A selenocysteine containing αHL for Single Molecule Studies

Sarah Elizabeth Rogers

Chemistry Department and Balliol College, Oxford



A thesis submitted for the degree of Doctor of Philosophy
in the University of Oxford.

Michaelmas Term 2011

For my family.

Declaration

I declare that no part of this thesis has been accepted, or is currently being submitted, for any degree or 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. The following is not my own work:

Chapter 3: The single channel recording data for the reaction between TCEP and selenocystine was gathered by Lai-Sheung Choi.

Sarah E. Rogers

(July 29, 2012)

Acknowledgements

I would like to thank my supervisor, Prof. Hagan Bayley for giving me the opportunity to work on such a fascinating project. Many thanks to Dr. Arijit Bannerjee and Dr. Marc Boudreau for their support and advise at the beginning of my DPhil. My thanks go to Dr. Steve Cheley for sharing his molecular biology expertise and to Lai Sheung Choi for her advice and support regarding synthetic procedures and electrical recording analysis. I would like to thank Dr. Gökçe Su Pulcu for generously giving her time to guide, support and advise me during the final year of my DPhil and to Dr. Arnold Boersma and Dr. Holger Gerlach for spending the time to each read a chapter of this thesis. My thanks also go to my examiners for finding the time to read this thesis.

I would like to express my gratitude to my parents Susan and David Rogers for their continuing encouragement and support, my sister Amy Rogers for always being there for me and my sister-in-law, Harriet Allison and parents-in-law Sheena and Robert Allison for welcoming me into their family. Lastly, I thank my husband James Allison for his constant support and companionship, especially during the writing of this thesis.

Abstract

Proteins containing selenocysteine (selenoproteins) have been found to exist in organisms from all domains of life. Selenoproteins are important for many *in vivo* processes such as the removal of reactive oxygen containing species (ROS), redox disulfide shuffling reactions, and pro-hormone activation. Structurally and functionally analogous to cysteine, selenocysteine's lower pKa appears to be the defining chemical difference between these two amino acids. Using a single-molecule electrical recording technique, rate constants for the reaction of selenocysteine with small molecule disulfides were obtained over a pH range of 6 - 10. Analogous single molecule experiments carried out using cysteine, revealed that, after correcting for the ratio of selenolate to selenol and thiolate to thiol based on the pKa of each amino acid, the nucleophilicity of selenocysteine was approximately 40 times that of cysteine. The selenium atom of the selenylsulfide bond was found to be substantially more electrophilic than a sulfur atom of the analogous disulfide bond and the leaving group ability of the selenolate of selenocysteine compared to the thiolate of cysteine were found to be comparable. Another biologically relavant interaction that occurs *in vivo* is the reaction between selenocysteine and organoarsenic (III) molecules. It is known that arsenic (III) compounds are toxic to organisms, and that this toxicity stems from the ability to coordinate to the thiol and selenol groups of the cysteine and selenocysteine residues within proteins. The reaction of selenocysteine with an organoarsenic species, was investigated at the single molecule level over the pH range 6.5 - 8.5. By carrying out an analogous reaction between cysteine and the organoarsenic (III) species, it was found that selenocysteine and cysteine exhibit similar reaction rates. The

organoarsenic reagent could exist in a range of different protonation states in solution and it was concluded that the rate of reaction was governed by the equilibrium of the arsenic molecule, where only some of the forms were reactive towards the selenocysteine and cysteine groups.

Contents

Declaration iii

Acknowledgements iv

Abstract v

Abbreviations xi

1 Introduction 1
- 1.1 The importance of selenocysteine-containing proteins *in vivo* . . . 1
  - 1.1.1 Selenylsulfide exchange reactions *in vivo* . . . 2
- 1.2 The bacterial selenocysteine incorporation pathway . . . 4
  - 1.2.1 Translation of the canonical 20 amino acids . . . 4
  - 1.2.2 The components required for selenocysteine translation into proteins . . . 5
- 1.3 Single molecule experiments . . . 10
- 1.4 A comparison between single molecule and ensemble experiments . . . 11
  - 1.4.1 Conformational fluctuations during the turnover of a single cholesterol oxidase enzyme . . . 11
  - 1.4.2 The effect of Au nano-particle topology on the catalytic conversion of resazurin to resorufin . . . 12
  - 1.4.3 The observation of a primary hydrogen-deuterium isotope effect . . . 13
  - 1.4.4 Rate fluctuations during enzymatic turnover of λ-exonuclease . . . 13
- 1.5 Single molecule techniques . . . 14
  - 1.5.1 Fluorescence spectroscopy . . . 14
  - 1.5.2 Force microscopy . . . 15
  - 1.5.3 Single channel electrical recordings . . . 16
- 1.6 αHL as a nanoreactor for single molecule studies . . . 17
  - 1.6.1 Properties of α-hemolysin . . . 17
  - 1.6.2 Advantages of using αHL as a nanoreactor . . . 19
  - 1.6.3 Covalent reversible and irreversible reactions carried out using αHL . . . 20
- 1.7 Thesis summary . . . 22

2 Incorporation of selenocysteine into the β barrel of αHL 24

2.1 Introduction . . . 24
2.1.1 Methods to incorporate selenocysteine into proteins . . . 24
2.1.2 *In vivo* selenoprotein expression . . . 25
2.1.3 *In vitro* selenoprotein expression . . . 27
2.1.4 Semisynthetic incorporation methods . . . 28
2.1.5 Incorporation methods summary . . . 29
2.2 Results . . . 29
2.2.1 Engineering the SECIS element and TGA codon into the αHL gene . . . 29
2.2.2 Protein Expression and Characterisation . . . 34
2.2.3 *In vitro* translation of a selenocysteine containing αHL monomer . . . 40
2.3 Discussion . . . 44
2.3.1 SECIS design . . . 44
2.3.2 Selenocysteine containing αHL expression in *Escherichia coli* . . . 46
2.3.3 Characterisation of the 117SecSECIS αHL monomer . . . 47
2.3.4 *In vitro* transcription translation of the 117SecSECIS αHL monomer . . . 48
2.3.5 Oxidation of selenocysteine during heteroheptamer preparation . . . 49
2.4 Conclusions . . . 50
2.5 Materials and Methods . . . 50
2.5.1 117CysSECISpT7 plasmid construction . . . 50
2.5.2 Recombinant Protein Expression . . . 53
2.5.3 *In vitro* transcription translation Protein Expression . . . 58
2.5.4 Miscellaneous . . . 64

3 Selenylsulfide and disulfide kinetics studied using single channel recordings 65
3.1 Introduction . . . 65
3.1.1 The importance of selenocysteine *in vivo* . . . 65
3.1.2 The physical properties of selenocysteine and cysteine . . . 68
3.1.3 Summary . . . 69
3.2 Results . . . 70
3.2.1 Characterisation of 117CysSECIS$_1$WT$_6$ and 117SecSECIS$_1$WT$_6$ αHL heptamers . . . 70
3.2.2 Reaction of the cysteine containing αHL pore with *trans*-4,5-dihydroxy-1,2-dithiane . . . 73
3.2.3 Reduction of selenylsulfide and disulfide bonds . . . 77
3.2.4 Reduction of a selenylsulfide bond using dithiothreitol . . . 84
3.3 Discussion . . . 85
3.3.1 Determining the pKa of the 117 cysteine residue situated within the β barrel of αHL . . . 85
3.3.2 Comparing the reactivity of selenium and sulfur . . . 86
3.3.3 The reduction of selenylsulfide and disulfide bonds by TCEP . . . 88
3.3.4 The phosphonium cationic intermediate . . . 91
3.3.5 Conclusion . . . 93
3.4 Materials and Methods . . . 94

3.4.1 Electrophysiology set up . . . 94
3.4.2 Chamber . . . 94
3.4.3 Electrodes . . . 95
3.4.4 Bilayer formation . . . 96
3.4.5 Bilayer capacitance . . . 96
3.4.6 Pore insertion . . . 97
3.4.7 Data acquisition . . . 97
3.4.8 Electrical recording set up . . . 97
3.4.9 Data processing . . . 99
3.4.10 Boc deprotection of Cys(Npys)-OH . . . 100
3.4.11 Solution preparation . . . 100
3.4.12 DTTox data analysis to obtain the kinetic rate constants . . . 102
3.4.13 TCEP reduction data analysis . . . 102

4 Reactions of selenocysteine and cysteine with an organoarsenic (III) compound 104
4.1 Introduction . . . 104
4.1.1 Arsenic *in vivo* . . . 104
4.1.2 The interaction of arsenic with selenium . . . 106
4.2 Results . . . 107
4.2.1 The pH dependence for the reaction between SPAO and 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ . . . 108
4.2.2 Determining the pKa values of SPAO . . . 113
4.3 Discussion . . . 115
4.3.1 The pH dependence of the forward reaction of 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ with SPAO . . . 116
4.3.2 The dependence of the reverse reaction between SPAO with 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ with pH . . . 118
4.3.3 The pH dependence of the rate of interconversion between diastereomers of SPAO . . . 120
4.3.4 Summary . . . 122
4.4 Materials and Methods . . . 122
4.4.1 Synthesis of sodium 4-sulfophenyldiiodoarsine (SPAO) . . . 122
4.4.2 Titration experiments . . . 124
4.4.3 ^{13}C NMR experiments . . . 124
4.4.4 Kinetic analysis . . . 125

5 Future Work 127
5.1 Enzyme mimics . . . 127
5.2 Selenocysteine containing enzymes . . . 128

Appendices

A Additional data 131

Bibliography **154**

Abbreviations

ACh	acetylcholine
AFM	atomic force microscopy
ATP	adenosine triphosphate
BSA	bovine serum albumin
Bis Tris Propane	bis(tris(hydroxymethyl)methylamino)
Beta-CD	beta-cyclodextrin
COx	colesterol oxidase
cGPx	cytosolic glutathione peroxidase
Cys(Npys)-OH	(3-nitro-2-pyridylthio)-L-cysteine
Cys	cysteine
Cys_2	cystine
Dbs	disulfide bond shuffling
DNA	deoxyribonucleic acid
DTTox	trans-4,5-dihydroxy-1,2-dithiane
DTNB	5,5'-dithio(2-nitrobenzoic acid)
DPhPC	1,2-diphytanoyl-sn-glycero-3-phosphocholine
DTT	dithiothreitol
dsDNA	double stranded DNA

ES-MS	electrospray mass spectroscopy
FAD	flavin adenine dinucleotide
FRET	fluorescence resonance energy transfer
FRAP	fluorescence recovery after photobleaching
GSH	glutathione
GTP	guanosine triphosphate
GSSG	oxidised glutathione
Gpx	glutathione peroxidase
GnCl	guanidinium chloride
HL	hemolysin
IASD	4-acetamido-4'-[(iodoacetyl)amino]stilbene-2,2'-disulfonate
ID	iodothyronine deiodinase
IPTG	isopropyl-D-thiogalactoside
IVTT	in vitro transcription and translation
LC-MS	liquid chromatography mass spectroscopy
MISD	4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid
MOPS	3-(N-morpholino)propanesulphonic acid
NADP	Nicotinamide adenine dinucleotide phosphate
NCL	native chemical ligation
NMR	nuclear magnetic resonance
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDI	protein disulfide isomerase
PEG	poly(ethylene glycol)
PLP	pyridoxal-5-phosphate
rRBC	rabbit red blood cell

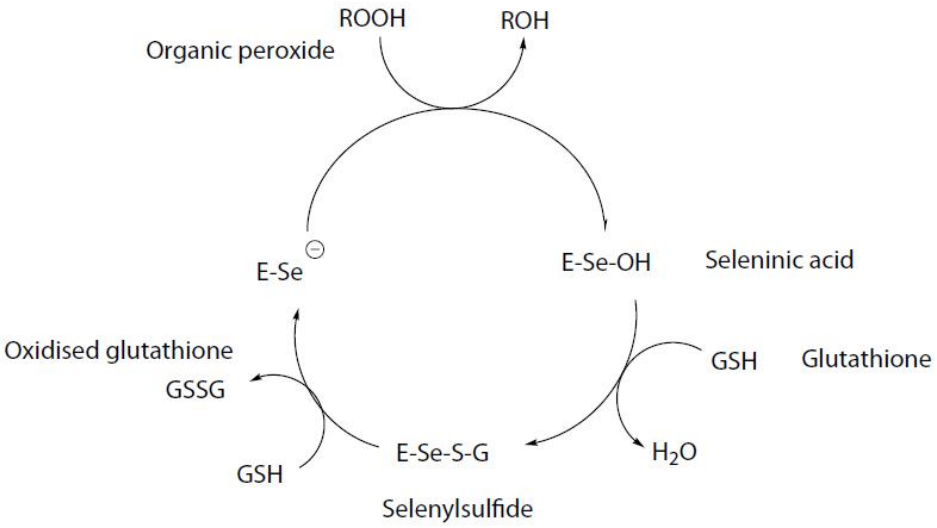
RNA	ribonucleic acid
ROS	reactive oxygen species
SDS	sodium dodecyl sulphate
Sec	selenocysteine
Sec_2	selenocystine
SECIS	selenocysteine inserting sequence
SMD	single molecule detection
SPAO	4-sulfophenylarsine oxide
TCEP	tris(2-carboxyethyl)phosphine
TIRF	Total internal reflection fluorescence
TNB	2-nitro-4-thiobenzoic acid
TOF	time of flight
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
Trx	thioredoxin
TrxR	thioredoxin reductase

Chapter 1

Introduction

1.1 The importance of selenocysteine-containing proteins *in vivo*

Selenium was discovered to be essential for the functioning of the *Escherichia coli* formate dehydrogenase enzyme (Pinsent, 1954) and mammalian (Schwarz & Flotz, 1957; Oh et al., 1974) and avian (Noguchi et al., 1973) glutathione peroxidase enzymes. Chemical characterisation and amino acid analysis of selenium containing mammalian glutathione peroxidase (Flohe et al., 1973; Forstrom et al., 1978) and glycine reductase (Cone et al., 1976) confirmed the presence of the amino acid selenocysteine (Sec) within the polypeptide of each protein. Many different selenoproteins have been so far discovered and those of known function can be grouped into those which are involved with antioxidant defence by catalytic removal of reactive oxidative species, those that catalyse the reduction of oxidised cysteine or methionine residues and the enzymes involved in deiodination during prohormone activation. The utilisation of selenoproteins for the maintenance and control of a cell's redox balance spans many different physiological pathways and so investigations into the mechanisms of these enzymes is of high importance.



Scheme 1.1: The mechanism for the reduction of organic peroxides by glutathione peroxidase (E-Se$^-$) adapted from Roy et al. (2005).

1.1.1 Selenylsulfide exchange reactions *in vivo*

The three most characterised selenocysteine containing proteins are glutathione peroxidase (Gpx), thioredoxin reductase (TrxR) and iodothyronine deiodinase (ID). The proposed mechanisms for the Gpx and TrxR enzymes both involve the formation and reduction of selenylsulfide bonds and are discussed below.

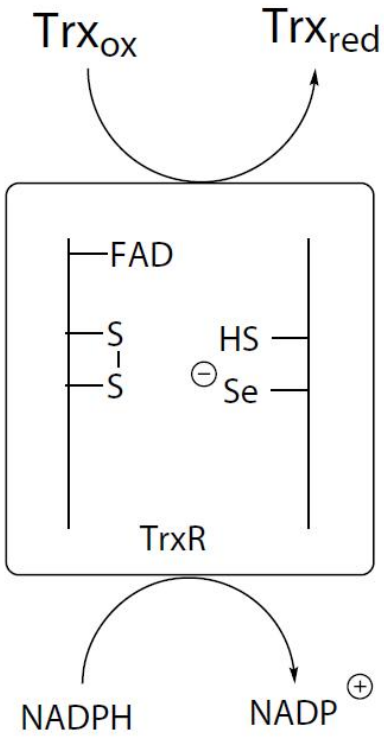
Glutathione peroxidase

Glutathione peroxidase (GPx) enzymes are responsible for the removal of reactive oxidative species. There are four distinct groups of mammalian glutathione peroxidase enzymes, all containing selenocysteine (Brigelius-Flohe, 1999). The cytosolic glutathione peroxidase enzyme (cGPx was the second selenoprotein to be discovered (Flohe et al., 1973; Ursini et al., 1995)) and is the most widely distributed selenoprotein throughout cell tissues. Bacterial homologues have been identified, but none of these are selenoproteins. The selenocysteine-containing GPx enzymes are all

homotetramers (Ladenstein & Wendel, 1983) of approximately 85 kDa (apart from the phospholipid hydroperoxide GPx), which is a monomeric protein of 19.5 kDa (Scheerer et al., 2007). The proposed catalytic mechanism by which the GPx enzymes function is shown in Scheme 1.1. The selenocysteine residue is situated in the active site stabilised in its deprotonated selenolate (Se^-) form by hydrogen bonding interactions with surrounding residues. The selenolate ion, attacks the peroxide to form seleninic acid (Se-OH) which is then attacked by glutathione (GSH) to form a selenylsulfide bond (Se-SG). A second GSH molecule regenerates the active selenolate by formation of GSSG (Mugesh & Singh, 2000). By reducing, for example, hydrogen peroxide to CO_2 and H_2O the Glutathione peroxidase enzymes protect an organism from the damaging oxidative effect of hydrogen peroxide.

Thioredoxin reductase

Thioredoxin (Trx) is an essential protein present in all organisms and maintains the redox state of other proteins. Thioredoxin reductase (TrxR) is responsible for converting Trx back to its original reduced state. Mammalian TrxR is a member of the pyridine nucleotide-disulfide oxidoreductase family and is a homodimer with one FAD bound per 55kDa subunit (Holmgren & Lu, 2010). Both selenocysteine-containing and non-selenocysteine containing forms of the mammalian TrxR enzyme exist (Zhong et al., 2000). The conserved sequence at the C-terminus of all mammalian TrxR's is Gly-Cys-Sec(or Cys)-Gly. The selenocysteine residue within this sequence is essential for function and a proposed mechanism of action involves a series of selenylsulfide and disulfide exchange reactions as illustrated in Scheme 1.2.



Scheme 1.2: The conversion of thioredoxin (Trx) from its oxidised to reduced form by thioredoxin (TrxR) adapted from Holmgren & Lu (2010)

1.2 The bacterial selenocysteine incorporation pathway

1.2.1 Translation of the canonical 20 amino acids

The 21st amino acid (Söll, 1988), selenocysteine, is cotranslationally incorporated into proteins (Zinoni et al., 1986; Leinfelder et al., 1988a) in response to the codon TGA (Zinoni et al., 1986; Chambers et al., 1986). The canonical 20 amino acids, as well as selenocysteine (and the more recently discovered 22nd amino acid, pyrrolysine), are incorporated into a growing polypeptide chain by the ribosome (Schmeing & Ramakrishnan, 2009). The ribosome is a giant structure (MW $2.6x10^6$) made up from two thirds RNA and one third proteins (Ban et al., 2000; Wimberly et al., 2000). It is divided into two domains, one known as the small subunit (S30) and the other the large subunit (S50). Translation involves the use of a tRNA species, a synthase (to attach that particular

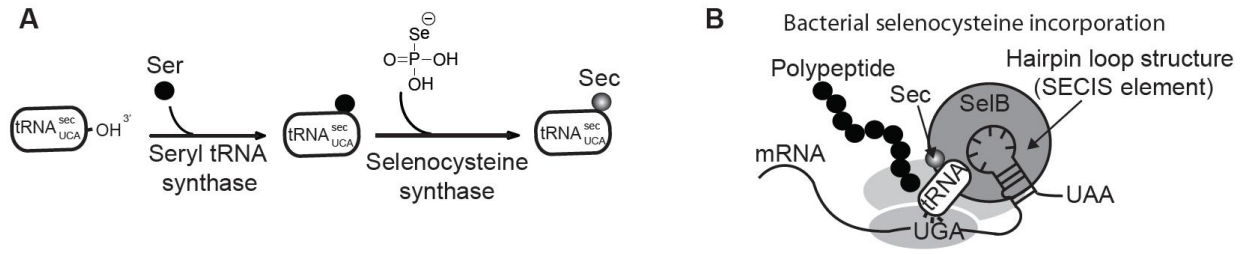


Figure 1.1: A; Aminoacylation of serine onto the tRNA^{sec} by the seryl-tRNA synthetase followed by conversion to selenocysteine carried out by selenocysteine synthase. B; Translation of selenocysteine (Sec) into the growing polypeptide chain at the ribosome.

amino acid to the 3'OH of the tRNA) and a GTP-binding chaperone protein (called EF-Tu for the canonical 20 amino acids) to guide the aminoacylated tRNA (aa-tRNA) to a specific site in the ribosome (A-site). The aa-tRNA binds to EF-Tu and enters the ribosome. Domain closure of the ribosomal subunits causes the hydrolysis of GTP resulting in the departure of EF-Tu (Selmer et al., 2006). The aa-tRNA is moved into the P-site of the ribosome where the incoming amino acid forms a peptide bond with the polypeptide chain being translated (Schmeing et al., 2005). Once the new peptide bond has been made, the tRNA exits via the ribosomal E-site. The TGA codon, encodes Sec but also acts as a translation termination sequence (stop codon) and so additional components are required to distinguish between protein termination and Sec incorporation. These additional components are summarised below.

1.2.2 The components required for selenocysteine translation into proteins

This section focuses on the bacterial selenocysteine incorporation pathway (see Figure 1.1 for an overview). For reviews detailing the translation of selenocysteine in eukaryotes see Lobanov et al. (2009); Caban & Copeland (2006); Tujebajeva et al. (2000); Low & Berry (1996). The selenocysteine tRNA (tRNA^{sec}) transcribed from the selC gene, (Leinfelder et al., 1988a,b) is aminoacylated with serine by the seryl-tRNA synthase to form Ser-tRNA^{sec}. The aminoacylated serine residue is then converted into selenocysteine on the tRNA^{sec} by the PLP-dependent selenocysteine synthase

enzyme (Leinfelder et al., 1989), using selenophosphate as a substrate. The selenophosphate is generated by the ATP-dependent selenophosphate synthase enzyme (SelD gene product). Once the Sec-tRNA^{sec} species has been created, it is recognised by the protein factor, SelB and a tertiary complex of GTP·SelB·Sec-tRNA^{sec} is formed (Leibundgut et al., 2005). This tertiary complex then binds to a hairpin loop structure (SECIS element) in the mRNA and forms a quaternary complex (GTP·SelB·Sec-tRNA^{sec}·SECIS). This quaternary complex guides the Sec-tRNA^{sec} into the correct position within the ribosome, allowing selenocysteine to be incorporated into the growing polypeptide chain (Hendrickson, 2007).

tRNA^{sec} (SelC gene product)

The nucleotide sequence of the *Escherichia coli* tRNA^{sec} (Schö et al., 1989) is shown in Figure 1.2. The tRNA^{sec} consists of 95 bases whereas other known tRNAs are approximately 75 bases long. The amino acid acceptor stem contains an extra base (eight instead of seven). A critical seryl-tRNA synthetase recognition base pairing region, is conserved, which enables seryl aminoacylation of the tRNA^{sec} by the seryl-tRNA synthetase. An additional base pair is located after the conserved synthetase recognition region ($G_1 \cdot C_{72}$ and $A_3 \cdot U_{70}$) and the anticodon sequence of tRNA^{sec} is complementary to the UGA codon (i.e. UCA). Figure 1.2 also illustrates the important recognition regions of tRNA^{sec} required for binding to the selB protein factor (see Section 1.2.2) and recognition by the selenocysteine synthase enzyme (see Section 1.2.2).

Selenophosphate synthetase (SelD gene product)

The selD gene product is a 37 kDa selenophosphatase enzyme which is responsible for the conversion of a reactive selenium containing species to selenophosphate. Selenophosphate is the substrate for conversion of serine to Sec by the selenocysteine synthase enzyme (see below). The mechanism of selenophosphate synthesis requires Mg^{2+} for the metal-ion-mediated binding of ATP. Hydrol-

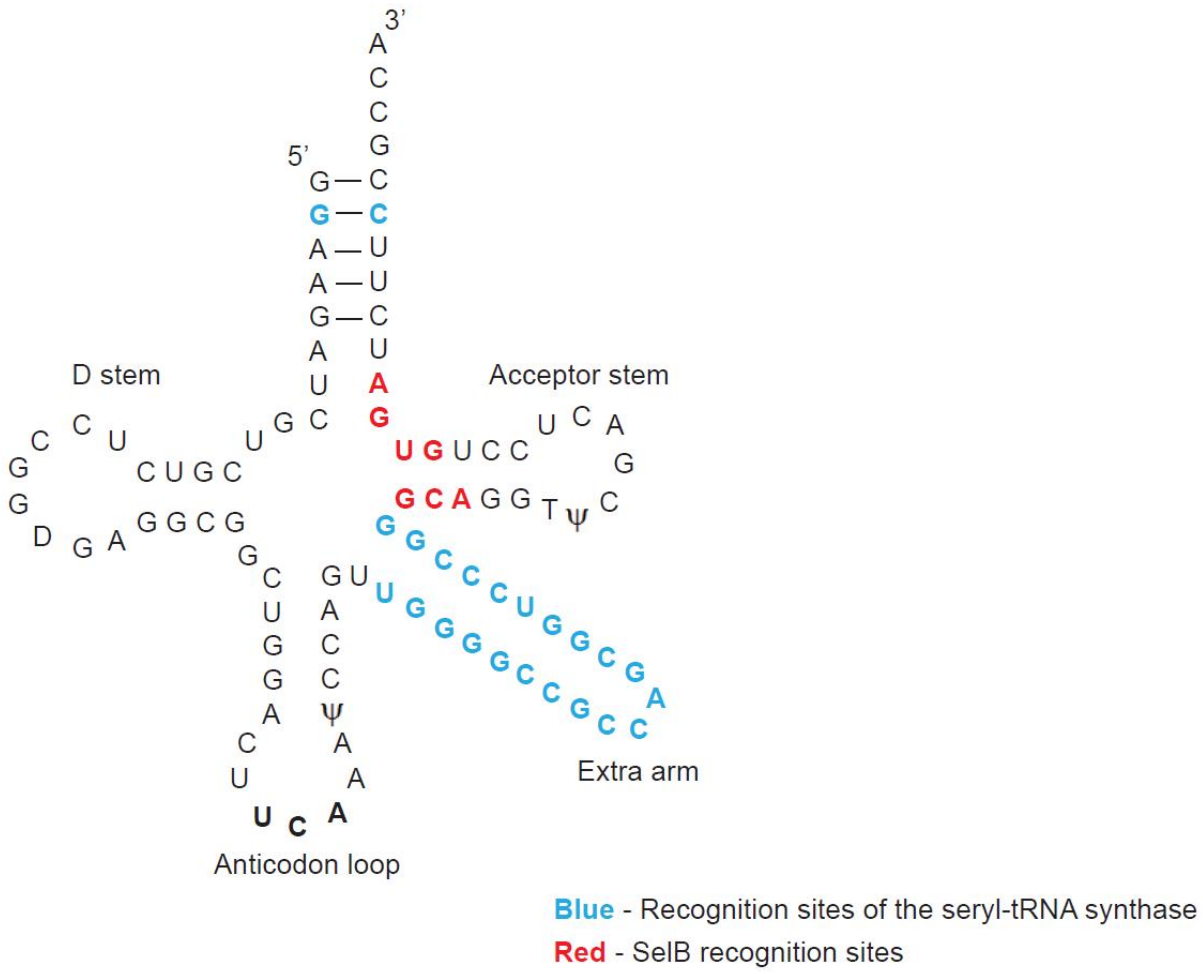
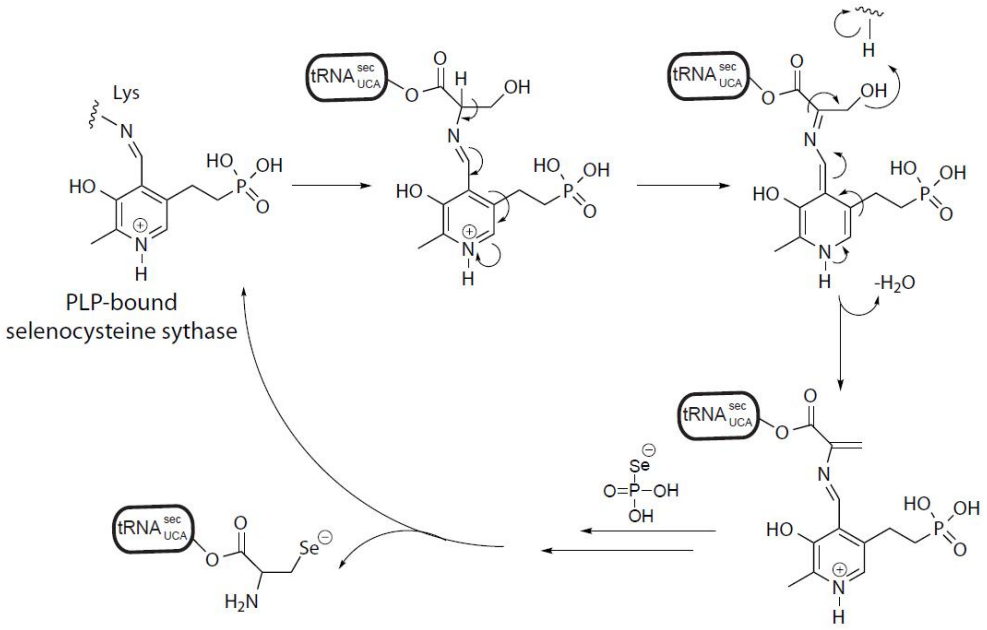


Figure 1.2: The selenocysteine tRNA ($tRNA^{sec}$) with complementary anticodon bases to the UGA codon.



Scheme 1.3: The mechanism by which selenocysteine synthase converts the aminoacylated serine residue (Ser-tRNA^{sec}) into selenocysteine. Adapted from Roy et al. (2005)

ysis of the α-β and β-γ phosphate groups of ATP, results in the addition of the γ-phosphate onto selenium as illustrated below (Ehrenreich et al., 1992):

$$\mathrm{ATP} + \mathrm{SeH} + \mathrm{Mg}^{2+} \rightarrow \mathrm{AMP} + \mathrm{P_i} + \mathrm{SeP_i} + \mathrm{H_2O} \tag{1.1}$$

Selenocysteine synthase (SelA)

The SelA gene product is a homodecamer made up of 50.7 kDa monomeric subunits arranged with pentameric symmetry in two rings (Engelhardt et al., 1992). Each subunit contains a single PLP molecule required for its function and two subunits are required to bind a single Ser-tRNA^{sec} via interaction with the acceptor arm of tRNA^{sec} (Commans & Bock, 1999). The proposed mechanism by which the 500 kDa selenocysteine synthase catalyses the conversion of serine to selenocysteine starts with the 2,3 dehydration of Ser-tRNA^{sec} by formation of the aminoacrylyl-tRNA^{sec} species, followed by nucleophilic addition of selenophosphate. The phosphate group is thought to be re-

moved by hydrolysis, resulting in the formation of Sec-tRNA^{sec}, which is then released from the enzyme (see Scheme 1.3).

Elongation factor (SelB) specific for Sec-tRNASec

The SelB gene product is a 68.8 kDa protein consisting of 614 amino acids. On binding GTP, the SelB protein undergoes a conformational change resulting in the ability to bind the Sec-tRNA^{sec} and forms the tertiary complex, SelB·GTP·Sec-tRNA^{sec}. Although the N-terminal region of the SelB protein shares sequence similarity with the elongation factor, EF-Tu (Hilgenfeld et al., 1996), SelB (unlike EF-Tu which binds all amino acids indiscriminately), will only bind tRNA^{sec}. Additionally, the SelB protein is sensitive to the aminoacylated tRNA^{sec} species and binds with higher affinity to Sec-tRNA^{sec} than Ser-tRNA^{sec} (Leibundgut et al., 2005).

The selenocysteine insertion sequence (SECIS)

The bacterial selenocysteine insertion sequence (SECIS element) situated in the mRNA coding region, 11 nucleotides downstream of the UGA (Sec) codon, is essential for selenocysteine incorporation (Zinoni et al., 1990). The SECIS element is a hairpin loop structure that allows the ribosome to discriminate between a UGA termination codon and a UGA codon encoding selenocysteine. The SECIS element interacts with the SelB·GTP·Sec-tRNA^{sec} complex to form a quaternary complex. The structure of the SECIS element from the mRNA encoding the *Escherichia coli* formate dehydrogenase H enzyme is displayed in Figure 1.3. The uracil bulge (U17) of the SECIS element, is essential for the binding of SelB to the mRNA and is thought to disrupt many important salt bridges within the SelB structure so as to initiate required conformational changes. G23 and U24 were also found to be essential for SelB recognition of the SECIS element (Leibundgut et al., 2005). By using the knowledge of which bases are essential, a minimal bacterial SECIS element was constructed Liu et al. (1998) (Liu et al., 1998; Sandman et al., 2003) and the uses of this to

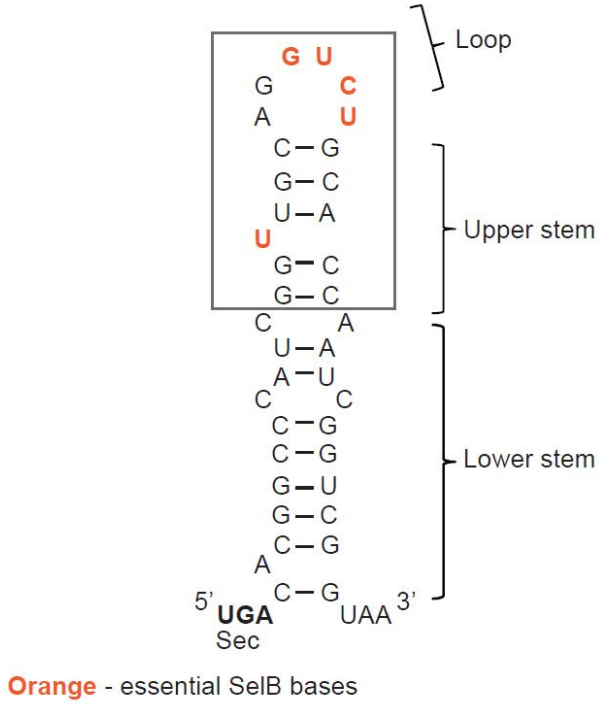


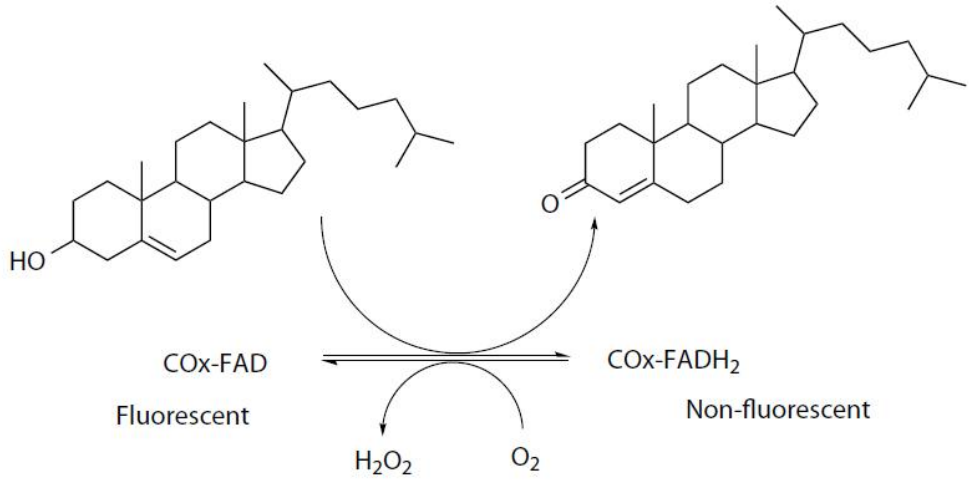
Figure 1.3: The formate dehydrogenase H selenocysteine inserting sequence (SECIS element). Adapted from Schö et al. (1989)

incorporate selenocysteine into naturally non-Sec containing proteins is discussed in Chapter 2.

The next part of this chapter introduces the concept of single molecule experiments.

1.3 Single molecule experiments

During both a single molecule and ensemble measurement, the progress of a particular reaction is monitored over time. The ensemble measurement records the proportion of reacted species to non-reacted whereas the single molecule experiment, records the time it takes for individual species to react, where the length of time to react is stochastic and discrete, i.e. it has either reacted or not (Kapanidis & Strick, 2009). The types of single molecule studies are diverse, ranging from atomic resolution of a carbon atom within a carbon nanotube (Zuo et al., 2003) to real-time imaging in



Scheme 1.4: The oxidation of cholesterol by cholesterol oxidase, adapted from Lu et al. (1998).

living bacteria (Chiu & Leake, 2011).

1.4 A comparison between single molecule and ensemble experiments

This section describes four examples to illustrate the additional information that can be obtained from a single molecule compared to an ensemble measurement.

1.4.1 Conformational fluctuations during the turnover of a single cholesterol oxidase enzyme

An advantage of studying single molecules over time compared to ensemble measurements, is that the distribution of the molecular properties can be obtained, as opposed to an average value (Deniz et al., 2008). The shape of the distribution can provide information regarding the heterogeneity (i.e. the number of pathways going from A to B) of the system and also allows data to be gathered in a time resolved manner. Cholesterol oxidase is a 52 kDa flavoenzyme that converts the alcohol

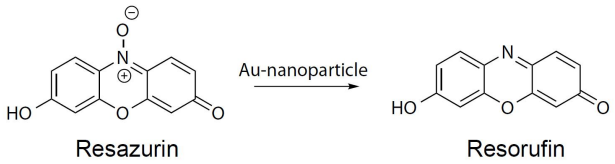
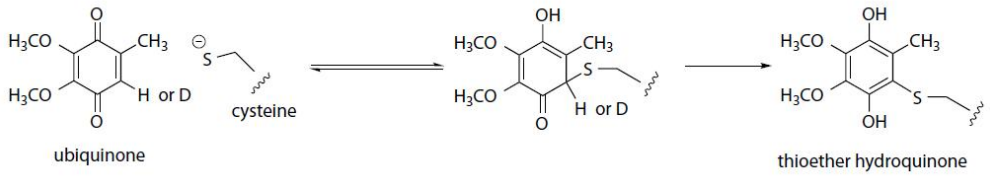


Figure 1.4: The conversion of resazurin to resorufin catalysed on the surface of Au-nanoparticles. Adapted from Xu et al. (2008)

group of cholesterol to a ketone via the utilisation of FAD (see Scheme 1.4). The flavin adenine dinucleotide (FAD) co-factor is naturally fluorescent in its oxidised form but not in its reduced form (Lu et al., 1998). By following the fluorescence blinking signals of individual COx's (i.e. the enzymatic turnover of a single oxidase), translationally immobilised in agarose gel (99% water), a frequency distribution histogram was obtained for the rate of turnover. The histograms were found to exhibit a non-exponential distribution, indicating the presence of an intermediate within the enzymatic turnover pathway. As the individual turnover events were recorded, it was possible to evaluate whether the adjacent events were independent. The adjacent turnover events were in fact not independent and exhibited a 'memory effect'. But in contrast to this, events 10 or more apart were shown to be independent of each other. It was concluded that conformational fluctuations of COx were the cause of these time-dependent, rate alterations.

1.4.2 The effect of Au nano-particle topology on the catalytic conversion of resazurin to resorufin

A more recent example where the heterogeneity of a system was observed, which would be otherwise hidden due to ensemble averaging, was the study by Xu et al. (2008) of 6 nm diameter Au-nanoparticles catalysing the conversion of resazurin to resorufin (see Figure 1.4) with NH_2OH. Total internal reflection fluorescence (TIRF) was used to monitor the fluorescence bursts that occurred on conversion of resazurin to the fluorescent resorufin product during catalysis, on the Au-nanoparticles. From single Au-nanoparticle kinetic analysis three distinct nanoparticle types



Scheme 1.5: The reaction of a cysteine residue within the β barrel of αHL, with ubiquinone. Adapted from (Lu et al., 2010)

were identified by their distinct frequency histogram distributions (i.e. the rate of conversion was different, depending on the nanoparticle type).

1.4.3 The observation of a primary hydrogen-deuterium isotope effect

A second advantage of single molecule experiments is that the lifetime of non-rate determining events or short lived steps, obscured during ensemble measurements, can be obtained. For example Lu et al. (2010) were able to observe a primary hydrogen-deuterium isotope effect by following the reaction of a single cysteine thiol reacting with a quinone molecule (see Scheme 1.5). In contrast to the ensemble ^{1}H NMR experiments, the isotope effect was observable from the single molecule experiments, regardless of whether this step was rate determining. The ensemble experiment only recorded the overall reaction rate making the equilibrium between unreacted ubiquinone and thiolate invisible unless rate determining, whereas at the single molecule level, all reaction pathways could be directly observed.

1.4.4 Rate fluctuations during enzymatic turnover of λ-exonuclease

Oijen et al. (2003) showed how the procession rate of λ-exonuclease along a piece of DNA was not constant with respect to time. Double stranded DNA (dsDNA) was attached to a poly(ethylene glycol) coated glass surface. A polystyrene bead was attached onto the other end of the DNA

and a flow cell set-up was employed to apply a low force onto the bead (0.1 - 1 pN) and stretch out the DNA strands in a linear fashion against the slide. As the dsDNA was converted into ssDNA by λ-exonuclease, the DNA strands coiled up (because ssDNA requires more force to remain stretched out than dsDNA). By calibration of this coiling on conversion to ssDNA at a constant applied force, changes in the position of the bead with respect to time were recorded and correlated to the processivity (i.e. number of 5' mononucleotides removed from the dsDNA per second) of the λ-exonuclease enzyme. Additional to the rate fluctuations with respect to time, it was found that the average processivity of the λ-exonuclease was 32 nt/s which is higher than ensemble measurements (i.e. 12 nt/s, Subramanian et al. (2003)). The difference was attributed to the occurrence of multiple enzymes on one strand of DNA, slow binding of λ-exonuclease and enzymatic processivity distance variations, all of which cannot be controlled during a bulk experiment in contrast to the single molecule experiments (Oijen et al., 2003).

1.5 Single molecule techniques

The three main single molecule techniques that have been used to study small molecules at ambient temperatures under aqueous conditions are fluorescence spectroscopy, atomic force microscopy and single channel electrical recordings. Examples have been provided for each of these techniques.

1.5.1 Fluorescence spectroscopy

One of the first single molecule fluorescence experiments was carried out by Shera et al. (1990), where individual rhodamine-6G molecules were imaged in a fluid flow cell. Now, there are many different single molecule fluorophores that exist for the imaging of single molecules, including fluorescent dyes (Molecular Probes, 2011), autofluorescent proteins (Chudakov et al., 2010),

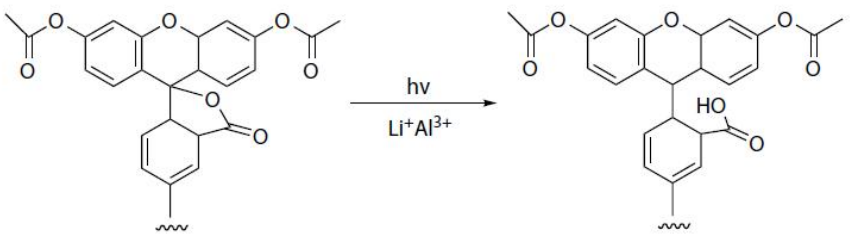


Figure 1.5: Esterolysis of a fluorocein ester. Adapted from Roeffaers et al. (2006)

nanocrystals (Rosenthal et al., 2011) and small genetic tags complexed with fluorochromes (Chiu & Leake, 2011). There are also a diverse number of ways to image using these fluorophores, including FRET (Roy et al., 2008, förster resonance energy transfer) , FRAP (Sprague & McNally, 2005, Fluorescence recovery after photobleaching) and TIRF (Schneckenburger, 2005, total internal reflection fluorescence). A second example of single molecule fluorescence was the real-time imaging of an organic reaction, carried out by Roeffaers et al. (2006). The esterolysis of a fluorescein ester (see Figure 1.5) by a [Li^{+}-Al^{3+}] layered, double hydroxide catalyst, was followed by monitoring the fluorescence emitted on hydrolysis of the ester. Roeffaers et al. (2006) observed which regions of the hydroxide surface contributed towards catalysis. They also obtained diffusion parameters and rate constants for both the trans-esterification reaction and the ester hydrolysis that took place after submerging the crystals in aqueous media. The ability to study catalytic surfaces, commonly employed during industrial organic transformations, under ambient conditions, is one of the advantages of fluorescence spectroscopy. Fluorescence microscopy is however limited by the necessity for some sort of fluorescing tag which could potentially perturb the dynamics of a small molecule.

1.5.2 Force microscopy

Atomic force microscopy (AFM) was developed in the 1980's (Binnig et al., 1986) and has provided valuable information regarding the topology of surfaces. Over the last decade, single molecule

atomic force microscopy has increasingly been employed to investigate the effect of mechanical strain on chemical reactivity and molecular conformation. The application of atomic force microscopy (AFM) to study and track single molecules has diversified since the first room-temperature images of single molecules immobilized on a surface (Betzig & Chichester, 1993), and it has been possible to image DNA molecules in solution (Lindsay et al., 1989), observe the unfolding of titin immunoglobulin domains (Rief et al., 1997), follow the motion of a single kinesin moving along a microtubule (Svoboda et al., 1993) and investigate the complete unfolding trajectory of a single ubiquitin molecule (Fernandez & Li, 2004). More recent examples include observing conformational changes of single (dextran) molecules (Walther et al., 2006), probing the energetics of host-guest binding using a series of immobilised macrocyclic compounds and measuring the force required to remove the host ion, for example ammonium (Eckel, 2005). At the single bond level, atomic force microscopy was used by Grandbois et al. (1999) to measure the force required to rupture single covalent Si-C and S-Au bonds of an amalyose polymer, anchored between the AFM tip and either a glass or gold surface respectively and submerged in a phosphate buffered saline solution at pH 7.4. More recently, AFM has been employed to investigate the effect of mechanical strain on the reactivity of single disulfide bonds located at different within a polypeptide. In these experiments the rate of reduction of the disulfide bonds was calculated over a range of applied forces. It was discovered that past a certain applied force, an abrupt change in the sensitivity of the reaction rate towards force (Garcia-Maynes et al., 2009). Single molecule AFM techniques are versatile in terms of the range of forces that can be applied or measured and also in terms of spatial resolution and sensitivity (Neuman & Nagy, 2008).

1.5.3 Single channel electrical recordings

The single channel electrical recording technique was developed in the 1960's. This technique involves placing a single membrane protein channel into a planar bilayer, applying a potential

difference across the bilayer, and monitoring the behaviour of the protein via changes in the flow of ions (the current) through the channel. The current fluctuations (events) over time, provide information regarding the properties of the ion channel. The single channel recording set-up was adapted for the study of single molecules interacting non-covalently with ion channels. Instead of observing current fluctuations due to the protein ion channel, the channel was used to create an environment in which single molecules, interacting with the protein either covalently or non-covalently, could be observed. Many different biological ion channels have been used to monitor single molecules (Grosse et al., 2011) and an early example was the observation of amino acid side chain protonation states in an L-type Ca^{2+} channel (Prod'hom et al., 1987). One of the first single molecule experiments was carried out using the channel forming peptide gramicidin A, where the cis-trans thermal isomerisation of a covalently attached carbamate was observed by recording the stochastic changes in current amplitude over time (Jaikaran & Woolley, 1995). An advantage of single channel electrical recordings is that they allow the monitoring of reactions/ interactions in an aqueous environment with non-perturbed small molecules (ie no tags). The detection of antibodies using the porin OmpF (Nestorovich et al., 2002), monitoring of a chemical reaction (Blake et al., 2006), drug sensing and porphyrin detection (Heins et al., 2005) using the ion channel Kapton are a selection of experiments utilising ion channels for single molecule studies.

1.6 αHL as a nanoreactor for single molecule studies

1.6.1 Properties of α-hemolysin

αHL is secreted by *Staphylococcus aureus* as a 33.2 kDa water soluble monomer. The monomeric subunits assemble in the presence of cellular and/or synthetic membranes to form a 232.4 kDa heptameric transmembrane spanning pore (Gouaux et al., 1994). *In vivo*, αHL lyses cells by forming channels in the host cell membranes. As illustrated in Figure 1.6, the heptameric αHL protein has a

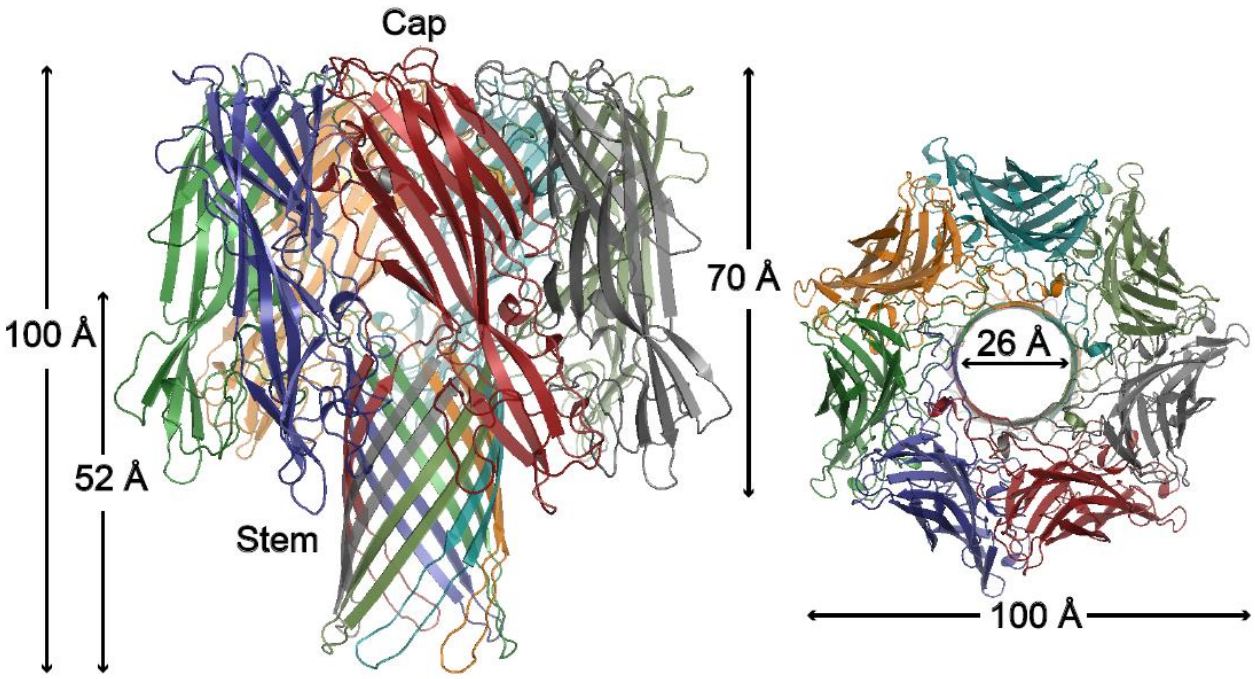


Figure 1.6: The αHL heptameric structure. Adapted from Song et al. (1996)

cross-sectional area of 100 Å^2 and the bilayer spanning, β barrel (or stem), portion of the heptamer is 52 Å in length. The crystal structure of the αHL heptamer (Song et al., 1996, 1.9 Å resolution) allowed the identification and subsequent mutation of residues facing inwards, within the β barrel. The central aqueous channel through which ions can flow has an internal diameter ranging from 13 Å to 26 Å. The effective channel diameter based on conductance measurements is thought to be around 11 Å meaning that the ions within the pore behave as they would in bulk solution. Protruding above the bilayer is the cap domain. Inside the cap, there exists a bowl shaped volume leading to the narrowest part of the pore (the constriction) situated at the top of the β barrel. Single channel recordings are carried out by insertion of a single αHL heptameric pore into a planar bilayer. The voltage applied across the bilayer creates a net flow of ions through the αHL pore and this current is recorded with respect to time. The monomeric and heptameric αHL proteins are stable at room temperature where the monomer is still hemolytically active after multiple freeze thaw cycles. Single channel electrical recordings require only nanogram amounts of αHL heptamer, allowing

for time-efficient expression using *Escherichia coli in-vitro* transcription/translation (IVTT). The heptameric αHL protein is stable in the presence of SDS allowing for facile purification by gel electrophoresis.

1.6.2 Advantages of using αHL as a nanoreactor

The use of the αHL nanoreactor for sensing small molecule analytes was first demonstrated by Walker et al. (1994). Site directed mutagenesis altering residues 123, 126, 133 and 135 to histidines (Braha et al., 1997), created a divalent metal ion binding site, or 'sensor' within the β barrel of αHL. There are a number of reasons for choosing to use αHL as the ion channel in which single molecule interactions can be studied. Intrinsic properties of the pore render it an attractive choice. For example, the conductance of the αHL pore is only very slightly rectifying, meaning that it exhibits almost no preference for cations or anions over a wide applied-voltage range. This is important when studying the kinetics of reactions, so as not to create a bias in the diffusion of small molecules through the pore. The heptamer remains stable and electrically quiet over a wide range of both pH (Gu & Bayley, 2000) and temperatures (Kang et al., 2000) and exhibits low background noise, allowing even small current amplitude, short lived events, to be detected. When a small molecule analyte in solution interacts with residues within the β barrel, the current flow through the pore is perturbed. If the pore were of larger or smaller dimensions this sensitivity would be lost and either the interactions would be unobservable or the small pore diameter would sterically hinder the reaction taking place. Examples of molecules in solution, observed at the single molecule level via interaction with αHL include, the binding of cyclodextrin molecules (Gu et al., 1999, 2001), the use of cyclic peptides as adapters within the β barrel for the binding of analytes (Sanchez-Quesada et al., 2000), divalent metal cation binding (Kasianowicz et al., 1999), single molecule nanocavity trapping (Gu & Cheley, 2001), triphosphate sensing (Cheley et al., 2002), protein-ligand interactions (Howorka et al., 2004), a DNA tether as a distance probe

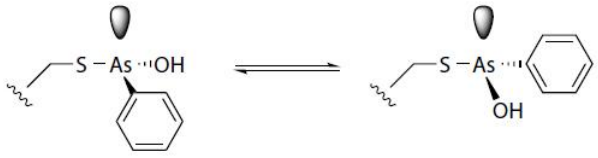


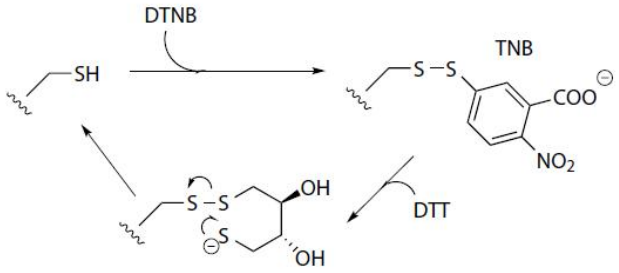
Figure 1.7: Diastereomeric species formed on reaction of the organoarsenic reagent SPAO with a cysteine residue situated within the β barrel of αHL. Diastereomeric species are generated due to the chiral nature of the β barrel of αHL (Shin et al., 2007).

(Howorka & Bayley, 2002), metal cation chelation by EDTA (Hammerstein et al., 2010) and the application of αHL towards the sequencing of DNA (Branton et al., 2008; Gouaux et al., 2006).

1.6.3 Covalent reversible and irreversible reactions carried out using αHL

The reaction of a single cysteine residue situated in the β barrel of αHL, with an organoarsenic(III) species

The first observed single molecule reaction of a cysteine residue within a protein environment (the diphtheria toxin channel-forming domain) was the work of Mindell et al. (1994) where the reaction of cysteine with methyl sulfonate was observed and kinetic rate constants obtained for this process. Shin et al. (2002) demonstrated the feasibility of studying reversible covalent-bond forming chemistry within the β barrel of αHL. The rate constants for the reaction between a cysteine thiol group and an organoarsenic (III) compound were obtained. The ability to observe the As (III) compound reacting with just one thiol group illustrated how the covalently attached organoarsenic compound produced two distinct current blockage levels (Shin et al., 2007, see Figure 1.7). These were attributed to two diastereomeric intermediates, formed on reaction with the cysteine residue, due to the chirality of the organoarsenic (III) molecule attached within the β barrel of αHL.



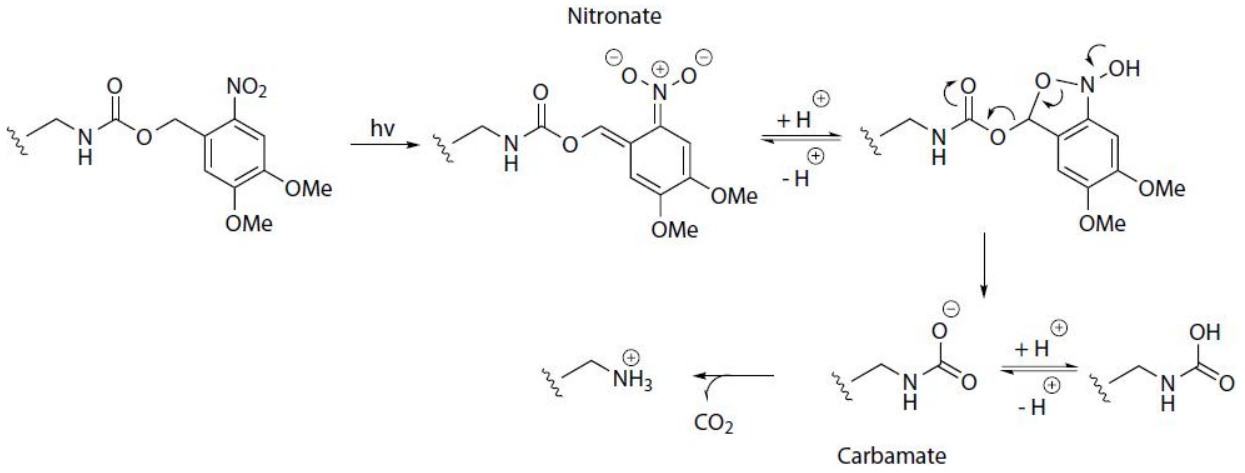
Scheme 1.6: The reaction of a cysteine residue within the β barrel of αHL with DTNB and DTT, where the reagents are spatially separated (Luchian et al., 2003b).

Multiple turnovers observed for irreversible reaction steps

Luchian et al. (2003b) showed how it was possible to use the intrinsic compartmentation of the electrical recording set up to observe the irreversible covalent reaction of the cysteine residue within the β barrel of αHL, with the aromatic disulfide, DTNB, followed by the reduction with DTT (see Scheme 1.6). By spatially separating the two reagents, it was possible to observe multiple turnovers of what would be an irreversible reaction in solution. Additionally, the short lived intermediate formed after attack of DTT on the disulfide bond between the cysteine residue and TNB (2-nitro-4-carboxybenzoic acid) was observed for the first time.

Observation of short lived intermediates regardless of where they appear in the reaction pathway

The covalent reactions observed using αHL have not been limited solely to disulfide chemistry. Luchian et al. (2003a) followed the three-step re-arrangement and cleavage of a 3, 4-dimethoxy-6-nitrobenzylcarbamate group tethered inside the αHL pore, where two short-lived intermediates (nitronate and carbamate) were observed after photolysis (See Figure 1.7). Rate constants were obtained for the reactive intermediates demonstrating how single molecule measurements for a fast step that follows one or more slow steps, could be obtained using the αHL pore. Other reactions investigated using the αHL pore as a nanoreactor include, the first single molecule observation



Scheme 1.7: The photolysis of a carbamate group tethered within the β barrel of αHL(Luchian et al., 2003a).

of step-wise polymer growth (Shin & Bayley, 2005), the cis-trans isomerisation of azobenzene (Loudwig & Bayley, 2006) and the rate alterations arising from a primary hydrogen-deuterium effect on the reaction of quinone with a cysteine residue (Lu et al., 2010, see Section 1.4.3).

1.7 Thesis summary

In chapter one, this thesis firstly explains the importance and origin of single molecule experiments. The next sections focus on the application of, and types of, single molecule experiments used to investigate small molecules in aqueous media. The final section describes the use of single channel electrical recordings as a method for stochastic sensing, with a focus on the application of αHL as a 'nanoreactor' in which both reversible and non-reversible covalent reactions can be observed. Chapter 2 describes the method by which the β barrel of αHL was engineered so as to contain a single selenocysteine residue and details the various characterisation methods employed to verify the presence of this residue. In Chapter 3, single channel electrical recording experiments were used to study the reaction of the single selenocysteine, within the β barrel of αHL, with small

molecule disulfide and thiol containing species in solution. By carrying out complementary single molecule experiments using an αHL pore containing a single cysteine residue, it was possible to compare the nucleophilicity, electrophilicity and leaving group ability of selenocysteine and cysteine. In Chapter 4 single channel electrical recording experiments were carried out to obtain rate constants for the reaction between an organoarsenic (III) reagent and the single selenocysteine residue within the β barrel of αHL. Rate constants for an analogous reaction between a single cysteine containing αHL pore and the organoarsenic (III) species were compared to those acquired for the selenocysteine-containing pore over the pH range 6 - 8.5. Chapter 5 summarises future single molecule experiments that could be carried out using the selenocysteine containing αHL pore.

Chapter 2

Incorporation of selenocysteine into the β barrel of αHL

2.1 Introduction

2.1.1 Methods to incorporate selenocysteine into proteins

The incorporation of unnatural amino acids into proteins has applications across many fields of research, for example to attach fluorescent tags (Turcatti et al., 1996) or to introduce an orthogonal reactive group for selective chemical modification (Wang 2006). Various genetic code reassignment technologies exist (Anderson et al., 2005; Normanly et al., 1990; Rodriguez & Dougherty, 2006; Xiao et al., 2008). The applications for selectively modifying an amino acid are broad, ranging from imaging (Wang & Xie, 2006; Pantoja et al., 2009; Laverman et al., 2002; Lee & Schultz, 2008), to the mechanistic probing of protein structure and function (Lee et al., 2009; Seyedsayamdost et al., 2007). A summary of the methods that were considered as viable for selenocysteine incorporation into αHL are detailed in the following sections.

2.1.2 *In vivo* selenoprotein expression

Auxotrophic expression

An auxotrophic system is one in which an organism is unable to produce a certain molecule or species itself. The auxotrophic bacterial strain would be chosen so as to be incapable of synthesising cysteine and so supplementation with selenocysteine would result in expression of selenoproteins (Jiang et al., 2009; Bock, 2005). If this method were used to incorporate selenocysteine into αHL it would involve the mis-acylation of $tRNA^{cys}$ with selenocysteine allowing selenocysteine incorporation in response to the cysteine codon. A limitation of this approach is that trace amounts of cysteine may cause the production of protein containing a mixture of selenocysteine and cysteine, for example the percentage of selenocysteine replacement achieved by Muller et al. (1994) was between 73 - 80 %. It was imperative that the αHL selenocysteine containing protein was not contaminated with the cysteine containing αHL analogue and separation of a heterogeneous mixture would be difficult due to the reactivity similarity between selenocysteine and cysteine. It was concluded that this method was feasible but a separation method would need to be devised so as to remove any contaminating cysteine containing αHL protein. A failure rate of approximately 20 - 30 % of proteins containing an incorrect amino acid at the chosen position was thought to be reasonable seeing as the protein would be used for single molecule experiments.

Recombinant expression utilising the SECIS element

By engineering a gene to contain the selenocysteine, TGA codon and special mRNA secondary structure (SECIS element), Aldag et al. (2009) were able to recombinantly express a selenocysteine containing protein. A selenocysteine containing p450 enzyme was produced by the co-expression of a plasmid containing the engineered p450 gene, and one containing the bacterial selenocysteine incorporation machinery (ie the SelA-D genes, see Introduction Section 1.2). One limitation of the

method employed by Aldag et al. (2009) is that the conserved base sequence encoding the SECIS element must be located within the reading frame of the gene, placing restrictions on the base sequence, potentially resulting in unwanted additional amino acid alterations. The engineering of a bacterial SECIS element into the rat Thioredoxin gene was carried out by Arnér (2002) so as to express the mammalian protein in *Escherichia coli* but the close proximity of the SECIS element to the TGA codon was not a problem in this case as the TGA codon was at the N-terminus of thioredoxin and so the SECIS element resided outside of the reading frame. A concern with the use of this method was that if the αHL gene were engineered to contain the SECIS element, these unwanted amino acid alterations potentially could introduce folding, stability or expression complications. An alternative method was investigated by Su et al. (2005) where the bacterial SECIS element was positioned outside of the reading frame by creating a very long hairpin loop stem region. The SECIS element was then placed at the end of this stem, 36 base pairs away from the TAG protein translational stop codon. It was found that the combination of TGA and SECIS element were still able to direct selenocysteine incorporation but only approximately 3 % of the expressed protein actually contained the selenocysteine amino acid. Opal suppression of the TGA codon by $tRNA^{Trp}$ was found to have dominated over translation termination, resulting in approximately 97 % of the protein containing tryptophan instead of selenocysteine. This has been previously reported to occur when the SECIS element is not able to function correctly (Sandman et al., 2003). An advantage of engineering a SECIS element into the αHL gene would be the elimination of non-specific amino acid incorporation as long as the SECIS element was functioning correctly. It would also ensure that a heterogeneous mixture of selenocysteine and cysteine containing αHL protein could not occur.

2.1.3 *In vitro* selenoprotein expression

The PURE system

The PURE system (Promega) or reconstituted *Escherichia coli* cell-free translation system consists of purified components (minus the DNA template and nuclease-free/ protease-free) necessary for transcription and translation of plasmid DNA (Shimizu et al., 2005). By employing the PURE system, it would theoretically be possible to omit cysteine and add selenocysteine instead, whereby the $tRNA^{Cys}$ would be aminoacylated with selenocysteine by the $tRNA^{Cys}$ synthase. As with the axotrophic strain method (see Section 2.1.2) a limitation is the reliance on complete eradication of cysteine.

In Vitro transcription translation (IVTT)

In vitro transcription translation (IVTT) systems are similar to the PURE system, but employ the use of concentrated cell extract as opposed to purified components. All the required components for protein transcription and translation are present. After addition of the gene of interest, the system is incubated for 1.5 hours at 37 ^{o}C and during this time protein translation takes place. The IVTT system has been previously used in conjunction with unnatural amino acid incorporation via addition of an aminoacylated $tRNA^{TAG}$ (Banerjee et al., 2010, for example). An advantage of these systems is the speed at which the protein can be produced. A limitation is that the quantity of protein produced is usually 5 μg per 100 μl IVTT reaction and so this method is impractical for large scale protein production. The IVTT system is designed to incorporate the twenty canonical amino acids and it was not known if the kit would be able to recognise the SECIS element or whether it contained the additional components required for selenocysteine incorporation (see Introduction Section 1.2).

2.1.4 Semisynthetic incorporation methods

Semisynthetic unnatural amino acid incorporation involves a combination of synthetic and recombinant approaches. There are a number of semisynthetic methods and those considered for selenocysteine incorporation into αHL have been summarised below.

Native Chemical Ligation

Native chemical ligation (NCL) is a semi-synthetic method for making a protein. It involves the chemical joining together of synthetic and recombinant polypeptides. NCL has been used to incorporate a variety of unnatural amino acids (for reviews see Kent (2009); Hofmann & Muir (2002)). This technique is usually carried out on smaller proteins due to limitations of solid phase peptide synthesis and re-folding complications associated with larger proteins. The joining of recombinant and synthetic polypeptides was carried out by Berry et al. (2002) to make a selenocysteine containing analogue of azurin, which is a type I copper binding protein. An advantage of NCL is that it can be used to generate milligram amounts of protein, allowing for characterisation via mass spectroscopy and contains a synthetic peptide, allowing unnatural amino acids, unsuited to conditions required for recombinant expression, to be incorporated into the protein.

Chemical post-translation modification

Post-translational modification of a serine residue to a selenocysteine residue inside the active site of Subtilisin via chemical activation followed by addition of SeH_2, was carried out by Wu & Hilvert (1989). Ren et al. (2002) similarly modified a serine residue to selenocysteine in a mammalian (rat) glutathione peroxidase, where NaHSe was used. The main drawback to post-translational incorporation via activation of a serine, is that there are many serine residues in αHL and so selective modification of just one of these may be challenging.

2.1.5 Incorporation methods summary

After considering the advantages and limitations of the above methods, it was decided that the recombinant method, involving the engineering of a SECIS element into the αHL gene was most viable for the incorporation of selenocysteine into the β barrel.

2.2 Results

2.2.1 Engineering the SECIS element and TGA codon into the αHL gene

So as to express the selenocysteine containing αHL monomer, the bacterial SECIS element (originating from the *Escherichia Coli* formate dehydrogenase H) was incorporated into the αHL gene. The SECIS element nucleotide sequence chosen, was that of the minimal SECIS element. The minimal SECIS element used here, was devised by Liu et al. (1998) and contains only the most essential nucleotides and base pairing interactions for selenocysteine incorporation in response to the TGA codon. The TGA codon was always engineered 11 nucleotides upstream of the SECIS element. This is the optimum distance for selenocysteine incorporation in response to the TGA and SECIS element during bacterial selenoprotein translation *in vivo*. The requirements for selenocysteine incorporation into proteins can be found in Introduction Section 1.2.

Construct design

A pT7, 4.6 kb plasmid containing the monomer αHL gene under control of the T7 promoter was the template DNA used to engineer the SECIS element and TGA codon into αHL (see Figure A.1). Restriction sites have been previously engineered at strategic positions within the αHL gene. The NdeI (NEB) and HindIII (NEB) as shown in Figure A.1 are situated at the N-terminus

αHL residue number	Sequence alterations
113 M113Stop S114L TGA CTC	T117 Y118L G119A F120G N121G G122Q N123G ACG CTT GCA GGT CTG CAG CGT
117 T117Stop Y118L TGA CTC	G122L N123A V124G T125L G126Q CTT GCA GGT CTG CAG
119 G119Stop F120L TGA CTC	N123Q V124 T125A G126 D127L D128H T129L CAG GTT GCA GGT CTG CAC CTG
137 G137Stop A138L TGA CTC	S141 I142L G143A H144A T145L L146Q K147R TCG CTT GCA CGT CTG CAG CGA

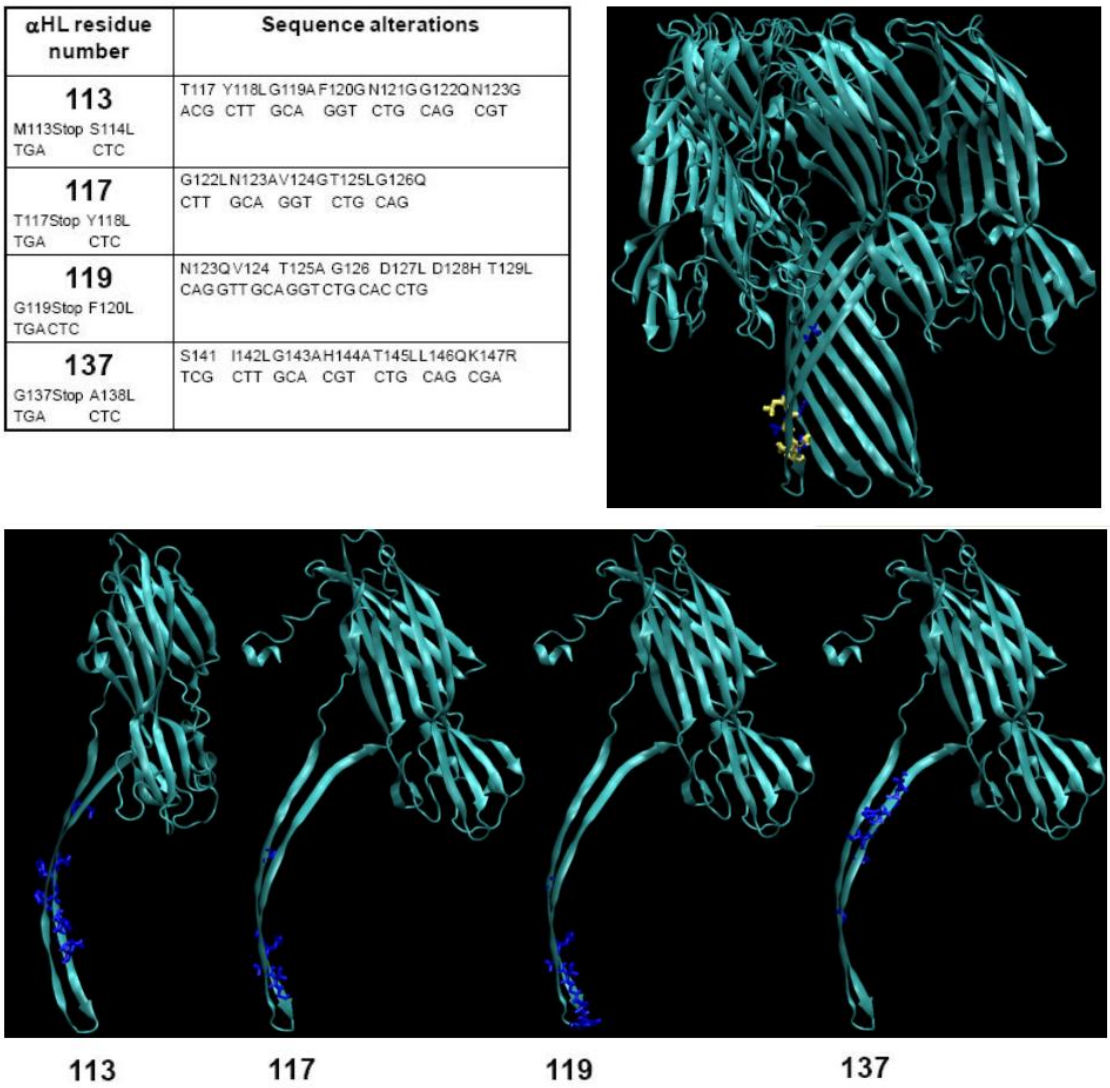


Figure 2.1: Alterations in the base sequence of the αHL gene so as to introduce the Sec codon and SECIS element resulted in non-silent codon changes. The amino acid alterations have been tabulated for each construct. The structures shown illustrate the position of these amino acid alterations and also the positioning of the SECIS element within the β-barrel of αHL.

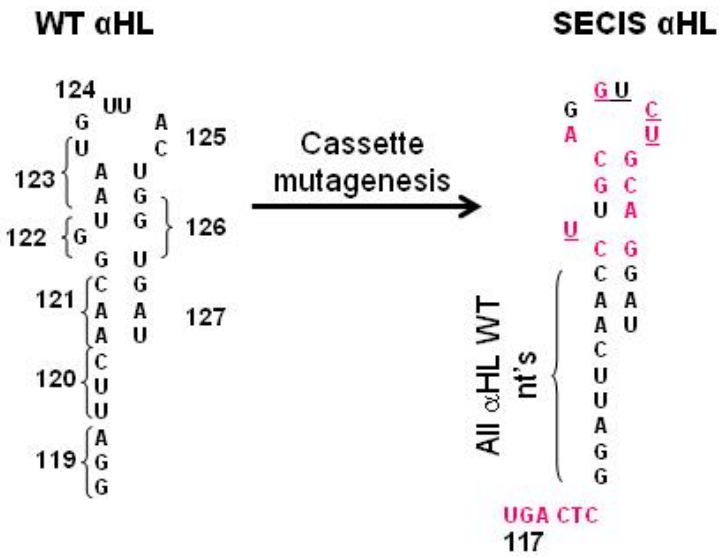


Figure 2.2: Engineering of the αHL gene so as to contain the minimal SECIS element. Bases that were altered from the original sequence has been hi-lighted in pink. Essential bases required for SelB recognition and/or binding have been underlined.

and C-terminus of the αHL gene respectively. Other restriction sites that were previously engineered into the αHL gene are SacII (NEB), StuI (NEB) and XhoI (NEB) and these are located along the β barrel region. This plasmid construct will be referred to as pT7WTD8. The restriction sites SacII, StuI and XhoI were used to introduce the SECIS element via over digestion of the pT7 plasmid containing the αHL gene followed by 3-way ligation using primers containing the SECIS element bases and also the ACA (cysteine) codon, 11 bases upstream of the SECIS element. The TGA codon was introduced separately via site directed mutagenesis and instead, the ACA codon (cysteine) replaced the TGA codon. The reason for initially using the cysteine codon instead of selenocysteine was to make sure that the monomeric αHL protein would still express, fold and oligomerise despite the amino acids alterations introduced due to the SECIS element nucleotide sequence requirements. Cassette mutagenesis was used to incorporate the SECIS element sequence and cysteine encoding codon into the αHL gene. The general method used to construct each mutant has been illustrated in Figure A.2. Three way ligation using 4 primers of lengths ranging from 35-45 bases (see Figure A.6) was carried out. Primers 2 and 3 were initially 5'

phosphorylated using T4 polynucleotide kinase (Promega). The pT7 plasmid containing the αHL gene was linearised by over digestion with StuI (NEB) and SacII (NEB) and the doubly cut vector, gel purified. The ligation reaction was carried out by incubation of primers 1-4 (see Figure A.2) and the StuI/SacII linearised vector. Transformation into XL10 Gold competent cells was carried out, followed by DNA sequence verification (Source Bioscience DNA sequencing service). The three-way ligation reaction used four different sets of primers so as to place the cysteine codon at positions 113, 117, 119 and 137 and the SECIS element 11 bases downstream from these positions. The plasmids containing each of these αHL constructs will be referred to as pT7113CysSECISD8, pT7117CysSECISD8, pT7119CysSECISD8 and pT7137CysSECISD8 where the "D8" refers to the eight aspartate residues encoded at the 3' end of the αHL genes. Using standard overlap extension PCR, site directed mutagenesis was carried out on each of the pT7CysSECISD8 constructs to replace the cysteine codon with a TGA (selenocysteine) codon i.e. at positions 113, 117, 119 and 137 in the β barrel (see Figure 2.2). As illustrated in Figure 2.1, a total of four new constructs were designed with the TGA codon at positions 113, 117, 119 and 137 in the β-barrel followed by the SECIS element 11 nucleotides downstream. These constructs were named pT7113SecSECISD8, pT7117SecSECISD8, pT7119SecSECISD8 and pT7137SecSECISD8. For purification reasons, six codons encoding a histidine tag (H6) were later added to each of the constructs directly after the D8 tail, via homologous recombination PCR (see Figure A.3). Homologous recombination was carried out on the 117 and 119 constructs only, generating D8H6 tagged constructs called pT7117SecSECISD8H6, pT7119SecSECISD8H6, pT7117CysSECISD8H6 and pT7119CysSECISD8H6. A H6 tagged version of the pT7WTD8 plasmid will be referred to as pT7WTD8H6. The plasmids were all sequenced verified (Source BioScience).

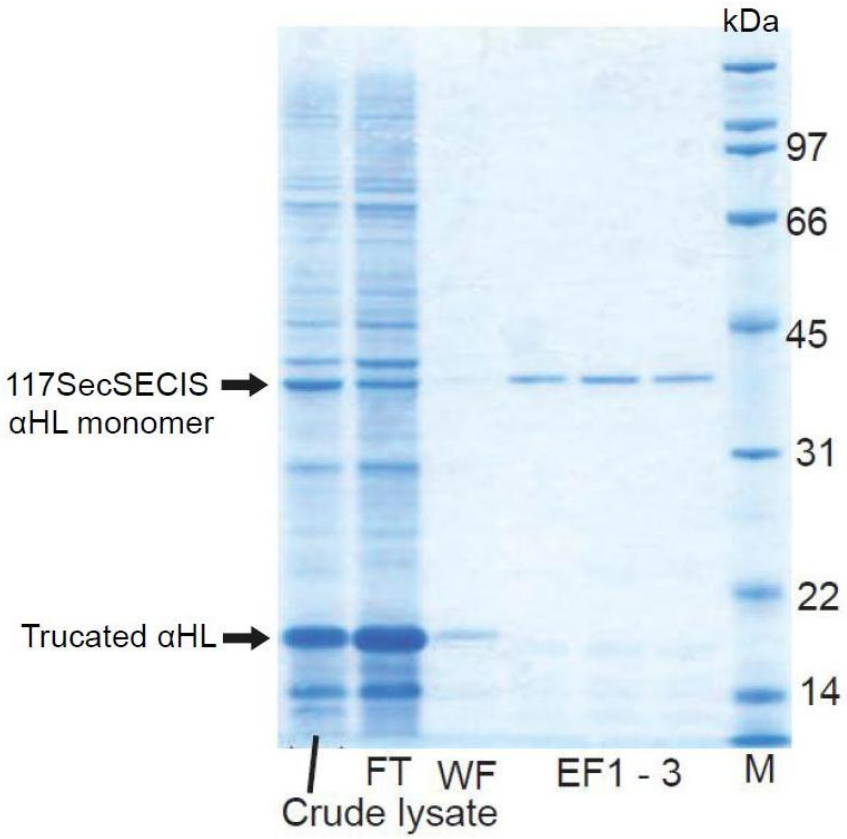


Figure 2.3: The selenocysteine containing αHL monomer was purified via His-tag technology. The flow through (FT), wash fraction (WF) and elution fractions (EF's) can be seen. The ratio of full length selenocysteine containing αHL to truncated αHL can be seen to be 1:5 before separation.

2.2.2 Protein Expression and Characterisation

Recombinant protein expression

The pT7117SecSECISD8H6, pT7117CysSECISD8H6 and pT7WTD8H6 plasmids were transformed into BL21(DE3) (Stratagene) *Escherichia coli* competent cells and used to express αHL monomeric protein containing a cysteine, selenocysteine or tyrosine (WT sequence amino acid) residue at position 117. These monomeric proteins will be referred to as 117CysSECIS, 117SecSECIS and WT respectively. The 117SecSECIS expression culture was supplemented with cysteine and selenite on induction with IPTG (Isopropyl β-D-thiogalactoside). See 2.3.1 for further details. All three monomeric αHL proteins were purified under denaturing conditions using Ni-NTA resin technology (See Figure 2.3). The WT, 117CysSECIS and 117SecSECIS αHL monomers were either oligomerised to form heptameric αHL protein for use during the single channel experiments (see Chapter 3) or employed for characterisation experiments (see Section 2.2.2). Oligomerisation of seven monomeric αHL units was carried out in the presence of rabbit red blood cell membranes (see Figure A.4). By mixing different ratios of each monomeric αHL construct, it was possible to form heteroheptamers consisting of between none and seven of each monomer (depending on the ratio of monomers mixed together during oligomerisation). By mixing non-D8H6-tagged WT αHL monomer with the 117SecSECISD8H6 monomer, in a ratio of 6:1, heteroheptamers were obtained which contained different ratios of 117SecSECIS monomer with respect to WT monomer. As the single channel experiments required only a single selenocysteine or cysteine residue to be situated within the β barrel of the αHL pore, gel electrophoresis (5% SDS PAGE gel) was used to separate the different ratios of heteroheptamer, where the charge of the D8H6 tag, (only present on monomers containing the selenocysteine or cysteine residue) caused the protein to migrate through the gel further than a non-tagged αHL heptamer. The non-tagged WT monomeric units were radiolabeled with methionine S^{35} and after oligomerisation and gel electrophoresis, the gel was dried and exposed so as to locate the protein on the gel. An overview of the expression and purification

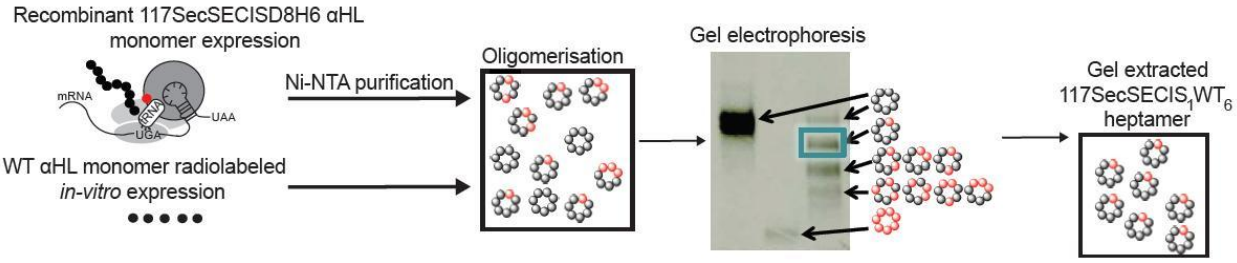


Figure 2.4: Purification protocol used to obtain a single selenocysteine-containing αHL heptamer. Red circles depict the D8H6 tagged monomers containing a single selenocysteine and black circles represent the non-tagged WT monomer units.

procedure described above, can be found in Figure 2.4. A typical picture of an exposed film has been shown in Figure 2.4 to illustrate heteroheptameric separation after gel electrophoresis. The heteroheptameric protein, containing a single selenocysteine or cysteine and six WT monomer subunits were named 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ respectively. By re-aligning the exposed film onto the dried gel, gel extraction of the 117SecSECIS$_1$WT$_6$ (or 117CysSECIS$_1$WT$_6$) heptamer was carried out and the protein samples stored under reducing conditions at -80 ^{o}C ready for use during the electrophysiology experiments.

LC-MS mass analysis

The Ni-NTA purified 117CysSECIS, 117SecSECIS and WT αHL monomer were de-salted and buffer exchanged via dialysis after which samples were spin-column concentrated to at least 0.5 mg/ml (14.3 μM). The selenocysteine residue of the 117SecSECIS monomer is known to readily form the seleninic acid (Se(O)(OH)) derivative of selenocysteine in solution and on exposure to oxygen and so an excess of DTT (approximately 50 mM) was added prior to mass analysis. Figure 2.5 shows the mass spectra obtained for the WT (expt. 34990, theoret. 34992), 117CysSECIS (expt. 34983, theoret. 34982) and 117SecSECIS (expt. 35032, theoret. 35029) αHL monomers, where the small peaks at 35073 and 35114 in the 117SecSECIS LC-MS spectra were attributed to salt adducts i.e. +K^+ (35032 + 39 = 35071) and +2K^+ (35071 + 39 = 35110) respectively as opposed

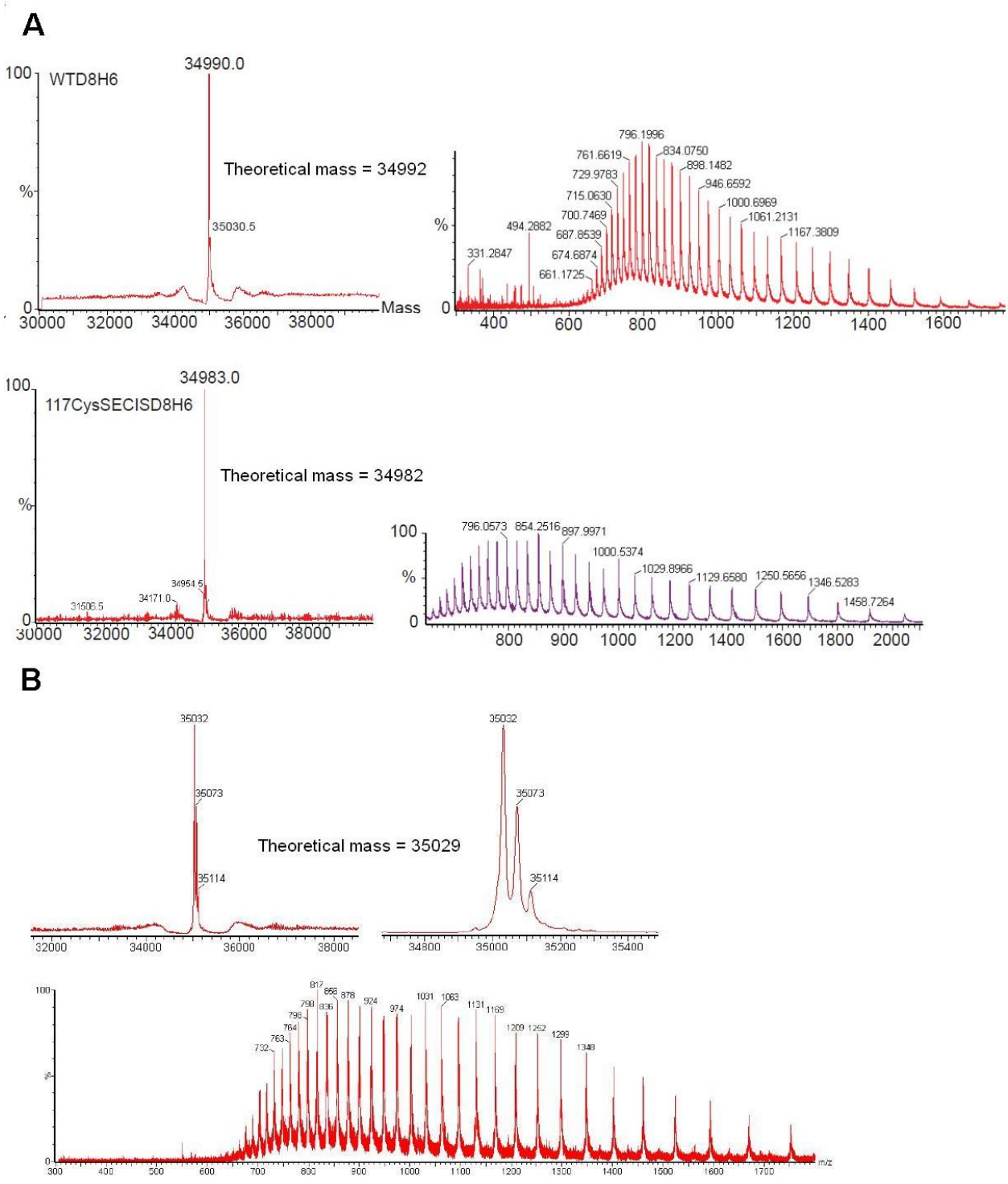


Figure 2.5: **A** The LC-MS spectra of the 117CysSECIS monomeric protein (Top). The LC-MS spectra of the WT monomer (Bottom). **B** The LC-MS spectra of the 117SecSECIS αHL monomer. The shoulders observed are thought to be due to salt adducts of K^+ and $2K^+$ respectively.

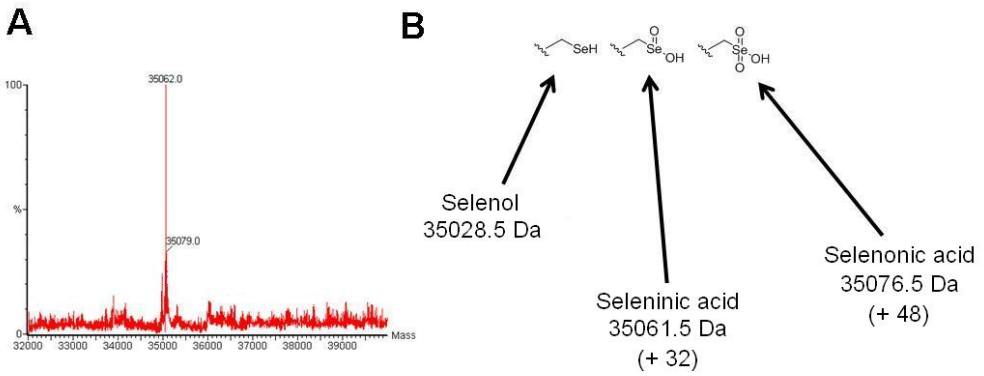


Figure 2.6: **A** The mass spectrum of 117SecSECIS before dialysis. **B** The seleninic, selenonic and selenonic acids and the corresponding 117SecSECIS mass that would be observed for each oxidation state respectively.

to oxidation of the selenol group (see Figure 2.6). The initial LC-MS spectra, obtained for the 117SecSECIS αHL monomer prior to dialysis showed a peak at 35062.0 ± 1 Da (see Figure 2.6). After dialysis, overnight at 4^oC into the solution containing 10 mM DTT, the same sample yielded the mass spectra shown in Figure 2.5 with a prominent peak at 35030 ± 4 Da. From these two mass spectra it was concluded that the peak at 35062.0 was due to oxidation of the selenocysteine residue (see Figure 2.6, **B**). The LC-MS for both the WTD8H6 and 117CysSECIS *Escherichia coli* expressed αHL monomer are shown in Figure 2.5. In contrast to the selenocysteine containing αHL monomer, only very low oxidation of the cysteine residue was observed both before and after dialysis. Similarly, a peak at 35030.5 was observed in the WTD8H6 spectrum and was attributed to an adduct of K^+ salt (i.e. 35030.5 - 34990.0 = 40.5).

Modification of 117SecSECIS and 117CysSECIS using IASD and MISD

So as to verify that the selenocysteine residue existed in its reduced selenol form, two thiol modifying reagents (IASD, 4-acetoamido-4'-((iodoacetyl)amino)stilbene-2,2'-disulfonate and MISD, 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid), see Figure 2.7 for structures) were reacted with the monomeric 117SecSECIS protein. The extent of modification was analysed by

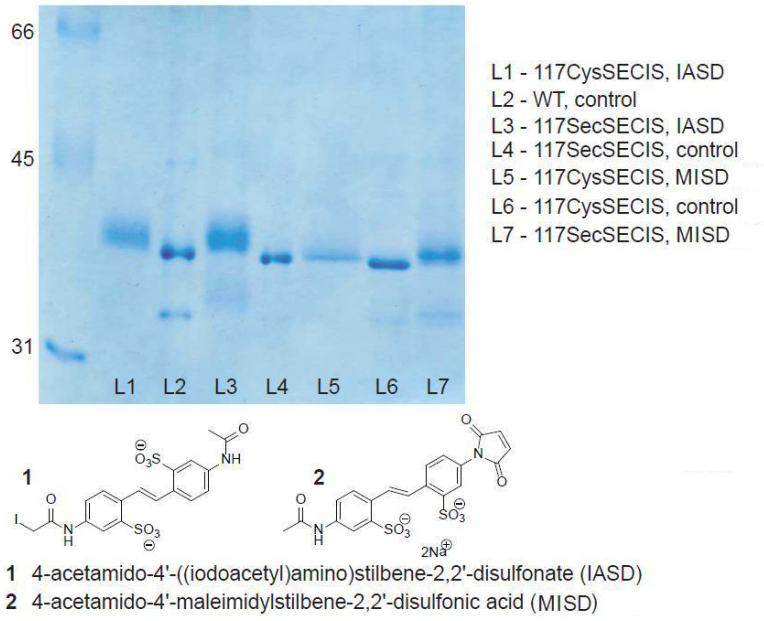


Figure 2.7: The selenocysteine and cysteine residues were reacted with both MISD and IASD so as to assess the activity of the selenol and thiol groups respectively. Lanes L1-7 illustrate how, on incubation with either MISD or IASD the 117SecSECIS and 117CysSECIS αHL monomer band shifts upwards relative to the unreacted control lanes.

12 % SDS PAGE gel. Figure 2.7 displays the gel shifts observed on reaction of the selenocysteine residue with MISD and IASD. Control reactions involved incubation of the 117CysSECIS and WT monomers with each of the modifying reagents. From comparison with the control lanes (see Figure 2.7), it can be concluded that the selenocysteine residue was in its reduced, selenol form directly after protein purification and so the oxidation observed in the mass spectrum prior to dialysis (see Figure 2.5) must have occurred during storage.

Oligomerisation and hemolytic activity of recombinantly expressed 117SecSECIS and 117CysSECIS αHL

A total of six amino acid alterations were introduced due to the SECIS element mRNA base sequence requirements (see Figure 2.1). Where possible, for example in the lower part of the stem of the hairpin structure, bases were chosen so as to maintain the original amino acid type (ie po-

lar/ non-polar, acidic/basic, size etc). By carrying out a hemolytic assay on the 113CysSECIS, 117CysSECIS, 119CysSECIS and 137CysSECIS monomeric protein expressed using the IVTT system, it was found that the new constructs were not able to lyse rabbit red blood cells (see Figure 2.8).

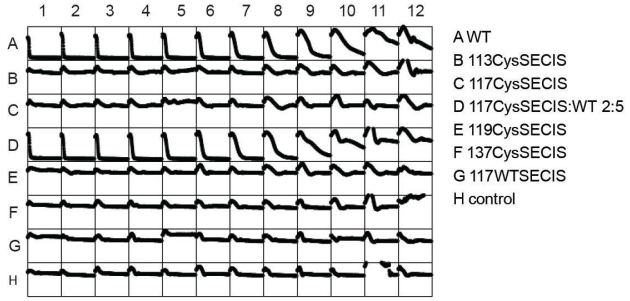


Figure 2.8: WT αHL heptameric pores are known to lyse rabbit red blood cells. The αHL monomeric subunits assemble on the surface of the cell membrane and form pores consisting of seven subunits and then insert across both membrane leaflets causing lysis. The lysis of rabbit red blood cells by addition of 5 μl of IVTT prepared WT, 113CysSECIS, 117CysSECIS, 117CysSECIS:WT (2:5 ratio), 119CysSECIS, 137CysSECIS and 117WTSECIS was investigated by recording the scattering of light by the red blood cells at 595 nm. Each well (for example, 1A) represents a single lysis experiment on the 96-well plate. For each well, the amount of scattered light (y-axis) was recorded at regular time intervals and plotted against time (x-axis). Red blood cells lysis was observed as a decrease in light scattering at 595 nm over time (as seen in 1A - 9A). The rate and extent of lysis by each of the mutant αHL monomers (i.e. rows B - G) was compared to cell lysis by WT αHL monomer (row A). Serial dilution was carried out across wells 1 - 12 for each of the A - H rows (see Methods 2.5.2 for further details). The control (row 1H - 12H) consisted of the addition of 5 μl of buffer.

This lack of lysis, was attributed to the number of consecutive amino acid residue alterations including the introduction of glycine residues and also hydrophilic residues (e.g. glutamine) facing out of the β barrel towards the hydrophobic hydrocarbon chains of the rabbit red blood cell lipid bilayer. To verify that it was the amino acid alterations, introduced due to the SECIS element base specifications, a 117WTSECISD8H6 construct was created. This mutant was also not hemolytically active and so ruled out aggregation of the 117SecSECIS or 117CysSECIS αHL monomers occurring i.e. due to disulfide or diselenide bond formation. A further hemolytic study using more concentrated 117CysSECIS monomeric, recombinantly expressed αHL protein, showed limited

lysis, suggesting that homoheptamer formation and insertion into the rabbit red blood cell membranes was possible, but was very slow and/or required high concentrations of monomer to assemble compared to lysis by homoheptameric WT protein. The β barrel of αHL has previously been the subject of extensive mutagenesis including work by Cheley et al. (1999) where the entire barrel peptide sequence was reversed and lack of hemolytic activity has been observed by (Walker & Bayley, 1995) due to mutagenesis of residues in the β barrel. The aim of incorporating selenocysteine into the β barrel of αHL was so as to investigate reactions of this residue with small molecules using single channel recordings. The heteroheptamer of monomeric ratio 1:6 (ie 117CysSECIS$_1$WT$_6$ and 117SecSECIS$_1$WT$_6$, see Figure A.4 and 2.4) was gel extracted and found to be capable of inserting into a synthetic DPhPC (diphytanoylphosphatidylcholine) bilayer (see Chapter 3).

2.2.3 *In vitro* translation of a selenocysteine containing αHL monomer

An *in vitro* method for the incorporation of selenocysteine into the β barrel of αHL was investigated. The aim of this work was to develop a fast method for selenoprotein synthesis, using the IVTT protein translation kit. The incorporation of selenocysteine into the growing polypeptide during protein translation requires additional components (see Chapter 1), to those utilised for incorporation of the twenty canonical amino acids. Initial experiments focused on ascertaining whether or not selenoprotein synthesis was possible using the IVTT system (E. coli S30 Extract System, Promega, Cat. L1130, see Introduction 2.1.3). Plasmids pT7113SecSECISD8, pT7117SecSECISD8 and pT7119SecSECISD8 were added to the IVTT system and incubated at 37 ^{o}C for 1.5 hours. Figure 2.9, **B** illustrates how no αHL protein expression was observed for any of the pT7113SecSECISD8, pT7117SecSECIS and pT7119SecSECIS constructs when incubated at 37^{o}C and only a truncated protein species was observed. A control pT7 plasmid containing an αHL construct with a TAG stop codon at position 117 was used (pT7117TAG). The observed

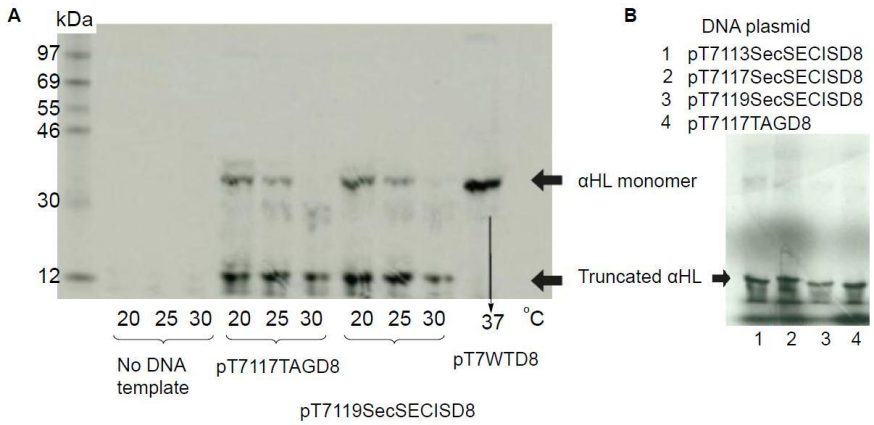


Figure 2.9: **A**; IVTT experiments using the 119SecSECIS DNA plasmid. IVTT reactions were carried out at 20, 25, 30 and 37^{o}C with either no DNA plasmid (controls), 117TAG αHL, or 119SecSECIS plasmid DNA added. Both the 117TAG and 119SecSECIS IVTT reactions exhibited read-through at temperatures 20^{o}C and 25^{o}C, however no full length protein was expressed at 30^{o}C or 37^{o}C. The WT αHL DNA plasmid was incubated at 37^{o}C. All reactions were carried out for 3 hrs. **B** 12 % Bis-Tris gel showing the IVTT reactions after 1.5 hrs incubation at 37^{o}C. It can be seen that only truncated protein was generated for all SecSECIS constructs (113SecSECIS, 117SecSECIS, 119SecSECIS and 117TAG).

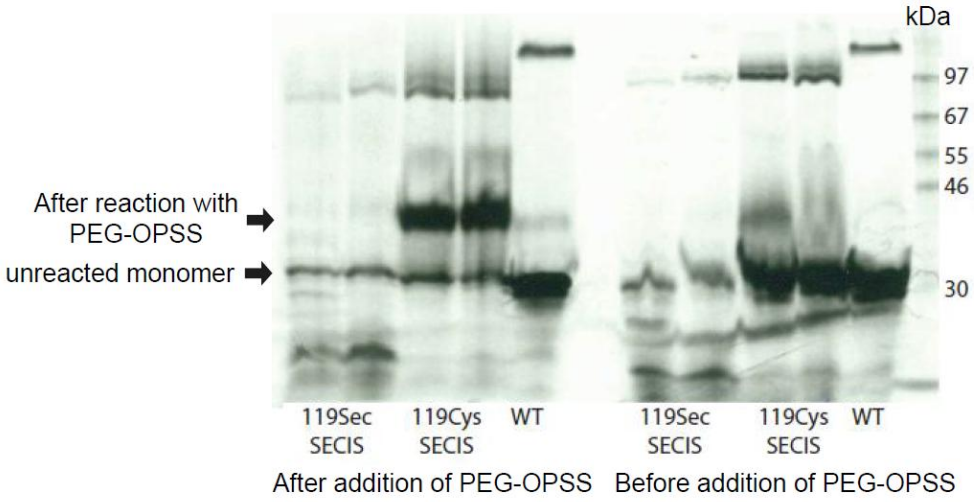
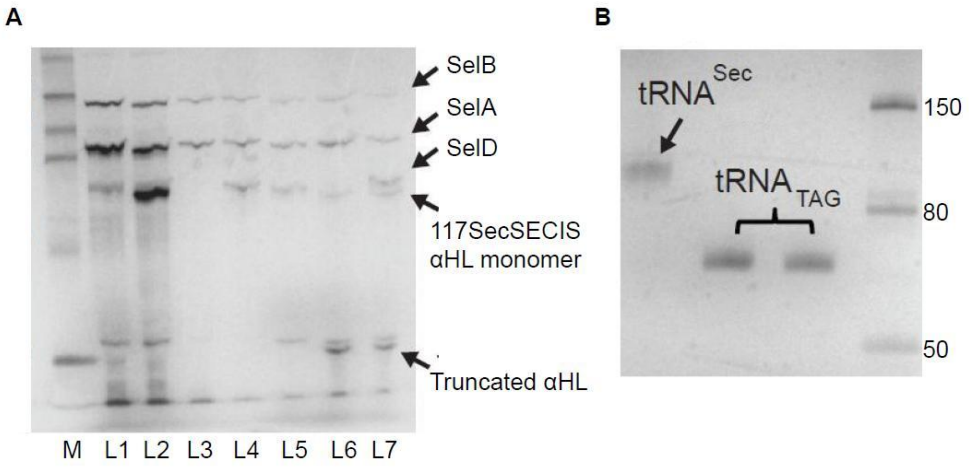


Figure 2.10: Before (RHS) and after (LHS) addition of PEG-OPSS 5kDa. By comparing the gel shifts it can be seen that only the 119CysSECIS monomer reacted. Both WT and 119SecSECIS monomers remained unreacted.

truncated protein indicated that the TGA codon, which also functions as a stop codon, was not being recognised as encoding selenocysteine during the translation of the 113SecSECIS, 117SecSECIS and 119SecSECIS monomers and so protein translation was terminated on reaching the TGA codon. On lowering the IVTT incubation temperature, to 20 ^{o}C, αHL monomer was generated from incubation of the IVTT components with the pT7119SecSECISD8 construct (see Figure 2.9). The αHL monomer, translated at 20^{o}C was incubated with a thiol modifying reagent 5 k PEG-OPSS (PEG-OPSS, Pegworks), however no reaction was observed from gel shift analysis (see Figure 2.10). IVTT generated 119CysSECIS monomer can be seen to have reacted with the PEG-OPSS reagent (approximately 90 %) producing the observed gel shift. An unexpected result was the observation that the αHL monomer was translated from the pT7117TAG plasmid at temperatures 20^{o}C and 25^{o}C, meaning that at these temperatures, protein translation was more favoured than termination. From these observations, it was concluded that the αHL monomer was observed as a result of non-specific amino acid incorporation at the 119TGA (pT7119SecSECISD8 construct) and the 117TAG (pT7117TAGD8 construct) respectively. No αHL monomer was observed from translation of any of the plasmids when the IVTT reaction was incubated at 30^{o}C. As non-specific amino acid incorporation was not desirable, 30^{o}C was subsequently used for all further IVTT experiments (see Discussion for an explanation). The lack of αHL monomer expressed at 30^{o}C was proposed to be due to the TGA codon being interpreted as a stop codon by the translational components of the IVTT system. The proteins required for selenocysteine incorporation were cloned from XL10 Gold genomic DNA (isolated via standard procedure) and placed, each (apart from the SelAB genes which overlap by 3 bases and so were kept together) into either the pT7 vector or the pUC57 vector (SelC, see Section 2.2.3). An overview of each construct is shown in Figure A.5. Briefly, the selAB gene, encoding the selenocysteine synthase and selB protein factor (see Introduction Chapter, section 1.2.2) was cloned into the pT7 vector. The pT7 vector was doubly digested with NdeI and HindIII followed by gel purification of the resulting linear vector. The SelAB containing PCR product was designed so as to have complementary NdeI and HindIII

sites. This PCR product (containing the SelAB gene) was similarly digested with NdeI/HindIII and gel purified. The plasmid containing the SelAB genes will be referred to as pT7SelAB. The SelD gene encoding selenophosphate synthase (see Introduction Chapter, section 1.2.2 was cloned into a separate NdeI/HindIII digested pT7 vector and purified and ligated together in an analogous way to the SelAB construct. The plasmid containing the SelD gene was named pT7SelD. The SelC gene encoding the $tRNA^{sec}$ (see Introduction Chapter 1.2.2) was designed so as to place a SapI restriction site at the 3' terminus of the RNA. The pUC vector and SelC PCR product were doubly digested with StuI and HindIII. After gel purification the SelC PCR product was ligated into the pUC vector. The SelC PCR product was also digested using SapI and used in this form for *in vitro* transcription of $tRNA^{sec}$ followed by purification (see Figure 2.11). The $tRNA^{sec}$ was incubated for 30 minutes at 37^oC with IVTT kit premix (containing the seryl-tRNA synthase) and serine (present in the amino acid mix) so as to create the $Ser\text{-}tRNA^{sec}$. The pT7SelAB and pT7SelD plasmids were translated in separate IVTT reactions, both incubated at 37^oC for 30 minutes. The SelA gene product is a PLP (pyridoxal-5-phosphate) containing enzyme, and so PLP was also added to this reaction mix. After 30 minutes the $Ser\text{-}tRNA^{sec}$-containing reaction was added to the IVTT reaction containing the pTSelAB and pT7SelD plasmids, along with selenite, so as to form GTP·SelB·$Sec\text{-}tRNA^{sec}$ (see Introduction Chapter section 1.2.2). An aliquot of this reaction was taken and added to an IVTT system containing the pT7117SecSECISD8 plasmid and incubated at 30 ^oC for 3 hours. Figure 2.11 shows the gel analysis of this experiment. The appearance of a band below that corresponding to the SelD gene product, was attributed to translation of the αHL monomer containing a selenocysteine residue at position 117. To verify whether this protein band contained selenocysteine, the protein was reacted with 5K PEG-OPSS, however these reactions were not conclusive and no clear band shift was observed.



L1 - pT7SelAB
L2 - pT7SelAB and SelD
L3 - diluted L1
L4 - diluted L2
L5 - pT7SelAB, pT7SelD and tRNA
L6 - pT7SelAB, $tRNA^{sec}$ and pT7117SecSECISD8H6
L7 - pT7SelAB, pT7SelD, $tRNA^{sec}$ and pT7117SecSECISD8H6

Figure 2.11: **A** Expression of the 117SecSECIS αHL by addition of the SelAB, SelD and SelC gene products. Lanes L6 and L7 show a new band attributed to the translation of full length αHL monomer. **B** A comparison of the transcribed and purified $tRNA^{sec}$ with similarly prepared $tRNA_{TAG}$. The RHS gel illustrates the SelB and SelA gene products after translation using the IVTT system.

2.3 Discussion

2.3.1 SECIS design

The SECIS element is a hairpin structure that, forms in the mRNA, 11 bases upstream of the UGA codon and is one of the essential requirements for selenocysteine incorporation. The next paragraph describes the rationale behind the chosen SECIS element engineered into the αHL gene. For information regarding the selenocysteine bacterial incorporation pathway and details regarding each of the required incorporation machinery see the Introduction chapter, section 1.2.2. The bacterial SECIS element contains five essential bases for interaction and recognition with the SelB

transcription factor (see Results Figure 2.2). Kromayer et al. (1996) have shown that selenocysteine incorporation cannot occur if any of these essential bases are missing. Similarly, studies have been carried out to investigate other important features of the SECIS hairpin loop. Liu et al. (1998) altered the number and location of the base pairing interactions along the hairpin and monitored the level of expression produced for each construct. It was noted that some base pairing positions were not only essential for selenocysteine incorporation, but even switching the bases from a purine to a pyrimidine could cause a significant decrease in protein expression. By carrying out mutagenesis studies on the fdhH (selenocysteine containing formate dehydrogenase H) gene SECIS element, Liu et al. (1998) devised a minimal SECIS structure that still enabled incorporation of selenocysteine into the *Escherichia coli* formate dehydrogenase but was simplified in the number of conserved bases and had a shortened stem. It was this minimal structure on which the SECIS element placed into the αHL gene, was based. Sandman et al. (2003) showed that the codon directly after the UGA codon was important in determining the specificity of selenocysteine insertion in response to the UGA codon. A cytosine or uricil base directly after the UGA, ie UGAC or UGAU resulted in only selenocysteine incorporation, where as a purine (for example adenine or guanine) immediately downstream of the UGA codon caused selenium-independent tryptophan insertion (Sandman & Noren, 2000). Aldag et al. (2009) incorporated selenocysteine into an endogenously selenocysteine-free P450 enzyme, in place of cysteine, by placing a minimal SECIS element into the reading frame of the P450 gene. Two approaches were considered viable for the incorporation of the minimal SECIS element into the αHL gene. As was carried out for the P450 enzyme by Aldag et al. (2009), a region of the base sequence 11 nucleotides downstream from the selenocysteine codon was mutated to those of the minimal SECIS element as reported by Liu et al. (1998). A second approach involved searching the αHL gene for regions of base sequence that matched those required for the SECIS element. The rational for this being to minimise the number of non-silent base changes due to the specific SECIS sequence. Two regions were located in which the base sequence partially matched (ie 6 bases out of 21 were the same). Unfortunately these

regions in the αHL gene would have placed the selenocysteine residue outside of the β barrel and so the design of these constructs was not pursued.

2.3.2 Selenocysteine containing αHL expression in *Escherichia coli*

The pT7117SecSECISD8H6 plasmid was used to express the selenocysteine containing αHL monomer recombinantly in *Escherichia coli* BL21(DE3) cells. As the TGA codon is degenerate and also codes for translation termination, there was competition between translation and termination of the polypeptide due to recognition of the TGA codon by release factor 2 (RF2) (Mansell et al., 2001). The "full length" αHL monomer refers to translational termination at the TAA stop codon after the D8H6 tag. Due to this competition between RF2 and the selenocysteine incorporation machinery, a large proportion (approximately 95 %) of the protein translated was truncated as a result of termination at the TGA selenocysteine (see Results, Figure 2.3). Previous experiments carried out by Tormay et al. (1996) have highlighted the importance of having not only elevated levels of the selenocysteine incorporation machinery proteins for selenocysteine de-coding, but also the balance of each component during protein expression (Rengby et al., 2004; Gursinsky et al., 2008). Rengby et al. (2004) noted that selenoprotein expression yields were higher when cultures were induced in the late exponential phase (attributed to lower RF2 levels), which was in agreement with previous findings by Mansell et al. (2001), so for expression of the 117SecSECIS αHL monomer, an OD_{600} of around 1.5 - 2 was reached before addition of IPTG (Isopropyl β-D-1-thiogalactopyranoside). The amount of IPTG was half the usual induction value (see Methods) because previous experiments by Rengby et al. (2004) showed that selenocysteine incorporation is favoured (over translation termination at the UGA codon by RF2) when both the concentration of the GTP·SelB·Sec-$tRNA^{sec}$ and mRNA are high. To provide the cells with a plentiful supply of selenium, cultures were supplemented with sodium selenite. Cysteine supplementation was also carried out so that the cysteine dependent cellular pathways would not utilise selenocysteine in-

stead of cysteine. The final concentrations of these reagents were not optimised for the expression of the selenocysteine containing αHL, but were taken from previous experiments by Cheng et al. (2006). Arnér et al. (1999) reported a significant increase in selenoprotein yield when expression was carried out in the presence of a second plasmid (pSUABC) containing the SelAB and SelC genes under the control of their endogenous promoters (Arnér et al., 1999). Arnér et al. (1999) co-transformed the pSUABC plasmid with another plasmid containing the selenoprotein gene, which increased the levels of SelAB and SelC gene products present within the bacterial cells allowing the selenocysteine incorporation pathway to compete more effectively with RF2 protein synthesis termination, resulting in higher selenoprotein expression yields. Attempts were made to doubly transform both the pT7117SecSECISD8H6 plasmid and pSUABC plasmid, but colonies of doubly transformed BL21(DE3) cells were never obtained. Sequential transformation of the pSUABC and pT7117SecSECISD8H6 plasmids was also carried out. Competent cells, transformed with either pSUABC or pT7117SecSECISD8H6 were made and then the second transformation performed with these cells. Doubly transformed BL21(DE3) colonies were not obtained via this method either. The double transformation was not pursued further, however other approaches would be, for example, to alter the cell strain or place all the genes onto a single plasmid. The selenocysteine-containing αHL monomer was produced without the pSUABC plasmid, at an expression level of approximately 5% that of WT αHL, so approximately 5-10 mg of 117SecSECIS αHL monomer was obtained per 1 L culture. Both the WT and 117SecSECIS was expressed in inclusion bodies and so was purified under denaturing conditions. It can be seen in Figure 2.3 that the proportion of αHL 117SecSECIS monomer compared to truncated protein was never greater than 5-10%.

2.3.3 Characterisation of the 117SecSECIS αHL monomer

Characterisation of the pT7117SecSECISD8H6 gene product was carried out via three methods. Initially, a gel-shift assay was carried out, where the selenocysteine residue was reacted with IASD

and MISD (see Figure 2.7). This assay showed how the 117SecSECIS monomer reacted with both these thiol modifying reagents and produced an observable gel shift. So as to verify the identity of the residue incorporated at the TGA codon, LC-MS analysis was carried out and confirmed the presence of the selenocysteine residue. These experiments were carried out in a similar way to that of House et al. (1992) where excess DTT was used to ensure that the selenocysteine residue was maintained in its reduced, selenol form.

2.3.4 *In vitro* transcription translation of the 117SecSECIS αHL monomer

From 12 % polyacrylamide gel analysis, it was not clear that the $tRNA^{sec}$ was present in the IVTT premix, supplied with the kit (Promega). By cloning the SelC gene from genomic DNA, the $tRNA^{sec}$ was transcribed, purified and added as a separate component to the IVTT reaction. Previous experiments investigating the requirements for selenocysteine synthesis on the $tRNA^{sec}$ have been carried out by Baron et al. (1990) and similar incubation conditions were employed to incubate the transcribed $tRNA^{sec}$ with premix (containing the seryl synthetase), serine, ATP, PLP, selenocysteine synthetase and sodium selenite. The idea was to form a surplus of selenocysteine aminoacylated $tRNA^{sec}$ to allow full length protein expression to compete with translational termination. Tormay et al. (1996) reported that the SelB gene product was able to distinguish between serylated and selenocysteine aminoacylated $tRNA^{sec}$ and so the formation of $tRNA^{sec}$ prior to addition to the main IVTT reaction was thought to enhance the concentration of the GTP:SelB:Sec-$tRNA^{sec}$ tertiary complex in the IVTT reaction mix. During IVTT expression of 117SecSECIS, the SelAB gene was expressed prior to addition of the pT7117SecSECISD8H6 plasmid DNA. Incubation of the pT7SelAB was carried out firstly, so as to generate the SelB protein so that on addition of the Sec-$tRNA^{sec}$ species, a ternary complex between the SelB:GTP:Sec-$tRNA^{sec}$ would be formed. For expression of pT7WTD8H6, 1.6 μg of plasmid DNA (per 25 μl IVTT reaction) is usually used, but addition of 40ng of pT7117SecSECISD8H6 was added per 25 μl

IVTT reaction volume, so as to control the amount of mRNA transcribed, relative to the amount of GTP:SelB:Sec-tRNA^{sec} complex. The IVTT reaction was incubated at 30 ^{o}C instead of 37^{o}C. This was also to aid full length protein synthesis over TGA truncated. 30^{o}C was chosen because temperatures lower than 30^{o}C generated non-specific read-through (see Results 2.2.3) but selenocysteine protein yields were reported as lower when expression was carried out at 37^{o}C (Rengby et al., 2004). Incubation of the pT7117SecSECISD8H6 IVTT reaction at 30^{o}C for 3 hours generated the appearance of a new band of approximately 35 kDa which was attributed to synthesis of the 117SecSECIS αHL monomer. Future experiments to verify the identity of this protein band would include trypsin digest followed by mass spectroscopy analysis and a gel-shift experiment to react the selenocysteine residue with MISD or IASD.

2.3.5 Oxidation of selenocysteine during heteroheptamer preparation

Following oligomerisation of the monomeric αHL subunits in a ratio of 1:6 (117SecSECIS$_1$WT$_6$) it was possible to isolate the heteroheptamer containing a single selenocysteine at position 117 in the β barrel (see Section 2.1.2). Various optimisation steps were carried out including increasing the amount of DTT present during oligomerisation from 5-10 mM, using an alternative reducing agent such as TCEP during oligomerisation and/or in the sample buffer, capping the selenol by formation of a selenylsulfide bond, increasing the amount of reducing agent in the sample buffer during gel purification, running the gel at 4 ^{o}C, negation of gel purification by utilising Ni-magnetic beads (to remove WT, non-H6-tagged heptamer) and finally, use of oligomerised heptamer without any purification. None of the above methods were found to increase the proportion of "active" (ie non-oxidised selenocysteine containing heptamers) to inactive pores present in the protein batches. The inactivity of the 117SecSECIS$_1$WT$_6$ was thought to arise primarily due to the gel purification step and is still currently an unresolved challenge.

2.4 Conclusions

The motivation for placing a selenocysteine residue into the β barrel of αHL was to compare the reactivity of the selenolate and thiolate functional groups. Selenium is an essential element for eukaryotic, prokaryotic and archea and it is most commonly used in proteins in the form of selenocysteine. A variety of different types of enzymes require selenocysteine for catalytic function, many of which are redox enzymes involved in the reduction of reactive organic species (for example Glutathione peroxidase) or disulfide bonds (for example Thioredoxin reductase). The results of this chapter describe the incorporation of selenocysteine in the β barrel of αHL by utilising the endogenous machinery of *Escherichia coli*. An *in vitro* method for selenocysteine containing αHL expression was also investigated. Results suggested that expression of full length selenocysteine containing αHL monomer was possible using the IVTT system, on addition of components usually present during selenoprotein bacterial expression *in vivo*.

2.5 Materials and Methods

2.5.1 117CysSECISpT7 plasmid construction

Double restriction digest of the pT7 vector

The WT αHL-containing plasmid (22.5 μl, 442 ng/μl) with an 8 aspartate tail (pT7WTD8) was added to an eppendorf containing 10X NEB buffer (4 μl, New England Biolabs), nuclease free water (14.5 μl, Ambion), SacII (3 μl, New England Biolabs, 20,000 U/mL) and StuI (6 μl, New England Biolabs, 10,000 U/mL). The 50 μl reaction was mixed and left to incubate at 37 ^{o}C overnight. The linearised plasmid was gel purified (0.5 % agarose gel, 100V, 45 minutes) using the QIAquick gel extraction kit (Qiagen). The concentration of DNA recovered was approximately

39 ng/μl and in total, 3.8 μg was obtained.

Phosphorylation

Primers 2 and 3 (1 μl, 10 μM final concentration) were 5' phosphorylated using T4 polynucleotide kinase (1 μl, 10,000 U/ml, New England Biolabs) and buffered with the 10X NEB T4 DNA ligase buffer (1 μl, New England Biolabs), and nuclease-free water (6 μl, Ambion). This reaction was incubated at 37 ^{o}C for 1 hour. The kinase was heat inactivated for 20 minutes at 70 ^{o}C.

3-way ligation to construct the pT7117CysSECISD8 plamsid

Primers 1 and 4 (1 μl, 10 μM of each) were added to an eppendorf containing the doubly digested pT7WTD8, linearised plasmid (1 μl, 39 ng/μl), 10X T4 DNA ligase buffer (1 μl, New England Biolabs) and nuclease-free water (4 μl). After addition of the phosphyorylation reaction containing primers 2 and 3 (1 μl, 10 μl/M), the T4 DNA ligase enzyme was added (1 μl, 400,000 cohesive end units/ml, New England Biolabs). Note that all primers were added to a final concentration of 1 μM. The reaction was mixed gently and incubated at rt for 1 hour. Heat inactivation was not necessary after this reaction.

Transformation of the ligation reaction into XL10 Gold competent cells

Competent cells (95 μl, XL10 Gold) were thawed at 4 ^{o}C. 8 μl of the above ligation reaction mix was added to the competent cells, stirred briefly and then left at 4 ^{o}C for 30-50 minutes. Heat shock was carried out for 35 seconds at 42 ^{o}C followed by streaking onto LB-ampicillin agar plates (100 mg/ml ampicillin, 10 μl per 10mL of LB).

Site directed mutagenesis

The procedure detailed here was carried out using the following template DNA plasmids: pT7113CysSECISD8 pT7117CysSECISD8, pT7119CysSECISD8 and pT7137CysSECISD8. The forward (F) and reverse (R) primer sequences used to carry out site directed mutagenesis were GAGTATTGACTCACGTTAACG (F), CGTTAACGTGAGTCAATACTC (R) (113Cys to 113Sec), CGTTATGACTCGGATTCAAC (F), GTTGAATCCGAGTCATAACG (R) (117Cys to 117Sec), GTTAACGTACTGACTCAACG (F), CGTTGAGTCAGTACGTTAAC (R) (119Cys to 119Sec), CCTTATTTGACTCAATGTTTC (F), GAAACATTGAGTCAAATAAGG (R) (137Cys to 137Sec).

The cysteine codon (ACA) was mutated to that encoding selenocysteine (TGA) using the Quikchange Site-directed mutagenesis kit II XL (Strategene). Briefly, the forward and reverse primers were diluted (TE buffer, 10 mM Tris.HCl, 1 mM EDTA, pH 8) to a concentration of 150 ng/μl. The PCR reaction mixture was prepared by addition of 5 μl reaction buffer (10X, Strategene), 1 μl template DNA (10 ng/μl, e.g. pT7117CysSECISD8), 1 μl forward primer (150 ng/μl), 1 μl reverse primer (150 ng/μl), 1 μl dNTP mix (Strategene), 3 μl Quikchange solution (Strategene), 38 μl H_2O (nuclease free) and 1 μl DNA polymerase (Strategene). This reaction mix was thermally cycled using the following PCR protocol; 95^{o}C, 1 minute followed by 18 cycles of 95^{o}C, 0.5 minutes, 60^{o}C, 0.5 minutes and 68^{o}C, 9 minutes. After these cycles, a temperature of 68^{o}C was held for 7 minutes followed by storage at 4^{o}C. Dpn1 (Strategene, 1 μl) was then added and the PCR reaction mix incubated for 1 hr at 37^{o}C. Transformation was then carried out into XL10 Gold competent cells (see 2.5.1) by incubating (on ice) 5 μl of PCR mix with 95 μl of competent cells. After miniprep (Qiagen) and sequence verified, the newly created plasmids were named pT7113SecSECISD8, pT7117SecSECISD8, pT7119SecSECISD8 and pT7137SecSECISD8.

Homologous recombination

Homologous recombination was carried out to introduce a six amino acid long, Histidine tag onto the C-terminus of the αHL monomer. Two digestion reactions were carried out as illustrated in Figure A.3. The plasmid DNA (1 μg, pT7SecSECISD8) was digested with HindIII (1μl, 20,000 U/ml, New England Biolabs) and the pT7WTD8H6 plasmid digested with NdeI (1 μl, 20,000 U/ml, New England Biolabs). Both digestion reactions were incubated at 37 ^{o}C for 1 hour 30 minutes followed by heat inactivation (65 ^{o}C, 20 minutes). The two separate PCR reactions were carried out using the Phusion Flash enzyme mix (Finzymes) where the sequence of the four primer sets used, are shown in Figure A.6. The ampicillin forward and reverse primers are ATAAAGTTGCAGGACCACTTCTG (F, T_m = 63.6 ^{o}C) and CAGAAGTGGTCCTGCAACTTTAT (F, T_m = 63.6 ^{o}C) Two 50 μl PCR reactions were set up, where 1 μl of each primer (50 pmol/μl) and linearised DNA template (10 ng/μl) were added to a nuclease-free water (22 μl, Ambion) and Phusion flash PCR master mix (Finymes). The PCR protocol was carried out according to the Finzymes information sheet. Analytical 0.5% agarose gels (TAE running buffer, 100 V, 30 minutes) of the PCR reaction products are shown in Figure A.3. Transformation into XL10 Gold competent cells was carried out by addition of 5μl of each PCR reaction (see 2.5.1).

2.5.2 Recombinant Protein Expression

Protein Expression and Purification

The pT7117SecSECISD8H6 and pT7WTD8H6 plasmids were transformed into BL21(DE3) bacterial cells and starter cultures (20 ml Luria Broth, Sigma Aldrich) grown over night (37 ^{o}C, 200 rpm). The expression was carried out using Terrific Broth (300 ml) where 5 ml (per 100 ml culture) of the starter culture was added initially. On reaching an OD_{600} = 0.7 (pT7WTD8H6) and 0.9 - 1.2 (pT7117SecSECISD8H6) were induced with IPTG (final concentration of 1 mM for

pT7WTD8H6 and 0.5 mM for pT7117SecSECISD8H6 respectively) and expression carried out for 19-22 hours at 20 ^{o}C, 200 rpm. Selenite (final concentration 5 nM, Sigma Adrich) and cysteine (final concentration 5 μM, Sigma Aldrich) solutions were added to the pT7117SecSECISD8H6 culture prior to IPTG addition. Cultures were pelleted at 4 ^{o}C, 3800 rpm for 20 minutes and then placed at -80 ^{o}C so as to freeze. The cell pellets were thawed and resuspended (3.8 ml pT7117SecSECISD8H6 and 4 ml/100 ml pT7WTD8H6 original culture volume) in lysis buffer (50 mM Tris-HCl, 300 mM NaCl, 10 mM Imidazole, pH 8.3). TCEP (5 mM, Fluka) was added to the lysis buffer of the pT7117SecSECISD8H6 pellet and in all subsequent buffers. Lysozyme was added (560 μl, 10 mg/ml) and Benzonase (6 μl, 250 U/μl) and the cultures left on ice for 30 minutes. Sonication was carried out (30 s, 10 kHz then 30 s off) and this sonication procedure repeated five times per lysate. The solutions were then centrifuged (4 ^{o}C, 25000 rcf, 20 minutes) and the supernatant discarded. The pellet was redissolved in a binding buffer (50 mM Tris-HCl, 8 M urea, 500 mM NaCl, 30 mM Imidazole, pH 8.5). The Ni-NTA resin (2 mL of resin, Qiagen) was equilibrated with the binding buffer (10 bed volumes). The sample was then loaded onto the column, sealing and left to rotate at 4 ^{o}C for 2 hours. The Ni-NTA column was then washed with 10 bed volumes of the binding buffer before eluting with the elution buffer (3:1, binding buffer: 50 mM Tris-HCl, 1 M Imidazole, pH 8.5). Approximately 8 bed volumes of elution buffer was used. The above purification method for 117SecSECIS αHL monomer was also used for the 117CysSECIS αHL monomer also.

117SecSECISD8H6, WTD8H6 and 117CysSECISD8H6 dialysis, de-salting and concentration

Ni-NTA purified αHL monomeric protein (WT, 117CysSECIS, 117SecSECIS) was buffer exchanged from the elution buffer (see above) into a lower salt buffer (10 mM Tris-HCl, 150 mM NaCl, 5 mM TCEP, pH 8.5). Float-Alyzers G2 (8-10 kDa cut off, Spectrum labs) were used

and the protocol was as documented. Briefly, the Dialysis membranes were washed with double deionised water for 40 minutes at room temperature. Prior to this, the protein samples were slowly diluted at 4 ^{o}C with a non-denaturing version of the elution buffer (3:1, 50 mM Tris-HCl, 8 M urea, 500 mM NaCl, 30 mM Imidazole, pH 8.5 : 50 mM Tris-HCl, 1 M Imidazole, pH 8.5) so that the urea concentration was 4 M. The protein samples were added into the dialysis bag and stirred at 4 ^{o}C for 3 hours. The external buffer was then renewed and the dialysis allowed to continue for a further 3 hours. A final external buffer renewal was carried out and left stirring overnight (always at 4 ^{o}C). The WT and 117CysSECIS monomeric αHL protein samples were lypholysed and stored at -80 ^{o}C. The 117SecSECIS monomer precipitated during dialysis and was centrifuged (4 ^{o}C, 12000 rcf, 10 minutes) and washed with the lower salt buffer. This process was repeated 3 times after which the protein was stored at -80 ^{o}C. The monomeric protein was dissolved (Buffer: 3 M GnCl, 20 mM NaP_i, 50 mM NaCl, 10 mM DTT, pH 7.6) to a final concentration of WT; 1 mg/ml, 117CysSECIS; 2 mg/ml and 117SecSECIS; 2 mg/ml. These concentrations were determined spectroscopically at 280 nm (Nanodrop) and 35 μl of each protein sample was used. An alternative procedure was also carried out where by after dialysis, samples were concentrated and buffer exchanged using the vivaspin columns (MWCO 10 kDa, Fisher). The buffer exchange was carried out into TE buffer (10 mM Tris.HCl, 100 μM EDTA, pH 8.0) containing 20 mM DTT. The LC-MS spectra can be seen in Figure 2.5.

Sample preparation for LC-MS analysis

The monomeric protein samples were concentrated to between 2-5 mg/ml (see above). The LC-MS set up involved the de-salting of the protein sample by passing through a C4 reverse phase column (Jupiter 5 μm C4 300 Å New Column 250 x 4.6 mm, Phenomenex) followed by entering the electrospray ionisation time of flight (TOF) mass spectrum machine (Waters LCT). The HPLC protocol employed a gradient of 5-95% aqueous solution over 60 minutes (water (acetonitrile, 5%

and 0.1 % formic acid): acetonitrile (water, 5% and 0.1 % formic acid)). A flow rate of 0.3 ml/min was employed and the sample analysed at 210 nm. The main mass spectrum parameters were typically as follows: Polarity ES+, Analyser V mode, Capillary (V) 3000.0, Sample Cone (V) 45.0, Desolvation Temp (^{o}C) 250.0, Source Temp (^{o}C) 120.0, Cone Gas Flow 18.0, Desolvation Gas Flow 587.0, Syringe Type SGE 250 μl, Ion Guide One 18.6 and Acceleration (V) 200.0.

Oligomerisation and isolation of a heteroheptamer containing a single selenocysteine residue

The general procedure by which the monomeric αHL protein was oligomerised has been described. Typically, 5 μg of monomeric αHL protein is generated per 100 μl IVTT reaction. Based on this estimation, ratios of 1:6, 2:5 and 3:4 of 117SecSECIS monomer: WT (IVTT generated) monomer were mixed together and MBSA buffer (150 mM NaCl, 10 mM MOPS, 0.1 % Bovine Serium Albumin (1 mg/ml)), added so as to dilute the 8 M urea present in the 117SecSECIS monomer solution to 0.1 M or less. Sometimes *E. Coli* expressed WT was added so as to keep the concentration of monomer high enough to oligomerise after addition of the MBSA buffer. An alternative method that was also employed, was the partial dilution of the 8 M urea using the viva spin columns (MWCO 10 kDa, Fisher). DTT (Sigma Aldrich) was added to a final concentration of 5 mM. Rabbit red blood cell membrane (2μl per 100 μl IVTT reaction, of approx. 1 mg/ml, see 2.5.4 for the preparation procedure) was added, mixed gently and left to incubate at 37 ^{o}C for 1 hour. After pelleting (centrifugation, 25,000 rcf, 10 minutes, 4^{o}C), the pellet was resuspended in laemmeli sample buffer (30 μl, Biorad) containing between 5-30% β-mercaptoethanol (Biorad). Samples were loaded onto an SDS-PAGE 5% gel and run overnight (roughly 12-16 hours) at an applied voltage of between 45-60 V (depending on the length of the gel setup used). The gel was then dried under vacuum (approximately 6 hours). As the WT IVTT generated monomeric protein was radiolabeled (see 2.1.3) the dried gel stapled to the film and exposed overnight in a cassette. The film was developed and the gel band of radioactive hetroheptamer containing a single seleno-

cysteine residue i.e. ($117SecSECIS_1WT_6$) was cut and rehydrated in TE buffer (300 μl, 10 mM Tris.HCl, 100 μM EDTA, pH 7.4) containing 7 mM DTT for 30 minutes on ice. A pestle was used to crush the gel into small pieces, followed by SDS removal and shredding by using a Qiashredder column (Qiagen). The $117SecSECIS_1WT_6$ sample was then aliquoted (10 μl aliquots) and stored at -80 ^oC ready for use in the single channel electrical recording experiments.

Gel-shift assay

Listed below are the protocols used with the WTD8H6, 117CysSECISD8H6 and 117SecSECISD8H6 monomeric αHL protein samples prior to loading (approximately 20μl of sample per lane with 40 μl laemmeli sample buffer) onto a 12.5 % Tri-HCl SDS PAGE gel (see Figure 2.7). Monomeric αHL protein (30μl, 0.2-0.5 mg/ml), 50 mM NaP_i buffer (50 μl, 300 mM NaCl, pH 8) and IASD solution (20 μl, 4-acetamido-4'-((iodoacetyl)amino)stilbene-2,2'-disulfonate, 100 mM in nuclease-free water) were mixed in an eppendorf. The reaction was wrapped in foil and left for 2.5 hours at rt before being quenched with DTT (10 μl, 1 M solution in nuclease-free water) and placed on ice. Protein (20 μl), 50 mM NaP_i buffer (60 μl, 300 mM NaCl, pH 8), MISD (20 μl, 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid, 100 mM solution in nuclease-free water) was mixed in an eppendorf. The reaction was allowed to proceed for 10 minutes at 4 ^oC before being quenched with DTT (10 μl, 1 M solution in nuclease-free water).

Hemolytic assay

MBSA buffer (900 μl, 10 mM MOPS, 150 mM NaCl, 0.1 % BSA, pH 7.4) was pipetted into an eppendorf containing rabbit red blood cells (450 μl, see 2.5.4). This was centrifuged (2 minutes, 12000 rpm) and the supernatant discarded. The above procedure was repeated until the supernatant was colourless (usually 3-4 washes). The red blood cell pellet was gently resuspended with MBSA buffer (1000 μl) and transferred to a 15 ml falcon tube where it was further diluted with

MBSA buffer to a final volume of 10 ml. The hemolytic assay was carried out by monitoring the absorbance at a wavelength of 595 nm. A serial dilution method was used across the 96-well plate, where the total volume in each well of the plate reader was 100 μl. Briefly, MBSA buffer (95 μl in the first column and 50 μl subsequently with the last column left blank) was pipetted into all wells using a multichannel pipette. Monomeric αHL protein sample was added (5 μl, 50 ng/μl from IVTT) to the first column. Using the multichannel pipette, the well contents were mixed seven times before removing 50 μl and placing it into the next column of wells. This process was repeated across the 96-well plate until all wells contained 50 μl of MBSA buffer resulting in a 2X serial dilution of the protein. Starting from the least concentrated protein column, the 1 % red blood cell solution was added (50 μl) without mixing. A total of 200 absorbance measurements (595 nm) were taken with 30 seconds between each measurement (apart from 15 s between the first and second readings). Mixing was carried out for 1 second before the readings (2 s between first and second).

2.5.3 *In vitro* transcription translation Protein Expression

Cloning of the Sel A-D genes

The SelA-D genes, responsible for the incorporation of selenocysteine were cloned into the pT7 vector in the following way. The general method firstly involved PCR amplification of the gene where restriction sites were placed at each end of the primers. These restriction sites were chosen so as to be complementary to those present on the pT7 vector. Once gel purified, the PCR product was digested, gel purified for a second time and then ligated into the pT7 vector using T4 DNA ligase.

Cloning the SelD gene into the pT7 vector

XL10 Gold genomic DNA (1 μl, 100 ng/μl), SelDF primer (1 μl, 50 pmol/μl, ATGAGCGAGAACTCGATTCGTTTGACCCAATA, Tm = 76.4 ^{o}C), SelDR primer (1 μl, 50 pmol/μl, TTAACGAATCTCAACCATGGCACGACCGC, Tm = 79.0 ^{o}C), nuclease-free water (22 μl, Ambion) and Phusion flash PCR master mix (25 μl, Finnzymes) were mixed in a PCR eppendorf. The protocol for PCR was as follows. Initial (98 ^{o}C 10 s) heat activation, denaturation (98 ^{o}C 1 s), annealing gradient (60 - 70^{o}C, 5s), extension (72 ^{o}C, 19 s, SelD gene: 1.044 kb, Enzyme rate: 15 s/kb). These three steps were cycled 40 times. Lastly, the final extension (72 ^{o}C, 60 s) was carried out and then the PCR reactions stored at 4 ^{o}C. The PCR product was gel purified (2 % Agarose, 100 V, approx. 45 minutes) using the Qiaquick Gel Extraction Kit (Qiagen) with a DNA recovery of 5 μg (100 ng/μl). The PCR product (10 μl) was then doubly digested by addition of NdeI (3 μl, 20,000 U/ml, New England Biolabs) and HindIII (3 μl, 20,000 U/ml, New England Biolabs) to an eppendorf containing NEB buffer 2 (4 μl, New England Biolabs) and nuclease-free water (30 μl, Ambion). The reaction was incubated at 37 ^{o}C for 1 hour followed by gel purification (2 % Agarose, 100 V, approx. 45 minutes) using the Qiaquick Gel Extraction Kit (Qiagen). The ligation reaction was set up in the following way. The digested and gel purified, PCR product (4 μl), linearised pT7 vector (1 μl), nuclease-free water (3μl, New England Biolabs), ligase buffer (1 μl, Promega) and T4 DNA ligase (1 μl, Promega) were mixed in an eppendorf and incubated for 1 hour, rt. 8 μl of the crude reaction mix was then used to transform into XL10 Gold competent cells (see Section 2.5.4 for the transformation procedure).

Cloning the SelAB genes into the pT7 vector

Genomic DNA (1 μl, 100 ng/μl), SelABF primer (1 μl, 50 pmol/μl, AATTATCATATGATGACAACCGAAACGCGTTCC, Tm = 75.1 ^{o}C), SelABR primer (1 μl, 50 pmol/μl, TTTTAAAAGCTTTTATTTTTCCGGAAATAATAATGCGTCGC, Tm = 75.3 ^{o}C), nuclease-free water (22 μl, Am-

bion) and Phusion flash PCR master mix (25 μl, Finnzymes) were mixed in a PCR eppendorf. The protocol for PCR was as follows. Initial (98 ^{o}C 10 s) heat activation, denaturation (98 ^{o}C 1 s), annealing (68 ^{o}C, 5 s), extension (72 ^{o}C, 25 s, SelAB gene: 3.2 kb, Enzyme rate: 15 s/kb). These three steps were cycled 40 times. Lastly, the final extension (72 ^{o}C, 60 s) was carried out and then the PCR reactions stored at 4 ^{o}C. The PCR product was gel purified (2 % Agarose, 100V, 45 minutes) using the Qiaquick Gel Extraction Kit (Qiagen) with a DNA recovery of approximately 1.15 μg (41 ng/μl). The PCR product (28 μl, 1 μg) was then doubly digested by addition of NdeI (3 μl, 20,000 U/ml, New England Biolabs) and HindIII (3 μl, 20,000 U/ml, New England Biolabs) to an eppendorf containing NEB buffer 2 (4 μl, New England Biolabs) and nuclease-free water (12 μl, Ambion). The reaction was incubated at 37 ^{o}C for 1 hour followed by gel purification (2 % Agarose, 100 V, 45 minutes) using the Qiaquick Gel Extraction Kit (Qiagen).

Cloning SelC gene into the pUC57 vector

Genomic DNA (1 μl, 100 ng/μl), SelCF primer (1 μl, 50 pmol/μl, CCTTAATACGACTCACTATAGGAAGATCGTCGTCTCCGGT, Tm = 77.0 ^{o}C), SelCR primer (1 μl, 50 pmol/μl, TAAGTTAAGCTTGCTCTTCAAGTTGGCGGAAGATCACAGGAGTC, Tm = 80.6 ^{o}C) nuclease-free water (22 μl, Ambion) and Phusion flash PCR master mix (25 μl, Finnzymes) were mixed in a PCR eppendorf. A 2-step PCR reaction was carried out as follows. Initial (98 ^{o}C 10 s) heat activation, denaturation (98 ^{o}C 1s), extension (72 ^{o}C, 7 s, SelC gene: 95 b, Enzyme rate: 15 s/kb). These three steps were cycled 40 times. Lastly, the final extension (72^{o}C, 7s) was carried out and then the PCR reactions stored at 4 ^{o}C.The PCR product was gel purified (2 % Agarose, 100 V, approx. 45 minutes) using the Qiaquick Gel Extraction Kit (Qiagen) with a DNA recovery of approximately 1.04 μg (37 ng/μl). The PCR product (28 μl) was then doubly digested by addition of StuI (6 μl, 10,000 U/ml, New England Biolabs) and HindIII (3 μl, 20,000 U/ml, New England Biolabs) to an eppendorf containing NEB buffer 2 (4 μl, New England Biolabs) and nuclease-free

water (9 μl, Ambion). The reaction was incubated at 37 ^{o}C for 1 hour followed by gel purification (2 % Agarose, 100 V, 45 minutes) using the Qiaquick Gel Extraction Kit (Qiagen).

Ligation reactions for each of the SelA-D genes

The ligation reaction to place SelAB and SelD digested and gel purified PCR products into the pT7 vector was set up in the following way. The PCR product (4 μl), linearised pT7 vector (1 μl, 111 ng/μl), nuclease-free water (3 μl, New England Biolabs), ligase buffer (1 μl, New England Biolabs) and T4 DNA ligase (1 μl, Promega) were mixed in an eppendorf and incubated for 1 hour, rt. 8 μl of the crude reaction mix was then used to transform into XL10 Gold competent cells (see Section 2.5.1 for the transformation procedure). The procedure for ligating the SelC PCR product was identical apart from this gene was placed into the pUC57 linearised vector (1 μl, 21.2 ng/μl).

Linearisation of the pT7 vector

pT7 template DNA (22.1 μl, 442 ng/μl, pT7WTD8 DNA template), NEB buffer 2 (4 μl, New England Biolabs), HindIII (3 μl, 20,000 U/ml, New England Biolabs), NdeI (3 μl, 20,000 U/ml, New England Biolabs) and nuclease-free water (17.5 μl, Ambion) were mixed and incubated at 37 ^{o}C overnight. Gel purification (Qiaquick Gel Extraction Kit, Qiagen) yielded 3.1 μg of linearised vector (28 μl, 111 ng/μl).

Linearisation of the pUC57 vector

pUC57 template DNA (10 μl, 1 μg/μl), NEB buffer 2 (4 μl, New England Biolabs), HindIII (3 μl, 20,000 U/ml, New England Biolabs), StuI (6 μl, 10,000 U/ml, New England Biolabs) and nuclease-free water (27 μl, Ambion) were mixed and incubated at 37 ^{o}C overnight. Gel purification (Qiaquick Gel Extraction Kit, Qiagen) yielded the linearised vector (12 μl, 21.2 ng/μl).

tRNA^{sec} transcription

Firstly the pUC57SelC (3 μl, 300 ng/μl)was digested with SapI (2 μl, 2,000 U/ml New England Biolabs) by incubating at 37 ^{o}C in NEB buffer 4 (1 μl, New England Biolabs), and nuclease-free water (5 μl, Ambion) for 2.5 hours (or overnight) followed by heat inactivation (65 ^{o}C, 25 minutes). The transcription reaction was carried out using the Ribomax^{TM} Express T7 RNA polymerase kit (Promega). Briefly, Ribomax^{TM} Express T7 2X buffer (10 μl), linearised pUC57SelC DNA template (8 μl, approx. 0.8 μg), nuclease-free water (2 μl, Ambion) and T7 Express RNA polymerase Enzyme Mix (2 μl, Promega). After incubation (1 hr, 37 ^{o}C) DNase (2 μl, 1 U/μg, Promega) was added to the reaction and left for a further 15 minutes. Purification of the 95 base tRNA^{sec} was carried out in accordance to the RNeasy (Promega) protocol for short RNA species with a final recovery of 24.7 μg (30 μl, 824.5 ng/μl) which was stored at -80 ^{o}C.

117SecSECIS αHL monomer generated using *In-vitro* translation transcription (IVTT)

The general protocol for the IVTT reactions carried out to generate the 117SecSECIS monomer are detailed. Any variations to this protocol are documented in the Discussion section. pT7SelAB (0.5 μl, 100 ng/μl) and pT7SelD (0.5 μl, 200 ng/μl) were added along with the usual IVTT components (3 μl Premix, 1 μl amino acid mix, 0.2 μl $S^{3}5$ labelled Methionine, 2.4 μl T7 S30 extract). An additional component was added; Pyridoxyl-5-phosphate (PLP, 0.4 μl, 10 μM final concentration). This reaction mix (reaction A) was incubated at 37 ^{o}C for 30 minutes. A second IVTT reaction was carried out by taking 2 μl of the above crude reaction mix (reaction A) and adding it to a PCR sized eppendorf containing tRNA^{sec} (4 μl, 207 ng/μl), premix (3 μl) and sodium selenite (1 μl, 20 μM final concentration). This 10 μl reaction (reaction B) was incubated at 37 ^{o}C for another 30 minutes. Finally, 6 μl of reaction B was added to a third PCR tube containing premix (6 μl), amino acid mix (2 μl), $S^{3}5$ methionine (0.2 μl), T7 S30 extract (4 μl) and 117SecSECISD8 DNA template (2 μl, 20 ng/μl). This third reaction mix (reaction C) was

incubated at 30 ^{o}C for 3 hours, centrifuged (4 ^{o}C, 25,000 rcf, 10 minutes) and the supernatant kept for analysis (precast Bis-Tris 12% gel, criterion).

In-vitro translation transcription (IVTT)

The standard procedure detailed here was used to translate WT αHL monomer, 117CysSECIS monomer and also for the temperature experiments (see section 2.1.3). For a 50 μl scale IVTT reaction, 8 μl DNA template (400 ng/μl), 5 μl amino acid mix (Promega IVTT kit), 20 μl premix (Promega IVTT kit), 2 μl $S^{3}5$ methionine and 15 μl S30 T7 *E. Coli* extract (containing rifampicin - 1 μl of 500 μg/ml into 150 μl of the S30 extract) were mixed and incubated at 37^{o}C for 1.5 hours. Alterations to this protocol were that the duration of the temperature experiments was 3 hours.

PEG-OPSS gel shift assay

PEG-OPSS (2 μl, 100 mM solution, polyethylene glycol 5 k, ortho-pyridyldisulfide, creativepegworks.com) was added to 20 μl IVTT reaction mix and incubated for 40 minutes at room temperature (i.e. 22 ^{o}C). After this time, the reaction was diluted with 10 μl sodium phosphate buffer (200 mM NaCl, 500 μM EDTA, 1 mM DTT, pH 8.3) and buffer exchanged using a p6 spin column (Biorad). 6 μl of XT (4X) loading buffer (Biorad) was then added to 20 μl of the reaction mixture and run on a 12 % Bis-Tris gel (Criterion, running buffer XT MOPS, 200 V, 40 minutes). The gel was then dried (under vacuum), exposed overnight and developed.

2.5.4 Miscellaneous

Making Competent cells

Transformation of an aliquot of XL10 Gold competent cells (or BL21 competent cells) was carried out without the presence of an antibiotic and using a dilution streaking method so as to obtain single colonies for inoculation. Luria Broth media (110mL, Sigma Aldrich) was placed into an autoclaved 200mL culture flask and a single colony used to inoculate the media. The culture was incubated at 37^{o}C and aerated at 200rpm in a NBS shaker (New Brunswick) until reaching an OD_{600} of 0.5 after which it was placed on ice. The culture was pelleted (2500rpm, 4^{o}C, 20 minutes, put centrifuge details here). A CO_2/Ethanol bath was made and the TSS buffer prepared as follows. PEG 800 (2g), DMSO (1ml), $MgCl_2$ (0.5ml, 1M solution) were dissolved in LB (final volume of 20mL) and filter sterilised. The culture pellet was then resuspended in the TSS buffer (10ml, 1/10th the original culture volume) and aliquoted into eppendorfs, pre-chilled in the CO_2/Ethanol bath.

Preparation of the rabbit red blood cell membrane (rbcm)

The rabbit red blood cells (10 ml) were centrifuged (700 rcf, 10 minutes) and the supernatant removed. Four to five wash steps were then carried out by gently resuspending the cells in MBSA buffer (40 ml, 150 mM NaCl, 10 mM MOPS, 0.1 % Bovine Serium Albumin (1 mg/ml), pH 7.4) followed by centrifugation. The pellet was resuspended in approximately 20 ml of MOPS buffer (the lysis buffer, 10 mM MOPS pH 7.4) and then centrifuged at 20,000 rcf for 30 minutes. The "soft pellet" or globular looking pellet part was collected. This pellet was washed 4-5 times with chilled MOPS buffer until the pellet colour was very pale pink. The rabbit red blood cell membrane was isolated at a concentration of approximately 3 mg/ml. Aliquots of 150 μl were made and stored at -80^{o}C. All buffers were chilled prior to use.

Chapter 3

Selenylsulfide and disulfide kinetics studied using single channel recordings

3.1 Introduction

3.1.1 The importance of selenocysteine *in vivo*

Evolutionary pressure for the maintenance of selenocysteine-containing proteins in the genome

Selenoprotein genes have been found to exist in the three domains of life (Eukaryota, Eubacteria and Archaea), and it has been proposed that selenocysteine was a late addition to the genetic code, occurring before the division into the three domains (Yuan et al., 2006). In support of this, it has been found that for all but one class of selenoproteins, a cysteine homologue exists (Fomenko et al., 2007; Baltazar et al., 2011) suggesting that these cysteine-containing proteins pre-date the selenocysteine-containing ones. Selenocysteine is widely distributed across the organisms in each domain of life. So far 25 different selenoprotein genes have been found to exist in humans and 24 in mice, with 15 different types of selenoprotein genes found in bacteria and archaea (Kryukov

et al., 2003). Out of the known bacterial organisms, approximately 25% use selenoproteins with the most selenocysteine-rich being the *Deltaproteobacteria* and *Clostridia* (Zhang & Gladyshev, 2010). Approximately 13% of archaeal microbes (6 organisms) synthesise selenoproteins with the most selenocysteine abundant are those of the *Methanococcale* and *Methanopyrale* orders (Stock & Rother, 2009). Even with the energetic cost associated with selenocysteine incorporation into proteins (see Introduction Chapter, Section 1.2), there is a low exchange rate between cysteine and selenocysteine in evolutionary terms (Wessjohann et al., 2007). Arnér (2010) interpreted this as pointing to an important functional difference between selenocysteine and cysteine, however the crucial advantage of selenocysteine over cysteine has yet to be discovered. A common evolutionary event is the replacement of cysteine by selenocysteine but that this has overall been balanced, by the loss of selenoprotein genes (Zhang et al., 2006). Neither the loss or acquisition of a selenoprotein gene was an irreversible event and Baltazar et al. (2011), concluded from this finding that most selenoproteins evolved from cysteine-containing analogues, again suggesting that the selenocysteine residue offers some sort of advantageous reactivity over cysteine, for these enzymes.

A catalytic comparison between selenocysteine and cysteine

The prokaryotic selenoproteins typically exhibit higher catalytic activity in terms of the rate at which they catalyse the conversion of substrate into product, compared to their cysteine homologues (Axley et al., 1991; Hatfield & Gladyshev, 2002). The reason for this is not known, but speculated to be due to the higher nucleophilicity of the selenocysteine residue as compared to cysteine (Huber & Criddle, 1967; Bunnett, 1963; Jaramillo et al., 2007). The cysteine analogues of selenocysteine-containing proteins are not always less efficient in catalysis however, (Kanzok et al., 2001; Gromer et al., 2003) and Arnér (2010) pointed out that an increased catalytic efficiency is not necessarily enough of an advantage to justify the use of selenocysteine in place of cysteine, as this could be compensated by increasing the expression level of the cysteine homologue (Arnér,

Property	**Se**	**S**
Electron configuration	$[Ar]3d^{10}.4s^2.4p^4$	$[Ne]3s^2.3p^4$
Atomic radius (in pm, emp.)[1]	115	100
Atomic radius (in pm, calc.)[2]	103	88
Covalent radius (in pm, emp.)[3–7]	116	102
Van der Waals radius (in pm)[3,5,8,9]	190	180
Ion radius (Pauling)[3,8] ox. state (I)	66	53
(VI)	42	29
(-I)	232	219
(-II)	198	184
Bond length (in pm) of dichalcogen[10]	232 (Se-Se)	205 (S-S)
Bond length (in pm) carbon-chalcogen	196 (C-Se)	18 (C-S)
Molecular volume (in cm^3)[7,10]	16.42	15.53
Diatomic bond energies (kJ mol^{-1})	332.6 (Se-Se)[12]	425.3 (S-S)[13]
	314.5 (Se-H)[14]	344.3 (S-H)[15]
	590.4 (Se-C)	714.1 (S-C)[17]
		371.1 (Se-S)[12]
Homolytic bond cleavage energies (kJ mol^{-1})		
HX-H	334.9[18]	381.6[19]
CH^3X-XCH^3	197.6[21]	204.7[21], 273.6[19]
Electron affinity (in kJ/mol^{-1})[3,5,22]	195.0	200
Ionisation energies (in kJ/mol^{-1})[10,23,24]		
1st	914.0	999.6
2nd	2045	2252
3rd	2973.7	3357
Redox potentials (V)[25]		
$XO_4^{2-} + H_2O + 2e^- \Rightarrow XO_3^{2-} + 2OH^-$ (pH = 7)	0.05	-0.93
$XO_3^{2-} + 6H^+ + 6e^- \Rightarrow X^{2-} + 3H_2O$ (pH < 7)	0.276	0.231
$X + 2e^- \Rightarrow X^{2-}$	-0.924	-0.476 (vs SCE)
$RX\text{-}XR + 2e^- \Rightarrow 2RX^-$; R=$CH_2CH(NH_2)COOH$	-0.488 (vs. NHE)[26]	-0.233 (vs. NHE)[26]
Polarisability (in $Å^3$)[27]	3.8	2.9
Pauling electronegativity[3–5,8,28]	2.55	2.58
pKa[29]	5.2	8.3

Table 3.1: Properties of selenium and sulfur adapted from Wessjohann et al. (2007). [1]Slater (1964), [2]Clementi et al. (1963), [3]Huheey et al. (1993), [4]Porterfield (1984), [5]James & Lord (1992), [6]Sanderson (1962), [7]Sutton (1965), [8]Pauling (1945), [9]Bondi (1964), [10]Singman (1984), [11]Allen (2002), [12]Drowart & Smoes (1977), [13]Huber & Herzberg (1979), [14]Gibson et al. (1986), [15]Johns & Ramsay (2007), [16]Smoes & Drowart (1977), [17]Coppens et al. (1979), [18]Berkowitz et al. (1994), [19]Nicovich et al. (1992), [20]Guziec (1979), [21]Gonzales et al. (2005), [22]Wheeler (1997), [23]Lide (1998), [24]Dean (1992), [25]Milazzo et al. (1978), [26]Jacob et al. (2003), [27](http://www.chemicool.com/elements/selenium.html, and /sulfur.html), [28]Allred (1961), [29]Huber & Criddle (1967)

2010).

3.1.2 The physical properties of selenocysteine and cysteine

Comparing the pKa of selenocysteine and cysteine

As it can be seen in Figure 3.1, selenium and sulfur exhibit comparable properties, with a pKa difference of 3 pKa units between selenocysteine and cysteine being the most contrasting feature (Wessjohann et al., 2007). An expected pKa value of approximately 5.5 for a selenocysteine amino acid within a protein (pKa = 5.2 for free selenocysteine), renders the selenol group (-SeH) in it's deprotonated, selenolate form at physiological pH. The necessity for selenocysteine solely based on this low pKa value does not explain why organisms prefer selenium over sulfur. For example, active site cysteines have been found which exhibit comparably low pKa values (Nelson & Creighton, 1994, eg Dsb enzymes). The tuning of the pKa value of active site cysteine by the nature of the adjacent amino acids, is observed *in vivo* (Harris & Turner, 2002). Bulaj et al. (1998) carried out an investigation into the effect of different adjacent amino acids on the pKa and disulfide reactions for a series of model peptides containing a single cysteine. They concluded that the pKa was important in dictating the rate of disulfide exchange, with sterics also playing an important role. Residues flanking low pKa cysteines are often polar and able to form hydrogen-bonds with surrounding residues. As the pKa of selenocysteine is intrinsically lower, there may be less of a constraint on which residues must be positioned nearby. From an evolutionary point of view, an intrinsically lower pKa may be a useful attribute if the enzyme were to require two conflicting properties to function e.g. a hydrophobic binding pocket and low pKa.

The nucleophilicity of selenocysteine vs cysteine

Selenium and sulfur both belong to group 16 of the periodic table. Both atoms exhibit the same number of electrons in the outer most shell ($3s^2$, $3p^6$ for sulfur, $4s^2$, $4p^6$ for selenium), however selenium has 18 more electrons than sulfur. These 18 electrons reside in the third orbital shell ($3s^2$, $3p^6$, $3d^{10}$). The nucleophilicity of selenocysteine is described as much better than that of cysteine (Jaramillo et al., 2007). From NMR experiments (Pleasants et al., 1989) it has been shown that the selenol group can undergo a reaction with diselenides with reaction rates up to 10^7 times faster than the analogous thiol (with disulfides), at a given pH. Another important difference between selenium and sulfur is the ease with which they can move between oxidation states (Axley et al., 1991). Sulfur, for example, once doubly oxidised to it's sulfinic acid form cannot be reduced back using other thiol containing molecules (eg DTT) whereas the seleninic acid is reported to be easily reduced back to its selenol form using excess DTT (Wu & Hilvert, 1990).

3.1.3 Summary

The incorporation pathway of selenocysteine *in vivo*, utilises ATP during the production of selenophosphate by selenophosphate synthetase (see Introduction Chapter, Section 1.2.2) and proceeds via a rather inefficient pathway due to competition between the SelB·GTP·Sec-tRNA^{sec} complex and the translational termination factor, RF2 (See the Introduction and Chapter 2 for further details). Although both selenium and sulfur are similar, in terms of physical properties (Huber & Criddle, 1967), they exist in very different ratios of selenolate to selenol and thiolate to thiol at physiological pH. For example, the pKa values of the free amino acids selenocysteine and cysteine, are 5.2 and 8.3 respectively. At physiological pH (pH 7.4), the selenol group is in its deprotonated, selenolate form but the thiol group of cysteine is protonated.

3.2 Results

By using αHL heteroheptameric pores containing a single selenocysteine or cysteine residue within the β barrel (117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ respectively) in conjunction with single channel recordings, it was possible to compare the nucleophilicity, electrophilicity and leaving group ability of selenocysteine compared to cysteine, within a protein environment. Due to the ease at which the selenocysteine residue became oxidised, almost all of the experimental single channel recording data presented in this chapter was obtained using the 117CysSECIS$_1$WT$_6$ αHL pore.

3.2.1 Characterisation of 117CysSECIS$_1$WT$_6$ and 117SecSECIS$_1$WT$_6$ αHL heptamers

A single selenocysteine (Sec) or cysteine (Cys) residue was incorporated into the β barrel of αHL at position 117 (see Chapter 2 for method details regarding selenocysteine incorporation), by formation of a heptameric pore where only one monomeric unit contained a Sec or Cys (117SecSECIS$_1$WT$_6$ or 117CysSECIS$_1$WT$_6$ respectively). The 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ pores were used to carry out single channel recording experiments. The electrophysiology set up is illustrated in Figure 3.1. Briefly, two buffer-filled chambers (1 mL) are separated by a piece of 25 μm-thick teflon with an aperture of approximately 100 μm diameter at the centre of the teflon. The aperture was carefully coated with hexadecane and a planar lipid bilayer of DPhPC was formed across the hole (see Figure 3.1, **A**) by slow pipetting of the buffer solution up and down, in each chamber. On addition of the 117SecSECIS$_1$WT$_6$ (PSe$^-$) αHL heptamer to the *cis* (ground) chamber, it was possible to insert a single heptameric protein into the bilayer between the two chambers. Agarose covered, Ag/AgCl electrodes were placed into each chamber prior to addition of the 117SecSECIS$_1$WT$_6$. On insertion of an αHL pore, ions were able to move

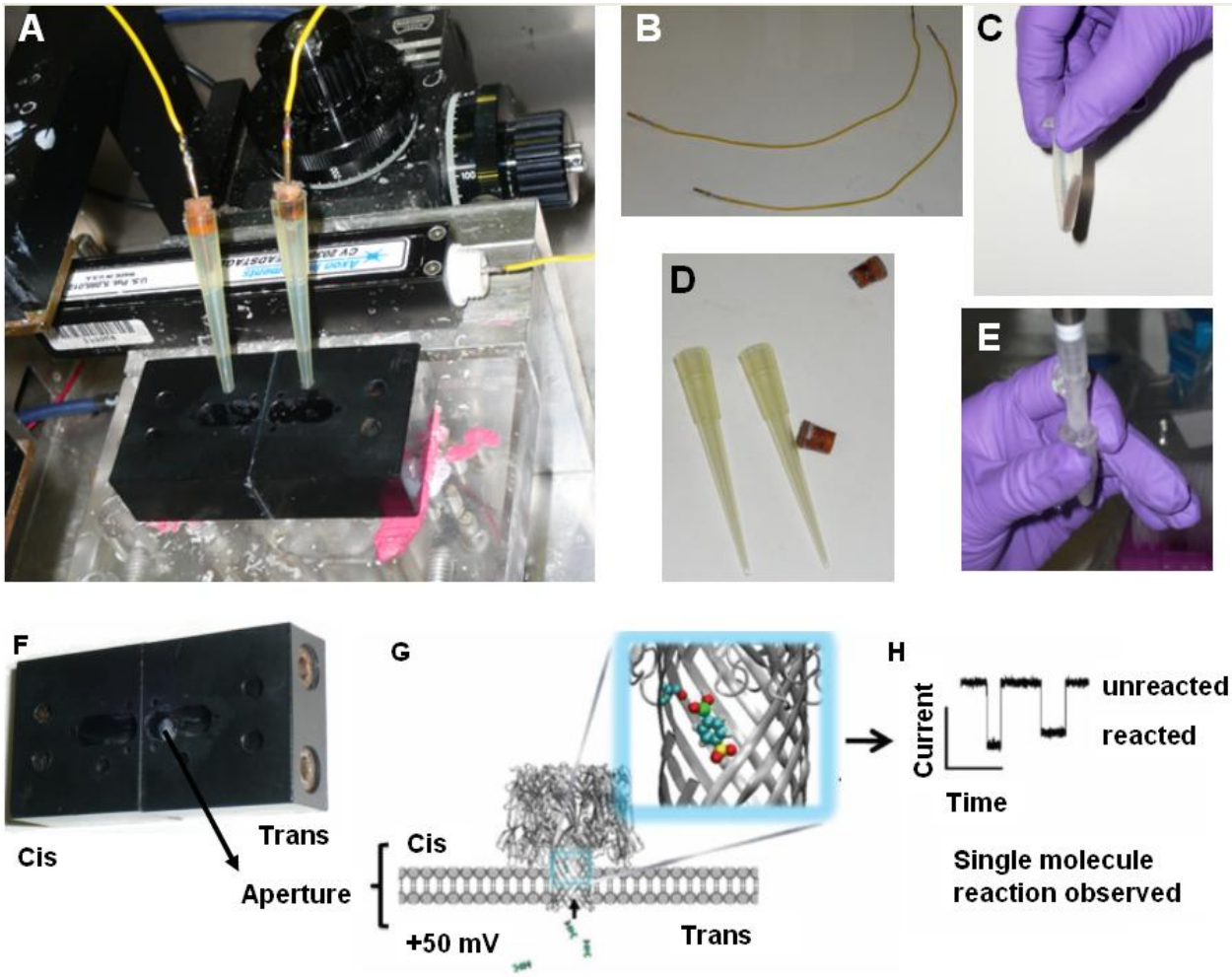


Figure 3.1: A; The electrical set up used to carry out all single channel recordings. B; The electrodes attached to the electrical wires. C; The Ag/AgCl coated silver electrodes. D; The pipette tips and rubber stoppers used to make the agarose bridges. E; Pipetting the agarose into the pipette tips to make the agarose bridges. F; The electrical recording chamber used to carry out the single-molecule experiments. A voltage was applied via the pair of electrodes placed in the cis (ground) and trans chambers. A single αHL heptameric pore was inserted into a planar lipid bilayer. G; The interaction of single molecules in solution, reacting with the selenocysteine or cysteine residue within the β barrel of αHL was observed by monitoring the current amplitude changes with respect to time. H; An example of the current recorded over time. The reaction of a cysteine residue with a small molecule in solution produces a drop in current from the "unreacted" level to the "reacted" level. By recording many of these events over time, it was possible to obtain kinetic rate constants.

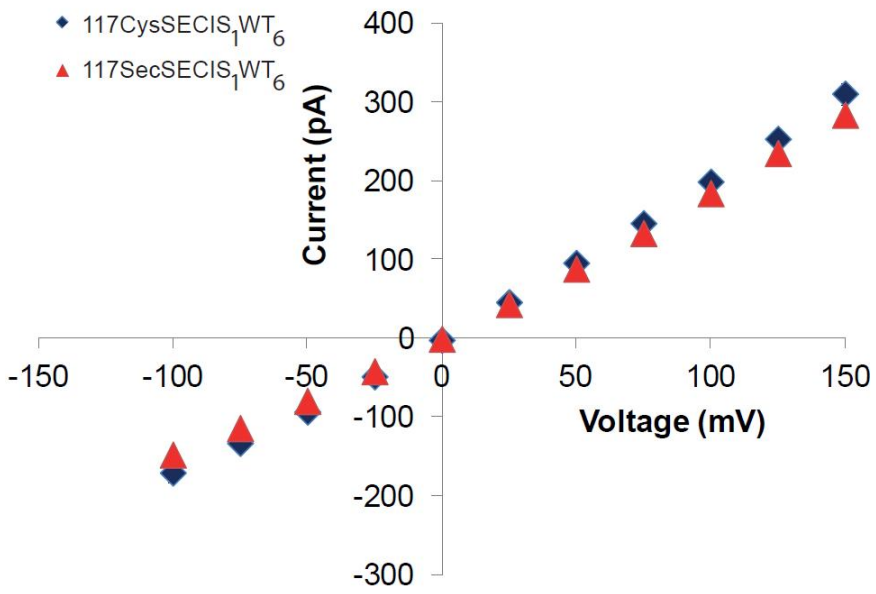


Figure 3.2: The relationship between current and voltage for both the 117CysSECIS$_1$WT$_6$ and 117SecSECIS$_1$WT$_6$ αHL heteroheptameric pores. Error bars reside within the data points.

through the pore between the *cis* and *trans* chambers. Due to the applied potential, a net flow of ions through the pore produced a current. The typical current observed on applying a particular voltage has been plotted for both the 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ pores (Figure 3.2). Any perturbation in the flow of ions through the pore was observed as an alteration in the current magnitude. For example a covalent bond formed between the single 117Sec or 117Cys residue within the β barrel would be observed as a decrease in current (see Figure 3.1, **C**). In this way, reversible covalent reactions were observed between the selenocysteine or cysteine residue, and small molecules in solution. The unitary conductance of the 117SecSECIS$_1$WT$_6$ heptameric pore in 2 M KCl at +50 mV (22 ^{o}C) was 1.8 $\pm$ 0.1 nS (3 separate experiments). Similarly, the conductance of the 117CysSECIS$_1$WT$_6$ (PSH) heptameric pore was 1.7 $\pm$ 0.1 nS (n = 3). The current-voltage relationship for both of these heteroheptamers shown in Figure 3.2, were similar to that of an αHL heptamer made from six WT monomer units and one 117Cys containing unit, without the SECIS element amino acid alterations in the β barrel (1.54 nS, Shin et al. (2002)).

3.2.2 Reaction of the cysteine containing αHL pore with *trans*-4,5-dihydroxy-1,2-dithiane

The reaction of oxidised DTT (DTTox, *trans*-4,5-dihydroxy-1,2-dithiane) with the 117Cys -SECIS$_1$WT$_6$ heptameric αHL pore was studied by addition of DTTox to the *cis* chamber. The reaction of DTTox with the 117Cys residue within the β barrel of αHL is illustrated in Scheme 3.1, **A**. Initial attack of the 117Cys thiolate group on the disulfide bond of DTTox formed a short lived disulfide intermediate, PS-DTT (where "PS" refers to the disulfide bond between 117Cys and DTT within the β barrel of 117CysSECIS$_1$WT$_6$).

By applying a constant voltage of +50 mV the reaction between DTTox and 117CysSECIS$_1$WT$_6$ was investigated between pH 7 - 9.5. DTTox diffused from the *cis* chamber (see Figure 3.1) into the β barrel and on reaction with the 117Cys residue, caused a current decrease (an event) of 2.2 ±0.1 pA (n = 7) (at +50 mV in 2 M KCl, 20 mM BisTris Propane,0.1 mM EDTA, pH 7.0 - 9.5, where this buffer was used throughout and all experiments were carried out at 22^{o}C if not otherwise stated). To verify that the events observed were due to the reaction of the 117Cys thiol group with DTTox, DTNB was added to the *trans* chamber after each experiment and an irreversible current decrease was observed due to the formation of PS-TNB (TNB; 2-nitro-5-thiobenzoic acid). In some experiments this control experiment was not possible, for example if the bilayer broke or became too unstable, then the rate constants were compared to other experiments. Control single channel recordings were carried out by inserting a single WT$_7$ heptamer into the bilayer and adding 80 mM DTTox to the *cis* chamber. No events were present, verifying that the current blockages (events) observed during the 117CysSECIS$_1$WT$_6$ single channel recordings, were due to the reaction of the 117Cys residue with DTTox. It was expected that two current blockage levels would exist, one for the "open pore" or 117CysSECIS$_1$WT$_6$ pore prior to reaction with a DTTox molecule, and the other for the short lived PS-DTT intermediate. The DTTox used during the the electrical recordings was enantiomeric, with both the *D* and *L* forms present as a mixture.

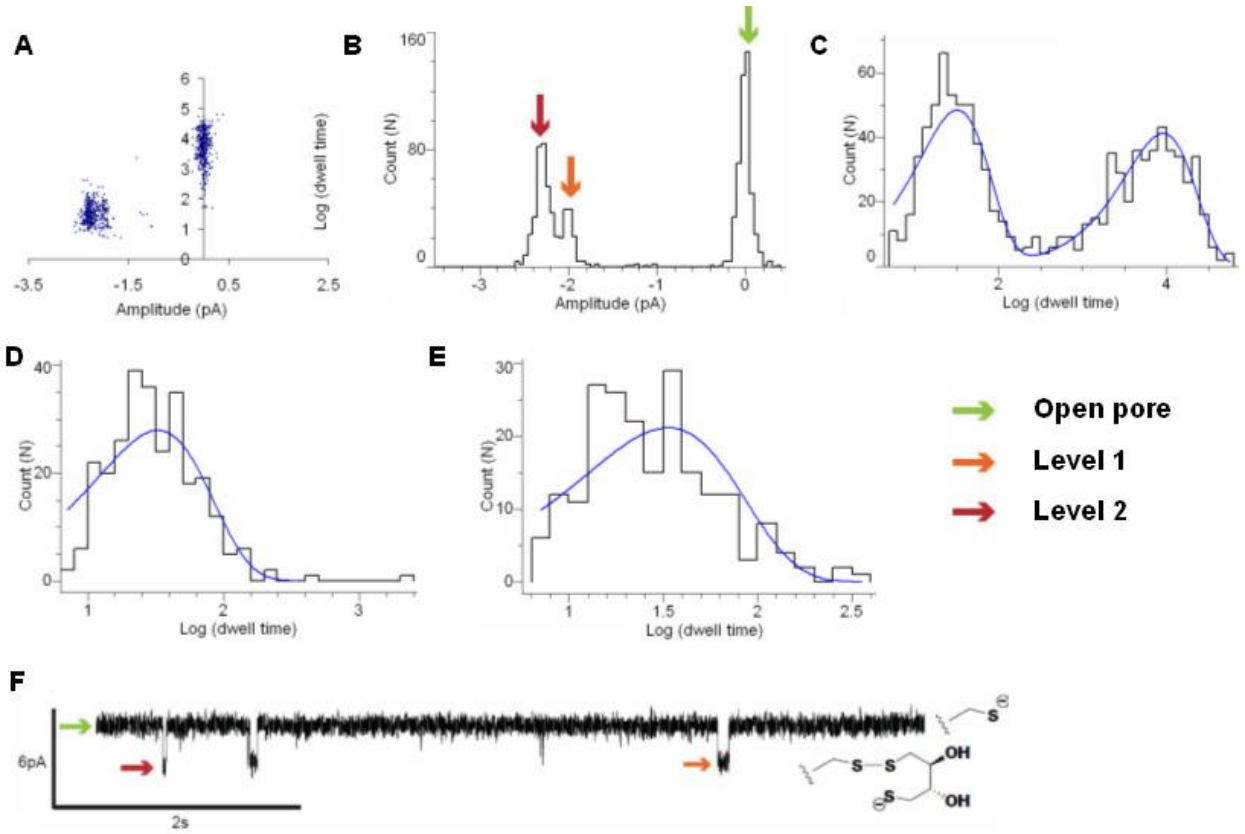
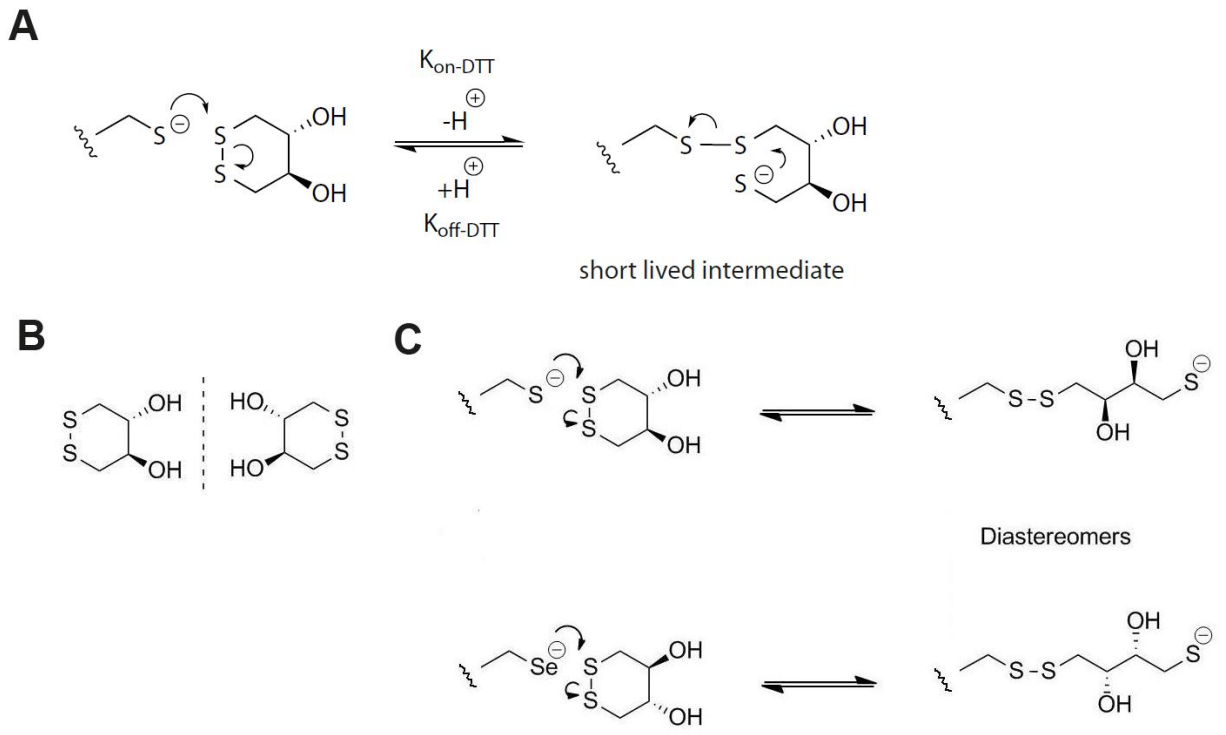


Figure 3.3: A; The relationship between the event amplitude and the dwell time. It can be seen that the two populations of amplitude are very close. B; The current amplitude of the short lived intermediate formed on reaction of DTTox with 117CysSECIS$_1$WT$_6$ at pH 8.5. The occurrence of level 1 events compared to that of level 2 events is approximately 32 %. C; An example histogram illustrating two dwell time distributions representing the unreacted open pore level (open pore, green) and the PS-DTT intermediate level (level 1 and 2, orange and red). The number of events (N) represents the number of dwell times within a particular bin. D; A histogram representing the distribution of dwell times for level 1. E; A histogram representing the distribution of dwell times for level 2. Levels 1 and 2 are so close in amplitude and dwell time distributions that the data was very difficult to resolve satisfactorily. F; The reaction of DTTox with 117Cys at pH 8.5. The current block levels 1 (orange) and 2 (red) are due to formation of the PS-DTT diastereomeric intermediates.

Because the difference in structure between the *D* and *L* -DTTox species is small, it was hypothesised that the current blockage difference between these isomers, would be non-distinguishable. However, upon analysis three distinct current levels were observed. Figure 3.3, **B** illustrates the two current blockage levels produced on reaction of DTTox with 117Cys to form diastereomeric PS-DTT species. The amplitudes of these two blocking levels are 2.2 $\pm$ 0.1 pA (denoted level 2, n = 7) and 1.94 $\pm$ 0.04 pA (denoted level 1, n = 5) giving a difference between the two levels of 0.3 pA. The two distinct levels were proposed to be observed (see Figure 3.3), due to the non-symmetrical nature of the β barrel surface, creating diastereomeric intermediates (see Scheme 3.1, **C** and Discussion, Section 3.3.2) however to fully confirm this hypothesis, experiments using a single enantiomer would need to be carried out. Just over 70 % of the 117CysSECIS$_1$WT$_6$ αHL pores were able to distinguish between the two DTT bound (PS-DTT) levels. Figure A.8 illustrates the amplitude graph for the reaction of 117CysSECIS$_1$WT$_6$ with DTTox at pH 8.5 in which the two PS-DTT current levels were not distinguishable. This was thought to be due to the population of the folding of the heptamers, which would affect the magnitude of the current blockages due to changes in the β barrel conformation. The dwell times for the PS-DTT intermediate have been tabulated and plotted against pH (see Figure A.9). The dwell times for each of the PS-DTT levels, were found to be comparable at pH 8.5 with rate constants of k_{off} = 45 $\pm$ 5 s^{-1} and k_{off} = 34 $\pm$ 7 s^{-1} respectively (see Scheme 3.1, **A** and Figure A.7). It was noted that, at pH 8.5, the frequency of transitions to level 1 was lower (approx. 30 % of the time) than to level 2 and has been illustrated in Figure 3.3, **A** by plotting the amplitude against the count number (the number of events of that amplitude).

Rate constants for the reaction of DTTox with 117Cys

The rate constants for the forward (k_{on-DTT}) and reverse ($k_{off-DTT}$) reactions (see Scheme 3.1, **A**) were determined by analysis of the mean inter-event intervals (τ_{on}) and mean event dwell times



Scheme 3.1: **A** Attack of 117Cys on the disulfide of DTTox generated a short lived intermediate, observable via single channel recordings using $117Cys_1WT_6$. **B** The *trans*-4,5-dihydroxy-1,2-dithiane was racemic where both the above structures were in existence. **C** *Trans*-4,5-Dihydroxy-1,2-dithiane (DTTox) exists as two enantiomers. On reaction of DTTox with the $117SecSECIS_1WT_6$ (or $117CysSECIS_1WT_6$), a diastereomeric intermediate would be formed due to the non-symmetrical nature of the β barrel surface.

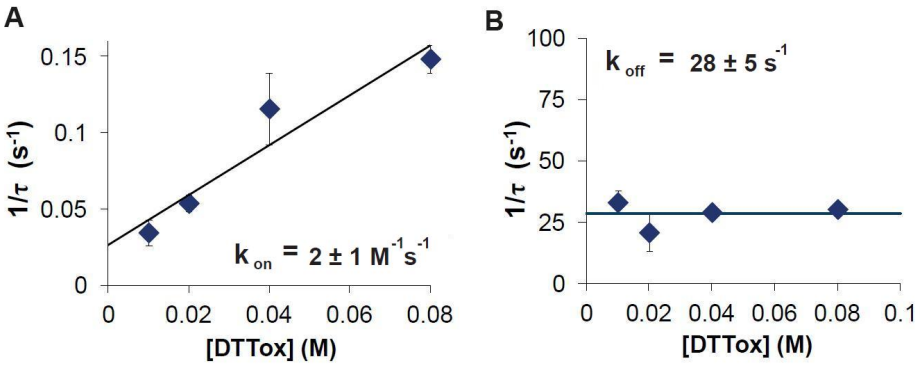


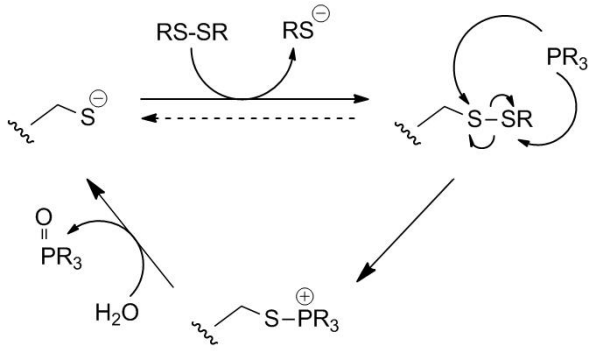
Figure 3.4: The reaction of DTTox with $117CysSECIS_1WT_6$ at pH 8.0. **A** The relationship between k_{on} and DTTox concentration. **B** The relationship between the life time of the PS-DTT intermediate and DTTox concentration.

(τ_{off}). τ_{on} is the amount of time before a molecule of DTTox reacts to form a covalent bond with the 117Cys residue within the β barrel (i.e. PS-DTT formation). Each τ_{off} is the amount of time that the PS-DTT intermediate exists, before cyclisation of DTT to DTTox occurs. For further details regarding the data analysis, see Methods, Section 3.4.7. The concentration dependent rate constants ($1/\tau_{on}$ and $1/\tau_{off}$) were plotted against concentration of DTTox as shown in Figure 3.4. The forward reaction (k_{on}) is pseudo-first order with respect to the DTTox concentration and the reverse reaction, involving cyclisation of the DTT intermediate back to DTTox, is unimolecular and therefore concentration independent. Rate constants for the reaction of 117CysSECIS$_1$WT$_6$ with DTTox at pH 8.5 are $k_{on} = 2 \pm 1$ $M^{-1}s^{-1}$ and $k_{off} = 28 \pm 5$ s^{-1} (see Table A.1). These rate constants were obtained from the reaction of 117CysSECIS$_1$WT$_6$ with increasing concentrations of DTTox (*trans*). Data was acquired at a sampling rate of 20 kHz and filtered at 5 kHz. Further filtering at 1000 Hz was carried out after the data were acquired. A further experiment, to verify that the two current blockage levels are the result of the reaction between 117Cys and the different enantiomers of DTTox, would be to react the 117CysSECIS$_1$WT$_6$ with only one of the enantiomers of DTTox.

3.2.3 Reduction of selenylsulfide and disulfide bonds

Reduction of disulfide bonds by TCEP

The 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ pores were used to study the reduction of selenylsulfide and disulfide bonds using the water soluble trialkylphosphine reagent, TCEP (tris(2-carboxyethyl)phosphine). Perfusion of the *cis* chamber (exchange of buffer for fresh buffer) was carried out using three times the chamber volume so as to remove any DTT from the heptamer solution (this was always carried out before all single channel experiments). To form a disulfide bond (i.e. PS-SR) within the β barrel of the 117CysSECIS$_1$WT$_6$ pore, either DTNB, cystine or (3-nitro-2-pyridylthio)-L-cysteine (Cys(Npys)-OH) were added to the *trans* chamber. The reaction



Scheme 3.2: The reaction of 117CysSECIS$_1$WT$_6$ (PS^-) with TCEP and a small molecule disulfide. Initial attack of the 117Cys thiolate group on the disulfide (represented by RS-SR ie either DTNB, cystine or Cys(Npys)-OH) in solution, formed a mixed disulfide. TCEP (PR_3) present in the *cis* chamber reduced the disulfide bond back to 117Cys (i.e. PS^-) via formation of a phosphonium (V) cation intermediate (PS-P^+R_3). The dashed arrows represent the alternative pathway where by TCEP attacked the other sulfur atom of the disulfide bond.

between the disulfide reagent, DTNB (cystine or Cys(Npys)-OH) and the 117Cys residue was observed as a decrease in current of 4.2 $\pm$ 0.5 pA (n = 3) for DTNB and 1.3 $\pm$ 0.3 (n = 3) for cystine or Cys(Npys)-OH, at +50 mV. TCEP was added to the *cis* chamber and the single channel current recorded at different TCEP concentrations. The overall reaction mechanism for the reduction of a disulfide situated within the β barrel of 117CysSECIS$_1$WT$_6$ by TCEP, is shown in Scheme 3.2. The initial attack of TCEP onto the disulfide bond (PS-SR) generates a phosphonium cation species and thiolate. This step is followed by the formation of R_3P=O, initiated by attack of water or OH^- onto the phosphorous (V) centre of the phosphonium cation, resulting in the formation of a thiolate group and the trialkylphosphine oxide (R_3P=O). Cys(Npys)-OH was employed as an alternative reagent to cystine because the reaction rate of cysteine with 117CysSECIS$_1$WT$_6$ was very slow with respect to the practical time scale for monitoring, using the single channel electrical recording set up. Typical single channel current traces for the reaction of DTNB and TCEP with 117CysSECIS$_1$WT$_6$ at pH 8, 8.5 and 9.0 are shown in Figure 3.5. It can be observed from these data that the phosphonium cation was formed by nucleophilic attack of TCEP onto the sulfur atom of the 117Cys residue (of the PS-TNB) and caused a current blockage (relative to the unreacted

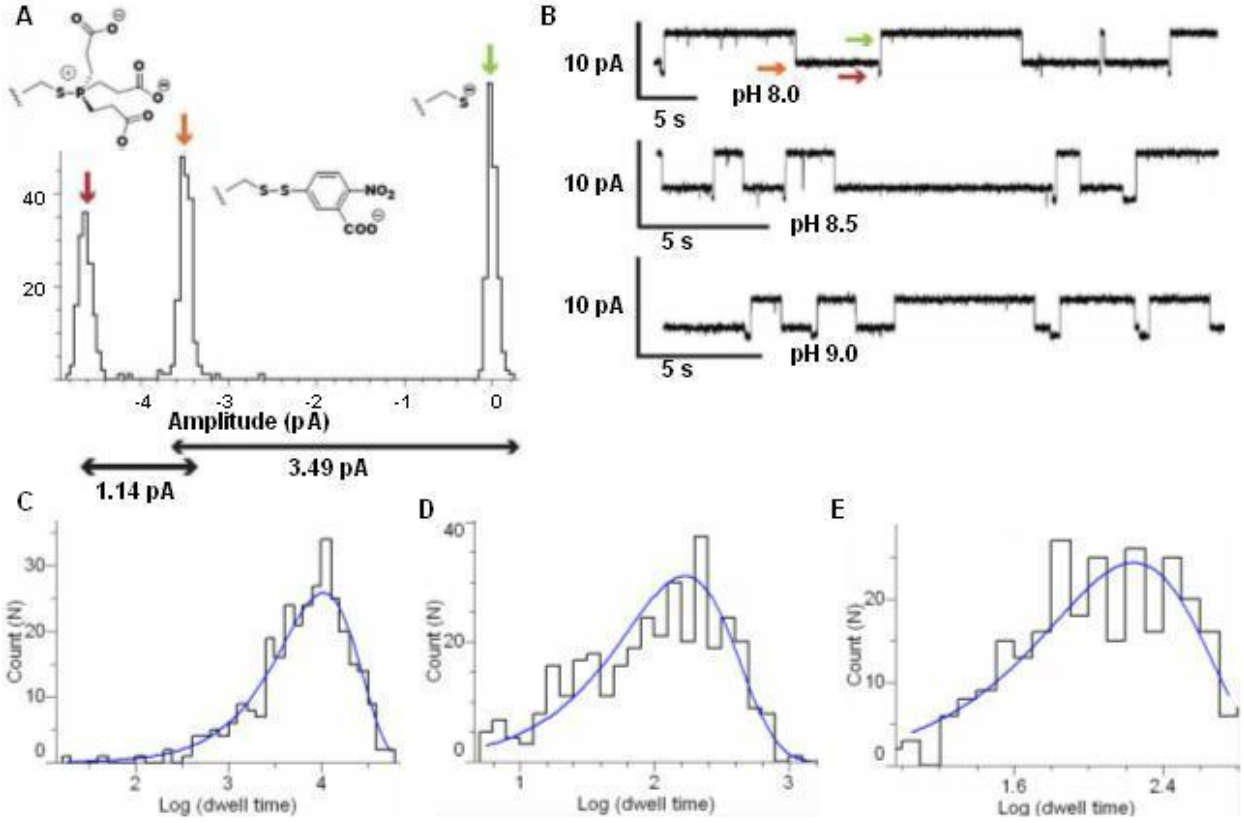


Figure 3.5: A; The current blockage amplitudes observed on reaction of 117CysSECIS$_1$WT$_6$ with DTNB, followed by reduction of the PS-TNB disulfide bond using TCEP. The green, orange and red arrows depict the levels for the unreacted pore, PS-TNB and PS-P$^+$R$_3$ (phosphonium cation) respectively. B; Single channel current trace examples for the reaction of 117CysSECIS$_1$WT$_6$ with TCEP and DTNB at pH 8.0, 8.5 and 9.0. C; The dwell time histogram for the reaction of the single cysteine residue situated within the β barrel of αHL with DTNB (*trans*). D; The dwell time histogram for the reduction of the mixed disulfide species formed after the reaction of DTNB with the single cysteine residue, by TCEP. E; The dwell time histogram for the short lived phosphonium intermediate.

pore level) of 5.1 $\pm$ 0.5 pA (n = 3) (see Figure 3.5, **A**). Rate constants for the reduction of the disulfide bonds formed within the 117CysSECIS$_1$WT$_6$ pore (i.e. PS-TNB and PS-Cys) were obtained and are discussed in Section 3.3.3.

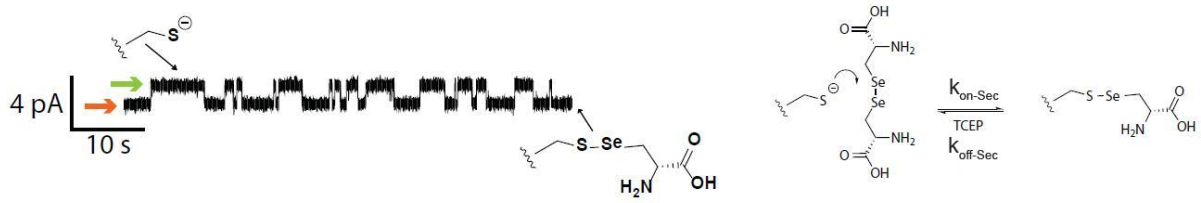


Figure 3.6: Single channel current trace for the reaction of 117CysSECIS$_1$WT$_6$ with selenocystine and TCEP at pH 7.0. Only two current blockage levels were observed because the TCEP exclusively attacked the selenium atom of the selenylsulfide bond (PS-Sec) formed within the β barrel of the 117CysSECIS$_1$WT$_6$ αHL pore.

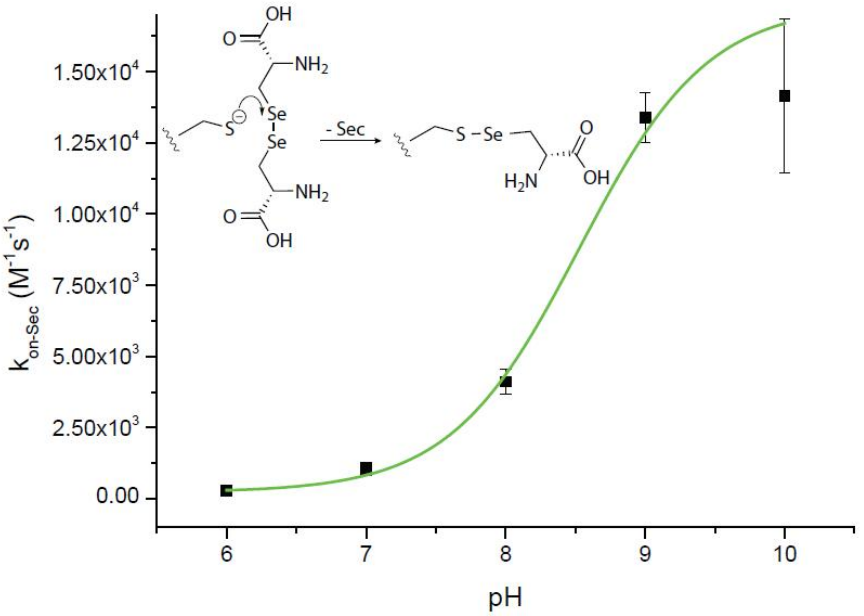


Figure 3.7: The relationship between the k$_{on-Sec}$ and pH for the reaction of 117CysSECIS$_1$WT$_6$ with Sec$_2$

Reduction of a selenylsulfide bond using TCEP

The reaction of 117CysSECIS$_1$WT$_6$ with selenocystine (Sec$_2$) was carried out in an analogous way to the reaction of 117CysSECIS$_1$WT$_6$ with cystine or Cys(Npys)-OH (see Figure 3.6). Selenocystine was pipetted into the *trans* chamber and reacted with the 117Cys residue to form a selenylsulfide species which was observed as a current decrease of 1.6 $\pm$ 0.3 pA (n = 5, +50 mV) relative to the unreacted 117CysSECIS$_1$WT$_6$ current level. This current blockage is the same (within error) as the current blockage observed for the reaction of 117CysSECIS$_1$WT$_6$ with cystine or Cys(Npys)-OH (1.3 $\pm$ 0.3 pA). TCEP was added to the *cis* chamber to reduce the selenylsulfide bond in an analogous way to that described for reduction of the disulfide (PS-SR) bonds (see Section 3.2.3). Rate constants were calculated for the reaction of 117CysSECIS$_1$WT$_6$ with Sec$_2$ (k_{on-sec}), and the reduction of the selenylsulfide bond formed between the 117Cys residue and selenocysteine by TCEP ($k_{off-sec}$). The addition of increasing amounts of TCEP to the *cis* chamber allowed the mean dwell times (τ) to be obtained at each concentration. These titration data were used to calculate the rate constants and data can be found in Figure A.11. The phosphonium cation intermediate was never observed during the reaction of TCEP and PS-Sec showing that TCEP exclusively attacked the selenium atom of the selenylsulfide bond (PS-Sec). The rate constant for the reaction of Sec$_2$ with the 117Cys residue (k_{on-Sec}) was plotted against pH over the pH range 6 - 10 (see Figure 3.7) and illustrates how the k_{on-Sec} increases in a sigmoidal fashion with increasing pH. By considering the equilibrium between the protonated thiol form of 117Cys (i.e. PSH) and the thiolate (i.e. PS^-), a relationship between this equilibrium constant (Ka) the rate of reaction with selenocystine was derived. The k_{on-sec} values obtained at pH 6, 7, 8, 9 and 10, were fitted to $\frac{k_{on-max}K_a}{(K_a+[H^+])}$ where the pKa of the 117Cys residue was obtained as pKa = 8.5 $\pm$0.3 and the k_{on-max} as (17 $\pm$ 4) x 10^3 $M^{-1}s^{-1}$ (see Figure 3.7). The $k_{off-Sec}$ was similarly plotted against pH (see Figure 3.8) and the relationship between $k_{off-Sec}$ and pH was also sigmoidal. By considering the proposed mechanism for the reduction of disulfide bonds by TCEP, this sigmoidal

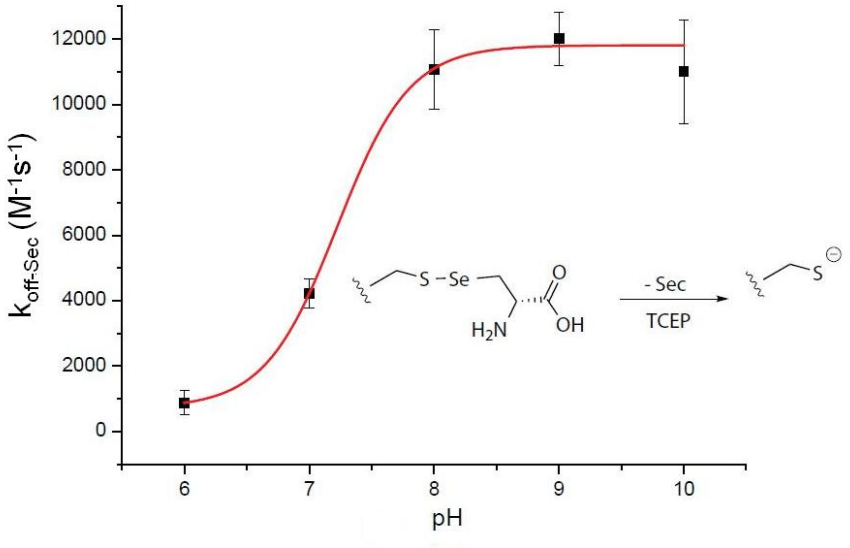
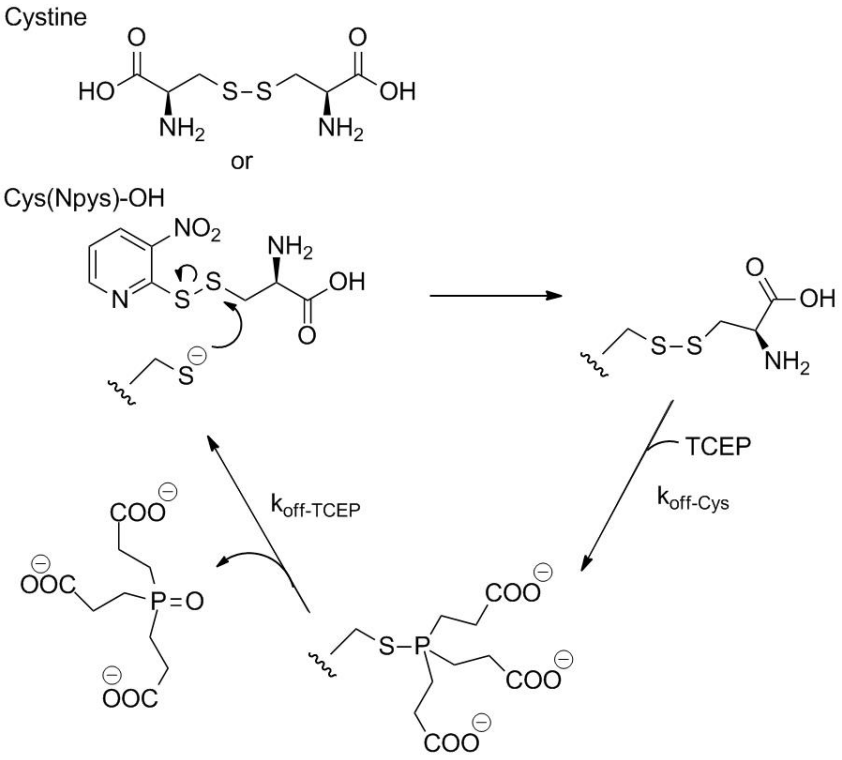


Figure 3.8: The relationship between the $k_{off-Sec}$ and pH for the reaction of 117CysSECIS$_1$WT$_6$ with Sec$_2$

relationship was proposed to be due to the pKa of the TCEP phosphine (pKa = 7.6 Burns et al. (1991)). The sigmoidal fit gave the pKa of the phosphine group as 7.3 $\pm$ 0.5 and the maximum $k_{off-Sec}$ as (1.2 $\pm$ 0.4) x 10^4 $M^{-1}s^{-1}$. Analysis of the rate constant for the PS-P^+R_3 phosphonium cation species ($k_{off-TCEP}$) was only possible from reduction of the mixed disulfide formed between the 117Cys residue and DTNB (PS-TNB). The lifetime of this intermediate (PS-P^+R_3) was found to have no pH dependence over the pH range 6 - 10, with the $k_{off-TCEP}$ = 5.9 $\pm$ 0.7 s^{-1} (n = 3).

Selenylsulfide and disulfide reduction using TCEP

The rate at which TCEP reduced the selenylsulfide bond formed between the 117Cys residue within the 117CysSECIS$_1$WT$_6$ pore and selenocysteine (PS-Sec) was compared with the rate of reduction of an analogous disulfide bond (PS-Cys). The disulfide bond between the 117Cys residue and cysteine, was formed using cystine or Cys(Npys)-OH ((3-nitro-2-pyridylthio)-L-cysteine, see Scheme 3.3). Reduction of the disulfide bond, PS-Cys, by TCEP was two orders of magnitude



Scheme 3.3: The reaction of 117CysSECIS$_1$WT$_6$ with cystine or cys$_2$ or (3-nitro-2-pyridylthio)-L-cysteine (cys(Npys)-OH) followed by reduction by TCEP.

slower than reduction of the selenylsulfide, PS-Sec bond at pH 9.0, and rate constants of $k_{off-Cys}$ = 1.85 $\pm$ 0.21 $M^{-1}s^{-1}$ (n = 2) and $k_{off-Sec}$ = (1.2 $\pm$ 0.4) x 10^4 $M^{-1}s^{-1}$ were obtained respectively. The phosphonium cation intermediate, PS-P^+R_3 was only very rarely observed during the reduction of the PS-Cys disulfide bond. Two possible explanations for this were that attack of the trialkylphosphine (TCEP) was affected by steric hinderance due to residues surrounding the PS-Cys bond and also by the leaving group ability of the sulfur atoms in the PS-Cys disulfide bond.

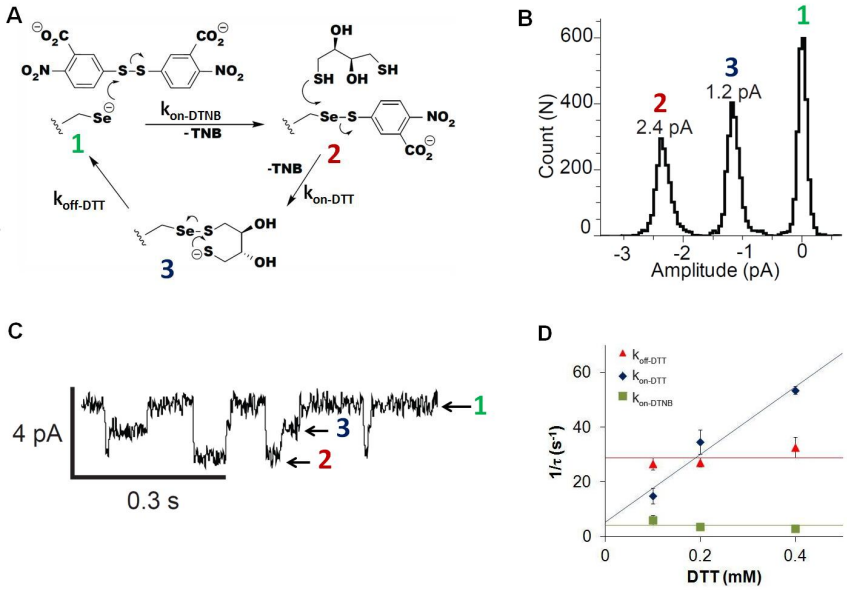


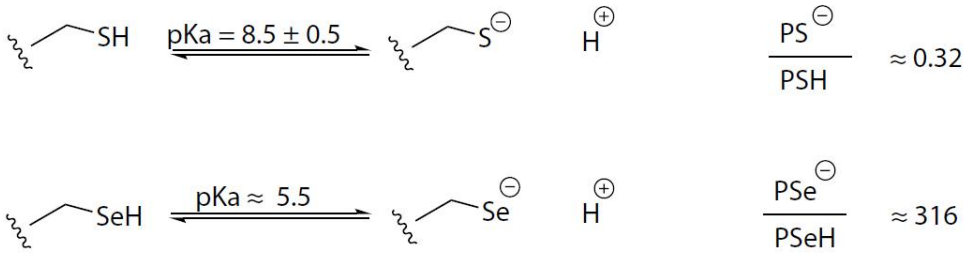
Figure 3.9: **A** The reaction of 117SecSECIS$_1$WT$_6$ with DTNB and DTT. **B** The current amplitude of the events. **C** An example trace illustrating the different current levels observed. **D** The relationship between $1/\tau$ and DTT concentration.

3.2.4 Reduction of a selenylsulfide bond using dithiothreitol

The reaction between the selenocysteine residue, situated within the β barrel of 117SecSECIS$_1$WT$_6$, and DTNB (added to the *trans* chamber) to form the selenylsulfide bond PSe-TNB, was observed from single channel current recordings. Reduction of the selenylsulfide bond, PSe-TNB was carried out by addition of DTT (DL-dithiothreitol) to the *cis* chamber. The selenylsulfide bond formed between the 117Sec residue and TNB (PSe-TNB) produced a current blockade of 2.4 ± 0.5 relative to the unreacted pore current level (see Figure 3.9, **B**). The PSe-DTT intermediate formed on attack of DTT on the PSe-TNB bond was observed as a current decrease of 1.2 ± 0.3 at pH 8.0 using an applied voltage of +50 mV. The overall reaction scheme has been illustrated in Figure 3.9, **A**. By addition of increasing amounts of DTT (*cis*), the mean dwell time (τ) for each selenylsulfide bond formed during the reaction pathway was obtained. By plotting $1/\tau$ against DTT concentration, it was possible to obtain the rate constants for the formation of PSe-TNB ($k_{on-DTNB}$), PSe-DTT (k_{on-DTT}) and regeneration of the unreacted pore, PSe^- ($k_{off-DTT}$, see Figure 3.9, **D**).

3.3 Discussion

3.3.1 Determining the pKa of the 117 cysteine residue situated within the β barrel of αHL



Scheme 3.4: The protonation-deprotonation equilibria for the 117Sec and 117Cys residues within the αHL β barrel. The ratio of selenolate to selenol and thiolate to thiol have been approximated at pH 8 using the experimentally determined pKa of 117Cys (pKa = 8.5 $\pm$ 0.5) and the assumption that the pKa of 117Sec would be 3 units lower.

The rate constant (k_{on-Sec}) for the reaction of selenocystine (Sec_2) with 117Cys was obtained over the range of pH 6 - 10. By considering the protonation-deprotonation equilibrium (see Scheme 3.4) of the 117Cys thiol group, it was possible to define k_{on-Sec} as a function of pH. The function ($k_{on-Sec} = \frac{k_{on-max}K_a}{(K_a+[H^+])}$, see section 3.2.3) was used to fit to the sigmoidal relationship between the k_{on-Sec} and pH, yielding a pKa = 8.5 $\pm$ 0.5 for the 117Cys residue within the $117CysSECIS_1WT_6$ pore (see Figure 3.7). By assuming that the pKa of the 117Sec residue is 3 pKa units lower than that of the 117Cys (Metanisa et al., 2006), the ratio of selenolate to selenol (of the 117Sec) at pH 8 was calculated to be 316 (from $10^{(pH-pKa)}$ = [Se^-/SeH]) and the ratio of thiolate to thiol (using the pKa = 8.5 for 117Cys) is approximately 0.32. This means that there exists 988 times more selenolate to thiolate at pH 8.

3.3.2 Comparing the reactivity of selenium and sulfur

Comparison between the nucleophilicity of selenocysteine and cysteine

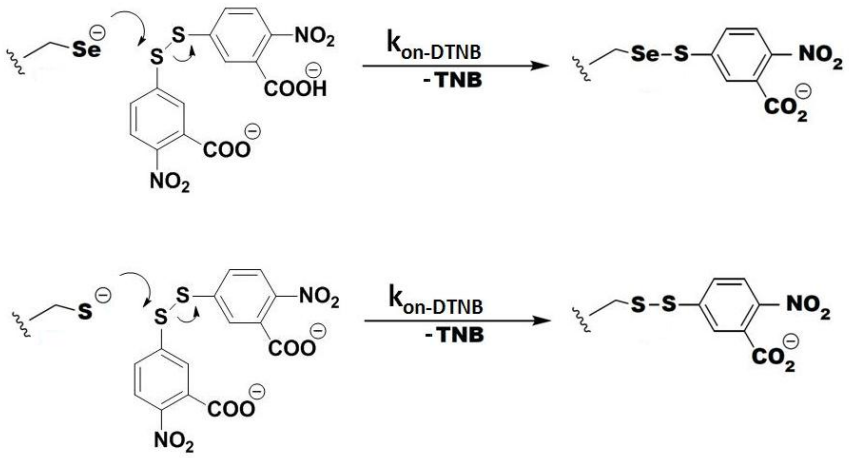
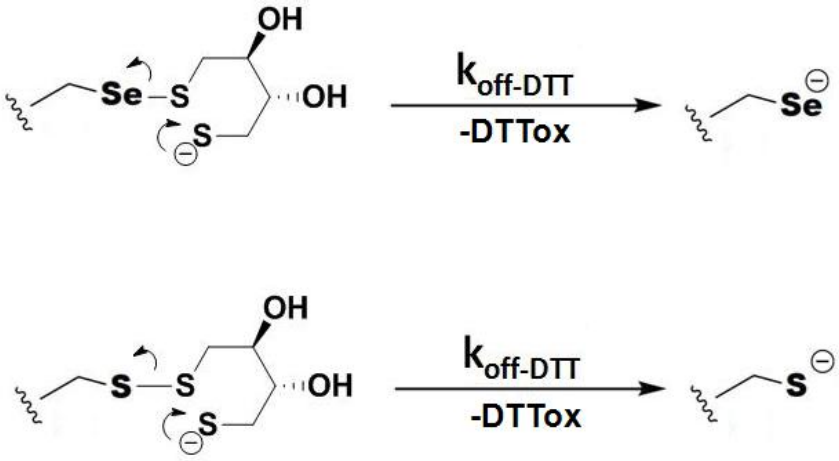


Figure 3.10: Comparison between the rate of selenolate and thiolate attack onto the disulfide of DTNB.

The rate constant ($k_{on-DTNB}$) obtained for the reaction of 117SecSECIS$_1$WT$_6$ with DTNB at pH 8.0 was (2.5 $\pm$ 0.8) x 10^4 $M^{-1}s^{-1}$. The rate constant for the reaction of 117CysSECIS$_1$WT$_6$ with DTNB was (1.3 $\pm$ 0.3) x 10^2 $M^{-1}s^{-1}$ at pH 8.0 (see Figure 3.10). The rate constants were corrected for the pKa ie the amount of reactive deprotonated thiolate vs selenolate species as at pH 8 the ratio of selenolate to thiolate was 988/0.32 $\approx 10^3$. The rate constants were found to be 2.5 x 10^4 $M^{-1}s^{-1}$ and 5.4 x 10^2 $M^{-1}s^{-1}$ for 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ respectively (calculated from $k = k_{on-DTNB}(10^{pKa-pH}+1)$). It was concluded that although the reaction ($k_{on-DTNB}$) between the 117Sec residue and DTNB is almost 200 times faster than the rate reaction between 117Cys and DTNB, after correction for the pKa of the selenol and thiol groups, the reaction of the selenolate is 46 $\pm$ 17 times faster than reaction of the thiolate with DTNB. The rate constants for the reaction of cysteamine (Mezyk, 1995, pKa = 8.6) and selenocysteamine (Arnold et al., 1986, pKa = 5.5) with cystamine, at pH 7.43 and pH 7.00 are 3.6 $\pm$ 0.1 $M^{-1}s^{-1}$ and (1.4 $\pm$ 0.3) x 10^3 $M^{-1}s^{-1}$ respectively (Steinmann et al., 2010). The ratio of thiolate/thiol and seleno-

late/selenol for cysteamine and selenocysteamine at pH 7 are 0.025 and 31, meaning that there exists 1.2 x 10^3 times more selenolate. After correcting for the pKa and pH, the rate of thiolate attack onto the disulfide of cystamine compared to attack of the analogous selenolate, was found to be approximately 25 times slower, further supporting the finding that the nucleophilicity of the selenolate group is enhanced by less than 50 decades compared to that of the thiolate group.

The leaving group ability of the selenolate compared to thiolate anion



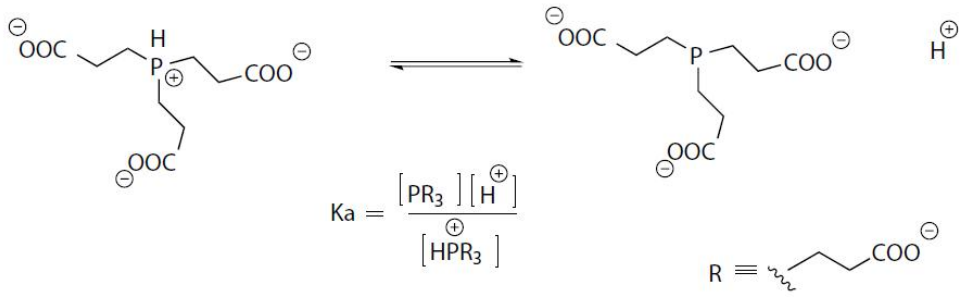
Scheme 3.5: The rate at which the PSe-DTT and PS-DTT intermediates returned back to PSe^- and PS^- respectively, were compared.

Scheme 3.5 illustrates the comparison between the rate of cyclisation of DTT to DTTox after formation of PS-DTT and PSe-DTT respectively. At pH 8.0 the rate constants for this reaction were 29 $\pm$ 5 s^{-1} (PSe-DTT, $k_{off-DTT}$) and 28 $\pm$ 4 s^{-1} (PS-DTT, $k_{off-DTT}$). It was concluded that at pH 8.0, the leaving group ability of the selenolate ion was comparable to that of the thiolate.

Comparison between the electrophilicity of the selenium and sulfur atoms of a selenylsulfide bond

The rate constants ($k_{off-Sec}$ and $k_{off-Cys}$) for the reduction of the selenylsulfide or disulfide bonds formed between 117Cys and selenocysteine or cysteine respectively were found to be (1.2 $\pm$ 0.8) x 10^4 $M^{-1}s^{-1}$ and 1.2 $\pm$ 0.3 $M^{-1}s^{-1}$ at pH 9 as the only difference between the reduction of PS-Cys and PS-Sec using TCEP was the presence of a selenium atom. The difference in magnitude of 10^4 between reduction of PS-Sec and PS-Cys at pH 9, by TCEP was attributed to the increased electrophilicity of the selenium atom (see Scheme 3.7). Stopped-flow spectrophotometric experiments were used by Steinmann et al. (2010) to investigate disulfide and selenylsulfide exchange reactions at pH 7. Scheme 3.8 compares the rate constants (obtained by Steinmann et al. (2010)) for the attack of a thiolate species onto a disulfide bond, with attack of the same thiolate onto an analogous selenylsulfide bond. Scheme 3.8 additionally compares the rate of reduction (by a thiolate species) of a diselenide bond and analogous selenylsulfide bond.

3.3.3 The reduction of selenylsulfide and disulfide bonds by TCEP



Scheme 3.6: The protonation- deprotonation equilibria for the phosphine group of TCEP, determined to be 7.66 (Burns et al., 1991)

TCEP is a water soluble trialkylphosphine reagent capable of reducing disulfides. TCEP and DTT are commonly employed for the reduction of disulfide bonds in proteins and peptides during bio-

chemical protein purification or cysteine modification reactions *in vivo*. The use of TCEP as a disulfide reducing agent was introduced by Burns et al. (1991) after the development of a convenient synthetic procedure. An estimation for the pKa of the phosphine group of TCEP, obtained from titration experiments, is 7.66 (Burns et al., 1991) (see Scheme 3.6). This pKa value has more recently been re-evaluated via stop flow experiments and ^{31}P NMR using a small peptide containing a disulfide bond and also found to be 7.6 (Krezel et al., 2003; Cline et al., 2004). Protonation of the phosphine of TCEP was considered in a similar way to the 117Cys thiol/thiolate equilibrium, where the relationship between $k_{off-Sec}$ and pH was sigmoidal. The pKa value for the protonation-deprotonation equilibrium of the phosphorous of TCEP was obtained from the electrical recording experiments (see Figure 3.8) to be 7.2 $\pm$0.5, which is in agreement with the literature value of pKa = 7.6. There are various rate constants quoted for the reduction of disulfides by TCEP, ranging from small molecule to peptides to proteins (Cline et al., 2004). Some earlier studies by Burns et al. investigated the effect of disulfide bond strain on the reduction rate by TCEP, however this was qualitative, providing a reactivity trend (Burns et al., 1991; Singh & Whitesides, 1990) as opposed to absolute rate constants. Example rate constants for the reduction of cystine, DTTox, DTNB, a short peptide and thioredoxin, by TCEP are shown in Table 3.2.

Disulfide bond reduction carried out by TCEP in the β barrel of αHL was compared to these rate constants, where the pseudo-first order rate constant for PS-TNB (TNB; 2-nitro-5-thiobenzoic acid) reduction at pH 7 was (2.4 $\pm$ 0.4) x 10^2 $M^{-1}s^{-1}$. This rate is of the same order of magnitude as the rate quoted for the reduction of a peptide disulfide (see Table 3.2), however no rate constant for a mixed disulfide reduction by TCEP between a protein cysteine residue, was found in the literature and so a direct comparison of this rate constant with existing data was not possible. It was hypothesised that the main factors that would affect the rate of disulfide bond reduction by TCEP, within the αHL pore, would be steric hinderance from surrounding residues on the β barrel surface and the increased leaving group ability of TNB compared to cysteine.

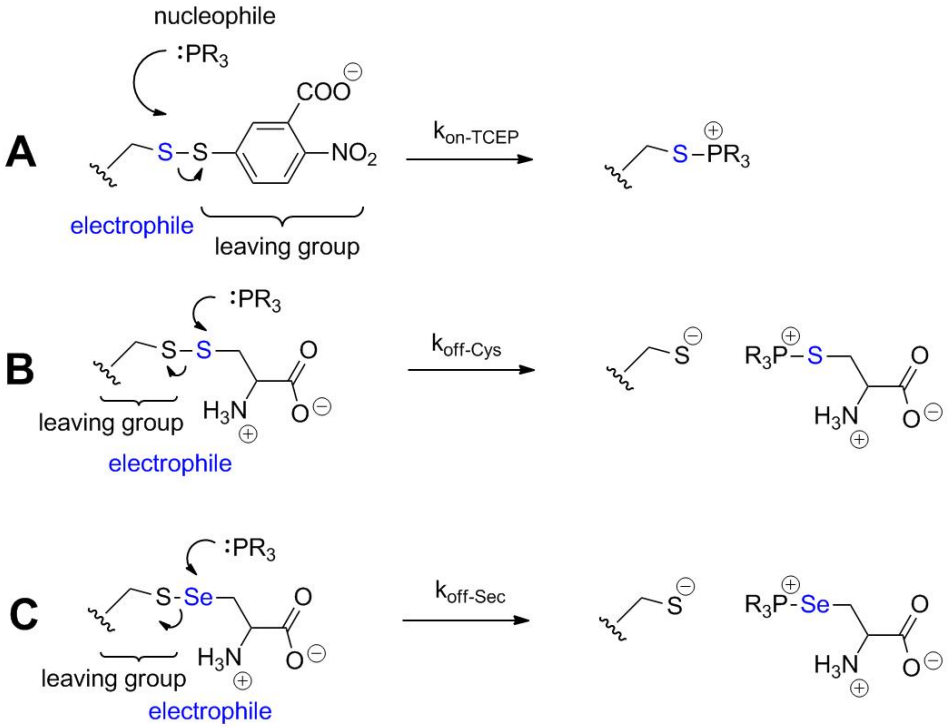
	k_{off} ($M^{-1}s^{-1}$)
[1]Cystine	2.22 ± 0.03
[2]$Peptide_{ox}$	6.5 x 10^2
[2]$Thioredoxin_{ox}$ (Trx_{ox})	2.4
[2]DTTox	4.3 x 10^1
[2]DTNB	1.8 x 10^4

Table 3.2: [1]The second order rate constant, obtained from bulk solution experiments, for the reduction of cystine in a 50% dioxane-H_2O (0.05 M HCl, 0.1 M KCl, 0.05 M EDTA) by TCEP (Overman & O'Connor, 1976). [2]Rate constants obtained at pH 7.5 under aqueous conditions (100 mM potassium phosphate, 0.3 mM EDTA) (Cline et al., 2004) for the reduction of disulfide bonds by TCEP. The peptide sequence was reported to be H_3N^+-VTWCGACKM-NH_2.

pH	[1]$K_{off\text{-}Sec}$ ($M^{-1}s^{-1}$)	[2]$K_{off\text{-}Cys}$ ($M^{-1}s^{-1}$)	[3]$K_{on\text{-}TCEP}$ ($M^{-1}s^{-1}$)
9	(1.2 ± 0.8) x 10^4	1.2 ± 0.3	(2 ± 0.6) x 10^2

Table 3.3: [1]The rate constant for the reaction between PS-Sec and TCEP at pH 9. [2]The rate constant for the reaction between PS-Cys and TCEP at pH 9. [3]The rate constant for the reaction between PS-TNB and TCEP at pH 9.

The rate of attack ($k_{on-TCEP}$) of TCEP onto the disulfide bond between 117Cys and cysteine (i.e. PS-Cys, see Table 3.3) was compared to the rate of attack of TCEP onto the disulfide bond between 117Cys and TNB, at pH 9 (i.e. PS-TNB, see Table 3.3). As the only difference between the PS-Cys and PS-TNB species was the Cys and TNB moieties respectively, the increased leaving group ability of TNB was proposed as the explanation for a 10^2 times faster reduction rate of the PS-TNB by TCEP compared to reduction of PS-Cys.



Scheme 3.7: **A** Reduction of the PS-TNB bond by TCEP. **B** Reduction of the PS-Cys disulfide bond by TCEP. **C** Reduction of the PS-Sec selenylsulfide bond by TCEP.

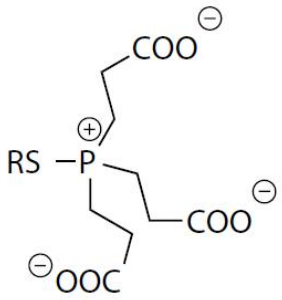
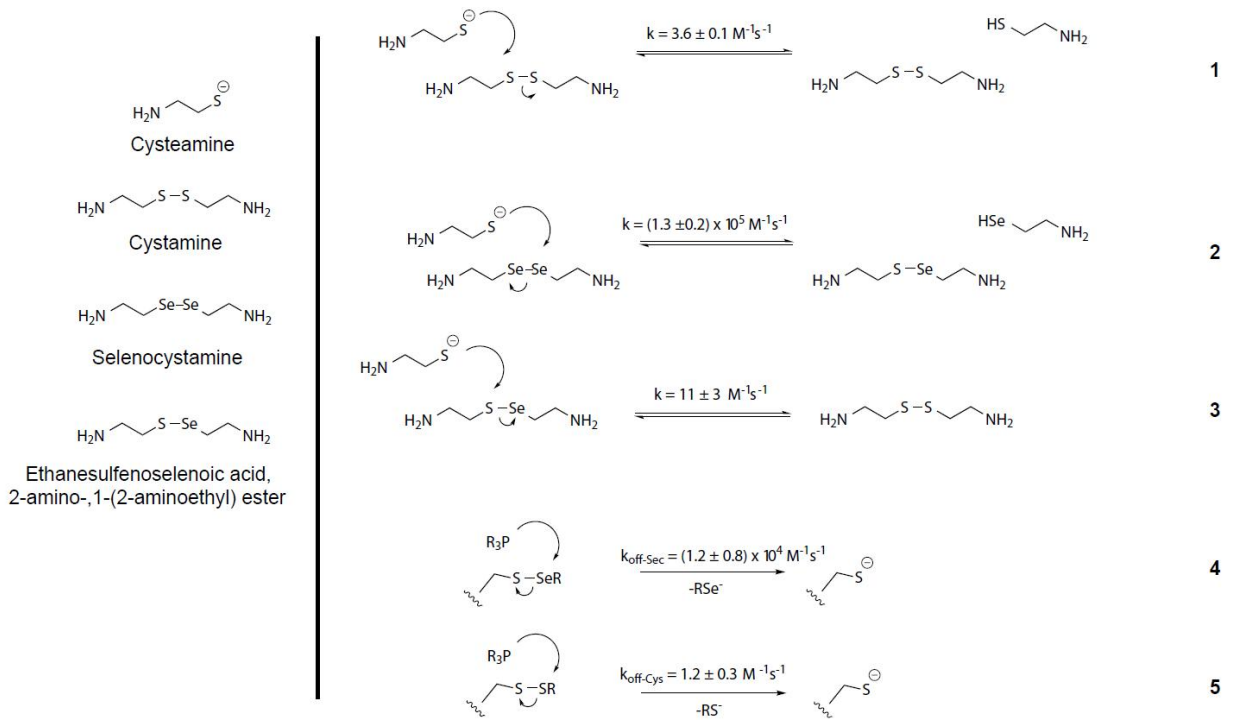


Figure 3.11: Phosphonium cationic intermediate formed between TCEP and the single cysteine residue at position 117 within the β barrel of αHL.

3.3.4 The phosphonium cationic intermediate

The short lived intermediate, proposed to be the previously hypothesised phosphonium cation (see Figure 3.11), was observed after almost every reduction event of PS-TNB. Very occasionally (i.e.

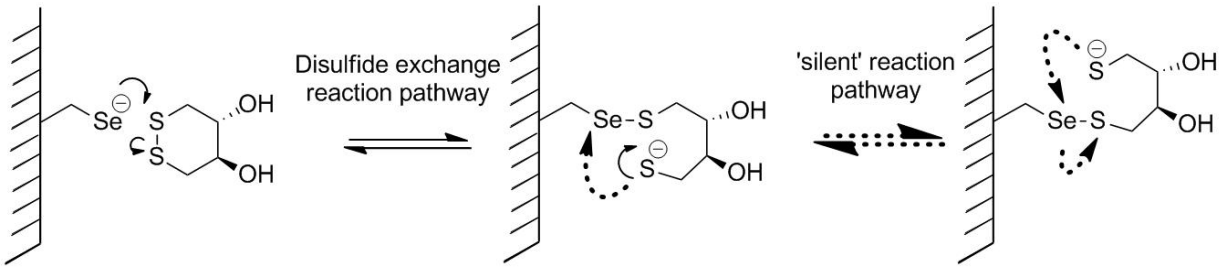


Scheme 3.8: Disulfide and selenylsulfide exchange rate constants measured at pH 7. The rate constants for attack of the cysteamine thiolate on the disulfide of cystamine (3.6 ± 0.1 $M^{-1}s^{-1}$, **1**) and the selenylsulfide of selenocystamine ((1.3 ± 0.2) x 10^5, **2**) illustrate an increase in rate of 3.6 x 10^4 $M^{-1}s^{-1}$. The rate constants for attack of the cysteamine thiolate on the disulfide of cystamine (3.6 ± 0.1 $M^{-1}s^{-1}$, **1**) was also compared to the rate of attack at the selenium atom of ethanesulfenoselenoic acid, 2-amino-,1-(2-aminoethyl) ester (11 ± 3 $M^{-1}s^{-1}$, (**3**)). Rate constants for **1** - **3** taken from Steinmann et al. (2010). Rate constants (measured at pH 9) for the reduction (by TCEP) of a selenylsulfide ((1.2 ± 0.8) x 10^4 $M^{-1}s^{-1}$, **4**) and disulfide (1.2 ± 0.3 $M^{-1}s^{-1}$, **5**) bond formed within the β barrel of $117CysSECIS_1WT_6$. The rate of attack of the phosphine of TCEP onto the selenium atom compared to sulfur atom was found to be almost 10^4 times faster and attributed to the increased electrophilicity of the selenium atom compared to the sulfur atom.

twice per 1 hour of recording) this intermediate was observed when reducing PS-Cys, and never, on reduction of PS-Sec. It was thought that the phosphorus of TCEP would preferentially attack the least sterically hindered side of the disulfide bond. No known literature rate constants were found regarding the life time of the phosphonium cation intermediate (ie PS-$P^+(CH_2CH_2COO^-)_3$), and a value of 7 ± 1 s^{-1} was obtained assuming the reaction to be unimolecular and this intermediate showed no pH dependance (146 ± 22 ms) over the pH range 6 - 9.

3.3.5 Conclusion

The 117CysSECIS$_1$WT$_6$ pore was used to investigate the disulfide exchange reaction between the 117Cys residue and DTTox. These experiments allowed the calculation of a forward (k_{on}) and reverse (k_{off}) rate constant at pH 8.5. The two diastereomeric compounds formed on reaction of DTTox with the 117Cys residue were observable, with a 0.3 pA difference in amplitude between the current blockage levels. From counting the number of events recorded due to each diastereomer, the ratio of diastereomeric species was calculated to be 3:7, although it was not possible to assign which structure corresponds to which PS-DTT covalent adduct. Future experiments would involve investigating the reaction of 117SecSECIS$_1$WT$_6$ with DTTox. From these experiments it is expected that the diastereomeric intermediates (ie PSe-DTT) would be observable. Another interesting question would be whether the intramolecular reaction, or "silent reaction" (see 3.9) would be fast enough to dominate the kinetics so as to observe longer blockades than observed for ring closure to form the DTTox species.



Scheme 3.9: The protonation- deprotonation equilibria for the phosphine group of TCEP.

The reduction of mixed disulfides and selenylsulfide bonds by TCEP was studied using the 117Cys-SECIS$_1$WT$_6$ αHL pore. Rate constants over a pH range of 6 - 10 were obtained for the reaction of selenocystine (Sec_2) with the 117Cys residue and similarly for reduction of PS-TNB. The rate constant for the reduction of PS-Cys was obtained at pH 9. The lifetime of the phosphonium cation intermediate formed on attack of the phosphine of TCEP onto the disulfide bond of PS-TNB, was analysed over a pH range 7 - 9 and found not to vary significantly with pH. The nucleophilic-

ity of the selenocysteine and cysteine residues within the β barrel of the 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ pores, was compared at pH 8.0. The rate constants for the reaction of 117Sec and 117Cys with DTNB were found to differ by 46 $\pm$ 17 orders of magnitude, after correction of the pKa difference between the selenolate and thiolate groups respectively. The leaving group ability of the selenolate and thiolate groups were found to be comparable at pH 8.0. The electrophilicity of the selenium atom of the selenylsulfide bond formed between the 117Cys residue and selenocysteine (PS-Sec), within the 117CysSECIS$_1$WT$_6$ pore, was 10^4 times higher than that of an analogous sulfur atom of the disulfide bond between the 117Cys residue and cysteine (PS-Cys).

3.4 Materials and Methods

3.4.1 Electrophysiology set up

The general principle of single channel electrical recordings was to measure the current flow through an αHL pore and monitor the reaction of a cysteine or selenocysteine residue situated within the β barrel, with analytes in solution, via changes in the current flow. From analysing the amplitude of the current changes and the duration of each amplitude, it was possible to obtain kinetic rate constants. This section provides details regarding how the electrophysiology experiments were performed and also how the data were recorded, stored and analysed.

3.4.2 Chamber

The chamber used for the electrical recordings consisted of two compartments, denoted cis and trans and was machined from Delrin. Each compartment was filled with 1 mL of buffer solution, separated by a 25 μl thick piece of teflon (polytetrafluoroethylene, PTFE, Goodfellow, Cat. #

FP301200/10). An aperture of 100 μm diameter was formed in the teflon using a high-voltage spark generator. The teflon was first indented with a pin and the spark produced by the generator would target this position creating the aperture/ hole. Across this aperture, a bilayer was formed and a single αHL pore inserted. The chamber was cleaned between experiments by firstly washing with water and then 100% ethanol and dried under a stream of nitrogen. An additional cleaning step was carried out prior to experiments each day, by soaking the compartments in 3 M KOH solution.

3.4.3 Electrodes

Ag/AgCl electrodes were used to apply the voltage across the single channel set up. One electrode functioned as the ground (situated in the "cis" chamber) and the other as the active electrode (situated in the "trans" chamber). The electrodes consisted of a piece of Ag wire (1.5mm diameter, Sigma-Aldrich). After abrading the surface of the wire with glass paper, the wire was submerged in 50% hypochlorite solution (NaClO, Fischer Scientific) and left to coat (AgCl coating) overnight. During electrical recordings, "agarose bridges" were used to separate the AgCl electrodes from the buffer solution in the cis and trans chambers. These bridges were made by dissolving 3% (w/v) low-melt agarose (Sigma-Aldrich) in 3 M KCl (Sigma-Aldrich). The bridges were prepared by melting the KCl/agarose solution and pipetting it into 200 μl pipette tips (Starlab Ltd.). Before the KCl/agarose solution solidified, the AgCl electrode was inserted into the top of the KCl/agarose-filled pipette tip and secured in position by utilising a rubber stopper through which the electrode was placed. The agarose bridges were used during all experiments so as to prevent Ag^+ ions from the Ag/Cl electrodes leaching into the buffer filled chambers during electrical recordings and binding to the cysteine or selenocysteine residues.

3.4.4 Bilayer formation

The aperture was firstly coated with hexadecane (10% (v/v) hexadecane in pentane) and the pentane evaporated leaving white crystalline hexadecane visible around the aperture. Buffer (typically Bis Tris Propane 20 mM, 2 M KCl, 100 μM EDTA, pH 6-10, 1 mL) was added to each compartment of the chamber and approximately 2 drops of 1,2-diphytanoyl-sn-glycero-3-phosphocholine solution (DPhPC, Avanti Polar Lipids, 10% (w/v) in pentane) were added onto the surface of the buffer filled compartments to form a monolayer of lipid. After allowing the pentane to evaporate (approx. 30 s) the solution in each chamber was pipetted up and down slowly to form a bilayer across the aperture (Montal & Mueller, 1972). Bilayer formation was monitored via the current output signal of the electrodes. The Ag/AgCl electrodes were placed into each of the compartments. When no bilayer existed ions were able to flow through the aperture but the bilayer acted as a capacitor and so electrically separated the cis (ground electrode side) and trans (live electrode side) compartments.

3.4.5 Bilayer capacitance

The capacitance of the bilayer was primarily used to assess the quality of the bilayer formed. The relationship between current (I) and capacitance (C, units are Farads) is given by $I = C\frac{\delta V}{\delta t}$.

As the current, $I \propto dV/dt$, the application of a triangle wave voltage (ie constant gradient, dV/dt) across a bilayer, generates a square-wave current output. The triangle wave was generated via the patch-clamp amplifier and was calibrated using a 100pF model bilayer cell (MBC-1U, Molecular Devices) so that 100 pA = 100 pF. It was the magnitude of the square-wave that was used to gauge the quality of the bilayer in terms of stability (Maglia et al., 2010). The bilayer was subjected to high voltages, typically $\pm$250mV as an additional verification of stability. For a 100 μm diameter aperture, a capacitance of 80 - 100 pF was expected.

3.4.6 Pore insertion

On formation of the bilayer, 0.5-1 μl of aliquoted protein ($\tilde{1}$ ng in total) was added to the cis chamber, close to the teflon. A potential of between +150-300mV was then applied and the solution stirred with a stirrer flea (Fisher). On insertion of a pore at +50 mV for example, the current level would jump from 0 pA to $\tilde{8}$0 pA (or 1.6 nS). Replacement of the buffer solution in the cis chamber with new (non-protein containing) buffer solution and was carried out for at least three chamber volumes so as to reduce the probability of a second pore insertion and also to remove any contamination from the protein storage buffer. For example, the cysteine and selenocysteine containing heteroheptameric αHL pores were always aliquoted in buffer containing 5mM DTT.

3.4.7 Data acquisition

Data were acquired by firstly applying an external voltage of +50 mV, the resulting signal was firstly amplified, filtered with a low-pass Bessel filter before being digitized at a chosen sampling frequency (usually 5-10 kHz). These transmembrane current data were then stored on a computer ready for further analysis. Post-acquisition analysis was carried out using the program Clampfit 10.2 (Axon Instruments), which allowed further Bessel filtering. The same program was also used to select events based on current amplitude changes, where dwell-time (event length) histograms were created so as to obtain the average dwell time of a particular current level. The following section describes in further detail, each part of the electrical recording set up.

3.4.8 Electrical recording set up

An Axopatch 200B patch-clamp amplifier (Molecular Devices) was used for all electrical recordings. The Axopatch 200B amplifier used cooling elements in the headstage so as to reduce electrical noise to <15fA and was used to transform the analogue current measured, into an analogue

voltage signal with a recording bandwidth of up to 100kHz. The output gain of the amplifier works by converting every 1 pA into a corresponding amplified voltage dependent on the gain chosen. For example, an output gain of 1X would convert 1 pA to 1 mV (ie 1 nA to 1 V), an output gain of 2X would convert every, 1 pA to 2 mV and so on. From the amplifier, the signal continued into the 1440A 16-bit digitizer (Molecular Devices). The digitizer converted the analogue signal into 2^{16} (ie 16-bit) discrete current values (bins). The dynamic input range or "voltage resolution" of the digitizer was $\pm$10 V. Being a 16-bit digitizer, this corresponded to a resolution of 20,000/(2^{16}) = $\pm$0.305 mV, per bin. So as the amplifier converts 1 nA to 1 V when the output gain of the amplifier is set to 1X, this corresponds to a dynamic range (in terms of the current signal) of $\pm$10 nA. The output gain was therefore important when resolving small current amplitude alterations and was typically set to 10X or 20X throughout the electrical recordings. At a 10X output gain, this means that 1 nA was converted to 10 V and so the input range was $\pm$1 nA with a resolution of $\pm$0.0305 mV respectively. Before digitisation, the current signal was filtered using the amplifiers in-built low-pass, 4-pole Bessel filter. The cut-off frequency (ie signals above this cut-off are filtered out) was set to 5 or 10 kHz in accordance with the Nyquist theorem which states that data should be sampled at a frequency equal to twice the cut-off frequency of the of the applied filter (The Axon Guide).

The software used to control the digitizer and apply the voltage was the Clampex program of the pClamp software v10.1 (Molecular Devices). Apart from when gathering I-V curve data, the "gap-free" mode was used, which applied a constant voltage signal from the digitizer through the amplifier, via the headstage, to the electrodes. During the I-V data gathering, the "episodic stimulation" mode was used. This mode applied voltage steps of both user defined, magnitude and duration.

3.4.9 Data processing

Data analysis was carried out using Clampfit (pClamp software, Molecular Devices). The Clampfit program allowed further low-pass Bessel (8-pole) filtering, typically 100Hz. The Bessel filter was chosen due to its ability to maintain the shape of the original signal (pClamp10 User guide, Molecular Devices). The amplitude of events was obtained by gaussian fitting as a histogram of constant bin width, where each bin contained values less than or equal to the upper limit of each bin. Each event recorded, signified the reaction of a single selenocysteine or cysteine residue within the β barrel of αHL with a reagent present in the *cis* or *trans* compartment. The Clampfit 'Single channel search' mode was used to obtain information of both the amplitude and dwell time of each event. Typically between 80 - 200 events were gathered per electrical recording so as to calculate an average dwell time. Log interval binning was used to create a histogram of event dwell times, where a bin size of 10 (bins/decade) was used. After obtaining the average dwell time (τ) from the histogram, the reciprocal of this value was used to calculate the concentration dependent rate constant (i.e. $1/\tau$ = k). A concentration dependent rate constant was obtained by addition of the reagent to either the *cis* or *trans* compartment during the electrical recording and calculating the average dwell time of a particular event level at each concentration. By graphically plotting these concentration dependent rate constants against the concentration of the reagent present in the *cis* or *trans* compartments, it was possible to elucidate the dependence of the rate of reaction on the concentration of a particular reagent. By gathering electrical recording data over at least 3 different reagent concentrations, it was possible to see the order of the reaction with respect to the reagent. For example zero order dependence in a certain reagent would produce rate constants of the same magnitude and be concentration independent. In this case a straight line fit was plotted on the graph in the form $1/\tau = \mathrm{k\ s^{-1}}$. A first order dependence in the reagent concentration produced concentration dependent rate constants which were directly proportional to the reagent concentration. For example, $1/\tau$ = k[A] where [A] is the concentration of species 'A' present in the *cis* or *trans*

compartment. On plotting the concentration dependent rate constant ($1/\tau$) against [A] calculated over at least 3 different concentrations, a straight line fit was used to obtain the overall rate constant in units of $s^{-}1M^{-}1$. A minimum of 3 repeats (n = 3) were obtained to calculate each rate constant unless otherwise stated, and the standard deviation between these three repeats was given as the error associated with each rate constant.

3.4.10 Boc deprotection of Cys(Npys)-OH

N-Boc-*S*-(3-nitro-2-pyridylthio)-L-cysteine (100 mg, 0.266 mmol) was dissolved in dichloromethane (2 mL) and 50% aqueous TFA (4 mL) was added. The reaction was stirred, under Ar(g) overnight at room temperature. After concentration under vacuum, toluene was used to co-evaporate the TFA. The residue was then re-dissolved in acetonitrile and precipitated using diethyl ether. The pale yellow precipitate was then collected and dried under vacuum (63 mg, 0.23 mmol). TLC (Petroleum ether:ethyl acetate, 1:3 with MeOH, 0.5 mL in 4 mL (and 5 drops of glacial acetic acid)), R_f = 0.3, *m/z* (TOF-ES) 276 $(M + H)^+$, ^{1}H NMR δ_H(400 MHz, $(CD_3)_2CO$/ DCl) 3.66 (2H, s, CH_2CS), 4.43 (1H, m, CHN), 7.92 (1H, s, C=CH), 8.95 (1H, d, *J* 7.74 Hz, HC=N), 9.19 (1H, s, $HC=CNO_2$), 9.32 (4H, m, N^+H_3 and COOH). Method taken from Singh et al. (2006).

3.4.11 Solution preparation

The buffer used during all of the electrical recordings consisted of Bis-Tris propane (20 mM, 1,3-Bis[tris(hydroxymethyl)methylamino]propane, Sigma), KCl (2 M, potassium chloride, Sigma) and EDTA (2 M, 2-(2-[bis(carboxymethyl)amino]ethyl(carboxymethyl)amino)acetic acid, Sigma). The pH of the solution was adjusted between the range of 7 - 10. The buffer used for experiments carried out at pH 6 was a KOAc buffer (10 mM KOac, 2 M KCl, 100 μM EDTA).

DTTox (*trans*-4,5-Dihydroxy-1,2-dithiane, Sigma) was dissolved in water/acetonitrile (7:3 v/v)

to a final concentration of 0.8 M and stored at room temperature. A new aliquot was prepared each day. During the titration experiments, DTTox was pipetted into the *trans* chamber to final concentrations ranging from 10 - 120 mM.

TCEP (Tris(2-carboxyethyl)phosphine hydrochloride, Sigma) was dissolved in water to a final concentration of 0.5 M and the pH was adjusted between the range of 6 - 10 using KOH (aq) to match that of the buffer. Fresh aliquots were prepared daily. TCEP was always pipetted into the *cis* chamber to give final concentrations ranging from 0.2 mM - 40 mM. At concentrations higher than 20 mM the bilayer remained stable for 30 minutes maximum.

DTNB (5,5'-Dithiobis(2-nitrobenzoic acid), Sigma) was dissolved in a sodium phosphate buffer (500 mM, pH 8.5) to a final concentration of 50 mM. Fresh aliquots were prepared daily. DTNB was always added into the *trans* chamber to final concentrations between 500 μM - 10 mM.

DTT (*DL*-Dithiothreitol, Sigma) was dissolved in water to a final concentration of 100 mM. Aliquots were replaced each day. DTT was added into the *cis* chamber to final concentrations in the range of 50 - 400 μM.

Cystine ((*S,S*)-3,3'-Dithiobis(2-aminopropionic acid), Sigma) was dissolved in Bis-Tris propane buffer, pH 8.5 (same as above) to a final concentration of 0.5 M. Cystine was added to the *trans* chamber to a final concentration ranging from 20 - 80 mM.

Selenocystine ((*S,S*)-3,3'-Diselenoobis(2-aminopropionic acid), Sigma) was dissolved in 0.05 M KOH (aq) solution, to a final concentration of 0.5 M. Selenocystine was added to the *trans* chamber to a final concentration ranging from 100 - 800 μM. Fresh solutions were prepared each day.

The **Cys(Npys) – OH** ((3-nitro-2-pyridylthio)-L-cysteine) solution was prepared to a final concentration of 100 mM by dissolving in acetonitrile/buffer. Cys(Npys)-OH was added the the *trans* chamber to a final concentration ranging from 5 - 20 mM at pH 9.

3.4.12 DTTox data analysis to obtain the kinetic rate constants

Titration experiments were carried out by pipetting the stock solutions (see above) into either the *cis* or *trans* chamber.

The events observed for the bimolecular reaction between DTTox and the 117Cys residue situated within the αHL β barrel were analysed using Clampfit (pClamp software, Molecular Devices). The raw data collected using a sampling frequency of 20 kHz and a filter of 5 Hz, was further filtered (lowpass Bessel filter) at 400 Hz. Using the "Single channel search" option in Clampfit, the three (or sometimes two, see Results) current levels were assigned (PS or open pore and PS-DTT) and the event durations (i.e. τ_{on} and τ_{off}) tabulated. The event dwell times for each were plotted and fitted using the Log probability function of Clampfit (bin size 10 bins/decade). As the reaction between DTTox and 117Cys is bimolecular, the τ_{on} = 1/k_{on-DTT}[DTTox] and so data were processed as described above at three (or four) different concentrations of DTTox, (for example 20, 40 and 80 mM DTTox with each recording, approximately 40 minutes in length). By plotting $1/\tau_{on}$ against concentration (see Results, Figure 3.4), the k_{on-DTT} was obtained. As the $k_{off-DTT}$ for the PS-DTT intermediate was independent of the concentration of DTTox, the dwell time remained constant (see Figure 3.4). The current amplitude of the three event levels was analysed using the "Single channel search" mode in Clampfit. To verify whether or not the two diastereomeric PS-DTT intermediates were distinguishable, a histogram of the events was plotted (see Results, Figure 3.3, **B**).

3.4.13 TCEP reduction data analysis

By spatially separating the reducing agent, TCEP from the disulfide reagent (i.e. DTNB, Selenocystine or Cystine), it was possible to observe multiple turnovers of the reduction of the PS-SR disulfide by TCEP. Data analysis was carried out using Clamfit and the "Single channel search"

option as explained above. Raw data were acquired at a sampling rate of 10 kHz with a filter of 2 Hz (and 5x output gain). Data were further filtered (200 Hz, Bessel 8-pole) before event dwell time and amplitude data were collected. Both the reaction of the disulfide reagent with the 117Cys residue within the β barrel of αHL and the reduction of the new disulfide (PS-SR) by TCEP, are separate bimolecular reactions and so the kinetic analysis employed to calculate the rate constants k_{on-SR}. Figure A.12 and Figure A.13 illustrate titration experiments carried out for the reduction of DTNB, and selenocystine by TCEP respectively. Titration experiments were carried out either by keeping the TCEP concentration constant and sequentially increasing the DTNB or selenocystine concentration or vice versa.

The pKa of the 117Cys residue was obtained by fitting to

$$k_{on} = \frac{k_{on-max} K_a}{(K_a + [H^+])} \tag{3.1}$$

By plotting the $k_{on-TCEP}$ against pH it was found that the rate of TCEP attack at the selenylsulfide bond was dependent on the protonation equilibrium of the phosphine group of TCEP. Using the same equation as above, the pKa of TCEP was obtained from the fit of $k_{on-TCEP}$ against pH.

Chapter 4

Reactions of selenocysteine and cysteine with an organoarsenic (III) compound

4.1 Introduction

4.1.1 Arsenic *in vivo*

The toxicity of arsenic

Arsenic is toxic in humans (Mandal & Kazuo, 2002) and most commonly enters organisms via oral consumption, for example contaminated water or food. Two inorganic arsenic-containing compounds are arsenate ($HAs^{(V)}O_4^-$) and arsenite ($As^{(III)}(OH)_3$) and a broad range of foods contain these arsenic species, although the ratios of inorganic and organic arsenic species and the absolute levels of each, vary. In mammals, arsenic compounds are transported via the blood stream, for example $As(OH)_3$ enters cells via aqua-glycoporins which are usually used for transportation of glycerol, urea and other small neutral molecules. The organoarsenic compounds are often described as being less toxic than inorganic arsenic species, however methylarsenic acid (MMA) is

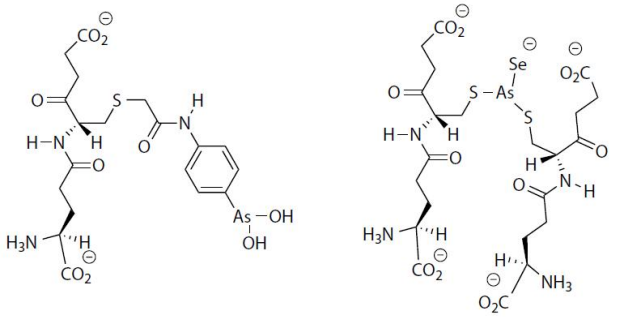


Figure 4.1: The chemical structures of *p*-4-[N-(S-glutathionylacetyl)amino]phenylarsenoxide (*p*-GSAO) and seleno-bis(S-glutathionyl)arsinium ($[(GS)_2AsSe]^-$).

now thought to be more toxic than $As(OH)_3$ (Mandal & Kazuo, 2001). Arsenic is known to detrimentally affect the respiratory tract, cardiovascular system, hematopoitetic system, nervous system and can cause skin cancers and can integrate into the phosphate pathway in its $As^{(V)}$ oxidised form. Arsenic contamination and subsequent poisoning is a world wide issue and so an understanding of arsenic metabolism and its physiological action *in vivo*, is essential. The toxicity of arsenic towards organisms has been exploited in the creation of arsenic containing drugs. Arsenite ($As(OH)_3$) is currently administered for acute promyelocytic leukemia, with 86% remission success. Derivatives of phenylarsenoxide, for example *p*-4-[N-(S-glutathionylacetyl)amino]phenylarsenoxide (p-GSAO, see Figure 4.1) being investigated as potential treatments for breast cancer and leukemia (Ralph, 2008).

The *in vivo* targets of arsenic (III) containing compounds

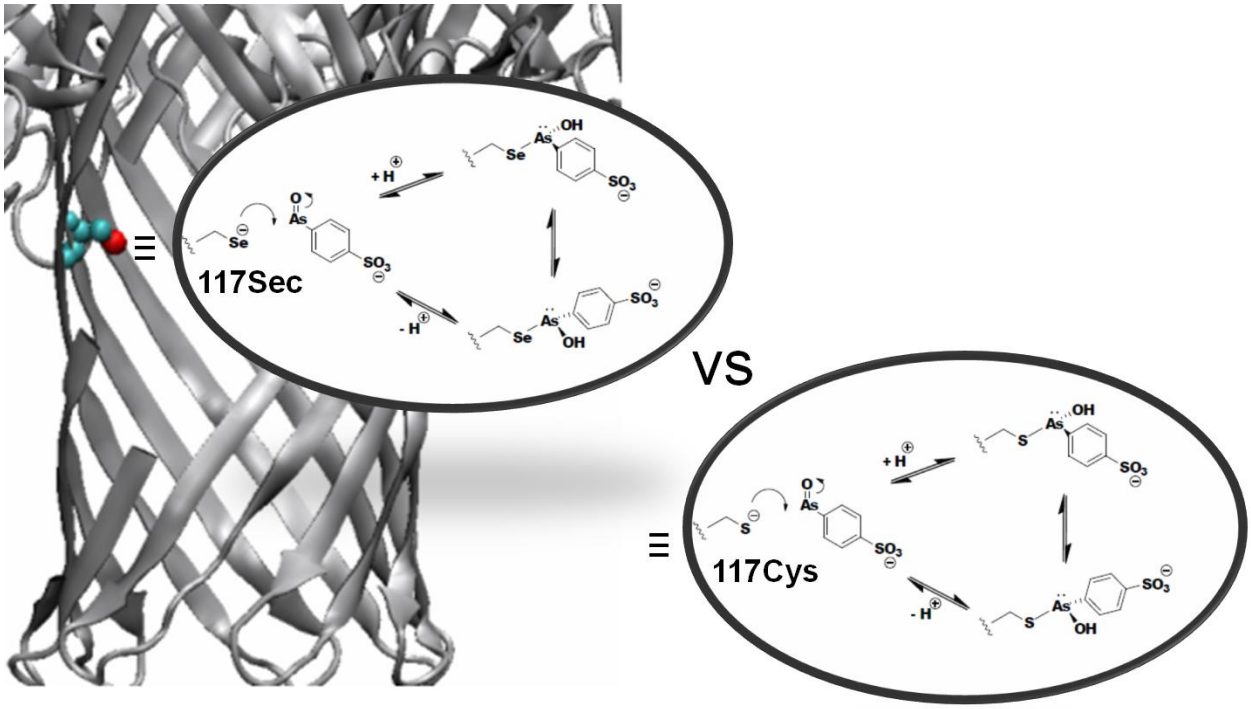
The toxicity of arsenic (III) compounds originates from the ability of arsenic to covalently attach to thiol groups within proteins. The pyruvate dehydrogenase (PD), multienzyme complex of *Escherichia Coli* responsible for the generation of acetyl-coA becomes inactive on exposure to *p*-aminophenylarsine oxide *in vitro* (Stevenson et al., 1978). The acetyl-coA generated by the PD complex is required for a range of functions within organisms, such as metabolism of car-

bohydrates, the synthesis of the neurotransmitter acetylcholine, the generation of citrate (and so indirectly affects ATP production by mitochondria) and various post-translational protein modifications. Therefore impairment of the PD multienzyme complex impacts on many signalling pathways. By utilising NADPH, glutathione reductase (GR) catalyses the reduction of oxidised glutathione and so when inhibited by $As(OH)_3$ and $CH_3As(OH)_3$, an imbalance in cellular levels of GSH arises (Chouchane & Snow, 2001). Glutathione is the most abundant cellular thiol and has many essential functions such as an involvement in the removal of toxic metals (Finney & O'Halloran, 2003), NO signalling (Gaston et al., 1993) and selenium transport (Fairweather-Tait & Hurst, 2010). Another important redox protein, targeted by arsenic containing compounds, is thioredoxin. Thioredoxin is responsible for the maintenance of the redox state of a range of proteins and is important for DNA synthesis (Mark & Richardson, 1976) and cell proliferation (Biaglow & Miller, 2005) and is itself maintained in its reduced form by thioredoxin reductase (TrxR). Arsenic (III) species are also known to bind and inhibit TrxR and so these essential cellular processes are disrupted. TrxR also reduces other small molecules such as selenite (Arnér & Holmgren, 2000), lipid peroxides (Bjornstedt et al., 1995), dehydroascorbate (May et al., 1997) and protein disulfide isomerases (Lundrom & Holmgren, 1990, PDIs) and so inhibition of this enzyme has knock-on effects for the pathways associated with these small molecules.

4.1.2 The interaction of arsenic with selenium

There currently exists no quantitative kinetic data on the reaction of arsenic compounds with small molecule organic selenols in solution. The interaction of arsenic (III) with selenite *in vivo* has been found to produce seleno-bis(S-glutathionyl) arsinium (Manley et al., 2006, see Figure 4.1) and the toxic effects of arsenic (V) species have been counter balanced by addition of selenite *in vivo* (Stolz et al., 2006). Arsenite is also proposed to inhibit the selenocysteine residue within TrxR (Arnér & Holmgren, 2000) and is known to form species such as $As(SG)_3$ in solution.

4.2 Results



Scheme 4.1: The reaction of the organoarsenic (III) compound, SPAO with the 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ pores respectively. The diasteromeric intermediates formed within the β barrel and the interconversion pathway between them, have been illustrated.

Using the 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ αHL pores, the reaction of the organoarsenic (III) reagent, SPAO with the selenocysteine and cysteine residues was investigated. Single channel electrical recordings were used to observe the reaction between SPAO and the single 117Sec residue or 117Cys residue within the β barrel of αHL of 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ respectively. The reaction scheme between SPAO and the 117Sec or 117Cys residues is illustrated in Scheme 4.1. Initial attack of the 117Sec/117Cys, forms a covalent selenoarsenic/thioarsenic bond, observed during the electrical recordings as a decrease in current of 4.5 $\pm$ 0.4 and 3.1 $\pm$ 0.2 (n = 3) and 4.7 $\pm$ 0.2 and 3.5 $\pm$ 0.1 (n = 3) respectively at an applied voltage of +50 mV. As depicted in Figure 4.2, the covalently attached SPAO within the β barrel, creates two blockades with different current amplitudes (see Figure 4.2). Shin et al. (2007) carried out NMR experiments

to investigate the diastereomeric adducts formed on reaction of SPAO and glutathione in solution, and by carrying out complementary single molecule electrical recordings, showed how these two different current amplitudes are due to the diastereomers (PS-SPAO) within the β barrel of αHL. From single channel recording data, it was not possible to assign a specific diastereomeric species to a particular current level. Diastereomeric adducts in the Figures have therefore been assigned arbitrarily. The reaction between SPAO and 117SecSECIS$_1$WT$_6$ or 117CysSECIS$_1$WT$_6$ was carried out over the pH range 6.5 - 8.5. A Bis-Tris propane buffer (20 mM Bis-Tris propane, 2 M KCl, 0.1 mM EDTA) was used for all experiments.

4.2.1 The pH dependence for the reaction between SPAO and 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$

After insertion of a single 117SecSECIS$_1$WT$_6$ or 117CysSECIS$_1$WT$_6$ pore into the planar DPhPC bilayer formed between the two chambers of the electrophysiology set-up (see Chapter 3, Materials and Methods for further set-up information), SPAO was added to the *trans* chamber. Rate constants for the forward reaction, that is, the attack of the 117Sec and 117Cys residues on the arsenic centre of SPAO were obtained at pH 6.5, 7.5, 8.0 and 8.5. Titration experiments were carried out by the addition of SPAO (*trans*) to a final concentration of between 100 μM and 16 mM. Rate constants were obtained from data recorded using at least three different concentrations of SPAO for each pH value. As the covalent diastereomeric adducts produced different current blockages, it was possible to obtain separate rate constants (k_{1-2} and k_{1-3}) for the formation of each adduct and these rate constants are denoted k_{1-2} and k_{1-3} (see Scheme 4.2).

Titration data for the reaction of SPAO with both 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ over the pH range 6.5 - 8.5 can be found in Figures A.15 and A.14 respectively. These graphs illustrate the relationship between the concentration of SPAO and $1/\tau$ where τ is the mean duration of a particular event level. It can be seen that the reaction of SPAO with either the 117SecSECIS$_1$WT$_6$

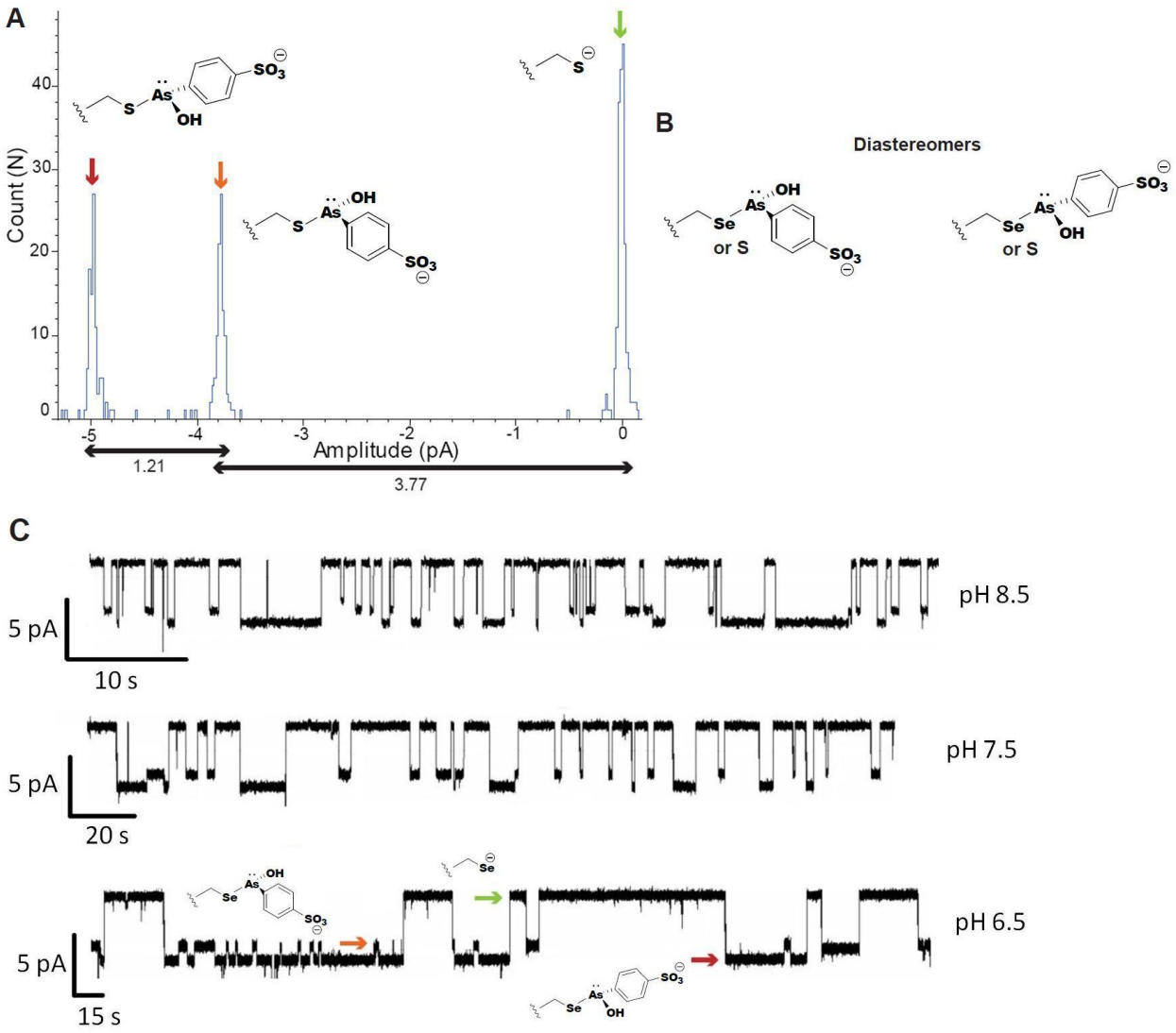
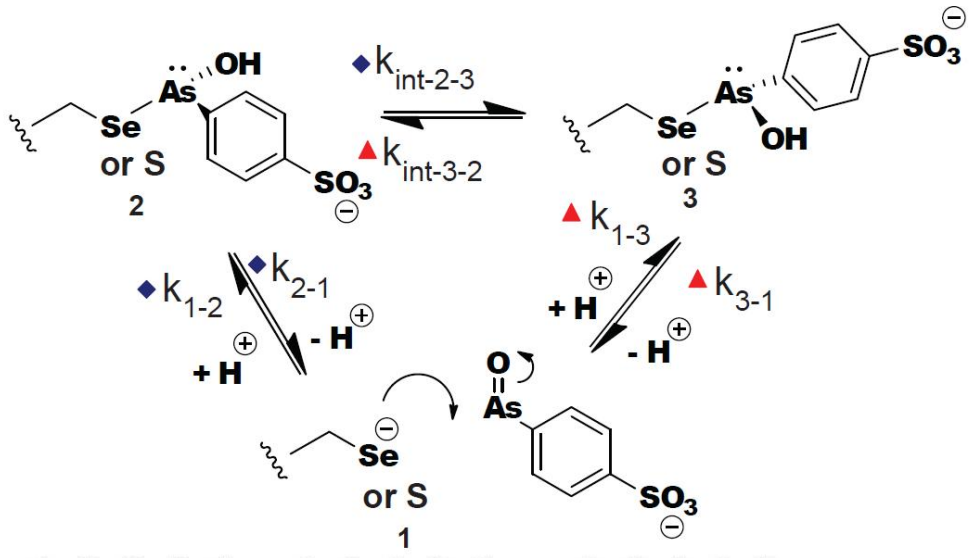


Figure 4.2: **A** The covalently attached organoarsenic (III) compound. The SPAO within the β barrel of 117CysSECIS$_1$WT$_6$ generates two current blockades of different amplitude. These blockades are due to the formation of diastereomeric adducts as illustrated above (although it was not possible to assign which orientation caused which blockage amplitude). Data obtained at pH 8.5 using a Bis-Tris propane buffer (20 mM Bis-Tris propane, 2 M KCl, 0.1 mM EDTA). **B** The diasteromers formed on reaction of SPAO with 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$. **C** The reaction between 117SecSECIS$_1$WT$_6$ and SPAO observed at the single molecule level. Single channel electrical recording data obtained at pH 8.5, 7.5 and 6.5 using a Bis-Tris propane buffer (20 mM Bis-Tris propane, 2 M KCl, 0.1 mM EDTA)



Scheme 4.2: Rate constants for the reaction between 117SecSECIS$_1$WT$_6$ or 117CysSECIS$_1$WT$_6$ with SPAO.

or 117CysSECIS$_1$WT$_6$ pores is pseudo-first order (i.e. d(PS-SPAO)/dt = k_{on}[SPAO]) as $1/\tau$ for the forward reaction (i.e. formation of PS-SPAO or PSe-SPAO) is directly proportional to the concentration of SPAO. By fitting a straight line through these data points (see Figures A.14 and A.15), the rate constants k_{1-2} and k_{1-3} were obtained at each pH. The rate of interconversion between the diastereomeric species (2 and 3 in Figure 4.2) refers to the stereochemical inversion of the arsenic centre whilst the SPAO molecule was covalently attached to the pore (see Figure 4.2, **C**). By plotting the $1/\tau$ against SPAO concentration it was shown that the rate constants for the interconversion between diastereomers was independent of SPAO concentration (see Figures A.14 and A.15). Rate constants for the reverse reaction from the covalent adduct (i.e. PSe-SPAO or PS-SPAO) back to the free 117Cys or 117Sec (denoted as 1 in Figure 4.2 with rate constants named k_{2-1} and k_{3-1}) were obtained by plotting $1/\tau$ against SPAO concentration. The rate of the reverse reaction was not dependent on the concentration of SPAO. The rate constants for the reaction of SPAO with 117CysSECIS$_1$WT$_6$ and 117SecSECIS$_1$WT$_6$ over the pH range 6.5 - 8.5 have been tabulated (see Table 4.1).

117SecSECIS$_1$WT$_6$

pH	k_{1-2} ($M^{-1}s^{-1}$)	k_{1-3} ($M^{-1}s^{-1}$)	k_{2-1} (s^{-1})	k_{3-1} (s^{-1})	k_{2-3} (s^{-1})	k_{3-2} (s^{-1})
6.5	21 ± 3	7 ± 6	0.02 ± 0.007	0.01 ± 0.003	0.21 ± 0.01	0.11 ± 0.01
7.5	193 ± 95	8.8 ± 0.4	0.06 ± 0.01	0.02 ± 0.002	0.05 ± 0.01	0.17 ± 0.04
8.0	1134 ± 131	1217 ± 72	1 ± 0.1	0.62 ± 0.06	0.08 ± 0.01	0.05 ± 0.005
8.5	995 ± 66	2162 ± 906	0.9 ± 0.2	0.3 ± 0.1	0.05 ± 0.01	0.03 ± 0.005

117CysSECIS$_1$WT$_6$

pH	k_{1-2} ($M^{-1}s^{-1}$)	k_{1-3} ($M^{-1}s^{-1}$)	k_{2-1} (s^{-1})	k_{3-1} (s^{-1})	k_{2-3} (s^{-1})	k_{3-2} (s^{-1})
6.5	25 ± 8	16 ± 7	0.02 ± 0.01	0.01 ± 0.005	0.25 ± 0.02	0.11 ± 0.01
7.5	102 ± 17	81 ± 22	0.11 ± 0.02	0.05 ± 0.01	0.04 ± 0.01	0.03 ± 0.01
8.0	788 ± 114	638 ± 145	0.16 ± 0.03	0.32 ± 0.04	0.02 ± 0.006	0.04 ± 0.02
8.5	985 ± 50	911 ± 46	0.99 ± 0.02	0.62 ± 0.04	0.07 ± 0.02	0.05 ± 0.01

Table 4.1: Tabulated rate constants obtained from the reaction of 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ with SPAO (*trans*) over the pH range 6.5 - 8.5.

All data were analysed using the single channel analysis software QuB (www.qub.buffalo.edu, see Materials and Methods for further details). The rate constants for the reaction of SPAO with 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ were plotted against pH, as illustrated in Figure 4.3. It was expected that the rate of the forward reaction would depend on the pKa values of selenocysteine and cysteine. From this assumption, the rate constants plotted against pH were initially fitted using

$$\frac{k_{SH} + k_{S-}10^{pH-pKa}}{1 + 10^{pH-pKa}} \tag{4.1}$$

where k_{SH} is the minimum rate, k_{S-} the maximum rate for the cysteine (or selenocysteine) residue in its fully deprotonated form. This function did not fit the data satisfactorily meaning that the forward rate of reaction was either not dependent on the pKa of the selenol/ thiol or was dependent on other pKa values, for example those associated with the SPAO molecule (see Discussion). The function used to fit the k_{on} vs pH data (as shown in Figure 4.3) was

$$\frac{a + k_1 \exp^{\frac{(pH-pKa)H}{RT}}}{1 + \exp^{\frac{(pH-pKa)H}{RT}}} \tag{4.2}$$

where k_1 is the maximum rate constant, pKa the apparent pKa or midpoint and H is related to the spread of the sigmoidal curve with respect to pH. The constant "a" is the apparent minimum rate. Equation 4.2 is a modified form of that used to fit titration data with three or more pKa values that are cooperatively coupled (Lewin et al., 2006) (i.e. the equilibria are all inter-related) and was thought appropriate due to the many equilibria possible for the reactive SPAO species, in aqueous solution (see Discussion). The rate constants for the reaction between SPAO and 117CysSECIS$_1$WT$_6$ and 117SecSECIS$_1$WT$_6$ αHL were found to be similar over the pH range 6.5 - 8.5 (see Table 4.1). This was an unexpected result based on the pKa difference between selenocysteine and cysteine (see Discussion).

The rate constants for the reverse reaction (k_{2-1} and k_{3-1}) for both 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ were plotted against pH as illustrated in Figure 4.4, **B**. It can be seen in Table 4.1 that the rate of the reverse reaction increases by a factor of approximately 40 for PSe-SPAO and by a factor of 56 for PS-SPAO as the pH is raised from 6.5 - 8.5. Additionally, the 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ rate constants (k_{2-1} and k_{3-1}) at pH 6.5 are comparable (within error). The rate of interconversion (k_{2-3} and k_{3-2}) between the diastereomeric covalent adducts, formed within the β barrel of 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ respectively, have been plotted against pH (see Figure 4.4, **A**). The diastereomeric interconversion rate for the PS-SPAO and PSe-SPAO are similar and do not show a strong pH dependence over the

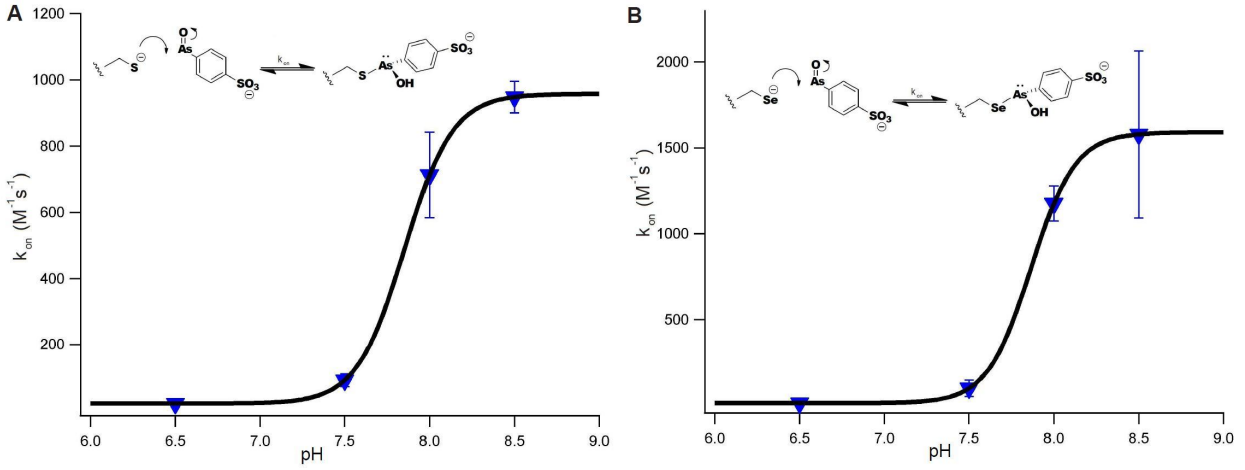


Figure 4.3: **A** The relationship between pH and the rate constants for the forward reaction of SPAO and 117CysSECIS$_1$WT$_6$. **B** The relationship between pH and the rate constants for the forward reaction of SPAO and 117SecSECIS$_1$WT$_6$

pH range 6.5 - 8.5.

4.2.2 Determining the pKa values of SPAO

The similarity between the kinetic data gathered using the 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ pores, was unexpected based on the difference between the pKa values of the 117Sec (pKa = 5.5) and 117Cys (pKa = 8.5 $\pm$ 0.5) residues (see Chapter 3). From observing how the sigmoidal relationship between the forward rate constants and pH (for the reaction of SPAO with 117Sec compared to with 117Cys) were so similar, another factor was deduced to be dominating the rate of the forward reaction, instead of the pKa difference between the 117Sec and 117Cys residues. The SPAO molecule, itself exists in an equilibrium between two structures as shown in Figure 4.5 **D**, however the position or the effect of pH on this equilibrium are not known. To investigate the properties of SPAO, a pH titration was carried out over the pH range 3 - 10.5 (see Figure 4.5, **A**) by addition of KOH to a solution of SPAO. The pH after each volume of KOH solution added, was recorded along with the volume of KOH solution added. By calculating the derivative of the change in volume, the pH at maximum gradient (the mid-point of the sigmoidal curve in Figure

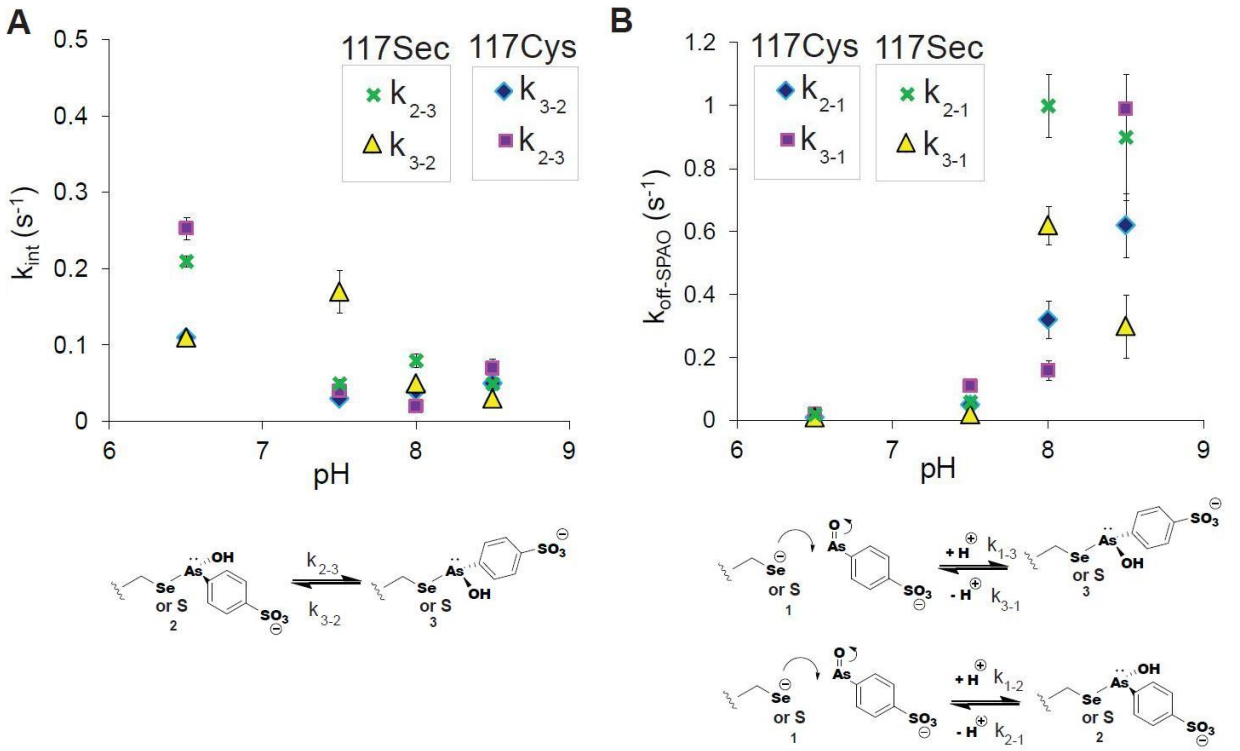


Figure 4.4: **A** The relationship between interconversion rate and pH for the reaction between SPAO and $117CysSECIS_1WT_6$ and $117SecSECIS_1WT_6$. **B** The relationship between pH and the rate constants for the reverse reaction of SPAO with $117CysSECIS_1WT_6$ and $117SecSECIS_1WT_6$.

4.5, **A**) was found to be 6.8 $\pm$ 0.2 (n = 3). This value was assigned to deprotonation of the first hydroxyl group (see Figure 4.5, **C**). The pKa for the second hydroxyl group of the hydrated SPAO form, was obtained by carrying out ^{13}C NMR experiments over the pH range 5 - 13. By plotting the difference (in Hz) between the carbon atoms ipso and para to the arsenic (III) centre, it was possible to obtain the second deprotonation equilibrium constant, pKa = 11.57 $\pm$ 0.04 (see Figure 4.5). The pKa of 6.8 obtained from the pH titration experiments was not observed from the ^{13}C NMR experiments. A possible explanation for this was that there was a negligible effect, of deprotonation of just one of the $RAs(OH)_2$ hydroxyl groups, on the ^{13}C NMR shifts, however when both OH groups were deprotonated (i.e. $RAs(O^-)_2$) the difference between the chemical shift values for the ipso carbon relative to the para carbon became greater. The ^1H NMR and ^{13}C NMR data are shown in Figures A.18 - A.28 (Appendix).

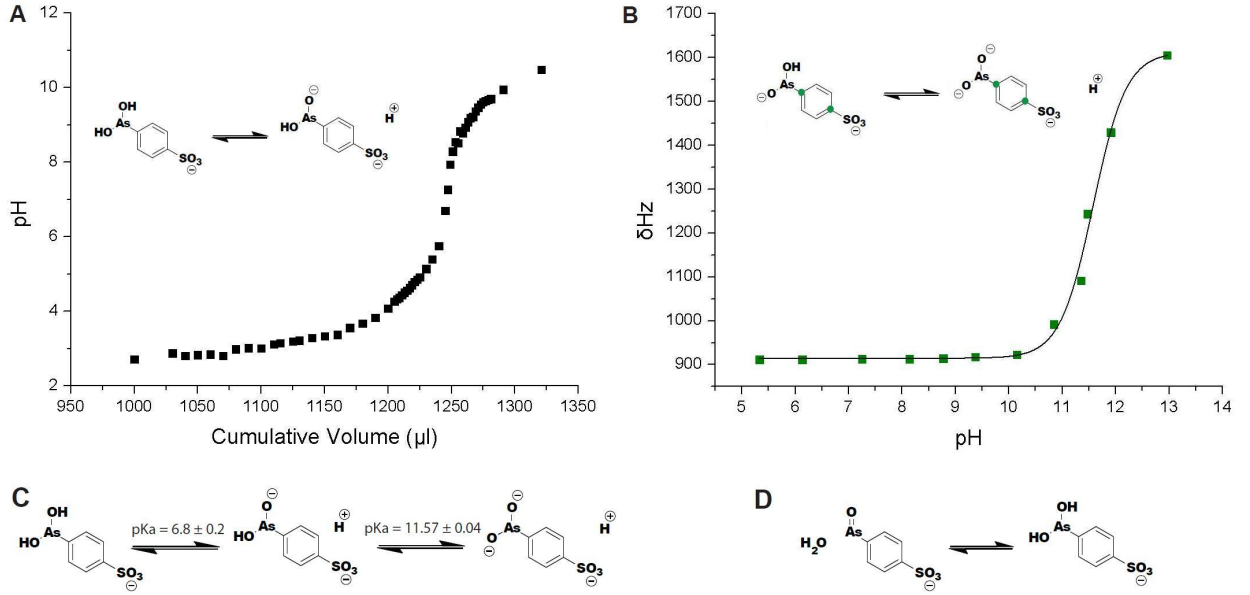


Figure 4.5: **A** The pH titration curve for SPAO over a pH range of 3 - 10.5. The volume of KOH added was recorded along with the pH after each addition. **B** The difference, in Hz between the ^{13}C NMR peaks of the ipso and para carbon atoms of the aromatic ring of SPAO was plotted against pH. **C** The pKa values obtained with the corresponding equilibria. **D** The equilibrium for the dehydrated and hydrated forms of SPAO.

4.3 Discussion

The reversible reaction between SPAO and a single cysteine residue situated at position 137 within the β barrel of αHL, was previously investigated (at pH 8.4, 80 mM MOPS, 2 M KCl, 100 μM EDTA) by Shin et al. (2007). Shin et al. (2007) reported k_{1-2} and k_{1-3} values of (14 $\pm$ 1) x 10^3 $M^{-1}s^{-1}$ and (5.9 $\pm$ 0.8) x 10^3 $M^{-1}s^{-1}$ which are approximately 5-10 times faster than those observed using the 117CysSECIS$_1$WT$_6$ pore. The difference in rate between the values obtained by Shin et al. (2007) and those obtained from the reaction of SPAO with 117CysSECIS$_1$WT$_6$, was thought to be due to the electrophoretic translocation of the negatively charged SPAO molecule (placed into the *trans* chamber) through the pore, being affected by the applied potential (which was +50 mV for the experiments using the 117CysSECIS$_1$WT$_6$ pore). The effective concentration

of SPAO in the pore was expected to be lower than when applying a potential of -50 mV (as used by Shin et al.). The rate constants k_{2-1} and k_{3-1}, were found to be comparable to those observed by Shin et al. (2007) using the αHL pore containing a single cysteine residue at position 137 in the β barrel.

4.3.1 The pH dependence of the forward reaction of 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ with SPAO

Correcting the forward rate constants with respect to the pKa of the selenol and thiol groups of 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ for the reaction with SPAO

As mentioned previously, the rates for the reaction of selenocysteine with SPAO and for the reaction of cysteine with SPAO were expected to be governed by the pKa of the selenol and thiol groups respectively. The comparable trend and magnitude of the forward rate constant values obtained for the reaction between SPAO and 117SecSECIS$_1$WT$_6$ or 117CysSECIS$_1$WT$_6$, over the pH range 6.5 - 8.5, indicated that the forward rate does not depend solely on the pKa of the selenocysteine or cysteine residue and that the equilibria of the SPAO molecule were determining the apparent rate of reaction.

The significance of the apparent pKa observed for the forward reaction between SPAO and 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$

The apparent pKa obtained for the reaction of SPAO with 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ was 7.9 $\pm$0 0.1 and this value was obtained from plotting the forward rate constant values against pH (see Figure 4.3). It is known that RAs(III)$(OH)_2$ species can exist in equilibrium with their dehydrated form (Sokolowski & Konopski, 2007). As mentioned previously, the position of the equilibrium between RAs$(OH)_2$ and RAs(O) shown in Figure 4.5, **D** is not known, and so it can-

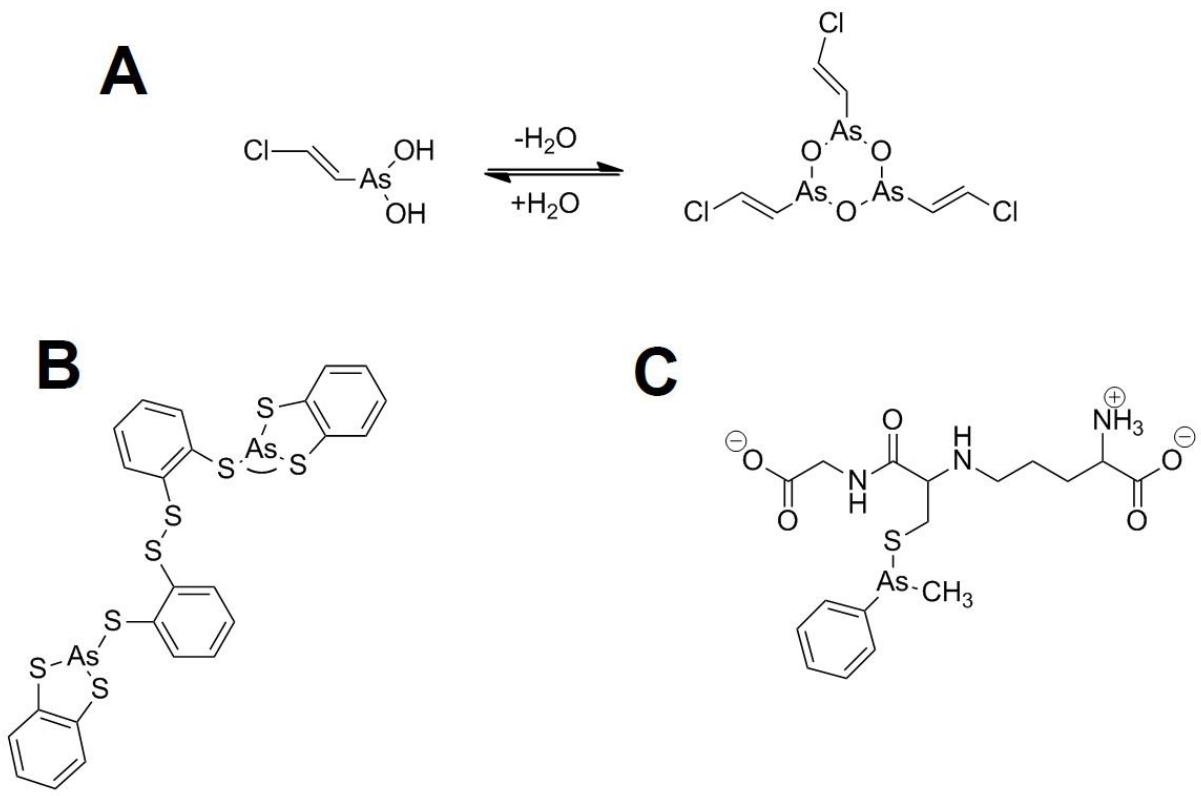


Figure 4.6: **A** The polymerisation/oligomerisation of lewisite on increasing its concentration above 10 mM (adapted from Sokolowski & Konopski (2007)). These oligomerised forms of lewisite were reported to be not very soluble. **B** The bond angle (α) which is 120^o in the transition state geometry during pyramidal inversion about the arsenic centre. **C** Methylphenylarsinic acid-(*L*)-glutathione.

not be assumed that the selenocysteine or cysteine residues are not reactive towards the RAs(O) form of SPAO. It is also not known with which of these two forms of SPAO the selenocysteine and cysteine residue prefers to react, or even if there is a preference. Waters & Williams (1950) hypothesised for lewisite (See Figure 4.6, **A**) that the backward reaction from the RAs(O) to the hydrated form ($RAs(OH)_2$) occurs at a "slow" rate, however no quantitative rate constants were provided. The relationship between pH and the As(III) species equilibrium ($RAs(OH)_2$/RAs(O)) could affect the forward rate of the reaction between 117Sec or 117Cys and SPAO if the selenocysteine and cysteine residues only reacted, for example with the hydrated, $RAs(OH)_2$ form of SPAO. Sokolowski & Konopski (2007) studied the reaction of lewisite with thiols under aqueous conditions and reported the same "slow" rate, but similarly no quantitative data was provided regarding

the equilibrium between the hydrated and dehydrated forms of lewisite. Another consideration was whether polymerisation of SPAO was another equilibrium that needed to be considered. Polymerisation of As(III) species has been reported to occur (Sokolowski & Konopski, 2007) in a concentration as opposed to pH, dependent manner for lewisite (see Figure 4.6, **A**). ^{1}H NMR spectra of SPAO over the pH range 5 - 12, were obtained and as no obvious peak broadening was observed, it was concluded that no polymer species existed in solution. In support of this, no precipitate was observed during any of the electrophysiology experiments and the ^{1}H NMR samples for SPAO over the pH range 5 - 12 (see Figure A.16), where precipitate was noted for lewisite at concentrations above 10 mM, due to insoluble polymerised species. The concentration of SPAO used during the NMR experiments was around 20 mM and so based on these experiments, it was concluded that polymerisation of SPAO was not occurring during the single channel electrical recordings.

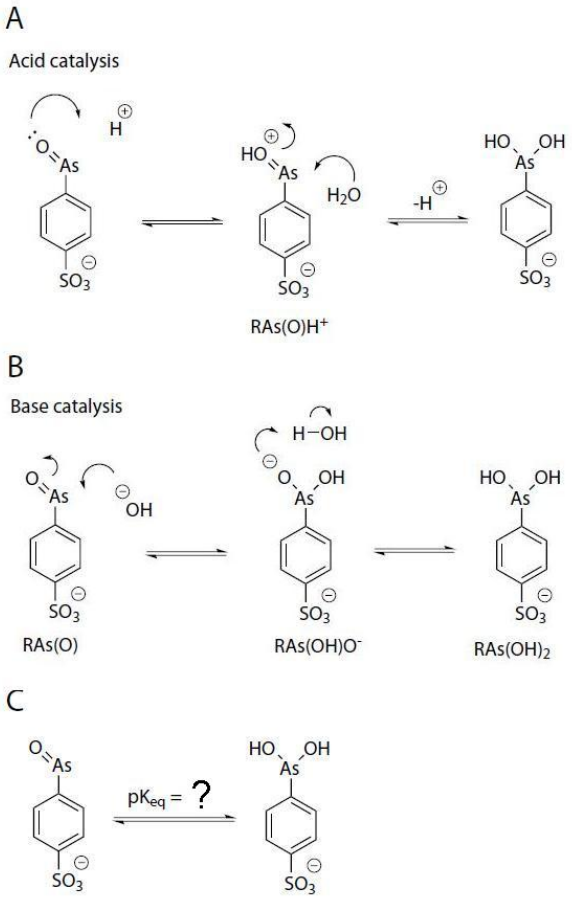
Rationalising the relationship between the rate constants and pH for 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$

The equilibrium between RAs(O) and RAs(OH)$_2$ as shown in Scheme 4.3, **C**, can be acid or base catalysed. The reactivity between 117CysSECIS$_1$WT$_6$ and 117SecSECIS$_1$WT$_6$ with SPAO over the pH range 6.5 - 8.5 have been summarised in Table 4.1.

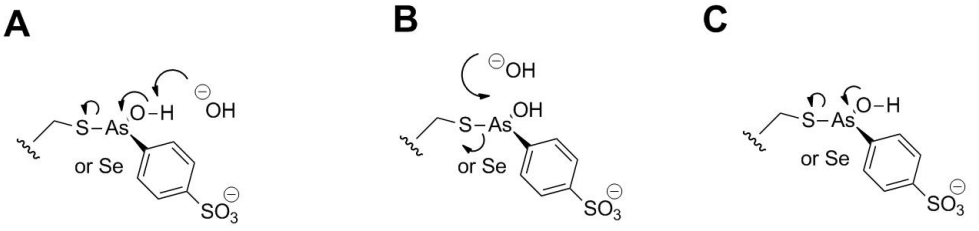
4.3.2 The dependence of the reverse reaction between SPAO with 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ with pH

The pH dependence of the reverse reaction rate (k_{2-1} and k_{3-1})

Two possible mechanisms by which the reverse reaction between SPAO and 117CysSECIS$_1$WT$_6$ proceeds at pH > 7, have been illustrated in Scheme 4.4. The k_{2-1} and k_{3-1} values increased by approximately 50 times on increasing the pH from 6.5 to 8.5. The rate constants, k_{2-1} and k_{3-1} for



Scheme 4.3: **A** The acid catalysed and **B** base catalysed pathway between the dehydrated RAs(O) form of SPAO and the hydrated $RAs(OH)_2$ form. **C** The equilibrium between the dehydrated and hydrated forms of SPAO.



Scheme 4.4: **A** Deprotonation of the PS-As(R)(OH) species followed by resonance of the electrons and loss of the thiolate. **B** Attack of the ^{-}OH onto the arsenic centre to form the thiolate and $RAs(OH)_2$ species. **C** Loss of arsenic via formation of an arsenium cation ($RAs^{+}(OH)$) in resonance with the $RAs{=}O^{+}H$ species.

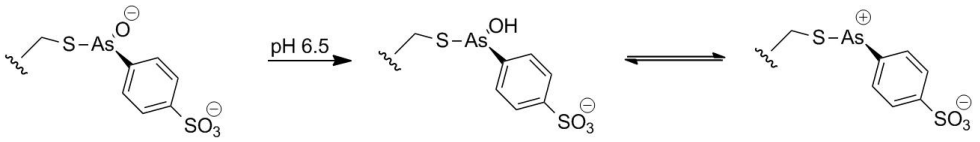


Figure 4.7: A mechanism for stereochemical inversion of the SPAO molecule (PS-SPAO) within the β barrel of αHL, via formation of an arsenium cation (PS-As^{+}R).

the reaction of 117Cys and 117Sec with SPAO were comparable suggesting that the mechanism governing the reverse reaction is the same. It was not possible to ascertain any further mechanistic insight, for example into whether either of the pathways illustrated in Scheme 4.4 are favoured over the other.

4.3.3 The pH dependence of the rate of interconversion between diastereomers of SPAO

Single channel experiments carried out by Shin et al. (2007) using a single cysteine containing αHL pore, first noted the appearance of two current blockade levels using an applied potential of -50 mV, on addition of SPAO to the cis chamber. These two blockade levels were proposed to be the formation of diastereomers of the PS-SPAO species. The interconversion between these two diastereomers was proposed to proceed via pyramidal inversion at the arsenic centre. To confirm the formation of diastereomeric species, ^{1}H NMR experiments were carried out using glutathione and SPAO in ratios of 1:1 and 2:1. The -SCH_2- group, hydrogen atoms of each diastereomer produced a pair of double doublets in the ^{1}NMR spectrum and so the observation of 16 lines confirmed the presence of both diastereomers (where the -SCH_2- resonances were assigned in the 2:1 adduct and are due to the chemically non-equivalent protons i.e. due to the tetrahedral nature of the As centre). Interconversion between the two blockade levels was also found to be independent of the concentration of SPAO, indicating that the mechanism by which the diastereomers interconvert is unimolecular. A possible mechanism for the stereochemical interconversion would be

via a arsenium cation (see Figure 4.7). The inversion rates between the diastereomeric species for both the 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ remained almost constant between pH 7.5 - 8.5, increasing by a factor of 10 when the pH was lowered from 7.5 to 6.5. The mechanism for pyramidal inversion proceeds via a trigonal planar transition state with bond angles of 120^o between the substituents surrounding the central atom (Lambert et al., 1968). Trigonal pyramidal compounds exhibiting bond angles closest to the transition state geometry (i.e. 120^o around the central As atom) will invert more easily. The bond angles quoted for an R-As-R species, where the R substituent is a second period element (for example S or P), typically range between 105 - 115^o (Lambert et al., 1968). For example a value of 103^o for the bis(2-(1,3,2-benzodithiarsol-2-ylsulfanyl)-benzenesulfide), RS$_1$-As-SR$_2$ bond angle was reported by Shaikh et al. (2006), see Figure 4.6, **C**. The PS-As-OH and PS-As-Ar bond angles of the diastereomeric SPAO species are expected to be dictated by the repulsion of the lone pair of the arsenic (III) centre and the steric hinderance of the three substituents (Rauk et al., 1969). The ability of the substituents to accomodate the As lone pair, for example via π-*d* overlap or $\pi - \pi$ interactions has been found to strongly affect the inversion rate. For example the pyramidal inversion of methylphenylarsinic acid-(*L*)-glutathione was observed using ^{1}H NMR by Edmonds et al. (2006) giving an estimated half life of 2 - 3 hours at pH 4. The comparable magnitudes of the inversion rates obtained using the 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ pores suggests that the rate of inversion at the arsenic centre was not affected by the alteration from a selenium to sulfur substituent. Inversion rate constants of 0.05 s^{-1} were obtained at pH 8.5, giving a half life of 10 seconds. Methylphenylarsinic acid-(*L*)-glutathione differs in structure from PS-SPAO (see Figure 4.6) where a methyl group is replaced by an OH group (in the case of SPAO) and the aromatic ring of SPAO contains a *para* SO_3^- group which is not present on the phenyl ring of Methylphenylarsinic acid-(*L*)-glutathione. This suggests that the presence of the OH and electron withdrawing SO_3^- groups significantly effect the inversion rate.

4.3.4 Summary

The reaction of 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$ with SPAO was carried out over the pH range 6.5 - 8.5. Rate constants for the forward, reverse and interconversion equilibria, were obtained. In summary, the pKa values of the hydrated form of SPAO species were obtained from pH titration and ^{13}C NMR experiments, yielding values of pKa = 6.8 $\pm$ 0.2 and pKa = 11.6 $\pm$ 0.04. The relationship between the forward reaction of SPAO and 117SecSECIS$_1$WT$_6$ was hypothesised to be dependent on the equilibrium of a particular SPAO species and that the pKa of 7.9 observed was the apparent pKa for the species of SPAO that was reactive towards the selenocysteine and cysteine residues. Comparisons between the magnitudes and trends of these rate constants, with respect to pH, were carried out providing mechanistic insight that can be used to further understand the action of organoarsenic (III) compounds *in vivo*.

4.4 Materials and Methods

4.4.1 Synthesis of sodium 4-sulfophenyldiiodoarsine (SPAO)

The synthesis of sodium 4-sulfophenyldiiodoarsine is described below. When dissolved in an aqueous solution, the hydrolysis product, 4-sulfophenyarsine (SPAO) is formed. The synthetic procedure was obtained from Oneto (1938).

Sodium 4-sulfophenylarsonic acid

To a solution of sulfanilic acid (5 g, 29 mmol) in 35 mL H_2O (5% Na_2CO_3). The diazonium species was formed *in situ* by simultaneous, dropwise addition of both a 12 M HCl (10 mL) solution and an aqueous sodium nitrite solution ($NaNO_2$, 2 g, 30 mmol, 35 mL), to the solution of sulfanilic acid. In a separate 500 mL round bottom flask, an excess of arsenic trioxide ($\approx$ 100 mM) was

dissolved in 100 mL of an aqueous solution (2 M NaOH, 7.5 mM $CuSO_4$). This solution was chilled on ice and remained on ice during the addition of the solution containing the diazonium species. The solution containing the diazonium species was added dropwise into the 500 mL flask containing the arsenic trioxide solution, resulting in a colour change of from blue to dark green and evolution of N_2 (g). The solution was then allowed to warm to room temperature and was left stirring for 3 days. The volume of the solution was reduced to approximately 25 mL via rotary evaporation followed by boiling and then acidification with concentrated HCl resulting in a brown precipitate. The solution was left to precipitate for 4 days at room temperature before filtration and washing with solutions of saturated KCl, 80% ethanol and finally acetone. The light brown, crude product was dried under vacuum to yield (4 g) and used without further purification.

Sodium 4-sulfophenyldiiodoarsine

Sodium 4-sulfophenylarsonic acid (1 g, 3.4 mmol) was dissolved in water (10 mL) and warmed to approximately 55 - 60^{o}C. An aqueous solution of HI (hydroiodic acid, 57%, 3 mL) was added and the resulting mixture (black/ dark in colour) stirred (in the dark) for approximately 5 minutes. After cooling the solution, the resulting precipitate was collected and washed with acetic acid followed by acetone. The yellow precipitate was dried under vacuum. Recrystallisation from 80% acetic acid was carried out by dissolving into the minimal amount of warmed solution followed by filtration of any insoluble precipitate. After the solution was cooled, sodium 4-sulfophenyldiiodoarsine was isolated as a yellow crystalline precipitate (200 mg, 0.8 mmol). ^{1}H NMR δ_H(400 MHz, D_2O, Sodium Phosphate 100 mM) 7.82 (2H, d, *J* 7.97, ^-O_3SC=CH), 7.78 (2H, d, *J*, 7.97, AsC=CH); δ_C(500 MHz, D_2O, Sodium Phosphate 100 mM) 125.48 (2C, **C**=C(SO_3^-)=**C**), 129.78 (2C, **C**=C($As(OH)_2$)=**C**), 144.12 (1C, C(SO_3^1)), 151.37 (1C, C($As(OH)_2$)).

4.4.2 Titration experiments

The pH titration experiments were carried out over the pH range 3 - 11. The pKa of the first deprotonation of 4-sulfophenyarsine (SPAO) involved sequential addition of 1 - 5 μl of KOH (aq) solution (0.2 M) to a solution of 20 mM SPAO dissolved in H_2O (500 μl starting volume). The pH after each volume addition was recorded and the relationship between the KOH solution volume and pH plotted.

4.4.3 ^{13}C NMR experiments

4-sulfophenyarsine (SPAO) was dissolved to a concentration of 24 mM (3.5 mg in 0.6 mL of deuterated 100 mM sodium phosphate buffer). The solution pH was altered by addition of KOD. ^{1}H and ^{13}C NMR spectra were recorded over the pD range 5 - 12 and the difference between the chemical shift for the carbons ipso to the arsenic centre and SO_3^- groups was measured (in Hz) at each pH.

Electrical recording set up

The electrophysiology set up has been described previously so for further information regarding the single channel recording experimental set up see Chapter 3. An applied voltage of +50 mV was employed to carry out the electrical recordings. Briefly, the electrophysiology set up used a chamber consisting of two compartments denoted *cis* and *trans* with a piece of teflon separating the compartments. The compartments were filled with buffer (20 mM Bis Tris Propane, 2 M KCl, 100 μM EDTA, pH 6 - 10) and a DPhPC bilayer was formed across a roughly 100 μm diameter aperture in the teflon. Between 0.5 - 1 μl of either 117SecSECIS$_1$WT$_6$ or 117CysSECIS$_1$WT$_6$ was added to the *cis* chamber and a high positive voltage of between 150 mV and 300 mV applied so as to encourage a pore into the bilayer.

Experimental

Sodium 4-sulfophenyldiiodoarsine SPAO was dissolved in H_2O to a final concentration of 100 mM and the pH of the solution raised from around pH 3 to pH 7. During the single channel recording, increasing amounts of this SPAO solution was added to the *trans* chamber and data recorded at each SPAO concentration so as to gather at least 50 events for each level.

4.4.4 Kinetic analysis

Kinetic analysis was carried out using QuB 2.0.0.2. An input file (.dwt) for this program was created and the level and dwell time information for each event was obtained by carrying out "single channel search" analysis in Clampfit (pClamp software, Molecular Devices) in an analogous way to that described in Chapter 3. A three state kinetic model was created as illustrated in Figure 4.8. The initial rate constants were defined as "1". QuB iteratively performs maximum interval likelihood analysis and fits the data by considering the duration of the events at each of the three levels. This process was carried out for all data and yielded the rate constants at a particular concentration of reagent. By plotting these values against SPAO concentration, it was possible to obtain the rate constant for the bimolecular reaction between SPAO and 177SecSECIS$_1$WT$_6$ or 117CysSECIS$_1$WT$_6$ i.e. τ = 1/k[SPAO]. The titration data can be seen in Figure A.14 and Figure A.15. The kinetic profile for this data has been illustrated in Scheme 4.2.

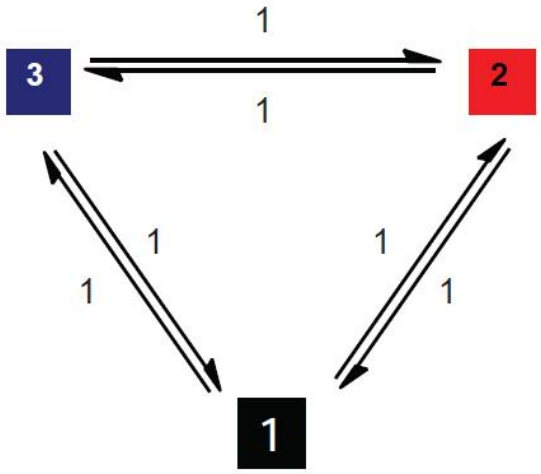


Figure 4.8: The three state kinetic model used for the analysis of the data gathered from reaction of SPAO with 117SecSECIS$_1$WT$_6$ and 117CysSECIS$_1$WT$_6$.

Chapter 5

Future Work

5.1 Enzyme mimics

One of the applications of the 117SecSECIS$_1$WT$_6$ αHL pore is to be engineered as an enzyme mimic or catalytic centre involving selenocysteine and or a cysteine residue. There are a broad range of methods by which catalytic sites have been created, including the use of proteins as scaffolds on which to attach the catalytic moiety (Pordea et al., 2008), supramolecular nanocavities (Hastings et al., 2010), DNA via non-covalent interactions (Boersma et al., 2010), active-site redesign (Toscano et al., 2007, e.g.) so that the enzyme carries out a different, modified function, embedding of an enzyme within a hydrogel (Wang et al., 2008), or the use of micelles (Huang et al., 2007) or βCD (Tang et al., 2010) as supramolecular scaffolds on which to attach and tune catalytic turnover. The ability to mimic enzymes has applications ranging from uses for medicinal and biocatalysis purposes to providing mechanistic insight into the reaction pathway itself. The ability to engineer functionality also relays information regarding structure-function relationships. Important factors to be considered when designing or modifying a catalytic site include the shape and polarity of the active site, the presence (or not) of water and recently it was noted that the ionic

strength could greatly enhance the catalytic reaction rate (Wang & Xie, 2006).

5.2 Selenocysteine containing enzymes

A second application of the 117SecSECIS$_1$WT$_6$ αHL pore would be to study the binding of metal cations to the selenolate functional group. αHL creates an easily controlled environment in which the pH and ionic strength can be controlled. The selenoenzyme, formate dehydrogenase H contains an active site Molybdenum which binds to a selenocysteine residue within the catalytic site (Enemark et al., 2004). The binding of metal ion such as silver (Ag^+) or zinc would also be of biological relevance, for example to understand the interaction of metal ions with selenocysteine-containing proteins *in vivo*.

The creation of a selenoprotein enzyme mimic within the β barrel of αHL would allow the study of catalytic cycles proposed for the iodothyronine (ID) enzymes. The iodothyronine enzymes involve the activation of an iodinated prohormone to the biologically active hormone. The ID-1 enzyme is a 28 kDa integral membrane protein found mostly in the liver, kidney, thyroid and pituitary. As well as an active site selenocysteine residue, the histidine 158 is known to be essential for enzyme activity. The proposed catalytic mechanism involves the formation of a selenenyl iodide bond by removal of one of the aromatic iodine atoms followed by regeneration of the active site selenocysteine by thiol containing co-factors. Questions still exist as to exactly how the selenocysteine residue catalyses the deiodination process. For example what oxidation state the selenol group is in or whether the deiodination pathway generates or proceeds via a radical mechanism. For example, Vasilev & Engman (1998) carried out a series of small molecule deiodination reactions using different selenolate and tellurolate nucleophiles and concluded that a radical mechanism could not be ruled out. Similarly, the identity of the thiol cofactor has evaded determination. Sun et al. (1997) proposed that one of the cysteine residues situated close to the active site selenocysteine

residue of ID-1 forms an internal selenylsulfide bond that could then be reduced by DTT, however Croteau et al. (1998) reported that the conserved cysteines near the ID-1 active site are not essential for catalytic activity. Mugesh et al. (2002) investigated whether the selenol group can be reformed in one step from the selenenyl iodide bond or whether this process requires two steps. It was concluded that the concept of an intermediate selenylsulfide bond was "more reasonable" than a one step reduction of the selenenyl iodide bond. The deiodination equilibria have been difficult to study due to the short lived nature of the selenenyl iodide bond in the presence of thiols (and water). By using a nanosized hydrophobic cavity, Goto et al. (2010) were able to isolate and characterise a selenenyl iodide bond within the hydrophobic interior. The decomposition rate of an aromatic selenenyl iodide compound was studied by du Mont et al. (1987) where from NMR analysis, the reaction was said to be complete within several minutes. By employing the nanocavity and taking ^{1}H NMR measurements over a period of a week, it was found that the selenenyliodide was stable throughout this time in $CDCl_3$ at room temperature. The proposed mechanism for the formation of the selenenyl iodide bond involved attack of the selenol group onto the iodine of the keto resonance structure. Support of a nucleophilic as opposed to radical mechanism came from the observation that no selenenyl iodide species was formed when a methyl ester was placed ipso instead of a hydroxyl group. Other interesting reactions that could be investigated within the selenocysteine containing αHL pore include the creation of an internal selenylsulfide bond by mutating a cysteine residue into the β barrel of αHL. The reaction of small disulfide molecules or peptides containing three cysteine residues with the selenylsulfide bond can be imagined to create a TrxR or disulfide shuffling center within αHL where it may be possible to tune the pKa values of the cysteine residues to favour the formation of a certain peptide conformation over another in an analogous way to how the conotoxins are arranged. The $117SecSECIS_1WT_6$ pore was challenging to use for single channel electrical recordings because of oxidation of the selenolate group. It may be possible to study the reduction of the oxidised species by addition of sodium borohydride or excess DTT. By incorporation of ebselen, a known Gpx mimic, the mechanism of the reaction of

ebselen with hydrogen peroxide and a dithiol containing species, could be studied. Intermediates would be observable, for example the singularly oxidised selenol (Se-OH) and the selenylsulfide formed between the selenolate and dithiol.

Appendix A

Additional data

Trypsin digest

The WTD8H6, 117CysSECISD8H6 and 117SecSECISD8H6 protein samples were cut from a pre-cast 12 % Bis-Tris gel (Criterion). The gel slices were cut into small cubes and placed in an eppendorf. An ammonium bicarbonate solution (100 μl, 25 mM, 10 mM DTT) was used to wash the gel slices followed by removal of the supernatant. This washing step was repeated until all of the instant blue stain (Expedeon) had been removed and the gel slices were colourless. Dehydration was carried out using acetonitrile (100 μl, Sigma Aldrich). The acetonitrile supernatant was removed and the white gel pieces centrifuged in a SpeedVac so as to remove any residual acetonitrile. Two further wash steps were carried out using the ammonium biocarbonate solution (100μl) and a final acetonitrile wash (100 μl). The iodoacetamide reaction (100 μl, 55 mM iodoacetamide, 25 mM ammonium bicarbonate) was carried out for an hour at rt (wrapped the eppendorfs in foil). After removal of the supernatant, two ammonium bicarbonate washes and a final acetonitrile wash, were carried out in an identical way as before. Once the gel pieces were dry, the trypsin solution (approximately 50 μl, see below for details) was added and the reaction left at 37 ^{o}C overnight. 1 μl of formic acid was added to quench the digest and the peptide containing supernatant retained.

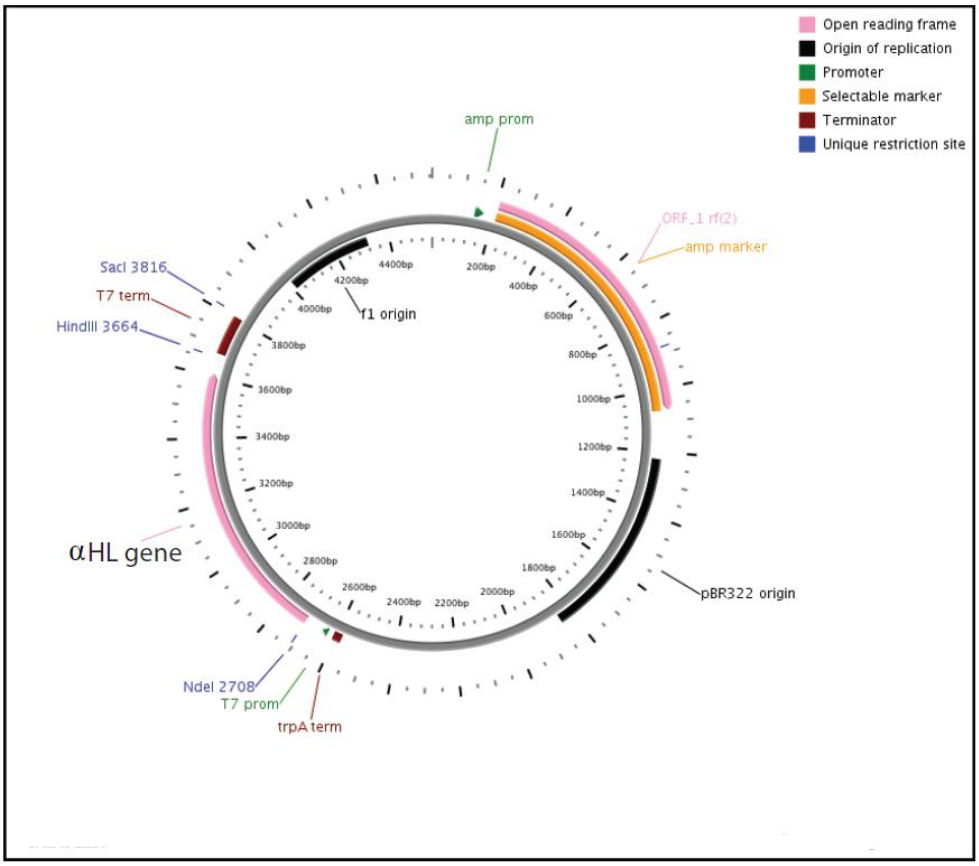


Figure A.1: The 4.6kb pT7 plasmid containing the αHL gene that was used to express the selenocysteine containing monomeric αHL protein. The restriction sites used to engineer a Sec codon and SECIS element into the αHL gene are shown. The position of the β-lactamase (AMP) gene has also been indicated. The above plasmid map was created by Dong et al. (2004).

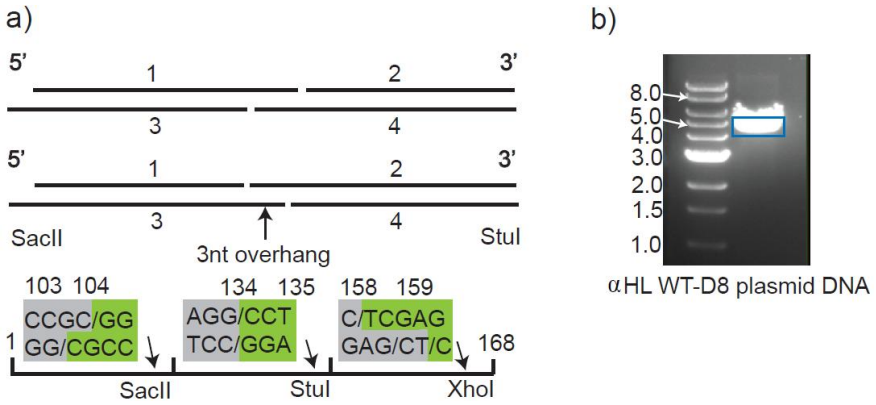


Figure A.2: **A** The primers used for three-way ligation so as to engineer the Sec codon and SECIS element into the αHL gene. The restriction sites chosen for this ligation are also shown. **B** The doubly digested pT7 plasmid prior to gel extraction ready for the three-way ligation.

A C4 membrane was used to desalt. Briefly, the supernatant was passed across the C4 membrane using an vacuum. Washing steps were carried out followed by elution using a 30 - 50 % acetonitrile solution (0.1 % formic acid). The volume of each sample was then reduced to 5 μl via SpeedVac and stored at -20 ^{o}C. The LC-MSMS was carried out at the Central Proteomics Facility (Sir William Dunn Pathology School, South Parks Rd., Oxford, OX1 3RE). The program Mascot was employed to analyse the mass spectra. Although the 117SecSECISD8H6 monomeric protein was observed, it was only possible to isolate a peptide fragment containing the selenocysteine residue on one out of six repeat experiments. The other five trypsin digests and subsequent mass spectroscopy analysis revealed all other fragments of the protein but were missing the fragment containing the selenocysteine residue. This was thought to be due to two main reasons. Firstly, although iodoacetamide protected, the selenium functional group was still vulnerable to oxidation or even elimination. Secondly, the peptide fragment (8 amino acids long) was originally part of the hydrophobic β barrel portion of α and so may not have been very soluble.

Trypsin solution preparation

Resuspension buffer (100 μl, Promega)was added to Trypsin (promega sequencing grade modified trypsin), aliquoted (5 μl) and stored at -20 ^{o}C. The trypsin solution referred to above was prepared

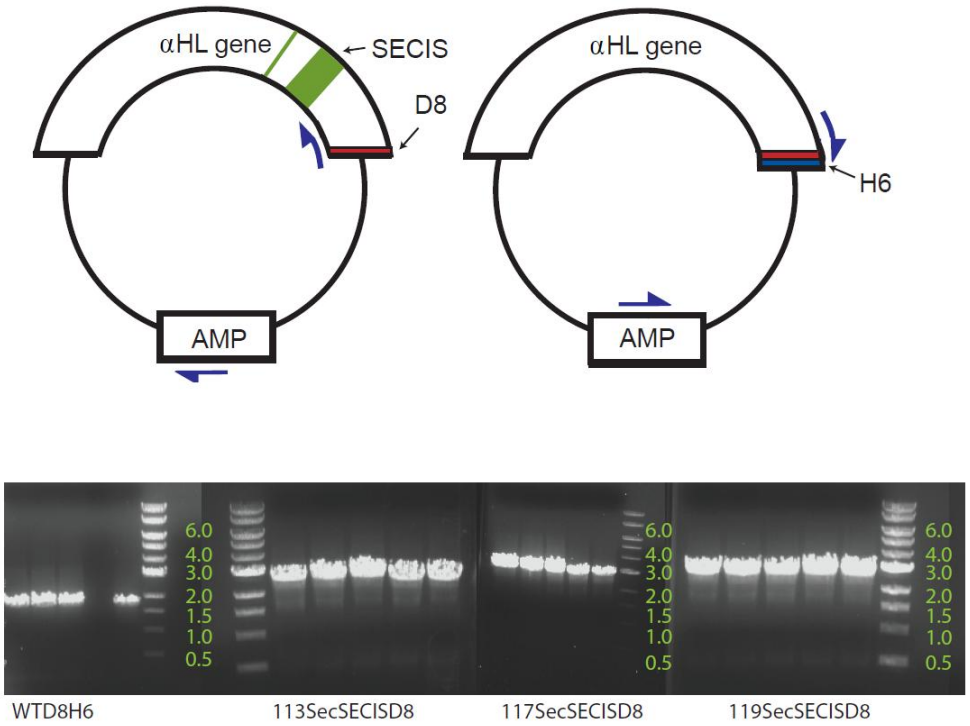


Figure A.3: Homologous recombination was used to add an addition six histidine residues onto the C-terminus of each of the selenocysteine containing αHL constructs. The primers used for the two PCR reactions have been depicted (arrows) along with the pT7 plasmids used for each half of the recombination. The PCR homologous recombination products can be seen on the analytical gels (below).

by addition of ammonium bicarbonate (95 μl, 25 mM ammonium bicarbonate) immediately prior to use.

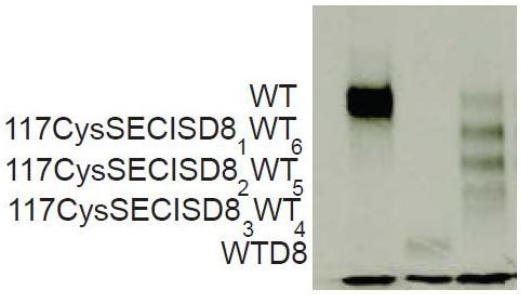


Figure A.4: On incubation of the αHL monomers in the presence of rabbit red blood cell membranes, the monomers will assemble into heptamers. By varying the ratio of WT non-selenocysteine containing monomeric protein to selenocysteine or cysteine containing monomeric protein it is possible to create heptamers containing different numbers of each type of monomer. The heteroheptamer consisting of a single selenocysteine/cysteine-containing monomer and six WT monomers was gel purified. The D8 tail of the selenocysteine/cysteine-containing monomer caused the heteroheptamers to run at varying speeds relative to the number of D8 tails. In this way a heteroheptamer containing 6 WT and 1 selenocysteine/cysteine, were isolated and purified.

pH	$k_{on\text{-}DTTox}$ ($M^{-1}s^{-1}$)	$k_{off\text{-}DTT}$ (s^{-1})
8.0	2 ± 1	28 ± 5
8.5	4.7 ± 0.01	40 ± 6

pH	$k_{off\text{-}DTT}$ (s^{-1})
6.95	0.4 ± 0.5
7.5	7 ± 2
8.34	35 ± 10
9.22	119 ± 3
9.5	250 ± 43

Table A.1: The rate constants obtained from the reaction of DTTox with 117CysSECIS$_1$WT$_6$.

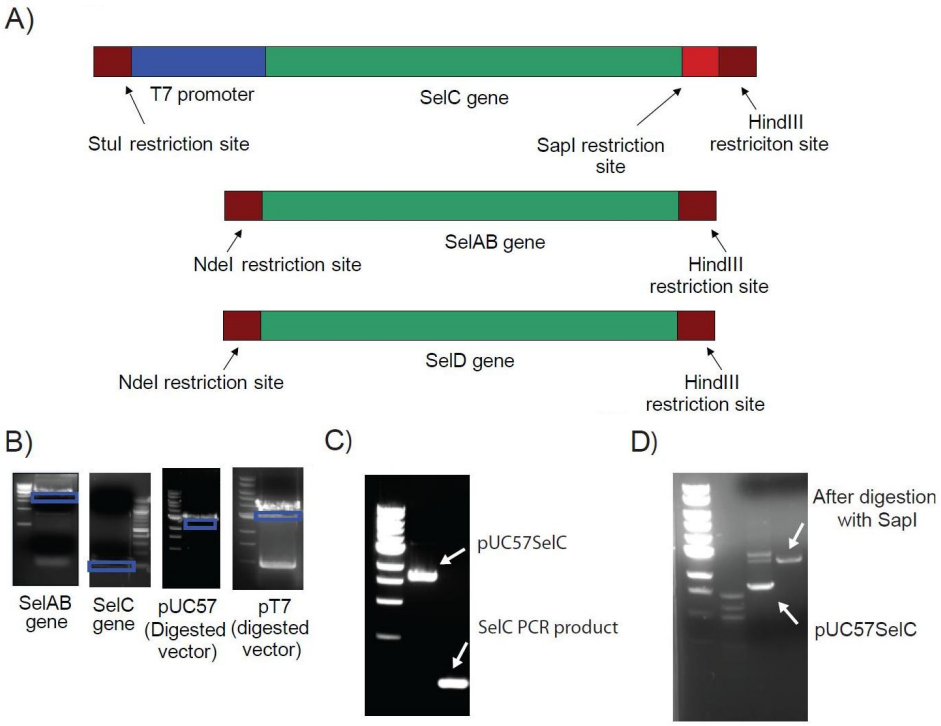


Figure A.5: A The gene constructs that were designed for each of the additional selenocysteine incorporation proteins (SelAB, selenocysteine synthase and selB and SelD, selenophosphate synthase) and tRNA^{sec} (selC) B Gel extraction was carried out after each PCR reaction to purify the DNA. Both PCR products and vectors are shown. C The final pUC57 vector containing the SelC gene (pUC57SelC) and the SelC PCR product containing a T7 promoter. D Digestion of the pUC57SelC using SapI. This was required so as to linearise the plasmid in a specific place prior to *in vitro* transcription.

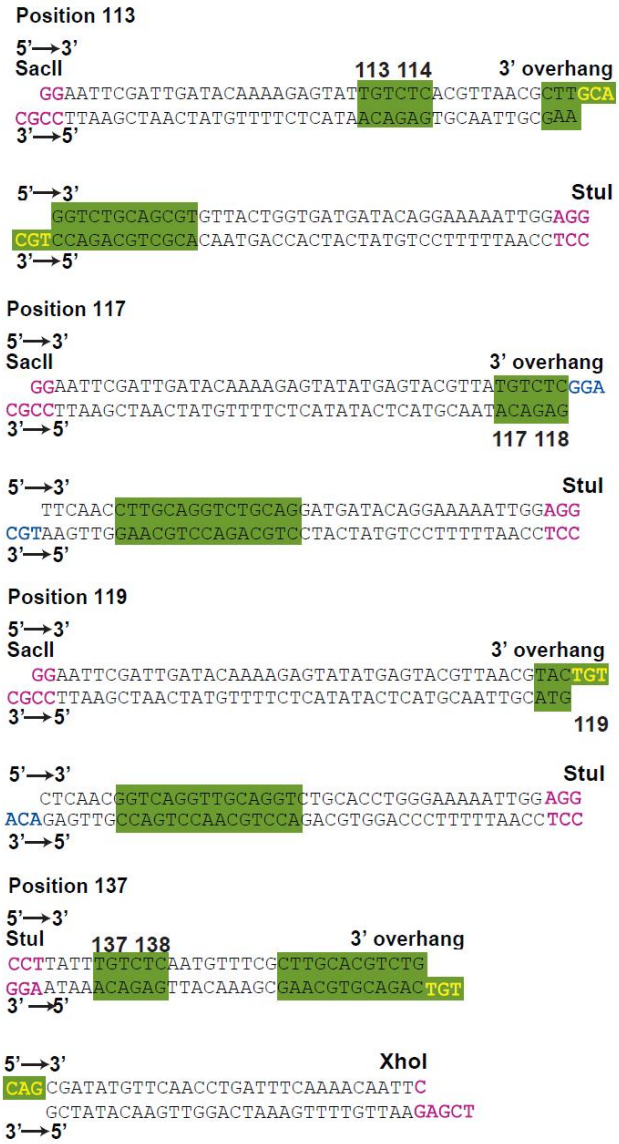


Figure A.6: The primer sequences used for three-way ligation. All the primers were designed so as to contain the Sec codon and SECIS element sequence (green)

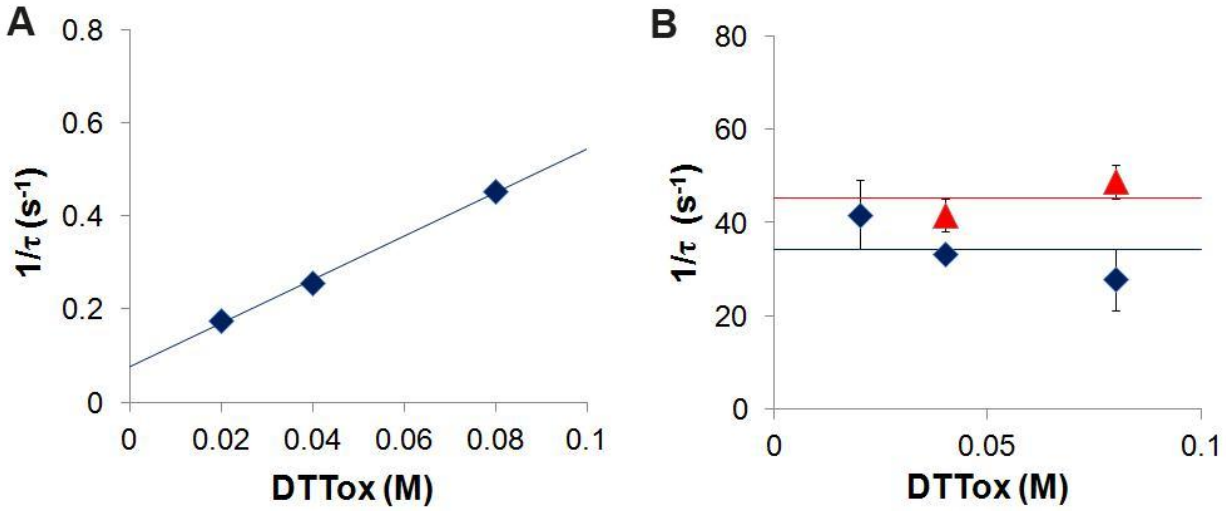


Figure A.7: The reaction of DTTox and 117CysSECIS$_1$WT$_6$ at pH 8.5. **A** The relationship between reaction rate and DTTox concentration. The $k_{on} = 4.70 \pm 0.05$ $M^{-1}s^{-1}$ **B** The relationship between the PS-DTT life time and the concentration of DTTox. $k_{off} = 45 \pm 5$ s^{-1} (level 1) and $k_{off} = 34 \pm 7$ s^{-1} (level 2).

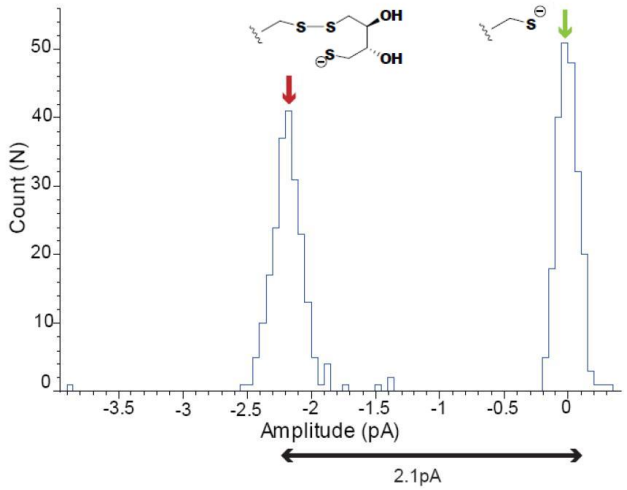


Figure A.8: The reaction of DTTox with 117CysSECIS$_1$WT$_6$ at pH 8.5, where only one PS-DTT level was observable.

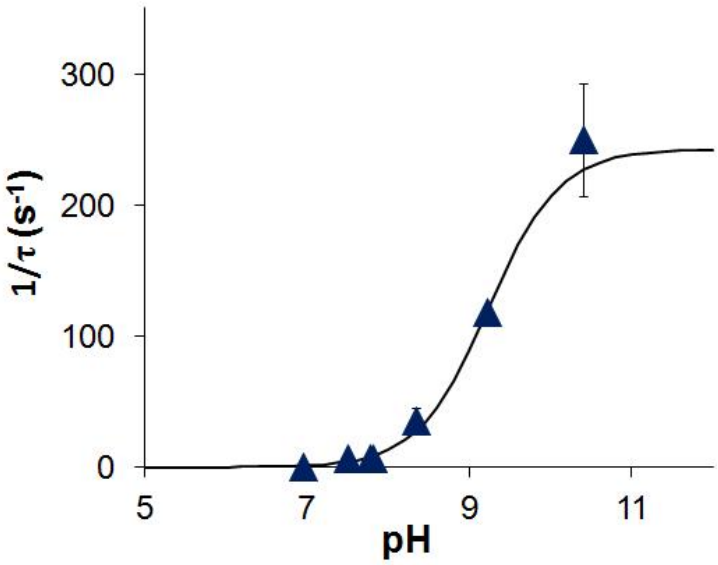


Figure A.9: The rate at which the PS-DTT intermediate cyclises to form PS^- and DTTox plotted against pH. The data were fitted using the relationship $k_{off-DTT} = \frac{k_{on-max}K_a}{(K_a+[H^+])}$ where the pKa = 9.2 $\pm$ 0.4 and k_{max} = 240 $\pm$ 50.

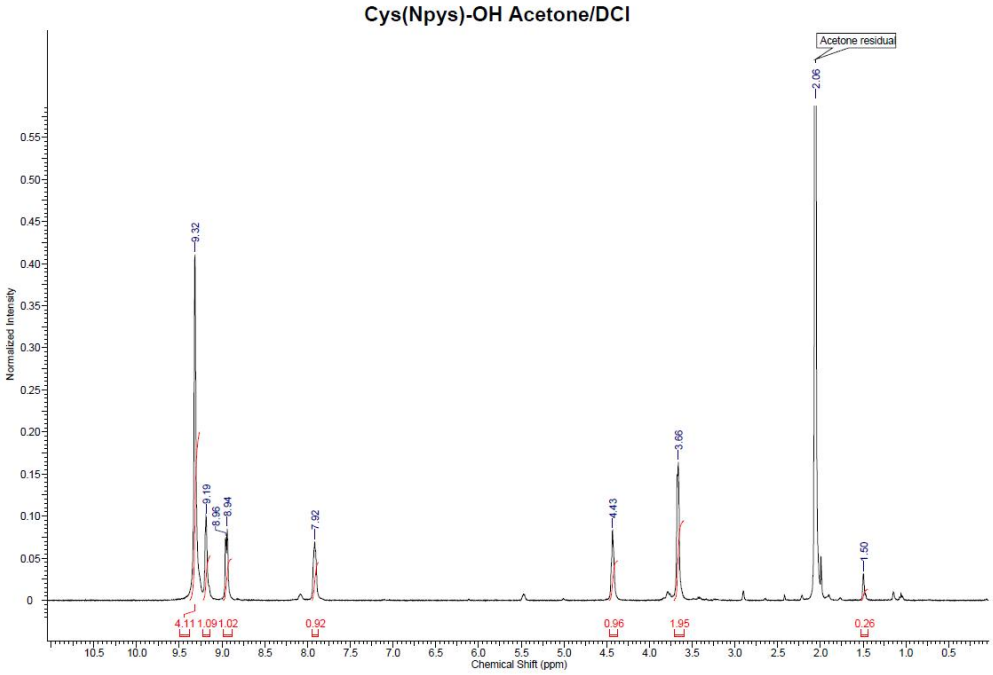


Figure A.10: Cys(Npys)-OH in deuterated acetone/ DCl. Due to limited solubility, DCl was added, however the Cys(Npys)-OH was not stable in this solution and decomposed quickly. Improved solubility was later discovered in a acetonitrile/ water mixture. The NH_3 and COOH groups were observable at this low pD, and these peaks verified by carrying out a ^{1}H NMR of cysteine in deuterated acetone/ DCl.

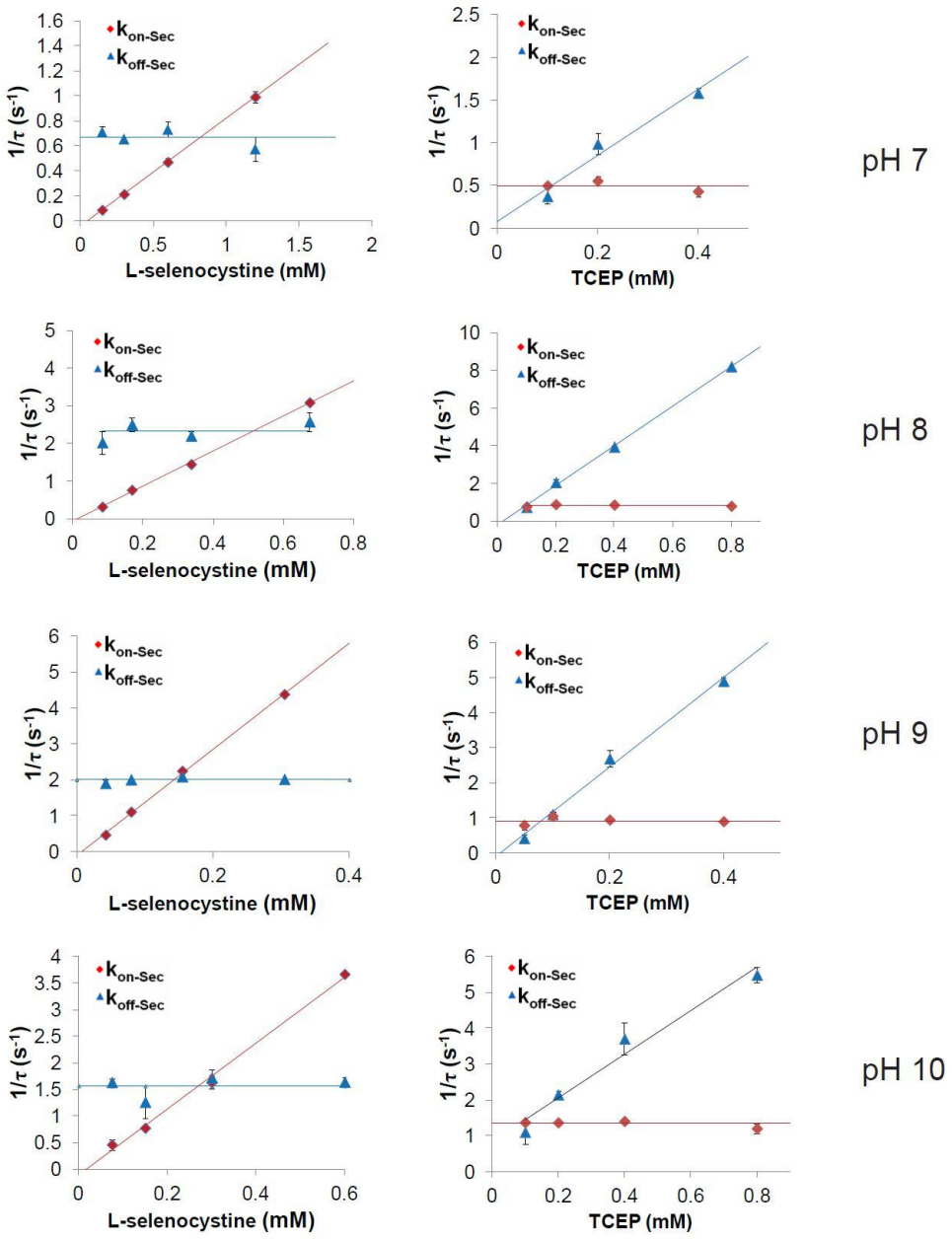


Figure A.11: The relationship between $1/\tau$ (i.e. 1/(dwell time)) and the concentration of selenocystine (Sec_2) and the concentration of TCEP.

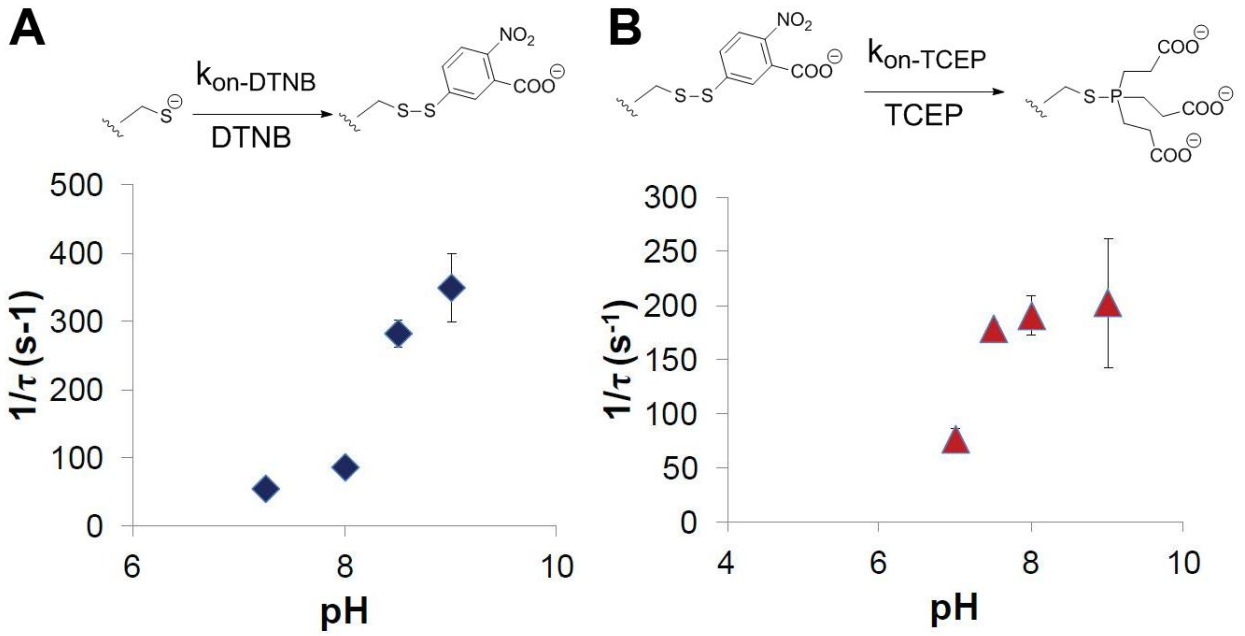


Figure A.12: **A** The rate constants ($k_{on-DTNB}$) for the reaction of DTNB (*trans*) with 117CysSECIS$_1$WT$_6$ plotted against pH. **B** The rate constants ($k_{on-TCEP}$) for the reaction of TCEP (*cis*) with 117CysSECIS$_1$WT$_6$ over the pH range 7 - 9.

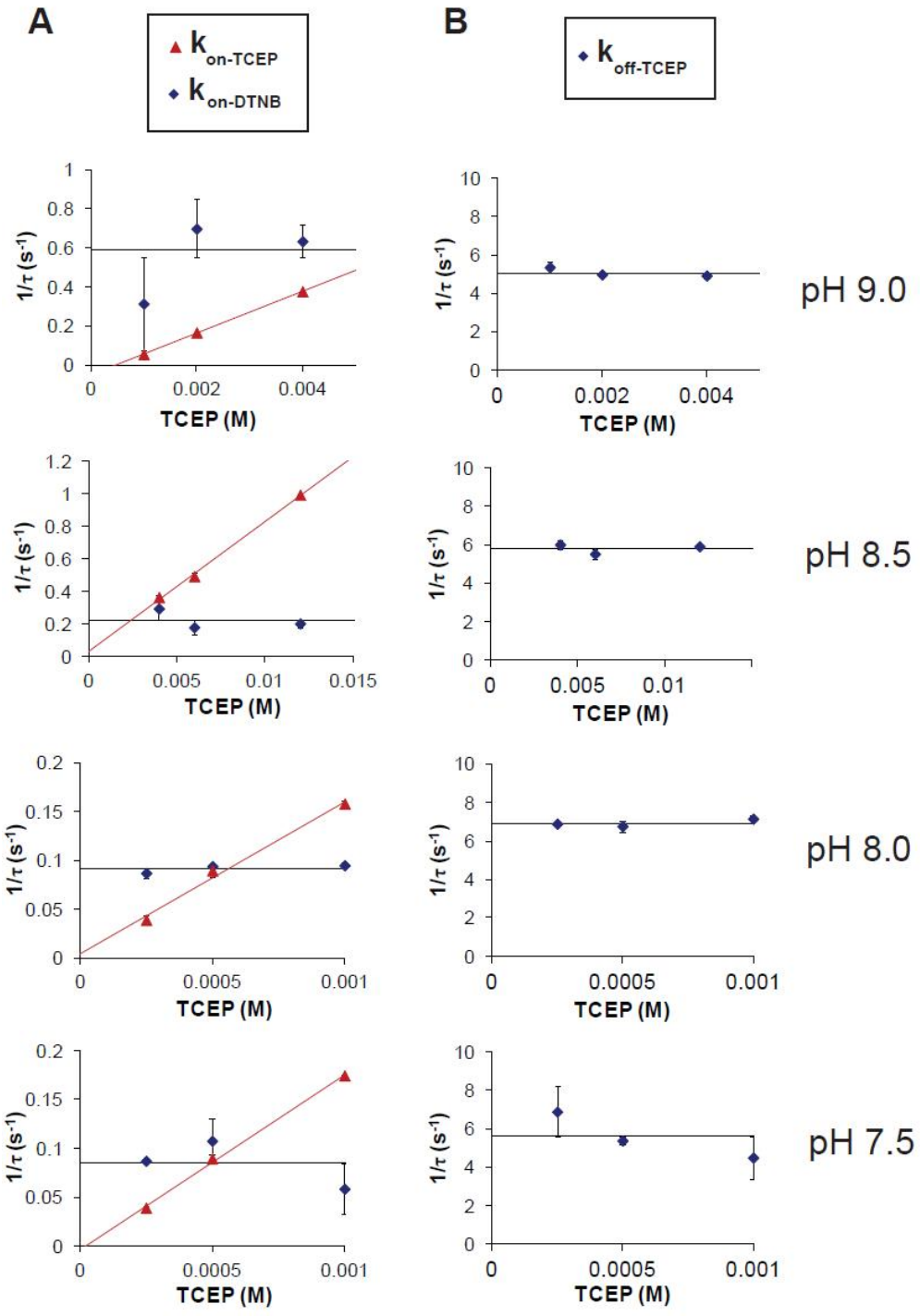


Figure A.13: Titration experiments carried out over the pH range 7 - 9. **A** The relationship between $1/\tau$ (s^{-1}) and concentration of TCEP at each pH. The rate constant for the reaction of $117CysSECIS_1WT_6$ and DTNB ($k_{on-DTNB}$) was calculated from and average of the individual $1/\tau$ values divided by the concentration of DTNB present (200μM at pH 8,5 and 9.0, 1 mM at pH 7.5, *trans*). **B** The phosphonium cation life time ($1/\tau$ s^{-1}) plotted against concentration of TCEP. It can be seen that the life time of the cationic intermediate was not dependent on the TCEP concentration and also the $k_{off-TCEP}$ was comparable over the pH range 7.5 - 9.0.

pH	$k_{on\text{-}sec}$ ($M^{-1}s^{-1}$)	$k_{off\text{-}sec}$ ($M^{-1}s^{-1}$)
6	(2.8 ± 0.7) x10^2	(8.8 ± 3.7) x 10^2
7	(1.1 ± 0.2) x 10^3	(4.2 ± 0.4) x 10^3
8	(4.1 ± 0.4) x 10^3	(1.1 ± 0.1) x 10^4
9	(1.34 ± 0.09) x 10^4	(1.20 ± 0.08) x 10^4

pH	$k_{on\text{-}Cys(Npys)}$ ($M^{-1}s^{-1}$)	$k_{off\text{-}Cys}$ ($M^{-1}s^{-1}$)
9	(1.9 ± 0.2) x 10^2	1.2 ± 0.3

pH	$k_{on\text{-}DTNB}$ ($M^{-1}s^{-1}$)	$k_{on\text{-}TCEP}$ ($M^{-1}s^{-1}$)
7	(2.7 ± 0.3) x 10^1	(8 ± 1) x 10^1
8	(9 ± 1) x 10^1	(1.9 ± 0.2) x 10^2
8.5	(2.8 ± 0.2) x 10^2	n/a
9	(3.5 ± 0.5) x 10^2	(2 ± 0.6) x 10^2

pH	$k_{off\text{-}TCEP}$ (s^{-1})
7 - 9	7 ± 1

Table A.2: A summary of the rate constants obtained from reaction of 117CysSECIS$_1$WT$_6$ with DTNB or selenocystine and TCEP over the pH range 7 - 10.

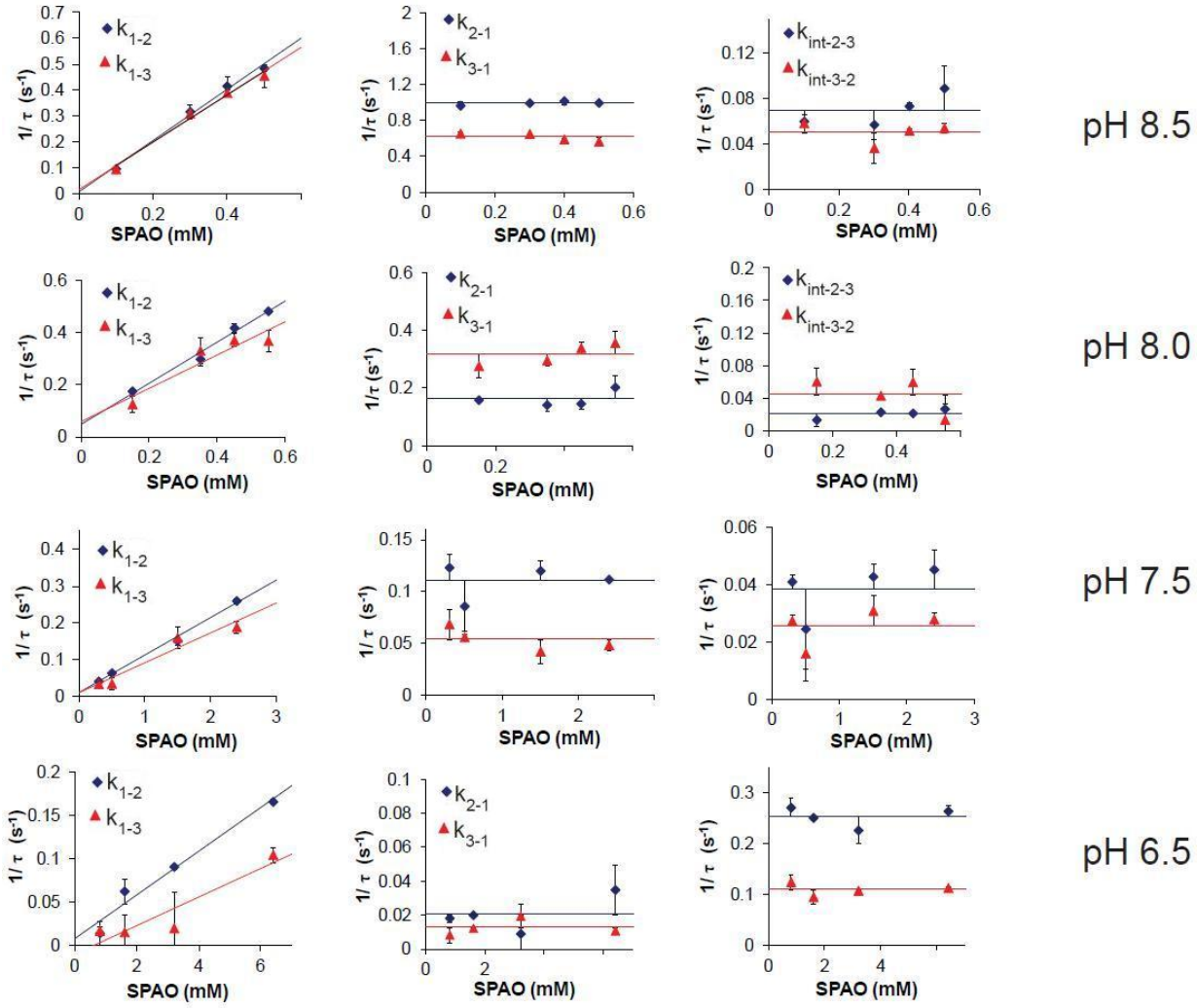


Figure A.14: Titration experiments for the reaction between SPAO and 117CysSECIS$_1$WT$_6$ over the pH range 6.5 - 8.5. The k_{on} for the formation of each diastereomeric PS-SPAO species have been coloured red and blue (LHS graph). The k_{off} is the rate constant for the reverse reaction from the PS-SPAO to PS^- (middle). The interconversion between the two diastereomeric species was observable and so the rate constants k_{int} were calculated (RHS).

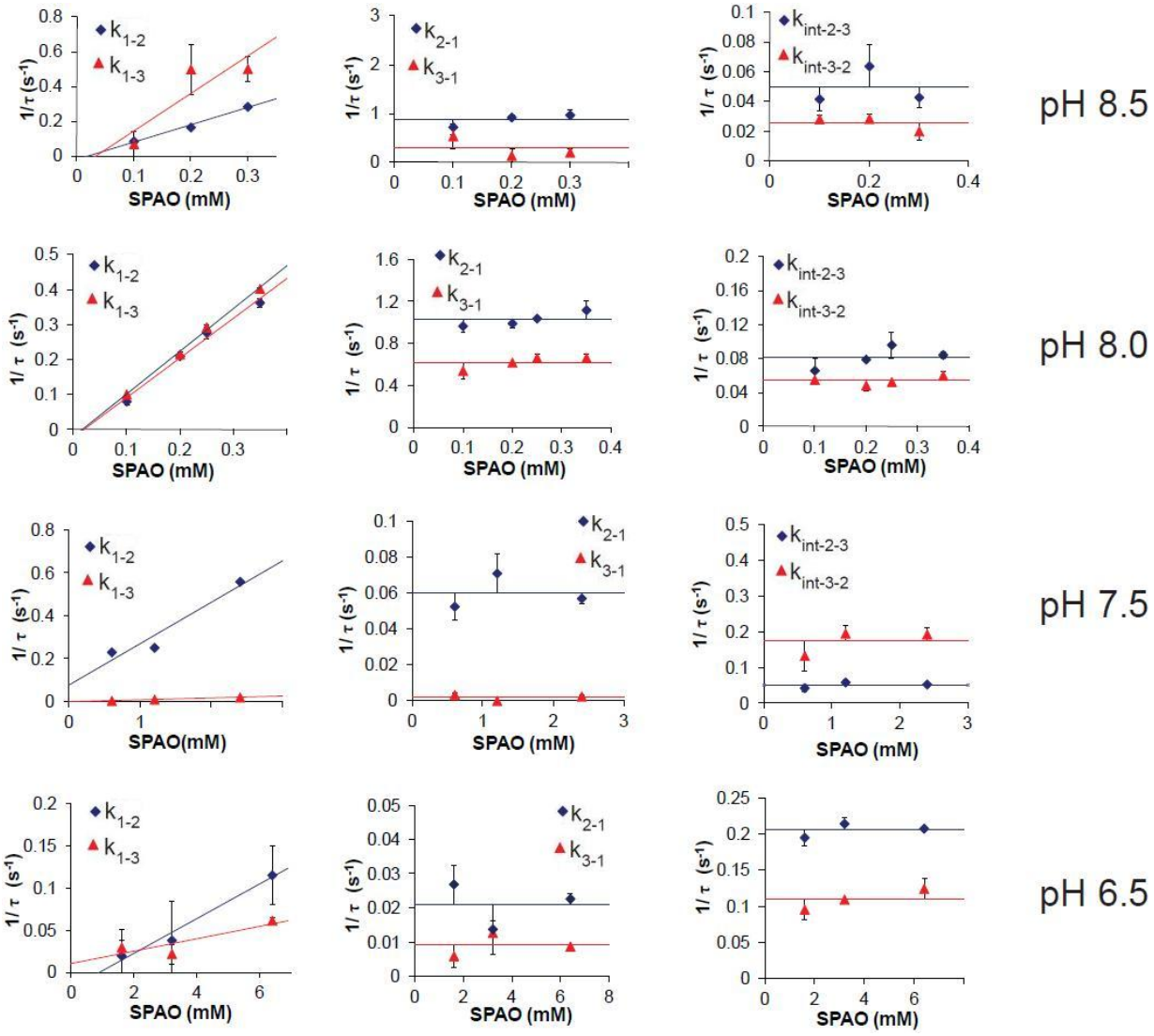


Figure A.15: Titration experiments for the reaction between SPAO and 117SecSECIS$_1$WT$_6$ over the pH range 6.5 - 8.5. The k_{on} for the formation of each diastereomeric PSe-SPAO species have been coloured red and blue (LHS graph). The k_{off} is the rate constant for the reverse reaction from the PSe-SPAO to PSe^- (middle). The interconversion between the two diastereomeric species was observable and so the rate constants k_{int} were calculated (RHS).

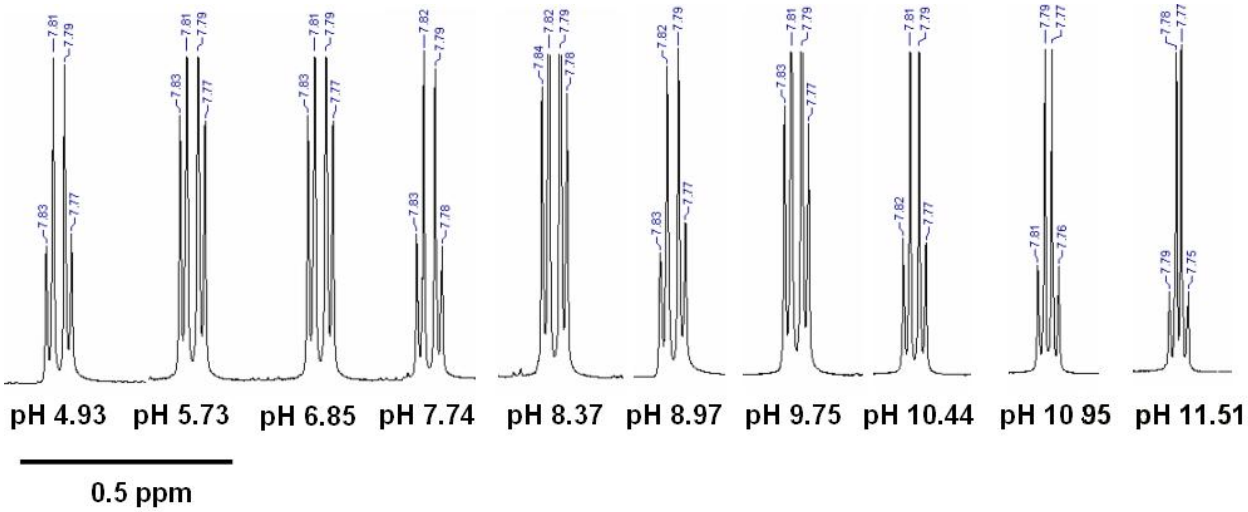


Figure A.16: The ^{1}H NMR spectral data gathered for SPAO over the pD range 5 - 13.

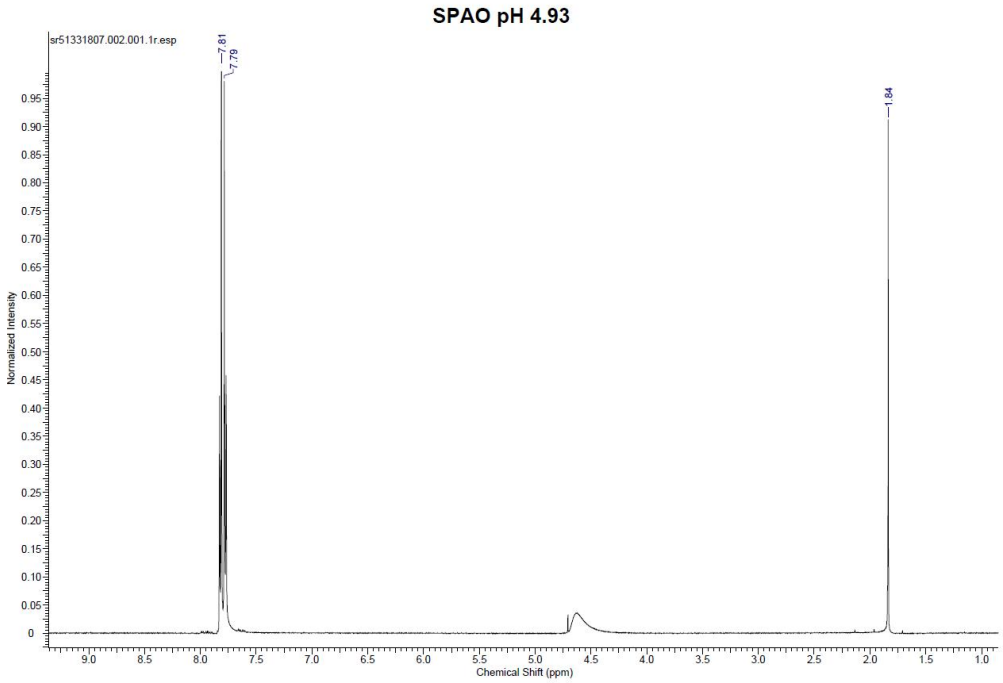


Figure A.17: The ^{1}H NMR spectrum of SPAO at pD 5.3.

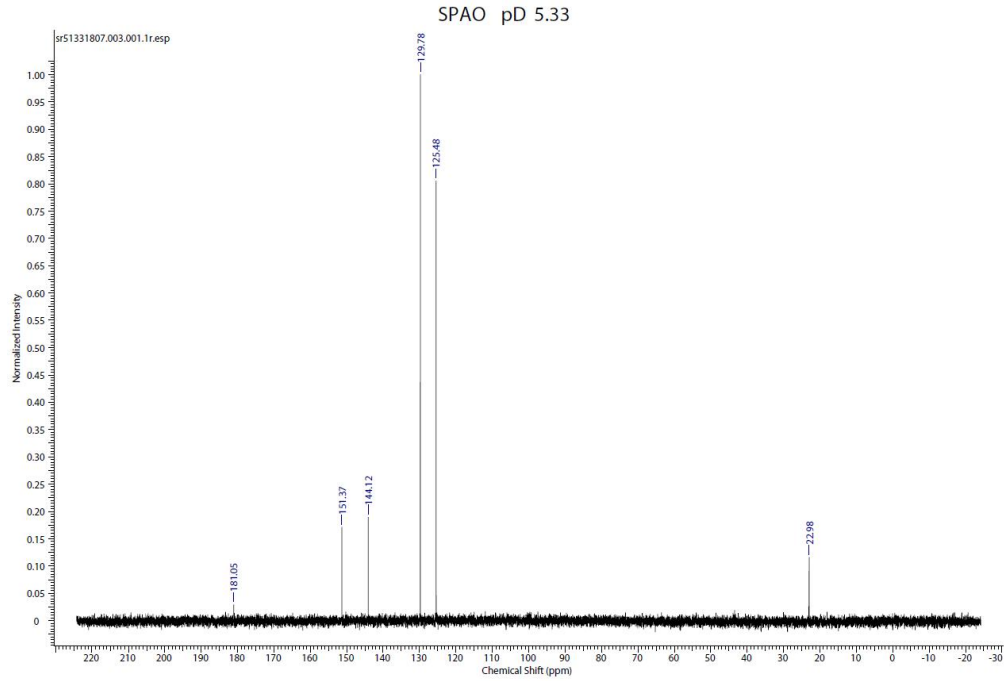


Figure A.18: The ^{13}C NMR spectral data gathered for SPAO at pD 5.33. (Acetic acid was used as the reference peak).

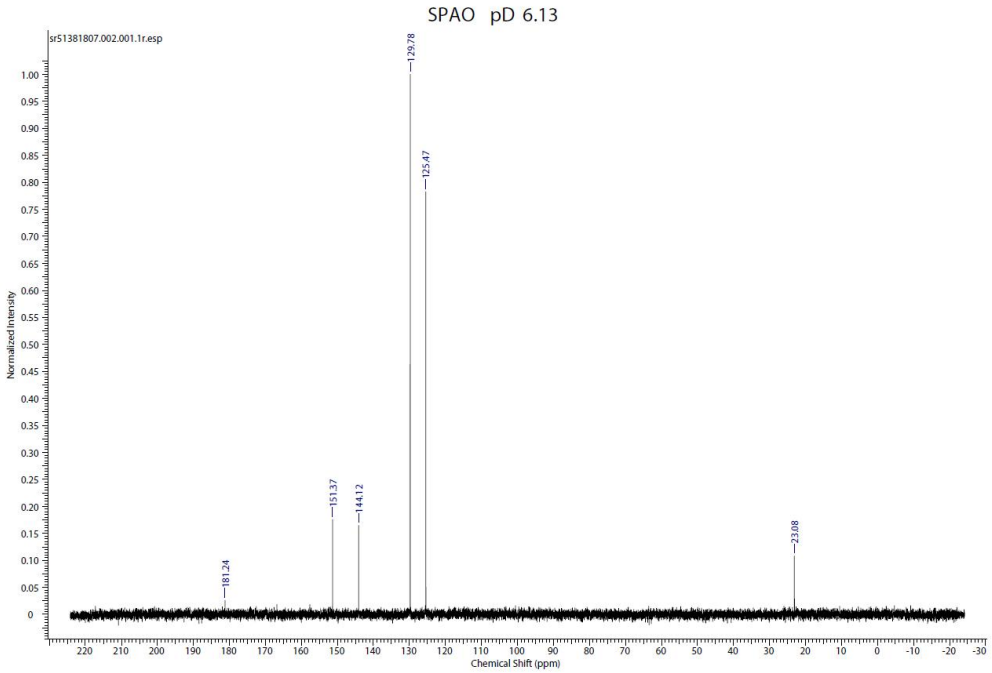


Figure A.19: The ^{13}C NMR spectral data gathered for SPAO at pD 6.13. (Acetic acid was used as the reference peak).

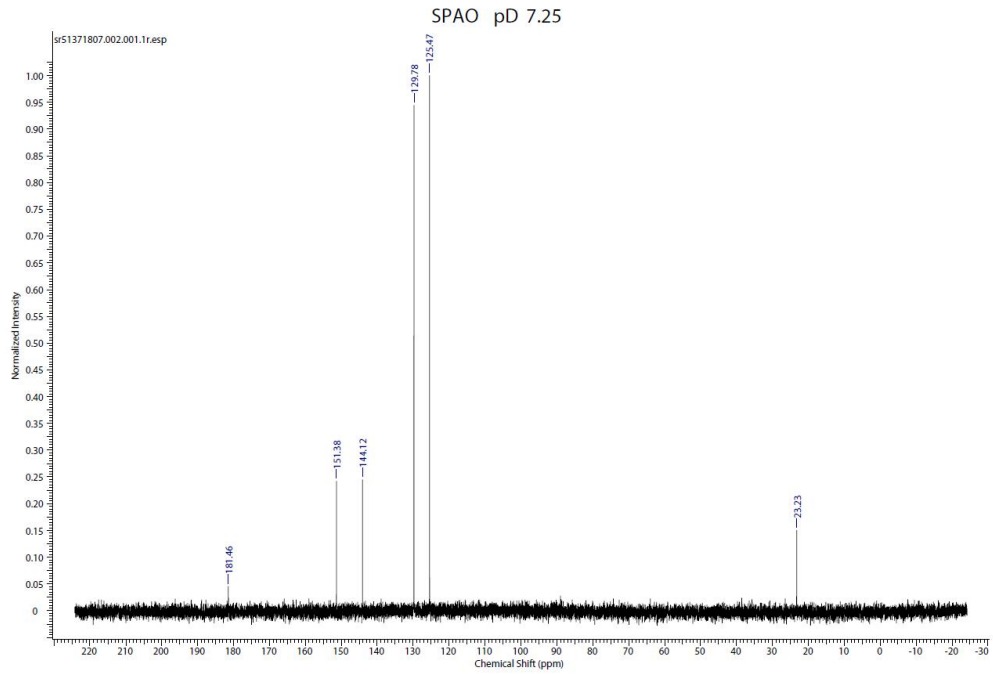


Figure A.20: The ^{13}C NMR spectral data gathered for SPAO at pD 7.25. (Acetic acid was used as the reference peak).

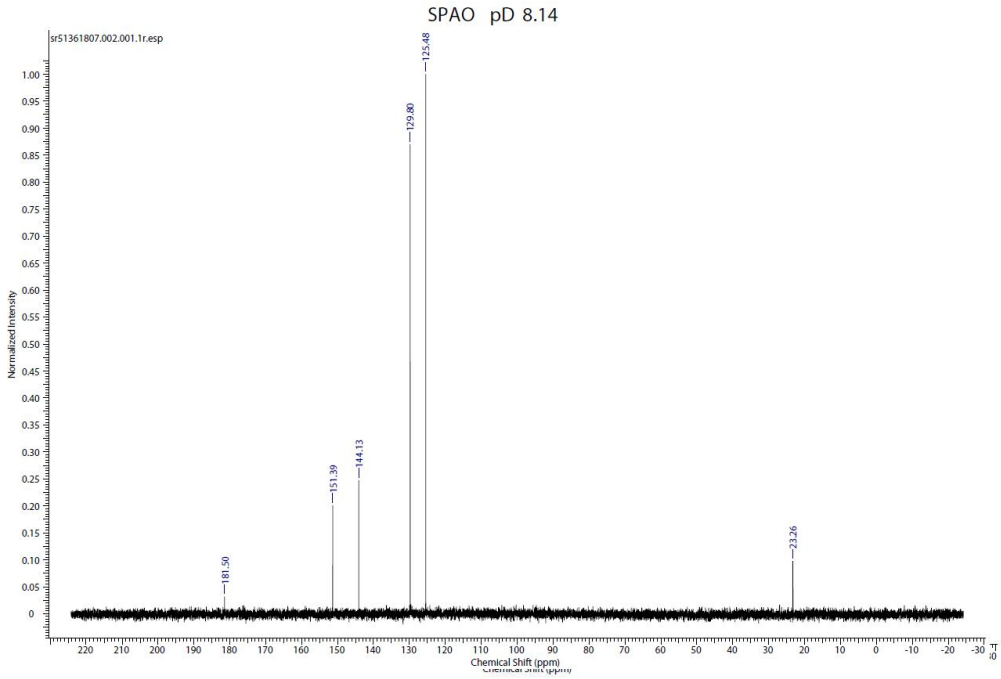


Figure A.21: The ^{13}C NMR spectral data gathered for SPAO at pD 8.14. (Acetic acid was used as the reference peak).

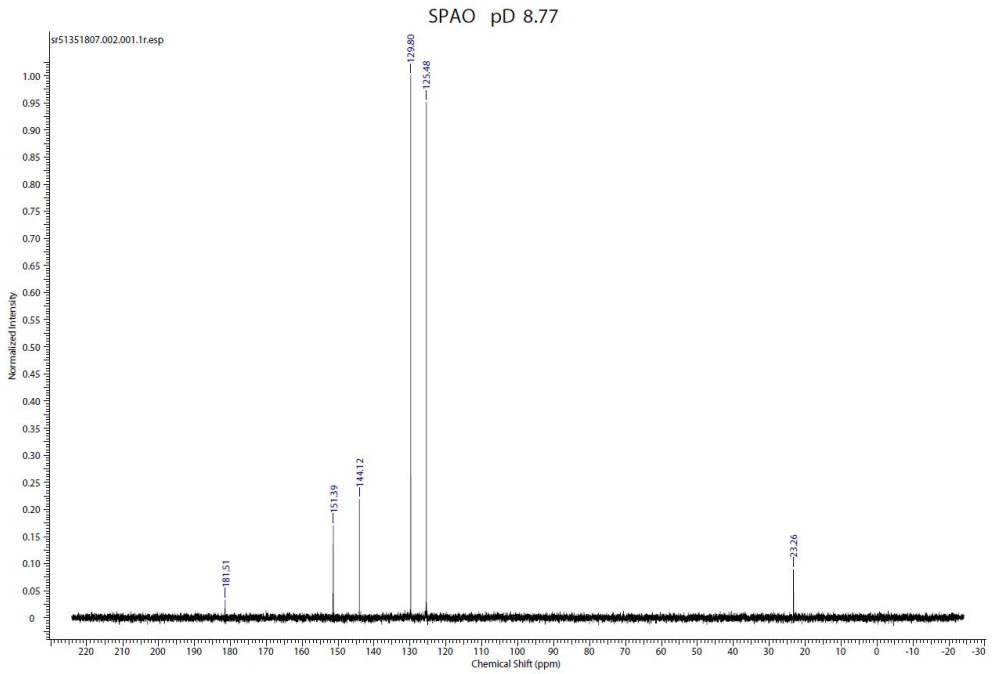


Figure A.22: The ^{13}C NMR spectral data gathered for SPAO at pD 8.77. (Acetic acid was used as the reference peak).

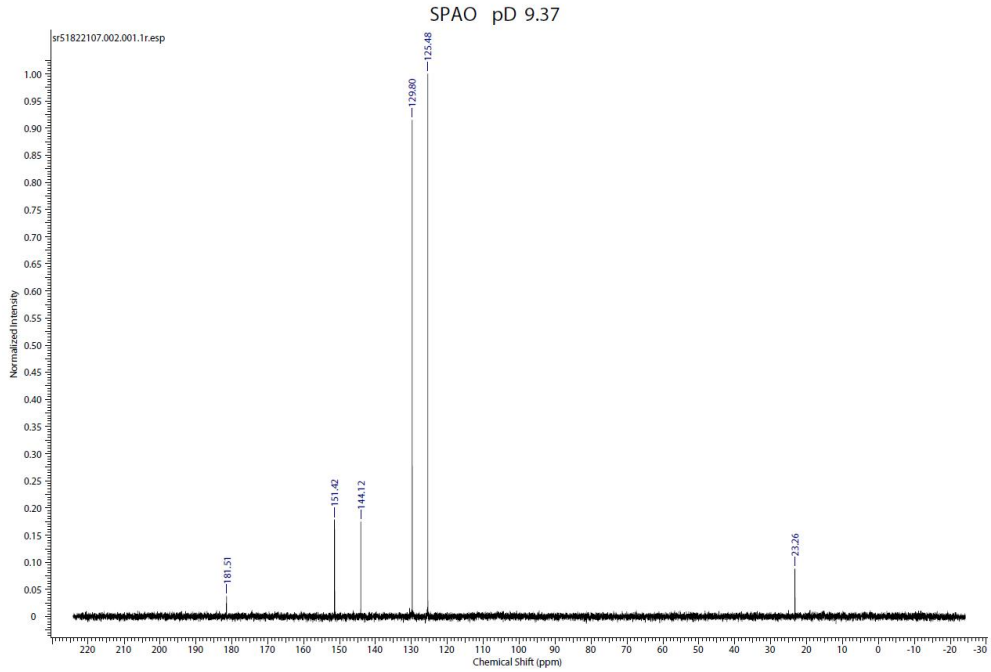


Figure A.23: The ^{13}C NMR spectral data gathered for SPAO at pD 9.37. (Acetic acid was used as the reference peak).

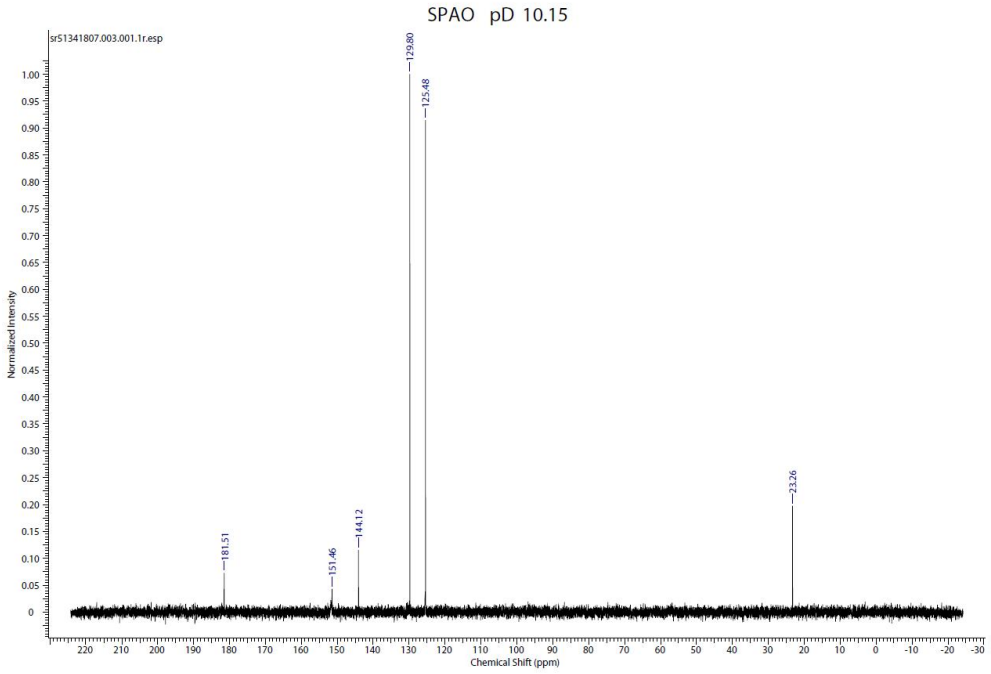


Figure A.24: The ^{13}C NMR spectral data gathered for SPAO at pD 10.15. (Acetic acid was used as the reference peak).

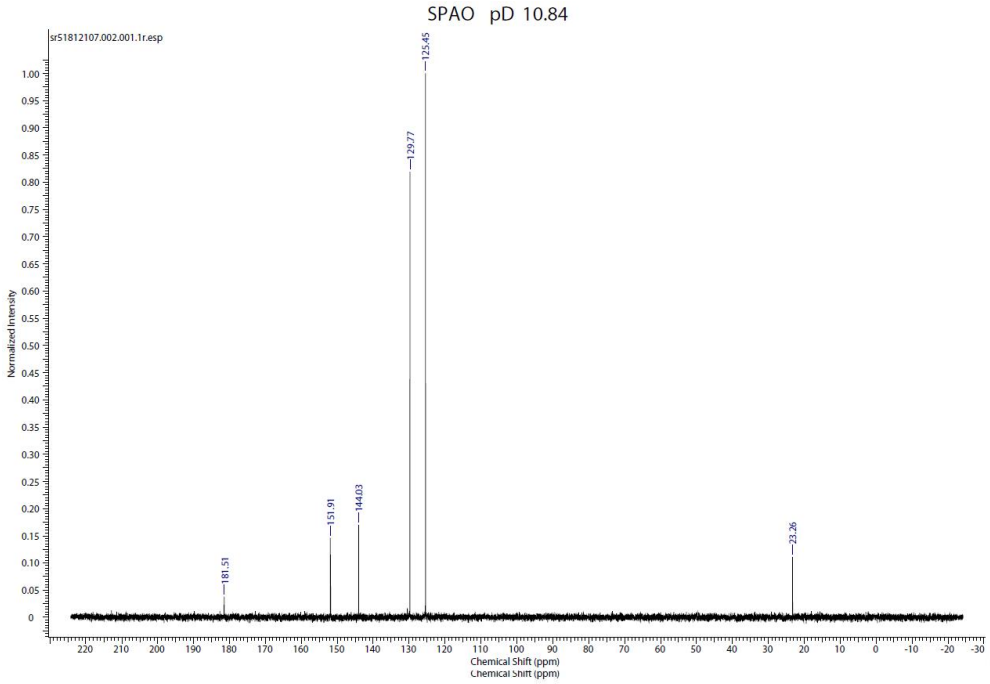


Figure A.25: The ^{13}C NMR spectral data gathered for SPAO at pD 10.84. (Acetic acid was used as the reference peak).

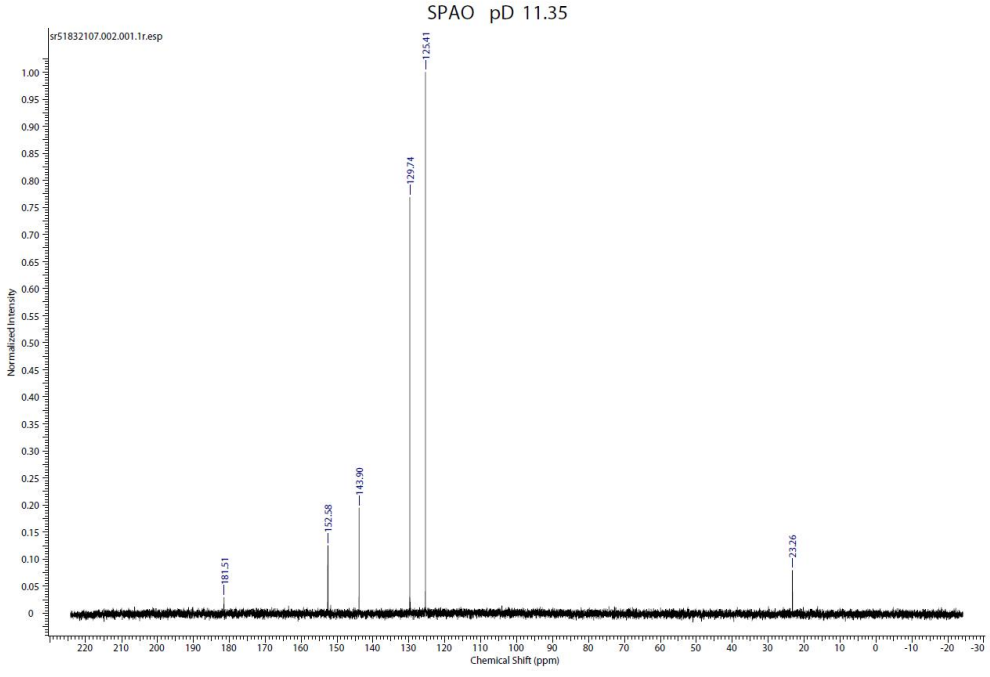


Figure A.26: The ^{13}C NMR spectral data gathered for SPAO at pD 11.35. (Acetic acid was used as the reference peak).

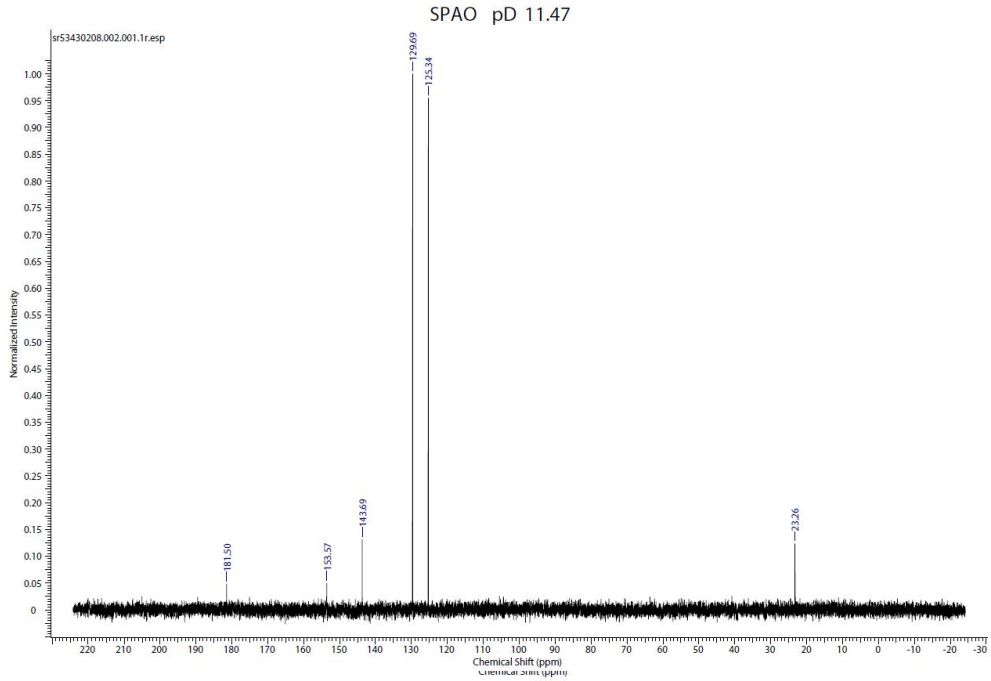


Figure A.27: The ^{13}C NMR spectral data gathered for SPAO at pD 11.47. (Acetic acid was used as the reference peak).

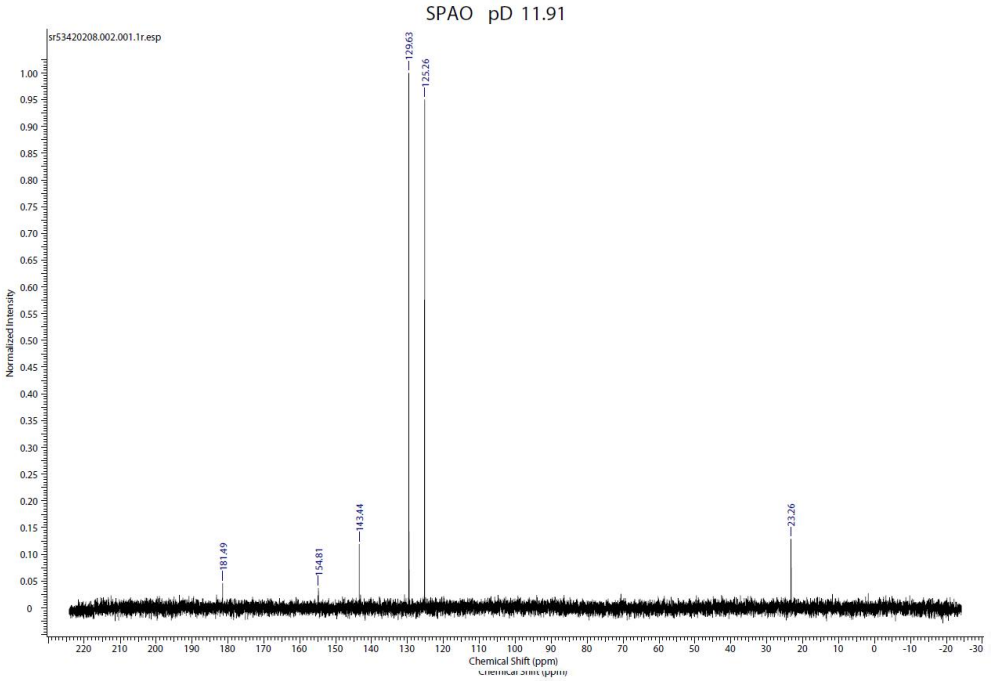


Figure A.28: The ^{13}C NMR spectral data gathered for SPAO at pD 11.91. (Acetic acid was used as the reference peak).

Bibliography

ALDAG, C., GROMOV, I. A., GARCIA-RUBIO, I., VON KOENIG, K., SCHLICHTING, I., BERNHARD, J. & HILVERT, D., 2009. Probing the role of the proximal heme ligand in cytochrome P450cam by recombinant incorporation of selenocysteine. *Proc. Natl. Acad. Sci.*, **106**, 5481–5486.

ALLEN, F. H., 2002. The Cambridge structural database: a quarter of a million crystal structures and rising. *Acta Crystallogr.*, 380–388.

ALLRED, A. L., 1961. Electronegativity values from thermochemical data. *J. Inorg. Nucl. Chem.*, **17**, 215–221.

ANDERSON, J. C., WU, N., SANTORO, S. W., LAKSHMAN, V., KING, D. S. & SCHULTZ, P. G., 2005. An expanded genetic code with a functional quadruplet codon. *Proc. Nat. Acad. Sci.*, **101**, 7566–7571.

ARNÉR, E. S. G., 2010. Selenoproteins-What unique properties can arise with selenocysteine in place of cysteine? *Experimental Cell Research*, **316**, 1296–1303.

ARNÉR, E. S. J., 2002. Selenoproteins. *Meth. Enzymol.*, **347**, 226–235.

ARNÉR, E. S. J. & HOLMGREN, A., 2000. Physiological functions of thioredoxin and thioredoxin reductase. *Eur. J. Biochem.*, **267**, 6102–6109.

ARNÉR, E. S. J., SARIOGLU, H., LOTTSPEICH, F., HOLMGREN, A. & A., B., 1999. High-level Expression in Escherichia coli of Selenocysteine-containing Rat Thioredoxin Reductase Utilizing Gene Fusions with Engineered Bacterial-type SECIS Elements and Co-expression with the SelA, SelB and SelC Genes. *J. Mol. Biol.*, **292**, 1003–1016.

ARNOLD, A. P., TAN, K.-S. & RABENSTEIN, D. L., 1986. Nuclear Magnetic Resonance Studies of the Solution Chemistry of Metal Complexes. 23. Complexation of Methylmercury by Selenohydryl-Containing Amino Acids and Related Molecules. *Inorg. Chem.*, **25**, 2433–2437.

AXLEY, M. J., BÖCK, A. & STADTMAN, T. C., 1991. Catalytic properties of an Escherichia coli formate dehydrogenase mutant in which sulfur replaces selenium. *Proc. Natl. Acad. Sci.*, **88**, 8450–8454.

BALTAZAR, C. S. A., MARQUES, M. C., SOARES, C. M. andDELACEY, A. M., PEREIRA, I. A. C. & MATIAS, P. M., 2011. Nickel-Iron-Selenium Hydrogenases - An Overview. *Eur. J. Inorg. Chem.*, **2011**, 948–962.

BAN, N., NISSEN, P., HANSEN, J., MOORE, P. B. & STEITZ, T. A., 2000. The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4Å Resolution. *Science*, **289**, 905–920.

BANERJEE, A., MIKHAILOVA, E., CHELEY, S., GUB, L. Q., MONTOYA, M., NAGAOKA, Y., GOUAUX, E. & BAYLEY, H., 2010. Molecular bases of cyclodextrin adapter interactions with engineered protein nanopores. *Proc. Natl. Acad. Sci.*, **107**, 8165Ű8170.

BARON, C., HEIDER, J. & BÖCK, A., 1990. Mutagenesis of selC, the gene for the selenocysteine inserting tRNA-species in *E.coli*: effects on in vivo function. *Mol. Microbiol.*, **18**, 6761–6766.

BERKOWITZ, J., ELLISON, G. B. & GUTMAN, D., 1994. Methods to measure Rh bond energies. *J. Phys. Chem.*, **98**, 2744–2765.

BERRY, S. M., GIESELMAN, M. D., NILGES, M. J., VAN DER DONK, W. A. & LU, Y., 2002. An Engineered Azurin Variant Containing a Selenocysteine Copper Ligand. *J. Am. Chem. Soc.*, **124**, 2084–2085.

BETZIG, E. & CHICHESTER, R. J., 1993. Single Molecules Observed by Near-Field Scanning Optical Microscopy. *Science*, **262**, 1422–1425.

BIAGLOW, J. E. & MILLER, R. A., 2005. The Thioredoxin Reductase/Thioredoxin System. *Cancer Biology and Therapy*, **4**, 6–13.

BINNIG, G., QUATE, C. F. & GERBER, C., 1986. Atomic Force Microscopy. *Phys. Rev. Lett.*, **56**, 930–934.

BJORNSTEDT, M., HAMBERG, M., KUMAR, S., XUE, J. & HOLMGREN, A., 1995. Human Thioredoxin Reductase Directly Reduces Lipid Hydroperoxides by NADPH and Selenocystine Strongly Stimulates the Reaction via Catalytically Generated Selenols. *J. Biol. Chem.*, **270**, 11761–11764.

BLAKE, S., MAYER, T., MAYER, M. & YANG, J., 2006. Monitoring Chemical Reactions by Using Ion-Channel-Forming Peptide. *ChemBioChem*, **7**, 433Ű435.

BOCK, A., 2005. Selenocysteine. *Encyclopedia of Life Sciences*, **385**, 1–7.

BOERSMA, A. J., COQUIÈRE, D., GEERDINK, D., ROSATI, F., FERINGA, B. L. & ROELFES, G., 2010. Catalytic enantioselective syn hydration of enones in water using a DNA-based catalyst. *Nature Chem.*, **2**, 991–995.

BONDI, A., 1964. van der Waals volumes and radii. *J. Phys. Chem.*, **68**, 441–451.

BRAHA, O., WALKER, B., CHELEY, S., KASIANOWICZ, J. J., SONG, L., GOUAUX, J. E. & BAYLEY, H., 1997. Designed protein pores as components for biosensors. *Chemistry & Biology*, **4**, 497–505.

BRANTON, D., DEAMER, D. W., MARZIALI, A., BAYLEY, H., BENNER, S. A. ET AL., 2008. The potential and challenges of nanopore sequencing. *Nature Biotech.*, **28**, 1146–1153.

BRIGELIUS-FLOHE, R., 1999. Tissue-specific functions of individual Glutathione peroxidases. *Free Rad. Biol. Med.*, **27**, 951–965.

BULAJ, G., KORTEMME, T. & GOLDENBERG, D. P., 1998. Ionization-Reactivity Relationships for Cysteine Thiols in Polypeptides. *Biochemistry*, **37**, 8965–8972.

BUNNETT, J. F., 1963. Nucleophilicity Reactivity. *Annu. Rev. Phys. Chem.*, **14**, 271–290.

BURNS, J. A., BUTLER, J. C., MORAN, J. & WHITESIDES, G. M., 1991. Selective Reduction of Disulfides by Tris(2-carboxyethy1)phosphine. *J. Org. Chem.*, **56**, 2648–2650.

CABAN, K. & COPELAND, P. R., 2006. Size matters: a view of selenocysteine incorporation from the ribosome. *Cell. Mol. Life Sci.*, **63**, 0073–0081.

CHAMBERS, I., FRAMPTON, J., GOLDFARB, P., AFFARA, N., MCBAIN, W. & HARRISON, P. R., 1986. The structure of the mouse glutathione peroxidase geen: the selenocysteine in the active site is encoded by the 'termination' codon, TGA. *EMBO J.*, **5**, 1221–1227.

CHELEY, S., BRAHA, O., LU, X., CONLAN, S. & BAYLEY, H., 1999. A functional protein pore with a retro transmembrane domain. *Protein Science*, **8**, 1257–1267.

CHELEY, S., GU, L.-Q. & BAYLEY, H., 2002. Stochastic Sensing of Nanomolar Inositol 1,4,5-Trisphosphate with an Engineered Pore. *Chemistry & Biology*, **9**, 829–838.

CHENG, Q., STONE-ELANDER, S. & ARNÉR, E. S. J., 2006. Tagging recombinant proteins with a Sel-tag for purification, labeling with electrophilic compounds or radiolabeling with 13C. *Nature Protocols*, **1**, 604–613.

CHIU, S.-W. & LEAKE, M. C., 2011. Functioning Nanomachines Seen in Real-Time in Living Bacteria Using Single-Molecule and Super-Resolution Fluorescence Imaging. *Int. J. Mol. Sci.*, **12**, 2518–2542.

CHOUCHANE, S. & SNOW, E. T., 2001. In Vitro Effect of Arsenical Compounds on Glutathione-Related Enzymes. *Chem. Res. Toxicol.*, **14**, 517–522.

CHUDAKOV, D. M., MATZ, M. V., LUKYANOV, S. & LUKYANOV, K. A., 2010. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiol. Rev.*, **90**, 1103–1163.

CLEMENTI, E., RAIMONDI, D. L. & REINHARDT, W. P., 1963. Atomic screening constants from SCF functions. *J. Chem. Phys.*, **38**, 2686–2689.

CLINE, D. J., REDDING, S. E., BROHAWN, S. G., PSATHAS, J. N., SCHNEIDER, J. P. & THORPE, C., 2004. New Water-Soluble Phosphines as Reductants of Peptide and Protein Disulfide Bonds: Reactivity and Membrane Permeability. *Biochemistry*, **43**, 15195–15203.

COMMANS, S. & BOCK, A., 1999. Selenocysteine inserting tRNAs: an overview. *FEMS Microbiol. Rev.*, **23**, 335–351.

CONE, J. E., DEL RIO, R. M., DAVIS, J. N. & STADMAN, T. C., 1976. Chemical characterization of the selenoprotein component of clostridial glycine reductase: Identification of selenocysteine as the organoselenium moiety. *Proc. Natl. Acad. Sci*, **73**, 2659–2663.

COPPENS, P., REYNAERT, J. C. & DROWART, J., 1979. Mass spectrometric study of the photoionization of carbon-disulfide in the wavelength interval 125-60 nm. *J. Chem. Soc. Faraday Trans.*, **75**, 292–301.

CROTEAU, W., BODWELL, J. E., RICHARDSON, J. M. & ST. GERMAIN, D. L., 1998. Conserved Cysteines in the Type 1 Deiodinase Selenoprotein Are Not Essential for Catalytic Activity. *J. Biol. Chem.*, **273**, 25230–25236.

DEAN, J. A., 1992. *Langes Handbook of Chemistry*.

DENIZ, A. A., MUKHOPADHYAY, S. & LEMKE, E. A., 2008. Single-molecule biophysics: at the interface of biology, physics and chemistry. *J. R. Soc. Interface*, **5**, 15–45.

DONG, X., STOTHARD, P., FORSYTHE, I. J. & S., W. D., 2004. PlasMapper: a web server for drawing and auto-annotating plasmid maps. *Nucleic Acids Res.*, **1**, 660–664.

DROWART, J. & SMOES, S., 1977. Determination by the mass spectrometric Knudsen cell method and discussion of the dissociation energies of the molecules Se2(g), SSe(g) and SeTe(g). *J. Chem. Soc. Faraday Trans.*, **73**, 1755–1767.

DU MONT, W.-W., KUBINIOK, S., PETERS, K. & VON SCHNERING, H.-G., 1987. Synthesis and Structure of a Stable Iodoselenide. *Angew. Chem. lnt. Ed. Engl.*, **26**, 780–781.

EDMONDS, J. S., NAKAYAMA, T., KONDO, T. & MORITA, M., 2006. Diastereoisomerism of thiol complexes of arsenic acids and pseudoasymmetry of arsenic: a ^{1}H and ^{13}C NMR study. *Mag. Res. Chem.*, **44**, 151–62.

EHRENREICH, A., FORCHHAMMER, K., TORMAY, P., VEPREK, B. & BÖCK, A., 1992. Selenoprotein synthesis in E. coli. *Eur. J. Biochem.*, **206**, 767–773.

ENEMARK, J. H., COONEY, J. A., WANG, J.-J. & HOLM, R. H., 2004. Synthetic Analogues and Reaction Systems Relevant to the Molybdenum and Tungsten Oxotransferases. *Chem. Rev.*, **104**, 1175–1200.

ENGELHARDT, H., FORCHHAMMER, K., MÜLLER, S., GOLDIE, K. N. & BÖCK, A., 1992. Structure of selenocysteine synthasefrom *Escherichia coli* and location of tRNA in the seryl-tRNA^{sec}-enzyme complex. *Mol. Microbiol.*, **6**, 3461–3467.

FAIRWEATHER-TAIT, S. J. COLLINGS, R. & HURST, R., 2010. Selenium bioavailability: current knowledge and future research requirements. *Am. J. Clin. Nutr.*, **91**, 1484–1491.

FERNANDEZ, J. M. & LI, H., 2004. Force-Clamp Spectroscopy Monitors the Folding Trajectory of a Single Protein. *Science*, **303**, 1674–1678.

FINNEY, L. A. & O'HALLORAN, T. V., 2003. Transition Metal Speciation in the Cell: Insights from the Chemistry of Metal Ion Receptors. *Science*, **300**, 931–936.

FLOHE, L., GÜNZLER, W. A. & SCHOCK, H. H., 1973. Glutathione Peroxidase:A Selenoenzyme. *FEBS Lett.*, **32**, 132–134.

FOMENKO, D. E., XING, W., ADAIR, B. M., THOMAS, D. J. & GLADYSHEV, V. N., 2007. High-Throughput Identification of Catalytic Redox-Active Cysteine Residues. *Science*, **315**, 387–389.

FORSTROM, J. W., ZAKOWSKI, J. J. & TAPPEL, A. L., 1978. Identification of the catalytic site of rat liver glutathione peroxidase as selenocysteine. *Biochemistry*, **17**, 2639–2644.

GARCIA-MAYNES, S., LIANG, J., SZOSZKIEWICZ, R., KUO, T. L. & FERNANDEZ, J. M., 2009. Force-activated reactivity switch in a bimolecular chemical reaction. *Nature Chemistry*, **1**, 236–242.

GASTON, B., REILLY, J., DRAZEN, J. M., FACKLER, J., RAMDEVS, P., ARNELLE, D., MULLINS, M. E., SUGARBAKER, D. J., CHEE, C., SINGEL, D. J., LOSCALZO, J. & STAMLER, J. S., 1993. Endogenous nitrogen oxides and bronchodilator S-nitrosothiols in human airways. *Proc. Natl. Acad. Sci.*, **90**, 10957–10961.

GIBSON, S. T., GREENE, J. P. & BERKOWITZ, J., 1986. A photoionization study of SeH and H_2Se. *J. Chem. Phys.*, **85**, 4815–4824.

GONZALES, J. M., MUSAEV, D. G. & MOROKUMA, K., 2005. Theoretical studies of oxidative addition of E-E bonds (EsS, Se, Te) to palladium(0) and platinum(0) complexes. *Organometallics*, **24**, 4908–4914.

GOTO, K., SONODA, D., SHIMADA, K., SASE, S. & KAWASHIMA, T., 2010. Modeling of the 5Š-Deiodination of Thyroxine by Iodothyronine Deiodinase: Chemical Corroboration of a Selenenyl Iodide Intermediate. *Angew. Chem.*, **122**, 555–557.

GOUAUX, J. E., BRAHA, O., HOBAUGH, M. R., SONG, L., CHELEY, S., SHUSTAK, C. & BAYLEY, H., 1994. Subunit stoichiometry of staphylococcal alpha-hemolysin in crystals and on membranes: a heptameric transmembrane pore. *Proc. Natl. Acad. Sci.*, **91**, 12828–12831.

GOUAUX, J. E., BRAHA, O., HOBAUGH, M. R., SONG, L., CHELEY, S., SHUSTAK, C. & BAYLEY, H., 2006. Sequencing single molecules of DNA. *Curr. Op. Chem. Biol.*, **10**, 628–637.

GRANDBOIS, M., BEYER, M., RIEF, M., CLAUSEN-SCHAUMANN, H. & GAUB, H. E., 1999. How Strong Is a Covalent Bond? *Science*, **283**, 1727–1730.

GROMER, S., JOHANSSON, L., BAUER, H., ARSCOTT, L. D., RAUCH, S., BALLOU, D. P., WILLIAMS, C. H., SCHIRMER, R. H. & ARNER, E. S. J., 2003. Active sites of thioredoxin reductases: Why selenoproteins? *Pro. Natl. Acad. Sci.*, **100**, 12618–12623.

GROSSE, W., ESSEN, L.-O. & KOERT, U., 2011. Strategies and Perspectives in Ion-Channel Engineering. *ChemBioChem*, **12**, 830–839.

GU, L.-Q. & BAYLEY, H., 2000. Interaction of the Noncovalent Molecular Adapter, β-Cyclodextrin, with the Staphylococcal α-Hemolysin Pore. *Biophys. J.*, **79**, 1967–1975.

GU, L.-Q., BRAHA, O., CONLAN, S., CHELEY, S. & BAYLEY, H., 1999. Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature*, **398**, 686–690.

GU, L.-Q. & CHELEY, S. BAYLEY, H., 2001. Capture of a Single Molecule in a Nanocavity. *Science*, **291**, 636–640.

GU, L.-Q., CHELEY, S. & BAYLEY, H., 2001. Prolonged Residence Time of a Noncovalent Molecular Adapter, β-Cyclodextrin, within the Lumen of Mutant α-Hemolysin Pores. *J. Gen. Physiol.*, **118**, 481–493.

GURSINSKY, T., GRÖBE, D., SCHIERHORN, A., JANA JÖGER, J., ANDREESEN, J. R. & SÖHLING, B., 2008. Factors and Selenocysteine Insertion Sequence Requirements for the Synthesis of Selenoproteins from a Gram-Positive Anaerobe in *Escherichia coli*. *Appl. Environ. Microbiol.*, **74**, 1385–1393.

GUZIEC, F. S., 1979. *In: The Chemistry of Organic Selenium and Tellurium Compounds*.

HAMMERSTEIN, A. F., SHIN, S. H. & BAYLEY, H., 2010. Single-Molecule Kinetics of Two-Step Divalent Cation Chelation. *Angew. Chem. Int. Ed.*, **49**, 5085–5090.

HARRIS, T. K. & TURNER, G. J., 2002. Structural Basis of Perturbed pKa Values of Catalytic Groups in Enzyme Active Sites. *Life*, **53**, 85–98.

HASTINGS, C. J., PLUTH, M. D., BERGMAN, R. G. & RAYMOND, K. N., 2010. Enzyme like catalysis of the Nazarov cyclization by supramolecular encapsulation. *J Am Chem Soc*, **132**, 6938–6940.

HATFIELD, D. L. & GLADYSHEV, V. N., 2002. How Selenium Has Altered Our Understanding of the Genetic Code. *Mol. Cell Biol.*, **22**, 3565–3576.

HEINS, E. A., SIWY, Z. S., BAKER, L. A. & MARTIN, C. R., 2005. Detecting Single Porphyrin Molecules in a Conically Shaped Synthetic Nanopore. *Nano Lett.*, **5**, 1824–1827.

HENDRICKSON, T. L., 2007. Easing selenocysteine into proteins. *Nature*, **14**, 100–101.

HILGENFELD, R., BÖCK, A. & WILTING, R., 1996. Structural model for the selenocysteine-specific elongation factor SelB. *Biochimie*, **78**, 971–978.

HOFMANN, R. M. & MUIR, T. W., 2002. Recent advances in the application of expressed protein ligation to protein engineering. *Curr. Op. Biotech.*, **13**, 297–303.

HOLMGREN, A. & LU, J., 2010. Thioredoxin and thioredoxin reductase: Current research with special reference to human disease. *Biochem. Biophys. Res. Comm.*, **396**, 120–124.

HOUSE, K. L., DUNLAP, R. B., ODOM, J. D., WU, Z.-P. & HILVERT, D., 1992. Structural Characterization of Selenosubtilisin by ^{77}Se NMR Spectroscopy. *J. Am. Chem. Soc.*, **114**, 8573–8579.

HOWORKA, S. & BAYLEY, H., 2002. Probing Distance and Electrical Potential within a Protein Pore with Tethered DNA. *Biophys. J.*, **83**, 3202–3210.

HOWORKA, S., NAM, J., BAYLEY, H. & KAHNE, D., 2004. Stochastic Detection of Monovalent and Bivalent Protein-Ligand Interactions. *Angew. Chem. Int. Ed.*, **43**, 842–842.

HUANG, X., DONG, Z. Y., LIU, J. Q., MAO, S. Z., XU, J. Y., LUO, G. M. & SHEN, J. C., 2007. Selenium-mediated micellar catalyst: an efficient enzyme model for glutathione peroxidase-like catalysis. *Langmuir*, **23**, 1518–1522.

HUBER, K. P. & HERZBERG, G., 1979. *(1979). In: Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules*.

HUBER, R. E. & CRIDDLE, R. S., 1967. Comparison of the Chemical Properties of Selenocysteine and Selenocystine with Their Sulfur Analogs. *Archive Biochem. Biophys.*, **122**, 164–173.

HUHEEY, J. E., KEITER, E. A. & KEITER, R. L., 1993. Inorganic Chemistry: Principles of Structure and Reactivity.

JACOB, C., GILES, G. L., GILES, N. M. & SIES, H., 2003. Sulfur and selenium: the role of oxidation state in protein structure and function. *Angew. Chem. Int. Ed.*, **42**, 4742–4758.

JAIKARAN, D. C. J. & WOOLLEY, G. A., 1995. Characterization of Thermal Isomerization at the Single Molecule Level. *J. Phys. Chem.*, **99**, 13352–13355.

JAMES, A. M. & LORD, M. P., 1992. Macmillans Chemical and Physical Data.

JARAMILLO, P., PEREZ, P. & FUENTEALBA, P., 2007. Relationship between basicity and nucleophilicity. *J. Phys. Org. Chem.*, **20**, 1050–1057.

JIANG, Y., TRNKA, M. J., MEDZIHRADSZKY, K. F., OUELLET, H., WANG, Y. & ORTIZ DE MONTELLANO, P. R., 2009. Covalent heme attachment to the protein in human heme oxygenase-1 with selenocysteine replacing the His25 proximal iron ligand. *J. Inorg. Biochem.*, **103**, 316–325.

JOHNS, J. W. C. & RAMSAY, D. A., 2007. The absorption spectrum and dissociation energy of SH. *Can. J. Phys.*, **39**, 210–217.

KANG, X.-F., GU, L.-Q., CHELEY, S. & BAYLEY, H., 2000. Single Protein Pores Containing Molecular Adapters at High Temperatures. *Angew. Chem. Int. Ed.*, **117**, 1519–1523.

KANZOK, S. M., FECHNER, A., BAUER, H., ULSCHMID, J. K., MULLER, H.-M., BOTELLA-MUNOZ, J., SCHNEUWLY, S., SCHIRMER, R. H. & BECKER, K., 2001. Substitution of the Thioredoxin System for Glutathione Reductase in Drosophila melanogaster. *Science*, **291**, 643–646.

KAPANIDIS, A. N. & STRICK, T., 2009. Biology, one molecule at a time. *Cell*, **34**, 234–243.

KASIANOWICZ, J. J., BURDEN, D. L., HAN, L. C., CHELEY, S. & BAYLEY, H., 1999. Genetically Engineered Metal Ion Binding Sites on the Outside of a Channels Transmembrane β-Barrel. *Biophys. J.*, **76**, 837–845.

KENT, S. B. H., 2009. Total chemical synthesis of proteins. *Chem. Soc. Rev.*, **38**, 338–351.

KREZEL, A., LATAJKA, R., BUJACZ, G. D. & BAL, W., 2003. Coordination Properties of Tris(2-carboxyethyl)phosphine, a Newly Introduced Thiol Reductant, and Its Oxide. *J. Org. Chem.*, **42**, 1994–2003.

KROMAYER, M., WILTING, R., TORMAY, P. & BÖCK, A., 1996. Domain Structure of the Prokaryotic Selenocysteine-specific Elongation Factor SelB. *J. Mol. Biol.*, **262**, 413–420.

KRYUKOV, G. V., CASTELLANO, S., NOVOSELOV, S. V., LOBANOV, A. V., ZEHTAB, O., GUIGO, R. & GLADYSHEV, V. N., 2003. Characterization of Mammalian Selenoproteomes. *Science*, **300**, 1439–1449.

LADENSTEIN, R. & WENDEL, A., 1983. The Refined Structure of the Selenoenzyme Glutathione Peroxidase at 0.2-nm Resolution. *Eur. J. Biochem.*, **133**, 51–69.

LAMBERT, J. B., JACKSON, G. F. & MUELIER, D. C., 1968. The stereochemical lability of diphosphines and diarsines. *J. Am. Chem. Soc.*, **90**, 6401–6405.

LAVERMAN, P., BOERMAN, O. C., CORSTENS, F. H. M. & OYEN, W. J. G., 2002. Fluorinated amino acids for tumour imaging with positron emission tomography. *Eur. J. Nucl. Med.*, **29**, 681–690.

LEE, H. S., GUO, J., LEMKE, E. A., DIMLA, R. D. & SCHULTZ, P. G., 2009. Genetic Incorporation of a Small, Environmentally Sensitive, Fluorescent Probe into Proteins in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.*, **131**, 12921–12923.

LEE, H. S. & SCHULTZ, P. G., 2008. Biosynthesis of a Site-Specific DNA Cleaving Protein. *J. Am. Chem. Soc.*, **130**, 13194–13195.

LEIBUNDGUT, M., FRICK, C., THANBICHLER, M., BÖCK, A. & BAN, N., 2005. Selenocysteine tRNA-specific elenogation factor SelB is a structural chimaera of elnogation and initiation factors. *EMBO J.*, **24**, 11–22.

LEINFELDER, W., FORCHHAMMER, K., ZINONI, F., SAWERS, G., MANDRAND-BERTHELOT, M. & BOCK, A., 1988a. Escherichia coli Genes Whose Products Are Involved in Selenium Metabolism. *J. Bacteriol.*, **170**, 540–546.

LEINFELDER, W., STADTMAN, T. C. & BÖCK, A., 1989. Occurence in Vivo of Selenocysteyl-tRNA^{Ser} in *Escherichia coli*. *J. Biol. Chem.*, **264**, 9720–9723.

LEINFELDER, W., ZEHELEIN, E., MANDRAND-BERTHELOT, M. A. & BÖCK, A., 1988b. Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature*, **331**, 723–725.

LEWIN, A., CROW, A., OUBRIE, A. & LE BRUN, N. E., 2006. Molecular basis for specificity of the extra-cytoplasmic thioredoxin ResA. *J. Biol. Chem.*, **281**, 35467–35477.

LIDE, D. R., 1998. *Chemical Rubber Company Handbook of Chemistry and Physics*.

LINDSAY, S. M., THUNDAT, T., NAGAHARA, L., KNIPPING, U. & RILL, R. L., 1989. Images of the DNA Double Helix in Water. *Science*, **244**, 1063–1064.

LIU, Z., RECHES, M., GROISMAN, I. & ENGELBERG-KULKA, H., 1998. The nature of the minimal 'selenocysteine insertion sequence' (SECIS) in Escherichia coli. *Nucleic Acids Res.*, **26**, 896–902.

LOBANOV, A. V., HATFIELD, D. L. & GLADYSHEV, V. N., 2009. Eukaryotic selenoproteins and selenoproteomes. *Biochimica et Biophysica Acta*, **1790**, 1424–1428.

LOUDWIG, S. & BAYLEY, H., 2006. Photoisomerisation of an individual azobenzene molecule in water: an on-off switch triggered by light at a fixed wavelength. *J. Am. Chem. Soc.*, **128**, 12404–12405.

LOW, S. C. & BERRY, M. J., 1996. Knowing when not to stop: selenocysteine incorporation in eukaryotes. *Trends in Biochem. Sci.*, **21**, 203–208.

LU, H. P., XUN, L. & XI, X. S., 1998. Single-Molecule Enzymatic Dynamic. *Science*, **282**, 1877–1882.

LU, S., LI, W. W. ROTEM, D., MIKHAILOVA, E. & BAYLEY, H., 2010. A primary hydrogenŰdeuterium isotope effect observed at the single-molecule level. *Nature Chemistry*, **2**, 921–928.

LUCHIAN, T., SHIN, S.-S. & BAYLEY, H., 2003a. Kinetics of a three-step reaction observed at the single molecule level. *Angew. Chem. Int. Ed.*, **42**, 1926–1929.

LUCHIAN, T., SHIN, S.-S. & BAYLEY, H., 2003b. Single-Molecule Covalent Chemistry with Spatially Separated Reactants. *Angew. Chem. Int. Ed.*, **42**, 3766–3771.

LUNDROM, J. & HOLMGREN, A., 1990. Protein Disulfide-isomerase Is a Substrate for Thioredoxin Reductase and has Thioredoxin-like Activity. *J. Biol. Chem.*, **265**, 9114–9120.

MAGLIA, G., HERON, A. J., STODDART, D., JAPRUNG, D. & BAYLEY, H., 2010. Chapter 22 Ű Analysis of Single Nucleic Acid Molecules with Protein Nanopores. *Meth. Enzymol.*, **457**, 591Ű623.

MANDAL, B. K. & KAZUO, T. S., 2001. Monomethylarsonous Acid (MMA^{III}) and Arsenite: LD50 in Hamsters and In Vitro Inhibition of Pyruvate Dehydrogenase. *Chem. Res. Toxicol.*, **14**, 651–656.

MANDAL, B. K. & KAZUO, T. S., 2002. Arsenic around the world: a review. *Talanta*, **58**, 201–235.

MANLEY, S. A., GEORGE, G. N., PICKERING, I. J., GLASS, R. S., PRENNER, E. J., YAMDAGNI, R., WU, Q. & GAILER, J., 2006. The Seleno Bis(S-glutathionyl) Arsinium Ion Is Assembled in Erythrocyte Lysate. *Chem. Res. Toxicol.*, **19**, 601–607.

MANSELL, J. B., GUEVREMONT, D., POOLE, E. S. & TATE, W. P., 2001. A dynamic competition between release factor 2 and the tRNAŜec decoding UGA at the recoding site of *Escherichia coli* formate dehydrogenase H. *The EMBO J.*, **20**, 7284–7293.

MARK, D. F. & RICHARDSON, C. C., 1976. Escherichia coli thioredoxin: A subunit of bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci.*, **73**, 780–784.

MAY, J. M., MENDIRATTA, S., HILL, K. E. & BURK, R. F., 1997. Reduction of Dehydroascorbate to Ascorbate by the Selenoenzyme Thioredoxin Reductase. *J. Biol. Chem.*, **272**, 22607–22610.

METANISA, N., KEINANA, E. & DAVISON, P. E., 2006. Synthetic Seleno-Glutaredoxin 3 Analogs are Highly Reducing Oxidoreductases with Enhanced Catalytic Efficiency. *J. Am. Chem. Soc.*, **128**, 16684Ű16691.

MEZYK, S. P., 1995. Rate Constant Determination for the Reaction of Sulfhydryl Species with the Hydrated Electron in Aqueous Solution. *J. Phys. Chem.*, **99**, 13970–13975.

MILAZZO, G., CAROLI, S. & SHARMA, V. K., 1978. *Tables of Standard Electrode Potentials*.

MINDELL, J. A., ZHAN, H., HUYNH, P. D., COLLIER, R. J. & FINKELSTEIN, A., 1994. Reaction of diphtheria toxin channels with sulfhydryl-specific reagents: Observation of chemical reactions at the single molecule level. *Pro. Natl. Acad. Sci.*, **91**, 5272–5276.

MOLECULAR PROBES, I., 2011. *Introduction to fluorescence techniques*.

MONTAL, M. & MUELLER, P., 1972. Formation of Biomolecular Membranes from Lipid Monopayers and a Study of Their Electrical Properties. *PNAS*, **69**, 3561–3566.

MUGESH, G., DU MONT, W.-W., WISMACH, C. & JONES, P. G., 2002. Biomimetic Studies on Iodothyronine Deiodinase Intermediates: Modeling the Reduction of Selenenyl Iodide by Thiols. *ChemBioChem*, **3**, 440–447.

MUGESH, G. & SINGH, H. B., 2000. Synthetic organoselenium compounds as antioxidants: glutathione peroxidase activity. *Chem. Soc. Rev.*, **29**, 347–357.

MULLER, S., SENN, H., GSEL1, B., VETTER, W., BARON, C. & BÖCK, A., 1994. The Formation of Diselenide Bridges in Proteins by Incorporation of Selenocysteine Residues: Biosynthesis and Characterization of $(Se)_2$-Thioredoxin. *Biochemistry*, **33**, 3404–3412.

NELSON, J. W. & CREIGHTON, T. E., 1994. Reactivity and Ionization of the Active Site Cysteine Residues of DsbA, a Protein Required for Disulfide Bond Formation in Vivo. *Biochemistry*, **33**, 5974–5983.

NESTOROVICH, E. M., DANELON, C., WINTERHALTER, M. & BEZRUKOV, S. M., 2002. Designed to penetrate: Time-resolved interaction of single antibiotic molecules with bacterial pores. *Proc. Natl. Acad. Sci.*, **99**, 9789–9794.

NEUMAN, K. C. & NAGY, A., 2008. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature Methods*, **5**, 491–505.

NICOVICH, J. M., KREUTTER, K. D., VANDIJK, C. A. & WINE, P. H., 1992. Temperature dependent kinetics studies of the reactions $Br(^2P_{3/2})$ + H_2S reversible SH + HBr and $Br(^2P_{3/2})$ + CH_3SH reversible CH_3S + HBr - heats of formation of SH and CH_3S radicals. *J. Phys. Chem.*, **96**, 2518–2528.

NOGUCHI, T., CANTOR, A. H. & SCOTT, M. L., 1973. Mode of Action of Selenium and Vitamin E in Prevention of Exudative Diathesis in Chicks. *J. Nutr.*, **103**, 1502–1511.

NORMANLY, J., KLEINA, L. G., MASSON, J. M., ABELSON, J. & MILLER, J. H., 1990. Construction of *Escherichia coli* Amber Suppressor tRNA Genes. *J. Mol. Biol.*, **213**, 719–726.

OH, S. H., GANTHER, H. E. & WILLIAM, G., 1974. Selenium as a component of glutathione peroxidase isolated from ovine erythrocytes. *Biochemistry*, **13**, 1825–1829.

OIJEN, A. M., BLAINEY, P. C., CRAMPTON, D. J., RICHARDSON, C. C., ELLENBERGER, T. & XIE, X. S., 2003. Single-Molecule Kinetics of λ-Exonuclease Reveal Base Dependence and Dynamic Disorder. *Science*, **301**, 1235–1238.

ONETO, J. F., 1938. Sulfophenylarsonic Acids and Certain of their derivatives. I. *p*-Sulfophenylarsonic Acid. *J. Am. Chem. Soc.*, **60**, 2058–2059.

OVERMAN, L. E. & O'CONNOR, E. M., 1976. Degenerate Intermolecular Thiolate-Disulfide Interchange Involving Cyclic Five-Membered Disulfides Is Faster by 10^3 Than That Involving Six- or Seven-Membered Disulfides. *J. Am. Chem. Soc.*, **98**, 771–775.

PANTOJA, R., RODRIGUEZ, E. A., DIBAS, M. I., DOUGHERTY, D. A. & LESTER, H. A., 2009. Single-Molecule Imaging of a Fluorescent Unnatural Amino Acid Incorporated Into Nicotinic Receptors. *Biophys. J.*, **96**, 226–237.

PAULING, L., 1945. *The Nature of the Chemical Bond*.

PINSENT, J., 1954. The Need for Selenite and Molybdate in the Formation of Formic Dehydrogenase by Members of the Coli-aerogenes Group of Bacteria. *Biochem. J.*, **57**, 10–16.

PLEASANTS, J. C., GUO, W. & RABENSTEIN, D. L., 1989. A Comparitive Study of the Kinetics of Selenol/Diselenide and Thiol/Disulfide Exchange Reactions. *J. Am. Chem. Soc.*, **111**, 6553–6558.

PORDEA, A., CREUS, M., PANEK, J., DUBOC, C., MATHIS, D., NOVIC, M. & WARD, T. R., 2008. Artificial metalloenzyme for enantioselective sulfoxidation based on vanadyl-loaded streptavidin. *J Am Chem Soc*, **130**, 8085–8988.

PORTERFIELD, W. W., 1984. Inorganic Chemistry, A Unified Approach.

PROD'HOM, B., PIETROBON, D. & HESS, P., 1987. Direct measurement of proton transfer rates to a group controlling the dihydropyridine-sensitive Ca^{2+} channel. *Nature*, **329**, 243–246.

RALPH, S. J., 2008. Arsenic-Based Antineoplastic Drugs and Their Mechanisms of Action. *Metal-Based Drugs*, **3**, 1–13.

RAUK, A., ALLEN, L. C. & MISLOW, K., 1969. Pyramidal Inversion. *Angew. Chem. Int. Ed.*, **9**, 400–414.

REN, X., JEMTH, P., BOARD, P. G., LUO, G., MANNERVIK, B., LIU, J., ZHANG, K. & SHEN, J., 2002. A Semisynthetic Glutathione Peroxidase with High Catalytic Efficiency: Selenoglutathione Transferase. *Chemistry & Biology*, **9**, 789–794.

RENGBY, O., JOHANSSON, L., CARLSON, L. A., SERINI, E., VLAMIS-GARDIKAS, A., P., K. & ARNER, E. S. J., 2004. Assessment of Production Conditions for Efficient Use of Escherichia coli in High-Yield Heterologous Recombinant Selenoprotein Synthesis. *Appl. Environ. Microbiol.*, **70**, 5159–5167.

RIEF, M., GAUTEL, M., OESTERHELT, F., FERNANDEZ, J. M. & GAUB, H. E., 1997. Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM. *Science*, **276**, 1109–1112.

RODRIGUEZ, E. A. LESTERĘ, H. A. & DOUGHERTY, D. A., 2006. *In vivo* incorporation of multiple unnatural amino acids through nonsense and frameshift suppression. *Proc. Nat. Acad. Sci.*, **103**, 8650–8655.

ROEFFAERS, M. B. J., SELS, B. F., UJI-I, H., DE SCHRYVER, F. C., JACOBS, P. A., DE VOS, D. E. & HOFKENS, J., 2006. Spatially resolved observation of crystal-facedependent catalysis by single turnover counting. *Nature Lett.*, **439**, 572–575.

ROSENTHAL, S. J., CHANG, J. C., KOVTUN, O., MCBRIDE, J. R. & TOMLINSON, I. D., 2011. Biocompatible Quantum Dots for Biological Applications. *Chemistry & Biology*, **18**, 10–24.

ROY, G., SARMA, B. K., PHADNIS, P. P. & MUGESH, G., 2005. Selenium-containing enzymes in mammals: Chemical perspectives. *J. Chem. Sci.*, **117**, 287–303.

ROY, R., HOHNG, S. & HA, T., 2008. A practical guide to single-molecule FRET. *Nature Methods*, **5**, 507–516.

SANCHEZ-QUESADA, J., GHADIRI, M. R., BAYLEY, H. & BRAHA, O., 2000. Cyclic Peptides as Molecular Adapters for a Pore-Forming Protein. *J. Am. Chem. Soc.*, **122**, 11757–11766.

SANDERSON, R. T., 1962. Chemical Periodicity.

SANDMAN, K. E. & NOREN, C. J., 2000. The efficiency of *Escherichia coli* selenocysteine insertion is influenced by the immediate downstream nucleotide. *Nucleic Acids Res.*, **28**, 755–761.

SANDMAN, K. E., TARDIFF, D. F., NEELY, L. A. & NOREN, C. J., 2003. Revised Escherichia Coli selenocysteine insertion requirements determined by in vivo screening of combinatorial libraries of SECIS variants. *Nucleic Acids Res.*, **31**, 2234–2241.

SCHEERER, P., BORCHERT, A., KRAUSS, N., WESSNER, H., GERTH, C., HOHNE, W. & KUHN, H., 2007. Structural Basis for Catalytic Activity and Enzyme Polymerization of Phospholipid Hydroperoxide Glutathione Peroxidase-4 (GPx4). *Biochemistry*, **47**, 9041–9049.

SCHMEING, T. M., HUANG, K. S., STROBEL, S. A. & STEITZ, T. A., 2005. Structure of the 70S Ribosome Complexed with mRNA and tRNA. *Science*, **438**, 520–524.

SCHMEING, T. M. & RAMAKRISHNAN, V., 2009. What recent ribosome structures have revealed about the mechanism of translation. *Nature*, **461**, 1234–1242.

SCHNECKENBURGER, H., 2005. Total internal reflection fluorescence microscopy: technical innovations and novel applications. *Curr. Opp. Biotech.*, **16**, 13–18.

SCHÖ, A., BÖCK, A., OTT, G., SPRINZL, M. & SÖLL, D., 1989. The selnocysteine-inserting opal suppressor serine tRNA from *E. Coli* is highly unusual in structure and modification. *Nucleic Acids Res.*, **17**, 7159–7156.

SCHWARZ, K. & FLOTZ, C. M., 1957. Selenium as an integral part of Factor 3 aganist Dietary Necrotic Liver Degeneration. *J. Am. Chem. Soc.*, **79**, 3292–3293.

SELMER, M., DUNHAM, C. M., MURPHY, F. V., WEIXLBAUMER, A., PETRY, S., KELLEY, A. C., WEIR, J. R. & RAMAKRISHNA, V., 2006. Structure of the 70S Ribosome Complexed with mRNA and tRN. *Science*, **313**, 1935–1942.

SEYEDSAYAMDOST, M. R., XIE, J., CHAN, C. T. Y., SCHULTZ, P. G. & STUBBE, J., 2007. Site-Specific Insertion of 3-Aminotyrosine into Subunit r2 of *E. coli* Ribonucleotide Reductase: Direct Evidence for Involvement of Y_730 and Y_731 in Radical Propagation. *J. Am. Chem. Soc.*, **129**, 15060–15071.

SHAIKH, T. A., PARKIN, S. & ATWOOD, D. A., 2006. Synthesis and characterization of a rare arsenic trithiolate with an organic disulfide linkage and 2-chloro-benzo-1,3,2-dithiastibole. *J. Organometallic Chem.*, **691**, 4167–4171.

SHERA, E. B., SEITZINGER, N. K., DAVIS, L. D., KELLER, R. A. & SOPER, S. A., 1990. Detection of single fluorescent molecules. *Chem. Phys. Lett.*, **174**, 553–557.

SHIMIZU, Y., KANAMORI, T. & UEDA, T., 2005. Protein synthesis by pure translation systems. *Methods*, **36**, 299–304.

SHIN, S.-H. & BAYLEY, H., 2005. Stepwise growth of a single polymer chain. *J. Am. Chem. Soc.*, **127**, 10462–10463.

SHIN, S.-H., LUCHIAN, T., CHELEY, S., BRAHA, O. & BAYLEY, H., 2002. Kinetics of a Reversible Covalent-BondForming Reaction Observed at the Single Molecule Level. *Angew. Chem. Int. Ed.*, **41**, 3707–3709.

SHIN, S.-H., STEFFENSEN, M. B., CLARIDGE, T. D. W. & BAYLEY, H., 2007. Stepwise growth of a single polymer chain. *Angew. Chem. Int. Ed.*, **46**, 7412–7416.

SINGH, R. & WHITESIDES, G., 1990. Degenerate Intermolecular Thiolate-Disulfide Interchange Involving Cyclic Five-Membered Disulfides Is Faster by 10^3 Than That Involving Six- or Seven-Membered Disulfides. *J. Am. Chem. Soc.*, **112**, 6304–6309.

SINGH, Y., SPINELLI, N., DEFRANCQ, E. & DUMY, P., 2006. A novel heterobifunctional linker for facile access to bioconjugates. *Org. Biomol. Chem.*, **4**, 1413–1419.

SINGMAN, C. N., 1984. Atomic volume and allotropy of the elements. *J. Chem. Educ.*, **61**, 137–142.

SLATER, J. C., 1964. Atomic radii in crystals. *J. Chem. Phys.*, **39**, 3199–3204.

SMOES, S. & DROWART, J., 1977. Determination of the atomization energies of the molecules CSe(g) and CSe2(g) by the mass spectrometric Knudsen cell method. *J. Chem. Soc. Faraday Trans.*, **73**, 1746–1754.

SOKOLOWSKI, M. S. & KONOPSKI, L., 2007. Reaction of Lewisite-1 with Alcohols, Diols, and Thiols in Water-A Simple Method of Derivatization of Thiodiglycol. *Phosphorus, Sulfur, and Silicon*, **182**, 2311–2327.

SÖLL, D., 1988. Enter a new amino acid. *Nature*, **331**, 662–663.

SONG, L., HOBAUGH, M. R., SHUSTAK, C., CHELEY, S., BAYLEY, H. & GOUAUX, 1996. Structure of Staphylococcal αup-Hemolysin, a Heptameric Transmembrane Pore. *Science*, **274**, 1859–1866.

SPRAGUE, B. L. & MCNALLY, J. G., 2005. FRAP analysis of binding: proper and fitting. *Trends in Cell Biol.*, **15**, 84–91.

STEINMANN, D., NAUSER, T. & KOPPENOL, W. H., 2010. Selenium and Sulfur in Exchange Reactions: A Comparative Study. *J. Org. Chem.*, **75**, 6696–6699.

STEVENSON, K. J., HALE, G. & PERHAM, R. N., 1978. Inhibition of Pyruvate Dehydrogenase Multienzyme Complex from Escherichia coli with Mono- and Bifunctional Arsenoxides. *Biochemistry*, **17**, 2189–2192.

STOCK, T. & ROTHER, M., 2009. Selenoproteins in Archea and Gram-positive bacteria. *Biochim. Biophys. Acta*, **1790**, 1520–1532.

STOLZ, J. F., BASU, P., SANTINI, J. M. & OREMLAND, R. S., 2006. Arsenic and Selenium in Microbial Metabolism. *Annu. Rev. Microbiol.*, **60**, 107–130.

SU, D., LI, Y. & GLADYSHEV, V. N., 2005. Selenocysteine insertion directed by the 3'-UTR SECIS element in *Escherichia coli*. *Nucleic Acids Res.*, **33**, 2486–2492.

SUBRAMANIAN, K., RUTVISUTTINUNT, W., SCOTT, W. & MYERS, R. S., 2003. The enzymatic basis of processivity in λ exonuclease. *Nucleic Acids Res.*, **31**, 1585–1596.

SUN, B. C., HARNEY, J. W., BERRY, M. J. & LARSEN, P. R., 1997. The Role of the Active Site Cysteine in Catalysis by Type 1 Iodothyronine Deiodinase. *Endocrinology*, **138**, 5452–5458.

SUTTON, L. E., 1965. *Table of Interatomic Distances and Configuration in Molecules and Ions*.

SVOBODA, K., SCHMIDT, C. F., SCHNAPP, B. J. & BLOCK, S. M., 1993. Direct observation of kinesin stepping by optical trapping interferometry. *Nature*, **365**, 721–727.

TANG, Y., ZHOU, L. P., LI, J. X., LUO, Q., HUANG, X., WU, P., WANG, Y. G., XU, J. Y. SHEN, J. C. & LIU, J. Q., 2010. Giant nanotubes loading artificial peroxidase centers: self-assembly of supramolecular amphiphiles as a tool to functionalize nanotubes. *Angew Chem Int Ed*, **49**, 3920–3924.

TORMAY, P., SAWERST, A. & BÖCK, A., 1996. Role of stoichiometry between mRNA, translation factor SelB and selenocysteyl-tRNA in selenoprotein synthesis. *Mol. Microbiol.*, **21**, 1253–1259.

TOSCANO, M. D., WOYCECHOWSKY, K. J. & HILVERT, D., 2007. Minimalist active-site redesign: teaching old enzymes new tricks. *Angew. Chem. Int. Ed.*, **46**, 3212–3236.

TUJEBAJEVA, R. M., COPELAND, P. R., XU, X. M., CARLSON, B. A., HARNEY, J. W., DRISCOLL, D. M., HATFIELD, D. L. & BERRY, M. J., 2000. Decoding apparatus for eukaryotic selenocysteine insertion. *EMBO rep.*, **1**, 158–163.

TURCATTI, G., NEMETH, K., EDGERTON, M. D., MESETH, U., TALABOT, F., PEITSCH, M., KNOWLES, J., VOGEL, H. & CHOLLET, A., 1996. Probing the Structure and Function of the Tachykinin Neurokinin-2 Receptor through Biosynthetic Incorporation of Fluorescent Amino Acids at Specific Sites. *J. Biol. Chem.*, **271**, 1991–19998.

URSINI, F., MAIORINO, M., BRIGELIUS-FLOHE, R. & AUMANN, K. D., 1995. Diversity of glutathione peroxidase. *Meth. Enzymol.*, **252**, 38–53.

VASILEV, A. A. & ENGMAN, L., 1998. Iodothyronine Deiodinase Mimics. Deiodination of o,o'-Diiodophenols by Selenium and Tellurium Reagents. *J. Org. Chem.*, **63**, 3911–3917.

WALKER, B. & BAYLEY, H., 1995. Restoration of pore-forming activity in staphylococcal a-hemolysin by targeted covalent modification. *Protein Engineering*, **8**, 491–495.

WALKER, B., KASIANOWICZ, J., KRISHNASASTRY, M. & BAYLEY, H., 1994. A pore-forming protein with a metal-actuated switch. *Protein Engineering*, **7**, 655–662.

WALTHER, K. A., BRUJIĆ, J., LI, H. & FERNÁNDEZ, J. M., 2006. Sub-Angstrom Conformational Changes of a Single Molecule Captured by AFM Variance Analysis. *Biophysical J.*, **90**, 3806–3812.

WANG, L. & XIE, J. SCHULTZ, P. G., 2006. Expanding the Genetic Code. *Annu. Rev. Biophys. Biomol. Struct.*, **35**, 225–249.

WANG, Q. G., YANG, Z. M., GAO, Y., GE, W. W., WANG, L. & XU, B., 2008. Enzymatic hydrogelation to immobilize an enzyme for high activity and stability. *Soft matter*, **4**, 550–553.

WATERS, W. & WILLIAMS, J. H., 1950. Hydrolyses and Derivatives of Some Vesicant Arsenicals. *J. Chem. Soc.*, 18–22.

WESSJOHANN, L. A., SCHNEIDER, A., ABBAS, M. & BRANDT, W., 2007. Selenium in chemistry and biochemistry in comparison to sulfur. *Biol. Chem.*, **388**, 997–1006.

WHEELER, J., 1997. Electron affinities of the alkaline earth metals and the sign convention for electron affinity. *J. Chem. Educ.*, **74**, 123–127.

WIMBERLY, B. T., BRODERSEN, D. E., CLEMONS JR, W. M., MORGAN-WARREN, R. J., CARTER, A. P., VONRHEIN, C., HARTSCHK, T. & RAMAKRISHNAN, V., 2000. Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.

WU, Z. P. & HILVERT, D., 1989. Conversion of a Protease into an Acyl Transferase: Selenolsubtilisin. *J. Am. Chem. Soc.*, **111**, 4513–4514.

WU, Z.-P. & HILVERT, D., 1990. Selenosubtilisin as a Glutathione Peroxidase Mimic. *J. Am. Chem. Soc.*, **112**, 5647–5648.

XIAO, H., MURAKAMI, H., SUGA, H. & FERRE-D'AMARE, A., 2008. Structural basis of specific tRNA aminoacylation by a small *in vitro* selected ribozyme. *Nature*, **454**, 358–362.

XU, W., KONG, J. S., YEH, Y. T. E. & CHEN, P., 2008. Single-molecule nanocatalysis reveals heterogeneous reaction pathways and catalytic dynamics. *Nature Materials*, **7**, 992–996.

YUAN, J., PALIOURA, S., SALAZAR, J. C., SU, D., O'DONOGHUE, P., HOHN, M. J., CARDOSO, A. M. WHITMAN, W. B. & S*ö*, D., 2006. RNA-dependent conversion of phosphoserine forms selenocysteine in eukaryotes and archaea. *Pro. Natl. Acad. Sci.*, **103**, 18923–18927.

ZHANG, Y. & GLADYSHEV, V. N., 2010. General Trends in Trace Element Utilization Revealed by Comparative Genomic Analyses of Co, Cu, Mo, Ni, and Se. *J. Biol. Chem.*, **285**, 3393–3405.

ZHANG, Y., ROMERO, H., SALINAS, G. & GLADYSHEV, V. N., 2006. Dynamic evolution of selenocysteine utilization in bacteria: a balance between selenoprotein loss and evolution of selenocysteine from redox active cysteine residues. *Genome Biol.*, **7**, 1–17.

ZHONG, L., ARNÉR, E. S. J. & HOLMGREN, A., 2000. Structure and mechanism of mammalian thioredoxin reductase: The active site is a redox-activeselenolthiolyselenenylsulfide formed from the conserved cysteine-selenocysteine sequence. *Pro. Nat. Acad. Sci.*, **97**, 5854–5859.

ZINONI, F., BIRKMANN, A., STADTMAN, T. C. & BÖCK, A., 1986. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*. *Proc. Natl. Acad. Sci.*, **83**, 4650–4654.

ZINONI, F., HEIDER, J. & BÖCK, A., 1990. Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine. *Proc. Natl. Acad. Sci.*, **87**, 4660–4664.

ZUO, J. M., VARTANYANTS, I., GAO, M., ZHANG, R. & NAGAHARA, L. A., 2003. Atomic Resolution Imaging of a Carbon Nanotube from Diffraction Intensities. *Science*, **300**, 1419–1421.