X-RAY COMPATIBLE MICROFLUIDIC PLATFORMS
FOR STUDYING PROTEIN STRUCTURE AND FUNCTION

BY
SUDIPTO GUHA

DISSERTATION
Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Doctoral Committee:
Professor Paul J.A. Kenis, Chair
Professor Robert B. Gennis
Professor Satish K. Nair
Assistant Professor Charles M. Schroeder
Abstract

If you can look into the seeds of time,
And say which grain will grow and which will not.

Knowledge of the three dimensional structure of a protein provides insight into its mechanism, as well as into potential sites for drug targeting. Often these insights provide functional information about an unknown protein, lead to a deeper understanding of how damaged proteins cause disease, and potentially reveal a path to new drug treatments. The most common method to determine protein structure and function is to crystallize them and probe them with X-rays to determine their structure. Protein crystallization still remains an art rather than an exact science, and it is extremely difficult to predict the conditions that will ultimately result in crystals resulting in difficulties in structure determination. Recent efforts in structural biology have helped to lessen bottlenecks using robotics and/or microfluidics, leading to improved, and often automated, high throughput methodologies for screening, crystallization, and X-ray diffraction analysis of novel protein targets. Even with these improvements, the harvesting and mounting of protein crystals for X-ray analysis remains a cumbersome, manual step in the structural biology pipeline that especially hampers collection of data from small/fragile crystals.

Obtaining high-quality protein crystals is the main obstacle for structure elucidation because crystallization is a complex, multi-parametric process that involves setting up thousands of crystallization trials to screen a vast chemical space. Current data collection strategies involve harvesting a single crystal from the crystallization droplet in which it was grown, a process that often damages the crystal, followed by cryocooling the crystal, before subjecting it to monochromatic X-rays to elucidate the structure from X-ray data. For de novo structure determination, phase information also needs to be collected to convert diffraction data (that records intensities) to structural information. Phase information is most commonly obtained experimentally by measuring anomalous diffraction signal. The collection of high quality phasing data is dependent on having diffraction data with a high signal-to-noise ratio while minimizing radiation-induced damage and crystal non-isomorphism.
The microfluidic crystallization platform developed in this work serves to overcome issues with current methods in protein structure determination. Specifically, it allows for screening (crystallization) conditions together with on-chip X-ray analysis of crystals formed, and can be further used for relevant functional studies of proteins via time resolved crystallography. Furthermore, this is the first example of a microfluidic chip being used for room temperature phasing from diffraction data collected and merged from multiple crystals.

Microfluidics offers more precise control over the composition and the kinetics of a crystallization trial than what is possible using conventional methods such as vapor diffusion and microbatch. The ability to control fluid flow and transport at such a small preparative scale makes it an ideal platform for crystallization since microfluidic devices use minimal amounts of protein solution, which are often times not available in large quantities for medically relevant human targets. The use of appropriate polymers in chip fabrication enables on-chip analysis thus eliminating the need for crystal harvesting prior to analysis. Identification of these materials and development of protocols to fabricate X-ray compatible microfluidic platforms is discussed in Chapter 2. Proof-of-principle studies using model systems that validate the microfluidic chip and an alternate data collection strategy for merging datasets from multiple crystals is covered in Chapter 3. Screening of conditions and de novo structure determination via on-chip collection of phase information is discussed in further detail in Chapter 4. In Chapter 5, the application of microfluidic platforms for use in seeding studies is discussed. Seeding is a method to improve chances of obtaining high quality crystals by separating the nucleation and growth phases in a crystallization trial, and this chapter focusses on how it was achieved using a microfluidic platform.

Laue crystallography allows us to look at protein molecules in motion, helping us understand the complex pathways and structural changes that accompany protein function. Chapter 6 describes how X-ray transparent microfluidic chips can be used for collecting time resolved Laue data thus enabling functional studies on protein targets that would be cumbersome using traditional methods. The platform developed here can also be used to enable structural studies that shed light on the change in structure during protein function, ranging from proteins that respond to stimuli such as light, pH, or the presence of ligands, to those that respond to temperature changes, thus allowing an unprecedented opportunity for a range of previously elusive biological studies.
Acknowledgements

I am extremely grateful to a number of people who have supported and guided me throughout my doctoral program and I would like to acknowledge them here.

First and foremost, I would like to thank my PhD advisor, Dr. Paul Kenis, for his constant support and encouragement throughout my stay here in the doctoral program at the University of Illinois. He has always encouraged me to work and think independently and I am grateful to have him as my PhD advisor. Besides a focus on research, he has also supported me tremendously in other endeavors that I have undertaken during my time as a graduate student. The skillsets I have acquired under his guidance have made my PhD experience a very enriching one.

We spend the maximum time during our PhD with our group members and it is from the constant discussions amongst our fellow researchers that we learn the most. I would like to thank all the members in the Kenis group, both past and present for their constant support and guidance. In particular, I would like to thank Dr. Ash Pawate for his constant encouragement and expert opinion on various aspects of my work. I would also like to thank Dr. Sarah Perry who started the crystallization project and helped a lot with training and guidance in my initial years here. Amongst group members past and present, I would specially like to thank Dr. Amit Desai, Dr. Daria Khvostichenko, Dr. Mike Thorson, Dr. Joshua Tice, Sachit Goyal, Dr. Adam Hollinger, Dr. Matt Naughton and Vivek Kumar for their invaluable input and thoughtful discussion throughout my time here at graduate school. I have also had two fantastic undergraduates, Mike Varel and Amrit Bhaskarla who have worked with me on various projects. Mike Varel in particular has contributed a lot to the seeding project.

My research has been a highly collaborative effort and would not have been possible without the support of many people in various labs and other universities. I would like to thank Dr. Satish Nair and Dr. Vinayak Agarwal for supplying PhnA protein as well as assisting in data collection. I also wish to acknowledge Dr. Bob Gennis and all his group members for their constant help and providing membrane proteins for my experiments. I am deeply grateful to the
staff at LS-CAT and BioCARS, without whom data collection at Argonne National Lab would have been impossible. In particular I would like to thank Dr. Joe Brunzelle, Dr. Spencer Anderson, Dr. Elena Kondrashkina and Dr. Keith Brister at LS-CAT for their constant help and support in the monochromatic data collection. I would like to acknowledge Dr. Zhong Ren and Dr. Vukica Srajer at BioCARS for their guidance and support in Laue data collection. The seeding efforts were aided by proteins supplied by Dr. Marius Schmidt from the University of Wisconsin, Milwaukee for which I am grateful.

My project was funded by NIH R01 grant (R01 GM086727). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor for the support of this research program (Grant 085P1000817).

Last but not the least, I would like to thank my family, particularly my parents and brother, for their constant support and encouragement not only during graduate school but throughout my life.
Table of Contents

Chapter 1 Introduction ........................................................................................................................................... 1
  1.1 Protein crystallization ....................................................................................................................................... 3
  1.2 Conventional methods for protein crystallization ............................................................................................... 8
    1.2.1 Microbatch crystallization ............................................................................................................................. 12
    1.2.2 Vapor diffusion .................................................................................................................................................. 13
    1.2.3 Free interface diffusion ..................................................................................................................................... 14
    1.2.4 Dialysis ............................................................................................................................................................ 15
    1.2.5 Seeding ............................................................................................................................................................ 15
    1.2.6 The “ideal” crystallization platform ........................................................................................................... 16
  1.3 Microfluidic platforms for protein crystallization ............................................................................................... 19
  1.4 X-ray compatible microfluidic platforms for protein crystallization ................................................................. 23
  1.5 Conclusions ........................................................................................................................................................ 27
  1.6 References ........................................................................................................................................................ 29

Chapter 2 Fabrication of X-ray compatible microfluidic platforms for protein crystallization .................................................... 35
  2.1 Introduction ........................................................................................................................................................ 35
  2.2 Materials and methods ....................................................................................................................................... 37
    2.2.1 Fabrication of microfluidic platforms ............................................................................................................ 37
  2.3 Results and discussion ......................................................................................................................................... 38
    2.3.1 X-ray compatibility and materials characterization ......................................................................................... 38
    2.3.2 Design and operation of microfluidic devices ............................................................................................... 41
    2.3.3 Array chip operation .......................................................................................................................................... 47
    2.3.4 Single-well chip design, operation, and proof-of-concept experiments ......................................................... 48
    2.3.5 Interfacing with beamline infrastructure and data collection strategy ......................................................... 50
    2.3.6 Screening and on-chip protein structure determination .................................................................................. 50
  2.4 Conclusions ........................................................................................................................................................ 52
  2.5 References ........................................................................................................................................................ 54

Chapter 3 Proof of concept studies for on-chip X-ray diffraction data collection from soluble proteins ................................................ 57
  3.1 Introduction ........................................................................................................................................................ 57
Chapter 6 Microfluidic approaches for time resolved studies of proteins .................. 112

6.1 Introduction ........................................................................................................... 112

6.2 Materials and methods ....................................................................................... 117
   6.2.1 Device fabrication and operation ................................................................. 117
   6.2.2 Protein and precipitant solutions ............................................................... 117
   6.2.3 Crystallization experiments ....................................................................... 118
   6.2.4 X-ray diffraction experiments .................................................................. 120
   6.2.5 Analysis of X-ray diffraction data ............................................................. 121

6.3 Results and discussion ....................................................................................... 122
   6.3.1 Microfluidic platforms for Laue crystallography ....................................... 122
   6.3.2 Automated data collection strategy at BioCARS ....................................... 126
   6.3.3 On-chip single shot and time resolved crystallography .............................. 127

6.4 Conclusions ......................................................................................................... 129

6.5 References .......................................................................................................... 131

Chapter 7 Concluding remarks and future work .................................................... 132

7.1 Summary ............................................................................................................. 132

7.2 Future directions ................................................................................................. 134
   7.2.1 Membrane protein crystallization via bicelles ......................................... 134
   7.2.2 Crystallization of proteins in different space groups ................................. 136
   7.2.3 Microseeding efforts for heme-copper oxidases ....................................... 136

7.3 References .......................................................................................................... 137
Chapter 1

Introduction

Proteins are a large class of complex organic molecules that are essential for life. They are fundamental components of all living cells and include many substances such as enzymes, hormones and antibodies. Proteins play a central role in biological processes and are the true workhouses of all living systems. They play an essential role in growth and repair of body tissues, material transport, cell-cell signaling, energy storage and transduction, immune response and interaction between intra and extra cellular environment; in essence they participate in virtually every process within cells. Many proteins catalyze reactions as enzymes, and are vital to metabolism. They can also have structural or mechanical functions such as proteins in the cytoskeleton, a form of scaffolding that maintains cell shape.

Proteins consist of one or more long chains of amino acids which are linked through peptide bonds and are maintained in distinct and varied 3-dimensional structures. They fold into structural units consisting of alpha helices, beta sheets and other folded shapes with structural integrity being provided by chemical bonding between these structures (Figure 1). These folded shapes and structural units are immensely important because they define the protein's function in the cell. Some protein shapes fit perfectly in cell receptors, turning chemical processes on and off, like a key in a lock; whereas others work to transport molecules (e.g., oxygen, nutrients) throughout the body (e.g., hemoglobin's shape is ideal for carrying oxygen). When proteins fail
to take on their preordained shapes, there can be serious consequences: misfolded proteins have been implicated in diseases such as Alzheimer's, mad cow, and Parkinson's, among others. Proteins can be informally divided into three main classes, which correlate with typical tertiary structures: globular (soluble) proteins, fibrous proteins, and membrane proteins. Almost all globular proteins are soluble and many are enzymes. Fibrous proteins are often structural, such as collagen, the major component of connective tissue, or keratin, the protein component of hair and nails. Membrane proteins often serve as receptors or provide channels for polar or charged molecules to pass through the cell membrane. In this dissertation, I will mainly focus on soluble
and membrane proteins. In terms of structure determination, soluble proteins are easier to crystallize than membrane proteins, which are way underrepresented in the Protein Data Bank (PDB).

Gaining a fundamental understanding into these complex structures and how proteins are able to fold into their requisite shapes is still poorly understood and remains a central question in biochemistry. To understand protein function we need to know their structure and this is the ultimate aim of the work presented here.

### 1.1 Protein crystallization

The three dimensional structure of a protein determines how it functions and interacts with other molecules. To fully understand these processes, pathways and eventually how organisms function, researchers need to know the structure of the protein. Whereas genomics can reveal the sequence of amino acids in a protein, structural biology tells us how that sequence folds up into a particular shape. The different functions of the protein molecule and its interaction with other molecules are determined by its three dimensional structure. The human proteome is estimated to contain at least a million proteins, so in order to map out the functionality of all the proteins an efficient, high throughput technological effort is required to determine structural information for all human proteins. The Protein Structure Initiative (PSI) was launched by the National Institute of General Medical Sciences (NIGMS), part of the National Institutes of Health (NIH) in 2000 with the aim of obtaining high resolution structures of 10000 proteins in a decade. To achieve this goal, the PSI also focused on the development of innovative approaches and tools that streamline and speed many steps involved in generating protein structures, and to incorporate these new methods into process pipelines that turn DNA sequence information into protein
Figure 2: (a) Progress in protein structure determination by the protein structure initiative (PSI). Figure adapted from R. F. Service, Science, 2008, 319, 1610-1613. (b) Plot showing the number of unique alpha helical integral membrane protein structures deposited in the PDB each year, highlighting the difficulty in obtaining membrane protein structures. Figure used with permission from R. M. Bill, P. J. F. Henderson, S. Iwata, E. R. S. Kunji, H. Michel, R. Neutze, S. Newstead, B. Poolman, C. G. Tate and H. Vogel, Nat Biotech, 2011, 29, 335-340.
structures. Figure 2 shows the rapid development achieved by the PSI in solving and depositing structures of proteins, while highlighting the difficulty in obtaining membrane protein structures\(^5\). Similar to this, the Membrane Protein Initiative Roadmap has been instrumental in driving research for structure determination of membrane proteins.

Protein structures are considered to contain a treasure trove of information about life's molecular machines. The use of X-ray crystallography, nuclear magnetic resonance spectroscopy and cryo-electron microscopy reveal structural information which provides insight into how these complex machines operate. Often these insights can help to discover the likely function of an unknown protein, lead to a deeper understanding of how misshaped proteins cause disease, and potentially reveal a path to new drug treatments\(^2\).

Protein crystallography is an important aspect of structural biology that provides information on the three dimensional structure of proteins and insight into their function. This intensive, multi-step process includes expression, purification, and crystallization of protein material, followed by harvesting of the crystals for X-ray analysis (Figure 3). Some of these steps have been improved recently, but major bottlenecks still exist in identifying crystallization conditions with limited material available and in avoiding crystal damage while manually harvesting the oftentimes small, delicate crystals for X-ray analysis\(^6\).

Current methods in crystallography rely on a single protein crystal to give the entire structural data at cryogenic temperatures to avoid radiation damage\(^7\). This requires physically harvesting individual crystals from crystallization droplets. The manual handling, exposure to the environment during harvesting and cryo-cooling has the potential to lead to degradation of the crystal thereby affecting the quality of data collected. As a result of these issues, researchers try to find alternative crystallization conditions yielding more robust crystals of the same protein
or the protein is abandoned in favor of crystallization of other, hopefully more robust proteins.

Structural data from a protein crystal needs to be coupled with phase information which can be obtained experimentally through methods like single wavelength anomalous diffraction (SAD) or multi-wavelength anomalous diffraction (MAD), or through computational methods like molecular replacement. For de novo structure determination, experimental methods are more useful since homologous structures of novel targets, as needed for molecular replacement, may not be readily available.

While the success of structural biology efforts has been limited by bottlenecks associated with protein expression, purification, and
crystallization, \cite{15-21} the development of high throughput strategies utilizing robotics \cite{4,16,22-25} and microfluidics \cite{17,26-76} have helped to overcome some of these challenges. Automated methods \cite{21,73} also exist for the screening and collection of X-ray diffraction data from protein crystals. Despite this level of automation, the harvesting and mounting of crystals for X-ray analysis \cite{22,65,67,70,77} remains the only fully manual step in the structural biology pipeline.

Once the protein has been expressed and purified satisfactorily, the next and most crucial step to obtaining structural information about the protein is growing diffraction quality crystals. However, methods for predicting protein crystallization conditions \textit{a priori} have do not exist, so obtaining high quality crystals involves searching a vast multidimensional chemical space which could involve screening thousands of different precipitants, buffers, salts etc. in separate crystallization trials. \cite{70,71,78-80} Traditional vapor diffusion \cite{18,19,81-84} and microbatch \cite{24,85,86} methods continue to be the most commonly used crystallization techniques. However for sparse matrix screening a sizeable quantity of protein solution is required which is not always feasible in the initial stages of protein crystallization because of either difficulty in the expression or purification of a hitherto unknown protein. Scalability is also an issue with current platforms. Hence technology that could screen hundreds of conditions using minimal amount of protein solution would significantly benefit the field.

Once optimal crystallization conditions have been identified, traditional methods in X-ray crystallography require manual harvesting of the crystal from the crystallization chamber, cryo-freezing the crystal, and mounting of the crystal onto a loop. Handling of crystals can be particularly challenging if the crystals are small and fragile and can also lead to damage of the crystal due to handling and environmental shock that can occur when the crystal is retrieved from the environs in which it was grown. \cite{36,82} In case of many hard to crystallize proteins, micro
crystals can be grown, but mounting and collecting data from such crystals presents a challenge. Advances in synchrotrons have led to the development of micro beams which can obtain structural information from such crystals, but harvesting and mounting these tiny crystals manually is in extremely challenging. A crystallization platform which enables in situ analysis would allow collecting slices of data from many such micro crystals and merging them to get a complete dataset. Removal of this final, manual step of harvesting and ability to perform in situ analysis will open up the possibility of getting structural information from recalcitrant protein targets.

1.2 Conventional methods for protein crystallization

Determining the three dimensional structure of a protein involves many steps, starting from cloning, expression and purification of the protein, followed by crystallization and then collection of diffraction data. Even when high quality, protein solution is available in plenty, producing diffraction quality crystals is the bottleneck of the protein structure determination pipeline. My work is focused on developing tools to accelerate and simplify the crystallization process and make the data collection process easier. To justify the methods and tools I am going to be using, I will first describe the methods which are most prevalent today for protein crystallization.

There are two phases in growing high quality protein crystals for collecting X-ray diffraction data. These are (a) the screening of a large number of chemical, biochemical and physical conditions that yield crystalline material, and (b) the systematic adjustment of these conditions to obtain crystals of high enough quality to collect diffraction data and pursue subsequent analysis. During screening, we may or may not get good quality crystals, but even if we get a shower of
microcrystals, that gives us some idea as to what the likely crystallization condition could be, and we move to the second step, i.e., optimization. Every component in the crystallization cocktail (precipitant, salt, buffer, temperature, pH etc.) might have an impact on the final quality of the crystal. Hence each of these parameters needs to be fine-tuned during the optimization step in order to get diffraction quality crystals. Another way of optimization is to influence and control the very environment in which crystallization is taking place, for example by seeding, and this will be discussed later.

When the objective is to grow protein crystals, a solution of the protein must be transformed or brought into a supersaturated state, whereby its return to equilibrium forces exclusion of protein molecules into the crystal. If, from a saturated solution, for example, solvent is gradually

\begin{figure}
\centering
\includegraphics[width=\textwidth]{phase_diagram}
\caption{Schematic of the phase diagram for protein crystallization. Adjustable parameters on the x-axis could include precipitant concentration, pH, temperature etc. Four major crystallization methods are shown, (i) microbatch, (ii) vapor diffusion, (iii) dialysis and (iv) FID. The filled grey circles represent starting points. Figure used with permission from N. E. Chayen and E. Saridakis, \textit{Nat Meth}, 2008, 5, 147-153.}
\end{figure}
withdrawn by evaporation, temperature is lowered or raised appropriately, the pH is altered, or additives change the solvent concentration, or some other property of the system is altered, then the solubility limit may be exceeded and the solution will become supersaturated. If a solid phase is present, or introduced, then strict saturation will be reestablished as protein molecules leave the solvent, join the solid phase (crystal), and equilibrium of the system is regained. If no solid (nuclei, previous protein crystal) is present as conditions are changed, then solute will not immediately partition into two phases. The solid state does not necessarily develop spontaneously as the saturation limit is exceeded because energy, analogous to the activation energy of a chemical reaction, is required to create the second phase, the stable nucleus of a crystal or a precipitate. A kinetic, or energy barrier has to be crossed to go further and further from equilibrium, into the zone of supersaturation. On a phase diagram, (Figure 4), the line indicative of saturation is also a boundary that marks the requirement for energy-requiring events to occur in order for a second phase (ordered nucleus of a crystal) to be established, or the nonspecific aggregate that characterizes a precipitate. Once a stable nucleus has formed in a supersaturated solution, it will continue to grow until the system regains equilibrium. So long as non-equilibrium forces prevail and some degree of supersaturation exists to drive events, a crystal will grow or precipitate will continue to form.

There are three main zones in the phase diagram (Figure 4): undersaturated, saturated and supersaturated. The area under the solubility curve is the undersaturated region where spontaneous nucleation cannot occur and any crystal placed in this region will dissolve. The next region, which is the saturated region, is defined by the experimentally defined solubility curve. Along this curve, spontaneous nucleation will not occur, neither will crystals dissolve nor increase in size. The supersaturated region above the solubility curve is divided into three
Figure 5: Schematic representation of the conventional methods of protein crystallization (a) Microbatch-under-oil method, showing how a crystallization trial is dispensed under oil and a standard microbatch well plate. Used with permission from N. E. Chayen, Structure, 1997, 5, 1269-1274. Vapor-diffusion method, with schematic of (b1) hanging drop method and (b2) sitting drop method. Figure from en.wikipedia.org. (c) Dialysis method for crystallization, figure from en.wikipedia.org
further regions, the metastable zone, nucleation zone and precipitation zone. The metastable zone which is directly next to the solubility curve is where nucleation will not occur, but crystals present in this zone will grow in size. The second region, the nucleation zone, is where spontaneous homogenous nucleation can occur and crystals added to this solution will grow in size. The final zone is the precipitation zone, which is extremely supersaturated with respect to crystal growth and will lead to amorphous precipitation.

All methods of crystallization involve a phase transition, in which the protein is in the solution at the start of the experiment, and is forced to come out of solution and form crystals as the experiment progresses and the solution is driven towards supersaturation. There are many methods used for bringing the protein solution into a supersaturated state, amongst them the most common ones are (a) microbatch, (b) vapor diffusion, (c) free interface diffusion and (d) dialysis.

1.2.1 Microbatch crystallization

Microbatch method\textsuperscript{24,85,87,88} is a scaled down method of batch crystallization method, where the protein and crystallizing agents are mixed at the final concentration into a droplet at the start of the experiment. Supersaturation is achieved on mixing and conditions in the droplet change only when the protein leaves the solution to form crystals. In microbatch experiments (Figure 5a) crystallization trials are dispensed and incubated under low-density paraffin oil (0.87 g/ml)\textsuperscript{88}. The aqueous crystallization drops are denser than the oil, so they remain immersed, where they are protected from evaporation, contamination and shock, thus making the plates more easily transportable. The fundamental difference between diffusion based methods (vapor diffusion, FID) and microbatch is that diffusion based systems are dynamic in which conditions change throughout the crystallization process, whereas in microbatch, the samples are mixed at their final concentration at the start of the experiment; thus, conditions are constant within a normal
time frame of a crystallization experiment (1–3 weeks). Thus microbatch experiments do not generally span as large a chemical space as is possible with diffusion based experiments. A large number of microbatch experiments need to be set up to cover a range of conditions (in a phase diagram) that can be covered with a single vapor diffusion trial. However, there are usually no changes in drop volume or pH, and crystals do not usually dissolve.

Robotic systems can dispense nanoliter sized droplets for setting up hundreds of microbatch trials under oil. Depending on the type of oil used\textsuperscript{88,89} microbatch technique can be used for screening as well as optimization. It is compatible with almost all known precipitants, buffers salts and additives, with the exception of small molecule organic volatiles (phenol or dioxane) as these dissolve in the oil.

1.2.2 Vapor diffusion

Vapor diffusion\textsuperscript{87,88,90} is the most commonly used method for protein crystallization. In this setup, a drop containing the protein solution and precipitant (at a lower concentration than what is needed for crystal formation) is equilibrated against a reservoir solution. The precipitant and protein in the droplet get concentrated over time due to dehydration-driven reduction of solution volume caused by equilibration of water vapor from the protein-containing droplet to a more hygroscopic reservoir solution. During the crystallization process (see phase diagram in Figure 4), the droplet passes through a range of conditions, thereby undergoing a self-screening process. Vapor diffusion trials can be set up as a hanging drop or a sitting drop trial (Figure 5b).

Trials are usually conducted in 24-well plates (for example, Hampton Research or Molecular Dimensions), where each well is filled with 0.5–1 ml of reservoir solution and sealed with a siliconized glass or plastic coverslip. The wells are fitted with a support for the sitting drop, and
in the case of hanging drop; the crystallization droplets are pipetted onto the coverslip and inverted on to the reservoir well.

There are instances when vapor diffusion is not an ideal method\textsuperscript{88} because of a number of problems. These include (i) uncontrollable changes in drop volume and lack of control over exact droplet composition during crystallization, (ii) changes in pH due to volatile ions and (iii) changes in temperature that can cause dissolution of crystals. In addition, the trials are not easily transportable, and as compared to microbatch trials a large amount of both protein and precipitant are required.

1.2.3 Free interface diffusion

Free interface diffusion (FID)\textsuperscript{33,82,88} is a relatively new technique where the protein and the precipitant are juxtaposed in a microchannel or capillary. Due to diffusion over time, a concentration gradient develops in the system, and the system self-selects the optimum nucleation and growth conditions. This technique is most powerful when suitable precipitant(s), additive(s), pH, and buffers are already known and tested, that is, as a fine-tuning screening and optimization technique. Recently, several alternatives have been developed that has led to further exploration of this technique as a screening and structure determination tool. An initial screening version using the method has been developed\textsuperscript{65}, as a crystallization cassette in which multiple capillaries are simultaneously filled with the protein solution and then come into contact each with a different precipitant solution well and thus combines screening and optimization. FID based crystallization techniques have also been widely used in development of microfluidic chips for high throughput crystallization screening and the advantages of using microfluidics for implanting this technique are discussed later.
1.2.4 Dialysis

In the dialysis\(^{88,90}\) method, the protein is sequestered from the precipitant solution by a semipermeable membrane that allows the precipitant solution to slowly mix with the protein molecules (which cannot cross the membrane). Dialysis provides an altered route across the phase diagram (Figure 5c) but requires expertise to set up. In this method, solution composition is altered by diffusion of low molecular weight components through a semipermeable membrane. The ability to change the protein solution composition accurately any number of times and with small incremental changes makes this one of the more versatile crystallization methods. Dialysis is also uniquely suited to crystallizations at low ionic strength, an approach shown to succeed in nearly 10% of all cases.\(^ {91}\) It is also a convenient method for crystallization in the presence of volatile reagents such as alcohols.

1.2.5 Seeding

From the phase diagram of crystallization it is clear that the optimal conditions for nucleation and growth differ. Therefore a method which can separate these two phenomena and can
optimize them independent of one another is particularly appealing. Seeding is a one such

technique where crystals are transferred from conditions that support only nucleation to those

which support crystal growth\textsuperscript{92,93}. There are two basic seeding methods, macroseeding and

microseeding. In macroseeding a whole crystal is transferred from a condition where nucleation

occurs to an area of lower supersaturation that supports growth. In contrast, in microseeding,

nuclei of crystals are added to the crystallization droplet at a growth inducing supersaturation.

These nuclei are prepared by crushing crystals and preparing dilution series with different

number of nuclei present in each series for further optimization. Figure 6 shows examples of

protein crystals grown via seeding. A more detailed description of the seeding process is present

in Chapter 5.

1.2.6 The “ideal” crystallization platform

All the approaches described above for protein crystallization have limitations. Most of them

require large amounts of proteins, which is often times hard to obtain for membrane proteins, or

targets that are difficult to purify, especially when screening and optimization for a novel protein. 

There have been developments especially in robotics\textsuperscript{16} that allows for using nanoliter amounts of

protein solution, but these methods require expensive and complicated ancillaries and robotic

systems, which are often times out of reach for most biochemical laboratories. A low-cost,

simple to use, technology that uses minimal quantities of protein and reagents can prove to be a

disruptive innovation in the protein crystallization community, which still uses plastic well-plates

for setting up majority of the trials. Microfluidics can also eliminate crystal harvesting and

looping, which tends to damage fragile crystals or proves difficult in the case of tiny crystals, by

allowing for on-chip crystal analysis through the use of suitable materials for chip fabrication.

Also, traditional methods rely on collecting all data from a single cryo-cooled crystal, which
doesn’t allow the study of those proteins that are not amenable to cryo-crystallography or when the structure and function is desired to be studied at biologically relevant temperatures.

Based on the above discussed issues in the current approach to perform X-ray crystallography on protein crystals, the attributes of an ideal crystallization platform can be enumerated as follows:

1. Ability to screen multiple conditions, in high throughput fashion for protein crystallization
2. Requirement of minimal amount of protein solution for crystallization trials
3. Capability of on-chip analysis (eliminating manual harvesting, cryocooling and handling)
4. Capability of room temperature data collection from multiple crystals
5. Interfacing with current infrastructure at beamlines in synchrotrons

Microfluidics has the capability of manipulating fluids on a very small scale ranging from microliters down to picoliters, and this has been used for a variety of applications ranging from chemical synthesis to biological studies. The ability to control fluid flow at such a small scale makes it an ideal platform for crystallization since microfluidic devices use minimal amounts of protein and precipitant which need to be mixed by free interface diffusion (FID) or counter diffusion in order for crystals to grow. Ease of scaling out of microfluidic systems will enable setting up screens of hundreds of crystallization conditions while maintaining control over the conditions in each individual trial. Traditional microfluidic chips are usually made out of materials which are not suited for on-chip analysis, which is discussed in further detail in the Chapter 2. To overcome this issue, I use a material which is compatible with in situ X-ray analysis and also convenient for fabrication of microfluidic chips. The material selection and subsequent design and fabrication of an X-ray compatible microfluidic device is one of the main focus areas of my research. As discussed earlier, diffraction quality crystals of some novel
proteins have been challenging to grow, but showers of micro crystals which are individually too small to solve a structure completely have been grown. Since microfluidics offers such precise fluid control and handling at a small, micro crystals of difficult proteins can easily be grown once the crystallization condition is known, and subsequently the protein structure can be obtained by collecting multiple datasets from many crystals and merging them together. In summary, with its small scale, microfluidics addresses the current problems in protein crystallization in two ways: firstly, by allowing experiments to be carried out on novel proteins which are not available in large amounts and secondly, by enabling us to perform experiments which are difficult to implement on a large scale. In addition to these scale advantages, the use of innovative materials that are transparent to X-rays in the chip fabrication would allow on-chip diffraction data collection, and thus remove manual handling of fragile protein crystals for the structure determination process.
1.3 Microfluidic platforms for protein crystallization

In the same way in which integrated circuits brought about many technical advances in the electronics and computation industry, the ability of microfluidics to integrate multiple laboratory functionalities on a single chip significantly impact the field of chemical synthesis and biological studies. This ability of integration of massively parallel operations on a single chip has not only the potential to increase the speed and reliability of many conventional processes, but also bring down the cost of manufacturing and operation due to economies of scale. Microfluidics has been gaining popularity in structural biology, and has shown promise in many related fields ranging from studying protein folding with time resolved studies to growing high quality protein crystals for structure determination.\textsuperscript{56,70,95,96}

Microfluidics devices for protein crystallization can be categorized into three main categories,

(i) devices with active valves (ii) droplet based devices and (iii) systems based on mimicking well plates, e.g. SlipChip.\textsuperscript{80} All these systems deal with nanoliters volumes of protein and precipitant solution but differ in the way of operation and use and in terms of the materials used for fabrication of the chip itself.

In microfluidic systems of the first type, valves control the fluid flow of the solutions on the device. The most prevalent valves in microfluidics are based on pneumatic actuation and can be broadly classified as (i) actuate-to-open\textsuperscript{94,97} (Figure 7a) and (ii) actuate-to-close\textsuperscript{98} (Figure 7b). Typically, the actuation of pneumatic microfluidic valves is based on the actuation of a thin membrane by pressurized air in a control layer that is positioned over a network of microchannels embedded in a fluid layer. Actuate-to-close valves are operated by applying a positive pressure to collapse the membrane over the fluid layer into the fluid channel and thus
Figure 8: (a) Microfluidic device for protein crystallization based on pneumatic valves. (a1-3) depict the loading of protein and precipitant into respective chambers by valve action. (a4) Optical micrograph of the entire chip. Crystals of (a5) bacterial 70S ribosome and (a6) type II topoisomerase grown on-chip. Figure used with permission from C. Hansen and S. R. Quake, Current Opinion in Structural Biology, 2003, 13, 538-544. (b) On-chip generation of lipidic cubic phase for membrane protein crystallization. (b1) Schematic of the chip showing the lipidic cubic phase mixer and the precipitant and crystallization chambers. (b2) Micrograph of the chip with crystals of bacteriorhodopsin grown on-chip. Reprinted with permission from S. L. Perry, G. W. Roberts, J. D. Tice, R. B. Gennis and P. J. A. Kenis, Crystal Growth & Design, 2009, 9, 2566-2569. Copyright (2009) American Chemical Society.
closing the channel, and need rounded fluid channels for the membrane to completely stop the flow or drive fluid out of the fluid chamber. This effect is similar to pressing down on a ketchup packet and the ketchup squirting out from the mouth of the packet. When the valves are not actuated, these valves are open, so transporting these chips from the fluid handling station to the X-ray chamber can lead to unwanted mixing between the various chambers. Microfluidic systems based on these valves need external ancillaries to operate the valves and thus are not suitable for portability.

Alternatively, actuate-to-open valves are closed in their rest state. These valves are operated by applying vacuum in the control layer, which causes the thin membrane to deflect upwards, allowing flow or mixing of the fluids in the fluid layer beneath. These valves are closed at rest so devices based on these valves can be transported easily without leading to unwanted mixing. However they require selective reversible bonding of the fluid layer to the substrate so as to allow the multiple actuations of the valves. These valves can be patterned much more densely than the actuate-to-close valves and thus are useful for fabrication of dense microfluidic networks.\textsuperscript{94,97} Figure 8 show some of the crystallization platforms which are based on the pneumatic valve networks. These valves have significant potential in high throughput screening applications where portability is desired.

In droplet based methods, a second type of microfluidic platforms that relies on two phase flow, protein crystallization takes place inside aqueous volumes of a mixture of protein and precipitant which are surrounded by an immiscible carrier fluid (e.g. oil) as shown in Figure 9\textsuperscript{29,31,80}. The crystallization trials can be conducted in tubes or capillaries made out of plastic or glass. These methods allow for trials to be set up to mimic microbatch and vapor diffusion trials and allow for \textit{in situ} analysis of crystals grown in the case where the capillaries are made out of
X-ray transparent materials. Recent advances have also allowed lipidic cubic phase generation on these droplet based devices for carrying out in meso crystallization trials.

The SlipChip is an example of the third type of microfluidic platform in which the protein and precipitant are loaded into two different wells on different substrates which can slide relative to each other (Figure 9d). The motion of the plates brings the two solutions in contact,
thus driving protein crystallization. In these chips, this relative motion of the wells is comparable to the valve actuation in the valve based devices. This method enables exact metering of nanoliters of solution into the wells on each of the plates which are then brought into contact. One of the main advantages of this method is the simple operation of these chips, since the operation doesn’t need any external ancillary system to set up the actual trial.

1.4 X-ray compatible microfluidic platforms for protein crystallization

The challenge in implementing on-chip analysis on a microfluidic chip is finding a material which offers minimum scattering and attenuation from X-rays. Once a suitable material has been identified design and fabrication of a microfluidic chip has to be developed to implement on-chip analysis on microfluidic platforms. Traditional methods for protein crystallography involve looping the crystal out of the mother liquor and flash freezing it in liquid nitrogen. This prevents formation of ice on the crystal. The looped crystal can be used for cryogenic data collection. While the mother liquor will still be present in most microfluidic systems, addition of cryoprotectant is possible to allow for cryogenic data collection. Additional challenges in development of an on-chip microfluidic system pertain to device geometry and size so that it can be accommodated on present beamlines setup for X-ray data collection. A chip which can be mounted on the current standardized setup on most synchrotrons would be highly beneficial as it can use the technology currently in place right away.

X-ray capillaries have been traditionally associated with crystallography. These capillaries are not always suited for protein crystallography especially under cryogenic conditions because of their length which can be several centimeters. The long tubes are necessary to obtain the required distance for diffusion to occur and consequently grow many droplets of crystals.
However, during cryocooling, not all the drops in the long capillary can be cooled with the narrow cryostream which is only a few millimeters in diameter. The remaining portion of the capillary will still be at room temperature and will act as a heat sink for the cooled part.

Therefore, these conventional systems suffer from sub optimal cryocooling and ice formation. Microfluidic systems based on droplets in capillaries or tubes suffer from the same problems and require cutting out the region of interest or harvesting the crystal out of the device in order to collect data. Microcapillary systems have been developed recently, which use X-ray transparent plastics so that on-chip data can be collected. However, the only reported structure from on-chip data is for model proteins like lysozyme, for other proteins, they still have to harvest the
crystal form the capillary, and collect data under cryogenic conditions. Figure 10 illustrates some of the platforms which are based on capillaries that have been developed for on-chip analysis.

Figure 10 illustrates some of the platforms which are based on capillaries that have been developed for on-chip analysis.

Figure 11: (a) A chip for protein crystallization by counter diffusion made out of X-ray transparent materials, shown here are eight crystallization channels connected through a tree-like network between the inlet for protein sample and the inlets for the crystallization agent. Crystals of (b) bovine insulin (c) plant virus, and (d) turkey egg white lysozyme. Reused with permission from K. Dhouib, C. Sauter et al., Lab on a Chip, 2009, 9, 1412-1421. (e) The MPCS crystal card mounted on the goniometer, for on-chip data collection. Reused with permission from C. J. Gerdts, M. Elliott, S. Lovell, M. B. Mixon, A. J. Napuli, B. L. Staker, P. Nolliert and L. Stewart, Acta Crystallographica Section D, 2008, 64, 1116-1122, http://dx.doi.org/10.1107/S0907444908028060 (f) The 24-well X-CHIP with base that attaches to the goniometer mount. (g) Close-up of the wells of the X-CHIP with protein crystals grown on-chip. Used with permission from G. Kisselman, W. Qiu, V. Romanov, C. M. Thompson, R. Lam, K. P. Bataille, E. F. Pai and N. Y. Chiragadze, Acta Crystallographica Section D, 2011, 67, 533-539, http://dx.doi.org/10.1107/S0907444911011589
Microfluidic chips for on-chip analysis require a different choice of material from traditional microfluidics. The majority of microfluidic devices reported in the literature have been fabricated out of relatively thick layers of polydimethylsiloxane (PDMS) and glass, materials which result in significant attenuation of X-rays,\(^{101}\) thus rendering the devices unsuitable for \textit{in situ} analysis.\(^{27-32,34,35,74,80,102-104}\) While materials like polyimide (PI), polycarbonate (PC), poly methyl methacrylate (PMMA), polystyrene (PS) and cyclic olefin copolymer (COC) which attenuate X-rays to a lesser extent have been used to make microfluidic devices,\(^{49,63,66,69,105-110}\) these devices have been limited to single layer devices with little or no integrated fluid handling capabilities (Figure 11). Simple counter diffusion and free interface diffusion based microfluidic chips have been fabricated and tested for on-chip protein crystallography, these devices usually comprise of microchannels molded into a plastic substrate which is X-ray transparent\(^{56,63,111,112}\). Figure 11 gives examples of various chips which have simple geometries and carry out protein crystallization by either counter diffusion or free interface diffusion. In all these examples, chips have been fabricated from thin layers of an organic based polymer to minimize the scattering and attenuation from X-rays. The main issue with these devices is that they consist of relatively simple networks and are single layered devices that don’t have active fluid control which would be necessary for more involved experiments. Apart from this, most of the experiments performed with these chips have involved data collection on a model protein, of which the crystallization condition and structure is already known. Halazonetis et al., have created chips out of cyclic olefin homopolymer that can carry out microbatch and vapor diffusion trials on chip and is X-ray transparent\(^{113}\). They have also successfully crystallized two novel proteins on-chip but do not report on-chip diffraction data from the same, the only on-chip data they show is again from model systems like lysozyme. X-CHIP\(^{114}\) (Figure 11) is another recent platform that allows
for data collection without manual harvesting of the crystals. While they have successfully collected phasing information from multiple crystals at room temperature, they have not utilized it for solving the structure of a protein or commented on the phasing data quality. Therefore, it still remains to be seen whether these devices can be used for screening conditions for new proteins and solve the structure from high resolution diffraction data collected on-chip. The few instances in which a multi-layer device was used for novel protein structure determination, involved punching out the area of interest for further analysis, which is not suited for high throughput crystallography$^{56,112,115}$.

### 1.5 Conclusions

Protein crystallization still remains an art instead of a science. Currently, no methods exist for predicting crystallization conditions $a$ $priori$ and this makes it necessary to set up of thousands of crystallization screening trials before identifying one or more suitable condition. Even after obtaining such a condition, a protein structure is not guaranteed through traditional methods as the crystals grown may not be of high quality or large enough to yield good quality data. Methods to carry out further studies of protein dynamics and interaction have not been fully developed yet. Manual handling of crystals is still required to harvest crystals out of the well or plates the trial has been set up in and subsequently loop them and cryo-cool the crystal, so that X-rays can be shot at them for structure determination. Crystals that are fragile and small have the potential to be damaged during this step. Automation of the crystal handling, the only manual step in the current process for protein structure determination, will obviate any human intervention once the trial has been set up. Microfluidics has the potential to impact the field by facilitating on-chip screening, crystallization and eliminating handling of the crystals, followed
by on-chip X-ray analysis at biologically relevant temperatures thus facilitating studies of those targets for which cryo-compatible crystallization conditions cannot be found.

The work described here hopes to address these challenges. In Chapter 2, new protocols for fabrication of a hybrid microfluidic platform fabricated out of COC, which is X-ray transparent, has been described. This platform would allow on-chip analysis of protein crystals while retaining the functionality and fluid control of a complex microfluidic network on-chip. Integrating this platform with current synchrotron facilities has also been discussed. In Chapter 3 validation of this platform with several model proteins is discussed. The data collection strategy for collecting data from multiple crystals and merging them to get a complete dataset is also discussed. Chapter 4 discusses the application of this platform for screening, crystallization and de novo structure determination via on-chip phasing of a novel protein PhnA. In Chapter 5, the development of a microfluidic approach to microseeding is presented. Crystals of photoactive yellow protein are grown via microseeding on-chip and diffraction data from the crystals are collected. Chapter 6 discusses a method to collect functional information from protein crystals via Laue crystallography. This is the first time that a microfluidic chip has been used for collecting time resolved diffraction data on multiple proteins.

Microfluidic platforms have tremendous potential to address many problems with respect to studying difficult systems in structural biology. The first step to realize this potential is to validate a robust platform that can screen multiple crystallization conditions and allow in situ structure determination of novel proteins. Later, this platform can be adapted to study various aspects of dynamic crystallography, including protein folding and kinetics, and interaction with other biomolecules.
1.6 References


Chapter 2

Fabrication of X-ray compatible microfluidic platforms for protein crystallization

2.1 Introduction

This chapter reports a method for fabricating multilayer microfluidic protein crystallization platforms using different materials to achieve X-ray transparency and compatibility with crystallization reagents. Design and fabrication of a microfluidic chip with an array of crystallization wells for screening different conditions that consume a minimal amount of protein solution as compared to traditional screening methods is described. A large number of high quality isomorphous protein crystals can be grown in the wells, after which slices of X-ray data can be collected from many crystals still residing within the wells. Complete protein structures can be obtained by merging these slices of data followed by further processing with crystallography software. This approach of using an x-ray transparent chip for screening, crystal growth, and X-ray data collection enables room temperature data collection from many crystals mounted in parallel, which thus eliminates crystal handling and minimizes radiation damage to the crystals.

Microfluidic platforms have gained widespread use in diverse fields like chemical synthesis, enzymatic and DNA analysis ¹,², proteomics ³,⁴, point-of-care, medical and clinical

¹ Part of this work has been published in S. Guha et al., Sensors and Actuators B, 2012, 174, 1-9.
diagnostics \cite{5-7}, energy conversion \cite{8,9}, and protein / pharmaceutical crystallization \cite{10-14}. Microfluidics offers several advantages in terms of reduced sample size and easy preparation, fine control over transport phenomena on the microscale, ease of scalability, detection and sample analysis on a single, integrated platform \cite{15,16}. A large fraction of microfluidic devices reported in the literature to date have been fabricated using poly(dimethylsiloxane) (PDMS) and glass, which allow for rapid prototyping \cite{13,17,18}. Issues still exist, however, with respect to compatibility with different chemistries, temperature and pressure variation, as well as amenability to integration of analysis techniques. Many alternative materials and combination of materials have been explored for the fabrication of microfluidic platforms with the goal to overcome these issues for specific applications \cite{19,20}. Typical challenges encountered in these efforts are bonding of different materials, interfacing with ancillary equipment, and the assembly of chips comprised of highly integrated, complex designs.

The focus of this work is on the application of microfluidic platforms in structural biology, specifically for protein crystallography. Protein structure determination involves handling of small sample volumes, fine control over transport properties during crystallization, and requires the collection and analysis of X-ray diffraction data from the crystals formed \cite{21-23}. A microfluidic approach has the potential to meet these requirements. For example, our group recently reported a microfluidic array chip for solid form screening of candidate pharmaceuticals \cite{14}, but these chips lacks X-ray transparency. The first step in protein crystallization involves screening a protein against a wide range of precipitants to identify suitable crystallization condition(s). Once one or more conditions have been identified, traditional methods require the crystals to be manually harvested, cryo-cooled and mounted in an X-ray beam for structure determination, all one crystal at a time. These challenging tasks, especially when crystals are small and fragile, often damage the crystal, affecting the quality of data collected. Several of the microfluidic platforms reported in the literature have tried to address these challenges, however most of them still require the crystal to be manually harvested before data collection \cite{12,13,24,25}. X-ray transparent platforms have also been reported, but they have been limited to relatively simple, single-layer designs \cite{11,19,26,27} that do not take advantage of the integrated fluid handling capabilities of multilayer microfluidics or require cutting out the section containing the crystal from the whole chip for further analysis \cite{28}.
Here I report on the fabrication of a microfluidic platform that allows for (i) screening up to 100 crystallization conditions while consuming minimal amount of protein solution, (ii) on-chip X-ray data collection from the protein crystals grown, while (iii) retaining the fluid routing and manipulation capabilities of integrated, valve-based high throughput microfluidic systems. This is achieved by assembling a thin hybrid microfluidic chip comprised of layers of cyclic olefin copolymer (COC) and PDMS. I later will validate the utility of this chip and its use by crystallizing various proteins on-chip, followed by X-ray diffraction data collection from the crystals grown while they still reside within the microfluidic platform mounted as a whole in the X-ray beam (subsequent chapters).

2.2 Materials and methods

2.2.1 Fabrication of microfluidic platforms

Silicon wafers were patterned using negative photoresist SU8-25 and SU8-2050 (MicroChem). The microfluidic platforms consisted of different layers of polymers. The control layer was fabricated with COC (TOPAS Advanced Polymers Inc, 4 mil). Patterning of the COC was performed via hot embossing at 175°C (Tg+50°C), using a master made out of a high temperature epoxy resin (Conapoxy FR 1080, 83:100 hardener: epoxy, by mass) using a hot press (Carver hot press, model 3851-0).

The fluid layer was fabricated out of PDMS (General Electric RTV 615, Part A/B). Inlets were drilled using a 750 µm drill bit (McMaster Carr). Bonding of the fluid and control layer (see SI for further details) was done using a 1% (v/v) solution of 3-glycidoxypropyltrimethoxysilane (GPTMS, Sigma Aldrich) and a 1% (v/v) solution of 3-aminopropyltrimethoxysilane (APTMS, Sigma Aldrich). The completed assembly was placed on a flat COC substrate prior to setting up of crystallization trials. The use of vacuum for the operation of the devices eliminated the need to bond the fluid layer irreversibly to the COC substrate.

All solutions were introduced on-chip by pipetting 2-4 µL of protein/precipitant solution on the inlet ports, then pulling the fluid into the chip by actuation of the appropriate valves by applying vacuum through a manifold (Cole Parmer) attached to a vacuum pump (GAST; Model DOA-P704-AA). The vacuum was applied onto the chip through 24 gauge AWG, thin walled
PTFE tubing (Cole Parmer) connected to a block of PDMS, which was positioned over the appropriate control line inlet. Once the chip was filled, all the inlets were covered with Crystal Clear Tape (Hampton HR4-511) to prevent evaporation and the trials were incubated. Even though there is some loss of solution through the thin layer of PDMS, experiments have shown that trials in these chips can be incubated for up to 2 weeks.

2.3 Results and discussion

2.3.1 X-ray compatibility and materials characterization

Attenuation of X-rays can be calculated for a particular energy based on the exponential decay in intensity of a narrow beam of monochromatic photons from an incident intensity $I_0$ as it passes through a material of thickness $x$ with a linear attenuation coefficient of the material $\mu$.

$$I = I_0 \exp(-\mu x)$$

(1)

Attenuation coefficients have been well studied and documented for elemental materials. For a compound containing multiple elements, a linear attenuation coefficient can be calculated based on the sum of the contribution to attenuation from each of the individual elements $i$, weighted based on their mass fraction $w_i$.

$$\mu = \sum \mu_i w_i$$

(2)

Table 1 lists the atomic mass fractional compositions of various materials commonly used in microfluidic device manufacture and shows calculated values for the linear attenuation coefficients for SiO$_2$, PDMS, COC, PMMA and PI as a function of photon energy. Soft X-rays (lower energy) attenuate much more strongly than do harder X-rays (higher energy), thus the energy of X-rays used for an experiment can have a significant effect on the signal observed from a device. Using values for the attenuation coefficient for X-rays with a wavelength of 1Å (12.4 keV), the transmission factor $I/I_0$ can be calculated as a function of material thickness using Eq. 1.
Table 1. Atomic mass fraction density and a calculated value for linear attenuation coefficient $\mu$ at 1Å (12.4keV) for various materials used in microfluidic devices. SiO$_2$ = quartz $^{31}$, PDMS = polydimethylsiloxane (Si$_{60}$O$_{60}$C$_{124}$H$_{368}$) $^{30}$, COC = cyclic olefin copolymer (C$_9$H$_{14}$) $^{32}$, PMMA = polymethylmethacrylate (C$_3$H$_8$O) $^{30}$, PI = polyimide (C$_{22}$H$_{10}$N$_2$O$_5$) $^{33}$.

<table>
<thead>
<tr>
<th>Element</th>
<th>SiO$_2$</th>
<th>PDMS</th>
<th>COC</th>
<th>PMMA</th>
<th>PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>--</td>
<td>0.081</td>
<td>0.116</td>
<td>0.096</td>
<td>0.026</td>
</tr>
<tr>
<td>C</td>
<td>--</td>
<td>0.329</td>
<td>0.885</td>
<td>0.714</td>
<td>0.692</td>
</tr>
<tr>
<td>N</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0.073</td>
</tr>
<tr>
<td>O</td>
<td>0.533</td>
<td>0.212</td>
<td>--</td>
<td>0.190</td>
<td>0.209</td>
</tr>
<tr>
<td>Si</td>
<td>0.467</td>
<td>0.378</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Density (g/cm$^3$)</td>
<td>2.65</td>
<td>0.92</td>
<td>1.02</td>
<td>0.94</td>
<td>1.42</td>
</tr>
<tr>
<td>$\mu$ at 1Å (cm$^{-1}$)</td>
<td>9.330</td>
<td>7.334</td>
<td>1.131</td>
<td>1.472</td>
<td>1.618</td>
</tr>
</tbody>
</table>

Figure 12: Diffractogram of on-chip data collected from a lysozyme crystal, with the inset showing diffraction spot quality. The black ring shows the upper limit of the diffraction data collected (high resolution). The annulus enclosed within the red circles represents the scatter from the device materials (7.5 Å for PDMS and 5.2 Å for COC), which is at a much lower resolution and doesn’t affect the high resolution data.

Several interesting observations can be made from the data presented in Table 1. PDMS and PMMA have very similar densities; however the linear attenuation coefficient for PDMS is significantly higher. This difference arises from the silicon content in PDMS. Heavier atoms present a larger cross-section for interacting with photons and will thus cause a larger degree of attenuation. The density of a material also plays a role in the degree of attenuation observed,
with higher density materials increasing the number of atoms which can interact with a photon for a given path length, though this effect is less significant than elemental composition.

Knowing the attenuation coefficient for various materials, an expression for the attenuation through a series of films of thickness \( j \) can be calculated based on Eq. (1).

\[
I = I_0 \exp(-\sum \mu_j x_j) \quad \text{(3)}
\]

A typical device used here has an X-ray path length of 145 µm of COC and 20 µm of PDMS (50 µm COC substrate, 20 µm PDMS membrane, 75 µm COC control layer). PDMS and COC show a characteristic scattering pattern that can be observed in the diffraction data collected on-chip.

However this scattering pattern occurs at relatively low angles of \( q \)-spacing or areas of low resolution diffraction (5.2 Å for COC and 7.5 Å for PDMS) and does not affect the data collected at higher resolution (Figure 12).
In designing an X-ray compatible microfluidic device for use in protein crystallography three main considerations with respect to the interaction between materials and X-rays must be taken into account: (i) attenuation and (ii) scattering of X-rays passing through device materials, and (iii) the strength of diffraction resulting from a crystal. Attenuation originates from the absorption of photons by the material, thereby decreasing the intensity of both the incident X-ray beam and the resulting diffracted X-rays. Scattering is an elastic redirection of photons based on the internal structure of the material and can affect the signal-to-noise ratio. The strength of the diffraction signal from a crystal is related to not only the degree of order within the crystal, but also the packing density and the size of the crystal\textsuperscript{34-38}.

Figure 13a shows the variation of the linear attenuation coefficient with the X-ray energy. In general, the linear attenuation coefficient is lower at higher X-ray energy. At relevant energies for X-ray data collection (here 12.4 keV), COC attenuates X-rays seven times less than PDMS and SiO\textsubscript{2}. Figure 13b shows the transmission factor for COC and PDMS as a function of film thickness. A thickness exceeding 3 mm is typical for conventional microfluidic devices, at which point the transmission factor is less than 10% (Figure 12b). However for the hybrid COC-PDMS device as reported here, the transmission factor was more than 90% (Figure 13a, dotted line shows data collection energy of 12.4 keV). An analysis of both PDMS and COC shows that these materials produce a characteristic scattering pattern, seen in the form of rings in magnified diffraction data. However this scattering occurs at relatively low resolution and does not impact the collection of high resolution diffraction data.

2.3.2 Design and operation of microfluidic devices

The microfluidic device presented here enables the screening of protein solution against different precipitants, and also allows for structure determination of the protein once a suitable crystallization condition has been found. The main challenge here is to create a microfluidic chip that is still capable of fluid routing and compartmentalization to allow screening of many conditions, while also being X-ray transparent.

Array chip fabrication

As discussed above, I chose to replace the traditionally thick layers of PDMS with thin films of COC because of its lower X-ray attenuation profile and beneficial material properties. However, COC is not suitable for the integration of valves into a microfluidic device because it does not have the same degree of flexibility as PDMS, an essential property for typical
pneumatic valves\textsuperscript{39}. To retain the ability to route and compartmentalize fluids on-chip, I use a thin layer of PDMS sandwiched between COC layers. This hybrid approach eliminates the bulk of PDMS usually found in traditional devices. Using a 20 µm PDMS layer sandwiched between two COC layers has only a minimal effect on the transmission of X-rays at energies relevant for X-ray data collection (Figure 13a).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure14.png}
\caption{Fabrication scheme for the hybrid COC-PDMS chip. (a) The inverse pattern for the control layer is patterned onto a Si-wafer using negative photore sist to create a master. (b) PDMS (10:1) was poured over the master mould for the control layer. (c) High temperature epoxy (FR-1080) was poured on the PDMS mould and cured at 120 °C for 4 hours, after which it was gently lifted off the PDMS mould. (d) The epoxy master was placed on top of a COC sheet (4 mils) and sandwiched between a PDMS backing layer and a glass slide and the assembly is placed inside a hot press. (e) The entire assembly was heated up to 175°C in the hot press and maintained at that temperature for 5 minutes to pattern the control layer. (f) A thin fluid layer was made out of PDMS (15:1) and patterned using standard soft lithographic techniques. (g) Holes were drilled for the control layer inlets (not shown) following which, the control and fluid layer were aligned and bonded using a silane based chemical treatment (see Figure 5). (h) Holes were drilled for the fluid layer inlets (not shown) and the assembled device was placed on a thin COC or Duralar substrate. (i) A 3-dimensional view of the control and fluid layer being aligned with the colored lines corresponding to the points in the device as shown in (g).}
\end{figure}
The hybrid microfluidic chips are comprised of a thin PDMS fluid layer sandwiched between and bonded to a COC control layer and flat COC substrate (Figure 14a). The control layer is fabricated via hot embossing against an epoxy master. This epoxy master is a fabricated by first creating a photoresist-on-silicon master using standard photolithography, followed by replication of this silicon master into a PDMS mold, which in turn is replicated in epoxy.\textsuperscript{20} Sheets of COC (100 µm) were patterned with this epoxy master by hot embossing at 175°C (Figure 14a). The COC control layers have 25 µm deep and 50 µm wide negative relief patterns as the control lines. Figure 15 shows the fidelity of pattern transfer into COC. Inlet holes for actuation of valves were then drilled in the COC control layer.

Figure 15: SEM images of (a) silicon wafer patterned with SU-8-25 with 25 µm tall features, (b) the inverse PDMS mold created by replica molding from the SU-8-25 on silicon pattern, (c) The epoxy master molded from the PDMS inverse master, (d) The final COC control layer fabricated via hot embossing using the epoxy master. The magnified view shows the steep sidewalls in the features. All scalebars correspond to 200 µm except where otherwise mentioned.

The fluid layer is fabricated in PDMS using standard soft lithography\textsuperscript{39}, with the COC control layer being bonded to the 70 µm thick fluid layer before being lifted from its photoresist mater\textsuperscript{40}. A perspective view of this assembly is shown in Figure 14b. The COC control layer was chemically bonded to the thin PDMS fluid layer. Initial simple well devices were bonded using an oxygen plasma treatment while array chips with dense valve networks, utilized cross-linking between an epoxy-terminated silane on the surface of the control layer and an amine-terminated silane on the surface of the fluid layer (Figure 16)\textsuperscript{40}. A 60 second oxygen plasma
treatment was used to activate the PDMS and COC surfaces for the formation of a silane bond. The COC control layer was immersed in a 1% (v/v) solution of 3-glycidoxypropyltrimethoxysilane (GPTMS) and the PDMS fluid layer was immersed in a 1% (v/v) solution of 3-aminopropyltrimethoxysilane (APTMS) for 20 minutes. After rinsing with DI water and blowing dry with nitrogen the two layers were aligned and brought into contact. A strong bond forms almost immediately and the assembled structure was allowed to cure for 1 hour at room temperature \(^{40}\). Because of the nature of the normally closed valves, it was not necessary to bond the device to the COC substrate \(^{41,42}\). To further decrease the thickness of the chip, Duralar (0.5 mils) was used as a substrate. Inlet holes for the fluid layer were then drilled through this composite assembly. Lastly, this COC-PDMS assembly was placed on a 50 µm thick COC sheet to close the fluid channels.

The fabrication material described above was successfully used to create and test 24- and 96-well array chips for on-chip collection of X-ray diffraction data. However, the success rate in fabricating denser and larger array chips with this fabrication protocol was low. This was mainly due to the challenges of aligning the hot pressed control layer successfully over the fluid layer. The two main issues were, (a) the APTMS-GPTMS bonding strategy meant that only one attempt could be made at aligning the two layers, since the bond was immediate and irreversible, and (b) the hot embossing created a factor of expansion in the COC control layer which made alignment of the 96-well array chips very cumbersome. Though the 24-well devices had an excellent success rate (>90%), the number of working 96-well chips was low. For high throughput screening it was
essential that the 96-well array chip be easy to fabricate, so I worked on an alternative fabrication protocol.

To address the two main issues, hot embossing and irreversible bonding, one method that could be used was to make the control and fluid layer both out of PDMS. These PDMS layers would have to be very thin so as to ensure that the new device geometry was still X-ray transparent. Conventional soft lithographic procedures utilize a thick PDMS control layer to ensure easy bonding to a thin fluid layer, a strategy that would not work in this case. Some other material would have to be used to lift off the control layer and provide the requisite stiffness to the control layer for it to be bonded to the fluid layer. I used COC as that material, because it was already proven to be X-ray transparent and was much more rigid than PDMS.

![Diagram of fabrication process](image)

**Figure 17:** Fabrication scheme for the hybrid COC-PDMS-PDMS-COC chip. PDMS is spun coat on the FL and CL masters respectively and baked in the oven. COC sheet is bonded to the CL PDMS and then the assembly is lifted up from the Si-wafer and aligned and bonded to the FL as in the previous method. The combined assembly is baked for four hours before placing it on a flat COC substrate. The dimensions for the complete device are: $a = 50 \ \mu m$, $b = 45 \ \mu m$, $c = 25 \ \mu m$, $d = 80 \ \mu m$, $e = 60 \ \mu m$, $f = 50 \ \mu m$.

In the new fabrication scheme (Figure 17), the features for the control layer were fabricated via patterning photoresist on silicon wafer followed by spin coating a thin layer of PDMS (5:1) onto it. This PDMS control layer was baked at 70°C for 10 minutes (until the PDMS was just slightly tacky). A 2 mil COC sheet was then plasma treated along with this PDMS coated wafer for a minute and then the two were brought into conformal contact and baked for an hour at 70°C to ensure bonding. The COC-PDMS control layer was then peeled from the control layer master and the exposed PDMS side was covered with scotch tape and holes were drilled for the inlets. The fluid layer was prepared in the same manner as before. The partially cured 15:1 PDMS on the fluid layer master was then placed under the microscope and the COC attached control layer was aligned over it. Since this was bonding between two PDMS layers, the bonding was reversible and allowed greater ease in alignment of the features, especially for larger array chips. Other lab members have successfully aligned 192-well array chips using this technique proving
its reliability and robustness. The aligned fluid and control layer assembly was then baked in the oven for four hours to ensure robust bonding. Following this, the assembly was carefully peeled
off the wafer, and holes were drilled for the fluid inlets. This completed assembly was then placed on a thin COC substrate to complete the fabrication.

**Design of array chips**

The array chips consist of a series of separate half-wells for protein and precipitant solutions arranged in columns. Each of these individual wells is a separate crystallization trial (Figure 18a) and is isolated from the rest of the wells using a series of normally-closed valves (Figure 18a, green and blue valves)\(^{43}\). Since these valves are closed at rest, the chip can be transported easily after filling without disturbing the crystallization trials. Each half-well contains 50 nL of solution and the entire chip uses just 1.4 µL of protein solution for a 24-well design. These adjacent half wells are separated by a normally closed valve that allows mixing between the protein and precipitant solution through free interface diffusion. The time required to mix two solutions in adjacent wells can be estimated using Fick’s law, \( t = \frac{x^2}{4D} \), where \( x \) is the combined length along which the two solutions are mixing and \( D \) is the diffusivity of the diffusing species (protein in this case, approximately \(~10^{-6}\) cm\(^2\)/s). For a mixing length of around 1 mm, the time needed to mix two solutions is approximately 20 minutes. The mixing time can be adjusted for other protein solutions with different diffusivities.

**2.3.3 Array chip operation**

The two half-wells are filled independently of each other using dedicated valve lines (blue and green valves for the aqueous protein and precipitant solutions respectively; see Figure 14a) for each set of half-wells arranged in columns. These valves are closed at rest thus isolating the various chambers. The mixing valves (pink in Figure 18) between adjacent half-wells enable the protein and precipitant solution to mix by diffusion. The entire filling process is illustrated in Figure 18. First 2 µL of protein solution is pipetted over the inlet of the protein line and vacuum is applied to the protein valve line (blue) which allows for dead-end filling of the protein solution into the series of protein half-wells (Figure 18b). Once the protein solution has been introduced, the protein valve set is closed, isolating the solutions in the protein chambers. Next, 1 µL droplets of precipitant solution are pipetted over the inlets of each of the six precipitant. Vacuum is then applied via the precipitant valve lines (green) and the precipitant solutions flow into the appropriate half-wells (Figure 18c). After closing the precipitant valve set, thus locking the precipitants in their respective chambers, the inlets for filling and actuation are sealed with
Crystal clear tape, preventing evaporation of the filled solutions. Next, the mixing valves (pink) located between the protein and precipitant chambers are actuated, allowing sets of two adjacent solutions to mix by free interface diffusion (Figure 18d). I allowed the solutions to mix for 20 minutes for all experiments performed here. After mixing, the mixing valves are closed, the inlets of these control lines are sealed with Crystal Clear tape, and the chips are incubated at either 4°C or room temperature for 8-48 hours. Previously, we have used a similar microfluidic array chip for solid form screening of candidate pharmaceuticals\(^{14}\), but this platform lacked X-ray transparency.

2.3.4 Single-well chip design, operation, and proof-of-concept experiments

Before completing the full design and fabrication of the 24-well array chip, I wished to confirm that indeed high quality diffraction data can be collected from crystals residing inside a microfluidic well comprised of COC layers as well as a thin PDMS layer (Figure 19). These chips for validating the X-ray compatibility of the materials are made using a similar fabrication scheme described earlier (Section 3.2.1). They consist of a large rectangular 2.36 mm by 3.36 mm by 25 µm wells with a volume of 0.2 µL. Six posts were fabricated within the fluidic well chamber to provide support for the thin PDMS membrane. The COC control layer consisted of

---

**Figure 19:** (a) Design schematic of single large well microfluidic chip. Numbers in brackets represent the total height of each layer. The feature size in both the COC and PDMS layers is 25 µm. (b) A single well chip mounted on a modified magnetic goniometer mount for data collection.
an identical sized chamber aligned directly above the fluid well (Figure 19). Pre-mixed solutions of protein and precipitant (see Materials and Methods) were introduced into the chip by pipetting a drop over the inlet of the fluid layer and pulling it into the fluid layer well by applying vacuum to the inlet of the control layer (Figure 19). The chip was allowed to incubate for 48-72 hours.

Figure 20: (a) A standard magnetic goniometer mount for macromolecular crystallography (b) A modified version of (a) to mount microfluidic chips. (c) The modified magnetic chip holder mounted on a beamline at LS-CAT, Argonne National Lab. The inset shows the inline imaging capabilities of the beamline, which allows us to identify and aim X-rays at individual crystals.
To test the effectiveness of the chip materials for on-chip X-ray data collection, I first grew crystals of model proteins lysozyme, ribonuclease A, and thaumatin in single well chips. The whole chips were then mounted onto a goniometer mount and taken to Argonne National Lab (LS-CAT). Further discussion of the on-chip data quality is discussed in Chapter 3.

2.3.5 Interfacing with beamline infrastructure and data collection strategy

The chips described above are X-ray transparent and can be used for data collection from crystals grown on-chip. In order to collect data from the 3rd generation synchrotron sources, which enable data collection at extremely high speeds, it is necessary that the chip be compatible with the infrastructure already present there with minimum modifications to the current setup. The University of Illinois has collaboration with LS-CAT (Life Sciences Collaborative Access Team) at the Advanced Photon Source in the Argonne National Lab in Chicago. LS-CAT is set up for macromolecular crystallography and the crystals are mounted on the beamline using a magnetic goniometer mount that has a loop for scooping out the crystal. To utilize the same setup with minimum modification, I modified the goniometer mount for single crystal studies to fit the microfluidic chip. To achieve this, I cut off the stem which holds the crystal loop, and glued a metallic cylinder to hold the chip. This cylinder was machined in order to have a slit running down the top, with an attached set screw to hold the chip in place. Figure 20 shows in further detail the modified goniometer mount for holding the chip and how it interfaces with the setup at LS-CAT. Being able to mount the chips on to the beamline with minimum alteration to the setup at LS-CAT ensures that we can use the excellent sample visualization aids already present to locate and shoot individual crystals on the chip with great precision.

2.3.6 Screening and on-chip protein structure determination

Identification of suitable crystallization conditions by screening protein solution against a wide variety of crystallization mixtures (cocktail of salts, buffers and precipitants) is the first step in the process of protein structure determination. Chapters 3 and 4 will discuss this further. In general, a series of array chips was used to set up to conduct screening experiments where a protein was screened against a library of different conditions. These experiments were generally carried out in larger 96-well array chips where the size of the protein and precipitant chambers were varied along the vertical to cover a larger chemical space in terms of the concentration of the protein and the precipitant. The fabrication of these chips is identical to the method
described earlier in the chapter. Figure 21 shows schematics of the both 24- and 96-well variants of the array chip. Once a suitable condition had been identified, the same condition was set up in replicate on the 24-well array chip (with identical wells) to grow a large number of isomorphous crystals for diffraction data collection. The 24-well chips were small enough to fit in the X-ray beam without further modification and were perfect for on-chip data collection.

Note that typical microbatch wellplate experiments need 2 µL of protein solution per condition whereas a whole 24-well array chip can be filled using less than 2 µL of protein solution. These chips truly scale down the crystallization trials that are performed in usual biochemistry and structural biology labs.
The microfluidic array chips were validated for crystallization screening experiments using various model and novel proteins that will be discussed in later chapters. X-ray compatibility of the chip allowed us to easily identify whether the hits were indeed protein crystals or just salt crystals using either a local X-ray source (Bruker APEX) or the APS. The variability between results from the microbatch well plate and our chip can be explained by (i) differences in the method of mixing of protein and precipitant solutions; (ii) slow concentration of solutions in the microfluidic chips over time; and (iii) the stochastic nature and variability of crystallization trials.

2.4 Conclusions

In summary, the microfluidic platform reported here allows for crystallization screening and subsequent on-chip X-ray data collection of crystallized proteins. The fabrication scheme uses a combination of COC and PDMS layers, and yields a platform that retains fluid routing capabilities via pneumatic valving while also achieving X-ray transparency. The dense network of integrated pneumatic valves allows for fast and easy setup of the crystallization screens, which in turn facilitates the reproducible growth of a large number of isomorphous crystals. The valves are only opened by vacuum actuation during filling and mixing. The valves are closed in rest, so the individual crystallization trials are isolated during crystallization and data collection in the absence of connections with ancillary equipment. The whole chip can be transported to and mounted in the X-ray beam for collection of diffraction data wedges from many crystals still residing on chip at room temperature. By merging these wedges of data a complete data set for full structure determination is obtained. This approach eliminates the need for crystal harvesting and cryo-cooling, while it simultaneously avoids the detrimental effects of the effects of radiation damage associated with room temperature data collection.

Going forward, the X-ray compatible microfluidic chip reported here can be used to solve structure of novel proteins, especially those that (i) yield small, fragile crystals, (ii) have no known cryogenic crystallization conditions, and/or (iii) have been known to yield low resolution data using traditional cryogenic data collection. In addition to allowing for the routine collection of static structural information of proteins at biologically relevant temperatures as reported above, the fine control over transport phenomena coupled with on-chip analysis capabilities offered by these microfluidic platforms could enable a range of other biological studies. For
example, precise microfluidic fluid handling could be utilized to enable *dynamic* structural studies based on the addition of various stimuli such as ligands, electrochemical agents, acid/base to affect the pH, or even changes temperature. Such studies have the potential to shed light on structural changes associated with protein function in unprecedented ways.
2.5 References


33. V. Ratta, PhD, Virginia Polytechnic Institute, 1999.


Chapter 3

Proof of concept studies for on-chip X-ray diffraction data collection from soluble proteins

3.1 Introduction

Protein structure determination via X-ray diffraction is a critical aspect of structural biology. Knowledge of the three dimensional structure of a protein provides insight into the mechanism whereby these complex macromolecules function, as well as into potential sites for drug targeting. Recent efforts in structural biology have produced not only a large number of protein structures, but have helped to lessen bottlenecks, leading to improved, and often automated, high throughput methodologies for the cloning, expression, purification, crystallization, and X-ray diffraction analysis of novel protein targets. These approaches frequently take advantage of robotics and/or microfluidics. With these improvements, the harvesting and mounting of crystals for X-ray analysis remains the only manual step in the structural biology pipeline.

Protein crystallization and structure determination is an intensive multi-step process. Each of the steps of identification and optimization of crystallization conditions, determination of appropriate cryoprotection conditions, and examination of heavy atom derivatives or ligand binding represents a significant screening effort in the lab coupled with harvesting of individual crystals for X-ray diffraction analysis. However, the scale of these efforts can be hampered both by limitations on the quantity of available protein and the difficulties associated with manual harvesting of fragile protein crystals.

---

2 Part of this work has been published in S. Guha et al., Sensors and Actuators B, 2012, 174, 1-9.
Current methodologies in protein crystallography rely on collecting structural data from a single crystal at cryogenic temperatures to minimize the effects of radiation damage while avoiding concerns of crystal-to-crystal variability. This requires harvesting individual protein crystals from the crystallization droplet in which they were grown. However, both the physical handling associated with this process and exposure of the crystal to the ambient environment have significant potential to damage the crystal. Furthermore, the cryocooling process has been shown to induce crystal damage, although cryoprotection strategies attempt to minimize these effects. As a result of this single-crystal paradigm, crystals that are not amenable to cryo-crystallography or are highly sensitive to X-ray radiation may prove so difficult to work with that they are abandoned in favor of searching for new crystallization conditions that do not suffer from these limitations.

Microfluidic approaches for crystallization provide more precise control over the composition and kinetic parameters of a crystallization experiment than what is accessible at larger scales. However, difficulties associated with harvesting crystals from these devices hamper their use. The crystal harvesting challenge could be mitigated by mounting X-ray transparent microfluidic chips directly in the X-ray beam. Unfortunately, the typical size and planar design of microfluidic chips limit rotational freedom and are incompatible with current methods for cryocooling. Data collection strategies, including how anomalous diffraction data is collected, will have to be adjusted to overcome the latter two chip-related challenges.

The majority of microfluidic chips reported in the literature have been fabricated out of relatively thick layers of polydimethylsiloxane (PDMS) and/or glass, which are incompatible with the requirements for high quality on-chip X-ray diffraction analysis. The designs of X-ray compatible microfluidic chips reported in the literature that are based on materials like polyimide, poly(methyl methacrylate), and cyclic olefin copolymer (COC) have been relatively simple and typically do not take advantage of the high throughput integrated fluid handling capabilities characteristic of multi-layer PDMS chips. In those cases where X-ray data was collected from a multi-layer chip, the wells with crystals were punched out from a larger chip prior to analysis, a strategy that is not suited for high throughput crystallography. Furthermore, none of the reported examples have been utilized for the collection of anomalous data, most likely due to the more stringent signal-to-noise requirements.
In this chapter, I utilize the microfluidic chips described in Chapter 2 that retain the integrated fluid handling capabilities of traditional multi-layer devices, necessary for high throughput crystallization, while achieving a high level of X-ray transparency. A comparison of traditional methods as well as on-chip methods for performing standard crystallization screens is presented. I validated the utility of this chip and its use by crystallizing the soluble proteins lysozyme, thaumatin and ribonuclease A on-chip, followed by X-ray diffraction data collection from the crystals grown while they still reside within the microfluidic platform mounted as a whole in the X-ray beam.

3.2 Materials and methods

3.2.1 Protein solutions

Hen egg white lysozyme (Sigma) was dissolved in 50 mM sodium acetate (Sigma-Aldrich) at pH 4.6 with 20% (w/v) glycerol (Fisher Scientific) at a concentration of ~100 mg/mL. Lysozyme concentrations were determined by UV absorbance measurements (Lambda 650 UV-Vis spectrophotometer, Perkin Elmer) at 280 nm using an extinction coefficient of 2.64 mL/(mg-cm) \textsuperscript{45}. For proof-of-concept crystallography experiments precipitant solutions of 1M and 2M NaCl (Aldrich) in 50 mM sodium acetate, pH 4.6 with 20% (w/v) glycerol were prepared. For screening experiments Hampton Crystal Screen chemicals were used directly (Hampton Research).

Thaumatin from Thaumatococcus daniellii (Sigma) was dissolved in 100 mM NaH\textsubscript{2}PO\textsubscript{4} (EMD Chemicals) at pH 6.5 at a concentration of 82 mg/mL. The protein concentration was determined by UV absorbance measurements at 280 nm using an extinction coefficient of 1.25 mL/(mg-cm) \textsuperscript{46}. A precipitant solution of 30% (w/v) Na/K tartrate (Malinckrodt) and 20% w/v glycerol in 100 mM NaH\textsubscript{2}PO\textsubscript{4} pH 7.0 was used \textsuperscript{47}.

Ribonuclease A (R-5500, Sigma) from bovine pancreas was dissolved in 100 mM sodium acetate at pH 4.5 at a concentration of 229 mg/mL. The protein concentration was determined by UV absorbance measurements at 280 nm using an extinction coefficient of 0.70 mL/(mg-cm). \textsuperscript{48} A precipitant solution of saturated NaCl in 100 mM sodium acetate at pH 4.5 was used \textsuperscript{49}.
3.2.2 Fabrication and operation of microfluidic platforms

Silicon wafers were patterned using negative photoresist SU8-25 and SU8-2050 (MicroChem). The microfluidic platforms consisted of different layers of polymers. The control layer was fabricated with COC (TOPAS Advanced Polymers Inc, 4 mil). Patterning of the COC was performed via hot embossing at 175°C (T<sub>g</sub>+50°C), using a master made out of a high temperature epoxy resin (Conapoxy FR 1080, 83:100 hardener: epoxy, by mass) using a hot press (Carver hot press, model 3851-0).

The fluid layer was fabricated out of PDMS (General Electric RTV 615, Part A/B). Inlets were drilled using a 750 µm drill bit (McMaster Carr). Bonding of the fluid and control layer (see SI for further details) was done using a 1% (v/v) solution of 3-glycidoxypropyltrimethoxysilane (GPTMS, Sigma Aldrich) and a 1% (v/v) solution of 3-aminopropyltrimethoxysilane (APTMS, Sigma Aldrich). The completed assembly was placed on a flat COC substrate prior to setting up of crystallization trials. The use of vacuum for the operation of the devices eliminated the need to bond the fluid layer irreversibly to the COC substrate.

All solutions were introduced on-chip by pipetting 2-4 µL of protein/precipitant solution on the inlet ports, then pulling the fluid into the chip by actuation of the appropriate valves by applying vacuum through a manifold (Cole Parmer) attached to a vacuum pump (GAST; Model DOA-P704-AA). The vacuum was applied onto the chip through 24 gauge AWG, thin walled PTFE tubing (Cole Parmer) connected to a block of PDMS, which was positioned over the appropriate control line inlet. Once the chip was filled, all the inlets were covered with Crystal Clear Tape (Hampton HR4-511) to prevent evaporation and the trials were incubated. Even though there is some loss of solution through the thin layer of PDMS, I have shown that trials in these chips can be incubated for up to 2 weeks.

3.2.3 Visualization of crystallization experiments and setup of traditional crystallization trials

Crystallization experiments were set up and visualized using either a stereomicroscope (Leica, MZ12.5) equipped with a digital camera (Leica, DFC295) or a computer controlled imaging system comprised of an optical microscope (Leica Z16 APO) equipped with an auto-zoom lens (Leica 10447176), a digital camera (Leica DFC280), and a motorized x-y stage.
(Semprex KL66) controlled by Image Pro Plus (Media Cybernetics). Periodically birefringent images of the wells were taken.

Traditional microbatch-under-oil crystallization trials were set up combining 2 µL each of protein and precipitant solutions in a Greiner well plate (Hampton Research) at room temperature. Crystals were harvested using crystal mounts (Mitegen). Crystallization trials of lysozyme were performed in traditional well plates, large well devices, and array chips at 4°C or at room temperature. Crystallization trials for thaumatin and ribonuclease A were performed at room temperature.

3.2.4 X-ray data collection

The microfluidic chips were mounted on a standard magnetic goniometer mount (Hampton Research) with an attached metal tube into which a slit was cut. A set-screw was used to secure samples. Diffraction data was collected either at room temperature or under cryogenic conditions. Cryo-cooling of samples was achieved by direct immersion into liquid nitrogen. Various sample-to-detector distances were used based on the quality of the crystal present. Typically data was collected using 1° steps with a 1s exposure at an X-ray energy of 12.7 keV (λ = 0.979Å) at Argonne National Lab (LS-CAT). Data from multiple crystals in the microfluidic devices was collected over a range of 10° (-5° to +5° from the normal) and an optimal subset of the frames was subsequently merged to obtain a complete dataset. Bench-top diffraction experiments were performed at the George L. Clark X-ray Facility at the University of Illinois using a General Area Diffraction Detector System (GADDS; Bruker) equipped with a four circle diffractometer and HiStar multiwire area detector. A rotating anode generator (Bruker M18XHF22) operating at 40 kV and 60 mA was used with a graphite monochromator supplying a Cu Kα radiation beam (λ = 1.54Å or 8.048 keV). The sample to detector distance is ~18 cm. Data was collected in a coupled mode where 2ω = 2θ such that multiple frames could be collected over a wider range of 2θ. Typically two such frames were collected, spanning the range of 2θ from 0° to 40°, up to a resolution of 2.3 Å.

Analysis of X-ray diffraction data collected at the synchrotron was performed using HKL2000 software for indexing, refinement, integration, and scaling (HKL Research) 50. Diffraction data collected at the University of Illinois was analyzed using GADDS software (version 4.1.08, Bruker AXS) and Topas 3 (Bruker AXS). Subsequent processing of
crystallography datasets was done using the CCP4 suite of programs\textsuperscript{51}. Electron density maps were displayed using COOT\textsuperscript{52}. Molecular replacement\textsuperscript{53} for lysozyme was done using PDB structure 193L as the model\textsuperscript{54}.

### 3.3 Results and discussion

#### 3.3.1 Single-well chip design, operation, and proof-of-concept experiments

Before completing the full design and fabrication of the 24-well array chip, I wished to confirm that indeed high quality diffraction data can be collected from crystals residing inside a microfluidic well comprised of COC layers as well as a thin PDMS layer. These chips for validating the X-ray compatibility of the materials are made using a similar fabrication scheme described earlier (See Chapter 2). They consist of a large rectangular 2.36 mm by 3.36 mm by
25 μm wells with a volume of 0.2 μL. Six posts were fabricated within the fluidic well chamber to provide support for the thin PDMS membrane. The COC control layer consisted of an identical sized chamber aligned directly above the fluid well (Figure 22a). Pre-mixed solutions of protein and precipitant were introduced into the chip by pipetting a drop over the inlet of the fluid layer and pulling it into the fluid layer well by applying vacuum to the inlet of the control layer (Figure 22a). The chip was allowed to incubate for 48-72 hours.

To test the effectiveness of the chip materials for on-chip X-ray data collection, I first grew crystals of model proteins lysozyme (Figure 22c), ribonuclease A (Figure 22d), and thaumatin (Figure 22e) in single well chips. The whole chips were then mounted onto a goniometer mount (Figure 22b) and taken to Argonne National Lab (LS-CAT). The inlets in Figure 25c-e shows on-chip diffraction data collected from these chips from the different model proteins. Even though a scatter ring from the materials is present, good quality, circular diffraction spots are still clearly visible, sufficient for structure determination.

### 3.3.2 Screening and on-chip protein structure determination

Identification of suitable crystallization conditions by screening protein solution against a wide variety of crystallization mixtures (cocktail of salts, buffers and precipitants) is the first step in the process of protein structure determination. Figure 23 shows a 24-well array chip that allows for the testing of six different precipitant solutions in quadruplicate. A series of array chips was used to test each of the individual conditions multiple times and the results were compared to those obtained using a traditional microbatch well plate.

**Table 1.** Crystallization results of 100 mg/mL lysozyme in 50mM sodium acetate with 20% glycerol against the 50 condition Hampton crystal screen at RT.

<table>
<thead>
<tr>
<th>Condition number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-10</td>
<td>-/+</td>
<td>-/+</td>
<td>/-</td>
<td>-/-</td>
<td>/-</td>
<td>+/+</td>
<td>+/+</td>
<td>+/+</td>
<td>+/+</td>
<td>+/+</td>
</tr>
<tr>
<td>11-20</td>
<td>+/+</td>
<td>+/-</td>
<td>/-</td>
<td>-/+</td>
<td>+/+</td>
<td>+/+</td>
<td>+/+</td>
<td>+/+</td>
<td>+/+</td>
<td>+/+</td>
</tr>
<tr>
<td>21-30</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>31-40</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>41-50</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
</tbody>
</table>

*Crystallization trials set up on a 24-well array chip. Trials set up on a microbatch tray under oil. "-" indicates absence and "+" indicates presence of crystals.
The microfluidic array chips were validated for crystallization screening experiments by testing solutions of lysozyme against the 50 condition Crystal Screen kit (Hampton Research) at room temperature. After one week crystals were observed in 32 out of the 50 conditions in the array chips (Table 1) compared to only 26 hits in the microbatch wellplate. X-ray compatibility of the chip allowed us to easily identify whether the hits were indeed protein crystals or just salt crystals using a local X-ray source (Bruker APEX, Table 2). A comparison of the results obtained in the well plates versus the traditional microbatch method shows that 21 of the conditions produced crystals on both platforms while 5 conditions yielded crystals uniquely in the microbatch wellplates and 11 hits were observed uniquely in the microfluidic chips. The variability between these results can be explained by (i) differences in the method of mixing of protein and precipitant solutions; (ii) slow concentration of solutions in the microfluidic chips over time; and (iii) the stochastic nature and variability of crystallization trials\textsuperscript{55-57}. Note that typical microbatch wellplate experiments need 2 µL of protein solution per condition whereas a whole 24-well array chip can be filled using less than 2 µL of protein solution.

\textbf{Figure 23}: Optical micrograph of crystals of lysozyme grown on a 24-well hybrid array chip, with a magnified view of one of the crystallization chambers.
Once crystals have grown on-chip, I want to collect X-ray data with the crystals still residing in the microfluidic wells. In tradition crystallography, crystals are manually looped out, cryo-cooled, mounted in a cryo-stream, and rotated in the X-ray beam for data collection. The dimensions and design of the array chips makes rotation of the chip for the collection of a complete dataset very difficult. The chip is also too large to fit within typical cryo-streams on crystallography beamlines, necessitating data collection at room temperature. Figure 24 shows a typical array chip mounted on one of the beamlines at the Advanced Photon Source (APS).

The inability to utilize cryo-cooling to mitigate the effects of radiation damage led us to adopt an alternative strategy where I collect small wedges of data (10 degrees) from a large number of crystals at room temperature and later merge these wedges to form a complete dataset. The use of small wedges of data limits the time of exposure, and thus the extent of radiation damage. These wedges can be merged to form a

---

Table 2. Crystallographic data statistics from the analysis of a subset of the lysozyme crystals obtained in wells of the microfluidic array chip via crystallization screening using Hampton Crystal Screen. Data was collected on-chip at room temperature. Reported values are for all hkls. Values shown in parenthesis represent the value for the highest resolution shell.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Crystal Screen Reagent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
</tr>
<tr>
<td>Resolution</td>
<td>26 – 1.83 Å</td>
</tr>
<tr>
<td>Linear R-factor ($R_{sym}$)</td>
<td>0.056 (0.244)</td>
</tr>
<tr>
<td>Mosaicity</td>
<td>0.08° – 0.26°</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.4 (3.0)</td>
</tr>
<tr>
<td>Completeness</td>
<td>85.1% (87.2%)</td>
</tr>
<tr>
<td>Frames collected</td>
<td>39</td>
</tr>
</tbody>
</table>

---

### 3.3.3 Strategy for on-chip data collection

Once crystals have grown on-chip, I want to collect X-ray data with the crystals still residing in the microfluidic wells. In tradition crystallography, crystals are manually looped out, cryo-cooled, mounted in a cryo-stream, and rotated in the X-ray beam for data collection. The dimensions and design of the array chips makes rotation of the chip for the collection of a complete dataset very difficult. The chip is also too large to fit within typical cryo-streams on crystallography beamlines, necessitating data collection at room temperature. Figure 24 shows a typical array chip mounted on one of the beamlines at the Advanced Photon Source (APS).

The inability to utilize cryo-cooling to mitigate the effects of radiation damage led us to adopt an alternative strategy where I collect small wedges of data (10 degrees) from a large number of crystals at room temperature and later merge these wedges to form a complete dataset. The use of small wedges of data limits the time of exposure, and thus the extent of radiation damage. These wedges can be merged to form a

---

Figure 24: An X-ray transparent, 24-well microfluidic array based device mounted on the beamline at Sector 21, LS-CAT, Argonne National Lab.
complete dataset, provided that non-isomorphism between crystals is minimal and the orientation of the grown crystals is mostly random. This method has been used previously to obtain structural information from tiny or fragile crystals or crystals which suffer from excessive radiation damage\textsuperscript{58,59}.

The advantage of applying this data collection strategy to crystals grown in a microfluidic chip is the ease by which a large number of crystals can be generated. Furthermore, the fine control of transport on the microfluidic scale allows for improved reproducibility both in crystal quality and isomorphism between crystals grown in different wells or even in different chips. Analysis of lysozyme crystals grown on-chip demonstrated that 98% of the crystals (47 out of 48) showed a coefficient of variation (standard deviation divided by the average) in the unit cell dimensions of \(~0.1\%\).

3.3.4 Analysis of on-chip collected X-ray data

I collected and analyzed room temperature X-ray data of lysozyme crystals grown in 24-well chips using the same crystallization conditions as used for the proof-of-principle experiments. For comparison, I also collected X-ray data at cryogenic temperatures on a lysozyme crystal grown in a wellplate and mounted using a loop with Table 3 showing a comparison of the important crystallographic statistics between the two methods.

Analysis of the crystallographic parameters and statistical measures of quality for both data collection and structural refinement via molecular replacement show that the quality of the data was not significantly affected either by collection at room temperature or because of the merging of data from multiple crystals. The diffraction data was analyzed over the range of 50 – 1.55 Å (the diffraction limit for the on-chip data) to enable direct comparison between the other parameters. Other data collection parameters such as $R_{sym}$, completeness, redundancy, and $I/\sigma$
and structural refinement parameters such as $R/R_{\text{free}}$ are comparable for the two cases, despite differences in crystal size, resolution, and signal-to-noise resulting from the presence of device materials. Interestingly, due to the absence of both physical handling and cryo-cooling, the mosaicity for the room temperature merged data is nearly an order of magnitude lower (i.e., better) than the mosaicity of the single-crystal cryogenic datasets. Figure 25 compares the diffraction patterns and electron density maps obtained from the traditional, single-crystal cryogenic data and the on-chip, multi-crystal, room temperature data. The high resolution data obtained in each of these cases is evident from the electron density maps of, e.g., aromatic amino acid side chains with the rings clearly visible.

Table 3. Crystallographic data statistics from the analysis of lysozyme crystals.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>On-Chip(^a)</th>
<th>Traditional (Cryogenic)(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit Cell Dimensions</td>
<td>$a = b = 79.693,\text{Å}$</td>
<td>$a = b = 78.817,\text{Å}$</td>
</tr>
<tr>
<td></td>
<td>$c = 37.781,\text{Å}$</td>
<td>$c = 37.025,\text{Å}$</td>
</tr>
<tr>
<td>Space Group</td>
<td>P4$_2$12</td>
<td>P4$_2$12</td>
</tr>
<tr>
<td>Observations (Unique)</td>
<td>783,994 (18,352)</td>
<td>223,433 (17,510)</td>
</tr>
<tr>
<td>Resolution</td>
<td>50 – 1.55Å</td>
<td>50 – 1.55Å</td>
</tr>
<tr>
<td>$R_{\text{sym}}$</td>
<td>0.064 (0.362)</td>
<td>0.052 (0.102)</td>
</tr>
<tr>
<td>Mosaicity</td>
<td>0.03° – 0.08°</td>
<td>0.21° – 0.34°</td>
</tr>
<tr>
<td>Redundancy</td>
<td>22.9 (5.7)</td>
<td>7.7 (7.7)</td>
</tr>
<tr>
<td>Completeness</td>
<td>98.1% (83.4%)</td>
<td>99.7% (100%)</td>
</tr>
<tr>
<td>$I/\sigma$</td>
<td>51.4 (3.9)</td>
<td>42.3 (19.4)</td>
</tr>
<tr>
<td># of Frames</td>
<td>363</td>
<td>100</td>
</tr>
</tbody>
</table>

$^a$ Merging of small datasets from "multiple crystals" analyzed on-chip within a 24-well device at room temperature. $^b$ The "traditional" sample was grown using microbatch techniques and mounted using a standard crystal mount for cryogenic data collection. Reported values are for all hkls. Values in parenthesis represent the value for the highest resolution shell except where indicated for the number of observations as compared to unique reflections and $R$ ($R_{\text{free}}$) and for the Ramachandran statistics where the number in parenthesis indicates the number of residues in a given region. Data was analyzed over the range of 50 – 1.55Å to enable a direct comparison with the data collected from merging the diffraction data taken from multiple crystals at RT.
3.4 Conclusions

In summary, the microfluidic platform reported here allows for crystallization screening and subsequent on-chip X-ray data collection of crystallized proteins. The fabrication scheme uses a combination of COC and PDMS layers, and yields a platform that retains fluid routing capabilities via pneumatic valving while also achieving X-ray transparency. The dense network of integrated pneumatic valves allows for fast and easy setup of the crystallization screens, which in turn facilitates the reproducible growth of a large number of isomorphous crystals. The valves are only opened by vacuum actuation during filling and mixing. The valves are closed in rest, so the individual crystallization trials are isolated during crystallization and data collection in the absence of connections with ancillary equipment. The whole chip can be transported to and mounted in the X-ray beam for collection of diffraction data wedges from many crystals still residing on chip at room temperature. By merging these wedges of data a complete data set for full structure determination is obtained. This approach eliminates the need for crystal harvesting and cryo-cooling, while it simultaneously avoids the detrimental effects of the effects of radiation damage associated with room temperature data collection.

The screening capabilities of the platform were highlighted using a standard screen, followed by comparison of the results with the same screen performed via a traditional crystallization method. The platform was validated by collecting data from several proteins like lysozyme, thaumatin and ribonuclease A. By solving the structure of lysozyme using only diffraction data from many crystals residing on-chip I confirmed that the strategy of merging wedges of data from many crystals to obtain a complete dataset can be applied to these microfluidic platforms.

Going forward, the methodology reported here can be used to solve structure of novel proteins, especially those that (i) yield small, fragile crystals, (ii) have no known cryogenic crystallization conditions, and/or (iii) have been known to yield low resolution data using traditional cryogenic data collection. While proof-of-concept studies have shown that data can be collected from model systems, the real test will be whether it can screen conditions for novel proteins and collect enough high quality data for structure determination. This has been discussed in the next chapter. In addition to allowing for the routine collection of static structural information of proteins at biologically relevant temperatures as reported above, the fine control over transport phenomena coupled with on-chip analysis capabilities offered by
these microfluidic platforms could enable a range of other biological studies. For example, precise microfluidic fluid handling could be utilized to enable dynamic structural studies based on the addition of various stimuli such as ligands, electrochemical agents, and acid/base to affect the pH, or even change temperature. Such studies have the potential to shed light on structural changes associated with protein function in unprecedented ways.
3.5 References


Chapter 4

De novo structure determination of proteins via X-ray compatible microfluidic platforms

4.1 Introduction

Protein structure determination via X-ray diffraction is a critical aspect of structural biology. Knowledge of the three dimensional structure of a protein provides insight into the mechanism whereby these complex macromolecules function, as well as into potential sites for drug targeting\(^1\)-\(^3\). Recent efforts in structural biology have produced not only a large number of protein structures, but have helped to lessen bottlenecks\(^4\),\(^5\), leading to improved, and often automated, high throughput methodologies for the cloning, expression, purification, crystallization, and X-ray diffraction analysis of novel protein targets\(^6\)-\(^15\). These approaches frequently take advantage of robotics\(^8\)-\(^12\),\(^14\),\(^16\) and/or microfluidics\(^17\)-\(^29\). With these improvements, the harvesting and mounting of crystals for X-ray analysis remains the only manual step in the structural biology pipeline\(^10\),\(^30\),\(^31\).

Protein crystallization and structure determination is an intensive multi-step process. Each of the steps of identification and optimization of crystallization conditions, determination of appropriate cryoprotection conditions, and examination of heavy atom derivatives or ligand binding represents a significant screening effort in the lab coupled with harvesting of individual crystals for X-ray diffraction analysis. However, the scale of these efforts can be hampered both by limitations on the quantity of available protein and the difficulties associated with manual harvesting of fragile protein crystals.

\(^{3}\) Part of this work has been submitted in a manuscript to Lab on a Chip and is under review.
Current methodologies in protein crystallography rely on collecting structural data from a single crystal at cryogenic temperatures to minimize the effects of radiation damage while avoiding concerns of crystal-to-crystal variability.\textsuperscript{15,32-37} This requires harvesting individual protein crystals from the crystallization droplet in which they were grown. However, both the physical handling associated with this process and exposure of the crystal to the ambient environment have significant potential to damage the crystal.\textsuperscript{21,35,38,39} Furthermore, the cryocooling process has been shown to induce crystal damage, although cryoprotection strategies attempt to minimize these effects.\textsuperscript{15,32,33} As a result of this single-crystal paradigm, crystals that are not amenable to cryo-crystallography or are highly sensitive to X-ray radiation may prove so difficult to work with that they are abandoned in favor of searching for new crystallization conditions that do not suffer from these limitations.

Protein structure determination involves not only the collection of a single diffraction dataset, but also generation of phase information. While computational methods such as molecular replacement can be used to generate phase information, these strategies cannot be applied to novel proteins for which a homologous structure is unavailable.\textsuperscript{37,40} Instead, phase information can be obtained experimentally by measuring the anomalous diffraction signal resulting from heavy atoms within a protein crystal, either at a single wavelength (single-wavelength anomalous diffraction, SAD) or at multiple wavelengths (multiple-wavelength anomalous diffraction, MAD).\textsuperscript{34,37} Because these anomalous diffraction measurements require the collection of symmetry-related pairs of data to observe a small anomalous signal, obtaining data with a high signal-to-noise ratio while minimizing radiation-induced damage and crystal non-isomorphism is critical. Historically these stringent requirements have meant that anomalous data must effectively be collected from a single protein crystal at cryogenic conditions, with only very few examples of anomalous data having been collected from multiple crystals.\textsuperscript{34,41-43}

Microfluidic approaches for crystallization provide more precise control over the composition and kinetic parameters of a crystallization experiment than what is accessible at larger scales.\textsuperscript{17-29} However, difficulties associated with harvesting crystals from these devices hamper their use. The crystal harvesting challenge could be mitigated by mounting X-ray transparent microfluidic chips directly in the X-ray beam. Unfortunately, the typical size and planar design of microfluidic chips limit rotational freedom and are incompatible with current
methods for cryocooling. Data collection strategies, including how anomalous diffraction data is collected, will have to be adjusted to overcome the latter two chip-related challenges.

The majority of microfluidic chips reported in the literature have been fabricated out of relatively thick layers of polydimethylsiloxane (PDMS) and/or glass, which are incompatible with the requirements for high quality on-chip X-ray diffraction analysis. The designs of X-ray compatible microfluidic chips reported in the literature that are based on materials like polyimide, poly(methyl methacrylate), and cyclic olefin copolymer (COC) have been relatively simple and typically do not take advantage of the high throughput integrated fluid handling capabilities characteristic of multi-layer PDMS chips. In those cases where X-ray data was collected from a multi-layer chip, the wells with crystals were punched out from a larger chip prior to analysis, a strategy that is not suited for high throughput crystallography. Furthermore, none of the reported examples have been utilized for the collection of anomalous data, most likely due to the more stringent signal-to-noise requirements.

In this chapter, approaches (i) for screening hundreds of crystallization conditions, and (ii) strategies to collect high quality structural data using these microfluidic chips, including anomalous diffraction data, which is necessary for the phasing and de novo structure determination of novel protein targets, are discussed. The microfluidic chips retain the integrated fluid handling capabilities of traditional multi-layer devices necessary for high throughput crystallization, while achieving a high level of X-ray transparency. A large number of high quality isomorphous crystals can be grown reproducibly and subsequent mounting of the chip in the X-ray beam enables the collection of diffraction data, including anomalous datasets for phasing, by merging small wedges of data collected from each of these crystals at room temperature. This approach bypasses concerns associated with crystal damage during handling, cryocooling, and radiation exposure. The approach is validated by solving the structure of a novel bacterial phosphonate hydrolase (PhnA) to 2.30 Å using SAD phasing methods that make use of data collected on crystals of selenomethionine labeled samples, entirely on-chip at room temperature. As such, this is the first example of an entirely hands-free crystallization and de novo structure determination performed within a microfluidic chip.
4.2 Materials and methods

4.2.1 Protein solutions

Native and Selenomethionine labeled PhnA from *Sinorhizobium meliloti* was dissolved in 20 mM HEPES, pH 7.5, at a range of concentrations: 15, 20, 25 and 40 mg/mL. For screening experiments, the 96 condition Hampton Index Screen (Hampton Research) was used and the condition I-80 (0.2M ammonium acetate, 0.1M HEPES pH 7.5 and 25% (w/v) PEG 3350) was used for further optimization. The crystallization trials were set up with the protein solution at 1:1 v/v ratio and were incubated at 9°C.

Traditional microbatch-under-oil crystallization trials were set up combining 2µL each of protein and precipitant solutions in a Greiner wellplate (Hampton Research) at 9°C. These were harvested using Mitegen crystal mounts.

4.2.2 Chip materials and fabrication

Hybrid microfluidic chips were fabricated out of a thin PDMS (General Electric RTV 650) fluid layer bonded to both a molded COC (4 mil, 5013 and 6013, TOPAS Advanced Polymers Inc.) control layer and either a 2 mil COC or a 0.5 mil DuraLar (McMaster Carr) substrate. Molding of the PDMS fluid layer was done using standard soft lithographic procedures while molding of COC control layers was done by hot embossing against an epoxy master (Conapoxy FR 1080, 83:100 w/w hardener:epoxy) prepared by soft lithography. Inlet holes were drilled using a 750 µm drill bit. The fluid and control layers were bonded using a chemical surface treatment, no bonding to the substrate was necessary. More details about the fabrication procedure can be found in Chapter 2.

4.2.3 Setup and visualization of crystallization experiments

Crystallization experiments were set up and visualized using either a stereomicroscope (Leica, MZ12.5) with an attached digital camera (Leica, DFC295) operated using Leica Application Suite software or a computer controlled imaging system comprised of an optical microscope (Leica Z16 APO) equipped with an auto-zoom lens (Leica 10447176), a digital camera (Leica DFC280), and a motorized x-y stage (Semprex KL66) controlled by Image Pro Plus (Media Cybernetics). For visualization of protein crystals, images were occasionally taken with the use of a cross-polarizer.
Traditional microbatch-under-oil crystallization trials were set up combining 2 μL each of protein and precipitant solutions in a Greiner wellplate (Hampton Research) at room temperature. These were harvested using Mitegen crystal mounts. Incubation of array chip screening experiments for PhnA was performed in a dark box at room temperature.

4.2.4 On-chip X-ray data collection

Samples were mounted on a modified magnetic goniometer mount (Hampton Research) with an attached metal tube into which a slit was cut and set-screw was used for securing samples. Data collection was performed at room temperature.

Synchrotron diffraction experiments were carried out at the LS-CAT (Life Sciences Collaborative Access Team) facility at the APS (Advanced Photon Source), Argonne National Lab in Lemont, IL. Beamlines 21ID-F and G were used for most of the data collection. Sample-to-detector distances were varied between 150 to 300 mm based on the quality of the crystal present in the chips. Typical data collection was done using a 1° oscillation with a 1 s exposure at an X-ray energy of 12.7 keV (λ = 0.979 Å). For merging slices of data from multiple crystals, 10° of data (-5° to +5° from the normal) was collected from various crystals in the device and an optimal subset of the frames was subsequently merged to obtain a complete dataset. For anomalous datasets, two datasets were collected, one straight on (5° of data, 0° to +5° from the normal) and the other after rotating the chip by 180° (5° of data, -5° to 0° from the normal) to give a total of 10° of data.

Analysis of X-ray diffraction data collected at the synchrotron was performed using HKL2000 software for indexing, refinement, integration, and scaling (HKL Research Inc.). The resolution range of the data was established based on the resolution shell at which I/σ fell above 2.0. Subsequent processing of crystallography datasets was done using the CCP4 suite of programs. MTZ files were generated using Scalepack2mtz from the CCP4 software package. Anomalous data first scaled using SCALEPACK in HKL2000 with the "Scale Anomalous" flag and truncated at Rmerge ≤ 0.25 (resolution 2.75 Å) for obtaining phase information. The scaling data was extended to Rmerge ≤ 0.5 (resolution 2.31 Å) and used for building the complete model. Subsequent processing of crystallography datasets was done using the CCP4 suite of programs. Electron density maps were generated by Phenix and images were generated using PyMOL.
The phase problem was solved using the Phenix.Autosol wizard in Phenix\textsuperscript{57} to locate the selenium atoms of the selenomethionine residues, followed by heavy atom refinement (FOM=0.394). The initial phases were further improved by density modification in the Phenix.Autosol wizard and a model was built with the Phenix.Autobuild wizard with 72% completeness. This was used as a starting model in the Phenix.Autobuild wizard for the 2.11 Å resolution experimental data and a model that was 79% complete was obtained. The model was further extended to 97% completeness using ARP/wARP\textsuperscript{54} and refined using Phenix.Refine wizard to $R_{\text{work}}$ of 0.176 and $R_{\text{free}}$ of 0.211.

4.3 Results and discussion

In the subsequent sections, I will first briefly discuss the X-ray transparent architecture utilized in these crystallization platforms. I then apply these X-ray transparent microfluidic array chips to the screening and optimization of crystallization conditions of PhnA, a novel bacterial hydrolase. Here I validate the utility of these chips for both cryogenic and room temperature on-chip X-ray data collection. In particular I demonstrate the ability of these chips to reproducibly generate a large number of crystals without suffering from variations in crystal quality or non-isomorphism. Finally, I discuss the \textit{de novo} structure determination of PhnA utilizing a completely on-chip approach and discuss the potential for extending this method to other novel protein targets.

4.3.1 Design and fabrication of X-ray compatible microfluidic crystallization array chips

Considerations that must be addressed when designing an X-ray transparent chip are (i) X-ray attenuation and scattering from the chip itself, and (ii) the ratio of signal-to-noise resulting from the diffraction from a protein crystal as compared to the level of background scatter. The first point can be addressed through careful consideration of the composition, properties, and thickness of the materials used, while the second point asks for consideration of the relative path length for diffraction through the crystal compared to the path length through the device materials.

The integration of pneumatic valves on the chips is essential to enable active control for complex microfluidic fluid handling operations for formulation of crystallization trials, and typically requires the use of a flexible elastomer like PDMS\textsuperscript{49,58,59}. Actuate-to-open valves were
fabricated by aligning a control layer made from COC, which defines the location of valves, with
a fluid layer made from PDMS, which defines the location of fluid channels and compartments (Figure 26a). This hybrid structure enables valve actuation while leaving only a thin PDMS valve membrane in the X-ray path, regardless of the total height of the fluid layer. Additional window-like structures present in the COC control layer further decouple the device material thickness from feature height, thus enabling the use of thicker control layers to provide device stability without impacting performance. This device structure is then placed onto a thin film of either COC or the polyethylene terephthalate-based DuraLar to seal the device. A typical device used here requires the X-rays to pass through a 75-µm COC control layer, a 20-µm PDMS membrane, and either a 50-µm COC substrate or a 12-µm DuraLar substrate, resulting in the transmission of 75-78% of the incident X-rays at an energy of 12.4 keV (λ = 1 Å).

The device materials also produce an X-ray scatter ring (Figure 26d) that contributes to the level of background noise. This ring can be minimized by decreasing the thickness of the materials present. Furthermore, the scatter appears at a location where it does not affect the signals needed for high resolution structure determination (q ≈ 0.161 Å⁻¹ (7.5 Å) for PDMS and q ≈ 0.232 Å⁻¹ (5.2 Å) for COC). For example, examination of on-chip diffraction data from a ~100-µm thick PhNA crystal grown on-chip shows that diffraction signals are clearly observable over the background resulting from the chip materials, even at the peak of the observed scattering ring (Figure 26d).

I fabricated 24- and 96-well microfluidic array chips with automated fluid handling capabilities for use in screening crystallization conditions, as well as in subsequent structure determination (See Chapter 2 for further details). These chips utilize actuate-to-open valves to create densely packed arrays of microfluidic wells. A single well on the chip consists of two sets of half-wells, each loaded respectively with protein and precipitant solutions. For a typical 50-µm chamber height this corresponds to ~50 nL per half-well and thus a total chip volume of ~1.4 µL for a 24-well chip and ~5.6 µL for a 96-well chip. This chip design can be configured such that the ratio of protein-to-precipitant varies for each precipitant solution tested (screening phase, Figure 27), or in a manner that all wells are identical (for X-ray data collection) composition. The individual compositions are formulated using three sets of valves: for introducing the protein (shown in blue), for introducing the precipitant (green, Figure 27), and for mixing two adjacent half-wells (pink, Figure 27). Solution droplets are pipetted over inlets, and then the solution is pulled into the rows or columns of wells by actuation of the appropriate
set of valves. Subsequent actuation of the mixing valves between adjacent half-wells creates the individual crystallization compositions. A detailed description of the chip operation can be found in Chapter 2.

4.3.2 On-Chip crystallization screening and structure determination

First, I validated my approach to crystallization screening by comparing the results obtained in these microfluidic chips with those for traditional microbatch-under-oil in a wellplate using the soluble protein lysozyme and the 50-condition Crystal Screen kit (Hampton Research) at room temperature (Chapter 3).

![Figure 27: (a) A 3D exploded view of 3 chambers of an array based microfluidic chip for protein crystallization. The blue chambers in the fluid layer are filled with protein solution and the green chambers with precipitant solution. The ratio of protein to precipitant is varied along the chip to allow for screening of different condition using the same chip. The corresponding valves in the control layer are used to fill in the protein and precipitant respectively using a vacuum line. (b) Design schematic of a 96-well array chip with ratio of the volume of protein to precipitant chambers varying from 1:4 to 4:1 moving vertically down each column.]

I pursued the de novo structure determination of phosphonacetate hydrolase (PhnA) from *Sinorhizobium meliloti*. The first step in structure determination of a new protein is screening against various potential crystallization reagents. Screening was performed using the 96-condition Index Screen (Hampton Research) on a 96-well chip where the ratio of protein-to-precipitant was varied (4:1 to 1:4, Figure 27). Each 96-well screening chip allows 12 different precipitants to be screened against a single protein solution. A total of eight chips were used for screening the entire Index Screen. Each chip requires just 6 µL of protein and 2 µL of each precipitant solution to screen a total of 96 different conditions. Since the chamber volume ratio is varied along the vertical (Figure 27a,b), these chips are screening each of these 12 conditions

81
over 8 different protein-to-precipitant ratios, effectively increasing the experimental range of the experiment. This type of screen, varying the protein-to-precipitant ratio, can be set up automatically on these microfluidic chips, while it would be much harder to implement at the traditional well plate scale where each condition would need to be set up independently. Each of these chips takes < 5 minutes to set up, thus enabling a large number of crystallization trials to be set up in a high throughput fashion. The microfluidic chips are extremely simple to set up, requiring only a pipette to meter the protein and precipitant solutions and a vacuum pump with a Teflon-tubing based connector to fill in the solutions into the chip.

Table 4. Summary of crystallization results of a microfluidic screen of 25 mg/mL PhnA in 20mM HEPES at pH 7.5 against the 96-condition Hampton index screen.

<table>
<thead>
<tr>
<th>Condition</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-16</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>17-32</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>33-48</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>49-64</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>65-80</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>81-96</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 notes:
- No crystals observed.
- + Crystals observed

Screening of a selenomethionine (SeMet) derivative of PhnA (10mg/mL) with the Index Screen gave hits in 21 out of the 96 conditions (Figure 28, Table 4), in comparison to just 3 hits obtained off-chip using traditional methods. The hits obtained in the first round of screening produced showers of tiny crystals, which diffracted to 3Å due to small size of the crystals resulting in poor signal-to-noise (Figures 3b,c). The condition I-80 produced the best crystals in the first round of screening and all subsequent optimization was done on that condition only. The next series of optimization was done with respect to the protein concentration (15, 20 and 25 mg/mL) as well as device geometry. In the first round of screening it was observed that the thickness of the crystals was limited by the height of the microfluidic chamber. This in turn meant that the X-ray beam would pass through more of the chip material than crystal, affecting the signal-to-noise quality of the resultant data. Consequently, in subsequent designs the height of the fluid layer was increased to 100 µm to allow the crystals to grow thicker in the direction of the X-ray beam and improve signal-to-noise. The mixing time between the protein and precipitant was also increased (from 20 minutes to 60 minutes) to allow better mixing. A longer
incubation time for the crystallization trials was achieved by equilibrating the finished crystallization trials within a petri dish with a vial of the precipitant solution, thus preventing dehydration of the chip. The highest quality crystals were grown using a higher protein
concentration (20 mg/mL), taller fluid channels and longer mixing time (Figure 28c).

4.3.3 On-chip phasing via single wavelength anomalous diffraction (SAD)

For novel proteins, once high quality crystals have been grown, obtaining phase information needed to solve the protein structure is one of the key remaining bottlenecks. This is particularly true for targets with low sequence identity to proteins of known structure, where the addition of structural knowledge could provide a wealth of information. The difficulties associated with obtaining phase information from anomalous diffraction methods such as SAD center around the need to collect a large amount of high quality data that is unaffected by either radiation damage or crystal-to-crystal variation. Because of these limitations, anomalous data collection has mostly been limited to traditional single-crystal studies at cryogenic temperatures. Furthermore, the need to observe very small differences in a large signal has thus far precluded the on-chip collection of anomalous data from crystals grown within microfluidic chips.

Table 5. Summary of crystallographic statistics for PhnA crystals grown on-chip

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SeMet [a]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
</tr>
<tr>
<td>Unit Cell Dimensions</td>
<td>a = b = 113.19 Å c = 73.87 Å</td>
</tr>
<tr>
<td>Space Group</td>
<td>P4_2_2_2</td>
</tr>
<tr>
<td>Total Observations</td>
<td>412,491</td>
</tr>
<tr>
<td>Unique Observations</td>
<td>28,002</td>
</tr>
<tr>
<td>Resolution</td>
<td>50 Å – 2.11 Å</td>
</tr>
<tr>
<td>R_{sym}</td>
<td>0.111 (0.508)</td>
</tr>
<tr>
<td>Mosaicity</td>
<td>0.04°</td>
</tr>
<tr>
<td>Redundancy</td>
<td>7.9 (6.7)</td>
</tr>
<tr>
<td>Figure of Merit</td>
<td>0.394</td>
</tr>
<tr>
<td>Completeness</td>
<td>99.8% (99.8%)</td>
</tr>
<tr>
<td>I/\sigma</td>
<td>15.4 (6.6)</td>
</tr>
<tr>
<td># of Frames</td>
<td>188</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
</tr>
<tr>
<td>R (R_{free})</td>
<td>0.176 (0.211)</td>
</tr>
<tr>
<td><strong>Ramachandran Statistics</strong></td>
<td></td>
</tr>
<tr>
<td>Most Favored</td>
<td>95.61% (392)</td>
</tr>
<tr>
<td>Allowed</td>
<td>2.68% (11)</td>
</tr>
<tr>
<td>Disallowed</td>
<td>1.71% (7)</td>
</tr>
</tbody>
</table>

[a] Selenomethionine derivative of PhnA for phasing. Reported values are for all hkls. Values shown in parenthesis represent the value for the highest resolution shell except where indicated. For the Ramachandran statistics the number in parenthesis indicates the number of residues in a given region. R-factor = \( \frac{\sum |F_{\text{obs}}| - k |F_{\text{calc}}|}{\sum |F_{\text{obs}}|} \) and R_{free} is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement. R_{sym} = \( \frac{\sum I_i <I_i>}{\sum I_i} \), where I_i is intensity of the i^{th} reflection and <I_i> = mean intensity.

In a strategy similar to what was applied for the collection of room temperature diffraction data for lysozyme, small wedges of data were collected from 86 PhnA crystals. However, because anomalous diffraction experiments require a comparison of the diffraction signal from
Figure 29. (a) Electron density map and theoretical model of PhnA determined from merged diffraction data collected on-chip. The PhnA carbon atoms are shown in green in stick representation. Superimposed is a 2Fo-Fc Fourier electron density map (contoured at 2.0σ over background in blue). (b) Cartoon representation of the protein. (c) χ² values from merging procedure in SCALEPACK for Friedel mates treated as independent (filled squares) and equivalent (open squares). Data shown is to the resolution of 2.75Å as used in phasing.
symmetry-related pairs of data, an inverse-beam mode of data collection was used. The crystal was rotated 180° after the collection of 5 frames (1s, 1° oscillation) to collect data on the corresponding Bijvoet pairs. This strategy helped to mitigate the effects of radiation damage due to extended exposure to the X-ray beam, and maintained crystal quality during data collection of symmetry-related signals.

Anomalous diffraction experiments require the comparison of the diffraction signals from symmetry-related pairs of data, therefore, an inverse-beam mode of data collection was used. The crystal was rotated 180° after the collection of 5 frames (1 second, 1° oscillation, 19 crystals) to collect data on the corresponding Bijvoet pairs. This strategy helped to minimize radiation damage due to extended exposure to the X-ray beam, and maintain crystal quality during data collection of symmetry-related signals. SeMet PhnA crystals diffracted to a resolution limit of 1.80Å. These small wedges of data were merged to obtain a complete dataset and then used to solve the structure of PhnA to a resolution of 2.11Å using only on-chip data (Table 5). As seen in Figure 29c, the anomalous signal is clearly observed beyond 2.75Å, indicated by the difference in $\chi^2$ for Friedel mates treated as equivalent and independent. Other anomalous indicators including $R_{\text{merge}}$ and $<\Delta F^+/\sigma(F)>$ showed similar patterns (data not shown).

For comparison, PhnA crystals were also grown in a traditional microbatch setup and analyzed using traditional single-crystal cryo-crystallographic techniques. As in the case of lysozyme, the single crystal analyzed via cryogenic methods diffracted to a higher resolution (1.70 Å vs. 2.11 Å), a difference that is probably due to the difference in crystal size. Both datasets were very complete, despite the lack of orientational control available in collecting small wedges of data. Variations in $R_{\text{sym}}$ were within a range typical for good diffraction data, and $I/\sigma$ and structural refinement parameters such as $R/R_{\text{free}}$ and Ramachandran statistics are comparable for both cases. As in the case of lysozyme, analysis of mosaicity shows again that on-chip analysis at room temperature results in a nearly one order of magnitude decrease as compared to single-crystal cryogenic data. The two structures matched to an RMSD < 0.4 Å when aligned in PyMOL. The primary aim of this study is to ascertain that high quality structural information can be obtained from crystals grown on the chip and detailed analysis of the structure and function of this protein is beyond the scope of this work.
4.4 Conclusions

In summary, this is the first example of an entirely hands-free crystallization and *de novo* structure determination for a novel protein target in a microfluidic chip. This achievement was enabled by the coupling of novel approaches for the construction of X-ray compatible multi-layer microfluidic chips with non-traditional strategies for the collection of high quality on-chip X-ray diffraction data at room temperature. The fine control over transport phenomena achieved within microfluidic chips facilitates the reproducible growth of a large number of isomorphous crystals, allowing the collection and merging of small sets of diffraction data from many crystals thereby avoiding the effects of radiation damage associated with room temperature data collection. I validated this strategy for the on-chip collection of single-wavelength anomalous diffraction data for PhnA, a novel bacterial phosphonate hydrolase, resulting in a structure to a resolution of 2.11 Å.

The current paradigm in structural biology has been focused on the collection of X-ray diffraction data under cryogenic conditions from a single crystal. Because of the challenges and variability introduced by the manual harvesting of crystals, all of the data needed would ideally be collected from a single crystal to avoid challenges with crystal-to-crystal variability. Unfortunately these data collection strategies have tended to preclude the use of novel crystallization platforms such as microfluidic chips that are less compatible with cryogenic data collection. Moreover, the traditional approaches lead to wasted time and material in search of crystal forms that are sufficiently robust to withstand X-ray radiation damage. The ability of my microfluidic platforms to generate a large number of reproducible crystals from which high quality on-chip X-ray data can be collected at room temperature has the potential to shift this paradigm, particularly in the case of recalcitrant targets. One can envision overcoming the challenge of collecting anomalous data from crystals that degrade in quality after only a few frames of data are taken by collecting only a single set of symmetry-paired frames from a large number of crystals. Efforts on extending this approach to other novel protein targets are currently underway.

Our microfluidic approach is especially directed towards instances where traditional methods yield protein crystals that are too small or fragile, because it eliminates crucial manual steps of crystal harvesting and mounting that are often the principal obstacles in obtaining good
quality diffraction data. Our approach only requiring a vacuum pump and microscope, while expending minimal amounts of valuable protein and reagents, provides researchers, especially those working on membrane proteins, the capabilities to successfully screen a wide variety of conditions in crystallization space. This ability to screen at nanoliter volumes per condition is typically only available to well-funded, specialized research laboratories that can afford crystallization robots. Our technology enables the opportunity for researchers to re-visit projects where structure determination was abandoned due to the inability of the protein crystals to survive manual harvesting steps, survive cryo-cooling, or radiation damage, and successfully couple them. The platform reported here potentially can also be used to facilitate dynamic structural studies that shed light on the change in structure during protein function, ranging from proteins that respond to stimuli such as light, pH, or the presence of ligands, to those that respond to temperature or electrical changes, thus opening up opportunity for a large variety of previously elusive biological studies.
4.5 References


20. J. D. Ng, P. J. Clark, R. C. Stevens and P. Kuhn, Acta Crystallographica Section D, 2008, 64, 189-197.


56. Schrodinger, LLC, unpublished work.


Chapter 5

Microseeding approaches in crystallization using microfluidic platforms

5.1 Introduction

Advances in structural genomics over the past decade have greatly succeeded in streamlining the entire structure determination pipeline process for proteins. Massive undertakings like the Protein Structure Initiative (PSI) are trying to replicate the success the Genome project had in sequencing, in unearthing the structures of as many novel proteins as possible\textsuperscript{1-3}. Even with a huge financial backing, great improvement in protein expression and purification methods, and the advent of high throughput screening the progress has been excruciatingly slow. There are nearly 600,000 high resolution small molecule crystallographic structures deposited in the Cambridge Crystallographic Data Center (CCDC) as of January 2011, while the number of protein structures deposited in the protein data bank (PDB) is relatively fewer (82224 as of March 2013). This is because of the difficulty associated in obtaining high diffraction quality crystals from protein solutions, and has led to researchers looking into various methods to increase the probability of obtaining crystals as well as improving crystal quality. Seeding is amongst the most dominant of these techniques and has proved extremely successful in unraveling the structure of a large number of proteins\textsuperscript{4-6}.

Protein crystallization is more an art than a science. The fact is that the optimal conditions which lead to nucleation of crystals are not ideal for their subsequent growth. Generation of nuclei is possible when the crystallization solution is supersaturated, while growth of crystals is
Figure 30: Phase diagram of a crystallization experiment with and without addition of crystal seed. There are three main zones in the diagram: undersaturated, saturated and supersaturated. The area under the solubility curve (thick black line) is the undersaturated region where spontaneous nucleation cannot occur and any crystal placed in this region will dissolve. The next region, which is the saturated region, is defined by the experimentally defined solubility curve. Along this curve, spontaneous nucleation will not occur, neither will crystals dissolve or increase in size. The supersaturated region above the solubility curve is divided into three further regions. The metastable zone which is directly next to the solubility curve is where nucleation will not occur, but crystals present in this zone will grow in size. This is the ideal zone for seeding. The second region, the labile zone, is where spontaneous homogenous nucleation can occur and crystals added to this solution will grow in size. Seed crystals added to this solution can induce excessive nucleation which is not desired. The final zone is the precipitation zone, which is extremely supersaturated with respect to crystal growth and will lead to amorphous precipitation. Adapted with permission from J. R. Luft and G. T. DeTitta, Acta Crystallogr. D, 1999, 55, 988-993, http://dx.doi.org/10.1107/S0907444999002085
found to occur at much lower supersaturation levels without further nucleation\textsuperscript{4,5}. It is ideal to separate the nucleation and growth stages of a crystallization trial so as to enable both processes to occur completely (Figure 30). Seeding is an effective technique to separate the nucleation and growth phases of protein crystallization. In this method, protein crystals that have been previously grown are used as seeds and introduced into fresh crystallization trials at lower levels of supersaturation. Based on the size of the seeds, the seeding method can be classified into two classes\textsuperscript{4}:

(a) Microseeding: transfer of seeds that are so small that they cannot be distinguished individually

(b) Macroseeding: transfer of a whole, single crystal, usually 5-50 µm in size

It is also not necessary that the same protein be used to seed for its own growth trials, cross seeding is also practiced where seeds from a different family of proteins are introduced into trials with a different protein (heterogeneous seeding). Also, it is not necessary that protein crystals are always used as seeds. Heterogeneous seeding is a technique where in other materials like minerals, crushed animal hairs, porous nucleants and thin films etc. have been used to successfully induce crystal nucleation\textsuperscript{5}.

Preparation of the seeds for implementation of seeding techniques is a labor intensive process. Macroseeding, also called seed transfer, requires tedious transfer of a parent crystal through multiple washes before it can be used for seeding. It is mostly used in cases where a great increase in the size of the crystal is needed. In most cases however, a whole crystal is not required. In microseeding, submicroscopic crystals, or seeds, are prepared by pulverizing the existing crystals to prepare seedstock, and then diluting these seedstocks for seeding. Streak seeding is a variation of this method, where an animal whisker is stroked against a protein crystal and picks up tiny fragments of the crystal and is then run through a crystallization drop. The problem with microseeding methods is that it is difficult to control the specific amount of seeds in a trial. One of the latest seeding techniques is microseed matrix screening\textsuperscript{7}, in which microseeds from the initial nucleation step are subsequently seeded into various different conditions to induce crystal growth (Figure 31).

Conventional methods for crystallizing proteins (such as vapor diffusion, microbatch) have been very successful in determining the structures of proteins with disparate topologies and
subunit compositions, and from different phyla and membrane types. These methods have their advantages for many proteins and structural biology laboratories such as ease of set up and well-studied experimental protocols. The reason why seeding is not popular yet, is that it has not been able to integrate well into current high throughput crystallization platforms.

Seeding techniques in the past few decades have been focused on manually extracting seeds using horse or cat hairs and transferring them to crystallization drops. However, recently there have been attempts to automate the process with some success. Microseeding is one of the most powerful tools for protein crystallization, but it is rarely used due to the complexity of the procedure and the skill required. It is particularly useful in cases of membrane protein crystallization, specifically the *in surfo* method. It is one of the best methods to accelerate identification and optimization of crystallization conditions for the *in surfo* approach. Many
attempts to simplify the procedure have been reported\textsuperscript{9-13}, however they are still labor intensive and have not significantly reduced the setup time of the microseeding method.

Taking advantage of the two-step process of crystal formation, crystal nucleation and growth, microseeding transfers a submicroscopic crystal from a condition where supersaturation is high to one where the supersaturation is in the metastable zone which aids the growth of crystals (Figure 30,2)\textsuperscript{14}. Considerable research has been done to study the kinetics and dynamics of nucleation and crystal growth, along with many approaches to decouple the nucleation and growth processes to enhance the rate of crystallization success,\textsuperscript{15-24} including a microfluidic droplet-based approach\textsuperscript{25}. During the initial screening for crystallization conditions, often conditions are identified that give rise to microcrystals, spherulites, fine needles, or irregular poorly formed crystals (Figure 32). Microseed matrix screening is a procedure where such microcrystals or spherulites are crushed to form microseeds that then are used to re-screen the initial matrix as nuclei for conditions that support ordered crystal growth\textsuperscript{10,26-28}. The best conditions from the resulting screen are harvested and the process is iterated to produce diffraction-quality crystals. Microseed matrix screening has been used successfully in cases where the best conditions for crystal growth without seeding lead to precipitation,\textsuperscript{10} to accelerate crystal growth of proteins complexed with hydrolysable substrates where the absence of seeding resulted in a hydrolyzed substrate,\textsuperscript{14} to crystallize antibody-antigen complexes,\textsuperscript{11} and many more cases\textsuperscript{29-32}.

![Figure 32: Improvements through seeding. (a) Crystals of CBH\textsubscript{2} resulting from spontaneous nucleation after 1 month. (b) Microseeding with the crystals of (a) into a new drop at lower protein concentration gave crystals that grew singly. (c) Result from the second round of seeding, using the crystals in (b) as microseeds. Magnification is the same in across all figures. Reused with permission from T. Bergfors, Journal of Structural Biology, 2003, 142, 66-76.](image)

D’Arcy et al. have reported an automated matrix microseeding method wherein seeding has been integrated into the screening method\textsuperscript{13}. This method involves adding the microseeds during the screening phase using a standard crystallization robot and has provided some success in
increasing hits in the initial screening runs. However this method is a departure from traditional screening philosophy, where seeds are grown independently in a supersaturated solution, so that they remain out of solution, otherwise they are expected to re-equilibrate and dissolve. To implement this technique in a high throughput manner requires the use of crystallization robots that are at present not accessible to a lot of structural biology laboratories. There is a distinct need for a low cost technology that can allow setting up of seeding trials in a high throughput manner.

Many efforts to simplify microseeding and automate the process have been successfully implemented for various proteins. Microseeding robots pre-mix the protein, precipitant and microseed solution during crystallization set up. Femtosecond laser ablation has been used to eject crystal fragments that serve as seed in the same crystallization drop. Acoustic matrix microseeding utilizes acoustic waves to deliver nanoliter volumes of seed suspension into protein drops. Despite the promise of these improved crystallization success with microseeding, it remains an under-utilized tool and is often chosen as the last resort when all other attempts to grow high quality crystals have failed. The procedure of making microseed stock solutions is the easy part, but introducing the microseed to the crystallization droplet requires skill and experience when performed manually. Microseeding robots can perform routine microseed matrix screening but are not widely available to biochemical labs that would benefit from using this technique.

We recently have succeeded in simplifying the process of microseeding using the X-ray transparent array chips described earlier. Our approach has the strengths of being simple, easily accessible to researchers of any skill level, and incorporates a time variable for seeding that is absent even in the most recent robotic microseeding machines.

5.2 Materials and methods

5.2.1 Protein purification

Photoactive yellow protein (PYP) from *E. Halophila* was expressed as an exogenous product in *E. Coli* and then purified to give holo-PYP that was used for crystallization trials. The protein was supplied by Dr. Marius Schmidt from the University of Wisconsin, Milwaukee. Following is the protocol for the protein purification:
Day 1: 5mL ampicillin containing LB media was inoculated in a disposable culture tube with a single colony from *E. halophila* BN9626 plate. It was then grown overnight (~10hrs) with shaking at 37°C.

Day 2: 100mL of ampicillin containing LB media was inoculated in a 250mL Baffle flask with the above 5mL culture (1 in 20 dilution). It was subsequently grown overnight with shaking at 37°C.

Day 3: 4 × 1L each of ampicillin containing LB media was inoculated in 2L baffle flasks with 20mL each of the 100mL overnight culture (1 in 50 dilution). It was grown with shaking at 37°C until an OD$_{600}$ (optical density measured at a wavelength of 600 nm) of 0.6 – 0.8 is reached. At optimum OD$_{600}$ the flasks were removed from the shaker and kept on ice for 15 minutes to stop growth. In the meantime, the shaker was set at 16°C. The 4 × 1L cultures were induced with IPTG (Isopropyl β-D-1-thiogalactopyranoside) to a final concentration of 1mM. The induction was carried out with shaking at 16°C for 20 - 24hrs.

Day 4: The cells were harvested by centrifugation at 8,000 g for 25 min at 4°C (JA-10 rotor, 500mL bottles). The cells were then resuspended in ~80mL Lysis Buffer and stirred at 4°C for 15min. Lysozyme (200µg/mL) and 2 aliquots (500µL) of protease inhibitors cocktail were added just before sonication at 70% amplitude for 1min (5 – 6 times). To clarify the cell debris, after sonication, the mixture was centrifuged at 22,000 rpm for 20min at 4°C (JA - 25.50 rotors, 50mL tubes. The supernatant was pooled into a beaker and activated pCA anhydride was slowly added while stirring. The beaker was stirred overnight at 4°C.

Day 5: The column was packed with 10mL of pre-charged IMAC Ni2+ resin (take 20mL since resin stored in 50% v/v 20% ethanol). Let the resin settle by adding d$_{H2O}$ (de-ionized water) at regular intervals (do not let the resin dry out). The flow adapter was washed with d$_{H2O}$ using a syringe, making sure there was no air bubbles in the tubing. The flow adapter was attached to the peristaltic pump and the pump was run with d$_{H2O}$ till there were no air bubbles in the tubing. While running the pump, the flow adapter remained attached to the column to make sure there were no air bubbles at the adapter-resin interface.

The column was equilibrated with Lysis buffer (100mL). In the meantime, the protein-pCA mixture was centrifuged at 20,000 rpm for 25min (JA-25.50 rotor, 50mL tubes) to get rid of the excess pCA. The supernatant was pooled into a fresh beaker and labeled as LOAD. The
supernatant was loaded on the column at a steady flow rate. An empty beaker was labeled as FLOW THROUGH (F.T.) and the flow through was collected in it. 15µL of LOAD was collected for SDS-PAGE analysis. After loading the column, it was washed with Lysis buffer (50mL) followed by Wash buffer (50mL) and the flow through was collected in a fresh beaker labeled as RINSE. The protein was eluted with Elution buffer and the elute was collected in a fresh beaker. After elution, the column and tubings were washed with 20% ethanol and dH2O. 15µL of F.T., RINSE and ELUTE were collected for SDS-PAGE analysis.

The purification was analyzed by running a SDS-PAGE gel. The ELUTE was dialyzed against Overnight Dialysis buffer -1 in 2L graduated cylinder to remove excess Imidazole.

Day 6: The ELUTE is concentrated to ~ 10mL in Millipore Amicon 10K tubes. The protein concentration was checked using spectrophotometer and the Abs 280nm/446nm ratio was recorded. 15µL of concentrated ELUTE was collected for SDS-PAGE analysis. Enterokinase was added to the concentrated ELUTE and rocked overnight (16 hours) at room temperature to cleave the His tag.

Day 7: The Ni2+ affinity column was washed with dH2O and equilibrated with Lysis buffer. The elute+enterokinase mixture was loaded and the F.T. was collected in a fresh beaker. 15µL of cut concentrated ELUTE was collected for SDS-PAGE analysis. Ideally, all the cut PYP will flow through while the His tag fragments and uncut PYP will cling to the column. The column was washed with the lysis and wash buffers and the flow through were collected as WASH. The column was further washed with Elution buffer to get rid of the uncut PYP from the column. 15µL of F.T. and WASH were collected respectively for SDS-PAGE analysis. A SDS-PAGE gel was run to analyze the extent of cleavage. The PYP was concentrated to ~ 1mL in Millipore Amicon 10K tubes for Ion-Exchange Chromatography next day.

Day 8: The Ion-Exchange Chromatography was run using AKTA FPLC system. The tubings were washed with Buffer A (20mM Tris at pH 8.0) and Buffer B (20mM Tris, 1M NaCl at pH 8.0) to get rid of any air bubbles. The column was washed with Buffer A for 10 – 15min and the loop was washed with Buffer A as well.

The tubes that appeared yellow were labeled and absorption spectrum for each of them was measured. The fractions corresponding to the first peak in the elution profile were pooled. The fractions corresponding the shoulders of the peak were avoided. The second peak may represent
apo-PYP i.e. without the chromophore. This explains the absorption spectrum for the fractions representing the second peak where the Abs 280nm/ 446nm is relatively higher compared to the first peak fractions.

If the peak separation is bad and the majority of protein still exists as holo and apo- pyp mixture, another ATKA can be run with smaller fraction volume and longer gradient to improve the resolution.

Day 9: The pooled fractions were concentrated to 15 – 30 mg/mL in Millipore Amicon 10K tubes. They were dialyzed overnight against the final storage buffer at 4°C. The final protein was then filter sterilized. Aliquots of 20 µL aliquots were flash frozen in liquid nitrogen and store at -80°C.

5.2.2 Protein crystallization

Vapor diffusion (hanging drop) trials were set up for crystallizing PYP, the precipitant used was 2.8M ammonium sulfate and 0.9M sodium chloride at pH 7.036. Trials were setup at room temperature. Crystals were obtained in the P63 space group within a week. The crystals however grew in bunches (Figure 33) and single, isolated crystals were not obtained. This is a well-known problem in crystallizing PYP. The crystal clusters were harvested and used for microseed preparation.

5.2.3 Microseed preparation

Microseed stock solution was prepared using a tissue homogenizer and seeds harvested from hanging drop experiments described in section 2.2. Crystals were harvested from the hanging drop and placed carefully in the glass tube of the homogenizer. Seed stock was made in the same condition that produced crystals (Ammonium sulfate, pH 7.0 and NaCl), at a 20% higher concentration than what induced crystallization so that the crystals do not dissolve. The crystals were then pulverized using the homogenizer34,37, until no crystal fragments could be seen in the glass tube under the microscope. This formulation became the stock solution for the PYP seeds from which serial dilutions were made with the stock solution. The dilutions made were 1:1, 1:2, 1:10, 1:100, 1:1000 and 1:10000, with each successive dilution containing fewer and fewer nuclei. Figure 34 describes the procedure of making seed stock solutions.
Microfluidic chip fabrication

Fabrication of the microfluidic chips used for microseeding was carried out using the modified approach which eliminates hot embossing. Briefly, PDMS mixed with the crosslinker in ratios of 5:1 and 15:1 was spun onto control layer (CL) and fluid layer (FL) silicon masters. A thin sheet of 2 mil COC was plasma bonded onto the partially cured PDMS on the CL master and baked for complete adhesion. After baking at 60°C for 1 hour, the COC sheet with the CL was peeled off, and aligned over the FL master. This combined assembly was baked at 60°C for 4 hours to allow the PDMS layers of different ratios to bond together. Finally the combined assembly was peeled of the FL master, holes were drilled for the CL and FL inlets, and the chip was placed on a flat COC substrate to complete the assembly.
5.2.5 Setting up and visualization of seeding trials

Reagents were filled in the microfluidic chip using a vacuum pump coupled with a modified manifold with Teflon tubes and PDMS blocks to apply vacuum on the inlets. The various reagent solutions (protein, precipitant seed stock etc.) were pipetted onto the respective inlet holes and sucked in through vacuum. The solutions in the adjacent chambers were mixed using vacuum actuated valves between the chambers. After setting up of the trials, the chips were placed in sealed petri-dishes and incubated at room temperature.
Visualization and setup of the crystallization experiments were done with the help of a stereomicroscope (Leica, MZ12.5) with an attached digital camera (Leica, DFC295) operated using Leica Application Suite software or a computer controlled imaging system comprised of an optical microscope (Leica Z16 APO) equipped with an auto-zoom lens (Leica 10447176), a digital camera (Leica DFC280), and a motorized x-y stage (Semprex KL66) controlled by Image Pro Plus (Media Cybernetics). Images were periodically taken with the help of a cross-polarizer.

5.3 Results and discussion

Seeding during crystallization trials is a technique to improve the chances of getting diffraction quality crystals, when normal crystallization methods are yielding poor results. Here, we demonstrate how we can use the array chip developed for screening and on-chip structure determination for microseeding trials.
Figure 36: Second round of microseeding trials carried out in 24-well array chip. The precipitant inlets were filled with 2.8M(NH₄)₂SO₄, while the usual protein inlet was filled with a mixture of PYP (20 mg/mL) and seeding solutions in 3M(NH₄)₂SO₄, different dilutions were used in different chips.

Seeding solution (numbers in red indicate dilution)

1:1
1:4
1:100

No dilution
1:2
1:10

Figure 37: Third round of microseeding trials carried out in 24-well array chip. The usual precipitant inlets were filled with seed stock in 3M(NH₄)₂SO₄, with different dilutions introduced in each of the six lines as shown in the figure. PYP (20 mg/mL) plus the precipitant (2.8M was (NH₄)₂SO₄) was filled in the protein inlet.
5.3.1 Microfluidic chip for seeding trials

The microfluidic chip described in the previous chapters is used as is for the microseeding trials. The advantage in using the same chip is that the procedure for fabrication and setting up trials was already well laid out and we could modify it to use it for seeding trials. The seedstock solution was prepared from crystals of PYP grown in hanging drop trials as described in the Materials and Methods section. In traditional seeding experiments, the seed is physically introduced by adding a hair or whisker with seeds already present on them and allowing crystals to grow along the hair. However in order to automate the seeding process and make it more user friendly, we had to devise a way to introduce seeds into the microfluidic wells. The array chips we are using here have two half wells in each crystallization chamber on-chip; one houses the protein and the other the precipitant. Here we have three solutions to fill into the chip, the protein, precipitant, and, the seeding solution. We could have added a third chamber to house the seeding solution, but that would have introduced complexity into the device in terms of filling and inducing mixing. This is something to add when we are looking to optimize the seeding process on-chip, but for proof of principle studies, we wanted to utilize an identical chip to see whether seeding was possible or not.

The 24-well array chip used here has two sets of inlets, one for the protein and one for the precipitant, and there are three sets of solutions we need to fill in. There are three combinations we tried on-chip, first adding the seed solution with the precipitant and adding it to the precipitant inlet and adding the protein solution as before through the protein inlet (Figure 35). A dilution series of seedstock solution was prepared (see Materials and Methods) with each series containing different number of nuclei, and each of the six inlet lines contained a different dilution of the seed stock solution. The second combination (Figure 36) we tried was adding the seed solution along with the protein solution and the precipitant solution was added as is in the precipitant inlets. This strategy meant that we needed a whole chip for each dilution we wanted to screen. To get around this, we can also switch the protein and precipitant filling inlet, and fill the same protein with different dilution series of seed stock in the six (usual) precipitant inlets and fill the precipitant through the (usual) protein inlet. The third combination (Figure 37) involves premixing the protein and precipitant and adding them through the protein inlet and
adding the seed solution (in dilution series) through the precipitant inlets. On close inspection, the third combination is the one which closely mimics the setting up of a conventional seeding trial where the seed solution is introduced after the protein and precipitant have been mixed.
(generally at a lower supersaturation which favors growth rather than nucleation). However, to be thorough we tried all three combinations possible.

The results of the initial seeding trials (Figure 38) show that the third combination was the most successful one. Showers of tiny crystals were obtained in the first and second combination as well, but there were no large, isolated crystals nor was there any change in crystal morphology when changing the dilution of the seed stock solution. In stark contrast, crystals were obtained in the third combination and, there was a clear differentiation between the crystals between different dilution series of the seed solution. Figure 38 clearly shows the variation in the protein crystals from a shower of tiny crystals at low dilution, which is expected as these contain a huge number of nuclei, to large, isolated crystals at higher dilutions that contain far fewer nuclei to begin with. The dilution series of 1:10 gave the best crystals in multiple trials, consistently producing single, isolated crystals, and were used for all further experiments involving PYP.

5.3.2 On-chip X-ray analysis of PYP crystals grown via microseeding

Once crystals were obtained, the next step was to ascertain the quality of the crystals grown via seeding. Taking advantage of the fact that the array chips allow for on-chip data collection, the chips were taken to LS-CAT for diffraction data collection. A similar strategy to what was used earlier for the case of PhnA was used for data collection. Multiple chips were setup with identical 1:10 dilution of the seeds in all chambers thus allowing for a large number of crystals of similar morphology to be grown on a single chip. The chips were mounted on the beamline as
described in the earlier chapters, and wedges of data were collected from each of the crystals at room temperature. Table 6 lists the important crystallographic parameters from the diffraction data collected. The maximum resolution diffraction spots obtained from a PYP crystal was 1.19 Å, indicative of the high quality of the crystals obtained, while the final structure was solved to 1.32 Å. Figure 39 shows on-chip diffraction data obtained at room temperature from a PYP crystal with the inset showing the high quality of the diffraction spots and good signal-to-noise of the diffraction data. The main purpose of crystallizing PYP was to use it as a model system for carrying out time-resolved studies using a microfluidic platform. Therefore, the data of PYP collected on the monochromatic beamline was not analyzed to full extent in terms of excluding frames to bring down the mosaicity, but it was still comparable to previously collected on-chip data. Integration, merging and scaling of the X-ray data was done using HKL 2000 (HKL Research)\textsuperscript{38}. Subsequent processing of the data was done using the CCP4 crystallography suite\textsuperscript{39} as described in earlier chapters. Molecular replacement for PYP was done using the 1OTA\textsuperscript{40} structure from the RSCB Protein Data Bank. The final model (Figure 40) was made using ARP/wARP\textsuperscript{39}.

Table 6. Summary of crystallographic statistics for PYP crystals grown on-chip

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Data Collection</th>
<th>PYP data\textsuperscript{[a]}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit Cell Dimensions</td>
<td>(a = b = 66.76) Å (c = 40.89) Å</td>
<td></td>
</tr>
<tr>
<td>Space Group</td>
<td>P6(_3)</td>
<td></td>
</tr>
<tr>
<td>Total Observations</td>
<td>804,824</td>
<td></td>
</tr>
<tr>
<td>Unique Observations</td>
<td>24,364</td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>50 Å – 1.32 Å</td>
<td></td>
</tr>
<tr>
<td>(R\text{_{sym}})</td>
<td>0.111 (0.508)</td>
<td></td>
</tr>
<tr>
<td>Mosaicity</td>
<td>0.02 – 0.54(^{\circ})</td>
<td></td>
</tr>
<tr>
<td>Redundancy</td>
<td>7.9 (6.7)</td>
<td></td>
</tr>
<tr>
<td>Completeness</td>
<td>99.9% (98.4%)</td>
<td></td>
</tr>
<tr>
<td>(I/\sigma)</td>
<td>11.1 (2.3)</td>
<td></td>
</tr>
<tr>
<td>Refinement</td>
<td>(R(R_{\text{free}})) = 0.3228 (0.372)</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{[a]} Reported values are for all(hkl)s. Values shown in parenthesis represent the value for the highest resolution shell except where indicated.
5.4 Conclusions

We have successfully demonstrated the ability to use the microfluidic array chips to optimize crystallization of photoactive yellow protein (PYP) by screening for the optimal microseed dilution necessary to seed and obtain a single crystal in each crystallization chamber. In our lab, we have also had initial success in screening for crystallization of succinate quinone reductase (SQR) from *C. maquiltingensis* and cytochrome *c* oxidase. The current chips are designed to screen for crystallization conditions using a two-component approach: the protein being crystallized and the precipitant being screened. In future, the chip design can be altered to include a third chamber for the microseed solution, giving greater control over the crystallization trial. In addition, we will incorporate on-chip serial dilution to enable efficient, reproducible microseed matrix screening. It is estimated that the optimized chip will consume 5-20 nL protein, 10 nL microseed solution and 5-20 nL precipitant per crystallization well, in a design where the protein: precipitant ratio is varied for each precipitant. Upon optimization of crystallization conditions, X-ray data collection can be completed without manual handling of the crystals as demonstrated for the array based chips. We plan to use new 3-compartment chip designs for crystallization of membrane proteins of the heme-copper oxidase superfamily and the membrane associated members of the electron-transport chain.

Seeding is a powerful tool to optimize the chances of obtaining high quality crystals of targets that do not yield results with conventional methods. With the array based chip we have implemented microseeding on a target protein (PYP), which does not crystallize using standard crystallography methods. The fine control over mixing that microfluidics provides, can allow for precise mixing of the seed solution with the protein and precipitant at a given time, thus making it easier to control the time and duration of seeding. These give our chips added advantage in terms of flexibility and optimization, making it a better platform for seeding and crystallization.
5.5 References


Chapter 6

Microfluidic approaches for time resolved studies of proteins

6.1 Introduction

While in situ diffraction analysis of protein crystals grown in microfluidic devices has been reported on previously, these efforts have been focused on the more mainstream approach of monochromatic X-ray diffraction. However, these studies produce only a static picture of the protein structure and may be affected by structural artifacts resulting from data collection at cryogenic temperatures. Laue crystallography, takes advantage of polychromatic X-rays, frequently performed at room temperature, to increase the rate of data collection and thereby facilitate dynamic structural and functional studies. However, these methods are very sensitive to crystal quality and are even more damaging to the crystal than monochromatic X-ray diffraction. X-ray compatible microfluidic crystallization platforms can be used to generate a large number of high quality crystals which can be analyzed in situ without the need for manual handling. These capabilities facilitate the analysis of many crystals in order to overcome difficulties associated with the severe radiation damage typical of Laue methods. This chapter presents a method for coupling these microfluidic crystallization platforms with automated procedures for crystal centering, data collection, and the subsequent analysis of polychromatic data taken from many crystals to minimize the effects of radiation damage on the resultant data. As proof-of-concept, single polychromatic diffraction images from multiple lysozyme and PhnA crystals grown on-chip were collected and merged into a single dataset. Following this, laser mediated on-chip studies of the photo cycle of photoactive yellow protein (PYP) were carried out, being the first instance of on-chip time resolved studies.

X-ray diffraction studies of biomolecules over the last few decades have resulted in a wealth of information about three dimensional, static structures of proteins, DNA and RNA molecules. These have provided significant insights into the structure and function of these molecules. However, a precise understanding of how a protein functions requires not only its three dimensional static structure, but also the subtle conformational changes that takes place during
protein function\(^1\). Therefore to fully understand how these biomolecules function, it is necessary to watch them function in real time, along a reaction path that often times involve short-lived reaction intermediates. Many such reactions are possible in crystals as they typically contain a large percentage of solvent. The solvent forms channels and hydration shells around protein molecules and facilitates dynamic processes such as diffusion and binding of substrates and other ligands, turn-over in enzyme crystals, and conformational change in response to absorption of light in photoreceptors\(^2\). Time resolved X-ray crystallography (TRX) is a unique tool that facilitates the study of these structural intermediates allowing the real time investigation of structural changes that molecules undergo while performing their function. TRX can, therefore, play an important role in visualizing structures of reaction intermediates thereby elucidating reaction mechanism\(^2\).

TRX is a term used in a broad sense to describe methods that are used to study the structure of intermediates that a macromolecule goes through during a chemical reaction. These approaches can be divided into two main classes, in the first case; the lifetime of the intermediate to be studied is artificially extended to match the typical X-ray data collection time available from monochromatic synchrotron sources. This is accomplished by generating a steady state accumulation of the intermediate in question by continuous reaction triggering, or by physical trapping of the intermediate using fine cryo-temperature control\(^1,3-5\). Trapping methods extend either the lifetime of the intermediate or increase their concentration, thus making it amenable to probe these structures using monochromatic radiation. Although trapping methods provide valuable insight into structures of intermediates, they at the same time perturb the reaction they are probing, thus making it necessary to consider the artifacts which these methods may have induced. In stark contrast, under the second approach, instead of extending the lifetime of the intermediate, the X-ray exposure time is matched to the lifetime of the
intermediates. This approach allows us to observe the reaction in its natural state, without any external perturbations that might have side effects to the changes we have to be considered. This approach is based on employing Laue crystallography, where X-ray diffraction data is obtained.
by exposing a static crystal to a polychromatic X-ray beam (“pink” beam) at room temperature, contrary to the standard X-ray crystallographic set-up where a cryo-cooled, rotating protein crystal is exposed to a monochromatic beam. In this way, exposure times are shortened by three to four orders of magnitude relative to monochromatic experiments, opening access to the microsecond time domain at which most reactions occur. A microsecond is about the time it takes for an electron bucket to circulate around a synchrotron ring, thus making it possible to take advantage of the time-structure of synchrotron radiation to further reduce the effective X-ray exposure time: if the full synchrotron current is loaded into a single electron bundle, each ~100 ps X-ray pulse generated by this bundle becomes sufficient to provide an interpretable X-ray diffraction pattern. Thus Laue crystallography is a more direct, yet more challenging method to study reaction intermediates and understand biomolecule function as it evolves at room temperature. This chapter focuses on Laue crystallography as a method to better understand protein function.

In TRX experiments, supplying the trigger for a chemical reaction to occur is a key step to probe structural changes in proteins. The reaction being studied is triggered in the crystal and X-rays are used to probe the structural changes at various time delays after the start of the reaction. The time resolution is determined by either by the duration of triggering process or the X-ray probe pulse, whichever is longer. Ideally, the trigger should be applied in a time resolution which is much smaller than the time of the X-ray pulse and the lifetime of the intermediate. A very high X-ray flux is needed to obtain diffraction data of sub-nanosecond resolution, which is possible at the latest X-ray synchrotron facilities (Advanced Photon Source (APS) USA, European Synchrotron Radiation Facility (ESRF) France, and Spring 8 (Japan)). It is impossible to record the integral intensity of reflection in such short exposure times using monochromatic X-rays, polychromatic X-rays are needed in such cases. It is ideal if the reaction being studied is reversible and the system restores to the original state after a short period of time, thus allowing complete data collection from a single crystal. If the reaction is irreversible, the data needs to be collected from a new crystal each time.

There are many challenges associated with Laue crystallography. The sub-microsecond time resolution comes at a price – the pink, polychromatic X-ray beam is extremely intense, and causes excessive radiation damage to protein crystals and is susceptible to induce unwanted reactions and chemistries in biomolecules. Since most of Laue data collection is done at room
temperature radiation damage is even more severe and a big issue. The major reason why Laue studies are done at room temperature is because in most cases, the experiment of interest is observing a chemical reaction or structural change at biologically relevant temperatures. Cryocooling in monochromatic X-ray data collection is used for reducing radiation damage, but if used in Laue data collection it has the potential to freeze out potential protein dynamics which may be of interest as well as inducing artifacts in the crystal structure that affect the quality of data collected.

The second major issue is that of simultaneous triggering of all the molecules in the protein crystal, to ensure that the majority of the molecules are activated and follow the same reaction path in a synchronous manner. When conducted at the sub-nanosecond level, Laue experiments are also affected by low signal-to-noise, because short exposure times mean fewer available X-ray photons, and small exposure times result in small structural changes which are hard to detect. This is directly linked to the difficulty in data analysis of Laue diffraction data. Laue data is also extremely sensitive to the quality of the protein crystal, a property measured by the mosaicity of the crystal as well by substantial background scattering generated by the polychromatic radiation.

The sensitivity of Laue crystallography to crystal quality, and in particular mosaicity is the result of spot elongation during data collection. During data analysis it is necessary to identify the exact location of the various diffraction spots observed. However, with an elongated or streaky spot identification of the spot location is much more difficult. This problem can be further compounded by high levels of background scatter which decrease the signal-to-noise ratio for weak spots. These limitations have typically restricted Laue data collection to crystals of relatively high quality and provided an opportunity for a disruptive innovation in this field.

Efforts in microfluidics for protein crystallization have advanced dramatically in the past ten years including a variety of strategies for performing in situ X-ray crystallographic analysis of crystals grown on-chip, however the efforts thus far have been limited to the use of monochromatic X-ray diffraction studies. Monochromatic X-ray diffraction is used for the vast majority of structural biology studies, but these results are limited to a static analysis of the protein structure and may be affected by structural artifacts resulting from data collection at cryogenic temperatures. Laue or time resolved crystallography relies on exposing a single
stationary crystal to a polychromatic X-ray beam, contrary to monochromatic methods where a monochromatic beam of X-rays is shot at a rotating crystal\(^1\). Exposure times in Laue are three to four orders of magnitude shorter than traditional methods, providing access to microsecond time domains in third generation synchrotron sources. This allows us to probe the structures of intermediates while a protein is undergoing structural changes. Microfluidic platforms have immense potential in Laue crystallography because they can address many of the challenges in the field today.

The array chips described in previous chapters have fine control over transport at the microscale and can be used to generate a large number of isomorphic crystals in each trial. This allows for data collection from multiple crystals simultaneously, thereby reducing radiation damage, without having to mount each crystal individually thus making the process much more user friendly. Using microfluidic platforms it is also probe reactions which are irreversible much more easily as multiple crystals can be analyzed in a stress-free manner. I have already demonstrated that the data collected from crystals grown on microfluidic array chips have an order of magnitude less mosaicity than traditionally grown crystals, which make it a suitable candidate for Laue experiments.

### 6.2 Materials and methods

#### 6.2.1 Device fabrication and operation

24-well and 96-well array chips (Figure 41a) used here were fabricated and filled as described previously\(^7\). Briefly, microfluidic chips were fabricated by bonding a thin polydimethylsiloxane (PDMS, General Electric RTV 650) fluid layer with either 25 or 50 \(\mu\)m features covered with a membrane thickness of ~20 \(\mu\)m. This fluid layer was chemically bonded to a molded cyclic olefin copolymer (COC, 2 mil, 4 mil, 5013 and 6013, from TOPAS Advanced Polymers Inc.) control layer and a flat COC substrate.

#### 6.2.2 Protein and precipitant solutions

Hen egg white lysozyme (Sigma) was dissolved in 50 mM sodium acetate (Sigma-Aldrich) at pH 4.6 with 20\% (w/v) glycerol (Fisher Scientific) at a concentration of either 88 mg/mL for simple large well experiments or 120 mg/mL for 96-well array chip experiments. Lysozyme concentrations were determined by UV absorbance measurements (Lambda 650 UV-Vis
spectrophotometer, Perkin Elmer) at 280 nm using an extinction coefficient of 2.64 mL/(mg-cm)$^8$. A precipitant solution of 1M NaCl (Aldrich) in 50 mM sodium acetate, pH 4.6 with 20% (w/v) glycerol was prepared. Prior to setting up a crystallization experiment, protein solutions were filtered through 0.1 µm (Ultrafree-MC, Millipore) filters. Precipitant solutions were filtered through 0.22 µm (Steriflip, Millipore) filters.

Selenomethionine labeled PhnA from Sinorhizobium meliloti was dissolved in 20 mM HEPES, pH 7.5, at a concentration of 20mg/mL. From screening experiments with the 96 condition Hampton Index Screen (Hampton Research), the condition I-80 (0.2M ammonium acetate, 0.1M HEPES pH 7.5 and 25% (w/v) PEG 3350) worked best was used for all further crystallization trials. The crystallization trials were set up with the protein solution at 1:1 v/v ratio and were incubated at 9°C.

Photoactive yellow protein (PYP) from E. Halophila was expressed as an exogenous product in E. Coli and then purified to give holo-PYP that was used for crystallization trials. Vapor diffusion (hanging drop) trials were set up for crystallizing PYP (for obtaining seeds), the precipitant used was 2.8M ammonium sulfate and 0.9M sodium chloride at pH 7.0$^9$. Seedstock solution was prepared as described in Chapter 5 and the same crystallization condition was used to setup trials in the array chips. A dilution of 1:10 of the seedstock solution was used in all cases. The chips were incubated at room temperature.

6.2.3 Crystallization experiments

Crystallization experiments were set up and visualized using either a stereomicroscope (Leica, MZ12.5) with an attached digital camera (Leica, DFC295) operated using Leica Application Suite software or a computer controlled imaging system comprised of an optical microscope (Leica Z16 APO) equipped with an auto-zoom lens (Leica 10447176), a digital camera (Leica DFC280), and a motorized x-y stage (Semprex KL66) controlled by Image Pro Plus (Media Cybernetics). For visualization of protein crystals, images were occasionally taken with the use of cross-polarizers.

Filling and operation of the microfluidic devices (Figure 41a) was achieved by the use of actuate-to-open valves$^{10,11}$. A small vacuum pump (Gast) connected to the device through a plastic gas manifold (Cole-Parmer Instrument Co.) and 24 AWG PTFE tubing coupled with a thin metal tube. Fluids (protein and precipitant solutions) were supplied to the device by
pipetting 1 – 5 µL of solution onto the inlet hole. Vacuum within the chip from actuation of
valves then pulled fluid into the chambers. After filling, the inlet holes for both the fluid and control layers were sealed with Crystal Clear tape (Hampton Research).

6.2.4 X-ray diffraction experiments

Diffraction experiments were carried out at beamline 14-ID-B at the Advanced Photon Source at Argonne National Laboratories with the assistance of the BioCARS team. 12-14-ID-B is an insertion device station with two in-line undulators (U27 and U23) that can operate over a wide energy range of 7 keV to 20 keV (1.77 Å to 0.62 Å) and can provide both a monochromatic and polychromatic X-ray beams. Data was collected with both undulators set to a peak energy of 12 keV (1.03 Å). Experiments were also performed with the energies of the undulators offset by 2 keV (U27 at 11.8 keV (1.05 Å), U23 at 13.8 keV (0.898 Å), thereby increasing the polychromatic range of the exposure and increasing the density of diffraction data obtained per image. A MARCCD-165 detector was used, with optical visualization provided by off-axis cameras.

For the collection of multiple frames from a single crystal, the microfluidic chips were mounted using a set-screw into a slit cut in a metal tube. This tube was then glued to a standard magnetic goniometer mount. X-ray exposures of 1 or 10 pulses of 11 electron bunches each were used. A sample-to-detector distance of 100 mm, 120 mm, or 150 mm was used depending on the strength of the diffraction signal observed. Collection of a complete dataset from a single crystal was attempted. 90° of data was collected using either a single pass method in steps of 8°, or in a gap filling mode in steps of 6°. All data collection was performed at room temperature.

For the collection of single-shot data, a 96-well array chip was mounted in a plastic frame attached to an automated x-y-z translational stage (Eastern Air Devices) (Figure 42b). An automated python script coupled with two off-axis viewing cameras was used to identify and mark the location of individual crystals within the array chip (Figure 42c). Once identified, single-shot data collection on each crystal occurred in an automated fashion. X-ray exposures of 10 pulses of 11 electron bunches were used. A sample-to-detector distance of 110 mm was used. All data collection was performed at room temperature.

Collection of time resolved data on PYP involved shooting the crystal with a laser pulse, before striking it with the polychromatic X-ray beam at fixed time intervals to study the photocycle. Sample to detector distance was maintained at 150mm. Wavelength used was 1.02
– 1.16 Å (undulators peaked at 12keV). A nanosecond laser at 485 nm with a spot size of approximately 215 microns ad density of 4.8 mJ/mm² was used, with the duration of a single pulse being 7 nanoseconds. A dark image was recorded initially, then after the laser had been shined images were recorded at 500 nanoseconds, 2 microseconds and 1 millisecond.

At the time of performing the time resolved data collection, the synchrotron was operating in the standard 11 bunch mode (1 pulse of X-rays contains 11 bunches). The time for an entire 11 bunch pulse to go through is about 1.5us, which was too slow for our purposes. For static data collection we had found that 10 or so 11 bunch pulses gave great signal to noise. However, since we needed faster time resolution we isolated a single bunch out of the pulse-train and then accumulated 80 or so laser pulse/probe sequences to generate a single frame of data. While we only needed to do this accumulation strategy for the two fastest time points, it was easier to just keep doing it this way for all of our data, rather than switching back and forth from single bunch to a complete pulse-train.

The samples were mounted directly on the goniometer using the standard pin mount. We aligned the sample at 30 degrees, straight-on with the high resolution camera. We tried to do automated rotation for data collection from 20-70 degrees but had trouble with the sample going out of the beam. Instead, we aligned and tested at each point. Data was collected in 5 degree increments from 20-70, but typically only 20-35 ended up being used in the data refinement.

6.2.5 Analysis of X-ray diffraction data

Analysis of Laue X-ray diffraction data was performed using Precognition analysis software and Epinorm for indexing, geometric refinement, integration, and scaling (Renz Research Inc.). C-shell scripting was used to automate the analysis of single-shot diffraction data from multiple crystals. The resolution range for the data was established based on the level at which the completeness in the highest resolution shell fell below 25%. Subsequent processing of crystallography datasets was done using the CCP4 suite of programs. Electron density maps were displayed using COOT. Molecular replacement for lysozyme was done using PDB structure 193L as a model. All images were indexed and refined separately before merging. The resulting .hkl file was converted to .mtz format using f2mtz (CCP4). Data refinement was done using Phenix.refine (PDB structure 3SZY for PhnA). Simulated annealing was used to improve map quality.
6.3 Results and discussion

6.3.1 Microfluidic platforms for Laue crystallography

Time resolved or Laue crystallography is extremely sensitive to crystal quality, and to ensure the clear capture of a protein molecule in motion it is essential that high quality diffraction data be collected. It is also not always possible to collect all the data from a single protein crystal because of reaction limitation discussed previously, and hence it may be required to loop and mount hundreds of crystals in order to collect a complete dataset. Since we are interested in looking at the protein function at biologically relevant temperatures, Laue crystallography is more often than not carried out at room temperature.

A combination of the above factors make the microfluidic array chips described in previous chapters a promising alternative to collect time resolved data. We have demonstrated the ability to grow a large number of isomorphous crystals on-chip and have collected high quality diffraction data to solve protein structures. The mosaicity of crystals grown on-chip is an order of magnitude lower than conventionally grown crystals, a further indicator of high quality crystals. Since we can collect wedges of data from multiple crystals, in effect each crystal is only exposed to the radiation for a short amount of time, thus allowing for room temperature data collection. The major concern remains whether the signal-to-noise resulting from the background scatter of the chip material and the signal from the relatively smaller crystals hampers on-chip Laue data collection.

To assess the effects of radiation damage and background scatter for polychromatic Laue data collection we first grew lysozyme crystals in the microfluidic chip. We then tested different data acquisition strategies in an attempt to collect complete single-crystal Laue datasets in situ. Data was first collected in a gap-filling mode at 6° increments with the energy of the X-ray beam peaked at 12 keV. While high quality diffraction spots were observed initially (Figure 43a), the effects of radiation damage on the relatively small 25 µm thick crystal used in this experiment were severe, quickly resulting in poor quality data as evidenced (evident) by the streaky spots and the loss of higher resolution diffraction spots (Figure 43b). Of the 31 diffraction images originally intended to comprise the complete dataset, only the first 7 proved to be usable for subsequent data analysis, resulting in an incomplete dataset.
An alternate data collection strategy was also tried. Data was also collected with the energy of the two undulators offset by 2 keV (11.8 keV (1.05 Å) and 13.8 keV (0.898 Å)). This offset resulted in an X-ray beam with a wider range of energies and thus a higher density of diffraction.
spots per image. As the previous attempt at data collection indicated that the lifetime of the crystal was not long enough to take advantage of the gap-filling data collection strategy we used simple data collection with a larger increment of 8° to try and cover a larger range of diffraction space. Again, high quality diffraction was observed initially (Figure 43c), but radiation damage allowed for the analysis of only 9 of the 12 frames collected. However, the higher density of data per frame obtained from using the offset undulators resulted in a higher overall completeness of the dataset and an improvement in the resolution.

While the level of background scatter for the microfluidic chip architecture used in these experiments had been tested previously for monochromatic data collection, the use of a pink X-ray beam had the potential to increase the level of background scatter above that observed previously. The size of crystals is limited by the height of the microfluidic chamber, resulting in a path length through the crystal of ~25 µm for the chips used here. This path length is relatively short in comparison to the ~145 µm path length of the beam through the device materials for devices fabricated from 2 mil 5013 COC (total thickness includes ~ 20 µm of PDMS. Background scatter from the device materials can be clearly seen as a diffuse ring in the lower resolution shells of the diffraction images. However, the intensity of this scatter did not occlude the clear observation of subsequent diffraction spots. While this ratio of the diffraction path through the crystal compared to background through the device proved to be adequate for the strongly diffracting crystals of lysozyme used in these experiments a longer path length through the crystal would facilitate the analysis of more weakly diffracting crystals. To test this, we conducted single shot Laue data collection on crystals of PhnA as well. These crystals were not as robust as Lysozyme so would be a better test if the chip could be used to collect useful data. To ensure the path length of the X-ray through the crystal was maximized, we switched to a much thinner substrate made out of Duralar. The fluid layer that houses the crystallization chambers were made taller (60 µm instead of 25 µm) allowing the crystal more space to grow, thus increasing the path length of the X-ray traversed through the crystal in comparison to the chip materials.

Due to the significant radiation damage resulting (Figure 43d) from the exposure of a crystal to a polychromatic X-ray beam at room temperature; it was difficult to obtain a complete high quality diffraction dataset from a single small crystal. Considering that the model system lysozyme forms robust and strong crystals, this issue would be further exacerbated in other
protein targets. The radiation damage and increases in crystal mosaicty that we observe here would be expected to be much more significant for crystals of most other proteins. However, our approach has the advantage of being able to produce a large number of crystals in the array chip. Therefore a large array chip with 96 wells was used to generate a large number of lysozyme and PhnA crystals, enabling "single-shot" structure determination. In the optimized devices the crystals were able to grow to ~50 µm in size and did not appear to have been significantly limited in their growth by the height of the chamber.

Data collection via a single shot of large number of crystals is particularly attractive for microfluidic applications, provided that crystals can be grown in random orientation on-chip. One of the significant challenges for data collection in a planar microfluidic device is the need to maintain alignment of the crystal during rotation, coupled with geometric limitations for rotation of the device. For single-shot analysis the microfluidic chip can be mounted perpendicular to the X-ray beam and simply translated from well-to-well. The random orientation of crystals within
the chip would then allow for a full survey of rotational space without having to rotate the sample itself. This type of a setup avoids difficulties with rotational and geometric limitations and eases difficulties with optically aligning a rotating crystal in the beam. This method was further enabled by the development of an automated method described in Section 6.3.2 for marking the location of crystals within a microfluidic chip such that data collection could be performed automatically once crystals have been identified.

6.3.2 Automated data collection strategy at BioCARS

The planar nature and large size of the microfluidic chips make it difficult to center each crystal individually and therefore single shot data collection from randomly oriented crystals in each well is a well-suited approach. As discussed before, single-shot data collection from multiple crystals on a single chip also makes it easier for collecting time resolved data because the radiation damage can be negated to a great extent as a single crystal is not subjected to radiation for an extended time. The 96-well array chips however can have close to three to four hundred crystals and collecting data from each of the crystals one at a time, and then translating over to the next one can be laborious. Dr. Zhong Ren at BioCARS has developed an automated x-y-z stage that is motor controlled and coupled to software running on the X-ray data collection computer. How this software coupled to the motor works is as follows. The array chip is mounted in the beam (Figure 42), and the crystals can be identified with the help of two high resolution cameras mounted on the beamline. The software allows us to mark each crystal position, translate the chip to the next well (or same well if there are multiple crystals) and sequentially mark all the crystals grown on the chip. The accuracy of the program is such that it can successfully target crystals that are close together in a single well. The program stores the location of all the crystals thus marked. There is also a provision to record detector distance and exposure time for each crystal. When all crystals have been identified and marked, the program sequentially shoots each crystal according to the data stored, and automatically completes data collection from all the crystals on the chip.

The automated stage, along with the accompanying software makes data collection from a large number of crystals very straightforward. Data collection for Lysozyme and PhnA was performed using this method. Also, the accuracy of crystal marking coupled with fine control of sample translation over the entire dimensions of the chip (20 mm x 10 mm) were validated in
that only a single frame of data was collected which showed overlapping diffraction patterns from multiple crystals, despite the close proximity of many crystals within wells. For PYP a different approach was used since the reaction that was being captured had to be triggered using a laser pulse prior to shooting the crystal with X-rays and wedges of data needed to be collected from each crystal so single shot data collection would not work. The sample was therefore mounted on the goniometer (similar to the monochromatic data collection) and data was collected from each crystal manually. The time delay between these two events was controlled to 500 nanoseconds, 2 microseconds and 1 millisecond on different crystals so the entire reaction pathway could be studied.

6.3.3 On-chip single shot and time resolved crystallography

Table 7 shows the important crystallographic statistics from single-shot Laue data collection on lysozyme and PhnA. In comparison to monochromatic data collection, we see that the resolution and completeness are not as good, which is due to the fact that radiation damage is much stronger for polychromatic data collection. However, for Laue data collection what is observed is that data obtained from merging wedges from multiple crystals at room temperature is superior to data obtained from a similar sized single crystal on-chip. This is due to minimizing the effects of radiation damage by merging of datasets. In addition to obtaining a more complete dataset, merging also gave data to a far better resolution than that obtained using a single crystal.
Table 7. Summary of crystallographic statistics for lysozyme and PhnA crystals grown on-chip

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Lysozyme</th>
<th>PhnA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit Cell Dimensions</td>
<td>a = b = 113.19Å c = 73.87Å</td>
<td>a = b = 113.23Å c = 74.261Å</td>
</tr>
<tr>
<td>Space Group</td>
<td>P4_3_2_1</td>
<td>P4_3_2_2</td>
</tr>
<tr>
<td>Resolution</td>
<td>100Å – 1.61Å</td>
<td>100Å – 2.06Å</td>
</tr>
<tr>
<td>Rmerge on</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Rmerge on</td>
<td>F</td>
<td>^2</td>
</tr>
<tr>
<td>Completeness</td>
<td>87.47% (30.00%)</td>
<td>90.33% (36.29%)</td>
</tr>
<tr>
<td>Mean F/σ(F)</td>
<td>43.06 (18.80)</td>
<td>35.14 (13.26)</td>
</tr>
<tr>
<td># of Frames</td>
<td>67</td>
<td>115</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R (R_free)</td>
<td>0.187 (0.253)</td>
<td>0.282 (0.381)</td>
</tr>
<tr>
<td><strong>Ramachandran Statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most Favored</td>
<td>97.6% (124)</td>
<td>93.1% (377)</td>
</tr>
<tr>
<td>Allowed</td>
<td>2.4% (3)</td>
<td>4.9% (20)</td>
</tr>
<tr>
<td>Disallowed</td>
<td>0.0% (0)</td>
<td>2.0% (8)</td>
</tr>
</tbody>
</table>

Reported values are for all hkl. Values shown in parenthesis represent the value for the highest resolution shell except where indicated. For the Ramachandran statistics the number in parenthesis indicates the number of residues in a given region.

R-factor = Σ (|F_{obs}|-|F_{calc}|)/Σ |F_{obs}| and R_{free} is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.

R_{sym} = Σ |I_i - <I>| / Σ I_i where I_i = intensity of the i^{th} reflection and <I> = mean intensity.

Figure 44 shows the electron density map of lysozyme obtained by merging data from multiple crystals using (a) polychromatic and (b) monochromatic diffraction. While it is clear that the monochromatic data is much more complete and of higher quality, for proof-of-principle, merging datasets using polychromatic diffraction provides acceptable structural information. Figure 45 shows the electron density map for PhnA obtained from single-shot diffraction and merging datasets from multiple crystals.

Preliminary crystallographic data from time resolved studies of PYP are shown in Table 8. From the resolution values and completeness, it is evident that the data is of good quality and that we were able to collect diffraction data on each of the time points. The R-values are comparatively high because the space group P6_3 can be indexed in two ways. In preliminary data analysis, we merged the data from crystals indexed in two different ways. This error is being fixed and more stringent parameters and optimization is being applied to the indexing and refining of data, which has brought down the R-values from 0.08(0.13) to a more acceptable 0.03(0.05) (Data analysis in progress, not shown).
### Table 8. Summary of time-resolved crystallographic statistics for PYP crystals grown on-chip

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Dark 1</th>
<th>500 ns</th>
<th>2 us</th>
<th>1 ms</th>
<th>Dark 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>100Å – 1.82Å</td>
<td>100Å – 1.84Å</td>
<td>100Å – 1.80Å</td>
<td>100Å – 1.80Å</td>
<td>100Å – 1.78Å</td>
</tr>
<tr>
<td>$R_{merge}$ on</td>
<td>0.083</td>
<td>0.078</td>
<td>0.087</td>
<td>0.082</td>
<td>0.075</td>
</tr>
<tr>
<td>$R_{merge}$ on</td>
<td>0.136</td>
<td>0.124</td>
<td>0.141</td>
<td>0.134</td>
<td>0.118</td>
</tr>
<tr>
<td>Completeness</td>
<td>94.9% (87.6%)</td>
<td>86.1% (67.6%)</td>
<td>86.3% (64.9%)</td>
<td>88.0% (64.5%)</td>
<td>89.9% (71.2%)</td>
</tr>
<tr>
<td>Mean F/$\sigma$(F)</td>
<td>25.3 (15.4)</td>
<td>24.3 (11.1)</td>
<td>23.2 (9.1)</td>
<td>24.0 (18.1)</td>
<td>25.8 (12.3)</td>
</tr>
<tr>
<td># of Frames</td>
<td>36</td>
<td>24</td>
<td>25</td>
<td>29</td>
<td>29</td>
</tr>
</tbody>
</table>

Space group: P6₃. Unit cell dimensions: $a = b = 66.76\text{Å}$, $c = 40.893\text{Å}$. Reported values are for all hkl.s. Values shown in parenthesis represent the value for the highest resolution shell except where indicated.

### 6.4 Conclusions

Time-resolved crystallography therefore has a distinct advantage as it provides direct and structural information, for the entire molecule, in atomic detail as a function of time. As such, it is ideally suited for following the propagation of structural changes throughout the protein from the active site (heme in heme proteins, for example), for exploring the events involved in an allosteric mechanism and for tracking the ligands on their migration pathways throughout a protein. However, difficulties associated with data collection at biologically relevant temperatures, extremely high radiation damage from the polychromatic beam, and difficulty in sample preparation, mounting and data analysis have slowed down the widespread acceptance of this method. There is a strong case for using microfluidics to address some of these challenges.

The potential of using microfluidic platforms for protein crystallization with on-chip X-ray analysis capabilities has been demonstrated for polychromatic Laue diffraction methods. The effects of sample and device geometry have been investigated and an automated "single-shot" method for obtaining a complete dataset from many crystals grown on-chip has been developed in collaboration with the BioCARS team at the Advanced Photon Source at Argonne National Laboratory. Further, collection of on-chip time resolved data on PYP has also been demonstrated strengthening the case for using microfluidics for performing functional studies on protein targets.

While the proof-of-concept experiments performed here have merely validated the approach, the application of microfluidic crystallization platforms with on-chip diffraction analysis capabilities for Laue crystallography has tremendous potential for further development. The majority of Laue experiments investigate time resolved structural changes, but have been limited thus far to systems for which robust protein crystals can be grown. Eliminating the need for
manual handling of crystals not only avoids the potential for physically damaging fragile crystals, it also enables the analysis of a significantly larger number of crystals. Further, microfluidics has the advantage of having precise control over transport at the microscale that can be used to trigger reactions in proteins by addition of a ligand, or by changing the pH or temperature. The further development and validation of this technology for time resolved and dynamic crystallography studies would be extremely powerful.
6.5 References

Chapter 7

Concluding remarks and future work

7.1 Summary

In the course of my Ph.D. work, I have developed a microfluidic platform that serves to address some of the major bottlenecks faced by the protein crystallization and structure determination community. The first advantage of using the microfluidic platform is the ease of screening tens of thousands of crystallization conditions using a fraction of the protein solution needed in conventional methods. Secondly, the material used in chip fabrication, as well as the chip design, is such that it is X-ray transparent allowing on-chip analysis of the protein crystals grown. Further, the ability to optimize and get an initial estimate of crystal quality via on-chip X-ray diffraction data collection in the screening phase is an added advantage. We have also demonstrated the ability to collect complete datasets at room temperature by collecting and merging wedges of data from multiple crystals. This allows us to probe the structure and function of crystals that are not amenable to cryo-cooling. Further, the platform\(^1\) eliminates manual harvesting of crystals since it is X-ray transparent, thus allowing the analysis of small and/or fragile crystals. Because we do not handle the crystals or cryo-cool them, the quality of data collected (indicated by mosaic spread) is an order of magnitude better than what is seen in traditional methods.

De novo structure determination of PhnA was accomplished by first collecting SAD phase information from a SeMet derivative and then using the phase information for structure building and obtaining the final 3D structure. This is the first reported work of on-chip phasing from multiple crystals using a microfluidic platform. Further, I implemented a crystallization optimization technique, microseeding on our microfluidic platform. As proof-of-concept, I took PYP, a protein of which it is extremely difficult to grow single large crystals. We successfully
implemented the microseeding technique on-chip and were able to grow and collect data from single, isolated crystals. Finally, our platform has also been utilized for carrying out dynamic studies on various proteins. We have successfully collected time-resolved Laue data from protein crystals (PYP) grown on-chip. Currently we are processing data that will allow us to
observe the various steps of the photo-cycle PYP undergoes when it is subject to a laser pulse. This will be the first time that a microfluidic platform has been used to study these ultra-fast reactions. What I have developed over the course of my PhD is a multi-faceted microfluidic platform capable of not only looking at static structure of various proteins but also investigating protein dynamics in the sub-nanosecond scale. It can be utilized for the complete understanding of a protein structure, right from the start where we screen for conditions to get the optimum crystallization condition, followed by static structure determination using a monochromatic light source, and finally being able to visualize the protein molecule in motion using polychromatic Laue diffraction.

7.2 Future directions

To further expand the capabilities of the microfluidic crystallization platform our group is currently looking at the following different areas for further studies.

7.2.1 Membrane protein crystallization via bicelles

Membrane proteins account for nearly a third of all proteins; however they are vastly underrepresented in terms of total structures in the protein data bank (www.pdb.org).2 Of the over 80,000 structures available in the data base, only 1200 are membrane proteins, which amounts to about one percent of total structures. This is because membrane proteins are notoriously hard to crystallize primarily due to their amphiphilic nature. There exist a number of methods for crystallizing membrane protein, (a) in surfo (b) in meso (lipidic cubic phase) and (c) bicelle based methods.3 Of these, the in surfo method is the oldest and is responsible for the maximum number of structures. However this method relies on using detergents to solubilize the protein, which often causes denaturing and leads to poor crystal quality. The type of crystal packing associated with this method is also not amenable to giving the best quality of diffraction from the crystals. The in meso and the bicelle based method have been recently developed and they both stabilize the protein in a more membrane-like environment, leading to better crystal packing and higher quality crystals. Another graduate student in our group is working to implement the in meso method on a microfluidic platform. I have conducted preliminary experiments to show that the bicelle based method can work on the microfluidic platform discussed in this dissertation. Crystals of the membrane protein bacteriorhodopsin have been grown via a modified crystallization protocol and on-chip data has been collected to 3.1 Å
We are in the process of optimizing the crystallization conditions to grow better quality crystals and collect enough data for a complete dataset. The initial success in implementing the bicelle based method on the existing platform holds great promise. We have access to a number of membrane proteins, which has either not been crystallized or whose high resolution structures are not available due to inability to grow large, high quality crystals using conventional methods. Some of these proteins include the ba3, bo3,
bd and aa3 type heme-copper oxidases. Currently efforts are underway to screen these proteins using the bicelles with commercially available screens.

### 7.2.2 Crystallization of proteins in different space groups

The proteins discussed in the dissertation all belong to high symmetry space groups, namely tetragonal and hexagonal. These high symmetry space groups mean that relatively lesser number of random orientations is needed to get the complete structure. It would be an interesting study to crystallize proteins that crystallize in low symmetry space groups, for e.g., monoclinic or triclinic. This would require data from a much larger number of randomly oriented crystals to get a complete dataset. Ensuring that the microfluidic platform is capable of collecting data from these low symmetry based crystals would further strengthen its versatility.

I have carried out preliminary experiments with lysozyme, to grow crystals in the triclinic space group by varying the crystallization conditions. Lysozyme is known to have difficulty in crystallizing in this low symmetry space groups and prefers the tetragonal geometry. Since the chips are X-ray transparent, it is much easier to check on various space groups after crystals have been obtained and to verify that the data being collected is indeed from the desired space group. Figure 47 shows preliminary results from such lysozyme crystals with excellent diffraction data. Analysis of the diffraction data in ongoing, and a full structure of the protein in the triclinic space group will be obtained shortly.

### 7.2.3 Microseeding efforts for heme-copper oxidases

Microseeding\(^4\) as discussed in Chapter 5 is an optimization method using which the nucleation and growth phases of crystallization can be decoupled, leading to the growth of single, large crystals. After having demonstrated that on-chip seeding can be used to grow large, well diffracting crystals of PYP, we are currently working on applying this technique to crystallize several members of the heme-copper oxidase family (ba3, bd, bo3, aa3) as well as quinol reductase (SQR). We have currently obtained showers of microcrystals from some of these targets and are looking to utilize these microcrystals as seed stock for further experiments.
7.3 References


