The University of Manchester

Biochemical and Biophysical Characterisation of Silicatein

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering

2017

Stephanie Caslin

School of Chemistry

List of Contents

List of Contents 2
List of Figures 8
List of Tables 14
Abstract 16
Declaration 17
Copyright Statement 18
Acknowledgements 19
List of abbreviations 20
1. Chapter 1: General Introduction 23
1.1 Silicon in nature 23
1.2 Properties of silicon 23
1.3 Hypervalent silicon species 24
1.4 Discovery of organosilicon compounds 25
1.5 Use of organosilicon compounds in synthetic chemistry 27
1.5.1 Cross coupling reactions with organosilicon compounds using chemocatalysts 27
1.5.2 Silyl protecting groups 28
1.5.3 Protection of silyl ethers 29
1.5.4 Deprotection of silyl ethers 31
1.6 Marine biosilification 32
1.7 Discovery of silicatein 33
1.8 Proposed mechanism of silicatein 34
1.9 Hydrolases 37
1.9.1 Serine hydrolases 37
1.9.2 Mechanism of serine hydrolase chymotrypsin 38
1.10 Biocatalysts in synthetic chemistry 41
1.11 Silicatein and metal oxides 43
1.12 Condensation of siloxanes using enzymes 44
1.13 Hydrolysis of siloxanes 46

1.14 Enzymes in organic solvents 46

1.15 Biochemical methods for analysing silicatein-α 48

1.16 Catalytic efficiency of biocatalysts using Michaelis-Menten kinetic parameters 48

1.17 Active site investigations through mutagenesis 50

1.18 Aims and Objectives 51

2. Chapter 2: Optimisation of recombinant protein expression and production of silicatein-α in *Escherichia coli* 53

2.1 Introduction 53

2.1.1 Recombinant protein production 53

2.1.2 Fusion protein technology 54

2.1.3 Co-expression with molecular chaperones 56

2.1.4 Optimum buffer conditions 57

2.1.5 Optimum purification methods 57

2.1.6 Aims and objectives 58

2.2 Results and discussion 59

2.2.1 Bioinformatic search using orthologs of silicatein-α from *S. domuncula* 59

2.2.2 Expression of fusion protein GST-silicatein-α 60

2.2.3 Expression of fusion protein SUMO-silicatein-α 62

2.2.4 Expression of fusion protein TF-silicatein-α 62

2.2.5 Purification of fusion protein TF-silicatein-α using immobilised metal affinity column (IMAC) 64

2.2.6 Purification of fusion protein TF-silicatein-α using Ammonium sulphate precipitation 65

2.2.7 Purification of fusion protein TF-silicatein-α using ammonium sulphate precipitation and immobilised metal affinity column (IMAC) 66

2.2.8 Cleavage of fusion protein TF-silicatein-α with Factor Xa at 25°C and 4°C 67

2.2.9 Cleavage of fusion protein TF-silicatein-α with HRV3C and GST-HRV3C at 4°C 68

2.2.10 Solubility analysis of co-expression of GST-silicatein-α and SUMO-silicatein-α with chaperone vector 69

2.2.11 Refolding of silicatein-α 69

2.2.12 Buffer optimisation for the purification of TF-silicatein-α and silicatein-α 71

2.2.13 Expression, purification and protease cleavage of TF-silicatein-α-Strep 73

2.2.14 Optimisation of protein purification with IMAC and SEC with TF-silicatein-α 76

2.3 Conclusion 77

3. Chapter 3: Biochemical analysis of silicatein-α and variants 79

3.1 Introduction 79

3.2 Biochemical analysis 79

3.2.1 Silicomolybdic acid assay 79

3.2.2 Hydrolysis of TBDMS-4np 80

3.2.3 Exploring active site residues through point mutagenesis 81

3.2.4 Aims and objectives 83

3.3 Results and Discussion 84

3.3.1 Biochemical analysis of silicatein-α using SMAA hydrolysis 84

3.3.2 Biochemical analysis of silicatein-α using SMAA condensation 84

3.3.3 Biochemical analysis of silicatein-α using hydrolysis of TBDMS-4np 85

3.3.4 Michaelis Menten Kinetic Parameters using TBDMS -4- nitrophenyl silyl ether 87

3.3.5 Esterase and Protease activity of silicatein-α and variants 90

3.4 Conclusions 93

4. Chapter 4: Biophysical analysis of silicatein-α and variants 95

4.1 Introduction 95

4.2 Biophysical measurements for protein characterisation 95

4.3 Circular Dichroism (CD) for secondary structural information of proteins 96

4.4 Theory of Light Scattering 98

4.4.1 Static light scattering 98

4.4.2 Dynamic light scattering 100

4.4.3 Aims and objectives 102

4.5 Results and discussion 102

4.5.1 Circular Dichroism (CD) measurements and secondary structure prediction 102

4.5.2 Dynamic Light Scattering 105

4.5.3 Size Exclusion Chromatography with Multi-Angled Light Scattering 115

4.6 Conclusions 119

5. Chapter 5: Investigation of substrate scope, thermostability and enantioselectivity of silicatein-α with a coupled biotransformation using ADH 120

5.1.1 Traditional methods used for silyl protection 120

5.1.2 Traditional methods used for silyl deprotection 121

5.1.3 The application of enzymes as biocatalysts in organic media with silyl protecting groups 121

5.1.4 Oxidoreductases 122

5.1.5 Alcohol dehydrogenase 123

5.1.6 Enzymatic reduction of ketones and oxidation of alcohols to obtain chiral compounds 123

5.1.7 The application of Alcohol Dehydrogenase in organic solvent 125

5.2 Aims and objectives 126

5.3 Results and Discussion 128

5.3.1 Thermostability of silicatein-α as a biocatalyst in organic solvent 128

5.3.2 Enantioselectivity of silicatein-α in organic solvent 132

5.3.3 Coupled organic synthesis in organic media with silicatein-α and ADH 144

5.4 Conclusions 145

6. Chapter 6: Conclusions 147

7. Chapter 7: Materials and Methods 151

7.1 General materials and equipment 151

7.2 General methods 152

7.2.1 Polymerase Chain Reaction (PCR) 152

7.2.2 DpnI Treatment 154

7.2.3 DNA electrophoresis 155

7.2.4 DNA extraction 155

7.2.5 SDS-PAGE gels 155

7.2.6 Gibson assembly 156

7.2.7 Chemically competent bacterial cells transformation 156

7.3 Experimental Methodology for Chapter 2 156
7.3.1 Bioinformatic search using orthologs of silicatein-α from *S. domuncula* 156
7.3.2 Plasmid construction 157
7.3.3 Cell cultures 159
7.3.4 Cell lysis, SDS PAGE and Western blot analysis 162
7.3.5 Protein purification TF-silicatein-α 162
7.3.6 Refolding silicatein-α 163
7.3.7 Protein Purification of silicatein-α 164
7.3.8 Protein Purification of TF-Silicatein-α-Strep 164
7.3.9 Proteolytic Cleavage of TF-silicatein-α and TF-silicatein-α-Strep 165
7.3.10 Size Exclusion Chromatography 165
7.4 Experimental Methodology for Chapter 3 165
7.4.1 Plasmid construction of protein variants 165
7.4.2 Biochemical analysis of silicatein-α using SMAA hydrolysis 167
7.4.3 Biochemical analysis of silicatein-α using SMAA condensation 167
7.4.4 Biochemical analysis of silicatein-α using hydrolysis of TBDMS-4np 167
7.5 Experimental Methodology for Chapter 4 168
7.5.1 Circular Dichroism (CD) measurements and secondary structure prediction 168
7.5.2 Dynamic Light Scattering 168
7.5.3 Static Light Scattering 169
7.5.4 Multi-Angled Light Scattering 170
7.6 Experimental Methodology for Chapter 5 170
7.6.1 Condensation reaction in organic solvent, silyl etherification of and transetherification with TMS-OH and TES- OH 170
7.6.2 Oxidation reaction with ADH in aqueous buffer 171
7.6.3 Reduction reaction with ADH in aqueous buffer 171
7.6.4 Oxidation reaction with ADH in organic solvent 171
7.6.5 Reduction reaction with ADH in organic solvent 172

7.6.6 GC-MS parameters 172

8. Chapter 8: References 174

List of Figures

Figure 1.1 Representation of the Pimentle-Rundle model of three centre four electron (3c-4e) theory 25

Figure 1.2 Examples of Organosilicon compounds (A) tetraethylsilane and (B) tetramethylsilane 25

Figure 1.3 The production of chlorosilanes using the Rochow process 26

Figure 1.4 Catalytic cycle of Hiyama coupling mechanism 28

Figure 1.5 an example of some silyl protecting group used in organic synthesis 29

Figure 1.6 Enantioselective silylation of asymmetric and cyclic diols 30

Figure 1.7 Condensation reaction of silicic acid mediated by silicatein 32

Figure 1.8 Truncated sequence alignment of human cathepsin L (UniProtKB P07711) and silicatein-α from *S. domuncula* (UniProtKB Q2MEV3) and predicted model of silicatein- α showing the catalytic residues 34

Figure 1.9 Reaction scheme with proposed mechanism of silicatein-α for hydrolysis and spontaneous condensation of alkoxysilane TEOS 36

Figure 1.10 Proposed mechanism of silicatein-α for silicic acid condensation determined from molecular docking studies using the crystal structure of cathsilicatein (cathepsinL/silicatein chimera) 37

Figure 1.11 Representation of catalytic mechanism of serine hydrolase, chymotrypsin 40

Figure 1.12 Enzyme-catalysed reaction with p-nitrophenylacetate to p-nitrophenolate 41

Figure 1.13 Enantioselective esterification of R-ibuprofen with lipase and 1-propranol 42

Figure 1.14 Schematic of hydrolysis and condensation of siloxanes mediated by enzymes 45

Figure 2.1 Sequence alignment of silicatein-α from *Suberites domuncula* (Q2MEV3), *Spongilla lacustris* (D3DEQ1), *Baikalospongia fungiformis* (G3EAH4), *Lubomirskia baicalensis* (Q2PC18), and *Ephydatia fluviatilis* (B5U9F0) using Clustal Omega 60

Figure 2.1. Illustrative analysis of fusion protein GST-silicatein-α 61

Figure 2.3 Illustrative analysis of fusion protein SUMO-silicatein-α 62

Figure 2.4 Solubility analysis of fusion protein TF-silicatein-α 63

Figure 2.5 Purification of fusion protein TF-silicatein-α using immobilised metal affinity column (IMAC) 64
Figure 2.6 Purification of fusion protein TF-silicatein-α using Ammonium sulphate precipitation 0-100% saturation 65
Figure 2.7 Purification of fusion protein TF-silicatein-α using Ammonium sulphate precipitation at 60, 65 and 70% saturation 66
Figure 2.8 Purification of fusion protein TF-silicatein-α using ammonium sulphate precipitation and immobilised metal affinity column (IMAC) 67
Figure 2.9 Cleavage of fusion protein TF-silicatein-α with Factor Xa at 25°C and 4°C 68
Figure 2.10 Refolding of silicatein-α 70
Figure 2.11 Buffer optimisation for silicatein-α 71
Figure 2.12 SEC chromatograms and SDS-PAGE analysis of TF-silicatein-α 72
Figure 2.13 SEC analysis of silicatein-α and TF-silicatein-α 73
Figure 2.14 Illustrative analysis of fusion protein TF-silicatein-α-Strep 74
Figure 2.15 purification of TF-silicatein-α-Strep 75
Figure 2.16 Protease cleavage of TF-silicatein-α-Strep 75
Figure 2.17 SDS-PAGE analysis of Strep-Tactin affinity column following thrombin cleavage of TF-silicatein-α-Strep 76
Figure 2.18 Optimisation of the overall purification process of fusion protein TF-silicatein-α using a continuous buffer system 77
Figure 3.1 Scheme showing the reaction for the formation of molybdenum blue 80
Figure 3.2 Precipitated silica formed by silicatein-α and variants of the enzyme after 1 hour incubation with TEOS to investigate hydrolysis and specific activity 84
Figure 3.3 Precipitated silica formed by silicatein-α and variants of the enzyme after 1 hour incubation with pre-hydrolysed TEOS to investigate condensation and specific activity 85
Figure 3.4 The hydrolysis of TBDMS-4-nitrophenol against time by TF-silicatein-α and enzyme variants 87
Figure 3.5 The concentration of 4-nitrophenol formed against time from hydrolysis catalysed by TF-silicatein-α (A), His165Ala (B), Asn185Ala (C), and Ser26C (D) 89
Figure 3.6 The concentration of 4-nitrophenol formed against time from hydrolysis catalysed by Ser26Ala (A) and Asn185Asp (B), MM plots for TF-silicatein-α and variants 90

Figure 3.7 Product formation against time to investigate esterase and protease activity of TF-Silicatein-α and variants. (A) p-nitrophenyl pivalate, (B) Cbz-Phe-Arg –NHNp, (C) TBDMS-4-nitrophenyl 91
Figure 4.1 CD spectra of proteins and polypeptides to show characteristic secondary structure 97
Figure 4.2 Schematic representation of Multi-angled light scattering 99
Figure 4.3 Homology model of Silicatein-α 101
Figure 4.4 CD spectra for TF-silicatein, Ser26Ala, Ser26Cys, His165Ala, Asn185Ala and Asn185Asp plotting molar ellipticity (MRE) against wavelength 102
Figure 4.5 CD Spectra for TF-silicatein, Ser26Ala, Ser26Cys, His165Ala, Asn185Ala, Asn185Asp, TF and silicatein-α plotting mean residue molar ellipticity (MRME) against wavelength 103
Figure 4.6 Dynamic Light Scattering experiment with TF-silicatein-α in 50 mM Tris, 100 mM NaCl pH 8.5 at 4 °C, 22 °C, 30 °C and 37 °C (A) estimated hydrodynamic radius (B) estimated molecular weight 108
Figure 4.7 Light scattering experiments with varying concentrations of NaCl 109
Figure 4.8 Light scattering experiments with varying pH levels of buffer 111
Figure 4.9 Light scattering experiments comparing samples with and without Arginine and Glutamate 113
Figure 4.10 Light scattering experiments with various buffer conditions 114
Figure 4.11 Size exclusion chromatography (SEC) chromatogram (A) TF-silicatein-α 1 mg/mL eluting in the void volume (B) TF-silicatein-α 0.3 mg/mL eluting in the void volume 116
Figure 4.12 Multi angled light scattering data (MALS) with TF-silicatein-α in 50 mM Tris 100 mM NaCl 50 mM Arg + Glu pH 8.5. Estimated molecular weight and MALS calculated molecular weight are shown along with polydispersity index 117
Figure 4.13 Multi angled light scattering data (MALS) with silicatein-α in 100 mM EPPS (blue) or 50 mM Bis-Tris (red) ,100 mM NaCl 50 mM Arg + Glu pH 8.5. Estimated molecular weight and MALS calculated molecular weight are shown along with polydispersity index 117
Figure 5.1 An example of silylating agents used for silyl protection 121
Figure 5.2 Batchwise enzymatic alcohol oxidation with two enantioselective ADH with enzyme cascade 124

Figure 5.3 Biotransformation of (*S*)-3,5-bistrifluoromethylphenyl ethanol with ADH pH 7 phosphate buffer at 30 °C using FDH as a regeneration system 125
Figure 5.4 Enzymatic ketone reduction of (2,5)-hexanedione to (2*R*,5*R*)-hexanediol by *L. kefir* ADH 125
Figure 5.5 Proposed coupled enzyme system for silylation using TF-silicatein-α and ADH 127
Figure 5.6 Etherification catalysed by TF-silα at 75 °C in octane with trimethylsilanol (1b) and 1-octanol (2b) to form trimethyl(octyloxy)silane (3b) 128
Figure 5.7 The percentage of product formation of trimethyl(octyloxy)silane (3b) from trimethylsilanol (1b) and 1-octanol (2b) in *n*-octane at 75 °C analysed by GC-MS 129
Figure 5.8 Scheme 4.2 Trans-Etherification catalysed by TF-silα at 75 °C in octane with trimethylsilane (4b) and 1-octanol (2b) to form trimethyl(octyloxy)silane (3b) 129
Figure 5.9 The percentage of product formation trimethyl(octyloxy)silane (3b) from trimethylsilane (4b) and octanol (2b) in *n*-octane at 75 °C analysed by GC-MS 129
Figure 5.10 Etherification catalysed by TF-silα at 75 °C and 90°C in n-octane with triethylsilanol (5b) and 1-octanol (2b) to form triethyl(octyloxy)silane (6b) 130
Figure 5.11 The percentage of product formation from triethylsilanol (5b) and octanol (2b) in *n*-octane at 75 °C and 90 °C to produce triethyl(octyloxy)silane (6b) 130
Figure 5.12 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in n-octane with triethylsilanol (7b) and 1-octanol (2b) to form triethyl(octyloxy)silane (6b) 131
Figure 5.13 The percentage of product formation, triethyl(octyloxy)silane (6b), from triethylsilane (7b) and octanol (2b) in *n*-octane at 75 °C and 90 °C 131
Figure 5.14 Etherification catalysed by TF-silα at 75 °C and 90°C in n-octane with triethylsilanol (5b) and (*S*)-2-octanol (8b) to form (S)-triethyl(octan-2-yloxy)silane (9b) 132
Figure 5.15 The percentage of product formation, (*S*)-triethyl(octan-2-yloxy)silane (9b), from triethylsilanol (5b) and (*S*)-2-octanol (8b) in *n*-octane at 75 °C and 90 °C 133
Figure 5.16 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in n-octane with triethylsilane (7b) and (S)-2-octanol (8b) to form (S)-triethyl(octan-2-yloxy)silane (9b) 133
Figure 5.17 The percentage of product formation, (*S*)-triethyl(octan-2-yloxy)silane (9b), from triethylsilane (7b) and (*S*)-2-octanol (8b) in *n*-octane at 75 °C and 90 °C 134
Figure 5.18 Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with (5b) triethylsilanol and (*R*)-2-octanol (10b) to form (*R*)-triethyl(octan-2-yloxy)silane (11b) 134

Figure 5.19 The percentage of product formation, (*R*)-triethyl(octan-2-yloxy)silane (11b), from triethylsilanol (5b) and (*R*)-2-octanol) (10b) in *n*-octane at 75 °C and 90 °C 135
Figure 5.20 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with (7b) triethylsilane and (*R*)-2-octanol (10b) to form (*R*)-triethyl(octan-2-yloxy)silane (11b) 135
Figure 5.21 The percentage of product formation, (*R*)-triethyl(octan-2-yloxy)silane (11b), from triethylsilane (7b) and (*R*)-2-octanol (10b) in *n*-octane at 75 °C and 90 °C 136
Figure 5.22 Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with (5b) triethylsilanol and (*S*)-1-phenylethanol (12b) to form (*S*)-triethyl(1-phenylethoxy)silane (13b) 137
Figure 5.23 The percentage of product formation, (*S*)-triethyl(1-phenylethoxy)silane (13b), from triethylsilanol (5b) and (*S*)-1-phenylethanoll (12b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS 137
Figure 5.24 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with triethylsilane (7b) and (*S*)-1-phenylethanoll (12b) to form (*S*)-triethyl(1-phenylethoxy)silane (13b) 137
Figure 5.25 The percentage of product formation, (*S*)-triethyl(1-phenylethoxy)silane (13b), from triethylsilane (7b) and (*S*)-1-phenylethanol (12b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS 138
Figure 5.26 Etherification catalysed by TF-silα at 75 °C and 90°C in n-octane with triethylsilanol (5b) and (R)-1-phenylethanoll (14b) to form (R)-triethyl(1-phenylethoxy)silane (15b) 138
Figure 5.27 The percentage of product formation, (R)-triethyl(1-phenylethoxy)silane (15b), from triethylsilanol (5b) and (R)-1-phenylethanol (14b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS 139
Figure 5.28 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with triethylsilane (7b) and (R)-1-phenylethanoll (14b) to form (R)-triethyl(1-phenylethoxy)silane (15b) 139
Figure 5.29 The percentage of product formation, (R)-triethyl(1-phenylethoxy)silane (15b), from triethylsilane (7b) and (R)-1-phenylethanoll (14b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS 139

Figure 5.30 Absorbance at 340 nm measured with UV-vis spectrophotometer from ADH1, ADH2 and ADH3 with cofactor NAD+ with or without 1,4-dioxane and (A) ethanol (B) *S*-1-phenylethanol (C) *R*-1-phenylethanol 142
Figure 5.31 Absorbance at 340 nm measured with UV-vis spectrophotometer from ADH1, ADH2 and ADH3 with cofactor NAD+ with or without 1,4-dioxane and (A) *R*-2-octanol (B) *S*-2-octanol 143

List of Tables

Table 2.1 Characterisation of chaperone vectors co-expressed with GST-silicatein-α and SUMO-silicatein-α 69

Table 3.1. Table of estimated kinetic parameters (Michaelis Menten constants) for the wild type enzyme and variants for the hydrolysis of TBDMS-4-nitrophenyl 88

Table 3.2. Table showing concentration of product formation for each enzyme against each substrate used to investigate esterase and protease activity 92

Table 4.1. Estimated proportions of secondary structure in TF-silicatein calculated by CDNN algorithm using CD spectrum in Figure 12.2 104

Table 4.2. Estimated proportions of secondary structure in TF-silicatein (Ser26Ala) calculated by CDNN algorithm using CD spectrum in Figure 12.2 104

Table 4.3. Estimated proportions of secondary structure in TF-silicatein (His165Ala) calculated by CDNN algorithm using CD spectrum in Figure 12.2 104

Table 4.4. Estimated proportions of secondary structure in TF-silicatein (Asn185Ala) calculated by CDNN algorithm using CD spectrum in Figure 12.2 105

Table 4.5. Estimated proportions of secondary structure in TF-silicatein (Asn185Asp) calculated by CDNN algorithm using CD spectrum in Figure 12.2 105

Table 4.6. Estimated proportions of secondary structure in TF-silicatein (Asn185Asp) calculated by CDNN algorithm using CD spectrum in Figure 12.2 105

Table 4.7. Comparative analysis of secondary structure in TF-silicatein and variants 105

Table 4.8. Percentage polydispersity estimated by Dynamic Light Scattering with TF-silicatein at a concentration of 2 mg/mL. Samples at pH 3.5 were in 50 mM Citrate buffer with the remaining samples in 50 mM Tris 100 mM NaCl 107

Table 4.9. Table of polydispersity index values of silicatein-α and TF-silicatein-α with varying concentrations of NaCl at 22 °C 109

Table 4.10. Table of polydispersity index values of silicatein-α and TF-silicatein-α with varying pH levels at 22 °C 111

Table 4.11. Table of polydispersity index values of silicatein-α and TF-silicatein-α in buffer pH 8.5 and pH 8.5 (Arg + Glu) at 22 °C 112

Table 4.12. Table of polydispersity index values of silicatein-α and TF-silicatein-α in buffer pH 8.5 and pH 8.5 (Arg + Glu) at 22 °C 115
Table 5.1 Percentage product conversion at 72 hours from etherification and trans-etherification reactions at 75 °C and 90 °C with 1-octanol 132
Table 5.2 Percentage product conversion at 72 hours from etherification and trans-etherification reactions at 75 °C and 90 °C with *R*- and *S*-2 octanol 136
Table 5.3 Percentage product conversion at 72 hours from etherification and trans-etherification reactions at 75 °C and 90 °C with R- and S-1-phenylethanol 140
Table 7.1 PCR reagents for construction of fusion protein plasmids 153
Table 7.2 PCR program conditions for construction of fusion protein plasmids 153
Table 7.3 PCR reagents for protein variants (megaprimer) 153
Table 7.4 PCR conditions for protein variants (megaprimer) 154
Table 7.5 Reagents of DpnI Treatment 154
Table 7.6 SDS-PAGE gel components for a single gel at 10% or 15% 155
Table 7.7. Buffer conditions screened for optimisation experiments using Dynamic Light Scattering 169
Table 7.8. Buffer conditions screened for optimisation experiments using Static Light Scattering thermal ramp and Dynamic Light Scattering at 22 °C 170
Table 7.9 GC-MS parameters used to analyse silylation of various compounds with TES 172
Table 7.10 GC-MS parameters used to analyse TMS silylation of 1-octanol 173

Abstract

The use of biocatalysts in organosilicon chemistry is not widely studied. The application of a catalyst that operates under ambient conditions would offer a more sustainable method in the production of organosilicon compounds and silicon based materials, in particular silicone, which is almost universally used in many consumer products and industrial processes. Silicatein, an enzyme found exclusively in the marine sponge catalyses the polycondensation of silica and would be a good candidate for the biocatalysis of silicon containing organic compounds. Recombinant silicatein-α has been previously shown to accept simple compounds as substrates such as silicon ethoxide and TEOS[1] along with more complex silanes and organosiloxanes.[2]

A steady and reliable supply of active, soluble protein was achieved by screening a range of protein production parameters, such as a variety of *E. coli* strains and fusion partners and a range of incubation temperatures. Further investigations led to the optimisation of purification and isolation methods to obtain fusion protein and native silicatein-α for structural and functional analysis. Therefore, the existing methods were refined to produce both recombinant enzymes more efficiently, reducing the overall time of protein purification.

A range of enzyme variants was used to characterise and explore the functionality of each catalytic residue on the overall catalytic ability. Kinetic data showed although similar K_M values were recorded the variants, the wild type enzyme has the highest catalytic efficiency in the region of 46 min^{-1} μM^{-1}. Therefore, the presence of all three catalytic residues are necessary for full catalytic activity towards this substrate. TF-silicatein-α and Ser26Cys revealed no esterase and protease activity, however, Asn185Asp displayed a small level of esterase activity (1.24 % conversion).

CD and DLS analysis confirmed structural integrity under the standard assay conditions with secondary structural predicted to be similar to theoretically calculated data from crystallographic data in the PDB. DLS and MALS data also highlighted two potential buffer conditions that are suitable for maintaining protein homogeneity in solution. MALS data shows silicatein-α to be monomeric with a Mw of ~ 30 kDa.

Silyl etherification and trans-etherification was catalysed by TF-silicatein-α at a higher rate of conversion at 90 °C and showed preference towards *S*-enantiomers. In addition, preliminary data indicated the possibility of a coupled enzyme synthesis with an *S*-specific ADH.

In summary this work confirms TF-silicatein-α as a robust, highly thermostable and enantioselective biocatalyst with potential applications in many industries that use organosilicon compounds.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

First of all, I would like to express my appreciation to my supervisors, Dr Lu Shin Wong and Dr Peter Quayle, for their expertise, patience and support throughout my PhD. My confidence and knowledge has greatly developed with their continued guidance and invaluable advice throughout this project.

My sincere thanks to Dr Wong's group past and present for all their support, particularly Seyed Yasin Tabatabaei Dakhili, Emily Sparkes, Teh Ser Huy and Joseph Hosford for their assistance and training on various aspects of this work.

I would like express my warmest gratitude to my loving husband (Neil Caslin) and my two beautiful daughters (Ellie Caslin and Megan Caslin) for all their continuing love, support and inspiration. My heartfelt appreciation for all the sacrifices they have made, especially over the last ten years, on my quest to provide a better life for our family.

Lastly, I would like to thank my parents (Neil Powney and Barbara Powney) and grandparents (Ann Powney and Henry Powney) for their love, support and motivation to achieve my goals.

List of abbreviations

ADH	Alcohol dehydrogenase
Ala	Alanine
APS	Ammonium persulphate
Arg	Arginine
Asn	Aspargine
Asp	Aspartic acid
ATP	Adenosine triphosphate
BICINE	2-(Bis(2-hydroxyethyl)amino)acetic acid
Bis-Tris	2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol
C	Carbon
CAST	Combinatorial active-site saturation testing
CD	Circular dichroism
$CDCl_3$	Deuterated chloroform
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
Cys	Cysteine
DLS	Dynamic light scattering
DMPS	Dimethylphenylsilyl
dNTP	Deoxynucleoside triphosphate
E	Enzyme
EDTA	Ethylenediaminetetraacetic acid
EP	Enzyme-product complex
EPC	Enantiopure compounds
EPPS	4-(2-Hydroxyethyl)-1-piperazinepropanesulfonic acid
ES	Enzyme-substrate complex
FPLC	Fast performance liquid chromatography
GC-MS	Gas chromatography mass spectrometry

GFC	Gel filtration chromatography
Glu	Glutamate
Gln	Glutamine
Gly	Glycine
GST	Glutathione S-transferase
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
His	Histidine
IMAC	Immobilised metal affinity column
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
ISM	Iterative saturation mutagenesis
LALS	Low angled light scattering
LB	Luria Broth
Lys	Lysine
MALS	Multi-angled light scattering
MBP	Maltose binding protein
MM	Michaelis-Menten
MOPS	3-Morpholinopropane-1-sulfonic acid
MWCO	Molecular weight cut off
MRE	Mean residue ellipticity
NaCl	Sodium chloride
NAD+/NADH	Nicotinamide adenine dinucleotide
NaOH	Sodium hydroxide
Ni^{2+}-NTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PCS	Photon correlation spectroscopy
PDB	Protein data bank

PDMS	Polydimethylsiloxane
Phe	Phenylalanine
Pro	Proline
PVDF	Polyvinylidene fluoride
QELS	Quasielastic light scattering
R_g	Radius of gyration
R_h	Radius of hydration
RMS	Root mean squared
S	Substrate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
Ser	Serine
Si	Silicon
SiH_4	Silicon tetrahydride
SLS	Static light scattering
SMAA	Silicomolybdic acid assay
SUMO	Small ubiquitin-related modifier
T_{agg}	Aggregation onset temperature
TB	Terrific Broth
TBDMS	*tert*-butyldimethylsilyl
TEOS	Tetraethylorthosilane
TEPS	Triethoxyphenylsilane
TES	Triethylsilyl
TF	Trigger factor
TIPS	Triiospropylsilyl
T_m	Melting temperature
TMS	Trimethylsilyl
TRIS	2-Amino-2-(hydroxymethyl)propane-1,3-diol
Trp	Tryptophan

1. Chapter 1: General Introduction

1.1 Silicon in nature

Silicon (Si) is an element with a diverse range of applications for humankind and the surrounding environment. In nature, the group IV metalloid exists as an oxide in sand, flint and quartz, or a silicate in clay and feldspar, rather than in its elemental form. It is the seventh most abundant element in the universe and is second to oxygen as the most abundant element on the earth's surface. Additionally it constitutes to over a quarter of the earth's crust by mass, with its most common compound, silicon dioxide occurring the most.[3–5] Silicon has three naturally occurring isotopes ^{28}Si, ^{29}Si and ^{30}Si and over twenty radioactive isotopes. ^{28}Si is the most abundant natural isotopes and ^{29}Si is exploited today in solid state silicon NMR Spectroscopy.[6,7] Silicon exists in two allotropic forms, amorphous and crystalline. Historically, the element has been a main component in tools used during the Stone Age, in the form of flint, and clay has been utilised to make ceramics for centuries. In our modern times it is still a main constituent in the basic building blocks used in construction and is found in many of the electronic products due to the semiconducting properties of silicon polymers. The importance of silicon is attributable to its unique and specific physical and chemical characteristics.[8,9]

1.2 Properties of silicon

Silicon has very unique physical and chemical properties in comparison to other group IV elements in the period table. At room temperature the element is a solid, noted to have a high melting point of 1420 °C and boiling point of ~3280 °C. The electron configuration of silicon, in its ground state, is [Ne]$3s^2 3p^2$, thus allowing four valence electrons to form a tetravalent structure.[9,10] Silicon is found in the same period group as carbon and have the same number of valence electrons to form a tetrahedral structure with other elements, hypervalent silicon species have also been described in some cases.[9,10] This ability to increase its covalency is thought to be attributable to the size of the silicon atom which is larger than

carbon, (silicon atomic radius of 1.17 Å, carbon atomic radius of 0.77 Å) and due to silicon being less electronegative (silicon, Pauling electronegativity value 1.9, carbon, Pauling electronegativity value 2.5)[10,11] Additionally, the electronegativity of silicon makes the atom more susceptible to nucleophilic attack when bonded to carbon, as the polarisation is shifted towards silicon. However, silicon has higher bond dissociation energies and longer bond lengths, thus forming very strong bonds to halides and oxygen in comparison to carbons bonding properties. Silicon can also form strong polymers which have a backbone of alternating oxygen and silicon atoms that can bind other atoms and functional groups, such as organic groups.[12] Recent studies have reported double and triple bond chemistry in the formation of silene, disilene, disilyne and silynes, although aided by the addition of phosphate ligands it was originally thought to be unstable in silicon due to its low value of electronegativity.[13]

1.3 Hypervalent silicon species

Since the 19th century when the first hypercoordinated silicon species were reported, SiF_6^{2-} and SiF_4-NH_3,[14] there has been much debate over a plausible explanation to the formation of these intermediate species. One of the main reasons these discoveries gained attention is that it contradicted the theory proposed by Gilbert Lewis, a concept we know as the octet rule.[15] Some of these potential explanations have deemed to be invalid, in particular the notion of the participation of the empty, high energy d orbitals in bonding. As computational chemistry calculations showed the hypervalency for silicon was provided mainly by 3s and 3p atomic orbitals with d orbital acting only to correct polarisation.[11,16] Pimentle and Rundle proposed a model known as the three centre four electron (3c-4e model) in 1951 which can be used to understand the formation of a pentavalent silicon species. This suggests hybridisation of three sp^2 orbitals (s, and p_x, p_y) in the central atom participates in equatorial bonds with ligand atoms and the remaining ligand orbitals hybridises with the axial p_z orbital to form three molecular orbitals, bonding, non-bonding and anti-bonding. With two electrons occupying the lower energy bonding orbitals, the remaining two electrons are in the non-bonding orbitals of the ligand atoms (Figure 1.1).[17–19] In this instance, as the two electrons in

the non-bonding orbitals are not part of the valence orbital of the central atom it agrees with the Lewis octet rule.[11,16]

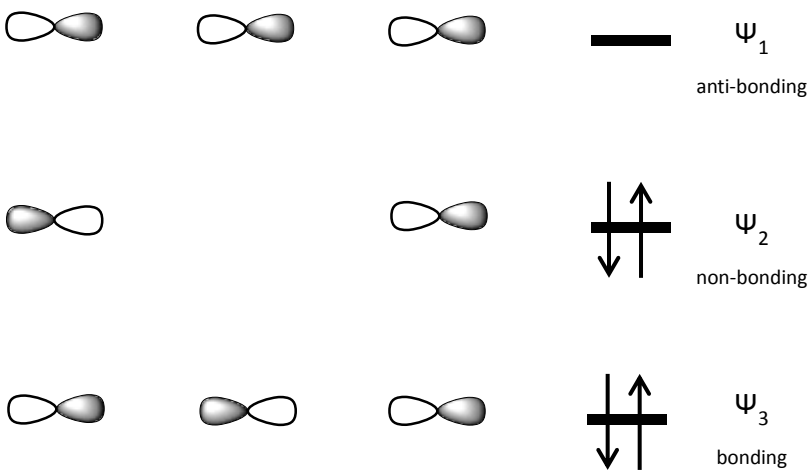


Figure 1.1 Representation of the Pimentle-Rundle model of three centre four electron (3c-4e) theory

1.4 Discovery of organosilicon compounds

The synthesis of organosilicon compounds has had an enormous impact on industry. Generally, the term organosilicon refers to the presence of one or more Si-C bond in the molecule, however within this overarching classification is a group of molecules called siloxanes with the distinctive motif C-Si-O. Tetraethylsilane (Figure 1.2A), was the first organosilicon compound formed by Charles Friedel and James Craft in 1863. The method used was adjusted to synthesise tetramethylsilane (Figure 1.2B).

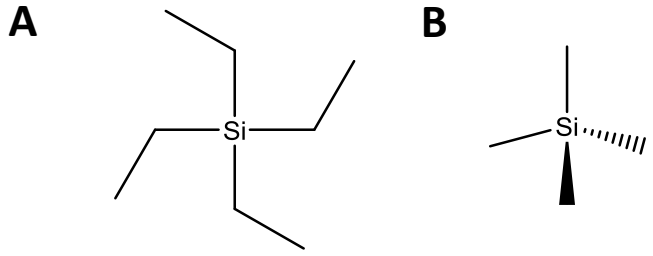


Figure 1.2 Examples of Organosilicon compounds (A) tetraethylsilane and (B) tetramethylsilane

In these reactions the silicon containing compound was reacted with diethylzinc and dimethlyzinc, to produce the relevant silane.[12] Subsequently, the utilisation of Grignard reagents by Frederic Stanley Kipping occurred in the 1904 to form organosilane compounds.[20] In the 1940s organosilanes were hydrolysed by James Franklin Hyde at Dow Corning Corporation to develop polysiloxane polymers, which were termed silicones, along with large scale production designed by Dr Eugene Rochow. In general terms, the distinguishing feature present in siloxanes are C-Si-O motifs. The formation of these silicones confirmed their importance as a valuable material with versatile applications owing to some unique properties such as oxidative and thermal stability, water resistance and low freezing point. To date they are still used in many consumer products and have a diverse range of industrial applications.[21] The chemical synthesis of silicones in the form of polydimethylsiloxane (PDMS), a polymer containing a silicon-oxygen backbone with two methyl groups attached to each silicon atom (Figure 1.3), is achieved by reaction of elemental silicon and methyl chloride to produce methyl chlorosilanes with subsequent hydrolysis in water to form siloxanes. However, these methods are costly requiring high temperatures of between 250-300 °C and pressures of 1-5 atm. They involve the use of chlorosilanes thus producing a high volume of waste and unwanted by products.[8,22] Therefore, a more sustainable and environmentally sound process of producing organosilicon compounds. Additionally, a route by where the current unwanted organosilicon by-products are recycled would be a huge asset with regards to environmental degradation issues and conform to green chemistry principles.

1. Isolation of silicon $SiO_2 + 2\,C \longrightarrow Si + 2\,CO$

2. Production of chloromethane $CH_3OH + HCl \longrightarrow CH_3Cl + H_2O$

3. Synthesis of chlorosilanes $2\,CH_3Cl + Si \longrightarrow Me_2SiCl_2$ +other silanes

Figure 1.3 The production of chlorosilanes using the Rochow process.

1.5 Use of organosilicon compounds in synthetic chemistry

1.5.1 Cross coupling reactions with organosilicon compounds using chemocatayst

Since their discovery, organosilicon compounds have been used in many synthetic chemical reactions with a variety of applications. One example is transition metal catalysed cross coupling reactions to form carbon-carbon bonds and silyl protecting groups. The use of organosilanes and organic halides in palladium mediated coupling reactions were first described in 1988 and has since been known as Hiyama coupling.[23] Organosilicon chemistry is attractive in comparison to other organometallic reagents as they are chemically stable, low in toxicity and very cost effective. The first stage in this reaction is the oxidative addition of the organic halide to palladium. This reaction requires a fluoride ion in the form of TASF or TBAF to form a pentavalent silicon species through the activation of the organosilane which allows transmetalation thus forming diorganopalladium(II) complex. Subsequently, the oxidative state of the transition metal is reduced with the formation of the carbon-carbon bond and the reaction cycle returns to its original starting point for successive oxidative additions (Figure 1.4). The use of fluoride in this reaction has been viewed as a limitation due to its high reactivity and many variations of this reaction have been investigated using different bases or coupling agents and thus the requirement of a pentavalent species has been in question.[24,25] Hiyama-Denmark coupling makes use of organosilanols as a coupling agent, Brønsted base as the activator and phosphate ligands to stabilise palladium. The mechanism of action works in a cycle with oxidative addition of the organic halide to the transition metal and generates the formation of a silanolate which activates the formation of a diorganopalladium(II) complex through an oxygen-palladium bond which accommodates transmetalation and subsequent release of polysiloxanes. Reductive elimination reduces palladium to its original valance state and allows the cycle to proceed again.[24] Further reactions have been completed using silylcyclobutane ring as coupling agents in the presence of fluoride ions and fluoride free methods using Lewis acids to allow milder bases such as K_3PO_4[26] and NaOH[25,27] to be used. The latter methods excluding fluoride are compatible with compounds that have silyl-protecting groups or other functional groups that may react readily with fluoride.

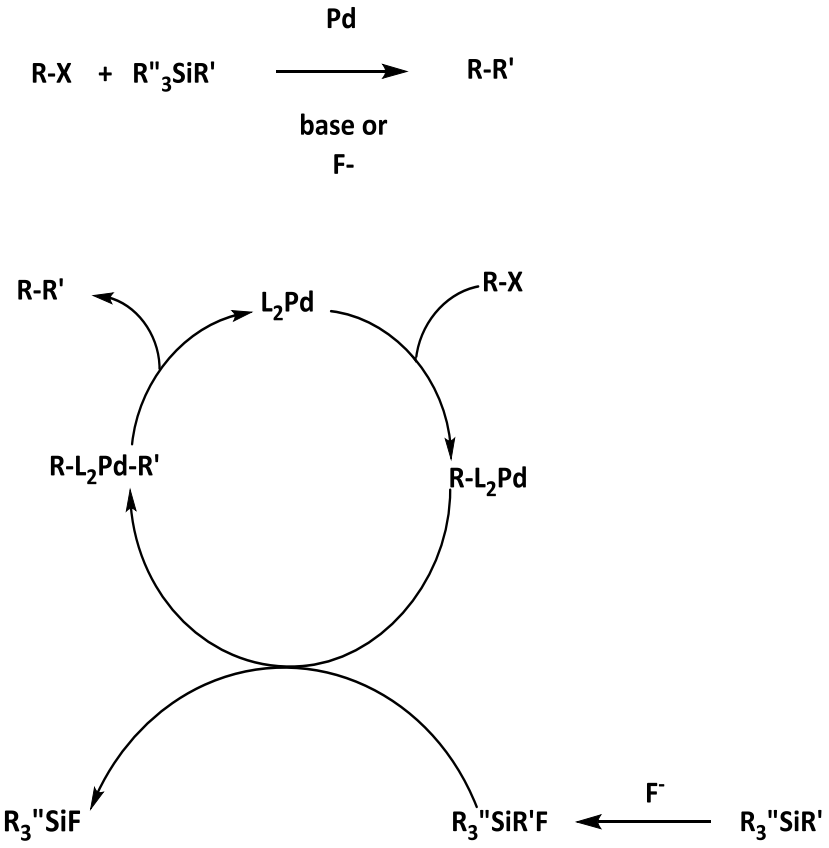


Figure 1.4 Catalytic cycle of Hiyama coupling mechanism.

1.5.2 Silyl protecting groups

The silylation of compounds in a reaction is a further example of the importance of organosilicon compounds in synthetic chemistry. The addition of silyl protecting groups can shield a particular functional group at specific stage in a reaction and be subsequently removed when required. The latter point occurs due to the shift in polarisation towards silicon in a Si-C bond.[12] Silyl protecting groups are important components in multi-step synthetic chemistry involving polyhydroxylated intermediates. They serve to selectively protect hydroxyl groups with subsequent deprotection and stop unwanted reactions or products in the intermediate stages of the synthesis. With a wide range of silyl protecting groups varying due to the electronic and steric properties of their substituents, thereby offers selectivity in protection of a particular hydroxyl functional group in a compound where several hydroxyl groups are present (Figure 1.5).[28]

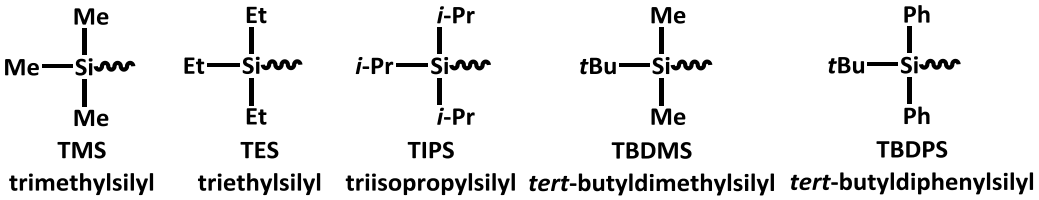


Figure 1.5 An example of some silyl protecting groups used in organic synthesis

1.5.3 Protection of silyl ethers

The traditional methods of silylation used silyl chloride and silyl triflates to protect hydroxyl functional groups[29–32] using many different catalysts. Enantioselective and site-selective silyl protection of hydroxyl group with a compound bearing multi- hydroxyl groups is imperative in many synthetic reactions.[28] With observations of silyl protection of 5' hydroxyl groups in nucleosides[33,34], asymmetric selective silylation of secondary alcohols[35–38] and poly-hydroxyl alcohols[39–41] However, some of the major drawbacks of using these halogenated compounds are the adverse environmental impact due to their corrosive properties and unwanted side products produced in the reaction. Therefore, there is a requirement for approaches that better conform to the principles of green chemistry.[42–44] Methods for protecting group free synthesis[44–46] have been described and have the advantage of omission of protection and deprotection steps, although, protecting groups are very much still pivotal components in chemical synthesis. Several groups have reported on enantioselective silyl protection of alcohols using very different and improved methods to circumvent the use of silyl chloride and silyl triflates, with improved separation and kinetic resolution of enantiomers.[28] Snapper *et al.* reported several instances of enantioselective silylation of asymmetric and cyclic diols catalysed by chiral and achiral Lewis bases and further optimisation of this with the utilisation of a co-catalyst. This combination served two separate functions during the reaction through nucleophilic activation of the silyl substrate and Brønsted base activation of the hydroxy group. However, the limitation to these reactions were a high concentration of the chemocatalysts, although the addition of the co-catalyst did reduce the overall reaction time[41] (Figure 1.6).

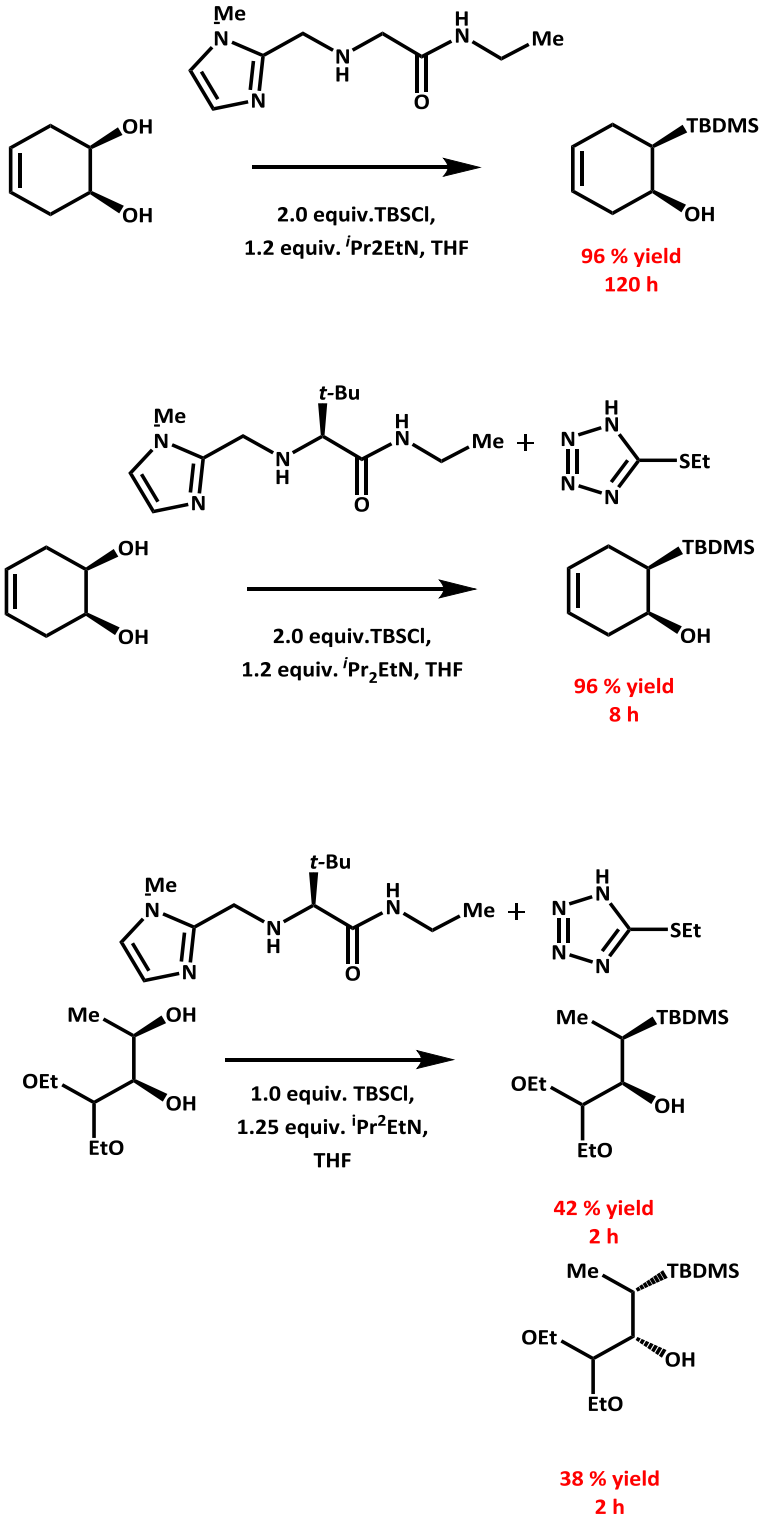


Figure 1.6 Enantioselective silylation of asymmetric and cyclic diols[41]

Hydrosilanes have been utilised in the silylation of alcohols both with metal catalysts such as Rh (II) carboxylates[28,47] and Lewis acids such as $InBr_3$[42] with the benefit of only hydrogen as effluent from the reaction. Other groups have reported kinetic resolution of racemic mixtures, in particular secondary alcohols, through stereoselective silylation using a copper catalyst and hydrosilanes.[48–50] Recent work has been aimed at enzyme mediated silyl protection using a variety of different hydrolases whose native function is not inclusive of silicon oxygen bond formation[51,52] and silicatein-α previously described to catalyse polycondensation of biosilica[53] Experimental data produced from the work involving silicatein-α is of particular interest as is shows promising application of the enzyme as a biocatalyst in organosilicon chemistry and as a potential in catalyst in silyl protection.

1.5.4 Deprotection of silyl ethers

Selective silyl deprotection methods used within multi-step organic synthesis have been extensively reviewed[54,55]. The selectivity is based upon electronic and steric effects of the substituents of the silyl protecting group around the silicon atom under certain reaction conditions and also considering the presence of other hydroxyl groups. Generally the steric bulk of the substituents can decrease silyl ether cleavage under both acidic and basic conditions.[28] Additionally, in the case of the same protecting group used for more than one hydroxyl group in the compound, the less sterically hindered will be cleaved first. However, selective deprotection of bulkier silyl protecting groups in the presence of labile and smaller silyl protecting groups can be manipulated through electronic effects. The rate of hydrolysis of silyl ethers under acidic conditions can be decreased with electron withdrawing groups on the silicon atom (TMS<TES<TBDMS<TIPS<TBDPS) while under basic conditions electron donating groups will decrease the rate of hydrolysis (TMS<TES<TBDMS~TBDPS<TIPS)[28]

Though a major limitation to these processes are the reagents and chemical catalysts involved in these reactions are harsh, costly and non-eco-friendly plus removal of protecting groups can be a rate limiting factor. The latter limitation has been addressed with the development of techniques aimed to improve removal of excess reagents following deprotection.[56,57] In addition, as previously mentioned, protecting group free methods have been documented for certain synthetic reactions offering a more ‘green’ approach.[44–46]

However, the application of an enzyme that could mediate in the formation of organosilicon compounds and attachment of silyl protecting groups would be very beneficial. The main advantages biocatalysts offer are regio-, stereo- and enantioselectivity in addition to operating under more ambient conditions, therefore, being more favourable to chemocatalysts.

1.6 Marine biosilification

Biosilification is a structure building process in which the polymerisation of inorganic silicic acid is incorporated into organic structures of unicellular diatoms and multi-cellular marine sponges. Diatoms, unicellular algae, takes up silicic acid from the aqueous environment with the successive incorporation into the cell wall, termed frustules, with the aid of proteins, silaffins and long chain polyamines.[58] In contrast, marine sponges are multicellular, sessile organisms from the phylum Porifera (meaning pore bearer). They are divided into four classes, *Calcarea*, *Hexactinellid*, *Demospongiae* and *Homoscleromorpha*, classified by anatomical and physiological structure. The latter three aforementioned classes produce small needle like structures known as spicules that are comprised of silica (silicon dioxide SiO_2) and protein.[58,59]The rigid, siliceous spicules formed by specialised cells, sclerocytes, function as a support to these otherwise pliable invertebrates, protecting them from the undulatory motion of oceanic waves, aid in reproduction and direction toward water column. The spicules are diverse in their chemical configuration depending upon species and the levels of soluble silicic acid in the surrounding environment but mainly comprise of silica, water and other trace elements[60] The formation of these spicules is facilitated by the enzyme silicatein which functions in the polycondensation of biosilica (Figure 1.7). However, the process occurs via different mechanisms in each organism.[58–60] The identification of silicatein is of particular interest due to the potential use as a biocatalyst in organosilicon chemistry.

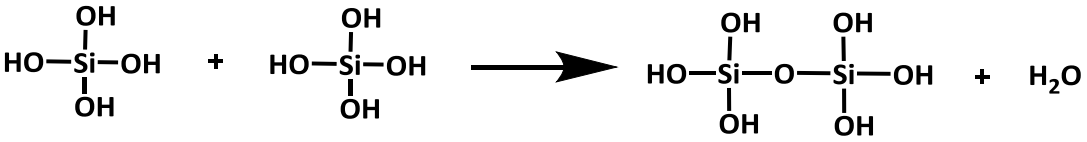


Figure 1.7 Condensation reaction of silicic acid mediated by silicatein.

1.7 Discovery of silicatein

The discovery of silicatein, a key catalyst in the process of biosilification, occurred in 1998 when the protein was isolated from the Pacific sponge *Tethya aurantia* (*T. aurantia*). It was found to exist in this organism in three isoforms silicatein α, β, γ in a ratio of 12:6:1. The most abundant isoform, silicatein-α, was identified to have sequence homology of 45% to the human protease cathepsin L with conservation of 3 disulphide bonds. However, a difference lay in the catalytic site where a cysteine (Cys) residue found in cathepsin L is replaced by a serine (Ser) residue in silicatein-α. The other residues that comprise the catalytic triad in both proteins remain the same, histidine (His) and asparagine (Asn) (Figure 1.8A; Figure 1.8B). An additional difference noted was silicatein-α contained more hydroxyl amino acids.[61] More recent studies have shown that some silicatein-α may exist with catalytic Cys instead of Ser.[62] Since its discovery in *T. aurantia*, silicatein-α has been found in many other sponges, such as a dimerising silicatein in *Petrosia ficiformis*[63] and silicatein-α from *Suberites domuncula* (*S. domuncula*).[64] The translated protein of silicatein-α gene from *S. domuncula* was characterised as being a pre pro-enzyme of 36 kDa. This comprises of a signal peptide sequence, a pro-enzyme and an enzymatically active mature protein. It has been suggested that the protein undergoes autocatalysis for the production of its mature and active form, however, the exact molecular mechanism by which this occurs remains unknown at this present moment.[64,65] To date no cofactors that aid catalytic activity have been identified, although, concurrent upregulation of the collagen gene was observed with transcription of the silicatein gene in response to an increase in silicate[64,66] and Fe^{2+} has been reported to promote the growth of cells involved in spiculogenesis, while Mg^{2+} and EDTA have no effect.[59,65,67] Previous work suggests the enzyme may undergo post-translational modifications through the identification of potential phosphorylation sites.[67,68] Additionally, interactions with a family of Ca^{2+} binding proteins, silintaphins, have been proposed to aid in the structure directing process of biosilification mediated by silicatein.[69,70] However, the mechanism by which silicatein functions in biosilification and the role of components has only been proposed thus far.

A

```
Cathepsin L   QGQCGSCWAF SSEDMDHGVL NKYWLVKNSWG
Silicatein    QGDCGASYAF SSSSLNHAMV KKYWLAKNSWG
```

B

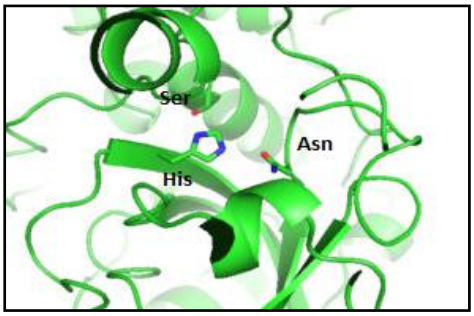


Figure 1.8 (A) Truncated sequence alignment of human cathepsin L (UniProtKB P07711) and silicatein-α from *S. domuncula* (UniProtKB Q2MEV3) to show the conserved residues, histidine and asparagine and serine/cysteine in the catalytic triad. (B) predicted model of silicatein-α showing the catalytic residues.

1.8 Proposed mechanism of silicatein

The molecular mechanism of silicatein remains elusive due to the lack of structural information using crystallography or NMR, although, several have been proposed using various different modelling methods.[1,2,71] One of the proposed catalytic mechanisms for silicatein is based on a SN_2 reaction which initially involves nucleophilic attack between the electronegative oxygen atom of serine hydroxyl group (OH) and the electropositive Silicon (Si) atom of the model substrate, silicic acid. This step in the process is thought to be enhanced by a hydrogen bond between the OH of Serine and the imidazole nitrogen of Histidine (His) which makes the serine OH more nucleophilic. Subsequently the transfer of a proton from His nitrogen to an OH group of silicic acid results in the formation of a pentavalent intermediate and the release of water (H_2O). At this point the silicic acid has formed a covalent bond with serine and proceeds in a nucleophilic attack on a Si atom of another silicic acid that enters the catalytic site. This stage in the reaction occurs due to a hydrogen bond between His imidazole nitrogen and the covalently bound substrate. A water molecule is subsequently released after the transfer of a proton from the imidazole nitrogen resulting in a disilicic acid bound to the enzyme through an ester like bond. The rotation of this bond facilitates a hydrogen bridge between His imidazole nitrogen and hydroxyl group of silicic acid thus aiding further nucleophilic attack to another silicic acid molecule to form a trisilicic acid. As the reaction

cycle continues involving nucleophilic attack, transfer of a proton, loss of H_2O and rotation of the ester like bond, silica is polymerised with the possible formation of siloxane ring structures.[2] The third member of the catalytic triad functions in the transfer of protons and to stabilise the substrate in the catalytic centre through the formation of a hydrogen bridge.[2] However, this mechanism was based upon data collated from molecular docking experiments and therefore ambiguous as the formation of siloxane ring structures is unlikely due to steric reasons.

Earlier studies explaining the mechanism of silicatein described silicon ethoxide condensation which was based on known mechanisms of serine and cysteine proteases. This proposal also suggests the nucleophilicity of the oxygen atom of serine is increased by the hydrogen bond between the imidazole nitrogen of histidine and the hydroxyl group of serine, thus causing nucleophilic attack of silicon and displacement of ethanol to form a covalent enzyme-O-Si intermediate. Following this an alkoxide bond is formed through hydrolysis by the addition of H_2O. A reactive Si-O is released and the oxygen can undergo a further nucleophilic attack on another substrate molecule resulting in the formation of disiloxane.[1] However, this proposed mechanism is illustrative of silicatein-α function in the formation of a disiloxane and does not elaborate on how it could mediate condensation reaction to form polysiloxanes. Additionally, only one substrate was screened which is analogous to the proposed natural substrate of the enzyme. An adaptation of this is illustrated in Figure 1.9 showing silicatein-α as a hydrolytic enzyme and subsequent spontaneous condensation. During this work, single mutations of the active site serine residue and histidine residue to alanine residues were investigated for the causal effect on catalytic activity. The results revealed a decrease in activity to 8 % for Ser26Ala and 10 % for His165Ala in comparison to the relative catalytic specific activity of the wild type enzyme, thus indicating the residues to be fundamental in overall enzyme activity and efficiency.

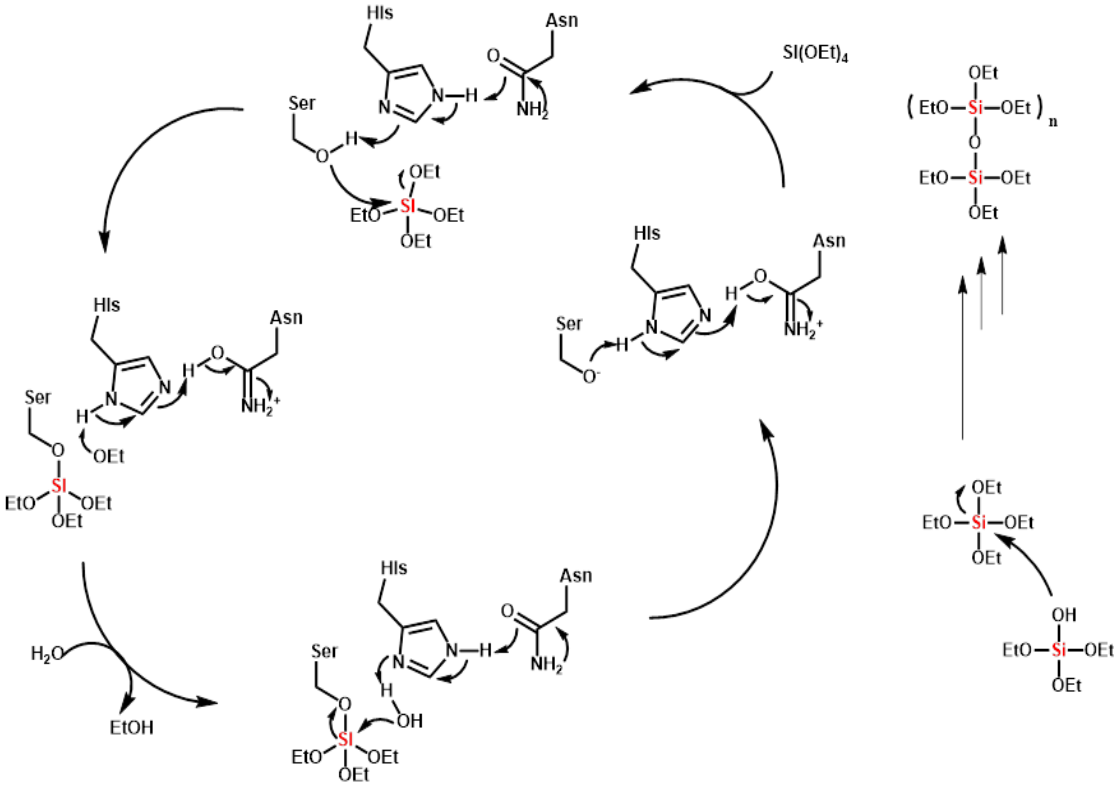


Figure 1.9 Reaction scheme with proposed mechanism of silicatein-α for hydrolysis and spontaneous condensation of alkoxysilane TEOS

In another study which used the crystal structure of mutant silicatein-cathepsin L chimera led to the proposal that histidine in the catalytic site deprotonates silicic acid with proton transfer being facilitated by a network of water molecules leading to increased nucleophilicity of the oxygen atom on the silicic acid, which may in turn result in an attack on a second substrate molecule leading to polymerisation. The reaction is believed to be promoted by proton shuffling to water and elimination of H_2O leading to a loss of the hydroxyl group on silicic acid (Figure 1.10).[71] However, the data lacked exploration of a range of different experimental conditions that would identify optimum working ranges and conditions.

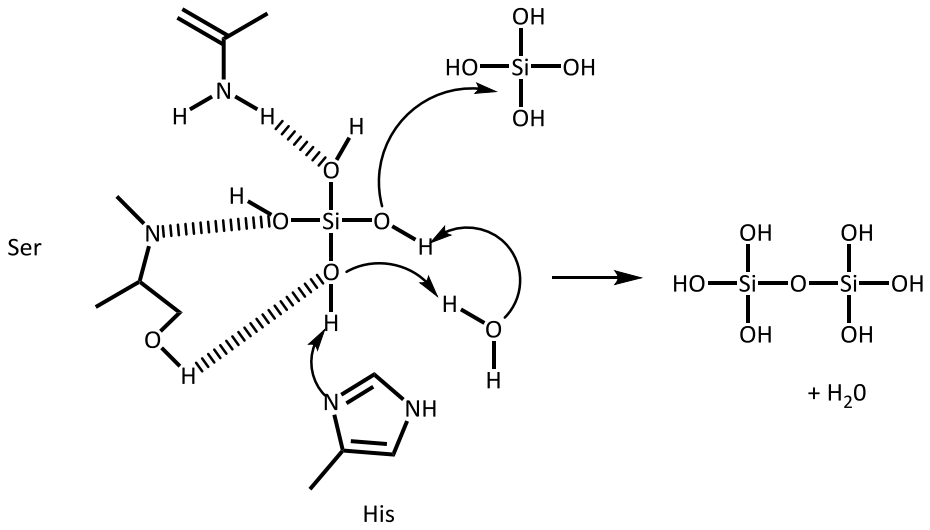


Figure 1.10 Proposed mechanism of silicatein-α for silicic acid condensation determined from molecular docking studies using the crystal structure of cathsilicatein (cathepsinL/silicatein chimera).

However, the lack of systematic mechanistic and structural studies means there is little evidence for any of these proposed mechanisms. Site-directed mutagenesis of the serine and histidine residues that comprise the catalytic triad, however, where shown to have a significant reduction in enzymatic activity, denoting the importance of these particular residues in the reaction mechanism. [72]

1.9 Hydrolases

Hydrolases are EC Class 3 enzymes abundant in many organisms and function in many essential physiological hydrolysis reactions. This superfamily of enzymes can be further sub-classified according to the particular bond they hydrolyse. For example, sub classes for ester bond hydrolysis (EC 3.1), ether bond hydrolysis (EC 3.3) and peptide bond hydrolysis (EC 3.4). Enzymes can then be further sub divided according to structural, mechanistic and amino acid sequence features.[73]

1.9.1 Serine hydrolases

Serine hydrolases are aptly named due to the presence of a nucleophilic Ser residue in the active site, along with conserved catalytic residues His and Asp. Another conserved feature is

the oxyanion hole, which participates in enzymatic function, through stabilisation of substrate binding and the transition state. They comprise one third of all known hydrolases and mediate many essential processes including homeostasis and digestion, through hydrolysis of peptide and ester bonds.[74–76]

Serine proteases in particular chymotrypsin have been extensively studied and used as a model enzyme for mechanistic studies due to the various catalytic techniques they utilise. The catalytic process uses hydrogen bonding for charge stabilization, acid-base catalysis, covalent catalysis, deprotonation of Ser195 by His57 for nucleophilic attack with protonation by His57 to enable release of C-terminus peptide. Also the active site provides geometric fit to accommodate the tetrahedral alkoxide intermediate for reduction of energy of the transition state.[74–76] As the proposed mechanism of silicatein-α was postulated to be similar to the mechanism of serine and cysteine proteases[1] it is necessary to understand the catalytic function of these enzymes.

1.9.2 Mechanism of serine hydrolase chymotrypsin

The catalytic mechanism of serine hydrolases is illustrated by Figure 1.11. The substrate binds, due to a recognition site within the active site of the enzyme, and the substrate is stabilised through hydrogen bonding by the oxyanion hole, formed from amide protons in the peptide backbone of residues Gly193 and catalytic Ser195.[73]

Subsequent nucleophilic attack of the carbonyl carbon by Ser195 is facilitated through the serine hydrolases characteristic ‘charge relay system’ of the catalytic triad residues, which generates an alkoxide ion on the serine activated through deprotonation by histidine. The resultant negative charge on the carbonyl oxygen of the substrate is stabilised by hydrogen bonding with oxyanion residues forming a tetrahedral intermediate. The unstable negative charge causes reformation of the double bond and collapse of the intermediate.[73]

Cleavage of the peptide bond through protonation of the amide nitrogen by His 57 occurs thereafter with release of C-terminal of the peptide substrate. This also results in formation of covalent acyl-enzyme intermediate with the remaining N-terminal of the peptide substrate. A water molecule enters the active site for ester hydrolysis through attack of the carbonyl carbon. This leads to formation of a second oxyanion tetrahedral intermediate. Acid catalysis

breaks the covalent acyl-enzyme bond leading to dissociation of the second peptide produce and regeneration of the free enzyme state (Figure 1.11).[73]

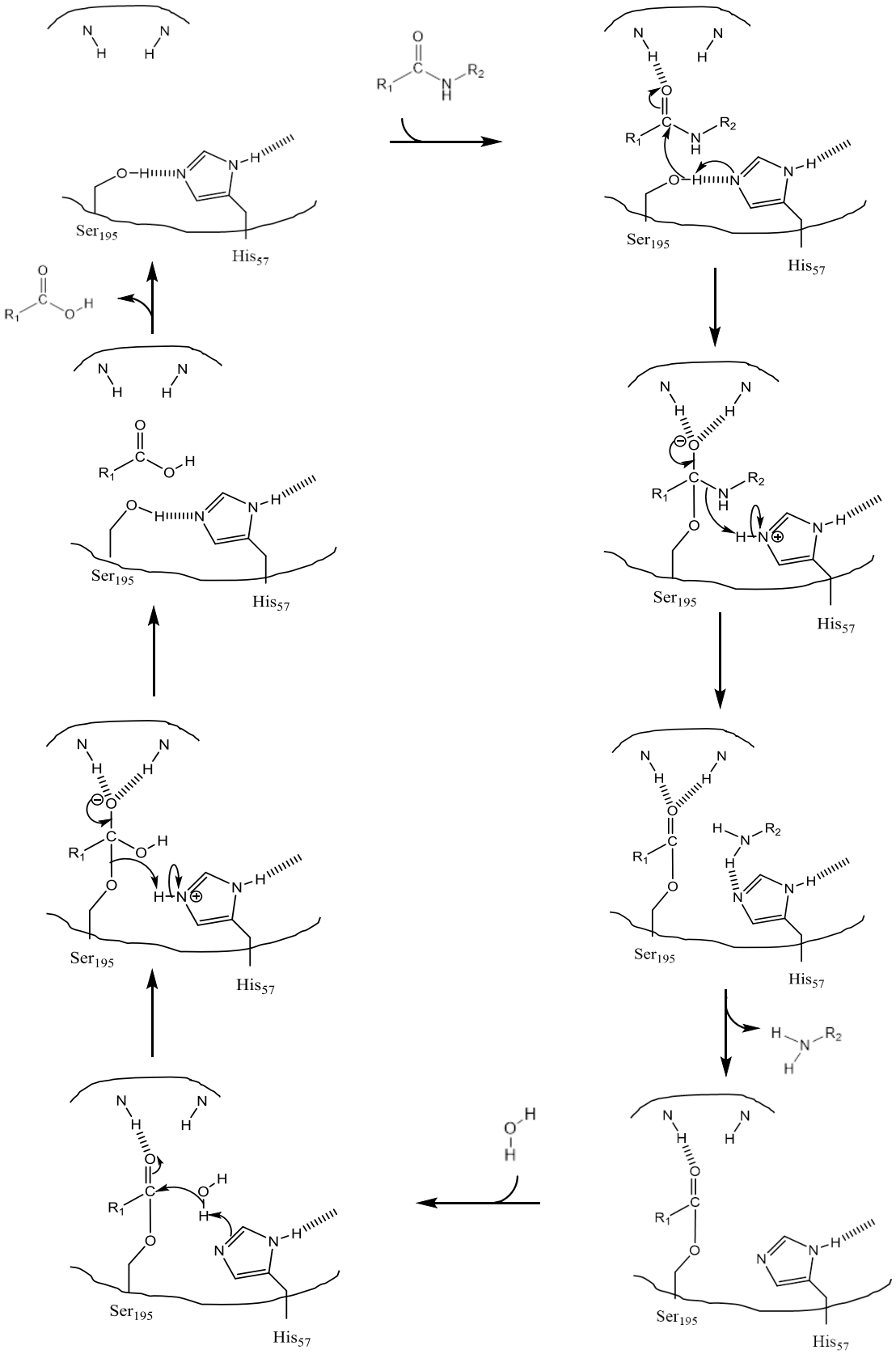


Figure 1.11 Representation of catalytic mechanism of serine hydrolase, chymotrypsin

In 1954, B. S Harley and B. A. Kilby proposed the formation of a covalent enzyme-substrate intermediate during the catalytic mechanism of serine hydrolases. A colorimetric assay using chymotrypsin and a non-natural substrate, p-nitrophenylacetate, enabled product formation to be monitored due to the release of p-nitrophenolate (Figure 1.12). In the initial stages of the reaction they observed a burst phase in product formation prior to steady state being reached. It was suggested the rapid release of product was equivalent to 1 mole of product formed by 1 mole of enzyme. The rate of reaction of reaction was then observed to decrease in comparison due to formation of the enzyme-substrate complex and the subsequent hydrolysis and release of the product at this stage being rate limiting. The use of this particular substrate allowed for observation of the burst phase which may not be noticeable during more efficient catalysis of the natural substrate.[77] This enabled Michaelis-Menton kinetic parameters to be obtained for the enzyme using these substrates.[53,78]

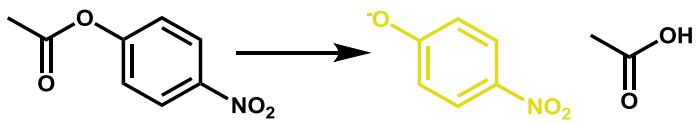


Figure 1.12 Enzyme-catalysed reaction with p-nitrophenylacetate to p-nitrophenolate

1.10 Biocatalysts in synthetic chemistry

Biocatalysis can be defined as the use of natural or modified isolated catalysts such as protein enzymes in the use of chemical or biological transformations. The utilisation of enzymes in such processes stem back centuries, for example enzymes in yeast are utilised for the production of ethanol and carbon dioxide though fermentation.[79] They are more favourable over chemical catalysts as they are often less energy demanding and produce less waste thus being overall more economically and environmentally friendly conforming to the principles of green chemistry.[80] Due to major advances, from the 1970s onwards, in the knowledge of protein structure-function relationships, protein expression through recombinant technology and of protein modification techniques such as directed evolution, the limitations that once prevented the use of some potential biocatalysts are now being addressed. Consequently, the numbers of biocatalysts have significantly increased in the last 40 years due to increased

stereo-, regio- and chemoselectivity and have a diverse range of applications in the pharmaceutical industry and large scale production of fine chemicals.[79–82] Pharmaceutical companies have utilised asymmetric biocatalysis to manufacture and obtain enantiopure drugs that may reduce unwanted side effects and aid efficacy. This can be obtained through kinetic resolution of a racemic mixture in which a biocatalyst with a chiral centre will react with one enantiomer in the mixture to produce the pure form of the desired drug with a 50% maximum yield. [83] However, this limitation of a 50% yield can be resolved by deracemization which gives a yield of up to 100% by converting a racemic mixture into one enantiomer. An example of this method was observed in the resolution of a racemic mixture containing (*R,S*)-ibuprofen using lipase from *Candida rugosa* and 1-propanol. Esterification of (*R*)-ibuprofen proceeded preferentially under optimum conditions and therefore allowed conversion of nearly 100% to the pharmacologically active (*S*)-ibuprofen (Figure 1.13).[84] As kinetic resolutions and deracemizations may require additional steps, desymmetrization is one method that can produce high titres in a one step process, and produce chiral end products from symmetric substrates.[83,85] Asymmetric biocatalytic reactions can be coupled to ultimately produce the desired product without generating a copious amount of waste products and thereby maximising efficiency of the overall reaction.[86] The continuing advancement in current knowledge of genomic information and the characterisation of catalysts that mediate molecular processes has led to the identification of potential biocatalysts.

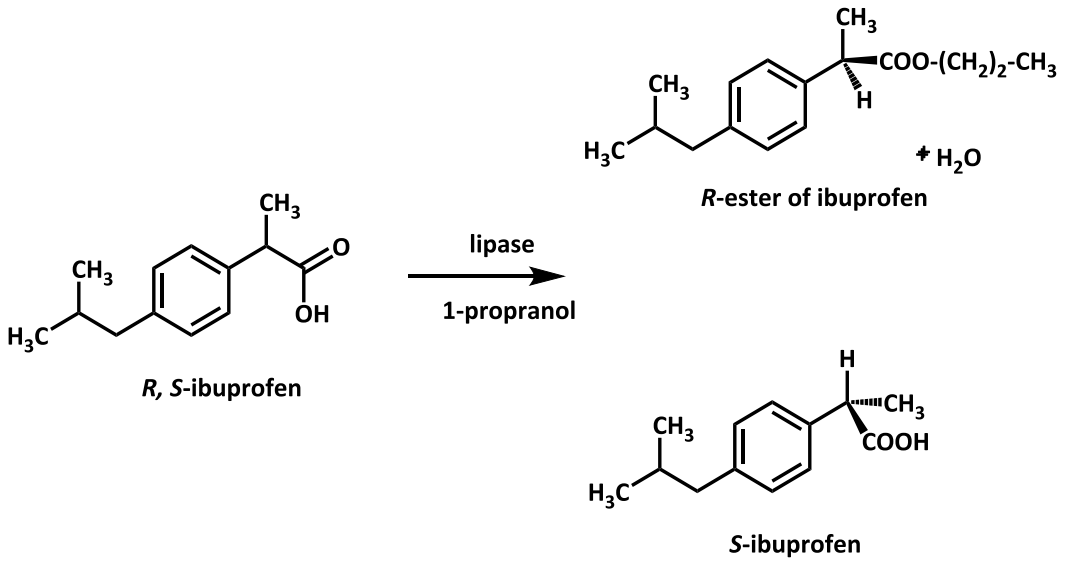


Figure 1.13 Enantioselective esterification of R-ibuprofen with lipase and 1-propranol

1.11 Silicatein and metal oxides

The development of new biocatalytic nanotechnology methods potentially offers low cost, low energy, near neutral pH and often chemical free way of depositing nanosized particles onto the desired surface. This is of particular interest for companies that manufacture silica based materials used in many electronic devices, biosensors and optics. Silicatein-α has been shown to be a good candidate in this process due to its versatility in catalysing a number of other substrates,[87–93] in addition to precursors of silica.[94–96] Immobilised silicatein-α from *S. domuncula* retained its catalytic activity and was observed to deposit biosilica from a precursor, tetraethoxysilane (TEOS), onto a gold surface. It was noted that silicatein had a dual purpose through this process as it catalysed the polycondensation and directed the structure of the silica nanoparticles onto the solid surface. The silica nanoparticles were observed to be deposited in close proximity to the active site of the immobilised enzyme, however, the overall coverage of the solid surface was described to be heterogeneous with only 70% containing a biosilica film. Therefore, this method lacked control over the consistency of the deposited silica, concluded to be due to the length of the reaction time or the immobilisation of silicatein on the surface being heterogeneous.[94] This was addressed and somewhat resolved by further investigations with cystamine or cysteamine and glutaraldehyde linked to the gold surface which immobilised silicatein and enabled fabrication of biosilica films onto the surface with the desired uniformity.[95] In addition to this, the assessment of biosilica formed *in vitro* and its interaction with a source of light, exposure to light and optical properties suggest a strong compatibility with photolithography methods and its use to coat analytical devices.[96] In a number of other investigations silicatein has been observed to catalyse reactions using other substrates to form a number of different products, thereby suggesting that it is not exclusive in acting on precursors of silica. The enzyme is able to hydrolyse and conduct polycondensation reactions involving metal oxides with semiconducting properties at ambient temperatures and neutral pH. Silicatein from *T. aurantia* was shown to produce spinel gallium oxide or gallium oxohydroxide, from the precursor gallium oxide *in vitro*. This could have potential uses in nano patterning of these substrates for the use in electronic devices.[87,88] In addition, the enzyme was observed to catalyse and yield titanium dioxide (TiO_2) from an alkoxide of titanium and then produce anatase and rutile, varying crystal structures of TiO_2[90] and produce a bimetallic perovskite-

like nanostructure composed of crystalline barium oxofluorotitanate $BaTiOF_4$ along silicatein filaments using the precursor $BaTiF_6$.[90] More recently, *in vivo* studies using silicatein from *S. domuncula* incorporated TiO_2 into newly synthesised spicules alongside SiO_2.[92] Silicatein can be immobilised on to TiO_2 nanowires, which are modified with Ni^{2+}-NTA ligand using His tags to create hybrid nanoparticles of TiO_2/Au through the reduction of tetrachloroauric acid ($HAuCl_4$)[92] or synthesise nanopatterning of TiO_2 and ZrO_2 from alkoxide precursors that are stable in water, titanium bis-(ammonium-lactato)-dihydroxide and hexafluorozirconate respectively.[93] The versatility of silicatein to act as a biocatalyst in the hydrolysis and polycondensation of these semiconducting metal oxides in aqueous solution at low temperatures, neutral pH and without the addition of harsh chemicals, in addition to its natural structure directing ability, makes it an attractive alternative in manufacturing electronic devices and bioelectronics.

1.12 Condensation of siloxanes using enzymes

As previously discussed, the current methods used for selective silyl protection and deprotection are detrimental to the environment due the harsh chemical catalysts and reaction conditions used. Therefore, research is often focused on a 'greener' approach with the possibly of employing enzymes as catalysts as they operate under more ambient and less toxic and hazardous conditions.

The hydrolysis of organosiloxanes using enzymes was conducted by Peter Taylor's group in which several serine hydrolases were used and screened with commercially sourced silyl ether bond compounds commercially sourced with silyl ether bond. However, none of the enzymes screened showed hydrolytic activity towards the substrates. They concluded that this may be due to steric hindrance of the bulky moieties of the substrates or could other factors not documented.[51,52]

Peter Taylor's research also investigated etherification and trans-etherification of silyl protection groups with a primary alcohol (octan-1-ol) using a variety of hydrolases whose native function is not inclusive of silicon oxygen bond formation (Figure 1.14).[51,52]

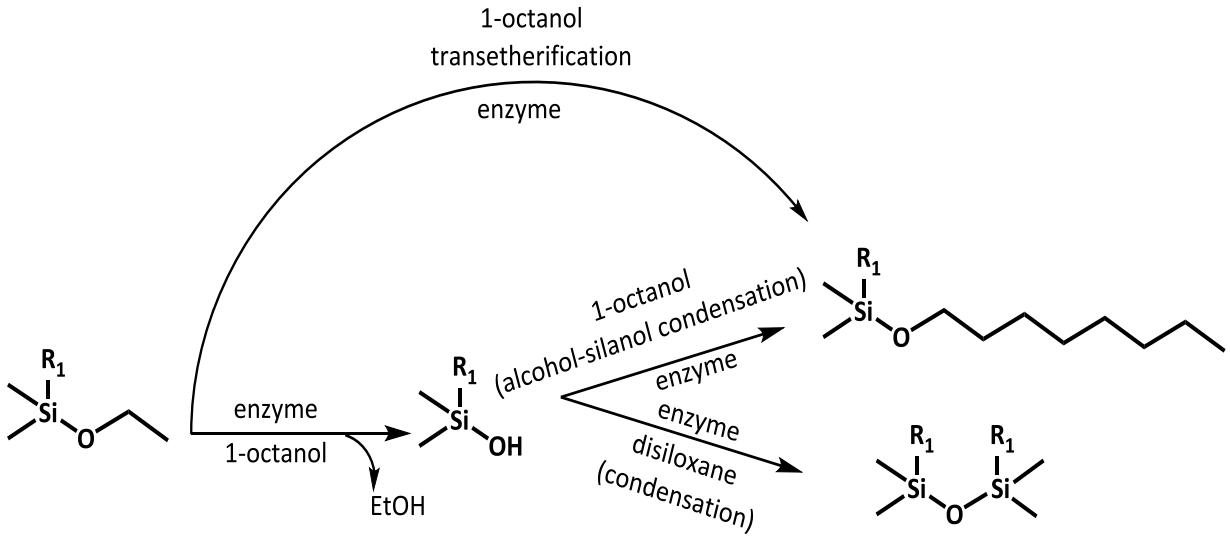


Figure 1.14 Schematic of hydrolysis and condensation of siloxanes mediated by enzymes. The diagram was redrawn without alteration.[52]

As silicatein-α was previously postulated to catalyse the polycondensation of biosilica in its native surroundings research within our group was aimed at developing this enzyme as a potential biocatalyst.[53,78] Recently published experimental data revealed that heterologously produced silicatein-α was able to catalyse silyl etherification and transetherification of a primary alcohol with three model silanols (TMS-OH, TES-OH and DMPS-OH) and ethoxysilanes (TMS-OEt, TES-OEt and DMPS-OEt)in *n*-octane.[53,78] These bulky silyl protection groups were accepted by the enzyme in organic solvent suggesting the enzyme to be a robust, versatile and stable catalyst under these conditions. Regioselectivity of silyl etherification mediated by silicatein-α was also investigated using multi-hydroxylated substrates with both aliphatic and aromatic alcohol functional groups. This data showed a clear preference and selectivity in enzymatic silyl protection catalysed by silicatein-α of aromatic hydroxyl function groups in the substrates that were screened.[53,78] The results produced from the work involving silicatein-α are of particular interest as they shows promising application of the enzyme as a biocatalyst in organosilicon chemistry and as a potential in catalyst in silyl protection.

1.13 Hydrolysis of siloxanes using enzymes

Biotransformations aimed at the Si-O bond cleavage of varying organosiloxane substrates have been the focus of several research groups.[97–100] The utilisation of microbial cultures to enhance degradation of several silicon based substrates was first demonstrated by Semprini's group.[97] Hydrolytic reactions mediated by enzymes, in particular hydrolases such as esterases, lipases and trypsin like enzymes were investigated on a range of different silicon containing substrates with dissimilar results. In contrast, the application of an serine hydrolase from *Rhodotorula mucilaginosa* was shown to catalyse the hydrolysis of silatrane substrates.[99] However, research in this area completed by Grogan and co-workers proved to be unsuccessful. This work was aimed at screening numerous enzymes, predominantly serine proteases, and assessing their catalytic activity using organosilane compounds as substrates. As no positive results were obtained they concluded this to be due to steric hindrance from the bulky substituents of the substrates.[98] Additionally Bassindale *et al.* produced experimental data that did reveal hydrolysis of such substrates, although, this was documented to be from the effect of non-specific catalysis rather than enzymatic.[100]

1.14 Enzymes in organic solvents

The use of enzymes for chemical transformations in aqueous media with academic and industrial applications comes with many benefits. They are often conducted at ambient temperatures and pressures, and are, therefore, generally more energy and economically efficient. However, the downside of emulating the natural environment is the consideration of several factors that may affect the optimum conditions required for enzyme activity and maintain structural stability. These factors include optimum pH, temperature, and correct concentrations of essential co-factors, substrate and enzyme.[101] This also restricts the use of many organic compounds as potential enzyme substrates due to incompatibility with aqueous environments. In addition, unfavourable thermodynamic equilibria such as condensation reactions mediated by hydrolases and problematic product recovery are well known limiting factors associated with aqueous media.[102–105]

Enzymatic reactions conducted in anhydrous organic solvent have been the focus of research over the past 30 years. This was due to the revelation of many advantages, most of which

were the aforementioned limitations associated with reactions conducted in aqueous media. Other advantages included enzyme insolubility facilitating retrieval and reuse plus an increase in enzyme stability, thus preventing structural distortion ensuing to denaturation and aggregation. This contradicted the belief that the structural integrity of the enzyme would be lost and its catalytic ability detrimentally affected in anhydrous solvents. This was based on the notion that enzymes would lose the hydrative effect of water molecules surrounding it which act to make it move more dynamically and give it more conformational flexibility. It was, therefore, thought that this loss would cause the enzyme to unfold. However, enzymes have been shown to rigidify and hold native structure in little or no water (~10%) in an organic solvent media.[105,106] Enzyme conversions performed under these conditions have also shown to have an increased tolerance to temperature with an increased thermostability far above the optimum temperature under aqueous conditions. Another observation has been termed pH memory, in which a lyophilised protein would retain its original ionisation state when applied to a solvent system.[105,107,108]

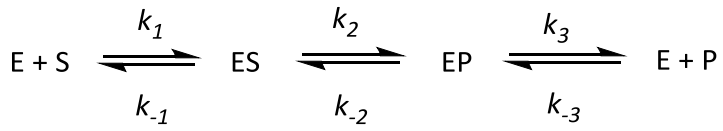
These advantages expanded the number of biotransformations particularly with hydrophobic substrates that are often incompatible in aqueous solutions due to immiscibility. In addition as the thermodynamic equilibrium could be shifted towards the reverse reaction as this was more favourable, early research focused on enzyme mediated etherification and transetherifications with hydrolases.[108–110] This was later expanded to asymmetric oxidoreductases with ADH.[111] However, the system still required optimisation mainly concerning the volume of water needed for the enzyme to retain its original structure and function. For example, in the case of lipases they required the addition of very little water (<0.02% vol)[112–114] for retention of functionality, however, in contrast the activity of chymotrypsin increased when applied to a biphasic system or in a more water-miscible solvent.[105,108,113,114] Ryu and Dordick concluded that the solvent system can enhance enzyme-substrate interaction through (1) stabilising the ground state of the substrate or destabilisation of the interaction, (2) increasing hydrogen bonding by entering the active site of the enzyme thus lowering local polarity and (3) indirectly altering the tertiary structure favourably so the active site is more accessible. [109,114–116]

1.15 Biochemical methods for analysing silicatein-α activity

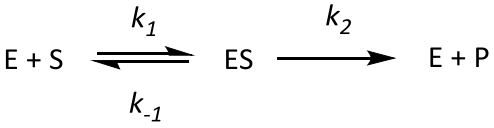
Since the discovery of silicatein-α was reported in 1998, the published method of biochemical analysis used to assess catalytic ability was the silicomolybdic acid assay.[1,61] However, this method is not high-throughput and only allows a semi-quantitative measurement of the enzymes catalytic ability. In addition, it is unsuitable for assessing and screening organosilicon substrates. Owing to the recent development of a high-throughput colorimetric assay, quantification of kinetic parameters can be obtained using organosilicon compounds.[53,78]

1.16 Catalytic efficiency of biocatalysts using Michaelis-Menten kinetic parameters [117–119]

Enzymes function to lower the activation energy required for a chemical reaction to proceed and increase the rate at which it reaches equilibrium. The catalytic efficiency of enzymes in mediating a particular chemical reaction can be assessed by studying its kinetic parameters. One of the most established and well used methods is Michaelis-Menten (MM) kinetic model which used a general model of a first order reaction proposed by Victor Henri. This model describes a reaction in which free enzyme (E) and substrate (S) interact to form and enzyme-substrate complex (ES), which progresses to an enzyme-product complex (EP), and subsequent release of free enzyme (E) and product (P), with all stages of reaction being reversible.



The equation can be further simplified by assuming that the rate constant k_{-2} would be negligible during the initial stages of the reaction and that product formation would be a rapid step. Therefore the equation can be simplified.



A further assumption used in the MM model was developed by Briggs and Haldane in 1925, in which the concentration of ES would be constant under steady state conditions. This follows an initial burst phase or pre-steady state where substrate concentrations are in excess. Therefore, initial velocity of the reaction (V_0) is measured the when product formation is low and the reverse reaction is negligible. As product formation follows the dissolution of the ES then V_0 is defined as

$$V_0 = k_2[ES]$$

The total amount of enzyme [E_t] is also assumed to be constant but as the reaction starts some free enzyme [E] binds [S] to form [ES] so that [E_t] is equal to [E] +[ES], and in terms of [E] this would be [E_t] -[ES]. Considering the steady state assumption the rate of formation and dissolution of [ES] complex can be considered to be in equilibrium and expressed as

$$k_1([Et] - [ES])[S] = (k_{-1} + k_2)\,[ES]$$

Rearranging equation 1.4 defines the Michaelis constant, K_m as $(k_{-1}+k_2)/k_1$ which is the substrate concentration when the reaction is at half its maximum and is not dependent on the concentrations of substrate and enzyme and can be substituted into the equation to give

$$[ES] = \frac{[Et]\,[S]}{K_M + [S]}$$

As product formation only occurs through the breakdown of ES complex, initial velocity is $V_0 = k_2[ES]$ and can, therefore, be substituted into the expression to yield

$$V_0 = \frac{k_2\,[Et] + [S]}{K_M + [S]}$$

A final consideration is that the amount of enzyme is a limiting factor of the reaction as the maximum velocity will only be reached when the entire free enzyme is at ES. Therefore, Vmax is described as Vmax = k_2[Et] and substitution of this equation gives the final Michaelis-Menten equation. This equation shows the relationship between initial velocity and maximum velocity as a function of the Michaelis constant and the concentration of substrate.

$$V_0 = \frac{V_{max} + [S]}{K_M + [S]}$$

The turnover number of the enzyme, also known as k_{cat}, is the value of the rate limiting step ($k_{cat} = V_{max}/[E_t]$) and gives the concentration of substrate converted to product by enzyme per unit time. This can be used along with K_M to assess the substrate specificity and catalytic efficiency of enzyme by the ratio k_{cat}/K_M.

1.17 Active site investigations through mutagenesis

In order to investigate the significance of the active site residues and their role in the overall enzymatic function, it is necessary to conduct site directed mutagenesis experiments to reveal this information. In the case of serine proteases, several studies have been aimed at kinetic analysis of variants with single, double and triple mutations of the catalytic triad (Ser>His>Asp), substituting these residues with alanine. These studies all show measured activity that is far below the activity of the wild type, with specificity constants that are ~10^5 lower. With minor effects to the Michaelis constant (maximum of two fold increase) noted for each variant, suggesting the substrate was still able to bind within the active site. In

addition to this, comparison to the non-enzymatic reaction showed the variants retained some activity in the region of ~10^3. In conclusion, the mutagenesis studies using serine proteases showed that the active site residues were fundamental in the full catalytic activity. However, with the variants still displaying a low level of activity in comparison to the non-enzymatic control, other factors still present, which are beyond the catalytic triad, have a contribution to overall activity. These factors were assumed to be the oxyanion hole, nucleophilic water molecule and other amino acids in the vicinity fulfilling the role of the substituted catalytic residue.[31,120–125] The catalytic activity and the role of proposed catalytic residues Ser26 and His165 in silicatein-α were assessed by individual mutations of these residues to alanine. The relative catalytic specific activity of the variants was reported as a percentage of the specific catalytic activity displayed by the wild type enzyme noted to be 100%. As the Ser26Ala and His165Ala mutant were recorded to have 8% and 10% relative specific activity it was concluded that these residues are essential for full enzymatic functionality and contribute largely towards catalysis of the substrate TEOS.[1]

1.18 Aims and Objectives

The use of biocatalysts in organosilicon chemistry is not widely studied at present. The application of such a catalyst that could operate under ambient conditions would offer a more sustainable method to the production of organosilicon compounds and silicon based materials, in particular silicone, which is almost universally used in many consumer products and industrial processes. Additionally, a route in which the current unwanted organosilicon by-products are recycled would be a huge asset with regards to environmental degradation issues and conform to green chemistry principles. Silicatein, an enzyme found exclusively in the marine sponge catalyses the polycondensation of silica and would be a good candidate for the biocatalysis of silicon-containing organic compounds as they play a pivotal role in fine chemical production, pharmaceuticals and materials science. Recombinant silicatein-α has been previously shown to accept simple compounds as substrates such as silicon ethoxide and TEOS[1] along with more complex silanes and organosiloxanes.[53,782] However, due to the hydrophobic nature of silicatein-α, isolation has been noted to be very challenging producing low yields and a very lengthy and time-comsuming purification processes. Therefore, one of the main aims of this thesis is to develop and optimise reliable methods for the production of

recombinant silicatein-α in *Escherichia coli* (*E. coli*) using a variety of fusion protein tags which have been utilised to increase the solubility of heterologously produced recombinant silicatein-α in *E. coli*. In addition, co-expression with molecular chaperones and optimisation of buffer components and conditions will be investigated to assess the effects on solubility of silicatein-α. Following this, biochemical assays will be conducted to determine if it is catalytically active against some model silicon containing small molecule substrates for both condensation and hydrolysis reactions. The variants Ser26Ala, His165Ala, Asn185Ala, Asn185Asp and Ser26Cys will then be subjected to the same assay to investigate the role of each of the proposed catalytic residues with Asn185Asp and Ser26Cys variants tested for esterase and protease activity and compared to the activity of commercially sourced serine and cysteine proteases.

As no structural data exists thus far for silicatein-α biophysical techniques such as circular dichroism and light scattering will be applied to aid in expanding structural and conformational knowledge and investigate optimum conditions for stability. These methods will aid in the identification of suitable buffer formulations that maintain a non-aggregated, homogenous sample that can progress to crystallography trials.

As silicatein-α is a promising candidate for use as a biocatalyst in organosilicon chemistry and has been shown to catalyse the silylation of 1-octanol with a range of silyl protecting groups. The work in this thesis aims to expand our current knowledge with regards to substrate scope, the catalytic environment in which it functions and investigate the possibility of a coupled biotransformation with another enzyme. Experiments will be conducted using silicatein-α in an organic solvent system at a temperature of 75 °C and 90 °C with a variety of substrates. Enantioselectivity of the enzyme in this system will be screened with a selection of the aforementioned substrates at the temperatures described. Subsequent research will then be aimed at highlighting a potential enzymatic partner for biotransformations using a coupled enzyme system with ADH as a potential catalytic partner in silyl protection and deprotection in one system.

This will be through racemic resolution with an *R/S*-specific ADH or specific silylation of *S*/*R* enantiomers.

2. Chapter 2: Optimisation of recombinant protein expression and production of silicatein-α in *Escherichia coli*

2.1 Introduction

2.1.1 Recombinant protein production

Recombinant protein production for functional and structural studies requires the protein of interest to be soluble, producible in high yields and easy to purify.[126] Additionally, the protein is required to be biologically active for further analysis and potential application as a biocatalyst or therapeutic protein. There are many well established protein expression hosts that can be utilised for heterologous protein production, for example mammalian and yeast cell expression systems.[127,128] The choice of expression host can depend upon specific requirements that facilitate isolation of a biologically active protein such as co-expression of cofactors or essential post-translational modifications. The enterobacterium *Escherichia coli* (*E. coli*) is often the preferred organism for recombinant protein production due to its cost efficiency, well-established genetic and metabolic behaviour and simple fermentation procedures. One major limitation, however, is the formation of inclusion bodies or insoluble protein aggregates in *E. coli,* thought to accumulate through high expression levels due to strong promoters, high concentration of inducer and increased temperatures during incubation. This leads to an overload of the folding pathways during overexpression of the protein, reducing cytosolic space, thus producing misfolding and exposure of hydrophobic regions, ultimately leading to aggregation and formation of inclusion bodies.[128–130] Processing inclusion bodies to produce active proteins can be a laborious task as they require renaturation. This involves isolation, solubilisation with denaturants, refolding and purification during which a significant amount of protein can be lost.[131] It is, therefore, preferable to avoid the need for renaturation where possible. There are several examples of proteins, previously described as insoluble, that have been isolated by optimising culture conditions. These include varying the incubation temperatures of induced cultures, alternative culture media and different *E. coli* strains with unique features that may overcome solubility problems.[128, 129,132]

In previous reports of silicatein-α, isolation of the protein was acquired either directly from the marine sponge in its native form[133,134] or using a recombinant protein expressed in yeast, insect or bacterial host systems.[135,136] In addition to this, recombinant zymogenic forms of the enzyme were expressed in an effort to increase the yield of soluble protein.[65] However, the propensity for aggregation of recombinant silicatein-α has been well documented and thought to be due to the presence of four hydrophobic regions on the solvent accessible exterior of the protein, identified through homology modelling with cathepsin L.[61,137] It has been proposed that the presence of these hydrophobic patches may aid in the self-assembly of the protein during spiculogenesis.[133] Due to its hydrophobicity and tendency for aggregation soluble heterologous silicatein-α has primarily been acquired by renaturation.[96,138] To address this issue fusion proteins or co-expression of chaperones are often used to increase solubility and take advantage of their inherent features for more selective purification procedures.[139,140]

2.1.2 Fusion protein technology

Fusion protein technology applies the advancement of genetic and protein engineering strategies to link short amino acid sequences or large proteins, known as fusion tags, to the *N*- or *C*-terminus of the target protein. This can improve target protein expression and solubility, assist correct folding and/or facilitate in purification strategies. Fusion tags can be categorised based on their function, as an affinity tag aiding in protein purification, a solubility tag to assist in producing soluble protein and fusion tags that have both affinity and solubility functions. Therefore, the design of the fusion protein construct must be judicious to gain the desired outcome. Affinity tags are employed to enhance purification in chosen downstream chromatographic methods, as they are generally small amino acid or peptide sequences that have a high affinity to a particular ligand or antibody/antigen pair used for the selective binding and purification. This allows separation of the target protein reducing non-specific binding of contaminants, examples include polyhistidine tag and its affinity for metal ions and FLAG tag which binds specifically to anti-FLAG monoclonal antibodies.[141,142]

Fusion tags like trigger factor (TF)[143] and small ubiquitin-related modifier (SUMO), [144] have been utilised to produce soluble proteins in bacterial host systems, such as *E. coli.* TF, a ribosome-associated chaperone endogenous to *E. coli*, functions in an ATP independent

manner to assist in the folding of nascent proteins.[145,142] SUMO has been shown to enhance solubility and has a unique feature being cleavable by the SUMO protease thus generating the target protein with its native *N*-terminus after isolation from the expression host.[144] Many other fusion proteins have been employed as they offer both solubility and affinity such as glutathione S-transferase (GST).[146] Each fusion tag possesses their own unique properties that assist in augmentation of target protein expression, solubility and/or facilitation of more specific purification methods. For example, when used as a fusion protein to enhance solubility, GST, an enzyme with high affinity for reduced glutathione, has the additional benefit of efficient non-denaturing purification through its selective binding with reduced glutathione.[140,146] Maltose Binding Protein (MBP) has previously been shown to promote protein folding also offering advantages as a solubilising agent and affinity tag due to a specificity for amylose.[147,148] Additionally, many fusion proteins are fused to a hexa-histidine tag for purification using immobilised metal affinity chromatography. This enables a combination of methods and techniques to be used sequentially for maximum gain and to increase yields of soluble protein.[140] In order to obtain the correct protein conformation and functionality it is often necessary to remove the fusion tag by enzymatic or chemical cleavage to prevent interference with protein function or the crystallography trials.[43,149,150] However, with different fusion proteins available, selecting the correct one that will produce high yields of recombinant soluble and active protein can often be labour intensive and require many strategies to be tested before the optimal method is formulated.

The use of many different fusion tags for heterologous production of silicatein-α has been documented in the literature. Morse applied the outer membrane protein A (OmpA) and MBP as fusion tags in two separate studies to utilise silicatein as a cell surface biocatalyst and in functional studies, respectively.[72,151] Trigger factor has previously been used to isolate silicatein-α through genetic linkage of the coding sequence for TF to the DNA sequence translating to the proenzyme and pre-proenzyme containing the signal peptide.[65] A further use of fusion tags with silicatein-α was with GST for surface immobilisation and subsequent biocatalysis with an analogue of the enzymes natural substrate.[152]

2.1.3 Co-expression with molecular chaperones

There are several families of molecular chaperones that function to facilitate protein transport or translocation, protein folding of polypeptides, assembly of protein complexes and stabilisation to prevent aggregation of denatured proteins under cellular stress. This native function has been utilised to improve protein expression in a heterologous host such as *E. coli* by co-expression with the recombinant gene of interest. As many chaperone proteins are available, screening is required due to find the complimentary chaperone system that will assist the production of soluble and active protein. Cytoplasmic chaperones from *E. coli* have been widely characterised, in particular the synergistic systems: ATP-dependent GroEL-GroES and DnaJ-DnaK-GrpE chaperone protein complexes and ATP-independent TF. They have been extensively used to promote the correct folding and solubilisation of recombinant target protein through recognition binding of exposed hydrophobic regions of misfolded or unfolded polypeptides. The GroEL-GroES and DnaJ-DnaK-GrpE chaperone protein complexes function in a similar manner to each other, shielding the exposed regions in cage-like structures, utilising energy from the hydrolysis of ATP to 'unfold' misfolded polypeptides and release unfolded intermediates that fold to their native state spontaneously.[153–157] TF consists of three domains, *N*-terminal domain, peptidyl-prolyl *cis-trans* isomerase (PPIase) domain and a *C*-terminal domain. It natively functions as a chaperone by scaffolding the nascent polypeptides emerging from the ribosome in a 1:1 stoichiometry, to protect them from aggregating or misfolding. Interaction with the ribosome is facilitated by the N-terminal domain, with main chaperone function represented by the C-terminal and a catalytically active PPIase domain to aid in *trans* to *cis* proline isomerisation during protein folding.[145,154,158–160] However, the exact mechanism by which the PPIase domain assists in protein folding is still unknown and site directed mutagenesis studies have shown it does not have this exclusive role. As studies show that protein folding can still occur in the absence of this domain.[161,162] Co-expression of molecular chaperones with the recombinant gene can be advantageous to the overall objective of producing a high yield of soluble and active protein, the co-expression of chaperones comes with limitations such as host cell toxicity, production of soluble aggregates and difficulty removing them during purification and isolation of the target protein.[163–165]

2.1.4 Optimum buffer conditions

The medium under which the protein is isolated is critical to maintaining or aiding the solubility. The composition of lysis buffer, therefore, is a key consideration, Leibly *et al*. tested 144 conditions in a solubility screen for proteins that had previously been noted as insoluble. The conditions included varying buffers and salt concentrations and the addition of natural ligands, non-denaturing detergents and chemical compounds. Their results suggest that soluble proteins can aggregate during and after cell lysis and purification, optimising cell lysis and purification buffer conditions, however, can aid in retaining their solubility.[166–170] This contradicts the notion of Williams *et al.* that insoluble proteins are sequestered within inclusion bodies and cannot be recovered.[130] Leibly recorded 33 of 41 proteins had improved solubility during the screening which used buffer additives that would work as a ligand to retain a soluble confirmation, reduce inter-molecular bonds or enhance favourable intramolecular interactions. In addition to this, additives that are known to affect protein stability and those that alter buffer or salt conditions were used.[166] Other groups have looked at amino acid additives, such as L-arginine and L-glutamic acid individually or in combination, to stabilise proteins.[168–171]

Recent work in our group to improve the solubility of silicatein-α, by optimising buffer conditions was successful, albeit resulting in a low yield. This was achieved by using non-denaturing detergents, CHAPS and Triton X-100, typically used to isolate membrane proteins as they interact with exposed hydrophobic regions mimicking the lipid bilayer, to aid in the solubilisation of the hydrophobic silicatein-α.[172–175] The inclusion of non-denaturing detergents and amino acid additives have been documented in the renaturation of silicatein-α but they have not been used as solubility enhancing components during cell lysis and protein purification.[70,176]

2.1.5 Optimum purification methods

As discussed in the introduction, protein studies require a high level of homogeneity and purity. Once a satisfactory expression level is acquired, the researcher needs to consider the optimal purification methods that will maintain the integrity of the protein and reduce loss of protein. This process can be trial and error to find the best method and often involves a

number of techniques to be applied sequentially. Usual methods can include a pre-purification step to reduce major contaminants such as the majority of the endogenous proteins expressed with the host alongside expression of the target protein. Following this a number of purification steps can be used that take advantage of particular features such as affinity chromatography or classical chromatographical methods like ion-exchange chromatography to reduce non-specific binding and further reduce contamination. For experiments requiring very high protein purity a final "polishing" purification step is performed. This is often size exclusion chromatography which separates proteins based on size, shape and molecular weight. Again the steps involved must be chosen judiciously based upon size and physiochemical properties of the protein of interest with assessment of contaminants present at each stage and assessment of protein sample purity, by methods like mass spectroscopy or SDS-PAGE.

2.1.6 Aims and objectives

There are several factors that could detrimentally affect protein expression and isolation resulting in little or no protein, aggregation or truncation of the protein of interest. Therefore, careful planning of potentially suitable methods can equate to a successful outcome of obtaining high yield and purity of the protein of interest. The main aim of this chapter is to optimise the production of recombinant silicatein-α in *E. coli* that will enable further downstream biochemical and biophysical analysis. Once a good yield of protein is obtained the secondary aim is the development of an optimum and reproducible purification method that will enrich purity and homogeneity of the sample.

Herein we describe the screening of a variety of fusion proteins, which have been utilised to increase the solubility of heterologously produced recombinant silicatein-α in *E. coli*. In addition, co-expression with molecular chaperones and optimisation of buffer components and conditions will be investigated to assess the effects on solubility of silicatein-α.

2.2 Results and discussion

2.2.1 Bioinformatic search using orthologs of silicatein-α from *S. domuncula*

As previous research shows, the isolation of silicatein-α from *S. domuncula, P. ficiformis* and *T. aurantia* has proven to be problematic due to several hydrophobic areas on the exterior of the protein. Therefore, a protein sequence alignment was conducted with silicatein-α from four other species of sponge to determine if a less hydrophobic alternative can be isolated and applied for further biochemical and biophysical analysis. The protein sequences for silicatein-α from *Suberites domuncula* (Q2MEV3), *Spongilla lacustris* (D3DEQ1), *Baikalospongia fungiformis* (G3EAH4), *Lubomirskia baicalens*is (Q2PC18), and *Ephydatia fluviatilis* (B5U9F0) were obtained from UniProt[177] and aligned using Clustal Omega[178,179] (Figure 2.1). As the sequence alignment revealed no noticeable differences subsequent structural analysis of each protein was completed by homology modelling and visualisation with SWISS-MODEL[180]and PyMol. However, the data revealed similar hydrophobic areas on the surface of each protein and, therefore, was not successful in identifying an alternative that could have a lower propensity to aggregate. Consequently, all further experiments were conducted using silicatein-α from *S. domuncula*.

```
Q2MEV3      DYPEAVDWRTKGAVTAVKDQGDCGASYAFSAMGALEGANALAKGNAVSLSEQNIIDCSIP
D3DEQ1      SYADSMDWRTKGVVTSVKTQSQCGSSYAFAAVGALEGASALATDKLVALSEQNIIDCSVP
G3EAH4      QYAESIDWRTKGAVTSFQYQGQCGASYAFAATGALEGASALANDKQVTLSEQNIIDCSVP
Q2PC18      QYAESIDWRTKGAVTSVKYQGQCGASYAFAATGALEGASALANDKQVTLSEQNIIDCSVP
B5U9F0      QYAESIDWRTKGAVTSVKYQGQCGASYAFAATGALEGASALANDKQVTLSEQNIIDCSVP
            .* :::******.**:.: *.:**:****:* ******.***..: *:**********:*

Q2MEV3      YGNHGCHGGNMYDAFLYVIANEGVDQDSAYPFVGKQSSCNYNSKYKGTSMSGMVSIKSGS
D3DEQ1      YGNHGCSGGDTYTAFKYVVDNGGIDTESSYPYKGKQSSCQYNSKNAGATATGVVKIASGS
G3EAH4      YGNHGCSGGDTYTAFKYVIDNGGIDTESSYSFKGKQSSCQYNNKTSGASATGVVSIGYGS
Q2PC18      YGNHGCSGGDTYTAFKYVIDNGGIDTESSYSFKGKQSSCQYNNKTSGASATGVVSIGYGS
B5U9F0      YGNHGCSGGDTYTAFKYVIDNGGIDTESSYSFKGKQSSCQYNNKTSGASATGVVSIAYGS
            ****** **: * ** **: * *:* :*:* : ******:**.*  *:: :*:*.*  **

Q2MEV3      ESDLQAAVSNVGPVSVAIDGANSAFRFYYSGVYDSSRCSSSSLNHAMVVTGYGSYNGKKY
D3DEQ1      ESDLMSAVASGGPVAVAVDASVNSFMFYQSGVFDSSTCSNTKLNHAMLVTGYGSVNGKDY
G3EAH4      ESDLLAAVATVGPVAVAVDANTNAFRFYQSGVFDSSSCSSTKLNHAMLVTGYGSYNGKDY
Q2PC18      ESDLLAAVATVGPVAVAVDANTNAFRFYQSGVFDSSSCSSTKLNHAMLVTGYGSYNGKDY
B5U9F0      ENDLLAAVATVGPVAVAIDANTNAFRFYQSGVFDSSSCSSTKLNHAMLVTGYGSYNGKDY
            *.** :**:. ***:**:*.  .:* ** ***:*** **.:.*****:****** ***.*

Q2MEV3      WLAKNSWGTNWGNSGYVMMARNKYNQCGIATDASYPTL
D3DEQ1      WLVKNSWGTSWGESGYIRMVRNKYNQCGIASDALIPML
G3EAH4      WLIKNSWSKNWGDSGYILMVRNKYNQCGIASDALYPML
Q2PC18      WLVKNSWSKNWGDSGYILMVRNKYNQCGIASDALYPML
B5U9F0      WLVKNSWSKNWGDSGYILMVRNKYNQCGIASDALYPML
            ** ****...**:***: *.**********:**  * *
```

Figure 2.1 Sequence alignment of silicatein-α from *Suberites domuncula* (Q2MEV3), *Spongilla lacustris* (D3DEQ1), *Baikalospongia fungiformis* (G3EAH4), Lubomirskia baicalensis (Q2PC18), and *Ephydatia fluviatilis* (B5U9F0) using Clustal Omega. Sequences were obtained from UniProt.

2.2.2 Expression of fusion protein GST-silicatein-α

As discussed in the introduction, fusion protein technology can be employed to improve the solubility and purification of a target protein. Silicatein-α was expressed as a fusion protein with GST in *E. coli*, and was first tested using the pET 28a vector (Figure 2.2A). A range of different culture conditions were applied to investigate and achieve the optimal method of producing soluble and active GST-silicatein-α fusion protein.

Firstly, following induction with 1 mM Isopropyl-β-D-1-thiogalactopyranoside (IPTG), the bacteria were incubated at 37°C for 4 hours, 30°C for 6 hours, 28°C, 25°C or 16°C overnight to enable comparison of the various incubation temperatures. After the appropriate time period, the cells were lysed by sonication and the lysate centrifuged. The supernatant and pellet were analysed by SDS-PAGE (Figure 2.2B). It revealed that the fusion protein was overexpressed but insoluble at all cultivation temperatures, as evidenced by a strong band on the gel of the appropriate molecular weight, but primarily only present in the insoluble fractions (i.e the pellet).

Secondly, a method was used to induce the cold shock response in bacteria and produce cold shock chaperones that may aid in the folding of the recombinant fusion protein and prevent aggregation.[181,182] Prior to induction with 0.5 mM or 1 mM IPTG at an OD_{600} of 0.1 or 0.3 the cultures were put on an ice bath for 30 minutes to induce expression of endogenous chaperones and subsequently incubated at 18°C overnight. SDS-PAGE analysis showed the fusion protein to be overexpressed but insoluble at all the conditions tested (results not shown, but similar to Figure 2.2). As GST is a soluble cytosolic protein, it is believed that it promotes solubility and prevents aggregation of the fusion partner. However, with several research groups noting problems with oxidative aggregation, degradation products and accumulation of inclusion bodies, the only solution would be to continue with further optimisation or conduct renaturation.[139,183–185] As many different culture conditions have been altered already in this study and with the fusion protein being larger and therefore more likely to require complex refolding, expression with other fusion tags was investigated.

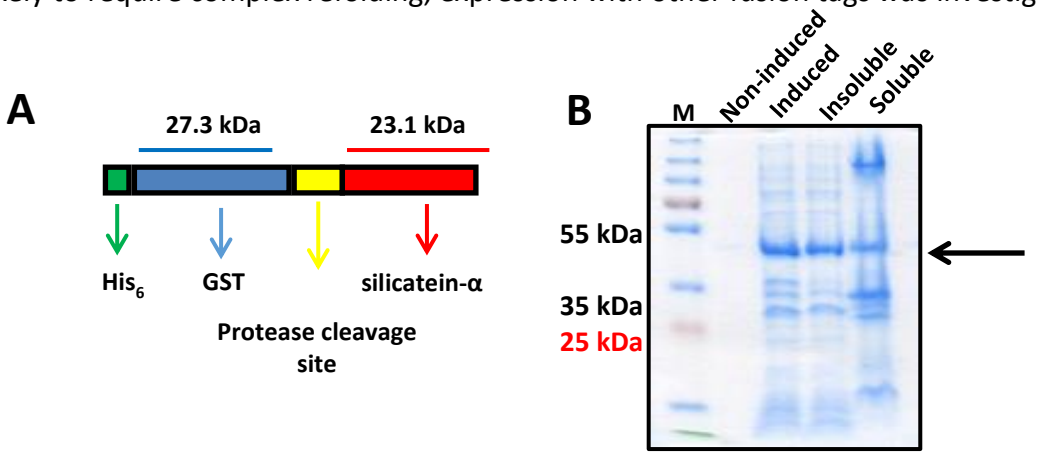


Figure 2.2. Illustrative analysis of fusion protein GST-silicatein-α. (A) Schematic diagram of fusion protein construct (B) SDS PAGE analysis, image of 10 % polyacrylamide gel showing molecular weight protein marker (M), non-induced, induced insoluble and soluble fractions of culture incubated at 16°C overnight.

2.2.3 Expression of fusion protein SUMO-silicatein-α

As expression with the GST fusion tag yielded insoluble protein silicatein-α was expressed in *E. coli* fused to SUMO in the pSUMO vector (Figure 2.3A). Following induction with 1 mM IPTG the bacterial cultures were incubated at 37°C for 4 hours (Figure 2.3B), 30°C for 6 hours, 25°C, 20°C overnight to enable comparison of the effects various incubation temperatures have on the expression and solubility of the fusion protein. Analysis by SDS PAGE indicated low expression of an insoluble fusion protein at all cultivation temperatures. This result was essentially equivalent to the results with the GST fusion tag, therefore, a different fusion tag was investigated.

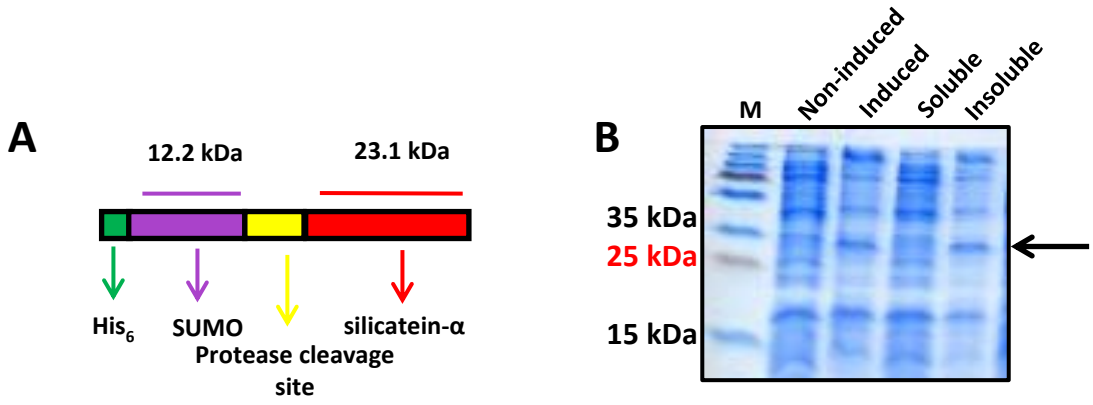


Figure 2.3. Illustrative analysis of fusion protein SUMO-silicatein-α. (A) Schematic diagram of fusion protein construct (B) SDS PAGE analysis, image of 15 % polyacrylamide gel showing molecular weight protein marker (M), non-induced, induced, soluble and insoluble fractions of culture incubated at 37°C for 4 hours.

2.2.4 Expression of fusion protein TF-silicatein-α

In the continuation of our investigations to find a compatible fusion tag that will solubilise our protein, a further fusion tag was selected. Silicatein-α was expressed in *E. coli* fused to trigger Factor in the pCold vector. The recombinant protein produced has a hexahistidine tag at the N-terminus and three protease cleavage sites between the fusion protein and silicatein-α (Figure 2.4A). Following induction with 1 mM IPTG the bacterial cultures grown in Luria Broth (LB) and Terrific Broth (TB) media were incubated at 15°C overnight, as this was the optimal temperature suggested for the pCold-TF vector, allowing for a slower expression rate of the protein and hindering the production of endogenous bacterial proteins. Analysis by SDS PAGE confirmed the overexpression of soluble recombinant TF-silicatein-α fusion protein cultured

in both LB and TB media (Figure 2.4B). The presence of TF-silicatein-α fusion protein in the gel was confirmed by Western blot using antibodies against the polyhistidine tag (results not shown). A range of different cultivation temperatures were then tested to compare the effect of an increase in temperature on the solubility of the fusion protein. SDS PAGE analysis indicated the overexpression of a soluble protein when the cultures were incubated at 37°C for 4 hours, 30°C for 6 hours, 20°C, 25°C overnight (Figure 2.4C). Therefore, growth and incubation of the bacteria with the pCold vector is not restricted to the suggested temperature of 15°C. However, a rise in temperature led to more endogenous protein being produced which may, in turn, affect the amount of fusion protein expressed. Further to this, pCold-TF vector containing our target protein was transformed into *E. coli* Artic Express. After induction with 1 mM IPTG the cultures were incubated at 15°C overnight and analysed by SDS PAGE. This *E. coli* strain was also shown to overexpress soluble fusion protein (Figure 2.4D). As the results show this fusion tag aided in the expression soluble protein, subsequent protein purification was undertaken.

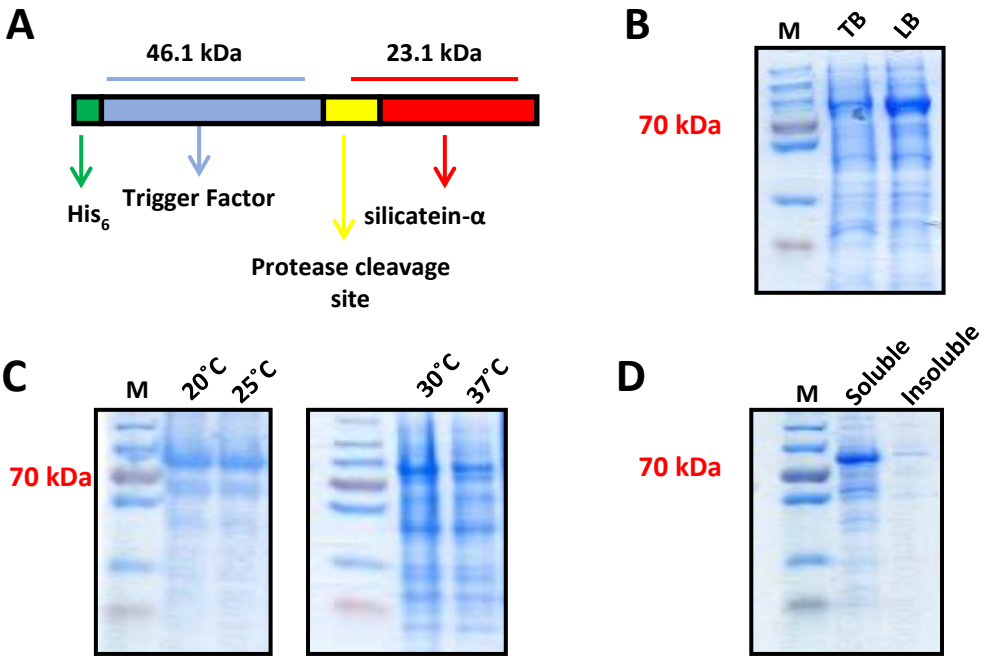


Figure 2.4. Solubility analysis of fusion protein TF-silicatein-α. (A) Schematic diagram of fusion protein construct. (B) SDS-PAGE analysis, image of 10% polyacrylamide gel with molecular weight protein marker (M), soluble recombinant TF-silicatein-α fusion protein cultured in both TB and LB media. (C) SDS-PAGE analysis 10% polyacrylamide gel with molecular weight protein marker (M), soluble recombinant TF-silicatein-α fusion protein incubated at 20°C, 25°C, 30°C and 37°C. (D) SDS-PAGE analysis, 10% polyacrylamide gel with molecular weight protein marker (M), soluble recombinant TF-silicatein-α fusion protein expressed using *E. coli* Arctic Express, soluble and insoluble fractions.

2.2.5 Purification of fusion protein TF-silicatein-α using immobilised metal affinity column (IMAC)

Immobilised metal affinity chromatography was used as the first mode of protein purification due to the presence of hexahistidine tag at the *N*-terminus of the fusion protein. Clarified lysate containing soluble proteins were loaded on to a column with Ni^{2+}-NTA agarose resin to capture the hexahistidine tagged fusion protein. Using a three step gradient with increasing concentration of imidazole, the protein was finally eluted with 250 mM imidazole and analysed by SDS PAGE (Figure 2.5A and Figure 2.5B). As this showed a number of impurities eluted along with the fusion protein, the imidazole concentration in the sample loaded on to the column was raised to reduce non-specific binding of untagged protein. However, even with 30 mM imidazole added to the loading buffer and the initial step of the wash stages it still showed possible binding of contaminating proteins (Figure 2.5C), Therefore, further optimisation to determine the optimum concentration range of imidazole that would maintain purity of the histidine tagged fusion protein and reduce non-specific binding of contaminating proteins, or a preliminary purification step that would remove im purities is required.

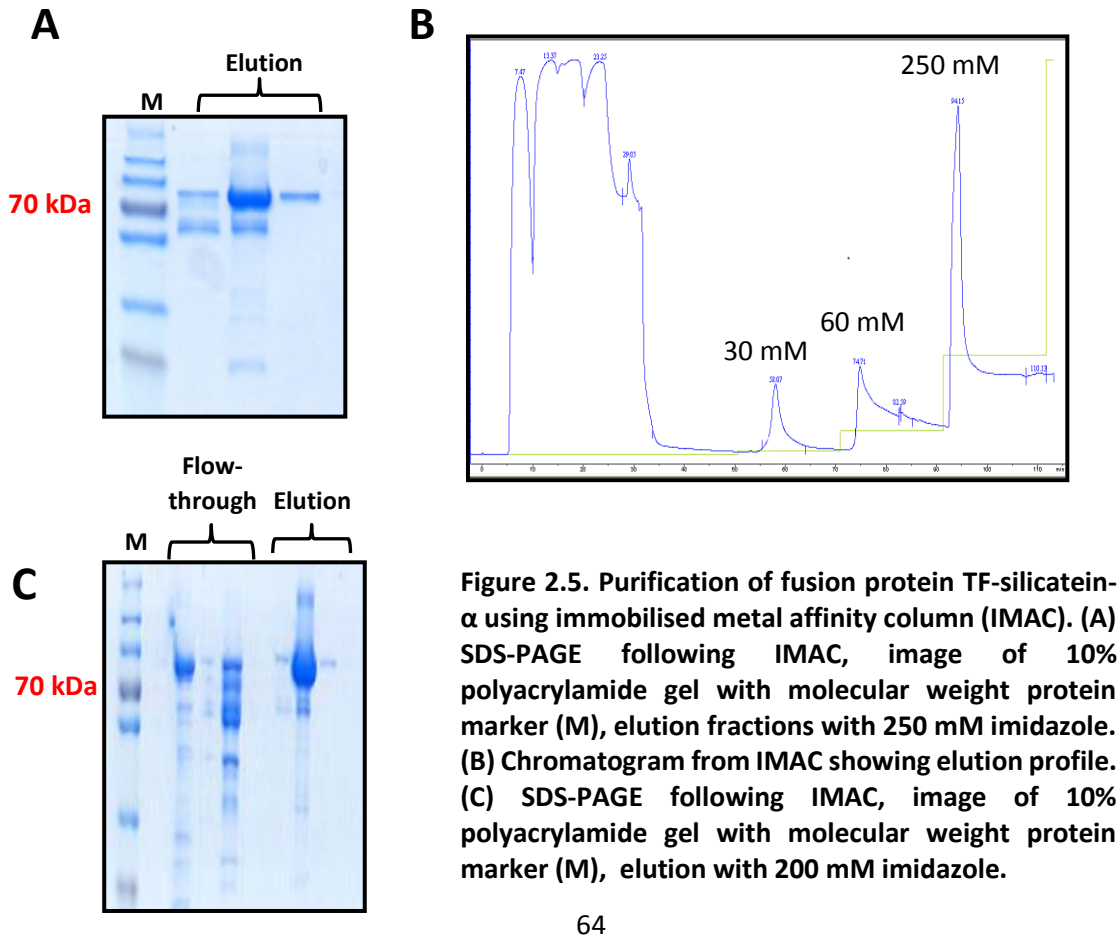


Figure 2.5. Purification of fusion protein TF-silicatein-α using immobilised metal affinity column (IMAC). (A) SDS-PAGE following IMAC, image of 10% polyacrylamide gel with molecular weight protein marker (M), elution fractions with 250 mM imidazole. (B) Chromatogram from IMAC showing elution profile. (C) SDS-PAGE following IMAC, image of 10% polyacrylamide gel with molecular weight protein marker (M), elution with 200 mM imidazole.

2.2.6 Purification of fusion protein TF-silicatein-α using Ammonium sulphate precipitation

In order to improve the quality of purification during IMAC, ammonium sulphate precipitation[186,187] was used as a preliminary purification step to attempt the depletion of unwanted proteins. To establish the ammonium sulphate concentration that effectively precipitates the maximum amount of fusion protein, parallel precipitation experiments were conducted. Varying concentrations of ammonium sulphate were added to samples of clarified lysate to achieve percentage saturation of ammonium sulphate solutions between 10-100% at 4°C and 25°C. Following centrifugation, the precipitate was resuspended in PBS and total protein concentration was obtained from the resuspended precipitate and the supernatant. The experiments were completed in triplicate and the total protein recovered was calculated as a percentage of the total protein in a sample of clarified lysate (Figure 2.6A). Each sample was also analysed by SDS-PAGE (Figure 2.6B and Figure 2.6C).

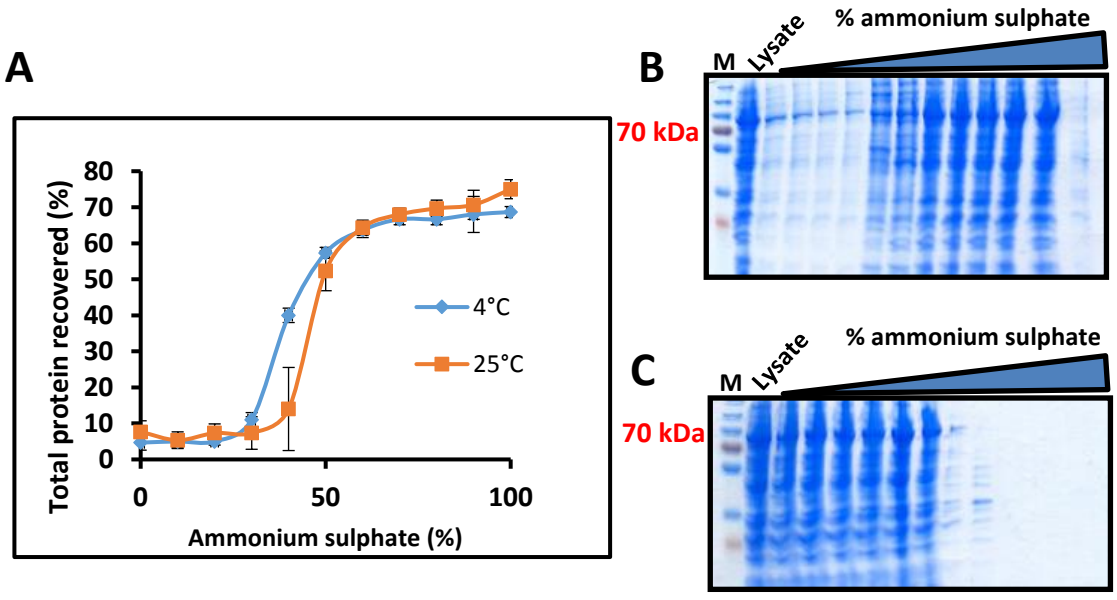


Figure 2.6. Purification of fusion protein TF-silicatein-α using Ammonium sulphate precipitation. (A) Percentage of total protein recovered by precipitation against amount of ammonium sulphate added to clarified lysate at 4°C and 25°C. (B) SDS PAGE analysis, 10% polyacrylamide gel of parallel ammonium sulphate precipitation experiments at 25°C showing samples from the resuspended precipitate, molecular weight protein marker (M), clarified lysate, 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100% and (C) samples from the supernatant, molecular weight protein marker (M), clarified lysate, 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100%.

SDS PAGE analysis of the supernatant from the parallel ammonium sulphate experiments show that a high number of contaminating proteins remained in solution, without a significant loss of the fusion protein, at concentrations of 60 and 70% saturation. Therefore, a large scale experiment was conducted using saturated ammonium sulphate solutions of 60%, 65% and 70% to obtain the optimum concentration and samples were analysed by SDS PAGE (Figure 2.7A) with total protein concentration measured and calculated (results not shown). The samples with 60% (Figure 2.7B) saturation appeared to remove the most contaminating proteins in solution without great loss of the fusion protein. Therefore, this concentration will be used as a pre-purification step followed by IMAC.

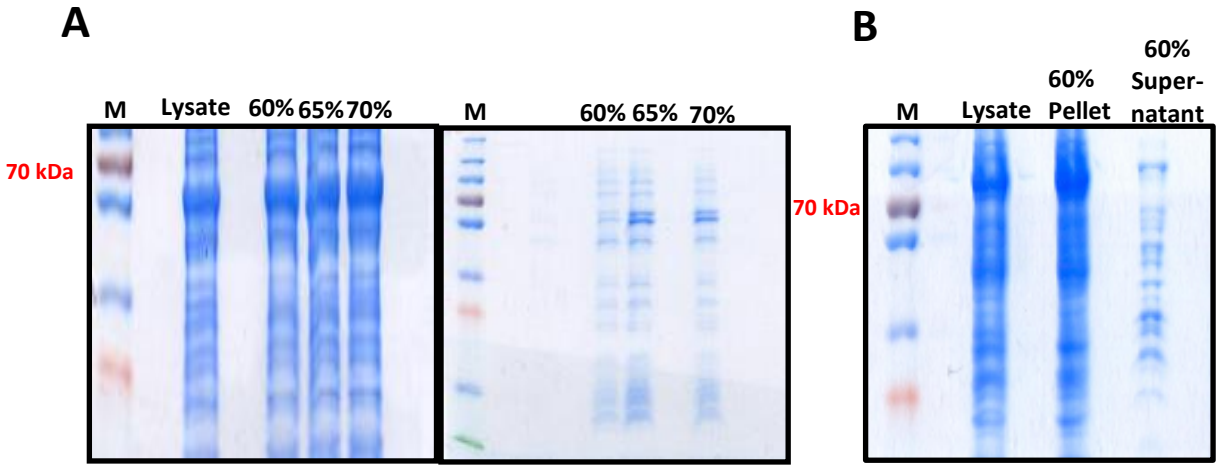


Figure 2.7. Purification of fusion protein TF-silicatein-α using ammonium sulphate precipitation. (A) SDS PAGE analysis of ammonium sulphate cut with 60, 65 and 70% saturation, image of 10% polyacrylamide gel showing molecular weight protein marker (M), clarified lysate, resuspended pellet with 60%, 65% and 70% saturation, samples from supernatant with 60%, 65% and 70% saturation; (B) SDS PAGE analysis of ammonium sulphate with 60% saturation, image of 10% polyacrylamide gel with molecular weight protein marker (M), clarified lysate, resuspended pellet with 60% saturation and supernatant with 60% saturation.

2.2.7 Purification of fusion protein TF-silicatein-α using ammonium sulphate precipitation and immobilised metal affinity column (IMAC)

The optimised ammonium sulphate precipitation was then combined with IMAC to attempt the isolation of TF-silicatein-α. Clarified lysate containing soluble proteins was brought to a final concentration of 60% ammonium sulphate saturation and subsequently purified by immobilised metal affinity chromatography. Using a three step gradient with increasing concentration of imidazole, proteins were finally eluted with 250 mM imidazole and analysed by SDS PAGE (Figure 2.8). This result clearly showed that the added purification step reduced contaminants significantly.

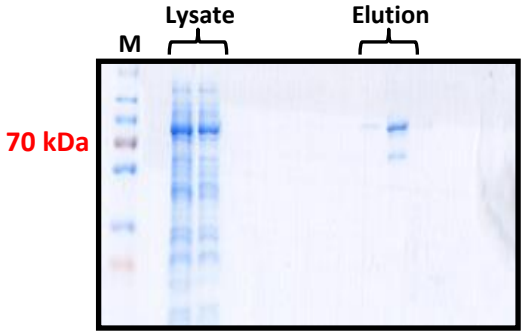


Figure 2.8. Purification of fusion protein TF-silicatein-α using ammonium sulphate precipitation and immobilised metal affinity column (IMAC). SDS-PAGE following IMAC, image of 10% polyacrylamide gel with molecular weight protein marker (M) and elution with 250 mM imidazole.

2.2.8 Cleavage of fusion protein TF-silicatein-α with Factor Xa at 25°C and 4°C

Following purification of the fusion protein, removal of the fusion tag by proteolytic cleavage with Factor Xa was investigated to enable separation of TF and unmodified silicatein-α. Factor Xa protease was used according to the small scale optimisation in the manufacturer's protocol. The samples were then analysed by SDS-PAGE, which showed there was more efficient cleavage with a higher concentration of Factor Xa protease due to a decrease in the amount of fusion protein and an increase in the amount of Trigger factor (Figure 2.9A). However, due to the reducing and denaturing condition of SDS-PAGE, silicatein-α and Factor Xa could not be confidently identified on the gel as Factor Xa is reduced and denatured to its two constituent polypeptide chains, with molecular weights of 30 kDa and 20 kDa, which are close to the molecular weight of silicatein, 23.1 kDa. Therefore, native PAGE was conducted to analyse the proteins in their native state. However, the results suggested incomplete cleavage of the fusion protein in both reactions carried out at 4°C and 25°C at various time points (Figure 2.9). At 4°C there appears to be no change in the amount of fusion protein and trigger factor in comparison to the uncleaved control and also no apparent band is observed which could be silicatein-α. In contrast, at 25°C the amount of fusion protein appears to be reduced in comparison with the uncleaved control sample, although, in a similar manner there is no apparent band for silicatein. However, there are bands above and below the band thought to be Factor Xa which are not present in the reaction at 4°C (Figure 2.9B). As the results were inconclusive in assessing Factor Xa as an effective protease for cleavage of the fusion tag, further investigations were aimed at evaluating the another protease sites.

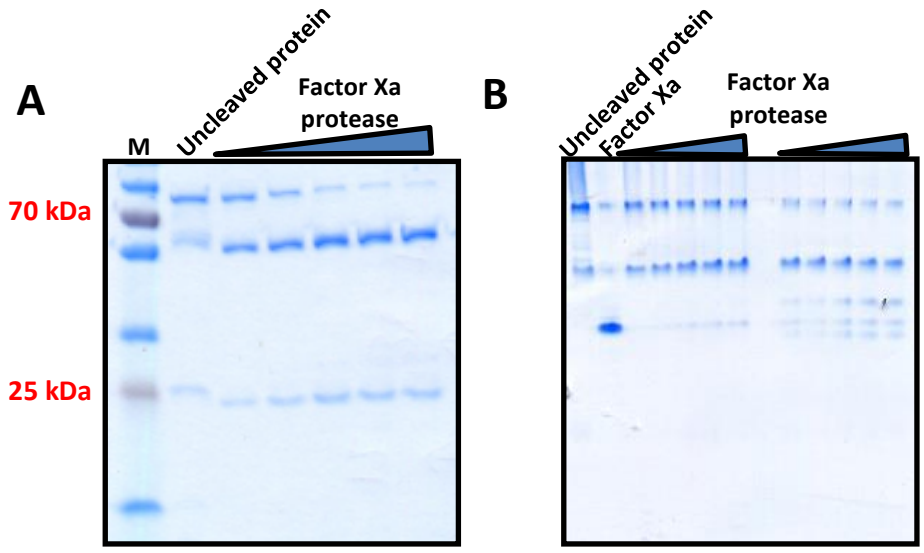


Figure 2.9. Cleavage of fusion protein TF-silicatein-α with Factor Xa at 25°C and 4°C. (A) SDS-PAGE analysis of Factor Xa assay at 4°C, image of 10% polyacrylamide gel with molecular weight protein marker (M), uncleaved fusion protein, 0.5 µl, 1 µl, 2 µl, 3 µl and 3.5 µl of Factor Xa protease. (B) native PAGE Factor Xa assay at 4°C, uncleaved fusion protein, factor Xa, 0.5 µl, 1 µl, 2 µl, 3 µl and 3.5 µl of Factor Xa protease, Factor Xa assay at 25°C, 0.5 µl, 1 µl, 2 µl, 3 µl and 3.5 µl of Factor Xa protease.

2.2.9 Cleavage of fusion protein TF-silicatein-α with HRV3C and GST-HRV3C at 4°C

As proteolytic removal of the fusion tag with Factor Xa was inconclusive further investigations were carried out using the protease HRV3C. A small scale optimization experiment to estimate the appropriate ratio of enzyme: target protein was conducted with ratios of 1:5, 1:25 and 1:50 at increasing incubation times of 1, 3 and 16 hours as per the manufacturer's protocol. The results showed that, although HRV3C was very efficient at cleaving the fusion protein, as evidenced from the loss of the original fusion protein band, (by SDS-PAGE) the results were considered to be inconclusive, as the protease of 22 kDa is close in molecular weight to our target protein, silicatein, 23.1 kDa (results not shown). Consequently, this small size difference would also make subsequent separation and purification by IMAC or size exclusion chromatography (SEC) difficult. Cleavage with a GST tagged HRV3C protease would provide a clearer identification by SDS-PAGE, due to a greater difference in molecular weight between the target protein and protease, with successive purification and isolation possibly less troublesome with the additional GST tag present on the protease. GST tagged HRV3C was

been expressed and purified but also proved to be unsuccessful, with incomplete cleavage of the fusion protein was maybe attributable to poor proteolytic activity (results not shown)

2.2.10 Solubility analysis of co-expression of GST-silicatein-α and SUMO-silicatein-α with chaperone vector

As an alternative to using fusion tags, the vector containing either GST-silicatein-α or SUMO-silicatein-α was co-expressed with each of the plasmid expressing various chaperones (Table 2.1) in *E. coli* to investigate their effect on solubility. Successful transformation of each plasmid was screened using the relevant antibiotic combination in accordance to the resistance gene they contained. Following induction with IPTG, the bacteria were incubated at 30°C for 6 hours and 25°C overnight to enable comparison of the effects various incubation temperatures have on the expression and solubility of the fusion protein. However, upon analysing each culture by SDS PAGE there was either no or minimal expression in the soluble fraction for GST-silicatein-α and SUMO-silicatein-α co-expressed with pGTf2 vector or SUMO-silicatein-α co-expressed with pGro7 vector. A slight expression of a protein at the expected molecular weight was observed for GST-Silicatein-α co-expressed with PKJE7 and pGro7 and SUMO-silicatein-α PKJE7. However, when it was further analysed by Western blot it confirmed no soluble fusion protein was present (results not shown).

Table 2.1 Characterisation of chaperone vectors co-expressed with GST-silicatein-α and SUMO-silicatein-α

		Protein solubility	
Plasmid	Chaperone	SUMO-silicatein-α	GST-silicatein-α
pGro7	groES-groEL	-	+
pKJE7	dnaK-dnaJ-grpE	+	+
pGTf2	groES-groEL-Tf	-	-

2.2.11 Refolding of silicatein-α

As Muller *et al*. [69] described a technique by which insoluble fraction of silicatein-α can be solubilised and subsequently refolded, the method was followed in an attempt to produce

protein that may be used for crystallisation studies. The rationale behind this being that structural and functional investigation of the refolded protein could be used in comparison studies with information gained from the same analysis using the fusion protein TF-silicatein-α, and natively folded silicatein-α. A hexahistidine tagged silicatein-α, was expressed in *E. coli* with SDS-PAGE analysis showing the majority of the protein in the insoluble fraction (Figure 2.10). The pellet was solubilised with buffer containing 6 M urea and applied to IMAC to isolate silicatein-α (Figure 2.10). Following buffer exchange, refolding by dialysis was attempted using buffer described by Müller *et al* (50 mM Tris/HCl, pH 8.5, 0.5 M L-arginine, glutathione (9 mM glutathione:1mM oxidized glutathione) , 0.3 M NaCl, 1 mM KCl).[69] However, protein precipitation was observed under both refolding conditions so the supernatant and precipitated material was further analysed by SDS PAGE. The results from SDS-PAGE analysis show that refolding conditions did not produce soluble protein (Figure 2.10) and the approach described in the literature is not reproducible. Renaturation of the protein was then attempted using a method known as slow drop-wise dilution.[188,189] This allows a gradual reduction of the solubilisation agents for a slower refolding of the protein and small increases in redox system reagents to enable disulphide bond formation. Refolding buffer C1 was added slowly in a dropwise manner, yet this method was not successful in producing soluble protein as aggregation was observed.

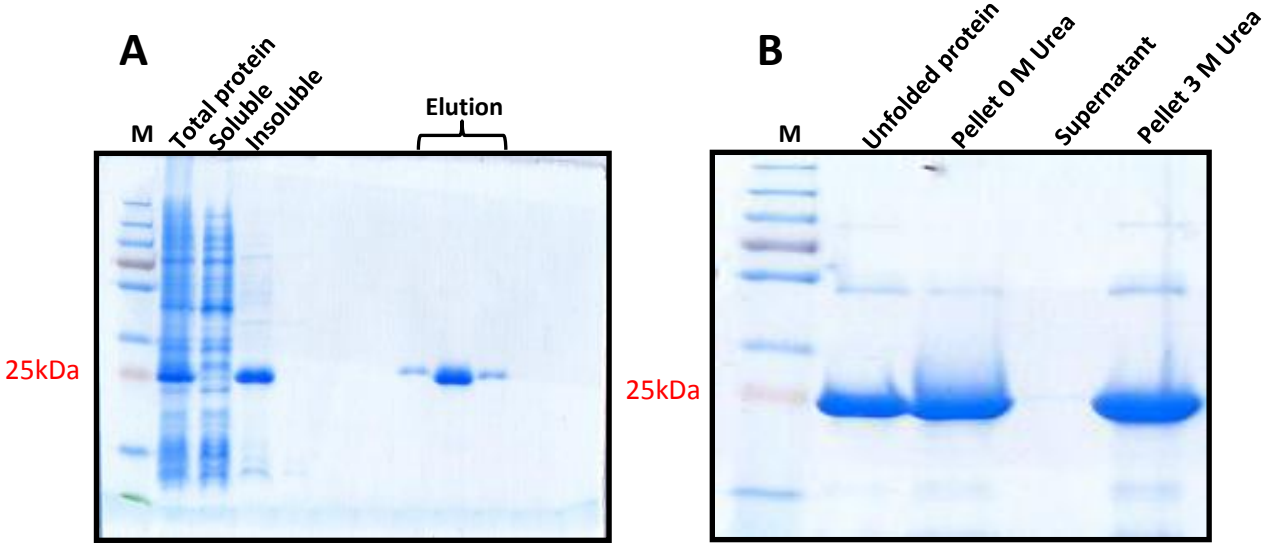


Figure 2.10. Refolding of silicatein-α. (A) SDS-PAGE analysis, image of 10% polyacrylamide gel with molecular weight protein marker (M), total protein, soluble fraction, insoluble fraction, elution with 250 mM during IMAC. (B) SDS-PAGE analysis, image of 10% polyacrylamide gel with molecular weight protein marker (M), unfolded pHis silicatein-α purified by IMAC, pellet following 0 M urea, supernatant following 0 M urea (2) and pellet following 3 M urea.

2.2.12 Buffer optimisation for the purification of TF-silicatein-α and silicatein-α

Previous work using lysis buffers containing non-denaturing detergents such as Triton-X 100 and CHAPS enabled a small amount (<1 mg/mL) of soluble silicatein-α to be purified (Figure 2.11A) and applied to biochemical activity assays confirming a catalytically active protein. However, purification of this protein requires several buffer exchange steps to a more compatible buffer for IMAC. Therefore, a buffer containing 500 mM L-arginine and 500 mM L-glutamic acid (50 mM Tris, 100 mM NaCl) was used as a lysis buffer. These additives have been used as a solubilising agent to prevent aggregation, although the exact mechanism of action is unclear it has been suggested that they function by reducing protein-protein interactions.[168–170,190] The results show that the addition of amino acid additives to the lysis buffer enabled soluble silicatein-α to be produced albeit to a low concentration similar to that of the buffer containing non-denaturing detergents (Figure 2.11B). However, the advantage of this new buffer is that it reduces several buffer exchange steps before and after IMAC purification, thereby streamlining the purification process.

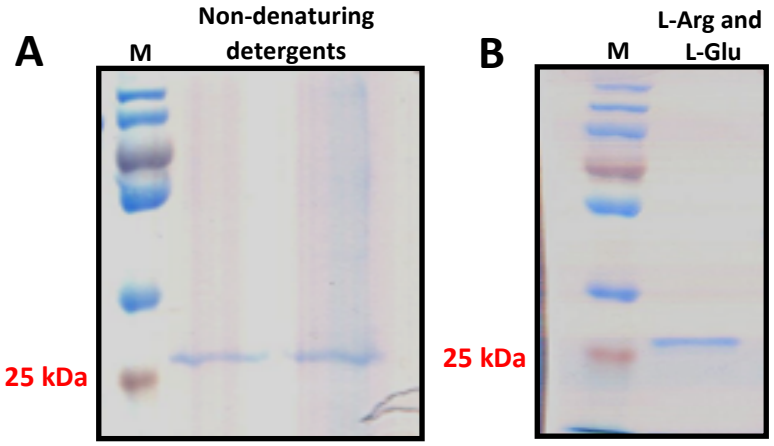


Figure 2.11. Buffer optimisation for silicatein-α. (A) SDS –PAGE analysis image of 15 % polyacrylamide gel with molecular weight protein marker (M) and protein sample in buffer containing non-denaturing detergents. (B) SDS –PAGE analysis image of 15 % polyacrylamide gel with molecular weight protein marker (M) and protein sample in buffer containing L-arginine and L-glutamate.

This buffer was also applied to the purification and isolation of TF-silicatein-α. During the final purification step using size exclusion chromatography (SEC) with the standard buffer (50 mM

Tris 100 mM NaCl, pH 8.5) one single peak is observed (Figure 2.12A). However, upon the addition of the amino acids (L-arginine and L-glutamate) to the buffer a second peak in the chromatogram is observed (Figure 2.12B). As SEC separates proteins on the basis of size and shape, calibration using globular protein standards of known molecular is not always reliable. This was because it is an uncharacterised fusion protein and we cannot assume this protein has a globular shape, therefore other methods will be utilised to determine the molecular weight and shape of the protein. Nevertheless, SEC shows the protein elutes after the void volume and analysis by SDS-PAGE shows a single band observed at the correct molecular weight. The fusion proteins may be forming a highly ordered complex in the buffer without amino acids additives that is reduced to a lower order in the presence of these extra components. A higher concentration of amino acids additives may be required to produce one single species of protein at this conformation, therefore, further optimisation would be needed. These aspects are further investigated subsequently (see Chapter 4).

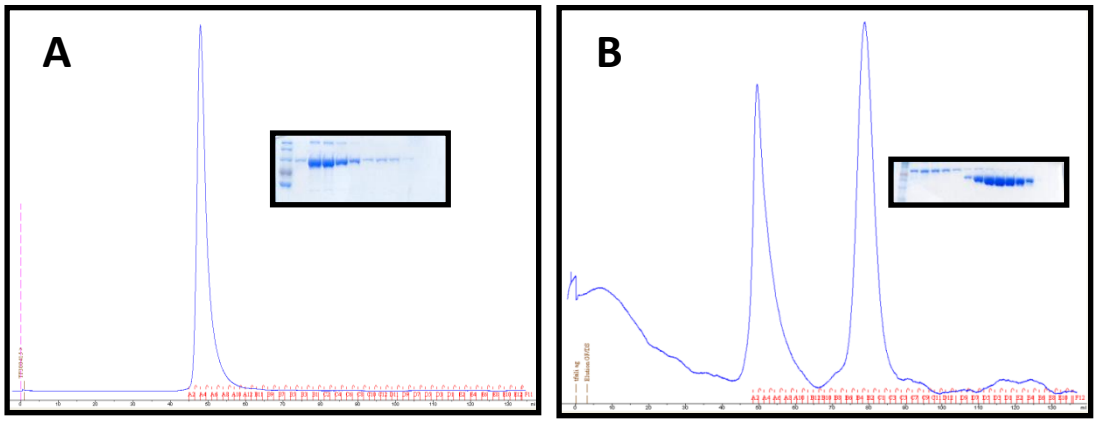


Figure 2.12 SEC chromatograms and SDS-PAGE analysis of TF-silicatein-α. (A) TF-silicatein-α in 50 mM Tris, 100 mM NaCl. (B) TF-silicatein-α in 50 mM Tris, 100 mM NaCl, 50 mM L-arginine, 50 mM L–glutamic acid

Silicatein-α and TF-silicatein-α were purified with SEC (Figure 2.A) in 50 mM Tris, 100 mM NaCl pH 8.5 and subsequent Western Blot analysis was conducted for each protein to confirm the presence of the hexahistidine tag. (Figure 2.B and Figure 2.C) Native PAGE analysis confirmed non aggregation of silicatein-α, as evidenced by a well-defined single band (Figure 2.D).

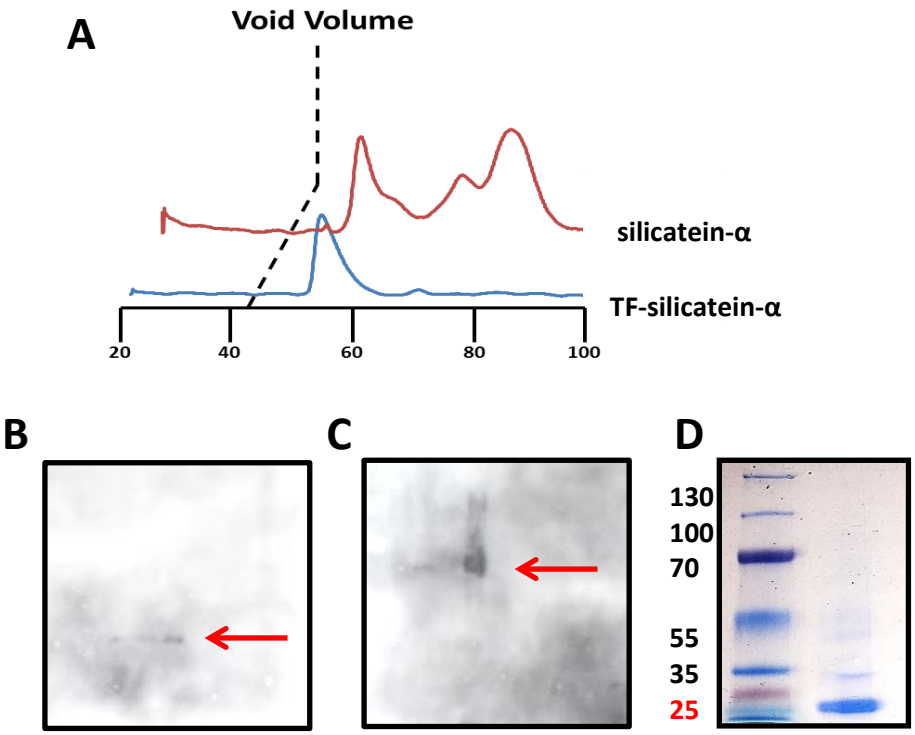


Figure 2.13 (A) SEC analysis of silicatein-α and TF-silicatein-α. (B) Western Blot analysis of silicatein-α. (C) Western Blot analysis of TF-silicatein α. (D) Native PAGE analysis of silicatein-α

2.2.13 Expression, purification and protease cleavage of TF-silicatein-α-Strep

As previously discussed, the expression and purification of TF-silicatein-α produces a high yield of protein but efforts to separate silicatein-α from the fusion protein have thus far proven problematic. In order to obtain a high yield of silicatein-α, a further affinity tag, Strep tag (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys), was added onto the C-terminus of the TF-silicatein-α fusion protein with a thrombin cleavage protease site. (Figure 2.14) This would enable an additional purification method, as the tag has affinity to immobilised streptavidin (Strep-Tactin) and facilitates in the capture of silicatein-α through one-step affinity chromatography[191] and subsequent isolation from proteolytic cleavage using thrombin. Thereby a method using a combination of IMAC and affinity chromatography plus HRV3C and Thrombin proteases was envisaged.

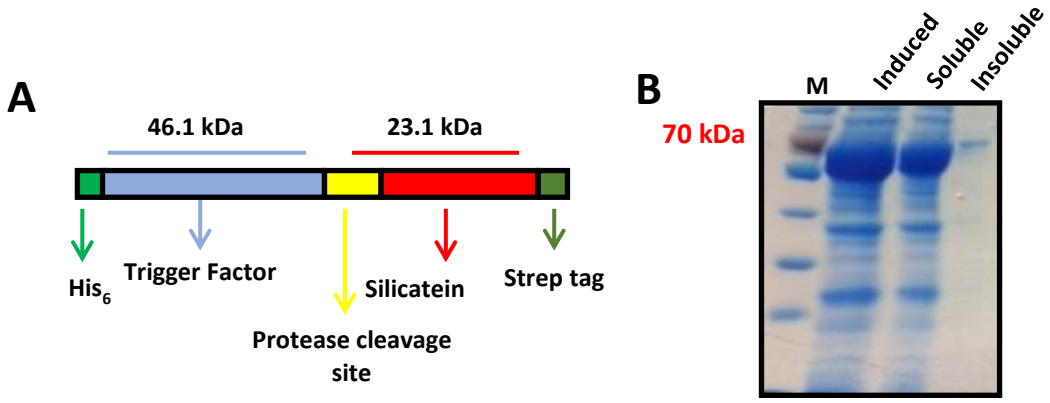


Figure 2.14 Illustrative analysis of fusion protein TF-silicatein-α-Strep. (A) Schematic diagram of fusion protein construct (B) SDS PAGE analysis, image of 10% polyacrylamide gel showing protein marker (M), induced, soluble and insoluble fractions incubated at 15˚C

Initial experiments were aimed at confirming the presence, accessibility and functionality of the hexahistidine tag in binding to Ni^{2+}-NTA column and Strep tag in binding to the Streptavidin column (Figure 2.15). The results of these initial tests also revealed that fusion protein of a higher yield and purity was obtained through purification using IMAC and subsequent strep-tag affinity chromatography. This order of purification methods removed co-eluting contaminants from IMAC and prevented loss of protein due to non-specific binding of proteins to the strep column.

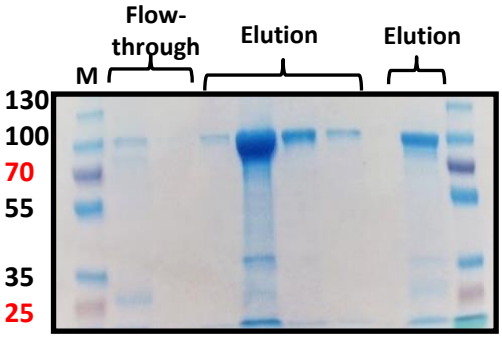


Figure 2.15. Purification of TF-silicatein-α-Strep. SDS PAGE analysis, image of 10 % polyacrylamide gel showing molecualr weight protein marker (M), IMAC elution fractions containing TF-Silicatein-α-Strep and elution fraction from strep-tag affinity chromatography.

The isolated fusion protein was then incubated with HRV3C and thrombin in separate assays to assess protease cleavage (Figure 2.16A and Figure 2.16B respectively). The results show that each protease has incomplete cleavage even at such low ratios of protease to target protein. This preliminary result suggests a higher concentration of protease may be required for complete cleavage of the fusion protein.

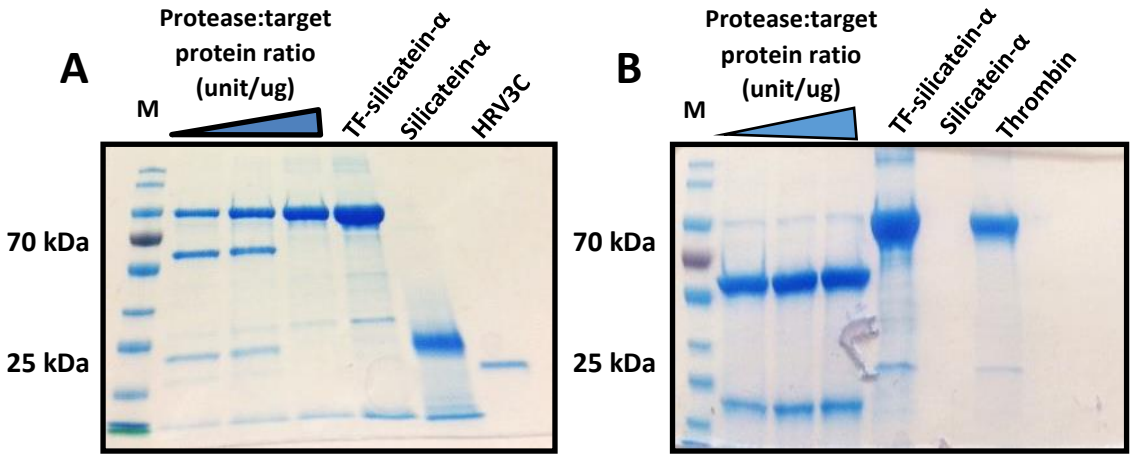


Figure 2.16. Protease cleavage of TF-silicatein-α-Strep. (A) SDS PAGE analysis Protease Cleavage of TF-silicatein-α-Strep with HRV3C. Molecular weight protein marker (M), 1:5 Protease:target protein ratio (unit/µg), 1:25 Protease:target protein ratio (unit/µg), 1:50 Protease:target protein ratio (unit/µg), Uncleaved fusion protein, silicatein-α, HRV3C. (B) SDS PAGE analysis Protease Cleavage of TF-silicatein-α-Strep with Thrombin. Molecular weight protein marker (M), 1:5 Protease:target protein ratio (unit/µg), 1:25 Protease:target protein ratio (unit/µg), 1:50 Protease:target protein ratio (unit/µg), Uncleaved fusion protein, silicatein-α and thrombin .

Alternatively, incomplete cleavage may be due to the protease cleavage sites being inaccessible particularly for HRV3C, especially considering the addition of the strep-tag which may affect access to the cleavage sequence. Optimisation to reduce the ratio of target protein to protease or increase the ratio of protease would be futile as this would not be cost effective upon scale up of the experiment. Thrombin appears to be more efficient even at a higher ratio of target protein to protease. Therefore, this could be a good protease to use for one stage of the separation. There are two thrombin cleavage sites, one between trigger factor and silicatein and one between silicatein and the strep tag. TF-Silicatein-α-Strep was incubated with thrombin for 24 hours after which it was eluted firstly through the Strep-Tactin affinity column, to capture the strep tag. The flow-through was subsequently applied to the IMAC column, to capture hexahistidine tagged trigger factor and enable isolation of silicatein-α. All fractions collected were analysed by SDS PAGE. The elution fractions from the Strep-Tactin affinity column, show a strong intensity band of around 70 kDa which may by indicative of incomplete proteolytic cleavage of the fusion protein by thrombin (Figure 2.17A Lanes 4-7). However, SDS PAGE analysis of samples following thrombin cleavage appeared to show no band that would correspond to the uncleaved fusion protein (results not shown).

The flow-through fractions collected from the Strep-Tactin affinity column were subsequently applied to the IMAC column and the flow-through and elution fractions were analysed by SDS PAGE. However, upon analysis of these fractions no bands were observed in the flow-through fractions and the elution fractions only presented a band at around 70 kDa. A band of this size corresponds to the fusion protein and could suggest incomplete cleavage by thrombin or incomplete separation of the fusion protein. In addition to the absence of ~23 kDa bands in the SDS PAGE analysis of the fractions from the flow-through which may be due to precipitation of silicatein-α, an observation noted in previous proteolytic cleavage of TF-silicateinα using thrombin protease.[65] Further optimisation is required for the proteolytic cleavage and isolation of silicatein-α, which could be different buffer conditions or components during the protease assay or subsequent purification steps. Alternatively, other protease cleavage sites could be cloned in place of the thrombin site between silicatein-α and strep tag as this would enable the use of another protease which could ultimately enhance efficiency and produce a more cost effective method for isolation of silicatein-α.

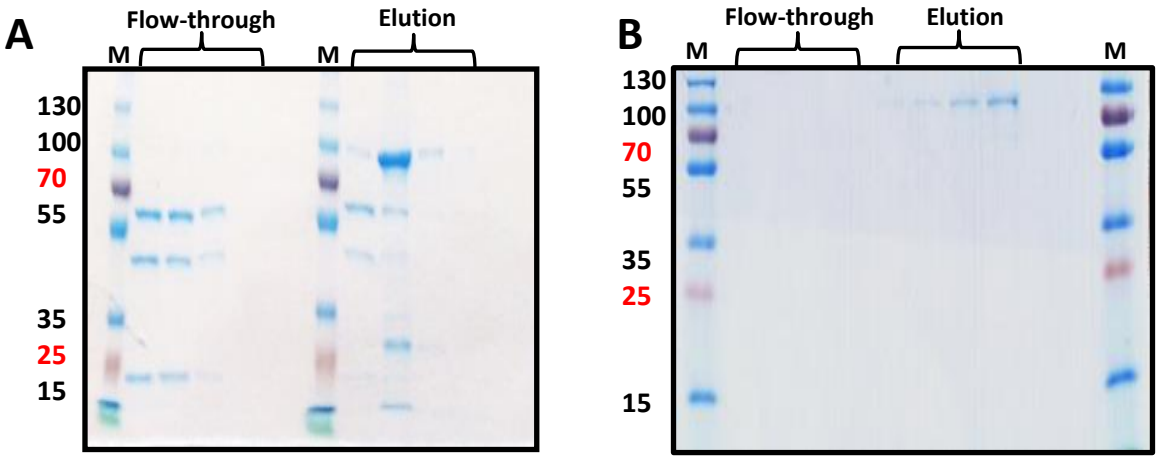


Figure 2.17. (A) SDS-PAGE analysis of Strep-Tactin affinity column following thrombin cleavage of TF-silicatein-Strep with molecular weight protein marker (M), flow-through fractions and elution fractions. (B) SDS-PAGE analysis of IMAC column following injection of flow-through fractions from Strep-Tactin column with molecular weight protein marker (M), flow-through and elution fractions.

2.2.14 Optimisation of protein purification with IMAC and SEC with TF-silicatein-α

As previously mentioned Immobilised metal affinity chromatography was utilised as a purification technique due to the presence of a hexahistidine tag with subsequent size

exclusion chromatography to isolate TF-silicatein-α. However, due to different buffers being used for lysis, IMAC purification and SEC, buffer exchange steps are usually required in between, which can be time consuming if dialysis is applied or risk protein loss by ultrafiltration. Therefore to speed up the process, reduce buffer exchange steps and circumvent the use of ultrafiltration, 50 mM Tris buffer 100 mM NaCl pH 8.5 was used for lysis, IMAC and SEC with TF-silicatein-α. The main concern with the optimisation and streamlining of the process was compatibility of the Tris buffer with IMAC, however, analysis by SDS-PAGE following IMAC showed the protein eluted to a good yield and purity. As soluble silicatein-α was also obtained through the use of Tris in the lysis buffer (Figure 2.11B) this method was also applied successfully. Therefore, the purification method was optimised and streamlined by reducing the number of buffer exchange steps thus reducing overall time and increasing efficiency without a detrimental effect on yield and purity (Figure 2.18).

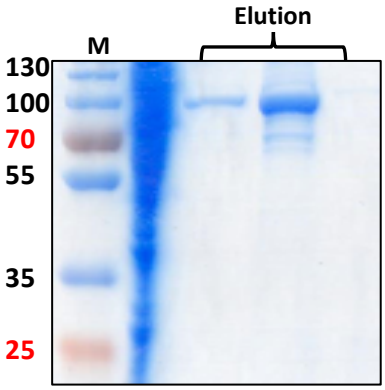


Figure 2.18 Optimisation of the overall purification process of fusion protein TF-silicatein-α using a continuous buffer system. SDS-PAGE following IMAC, image of 10% polyacrylamide gel with molecular weight protein marker (M) and elution with 250 mM imidazole.

2.3 Conclusion

The aim of this work was to develop an optimal method for the expression and purification of silicatein-α, which will produce a concentration of soluble protein to a high purity required for further downstream assays and structural analysis. Several of the fusion proteins investigated were unsuccessful in the production of soluble protein, in spite of efforts to optimise culture and expression conditions, however, the addition of trigger factor significantly enhanced the soluble yield. In the case of silicatein-α, previous studies have used solubilisation and renaturation of protein expressed as inclusion bodies, and more recently soluble silicatein-α was obtained through the addition of non-denaturing detergents in the lysis buffer. As the well-established techniques used in the approach to renaturation did not

produce soluble protein, the focus was aimed at optimising buffer conditions for silicatein-α with the addition of amino acids (L-arginine and L-glutamate). This was shown to be beneficial in both proteins, TF-silicatein-α and silicatein-α, in terms of expression levels and reducing the amount of processing steps during purification. The isolation of silicatein-α from the fusion protein was attempted; the addition of Strep tag to the C-terminal offered more positive results but will need further optimisation.

In summary, the work detailed in this chapter describes how the purification and isolation of two proteins, TF-silicatein-α and silicatein-α was optimised successfully. The purity of TF-silicatein-α was increased with ammonium sulphate as a pre-purification step, more efficient buffer conditions for both proteins were developed and the purification process was streamlined to reduce overall time.

3. Chapter 3: Biochemical analysis of silicatein-α and variants

3.1. Introduction

For a heterologously produced protein to serve its proposed applications, for example as a biocatalyst or therapeutic protein, it must retain its original biological functionality. Therefore, following isolation it is essential to confirm this by conducting specific functional tests or assays, which are available to establish catalytic efficiency and viability of the enzyme. As no crystal structure has been resolved for silicatein-α due to a high propensity to aggregate, theoretical structural data has so far been obtained by homology modelling using cathepsin L and chimeric protein, cathsilicatein.[71] Mutations of the proposed active site residues have been investigated to confirm their role in enzyme catalysed hydrolysis and condensation of alkoxysilanes[1] but to date extensive studies on each residue has not been documented for organosiloxanes.

3.2 Biochemical analysis

3.2.1 Silicomolybdic acid assay

As silicic acid is presumed to be the natural substrate of silicatein-α, the silicomolybdic acid assay (SMAA) can be applied to assess and confirm catalytic ability of the enzyme. This well established and long standing method is based upon the formation of silico-12-molybidic acid, a heteropolyacid, from silicic acid and acidified ammonium heptamolybdate. This product is silicomolybdic acid (λ_{max} = 400 ± 10) and can be quantified spectrophotometrically. However, the limit of detection at this wavelength is approximately 10^{-4} mol L^{-1} due to the relatively low molar extinction coefficient of (1500 ± 20 L mol^{-1} cm^{-1}). An additional limitation is the formation of phosphomolybdic polyacid from phosphate ion present in some biological buffers, as this also produces a yellow compound that strongly absorbs at 400 nm. Therefore, by reducing silicomolybdic acid to molybdenum blue more accurate measurements can be

made. This is due to limit of detection in the order of 5 x 10^{-6} mol L^{-1} as the complex has λ_{max} = 810 nm and molar extinction coefficient of 44700 ± 150 L mol^{-1} cm^{-1} [192]

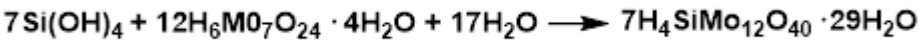


Figure 3.1 Scheme showing the reaction for the formation of molybdenum blue.

This method has been used for the confirmation of silicatein-α's activity as a hydrolytic enzyme in hydrolysis and condensation reactions with Tetraethyl orthosilicate (TEOS).[1,65,72,193] TEOS is an alkoxysilane substrate, which is analogous to the proposed natural substrate of the enzyme, orthosilicic acid. However, limitations of this method include it being insufficient for collation of kinetic data, several steps involved in the analysis making it prone to error and incompatibility for organosilicon compounds.

3.2.2 Hydrolysis of TBDMS-4np

With the potential application of silicatein-α as a biocatalyst in organosilicon chemistry, it is essential to assess its catalytic ability towards such substrates. However, as no high throughput assay existed previous research conducted in our group resulted in the development of a rapid colorimetric method to determine the Michaelis-Menten kinetic constants for this class of enzymes.[53] The data required to calculate these parameters is derived experimentally through measuring product formation as a function of time, using a variety of known enzyme and substrate concentrations.[194–197] Therefore, the design of this assay was based on long established methods used to determine efficiency of hydrolases.[77,198–201] The assay uses a 4-nitrophenol moiety linked to silyl protecting group (i.e. a 4-nitrophenoxy silyl ether) as substrate for silicatein-α. Subsequent hydrolysis leads to the release of a chromogenic product, 4-nitrophenoxylate ion, which can be spectrophotometrically quantified at 405 nm. The convenient and quantitative nature of this

assay means that it can be used as a tool for comparative site directed mutagenesis studies with variants of the enzyme. [53]

3.2.3 Exploring active site residues through site directed mutagenesis

As no crystal structure exists thus far for silicatein-α, the exact reaction mechanism has only been proposed. As discussed in Chapter 1, this proposal is based on the mechanism of homologous protease cathepsin L as the active site of silicatein differs only by one residue in comparison (Cys>Ser) and the mechanism of serine proteases which similarly also differs by one reside in the catalytic triad (Asp>Asn)[1,2,71] Comparative analysis using single amino acid mutations is usually undertaken to study variants of the enzyme to identify the importance of proposed catalytic residues or their roles in catalytic ability. Alanine substitutions are often employed to substitute particular residues with small and chemically inert side chains. Therefore, the importance of the substituted residue towards activity and stability can be assessed as alanine offers no steric hindrance or possibility of participation in the chemical reactions in the assay. This method has been used to investigate catalytic residues serine and cysteine from serine and cysteine hydrolases, respectively.[120–123,125,202,203] Previous work by Morse using silicatein-α from *Tethya aurantia* investigated the active site residues serine 25 and histidine 165 through mutagenesis and substitution of these residues with alanine. This reported of a reduction of relative catalytic specific activity to 10% and 8% for His165Ala and Ser25Ala respectively,[1,72,193] implying a pivotal role of these residues in catalytic activity. Notably, these activity tests were conducted using TEOS as a substrate; therefore, comparative analysis of silicatein-α active site variants, using organosilicon compounds as substrate, would be beneficial and more incisive in revealing the catalytic role of these residues.

3.2.4 Aims and objectives

The aim of this chapter is to conduct biochemical tests to confirm the fusion protein is catalytically active. Biochemical assays will be carried out to determine if it is catalytically competent against some model silicon containing small molecule substrates for both condensation and hydrolysis reactions. The variants will then be subjected to the same assay to investigate the role of each of the proposed catalytic residues. Further to this, variants

Ser26Ala, His165Ala, Asn185Ala, Asn185Asp and Ser26Cys with the conserved active site residues of serine and cysteine proteases will also be assessed as a comparative study.

The enzyme variants Asn185Asp and Ser26Cys will also be tested for esterase and protease activity and compared to the activity of commercially sourced serine and cysteine proteases.

3.3. Results and Discussion

3.3.1 Biochemical analysis of silicatein-α using SMAA hydrolysis

To assess the catalytic ability of silicatein-α as a hydrolytic enzyme, TEOS was incubated with the enzyme and for one hour at 22°C, pH 8.0. This process was also conducted using variants of the enzyme, Ser26Ala, His165Ala, Asn185Ala, Asn185Asp and Ser26Cys along with samples containing no enzyme or heat denatured enzyme. As the results show, (Figure 3.2) both silicatein-α and TF-silicatein-α have significant activity (100% relative catalytic specific activity) in comparison to the samples containing no and heat denatured enzyme which have a relative catalytic specific activity of 1.4%-1.6% and 14%-17% respectively. The increase in precipitated silica within enzyme samples confirms catalytic ability of the enzyme both in its assumed native state and as a fusion protein. However, it is noted that non-specific acid-base catalysis would explain the product formed in the samples with heat denatured protein and the control. To explore and confirm specific catalysis by the catalytic triad amino acid individual site directed mutagenesis of these residues was conducted replacing them with alanine, Ser26Ala, His165Ala, Asn185Ala. In all cases activity was reduced in comparison to the enzyme sample, upon removal of any of these proposed catalytic residues thereby suggesting the importance of each residue to the overall function; it is worth noting that the replacement of Ser26 to alanine had the most reduction in activity which may be due to the loss of the proposed initiating nucleophilic residue. These results are in agreement with past literature in which Ser26Ala and His165Ala variants of silicatein-α were shown to display a reduced relative catalytic specific activity in comparison to the wild type enzyme. This is further confirmed when Ser26 is replaced by a nucleophilic cysteine residue, as activity is then increased in comparison with the Ser26Ala variant to a value similar to that of the other variants. An interesting result came from the Asn185Asp with 72% relative catalytic specific activity also suggesting that all three of the natural catalytic residues are collectively important for full catalytic efficiency (Figure 3.2). Another explanation of non-specific catalysis observed in the reactions containing the variants is that other factors, such as water molecules or other residues in the vicinity, maybe participating in the turnover of the substrate albeit to a slower rate than the reactions with the unmodified enzyme.

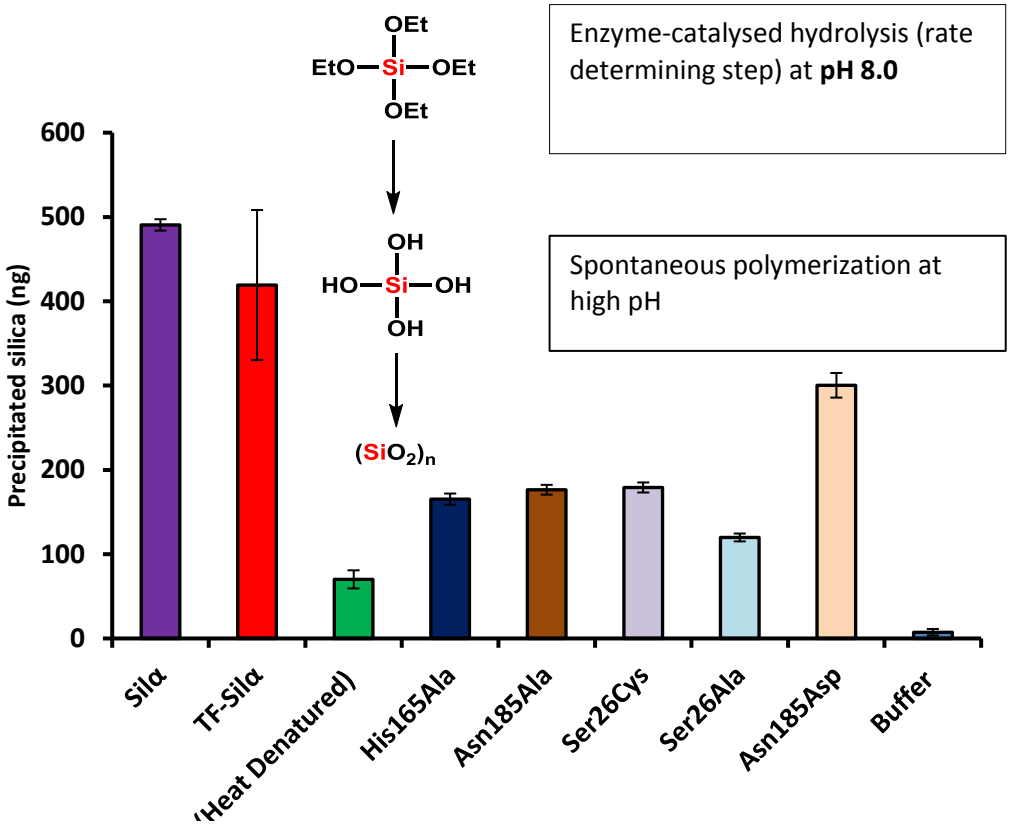


Figure 3.2. Precipitated silica formed by silicatein-α and variants of the enzyme after 1 hour incubation with TEOS and quantification by Silicomolybdic acid assay and subsequent analysis with UV-Vis spectrophotometer to investigate hydrolysis and specific activity.

3.3.2 Biochemical analysis of silicatein-α using SMAA condensation

To investigate catalytic ability of silicatein-α in the condensation of orthosilicic acid monomers to polymerised silica, pre-hydrolysed TEOS was used as a substrate. As spontaneous polymerisation has been observed at high pH levels, the pH of the samples were adjusted to pH 7.0 to circumvent this.[204] The remaining conditions of the experiment and analysis of samples following termination of the reaction were exactly the same as TEOS hydrolysis. Similar results to the hydrolysis reaction were also observed. A significant amount of precipitated silica was produced in samples containing silicatein-α and TF-silicatein-α in comparison to the heat denatured and control sample, containing only buffer and TEOS, suggesting these enzymes enhance product formation (Figure 3.3). Additionally, specific activity studies using individual site directed mutagenesis of active site residues produced results that show a reduced activity in comparison to the wild type enzymes that was again

higher than the results from samples containing heat denatured protein or the control reaction. In a similar manner to the hydrolysis investigations, the Ser26Ala has the lowest product formation which can also be attributable to the absence of nucleophilic serine residue. However, when Ser26 is replaced by cysteine residue this activity increases slightly. The cysteine residue also functions as a catalytic nucleophile in cysteine proteases, however, the Si-S bond energy is noted to be much weaker (293 kJ/mol) in comparison to the Si-O bond energy (452 kJ/mol) so therefore it may not have preference or good specificity to attack the silicon atom. The variant Asn185Asp was able to produce more than double precipitated silica levels compared to the other variants but its relative catalytic specific ability was only 70% of the wild type enzyme (Figure 3.3). This once again implies that all three residues within the catalytic site have importance and function in tandem. Additionally as mentioned before, as activity is not completely abolished other elements may function to catalyse turnover in the absence of an active site residue albeit to a slower rate.

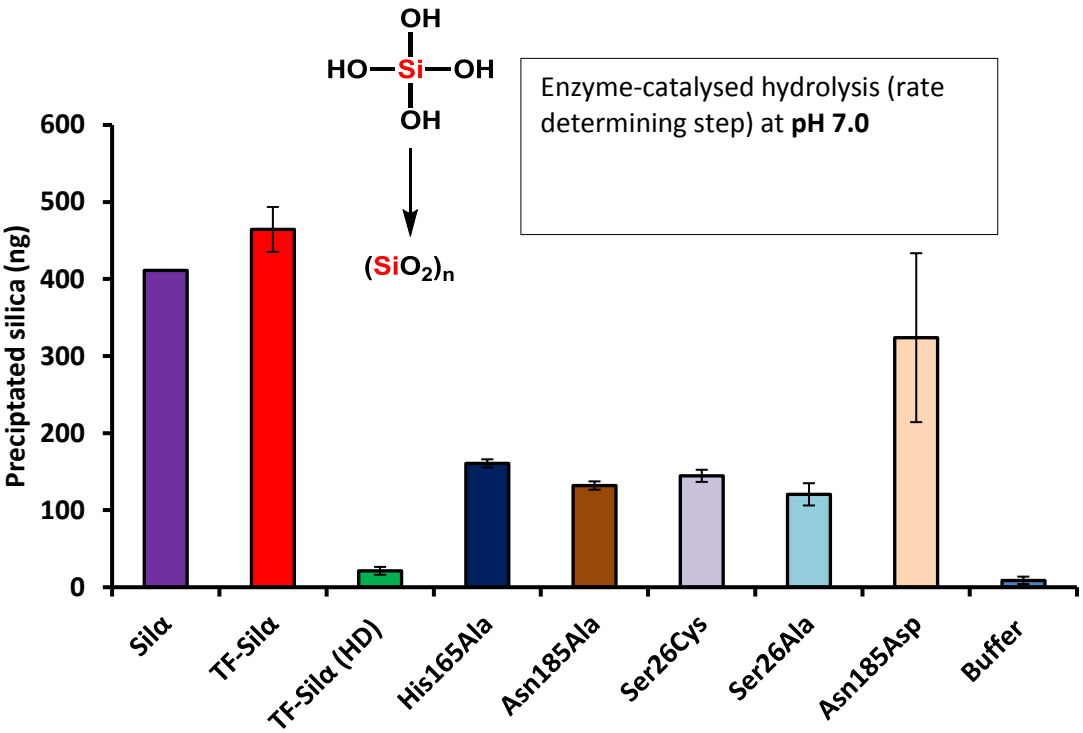


Figure 3.3. Precipitated silica formed by silicatein-α and variants of the enzyme after 1 hour incubation with pre-hydrolysed TEOS and quantification by Silicomolybdic acid assay and subsequent analysis with UV-Vis spectrophotometer to investigate condensation and specific activity.

3.3.3 Biochemical analysis of silicatein-α using hydrolysis of TBDMS-4np

The recent development of a high throughput assay to investigate catalytic activity of silicatein-α with organosilicon substrates has the advantage of enabling the ability to obtain

kinetic parameters to aid a more quantitative investigation into the catalysis of Si-O bond hydrolysis. Therefore to further add to our knowledge of the enzyme and investigate the importance of the catalytic triad residues, the variants were also applied to this assay to assess catalytic activity of silyl bond hydrolysis in aqueous solution.[53] The substrate used was TBDMS tethered to 4-nitrophenol moiety, (Figure 1.5) one of the three substrates synthesised and tested during the development of the assay. However, although this organic substrate was shown to be stable in buffer it requires the addition of a water miscible organic solvent, 1,4-dioxane to aid its solubility in aqueous buffer. Time course experiments measuring the hydrolysis of the substrate over time, at a range of known concentrations, were carried out. The progress curve of product formation over time (Figure 3.4) shows there is a significantly higher concentration of product formed in the sample containing the enzyme compared to the control sample. We do note that there is a significant background hydrolysis in the control sample possibly due to non-specific acid-base catalysis and because the substrate is labile due to the leaving group properties of the *p*-nitrophenyl group. Previous comparisons with samples containing heat denatured enzyme and Ser26Ala variant have also shown results that conclude the enzyme to be essential in the accelerated rate of product formation over time. However, to explore the contributory effect of the other active site residues, reactions containing the enzyme variants His165Ala and Asn185Ala were also conducted. Comparing the results for Ser26Ala, His165Ala and Asn185Ala shows the activity of each is around 50% higher than that of the control but around 50% lower that the wild type. This suggests that each of these residues is equally as important for the full catalytic activity of the enzyme. This trend was also observed during the silicomolybdic acid assay with the wild type and enzyme variants shown above. Additionally, comparing the Ser26Ala, with chemically inert alanine, against Ser26Cys, thiol group present in cysteine, shows an increase in product formation by around 8% in the latter. This thereby suggests that the presence of a nucleophilic residue alongside the other native catalytic residues raises the activity slightly but is still not specific enough to increase the activity to that of the wild type. A further observation which supports that notion of equal importance of catalytic triad residues is seen in the Asn185Asp variant. This reaction shows an increase in product formation of around 27% in comparison the other variants but has around 22% lower activity against the wild type enzyme. Therefore, substitution of the Asn residue to derivative form is not specific enough to restore 100 % activity but it does however, have a comparable increase observed over the other variants.

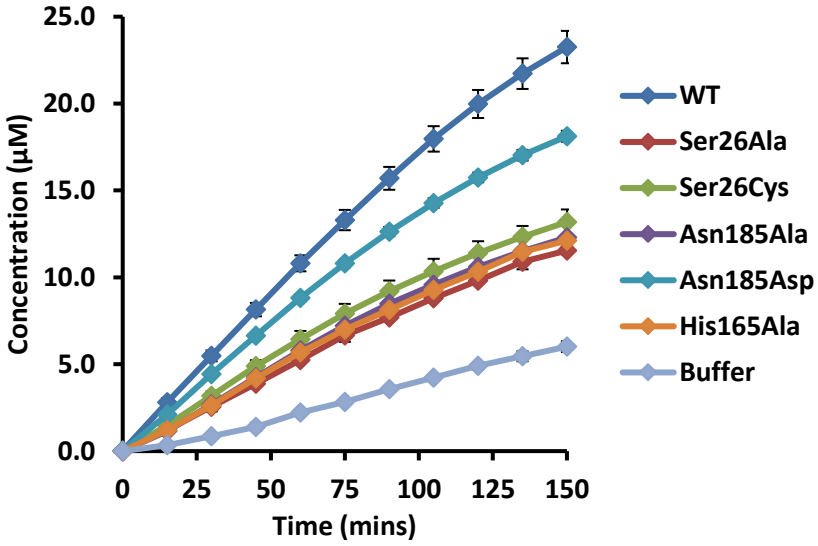


Figure 3.4 The hydrolysis of TBDMS-4-nitrophenol against time by TF-silicatein-α and enzyme variants using 0.0006 mol. equiv. of enzyme relative to the substrate in 5% p-dioxane, 50 mM Tris, 100 mM NaCl pH 8.5 at 22°C.

3.3.4 Michaelis Menten Kinetic Parameters using TBDMS -4- nitrophenyl silyl ether

The initial rates of reaction were measured using known concentrations of substrate, TBDMS-4-nitrophenyl ether, and known concentration of each enzyme and used to determine the kinetic parameters using the Michaelis Menten model and equation (Figure 3.5; Figure 3.6; Table 3.1). Looking at the binding constants for each enzyme, it is noted that they are in the µM range and therefore have a relatively weak affinity for this substrate. The K_m are within the range of ~60% of enzymes that have been analysed through data mining to determine trends in kinetic parameters, as are the k_{cat} for each enzyme tested here.[205] The results show the K_M values are mainly unaffected by the mutations with the exception of Asn185 mutation to both Ala and Asp when a minor decrease in value is observed with both. However, the effect of the mutations on turnover rate is not advantageous for any of the mutants that were tested in comparison to the wild type enzyme. As Asn185Ala has a k_{cat} ~5.5 fold lower than the wild type, and the substitution of Asn with Asp leads to a k_{cat} ~3 fold lower. As it was thought that Asp residue, in the catalytic triad of serine proteases, functions to stabilise the ion pair generated on His residue and orient the His tautomer in the correct conformation, which could explain the reduction in activity in Asn185Ala variant.[121,206] Substitution of Asn to Asp does lead to a 1.8 fold increase in turnover when comparing it with the Asn185Ala

mutant. Replacing the nucleophilic Ser with Ala has ~5.3 decrease in turnover number with a small observable increase when the Ser replaced by another nucleophilic residue, Cys. Reduction of activity is also observed in the His165Ala variant compared to the WT enzyme. Collectively the kinetic parameters imply that for full catalytic activity for this substrate to occur all three active site residues are required to be present. As the single mutations do not completely abolish activity, there may be other factors that are also responsible and contribute to substrate turnover as is the case for serine protease. Mutations of these active site residues have also shown the K_M is mainly unaffected and although turnover is reduced to a certain degree it is not completely abolished. This may be due to other amino acids or water fulfilling the role of the substituted residues and oxyanion hole for stabilisations.[74,121,122,203,207,208] When assessing k_{cat}/K_M values the data shows WT has the highest catalytic efficiency. There is some reported controversy over the suitability of k_{cat}/K_M data derived from experiments comparing different enzymes with one substrate and those with a number of substrates.[196,209] As the K_M values are similar the k_{cat}/K_M, the data can be assumed to be reliable in evaluating the contribution of each catalytic residue to overall enzymatic performance. Therefore an important observation to note is the WT has the highest turnover which agrees with previous data collated to assess the feasibility of silicatein-α as a good enzyme for application in organosilicon chemistry.[53]

Table 3.1. Table of estimated kinetic parameters (Michaelis Menten constants) for the wild type enzyme and variants for the hydrolysis of TBDMS-4-nitrophenyl.

Enzyme	$K_{M\ (\mu M)}$	$k_{cat\ (min^{-1})}$	$k_{cat}/K_{M\ (min^{-1}\mu M^{-1})}$
WT	21.6 ± 2.8	972.1 ± 223.9	46.0 ± 15.8
His165Ala	20.6 ± 5.4	257.2 ± 127.2	12.5 ± 2.7
Ser26Ala	20.3 ± 2.7	182.8 ± 43.6	9.0 ± 1.0
Asn185Ala	13.9 ± 2.7	175.4 ± 41.5	12.6 ± 4.9
Ser26Cys	19.1 ± 6.2	216.5 ± 115.2	11.1 ± 3.3
Asn185Asp	15.5 ± 1.34	319.4 ± 5.13	20.8 ± 1.5

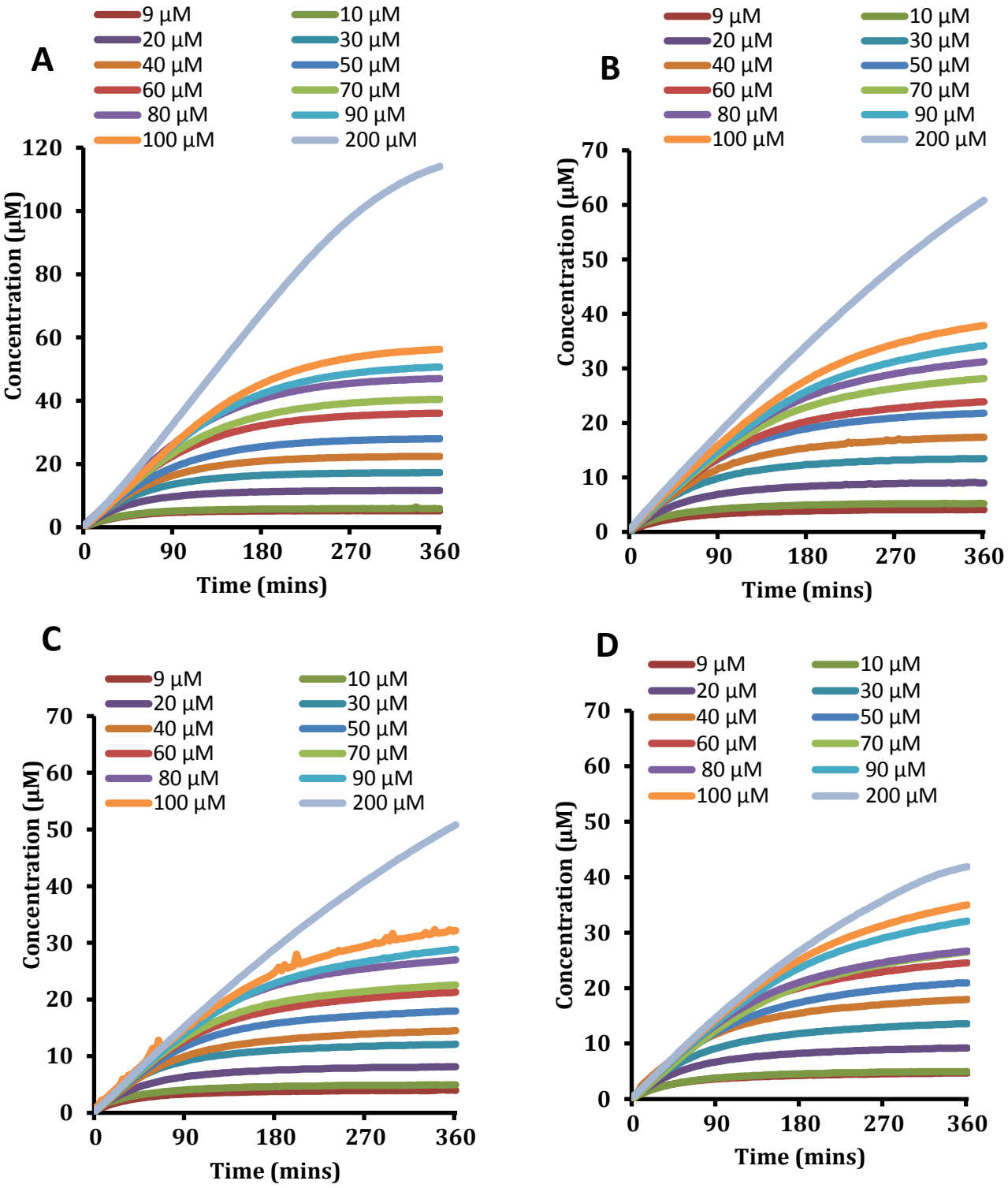


Figure 3.5 The concentration of 4-nitrophenol formed against time from hydrolysis catalysed by TF-silicatein-α (A), His165Ala (B), Asn185Ala (C), and Ser26C (D).

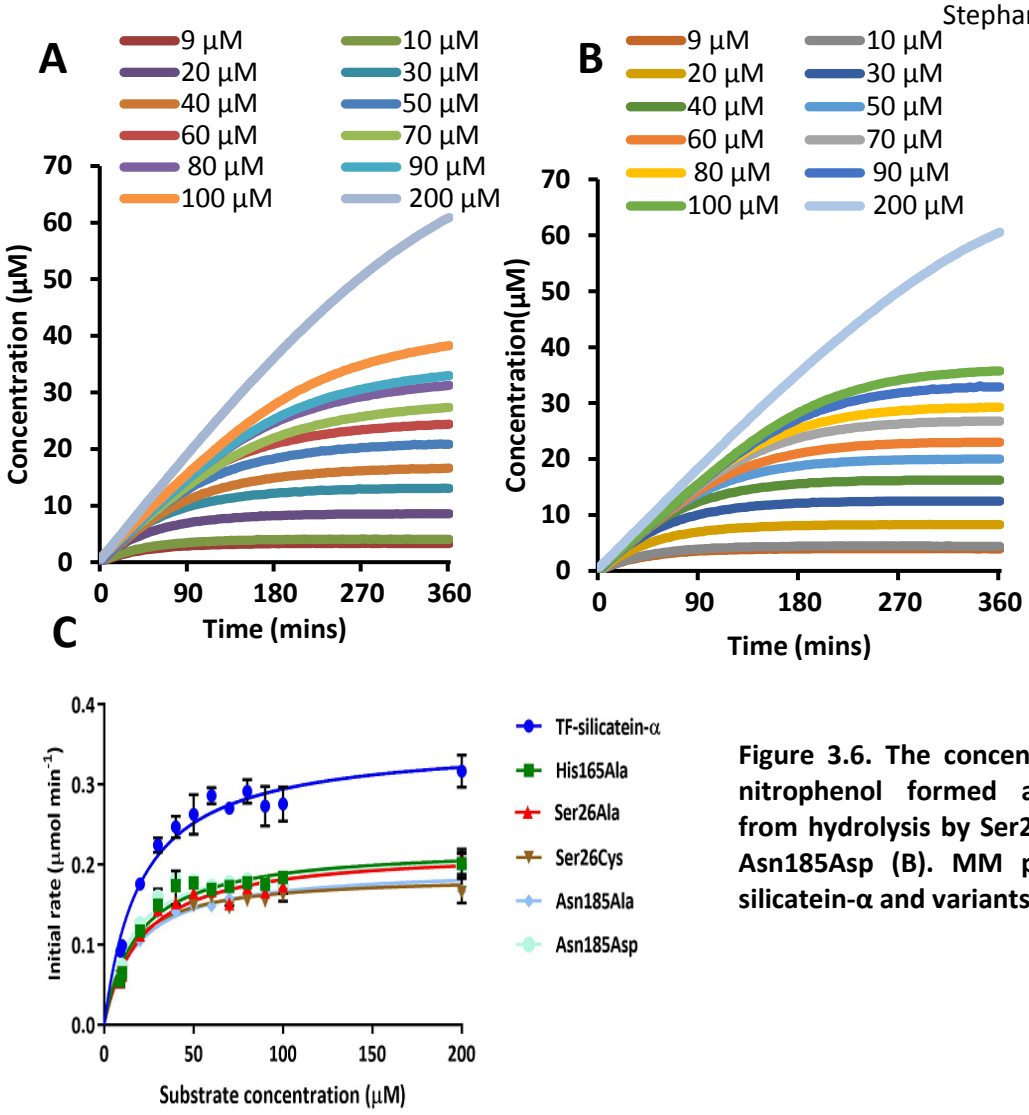


Figure 3.6. The concentration of 4-nitrophenol formed against time from hydrolysis by Ser26Ala (A) and Asn185Asp (B). MM plots for TF-silicatein-α and variants (C).

3.3.5 Esterase and Protease activity of silicatein-α and variants

The esterase and protease activities of silicatein-α were investigated due to its homology to the protease cathepsin L and its similarities with the conserved active site residues of other hydrolases. Here, the esterase and protease activity of the variants Ser26Cys and Asn185Asp mutants were studied as these mutations made the active site residues analogous to cysteine and serine proteases respectively. The substrates synthesised and used in this assay were analogous to TBDMS-4nitrophenyl but contained either an ester (4-nitrophenyl pivalate) or amide (4-nitrophenyl pivalamide) bonds in place of the dimethylsiloxy group. Additionally, a dipeptide (Cbz-Phe-Arg-NHNp) containing the cleavage site specific for Cathepsin L was also applied to the variants. As a measure of control and comparison TF-silicatein-α, cathepsin L and chymotrypsin was also used in the assay. Analysis of the results confirms previous

results[53] that TF-silicatein does not display esterase and protease activity against the three substrates used in these assays (Figure 3.6A and Figure 3.6B, Table 3.2). It also confirms that the serine protease chymotrypsin and cysteine protease cathepsin L were able to hydrolyse *p*-nitrophenyl pivalate and the commercially sourced dipeptide (Figure 3.6A and Figure 3.6B respectively), none of the proteins tested demonstrated activity against p-nitrophenyl pivalamide. An interesting observation was made with the Asn185Asp variants as it displayed someesterase activity. (Table 3.2).

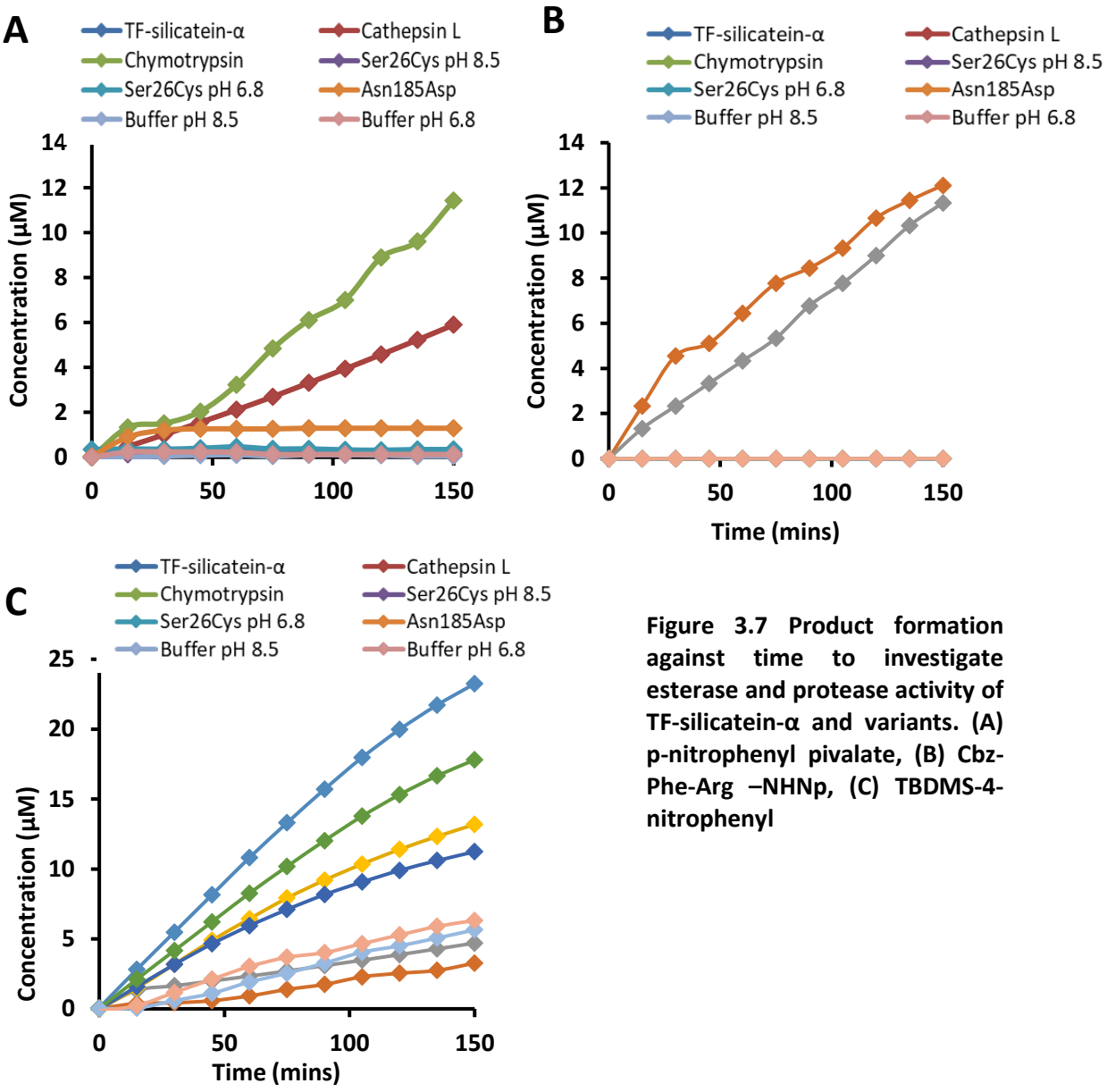


Figure 3.7 Product formation against time to investigate esterase and protease activity of TF-silicatein-α and variants. (A) p-nitrophenyl pivalate, (B) Cbz-Phe-Arg –NHNp, (C) TBDMS-4-nitrophenyl

Table 3.2. Table showing concentration of product formation for each enzyme against each substrate used to investigate esterase and protease activity. All reactions were performed for 6 hours with 0.0006 mol. equiv of enzyme and 0.1 mM of substrate in p-dioxane (5%), 50 mM Tris, 100 mM NaCl at the relevant pH. Conversions were calculated following subtraction of background hydrolysis.

Enzyme	Product formation (μM)				pH
	4-Nitrophenyl TBDMS	4-Nitrophenyl Pivalate	4-Nitrophenyl Pivalamide	4-Nitrophenyl Cbz-Phe-Arg	
TF-silicatein-α	**29.27 ± 1.151**	<0.01 ± 0.00	<0.01 ± 0.00	<0.01 ± 0.00	8.5
Chymotrypsin	**7.02 ± 0.003**	**27.47 ± 1.568**	<0.01 ± 0.00	**22.33 ± 2.520**	8.5
Cathepsin L	**10.39 ± 0.006**	**12.58 ± 0.613**	<0.01 ± 0.00	**19.00 ± 0.951**	8.5
Asn185Asp	**22.47 ± 0.992**	**1.24 ± 0.072**	<0.01 ± 0.00	<0.01 ± 0.00	8.5
Ser26Cys	**21.21 ± 0.007**	<0.01 ± 0.00	<0.01 ± 0.00	<0.01 ± 0.00	8.5
Ser26Cys	**15.24 ± 0.417**	<0.01 ± 0.00	<0.01 ± 0.00	<0.01 ± 0.00	6.8

Experimental analysis of Asn185Asp shows that despite this variant having catalytic triad residues identical to chymotrypsin (Ser>His>Asp), it did not display any protease activity but a low level of esterase activity was observed. The Ser26Cys variant, which has a catalytic triad identical to cathepsin L (Cys>His>Asn), did not hydrolyse the substrates used to test esterase and protease activity. The most likely explanation for this is other factors play a role in catalytic ability such as stabilisation of the oxyanion intermediate. Therefore, one single amino acid mutation is not sufficient to change substrate specificity.

3.4. Conclusion

This chapter aimed to further expand the mechanistic and structural knowledge of silicatein-α through biochemical tests. Following the development and optimisation of a reliable and reproducible expression and purification method to isolate an acceptable yield of protein, described in Chapter 2, further analysis could be applied. The work in this project was aimed at confirming the activity of the protein through two biochemical assays, a well-established method to investigate enzymatic role in the hydrolysis and condensation using an alkoxysilane substrate and a recently developed high-throughput, colorimetric technique to screen organosilicon substrates. The results for hydrolysis and condensation reactions were analysed by the silicomolybdic acid assay and hydrolysis of TBDMS-4-nitrophenyl confirmed catalytic activity of silicatein-α to the same level as previously reported.

To explore and confirm specific catalysis of the substrates by the active site residues and to study the contribution of each residue, site directed mutagenesis of each one was conducted replacing them with Alanine and applying them to the same biochemical assays; silicomolybdic acid assay and hydrolysis of TBDMS-4-nitophenyl ether. Experimental analysis of the results revealed a similar trend observed in all the assays. The activity of the active site variants was reduced in comparison to the wild type protein but showed higher activity compared to the control samples. The increase in activity over the control sample, plus the reduced activity against the wild type suggests these amino acids facilitate full catalytic activity. As turnover is not completely abolished or at the same level as the control, there may be other factors that contribute to activity such as other amino acids in the periphery or water molecules.

The colorimetric assay to assess catalytic activity of silyl bond hydrolysis in aqueous solution enabled kinetic parameters of silicatein-α and variants to be estimated and compared. Similar K_m values for all proteins were recorded suggesting substrate-enzyme binding was unaffected by the mutations although the k_{cat} value was higher for the wild type which implies the presence of all three catalytic residues are essential for full activity. In a similar manner to silicomolybdic acid assay results, replacing Ser with Cys, a nucleophilic residue, did increase the turnover slightly in comparison to the Ala variant, and the variant with the highest turnover was Asn185Asp but activity was still lower than the WT. Further notable

observations came from Ser26Cys and Asn185Asp, mutated active site replicates of cysteine and serine proteases, displaying no protease activity, however, Asn185Asp was observed to have a low level of esterase activity. This result could due to only one single mutation being made, and other amino acids that stabilise the substrate and contribute to the overall function are not present, hence exhibiting no activity towards these particular substrates.

4. Chapter 4: Biophysical analysis of silicatein-α and variants

4.1 Introduction

For a recombinant protein to serve its proposed applications, for example as a biocatalyst or therapeutic protein, it must retain stability in solution. Therefore, following isolation the structural integrity must be examined (using a number of biophysical techniques) as the target protein needs to be homogeneous and in its correct conformation. These methods also tend to be perceived as prerequisite assessments for candidates suitable for structural studies such as x-ray crystallography trials. Crystal structure determination is the gold standard in structure-function studies as it enables visualisation of protein structures at the molecular level. It can also advance our knowledge of protein function as conformational changes can be observed through co-crystallisation with substrates or inhibitor molecules. It can also indicate information such as bond lengths, position and importance of amino acids that function in the activity and enable the researcher to make well-informed proposals on mechanisms and decisions on further mutagenesis and rational engineering of the protein.[168,210–213] As no crystal structure has been resolved for silicatein-α due to a high propensity to aggregate, theoretical structural data has so far been obtained by homology modelling using cathepsin L and chimeric protein, cathsilicatein.[71] However, with the advancement of high throughput light scattering technologies as a technique to screen various buffering conditions as potentials for crystallisation trials resolving the structure of silicatein-α could be a possibility.

4.2 Biophysical measurements for protein characterisation

To characterise the protein of interest several biophysical techniques can be applied to ensure the sample is homogeneous. Several methods allow for high throughput screening of buffer conditions to ensure optimum conformational stability of the target protein in solution.

4.3 Circular Dichroism (CD) for secondary structural information of proteins

Circular dichroism (CD) offers a biophysical estimation of secondary structure composition. It works on the principle of differential absorption of left and right components of circularised polarised light by the presence of a chiral chromophore in the sample. Due to several chromogenic features within proteins such as aromatic amino acid side chains and the peptide bond, spectral bands can be assigned to particular secondary structures within the molecule.[214,215] This can be used to confirm correct conformation of the protein of interest using software algorithms such as CDNN that estimates the percentage of secondary structure features using the CD spectrum.[216,217] CD is an extensively used structural biology technique which enables researchers to examine proteins in their operable environments, i.e. in aqueous solution and monitor changes in structure.

$$\Delta A = A_L - A_R$$

Equation 4.1 difference of absorbance of left and right circularly polarised light measured by CD spectropolarimeters

The differential absorption of these components occurs upon the presences of an optically active, chiral chromophore and results in radiation that has elliptical polarisation with subsequently measureable CD signal. Proteins have intrinsic chirality due to Carbon atom in the peptide backbone or the dihedral angles of disulphide bonds. The resultant CD signal produced is described as ellipticity ϑ (in degrees) calculated with $\vartheta = 32.98\ \Delta A$ (the numerical relationship between ΔA and ellipticity (in degrees)) and successive CD spectrum is acquired from dichroism measured as a function of wavelength. For its application in the study of proteins, spectral bands can be assigned to distinct structural features; therefore a number of spectral regions can give rise to structural data.

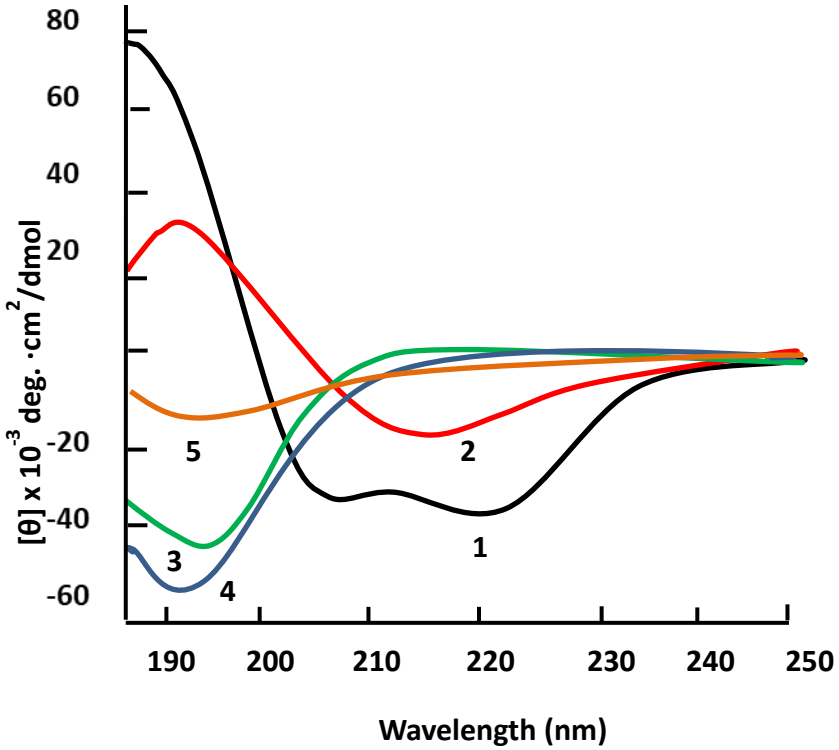


Figure 4.1 CD spectra of proteins and polypeptides to show characteristic secondary structure. 1 (black) is representative of α-helical conformations, 2 (red) is antiparallel β-sheet, 3 (green) is indicative of disordered proteins and CD spectra 4 (blue) and 5 (orange) show collagen in the native triple helix and denatured form, respectively.[268]

The chromophores present in proteins that absorb are the peptide bond (below 240 nm), disulphide bonds aromatic amino acid side chains (260 – 320 nm) and disulphide bonds (around 260 nm). Generally, secondary structure composition of a protein can be obtained from absorption by the peptide bond with a broad albeit weak transition around 220 nm

$(n \rightarrow \pi^*)$ and a stronger more intense signal around 190 nm $(\pi \rightarrow \pi^*)$

The different secondary structure features produce distinctive CD spectra allowing for characterisation of these features and estimation of percentage content of secondary structure composition of the protein of interest using an algorithm. A limitation to this is the analysis of oligopeptides which can only be studied reliably on the presence of a pre-dominant secondary structural feature.

4.4 Theory of Light Scattering

Light Scattering is a widely used biophysical technique in academic research and industry to measure and characterise a protein of interest in solution. It is non-destructive and provides information on protein sizing, molar mass, oligomeric form and homogeneity of the sample. The scattered light signal is collected by various detectors within the instrument which are placed at specific angles in relation to the incident light source as the light is scattered isotropically and anisotropically depending on the size of the molecule.[218–221] This technique provides important data about protein stability and conformation in solution.

4.4.1 Static light scattering

Static light scattering, also known as Rayleigh scattering or Multi-angled light scattering (MALS) enables measurement of a proteins molar mass and size from the time averaged intensity of scattered light, at a specific angle relative to an incident beam of light. The Rayleigh equation shows the directly proportional relationship between the intensity of scattered light and molecular weight. However, as this equation shows this theoretical relationship is based on the scattered light at 0° scattering angle, which is impossible to measure directly, due to the path of the incident beam of light. Therefore, in the case of MALS, simultaneous measurements are recorded by detectors placed at a number of specific angles and then transposed to zero angle (Figure 4.2).

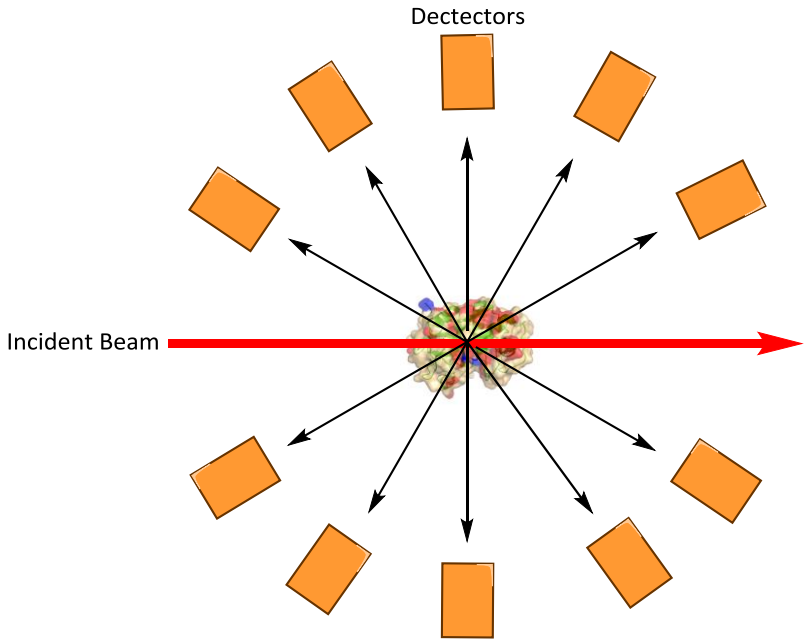


Figure 4.2 Schematic representation of Multi-angled light scattering

MALS equipment is often used in combination with size exclusion chromatography columns to fractionate protein samples and analyse them individually, independent of elution volume. The latter point accounts for elongated proteins or so called 'sticky' proteins that elute at uncharacteristic SEC volumes thereby making mass determination difficult by SEC alone. The technique also permits analysis of the native state of a protein in solution, whether this is monomeric or oligomeric, the presence of aggregates and the size of the molecules from the root mean squared (RMS) radius or radius of gyration (Rg). However, Rg (the mass distribution around a macromolecules centre of mass) is only detectable when the scattering signal is anisotropic and a macromolecule ~10 nm displays angular dependence. The limitation of this method comes from the quality of data as a result of locational placement of the detectors in relation to 0° scattering angle due to the relationship between this angle and molecular weight described in the Rayleigh equation. Therefore, this method is stated to have an accuracy of ~3-5 %. [222]

4.4.2 Dynamic light scattering

Dynamic Light Scattering (DLS), also referred to as quasielastic light scattering (QELS) or photon correlation spectroscopy (PCS), is based upon the same light scattering theory as SLS but disparately measures the rate of diffusion of macromolecules in solution. The method looks at fluctuations in the intensity of scattered light based upon Brownian motion, the random motion of particles in solution, as larger particles will move slower than smaller particles with temperature being directly proportional to the velocity of motion. This motion data is then processed to obtain size distribution of particles described as Stokes radius or hydrodynamic radius of the protein Rh. This value depends on the mass and shape and is an imperative valuable in comparing stability of different formulations and changes at elevated temperatures fluctuation of intensity of scattered light with time. However, limitations associated with this method are as the Rh is based upon the shape of the molecule, the method cannot discriminate between differences in oligomeric state and/or different protein shapes such as non-globular as this may be perceived as oligomerisation. The availability of instruments with high-throughput capabilities allows the rapid screening of buffer formulations to determine the optimum aqueous environments for protein stability. Variable temperature light scattering techniques also enable experiments to be conducted to gain insight into the range of temperatures in which the protein retains structural integrity (i.e. structural stability).[218–221] DLS measurements of silicatein-α have been reported by two groups, with the aim of characterising the function of self-assembly into natively formed filaments that comprise the exoskeleton of the marine sponge, and the effect of varying concentrations of urea on aggregation state. This technique was applied to characterise the oligomeric state of the protein sample through the recorded radius of hydration measurements. These studies show that due to the presence of five hydrophobic areas on the exterior of the protein it displayed a tendency to form oligomers with a radius of hydration measured to be 7 nm (Figure 4.3).[133,138]

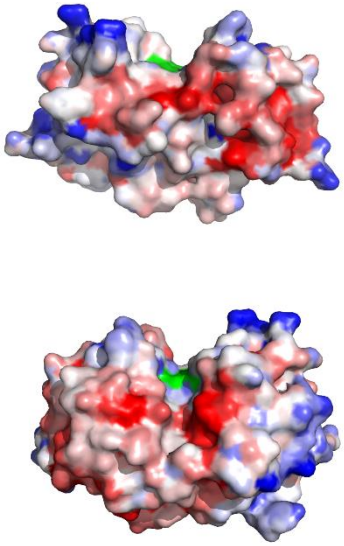


Figure 4.3 Homology model of Silicatein-α constructed using (PDB:2VHS) in PyMol showing the active site (green), negatively charged (red), positively charged (blue) and uncharged (white) surface residues.

4.4.3 Aims and objectives

To gain conformational knowledge and investigate optimum conditions for stability, biophysical techniques such as circular dichroism and light scattering will be applied. These methods will aid in the identification of suitable buffer formulations that maintain a non-aggregated, homogenous sample that can progress to crystallography trials.

CD measurement will allow for the analysis and estimation of secondary structure composition and subsequent comparison to theoretically calculated data using crystallographic information in PDB. In, addition to this, a range of light scattering techniques will be applied to investigate protein stability in a variety of buffering conditions.

4.5 Results and discussion

4.5.1 Circular Dichroism (CD) measurements and secondary structure prediction

To validate the aforementioned functional assays, structural analysis of the protein sample must be measured to eliminate possibility of protein aggregation or misfolding. As discussed in the introduction to this chapter, Circular Dichroism (CD) is a widely used technique to analyse secondary structure content.[215,216]

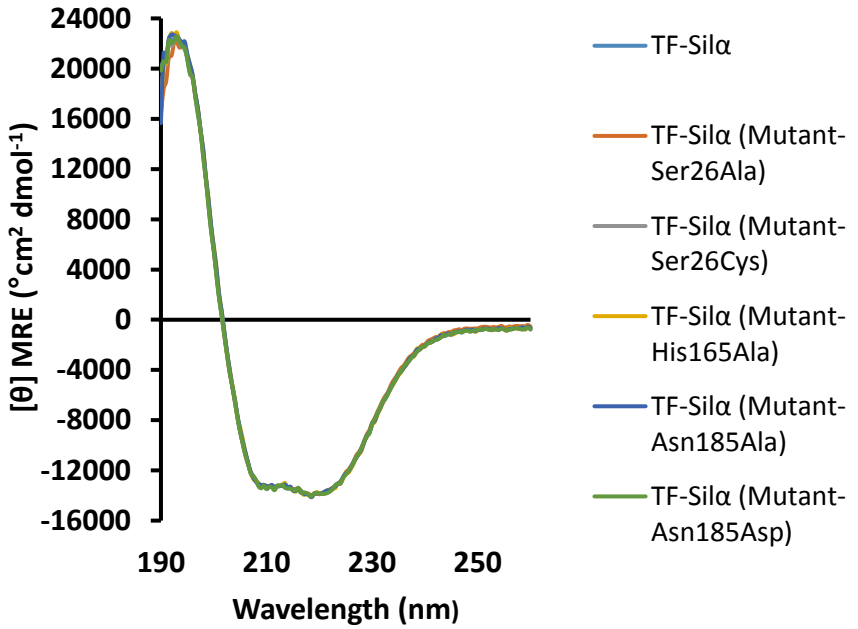


Figure 4.3. CD spectra for TF-silicatein, Ser26Ala, Ser26Cys, His165Ala, Asn185Ala and Asn185Asp plotting mean residue ellipticity (MRE) against wavelength

Experimental analysis of CD spectra comparing the wild type enzyme and TF-silicatein, with the variants with the single active site mutations, reveals there is no significant difference that would imply these mutations cause large structural changes. The peaks corresponding to alpha helix (positive at 193 nm, and negative at 222 and 208 nm) are present (Figure 4.3) suggesting also that the TF-silicatein-α and variants are mainly alpha helical. Plotting mean residue ellipticity (MRE) against wavelength enables comparison of other proteins to confirm non-aggregation and structural integrity. In this case, TF-silicatein and its variants were compared with silicatein-α and TF (Figure 4.4) and the results concur with the CD spectra of

the variants plotting mean residue ellipticity against wavelength. The peaks corresponding to an alpha helical protein are observed and the spectra show no disordered regions or aggregation.

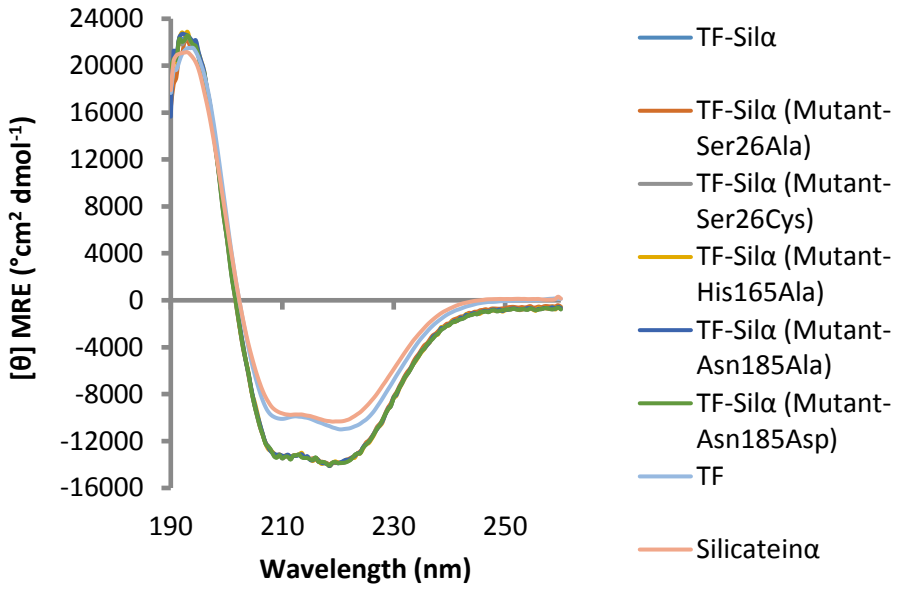


Figure 4.4 CD Spectra for TF-silicatein, Ser26Ala, Ser26Cys, His165Ala, Asn185Ala, Asn185Asp, TF and silicatein-α plotting mean residue molar ellipticity (MRE) against wavelength.

The data from Figure 4.3 was used to predict percentage of secondary structure feature that comprises each variant and the wild type TF-silicatein-α. (Table 4.1; Table 4.2; Table 4.3; Table 4.4; Table 4.5; Table 4.6;). This was compared to theoretical values for the fusion protein calculated from crystallographic data of homologue cathsilicatein (PDB ref. 2vhs) and TF (PDB ref. 1w26) (Table 4.7). This revealed the data calculated using CDNN was within 5% of the theoretical values so can be deemed reliable in terms of the output.

Table 4.1. Estimated proportions of secondary structure in TF-silicatein calculated by CDNN algorithm using CD spectrum in Figure 4.3

Wavelength (nm)	190-260	195-260	200-260	205-260	210-260	Average
α-Helix	31.70%	31.25%	31.15%	32.11%	31.22%	**31.49%**
Antiparallel	9.97%	10.12%	9.03%	8.05%	8.62%	**9.16%**
Parallel	8.67%	8.93%	9.23%	9.24%	9.02%	**9.02%**
β-Turn	17.15%	16.96%	17.06%	16.70%	17.05%	**16.98%**
Rndm. Coil	32.50%	32.74%	33.53%	33.90%	34.09%	**33.35%**
Total Sum	100.00%	100.00%	100.00%	100.00%	100.00%	**100.00%**

Table 4.2. Estimated proportions of secondary structure in TF-silicatein (Ser26Ala) calculated by CDNN algorithm using CD spectrum in Figure 4.3

Wavelength (nm)	190-260	195-260	200-260	205-260	210-260	Average
α-Helix	31.31%	31.25%	31.15%	32.11%	31.22%	**31.41%**
Antiparallel	10.34%	10.12%	9.03%	8.05%	8.62%	**9.23%**
Parallel	8.75%	8.93%	9.23%	9.24%	9.02%	**9.03%**
β-Turn	17.20%	16.96%	17.06%	16.70%	17.05%	**16.99%**
Rndm. Coil	32.41%	32.74%	33.53%	33.90%	34.90%	**33.33%**
Total Sum	100.00%	100.00%	100.00%	100.00%	100.00%	**100.00%**

Table 4.3. Estimated proportions of secondary structure in TF-silicatein (His165Ala) calculated by CDNN algorithm using CD spectrum in Figure 4.3.

Wavelength (nm)	190-260	195-260	200-260	205-260	210-260	Average
α-Helix	34.31%	32.35%	31.57%	32.22%	31.18%	**32.33%**
Antiparallel	12.24%	9.21%	10.76%	9.71%	10.49%	**10.48%**
Parallel	5.86%	9.53%	9.17%	8.71%	8.29%	**8.31%**
β-Turn	15.23%	15.18%	17.11%	15.83%	17.57%	**16.18%**
Rndm. Coil	32.36%	33.73%	31.39%	33.54%	32.46%	**32.70%**
Total Sum	100.00%	100.00%	100.00%	100.00%	100.00%	**100.00%**

Table 4.4. Estimated proportions of secondary structure in TF-silicatein (Asn185Ala) calculated by CDNN algorithm using CD spectrum in Figure 4.3

Wavelength (nm)	190-260	195-260	200-260	205-260	210-260	Average
α-Helix	32.16%	33.73%	32.03%	32.83%	31.48%	**32.45%**
Antiparallel	9.41%	9.75%	9.67%	9.51%	8.55%	**9.37%**
Parallel	9.35%	10.08%	9.81%	8.31%	9.55%	**9.42%**
β-Turn	17.03%	16.14%	17.44%	16.33%	17.68%	**16.92%**
Rndm. Coil	32.05%	30.31%	31.06%	33.02%	32.74%	**31.84%**
Total Sum	100.00%	100.00%	100.00%	100.00%	100.00%	**100.00%**

Table 4.5. Estimated proportions of secondary structure in TF-silicatein (Asn185Asp) calculated by CDNN algorithm using CD spectrum in Figure 4.3

Wavelength (nm)	190-260	195-260	200-260	205-260	210-260	Average
α-Helix	31.61%	31.26%	30.66%	31.06%	30.43%	**31.00%**
Antiparallel	9.64%	9.89%	9.10%	8.21%	8.79%	**9.13%**
Parallel	8.85%	9.00%	9.30%	9.50%	9.19%	**9.17%**
β-Turn	17.10%	16.91%	17.11%	16.92%	17.09%	**17.03%**
Rndm. Coil	32.80%	32.94%	33.83%	34.32%	34.49%	**33.68%**
Total Sum	100.00%	100.00%	100.00%	100.00%	100.00%	**100.00%**

Table 4.6. Estimated proportions of secondary structure in TF-silicatein (Asn185Asp) calculated by CDNN algorithm using CD spectrum in Figure 4.3

Wavelength (nm)	190-260	195-260	200-260	205-260	210-260	Average
α-Helix	30.04%	31.00%	31.00%	31.29%	30.89%	**30.84%**
Antiparallel	10.27%	9.74%	8.88%	8.22%	8.71%	**9.17%**
Parallel	9.00%	9.06%	9.28%	9.50%	9.11%	**9.19%**
β-Turn	17.03%	16.83%	16.88%	16.63%	17.03%	**16.88%**
Rndm. Coil	33.66%	33.36%	33.96%	34.36%	34.26%	**33.92%**
Total Sum	100.00%	100.00%	100.00%	100.00%	100.00%	**100.00%**

Table 4.7. Comparison of proportions of secondary structure in TF-silicatein and variants.

	TF-Silα Theoretical values	TF-Silα (Table3.2)	TF-Silα Ser26Ala (Table 3.3)	TF-Silα His165Ala (Table3.4)	TF-Silα Asn185Ala (Table 3.5)	TF-Silα Asn185Asp (Table 3.6)	TF-Silα (Ser26Cys) (Table 3.7)
α-Helix	33.18%	31.49%	31.41%	32.33%	32.45%	31.00%	30.84%
Antiparallel	9.29%	9.16%	9.23%	10.48%	9.37%	9.13%	9.17%
Parallel	5.22%	9.02%	9.03%	8.31%	9.42%	9.17%	9.19%
β-Turn	16.36%	16.98%	16.99%	16.18%	16.92%	17.03%	16.88%
Rndm. Coil	32.95%	33.35%	33.33%	32.70%	31.84%	33.68%	33.92%
Total Sum	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

4.5.2 Dynamic Light Scattering

As the CD spectra and the predicted secondary structure content for the fusion protein is in agreement with data collected for the native proteins, and thereby postulates no aggregation is present in the buffer conditions used for the hydrolysis reactions with TBDMS-4-nitrophenyl substrate, further structural characterisation of silicatein-α was considered using additional

biophysical techniques. Dynamic light scattering enables high throughput screening of a multitude of buffer environments, ionic strengths and temperatures to find optimum conditions for further downstream structural investigations. The data obtained from this technique can also determine quaternary structure for example it can show if the protein is monomeric or oligomeric under set conditions. Initial experiments were aimed at determining an acceptable concentration range that avoids aggregation. Stock concentrations of enzyme 0.5, 1, 1.5 and 2 mg/mL in 50 mM Tris, 100 mM NaCl, pH 8.5 were screened in a variable temperature experiment with stepped temperature increases of 4, 22, 30, 37, 45, 60 to 85 °C. Analysis of the results suggested a concentration of 2 mg/mL was the most reliable as other concentrations screened produced ambiguous results and light scattering signal above 37 °C suggested the presence of large aggregates. This effect above 37 °C under these buffer conditions may be due to the protein denaturing under the stress of increased temperature. Further screening of buffer additives known to improve protein solubility, various ionic strengths and different pH levels[166,168] was conducted using 2 mg/mL of protein and the same stepped temperature range as before (Table 4.8). As discussed in Chapter 1, the percentage polydispersity can be used as an indicator for homogeneity within the sample, % polydispersity ≤ 15 suggesting a high level of homogeneity and a potential candidate for crystallography trials and % polydispersity ≥ 30 conveys a low level of homogeneity and possible aggregation. Of the buffering conditions screened samples containing 100 mM Arg and Glu and standard assay conditions (50 mM Tris 100 mM NaCl pH 8.5) had percentage polydispersity value below 15 with monomodal size distribution across all temperatures 4-37 °C (Table 4.8).

Table 4.8. Percentage polydispersity estimated by Dynamic Light Scattering with TF-silicatein-α at a concentration of 2 mg/mL. Samples at pH 3.5 were in 50 mM Citrate buffer with the remaining samples in 50 mM Tris 100 mM NaCl

Additive	4°C	22°C	30°C	37°C
100 mM Ammonium Sulphate	48.0	20.9	15.9	25.7
200 mM Ammonium Sulphate	25.6	**5.6**	**13.8**	**13.3**
300 mM Ammonium Sulphate	22.3	21.1	18.2	**13.1**
100 mM Arginine + 100 mM Glutamate	**12.4**	**6.1**	**14.3**	**14.6**
300 mM Arginine + 300 mM Glutamate	46.5	27.4	**13.7**	18.1
500 mM Arginine + 500 mM Glutamate	68.5	26.4	27.6	16.2
NaCl omitted	23.9	**13.9**	**9.6**	**12.4**
300 mM NaCl	20.8	21.7	38.5	**12.7**
pH 3.5	**12.3**	38.5	40.0	31.0
pH 7.5	27.2	36.8	38.5	46.0
pH 8.5	**12.7**	**12.3**	**11.2**	**12.1**

However, analysing the estimated hydrodynamic radius and molecular weight for the samples under the standard assay buffer conditions shows values between 10.8-12.0 nm and 873.7-1125.3 kDa (Table 4.8). The estimated molecular weight and hydrodynamic radius from DLS measurements appears to be unrealistically high, as the theoretical molecular weight of a monomeric fusion protein is around 76.1 kDa. However, the estimates must be considered carefully, as they are derived using an empirical curve of known globular proteins and their hydrodynamic radius, which means it cannot be readily applied if the protein is not globular. It is worth considering the possibility of the fusion protein being highly elongated or having a highly ordered oligomeric structure due to native silicatein-α preference to self-assemble[133,138] and trigger factor natively forming oligomeric structures.[223,224] Allowing for all these possibilities and given thyroglobulin, a dimeric protein, has a molecular weight of 670 kDa and a measured hydrodynamic radius of 8.7 nm,[225] the molecular estimate molecular weights and hydrodynamic radius (Figure 4.5) can be regarded as realistic values.

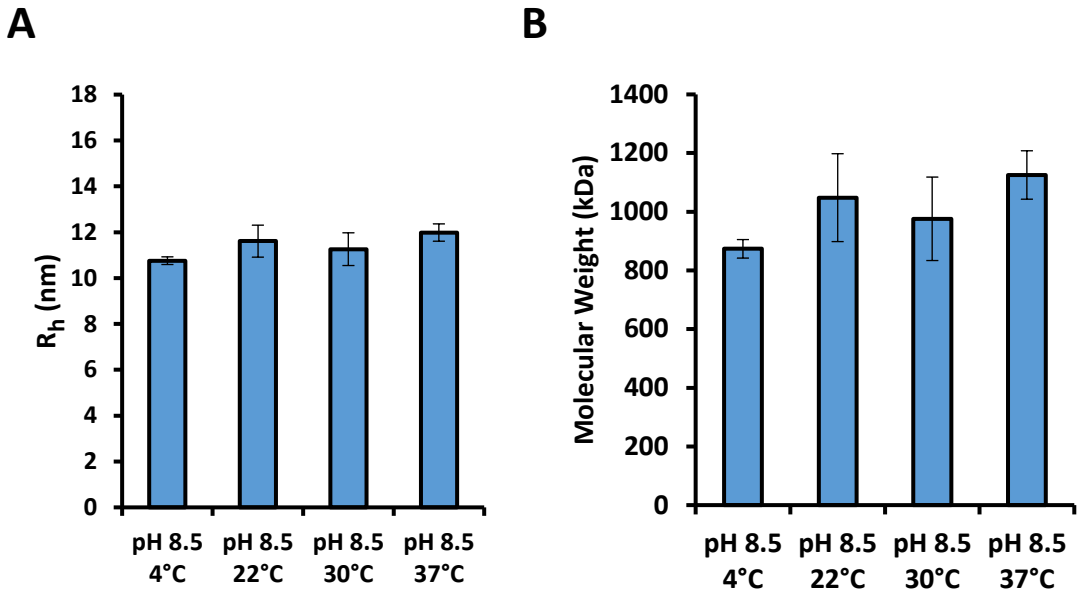


Figure 4.5 Dynamic Light Scattering experiment with TF-silicatein-α in 50 mM Tris, 100 mM NaCl pH 8.5 at 4 °C, 22 °C, 30 °C and 37 °C (A) estimated hydrodynamic radius (B) estimated molecular weight.

To further characterise the protein and determine optimum conditions, biophysical investigations were made using an alternative instrument that measures melting temperature (T_m) and the aggregation onset temperature (T_{agg}) alongside DLS measurements before and after the thermal ramp. Homogeneity was measured through the reported polydispersity index (PDI) for which a value below 0.7 indicates a high degree of homogeneity. Conformational stability under the various conditions screened was assessed from the T_m and T_{agg} which enables measurement of the maximum temperature the proteins can be exposed to before thermal denaturation and the onset of protein aggregation occur. This coupled with the insight into the homogeneity of the sample, under certain conditions, can aid in finding the optimum buffer conditions for protein stability and solubility. Firstly, the effect of salt concentration on homogeneity and protein structure stability with increasing temperatures was measured. Three different concentrations of NaCl (100 mM, 300 mM and 500 mM) were chosen to assess the chaotropic ('salting in') or precipitation ('salting out') effects of the resulting ions, based on the well described association in the Hofmeister series.[226]

Table 4.9. Table of polydispersity index values of silicatein-α and TF-silicatein-α with varying concentrations of NaCl at 22 °C.

	PDI Silicatein-α	PDI TF-Silicatein-α
100 mM NaCl	0.681	0.191
300 mM NaCl	0.634	0.229
500 mM NaCl	0.619	0.028

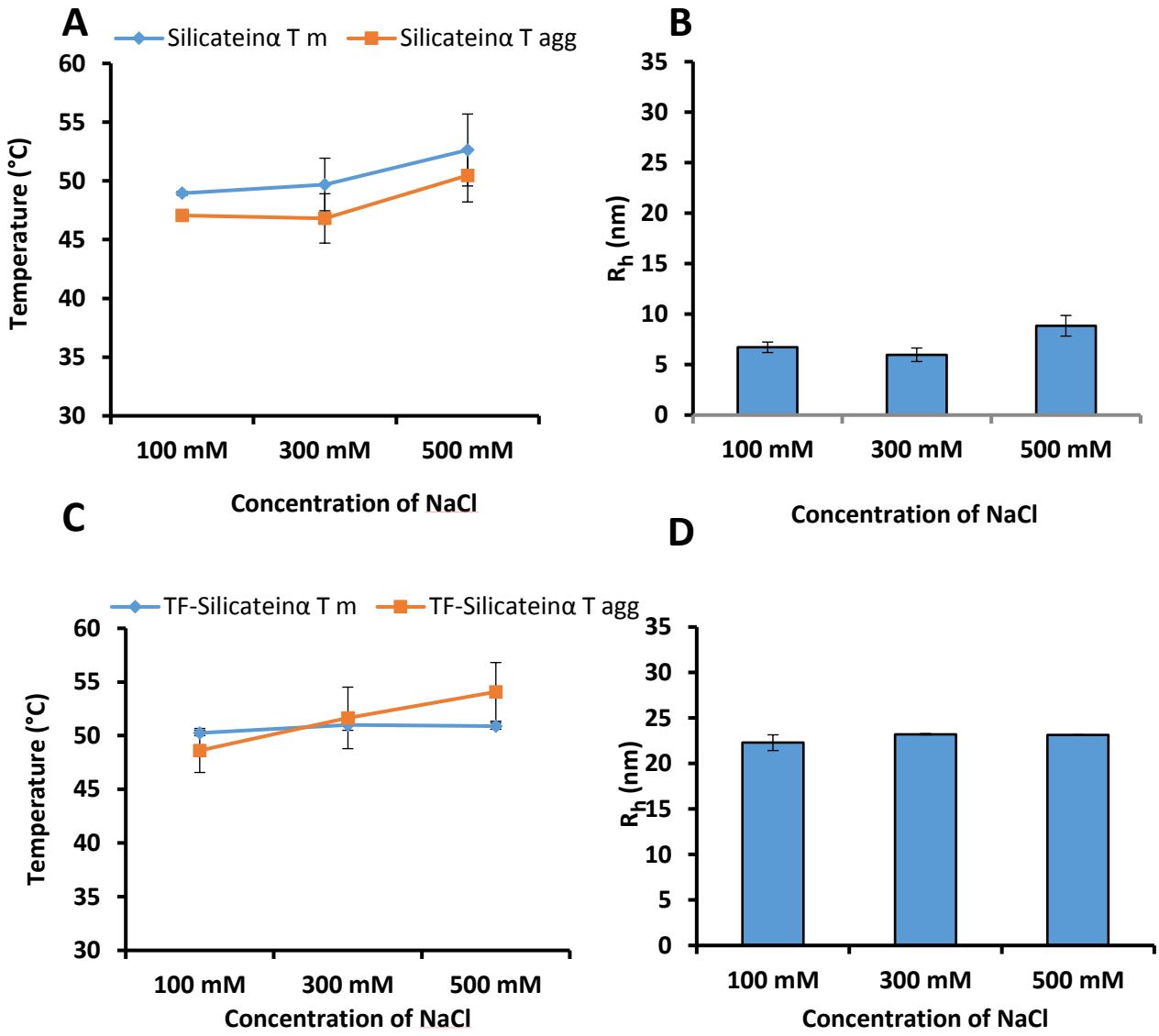


Figure 4.6. Light scattering experiments with varying concentrations of NaCl. (A) melting temperature of and aggregation onset temperature of silicatein-α, (B) estimated hydrodynamic radius of silicatein-α (C) melting temperature of and aggregation onset temperature of TF-silicatein-α, (D) estimated hydrodynamic radius of TF-silicatein-α.

Analysis of the results show PDI for both proteins are below 0.7 implying the sample has low level of polydispersity, however, PDI value for TF-silicatein-α at 500 mM suggests the sample

is monodisperse (Table 4.9). In thermal ramp experiments the increase in NaCl concentration from 100 mM to 500 mM also increases the melting temperature by 3.7 °C and aggregation onset temperature by 3.4 °C (Figure 4.6A). This may be due to the increased salt concentration having a stabilising effect on the protein compared to lower concentrations.[226] An interesting observation is the estimated hydrodynamic radius of silicatein-α under increasing salt concentration as it increases from 6.7 nm in 100 mM NaCl to 8.8 nm at 500 mM. (Figure 4.6B) Given the theoretical hydrodynamic radius for silicatein-α is ~3 nm [94] these values would suggest formation of an oligomeric conformation. In the experiments with TF-silicatein-α increasing salt raised the T_{agg} by 5 °C but has no effect on T_m. (Figure 4.6C). The estimated Rh also remained constant at a value of ~23 nm (Figure 4.6)

The effect of pH on melting temperature, aggregation onset and hydrodynamic radius was investigated to confirm protein stability and homogeneity. Samples containing silicatein-α and TF-silicatein-α in 50 mM Tris, 100 mM at pH 7.5 and pH 8.5 and 50 mM citrate 100 mM NaCl pH 3.5 were analysed. With the theoretical pI of silicatein-α and TF-silicatein-α calculated to be 6.73 and 5.13 respectively, pH 8.5 and pH 3.5 were chosen as they have acceptable divergence from the pI. The decision to use pH 7.5 was based on this being close to physiological pH and is usually deemed a neutral and ambient condition for biochemical reactions. The data recorded for T_m and T_{agg} for both silicatein-α and TF-silicatein-α show lower temperatures, of 36 ° C and 39 ° C respectively, for the proteins under pH 3.5 (Figure 4.7A; Figure 4.7C). Additionally, the R_h values are ~7 fold higher for silicatein-α and ~5 fold high for TF-silicatein-α at pH 3.5 suggesting the protein is aggregated under these conditions. (Figure 4.7B; Figure 4.7C).

Table 4.10. Table of polydispersity index values of silicatein-α and TF-silicatein-α with varying pH levels at 22 °C.

	PDI silicatein-α	PDI TF-silicatein-α
pH 3.5	1.4	3.384
pH 7.5	0.764	0.98
pH 8.5	0.619	0.634

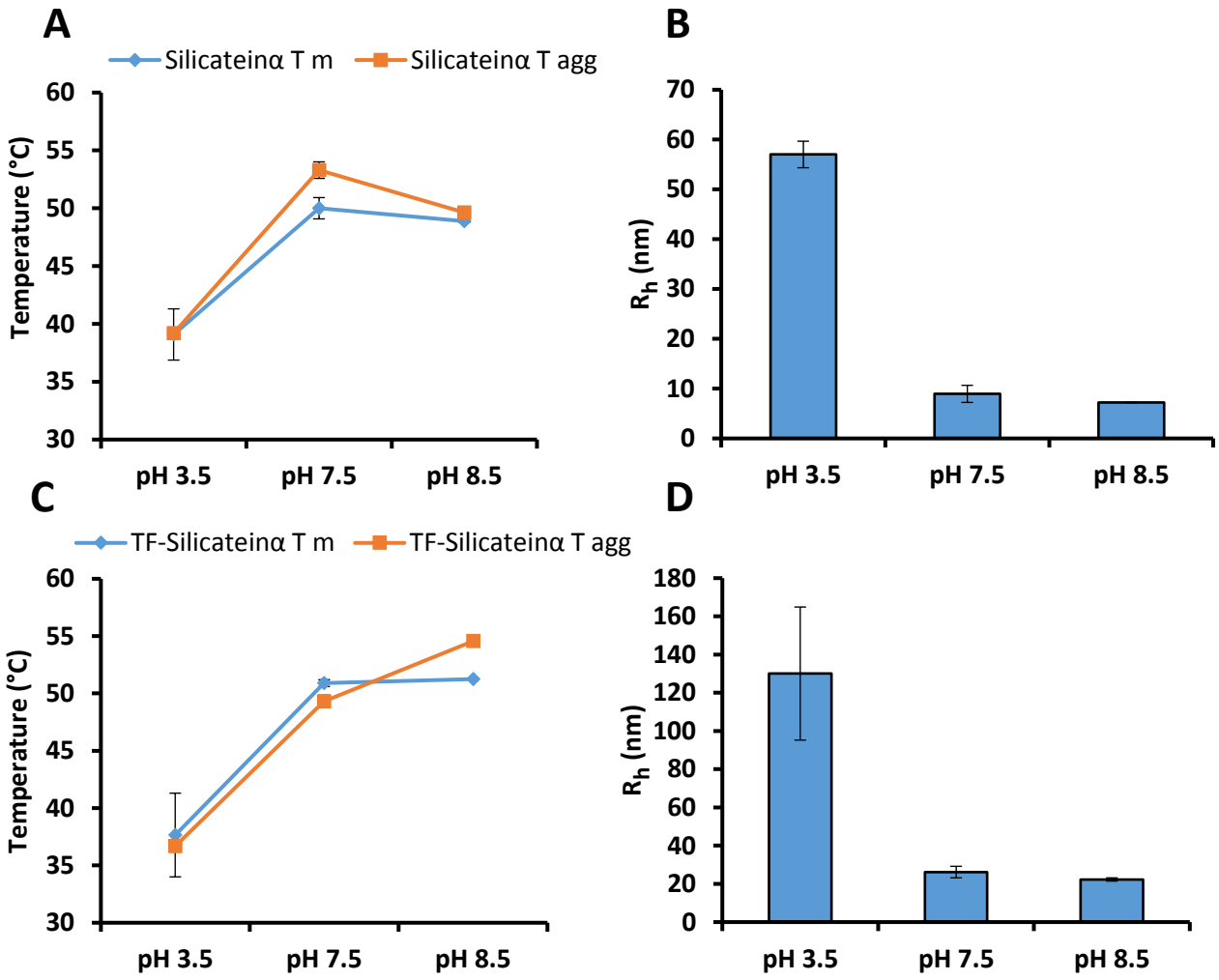


Figure 4.7 Light scattering experiments with varying pH levels of buffer (A) melting temperature and aggregation onset temperature of silicatein-α, (B) estimated hydrodynamic radius of silicatein-α (C) melting temperature and aggregation onset temperature of TF-silicatein-α, (D) estimated hydrodynamic radius of TF-silicatein-α.

This is further validated with PDI values >0.7 inferring a high level of polydispersity and broad distribution of particle size within the sample at that pH. (Table 4.10) As the T_m and T_{agg} and Rh values for silicatein-α and TF-silicatein-α (Figure 4.7 A-D; Table 4.10) are similar to those recorded under varying salt conditions (Figure 4.6 A-D) only PDI values for proteins at pH 8.5

suggest homogeneity (Table 4.10). Therefore, to further screen the effectiveness of buffer additives Arg + Glu in 50 mM Tris, 100 mM NaCl at pH 8.5 a thermal ramp and dynamic light scattering measurements were made. Previous results suggest homogeneity at a concentration of 100 mM Arginine and 100 mM Glutamate (Table 4.8) and also the positive results from the inclusion in the lysis buffer using in the purification of silicatein-α as described in Chapter 2.

Table 4.11. Table of polydispersity index values of silicatein-α and TF-silicatein-α in buffer pH 8.5 and pH 8.5 (Arg + Glu) at 22 °C.

	PDI silicatein-α	PDI TF-silicatein-α
pH 8.5	0.619	0.764
pH 8.5 (Arg +Glu)	0.24	0.229

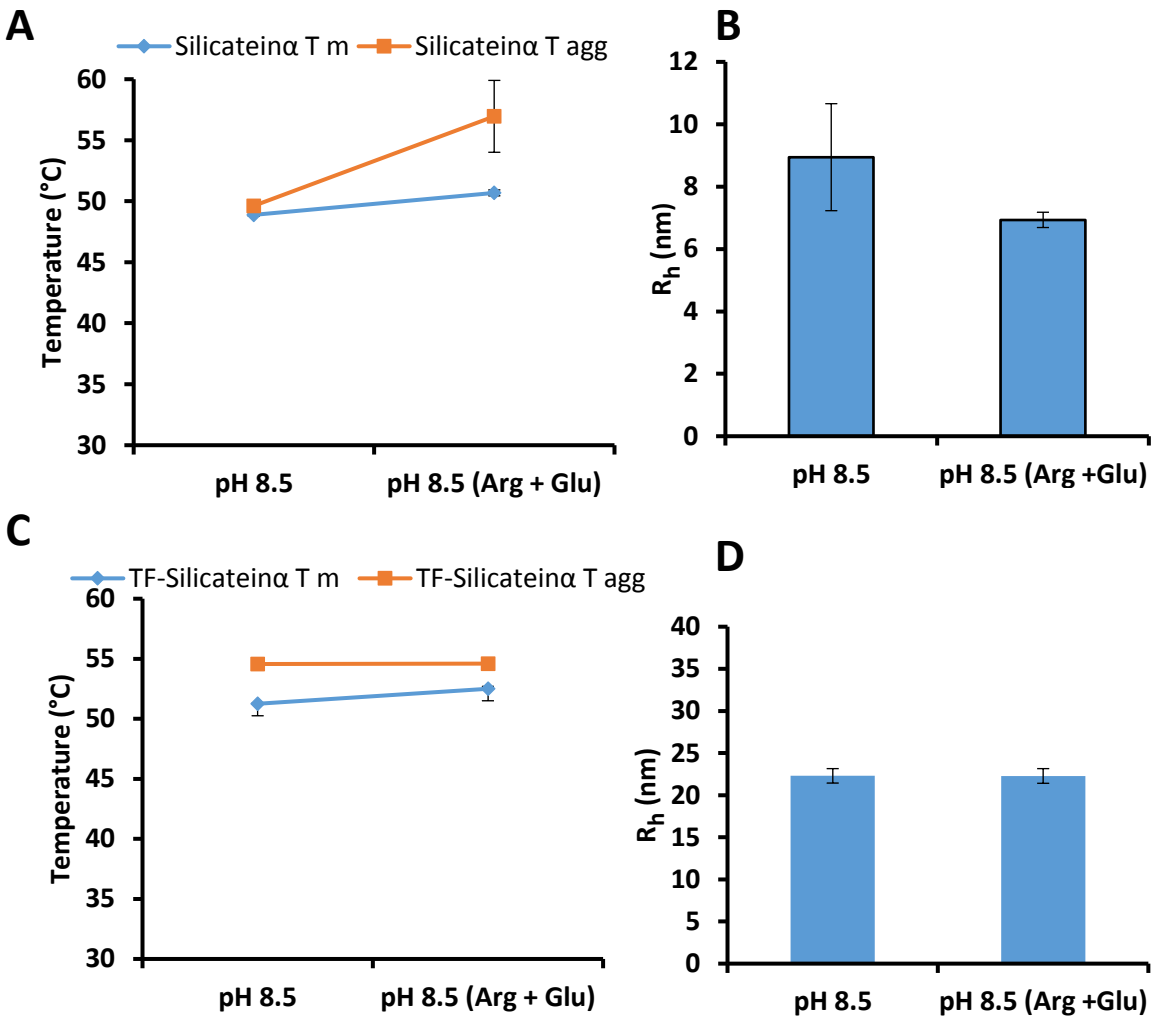


Figure 4.8 Light scattering experiments comparing samples with and without Arg and Glu. (A) melting temperature and aggregation onset temperature of silicatein-α, (B) estimated hydrodynamic radius of silicatein-α (C) melting temperature and aggregation onset temperature of TF-silicatein-α, (D) estimated hydrodynamic radius of TF-silicatein-α.

Experimental analysis revealed an increase in melting temperature and aggregation onset temperature with the addition of Arg + Glu for silicatein-α (Figure 4.8A) but no change was observed for TF-silicatein-α. (Figure 4.8C). A decrease of ~ 2 nm in R_h was also observed with the addition of buffer additives with silicatein-α (Figure 4.8B) but as before no change was observed for TF-silicatein-α. (Figure 4.8D) The results imply silicatein-α is more stable with the addition of the amino acid additives which also compliments the inclusion within the lysis buffer used in the purification process.

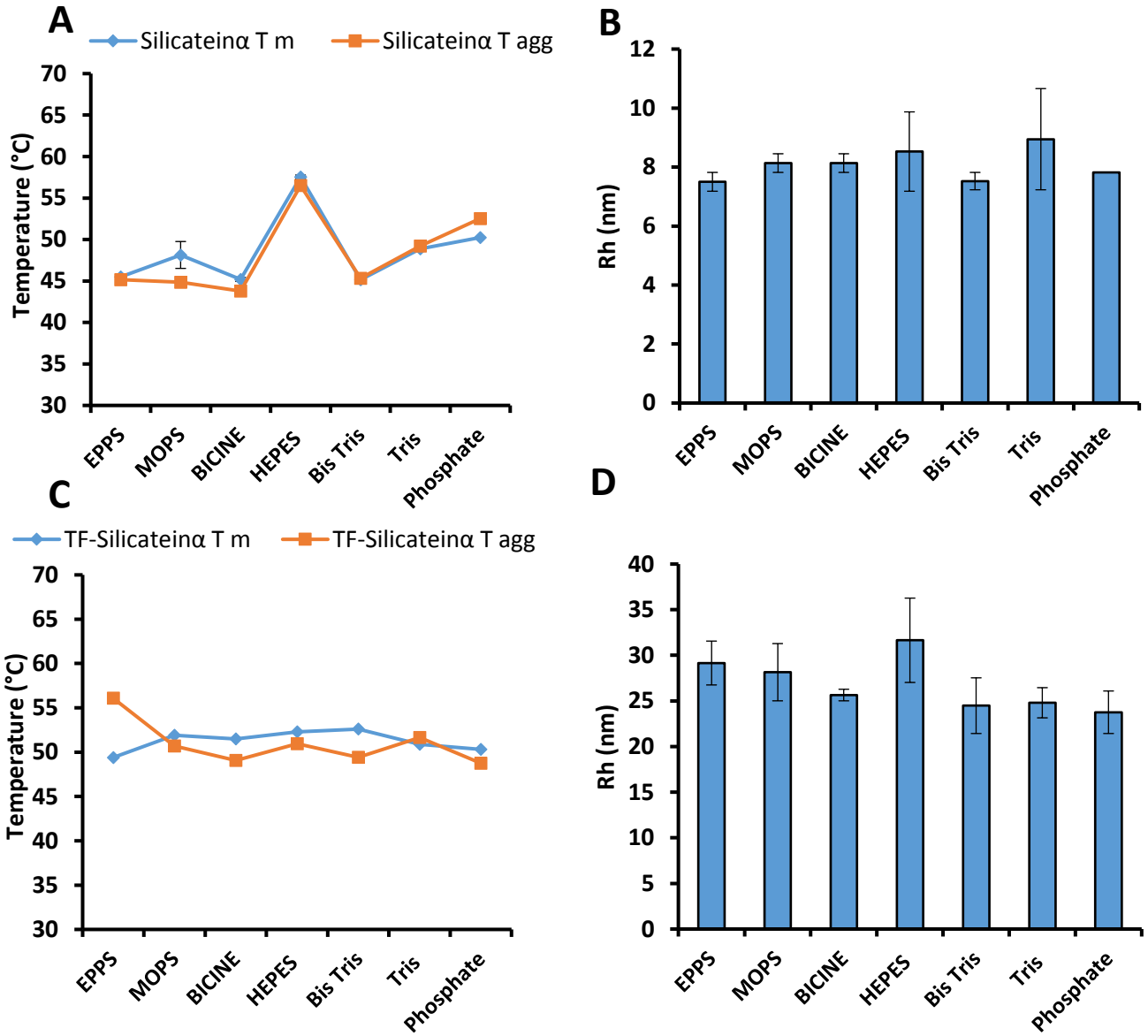


Figure 4.9 Light scattering experiments with various buffer conditions. (A) melting temperature and aggregation onset temperature of silicatein-α, (B) estimated hydrodynamic radius of silicatein-α (C) melting temperature and aggregation onset temperature of TF-silicatein-α, (D) estimated hydrodynamic radius of TF-silicatein-α.

Good's buffers have certain characteristics such as a working pH range between 6.0 and 8.0, non-toxic and high solubility which makes their application fundamental for many biological and biochemical research purposes.[227] However, it is often a trial and error process to identify the buffer that will provide optimum conditions for the protein of interest. As this method offers the advantage of high-throughput screening, seven Good's buffers were chosen, EPPS, MOP, BICINE, Bis-Tris, Tris, Phosphate and HEPES to assess protein stability and homogeneity. Further analysis of these various buffers revealed HEPES raised T_m and T_{agg} by around 10 °C

for silicatein-α (Figure 2.1310A) but the PDI is 3.358 (Table 4.12) which imply a very low level of homogeneity as previously mentioned a value >0.7 indicates the particle size has a broad distribution. The estimated Rh values for silicatein-α (Figure 2.13B) could be indicative of a dimer or trimer if the theoretical value is 3 nm[94]. A notable observation for TF-silicatein-α is the increase of ~5 °C with buffer containing EPPS (Figure 2.13C) although the Rh values for all buffers with TF-silicatein-α (Figure 2.13D) are still high in comparison to R_h of Thyroglobulin, a globular protein with a molecular weight of 670 kDa and Rh of 8.71 nm.[225] However, the estimates must be considered carefully, as they are derived using an empirical curve of known globular proteins and their hydrodynamic radius, which means it cannot be readily applied if the protein is not globular. As the conformation of the fusion protein is still unknown, It is worth considering the possibility of the fusion protein being highly elongated or having a highly ordered oligomeric structure due to native silicatein-α preference to self-assemble[133,138] The DLS measurements of the buffers containing Tris, EPPS and BIS-TRIS produced a single peak and PDI values are below 0.7 (Table 4.12) therefore these samples will be sent for further analysis to assess whether they can be progressed to crystallography trials.

Table 4.12. Table of polydispersity index values of silicatein-α and TF-silicatein-α in buffer pH 8.5 and pH 8.5 (Arg + Glu) at 22 °C.

	PDI silicatein-α	**PDI TF-silicatein-α**
EPPS	0.369	0.988
MOPS	3.094	1.708
BICINE	3.023	0.094
HEPES	3.358	0.669
Bis Tris	0.681	0.187
Tris	0.240	0.191
Phosphate	2.326	1.057

4.5.3 Size Exclusion Chromatography with Multi-Angled Light Scattering

Multi-angled light scattering (MALS) is a technique applied to determine absolute molecular mass of proteins in solution based upon the scattering light signal at different angles relative to the incident laser light source. These measurements can be useful in assessing oligomeric state and homogeneity of the protein sample in solution. Fractions of the sample can be

analysed by light scattering individually when it is used in combination with size exclusion chromatography.

Prior to submission of samples for light scattering analysis, SEC was used to confirm no co-eluting contaminants were present following purification by IMAC. Protein concentrations in the range of 1 mg/mL to 0.2 mg/mL were loaded onto the SEC column. However, the results show a concentration dependent aggregative state with protein concentrations above 0.5 mg/mL, eluting in the void volume (Figure 4.10A and Figure 4.10B). Therefore, further analysis of TF-silicatein-α in different buffers by MALS was conducted using a protein concentration of 0.3 mg/mL.

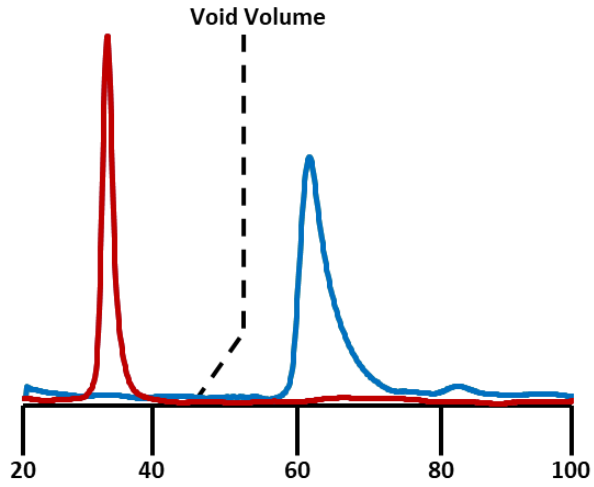


Figure 4.10. Size exclusion chromatography (SEC) chromatogram. TF-silicatein-α 1 mg/mL eluting in the void volume (red) TF-silicatein-α 0.3 mg/mL eluting in the void volume (blue).

Experimental analysis of MALS results for TF-silicatein-α in various buffers and pH were inconclusive as the protein eluted in the void volume with a calculated molecular weight on the order of magnitude of 10^6 Daltons, suggesting aggregation. This could be due to handling error, aggregation due to protein unfolding under the heat of the laser source or insufficient chromatographic conditions that requires further optimisation.

TF-silicatein-α, with amino acid additives, Arg + Glu from the second peak in SEC chromatogram (Figure 2.12B) were sent for MALS analysis. The molecular mass was estimated to be 48.9 kDa which could be Trigger factor alone, but there was no evidence of the remaining portion of the fusion protein, silicatein-α, in the SEC chromatogram or from MALS data. (Figure 4.11) Further investigation of this would need to be conducted.

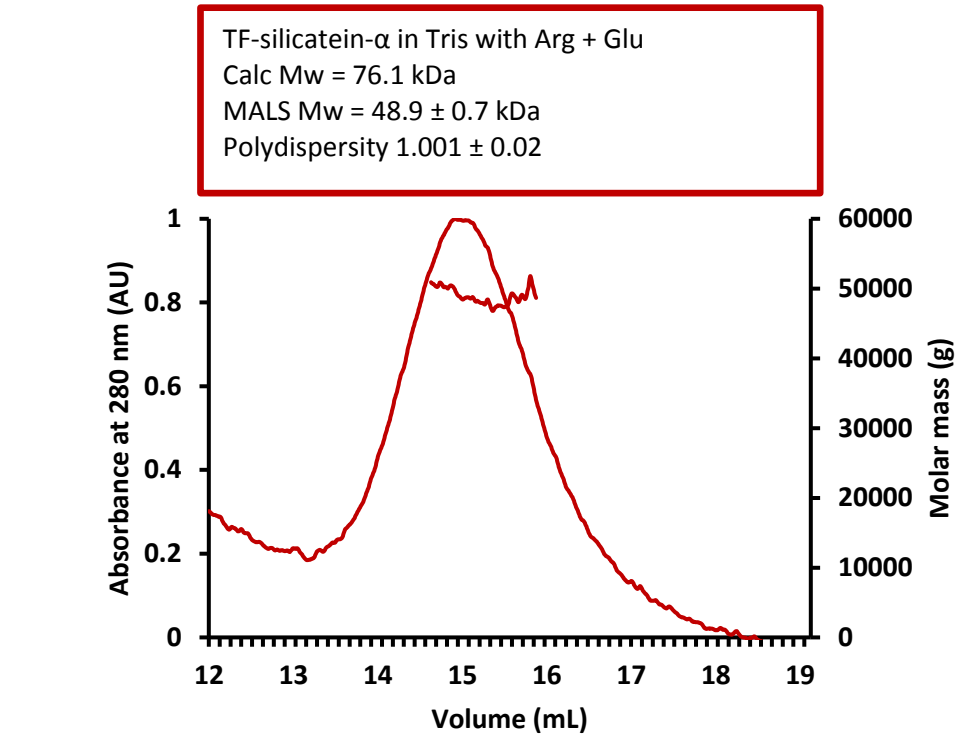


Figure 4.11 Multi angled light scattering data (MALS) with TF-silicatein-α in 50 mM Tris 100 mM NaCl 50 mM Arg + Glu pH 8.5. Estimated molecular weight and MALS calculated molecular weight are shown along with polydispersity index

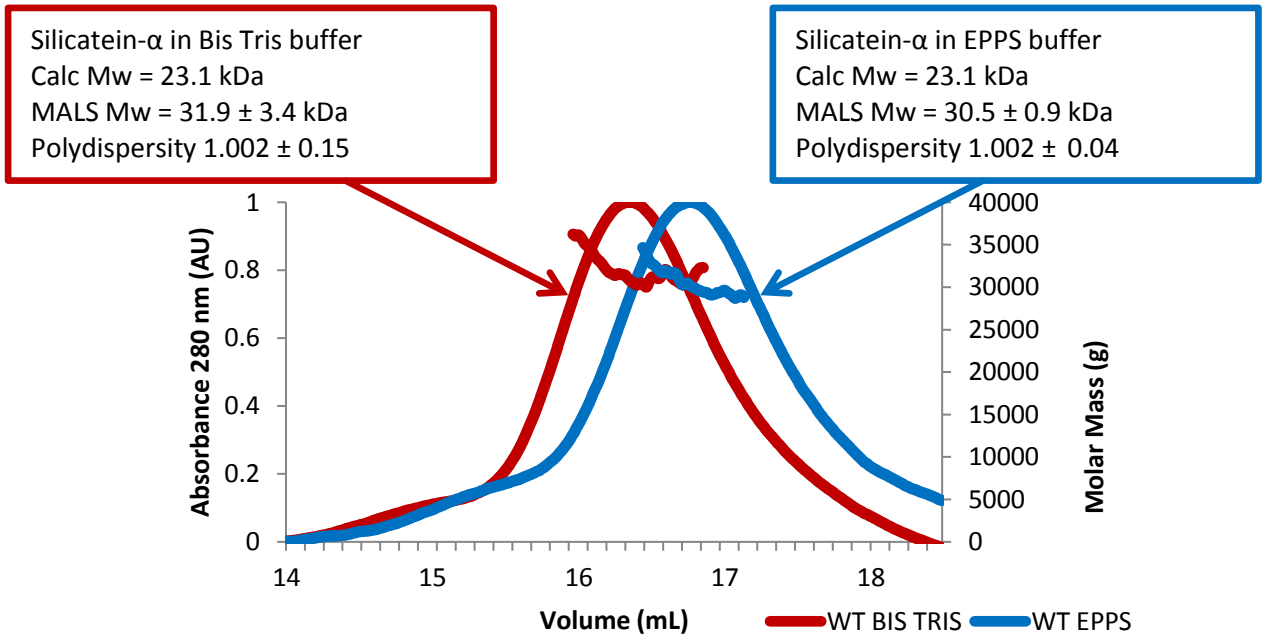


Figure 4.12 Multi angled light scattering data (MALS) with silicatein-α in 100 mM EPPS (blue) or 50 mM Bis-Tris (red) ,100 mM NaCl 50 mM Arg + Glu pH 8.5. Estimated molecular weight and MALS calculated molecular weight are shown along with polydispersity index

MALS data could not be obtained for silicatein-α in Tris due to protein precipitation in the sample, however, analysis of MALS data of silicatein-α in Bis-Tris and EPPS buffer with amino acids additives, Arg +Glu produced estimated molecular masses of ~30 kDa. (Figure 4.12). This value is ~7 kDa higher than the theoretical molecular weight of 23.1 kDa which could be due to a certain margin of error of the data output.

However, the data suggests the protein could be monomeric and stable with these buffers and offer two potential conditions that can be progressed forward for further analysis in preparation for crystallography trials.

4.6 Conclusions

This chapter aimed to further expand the structural knowledge of silicatein-α through biophysical tests. Following the development and optimisation of a reliable and reproducible expression and purification method to isolate an acceptable yield of protein, described in Chapter 2, and confirmation of the catalytic activity, described in Chapter 3 further analysis could be applied.

Biophysical tests were required to assess the structural integrity and stability of the enzyme under general assay conditions and screen a variety of buffer conditions to find the optimum conditions for further analysis by crystallography trials. CD confirmed that there was no aggregation in the wild type and fusion protein under the conditions previously published. Additionally, mutations of the active site residues did not affect the overall structure as the same trace was observed. The secondary structure content estimation for each protein shows close values (~5%) to the calculated theoretical values, obtained from crystallographic data in the PDB for trigger factor and homologous chimeric protein cathsilicatein.

DLS data for silicatein-α and TF-silicatein-α showed a large hydrodynamic radius, which could be attributable to the conformation of the protein being elongated and not globular or a highly ordered oligomeric form. The MALS measurements were inconclusive for the fusion protein even at a variety of concentrations tested. This was contrasting to SEC data which suggested a concentration dependent factor that influenced elution in or after the void volume. As this was not reflected in the MALS data for TF-silicatein-α, this may be due to handling error, aggregation due to protein unfolding under the laser source or insufficient chromatographic conditions that requires optimising. However, buffer screening experiments indicated that TF-silicatein-α and silicatein-α had a low level of polydispersity in the conditions used for the biochemical assays and this value was reduced further upon the addition of L-arginine and L-glutamate. The screening also aided in identifying two different buffers in which the silicatein-α displayed an acceptable level of homogeneity. MALS measurements for silicatein-α under these two different buffer conditions were successful, highlighting potential to progress to crystallography trials.

5. Chapter 5: Investigation of substrate scope, thermostability and enantioselectivity of silicatein-α with a coupled biotransformation using ADH.

The production of fine chemicals and pharmaceuticals is often complex and consists of many intermediate stages before the desired, final product is synthesised. The utilisation of protecting groups is common practice to ensure chemoselective modification of a particular functional group in the presence of others. During multi-step chemical synthesis involving polyhydroxylated compounds or intermediates, protection of a specific hydroxy group has been achieved using silyl ethers protecting groups.[28,30,54,55,228] Selective silylation and subsequent desilylation of specific alcohols or phenols can prevent undesirable and adverse side reactions or products. As discussed in Chapter 1, a plethora of silyl ether protecting groups exist which allow exploitation of the variety of electronic and steric properties that aid in selective deprotection.[28,54,55,228]

5.1.1 Traditional methods used for silyl protection

In many synthetic chemical reactions involving multi-hydroxy bearing compounds enantio- and site selective silylation and de-silylation of specific hydroxyl groups is often required.[43,44] The traditional methods of silylation used silyl chloride (Figure 5.1) and silyl triflates to protect hydroxyl functional groups.[44–47] However, these halogenated compounds have an adverse environmental impact due to their corrosive properties and unwanted side products produced in the reaction. Therefore, there is a requirement for approaches that better conform to the principles of green chemistry.[57,79,80] Several groups have reported on enantioselective silyl protection of alcohols using very different and improved methods to circumvent the use of silyl chloride and silyl triflates, with improved separation and kinetic resolution of enantiomers.[28,30,54,55] Hydrosilanes have been utilised in the silylation of alcohols both with metal catalysts such as Rh (II) carboxylates[41,83] and Lewis acids such as $InBr_3$[57] with the benefit of only hydrogen as the by-product from the reaction. Other groups have reported kinetic resolution of racemic mixtures, in particular secondary alcohols, through stereoselective silylation using a copper catalyst and hydrosilanes.[84–86]

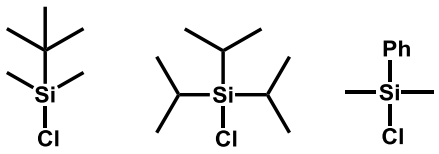


Figure 5.1 An example of silylating agents used for silyl protection

5.1.2 Traditional methods used for silyl deprotection

Approaches for subsequent selective silyl deprotection within multi-step organic synthesis have been extensively reviewed.[54,55,228] This generally considers that under certain reaction conditions and in the presence of other hydroxyl groups, selective desilylation can be achieved based upon electronic and steric effects of the substituents of the silyl protecting group around the silicon atom. The steric bulk of the substituents can, therefore, decrease silyl ether cleavage under both acidic and basic conditions.[28,54,55] Additionally, in the case of the same protecting group used for more than one hydroxyl group in the compound, the less sterically hindered will be cleaved first. Selective deprotection of bulkier silyl protecting groups in the presence of labile and smaller silyl protecting groups can, however, be manipulated through electronic effects. As discussed in Chapter 1, the rate of hydrolysis of silyl ethers under acidic conditions can be decreased with electron withdrawing groups on the silicon atom while under basic conditions electron donating groups will decrease the rate of hydrolysis. This step in organic synthesis also has limitations, being the use of harsh, expensive and environmentally adverse reagents and chemical catalysts in addition to problematic removal of the protecting group depending upon the reaction conditions.[28,54,55] The utilisation of enzymes as biocatalysts to mediate silylation and subsequent desilylation would be desirable as it would provide a more environmentally sustainable method.

5.1.3 The application of enzymes as biocatalysts in organic media with silyl protecting groups

As discussed in Chapter 1, chemical synthesis mediated by enzymes and conducted in organic solvents offers many benefits. This area has been widely studied because of enhanced features such as increased thermostability, easy enzyme recovery following the reaction and

the reaction equilibria can be directed to favour the reverse reaction.[229,230] Another advantage to conducting biotransformations in organic solvent systems is that it offers the potential to widen substrate scope of the enzyme opening up a lot more potential reactions.[112,229–231] Taylor *et al.* screened 80 enzymes, predominantly hydrolase enzymes, to assess their ability to mediate condensation and hydrolysis of alkoxysilanes in various solvent systems. Specifically they investigated enzyme-mediated etherification and trans-etherification of silyl protection groups with a primary alcohol (octan-1-ol) using a variety of hydrolases whose native function is not inclusive of silicon oxygen bond formation.[51,52] As silicatein-α was previously postulated to catalyse the polycondensation of biosilica in its native surroundings research within our group was aimed at developing this enzyme as a potential silylation biocatalyst.[53,78] Recently published experimental data revealed heterologously produced silicatein-α was able to catalyse silyl etherification and transetherification of a primary alcohol with three model silanols (TMS-OH, TES-OH and DMPS-OH) and ethoxysilanes (TMS-OEt, TES-OEt and DMPS-OEt) in *n*-octane.[53,78] Regioselectivity of silyl etherification mediated by silicatein-α was also investigated using hydroxylated substrates with both aliphatic and aromatic alcohol functional groups. This data showed a clear preference and selectivity in enzymatic silyl protection catalysed by silicatein-α of aromatic hydroxy function groups in the substrates that were screened.[53,78] The results produced from the work involving silicatein-α is of particular interest as it shows application of the enzyme as a potential in catalyst in silyl protection. To date investigations have not been conducted to screen the enantioselectivity of silicatein-α.

5.1.4 Oxidoreductases

Oxidoreductases are EC Class I enzymes ubiquitous in all kingdoms of life. Sub-classification of enzymes within this family is dependent upon the substrate and the type of reaction they catalyse. Commonly, they function in the transfer of electrons from a reductant (the donor) to an oxidant (the acceptor) in two half reaction, the reduction and the oxidation. Dehydrogenases are a subclass of the superfamily Oxidoreductases, that mediate oxidation through exchange of one or more proton (hydride ion) using a cofactor.[73]

5.1.5 Alcohol dehydrogenase

Alcohol dehydrogenases (ADH) have been widely studied since its isolation from yeast in 1937 by Negelein and Wulff,[232,233] and therefore, extensive literary knowledge exists for this subclass of oxidoreductases. They are ubiquitously found in many organisms within the kingdoms of life with many ADH genes and proteins have been characterised. With a wealth of structural and mechanistic data available, both native and engineered ADH enzymes have been applied to study the conversion of many aromatic and aliphatic primary and secondary alcohol substrates into their respective aldehyde and ketone products through the reduction of cofactor NAD(P).[234] Resultantly, ADH derived from many different organisms have been utilised as biocatalysts to synthesise chiral compounds including its use in organic solvents and as an important factor to determine the presence of ethanol in clinical samples.[232,235,236]

5.1.6 Enzymatic reduction of ketones and oxidation of alcohols to obtain chiral compounds

The pharmaceutical industry uphold the importance of chirality in the production of active pharmaceutical ingredients and their intermediates as a enantiomerically pure compounds (EPC) can be fundamental to the safety and efficacy of the active product. Chiral synthesis can be mediated by both chemical catalysts and biocatalysts. However, biocatalytic strategies have been described as advantageous in comparison to stoichiometric routes due to selectivity and enhanced conversion rates, reduced reliance on solvents and generally operation under more ambient reaction conditions.[237] The development of multistep synthetic chemical reactions mediated by enzymes has increased significantly as more enzymes are characterised and also owing to innovations in protein engineering. Extensive research has been conducted into employing ADH for the enzymatic reduction and oxidation of ketones and alcohols into their respective products with industrial applications or currently used in such processes with isolated enzymes or whole cell systems.[238–240] With the expansion in genomic sequencing and protein engineering many enantioselective ADH have been uncovered and utilised for racemic resolution.[239,241–243] Additionally, development of sustainable synthesis methods have been achieved through coupled enzymes reactions that facilitate co-factor regeneration.[239,242–247]

Utilising the inherent properties of these enantiomeric enzymes, *R*-specific ADH from *Lactobacillus kefir* (*L.kefir*) and *S*-specific ADH from to manipulate the desired product. This was achieved by stepwise racemic resolution with the required enzymes as illustrated by (Figure 5.2) to produce an EPC of *R*-Phenylethanol to 100% yield.[241] Asymmetric ketone reduction by ADH has been demonstrated in the synthesis of an important intermediate in the production of NK-1receptor antagonist (Figure 5.3). (*S*)-3,5-bistrifluoromethylphenyl ethanol was synthesised on a pilot scale with yieds of (>90 %).[248] A further example of enzymatic ketone reduction used on an industrial scale is conversion of (2,5)-hexanedione to (2*R*,5*R*)-hexanediol by *L.kefir* ADH with yields of >99 % (Figure 5.4). This study was particularly interesting as it showed this enzyme to be diastereoselective as well as enantiomeric for this substrate.[249]

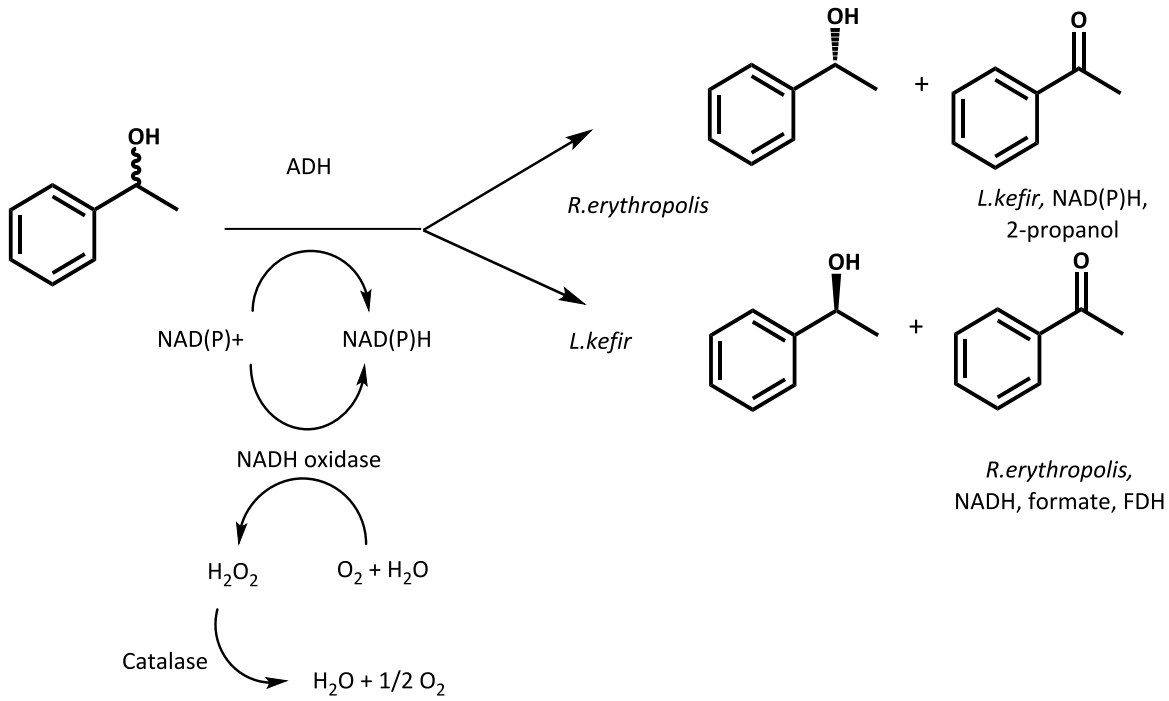


Figure 5.2 Batchwise enzymatic alcohol oxidation with two enantioselective ADH with enzyme cascade

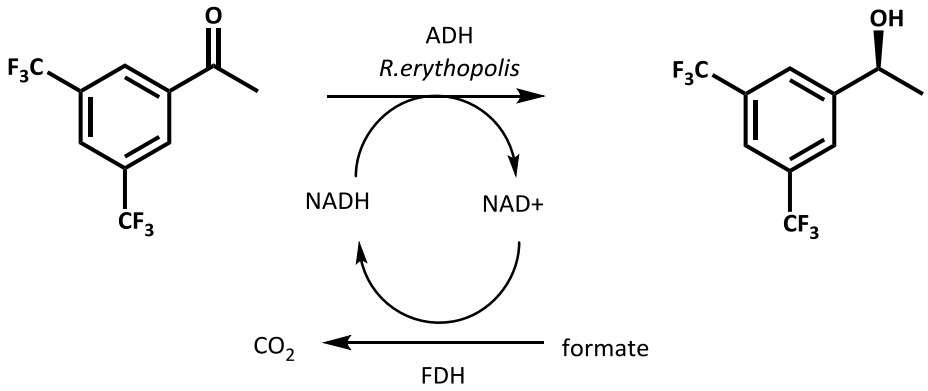


Figure 5.3 Biotransformation of (*S*)-3,5-bistrifluoromethylphenyl ethanol with ADH pH 7 phosphate buffer at 30 °C using FDH as a regeneration system

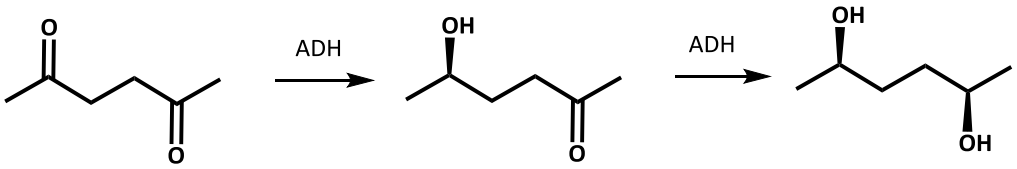


Figure 5.4 Enzymatic ketone reduction of (2,5)-hexanedione to (2*R*,5*R*)-hexanediol by *L.kefir* ADH

5.1.7 The application of Alcohol Dehydrogenase in organic solvent

Over the last few years a growing area of interest has been the use of tandem or coupled enzymatic systems or cascades for multi-step synthesis. This circumvents the use of harsh chemicals, with the possibility of running reactions under more ambient conditions and also has the potential benefits of being more regio- stereo- and enantioselective.[250–252] As more enzymes are structurally and mechanistically characterised, their application to a specific process have become more widespread.[79,253,254] Additionally the prospect of conducting biotransformations via non-conventional methods such as non-aqueous milieu has opened up even more possibilities. Early research in this area focused mainly on hydrolases, as a relatively simple system often requiring no cofactor. Grunwald *et al*. pioneered the use of alcohol dehydrogenase for asymmetric oxidoreductions in organic solvent.[111]. This is a more complex task in comparison to the application of hydrolases as they require a cofactor for

catalytic activity. To date ADH has been employed in numerous biotransformations conducted in solvent systems both biphasic and monophasic and also tandem enzyme systems.[103,108,255–257]

5.2 Aims and objectives

To date the experimental data produced within our group has revealed silicatein-α to be a promising candidate for use as a biocatalyst in organosilicon chemistry. The work in this chapter, therefore, aims to expand our current knowledge with regards to substrate scope, the catalytic environment in which it functions and investigate the possibility of a coupled biotransformation with another enzyme.

The thermostability of the enzyme in organic media will be investigated by increasing the reaction temperature to 90 °C for etherification and transetherification using model substrates octan-1-ol and the silanol TES. Further experiments will be conducted to assess enantioselectivity through repetition of the reactions using TES with *R*/*S*-octanol and *R*/*S*-phenylethanol at 75 °C and 90 °C.

An investigation of the hydrolytic function of the enzyme in silyl deprotection reactions in aqueous buffer will also be conducted using a model system initially. Experiments will be conducted to determine if silyl deprotection of triethyl(octyloxy)silane can be mediated by silicatein-α in aqueous media under ambient conditions. Work will then be aimed at progressing to the use of substrates that will allow for the assessment of enantioselectivity of the enzyme.

Finally, biotransformations using a coupled enzyme system will be investigated using ADH as a potential catalytic partner in silyl protection and deprotection in one system. This will be through racemic resolution with an *R*-specific ADH or specific silylation of *S*/*R* enantiomers (Figure 5.5). Initial work will be aimed at assessing the viability of ADH in organic solvent and aqueous media whilst simultaneously screening the aforementioned secondary aliphatic and primary aromatic alcohols.

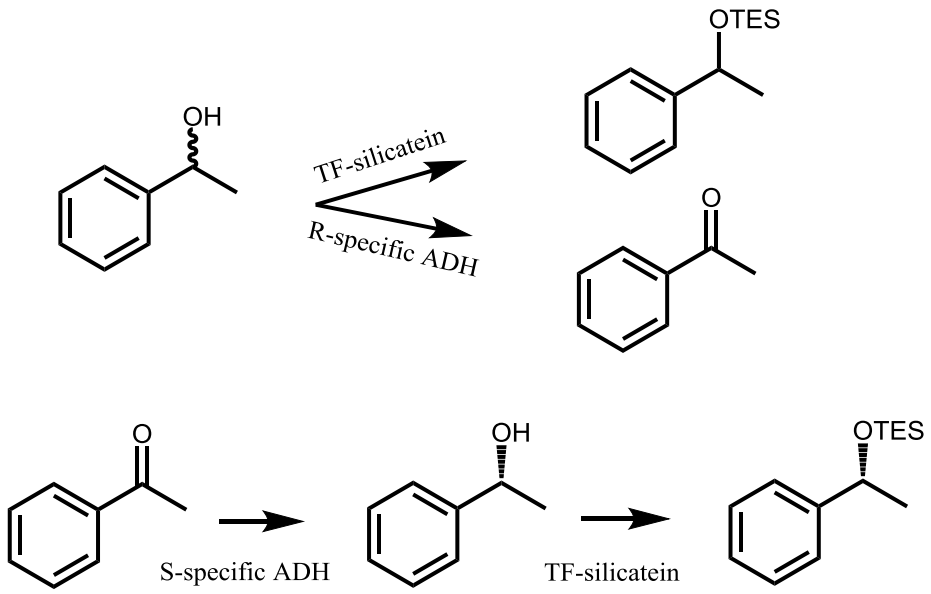


Figure 5.5. Proposed coupled enzyme system for silylation using TF-silicatein-α and ADH.

5.3 Results and Discussion

5.3.1 Thermostability of silicatein-α as a biocatalyst in organic solvent

Previous results suggested that the rate of reaction and production conversion to be significantly greater at the higher temperature of 75 °C for both etherification and transetherification using TMS-OH and 1-octanol. Therefore, initial experiments were completed at this temperature and under the same conditions to validate these results and as a quality control measurement for the stock of lyophilised enzyme. (Figure 5.6; Figure 5.7; Figure 5.8; Figure 5.9). Experimental data confirmed that silicatein-α is an efficient and stable biocatalyst, increasing the rate of etherification and transetherification reactions in comparison to the control sample where the enzyme was omitted. The enzyme's stability and ability to function within this environment is thought to be attributable to the hydrophobic nature of the solvent system and the amino acid composition on the exterior of the protein in addition to their electrostatic interactions.[133] The monolayer of essential water molecules that contribute to conformational flexibility and activity of some enzymes is retained in more hydrophobic solvents than hydrophilic solvents. The presence of disulphide bonds and electrostatic interactions and aid in preservation of the protein's functional three dimensional structure.[113,258,259] In the case of silicatein-α and its sequence homology to cathepsin L, it is has been postulated to have three disulphide bridges[61] which could aid in the stability of the protein in organic media.

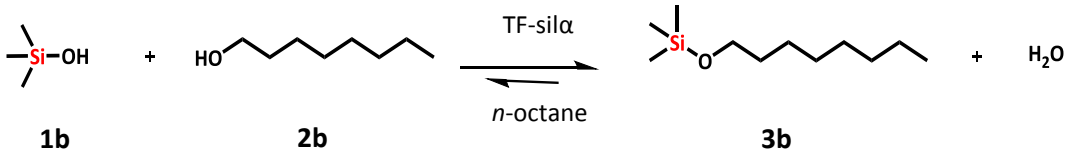


Figure 5.6 Etherification catalysed by TF-silα at 75 °C in octane with trimethylsilanol (1b) and 1-octanol (2b) to form trimethyl(octyloxy)silane (3b)

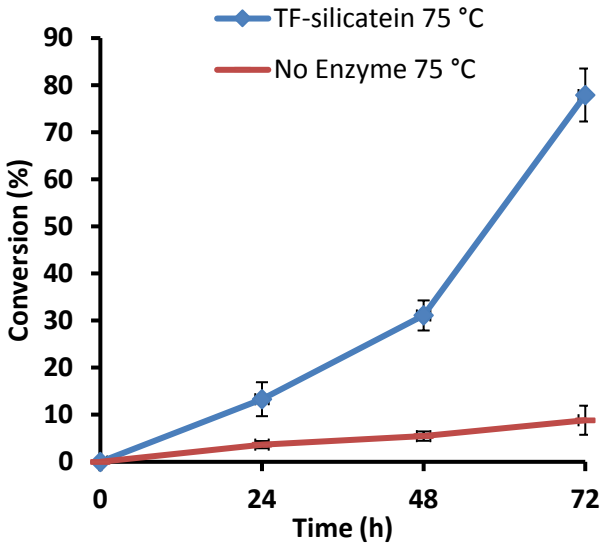
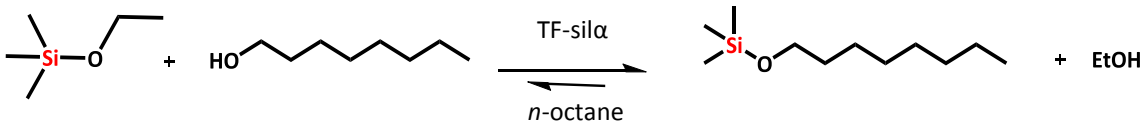


Figure 5.7 The percentage of product formation of trimethyl(octyloxy)silane (3b) from trimethylsilanol (1b) and 1-octanol (2b) in *n*-octane at 75 °C analysed by GC-MS



4b **2b** **3b**

Figure 5.8. Trans-Etherification catalysed by TF-silα at 75 °C in octane with trimethylsilane (4b) and 1-octanol (2b) to form trimethyl(octyloxy)silane (3b)

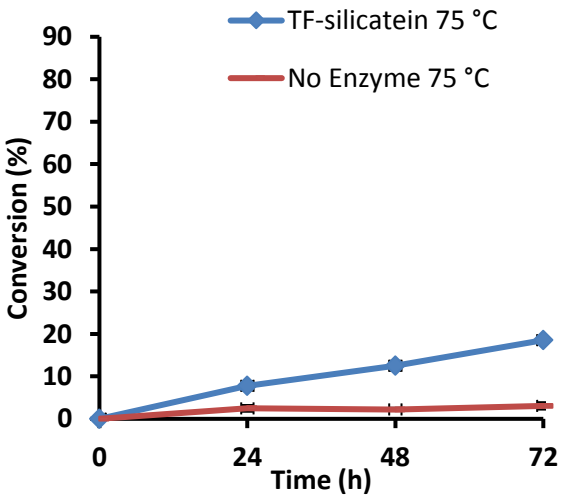


Figure 5.9 The percentage of product formation trimethyl(octyloxy)silane (3b) from trimethylsilane (4b) and octanol (2b) in *n*-octane at 75 °C analysed by GC-MS

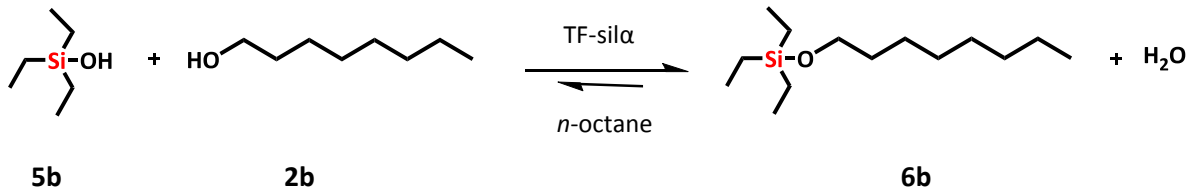


Figure 5.10 Etherification catalysed by TF-silα at 75 °C and 90˚C in octane with triethylsilanol (5b) and 1-octanol (2b) to form triethyl(octyloxy)silane (6b)

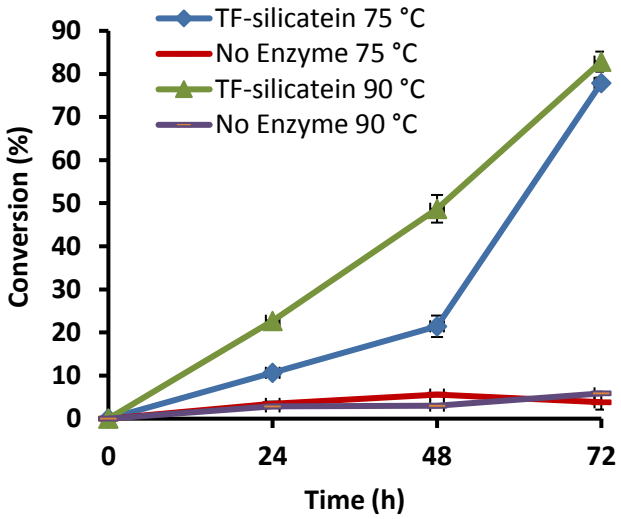


Figure 5.11 The percentage of product formation from triethylsilanol (5b) and octanol (2b) in n-octane at 75˚C and 90˚C to produce triethyl(octyloxy)silane (6b) analysed by GC-MS

Due to the presence of an anomalous peak in the GC-MS chromatogram thought to be dimerisation of TMS-OH, as observed by Taylor *et al.*,[51] further experiments, therefore, were conducted using TES-OH and model alcohol octanol. To test the thermostability and activity of silicatein-α, etherification and transetherification reactions in n-octane were repeated at an increased temperature of 90 °C in addition to 75 °C (Figure 5.10; Figure 5.11; Figure 5.12; Figure 5.13).

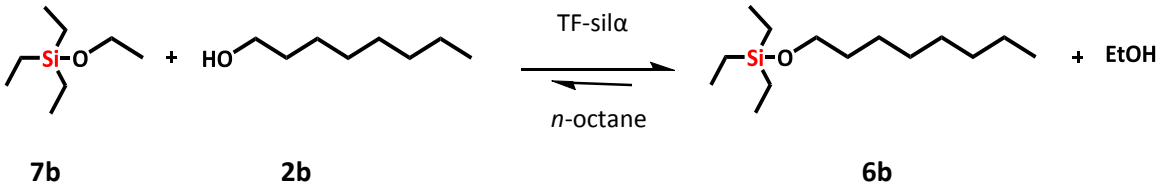


Figure 5.12 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in n-octane with triethylsilanol (7b) and 1-octanol (2b) to form triethyl(octyloxy)silane (6b)

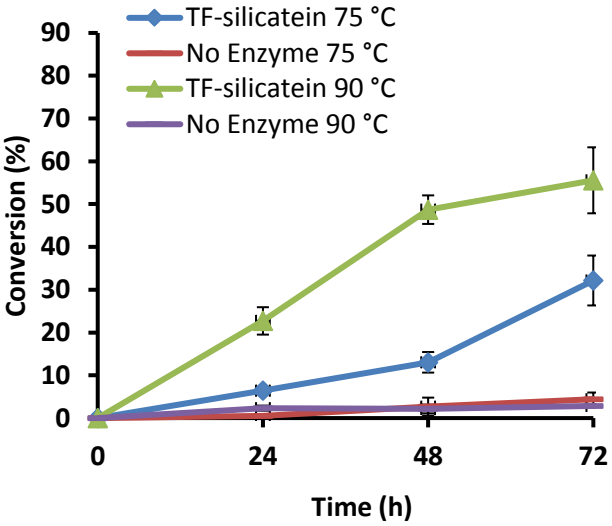


Figure 5.13 The percentage of product formation, triethyl(octyloxy)silane (6b), from triethylsilane (7b) and octanol (2b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS

The experimental data from the etherification and trans-etherification reactions with **2b** reveals a trend in which the percentage of product conversion is higher at 90 °C at the end time point and each time point taken in between (Figure 5.12; Figure 5.13). The control sample containing no enzyme does not appear to follow this trend; therefore, TF-silicatein-α

displays good thermostability at the higher temperature. This could be due to the effects of the anhydrous solvent environment maintaining the most effective three dimensional structure that would allow substrate access to the active site of the enzyme.[113,258,259] The percentage of product conversion is much higher for the etherification reaction which is in agreement with published data and has been described to be an effect of the side product of the trans-etherification reaction, ethanol, possibly inhibiting turnover or affecting the equilibrium (Table 5.1).[53,78] In both cases the samples containing TF-silicatein displays a high percentage of product conversion in comparison to the control implying the enzyme is catalysing the reaction.

Table 5.1 Percentage product conversion at 72 hours from etherification and trans-etherification reactions at 75 °C and 90 °C with 1-octanol

Substrate 1	Substrate 2 [a b]	Abbreviation	Product	Time (hrs)	Temperature (°C)	Conversion (%)
2b	1b	TMS-OH	3b	72	75	77.9 ± 5.6
2b	4b	TMS-Et	3b	72	75	18.5 ± 0.3
2b	5b	TES-OH	6b	72	75	77.9 ± 1.3
2b	5b	TES-OH	6b	72	90	82.9 ± 2.3
2b	7b	TES-Et	6b	72	75	32.1 ± 5.8
2b	7b	TES-Et	6b	72	90	55.6 ± 7.7

[a] 5 mol. equiv. of substrate used relative to substrate 1
[b] TMS = trimethylsilyl, TES = triethylsilyl

5.3.2 Enantioselectivity of silicatein-α in organic solvent

To investigate whether TF-silicatein-α has a preference between *R* or *S* enantiomers four substrates were chosen for enantioselective screening purposes, *R*/*S*-2-Octanol and *R*/*S*-1-phenylethanol. The potential substrates depicted different features, such as a long aliphatic chain and an aromatic ring structure.[53] The reactions were conducted over 72 hours at 75 °C and 95 °C and time points were taken every 24 hours and analysed by GC-MS.

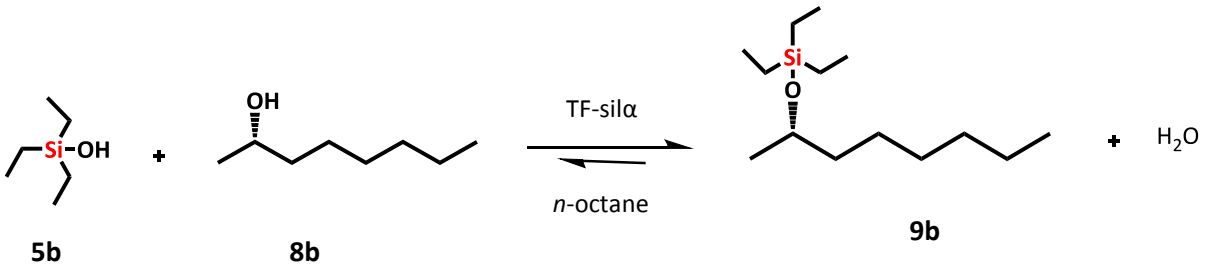


Figure 5.14 Etherification catalysed by TF-silα at 75 °C and 90°C in n-octane with triethylsilanol (5b) and (*S*)-2-octanol (8b) to form (S)-triethyl(octan-2-yloxy)silane (9b)

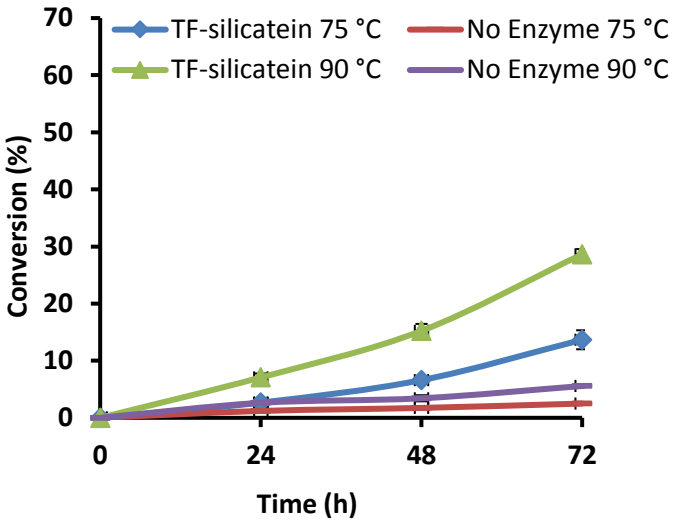


Figure 5.15 The percentage of product formation, (*S*)-triethyl(octan-2-yloxy)silane (9b), from triethylsilanol (5b) and (*S*)-2-octanol (8b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS

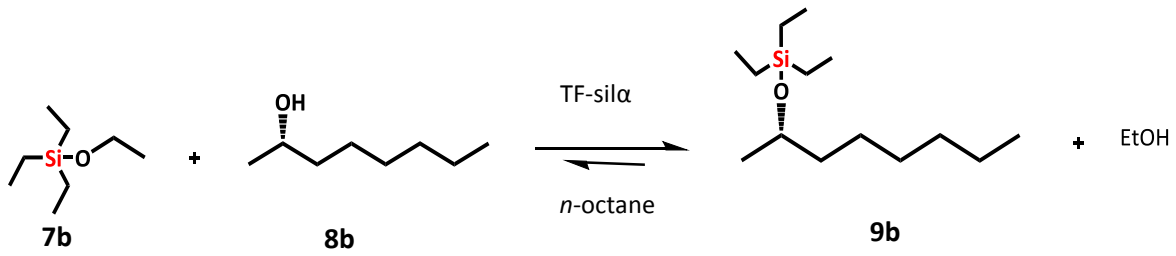


Figure 5.16 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in n-octane with triethylsilane (7b) and (S)-2-octanol (8b) to form (S)-triethyl(octan-2-yloxy)silane (9b)

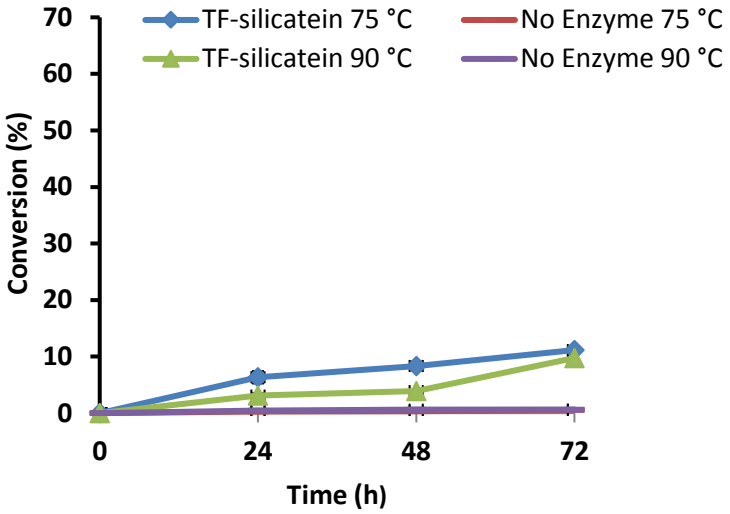


Figure 5.17 The percentage of product formation, (*S*)-triethyl(octan-2-yloxy)silane (9b), from triethylsilane (7b) and (*S*)-2-octanol (8b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS

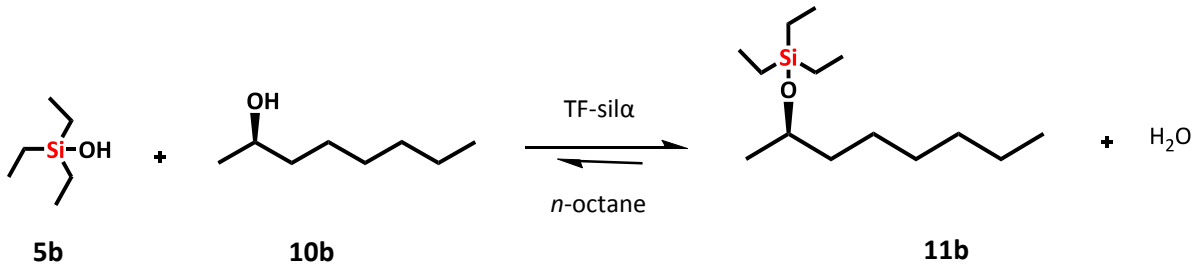


Figure 5.18 Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with triethylsilanol (5b) and (*R*)-2-octanol (10b) to form (*R*)-triethyl(octan-2-yloxy)silane (11b)

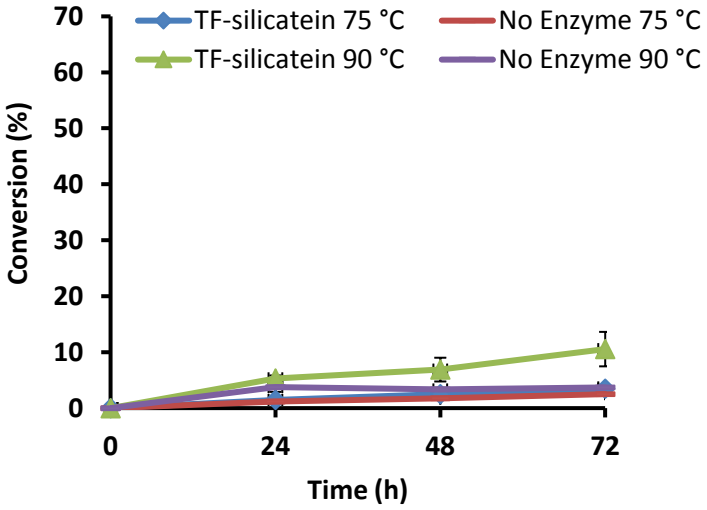


Figure 5.19 The percentage of product formation, (*R*)-triethyl(octan-2-yloxy)silane (11b), from triethylsilanol (5b) and (*R*)-2-octanol) (10b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS

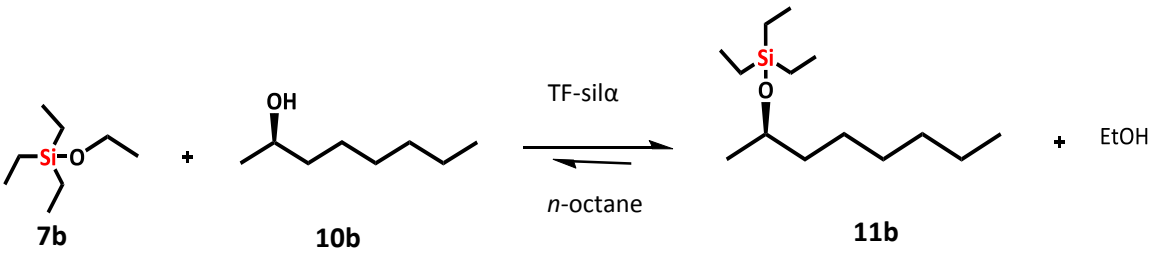


Figure 5.20 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with triethylsilane (7b) and (*R*)-2-octanol (10b) to form (*R*)-triethyl(octan-2-yloxy)silane (11b)

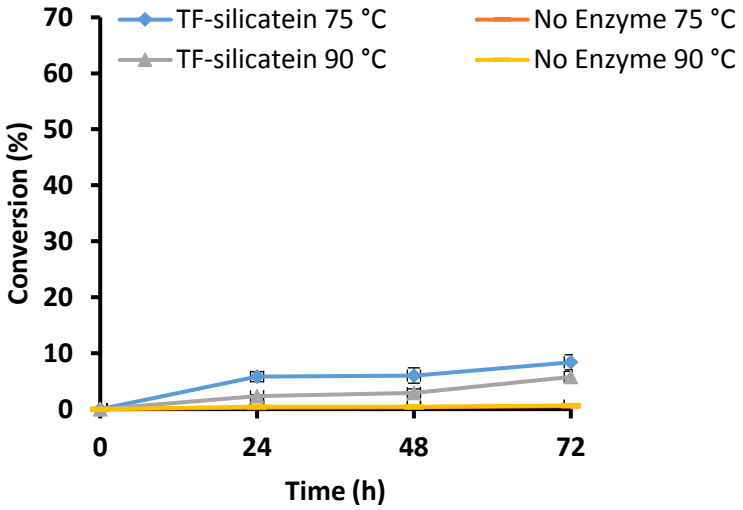


Figure 5.21 The percentage of product formation, (*R*)-triethyl(octan-2-yloxy)silane (11b), from triethylsilane (7b) and (*R*)-2-octanol (10b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS

Table 5.2 Percentage product conversion at 72 hours from etherification and trans-etherification reactions at 75 °C and 90 °C with *R*- and *S*-2 octanol

Substrate 1	Substrate 2 [a b]	Abbreviation	Product	Time (hrs)	Temperature (°C)	Conversion (%)
8b	5b	TES-OH	9b	72	75	13.7 ± 1.7
8b	5b	TES-OH	9b	72	90	28.6 ± 0.9
8b	7b	TES-Et	9b	72	75	11.2 ± 0.1
8b	7b	TES-Et	9b	72	90	9.8 ± 0.9
10b	5b	TES-OH	11b	72	75	3.5 ± 0.1
10b	5b	TES-OH	11b	72	90	10.5 ± 3.1
10b	7b	TES-Et	11b	72	75	8.4 ± 1.3
10b	7b	TES-Et	11b	72	90	5.6 ± 1.0

[a] 5 mol. equiv. of substrate used relative to substrate 1
[b] TES = triethylsilyl

GC-MS analysis of samples from silyl etherification and trans-etherification reactions with enantiomers *S*/*R*-2-octanol indicate TF-silicatein-α has a preference for *S*-2-octanol. The observed percentage product conversion for etherification and transetherification with *S*-2-octanol at 72 °C after 72 hours was 13.7% and 11.2% respectively which was marginally higher than 3.5% and 8.4% observed for *R*-2-octanol. As with 1-octanol the increase in temperature from 75 °C to 90 °C resulted in an increase in conversion (Figure 5.14; Figure 5.15; Figure 5.16; Figure 5.17; Figure 5.18; Figure 5.19; Figure 5.20; Figure 5.21). This may be due to the effects

of organic solvent on the 3D conformation of the enzyme such that it favours *S* enantiomer and facilitates correct orientation to accept the substrates and proceed with catalytic product conversion.[113,258,259]

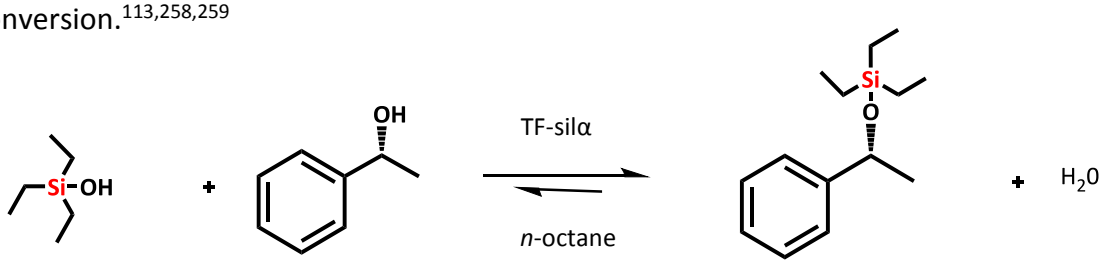


Figure 5.22 Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with triethylsilanol (5b) and (*S*)-1-phenylethanol (12b) to form (*S*)-triethyl(1-phenylethoxy)silane (13b)

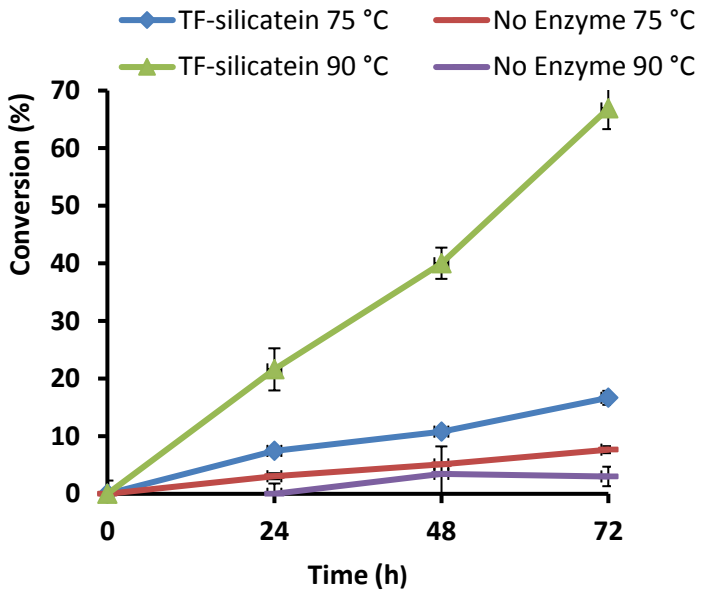


Figure 5.23 The percentage of product formation, (*S*)-triethyl(1-phenylethoxy)silane (13b), from triethylsilanol (5b) and (*S*)-1-phenylethanoll (12b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS

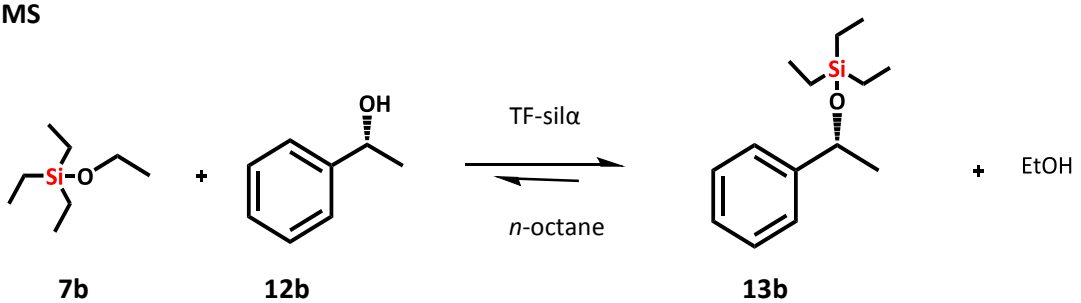


Figure 5.24 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with triethylsilane (7b) and (*S*)-1-phenylethanoll (12b) to form (*S*)-triethyl(1-phenylethoxy)silane (13b)

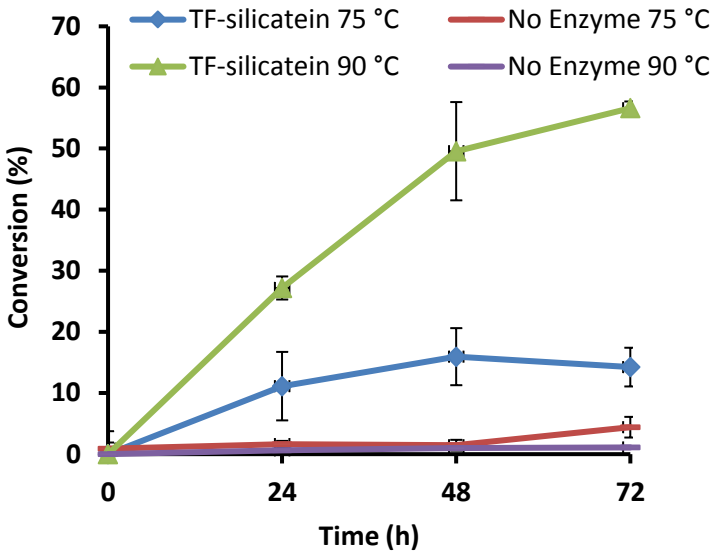


Figure 5.25 The percentage of product formation, (*S*)-triethyl(1-phenylethoxy)silane (13b), from triethylsilane (7b) and (*S*)-1-phenylethanol (12b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS

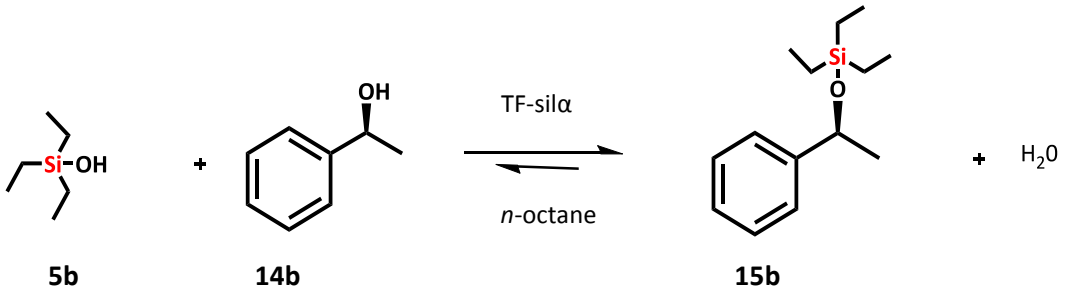


Figure 5.26 Etherification catalysed by TF-silα at 75 °C and 90°C in n-octane with triethylsilanol (5b) and (R)-1-phenylethanoll (14b) to form (R)-triethyl(1-phenylethoxy)silane (15b)

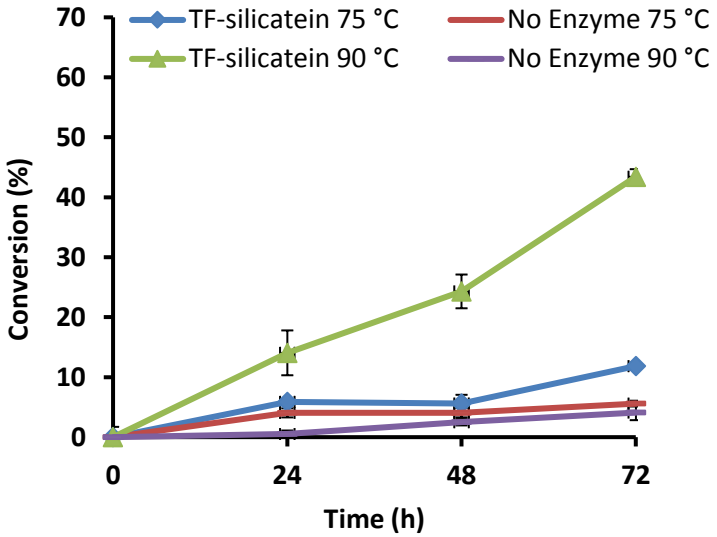
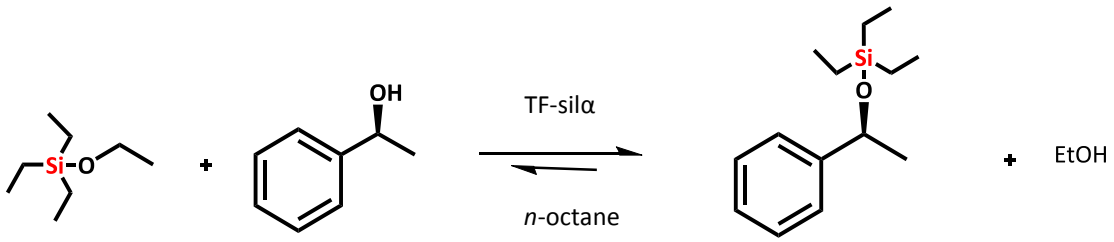


Figure 5.27 The percentage of product formation, (R)-triethyl(1-phenylethoxy)silane (15b), from triethylsilanol (5b) and (R)-1-phenylethanol (14b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS



7b **14b** **15b**

Figure 5.28 Trans-Etherification catalysed by TF-silα at 75 °C and 90°C in *n*-octane with triethylsilane (7b) and (R)-1-phenylethanoll (14b) to form (R)-triethyl(1-phenylethoxy)silane (15b)

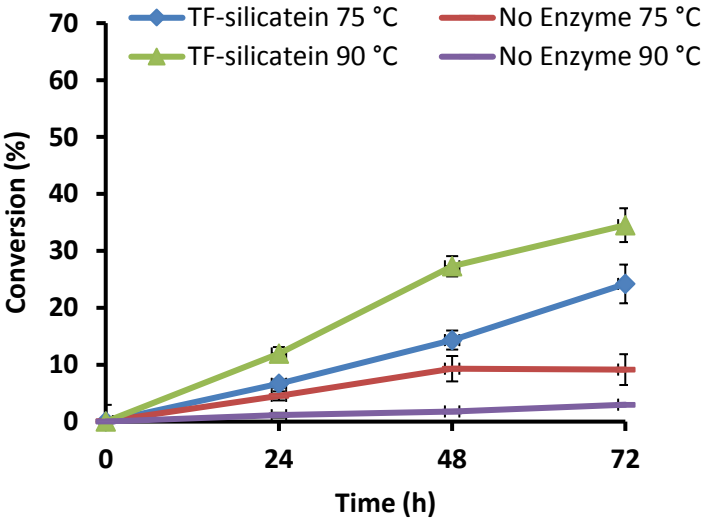


Figure 5.29 The percentage of product formation, (R)-triethyl(1-phenylethoxy)silane (15b), from triethylsilane (7b) and (R)-1-phenylethanoll (14b) in *n*-octane at 75 °C and 90 °C analysed by GC-MS

Table 5.3 Percentage product conversion at 72 hours from etherification and trans-etherification reactions at 75 °C and 90 °C with R- and S-1-phenylethanol

Substrate 1	Substrate 2 [a b]	Abbreviation	Product	Time (hrs)	Temperature (°C)	Conversion (%)
12b	5b	TES-OH	13b	72	75	16.7 ± 1.2
12b	5b	TES-OH	13b	72	90	66.9 ± 1.7
12b	7b	TES-Et	13b	72	75	14.2 ± 3.2
12b	7b	TES-Et	13b	72	90	56.6 ± 1.1
14b	5b	TES-OH	15b	72	75	11.8 ± 0.6
14b	5b	TES-OH	15b	72	90	43.4 ± 6.5
14b	7b	TES-Et	15b	72	75	24.2 ± 3.4
14b	7b	TES-Et	15b	72	90	34.5 ± 9.1

[a] 5 mol. equiv. of substrate used relative to substrate 1
[b] TMS = trimethylsilyl, TES = triethylsilyl

Experimental results from GC-MS analysis of samples from silyl etherification and trans-etherification reactions with enantiomers *S* and *R*- 1-phenylethanol indicate TF-silicatein-α has a preference for *S*-1-phenylethanol. The observed percentage product conversion for etherification and transethfication with *S*-1-phenylethanol at 90 °C after 72 hours was 66.9 % and 56.6 % respectively which was marginally higher than 43.4 % and 34.5 % observed for *R*-1-phenylethanol. As with 1-octanol and R/S octanol the increase in temperature from 75 °C to 90 °C resulted in an increase in product conversion and the percentage of product conversion was at a higher rate with the *S* enantiomer (Figure 5.22; Figure 5.23; Figure 5.24; Figure 5.25; Figure 5.26; Figure 5.27; Figure 5.28; Figure 5.29). In Chapter 3 the hydrolysis experiments with TBDMS-4np suggest the active site is large and can accept bulky substrates. However, in contrast the results in this chapter convey enantioselectivity which generally requires a binding site that is conformed to select between the enantiomers. However, plausible explanation for this could be loss of dynamic flexability and rigidity of the 3D structure of the enzyme in organic solvent being in such a form that would allow acceptance of the substrates into the active site and subsequent enzymatic condensation with silanol or transetherification with the ethoxysilane. This may also facilitate substrate binding and stabilisation particularly favouring the *S* enantiomer with the aromatic functional group in comparison to the other substrates (Table 5.2; Table 5.3) These results are in agreement with recently published data with suggested TF-silicatein-α selectively protected hydroxy groups present on phenolic rather than aliphatic compounds.[53,78] The results gained from these experiments confirm TF-silicatein-α to be stable and active in organic media at high

temperatures. It also shows enantioselectivity is initiated even at high temperatures. This makes it an attractive potential biocatalyst for several applications such as organosilicon synthesis and silyl protection of hydroxy groups in multistep synthesis.

5.3.3 Coupled organic synthesis in organic media with silicatein-α and ADH

As discussed in detail in Chapter 1, coupled or tandem enzyme reactions in organic synthesis are advantageous for many reasons and have been the focus of research for many years. The aim of this work was to investigate and identify an enzyme that could couple in a reaction to provide enantiomeric alcohol feedstocks required for silicatein-α to then complete silylation or function in partnership for racemic resoltuioion using the model silanol or ethoxy silane.

ADH was chosen as a prospective partner as it functions to reduce aldehydes and ketones into their corresponding alcohols using the cofactor NADH as a hydride donor. This would potentially produce substrates for TF-silicatein-α for silyl etherification and trans-etherification with silanol or ethoxysilane respectively. Three commercially sourced ADH varying in substrate specificity, one with a broad specificity, (ADH1), an *S* enantiomer specific, (ADH2), and an *R* enantiomer specific, (ADH3). The substrates used to investigate thermostability and enantioselectivity of TF-silicatein-α were used to screen substrate compatibility with ADH and aid in the identification of a candidate for coupled synthesis. The oxidation reaction with ethanol, *S*- and *R*-2-octanol and *S*- and *R*-1-phenylethanol was conducted in aqueous media to assess selectivity and catalytic activity with these substrates. Reactions containing 10% 1,4-dioxane were also examined due to immiscibility and poor solubility of some substrates in aqueous environments. The addition of the solubilising agent on enzyme activity was also assessed. The increase in absorbance at 340 nm was measured as it suggests reduction of NAD+ cofactor to NADH with product conversion.

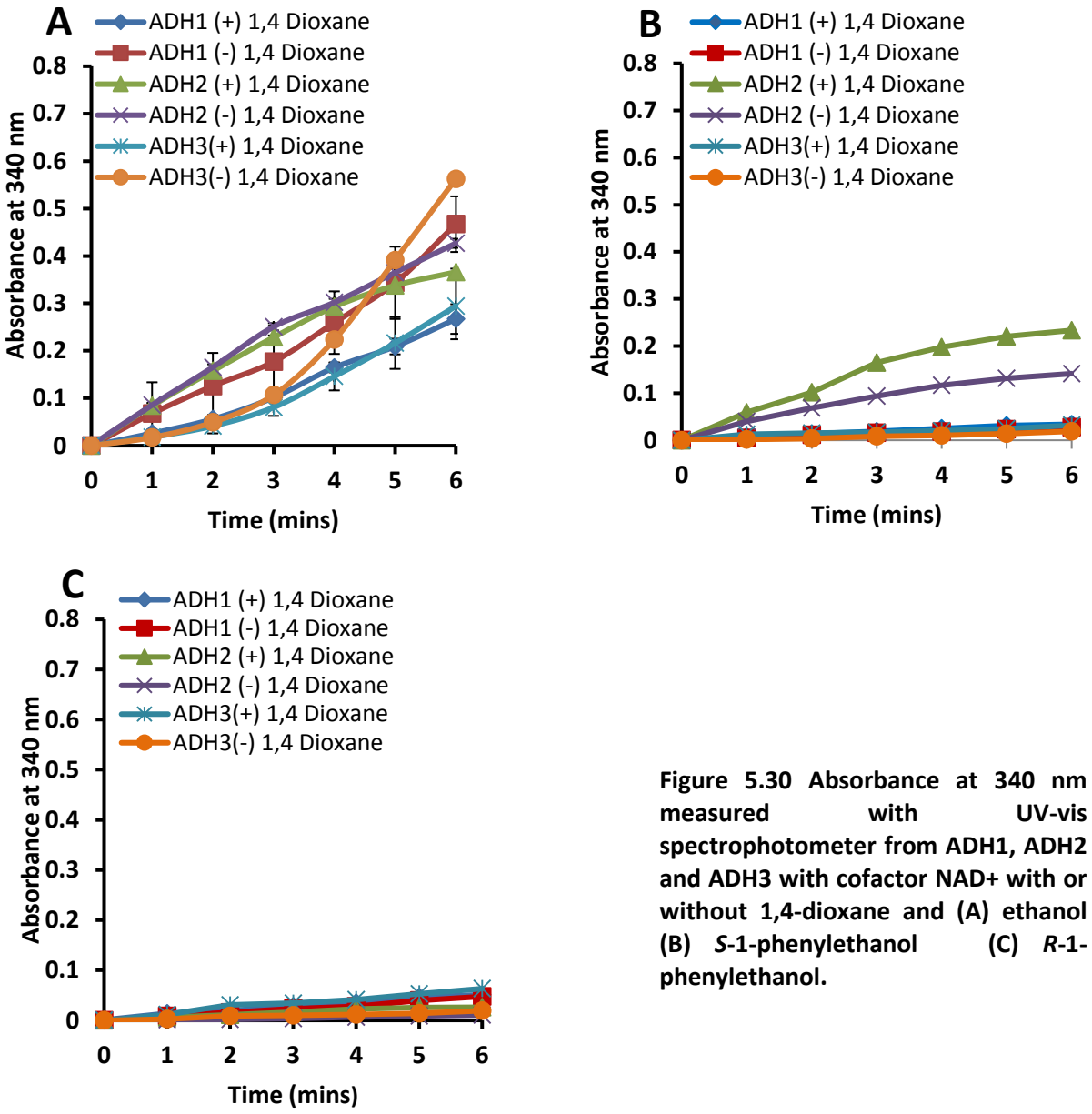


Figure 5.30 Absorbance at 340 nm measured with UV-vis spectrophotometer from ADH1, ADH2 and ADH3 with cofactor NAD+ with or without 1,4-dioxane and (A) ethanol (B) *S*-1-phenylethanol (C) *R*-1-phenylethanol.

The activity of the three enzymes were initially screened with ethanol to test the viability of each one under the reaction conditions used. With an increase in absorbance at 340 nm observed in all samples, due to the oxidation of NADH to NAD^+ the enzymes were assumed to be active and compatible with the applied conditions. Assessing the effect of addition of 1,4-dioxane with ADH and ethanol, it shows that the increase in absorbance is slightly lower in samples with 1,4-dioxane. This additive may have a slight inhibitory effect on the enzymes catalytic efficiency or act as a denaturant of the enzyme in this instance (Figure 5.30A). Reactions with *S*-1-phenylethanol and ADH2 also displayed this effect with a lower absorbance measurements displayed by samples without the solubilising agent. However, this may be due to the poor solubility of the substrate in aqueous media due to hydrophobicity of the molecule. An interesting observation is the significant increase in

absorbance with *S*-specific ADH2 (+)-1,4-dioxane (Figure 5.30B) suggesting it is able to accept this compound as a substrate. No significantly relevant results were observed in oxidation reactions with *R*-1-phenylethanol (Figure 5.30C) and *R*-2-octanol (Figure 5.31A). However, oxidation of *S*-2-octanol revealed similar results to *S*-1-phenylethanol as it suggests ADH2 also accepts this compound as a substrate (Figure 5.30B and Figure 5.31B).

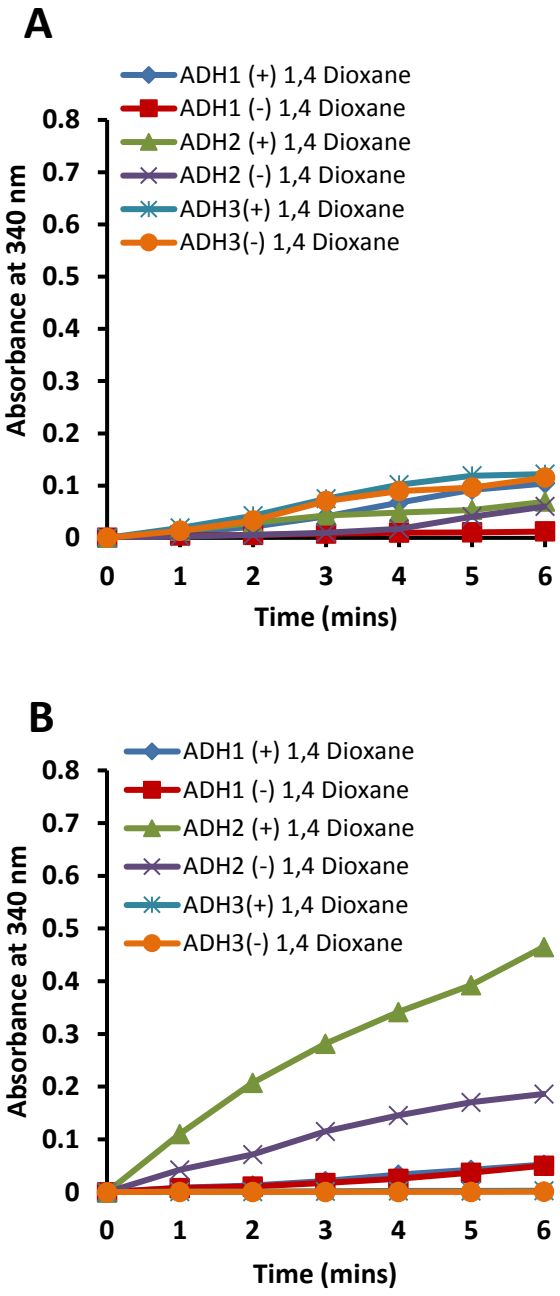


Figure 5.31 Absorbance at 340 nm measured with UV-vis spectrophotometer from ADH1, ADH2 and ADH3 with cofactor NAD+ with or without 1,4-dioxane and (A) *R*-2-octanol (B) *S*-2-octanol.

The reduction of acetaldehyde, octanone and acetophenone by ADH1, ADH2 and ADH3 with NADH as a cofactor in aqueous media was also conducted to screen and assess catalytic

activity and specificity towards potential substrates in a coupled reaction. As with the oxidation reactions all the substrates were screened and duplicate reactions were analysed either with or without 1,4-dioxane. A decrease in absorbance at 340 nm was measured as it is indicative of oxidation of NADH to NAD^+ and product conversion. Data from these experiments (not shown) confirmed ADH2 as a candidate for coupled synthesis as a decrease in absorbance at 340 nm was observed.

Analysis of the experimental data from the oxidation and reduction assays revealed ADH2 displayed a preference towards both *S*-enantiomeric substrates due to the increase in absorbance from the reduction of NAD^+ to NADH. Therefore, the next step was to progress with the potential candidate, ADH2, and screen the two *S* enantiomers, *S*-2-octanol and *S*-1-phenylethanol in an organic solvent system. The same conditions of the condensation reaction were applied as this was the optimum conditions for TF-silicatein-α. This approach was taken initially to determine whether the reaction could proceed in anhydrous organic solvent and a 'one pot system' for silylation or racemic resolution could be achieved or whether further optimisation would be required. The standards for each reaction were analysed by GC-MS, Adjustments to the GC method allowed separation of the retention times for octanone and *R*- and *S*-2-octanol. However, despite many alterations of the GC method differentiation of acetophenone and *R*- and *S*-1-phenylethanol could not be achieved. Therefore, this would need further optimisation for this method to be used as an analytical technique for the reaction or utilisation of an alternative method such as IR spectroscopy.

ADH2 was lyophilised with the relevant cofactor and added to the organic solvent containing *S*-2-octanol, *S*-1-phenylethanol, octanone or acetophenone. The samples taken every 24 hours and were analysed by IR and GC-MS but did not show any positive results as no product peak in the chromatogram was observed. A plausible explanation for this could be that the conditions used have a detrimental effect on the activity of the enzyme and therefore optimisation is required such as a lower temperature or a biphasic system.

5.4 Conclusion

The work in this chapter aimed to expand our current knowledge and recently published data that revealed TF-silicatein-α to be a promising candidate as a biocatalyst with industrial applications catalytically active in organic solvent.[53,78] Etherification and trans-etherification reactions in *n*-octane with TES-OH as a model silanol and primary alcohol (1-octanol) were conducted initially at temperature of 75 °C and 90 °C. The results show an increase in product conversion at the higher temperature suggesting TF-silicatein-α to be thermostable and catalytically active under these conditions.

Previous work also assessed the regioselectivity of TF-silicatein-α using compounds with multi-hydroxy groups showing it to have a preference for phenolic alcohols over aliphatic functional groups.[53,78]Therefore, enantioselectivity of the enzyme was explored with the substrates *R*- and *S*-2-octanol and *R*- and *S*-1-phenylethanol and model silanol TES-OH or ethoxy silane TES-OEt in *n*-octane at 75 °C and 90 °C. The main observations of these experiments were a significant increase in percent product conversion at 90 °C with most of substrates used in the screening process. In addition to displaying a preference towards *S* enantiomers compared to *R* enantiomers. An overall preference towards aromatic compounds was noted with product conversions of 66.9% and 56.6% for etherification and transetherification respectively with *S*-1-phenylethanol in comparison to 13.7% and 11.2% observed for *S*-2-octanol also validates data collated previously within our group.[53,78]

Additional experiments were aimed at developing a tandem enzymatic one pot synthesis for silyl etherification that could have applications in the pharmaceutical industry due to the importance of chiral compounds. Therefore, enzyme-mediated selective silyl protection would be advantageous offering a 'greener' approach. Preliminary oxidation and reduction reactions with three potential enzyme partners for TF-silicatein-α, ADH1, ADH2 and ADH3 were screened with the substrates used in the enantioselective experiments. Firstly this was conducted in aqueous media due to the possibility of a coupled-enzyme hydrolytic reaction, however, TF-silicatein-α was unable to hydrolysis the model substrate used under the conditions tested. The results from the substrate screening with ADH 1-3 did, however, reveal ADH2 to be a promising candidate for coupled synthesis. As ADH2 is a *S*-specific ADH it was shown to accept and convert *S*-2-octanol and *S*-1-phenylethanol, based on the observed

increase in absorbance due to the oxidation of the co-factor NADH. Therefore, this enzyme applied to the conditions used for etherification and trans-etherification by TF-silicatein-α in organic solvent. Experimental data from this suggested ADH2 was not suitable for that set of conditions and would, therefore, need extensive optimisation and screening of various different conditions and factors beyond the timescale allow for this study.

In summary, the data reported in this chapter shows TF-silicatein-α to be a robust, thermostable and catalytically active enzyme for silyl etherification and trans-etherification with a variety of different compounds. The enzyme also shows enantioselectivity towards *S*-enantiomers, in particular, aromatic alcohol functional groups. Therefore, the results collated in this chapter could be of great benefit and advantageous with respect to TF-silicatein-α as a potential biocatalyst with industrial applications considering especially no modification or protein engineering has been undertaken thus far.

6. Chapter 6: Conclusions

Siloxane based materials and silyl protective groups are nearly universally applied in many areas of industry. However, the production of these compounds has adverse environmental impacts due to the harsh conditions under which they are produced. Many areas of industry now focus on utilising natural or modified catalysts such as enzymes for chemical transformations. In this respect, the application of biocatalysts has many benefits as they can function under more ambient conditions, and be exploited through inherent specificity towards their substrates. Consequently, the numbers of biocatalysts have significantly increased in the last 40 years because of the benefit offered in terms of increased stereo-, regio- and chemoselectivity and have a diverse range of applications in pharmaceutical industry and large scale production of fine chemicals. The identification of silicatein is of particular interest due to the potential use as a biocatalyst in organosilicon chemistry.

In the heterologous production of silicatein-α many efforts were made to improve the solubility of the native, "wild type" enzyme, silicatein-α. Fusion protein technology was employed using various fusion tags, (TF, GST, MBP, SUMO) tethered at the N-terminal and cultured using a variety of conditions. Significant improvement in the solubility of silicatein-α was observed using TF as a fusion protein (Figure 2.4). The purification of the fusion protein was optimised with a more streamline process using the same buffer, thus reducing the number of buffer exchange steps in the process (Figure 2.18). Additionally, implementing a pre-purification step, ammonium sulphate precipitation also improved purity of the protein (Figure 2.8).

Earlier research within our group produced soluble silicatein-α in its native form though the addition of non-denaturing detergents in the lysis buffer, albeit to a small yield. However, this method was time-consuming and used a lot of expensive additives. Therefore, optimisation was achieved through the use of amino acid additives (L-arginine and L-glutamate) in the standard buffer (Figure 2.11), which have been documented to enhance solubility of many other proteins. This method was more cost and time efficient in addition to having no detrimental effect on the purity or yield of protein.

Several measures to increase the yield of the wild type protein were attempted through various refolding techniques (Figure 2.10) and protease cleavage of the fusion protein (Figure

2.16). However, this area of work requires further optimisation. The addition of the strep-tag and thrombin cleavage site to the C-terminal of TF-silicatein-α and subsequent protease cleavage may benefit from further molecular cloning to add an alternative protease cleavage site. As the results revealed thrombin to be most efficient protease, the presence of an additional cleavage site in between trigger factor and silicatein-α hindered isolation of the protein.

The catalytic ability of silicatein-α and TF-silicatein-α was assessed using established biochemical assays to confirm activity against substrates TEOS and silicic acid, in addition to a recently developed high throughput assay for organosilanes with the substrate TBDMS-4np. Experimental results confirmed catalytic activity through the accelerated rate of product formation in comparison to the control samples tested (Figure 3.2; Figure 3.3; Figure 3.4). To explore the function of each residue in the catalytic triad three variants of the enzyme were constructed by individual mutation of each residue to chemically inert alanine (Ser26Ala, His165Ala and Asn185Asp) and screened using the aforementioned biochemical assays. A general trend was observed with the enzyme variants as each of the mutants tested displayed a reduced catalytic activity in comparison to the wild type enzymes. However, the concentration of product formed in the enzyme variants samples was higher than that observed in the control samples (Figure 3.2; Figure 3.3;Figure 3.4;Table 3.1). It is thought that this result could be due to some other factors such as water molecules or other residues functioning in the absence of the nucleophilic or basic residue. Therefore, it appears that each residue is important for full catalytic activity but not essential as some activity is still observed. Additional enzyme variants were constructed with mutations of active site resdiues similar to the catalytic triad of serine and cysteine proteases, Asn185Asp, Ser26Cys respectively and screened using the same biochemical assays. The substrates contained either an ester (4-nitrophenyl pivalate) or amide (4-nitrophenyl pivalamide) bonds in place of the dimethylsiloxy group and a dipeptide (Cbz-Phe-Arg-NHNp) containing the cleavage site specific for cathepsin L. Experimental results confirmed previous data showing TF-silicatein-α displays no esterase or protease activity, this result was also true for the Ser26Cys variant, however, Asn185Asp was observed to show some esterase activity (1.24 % conversion) (Figure 3.6). Michaelis Menten kinetic data was obtained for all enzyme variants and the wild type enzyme using TBDMS-4-nitrophenyl silyl ether assay which shows K_M was essentially the same, however,

turnover is reduced in the enzyme variants. These results imply that substrate binding was unaffected and therefore not the cause of the reduced activity. In addition, for full catalytic activity in the hydrolysis of this substrate, all three native residues are required to be present but the absence of each one is not essential for overall catalytic function.

To gather some structural information, assess homogeneity and to screen potential conditions for crystallography trials biophysical analysis was conducted using several techniques. CD and light scattering experiments confirmed homogeneity of the enzymes under the standard assay conditions. In addition the mutations did not affect the overall structure as CD spectra for all proteins were observed to be the same (Figure 4.3; Figure 4.4). A large hydrodynamic radius was recorded from DLS for TF-silicatein-α, which could be due to the shape being non-globular (Figure 4.5). The DLS results for standard buffer conditions with and without L-arginine and L-glutamate revealed the addition of the additives promoted homogeneity due to the decrease in PDI (Table 4.11).

Analysis of the MALS results for TF-silicatein-α in various buffers and diferent pH were inconclusive. The protein eluted in the void volume with a calculated molecular weight on the order of magnitude of 10^6 Daltons, suggesting aggregation. This was concluded to be due to handling error, aggregation due to protein unfolding under the heat of the laser or insufficient chromatographic conditions that would require further optimisation. An alternative method of analysis for this protein could be low angle light scattering (LALS) especially if the results we observed are due to TF-silicatein-α forming highly ordered oligomeric structures. As LALS measures close to zero angle which is favoured in measuring large angular dependent molecules thereby reducing data fitting or extrapolation of data and producing more accurate measurement.[219,220] The DLS results of the buffer screening experiment identified two buffers (EPPS and Bis-Tris) in which silicatein-α exhibited a satisfactory level of homogeneity due to the low value of the PDI (Table 4.12). MALS measurements conducted with silicatein-a in each buffer were successful signifying potential progress to crystallography trials using these conditions.

To further expand the parameters in which the enzyme operates efficiently in non-aqueous solutions the condensation reactions were repeated at 75 °C and 90 °C. An increase in product formation was observed at the higher temperature in comparison to the control sample which confirms the thermostability and catalytic activity of TF-silicatein-α under these conditions

(Figure 5.). These reaction conditions were also use to test enantioselectivity using *R*-2 and *S*-2-octanol and *R*-1 and *S*-1-phenylethanol and model silanol TES-OH or ethoxy silane TES-OEt in n-octane at 75 °C and 90 °C. Product formation was observed to increase significantly at the higher temperature and the enzyme displayed a preference towards *S* enantiomers in particular *S*-1-phenylethanol (66.9% and 56.6%) (Table 5.2; Table 5.3). The enantioselectivity observed suggests there could be a change in the enzymes conformation which would facilitate this selectivity towards the S enantiomer, as previous hydrolysis reactions show the enzyme can accept bulky substrates. Further work was aimed at developing a tandem enzymatic one pot synthesis for silyl etherification using TF-silicatein-α and ADH. Preliminary data identified a potential ADH candidate, although, the conditions of the reaction in organic solvent using this enzyme did not produce positive results. Therefore, this step would need extensive optimisation and screening of various different conditions and factors beyond the timescale allowed for this study.

Overall, the work described in this thesis further confirms the potential of silicatein-α as a robust, thermostable and catalytically active enzyme that could be utilised as a biocatalyst with applications in organosilicon chemistry. The purification process was optimised reducing the overall time and also the number of steps involved thus reducing the risk of protein loss at each stage. The biochemical analysis with enzyme variants implies each residue is important to the overall function of the enzyme. The biophysical analysis confirmed no change to the overall structure of the protein with mutations and highlighted buffers that convey protein stability which is a great advantage to future crystallography trials. In organic solvent systems the enzyme was shown to have increased thermostability and enantioselectivity towards *S* enantiomers. In addition, preliminary data derived from the ADH experiments suggest a potential enzymatic partner for a coupled reaction to produce chiral compounds with industrial applications.

7. Chapter 7: Materials and Methods

7.1 General materials and equipment

All chemicals were purchased from Sigma Aldrich (UK), VWR International (UK) and Alfa Aesar (UK). All solvents and reagents were of analytical grade and ultrapure water (18.2 MΩ•cm resistivity at 25°C) used for assays and buffer preparation was prepared by Milli-Q® filtration (Merck Millipore, Germany). Silicate test kit was purchased from Merck (Germany) and molecular biology kits supplied by Qiagen (UK). Gibson assembly reagents, DNA ladders, 6x DNA loading dye and restriction enzymes were purchased from New England Biolabs. Prestained protein molecular weight markers were from Thermo Fisher Scientific (UK) and Instant Blue from Expedeon (UK) used as protein staining solution. GST-HRV3C was obtained from Mark Dunstan, University of Manchester and molecular chaperone plasmids were obtained for Takara.

E. coli strains BL21(DE3) and Arctic Express (DE3) (Agilent Technologies, UK) were used during this research. Cell lysis was conducted using lysozyme (Sigma Aldrich, UK) and a Bandelin Sonoplus HD2070 probe sonicator. HisTrap HP column (GE Healthcare, UK), GSTrap FF column (GE Healthcare, UK) and StrepTrap HP (GE Healthcare, UK) were used for Fast protein liquid chromatography (FPLC) based protein purification using an ÄKTApurifier system (GE Healthcare, UK). Protein purification was also carried out using gravity flow column (Qiagen, UK) loaded with Ni-NTA agarose resin (Thermo Scientific, UK). HiLoad 16/600 Superdex 200 column was used for Size exclusion chromatography (SEC) with an ÄKTApure system. DNA and protein concentrations were measured using Nanodrop 200 UV-Vis spectrophotometer (Thermo Scientific) measuring at 260 nm and 280 nm respectively. All other UV-visible data was collated using a Synergy HT Multi-Mode Microplate Reader (BioTek Instruments, USA) using a quartz 96-well plate (Hellma Analytics, Germany). Circular Dichroism (CD) measurements were made using Chirascan CD spectrometer (Applied Photophysics, UK), Multi angled laser light scattering MALLS conducted with DAWN HELEOS (Wyatt Technology) and GC-MS analysis using Thermo Finnigan PolarisQ Ion Trap Mass Spectrometer or Varian Saturn 2000 Ion Trap Mass Spectrometer. The dynamic light scattering experiments used DynaPro Plate Reader (Wyatt Technology) and UNcle (Unchained Lab, USA) which also gave static light scattering results for T_m and T_{agg}.

Growth media

LB agar: 1 % tryptone, 0.5 % yeast extract, 1 % NaCl, 15 % agar

LB broth: 1 % tryptone, 0.5 % yeast extract, 1 % NaCl

TB broth: 1.2% tryptone, 2.4% yeast extract 0.4% glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4

Buffers

Phosphate-buffered saline (PBS) buffer: 137 mM NaCl, 2.7 mM KCl, 8.1 mM Na_2HPO_4, 1.47 mM KH_2PO_4, pH 7.4

SUMO lysis buffer: 50 mM potassium phosphate, pH 7.8, 400 mM NaCl, 100 mM KCl, 10% glycerol, 0.5% Triton X-100, 10 mM imidazole

pHis lysis buffer 1: 5% Glycerol, 750 mM NaCl, 1mM $CaCl_2$, 1 mM $MgCl_2$, 5 mM CHAPS, 1 % Triton X-100, 25 mM Tris, 10 mM Imidazole.

pHis lysis buffer 2: 50 mM Tris, 100 mM NaCl, 500 mM L-arginine, 500 mM L-glutamate.

Western blot transfer buffer: 25 mM Tris, 192 mM glycine and 10% methanol

IMAC buffer A: 50 mM KH_2PO_4, 500 mM NaCl

IMAC buffer B: 50 mM KH_2PO_4, 500 mM NaCl, 1M imidazole

Refolding buffer A: 50 mM Tris-HCl, 1 mM EDTA, 10 mM DTT, pH 8.0

Refolding buffer B: 50 mM Tris-HCl,1 mM EDTA, 10 mM DTT, 2 M urea pH 8.0

Refolding buffer C: 50 mM Tris-HCl,1 mM EDTA, 10 mM DTT, 6 M urea pH 8.0

Refolding buffer C1: 50 mM Tris-HCl, 0.5 M L-arginine, 9 mM glutathione, 1 mM oxidized glutathione 0.3 M NaCl, 1 mM KCl, pH 8.5

TAE buffer: 40 mM Tris (pH 7.6), 20 mM acetic acid, 1 mM EDTA

SDS PAGE running buffer: 25 mM Tris, 192 mM glycine, 0.1 % SDS, pH 8.6

5x Laemmli sample buffer: 60 mM Tris-Cl pH 6.8, 2 % SDS, 10 % glycerol, 5 % β-mercaptoethanol, 0.01 % bromophenol blue

7.2 General methods

7.2.1 Polymerase Chain Reaction (PCR)

PCR for the construction of the fusion protein plasmids reported in Chapter 2 and Chapter 3 was conducted using the following conditions (Table 7.1; Table 7.2; Table 7.3; Table 7.4); however, annealing temperatures varied or each individual reaction was part of a temperature gradient with the range according to the manufacturer's protocol.

Table 7.1 PCR reagents for construction of fusion protein plasmids

	Final Concentration	Volume (µL)
Reaction buffer	1X	5
10 mM dNTPs	200 µM	0.5
10 µM Forward Primer	0.5 µM	1.25
10 µM Reverse Primer	0.5 µM	1.25
DNA template	Variable	Variable
Polymerase	0.02 U/µL	0.25
Nuclease-Free Water	to 25 µL	

Table 7.2 PCR program conditions for construction of fusion protein plasmids

	Temperature (°C)	Time (s)	Cycles
Initial denaturation	98	120 mins	1
Denaturation	98	10 secs	30
Annealing	Variable/gradient	30 secs	
Extension	72	120 mins	
Final extension	72	360 mins	1
Final Hold	4		

Table 7.3 PCR reagents for protein variants (megaprimer)

	Final Concentration	Volume (µL)
Reaction buffer	1X	5
5 mM dNTPs	200 µM	4
5 µM Forward Primer	0.5 µM	1
5 µM Reverse Primer	0.5 µM	1
DNA template	25 ng/ µL	1
Polymerase	1 U/µL	0.5
Nuclease-Free Water	to 25 µL	

Table 7.4 PCR conditions for protein variants (megaprimer)

	Temperature (°C)	Time (s)	Cycles
Initial denaturation	95	180	1
1st stage PCR	95	30	5
	60	60	
	72	330 (60 secs/kb template)	
2nd stage PCR	95	30	20
	68	480 (60 secs/kb template)	
Final extension	68	840 (120 secs/kb template)	1
Final Hold	4		

7.2.2 DpnI Treatment

Table 7.5 Reagents of *Dpn*I Treatment

	1x (µL)
Nuclease-free water	1.5
Buffer	2.5
*Dpn*I	1.0
PCR product	20
TOTAL	25

All reagents were added (Table 7.5) mixed well and incubated at 37°C for 1 hour, a further 1 µL of *Dpn*I was added to the reaction mixture and incubated for 1 hour. 5 µL of this mixture was used for transformation.

7.2.3 DNA electrophoresis

1g of Agarose was dissolved in 100 mL of TAE by heating then subsequently cooling to approximately 50°C. Ethidium bromide (2 µL, 10 mg/mL) was added and mixed thoroughly before being poured into a gel casting cassette to set with the required well comb. After setting the gel was placed in a gel electrophoresis tank and covered with TAE buffer. DNA samples and 1 kb DNA ladder were mixed with 6x DNA loading dye and loaded into each well. Electrophoresis was run at 110 V for 30-40 minutes after which gels were visualised by UV transilluminator.

7.2.4 DNA extraction

DNA fragments were removed from the gel using a scalpel then extracted using a QIAquick gel extraction kit (Qiagen) as per manufacturer's protocol.

7.2.5 SDS-PAGE gels

10% and 15% SDS –PAGE gels were made up according to the recipe below (Table 7.6) or pre-cast gels sourced from Bio-Rad

Table 7.6 SDS-PAGE gel components for a single gel at 10% or 15%

	Resolving gel 10%	Resolving gel 15%	Stacking gel 6%
ddH_2O	3.8 Ml	2.8 mL	2.9 mL
40% Acrylamide	2 Ml	3 Ml	0.75 mL
1.5M Tris pH 8.8	2 Ml	2 Ml	-
0.5M Tris pH 6.8	-	-	1.25 mL
10% SDS	80 µL	80 µL	50 µL
10% APS	80 µL	80 µL	50 µL
TEMED	8 µL	8 µL	5 µL

5x Laemmli sample buffer was added to protein samples, mixed and heated for 7-10 minutes at 95°C. SDS-PAGE gels were added to the SDS-PAGE tank containing 1x SDS buffer. The protein samples were then added to each individual well and SDS electrophoresis was conducted at 70 V until the sample reached the resolving gel then 110 V until the dye front reached the bottom of the gel. Pre-cast gels were run as per manufacturer's protocol for the time and voltage advised. Following electrophoresis the gels were stained with Instant blue and placed on a roller mixer for at least 1 hour then subsequently de-stained with water.

7.2.6 Gibson assembly

Gibson assembly reaction mixture was prepared on ice using 0.02-0.5 pmols of linearised vector and PCR fragments as per manufacturer's protocol. To this 10 µL of Gibson Assembly Master Mix was added and a volume of water to make the total volume 20 µL. The mixed was incubated at 50°C for 60 minutes and then used for transformation.

7.2.7 Chemically competent bacterial cells transformation

30-50 µL transformation of frozen competent cells were thawed on ice and added to 2-5 µL of plasmid DNA or 4 µL of Gibson assembly product then put on ice for 30 minutes. This was subsequently incubated at 42°C for 30 seconds and then put back on ice for 2 minutes, after which 200 µL of SOC media was added a further incubation at 37°C for 1 hour shaking at 200 rpm. Aliquots of incubated cells (25 µL, 50 µL and 100 µL) were spread on to pre-warmed LB agar plates which contained the necessary antibiotic for selection then incubated overnight at 37°C.

7.3 Experimental Methodology for Chapter 2

7.3.1 Bioinformatic search using orthologs of silicatein-α from *S. domuncula*

The protein sequences for silicatein-α from *Suberites domuncula* (Q2MEV3), *Spongilla lacustris* (D3DEQ1), *Baikalospongia fungiformis* (G3EAH4), *Lubomirskia baicalensis* (Q2PC18),

and *Ephydatia fluviatilis* (B5U9F0) were obtained from UniProt[177] and aligned using Clustal Omega.[178,179] Subsequent structural analysis was conducted using with SWISS-MODEL and Swiss PdbViewer[180] using cathsilicatein (PDB ref. 2vhs) as a homology model and visualisation using PyMol.

7.3.2 Plasmid construction

GST-silicatein-α: Cloning by Gibson assembly was carried out by Joseph Hosford. The full length open reading frame of silicatein-α and GST cDNA was amplified by PCR using oligonucleotide primers (Sigma-Aldrich)

SILI-FWD: 5' -GTTTCAGGGTCCGTCGGACTATCCGGAAGCC - 3',

SILI-REV: 5' – ATCTCAGTGGTGGTGGTGGTGGTGCACCAGGGCGGATA - 3'

GST-FWD: 5' - ACTGGTGGACAGCAAATGGGTCGCGTGTCCCCTATACTAGGTTATTG - 3'

GST-REV: 5' - CCGGATAGTCCGACGGACCCTGAAACAGAAC - 3' and subcloned into BamHI and Xho1 sites of the pET 28a vector by Gibson assembly and confirmed by sequencing. The protein of interest was expressed as a fusion protein with a hexahistidine tag followed by glutathione-S-transferase (GST) at the N-terminus and a further hexahistidine tag at the C-terminus.

SUMO-silicatein-α: The full length open reading frame of silicatein-α was amplified by PCR using oligonucleotide primers (Sigma-Aldrich)

SILI TA -FWD: 5' - GACTATCCGGAAGCC - 3',

SILI TA-REV: 5' -CAGGGTCGGATAGCT- 3'

and subcloned into pET SUMO vector (Life Technologies) according to the supplier's instructions and confirmed by sequencing. The protein of interest was expressed as a fusion protein with a hexahistidine tag and SUMO at the N-terminus.

Trigger factor-silicatein-α: This work was carried out by Seyed Yasin Tabatabaei Dakhili. The full length open reading frame of silicatein-α was amplified by PCR using oligonucleotide primers, (Takara)

pCold F1 Silicatein-FWD: 5' -CTCGAGGACTATCCGGAAGCCGTTGA- 3'

pCold R1 Primer-REV: 5’ - CAGAATCTAAGATCCCTGCCA - 3’

and subcloned into XhoI and EcoRI sites of the pCold vector and confirmed by sequencing. The protein of interest was expressed as a fusion protein with hexahistidine tag and trigger factor at the N-terminus.

MBP-silicatein-α: The full length open reading frame of silicatein-α and MBP cDNA was amplified by PCR using oligonucleotide primers (Sigma-Aldrich)

SILI-pet28aFWD: 5’ -CCTTCTTCGATTTTCATGCCGCGACCCATTTGCTGTCCACCAGT - 3’,

SILI-REV: 5’ - CCGGATAGTCCGACGGACCCTGAAACAGAAC – 3’

mbp-FWD: 5’ – GTTTCAGGGTCCGTCGGACTATCCGGAAGCC – 3’

mbp-pet28aREV: 5’ – CAGCTATCCGACCCTGGTGCACCACCACCACCACCACTGAGAT - 3’

and subcloned into the NdeI site of the pET 28a vector by Gibson assembly and confirmed by sequencing. The protein of interest was expressed as a fusion protein with a hexahistidine tag followed by maltose binding protein (MBP) at the N-terminus and a further hexahistidine tag at the C-terminus.

Silicatein-α: The codon optimised gene from *S. domuncula* equivalent to silicatein was synthesised (Genscript, Piscataway, NJ) and cloned into pHis vector using XhoI and EcoRI sites this work was completed by Protein Expression Facility, University of Manchester using oligonucleotide primers,

FORWARD: AAAAAACTCGAGGACTATCCGGAAGCCGTTGA

REVERSE: AAAAAAGAATTCTTACAGGGTCGGATAGCTTGG

TF-silicatein-α-Strep: This work was carried out by Ser Huy Teh. Plasmid was constructed by site directed mutagenesis using two step PCR with TF-silicatein-α in pCold vector for DNA template and primers:

P1_FORWARD: 5’ – CCACGCGGTAGTGGTGGTATC – 3’

Sil_Throm_Strep_REVERSE:5'-
ACAAGCTTGAATTCTTACTTTTCGAACTGCGGGTGGCTCCAACTGCCGCGCGGCACCAGCAGGGTCG
GATAG-3'

and confirmed by sequencing. The protein of interest was expressed as a fusion protein with a hexahistidine tag followed by trigger factor at the N-terminus and a Strep tag at the C-terminus.

7.3.3 Cell cultures

GST-silicatein-α: The constructed vectors were transformed into *E. Coli* BL21(DE3) and grown overnight at 37°C on LB agar plates supplemented with 50 µg/mL kanamycin. 5 mL cultures from a single colony were grown overnight at 37°C in LB broth containing 50 µg/mL kanamycin, shaking at 180 rpm. Cultures were diluted 50-fold into fresh LB medium supplemented with 50 µg/mL kanamycin, grown to OD_{600} of 0.5-0.6 at 37°C shaking at 180 rpm. Gene expression was induced by addition of isopropyl-β-D-galactoside (IPTG) to a final concentration of 1 mM and incubated at 37°C for 4 hours, 30°C for 6 hours, 25°C or 16°C overnight, shaking at 180 rpm, after which cells were harvested by centrifugation (3000 g, 20 minutes, 4°C) and pellets frozen at -20°C. Cold shock response was induced by placing cultures in an ice bath prior to the addition of 0.5 mM or 1 mM IPTG at an OD600 of 0.1 or 0.3 then subsequent incubation at 18°C overnight, shaking at 180 rpm. Cells were then harvested by centrifugation (3500 g, 20 minutes, 4°C) and pellets frozen at -20°C.

SUMO-silicatein-α: The constructed vectors were transformed into *E. Coli* BL21(DE3) (Life Technologies) and grown at 37°C on LB agar plates supplemented with 50 µg/mL kanamycin. 5 ml cultures from a single colony were grown overnight at 37°C in LB broth containing 50 µg/mL kanamycin, shaking at 180 rpm. Cultures were diluted 50-fold into fresh LB medium supplemented with 50 µg/mL kanamycin, grown to OD_{600} of 0.5-0.6 at 37°C shaking at 180 rpm. Protein expression was induced by addition of IPTG to a final concentration of 1 mM and incubated at 37°C for 4 hours, 30°C for 6 hours, 25°C or 16°C overnight, shaking at 180 rpm, after which cells were harvested by centrifugation (4000 rpm, 20 minutes, 4°C) and pellets frozen at -20°C.

Trigger factor-silicatein-α : The constructed vectors were transformed into *E. Coli* BL21(DE3) (Agilent Technologies) or Arctic Express (Agilent Technologies) and grown at 37°C on LB agar plates supplemented with 100 µg/mL ampicillin. 5 mL cultures from a single colony were grown overnight at 37°C in LB broth containing 100 µg/mL ampicillin, shaking at 180 rpm.

Cultures were diluted 50-fold into fresh LB medium or TB broth supplemented with 100 μg/mL ampicillin, grown to OD_{600} of 0.4-0.5 at 37°C shaking at 180 rpm. Gene expression was induced by addition of IPTG to a final concentration of 1 mM and incubated at 37°C for 4 hours, 30°C for 6 hours, 20°C, 25°C or 15°C overnight, shaking at 180 rpm, after which cells were harvested by centrifugation (3500 g, 20 minutes, 4°C) and pellets frozen at -20°C.

MBP-silicatein-α: The constructed vectors were transformed into *E. Coli* BL21(DE3) (Agilent Technologies) and grown at 37°C on LB agar plates supplemented with 100 μg/mL ampicillin. 5 mL cultures from a single colony were grown overnight at 37°C in LB broth containing 100 μg/mL ampicillin, shaking at 180 rpm. Cultures were diluted 50-fold into fresh LB medium or TB supplemented with 100 μg/mL ampicillin and grown to OD_{600} of 0.4-0.5 at 37°C, shaking at 200 rpm. Gene expression was induced by addition of IPTG to a final concentration of 1 mM and incubated at 37°C for 4 hours, 30°C for 6 hours, 20°C, 25°C or 15°C overnight, shaking at 180 rpm shaking at 180 rpm, after which cells were harvested by centrifugation (3500 g, 20 minutes, 4°C) and pellets frozen at -20°C.

GST-HRV3C: The vector containing GST-HRV3C was transformed into *E. Coli* BL21(DE3) (Agilent Technologies) and grown at 37°C on LB agar plates supplemented with 100 μg/mL ampicillin. 5 mL cultures from a single colony were grown overnight at 37°C in LB broth containing 100 μg/mL ampicillin, shaking at 180 rpm. Cultures were diluted 50-fold into fresh LB medium supplemented with 100 μg/mL ampicillin and grown to OD_{600} of 0.5-0.6 at 37°C, shaking at 180 rpm. Gene expression was induced by addition of IPTG to a final concentration of 1 mM and incubated at 30°C for 6 hours shaking at 180 rpm, after which cells were harvested by centrifugation (3500 g, 20 minutes, 4°C) and pellets frozen at -20°C.

Co-expression of GST-silicatein-α and SUMO-silicatein-α with molecular chaperones: The constructed vectors containing GST-silicatein-α and SUMO-silicatein-α were transformed separately into *E. Coli* BL21(DE3) along with either pKJE7, pGTf2 or pGro7 and grown overnight at 37°C on LB agar plates supplemented with 50 μg/mL kanamycin and 20 μg/mL chloramphenicol. 5 mL cultures from a single colony were grown overnight at 37°C in LB broth containing 50 μg/mL kanamycin and 20 μg/mL chloramphenicol shaking at 180 rpm. Cultures were diluted 50-fold into fresh LB medium supplemented with 50 μg/mL kanamycin and 20

μg/mL chloramphenicol, grown to OD_{600} of 0.3 at 37°C, shaking at 180 rpm. Gene expression of chaperones were induced by addition of L-arabinose (1 mg/mL) for cultures containing pGro7 or pKJE7 vectors and Tetracycline (10 ng/mL) for cultures containing pGTf2 vectors. Cultures grown to OD_{600} of 0.6 at 37°C, shaking at 180 rpm with gene induction of target fusion protein by addition of IPTG to a final concentration of 1 mM and incubated at 30°C for 6 hours or 25°C overnight, shaking at 180 rpm, after which cells were harvested by centrifugation (3500 g, 20 minutes, 4°C) and pellets frozen at -20°C. Cold shock response was induced by placing cultures in an ice bath prior to the addition of 0.5 mM or 1 mM IPTG at an OD_{600} of 0.1 or 0.3 then subsequent incubation at 18°C overnight, shaking at 180 rpm. Cells were then harvested by centrifugation (3500 g, 20 minutes, 4°C) and pellets frozen at -20°C.

Silicatein-α: The pHis vector containing the gene encoding for the mature portion of silicatein-α was transformed into *E. Coli* BL21(DE3) and grown at 37°C on LB agar plates supplemented with 100 μg/mL ampicillin. 5 mL cultures from a single colony was grown overnight at 37°C in LB containing 100 μg/mL ampicillin, shaking at 180 rpm. Cultures were diluted 50-fold into fresh LB medium supplemented with 100 μg/mL ampicillin and grown to OD600 of 0.5-0.6 at 37°C, shaking at 180 rpm. Gene expression was induced by addition of IPTG to a final concentration of 1 mM and incubated at 37°C for 4 hours shaking at 180 rpm, after which cells were harvested by centrifugation (3500 g, 20 minutes, 4°C) and pellets frozen at -20°C.

TF-Silicatein-α-Strep: The constructed vectors were transformed into *E. Coli* BL21(DE3) and grown at 37°C on LB agar plates supplemented with 100 μg/mL ampicillin. 5 mL cultures from a single colony were grown overnight at 37°C in LB broth containing 100 μg/mL ampicillin, shaking at 180 rpm. Cultures were diluted 50-fold into fresh LB medium or TB broth supplemented with 100 μg/mL ampicillin, grown to OD600 of 0.4-0.5 at 37°C shaking at 180 rpm. Gene expression was induced by addition of IPTG to a final concentration of 1 mM and incubated at 15°C overnight, shaking at 180 rpm, after which cells were harvested by centrifugation (3500 g, 20 minutes, 4°C) and pellets frozen at -20°C.

7.3.4 Cell lysis, SDS PAGE and Western blot analysis

Frozen pellets from overnight bacterial cultures were suspended in PBS buffer, SUMO lysis buffer for SUMO-silicatein-α and pHis lysis buffer 1 or pHis lysis buffer 2 for pHis Silicatein-α (Section 7.1). The cells were disrupted by sonication at 4°C for 8 cycles of 1 minute pulse and 30 seconds rest at 65% amplitude. A sample of lysate (total protein content) was reserved for protein solubility analysis by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE), the remaining lysate was separated into soluble and insoluble fractions by centrifugation (48200 g, 30 minutes, 4°C) with a sample of each prepared for SDS-PAGE analysis.[260] Following SDS-PAGE, the gel was placed in transfer buffer; proteins were then blotted onto PVDF membrane. The membrane was washed with PBS Tween (1X PBS, 0.05% Tween-20) and blocked with PBS Tween with 5% skimmed milk powder for 30 minutes then subsequently washed with PBS Tween. The washed blot was then incubated with 1:10000 dilution of monoclonal anti-poly histidine antibody (Sigma-Aldrich) for 1 hour. The membrane was washed three times in PBS Tween and incubated for 1 hour with the secondary antibody, HRP conjugated goat anti-mouse IgG at a dilution of 1:8000. The membrane was washed three times with PBS Tween and incubated with enhanced chemoiluminescence detection reagents (GE Healthcare) for 5 minutes then scanned using an X-ray film processor.

7.3.5 Protein purification TF-silicatein-α

Immobilised metal affinity chromatography (IMAC)

Soluble protein fractions from cell lysis were clarified using ultra-centrifugation as described in the previous section and filter sterilised with 0.22 µm filter. The clarified lysate was loaded onto a 5 mL His Trap Column (GE Healthcare) using ÄKTA FPLC and a step wise gradient was used with concentrations of 10 mM, 60 mM and 250 mM Imidazole in phosphate buffer (Section 4.1). A further step wise gradient was used with 30 mM imidazole in the loading buffer and increasing concentrations of 60 mM and 200 mM imidazole. Eluted fractions were identified from the chromatogram and analysed by SDS-PAGE. Buffer exchange to remove imidazole was conducted by dialysis with MWCO of 10K into the relevant buffer at 4°C overnight.

Ammonium sulphate precipitation optimisation

Precipitation experiments were conducted using varying concentrations of ammonium sulphate which were added to samples of clarified lysate to achieve percentage saturated ammonium sulphate solutions between 10-100% at 4°C and 25°C. The samples were placed on the roller mixer for 1 hour and centrifuged (8000 rpm, 30 mins). The supernatant was aspirated and the precipitate was resuspended in 1 X PBS. Samples from the supernatant and resuspended precipitate of each saturated ammonium sulphate concentration were analysed by SDS-PAGE.

Purification of fusion protein TF-silicatein-α using ammonium sulphate precipitation and IMAC.

Ammonium sulphate was added to clarified lysate containing soluble protein to achieve saturation of 60% and placed on the roller mixer for 1 hour then centrifuged (7720 g, 30 mins). The supernatant was aspirated and the precipitate resuspended in 1 X PBS. Following buffer exchange using dialysis at 4°C overnight the sample was loaded onto a 5 mL His Trap Column (GE Healthcare) using ÄKTA FPLC and a step wise gradient with buffer A and buffer B was used with concentrations of 10 mM, 60 mM and 250 mM imidazole. Eluted fractions were identified from the chromatogram and analysed by SDS-PAGE. Following buffer exchange by dialysis (4°C, overnight, 10 K MWCO), a second IMAC purification step was conducted with elution fraction from the first step using the aforementioned stepwise gradient with the addition of 30 mM imidazole in the loading buffer .

7.3.6 Refolding silicatein-α

Frozen pellets from overnight bacterial cultures were resuspended in buffer A (Section 7.1). Cells were disrupted by sonication at 4°C for 8 cycles of 1 minute pulse and 30 seconds rest at 65% amplitude. The lysate was separated into soluble and insoluble fractions by centrifugation (4000 g, 20 minutes, 4°C) soluble fraction discarded and the insoluble fraction washed in buffer B then centrifuged (4000g rpm, 20 minutes, 4°C), supernatant was discarded and the pellet solubilised in buffer C. Sample was loaded onto a 5 mL HisTrap Column (GE

Healthcare) using ÄKTA FPLC and a step wise gradient was used with concentrations of 10 mM, 30 mM, 60 mM and 250 mM imidazole. Elution fraction along with total protein, soluble and insoluble fractions following cell lysis was analysed by SDS-PAGE.[260] Refolding by dialysis and inside HisTrap column was attempted with refolding buffer C1 without urea[69] or with refolding buffer containing 3 M urea then without urea in a stepwise manner. A sample from each of the refolding conditions was taken and centrifuged (7700 rpm, 15 minutes) with SDS-PAGE analysis of the supernatant and pellet.[260] The method described above was used for slow dilution with the exception of the stepwise reduction in the concentration of urea using refolding buffer C1. For slow dilution this buffer without urea was added to the solubilised protein in a slow linear gradient (0.5 ml min^{-1}) and analysis of samples was conducted in the same manner as before.

7.3.7 Protein Purification of silicatein-α

Following cell lysis with pHis lysis buffer 1 or pHis lysis buffer 2 the soluble fraction was loaded onto a gravity flow column with 5 mL Ni^{2+}NTA agarose gel. A step wise gradient was used to wash and elute proteins with IMAC buffer A and concentrations of 10 mM, 60 mM and 250 mM Imidazole in phosphate buffer (Section 7.1)

7.3.8 Protein Purification of TF-Silicatein-α-Strep

Clarified lysate containing soluble protein was loaded on to 5 mL StrepTrap™HP in binding buffer (100 mM Tris-HCl, 150 Mm NaCl, 1 mM EDTA, pH 8) and washed with 10 column volumes of binding buffer. Proteins were eluted with 6 column volumes of elution buffer (Binding buffer with 2.5 mM desthiobiotin). Eluted fractions containing the desired protein were then exchanged into the relevant buffer for protease cleavage studies by dialysis (10 K MWCO).

7.3.9 Proteolytic Cleavage of TF-silicatein-α and TF-silicatein-α-Strep

Cleavage of TF-silicatein-α using Factor Xa at 4°C and 25°C

100 µg of purified fusion protein was incubated with 0.5 µL, 1 µL, 2 µl, 3 µL and 3.5 µL of Factor Xa protease (New England Biolabs) in reaction buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl) for 16 hours at 4°C or 25°C. Following this, SDS-PAGE analysis and native PAGE analysis of each reaction was conducted.

Cleavage of TF-silicatein-α and TF-silicatein-α-Strep using HRV3C and GST-HRV3C

Small scale optimization experiments to estimate the appropriate ratio of HRV3C protease were preformed by using enzyme:target protein with ratios (unit/µg) of 1:5, 1:25 and 1:50 and recombinant GST-HRV3C protease enzyme:target protein with ratios 1:20, 1:30, 1:40, 1:50 and 1:60 in protease cleavage buffer (150 mM NaCl, 50 mM Tris-HCl, pH 7.5) at increasing incubation times of 1, 3 and 16 hours. Samples were analysed by SDS-PAGE.[25]

Cleavage of thrombin TF-silicatein-α and TF-Silicatein-α-Strep

Small scale optimization experiment to estimate the appropriate ratio of Thrombin protease enzyme:target protein with ratios (unit/µg) of 1:5, 1:25 and 1:50 in protease cleavage buffer (20 mM Tris-HCl pH 8.5, 150 mM NaCl, 2.5 mM $CaCl_2$) at increasing incubation times of 1, 3 and 16 hours. Samples were analysed by SDS-PAGE.

7.3.10 Size Exclusion Chromatography

Following purification by IMAC further purification by size exclusion chromatography (SEC) with Superdex HiLoad 16/600 200 pg (GE Healthcare) at a flow rate of 0.2 mL min-1 with relevant buffers (Section 7.1)

7.4 Experimental Methodology for Chapter 3

7.4.1 Plasmid construction of protein variants

Plasmid was constructed by site directed mutagenesis using two step PCR [261]with TF-Silicatein-α in pCold vector for DNA template

TF-silicatein-α (Ser26Cys):

S26C Forward 5’ –GATCAGGGTGATTGCGGGGCCTGTTATGCTTTTTCTGCAATGGGCG - 3’ and GA-REVERSE: 5’ - GGCACCGGGATCTCGACC– 3’

TF-silicatein-α (His165Ala):

H165A Forward 5’ – GCTGTTCATCATCAAGCCTGAATGCAGCAATGGTCGTAACGGGC - 3’

GA-REVERSE: 5’ - GGCACCGGGATCTCGACC– 3’

TF-silicatein-α (Asn185Ala):

N185A Forward 5’ –GCAAGAAATACTGGCTGGCCAAAGCTTCGTGGGGGACCAATTGG – 3’

GA-REVERSE: 5’ - GGCACCGGGATCTCGACC– 3’

TF-silicatein-α (Asn185Asp) :

N185D Forward 5'-GCAAGAAATACTGGCTGGCCAAAGATTCGTGGGGGACCAATTGG-3'

GA-REVERSE: 5’ - GGCACCGGGATCTCGACC– 3’

Mutations were confirmed by sequencing. This work was carried out by Teh Ser Huy . The protein of interest was expressed as a fusion protein with a hexahistidine tag followed by Trigger Factor at the N-terminus.

TF-silicatein-α (Ser26Ala): Molecular cloning of this variant was by overlap extension method [262] with primers:

Non-mutagenic Forward: 5’ - CCACGCGGTAGTGGTGGTATC -3’

Non-mutagenic Reverse: 5’ – TACCTATCTAGACTGCAGGTC -3’

Mutagenic Forward: 5’ - GGGGCCGCATATGCTTTTTCT -3’

Mutagenic Reverse: 5’ - AGAAAAAGCATATGCGGCCCC -3’

Mutations were confirmed by sequencing. This work was carried out by Seyed Yasin Tabatabaei Dakhili.

7.4.2 Biochemical analysis of silicatein-α using SMAA hydrolysis

TF-silicatein-α heat denatured sample was prepared by heating to 85°C for 10 minutes and cooled to room temperature. Samples containing silicatein-α, variants or heat denatured protein (2.7 µM) in 50 mM Tris, 100 mM NaCl, pH 8 were equilibrated to 22°C before the start of the reaction. 0.5 mmol of TEOS was added to the samples, mixed thoroughly and shaken at 22°C for 1 hour. The samples were then centrifuged (13000 g) for 30 minutes to pellet the precipitated silica and the supernatant was discarded. Precipitated silica was subsequently washed four times with 200 µL of 100% EtOH and 2M NaOH was added to a final volume of 500 µL. The amount of precipitated silica was quantified using the Silicate (Silicic Acid) Test (Merck, UK) according to the manufacturer's protocol. Calibration of the assay was conducted using Na_2SiF_6 (0.960 g L^{-1}, equivalent to 5 µg mL^{-1} of silica) as a standard. [192]

7.4.3 Biochemical analysis of silicatein-α using SMAA condensation

Acidic hydrolysis of TEOS to produce silicic acid solution performed according to previously described methods.[69,263] 5 mM of this prehydrolysed TEOS was added to all samples, adjusted to pH 7 and mixed thoroughly. The reaction proceeded for 1 hour at 22°C, shaking at 850 rpm. The samples were then processed using the method described above to collect and quantify the precipitated silica.

7.4.4 Biochemical analysis of silicatein-α using hydrolysis of TBDMS-4np

Substrate stock solutions (1 mM) of TBDMS-4-nitrophenoxysilane, TBDMS-4-nitrophenyl ester and TBDMS-4-nitrophenylamide were prepared in 50 mM Tris, 100 mM NaCl, 10% v/v 1,4-dioxane and adjusted to the relevant pH. The desired concentration of substrate was added to a microtitre plate with the appropriate volume of the same buffer at the required pH added to a final volume of 100 µL. 100 µL of enzyme sample at the desired concentration was then added to each well. Absorbance was measured at 405 nm by UV-Vis spectrophotometer every three minutes over 6 hours at 22°C. The plate was shaken continuously throughout the reaction. Calibration of the assay used 4-nitrophenol or 4-nitroaniline as standards at the relevant pH. Each assay was performed in triplicate and the

background hydrolysis from the control was substrated from the enzyme samples. The time course data for each experiment was averaged. Initial velocities obtained from the tangent of the progress curve using the linear regression technique to produce the tangent of the slope. SigmaPlot 12.0 software was used to process the data and calculate V_{max}, K_M and k_{cat} using the Michaelis-Menten model of enzyme kinetics algorithm.

7.5 Experimental Methodology for Chapter 4

7.5.1 Circular Dichroism (CD) measurements and secondary structure prediction

The spectra were recorded using Chirascan CD Spectrometer (Applied Photophysics Limited, UK). Quartz cells with 0.1 mm pathlength (Starna Scientific, UK) were used for all measurements. Data was collected over a wavelength range of 190 to 260 nm, at a scan speed of 0.5 nm^{s-1}. The temperature was maintained at 22°C using a temperature controlled chamber. Each protein was analysed in buffered solution (50 mM Tris, 100 mM NaCl, pH 8.5) with baseline measurements using buffer alone. In all cases, the proteins were analysed at a concentration of 9.5 mM. An averaged spectrum of each sample was obtained by 20 repeat scans and subtraction of the mean baseline. The raw CD data was then used to calculate MRE. The secondary structure content of each protein was estimated with CDNN algorithm, [264] using MRE data.

The theoretical values for TF-silicatein-α secondary structure prediction in Table S3 were calculated using SWISS-MODEL and Swiss PdbViewer.[265] The data was obtained using the values given in PDBsum (www.ebi.ac.uk/pdbsum/) for trigger factor (PDB ref. 1w26) and a homology model constructed using cathsilicatein (PDB ref. 2vhs) for silicatein separately. The theoretical percentage values for each secondary structure in TF-silicatein-α were then calculated from these separate values as a proportion of their relative molecular weight within the fusion protein.

7.5.2 Dynamic Light Scattering

All protein samples were prepared in the relevant buffer by dialysis and filtered through a 0.02 µm syringe (Whatman Anotop, Sigma Aldrich) then diluted to the desired concentration with relevant buffer (Table 7.7). For the concentration dependent measurements samples to

the concentration of 0.5 mg/mL, 1 mg/mL, 1.5 mg/mL and 2 mg/mL of TF-Silicatein-α were prepared and the thermal ramp experiment used samples of 2 mg/mL. 40 µL of each sample and control were plated in a 384 well plate (Nunc DeepWell, ThermoFisher Scientific), and repeated in triplicate. Dynamic light scattering measurements (diffusion coefficient and radius of hydration) were made using a Dynapro plater reader with a laser wavelength of 830 nm and scattering angle of 158°. Acquisition time was 5 seconds with 10 acquisitions taken for each sample during this time. Data analysis was conducted with DYNAMICS 7.0 software which uses the DYNAL algorithm to determine correlation functions using cumulant analysis and regularisation analysis and subsequent polydispersity values.[266,267]

Table 7.7. Buffer conditions screened for optimisation experiments using Dynamic Light Scattering

Buffer	Additive	pH
50 mM Tris	100 mM Ammonium Sulphate + 100 mM NaCl	pH 8.5
50 mM Tris	200 mM Ammonium Sulphate + 100 mM NaCl	pH 8.5
50 mM Tris	300 mM Ammonium Sulphate + 100 mM NaCl	pH 8.5
50 mM Tris	10 mM Arginine + Glutamate + 100 mM NaCl	pH 8.5
50 mM Tris	30 mM Arginine + Glutamate + 100 mM NaCl	pH 8.5
50 mM Tris	50 mM Arginine + Glutamate + 100 mM NaCl	pH 8.5
50 mM Tris	No NaCl	pH 8.5
50 mM Tris	300 mM NaCl	pH 8.5
50 mM Tris	100 mM NaCl	pH 8.5
50 mM Tris	100 mM NaCl	pH 7.5
50 mM Citrate	100 mM NaCl	pH 3.5

7.5.3 Static Light Scattering

Silicatein-α and TF-silicatein-α were dialysed into the relevant buffer (Table 7.8) and adjusted to a concentration of 0.3 mg/mL. 9 µL of each sample was loaded into the multi-micro cell array in triplicate. The thermal ramp range was 20-90 °C with an increase of 0.3 °C per minute. Initial and final DLS measurements were also taken at four acquisitions at every 5 seconds. The intensity of scattered light at 266 nm and 473 nm was used to calculate T_{agg} and the first derivative of the barycentric mean (BCM) was used to calculate T_m with UNcle software with DLS measurements (R_h and PDI) calculated with the software correlation function.

Table 7.8. Buffer conditions screened for optimisation experiments using Static Light Scattering thermal ramp and Dynamic Light Scattering at 22 °C

Buffer	Additive	pH
50 mM Bis Tris	50 mM Arg + Glu + 100 mM NaCl	pH 8.5
100 mM HEPES	50 mM Arg + Glu + 100 mM NaCl	pH 8.5
100 mM MOPS	50 mM Arg + Glu + 100 mM NaCl	pH 8.5
100 mM EPPS	50 mM Arg + Glu + 100 mM NaCl	pH 8.5
100 mM BICINE	50 mM Arg + Glu + 100 mM NaCl	pH 8.5
50 mM Phosphate	50 mM Arg + Glu + 300mM NaCl	pH 8.5
50 mM Tris	50 mM Arg + Glu + 100 mM NaCl	pH 8.5
50 mM Tris	50 mM Arg + Glu + 300 mM NaCl	pH 8.5
50 mM Tris	50 mM Arg + Glu + 500 mM NaCl	pH 8.5
50 mM Tris	100 mM NaCl	pH 8.5
50 mM Tris	100 mM NaCl	pH 7.5
50 mM Citrate	100 mM NaCl	pH 3.5

7.5.4 Multi-Angled Light Scattering

TF-silicatein-α and silicatein-α were dialysed into the relevant buffer and adjusted to the required concentration and loaded onto Superdex S200 10/300 GL Column at a flow rate of 0.750 mL/min. The elution fractions were then passed through DAWN Heleos II EOS 18 angle laser photometer with QELS detector (Wyatt Technologies) and Optilab rEX refractive index detector. ASTRA 5.6 software was used to analyse data and calculate molecular mass, concentration and polydispersity.

7.6 Experimental Methodology for Chapter 5

7.6.1 Condensation reaction in organic solvent, silyl etherification of and transetherification with TMS-OH and TES- OH

Purified TF-silicatein-α from section 2.2.13 was buffer exchanged by unltracentrifugation into 20 mM KCl in 100 mM potassium phosphate buffer at pH 7 and adjusted to the concentration of 5 mg mL^{-1}. 18-Crown-6 was added to give an enzyme solution containing 2 mM of this compound. 5 mL of the solution was subsequently flash frozen in liquid nitrogen and freeze-dried overnight to produce lyophilsed enzyme powder. Negative control samples were prepared in the same way omitting the enzyme. The lyophilised samples were added to 5 mL

n-octane, heated to required temperature and mixed until resuspended. 1.26 mmol of the relevant alcohol was added followed by 5 mol equiv. of silanol or ethoxysilane relative to the alcohol. The mixture was continuously stirred and sampled (100 µL) in triplicate at 24 hour time points over 72 hours. Each sample was centrifuged (13000 g for 10 min) and the supernatant was analysed by GC-MS (Table 7.9;Table 7.10).

7.6.2 Oxidation reaction with ADH in aqueous buffer

Each ADH enzyme that was tested was reconstituted in 20 mM Tris pH 7.4 and adjusted to a final concentration of 2 mg mL^{-1}. 1.5 mM of NAD^+ was added to the samples followed by 35 mM of the relevant alcohol. Samples were continuously shaken for 6 minutes at 25 °C. 1 mL samples were taken at one minute time intervals and analysed by UV-vis spectrophotometer. The increase in absorbance at 340 nm was recorded. Negative controls were prepared in the same manner omitiing the enzyme. Background absorbance signals were subtracted from the enzyme samples.

7.6.3 Reduction reaction with ADH in aqueous buffer

Each ADH enzyme that was tested was reconstituted in 20 mM sodium phosphate pH 6.0 and adjusted to a final concentration of 2 mg mL^{-1}. 1.5 mM of NADH was added to the samples plus 35 mM of the relevant aldehyde or ketone. 1 mL samples were continuously shaken for 6 minutes at 25 °C. Samples were taken at one minute time intervals and analysed by UV-vis spectrophotometer. The increase in absorbance at 340 nm was recorded. Negative controls were prepared in the same manner omitiing the enzyme. Background absorbance signals were subtracted from the enzyme samples.

7.6.4 Oxidation reaction with ADH in organic solvent

Each ADH enzyme that was tested was reconstituted in 20 mM Tris pH 7.4 and adjusted to the concentration of 2 mg mL^{-1} with 1.5 mM NAD^+ The solution was subsequently flash frozen in liquid nitrogen and freeze-dried overnight to produce lyophilsed enzyme powder. Negative control samples were prepared in the same way omitting the enzyme. The lyophilised samples

were added to 5 mL n-octane at the required temperature and mixed until resuspended. 35 mM of the relevant alcohol was added. The mixture was continuously stirred and sampled (100 µL) in triplicate at 24 hour time points over 72 hours. Each sample was centrifuged (13000 g for 10 min) and the supernatant was analysed by GC-MS.

7.6.5 Reduction reaction with ADH in organic solvent

Each ADH enzyme that was tested was reconstituted in 20 mM sodium phosphate pH 6.0 and adjusted to the concentration of 2 mg mL^{-1} with 1.5 mM NADH The solution was subsequently flash frozen in liquid nitrogen and freeze-dried overnight to produce lyophilsed enzyme powder. Negative control samples were prepared in the same way omitting the enzyme. The lyophilised samples were added to 5 mL n-octane at the required temperature and mixed until resuspended. 35 mM of the relevant aldehyde or ketone was added. The mixture was continuously stirred and sampled (100 µL) in triplicate at 24 hour time points over 72 hours. Each sample was centrifuged (13000 g for 10 min) and the supernatant was analysed by GC-MS.

7.6.6 GC-MS parameters

GC-MS analysis were performed using the following parameters (Table 7.9;Table 7.10).

Table 7.9 GC-MS parameters used to analyse silylation of various compounds with TES.

Parameter	Setting
Instrument	Thermo Finnigan PolarisQ Ion Trap Mass Spectrometer
Carrier gas	99.9995% Ultra high purity (UHP) helium
GC inlet, split	200 °C, split flow 10 (min mL^{-1}), split ratio 10
Detector	PolarisQ positive ion source (260 °C), full scan (35-650 amu), 0.59 scan/s
MS ionisation	Electron ionisation
GC temperature program	60 (2 min) – 250 (4 min) at 20 °C min^{-1}, 15 min total run time
GC column	Thermo Scientific TraceGold TG-1MS (30 m x 0.25 mm, 0.25 µm film)
GC transfer line temperature	275 °C
Volume injected	1 µL

Table 7.10 GC-MS parameters used to analyse TMS silylation of 1-octanol.

Parameter	Setting
Instrument	Varian Saturn 2000 Ion Trap Mass Spectrometer
Carrier gas	99.998% Ultra high purity (UHP) helium
GC inlet, split	200 °C, split flow 10 (min mL^{-1}), split ratio 10
Detector	Saturn 2000 positive ion source (260 °C), full scan (35-650 amu), 0.59 scan/s
MS ionisation	Electron ionisation
GC temperature program	60 (2 min) – 250 (4 min) at 20 °C min^{-1}, 15.5 min total run time
GC column	Zebron ZB-Semi Volatiles (30 m x 0.25 mm, 0.25 µm film)
GC transfer line temperature	275 °C
Volume injected	1 µL

8. Chapter 8: References

1. Cha, J. N., Shimizu, K., Zhou, Y., Christiansen, S. C., Chemlka, B. F., Stucky, G. D. and Morse, D. E. Silicatein filaments and subunits from a marine sponge direct the polymerization of silica and silicones in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 361–5 (1999).

2. Schröder, H. C., Wiens, M., Schloßmacher, U., Brandt, D. & Müller, W. E. G. Silicatein-Mediated Polycondensation of Orthosilicic Acid: Modeling of a Catalytic Mechanism Involving Ring Formation. *Silicon* **4,** 33–38 (2012).

3. Graiver, D., Farminer, K. W. & Narayan, R. A Review of the Fate and Effects of Silicones in the Environment. *J. Polym. Environ.* (2003).

4. Sommers, M. A. *Silicon*. (The Rosen Publishing Group, 2008).

5. Exley, C. Silicon in life:A bioinorganic solution to bioorganic essentiality. *J. Inorg. Biochem.* **69,** 139–144 (1998).

6. Itoh, K. M., Kato, J., Uemura, M., Kaliteevskii, A. K., Godisov, O. N., Devyatych, G. G., Bulanov, A. D., Gusev, A. V., Kovalev, I. D., Sennikov, P. G., Pohl, H., Abrosimov, N. V. and Riemann, H. High Purity Isotopically Enriched ^{29}Si and ^{30}Si Single Crystals: Isotope Separation, Purification, and Growth. *Japanese J. Appl. Physics, Part 1 Regul. Pap. Short Notes Rev. Pap.* (2003).

7. Delak, K. M., Farrar, T. C. & Sahai, N. ^{29}Si NMR sensitivity enhancement methods for the quantitative study of organosilicate hydrolysis and condensation. *J. Non. Cryst. Solids* (2005).

8. Voronkov, M. G. History of Chemistry Silicon Era. *ISSN Russ. J. Appl. Chem.* **80,** 1070–4272 (2007).

9. Greenwood, N. N. and Earnshaw, A. *Chemistry of the Elements*. (Elsevier Butterworth-Heinemann, 2003).

10. Astruc, D. *Organometallic Chemistry and Catalysis*. (Springer-Verlag, 2007).

11. Pierrefixe, S. C. A. H., Fonseca Guerra, C. & Bickelhaupt, F. M. Hypervalent Silicon versus Carbon: Ball-in-a-Box Model. *Chem. A Eur. J.* **14,** 819–828 (2008).

12. Brook, M. A. *Silicon in organic, organometallic, and polymer chemistry*. (Wiley & Sons Inc., 2000).

13. Gau, D., Kato, T., Saffon-Merceron, N., Cózar, A., Cossio, F. P., and Baceiredo, A. Synthesis and structure of a base-stabilized C-phosphino-Si-amino silyne. *Angew. Chemie-Int. Ed.* **49,** 6585–6588 (2010).

14. Davy, J. An Account of Some Experiments on Different Combinations of Fluoric Acid. *Philos. Trans. R. Soc. London* **102,** 352–369 (1812).

15. Lewis, G. N. The Atom and the Molecule. *J. Am. Chem. Soc.* **38,** 762–785 (1916).

16. Pierrefixe, S. C. A. H. & Bickelhaupt, F. M. Hypervalence and the delocalizing versus

localizing propensities of H_3, Li_3, CH_5 and SiH_5. *Struct. Chem.* **18,** 813–819 (2007).

17. Pimentel, G. C. The Bonding of Trihalide and Bifluoride Ions by the Molecular Orbital Method. *J. Chem. Phys.* **19,** 446–448 (1951).

18. Hach, R. J. & Rundle, R. E. The Structure of Tetramethylammonium Pentaiodide1,1a. *J. Am. Chem. Soc.* **73,** 4321–4324 (1951).

19. Hoffmann, R., Howell, J. M. & Muetterties, E. L. Molecular orbital theory of pentacoordinate phosphorus. *J. Am. Chem. Soc.* **94,** 3047–3058 (1972).

20. Thomas, N. R. Frederic Stanley Kipping—Pioneer in Silicon Chemistry: His Life & Legacy. *Silicon* **2,** 187–193 (2010).

21. Moretto, H. H., Schulze, M. & Wagner, G. in *Ullmann's Encyclopedia of Industrial Chemistry* (Wiley-VCH Verlag GmbH & Co. KGaA, 2000).

22. Hyde, J. F. & DeLong, R. C. Condensation Products of the Organo-silane Diols. *J. Am. Chem. Soc.* **63,** 1194–1196 (1941).

23. Hatanaka, Y. & Hiyama, T. Cross-coupling of organosilanes with organic halides mediated by a palladium catalyst and tris(diethylamino)sulfonium difluorotrimethylsilicate. *J. Org. Chem.* **53,** 918–920 (1988).

24. Denmark, S. E. & Regens, C. S. Palladium-Catalyzed Cross-Coupling Reactions of Organosilanols and Their Salts: Practical Alternatives to Boron and Tin Based Methods. *Acc. Chem. Res.* **41,** 1486–1499 (2008).

25. Shi, S. & Zhang, Y. Pd(OAc)2-Catalyzed Fluoride-Free Cross-Coupling Reactions of Arylsiloxanes with Aryl Bromides in Aqueous Medium. *J. Org. Chem.* **72,** 5927–5930 (2007).

26. Nakao, Y., Takeda, M., Matsumoto, T. & Hiyama, T. Cross-Coupling Reactions through the Intramolecular Activation of Alkyl(triorgano)silanes. *Angew. Chemie* **122,** 4549–4552 (2010).

27. Wolf, C. & Lerebours, R. Palladium−Phosphinous Acid-Catalyzed NaOH-Promoted Cross-Coupling Reactions of Arylsiloxanes with Aryl Chlorides and Bromides in Water. *Org. Lett.* **6,** 1147–1150 (2004).

28. Crouch, R. D. Recent Advances in Silyl Protection of Alcohols. *Synth. Commun.* **43,** 2265–2279 (2013).

29. Corey, E. J. & Venkateswarlu, A. Protection of hydroxyl groups as tert-butyldimethylsilyl derivatives. *J. Am. Chem. Soc.* **94,** 6190–6191 (1972).

30. Corey, E. J., Cho, H., Rucker, C. & Hua, D. H. Studies with trialkylsilyltriflates: new syntheses and applications. *Tetrahedron Lett.* **22,** 3455–3458 (1981).

31. Corey, E. J., Tius, M. A. & Das, J. Total synthesis of (+)-aphidicolin. *J. Am. Chem. Soc.* **102,** 1742–1744 (1980).

32. Chaudhary, S. K. & Hernandez, O. 4-dimethylaminopyridine: an efficient and selective catalyst for the silylation of alcohols. *Tetrahedron Lett.* **20,** 99–102 (1979).

33. Ogilvie, K. K. The tert-Butyldimethylsilyl Group as a Protecting Group in Deoxynucleosides. *Can. J. Chem.* **51,** 3799–3807 (1973).

34. Ogilvie, K. K., Thompson, E. A., Quilliam, M. A. & Westmore, J. B. Selective protection of hydroxyl groups in deoxynucleosides using alkylsilyl reagents. *Tetrahedron Lett.* **15,** 2865–2868 (1974).

35. Park, S. Y., Lee, J.-W. & Song, C. E. Parts-per-million level loading organocatalysed enantioselective silylation of alcohols. *Nat. Commun.* **6,** 7512 (2015).

36. Isobe, T., Fukuda, K., Araki, Y. & Ishikawa, T. Modified guanidines as chiral superbases: the first example of asymmetric silylation of secondary alcohols. *Chem. Commun.* 243–244 (2001).

37. Zhao, Y., Rodrigo, J., Hoveyda, A. H. & Snapper, M. L. Enantioselective silyl protection of alcohols catalysed by an amino-acid-based small molecule. *Nature* **443,** 67–70 (2006).

38. Rendler, S., Auer, G. & Oestreich, M. Kinetic Resolution of Chiral Secondary Alcohols by Dehydrogenative Coupling with Recyclable Silicon-Stereogenic Silanes. *Angew. Chemie Int. Ed.* **44,** 7620–7624 (2005).

39. Rodrigo, J. M., Zhao, Y., Hoveyda, A. H. & Snapper, M. L. Regiodivergent Reactions through Catalytic Enantioselective Silylation of Chiral Diols. Synthesis of Sapinofuranone A. *Org. Lett.* **13,** 3778–3781 (2011).

40. You, Z., Hoveyda, A. & Snapper, M. Catalytic Enantioselective Silylation of Acyclic and Cyclic Triols: Application to Total Syntheses of Cleroindicins D, F, and C. *Angew. Chemie Int. Ed.* **48,** 547–550 (2009).

41. Manville, N., Alite, H., Haeffner, F., Hoveyda, A. H. & Snapper, M. L. Enantioselective silyl protection of alcohols promoted by a combination of chiral and achiral Lewis basic catalysts. *Nat Chem* **5,** 768–774 (2013).

42. Sridhar, M., Raveendra, J., China Ramanaiah, B. & Narsaiah, C. An efficient synthesis of silyl ethers of primary alcohols, secondary alcohols, phenols and oximes with a hydrosilane using InBr3 as a catalyst. *Tetrahedron Lett.* **52,** 5980–5982 (2011).

43. Young, C. L., Britton, Z. T. & Robinson, A. S. Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnol. J.* **7,** 620–634 (2012).

44. Young, I. S. & Baran, P. S. Protecting-group-free synthesis as an opportunity for invention. *Nat Chem* **1,** 193–205 (2009).

45. Hoffmann, R. W. Protecting-Group-Free Synthesis. *Synthesis (Stuttg).* **2006,** 3531–3541 (2006).

46. Baran, P. S., Maimone, T. J. & Richter, J. M. Total synthesis of marine natural products without using protecting groups. *Nature* **446,** 404–408 (2007).

47. Biffis, A., Braga, M. & Basato, M. Solventless Silane Alcoholysis Catalyzed by Recoverable Dirhodium(II) Perfluorocarboxylates. *Adv. Synth. Catal.* **346,** 451–458

(2004).

48. Rendler, S. *et al.* Stereoselective Alcohol Silylation by Dehydrogenative Si-O Coupling: Scope, Limitations, and Mechanism of the Cu-H-Catalyzed Non-Enzymatic Kinetic Resolution with Silicon-Stereogenic Silanes. *Chem. - A Eur. J.* **14,** 11512–11528 (2008).

49. Weickgenannt, A., Mewald, M., Muesmann, T. W. T. & Oestreich, M. Catalytic Asymmetric Si-O Coupling of Simple Achiral Silanes and Chiral Donor-Functionalized Alcohols. *Angew. Chemie Int. Ed.* **49,** 2223–2226 (2010).

50. Dong, X., Weickgenannt, A. & Oestreich, M. Broad-spectrum kinetic resolution of alcohols enabled by Cu–H-catalysed dehydrogenative coupling with hydrosilanes. **8,** 15547 (2017).

51. Abbate, V., Bassindale, A. R., Brandstadt, K. F. & Taylor, P. G. A large scale enzyme screen in the search for new methods of silicon-oxygen bond formation. *J. Inorg. Biochem.* **105,** 268—275 (2011).

52. Abbate, V., Brandstadt, K. F., Taylor, P. G. & Bassindale, A. R. Enzyme-Catalyzed Transetherification of Alkoxysilanes. *Catalysts* **3,** 27–35 (2013).

53. Tabatabaei Dakhili, S. Y. Biocatalysis in Organosiloxane Chemistry using Silicateins. (University of Manchester, 2017).

54. Nelson, T. D. & Crouch, R. D. Selective Deprotection of Silyl Ethers. *Synthesis (Stuttg).* **1996,** 1031–1069 (1996).

55. Crouch, R. D. Selective deprotection of silyl ethers. *Tetrahedron* **69,** 2383–2417 (2013).

56. Kaburagi, Y. & Kishi, Y. Operationally Simple and Efficient Workup Procedure for TBAF-Mediated Desilylation: Application to Halichondrin Synthesis. *Org. Lett.* **9,** 723–726 (2007).

57. Fustero, S., García Sancho, A., Aceña, J. L. & Sanz-Cervera, J. F. Fluorous TBAF: A Convenient and Selective Reagent for Fluoride-Mediated Deprotections. *J. Org. Chem.* **74,** 6398–6401 (2009).

58. Perry, C. C. & Keeling-Tucker, T. Biosilicification: The role of the organic matrix in structure control. *Journal of Biological Inorganic Chemistry* **5,** 537–550 (2000).

59. Müller, W. E. G., Krasko, A., Le Pennec, G. & Schröder, H. C. Biochemistry and cell biology of silica formation in sponges. *Microsc. Res. Tech.* **62,** 368–377 (2003).

60. Uriz, M. J., Turon, X., Becerro, M. A. & Agell, G. Siliceous spicules and skeleton frameworks in sponges: Origin, diversity, ultrastructural patterns, and biological functions. *Microsc. Res. Tech.* **62,** 279–299 (2003).

61. Shimizu, K., Cha, J., Stucky, G. D. & Morse, D. E. Silicatein α: Cathepsin L-like protein in sponge biosilica. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 6234–6238 (1998).

62. Veremeichik, G. N., Shkryl, Y., Bulgakov, V., Shedko, S., Kozhemyako, V., Kovalchuk, S., Krasokhin, V., Zhuravlev, Y and Kulchin, Y. Occurrence of a Silicatein Gene in Glass Sponges (Hexactinellida: Porifera). *Mar. Biotechnol.* **13,** 810–819 (2011).

63. Pozzolini, M., Sturla, L., Cerrano, C., Bavestrello, G., Camardella L., Parodi, A. M., Raheli, F., Benatti, U., Müller, W. E. G and Giovine, M. Molecular Cloning of Silicatein Gene from Marine Sponge Petrosia ficiformis (Porifera, Demospongiae) and Development of Primmorphs as a Model for Biosilicification Studies. *Mar. Biotechnol.* **6,** 594–603 (2004).

64. Krasko, A., Lorenz, B., Batel, R. Schröder, H and Müller, W. E. G. Expression of silicatein and collagen genes in the marine sponge Suberites domuncula is controlled by silicate and myotrophin. *Eur. J. Biochem.* **267,** 4878–4887 (2000).

65. Schröder, H. C., Wang, X., Manfrin, A., Yu, S., Grenbenjuk, V., Korzhev, M., Wiens, M., Schlossmacher, U. and Müller. W. E. G. Acquisition of structure-guiding and structure-forming properties during maturation from the pro-silicatein to the silicatein form. *J. Biol. Chem.* **287,** 22196–205 (2012).

66. Schröder, H. C., Boreiko, A., Korzhev, M., Tahir, M., Tremel, W., Eckert, C., Ushi, H., Müller, I. and Müller, W. Co-expression and functional interaction of silicatein with galectin: matrix-guided formation of siliceous spicules in the marine demosponge Suberites domuncula. *J. Biol. Chem.* **281,** 12001–9 (2006).

67. Armirotti, A., Damonte, G., Pozzolini, M,. Mussino, F., Cerrano, C., Salis, A., Benatti, U. and Glovine, M. Primary Structure and Post-Translational der Modifications of Silicatein Beta from the Marine Sponge Petrosia ficiformis (Poiret, 1789). *J. Proteome Res.* **8,** 3995–4004 (2009).

68. Müller, W. E. G., Rothenberger, M., Boreiko, A., Tremel, W., Reiber, A. and Schröder, H. Formation of siliceous spicules in the marine demosponge Suberites domuncula. *Cell Tissue Res.* **321,** 285–297 (2005).

69. Schloßmacher, U., Wiens, M., Schröder, H., Wang, X., Jochum, K. and Müller Silintaphin-1 interaction with silicatein during structure-guiding bio-silica formation. *FEBS J.* **278,** 1145–1155 (2011).

70. Wiens, M., Schröder, H., Wang, X., Link, T., Steindorf, D. and Müller, W. Isolation of the Silicatein-α Interactor Silintaphin-2 by a Novel Solid-Phase Pull-Down Assay. *Biochemistry* **50,** 1981–1990 (2011).

71. Fairhead, M. Johnson, K., Kowatz, T., McMahon, S., Carter, L., Oke, M., Liu, H., Naismith, J.and van der Walle C. Crystal structure and silica condensing activities of silicatein [small alpha]-cathepsin L chimeras. *Chem. Commun.* 1765–1767 (2008).

72. Zhou, Y., Shimizu, K., Cha, J. N., Stucky, G. D. & Morse, D. E. Efficient Catalysis of Polysiloxane Synthesis. *Angew. Chem. Int. Ed. Engl.* 779–782 (1999).

73. Miesfeld, R. L. & McEvoy, M. M. *Biochemistry*. (W.W.Norton & Company Ltd, London, 2017).

74. Page, M. J. & Di Cera, E. Serine peptidases: Classification, structure and function. *Cellular and Molecular Life Sciences* **65,** 1220–1236 (2008).

75. Di Cera, E. Serine proteases. *IUBMB Life* **61,** 510–515 (2009).

76. Hedstrom, L. Serine protease mechanism and specificity. *Chem. Rev.* **102,** 4501–4523

(2002).

77. Hartley, B. S. & Kilby, B. A. The reaction of p-nitrophenyl esters with chymotrypsin and insulin. *Biochem. J.* **56,** 288–297 (1954).

78. Tabatabaei Dakhili, S. Y. *et al.* Recombinant silicateins as model biocatalysts in organosiloxane chemistry. *Proc. Natl. Acad. Sci.* **114,** 5285–5291 (2017).

79. Reetz, M. T. Biocatalysis in Organic Chemistry and Biotechnology: past, Present, and Future. *Jacs* **135,** 12480–12496 (2013).

80. Ran, N., Zhao, L., Chen, Z. & Tao, J. Recent applications of biocatalysis in developing green chemistry for chemical synthesis at the industrial scale. *Green Chem.* **10,** 361 (2008).

81. Straathof, A. J., Panke, S. & Schmid, A. The production of fine chemicals by biotransformation. *Curr. Opin. Biotechnol.* **13,** 548–556 (2002).

82. Tao, J., Zhao, L. & Ran, N. Recent advances in developing chemoenzymatic processes for active pharmaceutical ingredients. *Org. Process Res. Dev.* **11,** 259–267 (2007).

83. Wohlgemuth, R. Asymmetric biocatalysis with microbial enzymes and cells. *Curr. Opin. Microbiol.* **13,** 283–292 (2010).

84. Carvalho, P. de O., Contesini, F. J., Bizaco, R., Calafatti, S. A. & Macedo, G. A. Optimization of enantioselective resolution of racemic ibuprofen by native lipase from Aspergillus niger. *J. Ind. Microbiol. Biotechnol.* **33,** 713–718 (2006).

85. Köhler, V., Bailey K., Znabet, A., Raftery, J., Helliwell, M. and Turner, N. Enantioselective biocatalytic oxidative desymmetrization of substituted pyrrolidines. *Angew. Chem. Int. Ed. Engl.* **49,** 2182–2184 (2010).

86. Koszelewski, D., Pressnitz, D., Clay, D. & Kroutil, W. Deracemization of Mexiletine Biocatalyzed by ω-Transaminases. *Org. Lett.* **11,** 4810–4812 (2009).

87. Kisailus, D., Choi, J. H., Weaver, J. C., Yang, W. & Morse, D. E. Enzymatic Synthesis and Nanostructural Control of Gallium Oxide at Low Temperature. *Adv. Mater.* **17,** 314–318 (2005).

88. Kisailus, D., Truong, Q., Amemiya, Y., Weaver, J. C. & Morse, D. E. Self-assembled bifunctional surface mimics an enzymatic and templating protein for the synthesis of a metal oxide semiconductor. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 5652–7 (2006).

89. Sumerel, J. L. *et al.* Biocatalytically Templated Synthesis of Titanium Dioxide. *Chem. Mater.* **15,** 4804–4809 (2003).

90. Brutchey, R. L., Yoo, E. S. & Morse, D. E. Biocatalytic Synthesis of a Nanostructured and Crystalline Bimetallic Perovskite-like Barium Oxofluorotitanate at Low Temperature. *J. Am. Chem. Soc.* **128,** 10288–10294 (2006).

91. Natalio, F., Mugnaioli, E., Wiens, M., Wang X., Schröder H., Tahir, M., Tremel W., Kolb U. and Müller, W. Silicatein-mediated incorporation of titanium into spicules from the demosponge Suberites domuncula. *Cell Tissue Res.* **339,** 429–436 (2010).

92. Tahir, M. N., Eberhardt, M., Therese, H., Kolb, U., Theato, P., Müller, W., Schröder, H. and Tremel W. From Single Molecules to Nanoscopically Structured Functional Materials: Au Nanocrystal Growth on TiO2 Nanowires Controlled by Surface-Bound Silicatein. *Angew. Chemie Int. Ed.* **45,** 4803–4809 (2006).

93. Tahir, M. N., Theato, P., Müller, W., Schröder, H., Borejko, A., Fai, S., Janshoff, A., Huth, J. and Tremel W. Formation of layered titania and zirconia catalysed by surface-bound silicatein. *Chem. Commun.* 5533–5535 (2005).

94. Tahir, M. N., Theato, P., Müller, W., Schröder, H., Janshoff, A., Zhang, J., Huth, J. and Tremel, *W.* Monitoring the formation of biosilica catalysed by histidine-tagged silicatein. *Chem. Commun.* 2848–2849 (2004).

95. Rai, A. & Perry, C. C. Facile Fabrication of Uniform Silica Films with Tunable Physical Properties Using Silicatein Protein from Sponges. *Langmuir* **26,** 4152–4159 (2010).

96. Polini, A., Pagliara, S., Camposeo A., Cingolani R., Wang X., Schröder H., Müller W. and Pisignano D. Optical properties of in-vitro biomineralised silica. *Sci. Rep.* **2,** 607 (2012).

97. Vancheeswaran, S., Halden, R. U., Williamson, K. J., Ingle James D. & Semprini, L. Abiotic and Biological Transformation of Tetraalkoxysilanes and Trichloroethene/*cis*-1,2-Dichloroethene Cometabolism Driven by Tetrabutoxysilane-Degrading Microorganisms. *Environ. Sci. Technol.* **33,** 1077–1085 (1999).

98. Maraite, A., Ansorge-Schumacher, M. B., Ganchegui, B., Leitner, W. & Grogan, G. On the biocatalytic cleavage of silicon–oxygen bonds: A substrate structural approach to investigating the cleavage of protecting group silyl ethers by serine-triad hydrolases. *J. Mol. Catal. B Enzym.* **56,** 24–28 (2009).

99. Fattakhova, A. N., Ofitserov, E. N., Diyakov, V. M. & Naumova, R. P. Utilization of 1-chloromethylsilatrane by *Rhodotorula mucilaginosa*. *FEMS Microbiol. Lett.* **48,** 317–319 (1987).

100. Bassindale, A. R., Brandstadt, K. F., Lane, T. H. & Taylor, P. G. Enzyme-catalysed siloxane bond formation. *J. Inorg. Biochem.* **96,** 401–406 (2003).

101. Bisswanger, H. Enzyme assays. *Perspect. Sci.* **1,** 41–55 (2014).

102. Cantone, S., Hanefeld, U. & Basso, A. Biocatalysis in non-conventional media-ionic liquids, supercritical fluids and the gas phase. *Green Chem.* **9,** 954–971 (2007).

103. Deetz, J. S. & Rozzell, J. D. Enzyme-catalysed reactions in non-aqueous media. *Trends Biotechnol.* **6,** 15–19 (2017).

104. Wang, S. *et al.* Enzyme Stability and Activity in Non-Aqueous Reaction Systems: A Mini Review. *Catalysts* **6,** (2016).

105. Klibanov, A. M. Improving enzymes by using them in organic solvents. *Nature* **409,** 241–246 (2001).

106. Vinogradov, A. A., Kudryashova, E., Grinberg, V., Grinberg, N., Burova, T. and Levashov. The chemical modification of α-chymotrypsin with both hydrophobic and hydrophilic compounds stabilizes the enzyme against denaturation in water–organic media.

Protein Eng. Des. Sel. **14,** 683–689 (2001).

107. Klibanov, A. M. What is remembered and why? *Nature* **374,** 596 (1995).

108. Zaks, A. & Klibanov, A. M. The effect of water on enzyme action in organic media. *J. Biol. Chem.* **263,** 8017–8021 (1988).

109. Zaks, A. & Klibanov, A. M. Enzyme-catalyzed processes in organic solvents. *Proc. Natl. Acad. Sci.* **82,** 3192–3196 (1985).

110. Sharma, S. & Kanwar, S. S. Organic solvent tolerant lipases and applications. *Sci World J* **2014,** (2014).

111. Grunwald, J., Wirz, B., P. Scollar, M. & Klibanov, A. Asymmetric oxidoreductions catalyzed by alcohol dehydrogenase in organic solvents. *Journal of the American Chemical Society* **108,** (1986).

112. Gupta, M. N., Mukherjee, J. & Malhotra, D. Use of high activity enzyme preparations in neat organic solvents for organic synthesis. *Univers. Org. Chem.* **1,** (2013).

113. Gupta, M. N. Enzyme function in organic solvents. *Eur. J. Biochem.* **203,** 25–32 (1992).

114. Yang, Z. & A. Robb, D. Partition coefficients of substrate and products and solvent selection for biocatalysis under nearly anhydrous conditions. *Biotechnology and bioengineering* **43,** (1994).

115. Ryu, K. & Dordick, J. S. How do organic solvents affect peroxidase structure and function? *Biochemistry* **31,** 2588–2598 (1992).

116. Dordick, J. S. Enzymatic catalysis in monophasic organic solvents. *Enzyme Microb. Technol.* **11,** 194–211 (1989).

117. Johnson, K. A. & Goody, R. S. The Original Michaelis Constant: Translation of the 1913 Michaelis Menten Paper. *Biochemistry* **50,** 8264–8269 (2011).

118. Johnson, K. A. A century of enzyme kinetic analysis, 1913 to 2013. *FEBS Lett.* **587,** 2753–2766 (2013).

119. Briggs, G. E. & Haldane, J. B. S. A Note on the Kinetics of Enzyme Action. *Biochem. J.* **19,** 338–339 (1925).

120. Craik, C. S., Roczniak, S., Largman, C. & Rutter, W. J. The catalytic role of the active site aspartic acid in serine proteases. *Science (80-.).* **237,** 909 LP-913 (1987).

121. Sprang, S., Standing, T., Fletterick, R., Stroud, R., Finer-Moore, J., Xuong, N., Hamlin, R., Rutter, W. and Craik, C. The three-dimensional structure of Asn102 mutant of trypsin: role of Asp102 in serine protease catalysis. *Science (80-.).* **237,** 905 LP-909 (1987).

122. Carter, P. & Wells, J. A. Dissecting the catalytic triad of a serine protease. *Nature* **332,** 564–568 (1988).

123. Carter, P. & Wells, J. A. Engineering enzyme specificity by "substrate-assisted catalysis" *Science (80-.).* **237,** 394 LP-399 (1987).

124. Carter, P., Abrahmsen, L. & Wells, J. A. Probing the mechanism and improving the rate

of substrate-assisted catalysis in subtilisin BPN'. *Biochemistry* **30,** 6142–6148 (1991).

125. Corey, D. R. & Craik, C. S. An investigation into the minimum requirements for peptide hydrolysis by mutation of the catalytic triad of trypsin. *J. Am. Chem. Soc.* **114,** 1784–1790 (1992).

126. Goulding, C. W. & Jeanne Perry, L. Protein production in Escherichia coli for structural studies by X-ray crystallography. *J. Struct. Biol.* **142,** 133–143 (2003).

127. He, Y., Wang, K. & Yan, N. The recombinant expression systems for structure determination of eukaryotic membrane proteins. *Protein Cell* **5,** 658–672 (2014).

128. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in Microbiology* **5,** 172 (2014).

129. Sørensen, H. P. & Mortensen, K. K. Advanced genetic strategies for recombinant protein expression in Escherichia coli. *J. Biotechnol.* **115,** 113–128 (2005).

130. Williams, D. C., Van Frank, R. M., Muth, W. L. & Burnett, J. P. Cytoplasmic inclusion bodies in Escherichia coli producing biosynthetic human insulin proteins. *Science (80).* **215,** 687 LP-689 (1982).

131. Baneyx, F. Recombinant protein expression in Escherichia coli. *Curr. Opin. Biotechnol.* **10,** 411–421 (1999).

132. Makino, T., Skretas, G. & Georgiou, G. Strain engineering for improved expression of recombinant proteins in bacteria. *Microbial Cell Factories* **10,** 32 (2011).

133. Murr, M. M. & Morse, D. E. Fractal intermediates in the self-assembly of silicatein filaments. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 11657–62 (2005).

134. Brutchey, R. L. & Morse, D. E. Silicatein and the Translation of its Molecular Mechanism of Biosilicification into Low Temperature Nanomaterial Synthesis. *Chem. Rev.* **108,** 4915–4934 (2008).

135. Kowatz, T. Mechanisms of silicate polymerisation, carbohydrate epimerisation and metalloprotease inhibition. (University of St Andrews, Scotland, 2009).

136. Manfrin, A. Preparation and functional characterisation of recombinant silicatein-alpha from sponge Suberites domuncula. (Johannes Gutenberg of Mainz, 2014).

137. van der Walle, C. F. Towards a Bottom-up Approach for Mimicking Marine Sponge Spicules. *Silicon* **4,** 23–31 (2012).

138. Müller, W. E. G., Schröder, H., Muth, S., Gietzen, S., Korzhev, M., Grenbenjuk, V., Wiens, M., Schloßmacher, U. and Wang, X. The silicatein propeptide acts as inhibitor/modulator of self-organization during spicule axial filament formation. *FEBS J.* **280,** 1693–1708 (2013).

139. de Marco, A. & De Marco, V. Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expression tuning. *J. Biotechnol.* **109,** 45–52 (2004).

140. Waugh, D. S. Making the most of affinity tags. *Trends Biotechnol.* **23,** 316–320 (2005).

141. Zhao, Xinyu, Li, Guoshun, Liang, S. Several Affinity Tags Commonly Used in Chromatographic Purification. *J. Anal. Methods Chem.* **2013,** 1–8 (2013).

142. Costa, S., Almeida, A., Castro, A. & Domingues, L. Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system. *Frontiers in Microbiology* **5,** (2014).

143. Qing, G. *et al.* Cold-shock induced high-yield protein production in Escherichia coli. *Nat Biotech* **22,** 877–882 (2004).

144. Butt, T. R., Edavettal, S. C., Hall, J. P. & Mattern, M. R. SUMO fusion technology for difficult-to-express proteins. *Protein Expr. Purif.* **43,** 1–9 (2005).

145. Maier, T., Ferbitz, L., Deuerling, E. & Ban, N. A cradle for new proteins: trigger factor at the ribosome. *Curr. Opin. Struct. Biol.* **15,** 204–212 (2005).

146. Smith, D. B. & Johnson, K. S. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* **67,** 31–40 (1988).

147. Lebendiker, M. & Danieli, T. in *Protein Chromatography: Methods and Protocols* (eds. Walls, D. & Loughran, S. T.) 281–293 (Humana Press, 2011).

148. Kapust, R. B. & Waugh, D. S. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **8,** 1668–1674 (1999).

149. Ramos, R., Moreira, S., Rodrigues, A., Gama, M. & Domingues, L. Recombinant expression and purification of the antimicrobial peptide magainin-2. *Biotechnol. Prog.* **29,** 17–22 (2013).

150. Ramos, R., Domingues, L. & Gama, M. Escherichia coli expression and purification of LL37 fused to a family III carbohydrate-binding module from Clostridium thermocellum. *Protein Expr. Purif.* **71,** 1–7 (2010).

151. Curnow, P., Bessette, P., Kisailus, D., Murr, M., Daugherty, P. and Morse, D. Enzymatic Synthesis of Layered Titanium Phosphates at Low Temperature and Neutral pH by Cell-Surface Display of Silicatein-α. *J. Am. Chem. Soc.* **127,** 15749–15755 (2005).

152. Ki, M. R., Yeo, K. B. & Pack, S. P. Surface immobilization of protein via biosilification catalyzed by silicatein fused to glutathione S-transferase (GST). *Bioprocess Biosyst. Eng.* **36,** 643–648 (2013).

153. Hartl, F. U. & Hayer-Hartl, M. Molecular Chaperones in the Cytosol: from Nascent Chain to Folded Protein. *Science (80-.).* **295,** 1852 LP-1858 (2002).

154. Hartl, F. U. & Hartl, F. U. Molecular chaperones in protein folding and proteostasis. *Nature* **475,** 324–332 (2011).

155. Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A. and Hartl, F. Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. *Nature* **352,** 36–42 (1991).

156. Sharma, S. K., De Los Rios, P., Christen, P., Lustig, A. & Goloubinoff, P. The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. *Nat*

Chem Biol **6,** 914–920 (2010).

157. Diamant, S., Ben-Zvi, A. P., Bukau, B. & Goloubinoff, P. Size-dependent Disaggregation of Stable Protein Aggregates by the DnaK Chaperone Machinery. *J. Biol. Chem.* **275,** 21107–21113 (2000).

158. Liu, C. P., Zhou, Q. M., Fan, D. J. & Zhou, J. M. PPIase domain of trigger factor acts as auxiliary chaperone site to assist the folding of protein substrates bound to the crevice of trigger factor. *Int. J. Biochem. Cell Biol.* **42,** 890–901 (2010).

159. Ferbitz, L., Maier, T., Patzelt, H., Bukav, B., Deuerling, E. and Ban, N. Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* **431,** 590–596 (2004).

160. Nishihara, K., Kanemori, M., Yanagi, H. & Yura, T. Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in Escherichia coli. *Applied and Environmental Microbiology* **66,** 884–889 (2000).

161. Kramer, G., Patzelt, H., Rauch, T., Kurz, T., Vorderwülbecke, S., Bukau, B. and Deuerling, E. Trigger Factor Peptidyl-prolyl cis/trans Isomerase Activity Is Not Essential for the Folding of Cytosolic Proteins in Escherichia coli. *J. Biol. Chem.* **279,** 14165–14170 (2004).

162. Gupta, R., Lakshmipathy, S. K., Chang, H.-C., Etchells, S. A. & Hartl, F. U. Trigger factor lacking the PPIase domain can enhance the folding of eukaryotic multi-domain proteins in Escherichia coli. *FEBS Letters* **584,** (2010).

163. Haacke, A., Fendrich, G., Ramage, P. & Geiser, M. Chaperone over-expression in Escherichia coli: Apparent increased yields of soluble recombinant protein kinases are due mainly to soluble aggregates. *Protein Expr. Purif.* **64,** 185–193 (2009).

164. Bergès, H., Joseph-Liauzun, E. & Fayet, O. Combined effects of the signal sequence and the major chaperone proteins on the export of human cytokines in Escherichia coli. *Applied and Environmental Microbiology* **62,** 55–60 (1996).

165. Martínez-Alonso, M., García-Fruitós, E., Ferrer-Miralles, N., Rinas, U. & Villaverde, A. Side effects of chaperone gene co-expression in recombinant protein production. *Microbial Cell Factories* **9,** 64 (2010).

166. Leibly, D. J., Nguyen, T., Kao, L., Hewitt, S., Barrett, L. and Van Voorhis, W. Stabilizing Additives Added during Cell Lysis Aid in the Solubilization of Recombinant Proteins. *PLoS One* **7,** e52482 (2012).

167. Bondos, S. E. & Bicknell, A. Detection and prevention of protein aggregation before, during, and after purification. *Anal. Biochem.* **316,** 223–231 (2003).

168. Golovanov, A. P., Hautbergue, G. M., Wilson, S. A. & Lian, L.-Y. A Simple Method for Improving Protein Solubility and Long-Term Stability. *J. Am. Chem. Soc.* **126,** 8933–8939 (2004).

169. Arakawa, T., Ejima, D., Tsumoto, K., Obeyama, N., Tanaka, Y., Kita, Y. and Timasheff, S. Suppression of protein interactions by arginine: A proposed mechanism of the arginine effects. *Biophys. Chem.* **127,** (2007).

170. Ishibashi, M., Tsumoto, K., Tokunaga, M., Ejima, D., Kita, Y. and Arakawa, T. Is arginine a protein-denaturant? *Protein Expr. Purif.* **42,** (2005).

171. Arakawa, T., Tsumoto, K., Kita, Y., Chang, B. & Ejima, D. Biotechnology applications of amino acids in protein purification and formulations. *Amino Acids* **33,** 587–605 (2007).

172. Tao, H., Liu, W., Simmons, B., Harris, H., Cox, T. and Massiah, M. Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100 and CHAPS. *Biotechniques* **48,** 61–64 (2010).

173. Rodi, P. M., Bocco Gianello, M. D., Corregido, M. C. & Gennaro, A. M. Comparative study of the interaction of CHAPS and Triton X-100 with the erythrocyte membrane. *Biochim. Biophys. Acta - Biomembr.* **1838,** 859–866 (2014).

174. le Maire, M., Champeil, P. & Møller, J. V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta - Biomembr.* **1508,** 86–111 (2000).

175. Helenius, A. & Simons, K. Solubilization of membranes by detergents. *Biochim. Biophys. Acta - Rev. Biomembr.* **415,** 29–79 (1975).

176. Natalio, F., Link, T., Müller, W., Schröder, H., Cui, F., Wang, X. and Wiens, M. Bioengineering of the silica-polymerizing enzyme silicatein-α for a targeted application to hydroxyapatite. *Acta Biomater.* **6,** 3720–3728 (2010).

177. Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., Bansal, P., Bridge, A. J.. Poux, S., Bougueleret, L. and Xenarios, I. in *Plant Bioinformatics: Methods and Protocols* (ed. Edwards, D.) 23–54 (Springer New York, 2016).

178. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7,** (2011).

179. Goujon, M. *et al.* A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic Acids Res.* **38,** 695–699 (2010).

180. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis* **18,** 2714–2723 (1997).

181. Kandror, O. & Goldberg, A. L. Trigger factor is induced upon cold shock and enhances viability of Escherichia coli at low temperatures. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 4978–4981 (1997).

182. Goldstein, J., Pollitt, N. S. & Inouye, M. Major cold shock protein of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **87,** 283–287 (1990).

183. de Marco, A. Two-step metal affinity purification of double-tagged (NusA-His6) fusion proteins. *Nat. Protoc.* **1,** 1538–1543 (2006).

184. Harper, S. & Speicher, D. W. Purification of proteins fused to glutathione S-tranferase. *Methods in molecular biology (Clifton, N.J.)* **681,** 259–280 (2011).

185. Kaplan, W., Erhardt, J., Sluis-Cremer, N., Dirr, H., Hüsler, P. and Klump, H Conformational stability of pGEX-expressed Schistosoma japonicum glutathione S-transferase: A detoxification enzyme and fusion-protein affinity tag. *Protein Sci.* **6,** 399–406 (2008).

186. Green, A. A. & Hughes, W. L. in 67–90 (1955). 187. Wingfield, P. T. Protein Precipitation Using Ammonium Sulfate. *Curr. Protoc. Protein Sci.,* Appendix-3F (2001).

188. Vallejo, L. F. & Rinas, U. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb. Cell Fact.* **3,** 11 (2004).

189. Yang, Z., Zhang, L., Zhang, Y., Zhang, T., Feng, Y., Lu, X., Lan, W., Wang, J., Wu, H., Cao, C. and Wang, X. Highly Efficient Production of Soluble Proteins from Insoluble Inclusion Bodies by a Two-Step-Denaturing and Refolding Method. *PLoS One* **6,** 22981 (2011).

190. Shukla, D. & Trout, B. L. Understanding the Synergistic Effect of Arginine and Glutamic Acid Mixtures on Protein Solubility. *J. Phys. Chem. B* **115,** 11831–11839 (2011).

191. Schmidt, T. G. M. & Skerra, A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protoc.* **2,** 1528–1535 (2007).

192. Coradin, Thibaud, Eglin, David and Livage, J. The silicomolybdic acid spectrophotometric method and its pllication to silicate/biopolymer interaction studies. *Spectroscopy* **18,** 567–576 (2004).

193. Morse, D. E. Silicon biotechnology: harnessing biological silica production to construct new materials. *Trends Biotechnol.* **17,** 230–232 (1999).

194. Berg, J. M., Tymoczko, J. L. & Stryer, L. in *Biochemistry textbook* 1026 (2010).

195. Stein, M., Gabdoulline, R. R. & Wade, R. C. Calculating enzyme kinetic parameters from protein structures. *Biochem. Soc. Trans.* **36,** 51–54 (2008).

196. Eisenthal, R., Danson, M. J. & Hough, D. W. Catalytic efficiency and k_{cat}/K_M: a useful comparator? *Trends Biotechnol.* **25,** 247–249 (2007).

197. Koshland, D. E. The Application and Usefulness of the Ratio k_{cat}/K_M. *Bioorg. Chem.* **30,** 211–213 (2002).

198. Park, J. H., Meriwether, B. P., Clodfelder, P. & Cunningham, L. W. The hydrolysis of p-nitrophenyl acetate catalyzed by 3-phosphoglyceraldehyde dehydrogenase. *J. Biol. Chem.* **236,** 136–141 (1961).

199. Buß, O., Jager, S., Dold, S., Zimmermann, S., Hamacher, K, Schmitz, K., and Rudat, J. Statistical Evaluation of HTS Assays for Enzymatic Hydrolysis of β-Keto Esters. *PLoS One* **11,** 1–19 (2016).

200. Bornscheuer, U. T. High-Throughput-Screening Systems for Hydrolases. *Eng. Life Sci.* **4,** 539–542 (2004).

201. Schmidt, M. & Bornscheuer, U. T. High-throughput assays for lipases and esterases. *Biomol. Eng.* **22,** 51–56 (2005).

202. Adams-Cioaba, M. A., Krupa, J. C., Xu, C., Mort, J. S. & Min, J. Structural basis for the recognition and cleavage of histone H3 by cathepsin L. *Nat. Commun.* **2,** 197 (2011).

203. Polgár, L. The catalytic triad of serine peptidases. *Cell. Mol. Life Sci. C.* **62,** 2161–2172 (2005).

204. Belton, D. J., Deschaume, O. & Perry, C. C. An overview of the fundamentals of the

chemistry of silica with relevance to biosilicification and technological advances. *Febs J.* **279,** 1710–1720 (2012).

205. Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. and Milo, R. The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. *Biochemistry* **50,** 4402–4410 (2011).

206. Vernet, T., Tessier, D., Chatellier, J., Plouffe, C., Lee, T., Thomas, D., Storer, A. and Mėnard, R. Structural and Functional Roles of Asparagine 175 in the Cysteine Protease Papain. *J. Biol. Chem.* **270,** 16645–16652 (1995).

207. Rawlings, N. D. & Barrett, A. J. *Proteolytic Enzymes: Serine and Cysteine Peptidases*. *Methods in Enzymology* **244,** (1994).

208. Hedstrom, L. Serine Protease Mechanism and Specificity. *Chem. Rev.* **102,** 4501–4524 (2002).

209. Koshland Jr., D. E. The application and usefulness of the ratio k_{cat}/K_M. *Bioorg. Chem.* **30,** 211–213 (2002).

210. Pedelacq, J.D., Plitch, E., Liong, E., Berendzen, J., Kim, C., Rho, B., Park, M., Terwilliger, T., Waldo, G. Engineering soluble proteins for structural genomics. *Nat Biotech* **20,** 927–932 (2002).

211. Yang, H., Liu, L. & Xu, F. The promises and challenges of fusion constructs in protein biochemistry and enzymology. *Appl. Microbiol. Biotechnol.* **100,** 8273–8281 (2016).

212. Hauer, B. & Roberts, S. M. Biocatalysis and biotransformation. *Curr. Opin. Chem. Biol.* **8,** 103–105 (2004).

213. Trincone, A. Potential biocatalysts originating from sea environments. *J. Mol. Catal. B Enzym.* **66,** 241–256 (2010).

214. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1,** 2876–2890 (2006).

215. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochim. Biophys. Acta - Proteins Proteomics* **1751,** 119–139 (2005).

216. Micsonai, A., Wien, F., Kernya, L., Lee, Y., Goto, Y., Rėfrėgiers, M. and Kardos, J. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc. Natl. Acad. Sci.* **112,** E3095–E3103 (2015).

217. Poschner, B. C., Reed, J., Langosch, D. & Hofmann, M. W. An automated application for deconvolution of circular dichroism spectra of small peptides. *Anal. Biochem.* **363,** 306–308 (2007).

218. Some, D. & Kenrick, S. Characterization of protein-protein interactions via stativ and dynamic light scattering. (INTECH Open Access Publisher, 2012).

219. Some, D. Light-scattering-based analysis of biomolecular interactions. *Biophys. Rev.* **5,** 147–158 (2013).

220. Wyatt, P. J. Light scattering and the absolute characterization of macromolecules. *Anal.*

Chim. Acta **272,** 1–40 (1993).

221. Lorber, B., Fischer, F., Bailly, M., Roy, H. & Kern, D. Protein analysis by dynamic light scattering: Methods and techniques for students. *Biochem. Mol. Biol. Educ.* **40,** 372–382 (2012).

222. Folta-Stogniew, E. in *New and Emerging Proteomic Techniques* 97–112 (Humana Press, 2006).

223. Blaha, G., Wilson, D., Stoller, G., Fischer, G, Willumeit, R. and Nierhaus, K. Localization of the Trigger Factor Binding Site on the Ribosomal 50S Subunit. *J. Mol. Biol.* **326,** 887–897 (2003).

224. Hoffmann, A., Bukau, B. & Kramer, G. Structure and function of the molecular chaperone Trigger Factor. *Biochim. Biophys. Acta - Mol. Cell Res.* **1803,** 650–661 (2010).

225. Stetefeld, J., McKenna, S. A. & Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys. Rev.* **8,** 409–427 (2016).

226. Baldwin, R. L. How Hofmeister ion interactions affect protein stability. *Biophys. J.* **71,** 2056–2063 (1996).

227. Good, N. E., Winget, G., Wineter, W., Connolly, T., Izawa, S. and Singh, R. Hydrogen Ion Buffers for Biological Research*. *Biochemistry* **5,** 467–477 (1966).

228. David Crouch, R. Selective monodeprotection of bis-silyl ethers. *Tetrahedron* **60,** 5833–5871 (2004).

229. Kumar, A. & Venkatesu, P. Overview of the Stability of α-Chymotrypsin in Different Solvent Media. *Chem. Rev.* **112,** 4283–4307 (2012).

230. Kumar, A., Dhar, K., Kanwar, S. S. & Arora, P. K. Lipase catalysis in organic solvents: advantages and applications. *Biol. Proced. Online* **18,** 2 (2016).

231. Stepankova, V., Bidmanova, S., Koudelskova, T., Prokop, Z., Chaloupkova, R. and Damborsky, J. Strategies for Stabilization of Enzymes in Organic Solvents. *ACS Catal.* **3,** 2823–2836 (2013).

232. Raj, S. B., Ramaswamy, S. & Plapp, B. V. Yeast Alcohol Dehydrogenase Structure and Catalysis. *Biochemistry* **53,** 5791–5803 (2014).

233. Racker, E. Crystalline alcohol dehydrogenase from bakers' yeast. *J. Biol. Chem.* **184,** 313–320 (1950).

234. Dickinson, F. M. & Monger, G. P. A study of the kinetics and mechanism of yeast alcohol dehydrogenase with a variety of substrates. *Biochem. J.* **131,** 261–270 (1973).

235. Guinn, R. M., Skerker, P. S., Kavanaugh, P. & Clark, D. S. Activity and flexibility of alcohol dehydrogenase in organic solvents. *Biotechnol. Bioeng.* **37,** 303–308 (1991).

236. Deetz, J. S. & Rozzell, J. D. in *Biocatalysts for Industry* (ed. Dordick, J. S.) 181–191 (Springer US, 1991). doi:10.1007/978-1-4757-4597-9_9

237. Bommarius, A. S., Schwarm, M. & Drauz, K. Biocatalysis to amino acid-based chiral pharmaceuticals examples and perspectives. *J. Mol. Catal. B Enzym.* **5,** 1–11 (1998).

238. Breuer, M., Ditrich, K., Haibicher, T., Hauer, B., Kesseler, M., Stürmer, R. and Zelinski, T. Industrial Methods for the Production of Optically Active Intermediates. *Angew. Chemie Int. Ed.* **43,** 788–824 (2004).

239. Goldberg, K., Schroer, K., Lütz, S. & Liese, A. Biocatalytic ketone reduction---a powerful tool for the production of chiral alcohols---part I: processes with isolated enzymes. *Appl. Microbiol. Biotechnol.* **76,** 237 (2007).

240. Liese, A., Seelbach, K. & Wandrey, C. *Processes in Industrial Biotransformations*. (Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2000).

241. Hummel, W. & Riebel, B. Chiral Alcohols by Enantioselective Enzymatic Oxidation. *Ann. N. Y. Acad. Sci.* **799,** 713–716 (1996).

242. Weckbecker, A. & Hummel, W. Improved synthesis of chiral alcohols with Escherichia coli cells co-expressing pyridine nucleotide transhydrogenase, NADP+-dependent alcohol dehydrogenase and NAD+-dependent formate dehydrogenase. *Biotechnol. Lett.* **26,** 1739–1744 (2004).

243. Kizaki, N., Yasohara, Y., Hasegawa, J., Wada, M., Kataoka, M. and Shimizu, S. Synthesis of optically pure ethyl (S)-4-chloro-3-hydroxybutanoate by Escherichia coli transformant cells coexpressing the carbonyl reductase and glucose dehydrogenase genes. *Appl. Microbiol. Biotechnol.* **55,** 590–595 (2001).

244. Gröger, H., Hummel, W., Rollmann, C., Chamouleau, F, Hüsken, H., Werner, H., Wunderlich, C., Abokitse, K., Drauz, K. and Buchholz, S. Preparative asymmetric reduction of ketones in a biphasic medium with an (S)-alcohol dehydrogenase under in situ-cofactor-recycling with a formate dehydrogenase. *Tetrahedron* **60,** 633–640 (2004).

245. Ernst, M., Kaup, B., Müller, M., Bringer-Meyer, S. & Sahm, H. Enantioselective reduction of carbonyl compounds by whole-cell biotransformation, combining a formate dehydrogenase and a (R)-specific alcohol dehydrogenase. *Appl. Microbiol. Biotechnol.* **66,** 629–634 (2005).

246. Tao, J. & McGee, K. Development of a Continuous Enzymatic Process for the Preparation of (R)-3-(4-Fluorophenyl)-2-hydroxy Propionic Acid. *Org. Process Res. Dev.* **6,** 520–524 (2002).

247. Schmidt, E., Ghisalba, O., Gygax, D. & Sedelmeier, G. Optimization of a process for the production of (R)-2-hydroxy-4-phenylbutyric acid — an intermediate for inhibitors of angiotensin converting enzyme. *J. Biotechnol.* **24,** 315–327 (1992).

248. Pollard, D., Truppo, M., Pollard, J., Chen, C. & Moore, J. Effective synthesis of (S)-3,5-bistrifluoromethylphenyl ethanol by asymmetric enzymatic reduction. *Tetrahedron: Asymmetry* **17,** 554–559 (2006).

249. Haberland, J., Hummel, W., Daussmann, T. & Liese, A. New Continuous Production Process for Enantiopure (2R,5R)-Hexanediol. *Org. Process Res. Dev.* **6,** 458–462 (2002).

250. Sheldon, R. A. *Multi-Step Enzyme Catalysis: Biotransforamtions and Chemoenzymatic Synthesis*. (Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2008).

251. Parsons, P. J., Penkett, C. S. & Shell, A. J. Tandem Reactions in Organic Synthesis: Novel Strategies for Natural Product Elaboration and the Development of New Synthetic Methodology. *Chem. Rev.* **96,** 195–206 (1996).

252. Li, C. J. & Trost, B. M. Green chemistry for chemical synthesis. *Proc. Natl. Acad. Sci.* **105,** 13197–13202 (2008).

253. Reetz, M. T. & Carballeira, J. D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* **2,** 891–903 (2007).

254. Grunwald, P. *Industrial Biocatalysis*. (Taylor & Francis, 2015).

255. Snijder-Lambers, A. M., Vulfson, E. N. & Doddema, H. J. Optimization of alcohol dehydrogenase activity and NAD(H) regeneration in organic solvents. *Recl. des Trav. Chim. des Pays-Bas* **110,** 226–230 (2010).

256. Voss, C. V., Gruber, C., Faber, K., Knaus, T., Macheroux, P. and Kroutil, W. Orchestration of Concurrent Oxidation and Reduction Cycles for Stereoinversion and Deracemisation of sec-Alcohols. *J. Am. Chem. Soc.* **130,** 13969–13972 (2008).

257. Guinn, R. M., Skerker, P. S., Kavanaugh, P. & Clark, D. S. Activity and flexibility of alcohol dehydrogenase in organic solvents. *Biotechnol. Bioeng.* **37,** 303–308 (1991).

258. Iyer, P. V. & Ananthanarayan, L. Enzyme stability and stabilization—Aqueous and non-aqueous environment. *Process Biochem.* **43,** 1019–1032 (2008).

259. Simon, L. M., László, K., Vértesi, A., Bagi, K. & Szajáni, B. Stability of hydrolytic enzymes in water-organic solvent systems. *J. Mol. Catal. B Enzym.* **4,** 41–45 (1998).

260. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680–685 (1970).

261. Sanchis, J., Fernández, L., Carballeria, J., Drone, J., Gumulya, Y., Höbenreich, H., Kahakeaw, D., Kille, S., Lohmer, R., Petralans, J., Podtetenieff, J., Prasad, S., Soni, P., Taglieber, A., Wu, S., Zilly, F. and Reetz, M. Improved PCR method for the creation of saturation mutagenesis libraries in directed evolution: application to difficult-to-amplify templates. *Appl. Microbiol. Biotechnol.* **81,** 387–397 (2008).

262. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77,** 51–59 (1989).

263. Adamson, D. H., Dabbs, D., Pacheco, C., Giotto, M., Morse, D. and Aksay, I. Non-Peptide Polymeric Silicatein α Mimic for Neutral pH Catalysis in the Formation of Silica. *Macromolecules* **40,** 5710–5717 (2007).

264. Poschner, B. C., Reed, J., Langosch, D. & Hofmann, M. W. *An automated application for deconvolution of circular dichroism spectra of small peptides*. *Analytical Biochemistry* **363,** (2007).

265. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis* **18,** 2714–2723 (1997).

266. DYNAMICS, version 6.10.0.10; Wyatt technology Corp.: Santa Barbara, CA,. (2008).

267. Goldin, A. A. Dynals v1.0 white paper. (2002).

268. Greenfield, N. J. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* **1,** 2527–2535 (2006).