC GCATATATGG TGCACTCTCA GTACAATCTG CTCTGATGCC
2101 GCATAGTTAA GCCAGTATAC ACTCCGCTAT CGCTACGTGA CTGGGTCATG
2151 GCTGCGCCCC GACACCCGCC AACACCCGCT GACGCGCCCT GACGGGCTTG
2201 TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT
2251 GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGGCAGCGC
2301 TCTCCCTTAT GCGACTCCTG CATTAGGAAG CAGCCCAGTA GTAGGTTGAG
2351 GCCGTTGAGC ACCGCCGCCG CAAGGAATGG TGCATGCAAG GAGATGGCGC
2401 CCAACAGTCC CCCGGCCACG GGGCCTGCCA CCATACCCAC GCCGAAACAA
2451 GCGCTCATGA GCCCGAAGTG GCGAGCCCGA TCTTCCCCAT CGGTGATGTC
2501 GGCGATATAG GCGCCAGCAA CCGCACCTGT GGCGCCCGTG ATGCCGGCCA
2551 CGATGCGTCC GGCGTAGAGG ATCGAGATCT AGCCCGCCTA ATGAGCGGGC
2601 TTTTTTTTAG ATCTCGATCC CGCGAAATTA ATACGACTCA CTATAGGGAG
2651 ACCACAACGG TTTCCCTCTA GAAATAATTT TGTTTAACTT TAAGAAGGAG
2701 ATATACATAT CGGCAGATTCT GATATTAATA TTAAAACCGG CACTACAGAT
2751 ATTGGAAGCA ATACTACAGT AAAAACAGGT GATTTAGTCA CTTATGATAA
2801 AGAAAATGGC ATGCACAAAA AAGTATTTTA TAGTTTTATC GATGATAAAA
2851 ATCACAATAA AAAACTGCTA GTTATTAGAA CAAAAGGTAC CATTGCTGGT
2901 CAATATAGAG TTTATAGCGA AGAAGGTGCT AACAAAAGTG GTTTAGCCTG
2951 GCCTTCAGCC TTTAAGGTAC AGTTGCAACT ACCTGATAAT GAAGTAGCTC
3001 AAATATCTGA TTACTATCCG CGGAATTCGA TTGATACAAA AGAGTATATG
3051 TCGACGTTAA CGTACGGATT CAACGGTAAT GTCACCGGTG ATGATAGCAG
3101 CGGAAGCGGC AGTAGCGGCG CACGTAAACG TCGTCGTCAT CCGAGCGGCC
3151 CACCGACCGC GGGTAGCTCC GGGAGCGGCT CTAGCAAAAT TGGAGGCCTC
3201 ATTGGTGCAA ACGTCTCGAT TGGCCATACA CTTAAGTATG TACAACCTGA
3251 TTTCAAAACA ATTCTCGAGC CCAACTGATA AAAAAGTAGG CTGGAAAGTG
3301 ATATTTAACA ATATGGTGAA TCAAAATTGG GGACCATACG ATCGAGATTC
3351 TTGGAACCCG GTATATGGCA ATCAACTTTT CATGAAAACT AGAAATGGTT
3401 CTATGAAAGC AGCAGATAAC TTCCTTGATC CTAACAAAGC AAGTTCTCTA
3451 TTATCTTCAG GGTTTTCACC AGACTTCGCT ACAGTTATTA CTATGGATAG
3501 AAAAGCATCC AAACAACAAA CAAATATAGA TGTAATATAC GAACGAGTTC
3551 GTGATGATTA CCAATTGCAT TGGACTTCAA CAAATTGGAA AGGTACCAAT
3601 ACTAAAGATA AATGGACAGA TCGTTCTTCA GAAAGATATA AAATCGATTG
3651 GGAAAAAGAA GAAATGACAA ATGATGACGA TGATGACGAC GATGATTGAT
```

```
3701 AAGCTTGGAT CCGGCTGCTA ACAAAGCCCG AAAGGAAGCT GAGTTGGCTG
3751 CTGCCACCGC TGAGCAATAA CTAGCATAAC CCCTTGGGGC CTCTAAACGG
3801 GTCTTGAGGG GTTTTTTGCT GAAAGGAGGA ACTATATCCG GATAATTCGA
3851 GCTCGGTACC CACCCCGGTT GATAATCAGA AAAGCCCCAA AAACAGGAAG
3901 ATTGTATAAG CAAATATTTA AATTGTAAAC GTTAATATTT TGTTAAAATT
3951 CGCGTTAAAT TTTTGTTAAA TCAGCTCATT TTTTAACCAA TAGGCCGAAA
4001 TCGGCAAAAT CCCTTATAAA TCAAAAGAAT AGACCGAGAT AGGGTTGAGT
4051 GTTGTTCCAG TTTGGAACAA GAGTCCAGTA TTAAAGAACG TGGACTCCAA
4101 CGTCAAAGGG CGAAAAACCG TCTATCAGGG CGATGGCCCA CTACGTGAAC
4151 CATCACCCTA ATCAAGTTTT TTGGGGTCGA GGTGCCGTAA AGCACTAAAT
4201 CGGAACCCTA AAGGGATGCC CCGATTTAGA GCTTGACGGG GAAAGCCGGC
4251 GAACGTGGCG AGAAAGGAAG GGAAGAAAGC GAAAGGAGCG GGCGCTAGGG
4301 CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCACCAC ACCCGCCGCG
4351 CTTAATGCGC CGCTACAGGG CGCGTGGGGA TCCTCTAGAG TGGACCTGCA
4401 GGCATGCAAG CTATCCCGCA AGAGGCCCGG CAGTACCGGC ATAACCAAGC
4451 CTATGCCTAC AGCATCCAGG GTGACGGTGC CGAGGATGAC GATGAGCGCA
4501 TTGTTAGATT TCATACACGG TGCCTGACTG CGTTAGCAAT TTAACTGTGA
4551 TAAACTACCG CATTAAAGCT AGCTTATCGA TGATAAGCTG TCAAACATGA
4601 GAA
```

pT7-SMC-OmpG

Plasmid was a gift from Sean Conlan. The cloning sites, NdeI and HindIII, are highlighted in red.

```
   1 TTCTTGAAGA CGAAAGGGCC TCGTGATACG CCTATTTTTA TAGGTTAATG
  51 TCATGATAAT AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT
 101 GTGCGCGGAA CCCCTATTTG TTTATTTTTC TAAATACATT CAAATATGTA
 151 TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA
 201 GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTT
 251 GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT
 301 AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG
 351 ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT
 401 CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG
 451 TGTTGACGCC GGGCAAGAGC AACTCGGTCG CCGCATACAC TATTCTCAGA
 501 ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC
 551 ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC
 601 TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG
 651 CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA
 701 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
 751 TGCAGCAATG GCAACAACGT TGCGCAAACT ATTAACTGGC GAACTACTTA
 801 CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT
 851 GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
 901 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG
 951 GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT
1001 CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC
1051 ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT
1101 AGATTGATTT AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC
1151 CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA
1201 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT
1251 TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA
1301 GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT
1351 AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCCTT CTAGTGTAGC
1401 CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC
1451 GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
1501 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT
1551 CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC
1601 TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT
1651 TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA
```

```
1701 CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT
1751 AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
1801 CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTTT
1851 TACGGTTCCT GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCCTGCG
1901 TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA
1951 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG
2001 AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT
2051 ATTTCACACC GCATATATGG TGCACTCTCA GTACAATCTG CTCTGATGCC
2101 GCATAGTTAA GCCAGTATAC ACTCCGCTAT CGCTACGTGA CTGGGTCATG
2151 GCTGCGCCCC GACACCCGCC AACACCCGCT GACGCGCCCT GACGGGCTTG
2201 TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT
2251 GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGGCAGCGC
2301 TCTCCCTTAT GCGACTCCTG CATTAGGAAG CAGCCCAGTA GTAGGTTGAG
2351 GCCGTTGAGC ACCGCCGCCG CAAGGAATGG TGCATGCAAG GAGATGGCGC
2401 CCAACAGTCC CCCGGCCACG GGGCCTGCCA CCATACCCAC GCCGAAACAA
2451 GCGCTCATGA GCCCGAAGTG GCGAGCCCGA TCTTCCCCAT CGGTGATGTC
2501 GGCGATATAG GCGCCAGCAA CCGCACCTGT GGCGCCGGTG ATGCCGGCCA
2551 CGATGCGTCC GGCGTAGAGG ATCGAGATCT AGCCCGCCTA ATGAGCGGGC
2601 TTTTTTTTAG ATCTCGATCC CGCGAAATTA ATACGACTCA CTATAGGGAG
2651 ACCACAACGG TTTCCCTCTA GAAATAATTT TGTTTAACTT TAAGAAGGAG
2701 ATATACATAT CAAAAAGTTA TTACCCTGTA CCGCACTGGT GATGTGTGCG
2751 GGAATGGCCT GCGCACAGGC CGAGGAAAGG AACGACTGGC ACTTTAATAT
2801 CGGCGCGATG TACGAAATAG AAAACGTCGA GGGTTATGGC GAAGATATGG
2851 ATGGGCTGGC GGAGCCTTCA GTCTATTTTA ATGCCGCCAA CGGGCCGTGG
2901 AGAATTGCTC TGGCCTATTA TCAGGAAGGG CCGGTAGATT ATAGCGCGGG
2951 TAAACGTGGA ACGTGGTTTG ATCGCCCGGA GCTGGAGGTG CATTATCAGT
3001 TCCTCGAAAA CGATGATTTC AGTTTCGGCC TGACCGGCGG TTTCCGTAAT
3051 TATGGTTATC ACTACGTTGA TGAACCGGGT AAAGACACGG CGAATATGCA
3101 GCGCTGGAAA ATCGCGCCAG ACTGGGATGT GAAACTGACT GACGATTTAC
3151 GTTTCAACGG TTGGTTGTCG ATGTATAAAT TTGCCAACGA TCTGAACACT
3201 ACCGGTTACG CTGATACCCG TGTCGAAACG GAAACAGGTC TGCAATATAC
3251 CTTCAACGAA ACGGTTGCCT TGCGAGTGAA CTATTATCTC GAGCGCGGCT
3301 TCAATATGGA CGACAGCCGC AATAACGGTG AGTTTTCCAC GCAAGAAATT
3351 CGCGCCTATT TGCCGCTGAC GCTCGGCAAC CACTCGGTGA CGCCGTATAC
3401 GCGCATTGGG CTGGATCGCT GGAGTAACTG GGACTGGCAG GATGATATTG
3451 AACGTGAAGG CCATGATTTT AACCGTGTAG GTTTATTTTA CGGTTATGAT
3501 TTCCAGAACG GACTTTCCGT TTCGCTGGAA TACGCGTTTG AGTGGCAGGA
3551 TCACGACGAA GGCGACAGTG ATAAATTCCA TTATGCAGGT GTCGGCGTAA
3601 ATTACTCGTT CTGATAAAAG CTTGGATCCG GCTGCTAACA AAGCCCGAAA
3651 GGAAGCTGAG TTGGCTGCTG CCACCGCTGA GCAATAACTA GCATAACCCC
3701 TTGGGGCCTC TAAACGGGTC TTGAGGGGTT TTTTGCTGAA AGGAGGAACT
```

```
3751 ATATCCGGAT AATTCGAGCT CGGTACCCAC CCCGGTTGAT AATCAGAAAA
3801 GCCCCAAAAA CAGGAAGATT GTATAAGCAA ATATTTAAAT TGTAAACGTT
3851 AATATTTTGT TAAAATTCGC GTTAAATTTT TGTTAAATCA GCTCATTTTT
3901 TAACCAATAG GCCGAAATCG GCAAAATCCC TTATAAATCA AAAGAATAGA
3951 CCGAGATAGG GTTGAGTGTT GTTCCAGTTT GGAACAAGAG TCCAGTATTA
4001 AAGAACGTGG ACTCCAACGT CAAAGGGCGA AAAACCGTCT ATCAGGGCGA
4051 TGGCCCACTA CGTGAACCAT CACCCTAATC AAGTTTTTTG GGGTCGAGGT
4101 GCCGTAAAGC ACTAAATCGG AACCCTAAAG GGATGCCCCG ATTTAGAGCT
4151 TGACGGGGAA AGCCGGCGAA CGTGGCGAGA AAGGAAGGGA AGAAAGCGAA
4201 AGGAGCGGGC GCTAGGGCGC TGGCAAGTGT AGCGGTCACG CTGCGCGTAA
4251 CCACCACACC CGCCGCGCTT AATGCGCCGC TACAGGGCGC GTGGGGATCC
4301 TCTAGAGTCG ACCTGCAGGC ATGCAAGCTA TCCCGCAAGA GGCCCGGCAG
4351 TACCGGCATA ACCAAGCCTA TGCCTACAGC ATCCAGGGTG ACGGTGCCGA
4401 GGATGACGAT GAGCGCATTG TTAGATTTCA TACACGGTGC CTGACTGCGT
4451 TAGCAATTTA ACTGTGATAA ACTACCGCAT TAAAGCTAGC TTATCGATGA
4501 TAAGCTGTCA AACATGAGAA
```

pQELH1-OmpG-ΔD215-L7I(S)

pQELH1 was generated from pQE60L (a gift from Stephen Tucker) as described. The cloning sites of the OmpG-ΔD215-L7I(S) gene, NdeI and HindIII, are highlighted in red. The sites flanking the L7I(S) insertion, NcoI and SpeI, are highlighted in yellow.

```
   1 CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT
  51 AATAGATTCA ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG
 101 AGGAGAAATT ACATATGAAA AAGTTATTAC CCTGTACCGC ACTGGTGATG
 151 TGTGCGGGAA TGGCCTGCGC ACAGGCCGAG GAAAGGAACG ACTGGCACTT
 201 TAATATCGGC GCGATGTACG AAATAGAAAA CGTCGAGGGT TATGGCGAAG
 251 ATATGGATGG GCTGGCGGAG CCTTCAGTCT ATTTTAATGC CGCCAACGGG
 301 CCGTGGAGAA TTGCTCTGGC CTATTATCAG GAAGGGCCGG TAGATTATAG
 351 CGCGGGTAAA CGTGGAACGT GGTTTGATCG CCCGGAGCTG GAGGTGCATT
 401 ATCAGTTCCT CGAAAACGAT GATTTCAGTT TCGGCCTGAC CGGCGGTTTC
 451 CGTAATTATG GTTATCACTA CGTTGATGAA CCGGGTAAAG ACACGGCGAA
 501 TATGCAGCGC TGGAAAATCG CGCCAGACTG GGATGTGAAA CTGACTGACG
 551 ATTTACGTTT CAACGGTTGG TTGTCGATGT ATAAATTTGC CAACGATCTG
 601 AACACTACCG GTTACGCTGA TACCCGTGTC GAAACGGAAA CAGGTCTGCA
 651 ATATACCTTC AACGAAACGG TTGCCTTGCG AGTGAACTAT TATCTCGAGC
 701 GCGGCTTCAA TATGGACGAC AGCCGCAATA ACGGTGAGTT TTCCACGCAA
 751 GAAATTCGCG CCTATTTGCC GCTGACGCTC GGCAACCACT CGGTGACGCC
 801 GTATACGCGC ATTGGGCTGC GCTGGAGTAA CTGGGACTGG CAGGATGATA
 851 TTGAACGTGA AGGCCATGAT TTTAACCGTG TAGGTTTATT TTACGGTTAT
 901 GATTTCCAGA ACGGACTTTC CGTTTCGCTG GAATACGCGT TTGAGTGGCA
 951 GGATCACGAC GAATCCATGG CGAGCATGAC AGGCGGCCAG CAGATGGGCA
1001 CTAGTGGCGA CAGTGATAAA TTCCATTATG CAGGTGTCGG CGTAAATTAC
1051 TCGTTCTGAT AAAAGCTTAA TTAGCTGAGC TTGGACTCCT GTTGATAGAT
1101 CCAGTAATGA CCTCAGAACT CCATCTGGAT TTGTTCAGAA CGCTCGGTTG
1151 CCGCCGGGCG TTTTTTATTG GTGAGAATCC AAGCTAGCTT GGCGAGATTT
1201 TCAGGAGCTA AGGAAGCTAA AATGGAGAAA AAAATCACTG GATATACCAC
1251 CGTTGATATA TCCCAATGGC ATCGTAAAGA ACATTTTGAG GCATTTCAGT
1301 CAGTTGCTCA ATGTACCTAT AACCAGACCG TTCAGCTGGA TATTACGGCC
1351 TTTTTAAAGA CCGTAAAGAA AAATAAGCAC AAGTTTTATC CGGCCTTTAT
1401 TCACATTCTT GCCCGCCTGA TGAATGCTCA TCCGGAATTT CGTATGGCAA
1451 TGAAAGACGG TGAGCTGGTG ATATGGGATA GTGTTCACCC TTGTTACACC
1501 GTTTTCCATG AGCAAACTGA AACGTTTTCA TCGCTCTGGA GTGAATACCA
```

```
1551 CGACGATTTC CGGCAGTTTC TACACATATA TTCGCAAGAT GTGGCGTGTT
1601 ACGGTGAAAA CCTGGCCTAT TTCCCTAAAG GGTTTATTGA GAATATGTTT
1651 TTCGTCTCAG CCAATCCCTG GGTGAGTTTC ACCAGTTTTG ATTTAAACGT
1701 GGCCAATATG GACAACTTCT TCGCCCCCGT TTTCACCATG CATGGGCAAA
1751 TATTATACGC AAGGCGACAA GGTGCTGATG CCGCTGGCGA TTCAGGTTCA
1801 TCATGCCGTC TGTGATGGCT TCCATGTCGG CAGAATGCTT AATGAATTAC
1851 AACAGTACTG CGATGAGTGG CAGGGCGGGG CGTAATTTTT TTAAGGCAGT
1901 TATTGGTGCC CTTAAACGCC TGGGGTAATG ACTCTCTAGC TTGAGGCATC
1951 AAATAAAACG AAAGGCTCAG TCGAAAGACT GGGCCTTTCG TTTTATCTGT
2001 TGTTTGTCGG TGAACGCTCT CCTGAGTAGG ACAAATCCGC CGCTCTAGAT
2051 TACGTGCAGT CGATGATAAG CTGTCAAACA TGAGAATTGT GCCTAATGAG
2101 TGAGCTAACT TACATTAATT GCGTTGCGCT CACTGCCCGC TTTCCAGTCG
2151 GGAAACCTGT CGTGCCAGCT GCATTAATGA ATCGGCCAAC GCGCGGGGAG
2201 AGGCGGTTTG CGTATTGGGC GCCAGGGTGG TTTTTCTTTT CACCAGTGAG
2251 ACGGGCAACA GCTGATTGCC CTTCACCGCC TGGCCCTGAG AGAGTTGCAG
2301 CAAGCGGTCC ACGCTGGTTT GCCCCAGCAG GCGAAAATCC TGTTTGATGG
2351 TGGTTAACGG CGGGATATAA CATGAGCTGT CTTCGGTATC GTCGTATCCC
2401 ACTACCGAGA TATCCGCACC AACGCGCAGC CCGGACTCGG TAATGGCGCG
2451 CATTGCGCCC AGCGCCATCT GATCGTTGGC AACCAGCATC GCAGTGGGAA
2501 CGATGCCCTC ATTCAGCATT TGCATGGTTT GTTGAAAACC GGACATGGCA
2551 CTCCAGTCGC CTTCCCGTTC CGCTATCGGC TGAATTTGAT TGCGAGTGAG
2601 ATATTTATGC CAGCCAGCCA GACGCAGACG CGCCGAGACA GAACTTAATG
2651 GGCCCGCTAA CAGCGCGATT TGCTGGTGAC CCAATGCGAC CAGATGCTCC
2701 ACGCCCAGTC GCGTACCGTC TTCATGGGAG AAAATAATAC TGTTGATGGG
2751 TGTCTGGTCA GAGACATCAA GAAATAACGC CGGAACATTA GTGCAGGCAG
2801 CTTCCACAGC AATGGCATCC TGGTCATCCA GCGGATAGTT AATGATCAGC
2851 CCACTGACGC GTTGCGCGAG AAGATTGTGC ACCGCCGCTT TACAGGCTTC
2901 GACGCCGCTT CGTTCTACCA TCGACACCAC CACGCTGGCA CCCAGTTGAT
2951 CGGCGCGAGA TTTAATCGCC GCGACAATTT GCGACGGCGC GTGCAGGGCC
3001 AGACTGGAGG TGGCAACGCC AATCAGCAAC GACTGTTTGC CCGCCAGTTG
3051 TTGTGCCACG CGGTTGGGAA TGTAATTCAG CTCCGCCATC GCCGCTTCCA
3101 CTTTTTCCCG CGTTTTCGCA GAAACGTGGC TGGCCTGGTT CACCACGCGG
3151 GAAACGGTCT GATAAGAGAC ACCGGCATAC TCTGCGACAT CGTATAACGT
3201 TACTGGTTTC ACATTCACCA CCCTGAATTG ACTCTCTTCC GGGCGCTATC
3251 ATGCCATACC GCGAAAGGTT TTGCACCATT CGATGGTGTC GGAATTTCGG
3301 GCAGCGTTGG GTCCTGGCCA CGGGTGCGCA TGATCTAGAG CTGCCTCGCG
3351 CGTTTCGGTG ATGACGGTGA AAACCTCTGA CACATGCAGC TCCCGGAGAC
3401 GGTCACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA GCCCGTCAGG
3451 GCGCGTCAGC GGGTGTTGGC GGGTGTCGGG GCGCAGCCAT GACCCAGTCA
3501 CGTAGCGATA GCGGAGTGTA TACTGGCTTA ACTATGCGGC ATCAGAGCAG
3551 ATTGTACTGA GAGTGCACCA ATTATGCGGT GTGAAATACC GCACAGATGC
```

```
3601 GTAAGGAGAA AATACCGCAT CAGGCGCTCT TCCGCTTCCT CGCTCACTGA
3651 CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA GCTCACTCAA
3701 AGGCGGTAAT ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC
3751 ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT
3801 GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC
3851 GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG
3901 GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC
3951 GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT
4001 CTCATAGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC
4051 AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT
4101 ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC
4151 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC
4201 GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG
4251 GACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA
4301 GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT
4351 TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTCAAGA
4401 AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT
4451 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG
4501 ATCCTTTTAA ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA
4551 GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG GCACCTATCT
4601 CAGCGATCTG TCTATTTCGT TCATCCATAG TTGCCTGACT CCCCGTCGTG
4651 TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT
4701 GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
4751 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC
4801 TCCATCCAGT CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC
4851 AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC ATCGTGGTGT
4901 CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA
4951 AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT
5001 CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
5051 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA
5101 TGCTTTTCTG TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG
5151 TATGCGGCGA CCGAGTTGCT CTTGCCCGGC GTCAATACGG GATAATACCG
5201 CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG
5251 GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA
5301 ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
5351 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA
5401 AGGGCGACAC GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA
5451 TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTTGAAT
5501 GTATTTAGAA AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA
5551 GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT TAACCTATAA
5601 AAATAGGCGT ATCACGAGGC CCTTTCGTCT TCAC
```

pB24-OmpG-ΔD215-L7I(S)

The ORF is identical to that in pQELH1-OmpG-ΔD215-L7I(S). The cloning sites, NdeI and HindIII, are highlighted in red.

```
   1 ATCGATGCAT AATGTGCCTG TCAAATGGAC GAAGCAGGGA TTCTGCAAAC
  51 CCTATGCTAC TCCGTCAAGC CGTCAATTGT CTGATTCGTT ACCAATTATG
 101 ACAACTTGAC GGCTACATCA TTCACTTTTT CTTCACAACC GGCACGGAAC
 151 TCGCTCGGGC TGGCCCCGGT GCATTTTTTA AATACCCGCG AGAAATAGAG
 201 TTGATCGTCA AAACCAACAT TGCGACCGAC GGTGGCGATA GGCATCCGGG
 251 TGGTGCTCAA AAGCAGCTTC GCCTGGCTGA TACGTTGGTC CTCGCGCCAG
 301 CTTAAGACGC TAATCCCTAA CTGCTGGCGG AAAAGATGTG ACAGACGCGA
 351 CGGCGACAAG CAAACATGCT GTGCGACGCT GGCGATATCA AAATTGCTGT
 401 CTGCCAGGTG ATCGCTGATG TACTGACAAG CCTCGCGTAC CCGATTATCC
 451 ATCGGTGGAT GGAGCGACTC GTTAATCGCT TCCATGCGCC GCAGTAACAA
 501 TTGCTCAAGC AGATTTATCG CCAGCAGCTC CGAATAGCGC CCTTCCCCTT
 551 GCCCGGCGTT AATGATTTGC CCAAACAGGT CGCTGAAATG CGGCTGGTGC
 601 GCTTCATCCG GGCGAAAGAA CCCCGTATTG GCAAATATTG ACGGCCAGTT
 651 AAGCCATTCA TGCCAGTAGG CGCGCGGACG AAAGTAAACC CACTGGTGAT
 701 ACCATTCGCG AGCCTCCGGA TGACGACCGT AGTGATGAAT CTCTCCTGGC
 751 GGGAACAGCA AAATATCACC CGGTCGGCAA ACAAATTCTC GTCCCTGATT
 801 TTTCACCACC CCCTGACCGC GAATGGTGAG ATTGAGAATA TAACCTTTCA
 851 TTCCCAGCGG TCGGTCGATA AAAAAATCGA GATAACCGTT GGCCTCAATC
 901 GGCGTTAAAC CCGCCACCAG ATGGGCATTA AACGAGTATC CCGGCAGCAG
 951 GGGATCATTT TGCGCTTCAG CCATACTTTT CATACTCCCG CCATTCAGAG
1001 AAGAAACCAA TTGTCCATAT TGCATCAGAC ATTGCCGTCA CTGCGTCTTT
1051 TACTGGCTCT TCTCGCTAAC CAAACCGGTA ACCCCGCTTA TTAAAAGCAT
1101 TCTGTAACAA AGCGGGACCA AAGCCATGAC AAAAACGCGT AACAAAAGTG
1151 TCTATAATCA CGGCAGAAAA GTCCACATTG ATTATTTGCA CGGCGTCACA
1201 CTTTGCTATG CCATAGCATT TTTATCCATA AGATTAGCGG ATCCTACCTG
1251 ACGCTTTTTA TCGCAACTCT CTACTGTTTC TCCATACCCG TTTTTTTGGG
1301 CTAGCAGGAG GAATTCCATA TGAAAAAGTT ATTACCCTGT ACCGCACTGG
1351 TGATGTGTGC GGGAATGGCC TGCGCACAGG CCGAGGAAAG GAACGACTGG
1401 CACTTTAATA TCGGCGCGAT GTACGAAATA GAAAACGTCG AGGGTTATGG
1451 CGAAGATATG GATGGGCTGG CGGAGCCTTC AGTCTATTTT AATGCCGCCA
1501 ACGGGCCGTG GAGAATTGCT CTGGCCTATT ATCAGGAAGG GCCGGTAGAT
1551 TATAGCGCGG GTAAACGTGG AACGTGGTTT GATCGCCCGG AGCTGGAGGT
1601 GCATTATCAG TTCCTCGAAA ACGATGATTT CAGTTTCGGC CTGACCGGCG
1651 GTTTCCGTAA TTATGGTTAT CACTACGTTG ATGAACCGGG TAAAGACACG
```

```
1701 GCGAATATGC AGCGCTGGAA AATCGCGCCA GACTGGGATG TGAAACTGAC
1751 TGACGATTTA CGTTTCAACG GTTGGTTGTC GATGTATAAA TTTGCCAACG
1801 ATCTGAACAC TACCGGTTAC GCTGATACCC GTGTCGAAAC GGAAACAGGT
1851 CTGCAATATA CCTTCAACGA AACGGTTGCC TTGCGAGTGA ACTATTATCT
1901 CGAGCGCGGC TTCAATATGG ACGACAGCCG CAATAACGGT GAGTTTTCCA
1951 CGCAAGAAAT TCGCGCCTAT TTGCCGCTGA CGCTCGGCAA CCACTCGGTG
2001 ACGCCGTATA CGCGCATTGG GCTGCGCTGG AGTAACTGGG ACTGGCAGGA
2051 TGATATTGAA CGTGAAGGCC ATGATTTTAA CCGTGTAGGT TTATTTTACG
2101 GTTATGATTT CCAGAACGGA CTTTCCGTTT CGCTGGAATA CGCGTTTGAG
2151 TGGCAGGATC ACGACGAATC CATGGCGAGC ATGACAGGCC TCCAGCAGAT
2201 GGGCACTAGT GGCGACAGTG ATAAATTCCA TTATGCAGGT GTCGGCGTAA
2251 ATTACTCGTT CTGATAAAAG CTTGGCTGTT TTGGCGGATG AGAGAAGATT
2301 TTCAGCCTGA TACAGATTAA ATCAGAACGC AGAAGCGGTC TGATAAAACA
2351 GAATTTGCCT GGCGGCAGTA GCGCGGTGGT CCCACCTGAC CCCATGCCGA
2401 ACTCAGAAGT GAAACGCCGT AGCGCCGATG GTAGTGTGGG GTCTCCCCAT
2451 GCGAGAGTAG GGAACTGCCA GGCATCAAAT AAAACGAAAG GCTCAGTCGA
2501 AAGACTGGGC CTTTCGTTTT ATCTGTTGTT TGTCGGTGAA CGCTCTCCTG
2551 AGTAGGACAA ATCCGCCGGG AGCGGATTTG AACGTTGCGA AGCAACGGCC
2601 CGGAGGGTGG CGGGCAGGAC GCCCGCCATA AACTGCCAGG CATCAAATTA
2651 AGCAGAAGGC CATCCTGACG GATGGCCTTT TTGCGTTTCT ACAAACTCTT
2701 TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
2751 AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT
2801 CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC
2851 TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC
2901 AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG
2951 ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT
3001 TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGTTGAC GCCGGGCAAG
3051 AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC
3101 TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT
3151 ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC
3201 TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG
3251 GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC
3301 CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA
3351 CGTTGCGCAA ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCGGCAA
3401 CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC CACTTCTGCG
3451 CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG
3501 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC
3551 TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA
3601 ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT AAGCATTGGT
3651 AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTACGCGCC
3701 CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG CGCAGCGTGA
```

```
3751 CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCGC TTTCTTCCCT
3801 TCCTTTCTCG CCACGTTCGC CGGCTTTCCC CGTCAAGCTC TAAATCGGGG
3851 GCTCCCTTTA GGGTTCCGAT TTAGTGCTTT ACGGCACCTC GACCCCAAAA
3901 AACTTGATTT GGGTGATGGT TCACGTAGTG GGCCATCGCC CTGATAGACG
3951 GTTTTTCGCC CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT
4001 GTTCCAAACT TGAACAACAC TCAACCCTAT CTCGGGCTAT TCTTTTGATT
4051 TATAAGGGAT TTTGCCGATT TCGGCCTATT GGTTAAAAAA TGAGCTGATT
4101 TAACAAAAAT TTAACGCGAA TTTTAACAAA ATATTAACGT TTACAATTTA
4151 AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT
4201 TAACGTGAGT TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA
4251 AGGATCTTCT TGAGATCCTT TTTTTCTGCG CGTAATCTGC TGCTTGCAAA
4301 CAAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA
4351 CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA
4401 TACTGTCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG
4451 TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT
4501 GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT
4551 ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACACAGC
4601 CCAGCTTGGA GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG
4651 CTATGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC
4701 GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG
4751 GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCA CCTCTGACTT
4801 GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA
4851 CGCCAGCAAC GCGGCCTTTT TACGGTTCCT GGCCTTTTGC TGGCCTTTTG
4901 CTCACATGTT CTTTCCTGCG TTATCCCCTG ATTCTGTGGA TAACCGTATT
4951 ACCGCCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG
5001 CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CGGTATTTTC
5051 TCCTTACGCA TCTGTGCGGT ATTTCACACC GCATAGGGTC ATGGCTGCGC
5101 CCCGACACCC GCCAACACCC GCTGACGCGC CCTGACGGGC TTGTCTGCTC
5151 CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA GCTGCATGTG
5201 TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGAGGCAG CAAGGAGATG
5251 GCGCCCAACA GTCCCCCGGC CACGGGGCCT GCCACCATAC CCACGCCGAA
5301 ACAAGCGCTC ATGAGCCCGA AGTGGCGAGC CCGATCTTCC CCATCGGTGA
5351 TGTCGGCGAT ATAGGCGCCA GCAACCGCAC CTGTGGCGCC GGTGATGCCG
5401 GCCACGATGC GTCCGGCGTA GAGGATCTGC TCATGTTTGA CAGCTTATC
```

pT7-H6-hVDAC1

The sites flanking the ORF, NdeI and HindIII, are highlighted in red. The NcoI site at the 3'-end of the region encoding an N-terminal hexahistidine tag is highlighted in yellow.

```
   1 TTCTTGAAGA CGAAAGGGCC TCGTGATACG CCTATTTTTA TAGGTTAATG
  51 TCATGATAAT AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT
 101 GTGCGCGGAA CCCCTATTTG TTTATTTTTC TAAATACATT CAAATATGTA
 151 TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA
 201 GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTT
 251 GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT
 301 AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG
 351 ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT
 401 CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG
 451 TGTTGACGCC GGGCAAGAGC AACTCGGTCG CCGCATACAC TATTCTCAGA
 501 ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC
 551 ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC
 601 TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG
 651 CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA
 701 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
 751 TGCAGCAATG GCAACAACGT TGCGCAAACT ATTAACTGGC GAACTACTTA
 801 CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT
 851 GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
 901 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG
 951 GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT
1001 CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC
1051 ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT
1101 AGATTGATTT AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC
1151 CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA
1201 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT
1251 TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA
1301 GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT
1351 AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCCTT CTAGTGTAGC
1401 CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC
1451 GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
1501 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT
1551 CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC
1601 TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT
```

```
1651 TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA
1701 CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT
1751 AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
1801 CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTTT
1851 TACGGTTCCT GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCCTGCG
1901 TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA
1951 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG
2001 AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT
2051 ATTTCACACC GCATATATGG TGCACTCTCA GTACAATCTG CTCTGATGCC
2101 GCATAGTTAA GCCAGTATAC ACTCCGCTAT CGCTACGTGA CTGGGTCATG
2151 GCTGCGCCCC GACACCCGCC AACACCCGCT GACGCGCCCT GACGGGCTTG
2201 TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT
2251 GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGGCAGCGC
2301 TCTCCCTTAT GCGACTCCTG CATTAGGAAG CAGCCCAGTA GTAGGTTGAG
2351 GCCGTTGAGC ACCGCCGCCG CAAGGAATGG TGCATGCAAG GAGATGGCGC
2401 CCAACAGTCC CCCGGCCACG GGGCCTGCCA CCATACCCAC GCCGAAACAA
2451 GCGCTCATGA GCCCGAAGTG GCGAGCCCGA TCTTCCCCAT CGGTGATGTC
2501 GGCGATATAG GCGCCAGCAA CCGCACCTGT GGCGCCGGTG ATGCCGGCCA
2551 CGATGCGTCC GGCGTAGAGG ATCGAGATCT AGCCCGCCTA ATGAGCGGGC
2601 TTTTTTTTAG ATCTCGATCC CGCGAAATTA ATACGACTCA CTATAGGGAG
2651 ACCACAACGG TTTCCCTCTA GAAATAATTT TGTTTAACTT TAAGAAGGAG
2701 ATATACATAT GCACCATCAC CACCATCATT CCATGGCGGT TCCACCAACG
2751 TATGCCGACT TAGGCAAAAG CGCACGCGAT GTTTTCACGA AAGGCTACGG
2801 CTTCGGCTTA ATTAAGCTGG ACCTGAAGAC GAAGAGCGAG AACGGTCTGG
2851 AATTCACGAG CAGCGGCAGC GCGAATACCG AGACGACCAA AGTGACGGGC
2901 AGCCTGGAAA CCAAGTACCG TTGGACGGAG TACGGCCTGA CCTTCACGGA
2951 GAAGTGGAAC ACCGATAACA CGCTGGGCAC CGAGATTACG GTGGAGGACC
3001 AGTTAGCACG CGGCCTGAAG TTAACGTTCG ACTCTAGCTT CAGCCCGAAC
3051 ACGGGCAAGA AGAACGCCAA GATCAAGACG GGCTACAAGC GCGAGCATAT
3101 CAACCTGGGC TGCGACATGG ACTTCGACAT CGCAGGTCCG AGCATCCGCG
3151 GTGCACTGGT GCTGGGCTAT GAGGGCTGGC TGGCCGGTTA CCAGATGAAC
3201 TTCGAAACGG CCAAGTCTCG TGTCACGCAG AGCAACTTCG CCGTGGGTTA
3251 CAAGACCGAC GAGTTCCAAC TGCATACGAA CGTGAATGAC GGCACGGAGT
3301 TCGGCGGCAG CATCTACCAG AAGGTGAACA AGAAGCTGGA GACGGCGGTG
3351 AACCTGGCAT GGACCGCCGG CAACAGCAAT ACGCGCTTTG GCATCGCCGC
3401 GAAATACCAA ATCGACCCAG ACGCATGCTT TAGCGCCAAG GTCAACAACT
3451 CTAGCCTGAT TGGCCTGGGT TATACCCAAA CGCTGAAACC GGGCATTAAG
3501 CTGACCCTGA GCGCCTTACT GGACGGCAAA AACGTCAATG CGGGCGGTCA
3551 CAAATTAGGT CTGGGCCTGG AGTTTCAAGC GTGATAAGCT TGGATCCGGC
3601 TGCTAACAAA GCCCGAAAGG AAGCTGAGTT GGCTGCTGCC ACCGCTGAGC
3651 AATAACTAGC ATAACCCCTT GGGGCCTCTA AACGGGTCTT GAGGGGTTTT
```

```
3701 TTGCTGAAAG GAGGAACTAT ATCCGGATAA TTCGAGCTCG GTACCCACCC
3751 CGGTTGATAA TCAGAAAAGC CCCAAAAACA GGAAGATTGT ATAAGCAAAT
3801 ATTTAAATTG TAAACGTTAA TATTTTGTTA AAATTCGCGT TAAATTTTTG
3851 TTAAATCAGC TCATTTTTTA ACCAATAGGC CGAAATCGGC AAAATCCCTT
3901 ATAAATCAAA AGAATAGACC GAGATAGGGT TGAGTGTTGT TCCAGTTTGG
3951 AACAAGAGTC CAGTATTAAA GAACGTGGAC TCCAACGTCA AAGGGCGAAA
4001 AACCGTCTAT CAGGGCGATG GCCCACTACG TGAACCATCA CCCTAATCAA
4051 GTTTTTTGGG GTCGAGGTGC CGTAAAGCAC TAAATCGGAA CCCTAAAGGG
4101 ATGCCCCGAT TTAGAGCTTG ACGGGGAAAG CCGGCGAACG TGGCGAGAAA
4151 GGAAGGGAAG AAAGCGAAAG GAGCGGGCGC TAGGGCGCTG GCAAGTGTAG
4201 CGGTCACGCT GCGCGTAACC ACCACACCCG CCGCGCTTAA TGCGCCGCTA
4251 CAGGGCGCGT GGGGATCCTC TAGAGTCGAC CTGCAGGCAT GCAAGCTATC
4301 CCGCAAGAGG CCCGGCAGTA CCGGCATAAC CAAGCCTATG CCTACAGCAT
4351 CCAGGGTGAC GGTGCCGAGG ATGACGATGA GCGCATTGTT AGATTTCATA
4401 CACGGTGCCT GACTGCGTTA GCAATTTAAC TGTGATAAAC TACCGCATTA
4451 AAGCTAGCTT ATCGATGATA AGCTGTCAAA CATGAGAA
```