

**Development and application of novel
crystallisation methodology to
medically relevant proteins**

A thesis submitted for the Degree of Doctor of Philosophy

Sean Cecil Kassen

Imperial College London
Division of Computational and Systems Medicine
Department of Surgery and Cancer
Faculty of Medicine

2019

DECLARATION OF ORIGINALITY

I confirm that the work presented in this thesis is my own.
All other work in this thesis is appropriately referenced.

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial-No Derivatives International Licence.

Under this licence, you may copy and redistribute the material in any medium or format on the condition that; you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text.

ABSTRACT

X-ray crystallography is the most widely used technique for determining 3-D protein structures. Information gained from crystallography is vital to the success of rational drug design and other biotechnological applications. Crystallography relies on the production of high quality protein crystals but obtaining such crystals is a major obstacle to progress. This project tackles this problem by (i) designing, developing and validating new and improved crystallisation methodologies for obtaining high quality crystals; and (ii) applying the new techniques in addition to existing methods, to crystallise medically important proteins in order to facilitate their 3-Dimensional structure determination.

Presented here is the design and validation of an improved method to slow down protein crystallisation in order to enable the growth of fewer, larger single crystals of higher quality. The method called the oil-on-drop method is a variation of the Chayen method which places an oil barrier over precipitant reservoirs in hanging drop trials. The variation consists of dispensing the oil directly onto the protein drops. This method was successfully tested and validated using five proteins, three of which are of medical relevance including a methyltransferase and an antibody-peptide complex.

A comparison of the oil-on-drop method with existing standard and non-standard techniques for optimising crystal quality such as the application of nucleants and the Chayen method showed the oil-on-drop technique to be superior in many ways. Most importantly, it reproducibly yielded crystals diffracting to higher resolution of three of the proteins tested namely the methyltransferase, the complex and trypsin. In the case of the Roab13 antibody-peptide complex the crystals obtained were of the highest resolution than ever attained previously. In addition, the oil-on-drop method overcomes problems encountered by other methods. It can be used with a wider variety of precipitating agents and can be performed using robots in a high throughput mode.

ACKNOWLEDGEMENTS

A sincere and heartfelt thank you to Professor Naomi Chayen for making this project available in her laboratory and for offering her never-ending guidance, mentorship and support as I progressed through my studies.

I would also like to thank Dr. Lata Govada for lending her support, direction, and guidance when needed. Her teachings and expertise have been invaluable towards the progress made thus far. Thanks are also extended to Dr. Sahir Khurshid for his assistance and availability with using the Douglas Instruments Oryx 8 automated crystallisation liquid handler and to Dr. Marc Morgan the X-ray facilities manager at the Centre for Structural Biology for his help and guidance on using the TTP Labtech Mosquito liquid handler. Also, his support with crystal preparation and data collection during Diamond Light Source beam-times were greatly valued and appreciated.

And finally, gratitude to Professor Benjamin Chain from University College London for kindly preparing and supplying the anti-CCR5 (RoAb13) antibody and the associated cognate peptide; Dr Nikos Pinotsis from Birkbeck College London for producing and supplying methyltransferase *Legionella pneumophila* - lpg293; and Dr Peter Zagalsky from Royal Holloway, University of London for supplying the lobster shell alpha crustacyanin protein.

My heartfelt gratitude to my wife for her continued and unwavering support, encouragement and understanding throughout this long but fruitful journey.

TABLE OF CONTENTS

DECLARATION OF ORIGINALITY	II
COPYRIGHT DECLARATION	II
ABSTRACT	III
ACKNOWLEDGEMENTS	IV
INDEX OF FIGURES	VIII
INDEX OF TABLES	X
INTRODUCTION	12
1.1 PROTEIN CRYSTALLISATION IN THE POST GENOMIC ERA.....	12
1.2 GENERAL PRINCIPLES OF CRYSTALLISATION	15
1.2.1 DYNAMIC EQUILIBRIUM AND THE FREE ENERGY BARRIER	15
1.2.2 THE CRYSTALLISATION PHASE DIAGRAM	16
1.2.3 CRYSTALLISATION STRATEGIES.....	18
1.3 CRYSTALLISATION METHODS.....	19
1.3.1 BATCH CRYSTALLISATION	20
1.3.1.1 MICROBATCH UNDER OIL	20
1.3.2 DIFFUSION METHODS	21
1.3.2.1 VAPOUR DIFFUSION	21
1.4 CHEMICAL PRECIPITANTS.....	22
1.4.1 COMMERCIAL SCREENS.....	23
1.5 OPTIMISATION APPROACHES	25
1.5.1 SEPARATION OF NUCLEATION AND GROWTH.....	25
1.5.2 PREVENTING EXCESSIVE NUCLEATION	25
1.5.3 NUCLEANTS	26
1.5.3.1 CARBON NANOMATERIALS AS POTENTIAL NUCLEANTS	29
1.5.3.2 CONTROLLED PORE GLASS (CPG) AS A POTENTIAL NUCLEANT	34
1.5.3.3 MOLECULAR IMPRINTED POLYMERS (MIPs)	35
1.6 MEDICALLY RELEVANT PROTEINS WORKED ON	37
1.6.1 METHYLTRANSFERASE <i>LEGIONELLA PNEUMOPHILA</i> - LPG2936.....	37
1.6.2 LOBSTER SHELL ALPHA CRUSTACYANIN (α -C)	39
1.6.3 RoAb13 ANTIBODY-PEPTIDE COMPLEX.....	40

WORKPLAN.....	43
2.1 PROJECT AIMS	43
2.2 PROJECT APPROACH	44
2.2.1 PREVENTING EXCESSIVE NUCLEATION WITH THE OIL BARRIER METHOD	44
2.2.2 NUCLEANT EXPERIMENTS	45
2.2.3 WORKING PHASE DIAGRAMS	45
2.2.4 OTHER OPTIMISATION STRATEGIES	45
2.2.5 X-RAY DIFFRACTION ANALYSIS.....	46
2.2.6 STATISTICAL ANALYSIS	46
MATERIALS & METHODS	47
3.1 PROTEINS AND BUFFERS	47
3.2 CRYSTALLISATION CONDITIONS FOR THE OIL BARRIER METHODS AND THE NUCLEANT TRIALS	48
3.3 CRYSTALLISATION PLATES AND DISPENSING SYSTEMS	49
3.4 OILS	51
3.5 OIL BARRIER EXPERIMENTS.....	52
3.6 NUCLEANT TRIALS	52
3.7 CHARACTERISATION OF DROPS	54
3.8 X-RAY DIFFRACTION	54
3.9 STATISTICAL ANALYSIS	55
RESULTS.....	57
4.1 PREVENTING EXCESSIVE CRYSTAL FORMATION: A NOVEL OIL BARRIER METHOD	57
4.1.1 MANUAL AND AUTOMATED TRIALS ON BENCHMARK AND TARGET PROTEINS	58
4.1.1.1 LYSOZYME.....	58
4.1.1.2 TRYPSIN.....	63
4.1.1.3 ALPHA CRUSTACYANIN	70
4.1.1.4 METHYLTRANSFERASE <i>LEGIONELLA PNEUMOPHILA</i> - LPG2936.....	75
4.1.2 SUMMARY OF MEAN CRYSTAL SIZES: CHAYEN METHOD VS OIL-ON-DROP METHOD IN HDVD	82
4.1.3 SUMMARY OF MEAN CRYSTAL SIZES: A COMPARISON BETWEEN THE ORYX 8 & THE MOSQUITO	83
4.2 HETEROGENOUS NUCLEANT TRIALS	85

4.2.1 THAUMATIN	85
4.2.2 CATALASE, TRYPSIN AND LYSOZYME	88
4.2.3 TARGET PROTEIN: – METHYLTRANSFERASE <i>LEGIONELLA PNEUMOPHILA</i> – LPG2936	89
4.3 OPTIMISATION TRIALS ON THE TARGET ROAB13-PEPTIDE COMPLEX	90
4.3.1 CONVENTIONAL OPTIMISATION: VARYING THE PROTEIN AND PRECIPITANT CONCENTRATIONS.....	90
4.3.2 VARYING INCUBATION PERIODS OF THE ANTIBODY AND PEPTIDE	92
4.3.2.1 INCUBATION TRIALS AT 10MG/ML.....	92
4.3.2.2 INCUBATION TRIALS AT 8MG/ML	93
4.3.3 SLOWING CRYSTALLISATION WITH THE OIL-ON-DROP METHOD	94
DISCUSSION	97
5.1 A NOVEL OIL BARRIER METHOD & AUTOMATION	97
5.1.1 SLOWING DOWN CRYSTALLISATION: ENHANCING SINGLE CRYSTAL GROWTH AND IMPROVING X-RAY DIFFRACTION WITH THE OIL-ON-DROP METHOD	97
5.1.1.1 METHYLTRANSFERASE <i>LEGIONELLA PNEUMOPHILA</i> – LPG2936	98
5.1.1.2 TRYPSIN.....	99
5.1.1.3 α -C.....	100
5.1.1.4 RoAb13-PEPTIDE.....	100
5.1.1.5 INCREASED CRYSTAL SIZE	101
5.1.2 OIL RATIO MIXTURES	102
5.1.2.1 MANUAL TRIALS.....	102
5.1.2.2 AUTOMATED TRIALS	104
5.1.3 OIL RUN-OFF DURING TRIALS.....	105
5.2 HETEROGENOUS NUCLEANT TRIALS	106
5.2.1 NEW NUCLEANTS	106
5.2.2 EXISTING NUCLEANTS: A QUICK COMPARISON WITH THE OIL-ON-DROP METHOD	108
5.3 TARGET PROTEIN ROAB13-PEPTIDE COMPLEX	111
FUTURE WORK.....	115
FUTURE PERSPECTIVES	118
BIBLIOGRAPHY	121

INDEX OF FIGURES

Figure 1. 1 Flowchart depicting the process from protein selection to potential drug development	15
Figure 1. 2 Energy barrier to protein crystallisation.....	16
Figure 1. 3 Crystallisation phase diagram	17
Figure 1. 4 Schematic representation of the microbatch technique.....	21
Figure 1. 5 The most popular vapour diffusion methods.....	22
Figure 1. 6 Schematic representation of the HDVD technique with the conventional Chayen oil barrier method.....	26
Figure 1. 7 Crystal nucleation in a pore	28
Figure 1. 8 Graphene sheet rolled up into a single-wall carbon nanotube (SWNT).....	29
Figure 1. 9 Multi-wall carbon nanotubes	29
Figure 1. 10 Transmission Electron Microscopic images of various CNMs...	32
Figure 1. 11 A catalogue of commercial unmodified and functionalized CNMs (Govada et al., 2016)	33
Figure 1. 12 Fabrication of MIPs. Schematic illustration of the steps involved in preparing MIPs for crystallisation studies	35
Figure 1. 13 Functional studies revealing RoAb13 Fab fragment antibodies bound to the native CCR5 receptor via a synthetic chimeric peptide...	42
Figure 3. 1 15 well EasyXtal plate with X-seal screw cap.....	49
Figure 3. 2 Automated systems for crystallisation trials	50
Figure 3. 3 Example screenshots of the Oryx 8 Front Panel software settings allowing for microtip tip height and directional alignment	51
Figure 3. 4 Scanning electron microscopic image of an individual CPG grain	52
Figure 3. 5 Scanning electron microscopic image of a CPG surface (8nm pore size)	53
Figure 3. 6 Scanning electron microscopic image of a CPG surface (48nm pore size)	53
Figure 4. 1 Schematic representation of the HDVD technique with the new oil-on-drop oil barrier	57
Figure 4. 2 Lysozyme crystal formation across methods and oil ratios	59
Figure 4. 3 Lysozyme crystal images of control drops vs oil-on-drops.....	60
Figure 4. 4 Lysozyme crystal count: Controls vs Oil-on-drops set up by the Oryx 8	61
Figure 4. 5 Lysozyme crystal images of control drop vs oil-on-drop.....	62
Figure 4. 6 Lysozyme crystal count: Controls vs Oil-on-drop set up by the Mosquito robot.	62
Figure 4. 7 Graph depicting lysozyme crystals using the Oryx 8 and Mosquito: Controls vs Oil-on-drop. Mean sample size (n) = 16.	63
Figure 4. 8 Trypsin crystals across methods and oil ratios.....	65
Figure 4. 9 Trypsin crystal images of control drops vs oil-on-drop method	66
Figure 4. 10 Trypsin crystal count: Controls vs Oil-on-drop using the Oryx 8	67
Figure 4. 11 Trypsin crystal images of control drops vs oil-on-drop method	68

Figure 4. 12 Trypsin crystal count: Controls vs Oil-on-drop using the Mosquito	68
Figure 4. 13 Graph depicting trypsin crystals using the Oryx 8 and Mosquito: Control vs Oil-on-drop. Mean sample size (n) = 15.....	69
Figure 4. 14 Alpha crustacyanin crystallisation across methods and oil ratios	71
Figure 4. 15 Alpha-C crystal images of control drop vs oil-on-drop.....	72
Figure 4. 16 Alpha crustacyanin crystal count: Controls vs Oil-on-drop using the Oryx 8.....	72
Figure 4. 17 Images of alpha-C protein drops representing one.....	73
Figure 4. 18 Alpha-C crystal count: Controls vs Oil-on-drop using the Mosquito	73
Figure 4. 19 Graph depicting alpha crustacyanin crystals using the Oryx 8 and Mosquito: Control vs Oil-on-drop. Mean sample size (n) = 16.	74
Figure 4. 20 Methyltransferase <i>Legionella pneumophila</i> - lpg2936 crystal across methods and oil ratios.....	75
Figure 4. 21 Methyltransferase <i>Legionella pneumophila</i> - lpg2936 X-ray data quality indicators.....	77
Figure 4. 22 Methyltransferase <i>Legionella pneumophila</i> - lpg2936 crystal images of control drop vs oil-on-drop.....	79
Figure 4. 23 Methyltransferase <i>Legionella pneumophila</i> - lpg2936 crystal count: Controls vs Oil-on-drop using the Oryx 8.....	79
Figure 4. 24 Methyltransferase <i>Legionella pneumophila</i> - lpg2936 crystal images of oil drop vs control drop.....	80
Figure 4. 25 Methyltransferase <i>Legionella pneumophila</i> - lpg2936 crystal count: Controls vs Oil-on-drops using the Mosquito	81
Figure 4. 26 Graph depicting methyltransferase <i>Legionella pneumophila</i> - lpg2936 crystals using the Oryx 8 and the Mosquito: Control vs Oil-on-drop. Mean sample size (n) = 16.....	82
Figure 4. 27 Mean crystal length (a) and width (b) taken from both the Oryx 8 and Mosquito in control drops; the Chayen method; and the oil-on-drop method	83
Figure 4. 28 Graph depicting mean crystal length using the Oryx 8 and Mosquito: Controls vs Oil-on-drop.	84
Figure 4. 29 Graph depicting mean crystal width using the Oryx 8 and Mosquito: Controls vs Oil-on-drop.	85
Figure 4. 30 Working phase diagram for thaumatin.....	86
Figure 4. 31 Thaumatin crystals with and without the CB ink.	87
Figure 4. 32 Thaumatin crystals produced with CPG and CB ink.....	88
Figure 4. 33 Working phase diagram for the RoAb13-Peptide complex	91
Figure 4. 34 RoAb13-Peptide complex crystals after a 2 week incubation period.	93
Figure 4. 35 Example of the anti-CCR5 RoAb13-Peptide complex crystals produced with controls and oil barrier methods	95
Figure 4. 36 Anti-CCR5 RoAb13-Peptide complex mean crystal count using the Chayen oil barrier and oil-on-drop methods with 80/20 oil ratio with 10 replicates (n=10) per method.....	95

INDEX OF TABLES

<i>Table 1. 1 Illustration of the attrition rate in obtaining diffraction quality crystals in the pipeline from clones to structures defined (2017)</i>	<i>13</i>
<i>Table 1. 2 List of nano-carbon substrates (adapted from Govada et al., 2016)</i>	<i>31</i>
<i>Table 3. 3 Composition of carbon black ink</i>	<i>53</i>
<i>Table 4. 4 Results of CPG nucleants with catalase, trypsin and lysozyme ..</i>	<i>89</i>
<i>Table 4. 5 Nucleant and resolution results on <u>Legionella pneumophila</u> - lpg2936</i>	<i>90</i>
<i>Table 4. 6 RoAb13-Peptide complex optimisation</i>	<i>91</i>
<i>Table 4. 7 The effect of the various incubation periods on crystal formation of the 8mg/ml RoAb13-Peptide complex at 2M ammonium sulphate</i>	<i>94</i>

ABBREVIATIONS

Heterogeneous nucleants

CNMs – carbon nanomaterials

MIP – molecular imprinted polymers

CPG – controlled pore glass

NPG – nanoporous gold

SWNT – single-wall carbon nanotube

MWNT – multi-wall carbon nanotube

CB – carbon black

GO – graphene oxide

1-D; 2-D; 3-D – one, two and three dimensions

Aq – aqueous

Solutions, buffers and precipitants

HDVD – hanging drop vapour diffusion

SDVD – sitting drop vapour diffusion

PEG – varying molecular weight polyethylene glycols

$\text{KNaC}_4\text{H}_4\text{O}_6 \cdot 4\text{H}_2\text{O}$ (short name NaKT) - potassium sodium tartrate

TRIS - tris(hydroxymethyl)aminomethane ($(\text{HOCH}_2)_3\text{CNH}_2$)

NaOH – sodium hydroxide

NaCl – sodium chloride

$\text{C}_2\text{H}_3\text{NaO}_2$ (short name NaOA_c) – sodium acetate

$(\text{NH}_4)_2\text{SO}_4$ – ammonium sulphate

$\text{HOC}(\text{CO}_2\text{NH}_4)(\text{CH}_2\text{CO}_2\text{NH}_4)_2$ – tri-ammonium citrate

H_2O – water

$\text{C}_7\text{H}_9\text{ClN}_2$ - benzamidine hydrochloride

CaCl_2 – calcium chloride

HEPES – (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) ($\text{C}_8\text{H}_{18}\text{N}_2\text{O}_4\text{S}$)

$(\text{CH}_3)_2\text{AsO}_2\text{H}$ – cacodylic acid

MPD - 2-methyl-2,4-pentanediol

EDTA – Ethylenediaminetetraacetic acid

DTT – Dithiothreitol

MES - 2-(*N*-morpholino)ethanesulfonic acid

X-Ray and data collection

Å – Angstrom

$I/\sigma(I)$ - signal to background noise ratio

$\text{CC}_{1/2}$ – a dataset is randomly split in half and Pearson's correlation coefficient is determined

$R_{\text{rim}}(R_{\text{meas}})$ – Redundancy-Independent Merging R factor - the precision of individual measurements, independent of how often a given reflection has been measured

1

INTRODUCTION

1.1 PROTEIN CRYSTALLISATION IN THE POST GENOMIC ERA

Proteins are the major machinery of life responsible for many of the body's functions such as oxygen transport, food digestion, muscle movements and many more. Human diseases often occur due to the malfunction of proteins therefore proteins are often the targets of therapeutic interventions. Protein functions are determined by their 3-D structure (McPherson, 1999; Bergfors, 2009), thus knowledge of protein structures is critical to the success of rational drug design. In nature, proteins have a variety of biological roles. Their complex functions can be explained, in part, by the vast combinatorial diversity created by their amino acids building blocks. Amino acid sequences alone are not sufficient to infer protein function and the three-dimensional structure of proteins is of equal importance. An understanding of the three-dimensional structure of proteins, its biogenesis, folding, and interactions with other molecules is required.

X-ray crystallography is the most powerful technique for protein structure determination at high resolution but totally relies on the production of high-quality crystals. The results from all major Structural Genomics projects are summarised in Table 1.1 showing the attrition rate in obtaining diffraction quality crystals and structures from the initial clones. To date, the production of useful crystals has continually been a major bottleneck to 3-D structure determination and with more proteins being discovered as drug targets this problem is becoming increasingly acute (Khurshid *et al.*, 2014).

Table 1. 1 Illustration of the attrition rate in obtaining diffraction quality crystals in the pipeline from clones to structures defined (2017)

Cloned	231 276
Expressed	127 587
Soluble	62 067
Purified	61 960
Diffraction crystals	12 590
Structure defined	6 920
Results compiled from TargetTrack (Structural Genomics centres and Protein Science Initiatives worldwide) (http://sbkb.org/tt).	

The major techniques used for determining protein structures are X-ray crystallography, nuclear magnetic resonance (NMR), electron microscopy (EM), cryo-EM and neutron crystallography. All these techniques have limitations. As mentioned before, X-ray crystallography relies on the production of high quality protein crystals, but it does deliver atomic and subatomic-level resolution once this bottleneck is overcome. X-ray crystallography is to date, the most powerful and favoured technique for protein and macromolecular structure determination as it provides the most direct way of forming 3D images of molecules at the highest resolution. The vast majority (~90%) of protein structures defined and deposited in the protein data bank (PDB) used X-ray crystallographic methods (Krishnan and Rupp, 2012; Wang *et al.*, 2017).

Nuclear magnetic Resonance (NMR) occurs when the nuclei of certain atoms are immersed in a static magnetic field whilst being exposed to a second oscillating magnetic field. Some atomic nuclei within molecules experience nuclear magnetic resonance whilst others do not, dependent on whether they possess a spin. Protein images as seen with electron microscopy (EM), cryo-electron microscopy (cryo-EM) and X-ray crystallography are not produced directly with the use of NMR, but through indirect structural information gathered by extensive data analysis and computer calculations. NMR spectroscopy resolves 3-D structures of proteins in solution but it cannot be applied accurately to macromolecules larger than 50-60 kilo-Daltons (kDa) (Krishnan and

Rupp, 2012) and the protein must remain stable for days under experimental conditions.

EM has the advantage of protein structure determination in a form that is free from crystallographic constraints but its main limitation is that its resolution limit is typically 5–10 Å, although higher resolution limits (~2–5 Å) have been achieved recently with EM, it is the exception rather than the rule. Cryo-electron microscopy (cryo-EM) is a new approach for studying molecular assemblies, which are often too large and flexible to be amenable to X-ray crystallography (Frank, 2002) but when compared to X-ray crystallography, cryo-EM has not been able to consistently match or better high resolution diffraction images. With neutron crystallography, the major bottleneck is growing large crystals with a minimum size of 300x300x300µm, which has limited the use of neutron crystallography for solving 3-D protein structures (Blakeley *et al.*, 2015).

The steps from the selection of a target protein through to structure determination leading to potential inferred use in commercial, drug development and research, is illustrated in the flowchart presented in Fig 1.1. As highlighted, protein purification and crystallisation are the major bottlenecks in this process. Due to these major bottlenecks and the complexity and multi-parametric nature of crystallisation, a very large number of proteins have not translated into successful structure determination (Pusey *et al.*, 2005).

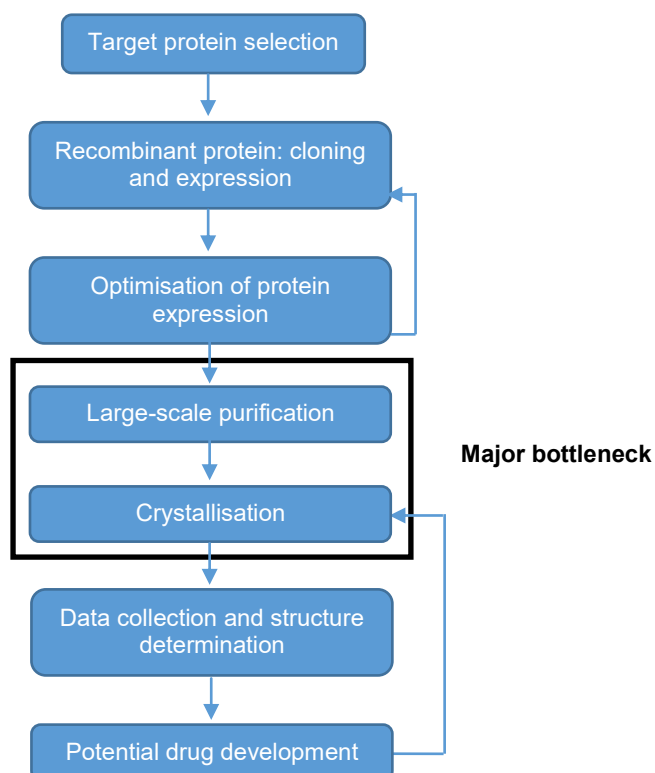


Figure 1. 1 Flowchart depicting the process from protein selection to potential drug development

1.2 GENERAL PRINCIPLES OF CRYSTALLISATION

1.2.1 DYNAMIC EQUILIBRIUM AND THE FREE ENERGY BARRIER

The crystallisation of molecules from a solution is an equilibrium event which is caused by lowering the free energy available in a particular system (Fig 1.2). Fully hydrated molecules in solution are at so-called system equilibrium. With the addition of precipitating agents, changes occur which leads to the dehydration of the molecules due to solvent insufficiency. This leads to a state of supersaturation and the system is no longer in equilibrium. Thermodynamics will create a new state of equilibrium with a new free energy minimum. Thus, an aggregation of molecules occurs as the solution becomes more concentrated and this leads to the formation of amorphous precipitate and/or crystal nuclei. Kinetically, amorphous precipitates are usually more favoured so it may dominate in the solid phase at the expense of crystal formation (Chirgadze, 2001). Typically, more energy is required to overcome the

free energy (ΔG) barrier to allow for the initial nuclei formation than is required for the eventual crystal growth from such nuclei.

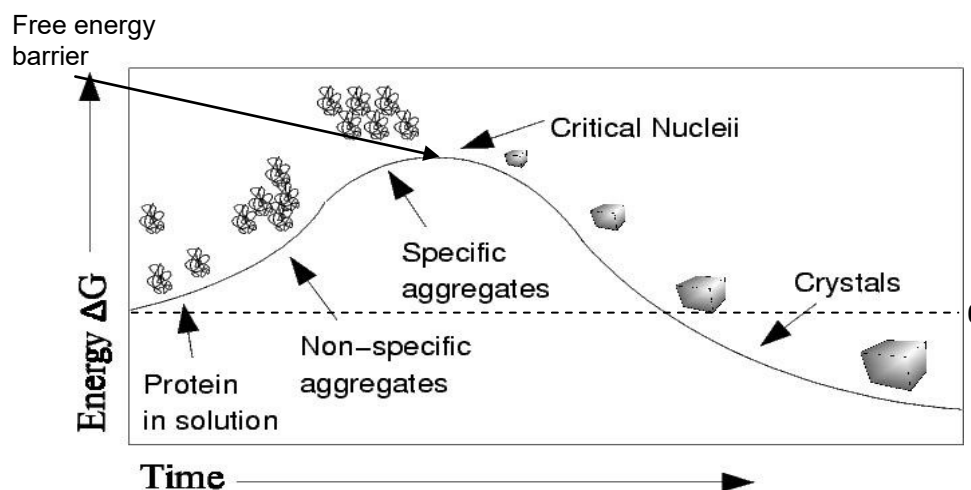


Figure 1. 2 Energy barrier to protein crystallisation

(Source: A.J. McCoy; © 1999-2005; University of Cambridge)

1.2.2 THE CRYSTALLISATION PHASE DIAGRAM

Crystallisation occurs as a phase transition phenomenon with different states (liquid, crystalline or precipitate) possible under a variety of crystallisation conditions (Fig 1.3). Chayen and Saridakis (2008) state that two main zones exist in a crystallisation phase diagram, namely undersaturation (where the protein is fully dissolved with no crystallisation events possible) and supersaturation where protein crystallisation can occur.

The supersaturated zone can be subdivided into a high supersaturation (also called the precipitation zone) – where the protein will precipitate; moderate supersaturation (nucleation zone) – where spontaneous nucleation occurs; and lower supersaturation (metastable zone) – found just below the nucleation zone, where crystals are stable and may grow but no further nucleation occurs. This zone contains the best conditions

for growing single, well-ordered crystals (Chayen, 2004; Chayen and Saridakis, 2008).

In an ideal situation, nuclei are formed when conditions are in the nucleation zone. Once a nucleus is formed, molecules from the solution attach to it leading to a reduced concentration of protein in the solution and to metastable conditions where the crystals continue to grow but no more nuclei can form. Excess nucleation (the formation of excess nuclei) is a problem in crystal growth, as this leads to large numbers of smaller crystals instead of a few single, high quality crystals for use in X-ray diffraction studies.

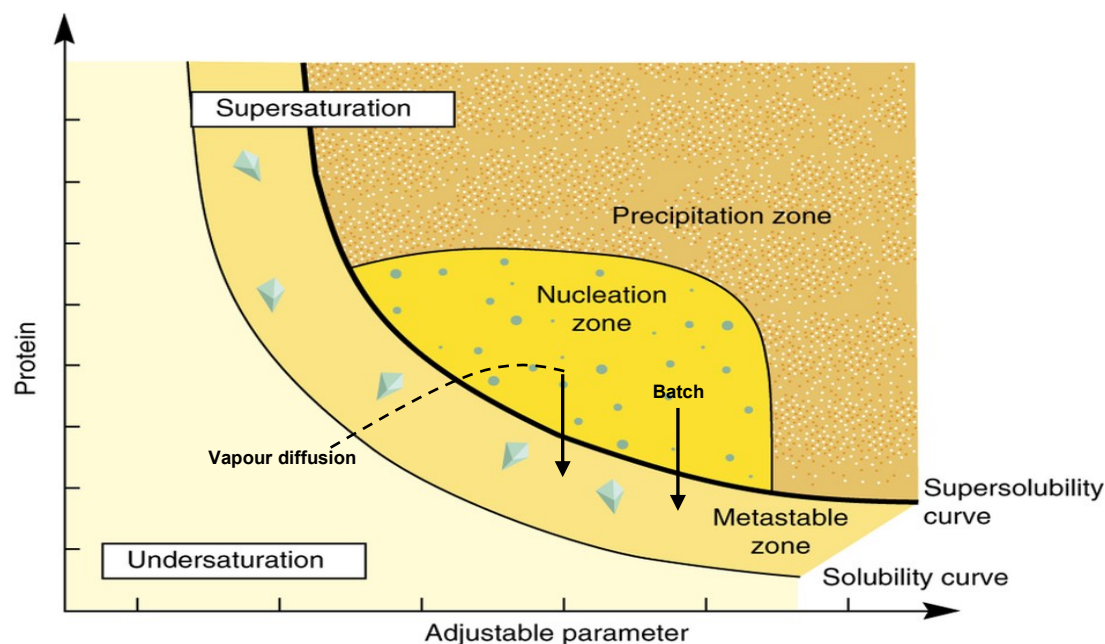


Figure 1. 3 Crystallisation phase diagram

The dashed line and black arrows represent the starting and finishing conditions for the vapour diffusion and batch crystallisation techniques. The solubility curve represents conditions where the concentration of the protein in the solute is in equilibrium with crystals. The super solubility curve separates the conditions where spontaneous nucleation (or phase separation, precipitation) occurs and conditions where the crystallisation solution remains clear if left undisturbed. Adjustable parameters can be precipitant concentration, buffer, pH, temperature, etc. Chayen (2004) *Current Opinion in Structural Biology* 14, 577–583.

Phase diagrams provide a means of quantifying the influence of parameters such as the protein and precipitant(s) concentrations, additive(s), pH and temperature on the production of crystals and form the basis for the design of crystal growth conditions for any crystallisation experiment (Ducruix and Giegé, 1992; Ataka, 1993). For practical purposes it is enough to generate working phase diagrams which involve the determination of the supersolubility curve (found between the nucleation and metastable zones) that separates the phases of crystal nucleation and crystal growth (Saridakis and Chayen, 2003). With these diagrams two parameters are varied (typically protein and precipitant concentrations) and their outcomes on crystallisation are plotted on a two-dimensional parameter grid. These experiments show where the crystallisation solution remains clear and where it spontaneously nucleates or precipitates (Fig 1.3).

1.2.3 CRYSTALLISATION STRATEGIES

The first step in crystallising a protein is initial screening where the protein of interest is exposed to a variety of conditions which may be conducive to crystallisation. Understanding the variety of variables affecting macromolecular crystallisation is critical to allow for effective and efficient crystallisation trials. The most important variables are listed below:

1. *Concentration of protein*
2. *Concentration of precipitant:* Both protein and precipitant are important variables used to bring the system into a state of relative supersaturation.
3. *pH:* a pH buffer is also an important factor involved in macromolecular crystallisation. Altering the pH indicates a change in the net charge of the protein. The most common buffers are intended to be effective in the pH range of 6.0 to 8.0, because the most physiologically important reactions occur near a neutral pH of 7.0-7.5. However, in some cases, proteins have been shown to crystallise under extreme pH conditions.

4. *Temperature:* Proteins have been documented to crystallise at a wide variety of temperatures. Protein solubility can increase, decrease, or remain constant as the temperature of the system fluctuates. It is always prudent to explore a range of temperatures if crystallisation does not occur at room temperature or at a set incubation temperature range (18-21°C).
5. *Additives:* Are generally small molecules that influence the intermolecular interactions of the sample and perhaps enhance crystallisation of macromolecules (Cudney and Patel, 1994). Crystal quality may improve in the presence of additives.

Any crystals, micro-crystals, precipitate and phase separation produced are seen as 'leads' and marked for further exploration by optimisation and fine-tuning trials. The multi-parametric nature of protein crystallisation allows for optimisation by varying protein and/or precipitant concentration, buffers, pH and temperature.

1.3 CRYSTALLISATION METHODS

Biological macromolecules are very sensitive to external conditions thus the usual methods (i.e. high pressure, drastic evaporation or drastic temperature variations) for crystallising organic and inorganic small molecules cannot be employed and must be replaced by gentler techniques. There are four major methods employed for growing protein crystals. Each of these methods aims to bring the solution of macromolecule to a supersaturated state and gets to the nucleation zone by a different route.

The two most commonly used techniques are vapour diffusion and microbatch. Other techniques that are used less frequently but worth a mention are dialysis and free interface diffusion (FID). Dialysis is based on diffusion and the equilibration of small precipitant molecules through a semi-permeable membrane that allows for the precipitant solution to

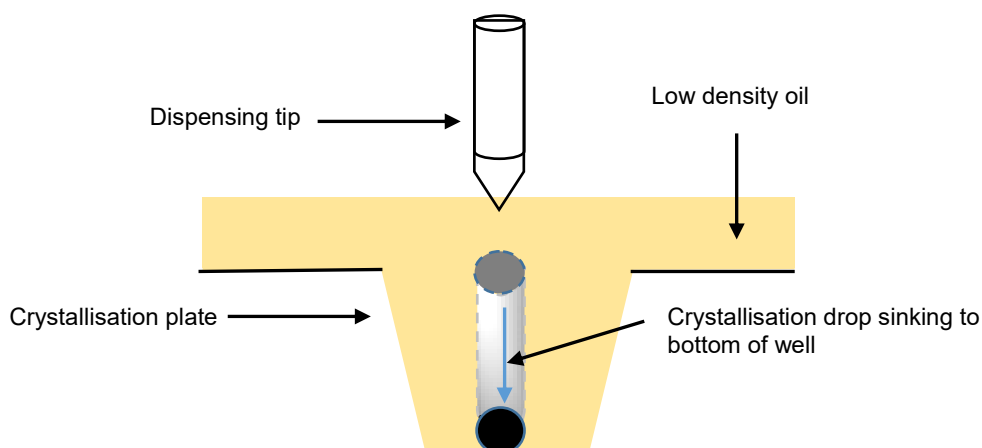
slowly mix with the protein molecules leading to supersaturation of the protein and crystallisation. FID consists of a protein and precipitating solution separated by a liquid interface in small capillaries, where diffusion across the interface saturates the protein solution leading to protein crystallisation (Chirgadze, 2001).

1.3.1 BATCH CRYSTALLISATION

Batch crystallisation is the oldest method used for crystallising proteins and involves direct mixing of the protein and precipitant solutions.

1.3.1.1 MICROBATCH UNDER OIL

Although not a method used in this study, it is worth mentioning the microbatch method as it is the second most used crystallisation method after vapour diffusion. In microbatch experiments, crystallisation trials are conducted under low-density paraffin oil or a combination of paraffin and silicon oil (also known as Al's Oil). Fig 1.4 illustrates the set up in which the protein of interest is mixed with a crystallising agent and then dispensed under the oil. As the drops are denser than the oil, they sink to the bottom of the wells and remains under the oil. This enables the solutions to be protected from evaporation, contamination and physical shock (Chayen *et al.*, 1990). Microbatch experiments are set-up by mixing protein and precipitant samples at supersaturated conditions, thus this is not a dynamic system such as vapour diffusion where equilibrium is achieved over time between the protein and precipitant solution (Chayen, 1997).



*Figure 1. 4 Schematic representation of the microbatch technique
(Adapted from Chayen, 1997)*

1.3.2 DIFFUSION METHODS

1.3.2.1 VAPOUR DIFFUSION

Vapour diffusion is the most widely used method for crystallising proteins and the primary technique used by X-ray crystallographers. In principle vapour diffusion is based on the diffusion and evaporation of water molecules between solutions of different concentrations. The protein solution is typically mixed with a crystallising agent to form a drop. The drop is equilibrated against a reservoir containing that crystallising agent at a higher concentration. The difference in precipitant concentration between the drop and the reservoir solution causes water molecules to evaporate from the drop until the concentration of the precipitant in the drop is equilibrated to that of the reservoir solution. This then leads to concentration of the protein from an undersaturated to a supersaturated state.

Ideally protein crystals will appear in the protein drop until the concentration of the protein reaches its solubility limit (Ducruix and Giegé, 1999). In practical terms, vapour diffusion can be performed as hanging drops, sitting drops and sandwich drops as illustrated in Fig 1.5. Protein drops are placed in different positions dependent on which vapour diffusion

method is utilised. The inevitable evaporation is minimised by sealing the wells.

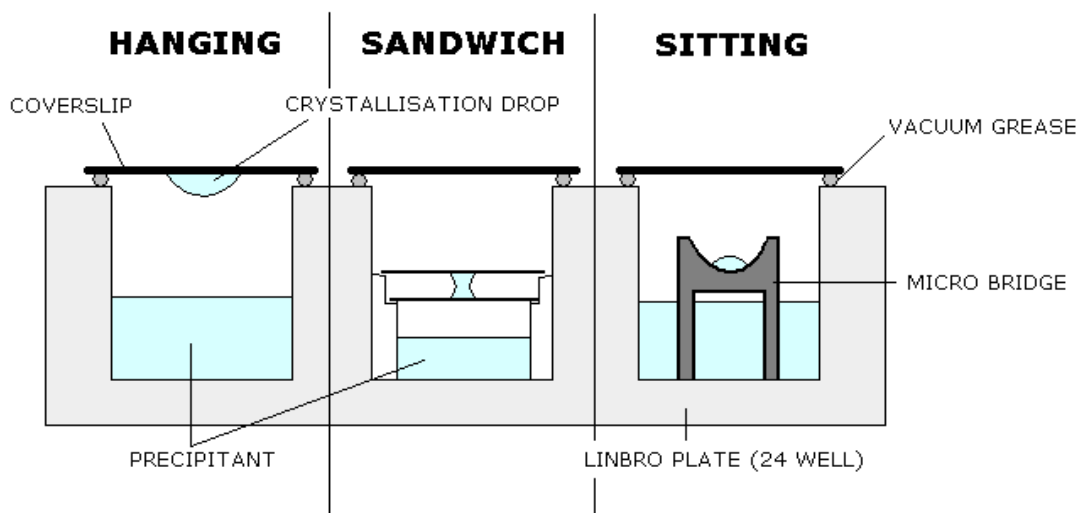


Figure 1. 5 The most popular vapour diffusion methods (adapted from Chirgadze, 2001)

The three variations are (i) hanging drop: where the crystallisation droplet is dispensed onto a siliconised glass coverslip and then inverted over a reservoir containing the precipitant solution; (ii) sandwich drop: which are created by placing the crystallisation drop in the middle of two glass coverslips of differing dimension; (iii) sitting drops: which involve the crystallisation droplet being placed on a platform surrounded by the precipitant solution.

1.4 CHEMICAL PRECIPITANTS

The general function of any precipitation reagent is the binding of water molecules, which leads to insufficient water content to maintain the hydration of the macromolecules. Chemical precipitants are the most widely used agents to assist with this dehydration process. These precipitants are broadly classified into the following categories:

1. *Salts:* Historically, salts have been the most effective precipitant and ammonium sulphate the most widely used. The effects of salts on protein solubility are complex and the change of protein solubility at increasing salt concentrations was studied in what was coined as 'salting-in' and 'salting-out' (Ducruix and Giegé, 1999). The ability of

salts to precipitate proteins was observed while carrying out experiments on hen egg white proteins. It was found that salts can either be lyotropic (when they reinforce the structures of water and biological macromolecules), or chaotropic (when they denature macromolecules) (Curtis *et al.*, 2002).

2. *Polymers*: High molecular weight polymers such as polyethylene glycol (PEG) and polyvinyl alcohol (PVA) are able to precipitate almost all proteins without causing denaturation (Polson *et al.*, 1964). PEGs are produced in a variety of molecular sizes ranging from 200 – 15 000 monomers, and also as mono- and di-methyl ethers. PEGs are the most extensively used precipitants in crystallisation due to their (a) solubility in both aqueous solutions, (b) capacity to bind water molecules, and (c) ability to link to macromolecules via chain end functional groups (Vert and Domurado, 2000). PEGs are also very cost effective.
3. *Organic solvents and non-volatile alcohols*: These solvents have the ability to precipitate molecules by lowering the chemical activity of water. They also have the ability to denature the proteins and thus require careful use during crystallisation trials. Common examples include ethanol, acetone, propanol and MPD.

Traditionally, most of the precipitants mentioned above were used independently during crystallisation trials, but recently, a large variety of commercial screens have become available which use a combination of these different precipitants as cocktails to determine the crystallisation potential of the protein during screening experiments.

1.4.1 COMMERCIAL SCREENS

Finding optimal conditions for macromolecular crystal nucleation and growth are difficult to predict and time consuming. Screening is an efficient and convenient tool used by researchers in this field. A range of commercial crystallisation screening kits are now available and have

become the norm in crystallisation trials as they provide a rapid screening method for determining pre-crystallisation conditions for macromolecules of interest. A few popular commercial screens are listed below:

1. *Sparse matrix Screen*: This screen uses known crystallisation conditions of previously successfully crystallised molecules. It allows for a quick test of a wide range of pH, salts and precipitants and it utilises a very small sample volume. This screen was first developed by Jancarik and Kim (1991).
2. *PEG/Ion Screen*: This screen is a crystallisation reagent kit from developed by Hampton Research (<https://www.hamptonresearch.com/>) with high purity PEG (3350) and 48 unique salts of varying ionic strengths and pH as its primary screening variables. It represents a complete range of cations and anions.
3. *Index Screen*: The index screen is another product from Hampton Research (<https://www.hamptonresearch.com/>) using a systematic approach containing 96 pre-formulated reagents which combines the strategies found in the sparse matrix, incomplete factorial and grid screens.
4. *ZetaSol™ Screen*: This screen determines the crystallisation potential of a protein based on its net charge and the chemical nature of the precipitant (<https://www.moleculardimensions.com/>).
5. *Stura Footprint Screen*: The footprint screen uses the concept of screening the protein of interest by analysing the protein's relative solubility with precipitants that have successfully crystallised many proteins (<https://www.moleculardimensions.com/products/2186-Stura-FootPrint-Screen/>).

1.5 OPTIMISATION APPROACHES

There are several optimisation approaches relevant to this project including the separation of nucleation and growth; the use of heterogeneous nucleants; and the use of oils to control the speed of crystallisation.

1.5.1 SEPARATION OF NUCLEATION AND GROWTH

The optimal conditions for protein crystal nucleation are far from ideal for their subsequent growth. This is due to the fact that spontaneous nucleation is more likely to occur at high levels of supersaturation, whereas a slow, more ordered growth of larger crystals are favoured by lower concentration levels. The ideal experiment therefore requires the separation of nucleation from growth to satisfy the different requirements of the two events (Bergfors, 2003).

The separation of nucleation and growth can be performed experimentally in several ways. The dilution of the crystallisation drops in microbatch experiments was described by Saridakis *et al.* (1994) and was extended to vapour diffusion through the dilution of the precipitant reservoir. Another example described by Saridakis and Chayen in vapour diffusion involved the transfer of cover slips from higher concentration reservoirs in crystallisation experiments to lower precipitant concentration reservoirs (which correspond to metastable conditions) (Saridakis and Chayen, 2003).

1.5.2 PREVENTING EXCESSIVE NUCLEATION

The formation of excess nuclei is a major problem in crystal growth, as this leads to a large number of small crystals or microcrystal showers instead of a few single, high quality crystals that would be suitable for X-ray analysis. This often happens because the crystallisation process takes place too rapidly. This process can be slowed down by dispensing a layer

of a mixture of paraffin and silicone oils (between 100-300 μ l) over the equilibrating precipitant reservoir (Fig 1.6).

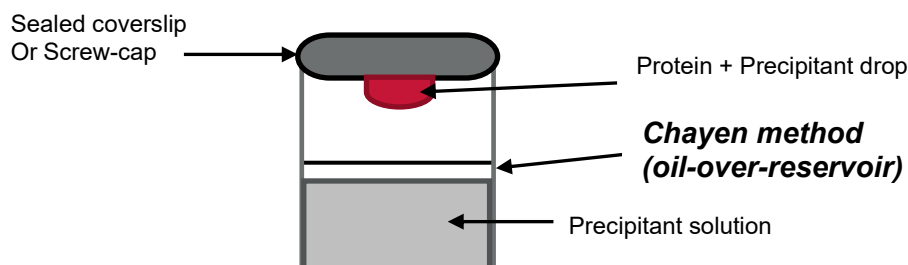


Figure 1. 6 Schematic representation of the HDVD technique with the conventional Chayen oil barrier method

This current method (Chayen, 1997) has been successful in limiting excessive crystal formation; leading to fewer, larger well-formed crystals resulting in better X-ray diffraction. However, this technique has several limitations: it requires the use of microlitre quantities of protein and reagents (rather than nanolitre quantities); it does not work when using organic solvents such as polyethylene glycols (PEGs) above 15% and 2-methyl-2,4-pentanediol (MPD) as the protein drops dry out within days of the initial trial setup; and it is not amenable to automation and high throughput trials.

1.5.3 NUCLEANTS

In attempts to control nucleation researchers have been introducing substances into protein solutions, called heterogeneous nucleants (Saridakis and Chayen, 2009). Nucleants work by encouraging nucleation of the protein molecules on the nucleant surface (Blow *et al.*, 1994), especially those nucleants with porous surfaces from which crystals can nucleate and grow. By triggering nucleation in this way, nucleants serve to bypass the requirement for the nucleation zone of the phase diagram (Fig 1.3) and are therefore added directly to metastable conditions (metastable zone).

The introduction of nucleants in protein drops assists in lowering the free energy (ΔG) barrier (i.e. the nucleation barrier) to initial crystal nucleation, which protein solutions without such nucleants would struggle to overcome under similar conditions (Fig 1.2). In practice, when setting up nucleant trials, protein drops are set up at metastable conditions where spontaneous nucleation does not occur (or as near to metastable conditions as possible). This allows for effective testing of potential nucleants which may facilitate crystal nucleation compared to control drops without any nucleants present. In addition, any nuclei formed find themselves in an ideal environment for slow ordered growth by accretion with growth units freely moving toward the formed crystal and attaching to its surfaces under appropriate conditions. This leads to the growth of larger single crystals which lead to improved crystal quality and better diffraction if impurities and imperfections are not incorporated into the crystal structure during the growth phase (García-Ruiz, 2003).

Over the last 30 years, researchers have used substances such as minerals (McPherson and Shlichta, 1988); etched silicon (Chayen *et al.*, 2001); bioglass with nanoscale pores (Chayen *et al.*, 2006); microporous synthetic zeolites (Sugahara *et al.*, 2008); polystyrene nanospheres (Kallio *et al.*, 2009); nanoporous gold (NPG) (Kertis *et al.*, 2012); carbon nanotube-based films (Asanithi *et al.*, 2009); molecularly imprinted polymers (MIPs) (Saridakis *et al.*, 2011; Khurshid *et al.*, 2015); and 3-D nanotemplates (Shah *et al.*, 2015). Numerous other materials such as equine and human hair (D'Arcy *et al.*, 2003) have also been tested.

In practice, an ideal nucleant should be able to:

- Act as a nucleant for many proteins (benchmark and target), not just one protein. Benchmark proteins are proteins that have previously been crystallised and are used to test and validate crystallisation conditions and methods),
- Allow control over the number of nuclei that form (promoting the growth of only a few crystals), and

- Nucleate crystals under conditions close to ideal growth conditions.

To date, the search continues for a 'universal nucleant' that would allow efficient nucleation of crystals across a wide range of macromolecules in a controlled way. Nanoporous bioglass (commercially available as "Naomi's nucleant") developed in the Chayen laboratory in collaboration with others has been successful in crystallising 14 proteins thus far. This is the highest number of proteins crystallised by a single nucleant and includes proteins which have previously been difficult to crystallise.

Crystallisation experiments using nucleants with non-porous surfaces have proven less successful at promoting nucleation, which have led to the belief that it is the pores which are promoting and accelerating nucleation (Chayen *et al.*, 2006). Furthermore, it is thought that a variety of pores sizes is required, as nucleants with a distribution of pores sizes have been effective inducers of crystal formation. Practical experiments and simulations indicated that nucleation occurs at a faster rate in pores compared to nucleants with smooth surfaces. It was further noted that nucleation occurs in two steps: (a) pore-filling; and (b) subsequent growth out of the pore (illustrated in Fig 1.7 a-c).

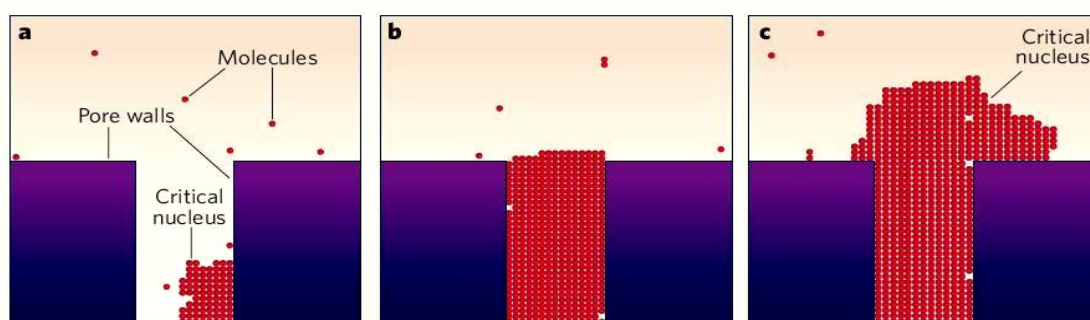


Figure 1. 7 Crystal nucleation in a pore

A model explaining how porous materials induce protein crystal nucleation. (a) The pore walls weakly attract the individual protein molecules. A critical nucleus is formed with subsequent crystal growth in the pore corners (b) Crystal growth continues and fills the pore (c) Another critical nucleus of molecules forms outside the pore, allowing for the growth of the 'bulk' crystal (Page & Sear, 2006).

1.5.3.1 CARBON NANOMATERIALS AS POTENTIAL NUCLEANTS

Carbon nanotubes and other nanomaterials (CNMs) such as graphene (Fig 1.8) and graphite have attracted major interest in the scientific, industrial and commercial industries due to their intrinsic structural, mechanical and electrical properties (Tran *et al.*, 2007; Singh *et al.*, 2011). These materials have the potential for a broad range of applications in engineering, bio-nanotechnology, catalysis (Menzel *et al.*, 2010), and even assisting with polymer crystallisation (Kim *et al.*, 2010). The use of CNMs has not yet fully translated to the field of macromolecular crystallisation.

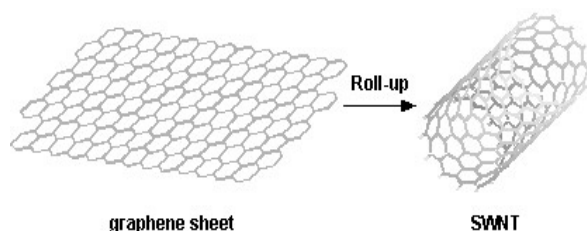


Figure 1. 8 Graphene sheet rolled up into a single-wall carbon nanotube (SWNT)

(Source: Odom *et al.*, (2002))

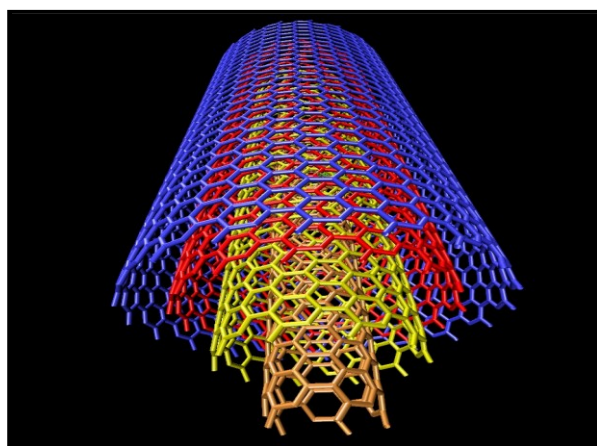


Figure 1. 9 Multi-wall carbon nanotubes

Four individual cylinders of CNTs make up this MWNT. (Source: Asst. Prof. Alain Rochefort; © 2013 (CERCA))

CNMs are particularly interesting as potential substrates for protein nucleation as they provide a highly accessible surface area coupled with large porosity through the formation of robust open networks. They also possess a highly tuneable degree of curvature, both positive (external surface) and negative (internal surface of opened CNMs), that can be

selected using diameters ranging from that of single polymer molecules (<1nm) to effectively flat at the molecular scale (>100nm) (Balasubramanian and Burghard, 2005).

Also, they have a linear morphology that has been shown to interact strongly with a wide range of polymers, including polypeptides, to create unusual ordered structures with a strong dependence on the specific nanotube type. From what we know from the work pioneered by the Chayen Group (Chayen *et al.*, 2006), the heterogeneity of many CNM samples may be a significant advantage for use as nucleants. In addition, a wide variety of chemical modifications can readily be introduced on the CNM surfaces (Menzel *et al.*, 2010), which may further enhance the potency for protein crystal nucleation.

Previous work conducted by the Chayen Group on the use of CNMs showed that CNM diversity (i.e. different CNM material types, geometries, diameter sizes, porosity, and chemical functionalisation) affected crystal nucleation differently when trialled with various proteins (Govada *et al.*, 2016). Some of the CNMs and their characteristics and chemical modifications used in that study are listed in Table 1.2 (a variety of nano-carbon substrates); in Fig 1.10 (transmission electron microscopic (TEM) images of functionalised individualised and bundling MWNTs and CB nanoparticles functionalised with PEG) and Fig 1.11 (a catalogue of commercial/unmodified/functionalized CNMs).

With reference to Fig 1.10, key results from that study showed CNM nucleant effectiveness as follows:

- **CNM 20** (CB - mPEG 5K) – was the most potent nucleant - effective with three model proteins (trypsin, thaumatin, lysozyme) and the antibody RoAb13
- **CNM 18** (GO) – was less effective than CNM20 with success in two proteins (lysozyme, trypsin). Induced nucleation in metastable zone

and promoted nucleation at borderline metastable conditions (thaumatin, RoAb13)

- **CNMs 6 and 8** (functionalised MWNTs) - induced nucleation at metastable conditions for two proteins (thaumatin and catalase for CNM 6; thaumatin and lysozyme for CNM 8)
- **CNMs 4, 10, 11, 13** (functionalised MWNTs) - effective for only one protein each (thaumatin for 4 and 10; lysozyme for 11, 13). From those, CNM 11 also had a marginal nucleation inducing effect for thaumatin and catalase, and CNM 4 for lysozyme)

Table 1. 2 List of nano-carbon substrates (adapted from Govada et al., 2016)

Carbon material	Description	Surface area before functionalisation (m ² /g)	Surface area after functionalisation (m ² /g)
Multi-wall carbon nanotubes MWNT (A)	CVD grown commercial Arkema ® CNTs, D approx. 10nm & several microns in length	220	180
Multi-wall carbon nanotubes MWNT (B)	Injection CVD grown carbon nanotubes D approx. 100nm & several tens of microns in length	30	<30
Graphene oxide GO	Heavily-oxidised, hydrophilic graphene	60	-
carbon black CB	Common form of amorphous carbon, agglomerate size 100-500nm, average primary particle size 10nm	270	220

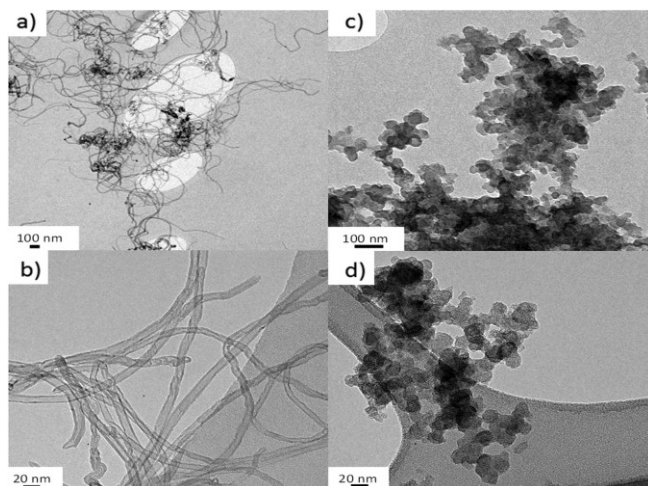


Figure 1.10 Transmission Electron Microscopic images of various CNMs (a) functionalised individualised and bundling MWNTs; (b) higher resolution image of the network formation of MWNTs; (c, d) CB nanoparticles functionalised with PEG forming agglomerates of nanoparticles between $1\mu\text{m}$ and 100nm (Govada et al., 2016)

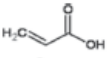
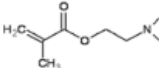
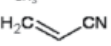
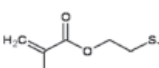
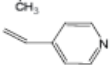
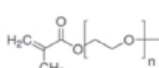
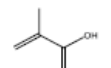
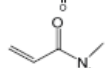
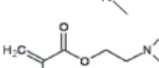
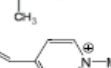
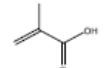
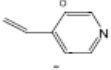
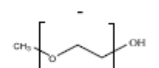
Label	Carbon Material	Grafted Molecules	Functionalising Group	Grafting Method ^a	Grafting conc. ^b ($\mu\text{mol m}^{-2}$)	IEP (pH)
CNM 1	MWNT (A)	ungrafted	-	U	n/a	4.9
CNM 2	MWNT (A)	thermally oxidised	-	T	n/a	-
CNM 3	MWNT (A)	acrylic acid (AA) ^c		T	6.6	-*
CNM 4	MWNT (A)	2-(dimethyl amino) ethyl methacrylate (DMAEMA) ^c		T	1.9	6
CNM 5	MWNT (A)	acrylonitrile (AN) ^c		T	8.2	-*
CNM 6	MWNT (A)	2-(methylthio) ethyl methacrylate (MTEMA) ^c		T	0.7	-*
CNM 7	MWNT (A)	4-vinyl pyridine (4-VP) ^d		T	2.3	7
CNM 8	MWNT (A)	acid oxidised	-	A	15.4 ^e	- ^s
CNM 9	MWNT (A)	poly(ethylene glycol) methacrylate (PEGMA) ^d		T	2.5	4.5
CNM10	MWNT (A)	methacrylic acid (MAA) ^d		T	2.3	4.3
CNM 11	MWNT (A)	N,N-dimethyl acrylamide (DMAA) ^d		T	1.7	5.2
CNM 12	MWNT (A)	2(dimethyl amino) ethyl methacrylate (DMAEMA) ^c		T	1.9	6.0
CNM 13	MWNT (A)	N-methyl-4-vinyl pyridine (MeVP) ^d		T	5.6	≥ 9.8
CNM 14	MWNT (B)	Ungrafted	-	U	n/a	3.6
CNM 15	MWNT (B)	methacrylic acid (MAA) ^d		T	3.1	-*
CNM 16	MWNT (B)	4-vinyl pyridine (4-VP)		T	9.8	-*
CNM 17	Graphite	ungrafted	-	U	n/a	-
CNM 18	GO	-	-	U	20.5 ^f	3.0
CNM 19	CB	ungrafted	-	U	n/a	5.1
CNM 20	CB	mPEG (5kDa)		R	1.2	4.2

Figure 1. 11 A catalogue of commercial unmodified and functionalized CNMs.

^achemical grafting methods used: U – ungrafted; T – thermochemical; R – reduction; A – acid oxidized. ^bgrafting concentration calculated from polymer grafting ratio and Brunauer–Emmett–Teller (BET) measurements. A range of oligomers were tested which included anionic, cationic, and non-ionic moieties, as well as mixtures of hydrophilic and hydrophobic species and different functional groups (such as thio-ether, amines, and hydroxyls). The functionalisation groups were grafted onto the CNMs via the various grafting methods listed. Isoelectric points (IEP) of each CNM are also listed representing each CNM's net electrical charge (Govada et al., 2016).

Following on from the pioneering work done by Govada *et al.* (2016) on using CNMs as substrates for protein crystallisation, the use of a new form of carbon black was investigated, namely CB ink as a potential nucleation-inducing agent. Carbon black ink is a soluble non-PEGylated derivative of CB carbon nanoparticles which is designed for inkjet printing systems (Table 1.3 in Chapter 3, subsection 3.4 lists the composition of CB ink). Key findings are presented in Chapter 4 (Results) and Chapter 5 (Discussion).

1.5.3.2 CONTROLLED PORE GLASS (CPG) AS A POTENTIAL NUCLEANT

Another nucleant of interest is controlled pore glass (CPG) which is 96% silica glass containing pores in the nanometre (nm) size ranges. It is produced through a phase separation in borosilicate glasses followed by a liquid extraction of the borate (Schnabel and Langer, 1991). Due to its high surface area, mechanical and thermal stability, efficient mass transfer and chemical inertness, it has effectively been used in chromatography columns (gas, thin layer and affinity) as a stationary phase to help separate compounds according to molecular weight (Schnabel and Langer, 1991) and also in polymer science as a vehicle to study different crystallisation outcomes and polymorphs in nanoscale confinement studies (Hamilton *et al.*, 2012). Commercially, CPG exists as dry white powder/grains with pore sizes ranging from a few nm up to 400nm (4000 Å). Another benefit of CPG is the ease at which the grains can be suspended and dispensed into solutions.

To the best of my knowledge, CPG has not been used in the field of protein crystallisation and affords a unique opportunity to test its viability as a nucleating agent in protein crystallisation trials. As noted before, previous research showing that protein nucleation can be facilitated by porous materials (Chayen, 2001; Chayen *et al.*, 2006; Asanithi *et al.*, 2009; Kertis *et al.*, 2012; Shah *et al.*, 2015) indicates that CPG is a worthwhile material to test in crystallisation trials.

1.5.3.3 MOLECULAR IMPRINTED POLYMERS (MIPs)

Molecularly imprinted polymers (MIPs), also referred to as “smart materials,” are polymers that are formed in the presence of a molecule that is extracted afterward, thus leaving complementary cavities (or ghost sites) behind. The molecular imprint that remains serves as a “memory effect” in the gel after the molecule is removed (Fig 1.12). Thus, the created cavities exhibit highly selective rebinding of the given molecule with which it was imprinted (Sellergren, 2000; Arshady and Mosbach, 1981).

MIPs were initially used for separation and purification of small molecules (Sellergren *et al.*, 1985); separation of carbohydrate derivatives (Nilsson *et al.*, 1995), and in thin layer chromatography (Suedee *et al.*, 1998). More recently MIPs have become an important tool in the preparation of artificial recognition materials capable of mimicking natural systems (Shi *et al.*, 1999). In the context of proteins, MIPs have been used for protein purification/isolation applications (Liao *et al.*, 1996); replacement of biological antibodies in immunoassays (Hansen, 2007); catalysis (Bruggemann, 2002), and biosensors for medicine (Hillberg and Tabrizian, 2008). MIPs, however, have never been used to facilitate protein crystallisation.

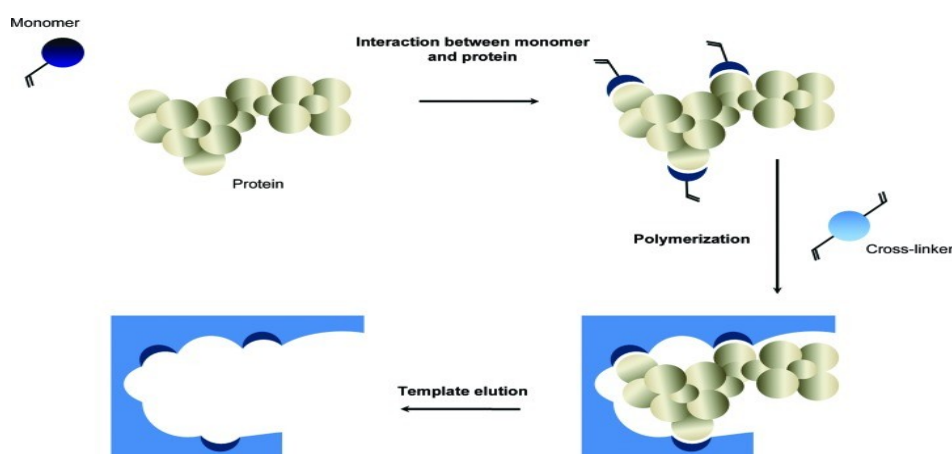


Figure 1. 12 Fabrication of MIPs. Schematic illustration of the steps involved in preparing MIPs for crystallisation studies

The initial assembly between the protein template and monomer is advanced through the presence of a polymerizing cross-linker. The protein is then eluted, leaving a protein-specific cavity known as a ‘ghost site’ (source: Khurshid *et al.*, 2015).

A collaboration between the Chayen and Reddy laboratories in 2011 led to the first macromolecular MIPs to be used in protein crystallisation. It was initially envisaged that these materials would be protein specific, meaning that a MIP would only be suitable for the cognate protein it was imprinted with, for example, a trypsin MIP would solely be used to crystallise the trypsin protein.

However, in practice there have been unexpected successes with other proteins of similar molecular weight to the cognate also being nucleated successfully. This reinforced the strong correlation between cavity/pore size of nucleants and the hydrodynamic radius of protein molecules in solution which has been a recurring theme in heterogeneous nucleant research (Page and Sear, 2006). Results demonstrated the induction of nucleation of nine proteins which included three target proteins under metastable conditions as well as increasing the number of crystal leads obtained from screening using MIPs on another three target proteins (Saridakis *et al.*, 2011).

Further to this, Khurshid *et al.* (2015) modified the MIPs to allow for high-throughput trials thus improving crystal quality of two target proteins and also increasing the probability of success when screening for suitable crystallization conditions during screening trials. With the success of MIPs in protein crystallisation within the Chayen laboratory since 2011 (Saridakis *et al.*, 2011; Khurshid *et al.*, 2015), the use of MIPs in optimisation and screening experiments has increased and has proven to be a valuable addition to our method development strategies and also towards structure determination of target proteins by improving X-ray diffraction of such proteins.

1.6 MEDICALLY RELEVANT PROTEINS WORKED ON

Finding crystallisation conditions for target proteins can be compared to searching for a needle in a haystack. As mentioned earlier in the text, the first step in the crystallisation process involves screening for 'hits' that are conducive to crystallisation. A lot of effort has been channelled into the automation of screening procedures in an attempt to streamline the process to minimise costs and the time that it takes to crystallise proteins. As such, the past fifteen years have seen some of the greatest achievements in the field of protein crystallisation by way of automation and miniaturisation.

However, the bottleneck of crystallisation still remains, where high throughput has not always led to high output. Crystallisation screening is still a 'trial and error' process and even though automation has proven invaluable for finding initial leads, the production of useful crystals for X-ray diffraction studies is still elusive (McPherson and Gavira, 2014). Addressing this issue remains crucial and it is increasingly clear that a more robust scientific approach needs to be applied.

1.6.1 METHYLTRANSFERASE *LEGIONELLA PNEUMOPHILA* - LPG2936

Methylation is the catalysis and transfer of a methyl group by enzymes from its substrate to an acceptor. In the case of ribosomal RNA (rRNA) such transfer is mediated by methyltransferases (MTases) and plays an important role in biogenesis and activity regulation of the ribosome (Decatur and Fournier, 2002). A variety of these enzymes act during ribosomal RNA maturation to modify nucleotides in a site-specific manner (Del Campo *et al.*, 2004). There are three types of rRNA modifications that can occur: base methylation, pseudouridylation, and 2'-O-methylations. Base methylation is the most frequent type of modification in the rRNAs from bacteria (Decatur and Fournier, 2002).

Legionella pneumophila is a Gram-negative bacterium that infects freshwater protozoa and replicates within them. Several different species of amoebae and protozoa have been shown to be environmental hosts for legionellae (Rolando *et al.*, 2015). When humans inhale water droplets or drink water contaminated by *L. pneumophila*, the bacterium can cause infection in host's lung and cause a severe form of pneumonia called Legionnaire's disease. *Legionella pneumophila* relies on a highly specialized type 4 secretion system (T4SS) to infect and persist in their hosts which can be fatal if undiagnosed or untreated by antibiotics. It has been shown that *Legionella pneumophila* targets the host cell nucleus and modifies the transcriptional response to the pathogens advantage (Rolando *et al.*, 2015).

It follows that methyltransferase *Legionella pneumophila* - lpg2936 belongs to a family of RNA methyltransferases (MTases) called RNA small subunit methyltransferase (RsmE). Based on amino-acid sequence conservation, MTases can thus be grouped into functional classes such as the ribosomal RNA small subunit methyltransferase A (RsmA class); the RsmB class; the RsmC class; or RsmE class (Pinotsis and Waksman, 2017; Schuhmacher *et al.*, 2018). Likely homologues of RsmE have been reported in several pathogenic bacteria including *Legionella pneumophila*. *Legionella* bacteria utilises its secretion system to secrete hundreds of proteins, also known as effectors, into the infected host. Effector proteins represent about 10% of the entire *Legionella pneumophila* genome made up of about 3000 genes. Therefore ribosomal activity and protein synthesis are expected to be essential for *Legionella pneumophila* pathogenic mechanisms (Pinotsis and Waksman, 2017; and references therein).

The first functional study of the RNA MTase RsmE family members based on their structures was published by Zhang *et al.*, (2012) which allowed them to gain critical insights into RNA MTase RsmE class catalytic process and thus providing a further understanding on the structure–function relationship of the conserved rRNA MTase (Zhang *et al.*, 2012). Pinotsis

and Waksman (2017) solved the structure of the methyltransferase *Legionella pneumophila* - lpg2936 to 1.5 Å and has allowed for further insights into the mechanism of RsmE family methyltransferases. Before the structure was solved, work on methyltransferase *Legionella pneumophila* - lpg2936 was performed as part of this thesis and presented within.

1.6.2 LOBSTER SHELL ALPHA CRUSTACYANIN (α -C)

Crustacyanins are members of the lipocalin family of hydrophobic ligand-binding proteins (Britton *et al.*, 1982). α -Crustacyanin is the carotenoid-protein complex responsible for the blue-black colouration of lobster *Homarus gammarus* carapace (carapace is a hard layer or shell that covers and protects animals such as crabs and turtles). Astaxanthin is one of the many cancer-protective carotenoids found naturally in orange and red fruits and vegetables as well as dark leafy greens and wild salmon. Astaxanthin is the carotenoid partner of α -crustacyanin (Britton *et al.*, 1982). As well as their principal role in photosynthetic processes, carotenoids provide bright colouration, serve as antioxidants, and can be a source for vitamin A. Astaxanthin belongs to a group of oxygenated derivatives of carotenoids known as xanthophylls (Ferrari *et al.*, 2012).

Both α -crustacyanin and astaxanthin are commonly found in lobster in a complex. In the assembly of α -crustacyanin, two genetically distinct apocrustacyanins (Chayen *et al.*, 2000; Cianci *et al.*, 2002; Habash *et al.*, 2004) each bind an astaxanthin molecule and form a heterodimer (β -crustacyanin). The crystal structure of β -crustacyanin (A1A3 dimer) has been determined and revealed two astaxanthin molecules held in close proximity (Cianci *et al.*, 2002). Eight β -crustacyanin dimers assemble to form α -crustacyanin, a 320 kDa complex containing 16 astaxanthin molecules.

Astaxanthin has absorption spectrum peak at wavelength 472 nm but upon binding to crustacyanin it undergoes a large shift towards longer wavelengths (632 nm) giving a blue-coloured protein complex (Cianci *et al.*, 2002). The mechanism of the additional wavelength shift and the function of the protein-carotenoid complex are not fully understood. The change in colour of α -crustacyanin from blue to red upon protein denaturation and from red to blue upon astaxanthin-complex reconstitution has been a subject of investigation for over 60 years (Wald *et al.*, 1948; Zagalsky, 1985).

Biochemical, crystallography, spectroscopy, solution X-ray scattering and microscopy have been applied to study the molecular basis of the colouration in the lobster shell (Chayen, 2003). The crystal structure of β -crustacyanin has been solved at 3.2 Å (Cianci *et al.*, 2002). Low resolution structure of α -crustacyanin was solved at 30 Å (Rhys *et al.*, 2011) and the best diffraction to date was achieved by Chayen *et al.* (2003) at 10 Å.

1.6.3 RoAb13 ANTIBODY-PEPTIDE COMPLEX

The chemokine receptor CCR5 is an important receptor in leukocyte activation and mobilisation in the immune systems' inflammatory responses to inflammation. It has also been identified as a potential target for HIV blocking antibodies against the HIV virus which uses this co-receptor as an entry point into the cell (Barmania and Pepper, 2013), and CCR5 deficiency is strongly linked in protecting against the HIV virus infection (Chain *et al.*, 2015). Thus, CCR5 has become a very attractive target for anti-HIV therapy (Kondru *et al.*, 2008). This has led to the development of a number of small molecule CCR5 antagonists (such as Maraviroc and Aplaviroc) which to date require long-term use as an antiviral treatment for HIV and may lead to the emergence of resistant strains. A more permanent solution is required and various research groups (Wu *et al.*, 2006; Chackerian *et al.*, 2004; Chain *et al.*, 2008) are exploring the possibility of developing a HIV vaccine with the use of

antibodies to bind the CCR5 co-receptor as it is one of the main routes of the HIV virus into the cell.

In recent years various research groups have investigated the possibility of raising antibodies (by auto-vaccination) against the CCR5 chemokine co-receptor and have used recombinant proteins, recombinant viruses or synthetic peptides to test its feasibility (Wu *et al.*, 2006; Chackerian *et al.*, 2004; Chain *et al.*, 2008). Chain *et al.* (2008) have demonstrated that this immunisation approach is possible as long as the cellular autoimmune responses against the CCR5 co-receptor (which is found on dendritic cells, macrophages, T cells and any other immune cells expressing this receptor) can be avoided. The anti-CCR5 Fab fragment (RoAb13) has previously been shown to block some HIV infection and also blocks monocyte migration.

To this end, Chain and associates constructed a chimeric linear epitope in the extracellular amino terminal (N-Terminal) domain of CCR5 which is recognised by the existing RoAb13 antibody and has shown proof of principle that the synthetic epitope (the RoAb13 cognate) can induce antibodies and act as a potential antibody immunogen, which recognise the intact CCR5 co-receptor when it is linked with a tetanus toxoid T helper cell epitope (Fig 1.13) (Chain *et al.*, 2015). They found that the RoAb13 Fab fragment binds to both the CCR5 and the synthetic peptide with moderate to high affinity.

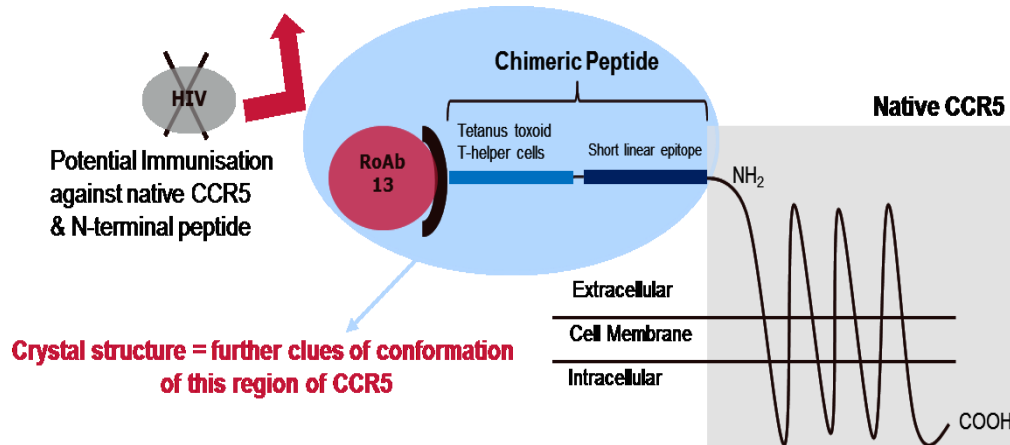


Figure 1. 13 Functional studies revealing RoAb13 Fab fragment antibodies bound to the native CCR5 receptor via a synthetic chimeric peptide.

Proof of principle: potential immunisation against the native CCR5 receptor using RoAb13 antibodies which binds to both the native CCR5 and the N-terminal peptide with the help of a synthetic epitope (the RoAb13 cognate) linked with a tetanus toxoid T helper cell epitope.

To understand and elucidate the interaction between the RoAb13 Fab fragment and the novel synthetic peptide we have been attempting to crystallise the RoAb13 antibody bound to the synthetic peptide. The 3-D structure of the RoAb13 antibody on its own has recently been determined to a reported 2.1Å resolution (Chain *et al.*, 2015).

2

WORKPLAN

2.1 PROJECT AIMS

The aim of this project focused on two main aspects of protein crystallisation which are interrelated:

- i. The first aspect was to reduce the crystallisation bottleneck by providing novel means of obtaining high quality crystals for X-ray analysis. This was done by designing and developing new and improved crystallisation methodologies and validating these using benchmark proteins.
- ii. The second was to apply these new techniques along with existing methods to crystallise target proteins which are of medical importance in order to facilitate their 3-D structures determination.
- iii. In addition, as automation and miniaturisation is critical in this post genomic era, efforts were made to automate any new methods where possible, with a view to adapting them to high-throughput.

Project hypothesis: The novel oil-on-drop vapour diffusion method will successfully crystallise and produce high quality diffracting crystals in medically important target proteins.

2.2 PROJECT APPROACH

2.2.1 PREVENTING EXCESSIVE NUCLEATION WITH THE OIL BARRIER METHOD

With reference to Chapter 1, subsection 1.5.2, I have devised a novel method to overcome the problems described in that section. Detailed results are presented in Chapter 4, subsection 4.1.

In terms of the approach taken, the work conducted focused on setting up manual crystallisation trials using this novel oil-on-drop method. After successful manual trials, the new method was tested, miniaturised and validated for automated use by two liquid handling systems for crystallisation which are widely used in academia and industry. They are the Douglas Instruments Oryx 8 and Labtech TTP Mosquito.

The method validation and automation was first conducted on two universally used benchmark proteins: lysozyme and trypsin, and thereafter on three target proteins gifted to us through key collaborations: alpha crustacyanin, an anti-CCR5 RoAb13-Peptide Complex and methyltransferase *Legionella pneumophila* - lpg2936. As the automated crystallisation liquid handling systems (Douglas Instruments Oryx 8 and the TTP Labtech Mosquito LCP) are able to perform both hanging drop and sitting drop experiments, the sitting drop technique was favoured because it is easier to set up on automated systems and quicker during the runtime of such trials. User intervention during a trial run is also not required as is with the hanging drop automated trials (refer to Fig 1.5 (iii)). Further to this, a comparison was only possible between oil-on-drops versus control drops (with no oil) in the automated sitting position, and thus a comparison with the conventional Chayen method (oil-over-reservoir) as performed during manual trials, was not possible.

The results in Chapter 4 are categorised by protein starting with the benchmark proteins then the target proteins. The manual and automated trials are included under each protein which also includes a comparison of crystal counts (number of crystals observed) in both automated systems. Lastly, comparisons were made on the mean crystal sizes (length and width)

between the Chayen method and the oil-on-drop method (subsection 4.1.2) and the differences in mean crystal sizes on the Oryx 8 and Mosquito (subsection 4.1.3) with regards to the new oil-on-drop method.

2.2.2 NUCLEANT EXPERIMENTS

As part of the repertoire of optimisation tools and techniques available in our laboratory, a range of optimisation experiments were conducted with a selection of available nucleants. These nucleants were tested on benchmark proteins such as thaumatin, lysozyme, catalase and trypsin and then on target proteins alpha crustacyanin, the RoAb13-Peptide Complex, and methyltransferase *Legionella pneumophila* - lpg2936. The nucleants used included CPGs (normal and functionalised), CB Ink, Bioglass, NPG, PEG GO and MIPs. Information of all the nucleants used is detailed in Chapter 3.

2.2.3 WORKING PHASE DIAGRAMS

In addition to the use of the aforementioned optimisation strategies, the inception of any crystallisation trial after screening for crystallisation “hits” requires the use of working phase diagrams for the proteins used in such trials.

Such diagrams help to determine the best and most optimal crystallisation conditions for crystal nucleation, such as lowering protein and precipitant concentrations. It also helps with the identification of the various zones (i.e. undersaturation, supersaturation, and the metastable zone, which is also important for nucleant crystallisation trials). This was used to great effect for all the nucleants reported in this thesis. The nucleant trial results are covered in Chapter 4 subsection 4.2.

2.2.4 OTHER OPTIMISATION STRATEGIES

During a critical period whilst working on target protein - RoAb13-Peptide complex and in search of producing diffraction quality crystals, there was a need to vary the period of incubation and complex bonding between the antibody (RoAb13) and its synthetic peptide before setting up any

crystallisation trials. Crystals of the RoAb13-Peptide complex were obtained at 10mg/ml and also at 8mg/ml during first phase optimisations. To save on the limited supply of this antibody, incubation experiments were conducted using the lower concentration of 8mg/ml. Incubation periods of 12, 24, 72, 96, 120, 144 hours of the complex were setup before crystallisation. Detailed results are presented in Chapter 4, subsection 4.3.2.

Further optimisation on the RoAb13-Peptide complex included applying the new oil-on-drop method to aid the production of fewer and larger crystals for X-ray diffraction studies (results are noted in subsection 4.3.3).

2.2.5 X-RAY DIFFRACTION ANALYSIS

Where suitable protein crystals were grown, they were harvested and shipped to the Diamond Light Source (DLS) synchrotron for X-ray diffraction studies (refer to Chapter 3, subsection 3.8 for method information). The true measure of whether a crystal is of high quality is when it is irradiated with X-rays and produces a diffraction pattern at a high enough resolution, in addition with other good data quality indicators, to allow for electron density maps to be produced towards 3-D structure determination. Crystals from target proteins alpha crustacyanin, methyltransferase *Legionella pneumophila* - lpg2936 and the RoAb13-Peptide complex were irradiated with X-rays as well as the benchmark protein trypsin – all used to validate the novel oil-on-drop method.

2.2.6 STATISTICAL ANALYSIS

The data and information from the crystallisation trials were collated and presented in Chapter 4. This includes qualitative, semi-qualitative and quantitative information gathered from all crystallisation experiments. The data was then analysed statistically to infer statistical significance (or not) between the Chayen and oil-on-drop methods. This is presented in Chapter 4 and discussed in Chapter 5.

3

MATERIALS & METHODS

3.1 PROTEINS AND BUFFERS

The benchmark proteins (i.e. proteins that have previously been crystallised and their 3-D structures solved and are used to test and validate crystallisation conditions and methods):

- Thaumatin (T-7638-25MG) from *Thaumatococcus daniellii* was prepared in de-ionised H₂O at concentrations of 10, 20, 25 and 30mg/ml.
- Lysozyme (L-6876) from hen egg was prepared in 0.1 M sodium acetate (NaOAc) pH 4.8 with concentrations ranging between 15-60mg/ml.
- Trypsin (T-9201) from bovine pancreas was prepared in 0.1M TRIS pH 7.4 with concentrations ranging between 30-60mg/ml.
- All the proteins purchased from Sigma-Aldrich, UK were analytical grade and highly pure lyophilised powders and kept refrigerated at -20°C to the manufacturer's instructions.
- Water used from Nanopure™ (Thermo Fisher Scientific, UK).

The target proteins:

- Anti-CCR5 antibody (RoAb13) in 0.1M NaCl and 20mM HEPES (pH 7.0) at 8-10mg/ml and its cognate antibody ligand - a chimeric linear N-Terminal domain epitope of CCR5 (peptide sequence: MDYQVSSPIYDINYTTSEPCQKINVKQIAA with a molecular weight of 3482.9 Da) at 100mM was received from Professor Benjamin Chain (UCL).
- Alpha crustacyanin with concentrations ranging between 6.5-10mg/ml in 0.1M Tris (pH 6.9) was received from Dr Peter Zagalsky (Royal Holloway, University of London).

- Methyltransferase (*Legionella pneumophila* - Lpg2936) in 25mM TRIS pH 8; 150mM NaCl; 1mM EDTA; 1mM DTT; 5% glycerol at concentrations ranging between 15-23.5mg/ml was received from Dr Nikos Pinotsis (Birkbeck College, University of London).
- The target proteins as received from collaborators were confirmed as highly purified batches which were then aliquoted into 100ul portions on receipt and stored at -80°C. Once thawed aliquots were stored at 4°C for use.

3.2 CRYSTALLISATION CONDITIONS FOR THE OIL BARRIER METHODS AND THE NUCLEANT TRIALS

Crystallisation solutions were freshly prepared and kept at room temperature unless indicated to be stored at 4°C. For both the Chayen and the oil-on-drop methods, the crystallisation conditions were the same.

Proteins

The RoAb13-Peptide complex was crystallised at 8 and 10mg/ml with ammonium sulphate at 1.9-2M. Existing crystallisation conditions are based on previous optimisations done on the complex.

Alpha crustacyanin was crystallised at 6.5-10mg/ml with 20-30% of precipitant stock consisting of 0.1M MES; 0.2M ammonium sulphate and 30% PEG 5000. Existing crystallisation conditions are based on previous optimisations done on the protein.

Legionella pneumophila - Lpg2936 was crystallised at 17-23.5mg/ml with 15-30% of condition 38 of the Morpheus screen - MD1-46 (Molecular Dimensions, UK) consisting of 0.1M MES/Imidazole pH6.5; 27% Ethylene glycol-PEG 8000 and 0.14M alcohols - for both Chayen oil barrier and oil-on-drop experiments as well as the nucleant trials. Optimal crystallisation conditions were received from the collaborator.

Thaumatococcus was crystallised at 10-30mg/ml with potassium sodium tartrate (NaKT) (pH 6.8) as the precipitating agent at concentrations of 0.2M-0.6M

NaKT.

Lysozyme was crystallised at 15-60mg/ml with 0.1 M NaOAc pH 4.8 and NaCl at concentrations of 0.2-0.8M.

Trypsin was crystallised at 30-60mg/ml with 0.1M TRIS pH 7.4 and 10-14% PEG 8000.

The crystallisation conditions for the benchmark proteins thaumatin, lysozyme and trypsin are widely known within the crystallisation and crystallography community and the conditions listed are optimal.

3.3 CRYSTALLISATION PLATES AND DISPENSING SYSTEMS

All manual Hanging Drop Vapour Diffusion (HDVD) crystallisation trials consisted of 0.75 μ l protein solution mixed with 0.75 μ l reservoir precipitant in a 1:1 ratio on X-seal screw caps inverted and sealed over a 300 μ l precipitant reservoir solution in 15-well EasyXtal plates (Qiagen, UK) (Fig 3.1).



Figure 3. 1 15 well EasyXtal plate with X-seal screw cap (Qiagen, UK)

Automated high-throughput sitting drop vapour diffusion (SDVD) trials were performed with two drop 96-well plates containing 80 μ l reservoir solution (Figure 3.2 (a)) and sealed with clear sealing tape (HD Clear™ Sealing Tape, Hampton Research). Experiments were performed on two automated platforms; the Douglas Instruments Oryx 8 (Figure 3.2 (b)) and the TTP Labtech Mosquito LCP (Figure 3.2 (c)) crystallisation dispensing systems.

For both systems, duplicate drops were setup with pre-programmed drop sizes consisting of 15-30nl protein solution mixed with 15-30nl from the 80 μ l reservoir solution. Oil was dispensed directly onto the drops in quantities of

between 100-200nl to cover the drops. Control drops without oil were set up in tandem. Drops were dispensed in paired groups, i.e. one control drop and one oil-on-drop per well sharing a single reservoir (Fig 3.2 a). The Chayen oil barrier method was not employed in the automated trials. Only 80/20 paraffin to silicone oil ratios were used for all automated trials. This was the set up for each replicate.

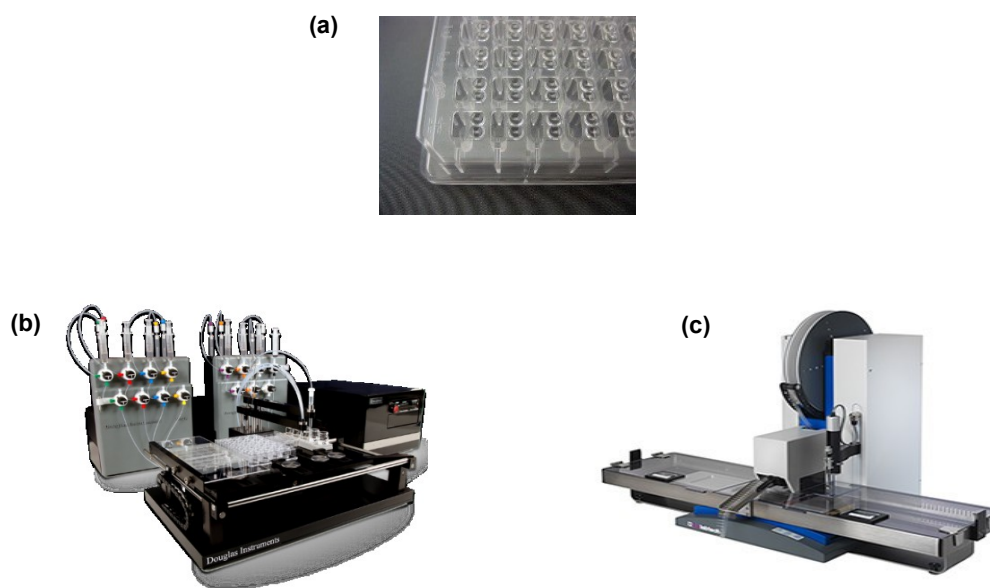


Figure 3. 2 Automated systems for crystallisation trials
(a) 96 well 2 drop MRC Swissci UVXPO plates (Douglas Instruments, UK)
(b) Oryx 8 system (Douglas Instruments, UK)
(c) Mosquito ® LCP system (TTP Labtech, UK)

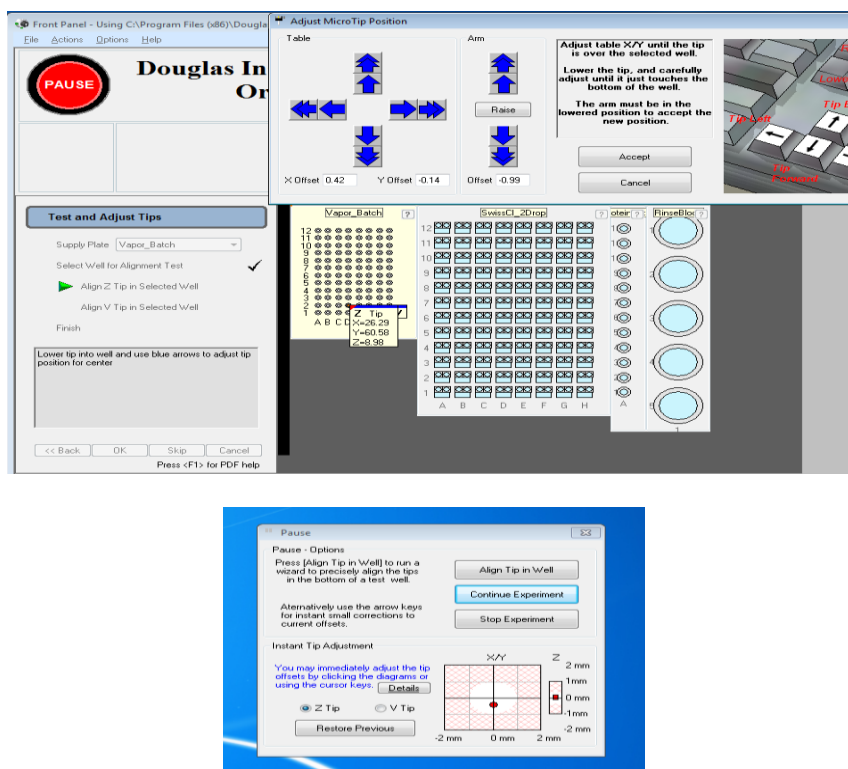


Figure 3.3 Example screenshots of the Oryx 8 Front Panel software settings allowing for microtip tip height and directional alignment

Z Tip allows for the protein and precipitant dispensing and the V Tip for the oil dispensing. Alignment settings are available for both Z and V Tips in relation to the crystallisation plate and well positions. There are system prompts for checking alignment before an experiment is executed & options to pause and continue runs after re-alignment.

3.4 OILS

Paraffin and silicone oil used in the oil barrier experiments were sourced from BDH Laboratory Supplies (294365H, Poole, UK). Al's oil, which is a mixture of 50% paraffin and 50% silicone, and other paraffin and silicone oil mixture ratios (paraffin/silicone: 60/40%; 70/30%; 80/20%; 90/10%) were made up from existing stock.

3.5 OIL BARRIER EXPERIMENTS

Al's oil and various paraffin and silicone oil mixtures were used for the Chayen barrier method. A layer of 200-250 μ l was dispensed on top of the precipitant reservoirs. For the oil-on-drop experiments, oil was dispensed directly onto the drops in quantities of between 0.2-0.5 μ l. The whole surface area of the protein-precipitant drop was covered by the oil mixture. Control drops were set up for both methods. The buffers, protein and precipitant concentrations are listed under subsections 3.1 Proteins and Buffers and 3.2 Crystallisation Conditions.

3.6 NUCLEANT TRIALS

All the nucleants: controlled pore glass (CPG) 8nm and 48nm (Figs 3.4 -3.6) (Ms. Clara Anduix Canto - Leeds University); haemoglobin-MIP (Dr Reddy - Surrey University); nanoporous gold (Kertis - Johns Hopkins University); PEGylated graphene oxide and carbon black ink (Dr. Hannah Leese - Imperial College London); and bioglass powder (Naomi's nucleant - Molecular Dimensions, UK) were manually inserted into crystallisation trials in hanging drops in EasyXtal tools™ (Qiagen, UK).

The liquid nucleant solutions were kept at 4°C and dispensed in quantities of 0.1-0.2 μ l into crystallisation trial drops using a standard micropipette. The solid nucleants were inserted into drops using fine-point forceps. The screw caps were then inverted and sealed over the reservoir solution and incubated at 20°C. Controls were set up for all respective nucleants trials.

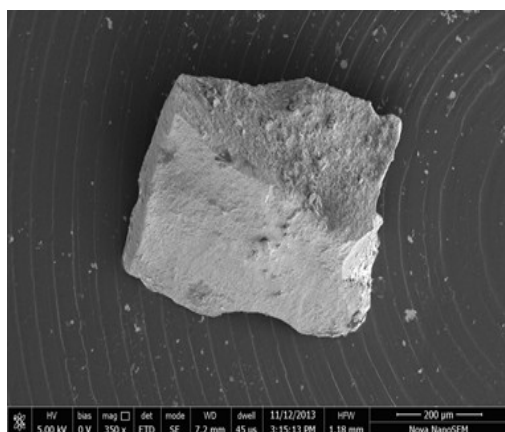


Figure 3. 4 Scanning electron microscopic image of an individual CPG grain

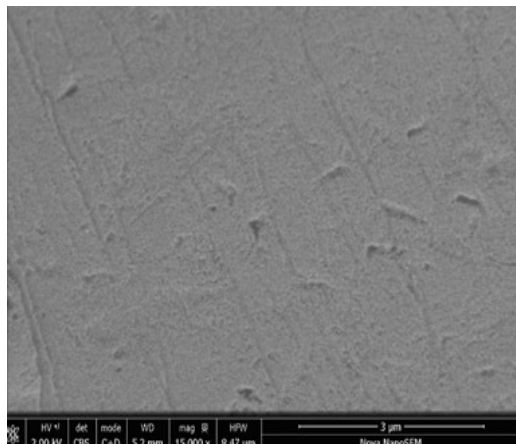


Figure 3. 5 Scanning electron microscopic image of a CPG surface (8nm pore size)

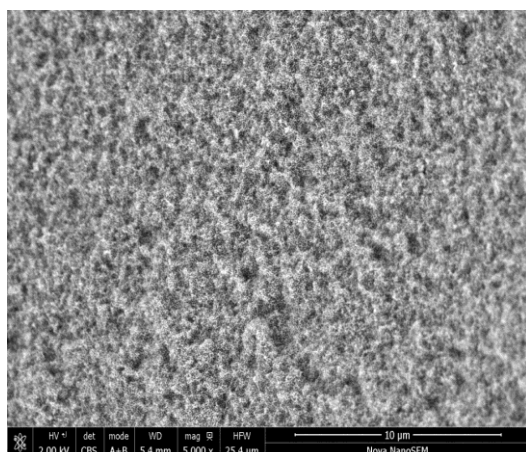


Figure 3. 6 Scanning electron microscopic image of a CPG surface (48nm pore size)

Carbon black ink (Table 3.3) was diluted with deionised water to a 1/80 working solution and then added to the drops.

Table 3. 3 Composition of carbon black ink

Ingredient name	% wt
Water	20 - 60
Carbon black	2.0 - 10.0
Ammonia salt of modified styrene acrylic polymers	3.0 - 30.0
Dipropylene glycol monomethyl ether	10 - 30
Copolymer with pigment affinic groups	0.3 - 6.0

Working phase diagrams were constructed for all proteins used for nucleant trials in order to find metastable conditions by varying protein and precipitant concentrations.

3.7 CHARACTERISATION OF DROPS

Observations and analysis of the trial drops was performed using a Leica light microscope (Model M165-C, Leica Microsystems, Germany) immediately after trial set-up, then at 24h, thereafter daily for the first two weeks, and subsequently at weekly intervals up to several months. Images of the drops were taken at various stages.

3.8 X-RAY DIFFRACTION

Crystals were harvested (i.e. scooped) from the protein drops with the aid of loops of varying sizes depending on the size of the crystals and frozen in liquid nitrogen in readiness for X-ray diffraction. Crystals were X-rayed at the Diamond Light Source (DLS) synchrotron in Oxfordshire via remote sessions. Frozen crystals were couriered in a travel dewar to the synchrotron a few days before the session to allow for the crystals to be logged and processed for the remote session. The i04 macromolecular crystallography beamline was used and has a typical working wavelength of 0.9795 Å (12.658 keV) but is tunable over the wavelength range 0.886 - 2.066 Å. The beam size can be focussed from 10 x 5 to 100 x 100 microns across the available energy range. The i24 MX micro-focus beamline was used for protein crystals too small for other beamlines. The i24 beamline has a typical working wavelength of 0.7 - 2.0 Å (keV - 6.4 - 20.0) and flux of 3.0×10^{12} ph/s - into full beam of 8x6µm @ 12800 eV and 300 mA ring current. The minimum focused beam size is 7 x 6 µm up to a maximum of 50 x 50 µm.

The mounted crystals were X-rayed at 0.1 oscillations with 0.1-0.2 second exposure time at beam strength of 50-100% and data collected by rotating the crystals 360° whilst being X-rayed. All experiments and data acquisition were run through the DLS GDA (Generic Data Acquisition) graphical user

interface software and analysed after data transfer from the Diamond servers onto local laboratory computer systems.

3.9 STATISTICAL ANALYSIS

In the manual HDVD trials, the sample size for each experimental trial were set at a minimum of 10 replicates ($n=10$) for each oil ratio tested i.e. $n=10$ for 50/50; 60/40; 70/30, etc. and across each method i.e. controls; Chayen method and oil-on-drop method.

Sample sizes for the automated trials were set at a minimum of 14 replicates ($n \geq 14$) per method i.e. comparing only controls to the oil-on-drop method.

For statistical analysis, protein drops containing visible crystals were counted at set intervals to collect data on the number and sizes of the crystals. Crystal counts were conducted manually by identifying a square area on the images using Fiji's ImageJ software (Schindelin, *et al.*, 2012), counting the number of crystals within that area. An estimate of the total number of crystals within the whole drop was calculated by extrapolating the total number of squares drawn per drop multiplied by the number of crystals counted (no. of squares drawn X crystals counted per square = total crystals per drop). Where micro-crystals and crystal showers were observed, ImageJ's automated counting settings were used.

The data was then analysed statistically to infer statistical significance to determine differences crystal count and crystal size (length and width) between control drops, the Chayen oil barrier method and the oil-on-drop method.

Statistical significance was tested by ANOVA (analysis of variance) in the case of comparing the three groups (controls; Chayen method and oil-on-drop method). The ANOVA test is a statistical test that can be used to compare the means of more than two populations and can be used to compare the mean crystal count between and within replicates across the

different methods and also across different oil ratios. Another test, the student's t-test was used to compare the means of two populations only, the control versus the oil-on-drop method in the automated trials to compare the mean crystal count between the two groups. Both tests were performed using the data analysis ToolPak package in Microsoft Excel version 2010. Statistical significance was reported in the ANOVA test by an F-test value and p-value and in the t-test by a p-value. The p-value for all the statistical analysis testing was set at a threshold alpha (α) value of 5% and an F-value that's greater than the F-crit value (a critical value). Thus, any ANOVA test results lower or equal to 5% ($\alpha \leq 0.05$) and F-values greater than the F-crit value indicates a statistical significant result and any t-test value lower or equal to 5% indicated a statistical significant result.

4

RESULTS

4.1 PREVENTING EXCESSIVE CRYSTAL FORMATION: A NOVEL OIL BARRIER METHOD

In order to overcome the limitations of the current Chayen oil barrier method (Fig 1.6), I conceived and designed a new oil-on-drop oil barrier method as depicted below (Fig 4.1). This new method overcomes the limitations experienced with the Chayen oil barrier method as follows:

- (i) Covering each drop with oil accentuates the localised system of controlled evaporation between each drop and the equilibrating precipitant reservoir
- (ii) In manual experiments, it allows for multiple drops to be set up per coverslip/screwcap with different oil ratios per drop which creates flexibility during trial set up and also eliminates consumable wastage
- (iii) And, unlike the traditional Chayen method it works with all PEGs and solvents such as MPD at all oil ratios

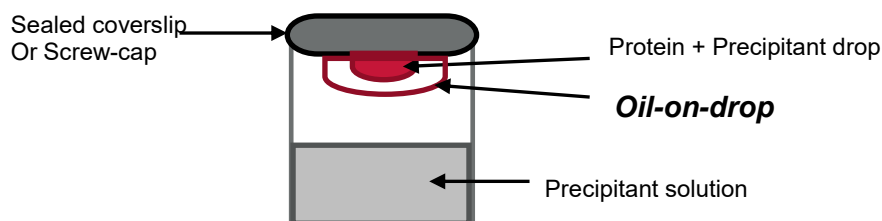


Figure 4. 1 Schematic representation of the HDVD technique with the new oil-on-drop oil barrier

The oil-on-drop method has been tested on two benchmark and two target proteins using the crystallisation conditions listed in Chapter 3 - Materials and Methods. Firstly, all trials were initially set up by hand using the hanging drop vapour diffusion method. Once successfully validated, the new method was trialled for automation and miniaturisation on the same proteins used for manual trials. The results follow below.

4.1.1 MANUAL AND AUTOMATED TRIALS ON BENCHMARK AND TARGET PROTEINS

In the manual HDVD trials, the sample size for each experimental trial were set at a minimum of 10 replicates ($n=10$) for each oil ratio tested i.e. $n \geq 10$ for 50/50; $n \geq 10$ for 60/40; $n \geq 10$ for 70/30, etc. (illustrated in Fig 4.2 - 4.5) and across each method i.e. controls; Chayen method and the oil-on-drop method.

Sample sizes for the automated trials were set at a minimum of 14 replicates ($n \geq 14$) per method i.e. comparing only controls to the oil-on-drop method. For all the automated trials, the Chayen method was not tested. The p-value for all the statistical analysis testing (ANOVA and 2-tail T-test) was set at a threshold alpha (α) value of 5% and in addition an F and F-crit (critical) value for ANOVA. Thus, any ANOVA test results lower or equal to 5% ($\alpha \leq 0.05$) and F-values greater than the F-crit value indicated a statistical significant result and any t-test results lower or equal to $\alpha \leq 0.05$ indicates a statistical significant result.

4.1.1.1 LYSOZYME

4.1.1.1.1 MANUAL TRIALS

Crystals formed 24-36 hours after setup in control drops; after 48 hours with the Chayen method; and ~ 72 hours with the oil-on-drop method. Fig 4.2 depicts the differences in both the crystal count (number of single crystals counted) and individual crystal sizes between (a) control; (b) the Chayen method; and (c) the oil-on-drop method. The oil-on-drop method (Fig 4.2 c) produced oil run-off where "excess" oil created a visible ring

around the drop, whilst still covering the drop. This was seen in some replicates but not all and the 80/20 ratio produced the least amount of run-off compared to the other ratios trialled. For a final crystal size comparison for each method after four months, typical mean sizes were $\sim 100 \times 80 \times 50 \mu\text{m}$ in the control drops; $\sim 120 \times 100 \times 80 \mu\text{m}$ in the Chayen method; and $\sim 200 \times 150 \times 100 \mu\text{m}$ in the oil-on-drop method.

Each oil ratio tested had a sample size of 10 replicates ($n=10$) per method. When looking at the number of individual crystals of lysozyme across a range of different oil ratio mixtures of paraffin and silicone, there was a statistically significant variation ($F > F_{\text{crit}} (11.35 > 3.89)$; $p = 0.002$) between the control drops and also between the two oil barrier methods.

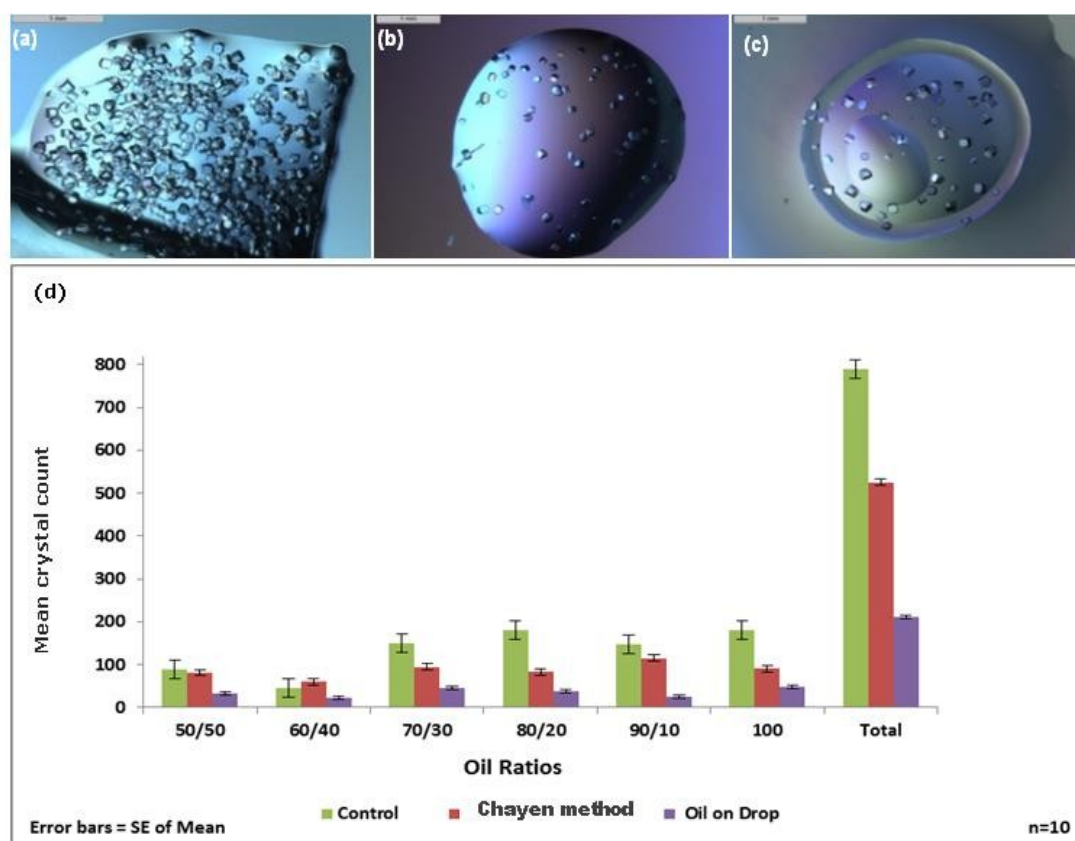


Figure 4. 2 Lysozyme crystal formation across methods and oil ratios
 (a) Control drop; (b) Chayen method; (c) Oil-on-drop method; and (d) Graph showing statistically significant variation ($F > F_{\text{crit}} (11.35 > 3.89)$ $p = 0.002 < \alpha = 0.05$) in crystals across the methods and oil ratios. Each oil ratio tested had a sample size (n) ≥ 10 per method.

When looking at the number of crystals of lysozyme across a range of paraffin to silicone oil ratio mixtures (50/50; 60/40; 70/30; 80/20; 90/10 and 100% paraffin oil), Figure 4.2 (d) shows the differences in the crystal count between the controls; the Chayen method; and the oil-on-drop method. In total the control drops produced the most crystals (~ 790), followed by the Chayen method (~ 525), and the new oil-on-drop method gave rise to the least number of crystals (~ 212).

4.1.1.1.2 AUTOMATED TRIALS

4.1.1.1.2.1 DOUGLAS INSTRUMENTS ORYX 8

Crystals formed 24-36 hours after setup in control drops and ~ 36 hours with the oil-on-drop method.

Final crystal size comparison was conducted for each method after 4 weeks and on average, the smallest crystal sizes were observed in the control drops ($\sim 180 \times 150 \times 100 \mu\text{m}$); and larger crystal sizes observed in the oil-on-drop method ($\sim 350 \times 300 \times 250 \mu\text{m}$).

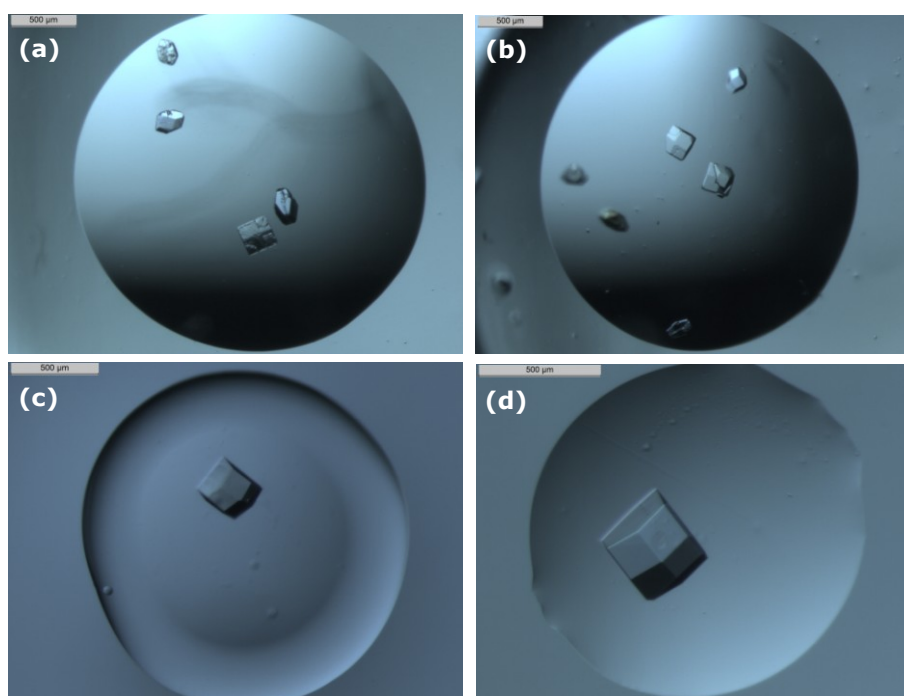


Figure 4. 3 Lysozyme crystal images of control drops vs oil-on-drops
(a & b) Control drops; (c & d) Oil-on-drop method

Fig 4.4 shows the differences in crystal count between the control drops compared to the oil-on-drop method for 16 replicates (n=16). When looking at the number of crystals of lysozyme there was a statistical significant variation (2-tail T-test unequal variances: $p = 0.000 < \alpha=0.05$) between the control drops and the oil-on-drop oil barrier method with the control drops producing a total of 70 crystals and the oil-on-drop method just above 20 crystals.

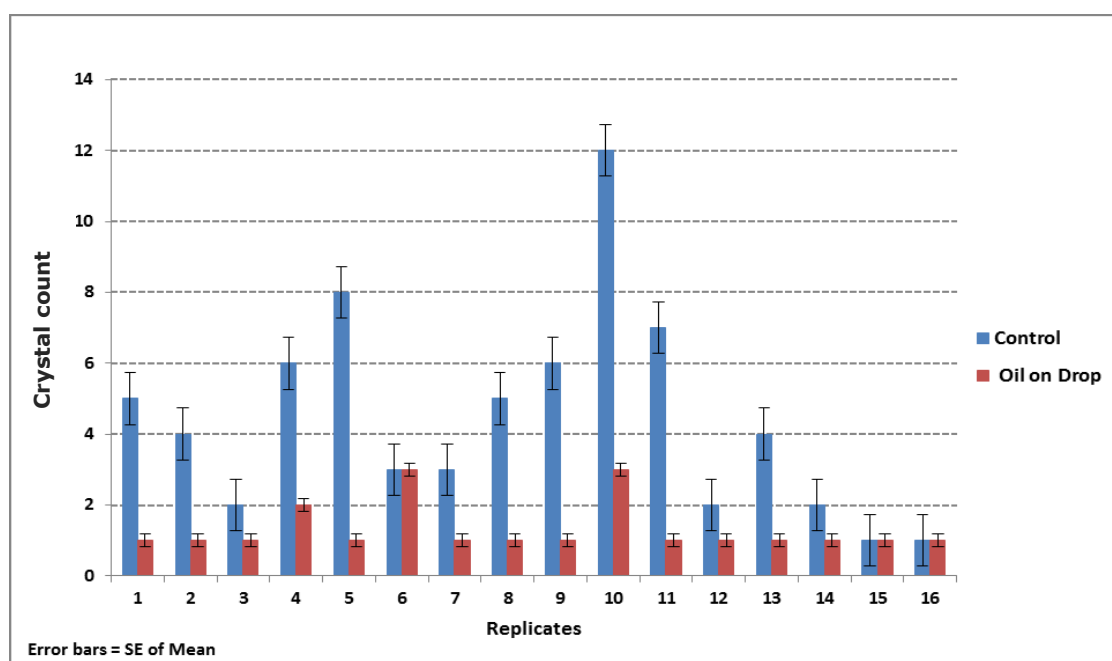


Figure 4. 4 Lysozyme crystal count: Controls vs Oil-on-drops set up by the Oryx 8

Statistical significance (2-tail T-test unequal variances: $p = 0.000 < \alpha=0.05$) in number of crystals between the controls and oil-on-drops.

4.1.1.1.2.2 TTP LABTECH MOSQUITO LCP

Crystals formed 24-36 hours after setup in control drops and ~36 hours with the oil-on-drop method (Fig 4.5 a-b). Comparison of crystal size was conducted for each method after 4 weeks, and on average, the smallest crystal sizes were observed in the control drops (~150x80x60 μm); and larger crystal sizes observed in the oil-on-drop method (~220x100x80 μm).

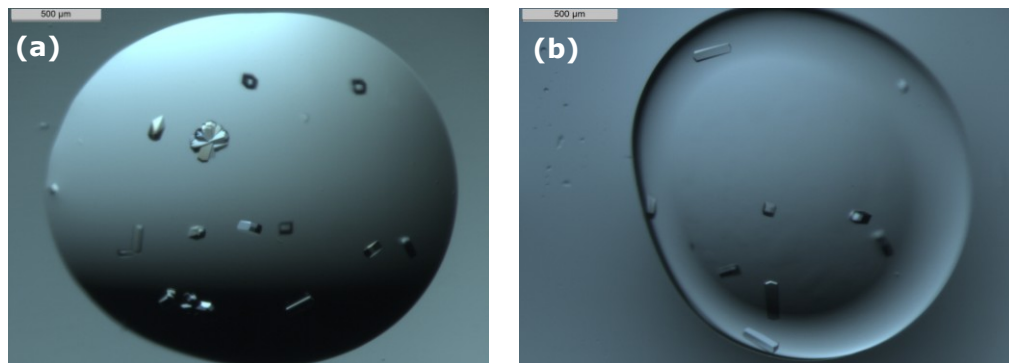


Figure 4. 5 Lysozyme crystal images of control drop vs oil-on-drop
(a) Control drop; (b) Oil-on-drop method

Fig 4.6 shows the differences in crystal count between the control drops and the oil-on-drop method with 14 replicates ($n=14$). When looking at the number of individual lysozyme crystals there was no statistical significant variation (2-tail T-test unequal variances: $p = 0.396 > \alpha=0.05$) between the control drops and the oil-on-drop oil barrier method with the control drops producing a total of 85 crystals and the oil-on-drop method 56 crystals.

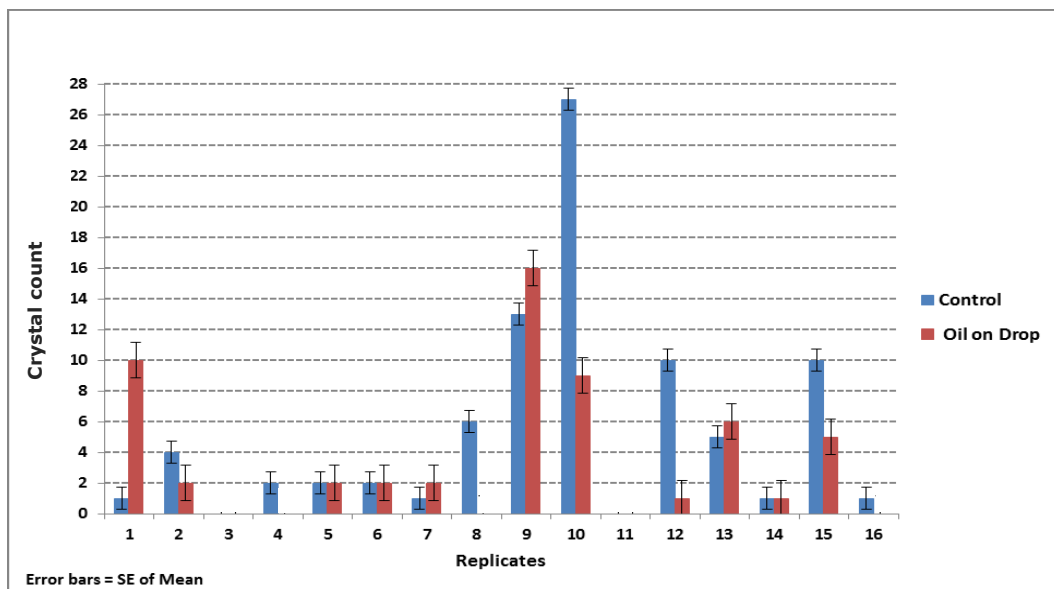


Figure 4. 6 Lysozyme crystal count: Controls vs Oil-on-drop set up by the Mosquito robot.

No statistical significance (2-tail T-test unequal variances: $p = 0.396 > \alpha=0.05$; $n=14$) in number of crystals between the control drops and the oil-on-drop method.

Fig 4.7 illustrates the differences in lysozyme crystal count when using the Oryx 8 and the Mosquito liquid handling systems commonly used in automated crystallisation trials.

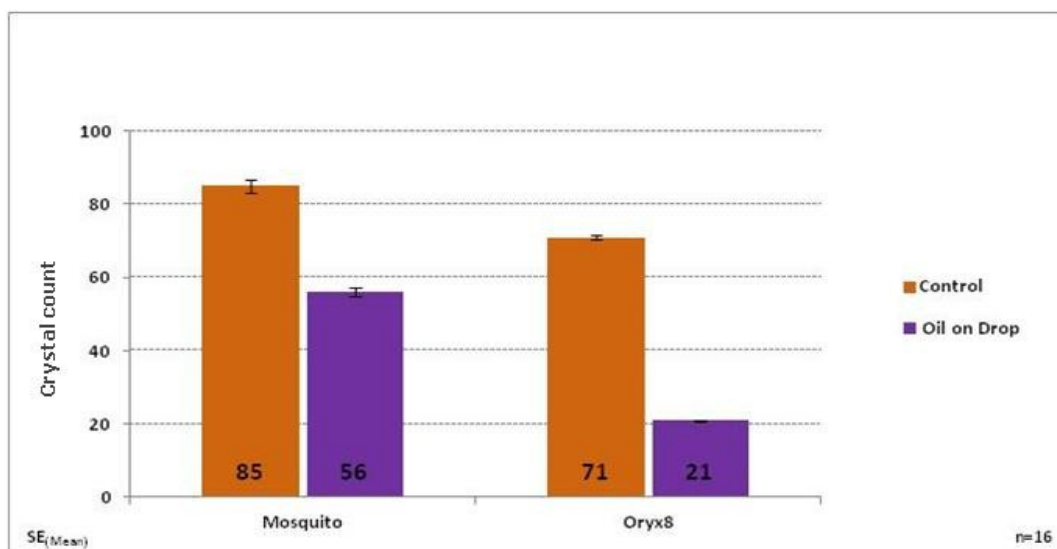


Figure 4. 7 Graph depicting lysozyme crystals using the Oryx 8 and Mosquito: Controls vs Oil-on-drop. Mean sample size (n) = 16.

Results from the controls and the oil-on-drop method trials using lysozyme across 16 replicates show a difference in the number of crystals formed across both the control drops (71 crystals with the Oryx 8 system vs. 85 crystals with the Mosquito) and in the oil-on-drop method (21 crystals with the Oryx 8 and 56 crystals with the Mosquito). The Oryx 8 had a bigger gap in crystals formed between the controls and the oil-on-drop method (a difference of 50) compared to the Mosquito (difference of 29). Also, the mosquito produced more crystals – both in controls and the oil-on-drop method.

4.1.1.2 TRYPSIN

4.1.1.2.1 MANUAL TRIALS

Crystals were seen after 3 days in control drops, no crystals were formed with the Chayen method, and crystals appeared after 5 days with the oil-on-drop method. The control drop gave rise to 8 crystals (Fig 4.8 (a)); the

Chayen method (Fig 4.8 (b)) dried out within 24-36 hours giving no crystals; and the oil-on-drop method gave 2 crystals (Fig 4.8 (c)).

To compare the crystal sizes obtained by each method, crystals were measured at week 12. The measurements at week 12 show that the control drops produced crystals with an average size of $\sim 400 \times 125 \times 80 \mu\text{m}$, the Chayen method produced no crystals, and the largest crystal sizes of $\sim 600 \times 200 \times 100 \mu\text{m}$ were observed in the oil-on-drop method. The images in Fig 4.8 (a-c) were taken at week 4.

Each oil ratio tested had a sample size of 10 replicates ($n=10$) per method. The student's T-test was performed on the two populations (controls versus oil-on-drop). The test results shows that there was no significant statistical variation in crystal count (2-tail T-test unequal variances: $p = 0.09 > \alpha=0.05$) between the two groups (Fig 4.8 d) and thus the null hypothesis of no difference must be accepted.

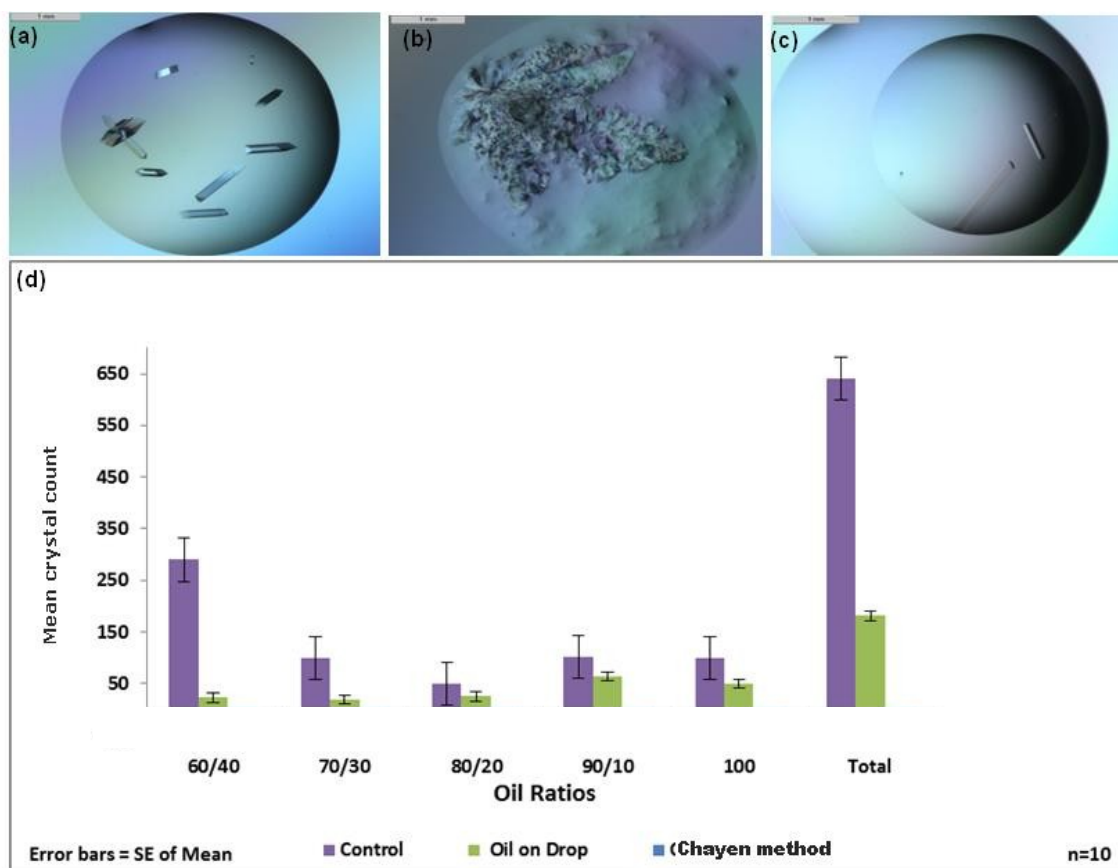


Figure 4. 8 Trypsin crystals across methods and oil ratios

(a) Control drop; (b) Chayen method; (c) Oil-on-drop method; (d) No statistical significance (2-tail T-test unequal variances: $p = 0.09 > \alpha = 0.05$) in the number of crystal across the methods and oil ratios. Each oil ratio tested had a sample size (n) ≥ 10 per method.

Looking at the oil-on-drop method in more detail in Fig 4.8 d, the 60/40 oil ratio showed the largest variation between the control and oil-on-drop method in terms of number of formed crystals (290 for controls versus 23 for oil-on-drop). The 80/20 oil ratio contributed to the smallest difference in formed crystals - 49 for the controls and 25 for the oil-on-drop. The total average crystal count across all oil ratios were 640 crystals (total in the graph) for the controls, zero for the Chayen method, and 181 for the oil-on-drop method. The oil-on-drop method produced a large amount of oil run-off clearly visible as a ring around the drops (Fig 4.8 c). The drops were still covered. This was seen in more than half of the replicates. The 80/20 ratio produced the least amount of run-off compared to the other ratios trialled.

4.1.1.2.2 AUTOMATED TRIALS

4.1.1.2.2.1 DOUGLAS INSTRUMENTS ORYX 8

Crystals formed 72 hours after setup in control drops and ~ 96 hours with the oil-on-drop method. Final size comparison was conducted for each method after 4 weeks incubation and on average, the smallest crystal sizes were observed in the control drops ($\sim 150 \times 60 \times 40 \mu\text{m}$); and larger crystal sizes observed in the oil-on-drop method ($\sim 280 \times 80 \times 50 \mu\text{m}$).

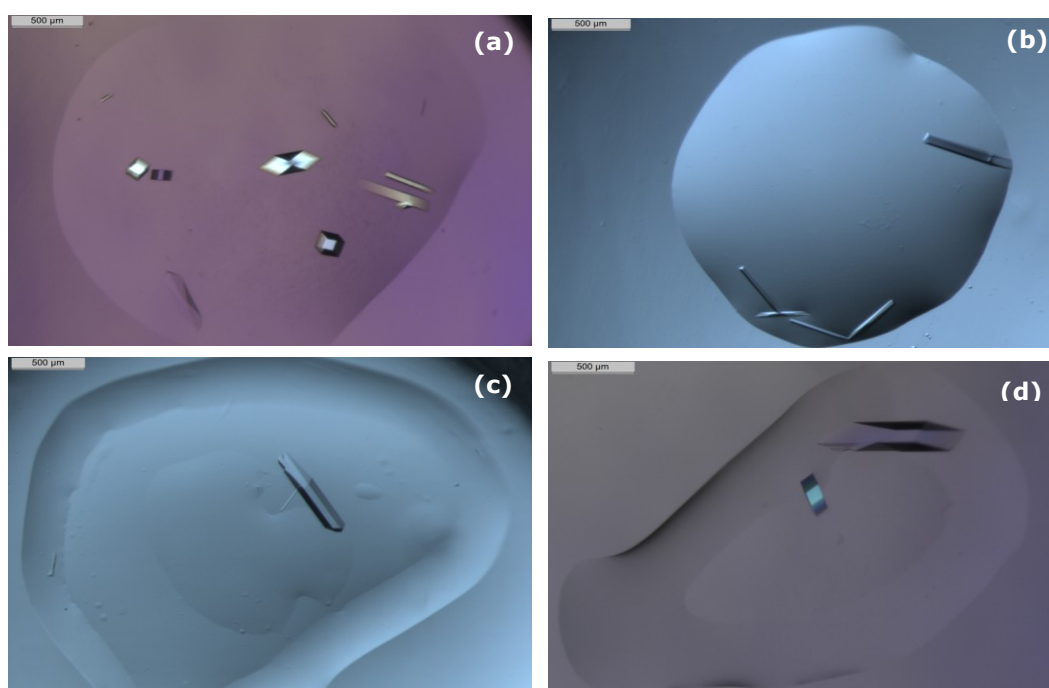


Figure 4. 9 Trypsin crystal images of control drops vs oil-on-drop method
(a & b) Control drops; (c & d) Oil-on-drop method

Fig 4.10, shows the differences in crystals counted between the control drops compared to the oil-on-drop method with 17 replicates ($n=17$) for trypsin. When looking at the number of individual trypsin crystals, there was a highly statistical significant variation ($p = 0.000 < \alpha=0.05$) between the control drops and the oil-on-drop method with the control drops producing a total of 156 crystals and the oil-on-drop method 48 crystals. The oil-on-drop method produced a large amount of oil run-off clearly visible as rings around the drops (Fig 4.9 c - d). The drops were still covered. This was seen in more than half of the replicates even in the 80/20 ratio using the Oryx 8.

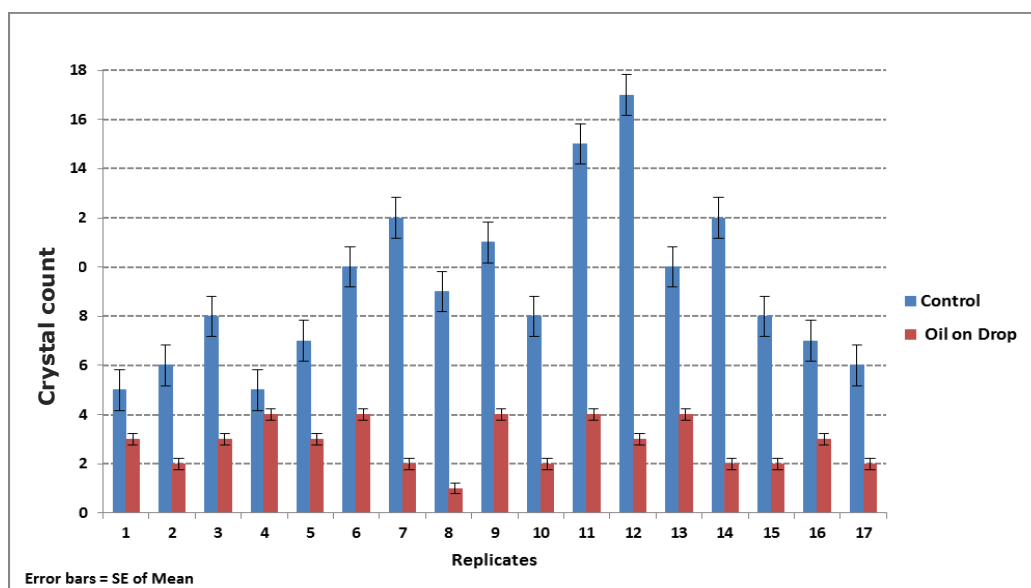


Figure 4. 10 Trypsin crystal count: Controls vs Oil-on-drop using the Oryx 8
 Statistical significance (2-tail T-test unequal variances: $p = 0.000 < \alpha = 0.05$; $n = 17$)
 in crystal number between controls and the oil-on-drop method.

4.1.1.2.2.2 TTP LABTECH MOSQUITO LCP

The crystals formed 72 hours after setup in control drops and ~96 hours with the oil-on-drop method. Crystal sizes were measured after 3 weeks of incubation from trial setup (images in Figure 4.11 a - b). A final size comparison was conducted for each method after 4 weeks, and on average, the control drops produced the smallest crystal sizes (~300x160x80 μm); and larger crystal sizes were observed in the oil-on-drop method (~500x210x100 μm). A reduced level of oil run-off was experienced with the oil-on-drop method using the Mosquito (Fig 4.11 b).

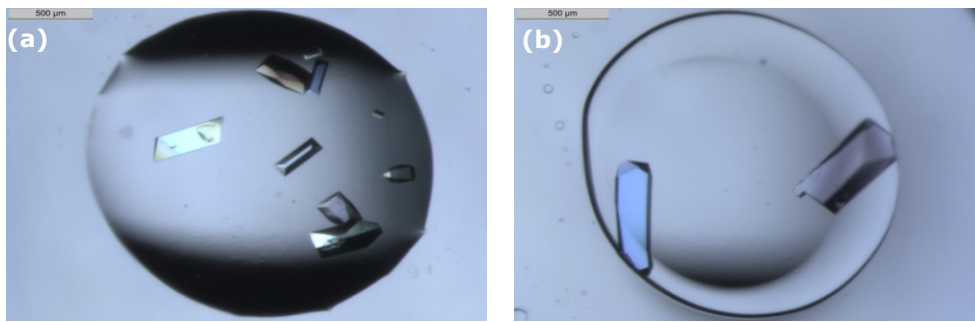


Figure 4. 11 Trypsin crystal images of control drops vs oil-on-drop method
(a) Control drop; (b) Oil-on-drop method

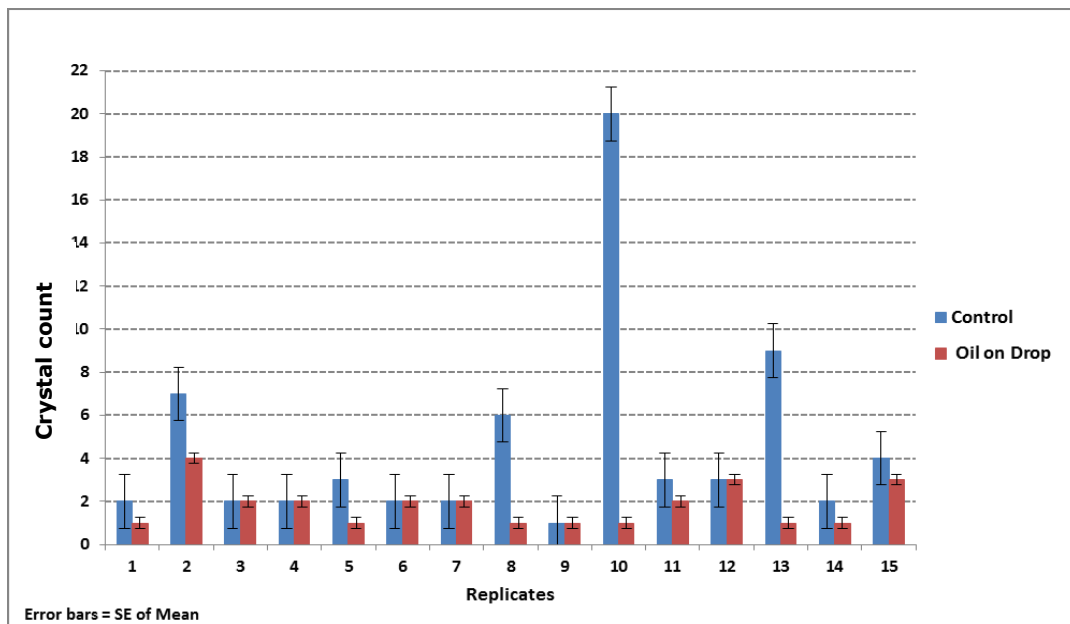


Figure 4. 12 Trypsin crystal count: Controls vs Oil-on-drop using the Mosquito

Statistical significance (2-tail T-test unequal variances: $p = 0.047 < \alpha = 0.05$; $n = 15$)
in crystal count between the control drops and the oil-on-drop.

Fig 4.12 shows the differences in crystals counted between the control drops compared to the oil-on-drop method for the 15 replicates ($n = 15$) trialled. When looking at the number of individual crystals, there was a statistical significant variation ($p = 0.047$) between the control drops and the oil-on-drop method with the control drops producing 68 crystals and

the oil-on-drop method 27 crystals. This equates to the controls drops producing 2.5 times the amount of crystals than the oil-on-drop method.

When comparing the Oryx 8 and the Mosquito when crystallising trypsin with 15 replicates, Fig 4.13 shows less crystals formed in controls (68 with the Mosquito vs. 156 with the Oryx 8) and with the oil-on-drop method (27 with the Mosquito vs. 48 with the Oryx 8). Overall, the Mosquito produced 109 fewer crystals than the Oryx 8 with trypsin across both controls and the oil-on-drop method.

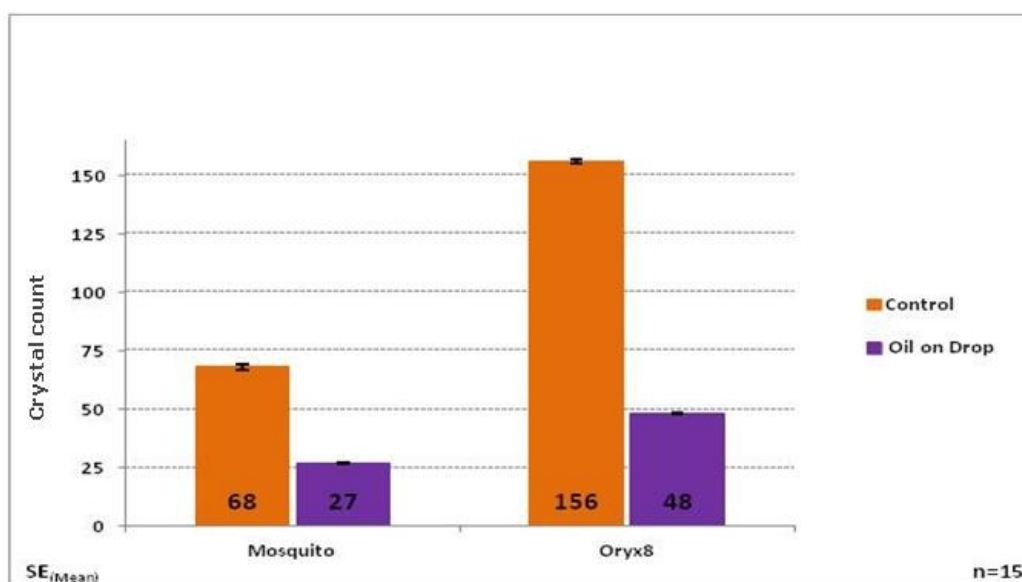


Figure 4. 13 Graph depicting trypsin crystals using the Oryx 8 and Mosquito: Control vs Oil-on-drop. Mean sample size (n) = 15

X-ray diffraction analysis of trypsin crystals showed the resolution limit for the best oil-on-drop crystals to 1.06 Å compared to 1.33 Å for the best control crystals. Diffraction resolution limit was improved by 0.27 Å. When comparing other data quality factors, the best performing oil crystal had an R_{meas} of 0.049; a signal to noise $I/\sigma(I)$ ratio of 16; a $CC_{1/2}$ of 0.999; and a dataset completeness of 92%. The best performing control crystal had an R_{meas} = 0.039; $I/\sigma(I)$ = 22; and dataset completeness = 89%; $CC_{1/2}$ = 0.999.

4.1.1.3 ALPHA CRUSTACYANIN

4.1.1.3.1 MANUAL TRIALS

The results for alpha crustacyanin followed a similar pattern as seen with the benchmark lysozyme and trypsin proteins, but with a more accentuated outcome as illustrated in Fig 4.14.

Crystals appeared after 4 days with the Chayen method and at ~5 days in the oil-on-drop method. Experiments with the Chayen oil barrier drops dried out after 5 days in Al's oil (50/50) and the 60/40 ratio and after 2 weeks in the 70/30 ratio (Fig 4.14 e). The 80/20 and 90/10 oil-on-drop ratios remained viable beyond 4 weeks. In the oil-on-drop method, drops did not dry at with any of the oil ratios trialled.

The final average crystal sizes for each method after 12 weeks of incubation showed the smallest crystals (~22.7x8x5µm) to be in the control drops; larger crystals (~120x12x10µm) in the Chayen method; and the largest crystals (~136x20x10µm) found in the oil-on-drop method.

Alpha crustacyanin was trialled with 10 replicates (n=10) per method. The result proved statistically significant ($F > F_{crit}$ (5.19 > 5.14); $p = 0.049$). Excessive crystal formation (>500 micro-crystals) was found in the controls (Fig 4.14 a) within 24hrs of incubation. Far fewer crystals were observed in the Chayen method (~50 crystals) (Fig 4.14 b) and oil-on-drop (~23 crystals) (Fig 4.4 c) methods.

The oil-on-drop method produced a small amount of oil run-off clearly visible as a ring around the drops (Fig 4.14 c). The drops were still adequately covered. This was seen in a few of the replicates and the 80/20 ratio had the least amount of run-off compared to the other oil ratios trialled.

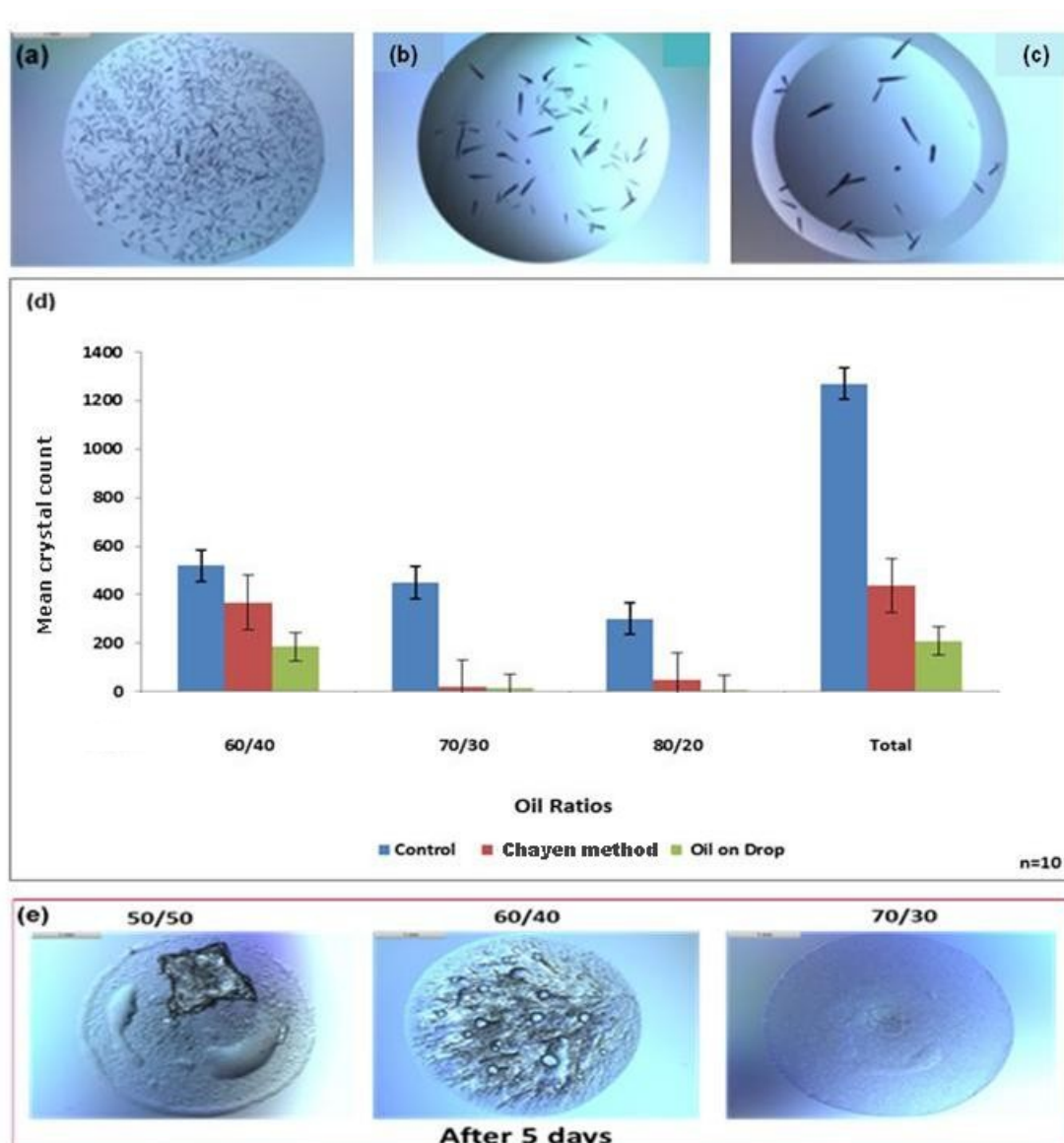


Figure 4. 14 Alpha crustacyanin crystallisation across methods and oil ratios (a) Control drop; (b) Chayen method; (c) Oil-on-drop method; (d) Graph showing statistically significant variation ($F > F_{crit}$ ($5.19 > 5.14$) $p = 0.049 < \alpha = 0.05$) in crystal counts across the methods and oil ratios; (e) Drops drying with the Chayen method using different oil ratios. Each oil ratio tested had a sample size (n) ≥ 10 per method.

4.1.1.3.2 AUTOMATED TRIALS

4.1.1.3.2.1 DOUGLAS INSTRUMENTS ORYX 8

Excessive crystal formation (>4000 crystals) was found in the control drops across all 16 replicates ($n=16$) after 24 hours of incubation (Fig 4.15 (a)). Far fewer crystals were observed in the corresponding replicates for the oil-on-drop method (~ 1960 crystals) (Fig 4.15 (b)) after 36 hours.

Crystal size measurements for each method was taken after 3 weeks of incubation, with final measurements showing the smallest crystals ($\sim 80 \times 20 \times 10 \mu\text{m}$) in the control drops and larger better formed crystals in the oil-on-drop method ($\sim 200 \times 40 \times 20 \mu\text{m}$). No oil run-off was experienced using the Oryx 8 (Fig 4.15 b).

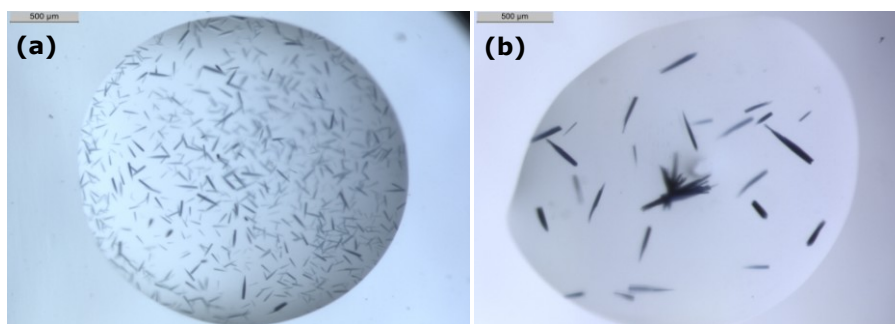


Figure 4. 15 Alpha-C crystal images of control drop vs oil-on-drop
(a) Control drop; (b) Oil-on-drop method

The results for alpha crustacyanin followed a similar pattern as seen with the benchmark proteins. The results proved statistically significant ($p=0.001$) (Fig 4.16).

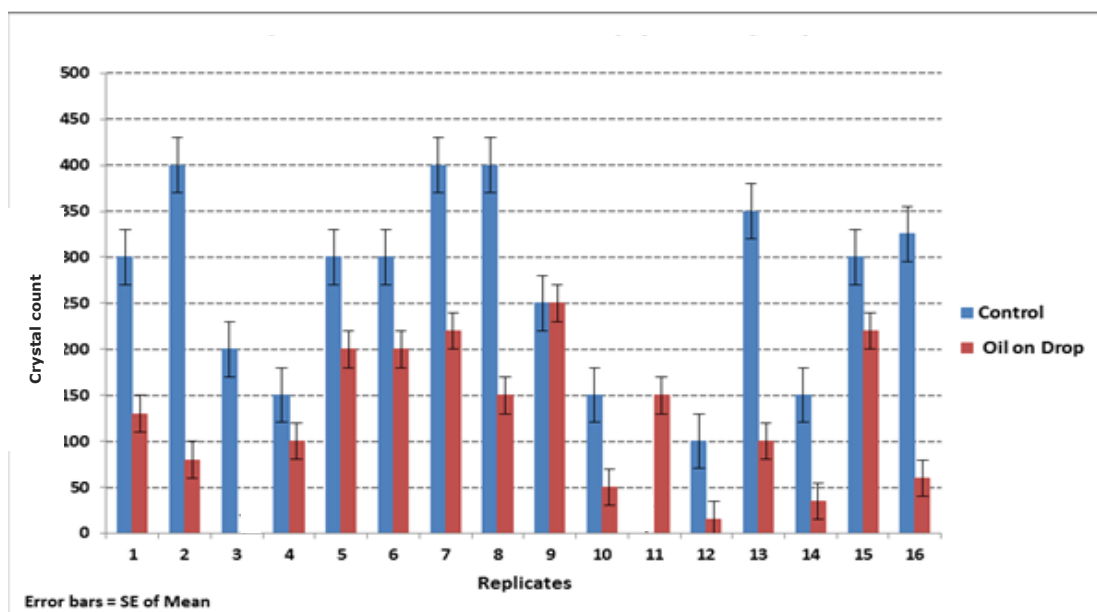


Figure 4. 16 Alpha crustacyanin crystal count: Controls vs Oil-on-drop using the Oryx 8

Statistical significance (2-tail T-test unequal variances: $p = 0.001 < \alpha=0.05$; $n=16$)
in crystal count between the controls and oil-on-drop method.

4.1.1.3.2.2 TTP LABTECH MOSQUITO LCP

The Mosquito liquid handler results for alpha crustacyanin followed a similar pattern as seen with the Oryx 8 system. 4950 crystals formed in the 11 replicates of control drops within 24 hours of incubation and 1740 crystals were observed in the oil-on-drop method where crystals appeared after 24 hours.

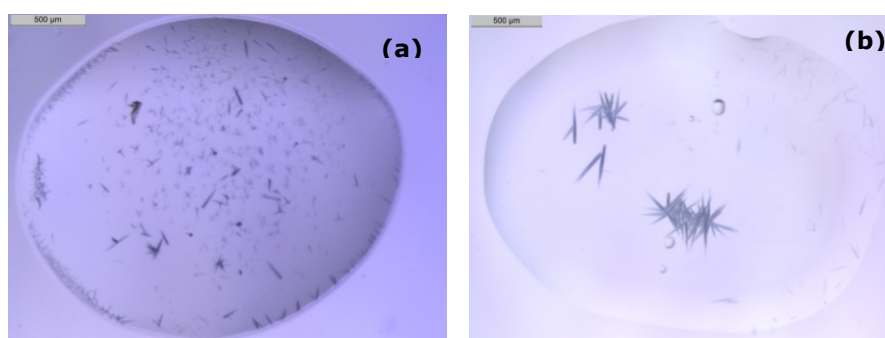


Figure 4.17 Images of alpha-C protein drops representing one paired replicate.

(a) Control drop; (b) Oil-on-drop method

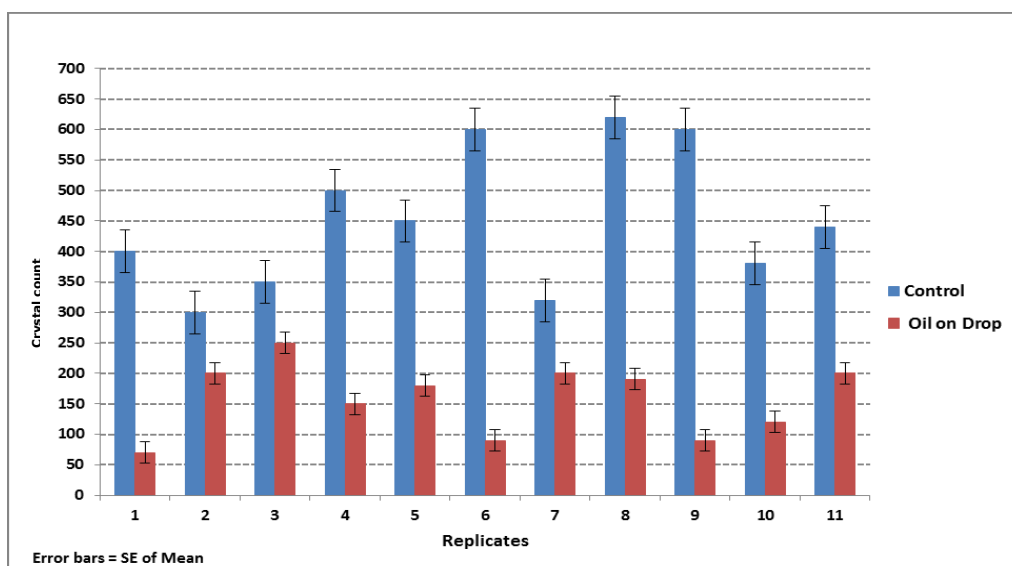


Figure 4.18 Alpha-C crystal count: Controls vs Oil-on-drop using the Mosquito

Statistical significance (2-tail T-test unequal variances: $p = 0.000 < \alpha = 0.05$; $n = 11$) in crystal count between the control and oil-on-drop drops.

Crystal size measurements for each method was taken after 3 weeks of incubation, with final measurements showing on average $\sim 40 \times 10 \times 5 \mu\text{m}$ in control drops and $\sim 180 \times 30 \times 10 \mu\text{m}$ for crystals from the oil-on-drop method. The results proved statistically significant ($p=0.000$) (Fig 4.18). The oil-on-drop method produced a small amount of oil run-off clearly visible as a ring around the drops (Fig 4.17 b).

When comparing the Oryx 8 and the Mosquito when crystallising alpha crustacyanin across 16 replicates, Fig 4.19 shows more crystals formed in the controls using the Mosquito (4960) compared to 4075 with the Oryx 8. The opposite effect was observed with the oil-on-drop method: 1740 crystals with the Mosquito vs. 1960 with the Oryx 8. Overall, with alpha crustacyanin, the Oryx 8 produced fewer crystals in total than the Mosquito.

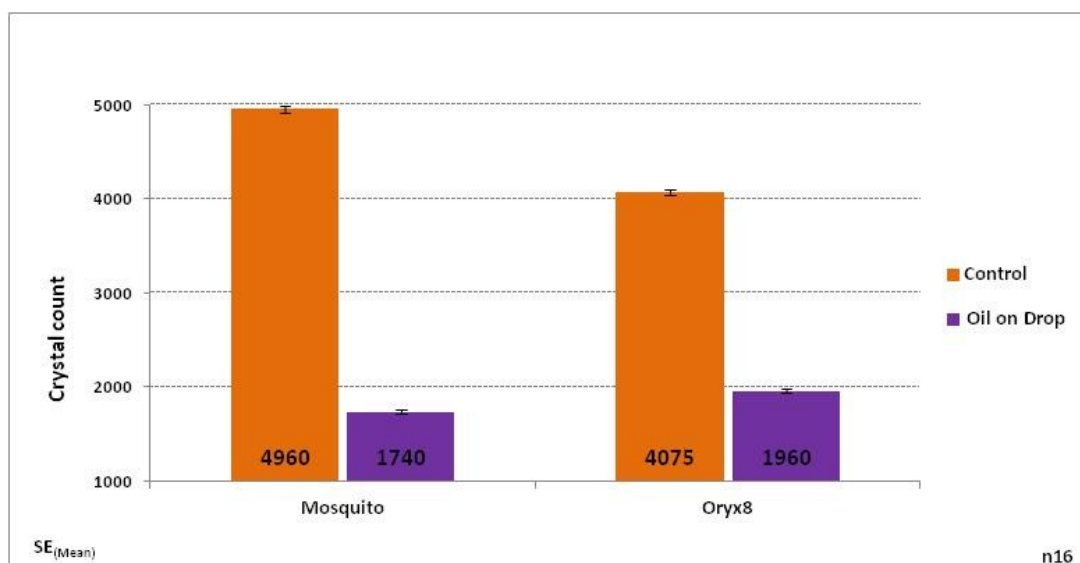


Figure 4. 19 Graph depicting alpha crustacyanin crystals using the Oryx 8 and Mosquito: Control vs Oil-on-drop. Mean sample size (n) = 16.

Although the crystals were fewer and larger using the oil-on-drop method, they did not produce any diffraction.

4.1.1.4 METHYLTRANSFERASE *LEGIONELLA PNEUMOPHILA* - LPG2936

4.1.1.4.1 MANUAL TRIALS

The results for methyltransferase *Legionella pneumophila* - lpg2936 followed a similar pattern as seen with the benchmark proteins and target protein alpha crustacyanin.

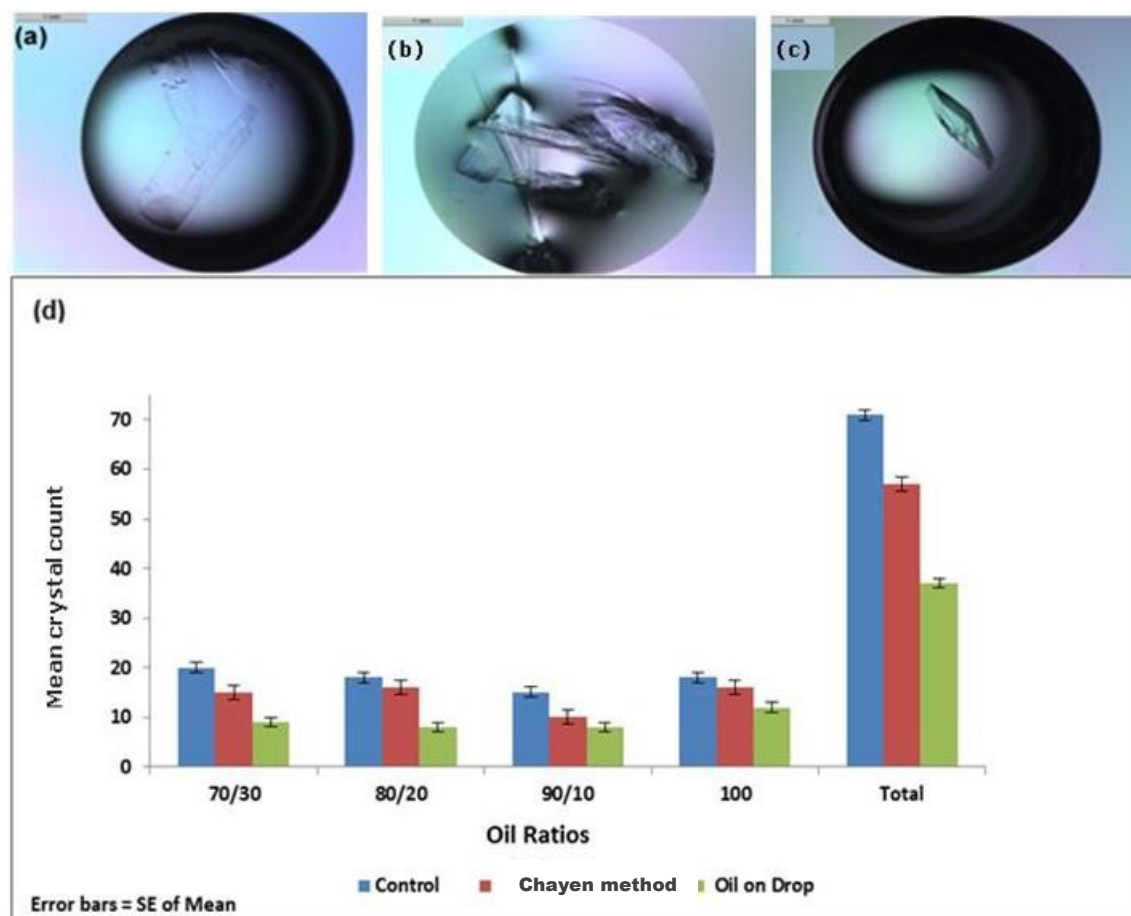


Figure 4. 20 Methyltransferase *Legionella pneumophila* - lpg2936 crystal across methods and oil ratios

(a) Control drop; (b) Chayen method; (c) Oil-on-drop method; (d) Graph showing statistically significant variation ($F > F_{crit}$ ($6.61 > 5.14$) $p = 0.030 < \alpha = 0.05$) in crystal count across the various methods and oil ratios. Each oil ratio tested had a sample size ($n \geq 10$ per method).

Crystals appeared within 24 hours after setup in control drops; after 24 hours with the Chayen method; and ~36 hours with the oil-on-drop method. Crystal growth was observed for 4 weeks and measured at the

end of that period. On average, the control drop crystals measured $\sim 800 \times 380 \times 100 \mu\text{m}$ (Fig 4.20 a); the Chayen method crystals $\sim 850 \times 350 \times 120 \mu\text{m}$ (Fig 4.20 b), and a single crystal in Fig 4.20 (c) for the oil-on-drop method measured $\sim 750 \times 300 \times 180 \mu\text{m}$. In this particular case, the oil-on-drop method was the only method to produce a single and well-formed crystal (Fig 4.20 c). The control and Chayen method produced crystals which were thin, plate-like and layered and some crystals had a needle-like morphology (Fig 4.20 a-b). Methyltransferase *Legionella pneumophila* - lpg2936 experienced very limited or no oil run-off during manual trials (Fig 4.20 c).

After additional trials were performed with 10 replicates per oil ratio and across all methods (Fig 4.20 (d)), cumulative crystal count data from these trials showed a statistical significance result ($F > F_{\text{crit}}$ ($6.61 > 5.14$); $p = 0.030$), highlighting a difference between controls, the Chayen method and the oil-on-drop method.

Additionally, data collected from crystals irradiated with X-rays showed that the crystals from the oil-on-drop method improved the X-ray resolution limit by 1.06 \AA when compared to the crystals from the controls. An improvement in diffraction resolution was achieved from a mean value of 2.85 \AA for the controls to a mean value of 1.79 \AA for the oils as illustrated in Figure 4.21 a-b. Crystals from the Chayen method did not produce good diffraction.

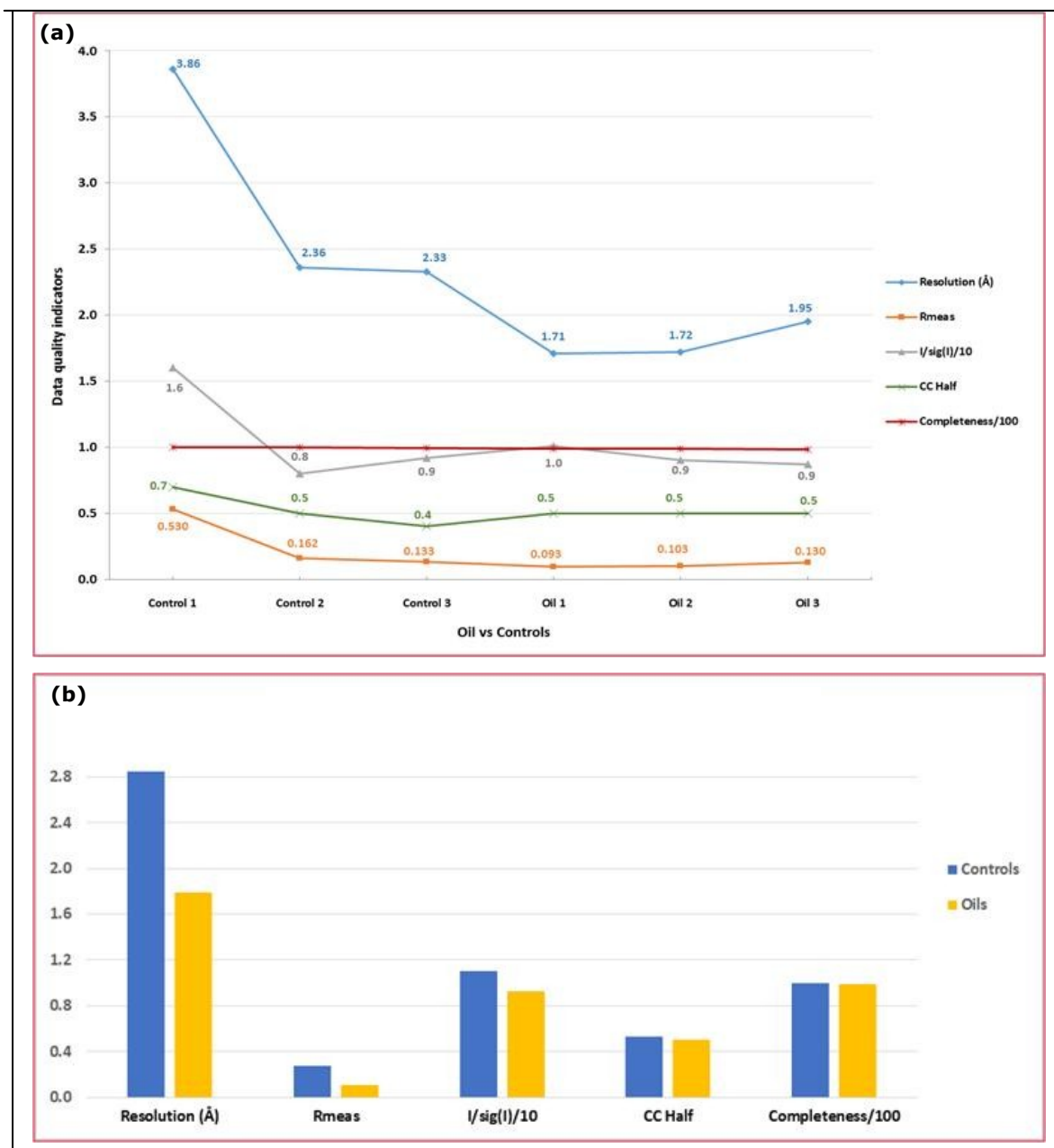


Figure 4. 21 Methyltransferase *Legionella pneumophila* - *lpg2936* X-ray data quality indicators

(a) Graph depicting controls vs oils; (b) Graph showing mean differences between quality indicators of controls vs oils. $I/\sigma(I)$ and $CC_{1/2}$ values are outer shell values. $I/\sigma(I)$ values have been divided by 10 and Completeness values by 100 to fit all the data quality indicators on the same graph's y-axis for ease of comparison.

There were very small differences in other X-ray data quality factors between the two groups (control and oil-on-drop) as shown in Fig 4.21 (b). These factors include (i) signal to noise ($I/\sigma(I)$); (ii) statistical significance

of the datasets measured by $CC_{1/2}$ ($CC_{1/2}$ - a primary indicator used for selecting high resolution cut-offs for data processing), (iii) and to a lesser extent how complete the datasets are.

In relation to the measure of how broad and complete the dataset distribution is, as measured by $R_{rim}(R_{meas})$, the mean oil-on-drop number was 0.11. This was less than half of the 0.28 value of the mean control drop number. Individually, the R_{meas} values recorded for each of the oil-on-drop datasets (oil 1 = 0.093; oil 2 = 0.103; and oil 3 = 0.130) came in at lower values than the controls (control 1 = 0.530; control 2 = 0.162; and control 3 = 0.133 (Figure 4.21 (a)).

Thus, all the individual datasets from the methyltransferase *Legionella pneumophila* - lpg2936 oil-on-drop method (with sample resolutions of 1.71 Å; 1.72 Å; and 1.95 Å) allows for a better starting point than those achieved by all the control samples with lower resolution values of 2.33 Å; 2.36 Å; and 3.86 Å.

4.1.1.4.2 AUTOMATED TRIALS

4.1.1.4.2.1 DOUGLAS INSTRUMENTS ORYX 8

The results for the target protein methyltransferase *Legionella pneumophila* - lpg2936 followed a similar pattern as seen with the benchmark proteins and target protein alpha crustacyanin.

From the 16 replicates that were set up, only 4 replicate sets did not yield any crystals. No obvious reason could be found for this other than possible experimental or human error. Crystals appeared within 24 hours after setup in control drops and after 24 hours with the oil-on-drop method. Crystal growth was observed for 4 weeks and measured at the end of that period. The number of crystals formed in the control drops amounted to 177 crystals, whilst only a total of 57 crystals were counted in the oil-on-drop method (Fig 4.23). Control drop crystals had a mean size of $\sim 400 \times 110 \times 40 \mu\text{m}$ (Fig 4.22 a) and the oil-on-drop method measured on

average $\sim 650 \times 150 \times 60 \mu\text{m}$ (Fig 4.22 b). Although it appears that run-off is present in Fig 4.22 b, it's not and the drop is covered with no run-off. The crystal in the drop has grown from within the solution and not from what appears to be oil.

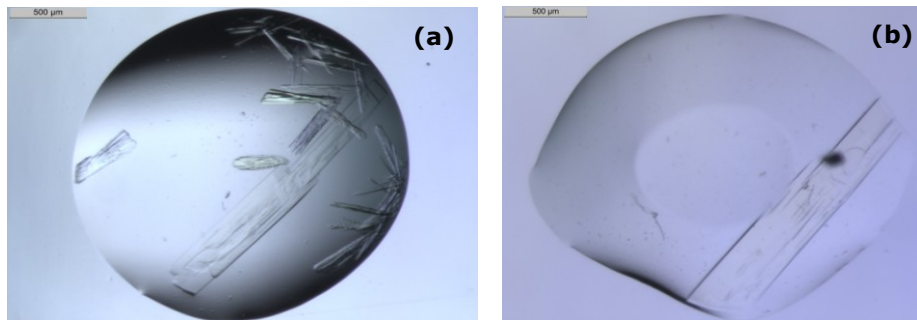


Figure 4. 22 Methyltransferase *Legionella pneumophila* - *lpg2936* crystal images of control drop vs oil-on-drop
(a) Control drop; (b) Oil-on-drop method

Fig 4.23 shows the statistical significance ($p=0.01$; $n=12$).

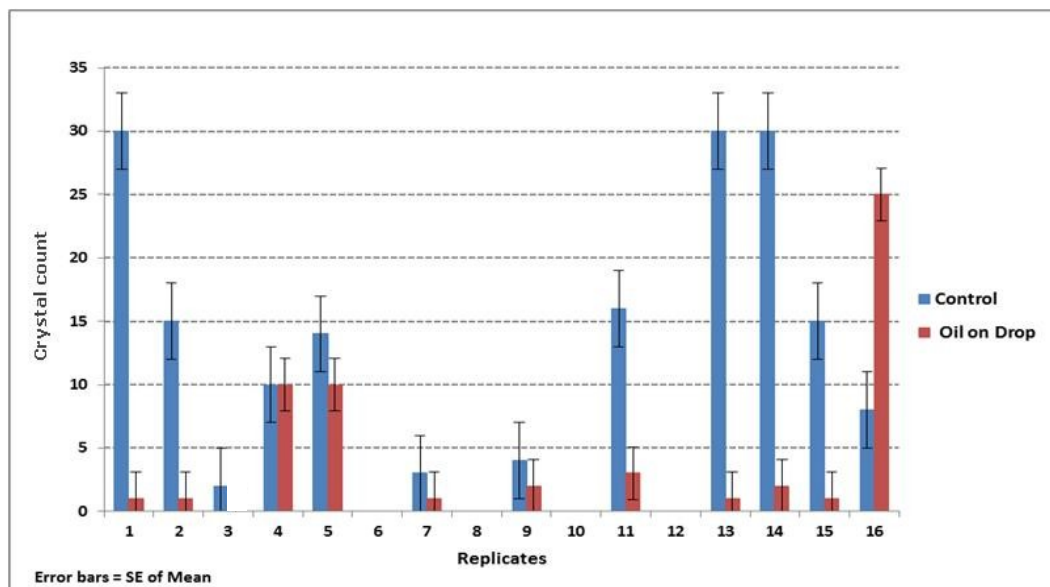


Figure 4. 23 Methyltransferase *Legionella pneumophila* - *lpg2936* crystal count: Controls vs Oil-on-drop using the Oryx 8
Statistical significance (2-tail T-test unequal variances: $p = 0.01 < \alpha=0.05$; $n=12$)
in crystals counted in controls and oil-on-drops.

4.1.1.4.2.2 TTP LABTECH MOSQUITO LCP

Similar to the Oryx 8, methyltransferase *Legionella pneumophila* - lpg2936 crystals appeared within 24 hours after setup in control drops and after 24 hours with the oil-on-drop method. Crystal growth was observed for 4 weeks and measured at the end of that period. The amount of formed crystals across the 16 replicates in the control drops amounted to 250 crystals, whilst only a total of 107 crystals were counted in the oil-on-drop method (Fig 4.25). Some oil run-off was experienced using the Mosquito (Fig 4.24).

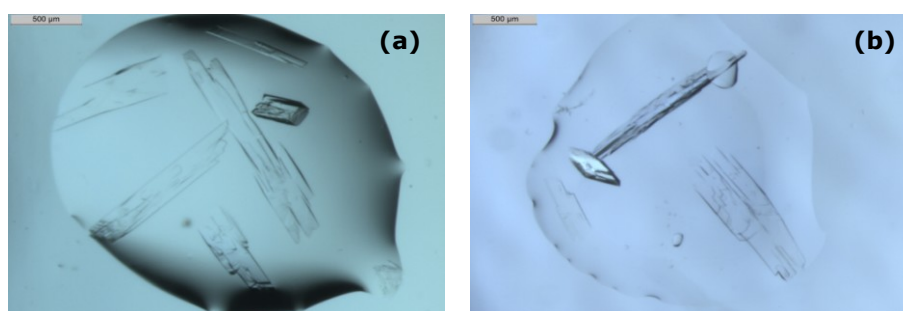


Figure 4. 24 Methyltransferase *Legionella pneumophila* - lpg2936 crystal images of oil drop vs control drop

(a) Control drop (b) Oil-on-drop method

The control drops had a mean crystal size of $\sim 600 \times 120 \times 40 \mu\text{m}$ and the oil-on-drop method measured on average at $\sim 600 \times 110 \times 40 \mu\text{m}$ which were very similar in size even though more crystals were produced in the control drops. The crystallisation results show statistical significance with a $p = 0.01$ value from crystals counted between control drops and oil-on-drop replicates (Fig 4.25).

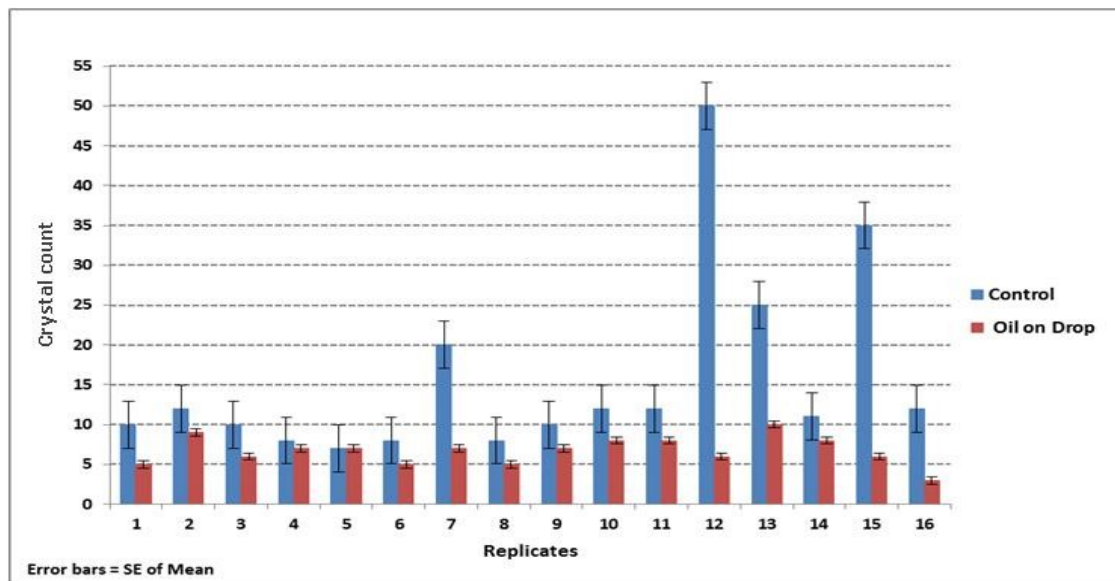


Figure 4. 25 Methyltransferase *Legionella pneumophila* - lpg2936 crystal count: Controls vs Oil-on-drops using the Mosquito

Statistical significance (2-tail T-test unequal variances: $p = 0.01 < \alpha = 0.05$; $n = 16$)
in crystal count between control and oil-on-drops.

A similar trend that was seen in lysozyme trials has been found with methyltransferase *Legionella pneumophila* - lpg2936, where in both methods, the Oryx 8 has produced fewer crystals compared to trials conducted with the Mosquito. With the oil-on-drop method, the Oryx 8 produced on average 50 crystals less than the Mosquito and when looking at the controls, the Oryx 8 produced on average 73 crystals less than the Mosquito (Fig 4.26).

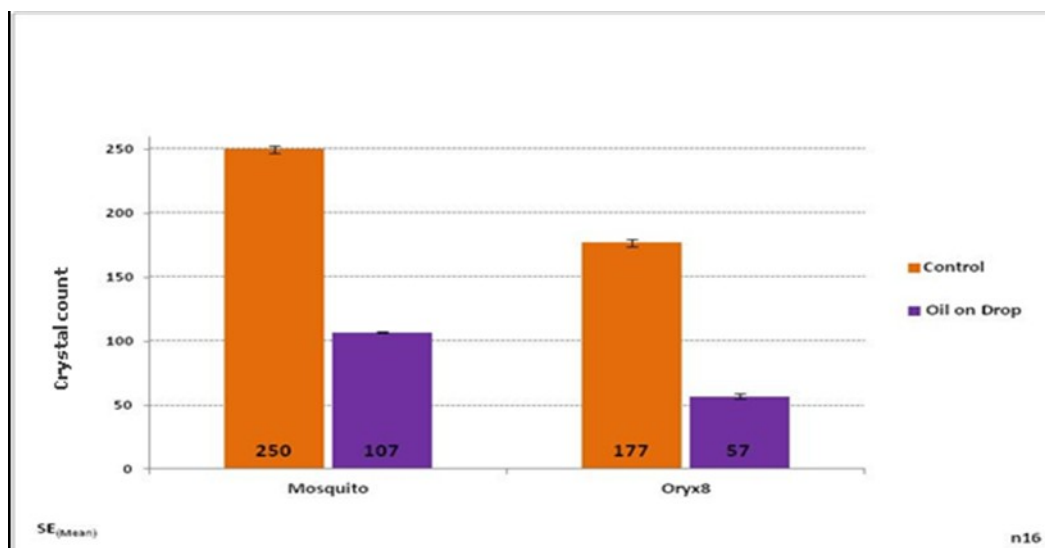


Figure 4. 26 Graph depicting methyltransferase *Legionella pneumophila* - lpg2936 crystals using the Oryx 8 and the Mosquito: Control vs Oil-on-drop. Mean sample size (n) = 16.

4.1.2 SUMMARY OF MEAN CRYSTAL SIZES: CHAYEN METHOD VS OIL-ON-DROP METHOD IN HDVD

The graphs in this section give a graphical snapshot of the *mean* crystal lengths (Fig 4.27 a) and widths (Fig 4.27 b) from manual HDVD trials across five proteins (lysozyme, trypsin, alpha crustacyanin, methyltransferase *Legionella pneumophila* - lpg2936 and the RoAb13-peptide complex – which is covered as a separate section in subsection 4.3) comparing controls, the Chayen method and the oil-on-drop method. A minimum of 5 replicates were set up per protein.

The Chayen method's crystal mean length and width were shorter than the controls (length 12.5% shorter and width 18% shorter). The oil-on-drop method produced the largest crystals (mean length of $\sim 427 \mu\text{m}$) compared to controls ($\sim 335 \mu\text{m}$) and the Chayen method ($\sim 298 \mu\text{m}$). Similarly, the oil-on-drop method's crystal mean width of width $\sim 170 \mu\text{m}$ was longer compared to controls ($\sim 149 \mu\text{m}$) and the Chayen method ($\sim 126 \mu\text{m}$) (Fig 4.27 a – b).

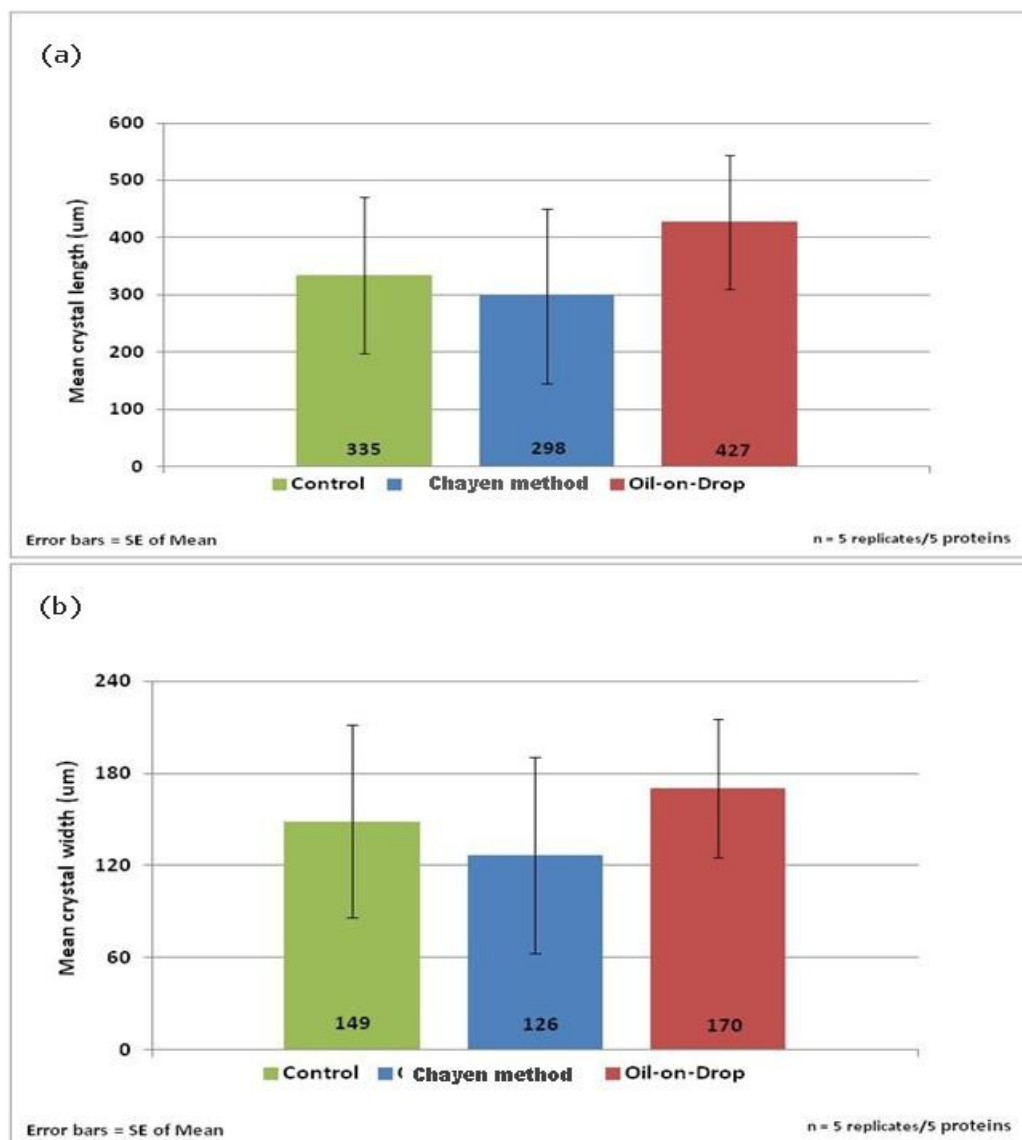


Figure 4.27 Mean crystal length (a) and width (b) taken from both the Oryx 8 and Mosquito in control drops; the Chayen method; and the oil-on-drop method tested using 5 proteins.

4.1.3 SUMMARY OF MEAN CRYSTAL SIZES: A COMPARISON BETWEEN THE ORYX 8 & THE MOSQUITO

Figs 4.28-4.29 illustrate the differences observed in average crystal sizes (length and width) between crystals found in the controls and oil-on-drop method when comparing the Oryx 8 and Mosquito systems.

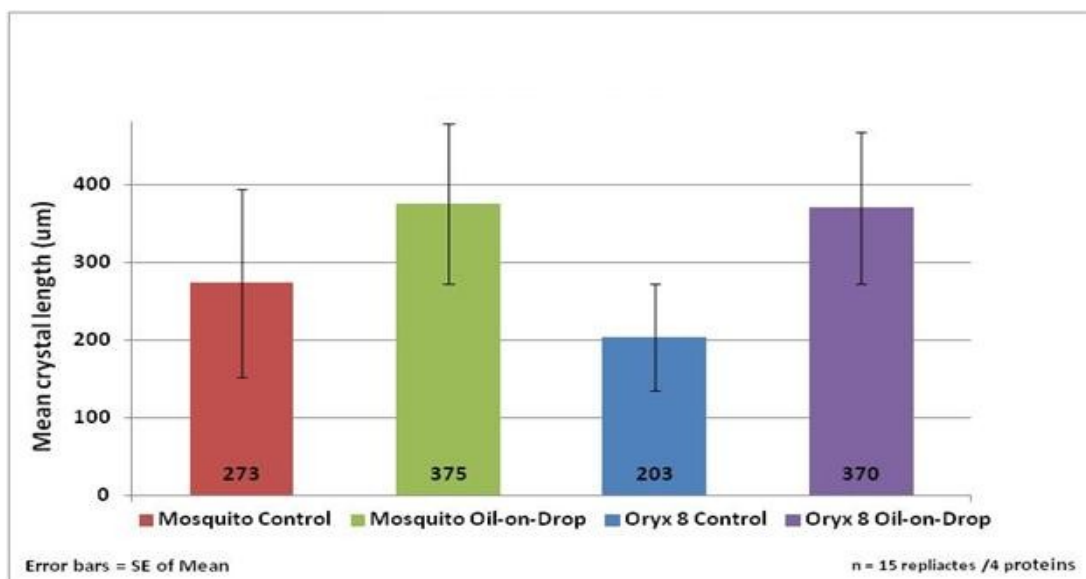


Figure 4. 28 Graph depicting mean crystal length using the Oryx 8 and Mosquito: Controls vs Oil-on-drop.

Results were statistically insignificant ($F < F_{crit}$ ($0.77 < 2.82$) $p = 0.517 > \alpha = 0.05$) in mean crystal length and width between the controls and the oil-on-drops across both automated systems tested using 4 proteins.

Across the four proteins and 15 replicates per protein, the average crystal length using the Mosquito increased by 37.5% from ~273 µm to ~375 µm when using the oil-on-drop method. On the Oryx 8 the average length increased by 82% from ~203 µm to ~370 µm with the oil-on-drop method compared to controls (Fig 4.28).

In terms of crystal width as seen in Fig 4.29, a similar result was observed with the oil-on-drop method, the average width using the Oryx 8 was larger than what was observed on the Mosquito. Crystal width increased by 68% on the Oryx 8 compared to 21.5% on the Mosquito.

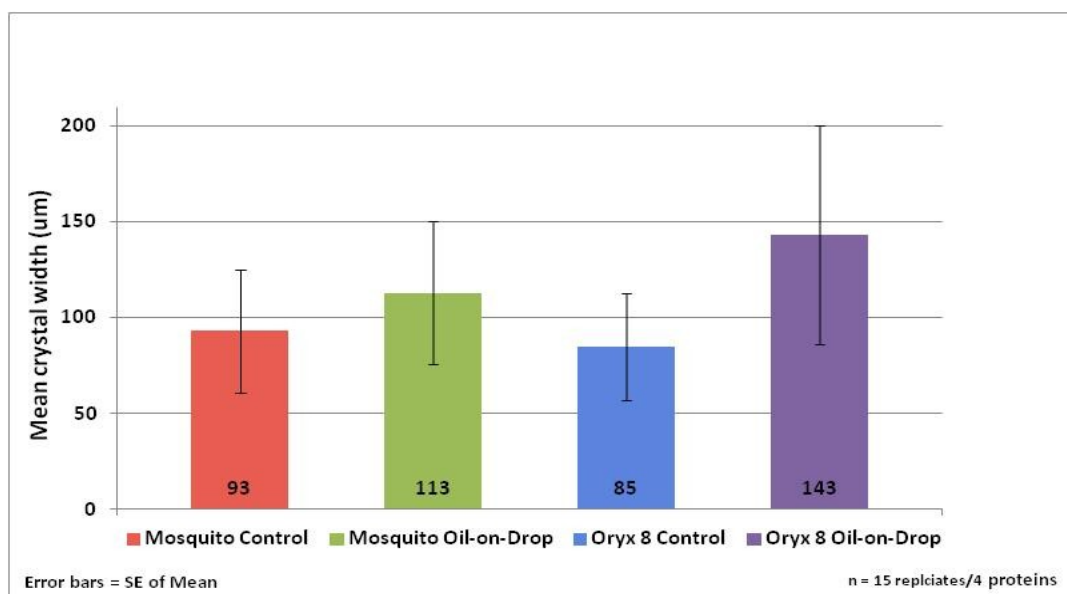


Figure 4. 29 Graph depicting mean crystal width using the Oryx 8 and Mosquito: Controls vs Oil-on-drop.

Results were statistically insignificant ($F < F_{crit}$ ($0.77 < 2.82$)) $p = 0.517 > \alpha = 0.05$) in mean crystal length and width between the controls and the oil-on-drops across both automated systems tested using 4 proteins.

In summary, when comparing the mean crystal sizes (length x width) of all the proteins tested (lysozyme, trypsin, α -C and methyltransferase *Legionella pneumophila* - lpg2936) of the controls versus the oil-on-drop method using both the Oryx 8 and Mosquito systems, it was found that results were not statistically different ($F < F_{crit}$ ($0.77 < 2.82$); $p = 0.517$) i.e. the difference in mean crystal size weren't big enough to infer a statistical significance between the two systems.

4.2 HETEROGENOUS NUCLEANT TRIALS

4.2.1 THAUMATIN

Carbon black ink and CPG were tested as nucleants with thaumatin. Fig 4.30 shows the working phase diagram constructed for thaumatin with protein concentrations ranging from 10-30mg/ml against potassium sodium tartrate (NaKT) precipitant concentrations of 0.2M-0.6M. The red squares represent clear drops and the green represent where crystals

formed. The nucleants were introduced into the clear drops at 10mg/ml of the protein at 0.2M and 0.3M NaKT concentrations and also at higher supersaturation conditions of 10mg/ml at 0.4M and 0.5M NaKT.

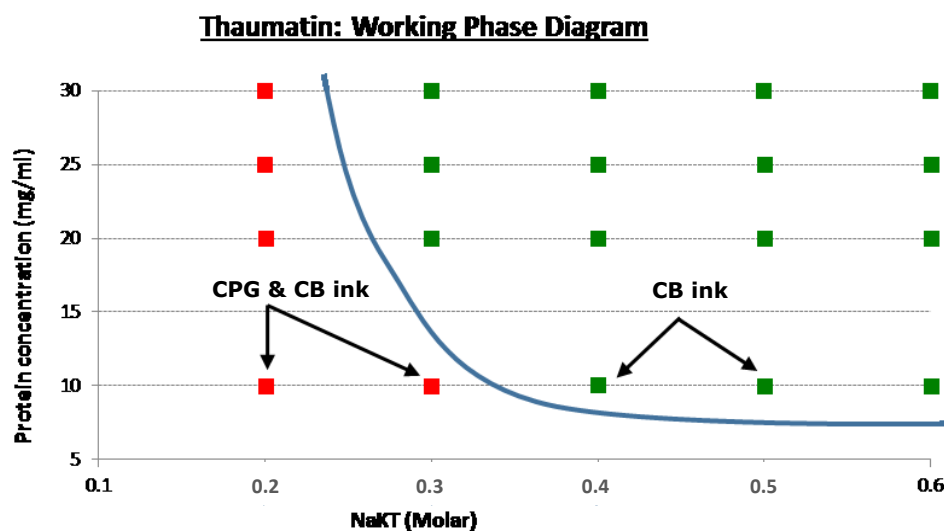


Figure 4. 30 Working phase diagram for thaumatococcus

The solid black arrows represent the protein and NaKT concentrations at which the CPG and CB ink were introduced into crystallisation drops. The red squares represent clear drops at 0.2M and 0.3M NaKT and the green squares represent crystals formation at higher supersaturation conditions (0.4M and 0.5M NaKT). The blue curved line represents the super solubility curve where spontaneous crystallisation occurs above the line (green squares) and conditions where the crystallisation solution remains clear if left undisturbed (red squares).

CB ink was initially tested at conditions within the nucleation zone (refer to Fig 1.3 in Chapter 1) where crystals would form spontaneously at 0.4M and 0.5M NaKT. The images in Fig 4.31 (c) and (d) show that the CB ink produced a few large crystals after 24 hours. At 24 hours the controls were clear but after 48 hours the controls caught up and ultimately produced showers of very small crystals (Fig 4.31 (a) and (b)).

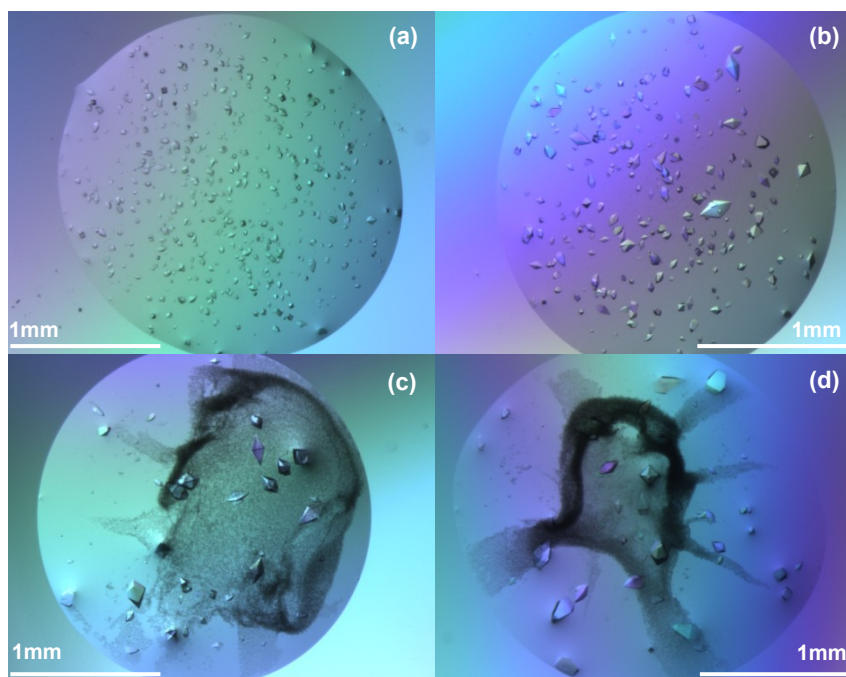


Figure 4.31 Thaumatin crystals with and without the CB ink. (a) & (b) Showers of small crystals after 48 hours without the CB ink set up with 0.4 and 0.5M NaKT at 10mg/ml; (c) & (d) Fewer, larger crystals after 24 hours with the CB ink set up with 0.4M and 0.5 NaKT at 10mg/ml.

Further to this, when adding CB ink and the CPG into crystallisation trials as shown in Figs 4.32 (a) and (b), crystals were obtained with CB ink at 0.2M NaKT at 10mg/ml after 7 days and the controls caught up after 8 days; whilst with 0.3M NaKT at 10mg/ml crystallisation occurred after 3 days with the controls catching up on day 8. Thus, the introduction of the carbon black ink at (0.2M - 0.3M NaKT) hastened crystal formation of thaumatin by 1 day at 0.2M NaKT and 5 days at 0.3M NaKT.

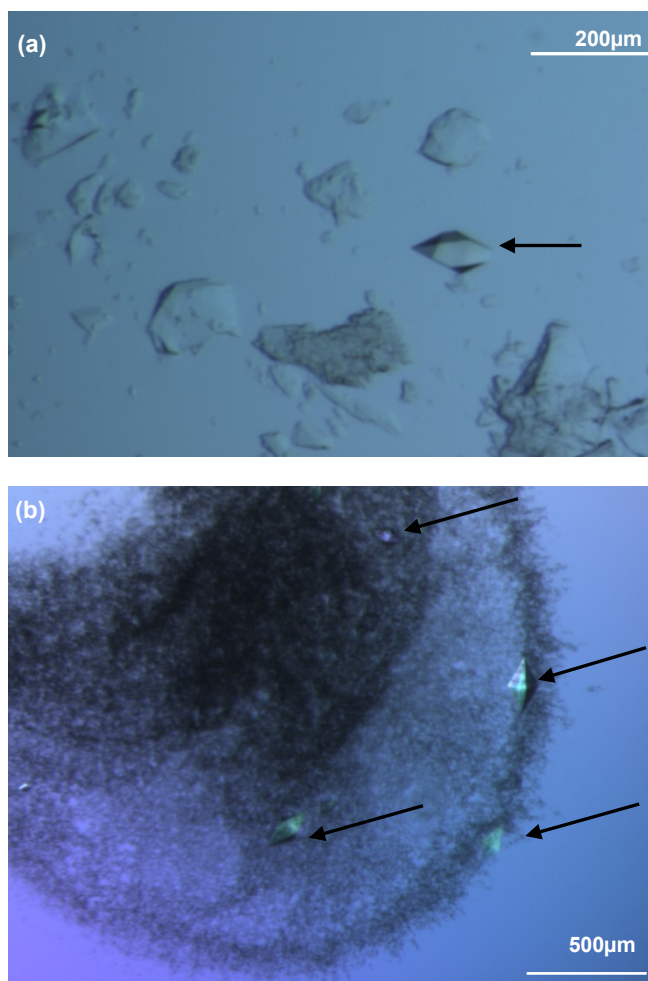


Figure 4. 32 Thaumatin crystals produced with CPG and CB ink
(a) A single crystal amongst the CPG grains (black arrow) with 0.2M and 0.3M NaKT at 10mg/ml; (b) A few crystals scattered within the CB ink (black arrows) with 0.2M and 0.3M NaKT at 10mg/ml.

The presence of the CPG nucleant with 0.2M and 0.3M NaKT at 10mg/ml of thaumatin produced crystals after 3 days whilst the control drops caught up after 8 days of trial setup. Thus, the introduction of CPG hastened crystal formation of thaumatin by 5 days.

4.2.2 CATALASE, TRYPSIN AND LYSOZYME

The CB ink and untreated CPG were then trialled on other benchmark proteins such as catalase, lysozyme and trypsin. Both the CB ink and the

CPG nucleants did not produce any positive results. No crystals were formed in their presence of any of these proteins.

In addition, chemically treated CPG nucleants were also tested with the detailed results shown in Table 4.4. As highlighted in the table, none of the chemically treated CPG nucleants and the untreated CPG produced any positive results.

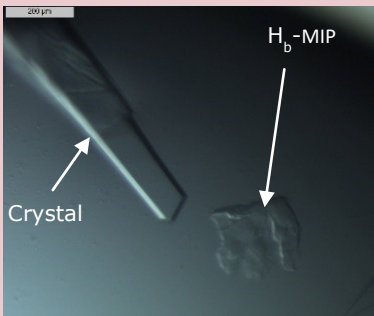
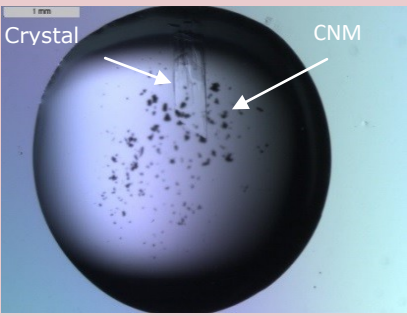
Table 4. 4 Results of CPG nucleants with catalase, trypsin and lysozyme

Nucleant: CPG chemical treatment	Pore size (nm)	Polar/non-Polar	Catalase 20mg/ml 24% PEG 3350 (>36 days)	Trypsin 50mg/ml 5% PEG 8000 (>21 days)	Lysozyme 10-15mg/ml 0.1-0.25M NaCl (>21 days)
Normal (untreated)	Control	---	1 crystal	crystalline material	clear drops
	8		1 crystal	clear drops	clear drops
	48		clear drops	clear drops	clear drops
Methyl Group (CH ₃)	Control	Light hydrophobic	1 crystal	light crystalline material	1 crystal
	8		1 crystal	clear drops	1 crystal
	48		clear drops	clear drops	clear drops
Amino Group (NH ₂)	Control	Light hydrophilic	1 crystal	light crystalline material	clear drops
	8		clear drops	clear drops	
	48		clear drops	light crystalline material	
Carboxyl Group (COOH)	Control	Heavy hydrophilic	clear drops	light crystalline material	1 crystal
	8		clear drops	clear drops	1 crystal
	48		clear drops	clear drops	1 crystal
Halogen Group (F)	Control	Heavy hydrophilic	clear drops	light crystalline material	clear drops
	8			clear drops	
	48			clear drops	

4.2.3 TARGET PROTEIN: – METHYLTRANSFERASE *LEGIONELLA PNEUMOPHILA* - LPG2936

Methyltransferase *Legionella pneumophila* - lpg2936 was tested with 5 nucleants: haemoglobin-MIP (H_b-MIP); PEGylated graphene oxide (PEG GO); CB ink; bioglass; and NPG. No nucleant-induced crystallisation was observed with CB ink, bioglass and NPG (Table 4.5). With haemoglobin-MIP, crystals appeared after 3 days and control drops produced crystals after 5 days. With PEG GO crystals appeared after ~2 days while control drops produced crystals after 4 days.

Table 4. 5 Nucleant and resolution results on *Legionella pneumophila* - lpg2936

Chayen Reddy H _b -MIP	Liquid CNM (PEG-GO)	CB Ink	Bioglass	NPG
 <p>Crystal</p> <p>H_b-MIP</p>	 <p>Crystal</p> <p>CNM</p>	No crystals	No crystals	No crystals
Control - 2.33 Å MIP - 1.91 Å	Control - 2.3 Å CNM - 2.1 Å			

X-ray analysis on available crystals from the H_b-MIP and PEG GO trials showed that the H_b-MIP-induced crystals diffracted to 1.91 Å with its controls to 2.33 Å while PEG GO-induced crystals diffracted to 2.1 Å with its controls also at 2.3 Å. H_b-MIP improved resolution by 0.42 Å and PEG GO by 0.23 Å.

4.3 OPTIMISATION TRIALS ON THE TARGET ROAB13-PEPTIDE COMPLEX

Previous work conducted during the early stages of my research focused on an anti-CCR5 (RoAb13 antibody) and its cognate peptide. Background information is covered in Chapter 1, subsection 1.6.3. A stepwise and methodical optimisation approach on the RoAb13-Peptide complex was followed and included the following:

4.3.1 CONVENTIONAL OPTIMISATION: VARYING THE PROTEIN AND PRECIPITANT CONCENTRATIONS

Crystals of the RoAb13-Peptide complex were obtained initially by the HDVD technique at 10mg/ml of the protein complex with 2M ammonium sulphate as the precipitant. In order to optimise the crystallisation

conditions, a working phase diagram was generated by varying the ammonium sulphate concentrations between 1.8-2M and the protein concentration at 8 and 10mg/ml (Fig 4.33).

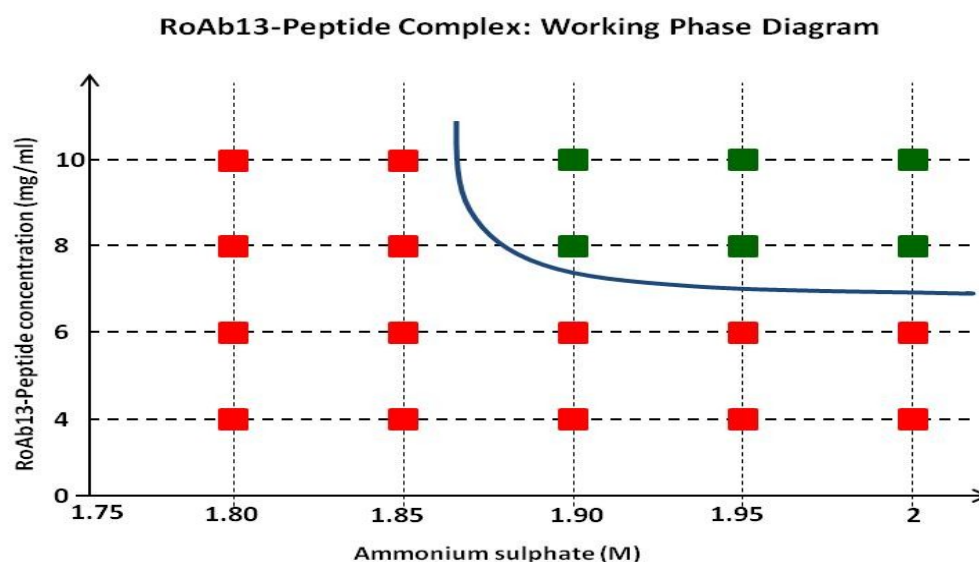


Figure 4. 33 Working phase diagram for the RoAb13-Peptide complex
The red squares represent clear drops and the green squares represent the formation of crystals. The blue curved line represents the super solubility curve where spontaneous crystal occurs above the line (green squares) and conditions where the crystallisation solution remains clear if left undisturbed (red squares).

Table 4.6 presents the crystallisation results of the RoAb13-Peptide complex.

Table 4. 6 RoAb13-Peptide complex optimisation

Protein concentration (mg/ml)	Ammonium sulphate precipitant concentrations (M)				
	1.8	1.85	1.9	1.95	2
10	X	X	P [†] C ^Φ	P; C ^h	P; C ^h
8	X	X	P [†] C ^Φ	P; C ^h	P; C ^h
6	X	X	X	X	X
4	X	X	X	X	X

On average a higher number of crystals were produced at 10mg/ml than at 8mg/ml and at higher precipitant concentrations, i.e. 2M produced more crystals at both 8mg/ml and 10mg/ml than with 1.9M of ammonium sulphate. Samples were incubated overnight.

X – Clear drops; P – precipitate; C- crystals; † - after 24hrs; h - at 24hrs; Φ - at 48hrs

The crystals produced at 10mg/ml of the complex were X-rayed yielding a diffraction resolution of 4.7 Å. All the other crystals irradiated with X-rays produced very low or no diffraction. Data analysis of the 4.7 Å crystals showed the presence of the RoAb13 antibody and possible space for the peptide. At such a low resolution, there was not enough information available to solve the structure of the complex; hence, further optimisation was required to obtain crystals which would diffract to higher resolution.

4.3.2 VARYING INCUBATION PERIODS OF THE ANTIBODY AND PEPTIDE

Varying the incubation period of the antibody with the peptide (10mg/ml and 8mg/ml) was done to test for effective binding of the complex before crystallisation trials.

4.3.2.1 INCUBATION TRIALS AT 10MG/ML

The samples at 10mg/ml of the RoAb13-Peptide complex were incubated for 2, 3 and 4 hours. All trials were carried out in duplicates. The 2 hour incubation gave crystals and precipitate with 2M ammonium sulphate 24 hours after trial set up, whilst the 3 and 4 hour incubation periods yielded more crystals 24 hours after trial set up.

Following these results, the samples were then incubated for two weeks at 10mg/ml with precipitant concentrations of 1.7M – 2M ammonium sulphate. The results are presented in Fig 4.34 which shows showers of crystals formed at higher concentrations of the precipitant solutions (1.8M – 2M); a single crystal was observed in two separate protein drops at 1.75M; and at 1.70M no crystals appeared and the drops remained clear.

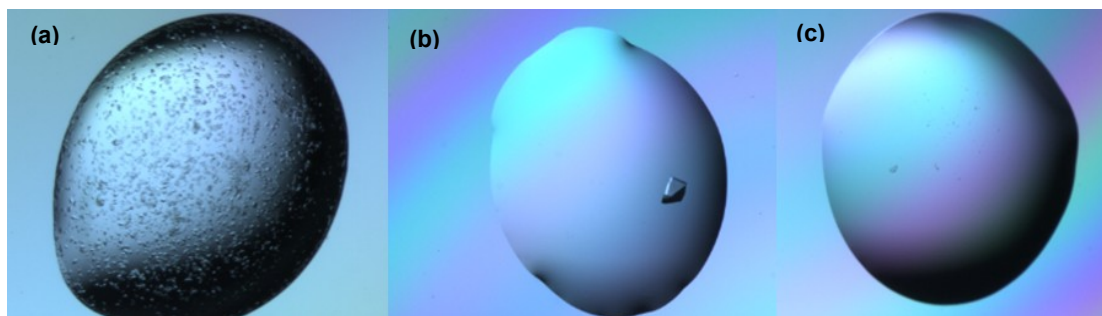


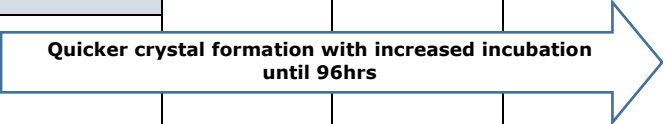
Figure 4. 34 RoAb13-Peptide complex crystals after a 2 week incubation period.

(a) Excessive crystal formation occurs at 10mg/ml with 1.8M-2M ammonium sulphate; (b) A single crystals formed at 10mg/ml with 1.75M; (c) At 10mg/ml with 1.70M and under the drops were clear.

4.3.2.2 INCUBATION TRIALS AT 8MG/ML

The results for the incubation trials at 8mg/ml are presented in Table 4.7. Trials were conducted at 2M ammonium sulphate with incubation periods of 12 hours, 24 hours, 72 hours, 96 hours, 120 hours and 144 hours. Crystals appeared after 4 days for the 12 hour incubation; 3 days for the 24 hour incubation; 2 days for the 72 hour incubation; and only 1 day when the RoAb13-Peptide complex was incubated for 96 hours. The 120 hour incubation produced crystalline material after 24 hours from which large crystals were produced after 3 weeks. The 144 hour incubation failed to produce any well-formed crystals, only crystalline material appeared 8 days after trial set up.

Table 4. 7 The effect of the various incubation periods on crystal formation of the 8mg/ml RoAb13-Peptide complex at 2M ammonium sulphate

Length of Incubation						Observations
12hrs	24hrs	72hrs	96hrs	120hrs	144hrs	
Clear	Clear	Clear	Crystals	Crystalline Material	Clear	Day 1
Clear	Clear	Crystals			Clear	Day 2
Clear	Crystals				Clear	Day 3
Crystals					Clear	Day 4
					Clear	Day 5
					Clear	Day 6
					Clear	Day 7
Crystals	Crystals	Crystals	Crystals	CM	CM	Day 8
						Day 9
					CM	Day 10
						>2Wks
Small Crystals	Smaller-sized Crystals	Large Crystals	Large Crystals	Large Crystals	CM	>3Wks

CM – Crystallise Material. Changes in colour-coding (light to dark) per incubation period accentuate the first observed crystals at different times (in days).

4.3.3 SLOWING CRYSTALLISATION WITH THE OIL-ON-DROP METHOD

Another strategy to slow down the crystallisation process of the RoAb13-Peptide complex to improve X-ray diffraction was by employing the novel oil-on-drop method used effectively on other benchmark and target proteins.

After trial set up, crystals formed within 24 hours in control drops; after 24 hours using the Chayen method; and ~36 hours with the oil-on-drop method. The Chayen and oil-on-drop methods produced fewer crystals than controls. In total, more crystals were observed in the control drops (~180) than with the Chayen and oil-on-drop methods (~138) (Fig 4.36). Crystals were observed for a total of 8 weeks and crystal sizes were measured at that time. Fig 4.35 (a-c) shows the control drops produced smaller crystals than the oil-on-drop method and Chayen method. Average

sizes for controls were $\sim 350 \times 150 \times 150 \mu\text{m}$; $400 \times 170 \times 170 \mu\text{m}$ for the Chayen method and $\sim 450 \times 180 \times 180 \mu\text{m}$ for the oil-on-drop method. Oil run-off occurred in the oil-on-drop method (Fig 4.35 c) with the 80/20 ratio but did not occur in other replicates.

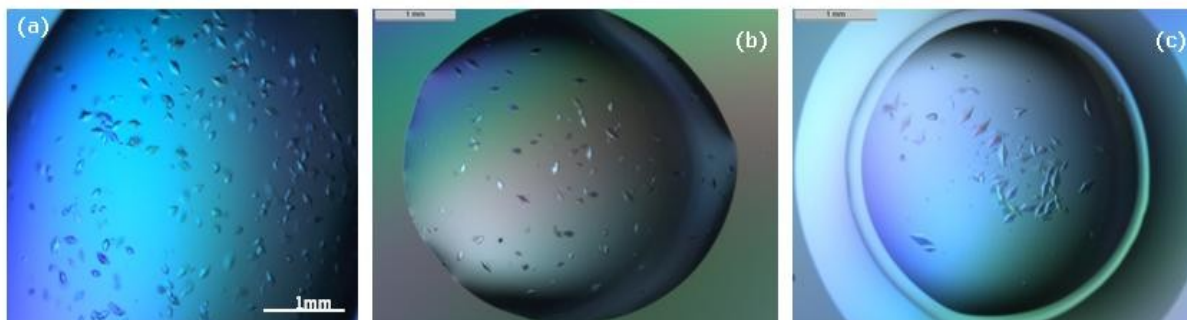


Figure 4. 35 Example of the anti-CCR5 RoAb13-Peptide complex crystals produced with controls and oil barrier methods

(a) Control drop; (b) Chayen oil barrier method; (c) Oil-on-drop drop with 80/20 paraffin/silicone oils.

No statistical tests were performed. Looking at the mean crystal count and mean crystal sizes in controls and both oil methods, it is reasonable to assume that no difference exist between the Chayen and oil-on-drop methods (Fig 4.36).

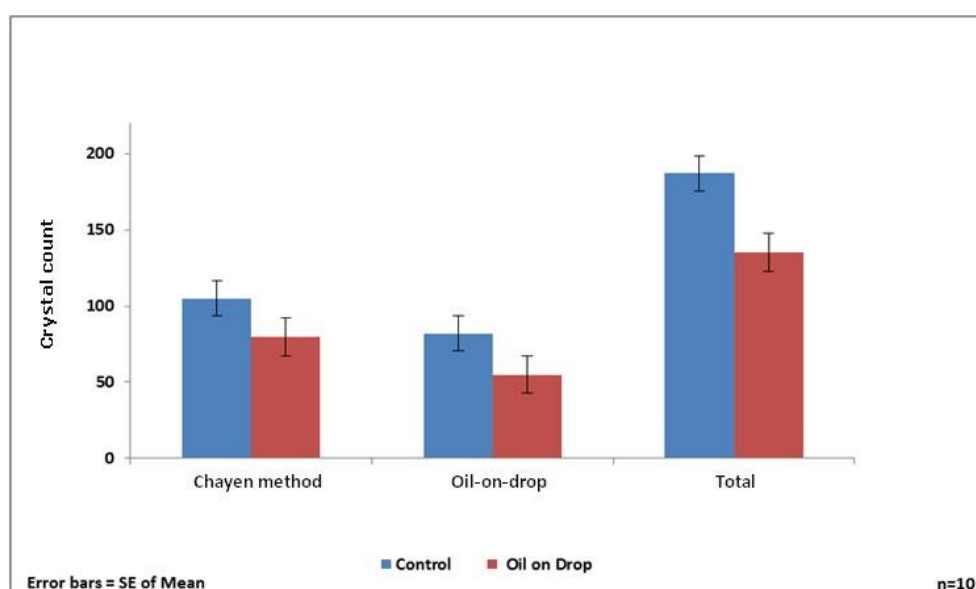


Figure 4. 36 Anti-CCR5 RoAb13-Peptide complex mean crystal count using the Chayen oil barrier and oil-on-drop methods with 80/20 oil ratio with 10 replicates ($n=10$) per method.

Crystals taken from the oil-on-drop method produced a resolution of 3.6 Å – a significant improvement on the ~4 Å achieved with the other optimisation strategies.

5

DISCUSSION

5.1 A NOVEL OIL BARRIER METHOD & AUTOMATION

5.1.1 SLOWING DOWN CRYSTALLISATION: ENHANCING SINGLE CRYSTAL GROWTH AND IMPROVING X-RAY DIFFRACTION WITH THE OIL-ON-DROP METHOD

The novel oil-on-drop method presented in this thesis has made a positive contribution to the field of protein crystallisation specifically in optimisation trials as evidenced by the results achieved. Another advantage of this method was that it lent itself to automation. The novel oil-on-drop method was conceived and successfully tested on two benchmark proteins: lysozyme and trypsin and three target proteins: alpha crustacyanin (α -C), methyltransferase *Legionella pneumophila* - lpg2936) and RoAb13-Petide complex. This method succeeded in producing fewer, larger single crystals of higher quality for use in X-ray diffraction studies.

The ultimate aim for the oil-on-drop method was to assist in achieving improved X-ray diffraction data. In this study four proteins were sent for X-ray analysis: *Legionella pneumophila* - lpg2936, trypsin, α -C and the RoAb13-Petide complex. The diffraction resolution of the crystals of *Legionella pneumophila* - lpg2936, trypsin and the RoAb13-Petide complex was improved using the oil-on-drop method. Furthermore, *Legionella pneumophila* - lpg2936 crystals grown with the oil-on-drop method was compared with crystals grown with a range of available nucleants, notably haemoglobin-MIP and PEGylated graphene oxide (PEG GO), which are the only two nucleants which successfully induced crystal formation at metastable conditions. Comparisons of the resolution from crystals grown with the oil-on-drop method and crystals grown with these nucleants show

that the highest resolution was found with the oil-on-drop method. This is covered in more detail in subsection 5.2.2.

The following subsections 5.1.1.1 – 5.1.1.4 cover a detailed discussion of the trialled proteins which had crystals sent for X-ray analysis.

Statistical analysis of the crystallisation results of lysozyme, trypsin, alpha crustacyanin (α -C) and methyltransferase *Legionella pneumophila* - lpg2936 proved statistically significant for mean crystal count and for mean crystal length and width, with a few exceptions (Fig 4.6; Fig 4.8 d and Figs 4.28-4.29).

5.1.1.1 METHYLTRANSFERASE *LEGIONELLA PNEUMOPHILA* – LPG2936

Methyltransferase *Legionella pneumophila* - lpg2936 produced single large crystal which diffracted to 1.71 Å (Fig 4.20 c) instead of crystal stacks and needles (Fig 4.20 a - b) with the oil-on-drop method. The diffraction of *Legionella pneumophila* - lpg2936 crystals grown with the oil-on-drop method significantly improved the mean diffraction compared to controls. This was seen across a range of replicates (Fig 4.21 a - b). For some unexplained reason crystals grown by the Chayen method did not produce good diffraction. While I was doing this work, the structure of *Legionella pneumophila* - lpg2936 was solved at 1.5 Å (with a higher resolution shell range of 1.57 Å – 1.49 Å) by Pinotsis and Waksman (2017).

Further to this, X-ray data quality factors (excluding resolution) for controls and the oil-on-drop method (Fig 4.21 (b)) were comparable except for *Rrim*(*Rmeas*) where the oil-on-drop method produced lower values which are favoured for data processing. Also, when comparing other data quality factors such as data completeness, signal to noise ($I/\sigma(I)$) ratio and statistical significance of the datasets ($CC_{1/2}$), the oil-on-drop method's best performing crystals were highly comparable to the best crystal dataset used by Pinotsis and Waksman (2017) for *Legionella pneumophila* - lpg2936 published structure. Pinotsis and Waksman (2017) quoted data completeness

at 95.0%; signal to noise ($I/\sigma(I)$) ratio at 10.6; and $CC_{1/2}$ at 0.725. The oil-on-drop method's best crystal dataset were 99% for data completeness; 10 for $I/\sigma(I)$; and 0.6 for $CC_{1/2}$. These are all higher resolution shell values. It would be plausible to assert that the structure of *Legionella pneumophila* - lpg2936 could as easily have been solved with the best crystals from the oil-on-drop method presented in this thesis.

5.1.1.2 TRYPsin

X-ray diffraction analysis of trypsin crystals showed the resolution limit for the best oil-on-drop crystals to 1.06 Å compared to 1.33 Å for the best control crystals. Diffraction resolution limit was improved by 0.27 Å. As the structure of trypsin is already known at 1.02 Å (<http://www.rcsb.org/structure/1S0Q>), performing X-ray analysis on trypsin crystals was merely a test of efficacy of the oil-on-method to see whether it could equal or better the diffraction compared to controls and the known structure. With trypsin, the crystal morphology between control drops and the two oil methods were aesthetically similar and choosing the "best" crystals for X-ray analysis was purely based on the crystal size. As shown, crystals from the oil-on-drop show the efficacy of this method by producing high quality crystals. In theory, a slower and more ordered aggregation of nuclei and packing of the crystal lattice allows for better X-ray diffraction patterns and higher resolution limits than poorer crystals (García-Ruiz, 2003), which the oil-on-drop method definitely delivers.

Although resolution is a very important factor in solving 3-D structures, other data quality factors are also considered. When looking specifically at X-ray data quality factors other than resolution, results were comparable between the best performing oil-on-drop crystal and the best performing control crystal. In this case priority would be given to the actual resolution for structure determination especially when other data quality factors were comparable between the best oil-on-drop crystal and the best control crystal (Chapter 4, subsection 4.1.1.2.2.2).

5.1.1.3 α -C

The oil-on-drop method was successful in producing crystals of α -C which were large enough for X-ray analysis as opposed to showers of tiny crystals observed in controls drops. Although the crystals were fewer and larger using the oil-on-drop method, they did not produce any diffraction. To date, researchers have not been able to obtain diffracting crystals to enable the determination of the structure of α -C. Historically, attempting to solve the structure of α -C by various technologies including X-ray crystallography, has always been problematic and has engaged researchers for decades (Chayen *et al.*, 2003; Dellisanti *et al.*, 2003; Nneji and Chayen, 2004; Rhys *et al.*, 2011 and references therein). Lower-resolution techniques have been used to identify the overall assembly of α -C (Rhys *et al.*, 2011). Problems with the crystallisation and diffraction studies with α -C may be attributed to the complexity of α -C complex as a 16mer heterodimer (Zagalsky and Jones, 1982). The best diffraction of 10 Å (within a range between 5 Å – 12 Å) was achieved by Chayen *et al.* (2003) which focused on crystallisation optimisation trials in both microbatch and VD techniques. Work on α -C continues.

5.1.1.4 RoAb13-PEPTIDE

Single large crystals from the oil-on-drop method diffracted to a resolution of 3.6 Å compared \sim 4 Å - 4.7 Å from smaller crystals using conventional methods, thus improving resolution. Additional optimisation trials were performed on the antibody peptide complex which is discussed in more detail in Section 5.3.

Overall, the X-ray data on *Legionella pneumophila* - Ipg2936, trypsin and the RoAb13-Peptide complex showed that the oil-on-drop method facilitated a slower crystallisation process leading to larger and better quality crystals. In the case of *Legionella pneumophila* - Ipg2936 and trypsin, results with the oil-on-drop method have produced very comparable X-ray data compared to their published structures and better results compared to their respective

controls. For the RoAb13-Peptide complex, the oil-on-drop method allowed for its best resolution to date to be achieved.

5.1.1.5 INCREASED CRYSTAL SIZE

Time and again crystallographers are faced with the same question of whether growing larger single crystals leads to better diffraction data for 3-D structure determination. Historically, the common consensus was that the larger the crystal, the greater the chance of solving a structure through the achievement of higher resolution data (Benvenuti and Mangani, 2007). The data taken from this study, especially when looking at the new oil-on-drop method appears to support this.

This study showed that the oil-on-drop method successfully reduced the number of crystals formed by slowing down the equilibration rate and thus approaching supersaturation slowly. This led to the formation of fewer, larger, better diffracting crystals, and also increased mean crystal size (length and width) in all proteins compared to controls and the Chayen method. The oil itself, as seen with the oil-on-drop method, played an important part in the outcome by affecting the crystallisation process throughout its various stages of nucleation and growth (Chayen, 1997).

When comparing mean crystal sizes across the various oil barrier methods for all proteins (Figs 4.27 (a) and (b)), the Chayen method skewed the overall results of trypsin trials. It was expected that mean crystal length in the Chayen method across all proteins would be larger than in controls but smaller than with the oil-on-drop method as the oil-on-drop method proved superior. Instead mean length and width in the Chayen method were smaller than in controls. This is due to the PEG issue with trypsin crystallisation. The Chayen method does not work with solutions containing high molecular weight PEGs such as PEG 5000 and PEG 8000 (used to crystallise α -C and trypsin respectively). The Chayen method also works less effectively with 2-methyl-2,4-pentanediol (MPD) as a precipitant (Chayen, 1997). The oil-on-drop method proved advantageous with trypsin crystallisation as it works

well with PEG as a crystallising solution. The trypsin drops with the oil-on-drop method did not dry after trial set up.

5.1.2 OIL RATIO MIXTURES

5.1.2.1 MANUAL TRIALS

Different combinations of oil mixtures were tested on all the proteins during manual trials. The use of different oil ratio combinations in microbatch trials were first proposed by D'Arcy *et al.* (1996) and the concept was further developed by Chayen (1997) in vapour diffusion trials. Whilst embarking on the oil-on-drop method development, testing different oil ratios seemed the most logical approach. After trying out different combinations as mentioned in the previous Chapters, the most consistent and effective oil ratio was the 80% paraffin, 20% silicone (80/20).

It is known that silicone is more permeable and less viscous than paraffin and allows for more evaporation to occur between protein-precipitant drops and their reservoirs leading to a faster rate of supersaturation and crystal formation when using more silicone. Thus, the less silicone (and more paraffin) used, the greater the effect on limiting and slowing down crystal growth leading to fewer and larger crystals. This has been borne out in the results where the least effective ratios were the paraffin-to-silicone ratios of 50/50 and 60/40. In order to avoid cutting off evaporation completely and preventing crystallisation from occurring, it was advisable to include a small amount of silicone oil in oil barrier experiments (D'Arcy *et al.*, 1996; Chayen, 1997).

Other oil combinations can also be explored. Experimentally, researchers may choose to set up and test different oil ratio combinations for their trials (e.g. 65/35, 75/25, 85/15, or any other combination for that matter). Due to the labour intensive nature of optimisation trials and the added time required to set up the oil-on-drop and Chayen methods, it was not feasible to explore all combinations during this study. Results from this study and the

choices made within are a good starting point for further work on exploring other oil ratio combinations and others should base decisions on their own crystallisation requirements and need. That said, based on the results presented within this thesis, there is a strong indication that the differences in results between say an 80/20 ratio versus a 75/25 ratio may not be significant enough to warrant tests at both ratios.

It has been shown that when comparing the 80/20 ratio to a 70/30 ratio, results supports 80/20 as the ratio which produced on average better, reproducible and more consistent results across all proteins. A good example of this was witnessed during the crystallisation of α -C with 60/40; 70/30 and 80/20 oil ratios in the oil-on-drop and Chayen methods. Smaller crystals were observed in ratios with more silicone oil in the oil mixture; and as the silicone portion in the ratio decreased and the paraffin portion increased fewer and larger crystals were produced (Fig 4.14 d). The 80/20 ratio produced the least amount of crystals in the oil-on-drop method. The 70/30 ratios produced more crystals in the Chayen method and even more in the control drops. The 60/40 ratio was the least effective in slowing the crystallisation process compared to the other two ratios. When looking at the 60/40 ratio specifically, the most crystals were found in controls, fewer in the Chayen method and the least amount of crystals in the oil-on-drop method.

On a final note on oil ratio combinations, there are two anomalies to highlight. In the case of lysozyme for some reason the 50/50 and 60/40 oil ratios produced fewer crystals than the more favoured 70/30 and 80/20 ratios. This clearly goes against the expected trend. In addition, using the 60/40 ratio, it was found that the control drops produced fewer crystals than what was observed in the Chayen method. This is also an anomalous result. In keeping with the results seen in this study, the oil-on-drop method produced the fewest crystals (Fig 4.2 d). The stated anomalies do not fit the trend observed with both oil barrier methods tested across the other proteins and oil ratios tested. Crystallisation especially with proteins is not a perfect science and as crystallisation is inherently multi-parametric, in some cases it can still be difficult to achieve the right result or keep to expected trends.

Even if variables that influence the crystallisation process are minimised, they cannot be completely eliminated (Pusey *et al.*, 2005).

5.1.2.2 AUTOMATED TRIALS

Given the ease, consistent and most reproducible results obtained with the 80/20 ratio during manual HDVD trials; the 80/20 ratio was used in setting up the automated oil-on-drop trials. As automated systems work seamlessly with the SDVD method (refer Fig 1.5 (iii)), and the oil-on-drop method is amenable to automation, a straightforward comparison and validation of oil-on-drops (i.e. "oil-on-sitting drops") to control drops was possible. A comparison with the Chayen method, which is not easily incorporated into automated trials, was not made.

The most frequently used automated systems for protein crystallisation trials are the Douglas Instruments Oryx 8 and Labtech TTP's Mosquito systems. It was essential to test the oil-on-drop method on both systems as they are widely used within the crystallisation and crystallography research communities.

In terms of crystal length and width (Fig 4.29), the oil-on-drop method allowed for the crystals to grow to larger sizes but it appears that the Oryx 8 was more effective in growing larger crystals than the Mosquito. This result could possibly be explained by the Mosquito's biggest drawback, which is contact dispensing. There was always contact between the dispensing tip and the protein-precipitant drops during runs. In a majority of cases (~90%), the contact tips did not appear to affect the sitting drops physically after contact, but in some instances (~10%) the appearance of the drops was affected as the dispensing tips "split" the protein-precipitant drops in two on contact. In effect, it created two smaller separate drops per well from the original single drops.

In contrast, as the Oryx dispensing system uses a dedicated oil dispensing tip and air to push out the required oil with no physical contact, it left the

drops undisturbed during oil-on-drop dispensing. It is not known how the splitting of the drops using the Mosquito had affected the rate of crystallisation and the quality of the crystals produced in these split drops. For statistical purposes, results from these split drops were excluded from any analysis and replacement replicates were set up to allow for a like for like comparison with results from the Oryx 8 system.

5.1.3 OIL RUN-OFF DURING TRIALS

In terms of the handling of the oils during trial set up, the 80/20 ratio was the most effective in reducing oil run-off. Limiting the quantity of silicone oil to 20% of the mixture (i.e. 80/20) enabled enough internal friction and “stickiness” for the oil to mostly sit on top of the protein-precipitant drops with minimal run-off compared to other oil ratios with a higher percentage of silicone present. This can be attributed to the viscosity of the paraffin oil (at 80% of the mixture). At a limit, the use of the 70/30 ratio may still be a useful alternative, but it does depend on experimental requirements. It can be speculated that using the 90/10 ratio would be even better than the 80/20 ratio in minimising oil run-off as it uses even more paraffin oil, but in terms of crystallisation, it is not as effective as the 80/20 oil ratio. Other variables that may also affect oil run-off are the protein and precipitant used. It is important to note that oil run-off occurred to a lesser extent with oil ratios with less than 30% silicone. This was seen with lysozyme (Fig 4.2 c); trypsin (Fig 4.8 c); α -C (Fig 4.14 c) and the RoAb13-Peptide complex (Fig 4.35 c). The least amount of oil run-off was experienced with *Legionella pneumophila* - lpg2936.

In addition, it appeared that the oil run-off phenomena occurred on average more readily with trypsin (Fig 4.8 c; Fig 4.9 c - d; and Fig 4.11 b), than it did with the other proteins. This may not be an entirely random occurrence or caused by a lack of experimental precision in manual trials, as the oil run-off was also visible during a few automated trials. The reason for the oil run-off behaviour with trypsin cannot be fully explained other than postulating that somehow the oils were “interacting” with the trypsin solution in a similar way to how oil interacts with water. Clearly, no homogeneous mixture occurs as

oil and water are immiscible and oils are inert. This run-off was merely a reaction that allowed for the excess oil to move to a position away from the protein-precipitant drop. This same reasoning could be used for all the other proteins used in these trials, but why does it have a more drastic affect with trypsin? No oil run-off was expected and just enough oil was dispensed to form a layer on top of the protein-precipitant drop. It could be due to the precipitant used as trypsin crystallises with PEG 8K. It is not known how the PEG interacts with the oils. No specific research evidence was found explaining such interactions which require further investigation.

5.2 HETEROGENOUS NUCLEANT TRIALS

5.2.1 NEW NUCLEANTS

To compare the efficacy of CPG and CB ink as heterogeneous nucleants, a range of other nucleants which have previously been successful in crystallisation trials have also been tested in this study (Chayen *et al.*, 2001; Chayen *et al.*, 2006; Saridakis *et al.*, 2011; Govada *et al.*, 2016). Conventionally, it is typical when testing different heterogeneous materials as potential nucleants, that some nucleants have a positive effect in crystallising certain proteins or none or a lesser effect on others.

From the results presented in this thesis, CB ink and untreated CPG only produced positive crystal-inducing results compared to controls on one benchmark protein thaumatin. Tests on other benchmark proteins catalase, lysozyme and trypsin did not show any success with CB ink, the untreated and the chemically treated CPGs. Interestingly, as illustrated in Fig 4.2 (a) and (b), instead of inducing many more crystals than controls at higher supersaturation levels it appears that the CB ink had an inhibitory effect (acting as an anti-nucleant) on crystal growth after initial crystal formation, thus leading to fewer and larger sized crystals. On the one hand, producing larger crystals for X-ray analysis is a positive, but, on the other hand, the CB ink appeared to have arrested further crystal formation where spontaneous nucleation would occur. At these levels, the controls produced showers of

smaller crystals as expected. It was expected that the nucleant drops would at least mirror what was seen in the controls. The crystal counts in the controls were at least 3-4 times higher than the number seen in the drops containing CB ink. Nucleants are conventionally inserted into protein drops at lower concentrations (metastable conditions) as demonstrated with the untreated CPG and additional trials with CB ink. The results at metastable conditions are what we were hoping to see where both nucleants induced crystal formation before this occurs in control drops (in some cases control drops remain clear).

Unlike other heterogeneous nucleants (such as Naomi's nucleant - i.e. bioglass, MIPs, PEG GO)) which were often successful as heterogeneous nucleants (Chayen *et al.*, 2001; Chayen *et al.*, 2006; Saridakis *et al.*, 2011; Khurshid *et al.*, 2015; Govada *et al.*, 2016), the untreated CPG as well as the CB ink as mentioned previously, did not show any success after tests on catalase, lysozyme and trypsin. After receiving chemically functionalised CPG nucleants from collaborators, these were tested in addition to the untreated CPG as a final test to see if the CPG nucleants could be effective. The functionalised CPGs failed to induce crystal formation at metastable conditions compared to controls. In some instances the CPG nucleant drops remained clear whilst control drops formed crystals. This is further evidence of the ineffectiveness of the whole range of CPG nucleants.

It was hoped that the functionalised CPGs would perform better at inducing crystal formation but this wasn't borne out in the results. I can only postulate that for some reason the length of time the functionalised CPGs were in possession may have affected their functionality i.e. the chemical treatments lost their potency and reacted as if untreated. The results show that in many cases across the two pore sizes (8nm and 48nm); the functionalised CPGs had similar or even worse outcomes than the untreated CPG and controls. This was evident across all the proteins tested. Also the differences in pore sizes of the CPG (8nm and 48nm) did not have an effect, results were equally disappointing.

Operationally, and as stated in the Materials and Methods section, initial trials with CPG were dispensed into the protein drops using the tip of a forceps to pick up individual grains of CPG. This proved very laborious and challenging as the grains were difficult to collect even after dipping the forceps tip into precipitant solutions to aid grain transfer into drops. The grains did not always find their way into the protein drops. A possible solution around this problem was to transform the nucleant grains into an aqueous suspension to allow for easy pipetting into drops. This made it easier to conduct the manual nucleant trials but may have also created a negative effect on the individual nucleants nucleation-inducing properties. To allow nucleants to be used more widely in optimisation trials it is critical to make them amenable to automation and high-throughput without losing their potency. The effect of such transformations from solid nucleants into liquid form and its effects on crystal induction, especially on automated systems, needs to be studied in much more detail and wasn't explored in this thesis.

5.2.2 EXISTING NUCLEANTS: A QUICK COMPARISON WITH THE OIL-ON-DROP METHOD

A comparison was made between the new oil-on-drop method and a range of existing nucleants on the crystallisation of *Legionella pneumophila* - Ipg2936 (Table 4.5). The main reason for this comparison was to further test the efficacy of the oil-on-drop method as an optimisation strategy and its ability to produce crystals of high quality for X-ray studies.

The focus was on utilising some of the most effective and readily available nucleants within the laboratory. The nucleants were: PEG GO (Govada *et al.*, 2016); molecular imprinted polymer haemoglobin-MIP (Saridakis *et al.*, 2011; Khurshid *et al.*, 2015); bioglass (Chayen *et al.*, 2006); and nanoporous gold (NPG) (Kertis *et al.*, 2012). CB ink was included as a final test of this nucleant on a target protein as previous trials on benchmark proteins failed. Nucleants are not random materials, but have been systematically produced and studied for their effects on protein crystallisation for decades. Naomi's nucleant and CNMs were developed and used based on the size of their surface pores (Chayen *et al.*, 2001; Chayen *et al.*, 2006;

Khurshid *et al.*, 2015; Govada *et al.*, 2016). MIPs work on imprinting which leaves a memory of the imprinted protein in the polymer gel (Saridakis *et al.*, 2011).

Crystal formation in *Legionella pneumophila* - lpg2936 was hastened with PEG GO and haemoglobin-MIP showing the effectiveness of these nucleants. Bioglass, NPG and CB ink did not induce crystallisation and proved ineffective as nucleants in *Legionella pneumophila* - lpg2936. These were solid nucleants and the reason for their failure cannot be explained.

The success of haemoglobin-MIP in inducing crystal formation in *Legionella pneumophila* - lpg2936 could be explained by its inherent characteristics as a protein memory cavity template. Thus, haemoglobin-MIP crystallises its cognate protein haemoglobin (Selligren, 2000; Arshady and Mosbach, 1981) and proteins of similar (size-compatible) molecular sizes to haemoglobin at 64.5kDa (Saridakis *et al.*, 2011; Khurshid *et al.*, 2015). *Legionella pneumophila* - lpg2936 with a molecular weight of 54 kDa could explain its successful crystallisation by haemoglobin-MIP at metastable conditions. Theoretically, the deeper into the metastable zone (i.e. the lowest level of supersaturation) the nuclei find themselves, the slower their growth, leading to improved crystal quality (Khurshid *et al.*, 2015).

Govada *et al.* (2016) reported that GO (the non PEGylated version) have been relatively successful in protein crystallisation with two proteins lysozyme and trypsin. In the case of lysozyme it induced crystal formation in the metastable zone and with thaumatin it promoted crystal formation at borderline metastable conditions. Further to this, the reported average pore size of GO is 10 - 15 nm and lysozyme is nearer to 2nm in diameter and may well fall into this wider pore size. It can be postulated that this GO pore size range stabilised nuclei containing a small number of lysozyme molecules and allowed for its success. PEG GO, a functionalised version of GO has further enhanced its potency as seen in this study.

To understand just how effective the oil-on-drop method is in producing very high quality crystals, one need to look at X-ray results from oil-on-drop

crystals. The best resolution was seen with the oil-on-drop method, even better than what was achieved by the very effective haemoglobin-MIP and PEG GO (Table 4.5). Haemoglobin-MIP improved resolution by 0.42 Å over controls and PEG GO by 0.2 Å. In comparison, the oil-on-drop method's best crystal diffracted to an even higher resolution than its controls improving resolution by 0.62 Å (Fig 4.21 a).

Why does the oil-on-drop method produce better diffracting crystals than haemoglobin-MIP and PEG GO? In the case of both nucleants, the physical introduction of a foreign substance does create a disturbance in the protein solution, even if minimal. It is generally believed that even external disturbances to protein drops such as vibration can cause excess nucleation and lead to the formation of smaller crystals or to crystal imperfections (Chayen, 1997). In microbatch trials drops are preserved from physical shock, because the nuclei and the forming crystals are protected and cushioned by the oil, making trials less susceptible to disturbance (Chayen, 1997).

With nucleants trials there is no oil to protect drops from such disturbances. Also, it is possible to conject that with these nucleants, the crystallisation process was hastened whereas the oil-on-drop method was effective in slowing down this process. The oil-on-drop method led to slower and better formed crystals whereas the lack of oil protection may have allowed the introduction of impurities and imperfections with the use of heterogeneous nucleants such as haemoglobin-MIP and PEG GO. It is likely that this may be responsible for the production of slightly lower quality diffracting crystals compared to the oil-on-drop method. This could be true for both solid grain and liquid nucleants. The non-invasive nature of the oil-on-drop method may have led to producing higher quality diffracting crystals compared to haemoglobin-MIP and PEG GO. Exploratory trials on more proteins across a wide range of nucleants and different oil ratios are needed to compare the efficacy of both methods to determine which optimisation strategy would yield the best crystals for X-ray diffraction studies. There was not enough time to explore this further during this thesis.

With reference to CB ink, as seen in the previous section of the nucleant trials, no successful nucleant-induced crystallisation was observed on catalase, lysozyme and trypsin and *Legionella pneumophila* - lpg2936. Further proof that CB ink is not an effective nucleant in these proteins. In a previous study (Govada *et al.*, 2016); CB was tested as part of a bigger study testing the efficacy of a variety of functionalised and untreated carbon nanomaterials (CNMs) as nucleating agents. In that study a 5K PEG-treated CB was found to be an effective and potent nucleant. The amorphous nature of the carbon black solid material makes it a low density and intrinsically porous material and combined with PEG functionalisation increased its potency (Govada *et al.*, 2016). CB ink, in contrast, is a liquid consisting of 2-10% CB and up to 60% water without any PEG functionalisation. This seems the likely reason why the transformation of CB into an aqueous solution did not yield positive results. The transformation may have in part diluted its potency.

5.3 TARGET PROTEIN ROAB13-PEPTIDE COMPLEX

Crystals of the RoAb13-Peptide complex were produced at both 8mg/ml and 10mg/ml concentrations. The oil-on-drop method produced the best diffraction of 3.6 Å of the RoAb13-Peptide complex to date; showing an improvement on the ~4 Å achieved with conventional optimisation techniques. Resolution was improved by at least 0.4 Å, as crystals from controls diffracted to 4 Å - 4.7 Å. The RoAb13 antibody structure on its own has previously been solved from crystals which diffracted to 2.1 Å (available in the RCSB protein data bank: PDB ID code 4S2S) (Chain *et al.*, 2015). Despite several attempts, the resolution of 3.6 Å could not be improved upon. The key question remains whether the RoAb13-Peptide complex remains bound after crystallisation. Analysis of X-ray data at 3.6 Å show the presence of the antibody and a space for the peptide (unpublished work and calculations), but at this low resolution, there's not enough detail to confirm the presence or absence of the peptide. Thus, more work on this complex is required.

The following optimisation strategies were employed in attempts to obtain higher quality crystals of the complex:

(i) Applying the oil-on-drop method to slow down the crystallisation process.

As seen in the case with other proteins (trypsin, lysozyme, α -C and methyltransferase *Legionella pneumophila* - lpg2936), the efficacy of the oil-on-drop method has also been demonstrated with the Roab13-Peptide complex.

Results for the RoAb13-Peptide complex have followed the same trend: the oil-on-drop method produced fewer and larger crystals than seen in the Chayen method and controls. In this case, and based on previous experience when looking at crystal count and crystal size, there was no need to perform a statistical test on the RoAb13-Peptide complex data, as it would not return a statistical significant result. The differences in crystal count and crystal sizes found in the RoAb13-Peptide complex were not big enough (i.e. there was a small effect). There are ways to increase the statistical power of the data and possibly create a bigger effect to enable statistical testing. This would include (i) increase the sample size by increasing the number of replicates used in the trials and/or (ii) lower the statistical significance criterion (i.e. from 0.05 to say 0.1) used for determining statistical significance. Changing the statistical significance to be less conservative (say to 0.1) would possibly mean that significance may be proven and reduces it the risk of a type II error (false negative - failing to reject a false null hypothesis). But it also increases the risk of obtaining a statistically significant result (i.e. rejecting the null hypothesis) when the null hypothesis is actually true - i.e. that there is no difference. This increases the risk of a type I error (false positive) (Sedgwick, 2014).

In short, if there is an unlimited amount of protein and time available, which wasn't the case with the antibody and peptide, then increasing the sample size would be the best course of action as long as experimental error (human or systemic) can be ruled out as a cause for the perceived low statistical power in the data which led to the lesser effect in crystal count and crystal

size found in RoAb13-Peptide complex. In this case though, the main goal was to obtain a significant improvement in resolution which was achieved.

(ii) Varying the RoAb13 antibody and peptide incubation times.

Although crystals grown using the oil-on-drop method increased the resolution limit, 3.6 Å was not high enough to solve its structure. Therefore the incubation time of the peptide and antibody was tested. The incubation time of the antibody and peptide proved to be a crucial factor in crystal formation and crystal growth irrespective of the concentration of the protein at 8mg/ml and 10mg/ml. Shorter incubation periods of 1-4 hours did not improve the diffraction resolution. Since RoAb13 concentration was not a factor and to save the precious and limited sample, the lower concentration of 8mg/ml was used to test the different incubation periods beyond 4 hours.

Longer incubation periods (72 hours and 96 hours) gave the best results with crystals being produced in the shortest time frame and leading to larger crystals (Table 4.7 subsection 4.3.2.2). Surprisingly, the least effective incubation period was 144 hours as it failed to produce any well-formed crystals. The protein crystallisation process and optimisation efforts can be laborious and time consuming which may include multiple rounds of trial and error, allowing the RoAb13-Peptide complex to incubate for longer periods before crystallisation can increase the success rate of producing fewer and larger crystals in a shorter period of time. It can be postulated that allowing the RoAb13 antibody to soak with the peptide for longer periods, has allowed sufficient binding to occur between the antibody and peptide (in a 8:1 ratio) to ensure an excess of the peptide for effective binding (personal communications with collaborators).

It is also possible that maximum affinity between RoAb13 and the peptide was achieved at around 96 hours of incubation. Protein solutions are very sensitive by nature; the complex would not be immune to instability or denaturing if conditions weren't ideal, thus affecting binding, before and during crystallisation. As a start, it is imperative to use stable and pure protein solutions for crystallisation (McPherson and Gavira, 2014). Protein

stability has long been recognized as an important feature in the tendency of a macromolecule to crystallise (McPherson, 1982). Functional studies have shown the RoAb13-Peptide complex to be quite stable at shorter incubation times up from a few hours and overnight incubation (Chain *et al.*, 2015; and references therein). In this study, the assumption was made that complex stability would be maintained for longer periods of incubation. Performing additional functional studies and purification tests to prove this was beyond the scope of this thesis.

6

FUTURE WORK

In order to increase the value of the oil-on-drop method and allow for its easy and widespread adoption within the crystallographic community, further work could include but is not limited to:

- Increase the range of benchmark proteins (to include catalase, alpha-lactalbumin, haemoglobin, etc.) to be tested with the oil-on-drop method both manually and on automated systems.
- Increase the range of new target proteins from existing and possibly new collaborators. Perform X-ray analysis on the crystals produced to further validate the oil-on-drop method's efficiency.
- Focus on increasing the power of the crystallisation data for statistical analysis by setting up more replicate samples per method (oil-on-drop method; Chayen method; and controls) to diminish the possibility of non-significant results as seen on three occasions in this thesis (Fig 4.6; Fig 4.8 d and Figs 4.28-4.29).
- Explore Oryx 8 and Mosquito system improvements and adjustments (if possible) to allow for the Chayen method to be set up on these systems. This will allow for a direct comparison of the oil-on-drop method and the Chayen method in sitting drop automated trials which could not be performed during this study.
- Continue crystallisation trials and X-ray diffraction efforts on the RoAb13-Peptide complex towards solving the complex's structure by building on what has been achieved to date.

- Continue crystallisation efforts towards solving the 3-D structure of target protein alpha crustacyanin by improving on the 10 Å achieved to date. Perform SDS-page analysis on the available protein to test its purity and stability for future crystallisation trials.
- A continuation of research and development into CNMs and the use of such porous materials in protein crystallisation efforts based on previous successes (Govada *et al.*, 2016).
- Additional work is required to further explore the RoAb13 antibody and peptide incubation times especially with the oil-on-drop method to determine whether the longer incubation periods up to 96 hours correlates positively with X-ray diffraction. It can be hypothesised that as longer incubation periods of the complex have produced fewer and larger single crystals in a shorter period of time as evidenced in this thesis, that higher resolution limits can thus be achieved. This hypothesis needs testing.
- The success of the haemoglobin-MIP nucleant in producing high resolution diffraction data which was then improved on by the oil-on-drop method, allows for testing of the whole range of available MIPs (lysozyme-MIP, trypsin-MIP) to also compare them against the oil-on-drop method.
- Another approach for the use of the oil-on-drop method is by combining it with nucleant trials. Convention dictates that specific optimisation techniques and tools be used independently to be able to attribute any crystallisation effects to that technique. One possibility is to use the nucleants (CB ink, NPG, bioglass) which failed in this study at metastable conditions by setting them up at conditions where spontaneous crystal formation occurs using the oil-on-drop method. In this way the nucleants may promote quicker crystal formation which will be balanced by slowing down the process due to the oil barrier. This may result in obtaining fewer larger crystals of diffraction quality. Also, the oil-on-drop method could be used with successful nucleants such as haemoglobin-MIP and PEG GO (at

borderline metastable/nucleation zone conditions) to achieve crystals of even high resolution than what has been produced by these methods separately.

7

FUTURE PERSPECTIVES

The driving force behind protein structure determination is the elucidation and understanding of its molecular function including functional pathways. The screening of thousands of crystallisation conditions is now feasible at nanolitre volumes by high throughput processes. With the focus on structural genomics, proteomics and beyond, crystallisation optimisation techniques based on the crystallisation principles, as presented in this thesis, has sadly been neglected by many researchers. It was hoped that large-scale screening would be sufficient to produce the desired results. However, as shown in this thesis, it is vital to continuously find ways to simplify, optimise, automate and miniaturise crystallisation techniques and methods in order to cope with the vast number of 'leads' resulting from protein screening procedures (Chayen, 2003).

On the other hand, an emerging field in macromolecular X-ray crystallography known as serial femtosecond crystallography (SFX) coupled with X-ray free electron lasers (XFELs) has been gaining popularity amongst crystallographers in recent times. These evolving technologies avoid the need for large, diffracting crystals which is a major crystallographic constraint and covered in this thesis. SFX uses femtosecond micro-focused serial X-ray pulses directed at sample streams of micro and nano-crystals. The datasets which are collected from this are made up of a collection of thousands of random diffraction image snapshots of these tiny crystals captured as they pass through the bursts of X-ray pulses aimed at the sample stream. Thus, 3-D structures can now be solved by piecing together all these individual snapshots (with its associated Bragg diffractions) (Chapman *et al.*, 2011; Boutet *et al.*, 2012; Martin-Garcia *et al.*, 2016).

A major drawback in using SFX XFEL is that it comes at a very high cost compared to traditional third generation synchrotron sources such as Diamond Light Source's macromolecular (MX) beam-lines. There are only a handful of XFEL sources, currently only 5 facilities that produce X-rays at short wavelengths called hard X-rays. This has led to the high costs, high demand and oversubscription with many research groups unable to secure beam-time (Grünbein *et al.*, 2018). In addition, given that there is short gaps between X-ray pulses hitting the micro-crystal sample stream; there is a concern of sample damage caused by shock waves created by the first XFEL pulse and causing subsequent pulses not being effective, leading to few or no image collection (in moments where the sample stream is disturbed) (Grünbein *et al.*, 2018).

In the short to medium term, the use of XFEL sources will remain a luxury until availability and costs can allow for general use. Also, experimental set up for SFX is very challenging and requires: (a) a very large sample volume; (b) the optimisation of crystal density and crystal size homogeneity; (c) the optimisation and analysis of the crystal diffraction datasets (Martin-Garcia *et al.*, 2016). As such, these technological advances are important in moving the field of crystallography forward but they certainly do not replace other methods of determining structures of important molecules. There still remains a critical need to produce large crystals of high quality for single crystal X-ray diffraction studies, which optimisation strategies such as the novel oil-on-drop method, delivers.

Other competing technologies such as NMR, EM and cryo-EM have remained slow in gaining ground on X-ray crystallography. Over the last decade a limited amount of structures have been solved at sub-3 Å resolutions by these technologies (Fischer *et al.*, 2015). When searching the PDB for deposited structures by experimental method, the following results were returned: X-ray crystallography (132 678); NMR (12 425); EM (2 794); neutron (151); hybrid (138); and other (32). It is clear from the search results of all the deposited structures in the PDB (<https://www.rcsb.org/pdb/search/advSearch.do?search=new>) that X-ray crystallography still account for ~89% of all structures. From the ~89% of X-

ray derived structures, the best deposited resolution was 0.48 Å compared to the best resolution of 1.8 Å for cryo-EM. Cryo-EM is gaining momentum and has received media attention when three researchers who were joint recipients of the 2017 Nobel Prize in Chemistry for their pioneering work in developing cryo-EM in the 1970s, 1980s and 1990s (Henderson and Unwin, 1975; Frank *et al.*, 1978; Lepault *et al.*, 1983; Adrian *et al.*, 1984; Radermacher *et al.*, 1987; Henderson *et al.*, 1990; Penczek *et al.*, 1994; Frank *et al.*, 1995). This laid the foundation in recent years for what many are calling the “cryo-EM resolution revolution”. That said, cryo-EM still has some way to go to consistently match atomic-level resolutions which is possible with X-ray crystallography.

In conclusion, there is not a single miracle method. They are all complementary methods that can be used in parallel, e.g. EM and cryo-EM solves big complexes that cannot always be done by X-ray. NMR can be used when no crystals can be formed even after a lot of effort and SFX XFEL is useful when single large crystals cannot be produced.

8

BIBLIOGRAPHY

Adrian, M., Dubochet, J., Lepault, J., McDowell, A. (1984). Cryo-electron microscopy of viruses. *Nature*, **308**, 32–36.

Arshady, R. and Mosbach, K. (1981). Synthesis of substrate-selective polymers by host-guest polymerization. *Makromolekul Chem*, **182**, 687–692.

Asanithi, P., Saridakis, E., Govada, L., Jurewicz, I., Brunner, E., Ponnusamy, R., Cleaver, J., Dalton, A., Chayen, N. & Sear, R. (2009). Carbon-Nanotube-Based Materials for Protein Crystallization. *ACS Appl. Mater. Interfaces*, **1** (6), 1203-1210.

Ataka, M. (1993). Protein crystal growth: an approach based on phase diagram determination. *Phase Transitions*, **45** (2-3), 205–219.

Balasubramanian, K. and Burghard, M. (2005). Chemically functionalised carbon nanotubes. *Small*, **1** (2), 180-92.

Barmania, F. and Pepper, M. (2013). C-C chemokine receptor type five (CCR5): An emerging target for the control of HIV infection. *Applied & Translational Genomics*, (2), 3–16.

Benvenuti, M. and Mangani, S. (2007). Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nature Protocols*, **2**, 1633–1651.

Bergfors, T. (2003). Seeds to Crystals. *J Struct Biol.*, **142**, 66-76.

Bergfors, T. (2009). *Protein crystallization* (International University Line, La Jolla, Calif., 2009), IUL biotechnology series. Edition: Second. Editor: Terese Bergfors, ISBN: 978-0-9720774-4-6).

Blakeley, M., Hasnain, S. and Antonyuk, S. (2015). Sub-atomic resolution X-ray crystallography and neutron crystallography: promise, challenges and potential. *IUCrJ*. **2** (Pt 4), 464–474.

Blow, D., Chayen, N., Lloyd, L., Saridakis, E. (1994). Control of Nucleation of Protein Crystals. *Prot Sci*. **10**, 1638-1643.

Boutet, S., Lomb, L., Williams, G., Barends, T., Aquila, A., Doak R.,... Schlichting, I. (2012). High-resolution protein structure determination by serial femtosecond crystallography. *Science*, **337** (6092), 362-364.

Britton, G., Armitt, G. and Lau, S. (1982). Carotenoproteins, Carotenoid Chemistry and Biochemistry, edited by G. Britton and T. W. Goodwin, pp. 1237–1253. *Oxford: Pergamon Press*.

Bruggemann, O. (2002). Molecularly imprinted materials—Receptors more durable than nature can provide. In: Freitag R, editor. Modern Advances in Chromatography: Advances in Biochemical Engineering Biotechnology. **76**. Berlin: Springer, pp. 127–163.

Chackerian, B., Briglio, L., Albert, P., Lowy, D., Schiller, J. (2004). Induction of Autoantibodies to CCR5 in Macaques and Subsequent Effects upon Challenge with an R5-Tropic Simian/Human Immunodeficiency Virus. *Journal of Virology*, **78** (8), 4037-4047.

Chain, B., Noursadeghi, M., Gardener, M., Tsang, J., Wright, E. (2008). HIV blocking antibodies following immunisation with chimaeric peptides coding a short N-terminal sequence of the CCR5 receptor. *Vaccine*, **26**, 5752–5759.

Chain, B., Arnold, J., Akthar, S., Brandt, M., Davis, D., Noursadeghi, M., Lapp, T., Ji, C., Sankuratri, S., Zhang, Y., Govada, L., Saridakis, E. & Chayen, N. (2015). A Linear Epitope in the N-Terminal Domain of CCR5 and Its Interaction with Antibody. *PLoS ONE*, **10**(6), <https://doi.org/10.1371/journal.pone.0128381>

Chapman, H., Fromme, P., Barty, A., White, T., Kirian, R., Aquila, A.,... Spence, J. (2011). Femtosecond X-ray protein nanocrystallography. *Nature*, **470**, 73–77.

Chayen, N., Shaw Stewart, P., Maeder, D., Blow, D. (1990). An automated system for microbatch protein crystallisation and screening. *J App Cryst.*, **23**, 297-302.

Chayen, N., Radcliffe, J. and Blow, D. (1993). Control of Nucleation in the Crystallization of Lysozyme. *Protein Science*, **2**, 113-118.

Chayen, N., Shaw Stewart, P., Baldock, P. (1994). New developments of the IMPAX small volume automated crystallization system. *Acta Cryst.*, **D50**, 456-458.

Chayen, N. (1996). A novel technique for containerless protein crystallization. *Protein Engineering*. **9**, 927-929.

Chayen, N. (1997). The role of oil in macromolecular crystallization. *Structure*, **5**, 1269-1274.

Chayen, N. and Helliwell, J. (1998). Protein crystallography: the human genome in 3-D. *Physics World*, **11**, 43-48.

Chayen, N. (1999). Recent advances in methodology for the crystallization of biological macromolecules. *Journal of Crystal Growth*, **199**, 649-655.

Chayen, N., Cianci, M., Olczak, A., Raftery, J., Rizkallah, P. J., Zagalsky, P. & Helliwell, J. (2000). *Acta Cryst.*, **D56**, 1064-1066.

Chayen, N. (2001). Porous silicon: an effective nucleation-inducing material for protein crystallization. 1 Edited by R. Huber. *Journal of Molecular Biolog.*, **312** (4), 591-595.

Chayen, N., Saridakis, E., El-Bahar, R., Nemirovsky, Y. (2001). Porous silicon: an effective nucleation-inducing material for protein crystallization. *J Mol Biol*, **312**, 591-595.

Chayen, N. and Saridakis, E. (2002). Protein crystallization for genomics: towards high-throughput optimization techniques. *Acta Cryst.*, **D58**, 921-927.

Chayen, N. (2003). Protein crystallization for genomics: throughput versus output. *J Funct Struct Gen.*, **4**, 115-120.

Chayen, N. (2004). Turning protein crystallisation from an art into a science. *Curr Opin Struct Biol*, **14**, 577-583.

Chayen, N. (2005). Methods for separating nucleation and growth in protein crystallisation. *Prog Biophys Mol Biol.*, **88**, 329-337.

Chayen, N., Saridakis, E. and Sear, R. (2006). Experiment and theory for heterogeneous nucleation of protein crystals in a porous medium. *Proc Natl Acad Sci U S A*, **103**, 597-601.

Chayen N. and Saridakis, E. (2008). Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods*, **5**, 147-153, ISSN: 1548-7091.

Chirgadze, D. (2001). *Protein Crystallisation in Action*. University of Cambridge, © 1999-2001.

Cianci, M., Rizkallah, P., Olczak, A., Raftery, J., Chayen, N., Zagalsky, P. & Helliwell, J. (2002). Structure of lobster apocrustacyanin A1 using softer X-rays. *Acta Cryst.*, **D57**, 1219-1229.

Cianci, M., Rizkallah, P., Olczak, A., Raftery, J., Chayen, N., Zagalsky, P. & Helliwell, J. (2002). The molecular basis of the coloration mechanism in lobster shell: β -Crustacyanin at 3.2-Å resolution. *Proc. Natl. Acad. Sci. USA*, **99**, 9795-9800.

Cudney, B. and Patel, S. (1994). Crystallization as a tool for bioseparation. *Am Biotechnol Lab*, **12** (7), 42.

Curtis, R., Ulrich, J., Montaser, A., Prausnitz, J., Blanch, H. (2002). Protein-protein interactions in concentrated electrolyte solutions. *Biotechnol Bioeng.*, **79** (4), 367-380.

D'Arcy, A., Elmore, C., Stihle, M., Johnston, J. (1996). A novel approach to crystallising proteins under oil. *Journal of Crystal Growth*, **168** (s1-4), 175-180.

D'Arcy, A., Mac Sweeney, A. and Haber, A. (2003). Using natural seeding material to generate nucleation in protein crystallization experiments. *Acta Crystallogr D Biol Crystallogr.*, (59), 1343-1346.

Decatur, W. and Fournier, M. (2002). rRNA modifications and ribosome function. *Trends in Biochemical Sciences*, **27** (7), 344-351.

Del Campo, M., Ofengand, J. and Malhotra, A. (2004). Crystal structure of the catalytic domain of RluD, the only rRNA pseudouridine synthase required for normal growth of *Escherichia coli*. *RNA*, **10** (2), 231-239.

Douglas Instruments. <https://www.douglas.co.uk/>

Dellisanti, C., Spinelli, S., Cambillau, C., Findlay, J., Zagalsky, P., Finet, S. & Receveur-Bréchet, V. (2003). Quaternary structure of alpha-crustacyanin from lobster as seen by small-angle X-ray scattering. *FEBS Lett.*, **544** (1-3), 189-93.

Ducruix, A. and Giegé, R. (Ed.). (1992). *Crystallisation of Nucleic Acids and Proteins, A Practical Approach*. Oxford University Press, Oxford, UK.

Ducruix, A. and Giegé, R. (1999). Crystallization of nucleic acids and proteins: a practical approach, *Oxford University Press*, Oxford.

Fischer, N., Neumann, P., Konevega, A., Bock, L., Ficner, R., Rodnina, M. & Stark, H. (2015). Structure of the *E. coli* ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature*, **520**, 567.

Feher, G. and Kam, Z. (1985). Nucleation and growth of protein crystals: general principles and assays. *Methods Enzymol*, **114**, 77-112.

Ferrari, M., Folli, C., Pincolini, E., McClintock, T., Rössle, M., Berni, R., & Cianci, M. (2012). Structural characterization of recombinant crustacyanin subunits from the lobster *Homarus americanus*. *Acta Cryst*, **F68**, 846-853.

Frank, J., Goldfarb, W., Eisenberg, D., Baker, T. (1978). Reconstruction of glutamine synthetase using computer averaging. *Ultramicroscopy*, **3**, 283-290.

Frank, J., Zhu, J., Penczek, P., Li, Y., Srivastava, S., Verschoor, A., Radermacher, M., Grassucci, R., Lata, R. & Agrawal, R. (1995). A model of protein synthesis based on cryo-electron microscopy of the E. coli ribosome. *Nature*, **376**, 441–444.

Frank, J. (2002). Single-Particle Imaging of Macromolecules by Cryo-Electron Microscopy. *Annual Review of Biophysics and Biomolecular Structure*, **31** (1), 303-319.

García-Ruiz, M. (2003). Counterdiffusion methods for macromolecular crystallization. *Methods Enzymol.*, **368**, 130-54.

Govada, L., Leese, H., Saridakis, E., Kassen, S., Chain, B., Khurshid, S., Menzel, R., Hu, S., Shaffer, M., & Chayen, N. (2016). Exploring Carbon Nanomaterial Diversity for Nucleation of Protein Crystals. *Scientific Reports*, **6**, ISSN: 2045-2322.

Grünbein, M., Bielecki, J., Gorel, A., Stricker, M., Bean, R., Cammarata, M., ... Schlichting, I. (2018). Megahertz data collection from protein microcrystals at an X-ray free-electron laser. *Nature Communications*, **9**: 3487.

Habash, J., Helliwell, J., Raftery, J., Cianci, M., Rizkallah, P., Chayen, N., Nneji, G. & Zagalsky, P. (2004). The structure and refinement of apocrustacyanin C2 to 1.3 Å resolution and the search for differences between this protein and the homologous apoproteins A1 and C1. *Acta Cryst.*, **D60**, 493–498.

Hamilton, B., Ha, J-M., Hillmyer, M., Ward, M. (2012). Manipulating Crystal Growth and Polymorphism by Confinement in Nanoscale Crystallization Chambers. *Accounts of Chemical Research*, **45** (3), 414–423.

Hampton Research, (2001). Crystallization Research Tools, 2001 catalogue.

Hansen, D. (2007). Recent developments in the molecular imprinting of proteins. *Biomaterials*. **28**, 4178–4191.

Henderson, R. and Unwin, P. (1975). Three-dimensional model of purple membrane obtained by electron microscopy. *Nature*, **257**, 28-31.

Henderson, R., Baldwin, J., Ceska, T., Zemlin, F., Beckmann, E., Downing, K. (1990). Model for the Structure of Bacteriorhodopsin Based on High-resolution Electron Cryo-microscopy. *J. Mol. Biol.*, **213**, 899-929.

Hillberg, A. and Tabrizian, M. (2008). Biomolecule imprinting: Developments in mimicking dynamic natural recognition systems. *IRBM*, **29**, 89–104.

Jancarik, J. and Kim, S. (1991). Sparse Matrix Sampling: A Screening Method for Crystallization of Proteins. *Journal of Applied Crystallography*, **24**, 409-411.

Ji, C., Brandt, M., Dioszegi, M., Jekle, A., Schwoerer, S., Challand, S., Zhang, J., Chen, Y., Zautke, L., Achhammer, G., Baehner, M., Kroetz, S., Heilek-Snyder, G., Schumacher, R., Cammack, N. & Sankuratri, S. (2007). Novel CCR5 monoclonal antibodies with potent and broad-spectrum anti-HIV activities. *Antiviral Research*, **74**, 125–137.

Kallio, J., Hakulinen, N., Kallio, J., Niemi, M., Kärkkäinen, S., Rouvinen, J. (2009). The Contribution of Polystyrene Nanospheres towards the Crystallization of Proteins. *Plos One.*, **4** (1), e4198.

Kertis, F., Khurshid, S., Okman, O., Kysar, J., Govada, L., Chayen, N. & Erlebacher, J. (2012). Heterogeneous nucleation of protein crystals using nanoporous gold nucleants. *J. Mater. Chem.*, **22**, 21928-21934.

Khurshid, S., Saridakis, E., Govada, L., Chayen, N. (2014). Porous nucleating agents for protein crystallization. *Nature Protoc.*, **9**, 1621–1633.

Khurshid, S., Govada, L., EL-Sharif, H., Reddy, S., Chayen, N. (2015). Automating the application of smart materials for protein crystallization. *Acta Crystallographica Section D: Biological Crystallography.*, **71**(Pt 3), 534–540. <http://doi.org/10.1107/S1399004714027643>

Kim, H., Abdala, A. and Macosko, C. (2010). Graphene/Polymer Nanocomposites. *Macromolecules*, **43** (16), 6515-6530.

Kondru, R., Zhang, J., Ji, C., Mirzadegan, T., Rotstein, D., Sankuratri, S. & Dioszegi, M. (2008). Molecular interactions of CCR5 with major classes of small-molecule anti-HIV CCR5 antagonists. *Mol Pharmacol.*, **73** (3), 789-800.

Krishnan, V. and Rupp, B. (2012). *Macromolecular Structure Determination: Comparison of X-ray Crystallography and NMR Spectroscopy*. In: eLS. John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0002716.pub2.

Lepault, J., Booy, F. and Dubochet, J. (1983). Electron microscopy of frozen biological suspensions. *J Microsc*, **129**, 89–102.

Liao, J-L., Wang, Y. and Hjerten, S. (1996). A novel support with artificially created recognition for the selective removal of proteins and for affinity chromatography. *Chromatographia.*, **42**, 259–262.

Odom, T., Huang, J-L. and Lieber, C. (2002). Single-Walled Carbon Nanotubes: From Fundamental Studies to New Device Concepts. *Ann. N.Y. Acad. Sci.*, **960**, 203–215.

Martin-Garcia, J., Conrad, C., Coe, J., Roy-Chowdhury, S., Fromme, P. (2016). Serial femtosecond crystallography: A revolution in structural biology. *Arch Biochem Biophys.*, **602**, 32-47.

Menzel, R., Tran, M., Menner, A., Kay, C., Bismarck, A., Shaffer, M. (2010). A versatile, solvent-free methodology for the functionalisation of carbon nanotubes. *Chem. Sci.* (5), 603-608.

McCoy, A. (2005). *Phase Diagrams*. University of Cambridge. 10 April 2019. (<<http://www-structmed.cimr.cam.ac.uk/Course/Crystals/Theory/phases.html>>)

McPherson, A. (1982). Preparation and analysis of protein crystals. *Kreiger publishing*, Florida.

McPherson, A. and Shlichta, P. (1988). Heterogeneous and epitaxial nucleation of protein crystals on mineral surfaces. *Science*, **239**, 385-387.

McPherson, A. (1995). Increasing the Size of Microcrystals by Fine Sampling of Ph Limits.

McPherson, A. (1999). Crystallization of biological macromolecules, Cold Spring Harbor Laboratory Press, *Cold Spring Harbor*, NY.

McPherson, A. and Cudney, B. (2014). Optimization of crystallization conditions for biological macromolecules. *Acta crystallographica*. Section F, Structural biology communications, **70** (Pt 11), 1445-67.

McPherson, A. and Gavira, J. (2014). Introduction to protein crystallization. *Acta Cryst*, Section **F70**, 2-20.

Molecular Dimensions. <https://www.moleculardimensions.com/>

Nilsson, K., Sakaguchi, K., Gemeiner, P., Mosbach, K. (1995). Molecular imprinting of acetylated carbohydrate derivatives into methacrylic polymers. *J Chromatogr A.*, **707**, 199-203.

Nneji, G. and Chayen, N. (2004). A crystallization plate for controlling evaporation in hanging drops. *J. Appl. Cryst.*, **37**, 502-503.

Page, A. and Sear, R. (2006). Heterogeneous Nucleation in and out of Pores. *Phys Rev Lett.*, **97**, 065701.

Penczek, P., Grassucci, R. and Frank, J. (1994). The ribosome at improved resolution: New techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy*, **53**, 251-270.

Pinotsis, N. and Waksman, G. (2017). Crystal structure of the Legionella pneumophila Lpg2936 in complex with the cofactor S-adenosyl-L-methionine reveals novel insights into the mechanism of RsmE family methyltransferases. *Protein Sci.*, **26** (12), 2381-2391.

Polson, A., Potgieter, G., Largier, J., Mears, G., Joubert, F. (1964). The Fractionation of Protein Mixtures by Linear Polymers of High Molecular Weight. *Biochim Biophys Acta*, **82**, 463-75.

Pusey, M. and Naumann, R. (1986). Growth-Kinetics of Tetragonal Lysozyme Crystals.

Pusey, M., Liu, Z., Tempel, W., Praissman, J., Lin, D., Wang, B., Gavira, J. & Ng, J. (2005). Life in the fast lane for protein crystallization and X-ray crystallography. *Progress in Biophysics and Molecular Biology*, **88** (3), 359-386.

Qiagen. <https://www.qiagen.com/gb/shop/sample-technologies/protein/crystallization/easyxtal-15-well-tools/>

Radermacher, M., Wagenknecht, T., Verschoor, A., Frank, J. (1987). Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of Escherichia coli. *J Microsc*, **146**, 113-136.

Rhys, N., Wang, M., Jowitt, T., Helliwell, J., Grossmann, J., Baldock, C. (2011). Deriving the ultrastructure of a-crustacyanin using lower-resolution structural and biophysical methods. *J. Synchrotron Rad.*, **18**, 79-83.

Rochefort, A. (2013). *Multi-wall carbon nanotubes.* © (Center for Research on Computation and its Applications (CERCA)).

Rolando, M., Gomez-Valero, L. and Buchrieser, C. (2015). Bacterial remodelling of the host epigenome: functional role and evolution of effectors methylating host histones. *Cell. Microbiol.*, **17**, 1098-1107.

Saridakis, E., Stewart, P., Lloyd, L., Blow, D. (1994). Phase diagram and dilution experiments in the crystallization of carboxypeptidase G₂. *Acta Cryst D.*, **50**, 293-297.

Saridakis, E. and Chayen, N. (2003). Systematic improvement of protein crystals by determining the supersolubility curves of phase diagrams. *Biophys J*, **84**, 1218-22.

Saridakis, E. and Chayen, N. (2009). Towards a 'Universal' Nucleant for Protein Crystallization. *Trends in Biotechnology*, **27**, 99-106.

Saridakis, E., Khurshid, S., Govada, L., Phan, Q., Hawkins, D., Crichton, G. & Chayen, N. (2011). Protein crystallization facilitated by molecularly imprinted polymers. *Proceedings of the National Academy of Sciences of the United States of America*, **108** (27), 11081-11086. <http://doi.org/10.1073/pnas.1016539108>.

Schindelin, J., Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T.,... Cardona, A. (2012). "Fiji: an open-source platform for

biological-image analysis". *Nature methods*, **9** (7), 676-682. PMID 22743772. <http://doi:10.1038/nmeth.2019>.

Schnabel, R. and Langer, P. (1991). Controlled-pore glass as a stationary phase in chromatography. *Journal of Chromatography*, (544), 137-146.

Schuhmacher, M., Rolando, M., Bröhm, A., Weirich, S., Kudithipudi, S., Buchrieser, C. & Jeltsch, A. (2018). The Legionella pneumophila Methyltransferase RomA Methylates Also Non-histone Proteins during Infection. *J Mol Biol.*, **22**; **430**(13), 1912-1925. doi: 10.1016/j.jmb.2018.04.032.

Sedgwick, P. (2014). Pitfalls of statistical hypothesis testing: type I and type II errors. *BMJ*, 349:g4287. <https://doi.org/10.1136/bmj.g4287>.

Sellergren, B. (2000). Imprinted polymers with memory for small molecules, proteins or crystals. *Angew Chem Int Ed Engl.*, **39**, 1031-1037.

Sellergren, B., Ekberg, B. and Mosbach, K. (1985). Molecular imprinting of amino acid derivatives in macroporous polymers: Demonstration of substrate- and enantio-selectivity by chromatographic resolution of racemic mixtures of amino acid derivatives. *J Chromatogr A.*, **347**, 1-10.

Shah, V., Parambil, J., Williams, D., Hinder, S., Heng, J. (2015). Preparation and characterization of 3D nanotemplates for protein crystallisation. *Powder Technology*, (282), 10-18.

Shi, H., Tsai, W., Garrison, M., Ferrari, S., Ratner, B. (1999). Molecular imprinted polymers for biorecognition of bioagents. *Nature*, **398**, 593-597.

Singh, V., Joung, D., Zhai, L., Das, S., Khondaker, S., Seal, S. (2011). Graphene based materials: Past, present and future. *Progress in Materials Science*, **56** (8), 1178-1271.

Suedee, R., Songkram, C., Petmoreekul, A., Sangkunakup, S., Sankasa, S., Kongyarnit, N. (1998). Thin-layer chromatography using synthetic polymers imprinted with quinine as chiral stationary phase. *J Planar Chromatogr-Mod TLC.*, **11**, 272-276.

Sugahara, M., Asada, Y., Morikawa, Y., Kageyama, Y., Kunishima, N. (2008). Nucleant-mediated protein crystallization with the application of microporous synthetic zeolites. *Acta Crystallographica. Section D.*, **64** (Pt 6), 686-695.

Tran, M., Tridech, C., Alfrey, A, Bismarck, A., Shaffer, M. (2007). Thermal oxidative cutting of multi-walled carbon nanotubes. *Carbon.*, **45** (12), 2341-2350.

TTP Labtech. <https://www.ttplabtech.com/products/liquid-handling/mosquito-lcp/>

Vert, M. and Domurado, D. (2000). Poly(ethylene glycol): protein-repulsive or albumin-compatible? *J Biomater Sci Polym Ed.*, **11** (12), 1307-17.

Wald, G., Nathanson, N., Jencks, W., Tarr, E. (1948). Crustacyanin, the blue carotenoid-protein of the lobster shell. *Biol Bull.*, **95** (2), 249.

Wang, H., Feng, L., Webb, G. I., Kurgan, L., Song, J., Lin, D. (2017). Critical evaluation of bioinformatics tools for the prediction of protein crystallization propensity. *Briefings in bioinformatics*, **19** (5), 838-852.

Wu, L. and KewalRamani, V. (2006). Dendritic-cell interactions with HIV: infection and viral dissemination. *Nat Rev Immunol.*, **6** (11), 859-68.

Zagalsky, P. and Jones, R. (1982). Quaternary structures of the astaxanthin-proteins of *Veella veella*, and of α -crustacyanin of lobster carapace, as revealed in electron microscopy. *Comparative biochemistry and physiology. B, Comparative biochemistry*, **71** (2), 237-242.

Zagalsky, P. (1985). Invertebrate Carotenoproteins. *Methods Enzymol.*, **111**, 216-247.

Zhang, H., Wan, H., Gao Z., Wei, Y., Wang, W., Liu, G., Shtykova, E., Xu, J. & Dong, Y. (2012). Insights into the Catalytic Mechanism of 16S rRNA Methyltransferase RsmE (m³U1498) from Crystal and Solution Structures. *Journal of Molecular Biology*, **423** (4), 576-589.