**Structural and Functional Studies of 3-Deoxy-
D-arabino-
Heptulosonate 7-Phosphate Synthase from
*Prevotella nigrescens***

BY

Yu Bai

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy in Chemistry

Victoria University of Wellington

2019

Abstract

Multifunctional enzymes, bearing two or more catalytic activities, provide exceptional contributions to the efficient and coherent function of metabolic pathways. Two main benefits of multifunctional enzymes have been clearly described. Firstly, linked catalytic modules can enhance the overall catalytic rate for consecutive reactions of a metabolic pathway due to substrate channelling. Secondly, the fusion of two protein domains can impart allosteric control, such that the catalytic function of one of the protein domains is altered by a ligand binding to the second, covalently linked domain. This study examines a bifunctional enzyme comprising a 3-deoxy-D-*arabino* heptulosonate 7-phosphate synthase (DAH7PS) domain covalently fused to a C-terminal chorismate mutase (CM) domain from *Prevotella nigrescens* (*Pni*DAH7PS). DAH7PS catalyses the first reaction of the shikimate pathway leading to the biosynthesis of aromatic amino acids, whereas CM functions at a pathway branch point, leading to the biosynthesis of tyrosine and phenylalanine. Through the investigation of *Pni*DAH7PS, a special functional interdependence between the two non-consecutive catalytic functionalities and the derived allosteric regulation was unravelled.

Chapter 2 generally characterises the biochemical and structural features of *Pni*DAH7PS. The two catalytic activities exhibit substantial hetero-interdependency and the separation of the two distinct catalytic domains results in a dramatic loss of both the DAH7PS and CM enzymatic activities. The structural investigation into this protein revealed a unique dimeric assembly and implicates a hetero-interaction between the DAH7PS and CM domains, providing a structural basis for the functional interdependence. Moreover, allosteric inhibition of DAH7PS by prephenate, the product of the CM-catalysed reaction, was observed. This allostery is accompanied by

a striking conformational change, as observed by SAXS, implying that a manipulation of the hetero-domain interaction is the mechanism underpinning the allosteric inhibition.

Chapter 3 looks into the mechanism underpinning the DAH7PS and CM functional interdependence. Rearrangements of the conformation of *Pni*DAH7PS following the addition of substrate combinations were observed. This indicates that a dynamic interaction between the DAH7PS and CM domains is important for catalysis. Furthermore, perturbation of these conformational variations by either a truncation mutation in the CM domain or the presence of a high concentration of NaCl interrupted the both the DAH7PS and CM catalytic activities, implying that a dynamic hetero-domain interaction is essential for the delivering the normal DAH7PS and CM functions. This work also reveals a dual role for the DAH7PS domain, exerting catalysis and allosteric activation on the CM activity simultaneously.

Chapter 4 investigates the mechanism of the allosteric inhibition of *Pni*DAH7PS by prephenate. The structural effect of prephenate on *Pni*DAH7PS, with the addition of substrate combinations, was inspected, and the results unravelled the same conformation of *Pni*DAH7PS under different conditions, exhibiting high compactness and rigidity. This finding indicates that the probable inhibitory effect of prephenate on *Pni*DAH7PS is realised by freezing the enzyme's structure in order to deprive *Pni*DAH7PS of the dynamic-dependent catalytic activity.

Chapter 5 describes the development of a method for producing segmentally isotopically labelled *Pni*DAH7PS using Expressed Protein Ligation (EPL). This chapter also details attempts to couple this method with small angle neutron scattering (SANS) and nuclear magnetic resonance spectroscopy (NMR) to gain more structural information regarding the catalytic and allosteric properties of *Pni*DAH7PS.

Table of content

<i>Abstract</i>	I
<i>Table of content</i>	III
<i>List of Figures</i>	IX
<i>List of Tables</i>	XVII
<i>Abbreviations</i>	XXI
<i>Acknowledgement</i>	XXV
Chapter 1 <i>Introduction</i>	1
1.1 Shikimate pathway	3
1.2 DAH7PS	4
1.2.1 The DAH7PS family	6
1.2.2 The effect of quaternary structure on DAH7PS functions	11
1.3 Chorismate Mutase	13
1.4 The role of multifunctional enzymes and enzyme complexes in the shikimate pathway	16
1.4.1 Increased catalytic efficiency	16
1.4.2 Allosteric regulation	20
1.4.2.1 A general theory of allosteric regulation	20
1.4.2.2 Non-covalent enzyme complexes and bifunctional enzymes confer allostery on shikimate pathway	23
1.5 C-terminal CM linked DAH7PS	27
1.6 Objectives	28
Chapter 2 <i>Preliminary characterisation of PniDAH7PS</i>	31
2.1 Introduction	33
2.2 Sequence alignment and CLANS reveals the distinction of the C-terminal CM linked DAH7PS subgroup in type 1 DAH7PS	34
2.3 Cloning, expression and purification of <i>PniDAH7PS</i>	37
2.4 Basic characterisation of the functions and structure of <i>PniDAH7PS</i> ^{WT}	39

2.4.1 Metal ion activation	39
2.4.2 Kinetic characteristics	40
2.4.3 Feedback inhibition.....	42
2.4.4 Characterisation of the quaternary structure of <i>Pni</i> DAH7PS ^{WT}	43
2.5 Cloning, expression and purification of truncated variants.....	46
2.5.1 Selection of truncation site	46
2.5.2 Cloning, expression and purification of <i>Pni</i> DAH7PS ^D	47
2.5.3 Cloning, expression and purification of <i>Pni</i> DAH7PS ^{CM}	49
2.6 Basic characterisation of the function and structure of <i>Pni</i> DAH7PS ^D	51
2.6.1 Metal ion activation and kinetic parameters	51
2.6.2 Feedback inhibition.....	52
2.6.3 Characterisation of the quaternary structure of <i>Pni</i> DAH7PS ^D	53
2.7 Basic characterisation of the function and structure of <i>Pni</i> DAH7PS ^{CM}	55
2.7.1 Kinetic parameters	55
2.7.2 Quaternary structure	55
2.8 Small angle X-ray scattering (SAXS) investigations on the global shape of <i>Pni</i> DAH7PS ^{WT} and its truncated variants	56
2.8.1 Scattering profiles of truncated variants	57
2.8.2 Conformational change of <i>Pni</i> DAH7PS ^{WT} underlying allostery.....	59
2.8.3 <i>Ab initio</i> modelling for <i>Pni</i> DAH7PS ^{WT}	61
2.9 Crystallography.....	63
2.9.1 <i>Pni</i> DAH7PS ^{WT}	63
2.9.2 <i>Pni</i> DAH7PS ^D	63
2.9.3 <i>Pni</i> DAH7PS ^{CM}	66
2.10 Summary and discussion.....	68
Chapter 3 <i>DAH7PS</i> and <i>CM</i> functional interdependence in <i>Pni</i>DAH7PS	71
3.1 Introduction	73
3.2 Hetero-domain interaction	73
3.2.1 DAH7PS catalysed reaction promotes CM activity	74

3.2.2 Conformational changes of <i>Pni</i> DAH7PS ^{WT} with added PEP/Mn ²⁺ or PEP/Mn ²⁺ /E4P	78
3.2.3 Rigid body modelling for apo- <i>Pni</i> DAH7PS ^{WT} , PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT} and PEP/Mn ²⁺ /E4P- <i>Pni</i> DAH7PS ^{WT}	81
3.2.4 Isothermal titration calorimetry (ITC) to investigate the role of Mn ²⁺ in PEP binding	83
3.3 The essential role of hetero-domain interactions in delivering the functions of <i>Pni</i> DAH7PS ^{WT}	85
3.3.1 Kinetic characterisation and SAXS profiles of <i>Pni</i> DAH7PS ^{Δ5AA}	86
3.3.2 The effect of NaCl on the catalytic activities and structure of <i>Pni</i> DAH7PS ^{WT}	94
3.4 Summary and Discussion	101
Chapter 4 <i>Allosteric inhibition of DAH7PS by prephenate</i>	107
4.1 Introduction	109
4.2 The structural effect of prephenate on <i>Pni</i> DAH7PS ^{WT}	110
4.3 Rigid body models of prephenate-inhibited <i>Pni</i> DAH7PS ^{WT}	116
4.4 The effect of prephenate on <i>Pni</i> DAH7PS ^{Δ5AA} : Presenting more details for explaining the mechanism of prephenate mediated inhibition	117
4.5 NaCl	121
4.6 The binding of PEP/Mn ²⁺ to <i>Pni</i> DAH7PS ^{WT} in the presence of prephenate ..	123
4.7 The allostery of <i>Gsp</i> DAH7PS by prephenate	124
4.8 Summary and discussion	131
Chapter 5 <i>Expressed protein ligation coupled SANS and NMR assays for investigating the structure of PniDAH7PS</i>	133
5.1 Introduction	135
5.1.1 Small angle neutron scattering	135
5.1.2 Expressed protein ligation	136
5.1.3 Nuclear magnetic resonance (NMR)	137
5.2 Development of the method for the synthesis and purification of ligated	

<i>Pni</i> DAH7PS ^{D-CM}	138
5.3 Examination of the quaternary structure of <i>Pni</i> DAH7PS ^{D-CM}	145
5.4 Ligation between non-deuterated truncated DAH7PS and CM variants restores the catalytic activities and allosteric function of <i>Pni</i> DAH7PS	146
5.5 Segmental deuteration and SANS profile of <i>Pni</i> DAH7PS ^{D-CM}	150
5.6 Rigid body modelling of <i>Pni</i> DAH7PS ^{D-CM} against SANS and SAXS data	153
5.7 SAXS examination on the structure of <i>Pni</i> DAH7PS ^{D-CM} in the presence of either PEP/Mn ²⁺ or prephenate.....	155
5.8 Isotopic labelling of <i>Pni</i> DAH7PS and preliminary NMR tests.....	158
5.9 Summary and discussion.....	163
Chapter 6 <i>Summary and remarks</i>	167
6.1 The significance of the distinctive quaternary structure on the biochemical nature of <i>Pni</i> DAH7PS	169
6.2 The highly dynamic global structure of <i>Pni</i> DAH7PS	171
6.3 Dynamic hetero-domain interaction and DAH7PS domain mediated allosteric activation.....	172
6.4 Allosteric inhibition via alleviating dynamics	173
6.5 The loop function	174
6.6 Expressed protein ligation.....	175
6.7 Conclusion.....	177
Chapter 7 <i>Experimental procedures</i>	179
7.1 General experimental details.....	181
7.1.1 Multiple sequence alignments and CLuster ANalysis of Sequences...181	
7.1.2 Culture media.....	181
7.1.3 Gene cloning	182
7.1.3.1 Gene synthesis	182
7.1.3.2 Polymerase chain reaction.....	182
7.1.3.3 Agarose gel electrophoresis.....	184
7.1.3.4 Infusion	184

7.1.3.5 Restriction digestion and DNA ligation	185
7.1.3.6 Gateway	185
7.1.3.7 DNA transformation	187
7.1.3.8 Colony PRC	188
7.1.3.9 Plasmid extraction and purification	189
7.1.3.10 DNA sequencing.....	189
7.1.3.11 Site-directed mutagenesis	190
7.1.4 Isopropyl β -D-thiogalactopyranoside induced expression	190
7.1.5 Harvesting cells	191
7.1.6 Cell lysis	191
7.1.7 Chromatography	192
7.1.7.1 Immobilised metal affinity chromatography	192
7.1.7.2 Immobilised glutathione affinity chromatography	192
7.1.7.3 Size exclusion chromatography.....	193
7.1.7.4 Anion exchange chromatography	193
7.1.7.5 Hydrophobic interaction chromatography	194
7.1.8 TEV protease preparation and TEV cleavage	194
7.1.8.1 Preparation of TEV protease	194
7.1.8.2 TEV cleavage reaction	194
7.1.9 Concentration and storage of enzymes	195
7.1.10 Sodium dodecyl sulphate polyacrylamide gel electrophoresis	195
7.1.11 Determination of protein concentration	196
7.1.12 Mass spectrometry	197
7.1.13 kinetics	197
7.1.14 Feedback inhibition.....	198
7.1.15 Analytical ultracentrifugation	199
7.1.16 Analytical size exclusion chromatography	199
7.1.17 Crystallography	200
7.1.17.1 Screening and optimisation of crystallisation conditions	200

7.1.17.2 X-ray diffraction and data process	201
7.1.18 Small angle X-ray scattering	201
7.1.19 Modelling	204
7.1.20 Isothermal titration calorimetry	206
7.2 Experimental details for Chapter 5	208
7.2.1 Expressed protein ligation.....	208
7.2.2 Small angle neutron scattering	208
7.2.3 Isotopic labelling	209
7.2.4 Nuclear magnetic resonance spectroscopy	212
Appendix	215
Supporting information for Chapter 2	217
Supporting information for Chapter 3	221
Supporting information for Chapter 5	225
References	231

List of Figures

Figure 1. 1: The shikimate pathway leads to the biosynthesis of the aromatic amino acids. The dashed rectangle and circle highlight the steps of the synthesis of DAH7P and prephenate respectively.	4
Figure 1. 2: The schematic diagram shows the reaction mechanism for the formation of DAH7P.	5
Figure 1. 3: Classification of DAH7PS. DAH7PSs are divided into two main groups, type I and type II DAH7PS; type I is further classified into I α and I β (including KDO8PS) subgroups.	7
Figure 1. 4: The architecture of the <i>Pfu</i> DAH7PS active site (PDB code 1ZCO) (23)..	9
Figure 1. 5: Quaternary assemblies of <i>Pfu</i> DAH7PS (PDB code 1ZCO) (23), <i>Nme</i> DAH7PS (PDB code 5DCB) (22) and <i>Mtu</i> DAH7PS (PDB code 2B7O) (21) with AB interface residues coloured in grey for C $^{\alpha}$ atoms and BC Interface residues coloured green for C $^{\alpha}$ atoms.....	12
Figure 1. 6: A schematic illustrating the reaction mechanism for the conversion of chorismate to prephenate.	13
Figure 1. 7: Structures of AroH (different subunits in separate colours) and AroQ (different subunits in separate colours) chorismate mutases.	15
Figure 1. 8: A schematic illustrating the pathway branch for tryptophan synthesis and the crystal structure of the bifunctional PRA isomerase-IGP synthase from <i>E. coli</i> , PDB code 1PII (PRA isomerase and IGP synthase domains are coloured yellow and green respectively) (51).....	18
Figure 1. 9: The structure of the channelling system embedded within the tryptophan synthase from <i>S. typhimurium</i> (52).	19
Figure 1. 10: Different modes explaining allosteric regulation.	21
Figure 1. 11: The cartoons coloured purple and blue corresponds to active enzymes with different conformations, whereas the orange one represents	

the inactive enzyme bound to an allosteric inhibitor (red pentagon).	22
Figure 1. 12: Crystal structures and homology model of <i>Lmo</i> DAH7PS and <i>Gsp</i> DAH7PS.	25
Figure 1. 13: Non-covalent interaction between DAH7PS and CM. The core DAH7PS catalytic tetramers are coloured blue, CM dimers are shown in red.	27
Figure 2. 1: CLANS clustering of type I DAH7PS sequences. Each dot represents a protein sequence.	35
Figure 2. 2: Sequence alignment of <i>Pni</i> DAH7PS with other type I β DAH7PS sequences, demonstrating the <i>Pni</i> DAH7PS sequence responsible for DAH7PS and CM activities.	36
Figure 2. 3: SDS-PAGE analysis shows the expression and purification of <i>Pni</i> DAH7PS.	38
Figure 2. 4: The extent of activation of <i>Pni</i> DAH7PS ^{WT} by different metal ions in comparison to Mn ²⁺ at a concentration of 100 μ M.	40
Figure 2. 5: The response of <i>Pni</i> DAH7PS ^{WT} to varying concentrations of prephenate.	42
Figure 2. 6: Inhibition of the CM activity of <i>Pni</i> DAH7PS ^{WT} by prephenate.	43
Figure 2. 7: Analytical SEC elution trace and standard curve of logMW against K _{av} for <i>Pni</i> DAH7PS ^{WT} at 1.5 mg/mL (blue line), 0.75 mg/mL (pink line) and 0.3 mg/mL (yellow line).	44
Figure 2. 8: Sedimentation velocity analysis of the quaternary structure of <i>Pni</i> DAH7PS ^{WT} . The AUC buffer contained 10 mM BTP, pH 7.4 and 150 mM NaCl.	45
Figure 2. 9: Multiple sequence alignment for selecting the cutting site in <i>Pni</i> DAH7PS (highlighted with a grey background).	47
Figure 2. 10: SDS-PAGE analysis shows the expression and purification of <i>Pni</i> DAH7PS ^D	48

Figure 2. 11: SDS-PAGE analysis shows the expression and purification of <i>PniDAH7PS^{CM}</i>	50
Figure 2. 12: The extent of activation of <i>PniDAH7PS^D</i> by different metal ions in comparison to Mn^{2+} at a concentration of 100 μ M.....	51
Figure 2. 13: The response of <i>PniDAH7PS^D</i> (red curve) and <i>PniDAH7PS^{WT}</i> (black curve) to varying concentrations of prephenate.	53
Figure 2. 14: The quaternary structure of <i>PniDAH7PS^D</i> determined by analytical SEC and AUC.....	54
Figure 2. 15: Analytical SEC elution trace and standard curve of logMW against K_{av} for <i>PniDAH7PS^{CM}</i> at 1.0 mg/mL (red line).....	56
Figure 2. 16: Analysis of the SEC-SAXS data obtained for <i>PniDAH7PS^D</i> (blue) and <i>PniDAH7PS^{CM}</i> (red).....	58
Figure 2. 17: Analysis of the SAXS profiles of <i>PniDAH7PS^{WT}</i> in the absence (blue) or presence (red) of prephenate.....	60
Figure 2. 18: <i>Ab initio</i> modelling for <i>PniDAH7PS^{WT}</i> in the absence or presence of prephenate using the experimental SAXS profiles.	62
Figure 2. 19: Crystals of <i>PniDAH7PS^D</i>	66
Figure 2. 20: Crystals of <i>PniDAH7PS^{CM}</i>	67
 Figure 3. 1: Schematics illustrating the kinetic assays for <i>PniDAH7PS^{WT}</i>	75
Figure 3. 2: UV absorption spectra of PEP (Blue) and chorismate (Red).....	76
Figure 3. 3: The promoting effect of the DAH7PS catalysed reaction on the CM activity of <i>PniDAH7PS^{WT}</i>	77
Figure 3. 4: Analysis of the SAXS data of apo- <i>PniDAH7PS^{WT}</i> (black) and PEP/ Mn^{2+} - <i>PniDAH7PS^{WT}</i> (blue) and PEP/ Mn^{2+} /E4P- <i>PniDAH7PS^{WT}</i> (red).	80
Figure 3. 5: Rigid body modelling for apo- <i>PniDAH7PS^{WT}</i> , PEP/ Mn^{2+} - <i>PniDAH7PS^{WT}</i> and PEP/ Mn^{2+} /E4P- <i>PniDAH7PS^{WT}</i> using SAXS profiles. DAH7PS and CM domains are respectively coloured blue and red.	83
Figure 3. 6: ITC experiments to detect the effect of Mn^{2+} on the binding of PEP to	

<i>Pni</i> DAH7PS ^{WT}	84
Figure 3. 7: Sequence and structural alignments predict a long loop within <i>Pni</i> CM.	87
Figure 3. 8: Loop truncation within the CM portion of <i>Pni</i> DAH7PS ^{WT}	88
Figure 3. 9: The loop truncation for generating <i>Pni</i> DAH7PS ^{Δ5AA} impairs the activating effect of the addition of PEP/Mn ²⁺ /E4P on the catalytic activity of CM.....	90
Figure 3. 10: The comparison of the SAXS data obtained for <i>Pni</i> DAH7PS ^{Δ5AA} with that obtained for <i>Pni</i> DAH7PS ^{WT}	93
Figure 3. 11: The analysis of the SAXS profiles of apo- <i>Pni</i> DAH7PS ^{WT} and PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT} in the absence or presence of NaCl.	97
Figure 3. 12: ITC assays showing the thermodynamics underlying the binding of PEP to <i>Pni</i> DAH7PS in absence or presence of NaCl.	100
Figure 3. 13: Rigid body models of <i>Pni</i> DAH7PS ^{WT} show the hetero-interaction between the DAH7PS (blue) and CM (red) domains respectively in the absence of any ligand (A) and in the presence of either PEP/Mn ²⁺ (B) or PEP/Mn ²⁺ /E4P (C).....	103
Figure 3. 14: Schematic representation of the proposed regulation loop mediated by <i>Pni</i> DAH7PS in the shikimate pathway.	105
Figure 4. 1: Comparison of SAXS profiles of <i>Pni</i> DAH7PS ^{WT} in the presence of prephenate (black), PEP/Mn ²⁺ (red) and PEP/Mn ²⁺ /E4P (blue).....	111
Figure 4. 2: Comparison of SAXS profiles for <i>Pni</i> DAH7PS ^{WT} in the presence of 350 μM prephenate respectively combined with 0 μM (black), 100 μM (red) and 500 μM (blue) of both PEP and MnSO ₄	113
Figure 4. 3: The analysis of SAXS profiles shows the combined effect of prephenate and PEP/Mn ²⁺ /E4P on the structure of <i>Pni</i> DAH7PS ^{WT}	115
Figure 4. 4: Rigid body modelling for PA- <i>Pni</i> DAH7PS ^{WT} , PA-PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT} and PA-PEP/Mn ²⁺ /E4P- <i>Pni</i> DAH7PS ^{WT} and fits to the corresponding	

experimental SAXS profiles.	116
Figure 4. 5: Response of the DAH7PS activity of either <i>Pni</i> DAH7PS ^{WT} (black) or <i>Pni</i> DAH7PS ^{Δ5AA} (red) to increasing concentrations of prephenate.	118
Figure 4. 6: Comparison between the SAXS profiles of PA- <i>Pni</i> DAH7PS ^{Δ5AA} , apo- <i>Pni</i> DAH7PS ^{Δ5AA} , PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{Δ5AA} and PA- <i>Pni</i> DAH7PS ^{WT}	120
Figure 4. 7: Response of the DAH7PS activity of <i>Pni</i> DAH7PS ^{WT} to increasing concentrations of prephenate in the absence (black) or presence of 150 mM NaCl (red).	121
Figure 4. 8: Comparison of the SAXS profiles for PA- <i>Pni</i> DAH7PS ^{WT} recorded in the presence of 0 mM (black), 150 mM (red) or 500 mM (blue) NaCl.	123
Figure 4. 9: ITC assays showing the thermodynamics of PEP binding to <i>Pni</i> DAH7PS ^{WT} and the effect of prephenate of PEP binding.	124
Figure 4. 10: Sedimentation velocity analysis for <i>Gsp</i> DAH7PS ^{WT} and <i>Gsp</i> DAH7PS ^D at pH 7.4 and 20 °C recorded at 50,000 rpm.	125
Figure 4. 11: Response of <i>Gsp</i> DAH7PS ^{WT} (blue) and <i>Gsp</i> DAH7PS ^D (red) DAH7PS activity to increasing concentrations of prephenate at 60 °C (20).	126
Figure 4. 12: Analysis of the SAXS data for <i>Gsp</i> DAH7PS in the absence (blue) or presence (red) of prephenate.	128
Figure 4. 13: The distribution of open and closed conformations in <i>Gsp</i> DAH7PS with or without prephenate in solution.	130
Figure 5. 1: The schematic showing the mechanism of chemical ligation, thioester—cysteine condensation and rearrangement.	137
Figure 5. 2: Schematics showing the production of <i>Pni</i> DAH7PS ^{CM-T263C} (A) and the EPL reaction (B).	139
Figure 5. 3: Reaction of EPL and purification of the ligated product, <i>Pni</i> DAH7PS ^{D-CM}	141
Figure 5. 4: Schematics illustrating the homodimeric and heterodimeric product produced by EPL.	143

Figure 5. 5: SEC-IMAC two-step purification of the ligated product, <i>Pni</i> DAH7PS ^{D-CM}	144
Figure 5. 6: Analytical SEC examining the oligomeric state of <i>Pni</i> DAH7PS ^{D-CM} ...	146
Figure 5. 7: A bar chart shows the values of k_{cat}/K_M for the DAH7PS and CM activities of <i>Pni</i> DAH7PS ^{WT} , truncated variants (<i>Pni</i> DAH7PS ^D and <i>Pni</i> DAH7PS ^{CM}) and <i>Pni</i> DAH7PS ^{D-CM} , indicating restoration of the functions of DAH7PS and CM by EPL.....	148
Figure 5. 8: The response of <i>Pni</i> DAH7PS ^D (blue curve), <i>Pni</i> DAH7PS ^D (red curve), and <i>Pni</i> DAH7PS (black curve) to varying concentrations of prephenate. ..	149
Figure 5. 9: Schematic showing the contrast matching and the segmentally deuterated <i>Pni</i> DAH7PS ^{D-CM} for SANS examinations.....	151
Figure 5. 10: SANS scattering profiles using contrast variation of ligand-free segmentally deuterated <i>Pni</i> DAH7PS ^{D-CM}	153
Figure 5. 11: Rigid body modelling for apo- <i>Pni</i> DAH7PS ^{D-CM} against the SAXS and SANS data.....	154
Figure 5. 12: Analysis of the SAXS profiles obtained for apo- <i>Pni</i> DAH7PS ^{D-CM} (black) and PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT} (red) and PA- <i>Pni</i> DAH7PS ^{WT} (blue).....	157
Figure 5. 13: Pre-examinations of ¹⁵ N singly labelled <i>Pni</i> DAH7PS ^{WT} (0.21 mM) using ¹ H- ¹⁵ N HSQC and TROSY. Spectra were collected at a proton resonance frequency of 600 MHz, at 298 K, in 10 mM BTP buffer at pH 7.3 containing 150 mM NaCl, 50 mM Glu, 50 mM Arg and 10 mM DTT.	160
Figure 5. 14: Examinations of ¹⁵ N singly labelled <i>Pni</i> DAH7PS ^{CM} (0.2 mM) and <i>Pni</i> DAH7PS ^D (0.3 mM) using ¹ H- ¹⁵ N HSQC at a proton resonance frequency of 600 MHz.	161
Figure 5. 15: The expression and purification of ² H/ ¹³ C/ ¹⁵ N <i>Pni</i> DAH7PS ^D	162
Figure 6. 1: Schematic illustrating the proposed evolutionary trajectory of the DAH7PS family.	171

Figure A.2. 1: Michaelis-Menten plots of the <i>Pni</i> DAH7PS ^{WT} for the K_M of PEP, E4P and chorismate respectively.	217
Figure A.2. 2: Michaelis-Menten plots of the <i>Pni</i> DAH7PS ^D for the K_M of PEP and E4P, and of the <i>Pni</i> DAH7PS ^{CM} for the K_M of chorismate respectively.	218
Figure A.2. 3: Crystals of <i>Pni</i> DAH7PS ^D	219
Figure A.5. 1: Schematic diagram of pTXB1-DAH7PS-reactant for the expression of His- <i>Pni</i> DAH7PS ^{D-H262} -intein-CBD.....	225
Figure A.5. 2: The isolations of <i>Pni</i> DAH7PS ^{D-CM} from <i>Pni</i> DAH7PS ^{CM-T263C} using hydrophobic interaction chromatography (A) and anion exchange chromatography (B).	226
Figure A.5. 3: Michaelis-Menten plots of the <i>Pni</i> DAH7PS ^{D-H262} for the K_M of PEP and E4P, and of the <i>Pni</i> DAH7PS ^{CM-T263C} for the K_M of chorismate respectively.	227
Figure A.5. 4: Michaelis-Menten plots of the <i>Pni</i> DAH7PS ^{D-CM} for the K_M of PEP, E4P and chorismate respectively.	227
Figure A.5. 5: Scattering intensities and distance distribution functions of geometrical bodies (164).	228
Figure A.5. 6: SANS scattering using contrast variation of ligand free <i>Pni</i> DAH7PS ^{D-CM} (black) and the <i>Pni</i> DAH7PS ^{D-CM} respectively in the presence of PEP/Mn ²⁺ (red) and prephenate (blue).	229

List of Tables

Table 1: Variations in type I β DAH7PS.....	10
Table 2. 1: Table shows the sequence identities between <i>Pni</i> DAH7PS and other homologous DAH7PS and CM enzymes.....	37
Table 2. 2: Table shows the kinetic parameters of the DAH7PS activity of <i>Pni</i> DAH7PS ^{WT} compared to other type I β DAH7PS enzymes.....	41
Table 2. 3: Table shows the kinetic parameters of the CM activity of <i>Pni</i> DAH7PS ^{WT} compared to that of <i>Gsp</i> DAH7PS, <i>Pgi</i> DAH7PS and <i>Pfu</i> CM	41
Table 2. 4: Table of the kinetic parameters for the DAH7PS activities of the wild-type <i>Pni</i> DAH7PS, <i>Pgi</i> DAH7PS, <i>Gsp</i> DAH7PS, <i>Bsu</i> DAH7PS and their truncated DAH7PS variants.....	52
Table 2. 5: Table of the kinetic parameters of the CM activity of the wild-type <i>Pni</i> DAH7PS, <i>Pgi</i> DAH7PS, <i>Gsp</i> DAH7PS, <i>Bsu</i> DAH7PS and their truncated CM variants.....	55
Table 2. 6: Table of the SEC-SAXS parameters of <i>Pni</i> DAH7PS ^D , <i>Pni</i> DAH7PS ^{CM} and <i>Pni</i> DAH7PS ^{WT} in the absence or presence of prephenate	59
Table 2. 7: Table of the conditions for the crystallisation of <i>Pni</i> DAH7PS ^D	64
Table 2. 8: Assessment of the quality of the diffraction data set of <i>Pni</i> DAH7PS ^D using aimless (127).....	65
Table 2. 9: Table of the conditions for the crystallisation of <i>Pni</i> DAH7PS ^{CM}	67
Table 3. 1: SAXS parameters of apo- <i>Pni</i> DAH7PS ^{WT} , PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT} and PEP/Mn ²⁺ /E4P- <i>Pni</i> DAH7PS ^{WT}	79
Table 3. 2: Contact information for rigid body modelling using BUNCH	82
Table 3. 3: Kinetic parameters of the DAH7PS and CM activities of <i>Pni</i> DAH7PS ^{Δ5AA} and a comparison to kinetic parameters of <i>Pni</i> DAH7PS ^{Δ5AA} , <i>Pni</i> DAH7PS ^{WT} , <i>Pni</i> DAH7PS ^D , and <i>Pni</i> DAH7PS ^{CM}	89
Table 3. 4: SAXS parameters for apo- <i>Pni</i> DAH7PS ^{Δ5AA} , PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{Δ5AA} and	

PEP/Mn ²⁺ /E4P- <i>Pni</i> DAH7PS ^{Δ5AA} compared to the parameters for <i>Pni</i> DAH7PS ^{WT}	92
Table 3. 5: Kinetic parameters for the DAH7PS and CM activities of <i>Pni</i> DAH7PS ^{WT} in the absence or presence of 150 mM or 500 mM NaCl	95
Table 3. 6: SAXS parameters for apo- <i>Pni</i> DAH7PS ^{WT} and PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT} in the presence of varying concentrations of NaCl	96
Table 4. 1: Comparison of the SAXS parameters of PA- <i>Pni</i> DAH7PS ^{WT} , PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT} and PEP/Mn ²⁺ /E4P- <i>Pni</i> DAH7PS ^{WT}	110
Table 4. 2: Comparison of SAXS parameters of PA-PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT} with that of PA- <i>Pni</i> DAH7PS ^{WT} and PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT}	112
Table 4. 3: Comparison of SAXS parameters for PA-PEP/Mn ²⁺ /E4P- <i>Pni</i> DAH7PS ^{WT} with that of PA- <i>Pni</i> DAH7PS ^{WT} , PEP/Mn ²⁺ /E4P- <i>Pni</i> DAH7PS ^{WT} and PA-PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{WT}	114
Table 4. 4: Comparison of the SAXS parameters of apo- <i>Pni</i> DAH7PS ^{Δ5AA} , PEP/Mn ²⁺ - <i>Pni</i> DAH7PS ^{Δ5AA} , PA- <i>Pni</i> DAH7PS ^{Δ5AA} and PA- <i>Pni</i> DAH7PS ^{WT}	119
Table 4. 5: SAXS parameters for PA- <i>Pni</i> DAH7PS ^{WT} in the presence of 0 mM, 150 mM or 500 mM NaCl	122
Table 4. 6: Kinetic parameters for <i>Gsp</i> DAH7PS ^{WT} , <i>Gsp</i> DAH7PS ^D , and <i>Gsp</i> DAH7PS ^{CM} (20)	126
Table 4. 7: SAXS parameters for <i>Gsp</i> DAH7PS ^{WT} in the absence or presence of prephenate	127
Table 5. 1: Comparison of the kinetic parameters for <i>Pni</i> DAH7PS ^{D-CM} , <i>Pni</i> DAH7PS ^{D-H262} and <i>Pni</i> DAH7PS ^{CM-T263C} , with that of <i>Pni</i> DAH7PS ^{WT} , <i>Pni</i> DAH7PS ^D and <i>Pni</i> DAH7PS ^{CM}	147
Table 5. 2: SANS and SAXS parameters for segmentally deuterated <i>Pni</i> DAH7PS ^{D-CM}	152
Table 5. 3: SAXS parameters determined for <i>Pni</i> DAH7PS ^{D-CM} and <i>Pni</i> DAH7PS ^{WT}	156

Table 7. 1: Primers for construction of expression vectors.....	183
Table 7. 2: The competent cells used in this project.....	188
Table 7. 3: Expression constructs generated in this project	189
Table 7. 4: Compositions of cell lysis buffer	191
Table 7. 5: Extinction coefficients (ϵ_{280}) for each protein	196
Table 7. 6: Contact information for the rigid body modelling of <i>Pni</i> DAH7PS ^{D-CM} using SASRAFCV	205
Table 7. 7: Cross-correlation information for the rigid body modelling of <i>Pni</i> DAH7PS ^{D-CM} using SASRAFCV	206
Table 7. 8: ITC experimental set-up.....	207
Table 7. 9: The recipe for the M9 minimal medium	210
Table 7. 10: The recipes for NG and N-5052 minimal medium.....	212
Table 7. 11: The recipes of the buffers for NMR samples	213
 Table A.3. 1: SAXS parameters for apo- <i>Pni</i> DAH7PS ^{WT} and the <i>Pni</i> DAH7PS ^{WT} in the presence of either 350 μ M PEP or 350 μ M MnSO ₄	 221

Abbreviations

ACT	aspartate kinase,chorismate mutase and TyrA
AEC	anion exchange chromatography
<i>Ape</i>	<i>Aeropyrum pernix</i>
ATCase	aspartate transcarbamoylase
AUC	analytical ultracentrifugation
BSA	bovine serum albumin
<i>Bsu</i>	<i>Bacillus subtilis</i>
BTP	1,3-bis[tris(hydroxymethyl)methylamino]propane
CBD	chitin binding domain
CdRP	1-(<i>o</i> -carboxyphenylamino)-1-desoxyribulose-5-phosphate
<i>Cgl</i>	<i>Corynebacterium glutamicum</i>
CLANS	CLuster ANalysis of Sequences
CM	chorismate mutase
DAH7PS	3-deoxy-D- <i>arabino</i> -heptulosonate 7-phosphate synthase
$D_{\max}$	the maximum particle dimension
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E4P	erythrose 4-phosphate
<i>Eco</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
EPL	expressed protein ligation
ESI-MS	electrospray ionisation mass spectrometry
G3P	glyceraldehyde 3-phosphate
<i>Gsp</i>	<i>Geobacillus sp</i>
GST	glutathione sulfur-transferase
HIC	hydrophobic interaction chromatography

HSQC	heteronuclear single quantum correlation
IGP	indole-3-glycerol phosphate
IMAC	immobilised metal affinity chromatography
IPTG	isopropyl β -D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
KDO8PS	3-deoxy-D- <i>manno</i> -2-octulosonate-8-phosphate synthase
<i>Lmo</i>	<i>Listeria monocytogenes</i>
MESNA	2-mercaptoethanesulfonic acid
<i>Mtu</i>	<i>Mycobacterium tuberculosis</i>
MWC	Monod-Wyman-Changeux model
<i>Nme</i>	<i>Neisseria meningitidis</i>
NMR	nuclear magnetic resonance
ORF	open reading frame
$P(r)$	distribution pair-wise distance distribution
PA	prehenate
<i>Pae</i>	<i>Pseudomonas aeruginosa</i>
PCR	polymerase chain reaction
PDB	Protein Data Bank
PDH	prephenate dehydrogenase
PDT	prephenate dehydratase
PEP	phosphoenolpyruvate
<i>Pfu</i>	<i>Pyrococcus furiosus</i>
<i>Pgi</i>	<i>Porphyromonas gingivalis</i>
PISA	Proteins, Interfaces, Structures and Assemblies
<i>Pni</i>	<i>Prevotella nigrescens</i>
<i>Pni</i> DAH7PS ^{CM}	the CM domain of <i>Pni</i> DAH7PS
<i>Pni</i> DAH7PS ^D	the DAH7PS domain of <i>Pni</i> DAH7PS
PRA	N-(5'-phosphoribosyl) anthranilate
R_g	the radius of gyration

SANS	small-angle neutron scattering
SAXS	small-angle X-ray scattering
<i>Sce</i>	<i>Saccharomyces cerevisiae</i>
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SLD	scattering length density
SOC	super optimal broth
TAE	tris-acetate-EDTA
TCEP	tris(2-carboxyethyl)phosphine
TIM	triosephosphate isomerase
<i>Tma</i>	<i>Thermotoga maritima</i>
Tris	tris(hydroxymethyl)aminomethane
TROSY	transverse relaxation optimised spectroscopy
TSA	transition state analogue
V_p	Porod volume

Acknowledgement

First and foremost, I would like to express my most sincere gratitude and appreciation towards my supervisor Professor - Emily Parker. Her outstanding professionalism, passion for science, profound knowledge and patience has guided me into the incredibly fascinating research area of protein. With her support and encouragement, I was able to fully concentrate on my project which has been accompanied by the joy of researching. I will be forever grateful for her mentorship and the opportunity of research training which she has provided.

Many thanks to Dr. Wanting Jiao and Dr. Yifei Fan who have provided the detailed reviews and valuable advices on my thesis. Their consideration and comfort helped me a lot in getting through the difficult times during my writing. I also do thoroughly appreciate all of our pleasant discussions on the research and other subjects, which really has broadened my perspective and provided me a lot of enjoyment. Thanks to Dr. Eric Lang for our collaboration on the SAXS analysis of *GspDAH7PS* as well as generously sharing his great ideas and research experiences related to protein dynamics and bioinformatics. Thanks to Dr Oliver Sterritt for reading my entire thesis and suggesting valuable corrections, and for our interesting talks on DAH7PS and crystallography. Thanks to Dr. Leyla Bustamante for lab maintenance. Thanks to Preeti Kundu for our pleasant conversations about many subjects and for your kind help on my chemical synthesis work. Also, thank you to all other current and past members in Parker's group, in particular, Dr Kyle Van de Bittner, Dr Rosannah Cameron, Candice Gu, Parastoo Khajeaian, Mohammad Firoozinia, Fiona Given and Michael Weusten, for being willing to help in any way possible.

Special thanks go to Dr Thomas Cotton, Dr Gerd Mittelstädt, Dr Nicola Blackmore, Dr Gert-Jan Moggré and Dr Logan Heye, for helping me to get started in the lab and

providing consistent strong support to me during my PhD years.

Considerable thanks go to Professor Joel Mackay and his research group for their efficient work, communication and advice on our collaborative investigation of *Pni*DAH7PS structure using NMR technology. I am grateful to Dr Anthony Duff for his great job of producing the deuterated sample on the SANS experiments, and also for patiently teaching me the proper method of triple isotopic labeling of protein. I would like to thank Dr Andrew Whitten for his significant help in collecting and analysing SANS data. Thanks to Professor Geoffrey Jameson and Dr Santosh Panjikar for providing me valuable knowledge and suggestions on crystallography.

I would also like to thank all my friends in New Zealand - Qingwei Li, Haizhou Huang, Tingting Xie, Hui Ye, Ning Wang and Yuxin Jin, for our sincere friendship. Your company really made my life easier and more colourful over the past years!

Finally, I would like to dedicate this thesis to my parent for their love, endless support and encouragement throughout my life! The words can never be enough to describe how thankful I am to the both of you. I'll do everything to repay for what you have done for me.

Chapter 1 *Introduction*

1.1 Shikimate pathway

Aromatic amino acids are essential building blocks for protein synthesis and are important precursors for a range of metabolites that are critical for maintaining life on the earth, from viruses to humans. Twenty standard amino acids are essential for basic physiological activities, and the ability to synthesise all these amino acids *de novo* is found in microorganisms, plants, and apicomplexan parasites, only, whereas most mammals lack the ability to synthesise nine of these amino acids, including the aromatic amino acids Phe and Trp (1-3).

In microorganisms and plants, the metabolic pathway that is responsible for the biosynthesis of the aromatic amino acids, and additionally folic acid and salicylate, is the shikimate pathway (Figure 1.1). The absence of this pathway from higher organisms makes the enzymes of this pathway potential targets for the development of novel antibiotics and herbicides (1-3).

Figure 1. 1: The shikimate pathway leads to the biosynthesis of the aromatic amino acids. The dashed rectangle and circle highlight the steps of the synthesis of DAH7P and prephenate respectively.

1.2 DAH7PS

The first committed step of the shikimate pathway is the aldol-like condensation reaction between phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) to produce 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAH7P). This step is catalysed by DAH7P synthase (DAH7PS) (Figure 1.2).

Figure 1. 2: The schematic diagram shows the reaction mechanism for the formation of DAH7P.

The biosynthesis of aromatic amino acids is regulated in response to metabolic demands in order to best utilise cellular resources, such as carbon and energy (4). Situated at the first step of the shikimate pathway, DAH7PS is a major regulatory control point for the pathway. DAH7PS is subject to allosteric inhibition by pathway end products (Phe, Tyr and Trp) or intermediates (chorismate and prephenate). This allosteric control allows for the regulation of the biosynthetic flux through the pathway in response to the cellular requirements for the pathway products (5-8). Allosteric regulation of enzyme activity occurs when the activity of an enzyme is remotely controlled by an effector, which binds at a position distant from the active site. More detailed information about allosteric regulation will be introduced in section 1.4.2. The crucial role that DAH7PS plays as both an initiator for the biosynthesis of key metabolites and a regulatory control point for the shikimate pathway means that this enzyme is a valuable target for both the development of antibiotics and for probing metabolic regulatory mechanisms.

1.2.1 The DAH7PS family

Two main DAH7PS groups have been identified. These groups are discriminated by amino acid sequence similarity and their molecular weights. Based on these sequence characteristics, DAH7PSs are divided as type I and type II DAH7PS (9). Generally, the molecular weight of type I DAH7PS enzymes is lower than 40 kDa, whereas type II DAH7PS enzymes are larger than 50 kDa. Type I and type II enzymes share less than 10% sequence identity (9).

Type I enzymes are further divided into two groups, denoted type I α and type I β , based on sequence identity (10-12) (Figure 1.3). Type I β DAH7PS enzymes share sequence similarity to a related enzyme, 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase (KDO8PS) (Figure 1.3) (9). KDO8PS catalyses an aldol-like condensation reaction between PEP and five-carbon sugar, arabinose 5-phosphate (A5P), resembling the reaction that is catalysed by DAH7PS. Through the analysis of multiple sequence alignments and the comparison of crystal structures, it has been proposed that type I DAH7PSs and KDO8PS enzymes share a common evolutionary origin (11,13-18).

Figure 1. 3: Classification of DAH7PS. DAH7PSs are divided into two main groups, type I and type II DAH7PS; type I is further classified into Ia and Iβ (including KDO8PS) subgroups. Structures of the monomeric units and homotetrameric assemblies, representative of each DAH7PS sub-type, are shown in Figure 1.3 (16,17,19-22). The α-helices and β-strands in the catalytic core barrels are coloured green and blue respectively and the α-helices and β-sheets in decorated elements, extensions and insertion, are shown in red and orange respectively.

All available structures of DAH7PS enzymes share a common (β/α)₈ barrel fold (TIM barrel, named after triosephosphate isomerase), which is responsible for catalysis and contains highly conserved key active site residues (Figure 1.3) (5,12). The active site of the DAH7PS family comprises a divalent metal, PEP and E4P binding sites. A divalent metal is essential for the catalytic activity of all DAH7PS enzymes. The metal coordination site is located close to the C-terminal end of the barrel (14). In the structure of the DAH7PS from *Pyrococcus furiosus* (*Pfu*DAH7PS), the metal ion, Mn^{2+} , is coordinated to the enzyme via residues located in the short loops $\beta 1-\alpha 1$ (Cys31) and $\beta 7-\alpha 7$ (His201), strand $\beta 8$ (Glu227) and on the long loop $\beta 8-\alpha 8$ (Asp238) (Figure 1.4 A). PEP is held by three Arg residues located on strand $\beta 2$, loops $\beta 4-\alpha 4$ and $\beta 6-\alpha 6$, and two Lys residues on loop $\beta 2-\alpha 2$ and strand $\beta 5$ (Arg55, 115, 166 and Lys60, 136 in *Pfu*DAH7PS) (Figure 1.4B). However, no plausible atomic structure showing the binding site of E4P is available yet. Through analysis based on the crystal structure of the DAH7PS from *Saccharomyces cerevisiae* (*Sce*DAH7PS) in complex with a metal ion (Co^{2+}), PEP and glycerol 3- phosphate (G3P), a probable E4P binding pocket has been identified (15), which is formed by a KPR(T/S) motif that is conserved across all DAH7PS enzymes (11,12). A structural model presenting the E4P binding mode to *Pfu*DAH7PS (numbering Lys60, Pro61, Arg62, Thr63) was subsequently generated (Figure 1.4 C) (23).

Figure 1. 4: The architecture of the *Pfu*DAH7PS active site (PDB code 1ZCO) (23). (A) The residues involved in coordinating Mn²⁺ (magenta sphere) are indicated as pink stick models. (B) The residues interacting with PEP (orange stick model) are shown as yellow stick models. (C) The conserved KPRT segment responsible for E4P binding is shown using cyan stick models.

Although sharing a common core structure, there are structural variations among different DAH7PS sub-types due to the variable additional structural elements that are fused to the core catalytic TIM barrel. These various additional elements include loop insertions and C- or N-terminal extensions that are responsible for the different patterns of allosteric control. The allosteric inhibitory machinery of type II DAH7PS proteins, such as the DAH7PS from *Mycobacterium tuberculosis* (*Mtu*DAH7PS), is a

consequence of extensive modification to the core DAH7PS barrel, with a double- α -helix insertion and an N-terminal extension consisting of a β -strand and three helices (Figure 1.3) (5,24). The single binding site for an allosteric effector is located in the N-terminal barrel extension of type I α DAH7PS enzymes and plays a crucial role in the inhibitory effect of aromatic amino acids on this DAH7PS subtype (Figure 1.3); such as the DAH7PS from *Neisseria meningitidis* (*NmeDAH7PS*) and the Phe-regulated DAH7PS from *Escherichia coli* (*EcoDAH7PS* (Phe)) (22,25,26).

Compared to type I α and type II DAH7PSs, more structural variations are observed in the I β subtype of the DAH7PS family (Table 1). *PfuDAH7PS* and the DAH7PS from *Aeropyrum pernix* (*ApeDAH7PS*) are unregulated enzymes without any additional structural decorations. In comparison, the DAH7PS from *Thermotoga maritima* (*TmaDAH7PS*) has a regulatory N-terminal ACT domain (named after three of the proteins that contain it: **A**spartate kinase, **C**horismate mutase and **T**yrA) for binding Tyr (7,18,23) (Figure 1.3). In addition, a functional chorismate mutase domain is found fused to the N-terminus of the DAH7PSs from *Bacillus subtilis* (*BsuDAH7PS*), *Geobacillus sp.* (*GspDAH7PS*) and *Listeria monocytogenes* (*LmoDAH7PS*) (8,27,28). This type of bifunctional chorismate mutase-DAH7PS is allosterically inhibited by prephenate (8,28) (Figure 1.3). More details about this type of bifunctional enzyme will be presented in section 1.4.2.2.

Table 1: Variations in type I β DAH7PS.

Decorated Structure	Representative	Allosteric Inhibitor
None	<i>PfuDAH7PS</i> and <i>ApeDAH7PS</i>	None
N terminal ACT domain	<i>TmaDAH7PS</i>	Tyr
N terminal CM domain	<i>BsuDAH7PS</i> , <i>LmoDAH7PS</i> and <i>GspDAH7PS</i>	Prephenate

1.2.2 The effect of quaternary structure on DAH7PS functions

The way in which protein subunits assemble to form an oligomeric enzyme determines a range of important characteristics of an enzyme including catalytic activity, thermal stability and allosteric regulation.

The quaternary structure of DAH7PS enzymes has been determined for a variety of organisms, including *Eco*DAH7PS (Phe) (type I α) (14), *Sce*DAH7PS (type I α) (15), *Nme*DAH7PS (type I α) (22), *Pfu*DAH7PS (type I β) (23), *Ape*DAH7PS (type I β) (18), *Lmo*DAH7PS (type I β) (28), *Gsp*DAH7PS (type I β) (20), *Tma*DAH7PS (type I β) (17), *Mtu*DAH7PS (type II) (29), the DAH7PS from *Corynebacterium glutamicum* (*Cgl*DAH7PS, type II) (30) and the DAH7PS from *Pseudomonas aeruginosa* (*Pae*DAH7PS, type II) (31). Without exception, all of these DAH7PS enzymes, including both type I and type II enzymes, adopt homotetrameric assemblies, which alludes to an essential role of this oligomeric organisation in the maintenance of the normal function of DAH7PS. The investigation on the quaternary structure of *Pfu*DAH7PS (type I β), using *PISA* (Proteins, Interfaces, Structures and Assemblies) (32), identified two distinct types of interfaces within the homotetrameric structure (Figure 1.5 A) (33). One interface of *Pfu*DAH7PS, between the chain A and B (or chain C and D, denoted as AB interface), is in the vicinity of the active site, and buries a number of residues which are extensively involved in catalysis; whereas another interface, formed by the chain B and C (or chain A and D, denoted as BC interface), is located further from the active site and is formed by fairly strong hydrophobic interactions. The perturbation of the BC interface, via single point mutagenesis, profoundly reduced the thermostability of *Pfu*DAH7PS (33) without disrupting the enzyme's catalytic activity.

Figure 1. 5: Quaternary assemblies of *Pfu*DAH7PS (PDB code 1ZCO) (23), *Nme*DAH7PS (PDB code 5DCB) (22) and *Mtu*DAH7PS (PDB code 2B7O) (21) with AB interface residues coloured in grey for C $^{\alpha}$ atoms and BC Interface residues coloured green for C $^{\alpha}$ atoms. PEP and Mn $^{2+}$ molecules bound to *Pfu*DAH7PS are shown in spheres which are coloured orange and purple respectively.

The similar differentiation of interfaces (AB interface and BC interface) was also observed in *Nme*DAH7PS (type Ia) (34) and *Mtu*DAH7PS (type II) (35) (Figure 1.5 B, C). The disruption of the tetramer interface of either *Nme*DAH7PS or *Mtu*DAH7PS, which corresponds to the BC interface of *Pfu*DAH7PS, slightly impaired the thermostability of *Nme*DAH7PS (34), whereas completely abolished the allosteric functionality of *Mtu*DAH7PS (35). In either case, just as what was detected in *Pfu*DAH7PS, the native catalytic function was not substantially affected by the perturbation of tetramer interface (34,35), which suggests that the basic catalytic unit of DAH7PS may be a dimer consisting of A and B subunits (or C and D subunits) or even a monomer containing an intact active site.

1.3 Chorismate Mutase

Chorismate mutase (CM), another pivotal enzyme in the shikimate pathway, catalyses the conversion of chorismate to prephenate by a pericyclic Claisen rearrangement (Figure 1.6). Subsequently, prephenate is converted to the precursors for Tyr and Phe biosynthesis by prephenate dehydrogenase (PDH) and prephenate dehydratase (PDT), respectively (36-38). In microorganisms, CMs are divided into two groups, AroQ and AroH, with the two groups having totally different structures despite catalysing the same chemical reaction (39,40) (Figure 1.7). The AroH class of CMs are unregulated monofunctional enzymes. They are mainly found in Gram-positive bacteria of the *Bacillus/Clostridia* group (41). The AroH type of CMs are homo-trimeric containing three symmetric pseudo- α/β barrels with a β -sheet core which is surrounded by α -helices. The active sites are formed in clefts between two adjacent chains (42) (Figure 1.7 A).

Figure 1. 6: A schematic illustrating the reaction mechanism for the conversion of chorismate to prephenate.

In contrast with AroH CMs, the AroQ class of CMs is widely spread in microorganisms, and these enzymes exhibit an all α -helical fold and function as homodimers. AroQ CM enzymes are further subclassified as AroQ $_{\alpha}$, AroQ $_{\beta}$, AroQ $_{\gamma}$ and AroQ $_{\delta}$ on the basis of different structural folds (43). AroQ $_{\alpha}$ is the most common type of CM, consisting of two identical polypeptide chains, each of which is folded in a triple helical structure. The active site of AroQ $_{\alpha}$ CM is formed by residues from both chains (44) (Figure 1.7 B), and the catalytic activity is unregulated. Some AroQ $_{\alpha}$ CMs are monofunctional, such as

the CMs from *A. pernix* and *P. furiosus* (*ApeCM* and *PfuCM*). Whereas, some others exist as bifunctional enzymes where the CM moiety is part of a larger protein – and linked to N-terminal or C-terminal domains with a different functionality. Some CM portions can act as an allosteric domain exerting regulation on the other enzymatic moiety coexpressed with CM in the same peptide chain, for instance, the CM/DAH7PS bifunctional enzymes from *L. monocytogenes*, *B. subtilis* and *Geobacillus sp.* (8,20,28) .

AroQ β CM is composed of two chains, each of which consists of 12 helices. Not like either AroH or AroQ α CM whose active sites are located at inter-chain interface(s), the active site of AroQ β CM is formed within each monomeric chain and constructed by a 4-helix bundle (45). AroQ β CM is the only type of CM possessing allosteric regulation of CM activity found to date, which is allosterically activated by Trp and inhibited by Tyr (45,46) (Figure 1.7 C, D). Take the CM from *S. cerevisiae* (*SceCM*) as a example, the allosteric effectors, Trp and Tyr, shares the same binding sites, which are located at the dimer interfaces (47). The transition from the binding of Tyr to the binding of Trp causes a 15° rotation of one of the two subunits of *SceCM* relative to the other. This conformational change at the dimer interface is transmitted to the active site and switch enzyme from inactive state to active site (47).

M. tuberculosis expresses two distinct CMs (*MtuCM*) – one classified as AroQ γ and one classified as AroQ δ . The AroQ γ subclass *MtuCM* is a secreted CM consisting of two chains, each of which forms an α -helical bundle with 10 helices (43). The active site of this type of CM is also located within each chain, and it does not contain a regulatory domain (43) (Figure 1.7 E). The AroQ δ subclass *MtuCM* has a very similar overall structure to enzymes of the AroQ α subclass. However, the third helix of this AroQ δ *MtuCM* is much shorter than that of AroQ α CMs, which means that this enzyme lacks two substrate-binding site residues and therefore exhibits weak catalytic activity (48) (Figure 1.4 F).

Figure 1. 7: Structures of AroH (different subunits in separate colours) and AroQ (different subunits in separate colours) chorismate mutases. (A) AroH CM is exemplified by the trimeric CM from *B. subtilis* (*BsuCM*) (PDB code 1COM) (42). The transition state analogues (TSA) for CM catalysis, bound at the active sites, are indicated by the orange stick models. **(B)** AroQ $_{\alpha}$ CM is exemplified by the dimeric CM domain of *GspDAH7PS* (PDB code 5J6F) (20). The prephenate molecules bound, at active sites, are indicated by the orange stick models. **(C)** AroQ $_{\beta}$ CM is exemplified by the dimeric CM from *S. cerevisiae* (*SecCM*) (PDB code 3CSM) (47). The TSA bound at the active sites and the Trp bound at the dimer interface are indicated by the orange and pink stick models respectively. **(D)** The structure of *SecCM* with TSA (orange stick) and Tyr (yellow stick) bound (PDB code 4CSM) (47). **(E)** AroQ $_{\gamma}$ CM is exemplified by the secreted, dimeric, CM from *M. tuberculosis* (*MtuCM*) (PDB code 2FP2) (43). The TSA bound at the active site is indicated by the orange stick model. **(F)** AroQ $_{\delta}$ CM is exemplified by the intracellular *MtuCM* dimer (PDB code 2W1A) (29). The TSA bound at active sites is indicated by the orange stick models.

1.4 The role of multifunctional enzymes and enzyme complexes in the shikimate pathway

Multifunctional enzymes and enzyme complexes are often found in metabolic pathways and have been described as evolutionary consequences that shape or improve the performance of metabolic pathways (49). The most characterised example of a beneficial consequence of a multifunctional enzyme is the boosting of the overall catalytic rate by co-localising enzymatic active sites that catalyse consecutive chemical reactions within a biosynthetic pathway (50-53). In recent years, more investigations have demonstrated that non-covalent interactions between two separate enzymes, or covalent connections between two distinct protein domains, play a significant role in the allosteric regulation of the shikimate pathway (8,20,28-30,54). These diverse multifunctional enzymes and multi-enzyme complexes are either committed to organising the catalytic entities for efficient substrate channelling or for delivering allosteric functionality. This provides an exceptional contribution to the efficient and coherent function of this important aromatic amino acid biosynthetic pathway.

1.4.1 Increased catalytic efficiency

As mentioned above, the biosynthesis of aromatic amino acids is a costly process for microorganisms, hence the appropriate subcellular distribution of pathway metabolites are crucial for the efficient maintenance of cells. However, due to the lack of internal membranes, enzymes for different physiological functions in prokaryotic cells cannot be localised in specific subcellular compartments as effectively as in eukaryotic cells. As a result, a product released by an enzyme may diffuse within the cytoplasm before it is delivered to the enzyme responsible for the subsequent chemical transformation. To avoid the diffusion of pathway intermediates, one important strategy adopted by prokaryotic cells is expressing multifunctional enzymes,

which contain multiple protein domains which exhibit activities of consecutive enzymes within the pathway (50-52). Due to the close proximity of the consecutive active sites within a multifunctional enzyme, the product of one reaction can be efficiently captured by the next active site to ensure a high utilisation rate of metabolites.

One of the most well-characterised examples of a multifunctional enzyme is phosphoribosylanthranilate isomerase-indole glycerolphosphate synthase (PRA isomerase-IGP synthase) found in the tryptophan biosynthetic branch of the shikimate pathway in *E. coli*. The biosynthesis of tryptophan from chorismate includes several steps, as shown in Figure 1.8. Initially, chorismate is transformed to *N*-(5'-phosphoribosyl) anthranilate (PRA), which is catalysed by anthranilate synthase and anthranilate 5-phosphoribosyl-1-pyrophosphate transferase. PRA is further converted to 1-(*o*-carboxyphenylamino)-1-deoxyribulose-5-phosphate (CdRP) and indole 3-glycerol phosphate (IGP) sequentially via two consecutive reactions catalysed individually by the phosphoribosylanthranilate isomerase (PRA isomerase) and indole glycerol phosphate synthase. In *E. coli*, these two enzyme catalysed reactions are carried out by one protein that contains two distinct functional domains – the bifunctional PRA isomerase-IGP synthase (Figure 1.8) (55,56). After cleavage of the two functional domains, the separated PRA isomerase and the IGP synthase maintain similar catalytic activities as compared to the full-length bifunctional protein (57,58). Although there are no direct interactions between the active sites of PRA isomerase and IGP synthase, the two active sites are located in close proximity to each other (Figure 1.8).

Figure 1. 8: A schematic illustrating the pathway branch for tryptophan synthesis and the crystal structure of the bifunctional PRA isomerase-IGP synthase from *E. coli*, PDB code 1PII (PRA isomerase and IGP synthase domains are coloured yellow and green respectively) (51).

Some multifunctional enzymes or multi-enzyme complexes form molecular tunnels that connect spatially separated catalytic active sites. This sophisticated channelling system directs the metabolite produced in the first active site to the next active site as a reaction substrate. An unambiguous example for demonstrating the advantage of a tunnel structure formed within a bifunctional enzyme complex is tryptophan synthase, found also in the shikimate pathway. Tryptophan synthase catalyses the final step of tryptophan synthesis (52). Tryptophan synthase is a heterotetramer comprising two α units and two β units. The last step for the synthesis of tryptophan includes two consecutive reactions – the cleavage of indole 3-glycerolphosphate to indole and glyceraldehyde 3-phosphate (G3P), catalysed by the α units, and the synthesis of tryptophan from indole and L-serine, catalysed by the β units (Figure 1.9 A) (59). The crystal structure of tryptophan synthase from *Salmonella typhimurium* consists of two

asymmetric units of an α/β heterodimer where a tunnel is formed to channel metabolites between the active sites of α and β subunits (Figure 1.9 B) (60). As a consequence, the indole produced by the α subunit is quickly and directly transferred to the active site of the β subunit for the final stage of tryptophan synthesis, which has been demonstrated by kinetic experiments (61,62).

Figure 1. 9: The structure of the channelling system embedded within the tryptophan synthase from *S. typhimurium* (52). (A) Hereotetrameric tryptophan synthase from *S. typhimurium* (PDB code 1TTQ) comprising two α subunits (Pink) and a homodimer of β subunits (Blue). **(B)** The tunnel that channels metabolites between the active sites of the α and β subunits are coloured in grey.

1.4.2 Allosteric regulation

As stated above, apart from enhancing catalytic efficiency, multifunctional enzymes and complexes also assist with the provision of allosteric regulation (8,20,30,54).

1.4.2.1 A general theory of allosteric regulation

Allosteric regulation, or allostery, is generally defined as the regulation of protein function induced by a ligand binding at a site which is distant from the active site (63,64). Allostery was initially described as a special phenomenon in which the binding of a ligand to one subunit of an oligomeric protein significantly promotes ligand binding in subsequent subunits (Figure 1.10 A). The classical examples illustrating this principle of allostery are aspartate transcarbamoylase (ATCase) and haemoglobin (65-69). ATCase is composed of two trimers of catalytic subunits and three dimers of regulatory subunits (70), while haemoglobin assembles as a heterotetramer consisting of two α and two β subunits (71). Both proteins exhibit cooperative effects upon binding of a ligand, mediated by the alteration of the interactions between monomeric units. This initial principle limits the concept of allostery to multimeric proteins. Owing to the development of new techniques and extensive studies of the regulation of enzyme function, allostery has now also been identified in monomeric proteins (Figure 1.10 B) (72-77).

Figure 1. 10: Different modes explaining allosteric regulation. (A) Cooperative model as illustrated by a cartoon representation of the Monod-Wyman-Changeux (MWC) model of allosteric transitions. Each unit consists of a symmetric multimeric protein that exists in an active or an inactive state and possesses an active site and a remote binding site for an allosteric effector. **(B)** The allosteric inhibition (left) and activation (right) of a monomeric protein. The binding site of the allosteric effector is distant from the active site. **(C)** The acquisition of allosteric inhibition (left) and activation (right) via protein fusion. The regulatory domain, fused to the catalytic domain, behaves like an allosteric switch, accommodating allosteric effectors and delivering allosteric control of enzymatic activity.

Proteins, even well folded, single-domain, globular proteins, are flexible. They can essentially be considered as dynamic objects due to their own intrinsic structural fluctuations (78-80) and those of water molecules surrounding the protein surface (81). As a result, rather than adopting a single conformation, a protein populates an ensemble of conformers that are able to interconvert on various time-scales to form a conformational equilibrium, although some conformational variations may be very

subtle (Figure 1.11 A) (77,82-86). The energy states of different protein conformers constitute the energy landscape of a protein and a specific enzymatic function is determined by the dominant population of a conformer that is the most energetically favourable; albeit other conformations may coexist (Figure 1.11 B) (87-89). Therefore, conformational equilibrium and protein function are fundamentally controlled by thermodynamics.

Figure 1. 11: The cartoons coloured purple and blue corresponds to active enzymes with different conformations, whereas the orange one represents the inactive enzyme bound to an allosteric inhibitor (red pentagon). (A) The schematic of the conformational preorganization and selection model to illustrate that allosteric inhibitors interfere with the conformational equilibrium and induce a conformational redistribution. **(B)** A further illustration of the conformational preorganization and selection model based on the variation of the energy landscape of a conformational ensemble. The beads in the energy basins represent the relative population of each conformation and the red arrows indicate population shifts. The energy surface prior to allosteric ligand binding is highlighted by a dashed line.

This concept of describing of protein population ensembles has been utilised to investigate and explain enzymatic properties. For a typical enzyme, a series of redistributions of conformational populations can occur along a catalytic cycle. Each catalytic step – such as substrate binding, transition state stabilisation and product release – requires a specific conformational state which facilitates the reaction to progress to the next step (77,89-91). The onward movement of a catalytic reaction, in turn, alters the thermodynamic state and influences the distribution of conformational populations in preparation for the next catalytic step, until the catalytic cycle is completed (77,89-92). With the concept of thermodynamics, investigations on allostery have revealed that allosteric regulation of an enzyme shares the same energetic basis as underlying catalysis; namely, allostery is driven by systematic thermodynamic properties and protein conformational equilibria (63,77). In allostery, the addition of an allosteric effector reshapes the entire energy landscape, which shifts the conformational equilibria towards the more energetically favourable conformation of the enzyme-allosteric effector complex (Figure 1.11 B). By means of redistribution of pre-existing conformational populations, and alteration of the rates of their interconversion, an allosteric effector can alter the local properties of the active site or other binding sites, in turn, to affect the function of a protein (63,83,93-95).

1.4.2.2 Non-covalent enzyme complexes and bifunctional enzymes confer allostery on shikimate pathway

Multidomain structural organisation is often found for allosteric enzymes. So far, the available reports have presented various examples showing how enzymes acquire allosteric functions by covalently linked with other proteins (Figure 1.10 C). In these cases, the extra protein portions work as switches – providing the binding sites for allosteric effectors, transmitting the allosteric signal, and interacting with the catalytic domain – to exert allosteric control on enzymatic activity (8,17,20,96,97). This regulatory strategy has a spatial separation between the catalytic and allosteric

functionalities of one enzyme molecule. As mentioned above, some regulated type I β DAH7PSs contain an ACT domain, which are well-characterised examples illustrating the delivery of allosteric functionality by multi-domain proteins (96-98). The ACT domain works solely as a ligand-binding motif to receive and transmit the inhibitory signal of Tyr to the DAH7PS catalytic site. Hence, it is generally considered as a regulatory element required for delivering an allosteric response in these type I β DAH7PS enzymes.

Notably, CM, a pivotal enzyme found at the branch point of the shikimate pathway, is also found fused to some type I β DAH7PS and in these cases the CM domain delivers allosteric regulation to DAH7PS activity. The first reported example of a fusion enzyme between CM and DAH7PS was the bifunctional enzyme from *B. subtilis* (*BsuDAH7PS*) (27). In this protein, a CM domain is fused to the N-terminus of the DAH7PS core catalytic barrel, and this protein is denoted as an N-terminal CM linked DAH7PS. At first, this enzyme was considered as a simple bifunctional enzyme, which possesses both CM and DAH7PS functions. But later, the CM domain was found to be able to bind prephenate molecules and prephenate binding was subsequently shown to inhibit the activity of the appended DAH7PS domain (8). In addition to the CM activity embodied in the bifunctional DAH7PS, an unregulated monofunctional AroH type of CM was also found in *B. subtilis* and this enzyme displayed efficient catalytic activity (42,99,100).

Figure 1. 12: Crystal structures and homology model of *LmoDAH7PS* and *GspDAH7PS*. (A) Crystal structure of ligand-free *LmoDAH7PS* (PDB code 3NVT) (28). (B) Binding of prephenate induces a large conformational change of *GspDAH7PS* resulting in inhibition of DAH7PS activity. *GspDAH7PS* (ligand-free) is a homology model generated using the crystal structure of *LmoDAH7PS* as a template while the structure showing prephenate (yellow spheres) bound *GspDAH7PS* is a crystal structure (PDB code 5J6F) (20).

The first crystal structure of an N-terminal CM linked DAH7PS was *LmoDAH7PS* (28). *LmoDAH7PS* possesses a CM domain fused to the N-terminus of an uninterrupted TIM barrel and shares a high sequence identity (51%) to *PfuDAH7PS*. The crystal structure of *LmoDAH7PS* shows that the DAH7PS domains assemble as a tetramer. Each of the

four CM domains connects to the N-terminus of each DAH7PS domain via a ~15 amino acid linker and form a pair of dimers at either side of the core DAHPS tetramer (Figure 1.12 A). However, due to the lack of a crystal structure of the prephenate bound enzyme, the investigations on *Lmo*DAH7PS were unable to characterise the allosteric mechanism. The mechanism underlying the CM mediated allostery was unclear until the study on *Gsp*DAH7PS was reported (20). *Gsp*DAH7PS is another N-terminal CM linked DAH7PS, which has high similarity to *Lmo*DAH7PS in terms of both sequence and structure. According to the crystal structure of prephenate bound *Gsp*DAH7PS, the catalytic core of this enzyme is also constructed in a homotetrameric arrangement. The CM domains of diagonally opposed protomers form a dimer at either side of the DAH7PS tetrameric core and are positioned above the entrances to the DAH7PS active sites. Furthermore, the homology model of apo-*Gsp*DAH7PS was generated based on the crystal structure of apo-*Lmo*DAH7PS, and compared with the prephenate-bound structure. The result of the comparison reveals that the conformation of the CM domains shifts towards a tighter association with the tetrameric DAH7PS core when prephenate is present and occludes the entrances of the two active sites and thereby inhibits the activity of DAH7PS (Figure 1.12 B) (20).

Interestingly, two complexes that are formed by non-covalent interactions between CM and type II DAH7PS (CM-DAH7PS complex) have been reported for *M. tuberculosis* (Figure 1.13 A) and *Corynebacterium glutamicum* (29,30) (Figure 1.13 B). Similar to the CM and DAH7PS covalent fusion enzyme, this CM-DAH7PS non-covalent complex is formed by a tetrameric DAH7PS core and a pair of CM dimers located at opposite ends of the tetrameric DAH7PS core. In contrast to the delivery of allostery to DAH7PS by CM exhibited by the N-terminal CM linked DAH7PSs, CM is activated via the non-covalent interaction with the DAH7PS in these CM-DAH7PS complexes. In addition, the catalytic activity of the CM becomes allosterically regulated by the binding of Phe and Tyr to binding sites located within the DAH7PS subunits (24,29,30,54).

Figure 1. 13: Non-covalent interaction between DAH7PS and CM. The core DAH7PS catalytic tetramers are coloured blue, CM dimers are shown in red. (A) Non-covalent complex between *Mtu*DAH7PS and the intracellular AroQ δ *Mtu*CM (PDB code 2W19) (29). **(B)** Non-covalent complex formed between DAH7PS and CM from *C. glutamicum* (PDB code 5HUD) (30).

1.5 C-terminal CM linked DAH7PS

Besides the DAH7PS with an N-terminal CM, a functioning CM domain has also been found fused at the C-terminus of some other type I β DAH7PS enzymes (8). So far, the information about this kind of DAH7PS is sparse, with only a single report on the C-terminal linked DAH7PS from *Porphyromonas gingivalis* (*Pgi*DAH7PS). Wu and Woodard reported that this DAH7PS, with a C-terminal CM domain, was also allosterically regulated by prephenate and chorismate. In addition, this enzyme was found to be much more sensitive to these inhibitors compared to its N-terminal CM linked counterparts. The catalytic activities of the separated individual DAH7PS and CM domains (*Pgi*DAH7PS^D and *Pgi*DAH7PS^{CM}) were both shown to decrease dramatically, relative to those found for the wild-type enzyme. To the contrary, the two separated domains of *Bsu*DAH7PS (an N-terminal CM linked DAH7PS) were reported to be a little more active than their covalently linked form(8).

Based on the fact that a monofunctional AroH CM and the bifunctional CM-DAH7PS

enzyme co-exist in *B. subtilis*, it was hypothesised that the fusion of DAH7PS with CM took place in order to acquire allosteric regulation for the DAH7PS to respond to the high level prephenate turned up by the monofunctional AroH CM. In this proposition, a CM domain appended to the N-terminus of a DAH7PS mainly functions as an allosteric regulatory module rather than a CM.

However, for the majority of microorganisms that contain a C-terminal CM linked DAH7PS, only a single gene encoding for a CM is found. Therefore, it appears likely that the CM domain of the C-terminal CM linked DAH7PS may take both the responsibility of producing prephenate and providing allosteric regulation for DAH7PS activity. In addition, *Pgi*DAH7PS was found to be extremely sensitive to inhibition by prephenate and chorismate. It was reported that 1.5 μ M of prephenate or 2.5 μ M of chorismate inhibits *Pgi*DAH7PS activity by 50%, whereas 100 μ M of prephenate or 1 mM of chorismate was required to achieve the same inhibition levels for *Bsu*DAH7PS (8). Such a high sensitivity to chorismate and prephenate may lead to relatively low concentrations of prephenate and pathway end products (e.g. Tyr and Phe) in *P. gingivalis*. What are the structural and functional features of C-terminal CM linked DAH7PS? What is the mechanism underlying allosteric regulation for these C-terminally modified DAH7PS enzymes? Moreover, what does this bifunctional organisation contribute to the metabolism of aromatic amino acids? There are many questions that remain to be answered regarding the true role of this bifunctional DAH7PS subgroup.

1.6 Objectives

The primary aim of this thesis is to characterise an enzyme belonging to the C-terminal CM linked DAH7PS subgroup and hence unravel the mechanisms underlying the catalysis and allostery of this bifunctional DAH7PS subtype. A comprehensive investigation of this obscure DAH7PS subtype presents a distinctive example for

illustrating how protein dynamics serves for catalysis to allostery. Such information will also contribute to the better understanding of the evolutionary significance of the bifunctional enzymes in metabolic pathways.

In this project, the DAH7PS from *Prevotella nigrescens* (*Pni*DAH7PS) was selected to study the nature of a C-terminal CM linked DAH7PS. *P. nigrescens* is a member of a group of anaerobic, rod-shaped bacteria, which can be isolated from the human oral and vaginal flora (101,102). The presence of this organism is associated with the increased incidence of many diseases such as periodontal disease, gingivitis and odontogenic infections (103-105). Recently it was reported that this organism is also significantly relevant to carotid atherosclerosis (106). According to sequence analysis, *Pni*DAH7PS shares a high overall sequence similarity with *Pgi*DAH7PS, and also the *Pni*DAH7PS domain shows close homology to that of other type I β DAH7PS enzymes. Until now, there has not been any information available on the structure or function of *Pni*DAH7PS.

The specific objectives of each chapter of this thesis are as follows.

Chapter 2 describes the general biochemical and structural characteristics of *Pni*DAH7PS; including kinetic parameters, allosteric inhibition, quaternary structure and a low-resolution structural information acquired using small angle X-ray scattering (SAXS). This chapter also presents studies that identify distinctive features of this enzyme for subsequent investigations.

Chapter 3 focuses on how the DAH7PS and CM domains of *Pni*DAH7PS communicate with each other to ensure efficient enzymatic activities during catalysis. In this chapter, the heterodomain cross-talk within the wild type *Pni*DAH7PS is investigated using both biochemical and biophysical methods, including a series of kinetic, isothermal titration calorimetric (ITC), and SAXS assays. A truncated variant of *Pni*DAH7PS and the effect

of NaCl on *Pni*DAH7PS are also examined for characterising the relationship between the functional and structural heterodomain interplays during the catalysis of *Pni*DAH7PS.

Chapter 4 probes the mechanism of allosteric inhibition of *Pni*DAH7PS by prephenate and unravels the inhibitory effect of prephenate on the DAH7PS catalysed reaction. The mechanisms underpinning the allosteric regulation of N-terminal and C-terminal CM linked DAH7PS enzymes, such as *Gsp*DAH7PS and *Pni*DAH7PS, are compared in this chapter.

The dynamic features of *Pni*DAH7PS impede the crystallisation of the full-length *Pni*DAH7PS, which poses significant challenges for the detailed inspection of the hetero-domain interactions and interfaces. To overcome the difficulty of crystallisation, and to obtain reliable data depicting *Pni*DAH7PS-CM domain interactions, a method of producing segmented isotopically labelled *Pni*DAH7PS, using expressed protein ligation (EPL), was explored and is described in **Chapter 5**. The segmental labelling of *Pni*DAH7PS was subsequently coupled with small angle neutron scattering (SANS) or nuclear magnetic resonance spectroscopy (NMR) for structural examinations, which are presented in **Chapter 5**.

Chapter 2 *Preliminary characterisation of PniDAH7PS*

2.1 Introduction

As introduced in Section 1.2.1, the subgroup comprising the type I β enzymes contains the largest variation in the extra-barrel decorations. The simplest type I β DAH7PS proteins comprise the core barrel alone and thus only support an enzymatic function that is not subject to feedback regulation by aromatic amino acids (18,23). More complex type I β DAH7PS proteins also contain the core catalytic barrel but additional discrete domains are appended at the N- or C-terminals (8,107). These N- or C-terminal extensions are often important for providing the allosteric machinery.

So far, the majority of studies on the regulated type I β DAH7PS have focused on the enzymes with an N-terminal extension, which provides either an ACT-like N-terminal domain or N-terminal CM domain (8,20,108), such as *Tma*DAH7PS and *Gsp*DAH7PS. The allosteric mechanisms whereby N-terminal ACT or CM domains control the catalytic function of the DAH7PS domain are relatively well understood (20,108,109). For *Tma*DAH7PS, the ACT domains from two neighbouring chains dimerise upon the binding of Tyr, blocking the entrance to the DAH7PS active site, thus, reducing the rate of DAH7PS catalysis (108,109). Similarly, prephenate binding to the CM domain in *Gsp*DAH7PS also results in the occlusion of the DAH7PS active site (20). Both of these two proteins, *Tma*DAH7PS and *Gsp*DAH7PS, share a common homotetrameric assembly and this quaternary structure is essential for the allosteric functionality of both enzymes. As the unregulated type I β forms of DAH7PS also share this homotetrameric assembly, it has been proposed that an unregulated type I β DAH7PS tetrameric scaffold may have been an ancestral species that acquired various allosteric machinery (8). This proposal is supported by domain fusion and domain swapping experiments, which reveal how readily allostery can be acquired, or interchanged, by type I β DAH7PS domains (96,97).

There is, however, one distinct group of type I β DAH7PS proteins for which very limited

information is available. Unlike other type I β DAH7PS proteins, this group possesses a C-terminal CM domain. Functional studies on this group of type I β DAH7PSs are limited to *Pgi*DAH7PS, which shows dual DAH7PS and CM activity and allosteric inhibition by prephenate (8). However, unlike the N-terminally linked proteins, on splitting the full-length *Pgi*DAH7PS protein into its individual DAH7PS and CM components, both enzymatic activities were lost (8,20,109). This characteristic may indicate that a unique architectural organisation of the DAH7PS and CM domains exists for these C-terminally modified proteins that is essential for catalysis and allosteric regulation. However, to date, no structural information has been available for these DAH7PS proteins and the mechanisms of catalysis and allostery are unknown.

To probe the properties of the C-terminal CM-fused DAH7PS group, and to elucidate more fully the effects arising from the fusion of these enzymes, we chose to examine the biochemical and structural characteristics of the DAH7PS from *Prevotella nigrescens* (*Pni*DAH7PS), which contains a core catalytic DAH7PS barrel and a covalently linked C-terminal CM domain.

2.2 Sequence alignment and CLANS reveals the distinction of the C-terminal CM linked DAH7PS subgroup in type 1 DAH7PS

To illustrate the diversity of type I DAH7PSs based on sequence and to explore more fully the sequence relationships, type I DAH7PS sequences were sorted and analysed using CLuster ANalysis of Sequences (CLANS) (110). In this way, as expected based on previous sequence analysis (10,11), two distinct DAH7PS clusters were identified: The type I α subgroup and the type I β subgroup. The type I β subgroup also included the closely related enzyme 3-deoxy-D-*manno*-2-octulosonate 8-phosphate synthase (KDO8PS), which catalyses a closely related reaction, as a part of lipopolysaccharide biosynthesis in Gram-negative bacteria (9) (Figure 2.1A). By using this analysis, the type I β subgroup was shown to comprise of one major cluster with an adjacent smaller

grouping exclusively containing C-terminal CM linked CM DAH7PSs (Figure 2.1B). All well-characterised type I β DAH7PS proteins to date reside in this main cluster, including *Gsp*DAH7PS, *Tma*DAH7PS and *Pfu*DAH7PS (111).

*Pni*DAH7PS was selected as a representative of the smaller grouping for a more detailed investigation. Sequence alignments of *Pni*DAH7PS with the known *Gsp*DAH7PS, *Tma*DAH7PS, *Pfu*DAH7PS and *Pfu*CM (from *P. furiosus*) indicated that *Pni*DAH7PS comprises of an N-terminal DAH7PS domain and a C-terminal CM domain, a feature which it shares with other proteins clustered into this smaller grouping (Figure 2.1). Although the alignments do not show great sequence similarity between either the DAH7PS or CM portion of *Pni*DAH7PS and the other aligned sequences (Table 2.1), the key active site residues are conserved (Figure 2.2 B).

Figure 2. 1: CLANS clustering of type I DAH7PS sequences. Each dot represents a protein sequence. **(A)** Two main clusters were identified by CLANS as indicated: The type I α subgroup and the type I β subgroup (also containing the related KDAO8PS proteins). **(B)** The CLANS analysis of the type I β DAH7PS sequences reveals three distinct clusters, the type I β C-terminal CM-fused subgroup (red) including *Pni*DAH7PS, the main type I β subgroup (green) including N-terminal CM linked DAH7PSs, such as *Gsp*DAH7PS, *Tma*DAH7PS, *Pfu*DAH7PS, and the KDO8PS group (blue).

Figure 2. 2: Sequence alignment of *Pni*DAH7PS with other type I β DAH7PS sequences, demonstrating the *Pni*DAH7PS sequence responsible for DAH7PS and CM activities. (A) Schematic diagrams showing the domain architectures of *Pni*DAH7PS, *Gsp*DAH7PS, *Tma*DAH7PS, *Pfu*DAH7PS and *Pfu*CM. (B) Sequence alignments of *Pni*DAH7PS and the well-studied type I β DAH7PSs (upper) and CMs (lower) respectively. The residues highlighted by the red star indicate conserved residues that form part of the active sites.

Table 2. 1: Table shows the sequence identities between *Pni*DAH7PS and other homologous DAH7PS and CM enzymes

Pairwise alignment between <i>Pni</i> DAH7PS and homologous DAH7PS sequences		
DAH7PS Sequence	UniportKB Entry No.	Pairwise Identity (%)
<i>Pfu</i> DAH7PS	Q8U0A9	35.0
<i>Gsp</i> DAH7PS	L8A208	30.5
<i>Tma</i> DAH7PS	Q9WYH8	34.1
Pairwise alignment between <i>Pni</i> DAH7PS and homologous CM sequences		
CM Sequence	UniportKB Entry No.	Pairwise Identity (%)
<i>Pfu</i> CM	Q8U098	27.6
<i>Gsp</i> DAH7PS	L8A208	27.2

2.3 Cloning, expression and purification of *Pni*DAH7PS

The gene of *Pni*DAH7PS, named *pheB*, is from the *P. nigrescens* ATCC 33563 strain. Due to codon usage bias, in order to achieve a more efficient and accurate translation, the *Pni*DAH7PS gene sequence was optimised according to the codon usage frequency of the expression host organism, *E. coli* BL21 (DE3), before it was synthesised by Life Technologies (Thermo Fisher Scientific Inc.). The synthesised gene was cloned into the pET28a vector to facilitate the expression of the recombinant protein containing an N-terminal 6 × His fusion tag (His-tag) for ease of detection and purification. The sequence for a Tobacco Etch Virus (TEV) protease cleavage site was inserted between the His-tag and the *Pni*DAH7PS gene to enable the removal of the tag using TEV protease after protein purification. The *Pni*DAH7PS gene insert of the prepared pET28a vector was verified by DNA sequencing before the expression construct was transformed into BL21 (DE3) chaperone combination 3 (CC3) cells following standard procedures described in section 7.3.7.

Figure 2. 3: SDS-PAGE analysis shows the expression and purification of *PniDAH7PS*. (A) The expression of His-*PniDAH7PS* in BL21 (ED3) CC3 cells. Lane 1: Protein marker. Lane 2: Whole cell lysate. Lane 3: Supernatant after centrifugation, the soluble fraction of the cell lysate. Lane 4: Cell debris, the insoluble fraction of the cell lysate. **(B)** TEV cleavage of His-*PniDAH7PS*^{WT}. lane 1 and 5: Protein marker. Lane 2 to 4: His-*PniDAH7PS*^{WT} before TEV cleavage. Lane 6 to 8: *PniDAH7PS*^{WT} without a His-tag after TEV cleavage. **(C)** The fractions of the elution peak collected from the SEC.

Protein expression was induced by adding isopropyl β -D-thiogalactopyranoside (IPTG) when the cells reached the mid-logarithmic growth phase ($OD_{600} = 0.4 - 0.8$). Cells were harvested after overnight incubation and lysed, on ice, by sonication before cellular debris was removed by centrifugation. *PniDAH7PS* was purified from the soluble fraction of the cell lysate (Figure 2.3 A). As *PniDAH7PS* was expressed as a

recombinant protein with an N-terminal His-tag, the cell lysate was applied to a HiTrap TALON® crude column (Talon column) (GE Healthcare) for purification. The protein was eluted from the column using an increasing imidazole concentration gradient. The fractions of interest were then pooled and desalted to remove the added imidazole. TEV protease was added into the desalted eluate, for an overnight incubation, to remove the His-tag (Figure 2.2 B). After His-tag cleavage, the eluate was reapplied to the Talon column to separate the cleaved His-tag or any uncleaved His-tagged protein from the untagged *PniDAH7PS* of interest. The final step of the purification process was size-exclusion chromatography (SEC) using a Hiload 26/60 Superdex 200 pg column (GE Healthcare). The protein eluted as a single peak and sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) indicated that the protein was pure (Figure 2.2 C).

The molecular weight of the purified wild-type *PniDAH7PS* (*PniDAH7PS*^{WT}), including four amino acid residues, located at the N-terminus, that were remnants of the TEV cleavage site (GSGA), was found to be 40,365.4 Da, as determined by electrospray ionisation mass spectrometry (ESI-MS). This is very close to the expected value of 40,364.7 Da, calculated from the amino acid sequence using ProtParam (112).

2.4 Basic characterisation of the functions and structure of *PniDAH7PS*^{WT}

2.4.1 Metal ion activation

A divalent metal ion is required for the catalysis of all other known DAH7PS enzymes (8,10,15,17,18,20,29-31,34,113,114). To investigate the metal ion dependency of *PniDAH7PS*, the DAH7PS activity was monitored in the presence of various divalent metal cations. To avoid the influence of other metal ions in this assay, buffer and substrate solutions were treated with Chelex® (an ion-exchange resin that removes polyvalent metal ions) prior to use.

Figure 2. 4: The extent of activation of *PniDAH7PS*^{WT} by different metal ions in comparison to Mn²⁺ at a concentration of 100 μ M. The inactivation of *PniDAH7PS*^{WT} in the presence of 100 μ M EDTA is also shown as a negative control. Reactions were carried out at 35 °C, containing 100 μ M metal ion, 150 μ M PEP, 225 μ M E4P and 3 μ g *PniDAH7PS*^{WT} in 50 mM BTP buffer (pH 7.4). Error bars represent the standard deviation from triplicate measurements.

It was found that Mn²⁺ and Co²⁺ activated the apo-enzyme significantly, with Mn²⁺ resulting in the fastest reaction rates (Figure 2.4).

2.4.2 Kinetic characteristics

The steady-state kinetic characteristics for DAH7PS activity of *PniDAH7PS*^{WT} were determined by measuring the consumption of PEP by monitoring absorbance at 232 nm (Appendix Figure A.2.1 A, B). All assays were initiated by the addition of E4P, rather than enzyme. Initial velocity rates for the DAH7PS catalytic activity of *PniDAH7PS*^{WT} were measured under the particular conditions as described in section 7.14, where one substrate concentration was fixed while the other substrate concentration was varied. The apparent K_M values for PEP and E4P were found to be 41 ± 4 μ M and 58 ± 6 μ M respectively and the k_{cat} value was calculated as 16.8 ± 0.5 s⁻¹ (Table 2.2).

Table 2. 2: Table shows the kinetic parameters of the DAH7PS activity of *PniDAH7PS*^{WT} compared to other type Iβ DAH7PS enzymes

Enzyme	K_M^{PEP} (μM)	K_M^{E4P} (μM)	k_{cat} (s ⁻¹)	$k_{\text{cat}}/K_M^{\text{PEP}}$ (μM ⁻¹ ·s ⁻¹)	$k_{\text{cat}}/K_M^{\text{E4P}}$ (μM ⁻¹ ·s ⁻¹)
<i>PniDAH7PS</i> ^{WT}	41 ± 4	58 ± 6	16.8 ± 0.5	0.41	0.29
<i>GspDAH7PS</i> ⁽¹¹⁵⁾	87 ± 9	95 ± 8	45 ± 4	0.52	0.47
<i>PgiDAH7PS</i> ⁽⁸⁾	421 ± 43	1238 ± 141	1.6 ± 0.1	3.8 × 10 ⁻³	1.3 × 10 ⁻³
<i>BsuDAH7PS</i> ⁽⁸⁾	139 ± 11.4	1760 ± 110	4.6 ± 0.1	3.3 × 10 ⁻²	2.6 × 10 ⁻³
<i>TmaDAH7PS</i> ⁽¹⁰⁸⁾	4.9 ± 0.4	13 ± 1	13 ± 1	2.4	0.9
<i>PfuDAH7PS</i> ⁽¹⁶⁾	120 ± 20	28 ± 4	1.5 ± 0.1	1.3 × 10 ⁻²	5.4 × 10 ⁻²

The apparent Michaelis constants for the DAH7PS activity are apparent values determined using a fixed concentration of the invariant substrate (as described in section 7.1.13). The uncertainty values for *PniDAH7PS*^{WT} represent the standard deviation from triplicate measurements.

The steady-state kinetic parameters for the CM domain of *PniDAH7PS*^{WT} were determined by measuring the consumption of chorismate at 274 nm at 35 °C (Appendix Figure A.2.1 C). All assays were initiated by adding chorismate. The K_M value for chorismate was determined to be 7.6 ± 0.6 μM and the k_{cat} value was calculated as 1.68 ± 0.06 s⁻¹ (Table 2.3).

Table 2. 3: Table shows the kinetic parameters of the CM activity of *PniDAH7PS*^{WT} compared to that of *GspDAH7PS*, *PgiDAH7PS* and *PfuCM*

Enzyme	K_M (μM)	k_{cat} (s ⁻¹)	k_{cat}/K_M (μM ⁻¹ ·s ⁻¹)
<i>PniDAH7PS</i> ^{WT}	7.6 ± 0.6	1.68 ± 0.06	0.22
<i>GspDAH7PS</i> ⁽¹¹⁵⁾	88 ± 6	1.9 ± 0.1	2.1 × 10 ⁻²
<i>PgiDAH7PS</i> ⁽⁸⁾	337 ± 15	0.45 ± 0.04	1.3 × 10 ⁻³
<i>BsuDAH7PS</i> ⁽⁸⁾	850.5 ± 97	0.41 ± 0.01	4.8 × 10 ⁻⁴
<i>PfuCM</i> ⁽¹¹⁶⁾	25 ± 2	6.6 ± 0.7	0.26
<i>EcoCM</i> ⁽¹¹⁷⁾	30 ± 1	41.4 ± 0.47	1.38

The uncertainty values for *PniDAH7PS*^{WT} represent the standard deviation from triplicate measurements.

2.4.3 Feedback inhibition

As previously reported, the DAH7PS activity of *Pgi*DAH7PS is very sensitive to the presence of either chorismate or prephenate with 50 % activity remaining in the presence of 1.5 μ M chorismate or 2.5 μ M prephenate (8). To determine if a similar sensitivity to prephenate also exists for the *Pni*DAH7PS^{WT} catalytic reaction, the extent of inhibition of *Pni*DAH7PS by prephenate was investigated using standard assay conditions. The DAH7PS catalytic activity was tested in the presence of varying concentrations of prephenate (0 – 312 μ M). It was found that *Pni*DAH7PS^{WT} is also very sensitive to prephenate and the relative half-maximal inhibition concentration (IC₅₀) of prephenate was calculated to be 4.7 μ M using a Logistic Fit (Origin Pro 8) (Figure 2.5).

Figure 2. 5: The response of *Pni*DAH7PS^{WT} to varying concentrations of prephenate.

Reactions were carried out at 35 °C in 50 mM BTP buffer (pH 7.4) containing 3 μ g of enzyme, 100 μ M Mn²⁺, 150 μ M PEP and 225 μ M E4P. Error bars represent the standard deviation from triplicate measurements.

To confirm that prephenate binds directly to the active site of the CM domain, inhibition by prephenate on CM activity of *Pni*DAH7PS was also assessed. Four groups of kinetic assays for the CM catalysed reaction were conducted in the presence of prephenate at different concentrations. The obtained kinetic data was analysed using Lineweaver–Burk plot (Figure 2.6 B), which fits a typical competitive inhibition of CM

activity with a goodness fit R^2 of 0.9915, calculating a K_i value of 1.2 μM (Figure 2.6 A).

Figure 2. 6: Inhibition of the CM activity of *PniDAH7PS*^{WT} by prephenate. (A) The Michaelis-Menten plots of the data characterising the inhibition of CM activity by prephenate. **(B)** The Lineweaver–Burk analysis for the prephenate mediated inhibition on CM.

2.4.4 Characterisation of the quaternary structure of *PniDAH7PS*^{WT}

To determine the quaternary structure of *PniDAH7PS*^{WT} in solution, analytical size-exclusion chromatography (analytical SEC) and analytical ultracentrifugation (AUC) experiments were carried out.

Analytical SEC

Analytical SEC separates proteins in solution based on their molecular mass. Proteins with a higher molecular weight are eluted earlier and elution volumes (V_e) and the void volume (V_o) can be transformed to a coefficient K_{av} , which decreases linearly with the logarithm of the molecular mass ($\log\text{MW}$). Thus, an analytical SEC column was calibrated with blue dextran (2000 kDa) and standard samples of known molecular weight (aldolase 158 kDa, conalbumin 75 kDa, ovalbumin 43 kDa and cytochrome 12.4 kDa) to determine the V_o and V_e values and to construct a calibration curve that relates K_{av} to $\log\text{MW}$ (Figure 2.7). The V_e of *PniDAH7PS*^{WT} at three concentrations (1.5 mg/mL, 0.75 mg/mL, and 0.3 mg/mL) was found to be approximately the same. The

corresponding molecular mass was calculated to be around 90 kDa, according to the calibration curve, indicating that the quaternary assembly of *Pni*DAH7PS in solution is dimeric (Figure 2.7).

Figure 2. 7: Analytical SEC elution trace and standard curve of logMW against K_{av} for *Pni*DAH7PS^{WT} at 1.5 mg/mL (blue line), 0.75 mg/mL (pink line) and 0.3 mg/mL (yellow line). The SEC buffer contained 10 mM BTP, pH 7.4, 150 mM NaCl and 200 μ M PEP.

AUC

AUC was utilised to further determine the quaternary structure of *Pni*DAH7PS^{WT} in solution. By applying a sufficiently high angular velocity, rapid sedimentation of particles occurs and boundaries are formed at the trailing edges of each particulate species. The velocity of the individual particles in sedimentation velocity experiments cannot be resolved, but the rate of movement of the boundary region can be measured. From this, the sedimentation coefficient (s) can be determined, which can be used to identify the quaternary structure of a protein. Following buffer exchange into freshly made AUC buffer, the *Pni*DAH7PS^{WT} solution was made up to three different concentrations: 1.5, 1.0 and 0.5 mg/mL. According to the absorbance versus radial position data, two sedimentation boundaries were present, which suggests that there are two oligomeric species of *Pni*DAH7PS^{WT} in solution. Fitting the experimental data to a continuous model showed a major species with a sedimentation coefficient

of 4.9 S and a minor species with a sedimentation coefficient of 7.3 S (Figure 2.8). The molecular mass of each species was determined as ~71 kDa and ~131 kDa respectively, based on the sedimentation coefficients. The species with 71 kDa, which is consistent with the expected molecular masses for the dimeric (80 kDa) *PniDAH7PS*^{WT}, accounts for the vast majority, and no concentration-dependent equilibrium between the two suspected species was observed. Thus, the species with a sedimentation coefficient of 7.3 S, and corresponds in calculated mass most closely to a trimeric species (120 kDa) was suspected to be an aggregation of *PniDAH7PS*^{WT}. To investigate if the prephenate presence influenced the distribution of the two observed oligomeric states of *PniDAH7PS*, AUC assays were repeated using 1.0 mg/mL protein in presence of 200 μ M prephenate. However, neither increasing protein concentration (to 1.5 mg/mL) nor the prephenate presence gave rise to significant changes in the distribution of species of *PniDAH7PS*^{WT} (Figure 2.8).

Figure 2. 8: Sedimentation velocity analysis of the quaternary structure of *PniDAH7PS*^{WT}. The AUC buffer contained 10 mM BTP, pH 7.4 and 150 mM NaCl. Data were collected at 50,000 rpm, at a wavelength of 280 nm. The species distribution curve represents *PniDAH7PS*^{WT} at concentrations of 1.5 mg/mL (blue line), 1.0 mg/mL (red line), 0.5 mg/mL (black line) and 1.1 mg/mL *PniDAH7PS*^{WT} supplemented with 200 μ M prephenate (green dashed line).

2.5 Cloning, expression and purification of truncated variants

As mentioned in the introduction, studies of *Pgi*DAH7PS found that both the DAH7PS and CM activities of *Pgi*DAH7PS dropped significantly after the wild-type protein was split to individual DAH7PS (*Pgi*DAH7PS^D) and CM (*Pgi*DAH7PS^{CM}) (8). This observation is in contrast to the results of previous studies on the N-terminal CM linked DAH7PS from *B. subtilis* and *Geobacillus. sp* where CM and DAH7PS activities remained similar, compared to the respective activities of the wild-type enzyme, when the DHA7PS or CM domains were expressed individually (20). This huge change in enzymatic activity for the individual DAH7PS or CM domains of *Pgi*DAH7PS, relative to that of the wild-type enzyme, may indicate that another function is provided by the combined protein complex rather than just the introduction of allosteric regulation. To clarify the significance of the protein fusion in the C-terminal CM linked DAH7PSs, truncated DAH7PS and CM domains of *Pni*DAH7PS (*Pni*DAH7PS^D and *Pni*DAH7PS^{CM}) were constructed and investigated individually from the aspect of biochemical characteristic and protein structure.

2.5.1 Selection of truncation site

As stated above, sequences of C-terminal CM linked DAH7PS enzymes are considerably different from that of the main group of type I β DAH7PSs. Moreover, there is no structural information available for this DAH7PS subgroup, thus it is difficult to accurately demarcate the regions of the DAH7PS and CM domains or the linker region in the primary structure of the protein. However, by selecting a truncation site within the predicted linker region it is likely that the individual domains were as intact as possible. Through multiple sequence alignment of *Pni*DAH7PS, *Pgi*DAH7PS, *Pfu*DAH7PS, *Lmc*DAH7PS and *Gsp*DAH7PS sequences, the selected truncation site of *Pni*DAH7PS was chosen to be between K261 and H262 (Figure 2.9).

Figure 2. 10: SDS-PAGE analysis shows the expression and purification of *PniDAH7PS^D*. (A) The expression of His-*PniDAH7PS^D* in pLysS cells. Lane 1: Protein marker. Lane 2: Whole cell lysate. Lane 3: The soluble fraction of the cell lysate. Lane 4: Cellular debris, the insoluble fraction of the cell lysate. **(B)** Lane 1: Protein marker. Lane 2: His-*PniDAH7PS^D* before TEV cleavage. Lane 3: *PniDAH7PS^D* without a His-tag after TEV cleavage. Lane 4 to 7: The fractions of the elution peak collected from SEC.

The sequence verified plasmid pDEST17-*PniDAH7PS^D* was transformed into BL21 (DE3) cells. Expression of *PniDAH7PS^D* was achieved by the addition of 0.5 mM IPTG and shaking incubation at 23 °C overnight (Figure 2.10A). The purification procedure for *PniDAH7PS^D* was similar to that of *PniDAH7PS*. After cell lysis by sonication, the soluble fraction was subjected to IMAC, TEV cleavage and a second round of IMAC to remove the majority of impurities, the cleaved His-tag and any remaining His-tagged *PniDAH7PS^D*. These steps were followed by further purification by SEC (Figure 2.10B).

The molecular weight of *PniDAH7PS^D*, including the additional four amino acid residues at the N-terminus from the TEV cleavage site (GSGA), was determined by ESI-MS. *PniDAH7PS^D* was found to have a molecular mass of 29,461.9 Da, which corresponds closely to the expected molecular weight of 29,464.2 Da calculated from the amino acid sequence using ProtParam (112).

2.5.3 Cloning, expression and purification of *PniDAH7PS*^{CM}

The ORF of *PniDAH7PS*^{CM} was amplified from the plasmid containing the full-length *PniDAH7PS* gene. For the purpose of improving the expression and solubility, the DNA of *PniDAH7PS*^{CM} was inserted into a pDEST15 expression vector, which encodes a glutathione S-transferase (GST) tag at the N-terminus of the inserted gene. The GST tagged protein can be purified by GSTrap affinity chromatography using a GSTrap column (GE Healthcare) containing immobilised glutathione which is able to bind to the GST tag of the tagged protein. A TEV protease cleavage site was introduced between the GST tag and *PniDAH7PS*^{CM} for removal of the GST tag. The sequence verified plasmid construct was transformed into BL21 (DE3) pLysS cells. Expression of *PniDAH7PS*^{CM} was achieved by the addition of 0.5 mM IPTG and shaking incubation at 23 °C overnight (Figure 2.11A).

Following cell lysis and centrifugation to remove the cellular debris, the supernatant was applied to a GSTrap column for purification of *PniDAH7PS*^{CM}. The protein was eluted from the column as a result of an increasing glutathione concentration gradient. The fractions of interest were then pooled and desalted to remove the glutathione before the protein was treated with TEV protease in order to cleave the GST tag (Figure 2.11B). After digestion with TEV protease, the eluate was subjected to a second round of GSTrap affinity chromatography to remove the cleaved GST tag and any remaining uncleaved GST tagged protein. However, the interactions between *PniDAH7PS*^{CM} and the cleaved GST tag or uncleaved GST- *PniDAH7PS*^{CM} were found to be very strong. Therefore GSTrap affinity chromatography was not able to remove the GST tag completely (Figure 2.11C). Following SEC, *PniDAH7PS*^{CM} and the cleaved GST tag or uncleaved protein were separated, resulting in pure *PniDAH7PS*^{CM} (Figure 2.11C,D).

Figure 2. 11: SDS-PAGE analysis shows the expression and purification of *PniDAH7PS^{CM}*. (A) The expression of GST-*PniDAH7PS^{CM}* in pLysS cells. Lane 1: Protein marker. Lane 2: Whole cell lysate. Lane 3: The soluble fraction of the cell lysate. Lane 4: Cellular debris, the insoluble fraction of the cell lysate. **(B)** Lane 1: Protein marker. Lane 2: GST-*PniDAH7PS^{CM}* before TEV cleavage. Lane 3: *PniDAH7PS^{CM}* without the GST tag, after TEV cleavage. **(C)** The elution trace of SEC for the purification of *PniDAH7PS^{CM}*, there were three peaks in the elution trace. **(D)** SDS-PAGE analysis shows that the peak 1 fraction from SEC purification was *PniDAH7PS^{CM}* binding to uncleaved GST-*PniDAH7PS^{CM}*, the peak 2 fraction was the complex of *PniDAH7PS^{CM}* and the cleaved GST tag, and the peak 3 fraction was the pure *PniDAH7PS^{CM}*.

The molecular weight of *PniDAH7PS^{CM}*, including the four additional amino acid residues, located at the N-terminus, from the TEV cleavage site (GSGA), was determined by ESI-MS. *PniDAH7PS^{CM}* was found to have a molecular mass of 11,191.1 Da, which corresponds closely to the expected molecular weight of 11,190.8 Da calculated from the amino acid sequence using ProtParam (112).

2.6 Basic characterisation of the function and structure of *Pni*DAH7PS^D

2.6.1 Metal ion activation and kinetic parameters

Figure 2. 12: The extent of activation of *Pni*DAH7PS^D by different metal ions in comparison to Mn²⁺ at a concentration of 100 μ M. The activity of *Pni*DAH7PS in the presence of 100 μ M EDTA is shown as a negative control. The reactions were carried out at 35 °C, in the presence of 100 μ M metal ion, 150 μ M PEP, 225 μ M E4P and 15 μ g enzyme in 50 mM BTP buffer (pH 7.4).

The characterisation of the divalent metal ion dependency of *Pni*DAH7PS^D was carried out as described in section 2.3.1. Mn²⁺ and Co²⁺ were found to activate the apo-enzyme significantly, with Mn²⁺ resulting in the fastest reaction rate (Figure 2.12). The kinetic characteristics of *Pni*DAH7PS^D were determined as described in section 2.3.2 (Appendix Figure A.2.2A, B). The apparent K_M values for PEP and E4P were found to be 367 ± 35 μ M and 239 ± 29 μ M respectively and the k_{cat} value was calculated to be 1.62 ± 0.11 s⁻¹ (Table 2.4).

Table 2. 4: Table of the kinetic parameters for the DAH7PS activities of the wild-type *PniDAH7PS*, *PgiDAH7PS*, *GspDAH7PS*, *BsuDAH7PS* and their truncated DAH7PS variants

Enzyme	K_M^{PEP} (μM)	K_M^{E4P} (μM)	k_{cat} (s^{-1})	$k_{\text{cat}}/K_M^{\text{PEP}}$ ($\mu\text{M}^{-1}.\text{s}^{-1}$)	$k_{\text{cat}}/K_M^{\text{E4P}}$ ($\mu\text{M}^{-1}.\text{s}^{-1}$)
<i>PniDAH7PS</i> ^D	367 ± 35	239 ± 19	1.62 ± 0.11	4.4 × 10 ⁻³	6.8 × 10 ⁻³
<i>PniDAH7PS</i> ^{WT}	41 ± 4	58 ± 6	16.8 ± 0.5	0.41	0.29
<i>PgiDAH7PS</i> ^D (8)	2724 ± 256	3638 ± 375	0.8 ± 0.1	2.9 × 10 ⁻⁴	2.2 × 10 ⁻⁴
<i>PgiDAH7PS</i> (8)	421 ± 43	1238 ± 141	1.6 ± 0.1	3.8 × 10 ⁻³	1.3 × 10 ⁻³
<i>GspDAH7PS</i> ^D (115)	98 ± 11	62 ± 5	63 ± 5	0.64	1.0
<i>GspDAH7PS</i> (115)	87 ± 9	95 ± 8	45 ± 4	0.52	0.47
<i>BsuDAH7PS</i> ^D (8)	141.5 ± 9.1	1400 ± 60	4.2 ± 0.2	3.0 × 10 ⁻²	3.0 × 10 ⁻³
<i>BsuDAH7PS</i> (8)	139.0 ± 11.4	1760 ± 110	4.6 ± 0.1	3.3 × 10 ⁻²	2.6 × 10 ⁻³

The apparent Michaelis constants for the DAH7PS activity are apparent values determined using a fixed concentration of the invariant substrate (as described in section 7.1.13). The uncertainty values for *PniDAH7PS*^{WT} represent the standard deviation from triplicate measurements.

Compared to *PniDAH7PS*^{WT}, both the K_M^{PEP} and K_M^{E4P} of *PniDAH7PS*^D were considerably increased and the k_{cat} value for the truncated variant was much smaller. The $k_{\text{cat}}/K_M^{\text{PEP}}$ and $k_{\text{cat}}/K_M^{\text{E4P}}$ values for *PniDAH7PS*^D were approximately 100 and 42-fold lower, respectively, compared to that for the *PniDAH7PS*^{WT}. Intriguingly, this truncated enzyme was substantially less active than *PniDAH7PS*^{WT}.

2.6.2 Feedback inhibition

In the previous research on *PgiDAH7PS*, the CM domain was demonstrated to act as a regulatory domain for the feedback inhibition of DAH7PS activity by prephenate; removal of the CM domain was shown to abolish this allosteric inhibition (8). To determine if the CM domain of *PniDAH7PS* exerts the same function, the extent of inhibition of *PniDAH7PS*^D by prephenate was investigated using standard kinetic assay conditions. The DAH7PS catalytic activity was tested in the presence of increasing concentrations of prephenate (0–315 μM) and the results show that *PniDAH7PS*^D did not respond to prephenate effectively without the CM domain (Figure 2.13).

Figure 2. 13: The response of *PniDAH7PS^D* (red curve) and *PniDAH7PS^{WT}* (black curve) to varying concentrations of prephenate. Error bars represent the standard deviation from triplicate measurements.

2.6.3 Characterisation of the quaternary structure of *PniDAH7PS^D*

The quaternary structure of *PniDAH7PS^D* in solution was determined by analytical SEC and AUC. Following the calibration of analytical SEC column as described in section 2.3.4, three concentrations (1.5, 0.75 and 0.3 mg/mL) of *PniDAH7PS^D* were applied to the analytical SEC column. A single peak was observed in the elution trace for each concentration and the elution volumes for each protein concentration were found to be similar and to correspond to a molecular weight of ~24 kDa. This indicated that *PniDAH7PS^D* exists as a monomer in solution (Figure 2.14A).

To further confirm the quaternary structure of *PniDAH7PS^D* in solution, *PniDAH7PS^D* was subjected to AUC. Following buffer exchange into fresh AUC buffer containing 10 mM BTP, 150 mM NaCl, pH 7.4, the *PniDAH7PS^D* solution was diluted to three different concentrations, 1.2, 0.4 and 0.1 mg/mL. According to the absorbance versus radial position data, two sedimentation boundaries were observed suggesting that there were two oligomeric species of *PniDAH7PS^D* present in solution. Fitting the experimental data to a continuous model showed that the two species had

sedimentation coefficients of 2.8 S and 4.7 S (Figure 2.14B). The molecular mass for each species was determined as ~29 kDa and ~61 kDa based on the sedimentation coefficients. This is consistent with the molecular masses for the monomeric (30 kDa) and dimeric (60 kDa) assemblies of *Pni*DAH7PS^D (Figure 2.14B). The sedimentation coefficient boundaries obtained from the three protein concentrations did not show an obvious concentration-dependent equilibrium between the monomer and the dimer (Figure 2.14B).

Figure 2. 14: The quaternary structure of *Pni*DAH7PS^D determined by analytical SEC and AUC.

(A) Analytical SEC elution trace and standard curve of logMW against K_{av} for *Pni*DAH7PS^D at 1.5 mg/mL (blue line), 0.75 mg/mL (pink line) and 0.3 mg/mL (yellow line). The SEC buffer contained 10 mM BTP, pH 7.2, 150 mM NaCl and 200 μ M PEP. **(B)** The distribution of the oligomeric species of *Pni*DAH7PS^D at 1.2 mg/mL (blue line), 0.4 mg/mL (red line) and 0.1 mg/mL (black line) obtained from sedimentation velocity analysis in AUC buffer containing 10 mM BTP, pH 7.4 and 150 mM NaCl. Data were collected at 50,000 rpm at a wavelength of 280 nm.

2.7 Basic characterisation of the function and structure of *Pni*DAH7PS^{CM}

2.7.1 Kinetic parameters

The steady-state kinetic characteristics of *Pni*DAH7PS^{CM} were determined following the standard process described in section 7.1.13 (Appendix Figure A.2.2C). The K_M value for chorismate was determined to be $84 \pm 5 \mu\text{M}$, and the k_{cat} value was calculated as $2.01 \pm 0.07 \text{ s}^{-1}$ (Table 2.5). The $k_{\text{cat}}/K_M^{\text{chorismate}}$ value for *Pni*DAH7PS^{CM} is approximately 10 fold lower compared to that of *Pni*DAH7PS^{WT}.

Table 2. 5: Table of the kinetic parameters of the CM activity of the wild-type *Pni*DAH7PS, *Pgi*DAH7PS, *Gsp*DAH7PS, *Bsu*DAH7PS and their truncated CM variants.

Enzyme	$K_M (\mu\text{M})$	$k_{\text{cat}} (\text{s}^{-1})$	$k_{\text{cat}}/K_M (\mu\text{M}^{-1}\text{s}^{-1})$
<i>Pni</i> DAH7PS ^{CM}	84 ± 5	2.01 ± 0.07	2.4×10^{-2}
<i>Pni</i> DAH7PS ^{WT}	7.6 ± 0.6	1.68 ± 0.06	0.22
<i>Pgi</i> DAH7PS ^{CM (8)}	NA	NA	NA
<i>Pgi</i> DAH7PS ⁽⁸⁾	337 ± 15	0.45 ± 0.04	1.3×10^{-3}
<i>Gsp</i> DAH7PS ^{CM (115)}	88 ± 6	1.9 ± 0.1	2.1×10^{-2}
<i>Gsp</i> DAH7PS ⁽¹¹⁵⁾	450 ± 40	5.5 ± 0.52	1.2×10^{-2}
<i>Bsu</i> DAH7PS ^{CM (8)}	1133.6 ± 66.6	0.73 ± 0.02	6.4×10^{-4}
<i>Bsu</i> DAH7PS ^{WT (8)}	850.5 ± 97	0.41 ± 0.01	4.8×10^{-4}

The uncertainty values for *Pni*DAH7PS^{WT} represent the standard deviation from triplicate measurements.

2.7.2 Quaternary structure

The quaternary structure of *Pni*DAH7PS^{CM} in solution was determined using analytical SEC. The analytical SEC column was calibrated following the procedures as described in section 2.3.4 using SEC buffer containing 10 mM BTP, pH 8.1, 150 mM NaCl. Sequentially, 0.5 mL of *Pni*DAH7PS^{CM}, with a concentration of 1.0 mg/mL, was applied to the analytical SEC column and eluted using the same SEC buffer as above. One peak eluted at around 17 mL after injection, corresponding to a molecular weight of

approximately 27 kDa. This indicates that *PniDAH7PS*^{CM} is approximately double the size of the theoretical molecular mass of a single chain of *PniDAH7PS*^{CM} (10 kDa), suggesting that *PniDAH7PS*^{CM} exists as a dimer in solution (Figure 2.15). Because of *PniDAH7PS*^{CM}'s extremely low extinction coefficient ($\epsilon_{280} = 1490 \text{ M}^{-1} \text{ cm}^{-1}$), AUC is not suitable for the determination of the quaternary structure of *PniDAH7PS*^{CM}.

Figure 2. 15: Analytical SEC elution trace and standard curve of logMW against K_{av} for *PniDAH7PS*^{CM} at 1.0 mg/mL (red line). The SEC buffer contained 10 mM BTP, pH 8.1, 150 mM NaCl.

2.8 Small angle X-ray scattering (SAXS) investigations on the global shape of *PniDAH7PS*^{WT} and its truncated variants

At present, the atomic structure for either *PniDAH7PS*^{WT} or the truncated variants is not available. Therefore, small angle X-ray scattering (SAXS) experiments were carried out to obtain the low resolution overall solution-state shapes of *PniDAH7PS*^D, *PniDAH7PS*^{CM}, and *PniDAH7PS*^{WT}, and to further determine the effect of prephenate on the wild-type protein conformation in solution. SAXS is a very powerful technique for investigating structural characteristics of biological macromolecules including proteins. It is able to provide structural information at low resolution, providing information on the average overall size and shape of a molecule in solution. In a typical

SAXS examination on a protein solution sample, the sample is exposed to a highly collimated monochromatic X-ray (wavelength λ) with the wave vector $k = 2\pi/\lambda$. The isotropic scattered intensity I is recorded as a function of the momentum transfer q ($\text{\AA}^{-1}$) $= 4\pi\sin\theta/\lambda$, where 2θ is the angle between the incident and scattered beam (121). The scattering data of protein particles is attained following the subtraction of background signal from the solvent, which is measured separately (121). SAXS experiments were performed using the BioSAXS beamline, with integrated size exclusion chromatography (SEC-SAXS), at the Australian Synchrotron.

2.8.1 Scattering profiles of truncated variants

Ligand-free *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} were examined by SEC-SAXS using a standard SAXS buffer consisting of 20 mM BTP at pH 7.4 (for *Pni*DAH7PS^D) or pH 8.1 (for *Pni*DAH7PS^{CM}), 150 mM NaCl and 3% (v/v) glycerol (protecting sample and buffer from radiation damage in the X-ray beam). The acquired scattering data for *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} exhibited a compact globular and an elongated flat shape respectively (Figure 2.16 and Table 2.6). The Porod volumes for the two variants, 49.7 nm³ for *Pni*DAH7PS^D and 35.8 nm³ for *Pni*DAH7PS^{CM}, combined with a pair-wise distance distribution ($P(r)$ distribution) analysis further verified the monomeric and dimeric assembly of *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} respectively (Figure 2.16E,F and Table 2.6), which are in agreement with the analytical SEC and AUC data.

Figure 2. 16: Analysis of the SEC-SAXS data obtained for *PniDAH7PS*^D (blue) and *PniDAH7PS*^{CM} (red). (A) SAXS profiles ($\log I(q)$ versus q). (B) Guinier plots ($\ln I(q)$ versus q^2). (C) Kratky plot ($q^2 I(q)$ versus q). (D) Porod-Debye plot ($q^4 I(q)$ versus q^4) limited to the range of the SAXS data for which the Guinier linearity was observed. The $P(r)$ and *ab initio* envelopes computed from *PniDAH7PS*^{CM} (E) and *PniDAH7PS*^D (F) experimental data. The fits of the *ab initio* envelopes of *PniDAH7PS*^{CM} and *PniDAH7PS*^D to their corresponding experimental data were $\chi^2 = 0.1$ and 0.6 respectively.

2.8.2 Conformational change of *PniDAH7PS*^{WT} underlying allostery

The wild-type enzyme was assessed following the investigation on the individual separated domains using the standard SAXS buffer described above, supplemented with 200 μ M PEP. The experimental profile of the prephenate-free *PniDAH7PS*^{WT} suggested an elongated protein particle with a Porod volume of 131.3 nm³ (Table 2.6) corresponding to an estimated molecular weight of $\sim$ 77 kDa for *PniDAH7PS*^{WT}, consistent with the proposed homodimeric state of *PniDAH7PS*^{WT} in solution. Remarkably, when prephenate was included, a striking change in protein conformation was observed (Figure 2.17A,B). A significant decrease in the radius of gyration (R_g), the maximum particle diameter (D_{max}) and the Porod volume (Table 2.6) clearly show a transformation of the *PniDAH7PS*^{WT} structure from an elongated shape to a smaller and more compact conformation on prephenate binding.

Table 2. 6: Table of the SEC-SAXS parameters of *PniDAH7PS*^D, *PniDAH7PS*^{CM} and *PniDAH7PS*^{WT} in the absence or presence of prephenate

	<i>PniDAH7PS</i> ^D	<i>PniDAH7PS</i> ^{CM}	<i>PniDAH7PS</i> ^{WT} ligand Free	<i>PniDAH7PS</i> ^{WT} prephenate
Guinier Analysis				
R_g (Å)	19.8 $\pm$ 0.11	22.2 $\pm$ 0.80	42.90 $\pm$ 0.57	33.65 $\pm$ 0.48
$I(0)$ (cm ⁻¹)	0.06 $\pm$ 0.00	0.07 $\pm$ 0.00	0.11 $\pm$ 0.00	0.09 $\pm$ 0.00
Correlation coefficient, R^2	0.99	0.99	0.99	0.99
Pair Wise Distribution Analysis				
R_g (Å)	19.35 $\pm$ 0.01	23.08 $\pm$ 0.01	43.59 $\pm$ 0.10	32.60 $\pm$ 0.09
$I(0)$ (cm ⁻¹)	0.06 $\pm$ 0.00	0.07 $\pm$ 0.00	0.11 $\pm$ 0.00	0.09 $\pm$ 0.00
D_{max} (Å)	57.0	74.0	158.0	105.0
V_p (nm ³)	49.7	35.8	131.3	114.5
<i>MM</i>_{porod}				
<i>MM</i> (kDa)	29.2	21.1	77.3	67.4
Number of Subunits	1	2	2	2

V_p : Porod volume; *MM*_{porod}: molecular mass estimated from the Porod volume. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

The Kratky plot of the experimental data of prephenate-free *PniDAH7PS* exhibits a

collapsed bell-shape curve, at a low q range, indicating a less compact structure containing some not well-folded local regions. However, this relatively irregular structure becomes far more compact and organized in the presence of prephenate, as demonstrated by the more regular bell-shaped plot for the SEC-SAXS data collected in the presence of prephenate (Figure 2.17C). Moreover, the Porod-Debye plot reveals a curve with a plateau in the presence of prephenate, whereas a distinct ascending curve is observed in the absence of prephenate at low q range (Figure 2.17D). This large difference in the Porod-Debye plots illustrates that the intrinsic flexibility of prephenate-free *Pni*DAH7PS^{WT} is notably attenuated by the binding of prephenate (122,123), which is consistent with conformational change indicated by the Kratky plots.

Figure 2. 17: Analysis of the SAXS profiles of *Pni*DAH7PS^{WT} in the absence (blue) or presence (red) of prephenate. (A) SAXS profiles ($\log I(q)$ versus q). (B) Guinier plots ($\ln I(q)$ versus q^2). (C) Kratky plots ($q^2 * I(q)$ versus q). (D) Porod-Debye plots ($q^4 * I(q)$ versus q^4) limited to the range of the SAXS data for which Guinier linearity is observed.

2.8.3 *Ab initio* modelling for *Pni*DAH7PS^{WT}

To visualise the structural information available from the scattering profiles, *ab initio* modelling was employed for the construction of molecular envelopes. *Ab initio* modelling for *Pni*DAH7PS^{WT}, and the truncated variants, was performed using *DAMMIF* (124) with imposed asymmetry of P1 (for the *Pni*DAH7PS^D monomeric barrel) or P2 (for the *Pni*DAH7PS^{CM} or *Pni*DAH7PS^{WT} homodimer). Both *ab initio* models for *Pni*DAH7PS^{WT} in the absence or presence of prephenate resulted in reasonable fits, $\chi^2 = 0.6$ and 0.8 respectively, to the corresponding experimental SAXS data (Figure 2.18A). To further validate these *Pni*DAH7PS^{WT} envelopes, and also specify the spatial relationship between the CM and DAH7PS portions within the wild-type enzyme, the low resolution models for *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} were generated and fitted into the *Pni*DAH7PS^{WT} envelopes. The predicted envelope for the prephenate-free *Pni*DAH7PS^{WT} was elongated, and exhibits symmetry of a dumb-bell shape, strongly suggesting a structure consisting with a central CM dimer with two flanking DAH7PS barrel domains (Figure 2.18B). Two potential interfaces between the TIM barrels and a CM dimer can be clearly identified from the fit between the envelopes of the full-length protein and the two truncated variants (Figure 2.18B). This spatial arrangement is substantially altered when prephenate is present; the model predicts a structure for which the two DAH7PS barrels are found closer to the CM dimer, consistent with a more intimate association between the two DAH7PS domains and the CM dimeric moiety (Figure 2.18B).

Figure 2. 18: *Ab initio* modelling for *PniDAH7PS*^{WT} in the absence or presence of prephenate using the experimental SAXS profiles. (A) The calculated scattering profiles (red line) of the *ab initio* models for the prephenate-free (left) and prephenate-bound (right) *PniDAH7PS*^{WT} overlaid with the corresponding experimental data (black circles, $\chi^2 = 0.6$ and 0.8 respectively). **(B)** The calculated *ab initio* envelopes for the prephenate-free (upper) and prephenate-bound (lower) *PniDAH7PS*^{WT}. The blue and red dummy models are the predicted DAH7PS and CM portions respectively.

2.9 Crystallography

2.9.1 *PniDAH7PS*^{WT}

In order to characterise the hetero-domain interaction and to gain deeper insight into the functional interdependence of the two domains, the crystallisation trials of *PniDAH7PS*^{WT} were carried out. Crystallisation conditions for *PniDAH7PS*^{WT} were screened using the sitting-drop vapour diffusion method and Clear Strategy I, Clear Strategy II, JCSG⁺, PACT, MIDAS, Morpheus (125) and Proplex screen suites (section 7.1.17.1). In addition, the condition of *PniDAH7PS*^{WT} was further varied including a range of protein concentrations (2, 4, 5, 6 and 10 mg/mL), supplemented with a ligand or ligand combinations (PEP, E4P, Mn²⁺, prephenate) and two growth temperatures (5 and 20 °C). Unfortunately, all the crystal screen trails failed to grow any crystals, which may be attributed to the large flexibility of *PniDAH7PS*^{WT}, as shown in the SAXS data (Figure 2.17C, D). In addition, it was found that *PniDAH7PS*^{WT} precipitated in screening drops once its concentration was higher than 4 mg/mL. Even when in the SEC buffer (20 mM BTP, 150 mM NaCl and pH 7.4), *PniDAH7PS*^{WT} is prone to precipitation at a concentration higher than 10 mg/mL and aggregation of *PniDAH7PS*^{WT} was seen in the AUC sedimentation plots (Figure 2.8). These observations suggest some probable intermolecular interactions between *PniDAH7PS*^{WT} dimers, which may be another factor interfering with the formation of an ordered crystal lattice of *PniDAH7PS*^{WT}.

2.9.2 *PniDAH7PS*^D

Crystallisation conditions for *PniDAH7PS*^D were screened using the sitting-drop vapour diffusion method and using the Clear Strategy I, Clear Strategy II, JCSG⁺, PACT screen suites. Crystals of apo-*PniDAH7PS*^D were found to grow in a number of conditions, and some of the conditions shared the same compositions, but differ in pH or the concentration of precipitant. After combining the highly similar conditions, four conditions were finally chosen due to the reasonable size and appearance of the

crystals they produced (Table 2.7). Optimisations of these conditions were performed for *Pni*DAH7PS^D, at 5 mg/mL and 10 mg/mL, by varying the concentrations of salts and precipitants in either the presence or absence of 500 μ M PEP/Mn²⁺.

Table 2. 7: Table of the conditions for the crystallisation of *Pni*DAH7PS^D

Condition	Salt	Buffer	Precipitant	Diffraction quality and resolution
E4 (JCSG⁺)	0.2 M Li ₂ SO ₄	0.1 M Tris, pH 8.5	1.26 M (NH ₃) ₂ SO ₄	Smeary 3.2 Å
A1 (Clear II)	1.5 M (NH ₃) ₂ SO ₄	0.1 M Sodium Acetate, pH 5.5	None	Smeary 2.6 Å
G7 (JCSG⁺)	0.1 M succinic acid	None, pH 7.0	15% w/v PEG 3350	No obvious diffraction
D6 (Clear I)	0.8 M Sodium formate	0.1M Sodium cacodylate, pH 6.5	10% w/v PEG 8000 10% w/v PEG 1000	Smeary 3.7 Å

The best crystals, grown either in the presence or absence of 500 μ M PEP/Mn²⁺, were looped out (Figure 2.19A-D) and examined using either the MX1 or MX2 beamline at the Australian Synchrotron. The crystals from condition E4 of JCSG⁺ and condition A1 of Clear II both gave rise to diffraction data sets with workable resolutions around 3.0 Å. However, all the diffraction patterns exhibited extremely streaky diffraction spots (Figure 2.19E, F). It appears that the X-ray Detector Software (XDS) (126), although very powerful in integrating diffraction data, failed to effectively identify single spots and integrated some split spots as single spots, which is indicated by the multiple twining fractions in the diffraction data sets. The best data set, in the C121 space group, was promising with an overall R_{merge} of 7.9% to a resolution of 2.6 Å (Table 2.8), and Phaser (127) was able to place the search model (the monomeric unit of *Pfu*DAH7PS, PDB code ID: 4C1L) in the asymmetric unit and provide an initial solution for the data set. However due to the smeary diffraction spots, neither the refinement nor the intensity based twin refinement of the structure using Refmac5 could be progressed

satisfactorily, and finally ended up with a R of 0.38 and a R_{free} of 0.39. The speculated explanation for the smeary diffraction pattern is an irregular crystal lattice, which could be caused by a number of factors, such as the high flexibility of some local structures within the DAH7PS barrel. To solve this issue, and to attain improved diffraction data, further attempts at optimising the crystallisation condition and cryoprotectant are required.

Table 2. 8: Assessment of the quality of the diffraction data set of *Pni*DAH7PS^D using aimless (128).

	Overall	InnerShell	OuterShell
Low resolution limit (Å)	43.86	43.86	2.72
High resolution limit (Å)	2.6	9.01	2.6
R_{merge} (within I+/I-)	0.079	0.056	1.158
R_{merge} (all I+ & I-)	0.083	0.057	1.213
R_{meas} (within I+/I-)	0.093	0.067	1.364
R_{meas} (all I+ & I-)	0.09	0.063	1.311
R_{pim} (within I+/I-)	0.049	0.037	0.715
R_{pim} (all I+ & I-)	0.035	0.026	0.494
R_{merge} in top intensity bin	0.048	-	-
Total number of observations	71342	1590	8890
Total number unique	10339	278	1244
Mean($\langle I \rangle / \sigma(I)$)	13.9	27.6	2.9
Mn(I) half-set correlation $CC_{1/2}$	0.998	0.997	0.842
Completeness	100	99.2	100
Multiplicity	6.9	5.7	7.1
Anomalous completeness (%)	99.9	100	100
Anomalous multiplicity	3.6	3.3	3.7
DelAnom correlation between half-sets	-0.158	-0.115	-0.091
Mid-Slope of Anom Normal Probability	0.815	-	-

Figure 2. 19: Crystals of *PniDAH7PS^D*. (A) Crystals grown from E4 JCSG⁺. (B) Crystals grown from A1 Clear II. (C) Crystals grown from G7 JCSG⁺. (D) Crystals grown from D6 Clear I. (E) X-ray diffraction pattern for a crystal grown from A1 Clear II. (F) X-ray diffraction pattern for a crystal grown from E4 JCSG⁺.

2.9.3 *PniDAH7PS^{CM}*

Crystallisation conditions for *PniDAH7PS^{CM}* were screened using the sitting-drop vapour diffusion method with Clear Strategy I, II JCSG⁺ and PACT screens. Although crystals grew in many conditions, the crystals were generally low quality, exhibiting

thin needle-like shape (Appendix Figure A.2.3). The conditions C8 and D9 from Clear I, B6 and C12 from Clear II and B9 from JCSG⁺ were repeated and optimised in either the presence or absence of 500 μ M prephenate (Table 2.9). However, all the attempts failed to reproduce or improve the quality of the crystals. Only crystals of apo-*Pni*DAH7PS^{CM}, grown in conditions C12 (Figure 2.20A) from Clear II or B9 (Figure 2.20B) from JCSG⁺, were successfully looped out and examined using the MX2 beamline. however, none of these crystals were able to produce an X-ray diffraction pattern.

Table 2. 9: Table of the conditions for the crystallisation of *Pni*DAH7PS^{CM}

Condition	Salt	Buffer	Precipitant
C8 (Clear I)	2 M Li ₂ SO ₄	0.1 M Sodium cacodylate, pH 6.5	15% w/v PEG 4000
C9 (Clear I)	0.2 M MgCl ₂	0.1 M Sodium cacodylate, pH 6.5	15% w/v PEG 4000
B6 (Clear II)	0.15 M Potassium thiocyanate	0.1 M Sodium acetate, pH 5.5	20% w/v PEG 1500
C12 (Clear II)	0.2 M Calcium acetate	0.1 M Sodium cacodylate, pH 6.5	8% w/v PEG 20,000, 8% v/v PEG 500 MME
B9 (JCSG⁺)	None	0.1 M Na citrate, pH 5.0	20% w/v PEG 6K

Figure 2. 20: Crystals of *Pni*DAH7PS^{CM}. (A) Crystals grown from B9 JCSG⁺. (B) Crystals grow from C12 Clear II.

2.10 Summary and discussion

As mentioned in chapter 1, multifunctional enzymes play diverse roles in improving the performance of metabolic pathways. In some cases, multiple catalytic capacities that are responsible for consecutive reactions are merged into a single enzyme and boost the overall reaction rate by forming a substrate channelling system between different active sites (50). Besides, some organisations of multifunctional enzymes aid in the construction of allosteric systems (8,20). The CM that is found at the N-terminus of some type I β DAH7PSs is one of the enzymes whose role for allostery has been well described (8,20).

In well-studied CM-DAH7PS fusion enzymes, such as *Gsp*DAH7PS and *Bsu*DAH7PS, the functional CM moiety accommodates prephenate and delivers allosteric control of the DAH7PS catalytic activity. When the allosteric effector is absent, the two functional portions exhibit little hetero-domain interplay and almost work independently, which implies a compartmentalisation between allostery and catalytic function (8,20). In other words, the N-terminal CM domain plays as a removable regulatory accessory for its fusion partner. However, this study on *Pni*DAH7PS presents an exceptional relationship between DAH7PS and its C-terminal CM extension (8). The removal of the CM domain does not just deprive DAH7PS of allostery, but also significantly impairs both DAH7PS and CM catalytic capacities, which suggests that the grafted CM is an essential partner for both the efficient and regulated catalysis of DAH7PS, in the other way round, DAH7PS also delivers the functionality to CM.

In addition to the functional interdependence between the two enzymatic moieties, *Pni*DAH7PS also displays a unique dimeric quaternary assembly. Moreover, the truncated *Pni*DAH7PS barrel exists as a monomer under experimental conditions, which is a dramatic difference from either the dominant homotetrameric or the sparse homodimeric assemblies found in all other well-known DAH7PS, and their variants, to

date (18,20,21,33,34,109). Protein structure determines protein function; thus, it is reasonable to propose that the altered quaternary structure of the C-terminal CM-fused DAH7PS facilitates its special interdomain functional interdependence.

Due to the lack of high resolution structures, the correlation between the unique quaternary structure and the functional characteristics of *Pni*DAH7PS currently cannot be probed easily. Alternatively, the self-associations of the DAH7PS barrels within the homotetrameric structure of DAH7PS enzymes, from the main type I β group, were analysed and referred to in order to deduce the 'structure-function' relationship for *Pni*DAH7PS. Take *Pfu*DAH7PS as a representative, the two distinct types of homo-interfaces were identified by *PISA* (Proteins, Interfaces, Structures and Assemblies) (32) and denoted as the tetramer interface and the dimer interface (Figure 9) (33). The tetramer interface, relatively distant from the active site, contains fairly strong hydrophobic interactions, which primarily contributes to the enzyme's thermostability rather than the enzymatic activity, as demonstrated by a mutagenesis assay where this hydrophobic interface was broken (33). In comparison, the dimer interface is located in the vicinity of the active site, burying a number of segments extensively involved in catalysis, which suggests a critical role for this interface in organising and maintaining the active site of DAH7PS. However, for *Pni*DAH7PS, the unusual quaternary structure implies that interaction between the DAH7PS barrels is absent. Meanwhile, the spatial fit of the DAH7PS and CM dummy models into the envelope of ligand-free *Pni*DAH7PS is suggestive of a potential DAH7PS-CM hetero-interface within this bifunctional enzyme (Figure 8B). Thus, we hypothesise that a hetero-interaction between a DAH7PS barrel and a CM dimer, as opposed to the homo-interaction, plays an equivalent buttressing role in the construction of both the DAH7PS and CM active sites.

If this speculated hetero-domain interaction provides the structural basis for the enzymatic functions of both DAH7PS and CM, interrupting this interaction could be an effective way to down-regulate both catalytic activities. Actually, the dramatic

conformational change of *Pni*DAH7PS upon prephenate binding verifies that a reorganisation of an interdomain interaction happens during allosteric regulation. Thus, it is reasonable to suggest that inducing a substitution of the hetero-interface to disrupt the functional interdomain interaction is the way by which prephenate allosterically affects the enzymatic function of DAH7PS.

Taken together, the distinguished characteristics of *Pni*DAH7PS suggests that the role of the C-terminal CM moiety is a functional partner of the DAH7PS domain, but is not an accessory to regulate substrate access to the active site of DAH7PS. In *B. subtilis* and *Geobacillus sp*, the N-terminal CM-linked DAH7PS coexists with an unregulated mono-functional CM, which shows a fairly high catalytic ability to increase the cellular levels of prephenate (8). This fact may explain why the N-terminal CM functions more specifically for allosteric inhibition on DAH7PS activity. Unlike *B. subtilis* and *Geobacillus sp*, *P. nigrescens* or *P. gingivalis* only carry one CM enzyme, embodied in the bifunctional DAH7PS. Therefore, for *Pni*DAH7PS and *Pgi*DAH7PS, securing essential CM function appears to have occurred prior to the introduction of allostery to the DAH7PS by the fusion of the DAH7PS and CM genes. In this case, the CM and DAH7PS portions play equal roles in supporting each other's activities. Moreover, the mechanism underlying the allostery of *Pni*DAH7PS is the interference of the cooperative association between DAH7PS and CM to deprive the fusion protein of the functional buttress. This novel relationship between two functional moieties, observed in *Pni*DAH7PS, unravels a distinctive evolutionary strategy for delivering enzymatic functionality and acquiring allosteric regulation simultaneously.

Chapter 3 *DAH7PS and CM functional interdependence in PniDAH7PS*

3.1 Introduction

An enhancement of catalytic efficiency is often presented as an important consequence of enzyme multifunctionality. This enhancement usually emerges from the associations between the enzymes that catalyse consecutive reactions within a biosynthetic pathway (50,52,59). In this type of multifunctional enzyme, the channelling systems that connect the two active sites effectively prevents the free diffusion of metabolites (129). However, for *Pni*DAH7PS^{WT}, the two functional domains catalyse non-consecutive reactions of the shikimate pathway, which means there is no substrate channelling advantage for the two domains to be structurally linked together. Thus, the molecular basis for the functional interdependence between the two domains in *Pni*DAH7PS is unclear. A previous report for *Pgi*DAH7PS proposed that the functional inter-dependence may be contributed to by a large buried area between the CM and DAH7PS domains (8); however, there is no experimental data to support this speculation. In chapter 2, a structural analysis of *Pni*DAH7PS^{WT}, *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} using SAXS, combined with *ab initio* models derived from SAXS profiles, suggested that a direct interaction existed between the DAH7PS and CM domains. However, direct experimental evidence to demonstrate this hetero-domain association has not been obtained. Therefore, the work described in this chapter aimed to uncover the mechanism underpinning the inter-dependence between the DAH7PS and CM functions. The results suggest that a dynamic spatial arrangement of the two domains along the DAH7PS catalytic cycle exists, demonstrating the essential role of these hetero-domain interactions in the enzymatic functions of both the DAH7PS and CM.

3.2 Hetero-domain interaction

Kinetic, SAXS and ITC assays were carried out to investigate the hetero-interactions between the DAH7PS and CM domains in *Pni*DAH7PS^{WT}.

3.2.1 DAH7PS catalysed reaction promotes CM activity

In chapter 2, data showing the kinetic parameters for either the DAH7PS or the CM domain alone, determined in the absence of the other catalytic reaction taking place, were presented (Figure 3.1A,B). However, in order to check whether substrate binding and/or the catalytic reaction taking place in one domain affects the enzymatic function of the other domain, the kinetic parameters should be measured when both reactions are taking place simultaneously (Figure 3.1C). Two main issues emerged when detecting DAH7PS activity in the presence of chorismate. Firstly, the absorption spectrum of chorismate partially overlaps with the PEP absorption spectrum, including at a wavelength of 232 nm (the λ_{max} of PEP and the wavelength for monitoring the DAH7PS-catalysed reaction) (Figure 3.2). This means that the reduction in absorption at 232 nm would reflect the combined consumption of PEP and chorismate rather than that of PEP alone. Secondly, the prephenate produced by the enzymatic activity of CM may inhibit the function of the DAH7PS domain. Therefore, measuring CM activity in the presence of the DAH7PS reaction is the only option for a direct absorbance-based assay.

Figure 3. 1: Schematics illustrating the kinetic assays for *Pni*DAH7PS^{WT}. (A) and (B) respectively show the assays measuring the individual catalytic activity of DAH7PS or CM and (C) illustrates the experiment for examining the effect of the DAH7PS catalysed reaction on CM activity.

The catalytic activity of the CM domain was determined by measuring the consumption of chorismate at 274 nm. To detect whether any changes were seen at this wavelength when the DAH7PS reaction was taking place, the absorptions of MnSO₄, PEP, E4P and DAH7P at 274 nm were checked. None of these additives showed significant absorption at this wavelength. To further confirm this, a DAH7PS catalysed reaction, comprising 350 μM PEP, 350 μM E4P, 100 μM MnSO₄ and 6 μg of enzyme was initiated and the absorbance at 274 nm was monitored. No significant absorption was observed over time.

Figure 3. 2: UV absorption spectra of PEP (Blue) and chorismate (Red).

The catalytic activity of the CM domain of *Pni*DAH7PS^{WT} was examined under a number of different conditions; in the presence of MnSO₄, PEP, E4P and DAH7P alone, or in the presence of the combinations of Mn²⁺/PEP, Mn²⁺/E4P, PEP/E4P and PEP/Mn²⁺/E4P. The results show a ~ two-fold increase in CM activity in the presence of the PEP/Mn²⁺/E4P combination (a DAH7PS reaction mixture) (Figure 3.3A); whereas there was no observable effect on the CM reaction rate under the other tested conditions. This observation implies that there is an interplay between occupancy and catalysis at the DAH7PS active site and the CM domains of *Pni*DAH7PS^{WT}.

Figure 3. 3: The promoting effect of the DAH7PS catalysed reaction on the CM activity of *PniDAH7PS*^{WT}. Error bars represent the standard deviation of triplicate measurements. **(A)** The catalytic rates of the CM domain in *PniDAH7PS*^{WT} were examined in the presence of various substrates, substrates combinations or DAH7P. The CM catalysed reactions, consisting of standard kinetic buffer (Section 7.1.13) and 6 μ g of *PniDAH7PS*^{WT}, were initiated by the addition of 100 μ M chorismate in the presence of 100 μ M MnSO₄, 350 μ M PEP, 350 μ M E4P and 350 μ M DAH7P alone, or combinations of Mn²⁺ (350 μ M)/PEP (350 μ M), Mn²⁺ (350 μ M)/E4P (350 μ M), PEP (350 μ M)/E4P (350 μ M) and PEP (350 μ M)/Mn²⁺ (350 μ M)/ E4P (350 μ M) respectively. **(B)** The effects of various substrates and substrates combinations of the DAH7PS catalysed reaction on the CM catalytic rate were examined for *GspDAH7PS*^{WT} using the same conditions that were used for *PniDAH7PS*^{WT}.

As an additional control, the bifunctional *GspDAH7PS* was also subjected to the same experimental measurements under all aforementioned conditions. Under the condition in the presence of PEP/Mn²⁺/E4P, which enhanced the CM activity of *PniDAH7PS*^{WT}, the CM activity of *GspDAH7PS* showed no rate enhancement (Figure 3.3B). This result provides strong support for the proposition that CM catalytic activity is enhanced by binding and/or catalytic events at the DAH7PS domain for *PniDAH7PS* specifically.

3.2.2 Conformational changes of *PniDAH7PS*^{WT} with added PEP/Mn²⁺ or PEP/Mn²⁺/E4P

The SAXS profile of *PniDAH7PS*^{WT} in the presence of PEP indicated a highly dynamic structure for this protein (Section 2.8.2). Given that the separate *PniDAH7PS*^D and *PniDAH7PS*^{CM} were characterised as a globular protein and a compact rod-shape entity, respectively by the SAXS analysis (Section 2.8.1), the observed flexibility of *PniDAH7PS*^{WT} is not likely to arise from the structural dynamics within the individual domains. Instead, relative motions between the DAH7PS and CM domains may account for the dynamic features of *PniDAH7PS*^{WT}. Therefore, SAXS analysis was employed to examine the structural changes of *PniDAH7PS*^{WT} in the presence of various DAH7PS substrates and substrate combinations aiming to capture information about the spatial rearrangement of the DAH7PS and CM domains with different ligands present.

Apo-*PniDAH7PS*^{WT} showed an identical SAXS profile compared to that of *PniDAH7PS*^{WT} in the presence of PEP or MnSO₄, indicating that neither PEP or MnSO₄ alone are not able to alter the enzyme's conformation (Appendix Appendix Table A.3.1 and Appendix Figure A.3.1). In the presence of PEP/Mn²⁺/E4P (*PEP/Mn*²⁺/*E4P-PniDAH7PS*^{WT}) or PEP/Mn²⁺ (*PEP/Mn*²⁺-*PniDAH7PS*^{WT}) *PniDAH7PS*^{WT} demonstrated conformational

rearrangements (Figure 3.4A). In the presence of PEP/Mn²⁺, *PniDAH7PS*^{WT} was more compact than the apo-protein, showing reductions of R_g and D_{max} values from 42.0 ± 1.2 and $152.7 \text{ \AA}$ to $39.4 \pm 0.8 \text{ \AA}$ and $142.7 \text{ \AA}$, respectively (Table 3.1).

Table 3. 1: SAXS parameters of apo-*PniDAH7PS*^{WT}, PEP/Mn²⁺-*PniDAH7PS*^{WT} and PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT}

	Apo-protein	PEP/Mn ²⁺	PEP/Mn ²⁺ /E4P
Guinier Analysis			
R_g (Å)	42.0 ± 1.2	39.4 ± 0.8	35.8 ± 1.0
$I(0)$ (cm ⁻¹)	0.08 ± 0.00	0.07 ± 0.00	0.07 ± 0.00
Correlation coefficient, R^2	0.99	0.99	0.99
Pair Wise Distribution Analysis			
R_g (Å)	44.1 ± 0.2	41.0 ± 0.1	37.7 ± 0.6
$I(0)$ (cm ⁻¹)	0.08 ± 0.00	0.07 ± 0.00	0.07 ± 0.00
D_{max} (Å)	152.3	142.7	140.2
V_p (nm ³)	137	139	121
MM_{porod}			
MM (kDa)	81	82	71
Number of Subunits	2	2	2

V_p : Porod volume; MM_{porod} : molecular mass estimated from Porod volume. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

As with the effect of prephenate on *PniDAH7PS*^{WT}, the addition of PEP/Mn²⁺ rigidified the structure of *PniDAH7PS*^{WT}. This was demonstrated by the bell-shaped Kratky plot with a better-defined maximum (Figure 3.4C). Moreover, the Porod-Debye plot presents a curve with a plateau in the low q region when PEP/Mn²⁺ is present (Figure 3.4D). This observation implies that the intrinsic flexibility of apo-*PniDAH7PS*^{WT} is notably attenuated upon the binding of PEP/Mn²⁺ and this is consistent with the well-organised conformation of *PniDAH7PS*^{WT} indicated by the Kratky plot.

Figure 3. 4: Analysis of the SAXS data of apo-*PniDAH7PS*^{WT} (black) and PEP/Mn²⁺-*PniDAH7PS*^{WT} (blue) and PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT} (red). (A) SAXS profiles ($\log I(q)$ versus q). (B) Guinier plots ($\ln I(q)$ versus q^2). (C) Kratky plot ($q^2 \cdot I(q)$ versus q). (D) Porod-Debye plot ($q^4 \cdot I(q)$ versus q^4) limited to the range of the SAXS data for which the Guinier linearity was observed.

Intriguingly, when PEP/Mn²⁺/E4P were present the scattering profile of *PniDAH7PS*^{WT} (PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT}) displayed a smaller average size of *PniDAH7PS*^{WT} compared to either apo-*PniDAH7PS*^{WT} or PEP/Mn²⁺-*PniDAH7PS*^{WT}. R_g , D_{max} and V_p values were reduced to 37.7 ± 0.6 Å, 140.2 Å and 121 nm³ respectively (Table 3.1). The Kratky and Porod-Debye plots indicate that the average overall conformation of PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT} is more rigidified and better folded compared to that of apo-*PniDAH7PS*^{WT}. In the presence of a full complement of substrates and metal ion (PEP, Mn²⁺ and E4P), the DAH7PS reaction is expected to be proceeding through the catalytic cycle and the reaction progresses of different enzyme molecules is unsynchronised. This makes it impossible to identify which catalytic step(s)

determined the features of the SAXS profile of PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT}. However, the smaller R_g value of this experimental data, compared to that in either the absence (ligand free) or presence of PEP/Mn²⁺, strongly suggests that it is impossible that this conformational ensemble is simply a mixture of the conformer populations of apo-*Pni*DAH7PS^{WT} and PEP/Mn²⁺-*Pni*DAH7PS^{WT}, and a conformational distribution, presenting a more compact average structure(s), appears to be formed during the DAH7PS catalytic cycle.

3.2.3 Rigid body modelling for apo-*Pni*DAH7PS^{WT}, PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT}

To visualize the conformational changes indicated by the scattering profiles described above, rigid body modelling was carried out using homology models for *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} and the real space information extracted from the SAXS data. Homology modelling of monomeric *Pni*DAH7PS^D and dimeric *Pni*DAH7PS^{CM} was performed with Modeller using *Pfu*DAH7PS (23) and *Pfu*CM (130) as template structures, respectively (131). Then, the two resultant homology models were connected together as a full length monomeric unit (*Pni*DAH7PS-CM) containing an undefined linker using the software pre_bunch (132). Subsequently, two *Pni*DAH7PS-CM monomeric models were docked into the SAXS *ab initio* envelope of apo-*Pni*DAH7PS^{WT}, PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} to generate a rigid body model using BUNCH (132), constrained by two-fold symmetry and distance information of the contacting residues obtained from the homology model of the *Pni*DAH7PS^{CM} dimer (Table 3.2).

Table 3. 2: Contact information for rigid body modelling using BUNCH

Residue of chain 1	Residue of chain 2	Distance between C _α atoms (Å)
R272	Q309	12.0
D276	R312	11.0
D279	R289	8.5
R289	D279	8.5
R312	D276	11.0
K319	D276	10.0
I340	I341	10.0
I341	I340	10.0

Rigid body models for apo-*Pni*DAH7PS^{WT}, PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} yield reasonable fits ($\chi^2 = 0.3, 1.7$ and 1.3 respectively) to their corresponding experimental SAXS data (Figure 3.5). Both the models of PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} adopt more compact spatial organizations compared to that of apo-*Pni*DAH7PS^{WT} (Figure 3.5B,C). Particularly, in the model for PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT}, two DAH7PS barrels form close contacts with the CM dimer, which might provide a structural basis for the increase in the CM activity induced by the addition of PEP/Mn²⁺/E4P (Figure 3.5C). These rigid body models hypothesise a diverse range of possible hetero-domain interactions along the DAH7PS catalytic cycle.

Figure 3. 5: Rigid body modelling for apo-*PniDAH7PS*^{WT}, PEP/Mn²⁺-*PniDAH7PS*^{WT} and PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT} using SAXS profiles. DAH7PS and CM domains are respectively coloured blue and red. (A) Rigid body model of apo-*PniDAH7PS*^{WT} and the fit to the experimental SAXS data ($\chi^2=0.3$). (B) Rigid body model of PEP/Mn²⁺-*PniDAH7PS*^{WT} and the fit to the experimental SAXS data ($\chi^2=1.7$). (C) Rigid body model of PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT} and the fit to the experimental SAXS data ($\chi^2=1.3$).

3.2.4 Isothermal titration calorimetry (ITC) to investigate the role of Mn²⁺ in PEP binding

DAH7PS enzymes are dependent on divalent metal ions for catalytic activity: Mn²⁺ was shown to be the most activating metal for *PniDAH7PS*^{WT} (Section 2.4.1). Moreover, SAXS analysis, examining the structural effect of PEP and Mn²⁺ on *PniDAH7PS*^{WT}, showed that the conformational change of *PniDAH7PS*^{WT}, caused by the addition of PEP, was observed only when Mn²⁺ was present. To determine whether Mn²⁺ affects

PEP binding directly, ITC experiments were carried out.

Figure 3. 6: ITC experiments to detect the effect of Mn^{2+} on the binding of PEP to *PniDAH7PS^{WT}*. Raw data from ITC experiments involving the titration of 1.5 mM PEP into the cell containing 65 μM *PniDAH7PS^{WT}* in the absence or presence of 500 μM MnSO_4 are shown as (A) and (B) respectively. (C) Comparison of K_D for the binding of PEP to *PniDAH7PS^{WT}* in the absence and presence of 500 μM Mn^{2+} . No binding was observed when Mn^{2+} was absent. (D) The thermodynamic parameters for the binding of PEP to *PniDAH7PS^{WT}* in the presence of 500 μM MnSO_4 .

ITC measurements directly determine the change enthalpy (ΔH) and the disassociation constant (K_D , the binding affinity) associated with a protein-ligand binding event in the whole system comprising solvent and colloidal solutes. The changes of the total Gibbs free energy (ΔG) and entropy (ΔS) are calculated using the equations $\Delta G = -RT \ln K_D = \Delta H - T\Delta S$ (where R is the gas constant; T is the absolute temperature). The total ΔG of a ligand binding event contains contributions from a number of molecular interactions including hydrogen bonding, hydration, electrostatic and ionisation effects and van

der Waals interactions, which contribute principally to ΔH , and conformational entropy (133). Two sets of ITC experiments were performed by titrating 1.5 mM PEP into 65 μM *Pni*DAH7PS^{WT} in either the absence or presence of 500 μM MnSO₄. The obtained data clearly showed that the affinity of *Pni*DAH7PS^{WT} for PEP was too weak to detect when Mn²⁺ was absent. However, following the addition of Mn²⁺ the binding of PEP was enhanced, showing a K_D of $52 \pm 6 \mu\text{M}$ (Figure 3.6C). The binding isotherms recorded in the presence of Mn²⁺ were consistent with an independent binding model, implying that the PEP binding events occurring at the two DAH7PS active sites within the homodimeric *Pni*DAH7PS^{WT} are independent (Figure 3.6B). This observation suggests that assisting PEP binding to the active site is one of the ways whereby Mn²⁺ aids the catalysis of *Pni*DAH7PS^{WT}. This result could also explain why neither PEP nor Mn²⁺ on their own were observed to alter the *Pni*DAH7PS^{WT} conformation in the SAXS analysis. Moreover, the thermodynamic parameters obtained from the binding isotherm indicate that PEP binding to *Pni*DAH7PS^{WT} in the presence of Mn²⁺ is an entropically driven process (Figure 3.6D).

3.3 The essential role of hetero-domain interactions in delivering the functions of *Pni*DAH7PS^{WT}

As discussed above, the intrinsic flexibility of *Pni*DAH7PS^{WT} appears to be mostly due to the relative motions between the DAH7PS and CM domains. The observations of a conformational change of *Pni*DAH7PS^{WT} in the presence of either PEP/Mn²⁺ or PEP/Mn²⁺/E4P and the enhanced CM activity by the addition of PEP/Mn²⁺/E4P suggest a close relationship between the structural rearrangements and enzymatic catalysis of *Pni*DAH7PS^{WT}. In order to check this proposed structural and functional relationship, two experiments were designed to disrupt the association between the DAH7PS and CM domains and to examine the changes in the catalytic abilities of DAH7PS and CM.

3.3.1 Kinetic characterisation and SAXS profiles of *Pni*DAH7PS^{Δ5AA}

Sequence analysis and generation of a truncated variant

A pairwise sequence alignment between *Pni*DAH7PS^{CM} and *Pfu*CM revealed an extra 10 amino acids in the C-terminal region of *Pni*DAH7PS^{CM} compared to *Pfu*CM (Figure 3.7A). The homology model of *Pni*DAH7PS^{CM} generated using *Pfu*CM as a template displayed a loop structure for this poorly aligned region between helices 2 and 3 (loop_{C320-F333}) (Figure 3.7B). A further sequence alignment was constructed for *Pfu*CM, *Ape*CM and the CM domains of *Pni*DAH7PS, *Pgi*DAH7PS, *Gsp*DAH7PS and *Lmo*DAH7PS (*Pni*DAH7PS^{CM}, *Pgi*DAH7PS^{CM}, *Gs*DAH7PS^{CM} and *Lmo*DAH7PS^{CM}). All of the six proteins are classified as members of the AroQ_α CM subclass. *Pfu*CM and *Ape*CM are monofunctional enzymes, whereas the other four enzymes are the N-terminal (*Gsp*DAH7PS^{CM} and *Lmo*DAH7PS^{CM}) or the C-terminal (*Pni*DAH7PS^{CM} and *Pgi*DAH7PS^{CM}) CM moieties of DAH7PS-CM bifunctional enzymes (8,20,28,134,135). This alignment suggests that the CM domains from the bifunctional enzymes, whether they are linked to the C-terminal or N-terminal of their companion DAH7PS, all have a ~10 amino acid residue unaligned region compared to the monofunctional CM sequences (Figure 3.7A). Protein structure alignment between the crystal structures of *Pfu*CM and *Gsp*DAH7PS^{CM} verified that the unaligned region is in the linker region connecting the 2nd and 3rd helices of *Gsp*DAH7PS^{CM} (Figure 3.7B). However, the previous study on *Gsp*DAH7PS did not characterise the function of this long linker for either CM or DAH7PS activity, or on allosteric function, and no obvious role for this linker was observed in the prephenate mediated allosteric inhibition (20).

Figure 3. 7: Sequence and structural alignments predict a long loop within *Pni*CM. (A) The sequence alignment of *Pni*CM with the CM portions of other bifunctional DAH7PS and monofunctional AroQ CMs as indicated. The red rectangle highlights the unaligned region that is suspected to be a long loop. **(B)** The structural alignments of *Pfu*CM (blue cartoon) with *Gsp*CM (orange cartoon) and *Pni*CM (red cartoon) respectively, showing the linker (ribbon) between helices 2 and 3.

Recent studies on protein dynamics have revealed crucial roles of disordered regions, such as loops and linkers, in the conformational dynamics associated with catalysis and allostery such as the role of loops in the functions of aldolase and dihydrofolate reductase (136-138). Due to the lack of any atomic structures for C-terminal CM-linked DAH7PS proteins, and functional information for the counterpart linker in N-terminal CM-linked DAH7PSs, it is hard to accurately predict the structure and role of the

unaligned region in *PniDAH7PS*^{WT}. Therefore, in order to investigate the contribution of this segment to the domain interactions and to the functional interdependence in *PniDAH7PS*, a variant enzyme with a shortened linker region was generated.

Five consecutive residues, Cys-Gly-Ile-Asp-Thr, at the centre of the predicted loop_{C320-F333} were removed by mutagenesis (Figure 3.8A). The homology model for the CM domain of this truncated variant (named as *PniDAH7PS*^{Δ5AA}) predicts normal folding of the three helices of the CM domain. The gene for *PniDAH7PS*^{Δ5AA} was cloned into the pET28a expression vector and transformed into BL21 (DE3) chaperone combination 3 cells (Table 7.2). An expression trial showed that even though *PniDAH7PS*^{Δ5AA} did not exhibit the same solubility as the wild type protein, still around half of the expressed *PniDAH7PS*^{Δ5AA} was observed in the soluble fraction following cell lysis (Figure 3.8B). Thereafter, cells containing the expressed protein were harvested, lysed, and purified using IMAC and SEC as described for the purification of *PniDAH7PS*^{WT} (Figure 3.8B).

Figure 3. 8: Loop truncation within the CM portion of *PniDAH7PS*^{WT}. **(A)** Structure alignment between the CM portions of the homology models of *PniDAH7PS*^{WT} (red cartoon) and *PniDAH7PS*^{Δ5AA} (yellow cartoon). The 2nd linkers are highlighted by ribbon representations. **(B)** SDS-PAGE analysis for the expression and purification of *PniDAH7PS*^{Δ5AA}. Lane 1: Whole cell lysate. Lane 2: Supernatant after centrifugation, the soluble fraction of the cell lysate. Lane 3: Cell debris, the insoluble fraction of the cell lysate. Lane 4–10: The fractions of the elution peak collected from the IMAC.

Kinetic characterisation of *PniDAH7PS*^{Δ5AA}

Kinetic assays for the DAH7PS and CM catalytic activities of *PniDAH7PS*^{Δ5AA} were performed under the same conditions used for *PniDAH7PS*^{WT} (Section 2.4.2) by monitoring the disappearance of PEP at 232 nm and chorismate at 274 nm respectively.

Table 3. 3: Kinetic parameters of the DAH7PS and CM activities of *PniDAH7PS*^{Δ5AA} and a comparison to kinetic parameters of *PniDAH7PS*^{Δ5AA}, *PniDAH7PS*^{WT}, *PniDAH7PS*^D, and *PniDAH7PS*^{CM}

Enzyme	DAH7PS activity					CM activity		
	K_M^{PEP}	K_M^{E4P}	k_{cat}	$k_{\text{cat}}/K_M^{\text{PEP}}$	$k_{\text{cat}}/K_M^{\text{E4P}}$	K_M	k_{cat}	k_{cat}/K_M
	(μM)	(μM)	(s^{-1})	($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	(μM)	(s^{-1})	($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)
<i>PniDAH7PS</i> ^{WT}	41 ± 4	58 ± 6	16.8 ± 0.5	0.41	0.29	7.6 ± 0.6	1.68 ± 0.06	0.22
<i>PniDAH7PS</i> ^{Δ5AA}	216 ± 20	51 ± 6	7.6 ± 0.3	3.5 × 10 ⁻²	0.15	21 ± 2	1.20 ± 0.02	5.7 × 10 ⁻²
<i>PniDAH7PS</i> ^D	367 ± 35	239 ± 19	1.62 ± 0.11	4.4 × 10 ⁻³	6.8 × 10 ⁻³	NA	NA	NA
<i>PniDAH7PS</i> ^{CM}	NA	NA	NA	NA	NA	84 ± 5	2.01 ± 0.07	2.4 × 10 ⁻²

The apparent Michaelis constants for the DAH7PS activity are apparent values determined using a fixed concentration of the invariant substrate (as described in Section 7.1.13). The uncertainty values represent the standard deviation from triplicate measurements. The parameters of *PniDAH7PS*^{Δ5AA} are highlighted in red.

Compared to the performance of the wild type protein, the CM catalytic activity of *PniDAH7PS*^{Δ5AA} was considerably reduced, as demonstrated by the increase in the K_M value for chorismate from 7.6 ± 0.6 μM to 21 ± 2 μM and a slight decrease in k_{cat} from 1.68 ± 0.06 s^{-1} to 1.20 ± 0.02 s^{-1} (Table 3.3). This impaired enzymatic function could be a direct consequence of the loop truncation, or a result of disrupted interactions between the CM and DAH7PS domains, or a combination of both.

Notably, in addition to the impaired CM function, the DAH7PS activity for *PniDAH7PS*^{Δ5AA} was also found to be significantly weakened. The apparent K_M value for PEP was 216 ± 20 μM , approximately 5 times higher compared to that of *PniDAH7PS*^{WT},

and the k_{cat} value was reduced from $16.8 \pm 0.5 \text{ s}^{-1}$ to $7.6 \pm 0.3 \text{ s}^{-1}$ (Table 3.3). However, no noticeable effect on the apparent K_M value for E4P was detected. The effect on both CM and DAH7PS functions indicates that loop_{C320-F333} may be an important element in maintaining the interdependence between the DAH7PS and CM functions.

Figure 3. 9: The loop truncation for generating *PniDAH7PS*^{Δ5AA} impairs the activating effect of the addition of PEP/Mn²⁺/E4P on the catalytic activity of CM. Reactions were carried out at 35 °C, containing 100 μM chorismate and 3 μg *PniDAH7PS*^{Δ5AA} in 50 mM BTP buffer (pH 7.4), in the absence (ligand free) or presence of PEP/Mn²⁺/E4P (100 μM Mn²⁺, 350 μM PEP, 350 μM E4P). Error bars represent the standard deviation from triplicate measurements.

Moreover, the kinetic assay for investigating the effect of the combination of PEP/Mn²⁺/E4P on CM catalytic activity was repeated using *PniDAH7PS*^{Δ5AA}. The significant boost of CM activity by the addition of PEP/Mn²⁺/E4P was not observed for this truncated variant. This observation further suggests that loop_{C320-F333} is involved in the interaction between the DAH7PS and CM domains (Figure 3.9).

SAXS experiments to investigate the structural effects of loop truncation on *PniDAH7PS*

Structural investigations on *PniDAH7PS*^{Δ5AA} were conducted using SAXS. The SAXS scattering profiles were recorded for apo-*PniDAH7PS*^{Δ5AA} and *PniDAH7PS*^{Δ5AA} in the presence of PEP/Mn²⁺ or PEP/Mn²⁺/E4P (denoted PEP/Mn²⁺-*PniDAH7PS*^{Δ5AA} and PEP/Mn²⁺/E4P-*PniDAH7PS*^{Δ5AA} respectively) and then compared with those of *PniDAH7PS*^{WT}.

The scattering data demonstrated that the loop truncation did not alter the basic spatial organisation as observed in *PniDAH7PS*^{WT} and that apo-*PniDAH7PS*^{Δ5AA} retains a flexible and elongated homodimeric structure (Table 3.4). However, the increased R_g and V_p values of apo-*PniDAH7PS*^{Δ5AA} (Table 3.4), compared to those for apo-*PniDAH7PS*^{WT}, indicate that the loop truncation resulted in a notable expansion of the average size of the *PniDAH7PS*^{WT} molecule (Figure 3.10D). PEP/Mn²⁺-*PniDAH7PS*^{Δ5AA} was shown to adopt a more compact conformation, with a scattering profile corresponding very closely to that of PEP/Mn²⁺-*PniDAH7PS*^{WT} at low q range (Figure 3.10G). Accordingly, the R_g and D_{max} values of PEP/Mn²⁺-*PniDAH7PS*^{Δ5AA}, 39.9 ± 1.0 Å and 140.1 Å respectively, were nearly identical to the parameters determined for PEP/Mn²⁺-*PniDAH7PS*^{WT}, 39.4 ± 0.8 Å and 142.7 Å, (Table 3.4). Whereas a discrepancy between the two scattering profiles was evident at the q range approximately from 0.075 to 0.125 Å⁻¹, which hints at a smaller V_p for *PniDAH7PS*^{Δ5AA} compared to that of *PniDAH7PS*^{WT} (Table 3.4 and Figure 3.10G). Moreover, the Kratky and Porod-Debye transformations suggest that the presence of PEP/Mn²⁺ was not able to attenuate the intrinsic flexibility of the structure of *PniDAH7PS*^{Δ5AA} as effectively as was observed for *PniDAH7PS*^{WT} (Figure 3.10H, I)

Table 3. 4: SAXS parameters for apo-*Pni*DAH7PS^{Δ5AA}, PEP/Mn²⁺-*Pni*DAH7PS^{Δ5AA} and PEP/Mn²⁺/E4P-*Pni*DAH7PS^{Δ5AA} compared to the parameters for *Pni*DAH7PS^{WT}

Enzyme	<i>Pni</i> DAH7PS ^{Δ5AA}			<i>Pni</i> DAH7PS ^{WT}		
Additive	Apo-protein	PEP/Mn ²⁺	PEP/Mn ²⁺ /E4P	Apo-protein	PEP/Mn ²⁺	PEP/Mn ²⁺ /E4P
Guinier Analysis						
R_g (Å)	45.6 ± 0.8	39.9 ± 1.0	42.3 ± 1.7	42.0 ± 1.2	39.4 ± 0.8	35.8 ± 1.0
$I(0)$ (cm ⁻¹)	0.08 ± 0.00	0.07 ± 0.00	0.08 ± 0.00	0.08 ± 0.00	0.07 ± 0.00	0.07 ± 0.00
Correlation coefficient, R^2	0.99	0.99	0.99	0.99	0.99	0.99
Pair Wise Distribution Analysis						
R_g (Å)	46.1 ± 0.6	41.1 ± 0.5	43.4 ± 0.4	44.1 ± 0.2	41.0 ± 0.1	37.7 ± 0.6
$I(0)$ (cm ⁻¹)	0.08 ± 0.00	0.07 ± 0.00	0.08 ± 0.00	0.08 ± 0.00	0.07 ± 0.00	0.07 ± 0.00
D_{max} (Å)	150.5	140.1	146.5	152.3	142.7	140.2
V_p (nm ³)	175	128	143	137	139	121
MM_{porod}						
MM (kDa)	103	75	84	81	82	71
Number of Subunits	2	2	2	2	2	2

V_p : Porod volume; MM_{porod} : molecular mass estimated from Porod volume. The columns coloured in grey present the SAXS data shown previously, provided for a convenient comparison. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

Unexpectedly, the addition of PEP/Mn²⁺/E4P resulted in an opposite structural effect on *Pni*DAH7PS^{Δ5AA} compared to that observed for *Pni*DAH7PS^{WT}, displaying a considerable increase in the size of the protein compared with that for either PEP/Mn²⁺-*Pni*DAH7PS^{Δ5AA} or PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} (Table 3.4 and Figure 3.10J). Both the Kratky and Porod-Debye plots demonstrated that when PEP, Mn²⁺, and E4P were added, the structure of the *Pni*DAH7PS^{Δ5AA} remained flexible (Figure 3.10K, L). However, due to the ongoing DAH7PS reaction, it is not certain whether the SAXS data of PEP/Mn²⁺/E4P-*Pni*DAH7PS^{Δ5AA} and PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} were collected under precisely the same conditions. Therefore, it cannot be confirmed that the loop truncation resulted in the observed differences between the SAXS profiles of *Pni*DAH7PS^{Δ5AA} and *Pni*DAH7PS^{WT} when PEP, Mn²⁺, and E4P were added.

Figure 3. 10: The comparison of the SAXS data obtained for *PniDAH7PS*^{Δ5AA} with that obtained for *PniDAH7PS*^{WT}. Panels **A - C** respectively show the SAXS profiles, Kratky and Porod-Debye plots of apo-*PniDAH7PS*^{Δ5AA} (black), PEP/Mn²⁺-*PniDAH7PS*^{Δ5AA} (Red) and PEP/Mn²⁺/E4P-*PniDAH7PS*^{Δ5AA} (Blue). Panels **D - F** respectively compare the SAXS profiles, Kratky and Porod-Debye plots of apo-*PniDAH7PS*^{Δ5AA} (Red) with that of apo-*PniDAH7PS*^{WT} (Black). Panels **G - I** respectively compare the SAXS profiles, Kratky and Porod-Debye plots of PEP/Mn²⁺-*PniDAH7PS*^{Δ5AA} (Red) with that of PEP/Mn²⁺-*PniDAH7PS*^{WT} (Black). Panels **J - L** respectively compare the SAXS profiles, Kratky and Porod-Debye plots of PEP/Mn²⁺/E4P-*PniDAH7PS*^{Δ5AA} (Red) with that of PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT} (Black).

In general, the structural information obtained from the SAXS experiments for *Pni*DAH7PS^{Δ5AA} implies that the shortened loop_{C320-F333} interfered with the rigidity of the PEP/Mn²⁺ bound protein structure and probably affects the conformational rearrangements associated with certain step(s) of the DAH7PS catalysed reaction. These observations are consistent with the decreased activities of DAH7PS and CM in *Pni*DAH7PS^{Δ5AA}. Combined with the altered kinetic properties of *Pni*DAH7PS^{Δ5AA}, these observations for the protein structure imply an important role of loop_{C320-F333} in the DAH7PS-CM interaction. More structural investigations on *Pni*DAH7PS^{WT} and *Pni*DAH7PS^{Δ5AA} in the presence of some substrate, product, or even transition state analogues might be required to fully characterise the role of loop_{C320-F333} in DAH7PS catalysis. In addition, an equivalent set of studies might be conducted for *Gsp*DAH7PS, which could help to understand the evolutionary association between N-terminal and C-terminal CM linked DAH7PS.

3.3.2 The effect of NaCl on the catalytic activities and structure of *Pni*DAH7PS^{WT}

NaCl was absent in the standard kinetic assays for examining the kinetic parameters of *Pni*DAH7PS^{WT}. When 500 mM NaCl was included in the buffer for these assays, a significant reduction in the activity of either the DAH7PS or CM was observed, compared to the activities examined in the standard assays (Section 2.4.2). Given that salt is a crucial factor affecting protein-protein interactions (139,140), the impairment of *Pni*DAH7PS^{WT} activities in the presence of 500 mM NaCl might be the result of the disruption of the DAH7PS-CM interaction by NaCl. To shed light on the effect of NaCl on the functions of *Pni*DAH7PS^{WT}, an investigation was carried out using kinetic assays, SAXS and ITC assays.

Kinetics characterisation of *Pni*DAH7PS^{WT} in the presence of NaCl

Kinetic assays for both the DAH7PS and CM catalytic activities of *Pni*DAH7PS^{WT} were repeated under the standard kinetic assay conditions, in buffer containing 50 mM BTP

(pH 7.4), in the presence of either 150 mM or 500 mM NaCl.

Table 3. 5: Kinetic parameters for the DAH7PS and CM activities of *PniDAH7PS*^{WT} in the absence or presence of 150 mM or 500 mM NaCl

Enzyme	DAH7PS activity					CM activity		
	K_M^{PEP}	K_M^{E4P}	k_{cat}	k_{cat}/K_M^{PEP}	k_{cat}/K_M^{E4P}	K_M	k_{cat}	k_{cat}/K_M
	(μ M)	(μ M)	(s ⁻¹)	(μ M ⁻¹ s ⁻¹)	(μ M ⁻¹ s ⁻¹)	(μ M)	(s ⁻¹)	(μ M ⁻¹ s ⁻¹)
0 mM NaCl						0 mM NaCl		
<i>PniDAH7PS</i> ^{WT}	41 ± 4	58 ± 6	16.8 ± 0.5	0.41	0.29	7.6 ± 0.6	1.68 ± 0.06	0.22
<i>PniDAH7PS</i> ^D	367 ± 35	239 ± 19	1.62 ± 0.11	4.4 × 10 ⁻³	6.8 × 10 ⁻³	NA	NA	NA
<i>PniDAH7PS</i> ^{CM}	NA	NA	NA	NA	NA	84 ± 5	2.01 ± 0.07	2.4 × 10 ⁻²
150 mM NaCl						150 mM NaCl		
<i>PniDAH7PS</i> ^{WT}	96 ± 7	132 ± 9	13.1 ± 0.6	0.14	9.9 × 10 ⁻²	32.6 ± 2.0	1.50 ± 0.03	4.6 × 10 ⁻²
500 mM NaCl						500 mM NaCl		
<i>PniDAH7PS</i> ^{WT}	NA	NA	NA	NA	NA	81.3 ± 7.5	1.26 ± 0.05	0.22

The apparent Michaelis constants for the DAH7PS activity are apparent values determined using a fixed concentration of the invariant substrate (as described in Section 7.1.13). The uncertainty values represent the standard deviation from triplicate measurements. The parameters determined in the presence of 150 mM and 500 mM NaCl are respectively highlighted in red and blue.

The interfering effect of high concentrations of NaCl on the catalytic functionality of *PniDAH7PS*^{WT} is evident (Table 3.5). Compared with the catalytic performance observed in the absence of NaCl, both the CM and DAH7PS catalytic activities of *PniDAH7PS*^{WT} were notably reduced in the presence of 150 mM NaCl. Compared to the kinetic parameters of *PniDAH7PS*^{WT} in the absence of NaCl, the apparent K_M values for PEP and E4P both increased by more than two-fold when 150 mM NaCl was present; the k_{cat} reduced from 16.8 ± 0.5 s⁻¹ to 13.1 ± 0.6 s⁻¹. Similarly, for the CM activity, an obvious increase in K_M for chorismate was observed, with a rise from 7.6 ± 0.6 to 32.6 ± 2.0 μ M. However, no substantial effect on the k_{cat} value for CM activity was observed in the presence of 150 mM NaCl. When the NaCl concentration was increased to 500 mM, the DAH7PS activity was almost abolished. However, the CM moiety still possessed detectable activity, although this activity was also severely attenuated. Compared to the kinetic parameters for CM recorded in the absence of NaCl, the K_M

value for chorismate increased from $7.6 \pm 0.6 \mu\text{M}$ to $81.3 \pm 7.5 \mu\text{M}$ in the presence of 500 mM of NaCl, whereas the k_{cat} value slightly dropped from $1.68 \pm 0.06 \text{ s}^{-1}$ to $1.26 \pm 0.05 \text{ s}$.

Examination of the structural effect of NaCl on *PniDAH7PS*^{WT} using SAXS

All SAXS data for both *PniDAH7PS*^{WT} and *PniDAH7PS* ^{$\Delta 5\text{AA}$} were collected using the standard SAXS buffer containing 150 mM NaCl. To determine how varying concentrations of NaCl affect the conformation of *PniDAH7PS*^{WT}, SAXS experiments were performed for *PniDAH7PS*^{WT} in the absence or presence of PEP/Mn²⁺ using the SAXS buffer containing either 0 mM or 500 mM NaCl.

Table 3. 6: SAXS parameters for apo-*PniDAH7PS*^{WT} and PEP/Mn²⁺-*PniDAH7PS*^{WT} in the presence of varying concentrations of NaCl

Enzyme	0 mM NaCl		150 mM NaCl		500 mM NaCl	
Additive	apo-protein	PEP/Mn ²⁺	apo-protein	PEP/Mn ²⁺	apo-protein	PEP/Mn ²⁺
Guinier Analysis						
R_g (Å)	33.1 ± 1.1	38.0 ± 0.6	42.0 ± 1.2	39.4 ± 0.8	37.8 ± 0.7	37.0 ± 0.7
$I(0)$ (cm ⁻¹)	0.07 ± 0.00	0.07 ± 0.00	0.08 ± 0.00	0.07 ± 0.00	0.07 ± 0.00	0.07 ± 0.00
Correlation coefficient, R^2	0.99	0.99	0.99	0.99	0.99	0.99
Pair Wise Distribution Analysis						
R_g (Å)	33.0 ± 0.4	38.8 ± 0.3	44.1 ± 0.2	41.0 ± 0.1	38.6 ± 0.2	37.9 ± 0.2
$I(0)$ (cm ⁻¹)	0.07 ± 0.00	0.07 ± 0.00	0.08 ± 0.00	0.07 ± 0.00	0.07 ± 0.00	0.07 ± 0.00
D_{max} (Å)	113.6	136.8	152.3	142.7	130.8	128.0
V_p (nm ³)	118	133	137	139	104	108
<i>MM</i>_{porod}						
MM (kDa)	69	78	81	81	61	63
Number of Subunits	2	2	2	2	2	2

V_p : Porod volume; *MM*_{porod}: molecular mass estimated from Porod volume. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

Figure 3. 11: The analysis of the SAXS profiles of apo-*PniDAH7PS*^{WT} and PEP/Mn²⁺-*PniDAH7PS*^{WT} in the absence or presence of NaCl. Panels **A - C** respectively show the SAXS profiles, Kratky and Porod-Debye plots of apo-*PniDAH7PS*^{WT} in the presence of 0 mM (black), 150 mM (red) or 500 mM (blue) NaCl. Panels **D - F** respectively compare the SAXS profiles, Kratky and Porod-Debye plots of PEP/Mn²⁺-*PniDAH7PS*^{WT} in the presence of 0 mM (black), 150 mM (red) or 500 mM (blue) NaCl. Panels **G - I** respectively compare the SAXS profiles, Kratky and Porod-Debye plots of apo-*PniDAH7PS*^{WT} (black) and PEP/Mn²⁺-*PniDAH7PS*^{WT} (red) when NaCl is absent. Panels **J - L** respectively compare the SAXS profiles, Kratky and Porod-Debye plots of apo-*PniDAH7PS*^{WT} (black) and PEP/Mn²⁺-*PniDAH7PS*^{WT} (red) in the presence of 150 mM NaCl. Panels **M - O** respectively compare the SAXS profiles, Kratky and Porod-Debye plots of apo-*PniDAH7PS*^{WT} (black) and PEP/Mn²⁺-*PniDAH7PS*^{WT} (red) in the presence of 500 mM NaCl.

A comparison between the SAXS profiles of apo-*Pni*DAH7PS^{WT} determined in the absence or presence of 150 mM NaCl revealed a substantial change in the average size of the protein molecule in the absence of NaCl compared to that in the presence of NaCl. This was evident as a decrease in R_g , D_{max} and V_p from 42.0 ± 1.2 Å, 152.3 Å and 137 nm³ to 33.1 ± 1.1 Å, 113.6 Å and 118 nm³ respectively (Table 3.6). Interestingly, in the presence of 500 mM NaCl, apo-*Pni*DAH7PS^{WT} appears to be more compact compared to that in the presence of 150 mM NaCl (Table 3.6). Apo-*Pni*DAH7PS^{WT} was well-folded and rigid in the absence of NaCl, as demonstrated by the Kratky and Porod-Debye plots (Figure 3.11B,C). In contrast, in the presence of either 150 mM or 500 mM NaCl, the protein structure became significantly more flexible (Figure 3.11B,C).

Subsequently, the effect of NaCl on the conformation of *Pni*DAH7PS^{WT} in the presence of PEP/Mn²⁺ binding was examined. A conformational change of *Pni*DAH7PS^{WT} due to the addition of PEP/Mn²⁺ was observed in the absence or presence of 150 mM NaCl. Although the scattering profiles of PEP/Mn²⁺-*Pni*DAH7PS^{WT} appeared to be highly similar for these two NaCl concentrations (Table 3.6 and Figure 3.11D-F), the structural changes of *Pni*DAH7PS^{WT} that resulted from the addition of PEP/Mn²⁺ showed different tendencies. In contrast to the effect of PEP/Mn²⁺, which compresses the structure of *Pni*DAH7PS^{WT} in the presence of 150 mM NaCl (Table 3.6 and Figure 3.11J-L), the addition of PEP/Mn²⁺ disturbed the compactness of *Pni*DAH7PS^{WT} when NaCl was absent (Table 3.6 and Figure 3.11G-I), increasing the R_g , D_{max} and V_p respectively from 33.1 ± 1.1 Å, 113.6 Å and 118 nm³ to 38.0 ± 0.6 Å, 136.8 Å and 133 nm³. Further to this, the presence of 500 mM NaCl appeared to completely abolish the effect of PEP/Mn²⁺ on modulating the conformation of *Pni*DAH7PS^{WT}, as evidenced by the highly similar scattering data of apo-*Pni*DAH7PS^{WT} and PEP/Mn²⁺-*Pni*DAH7PS^{WT} in the presence of 500 mM NaCl (Figure 3.11M-O). The interfering effect of NaCl on the conformation of *Pni*DAH7PS^{WT} is in line with the impaired catalytic activities that were detected at high NaCl concentrations (Table 3.5). This observation supports the tight correlation between the DAH7PS-CM interaction and both the DAH7PS and CM

enzyme reactions.

Investigating the effect of NaCl on PEP binding using ITC

Although the structural investigation using SAXS revealed that the presence of NaCl likely interferes with the conformation of *PniDAH7PS*^{WT}, this data raised several questions: Why does the apo-protein display a compact and rigid structure in the absence of NaCl (Table 3.6 and Figure 3.11G-I)? How does PEP/Mn²⁺ affect the well-organised *PniDAH7PS*^{WT} structure? Moreover, how does a high concentration NaCl abolish the conformational changes of *PniDAH7PS*^{WT} due to the addition of PEP/Mn²⁺? To gain an in-depth understanding of the effect of NaCl on the structure and function of *PniDAH7PS*^{WT}, ITC measurements were carried out to obtain thermodynamic information to complement the previously described kinetic and SAXS data.

The ITC parameters for PEP binding were obtained by titrating PEP into a solution of *PniDAH7PS*^{WT} containing MnSO₄, in the presence of 0 mM or 500 mM NaCl. These parameters were compared to the ITC parameters obtained in the presence of 150 mM NaCl, as described above (Figure 3.6C,D). The isotherms indicate that the binding of PEP to *PniDAH7PS*^{WT} became weaker as the concentration of NaCl was increased. Further to this, PEP binding to *PniDAH7PS*^{WT} was not completely disrupted by the presence of 500 mM NaCl (Figure 3.12A-C). The binding isotherms recorded in the absence of NaCl were fitted with an independent binding model, as was the case for the data obtained in the presence of 150 mM NaCl (Figure 3.12A,B). The calculated K_D values for PEP binding to *PniDAH7PS*^{WT} decreased from 51.8 ± 6.1 μ M to 38.3 ± 6.4 μ M following the removal of 150 mM NaCl (Figure 3.12D). The thermodynamic properties of PEP binding were not fundamentally altered by the absence of NaCl, which also showed an entropically driven and enthalpically unfavourable process (Figure 3.12E).

Figure 3. 12: ITC assays showing the thermodynamics underlying the binding of PEP to *PniDAH7PS* in absence or presence of NaCl. (A, B, C) Raw data from ITC experiments involving the titration of 1.5 mM PEP into a cell containing 65 μM *PniDAH7PS*^{WT} and 500 μM MnSO_4 in the presence of 0 mM, 150 mM or 500 mM NaCl respectively, illustrating the fit to the independent binding model. (D) Comparison between K_D for PEP in the presence of 0 mM (grey) and 150 mM (red) NaCl respectively. (E) Comparison between the thermodynamic parameters for the binding of PEP to *PniDAH7PS*^{WT} in the presence of 0 mM (grey) and 150 mM (red) NaCl respectively.

3.4 Summary and Discussion

Based on the investigations into the activities of the truncated individual DAH7PS and CM enzymes, described in the previous chapter, the functional interdependence between the DAH7PS and CM domains was proposed to be a distinctive feature of *PniDAH7PS^{WT}*. This speculation was supported by the boost in the CM activity of *PniDAH7PS^{WT}* promoted by the addition of PEP/Mn²⁺/E4P.

Dynamic DAH7PS-CM hetero-interaction provides the structural basis for the functions of both DAH7PS and CM

Besides the conformation of apo-*PniDAH7PS^{WT}*, two other distinct conformations of *PniDAH7PS^{WT}* were observed when PEP/Mn²⁺ and PEP/Mn²⁺/E4P were respectively added. Due to the DAH7PS reaction proceeding, the SAXS data of *PniDAH7PS^{WT}* determined for the condition containing PEP/Mn²⁺/E4P, likely represents a mixture of ligand-bound species and any measurements were done on a population of mixed conformations. In spite of this uncertainty, the higher compactness of PEP/Mn²⁺/E4P-*PniDAH7PS^{WT}*, indicated by the SAXS data, compared to those of apo-*PniDAH7PS^{WT}* and PEP/Mn²⁺-*PniDAH7PS^{WT}*, implies some other distinct conformations of *PniDAH7PS^{WT}* are accessible under these conditions.

These observations suggest that multiple conformations might be adopted by *PniDAH7PS^{WT}* during the DAH7PS catalytic cycle. Further to this, a hypothesis predicting a series of DAH7PS-CM hetero-interactions along the DAH7PS catalytic cycle is proposed, in which catalytic steps and hetero-domain rearrangements occur in concert to facilitate the catalytic reaction. This hypothesis is further supported by the investigations on the *PniDAH7PS^{Δ5AA}* variant and the effect of NaCl on *PniDAH7PS^{WT}*, in which the loop truncation and high NaCl concentration disrupted the hetero-domain interactions and enzymatic activities simultaneously. In addition, the loop_{C320-F333} appears to be likely involved in PEP/Mn²⁺ binding, since a disruption to the molecular

rigidity was observed for PEP/Mn²⁺-*Pni*DAH7PS^{Δ5AA}, compared to that of PEP/Mn²⁺-*Pni*DAH7PS^{WT} (Figure 3.10). In the rigid body model derived from the SAXS data of apo-*Pni*DAH7PS^{WT}, PEP/Mn²⁺-*Pni*DAH7PS^{WT} or PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT}, the loop_{C320-F333} is predicted to be in close proximity to loop β6-α6 of the DAH7PS domain involved in the PEP binding (Figure 3.13). This may correlate with the significantly increased K_M^{PEP} of *Pni*DAH7PS^{Δ5AA} (Table 3.3). This prediction from the models alludes to a possibility that some of the conformational changes of *Pni*DAH7PS^{WT} that occur during DAH7PS catalysis are consequences of the propagation of structural rearrangements at the DAH7PS active site via the DAH7PS-CM domain interactions. As discussed in Chapter 2, enzymes are flexible entities and exhibit substantial structural dynamics and may adopt varied conformations to fulfil all steps involved in catalytic reactions (88-90). Therefore, the correlation between the overall structural variations of *Pni*DAH7PS^{WT} and the DAH7PS catalytic function could demonstrate this concept.

All SAXS and biochemical data presented in this work provide the basic information for proposing this model to describe the correlation between the rearrangements of the DAH7PS-CM interaction and the DAH7PS catalytic cycle in *Pni*DAH7PS^{WT}. To demonstrate this model and characterise the hetero-domain interactions in detail, more structural investigations have to be carried out in the future, perhaps with the assistance of some substrate or product analogues or even a transition state analogue.

Figure 3. 13: Rigid body models of *PniDAH7PS*^{WT} show the hetero-interaction between the DAH7PS (blue) and CM (red) domains respectively in the absence of any ligand (A) and in the presence of either PEP/Mn²⁺ (B) or PEP/Mn²⁺/E4P (C). The red ribbon diagram indicates loop_{C320-F333} and the blue ribbon diagram indicates loop $\beta 6$ - $\alpha 6$. The yellow spheres represent the active site of DAH7PS.

A desolvation process may occur at the initial stage of the DAH7PS catalysed reaction

The ITC data demonstrate that the binding of PEP to *PniDAH7PS*^{WT} is an endothermic and entropic driven process in the presence of either 0 mM or 150 mM NaCl (Figure 3.12E). Generally, ligand binding is associated with the loss of translation, rotation and vibration motions of the ligand and the protein residues in the ligand binding site. Hence it is an entropically unfavourable and enthalpically favourable process. In spite of this, contradictory thermodynamic properties were also observed for the binding of the mouse pheromone 2-sec-butyl-4,5-dihydrothiazole to a mouse major urinary protein, which shows an entropically dominated stabilisation of the protein-ligand complex, accompanied by an increased flexibility of protein backbone (141). However, PEP binding to *PniDAH7PS*^{WT} appears to be a different case according to the analysis of the SAXS data for *PniDAH7PS*^{WT}. The Kratky and Porod Debye transformations of the SAXS data show that PEP/Mn²⁺ binding does not change or even reduces the flexibility

of the overall structure of *Pni*DAH7PS^{WT} in the presence of 0 mM and 150 mM NaCl (Figure 3.11G-L).

In order to comprehensively analyse the thermodynamic properties underpinning PEP binding, the solvent must be taken into account. If water molecules bind to the hydrophilic surface of *Pni*DAH7PS^{WT} and form a well-defined hydration shell, the desolvation process triggered by PEP binding would lead to a considerable reduction in the solvent enthalpy by disrupting hydrogen bonds and van de Waals interactions between the protein and water. This would result in a significant positive solvent entropic change via releasing the water molecules from hydration shell (142-144). Therefore, for entropically driven ligand binding, the thermodynamic changes of the solvent exert a dominant effect in the protein solution system and the gain of solvent entropy may be high enough to compensate for the penalty of solvent enthalpy. This theory could explain the thermodynamic characteristics of PEP binding to *Pni*DAH7PS^{WT} and suggests that PEP and water compete for association to *Pni*DAH7PS^{WT}. Moreover, the well-defined global structure of apo-*Pni*DAH7PS^{WT} in the absence of NaCl, and the notable re-organisation of protein structure following the addition of PEP/Mn²⁺, agrees with the speculated protein hydration and competitive relationship between PEP/Mn²⁺ and water (Figure 3.11G-I). Therefore, it is hypothesised that the conformational change upon PEP/Mn²⁺ binding is driven by a desolvation process occurring on the *Pni*DAH7PS^{WT} molecule.

DAH7PS domain mediated allosteric activation of CM activity

According to the definition of allostery, the up-regulation of the CM catalytic activity following the addition of PEP/Mn²⁺/E4P is a typical allosteric activation. Thus, simultaneously with catalysing the formation of DAH7P, the DAH7PS domain also functions as a regulatory domain to activate the CM function. This phenomenon, where catalysis and allostery are merged, seems to be rarely reported. Nevertheless, it supports the concept that catalysis and allostery take place through a common

mechanism (63). By taking advantage of the dual-functionality of the DAH7PS moiety, *Pni*DAH7PS is able to turn up its CM activity by sensing the synthesis of DAH7P and in turn to convert chorismate to prephenate and lead to the end products of the biosynthetic pathway more efficiently. Thereby, a potential feedforward control of CM function, mediated by DAH7PS, may be operating in *P. nigresens*. Cooperation between the feedforward activation of CM and the feedback inhibition on DAH7PS activity by prephenate creates a regulation loop allowing for the more precise modulation of metabolic flow through the shikimate pathway (Figure 3.14).

Figure 3. 14: Schematic representation of the proposed regulation loop mediated by *Pni*DAH7PS in the shikimate pathway. The blue point arrows represent the DAH7PS reaction (highlighted by blue dash rectangle) induced feedforward activation of CM activity, promoting the conversion of chorismate to prephenate; the red blunt arrows illustrate the feedback inhibition of the DAH7PS activity by prephenate (highlighted by red dash oval) via the CM domain.

Chapter 4 *Allosteric inhibition of DAH7PS by prephenate*

4.1 Introduction

The investigations of *Tma*DAH7PS and *Gsp*DAH7PS have demonstrated that the relative motions of the regulatory domains and the DAH7PS barrels disrupt substrate-binding and underpin allosteric inhibition of DAH7PS (20,96,109). The preliminary investigation on the allosteric inhibition of *Pni*DAH7PS^{WT} also indicates that a spatial re-organisation between the DAH7PS and CM domains takes place in the presence of prephenate (Sections 2.5.2 and 2.7). However, this conformational change is suspected to deliver the inhibitory effect of prephenate via interrupting the functional-buttressing interaction between the DAH7PS and CM domains, which appears to be basically different from the allosteric mechanism observed in either *Tma*DAH7PS or *Gsp*DAH7PS.

The work presented in Chapter 3 reveals that different conformations are adopted by *Pni*DAH7PS^{WT} when different ligand concentrations are, suggesting that the dynamics of the overall structure of *Pni*DAH7PS^{WT} correlate with the DAH7PS catalytic cycle. From this this new finding, the preliminary hypothesis of the mechanism underlying the allostery of *Pni*DAH7PS^{WT} by prephenate needs to be reconsidered. How does the binding of prephenate influence the dynamic structure of *Pni*DAH7PS^{WT} and ultimately exert its effect on the DAH7PS function? Clarifying the structural effects of prephenate on the conformational change of *Pni*DAH7PS^{WT} caused by the addition of either PEP/Mn²⁺ or PEP/Mn²⁺/E4P would help to characterise the allosteric inhibition of *Pni*DAH7PS^{WT} more comprehensively. Therefore, the studies described in this chapter aim to refine a model to explain the mechanism underpinning the prephenate-mediated allostery of *Pni*DAH7PS^{WT}. These findings allow a clearer comparison of the allosteric mechanisms of the N-terminal and C-terminal CM linked DAH7PS enzymes.

4.2 The structural effect of prephenate on *Pni*DAH7PS^{WT}

The scattering profile of prephenate-bound *Pni*DAH7PS^{WT} (PA-*Pni*DAH7PS^{WT}) was compared with the scattering profiles of PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT}. It is evident that in the presence of prephenate, the enzyme adopts a much more compact and rigid conformation compared to the other two conformations (Table 4.1).

Table 4. 1: Comparison of the SAXS parameters of PA-*Pni*DAH7PS^{WT}, PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT}

Additive	Prephenate	PEP/Mn ²⁺	PEP/Mn ²⁺ /E4P
Guinier Analysis			
R_g (Å)	32.7 ± 0.3	40.1 ± 0.7	35.8 ± 1.0
$I(0)$ (cm ⁻¹)	0.06 ± 0.00	0.07 ± 0.00	0.07 ± 0.00
Correlation coefficient, R^2	0.99	0.99	0.99
Pair Wise Distribution Analysis			
R_g (Å)	32.9 ± 0.2	41.0 ± 0.3	37.7 ± 0.6
$I(0)$ (cm ⁻¹)	0.06 ± 0.00	0.07 ± 0.00	0.07 ± 0.00
D_{max} (Å)	105.9	144.3	140.2
V_p (nm ⁻³)	115	146	121
MM_{porod}			
MM (kDa)	68	86	71
Number of Subunits	2	2	2

V_p : Porod volume; MM_{porod} : molecular mass estimated from Porod volume. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

Figure 4. 1: Comparison of SAXS profiles of *PniDAH7PS^{WT}* in the presence of prephenate (black), PEP/Mn²⁺ (red) and PEP/Mn²⁺/E4P (blue). (A) SAXS profiles (log $I(q)$ versus q). (B) Guinier plot ($\ln I(q)$ versus q^2). (C) Kratky plot ($q^2 I(q)$ versus q). (D) Porod-Debye plot ($q^4 I(q)$ versus q^4) limited to the range of the SAXS data for which Guinier linearity is observed.

The effect of prephenate on PEP/Mn²⁺-*PniDAH7PS^{WT}* (PA-PEP/Mn²⁺-*PniDAH7PS^{WT}*) was then examined by performing SAXS experiments in the presence of 350 μM prephenate in combination with either 100 μM or 500 μM PEP and MnSO₄ each, respectively. The scattering profiles recorded under both conditions were almost identical, without a notable variation caused by the difference in the concentration of PEP/Mn²⁺. Intriguingly, the structure of *PniDAH7PS^{WT}* in the presence of the combination of prephenate and PEP/Mn²⁺ was more compact and more rigid than that in the presence of prephenate alone, reducing the R_g , D_{max} and V_p values from 32.7 ± 0.3 Å, ~ 105.9 Å and ~ 115 nm³ to ~ 30 Å, ~ 86 Å and ~ 109 nm³, respectively (Table 4.2 and Figure 4.2C,D).

Table 4. 2: Comparison of SAXS parameters of PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} with that of PA-*Pni*DAH7PS^{WT} and PEP/Mn²⁺-*Pni*DAH7PS^{WT}

	350 μ M PA & 100 μ M PEP/Mn ²⁺	350 μ M PA & 500 μ M PEP/Mn ²⁺	350 μ M PA	350 μ M PEP/Mn ²⁺
Guinier Analysis				
R_g (Å)	29.6 $\pm$ 0.2	29.7 $\pm$ 1.1	32.7 $\pm$ 0.3	40.1 $\pm$ 0.7
$I(0)$ (cm ⁻¹)	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00	0.07 $\pm$ 0.00
Correlation coefficient, R^2	0.99	0.99	0.99	0.99
Pair Wise Distribution Analysis				
R_g (Å)	29.5 $\pm$ 0.4	29.3 $\pm$ 0.5	32.9 $\pm$ 0.2	41.0 $\pm$ 0.3
$I(0)$ (cm ⁻¹)	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00	0.07 $\pm$ 0.00
D_{max} (Å)	87.0	86.0	105.9	144.3
V_p (nm ³)	109	109	1159	146
MM_{porod}				
MM (kDa)	64	64	68	86
Number of Subunits	2	2	2	2

V_p : Porod volume; MM_{porod} : molecular mass estimated from Porod volume. The columns coloured in grey present the SAXS data shown previously, provided for convenient comparison. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

Figure 4. 2: Comparison of SAXS profiles for *PniDAH7PS*^{WT} in the presence of 350 μM prephenate respectively combined with 0 μM (black), 100 μM (red) and 500 μM (blue) of both PEP and MnSO₄. (A) SAXS profiles (log $I(q)$ versus q). (B) Guinier plots. (C) Kratky plots. (D) Porod-Debye plots.

To probe the structural effect of prephenate on *PniDAH7PS*^{WT} in the presence of a full complement of the DAH7PS substrates, SAXS analyses were carried out in the standard SEC-SAXS buffer with the addition of 350 μM PEP, 100 μM MnSO₄, 350 μM E4P and 350 μM prephenate (PA-PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT}) (Section 7.1.18). Compared with the SAXS data of PA-*PniDAH7PS*^{WT} and PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT}, the scattering profile of PA-PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT} implied that the enzyme adopted a more compact structure (Figure 4.3A,B), as indicated by the reduction of R_g , D_{max} and V_p from 32.7 ± 0.3 Å, 105.9 Å and 115 nm³ to 30.7 ± 0.5 Å, 93.4 Å and 108 nm³ respectively (Table 4.3). The flexibility observed for PEP/Mn²⁺/E4P-*PniDAH7PS*^{WT} was greatly restricted following the addition of prephenate, as demonstrated by the Kratky

and Porod-Debye plots (Figure 4.3C,D). A comparison between the scattering data of PA-PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} and PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} (Table 4.3 and Figure 4.3E–H) shows that *Pni*DAH7PS^{WT} adopts a similar conformation under the two conditions.

Table 4. 3: Comparison of SAXS parameters for PA-PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} with that of PA-*Pni*DAH7PS^{WT}, PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} and PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT}

Additive	350 μ M PA & PEP/Mn ²⁺ /E4P	350 μ M PA	PEP/Mn ²⁺ /E4P	350 μ M PA & 100 μ M PEP/Mn ²⁺	350 μ M PA & 500 μ M PEP/Mn ²⁺
<i>Guinier Analysis</i>					
R_g (Å)	30.7 $\pm$ 0.5	32.7 $\pm$ 0.3	35.8 $\pm$ 1.0	29.6 $\pm$ 0.2	29.7 $\pm$ 1.1
$I(0)$ (cm ⁻¹)	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00	0.07 $\pm$ 0.00	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00
Correlation coefficient, R^2	0.99	0.99	0.99	0.99	0.99
<i>Pair Wise Distribution Analysis</i>					
R_g (Å)	30.5 $\pm$ 0.1	32.9 $\pm$ 0.2	37.7 $\pm$ 0.6	29.5 $\pm$ 0.4	29.3 $\pm$ 0.5
$I(0)$ (cm ⁻¹)	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00	0.07 $\pm$ 0.00	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00
D_{max} (Å)	93.4	105.9	140.2	87.0	86.2
V_p (nm ³)	108	115	121	109	109
<i>MM_{porod}</i>					
MM (kDa)	63	68	71	64	64
Number of Subunits	2	2	2	2	2

V_p : Porod volume; MM_{porod} : molecular mass estimated from Porod volume. The columns coloured in grey present the SAXS data shown previously, to provide a convenient comparison. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

Figure 4. 3: The analysis of SAXS profiles shows the combined effect of prephenate and PEP/Mn²⁺/E4P on the structure of *PniDAH7PS^{WT}*. (A–D) individually show SAXS profiles, Guinier plots, Kratky plots and Porod-Debye plots respectively for *PniDAH7PS^{WT}* upon the combined effect of prephenate and the PEP/Mn²⁺/E4P (black), PA-*PniDAH7PS^{WT}* (red) and PEP/Mn²⁺/E4P-*PniDAH7PS^{WT}* (blue). (E–H) A comparison between the SAXS profiles, Guinier plots, Kratky plots and Porod-Debye plots for *PniDAH7PS^{WT}* affected by prephenate combined with the PEP/Mn²⁺/E4P (black) and PEP/Mn²⁺ (green).

4.3 Rigid body models of prephenate-inhibited *Pni*DAH7PS^{WT}

Following the procedures and set-up of rigid body modelling, as described in section 3.2.2, models of PA-*Pni*DAH7PS^{WT}, PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PA-PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} were generated using BUNCH (132). The resultant models exhibited reasonable fits ($\chi^2 = 0.6$, 1.5 and 0.6 respectively) to the corresponding experimental SAXS data.

Figure 4. 4: Rigid body modelling for PA-*Pni*DAH7PS^{WT}, PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PA-PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} and fits to the corresponding experimental SAXS profiles. DAH7PS and CM domains are coloured blue and red respectively. **(A)** Rigid body model of PA-*Pni*DAH7PS^{WT} and its fit to the experimental SAXS data ($\chi^2 = 0.6$). **(B)** Rigid body model of PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} and its fit to the experimental SAXS data ($\chi^2 = 1.5$). **(C)** Rigid body model of PA-PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} and its fit to the experimental SAXS data ($\chi^2 = 0.6$).

The models of PA-*Pni*DAH7PS^{WT}, PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PA-PEP/Mn²⁺/E4P-

*Pni*DAH7PS^{WT} display very similar spatial organisations, where the two DAH7PS barrels are tightly associated and loop α 2- α 3 of the CM (loop_{C320-F333}, described in Section 3.3.1) is in close proximity to the N-terminal random coil of DAH7PS barrel in the model of PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} (Figure 4.4B,C).

4.4 The effect of prephenate on *Pni*DAH7PS^{Δ5AA}: Presenting more details for explaining the mechanism of prephenate mediated inhibition

The truncation of Loop_{C320-F333} was previously shown to be able to influence the DAH7PS catalysed reaction (Section 3.3.1). To shed light on the role of this loop in the prephenate mediated allosteric regulation, the response of *Pni*DAH7PS^{Δ5AA} to prephenate was examined.

Kinetic assay

The inhibitory effect of prephenate on the DAH7PS activity of *Pni*DAH7PS^{Δ5AA} was observed and compared to that of *Pni*DAH7PS^{WT}. As is the case for the wild-type enzyme, *Pni*DAH7PS^{Δ5AA} is remarkably sensitive to prephenate with a calculated relative IC₅₀ of 3.3 μ M; similar to that shown for the inhibition of *Pni*DAH7PS (4.7 μ M) (Figure 4.5). However, the percentage of inhibition decreased substantially, from ~74% in *Pni*DAH7PS^{WT} to ~43% in *Pni*DAH7PS^{Δ5AA} (Figure 4.5).

Figure 4. 5: Response of the DAH7PS activity of either *PniDAH7PS*^{WT} (black) or *PniDAH7PS*^{Δ5AA} (red) to increasing concentrations of prephenate. Error bars represent the standard deviation from triplicate measurements.

SAXS examination

SAXS experiments were carried out to examine the structural changes related to the allosteric inhibition of *PniDAH7PS*^{Δ5AA} by prephenate. The comparison between the recorded SAXS data for apo-*PniDAH7PS*^{Δ5AA} and *PniDAH7PS*^{Δ5AA} in the presence of prephenate (PA-*PniDAH7PS*^{Δ5AA}) showed an evident conformational change upon the binding of prephenate. In the presence of prephenate, *PniDAH7PS*^{Δ5AA} was compressed and rigidified (Table 4.4 and Figure 4.6A–D). However, compared to PA-*PniDAH7PS*^{WT}, PA-*PniDAH7PS*^{Δ5AA} exhibited a much less compact and rigid conformation (Table 4.4 and Figure 4.6E–H). It appears that the removal of loop_{C320-F333} alters the conformational response of the protein to prephenate.

A further comparison between the SAXS data of PA-*PniDAH7PS*^{Δ5AA} and PEP/Mn²⁺-*PniDAH7PS*^{Δ5AA} revealed that these two enzyme-ligand complexes adopted very similar conformations, as evidenced by their similar SAXS profiles (Figure 4.6A–D) and parameters (R_g , D_{max} and V_p) (Table 4.4). Based on these observations, it is speculated

that, loop_{C320-F333} has a role in delivering the compact protein conformation associated with prephenate binding. This is to some extent supported by the close proximity between loop_{C320-F333} and the N-terminal random coil of the DAH7PS domain implied in the rigid body models of PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} (Figure 4.4B). It is speculated that loop_{C320-F333} may contribute to the formation of the rigid structures of prephenate bound *Pni*DAH7PS^{WT} via interacting with the N-terminal segment of the DAH7PS domain.

Table 4. 4: Comparison of the SAXS parameters of apo-*Pni*DAH7PS^{Δ5AA}, PEP/Mn²⁺-*Pni*DAH7PS^{Δ5AA}, PA-*Pni*DAH7PS^{Δ5AA} and PA-*Pni*DAH7PS^{WT}

Enzyme	<i>Pni</i> DAH7PS ^{Δ5AA}			<i>Pni</i> DAH7PS ^{WT}
Additives	Prephenate	Apo-protein	PEP/Mn ²⁺	Prephenate
Guinier Analysis				
R_g (Å)	39.7 ± 0.6	45.6 ± 0.8	39.9 ± 1.2	32.7 ± 0.3
$I(0)$ (cm ⁻¹)	0.07 ± 0.00	0.08 ± 0.00	0.07 ± 0.00	0.06 ± 0.00
Correlation coefficient, R^2	0.99	0.99	0.99	0.99
Pair Wise Distribution Analysis				
R_g (Å)	40.4 ± 0.4	46.1 ± 0.6	41.1 ± 0.5	32.9 ± 0.2
$I(0)$ (cm ⁻¹)	0.07 ± 0.00	0.08 ± 0.00	0.07 ± 0.00	0.06 ± 0.00
D_{max} (Å)	136.4	150.5	140.1	105.9
V_p (nm ⁻³)	148	175	128	115
MM_{porod}				
MM (kDa)	87	103	75	68
Number of Subunits	2	2	2	2

V_p : Porod volume; MM_{porod} : molecular mass estimated from Porod volume. The columns coloured in grey present the SAXS data shown previously, provided for a convenient comparison. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

4.5 NaCl

As a dramatic influence by NaCl on DAH7PS catalysis was observed for *Pni*DAH7PS^{WT} (section 3.3.2), kinetic and structural characterisations were performed to inspect whether the allosteric response of *Pni*DAH7PS, caused by prephenate, was affected by varying NaCl concentrations.

Kinetics

Feedback inhibition of the DAH7PS activity of *Pni*DAH7PS^{WT} by prephenate was tested in the presence of 150 mM NaCl and compared to the inhibition in the absence of NaCl. Due to the significant inhibition of 500 mM NaCl on DAH7PS activity of *Pni*DAH7PS^{WT} (Table 3.5), an examination of the inhibitory effect of prephenate on the DAH7PS activity in the presence of 500 mM NaCl was not performed.

Figure 4. 7: Response of the DAH7PS activity of *Pni*DAH7PS^{WT} to increasing concentrations of prephenate in the absence (black) or presence of 150 mM NaCl (red). Error bars represent the standard deviation from triplicate measurements.

*Pni*DAH7PS^{WT} retained its high sensitivity to prephenate in the presence of 150 μM NaCl; although the calculated relative IC₅₀ of 7.5 μM was higher than that observed in

the absence of NaCl (4.7 μ M) (Figure 4.7). The extent of the inhibition of DAH7PS activity in the presence or absence of 150 mM NaCl was similar, exhibiting ~74% and 77% inhibition respectively (Figure 4.7).

SAXS

The structural effect of NaCl on the allosteric inhibition by prephenate was characterised by SAXS. Scattering profiles for PA-*Pni*DAH7PS^{WT} were collected in the presence of 0 mM, 150 mM or 500 mM NaCl and the three profiles were almost identical (Table 4.5 and Figure 4.8), suggesting that NaCl had no effect on the conformation of PA-*Pni*DAH7PS^{WT}.

Table 4. 5: SAXS parameters for PA-*Pni*DAH7PS^{WT} in the presence of 0 mM, 150 mM or 500

mM NaCl			
Enzyme	PA- <i>Pni</i> DAH7PS ^{WT}		
NaCl concentration (mM)	0	150	500
Guinier Analysis			
R_g (Å)	31.7 $\pm$ 0.3	32.7 $\pm$ 0.3	31.5 $\pm$ 0.9
$I(0)$ (cm ⁻¹)	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00
Correlation coefficient, R^2	0.99	0.99	0.99
Pair Wise Distribution Analysis			
R_g (Å)	31.9 $\pm$ 0.2	32.9 $\pm$ 0.2	31.8 $\pm$ 0.2
$I(0)$ (cm ⁻¹)	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00
D_{max} (Å)	106.3	105.9	104.1
V_p (nm ³)	108	115	111
MM_{porod}			
MM (kDa)	64	68	65
Number of Subunits	2	2	2

V_p : Porod volume; MM_{porod} : molecular mass estimated from Porod volume. The column coloured in grey present the SAXS data shown previously, provided for convenient comparison. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

Figure 4. 8: Comparison of the SAXS profiles for PA-*PniDAH7PS*^{WT} recorded in the presence of 0 mM (black), 150 mM (red) or 500 mM (blue) NaCl. (A) SAXS profiles. (B) Guinier plots. (C) Kratky plots. (D) Porod-Debye plots.

4.6 The binding of PEP/Mn²⁺ to *PniDAH7PS*^{WT} in the presence of prephenate

Standard ITC assays were employed to explore the influence of prephenate on PEP binding to *PniDAH7PS*. PEP (1.5 mM) was titrated into the calorimeter cell containing 65 μ M *PniDAH7PS*^{WT} and 250 μ M prephenate in the presence of MnSO₄. The resultant binding isotherms were fitted to an independent model and the fits were similar to those observed when prephenate was absent (Figure 4.9A,B). The K_D value for PEP binding to *PniDAH7PS*^{WT} increased from 51.8 ± 6.1 μ M to 106.3 ± 15.5 μ M in the presence of prephenate (Figure 4.9C). Moreover, the thermodynamic parameters suggest that the binding of PEP remained as an endothermic process in the presence of prephenate; being solely driven by $-T\Delta S$ (Figure 4.9D).

Figure 4. 9: ITC assays showing the thermodynamics of PEP binding to *PniDAH7PS*^{WT} and the effect of prephenate of PEP binding. (A, B) Raw data from ITC experiments involving the titration of 1.5 mM PEP into the cell containing 65μM *PniDAH7PS*^{WT} in the presence of 0 μM or 250 μM prephenate respectively. The fits to the independent binding model are indicated by the solid line. (C) Comparison of the K_D value for PEP in the presence of 0 μM (grey) or 250 μM (Red) prephenate. (D) Comparison between the thermodynamic parameters for PEP binding to *PniDAH7PS*^{WT} in the presence of 0 μM (grey) or 250 μM (Red) prephenate.

4.7 The allostery of *GspDAH7PS* by prephenate

GspDAH7PS is a type Iβ DAH7PS which contains an N-terminal CM domain. *GspDAH7PS* and *PniDAH7PS* share two common features—the DAH7PS barrels are covalently connected to a CM domain via a long and flexible linker and both proteins are allosterically inhibited by prephenate. However, the CM domains of these two enzymes serve different functional purposes. This section aims to describe the allosteric inhibition of the DAH7PS function of *GspDAH7PS* exerted by CM domain, upon the binding of prephenate.

Figure 4. 10: Sedimentation velocity analysis for *GspDAH7PS*^{WT} and *GspDAH7PS*^D at pH 7.4 and 20 °C recorded at 50,000 rpm. (A) The normalised species distribution for *GspDAH7PS*^{WT} at concentrations of 0.16 mg/mL (red), 0.4 mg/mL (blue) and 1.2 mg/mL (black). (B) The normalised species distribution for *GspDAH7PS*^D at concentration of 0.5 mg/mL (blue), 1 mg/mL (red) and 1.8 mg/mL (black). This AUC work was conducted by Dr Ali Nazimi and Mohamad Othman.

GspDAH7PS is a typical type Iβ DAH7PS, consisting of a simple DAH7PS barrel and a functional CM extension at the N-terminus. The DAH7PS core forms a homotetramer and the four CM moieties assemble two pairs of CM dimers sitting on two opposite sides the tetrameric TIM barrel core (Figure 1.12 B). The crystal structure of prephenate bound *GspDAH7PS* was solved by Dr Ali Nazmi with assistance from Dr Santosh Panjikar, Dr Vickery Arcus and Dr Timothy Allison and a homology model for the ligand-free protein was generated by Dr Eric Lang. The biochemical characterisation, carried out by Ali Nazmi and Mohamad Othman, showed that in *GspDAH7PS*, the DAH7PS and CM domains were fairly independent from each other

in function and quaternary structure when prephenate was absent. Both the separated DAH7PS and CM domains (*GspDAH7PS^D* and *GspDAH7PS^{CM}*) were able to maintain their native quaternary structures, although the *GspDAH7PS^D* homotetramer exhibited less stability compared to that of *GspDAH7PS^{WT}*. In addition, *GspDAH7PS^D* was found to also exist as a homodimer (Figure 4.10). The catalytic activities of either domain were not greatly affected by the domain-separation. The majority of the two enzymatic activities were retained; *GspDAH7PS^D* even displayed an enhanced activity (Table 4.6). However, the removal of the CM domain completely abolished the inhibitory effect of prephenate on the DAH7PS activity (Figure 4.11).

Table 4. 6: Kinetic parameters for *GspDAH7PS^{WT}*, *GspDAH7PS^D*, and *GspDAH7PS^{CM}* (20)

Enzyme	DAH7PS activity					CM activity		
	K_M^{PEP} (μM)	K_M^{E4P} (μM)	k_{cat} (s^{-1})	$k_{\text{cat}}/K_M^{\text{PEP}}$ ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	$k_{\text{cat}}/K_M^{\text{E4P}}$ ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)
<i>GspDAH7PS</i>	87 ± 9	95 ± 8	45 ± 4	0.52	0.47	7.6 ± 0.6	1.68 ± 0.06	0.22
<i>GspDAH7PS^D</i>	98 ± 11	62 ± 5	1.62 ± 0.11	0.64	1	NA	NA	NA
<i>GspDAH7PS^{CM}</i>	NA	NA	NA	NA	NA	450 ± 40	5.5 ± 0.5	1.2 × 10 ⁻²

Note: Kinetic assays were conducted by Mohamad Othman and Dr Ali Nazimi.

Figure 4. 11: Response of *GspDAH7PS^{WT}* (blue) and *GspDAH7PS^D* (red) DAH7PS activity to increasing concentrations of prephenate at 60 °C (20). This work was conducted by Mohamad Othman and Dr Ali Nazimi.

A comparison between the crystal structure of prephenate bound *GspDAH7PS*^{WT} and the homology model for the apo-enzyme revealed a tight association between a CM dimer and the entrances of the proximate DAH7PS dimer's active sites in the presence of prephenate; acting like a switch from an 'open' conformation to a 'closed' conformation to occlude the binding of the substrates at the DAH7PS active site (Figure 1.12B). To verify this inferred mechanism underlying the allosteric inhibition of *GspDAH7PS*^{WT}, experimental SAXS data were collected for *GspDAH7PS* in the absence or presence of prephenate. A comparison between the recorded SAXS profiles of *GspDAH7PS*, in the absence or presence of prephenate, showed clear structural variation in the presence of prephenate, consistent with the conformational change predicted by the crystal structure and homology model (Figure 4.12 A). This change in conformation is associated with an increased compactness of the enzyme in the presence of prephenate as indicated by the smaller radius of gyration ($R_g = 34.8 \pm 0.6$ Å with prephenate compared to $R_g = 36.5 \pm 0.4$ Å without prephenate) (Table 4.7 and Figure 4.11 A).

Table 4. 7: SAXS parameters for *GspDAH7PS*^{WT} in the absence or presence of prephenate

	apo-protein	prephenate-bound
Guinier Analysis		
R_g (Å)	36.5 ± 0.4	34.8 ± 0.6
$I(0)$ (cm ⁻¹)	0.14 ± 0.00	0.13 ± 0.00
Correlation coefficient, R^2	0.99	0.99
Pair Wise Distribution Analysis		
R_g (Å)	36.6 ± 0.4	35.0 ± 0.5
$I(0)$ (cm ⁻¹)	0.14 ± 0.00	0.13 ± 0.00
D_{max} (Å)	147.2	148.5
V_p (nm ⁻³)	297	279
<i>MM</i>_{porod}		
<i>MM</i> (kDa)	175	164
Number of Subunits	4	4

V_p : Porod volume; ***MM*_{porod}**: molecular mass estimated from Porod volume. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

Notably, the bell-shaped curve observed in the Kratky plot (Figure 4.11B) shows that in both the presence and absence of prephenate, the protein exhibits a typical folded shape, without random coils. This verifies that both domains remain folded, even in the absence of prephenate; although the slightly more flattened curve for the prephenate-free sample may indicate an increased flexibility of the protein. The Porod-Debye plots (Figure 4.12C) present two different plateaus in each of the two cases, suggesting that the protein experiences a discrete conformational change upon ligand binding without any major change in its intrinsic flexibility (145,146). This conclusion is supported by the $q^3 \cdot I(q)$ vs. q^3 plot (Figure 4.12D) that displays a decrease, rather than plateau as one would expect in the case of a highly flexible enzyme.

Figure 4. 12: Analysis of the SAXS data for *GspDAH7PS* in the absence (blue) or presence (red) of prephenate. (A) SAXS profiles ($\log I(q)$ vs. q) with Guinier plots inset. (B) Kratky plot ($q^2 \cdot I(q)$ vs. q). (C) Porod-Debye plot ($q^4 \cdot I(q)$ vs. q^4) and (D) SIBYLS plot ($q^3 \cdot I(q)$ vs. q^3), limited to the range of the SAXS data for which Guinier linearity is observed.

The SAXS data, recorded in the absence or presence of prephenate, were compared to the prephenate free (open form) homology model and prephenate bound (closed form) crystal structures of the protein (Figure 4.13A,B). The theoretical scattering curves calculated from each model were in good agreement with the corresponding experimental data ($\chi^2 = 0.7$ and 0.4 for the closed and open forms respectively).

An OLIGOMER (147) analysis based on the open form and the closed form models of *GspDAH7PS* indicated that a mixture of these two conformers fitted the experimental SAXS data ($\chi^2 = 0.2$), collected in the absence of prephenate, better compared to the closed form model alone ($\chi^2 = 0.7$). This analysis predicted that ~73% of the protein adopted the open conformation and ~27% of the protein adopted the closed conformation (Figure 4.13A). On this basis, a conformational equilibrium may be established between the open and closed forms in the absence of a ligand binding to the CM domains; and intermediate states between the open and closed forms may be present. Nevertheless, the open structure, or an open-like structure, is predominant, suggesting that the closed form is less stable or more difficult to access in the absence of prephenate. In the presence of prephenate, the best fit to the experimental profile ($\chi^2 = 0.4$) was obtained using the closed model alone, indicating that the closed structure, or a set of conformationally similar closed-like structures, is the sole form adopted by the prephenate-bound protein (Figure 4.13 B). These data are consistent with the observed full loss of DAH7PS activity of *GspDAH7PS* on prephenate binding.

Figure 4. 13: The distribution of open and closed conformations in *GspDAH7PS* with or without prephenate in solution. Theoretical scattering profiles were generated from models of both the prephenate-free, open form, (*blue line*), and the prephenate-bound, closed form, (*red line*) using OLIGOMER. These theoretical profiles were compared to the experimental *GspDAH7PS* SAXS data (*open circles*) collected in both the absence (**A**) and the presence (**B**) of prephenate. The experimental SAXS data were fitted to a distribution of both open and closed forms using OLIGOMER (*yellow dashed line*) and the column charts show the percentages of open and closed forms of *GspDAH7PS* in solution with and without prephenate. (**C**) The crystal structure of the prephenate bound *GspDAH7PS* is docked into the molecular envelope derived from the experimental data of *GspDAH7PS* in the presence of prephenate.

An *ab initio* envelope of prephenate bound *GspDAH7PS* was generated using Gasbor (148) and the superimposition between the *ab initio* model and the atomic structure of the closed form *GspDAH7PS* (PDB ID: 5J6F) showed a reasonable consistency between the crystal and solution structures of the prephenate bound *GspDAH7PS*,

indicating that the crystal structure of the closed form *Gsp*DAH7PS is representative of the solution-state structure of *Gsp*DAH7PS in the presence of prephenate (Figure 4.13 C).

4.8 Summary and discussion

In Chapter 2, a structural characterisation of the allostery of *Pni*DAH7PS^{WT} by prephenate was carried out in the absence of any substrates of the DAH7PS reaction, whereas, the inhibition kinetic assays were performed by initiating the DAH7PS reaction in the presence of prephenate. Moreover, the variation in the conformation of *Pni*DAH7PS^{WT} was observed in the presence of PEP/Mn²⁺ or PEP/Mn²⁺/E4P. Therefore, the SAXS examinations for investigating the allostery of *Pni*DAH7PS^{WT} by prephenate were repeated when PEP/Mn²⁺ or PEP/Mn²⁺/E4P were added in the experimental condition in order to gain a more comprehensive understanding of the mechanisms underlying this prephenate-mediated inhibition.

The SAXS data of PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} and PA-PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} presented highly similar structural parameters (Table 4.3 and Figure 4.3E-H) and the rigid body models derived from them showed almost identical conformations (Figure 4.4B,C). The remarkably higher rigidities indicated by these measurements (Figure 4.3G,H), in comparison to conditions where prephenate was not present, suggested that the addition of prephenate significantly narrowed conformational distribution of the protein. Based on these experimental observations and inferences, it is hypothesised that prephenate disrupts the dynamic DAH7PS-CM spatial organisation correlated with the DAH7PS reaction by stabilising *Pni*DAH7PS^{WT} in a narrow conformational distribution principally consisting of highly compact structures, so as to deprive the enzyme of dynamic-dependent catalytic activity.

Further kinetic and SAXS examinations on the allostery of *Pni*DAH7PS^{Δ5AA} by

prephenate indicated that the truncation of loop_{C320-F333} interfered with the inhibitory effects of prephenate on the activity of *Pni*DAH7PS^{WT} (Figure 4.5) and also disrupted the compactness and rigidity of *Pni*DAH7PS^{WT} upon the binding of prephenate (Table 4.4 and Figure 4.6C,D). These experimental data suggest an important role of loop_{C320-F333} in the allosteric mechanism of *Pni*DAH7PS^{WT}. The rigid body model for either PA-PEP/Mn²⁺-*Pni*DAH7PS^{WT} or PA-PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} predicts that loop_{C320-F333} may participate in the allosteric inhibition by interacting with the loop β6-α6 of the DAH7PS domain (Figure 4.14), similar to what was observed in the model for PEP/Mn²⁺/E4P-*Pni*DAH7PS^{WT} (Section 3.2.2). In addition, unlike the structural effect of PEP/Mn²⁺ on *Pni*DAH7PS^{WT}, the allosteric conformational change, observed for *Pni*DAH7PS^{WT}, upon the binding of prephenate was not notably influenced by the variation in the concentration of NaCl (Table 4.5 and Figure 4.7). This implies that more stable interactions, which could be preserved in the presence of a reasonably broad range of ionic strength conditions, might occur between the DAH7PS and CM domains to maintain the compact and rigid structure.

In addition, an interruption of PEP binding to *Pni*DAH7PS^{WT} by prephenate was detected by ITC assays (Figure 4.8) However, due to the lack of further experimental data, it is hard to tell if this interference with PEP binding is the consequence of the rigidification of the structure of *Pni*DAH7PS^{WT}, or caused by other mechanisms, for instance, the ‘gating’ inhibitory machinery adopted by *Gsp*DAH7PS.

Although it is hard to make a full description of the process of allosteric conformational change for *Pni*DAH7PS^{WT}, the available data do illustrate a distinctive allosteric mechanism adopted by *Pni*DAH7PS^{WT} from the mechanism underpinning the allosteric inhibition of *Gsp*DAH7PS by prephenate. Compared to the ‘gating’ mechanism for controlling the accessibility of substrates, the allosteric binding of prephenate rigidifies *Pni*DAH7PS^{WT}, subsequently reducing the dynamic nature of the hetero-domain interactions that are required for catalysis.

Chapter 5 *Expressed protein ligation coupled SANS and NMR assays for investigating the structure of PniDAH7PS*

5.1 Introduction

High intrinsic protein flexibility, as exhibited by *Pni*DAH7PS, often affects protein crystallisation and makes structural investigation challenging. Owing to the development of alternative tools and newly emerging methods, more structural data for dynamic proteins have become attainable without the use of crystallography. The work described in this chapter aimed to explore how using a suite of complementary techniques, combining expressed protein ligation (EPL)-mediated segmented labelling with small-angle neutron scattering (SANS) and nuclear magnetic resonance (NMR) spectroscopy, could shed light on the structural features of *Pni*DAH7PS.

5.1.1 Small angle neutron scattering

As is the case for X-rays, neutrons can be scattered by macromolecules, interacting with neutrons rather than electrons. X-rays are sensitive to electrons, and this sensitivity increases with the atomic number. Hence X-ray diffraction is much more sensitive to heavy atoms compared to nitrogen, carbon and hydrogen. However, neutrons interact with the nucleus without any clear tendency related to the atomic number. Therefore, neutron scattering is as sensitive to light elements as to heavy elements. The common elements, from which proteins are composed, including carbon, nitrogen and oxygen, have similar scattering length densities (SLD) and the isotopic substitutions for these atoms do not incur a huge difference in the SLD. However, for hydrogen isotopes (hydrogen and deuterium), neutron scattering results in a negative SLD for hydrogen or a strongly positive SLD for deuterium. This broad range of the SLD for the hydrogen isotopes means that the average SLD of water comprising solutions for biological experiments, can be varied from a value of -0.6 to $6.4 \times 10^{10} \text{ cm}^{-2}$ by manipulating the ratio of H_2O to D_2O . As the SLD of both protonated and deuterated proteins falls within this range, along with the fact that the intensity of small-angle scattering is determined by subtracting the SLD of the solvent from that

of solute molecule, it is possible to sufficiently eliminate the scattering of specific parts of a protein, or protein complexes, through the selective deuteration of the protein and by varying D₂O percentage of the buffer. Furthermore, this makes it feasible to examine the scattering profiles of certain parts (or domains) of a protein separately by matching out other parts. Therefore, SANS is a very powerful technique to dissect the structure of a biomolecular system, including protein-protein complexes, protein-DNA/RNA complexes or multi-domain proteins.

5.1.2 Expressed protein ligation

Expressed protein ligation (EPL) is an engineering method for covalently connecting multiple proteins or peptide fragments together *in vitro* (149,150). This post-translational ligation allows for the specific labelling of proteins for biophysical assays such as SANS, NMR, and Förster resonance energy transfer (FRET). In the EPL method, a ligation reaction occurs between a protein with a C-terminal thioester, created by an intein-mediated self-splicing reaction, and a protein with an N-terminal Cys. Using a thiol reagent, the thiol group of the N-terminal Cys attacks the C-terminal thioester group generating a new thioester, a further rearrangement forms a peptide bond (Figure 5.1) (150). This novel method provides a chance to reconstruct the homodimer of *Pni*DAH7PS^{WT} from separated DAH7PS and CM proteins *in vitro* to probe the inter-dependent relationship between the DAH7PS and CM domains. As the use of EPL is for the assembly of multi-domain proteins *in vitro*, it can be used to ligate a deuterated CM domain to a protonated DAH7PS barrel to produce a segmentally ²H-labelled *Pni*DAH7PS for SANS examination. This EPL-enabled SANS experiment may reveal the spatial arrangement of the CM and DAH7PS portions within a *Pni*DAH7PS homodimer. This method may also assist the investigation into the relative movement between the two domains, resulting from the presence of PEP/Mn²⁺ or prephenate-induced allostery.

Figure 5. 1: The schematic showing the mechanism of chemical ligation, thioester—cysteine condensation and rearrangement.

5.1.3 Nuclear magnetic resonance (NMR)

In addition to SANS, the EPL mediated segmental labelling can be coupled with NMR spectroscopy. As large proteins cause transverse relaxation faster compared to smaller proteins, resulting in a loss of the NMR signal, protein samples for NMR spectroscopy are typically restricted to a molecular weight of less than 50 kDa (151). Another problem for large proteins in NMR is the large number of signals that may overlap. Segmental labelling has been developed to reduce the transverse relaxation and signal overlap, which makes it feasible to investigate the structure of a specific part, or domain, of a full-length protein (152-154). The DAH7PS barrel of *Pni*DAH7PS was selected as a target for NMR investigation, isotopically labelled and ligated to an unlabelled or deuterated CM domain, which could minimise the transverse relaxation caused by the CM domain and improve the signal of the DAH7PS barrel in *Pni*DAH7PS.

Furthermore, the residues involved in the dynamic DAH7PS-CM hetero-interaction could be probed using a segmentally labelled sample and an NMR technique called chemical shift perturbation (CSP). CSP is very sensitive to the electronic environment of a nucleus, and a perturbation of this signal by a binding event or by a conformational

change could cause changes in the observed chemical shifts. Thus, CSP has been used as an effective probe to identify alternations at a protein-protein interface induced by ligand binding events or during a catalytic cycle (155,156). Currently, the most common method for examining protein CSP is ^{15}N heteronuclear single quantum correlation (^{15}N -HSQC) (157). Through detecting the ^1H - ^{15}N J coupling within the amide groups of the peptide bond, ^{15}N -HSQC produces a spectrum containing NH peaks representing all of the amino acid residues of the protein, except for Pro, which makes a 2D ^{15}N -HSQC spectrum a good structural fingerprint of a protein. Therefore, based on the ^{15}N -HSQC spectrum, the CSP is able to target the amino acid residues involved in a protein interface. In this study, the ^{15}N -HSQC spectra for individual CM and DAH7PS domains were recorded. In the future work, the ^{15}N labelled protein will be ligated to its unlabelled partner and the resultant ligated protein was examined by ^{15}N -HSQC. By comparing the two spectra for each domain in the ligated and un-ligated form, residues mostly relevant to the hetero-domain interface were determined. Furthermore, the spectra for the proteins in the presence of PEP/ Mn^{2+} or prephenate will be recorded to elucidate the role of hetero-domain interactions in the catalysis and allostery of *Pni*DAH7PS. However, prior to performing the CSP assays, the chemical shift peaks of the protein backbone in a ^{15}N -HSQC spectrum have to be assigned to their corresponding residues, which involves a series of ^1H , ^{13}C and ^{15}N 3D NMR experiments (158,159).

5.2 Development of the method for the synthesis and purification of ligated *Pni*DAH7PS^{D-CM}

Truncated non-deuterated DAH7PS and CM variants were used to assemble a full-length *Pni*DAH7PS via an EPL reaction. The intein-mediated EPL reaction requires an intein C-terminus of one reactant protein and a Cys residue at the N-terminus of the second reactant protein. As there is no native Cys within the linker region connecting the CM and DAH7PS domains, a Cys residue was introduced using site-directed

mutagenesis for the creation of the N-terminal Cys for the CM reactant. To alleviate the influence of the Cys mutation on the native characteristics of the linker, Thr263 was selected as the mutation point to produce the N-terminal-Cys-CM reactant ($PniDAH7PS^{CM-T263C}$), for the EPL reaction (Figure 5.2 A). The open reading frame (ORF) corresponding to $PniDAH7PS^{CM-T263C}$, the segment from T263(C) to the C-terminus of $PniDAH7PS^{WT}$, was amplified and mutated from the $PniDAH7PS^{WT}$ gene. Subsequently, the DNA of $PniDAH7PS^{CM-T263C}$ was inserted into a pDEST15 expression vector, which encodes a GST tag at the N-terminus of the inserted gene. For the purposes of both purification and creation of the C263 N-terminus by TEV cleavage (Figure 5.2 A), a TEV cleavage site, ENLYFQ, was inserted between the GST tag and the $PniDAH7PS^{CM-T263C}$ (GST-TEV- $PniDAH7PS^{CM-T263C}$).

Figure 5. 2: Schematics showing the production of $PniDAH7PS^{CM-T263C}$ (A) and the EPL reaction (B).

The N-terminal reactant for the ligation is the DAH7PS barrel (*Pni*DAH7PS^{D-H262}). The ORF encoding *Pni*DAH7PS^{D-H262}, the segment from the N-terminus to H262 of *Pni*DAH7PS^{WT}, was amplified from the *Pni*DAH7PS^{WT} gene. Subsequently, the DNA of *Pni*DAH7PS^{D-H262} was inserted into a pTXB1 expression vector. The pTXB1 vector encodes a Mxe GyrA intein, a mutated intein with a higher splicing efficiency, at the C-terminus of the inserted gene (Appendix Figure A.5.1 and Figure 5.2B) (160,161). To simplify the purification of the intein-fused DAH7PS, a chitin binding domain (CBD) was encoded by pTXB1 vector as an affinity tag at the C-terminus of the intein domain to construct a *Pni*DAH7PS^{D-H262}-intein-CBD fusion protein (161).

The constructs for the expressions of *Pni*DAH7PS^{CM-T263C} (pDEST15-*Pni*DAH7PS^{CM-T263C}) and *Pni*DAH7PS^{D-H262}-intein-CBD (pTXB1-*Pni*DAH7PS^{D-H262}) were separately transformed into BL21 (DE3) chaperone combination 3 cells, which co-express GroEL and GroES for assisting protein folding (162,163). The expression of GST-TEV-*Pni*DAH7PS^{CM-T263C} and *Pni*DAH7PS^{D-H262}-intein-CBD was induced by the addition of 0.5 mM IPTG, following the standard procedures described in section 7.1.3.7. The purification of GST-TEV-*Pni*DAH7PS^{CM-T263C} and TEV cleavage for yielding *Pni*DAH7PS^{CM-T263C} were carried out as previously described for the truncated *Pni*DAH7PS^{CM} variant (Section 2.5.3).

Figure 5. 3: Reaction of EPL and purification of the ligated product, *PniDAH7PS*^{D-CM}. (A) SDS-PAGE analysis showing the completed EPL reaction; containing substrates and the ligated product. (B) The SEC elution trace obtained during the purification of the ligated product. The red rectangle highlights the eluted fractions containing *PniDAH7PS*^{D-CM}. (C) SDS-PAGE analysis showing the SEC fractions containing *PniDAH7PS*^{D-CM}. (D) SDS-PAGE analysis showing the elution fractions of *PniDAH7PS*^{D-CM} after the second round SEC.

As the initial step of EPL, the clarified cell lysate containing *PniDAH7PS*^{D-H262}-intein-CBD was loaded onto chitin beads until the beads were saturated by *PniDAH7PS*^{D-H262}-intein-CBD at a concentration of ~150 μ M. After washing out the residual cell lysate and unbound proteins, the chitin beads were further loaded with 200 μ M of purified *PniDAH7PS*^{CM-T263C} solution containing 50 mM 2-mercaptoethanesulphonic acid (MESNA) to initiate the EPL reaction. The reaction was carried out at 25 °C for 2 days. All components of the reaction were subsequently eluted from the chitin beads, and analysed by SDS-PAGE analysis (Figure 5.3A). The SDS-PAGE analysis showed that a

considerable amount of the ligated product (*Pni*DAH7PS^{D-CM}), with a similar molecular weight compared to that of *Pni*DAH7PS^{WT}, was obtained.

SEC, using a Hiperrep 16/60 Sephacryl s-200 HR column (GE Healthcare), was employed to isolate *Pni*DAH7PS^{D-CM} from the reaction mixture. Three peaks were observed in the elution trace (at a wavelength of 280 nm), in which *Pni*DAH7PS^{D-CM} eluted in the second peak according to the SDS-PAGE analysis (Figure 5.3B,C). However, a considerable amount of *Pni*DAH7PS^{CM-T263C} co-eluted with *Pni*DAH7PS^{D-CM} (Figure 5.3C), which could not be sufficiently removed by a second round of SEC purification (Figure 5.3D). The purpose of this EPL reaction is to synthesise the homodimer of *Pni*DAH7PS^{D-CM}, however a heterodimer consisting of a *Pni*DAH7PS^{D-CM} monomer and an un-ligated *Pni*DAH7PS^{CM-T263C} monomer may also be produced in an incomplete EPL reaction (Figure 5.4). Therefore, this difficulty in removing the *Pni*DAH7PS^{CM-T263C} contamination from the *Pni*DAH7PS^{D-CM} sample may have been caused by the intramolecular interaction between the *Pni*DAH7PS^{CM-T263C} and *Pni*DAH7PS^{D-CM} chains within a heterodimer, or the intermolecular interaction between the individual *Pni*DAH7PS^{CM-T263C} dimer and the *Pni*DAH7PS^{D-CM} homodimer. In addition, the co-elution of *Pni*DAH7PS^{CM-T263C} with *Pni*DAH7PS^{D-CM} may have been resulted from a possible aggregation of *Pni*DAH7PS^{CM-T263C} via the formation of disulfide bonds, which may have given rise to various oligomers exhibiting a broad range of retention times overlapping with the elution of *Pni*DAH7PS^{D-CM}. The possibility of the formation of *Pni*DAH7PS^{D-CM}/*Pni*DAH7PS^{CM-T263C} heterodimer will be checked in section 5.3.

Figure 5. 4: Schematics illustrating the homodimeric and heterodimeric product produced by EPL.

To attempt to disrupt the suspected DAH7PS-CM intermolecular interaction and disulfide bonds, another SEC purification was performed using an optimised SEC buffer containing increased concentrations of NaCl (500 mM) and TCEP (1 mM), based on the observation of the interference of the DAH7PS-CM intramolecular interaction by high NaCl concentrations (Section 3.3.2). However, the new buffer failed to improve the purification effectively and significant CM contaminant remained. Anion exchange and hydrophobic interaction chromatography were also used to try to remove the CM contamination, without any success (Appendix Figure A.5.2).

Figure 5. 5: SEC-IMAC two-step purification of the ligated product, *PniDAH7PS*^{D-CM}. (A) Elution trace (blue and red curves represent the UV 280 nm and 260 nm traces respectively) of the HisTrap column for purifying *PniDAH7PS*^{D-CM}. The red rectangle indicates the eluted peak containing *PniDAH7PS*^{D-CM}. (B) SDS-PAGE analysis showing the elution fractions collected from the HisTrap column. (C) SDS-PAGE analysis showing *PniDAH7PS*^{D-CM} purified by either SEC alone (lane 1) or SEC followed by IMAC (lane 2, 3).

Subsequently, an uncleavable His₆ purification-tag was added to the N-terminus of *PniDAH7PS*^{D-H262}-intein-CBD to form His-*PniDAH7PS*^{D-H262}-intein-CBD (Appendix Figure A.5.1), and IMAC was used for a further purification of ligation product. Following the completion of the ligation reaction, using the His-tagged reactant, the ligated product, His-*PniDAH7PS*^{D-CM}, was initially purified by SEC to remove the unreacted His-

PniDAH7PS^{D-H262}-intein-CBD, *PniDAH7PS*^{D-H262} and intein-CBD. The His-*PniDAH7PS*^{D-CM} along with the *PniDAH7PS*^{CM-T263C} contaminant, eluted from SEC, was subsequently subjected to IMAC purification, to attempt to remove the contamination (Figure 5.5A). SDS-PAGE analysis for the eluted fractions from the IMAC column showed that the band of *PniDAH7PS*^{CM-T263C} was still observable but significantly more faint compared to the fractions collected directly after elution from the SEC column (Figure 5.5B). The fractions containing *PniDAH7PS*^{D-CM} were combined, concentrated and examined. SDS-PAGE analysis revealed a significant improvement in the purity of the resultant *PniDAH7PS*^{D-CM} compared with the ligated product purified by SEC alone (Figure 5.5C). The final yield of *PniDAH7PS*^{D-CM} was ~3 mg, from ~36 mg of His-*PniDAH7PS*^{D-H262}-intein and ~8 mg of *PniDAH7PS*^{D-H262}. The molecular weight of the isolated ligation product was 40,109.9 Da, as determined by electrospray ionization mass spectrometry (ESI-MS). This is very close to the 40,108.6 Da molecular weight calculated from the amino acid sequence using the ProtParam online server (112).

5.3 Examination of the quaternary structure of *PniDAH7PS*^{D-CM}

To examine the quaternary structure of the purified product of the EPL reaction, analytical SEC assays were carried out. The analytical SEC column was calibrated with blue dextran (2,000 kDa) and standard samples of known molecular weight (aldolase 158 kDa, conalbumin 75 kDa, ovalbumin 43 kDa and cytochrome 12.4 kDa) (Figure 5.6A). *PniDAH7PS*^{WT} (1 mg/mL) and the ligation product (1 mg/mL) were applied to the analytical SEC column, and only one elution peak was observed (Figure 5.6A). The molecular mass was calculated to be 90 kDa for *PniDAH7PS*^{WT} and 77 kDa for the ligated protein, based on their elution volumes. Both of the molecular masses were close to the theoretical molecular mass for homodimeric *PniDAH7PS* (80 kDa) and significantly higher than the theoretical molecular mass for the heterodimeric intermediate (~50 kDa). Moreover, the SDS-PAGE analysis (Figure 5.6B) for the elution peak of the ligated protein from analytical SEC showed a clear *PniDAH7PS*^{D-CM} band

with only a small band corresponding to un-ligated *Pni*DAH7PS^{CM-T263C}. These observations imply that the isolated ligation product mainly consists of the homodimeric *Pni*DAH7PS^{D-CM}, with possible, minor contamination introduced by the presence of *Pni*DAH7PS^{CM-T263C} and/or *Pni*DAH7PS^{D-CM}/*Pni*DAH7PS^{CM-T263C} heterodimer.

Figure 5. 6: Analytical SEC examining the oligomeric state of *Pni*DAH7PS^{D-CM}. (A) Analytical SEC elution trace and standard curve of logMW against $(V_e - V_o)/(V_c - V_o)$ volume for *Pni*DAH7PS^{WT} (black line) and *Pni*DAH7PS^{D-CM} (red line) at 1 mg/mL. The SEC buffer contained 10 mM BTP at pH 7.4, 150 mM NaCl, and 1 mM DTT. (B) SDS-PAGE analysis shows *Pni*DAH7PS^{D-CM} before (lane 2) and after (lane 3) analytical SEC.

5.4 Ligation between non-deuterated truncated DAH7PS and CM variants restores the catalytic activities and allosteric function of *Pni*DAH7PS

The kinetic parameters of *Pni*DAH7PS^{D-CM} and the two reactants (*Pni*DAH7PS^{D-H262} and *Pni*DAH7PS^{CM-T263C}) were examined and compared to those of the *Pni*DAH7PS^{WT}, *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} (Appendix Figure A.5.3 and Appendix Figure A.5.4). All kinetics assays were performed using the standard buffer and following the

procedures described for characterising the DAH7PS and CM activities of *PniDAH7PS*^{WT} (Section 2.4.2).

Table 5. 1: Comparison of the kinetic parameters for *PniDAH7PS*^{D-CM}, *PniDAH7PS*^{D-H262} and *PniDAH7PS*^{CM-T263C}, with that of *PniDAH7PS*^{WT}, *PniDAH7PS*^D and *PniDAH7PS*^{CM}

Enzyme	DAH7PS activity					CM activity		
	K_M^{PEP}	K_M^{E4P}	k_{cat}	k_{cat}/K_M^{PEP}	k_{cat}/K_M^{E4P}	K_M	k_{cat}	k_{cat}/K_M
	(μM)	(μM)	(s^{-1})	($\mu M^{-1} \cdot s^{-1}$)	($\mu M \cdot s^{-1}$)	(μM)	(s^{-1})	($\mu M^{-1} \cdot s^{-1}$)
<i>PniDAH7PS</i> ^{D-CM}	45 ± 8	54 ± 8	7.5 ± 0.4	0.17	0.14	13 ± 2	1.82 ± 0.06	0.14
<i>PniDAH7PS</i> ^{D-H262}	352 ± 28	287 ± 33	1.79 ± 0.09	5.1 × 10 ⁻³	6.7 × 10 ⁻³	NA	NA	NA
<i>PniDAH7PS</i> ^{CM-T263C}	NA	NA	NA	NA	NA	86 ± 6	1.60 ± 0.14	1.9 × 10 ⁻²
<i>PniDAH7PS</i>	41 ± 4	58 ± 6	16.8 ± 0.5	0.41	0.29	7.6 ± 0.6	1.68 ± 0.06	0.22
<i>PniDAH7PS</i> ^D	367 ± 35	239 ± 19	1.62 ± 0.11	4.4 × 10 ⁻³	6.8 × 10 ⁻³	NA	NA	NA
<i>PniDAH7PS</i> ^{CM}	NA	NA	NA	NA	NA	84 ± 5	2.01 ± 0.07	2.4 × 10 ⁻²

The apparent Michaelis constants for the DAH7PS activity are apparent values determined using a fixed concentration of the invariant substrate (as described under “Materials and Methods”). The uncertainty values represent the standard deviation of triplicate measurements.

Compared to the truncation site (K261/H262) for separating *PniDAH7PS*^D and *PniDAH7PS*^{CM} (Section 2.5.1), the splitting point selected for generating the ligation reactants shifts towards the C-terminus of *PniDAH7PS*^{WT} only by one residue. It appeared that this small change did not cause any perceptible alterations in the two separated enzymatic functions, and both *PniDAH7PS*^{D-H262} and *PniDAH7PS*^{CM-T263C} exhibited weak catalytic activities, presenting highly similar kinetic characteristics compared to those of *PniDAH7PS*^D or *PniDAH7PS*^{CM}. After covalently linking *PniDAH7PS*^{D-H262} and *PniDAH7PS*^{CM-T263C} together to form a *PniDAH7PS*^{D-CM}, both the poor DAH7PS and CM enzymatic activities were boosted. The values of K_M^{PEP} and K_M^{E4P} for *PniDAH7PS*^{D-CM} were significantly reduced compared to those for the truncated *PniDAH7PS*^{D-H262}. Further comparisons between the kinetic parameters of *PniDAH7PS*^{D-CM} and *PniDAH7PS*^{WT} indicated that the ligation product and *PniDAH7PS*^{WT} have the similar affinities for both substrates of the DAH7PS reaction, as evidenced by their comparable K_M^{PEP} and K_M^{E4P} values (Table 5.1). The k_{cat} for the DAH7PS activity of

*Pni*DAH7PS^{D-CM} was $7.5 \pm 0.4 \text{ s}^{-1}$, which was notably improved compared to that for *Pni*DAH7PS^{D-H262} ($1.79 \pm 0.09 \text{ s}^{-1}$), but still lower than that for *Pni*DAH7PS^{WT} ($16.8 \pm 0.5 \text{ s}^{-1}$) (Table 5.1). Overall, approximately half of the native DAH7PS activity was restored via ligation (Figure 5.7). The lower enzymatic efficiencies of the ligation product, relative to those of *Pni*DAH7PS^{WT}, may be the result of the contamination by *Pni*DAH7PS^{CM-T263C} and/or *Pni*DAH7PS^{D-CM}/*Pni*DAH7PS^{CM-T263C} heterodimer, or other factors, such as the T263C mutation at the linker region of *Pni*DAH7PS^{D-CM}.

Similar to the DAH7PS activity, the enzymatic function of CM was also significantly restored after *Pni*DAH7PS^{CM-T263C} was ligated to *Pni*DAH7PS^{D-H262}, as indicated by the decreased K_M for chorismate, from $86 \pm 6 \text{ } \mu\text{M}$ (*Pni*DAH7PS^{CM-T263C}) to $13 \pm 2 \text{ } \mu\text{M}$ (*Pni*DAH7PS^{D-CM}) (Table 5.1). The overall effect of the ligation between *Pni*DAH7PS^{CM-T263C} and *Pni*DAH7PS^{D-H262} on the CM enzymatic function was that more than half of the CM activity of *Pni*DAH7PS^{WT} was restored (Figure 5.7).

Figure 5. 7: A bar chart shows the values of k_{cat}/K_M for the DAH7PS and CM activities of *Pni*DAH7PS^{WT}, truncated variants (*Pni*DAH7PS^D and *Pni*DAH7PS^{CM}) and *Pni*DAH7PS^{D-CM}, indicating restoration of the functions of DAH7PS and CM by EPL.

Subsequently, the inhibitory effect of prephenate on the DAH7PS function of *Pni*DAH7PS^{D-CM} was examined following the procedures described in section 7.1.14. Inhibition of DAH7PS activity by prephenate was observed for *Pni*DAH7PS^{D-CM} (Figure 5.8). Similar to the response of *Pni*DAH7PS^{WT} to prephenate, *Pni*DAH7PS^{D-CM} showed a high sensitivity towards prephenate and the relative IC₅₀ value for prephenate was calculated as 5.3 μ M, in close agreement to that of *Pni*DAH7PS (4.2 μ M). However, compared to *Pni*DAH7PS^{WT}, a higher percentage of remaining DAH7PS activity (~30%) was detected for *Pni*DAH7PS^{D-CM} in the presence of high prephenate concentrations. Taken together, the ligation between the truncated DAH7PS and CM variants effectively reconstructed both the catalytic and allosteric functions of *Pni*DAH7PS.

Figure 5. 8: The response of *Pni*DAH7PS^D (blue curve), *Pni*DAH7PS^D (red curve), and *Pni*DAH7PS (black curve) to varying concentrations of prephenate. Error bars represent the standard deviation of triplicate measurements.

5.5 Segmental deuteration and SANS profile of *Pni*DAH7PS^{D-CM}

Due to the difficulty in crystallising *Pni*DAH7PS^{WT} and the demand for more structural information for *Pni*DAH7PS^{WT} to get a deeper insight into the mechanisms underpinning the catalysis and allostery, a method of coupling segmental deuteration with SANS to probe the relative motions between the DAH7PS and CM domains of *Pni*DAH7PS was developed in this study.

Using the ligation protocol described above, perdeuterated *Pni*DAH7PS^{CM-T263C} (^DCM) and protonated *Pni*DAH7PS^{D-H262} (^HDAH7PS) were covalently linked together to yield *Pni*DAH7PS^{D-CM} with a selectively deuterated CM domain. Cell growth to produce the ^DCM was carried out by Dr. Anthony Duff at the National Deuteration Facility (NDF) of the Australian Nuclear Science and Technology Organisation (ANSTO). *E. coli* BL21 cells containing pDEST15-*Pni*DAH7PS^{CM-T263C} were used to inoculate three progressively larger flask cultures consisting of the ModC1 minimal media respectively containing 50%, 90% and 90% D₂O, which was followed by the inoculation of 100 mL pre-culture into 1 L bioreactor culture consisting of 90% ModC1 media (164). Protein production was induced with the addition of 1 mM IPTG to the bioreactor culture when the OD₆₀₀ reached ~16, then the growth was continued at 20 °C until exhaustion of the carbon source was indicated by a sudden rise in dissolved oxygen tension and pH (164). Cell harvest and the purification of ^DCM were followed by procedures described for *Pni*DAH7PS^{CM} (Section 2.5.3). The deuteration level of ^DCM was ~76% as determined by mass spectrometric analysis using molecular mass using matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS). Subsequently, the purified ^DCM was subject to the EPL reaction with non-deuterated *Pni*DAH7PS^{D-H262}. As this ligation reaction was carried out prior to the development of the IMAC-SEC combined purification method as described in section 5.2, the ligated product was purified only using SEC. The isolated segmentally labelled *Pni*DAH7PS^{D-CM} was concentrated to ~4 mg/mL in a buffer consisting of 20 mM BTP at pH 7.4, 500 mM NaCl

and 1 mM TCEP, followed by characterisation using contrast-matched SANS experiments at 4 °C using a neutron beam (QUOKKA) at ANSTO. The percentage of D₂O in the buffer was changed from 0% to ~42% and then to ~100% to vary the neutron scattering contrast. In 0% D₂O, the neutron scattering of the whole protein was observed; in ~42% D₂O, the neutron scattering was dominated by ^DCM; and in ~100% D₂O, the neutron scattering represented the scattering of the ^HDAH7PS component (Figure 5.9). These measurements allow for the evaluation of the spatial arrangements and conformations of the individual DAH7PS and CM domains, within the full-length protein.

Buffer	Sample	Labeling
0 % D ₂ O		^H DAH7PS + ^D CM
42 % D ₂ O
100 % D ₂ O

Figure 5. 9: Schematic showing the contrast matching and the segmentally deuterated *Pni*DAH7PS^{D-CM} for SANS examinations.

From the 42% D₂O SANS data, where only the CM portion scattered, structural parameters of $R_g = 24.7 \pm 2.4$ Å and $D_{max} = 77.9$ Å were determined, which are close to the values extracted from the SAXS scattering profile of *Pni*DAH7PS^{CM} ($R_g = 22.2 \pm 0.8$ Å and $D_{max} = 74.0$ Å) (Table 5.2 and Figure 5.10C). The consistency between the two scattering data sets validates the successful segmental deuteration of the CM domain. The 100% D₂O SANS data, for investigating the protonated DAH7PS domain, presented a $D_{max} = 159.9$ Å, which is in accordance with the maximum distance detected in apo-*Pni*DAH7PS^{WT} by SAXS (Table 5.2 and Table 3.1). Moreover, the $P(r)$ analysis of 100% D₂O SANS data showed two evident peaks (Figure 5.10B), indicating a dumbbell-

shaped multidomain protein (Figure A.5.5) (165). The peak at smaller r showed a pattern similar to the $P(r)$ distribution transformed from the SAXS data of $PniDAH7PS^D$, and its central position corresponds to a distance of 22 Å, consistent with the peak position of $PniDAH7PS^D$ at 24 Å (Table 5.2 and Figure 5.10D). These data imply that in $PniDAH7PS^{D-CM}$, two DAH7PS barrels are separated by a portion of CM dimer, which agrees with our prediction about the spatial arrangement of the DAH7PS and CM domains for wild-type protein and further verifies the homodimeric assembly of the ligated product. However, the 0% D₂O SANS data, which should characterise the whole $PniDAH7PS^{D-CM}$ dimer, indicated a significantly smaller size than that expected for $PniDAH7PS^{WT}$ (Table 2). Due to the insufficient purification by SEC, the ligated sample was partially contaminated by the un-ligated CM reactant (Figure 5.4C), which may account for the smaller size of $PniDAH7PS^{D-CM}$ indicated by the 0% D₂O SANS data, compared to that of $PniDAH7PS^{WT}$.

Table 5. 2: SANS and SAXS parameters for segmentally deuterated $PniDAH7PS^{D-CM}$

	R_g (Å) (Guinier)	R_g (Å) [$P(r)$]	D_{max} (Å)
SANS scattering			
$^HDAH7PS^{D-CM}$ (0% D ₂ O)	33.6 ± 1.0	34.6 ± 1.0	124.8
$^HDAH7PS^{D-CM}$ (42% D ₂ O)	24.7 ± 2.4	24.2 ± 0.6	77.9
$^HDAH7PS^{D-CM}$ (100% D ₂ O)	47.7 ± 2.4	45.3 ± 0.2	159.9
SAXS scattering			
$^HDAH7PS^{D-CM}$	43.9 ± 1.5	45.6 ± 0.6	153.3

The uncertainties represent standard deviations estimated by error propagation from the experimental data.

These neutron contrast variation experiments were repeated for $PniDAH7PS^{D-CM}$ in the presence of either PEP/Mn²⁺ (350 μM each) or prephenate (350 μM) to investigate the relative motions between the DAH7PS and CM domains responding to the allosteric effects (Figure A.5.6). However, no noticeable change was observed when comparing

the scattering data of $PniDAH7PS^{D-CM}$ recorded in the absence or presence of ligands.

Figure 5. 10: SANS scattering profiles using contrast variation of ligand-free segmentally deuterated $PniDAH7PS^{D-CM}$. (A) and (B) respectively show the scattering profiles and $P(r)$ distribution functions. After data scaling, the $P(r)$ distribution function transformed from the 42% D_2O SANS data was compared with that from the $PniDAH7PS^{CM}$ SAXS data (C) and the $P(r)$ distribution function transformed from the 100% D_2O SANS data was compared with that from the $PniDAH7PS^D$ SAXS data (D).

5.6 Rigid body modelling of $PniDAH7PS^{D-CM}$ against SANS and SAXS data

To visualise the structural information extracted from the SANS scattering data, rigid body modelling was employed. In addition to neutron scattering data, SAXS data for the $PniDAH7PS^{D-CM}$ sample were also recorded using static SAXS at ANSTO (Table 5.2 and Figure 5.11A). Due to the serious interruption caused by the strong neutron scattering length density of the D CM contaminant, the 0% D_2O SANS data were replaced by the SAXS profile for the rigid body modelling to weaken the contrast

between the scattering of the CM and DAH7PS domains and to provide more realistic information of the overall shape of the *Pni*DAH7PS^{D-CM} homodimer. Homology models of two individual DAH7PS barrels and a CM homodimer were spatially organised against the SAXS and the 42% and 100% D₂O SANS data simultaneously using SASREF7 (132). Finally, a rigid body model for apo-*Pni*DAH7PS^{D-CM} was obtained, which exhibited a reasonable fit to the SAXS and the 42% and 100% D₂O SANS data, $\chi^2 = 1.7$, 1.6 and 2.1 respectively (Figure 5.11A).

Figure 5. 11: Rigid body modelling for apo-*Pni*DAH7PS^{D-CM} against the SAXS and SANS data.

(A) The calculated scattering profiles (red line) of the rigid body models for *Pni*DAH7PS^{D-CM} (left), the CM domain (middle) and the DAH7PS domain (right) fit to the SAXS profile and 42% and 100% SANS data (black circles) respectively, presenting $\chi^2 = 1.7$, 1.6 and 2.1 respectively.

(B) The rigid body model of apo-*Pni*DAH7PS^{D-CM}. The DAH7PS and CM domains are coloured blue and red respectively. **(C)** The rigid body model of apo-*Pni*DAH7PS^{WT}.

Although the relative positions between the DAH7PS barrels and the CM dimer exhibited by the rigid body models of *Pni*DAH7PS^{D-CM} and apo-*Pni*DAH7PS^{WT} are different (Figure 5.11B,C), both models show the agreement in the basic spatial arrangement of the DAH7PS and CM domains, namely that the two DAH7PS barrels are separated and located either side of the CM dimer (Figure 5.11B).

5.7 SAXS examination on the structure of *Pni*DAH7PS^{D-CM} in the presence of either PEP/Mn²⁺ or prephenate

The application of EPL coupled SANS on probing the conformational changes underlying catalysis and allostery in *Pni*DAH7PS was one objective of this study. However, the addition of either PEP/Mn²⁺ or prephenate failed to trigger any observable conformational changes in *Pni*DAH7PS^{D-CM} in the SANS experiments as described above. Therefore, the investigation into the structure of *Pni*DAH7PS^{D-CM} was continued using SAXS, to attempt to find out the cause of the failure to observe a structural effect of PEP/Mn²⁺ or prephenate on *Pni*DAH7PS^{D-CM} in the SANS experiments.

The ligated sample used for the SAXS assays was purified following the IMAC-SEC two-step purification method, described in section 5.4, to remove the contamination of *Pni*DAH7PS^{CM-T263C}. The isolated *Pni*DAH7PS^{D-CM} was stored in standard SAXS buffer (20 mM BTP, 150 mM NaCl, 5% glycerol, pH 7.4) containing less NaCl compared to the buffer used for the SANS experiments. *Pni*DAH7PS^{D-CM} (4 mg/mL) was characterised by SEC-SAXS in the absence or presence of at 25 °C.

Table 5. 3: SAXS parameters determined for *PniDAH7PS*^{D-CM} and *PniDAH7PS*^{WT}

Enzyme	<i>PniDAH7PS</i> ^{D-CM}			<i>PniDAH7PS</i> ^{WT}		
Additive	Ligand-free	PEP/Mn ²⁺	prephenate	Ligand-free	PEP/Mn ²⁺	prephenate
Guinier Analysis						
R_g (Å)	36.1 ± 1.2	36.9 ± 0.4	32.9 ± 0.6	42.0 ± 1.2	40.1 ± 0.7	32.7 ± 0.3
$I(0)$ (cm ⁻¹)	0.07 ± 0.00	0.07 ± 0.00	0.07 ± 0.00	0.08 ± 0.00	0.07 ± 0.00	0.06 ± 0.00
Correlation coefficient, R^2	0.99	0.99	0.99	0.99	0.99	0.99
Pair Wise Distribution Analysis						
R_g (Å)	37.2 ± 0.7	38.0 ± 0.3	33.0 ± 0.2	44.1 ± 0.2	41.0 ± 0.3	32.9 ± 0.2
$I(0)$ (cm ⁻¹)	0.07 ± 0.00	0.07 ± 0.00	0.07 ± 0.00	0.08 ± 0.00	0.07 ± 0.00	0.06 ± 0.00
D_{max} (Å)	141.8	133.0	109.0	152.3	144.3	105.9
V_p (nm ³)	95	112	102	137	146	115
MM_{porod}						
MM (kDa)	56	66	60	81	86	68
Number of Subunits	2	2	2	2	2	2

V_p : Porod volume; MM_{porod} : Molecular mass estimated from Porod volume. The uncertainties represent standard deviations estimated by error propagation from the experimental data. The columns coloured in grey present the SAXS data shown previously, for a convenient comparison.

The SAXS scattering profiles were recorded for ligand free *PniDAH7PS*^{D-CM} (apo-*PniDAH7PS*^{D-CM}) and the *PniDAH7PS*^{D-CM} in the presence of either the combination of PEP/Mn²⁺ (PEP/Mn²⁺-*PniDAH7PS*^{D-CM}) or prephenate (PA-*PniDAH7PS*^{D-CM}). According to the experimental data, apo-*PniDAH7PS*^{D-CM} and PEP/Mn²⁺-*PniDAH7PS*^{D-CM} appeared to be smaller compared to apo-*PniDAH7PS*^{WT} and PEP/Mn²⁺-*PniDAH7PS*^{WT} respectively (Table 5.3). The reduced size of *PniDAH7PS*^{D-CM} may be partially due to the average of the sizes of *PniDAH7PS*^{D-CM} and the incompletely removed *PniDAH7PS*^{CM-T263C} and/or *PniDAH7PS*^{CM-T263C} heterodimer contaminant. Nevertheless, a comparison between the SAXS profiles of *PniDAH7PS*^{D-CM} under different conditions showed that the ligation product, *PniDAH7PS*^{D-CM}, clearly exhibited conformational changes when either PEP/Mn²⁺ or prephenate was present (Figure 5.12). It appeared that the presence of PEP/Mn²⁺ was not able to notably affect the compactness of *PniDAH7PS*^{D-CM}, according

to close similarity between the R_g values of apo-*Pni*DAH7PS^{D-CM} (36.1 ± 1.2 Å) and PEP/Mn²⁺-*Pni*DAH7PS^{D-CM} (36.9 ± 0.4 Å). However, the average overall shape of *Pni*DAH7PS^{D-CM} was changed due to the addition of PEP/Mn²⁺, as demonstrated by the discrepancy between the SAXS profiles of apo-*Pni*DAH7PS^{D-CM} and PEP/Mn²⁺-*Pni*DAH7PS^{D-CM} at the q range between $\sim 0.075 - 0.15$ Å⁻¹ (Figure 5.12A,C), and also the differentiation between their $P(r)$ distributions (Figure 5.12B). The examination of the structural effect of prephenate on *Pni*DAH7PS^{D-CM} showed a similar conformational change to that observed in *Pni*DAH7PS^{WT}, in which the ligated protein became more compact in the presence of prephenate as evidenced by a decrease in the R_g and D_{max} values, from 37.2 ± 0.7 Å and 141.8 Å to 32.9 ± 0.6 Å and 109 Å, respectively (Table 5.3). Moreover, the Kratky and Porod-Debye transformations revealed that prephenate effectively attenuates the flexibility of the structure of *Pni*DAH7PS^{D-CM} (Figure 5.12D), as its effect on *Pni*DAH7PS^{WT}.

Figure 5. 12: Analysis of the SAXS profiles obtained for apo-*Pni*DAH7PS^{D-CM} (black) and PEP/Mn²⁺-*Pni*DAH7PS^{WT} (red) and PA-*Pni*DAH7PS^{WT} (blue). (A) SAXS profiles ($\log I(q)$ versus q). (B) Guinier plots ($\ln I(q)$ versus q^2). (C) Kratky plot ($q^2 * I(q)$ versus q). (D) Porod-Debye plot ($q^4 * I(q)$ versus q^4) limited to the range of the SAXS data for which Guinier linearity was observed.

5.8 Isotopic labelling of *Pni*DAH7PS and preliminary NMR tests

BL21 cells carrying either the pET28a-*Pni*DAH7PS^{WT}, pDEST17-*Pni*DAH7PS^D or pDEST15-*Pni*DAH7PS^{CM} vectors were grown in M9 minimal growth medium supplemented with 1g/L ¹⁵NH₄Cl and the expression of the encoded proteins was induced by IPTG (Section 7.2.3). The auto-induction technique was also employed for increasing the yield of ¹⁵N-*Pni*DAH7PS^{CM} in the minimal medium (Section 7.2.3). ¹⁵N singly labelled proteins were purified following the protocols described above for the purification of their corresponding unlabelled counterparts (Section 2.3, 2.5.2 and 2.5.3). The HSQC experiment normally requires a protein concentration higher than 200 μM. However, both ¹⁵N-*Pni*DAH7PS^{WT} and ¹⁵N-*Pni*DAH7PS^{CM} showed poor solubilities and precipitated once the concentration exceeded ~150 μM. The DAH7PS-CM hetero-interaction within *Pni*DAH7PS^{WT} alluded to a possible intermolecular DAH7PS-CM interaction. In addition, the homology model of *Pni*DAH7PS^{CM} suggested that the C327 residue in loop_{C320-F333} was exposed to the solvent, which could cause an intermolecular cross-linking mediated by the formation of a disulfide bond. These two suspected factors may lead to the aggregation of *Pni*DAH7PS^{WT} or *Pni*DAH7PS^{CM} at high concentrations. Therefore, the protein storage buffer was optimised by adding Glu and Arg each to a concentration of 50 mM to diminish any electrostatic effects between the protein molecules to interfere with any intermolecular interactions and supplementing with 10 mM dithiothreitol (DTT) to prevent the formation of disulfide bonds. As a consequence, the solubility of *Pni*DAH7PS^{WT} and *Pni*DAH7PS^{CM} was significantly improved and samples of ~200 μM ¹⁵N-*Pni*DAH7PS^{WT}, ~200 μM ¹⁵N-*Pni*DAH7PS^{CM} and ~300 μM ¹⁵N-*Pni*DAH7PS^D were sent to Prof. Joel Mackay at Sydney University for NMR examination.

In the ¹H-¹⁵N HSQC spectrum of ¹⁵N-*Pni*DAH7PS^{WT}, few signals were observed (Figure 5.13A). Subsequently, 1.2 mM of chorismate was spiked into 200 μM *Pni*DAH7PS^{WT} with the expectation of regularising the protein conformation to improve the NMR

spectrum. On the contrary, all of the HSQC peaks vanished following the addition of chorismate (Figure 5.13B). In this situation, transverse relaxation optimised spectroscopy (TROSY) experiments were utilised and coupled with HSQC, which was developed to improve the sensitivity of the NMR experiments to proteins of a large size (166-168). A ^1H - ^{15}N TROSY-HSQC examination was carried out overnight for *PniDAH7PS*^{WT} in the presence of chorismate and many more signals were recorded compared to those recorded in the HSQC spectrum (Figure 5.13C). However, most of the cross peaks observed in the TROSY spectrum were so broad that they merged together and it was impossible to differentiate them from each other (Figure 5.13C). These observations from the preliminary NMR experiments suggested that in addition to its intrinsically large size (80 kDa homodimer), *PniDAH7PS*^{WT} may also form soluble aggregates, which makes the direct application of NMR techniques on the wild-type protein very difficult.

Figure 5. 13: Pre-examinations of ^{15}N singly labelled *PniDAH7PS^{WT}* (0.21 mM) using ^1H - ^{15}N HSQC and TROSY. Spectra were collected at a proton resonance frequency of 600 MHz, at 298 K, in 10 mM BTP buffer at pH 7.3 containing 150 mM NaCl, 50 mM Glu, 50 mM Arg and 10 mM DTT. (A) ^1H - ^{15}N HSQC spectrum of *PniDAH7PS^{WT}*. (B) ^1H - ^{15}N HSQC spectrum of *PniDAH7PS^{WT}* in the presence 1.6 mM chorismate. (C) ^1H - ^{15}N TROSY-HSQC spectrum of *PniDAH7PS^{WT}*.

^{15}N -*PniDAH7PS^{CM}* ($\sim 200 \mu\text{M}$) and ^{15}N -*PniDAH7PS^D* ($\sim 300 \mu\text{M}$) were examined by ^1H - ^{15}N

HSQC using a 600 MHz spectrometer at 320 K and 298 K respectively. The spectrum for *Pni*DAH7PS^{CM} was of higher quality compared to that obtained for *Pni*DAH7PS^{WT} but was still too poor to continue further analysis (Figure 5.14A). The low solubility and serious aggregation observed during the preparation of ¹⁵N-*Pni*DAH7PS^{CM} may account for the low quality of the resultant spectrum. In contrast to the unsatisfactory NMR spectra of *Pni*DAH7PS^{WT} and *Pni*DAH7PS^{CM}, the detection on ¹⁵N-*Pni*DAH7PS^D gave rise to a very nice spectrum, representing ~240 cross-peaks; excluding the Gln and Asn sidechains. *Pni*DAH7PS^D contains 265 residues, of which 20 Pro and the N-terminus were invisible under the ¹H-¹⁵N HSQC inspection. Thus, the number of the observed signals is extremely close to the expected maximum number of 244 (Figure 5.14B). Thus, the assignment of the cross-peaks to each residue of *Pni*DAH7PS^D appeared to be fairly promising.

Figure 5. 14: Examinations of ¹⁵N singly labelled *Pni*DAH7PS^{CM} (0.2 mM) and *Pni*DAH7PS^D (0.3 mM) using ¹H-¹⁵N HSQC at a proton resonance frequency of 600 MHz. (A) ¹H-¹⁵N HSQC spectrum of *Pni*DAH7PS^{CM} collected at 310 K, in 10 mM BTP buffer at pH 7.8 containing 150 mM NaCl, 50 mM Glu, 50 mM Arg and 10 mM DTT. (B) ¹H-¹⁵N HSQC spectrum of *Pni*DAH7PS^D collected at 298 K, in 10 mM BTP buffer at pH 7.2 containing 150 mM NaCl and 1 mM DTT.

To achieve a backbone assignment, a number of ¹H/¹³C/¹⁵N 3D NMR experimental data for *Pni*DAH7PS^D were required, including HNCA, HN(CO)CA, HNCACB, HN(CO)CACB,

HNCO and HN(CA)CO. For the purpose of increasing the sensitivity of all of the 3D resonances, samples were usually deuterated to exclude the disturbance from extra ^1H , in addition to the ^{13}C and ^{15}N labelling. Therefore, the BL21 cells containing the pDEST17-*PniDAH7PS*^D vector were grown in M9 minimal growth medium prepared with D₂O, supplemented with 1 g/L ^{15}N -NH₄Cl and ^{13}C -glucose and the expression of protein was induced by the addition of IPTG (Figures 5.15A) (Section 7.2.3). The triply labelled *PniDAH7PS*^D was purified using IMAC and gel filtration, as described above for the unlabelled protein (Figures 5.15B). The molecular weight of the triply labelled, purified protein was determined by ESI-MS. The isotopically labelled protein showed a molecular mass of 32,545.3 Da. According to the theoretical molecular weight of *PniDAH7PS*^D of 29,464.2 Da, it was estimated that the deuteration percentage of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ *PniDAH7PS*^D was around 68%. The 3D experiments for the backbone assignment using this triply labelled sample are currently being performed by Prof. Joel Mackay's group.

Figure 5. 15: The expression and purification of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ *PniDAH7PS*^D. (A) The expression of triply labelled His-*PniDAH7PS*^D in BL21 (ED3) cells. Lane 1: Whole cell lysate. Lane 2: Supernatant after centrifugation, the soluble fraction of the cell lysate. Lane 3: Cell debris, the insoluble fraction of the cell lysate. (B) The elution fractions from SEC for purifying the triply labelled *PniDAH7PS*^D.

5.9 Summary and discussion

The research work presented in this chapter developed a method for ligating truncated DAH7PS and CM variants using EPL to form a functional *Pni*DAH7PS homodimer *in vitro*. Several successful precedents for the *in vitro* assembly of proteins with a large molecular mass have been achieved by either EPL or by enzyme catalysed ligation (169-171). However, these synthetic proteins mainly exert functions as binding partners for other proteins and DNA/RNA (169,170), and in contrast to *Pni*DAH7PS, their multiple domains do not exhibit any extensive intramolecular interactions. With the concern of introducing too many non-native residues and disrupting the hetero-domain interactions within *Pni*DAH7PS, EPL was selected to reconstruct *Pni*DAH7PS, rather than using the enzymatic methods, which require the addition of a considerable number of specific residues or sequences in the C and N-terminal regions of reactants (150,172,173).

The recovered enzymatic activities and allosteric inhibition for *Pni*DAH7PS^{D-CM} demonstrated that the architecture of *Pni*DAH7PS^{WT} was essentially rebuilt by joining the individual DAH7PS and CM truncated variants together using EPL, which also validates the utility of the ligated protein, *Pni*DAH7PS^{D-CM}, for the study into the native hetero-domain interactions of *Pni*DAH7PS. The repaired catalytic and allosteric functions exhibited by the ligated *Pni*DAH7PS^{D-CM} also provide solid evidence to show that the enzymatic activities of this protein are delivered by a DAH7PS-CM hetero-interaction.

The application of neutron variation contrast on CM-deuterated *Pni*DAH7PS^{D-CM} gave rise to clear scattering data for both the DAH7PS and CM domains, based on which a rigid body model was generated which was consistent with the spatial arrangement of the DAH7PS and CM domains in the model derived from the SAXS profile of *Pni*DAH7PS^{WT}. However, the expected conformational changes of *Pni*DAH7PS^{D-CM}

relating to substrate binding and allostery were not observed in the SANS experiments. As mentioned above, the SANS sample contained contamination of *Pni*DAH7PS^{CM-T263C} and/or *Pni*DAH7PS^{CM-T263C}/*Pni*DAH7PS^{D-CM} heterodimer. Therefore, the absence of conformational change may be the result of the intermolecular interactions between the contaminant and the *Pni*DAH7PS^{D-CM} homodimer, which could interrupt the normal intramolecular interactions responding to the ligand binding. Also, the dimeric CM portion of the contaminant might bind to prephenate, and in turn to reduce the concentration of the available prephenate for the allostery of *Pni*DAH7PS^{D-CM}. In addition, high NaCl (500 mM) in the SANS buffer was highly likely to have muted the structural response of *Pni*DAH7PS^{D-CM} to PEP/Mn²⁺, just as the effect of high salt on *Pni*DAH7PS^{WT}, as described in section 3.3.2. Unfortunately, beamtime was not available to allow us to repeat the experiment under lower salt conditions.

The majority of the CM contaminant was removed using the IMAC-SEC two-steps purification method. Thereafter the polished *Pni*DAH7PS^{D-CM} was examined by SAXS in the standard buffer containing 150 mM NaCl. The conformational response of *Pni*DAH7PS^{D-CM} to PEP/Mn²⁺ or prephenate was detectable in the SAXS profiles, which exhibited similar structural variations as compared to *Pni*DAH7PS. However, the SAXS profile of apo-*Pni*DAH7PS^{D-CM} indicated that the protein was more compact compared to apo-*Pni*DAH7PS^{WT}. Some residual contaminating *Pni*DAH7PS^{CM-T263C} and/or *Pni*DAH7PS^{CM-T263C}/*Pni*DAH7PS^{D-CM} heterodimer may account for the structural discrepancy between the ligated and wild-type *Pni*DAH7PSs. It is noted that the maximum distance suggested by the 100% D₂O SANS data for the DAH7PS domains (159 Å) is in accordance with that of *Pni*DAH7PS^{WT} (152 Å). Hence, the real size of the ligated sample from the SAXS experiments could also be similar to the wild-type protein, but this was diminished in the SAXS profile by the contribution of the remaining contaminant. In addition to this speculation, other factors such as the effect of the T263C mutation may also contribute towards the difference between the sizes of *Pni*DAH7PS^{D-CM} and *Pni*DAH7PS^{WT}, as indicated by the scattering data. This could be

checked by exchanging the Thr263 residue of *Pni*DAH7PS^{WT} for Cys by mutagenesis in future work.

Chapter 6 *Summary and remarks*

This thesis presents a comprehensive characterisation of *Pni*DAH7PS. *Pni*DAH7PS is a representative of a subgroup of type I β DAH7PS enzymes, the C-terminal CM linked DAH7PSs. This is a group of DAH7PS proteins that was identified based on the analysis of protein sequence identity, but has not previously been extensively investigated. As the proteins of this group of type I β DAH7PS group are bifunctional, comprising DAH7PS and CM moieties, *Pni*DAH7PS and other C-terminal CM linked DAH7PSs were initially considered as the C-terminal version of the N-terminal CM-linked DAH7PSs, such as *Bsu*DAH7PS and *Gsp*DAH7PS (8,20). However, our investigation on *Pni*DAH7PS provides adequate experimental evidence to demonstrate that this underestimated DAH7PS subgroup adopts unique architecture, completely different to that observed in other well-known DAH7PS enzymes. This different structure allows *Pni*DAH7PS to fulfil its full catalytic and allosteric functionalities. Most importantly, in this study, a dynamic functional interdependence between the DAH7PS and CM domains was revealed as the key characteristic of *Pni*DAH7PS. This domain interaction underpins both the catalytic activities and allosteric functionality respectively for DAH7PS and CM.

The main findings with respect to the mechanisms underlying *Pni*DAH7PS functions and new techniques that were developed for investigating protein structure, presented in the previous chapters, are summarised in this chapter.

6.1 The significance of the distinctive quaternary structure on the biochemical nature of *Pni*DAH7PS

Although the DAH7PS family comprises a large number of subtypes, the quaternary structures of these enzymes have been found to be largely homotetrameric and homodimeric examples of DAH7PSs are relatively rare (17,18,20,21,28,30,33-35). As was introduced in Section 1.2.2, the homotetrameric assembly is responsible for maintaining thermostability (33,34) or allostery (35), whereas the homodimeric

assembly appears to be the basic functional unit, which may be closely related to the structural organisation of the active site of DAH7PS barrel (33-35). This study of *Pni*DAH7PS uncovered an exception to the common oligomeric organisation of the DAH7PS enzyme, demonstrating a structure consisting of a CM homodimer and two un-associated monomeric DAH7PS domains (Section 2.4.4). It would be surprising to observe that a DAH7PS enzyme maintains a reasonable catalytic activity as a monomer without self-association (Section 2.4.2, Section 2.6.3 and Section 2.8). However, kinetic examination of the separated DAH7PS domain showed that the individual monomeric DAH7PS barrel exhibited poor activity in the absence of the CM domain (Section 2.6.1). Therefore, we hypothesise that CM, as an alternative partner, forms a hetero-interaction with the DAH7PS barrel to exert an equivalent buttressing effect to support the functionality of the DAH7PS active site. Moreover, the truncated CM domain displays extremely poor catalytic activity as well (Section 2.7.1), indicating that the hetero-interaction plays an equivalent role in activating either the DAH7PS or CM functionality. In spite of lacking a crystal structure as direct proof to demonstrate the hetero-interaction, the rigid body models derived from both SAXS and SANS data for apo-*Pni*DAH7PS^{WT} strongly imply an extensive interface between the DAH7PS and CM domains (Figure 3.5 and Figure 5.11).

The divergence in the oligomeric organisation for acquiring DAH7PS activity complicates the prediction of the evolutionary trajectory of the DAH7PS family. Prior to this finding, the commonly accepted proposition describes that the ancestral DAH7PS enzymes might have been a group of primitive type I β DAH7PSs assembled as unregulated homodimers or homotetramers (18,23,33), which were subsequently grafted with extended elements by gene fusions to acquire allosteric regulation (8,20,28,96,97,107,108). Thereafter, a series of insertions, deletions and mutations on the subsequently regulated type I β DAH7PS enzymes led to the type I α and type II DAH7PSs in the proposed trajectory (21). Given that the homomeric association is not strictly a necessary requirement for a functional DAH7PS, an alternative evolutionary

trajectory can be proposed where the ancestral DAH7PS enzymes were monomeric and were subsequently endowed with catalytic and allosteric functions by fusion with a C-terminal CM (Figure 6.1). Unravelling the full evolutionary details for these heterogenic DAH7PS proteins awaits a deeper phylogenetic investigation and more experimental investigations.

Figure 6. 1: Schematic illustrating the proposed evolutionary trajectory of the DAH7PS family. The DAH7PS barrel is shown in green (α-helix) and blue (β-strand) and the extra-barrel structural elements and domains, are marked in red.

6.2 The highly dynamic global structure of *Pni*DAH7PS

Due to the difficulties encountered for obtaining any high-resolution structural information for *Pni*DAH7PS^{WT}, low-resolution structural techniques, including SAXS, were extensively used to probe the DAH7PS and CM hetero-interaction. Although unable to provide atomic-level structural details, SAXS has an outstanding advantage of dissecting conformational variation and assessing protein flexibility by monitoring the averaged overall shape of a protein in solution. Also, without the influence from crystal packing effects, the information about the solution-state structure may be more biologically relevant compared to a solid state structure. The SAXS profile of apo-*Pni*DAH7PS^{WT} suggested an elongated structure and indicates an intrinsically dynamic structure, as demonstrated by the Kratky and Porod-Debye transformations (Figure

3.4). Based on the SAXS investigations, which characterised *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} as a rigid globular protein and a compact rod-shape entity, respectively (Figure 2.17 and Table 2.6), the significant flexibility of the full-length protein (Figure 2.18) is not a feature of the individual domains. Instead, the relative motions between the DAH7PS and CM domains may principally account for the dynamic features of the full-length protein. These dynamic motions appear to be dependent on the nature of the linker between the DAH7PS and CM domains, with a length of ~12 residues, identified by multiple sequence alignments. Due to the crucial role of protein dynamics in enzyme function, allosteric regulation and even evolution (63,77,90,174), dissecting the domain motions and conformational changes that occur during catalysis and as a result of allosteric events within *Pni*DAH7PS could be critical for understanding the mechanisms that underlie catalysis and allostery.

6.3 Dynamic hetero-domain interaction and DAH7PS domain mediated allosteric activation

Besides the impairment on the enzymatic functions of *Pni*DAH7PS^{WT} caused by domain separation, more simple evidence for the hetero-domain interplay is provided by kinetic analysis, which showed a significant promotion of CM activity when a full complement of the substrates of DAH7PS catalysed reaction was added (Section 3.2.1). This communication between the two distant active sites is essentially an allosteric activation of the CM activity exerted by occupancy of the DAH7PS active site, in which DAH7PS displays dual roles as both a catalytic and an allosteric domain simultaneously. The subsequent SAXS investigations, using the SAXS buffer containing PEP/Mn²⁺/E4P, aimed to specify the structural basis of this special phenomenon, uncovered a dramatic conformational change of *Pni*DAH7PS^{WT}. This conformational change involved a significant decrease in the protein size, which manifested as a reduction of ~7 Å, ~12 Å and 15,000 Å³ in the R_g , D_{max} and V_p values respectively, compared to that of apo-*Pni*DAH7PS^{WT}. Considering the high rigidity and small size of an individual

DAH7PS domain (R_g of 19.3 ± 0.5 Å, $D_{\max}$ of 57.0 Å and V_p of 49,693 Å³) (Figure 2.17 and Table 2.6), the observed structural changes are likely a consequence of the interplay between DAH7PS and CM, rather than local conformational changes within the DAH7PS barrel.

In addition to the combination of PEP/Mn²⁺/E4P, the presence of PEP/Mn²⁺ exerts a distinct structural effect on *Pni*DAH7PS^{WT} (Figure 3.4 and Table 3.1) but does not affect the activity of the CM domain (Figure 3.3). The conformation of *Pni*DAH7PS^{WT} in the presence of PEP/Mn²⁺ also displays a higher compactness compared to the ligand-free conformation but differs from that caused by the addition of PEP/Mn²⁺/E4P (Figure 3.4 and Table 3.1). This strongly suggests that variations in the hetero-domain interaction within *Pni*DAH7PS^{WT} may take place in the DAH7PS catalytic cycle. Therefore, it is proposed that a dynamic DAH7PS-CM hetero-interaction provides the structural basis for buttressing both enzymatic functions and propagates an allosteric signal to the CM domain in response to certain catalytic steps of the DAH7PS reaction. This hypothesis is supported by the investigation of the *Pni*DAH7PS^{Δ5AA} variant and the effect of NaCl on the activity of *Pni*DAH7PS^{WT}, in which the loop truncation and high NaCl concentrations disrupted the hetero-domain interaction and both enzymatic activities simultaneously (Section 3.3).

6.4 Allosteric inhibition via alleviating dynamics

A 'gating' mechanism has been well characterised in a number of studies on the allosteric regulation of main group type Iβ DAH7PS enzymes (20,28,109). This thesis also reports an investigation on the allostery of *Gsp*DAH7PS in solution using SAXS (Section 4.7). The comparison between the SAXS profiles of ligand-free and prephenate bound *Gsp*DAH7PS reveals a distinct conformational change (Figure 4.12 and Table 4.7). This change satisfactorily verifies the proposed model in which prephenate-binding induced CM motions block the DAH7PS active sites and, thereby,

suppress the DAH7PS catalytic function.

Given that *Pni*DAH7PS exhibits a DAH7PS-CM functional interdependence that is absent in *Gsp*DAH7PS, it would not be surprising that *Pni*DAH7PS utilises a totally distinct mechanism to carry out allostery. The preliminary study, where *Pni*DAH7PS^{WT} was incubated with prephenate alone, showed substantial compression and rigidification of protein structure (Figure 2.18 and Table 2.6) compared to that in the absence of prephenate. Subsequently, the effect of prephenate on the conformation of *Pni*DAH7PS^{WT} in the presence of either PEP/Mn²⁺ or PEP/Mn²⁺/E4P was inspected. The results showed that under both conditions, the protein adopted a conformation that was similar to that observed in the presence of prephenate alone, but displayed a higher compactness and rigidity (Figure 4.3 and Table 4.3). Therefore, we speculate that the effect of prephenate, likely combined with that of PEP/Mn²⁺, fundamentally rigidifies the protein structure in order to deprive *Pni*DAH7PS of the dynamic-dependent catalytic activity. Although an interruption of PEP binding by prephenate (a ~50% decrease in K_D) was detected by ITC (Figure 4.9), it was seemingly not significant enough to account for the major impairment in the DAH7PS catalytic activity.

6.5 The loop function

The multiple sequence alignment for *Pfu*CM, *Ape*CM and the CM domains of *Pni*DAH7PS, *Pgi*DAH7PS, *Gsp*DAH7PS, and *Lmc*DAH7PS indicated an extra segment of ~10 AA in the C-terminal region of the CM moiety of *Pni*DAH7PS^{WT} and *Pgi*DAH7PS^{WT} (Figure 3.7). The homology model of *Pni*DAH7PS^{CM}, using *Pfu*CM as a template, predicted that this segment to be a long and poorly structured loop linking helices 2 and 3 (Figure 3.7). Generally, loops are more flexible compared to other well-folded structures; thereby they possess a higher diversity in structure compared to α -helices and β -sheets and are easier to manipulate according to variable functional requirements, such as catalysis and allostery (137,138,175,176). To investigate the role

of the suspected additional loop of the CM domain in the bifunctional system of *Pni*DAH7PS, a variant enzyme, *Pni*DAH7PS^{Δ5AA}, with a shortened linker region was generated (Figure 3.8). This variation in the CM linker significantly reduced not only the activity of CM, but also that of DAH7PS (Table 3.3). In addition, the conformational change in the presence of PEP/Mn²⁺/E4P was obviously disrupted, along with a disruption to the allosteric activation of the CM activity by the addition of PEP/Mn²⁺/E4P (Figure 3.9, Figure 3.10 and Table 3.4). Similarly, the prephenate mediated allosteric inhibition of DAH7PS was also notably attenuated by the loop truncation (Figure 4.5). The extremely compact and rigid structure of *Pni*DAH7PS in the presence of prephenate was loosened to give a conformation resembling that induced by PEP/Mn²⁺ (Figure 4.6 and Table 4.4). Although we are lacking high-resolution structural information to describe what happens during the catalytic functions of *Pni*DAH7PS^{Δ5AA} at the molecular level, it has been confirmed that the additional loop in the CM domain evidently contributes towards the functional interdependence and the allostery of both DAH7PS and CM.

6.6 Expressed protein ligation

As mentioned above, due to the high intrinsic flexibility of *Pni*DAH7PS^{WT}, traditional crystallographic methods were unable to provide structural information in this project. Therefore, a technique called expressed protein ligation (EPL) was employed to develop an *in vitro* assembly of the full-length protein using the separated DAH7PS and CM moieties, aiming to attain a segmentally labelled *Pni*DAH7PS^{WT} and probe its structural features using SANS and NMR techniques.

Prior to being subjected to structural investigations, the biochemical characteristics of the ligated protein *Pni*DAH7PS^{D-CM} were examined and it was verified that the reconstruction of the bifunctional system greatly restored both the DAH7PS and CM catalytic activities and restored the prephenate-induced allostery (Section 5.4). These

inspiring results further demonstrate that the full enzymatic function of *Pni*DAH7PS is delivered by a specific DAH7PS-CM hetero-interaction.

Thereafter, neutron variation contrast examinations were carried out using a CM-deuterated *Pni*DAH7PS^{D-CM} and resulted in clear scattering data for either the DAH7PS or CM domain (Figure 5.10 and Table 5.2). Analysis of the acquired SANS scattering data gave evidence for speculation on the spatial arrangement of *Pni*DAH7PS based on the SAXS profiles—namely two separated DAH7PS barrels are spanned by a CM dimer core. However, it was unfortunate that the expected conformational changes relating to substrate binding and allostery were not observed during the in SANS experiments. It is suspect that this may have been caused by a combination of an interaction between *Pni*DAH7PS^{D-CM} and the unseparated CM reactant and the high salt concentration (500 mM NaCl) in the SANS buffer that disrupts PEP/Mn²⁺ binding. Further efforts have attempted to probe this issue and a structural response of *Pni*DAH7PS^{D-CM} to PEP/Mn²⁺ and prephenate were subsequently observed in the SAXS profiles (Figure 5.12 and Table 5.3). A further investigation into the conformational changes of this ligated protein, in the presence of PEP/Mn²⁺ and prephenate, will be executed by SANS in future work.

Likewise, we are striving for coupling EPL-mediated segmentally labelling with NMR techniques to dissect the DAH7PS-CM interfaces and to gain deeper insight into how the hetero-domain interactions support the enzymatic functions. Currently, with the help of Prof. Joel Mackay, we are trying to complete the backbone assignment for the residues constituting the DAH7PS domain, which would be the basis for all subsequent NMR assays.

6.7 Conclusion

DAH7PS and CM respectively initiate the main pathway and one branch of the shikimate pathway, meaning that they are the metabolically appropriate enzymes for housing regulatory machinery for the modulation of the biosynthesis of aromatic amino acids. Besides the monofunctional, regulated, individual DAH7PS and CM enzymes (22,24,108,113), a number of covalently or non-covalently associated DAH7PS-CM complexes have been reported. This suggests bifunctional enzymes have evolved as a contemporary strategy for delivering catalysis or/and allostery (8,20,30,54,177). In all known bifunctional DAH7PS-CM enzymes, the CM domain shows a functional independence from the DAH7PS domain, but imposes feedback inhibition on DAH7PS activity (8,20). In comparison, as is seen in the well-studied DAH7PS-CM non-covalent complexes, the DAH7PS unit provides catalytic activity and bestows inhibitory regulation to its CM partner (30,54). In both cases, the DAH7PS enzymes possess normal catalytic activity independently and no bi-directional effects were observed between the DAH7PS and CM functions. However, *Pni*DAH7PS, which is a member of the C-terminal CM-linked type I β DAH7PSs, has been shown as an exception in this study—the DAH7PS and CM functions are fundamentally inter-dependent on one another and a dynamic hetero-interaction between the two domains was demonstrated to be the structural basis for buttressing the normal activities of both the DAH7PS and CM. Moreover, the dynamic hetero-domain interaction was shown to be manipulated by either the DAH7PS reaction or prephenate binding, which delivers allosteric activation or inhibition to the CM or the DAH7PS function. The combination of the feedforward activation by the DAH7PS reaction and the feedback inhibition by prephenate in *Pni*DAH7PS indicates that a sophisticated regulation cycle may be established in *P. nigrescens* (Figure 3.14). This suspected cycle makes it possible for *Pni*DAH7PS to modulate the metabolic flow of shikimate pathway more dynamically and precisely via sensing both the synthesis of DAH7P and the production of prephenate.

As well as a possible significance in metabolic control, the dynamic hetero-domain interaction of *Pni*DAH7PS itself is a very interesting phenomenon; displaying a distinctive molecular mechanism for acquiring catalytic functions and allosteric regulation. Although this thesis was not able to depict the dynamic DAH7PS-CM interaction at the atomic level, nevertheless the developed EPL coupled SANS and NMR techniques could potentially elucidate this hetero-domain cooperation in more depth in the future.

Chapter 7 *Experimental procedures*

7.1 General experimental details

7.1.1 Multiple sequence alignments and CLuster ANalysis of Sequences

Seed sequences of DAHP synthetase I (PF00793), representing type I β and type I α DAH7PS and KDO8PS, were extracted from the Pfam database (178) and aligned using Clustal Ω (179). The aligned seed sequence file was then submitted to jackhammer, a tool of the HMMER web server (180), to generate a hidden Markov model (HMM) for each multiple sequence alignment, which was in turn used to retrieve type I DAH7PS and KDO8PS sequences from four databases: UniProt (181), RefSeq (182), Pfamseq (178) and NR (182) using an E-value sequence cut-off of 10^{-15} . For each database, the scan was iterated until convergence. After removing 90% of the redundancy by Jalview (183), candidate sequences were processed with the clustering method implemented in CLuster ANalysis of Sequences (CLANS) (110). Following an all-against-all BLAST search of the sequences, the force directed pairwise similarities clustering algorithm was run for more than 5000 iteration cycles at a P-value of 10^{-40} , yielding three clusters that were identified as type I β , type I α , and KDO8PS, rendered as 3D graphs. After removal of the outliers, the sequences of type I β and KDO8PS were extracted and put into CLANS again in order to obtain more distinct clusters. Sequence alignments were generated with T-Coffee (184) and rendered with ESPript 3 (185).

7.1.2 Culture media

Lysogeny broth (LB) medium was used for the standard culture of *E.coli* cells for gene cloning and protein expression. LB medium and LB-agar were prepared respectively by dissolving 20 g/L LB base (Lennox L) and 37 g/L LB-agar blend (Miller's) in Milli-Q water and were sterilised by autoclaving. Appropriate antibiotic(s) were added before use. For the preparation of LB agar plates, the solidified LB agar was heated until thoroughly melted using a microwave oven and left in a 50 °C oven for cooling. The LB-agar solution was then supplemented with the appropriate antibiotic(s) and poured into

Petri dishes.

Super optimal broth with catabolite repression (SOC) is a very rich medium, which helps competent cell recover during transformation (184). SOC was prepared by two steps: Firstly, super optimal broth (SOB) medium consisting of 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl and 10 mM MgCl₂ was prepared and adjusted to pH 7.5 prior to sterilisation by autoclave; secondly, SOB was mixed with 20 mM filter sterilised glucose to form SOC.

7.1.3 Gene cloning

7.1.3.1 Gene synthesis

The gene (*PheB*) encoding *PniDAH7PS*^{WT} was codon-optimised for expression in engineered *E.coli* cells. An *NedI* restriction site and the nucleotide sequence encoding a TEV cleavage site were added at the 5' end of the optimised *PheB* gene, while an extra stop codon (TGA) and an *XhoI* restriction site were added at the 3' end (Table 7.1). The assembled DNA sequence was then synthesised by Geneart (Thermo Fisher Scientific, Inc.) and cloned into a pMA-T plasmid (pMA-T-*PheB*).

7.1.3.2 Polymerase chain reaction

Target genes were amplified using the polymerase chain reaction (PCR) technique. PCR experiments were performed using a Veriti® 96-well Thermal Cycler (Applied Biosystems). A typical PCR reaction volume was 50 µL and the components were approximately 0.02 ng/µl of template, usually purified plasmid containing the target gene, 0.5 µM of each of the forward and reverse primers (Table 7.1), 200 µM of each of the deoxyribonucleotide triphosphates (dNTPs), 10 µL of 5× Phusion HF Buffer, 0.02 U/µL Phusion DNA polymerase (Thermo Fisher Scientific Inc) and sterile Milli-Q water

for making up the remaining volume. PCR primers were synthesised by life technologies and dissolved in sterilised Milli-Q water to a concentration of 100 µM.

Table 7. 1: Primers for construction of expression vectors

Construction of pE28a-<i>Pni</i>DAH7PS
Forward primer: 5' CGCGCGGCAGC CATATG GAAAACCTGTATTTTCAGGG 3'
Reverse primer: 5' GGTGGTGGTG CTCGAG TCATCAATTCAGCATTGTC 3'
Construction of pDEST17-<i>Pni</i>DAH7PS^D 1st round PCR
Forward primer: 5' GGCAGCGGCGCGATGGAAGTGAAGTGCAG 3'
Reverse primer: 5' GAAAGCTGGGTGTCATCATTTATTACGAACAATCAGCTGGCTC 3'
Construction of pDEST15-<i>Pni</i>DAH7PS^{CM} 1st round PCR
Forward primer: 5' GGCAGCGGCGCGCATACCACAGTGAAGAA 3'
Reverse primer: 5' GAAAGCTGGGTGTCATCAATTCAGCATTGTCAGTTGCTG 3'
Constructions of pDEST17-<i>Pni</i>DAH7PS^D and pDEST15-<i>Pni</i>DAH7PS^{CM} 2nd round PCR
Forward primer: 5'GGGG ACAAGTTTGTACAAAAAAGCAGGCT TCGAAAACCTGTATTTTCAGGGCAGCGGCGCG 3'
Reverse primer: 5' GGGG ACCACTTTGTACAAGAAAGCTGGGT 3'
Construction of pTXB1-<i>Pni</i>DAH7PS^D for EPL
1st round Forward primer: 5' AACCTGTATTTTCAGGGCAGCGGCGCGATGGAAGTGAAGTGCAGCCGCTG 3'
2nd round Forward primer: 5' T CATATG CATCATCATCATCACGAAAACCTGTATTTTCAGGGCAGC 3
Reverse primer: 5' GGCTCTTC CGCAATGTTTATTACGAACAAT 3'
Construction of pTXB1-<i>Pni</i>DAH7PS^D for EPL
Forward primer: 5' GAAAACCTGTATTTTCAGTGTACAGTGAAGAACTGCACCAG 3'
Forward primer: 5' CTGGTGCAGTTCTTCACTGGTACACTGAAAATACAGGTTTTC 3'
Construction of <i>Pni</i>DAH7PS^{Δ5AA}
Forward primer: 5' AAAATGTGCACGTCAGGCAAATGCAAAATTTGCAGCACGTATTCTGGAAA 3'
Reverse primer: 5' TTTCCAGAATACGTGCTGCAAATTTGCATTGCTGACGTGCACATTTT 3'

CATATG: *Nde*I restriction site; **CTCGAG:** *Xho*I restriction site; **GCTCTTC:** *Sap*I restriction site

ACAAGTTTGTACAAAAAAGCAGGCT: *attB1* site;

ACCACTTTGTACAAGAAAGCTGGGT; *attB2* site.

A typical thermal cycling programme of the PCR consisted of an initial denaturation at 95 °C for 1 min and 30 cycles of denaturation for 30 sec, primer annealing for 30 sec at a proper temperature recommended by the primer supplier (Life technologies), and

extension at 72 °C for 45 sec, followed by a final extension at 72 °C for 10 min. The PCR product was purified using a NucleoSpin® Gel and a PCR Clean-up Kit (MACHEREY-NAGEL). The concentrations of PCR products were determined by measuring absorbance at 260 nm using a NanoDrop® ND-1000 spectrophotometer.

7.1.3.3 Agarose gel electrophoresis

Sizes of linear DNA were examined using agarose gel electrophoresis. Agarose gels were prepared by heating a mixture of 1% (w/v) LE agarose (Seakem) and Tris-acetate EDTA (TAE) buffer (50 mM Tris, 20 mM acetic acid, and 1 mM EDTA) until all agarose was dissolved. SYBR® Safe DNA gel stain (Invitrogen) was then added to the mixture before the gel was cast. Six times loading buffer, containing 60 mM Tris-HCl, 60 mM EDTA, 0.2% (w/v) Orange G, 0.05% (w/v) xylene cyanol FF, 60% (v/v) glycerol, was mixed with DNA to prepare the electrophoresis samples. The electrophoresis was conducted at 85 V in TAE buffer for 45 min.

7.1.3.4 Infusion

In-Fusion HD Cloning technology (Clontech, Inc) was employed to incorporate the amplified *PniDAH7PS* gene into a pET28a expression vector. In-Fusion is designed for fast and directional cloning, which requires a ~15 bp overlap between the target DNA and the linearised vector at each end. Via PCR, the respective overlapping sequences were added at each end of the PCR product for the *PniDAH7PS* gene. The linearised pET28a vector was prepared by performing a restriction digestion using *NdeI* and *XhoI* enzymes (NEB). In-fusion reactions contained 75 ng of purified PCR product of the *PniDAH7PS* gene, 75 ng of linearised pET28a vector, 2 µL 5× In-Fusion HD Enzyme Premix. The total reaction volume was adjusted to 10 µL using sterile Milli-Q water and the reaction was thoroughly mixed. The reaction was incubated at 50 °C for 15 min and then transformed into Stellar™ (Clontech) competent cells (Stellar cells).

7.1.3.5 Restriction digestion and DNA ligation

Double restriction enzyme digestions were used for generating the linearised vector and PCR product in both In-fusion and digestion-ligation cloning.

To generate the linear pET28a for infusion, the digestion reaction occurred in a 30 µL mixture containing approximately 50 ng/µL of pET28a plasmid, 3 µL of 10× CutSmart® Buffer (NEB), 20 U/µL of each of *NdeI* and *XhoI* restriction enzyme (NEB), 1 µL each. The remaining volume was made up with sterile Milli-Q water. The mixture was incubated at 37 °C for 2 hours, followed by inspection using agarose gel electrophoresis. Digested pET28a was purified using a NucleoSpin® Gel and was then subjected to NanoDrop for the determination of concentration.

The construction of pTXB1-*Pni*DAH7PS^D was carried out using digestion-ligation cloning. Approximately 2 mg of PCR product for *Pni*DAH7PS^D and ~3 mg of pTXB1 plasmid respectively were incubated with 20 U of *NdeI* and 10 U of *SapI* (NEB) in 20 µL × CutSmart® Buffer (NEB) at 37 °C for 2 hours. The digestion products were checked by agarose gel electrophoresis and purified using a NucleoSpin® Gel. After determining the concentrations, 36 ng of the digested PCR product of *Pni*DAH7PS^{CM-T263C} and 100 ng of pTXB1, ~3:1 insert:vector molar ratio, were mixed and incubated with 400 U of T4 DNA Ligase (NEB) in 20 µL 1× T4 DNA Ligase Buffer (NEB) at 16 °C overnight. Thereafter, 5 µL of the resultant reaction mixture was transformed into Stellar cells.

7.1.3.6 Gateway

Gateway® Technology was utilised for cloning the ORFs of the *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} truncated variants. Gateway® Technology uses phage lambda-based

site-specific recombination, instead of restriction endonucleases and ligase, to achieve the insertion of a DNA fragment into an expression vector (120,186).

Step 1: Amplification of the *PniDAH7PS^D* and *PniDAH7PS^{CM}* genes

In the Gateway cloning system, two rounds of PCR are required to attach an attB site (recombination site) at either end of the gene of interest. The primers used for gateway cloning are shown in Table 7.1.

The first round PCR reaction occurred in a 50 µL mixture consisting of approximately 0.02 ng/µl of the pMA-T *PniDAH7PS* plasmid, 0.5 µM of each 1st round PCR primer, 200 µM of each of the dNTPs, 10 µL of 5× Phusion HF Buffer, 0.02 U/µL Phusion DNA polymerase and sterile Milli-Q water for making up the remaining volume.

After the first round of PCR was completed, the product was treated with the restriction enzyme *DpnI* to digest the template pMA-T *PniDAH7PS* plasmid and then used as a template for a second round of PCR. The components of the second round PCR reaction were similar to those of the first round, except that the template and primers were changed to 0.5 µL 1st PCR product (~2 ng) and 2nd PCR primers (Table 7.1).

The thermal cycling conditions for the 1st round PCR for both truncated variants were the same, which consisted of an initial denaturation at 95 °C for 1 min and 30 cycles of denaturation for 30 sec, primer annealing at 58 °C for 30 sec and extension at 72 °C for 45 sec followed by a final extension at 72 °C for 10 min. The 2nd round PCR followed the same protocol for both variants, the only difference was that the annealing temperature was 55 °C.

Step 2: BP recombination reaction

The 2nd PCR product contained attB sites, which can recognize the attP sites of the entry vector pDONR221. Under the catalysis of the Gateway® BP Clonase® II enzyme

mix, the 2nd PCR product was ligated into the pDONR221 through recombination. The BP reaction contained a 10 µL mixture consisting of amplified DNA fragments (~112 ng for *PniDAH7PS^D* and ~133 ng for *PniDAH7PS^{CM}*), 150 ng pDONR221 vector, 2 µL BP Clonase® II enzyme mix and TE buffer (10 mM Tris/HCl, 1 mM EDTA pH 8) for the remaining volume. The reactions were incubated at 25 °C for 1 hour, followed by treatment of Proteinase K at 37 °C for 15 min to inactivate the BP Clonase.

Step 3: Transformation of the BP reaction and DNA sequencing

The BP recombination reaction was transformed into TOP 10 competent cells.

Step 4: The LR recombination reaction

After confirmation of the sequence of the gene carried by the pDONR221 vector, the plasmid was used to conduct an LR recombination reaction to transfer the cloned gene from the entry vector to an expression vector. The LR reaction consisted of the pDONR221 vector containing the cloned gene (~112 ng for *PniDAH7PS^D* and ~106 ng for *PniDAH7PS^{CM}*), 150 ng of the expression vector (pDEST17 for *PniDAH7PS^D* and pDEST15 for *PniDAH7PS^{CM}*), 2 µL LR Clonase® II enzyme mix and TE buffer for remaining volume. The reactions were incubated at 25 °C for 1 h, followed by treatment of Proteinase K at 37 °C for 15 min to inactivate the LR Clonase.

Step 5: Transformation of LR recombination reaction

The LR reactions for *PniDAH7PS^D* and *PniDAH7PS^{CM}* were respectively transformed into BL21 (DE3) CC3 cells and BL21 (DE3) pLysS cells for expression.

7.1.3.7 DNA transformation

Competent cells (Table 7.2) were taken out of storage at -80 °C and thawed on ice. One to 5 µL of each plasmid or cloning reaction mixture (such as In-fusion, BP and LR reaction mixture) was added into 50 µL of competent cells and gently mixed. The

competent cell/DNA mixture was incubated on ice for 20–30 min, followed by a heat shock at 42 °C water bath for 45 sec and put back on ice for 2 min. After the addition of 250 µL of SOC medium, the cells were recovered at 37 °C, with shaking, for 45 min. Subsequently, the transformed cells were plated onto LB agar plates containing the appropriate antibiotics and incubated at 37 °C overnight.

Table 7. 2: The competent cells used in this project

Strain	Species	Transformed plasmid	Antibiotics resistance	Characteristics
Strains for Cloning and Subcloning				
Stellar cells	<i>E. coli</i>	None	None	Lack the gene cluster for cutting foreign methylated DNA
TOP 10 cells	<i>E. coli</i>	None	None	Lack the gene cluster for cutting foreign methylated DNA
Strains for Protein Expression				
BL21 (DE3)	<i>E. coli</i>	None	None	T7 expression
BL21 (DE3) pLysS (pLysS)	<i>E. coli</i>	pLysS	chloramphenicol (CAM)	T7 expression; pLysS expresses low levels of T7 lysozyme, attenuating leaky expression.
BL21 (DE3) chaperone combination 3 (CC3)	<i>E. coli</i>	pBB528 pBB541	CAM, spectinomycin (Spec)	T7 expression; pBB528 expresses LacIq to achieve more stringent control of protein expression; pBB541 expresses GroEL and GroES to assist protein folding.

7.1.3.8 Colony PRC

Colony PCR is a simple way to check whether the plasmid construct in a transformant contains the gene of interest. This technique basically follows the same procedures of normal PRC, except that a small portion of a colony is used as the template instead of purified plasmids. To make sure that cell lysis occurs completely, an elongated denaturation step was added at the start of PCR thermocycle.

7.1.3.9 Plasmid extraction and purification

Five millilitres of cell culture containing the target plasmid were incubated overnight with the appropriate antibiotic(s). The cells were harvested by high-speed centrifugation (15,000 *g*) using 1.6 mL Eppendorf tubes. Plasmids (Table 7.3) were isolated using a High Pure Plasmid Isolation Kit (Roche) and following the kit instructions. The concentration of purified plasmid was measured by absorption at 260 nm using an appropriately blanked Nanodrop ND-1000 spectrophotometer.

Table 7. 3: Expression constructs generated in this project

Construct	Antibiotics resistance	Tag
pET28a- <i>Pni</i> DAH7PS ^{WT}	kanamycin (Kan)	6 x His (His)
pDEST17- <i>Pni</i> DAH7PS ^D	ampicillin (Amp)	His
pDEST15- <i>Pni</i> DAH7PS ^{CM}	Amp	glutathione S-transferase (GST)
pET28a- <i>Pni</i> DAH7PS ^{Δ5AA}	Kan	His
pTBX1-His- <i>Pni</i> DAH7PS ^{D-H262}	Amp	chitin binding domain (CBD)
pDEST15- <i>Pni</i> DAH7PS ^{CM-T263C}	Amp	GST

7.1.3.10 DNA sequencing

A DNA sequencing service was provided by the Massey Genome Service (MGS), or MacroGen DNA Sequencing Service. Approximate 100 ng/μL of purified plasmid DNA containing the gene of interest was supplied for each sample. All expression vectors in this project contain a T7 promoter and terminator, thus, the T7 forward and reverse primers were used for sequencing.

T7 Forward primer: 5' TAATACGACTCACTATAGGG 3'

T7 Reverse primer: 5' CTAGTTATTGCTCAGCGGTG 3'

7.1.3.11 Site-directed mutagenesis

Site-directed mutagenesis for generating constructs expressing *PniDAH7PS*^{Δ5AA} and *PniDAH7PS*^{CM-T263C} was performed following procedures similar to those of standard PCR. A typical PCR mutagenesis reaction had a volume of 50 μL and the components included ~1 ng/μL of purified plasmid template, 5 μM of each of the forward and reverse primers (Table 7.1), 200 μM of each of the deoxyribonucleotide triphosphates (dNTPs), 10 μL of 5× Phusion HF Buffer, 0.02 U/μL Phusion DNA polymerase and sterile Milli-Q water to make up the remaining volume. PCR primers were synthesised by life technologies and dissolved in sterilised Milli-Q water to a concentration of 100 μM.

The thermal cycling programme of the PCR consisted of an initial denaturation at 95 °C for 1 min and 18 cycles of denaturation for 30 sec, primer annealing for 30 sec at the recommended temperature and extension at 72 °C for 7 min, followed by a final extension at 72 °C for 10 min. The methylated parental DNA (template DNA) in the reaction mixture then was digested by *DpnI* (NEB) at 37 °C for 1 h. After the inactivation of *DpnI* at 80 °C for 20 min, 5 μL of the reaction mixture was transformed into Stellar cells.

7.1.4 Isopropyl β-D-thiogalactopyranoside induced expression

Expression cell strains were inoculated into fresh LB medium containing the appropriate antibiotic(s) and grown at 37 °C overnight with shaking at 180 rpm. Twenty millilitre of this overnight pre-culture was inoculated into 1 L of LB medium containing the appropriate antibiotic(s) and grown in 3 L baffled conical flasks at 37 °C with shaking at 180 rpm until the OD₆₀₀ value reached 0.4–0.8. Once the OD₆₀₀ reached ~0.5, 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) was added into each cell culture before the culture was transferred to 23 °C and grown with shaking at 180 rpm overnight for inducing protein expression.

7.1.5 Harvesting cells

Cell cultures having a large volume (> 500 mL) were harvested by centrifugation in 1 L centrifugation bottles using a Fiberlite™ F9-6x1000 LEX fixed-angle rotor (Thermo Fisher Scientific) at 12000 *g* for 20 min at 4 °C. Fifty millilitre Falcon tubes and a Fiberlite™ F14-14x50cy fixed-angle rotor (Thermo Fisher Scientific) were used to harvest relatively small cultures. For small cultures, less than 5 mL, centrifugation in 1.6 mL microcentrifuge tubes at 17000 *g* for 1 min was utilised. Harvested cell pellets were either immediately lysed or stored at -80 °C.

7.1.6 Cell lysis

Cells were lysed by sonication. Sonication was performed using an Omni-Ruptor 4000 Ultrasonic Homogenizer (Omni International). Cell pellets were resuspended in 20 mL lysis buffer (Table 7.4) containing half a tablet of a protease inhibitor cocktail tablet (Roche). The suspended cells were placed in a beaker surrounded with ice and sonicated in 4–5 rounds of 5 min, at 70% power using the 3/8 inch probe with the pulser set at 40.

Table 7. 4: Compositions of cell lysis buffer

Cell strain	Buffer composition
pET28a- <i>Pni</i> DAH7PS ^{WT} /CC3 pET28a- <i>Pni</i> DAH7PS ^{Δ5AA} /CC3 pDEST17- <i>Pni</i> DAH7PS ^D /CC3	50 mM Bis-tris propane (BTP) at pH 7.4, 200 mM KCl, 200 μM PEP and 2mM dithiothreitol (DTT)
pDEST15- <i>Pni</i> DAH7PS ^{CM} /pLysS pDEST15- <i>Pni</i> DAH7PS ^{CM-T263C} /pLysS	50 mM BTP at pH 7.3, 200 mM KCl and 2mM DTT
pTBX1-His- <i>Pni</i> DAH7PS ^{D-H262} /CC3	20 mM Tris-HCl at pH 8.0, 500 mM NaCl, 1 mM EDTA, 0.1% Triton X-100 and 200 μM PEP

7.1.7 Chromatography

7.1.7.1 Immobilised metal affinity chromatography

After centrifugation, the clarified cell lysate (*PniDAH7PS*^{WT}, *PniDAH7PS*^{Δ5AA} or *PniDAH7PS*^D) was filtered and loaded onto a HiTrap TALON[®] crude column (filled with a cobalt-based affinity resin, GE Healthcare) pre-equilibrated with the binding buffer (20 mM sodium phosphate at pH 7.4, 500 mM NaCl and 10 mM imidazole), then eluted with the elution buffer (20 mM sodium phosphate at pH 7.4, 500 mM NaCl and 250 mM imidazole). The eluted fractions were pooled together and desalted, which was followed by TEV protease cleavage overnight at 4 °C for the removal of the His tag. Then the mixture was subjected to IMAC again to remove the His tag and any uncleaved protein. The protein of interest was collected from the flow through and concentrated.

7.1.7.2 Immobilised glutathione affinity chromatography

The clarified cell lysate of either *PniDAH7PS*^{CM} or *PniDAH7PS*^{CM-T263C} was filtered and loaded onto a GSTrap[™] HP column (GE Healthcare) pre-equilibrated with the binding buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄ and 1.8 mM KH₂PO₄, pH 7.3), then eluted with the elution buffer (50 mM Tris-HCl, and 10 mM reduced glutathione, pH 8.0). The eluted fractions were collected and desalted and then were subjected to TEV protease cleavage overnight at 4 °C for cleaving the GST tag. Once the cleavage was completed, the protein mixture was reloaded onto the GSTrap column again to remove the GST tag and any uncleaved protein. The protein of interest was collected from the flow through and concentrated.

7.1.7.3 Size exclusion chromatography

Size exclusion chromatography (SEC) is a very important method for protein purification whereby proteins are separated based on their molecular weights. Larger proteins are eluted earlier, while smaller proteins are eluted later. Several types of SEC columns were used for protein purification in this project. *PniDAH7PS*^{WT}, *PniDAH7PS*^{Δ5AA} and *PniDAH7PS*^D were purified using a Hiloal 26/60 Superdex 200 pg column (GE Healthcare) at a flow rate of 1 mL/min using a buffer containing 10 mM BTP at pH 7.4, 150 mM NaCl and 200 μM PEP. *PniDAH7PS*^{CM} was purified using a HiLoad 16/60 Superdex 75 column (GE Healthcare) at a flow rate of 0.5 mL/min using a buffer containing 10 mM BTP at pH 8.1, 150 mM NaCl and 1 mM DTT. *PniDAH7PS*^{D-CM} and *PniDAH7PS*^{CM-T263C} were purified using a Hiprep 16/60 Sephacryl s-200 HR column (GE Healthcare) at a flow rate of 0.5 mL/min using a buffer containing 10 mM BTP at pH 7.4, 500 mM NaCl and 1 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl).

7.1.7.4 Anion exchange chromatography

Anion exchange chromatography (AEC) was adopted to try to separate ligated *PniDAH7PS*^{D-CM} from *PniDAH7PS*^{CM-T263C}. A Tricorn 10/100 column was packed with SOURCE™ 15Q resin (SOURCE-Q column) and equilibrated with a no-salt buffer (10mM BTP at pH 6.8 and 1 mM EDTA). The ligation product was transferred into the no-salt buffer to remove free Cl⁻ anions and generate a net negative charge on *PniDAH7PS*^{D-CM} (*PniDAH7PS*^{D-CM}, pI = 6.39; *PniDAH7PS*^{CM-T263C}, pI = 7.01). The pre-treated sample was then loaded onto the pre-equilibrated SOURCE-Q column. Elution from the column was performed using a linear gradient from 0 to 1 M NaCl at a flow rate of 1 mL/min over 5 column volumes (~ 40 mL).

7.1.7.5 Hydrophobic interaction chromatography

Hydrophobic interaction chromatography (HIC) was also used to separate ligated *Pni*DAH7PS^{D-CM} from *Pni*DAH7PS^{CM-T263C}. A Tricorn 10/100 column was packed with SOURCETM15Phe resin (SOURCE-Phe column) and equilibrated with a high-salt buffer (20 mM BTP at pH 7.4 and 1 M (NH₄)₂SO₄). (NH₄)₂SO₄ (1 M) was added into the ligation product with gentle stirring. The solution was then filtered and loaded onto the pre-equilibrated SOURCE-Phe column. Elution was performed using a linear gradient from 1 to 0 M (NH₄)₂SO₄ over 50 mL.

7.1.8 TEV protease preparation and TEV cleavage

7.1.8.1 Preparation of TEV protease

The protease from Tobacco Etch Virus (TEV) was used to cut off the affinity tags, such as His and GST, from the target proteins. His tagged TEV protease was expressed upon the induction by IPTG from the *E.coli* BL21 cells carrying the expression vector of TEV protease as described above. Subsequently, the protease was isolated using a HisTrapTM HP column (filled with a nickel-based affinity resin, GE Healthcare) using binding buffer (50 mM sodium phosphate at pH 7.4, 500 mM NaCl, 20 mM imidazole) and elution buffer (50 mM sodium phosphate at pH 7.4, 500 mM NaCl and 500 mM imidazole). After replacing the elution buffer with storage buffer (20 mM sodium phosphate at pH 8.0, 100 mM NaCl, 2 mM EDTA, 10 mM DTT and 20% (v/v) glycerol), the purified His-TEV protease was concentrated to 1 mg/mL and divided into 1 mL aliquots that were stored at -80 °C.

7.1.8.2 TEV cleavage reaction

The thawed TEV stock(s) was directly added into the target protein solution to make a

concentration ratio of 1:10–1:100 between TEV and the target protein. Following an overnight reaction at 4 °C (or a longer duration for sufficient cleavage), the protein solution was subjected to IMAC to remove the TEV protease.

7.1.9 Concentration and storage of enzymes

Protein solutions were concentrated using either a 10 kDa molecular weight cut off (MWCO) device Vivaspin® Turbo 15 (Sartorius) or a Macrosep Advance centrifugal device (PALL) with 3 kDa MWCO via centrifugation at 4000 *g*, 4 °C. Concentrated protein solutions were filtered and divided into 200 µL PCR tubes with a small volume (< 100 µL). Protein aliquots were snap frozen and stored at -80 °C.

7.1.10 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SDS-PAGE was performed using either Bolt® 10% Bis-Tris Plus 1.0 mm 12-well pre-caste protein gels (Life Technologies) or Bolt® 10% Bis-Tris Plus 1.0 mm 15-well pre-caste protein gels (Life Technologies) with Bolt® MES SDS Running Buffer (Life Technologies). Electrophoresis samples were prepared via boiling proteins mixed with 1× Bolt® LDS Sample Buffer (Life Technologies) and DTT for 5 min. Electrophoresis was carried out using a Bolt® Mini Gel Tank (Life Technologies) at 165 V for 35 min (MES). For the purpose of estimation of protein size, Novex® Sharp Pre-Stained Protein Standards (Invitrogen) were run in one lane together with the protein samples.

Visualisation of protein bands was attained by staining using a hot dye solution containing 0.1% (w/v) Coomassie brilliant blue R-250, 10% (v/v) glacial acetic acid, and 40% (v/v) methanol for 30 minutes with gentle shaking, followed by destaining using a hot solution containing 10% (v/v) glacial acetic acid, and 40% (v/v) methanol with gentle shaking until the protein bands become clearly visible.

7.1.11 Determination of protein concentration

Protein concentrations were determined using a Nanodrop® ND-1000 spectrophotometer by measuring and transforming the absorbance at 280 nm with the protein-specific molar extinction coefficient (ϵ_{280}) values. ϵ_{280} values for each protein were calculated using the ProtParam online tool (112) and are listed in Table 7.5.

Table 7. 5: Extinction coefficients (ϵ_{280}) for each protein

Protein	ϵ_{280}	Abs 0.1% (=1 g/L)
His- <i>Pni</i> DAH7PS ^{WT}	40450 M ⁻¹ cm ⁻¹	0.965
<i>Pni</i> DAH7PS ^{WT}	38960 M ⁻¹ cm ⁻¹	0.965
His- <i>Pni</i> DAH7PS ^D	43430 M ⁻¹ cm ⁻¹	1.318
<i>Pni</i> DAH7PS ^D	37470 M ⁻¹ cm ⁻¹	1.272
GST- <i>Pni</i> DAH7PS ^{CM}	52830 M ⁻¹ cm ⁻¹	1.329
<i>Pni</i> DAH7PS ^{CM}	1490 M ⁻¹ cm ⁻¹	0.133
His- <i>Pni</i> DAH7PS ^{Δ5AA}	38960 M ⁻¹ cm ⁻¹	0.957
<i>Pni</i> DAH7PS ^{Δ5AA}	38960 M ⁻¹ cm ⁻¹	0.984
His- <i>Pni</i> DAH7PS ^{D-H262} -intein-CBD	74370 M ⁻¹ cm ⁻¹	1.256
His- <i>Pni</i> DAH7PS ^{D-H262}	37470 M ⁻¹ cm ⁻¹	1.278
GST- <i>Pni</i> DAH7PS ^{CM-T263C}	52830 M ⁻¹ cm ⁻¹	1.343
<i>Pni</i> DAH7PS ^{CM-T263C}	1490 M ⁻¹ cm ⁻¹	0.138
His- <i>Pni</i> DAH7PS ^{D-CM}	38960 M ⁻¹ cm ⁻¹	0.972

Proteins with a low ϵ_{280} , such as *Pni*DAH7PS^{CM}, are vulnerable to inaccurate concentration readings due to the presence of other substances that absorb at 280 nm. Therefore, Bradford's assay is a more suitable method to measure the concentrations of proteins with a low ϵ_{280} , compared to UV spectrophotometry. Bovine serum albumin (BSA) solutions at 0, 0.2, 0.4, 0.6, 0.8, and 1 mg/mL were stained using the Bradford reagent and used as concentration standards. The

concentration of each standard was measured in triplicate at 595 nm to create a calibration curve and a regression equation. Target proteins were diluted to 0.2–0.8 mg/mL based on the measurement using Nanodrop and stained using the Bradford reagent as well. Following the measurement at 595 nm, the absorption values of the target proteins were put into the regression equation for the calculations of protein concentrations.

7.1.12 Mass spectrometry

Determination of protein molecular mass was performed by electrospray ionisation mass spectrometry (ESI-MS) using a Micromass LCT Classic, Bruker maXis 3G. Protein samples were prepared in MilliQ to a concentration of 1 mg/mL to reduce the influence of buffer components. Alternatively, 5 mM ammonium bicarbonate buffer was used for sample preparation.

7.1.13 Kinetics

A modified continuous assay for the spectrophotometric analysis of either DAH7PS or CM enzymatic activity (113) was employed to determine the kinetic parameters of the wild-type enzyme and the two truncated variants. Enzymatic reactions were carried out in 1 cm path length quartz cuvettes at 35 °C and the reaction rates were monitored by the disappearance of the substrate PEP at 232 nm or chorismate at 274 nm for the DAH7PS and CM activities respectively.

The determination of the metal dependency of *Pni*DAH7PS^{WT} and *Pni*DAH7PS^D was carried out using $\sim 7.3 \times 10^{-2}$ μ M enzyme and ~ 0.41 μ M enzyme respectively, using assays that contained 200 μ M PEP and 225 μ M E4P in the presence of 100 μ M of either MnSO₄, CdCl₂, CoCl₂, CuCl₂, CaCl₂, MgSO₄, NiCl₂, SrCl₂, PdCl₂ or EDTA.

The reaction mixture for determining the kinetic parameters for the DAH7PS domain of *PniDAH7PS^{WT}*, *PniDAH7PS^{Δ5AA}* and ligated *PniDAH7PS^{D-CM}* contained 100 μM MnSO₄, $\sim 7.3 \times 10^{-2}$ μM (*PniDAH7PS^{WT}*), $\sim 7.3 \times 10^{-2}$ μM (*PniDAH7PS^{Δ5AA}*) and $\sim 2.6 \times 10^{-2}$ μM (*PniDAH7PS^{D-CM}*) in 50 mM BTP buffer at pH 7.4. To determine the kinetic parameters, the PEP concentration was fixed at 338 μM and the concentration of E4P was fixed at 329 μM while conducting assays with varying concentrations of the other substrate.

The assays for determining the kinetic parameters for the CM domain of *PniDAH7PS^{WT}*, *PniDAH7PS^{Δ5AA}* and ligated *PniDAH7PS^{D-CM}* contained $\sim 7.3 \times 10^{-2}$ μM (*PniDAH7PS^{WT}*), $\sim 7.3 \times 10^{-2}$ μM (*PniDAH7PS^{Δ5AA}*) and $\sim 2.6 \times 10^{-2}$ μM (*PniDAH7PS^{D-CM}*) in 50 mM BTP buffer (pH 7.4) and were initiated by addition of chorismate at varying concentrations.

The assays for determining the kinetic parameters for *PniDAH7PS^D* contained 100 μM MnSO₄, 0.41 μM enzyme in 50 mM BTP buffer (pH 7.4). The kinetic parameters were examined by fixing the PEP concentration at 743 μM and the E4P concentration at 330 μM while varying the concentration of the other substrate.

The assays for determining the kinetic parameters of *PniDAH7PS^{CM}* contained 0.32 μM enzyme in 50mM BTP buffer at pH 7.4. To determine the apparent $K_M^{\text{chorismate}}$, varying concentrations of chorismate were used from 6 μM to 452 μM.

7.1.14 Feedback inhibition

The inhibition of the DAH7PS activity of *PniDAH7PS^{WT}*, *PniDAH7PS^D*, *PniDAH7PS^{Δ5AA}* and *PniDAH7PS^{D-CM}* was characterised by monitoring DAH7PS activity in the presence of varying concentrations of prephenate (0 – 315 μM). Reactions carried out at 35 °C and contained 100 μM Mn²⁺, 200 μM PEP, 225 μM E4P and 7.3×10^{-2} μM enzyme (*PniDAH7PS^{WT}*), 0.41 μM enzyme (*PniDAH7PS^D*), $\sim 7.3 \times 10^{-2}$ μM enzyme (*PniDAH7PS^{Δ5AA}*) or $\sim 2.6 \times 10^{-2}$ μM enzyme (*PniDAH7PS^{D-CM}*) in 50 mM BTP buffer (pH

7.4). The calculation of IC_{50} was carried out using the Logistic Fit (Origin Pro 8).

The product inhibition of prephenate on the CM activity of *PniDAH7PS^{WT}* was examined also using the continuous spectrophotometric assay in 50 mM BTP buffer (pH 7.4) at 35 °C. The K_i value for prephenate was determined by varying the prephenate concentration from 0–20 μ M, with a chorismate concentration range of 0–310 μ M.

7.1.15 Analytical ultracentrifugation

Sedimentation velocities of *PniDAH7PS^{WT}* and *PniDAH7PS^D* at varying concentrations were determined using a Beckman Coulter Model XL-I analytical ultracentrifuge equipped with a UV/Vis dual-beam UV/VIS spectrophotometer with a monochromator. The reference buffer solution (10 mM BTP at pH 7.4, 150 mM NaCl) and sample solutions were examined in 12 mm double sector aluminium centrepieces with quartz windows, which were mounted in a 4-hole An60Ti rotor. Samples (380 μ L/cell) and a buffer reference (400 μ L/cell) were centrifuged at 50,000 rpm at 20 °C and absorbance data were collected in continuous mode at 275 nm without averaging. The partial specific volume ($\bar{v}$) of the samples (0.7445 mL/g for *PniDAH7PS^{WT}* and 0.7479 mL/g for *PniDAH7PS^D*), buffer density (1.0049 g/mL) and buffer viscosity (1.0214 cP) were calculated using SEDNTERP (187). The data were fitted to a continuous size distribution model using the program SEDFIT (188).

7.1.16 Analytical size exclusion chromatography

Analytical size exclusion chromatography (Analytical SEC) was carried out using a Superdex 10/300 GL column (GE Healthcare) pre-equilibrated with the SEC buffer for *PniDAH7PS^{CM}*. Five hundred microliter of 0.5 mg/mL protein standards (GE Healthcare) in SEC buffer, Ribonuclease A (13.7 kDa), ovalbumin (43 kDa), conalbumin (75 kDa),

aldolase (158 kDa) and blue dextran (2000 kDa), were applied at a flow rate of 0.2 mL/min. Following the standard protein, a 500 μ L sample (*Pni*DAH7PS^{WT}, *Pni*DAH7PS^D, *Pni*DAH7PS^{D-CM} and *Pni*DAH7PS^{CM}) at concentrations in the range of 0.3-1.5 mg/mL was injected onto the column followed by elution with the SEC buffer. The elution volumes of blue dextran and other standard protein were respectively recorded as the void volume (V_o) and V_e . The column volume is denoted as V_c . Using the equation $K_{av} = (V_e - V_o) / (V_c - V_o)$, V_e values were transformed into K_{av} (a coefficient defining the proportion of the bead pores which can be occupied by a molecule) values, which were further plotted against log (standard protein mass (Da)) to generate a calibration curve and a regression equation. With the aid of this correlation between the elution volume and protein mass, the calculated mass of the sample was obtained, based on its elution volume.

7.1.17 Crystallography

7.1.17.1 Screening and optimisation of crystallisation conditions

Initial screening for crystallisation conditions was performed using a Mosquito[®] Crystal Robot (TTP Labtech) and 96-well sitting drop iQ plates (TTP Labtech) following the sitting drop vapour diffusion method. The crystallisation conditions screened in this project included the PACT premier[™] HT-96, JCSG-plus[™] HT-96, Clear Strategy[™] I HT-96 and Clear Strategy[™] II HT-96, MIDAS plus[™] HT-96, ProPlex[™] HT-96 and Morpheus[®] II screens (Molecular Dimensions). For each screening well, a drop of 400 nL protein solution and 400 nL screening condition sat beside a well containing 40 μ L of reservoir solution. Screening plates were sealed and incubated at either 20 °C or 5 °C. Crystal formation was checked and recorded regularly.

Once a hit condition was identified, the crystallisation was scaled up and optimised using 24-well VDX plates (Hampton Research) using the hanging drop vapour diffusion

method. For the purpose of optimisation, salt and precipitant concentrations, protein concentrations and pH values of a hit condition were varied and constituted a variety of reservoir solutions. Protein solutions and reservoir solutions were mixed at 1:1 ratio (v/v) to form 2 μ L or 4 μ L drops, which were subsequently hanged above their corresponding reservoir solutions (500 μ L). Plates were incubated at either 20 °C or 5 °C and examined regularly.

7.1.17.2 X-ray diffraction and data process

All X-ray diffraction data for PniDAH7PS^D were collected at the Australian Synchrotron (AS) using either the Macromolecular Crystallography (MX1) (189) or Micro Crystallography (MX2) beamlines (190) and were processed using X-ray Detector Software (XDS), Aimless, and Truncate (CCP4 suite) (126,128,191,192). Alternatively, reflections were integrated using iMosflm (193) instead of XDS. The initial phases were calculated by molecular replacement using MolRep (194) or Phaser (127) (CCP4 suite) with the aid of the *Pfu*DAH7PS crystal structure (PDB ID code: 4C1L; all ligand/waters were removed from this structure prior to molecular replacement). Once a reasonable phase solution was identified, the initial structural model was refined by rigid-body refinement and restrained refinement using Refmac5 (195,196) with iterative model building using Coot (197).

7.1.18 Small angle X-ray scattering

Small angle X-ray scattering (SAXS) experiments were conducted at the Australian Synchrotron (198,199), using the SAXS/WAXS beamline equipped with a Pilatus detector (1M, 170 $\times$ 170 mm, effective pixel size, 172 $\times$ 172 μ m). Samples were examined using X-rays with a wavelength of 1.0332, and the subsequently scattered X-rays intensities were recorded using a sample detector at a distance of 1.6 m from the sample, covering the q range of 0.0015-3.0 $\text{\AA}^{-1}$. Samples were eluted from a SEC

column (Superdex 200 5/150) using SEC buffers consisting of the conditions of interest which as listed below. The eluted sample passed into a 1.5 mm thin-walled glass capillary where the sample was examined using X-rays at 25 °C at 2 s intervals-

The reduction and buffer subtraction for the raw scattering data of protein sample were performed using Scatterbrain, developed at the Australian Synchrotron. Processed data were plotted as $I(q)$ vs. q ($q = (4\pi\sin\theta)/\lambda$, 2θ is the scattering angle), and analysed using Primus (200) and Origin 8.

The data sets for structural analyses were recorded with 447 data points over the range $0.005 \leq q \leq 0.35 \text{ \AA}^{-1}$. The radius of gyration (R_g) of the protein particle was evaluated from Guinier plots ($\ln I(q)$ vs. q^2) of scattering data. In addition, the linear Guinier plots in the q range of $q \geq 1.3/R_g$ was a crucial indicator of the monodispersity of the samples. The Porod volume (V_p) and the maximum dimension ($D_{\max}$) of the protein particle were evaluated respectively from the Porod plot and the pair-distance distribution ($P(r)$) function obtained via Fourier transformation. MM_{porod} was obtained by dividing the Porod volumes by 1.7.

The buffers for all SAXS experiments are listed below.

Buffer 1 (standard buffer): 20 mM BTP at pH 7.4, 150 mM NaCl and 3% (v/v) glycerol

Samples: *PniDAH7PS*^{WT}, *PniDAH7PS* ^{$\Delta 5\text{AA}$} and *PniDAH7PS*^{D-CM}

Buffer 2: 20 mM BTP at pH 7.4, 150 mM NaCl, 350 μM PEP and 3% (v/v) glycerol

Samples: *PniDAH7PS*^{WT}, *PniDAH7PS*^D and *GspDAH7PS*^{WT}

Buffer 3: 20 mM BTP at pH 7.4, 150 mM NaCl, 350 μM MnSO₄ and 3% (v/v) glycerol

Sample(s): *PniDAH7PS*^{WT}

Buffer 4: 20 mM BTP at pH 7.4, 150 mM NaCl, 350 μ M PEP, 350 μ M MnSO₄ and 3% (v/v) glycerol

Sample(s): *PniDAH7PS*^{WT}, *PniDAH7PS* ^{Δ 5AA} and *PniDAH7PS*^{D-CM}

Buffer 5: 20 mM BTP at pH 7.4, 150 mM NaCl, 350 μ M PEP, 350 μ M MnSO₄, 350 μ M E4P and 3% (v/v) glycerol

Sample(s): *PniDAH7PS*^{WT} and *PniDAH7PS* ^{Δ 5AA}

Buffer 6: 20 mM BTP at pH 7.4, 0 mM NaCl and 3% (v/v) glycerol

Sample(s): *PniDAH7PS*^{WT}

Buffer 7: 20 mM BTP at pH 7.4, 0 mM NaCl, 350 μ M PEP, 350 μ M MnSO₄ and 3% (v/v) glycerol

Sample(s): *PniDAH7PS*^{WT}

Buffer 8: 20 mM BTP at pH 7.4, 500 mM NaCl and 3% (v/v) glycerol

Sample(s): *PniDAH7PS*^{WT}

Buffer 9: 20 mM BTP at pH 7.4, 500 mM NaCl, 350 μ M PEP, 350 μ M MnSO₄ and 3% (v/v) glycerol

Sample(s): *PniDAH7PS*^{WT}

Buffer 10: 20 mM BTP at pH 7.4, 150 mM NaCl, 350 μ M prephenate and 3% (v/v) glycerol

Sample(s): *PniDAH7PS*^{WT}, *PniDAH7PS* ^{Δ 5AA}, *PniDAH7PS*^{D-CM} and *GspDAH7PS*^{WT}

Buffer 11: 20 mM BTP at pH 7.4, 150 mM NaCl, 350 μ M prephenate, 100 μ M PEP, 350 μ M MnSO₄ and 3% (v/v) glycerol

Sample(s): *PniDAH7PS*^{WT}

Buffer 12: 20 mM BTP at pH 7.4, 150 mM NaCl, 350 μ M prephenate, 500 μ M PEP, 500 μ M MnSO₄ and 3% (v/v) glycerol

Sample(s): *Pni*DAH7PS^{WT}

Buffer 13: 20 mM BTP at pH 7.4, 150 mM NaCl, 350 μ M prephenate, 350 μ M PEP, 350 μ M MnSO₄, 350 μ M E4P and 3% (v/v) glycerol

Sample(s): *Pni*DAH7PS^{WT}

Buffer 14: 20 mM BTP at pH 7.4, 0 mM NaCl, 350 μ M prephenate and 3% (v/v) glycerol

Sample(s): *Pni*DAH7PS^{WT}

Buffer 15: 20 mM BTP at pH 7.4, 500 mM NaCl, 350 μ M prephenate and 3% (v/v) glycerol

Sample(s): *Pni*DAH7PS^{WT}

Buffer 16: 20 mM BTP at pH 8.1, 150 mM NaCl and 3% (v/v) glycerol

Sample(s): *Pni*DAH7PS^{CM}

7.1.19 Modelling

***Ab initio* modelling**

The low resolution *ab initio* modelling for *Pni*DAH7PS^D, *Pni*DAH7PS^{CM}, *Pni*DAH7PS^{WT} and *Gsp*DAH7PS was performed using either *DAMMIF* (124) or *GSABOR* (148), based on the $P(r)$ function, to generate 20 dummy-residue models with imposed asymmetry (P1) (for *Pni*DAH7PS^D) or 2 fold (P2) symmetry (for *Pni*DAH7PS^{CM}, *Pni*DAH7PS^{WT} and *Gsp*DAH7PS). The models were subsequently averaged using *DAMAVAR* (201) to yield the most probable model for each protein.

Homology modelling

MODELLER (202) was utilised, using crystal structures of *Pfu*DAH7PS (PDB ID code: 4C1L) and *Pfu*CM (PDB ID code: 1YBZ) as templates, for the homology modelling of *Pni*DAH7PS^D and *Pni*DAH7PS^{CM}. The atomic structures of monomeric *Pfu*DAH7PS^D and dimeric *Pfu*DAH7PS^{CM} were generated in PyMOL. Homology models of monomeric *Pni*DAH7PS^D, and both monomeric and homodimeric *Pni*DAH7PS^{CM} were generated using MODELLER using the “automodel” routine to create 100 models. The best model was selected based on models exhibiting the lowest MODELLER objective function value and the lowest global DOPE score, followed by visual inspection and a residue-by-residue DOPE score analysis.

Rigid body modelling

Rigid body modelling against SAXS data was carried out by using BUNCH (132) and homology models of *Pni*DAH7PS^D and *Pni*DAH7PS^{CM}. Prior to running BUNCH modelling for the *Pni*DAH7PS^{WT} homodimer, an initial structural model of monomeric *Pni*DAH7PS^{WT} was preliminarily constructed by pre-bunch via linking the models of *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} together according to the sequence of *Pni*DAH7PS^{WT}. Sequentially, the initial structure was inputted into BUNCH together with the experimental SAXS data and calculated scattering amplitudes for the *Pni*DAH7PS^D and *Pni*DAH7PS^{CM} models. Then, 80–100 steps of simulated annealing, constrained by P2 symmetry and contact information for the *Pni*DAH7PS^{CM} dimer, were carried out (Table 3.2), which give rise to a rigid body model and a fit to the corresponding SAXS profile.

Table 7. 6: Contact information for the rigid body modelling of *Pni*DAH7PS^{D-CM} using

SASRAFCV		
Residue	Residue	Distance between C _α atoms (Å)
K261 (chain 1)	H1 (chain 3)	4.00
K261 (chain 2)	H94 (chain 3)	4.00

Rigid body modelling for visualising the combined information of the SAXS and SANS scatterings of *PniDAH7PS*^{D-CM} was performed using SASRAFCV (203). The SAXS and SANS (42% and 100% D₂O) scattering data and homology models of each subunit (two monomeric *PniDAH7PS*^D and a *PniDAH7PS*^{CM} homodimer) were used for running simulated annealing restrained by the distances of the inter-subunit contacts (Table 7.6) and the cross-correlation between the subunits and the scattering profiles (Table 7.7). In this modelling, the perdeuteration of the CM dimer was set at 75%. The maximum amount of annealing steps and symmetry were respectively set as 100 and P2, and the default values provided by SASREFCV were employed as were the other default parameters.

Table 7. 7: Cross-correlation information for the rigid body modelling of *PniDAH7PS*^{D-CM} using SASRAFCV

		Homology model of subunit		
		<i>PniDAH7PS</i> ^D	<i>PniDAH7PS</i> ^D	<i>PniDAH7PS</i> ^{CM} dimer
Scattering profile	SAXS data	0.0	0.0	0.0
	42% D ₂ O	0.0	0.0	0.75
	100% D ₂ O	0.0	0.0	0.75

7.1.20 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) assays were performed using a Nano Isothermal Titration Calorimeter (TA Instruments) at 278 K by directly measuring the heat change generated by each injection of a ligand into 400 µL of protein solution placed in the sample cell. Every ITC measurement in this project was carried out with 22 injections containing the first 0.42 µL injection and 21 subsequent 2.02 µL injections. Compositions and pH values of the buffers for preparing the ligand and protein solutions were well matched. All solutions were filtered and degassed, and the sample

cell and ligand syringe were washed several times with the degassed buffer immediately before the experiments. The experimental set-up for each ITC experiment is listed in Table 7.8.

Table 7. 8: ITC experimental set-up

Buffer	Protein (sample cell)	Ligand (syringe)
10 mM BTP (pH7.4), 150 mM NaCl	65 μ M <i>PniDHA7PS</i> ^{WT}	1.5 mM PEP
10 mM BTP (pH7.4), 150 mM NaCl and 0.5 mM MnSO ₄	65 μ M <i>PniDHA7PS</i> ^{WT}	1.5 mM PEP
10 mM BTP (pH7.4), 0 mM NaCl and 0.5 mM MnSO ₄	65 μ M <i>PniDHA7PS</i> ^{WT}	1.5 mM PEP
10 mM BTP (pH7.4), 500 mM NaCl and 0.5 mM MnSO ₄	65 μ M <i>PniDHA7PS</i> ^{WT}	1.5 mM PEP
10 mM BTP (pH7.4), 150 mM NaCl, 0.5 mM MnSO ₄ and 0.25 mM prephenate	65 μ M <i>PniDHA7PS</i> ^{WT}	1.5 mM PEP

The heat of dilution was measured independently for each experiment and subtracted from the integrated data using NanoAnalyze (TA Instruments). Curve fitting to an appropriate binding model and result evaluation were also performed using NanoAnalyze (TA Instruments).

7.2 Experimental details for Chapter 5

7.2.1 Expressed protein ligation

The expressed protein ligation (EPL) for generating *Pni*DAH7PS^{D-CM} was performed using His-*Pni*DAH7PS^{D-H262}-intein-CBD and *Pni*DAH7PS^{CM-T263C} with the induction of a thiol reagent, sodium 2-mercaptoethanesulfonate (MESNA), in a chitin affinity column. Chitin is a resin for the purification of target proteins fused to an intein-chitin binding domain (CBD) (149). His-*Pni*DAH7PS^{D-H262}-intein-CBD was expressed as described in chapter 7.1.4. After lysis of the cell pellet harvested from 1 L of cell culture by sonication using the buffer (20 mM Tris-HCl at pH 8.5, 500 mM NaCl, 1 mM EDTA and 1 mM TCEP), the soluble fraction of the cell lysate was filtered and loaded onto a 10 mL chitin column pre-equilibrated with the column buffer consisting of 20 mM Tris-HCl (pH 8.5), 500 mM NaCl and 1 mM EDTA. Subsequently, the column was washed using 20 bed volumes of column buffer, which was followed by an equilibration using 3 bed volumes of column buffer containing 50 mM MESNA. The purified *Pni*DAH7PS^{CM-T263C} was subjected to TEV cleavage, as described in chapter 7.1.7.2 and 7.1.7.3, and exchanged into column buffer supplemented with 50 mM MESNA and 1 mM TCEP and concentrated to ~2 mg/mL. The ligation reaction was initiated by the injection of *Pni*DAH7PS^{CM-T263C} solution into the MESNA pre-treated chitin column that was carrying His-DAH7PS-intein-CBD. The reaction was statically incubated at 25 °C for 2 days, followed by elution from the column using 3 bed volume of column buffer.

The harvested ligation product was isolated by SEC using a Hiperp 16/60 Sephacryl s-200 HR column (GE Healthcare) and IMAC using a HiTrap TALON[®] crude column (GE Healthcare) sequentially, following the protocols described in 7.1.7.

7.2.2 Small angle neutron scattering

Small angle neutron scattering (SANS) data were recorded using the SANS instrument

QUOKKA at the Australian Nuclear Science and Technology Organisation (ANSTO) (204,205) using a wavelength (λ) of 5.00 Å at 10 °C. Sample to detector distances of 2 m, 10 m and 12 m were adopted for covering a q range of ~ 0.003 – 0.6 Å⁻¹.

The *Pni*DAH7PS^{D-CM} with a segmentally deuterated CM domain was concentrated to ~ 3.5 mg/mL and exchanged into a buffer containing 20 mM BTP (pH 7.8), 500 mM NaCl, and 1 mM TCEP made from either 42% or 100% (v/v) D₂O. Data for the 0% (v/v) D₂O samples and buffers were collected for 2 h at all detector distances; data for the 42% (v/v) D₂O samples and buffers were collected for 4 h at all detector distances; data for the 100% (v/v) D₂O sample and buffer were collected for 1 h at all detector distances.

Data for each sample and buffer were combined and reduced to $I(q)$ versus q ($q = (4\pi\sin\theta)/\lambda$, 2θ is the scattering angle) using IGOR Pro 6.1 (WaveMetrics, Portland, OR) (206) with corrections for detector sensitivity, buffer and sample transmission, and background radiation. All data were in an absolute scale with the reference of the flux of the direct beam. The scattering data for proteins were obtained by subtracting the solvent scattering from the protein-solvent scattering. The scattering data analysis and manipulations were performed using PRIMUS and the $P(r)$ distributions were calculated using GNOM (207).

7.2.3 Isotopic labelling

IPTG induction

Uniform single labelling (¹⁵N), double labelling (¹³C, ¹⁵N), and triple labelling (²H, ¹³C, ¹⁵N) were performed via IPTG induced expression using an M9 minimal salts based growth medium (M9 medium, Table 7.9) (208).

Table 7. 9: The recipe for the M9 minimal medium

1.	978 or 977 mL M9 salt	per 978 or 977 mL
	KH ₂ PO ₄	13.0 g
	K ₂ HPO ₄	10.0 g
	Na ₂ HPO ₄	9.0 g
	K ₂ SO ₄	2.4 g
	¹⁵ NH ₄ Cl	1 g
	H ₂ O or D ₂ O (triple labelling)	978 or 977 mL
2.	1ml trace element solution (1000 ×)	per 100 mL
	FeSO ₄ (7 H ₂ O)	0.60g
	CaCl ₂ (2 H ₂ O)	0.60 g
	MnCl ₂ (4 H ₂ O)	0.12 g
	CoCl ₂ (6 H ₂ O)	0.08 g
	ZnSO ₄ (7 H ₂ O)	0.07 g
	CuCl ₂ (2 H ₂ O)	0.03 g
	H ₃ BO ₃	2 mg
	(NH ₄) ₆ Mo ₇ O ₂₄ (4 H ₂ O)	0.025 g
	EDTA	0.50 g
3.	10 ml 1 M MgCl ₂	
4.	6 ml thiamine hydrochloride (5 mg/mL)	
5.	D-Glucose or ¹³ C-D-Glucose (3 g)	
6.	Antibiotics: 1 mL CAM (25 mg/mL) and 1 mL Amp (100 mg/mL) OR 1 mL CAMl (25 mg/mL), 1 mL kan (100 mg/mL) and 1 mL Spec (100 mg/mL)	

Expression cell strains were inoculated into fresh M9 medium containing the appropriate antibiotic(s) and grown at 37 °C overnight with shaking at 180 rpm. For triple labelling, 300 µL of an overnight pre-culture was inoculated into 10 mL medium (containing appropriate antibiotics) of a mixture of LB and M9 medium (²H, ¹³C and ¹⁵N) with a ratio of 1: 1 (v/v) for the sake of growth adaption in D₂O. Cells were grown in this mixed medium at 37 °C, with shaking at 200 rpm, until the OD₆₀₀ value reached 0.4–0.8 and 1 mL cell culture was then inoculated into 10 mL 100% triple labelling M9 medium containing the appropriate antibiotic(s) and grown at 37 °C, with shaking at 200 rpm. Once the OD₆₀₀ value of the culture reached 0.4–0.8, 10 mL culture was enlarged into 1 L by the addition of 990 mL of fresh 100% triple labelling M9 medium

containing the appropriate antibiotic(s). The induction of expression was initiated by adding 0.5 mM IPTG into the cell culture when the OD₆₀₀ reached ~0.8. Subsequently, the culture was transferred to 20 °C and grown with shaking at 180 rpm for ~36 hrs. For single or double labelling, the adaption procedures were inessential, thus, 300 µL LB pre-culture was directly inoculated into 10 mL of isotopic labelling M9 medium. The subsequent procedures for the induction of expression were basically the same as that for the triple labelling, except that the over-expression was carried out overnight instead of for 36 h.

Autoinduction

Autoinduction (209) was employed for increasing the yield of ¹⁵N labelled *PniDAH7PS^{CM}* by using NG and N-5052 defined medium (Table 7.10). The glycerol stock of BL21 (DE3) pLysS cells carrying pDEST15-*PniDAH7PS^{CM}* was used to make a 5 mL overnight preculture in LB containing 25 mg/mL CAM and 100 mg/mL ampicillin (Amp). Twenty microlitres of overnight preculture were added into 2 mL of NG minimal medium containing 25 mg/mL CAM and 100 mg/mL Amp which was subsequently incubated at 37 °C for 6–8 h until the culture was visibly turbid. Subsequently, the entire 2 mL minimal culture was transferred into the remaining 48 mL of NG medium and growth was continued at 37 °C overnight. A 50 mL overnight NG culture was added into 1 L of N-5052 minimal medium containing 25 mg/mL CAM and 100 mg/mL Amp and grown at 30 °C for 24 h leading to protein expression.

Table 7. 10: The recipes for NG and N-5052 minimal medium

1.	NG defined medium for growth to saturation	per 50 mL
	20× N (minimal medium)	2.5 mL
	D-Glucose (50 w/v)	500 µL
	Trace element (1000×)	9.0 g
	1 M MgSO ₄	50 µL
	Thiamine hydrochloride (5 mg/mL)	0.3 ml
	CAM (25 mg/mL) and Amp (100 mg/mL)	50 µL each
	¹⁵ NH ₄ Cl	125 mg
	Milli-Q	50 mL
2.	N-5052 defined medium	per 1 L
	50 × 5052	20 mL
	20× N	50 mL
	Trace element (1000×)	1 mL
	1 M MgSO ₄	1 mL
	Thiamine hydrochloride (5 mg/mL)	6 mL
	CAM (25 mg/mL) and Amp (100 mg/mL)	1 mL each
	¹⁵ NH ₄ Cl	2.5 g
	Milli-Q	920 mL
3.	20× N minimal medium	per 1 L
	KH ₂ PO ₄	136 g
	Na ₂ HPO ₄	142 g
	K ₂ SO ₄	14.2 g
	Milli-Q	1 L
4.	50× 5052 solution	per 1 L
	Glycerol	250 g
	Glucose	25 g
	α-lactose	100 g

7.2.4 Nuclear magnetic resonance spectroscopy

Nuclear magnetic resonance (NMR) spectra were performed using a Bruker AvanceIII 600 or 800 MHz spectrometer (the School of Life and Environmental Sciences, Sydney University), both of which were equipped with a 5 mm triple resonance TCI cryogenic probe. Samples, with a concentration in the range of 0.2–0.25 mM, were examined in

their appropriate buffers (Table 7.11) in 10% D₂O/90% H₂O supplemented with 17 µM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as an internal reference.

Table 7. 11: The recipes of the buffers for NMR samples

Protein	Buffer composition
<i>Pni</i> DAH7PS ^{WT}	10 mM BTP, pH 7.4, 150 mM NaCl, 50 mM Glu&Arg and 10 mM DTT
<i>Pni</i> DAH7PS ^{CM}	10 mM BTP, pH 7.8, 150 mM NaCl, 50 mM Glu&Arg and 10 mM DTT
<i>Pni</i> DAH7PS ^D	10 mM BTP, pH 7.2, 150 mM NaCl and 1 mM DTT

Appendix

Supporting information for Chapter 2

Figure A.2. 1: Michaelis-Menten plots of the *PniDAH7PS*^{WT} for the K_M of PEP, E4P and chorismate respectively. Plotted and fitted globally using Graphpad Prism.

Figure A.2. 2: Michaelis-Menten plots of the *Pni*DAH7PS^D for the K_M of PEP and E4P, and of the *Pni*DAH7PS^{CM} for the K_M of chorismate respectively. Plotted and fitted globally using Graphpad Prism.

Figure A.2. 3: Crystals of *PniDAH7PS*^D (A) crystals growing in the conditions of Clear Strategy I screen suite; (B) crystals growing in the conditions of Clear Strategy II screen suite; (C) crystals growing in the conditions of JCSG plus screen suite.

Supporting information for Chapter 3

Table A.3. 1: SAXS parameters for apo-*Pni*DAH7PS^{WT} and the *Pni*DAH7PS^{WT} in the presence of either 350 μ M PEP or 350 μ M MnSO₄

	Apo-protein	PEP	Mn ²⁺
<i>Guinier Analysis</i>			
R_g (Å)	42.0 ± 1.2	42.9 ± 0.6	43.0 ± 0.5
$I(0)$ (cm ⁻¹)	0.08 ± 0.00	0.08 ± 0.00	0.08 ± 0.00
Correlation coefficient, R^2	0.99	0.99	0.99
<i>Pair Wise Distribution Analysis</i>			
R_g (Å)	44.1 ± 0.2	43.6 ± 0.1	44.6 ± 0.4
$I(0)$ (cm ⁻¹)	0.08 ± 0.00	0.08 ± 0.00	0.08 ± 0.00
D_{max} (Å)	152.3	158.0	155.5
V_p (nm ³)	137	131	142
<i>MM_{porod}</i>			
MM (kDa)	81	77	84
Number of Subunits	2	2	2

V_p : Porod volume; ***MM_{porod}***: molecular mass estimated from Porod volume. The uncertainties represent standard deviations estimated by the error propagation from the experimental data.

Figure A.3. 1: Analysis of the SAXS profiles of apo-*PniDAH7PS*^{WT} (black) and the *PniDAH7PS*^{WT} respectively in the presence of PEP (red) and Mn²⁺ (blue). (A) SAXS profiles ($\log I(q)$ versus q). (B) Guinier plots ($\ln I(q)$ versus q^2). (C) Kratky plot ($q^2 I(q)$ versus q). (D) Porod-Debye plot ($q^4 I(q)$ versus q^4) limited to the range of the SAXS data for which the Guinier linearity is observed.

Figure A.3. 2: Michaelis-Menten plots of the *PniDAH7PS*^{Δ5AA} for the K_M of PEP, E4P and chorismate. Plotted and fitted globally using Graphpad Prism.

Figure A.3. 3: Michaelis-Menten plots of the *PniDAH7PS*^{WT} for the K_M of PEP, E4P and chorismate in the presence of 150 mM or 500 mM NaCl respectively. Plotted and fitted globally using Graphpad Prism.

Supporting information for Chapter 5

Figure A.5. 1: Schematic diagram of pTXB1-DAH7PS-reactant for the expression of His-*Pni*DAH7PS^{D-H262}-intein-CBD.

Figure A.5. 2: The isolations of *PniDAH7PS*^{D-CM} from *PniDAH7PS*^{CM-T263C} using hydrophobic interaction chromatography (A) and anion exchange chromatography (B).

Figure A.5. 3: Michaelis-Menten plots of the *PniDAH7PS*^{D-H262} for the K_M of PEP and E4P, and of the *PniDAH7PS*^{CM-T263C} for the K_M of chorismate respectively. Plotted and fitted globally using Graphpad Prism.

Figure A.5. 4: Michaelis-Menten plots of the *PniDAH7PS*^{D-CM} for the K_M of PEP, E4P and chorismate respectively. Plotted and fitted globally using Graphpad Prism.

Figure A.5. 5: Scattering intensities and distance distribution functions of geometrical bodies (165).

Figure A.5. 6: SANS scattering using contrast variation of ligand free *Pni*DAH7PS^{D-CM} (black) and the *Pni*DAH7PS^{D-CM} respectively in the presence of PEP/Mn²⁺ (red) and prephenate (blue). (A) The 0% D₂O SANS scattering profiles (left) and Guinier plots (ln *I*(*q*) versus *q*²) (right). (B) The 42% D₂O SANS scattering profiles (left) and Guinier plots (ln *I*(*q*) versus *q*²) (right). (C) The 100% D₂O SANS scattering profiles (left) and Guinier plots (ln *I*(*q*) versus *q*²) (right).

References

1. Davis, B. D. (1955) Intermediates in amino acid biosynthesis. *Adv Enzymol Relat Subj Biochem* **16**, 247-312
2. Sprinson, D. B. (1960) The biosynthesis of aromatic compounds from D-glucose. *Adv Carbohydr Chem* **15**, 235-270
3. Herrmann, K. M., and Weaver, L. M. (1999) The Shikimate Pathway. *Annu Rev Plant Physiol Plant Mol Biol* **50**, 473-503
4. Kaleta, C., Schauble, S., Rinas, U., and Schuster, S. (2013) Metabolic costs of amino acid and protein production in *Escherichia coli*. *Biotechnol J* **8**, 1105-1114
5. Hartmann, M., Schneider, T. R., Pfeil, A., Heinrich, G., Lipscomb, W. N., and Braus, G. H. (2003) Evolution of feedback-inhibited β/α barrel isoenzymes by gene duplication and a single mutation. *Proc Natl Acad Sci U S A* **100**, 862-867
6. Shumilin, I. A., Zhao, C., Bauerle, R., and Kretsinger, R. H. (2002) Allosteric inhibition of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase alters the coordination of both substrates. *J Mol Biol* **320**, 1147-1156
7. Cross, P. J., and Parker, E. J. Allosteric inhibitor specificity of *Thermotoga maritima* 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. *FEBS Lett* **587**, 3063-3068
8. Wu, J., and Woodard, R. W. (2006) New insights into the evolutionary links relating to the 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase subfamilies. *J Biol Chem* **281**, 4042-4048
9. Jensen, R. A., Xie, G., Calhoun, D. H., and Bonner, C. A. (2002) The correct phylogenetic relationship of KdsA (3-deoxy-D-manno-octulosonate 8-phosphate synthase) with one of two independently evolved classes of AroA (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase). *J Mol Evol* **54**, 416-423
10. Walker, G. E., Dunbar, B., Hunter, I. S., Nimmo, H. G., and Coggins, J. R. (1996) Evidence for a novel class of microbial 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase in *Streptomyces coelicolor* A3(2), *Streptomyces rimosus* and *Neurospora crassa*. *Microbiology* **142** (Pt 8), 1973-1982
11. Subramaniam, P. S., Xie, G., Xia, T., and Jensen, R. A. (1998) Substrate ambiguity of 3-deoxy-D-manno-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* in the context of its membership in a protein family containing a subset of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthases. *J Bacteriol* **180**, 119-127
12. Gosset, G., Bonner, C. A., and Jensen, R. A. (2001) Microbial origin of plant-type 2-keto-3-deoxy-D-arabino-heptulosonate 7-phosphate synthases, exemplified by the chorismate- and tryptophan-regulated enzyme from *Xanthomonas campestris*. *J Bacteriol* **183**, 4061-4070
13. Sheflyan, G. Y., Sundaram, A. K., Taylor, W. P., and Woodard, R. W. (2000) Substrate ambiguity of 3-deoxy-D-manno-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* revisited. *J Bacteriol* **182**, 5005-5008
14. Shumilin, I. A., Kretsinger, R. H., and Bauerle, R. H. (1999) Crystal structure of phenylalanine-regulated 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Escherichia coli*. *Structure* **7**, 865-875
15. Konig, V., Pfeil, A., Braus, G. H., and Schneider, T. R. (2004) Substrate and metal complexes of

- 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae* provide new insights into the catalytic mechanism. *J Mol Biol* **337**, 675-690
16. Schofield, L. R., Patchett, M. L., and Parker, E. J. (2004) Expression, purification, and characterization of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Pyrococcus furiosus*. *Protein Expr Purif* **34**, 17-27
 17. Shumilin, I. A., Bauerle, R., Wu, J., Woodard, R. W., and Kretsinger, R. H. (2004) Crystal structure of the reaction complex of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Thermotoga maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation. *J Mol Biol* **341**, 455-466
 18. Zhou, L., Wu, J., Janakiraman, V., Shumilin, I. A., Bauerle, R., Kretsinger, R. H., and Woodard, R. W. (2012) Structure and characterization of the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Aeropyrum pernix*. *Bioorg Chem* **40**, 79-86
 19. Vainer, R., Belakhov, V., Rabkin, E., Baasov, T., and Adir, N. (2005) Crystal structures of *Escherichia coli* KDO8P synthase complexes reveal the source of catalytic irreversibility. *J Mol Biol* **351**, 641-652
 20. Nazmi, A. R., Lang, E. J. M., Bai, Y., Allison, T. M., Othman, M. H., Panjikar, S., Arcus, V. L., and Parker, E. J. (2016) Interdomain Conformational Changes Provide Allosteric Regulation en Route to Chorismate. *J Biol Chem* **291**, 21836-21847
 21. Webby, C. J., Baker, H. M., Lott, J. S., Baker, E. N., and Parker, E. J. (2005) The structure of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Mycobacterium tuberculosis* reveals a common catalytic scaffold and ancestry for type I and type II enzymes. *J Mol Biol* **354**, 927-939
 22. Cross, P. J., Pietersma, A. L., Allison, T. M., Wilson-Coutts, S. M., Cochrane, F. C., and Parker, E. J. (2013) *Neisseria meningitidis* expresses a single 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase that is inhibited primarily by phenylalanine. *Protein Sci* **22**, 1087-1099
 23. Schofield, L. R., Anderson, B. F., Patchett, M. L., Norris, G. E., Jameson, G. B., and Parker, E. J. (2005) Substrate ambiguity and crystal structure of *Pyrococcus furiosus* 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase: an ancestral 3-deoxyald-2-ulosonate-phosphate synthase? *Biochemistry* **44**, 11950-11962
 24. Webby, C. J., Jiao, W., Hutton, R. D., Blackmore, N. J., Baker, H. M., Baker, E. N., Jameson, G. B., and Parker, E. J. (2010) Synergistic allostery, a sophisticated regulatory network for the control of aromatic amino acid biosynthesis in *Mycobacterium tuberculosis*. *J Biol Chem* **285**, 30567-30576
 25. Hu, C., Jiang, P., Xu, J., Wu, Y., and Huang, W. (2003) Mutation analysis of the feedback inhibition site of phenylalanine-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase of *Escherichia coli*. *J Basic Microbiol* **43**, 399-406
 26. Ahmad, S., Rightmire, B., and Jensen, R. A. (1986) Evolution of the regulatory isozymes of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase present in the *Escherichia coli* genealogy. *J Bacteriol* **165**, 146-154
 27. Huang, L., Nakatsukasa, M., and Nester, E. (1974) Regulation of aromatic amino acid biosynthesis in *Bacillus subtilis* 168. Purification, characterization, and subunit structure of the bifunctional enzyme 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthetase-chorismate mutase. *J Biol Chem* **249**, 4467-4472
 28. Light, S. H., Halavaty, A. S., Minasov, G., Shuvalova, L., and Anderson, W. F. (2012) Structural

- analysis of a 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase with an N-terminal chorismate mutase-like regulatory domain. *Protein Sci* **21**, 887-895
29. Okvist, M., Sasso, S., Roderer, K., Kast, P., and Krengel, U. (2009) A novel noncovalent complex of chorismate mutase and DAHP synthase from *Mycobacterium tuberculosis*: protein purification, crystallization and X-ray diffraction analysis. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**, 1048-1052
 30. Burschowsky, D., Thorbjornsrud, H. V., Heim, J. B., Fahrig-Kamarauskaite, J. R., Wurth-Roderer, K., Kast, P., and Krengel, U. (2018) Inter-Enzyme Allosteric Regulation of Chorismate Mutase in *Corynebacterium glutamicum*: Structural Basis of Feedback Activation by Trp. *Biochemistry* **57**, 557-573
 31. Sterritt, O. W., Kessans, S. A., Jameson, G. B., and Parker, E. J. (2018) A Pseudoisostructural Type II DAH7PS Enzyme from *Pseudomonas aeruginosa*: Alternative Evolutionary Strategies to Control Shikimate Pathway Flux. *Biochemistry* **57**, 2667-2678
 32. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797
 33. Nazmi, A. R., Schofield, L. R., Dobson, R. C., Jameson, G. B., and Parker, E. J. (2014) Destabilization of the homotetrameric assembly of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from the hyperthermophile *Pyrococcus furiosus* enhances enzymatic activity. *J Mol Biol* **426**, 656-673
 34. Cross, P. J., Heyes, L. C., Zhang, S., Nazmi, A. R., and Parker, E. J. (2016) The functional unit of *Neisseria meningitidis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase is dimeric. *PLoS One* **11**, e0145187
 35. Jiao, W., Blackmore, N. J., Nazmi, A. R., and Parker, E. J. (2017) Quaternary structure is an essential component that contributes to the sophisticated allosteric regulation mechanism in a key enzyme from *Mycobacterium tuberculosis*. *PLoS One* **12**, e0180052
 36. Turnbull, J., and Morrison, J. F. (1990) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. 2. Evidence for two different active sites. *Biochemistry* **29**, 10255-10261
 37. Turnbull, J., Cleland, W. W., and Morrison, J. F. (1990) Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. 1. Kinetic characterization of the dehydrogenase reaction by use of alternative substrates. *Biochemistry* **29**, 10245-10254
 38. Cotton, R. G., and Gibson, F. (1965) The Biosynthesis of Phenylalanine and Tyrosine; Enzymes Converting Chorismic Acid into Prephenic Acid and Their Relationships to Prephenate Dehydratase and Prephenate Dehydrogenase. *Biochim Biophys Acta* **100**, 76-88
 39. MacBeath, G., Kast, P., and Hilvert, D. (1998) A small, thermostable, and monofunctional chorismate mutase from the archaeon *Methanococcus jannaschii*. *Biochemistry* **37**, 10062-10073
 40. Lee, A. Y., Stewart, J. D., Clardy, J., and Ganem, B. (1995) New insight into the catalytic mechanism of chorismate mutases from structural studies. *Chem Biol* **2**, 195-203
 41. Helmstaedt, K., Heinrich, G., Merkl, R., and Braus, G. H. (2004) Chorismate mutase of *Thermus thermophilus* is a monofunctional AroH class enzyme inhibited by tyrosine. *Arch Microbiol* **181**, 195-203
 42. Chook, Y. M., Gray, J. V., Ke, H., and Lipscomb, W. N. (1994) The monofunctional chorismate mutase from *Bacillus subtilis*. Structure determination of chorismate mutase and its complexes with a transition state analog and prephenate, and implications for the mechanism of the

- enzymatic reaction. *J Mol Biol* **240**, 476-500
43. Okvist, M., Dey, R., Sasso, S., Grahn, E., Kast, P., and Krengel, U. (2006) 1.6 Å crystal structure of the secreted chorismate mutase from *Mycobacterium tuberculosis*: novel fold topology revealed. *J Mol Biol* **357**, 1483-1499
 44. Angela Y. Lee, P. A. K., Bruce Ganem, Jon Clardy. (1995) Atomic structure of the buried catalytic pocket of *Escherichia coli* chorismate mutase. *J. Am. Chem. Soc.* **117**, 3627–3628
 45. Xue, Y., Lipscomb, W. N., Graf, R., Schnappauf, G., and Braus, G. (1994) The crystal structure of allosteric chorismate mutase at 2.2-Å resolution. *Proc Natl Acad Sci U S A* **91**, 10814-10818
 46. Schmidheini, T., Mosch, H. U., Evans, J. N., and Braus, G. (1990) Yeast allosteric chorismate mutase is locked in the activated state by a single amino acid substitution. *Biochemistry* **29**, 3660-3668
 47. Strater, N., Schnappauf, G., Braus, G., and Lipscomb, W. N. (1997) Mechanisms of catalysis and allosteric regulation of yeast chorismate mutase from crystal structures. *Structure* **5**, 1437-1452
 48. Kim, S. K., Reddy, S. K., Nelson, B. C., Robinson, H., Reddy, P. T., and Ladner, J. E. (2008) A comparative biochemical and structural analysis of the intracellular chorismate mutase (Rv0948c) from *Mycobacterium tuberculosis* H(37)R(v) and the secreted chorismate mutase (y2828) from *Yersinia pestis*. *Febs J* **275**, 4824-4835
 49. Yon-Kahn, J., and Hervé, G. (2010) Molecular and Cellular Enzymology.
 50. Singh, H., Arentson, B. W., Becker, D. F., and Tanner, J. J. (2014) Structures of the PutA peripheral membrane flavoenzyme reveal a dynamic substrate-channeling tunnel and the quinone-binding site. *Proc Natl Acad Sci U S A* **111**, 3389-3394
 51. Wilmanns, M., Priestle, J. P., Niermann, T., and Jansonius, J. N. (1992) Three-dimensional structure of the bifunctional enzyme phosphoribosylanthranilate isomerase: indoleglycerolphosphate synthase from *Escherichia coli* refined at 2.0 Å resolution. *J Mol Biol* **223**, 477-507
 52. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., and Davies, D. R. (1988) Three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ multienzyme complex from *Salmonella typhimurium*. *J Biol Chem* **263**, 17857-17871
 53. Raushel, F. M., Thoden, J. B., Reinhardt, G. D., and Holden, H. M. (1998) Carbamoyl phosphate synthetase: a crooked path from substrates to products. *Curr Opin Chem Biol* **2**, 624-632
 54. Blackmore, N. J., Nazmi, A. R., Hutton, R. D., Webby, M. N., Baker, E. N., Jameson, G. B., and Parker, E. J. (2015) Complex Formation between Two Biosynthetic Enzymes Modifies the Allosteric Regulatory Properties of Both: AN EXAMPLE OF MOLECULAR SYMBIOSIS. *J Biol Chem* **290**, 18187-18198
 55. Luger, K., Hommel, U., Herold, M., Hofsteenge, J., and Kirschner, K. (1989) Correct folding of circularly permuted variants of a beta alpha barrel enzyme in vivo. *Science* **243**, 206-210
 56. Eberhard, M., and Kirschner, K. (1989) Modification of a catalytically important residue of indoleglycerol-phosphate synthase from *Escherichia coli*. *FEBS Lett* **245**, 219-222
 57. Hommel, U., Eberhard, M., and Kirschner, K. (1995) Phosphoribosyl anthranilate isomerase catalyzes a reversible amadori reaction. *Biochemistry* **34**, 5429-5439
 58. Eberhard, M., Tsai-Pflugfelder, M., Bolewska, K., Hommel, U., and Kirschner, K. (1995) Indoleglycerol phosphate synthase-phosphoribosyl anthranilate isomerase: comparison of the bifunctional enzyme from *Escherichia coli* with engineered monofunctional domains.

59. Anderson, K. S., Miles, E. W., and Johnson, K. A. (1991) Serine modulates substrate channeling in tryptophan synthase. A novel intersubunit triggering mechanism. *J Biol Chem* **266**, 8020-8033
60. Burkhard, P., Rao, G. S., Hohenester, E., Schnackerz, K. D., Cook, P. F., and Jansonius, J. N. (1998) Three-dimensional structure of *O*-acetylserine sulfhydrylase from *Salmonella typhimurium*. *J Mol Biol* **283**, 121-133
61. Schlichting, I., Yang, X. J., Miles, E. W., Kim, A. Y., and Anderson, K. S. (1994) Structural and kinetic analysis of a channel-impaired mutant of tryptophan synthase. *J Biol Chem* **269**, 26591-26593
62. Anderson, K. S., Kim, A. Y., Quillen, J. M., Sayers, E., Yang, X. J., and Miles, E. W. (1995) Kinetic characterization of channel impaired mutants of tryptophan synthase. *J Biol Chem* **270**, 29936-29944
63. Goodey, N. M., and Benkovic, S. J. (2008) Allosteric regulation and catalysis emerge via a common route. *Nat Chem Biol* **4**, 474-482
64. Laskowski, R. A., Gerick, F., and Thornton, J. M. (2009) The structural basis of allosteric regulation in proteins. *FEBS Lett* **583**, 1692-1698
65. Howlett, G. J., Blackburn, M. N., Compton, J. G., and Schachman, H. K. (1977) Allosteric regulation of aspartate transcarbamoylase. Analysis of the structural and functional behavior in terms of a two-state model. *Biochemistry* **16**, 5091-5100
66. Jacobson, G. R., and Stark, G. R. (1973) Aspartate transcarbamylase. A study of possible roles for the sulfhydryl group at the active site. *J Biol Chem* **248**, 8003-8014
67. Kirschner, K. (1968) Allosteric regulation of enzyme activity. An introduction to the molecular basis of and the experimental approaches to the problem. *Curr Top Microbiol Immunol* **44**, 123-146
68. Ruckpaul, K., Blanck, J., Scheler, W., and Jung, F. (1971) Kinetic studies on the binding of alkylisocyanides to some hemoglobins. *Acta Biol Med Ger* **27**, 693-701
69. Shulman, R. G., Ogawa, S., and Hopfield, J. J. (1972) An allosteric model of hemoglobin. II. The assumption of independent binding. *Arch Biochem Biophys* **151**, 68-74
70. Macol, C. P., Tsuruta, H., Stec, B., and Kantrowitz, E. R. (2001) Direct structural evidence for a concerted allosteric transition in *Escherichia coli* aspartate transcarbamoylase. *Nat Struct Biol* **8**, 423-426
71. Paoli, M., Liddington, R., Tame, J., Wilkinson, A., and Dodson, G. (1996) Crystal structure of T state haemoglobin with oxygen bound at all four haems. *J Mol Biol* **256**, 775-792
72. Denisov, I. G., and Sligar, S. G. (2012) A novel type of allosteric regulation: functional cooperativity in monomeric proteins. *Arch Biochem Biophys* **519**, 91-102
73. Panagou, D., Orr, M. D., Dunstone, J. R., and Blakley, R. L. (1972) A monomeric, allosteric enzyme with a single polypeptide chain. Ribonucleotide reductase of *Lactobacillus leichmannii*. *Biochemistry* **11**, 2378-2388
74. Garel, J. R., and Dautry-Varsat, A. (1980) Sequential folding of a bifunctional allosteric protein. *Proc Natl Acad Sci U S A* **77**, 3379-3383
75. Cornish-Bowden, A., and Cardenas, M. L. (1987) Co-operativity in monomeric enzymes. *J Theor Biol* **124**, 1-23
76. Ludwig, M. L., and Matthews, R. G. (2002) Effector regulation in a monomeric enzyme. *Nat*

77. Wei, G., Xi, W., Nussinov, R., and Ma, B. (2016) Protein Ensembles: How Does Nature Harness Thermodynamic Fluctuations for Life? The Diverse Functional Roles of Conformational Ensembles in the Cell. *Chem Rev* **116**, 6516-6551
78. Cooper, A. (1984) Protein fluctuations and the thermodynamic uncertainty principle. *Prog Biophys Mol Biol* **44**, 181-214
79. Cooper, A. (2010) Protein Heat Capacity: An Anomaly that Maybe Never Was. *The Journal of Physical Chemistry Letters* **1**, 3298-3304
80. Gekko, K. (2002) Compressibility gives new insight into protein dynamics and enzyme function. *Biochim Biophys Acta* **1595**, 382-386
81. Ohmine, I., Tanaka, H., and Wolynes, P. G. (1988) Large local energy fluctuations in water. II. Cooperative motions and fluctuations. *The Journal of chemical physics* **89**, 5852-5860
82. Keedy, D. A., Kenner, L. R., Warkentin, M., Woldeyes, R. A., Hopkins, J. B., Thompson, M. C., Brewster, A. S., Van Benschoten, A. H., Baxter, E. L., Uervirojnangkoorn, M., McPhillips, S. E., Song, J., Alonso-Mori, R., Holton, J. M., Weis, W. I., Brunger, A. T., Soltis, S. M., Lemke, H., Gonzalez, A., Sauter, N. K., Cohen, A. E., van den Bedem, H., Thorne, R. E., and Fraser, J. S. (2015) Mapping the conformational landscape of a dynamic enzyme by multitemperature and XFEL crystallography. *Elife* **4**
83. Hilser, V. J., and Thompson, E. B. (2007) Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc Natl Acad Sci U S A* **104**, 8311-8315
84. Luque, I., Leavitt, S. A., and Freire, E. (2002) The linkage between protein folding and functional cooperativity: two sides of the same coin? *Annu Rev Biophys Biomol Struct* **31**, 235-256
85. Watt, E. D., Shimada, H., Kovrigin, E. L., and Loria, J. P. (2007) The mechanism of rate-limiting motions in enzyme function. *Proc Natl Acad Sci U S A* **104**, 11981-11986
86. Moscarello, J. M., Ben-Shahar, O., and Ettenberg, A. (2007) Dynamic interaction between medial prefrontal cortex and nucleus accumbens as a function of both motivational state and reinforcer magnitude: a c-Fos immunocytochemistry study. *Brain Res* **1169**, 69-76
87. Benkovic, S. J., Hammes, G. G., and Hammes-Schiffer, S. (2008) Free-energy landscape of enzyme catalysis. *Biochemistry* **47**, 3317-3321
88. Boehr, D. D., McElheny, D., Dyson, H. J., and Wright, P. E. (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638-1642
89. Kerns, S. J., Agafonov, R. V., Cho, Y. J., Pontiggia, F., Otten, R., Pachov, D. V., Kutter, S., Phung, L. A., Murphy, P. N., Thai, V., Alber, T., Hagan, M. F., and Kern, D. (2015) The energy landscape of adenylate kinase during catalysis. *Nat Struct Mol Biol* **22**, 124-131
90. Ma, B., and Nussinov, R. (2016) Protein dynamics: Conformational footprints. *Nat Chem Biol* **12**, 890-891
91. Nussinov, R., and Wolynes, P. G. (2014) A second molecular biology revolution? The energy landscapes of biomolecular function. *Phys Chem Chem Phys* **16**, 6321-6322
92. Kumar, S., Ma, B., Tsai, C. J., Sinha, N., and Nussinov, R. (2000) Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci* **9**, 10-19
93. Daily, M. D., and Gray, J. J. (2007) Local motions in a benchmark of allosteric proteins. *Proteins* **67**, 385-399
94. Chennubhotla, C., Yang, Z., and Bahar, I. (2008) Coupling between global dynamics and signal transduction pathways: a mechanism of allostery for chaperonin GroEL. *Mol Biosyst* **4**, 287-292

95. Daily, M. D., Upadhyaya, T. J., and Gray, J. J. (2008) Contact rearrangements form coupled networks from local motions in allosteric proteins. *Proteins* **71**, 455-466
96. Cross, P. J., Allison, T. M., Dobson, R. C., Jameson, G. B., and Parker, E. J. (2013) Engineering allosteric control to an unregulated enzyme by transfer of a regulatory domain. *Proc Natl Acad Sci U S A* **110**, 2111-2116
97. Fan, Y., Cross, P. J., Jameson, G. B., and Parker, E. J. (2018) Exploring modular allostery via interchangeable regulatory domains. *Proc Natl Acad Sci U S A* **115**, 3006-3011
98. Lang, E. J., Cross, P. J., Mittelstadt, G., Jameson, G. B., and Parker, E. J. (2014) Allosteric ACTION: the varied ACT domains regulating enzymes of amino-acid metabolism. *Curr Opin Struct Biol* **29**, 102-111
99. Lorence, J. H., and Nester, E. W. (1967) Multiple molecular forms of chorismate mutase in *Bacillus subtilis*. *Biochemistry* **6**, 1541-1553
100. Gray, J. V., Golinelli-Pimpaneau, B., and Knowles, J. R. (1990) Monofunctional chorismate mutase from *Bacillus subtilis*: purification of the protein, molecular cloning of the gene, and overexpression of the gene product in *Escherichia coli*. *Biochemistry* **29**, 376-383
101. Shah, H. N., and Gharbia, S. E. (1992) Biochemical and chemical studies on strains designated *Prevotella intermedia* and proposal of a new pigmented species, *Prevotella nigrescens* sp. nov. *Int J Syst Bacteriol* **42**, 542-546
102. Shah, H. N., and Gharbia, S. E. (1993) Proposal of a new species *Prevotella nigrescens* sp. nov. among strains previously classified as *Pr. intermedia*. *FEMS Immunol Med Microbiol* **6**, 97
103. Fukui, K., Kato, N., Kato, H., Watanabe, K., and Tatematsu, N. (1999) Incidence of *Prevotella intermedia* and *Prevotella nigrescens* carriage among family members with subclinical periodontal disease. *J Clin Microbiol* **37**, 3141-3145
104. Stingu, C. S., Schaumann, R., Jentsch, H., Eschrich, K., Brosteanu, O., and Rodloff, A. C. (2013) Association of periodontitis with increased colonization by *Prevotella nigrescens*. *J Investig Clin Dent* **4**, 20-25
105. Rams, T. E., Degener, J. E., and van Winkelhoff, A. J. (2014) Antibiotic resistance in human chronic periodontitis microbiota. *J Periodontol* **85**, 160-169
106. Yakob, M., Soder, B., Meurman, J. H., Jogestrand, T., Nowak, J., and Soder, P. O. (2011) *Prevotella nigrescens* and *Porphyromonas gingivalis* are associated with signs of carotid atherosclerosis in subjects with and without periodontitis. *J Periodontal Res* **46**, 749-755
107. Wu, J., Howe, D. L., and Woodard, R. W. (2003) *Thermotoga maritima* 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase: the ancestral eubacterial DAHP synthase? *J Biol Chem* **278**, 27525-27531
108. Cross, P. J., and Parker, E. J. (2013) Allosteric inhibitor specificity of *Thermotoga maritima* 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. *FEBS Lett* **587**, 3063-3068
109. Cross, P. J., Dobson, R. C., Patchett, M. L., and Parker, E. J. (2011) Tyrosine latching of a regulatory gate affords allosteric control of aromatic amino acid biosynthesis. *J Biol Chem* **286**, 10216-10224
110. Frickey, T., and Lupas, A. (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* **20**, 3702-3704
111. Light, S. H., and Anderson, W. F. (2013) The diversity of allosteric controls at the gateway to aromatic amino acid biosynthesis. *Protein Sci* **22**, 395-404
112. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., and

- Hochstrasser, D. F. (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* **112**, 531-552
113. Schoner, R., and Herrmann, K. M. (1976) 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. Purification, properties, and kinetics of the tyrosine-sensitive isoenzyme from *Escherichia coli*. *J Biol Chem* **251**, 5440-5447
 114. Sterritt, O. W., Lang, E. J., Kessans, S. A., Ryan, T. M., Demeler, B., Jameson, G. B., and Parker, E. J. (2018) The short-form type II DAH7PS from *Pseudomonas aeruginosa*, associated with phenazine biosynthesis, suggests reclassification of type II enzymes. *Bioscience reports*, BSR20181605
 115. Nazmi, A. R., and Othman, M. (2014) personal communication.
 116. Fan, Y. (2014) personal communication.
 117. Zhang, S., Kongsaree, P., Clardy, J., Wilson, D. B., and Ganem, B. (1996) Site-directed mutagenesis of monofunctional chorismate mutase engineered from the E. coli P-protein. *Bioorg Med Chem* **4**, 1015-1020
 118. Hartley, J. L., Temple, G. F., and Brasch, M. A. (2000) DNA cloning using in vitro site-specific recombination. *Genome Research* **10**, 1788-1795
 119. Liang, X., Peng, L., Baek, C. H., and Katzen, F. (2013) Single step BP/LR combined Gateway reactions. *Biotechniques* **55**, 265-268
 120. Katzen, F. (2007) Gateway® recombinational cloning: a biological operating system. *Expert Opin Drug Discov* **2**, 571-589
 121. Feigin, L., and Svergun, D. I. (1987) *Structure analysis by small-angle X-ray and neutron scattering*, Springer
 122. Rambo, R. P., and Tainer, J. A. (2011) Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95**, 559-571
 123. Hammel, M. (2012) Validation of macromolecular flexibility in solution by small-angle X-ray scattering (SAXS). *Eur Biophys J* **41**, 789-799
 124. Franke, D., and Svergun, D. I. (2009) *DAMMIF*, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* **42**, 342-346
 125. Gorrec, F. (2009) The MORPHEUS protein crystallization screen. *J Appl Crystallogr* **42**, 1035-1042
 126. Kabsch, W. (2010) Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132
 127. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674
 128. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72-82
 129. Easterby, J. S. (1989) The analysis of metabolite channelling in multienzyme complexes and multifunctional proteins. *Biochem J* **264**, 605-607
 130. Lee, D. C., L.; Nguyen, D.; Dillard, B.D.; Tempel, W.; Habel, J.; Zhou, W.; Chang, S.-H.; Kelley, L.-L.C.; Liu, Z.-J.; Lin, D.; Zhang, H.; Praissman, J.; Bridger, S.; Eneh, J.C.; Hopkins, R.C.; Jenney Jr., F.E.; Lee, H.-S.; Li, T.; Poole II, F.L.; Shah, C.; Sugar, F.JionAdom W W.; gosc, W JP; ang,

132. Petoukhov, M. V., and Svergun, D. I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J* **89**, 1237-1250
133. Perozzo, R., Folkers, G., and Scapozza, L. (2004) Thermodynamics of protein-ligand interactions: history, presence, and future aspects. *J Recept Signal Transduct Res* **24**, 1-52
134. Ward, D. E., de Vos, W. M., and van der Oost, J. (2002) Molecular analysis of the role of two aromatic aminotransferases and a broad-specificity aspartate aminotransferase in the aromatic amino acid metabolism of *Pyrococcus furiosus*. *Archaea* **1**, 133-141
135. Kawarabayashi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Haikawa, Y., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Nakazawa, H., Takamiya, M., Masuda, S., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Kikuchi, H., and et al. (1999) Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1. *DNA Res* **6**, 83-101, 145-152
136. Ma, B., Tsai, C. J., Haliloglu, T., and Nussinov, R. (2011) Dynamic allostery: linkers are not merely flexible. *Structure* **19**, 907-917
137. Pegan, S. D., Rukserree, K., Capodagli, G. C., Baker, E. A., Krasnykh, O., Franzblau, S. G., and Mesecar, A. D. (2013) Active site loop dynamics of a class IIa fructose 1,6-bisphosphate aldolase from *Mycobacterium tuberculosis*. *Biochemistry* **52**, 912-925
138. Chen, J., Dima, R. I., and Thirumalai, D. (2007) Allosteric communication in dihydrofolate reductase: signaling network and pathways for closed to occluded transition and back. *J Mol Biol* **374**, 250-266
139. Kunz, W. (2010) *Specific ion effects*, World Scientific
140. Zhang, J. (2012) *Protein-protein interactions in salt solutions*, INTECH Open Access Publisher Croatia, Europe
141. Zidek, L., Novotny, M. V., and Stone, M. J. (1999) Increased protein backbone conformational entropy upon hydrophobic ligand binding. *Nat Struct Biol* **6**, 1118-1121
142. Odriozola, G., and Lozada-Cassou, M. (2013) Statistical mechanics approach to lock-key supramolecular chemistry interactions. *Phys Rev Lett* **110**, 105701
143. Oleinikova, A., Smolin, N., Brovchenko, I., Geiger, A., and Winter, R. (2005) Formation of spanning water networks on protein surfaces via 2D percolation transition. *J Phys Chem B* **109**, 1988-1998
144. Schneider, H. J. (2015) Limitations and extensions of the lock-and-key principle: differences between gas state, solution and solid state structures. *Int J Mol Sci* **16**, 6694-6717
145. Rambo, R. P., and Tainer, J. A. (2011) Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95**, 559-571
146. Hammel, M. (2012) Validation of macromolecular flexibility in solution by small-angle X-ray scattering (SAXS). *Eur Biophys J* **41**, 789-799
147. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of applied crystallography* **36**, 1277-1282
148. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* **80**, 2946-2953
149. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) Expressed protein ligation: a general method for protein engineering. *Proc Natl Acad Sci U S A* **95**, 6705-6710
150. Severinov, K., and Muir, T. W. (1998) Expressed protein ligation, a novel method for studying

- protein-protein interactions in transcription. *J Biol Chem* **273**, 16205-16209
151. Kwan, A. H., Mobli, M., Gooley, P. R., King, G. F., and Mackay, J. P. (2011) Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS J* **278**, 687-703
 152. Evans, T. C., Jr., Benner, J., and Xu, M. Q. (1999) The in vitro ligation of bacterially expressed proteins using an intein from *Methanobacterium thermoautotrophicum*. *J Biol Chem* **274**, 3923-3926
 153. Otomo, T., Teruya, K., Uegaki, K., Yamazaki, T., and Kyogoku, Y. (1999) Improved segmental isotope labeling of proteins and application to a larger protein. *J Biomol NMR* **14**, 105-114
 154. Camarero, J. A., Fushman, D., Cowburn, D., and Muir, T. W. (2001) Peptide chemical ligation inside living cells: in vivo generation of a circular protein domain. *Bioorg Med Chem* **9**, 2479-2484
 155. Williamson, M. P. (2013) Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc* **73**, 1-16
 156. Liu, Y., Kahn, R. A., and Prestegard, J. H. (2014) Interaction of Fapp1 with Arf1 and PI4P at a membrane surface: an example of coincidence detection. *Structure* **22**, 421-430
 157. Rajagopal, P., Waygood, E. B., Reizer, J., Saier, M. H., Jr., and Klevit, R. E. (1997) Demonstration of protein-protein interaction specificity by NMR chemical shift mapping. *Protein Sci* **6**, 2624-2627
 158. Ikura, M., Kay, L. E., and Bax, A. (1990) A novel approach for sequential assignment of ^1H , ^{13}C , and ^{15}N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry* **29**, 4659-4667
 159. Friedrichs, M. S., Mueller, L., and Wittekind, M. (1994) An automated procedure for the assignment of protein ^1HN , ^{15}N , ^{13}C alpha, ^1H alpha, ^{13}C beta and ^1H beta resonances. *J Biomol NMR* **4**, 703-726
 160. Evans, T. C., Jr., Benner, J., and Xu, M. Q. (1999) The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins. *J Biol Chem* **274**, 18359-18363
 161. Bieri, M., Kwan, A. H., Mobli, M., King, G. F., Mackay, J. P., and Gooley, P. R. (2011) Macromolecular NMR spectroscopy for the non-spectroscopist: beyond macromolecular solution structure determination. *FEBS J* **278**, 704-715
 162. Gragerov, A., Nudler, E., Komissarova, N., Gaitanaris, G. A., Gottesman, M. E., and Nikiforov, V. (1992) Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **89**, 10341-10344
 163. Martin, J., Mayhew, M., Langer, T., and Hartl, U. (1993) The reaction cycle of GroEL and GroES in chaperonin-assisted protein folding. *Nature* **366**, 228
 164. Duff, A. P., Wilde, K. L., Rekas, A., Lake, V., and Holden, P. J. (2015) Robust high-yield methodologies for ^2H and $^2\text{H}/^{15}\text{N}/^{13}\text{C}$ labeling of proteins for structural investigations using neutron scattering and NMR. in *Methods in enzymology*, Elsevier. pp 3-25
 165. Koch, M. H., Vachette, P., and Svergun, D. I. (2003) Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q Rev Biophys* **36**, 147-227
 166. Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A*

167. Tugarinov, V., Choy, W. Y., Orekhov, V. Y., and Kay, L. E. (2005) Solution NMR-derived global fold of a monomeric 82-kDa enzyme. *Proc Natl Acad Sci U S A* **102**, 622-627
168. Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wüthrich, K. (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proceedings of the National Academy of Sciences* **95**, 13585-13590
169. Karagoz, G. E., Sinnige, T., Hsieh, O., and Rudiger, S. G. (2011) Expressed protein ligation for a large dimeric protein. *Protein Eng Des Sel* **24**, 495-501
170. Sonntag, M., Jagtap, P. K. A., Simon, B., Appavou, M. S., Geerlof, A., Stehle, R., Gabel, F., Hennig, J., and Sattler, M. (2017) Segmental, Domain-Selective Perdeuteration and Small-Angle Neutron Scattering for Structural Analysis of Multi-Domain Proteins. *Angew Chem Int Ed Engl* **56**, 9322-9325
171. Levary, D. A., Parthasarathy, R., Boder, E. T., and Ackerman, M. E. (2011) Protein-protein fusion catalyzed by sortase A. *PLoS One* **6**, e18342
172. Mao, H., Hart, S. A., Schink, A., and Pollok, B. A. (2004) Sortase-mediated protein ligation: a new method for protein engineering. *J Am Chem Soc* **126**, 2670-2671
173. Nguyen, G. K., Qiu, Y., Cao, Y., Hemu, X., Liu, C. F., and Tam, J. P. (2016) Butelase-mediated cyclization and ligation of peptides and proteins. *Nat Protoc* **11**, 1977-1988
174. Campbell, E., Kaltenbach, M., Correy, G. J., Carr, P. D., Porebski, B. T., Livingstone, E. K., Afriat-Jurnou, L., Buckle, A. M., Weik, M., Hollfelder, F., Tokuriki, N., and Jackson, C. J. (2016) The role of protein dynamics in the evolution of new enzyme function. *Nat Chem Biol* **12**, 944-950
175. Wade, R. C., Davis, M. E., Luty, B. A., Madura, J. D., and McCammon, J. A. (1993) Gating of the active site of triose phosphate isomerase: Brownian dynamics simulations of flexible peptide loops in the enzyme. *Biophys J* **64**, 9-15
176. Venkitakrishnan, R. P., Zaborowski, E., McElheny, D., Benkovic, S. J., Dyson, H. J., and Wright, P. E. (2004) Conformational changes in the active site loops of dihydrofolate reductase during the catalytic cycle. *Biochemistry* **43**, 16046-16055
177. Sasso, S., Okvist, M., Roderer, K., Gamper, M., Codoni, G., Krengel, U., and Kast, P. (2009) Structure and function of a complex between chorismate mutase and DAHP synthase: efficiency boost for the junior partner. *EMBO J* **28**, 2128-2142
178. Finn, R. D., Bateman, A., Clements, J., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J., and Punta, M. (2014) Pfam: the protein families database. *Nucleic Acids Res* **42**, D222-230
179. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539
180. Johnson, L. S., Eddy, S. R., and Portugaly, E. (2010) Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* **11**, 431
181. Consortium, U. (2014) Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res* **42**, D191-198
182. Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., Dicuccio, M., Federhen, S., Feolo, M., Fingerman, I. M., Geer, L. Y., Helmberg, W., Kapustin, Y., Krasnov, S., Landsman, D., Lipman, D. J., Lu, Z., Madden, T. L., Madej, T., Maglott, D. R., Marchler-Bauer, A., Miller, V., Karsch-Mizrachi, I., Ostell, J., Panchenko, A., Phan, L., Pruitt,

- K. D., Schuler, G. D., Sequeira, E., Sherry, S. T., Shumway, M., Sirotkin, K., Slotta, D., Souvorov, A., Starchenko, G., Tatusova, T. A., Wagner, L., Wang, Y., Wilbur, W. J., Yaschenko, E., and Ye, J. (2011) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **40**, D13-25
183. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191
 184. Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205-217
 185. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324
 186. Landy, A. (1989) Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu Rev Biochem* **58**, 913-949
 187. Brookes, E., Demeler, B., Rosano, C., and Rocco, M. (2010) The implementation of SOMO (SOLution MOdeller) in the UltraScan analytical ultracentrifugation data analysis suite: enhanced capabilities allow the reliable hydrodynamic modeling of virtually any kind of biomacromolecule. *Eur Biophys J* **39**, 423-435
 188. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606-1619
 189. Cowieson, N. P., Aragao, D., Clift, M., Ericsson, D. J., Gee, C., Harrop, S. J., Mudie, N., Panjikar, S., Price, J. R., Riboldi-Tunnicliffe, A., Williamson, R., and Caradoc-Davies, T. (2015) MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. *J Synchrotron Radiat* **22**, 187-190
 190. Aragao, D., Aishima, J., Cherukuvada, H., Clarken, R., Clift, M., Cowieson, N. P., Ericsson, D. J., Gee, C. L., Macedo, S., Mudie, N., Panjikar, S., Price, J. R., Riboldi-Tunnicliffe, A., Rostan, R., Williamson, R., and Caradoc-Davies, T. T. (2018) MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *J Synchrotron Radiat* **25**, 885-891
 191. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* **66**, 133-144
 192. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242
 193. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) *iMOSFLM*: a new graphical interface for diffraction-image processing with *MOSFLM*. *Acta Crystallogr D Biol Crystallogr* **67**, 271-281
 194. Vagin, A., and Teplyakov, A. (1997) *MOLREP*: an automated program for molecular replacement. *Journal of applied crystallography* **30**, 1022-1025
 195. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255
 196. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) *REFMAC5* for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367

197. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501
198. Kirby, N., Boldeman, J., Gentle, I., and Cookson, D. (2007) Conceptual design of the small angle scattering beamline at the Australian synchrotron. in *AIP Conference Proceedings*, AIP
199. Kirby, N. M., Mudie, S. T., Hawley, A. M., Cookson, D. J., Mertens, H. D., Cowieson, N., and Samardzic-Boban, V. (2013) A low-background-intensity focusing small-angle X-ray scattering undulator beamline. *Journal of Applied Crystallography* **46**, 1670-1680
200. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* **36**, 1295-1296
201. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J Appl Cryst.* **36**, 860-864
202. Webb, B., and Sali, A. (2016) Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci* **86**, 2 9 1-2 9 37
203. Petoukhov, M. V., and Svergun, D. I. (2006) Joint use of small-angle X-ray and neutron scattering to study biological macromolecules in solution. *Eur Biophys J* **35**, 567-576
204. Gilbert, E. P., Schulz, J. C., and Noakes, T. J. (2006) 'Quokka'—the small-angle neutron scattering instrument at OPAL. *Physica B: Condensed Matter* **385**, 1180-1182
205. Hamilton, W. A., Garrett, R., and Gilbert, E. P. (2009) QUOKKA: the small-angle neutron scattering instrument. *Neutron News* **20**, 24-28
206. Kline, S. R. (2006) Reduction and analysis of SANS and USANS data using IGOR Pro. *Journal of applied crystallography* **39**, 895-900
207. Svergun, D. (1992) Determination of the regularization parameter in indirect - transform methods using perceptual criteria. *Journal of applied crystallography* **25**, 495-503
208. Cai, M., Huang, Y., Sakaguchi, K., Clore, G. M., Gronenborn, A. M., and Craigie, R. (1998) An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*. *J Biomol NMR* **11**, 97-102
209. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234