Bacterial back-talk: Structural stability, unfolding and molecular interaction studies on Staphylococcal proteins

Sara Zandomeneghi, BSc

Thesis submitted to the University of Nottingham in part fulfilment of the Degree of Doctor of Philosophy

School of Life Sciences

Declaration

I, Sara Zandomeneghi, declare that the work described in this thesis is the result of my own investigation and any work carried out by other investigators is fully acknowledged in the text.

I declare that this thesis has not been submitted, nor is currently being submitted for any other degree, at this or any other institution.

Abstract

The structural characterisation of proteins is key to understanding their biological functions and interactions. Many different biophysical techniques have been developed in the last century to achieve protein structure determination, among which crystallography has become the most common and successful. The physical constraints imposed by crystallography, though, such as low temperatures and the tight packing of molecules, cause the loss of important biological information especially when it comes to soluble proteins. In a native environment, these move freely in an aqueous environment and can display varying degrees of flexibility to accommodate for molecular interactions. The structure obtained by crystallizing the protein might therefore not be truly representative of the conformation adopted under physiological conditions.

Different biophysical techniques, such as circular dichroism and small angle X-ray and neutron scattering were combined with computational tools to characterise the structure and molecular interactions of proteins in a way that allows to retain physiological conditions. This methodology was developed and applied to the proteins of the quorum-sensing system (or accessory gene regulator, *agr*) in *Staphylococcus aureus*, a Gram-positive pathogen which is often responsible for post-operative infections in hospital patients. The *agr* system is responsible for the collective behaviour and control of infection phases in *S. aureus* and an understanding of the mechanism of action of these not yet fully characterised proteins will inform on potential pathway inhibitors to improve patient prognosis without encouraging antibiotic resistance. Particular focus was on the *agr* two-component system. Two-component systems are a

signal transduction mechanism comprising of a membrane receptor and a response regulator widely found in bacteria.

Specifically, molecular dynamics simulations were used to create atomic-level models, small angle scattering to investigate oligomeric state, create low-resolution models and identify single or multiple conformations and CD and SAXS to identify molecular interactions. The cytoplasmic domains of AgrC, the histidine kinase in the two-component system of the *agr* system, were shown to retain a dimeric state even in absence of the transmembrane region, and an MD-generated model was validated by SAXS. An experimentally validated model of full-length AgrA, the response regulator, was also identified. The effect of ATP, which acts as phosphor-donor, was investigated by both CD and SAXS. A non-hydrolysable analogue, sATP, was found to alter the compactness of $AgrC_{206}$ (the cytoplasmic domain of AgrC) and affect its thermal response. SANS studies supplied information on the behaviour of $AgrC_{206}$ and AgrA while in a complex. The effect of ATP on AgrA was shown to be dependent on the interaction between the two proteins.

Publications

Bardelang P, Murray EJ, Blower I, Zandomeneghi S, Goode A, Hussain R, Kumari D, Siligardi G, Inoue K, Luckett J, Doutch J, Emsley J, Chan WC, Hill P, Williams P and Bonev BB (2023) Conformational analysis and interaction of the Staphylococcus aureus transmembrane peptidase AgrB with its AgrD propeptide substrate.

Front. Chem. 11:1113885. doi: 10.3389/fchem.2023.1113885

Acknowledgments

First and foremost, I would like to thank Dr Boyan Bonev for his constant guidance and wisdom throughout my PhD. His support has been unwavering, through the tough times as well as the good ones. He made sure that, as a lab, we never lost the sense of community during the periods of isolation, which provided great comfort to all of us. I will be forever grateful to him.

Big thanks to Dr James Doutch for the massive help he provided in all things scattering and for giving me the opportunity to explore the world of synchrotron radiation. His enthusiasm and expertise have been truly inspirational.

I am grateful to Prof Paul Williams for his guidance in the quorum-sensing system, the access to equipment and for bringing together the fantastic 'Staph' group of people. Through him I met some amazingly knowledgeable people, in particular Dr Phil Bardelang and Dr Ewan Murray, to whom I send massive thanks for the experimental help and expertise shared throughout my PhD.

To all the lab members, no words are enough to thank you for these four fantastic years. To Dr Alice Goode, for all the hot chocolates, the support in the lab and being an amazing person all round. To Dr Vivien Yeh, for being there through the rough times, sharing her expertise in the lab, and organising pretty much all the socials. To Gurdip Matharu, for going through the 'joys' of PhD together.

Big thanks to all the instrument scientists that supported us through the various experimental sessions: Katsuaki Inoue, Rohanah Hussain, Giuliano Siligardi, Tiberiu-Marius Gianga and Sam Burholt and to Diamond Light Source and the ISIS Neutron and Muon Source for granting the experimental time.

Huge thanks to my amazing boyfriend Alex Byrne for the massive moral support throughout the last 3 years, for helping through the late nights in the lab and listening to my endless ranting.

The greatest acknowledgment of all goes to my parents, who made it possible for me to move to the UK and start this massive journey. Thank you for never doubting that I could do it and for the advice, support and love you have always given me.

And finally, I could not end this without mentioning our family dog, Lilla, and our horse My Horse (haha). Thank you for the fun times, the frustrating times, and just for being you.

I would like to acknowledge the BBSRC Doctoral Training Programme and the ISIS Neutron and Muon Source for jointly funding this research.

Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

This work benefited from the use of the SasView application, originally developed under NSF Award DMR - 0520547. SasView also contains code developed with funding from the EU Horizon 2020 programme under the SINE2020 project Grant No 654000.

PDB structures were visualised using BIOVIA, Dassault Systèmes, Discovery Studio Visualizer, v21.1.0.20298, San Diego: Dassault Systèmes, 2020.

Contents

Declaration 2

Abstract 3

Publications 5

Acknowledgments 6

List of abbreviations 12

List of tables 25

1. Introduction 26

1.1. *S. aureus* and quorum sensing system 28

1.1.1. AgrB and the proteolysis of AgrD 32

1.1.2. Two-component systems – AgrA and AgrC 34

1.1.3. Inhibition of AgrA – IQS analogues 43

1.2. Small Angle Scattering 45

1.2.1. Small Angle X-ray Scattering (SAXS) 49

1.2.2. Small Angle Neutron Scattering (SANS) 50

1.3. Circular Dichroism 52

1.4. Phase transitions and the Landau free energy expansion analysis 55

1.5. Magnetically aligned bicelles in morphological studies of membrane proteins using SANS 65

1.6. Hypothesis and Aims 66

2. Materials and Methods 68

2.1. Materials 68

2.2. Homology modelling and MD simulations 68

2.3. *E. coli* BL21 electrocompetent cells 69

2.4. DNA gel electrophoresis 69

2.5. Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS PAGE) 70

2.6. Cloning 71

2.6.1. $AgrC_{206}$ 71

2.6.2. N-terminal AgrA 74

2.7. *E. coli* BL21 transformation by electroporation 75

2.8. Expression of recombinant AgrA in *E. coli* BL21 75

2.9. Expression of recombinant C-terminal AgrA and $AgrC_{206}$ in *E. coli* BL21 76

2.12. Expression of recombinant HmbRI D94N 78

2.12.1. ^{1}H-HmBR 78
2.12.2. ^{2}H-HmbR 78
2.13. Electrophoretic Mobility Shift Assay (EMSA) 79
2.14. Probes for Electrophoretic Mobility Shift Assays 81
2.15. Protein purification 83
2.15.1. Purification protocols for full-length, C-terminal and N-terminal AgrA and $AgrC_{206}$ 83
2.15.2. Purification protocol for HmbRI D94N 83
2.16. Incorporation of membrane proteins into lipid 84
2.17. Size exclusion chromatography 85
2.18. Bicelle preparation 85
2.19. Circular Dichroism 86
2.20. Landau free energy expansion analysis 86
2.21. Small Angle X-ray Scattering (SAXS) 88
2.21.1. Room temperature measurements 88
2.21.2. Temperature scans 89
2.22. Small Angle Neutron Scattering (SANS) 89
2.22.1. Protein SANS 89
2.22.2. Magnet SANS 90
2.23. Analysis of scattering data 91
2.23.1. *Ab initio* shape determination from SAXS data 91
2.23.2. Scattering pattern prediction from atomic resolution structures 91
2.23.3. Ensemble Optimisation Method 92
2.23.4. SasView 92
3. Expression, purification, activity and deuteration of AgrA and AgrC 94
3.1. Full-length AgrA 94
3.1.1. Expression 94
3.1.2. Purification 96
3.1.3. Activity assay 98
3.2. AgrC 99
3.2.1. Cloning 100
3.2.2. Expression 100
3.2.3. Purification 101

3.3. Size exclusion chromatography for AgrA/C complex 103
3.4. Deuteration of $AgrC_{206}$ 108
4. Investigation of AgrA and AgrC protein structures and molecular interactions at room temperature 112
4.1. Structural characterisation of $AgrC_{206}$ and AgrA 113
4.2. The molecular interaction between AgrA and AgrC and the role of ATP 131
4.2.1. Small Angle X-ray Scattering 132
4.2.2. Small Angle Neutron Scattering 148
4.3. Effect of IQS analogues on the tertiary structure of AgrA 151
4.4. Discussion 159
5. Thermal denaturation as a tool to probe molecular interactions in the *agr* two-component system 163
5.1. The role of ATP in the molecular interaction of AgrA and AgrC ... 165
5.2 Effect of the IQS analogues on the thermal stability of AgrA 186
5.3 Thermal studies of the *agr* two-component system by Small Angle Scattering 195
5.2. Discussion 213
6. Sample environment development for SANS magnetic studies on biological samples 220
6.1. Magnetically aligned bicelles in morphological studies of membrane proteins using SANS 220
6.2. Discussion 244
7. Additional work 246
7.1. C-terminal AgrA 247
7.1.1. Expression 247
7.1.2. Purification 248
7.1.3. Activity assay 249
7.2. N-terminal AgrA 250
7.2.1. Cloning 250
7.2.2. Expression 250
7.2.3. Purification 252
7.3. AgrB 257
8. Discussion 260
9. Conclusion and future work 267
References 269

Appendix..... 279

List of abbreviations

agr	Accessory gene regulator
AgrA	Accessory gene regulator protein A
AgrB	Accessory gene regulator protein B
AgrC	Accessory gene regulator protein C
AgrC cyto	cytosolic domain of AgrC
AgrD	Accessory gene regulator protein D
CA domain	Catalytic and ATP binding domain
CD	Circular Dichroism
DDM	n-Dodecyl-B-D-Maltoside
DHPC	1,2-diheptanoyl-sn-glycero-3-phosphocholine
DHp domain	Membrane proximal dimerisation and histidine phosphotransfer domain
DL-SAXS	Diamond-Leeds SAXS
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DSC	Differential scanning calorimetry
DTT	1,4-dithiothreitol
EM	Electron Microscopy

EMSA	Electrophoretic Mobility Shift Assay
EOM	Ensemble Optimisation Method
EPR	Electron Paramagnetic Resonance
FPLC	Fast Performance Liquid Chromatography
GAJOE	Genetic Algorithm Judging Optimisation of Ensembles
Hla	α-haemolysin
Hlb	β-haemolysin
HmBRI	*Haloarcula marismortui* Bacteriorhodopsin
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
IQS	Integrated Quorum Sensing
IQS3	IQS analogue number 3
IQS11	IQS analogue number 11
IQS15	IQS analogue number 15
ITC	Isothermal titration calorimetry
NMR	Nuclear Magnetic Resonance
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
ROS	Reactive oxygen species

sae	*S. aureus* exoprotein expression regulation
SANS	Small Angle Neutron Scattering
sar	Staphylococcal accessory regulator
SarA	Staphylococcal accessory regulator A
SAS	Small Angle Scattering
SAXS	Small Angle X-ray Scattering
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SLD	Scattering Length Density
SRCD	Synchrotron Radiation Circular Dichroism
TAE	Tris-acetate-EDTA
TEMED	Tetramethylethylenediamine
ToF	Time-of-flight

List of figures

Figure 1.1 **The accessory gene regulator system.**

Figure 1.2 **Suggested topology of AgrB.**

Figure 1.3 **Crystal structure of the ATP binding domain of AgrC.**

Figure 1.4 **Crystal structure of the C-terminal domain of AgrA in complex with its DNA-binding site in order to stabilise the protein.**

Figure 1.5 **Dimensionless Kratky plot.**

Figure 1.6 **CD profiles corresponding to different secondary structures.**

Figure 1.7 **Boltzmann analysis curve.**

Figure 1.8 **Change in cooperativity between protein and protein-partner**.

Figure 3.1.1 **Expression of C199S AgrA.**

Figure 3.1.2 **Purification of full-length AgrA.**

Figure 3.1.3 **Activity assay of full-length recombinant C199S AgrA.**

Figure 3.3.1 **N-terminal AgrA expression test ± sorbitol.**

Figure 3.4.1 **Expression of $AgrC_{206}$.**

Figure 3.4.2 **Purification of $AgrC_{206}$.**

Figure 3.5.1 **Size exclusion chromatography of AgrA+AgrC$_{206}$.**

Figure 3.5.2 **Size exclusion chromatography of AgrA+AgrC$_{206}$+ATP.**

Figure 3.5.3 **Size exclusion chromatography of AgrA+AgrC$_{206}$+sATP.**

Figure 3.5.4 **SEC fractions of AgrC+AgrA+nucleotides.**

Figure 3.6.1 **Expression of AgrC$_{206}$ in minimal medium for deuteration.**

Figure 3.6.2 **Purification of ^{2}H-AgrC$_{206}$.**

Figure 4.1.1 **Experimental SAXS data of AgrC$_{206}$.**

Figure 4.1.2 **Flexibility analysis of AgrC$_{206}$.**

Figure 4.1.3 **Particle characterisation of AgrC$_{206}$.**

Figure 4.1.4 **SAXS simulation of MD-annealed AgrC$_{206}$ monomer.**

Figure 4.1.5 **SAXS simulation of AgrC$_{206}$ dimer dominant cluster.**

Figure 4.1.6 **AgrC$_{206}$ ensemble optimisation method.**

Figure 4.1.7 **MD frame of AgrC$_{206}$ identified by GAJOE.**

Figure 4.1.8 **GCN4-AgrC$_{209}$ fit to experimental data.**

Figure 4.1.9 **Experimental SAXS data of C199S AgrA.**

Figure 4.1.10 **Flexibility analysis of C199S AgrA.**

Figure 4.1.11 **Particle characterisation of C199S AgrA.**

Figure 4.1.12 **Superimposition of the dominant clusters of MD simulated AgrA.**

Figure 4.1.13 **SAXS simulation of C199S AgrA dominant clusters.**

Figure 4.1.14 **C199S AgrA ensemble optimisation method.**

Figure 4.1.15 **MD frame of C199S AgrA identified by GAJOE.**

Figure 4.1.16 **Topology of AgrA N-terminus.**

Figure 4.2.1 **Experimental scattering data of AgrA, AgrC$_{206}$ and AgrA+AgrC$_{206}$.**

Figure 4.2.2 **Particle characterisation of the 1:1 AgrA and AgrC$_{206}$ mixture.**

Figure 4.2.3 **Particle characterisation of the 1:2 AgrA and AgrC$_{206}$ mixture.**

Figure 4.2.4 **Complexation determination plots of the AgrA and AgrC$_{206}$ mixture.**

Figure 4.2.5 **Complexation determination plots of the AgrA and AgrC$_{206}$ mixture in low salt.**

Figure 4.2.6 **MONSA-calculated bead model of AgrA and AgrC$_{206}$ complex.**

Figure 4.2.7 **Close-up of the residues predicted to be involved in the phosphorylation of AgrA.**

Figure 4.2.8 **MONSA-calculated bead model of AgrA and AgrC$_{206}$ mixture (1:2).**

Figure 4.2.9 **Effect of nucleotides on AgrC$_{206}$.**

Figure 4.2.10 **Particle characterisation of AgrC$_{206}$+ATP.**

Figure 4.2.11 **Particle characterisation of AgrC$_{206}$+sATP.**

Figure 4.2.12 **Effect of nucleotides on the AgrA and AgrC$_{206}$ mixture.**

Figure 4.2.13 **Flexibility analysis of AgrA+AgrC$_{206}$ with nucleotides.**

Figure 4.2.14 **Complexation determination plots for AgrA and AgrC mixture and ATP.**

Figure 4.2.15 **Complexation determination plots for AgrA and AgrC mixture and non-hydrolysable ATP.**

Figure 4.2.16 **BioSANS of C199S AgrA and ^{2}H-AgrC$_{206}$.**

Figure 4.3.1 **Experimental scattering data of AgrA and IQS analogues.**

Figure 4.3.2 **Experimental scattering data of AgrA and IQS analogues.**

Figure 4.3.3 **Flexibility plots of full-length AgrA (A), C-terminal AgrA (B) and IQS analogues.**

Figure 4.3.4 **IQS analogues P(r) distribution.**

Figure 4.4.1 **Proposed pipeline for the processing of Small Angle Scattering data of proteins.**

Figure 5.1.1 **CD spectra and secondary structure of AgrA.**

Figure 5.1.2 **Landau analysis of the thermal unfolding of AgrA.**

Figure 5.1.3 **CD spectra and secondary structure of AgrC.**

Figure 5.1.4 **Landau analysis of the thermal unfolding of AgrC.**

Figure 5.1.5 **CD spectra and secondary structure of AgrA+AgrC (1:1).**

Figure 5.1.6 **Landau analysis of AgrA and AgrC in 1:1 molar ratio.**

Figure 5.1.7 **CD spectra and secondary structure of AgrA+AgrC (1:2).**

Figure 5.1.8 **Landau analysis of AgrA and AgrC in 1:2 molar ratio.**

Figure 5.1.9 **Transition cooperativity.**

Figure 5.1.10 **Manual Boltzmann fit of the double transition in AgrC.**

Figure 5.1.11 **Manual Boltzmann fit of the double transition in AgrA+AgrC+ATP.**

Figure 5.1.12 **Landau analysis of AgrC.**

Figure 5.1.13 **Landau analysis of AgrA and AgrC in 1:1 ratio.**

Figure 5.1.14 **Landau analysis of AgrA and AgrC in 1:2 ratio.**

Figure 5.1.15 **Transition cooperativity.**

Figure 5.2.1 **Secondary structure content of AgrA with and without IQS analogues at 20 °C.**

Figure 5.2.2 **Far UV CD spectra and secondary structure content of AgrA with IQS analogues as a function of temperature.**

Figure 5.2.3 **Landau analysis of AgrA with IQS analogues.**

Figure 5.2.4 **Far UV CD spectra and secondary structure content of C-terminal AgrA with IQS analogues.**

Figure 5.2.5 **Landau analysis of C-terminal AgrA with IQS analogues.**

Figure 5.3.1 **Simulated scattering profiles of $AgrC_{206}$ at 293 and 333 K.**

Figure 5.3.2 **Simulated scattering profiles of $AgrC_{206}$ at increasing temperatures.**

Figure 5.3.3 **Landau analysis of AgrC$_{206}$ from the SAXS data.**

Figure 5.3.4 **Landau analysis of C199S AgrA from the SAXS data.**

Figure 5.3.5 **Landau analysis of C-terminal AgrA from the SAXS data.**

Figure 5.3.6 **Landau analysis of the C199S AgrA and AgrC$_{206}$ mixture from the SAXS data.**

Figure 5.3.7 **Landau analysis of the C-terminal AgrA and AgrC$_{206}$ mixture from the SAXS data.**

Figure 5.3.8 **Thermal unfolding transition cooperativity of the Agr proteins in SAXS.**

Figure 5.3.9 **Melted protein in DL-SAXS capillaries**.

Figure 5.3.10 **Temperature-dependent scattering profiles of GAPDH.**

Figure 5.3.11 **Temperature-dependent scattering profiles of AgrC$_{206}$.**

Figure 5.3.12 **Temperature-dependent scattering profiles of AgrC$_{206}$ with error bars.**

Figure 5.3.13 **Radius of gyration of GAPDH as a function of temperature.**

Figure 5.4.1 **Possible two-component system cascades.**

Figure 6.1 **Schematic of a bicelle.**

Figure 6.2 **Schematic representation of bicelle orientation**.

Figure 6.3 **BioSANS Magnet.**

Figure 6.4 **Effect of magnetic field on 'negative' bicelles.**

Figure 6.5 **Time evolution in the 1D scattering profile.**

Figure 6.6 **2D intensity maps of 'negative' bicelles.**

Figure 6.7 **1D scattering profile of negative bicelles in magnetic field.**

Figure 6.8 **Intensity difference profile.**

Figure 6.9 **Schematics of (A) axial orientational distribution of bicelles in a horizontal magnetic field, and (B) a "bulging" bicelle.**

Figure 6.10 **SANS profiles before, during and after the bicellar transition as temperature increases.**

Figure 6.11 **SANS profile of the system during the cooling process**.

Figure 6.12 **SANS profile of the system during the cooling process by axis.**

Figure 6.14 **2D intensity map of DMPC/DHPC bicelles in the presence of 9 mM $YbCl_3$.**

Figure 6.15 '**Positive' bicelles with 9 mM $YbCl_3$.**

Figure 6.16 **Schematic of stacked positive bicelles.**

Figure 6.17 **2D intensity map of positive bicelles with high $YbCl_3$.**

Figure 6.18 **Structure of the HmbRID94N trimer.**

Figure 6.19 **Scattering ^2H-HmbRID94N in positive bicelles.**

Figure 6.20 **Experimental set-up.**

Figure 7.1.1 **Expression of C-terminal AgrA.**

Figure 7.1.2 **C-terminal AgrA purification.**

Figure 7.1.3 **Activity assay of recombinant C-terminal AgrA.**

Figure 7.2.1 **N-terminal AgrA expression test ± sorbitol.**

Figure 7.2.2 **Expression of N-terminal AgrA in minimal medium.**

Figure 7.2.3 **N-terminal AgrA purification.**

Figure 7.2.4 **Size exclusion chromatography of N-terminal AgrA.**

Figure 7.2.5 **Size exclusion chromatography of N-terminal AgrA.**

Figure 7.3.1 **Scattering pattern of AgrB ± AgrD in DDM micelles**.

Figure A1 **Chemical structures of IQS analogue number 3 (A), 11 (B) and 15 (C).**

Figure A2 **Guinier analysis of C199S AgrA (blue), AgrC$_{206}$ (green), the mixture in a 1:1 ratio (red) and in a 1:2 ratio (purple) (AgrA:AgrC).**

Figure A3 **Guinier analysis of AgrC$_{206}$ alone (green), with ATP (blue) and non-hydrolysable ATP (purple).**

Figur A4 **Guinier analysis of the AgrA/AgrC$_{206}$ mixture (in a 1:1 molar ratio) alone (blue), with ATP (orange) and non-hydrolysable ATP (green).**

Figure A5 **Guinier analysis of AgrA alone (green), with IQS3 (yellow), IQS11 (red) and IQS15 (blue).**

Figure A6 **Guinier analysis of C-terminal AgrA alone (green), with IQS3 (yellow), IQS11 (red) and IQS15 (blue).**

Figure A7 **NMR spectra of DMPC/DHPC (3:1) bicelles.**

Figure A8 **NMR spectra of DMPC/DHPC (3:1) bicelles with 5 mM YbCl$_3$.**

List of tables

Table 2.1 Resolving gel composition.

Table 2.2 Stacking gel composition.

Table 2.3 PCR mix for EMSA P2-P3 probe synthesis.

Table 3.1.1 Growth rate of *E. Coli* BL21 (transformed with pCOLD I-C199S AgrA) in presence and absence of sorbitol (OD).

Table 4.3.1 R_g of AgrA and C-terminal AgrA in presence and absence of IQS analogues.

Table 4.3.2 Real space analysis of AgrA and IQS analogues.

Table 5.2.1 Melting temperatures of AgrA with IQS analogues.

Table 5.2.2 Melting temperatures of AgrA with IQS analogues.

Table 5.3.1 R_g values for C199S AgrA at different temperatures.

Table A1. Primers for cytosolic AgrC cloning.

Table A2. Primers for EMSA probes synthesis.

1. Introduction

"At this stage, such ideas are merely speculative, but now that a crystalline protein has been made to give X-ray photographs, it is clear that we have the means of checking them and, by examining the structure of all crystalline proteins, arriving at far more detailed conclusions about protein structure than previous physical or chemical methods have been able to give." (Bernal and Crowfoot, 1934)

These words mark the beginning of the long journey on the road to protein structure determination. Less than one hundred years ago the first X-ray profile of a protein was recorded. Since then, scientific discovery has been faster than ever before and protein crystallography became the most powerful and widely used method to investigate high-resolution structure, solving over 100,000 protein structures.

Although the principle "structure is key to function" was always the foundation of such work, there is yet another aspect that has been somewhat overlooked, and that is the due consideration to the physiological conditions of a protein. More specifically, soluble proteins are physiologically present in an aqueous environment which guarantees a certain degree of mobility; the same can be said for temperature, which is normally found between 0 and 40°C in nature. By locking the protein in a specific conformation influenced by a series of physical constraints, including keeping the systems at liquid nitrogen or helium temperatures, a significant portion of the biologically relevant information is lost. Indeed restricting protein mobility is one of the key principles of crystallography. The natural flexibility of a protein is a key factor in many

molecular interactions and in understanding the different biological roles it plays in a comprehensive manner.

The characterisation of protein structure and function is a challenging undertaking to this day. Many are notoriously difficult to produce and isolate due to the specific conditions required for correct folding and solubilisation. In addition, no biophysical technique presently available is free from challenges and limitations. Likewise, none of them, when used in isolation, can give a somewhat perfect representation of protein structure and behaviour. For this reason, it is important for these approaches to come together to each give their contribution towards structural characterisation and create a cohesive environment for protein biochemistry.

This is the overall aim of this work: to bring together complementary computational and biophysical techniques such as small angle X-ray and neutron scattering, circular dichroism, molecular dynamics simulations and nuclear magnetic resonance and apply them to the characterisation of the biology of the two-component system in *Staphylococcus aureus* quorum sensing system.

1.1. *S. aureus* and quorum sensing system

Staphylococcus aureus is a Gram-positive pathogen often known for its multi-drug resistance. It is commonly responsible for systemic infections in hospitals, biofilm formation on surgical equipment and implants and food spoilage. An example of a particularly severe consequence of infection is post-operative bacteraemia which kills approximately 10-30% of patients that develop it (van Hal et al., 2012, Bader, 2006, Yilmaz et al., 2016). The three main regulatory systems are responsible for *S. aureus* virulence: the accessory gene regulator (*agr*) system, the *S. aureus* exoprotein expression regulator (*sae*) and the staphylococcal accessory regulator (*sar*).

SaeRS is a two-component regulatory system involved in the production of exoproteins at the transcriptional level. *S. aureus sae* mutants show decreased virulence and lower production of haemolysins α and β (Hla and Hlb) (Giraudo et al., 1994, Giraudo et al., 1997). The accessory regulator system *sar* acts as a regulator of virulence in both an *agr*-dependent manner, acting as a transcriptional aid via the protein SarA, and an *agr*-independent manner, as a transcription factor of a series of target genes including that coding for collagen adhesin (Blevins et al., 1999).

The accessory gene regulator (*agr*) acts as quorum-sensing system in *S. aureus*. This class of systems is commonly found in bacteria. The quorum sensing allows the bacterial population to act as a community and coordinate response to external cues. The release of signalling peptides into the environment allows the colony to detect its own cell density and the reaching of a threshold triggers a change in the behaviour of the population (Williams et al., 2007).

As the bacterial population increases, so does the concentration of a specific quorum sensing signal molecule, called autoinducer, which is initially produced at a basal level and exported into the surrounding environment. In a growing population the QS signal molecule accumulates to reach a threshold concentration at which it autoinduces its own biosynthesis through binding to and activating a transcriptional regulator. This then drives the expression of the QS signal molecule synthase so that more QS signal molecules can be generated as well as activating or repressing target structural genes. For *S. aureus*, activation of the *agr* results in switching from a tissue colonisation stage to a cytotoxic lifestyle releasing the exotoxins that drive establishment of infection (Williams, 2007, Bassler, 2002); proteins promoting cell adhesion (e.g. protein A) are upregulated in the initial phase of colonisation and, as the exponential phase reaches its end, they are downregulated in favour of exotoxins and other virulence factors (Chan et al., 2004). This allows the systemic spreading of infection and consequent increased availability of nutrients and host immune response evasion (for example by upregulating leukocidins and α-toxin targeting blood cells) (DeLeo et al., 2009).

The quorum-sensing system in *S. aureus* is the *agr* system. It has four components, AgrA, B, C and D, and each of them is encoded by the corresponding genes *agrA, B, C* and *D*. These genes are part of the *agr* locus and are collectively regulated by the promoter P2 (Jensen et al., 2008). The rest of the locus encodes for virulence factors and exoproteins and is controlled by the promoter P3. The transcripts of the P2 and P3 promoters are called RNAII and RNAIII respectively. RNAIII represses adhesion proteins as well as encoding toxins and is therefore present in the later phases of bacterial growth (late and

post-exponential phases) while cell surface proteins are expressed during the early exponential phase.

The signal molecule secreted by *S. aureus* is called auto-inducing peptide, or AIP. The precursor of this auto-inducing peptide produced as a quorum sensing signal is AgrD, ribosomally synthesised and processed through a two-step proteolytic event by transmembrane proteases AgrB and MroQ (Stock et al., 2022).

The 7-9 residue mature AIP (Ji et al., 1997) is characterised by a thiolactone macrocycle formed *via* condensation of the C-terminal carboxyl group and the sulfhydryl group of an internal cysteine (Chan et al., 2004, Ji et al., 1997, McDowell et al., 2001). This specific structure is necessary for binding to and activating AgrC (Lyon et al., 2002, Chan et al., 2004). AgrC is the transmembrane receptor component of a two-component system present in the bacterial membrane. It is a histidine kinase and the only cellular protein capable of binding to AIP (Lina et al., 1998). Upon binding to AIP, AgrC phosphorylates the intracellular, soluble protein AgrA. AgrA is the effector protein, which binds to the *agr* P2 and P3 which drive the *agr* system. These promoters regulate the synthesis of the *agr* components as well as virulence and colonisation factors (Chan et al., 2004, Williams, 2007).

S. aureus strains are classified into four *agr* specificity groups which can either cross-activate or cross-inhibit the *agr* system (Ji et al., 1997, Jarraud et al., 2000, Jensen et al., 2008).

The *agr* regulome is extensive and mostly related to the transcript RNAIII. RNAIII has been linked to the up-regulation of expression of α-haemolysin and

β-haemolysin (Novick et al., 1993, Morfeldt et al., 1995) and to the down-regulation of *spa* (Novick et al., 1993), the staphylococcal protein A. Protein A is a B cell toxin able to cause cell death in its targets (Goodyear and Silverman, 2003), therefore aiding immune evasion. Alpha-haemolysin is a contributor in skin and soft tissue infections (Kobayashi et al., 2011, Kennedy et al., 2010) and pneumonia (Bubeck Wardenburg et al., 2007), with PSMα also being related to skin infections (Kobayashi et al., 2011). *psmα* and *psmβ* are positively upregulated by direct promoter binding of AgrA (Queck et al., 2008). RNAIII was also found to be involved in the regulation of a serine protease operon, the *spl* (Reed et al., 2001, Dunman et al., 2001). The *spl* is associated with biofilm formation, adhesion and immune evasion via its control over proteins involved in these processes (Paharik et al., 2016). Coagulase is an extracellular protein which leads to the conversion of fibrinogen to fibrin by associating with prothrombin (Baron, 1996). The coagulase gene, *coa*, is both down- and up-regulated by the *agr* system, based on the growth phase (Lebeau et al., 1994). RNAIII has also been reported as an antagonist of Rot, also a regulator in *S. aureus* (Geisinger et al., 2006). Rot and the *agr* have opposite effects on gene expression, with Rot upregulating genes for cell-surface proteins while downregulating certain virulence factors including *hlb* (Saïd-Salim et al., 2003). Another important protein for surface adhesion and immunomodulation in *S. aureus* is MAP (Harraghy et al., 2003); MAP expression was found to be up-regulated by RNAIII via an interaction with its mRNA (Liu et al., 2011). Other genes that are up-regulated by *agr* include *hlgB* and *C* (encoding γ-haemolysin components), *set8* (exotoxin 2), *clfB* (clumping factor B), *arc* and others

involved in amino acid metabolism, *epif* (ABC-type multidrug transport system) (Dunman et al., 2001).

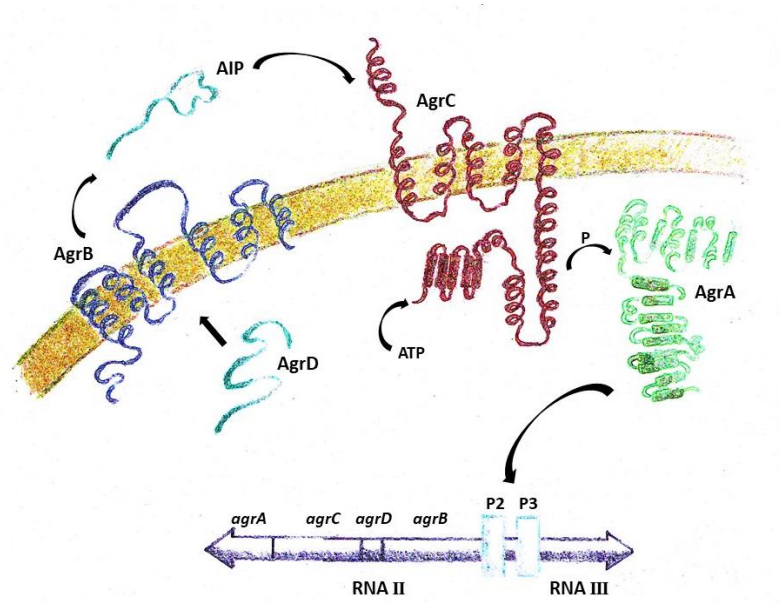


Figure 1.1 ***The accessory gene regulator system.*** *Synthesis of AgrD, like the rest of the agr proteins, is regulated by promoter P2. P2 allows the synthesis of the* agr *proteins upon AgrA binding at basal levels. AgrD inserts into the plasma membrane and is processed into AIP by AgrB. At some point in the process AgrD/AIP flips into the outside environment and the α-helix that keeps it attached to the membrane is cleaved off. AIP selectively binds to AgrC, which in turn phosphorylates AgrA. Phosphorylation of AgrA promotes its dimerisation.*

1.1.1. AgrB and the proteolysis of AgrD

AgrB is the transmembrane endopeptidase which processes AgrD into AIP (Wang et al., 2015). Topology studies revealed that both the N-terminus and the C-terminus of AgrB are in the cytoplasm and the N-terminus is rich in

hydrophilic residues (Zhang et al., 2002). Although the full structure of AgrB is not yet resolved, it is suggested to have 6 transmembrane regions.

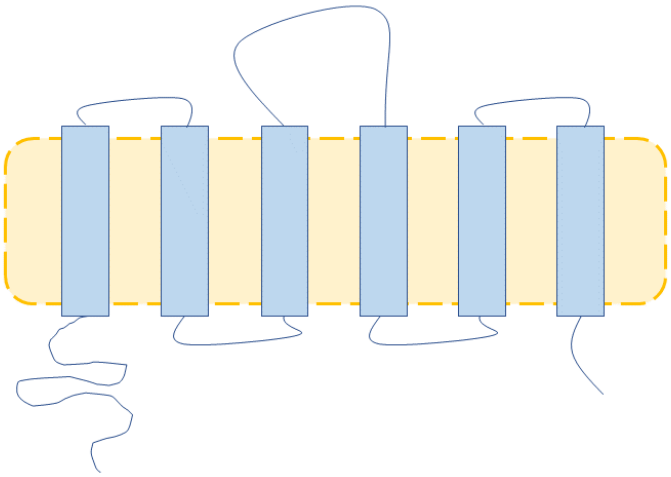


Figure 1.2 ***Suggested topology of AgrB.*** *Drawn as described in Zhang et al, Journal of Biological Chemistry, 2002.*

Later investigations reported AgrB to have 5 transmembrane regions instead, with both termini in the extracellular environment (Thoendel and Horswill, 2013).

Acccessory gene regulation in *S. aureus* is group specific. This means that AgrB is responsible for processing AgrD in a group specific way. It has been shown that the first transmembrane helix and the first loop on the N-terminus (residues 42 to 66) are key for the group-specific processing of AgrD in group I *S. aureus* (Zhang and Ji, 2004). However, in group II the specificity is given by segments in the second and fourth transmembrane domains plus the fourth loop (residues 67 to 75, 126 to 141) (Zhang and Ji, 2004). The second and fourth transmembrane domains are hydrophilic while the rest of the helices are hydrophobic. It is possible that the protein adopts a channel-like shape with the

two hydrophilic regions on the inside, which interact with AgrD (Zhang et al., 2002).

A key residue in the processing of AgrD by AgrB is C84 (Qiu et al., 2005). In particular, AgrB was shown to have a key role in the cyclisation (Wang et al., 2015).

Recently a new protein, MroQ, was found to be involved in the processing of AgrD. Its deletion was associated with an accumulation of AgrD intermediate therefore suggesting that MroQ and AgrB have a complementary activity (Marroquin et al., 2019, Stock et al., 2022). The process of export of the processed AIP remains unclear.

At present, a full structure of AgrB has not been obtained. SAXS was used in this study to complement MD simulations and wet-lab data (obtained by other members of the group, see Publications) in an effort to structurally and functionally characterise AgrB.

1.1.2. Two-component systems – AgrA and AgrC

Sensing signals is the basis for many biological processes. Whether we consider eukaryotic or prokaryotic systems, there will always be a way for cells to detect external stimuli and respond to them. Two-component systems play a fundamental role in this. They are a favourite signalling system in prokaryotes and represent on average 1-2% of the bacterial genome (Salvado et al., 2015). The typical bacterial two-component system comprises of a dimeric transmembrane histidine kinase receptor, which transduces the signal from the extracellular environment into the cell, and a DNA-binding response regulator,

which is the final recipient of the stimulus and provokes a change in the gene expression. They tend to share similar structural features: the histidine kinase receptor is characterised by a transmembrane sensory element, a dimerisation and histidine phosphotransfer domain (DHp) and a catalytic and ATP-binding domain (CA); dimerisation is key to the function of the histidine kinase receptor: the catalytic domain is responsible for acquiring the phosphate group from ATP and transferring it to a conserved histidine residue on the opposing monomer (auto phosphorylation). The phosphate group is then transferred from the histidine on the receptor to a highly conserved aspartate residue on the response regulator. Phosphorylation often causes a conformational change in the response regulators, which can be single-domain or multi-domain proteins. This change in conformation often regulates the affinity and binding of the regulator to DNA, thus driving gene expression (Capra and Laub, 2012).

The acceptor histidine is generally located in the same region which provides the interface for the dimerisation of the kinase, called the H-box. Other relevant sites for the phosphorylation process are the so-called G1 and G2 boxes (glycine residues), which are thought to be important for nucleotide binding, and the F and N boxes, whose function remains to be clarified (Swanson et al., 1994). The DHp domain is also thought to be responsible for substrate recognition and specificity (Tomomori et al., 1999, Laub and Goulian, 2007).

Two members of *S. aureus* accessory gene regulator system, AgrA and AgrC, constitute a typical two-component system and are responsible for the downstream activation of the *agr* system.

AgrC is the histidine kinase receptor in the staphylococcal quorum sensing system. It recognises and binds the autoinducer peptide AIP on the extracellular side of the membrane and phosphorylates AgrA in response. It presents the typical topology of a kinase, with a transmembrane domain and a cytosolic domain linked by a flexible region. The transmembrane domain, also called sensor module, consists of 5 transmembrane helices (Lina et al., 1998) corresponding to residues 1 to 205. The N-terminus of this domain contains the recognition site for AIP. The cytoplasmic domain has the histidine kinase activity which responds to ligand binding and is divided into the 'membrane proximal dimerization and histidine phosphotransfer' domain (DHp) and the catalytic domain which binds ATP (CA) at the C-terminus. AgrC shows many of the canonical characteristics of histidine kinase receptors: it is dimeric and has the typical 4 helix bundle. It is activated by *trans* autophosphorylation following interaction with the signal molecule AIP (Cisar et al., 2009, Xie et al., 2019). According to the structure published by Xie *et al.* (Xie et al., 2019), the phosphorylation in *trans* is unusual: the handedness of the DHp loops connecting the two helices is usually indicative of type of autophosphorylation, with left handed loops associated with *cis* and right handed loops with *trans* (Ashenberg et al., 2013*)*. AgrC, despite having a left-handed loop, was shown to autophosphorylate in *trans* (Xie et al., 2019). When active, AgrC phosphorylates AgrA, which is a requisite for its dimerisation with consequent increase in affinity for the P3 promoter (Khodaverdian et al., 2013).

The atomic-level structure of full-length AgrC is unknown. A crystal structure describes a shortened construct of the cytosolic domain of AgrC, corresponding to its ATP binding region (residues 234 to 430). Attempts to crystallise the whole

cytosolic domain were unsuccessful as proteolysis was observed once the protein was put into crystallisation drops (Srivastava et al., 2014). The structure of the CA domain (catalytic and ATP binding, residues 278-430) was determined and consists of 6 β-strands arranged in one β-sheet and stacked against 4 helices (Fig. 1.4) (Srivastava et al., 2014). A recently published crystal structure describes a chimeric construct, GCN4-AgrC$_{209}$, where a leucine zipper was attached to the end of the dimerisation helices (Xie et al., 2019) (Fig. 1.3).

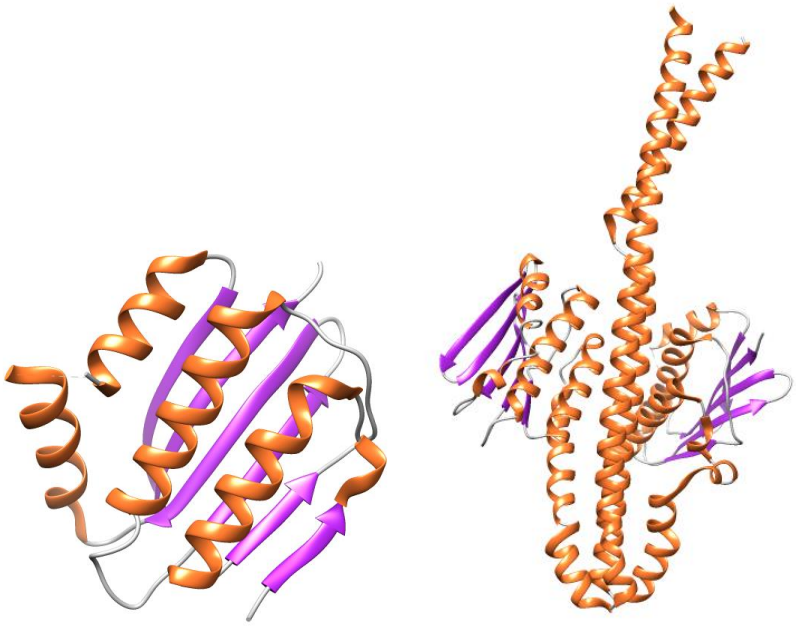


Figure 1.3 ***Crystal structures of AgrC.*** *Crystal structures of the catalytic domain (CA) of AgrC (residues 278-430), PDB entry: 4BXI DOI: Srivastava, S.K., Rajasree, K., Fasim, A., Arakere, G., Gopal, B. (2014) J Bacteriol 196: 2876 https://doi.org/10.2210/pdb4BXI/pdb Literature citation: Srivastava, S.K., Rajasree, K., Fasim, A., Arakere, G., Gopal, B. (2014) J Bacteriol 196: 2876 https://doi.org/10.1128/JB.01530-14 taken from (Srivastava et al., 2014) (left) and of the GCN4-AgrC$_{209}$ chimeric construct, PDB entry: 6E95 PDB DOI: Xie, Q., Zhao, A., Jeffrey, P.D., Kim, M.K., Bassler, B.L., Stone, H.A., Novick, R.P., Muir, T.W. (2019) Cell Chem Biol 26: 548-558.e4 https://doi.org/10.2210/pdb6E95/pdb Literature citation: Xie, Q., Zhao, A., Jeffrey, P.D., Kim, M.K., Bassler, B.L., Stone, H.A., Novick, R.P., Muir, T.W. (2019) Cell Chem Biol 26: 548-558.e4 https://doi.org/10.1016/j.chembiol.2019.01.006 taken from (Xie et al., 2019) (right).*

The chimeric construct was designed to investigate the effect of torsion angles of the helices on the activity of the protein at the level of the catalytic domain (Xie et al., 2019).

The binding of AIP to AgrC has an effect at the level of the AgrC autophosphorylation (Wang et al., 2014), while AIP binding does not seem to affect the phosphorylation of AgrA once AgrC is phosphorylated. There is also no evidence of AgrC acting as a phosphatase on AgrA (Wang et al., 2014). The lack of direct effect of AIP binding on AgrA phosphorylation likely explains the basal levels of AgrA binding to the P2 and P3 operons even when the *agr* system is in its inactive state (Wang et al., 2014).

Although the full structure of AgrC has not been solved, evidence suggests that the linker segment between the transmembrane region and the catalytic and DHp regions adopts a helical conformation (Wang et al., 2014). Mutagenesis experiments suggest that two residues, Leu 205 and Met 208, are particularly important in stabilising the helical shape of the linker (Wang et al., 2014). A GCN4-$AgrC_{209}$ chimera was recently published suggesting that the characteristics of the CA and DHp domains in this construct could resemble their *in vivo* behaviour (Xie et al., 2019). It is hypothesised that the CA domain is retained by two helices belonging to the DHp domain when the protein is inactive, blocking the phosphorylation process. According to this mechanism, a hydrogen bond is created between residues Arg 238 and Gln 305 to form a 'latch' and sequester the CA domain in its inactive conformation. Upon AIP binding, this hydrogen bond is broken and the histidine residue at position 239 becomes available for phosphorylation (Xie et al., 2019).

The outlines of each domain vary slightly between reports. For example, (Srivastava et al., 2014) indicate residues 278-430 as the catalytic and ATP binding domain, while (Xie et al., 2019) suggest the CA domain to start from residue 291 and the DHp domain to end at residue 285.

AgrA is the response regulator of the two-component signal transduction pathway in *S. aureus* quorum-sensing system (Morfeldt 1996) and was first recombinantly produced by Peng et al. in 1998 (Peng et al., 1988).

In its native environment, AgrA acts both as a monomer and as a dimer and is a DNA-binding protein. The DNA-binding domain of this class of proteins needs to have a specific conformation that allows it to interact with DNA. Initially it was hypothesised that AgrA required the assistance of SarA, another DNA-binding protein, to bind to the promoters (Novick, 2003). AgrA was subsequently shown to interact with the promoters autonomously and its DNA-binding domain was successfully crystallised in complex with them (Sidote et al., 2008).

The transcription factor AgrA regulates two promoters called P2 and P3, and these in turn control two operons. The P2 operon is responsible for the production of all components of the *agr* system (AgrA, B, C and D). The promoter P3 regulates the production of RNAIII, which is associated with the production of exotoxins and virulence factors (Chan et al., 2004, Williams, 2007, Gordon et al., 2013, Queck et al., 2008).

The binding of AIP to the receptor AgrC results in the autophosphorylation of AgrC from cellular ATP and ultimately in the phosphorylation of a conserved aspartate residue on AgrA (D59) (Khodaverdian et al., 2013). Phosphorylation

of AgrA makes it prone to dimerization and it has been shown that the affinity of dimeric AgrA for P3 is enhanced compared to monomeric AgrA, while there is no evidence that the affinity for P2 is affected by dimerisation (Khodaverdian et al., 2013).

According to the currently acknowledged mode of action of AgrA, its C-terminus interacts with the promoters and has DNA-binding properties, while the N-terminus is responsible for the interaction with AgrC. At present, the full structure of AgrA is not available; its C-terminus is the only domain fully characterised with its structure resolved by crystallography (Sidote et al., 2008, Leonard et al., 2012, Rajasree et al., 2016).

Structural studies of many of these DNA-binding proteins have revealed that the most common configuration adopted by their DNA-binding domains include the helix-turn-helix (HTH) and the 'winged HTH' motifs. AgrA, however, together with a few other proteins, was found to exhibit an unusual DNA-binding domain at its C-terminus. This domain was only observed in the AgrA/ComE family of proteins (ComE is a DNA-binding protein found in *Neisseria* that is thought to aid natural DNA transformation (Chen and Gotschlich, 2001)). This domain is found, with slight variations, in different species such as *B. subtilis* and *S. aureus*. The nomenclature for this domain is species specific: it was named LytTR ('litter') in *S. aureus.* This domain is 105 residues in length and responsible for the regulation of virulence factors and toxins. The structure of the LytTR domain of AgrA exhibits 10 β-strands, arranged into 3 antiparallel β-sheets, and a two-turn α-helix not involved in DNA binding (Sidote et al., 2008).

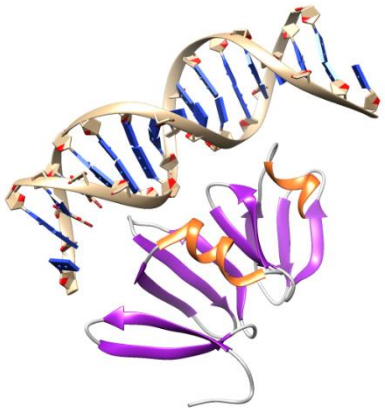


Figure 1.4 ***Crystal structure of the C-terminal domain of AgrA in complex with its DNA-binding site in order to stabilise the protein.*** *PDB entry 3BS1.DOI: Sidote, D.J., Barbieri, C.M., Wu, T., Stock, A.M. (2008) Structure 16: 727-735 https://doi.org/10.2210/pdb3BS1/pdb Literature citation: Sidote, D.J., Barbieri, C.M., Wu, T., Stock, A.M. (2008) Structure 16: 727-735 https://doi.org/10.1016/j.str.2008.02.011.*

Shown in Figure 1.4 are the three β-sheets, one in the middle made of 5 strands and the two flanking, smaller ones made respectively of 2 and 3 strands and the N-terminus and at the C-terminus. The middle sheet presents itself together with the only α-helix of the domain which, again, is not involved in the binding of DNA. This conformation is characterised by a two-fold symmetry between strands β1-5 and β6-10. A series of salt bridges stabilises the protein: D157-H208, R195-E141, D176-K223 and H174-E226. The final residues at the N-terminus of this domain (C-) of AgrA are extended, while they most likely adopt a different conformation in the full-length protein (Sidote et al., 2008).

Because of the lack of a full structural characterisation of AgrA, there is no evidence that the published structure of the C-terminus is maintained in the full-

length protein. Insights into the structure of the N-terminus will be of great importance in understanding the phosphorylation process and how AgrA interacts with AgrC.

1.1.3. Inhibition of AgrA – IQS analogues

Antibiotic and antimicrobial resistance is one of the biggest challenges to public health that science is facing in the modern age since the advent of antibiotic treatments. In light of this and the pathogenicity of *S. aureus*, it is particularly important to investigate potential ways of disrupting the infection cycles of the pathogen without instigating mutations leading to the creation of new antibiotic resistant strains to add to the ones already present within the community. This is achieved by preventing infection in a way that does not kill the bacterium but simply makes it harmless to the human body. As the *agr* system regulates virulence factor production, it is a primary target for this type of research.

Small molecule regulators of the *agr* system have been identified in the past with varying degrees of efficacy. Previous studies Yanin Jaiyen discovered that a specific metabolite of *Pseudomonas aeruginosa*, 2-(2-hydroxyphenyl)-thiazole-4-carbaldehyde (HPTC) and a series of closely related analogues, was capable of inhibiting the interaction between AgrA and its P2 and P3 DNA targets (Jaiyen, 2016). Indeed *P. aeruginosa* has been known to suppress staphylococcal infections in its own favour. HPTC was called the integrated quorum sensing (IQS) molecule and reported to be a quorum sensing signal molecule (Lee et al., 2013). However this has subsequently been shown to be incorrect (Paiva, 2019). IQS is in fact a well-known by-product formed during the biosynthesis of the siderophore, pyochelin (Cornelis, 2020).

Results suggest that this molecule carries out its inhibitory role by binding to AgrA, the response regulator, on or in the proximity of its DNA-binding region, thus disrupting its interaction with the promoters P2 and P3.

Jaiyen investigated the *agr* inhibitory efficacy of IQS and various chemically synthesised IQS analogues (Dr Alex Truman). Out of the >50 analogues, numbers 3, 11 and 15 were selected to conduct our study, IQS3 as an example of active AgrA inhibitor, IQS15 as an example of structurally related but inactive inhibitor and negative control, and IQS11 which is also structurally related to IQS but targets virulence factor production via an *agrP3* promoter-independent pathway.

Their structures are shown in the Appendix.

A combination of Circular Dichroism and Small Angle X-ray and Neutron Scattering are employed to provide insights into the structure and interactions of AgrA, AgrC and AgrB. These techniques and some elements of the associated data analysis are described in the next sections.

1.2. Small Angle Scattering

Scattering is a physical technique historically used to study a variety of systems relating to a range of disciplines. The diversity and versatility of this methodology make it possible to apply it to the study of biological. It made a first appearance in the 1930s and innovations in the instrumentation and analysis of data brought it to the forefront of science in the 1970s, when the number of facilities across the world boomed (Svergun et al., 2013).

In small angle scattering (SAS) a beam of particles is directed towards the sample which will scatter the particles according to its own structure and composition at small angles. The pattern of the scattering is recorded and processed to determine structure, state or properties of the sample. The radiation can be of different types, ranging from X-rays and neutrons to visible light. Scattering is normally used and most sensitive when applied to the study of molecules in the size range 10-1000 Å, therefore being particularly useful in the study of proteins (Cousin, 2015). The proteins of the *agr* system fall right into the size range that is detectable using small angle scattering, approximately in the range of 20-55 Å. Structural studies normally use X-rays around 10keV in energy, corresponding to a wavelength of 0.10-0.15, or neutrons in the wavelength range of 0.1-1.0 nm. Scattering exists in two forms, elastic and inelastic. For the purpose of this work, small angle scattering is implied to be elastic, which means that there is no energy transfer between the sample and the radiation (Svergun et al., 2013).

The starting point of SAS data analysis is the identification of a Guinier region, which refers to the initial slope of the spectrum. Defining this sets the starting parameters that will be used throughout the data processing of soluble proteins.

In fact, even just being able to identify a Guinier region suggests sample monodispersity. This portion refers to the region in which the Guinier law applies and corresponds to $0 < q < \frac{1}{R_g}$ (Svergun et al., 2013). The Guinier plot presents q^2 on the abscissa and $\ln[I(q)]$ on the ordinate, and also shows the distribution of residuals (also *vs* q^2). The latter is especially indicative of data quality. A flat, unbiased distribution of residuals represents a good, monodisperse sample. A 'smiley' distribution of residuals suggests aggregation in the sample while a 'frowny' distribution may indicate interparticle repulsion (Jacques and Trewhella, 2010). The limits of a Guinier region should be $q_{max} * R_g < 1.3$ (Jacques and Trewhella, 2010) and $q_{min} * R_g > 0.65$.

The conversion of data from reciprocal space to real space yields the pair-distance distribution function $p(r)$. The $p(r)$ is obtained via an Inverse Fourier Transform and provides the radius of gyration and the $I(0)$ in real space and the maximum distance between two points within the sample, called D_{max}. This parameter is especially important for downstream analysis as it informs on the size of the system. In the p(r) plot, *r* is plotted on the abscissa and *p(r)* on the ordinate. *p(r)* functions should end smoothly at D_{max}, and irregularities such as a trailing end or shoulder may indicate aggregation, while a negative tail suggests interparticle repulsive forces (Svergun et al., 2013).

The degree of folding and globularity of a protein sample can be qualitatively assessed using a Kratky plot. This plot is characterised by an abscissa with q and an ordinate with $q^2 * I(q)$ but can often be represented as a dimensionless Kratky plot ($q * R_g$ *vs* $\frac{I(q)}{I(0)} * (q * R_g)^2$). This has the advantage of allowing a more uniform comparison between proteins. The rate of decay is indicative of

folding: fully folded proteins are characterised by q^{-4} rate of decay, partially folded proteins by a q^{-2} decay and intrinsically disordered proteins or fully denatured proteins present an asymptotic behaviour (Fig. 1.5).

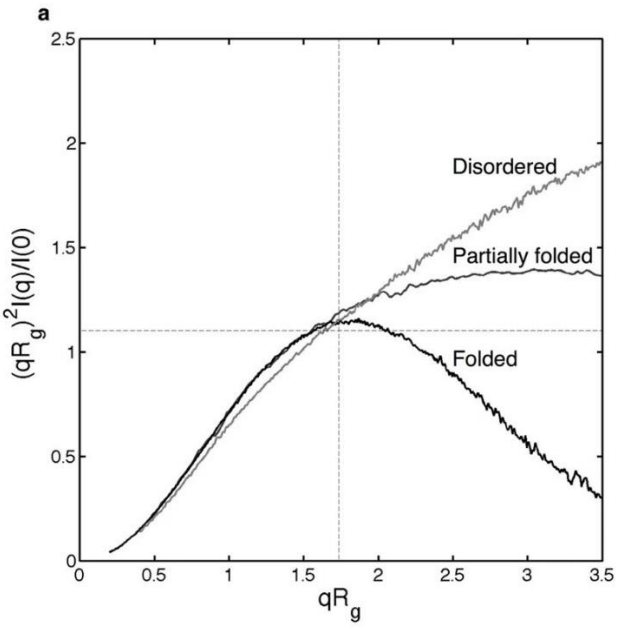


Figure 1.5 ***Dimensionless Kratky plot.*** *Reused from BURGER, V. M., ARENAS, D. J. & STULTZ, C. M. 2016. A Structure-free Method for Quantifying Conformational Flexibility in proteins. Scientific Reports, 6, 29040. Copyright © 2016, The Author(s). https://doi.org/10.1038/srep29040, https://creativecommons.org/licenses/by/4.0/ (Burger et al., 2016) (*note that only part of the original figure is reproduced here).*

The molecular weight of protein samples can either be extrapolated by automated processing or manually inferred from the Porod volume. Both methods are not very accurate in the calculations and therefore this feature is disregarded from further analysis. The Porod volume will be discussed as a measure of particle density, together with flexibility analysis. This feature is

particularly relevant for soluble proteins. In this respect, Kratky and normalised Kratky plots were also used as a qualitative indication of the folding of proteins.

Ab initio modelling of proteins from Small Angle X-ray Scattering (SAXS) was successfully performed and validated in a number of occasions, both for individual proteins and molecular complexes (Graebsch et al., 2009, Chacón et al., 2000, Svergun et al., 2013). DAMMIN (Svergun, 1999) and MONSA (Svergun, 1999, Svergun and Nierhaus, 2000) were specifically implemented in the creation of *ab initio* models of protein complexes (Diepholz et al., 2008). MONSA is bead modelling software operating using a genetic algorithm. The model is composed of spheres with radius much much smaller than the radius of the molecule under investigation and identifies the conformation that best fits the experimental data (Chacón et al., 2000). SAS, due to its comparatively low-resolution, is ideally placed for the study of flexible proteins and has been used in conjunction with other techniques (including modelling and EM) to characterise flexible structures (Tsutakawa et al., 2011, Yang et al., 2010, Svergun et al., 2013). Temperature-related investigations of proteins have traditionally been neglected, these techniques remaining more popular in the lipid community (Svergun et al., 2013). Until recently, SAS data were mostly used for *a posteriori* validation of otherwise obtained experimental structures (Svergun et al., 2013). Recent advances in software development enable more precise structural analysis, as described in Chapter 4.

1.2.1. Small Angle X-ray Scattering (SAXS)

Small Angle X-ray Scattering is probably the most widely used small angle scattering technique. X-rays are scattered by the electronic cloud surrounding the atoms and for this reason are not penetrating. X-ray sources are classified according to their brightness, or number of particles emitted per unit area (Svergun et al., 2013); the two main sources, both based on electron or positron acceleration, are synchrotron radiation or benchtop laboratory sources (Svergun et al., 2013).

X-rays are highly damaging to the sample at variable rates depending on the intensity of the beam. While this is important to consider when planning a SAXS experiment, the higher intensity also shortens the necessary acquisition time to record high quality data (Svergun et al., 2013) and B21 offers the option to keep the sample flowing throughout the measurement to avoid radiation damage.

Angles used in SAXS are in the range of 0.1 to 5 degrees and normally used for materials of 1-100nm, while Wide Angle X-ray Scattering is used for interrogating smaller length scales.

Conformational analysis of *agr* protein samples investigated by SAXS was combined with homology modelling and molecular dynamics simulations to obtain a motionally modulated atomistic narrative of proteins and protein complexes in solution. The protein response to temperature was also investigated through SAXS studies.

1.2.2. Small Angle Neutron Scattering (SANS)

Beams of neutrons interact with and are scattered by atomic nuclei thanks to strong nuclear force. In the field of protein studies, SANS present two major useful properties: the non-destructive behaviour of the beam (of particular interest when dealing with a particularly expensive or laborious-to-produce sample), and contrast matching. The neutron scattering length is different for different isotopes of the same element and this allows, specifically, for selective labelling. The most common application of this for scatterers is the use of deuterium oxide (^{2}H or 'heavy water') instead of the more naturally abundant hydrogen, ^{1}H. The coexistence of these two isotopes often produces large incoherent scattering acting as background noise, and for this reason researchers try to control the isotope concentration in their samples. Obtaining deuterated sample is often extremely expensive due to the cost of pure deuterium oxide, especially in the case of biological samples. Despite this, the advantages are undisputable: selective deuteration and contrast matching allow to study individual components of complex systems; for biological samples, this means being able to study a component in a state that more closely resembles native conditions, especially when this is part of a complex, pathway or simply requiring specific conditions to retain structure.

The two most common sources of neutron beams are nuclear reactors and spallation sources. These two present different advantages: reactors provide a constant flux of neutrons, are less prone to interruptions and generally considered more reliable for experiments whose measurements take several hours (e.g. reflectometry); spallation sources are more efficient in terms of

neutron production and the fact that they produce neutrons at different speeds can be exploited for experiment optimisation.

Contrast matching will be used to visualise each of the proteins of *agr* two-component system when interacting with the other and investigate any conformational rearrangement resulting from introducing the phosphor-donor ATP and its non-hydrolysable analogue sATP into the system. This will improve our understanding of the dynamics adopted by the two proteins with respect to their function and each other.

1.3. Circular Dichroism

The properties of our *agr* two-component system were elucidated by analysing its response to temperature. Circular Dichroism was used to quantify the variation of secondary structure as an indication of thermal denaturation.

Circular dichroism follows the difference in absorption of left-handed circularly polarised light (L-CPL) and right-handed circularly polarised light (R-CPL) as a function of frequency and it is measured in units of Δε, or difference in absorbance (Greenfield, 2006a). It occurs only when a molecule contains one or more chiral chromophores. If one or more of these chiral chromophores are present, then one of the two polarisations will be absorbed with preference at a given frequency. More specifically, the CD signal will be positive if (L-CPL) is greater than (R-CPL), while it will be negative when (R-CPL) is greater than (L-CPL).

CD spectra report on the secondary structure of a protein: the spectrum is different for α helices and β sheets, and again if the protein is unstructured (Whitmore and Wallace, 2008, Lopes et al., 2014). The spectrum of a protein is not obtained by summing the spectra of each residue but is also influenced by the 3-dimensional structure of the molecule.

The CD spectrum of a protein is highly representative of its secondary structure, and presents distinct characteristics associated with helices, strands and turns. A highly helical protein is characterised by negative minima at 208 and 222 nm, while β-strand dominated proteins exhibit a minimum at 218 nm (Greenfield, 2006b) (Fig. 1.6).

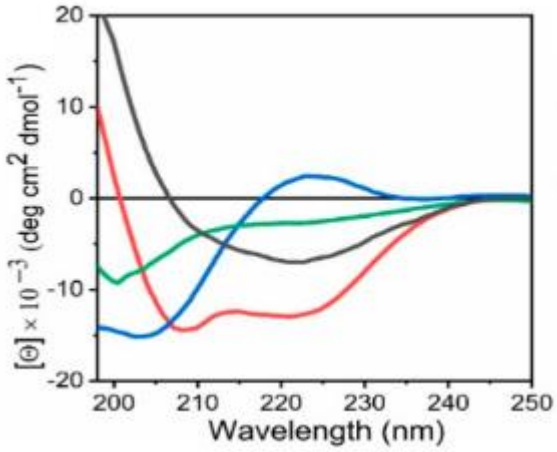


Figure 1.6 ***CD profiles corresponding to different secondary structures.*** *α-helix is shown in red, β strand in black, polyproline II in blue and random coil in green. Reused from PIGNATARO, M. F., HERRERA, M. G. & DODERO, V. I. 2020. Evaluation of Peptide/Protein Self-Assembly and Aggregation by Spectroscopic Methods. Molecules, 25, 4854. https://doi.org/10.3390/molecules25204854 © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations. (Pignataro et al., 2020) (*note that only part of the original figure is reproduced here).*

In the context of the thermal unfolding of proteins, if the change in the CD spectra as a function of temperature is reversible, then it is possible to derive the van't Hoff enthalpy (ΔH), and entropy (ΔS), of the unfolding, the free energy of unfolding (ΔG), the midpoint of the unfolding process (T_m), and the heat capacity change (ΔC_c), (Greenfield, 2006a). All of these define the thermodynamic characteristics of the unfolding process.

Changes in the levels of α-helices and β-sheets therefore become indicators of the level of structural degradation of the protein. It has been shown in previous studies that the amount of α-helices decreases as the protein approaches its melting temperature while the β-sheets increase, prior to the complete collapse of the structure and onset of an aggregated state (Hussain et al., 2018).

1.4. Phase transitions and the Landau free energy expansion analysis

This study explored the behaviour of proteins as they approach the transition temperature, to understand their molecular interactions and to develop a sensitive assay of such interactions. A recently developed method for studying protein interactions by monitoring the thermodynamic parameters of protein unfolding by Circular Dichroism (CD) was applied (Hussain et al., 2018).

Proteins retain their native, active conformation and secondary structure (α-helix and β-sheet) content upon heating until they reach their melting temperature, which is protein-specific. They are subjected to random thermal fluctuations as temperature increases, which eventually lead to protein unfolding. Once the melting temperature is reached, proteins undergo denaturation until their secondary structure is transformed into sheet-rich, amyloid-like aggregates (Hussain et al., 2018). This transition occurs at different rates which are dependent on the protein itself and on interaction partners, which may stabilise (or destabilise) its structure. These variations in thermal behaviour are indicative of molecular interaction.

Two states are different if there is at least one physical property that varies between the two, e.g. density. The observation of the physical properties of the system allows for identification of the transition point, for example the transition temperature, at which the system changes state. The transition normally happens at certain temperatures and pressures. The transition point is generally revealed by a discontinuity (opposed to the continuous variation during each state) in the observed physical property (Tolédano and Tolédano, 1987).

The transition point was first referred to by Thomas Andrews in 1869, who defined it as the ‘critical point’. In 1873 Josiah Willard Gibbs first gave a proper definition to the term ‘critical point’ (Jaeger, 1998). This was made possible by the introduction, also due to Gibbs, of the first phase diagram. By definition: the critical point is ‘the location in a phase diagram where the phase coexistence curve terminates’ and the latent heat vanishes as first order phase transitions become second order (Gibbs, 1873, Jaeger, 1998). As demonstrated by van der Waals, the critical point can be determined with something called an equation of state. An equation of state is the relationship between a set of parameters that altogether define the state of the system (Britannica).

In biological systems, the Landau theory (developed from the original classification of transitions) (Landau, 1936) has been applied to describe the phase transitions of lipids from the liquid to the crystalline phase and vice versa. But, because the first order transition is universal, it can also be applied to a variety of systems, including the structure of proteins. Proteins are very sensitive to heat, which can alter their structure, properties and activity, and a temperature above the critical point causes protein denaturation. This means that the protein loses its secondary structure completely and becomes inactive.

The Boltzmann analysis for protein melting, as described by Greenfield in 2006 (Greenfield, 2006a) and used for the analysis of thermal unfolding of proteins from Circular Dichroism data, relates to a two-state exchange process, best fitting equilibrium states far from the transition temperature. The transition is fitted mathematically to locate the melting temperature at midpoint and to characterise the width of the transition (related to cooperativity). The Landau free energy series expansion, by contrast, is performed at the melting point and

is valid immediately below the melt within the folded state of the system. As such, it complements the existing analysis reflecting the contribution from thermal fluctuations in protein structure on its approach to melting.

From data fitting, it was observed the proteins of our system undergo a first order transition while the Boltzmann analysis curve (Fig. 1.7) resembles a second order phase transition. First order phase transitions are those in which a change in the physical properties of the material is observed, e.g. a difference in the volume (Ehrenfest classification). Second order phase transitions were defined to accommodate a class of transitions that were not identifiable by a clear physical change, but only by a discontinuity in the heat capacity (Jaeger, 1998). In other terms, first order phase transitions present a discontinuous first derivative of the Gibbs free energy, while second order phase transitions have a continuous first derivative but a discontinuous second derivative (Jaeger, 1998).

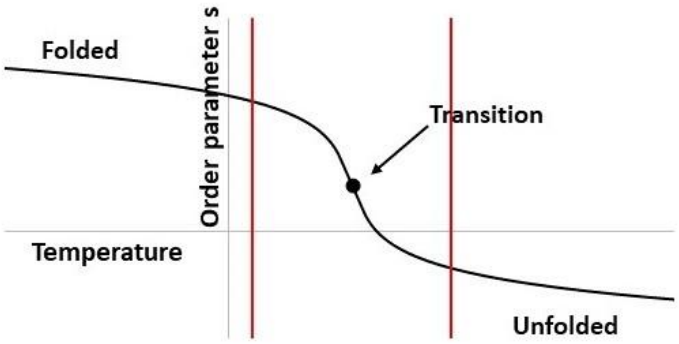


Figure 1.7 ***Boltzmann analysis curve.*** *The Boltzmann analysis is accurate outside the red lines, the transition point is placed in the middle and the curve is fitted based on that solely. The proposed method focuses on the proximity of the transition, i.e. between the red lines.*

The Gibbs free energy ('the greatest amount of mechanical work which can be obtained from a given quantity of a certain substance in a given initial state, without increasing its total volume or allowing heat to pass to or from external bodies, except such as at the close of the processes are left in their initial condition' (Gibbs, 1873)) is normally defined as the difference between the enthalpy and the product of the entropy times the temperature of the system but can be re-written to describe the process of protein unfolding:

(1.4)

$$\Delta G = -RTlnK$$

where R is the gas constant (1.98 cal/mol), T is the absolute temperature in Kelvin and K is the constant of folding K=[F]/[U] (Greenfield, 2006a).

The Gibbs-Helmholtz equation, fitted with CD at a single wavelength, describes the energetics of the thermal unfolding of proteins:

(1.5)

$$\Delta G = \Delta H(1 - \frac{T}{T_m}) - \Delta Cp[(T_m - T) + Tln(\frac{T}{T_m})]$$

where ΔH is the transition enthalpy, ΔC_p is the change in heat capacity, T is the absolute temperature and T_m is the denaturation temperature obtained when the protein is half folded and half unfolded.

The Gibbs free energy can be expanded in terms of a thermodynamic order parameter(Hussain et al., 2018, Morrow and Whitehead, 1988, Bonev and Morrow, 1997). The order parameter is a quantity characterising the two phases across the transition, or more specifically, the degree of freedom across the

phases in the system undergoing a phase transition (McNaught and Wilkinson, 1997).

(Hussain et al., 2018) introduced the reduced temperature τ, a dimensionless parameter that centres the thermodynamic description at the transition point:

(1.6)

$$\tau = 1 - \frac{T}{T_m}$$

describing the separation of a physical state from the transition. This parameter was chosen so that it is always positive below the transition and with values between [0;1].

Using the reduced temperature, the Gibbs-Helmholtz equation can be re-written following (Hussain et al., 2018):

(1.7)

$$\Delta G = \Delta H(\tau) - \Delta C_p \left[T_m \tau + T ln(1 - \tau)\right]$$

Now we want to consider a different approach and introduce an alternative definition of the free energy at the transition using the Landau formalism.

The free energy of the system at the transition can be expressed as a series in terms of a thermodynamic order parameter *s*:

(1.8)

$$G(s) = \sum_{n=0}^{\infty} G^n(s_0) \frac{s^n}{n!}$$

The order parameter of choice is the α-helical content of the protein. The behaviour of this parameter is monotonic and was normalised so that [0;1], 0 being no α-helical content i.e. completely unfolded, and 1 meaning that the protein is totally α-helical. The β-sheet content is also introduced as order parameter and defined as (1 – α helices).

Following Morrow & Whitehead (Morrow and Whitehead, 1988), Jahnig (Jahnig, 1981) and Bonev & Morrow (Bonev and Morrow, 1997), the Gibbs free energy is truncated after the fourth order term in terms of the thermodynamic order parameter s:

(1.9)

$$G = G_0 \left[\frac{s^4}{4} + \alpha(T - T_c) \frac{s^2}{2} + \beta(T - T_m)s \right]$$

The series was truncated so that n[0;4] but n=0 and n=3 were excluded in order to satisfy the symmetry conditions for the series expansion. Here T_c is the critical temperature and α and β <0 (Hussain et al., 2018).

The first derivative of the equation of state, defined in terms of the Gibbs energy, must vanish for a system at equilibrium:

(1.10)

$$\frac{\partial G}{\partial s} = s^3 + \alpha(T - T_c)s + \beta(T - T_m) = 0$$

The two spinodal points on the equation of state 1.9 (i.e. the borderline between metastable and unstable states described in Figure 1.8) are the maximum and minimum with s and are defined as vanishing second derivative of $G(s)$. These

allow us to determine the values of the remaining parameters in *G* in terms of the critical temperature T_c and the positive spinodal temperature T_+

(1.11)

$$\alpha = \frac{3s_+^2}{T_c - T_+} \qquad \beta = \frac{2s_+^3}{T_+ - T_m}$$

The transition is considered of first order when $T_m < T_c$ and of second order when $T_m > T_c$. For first order transitions, the equation of state becomes

(1.12)

$$T = T_+ + \frac{T_m - T_+}{2}\left[3\left(\frac{s - s_+}{s_+}\right)^2 + \left(\frac{s - s_+}{s_+}\right)^3\right]$$

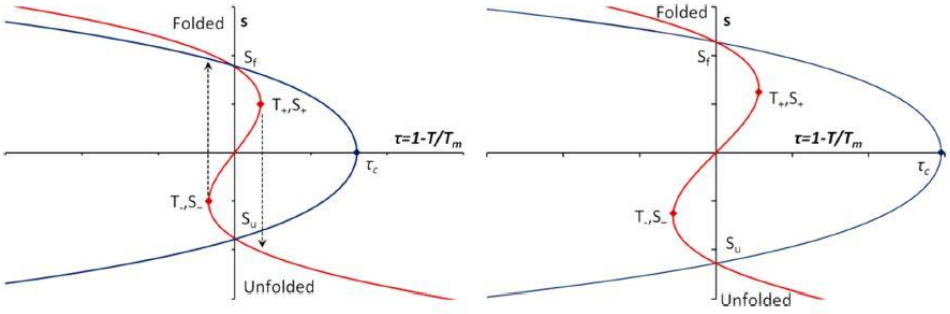


Figure 1.8 ***Change in cooperativity between protein and protein-partner****. The equation of state (red) and the coexistence curve (blue). s is the order parameter. T_+ and T_- are the spinodals. The spinodals are reached when the protein exists in a metaphase before collapsing. Reprinted with permission from HUSSAIN, R., HUGHES, C. S., JÁVORFI, T., SILIGARDI, G., WILLIAMS, P. & BONEV, B. B. 2018. To Boil an Egg: Substrate Binding Affects Critical Stability in Thermal Unfolding of Proteins. The Journal of Physical Chemistry. B, 122, 2213-2218. Copyright 2018 American Chemical Society.*

This analysis is used to quantify the 'strength' of the transition. Variation in this value between protein alone and protein with a potential binding partner indicate a molecular interaction is taking place and stabilises the protein secondary structure. The separation between T_+ and T_m can be used as a measure of the first order strength of the transition: the bigger $T_+ - T_m$ the more stable the protein is. We can also say that such transition is more cooperative.

The α-helical and β-sheet contents of a protein can be used to define a thermodynamic order parameter using Circular Dichroism, as shown by Hussain et al., 2018. The analysis described above was extended to small angle scattering

thermal studies. A related physical parameter can also be monitored as a function of temperature in SAS studies and this parameter was chosen to be the molecular radius of gyration R_g. The radius of gyration can be described broadly as the approximation of a particle of arbitrary shape with a sphere of radius R_g, which would scatter in a similar way. While CD monitors changes in the secondary structure content, R_g relates to the protein tertiary folding. Its value decreases with the increase in temperature as a result of increasing solvent access into the core of the protein. This is due to disappearance of hydrogen bonding, which is broken up by the increasing thermal fluctuations.

A good thermodynamic order parameter would be a normalised quantity, *s* being between [0;1). For this reason, *s* can be defined:

(1.13)

$$s = \frac{R_g}{R_{gmax}}$$

This makes it a dimensionless and normalised parameter which is also positive and [0;1). R_{gmax} is identified among the R_g values before the transition and normally found at lower temperatures, when the protein is in a stable state.

In both the CD and small angle scattering analysis, the parameter $\frac{T_m}{T_+}$ describes the transition strength and is related to the reduced temperature τ. It allows comparative analysis of all protein melts using a single scale limited by $T = 0$ from below and the spinodal point for each system from above. The parameter is dimensionless and takes values between 0 and 1. It is defined as

(1.14)

$$\tau = 1 - \frac{T_m}{T_+}$$

with weaker transitions appearing close to 0 and stronger transitions approaching 1.

1.5. Magnetically aligned bicelles in morphological studies of membrane proteins using SANS

The structural characterisation of membrane proteins is often a challenge due to their stability being dependent on a hydrophobic environment. In Small Angle Scattering, the detergent or lipidic environment surrounding membrane proteins represents a significant contribution to the scattering of the sample, therefore masking a lot of the characteristics of the protein and limiting its use. X-rays are unable to distinguish between lipid/detergent and protein and therefore, given that the scattering comes from their interaction with the electronic cloud making them non-penetrative, the profile shows a major contribution from the lipid or detergent. Neutron Scattering takes full advantage of contrast matching in this instance, allowing to contrast-match the lipid or detergent and obtaining a scattering profile entirely representative of the protein.

As part of this studentship, sample environment development for SANS magnetic studies on biological samples was carried out at the ISIS Neutron and Muon Source. An electromagnet suitable for biological samples was purposefully designed by the engineers to investigate the potential of magnetically aligned bicelles in the study of the morphology of membrane proteins by SANS. A more detailed introduction to these systems is provided in Chapter 6.

1.6. Hypothesis and Aims

The main objective of this work was to obtain a comprehensive insight into the secondary and tertiary structure characteristics and stability of the proteins of *S. aureus agr* quorum sensing system through independent observations.

SAXS was combined with *in silico* modelling methods to characterise the structure of AgrA and the cytoplasmic domains of AgrC ($AgrC_{206}$) in near-native conditions. Flexibility analysis was carried out and solution SAXS data were used to validate atomistic structures obtained by homology modelling and MD simulations. The aim was to describe different SAXS data analysis approaches in an effort to create a pipeline for soluble proteins structural characterisation.

The Landau formalism was applied to the proteins of the *agr* system, with the intention of expanding the methodology to include thermal unfolding data obtained from SAXS. This dual approach would allow to quantitatively monitor the unfolding of secondary structure through CD, and tertiary structure through SAXS.

To characterise the interaction between AgrC and AgrA *in vitro*, both thermal studies and room-temperature SAXS and SANS were used to evaluate the effect of ATP and its non-hydrolysable analogue sATP on this interaction. It is hypothesised that sATP will create a stable complex by preventing the phosphorylation of AgrC and, in turn, AgrA.

This work was aimed at showing how complementary biophysical, wet-lab and computational techniques can be used to structurally and functionally characterise soluble proteins.

Membrane proteins are notoriously difficult to study, due to low yields and hydrophobicity. One of the objectives of this work was to develop a new system for SANS studies of membrane proteins using magnetically aligned bicelles to form a uniform environment for dimensionality determination. This forms an independent sample environment development project for the ZOOM beamline at the ISIS Neutron and Muon Source, as required by the terms of this PhD funding.

2. Materials and Methods

2.1. Materials

All chemicals were purchased from either Sigma Aldrich or Thermo Fisher Scientific unless stated otherwise. All lipids were purchased from Avanti Lipids. The IQS analogues were synthesised by Dr Alex Truman. Spin concentrators were purchased from Merck Millipore. The NiNTA was purchased from Thermo Scientific. All primers were purchased from Sigma Aldrich.

2.2. Homology modelling and MD simulations

This work was carried out by Dr Boyan Bonev.

Homology modelling of AgrC has been reported previously (Sloan et al., 2019). The monomer model was obtained using I-TASSER (Zhang, 2008, Yang et al., 2015, Roy et al., 2010) and dimers were obtained using ClusPro (Kozakov et al., 2017). MD simulations were performed as previously described (Sloan et al., 2019). Homology modelling of full length AgrA was done using I-TASSER (Zhang, 2008, Yang et al., 2015, Roy et al., 2010) and the models were prepared for simulations using CHARMM-GUI (Jo et al., 2008). MD simulations were done using NAMD (Phillips et al., 2005) on an NVIDIA GPGPU Supermicro server for equilibration and shorter production runs. Scripts for larger systems and longer trajectory calculations were transferred to the regional scalar multicore HPC Midlands Plus cluster where 512 core parallel runs were roughly five times faster than the 16-core Supermicro 1U server fitted with K80 and K40 GPGPU accelerators from NVIDIA.

2.3. *E. coli* BL21 electrocompetent cells

LB broth (5 ml) was inoculated with *E. coli* BL21 and grown overnight at 37°C, 180rpm.

LB broth was inoculated with the overnight culture using a 1/100 dilution and incubated at 37°C, 180-200 rpm until it reached an OD of 0.4-0.6. The culture was cooled on ice for 30 minutes. The cells were harvested by centrifugation at 2500 *x g*, 4°C for 5 minutes. Cell pellet was washed with 10% sterile glycerol (v/v) three times with 5 minutes incubations. The final incubation was 20 minutes. The volume was reduced to 1 ml. Electrocompetent cells were aliquotted and stored in the -80°C freezer.

2.4. DNA gel electrophoresis

DNA gel electrophoresis was used to verify products of PCR and ligation reactions.

Agarose gels were made at 1% and 2% agarose. 0.5 g (1%) or 1 g (2%) of agarose were mixed with 50 ml 1x Tris-acetate-EDTA TAE buffer and microwaved until the solution was clear. The solution was poured onto the plate and mixed with 5 µl SYBR Safe DNA Gel Stain. The gel was allowed to set and 10 µl of sample were run at 80 V for ~30 minutes.

2.5. Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS PAGE)

The following tables describe the composition of the resolving and stacking gels for SDS PAGE.

Table 2.1 Resolving gel composition.

Reagent	**12%**	**15%**	**Or 15%**	**18%**
Water	2.19 ml	1.365 ml	1.1 ml	0.965 ml
Tris 1M pH 8	1.5 ml	1.5 ml	(1.5 M, pH 8.8) 1.3 ml	1.5 ml
SDS 10%	60 µl	60 µl	50 µl	60 µl
Acrylamide 30%	2.25 ml	3.075 ml	2.5 ml	3.45 ml
APS 10%	75 µl	75 µl	50 µl	75 µl
TEMED	7.5 µl	7.5 µl	5 µl	7.5 µl

Table 2.2 Stacking gel composition.

Reagent	**6%**	**4%**
Water	1.113 ml	1.5 ml
Tris 1M pH 6.8	0.47 ml	(0.5 M) 0.625 ml
SDS 10%	18.75 µl	25 µl
Acrylamide 30%	0.25 ml	0.335 ml
APS 10%	30 µl	25 µl
TEMED	3 µl	2.5 µl

Gel imaging was carried out on a LI-COR Odyssey Classic Imager.

2.6. Cloning

2.6.1. $AgrC_{206}$

The sequence of AgrC was retrieved from GenomeNet. The primers were designed using NEB Builder Assembly and can be found in the Appendix.

Two PCR reactions were set up, one to amplify the empty vector pCOLD-I and one for the insert.

The master mixes were prepared as follows:

AgrC insert:

Reagent	Volume	Final concentration
5x Q5 reaction buffer	40 µl	1x
Forward primer AgrC cyto	1 µl	0.5 µM
Reverse primer AgrC cyto	1 µl	0.5 µM
Generic template	5 µl	1 ng DNA/tube
Q5 Polymerase	2 µl	
dNTPs	4 µl	200 µM
Molecular grade water	150 µl	

pCOLD-I:

Reagent	Volume	Final concentration
5x Q5 reaction buffer	40 µl	1x
Forward primer pCOLD-I	1 µl	0.5 µM
Reverse primer pCOLD-I	1 µl	0.5 µM
pCOLD-I template	5 µl	1 ng DNA/tube
Q5 Polymerase	2 µl	
dNTPs	4 µl	200 µM
Molecular grade water	150 µl	

Ligation of AgrC cyto and the amplified vector was done using the HiFi method. Briefly, two master mixes for ligation were prepared, one with the AgrC cyto insert and one without.

Diluted vector (1/10)	1 µl
2x hi-fi mix	5 µl
Molecular grade water	3-4 µl
(AgrC cyto insert)	1 µl

The mixes were incubated at 50°C in the PCR block for 30 minutes and then desalted.

Both the + insert and the – insert reactions were transformed into DC10B *E.coli* competent cells by electroporation and plated onto ampicillin agar plates. The plates were incubated at 37°C overnight.

Single colonies from the positive plate were streaked onto ampicillin agar plates and incubated at 37°C overnight.

The plasmid was amplified by colony PCR using the pCOLD-I primers.

Colony PCR master mix set-up:

Reagent	Volume	Final concentration
Taq+dNTPs+buffer mix	225 µl	
pCOLD-I forward primer	15.75 µl	0.2-0.35 µM
pCOLD-I reverse primer	15.75 µl	0.2-0.35 µM
Water	193.5 µl	

The master mix was divided into 16 fractions and supplemented with some bacteria from each streak.

From 5 ml cultures, the plasmid was extracted from 4 colonies, using the Sigma Genelute Plasmid Miniprep Kit, and sent for sequencing together with a pCOLD-I forward primer.

2.6.2. N-terminal AgrA

This construct was created from the existing C199S AgrA construct by restriction digest reaction.

The inverse PCR reaction was set up as follows:

- Water 25 µl
- 5x Phusion buffer 10 µl
- 2 mM dNTPs 5 µl
- 10 µM Forward primer 4 µl
- 10 µM Reverse primer 4 µl
- 14 ng/µl Template 1 µl
- Phusion polymerase 1 µl

The PCR cycle consisted of 2 min at 98 °C, 19 cycles of 30 sec at 96 °C, 30 sec at 50-60 °C and 7 min at 72 °C, followed by 3 min at 72 °C. The product was cleaned using the Qiagen PCR Clean Up Kit.

The restriction digest reaction was set up as follows:

- Multicore 10x buffer 3 µl
- 170 ng/µl iPCR product 8 µl
- 10 mg/ml BSA 0.3 µl
- XbaI 1 µl
- Water 17.7 µl

The reaction was incubated at 37 °C, 60 rpm for 1.5 hours, heated to 65 °C for 20-25 minutes and cleaned using the Qiagen PCR Clean Up Kit.

The final construct was verified, after the ligation reaction, by colony PCR and sequencing.

2.7. *E. coli* BL21 transformation by electroporation

Up to 5 µl of plasmid was mixed with ~50 µl of competent BL21 *E. coli* cells (thawed on ice). The mixture (50-60 µl) was transferred into an electroporation cuvette. Electroporation was carried out using a BioRad MicroPulser Electroporator.

Following electroporation, 1 ml of LB was added to the mixture for recovery and the cells were incubated for 1 hour at 37°C, 200 rpm. 40-50 µl of the culture were plated onto LB agar plates containing 100 µg/ml of ampicillin. The plates were incubated at 37°C overnight.

2.8. Expression of recombinant AgrA in *E. coli* BL21

A single colony of BL21(pCOLD-I AgrA) was picked from the Amp LB agar plates and inoculated into ~5 ml terrific broth (containing 100 µg/ml ampicillin). The medium was supplemented with 1 M sorbitol. The culture was incubated at 37°C, 180-200 rpm overnight. Terrific broth (500 ml, with 100 µg/ml ampicillin) was inoculated with the overnight culture (1/100 dilutions). The culture was grown at 37°C until OD=0.6-0.9 was reached. The flask was put on ice for 30-40 minutes and expression was induced with 1 mM IPTG final concentration. The culture was incubated overnight at 16°C to express pCOLD. The culture was harvested (2500 x g, 4 °C, 15 minutes), washed with PBS, centrifuged at

3000 x g, 6 °C for 10 minutes and resuspended in lysis buffer (20 mM Tris pH 7.5, 10% glycerol, 5 mM $MgCl_2$). Stored at -80 °C.

The composition of terrific broth is as follows:

- Yeast extract 24 g
- Tryptone 12 g
- Glycerol 4 ml
- Water 900 ml
- D-sorbitol 180 g
- Potassium phosphate buffer 80 ml

Potassium phosphate buffer:

- KH_2PO_4 0.17 M
- K_2HPO_4 0.72 M

2.9. Expression of recombinant C-terminal AgrA and $AgrC_{206}$ in *E. coli* BL21

A single colony of *E. coli* BL21 (pCOLD-I AgrAc or AgrC cyto) was picked from the Amp LB agar plates and inoculated into ~5 ml terrific broth (containing 100 µg/ml ampicillin), and incubated at 37°C, 180-200 rpm overnight. Flask containing 500 ml terrific broth and ampicillin was inoculated with the overnight culture (1/100 dilutions). The culture was grown at 37°C until OD=0.6-0.9 was reached. The flask was put on ice for 30-40 minutes and expression was induced with 1 mM IPTG final concentration. The culture was incubated overnight at 16°C to express pCOLD. The culture was harvested (2500 x g, 4 °C, 15 minutes), washed with PBS, centrifuged at 3000 x g, 6 °C for 10 minutes and resuspended

in lysis buffer (20 mM Tris pH 7.5, 10% glycerol, 5 mM $MgCl_2$). Stored at -80 °C. For the terrific broth composition, see 2.8 (- sorbitol).

2.11. Expression of recombinant ^{2}H-AgrC206 in E. coli BL21

E. coli BL21 carrying the pCOLD-I_$AgrC_{206}$ plasmid were streaked onto LB agar plates supplemented with ampicillin (100 µg/ml final concentration) and grown overnight at 37 °C in a static incubator. Single colonies were picked and inoculated into 5 ml terrific broth (and ampicillin) cultures incubated at 37 °C, 180-200 rpm overnight. The overnight cultures were harvested by centrifugation, resuspended in 5 ml protonated M9 medium and incubated for 1.5 hours at 37 °C, 180 rpm for adaptation. Deuterated M9 medium flasks were inoculated using a 1:50 dilution and incubated at 37 °C, 180 rpm until OD=0.6-0.9. Once OD was reached, the cultures were incubated on ice for ~30 minutes before inducing expression with 1 mM IPTG final concentration. Cultures were then incubated at 16 °C, 180 rpm.

5x M9 salts – 500 ml

- Na_2HPO_4 15 g
- KH_2PO_4 7.5 g
- NH_4Cl 2.5 g
- NaCl 1.25 g

M9 medium – 1 L

- 5x M9 200 ml
- 1 M $MgSO_4$ 2 ml

- 1 M $CaCl_2$ 100 µl
- 20% pyruvate 12 ml
- 10 mg/ml Biotin/Thiamine 1 ml
- Vishniac elements 1x

2.12. Expression of recombinant HmbRI D94N

2.12.1. ^{1}H-HmBR

E. coli C41 were streaked onto LB agar plates supplemented with 50 µg/ml kanamycin from glycerol stocks and incubated at 37 °C overnight. Overnight cultures were grown in LB supplemented with 50 µg/ml kanamycin from single colonies identified on the plate, at 37 °C, 180-200 rpm.

LB flasks were inoculated using a 1:100 dilution and cultures incubated at 37 °C shaking at 180 rpm until OD=0.9-1.1 was reached. OD was measured using a Spectrophotometer. Protein expression was induced by adding 1 mM IPTG and 10 µM all-*trans* retinal (final concentrations) and incubating at 30 °C overnight.

Cultures were harvested by centrifugation at 2500 *x g* for 30 minutes, resuspended in ~30 ml LB and centrifugated at 3000 *x g*, 4 °C for 15 minutes. Pellet was stored at -80 °C.

2.12.2. ^{2}H-HmbR

E. coli C41 were streaked onto LB agar plates supplemented with 50 µg/ml kanamycin from glycerol stocks and incubated at 37 °C overnight. Overnight cultures were grown in LB supplemented with 50 µg/ml kanamycin from single colonies identified on the plate, at 37 °C, 180-200 rpm. LB was removed by

centrifugation and cultures resuspended in high salt M9 medium (see below) in D_2O, supplemented with 0.3% glucose and 50 µg/ml kanamycin.

1x High salt M9 medium (/L)

- KH_2PO_4 13 g
- K_2HPO_4 10 g
- Na_2HPO_4 9 g
- K_2SO_4 2.4 g
- NH_4Cl 1 g
- MgCl 10 mM
- Vishniac trace 1x
- Biotin/thiamine 1 ml

Cultures were equilibrated for ~2 hours before inoculating 500 ml with a 1/50 dilution into the same medium. Once OD = ~0.4 was reached, the cultures were centrifuged at 2500 x g, 22 °C for 30 minutes, resuspended in high salt M9 medium with 1.2% pyruvate instead of glucose and incubated at 37 °C, 180 rpm until OD = 0.9-1.1. Expression was induced with 1 mM IPTG and 10 µM all-*trans* retinal. OD was measured using a Genesys 10S VIS Spectrophotometer.

2.13. Electrophoretic Mobility Shift Assay (EMSA)

Electrophoretic Mobility Shift Assays (EMSAs) were performed as described by (Yakhnin et al., 2012). Briefly, native Tris-glycine gels were pre-run in the cold room until the current was invariant at 200 V (approximately 1-1.5 hour). The

recombinant AgrA/AgrAc was incubated in 1x reaction buffer with 5-10 nM P2 probe at 25°C for 25 minutes prior to the addition of 1x EMSA loading dye. Once mixed with the loading dye, the samples were immediately loaded into the Tris-glycine gels. The gels were run in the cold room at 200 V for 15 minutes and then at 120 V for 1.5 hour.

The results were analysed using an Odissey® Imaging System (LI-COR) at a wavelength of 700 nm.

Buffer, gel and loading dye compositions are described below.

5x reaction buffer (final concentrations):

- 50% glycerol
- 100 mM Tris pH 7.5
- 10 mM EDTA pH 7.5
- 500 mM NaCl
- 20 mM DTT
- Autoclaved Milli-Q water

6x EMSA loading dye:

- Autoclaved Milli-Q water 1 ml
- 0.5% bromophenol blue 21 µl

5x Tris-glycine running buffer:

- Tris-base 16.3 g
- Glycine 72 g

- EDTA (disodium salt) 1.8 g
- Distilled water 1 L

Diluted to 1x before use

Native Tris-glycine gel:

- HPLC water 2.7 ml
- 30% acrylamide 2 ml
- 50% glycerol 0.6 ml
- 1.5M Tris-base pH 8.8 1.2 ml
- 10% ammonium persulfate 60 µl
- TEMED 6 µl

2.14. Probes for Electrophoretic Mobility Shift Assays

The P2 (48 bp) and P2-P3 (~120 bp) probes were constructed using the sequences included in the Appendix. These were synthesised by Eurofins.

The synthesis of P2 only required an annealing reaction between the oligonucleotides. The ATTO700-labelled probe was made using oligonucleotides YJ65 and YJ66 at 10 µM final concentration. YJ66 and YJ70 were used to make the cold P2 probe. The construction of the P2-P3 probe was carried out with a PCR reaction using the YJ69 and YJ64 for the ATTO700-labelled probe and YJ62 and YJ64 for the cold probe.

The set-up was as follows:

Table 2.3 PCR mix for EMSA P2-P3 probe synthesis.

Reagent	**Stock**	**Final concentration**	**Volume**
Q5 buffer	5x	1x	10 µl
dNTPs	10 mM	200 µM	1 µl
YJ69 (YJ62)	10	0.5 µM	2.5 µl
YJ64	10	0.5 µM	2.5 µl
Q5 polymerase		1U	0.5 µl
Template	10 ng/µl	1 ng/µl	0.5 µl
Water			33 µl

The template used was P2-agrB in psk5630 amplified from SH1000 *agr* group 1, made by Dr Ewan Murray.

Below are listed the PCR conditions:

Denaturation	98°C	30 seconds
35 cycles	98°C	10 seconds
	58°C	20 seconds
	72°C	15 seconds
Extension	72°C	5 minutes

The P2-P3 PCR product was cleaned using the Wizard SV Gel and PCR Clean-up System.

2.15. Protein purification

2.15.1. Purification protocols for full-length, C-terminal and N-terminal AgrA and $AgrC_{206}$

Culture was thawed and sonicated using a macroprobe at amplitude 15 for 10-12 minutes, 15 sec on and 15 sec off. A 1/1000 dilution of protease inhibitor cocktail was added prior to sonication. The soluble fraction was isolated by centrifugation at 1000 x g, 8 °C for 12 minutes. Sodium chloride (300 mM) and imidazole (10 mM) were added to the supernatant. After column equilibration, protein binding was done on a Cytiva NiNTA HisTrap HP 5 ml.

After binding, the column was washed with high salt (50 mM NaP pH 7.8, 1.5 M NaCl) and buffer containing 50 mM NaP pH7.5, 300 mM NaCl, 10% glycerol v/v and 10 mM imidazole prior to elution. Elution was achieved by using gradient and fraction collection on either the AKTAPrime or AKTAStart. Protein was eluted in 50 mM NaP pH 7.5, 2 M NaCl, 20% glycerol v/v, 0.5 M imidazole and 10 mM DTT (no DTT for C-terminal AgrA), dialysed into 20 mM NaP pH 7.5, 0.5 M NaCl, 10% glycerol v/v, 1 mM EDTA and 5 mM DTT (1 mM DTT for C-terminal AgrA), quantified by absorbance spectroscopy and flash-frozen using liquid nitrogen.

2.15.2. Purification protocol for HmbRI D94N

Bacterial pellet was resuspended in lysis buffer (50 mM NaPi, 2 M NaCl, 1 mM PMSF, 14.7 mM β-mercaptoethanol, DNase and lysozyme, pH 8) and lysed by

sonication (amplitude 20, 1 s on 2 s off for 6 minutes, using a macroprobe). The insoluble fraction was isolated by centrifugation at 15000 rpm, 4 °C for 1 hour using an Avanti JA25.50 rotor. The insoluble fraction was solubilised in 50 mM NaPi, 2 M NaCl, 14.7 mM β-mercaptoethanol, DNase and 1% DDM, pH 8 at 37 °C overnight, shaking at 200 rpm. Solubilised sample was isolated by centrifugation at 3000 x g, 4 °C for 1 hour. Protein binding to NiNTA was carried out, following NiNTA equilibration in 50 mM NaPi, 2M NaCl at pH 8, by incubation at 4 °C on a rocker for 2 hours. Two washes were carried out prior to elution, in 50 mM NaPi, 2 M NaCl, 20 mM imidazole and 0.05 % DDM, pH 8 (1^{st} wash) and 50 mM NaPi, 2 M NaCl, 50 mM imidazole and 0.05% DDM, pH 8 (2^{nd} wash). Protein sample was eluted with 50 mM NaPi, 2 M NaCl, 500 mM imidazole and 0.1% DDM, pH 8 and dialysed into 250 mM NaPi, 1 M NaCl and 5 mM EDTA, pH 7.5. Sample concentration was done using 10000 MWCO spin concentrators and quantified by Nanodrop.

2.16. Incorporation of membrane proteins into lipid

Detergent and lipid stock was prepared using 25 mM DMPC and 100 mM DDM dissolved in 25 mM Tris, 100 mM NaCl and 0.5 mM EDTA at pH 7.5. The protein:lipid ratio used was 1:100. BioBeads were used to remove the DDM from the mixture (50-80 mg BioBeads per 100 µl sample). The sample was incubated with the BioBeads for 4 hours at 4 °C on a sample carousel spinning at 25 rpm. Protein incorporated into lipid was separated from the BioBeads by

centrifugation at 3000-5000 rpm for 5 minutes. Sample was further centrifugated at 15000 rpm for 10 minutes and supernatant discarded.

2.17. Size exclusion chromatography

Size exclusion chromatography to assess AgrA/AgrC complex formation was carried out as follows. A HiLoad Superdex S200pg preparative grade column was equilibrated in 20 mM NaP pH 7.5, 0.5 M NaCl, 5% glycerol v/v, 1 mM EDTA and 5 mM DTT. C199S AgrA and $AgrC_{206}$ were mixed in a 1:1 molar ratio. For the ATP and sATP (purchased from Sigma Aldrich) samples, 1 mM nucleotides and 1 mM $MgSO_4$ were added to the mixture. Samples were loaded using a 1 ml loop and eluted in 108 ml.

2.18. Bicelle preparation

Bicelles were prepared using DMPC and DHPC in a 3:1 molar ratio, respectively. Chloroform was removed from the mixture using a Rotavapor R-210 by Buchi. For SANS experiments, the lipid sample was hydrated with 25 mM NaPi, 0.1 M NaCl and 0.5 mM EDTA at pH 7.5 (in D_2O) obtaining 0.22 M (DHPC) and 0.66 M (DMPC) final concentrations.

Ytterbium chloride was added to bicelles as described in Chapter 6.

Bicelles containing ^{2}H-HmBRI D94N were prepared by using 4.39 mg of protein, 30 µmol of DMPC and 12 µmol of DHPC to compensate for lipid displacement by the protein. The protein was equilibrated in 12.5 % D_2O buffer prior to resuspension.

2.19. Circular Dichroism

Proteins were dialysed into 20 mM NaP pH 7.5, 0.5 M NaF, 5% glycerol v/v, 1 mM EDTA, 1 mM TCEP prior to CD measurements using PD10 columns (GE Healthcare) and the concentration quantified using Nanodrop.

Circular Dichroism experiments were performed using the Chirascan spectrophotometer at B23 Synchrotron Radiation CD Beamline at Diamond Light Source. A wavelength range of 185-260 nm and a temperature range of 20-89°C in 3°C increments with 60 s incubation were used, with 1 nm wavelength increment, 1 nm slit and 1 s interval time. Cuvettes with pathlengths of 0.01 or 0.02 cm were used in a 4-cell turret sample environment. CD spectra processing was done using CDApps (Siligardi et al., 2002). Unfolding transition temperature was estimated using a Boltzmann two state exchange model and secondary structure content using a CONTINLL algorithm.

For the IQS study, IQS analogues were dissolved in EtOH and added to the protein (50 μM final concentration).

ATP and its non-hydrolysable analogue were added in a final concentration of 1 mM, as well as $MgSO_4$ (also 1 mM).

2.20. Landau free energy expansion analysis

Landau free energy expansion analysis (Bonev and Morrow, 1997) of thermolytic degradation of proteins and protein complexes was done in Excel following the procedure described previously (Hussain et al., 2018). Secondary structure content obtained from CD as a function of temperature was used to define a normalized thermodynamic order parameter s [0,1] related to the helical

protein content (*cf*. (Hussain et al., 2018)). The free energy expressed in a Landau series expansion in terms of *s* is:

(2.1)

$$G = G_0\left[\frac{s^4}{4} + \alpha(T - T_c)\frac{s^2}{2} + \beta(T - T_m)s\right]$$

Minimising *G(s)* in stable systems yields the coordinates of the positive spinodal s_+,T_+:

(2.2)

$$T = T_+ + \frac{T_m - T_+}{2}\left[3\left(\frac{s - s_+}{s_+}\right)^2 + \left(\frac{s - s_+}{s_+}\right)^3\right]$$

Further, linear fitting of T *vs*. the quantity:

(2.3)

$$f(s) = 3\left(\frac{s - s_+}{s_+}\right)^2 + \left(\frac{s - s_+}{s_+}\right)^3$$

to the temperature dependence of *s(T)* allows us to obtain the equation of state for the protein system in the folded state. The normalised spinodal temperature $\tau = 1 - T_m/T_+$ is directly related the absolute strength of the first order phase transition where fully cooperative transition takes value of 0.5 and a critical transition with $T_m = T_+$ is characterized by τ of 0.

2.21. Small Angle X-ray Scattering (SAXS)

2.21.1. Room temperature measurements

SAXS measurements were performed at Diamond Light Source (Harwell Campus, UK) on the BioSAXS B21 beamline (Cowieson et al., 2020) using a temperature regulated robotic sample environment (Arinax, Grenoble, France). Samples were stored in a 96-well plate at a temperature set to 20°C prior to loading. Buffer solution and protein samples were exposed to the beam in a 1.6 mm quartz capillary washed between each measurement following a detergent wash/rinse/dry cycle. The sample environment was set to maintain the exposure capillary at a constant temperature of 27 °C (IQS study) and 30°C (AgrA-AgrC study). The beamline was configured with 2.5 m sample-detector distance with an Eiger 4M detector to cover a q range between 0.0034 and 0.42 $Å^{-1}$ (IQS study) and 3.6 m distance giving a q range of 0.0045-0.344 $Å^{-1}$ (AgrA-AgrC study).

AgrA and $AgrC_{206}$ samples were dialysed into 20 mM NaP pH 7.5, 0.5 M NaCl, 5% glycerol v/v, 1 mM EDTA, 5 mM DTT for measurements.

AgrA samples for IQS analogues study were dialysed into 2 mM NaP pH 7.5, 0.5 M NaCl, 5% glycerol v/v, 1 mM EDTA, 5 mM DTT for measurements.

ATP and its non-hydrolysable analogue were added in a final concentration of 1 mM, as well as $MgSO_4$ (also 1 mM).

Data were reduced to 1-D curves using DAWN (Filik et al., 2017, Basham et al., 2015) and background subtraction and initial data processing were performed using using DAWN and ScÅtter IV (Rob Rambo, Diamond Light Source, Harwell, http://www.bioisis.net/tutorial/9).

2.21.2. Temperature scans

See 2.15.1. The sample environment was set to collect at 20 °C and between 27 and 57°C in 3 °C increments.

Samples were dialysed into 2 mM NaP pH 7.5, 0.5 M NaCl, 5% glycerol v/v, 1 mM EDTA, 5 mM DTT for measurements.

Data were reduced to 1-D curves using DAWN (Filik et al., 2017, Basham et al., 2015) and background subtraction and initial data processing were performed using using DAWN and ScÅtter IV (Rob Rambo, Diamond Light Source, Harwell, http://www.bioisis.net/tutorial/9).

2.22. Small Angle Neutron Scattering (SANS)

2.22.1. Protein SANS

Protein samples were dialysed into 20 mM NaPi pH 7.5, 0.5 M NaCl, 5% glycerol v/v, 1 mM EDTA and 5 mM DTT. Three equivalent buffers were made: in 100% H_2O, 42% D_2O and 100% D_2O.

SANS measurements were carried out at the ZOOM beam line of the ISIS Neutron and Muon Source at the Rutherford Appleton Laboratory, Didcot, UK, using a sample changer with Julabo water bath for temperature control, and 1 mm path length quartz cuvettes. The beamline was configured with L1 = L2 = 4 m collimation and sample-detector distances to give a scattering vector $Q = (4\pi/\lambda)\sin(\theta/2)$ range of 0.004 to 0.8 Å-1, where θ is the scattering angle and neutrons of wavelengths λ of 1.75 to 16.5 Å were used simultaneously by time of flight. Samples were prepared in 100% D2O, 100% H2O and 42% D2O to make visible predominantly protonated, protonated and deuterated, and

deuterated protein samples respectively. Scattering measurements were collected for 3 hours. Data reduction was performed using Mantid.

SANS data were plotted and processed using ScÅtter IV (Rambo, Diamond Light Source, 2020, http://www.bioisis.net/tutorial/9). Selected data are displayed in the form of *log[I(q)] : q*, *ln[I(q)] : ln[q]* and *I(q)/I(0)*(q*Rg)2* plots to emphasise the quality of the data subtraction, the level of aggregation and the folding of the samples, respectively.

2.22.2. Magnet SANS

See 2.18 for bicelle preparation protocol.

SANS measurements were carried out at the ZOOM beam line of the ISIS Neutron and Muon Source at the Rutherford Appleton Laboratory, Didcot, UK, using a specifically designed magnet-sample holder set-up. The applied horizontal magnetic field was ~1.5 T and a 1 mm path length banjo-type quartz cell was used. Temperature control was linked to the magnet and provided either room temperature or ~45 °C. The beamline was configured with L1 = L2 = 8 m collimation and sample-detector distances to give a scattering vector Q = $(4\pi/\lambda)\sin(\theta/2)$ range of 0.0025 to 0.5 $Å^{-1}$, where θ is the scattering angle and neutrons of wavelengths λ of 1.75 to 15 Å were used simultaneously by time of flight.

2.23. Analysis of scattering data

2.23.1. *Ab initio* shape determination from SAXS data

Multiphase ab initio shape determination from experimental SAXS data was performed using the MONSA online service available as part of the European Molecular Biology Laboratory ATSAS web services (single repeat) (ATSAS 3.1.3 (r14636) Copyright (c) ATSAS Team, EMBL, Hamburg Outstation 1998-2022). Algorithms used are described in (Svergun, 1999, Svergun and Nierhaus, 2000). The system studied consists of two phases, AgrA and $AgrC_{206}$. The overall symmetry was set to P1 (no symmetry) and the search volume to sphere. The estimated $D_{max}/2$ for the AgrA/$AgrC_{206}$ complex was determined to be 48 Å. Phase 1 volume was defined to be 30289 and phase 2 was 27966, which are respectively AgrA and $AgrC_{206}$ molecular weights. The experimental curves correspond to full-length AgrA, $AgrC_{206}$ and AgrA/$AgrC_{206}$ at 30°C.

2.23.2. Scattering pattern prediction from atomic resolution structures

X-ray scattering curves were simulated from known atomic resolution structures of full-length AgrA and $AgrC_{206}$ and compared to experimental data using CRYSOL, part of the ATSAS package developed by the EMBL (Svergun et al., 1995). Experimental data were used to q_{max} of 0.2$Å^{-1}$. Default options were chosen throughout as detailed below:

- Program options: evaluate scattering amplitudes and envelope
- Maximum order of harmonics: 15
- Order of Fibonacci grid: 17
- Maximum *s* value: 0.2 (not default)

- Number of points: 21
- Account for explicit hydrogens: No
- Fit the experimental curve: Yes
- Subtract constant: No
- Angular units in the input file: 1
- Electron density of the solvent: 0.3340
- Plot the fit: Yes

2.23.3. Ensemble Optimisation Method

The frames generated from the MD simulations were used as inputs for the ensemble optimisation using the GAJOE (Tria et al., 2015, Bernadó et al., 2007) program, which allows the selection of an ensemble of models which best explain the small-angle scattering curve. This allows the characterisation of the solution state of protein molecules containing highly flexible or disordered regions, or have multiple conformations. The scattering curve of each frame is computed separately and a genetic algorithm is used to select an appropriate ensemble whose linear combination gives the best fit to the experimental SAXS curve. Four hundred and 100 frames were processed for $AgrC_{206}$ and C199S AgrA, respectively. For the experimental scattering data, see 2.20.1.

2.23.4. SasView

SasView 5.0.5 (Doucet et al., 2021)was used to determine the Rg used in the Landau series expansion analysis. The Guinier-Porod model (Hammouda, 2010b, Hammouda, 2010a) with the Levenberg-Marquardt fitting algorithm. The data fitting for the data obtained at the DL-SAXS beamline was done using the fitting algorithm DREAM.

The Rg determination of AgrA and the IQS analogues was done using the Guinier model (Guinier and Fournet, 1955) and the Levenberg-Marquardt algorithm.

The analysis of scattering profiles of membrane proteins was performed using SasView 5.0.4 ((Doucet et al., 2021) using a core shell ellipsoid model (Kotlarchyk and Chen, 1983, Berr, 1987). The scattering length densities (SLD) were fixed at $12.3x10^{-6}/Å^2$ (core – protein), $8.3x10^{-6}/Å^2$ (shell – detergent) and $9.44x10^{-6}/Å^2$ (solvent).

3. Expression, purification, activity and deuteration of AgrA and AgrC

Characterising the structural features of the *agr* proteins is pivotal to understanding their function and, in turn, being able to design strategies to prevent or inhibit infection. In recent years biophysical techniques such as circular dichroism, NMR and scattering are increasingly being recognised as a complementary alternative to crystallography as they provide structural insights into proteins in an environment more closely resembling their native conditions and subject to ambient thermal dynamics.

In order to biophysically characterise the two-component system of the *agr* quorum sensing, recombinant full-length and C-terminal AgrA, and $AgrC_{206}$ were expressed and purified. Complex formation between full-length AgrA and $AgrC_{206}$ was evaluated by Size Exclusion Chromatography.

3.1.Full-length AgrA

The investigation of *S. aureus* quorum sensing two-component system started with the only soluble protein in this system, AgrA. This is the response regulator in the system and interacts with the promoters P2 and P3 following phosphorylation by AgrC and dimerisation. The recombinant AgrA that was used carries a point mutation, C199S, aimed at reducing cross-linking and aiding solubility.

3.1.1. Expression

A protocol for purification of the recombinant protein expressed in *E. coli* BL21 was already optimised (previous unpublished work by Dr Phil Bardelang) involving the use of high concentrations (1 M) of sorbitol in order for this protein

to express in a soluble manner. One of the main problems to face when expressing recombinant proteins is the formation of inclusion bodies resulting from the high concentrations of the recombinant protein itself, which is therefore often found in the insoluble fraction of lysate. Sorbitol has been previously reported among the osmolytes able to stabilise the native folding of recombinant proteins, both *in vitro* (Xie and Timasheff, 1997, Filatova et al., 2010) and *in vivo* (Prasad et al., 2011, Blackwell and Horgan, 1991).

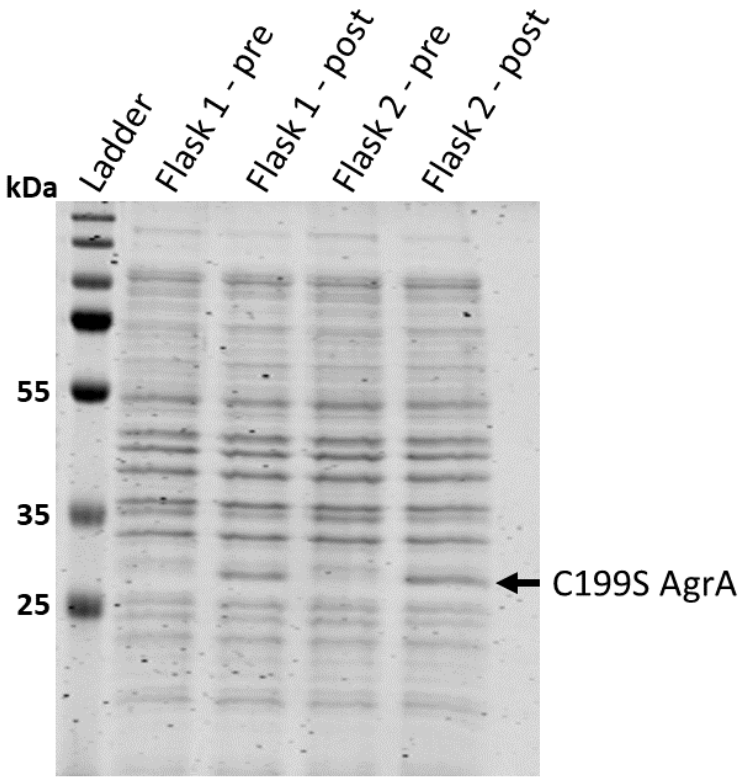


Figure 3.1.1 ***Expression of C199S AgrA.*** *C199S AgrA was expressed in* E. coli *BL21 in terrific broth with 1 M sorbitol. The 15% SDS PAGE gel shows samples before and after induction.*

The expression of full-length AgrA in presence of sorbitol is shown in the SDS PAGE gels in Figure 3.1.1. Overexpression of recombinant AgrA in absence of sorbitol results in mostly insoluble protein, but the downside to the use of

sorbitol is a reduction in the cell growth, or a slower growth than in absence of sorbitol (see Table 3.1.1).

Table 3.1.1 Growth of *E. Coli* BL21 (transformed with pCOLD I-C199S AgrA) in presence and absence of sorbitol (OD) (duplicates).

Time	+ sorbitol	Time	- sorbitol
0 h	Inoculation	0 h	Inoculation
3 h 35 min	0.114, 0.119	2 h	0.122, 0.116
6 h 5 min	0.428, 0.436	3 h	0.7, 0.614
7 h 10 min	0.508, 0.516	3 h 20 min	0.816, 0.72
7 h 40 min	0.560, 0.547	-	-
7 h 55 min	0.632, 0.622	-	-

The presence of 1 M D-sorbitol in the terrific broth slows the cell growth by approximately 4 and a half hours, to reach the correct OD for induction (0.6 – 0.9).

3.1.2. Purification

The purification of AgrA was carried out following the protocol described in the Methods, which had previously been optimised (previous unpublished work by Dr Phil Bardelang).

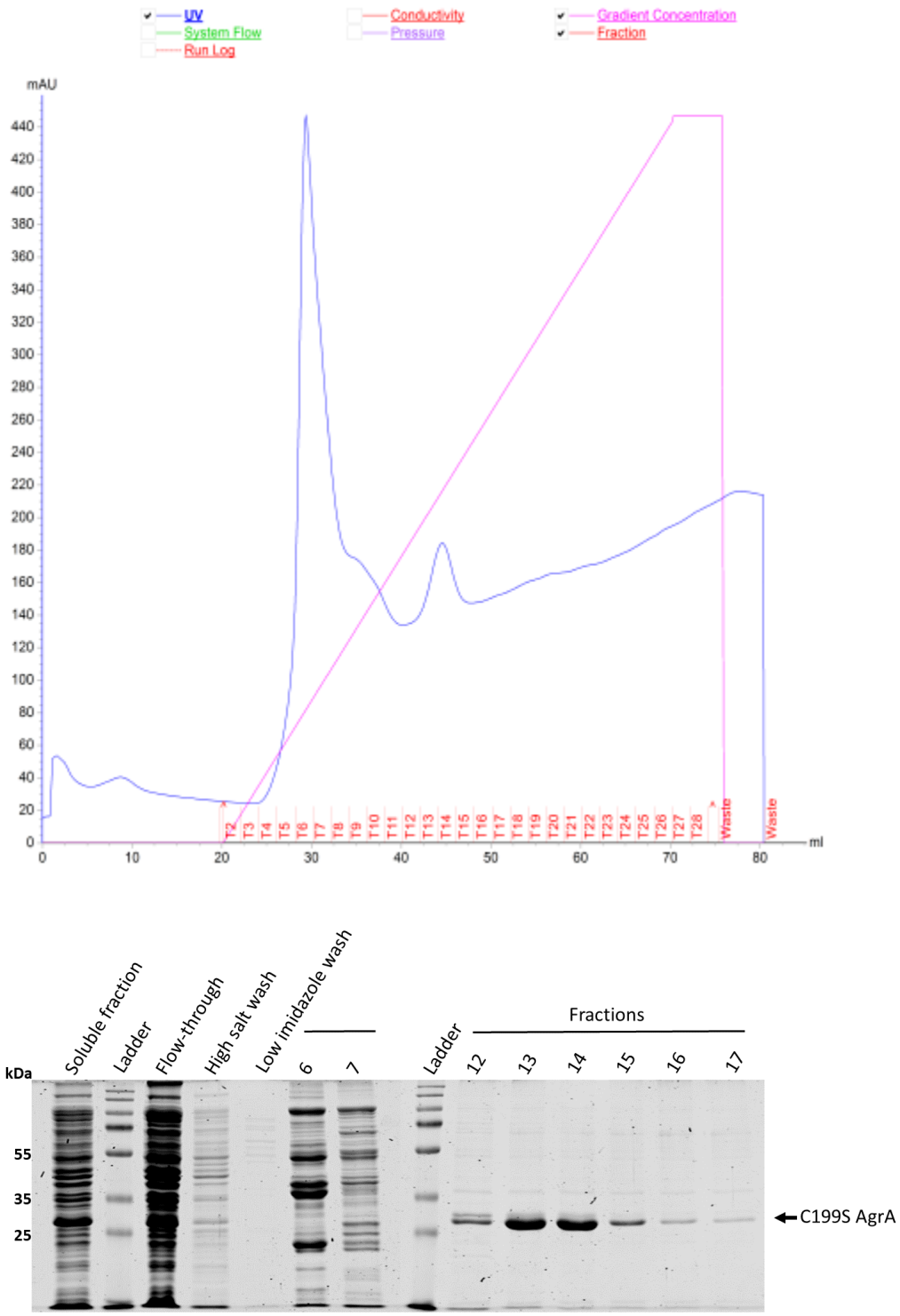


Figure 3.1.2 ***Purification of full-length AgrA.*** *UV trace from FPLC elution (top) and 15% SDS PAGE gels showing the fractions eluted by FPLC. AgrA was eluted in fractions 13 and 17.*

As shown in Figure 3.1.2, recombinant AgrA was eluted in approximately 45-50% high imidazole buffer (high imidazole buffer contains 0.5 M imidazole). The purity of the eluted sample was estimated at approximately 92-95% and this, together with the following dialysis step which ensured the reduction of contaminants, means the sample is judged of sufficient purity for low-resolution biophysical investigation.

3.1.3. Activity assay

To establish the native conformation adopted by the recombinant protein in relation to its function, its activity was assessed by Electrophoretic Mobility Shift Assay (EMSA). The principle behind the choice of this assay is the DNA-binding ability of AgrA: the progression of the fluorescently tagged, negatively charged P2 probe towards the positive electrode is monitored with and without protein. If recombinant AgrA is folded correctly and active, it will selectively recognise and bind P2, therefore increasing the size of the complex and slowing the shifting of the probe towards the positive electrode.

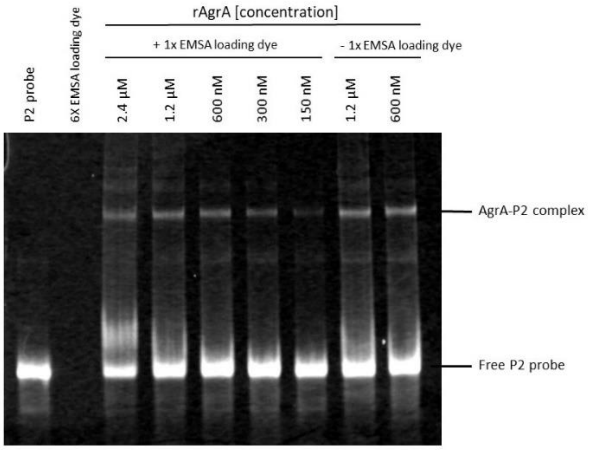


Figure 3.1.3 ***Activity assay of full-length recombinant C199S AgrA.*** *Electrophoretic mobility shift assay showing the shift in the migration of fluorescent P2 probe upon interaction with recombinant C119S AgrA.*

Figure 3.1.3 shows that the recombinant C119S AgrA is active and therefore assumed to be in its native conformation, as the shift in the P2 probe shows. The best shift is obtained at protein concentrations between 600 nM and 1.2 µM.

3.2.AgrC

AgrC has not yet been structurally characterised, due to the difficulties of working with membrane proteins in terms of their purification and stability. Previous efforts have been directed towards exploring the activity of AgrC by creating chimeric systems (Xie et al., 2019), focusing on the phosphorylation process (Srivastava et al., 2014) and exploring the effect of point mutations on AgrC activity (Wang et al., 2017a, Sloan et al., 2019). The cytoplasmic (soluble) domain of AgrC was recombinantly expressed and purified starting from residue 206 onwards, which comprises of the catalytic domain and the phosphotransfer

and dimerisation domain. The transmembrane sensory domain was excluded as the investigation focused on the interaction of AgrC with AgrA.

3.2.1. Cloning

Cloning of the cytoplasmic domain of AgrC was carried out using the NEB HiDi DNA Assembly kit. The *agrC* gene was incorporated into a pCOLD-I vector and the reaction verified by colony PCR and sequencing of the end product (see Methods 2.7.1 for more details of the protocol). The work was carried out with the help of Dr Ewan Murray.

3.2.2. Expression

$AgrC_{206}$ was cloned into the same expression vector that was used for the AgrA protein, pCOLD-I, therefore the same expression protocol was used as described in the Methods.

The expression of $AgrC_{206}$ was successful as shown in Figure 3.4.1.

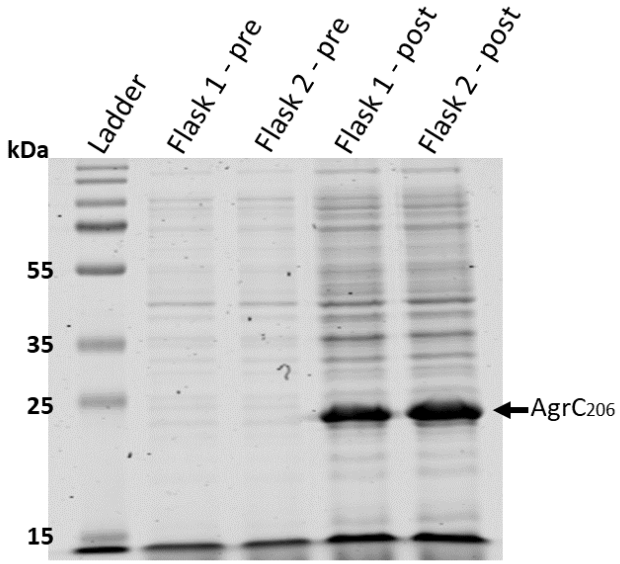


Figure 3.4.1 ***Expression of AgrC$_{206}$.*** *Samples pre- and post-induction are shown on a 15% SDS PAGE gel from an expression in duplicate. AgrC$_{206}$ is approximately 28 kDa but runs just below the 25 kDa mark.*

3.2.3. Purification

The initial purification protocol was the same used for AgrA. This had to be slightly adapted to use a protease inhibitor cocktail instead of PMSF as the purified product was broken down into fragments. This may explain the fragmentation observed by Srivastava *et al.* (Srivastava et al., 2014).

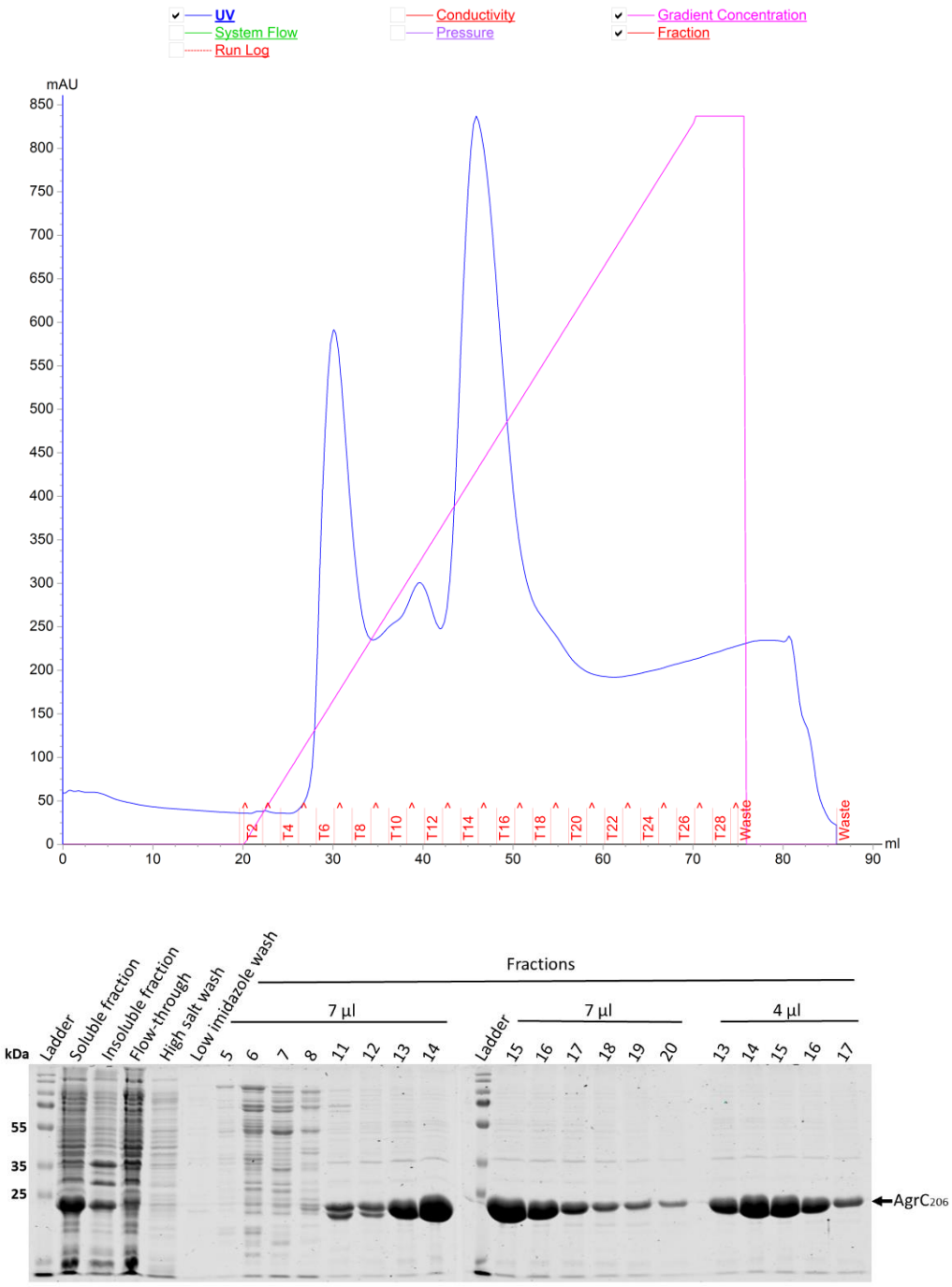


Figure 3.4.2 ***Purification of AgrC$_{206}$.*** *UV trace from FPLC elution (top) and 15% SDS PAGE gels showing the fractions eluted by FPLC. AgrC was eluted in fractions 13 to 20.*

The amount of purified protein vastly exceeded what was obtained from AgrA purifications, with up to 32 mg of protein purified from a 500 ml culture.

3.3. Size exclusion chromatography for AgrA/C complex

AgrA and AgrC are expected to interact with each other for AgrC to carry out its role in the phosphorylation and subsequent dimerisation of AgrA. We investigated whether this interaction was stable enough to be observed using Size Exclusion Chromatography.

The two proteins were mixed together in 1:1 molar ratio and run down a S200 SEC column. This experiment was repeated in presence of ATP and non-hydrolysable ATP to determine any possible stabilising effect the nucleotides may provide to the complex formed. This study complements SAXS and CD data presented in Chapter 4 and 5, respectively. A graph describing the possible interaction mechanisms can be found in 5.4.

The first SEC run, on the two proteins in absence of nucleotides, was carried out by Dr Phil Bardelang using an AktaPure, while the experiments with both nucleotides were carried out by me using an AktaStart.

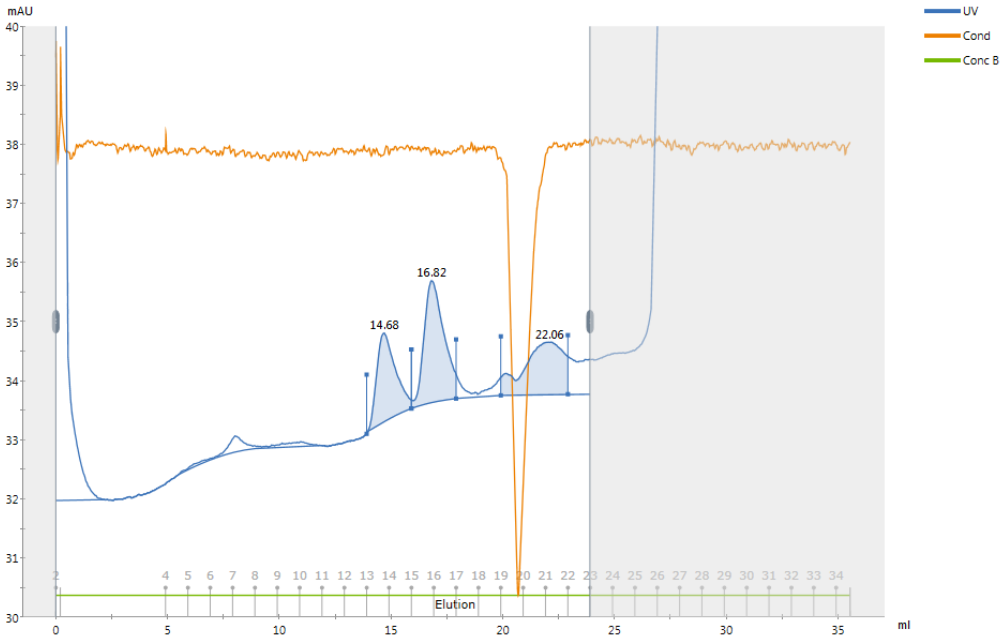


Figure 3.5.1 ***Size exclusion chromatography of AgrA+AgrC$_{206}$.*** *UV trace showing the elution of the AgrA and AgrC$_{206}$ mixture on an S200 column. The first peak around fraction 14 shows the elution of the dimeric AgrC$_{206}$ (see Chapter 4 for information on the dimer) and the second (fractions 15-16) of C199S AgrA. There is no peak corresponding to the earlier fractions which would represent a complex.*

The SEC experiment on the mixture alone does not evidence formation of a complex (Fig. 3.5.1). The two proteins eluted separately, with the dimeric AgrC$_{206}$ eluting first.

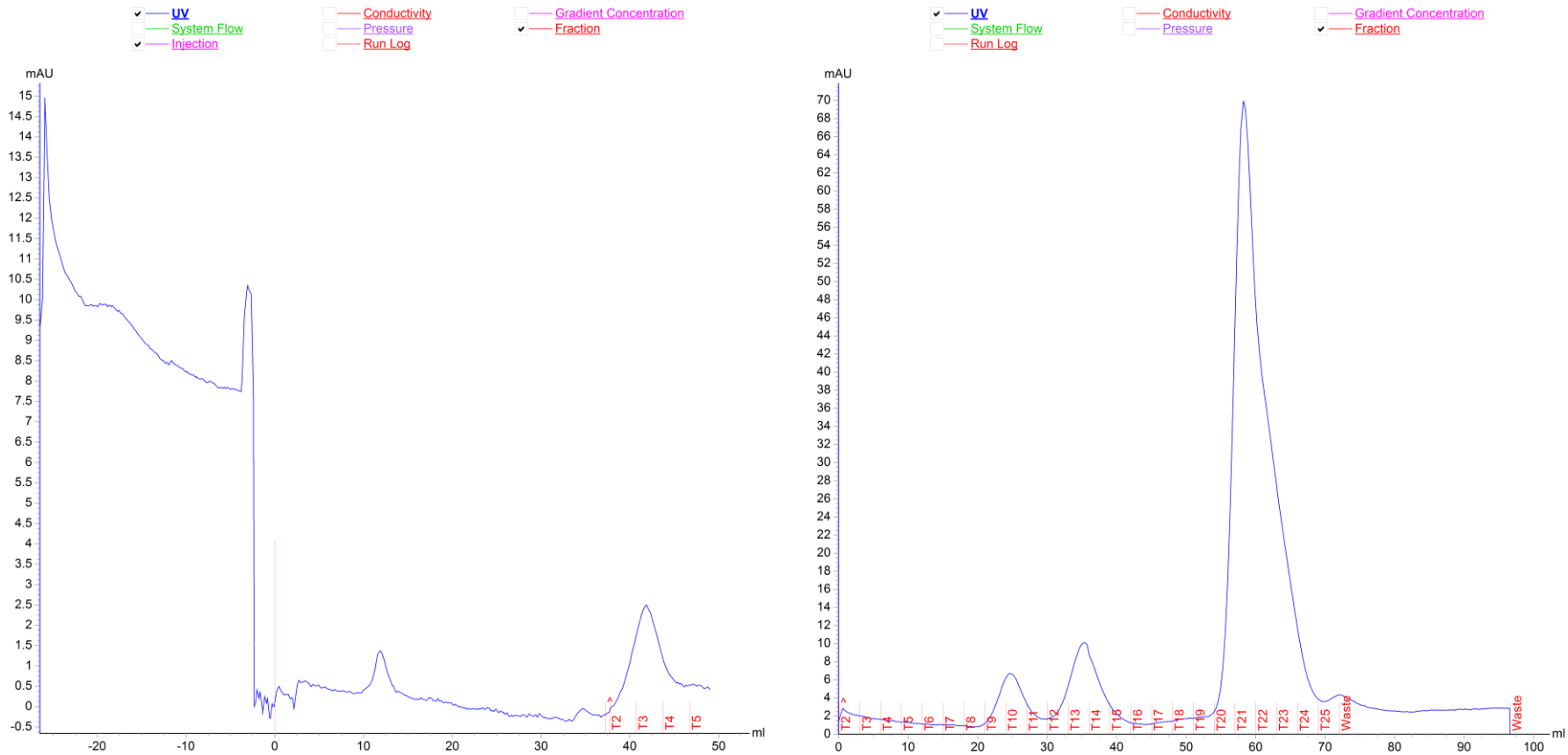


Figure 3.5.2 ***Size exclusion chromatography of AgrA+AgrC$_{206}$+ATP.*** *UV trace showing the elution of the AgrA and AgrC$_{206}$ mixture in presence of ATP on an S200 column. The SEC run is shown in two separate graphs due to equipment issues.*

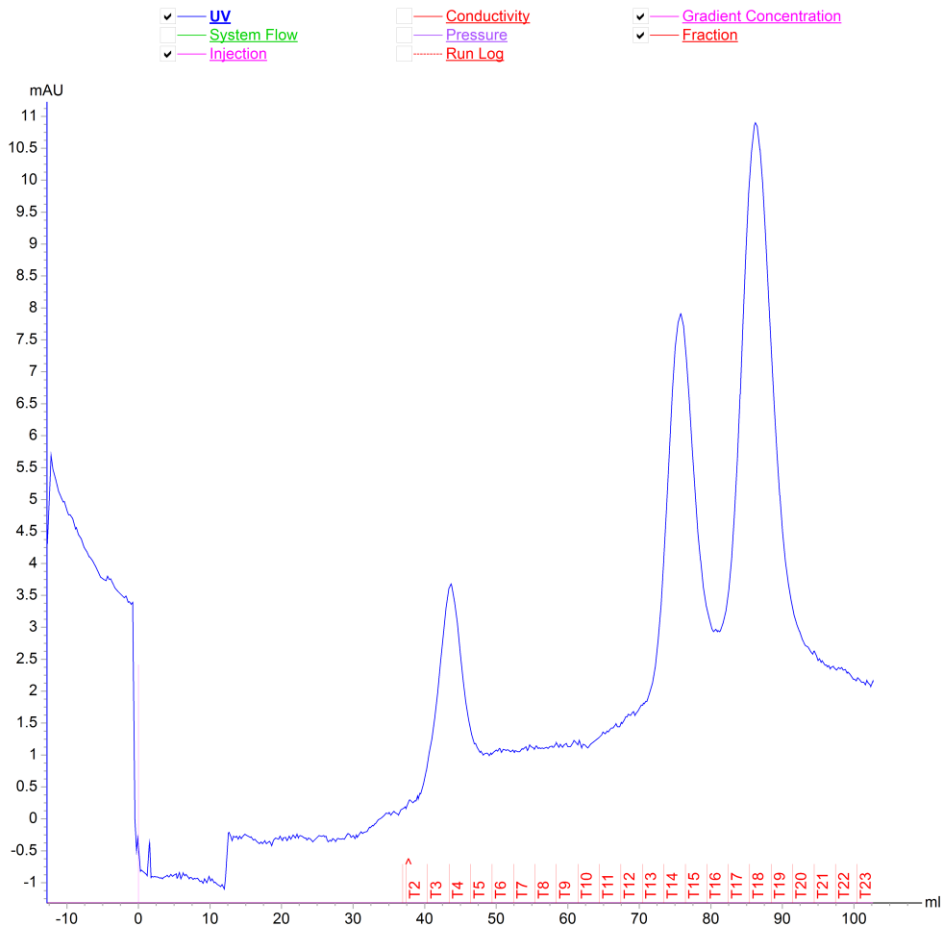


Figure 3.5.3 ***Size exclusion chromatography of AgrA+AgrC$_{206}$+sATP.*** *UV trace showing the elution of the AgrA and AgrC$_{206}$ mixture in presence of sATP on an S200 column.*

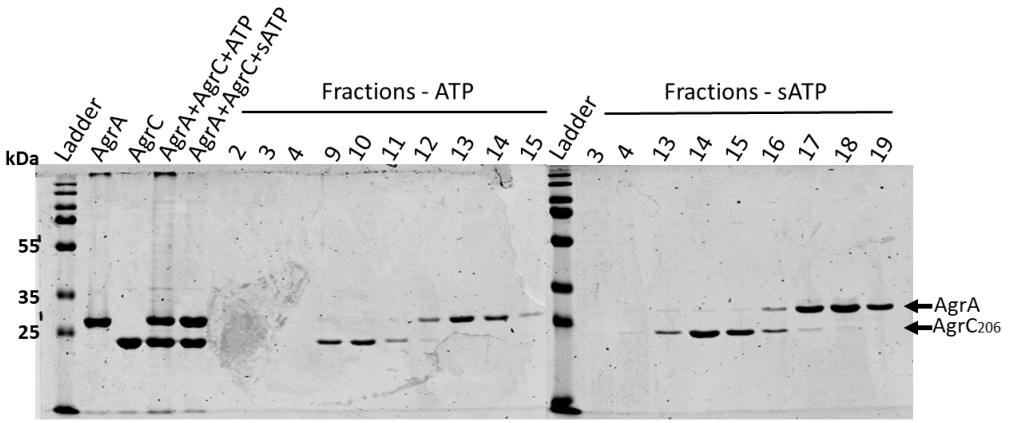


Figure 3.5.4 ***SEC fractions of AgrC+AgrA+nucleotides.*** *15% SDS PAGE gels showing the fractions eluted from an S200 column from two runs: the mixture of the two proteins with ATP and with sATP. None of the eluted fractions show the presence of a complex.*

In all cases the UV traces and SDS PAGE gels reveal that AgrA and $AgrC_{206}$ run as two separate systems. $AgrC_{206}$ elutes first in its dimeric form at 56 kDa, followed by AgrA at 30 kDa (the FPLC trace in Fig. 3.5.2 is broken into two due to equipment malfunction).

This is a sensory response that requires only a transient association of the two proteins. Weak or transient protein-protein interactions are notoriously difficult to isolate (Qin and Gronenborn, 2014). The mechanical forces involved in the passage of the samples through the SEC column may explain the absence of a detectable complex, especially in the case of a transient interaction.

The molecular interaction between AgrA and AgrC is further investigated in Chapters 4 and 5 using SAXS, SANS and CD.

3.4.Deuteration of $AgrC_{206}$

To gain insights into possible structural rearrangements in AgrC and AgrA upon interaction with each other, contrast-matching was used to carry out a SANS investigation to complement SAXS data. Protein deuteration is essential to exploit the potential of contrast matching in SANS. While the scattering profiles of the two-component system proteins alone and in complex were explored by SAXS, it is not possible to discriminate structural effects of the interaction on the two individual proteins. Contrast-matching allows to selectively deuterate one of the proteins, protonate the other and create three contrasts with buffer to mask the proteins in turn.

The need for sorbitol at high concentrations, coupled with minimal media, D_2O and pyruvate needed for the deuteration, creates very unfavourable conditions and prevents growth of *E. coli* BL21. AgrA could therefore not be deuterated.

$AgrC_{206}$ on the other hand is a very stable protein that is expressed with very high yields (see Fig. 3.4.2) without the need for sorbitol and is therefore a good candidate for deuteration. $AgrC_{206}$ expressed in deuterated M9 minimal medium is shown in Figure 3.6.1.

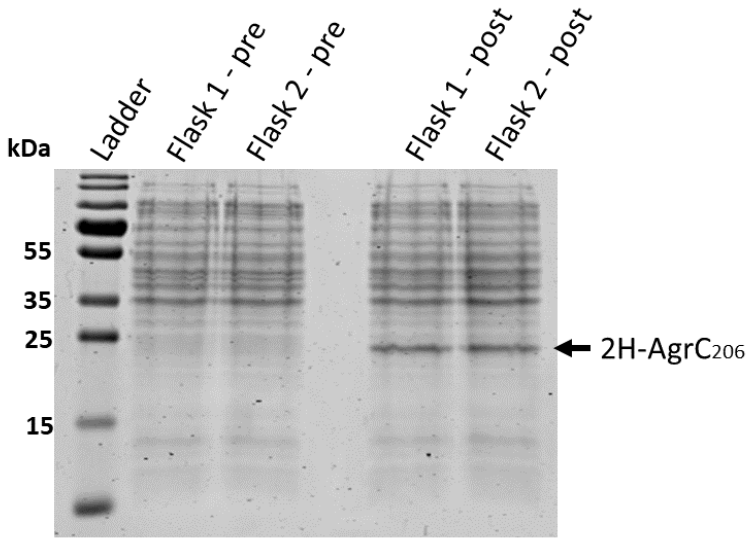


Figure 3.6.1 ***Expression of AgrC$_{206}$ in minimal medium for deuteration.*** *Samples pre- and post-induction are shown on a 15% SDS PAGE gel from an expression in duplicate.*

Following the successful expression of ^{2}H-AgrC$_{206}$, purification following the standard protocol could be carried out (Fig. 3.6.2)

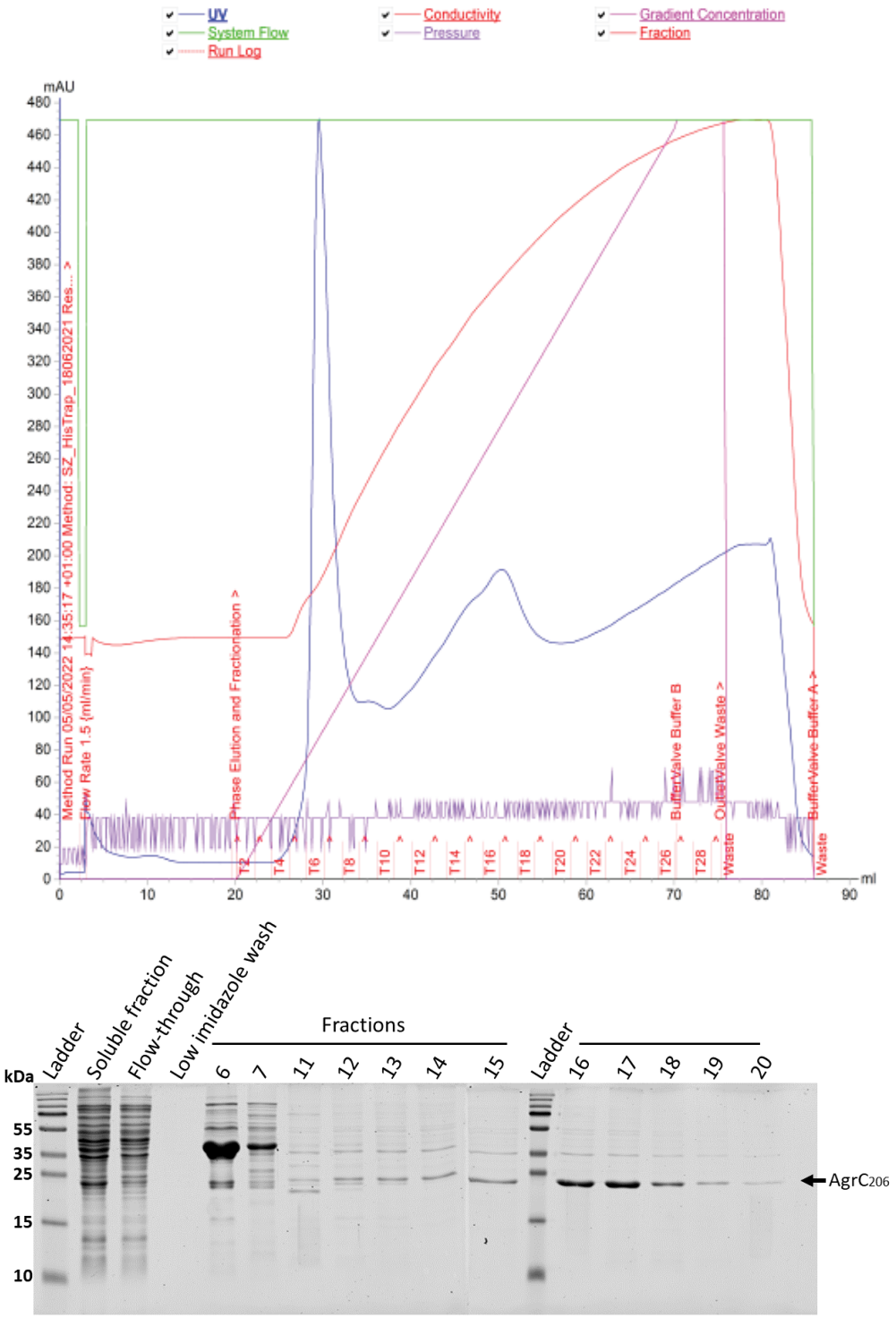


Figure 3.6.2 ***Purification of ^{2}H-AgrC$_{206}$.*** *UV trace from FPLC elution (top) and 15%SDS PAGE gels (bottom) showing samples from various stages of purification of the deuterated protein. ^{2}H-AgrC$_{206}$ was eluted in fractions 16-20.*

The purification of these proteins was successful and sample for further biophysical investigations were obtained. Activity assays of AgrA confirmed its activity and therefore the correct folding of this protein is assumed. Although the formation of a stable complex by AgrA and $AgrC_{206}$ was not detected by SEC, this may be due to a transient interaction which is not retained during the column run. These results are to be considered in relation to the findings described in the next two chapters.

4. Investigation of AgrA and AgrC protein structures and molecular interactions at room temperature

The molecular structure of proteins has always been central to the understanding of biological processes, and this knowledge allows us to take advantage of favourable mechanisms or interfere with them if they negatively impact us. Structure is the key to function, which means we cannot fully comprehend how a protein works until its structure is fully resolved. For example, being able to describe protein structures at atomic level has enabled the design of drugs able to stimulate or inhibit the action of a particular protein, therefore altering a whole pathway of reactions.

Historically, the most common approach to structure determination has been protein crystallography. Crystallography involves the creation of a crystal containing ultra-pure and concentrated sample, which will scatter x-rays. This allows for the creation of atomic-level models based on the scattering pattern obtained. The major criticism that crystallography faces is that the proteins are in conditions far from native, which means the structure obtained may not be a true representation of their native conformation due to the different environment and physical constraints.

Another possible approach is Nuclear Magnetic Resonance. While this technique also reaches atomic-level precision, the preparation of samples can be laborious and expensive. The concentration of the protein needs to be decent, which rules out proteins that have a tendency to aggregate or precipitate, and the labelling process can be expensive and have low yield.

Small angle scattering (SAS) is a low-resolution technique which gives the user comparatively good flexibility in terms of the buffer used and does not require the same level of purity needed for crystallography. SAS also allows for the sample to be at room temperature, in solution and in optimal buffer for stability. SAS data analysis software has improved significantly in recent years and, although at low resolution, allows to detect changes in conformation, formation of complexes and is informative on protein tertiary structure.

One immediate way to detect strong protein-protein complexes in the lab is Size Exclusion Chromatography. Chapter 3 described the results of this type of experiment on the mixture of the *agr* two-component system proteins, AgrA and $AgrC_{206}$. The two proteins eluted separately from the column, which suggests either no complex formation, or a weak or transient interaction. The same was true when ATP was added. This chapter describes how biophysical techniques, namely Small Angle Scattering and Circular Dichroism, were implemented together to characterise individual proteins of the *agr* system and their interactions.

4.1. Structural characterisation of $AgrC_{206}$ and AgrA

AgrC is the transmembrane histidine kinase responsible for receiving the quorum-sensing extracellular signal and transducing it into the cell. It comprises of three domains: a transmembrane domain, dimerisation and phosphotransfer domain and a catalytic domain; the transmembrane domain is excluded from this study as the molecular interaction between AgrA and the response regulator AgrA occurs in the soluble, cytoplasmic region of the protein. Structural

characterisation that is already available include two crystal structures, one of the catalytic domain (Srivastava et al., 2014) and the other of the chimeric GCN4-AgrC construct (Xie et al., 2019), both described in the introduction. The aim was to characterise the behaviour of the protein in close to native conditions, accounting for its flexibility.

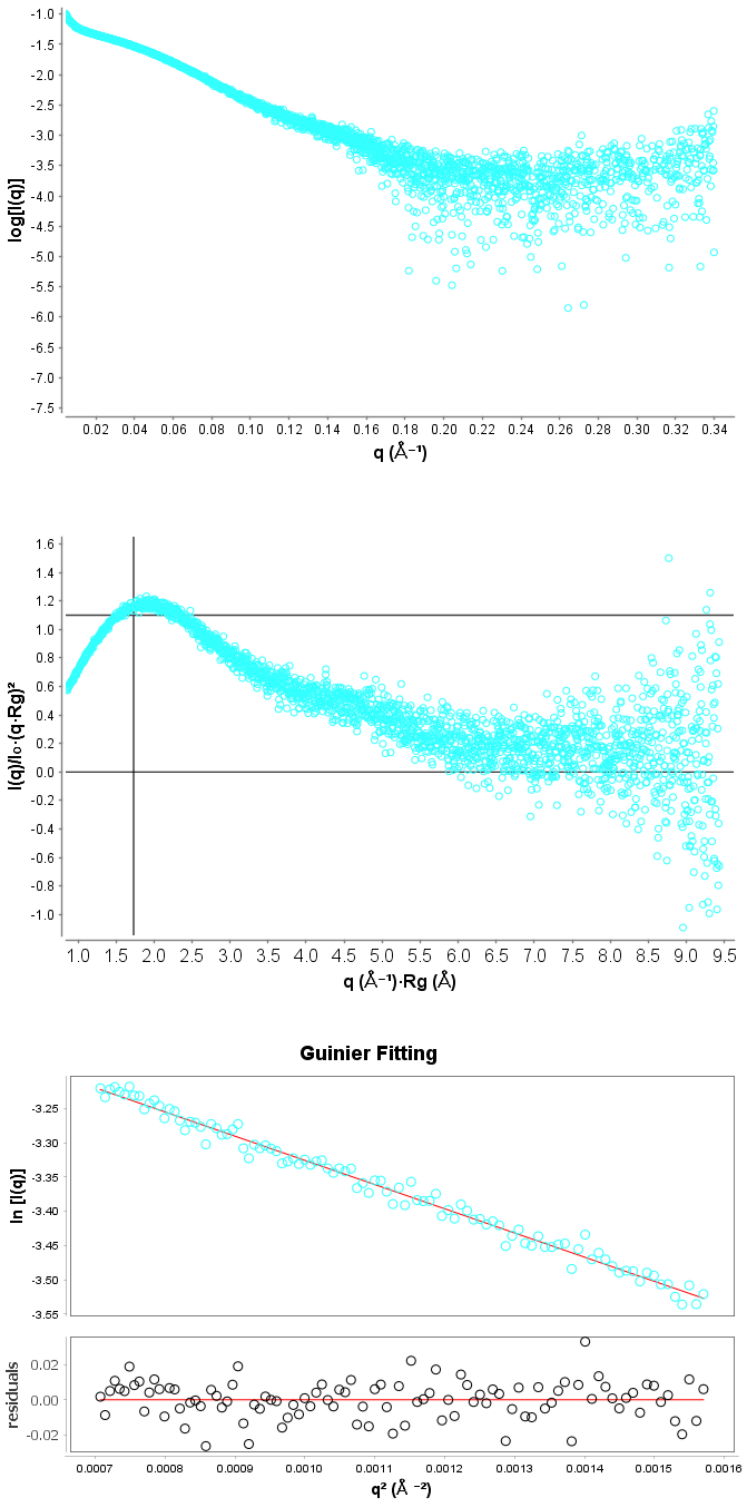


Figure 4.1.1 ***Experimental SAXS data of AgrC$_{206}$.*** *Logarithmic scale (top), dimensionless Kratky (middle) and Guinier analysis (bottom) plots of AgrC$_{206}$. The normalised Kratky shows a well folded protein. The $q * R_g$ limits of the Guinier analysis are 0.8557 and 1.2827.*

The quality of the data is good, with not excessive levels of noise and a comparatively small amount of aggregation visible at low *q* on the logarithmic plot (Fig. 4.1.1). The Kratky plot is indicative of the foldedness and globularity of a protein, where a folded or globular protein will exhibit a bell-shaped Kratky plot. $AgrC_{206}$ shows a normalised Kratky (Fig. 4.1.1) which is consistent with a well-folded protein.

While the Kratky plots give a qualitative account of the characteristics of the protein, the flexibility analysis allows to take a more quantitative approach to the determination of particle volume.

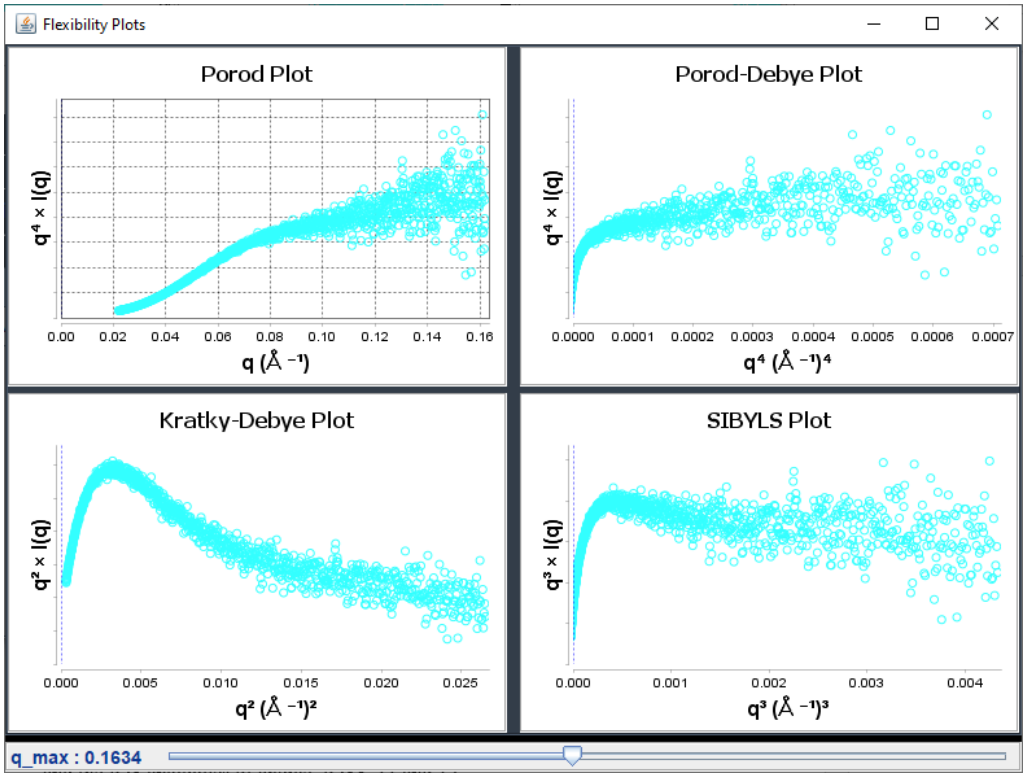


Figure 4.1.2 ***Flexibility analysis of $AgrC_{206}$.*** *The plateau in the Porod-Debye plot indicates a well-assembled, rigid protein.* q_{max} *was calculated at 0.1634.*

The Porod plateau is reached in the Porod-Debye Plot, which suggests $AgrC_{206}$ to be a fairly rigid protein. A plateau in the Kratky-Debye plot is instead characteristic of a flexible, less ordered protein (Brosey and Tainer, 2019). The determination of q_{max} ($q_{max} = 0.1634$) allows to calculate the particle volume and the Porod exponent (Rambo and Tainer, 2011).

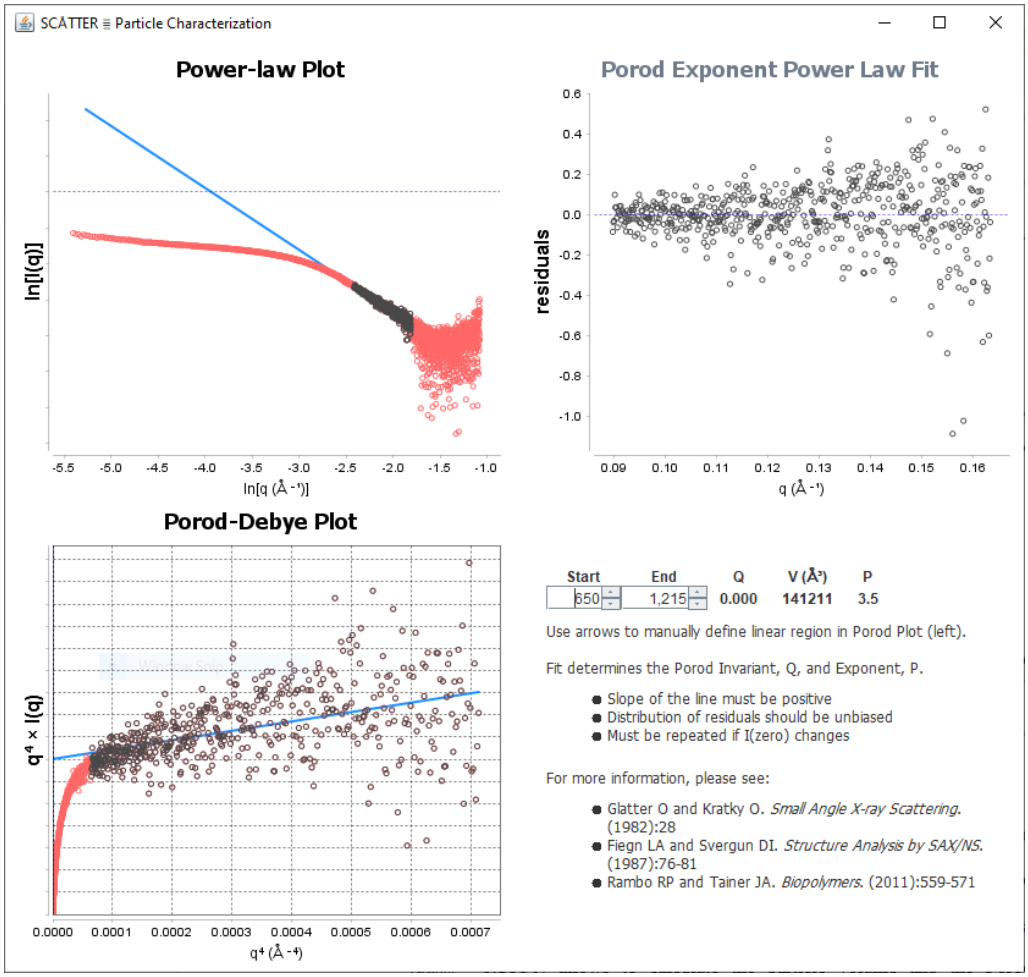


Figure 4.1.3 ***Particle characterisation of AgrC$_{206}$.*** *The linear fit in the Porod-Debye plot, with an unbiased distribution of residuals, is used to calculate the volume of the particle (=140000 Å^3) and the Porod exponent (=3.5).*

Based on the q_{max} extrapolated from the flexibility analysis, a linear fit in the Porod-Debye plot was obtained the volume of the particle $V = 140000$ Å^3 and the Porod exponent $P = 3.5$ calculated. The Porod exponent represents a

quantitative assessment of the behaviour observed in the flexibility analysis. For protein samples, it normally varies between 2 and 4 depending on the flexibility, with 2 representing flexible proteins and 4 extremely rigid systems (Brosey and Tainer, 2019). The calculated value for $AgrC_{206}$ is in agreement with what reported above and suggests a mostly rigid structure.

Using the particle volume, the density of the protein can be calculated using the equation (4.1.1)

$$\rho = \frac{10^{24} * M}{V * N_A}$$

where M is the molecular weight in Da, V the volume and N_A Avogadro's number (Quillin and Matthews, 2000). The calculated density is 0.66 $g.cm^{-3}$, which is a very low value compared to the average protein density 1.37 $g.cm^{-3}$.

This volume estimate is in accordance with the volume calculated from the radius of gyration, which is approximately 33 Å. As the radius of gyration corresponds to the size of a sphere which scatters in a similar way to the sample, the formula to calculate the volume of a sphere from its radius can be used here:

(4.1.2)

$$V = \frac{4}{3}\pi R_g^3$$

which equals 150000 $Å^3$.

This is in disagreement with the volume calculated from the MD model using UCSF Chimera, which is approximately 61000 $Å^3$, just below half what estimated from the SAXS. This would give a protein density of 1.52 $g.cm^{-3}$. These discrepancies may be due to the oblong shape of this protein, which does

not approximate well to the compact and globular shape assumed during these calculations.

While a crystal structure of both the catalytic domain and the GCN4-AgrC$_{209}$ chimeric construct are available, this study investigated the morphology of AgrC$_{206}$ at room temperature, to account for flexibility. These conditions also guarantee an unconstrained environment. This protein was also used to create a pipeline for structural analysis of soluble proteins using SAS data.

In this context structure prediction tools were used to create a homology model first and annealed simulated structures from the initial model to fit the scattering data. The modelling was carried out by Dr Boyan Bonev.

CRYSOL, part of the ATSAS suite of data analysis software, was used to evaluate the likeness of an MD-annealed model of monomeric AgrC$_{206}$.

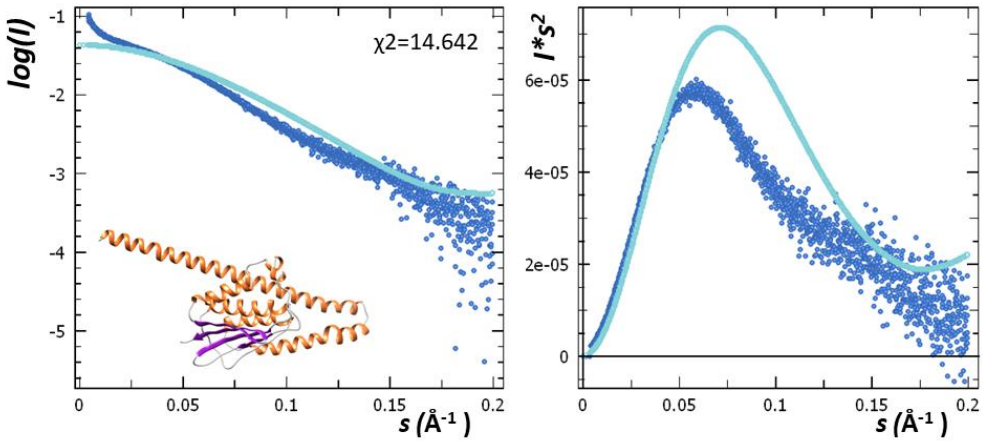


Figure 4.1.4 ***SAXS simulation of MD-annealed AgrC$_{206}$ monomer.*** *Simulated 1D X-ray scattering profile (light blue) using MD-annealed homology model of monomeric AgrC$_{206}$ is fitted to the corresponding AgrC$_{206}$ experimental data (dark blue) using CRYSOL (Svergun et al., 1995). The logarithmic plot (left) and Kratky plot (right) are shown.*

Figure 4.1.4 shows the disagreement of the simulated curve of the monomeric $AgrC_{206}$ with the experimental data. The discrepancy is more evident in the Kratky plot.

The mechanism of action of AgrC requires dimerisation. For this reason, a homology model of the dimer was constructed and annealed.

The MD simulation was run for 100 ns collecting a frame per ns. Cluster analysis revealed the frame corresponding to the dominant cluster and the scattering curve corresponding to this frame was simulated and fitted to the experimental data.

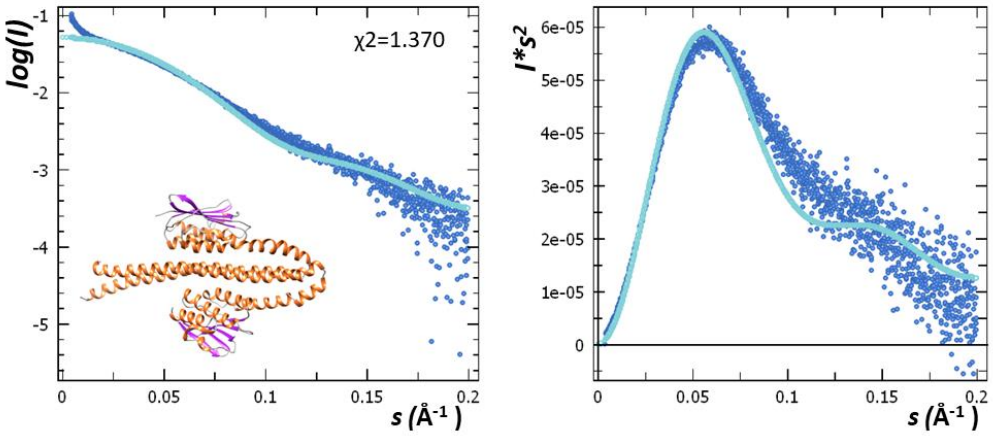


Figure 4.1.5 ***SAXS simulation of $AgrC_{206}$ dimer dominant cluster.*** *Simulated 1D X-ray scattering profile (light blue) using an MD-annealed homology model of dimeric $AgrC_{206}$ is fitted to the corresponding $AgrC_{206}$ experimental data (dark blue) using CRYSOL (Svergun et al., 1995). The logarithmic plot (left) and Kratky plot (right) are shown. The frame from the 100 ns MD simulation was selected by cluster analysis.*

The dimeric model shows an improved fit to the data (Fig. 4.1.5), especially in the Kratky plot and in the *q* region between 0.05 and 0.15. These results show

that the cytosolic fraction of AgrC is dimeric in absence of the transmembrane region.

Although the dominant cluster shows a much improved fit to the experimental data, a certain degree of rearrangement was observed in the MD simulation. In an attempt to further refine the fitting, the Ensemble Optimisation Method was implemented, more specifically GAJOE (Genetic Algorithm Judging Optimisation of Ensembles). GAJOE performs a selection of models that, combined, give the best fit to the scattering curve, indicating what percentage each of them contribute (Tria et al., 2015).

The MD simulation was extended to 400 ns, which generated 400 frames (Dr Boyan Bonev). The programme selected the representative frames from this pool of structures, based on the experimental data.

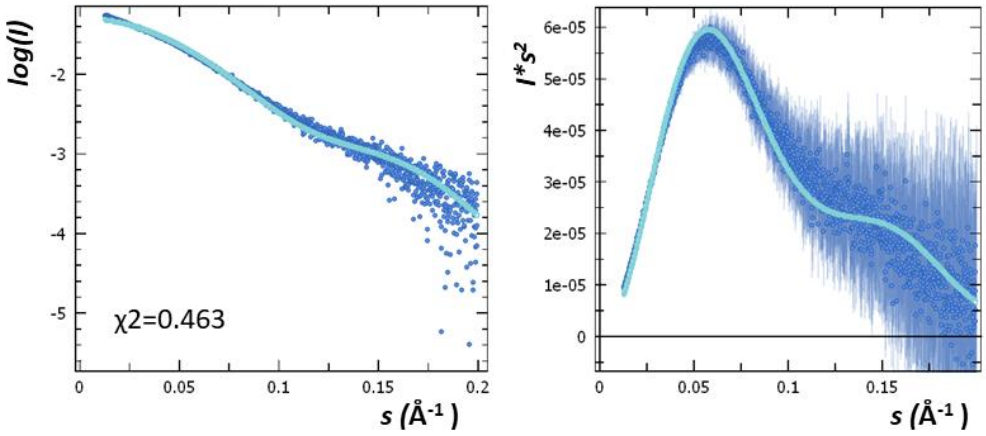


Figure 4.1.6 ***AgrC$_{206}$ ensemble optimisation method.*** *The logarithmic scale plot (left)and Kratky plot (right) show the fitting (light blue) of the selected MD-annealed homology model of dimeric AgrC$_{206}$ against the experimental data (dark blue).*

Frame 177 (Fig. 4.1.7) was identified as the best representative of the experimental data.

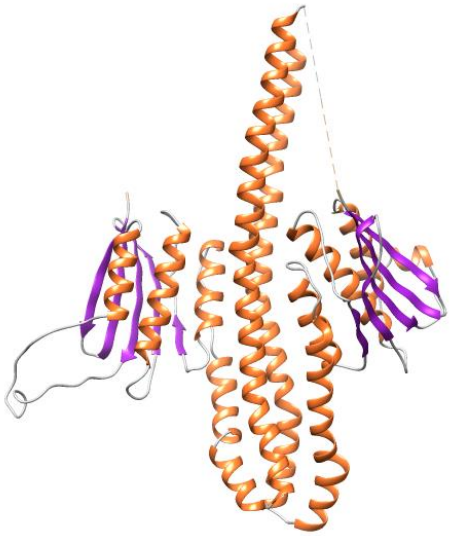


Figure 4.1.7 ***MD frame of AgrC$_{206}$ identified by GAJOE.*** *Frame 177 was indicated by GAJOE to be representative of the scattering profile of AgrC$_{206}$. Selected from a 400 ns MD simulation done by Dr Boyan Bonev.*

The analysis of the MD simulation highlights a degree of flexibility in the region of the long helices providing the interface for dimerisation.

The published crystal structure of the GCN4-AgrC$_{209}$ chimera was also fitted to the experimental data to compare constrained structure to the fit obtained using MD simulations. The crystal structure was cropped to have the same number of residues as our tagged AgrC$_{206}$.

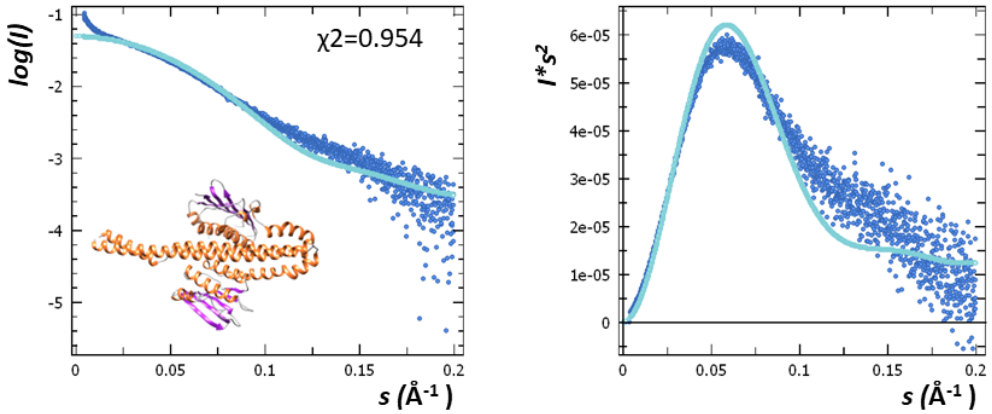


Figure 4.1.8 ***GCN4-AgrC$_{209}$ fit to experimental data.*** *The crystal structure of the GCN4-AgrC$_{209}$ chimeric construct was cropped to have similar length to our construct and fitted (light blue) to the experimental data (dark blue). The logarithmic (left) and Kratky (right) plots are shown together with the structure (inset, please see Fig. 1.3 for citations).*

The fit obtained from the crystal structure is not as precise as the ones shown in Figures 4.1.5 and 4.1.6. While the overall conformations of the crystal and the model are similar, there are a few important differences; the H239, which gets phosphorylated by the catalytic domain, is not solvent exposed but faces the CA domain, and the loop at residues 211 and 213 is absent. On the other hand, the splaying and asymmetry of the two α2 helices is retained. Another difference between our model and the crystal structure is the handedness of the loop connecting α1 and α2. While the crystal structure presents a left-handed loop, our model shows a right-handed loop, consistent with the classification provided by (Ashenberg et al., 2013).

AgrA, the response regulator of the system, is a multidomain protein of which only the DNA-binding domain has been structurally characterised. The N-terminus is thought to interact with AgrC and to drive dimerisation upon

receiving the phosphate group on Asp59. Following a similar protocol to the one described above for $AgrC_{206}$, the scattering data from our construct C199S AgrA were characterised and fitted to an MD-annealed model (generated by Dr Boyan Bonev).

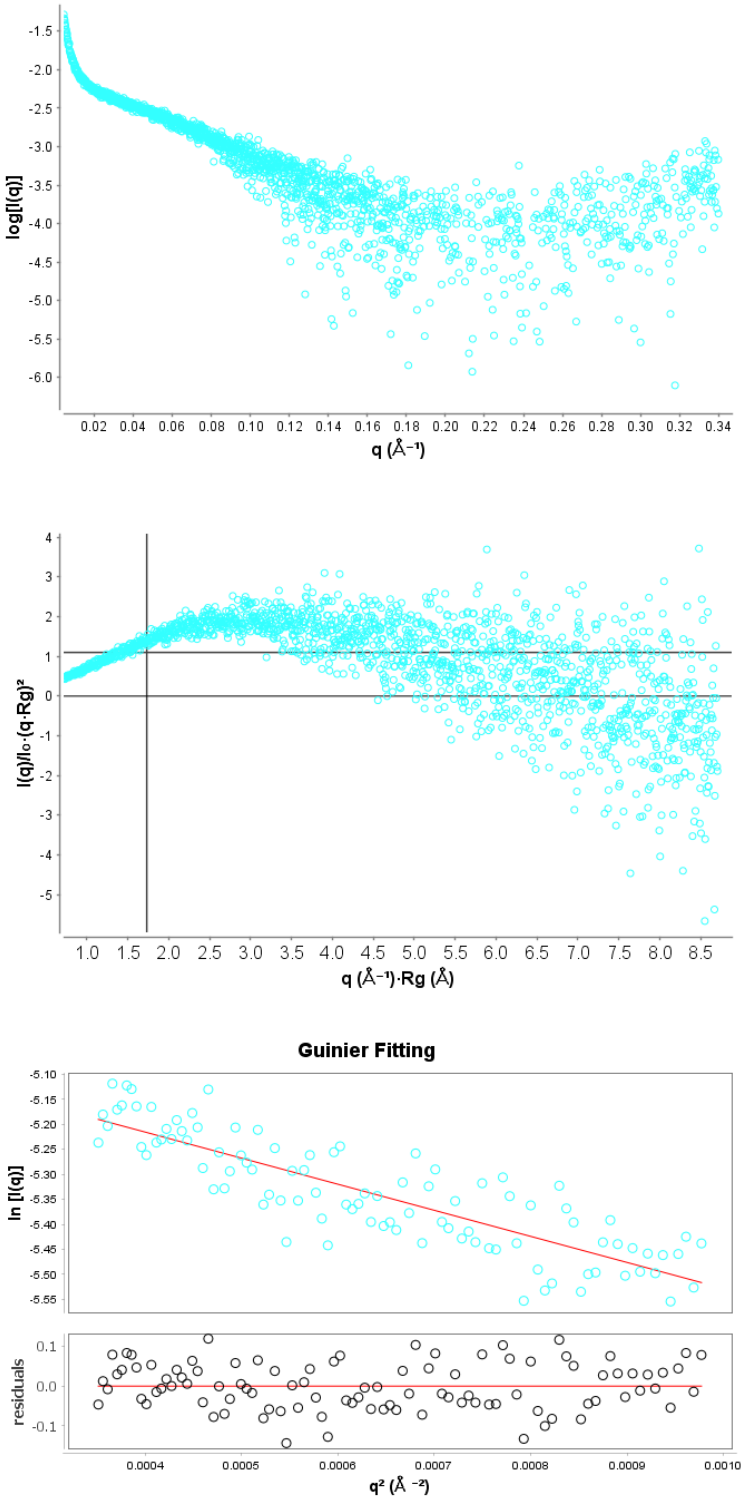


Figure 4.1.9 ***Experimental SAXS data of C199S AgrA.*** *Logarithmic scale (top), dimensionless Kratky (middle) and Guinier analysis (bottom) of C199S AgrA. The normalised Kratky shows an asymmetric protein. The* $q * R_g$ *limits of the Guinier analysis are 0.7318 and 1.2971.*

This protein shows an increased level of noise in the data and more aggregation at lower q (Fig 4.1.9). The corresponding normalised Kratky plot indicates an asymmetric structure which is consistent with the multi-domain nature of the protein.

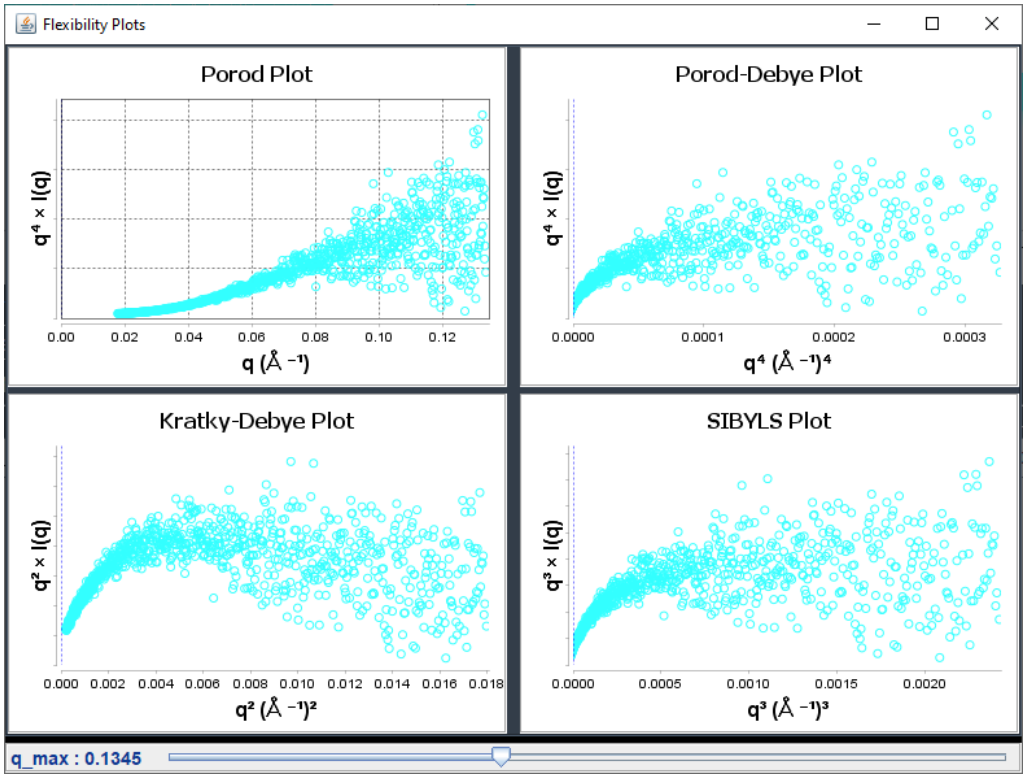


Figure 4.1.10 ***Flexibility analysis of C199S AgrA.*** *The plateau in the SYBILS plot indicates possibly a multi-domain protein.* q_{max} *was calculated at 0.1345.*

The plateau is reached in the SIBYLS plot, which is typical of more flexible or multi-domain proteins (Brosey and Tainer, 2019). The estimated $q_{max} = 0.1345$ was used to calculate the particle volume and Porod exponent.

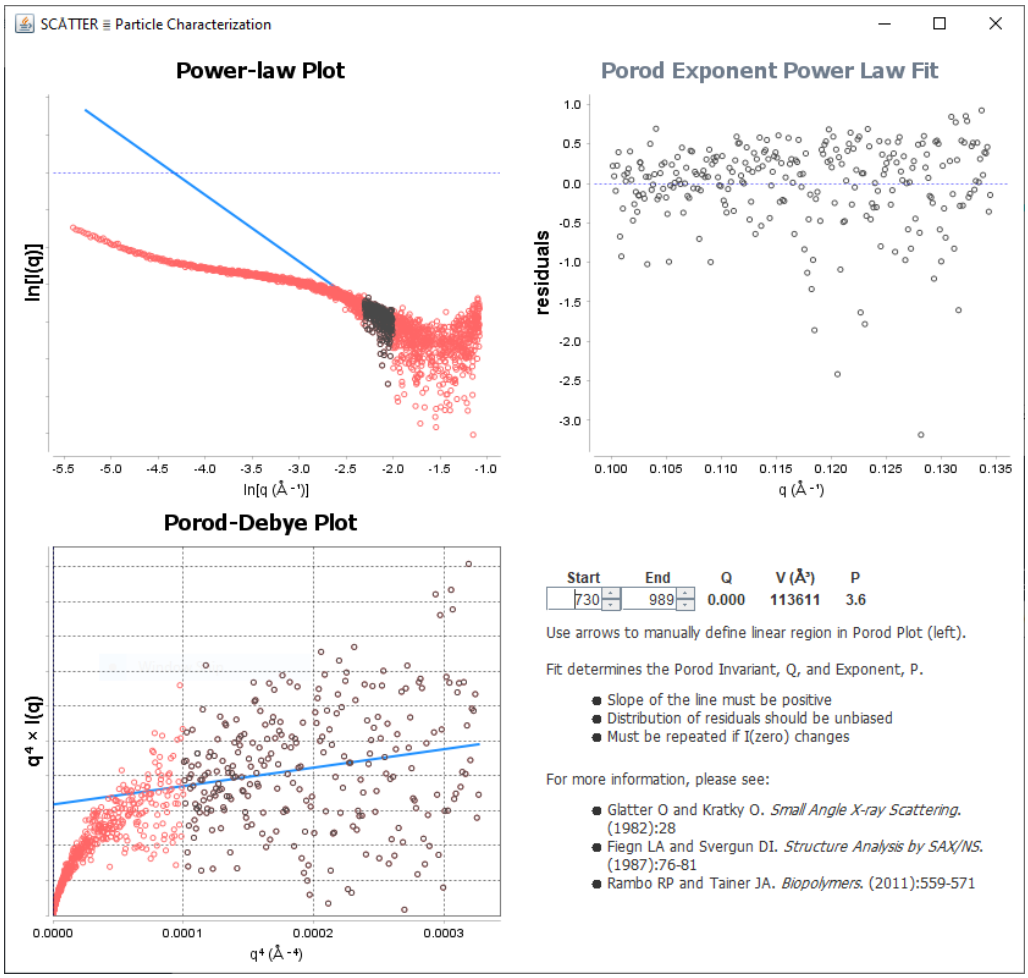


Figure 4.1.11 ***Particle characterisation of C199S AgrA.*** *The linear fit in the Porod-Debye plot, with an unbiased distribution of residuals, is used to calculate the volume of the particle (=115000 Å^3) and the Porod exponent (=3.6).*

The Porod exponent at 3.6 indicates a protein that is slightly more compact than AgrC$_{206}$. The volume 115000 Å^3 corresponds to a density of 0.44 g.cm^{-3}, even lower that the estimate for AgrC$_{206}$. The estimated Rg for this protein is approximately 25 Å, which gives a volume of 65000 Å^3 according to Eq 4.1.2. The volume estimate from a molecular dynamics model is 35 Å^3. It is possible that this protein assembles into higher order structures in solution which would increase the apparent size of the system.

The cluster analysis on 100 ns MD simulations of AgrA revealed three dominant clusters. Averaged structures for each of these clusters were generated and superimposed to assess the degree of variation (Fig. 4.1.12).

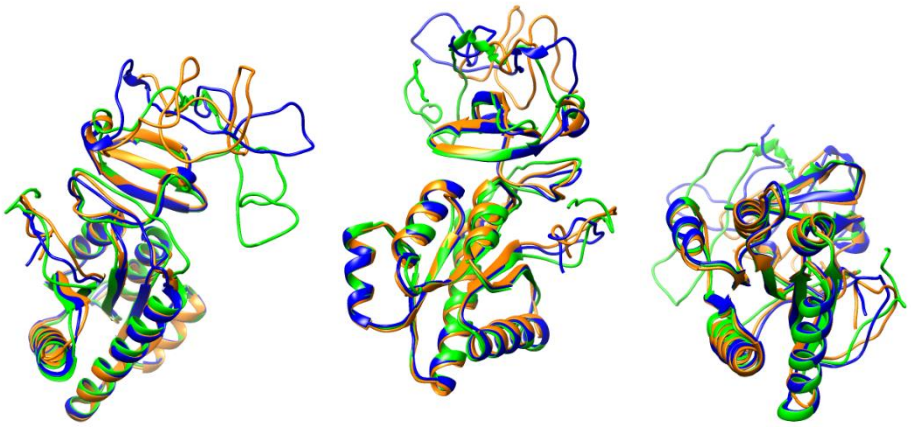


Figure 4.1.12 ***Superimposition of the dominant clusters of MD simulated AgrA.*** *The averaged structure for each dominant cluster in the 100 ns MD simulation of the C199S AgrA construct is shown in a different colour: cluster 1 in green, cluster 2 in orange and cluster 3 in blue. The three structures are superimposed and shown at different angles.*

The superimposition of the three structures reveals little variation in the ordered region of the protein, especially in the N-terminal domain. In the simulation, the DNA-binding C-terminus shows great flexibility, and it is also the only region of the protein that shows discrepancies in the different clusters. We hypothesise that the flexibility results in poorer sample scattering in this area.

The representative frames of the two dominant clusters were used to fit the experimental data using CRYSOL (Fig. 4.1.13).

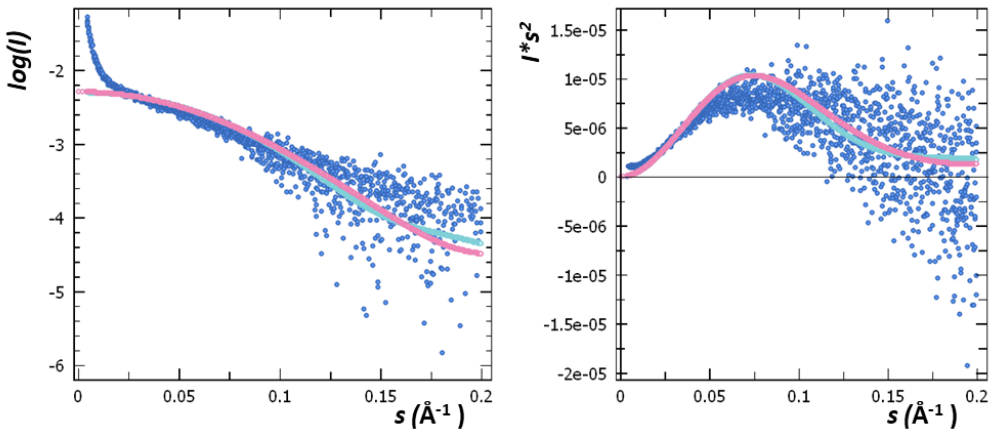


Figure 4.1.13 ***SAXS simulation of C199S AgrA dominant clusters.*** *The simulated scattering curves of the representative frames of the two dominant clusters (cluster 1 light blue and cluster 2 pink) are fitted to the experimental data (dark blue) using CRYSOL (Svergun et al., 1995). The logarithmic plot (left) and Kratky plot (right) are shown.*

There is comparatively little variation in the simulated scattering of the two frames. Neither of them shows a very precise fit to the experimental data, likely due to the quality of the data. Ensemble optimisation was performed on AgrA using GAJOE in the same rationale used for $AgrC_{206}$. The MD trajectory was 100 ns long and done by Dr Bonev.

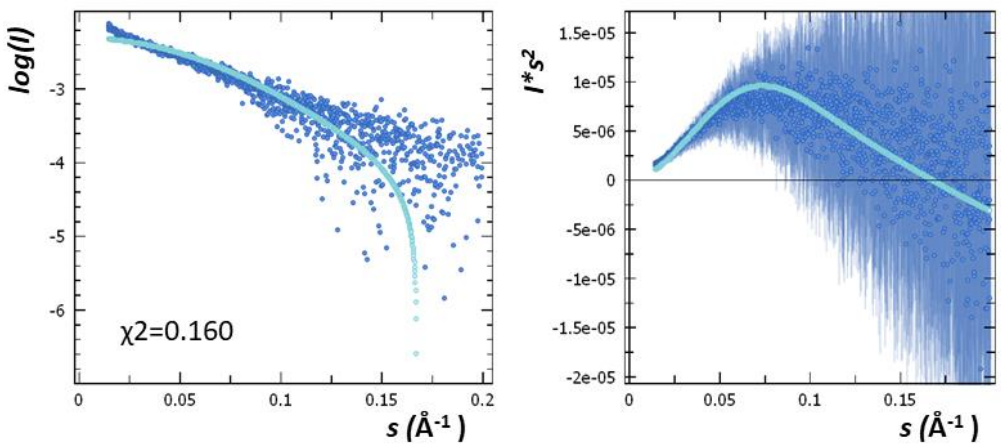


Figure 4.1.14 ***C199S AgrA ensemble optimisation method.*** *The logarithmic scale plot (left)and Kratky plot (right) show the fitting (light blue) of the selected MD-annealed homology model of C199S AgrA against the experimental data (dark blue).*

The fitting was less successful in this case, again due to the quality of the data. One frame, 93, was identified as the most representative (Fig. 4.1.15).

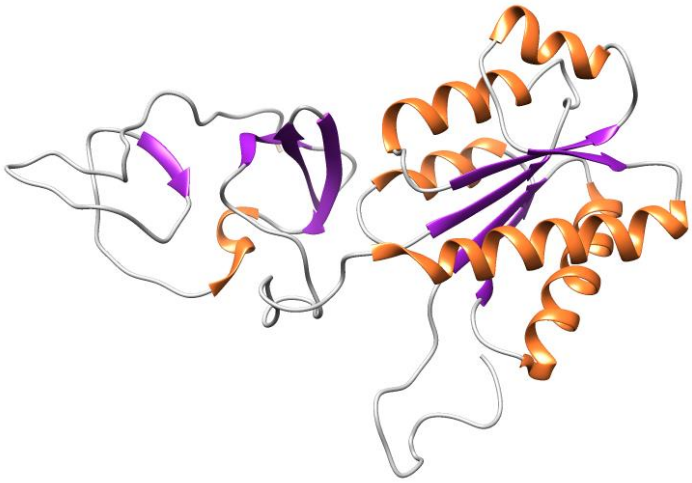


Figure 4.1.15 ***MD frame of C199S AgrA identified by GAJOE.*** *Frame 93 was indicated by GAJOE to be representative of the scattering profile of* $AgrC_{206}$*. Selected from a 100 ns MD simulation done by Dr Boyan Bonev.*

The dominant β-strand conformation is largely lost in the C-terminal domain and is replaced by a highly flexible region. The N-terminus is, however, well defined and is arranged with the five strands in the middle of the core and the four helices surrounding them (Fig. 4.1.15 and 4.1.16).

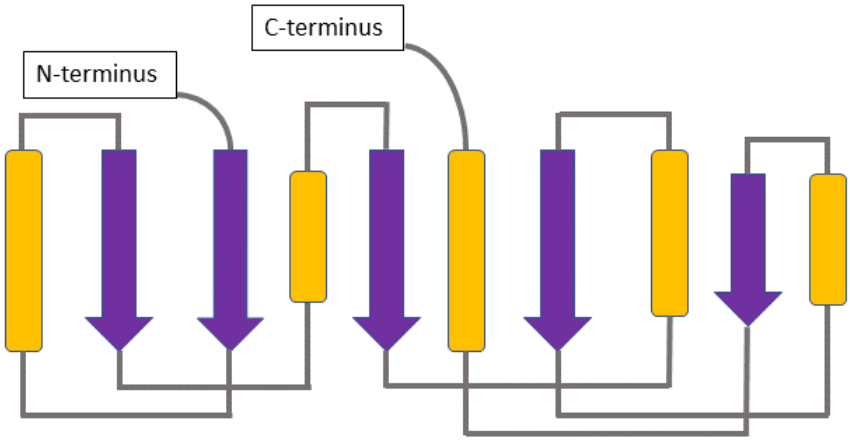


Figure 4.1.16 ***Topology of AgrA N-terminus.*** *The sheets are represented as purple arrows, the helices in orange and the turns in dark grey.*

4.2. The molecular interaction between AgrA and AgrC and the role of ATP

This study aims to characterise the molecular interaction between the two components of the quorum sensing signal transduction system, the histidine kinase AgrC and the transcription regulator AgrA. While interaction between these two proteins has been reported in kinetics studies (Srivastava et al., 2014), a complex formed by these two proteins is yet to be isolated in vitro.

Our constructs C199S AgrA and $AgrC_{206}$ were used to study the behaviour of this system with Small Angle Scattering.

The methodology was then expanded to investigate the effect of ATP and its non-hydrolysable analogue sATP on the system. ATP serves as phosphodonor to AgrC, which auto *trans*-phosphorylate before transferring the phosphate group to partner AgrA.

4.2.1. Small Angle X-ray Scattering

Figure 4.2.1 shows that the mixture of the two proteins in a 1:1 molar ratio gave good quality data, with little aggregation at lower q. The normalised Kratky plot shows a multi-domain structure, which is expected for a mixture of samples.

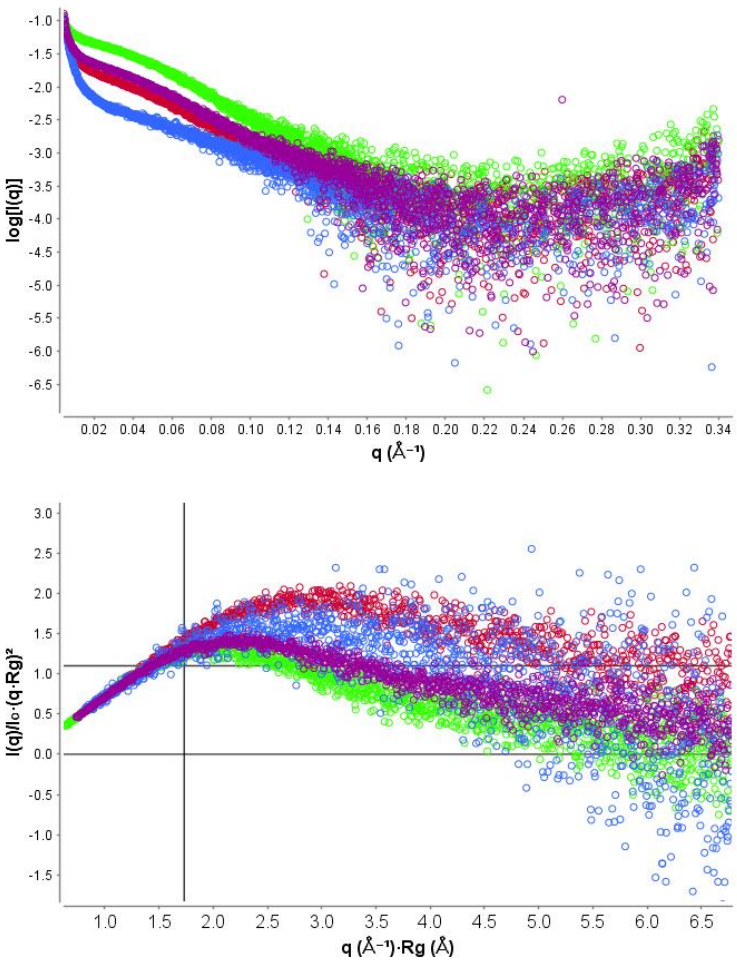


Figure 4.2.1 ***Experimental scattering data of AgrA, AgrC$_{206}$ and AgrA+AgrC$_{206}$.*** *AgrA (blue), AgrC$_{206}$ (green) and their mixture in a 1:1 (red) and 1:2 (purple) molar ratio are shown in a logarithmic (top) and Kratky (bottom) plots. See Appendix for Guinier analysis (Fig. A2).*

AgrA, as shown in the previous subchapter, gives lower quality data, due to the issues in reaching high sample concentrations. The protein mixture in 1:2 molar ratio shows a more compact structure than its 1:1 equivalent, probably due to the increased presence of the well-defined AgrC$_{206}$.

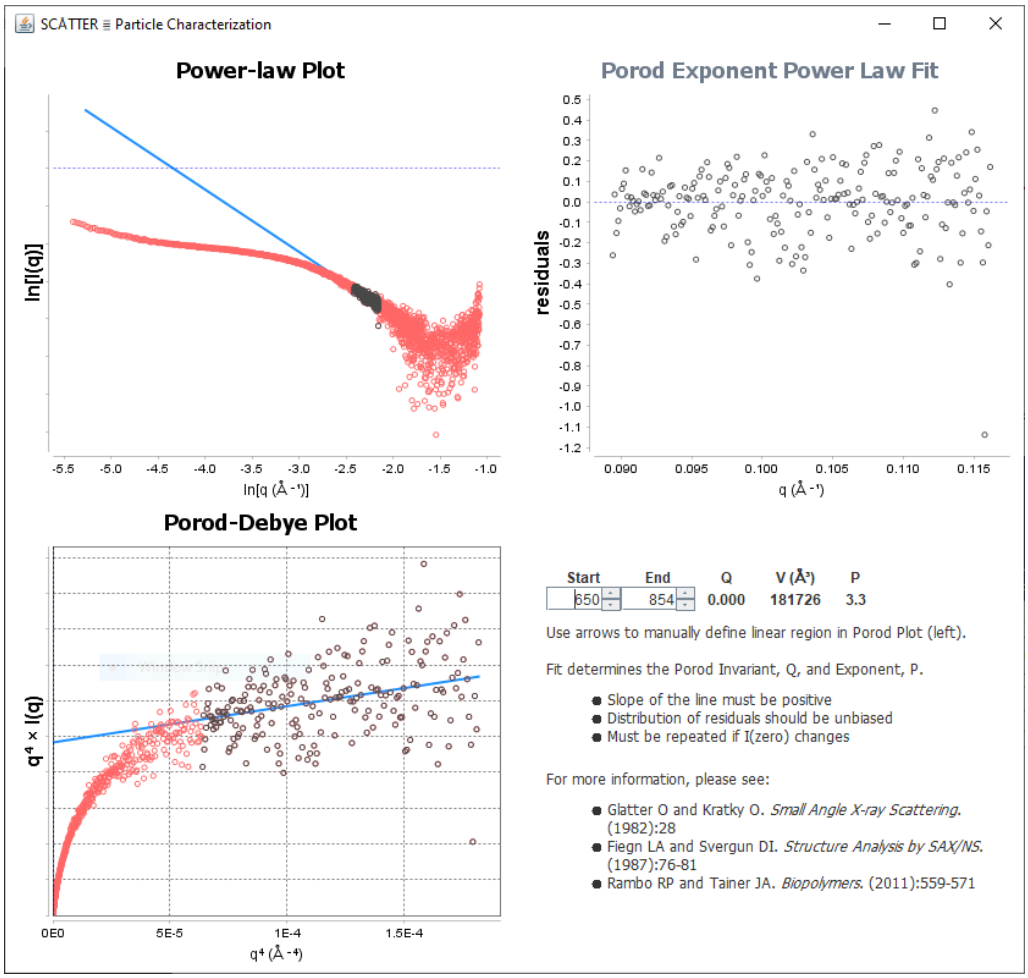


Figure 4.2.2 ***Particle characterisation of the 1:1 AgrA and AgrC$_{206}$ mixture.*** *The linear fit in the Porod-Debye plot, with an unbiased distribution of residuals, is used to calculate the volume of the particle (180000 Å^3) and the Porod exponent (3.3).*

The particle volume is calculated at 180000 Å^3. Given the dimeric conformation adopted by AgrC$_{206}$, the 1:1 ratio implies the interaction of two AgrAs with one AgrC$_{206}$ dimer. Using formula 4.1.1, the estimated protein density is 1.06 g.cm^{-3}.

In the 1:2 ratio system, an AgrA monomer would interact with an AgrC$_{206}$ dimer.

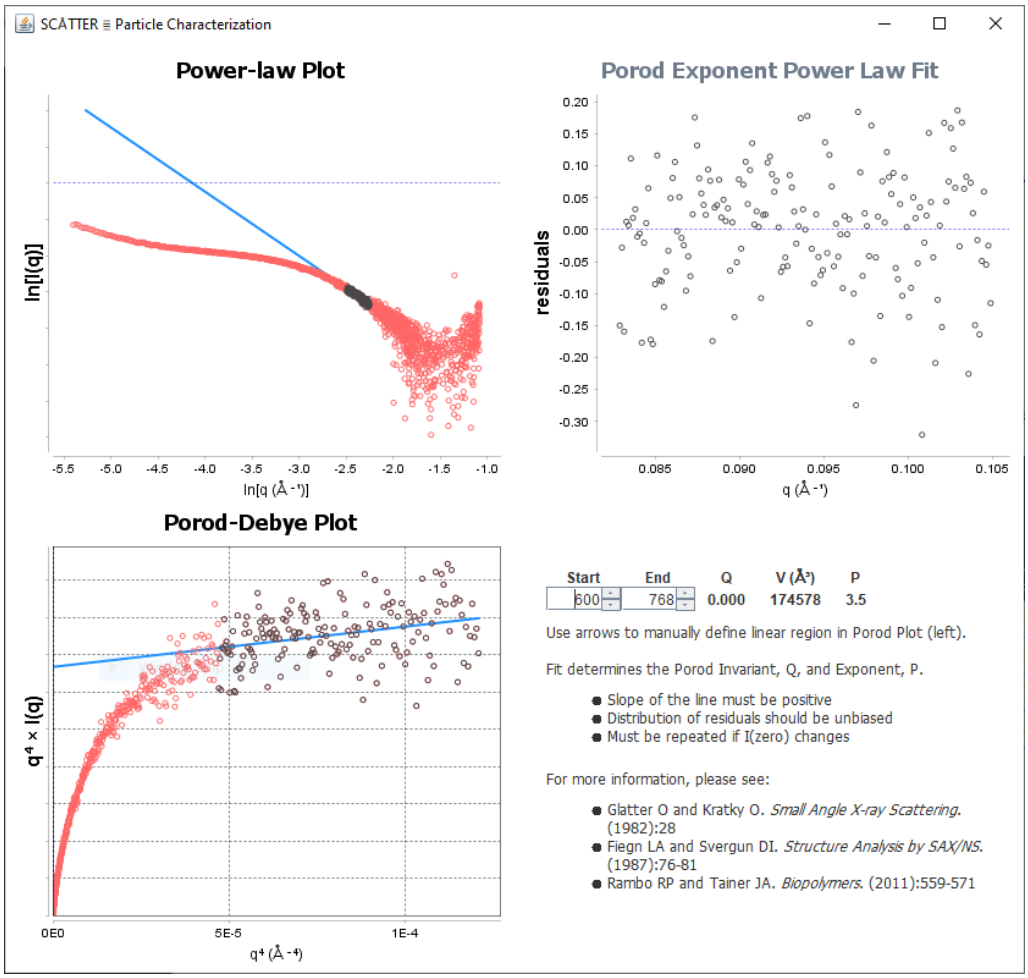


Figure 4.2.3 ***Particle characterisation of the 1:2 AgrA and AgrC$_{206}$ mixture.*** *The linear fit in the Porod-Debye plot, with an unbiased distribution of residuals, is used to calculate the volume of the particle (175000 Å^3) and the Porod exponent (3.5).*

The estimated particle density of the 1:2 molar ratio system is 0.82 g.cm^{-3}, given the volume of 175000 Å^3. A density of 1.06 g.cm^{-3} appears more plausible.

The Complex determination function on ScÅtterIV (Rob Rambo, Diamond Light Source, http://www.bioisis.net/tutorial/9) calculates the best linear combination of the two individual scattering curves which would match the scattering curve of the complex. A nearly perfect fit with flat residuals indicates

the complex is not formed, while an uneven distribution of residuals suggests the formation of a complex.

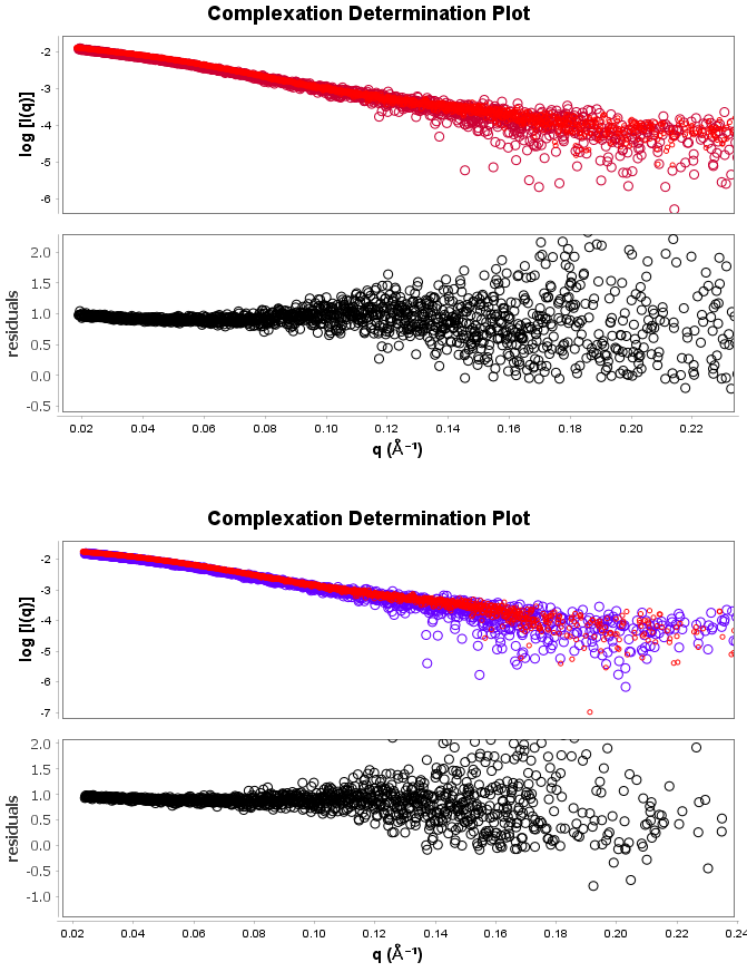


Figure 4.2.4 ***Complexation determination plots of the AgrA and AgrC$_{206}$ mixture.*** *Complexation plots for AgrA and AgrC$_{206}$ mixed in a 1:1 (top) and 1:2 (bottom) molar ratio.*

There is little evidence of the formation of a complex, especially in the case of the 1:2 ratio mixture (Fig. 4.2.4).

High salt concentration is needed for AgrA stability (data not shown). Samples in low salt (250 mM) were also assessed to exclude a screening effect of the high salt which might prevent interaction.

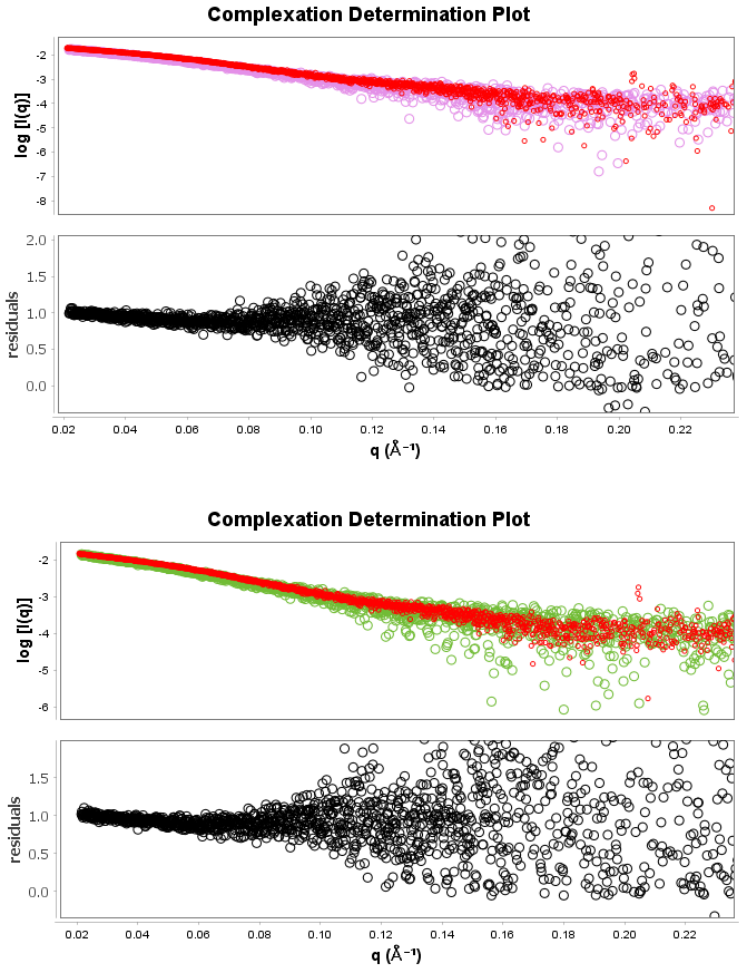


Figure 4.2.5 ***Complexation determination plots of the AgrA and AgrC$_{206}$ mixture in low salt.*** *Complexation plots for AgrA and AgrC$_{206}$ mixed in a 1:1 (top) and 1:2 (bottom) molar ratio, in low salt buffer.*

The lower salt concentration does not alter the complexation plots, which show no evidence of a complex in either molar ratio (Fig. 4.2.5). We therefore conclude that the high salt is not preventing the interaction.

X-rays do not allow for the separation of individual components of a complex as they cannot discriminate between them. The programme MONSA (Svergun, 1999, Svergun and Nierhaus, 2000), however, is able to create bead models of the separate components and arrange them by fitting the scattering patterns of the components and the one of the complex, with the aid of a few values such as volume (or molecular weight).

MONSA was used to create bead models of the AgrA and $AgrC_{206}$ mixture (in both molar ratios) and evaluate the relative arrangements of the two. Figure 4.2.6 shows the appearance of the AgrA-$AgrC_{206}$ mixture.

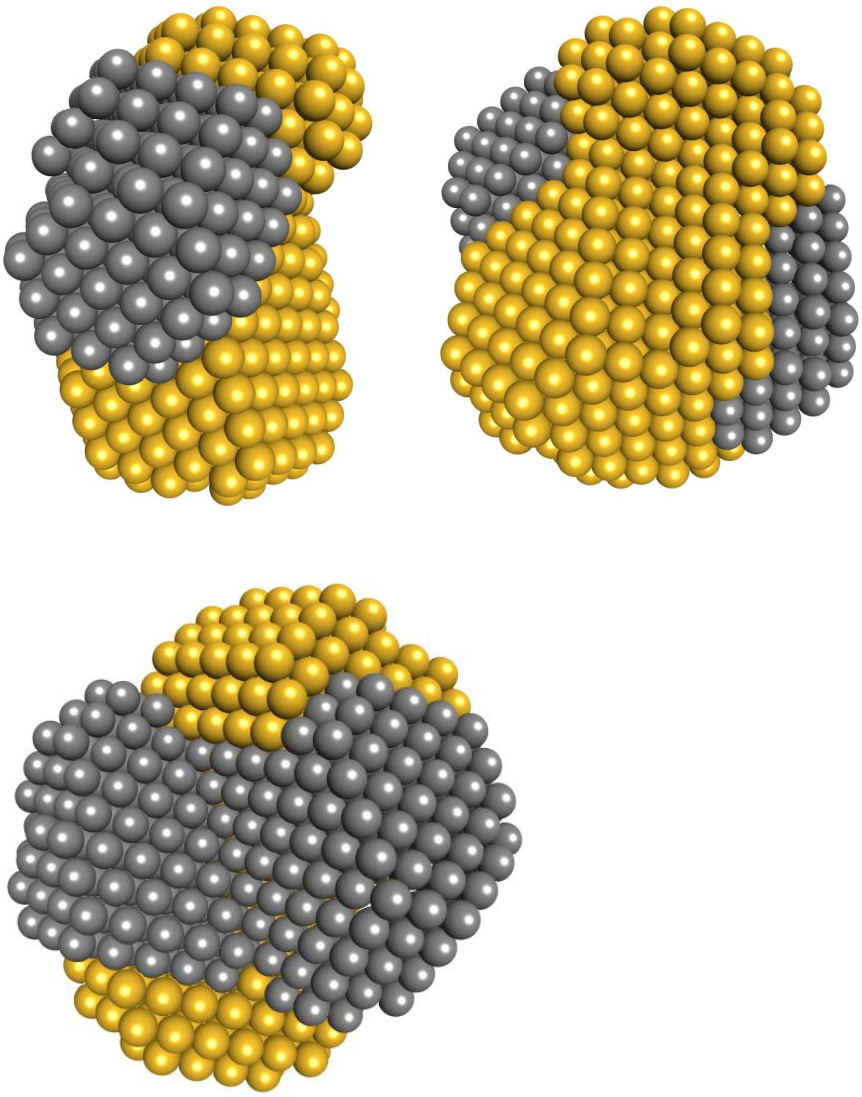


Figure 4.2.6 ***MONSA-calculated bead model of AgrA and AgrC$_{206}$ complex.*** *The relative arrangement of the two proteins (in a 1:1 molar ratio) is calculated using MONSA from the scattering profiles. AgrC$_{206}$ is represented in yellow and AgrA in grey. The complex is shown from different angles. Chi^2 = 0.835 (AgrA), 1.30 (AgrC$_{206}$), 1.21 (AgrA+AgrC$_{206}$).*

The assembly of the MONSA-simulated annealing models is satisfactory, with the two proteins well separated and no superimposition artifact. AgrC$_{206}$ appears to wrap around the barrel represented by AgrA, with AgrA positioning across the helical bundle in a perpendicular fashion. This model supports the proposed

mechanism which suggests H239 as the phosphorylation site for phosphate transfer to AgrA (Fig. 4.2.7) (Cisar et al., 2009).

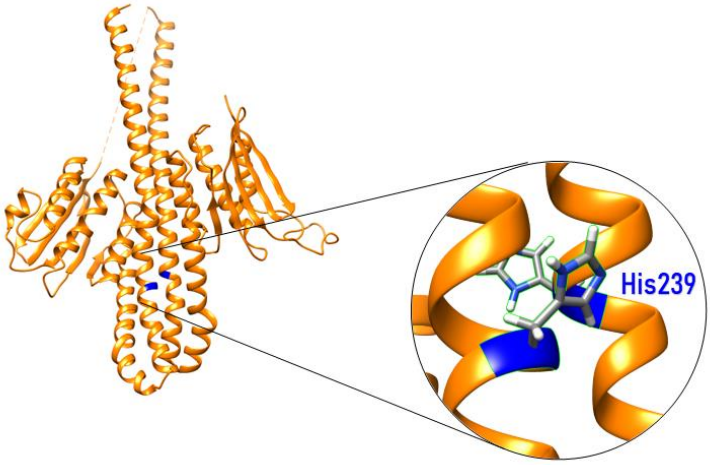


Figure 4.2.7 ***Close-up of the residues predicted to be involved in the phosphorylation of AgrA.*** *A representative MD-annealed model of $AgrC_{206}$ is represented in orange, with the histidine residues (H239) hypothesised to be responsible for AgrA phosphorylation highlighted in blue.*

The protein mixture in a 1:2 ratio (AgrA:$AgrC_{206}$) showed AgrA (represented in grey) creating a bridge through $AgrC_{206}$ (in orange), un unrealistic model (Fig. 4.2.8). This suggests the 1:1 molar ratio to be the correct stoichiometry.

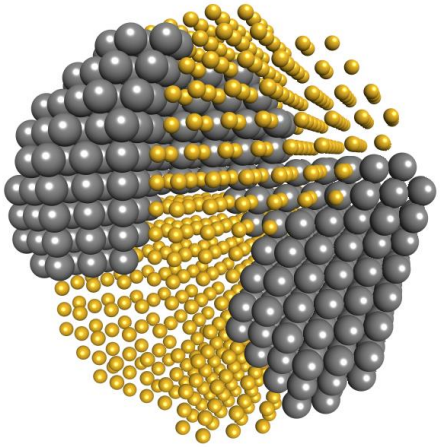


Figure 4.2.8 ***MONSA-calculated bead model of AgrA and AgrC$_{206}$ mixture (1:2).*** *The relative arrangement of the two proteins (in a 1:2 molar ratio) is calculated using MONSA from the scattering profiles. AgrC$_{206}$ is represented in orange and AgrA in grey.*

The effects of ATP and its non-hydrolysable analogue sATP on this system were also studied following a similar pipeline.

Firstly the scattering pattern of AgrC$_{206}$ upon addition of nucleotides was investigated, as this protein is responsible for accepting the phosphate group from ATP (Fig. 4.2.9).

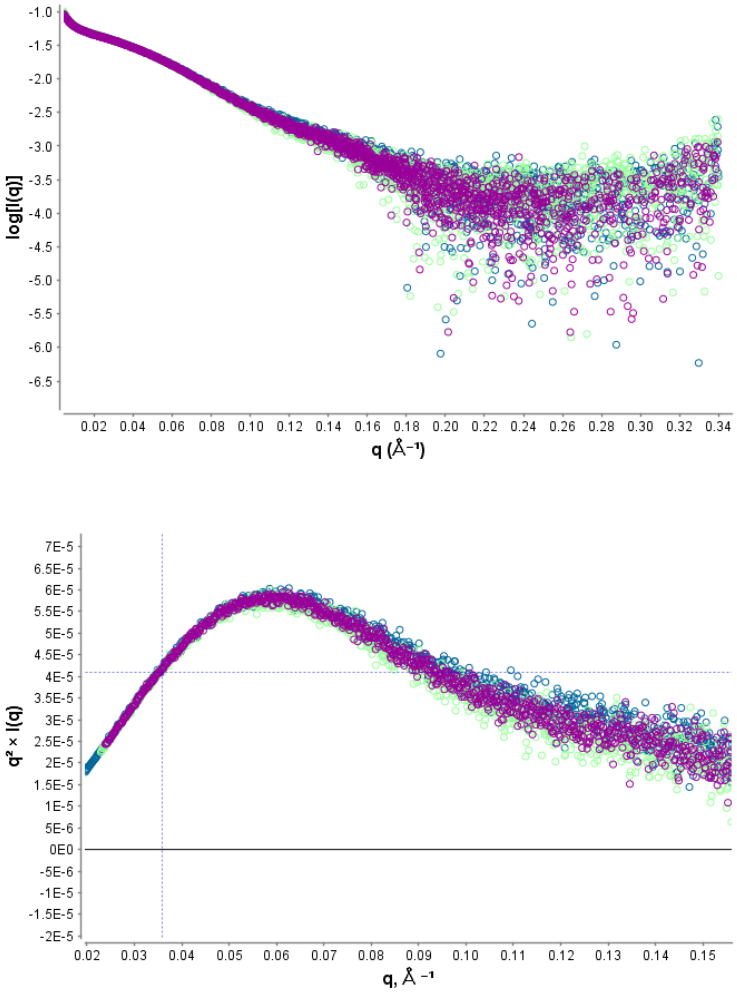


Figure 4.2.9 ***Effect of nucleotides on AgrC$_{206}$.*** *The logarithmic plots (top) and Kratky plots (bottom) are shown for AgrC$_{206}$ (green), AgrC$_{206}$+ATP (blue) and AgrC$_{206}$+sATP (purple). See Appendix for Guinier analysis (Fig. A3).*

The addition of ATP and non-hydrolysable ATP does not appear to alter the scattering pattern in the logarithmic and Kratky plots.

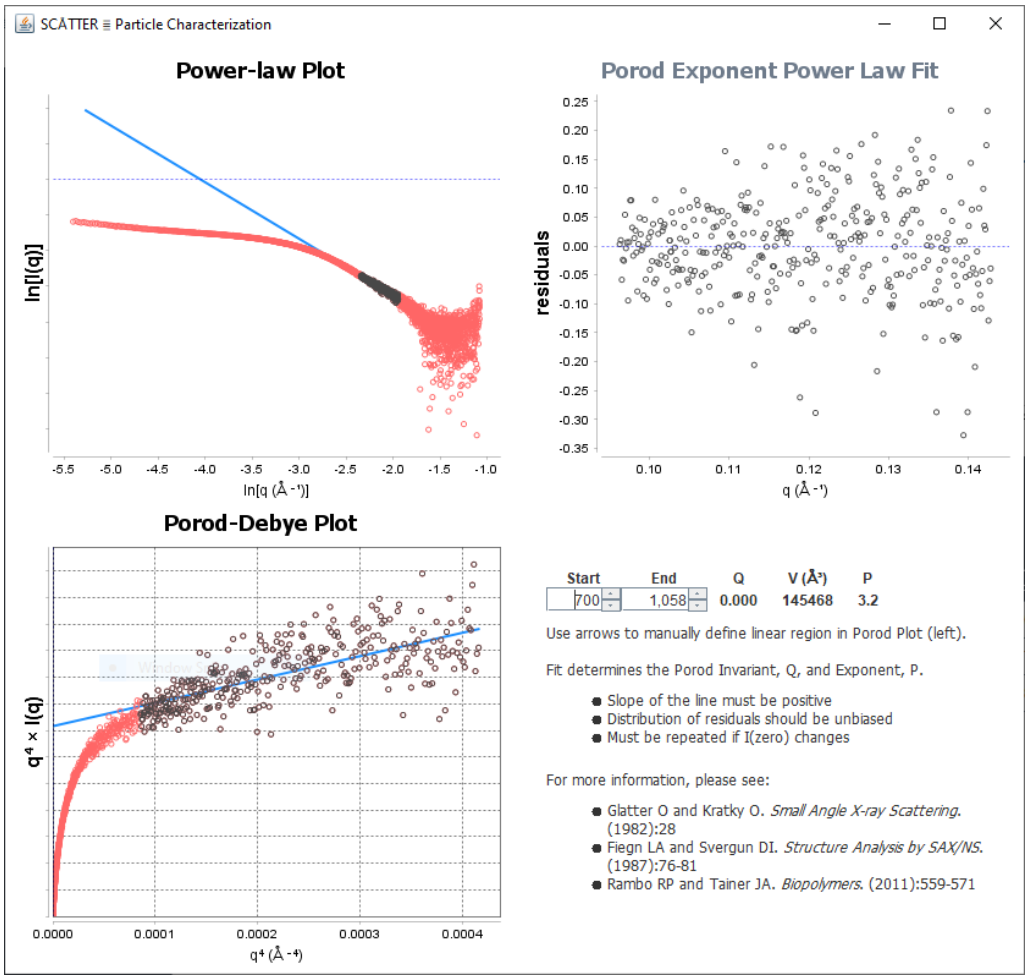


Figure 4.2.10 ***Particle characterisation of AgrC$_{206}$+ATP.*** *The linear fit in the Porod-Debye plot, with an unbiased distribution of residuals, is used to calculate the volume of the particle (145000 Å^3) and the Porod exponent (3.2).*

The addition of ATP (Fig. 4.2.10) does not significantly alter the estimated volume of AgrC$_{206}$. Non-hydrolysable ATP, however, triggers an increase in volume (160000 Å^3) which, coupled with a lower Porod exponent (3.2) suggests a less compact structure (Fig. 4.2.11).

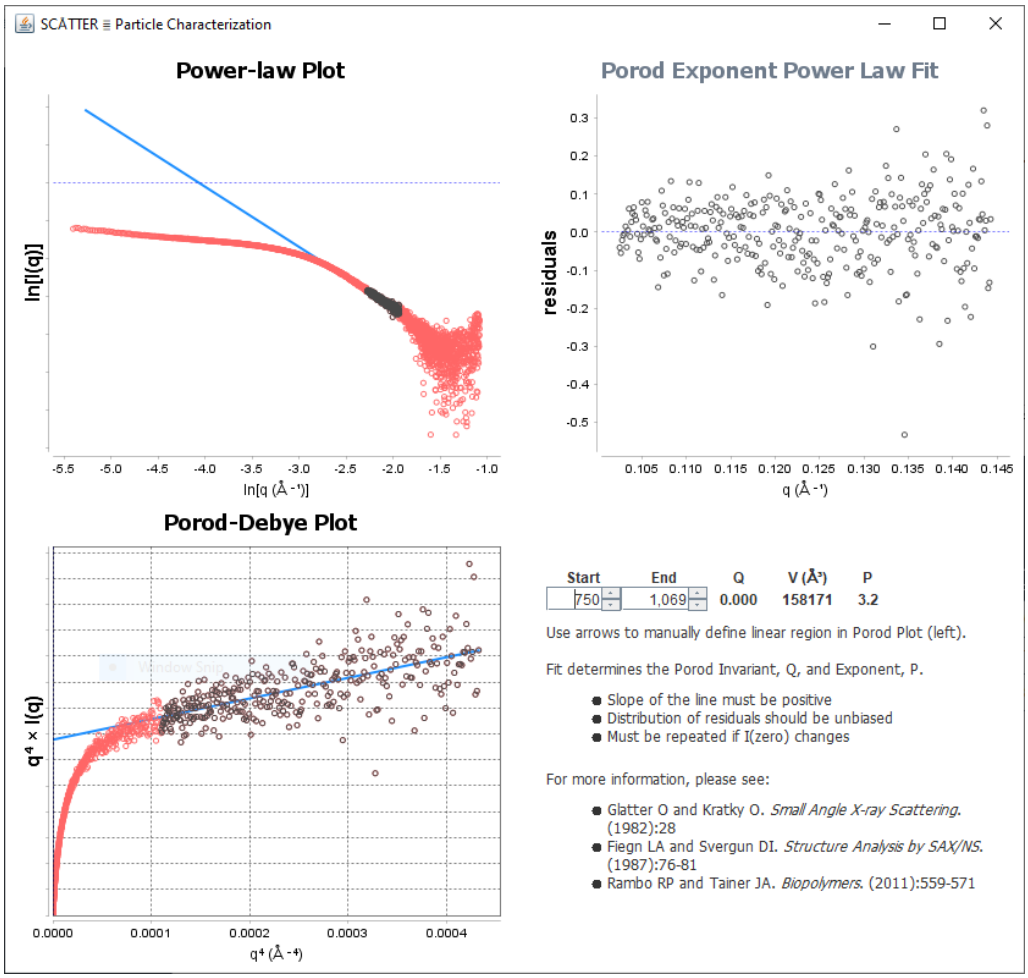


Figure 4.2.11 ***Particle characterisation of AgrC$_{206}$+sATP.*** *The linear fit in the Porod-Debye plot, with an unbiased distribution of residuals, is used to calculate the volume of the particle (160000 Å^3) and the Porod exponent (3.2).*

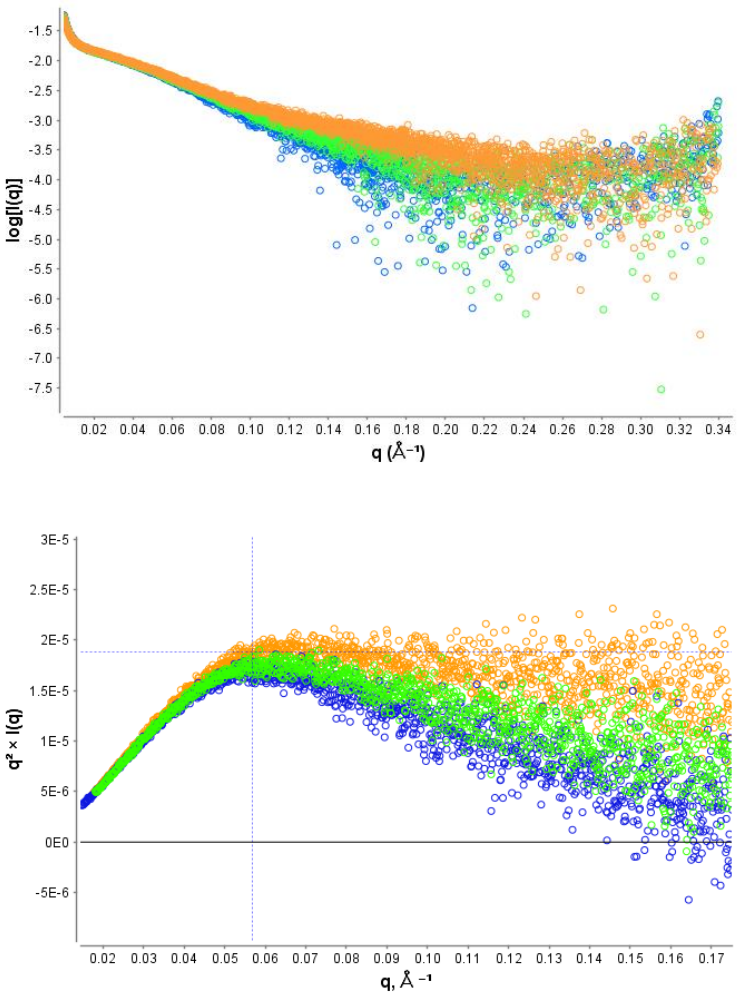
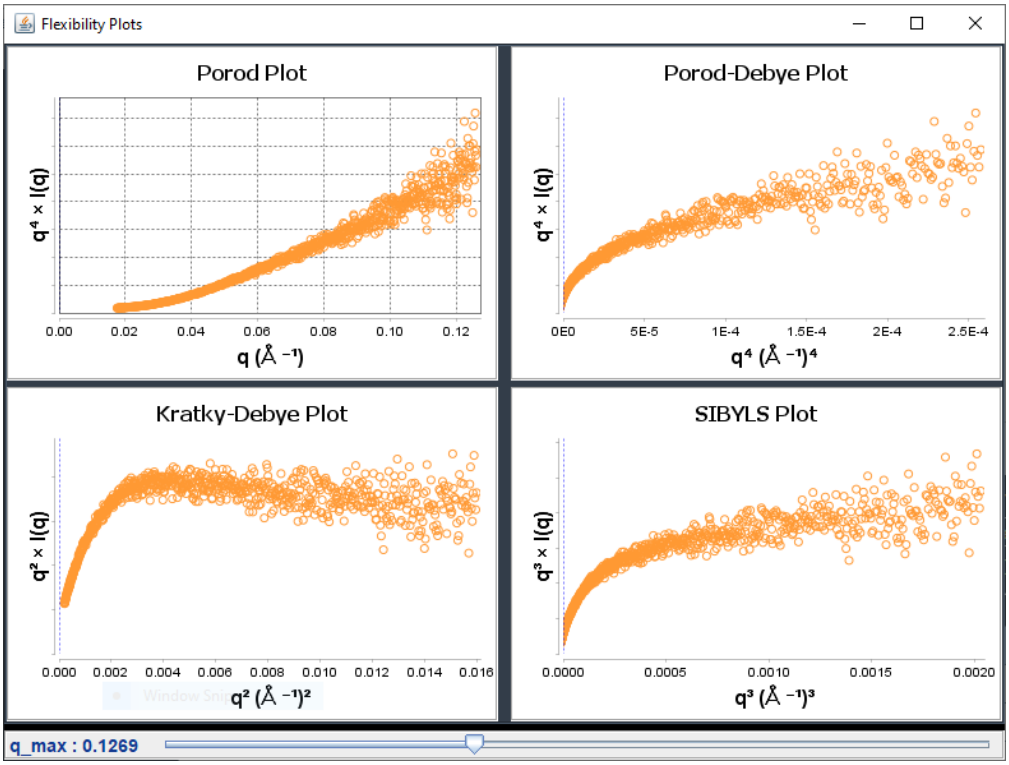


Figure 4.2.12 ***Effect of nucleotides on the AgrA and AgrC$_{206}$ mixture.*** *The protein mixture in absence of nucleotides is shown in blue, in presence of ATP in orange and sATP in green. The molar ratio of the two proteins is 1:1; the logarithmic plot (top) and Kratky plot (bottom) are shown. See Appendix for Guinier analysis (Fig. A4).*

The addition of nucleotides induces tertiary structure rearrangement in the complex. The Kratky shows that both ATP and sATP create a less well-folded structure compared to the system alone, especially ATP (Fig. 4.2.12). The Guinier region is similar across the three samples, while variation is observed at higher *q* (Fig. 4.2.12).



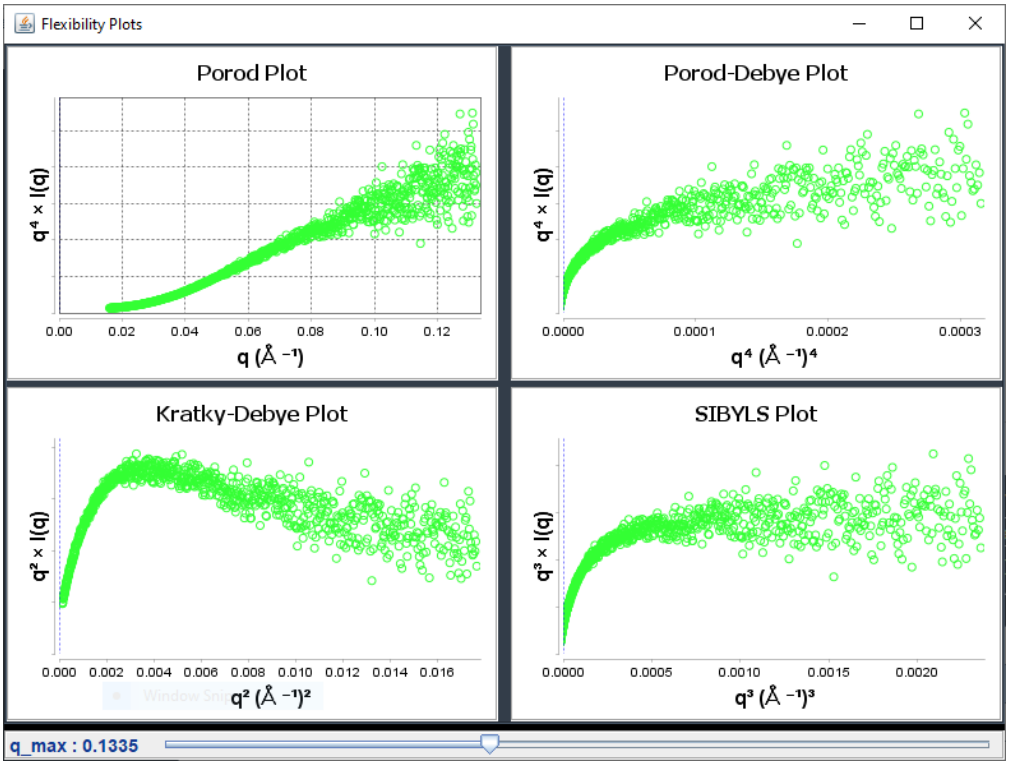


Figure 4.2.13 ***Flexibility analysis of AgrA+AgrC$_{206}$ with nucleotides.*** *The protein mixture with ATP is shown in orange, and with sATP in green. The latter*

shows a well-defined plateau in the SYBILS plot, while the plateau is less convincing in the sample with ATP.

While both samples have their plateaus in the SYBILS plot, the sample with sATP is much better defined.

The Complex determination function was used here to investigate whether the addition of nucleotides triggers a complex formation.

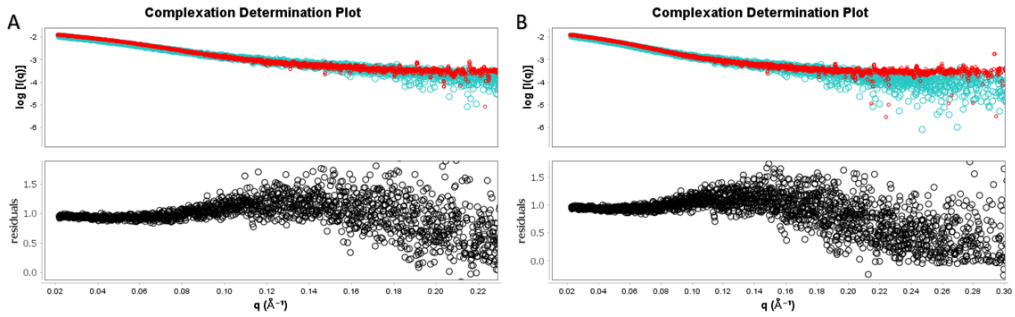


Figure 4.2.14 ***Complexation determination plots for AgrA and AgrC mixture and ATP.*** *Complexation plots for C199S AgrA and $AgrC_{206}$ (1:1 ratio) (A) with ATP, subtracted with AgrA and $AgrC_{206}$ and (B) with ATP, subtracted with AgrA and $AgrC_{206}$+ATP.*

The addition of ATP to the mixture of C199S AgrA and $AgrC_{206}$ shows marked differences to the complexation landscape (Fig. 4.2.14). The distribution of residuals suggests the formation of a complex between the two proteins, aided by the presence of excess ATP. This is more evident when the data is subtracted with $AgrC_{206}$ and ATP (Fig. 4.2.14 B).

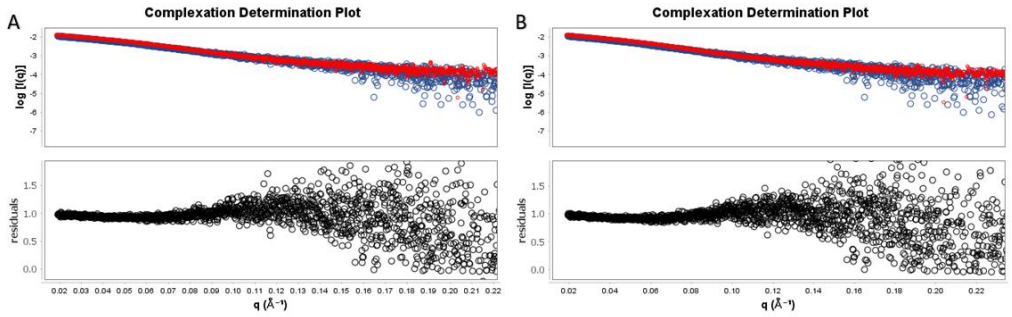


Figure 4.2.15 ***Complexation determination plots for AgrA and AgrC mixture and non-hydrolysable ATP.*** *Complexation plots for C199S AgrA and AgrC$_{206}$ (1:1 ratio) (A) with sATP, subtracted with AgrA and AgrC$_{206}$, (B) with sATP, subtracted with AgrA and AgrC$_{206}$+sATP.*

The addition of non-hydrolysable ATP does not suggest stabilisation of the complex as it was originally hypothesised. The residuals profile is not as flat as the mixture in absence of nucleotides but not distinctive enough to determine complex formation with certainty (Fig. 4.2.15 A and B).

4.2.2. Small Angle Neutron Scattering

Neutron scattering is a valuable asset in the study of molecular complexes as it allows to investigate the changes in the individual components while they are in a complex, by contrast matching. It is particularly useful when one or more proteins can be affected by molecular interactions, which can be difficult to resolve by SAXS.

One of the major advantages of neutron scattering is the possibility to use deuterium oxide to contrast match samples. Samples can be deuterated during production to increase their scattering and separate it from the scattering of the protonated surrounding environment or buffer. In terms of the study of molecular

complexes, this means it is possible to selectively deuterate one of the components and contrast-match the other to obtain the scattering profile of the deuterated element. It was applied to this system, using three contrasts: 100% H_2O, 100% D_2O and 42% D_2O. Deuteration of $AgrC_{206}$ is described in Chapter 3.

Small angle neutron scattering was used to investigate potential structural changes in $AgrC_{206}$ and AgrA in presence and absence of ATP and non-hydrolysable ATP. This was done by selectively deuterating $AgrC_{206}$, combining it with protonated AgrA and contrast matching the buffer to both components.

The D_2O contrast shows how the protonated component (C199S AgrA) is affected by the addition of ATP and sATP, in presence of $AgrC_{206}$. The linear plot shows marked differences in both cases as opposed to the proteins alone. A minimum appears at approximately 0.2 $Å^{-1}$, while at higher q the region characterised by a minimum exhibits a new peak. The lack of compact structure observed in the Kratky plot disappears and the system adopts a more globular shape.

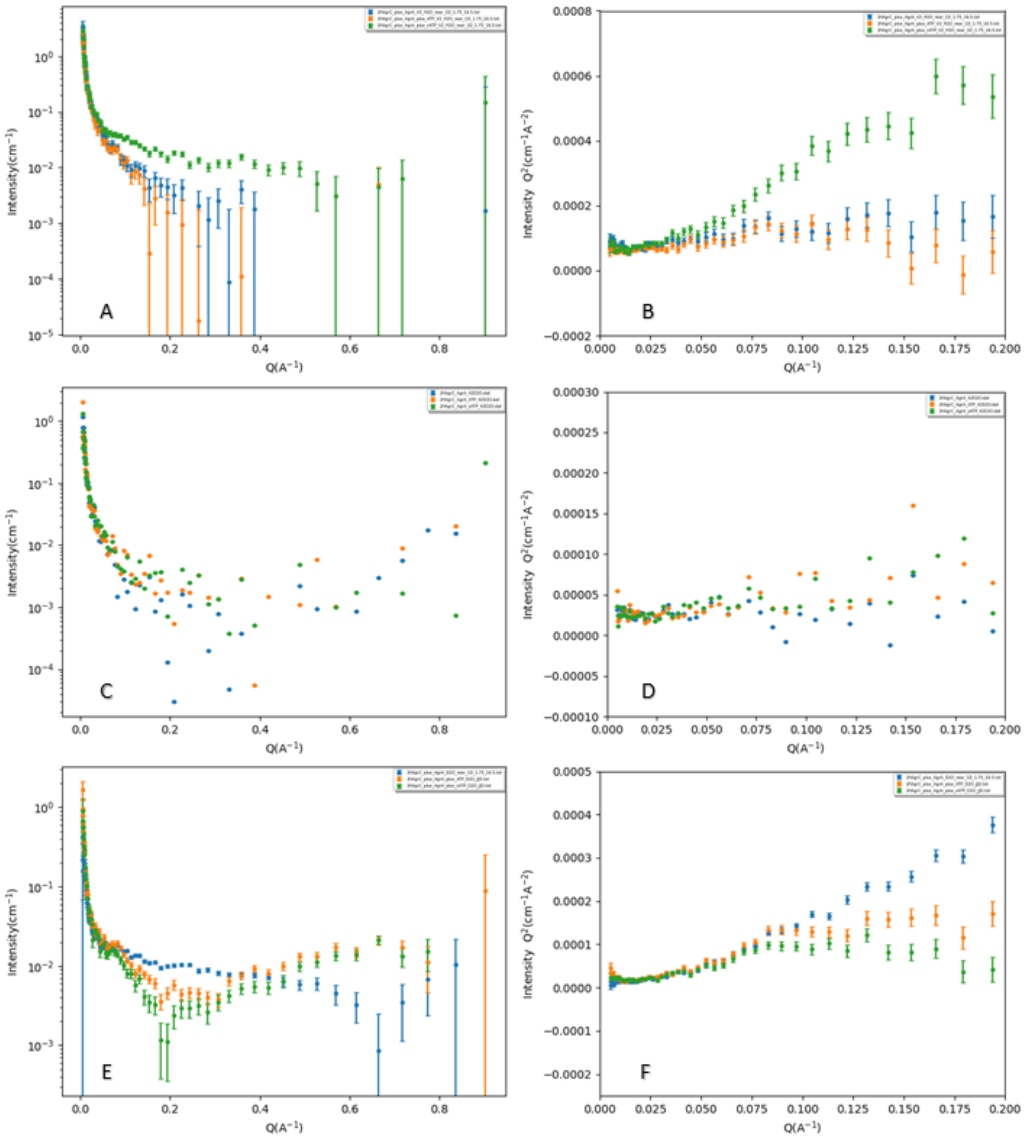


Figure 4.2.16 ***BioSANS of C199S AgrA and ^{2}H-AgrC$_{206}$.*** *(A) Linear plot of C199S AgrA and ^{2}H-AgrC$_{206}$ in 100% H_2O, (B) Kratky plot of C199S AgrA and ^{2}H-AgrC$_{206}$ in 100% H_2O, (C) Linear plot of C199S AgrA and ^{2}H-AgrC$_{206}$ in 42% D_2O, (D) Kratky plot of C199S AgrA and ^{2}H-AgrC$_{206}$ in 42% D_2O, I Linear plot of C199S AgrA and ^{2}H-AgrC$_{206}$ in 100% D_2O and (F) Kratky plot of C199S AgrA and ^{2}H-AgrC$_{206}$ in 100% D_2O. Samples in absence of nucleotides are shown in bluI, in presence of ATP in orange and in presence of non-hydrolysable ATP in green.*

The buffer containing 42% D_2O contrast-matches the protonated component, showing $AgrC_{206}$. This dataset doesn't match the quality of the other two and does not show any discrepancies resulting from the addition of nucleotides.

The H_2O buffer shows scattering from both components, with a preference for the deuterated protein. The addition of ATP does not trigger any conformational rearrangement, while the sample with s-ATP shows appreciable differences in globularity as well as overall shape.

These data need to be interpreted in conjunction with the corresponding Size Exclusion Chromatography (see Chapter 3) and thermolysis studies (see Chapter 5).

4.3.Effect of IQS analogues on the tertiary structure of AgrA

The experiments described in this section are aimed at detecting any possible structural alteration of full-length and C-terminal AgrA resulting from the interaction with their proposed inhibitors IQS3, 11 and 15. IQS3 inhibits exotoxin expression in *S. aureus* as well as AgrA-P2/P3 interactions as observed via EMSA assays (Jaiyen, 2016). IQS11 and IQS15 are structurally related to IQS3 but do not inhibit AgrA. Consequently, small detectable variations in the profiles of both full-length and C-terminal AgrA when studied by SAXS may be observed.

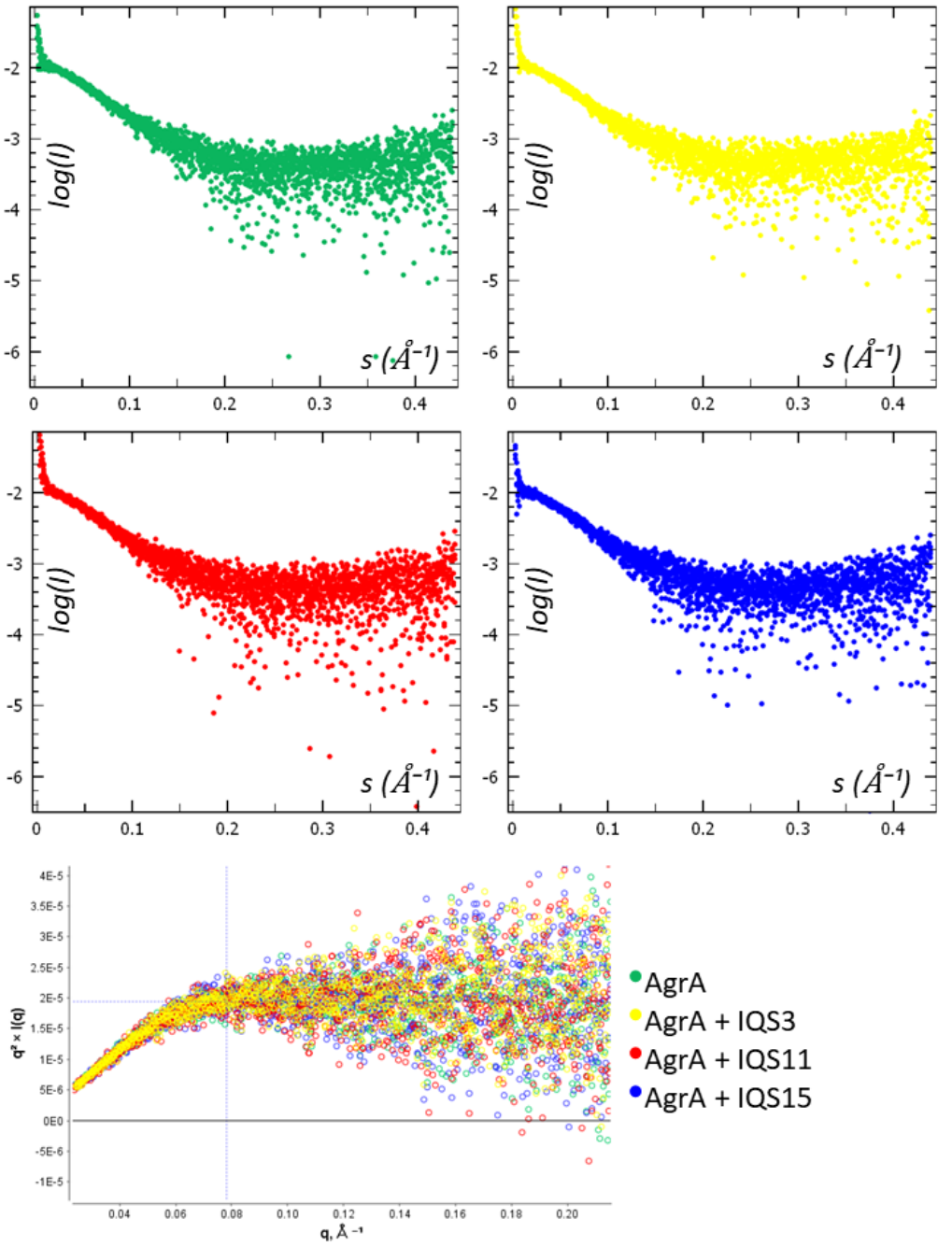


Figure 4.3.1 ***Experimental scattering data of AgrA and IQS analogues.*** *Logarithmic scale plots and Kratky plot of full-length AgrA and IQS analogues. All spectra were recorded at room temperature (27 °C). See Appendix for Guinier analysis (Fig. A5).*

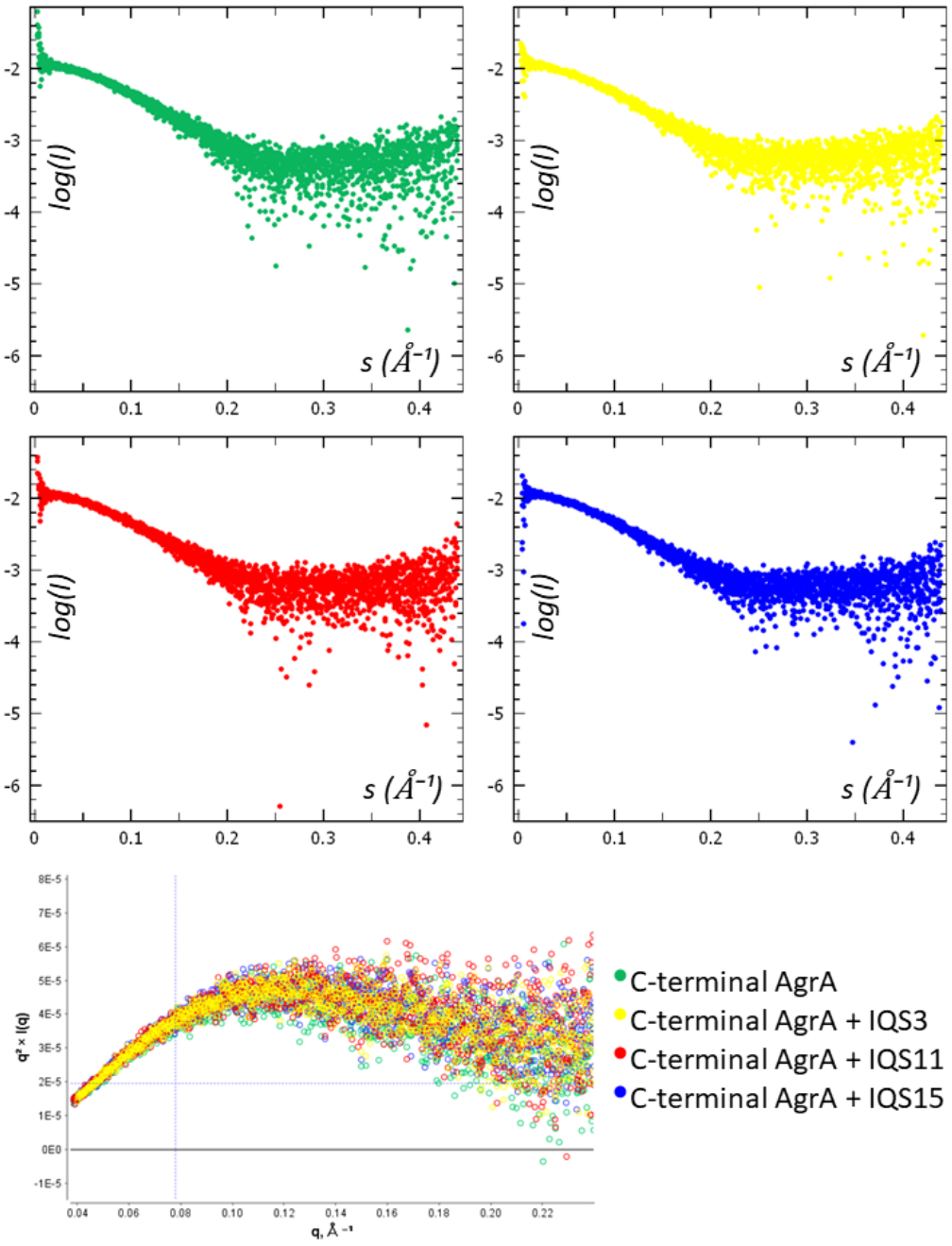


Figure 4.3.2 ***Experimental scattering data of AgrA and IQS analogues.*** *Logarithmic scale plots and Kratky plot of C-terminal AgrA and IQS analogues. All spectra recorded at room temperature (27 °C). See Appendix for Guinier analysis (Fig. A6).*

The logarithmic plots (Fig. 4.3.1 and 4.3.2) show very similar profiles when the protein is mixed with the analogues, in the case of both full-length AgrA and C-

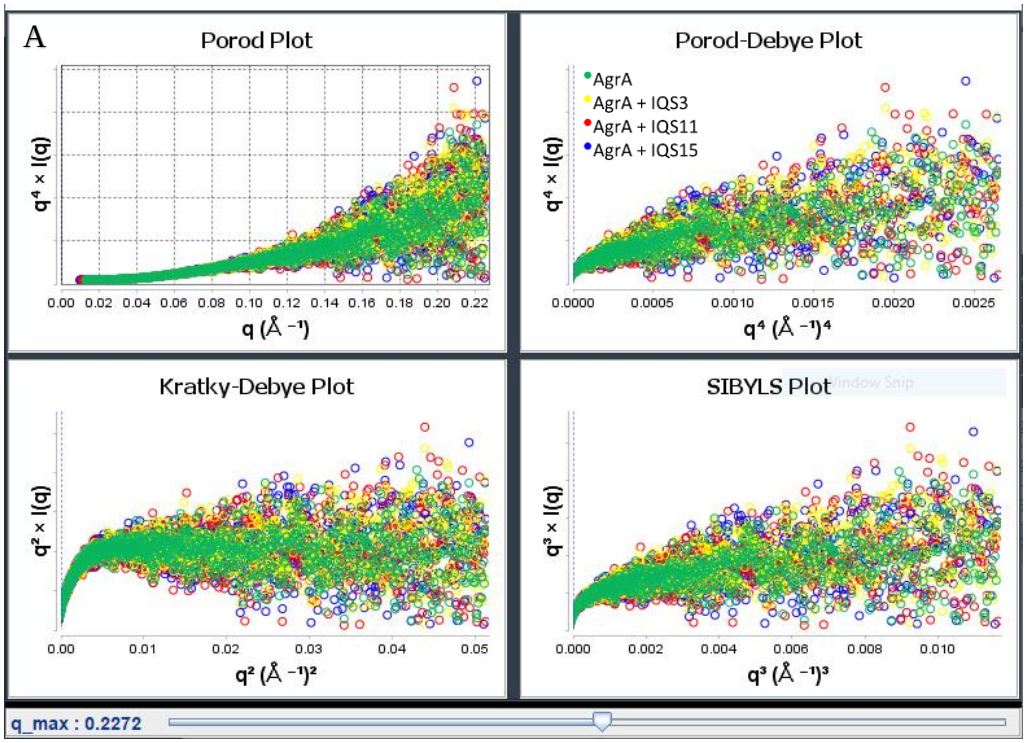
terminal AgrA. The normalised Kratky plots show very little difference in the globularity of the proteins in presence and absence of IQS. The radius of gyration was calculated and is shown in Table 4.3.1.

Table 4.3.1 R_g of AgrA and C-terminal AgrA in presence and absence of IQS analogues.

IQS analogue	Full-length AgrA	C-terminal AgrA
No IQS	24 Å	17 Å
IQS3	24 Å	17 Å
IQS11	24 Å	17 Å
IQS15	24 Å	17 Å

Both full-length and C-terminal AgrA show uniformity across the data sets, with no significant change in R_g.

The effect of the IQS analogues on the flexibility of the two proteins was also investigated. The flexibility plots are shown in Figure 4.3.3.



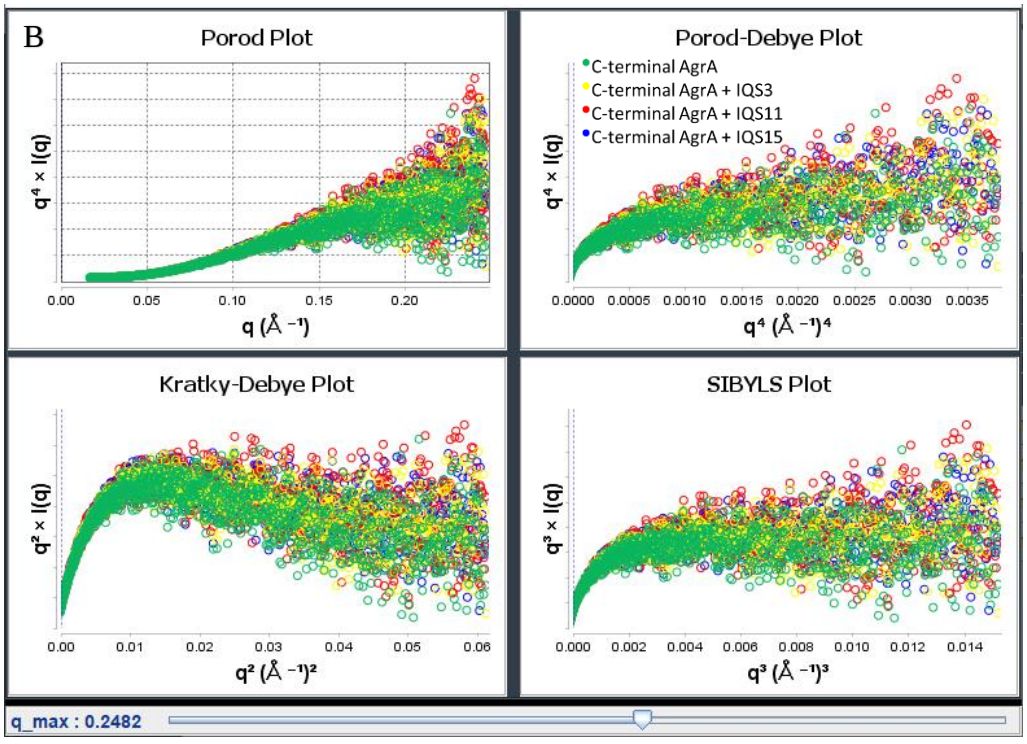


Figure 4.3.3 ***Flexibility plots of full-length AgrA (A), C-terminal AgrA (B) and IQS analogues.***

The flexibility plots reveal AgrA as a flexible protein, as the plateau is reached in the Kratky-Debye plot. The C-terminal domain, on the other hand, plateaus in

the Porod-Debye plot and is therefore compact. In both cases, the addition of IQS analogues does not elicit significant variation in the flexibility.

The pair-distance distribution function (P(r) distribution) is another parameter that can be indicative of structural change. It is measured in the real space (as opposed to the reciprocal space used in the previous analysis) and relates to the paired-set of distances between points within the molecule, making it sensitive to alterations in even just a few residues. It also describes the radius of gyration and the intensity in the real space. These two values should broadly match in the real and reciprocal space but imperfect Guinier or P(r) analyses can give slight variation. As the scattering profiles and the flexibility analysis of AgrA in presence and absence of IQS analogues fail to show any significant change, the P(r) distribution was used to investigate possible changes in the real space.

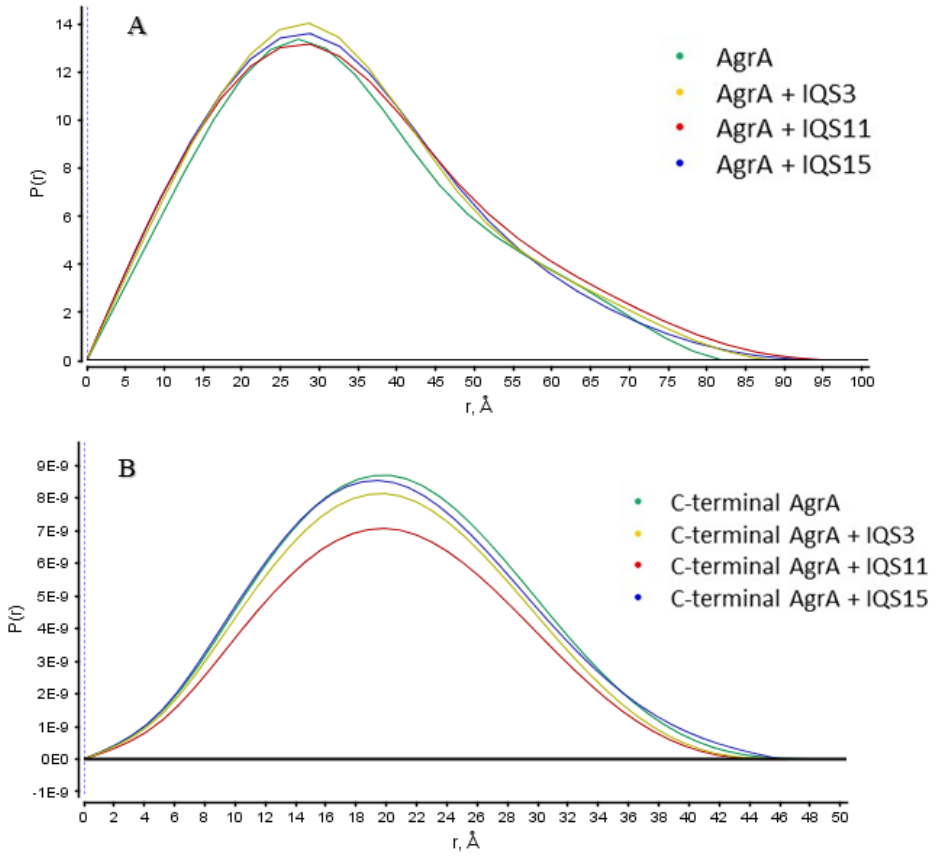


Figure 4.3.4 ***IQS analogues P(r) distribution.*** *Real space analysis of (A) full length AgrA and (B) the C-terminal domain of AgrA in presence of the IQS analogues.*

The P(r) distribution revolves around estimating a parameter called D_{max}. This represents the maximum intramolecular distance within the sample and is therefore indicative of the maximum dimension of the molecules. For this analysis, the sample is assumed to be dilute and self-dissociated. Table 4.3.2 below lists the D_{max} values extrapolated for full-length and C-terminal AgrA upon addition of the IQS analogue candidates.

Table 4.3.2 Real space analysis of AgrA and IQS analogues.		
Sample	**D_{max}**	**Score**
AgrA	91	0.897
AgrA+IQS3	95.5	0.984
AgrA+IQS11	96	1.067
AgrA+IQS15	96	1.051
C-term AgrA	48	1.081
C-term AgrA+IQS3	47	1.096
C-term AgrA+IQS11	47	1.097
C-term AgrA+IQS15	46	1.270

There is a noticeable difference in D_{max} of AgrA upon addition of IQS, while the C-terminal domain does not show significant variation.

Overall, the structural changes induced by the interaction with IQS3,11 and 15, if any, are too subtle to be detected using SAXS. The mechanism of action is therefore likely due to hindrance of the DNA-binding region rather than induced structural rearrangements. These results are to be interpreted in conjunction with Chapter 5.

4.4.Discussion

Small Angle Scattering has much unexplored potential in the field of protein structure determination. For soluble proteins, the possibility to keep the protein in solution and in a buffer that represents closely their native environment offers a truthful insight into the conformations they adopt in real life.

In this chapter, many ways to approach the determination of protein structure using small angle scattering were explored and applied to the two proteins of the *agr* signal transduction system, AgrA and $AgrC_{206}$. Soluble proteins are the ideal system for x-rays-based studies, while membrane proteins present more challenges due to the detergent or lipidic shell that surrounds them, challenges that can be addressed by using neutron scattering. A proposed pipeline for small angle scattering data analysis is described in Fig. 4.4.1.

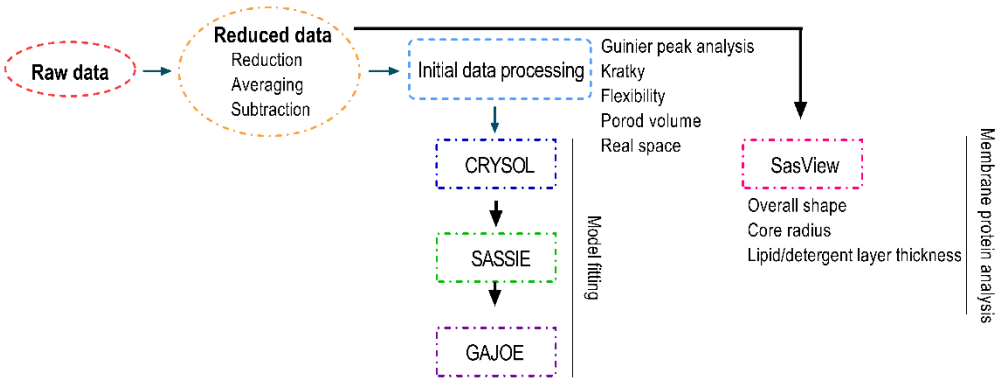


Figure 4.4.1 ***Proposed pipeline for the processing of Small Angle Scattering data of proteins.***

Our case study focuses on the proteins of the staphylococcal quorum sensing system, namely AgrC, AgrA and AgrB (see Chapter 7 for AgrB). AgrA and the cytosolic domain of AgrC, $AgrC_{206}$, are soluble, while AgrB provides a prime

example of a membrane protein whose structure cannot be easily resolved by crystallography.

AgrC was found to retain its dimeric state even in absence of its membrane component. $AgrC_{206}$ is a comparatively rigid protein which still presents a degree of flexibility in the region comprising the long helices involved in the dimerisation process. Molecular dynamics simulations were key in identifying the arrangements adopted by this protein, allowing selected scattering analysis software to identify a model based on the experimental data. There are currently two crystallographic entries in the Protein Data Bank for AgrC: Srivastava *et al.* resolved the catalytic domain (residues 278-430) in 2014 (Srivastava et al., 2014), while Xie *et al.* created a chimeric construct that attached a leucine zipper in place of the transmembrane domain and published its structure (Xie et al., 2019). The GCN4-$AgrC_{209}$ structure was compared to the scattering experimental data after being cropped to have helices of the same length as our his-tagged construct. This offered a comparatively poor fit to the data, showing how crystallographic data might not always be a true representation of proteins in solution.

The interaction between AgrC and AgrA has been previously reported (Srivastava et al., 2014), however the study aimed at providing kinetics information, therefore looking at very short timescales. Our scattering data suggest an interaction is indeed occurring at first glance, as the system appears to be monodispersed. MONSA supports this statement by providing plausible bead models of the two proteins as they are involved in the interaction. The Complex Determination function and the size exclusion chromatography data, however, do not show interaction. It is possible that while these two proteins do

find each other, the interaction is not retained over a significant period of time and the two cycle between bound and unbound states. The addition of ATP and non-hydrolysable ATP alters the interaction profile: the Complex Determination shows more encouraging results, especially in the case of ATP, while MONSA is unable to model the complex in both cases. The two proteins do not come off the size exclusion chromatography column together, which supports the hypothesis of a weak, transient complex, if indeed a complex is formed as suggested by the literature.

Neutron scattering can provide further insight into this interaction. Contrast matching each of the two components allows to observe variations in their tertiary structure as conditions vary while they are interacting with each other. This approach revealed that the addition of both ATP and non-hydrolysable ATP triggers conformational rearrangements in AgrA, in presence of $AgrC_{206}$. According to the literature, the response regulator AgrA is a phosphor-acceptor that requires AgrC, the phosphodonor, to acquire the phosphate from ATP. There is no evidence that AgrA has any kinase activity, which suggests that these rearrangements are triggered by some activity related to $AgrC_{206}$. The latter does not show any conformational changes upon addition of nucleotides, which is consistent with the SAXS data of $AgrC_{206}$ alone. The fully protonated contrast, however, shows variations between the mixture containing the non-hydrolysable ATP and the other two. This contrast, which preferentially shows scattering from the deuterated component but partially also from the protonated AgrA, suggests that the presence of non-hydrolysable ATP triggers a change in the globularity of the system as well as in the overall conformation. This is in contrast with the SAXS data for the same sample and could be explained by attributing the

changes to AgrC, which may be masked when not contrast matched, if the contrast with 42% D_2O showed consistent variations.

5. Thermal denaturation as a tool to probe molecular interactions in the *agr* two-component system

This Chapter describes the observations and quantitation of molecular interactions through quantitative analysis of protein structural (by CD) and conformational (by SAXS) stability to thermal denaturation. One of the main applications of thermal denaturation studies is to use thermal physical parameters to identify molecular interactions. One of the most popular methods used in response to the growing interest towards the thermal stability and unfolding of proteins, first developed in the 1960s and 70s, is differential scanning calorimetry (DSC). This technique is useful in determining the melting temperature and the enthalpy of the denaturation process (Johnson, 2013). The calculation of enthalpy, however, is heavily dependent on the assumptions that the estimated protein concentration is correct and the sample is completely homogeneous. The DSC data analysis also relies on the reversibility of the transition, which is often not the case in protein thermal studies. For this reason, most of the DSC studies, especially in the earlier days, were carried out using buffers at highly acidic pH (3-5), which are of course far from the native conditions of proteins (Schön et al., 2017, Johnson, 2013). Another, more commonly used technique is isothermal titration calorimetry (ITC), which measures released or absorbed heat at a constant temperature. This technique is more focused on molecular interactions than protein characterisation *per se* (Johnson, 2013, Ladbury, 2004).

Neither of these techniques allow to extrapolate details on the secondary structure of proteins. DSC analysis estimates the melting temperature but is

subjected to a series of assumptions as discussed above. Circular Dichroism, on the other hand, allows to follow the protein degradation process in terms of temperature-dependent secondary structure variation and to extrapolate the melting temperature.

The Boltzmann analysis is frequently used to quantify the thermal denaturation of proteins. It allows to calculate the melting temperature based on the acquired CD spectrum but is based on the average of folded and unfolded states. The Landau series expansion analysis (Hussain et al., 2018) is proposed as a tool not only to quantify the thermal stability of proteins, but also to identify molecular interactions by establishing the effect that potential interaction partners have on the cooperativity of the melting transition.

This protocol was developed to complement the Boltzmann analysis on CD data and used to investigate the role of ATP in the processes of the *agr* two-component system, AgrC and AgrA. The effect of potential AgrA inhibitors, the IQS analogues, on the protein structure and thermal stability is also investigated.

While the current methodology uses the thermal response of the protein secondary structure, the Landau series expansion analysis was adapted to allow the analysis of protein thermal melts by Small Angle Scattering by looking at the disruption of tertiary structure.

5.1. The role of ATP in the molecular interaction of AgrA and AgrC

ATP is required in the regulation of the *agr* system as it drives its upregulation by ultimately promoting dimerisation of AgrA. This study aims at characterising its involvement in the interaction between AgrA and AgrC *in vitro* by investigating the effect of ATP and non-hydrolysable ATP on the thermal stability of the two proteins. These results are directly linked to findings presented in Chapter 4.

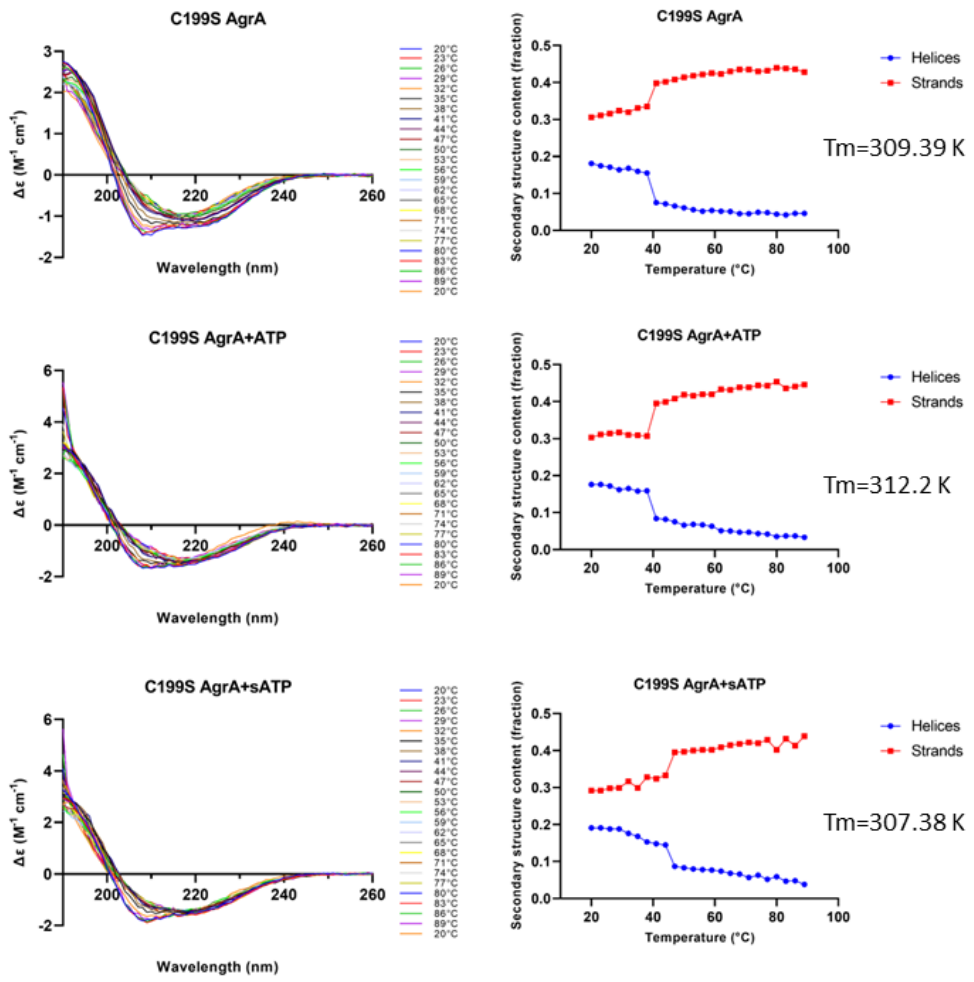


Figure 5.1.1 ***CD spectra and secondary structure of AgrA.*** *Far UV CD spectra (left) and secondary structure components (right) of C199S AgrA alone (top), with ATP (middle) and with non-hydrolysable ATP (bottom) as a function of temperature.*

The CD measurements were processed using an Average Amino Acid Molecular Weight of 113 and the single wavelength chosen for the Boltzmann analysis was 208 nm. The CD spectra and secondary structure plots (Fig. 5.1.1) reveal that the secondary structure profile of C199S AgrA is not altered by the addition of ATP or non-hydrolysable ATP. There is a similar drop in the helical content with consequent increase in β-strands across the three datasets. The melting

temperature, however, is affected, with the ATP providing a stabilising effect and increasing it by ~3 K and the sATP having the opposite effect and decreasing it by 2 K.

The Landau approach was applied to investigate potential changes to the cooperativity of the transition upon nucleotide addition. An increase in cooperativity is indicative of a system more stable against thermal denaturation, while a decrease indicates that the interaction partner has a destabilising effect on the system. Both suggest molecular interaction, while no variation indicates either that no binding is taking place or that the interaction does not alter the secondary structure or the stability of the protein.

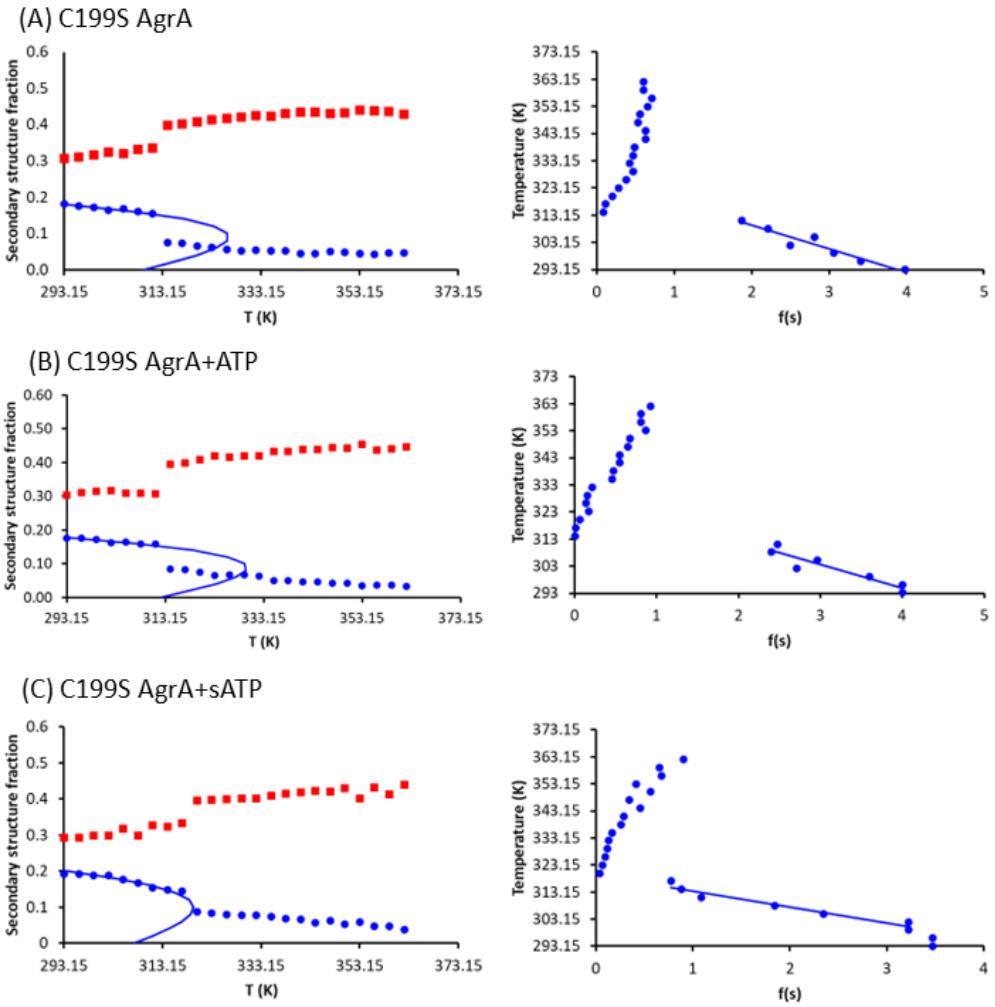


Figure 5.1.2 ***Landau analysis of the thermal unfolding of AgrA.*** *Secondary structure content with equation of state (left) and linear fits of T vs f(s) (right) of C199S AgrA alone (A), with ATP (B) and with sATP (C).*

The cooperativity of the transition is not altered by adding ATP, as seen in the respective equation of states which are nearly identical. The addition of non-hydrolysable ATP noticeably alters the transition making it less cooperative, which suggests that the sATP decreases the stability of the protein to temperature (Fig. 5.1.2).

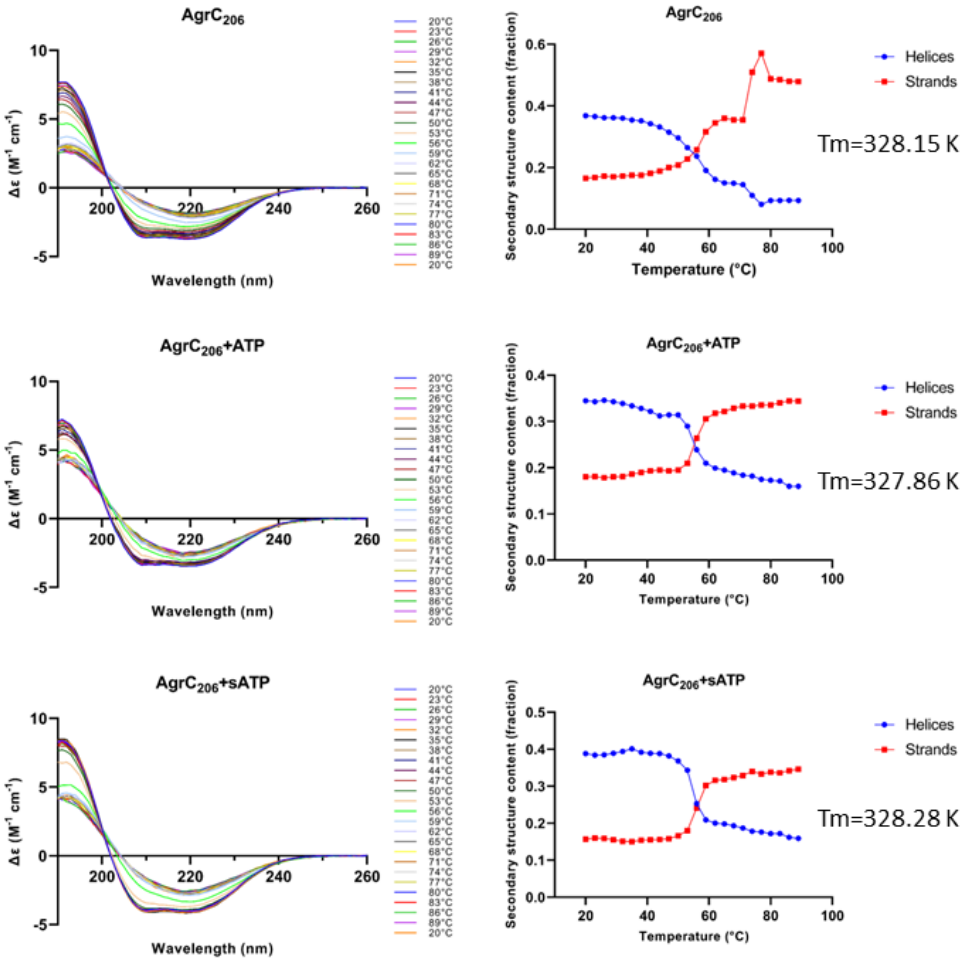


Figure 5.1.3. ***CD spectra and secondary structure of AgrC.*** *Far UV CD spectra (left) and secondary structure components (right) of $AgrC_{206}$ alone (top), with ATP (middle) and with non-hydrolysable ATP (bottom) as a function of temperature.*

According to the literature, AgrC is the actual interaction partner of ATP and uptakes the phosphate group thanks to its kinase activity. The transition temperature of $AgrC_{206}$ is constant across the datasets (Fig. 5.1.3). The drop in the helical content is more abrupt when ATP and sATP are added to the protein, but these dynamics are better visualised in the Landau analysis (Fig. 5.1.4).

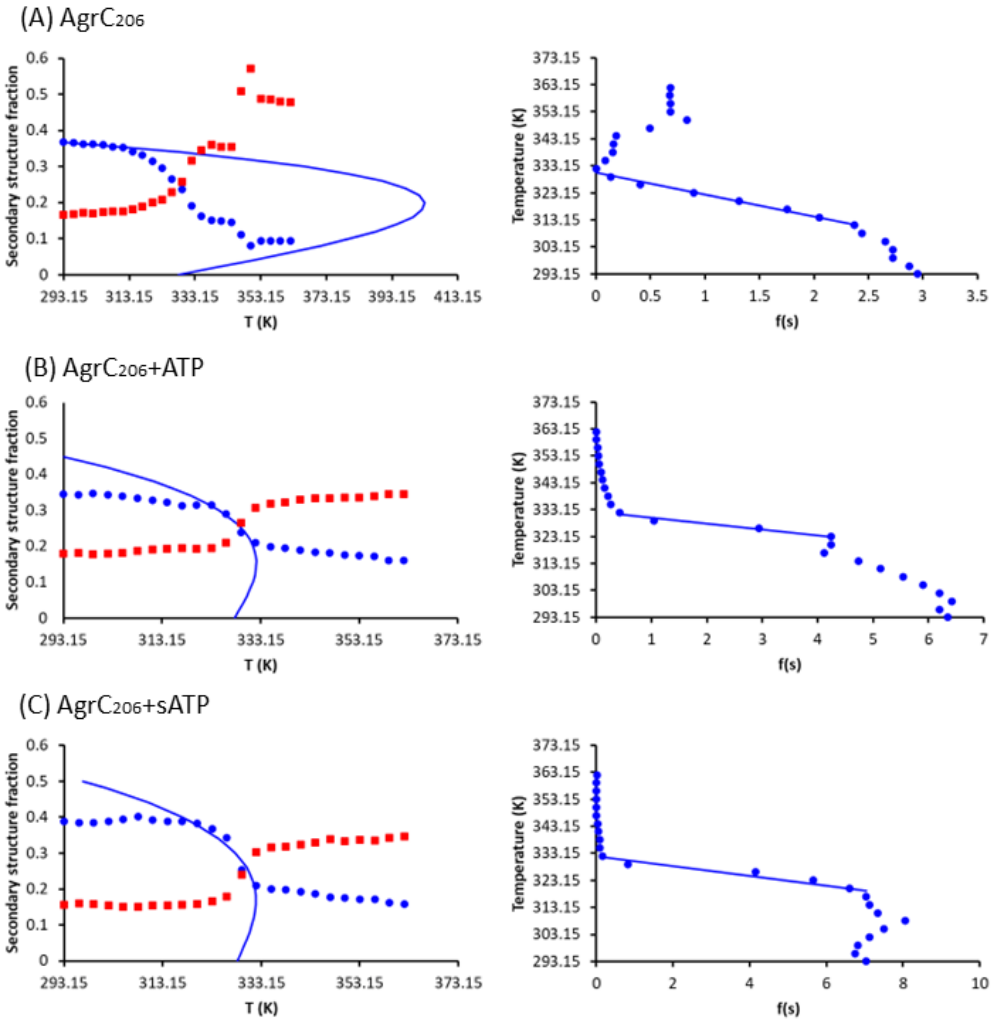


Figure 5.1.4 ***Landau analysis of the thermal unfolding of AgrC.*** *Secondary structure content with equation of state (left) and linear fits of T vs f(s) (right) of AgrC$_{206}$ alone (A), with ATP (B) and with sATP (C).*

AgrC$_{206}$ exhibits a very cooperative transition, which relates to its high temperature stability. This feature is noticeable already from its high melting temperature at 55 °C. The addition of both nucleotides makes the protein more prone to thermal denaturation, as shown by the narrowing of $T_+ - T_m$. This suggests that both analogues interact with AgrC and alter its thermal stability (Fig. 5.1.4).

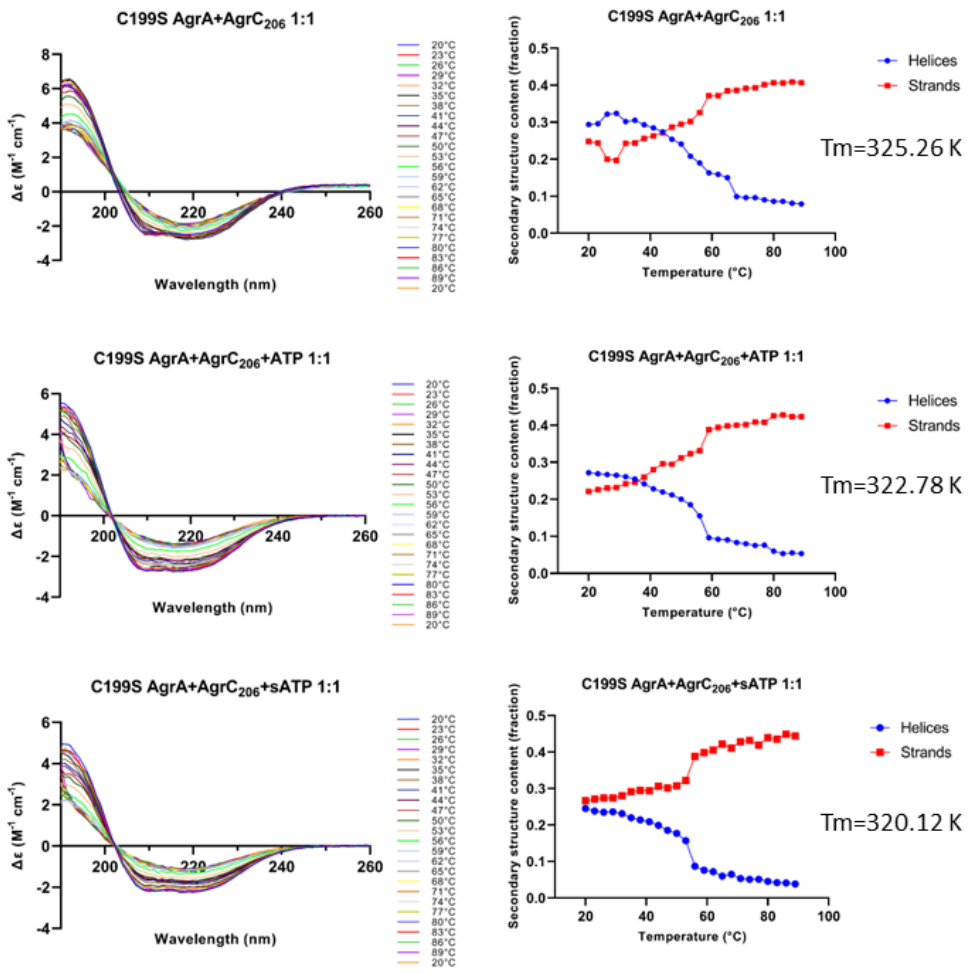


Figure 5.1.5 ***CD spectra and secondary structure of AgrA+AgrC (1:1).*** *Far UV CD spectra (left) and secondary structure components (right) of the AgrA/AgrC$_{206}$ mixture (1:1 molar ratio) alone (top), with ATP (middle) and with non-hydrolysable ATP (bottom) as a function of temperature.*

The AgrA/AgrC$_{206}$ mixture in a 1:1 molar ratio exhibits a melting temperature of 325.3 K, nearly as high as AgrC$_{206}$ alone which is a very stable protein. The transition here is very gradual. The addition of ATP significantly lowers the melting temperature, bringing it to 322.8 K. The non-hydrolysable ATP alters the starting secondary structure content of the system by increasing the strand content and consequently reducing the helical fraction. This system also exhibits

a lower melting point, 5 °C less than the equivalent in absence of nucleotides (Fig 5.1.5).

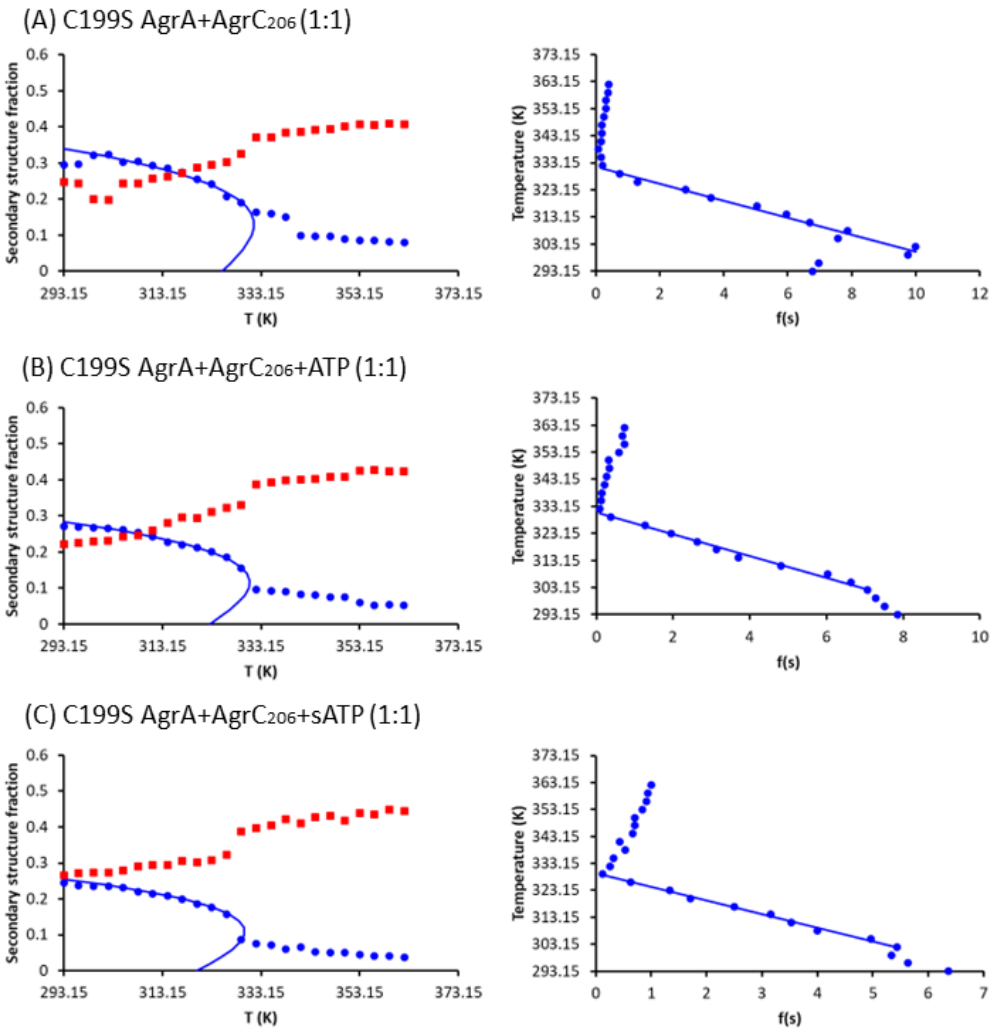


Figure 5.1.6 ***Landau analysis of AgrA and AgrC in 1:1 molar ratio.*** *Secondary structure content with equation of state (left) and linear fits of T vs f(s) (right) of the AgrA and AgrC$_{206}$ mixture in a 1:1 molar ratio (A), with ATP (B) and with sATP (C).*

The visual inspection of the equations of state does not suggest any major alteration in the cooperativity of the transition upon addition of nucleotides (Fig.

5.1.6). Any difference is better visualised in Figure 5.1.9, where the τ values are plotted.

Due to the dimeric state of $AgrC_{206}$ (see Chapter 4), a molar ratio of 1:2 (AgrA:AgrC) was investigated, as the mechanism of AgrA recruitment to AgrC is not yet characterised.

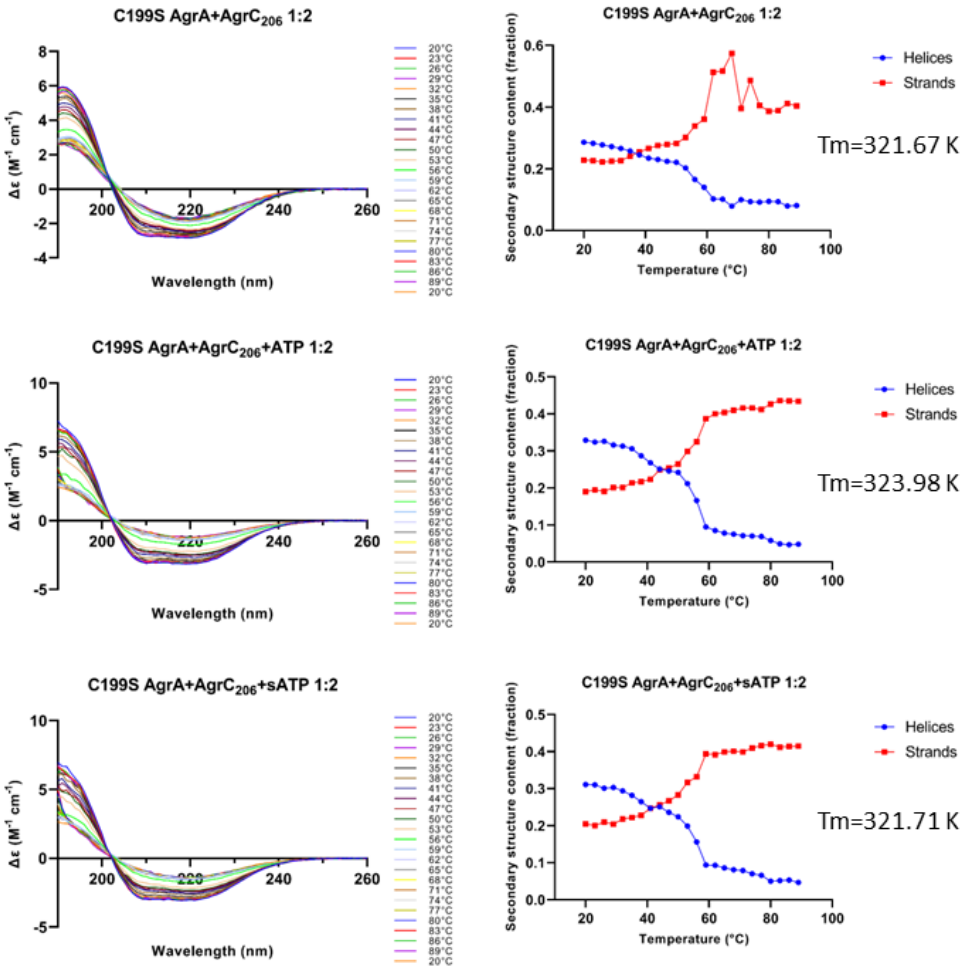


Figure 5.1.7 ***CD spectra and secondary structure of AgrA+AgrC (1:2).*** *Far UV CD spectra (left) and secondary structure components (right) of the AgrA/$AgrC_{206}$ mixture (1:2 molar ratio) alone (top), with ATP (middle) and with non-hydrolysable ATP (bottom) as a function of temperature.*

Despite more AgrC present in the system, the melting point of the mixture alone is lower than its 1:1 ratio equivalent. It is not clear whether the behaviour of the β-strands at ~65-70 °C is real or an artifact. In this molar ratio ATP has a stabilising effect, raising the melting temperature by 2 K, while the sATP does not alter the dynamics of the system. It is interesting that the effect of ATP, which is stabilising in this case, is reversed in the 1:1 molar ratio system; likewise, the destabilising influence of sATP is not present in the 1:2 ratio system.

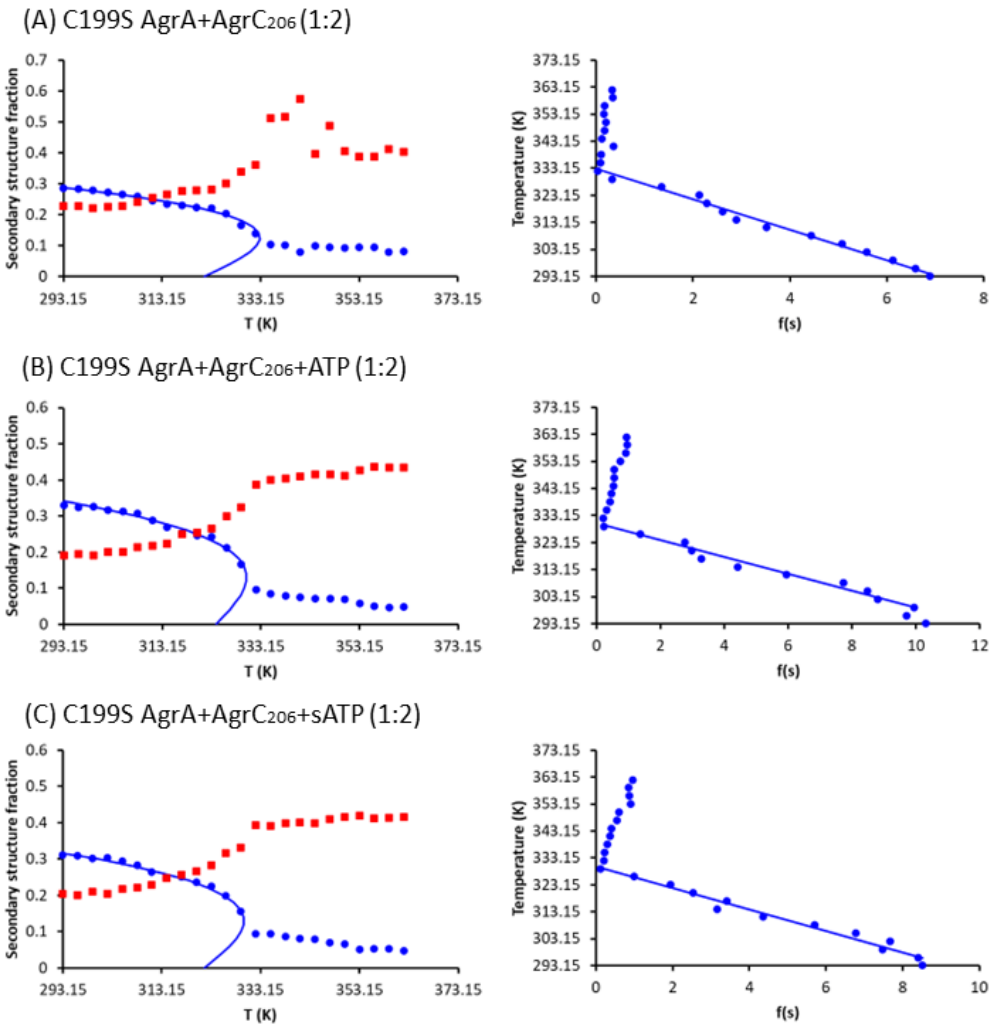


Figure 5.1.8 ***Landau analysis of AgrA and AgrC in 1:2 molar ratio.*** *Secondary structure content with equation of state (left) and linear fits of T vs f(s) (right) of the AgrA and AgrC$_{206}$ mixture in a 1:2 molar ratio (A), with ATP (B) and with sATP (C).*

The Landau analysis of this system reveals that both ATP and sATP have in fact a destabilising effect on the system, making it more prone to thermal unfolding (Fig. 5.1.8). This is a good example of the usefulness of this type of analysis as opposed to the standard Boltzmann analysis.

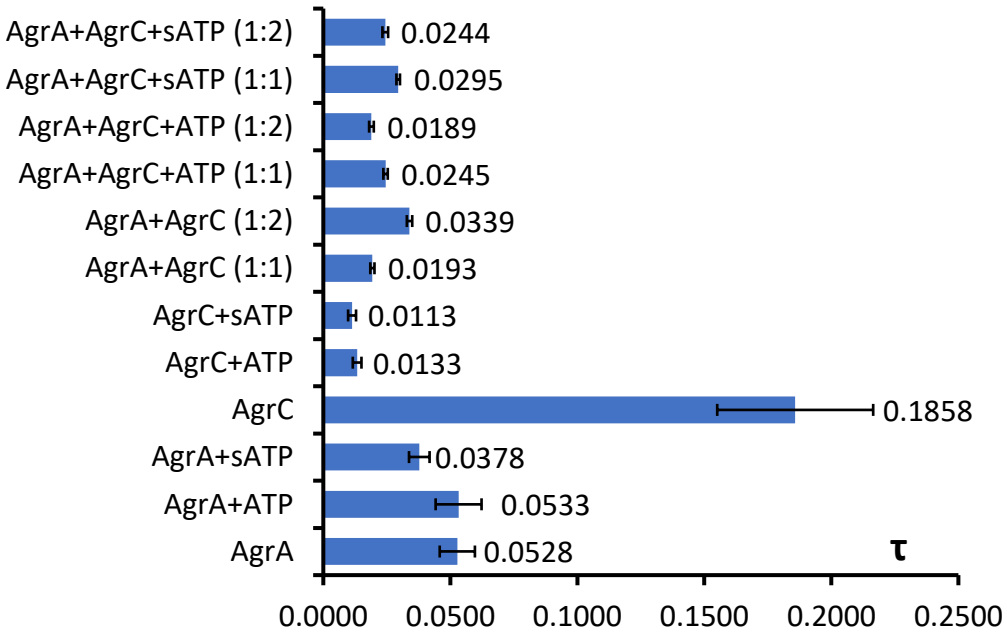


Figure 5.1.9 ***Transition cooperativity.*** *The τ values, corresponding to the cooperativity of the transition, are plotted for comparison. AgrC shows great stability to thermal unfolding, which is dramatically altered by addition of nucleotides.*

The plotting of τ values for each protein (Fig. 5.1.9) enables comparison of the behaviours of the various systems effectively. The temperature stability of $AgrC_{206}$ is confirmed, while the addition of nucleotides greatly destabilises its secondary structure response. $AgrC_{206}$ in presence of ATP and sATP shows in fact the lowest cooperativity in this study. AgrA is characterised by a comparatively low thermal stability, as highlighted by both the melting point identified by the Boltzmann analysis and the Landau analysis. The addition of ATP does not alter the AgrA profile, while the sATP lowers the cooperativity value and destabilises the protein's secondary structure. This suggests a possible molecular interaction between AgrA and sATP.

All of the AgrA/AgrC mixtures have a cooperativity value lower than either individual protein. While this suggests very thermally unstable systems, it is also indicative that some molecular interaction is occurring. When looking at the AgrA/AgrC mixture in a 1:1 ratio, the presence of both ATP and sATP has a stabilising effect on the system, raising τ from 0.0193 to 0.0245 and 0.0295 respectively.

The 1:2 molar ratio has a noticeably higher cooperativity value, at 0.0339. This can either suggest that the two proteins do in fact interact preferentially as an AgrC dimer and AgrA monomer, making the system more thermally stable, or that the higher presence of AgrC, which is inherently very stable, raises the overall cooperativity of the mixture without implying a preferential molecular interaction. The addition of ATP and sATP significantly weakens the transition, especially in the case of ATP.

When the Landau analysis was carried out, visual inspection of both the secondary structure content and the linear fits of some of these systems allowed to identify two separate transitions, one close to the identified melting point and one at lower temperatures. The melting temperature of the first transition was identified by carrying out a manual Botzmann fit on CDApps, as shown in Figure 5.1.10 and 5.1.11.

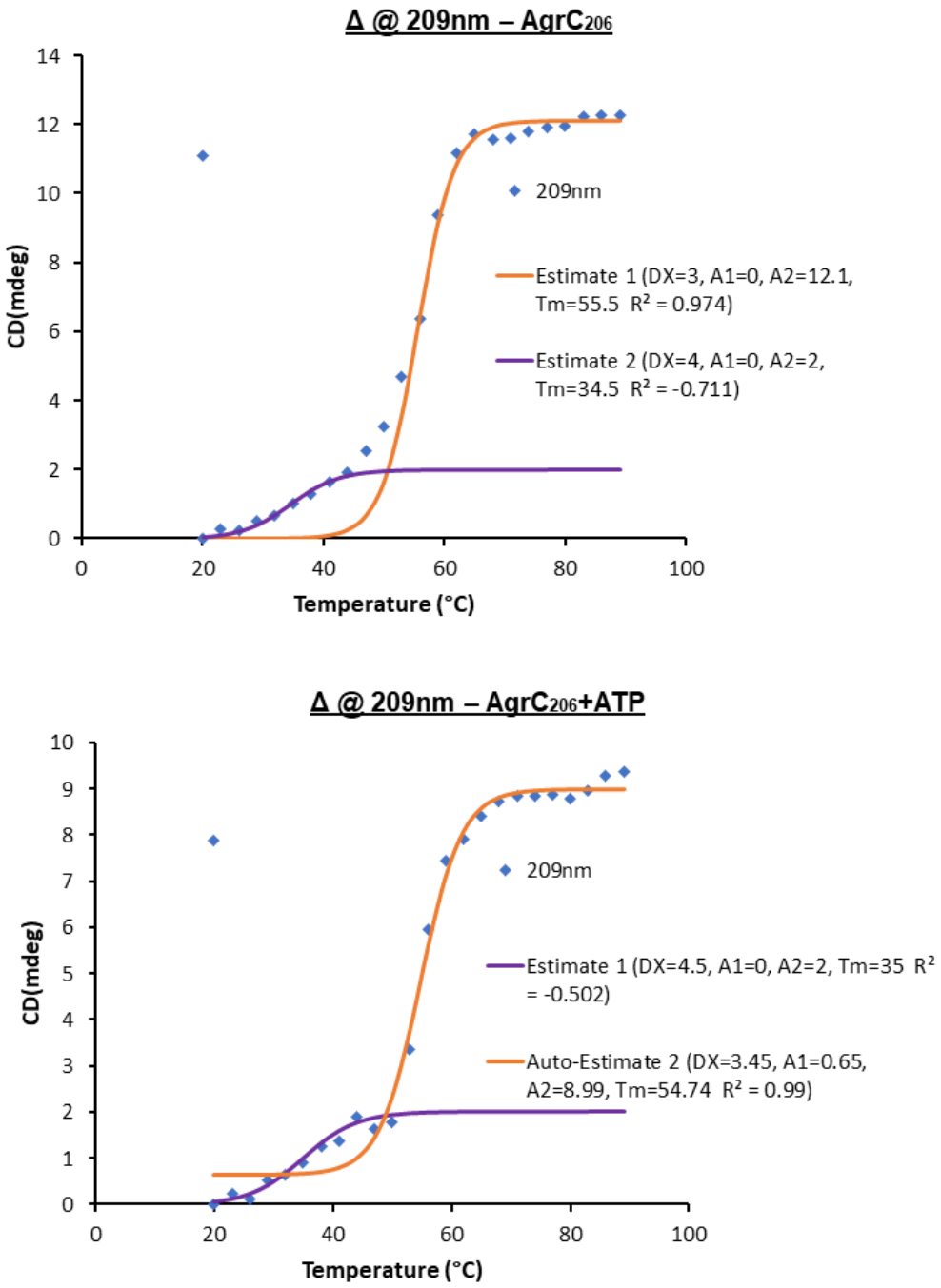


Figure 5.1.10 ***Manual Boltzmann fit of the double transition in AgrC.*** *Manual fit of the Boltzmann curve on the two transitions observed in $AgrC_{206}$ (top) and $AgrC_{206}$ + ATP (bottom). The low transition temperatures are identified at 34.5 and 35 °C, and the high melting temperatures at 55.5 and 54.74 °C respectively.*

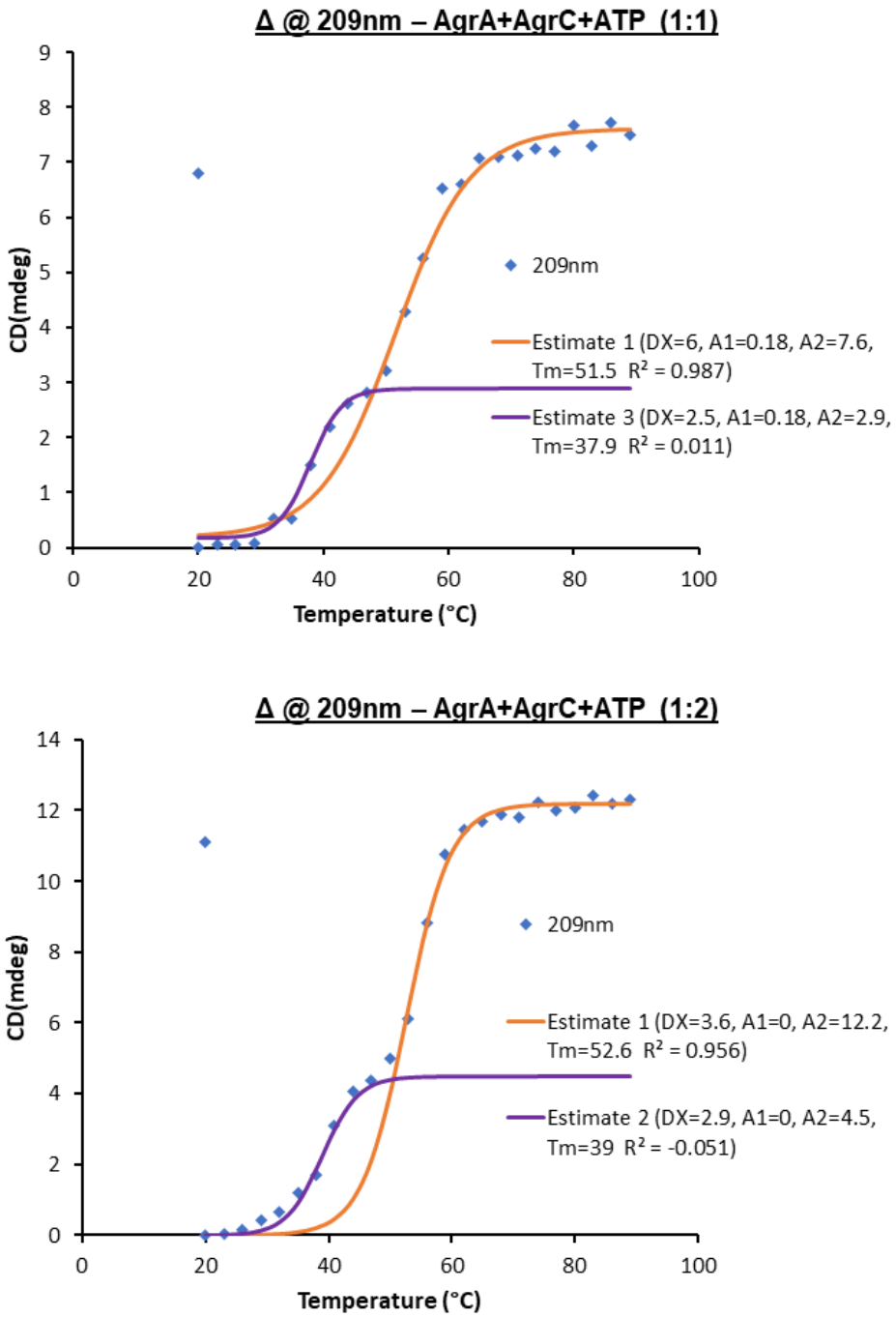


Figure 5.1.11 ***Manual Boltzmann fit of the double transition in AgrA+AgrC+ATP.*** *Manual fit of the Boltzmann curve on the two transitions observed in AgrA and AgrC in a 1:1 ratio (top) and 1:2 ratio (bottom), with ATP. The low transition temperatures are identified at 37.9 and 39 °C, and the high melting temperatures at 51.5 and 52.6 °C respectively.*

The low and high melting temperatures were used to carry out the double Landau analysis, corresponding to the low temperature phase and the high temperature phase. Both equations of state are plotted (Fig. 5.1.12-14) and two separate cooperativity values are calculated.

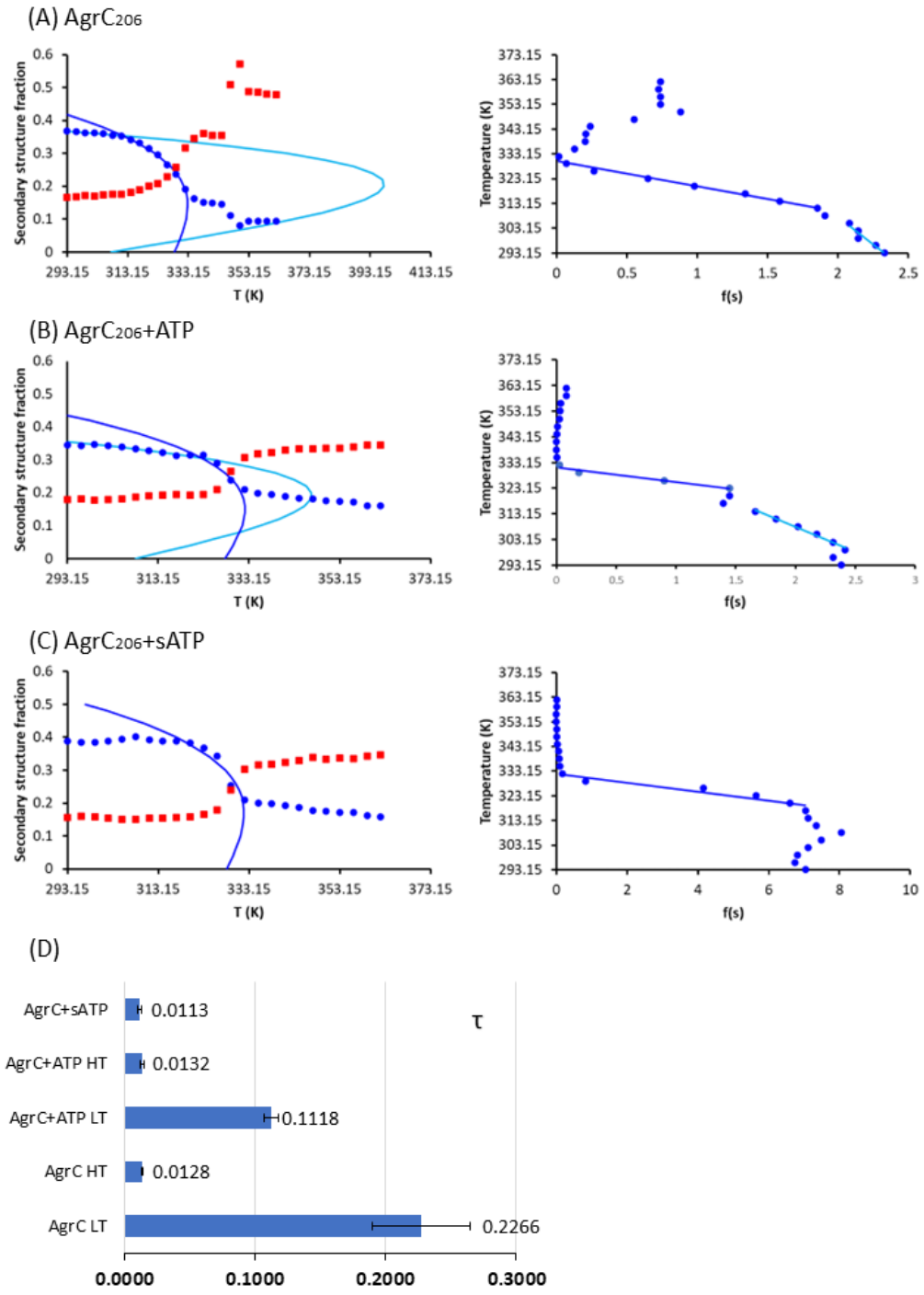


Figure 5.1.12 ***Landau analysis of AgrC.*** *Secondary structure content with equation of state (left) and linear fits of T vs f(s) (right) of* $AgrC_{206}$ *alone (A), with ATP (B) and with sATP (C); (D) transition cooperativity plot. Two transitions are highlighted in (A) and (B), one at low temperature (LT) and one at higher temperature (HT).*

$AgrC_{206}$, the most thermally stable component of this system, shows a two-step transition in the behaviour of the secondary structure components, although the abrupt variation around 70 °C could be an artifact of the measurements. Nonetheless, when separated the low temperature transition exhibits a much higher cooperativity, as expected for this protein, with a markedly lower cooperativity in the high temperature region. A similar behaviour is observed upon addition of ATP, although with a less pronounced difference. In this case, the presence of the two transitions is more evident in the linear fit plot. The addition of non-hydrolysable ATP alters the behaviour of the protein, which is now characterised by a single, weak transition (Fig. 5.1.12). The behaviour of the system upon addition of both nucleotides suggests a molecular interaction with the protein, manifesting in reduced thermal stability. This is in agreement with what observed in the original analysis.

The presence of a double transition is indicative of a protein that is very stable at low temperatures and has a very cooperative transition. After the first transition temperature the system becomes more susceptible to thermal unfolding, manifesting in a less cooperative, more gradual transition. The disappearance (or appearance) of a second transition can imply that either the system is more susceptible to temperature overall, or the system is indeed more stable overall as it is only affected by high temperatures.

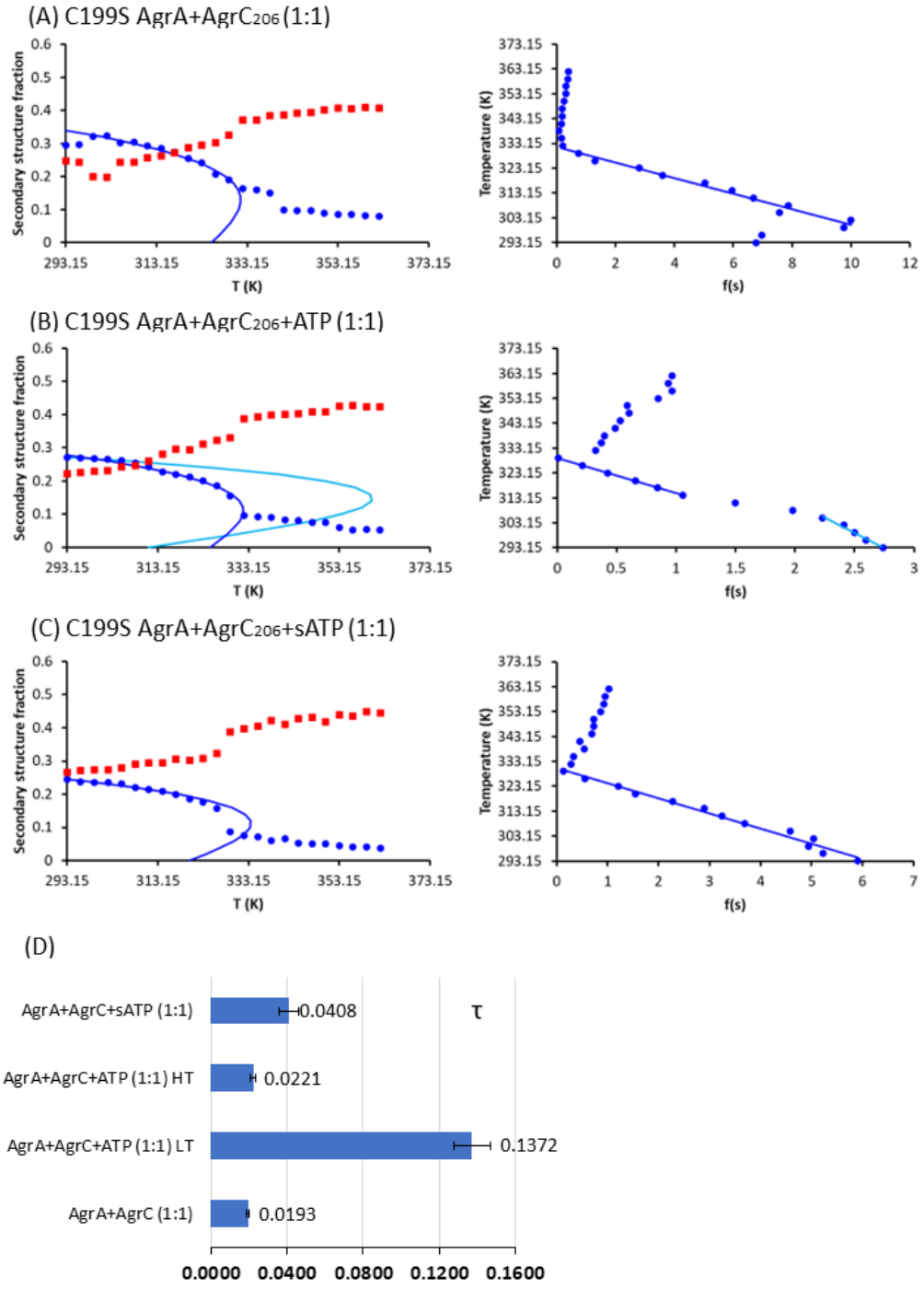


Figure 5.1.13 ***Landau analysis of AgrA and AgrC in 1:1 ratio.*** *Secondary structure content with equation of state (left) and linear fits of T vs f(s) (right) of the AgrA and AgrC$_{206}$ mixture in a 1:1 molar ratio (A), with ATP (B) and with sATP (C); (D) transition cooperativity plot. Two transitions are resolved in (B), one at low temperature (LT) and one at higher temperature (HT).*

The mixture of the two proteins in a 1:1 molar ratio exhibits a single, well-defined transition. This is true for the system in presence of sATP, while a second transition characterises this system in presence of ATP (Fig. 5.1.13). The non-hydrolysable ATP triggers an increase in the cooperativity of the system's unfolding. ATP, on the other hand, stabilises the system at low temperatures, resulting in two transitions. The cooperativity of the transition at high temperature is similar to the one of the system alone, while the low temperature transition is very cooperative.

The system in 1:2 molar ratio (Fig. 5.1.14) also exhibits a double transition upon addition of ATP, which is absent in the other two cases. While the low temperature transition is more cooperative than the single transition of the system without nucleotides, the high temperature transition shows a decrease in cooperativity. This means that, while the system is more thermally stable at low temperatures, once the first threshold is reached the system becomes more unstable.

The addition of sATP marginally destabilises the system, which shows a 0.0061 decrease in cooperativity.

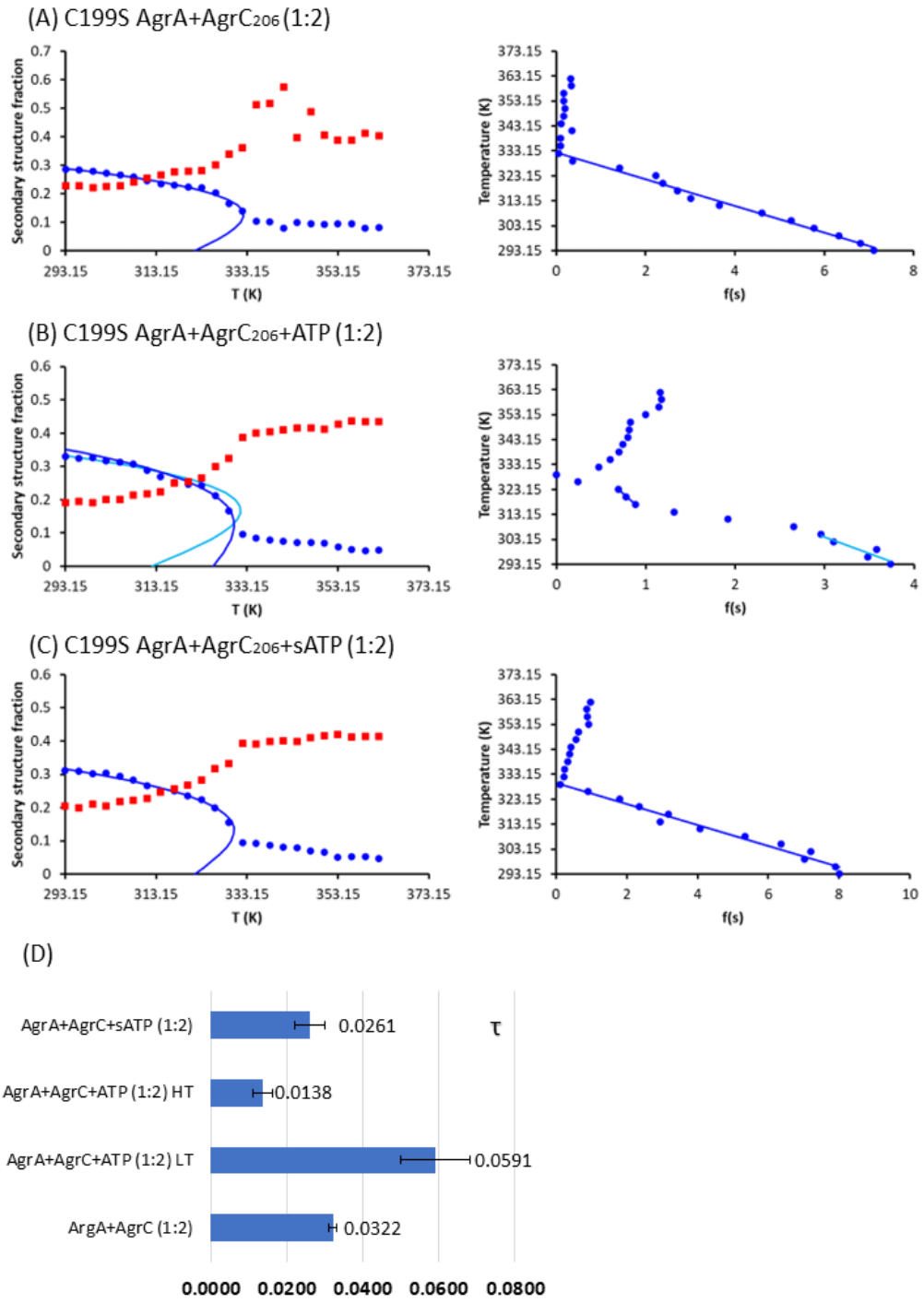


Figure 5.1.14 ***Landau analysis of AgrA and AgrC in 1:2 ratio.*** *Secondary structure content with equation of state (left) and linear fits of T vs f(s) (right) of the AgrA and AgrC$_{206}$ mixture in a 1:2 molar ratio (A), with ATP (B) and with sATP (C); (D) transition cooperativity plot. Two transitions are highlighted in (B), one at low temperature (LT) and one at higher temperature (HT).*

The distribution of τ values in Figures 5.1.13 and 5.1.14 is very similar.

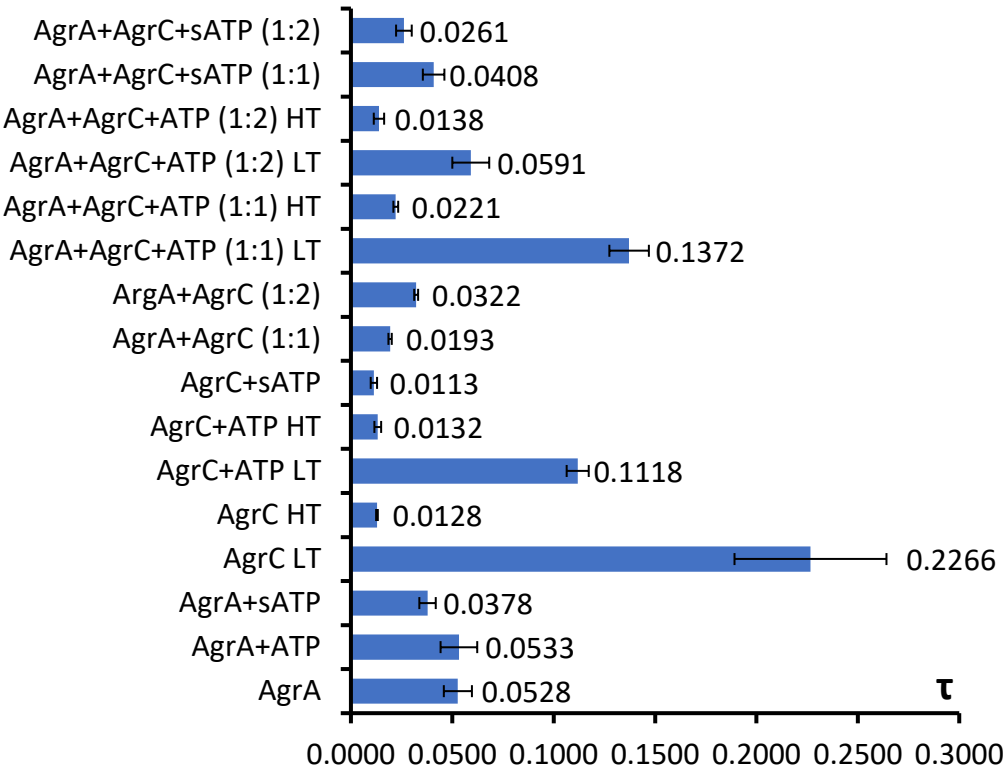


Figure 5.1.15 ***Transition cooperativity.*** *The τ values, corresponding to the cooperativity of the transition, are plotted for comparison. For the systems which exhibit two transitions, these are plotted separately and identified with LT (low temperature transition) and HT (high temperature transition).*

Figure 5.1.15 shows the τ values of all the various systems studied in a single plot, to better contextualise the cooperativity of each system against the others.

5.2Effect of the IQS analogues on the thermal stability of AgrA

The study of the IQS analogues, the potential inhibitors of AgrA, yielded more subtle results. IQS is a class of compounds which is suggested to act by preventing the interaction of AgrA with the DNA promoters, either by physically obstructing the interaction site or by altering the structure of the protein around the binding site. CD was therefore used to investigate the potential interaction and effect of the IQS analogues 3, 11 and 15 on the thermal stability of both full-length AgrA and C-terminal AgrA, the DNA-binding domain.

Figure 5.2.1 shows the helical and strand content of AgrA at room temperature alone and when mixed with IQS3, 11 and 15. For the most part the differences are marginal and, one could argue, within the error margins. Among the three analogues studied, IQS11 elicits the biggest structural change in full-length AgrA at room temperature, causing an increase in the helical content and the corresponding decrease in β-strands.

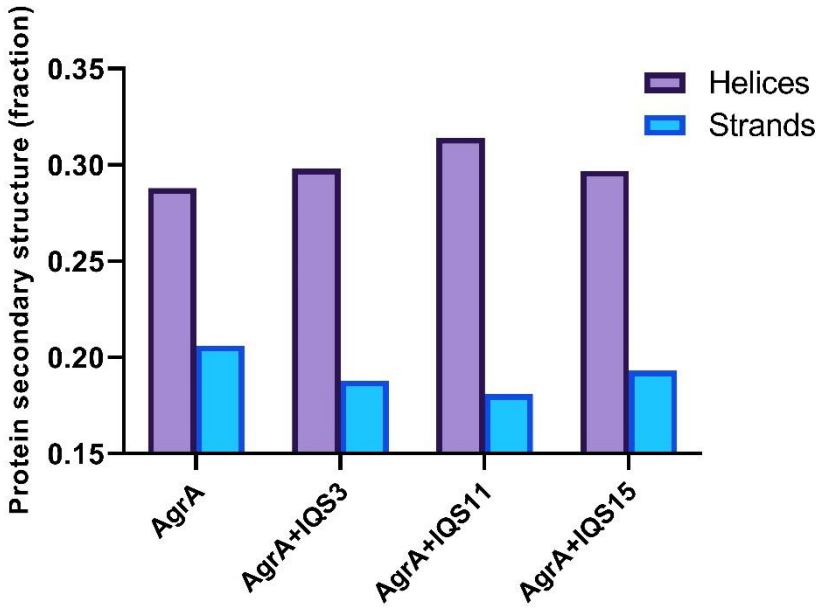


Figure 5.2.1 ***Secondary structure content of AgrA with and without IQS analogues at 20 °C.*** *The α-helical and β-strand content of AgrA alone is shown and compared to AgrA in presence of potential inhibitors IQS3, 11 and 15.*

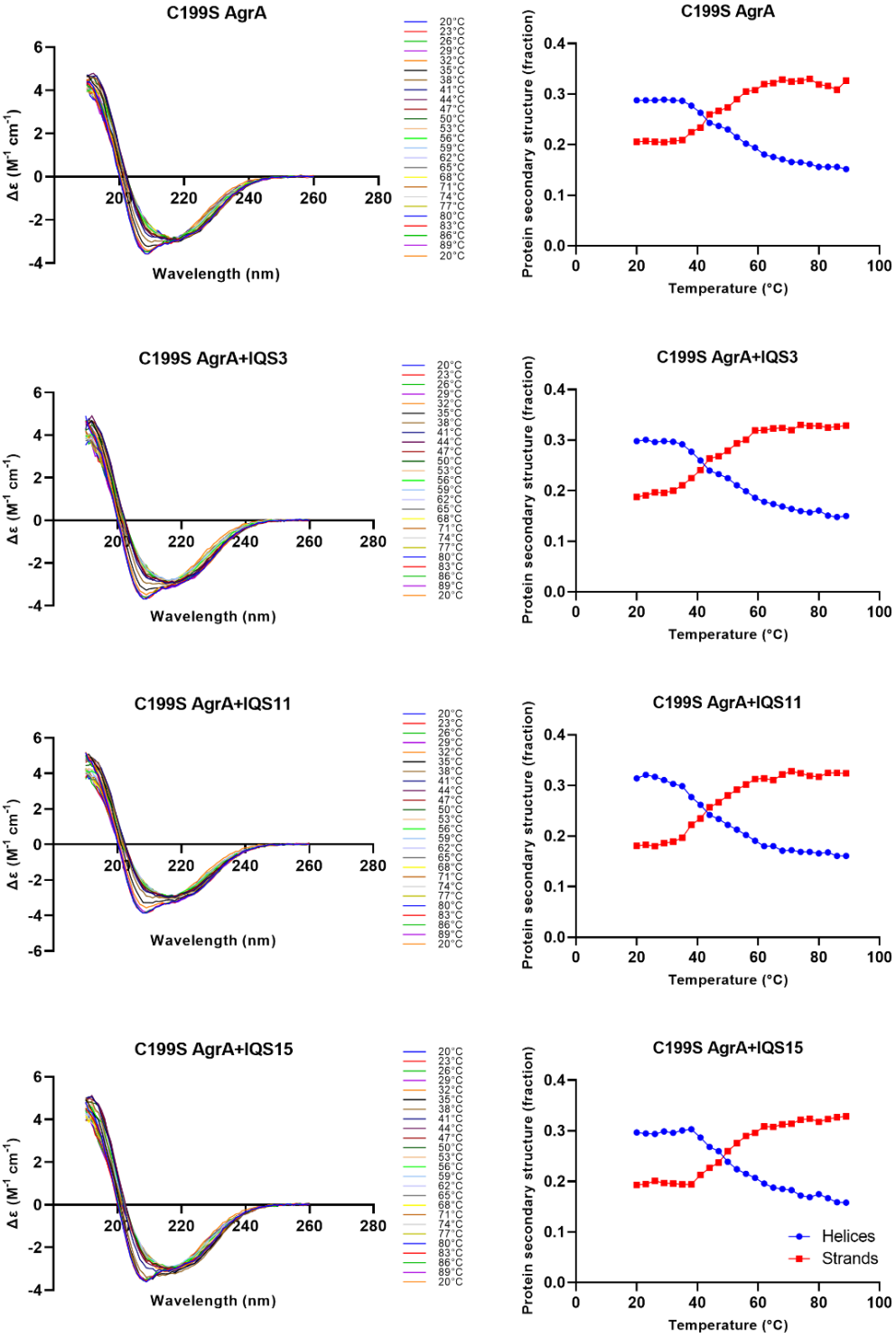


Figure 5.2.2 ***Far UV CD spectra and secondary structure content of AgrA with IQS analogues as a function of temperature.*** *The far UV CD spectra are shown on the left and the secondary structure content (helical in blue and strands in red) on the right, of AgrA alone and in presence of IQS analogues 3, 11 and 15.*

The amyloidogenic conversion of the predominantly helical AgrA is shown in the β-strand content increase. Obtained using the CONTINLL SMP56, SP43+13 Membrane SSE Fit algorithm.

Figure 5.2.2 shows the CD spectra resulting from thermal denaturation of full-length AgrA and the IQS analogues and the changes in the secondary structure content of full-length AgrA upon addition of the three IQS analogues. The data processing was done following the CD signal at wavelength 208 nm and using an Average Amino Acid Molecular Weight of 118. All samples show similar trends in the helical decrease and sheet increase. The addition of IQS11 appears to make the decrease in helical content more gradual compared to the other samples, in particular IQS15.

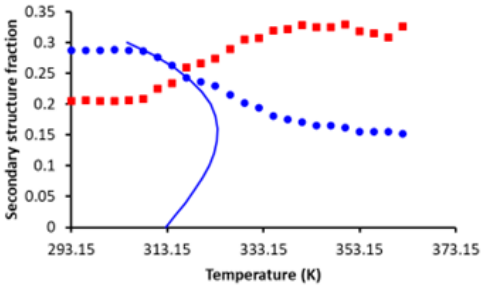
The melting temperatures extrapolated from CD are indicated in Table 5.2.1.

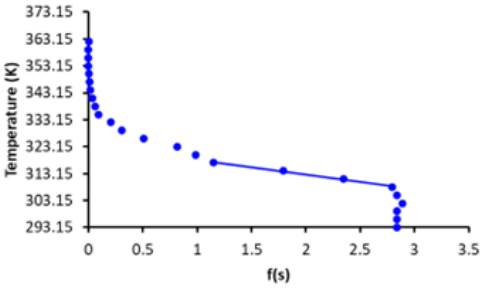
Table 5.2.1 Melting temperatures of AgrA with IQS analogues.

Sample	**Temperatures (°C)**
AgrA	39.57
AgrA + IQS3	38.41
AgrA + IQS11	36.62
AgrA + IQS15	41.22

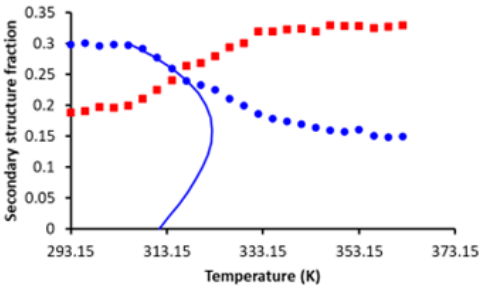
The changes in the melting temperature elicited by the analogues are quite subtle. The biggest variation, 2.95 °C, is seen in presence of IQS11.

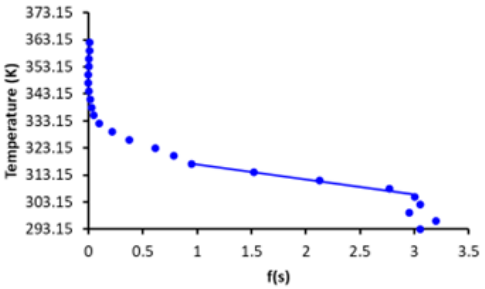
(A) C199S AgrA



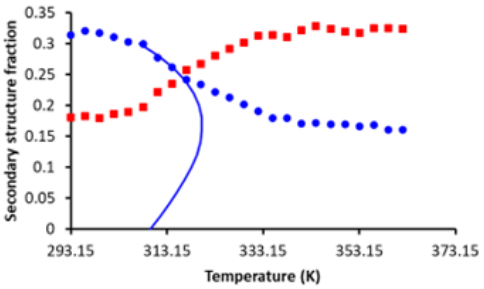


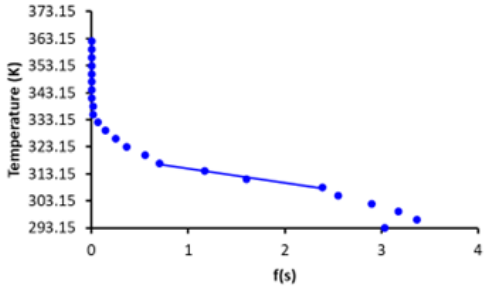
(B) C199S AgrA+IQS3



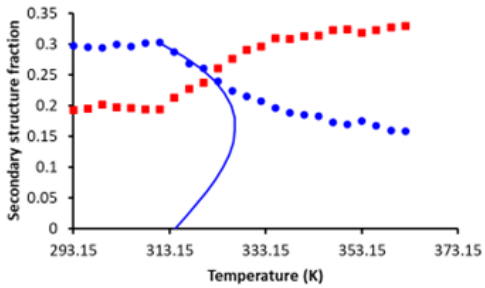


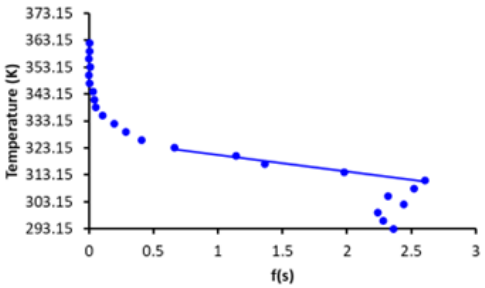
(C) C199S AgrA+IQS11





(D) C199S AgrA+IQS15





(E)

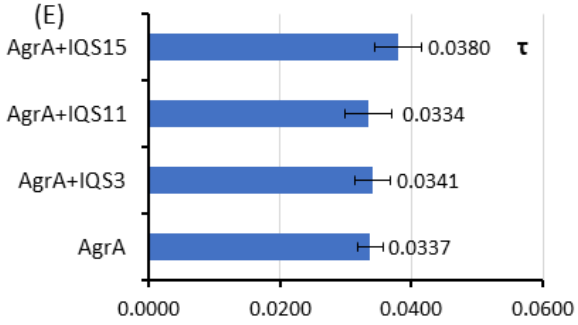


Figure 5.2.3 ***Landau analysis of AgrA with IQS analogues.*** *Secondary structure content with equation of state (left) and linear fits of T vs f(s) (right) of C199S AgrA alone (A), with IQS3 (B), with IQS11 (C) and with IQS15 (D); transition cooperativity plot (E).*

The Landau analysis reveals that IQS15 triggers an increase in the cooperativity of the transition of AgrA, making it more resistant to thermal denaturation. To note is that this difference is only 0.004 and therefore still quite small. Neither IQS3 nor 11 have an effect (Fig. 5.2.3).

The working hypothesis is that the AgrA inhibition by these compounds is related to the DNA binding site, located on the C-terminus. The potential effect of the analogues on the secondary structure and thermal stability of the C-terminal AgrA was therefore investigated.

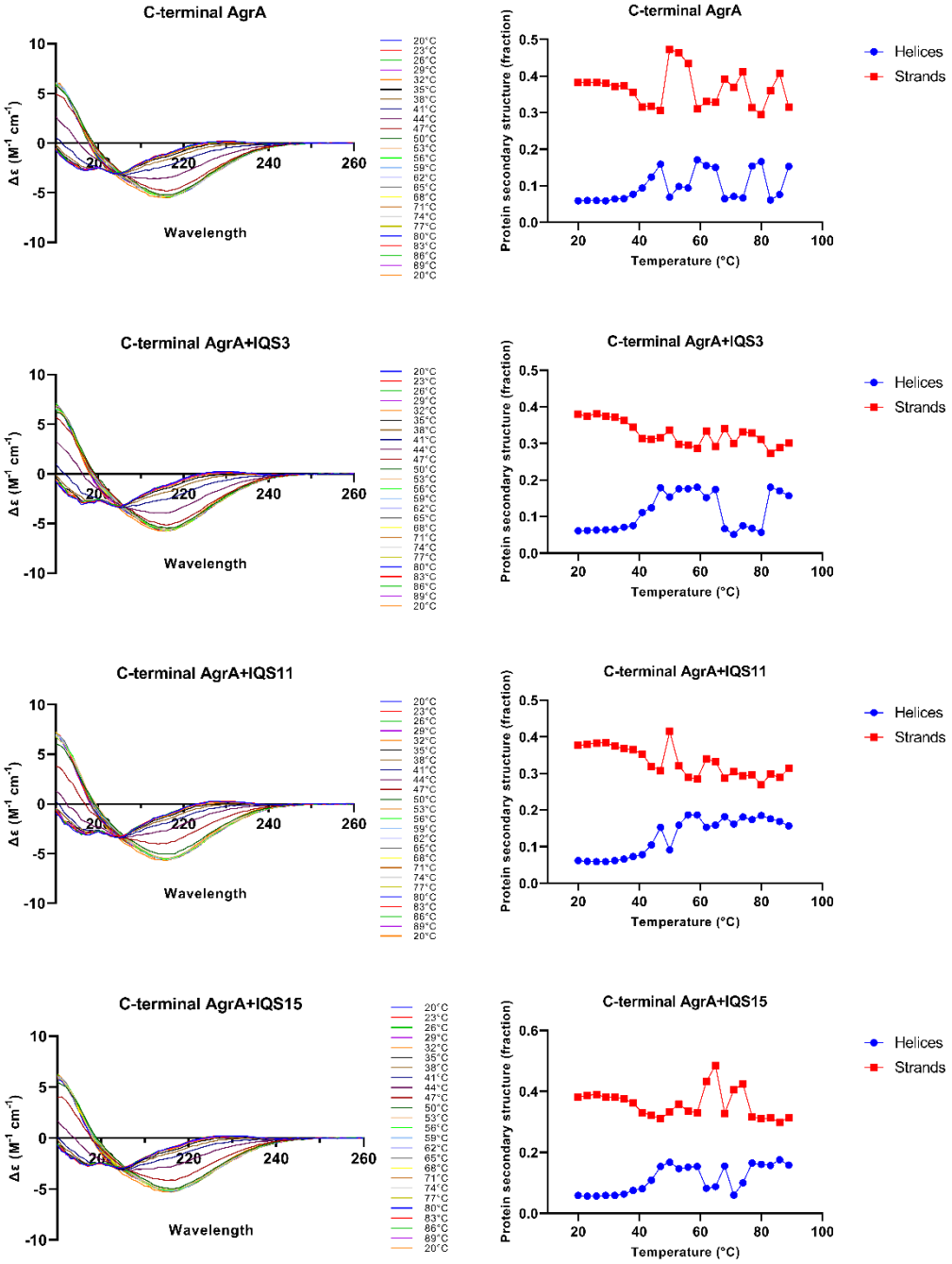


Figure 5.2.4 ***Far UV CD spectra and secondary structure content of C-terminal AgrA with IQS analogues.*** *The far UV CD spectra are shown on the left and the secondary structure content (helical in blue and strands in red) on the right, of C-terminal AgrA alone and in presence of IQS analogues 3, 11 and 15. The β-strand dominated structure transitions towards a slightly more helical*

conformation. Obtained using the CONTINLL SMP56, SP43+13 Membrane SSE Fit algorithm.

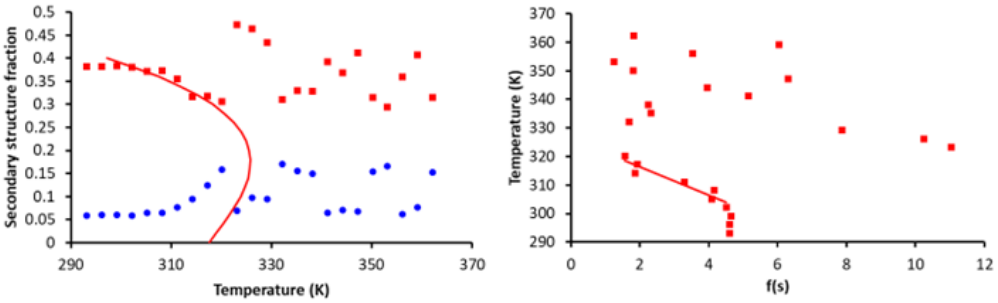
The calculated secondary structure fraction is quite noisy and makes the Landau analysis less precise. Figure 5.2.4 shows the CD spectra and the secondary structure. The data were processed using an Average Amino Acid Molecular Weight of 121 and at wavelength 216 nm.

Table 5.2.2 Melting temperatures of AgrA with IQS analogues.

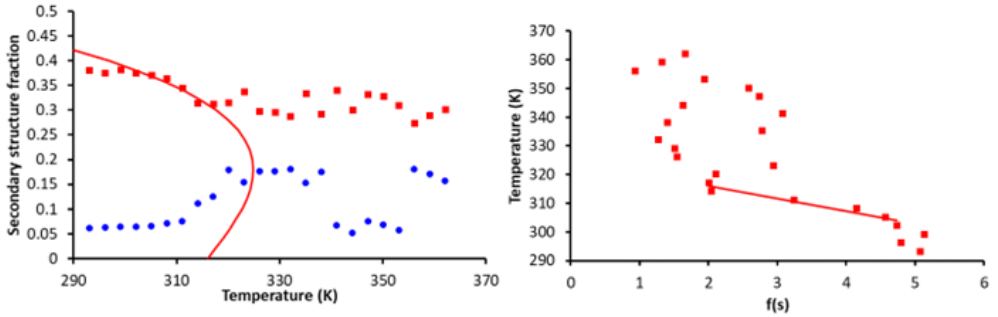
Sample	**Temperatures (°C)**
C-terminal AgrA	43.32
C-terminal AgrA + IQS3	42.87
C-terminal AgrA + IQS11	45.37
C-terminal AgrA + IQS15	44.38

The biggest change in the melting point is seen upon addition of IQS11 to the C-terminal AgrA, with a 2 °C increase. IQS15 triggers a 1 °C increase in transition temperature, while IQS3, the supposed active compound, decreases it by 0.5 °C.

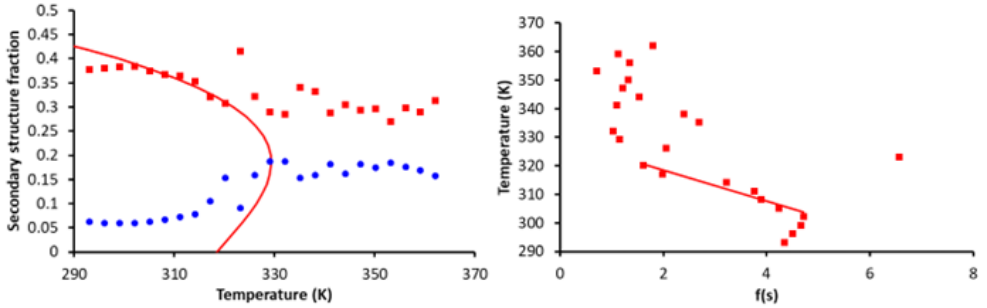
(A) C-terminal AgrA



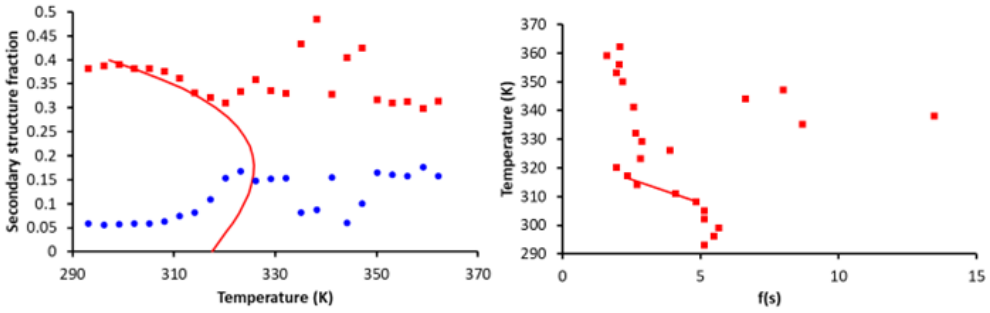
(B) C-terminal AgrA+IQS3



(C) C-terminal AgrA+IQS11



(D) C-terminal AgrA+IQS15



(E)

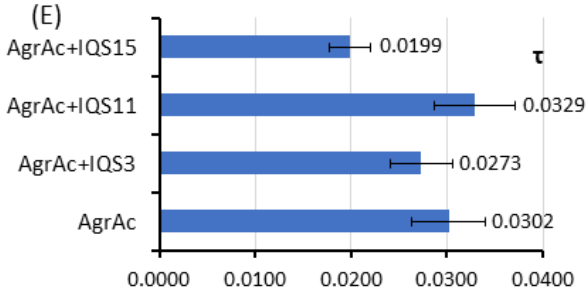


Figure 5.2.5 ***Landau analysis of C-terminal AgrA with IQS analogues.*** *Secondary structure content with equation of state (left) and linear fits of T vs f(s) (right) of C-terminal AgrA alone (A), with IQS3 (B), with IQS11 (C) and with IQS15 (D); transition cooperativity plot (E).*

The cooperativity of the transition is decreased by adding IQS3, the active compound, and IQS15, the inactive compound. IQS11 triggers a slight increase in τ, but all changes in this data set are very subtle (Fig. 5.2.5).

Overall, although some of the compounds do trigger a change in the proteins' sensitivity to temperature, these changes are very subtle and not decisive enough to draw conclusions on activity. This may be an indication that their inhibitory activity is not related to AgrA and its structure, or that a hindrance to the DNA binding site is the only effect with no effect on structure.

5.3 Thermal studies of the *agr* two-component system by Small Angle Scattering

The Landau series expansion analysis of CD data provides information on the thermal behaviour of proteins around their transition point on the basis of the evolution of their secondary structure. The increase in temperature breaks the hydrogen bonds unfolding the structure, which transitions into a β-strand rich, amyloidogenic arrangement. The melting temperature, calculated using the Boltzmann analysis, is used to calculate a transition cooperativity value, τ, and the comparison of this value allows to quantify the effect of molecular interactions on protein secondary structure stability.

This technique was extended and adapted to small angle scattering. This was done by identifying the radius of gyration R_g as thermodynamic order parameter instead of the helical or β-sheet content used in CD. The R_g of a particle is defined as the average radius of scattering density from the centre of mass (Wang et al., 2017b) but is more commonly described as a measure of the overall size of the particle (Alford et al., 2017, Boothroyd, 2020). The R_g reflects the apparent scattering of proteins at the hydration interface. As the temperature increases, so does the overall "size" of the protein in relation to the increased fluctuations induced by heat. These structural fluctuations allow hydration deeper into the core of the protein, which results in the scattering core, described by R_g, appearing smaller as temperature increases. The Gibbs free energy reflects the energy required in the making of the given system, which includes the energy necessary in the creation of bonds that give both secondary and tertiary structure of a protein. Similarly to the decrease in Gibbs free energy described in the methodology for CD spectroscopy and associated with the loss of secondary structure, the same is observed in SAXS as the tertiary structure unfolds.

The response of the protein $AgrC_{206}$ to temperature was assessed by molecular dynamics simulations (MD simulations carried out by Dr Boyan Bonev). The protein dimer was built in CHARMM-GUI (Jo et al., 2008, Brooks et al., 2009, Lee et al., 2016) and models annealed by MD at different temperatures, 77, 273, 293, 303, 310, 313, 323 and 333 K. Each simulation was carried out for 100 ns and one frame recorded for each ns. The last temperature, 333 K, is past the melting point of $AgrC_{206}$ as experimentally determined by CD and SAXS.

Cluster analysis was carried out using UCSF Chimera (Pettersen et al., 2004) for each trajectory and a SAXS profile simulation of the main clusters was carried out using CRYSOL (Svergun et al., 1995).

Plotting the simulated SAXS profiles of the room temperature and highest temperature simulations (Fig. 5.3.1) shows that protein conformation of the leading cluster varies with temperature.

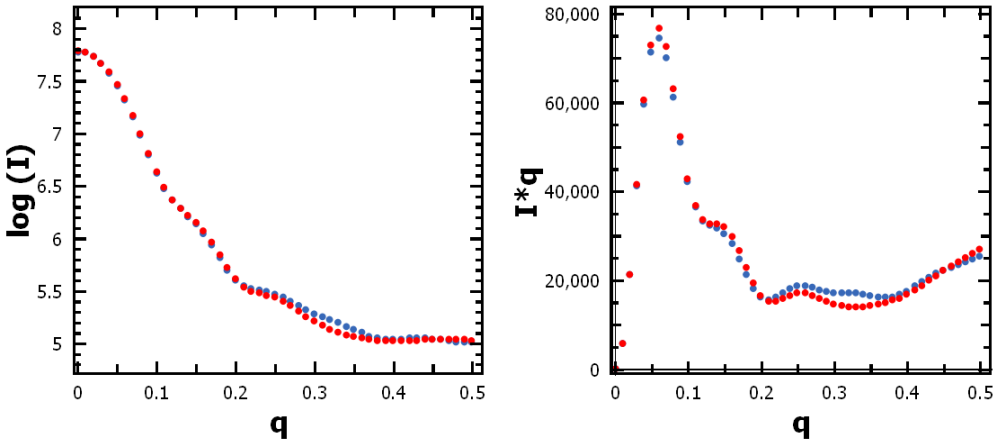


Figure 5.3.1 ***Simulated scattering profiles of AgrC$_{206}$ at 293 and 333 K.*** *Logarithmic (left) and Kratky (right) plots of the CRYSOL-simulated scattering profiles of AgrC$_{206}$ at 293 K (room temperature – blue) and 333 K (past melting point – red). The frame corresponding to the main cluster, used for the simulations, was identified using UCSF Chimera (Pettersen et al., 2004) from the 100 ns simulations.*

The conformational difference between the two temperatures is especially evident in the Kratky plot. Differences in the high *q* region indicate helical or strand rearrangement, rather than variations in the gross structure.

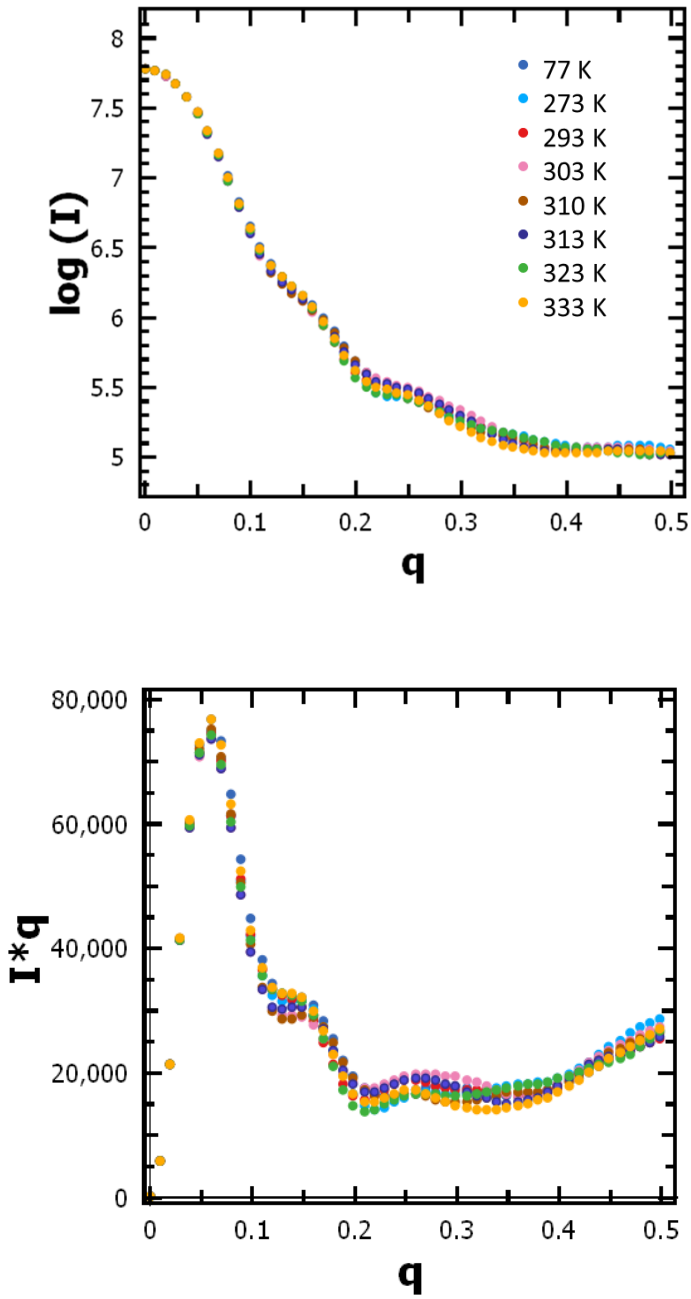


Figure 5.3.2 ***Simulated scattering profiles of AgrC$_{206}$ at increasing temperatures.*** *Logarithmic (top) and Kratky (bottom) plots of the CRYSOL-simulated scattering profiles of AgrC$_{206}$ at 77, 273, 293, 303, 310, 313, 323, 333 K. The frame corresponding to the main cluster, used for the simulations, was identified using UCSF Chimera from the 100 ns simulations. The Kratky plot in particular helps to visualise the conformational oscillations, especially evident in the* q*=0.25 region.*

Different options were screened to determine the most robust calculations to determine an accurate R_g, including ScÅtterIV, PRIMUS and a few models and algorithms from SasView 5.0.x. The Guinier-Porod model function, part of the Shape Independent models in SasView 5.0.5, was found to be the most accurate and consistent in the fitting when used with the Levenberg-Marquardt algorithm (Hammouda, 2010b, Hammouda, 2010a). This model was specifically developed to determine the special characteristics of a variety of shapes, including asymmetric objects.

The R_g values are normalised to the maximum value to obtain a thermodynamic order parameter of values between 0 and 1. These are then used, in conjunction with the positive spinodal, to calculate the function of the order parameter, in our case R_g (s is used in this case to refer to R_g):

(5.3.2)

$$f(s) = 3\left(\frac{s - s_+}{s_+}\right)^2 + \left(\frac{s - s_+}{s_+}\right)^3$$

The values of T vs *f(s)* giving a monotonic decrease are taken to calculate the intercept and slope of the spinodal temperature, which, once averaged, give the spinodal temperature. The resulting relation between the melting temperature T_m and the spinodal temperature T_+ represents the normalised value called τ as described below:

(5.3.3)

$$\tau = 1 - \frac{T_m}{T_+}$$

This value τ is dimensionless and allows comparison between any protein or peptide in terms of the cooperativity of their unfolding transition.

Data acquisition was set up to collect at 20 °C and ramp from 27 to 60 °C in 3 °C steps. The transition from folded to unfolded state is evident by observing the linear plots and manifests through an increase in the aggregation at low *q*. The melting temperature is assumed to correspond to the highest temperature at which a good fit is possible and the R_g estimation is not evidently the result of aggregates. C199S AgrA is shown as an example below (Table 5.3.1).

Table 5.3.1 R_g values for C199S AgrA at different temperatures.

Temperature (°C)	R_g (Å)	Temperature (°C)	R_g (Å)
20	22.811	45	21.697
27	23.538	48	19.097
30	23.558	51	97.27
33	22.906	54	162.16
36	22.88	57	271.54
39	22.997	60	232.63
42	21.883		

In this case, the melting temperature is estimated to be 48 °C as the R_g of the subsequent temperatures is not in the vicinity of any of the previous ones. The scale of the radii of gyration past the melting temperature is so different that these values are disregarded from the subsequent elements of the analysis as they would make the visualisation difficult. The previously described function $f(s)$ (or $f(R_g)$) is calculated from the R_g values up to the melting temperature and plotted against temperature. The *f(s)* values that show a linear correlation are used to determine the intercept, which equals T_+, and the slope which is

proportional to the distance between T_+ and T_m, by adjusting s_+. The spinodal temperature is given by the average of the intercept and slope. From T_+ and T_m, the parameter τ is derived as described in equation 5.3.3.

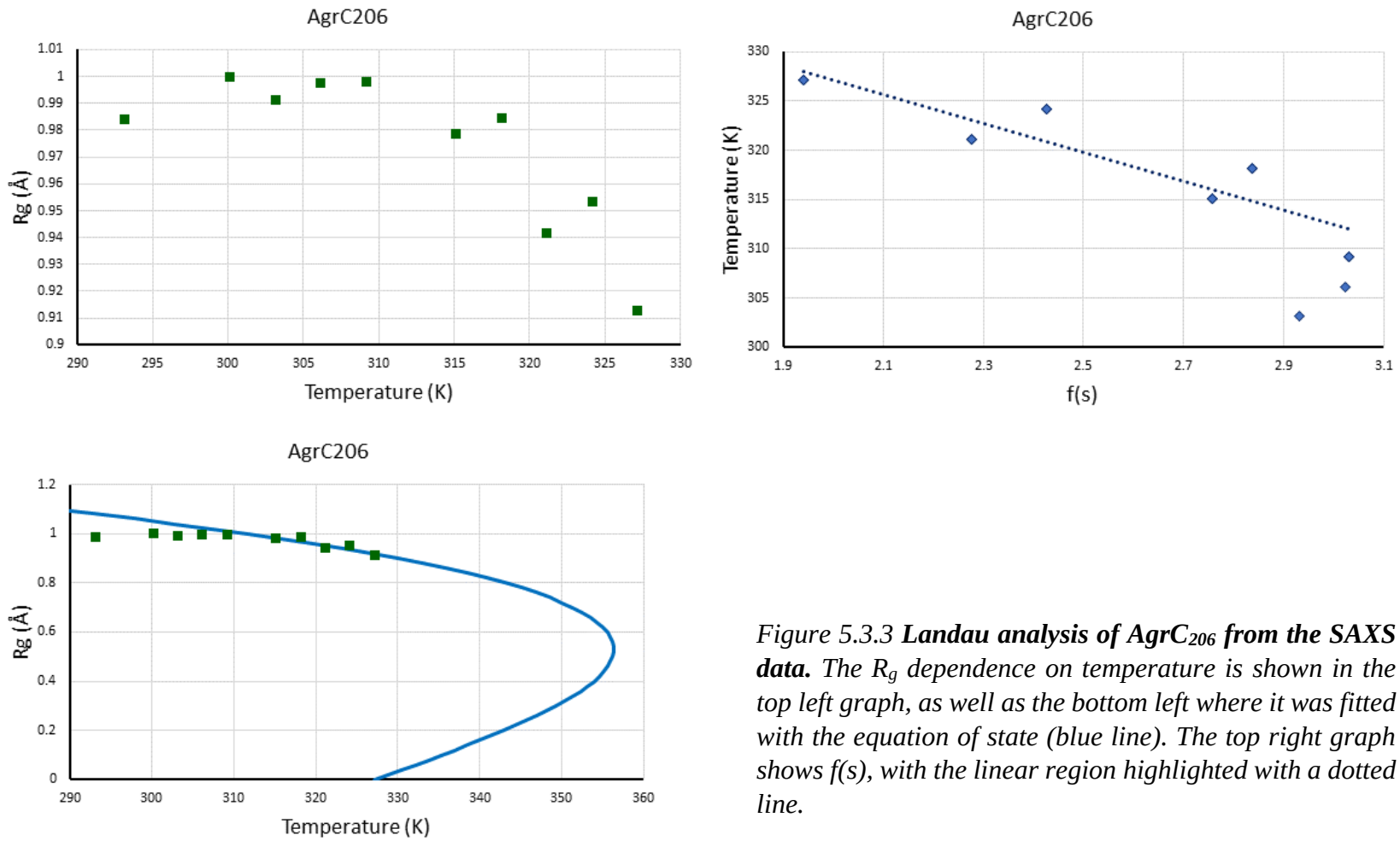


Figure 5.3.3 ***Landau analysis of AgrC$_{206}$ from the SAXS data.*** *The R_g dependence on temperature is shown in the top left graph, as well as the bottom left where it was fitted with the equation of state (blue line). The top right graph shows f(s), with the linear region highlighted with a dotted line.*

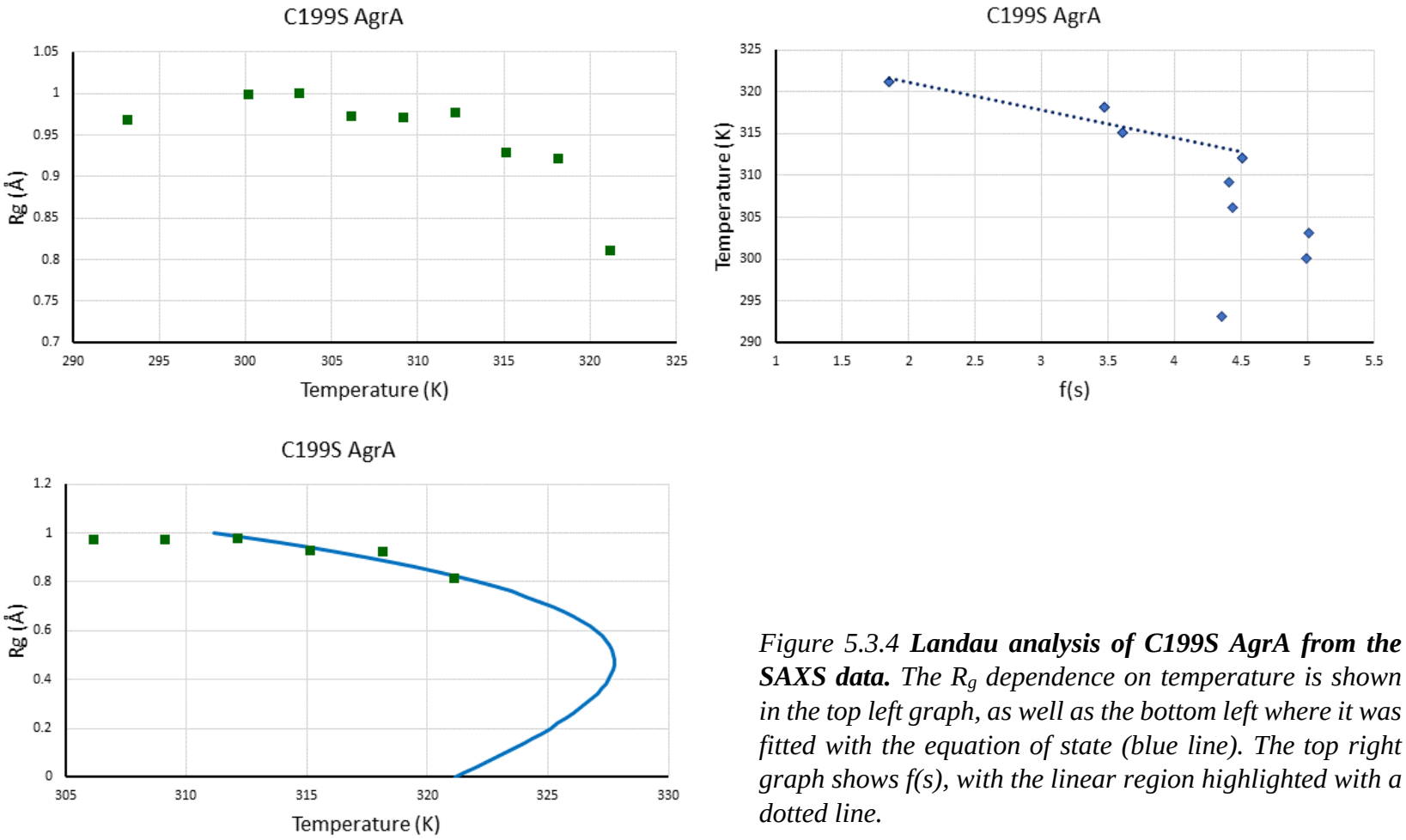


Figure 5.3.4 ***Landau analysis of C199S AgrA from the SAXS data.*** *The R_g dependence on temperature is shown in the top left graph, as well as the bottom left where it was fitted with the equation of state (blue line). The top right graph shows f(s), with the linear region highlighted with a dotted line.*

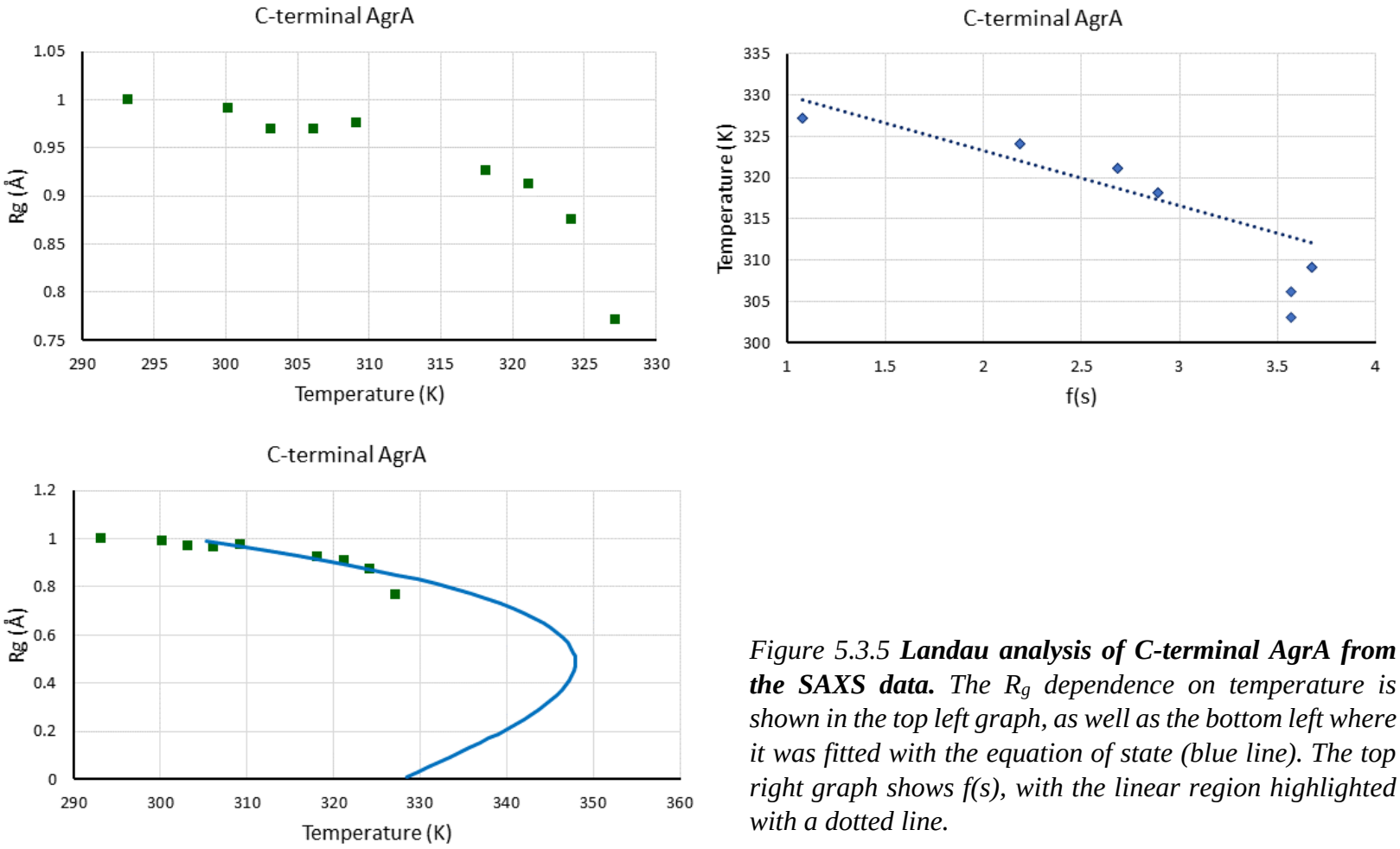


Figure 5.3.5 ***Landau analysis of C-terminal AgrA from the SAXS data.*** *The R_g dependence on temperature is shown in the top left graph, as well as the bottom left where it was fitted with the equation of state (blue line). The top right graph shows f(s), with the linear region highlighted with a dotted line.*

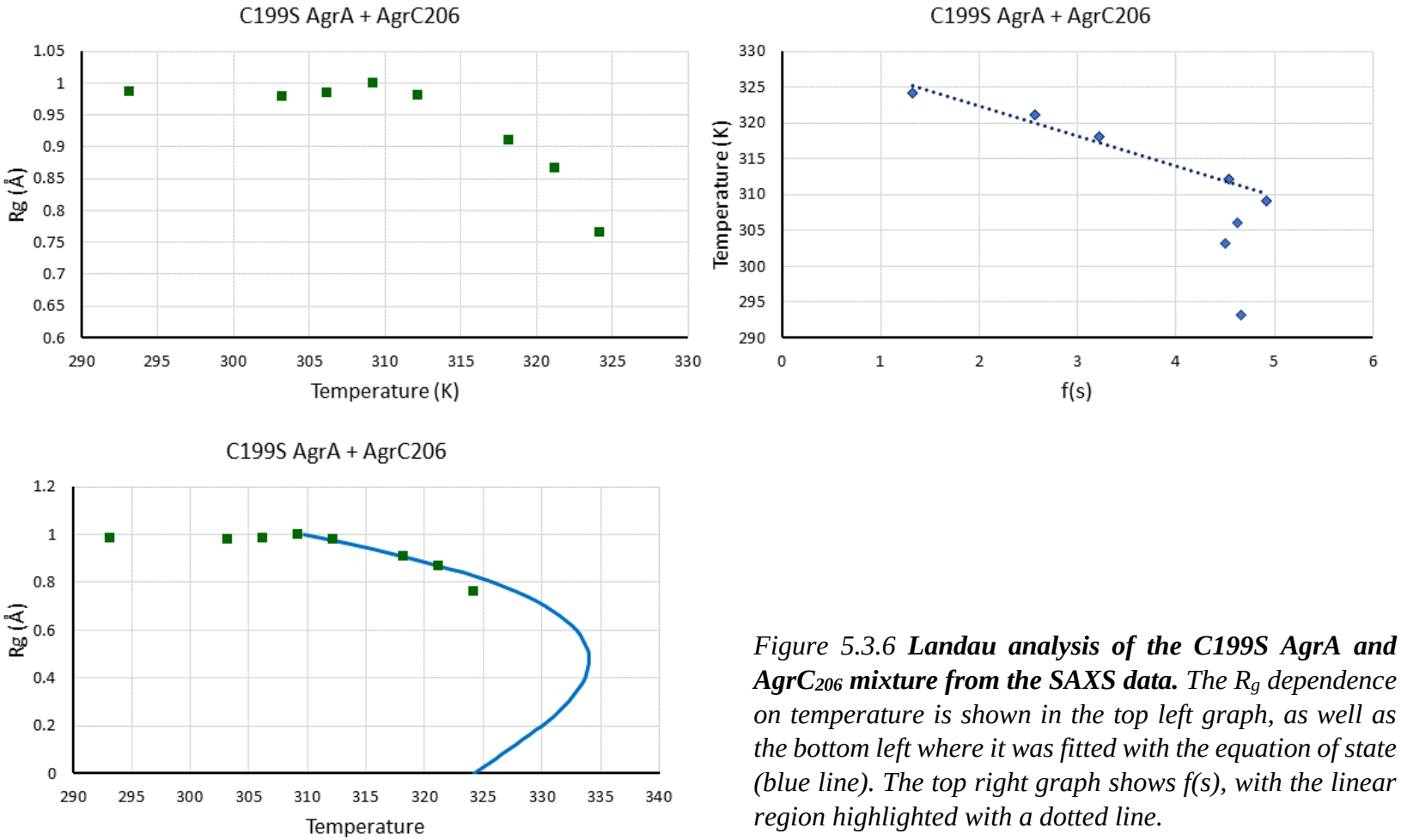


Figure 5.3.6 ***Landau analysis of the C199S AgrA and $AgrC_{206}$ mixture from the SAXS data.*** *The R_g dependence on temperature is shown in the top left graph, as well as the bottom left where it was fitted with the equation of state (blue line). The top right graph shows f(s), with the linear region highlighted with a dotted line.*

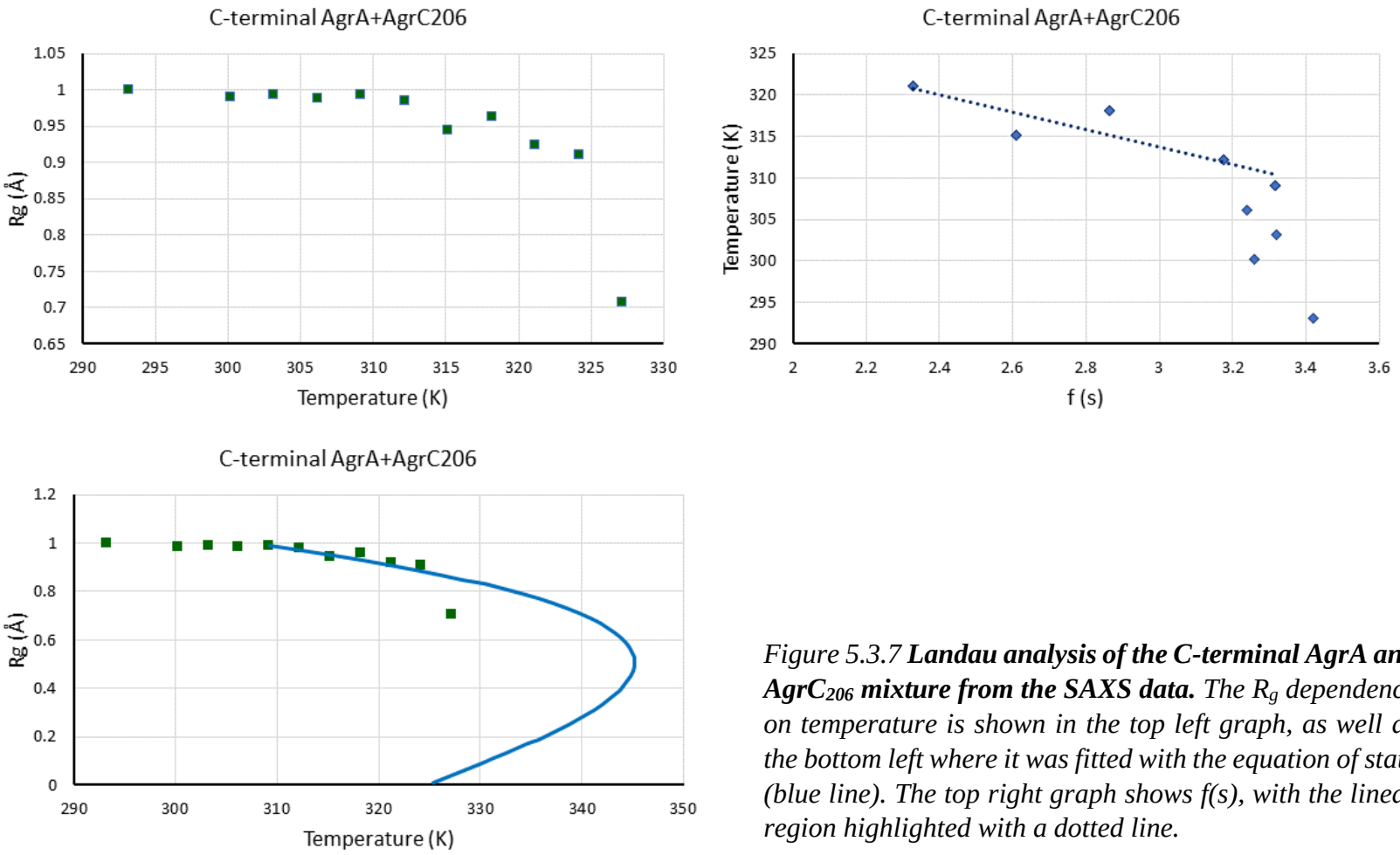


Figure 5.3.7 ***Landau analysis of the C-terminal AgrA and AgrC$_{206}$ mixture from the SAXS data.*** *The R_g dependence on temperature is shown in the top left graph, as well as the bottom left where it was fitted with the equation of state (blue line). The top right graph shows f(s), with the linear region highlighted with a dotted line.*

Figures 5.3.3 to 5.3.7 show how the calculated Rg decrease as a function of temperature. The data is only plotted up to the transition temperature, as the difference in the Rg range after the melting point would make visualisation very difficult. The plotting of *f(s)* against temperature allows to determine the linear region in the proximity of the transition which is used for the calculation of τ. Finally, the equation of state is plotted together with the Rg progression to show the cooperativity of the transition, represented by the difference between the spinodal *T+* and *Tm*. Tm is the intercept on the temperature axis, while the maximum temperature in the equation of state represents T+. The cooperativity is also shown in Figure 5.3.8, where τ is plotted for all proteins studied.

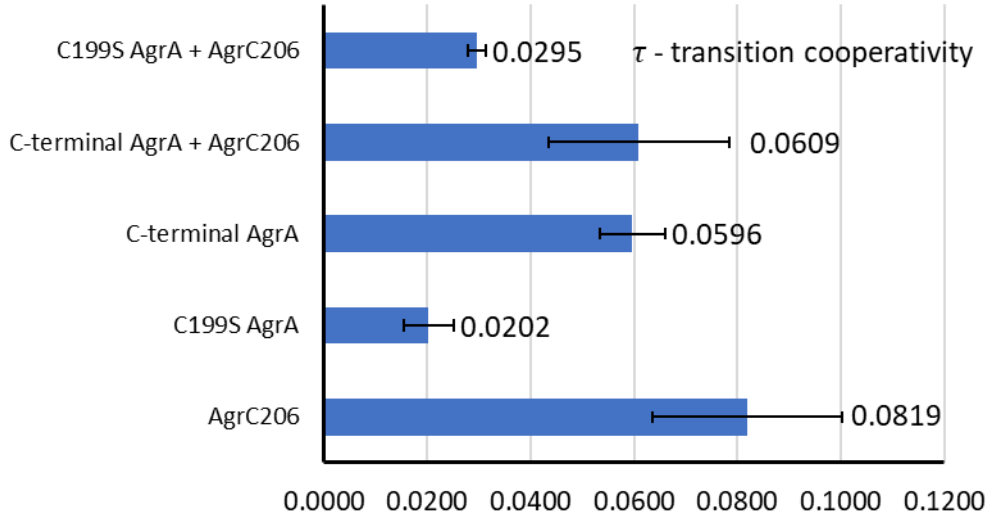


Figure 5.3.8 ***Thermal unfolding transition cooperativity of the Agr proteins in SAXS.*** *The cooperativity of the transition of C199S AgrA, C-terminal AgrA, $AgrC_{206}$ and the two mixtures is shown on a scale of 0-1.*

The scale of the values is in accordance with what is observed in CD. C199S AgrA shows the lowest cooperativity, indicating an unstable behaviour. $AgrC_{206}$ has the highest temperature stability in this set, which is consistent with the

corresponding CD data. The mixture of these two proteins in a 1:1 molar ratio shows a cooperativity in between the individual proteins, with a tendency towards AgrA. The C-terminal domain of AgrA has a noticeably higher stability compared to the whole protein, and the mixture of this domain with $AgrC_{206}$ sits very closely to the C-terminal AgrA.

The data acquired on B21 are well suited to this processing pipeline. They were compared with the data for two model proteins collected at the Diamond-Leeds offline SAXS beamline (DL-SAXS). This instrument has a much lower flux (1.4e9 photons per second) compared to the online beamlines at Diamond such as B21 (2×10^{12}photons pre second) but the Peltier temperature stage guarantees much finer temperature control than temperature-controlled quartz capillary used by the BioSAXS B21 beamline. It also can reach higher temperatures, therefore allowing analysis of proteins with higher transition points.

For this study two proteins which can be kept at high concentrations and present minimal handling issues were chosen: $AgrC_{206}$ and GAPDH. GAPDH (glyceraldehyde-3-phosphate dehydrogenase) is most commonly known for its role in glycolysis but has a multitude of additional functions. One of them is as an RNA-binding protein involved in RNA stability and translation. While the interactions between GAPDH and a series of RNA targets have been established, the precise biological relevance is still largely unknown. The function of GAPDH in relation to RNAs is varied, as it can have a positive or negative influence on stability, therefore also affecting RNA translation. The mechanism allowing such flexibility in interaction and function is yet to be elucidated. Conformational plasticity is hypothesised to be related to the ability of RNA-

binding proteins to interact with a variety of RNA structures and influence them so selectively.

The protein samples were heated from 20°C to 80°C in 3°C steps and 3 spectra per temperature recorded for 10 minutes. These were then averaged and subtracted as part of data reduction. Figure 5.3.9 represents the appearance of the samples at the end of the temperature ramp. They appear milky and viscous and the irreversible degradation is much more apparent in these photos than in the ones taken during the CD experimental sessions.

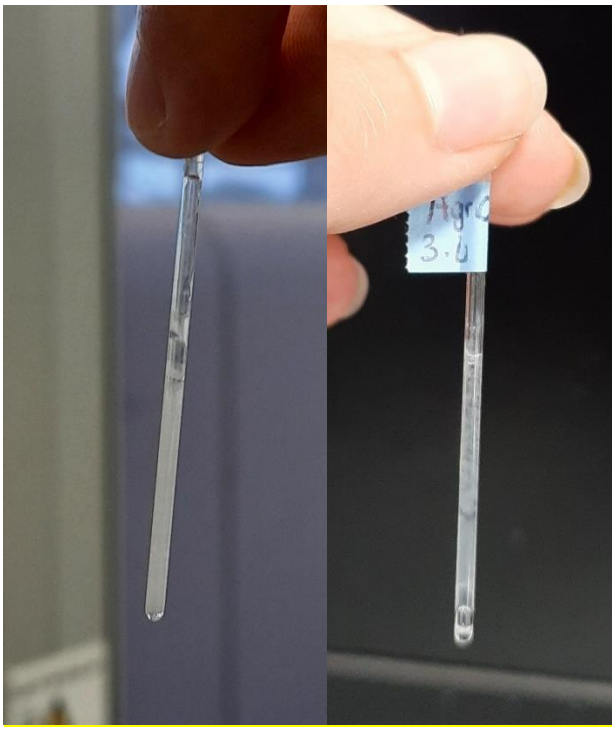


Figure 5.3.9 ***Melted protein in DL-SAXS capillaries****. Photos of the capillaries used at the DL-SAXS beamline containing GAPDH (left) and* $AgrC_{206}$ *(right) after the temperature ramp (20 to 80°C).*

The visual inspection of the data collected for GAPDH already shows the trend of the transition (Fig. 5.3.10).

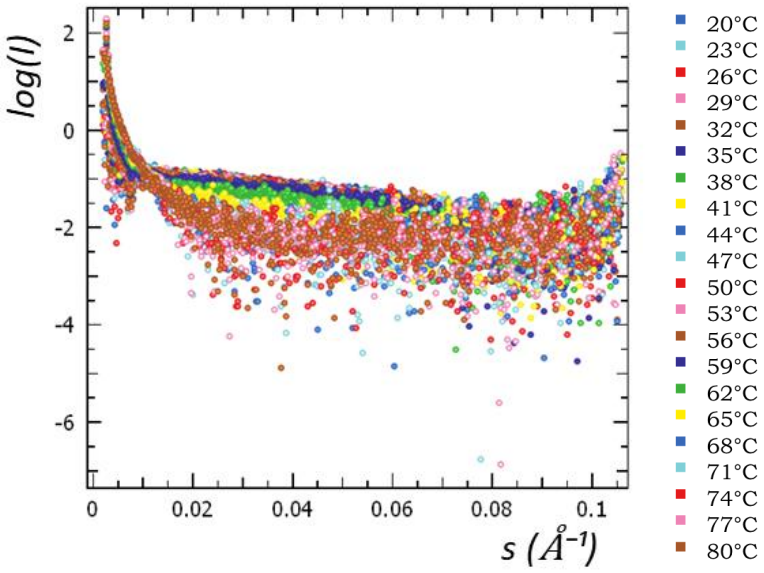


Figure 5.3.10 ***Temperature-dependent scattering profiles of GAPDH.*** *The scattering profile of GAPDH was collected at temperatures 20-80°C in 3°C steps. The logarithmic scale plots at different temperatures are shown.*

The levels of aggregation and the slope increase noticeably together with temperature.

The thermal scan of $AgrC_{206}$ shows a similar trend to GAPDH, where the level of aggregation consistently increases with temperature (Fig. 5.3.11).

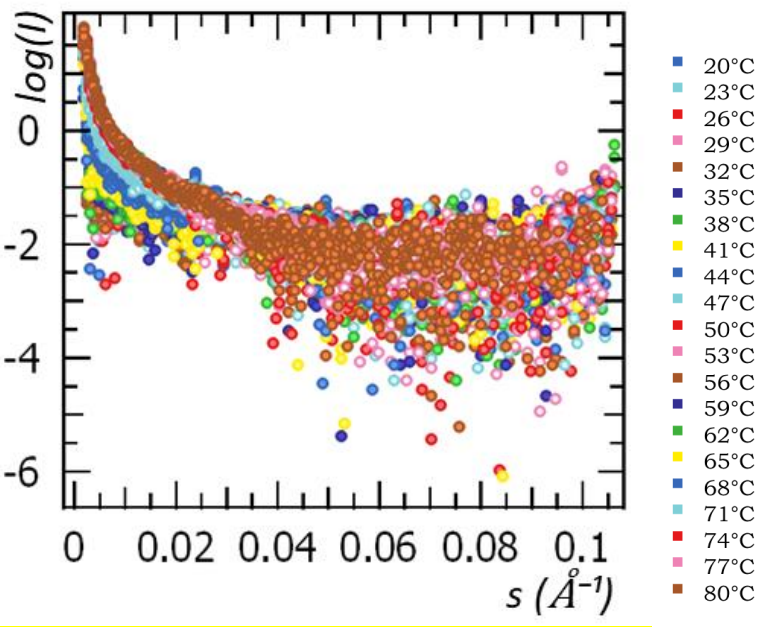


Figure 5.3.11 ***Temperature-dependent scattering profiles of AgrC$_{206}$.*** *The scattering profile of AgrC$_{206}$ was collected at temperatures 20-80°C in 3°C steps. The logarithmic scale plots at different temperatures are shown.*

Both proteins were kept at very high concentrations for this study, which may not be possible for all protein samples. There is also the risk of incurring in multiple scattering events when the sample is overly concentrated. Despite recording these samples at 9-10 mg/ml, the resolution is still comparatively poor, especially for AgrC$_{206}$. This protein did not perform as well as GAPDH due to its dimeric state, which corresponds to ~56 kDa, as opposed to tetrameric GAPDH at ~120 kDa. This is especially evident in Figure 5.3.12, where the error bars are shown.

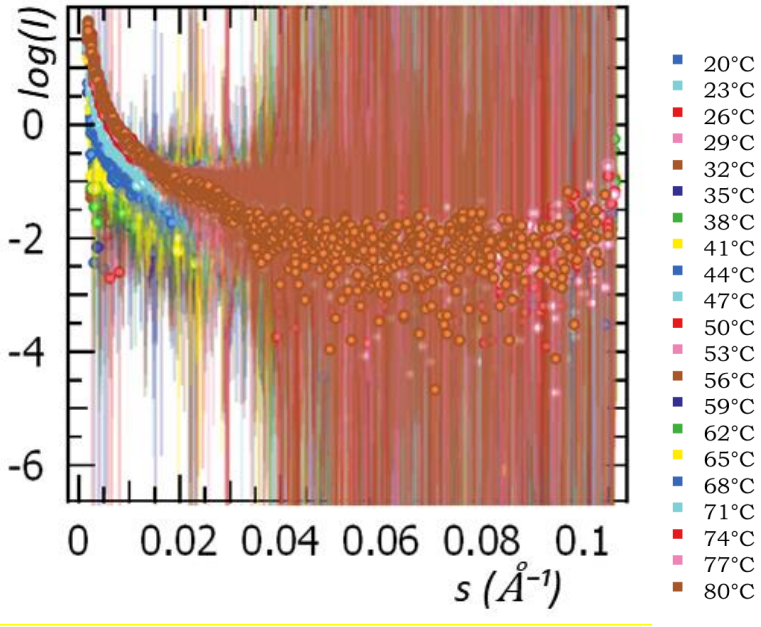


Figure 5.3.12 ***Temperature-dependent scattering profiles of $AgrC_{206}$ with error bars.*** *The scattering profile of $AgrC_{206}$ was collected at temperatures 20-80°C in 3°C steps. The logarithmic scale plots at different temperatures are shown and the error bars are highlighted.*

The *q* range recorded at DL-SAXS was only up to 0.1 $Å^{-1}$. While this increases the level of detail in the Guinier region, it also leads to a proportionally greater loss of data when they are processed to discard low quality data points and aggregation at low *q*. These reasons make the data fitting and analysis difficult and not fully reliable. The best fitting was obtained using the Guinier-Porod model and the fitting algorithm DREAM, but the R_g distribution (Fig. 5.3.13) is not characterised by a definite trend and does not allow to extrapolate a coherent $f(s)$.

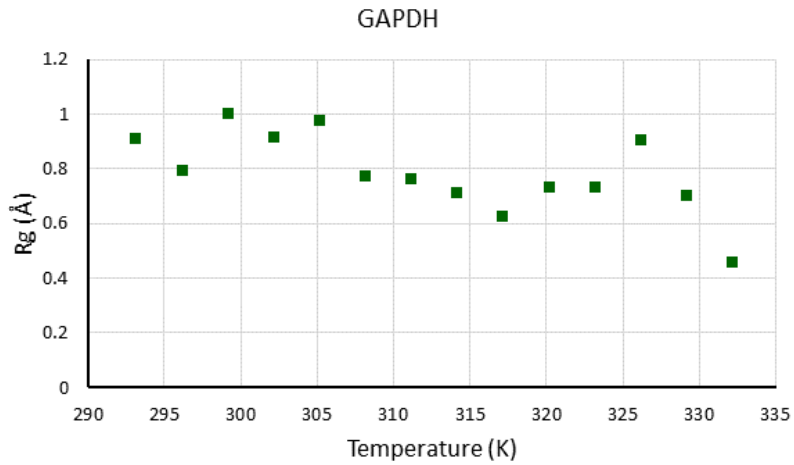


Figure 5.3.13 ***Radius of gyration of GAPDH as a function of temperature.***

While the downwards trend is somewhat apparent upon visual inspection, it was not possible to identify a spinodal point, probably due to the noise of the data.

While theoretically this beamline provides an optimal environment for this type of thermal studies, the quality of the data, even at such high concentrations, does not provide enough detail on the thermodynamics of the systems.

5.2. Discussion

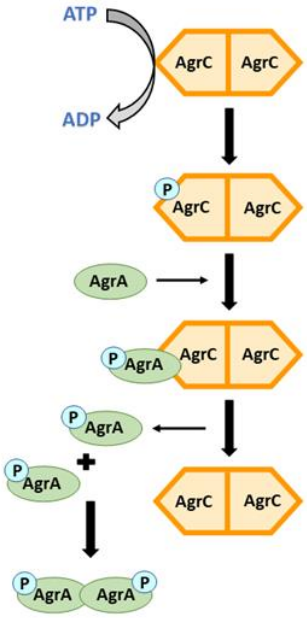
This Chapter aimed at describing the thermal unfolding of AgrA and $AgrC_{206}$ following their loss of both secondary structure (by CD) and tertiary fold (by SAXS), and using this information to characterise their molecular interaction.

Studying molecular interactions *in vitro* is notoriously difficult. Our proposed approach uses thermal stability studies as a way to explore stabilising or destabilising influences of an interaction partner on the protein studied. The Landau series expansion analysis (Hussain et al., 2018) uses CD data in the form of secondary structure content estimation and melting temperature as

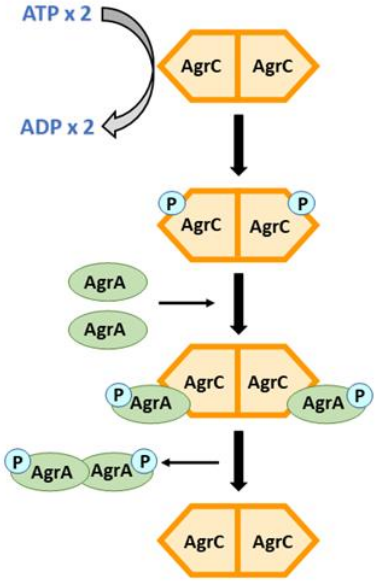
extrapolated using the Boltzmann analysis, to determine the cooperativity of the melting transition of the protein. High cooperativity indicates a system that is very stable against thermal denaturation, while low cooperativity suggests a more gradual unfolding and general sensitivity to temperature changes.

This methodology was used to explore the role of ATP and its non-hydrolysable analogue, sATP, in the molecular interaction between AgrA and AgrC, the two-component system of the *agr* quorum sensing system. It is well established that AgrC, as the histidine kinase, recruits ATP following extracellular stimulus by AIP and acquires a phosphate group. This is then passed onto AgrA, the response regulator, making it more prone to dimerisation and increasing its affinity for the promoter P2. As many histidine kinases, AgrC operates as a dimer by auto-*trans*-phosphorylating (Xie et al., 2019), and this oligomeric state was shown to be retained upon removal of its transmembrane domain (see Chapter 4). While the existence of the interaction between these two proteins has been established, the mechanism of recruitment of ATP and AgrA to AgrC is still poorly characterised. The possible scenarios are described in Figure 5.4.1.

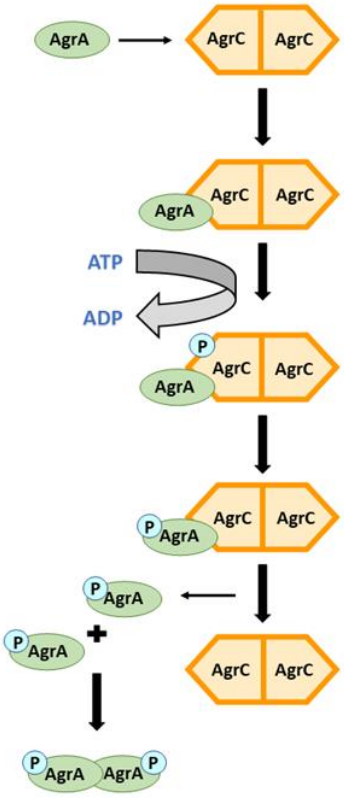
Scenario 1



Scenario 2



Scenario 3



Scenario 4

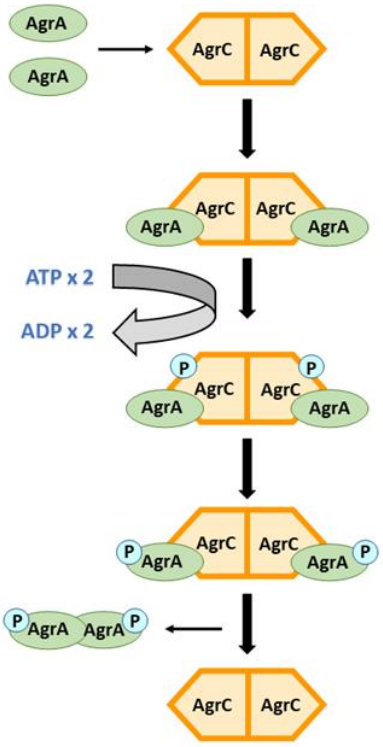


Figure 5.4.1 ***Possible two-component system cascades.*** *Four potential scenarios which show different recruitment orders of ATP and AgrA to AgrC.*

We sought to shed light on how ATP and non-hydrolysable ATP affect the secondary structure and thermal stability of the two proteins of the two-component system, and to try and elucidate the oligomeric state of AgrA when recruited to AgrC.

The Landau analysis of AgrA in presence and absence of nucleotides reveals that, while ATP does not affect the thermal denaturation, the non-hydrolysable analogue has a destabilising effect on the protein, suggesting molecular interaction. $AgrC_{206}$, on the other hand, is greatly affected by both analogues, which make it extremely more thermally unstable.

Two protein mixtures where investigate: 1:1 and 1:2 molar ratio (AgrA:AgrC). Due to the inherently dimeric state of AgrC, the 1:1 ratio refers to Scenarios 2 and 4, while the 1:2 ratio represents Scenarios 1 and 3 (Fig. 5.4.1).

The mixture of the two proteins in a 1:1 ratio shows altered secondary structure profile and an increase in the transition cooperativity when sATP is present. Addition of ATP also shows increased cooperativity, although not as much as its non-hydrolysable analogue.

The 1:2 molar ratio mixture has a higher cooperativity compared to its 1:1 equivalent. This can either be due to a real stabilising effect, or to the predominant presence of AgrC, which is an extremely thermally stable protein. Both analogues significantly lower the transition cooperativity of this system, ATP more so than sATP.

The study of these systems also allowed to further develop the Landau methodology by investigating the presence of two transitions. The manual Boltzmann fitting option in CD Apps allowed to identify two transition temperatures, the higher of which is similar to the one originally identified by the automatic function. These temperatures where then used to carry out a double Landau analysis on $AgrC_{206}$, $AgrC_{206}$+ATP and AgrA+AgrC+ATP both in 1:1 and 1:2 ratios. The presence of a second transition can be associated with multi-domain proteins, in which case the appearance or disappearance of a second transition would represent a stiffening (or loosening) of the overall structure resulting from molecular interaction. It means that the system is either stabilised at low temperatures, to then proceed to its normal transition, or that the system is overall more unstable at higher temperatures if the cooperativity of the transition at high temperature is reduced. The disappearance of the first transition in $AgrC_{206}$ is observed upon addition of sATP. As ATP (and sATP) reacts with the CA domain, which according to structure is quite separate from the DHp domain and can have its own responsiveness, this disappearance of the first, cooperative transition indicates that sATP has a strong interaction with $AgrC_{206}$ and suggests that first transition to be associated with the CA domain. The lower cooperativity of the first transition in presence of ATP suggests a transient molecular interaction, where the variation in stability is observed, as seen in the sATP case, but potentially not retained. The high temperature transition does not show any variation across the AgrC dataset and is therefore associated with the DHp domain.

The addition of ATP also triggers a second, cooperative transition in the 1:1 and 1:2 mixtures of proteins.

A similar approach was taken to investigate effects of the potential inhibitors of AgrA, the IQS analogues, on its structure and stability. The results of this study are however very subtle, showing that the activity of IQS may be more related to binding site hindrance than to any structural alteration.

Finally, the application of thermal studies of proteins in Small Angle X-ray Scattering was described. The temperature dependence of the model protein $AgrC_{206}$ was first investigated *in silico* by performing molecular dynamics simulations at different temperatures, ranging from 77 K to past the experimentally determined melting point of the protein. The scattering profiles simulated using the main cluster frames for each temperature show conformational rearrangements dependent on temperature. While the Kratky plots show a fully folded, rigid protein, this is due to the *in silico* origin of the file; the rearrangements are better visualised in terms of protein globularity despite the rigidity of the system.

The Landau analysis was adapted to be used on scattering data by extrapolating the Rg values from the thermal melts of a set of proteins from the *agr* system. The Rg is taken to represent the tertiary folding of the proteins and interpreted similarly to the secondary structure in CD. A decreasing trend in the Rg as temperature increases is observed across all datasets, due to the increasingly penetrating hydration as the thermal fluctuations of the molecule increase. This allowed to extrapolate τ values following a similar procedure to the one used for CD data. Overall these data sets provide a sound proof of principle for thermal studies to be conducted using SAXS (or SANS). This methodology will need further validation from a more varied collection of protein samples and could

benefit from more temperature points being collected, to provide a more convincing trend.

6. Sample environment development for SANS magnetic studies on biological samples

6.1. Magnetically aligned bicelles in morphological studies of membrane proteins using SANS

The main challenge in the structural investigation of membrane proteins is their hydrophobicity and consequent need for a surrounding lipidic environment. Such membrane mimics often mask the protein itself or are incompatible with the most common experimental designs for structure determination. Membrane-mimicking environments commonly used in structural biology include liposomes, bilayer micelles (or bicelles), micelles and more recently nanodiscs (Dufourc, 2021, Warschawski et al., 2011).

While bicelles present stability issues and can only be stored for ~3 days, their magnetic properties have made them very attractive to the Physical Chemistry community and NMR researchers in particular thanks to the possibility to obtain macroscopically oriented samples giving orientation-dependent constraints.

Bicelles originate from a mixture of short and long-chain lipids (Prosser et al., 2006). They arrange in a coin-like fashion, with the shorter-chain lipids or detergents positioning themselves around the rim of the coin as shown in Fig. 6.1 (Dufourc, 2021, van Dam et al., 2004, Sanders and Schwonek, 1992).

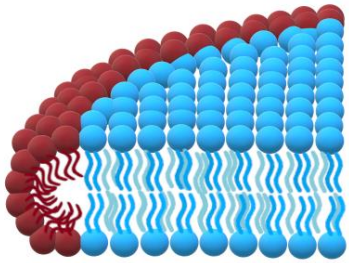


Figure 6.1 ***Schematic of a bicelle.*** *Portion of a negative bicelles, with the long-chain lipids forming a bilayer disk represented in blue and the short-chain lipids arranged to form the rim of the disc in red.*

Biological molecules are susceptible to aligning torques from the magnetic field due to their anisotropic diamagnetic susceptibility. For molecules with a negative diamagnetic anisotropy, including lipids, this means that the principal molecular axis is aligned with the magnetic field and the normal of the coin-like assembly is perpendicular to the field, as shown in Figure 6.2 (Sanders et al., 1994).

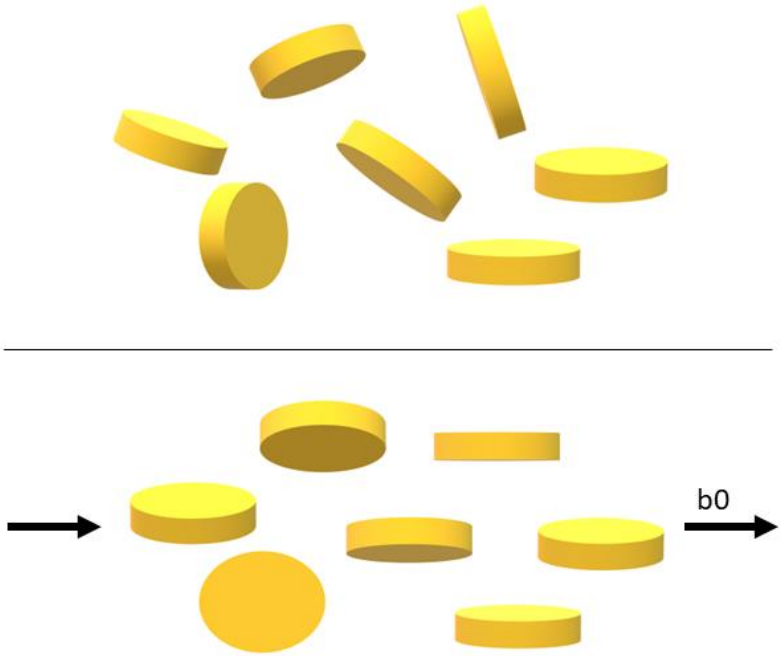
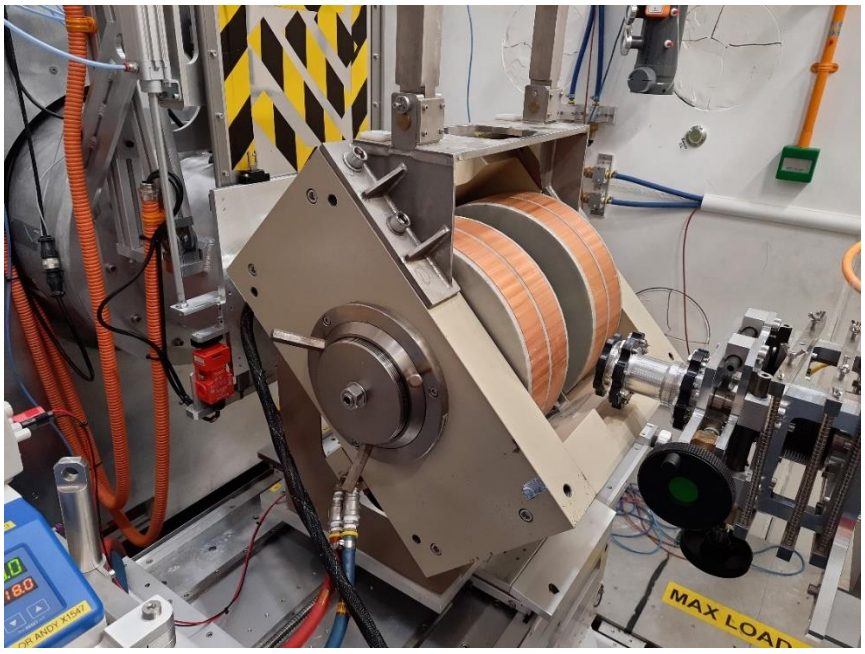


Figure 6.2 ***Schematic representation of bicelle orientation****. The top panel shows bicelles in no magnetic field; the bottom panel shows bicelles in a horizontal magnetic field.*

According to the commonly accepted model, lanthanide ions interact with the phospholipid headgroups (Prosser et al., 1998b) and their addition triggers a spatial re-orientation where the discoidal structures convert into positively ordered, perforated stacked lamellae (Nieh et al., 2002, Nieh et al., 2001, Prosser et al., 1998a), due to a positive magnetic susceptibility anisotropy (Prosser et al., 1996). The largest portion of literature regarding bicelles comes from the NMR community, who have the inherent advantage of an ever-present vertical magnetic field. As a result, the evidence of structural conformation of positive bicelles in absence of magnetic field is not very extensive, with only one study of positively ordered lamellae in a magnetic field reported by SANS (Nieh et al., 2001).

While unoriented bicelles have been characterised using small angle neutron scattering (Li et al., 2013), the characterisation and use of positively aligned bicelles represents a challenge for small angle scatterers. The use of magnets in the beamlines has largely been limited to cryo-magnets and magnetism-related experiments, which have a sample environment inherently incompatible with liquid samples. In this study the aim was to explore the behaviour of magnetically aligned bicelles in a small angle scattering beamline, with the help of the engineers at the ISIS Neutron and Muon Source who designed a temperature-controlled magnet that could be fitted on ZOOM and accommodate a banjo cell (Fig. 6.3).



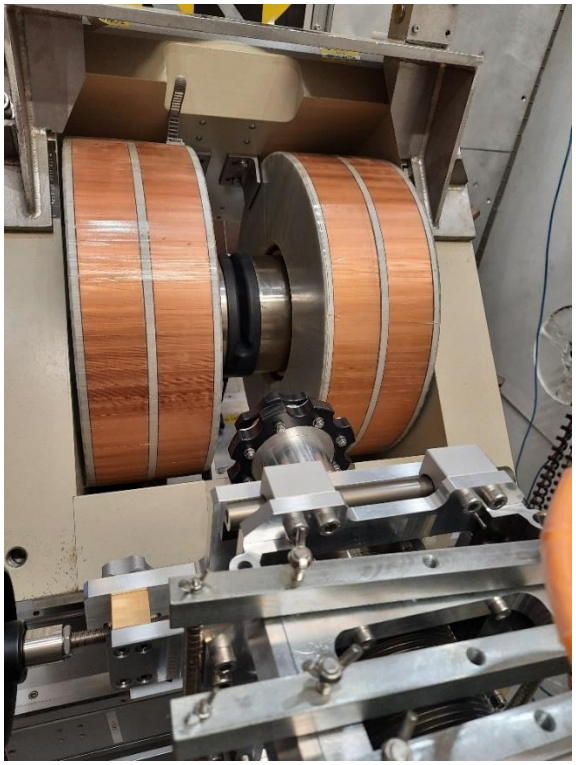


Figure 6.3 ***BioSANS Magnet.*** *Electro-magnet specifically designed to study aligned bicelles, fitted onto the ZOOM beamline. It is equipped with optional temperature control. The sample environment is at room temperature (18-20°C) when the temperature control is off, and at 45°C when it is on. The achievable magnetic field is approximately 1.5 T. The magnetic poles are positioned on either side of the sample holder creating a horizontal field.*

The magnet set-up in the ZOOM beamline showed in figure 6.3 allows to perform experiments with and without a magnetic field of approximately 1.5 T, at room temperature and at 45 °C, in all four combinations. The sample holder in the middle of the magnet was 3D-printed specifically to accommodate a 1 mm sample cell, which can hold ~300 µl of liquid sample.

The horizontal orientation of the magnetic field also allows to separate the effect of gravitational field from the magnetic field, with the two opposing each other.

Bicelles have been used as membrane mimetics to structurally characterise membrane proteins by NMR (Sanders and Landis, 1995, De Angelis and Opella, 2007, De Angelis et al., 2005, Park et al., 2006, Rajesh et al., 2016). The inherent magnetic field of NMR is an advantage in creating a uniform system of bicelles with the membrane protein encased within the membrane part of the discoidal structure. A bicellar system containing a membrane protein was constructed to ultimately adapt this approach to SANS experiments and extrapolate dimensional information on the conformation of membrane proteins.

The magnets used in alignment of bicelles as explored by NMR are characterised by high magnetic fields, normally a few T. The magnetic field used in NMR studies of bicelles is as low as 5.9 T (Prosser et al., 1998a, Vold and Prosser, 1996), while EPR spectroscopy has explored these systems in much lower fields such as 0.64 T (Inbaraj et al., 2006). The susceptibility of our system in the 1.5 T field generated by our magnet was verified.

A mixture of DMPC and DHPC was used in a 3:1 molar ratio (q = 3). The identified optimal alignment temperature for bicelles of this composition is 40-43 °C (De Angelis and Opella, 2007), which is very similar to our sample

environment temperature. A temperature of ≥30 °C (Sternin et al., 2001) and approximately 48% w/v total lipid were previously identified as optimal conditions for this phase by Dr Boyan Bonev and Dr Vivien Yeh (see Appendix for NMR data at increasing temperatures).

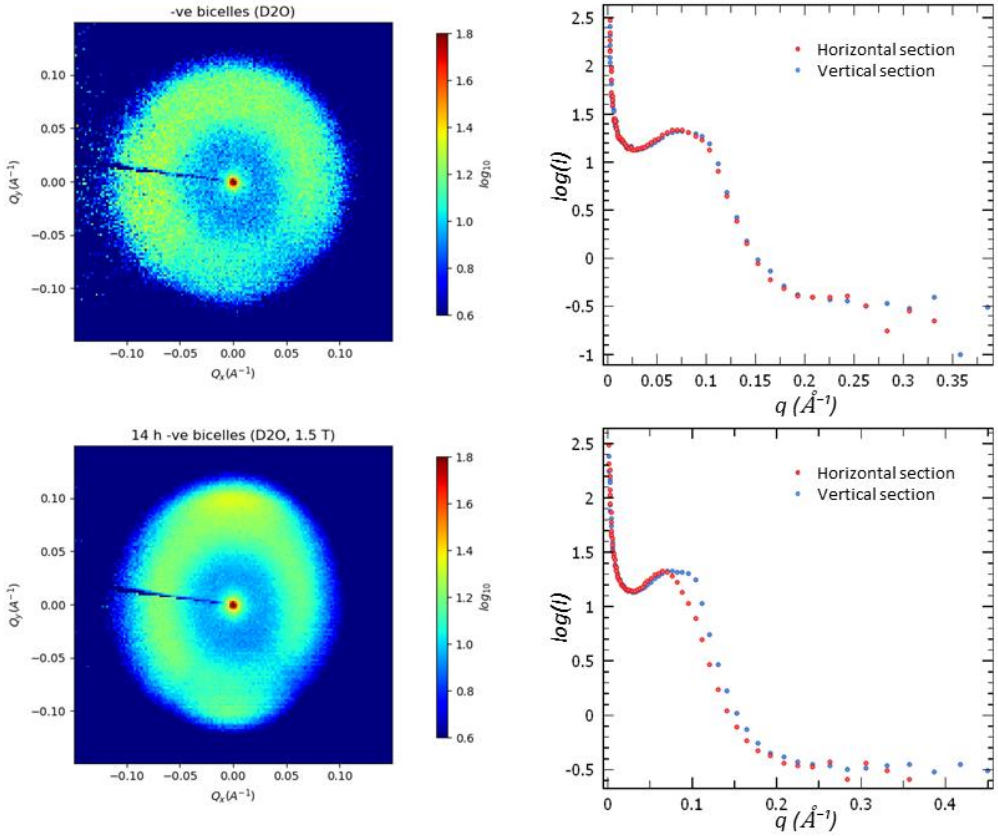


Figure 6.4 ***Effect of magnetic field on 'negative' bicelles.*** *DMPC/DHPC bicelles in absence of magnetic field in 2D (top left) and 1D (top right). Negative DMPC/DHPC bicelles after 14 hours in a 1.5 T horizontal magnetic field, in 2D (bottom left) and 1D (bottom right). The alignment shows as an intensity increase on the vertical axis and a narrowing of the horizontal plane.*

Bicelles in an external magnetic field align in a direction dictated by the negative value of lipid diamagnetic susceptibility anisotropy; we will refer to them as 'negative' bicelles from now on. A horizontal magnetic field of 1.5 T is sufficient align the DMPC/DHPC bicellar system (Fig. 6.4). In absence of magnetic field

the mixture exhibits an axially symmetric 2D pattern and no discrepancies in the scattering profiles between the horizontal and vertical planes (Fig. 6.4). Upon application of a 1.5 T magnetic field to the same sample, the 2D profile adopts a cross-like conformation, with the asymmetry confirmed by the 1D profile which shows a sharper peak on the horizontal axis.

To understand the timeframes required for system equilibration in this new set-up, a time-resolved analysis of the alignment kinetics was performed, following the evolution while separating the vertical and horizontal axes (Fig. 6.5).

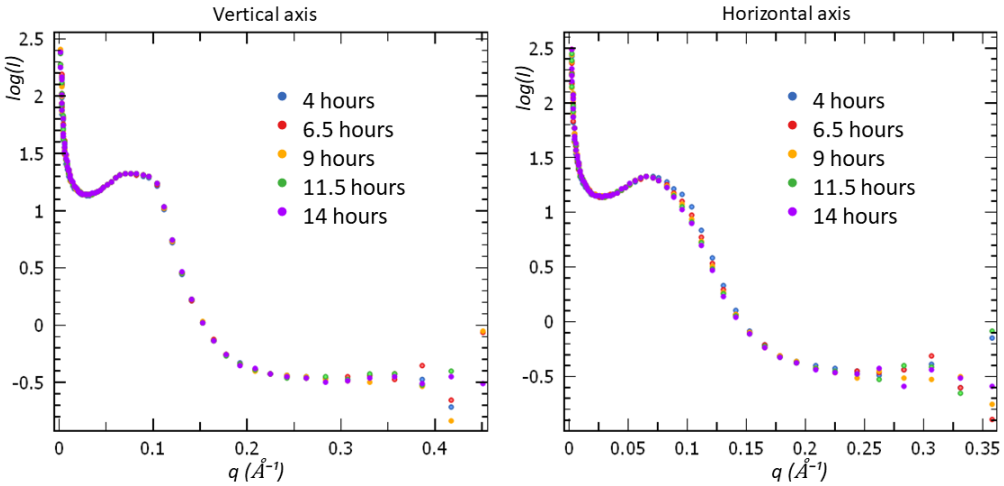


Figure 6.5 ***Time evolution in the 1D scattering profile.*** *The evolution of both the horizontal and vertical axes is shown separately, as a function of time. There are no variations in the vertical axis, while the peak in the horizontal axis gets narrower with time.*

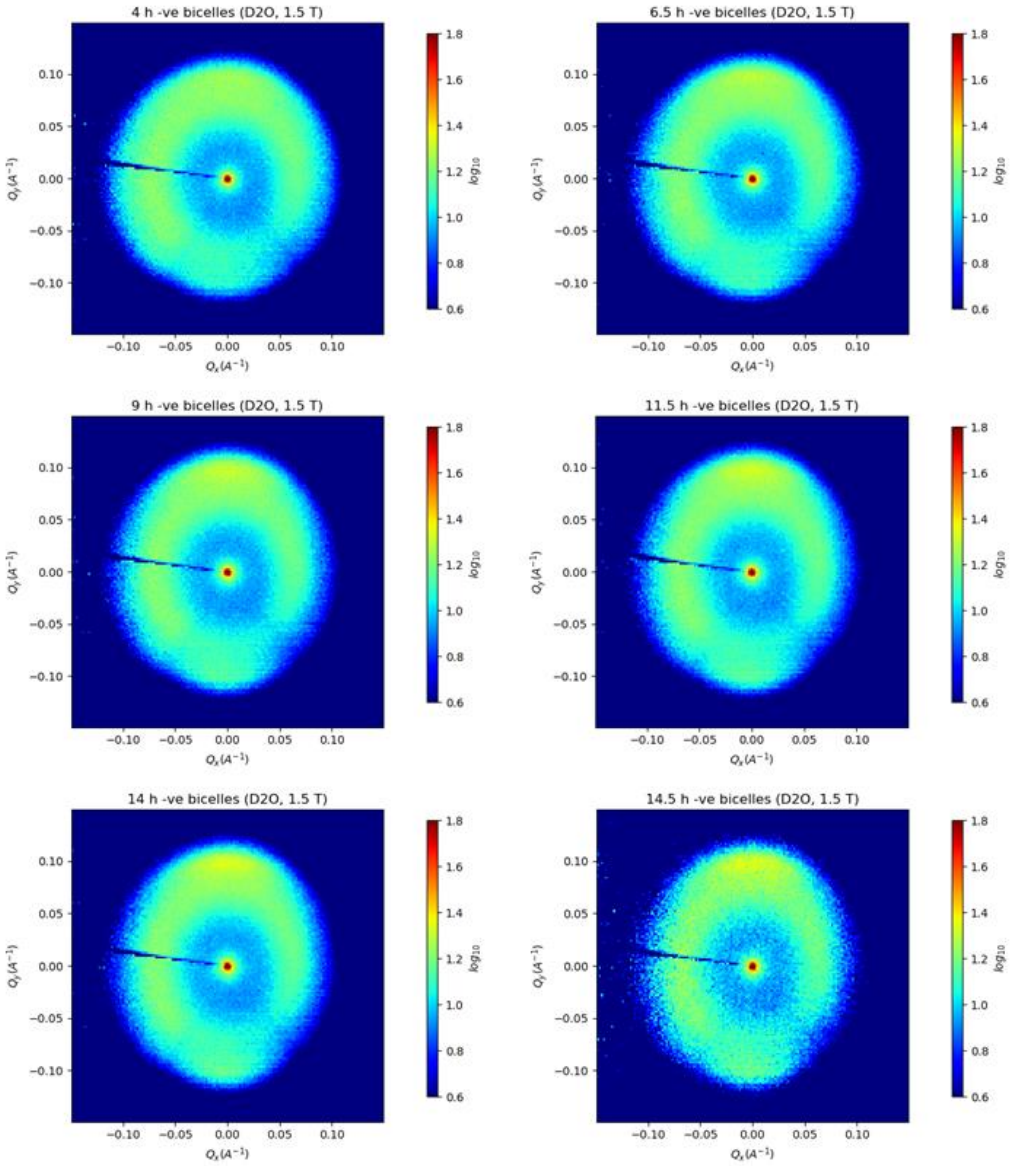


Figure 6.6 ***2D intensity maps of 'negative' bicelles.*** *The 2D intensity map is time resolved, showing the evolution of the system across 14 and a half hours. The DMPC/DHPC negative bicelles are subjected to a 1.5 T magnetic field and ~45 °C. Maximum alingment is reached after 11.5 hours of incubation in the external field.*

Figure 6.5 shows very little variation on the vertical axis, while horizontally the peak appears to narrow with time. This is also noticeable in the 2D profiles in Figure 6.6, where over time the intensity transitions from an axially symmetric

distribution to a more oblong shape as the horizontal axis narrows. To better evaluate this phenomenon, both axes were plotted at their corresponding time frame (Fig. 6.7) and the intensities on the horizontal axis were subtracted from the vertical axis, on an absolute scale (Fig. 6.8).

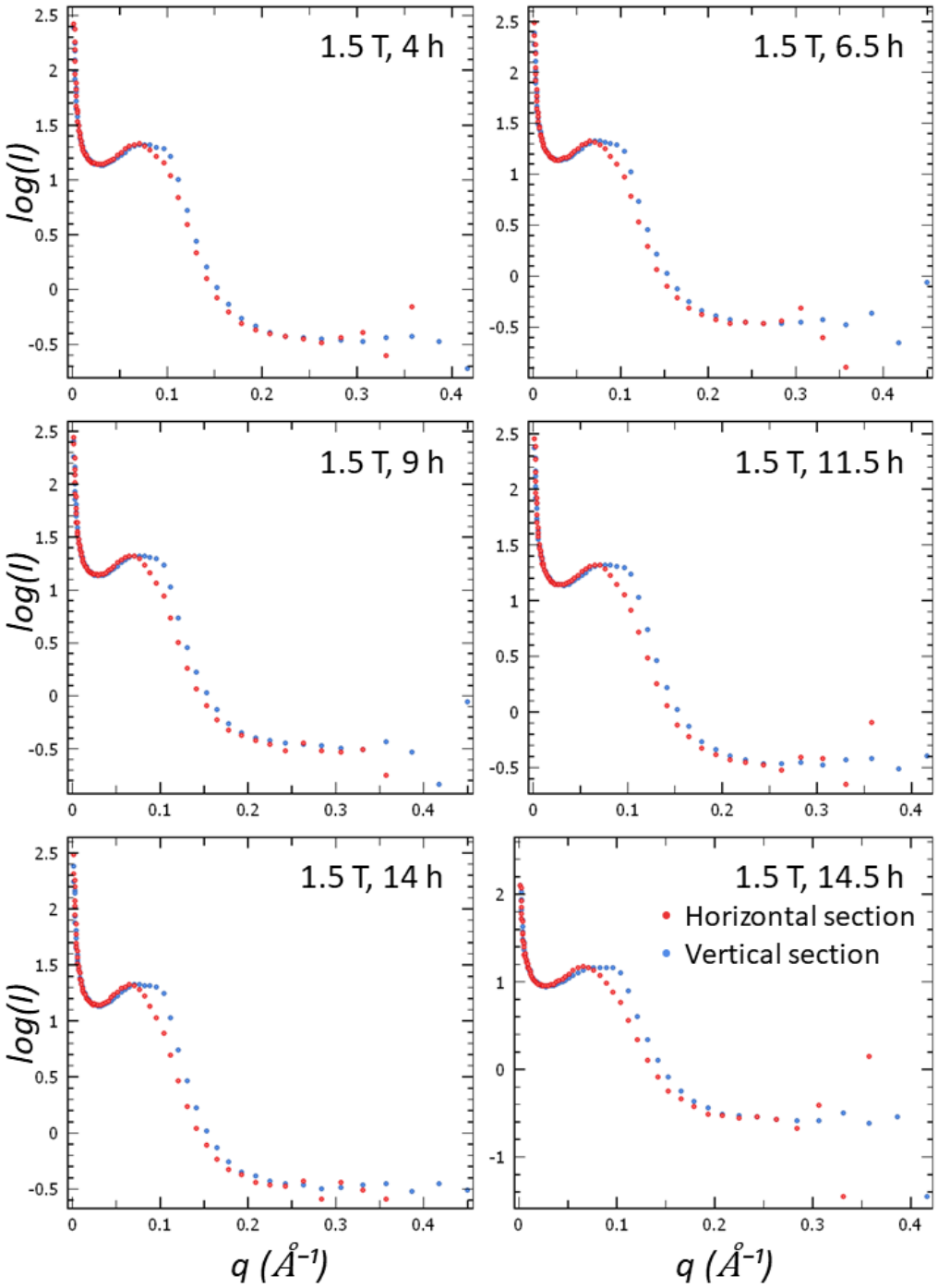


Figure 6.7 ***1D scattering profile of negative bicelles in magnetic field.*** *Evolution of the 1D scattering profile of negative bicelles in a 1.5 T magnetic field as a function of time. The vertical axis (blue) and horizontal axis (red) are shown individually.*

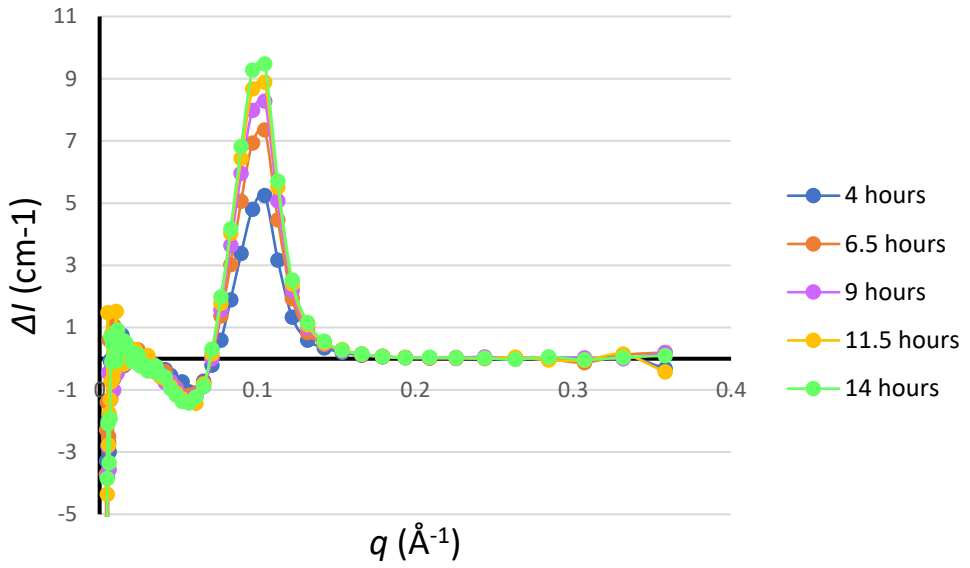


Figure 6.8 ***Intensity difference profile.*** *The absolute intensities on the horizontal axis were subtracted from the vertical axis to obtain the size distribution of the bicelle population. This allows to determine the thickness of the bilayer. The predominant response happens in the first 6.5 hours, with a slow but continuous increase from 9 hours onwards. The range is restricted to ≥-5 cm^{-1} to better visualise the peak of interest.*

With an horizontal magnetic field, the bicelles align as described in Figure 6.2. The measurements recorded along the horizontal axis are therefore representative of a single size population while the vertical axis represents a distribution of apparent sizes corresponding to all the arrangements the disc adopts with respect to its axis (as shown in Figure 6.9 Panel A). The diameter of the disc can be extrapolated from the maximum *q* on the vertical axis, which is constant after 9 hours at 0.06588 Å^{-1}. This corresponds to 95 Å. The peak in the vertical-horizontal subtraction, although varying in intensity, is consistently at 0.10455 Å^{-1}, which converts to 60 Å. This represents the thickness of the bicelle when laying "flat", or the thickness of the bilayer that forms it. The

bilayer thickness of DMPC as reported in the literature and determined by neutron reflectivity was 39 Å (Perino-Gallice et al., 2002) and 44.5 Å as determined by SANS (Kučerka et al., 2004), while observations of DMPC/CHAPSO bicelles by NMR reported a thickness of 51 Å (Sanders et al., 1994). This may imply that the bicelle has a slightly bulging structure instead of presenting a flat surface (Fig. 6.9 Panel B).

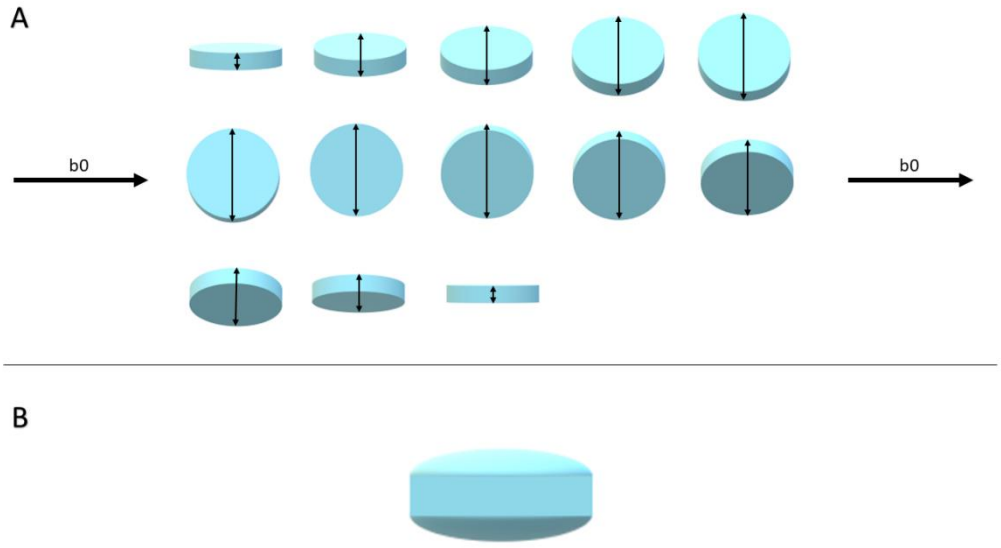


Figure 6.9 ***Schematics of (A) axial orientational distribution of bicelles in a horizontal magnetic field, and (B) a "bulging" bicelle.***

As part of the methodological development, the response of the system to temperature was also investigated. Data were acquired as the system reached the target temperature of ~45 °C and reversed to room temperature, to assess the equilibration time. Figure 6.10 shows the first (0-5 minutes) and last frame (95-100 minutes) of the scattering of negative bicelles which were cooled down to 4 °C prior to being inserted into the beamline. The temperature control was switched on as the sample was inserted , starting the process of heating from room temperature (estimated at approximately 18-20 °C) to 45 °C.

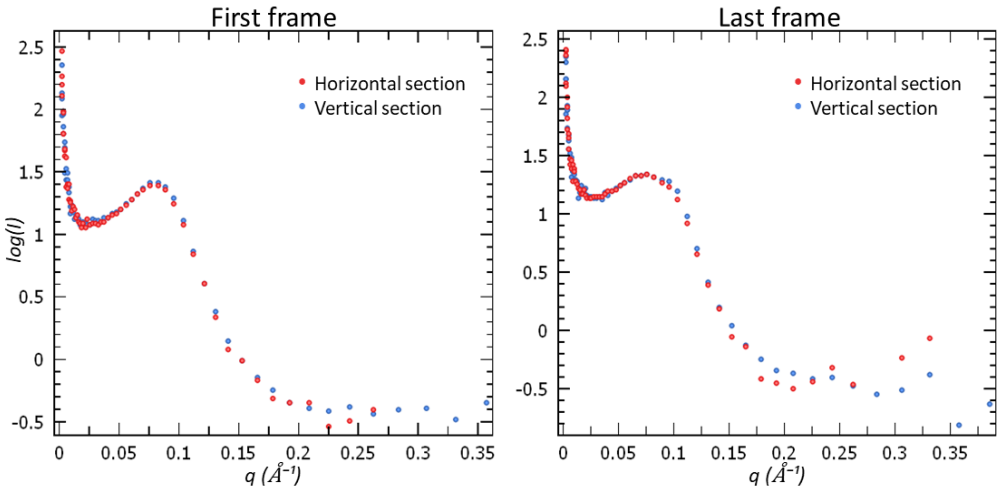


Figure 6.10 ***SANS profiles before, during and after the bicellar transition as temperature increases.*** *The first frame is the average scattering collected during the first 5 minutes following the start of the heating process. The last frame is the average of minutes 95-100. The horizontal and vertical axes are shown separately. The appearance of bicellar phase is evident in the last frame.*

Frames were collected every 5 minutes for a total of 100 minutes. The heating of the sample was complete after ~25 minutes (data not shown). This timeframe includes both factors involved: the heating of the sample environment and that of the sample itself; the latter is sample specific while the former entirely relates to the equipment. Another consideration is that the scattering profile only shows the time frame required for the sample to adopt a bicellar phase, which happens at ≥ 30 °C (Sternin et al., 2001), not reaching the temperature of 45 °C.

A similar acquisition was collected for the process of cooling down from 45 °C to ~20 °C (Fig. 6.11), which triggers the return to the liposomal phase, with the effect appearing in the first 10 minutes of the cooling process.

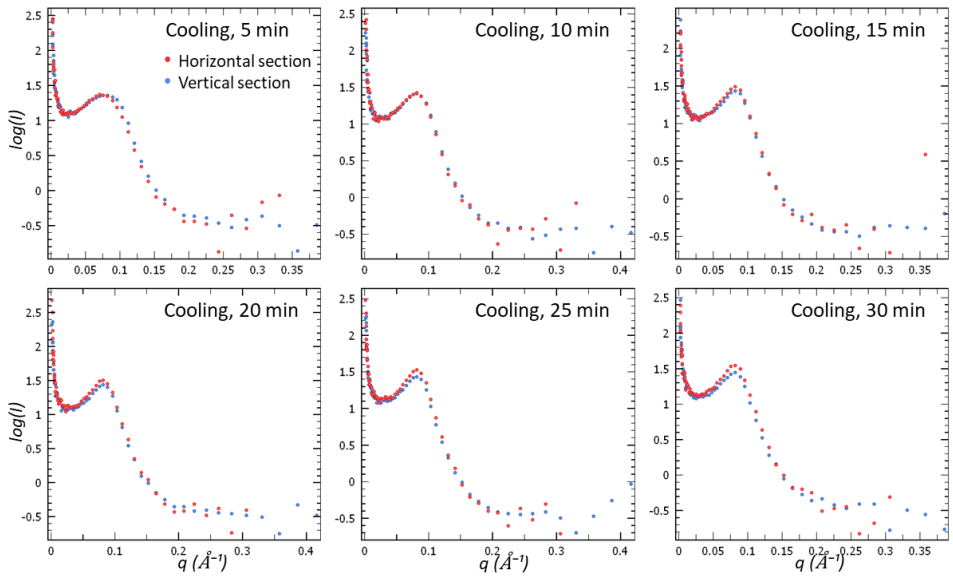


Figure 6.11 ***SANS profile of the system during the cooling process****. The process of cooling of negative bicelles is shown in 5 minute intervals. The horizontal and vertical axes are shown separately. There is partial structure loss already in the first 5 minutes, which is nearly complete at 10 minutes.*

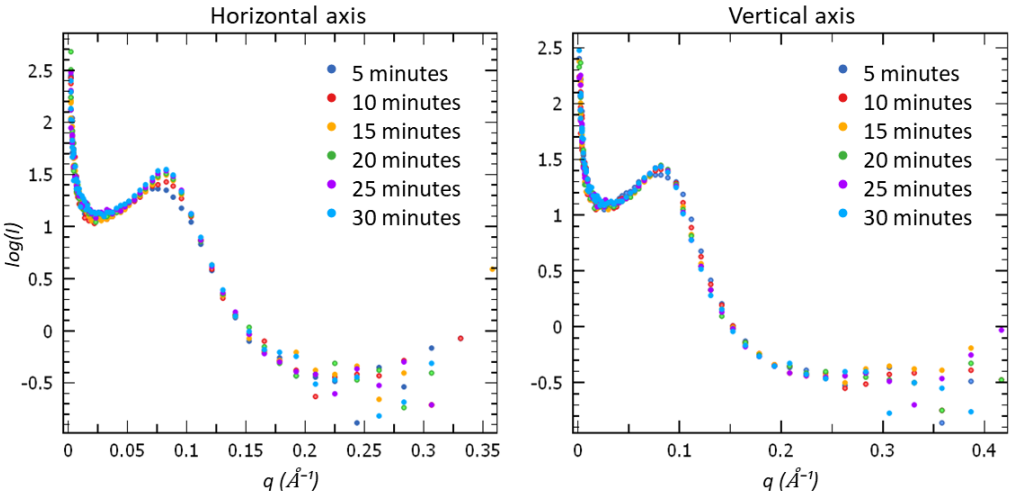


Figure 6.12 ***SANS profile of the system during the cooling process by axis.*** *The progression of the horizontal and vertical axes is followed individually as a function of time at 5 minute intervals. The horizontal axis, which is more characteristic of the bicellar environment (see above), is most affected by the heat loss.*

Figure 6.12 shows the evolution of the individual sections with cooling. The scattering in both axes changes with time, with the horizontal axis showing larger discrepancies. This is expected as the horizontal axis distinctly identifies the bicellar phase as opposed to the liposomal phase (as described above in Figures 6.4 and 6.5). The biggest variation happens in the first 15 minutes followed by a more stable state. This is consistent with what shown in Figure 6.11.

The next phase of the investigation looks at positive bicelles. The phase is flipped by the addition of lanthanides with positive magnetic susceptibility anisotropy, in our case Ytterbium chloride, which triggers the alignment of the plane normal with the magnetic field (Inbaraj et al., 2006). This behaviour has been extensively characterised by NMR and the working concentration of lanthanide was determined to be 5 mM. Our initial acquisition of a DMPC/DHPC/5 mM $YbCl_3$ system did not show alignment (Fig. 6.13).

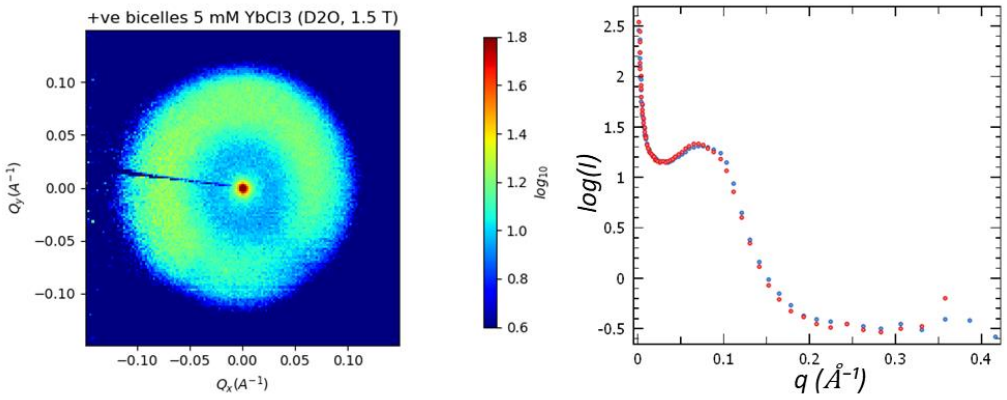


Figure 6.13 ***Effect of magnetic field on 'positive' bicelles.*** *DMPC/DHPC bicelles in a magnetic field in 2D (left) and 1D (right).*

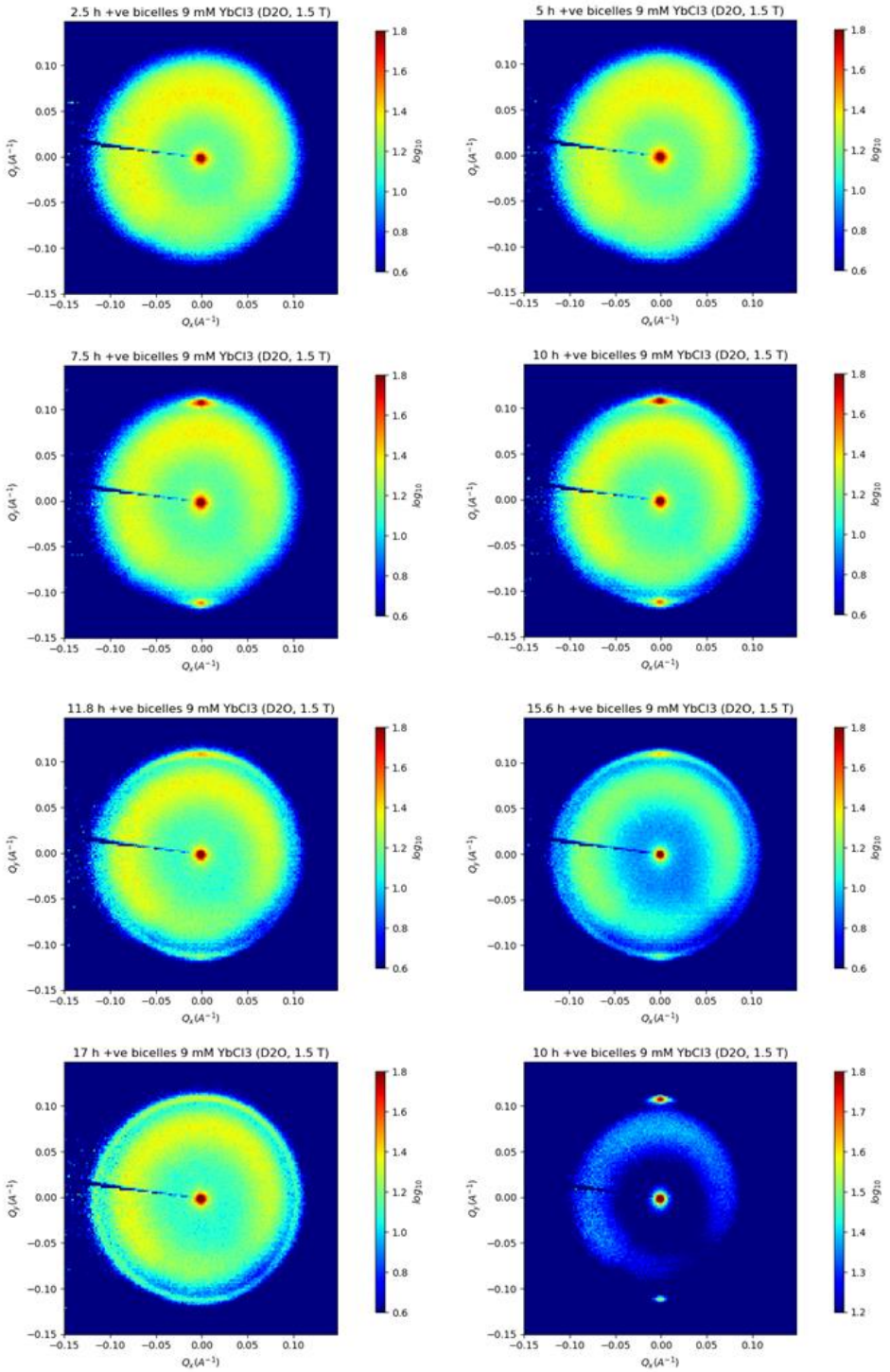


Figure 6.14 ***2D intensity map of DMPC/DHPC bicelles in the presence of 9 mM $YbCl_3$.*** *The 2D intensity map is time resolved, showing the evolution of the system across 17 hours. The DMPC/DHPC positive bicelles with 9 mM $YbCl_3$ are subjected to a 1.5 T magnetic field and ~45 °C. The lamellar stacking occurring from 7.5 hours manifests as increased intensity on the vertical axis, to*

then gradually disappear from 11.8 hours. The sample at 10 hours is shown in the last graph in a different colour scale for visualisation purposes.

Upon addition of 9 mM $YbCl_3$ (total concentration) the behaviour of the system in 1.5 T (and 45 °C) is very time dependent. There is no indication of bicellar phase in the first 5 hours of incubation in the magnetic field, while a marked asymmetry appears at 7.5 hours. This is representative of specular intensity given by Bragg reflections arising from ordered structures. The vertical axis shows an increase in intensity at both poles, and while the south pole does not appear as bright as the north, this can be explained by the masking present on the right hand corner due to instrumentation. This vertical dominance is retained after 10 hours but starts to fade at 11 hours and 50 minutes. A overall decrease in intensity is observed after 15 hours and 35 minutes, which transitions into a fairly uniform Bragg behaviour by 17 hours (Fig. 6.14).

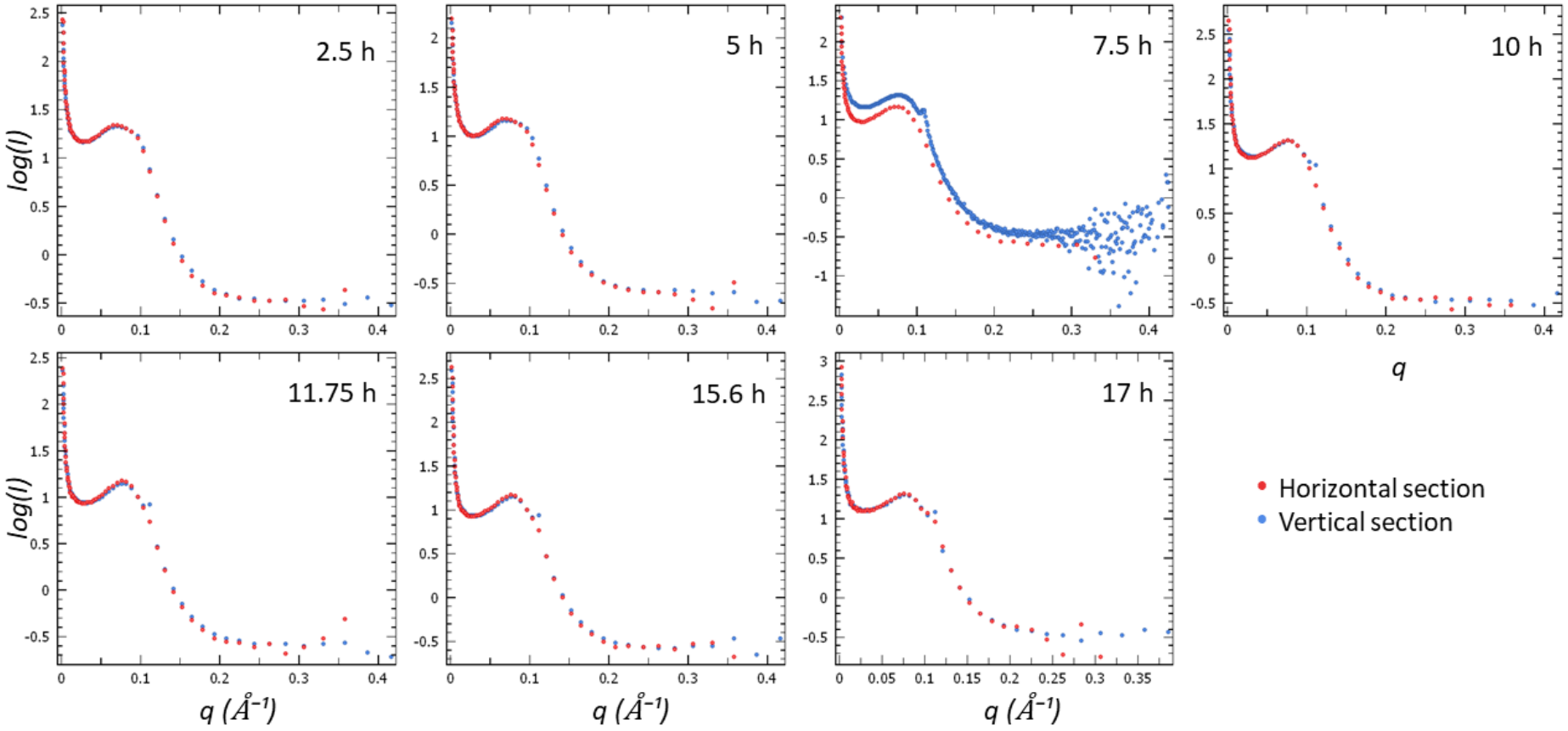


*Figure 6.15 '**Positive' bicelles with 9 mM $YbCl_3$.** Time resolved evolution of positive bicelles with 9 mM $YbCl_3$ in a 1.5 T magnetic field, 45 °C. The appearance of a Bragg peak at 7.5 hours shows the bilayer stacking on the vertical axis.*

The 1D profiles also show an event in the 7.5 hour measurement, with an intensity increase in the vertical axis and the formation of a secondary, Bragg peak at 0.1 $Å^{-1}$. While the behaviour is retained for at least another 2.5 hours in the 2D profiles, this observation in the 1D scattering is only present in the 7.5 hour window (Fig. 6.15).

While an asymmetry between the vertical and horizontal axis and the appearance of Bragg peaks is expected as the sample is subjected to a magnetic field, the orientation of intensity increase is opposite to the expectations (Fig. 6.16). In a horizontal field, a vertical arrangement is observed.

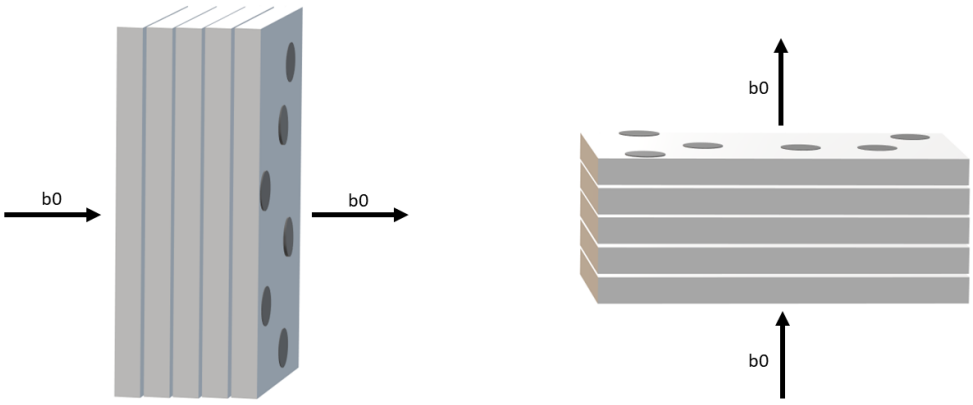


Figure 6.16 ***Schematic of stacked positive bicelles.*** *The expected orientation of the stacked positive bicelles in a horizontal field (left) and vertical field (right) is is shown, with the bilayers represented in light grey and the discoidal holes given by the short-chain lipids in dark grey.*

Upon further addition of Ytterbium chloride (17 mM), the system is subject to a total rearrangement (Fig. 6.17).

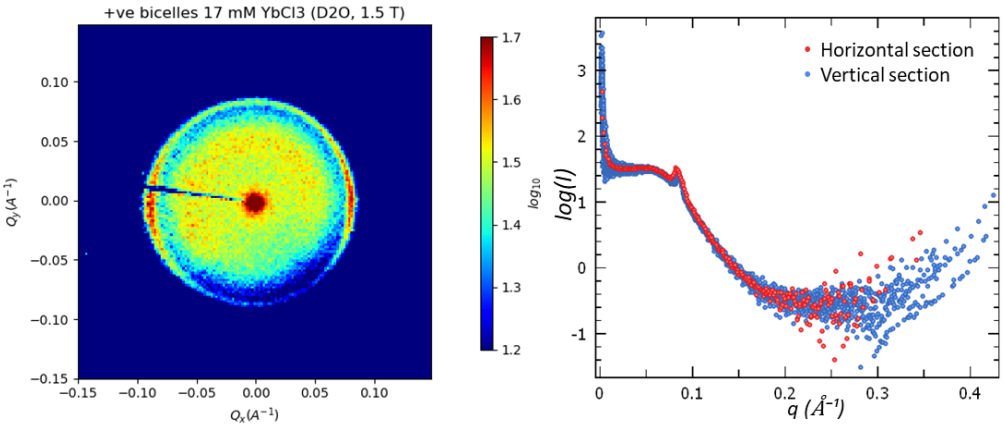


Figure 6.17 ***2D intensity map of positive bicelles with high $YbCl_3$.*** *A higher Ytterbium chloride concentration (17 mM) is applied to the positive bicellar system, subjected to 1.5 T, 45 °C. The strong diffraction on the horizontal plane indicates the stacked bilayers.*

The high concentration of $YbCl_3$ triggers a rearrangement in favour of the horizontal axis. Overall this behaviour suggests that, in relatively low concentrations of Ytterbium chloride, the lamellar arrangement is mostly due to gravitational forces instead of magnetic. Much higher concentrations of lanthanide ions are needed to overcome gravity and trigger magnetical alignment. This phenomenon has never been reported in the literature, likely because of the forced vertical orientation of NMR magnets which made it impossible to discriminate between magnetic and gravitational effects.

To test the orientation of protein within the positively aligned bicelles, HmbRID94N was incorporated into the bicellar system. *Haloarcula marismortui* bacteriorhodopsin I is a transmembrane protein that functions as a light-sensitive proton pump. It has 7 transmembrane helices and preferentially adopts a trimeric conformation (Fig. 6.18). HmbRID94N was chosen because of

its well-characterised structure and behaviour, relative ease of handling and its characteristic pink/purple colouration shown when correctly folded and active.

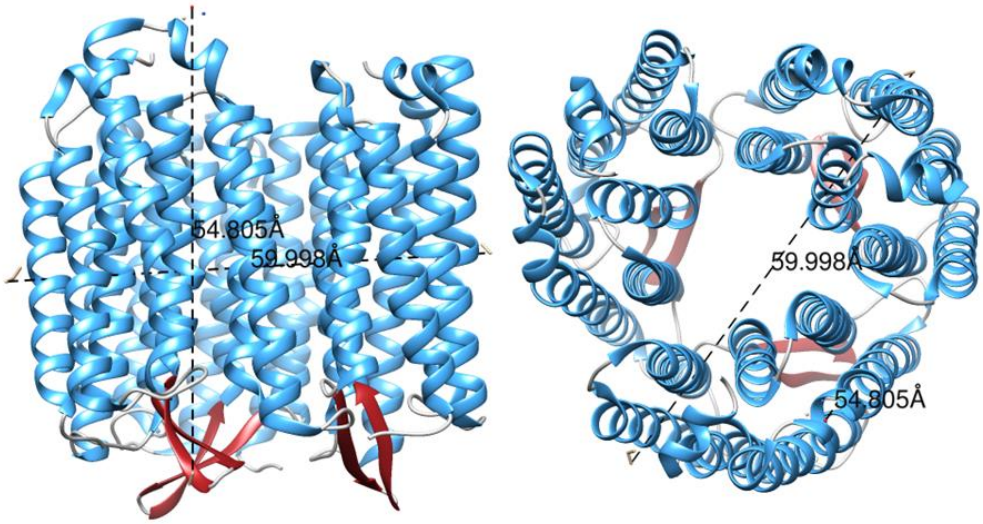


Figure 6.18 ***Structure of the HmbRID94N trimer.*** *Side (left) and top (right) view of trimeric HmbRID94N (PDB entry 4PXK). The vertical and lateral dimensions were calculated using UCSF Chimera 1.15 and are 55 Å and 60 Å respectively. DOI: Shevchenko, V., Gushchin, I., Polovinkin, V., Round, E., Borshchevskiy, V., Utrobin, P., Popov, A., Balandin, T., Buldt, G., Gordeliy, V. (2014) PLoS One 9: e112873-e112873* [*https://doi.org/10.2210/pdb4PXK/pdb*](https://doi.org/10.2210/pdb4PXK/pdb) *Literature citation: Shevchenko, V., Gushchin, I., Polovinkin, V., Round, E., Borshchevskiy, V., Utrobin, P., Popov, A., Balandin, T., Buldt, G., Gordeliy, V. (2014) Crystal Structure of Escherichia coli-Expressed Haloarcula marismortui Bacteriorhodopsin I in the Trimeric Form. PLoS One 9: e112873-e112873* [*https://doi.org/10.1371/journal.pone.0112873*](https://doi.org/10.1371/journal.pone.0112873)

Ytterbium chloride was once again used to flip the phase at a concentration of 5 mM. The protein was perdeuterated and the experiment performed in 12.5% D_2O

buffer to contrast match the lipidic component. The recorded scattering is therefore corresponding to the protein (Fig. 6.19).

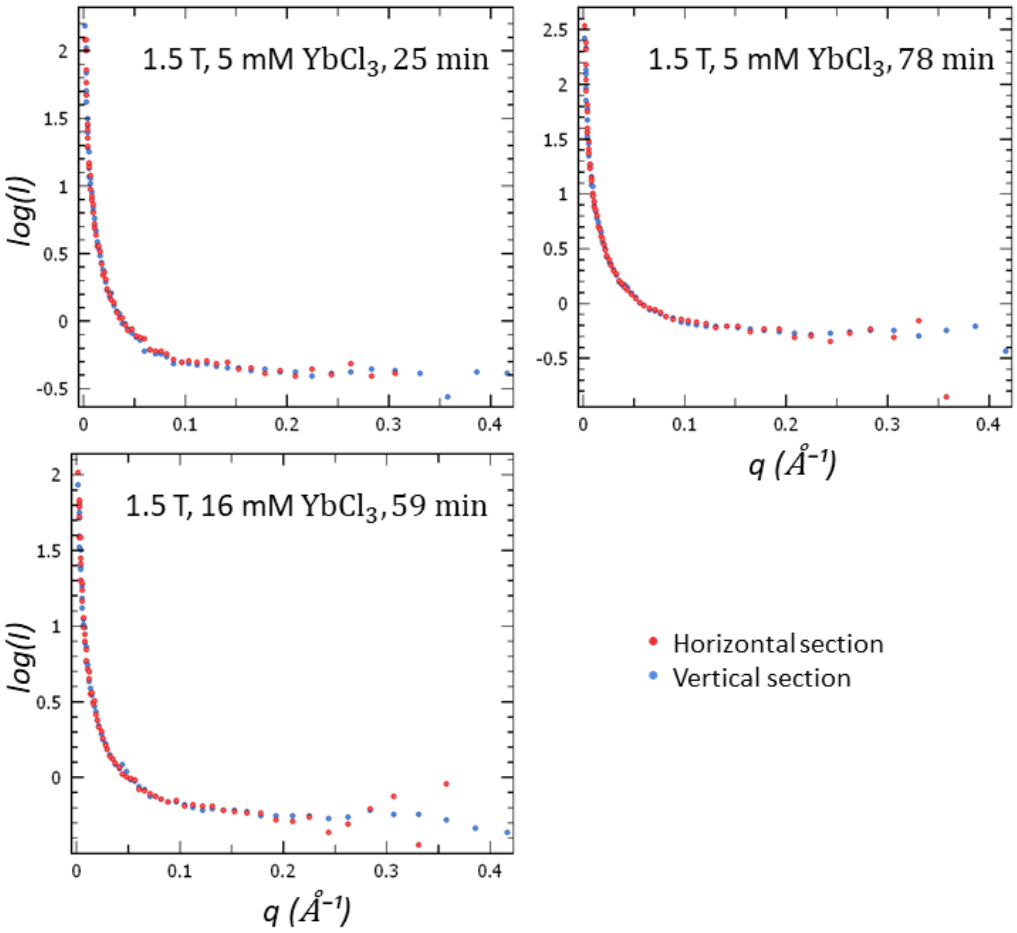


Figure 6.19 ***Scattering ^{2}H-HmbRID94N in positive bicelles.*** *^{2}H-HmbRID94N was incorporated into positive bicelles by adding 5 mM and 16 mM $YbCl_3$. The 1D scattering profiles recorded at 25 minutes and 78 minutes from inserting the sample into the beamline are shown on the top, while the sample with higher $YbCl_3$ concentration is at the bottom.*

Data were collected for approximately 78 minutes and averaged for the first 25 minutes and the subsequent 53 minutes (Fig. 6.19). Unfortunately the data collected are not of sufficient quality to carry out a comprehensive analysis. This is due to the region at low *q* showing larger structures or aggregates and therefore

masking important information for protein scattering processing, including size and molecular weight estimation. There is no appreaciable scattering variation between the horizontal and vertical axes but this cannot be ascertained due to the lack of information in the low *q* region.

Upon visual inspectionof the sample, it became apparent that the protein sample had degraded and lost the characteristic pink colouration (Fig. 6.20).

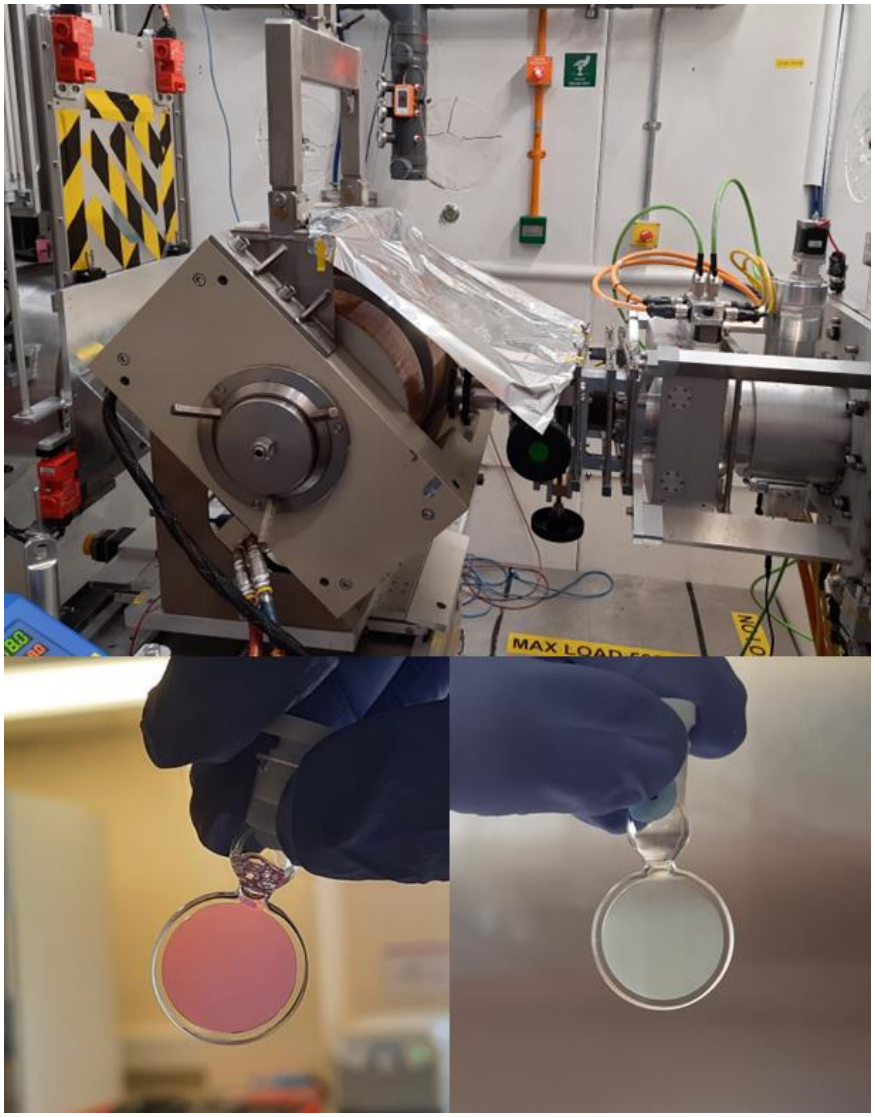


Figure 6.20 ***Experimental set-up*** *to protect the light sensitive sample (top) and the discolouration of the HmbRID94N sample following acquisition. The folded protein before (left) and unfolded after (right) acquisition.*

6.2.Discussion

The use of bicelles in NMR and EPR studies of membrane proteins has made them an important asset for NMR and EPR spectroscopy studies. The intrinsic vertical magnetic field of NMR makes for an ideal environment to exploit the magnetic properties of lipids and align embedded membrane proteins.

The possibility of extending this methodology to SANS studies was explored, to ultimately characterise the overall morphology of membrane proteins by contrast-matching the surrounding lipidic environment. A new sample environment, including a 1.5 T magnetic field and a heating system, was purposefully designed to for SANS studies of biological samples in solution. The formation, alignment and stability of DMPC/DHPC bicelles were characterised, before loading the membrane protein HmbRI into the system.

The alignment of negative bicelles in the 1.5 T field was found to be time-dependent, reaching its maximum after 14 hours of incubation at ~45 °C. This is evidenced by an increase in intensity on the vertical axis and a narrowing horizontal profile. This is due to the size distribution given by the rotational freedom of the bicelles along the horizontal axis of alignment, as shown in Panel A of Figure 6.9.

The vertical alignment of positive bicelles is not consistent with the literature, which suggests that addition of 5 mM lanthanide ions is sufficient to trigger alignment with the magnetic field. With 9mM $YbCl_3$, a sharp Bragg peak was observed on the vertical axis, which indicates a horizontal stacking of bilayers. This phenomenon was observed after 7.5 hours of incubation. Upon further addition of lanthanide ions (17 mM final concentration) the arrangement of the

system flipped into a vertical membrane stacking along the magnetic field, with a secondary population of vesicles.

The magnetic field is opposing the gravitational effect, and this phenomenon suggests that, while 9 mM $YbCl_3$ is enough to trigger the formation of a lamellar system, a field of 1.5 T is not enough to overcome the effect of gravity on the system. The addition of high concentration lanthanide ions aids the horizontal field in flipping the system to a vertical stacking. The timescale of events supports this theory: while the horizontal stacking was only achieved after a long incubation period, the addition of 17 mM $YbCl_3$ immediately flips the phase (< 1 hour).

Prolonged measurements, coupled with the brightness of the environment, denatured the light-sensitive protein despite our best efforts. Optimisation needs to be carried out and measurements on a bicelles-embedded membrane protein will have to be repeated to establish whether there is protein alignment.

In this experiment SANS measurements of a magnetically aligned bicellar suspension were carried out for the first time. Controlled magnetic alignment was achieved both across and along the field and observed and characterised a stacked bilayer assembly, previously associated with magnetic torques, now shown to be gravitationally self-assembled. Overall, these results are encouraging and show promise both in terms of further development of this methodology for the study of membrane proteins and for the use of this type of sample environment in other types of biological investigations by Small Angle Scattering.

7. Additional work

This section described additional work, relevant to the overall project but not the investigation described above. This includes the cloning, expression and purification of the N-terminal AgrA, which will serve for further CD and SAXS measurements and is a candidate for structure determination by solution-state NMR, and purification of the C-terminal AgrA (following Dr Bardelang's protocol), which was included in the CD investigation (not presented here), in the 'SAXS 4 Landau' analysis and in the IQS-related investigation.

The data analysis of SAXS data of AgrB with and without AgrD is also shown here and relates to the submitted manuscript (See Publications).

Other work includes CD measurements of GAPDH, the RNA binding protein discussed in Chapter 5.3, alone and in complex with total RNA extract and poly-A. The data are not presented here because have not yet been analysed as the presence of RNA introduces additional challenges to the processing of CD data.

7.1. C-terminal AgrA

The C-terminal domain of AgrA is at present the only domain to have been successfully crystallised and resolved (Sidote et al., 2008). This domain was investigated as method validation for the Landau free energy expansion of small angle scattering data.

Our approach towards the C-terminal AgrA is similar to the one adopted for the full-length C119S AgrA. The expression and purification protocols had been previously developed (previous unpublished work by Dr Phil Bardelang) and are described in the Methods.

7.1.1. Expression

Unlike the full-length protein, the expression of the C-terminal AgrA does not require high concentrations of sorbitol to ensure solubility.

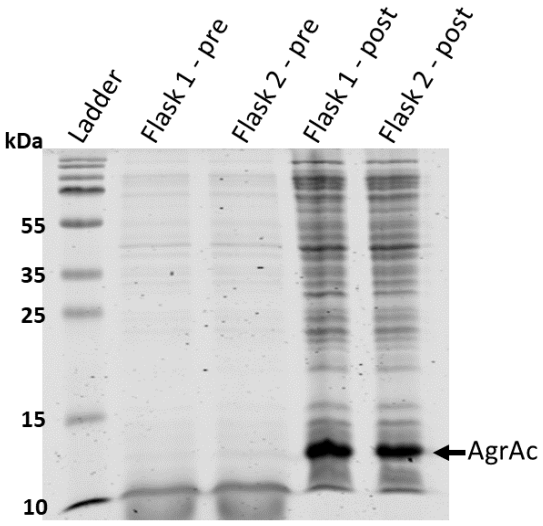


Figure 7.1.1 ***Expression of C-terminal AgrA.*** *Samples pre- and post-induction from an expression in duplicate are shown on a 18% SDS PAGE gel. The protein is approximately 15 kDa in size.*

The expression yield is much higher for this protein compared to its full-length equivalent and can be explained by the absence of sorbitol.

7.1.2. Purification

The purification of recombinant C-terminal AgrA was carried out as described in the Methods.

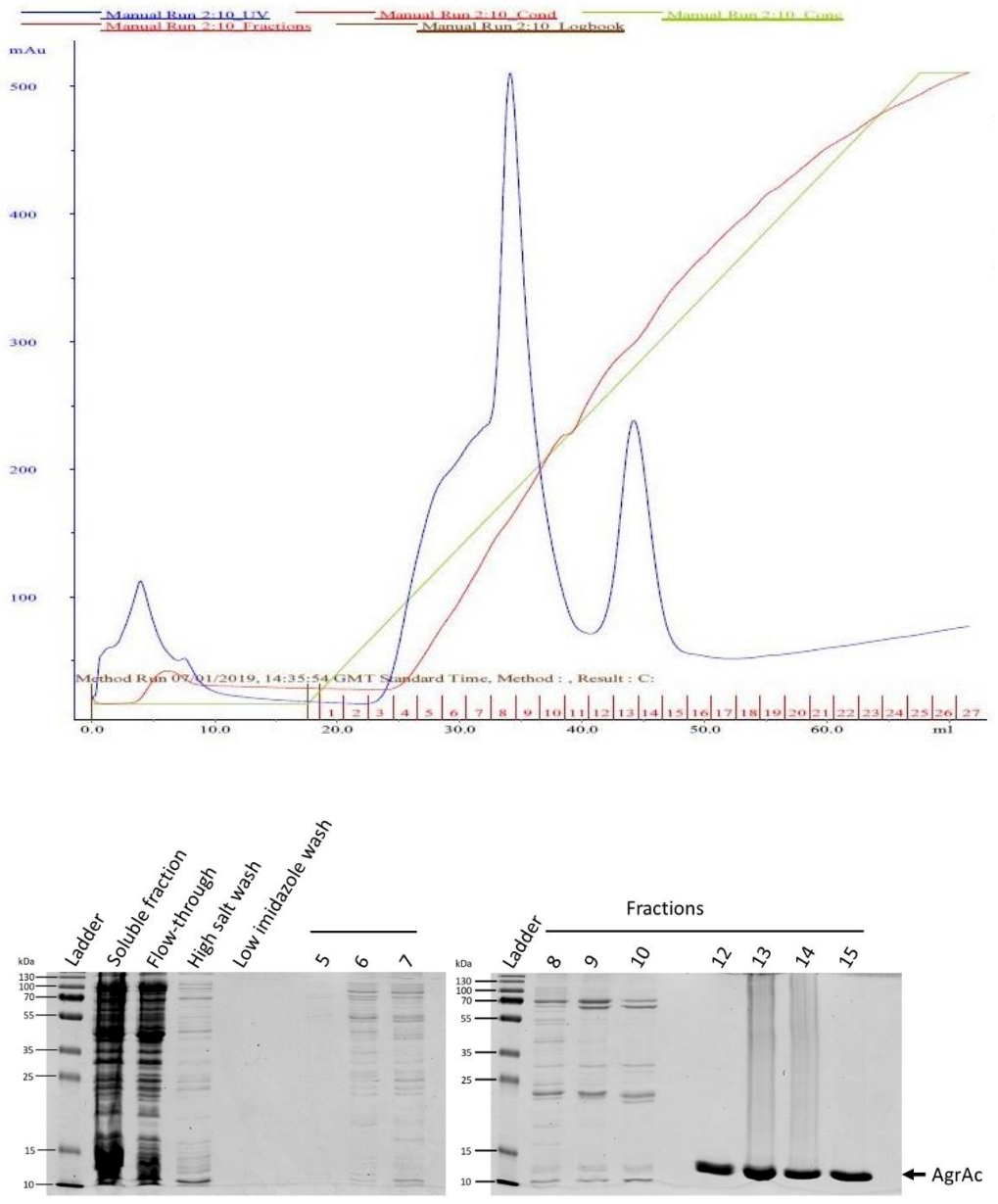


Figure 7.1.2 ***C-terminal AgrA purification.*** *18% SDS PAGE gels showing the fractions eluted by FPLC. C-terminal AgrA was eluted in fractions 12-15.*

As shown in Figure 7.1.2, the recombinant C-terminal AgrA elutes at approximately 50% high imidazole and in an even cleaner fashion than the full-length protein.

With this degree of purity, the eluted sample does not require further purification procedures such as size exclusion chromatography.

7.1.3. Activity assay

In assessing the activity of this recombinant protein, the same rationale described for the full-length protein was followed. According to the proposed quorum sensing mechanism, the domain interacting with the promoters P2 and P3 is the C-terminus of AgrA and therefore its binding to P2 is assessed as activity indicator.

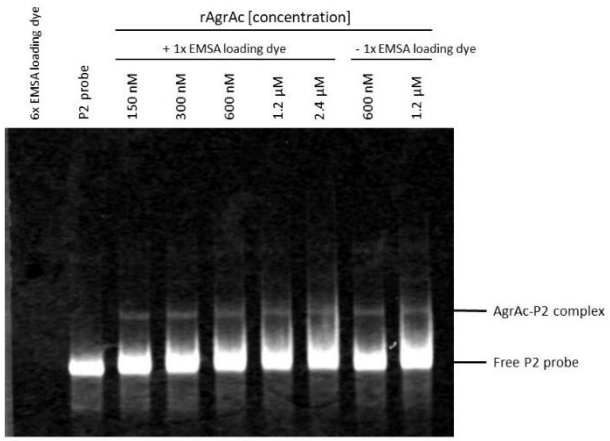


Figure 7.1.3 ***Activity assay of recombinant C-terminal AgrA.*** *Electrophoretic mobility shift assay showing the shift in the run of fluorescent P2 probe upon interaction with recombinant C-terminal AgrA.*

Figure 7.1.3 shows the shift in the size of the fluorescently-tagged P2 probe resulting from the interaction with the recombinant C-terminal AgrA. This confirms the protein is active.

7.2.N-terminal AgrA

The N-terminal domain of AgrA, which, according to the proposed mechanism, interacts with AgrC, receives the phosphate and drives dimerisation, has not yet been structurally characterised. This work is preparatory for future structural characterisation by solution state NMR.

7.2.1. Cloning

Cloning of the N-terminal domain of AgrA was carried out using the pCOLD-I vector already containing the full-length C199S AgrA as a template, using a restriction digest reaction. The end product was verified by colony PCR and sequencing (see Methods for more details of the protocol). The work was carried out with the help of Dr Phil Bardelang.

7.2.2. Expression

In light of the need for high concentrations of sorbitol of the full-length recombinant protein, but also lack thereof for the C-terminal domain, the recombinant N-terminal AgrA was expressed both in presence and absence of 1 M D-sorbitol to see how it affects expression and purification.

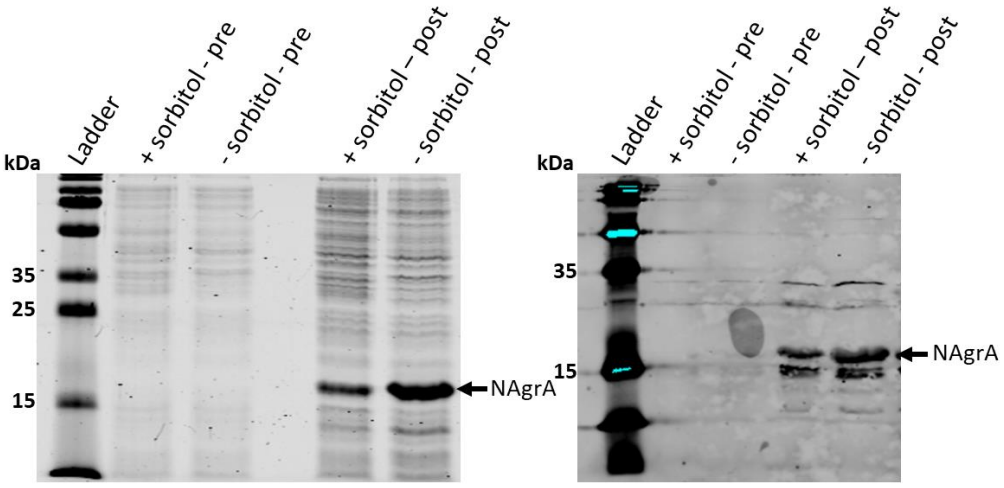


Figure 7.2.1 ***N-terminal AgrA expression test ± sorbitol.*** *(Left) 18% SDS PAGE gel showing pre- and post-induction samples in presence and absence of 1 M sorbitol. (Right) Western blot showing pre- and post-induction samples in presence and absence of 1 M sorbitol.*

The expression of N-terminal AgrA was successful both in presence and absence of sorbitol (Fig. 7.2.1). The amount of expressed recombinant protein is sensibly reduced in presence of high sorbitol due to the reduced cell growth under stress conditions.

Protein expression was also tested in M9 minimal medium, in preparation for possible future structure determination by solution-state NMR (Fig. 7.2.2).

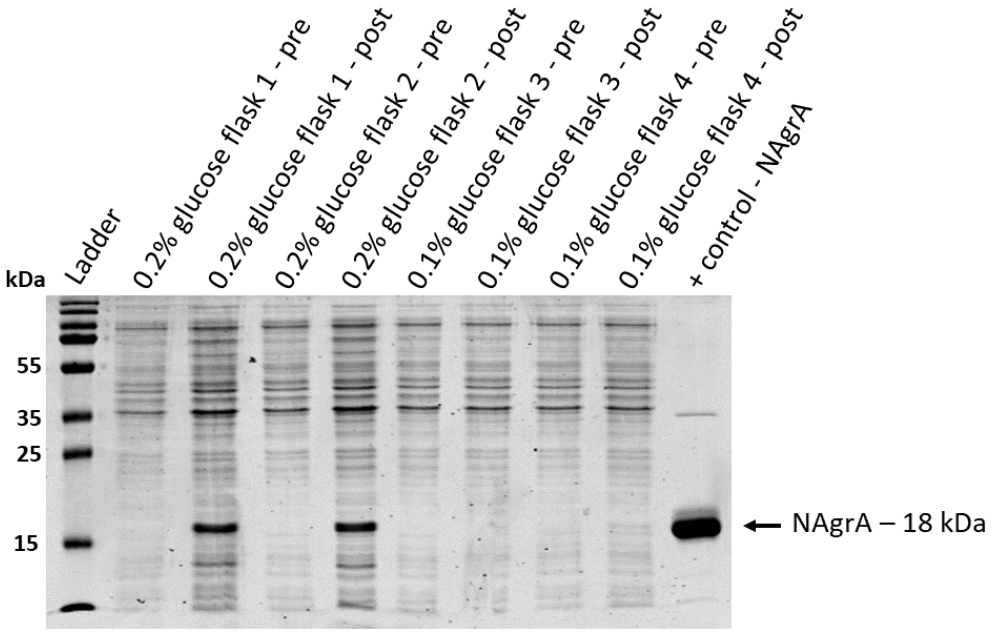


Figure 7.2.2 ***Expression of N-terminal AgrA in minimal medium.*** *18% SDS PAGE gel showing pre- and post-induction samples of N-terminal AgrA expression in M9 minimal medium. Two glucose concentrations, 0.2 and 0.1%, were tested. Expression was successful only with 0.2% glucose.*

The expression in M9 minimal medium was carried out using 0.1 and 0.2% glucose. Knowing sensitivity to glucose in minimal medium expression is useful in case deuteration is required. Deuteration involved the use of pyruvate as a glucose substitute and, as pyruvate carries fewer carbons and is therefore less efficient as a nutrient, the minimal required amount of glucose can be used as a reference to calculate how much pyruvate to use. N-terminal AgrA does not express at all in 0.1% glucose, while 0.2% gives good expression.

7.2.3. Purification

Figure 7.2.3 in the previous section shows how the expression of recombinant N-terminal AgrA was successful in both presence and absence of high sorbitol concentrations, with a reduced yield in presence of sorbitol likely due to the

effect of stress on the cell growth. There was still no evidence, however, of the solubility of the overexpressed recombinant protein. Therefore, N-terminal AgrA was purified from both lysates, with and without sorbitol.

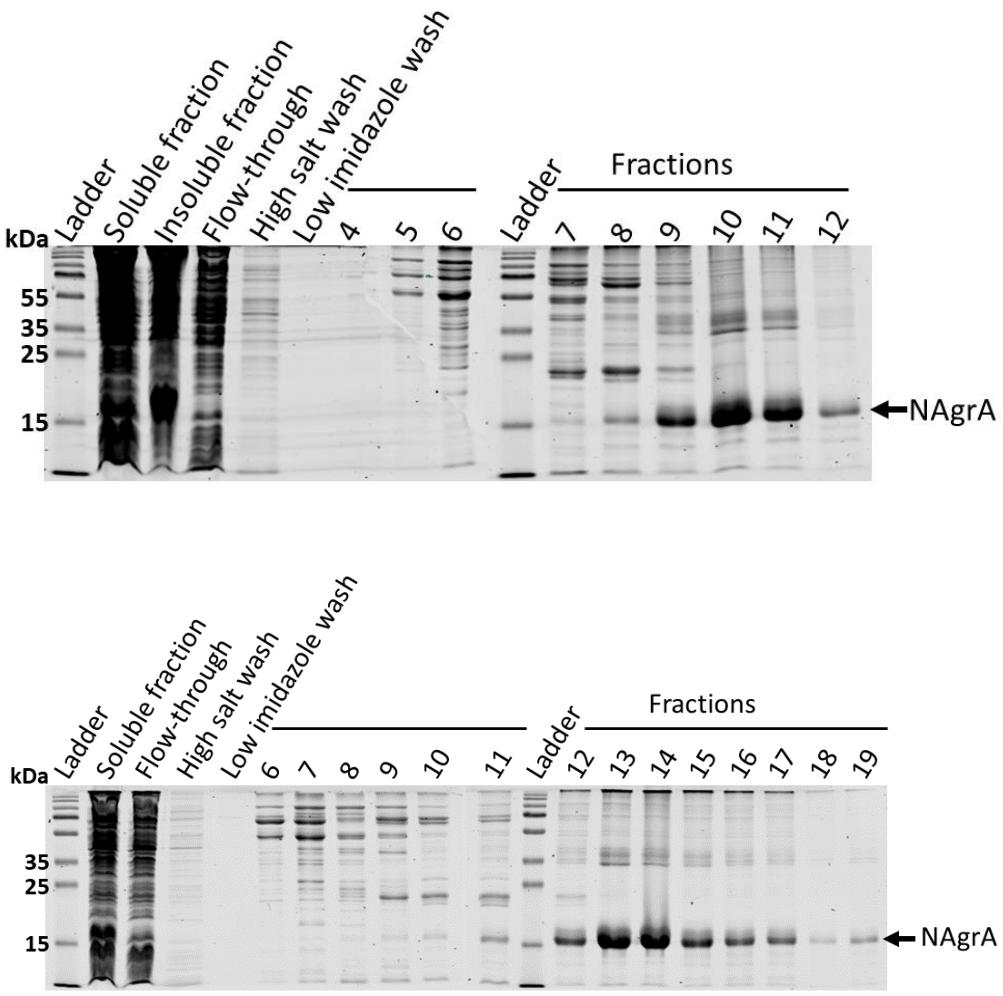


Figure 7.2.3 ***N-terminal AgrA purification.*** *The purification of N-terminal AgrA, expressed with (top) and without (bottom) sorbitol, does not yield a high purity sample. Samples were run on 18% SDS PAGE gels.*

The level of purity of the eluted fractions is not as high as the full-length and C-terminal equivalent. It is possible that some of the contaminants at ~35 kDa are the protein in dimeric form. One of the objectives of the study of this protein is to resolve its structure by solution-state NMR, and for these reasons the eluted

fractions of both ± sorbitol cultures were further purified by size exclusion chromatography (Fig. 7.2.4; size exclusion chromatography was performed by Dr Philip Bardelang).



Figure 7.2.4 ***Size exclusion chromatography of N-terminal AgrA.*** *(A) UV trace from the AKTAPure of size exclusion chromatography of N-terminal AgrA sample expressed in presence of sorbitol and initially purified by IMAC. (B) UV trace from the AKTAPure of size exclusion chromatography of N-terminal AgrA sample expressed in absence of sorbitol and initially purified by IMAC.*

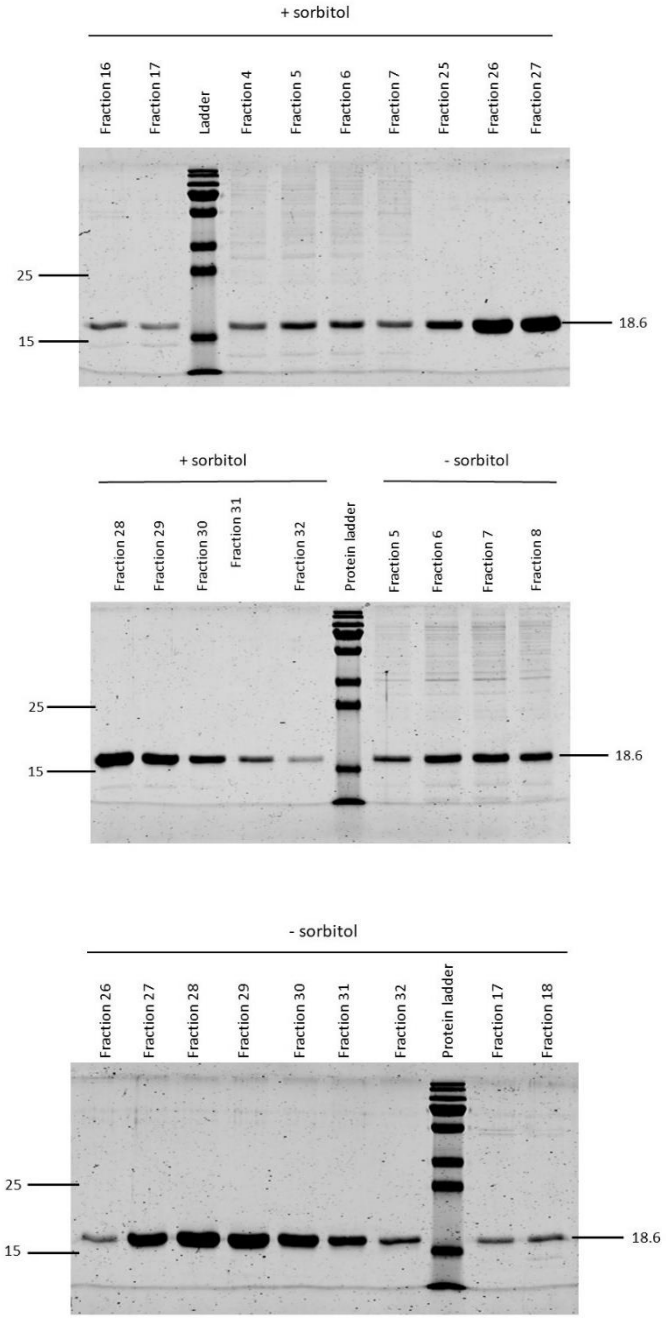


Figure 7.2.5 ***Size exclusion chromatography of N-terminal AgrA.*** *18% SDS PAGE gels showing the eluted fractions from the size exclusion chromatography, of recombinant N-terminal AgrA expressed in presence and absence of high sorbitol and initially purified by IMAC.*

The protein samples obtained from the size exclusion chromatography are pure in both cases to an appropriate level for solution-state NMR. There is some loss of sample eluted in the earlier fractions together with contaminants (Fig. 7.2.5).

7.3.AgrB

The membrane protein AgrB serves as transmembrane endopeptidase in the *agr* system and proteolytically processes the pro-peptide AgrD. In this section, a qualitative comparison of the dimensionality of AgrB in presence and absence of the interaction partner AgrD was performed by SAXS. Please note that this analysis was recently published here: (Bardelang et al., 2023). The data processing was carried out using SasView 5.0.5 (Doucet et al., 2021) but presents some limitations. The study of membrane proteins using X-rays presents a major obstacle represented by the detergent or lipid layer surrounding the protein. X-rays are non-penetrative and therefore the bulk of the signal will be representative of the detergent/lipid shell. Another factor to consider is the non-perfect buffer subtraction. In the case of membrane protein embedded in a detergent micelle, a portion of the detergent will reconstitute into empty micelles. It is impossible to know what percentage will contain the protein of interest and therefore the scatterer cannot identically replicate the ratio of empty micelles in the buffer used for subtraction.

With these considerations in mind, the data was fitted using a core shell ellipsoid model (Kotlarchyk and Chen, 1983, Berr, 1987) (Fig. 7.3.1). This allowed to extrapolate the dimensions of the core, which represents the protein pocket, and the thickness of the shell (or the micelle belt surrounding the protein). Sample

preparation was done by Dr Alice Goode with the aid of Dr Philip Bardelang (Bardelang et al., 2023).

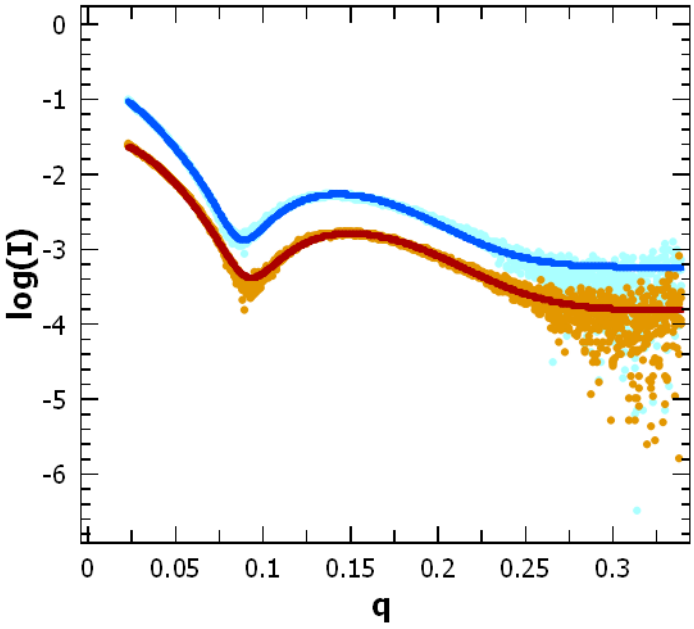


Figure 7.3.1 ***Scattering pattern of AgrB ± AgrD in DDM micelles****. The experimental data for AgrB alone are shown in light blue and corresponding data fitting in dark blue. These are compared to the AgrBD/AgrD complex, for which experimental data are shown in orange and data fitting in red. The fits were obtained using SasView5.0.5 and visualised using PRIMUS (Konarev et al., 2003). The data for the fit were restricted to a q-range starting from approximately 0.023715 A^{-1}. The data and corresponding fits are scaled for clarity. This figure was published in Bardelang P, Murray EJ, Blower I, Zandomeneghi S, Goode A, Hussain R, Kumari D, Siligardi G, Inoue K, Luckett J, Doutch J, Emsley J, Chan WC, Hill P, Williams P and Bonev BB (2023) Conformational analysis and interaction of the Staphylococcus aureus transmembrane peptidase AgrB with its AgrD propeptide substrate. Front. Chem. 11:1113885. doi: 10.3389/fchem.2023.1113885. Copyright © 2023*

Bardelang, Murray, Blower, Zandomeneghi, Goode, Hussain, Kumari, Siligardi, Inoue, Luckett, Doutch, Emsley, Chan, Hill, Williams and Bonev. https://creativecommons.org/licenses/by/4.0/ (Disclaimer in original publication: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.).

The SAXS data show that the oligomeric state of AgrB is not affected by the presence of AgrD. The radius of the core is reported at 10.60 ± 0.05 Å for AgrB alone and 10.51 ± 0.04 Å for the AgrB/D complex (Bardelang et al., 2023). These values suggest a monomeric protein but this is in disagreement with wet-lab data which suggest AgrB forms a stable dimer (Bardelang et al., 2023). It is possible that the scattering originating from the micelle altered the dimensionality values of the protein. For this reason, the analysis is reported from a comparative point of view.

AgrD also induces a decrease in the shell thickness, bringing it from 24.0 ± 0.1 to 22.3 ± 0.1 Å, which suggests partial deconstruction of the micellar belt (Bardelang et al., 2023).

These data were submitted for publication (see Publications).

8. Discussion

This work combines different biophysical techniques and especially focuses on developing data analysis to further characterise the proteins of *S. aureus* quorum-sensing system, the accessory gene regulator system. More specifically, the structure, mode of action and structural stability of AgrA and AgrC, which represent the two-component system, are investigated. AgrA, the transcriptional response regulator, has two domains, the C-terminal DNA-binding domain, for which a crystal structure is available (Sidote et al., 2008, Leonard et al., 2012, Rajasree et al., 2016), and the N-terminal domain, which is thought to interact with the phosphorylated histidine kinase AgrC and accept the phosphate group on Asp59. This results in dimerisation of AgrA, which in turn increases its affinity for the *agrP3* promoter (Khodaverdian et al., 2013). No topological or structural information is at present available for this domain. By using a combination of Small Angle X-ray Scattering and MD simulations, an all-atom model of full-length AgrA was obtained from MD trajectories and validated against experimental data. SAXS experiments allowed to obtain tertiary structure data while keeping the protein sample in solution at room temperature and thus providing a realistic representation of its state in near-native conditions. The DNA-binding domain shows a lot of flexibility, which in the MD simulations unravels the predominant β-sheet arrangement identified in the crystal structure. CD data of this domain alone (not shown) support the crystal structure in the amount of strands present, so it is possible that this flexibility is only observed when the protein is whole. The N-terminal domain adopts a conformation equally shared between helices and strands (topology shown in Fig. 4.1.16), in a helix-strand repeated arrangement. The strands are

preferentially positioned in the inside of the protein, surrounded by the helices. The model selection was carried out first by cluster analysis of the MD trajectories and then refined using the Ensemble Optimisation GAJOE (Tria et al., 2015, Bernadó et al., 2007). This functionality allows the selection of PDB files which best describe the experimental scattering data, in this instance representing each frame of the MD trajectory. While the proteins of the *agr* system are well-defined in structure, this feature is particularly useful when applied to the study of very flexible or intrinsically disordered proteins by giving an account of the various conformations adopted and in what percentage.

The same analysis was carried out on the histidine kinase AgrC. This construct ($AgrC_{206}$) is characterised only by the cytoplasmic portion of the protein, namely the dimerisation and histidine phosphotransfer domain (DHp) and the catalytic domain (CA). The transmembrane domain, which transduces the signal across the cell membrane, was excluded from this study due to the constraints and challenges involved in the study of membrane proteins. The cytoplasmic domains are directly involved in the interaction of AgrC with the phosphodonor ATP first and with AgrA. While the interaction between the two proteins has been inferred based on function and kinetic studies, an *in vitro* complex has never been isolated. The sequence of events in the acquiring of the phosphate group ATP and transfer to AgrA has not been elucidated: the potential options are shown in Figure 5.4.1 and relate to the oligomeric state of the two proteins throughout the process. Thanks to SAXS data, the dimeric state of $AgrC_{206}$ was established and an experimentally validated all-atom model proposed.

The identification of a molecular complex between AgrA and $AgrC_{206}$ has proven challenging. While SAXS data suggest monodispersity and MONSA is

able to model the complex, the complex determination function (ScÅtterIV) is not able to detect the complex. Size exclusion chromatography shows the two proteins elute separately from the SEC column. Based on the established function-derived assumption that AgrA and AgrC do indeed interact (Srivastava et al., 2014) and that the soluble nature of AgrA makes it impossible for the interaction to occur with the transmembrane domain of AgrC, we infer that the contact between the two proteins is transient and cyclic. The transient nature of protein-protein interactions in bacterial two-component systems is not unusual and has been observed in different organisms, including *Bacillus subtilis* phosphorelay mechanism (Zapf et al., 2000) and *Thermotoga maritima* (Casino et al., 2009). This explains why it is not retained when assessed by SEC. Based on this hypothesis, the role of ATP in this interaction was explored. A non-hydrolysable analogue of ATP, sATP, was also employed to elucidate on the kinetics of this exchange. If AgrA is recruited to AgrC together with the nucleotide, and unable to bind in absence of it, then the addition of either ATP or sATP should trigger a change. If AgrA is not released until phosphorylated, then this change is retained and measurable. If AgrA cannot be recruited until the ADP is released from AgrC, the recruitment would be halted in presence of sATP. Of course if this interaction is indeed fleeting and independent on the phosphorylation state of either protein, then no changes would be observed in any case. This investigation was approached using a combination of wet lab (SEC) and biophysical techniques, namely CD, SAXS and SANS. This type of study is the ideal application of contrast-matching in SANS for protein work, as it allows to detect any potential gross structural rearrangement occurring in either protein resulting from the interaction with the other. This overcomes the

limitations of SAXS in being able to discriminate scattering signal coming from two or more binding partners. MONSA provides an option to model the two proteins in complex but only provides a low-resolution representation of the interaction.

Deuteration of AgrA was not achievable due to the need for high concentrations of sorbitol to express the protein in soluble form, while $AgrC_{206}$ was successfully deuterated. SANS revealed that both ATP and its analogue affect the structural arrangement of AgrA when interacting with $AgrC_{206}$. These changes are observed both at low q (~0.2 $Å^{-1}$) and at higher q (0.6 $Å^{-1}$), suggesting that alterations occur both on the overall arrangement of the protein and at a smaller level. The H_2O contrast showed that the system, with preferential visualisation of $AgrC_{206}$, is affected by the presence of sATP but not by ATP. These results suggest that the non-hydrolysable ATP causes rearrangements in both proteins, while ATP only affects AgrA, perhaps causing dimerisation, and increasing its degree of folding. Circular Dichroism data were acquired under similar conditions. The change observed in AgrA in presence of sATP is substantiated by CD, which shows a decrease in transition cooperativity and a more extended transition. $AgrC_{206}$ is a highly stable protein with a comparatively high τ value. This dramatically changes in presence of ATP and sATP, both of which cause a drop in the transition stability. This compares well with the SAXS data (Chapter 4.2) in the case of the sATP, which shows a better folded but less compact protein, but is not substantiated for the ATP data set. The transition cooperativity value of the AgrA-$AgrC_{206}$ mixture lies outside of the bracket created by the two individual proteins, which in itself indicates some form of molecular interaction.

The addition of ATP affects this only with a marginal increase, which becomes more significant in presence of sATP. This corroborates the SANS observations.

To further elucidate the sequence of events in the interaction, as discussed above, the mixture of the two proteins was assessed in different molar ratios: 1:1 and 1:2 (AgrA:AgrC). The former implies that either the AgrC dimer is disrupted upon interaction with AgrA, or that two AgrA monomers bind AgrC at once. The latter accounts for the dimeric state of $AgrC_{206}$ and would allow an individual AgrA monomer to interact solely with $AgrC_{206}$. The samples in 1:2 ratio (1:1 is discussed above) have a higher τ value compared to the 1:1 equivalent, which most probably accounts for the increased presence of the stable $AgrC_{206}$. The combination in presence of ATP and sATP shows however a slightly lower value than the 1:1 mixture.

The tertiary conformational rearrangement in AgrA observed by SANS in presence of both ATP and sATP, taken together with the CD data, confirmed that AgrC is required for AgrA to acquire the phosphate donated by ATP. sATP, on the other hand, interacts with both AgrA and $AgrC_{206}$ both in a complex (SANS) and separately (CD).

While the CD investigation did not allow to establish the 'correct' assembly mode of the complex, it provided good data for the optimisation of the Landau series expansion formalism to account for systems that present two transitions: an initial slight structure alteration at lower temperatures and a total loss of structure at high temperatures which represents the more evident denaturation often identified by the Boltzmann analysis. The bead model of the AgrA/C complex obtained by MONSA suggests the correct stoichiometry to be 1:1,

which corresponds to Scenarios 2 or 4 in Figure 5.4.1. This finding would not be surprising as it would coincide with the shown stoichiometry of AgrC and AIP (Cisar et al., 2009, Wang et al., 2014).

Temperature denaturation studies were conducted on these systems by SAXS for the first time, with the aim to apply our temperature-induced transition model to these data. The adaptation of the Landau formalism to SAXS data was successful and implemented by using the radius of gyration R_g as order parameter, replacing the secondary structure content used in CD. While this method will require further validation on a wider array of proteins, these results are encouraging and may increase the use of Small Angle Scattering in the field of Structural Biology.

Chapter 6 describes a study that is not related to the system discussed throughout this work. It was conducted as a sample environment project linked to my iCASE studentship partially funded by the ISIS Neutron and Muon Source. While magnetism studies are routinely carried out on ISIS beamlines, the magnets implemented are incompatible with the sample environment required by protein samples. In this instance, a magnet was specifically designed to investigate the use of magnetically aligned bicelles in the structural characterisation of membrane proteins by SANS. The possibility to contrast-match the lipid component surrounding the protein makes SANS a very attractive possibility for the study of proteins which are difficult to crystallise and are too small for Cryo-EM studies. The magnetic properties of bicelles make this an alignable system allowing to extract information on the dimensionality of the embedded protein. The bicellar system was successfully oriented, in both its negative and positive forms. Incubation times were established to obtain optimal alignment and gravity was determined as an important factor. This was observable in our setup

for the first time as the main technique used for the investigation of bicelles is NMR, which has an inherent vertical field. These results are encouraging and warrant further investigation and method development.

This work shows an array of solutions for the structural characterisation of proteins in near-native conditions. It is our hope that some of this analysis will, in the future, be combined with more popular techniques such as crystallography and NMR to provide an all-rounded representation of protein characteristics and generally increase the application of Biophysics in the field of Structural Biology.

9. Conclusion and future work

This work showed how different biophysical techniques can be brought together to structurally characterise protein systems. A pipeline of sorts was created, encompassing SAXS and related data analysis models, SANS, CD, wet lab techniques and *in silico* molecular dynamics simulations in a novel, comprehensive way. These methodologies were applied to *S. aureus* accessory gene regulator system, a system which will still benefit from further characterisation.

This study would benefit from additional work. Areas to consider for the future direction of this research include:

- Structure determination of N-terminal (or full-length) AgrA by solution state NMR;
- Structure determination of full-length AgrC by solid-state NMR;
- Optimisation of his-tag cleavage on AgrA (data not shown) and further investigation into the buffer stability of AgrA;
- Investigating the phosphorylation of AgrC and AgrA by NMR – this might provide insights as to the inability to retain the complex of the two proteins;
- Binding site mapping by carbene footprinting of the AgrA-$AgrC_{206}$ complex;
- Additional MD modelling of the AgrA-AgrC-ATP system to untangle the order of events and identify timescales.

In terms of methodological development:

- Expanding the Landau formalism for SAXS to include a wider set of proteins;
- Collecting temperature-dependent SAXS data on I22 – potentially carry out some sample environment development to accommodate for single-use capillaries or cuvettes (similar to what is used on the DL-SAXS beamline), which would allow easier handling of proteins in the denatured state and prevent capillary clogging;
- The ‘Landau for SAXS’ could potentially be built into the software that is going to be written for CD data.

References

ALFORD, A., KOZLOVSKAYA, V. & KHARLAMPIEVA, E. 2017. Small Angle Scattering for Pharmaceutical Applications: From Drugs to Drug Delivery Systems. *In:* CHAUDHURI, B., MUÑOZ, I. G., QIAN, S. & URBAN, V. S. (eds.) *Biological Small Angle Scattering: Techniques, Strategies and Tips.* Singapore: Springer Singapore.

ASHENBERG, O., KEATING, A. E. & LAUB, M. T. 2013. Helix bundle loops determine whether histidine kinases autophosphorylate in cis or in trans. *J Mol Biol,* 425**,** 1198-209.

BADER, M. S. 2006. Staphylococcus aureus bacteremia in older adults: predictors of 7-day mortality and infection with a methicillin-resistant strain. *Infect Control Hosp Epidemiol,* 27**,** 1219-25.

BARDELANG, P., MURRAY, E. J., BLOWER, I., ZANDOMENEGHI, S., GOODE, A., HUSSAIN, R., KUMARI, D., SILIGARDI, G., INOUE, K., LUCKETT, J., DOUTCH, J., EMSLEY, J., CHAN, W. C., HILL, P., WILLIAMS, P. & BONEV, B. B. 2023. Conformational analysis and interaction of the Staphylococcus aureus transmembrane peptidase AgrB with its AgrD propeptide substrate. *Front Chem,* 11**,** 1113885.

BARON, S. 1996. *In:* BARON, S. (ed.) *Medical Microbiology.* Galveston (TX): University of Texas Medical Branch at Galveston

Copyright © 1996, The University of Texas Medical Branch at Galveston.

BASHAM, M., FILIK, J., WHARMBY, M. T., CHANG, P. C. Y., EL KASSABY, B., GERRING, M., AISHIMA, J., LEVIK, K., PULFORD, B. C. A., SIKHARULIDZE, I., SNEDDON, D., WEBBER, M., DHESI, S. S., MACCHEROZZI, F., SVENSSON, O., BROCKHAUSER, S., NARAY, G. & ASHTON, A. W. 2015. Data Analysis WorkbeNch (DAWN). *Journal of Synchrotron Radiation,* 22**,** 853-858.

BASSLER, B. L. 2002. Small talk. Cell-to-cell communication in bacteria. *Cell,* 109**,** 421-4.

BERNADÓ, P., MYLONAS, E., PETOUKHOV, M. V., BLACKLEDGE, M. & SVERGUN, D. I. 2007. Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering. *Journal of the American Chemical Society,* 129**,** 5656-5664.

BERNAL, J. D. & CROWFOOT, D. 1934. X-Ray Photographs of Crystalline Pepsin. *Nature,* 133**,** 794-795.

BERR, S. S. 1987. Solvent isotope effects on alkytrimethylammonium bromide micelles as a function of alkyl chain length. *The Journal of Physical Chemistry,* 91**,** 4760-4765.

BLACKWELL, J. R. & HORGAN, R. 1991. A novel strategy for production of a highly expressed recombinant protein in an active form. *FEBS,* 295**,** 10-12.

BLEVINS, J. S., GILLASPY, A. F., RECHTIN, T. M., HURLBURT, B. K. & SMELTZER, M. S. 1999. The staphylococcal accessory regulator (sar) represses transcription of the Staphylococcus aureus collagen adhesin gene (cna) in an agr-independent manner. *Molecular Microbiology,* 33**,** 317-326.

BONEV, B. B. & MORROW, M. R. 1997. Effect of pressure on the dimyristoylphosphatidylcholine bilayer main transition. *Physical Review E,* 55**,** 5825-5833.

BOOTHROYD, A. T. 2020. *Principles of Neutron Scattering from Condensed Matter*, Oxford University Press.

BROOKS, B. R., BROOKS III, C. L., MACKERELL JR., A. D., NILSSON, L., PETRELLA, R. J., ROUX, B., WON, Y., ARCHONTIS, G., BARTELS, C., BORESCH, S., CAFLISCH, A., CAVES, L., CUI, Q., DINNER, A. R., FEIG, M., FISCHER, S., GAO, J., HODOSCEK, M., IM, W., KUCZERA, K., LAZARIDIS, T., MA, J., OVCHINNIKOV, V., PACI, E., PASTOR, R. W., POST, C. B., PU, J. Z., SCHAEFER, M., TIDOR, B., VENABLE, R. M., WOODCOCK, H. L., WU, X., YANG, W., YORK, D. M. & KARPLUS, M. 2009. CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry,* 30**,** 1545-1614.

BROSEY, C. A. & TAINER, J. A. 2019. Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology. *Current Opinion in Structural Biology,* 58**,** 197-213.

BUBECK WARDENBURG, J., PATEL, R. J. & SCHNEEWIND, O. 2007. Surface proteins and exotoxins are required for the pathogenesis of Staphylococcus aureus pneumonia. *Infect Immun,* 75**,** 1040-4.

BURGER, V. M., ARENAS, D. J. & STULTZ, C. M. 2016. A Structure-free Method for Quantifying Conformational Flexibility in proteins. *Scientific Reports,* 6**,** 29040.

CAPRA, E. J. & LAUB, M. T. 2012. Evolution of Two-Component Signal Transduction Systems. *In:* GOTTESMAN, S., HARWOOD, C. S. & SCHNEEWIND, O. (eds.) *Annual Review of Microbiology, Vol 66.* Palo Alto: Annual Reviews.

CASINO, P., RUBIO, V. & MARINA, A. 2009. Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction. *Cell,* 139**,** 325-36.

CHACÓN, P., DÍAZ, J. F., MORÁN, F. & ANDREU, J. M. 2000. Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. *J Mol Biol,* 299**,** 1289-302.

CHAN, W. C., COYLE, B. J. & WILLIAMS, P. 2004. Virulence regulation and quorum sensing in staphylococcal infections: competitive AgrC antagonists as quorum sensing inhibitors. *J Med Chem,* 47**,** 4633-41.

CHEN, I. & GOTSCHLICH, E. C. 2001. ComE, a competence protein from Neisseria gonorrhoeae with DNA-binding activity. *Journal of Bacteriology,* 183**,** 3160-3168.

CISAR, E. A., GEISINGER, E., MUIR, T. W. & NOVICK, R. P. 2009. Symmetric signalling within asymmetric dimers of the Staphylococcus aureus receptor histidine kinase AgrC. *Molecular Microbiology,* 74**,** 44-57.

CORNELIS, P. 2020. Putting an end to the Pseudomonas aeruginosa IQS controversy. *Microbiologyopen,* 9**,** e962.

COUSIN, F. 2015. Small angle neutron scattering. *EPJ Web of Conferences,* 104**,** 49.

COWIESON, N. P., EDWARDS-GAYLE, C. J. C., INOUE, K., KHUNTI, N. S., DOUTCH, J., WILLIAMS, E., DANIELS, S., PREECE, G., KRUMPA, N. A., SUTTER, J. P., TULLY, M. D., TERRILL, N. J. & RAMBO, R. P. 2020. Beamline B21: high-throughput small-angle X-ray scattering at Diamond Light Source. *Journal of Synchrotron Radiation,* 27**,** 1438-1446.

DE ANGELIS, A. A., JONES, D. H., GRANT, C. V., PARK, S. H., MESLEH, M. F. & OPELLA, S. J. 2005. NMR experiments on aligned samples of membrane proteins. *Methods Enzymol,* 394**,** 350-82.

DE ANGELIS, A. A. & OPELLA, S. J. 2007. Bicelle samples for solid-state NMR of membrane proteins. *Nat Protoc,* 2**,** 2332-8.

DELEO, F. R., DIEP, B. A. & OTTO, M. 2009. Host Defense and Pathogenesis in Staphylococcus aureus Infections. *Infectious Disease Clinics of North America,* 23**,** 17-34.

DIEPHOLZ, M., VENZKE, D., PRINZ, S., BATISSE, C., FLÖRCHINGER, B., RÖSSLE, M., SVERGUN, D. I., BÖTTCHER, B. & FÉTHIÈRE, J. 2008. A different conformation for EGC stator subcomplex in solution and in the assembled yeast V-ATPase: possible implications for regulatory disassembly. *Structure,* 16**,** 1789-98.

DOUCET, M., CHO, J. H., ALINA, G., ATTALA, Z., BAKKER, J., BOUWMAN, W., BUTLER, P., CAMPBELL, K., COOPER-BENUN, T., DURNIAK, C., FORSTER, L., GONZALEZ, M., HEENAN, R., JACKSON, A., KING, S., KIENZLE, P., KRZYWON, J., MURPHY, R., NIELSEN, T., O'DRISCOLL, L., POTRZEBOWSKI, W., PRESCOTT, S., FERRAZ LEAL, R., ROZYCZKO, P., SNOW, T. & WASHINGTON, A. 2021. SasView version 5.0.4. *Zenodo*.

DUFOURC, E. J. 2021. Bicelles and nanodiscs for biophysical chemistry11A tribute to Prof Michèle Auger, University Laval, Canada. *Biochimica et Biophysica Acta (BBA) - Biomembranes,* 1863**,** 183478.

DUNMAN, P. M., MURPHY, E., HANEY, S., PALACIOS, D., TUCKER-KELLOGG, G., WU, S., BROWN, E. L., ZAGURSKY, R. J., SHLAES, D. & PROJAN, S. J. 2001. Transcription profiling-based identification of Staphylococcus aureus genes regulated by the agr and/or sarA loci. *J Bacteriol,* 183**,** 7341-53.

FILATOVA, L. Y., BECKER, S. C., DONOVAN, D. M., GLADILIN, A. K. & KLYACHKO, N. L. 2010. LysK, the enzyme lysing Staphylococcus aureus cells: Specific kinetic features and approaches towards stabilization. *Biochimie,* 92**,** 507-513.

FILIK, J., ASHTON, A. W., CHANG, P. C. Y., CHATER, P. A., DAY, S. J., DRAKOPOULOS, M., GERRING, M. W., HART, M. L., MAGDYSYUK, O. V., MICHALIK, S., SMITH, A., TANG, C. C., TERRILL, N. J., WHARMBY, M. T. & WILHELM, H. 2017. Processing two-dimensional X-ray diffraction and small-angle scattering data in DAWN 2. *Journal of Applied Crystallography,* 50**,** 959-966.

GEISINGER, E., ADHIKARI, R. P., JIN, R., ROSS, H. F. & NOVICK, R. P. 2006. Inhibition of rot translation by RNAIII, a key feature of agr function. *Mol Microbiol,* 61**,** 1038-48.

GIBBS, J. W. 1873. *A method of geometrical representation of the thermodynamic properties of substances by means of surfaces,* [New Haven], Transactions of the Connecticut Academy of Arts and Sciences.

GIRAUDO, A. T., CHEUNG, A. L. & NAGEL, R. 1997. The sae locus of Staphylococcus aureus controls exoprotein synthesis at the transcriptional level. *Archives of Microbiology,* 168**,** 53-58.

GIRAUDO, A. T., RASPANTI, C. G., CALZOLARI, A. & NAGEL, R. 1994. Characterization of a Tn551-mutant of Staphylococcus aureus defective in the production of several exoproteins. *Can J Microbiol,* 40**,** 677-81.

GOODYEAR, C. S. & SILVERMAN, G. J. 2003. Death by a B cell superantigen: In vivo VH-targeted apoptotic supraclonal B cell deletion by a Staphylococcal Toxin. *J Exp Med,* 197**,** 1125-39.

GORDON, C. P., WILLIAMS, P. & CHAN, W. C. 2013. Attenuating Staphylococcus aureus Virulence Gene Regulation: A Medicinal Chemistry Perspective *Journal of Medicinal Chemistry,* 56**,** 1389-1404.

GRAEBSCH, A., ROCHE, S. & NIESSING, D. 2009. X-ray structure of Pur-alpha reveals a Whirly-like fold and an unusual nucleic-acid binding surface. *Proc Natl Acad Sci U S A,* 106**,** 18521-6.

GREENFIELD, N. J. 2006a. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.,* 1**,** 2527-2535.

GREENFIELD, N. J. 2006b. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc,* 1**,** 2876-90.

GUINIER, A. & FOURNET, G. 1955. *Small Angle Scattering of X-Rays,* New York, Wiley.

HAMMOUDA, B. 2010a. Analysis of the Beaucage model. *Journal of Applied Crystallography,* 43**,** 1474-1478.
HAMMOUDA, B. 2010b. A new Guinier-Porod model. *Journal of Applied Crystallography,* 43**,** 716-719.
HARRAGHY, N., HUSSAIN, M., HAGGAR, A., CHAVAKIS, T., SINHA, B., HERRMANN, M. & FLOCK, J.-I. 2003. The adhesive and immunomodulating properties of the multifunctional Staphylococcus aureus protein Eap. *Microbiology,* 149**,** 2701-2707.
HUSSAIN, R., HUGHES, C. S., JÁVORFI, T., SILIGARDI, G., WILLIAMS, P. & BONEV, B. B. 2018. To Boil an Egg: Substrate Binding Affects Critical Stability in Thermal Unfolding of Proteins. *The Journal of Physical Chemistry. B,* 122**,** 2213-2218.
INBARAJ, J. J., CARDON, T. B., LARYUKHIN, M., GROSSER, S. M. & LORIGAN, G. A. 2006. Determining the Topology of Integral Membrane Peptides Using EPR Spectroscopy. *Journal of the American Chemical Society,* 128**,** 9549-9554.
JACQUES, D. A. & TREWHELLA, J. 2010. Small-angle scattering for structural biology--expanding the frontier while avoiding the pitfalls. *Protein Sci,* 19**,** 642-57.
JAEGER, G. 1998. The Ehrenfest Classification of Phase Transitions: Introduction and Evolution. *Archive for History of Exact Sciences,* 53**,** 51-81.
JAHNIG, F. 1981. Critical Effects from Lipid-Protein Interaction in Membranes. 1. Theoretical Description. *Biophys. J.,* 36**,** 329-345.
JAIYEN, Y. 2016. *Novel inhibitors of the Staphylococcus aureus quorum sensing system.* . PhD, University of Nottingham.
JARRAUD, S., LYON, G. J., FIGUEIREDO, A. M. S., GERARD, L., VANDENESCH, F., ETIENNE, J., MUIR, T. W. & NOVICK, R. P. 2000. Exfoliatin-producing strains define a fourth agr specificity group in Staphylococcus aureus. *Journal of Bacteriology,* 182**,** 6517-6522.
JENSEN, R. O., WINZER, K., CLARKE, S. R., CHAN, W. C. & WILLIAMS, P. 2008. Differential recognition of Staphylococcus aureus quorum-sensing signals depends on both extracellular loops 1 and 2 of the transmembrane sensor AgrC. *J Mol Biol,* 381**,** 300-9.
JI, G. Y., BEAVIS, R. & NOVICK, R. P. 1997. Bacterial interference caused by autoinducing peptide variants. *Science,* 276**,** 2027-2030.
JO, S., KIM, T., IYER, V. G. & IM, W. 2008. CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry,* 29**,** 1859-1865.
JOHNSON, C. M. 2013. Differential scanning calorimetry as a tool for protein folding and stability. *Archives of Biochemistry and Biophysics,* 531**,** 100-109.
KENNEDY, A. D., BUBECK WARDENBURG, J., GARDNER, D. J., LONG, D., WHITNEY, A. R., BRAUGHTON, K. R., SCHNEEWIND, O. & DELEO, F. R. 2010. Targeting of alpha-hemolysin by active or passive immunization decreases severity of USA300 skin infection in a mouse model. *J Infect Dis,* 202**,** 1050-8.
KHODAVERDIAN, V., PESHO, M., TRUITT, B., BOLLINGER, L., PATEL, P., NITHIANANTHAM, S., YU, G., DELANEY, E., JANKOWSKY, E. & SHOHAM, M. 2013. Discovery of antivirulence agents against methicillin-resistant Staphylococcus aureus. *Antimicrob Agents Chemother,* 57**,** 3645-52.
KOBAYASHI, S. D., MALACHOWA, N., WHITNEY, A. R., BRAUGHTON, K. R., GARDNER, D. J., LONG, D., BUBECK WARDENBURG, J., SCHNEEWIND, O., OTTO, M. & DELEO, F. R. 2011. Comparative analysis of USA300 virulence determinants in a rabbit model of skin and soft tissue infection. *J Infect Dis,* 204**,** 937-41.
KONAREV, P. V., VOLKOV, V. V., SOKOLOVA, A. V., KOCH, M. H. J. & SVERGUN, D. I. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography,* 36**,** 1277-1282.

KOTLARCHYK, M. & CHEN, S. H. 1983. Analysis of small angle neutron scattering spectra from polydisperse interacting colloids. *The Journal of Chemical Physics,* 79**,** 2461-2469.
KOZAKOV, D., HALL, D. R., XIA, B., PORTER, K. A., PADHORNY, D., YUEH, C., BEGLOV, D. & VAJDA, S. 2017. The ClusPro web server for protein-protein docking. *Nature Protocols,* 12**,** 255-278.
KUČERKA, N., KISELEV, M. A. & BALGAVÝ, P. 2004. Determination of bilayer thickness and lipid surface area in unilamellar dimyristoylphosphatidylcholine vesicles from small-angle neutron scattering curves: a comparison of evaluation methods. *European Biophysics Journal,* 33**,** 328-334.
LADBURY, J. E. 2004. Application of Isothermal Titration Calorimetry in the Biological Sciences: Things Are Heating Up! *BioTechniques,* 37**,** 885-887.
LANDAU, L. 1936. The Theory of Phase Transitions. *Nature,* 138**,** 840-841.
LAUB, M. T. & GOULIAN, M. 2007. Specificity in two-component signal transduction pathways. *Annu Rev Genet,* 41**,** 121-45.
LEBEAU, C., VANDENESCH, F., GREENLAND, T., NOVICK, R. P. & ETIENNE, J. 1994. Coagulase expression in Staphylococcus aureus is positively and negatively modulated by an agr-dependent mechanism. *J Bacteriol,* 176**,** 5534-6.
LEE, J., CHENG, X., SWAILS, J. M., YEOM, M. S., EASTMAN, P. K., LEMKUL, J. A., WEI, S., BUCKNER, J., JEONG, J. C., QI, Y., JO, S., PANDE, V. S., CASE, D. A., BROOKS, C. L., MACKERELL, A. D., KLAUDA, J. B. & IM, W. 2016. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation,* 12**,** 405-413.
LEE, J., WU, J., DENG, Y., WANG, J., WANG, C., WANG, J., CHANG, C., DONG, Y., WILLIAMS, P. & ZHANG, L. H. 2013. A cell-cell communication signal integrates quorum sensing and stress response. *Nat Chem Biol,* 9**,** 339-43.
LEONARD, P. G., BEZAR, I. F., SIDOTE, D. J. & STOCK, A. M. 2012. Identification of a Hydrophobic Cleft in the LytTR Domain of AgrA as a Locus for Small Molecule Interactions that Inhibit DNA Binding *Biochemistry,* 51**,** 10035-10043.
LI, M., MORALES, H. H., KATSARAS, J., KUCERKA, N., YANG, Y. K., MACDONAD, P. M. & NIEH, M. P. 2013. Morphological Characterization of DMPC/CHAPSO Bicellar Mixtures: A Combined SANS and NMR Study. *Langmuir,* 29**,** 15943-15957.
LINA, G., JARRAUD, S., JI, G., GREENLAND, T., PEDRAZA, A., ETIENNE, J., NOVICK, R. P. & VANDENESCH, F. 1998. Transmembrane topology and histidine protein kinase activity of AgrC, the agr signal receptor in Staphylococcus aureus. *Molecular Microbiology,* 28**,** 655-662.
LIU, Y., MU, C., YING, X., LI, W., WU, N., DONG, J., GAO, Y., SHAO, N., FAN, M. & YANG, G. 2011. RNAIII activates map expression by forming an RNA–RNA complex in Staphylococcus aureus. *FEBS Letters,* 585**,** 899-905.
LOPES, J. L., MILES, A. J., WHITMORE, L. & WALLACE, B. A. 2014. Distinct circular dichroism spectroscopic signatures of polyproline II and unordered secondary structures: applications in secondary structure analyses. *Protein Sci,* 23**,** 1765-72.
LYON, G. J., WRIGHT, J. S., MUIR, T. W. & NOVICK, R. P. 2002. Key determinants of receptor activation in the agr autoinducing peptides of Staphylococcus aureus. *Biochemistry,* 41**,** 10095-104.
MARROQUIN, S., GIMZA, B., TOMLINSON, B., STEIN, M., FREY, A., KEOGH, R. A., ZAPF, R., TODD, D. A., CECH, N. B., CARROLL, R. K. & SHAW, L. N. 2019. MroQ Is a Novel Abi-Domain Protein That Influences Virulence Gene Expression in Staphylococcus aureus via Modulation of Agr Activity. *Infect Immun,* 87.

MCDOWELL, P., AFFAS, Z., REYNOLDS, C., HOLDEN, M. T. G., WOOD, S. J., SAINT, S., COCKAYNE, A., HILL, P. J., DODD, C. E. R., BYCROFT, B. W., CHAN, W. C. & WILLIAMS, P. 2001. Structure, activity and evolution of the group I thiolactone peptide quorum-sensing system of Staphylococcus aureus. *Molecular Microbiology,* 41**,** 503-512.

MCNAUGHT, A. D. & WILKINSON, A. 1997. *IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"),* Oxford, Blackwell Scientific Publications

MORFELDT, E., TAYLOR, D., VON GABAIN, A. & ARVIDSON, S. 1995. Activation of alpha-toxin translation in Staphylococcus aureus by the trans-encoded antisense RNA, RNAIII. *Embo j,* 14**,** 4569-77.

MORROW, M. R. & WHITEHEAD, J. P. 1988. A Phenomenological Model for Lipid-Protein Bilayer with Critical Mixing. *Biochim. Biophys. Acta, Biomembr.,* 941**,** 271-277.

NIEH, M.-P., GLINKA, C. J., KRUEGER, S., PROSSER, R. S. & KATSARAS, J. 2001. SANS Study of the Structural Phases of Magnetically Alignable Lanthanide-Doped Phospholipid Mixtures. *Langmuir,* 17**,** 2629-2638.

NIEH, M.-P., GLINKA, C. J., KRUEGER, S., PROSSER, R. S. & KATSARAS, J. 2002. SANS Study on the Effect of Lanthanide Ions and Charged Lipids on the Morphology of Phospholipid Mixtures. *Biophysical Journal,* 82**,** 2487-2498.

NOVICK, R. P. 2003. Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol,* 48**,** 1429-49.

NOVICK, R. P., ROSS, H. F., PROJAN, S. J., KORNBLUM, J., KREISWIRTH, B. & MOGHAZEH, S. 1993. Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *Embo j,* 12**,** 3967-75.

PAHARIK, A. E., SALGADO-PABON, W., MEYERHOLZ, D. K., WHITE, M. J., SCHLIEVERT, P. M. & HORSWILL, A. R. 2016. The Spl Serine Proteases Modulate Staphylococcus aureus Protein Production and Virulence in a Rabbit Model of Pneumonia. *mSphere,* 1.

PAIVA, C. A. 2019. *Surface signalling and quorum sensing in Pseudomonas aeruginosa.* PhD, University of Nottingham.

PARK, S. H., DE ANGELIS, A. A., NEVZOROV, A. A., WU, C. H. & OPELLA, S. J. 2006. Three-dimensional structure of the transmembrane domain of Vpu from HIV-1 in aligned phospholipid bicelles. *Biophys J,* 91**,** 3032-42.

PENG, H. L., NOVICK, R. P., KREISWIRTH, B., KORNBLUM, J. & SCHLIEVERT, P. 1988. Cloning, Characterization, and Sequencing of an Accessory Gene Regulator (agr) in Staphylococcus aureus. *Journal of Bacteriology,* 170**,** 4365-4372.

PERINO-GALLICE, L., FRAGNETO, G., MENNICKE, U., SALDITT, T. & RIEUTORD, F. 2002. Dewetting of solid-supported multilamellar lipid layers. *Eur Phys J E Soft Matter,* 8**,** 275-82.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera-A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.,* 25**,** 1605-1612.

PHILLIPS, J. C., BRAUN, R., WANG, W., GUMBART, J., TAJKHORSHID, E., VILLA, E., CHIPOT, C., SKEEL, R. D., KALE, L. & SCHULTEN, K. 2005. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry,* 26**,** 1781-1802.

PIGNATARO, M. F., HERRERA, M. G. & DODERO, V. I. 2020. Evaluation of Peptide/Protein Self-Assembly and Aggregation by Spectroscopic Methods. *Molecules,* 25**,** 4854.

PRASAD, S., KHADATARE, P. B. & ROY, I. 2011. Effect of Chemical Chaperones in Improving the Solubility of Recombinant Proteins in Escherichia coli. *Applied and Environmental Microbiology,* 77**,** 4603-4609.

PROSSER, R. S., EVANICS, F., KITEVSKI, J. L. & AL-ABDUL-WAHID, M. S. 2006. Current Applications of Bicelles in NMR Studies of Membrane-Associated Amphiphiles and Proteins. *Biochemistry,* 45**,** 8453-8465.

PROSSER, R. S., HUNT, S. A., DINATALE, J. A. & VOLD, R. R. 1996. Magnetically Aligned Membrane Model Systems with Positive Order Parameter: Switching the Sign of Szz with Paramagnetic Ions. *Journal of the American Chemical Society,* 118**,** 269-270.

PROSSER, R. S., HWANG, J. S. & VOLD, R. R. 1998a. Magnetically Aligned Phospholipid Bilayers with Positive Ordering: A New Model Membrane System. *Biophysical Journal,* 74**,** 2405-2418.

PROSSER, R. S., VOLKOV, V. B. & SHIYANOVSKAYA, I. V. 1998b. Novel chelate-induced magnetic alignment of biological membranes. *Biophysical journal,* 75**,** 2163-2169.

QIN, J. & GRONENBORN, A. M. 2014. Weak protein complexes: challenging to study but essential for life. *Febs j,* 281**,** 1948-9.

QIU, R. D., PEI, W. H., ZHANG, L. S., LIN, J. Q. & JI, G. Y. 2005. Identification of the putative staphylococcal AgrB catalytic residues involving the proteolytic cleavage of AgrD to generate autoinducing peptide. *Journal of Biological Chemistry,* 280**,** 16695-16704.

QUECK, S. Y., JAMESON-LEE, M., VILLARUZ, A. E., BACH, T. L., KHAN, B. A., STURDEVANT, D. E., RICKLEFS, S. M., LI, M. & OTTO, M. 2008. RNAIII-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in Staphylococcus aureus. *Mol. Cell. ,* 32**,** 150-158.

QUILLIN, M. L. & MATTHEWS, B. W. 2000. Accurate calculation of the density of proteins. *Acta Crystallogr D Biol Crystallogr,* 56**,** 791-4.

RAJASREE, K., FASIM, A. & GOPAL, B. 2016. Conformational features of the Staphylococcus aureus AgrA-promoter interactions rationalize quorum-sensing triggered gene expression. *Biochemistry and Biophysics Reports,* 6**,** 124-134.

RAJESH, S., OVERDUIN, M. & BONEV, B. B. 2016. NMR of Membrane Proteins: Beyond Crystals. *In:* MORAES, I. (ed.) *The Next Generation in Membrane Protein Structure Determination.* Springer.

RAMBO, R. P. & TAINER, J. A. 2011. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers,* 95**,** 559-571.

REED, S. B., WESSON, C. A., LIOU, L. E., TRUMBLE, W. R., SCHLIEVERT, P. M., BOHACH, G. A. & BAYLES, K. W. 2001. Molecular characterization of a novel Staphylococcus aureus serine protease operon. *Infect Immun,* 69**,** 1521-7.

ROY, A., KUCUKURAL, A. & ZHANG, Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc,* 5**,** 725-38.

SAÏD-SALIM, B., DUNMAN, P. M., MCALEESE, F. M., MACAPAGAL, D., MURPHY, E., MCNAMARA, P. J., ARVIDSON, S., FOSTER, T. J., PROJAN, S. J. & KREISWIRTH, B. N. 2003. Global regulation of Staphylococcus aureus genes by Rot. *J Bacteriol,* 185**,** 610-9.

SALVADO, B., VILAPRINYO, E., SORRIBAS, A. & ALVES, R. 2015. A survey of HK, HPt, and RR domains and their organization in two-component systems and phosphorelay proteins of organisms with fully sequenced genomes. *Peerj,* 3**,** 27.

SANDERS, C. R., HARE, B. J., HOWARD, K. P. & PRESTEGARD, J. H. 1994. Magnetically-oriented phospholipid micelles as a tool for the study of membrane-

associated molecules. *Progress in Nuclear Magnetic Resonance Spectroscopy,* 26**,** 421-444.

SANDERS, C. R., II & LANDIS, G. C. 1995. Reconstitution of Membrane Proteins into Lipid-Rich Bilayered Mixed Micelles for NMR Studies. *Biochemistry,* 34**,** 4030-4040.

SANDERS, C. R., II & SCHWONEK, J. P. 1992. Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR. *Biochemistry,* 31**,** 8898-8905.

SCHÖN, A., CLARKSON, B. R., JAIME, M. & FREIRE, E. 2017. Temperature stability of proteins: Analysis of irreversible denaturation using isothermal calorimetry. *Proteins,* 85**,** 2009-2016.

SIDOTE, D. J., BARBIERI, C. M., WU, T. & STOCK, A. M. 2008. Structure of the Staphylococcus aureus AgrA LytTR Domain Bound to DNA Reveals a Beta Fold with a Novel Mode of Binding. *Structure,* 16**,** 727-735.

SILIGARDI, G., PANARETOU, B., MEYER, P., SINGH, S., WOOLFSON, D. N., PIPER, P. W., PEARL, L. H. & PRODROMOU, C. 2002. Regulation of Hsp90 ATPase activity by the co-chaperone Cdc37p/p50(cdc97). *Journal of Biological Chemistry,* 277**,** 20151-20159.

SLOAN, T. J., MURRAY, E., YOKOYAMA, M., MASSEY, R. C., CHAN, W. C., BONEV, B. B. & WILLIAMS, P. 2019. Timing Is Everything: Impact of Naturally Occurring Staphylococcus aureus AgrC Cytoplasmic Domain Adaptive Mutations on Autoinduction. *Journal of Bacteriology,* 201**,** 15.

SRIVASTAVA, S. K., RAJASREE, K., FASIM, A., ARAKERE, G. & GOPAL, B. 2014. Influence of the AgrC-AgrA Complex on the Response Time of Staphylococcus aureus Quorum Sensing. *Journal of Bacteriology,* 196**,** 2876-2888.

STERNIN, E., NIZZA, D. & GAWRISCH, K. 2001. Temperature Dependence of DMPC/DHPC Mixing in a Bicellar Solution and Its Structural Implications. *Langmuir,* 17**,** 2610-2616.

STOCK, M. R., FANG, L., JOHNSON, K. R., COSGRIFF, C., TEOH, W. P. & ALONZO, F., 3RD 2022. Characterization of MroQ-Dependent Maturation and Export of the Staphylococcus aureus Accessory Gene Regulatory System Autoinducing Peptide. *Infect Immun,* 90**,** e0026322.

SVERGUN, D. 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.***,** 2879-2886.

SVERGUN, D. I., BARBERATO, C. & KOCH, M. H. J. 1995. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Cryst.***,** 768-773.

SVERGUN, D. I., KOCH, M. H. J., TIMMINS, P. A. & MAY, R. P. 2013. *Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules*, Oxford University Press.

SVERGUN, D. I. & NIERHAUS, K. H. 2000. A map of protein-rRNA distribution in the 70S Escherichia coli ribosome. *J. Biol. Chem.***,** 14432-14439.

SWANSON, R. V., ALEX, L. A. & SIMON, M. I. 1994. Histidine and aspartate phosphorylation: two-component systems and the limits of homology. *Trends in Biochemical Sciences,* 19**,** 485-490.

THOENDEL, M. & HORSWILL, A. R. 2013. Random mutagenesis and topology analysis of the autoinducing peptide biosynthesis proteins in Staphylococcus aureus. *Mol Microbiol,* 87**,** 318-37.

TOLÉDANO, J. C. & TOLÉDANO, P. 1987. *The Landau Theory of Phase Transitions: Application to Structural, Incommensurate, Magnetic, and Liquid Crystal Systems,* Singapore, World Scientific.

TOMOMORI, C., TANAKA, T., DUTTA, R., PARK, H., SAHA, S. K., ZHU, Y., ISHIMA, R., LIU, D., TONG, K. I., KUROKAWA, H., QIAN, H., INOUYE, M. & IKURA, M. 1999. Solution structure of the homodimeric core domain of Escherichia colihistidine kinase EnvZ. *Nature Structural Biology,* 6**,** 729-734.

TRIA, G., MERTENS, H. D. T., KACHALA, M. & SVERGUN, D. I. 2015. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ***,** 207-217.

TSUTAKAWA, S. E., VAN WYNSBERGHE, A. W., FREUDENTHAL, B. D., WEINACHT, C. P., GAKHAR, L., WASHINGTON, M. T., ZHUANG, Z., TAINER, J. A. & IVANOV, I. 2011. Solution X-ray scattering combined with computational modeling reveals multiple conformations of covalently bound ubiquitin on PCNA. *Proc Natl Acad Sci U S A,* 108**,** 17672-7.

VAN DAM, L., KARLSSON, G. & EDWARDS, K. 2004. Direct observation and characterization of DMPC/DHPC aggregates under conditions relevant for biological solution NMR. *Biochim Biophys Acta,* 1664**,** 241-56.

VAN HAL, S. J., JENSEN, S. O., VASKA, V. L., ESPEDIDO, B. A., PATERSON, D. L. & GOSBELL, I. B. 2012. Predictors of mortality in Staphylococcus aureus Bacteremia. *Clin Microbiol Rev,* 25**,** 362-86.

VOLD, R. R. & PROSSER, R. S. 1996. Magnetically oriented phospholipid bilayered micelles for structural studies of polypeptides. Does the ideal bicelle exist? *Journal of Magnetic Resonance Series B,* 113**,** 267-271.

WANG, B., ZHAO, A., NOVICK, R. & MUIR, T. W. 2014. Activation and Inhibition of the Receptor Histidine Kinase AgrC Occurs Through Opposite Helical Transduction Motions. *Mol. Cell.,* 53**,** 929-940.

WANG, B., ZHAO, A., NOVICK, R. P. & MUIR, T. W. 2015. Key driving forces in the biosynthesis of autoinducing peptides required for staphylococcal virulence. *Proc Natl Acad Sci U S A,* 112**,** 10679-84.

WANG, B., ZHAO, A., XIE, Q., OLINARES, P. D., CHAIT, B. T., NOVICK, R. P. & MUIR, T. W. 2017a. Functional Plasticity of the AgrC Receptor Histidine Kinase Required for Staphylococcal Virulence. *Cell Chem Biol,* 24**,** 76-86.

WANG, Y., ZHOU, H., ONUK, E., BADGER, J. & MAKOWSKI, L. 2017b. What Can We Learn from Wide-Angle Solution Scattering? *In:* CHAUDHURI, B., MUÑOZ, I. G., QIAN, S. & URBAN, V. S. (eds.) *Biological Small Angle Scattering: Techniques, Strategies and Tips.* Singapore: Springer Singapore.

WARSCHAWSKI, D. E., ARNOLD, A. A., BEAUGRAND, M., GRAVEL, A., CHARTRAND, É. & MARCOTTE, I. 2011. Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochim Biophys Acta,* 1808**,** 1957-74.

WHITMORE, L. & WALLACE, B. A. 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers,* 89**,** 392-400.

WILLIAMS, P. 2007. Quorum sensing, communication and cross-kingdom signalling in the bacterial world. *Microbiology,* 153**,** 3923-38.

WILLIAMS, P., WINZER, K., CHAN, W. C. & CÁMARA, M. 2007. Look who's talking: communication and quorum sensing in the bacterial world. *Philos Trans R Soc Lond B Biol Sci,* 362**,** 1119-34.

XIE, G. F. & TIMASHEFF, S. N. 1997. Mechanism of the stabilization of ribonuclease A by sorbitol: Preferential hydration is greater for the denatured than for the native protein. *Protein Science,* 6**,** 211-221.

XIE, Q., ZHAO, A., JEFFREY, P. D., KIM, M. K., BASSLER, B. L., STONE, H. A., NOVICK, R. P. & MUIR, T. W. 2019. Identification of a Molecular Latch that Regulates Staphylococcal Virulence. *Cell Chem Biol,* 26**,** 548-558.e4.

YAKHNIN, A. V., YAKHNIN, H. & BABITZKE, P. 2012. Gel mobility shift assays to detect protein-RNA interactions. *Methods Mol Biol,* 905**,** 201-11.

YANG, J., YAN, R., ROY, A., XU, D., POISSON, J. & ZHANG, Y. 2015. The I-TASSER Suite: protein structure and function prediction. *Nat Methods,* 12**,** 7-8.

YANG, S., BLACHOWICZ, L., MAKOWSKI, L. & ROUX, B. 2010. Multidomain assembled states of Hck tyrosine kinase in solution. *Proc Natl Acad Sci U S A,* 107**,** 15757-62.

YILMAZ, M., ELALDI, N., BALKAN, İ., ARSLAN, F., BATıREL, A. A., BAKıCı, M. Z., GOZEL, M. G., ALKAN, S., ÇELIK, A. D., YETKIN, M. A., BODUR, H., SıNıRTAŞ, M., AKALıN, H., ALTAY, F. A., ŞENCAN, İ., AZAK, E., GÜNDEŞ, S., CEYLAN, B., ÖZTÜRK, R., LEBLEBICIOGLU, H., VAHABOGLU, H. & MERT, A. 2016. Mortality predictors of Staphylococcus aureus bacteremia: a prospective multicenter study. *Ann Clin Microbiol Antimicrob,* 15**,** 7.

ZAPF, J., SEN, U., MADHUSUDAN, HOCH, J. A. & VARUGHESE, K. I. 2000. A transient interaction between two phosphorelay proteins trapped in a crystal lattice reveals the mechanism of molecular recognition and phosphotransfer in signal transduction. *Structure,* 8**,** 851-62.

ZHANG, L. S., GRAY, L., NOVICK, R. P. & JI, G. Y. 2002. Transmembrane topology of AgrB, the protein involved in the post-translational modification of AgrD in Staphylococcus aureus. *Journal of Biological Chemistry,* 277**,** 34736-34742.

ZHANG, L. S. & JI, G. Y. 2004. Identification of a staphylococcal AgrB segment(s) responsible for group-specific processing of AgrD by gene swapping. *Journal of Bacteriology,* 186**,** 6706-6713.

ZHANG, Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics,* 9**,** 40.

Appendix

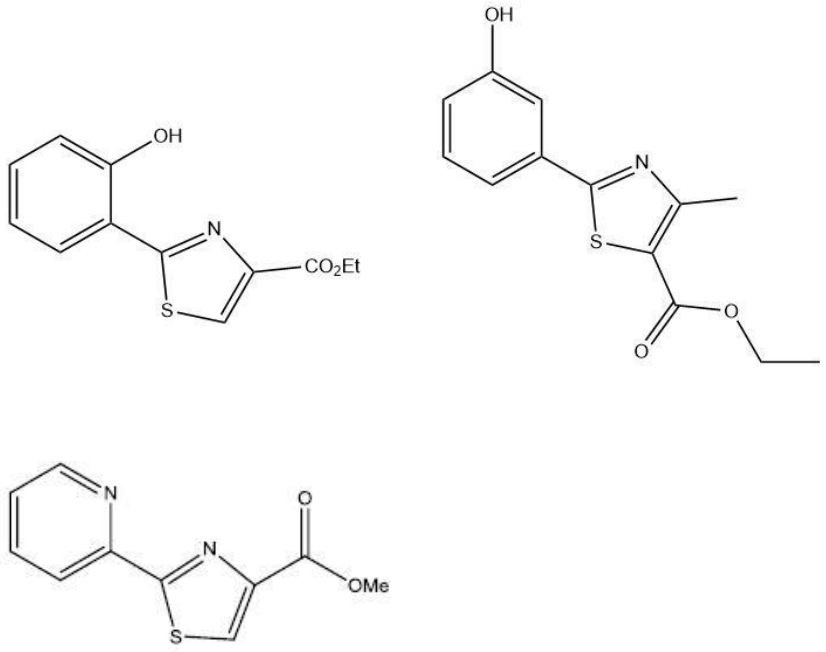


Figure A1 *Chemical structures of IQS analogue number 3 (A), 11 (B) and 15 (C). Drawn using Chem3D (PerkinElmer Informatics).*

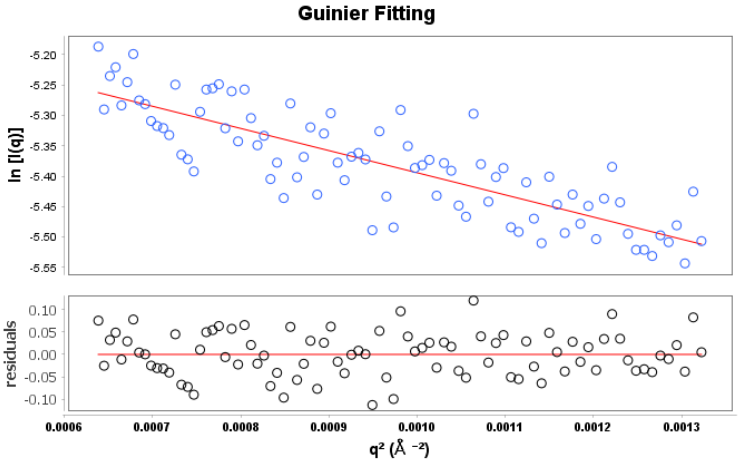
Table A1. Primers for cytosolic AgrC cloning.		
Primer name	**Construct**	**Sequence**
SZ1	pCOLD-I reverse	ATGCCTACCTTCGATATG

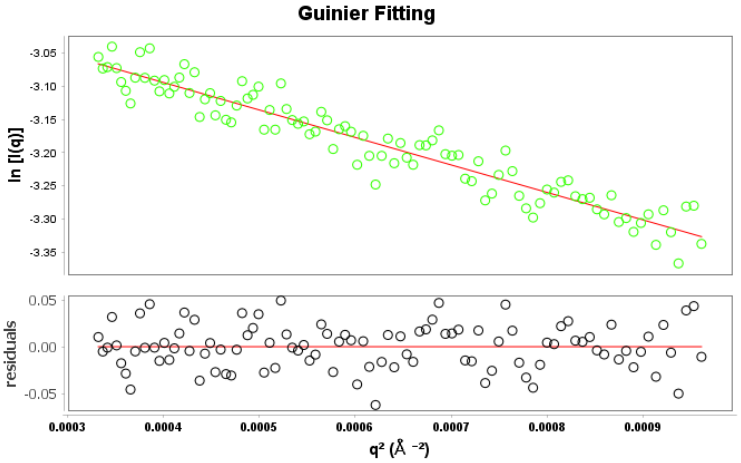
SZ2	pCOLD-I forward	ATGGAGCTCGGTACCCTC
SZ3	agrC cyto forward	ATCATATCGAAGGTAGGCATAAAGAGATG AAATATAAACGTAATC
SZ4	agrC cyto reverse	TCGAGGGTACCGAGCTCCATCTAGTTGTTA ATAATTTCAACTTTTTG

Table A2. Primers for EMSA probes synthesis.

Name	5’ modification	3’ modification	Purification	Sequence
YJ70	-	-	-	TACATTTAACAGTTAAG TATTTATTTCCTACAGTT AGGCAATATAATG
YJ66	-	-	-	CATTATATTGCCTAACT GTAGGAAATAAATACT TAACTGTTAAATGTA
YJ62	-	-	-	AATTTTTCTTAACTAGTCG TTTTTTATTC
YJ64	-	-	-	CATTATATTGCCTAACTGT AGG
YJ69	ATTO700		HPLC	AATTTTTCTTAACTAGTCG

				TTTTTTATTC
YJ65	ATTO700		HPLC	TACATTTAACAGTTAAGTA TTTATTTCCTACAGTTAGGC AATATAA TG


Guinier Fitting
ln [I(q)]
-5.20
-5.25
-5.30
-5.35
-5.40
-5.45
-5.50
-5.55
residuals
0.10
0.05
0.00
-0.05
-0.10
0.0006
0.0007
0.0008
0.0009
0.0010
0.0011
0.0012
0.0013
q² (Å ⁻²)


Guinier Fitting
ln [I(q)]
-3.05
-3.10
-3.15
-3.20
-3.25
-3.30
-3.35
residuals
0.05
0.00
-0.05
0.0003
0.0004
0.0005
0.0006
0.0007
0.0008
0.0009
q² (Å ⁻²)

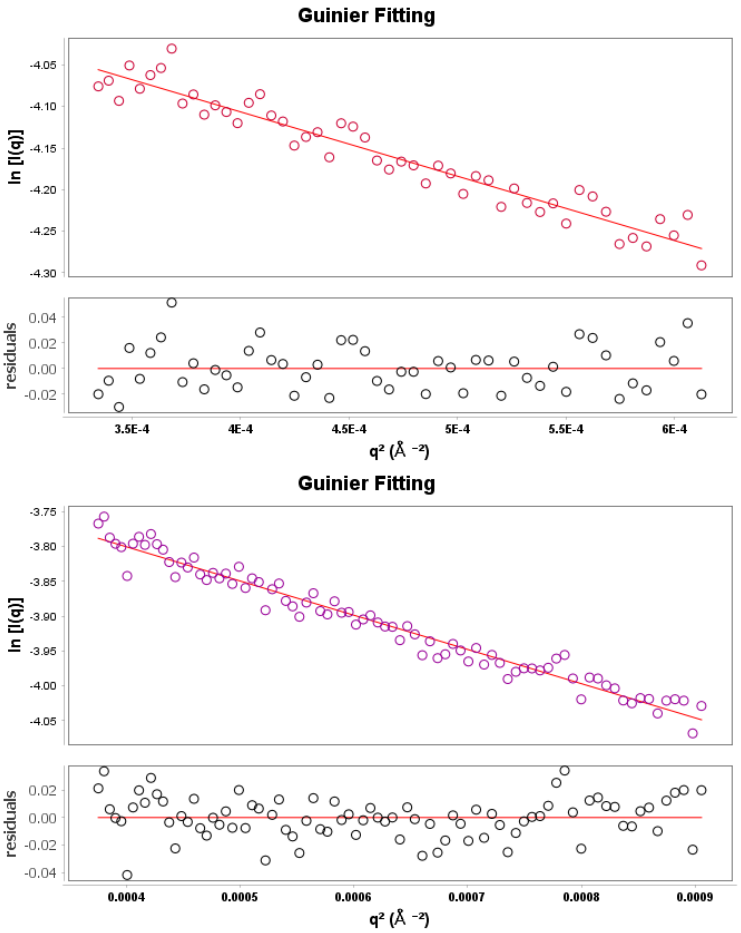


Figure A2 ***Guinier analysis of C199S AgrA (blue), AgrC$_{206}$ (green), the mixture in a 1:1 ratio (red) and in a 1:2 ratio (purple) (AgrA:AgrC). The $q * R_g$ limits are 0.8266-1.1961, 0.6338-1.0872, 0.8732-1.1864 and 0.7362-1.1507 respectively.***

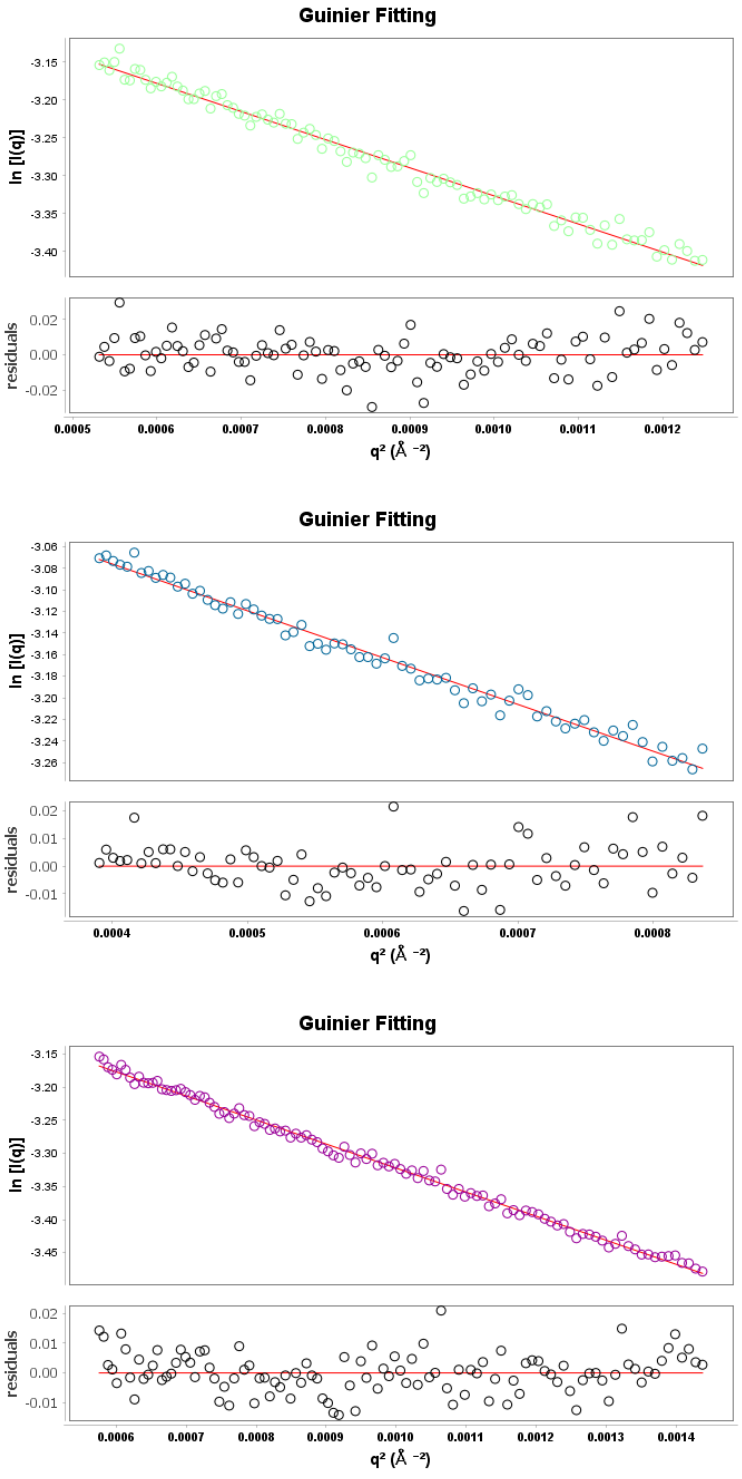


Figure A3 ***Guinier analysis of AgrC$_{206}$ alone (green), with ATP (blue) and non-hydrolysable ATP (purple). The $q * R_g$ limits are 0.7608-1.1728, 0.7050-1.0378 and 0.7824-1.2458 respectively.***

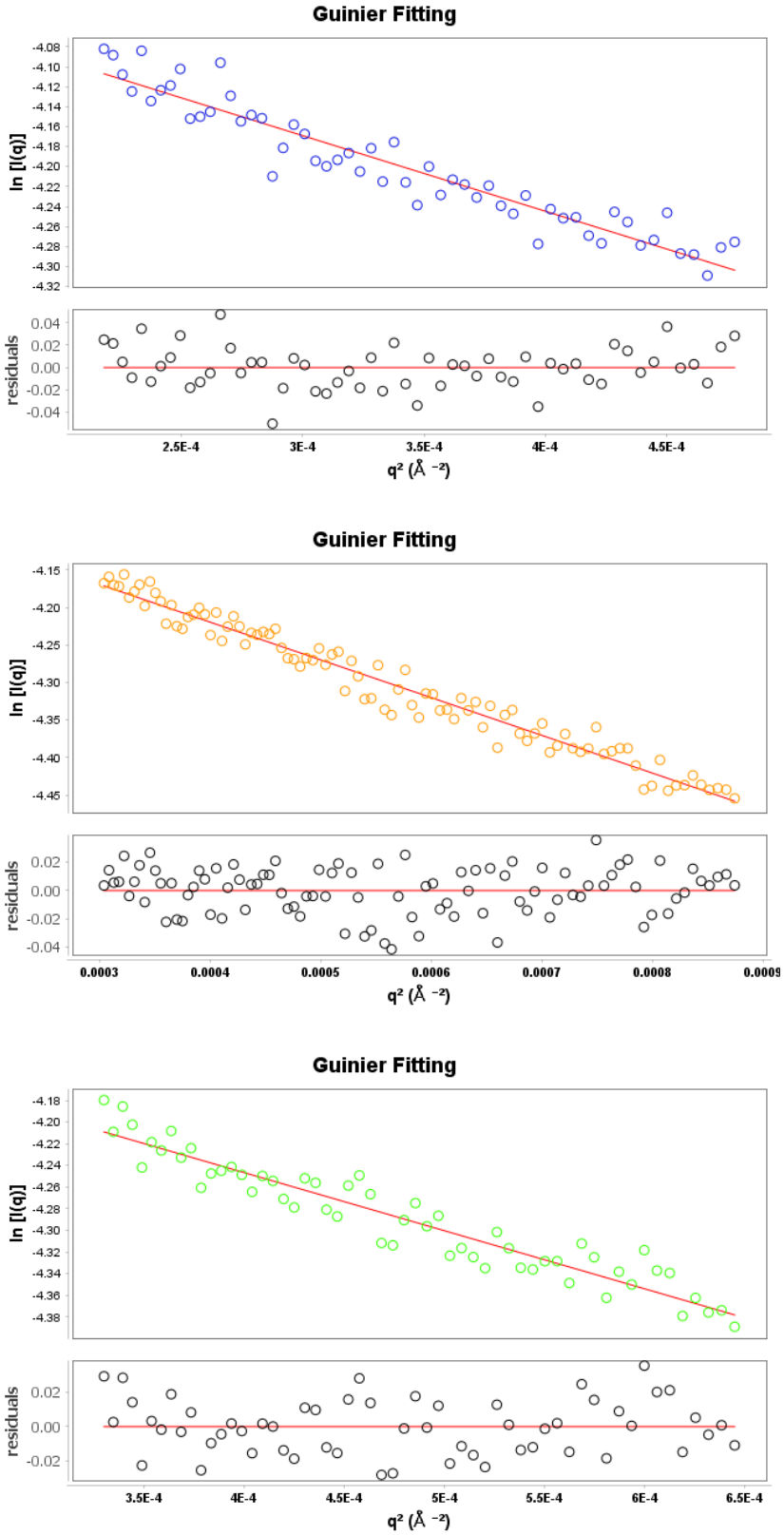
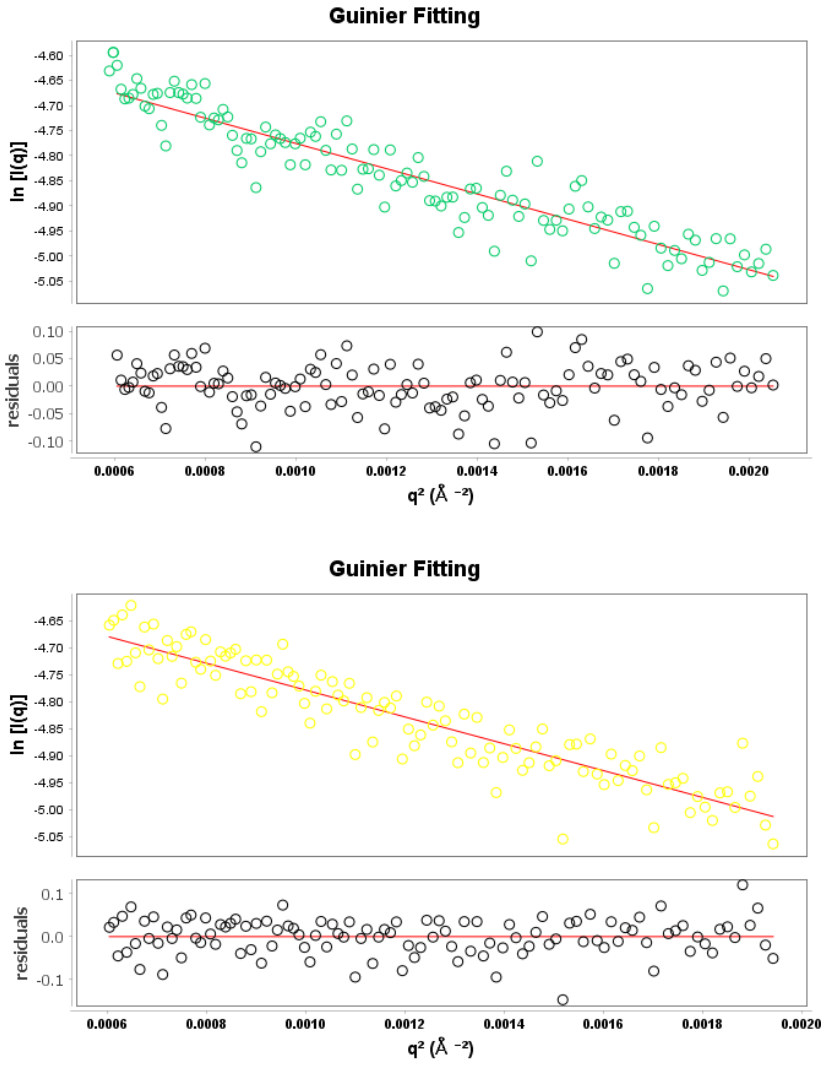


Figure A4 ***Guinier analysis of the AgrA/AgrC$_{206}$ mixture (in a 1:1 molar ratio) alone (blue), with ATP (orange) and non-hydrolysable ATP (green).***

The $q * R_g$ limits are 0.6959-1.0365, 0.6688-1.1436 and 0.7220-1.0141 respectively.



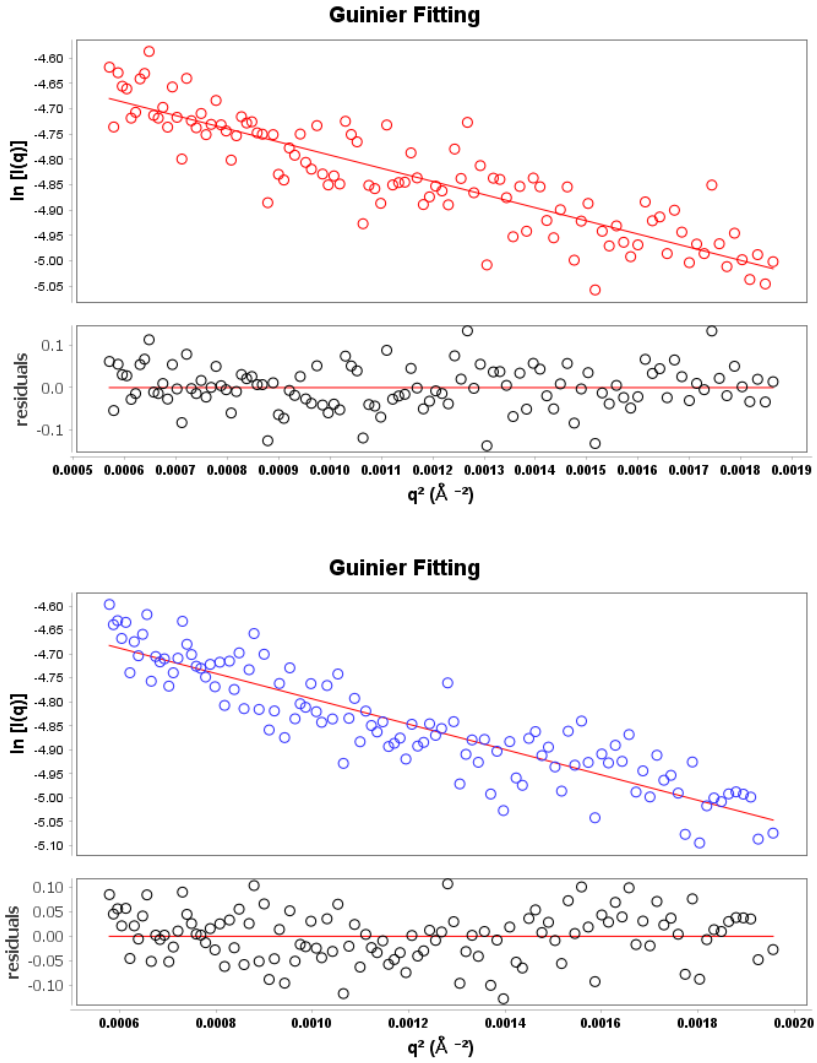
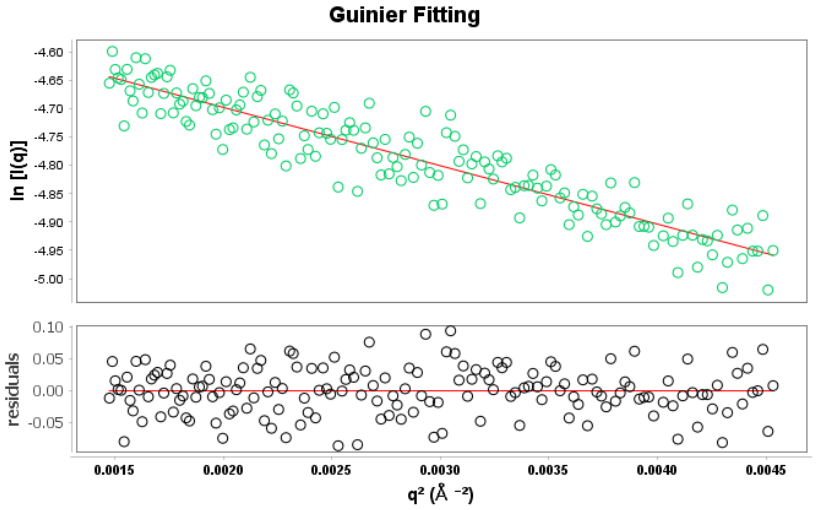
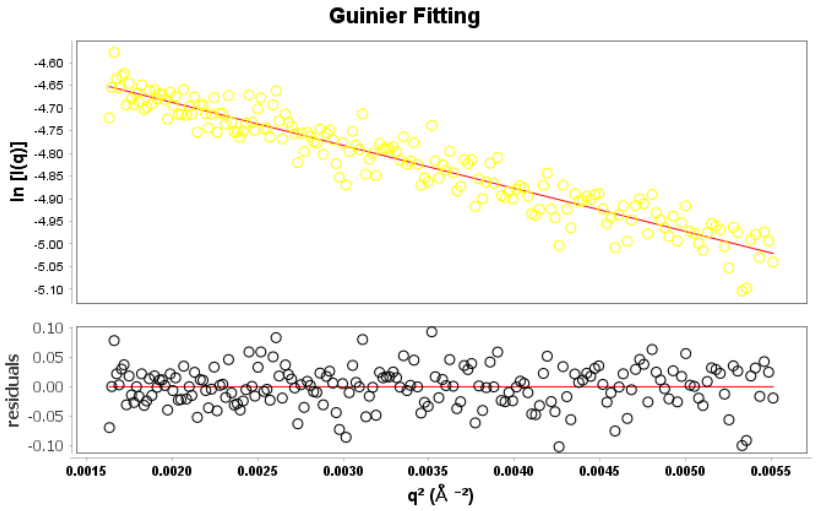


Figure A5 ***Guinier analysis of AgrA alone (green), with IQS3 (yellow), IQS11 (red) and IQS15 (blue). The $q * R_g$ limits are 0.6659-1.2395, 0.6608-1.1963, 0.6558-1.1982 and 0.6673-1.2401 respectively.***


Guinier Fitting
ln [I(q)]
-4.60
-4.65
-4.70
-4.75
-4.80
-4.85
-4.90
-4.95
-5.00
residuals
0.10
0.05
0.00
-0.05
0.0015
0.0020
0.0025
0.0030
0.0035
0.0040
0.0045
q² (Å ⁻²)


Guinier Fitting
ln [I(q)]
-4.60
-4.65
-4.70
-4.75
-4.80
-4.85
-4.90
-4.95
-5.00
-5.05
-5.10
residuals
0.10
0.05
0.00
-0.05
-0.10
0.0015
0.0020
0.0025
0.0030
0.0035
0.0040
0.0045
0.0050
0.0055
q² (Å ⁻²)

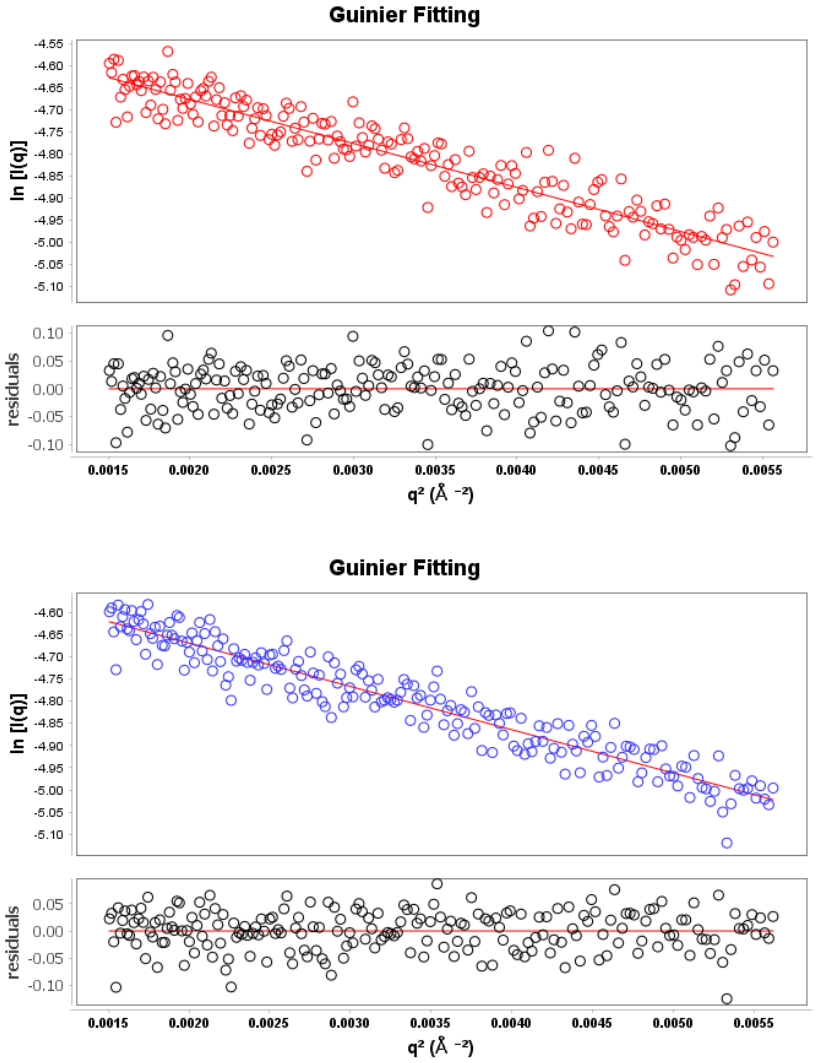


Figure A6 ***Guinier analysis of C-terminal AgrA alone (green), with IQS3 (yellow), IQS11 (red) and IQS15 (blue). The $q * R_g$ limits are 0.6649-1.1777, 0.6712-1.2471, 0.6614-1.2861 and 0.6531-1.2760 respectively.***

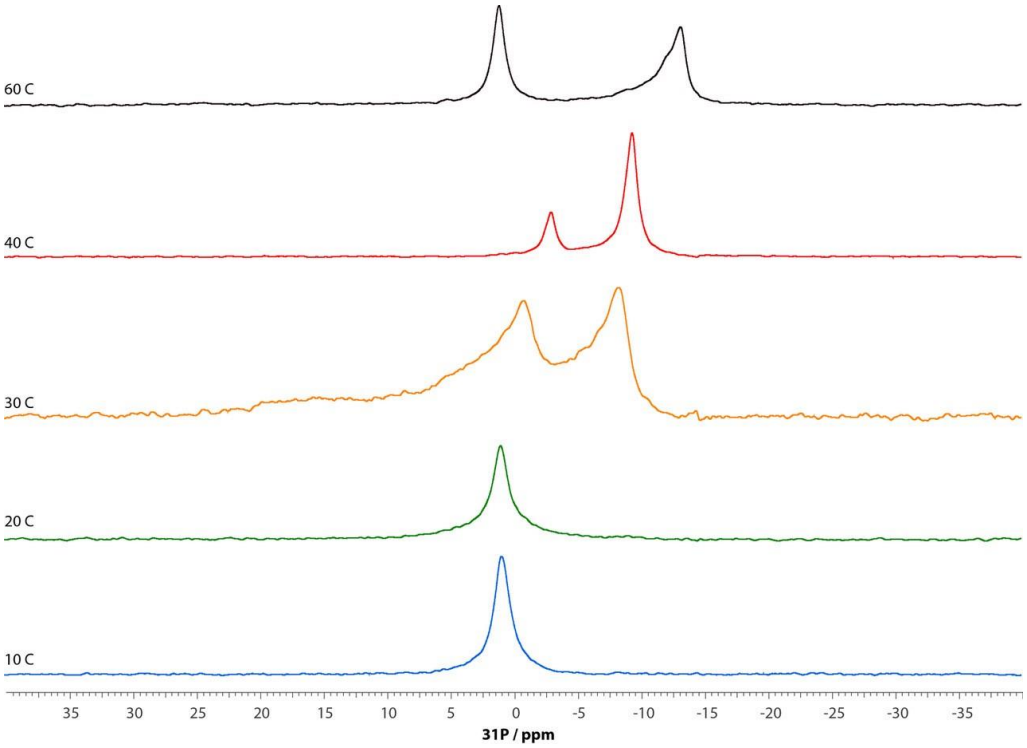


Figure A7 ***^{31}P static wideline NMR spectra of DMPC/DHPC (3:1) bicelles.***

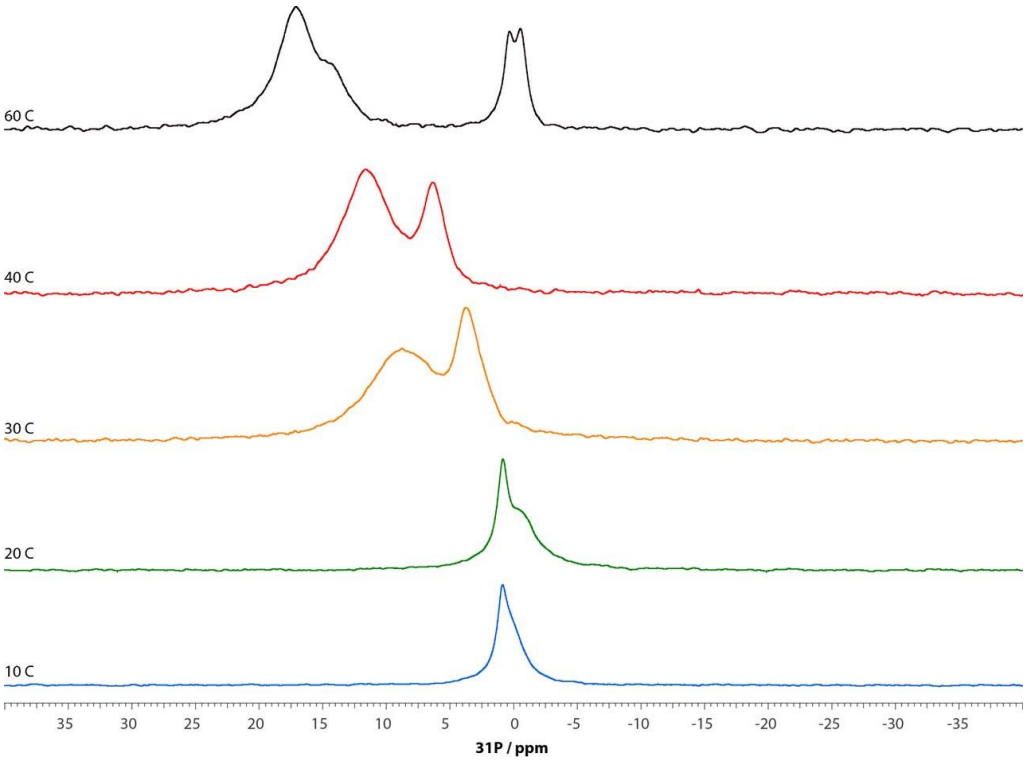


Figure A8 ***^{31}P static wideline NMR spectra of DMPC/DHPC (3:1) bicelles with 5 mM $YbCl_3$.***