Structural Based Drug Discovery: The Significance of Protein Structure

by

Chang Liu

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2021 by the
Graduate Supervisory Committee:

Wei Liu, Chair
Nicholas Stephanopoulos
Po-Lin Chiu

ARIZONA STATE UNIVERSITY

December 2021

ABSTRACT

Structural-based drug discovery is becoming the essential tool for drug development with lower cost and higher efficiency compared to the conventional method. Knowledge of the three-dimensional structure of protein targets has the potential to accelerate the process for screening drug candidates. X-ray crystallography has proven to be the most used and indispensable technology in structural-based drug discovery. The provided comprehensive structural information about the interaction between the disease-related protein target and ligand can guide the chemical modification on the ligand to improve potency and selectivity. X-ray crystallography has been upgraded from traditional synchrotron to the third generation, which enabled the surge of the structural determination of macromolecular. The introduction of X-ray free electron laser further alleviated the uncertain and time-consuming crystal size optimization process and extenuated the radiation damage by "diffraction before destruction". EV-D68 2A protease was proved to be an important pharmaceutical target for acute flaccid myelitis. This thesis reports the first atomic structure of the EV-D68 2A protease and the structures of its two mutants, revealing it adopting N-terminal four-stranded sheets and C-terminal six-stranded ß-barrels structure, with a tightly bound zinc atom. These structures will guide the chemical modification on its inhibitor, Telaprevir. Integrin $\alpha_M\beta_2$ is an integrin with the α I-domain, related to many immunological functions including cell extravasation, phagocytosis, and immune synapse formation, so studying the molecular ligand-binding mechanism and activation mechanism of $\alpha_M\beta_2$ is of importance. This thesis uncovers the preliminary crystallization condition of $\alpha_M\beta_2$-I domain in complex with its ligand Pleiotrophin and the initial structural model. The structural model shows

consistency with the previous hypothesis that the primary binding sites are metal ion-dependent adhesion sites on $\alpha_M\beta_2$-I domain and the thrombospondin type-1 repeat (TSR) domains of Pleiotrophin. Drug molecules with high potency and selectivity can be designed based on the reported structures of the EV-D68 2A protease and $\alpha_M\beta_2$-I domain in the future.

DEDICATION

To my mom and dad, who have loved me and supported me for 28 years and forever.

To the almighty God for his love and guidance.

ACKNOWLEDGMENTS

I would like to express my most profound appreciation to my advisor, Dr. Wei Liu, for his expertise, guidance, and assistance throughout all these years. Without his mentorship, this work would not be possible. I also highly appreciated the chances he provided me to collaborate with so many other talented scientists from different groups. These collaborations granted me tremendous opportunities to learn from and work with amazing people. I would like to thank my committee members, Dr. Nicholas Stephanopoulos, and Dr. Po-Lin Chiu, for their support, advice, and guidance during my Ph.D. journey.

Additionally, I would express my appreciation to my excellent collaborators. I would like to thank Dr. Xu wang, Dr. Di Shen, and Hoa Nguyen for being wonderful collaborators for the Integrin $\alpha_M\beta_2$ project. I would like to thank Dr. Jun Wang, Dr. Chunlong Ma, and Juliana Choza for their help in generating the protein sample of EV-D68 2A protease.

I would like to thank my colleagues in Dr. Wei Liu's lab. I would like to pay my appreciation to Dr. Lan Zhu, who has trained me at the beginning of my research career. I would like to thank Dr. Ming-Yue Lee, who has not only been an excellent collaborator but a mentor to me for several years. He had taught me a lot to be a structural biologist. I would like to thank Dr. James Geiger and Dr. Zina Al-Sahouri for being kind lab mates and supportive all the time.

Finally, I want to thank Liang, who is always by my side during my ups and downs. I cannot make this without you. And Fourier, my cat, for being the cutest creature in my life.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

1.1 Structural based drug discovery

In the drug discovery process, the development of novel drugs with potential interactions with therapeutic targets is significant (Batool et al., 2019). Traditionally, potentially lead was identified with experimental high-throughput screening (HTS), yet it's time-consuming and expensive (Cheng et al., 2012). It takes up to 14 years to discover a preclinical drug candidate and bring it to FDA approval (Song et al., 2009) with a typical approximate cost of 800 million dollars (A Lavecchia & Di Giovanni, 2013). However, the number of new drug approved by the FDA has been decreased because of the low success rate in different clinical trials (Antonio Lavecchia & Cerchia, 2016). Thus, developing an efficient and low-cost alternative to overcome the limitation of conventional drug discovery is of importance. Structural based drug design (SBDD) facilitates drug development with the increasing availability of a considerable number of target proteins with the completion of the Human Genome Project and advances in bioinformatics (Batool et al., 2019). Compared to the conventional drug discovery process, SBDD is more specific, efficient, and rapid to identify potential drug candidates and to optimize the best clinical molecule, because it utilizes the three-dimensional (3D) structural information of the target protein and molecular level knowledge of the related disease (Lionta et al., 2014) (Figure 1.1). There are four stages in SBDD: discovery, development, clinical trials, and registry (Batool et al., 2019).

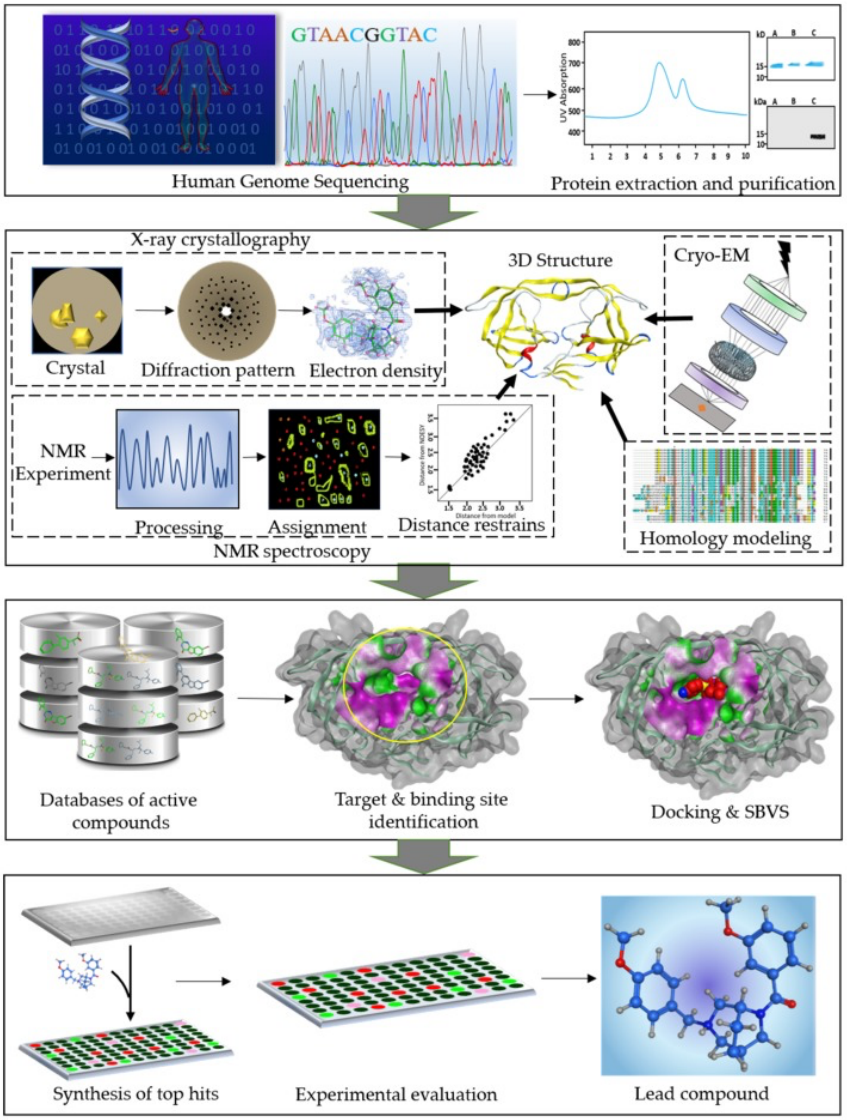


Figure 1.1 A workflow diagram of structure-based drug design (SBDD) process.

Reprinted with permission from Batool, M., Ahmad, B., & Choi, S. (2019). A Structure-Based Drug Discovery Paradigm. International journal of molecular sciences, 20(11), 2783. Copyright © 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

In the discovery stage, a disease-related protein target and potential ligands were proceeded cloning of target gene, followed by the expression, extraction, purification, and 3D structure determination of the protein. A tremendous number of small molecules or compound fragments library can be docked into the binding cavity of the protein by using computational algorithms (Batool et al., 2019). The ranking of these molecules reflects the electrostatic and steric interactions with the binding site, thus, suggesting the possibility of these molecules being good drug candidates. In the development stage, top-ranking molecules are synthesized and optimized. The molecules will be tested by in vitro biomedical essays to evaluate if the candidates can interfere with crucial cellular pathways to realize a desired therapeutic and pharmacological effect (de Graaf et al., 2017). Biological properties of protein-ligand interaction like efficacy, affinity, and potency of the selected compounds are evaluated by experimental methods (Fang, 2012). The 3D structure of the target protein in complex with the high-profile ligands will be determined by suitable technologies, including X-ray crystallography, nuclear magnetic resonance (NMR), and cryogenic electron microscopy (Cryo-EM). Intermolecular features, including binding pockets, ligand-protein interactions, and conformational changes, for molecular recognition and ligand binding, are revealed by the structural details to better elucidate the mechanism behind the interaction and instruct the further optimization of the drug candidates. The third stage will proceed the drug candidate to clinical trials for clinical endpoints, proof of pharmaceutical outcome, and potential side effects. Drug candidates passed the third stage will be thrown into the market to treat patients for a better-quality life and wellness.

1.2 X-ray crystallography for SBDD

The central theme of SBDD is the molecular recognition between a protein target and its ligand (Renaud, 2020). Understanding the molecular recognition and their interaction is essential to reveal their function and disfunction of target protein in related diseases and to design drugs with better efficacy and specificity. The toolbox of structural biology in drug discovery has been expanded to a wide range of computational techniques, including molecular dynamics simulation and novel imaging techniques like cryo-EM. Among different direct imaging techniques (Table 1.1), the most used technique of the single-crystal X-ray crystallography, mostly at synchrotrons.

Table 1.1

Direct imaging techniques used in drug discovery

Sample	Probe	Technique	Interaction	Outcome	Maximum resolution obtained with macromolecules
Solid (crystal)	Photons (X-rays)	X-ray crystallography	Electrons	Electron density map	~0.5 Å
Solid (crystal)	Electrons	Electron crystallography (including micro-electron diffraction [MicroED])	Electrons and protons	Electrostatic potential map	~1 Å
Solid (crystal)	Neutrons	Neutron crystallography	Nuclei	"Nuclear density" map	~1 Å
Solid (cryo-cooled solution)	Electrons	Cryo-electron microscopy (cryo-EM)	Electrons and protons	Electrostatic potential map	~1.5 Å
Solution	Photons (X-rays)	Small-angle X-ray scattering	Electrons	Molecular envelope	~1 nm
Solution	Neutrons	Small-angle neutron scattering	Nuclei	Molecular envelope	~1 nm
Cryo-cooled cell suspension	Electrons	Electron tomography	Electrons and protons	Electrostatic potential map	~5 nm
Cryo-cooled cell suspension	Photons (X-rays)	Soft X-ray tomography	Electrons	Electron density map	~25 nm

Reprinted with permission from Renaud, J.-P. (2020). The Evolving Role of Structural Biology in Drug Discovery. In Structural Biology in Drug Discovery, J.-P. Renaud (Ed.).

The first X-ray diffraction by protein crystals was reported in the 1930s (BERNAL & CROWFOOT, 1934; Clark & Corrigan, 1932), but after 30 years, the atomic crystallographic structure of myoglobin was published (STRANDBERG et al., 1960). Although the X-ray crystallography was believed to determine unambiguous protein structures, the idea of a protein structure could help the design and optimization of specific ligands, which is now a widely accepted obviousness, appeared in 1976 (Maveyraud & Mourey, 2020), six years after launching of the Protein Data Bank (PDB) (Bernstein et al., 1978). The idea then elaborated as "rational drug design cycle" mentioned in 1986 (Hol, 1987): the structural information of a protein complex with its ligand can guide the rational design and chemical optimization of the drug molecule. The central scheme is still appliable nowadays, but with more complexities (Figure 1.2). X-ray crystallography can be used in various stages of drug discovery, from target identification to lead optimization to improve drug candidate potency and selectivity.

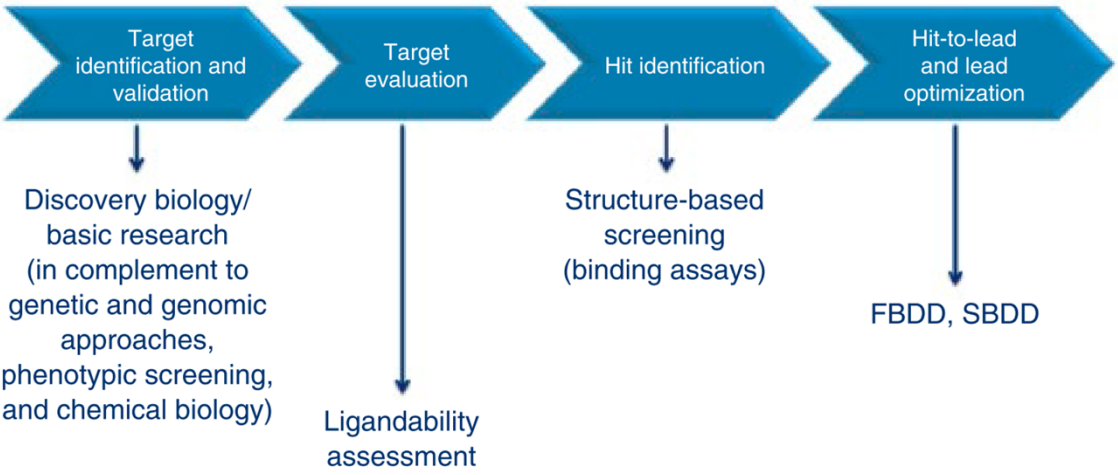


Figure 1.2 Various uses of X-ray crystallography in the drug discovery pipeline

Reprinted with permission from Renaud, J.-P. (2020). The Evolving Role of Structural Biology in Drug Discovery. In Structural Biology in Drug Discovery, J.-P. Renaud (Ed.). Copyright John Wiley and Sons through Copyright Clearance Center's RightsLink®service

X-ray crystallography data collection provides the electron density map of molecules in the crystals when interacting with electrons of atoms. The electron density map represents the time- and space average of the electron density of each molecule, part of which may be missing in the final model due to the poor electron density resulting from its high flexibility. Similarly, the low molecular weight compounds may also display weak electron densities when overlapping with the electron density of the surrounding solvent molecule (Maveyraud & Mourey, 2020). The interaction of an X-ray beam with a 3D crystal will generate a diffraction pattern where X-ray waves are diffracted in definite directions with definite amplitudes and phases (Maveyraud & Mourey, 2020). The structure of the molecules was reflected by the amplitudes and phases, which if they are known, can be computed to get a determined structure. The wave amplitudes are relatively easy to known since the diffracted beam intensities can be measured during the

data collection, and it is proportional to the square of the amplitudes. However, the wave phases are not directly measurable in diffraction experiments, requiring specific phasing methods. Molecular replacement was mainly used for phase determination when a homologous structure is available in PDB. Otherwise, experimental phasing is needed to provide unbiased phase information by the isomorphous introduction of a specific atom such as mercury or iodide or weak anomalous signal of sulfur atoms that natively exist in cysteine and methionine sidechains. Combined with calculated phase and observed amplitudes, an initial atomic model can be built out of the density map and then refined with computation methods. After rounds of model modification and refinement, the final best possible model given the experimental data will be determined. Two factors are often used to summarize the quality of refinement, R and R_{free}. R factor describes the ability of the model to explain the observed data used for building and refining the model, where else, R_{free} factor corresponds to a subset of data that were not used for model building. Not only protein or ligand molecules reside in the model, but also solvent molecules like water molecules. There are two types of water molecules in a structural model. One is discrete water molecules that firmly bound to the proteins, and the other type is the water molecules that are more loosely present in the surrounding shells. Examining water molecules or other solvent molecules may elucidate important information for drug development.

The determined atomic structural model of protein in complex with ligand and solvent molecules provides valuable information allowing the medicinal chemist to design chemical modifications that can improve the drug potency rationally.

1.3 2A protease of human Enterovirus D68

The human Enterovirus D68 (EV-D68) was first identified and characterized in California in 1962 (Schieble et al., 1967), as the infection was linked to the outbreak of 1153 cases with acute respiratory symptoms in the United States in 2014 (Midgley et al., 2015). In the same period, a similar infection was reported in Europe (Poelman et al., 2015), suggesting that the EV-D68 is spreading worldwide and affecting an increasing number of people. Among all the infection cases, respiratory illness was identified as the most common clinical signature of EV-D68 infections (Oberste et al., 2004). EV-D68 infections were also linked to acute flaccid myelitis (AFM), a polio-like neurological disorder, with symptoms including dysgeusia and muscle weakness (Messacar et al., 2018). Although EV-D68 is now considered to impact global health significantly, there is currently no vaccine or specific antiviral treatment against EV-D68 infection besides limited supportive care for severe cases. A recent study shows that the 2A protease on EV-D68 performs as a "security protein" by preventing stress granule (SG) formation and interferon- ß (IFN-ß) gene transcription in all human EV species (EV-A to EV-D) (Visser et al., 2019). Thus, 2A of EV-D68 is significant in the viral infection process of evading the host, escaping the host antiviral response, and viral replication.

An FDA-approved HCV drug, Telaprevir, was proved to be a potent EV-D68 2A protease inhibitor (Musharrafieh et al., 2019). To further uncover the detailed structure of the EV-D68 2A protease and guide chemical design modification on Telaprevir to achieve better potency and selectivity, I report and discuss in this thesis the first atomic structure of the first atomic structure EV-D68 2A protease and the structures of its two mutant C107A and N84T. The structural model reveals the detailed information of EV-

D68 2A protease, suitable crystallization conditions and model-building strategies, promising to lead the structure determination of EV-D68 2A protease with Telaprevir or other compound candidates.

1.4 Leukocyte Integrin $\alpha_M\beta_2$ with Cationic Ligand

Integrins are cell surface adhesion receptors composed of two non-covalently associated αβ heterodimers that play vital roles in lots of cellular functions, such as cell differentiation, cell adhesion, cell migration, and immune response (Hynes, 1992). Integrin $\alpha_M\beta_2$ is an integrin with the α I-domain, expressed on leukocytes like the myeloid, natural killer, and T cells (Shen, 2020). $\alpha_M\beta_2$ is associated with many immunological functions, including cell extravasation, phagocytosis, and immune synapse formation (Kinashi, 2007), so studying the molecular ligand-binding mechanism and activation mechanism of $\alpha_M\beta_2$ is of importance. Pleiotrophin (PTN) is a glycosaminoglycan-binding cytokine and growth factor with potent mitogenic and angiogenic activities (Shen, 2020). It's recently been discovered that PTN is the Cationic Ligand of $\alpha_M\beta_2$ (Shen, 2020). The interaction between $\alpha_M\beta_2$ and PTN was initially accessed by the ability of Mac-1-HEK293 to adhere to immobilized PTN, where demonstrates the fact that both HSPG and Mac-1 can act as receptors for PTN and confirmed by solution NMR spectroscopy (Shen, 2020). This thesis reports the preliminary crystallization condition of the PTN-αM I-domain complex and proposes an initial structure model with a 4.0 Å resolution that provides a closer understanding of PTN-αM I-domain interaction and sheds light on solving this complex structure in the future.

CHAPTER 2

PROTEIN CRYSTALLOGRAPHY AT X-RAY FREE ELECTRON LASERS

2.1 Abstract

Protein crystallography has been successful in structural biology studies for decades. Atomic details of biological macromolecules solved by X-ray diffraction reveal the mechanism of important biological activities, thereby facilitating related drug discovery. However, the surge of novel protein structures was slowed down by the requirement of conventional synchrotron X-ray diffraction for large-size crystals. The application of X-ray free electron laser (XFEL) alleviates the uncertain and time-consuming crystal size optimization process. Tens of thousands of micron-sized protein crystals are injected into an ultra-brilliant femtosecond beam, and the diffraction patterns from each crystal are merged and processed to build the complete model of the protein. In this chapter, the fundamental principle of XFEL is introduced, the successful examples of protein structure determination and the great potential in time-resolved studies are summarized and discussed.

2.2 Introduction

1. Conventional Protein Crystallography Study at Synchrotron Sources

With the increasing demand for functional investigation of targets such as membrane proteins, the structure determination of biological macromolecules becomes increasingly significant for drug development. High-resolution structures of macromolecules are crucial for characterizing biological functions. Currently, macromolecular

crystallography (MX), nuclear magnetic resonance (NMR), and cryogenic electron microscopy (cryo-EM) are the most widely employed methods to investigate the atomic level, three-dimensional (3D) protein structures. Among these technologies, so far, MX is still the most used method to obtain atomic structures of biological macromolecules (Figure 2A). In 1912, Max Von Laue suggested that X-ray could be used to explore the crystal structure, whereas it took about 30 years to prove his hypothesis with protein sample experimentally (BERNAL & CROWFOOT, 1934). Henceforth, X-ray crystallography became the central area for structural biology. With technology development, synchrotron radiation, generating bright, stable, and narrow X-ray, has become an indispensable tool for protein structure determination. Fewer than 500 protein crystal structures were deposited in Protein Data Bank (PDB) before the synchrotron radiation was used for crystallographic studies. In the 1990s, the introduction of third-generation synchrotron radiation (SR) and the improvement of tremendously advanced structure-analysis software (Evans et al., 2011) enabled the surge of macromolecular structures deposited in the PDB (Berman et al., 2003). By September 2021, there were over 178,000 protein structures deposited in the PDB, and more than 85% of them were determined by X-ray crystallography, among which above 80% of them were determined by SR. The brilliant beams achieved using SR have improved data quality and reduced the demand for larger crystal sizes for structure determination. The wavelength variability of SR makes it easy to obtain experimental phases for anomalous diffraction (Hendrickson, n.d.). SR MX beamline now provides the majority of X-ray structures deposition every year (Figure 2B).

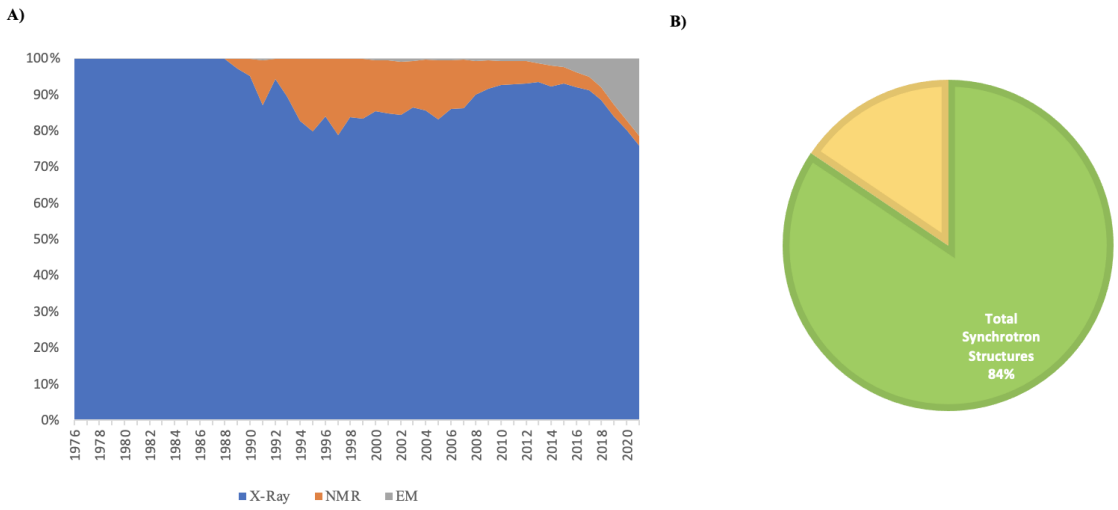


Figure 2.1. Percentage of released PDB structures by methods. A) As shown in the legend, the blue area represents the percentage of PDB structures determined by X-ray crystallography each year from 1976-2021 (09/21), the orange area represents the percentage of structures solved by NMR and the gray area represent the structures determined by EM. B) "Total X-Ray Structures" is the total number of X-ray structures deposited in the PDB; "Total Synchrotron Structures" is the total number of X-ray structures deposited in the PDB for which data was collected at synchrotron facilities. Data accumulated from 1995 to September 2021.

Despite the advancement of SR, there are still challenges remaining from the difficulty of obtaining crystals with a suitable size (ideally 50-100 µm). Especially for a membrane protein, the size of crystals was an inevitable problem while using conventional synchrotron radiation. Although the success rate of getting membrane protein crystals was considerably expedited by the discovery of a unique crystallization method called lipidic cubic phase (LCP), a membrane-mimetic medium (Caffrey, 2003), the typical

crystal size from LCP crystallization is only about 30 μm (Cherezov et al., 2009), far from the expectation. Due to the higher nucleation rate and slower diffusion rate in the lipid bilayer (W. Liu, Wacker, Wang, Abola, & Cherezov, 2014), crystals embedded in LCP tend to be smaller than those using traditional vapor diffusion crystallization techniques. Thus, rounds of time-consuming crystallization optimization were necessary to achieve a suitable size yet not guaranteed, for data collection at the synchrotron. Additionally, room-temperature collected data deterioration caused by X-ray radiation damage was also a negative factor in obtaining high-resolution structures.

Therefore, the exploitation of cryocooling and micro-crystallography has dramatically boosted the success of membrane protein *in-meso* crystallization (Ishchenko et al., 2014). The third-generation synchrotron uses cryocooling during data collection to reduce secondary damage due to the slowing diffusion of radiation-induced radical products. Thus, better quality data can be obtained without an apparent decrease in the radiation dose (Juers & Matthews, 2004). Micro-crystallography also benefited from third-generation synchrotron sources, which achieve an increased signal-to-noise ratio and the ability to conduct selective data collection to avoid the problems brought by inhomogeneous crystals. Yet, radiation damage remains a key limitation for high-resolution diffraction on small crystals until the X-ray free electron lasers (XFELs) sources are used to overcome the barrier.

2. Fundamental of XFEL (Diffraction Before Destruction)

XFEL generates an ultra-bright and ultrashort X-ray beam with tunable energy and femtosecond pulses. Although the power of the XFELs is high enough to destroy the crystals immediately in every encounter, the pulse duration enables to catch the diffraction patterns before the influence of radiation damage caused by scattered photons (Spence et al., 2012). Thus, XFELs allow better collections of high-resolution diffraction from small crystals compared to synchrotron radiation. The nature of "diffraction before destruction" also eliminates the need for cryocooling, which enables collections under room temperature that better mimic in vivo environments. Femtosecond pulses of XFEL beams also allow the recording of conformational changes in macromolecules within the femtosecond time frame, which may enable the elucidation of the dynamic functional mechanism behind the structure.

This 'diffraction before destruction' method lay a solid foundation for serial femtosecond crystallography (SFX). Thousands of microcrystals were injected into the synchronized X-ray bunch in random orientation so that enough patterns of various orientations could be collected to produce a complete data set for structure determination. In the early state, membrane protein microcrystals were delivered through a liquid matrix, which required a large consumption of membrane protein, and most of them went to waste. The microcrystals of photosystem I, the first membrane protein structure solved by SFX in 2011, were delivered with a liquid micro-jet that consumed 10 mg proteins to obtain a full data set (Chapman et al., 2011). In the following year, only 3 mg proteins of photosynthetic reaction center were required because of changing to a sponge phase micro-jet which retained a higher viscosity (Johansson et al., 2012). LCP was taken into

consideration not only because its high viscosity can decrease the flow rate with lower sample waste but also owing to its ability as a crystal-grown matrix so that crystal harvesting steps can be eliminated.

2.3 LCP- serial femtosecond crystallography (SFX)

Traditionally, the protein crystals were delivered to the X-ray beam through a gas dynamic virtual nozzle (GDVN), which delivers the crystals through the low-viscosity medium such as the mother liquor of the crystal (DePonte et al., 2008). However, the flow rate of this method is about 10 µl/min, and the speed is about 10 m/s. The sample consumption is relatively high, which requires about 100 mg of protein for a single complete data set, and thinking about the repetition rate of the hard XFEL sources currently, most of the crystals are not exposed to the beam which are wasted. LCP can not only provide a native-like environment for the protein to crystallize but is a good delivery medium for SFX. LCP micro-extrusion injector was designed because the GDVN injector was not compatible with the highly viscous material (Weierstall et al., 2014). The LCP injector can host crystal-laden LCP samples up to 20, 40, or 100 µL and consists of a narrow capillary with 10-50 µm diameter. The LCP sample is extruded through the capillary by application of high pressure (up to 10,000 psi) from a hydraulic plunger, which is driven by a high-performance liquid chromatography (HPLC) pump. A flow of nonreactive gas (helium or nitrogen) stabilizes the LCP stream from the capillary nozzle. With this injector, the usage of protein can dramatically be reduced to 0.3 mg per dataset. With the development of the LCP injector, high-viscosity, crystal embedded LCP

can be extruded to the beam at a lower speed at 1-300 μL/min. It reduces the sample consumption by 50-100 folds compared with the conventional jetting method.

Practically, because of the huge number of tiny protein crystals, harvesting the protein crystals individually is impossible and unnecessary. The protein crystallization occurs in syringe condition by injecting the protein-laden LCP into another Hamilton syringe filling with the precipitant solution (Cherezov, 2012). Before loading the crystals embedded LCP into the sample delivering nozzle, the precipitant solution must be extruded. Compared with the traditional crystal harvesting process, sample preparation for SFX is more oversimplified. In the experimental validation of LCP-SFX, the appearance of strong, sharp diffraction rings was observed due to the formation of lipidic lamellar crystalline, Lc, phase (Ishchenko et al., 2014). The transformation from cubic phase to Lc phase is attributed to the injection process to the vacuum chamber. This problem was subsequently solved by the reconstitution of the protein-laden LCP sample by adding a small volume of shorter chain lipidic homologous, 7.9 MAG and 9.7 MAG. The transition temperature of these two lipids, from Lc to cubic phase, is 6°C rather than 18°C for 9.9 MAG (Misquitta et al., 2004).

Besides using LCP as a delivery medium for membrane proteins, it is now considered as a suitable carrier of microcrystals of soluble proteins, which allows a decrease in the requirement of sample amount compared to the crystals delivered by liquid injectors. Two soluble proteins, lysozyme, and phycocyanin, have been used to validate the

application of LCP-SFX in soluble protein structure determination with less than 0.1 mg of protein usage and sub 2Å resolution (Fromme et al., 2015).

2.4 Successful cases of XFEL in membrane crystallization

Membrane protein Photosystem I was firstly used to prove the concept of single-crystal X-ray diffraction "snapshot" using XFEL (Chapman et al., 2011). The key innovation of the diffraction experiment included the vacuum-enabled X-ray detection environment and the unique approach where diffractions of nanocrystals were collected on a continuously liquid jet with a gas dynamic virtual nozzle (GDVN) (DePonte et al., 2008). The structure was solved with molecular replacement at 8 Å. The second membrane protein structure solved by SFX was the bacterial photosynthetic reaction center with lipidic sponge phase (LSP) crystallization (Johansson et al., 2012). By adding additives as jeffamine, dimethyl sulfoxide (DMSO), or PEG, the aqueous channels in LCP will swell and transform the cubic phase into a liquid-like phase (Wadsten et al., 2006). The usage of LSP made it possible to overcome the difficulty of jetting the semisolid LCP through a micrometer-sized nozzle to deliver microcrystals in SFX and decreased the total required protein sample amount (Chapman et al., 2011).

To achieve better sample utilization and simplify crystal handling procedures, an LCP micro-extrusion injector was developed to make LCP suitable for SFX (Weierstall et al., 2014). Datasets of several G protein-coupled receptors (GPCRs), including serotonin 5-HT_{2B} receptor bound with ergotamine (W. Liu et al., 2013), were collected using SFX with the LCP injector. Compared with the corresponding synchrotron cryo-structure (Wacker et al., 2013), the XFEL room-temperature structure has a comparable resolution

and similar quality, with a more native-like distribution of thermal motions and residual conformation. LCP-SFX also helped to solve the lacking density around the extracellular Cysteine-Rich domain of the smoothened receptor bound to the teratogen (Weierstall et al., 2014). Compared with structure solved at synchrotron sources with different ligands (W. Liu, Wacker, Wang, Abola, Cherezov, et al., 2014; Wang et al., 2013), the SFX structure uncovered the pose of cyclopamine inside the transmembrane part of the receptor, with lower mosaicity and higher diffraction. Later on, the angiotensin II receptor type 1 (AT_1R) became the first novel GPCR structure solved by XFEL (Haitao Zhang et al., 2015), followed by AT_2R (Haitao Zhang et al., 2017). These structures shed light on the unique function of the two angiotensin receptors and defined the conformational insights for ligand binding and selectivity, which coherently prove that XFELs can benefit GPCRs structural determination.

Despite mentioned advantages of XFELs, crystallographic phasing using S-SAD was still a challenge for XFELs (Stauch & Cherezov, 2018) since the extremely weak anomalous signal from endogenous sulfur atoms (Q. Liu et al., 2012). This has been demonstrated by the determination of the LCP-SFX human adenosine A_2A receptor (A_2AR) phased by S-SAD (Batyuk et al., 2016). This success validated that the beam energy can be tuned to achieve optimal anomalous scattering wavelength for a given element, suggesting that S-SAD should be a generally applicable method for solving the crystallographic phase problem by XFELs (Batyuk et al., 2016).

GPCRs ligand-complex studies and light-driven dynamic structural studies are the major contributions of XFEL in membrane protein structure study (Table 2.2a and 2.2b). As GPCRs take about 50% of the drug market, XFELs would facilitate the understanding of structure-function relationships in the GPCR superfamily and accelerate drug discovery. For example, the ligand switching study based on β_2AR (Ishchenko et al., 2019) reported 8 structures of β_2AR that co-crystallized with different ligands, and 2 of them were unreported before. This study revealed an alternative to address the bottleneck in GPCR related structure-based drug design (SBDD), which requires many structures of the target protein binding with different ligands. This study unlocked the potential of XFELs to become a high throughput method for GPCR ligand-complex structure determination.

Significant functions in the photosystem, such as the electron transport chain, were carried by membrane proteins, which makes them great targets for the light-driven dynamic study. Photosynthesis is considered the most important reaction on earth. It converts solar energy to chemical energy that could be consumed by aerobic organisms, including humans (Dismukes et al., 2001). The unprecedented structures determination of the photosystem by XFEL not only expanded the understanding of the structure insights, the time-resolved study also demonstrated the catalytic details of photosynthesis. For example, recently discovered 6 structures of the photosynthetic reaction center of Blastochloris Viridis show the process of light-induced structural changes on a timescale of picoseconds, which uncover the mechanism of how conformational dynamics stabilized the charge-separation steps of electron transfer reactions (Dods et al., 2021).

XFELs were also benefited other types of studies, including enzymatic mechanism studies, photosynthesis mechanism studies, and technology advancement studies (Table 2.2c). As mentioned in the previous paragraph, the challenge of using S-SAD to solve phase XFELs has been overcome and demonstrated by the determination of the LCP-SFX human adenosine A_2A receptor ($A_{2A}AR$) phased by S-SAD (Batyuk et al., 2016), which confirm the potential of S-SAD to be a potential alternative to solve phase for XFELs.

XFEL is a powerful tool for structural biologists and biochemists. Since the first XFEL facility began to operate, thousands of protein structures have been determined with it. With the development of the instrument, more exciting results will be discovered, more methods and protocols will be established in the foreseen future. However, the current availability of the XFEL facility is too scarce to satisfy the demand from all researchers worldwide. More facilities and alternatives, such as serial millisecond crystallography (SMX), are needed and will be discussed in the later sections.

Table 2.2

XFEL structures of membrane protein

Table 2.2a

XFEL structures of GPCRs ligand-complex study (The reference of each protein structure attached in Appendix A)

Year	Protein	Species	Sub-group	PDB ID	Resolution	Institution
2013	5-HT_{2B}	Human	GPCRs	4NC3	2.80	SLAC LCLS
2014	δ-OR	Human	GPCRs	4RWD	2.70	SLAC LCLS
2014	SMO	Human	GPCRs	4O9R	3.20	SLAC LCLS
2015	AT_1R	Human	GPCRs	4YAY	2.90	SLAC LCLS
2015	rhodopsin	Human	GPCRs	4ZWJ	3.30	SLAC LCLS
2016	OX_2R	Human	GPCRs	5WS3	2.30	SACLA
2017	GCGR	Human	GPCRs	5XEZ	3.00	SLAC LCLS
2017	AT_2R	Human	GPCRs	5UNG, 5UNF	2.80	SLAC LCLS
2017	SMO	Human	GPCRs	5V56,	2.90	SLAC LCLS
2018	MT_2	Human	GPCRs	6ME7, 6ME6, 6ME9, 6ME8	3.10-3.30	SLAC LCLS
2018	MT1	Human	GPCRs	6ME3, 6ME2, 6ME5, 6ME4	2.80-3.20	SLAC LCLS
2018	EP3	Human	GPCRs	6M9T	2.50	SLAC LCLS
2019	MT1	Human	GPCRs	6PS8	3.30	SLAC LCLS
2019	$CysLT_1R$	Human	GPCRs	6RZ5	2.53	SLAC LCLS
2019	β_2AR	Human	GPCRs	6PS1, 6PS4, 6PS3, 6PS6, 6PS0, 6PRZ, 6PS5	2.50-3.40	SLAC LCLS
2019	$A_{2A}AR$	Human	GPCRs	6PS7	1.85	SLAC LCLS
2020	D_2R	Human	GPCRs	7DFP	3.10	SACLA

Table 2.2b

XFEL structures of light-driven dynamic study

Year	Protein	Species	Sub-group	PDB ID	Resolution	Institution
2014	PS II	Thermosynechococcus elongatus BP-1	Photosystems	4PBU	5.00	SLAC LCLS
2016	PS II	Thermosynechococcus elongatus BP-1	Photosystems	5TIS, 5KAF, 5TIS	2.25-3.00	SLAC LCLS
2016	PS II	Thermosynechococcus elongatus BP-2	Photosystems	5KAI	2.80	SLAC LCLS
2016	PS II	Thermosynechococcus vulcanus	Photosystems	5GTH, 5WS6, 5WS5, 5GTI	2.35-50	SACLA
2016	bR	Halobacterium salinarum NRC-1	Bacterial Rhodopsin	5B6Z, 5B6Y, 5B6V, 5B6X, 5B6W, 5H2P, 5H2O, 5H2H, 5H2J, 5H2I, 5H2L, 5H2K, 5H2N, 5H2M	2.10	SACLA
2018	bR	Halobacterium salinarum NRC-1	Bacterial Rhodopsin	6GAH, 6GAA, 6GAC, 6GAB, 6GAE, 6GAD, 6GAG, 6GAF, 6GA9, 6GA8, 6GA1, 6GA3, 6GA2, 6GA5, 6GA4, 6GA7, 6GA6, 6RMK	1.70-2.10	SLAC LCLS
2018	bR	Halobacterium salinarum NRC-1	Bacterial Rhodopsin	6G7I, 6G7H, 6G7K, 6G7J, 6G7L	1.50-1.90	SLAC LCLS
2019	KR2	Dokdonia eikasta	Bacterial Rhodopsin	6TK6, 6TK5, 6TK7, 6TK2, 6TK1, 6TK4, 6TK3	1.60-2.50	SwissFEL ARAMIS
2020	PS II	Thermosynechococcus elongatus BP-1	Photosystems	6W1O, 6W1Q, 6W1P, 6W1R, 6W1U, 6W1T, 6W1V	2.01-2.80	SLAC LCLS
2020	*Bv*RC	Blastochloris viridis	Photosynthetic Reaction Centers	6ZIA, 6ZID, 6ZI6, 6ZI5, 6ZI9, 6ZI4	2.80	SLAC LCLS

Table 2.2c

XFEL structures of other types of studies

Year	Protein	Species	Sub-group	PDB ID	Resolution	Institution
2014	DgkA	Escherichia coli K-12	Phosphotransferases	4UYO	2.18	SLAC LCLS
2017	bCcO	Bos taurus	Electron Transport Chain Complexes	5W97	2.30	SLAC LCLS
2019	bCcO	Bos taurus	Electron Transport Chain Complexes	6NKN, 6NMP, 6NMF	2.50-2.90	SLAC LCLS
2017	M_2	Influenza A virus (A/Hickox/1940(H1N1))	Other Ion Channels	5TTC	1.40	SACLA
2017	M_3	Influenza A virus (A/Hickox/1940(H1N1))	Other Ion Channels	5JOO	1.41	SACLA
2017	M_4	Influenza A virus (A/Hickox/1940(H1N1))	Other Ion Channels	5UM1	1.45	SACLA
2014	PS II	hermosynechococcus vulcanus	Photosystems	4UB6, 4UB8	1.95	SACLA
2019	PS II	Thermosynechococcus vulcanus	Photosystems	6JLK, 6JLJ, 6JLM, 6JLL, 6JLP, 6JLO, 6JLN	2.15-2.50	SACLA
2015	PS II	Thermosynechococcus elongatus BP-1	Photosystems	5E7C, 5E79	3.50-4.50	SLAC LCLS
2015	SR Ca^{2+}-ATPase	Oryctolagus cuniculus	P-type ATPase	4XOU	2.80	SLAC LCLS
2016	bR	Halobacterium salinarum	Bacterial Rhodopsin	5B34, 5B35	2.10-2.35	SACLA
2017	RC_{vir}	Blastochloris viridis	Photosystems	5NJ4, 5O4C	2.40-2.80	SLAC LCLS
2019	hAQP2	Human	Channels	6QF5	3.70	SLAC LCLS
2019	PS I	Thermosynechococcus elongatus BP-1	Photosystems	6PGK	2.90	EuXFEL
2020	A_{2A}AR	Human	G Protein-Coupled Receptors	6WQA	2.00	SLAC LCLS
2020	A_{2A}AR	Human	G Protein-Coupled Receptors	6LPK, 6LPJ	1.80	SACLA

*The studies of DgkA and bCcO were categorized as enzymatic mechanism study. The studies of PS II were categorized as photosynthesis mechanism study. The studies of SR Ca^{2+}-ATPase, bR, $RC_{vir,}$ hAQP2, PS I and $A_{2A}AR$ were categorized as technology advancement studies.

2.5 LCP- serial millisecond crystallography SMX and LCP-Pink beam

The application of SFX facilitates protein structure determination and creates a new path of data collection. The remaining issue of SFX, however, is the scarce availability of XFEL globally. The theory of free electron laser was developed in the 1970s, and the first proposal to build an X-ray free electron laser facility was proposed in 1992 (Pellegrini, 2012). From there, it took about two decades for the first XFEL facility- LCLS, to start for operation.

So far, there are only six operational facilities worldwide. They are European X-Ray Free-Electron Laser (European XFEL) in Germany, Linac Coherent Light Source (LCLS) and the recently updated LCLS-II in the US, SPRing-8 Compact Free-Electron Laser (SACLA) in Japan, Swiss Free-Electron Laser (SwissFEL) in Switzerland, and Pohang Accelerator Laboratory X-Ray Free-Electron Laser (PAL-XFEL) in South Korea (Zhu et al., 2020). The LCLS was upgraded to LCLS-II provides a giant leap in capability by increasing the pulse interval from 120 pulse/sec to 1 million pulse/sec. In 2025, there will be one more light source available in China, Shanghai High Repetition Rate XFEL and Extreme Light Facility (SHINE) (Xu et al., 2020). The percentage of protein structures determined in different XFEL facilities is summarized in figure 2.2.

Since the facilities are in short supply and in great demand, it would be optimal if there were alternatives with wide accessibility and bright X-ray sources such as synchrotrons. To date, the third generation, storage-ring-based synchrotron source can provide microfocus monochromatic beamlines with high brightness (up to 10^{13} photons s^{-1}) and extremely focused X-ray beams which makes it become a readily available substitute for XFEL (Martin-Garcia et al., 2019). Besides, prescreening the protein crystal sample using LCP- serial millisecond crystallography (LCP-SMX) at synchrotron source can provide a preliminary diffraction result before collecting data at XFEL, which can make better use of the precious XFEL beam time and increase the success rate at XFEL.

Same as LCP-SFX, to collect a complete data set with LCP-SMX, thousands of micrometer-sized protein crystals need to be injected into the beam with the LCP nozzle, and LCP-SMX obviates the inefficient protein crystal harvesting. The advantages of SMX also include that the data collection is under room temperature, which is more relevant to the physiological condition. However, limitations of LCP-SMX include radiation damage, the need for more high-density protein crystals than LCP-SFX. Due to the bandwidth of SMX ($\Delta E/E \simeq 10^{-4}$) being about 1/5th of XFEL plus ($\Delta E/E \simeq 2\times10^{-3}$), more randomly oriented protein crystals are needed to sample the complete Bragg diffraction peaks (Meents et al., 2017).

So far, many SMX measurements have been performed at synchrotron sources with monochromatic radiation using millisecond exposure time (Jaeger et al., 2016; Martin-Garcia et al., 2017). The first serial crystallography study with synchrotron sources was

conducted in 2014, which yielded a 2.1 Å lysozyme structure (Stellato et al., 2014). This proof-of-concept study confirmed that serial crystallography is applicable at synchrotron sources. By merging snapshots from single microcrystals with random ordinations, protein structure can be determined. In 2015, the LCP injector was applied to the serial crystallography study at synchrotron sources (Nogly et al., 2015) and they determined the structure of a membrane bacteriorhodopsin (bR) at a resolution of 2.4 Å at room temperature. There is no noticeable difference between the room temperature structure and the cryocooling structure, although room temperature data collection is a huge advantage because cryocooling sometimes affects the dynamic behavior of protein (Fraser et al., 2011; Keedy et al., 2014).

However, to discover the bio-reactions that proceed on a shorter timescale, shorter exposure time, and brighter X-ray are needed. Laue diffraction (Keith Moffat et al., 1984), a method that employs the full polychromatic spectrum of an undulator harmonic at a synchrotron source, has been applied to the study of many pivotal biological reactions such as enzymatic reactions, ligand–enzyme interactions, viruses, and viral drug complex (Keith Moffat et al., 1986). By using the pink beam, the intensity of most reflections is fully recorded. Without using a monochromator and only using a mirror for high energy cut-off, pink-beam obtains a larger bandwidth ($\Delta E/E \simeq 5\times10^{-2}$) which reduces the sample consumption. Despite all this, a well-known drawback of Laue diffraction is the high-scattering background which typically leads to the low single-to-noise ratios, thereby reducing the resolution. Air is one of the contributors to the background. Embedding the protein sample is one of the solutions to eliminate the effect

from the air. Since LCP is an ideal sample deliver matrix, LCP-pink beam has been proved a successful method to obtaining atomic-resolution structure.

Figure 2.2. XFEL Facilities Worldwide. a) Distribution of XFEL facilities in the United States, Germany, Japan, and Switzerland; b) Percentage of membrane protein XFEL structure solved at different facilities, 65% at the US SLAC LCLS, 29% at Japan SACLA, 5% at Switzerland SwissFEL and 1% at Germany EuXFEL.

2.6 XFEL application in protein dynamic study

XFEL-based time-resolved studies create a new path to investigate the molecular insight of protein dynamic, which is not accessible with synchrotron source or single-particle cryo-electron microscopy (cryo-EM) (Westenhoff et al., 2010). The protein dynamic study using X-ray crystallography was limited by the electron pulse duration of 100 ps since 1996 (K Moffat, 1998; Šrajer et al., 1996). Time-resolved-SFX (TR-SFX) provides the opportunity of ultrafast time-resolved diffraction study with its extremely brilliant pulse and ~40 fs pulse duration.

The TR-SFX results are dominant with light-sensitive proteins, which are not common in the human. Therefore, the dynamic studies from TR-SFX do not show huge clinical and pharmaceutical impacts so far. However, the light-sensitive systems are easy to initiate or

trigger, which makes them an ideal target to prove the concept. Most of the TR-SFX studies employed micron-sized crystals, which are less dense compared to other SFX experiments, thus, the visible laser beam can perturb most of the protein molecules within the microcrystals and thereby trigger or induce the conformational changes. The snapshots of conformational changes are recorded by ultrafast pulses. The diffraction often exhibits various intermediates of the protein, which consequently lead to X-ray-induced "radiation damage" as analogous diffraction at synchrotron sources (de la Mora et al., 2020; Nass, 2019). TR-SFX performed on light-sensitive proteins has discovered many important molecular insights after absorbing a photon. A recent example of TR-SFX involving the photosystem II (PS-II) has revealed the mechanism of diffusion of energy absorbed with internal chromophores network (Ibrahim et al., 2020; Kern et al., 2018). Other photo-induced protein systems that have been studied by TR-SFX include rhodopsin (Nogly et al., 2018; Skopintsev et al., 2020), green fluorescent proteins (Woodhouse et al., 2020), phytochromes (Claesson et al., 2020), and photoactive yellow proteins (Pandey et al., 2020).

TR-SFX studies also include relatively slower biological processes such as enzymatic reactions whose reaction duration is at the millisecond range. Additional requirements for studying the non-light-driven enzymatic reaction include caged compounds or caged protein, ligand exchange, or temperature jump. This is challenging because these requirements must be high selectivity, high quantum yield, and temporal resolution. Takehiko et al. reported the capturing of intermediate of the enzymatic reaction of P450nor in 2017, which is a successful example of the usage of caged compounds (Tosha

et al., 2017). The ligand switching experiments with GPCR demonstrated the great potential of TR-SFX in drug discovery (Ishchenko et al., 2019).

2.7 Conclusion and future perspective

The development of XFEL offers a new powerful tool for structural biologists. It increased the efficiency and decreased the consumed time by using a smaller and smaller number of crystals to collect a complete data set for structure determination, compared to conventional synchrotron radiation. To cope with radiation damage, one of the major challenges in crystallography, XFEL minimized it with "diffraction before destruction." However, the data processing of serial crystallography became more complicated than single-crystal collection since all microcrystals are randomly oriented. It makes the phasing determination even harder. Molecular replacement is the most employed method for phase determination, but it highly relies on suitable search models generated from previously known protein structures. Studies showed that the beam energy at XFELs can be tuned to reach the optimal anomalous scattering wavelength for native sulfur, demonstrating that S-SAD could be a potential alternative for some XFELs targets whose phasing cannot be determined by molecular replacement. It also has been proved on lysozyme that de novo phasing with single-wavelength anomalous dispersion (SIR) phasing (Barends et al., 2014) and multiple wavelength anomalous diffraction can help the phase problem of XFEL data (Gorel et al., 2017). Yet, further development of data processing, especially the alternative methods of phasing, is of great interest.

Improvements in the crystal jetting and delivery medium made the data collection more efficient and less sample volume required. Practically, the current LCP injector often has clog issues due to larger size crystals in the LCP string. This could waste precious beam time and protein samples. The balance between crystal size and LCP injector nozzle diameter needs to be optimized. Besides, there are still plenty of crystals in the LCP stream that do not have a chance to be exposed to the X-ray. Even though the current consumption of protein is acceptable, less sample waste is always preferred. Thus, further studies of improving hit rate, discovering better delivery medium, optimizing jetting nozzle, and designing a better speed control system, are necessary.

The total number of the XFEL facility and available resources are still under demand, which makes the beam time very competitive. Better allocation of synchrotron sources for the serial crystallography and time-resolved studies will still be valuable. Room temperate data collection of serial crystallography at synchrotron source has a great potential for the time-resolved study. For the photo-sensitive protein target, synchrotron laser pauses can be used to activate the protein crystal, and with the rapid readout detector, millisecond resolution can be achieved already. The new generation synchrotron source and detector may yield higher resolution for the real-time study.

The TR-SMX facilitates protein dynamic studies. Although the majority of results are based on the light-sensitive proteins, the application of caged compounds, ligand switching, and substrate diffusion have a great potential to become generalized, which can be applied to induce most of the biochemical reactions, and this could be extremely

useful for the investigation about the pharmacologically relevant protein target such as GPCRs and transporters.

CHAPTER 3

Structural insights of the 2A protease from human Enterovirus D68

3.1 Abstract

The infection of Enterovirus D68 (EV-D68) is associated with severe respiratory illness and acute flaccid myelitis (AFM) in children. There is no vaccine for prevention or treatment available for EV-D68 infection. EV-D68 2A protease ($2A^{pro}$) has been reported to be the antiviral target for EV-D68 infection. This thesis reports the first structure of EV-D68 $2A^{pro}$ and the structures of its two mutants, N84T ($2A^{N84T}$) and C107A ($2A^{C107A}$). EV-D68 $2A^{pro}$ adopts a structure of N-terminal four-stranded sheet and C-terminal six-stranded ß-barrel, like the other enterovirus 2A structures with sub-one r.m.s.d. The surface rendering of EV-D68 2A presents an open cleft, with the active site being accessible, which promises the interaction between the active site on EV-D68 2A and different substrates. The structures of ligand-free 2A protease shed light on the detailed structural information and can guide to obtain the ligand-bound complex structure in the future.

3.2 Introduction

The genus *Enterovirus* of the family *Picornaviridae* contains many significant pathogens related to human and mammalian diseases. This genus contains fifteen species: there are four human enteroviruses (A-D) (J. Sun et al., 2019), eight animal enterovirus (E-L), and three rhinoviruses (A-C). The human Enterovirus D68 (EV-D68) was first identified and characterized in California in 1962 (Schieble et al., 1967). Only 26 sporadic confirmed

cases of EV-D68 were reported from 1970 to 2005, according to the statistics of the US National Enterovirus Surveillance System (NESS) (Midgley et al., 2015). Although infections associated with EV-D68 were considered rare, there was an outbreak in 2014 in the United States when 1153 cases with acute respiratory symptoms were confirmed to be EV-D68 infections (Midgley et al., 2015). A similar statistic of EV-D68 infection was reported in Europe during the same period (Poelman et al., 2015). Respiratory illness is the most common clinical signature of EV-D68 infections (Oberste et al., 2004). Although EV-D68 is a non-polio enterovirus, it was linked to acute flaccid myelitis (AFM), a polio-like neurological disorder, with symptoms including dysgeusia and muscle weakness (Messacar et al., 2018). Most patients had onset of AFM between August and November, with increases in AFM cases every two years since 2014 (Figure 3.1a). This special linkage suggests a possibility that EV-D68 may result in neurological diseases. Although EV-D68 is now considered to impact global health significantly, there is currently no vaccine or specific antiviral treatment against EV-D68 infection besides limited supportive care for severe cases.

The genome of EV-D68 contains a positive-sense single-stranded RNA of about 7.6 kb, encoding a precursor polyprotein which yields four structural proteins (VP1, VP2, VP3, and VP4) after autocatalytic cleavage, and seven non-structural proteins (2A, 2B, 2C, 3A, 3B, 3C, and 3D) (Opanda et al., 2014). (Figure 3.1b) The 2A protease ($2A^{pro}$), $3C^{pro}$, and $3CD^{pro}$ will cleave the polyprotein to produce mature viral proteins upon translation. Both $2A^{pro}$ and $3C^{pro}$ comprise a catalytic cysteine nucleophile and chymotrypsin-like fold (J. Sun et al., 2019). $2A^{pro}$ and $3C^{pro}$ perform different site-cleavage on polyprotein via

specific sequences recognition. The $2A^{pro}$ cleaves at the VP1-2A junction on a conserved amino acid sequence, whereas $3C^{pro}$ cleaves at all other junctions, including the 3C N-terminal and C-terminal (J. Sun et al., 2019). $2A^{pro}$ from human rhinovirus 2 (HRV2) (W. Lee et al., 2014), coxsackievirus B4 (CVB4) (Baxter et al., 2006), Coxsackievirus A16 (CVA16) (Y. Sun et al., 2013), human rhinovirus C15 (HRV-C15) (Ling et al., 2018) and EV-A71 (Cai et al., 2013) has an N-terminal four-stranded ß-sheet and a C-terminal ß-barrel with six ß-sheets, while $3C^{pro}$ from poliovirus 1 (PV-1) (Mosimann et al., 1997), hepatitis A virus (HAV) (Malcolm et al., 1994), human rhinovirus 14 (HRV14) (Matthews et al., 1994), EV-A71 (Cui et al., 2011), and EV-D68 (Tan et al., 2013) adopts an N-terminal ß-barrel with six ß-sheet. $2A^{pro}$ contains an additional noncatalytic Zn atom not found in the EV $3C^{pro}$ structures except for hepatitis C virus (HCV) NS3-4A serine protease (Stempniak et al., 1997).

Studies show that non-structural proteins can interact with the host to facilitate virus replication by suppressing interferon production and escaping immune response (Feng et al., 2014; Lei et al., 2011). A recent study shows that 2A performs as a “security protein” of enteroviruses by preventing stress granule (SG) formation and interferon- ß (IFN-ß) gene transcription in all human EV species (EV-A to EV-D) (Visser et al., 2019). Thus, 2A of EV-D68 is significant in the viral infection process of evading the host, escaping the host antiviral response, and maintaining viral replication.

Despite the importance of EV-D68 $2A^{pro,}$ and it has become a popular target for drug discovery, it turned out that it cannot be expressed in *E. coli.* on its own, until recently, a

cleavable SUMO tag realized its upscaled production (Musharrafieh et al., 2019). As shown in the fluorescence resonance energy transfer (FRET) assay, the 2A protease activity demonstrated its specific cleavage towards a peptide substrate designed based on VP1–2A junction (Yanmei Hu, Rami Musharrafieh, Madeleine Zheng, 2020). Telaprevir - an FDA-approved HCV NS3/NS4A serine protease inhibitor, was identified as a 2A protease inhibitor with an IC_{50} value of 0.2 μM after FRET-based high-throughput screening against the Selleckchem protease inhibitor library (Musharrafieh et al., 2019). Enzyme kinetic studies suggest, and the dialysis experiment suggested that the complex formation of telaprevir and EV-D68 $2A^{pro}$ is irreversible. Following cell studies supported telaprevir blocking EV-D68 $2A^{pro}$ replication during early stages in multiple cell lines, including RD, A549, HEK293, and neuronal cells A172 (Musharrafieh et al., 2019). The serial viral passage experiment (Ulferts et al., 2016) selected a drug-resistant mutation N84T in both the enzymatic assay and the cellular antiviral assay (Musharrafieh et al., 2019). The thermal shift assay also supports the single mutation resistance that telaprevir can significantly stabilize the EV-D68 $2A^{pro}$ by increasing the melting temperature (Tm) by 13.78°C at 30 μM, while for the N84T mutant, it has a minimal effect. Altogether, EV-D68 $2A^{pro}$ was proved to be an antiviral drug target, and telaprevir is a candidate of EV-D68 $2A^{pro}$ inhibitors.

Even though great opportunities lie in telaprevir to be a potent antiviral drug, some risks are still uncovered. It has been reported that as a discontinued HCV drug, dermatological side-effects were observed on patients getting telaprevir monotherapy (Torii et al., 2013), which is a red flag for pediatric patients' treatment. Since telaprevir was not discovered

for antiviral treatment for EV-D68 $2A^{pro}$, additional drug design to achieve higher potency and better selectivity is expected through a structure-based approach. Thus, solving the high-resolution structures of the EV-D68 $2A^{pro}$ and EV-D68 $2A^{N84T}$ is necessary to accelerate e the process.

Here, the first crystal structure of EV-D68 $2A^{pro}$ at 3.5 Å resolution (7MG0), along with two mutants' structure EV-D68 $2A^{C107A}$ mutant (7JRE) at 2.5 Å and N84T mutant (EV-D68 $2A^{N84T}$) at 2.6 Å (7LW2), are reported (Figure 3.2). Comparison of these structures to other $2A^{pro}$ from EV-A71, CVA16, and HRV-C15 was conducted since they are three of the $2A^{pro}$ structures with the highest sequence identity and structural similarity.

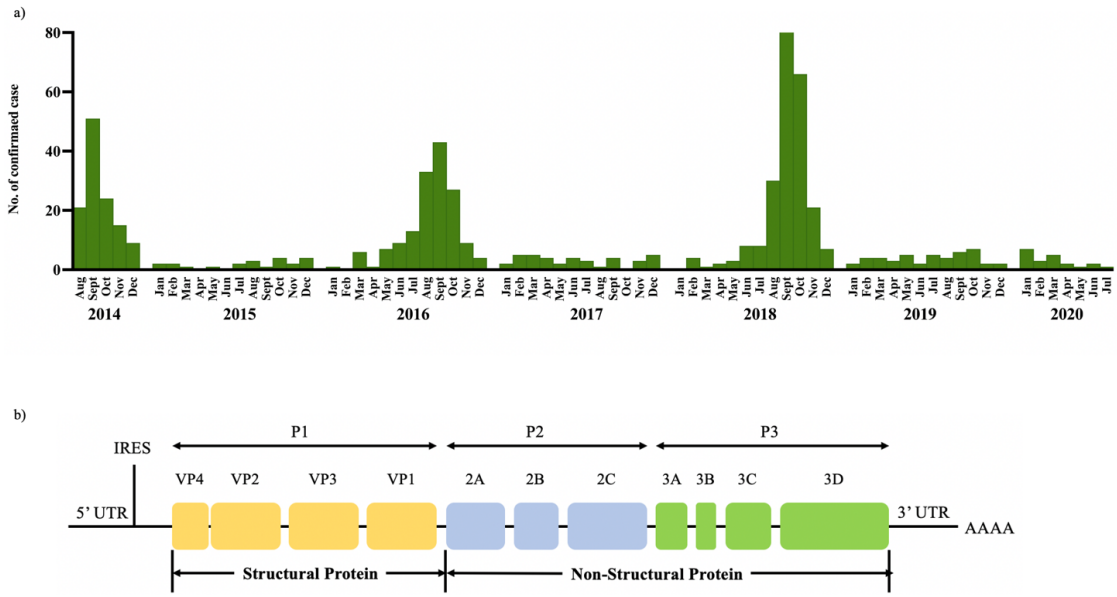


Figure 3.1 EV-D68 infection cased from 2014-2020 a) Number of confirmed cases of acute flaccid myelitis reported to the Centers for Disease Control and Prevention, United

States, August 1, 2014–August 31, 2020. Data as of August 31, 2020. b) Schematic representation of the enterovirus D68 genome.

Figure 3.1b Sun, J., Hu, X. Y., & Yu, X. F. (2019). Current Understanding of Human Enterovirus D68. Viruses, 11(6), 490. https://doi.org/10.3390/v11060490 Copyright © 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

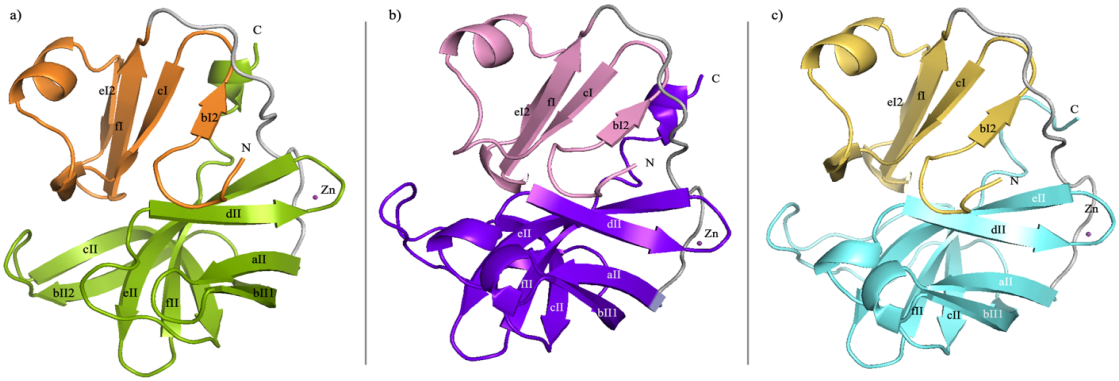


Figure 3.2 Overall structure of EV-D68 2A protease structures a) Overall structure of EV-D68 $2A^{C107A}$chain E. The N-terminal domain and the C-terminal domain are shown in orange and green, respectively. The zinc ion is represented as a magenta-colored sphere. b) Overall structure of EV-D68 $2A^{N84T}$ chain E. The N-terminal domain and the C-terminal domain are shown in pink and purple. The zinc ion is represented as a magenta-colored sphere. c) Overall structure of EV-D68 $2A^{pro}$ chain B. The N-terminal domain and the C-terminal domain are shown in yellow and cyan. The zinc ion is represented as a magenta-colored sphere.

3.3 Materials and Methods

1. Materials and Instruments

Commercialized crystallization screens PEG/Ion HT (HR2-139), Index HT (HR2-134), Crystal Screen HT (HR2-130), and Natrix HT (HR2-131) were purchased from Hampton Research. High-throughput crystallization robot Crystal Gryphon was purchased from Art Robbins Instruments. Crystallization plate storage and inspection robot Rigaku was purchased by Petra Fromme Group and was shared with the entire Biodesign Center for Applied Structure Discovery (BCASD).

Protein samples were overexpressed and purified by Chunlong Ma, Juliana Choza, and Tommy Szeto from Dr. Jun Wang's lab at the University of Arizona.

2. EV-D68 $2A^{pro}$ crystallization

A) First round of crystallization screening

Samples were received on dry ice, including EV-D68 $2A^{pro}$ #2 at 6.2 mg/ml and EV-D68 $2A^{pro}$ #6 at 13 mg/ml. Samples were centrifuged at 16,000*g* for 10 min at 4 °C before setting up crystallization. Crystallization was set up at room temperature using Crystal Gryphon. 6 commercialized crystallization screens were tried for each sample. For the EV-D68 $2A^{pro}$ #6, the volume was only enough for 3 screens.

Detailed information of crystallization setup for each sample shows in Table 3.1.

Table 3.1

First-round EV-D68 2A^{pro} crystallization screening

Sample	Concentration	Protein: Precipitant	Screen
EV-D68 2A^{pro} #2	6.2 mg/ml	2:1 in subwell1 1:1 in subwell2	HR2-139 PEG/ION HT HR2-134 Index HT HR2-131 Natrix HT HR2-137 MembFac HT HR2-248 Grid Screen Salt HT HR2-130 Crystal Screen HT

Sample	Concentration	Protein: Precipitant	Screen
EV-D68 2A^{pro} #6	13.0 mg/ml	2:1 in subwell1 1:1 in subwell2	HR2-139 PEG/ION HT HR2-134 Index HT HR2-131 Natrix HT

* For each well of the plate, 200 nL protein and 100 nL precipitant solution were dispensed in subwell1, and 100 nL protein and 100 nL precipitant solution were dispensed in subwell2.

All screens were incubated in 20 °C storage, and images were taken by Rigaku.

B) First round of crystallization optimization

EV-D68 $2A^{pro}$ sample was received on blue ice, 10.0 mg/ml. The sample was concentrated at 3,000 g at 4 °C until the concentration reached 13.0 mg/ml and then centrifuged at 16,000*g* for 10 min at 4°C before setting up crystallization. Crystallization was set up at room temperature using Crystal Gryphon.

Optimization screens were designed based on the result of previous screenings (Figure 3.3-3.5). The concentration of each component was optimized based on its initial concentration. Two commercialized crystallization screens were used as positive controls, and three optimization screens OPT-VP-1, OPT-VP-2 and OPT-VP-3 were tested (Table 3.2)

All screens were incubated in 20°C storage and images were taken by Rigaku.

		1	2	3	4	5	6	7	8	9	10	11	12
A	1M Sodium cacodylate pH 6.0	0	0	0	0	0	0	0	0	0	0	0	0
	1M Magnesium chloride	0	1	2	3	4	5	6	7	8	9	10	11
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	160.1	159.1	158.1	157.1	156.1	155.1	154.1	153.1	152.1	151.1	150.1	149.1
B	1M Sodium cacodylate pH 6.0	5	5	5	5	5	5	5	5	5	5	5	5
	1M Magnesium chloride	0	1	2	3	4	5	6	7	8	9	10	11
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	155.1	154.1	153.1	152.1	151.1	150.1	149.1	148.1	147.1	146.1	145.1	144.1
C	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	0	1	2	3	4	5	6	7	8	9	10	11
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	150.1	149.1	148.1	147.1	146.1	145.1	144.1	143.1	142.1	141.1	140.1	139.1
D	1M Sodium cacodylate pH 6.0	15	15	15	15	15	15	15	15	15	15	15	15
	1M Magnesium chloride	0	1	2	3	4	5	6	7	8	9	10	11
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	145.1	144.1	143.1	142.1	141.1	140.1	139.1	138.1	137.1	136.1	135.1	134.1
E	1M Sodium cacodylate pH 6.0	20	20	20	20	20	20	20	20	20	20	20	20
	1M Magnesium chloride	0	1	2	3	4	5	6	7	8	9	10	11
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	140.1	139.1	138.1	137.1	136.1	135.1	134.1	133.1	132.1	131.1	130.1	129.1
F	1M Sodium cacodylate pH 6.0	25	25	25	25	25	25	25	25	25	25	25	25
	1M Magnesium chloride	0	1	2	3	4	5	6	7	8	9	10	11
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	135.1	134.1	133.1	132.1	131.1	130.1	129.1	128.1	127.1	126.1	125.1	124.1
G	1M Sodium cacodylate pH 6.0	30	30	30	30	30	30	30	30	30	30	30	30
	1M Magnesium chloride	0	1	2	3	4	5	6	7	8	9	10	11
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	130.1	129.1	128.1	127.1	126.1	125.1	124.1	123.1	122.1	121.1	120.1	119.1
H	1M Sodium cacodylate pH 6.0	35	35	35	35	35	35	35	35	35	35	35	35
	1M Magnesium chloride	0	1	2	3	4	5	6	7	8	9	10	11
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	125.1	124.1	123.1	122.1	121.1	120.1	119.1	118.1	117.1	116.1	115.1	114.1

Figure 3.3 **OPT-VP-1** (optimization of Sodium cacodylate and magnesium chloride concentration) shows the detailed design of the optimization screen. A-H represent 8 rows on a 96 well screen plate, and 1-12 represent 12 columns it has. All components were labeled with their stock concentration, and each number in a cell represents the amount in volume (μl) of the component in each corresponding well.

		1	2	3	4	5	6	7	8	9	10	11	12
A	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0	0	0	0	0	0	0	0	0	0	0	0
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5
	dH2O	151.5	151	150.5	150	149.5	149	148.5	148	147.5	147	146.5	146
B	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5
	dH2O	151.3	150.8	150.3	149.8	149.3	148.8	148.3	147.8	147.3	146.8	146.3	145.8
C	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5
	dH2O	151.1	150.6	150.1	149.6	149.1	148.6	148.1	147.6	147.1	146.6	146.1	145.6
D	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5
	dH2O	150.9	150.4	149.9	149.4	148.9	148.4	147.9	147.4	146.9	146.4	145.9	145.4
E	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5
	dH2O	150.7	150.2	149.7	149.2	148.7	148.2	147.7	147.2	146.7	146.2	145.7	145.2
F	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5
	dH2O	150.5	150	149.5	149	148.5	148	147.5	147	146.5	146	145.5	145
G	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5
	dH2O	150.3	149.8	149.3	148.8	148.3	147.8	147.3	146.8	146.3	145.8	145.3	144.8
H	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5
	dH2O	150.1	149.6	149.1	148.6	148.1	147.6	147.1	146.6	146.1	145.6	145.1	144.6

Figure 3.4 **OPT-VP-2** (optimization of calcium chloride and spermine concentration) shows the detailed design of the optimization screen. A-H represent 8 rows on a 96 well screen plate, and 1-12 represent 12 columns it has. All components were labeled with their stock concentration, and each number in a cell represents the amount in volume (μl) of the component in each corresponding well.

		1	2	3	4	5	6	7	8	9	10	11	12
A	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	0	0	0	0	0	0	0	0	0	0	0	0
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	186.6	186.6	186.6	186.6	186.6	186.6	186.6	186.6	186.6	186.6	186.6	186.6
B	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	174.1	174.1	174.1	174.1	174.1	174.1	174.1	174.1	174.1	174.1	174.1	174.1
C	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	25	25	25	25	25	25	25	25	25	25	25	25
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6
D	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	149.1	149.1	149.1	149.1	149.1	149.1	149.1	149.1	149.1	149.1	149.1	149.1
E	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	50	50	50	50	50	50	50	50	50	50	50	50
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	136.6	136.6	136.6	136.6	136.6	136.6	136.6	136.6	136.6	136.6	136.6	136.6
F	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	62.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	124.1	124.1	124.1	124.1	124.1	124.1	124.1	124.1	124.1	124.1	124.1	124.1
G	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	75	75	75	75	75	75	75	75	75	75	75	75
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	111.6	111.6	111.6	111.6	111.6	111.6	111.6	111.6	111.6	111.6	111.6	111.6
H	1M Sodium cacodylate pH 6.0	10	10	10	10	10	10	10	10	10	10	10	10
	1M Magnesium chloride	1	1	1	1	1	1	1	1	1	1	1	1
	1M Calcium chloride	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
	80% 2-propanol	87.5	87.5	87.5	87.5	87.5	87.5	87.5	87.5	87.5	87.5	87.5	87.5
	0.1M Spermine	2	2	2	2	2	2	2	2	2	2	2	2
	dH2O	99.1	99.1	99.1	99.1	99.1	99.1	99.1	99.1	99.1	99.1	99.1	99.1

Figure 3.5 **OPT-VP-3** (optimization 2-propanol concentration) shows the detailed design of the optimization screen. A-H represent 8 rows on a 96 well screen plate and 1-12 represent 12 columns it has. All components were labeled with their stock concentration and each number in a cell represents the amount in volume (μl) of the component in each corresponding well.

Table 3.2

First-round D682A-WT crystallization optimization

Sample	Concentration	Protein: Precipitant	Screen
EV-D68 $2A^{pro}$	13.0 mg/ml	2:1 in subwell1 1:1 in subwell2	HR2-134 Index HT HR2-131 Natrix HT OPT-VP-1 OPT-VP-2 OPT-VP-3

C) Second round of crystallization optimization

EV-D68 $2A^{pro}$ sample was received on blue ice, 10.0 mg/ml. Sample was aliquot into three tubes with the same volume. Sample 1 was diluted into 8 mg/ml using the purification buffer. Sample 2 was kept as 10 mg/ml. Sample 3 was concentrated at 3,000 g at 4 °C until the concentration reached 13.0 mg/ml. All three samples were centrifuged at 16,000*g* for 10 min at 4 °C before setting up crystallization. Crystallization was set up at room temperature using Crystal Gryphon. In this round, desired protein concentration was optimized (Table 3.3). All screens were incubated in 20 °C storage and images were taken by Rigaku.

Table 3.3

Second-round EV-D68 2A^{pro} crystallization optimization

Sample	Screen	Protein: Precipitant	Concentration
1	OPT-VP- 4	2:1 in subwell1	8 mg/ml
2		1:1 in subwell2	10 mg/ml
3			13 mg/ml

D) Third round of crystallization optimization

EV-D68 2A^{pro} sample was received on blue ice, 10.0 mg/ml. Sample was aliquot into three tubes with the same volume. Sample 1 was diluted into 8 mg/ml using the purification buffer. Sample 2 was kept as 10 mg/ml. Sample 3 was concentrated at 3,000 g at 4 °C until the concentration reached 13.0 mg/ml. All three samples were centrifuged at 16,000*g* for 10 min at 4 °C before setting up crystallization. Crystallization was set up at room temperature using Crystal Gryphon. In this round, different crystallization volumes would be tested to find out if bigger crystals would grow. Additionally, different concentrations of the sample were tested again in this larger volume. Screen was incubated in 20 °C storage, and inspections were conducted manually. Eight crystallization conditions (named E3-E6, F3-F6) were chosen based on the second-round optimization result (Table 3.4)

Table 3.4

Third-round EV-D68 2A^{pro} crystallization optimization

	1	***2***	***3***	***4***	***5***	***6***
A	E3	F3	E3	F3	E3	F3
B	E4	F4	E4	F4	E4	F4
C	E5	F5	E5	F5	E5	F5
D	E6	F6	E6	F6	E6	F6

Detailed crystallization conditions are listed below

(E3) 1430.3 μl dH2O, 0.05 M Sodium cacodylate, 15% v/v 2-Propanol, 0.0025 M MgCl2, 21.8 μl CaCl2, 0.001 M Spermine.

(E4) 1422.1 μl dH2O, 0.05 M Sodium cacodylate, 342.9 μl 2-Propanol, 0.0025 M MgCl2, 30 μl CaCl2, 0.001 M Spermine.

(E5) 1412.1 μl dH2O, 0.05 M Sodium cacodylate, 342.9 μl 2-Propanol, 0.0025 M MgCl2, 40 μl CaCl2, 100 μl Spermine.

(E6) 1402.1 μl dH2O, 0.05 M Sodium cacodylate, 342.9 μl 2-Propanol, 0.0025 M MgCl2, 50 μl CaCl2, 100 μl Spermine.

(F3) 1425.3 μl dH2O, 0.05 M Sodium cacodylate, 342.9 μl 2-Propanol, 0.005 M MgCl2, 21.8 μl CaCl2, 100 μl Spermine.

(F4) 1417.1 μl dH2O, 0.05 M Sodium cacodylate, 342.9 μl 2-Propanol, 0.005 M MgCl2, 30 μl CaCl2, 100 μl Spermine.

(F5) 1407.1 μl dH2O, 0.05 M Sodium cacodylate, 342.9 μl 2-Propanol, 0.005 M MgCl2, 40 μl CaCl2, 100 μl Spermine.

(F6) 1397.1 μl dH2O, 0.05 M Sodium cacodylate, 342.9 μl 2-Propanol, 0.005 M MgCl2, 50 μl CaCl2, 100 μl Spermine

For each condition, 500 μl of precipitant will be pipette into the reservoir, and 2.5 μl of precipitant will be transferred onto the well. 5 μl of the protein sample will then be added to the well. After all conditions finished, the plate will be covered by Crystal Clear Sealing Tape and stored at 20° C for manual inspection.

3. EV-D68 $2A^{N84T}$ Crystallization

A) First round of crystallization screening

EV-D68 $2A^{N84T}$ was received on dry ice at 6.6 mg/ml and was centrifuged at 16,000*g* for 10 min at 4 °C before setting up crystallization. Crystallization was set up at room temperature using Crystal Gryphon. 6 commercialized crystallization screens were tried for each sample. Detailed information on crystallization setup for each sample shows in Table 3.5. All screens were incubated in 20 °C storage and images were taken by Rigaku.

Table 3.5

First-round of EV-D68 2A^{N84T} crystallization screening setup

Sample	Concentration	Protein: Precipitant	Screen
EV-D68 2A^{N84T}	6.6 mg/ml	2:1 in subwell1	HR2-139 PEG/ION HT
		1:1 in subwell2	HR2-134 Index HT
			HR2-131 Natrix HT
			HR2-137 MembFac HT
			HR2-248 Grid Screen Salt HT
			HR2-130 Crystal Screen HT

B) Second round of crystallization screening

EV-D68 2A^{N84T} was received on dry ice at 10 mg/ml and was concentrated at 3,000 g at 4 °C until the concentration reached 13.0 mg/ml. The sample was centrifuged at 16,000*g* for 10 min at 4 °C before setting up crystallization. Crystallization was set up at room temperature using Crystal Gryphon. 6 commercialized crystallization screens were tried. Detailed information of crystallization setup for each sample shows in Table 3.6. Crystallization plates were stored at 20 ºC, imaged, and inspected using the CrystalMation System (Rigaku) with visible, cross-polarized filter and UV imaging modes.

Table 3.6

Second-round of EV-D68 2A^{N84T} crystallization screening setup

Sample	Concentration	Protein: Precipitant	Screen
EV-D68 2A^{N84T}	13.0 mg/ml	2:1 in subwell1	HR2-139 PEG/ION HT
		1:1 in subwell2	HR2-134 Index HT
			HR2-131 Natrix HT
			HR2-137 MembFac HT
			HR2-248 Grid Screen Salt HT
			HR2-130 Crystal Screen HT

C) First round of EV-D68 2A^{N84T} crystallization optimization

EV-D68 2A^{N84T} was received on dry ice at 10 mg/ml and was concentrated at 3,000 g at 4 °C until the concentration reached 13.0 mg/ml. The sample was centrifuged at 16,000*g* for 10 min at 4 °C before setting up crystallization. Crystallization was set up using 24-well plates. The optimization plate (Table 3.6) was designed based on the crystallization condition of the second round of screening. For each condition, 500 µl of precipitant will be pipette into the reservoir, and 3 µl of precipitant will be transferred onto the well. 3 µl of protein sample will then be added onto the well. After all conditions finished, the plate will be covered by Crystal Clear Sealing Tape and stored at 20° C for manual inspection.

Table 3.7

Second-round of EV-D68 2A^{N84T} crystallization screening setup

		1	2	3	4	5	6
A	Magnesium chloride/M	0.005	0.018	0.025	0.005	0.018	0.025
	Sodium cacodylate pH 6.5/M	0.05	0.05	0.05	0.05	0.05	0.05
	2-propanol/% (v/v)	10	10	10	10	10	10
	Spermine/M	0.003	0.003	0.003	0.01	0.01	0.01
B	Magnesium chloride/M	0.005	0.018	0.025	0.005	0.018	0.025
	Sodium cacodylate pH 6.5/M	0.05	0.05	0.05	0.05	0.05	0.05
	2-propanol/% (v/v)	15	15	15	15	15	15
	Spermine/M	0.003	0.003	0.003	0.01	0.01	0.01
C	Magnesium chloride/M	0.005	0.018	0.025	0.005	0.018	0.025
	Sodium cacodylate pH 6.5/M	0.05	0.05	0.05	0.05	0.05	0.05
	2-propanol/% (v/v)	10	10	10	10	10	10
	Spermine/M	0.005	0.005	0.005	0.015	0.015	0.015
D	Magnesium chloride/M	0.005	0.018	0.025	0.005	0.018	0.025
	Sodium cacodylate pH 6.5/M	0.05	0.05	0.05	0.05	0.05	0.05
	2-propanol/% (v/v)	15	15	15	15	15	15
	Spermine/M	0.005	0.005	0.005	0.015	0.015	0.015

4. EV-D68 2A^{C107A} crystallization

A) First round crystallization screening

EV-D68 2A^{C107A} sample was received on blue ice, 10.4 mg/ml. The sample was concentrated at 3,000 g at 4 °C until the concentration reached 13.0 mg/ml and then centrifuged at 16,000*g* for 10 min at 4 °C before setting up crystallization. Crystallization was set up at room temperature using Crystal Gryphon. 6 commercialized crystallization screens were tried for each sample. Detailed information of crystallization setup for each sample shows in Table 3.7A. All screens were incubated at 20 °C storage and images were taken by Rigaku.

B) Second round crystallization screening

A new sample of EV-D68 $2A^{C107A}$ was requested and was received on blue ice at ~10.0 mg/ml. The sample was concentrated at 3,000 g at 4 °C until the concentration reached 13.0 mg/ml and then centrifuged at 16,000*g* for 10 min at 4 °C before setting up crystallization. 6 commercialized crystallization screens were repeated (Table 3.7A), and crystallization was also set up using the same condition where the $2A^{N84T}$ proteins grew into large crystals in a buffer of 0.05 M sodium cacodylate pH 6.5, 10%(v/v) 2-propanol, 0.005 M magnesium chloride and 0.005 M Spermine. The crystallization was set up using the 24-well crystallization plate (Table 3.7B).

Table 3.8

First/Second-round of EV-D68 $2A^{C107A}$ crystallization screening setup

A. 96-well plates

Sample	Concentration	Protein: Precipitant	Screen
EV-D68 $2A^{C107A}$	13.0 mg/ml	2:1 in subwell1	HR2-139 PEG/ION HT
		1:1 in subwell2	HR2-134 Index HT
			HR2-131 Natrix HT
			HR2-137 MembFac HT
			HR2-248 Grid Screen Salt HT
			HR2-130 Crystal Screen HT

B. 24-well plate

		1	2
A	Protein	5 μL	3 μL
	Precipitant	2.5 μL	3 μL

B	Protein	5 μL	3 μL
	Precipitant	2.5 μL	3 μL
C	Protein	5 μL	3 μL
	Precipitant	2.5 μL	3 μL

*For each condition, 500 µl of precipitant was added to the reservoir. Then, 2.5 µl or 3 µl of precipitant was transferred from the reservoir onto the well. 5 µl or 3 µl of protein solution was then added onto the well.

5. Crystallization for data collection

All crystals sent for beamline data collection and used for model building and structural determination, were produced on 24-well crystallization plates.

The EV-D68 $2A^{pro}$ crystals were produced under the condition of 0.05 M Sodium Cacodylate pH 6.0, 0.05 M Magnesium Chloride, 0.002 M Calcium Chloride, 15% v/v 2-Propanol, and 0.001 M Spermine. The ratio of protein to precipitant was 2:1, and the crystallization temperature was 20 °C.

The EV-D68 $2A^{N84T}$ crystals were produced under the condition of 0.05 M Sodium Cacodylate pH 6.5, 10%(v/v) 2-propanol, 0.005 M Magnesium Chloride and 0.005 M Spermine. The ratio of protein to precipitant was 1:1, and the crystallization temperature was 20 °C.

The EV-D68 $2A^{C107A}$ crystals were produced under the condition of 0.05 M Sodium Cacodylate pH 6.5, 10%(v/v) 2-propanol, 0.005 M Magnesium Chloride and 0.005 M Spermine. The ratio of protein to precipitant was 2:1, and the crystallization temperature was 20 °C.

6. Structure determination

The crystals of $2A^{C107A}$ were transferred into a mother liquor solution containing 40% of glycerol as a cryoprotectant and were incubated for 10 min under 20 ºC before flash freezing in liquid nitrogen for data collection. X-ray diffraction data were collected remotely at the National Synchrotron Light Source II beamline 17-ID-2 on an Eiger 16M detector.

Single crystal diffraction data were processed by XDS (Kabsch, 2010), and the space group was determined to belong to P4. Partial twining was detected by Xtriage in Phenix (Adams et al., 2010). Then, the phase of $2A^{C107A}$ was determined using molecular replacement (MR) by the program BALBES on the CCP4 online server (Krissinel et al., 2018). The crystal structure of human rhinovirus 2 (HRV2) $2A^{pro}$ (PDB code: 2HRV) was used as the search model. To achieve a model with better statistics and geometry, the second round of molecular replacement using the initial MR model using Phaser-MR in Phenix, along with the poly-Alanine model to reduce bias, which gave a better MR resolution with proper density fit. The model was further refined by multiple rounds of phenix.refine, followed by Phenix Rosetta refinement (twin operator applied with default setting) and Buster refinement (Smart et al., 2012). Ramachandran outliers and rotamer outliers were manually fixed by KING (Chen et al., 2009).

The crystals of $2A^{N84T}$ were transferred into a mother liquor solution containing 40% of glycerol as a cryoprotectant and were incubated for 10 min under 20 ºC before flash freezing in liquid nitrogen for data collection. X-ray diffraction data were collected

remotely at the Lawrence Berkeley National Laboratory beamline 5.0.2 on a Pilatus3 6M detector. The phase of $2A^{N84T}$ was determined using MR by the program Phaser in Phenix, using $2A^{C107A}$ as the search model. The model was further refined using the same method described above.

The crystals of $2A^{pro}$ were transferred into a mother liquor solution containing 40% of glycerol as a cryoprotectant and were incubated for 10 min under 20 ºC before flash freezing in liquid nitrogen for data collection. X-ray diffraction data were collected remotely at the Advanced Photon Source beamline 19-BM on an ADSC Quantum 315r detector. The phase of $2A^{pro}$ was determined using MR by the program Phaser in Phenix, using $2A^{C107A}$ as the search model. The model was further refined using the same method described above.

The statistics for data collection and refinement are given in Table 3.8. The graphical representations within this article were made with the PyMOL Molecular Graphics System, version 1.4 (Schrödinger LLC) and Chimera (Pettersen et al., 2004).

Table 3.9

Data collection and refinement statistics

		$2A^{pro}$	$2A^{C107A}$	$2A^{N84T}$
Data collection				
	Space group	P4	P4	P4
	Cell dimensions			
	a (Å)	118.157	118.77	119.255
	b (Å)	118.157	118.77	119.255
	c (Å)	81.735	80.38	80.326
	β (°)	90	90	90
	Resolution (Å)	44.38-3.2 (3.26-3.2)	47.77-2.50 (2.50-2.59)	39.75-2.57 (2.662-2.57)
	R_{pim}	0.044 (0.496)	0.044 (0.213)	0.032(0.698)
	R_{meas}	0.120 (1.335)	0.062 (0.302)	0.045 (0.988)
	$I/\sigma(I)$	25.4 (1.3)	21 (3.8)	16.31 (1.00)
	Completeness (%)	99.92 (99.93)	100.0 (100.0)	99.94 (99.75)
	Redundancy	3.9	2.0	2.0
Refinement				
	Resolution (Å)	3.5	2.5	2.6
	No. of reflections	14,386	38,883	36,100
	R_{work}/R_{free}	0.238/0.285	0.219/0.240	0.216/0.242
	No. of non-hydrogen atoms			
		5,980	6,150	6,128
	Macromolecules			
	Ion	6	6	6
	Solvent	17	59	35
	Average B factor (Å^2)	117.15	64.77	73.96
	Twin operator/fraction	h,-k,-l/ 0.468	h,-k,-l/ 0.347	h,-k,-l/0.478
	Ramachandran statistics (%)			
	Most favored	92.52	94.76	97.26
	Allowed	7.24	5.24	2.74
	Outliers	0.24	0.00	0.00
	RMSD			

Bond lengths (Å)	0.003	0.006	0.006
Bond angles (°)	0.71	0.77	0.83

3.4 Results

1. First round EV-D68 $2A^{pro}$ crystallization

EV-D68 $2A^{pro}$ #2 didn't give promising crystal hits during 20 days of inspection. EV-D68 $2A^{pro}$ #6 gave initial crystals in several crystallization wells.

Figure 3.6A shows the brightfield and UV image of crystals under the condition of 0.04 M Magnesium chloride hexahydrate, 0.05 M Sodium cacodylate trihydrate pH 6.0, 5% v/v (+/-)-2-Methyl-2,4-pentanediol. Figure 3.1B shows the brightfield and UV image of crystals under the condition of 0.005 M Magnesium chloride hexahydrate, 0.002 M Calcium chloride dihydrate, 0.05 M Sodium cacodylate trihydrate pH 6.0, 15% v/v 2-Propanol, 0.001 M Spermine. Figure 3.1C shows the brightfield and UV image of crystals under the condition of 0.018 M Magnesium chloride hexahydrate, 0.05 M Sodium cacodylate trihydrate pH 6.5, 10% v/v 2-Propanol, 0.003 M Spermine.

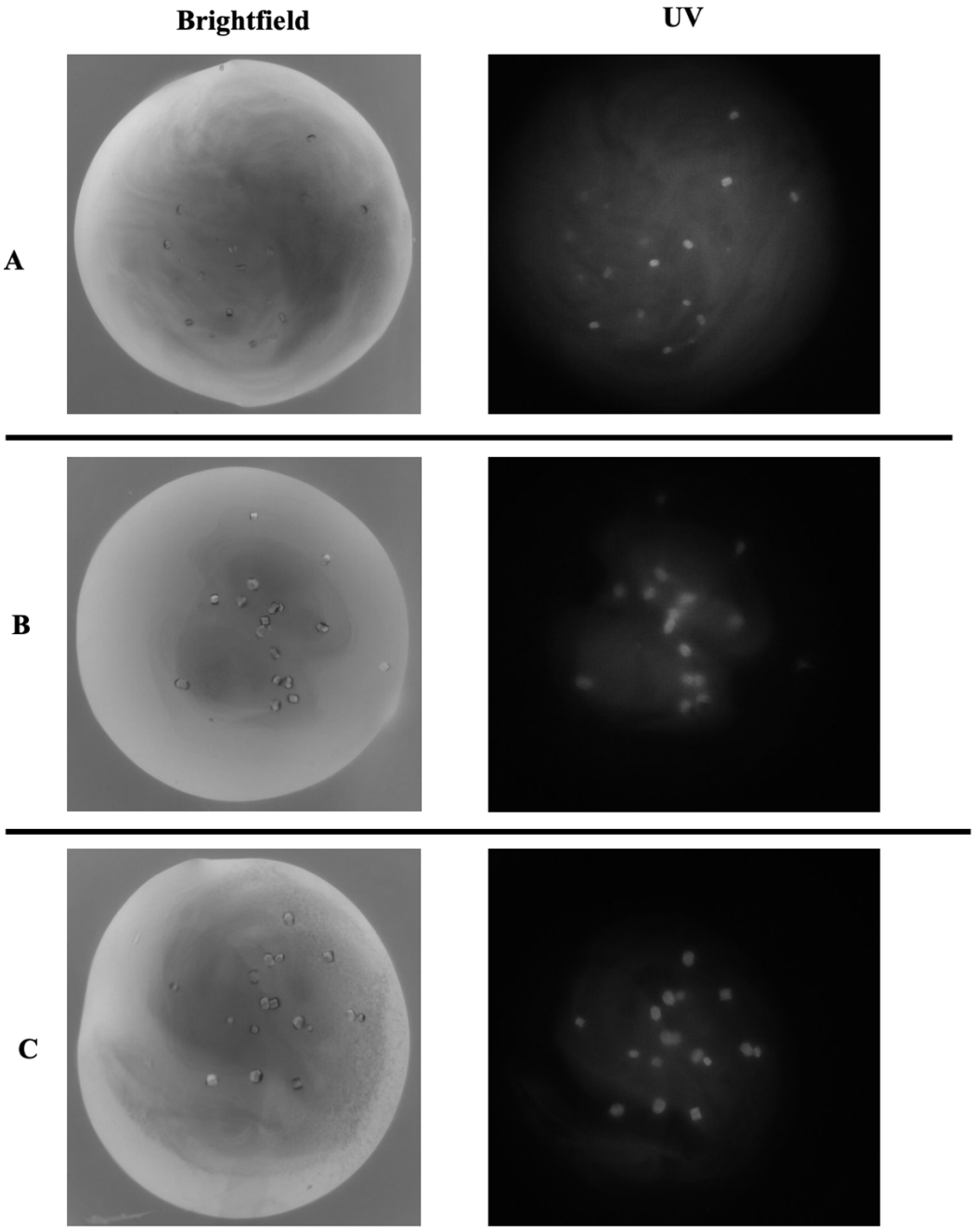


Figure 3.6. Brightfield and UV images of initial crystal hits of D682A-WT #6 under three crystallization conditions.

Compared to the result of EV-D68 $2A^{pro}$ #2 and EV-D68 $2A^{pro}$ #6, protein concentration was the determining factor of crystal formation. 13 mg/ml as EV-D68 $2A^{pro}$ #6 was chosen as the concentration for future optimizations rather than 6.2 mg/ml.

According to all crystallization conditions of first-round screening of EV-D68 $2A^{pro}$ #6, the optimization conditions should contain Magnesium chloride hexahydrate, Calcium chloride dihydrate, Sodium cacodylate trihydrate pH 6.0, 2-Propanol, and Spermine. The concentration of each component would be optimized based on their initial concentration.

2. First round of EV-D68 $2A^{pro}$ crystallization optimization

Crystals were observed in optimization screens, and crystal sizes were in the range of 10-40 μm. Two conditions gave the biggest crystals among all wells (Figure 3.7), and crystals were observed on Day1 and stopped growing bigger until Day5. According to the best two wells, the concentration of Sodium Cacodylate pH 6.0 can be set as 0.05 M, the concentration of Magnesium Chloride can be set as 0.005 M, and the concentration of Spermine can be set as 0.001 M. As for the rest of the two components, further optimization should be designed, and the concentration range of 2-propanol was set between 11.4% - 17.1% v/v and the concentration range of Calcium Chloride was set between 0.001-0.013 M.

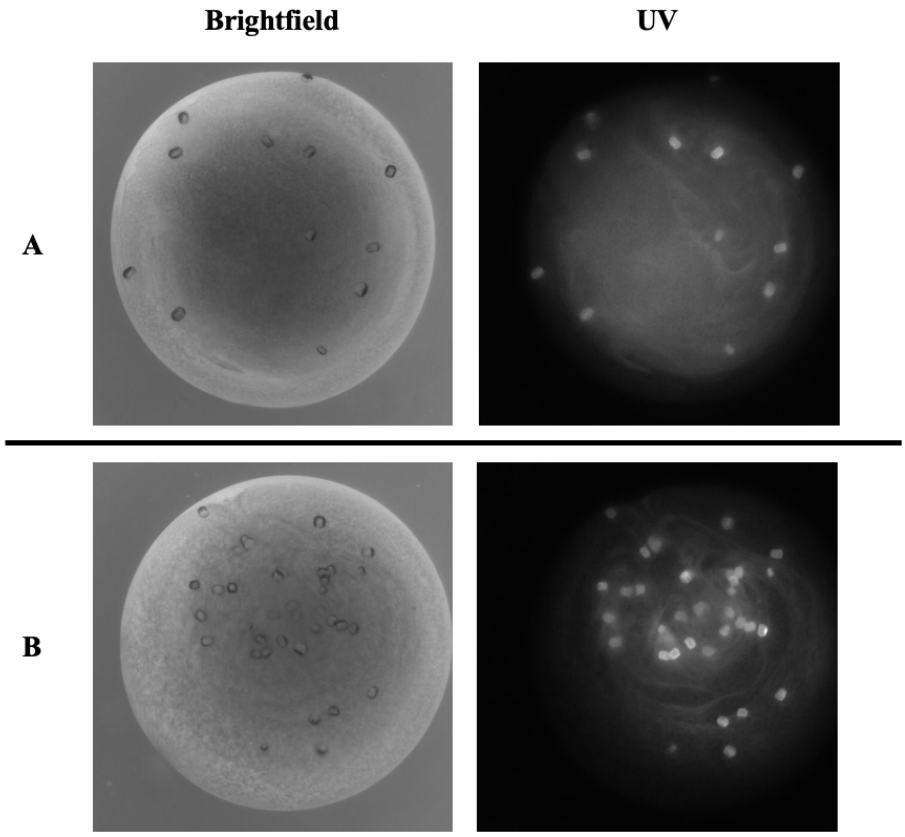


Figure 3.7. Brightfield and UV images of crystal of D682A-WT for first round optimization under three crystallization conditions. Crystallization condition for image A: 0.005 M Magnesium Chloride, 0.05 M Sodium Cacodylate pH 6.0, 0.001 M Calcium Chloride, 17.1% v/v 2-propanol, 0.001 M Spermine. Crystallization condition for image B: 0.005 M Magnesium Chloride, 0.05 M Sodium Cacodylate pH 6.0, 0.013 M Calcium Chloride, 11.4% v/v 2-propanol, 0.001 M Spermine.

3. Third-round of EV-D68 $2A^{pro}$ crystallization optimization

Crystals with the size around 50 μm were observed in the larger volume crystallization no matter the protein concentration and the condition with 0.05 M Sodium Cacodylate pH

6.0, 0.05 M Magnesium Chloride, 0.002 M Calcium Chloride, 15% v/v 2-Propanol, 0.001 M Spermine gave the relatively larger crystal among all wells (Figure 3.8).

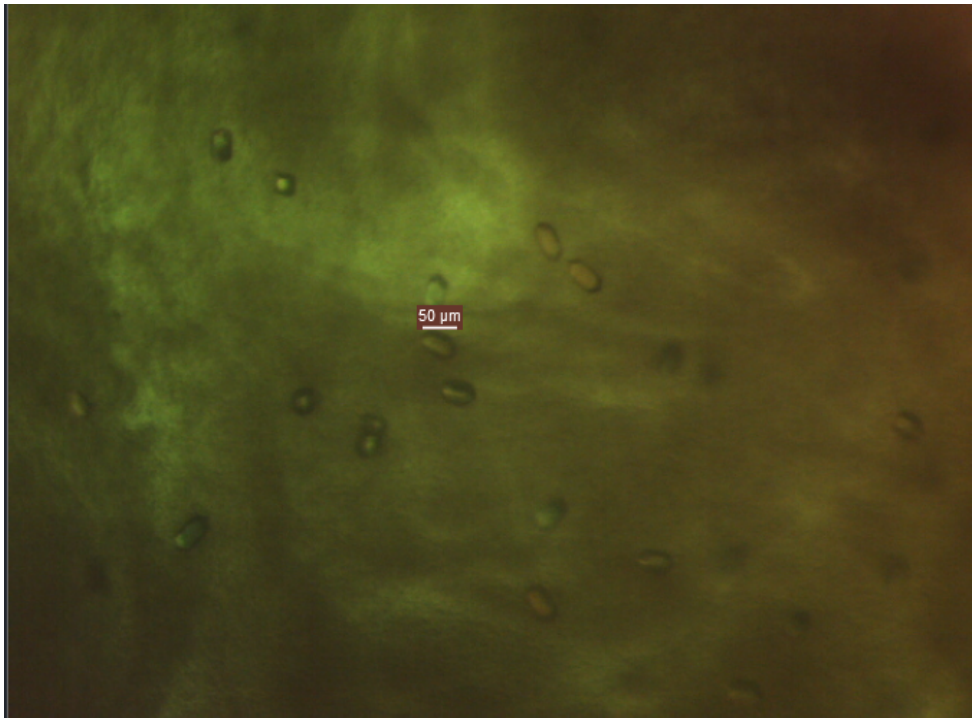


Figure 3.8. Brightfield image of crystal of D682A-WT for third-round optimization.

4. First round of EV-D68 $2A^{N84T}$ crystallization screening

No crystal was observed in this initial screening, which might be due to the relatively low protein concentration, considering the result EV-D68 $2A^{pro}$.

5. Second round of EV-D68 $2A^{N84T}$ crystallization screening

Only one single crystal (Figure 3.9) was observed in the condition containing 0.018 M Magnesium chloride, 0.05 M Sodium Cacodylate trihydrate pH 6.5, 10% 2-propanol, and 0.003 M Spermine. This condition is highly similar to the crystallization condition of EV-

D68 $2A^{pro}$. The significant difference is the ratio of protein to precipitant of EV-D68 $2A^{N84T}$ is 1:1, other than 2:1 as the EV-D68 $2A^{pro}$, which indicated that EV-D68 $2A^{N84T}$ needs less protein concentration to nucleate compared to EV-D68 $2A^{pro}$.

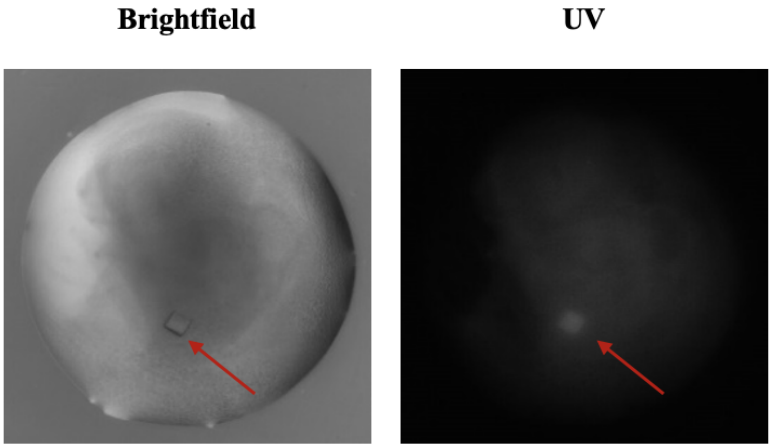


Figure 3.9. Brightfield and UV image of a single crystal for EV-D68 $2A^{N84T}$ in the second round of crystallization screening.

6. First round of EV-D68 $2A^{N84T}$ crystallization optimization

Crystal was reproduced in the optimized conditions, and the biggest crystals were found in the well (Figure 3.10) with 0.05 M Sodium Cacodylate pH 6.5, 10%(v/v) 2-propanol, 0.005 M Magnesium Chloride, and 0.005 M Spermine.

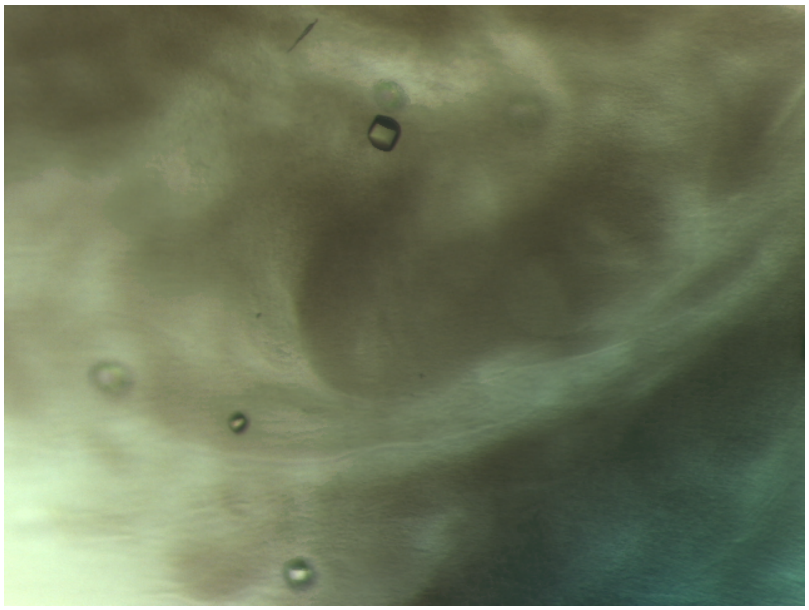


Figure 3.10. Brightfield image of crystals for EV-D68 $2A^{N84T}$ in the first round of crystallization optimization.

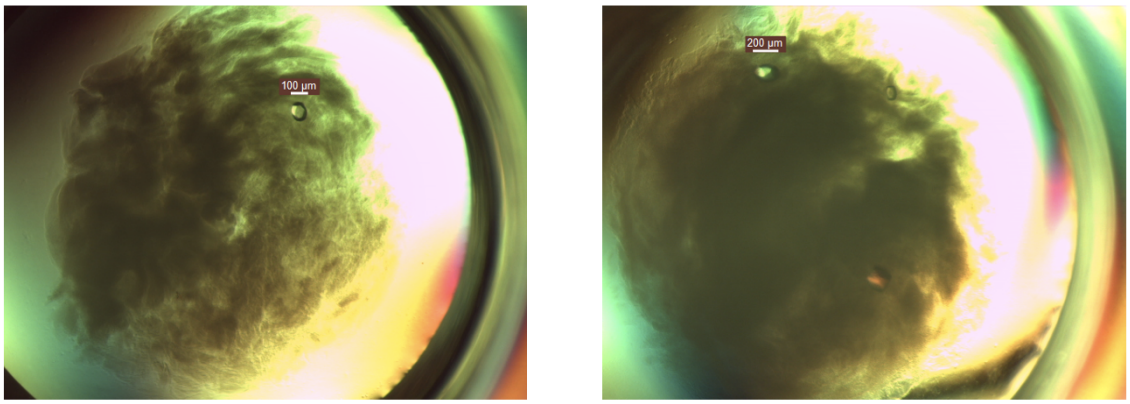
7. First round of EV-D68 $2A^{C107A}$ crystallization screening

No crystal was observed in neither of the plates. The image was taken on the received sample, and yellow color was observed (Figure 3.12), which is different from the more transparent color of previous samples.

8. Second round of EV-D68 $2A^{C107A}$ crystallization screening

Crystals were observed in the wells of the 24-well plate setup (Figure 3.11). The crystals in the 2:1 ratio of protein to precipitant were double the size of the crystals in the 1:1 ratio wells, which indicates that for EV-D68 $2A^{C107A}$, the higher concentration of the protein, the bigger the crystal yield. However, heavy protein precipitant was observed in the well

as well and occurred on the first day of setup, so higher protein concentration would not be perused.



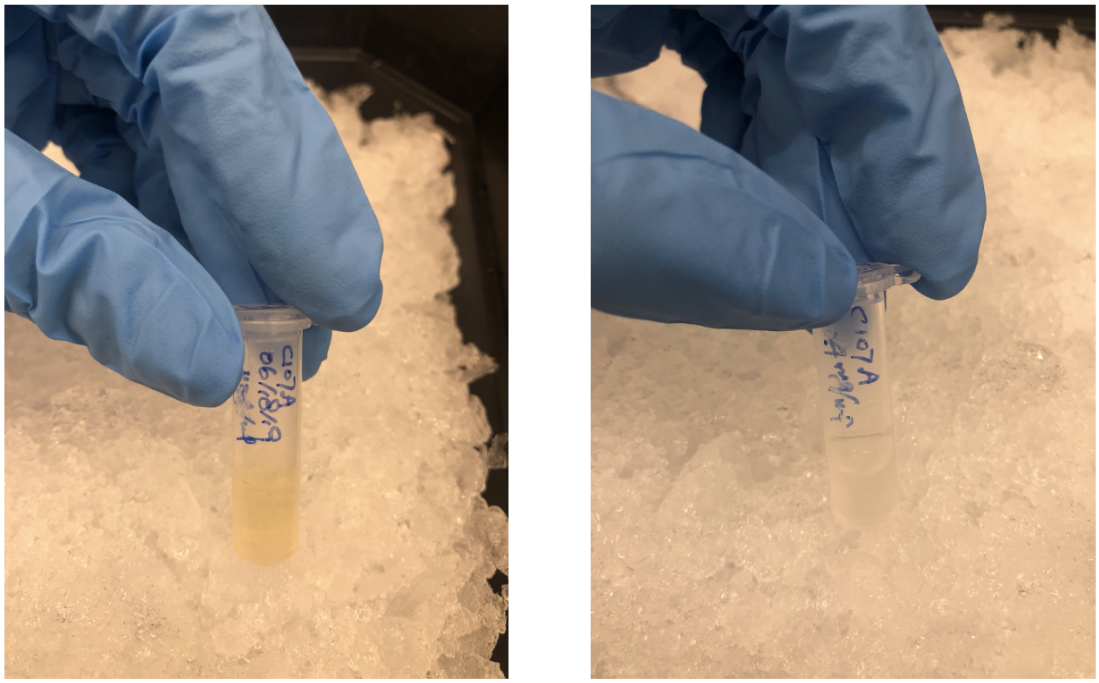
Protein: precipitant = 1:1 **Protein: precipitant = 2:1**

Figure 3.11. Brightfield image of crystals for EV-D68 $2A^{C107A}$ in the second round of crystallization screening. The size bar shows the approximate size of the crystals. For the well with a 1:1 ratio of protein to precipitant, the size of the crystal is about 100 μm. For the well with a 2:1 ratio of protein to precipitant, the size of the crystal is about 200 μm.

The sample image was taken of this batch. Compared to the image of two batches of EV-D68 $2A^{C107A}$ samples, it's obvious that the first batch of the sample had a bright yellow color. Additionally, an experiment was set up to compare the crystallization results for these two batches of protein (Figure 3.12). The crystallization condition used for each well was 0.05 M Sodium Cacodylate, 10% (v/v) 2-Propanol, 0.005 M Magnesium Chloride, 0.005 M Spermine.

The results demonstrate that the first batch sample was not able to produce crystal, while the second batch sample could. For the second batch sample, the higher ratio of protein to

precipitant wells was more likely to produce crystals, and higher protein concentration was pro to yield bigger crystals up to 200 μm and maintain higher reproducibility (Figure 3.13).



First batch **Second batch**

Figure 3.12. Images of different batches of EV-D68 2A^{C107A} upon receiving. The yellow color of the first batch of samples is obvious on the image. Where else, the second batch of samples is transparent, and no apparent color is observed.

A

	1	2	3	4	5	6
A	1st Batch 11 mg/ml Ratio 1:1	1st Batch 11 mg/ml Ratio 1:1	1st Batch 11 mg/ml Ratio 1:1	2nd Batch 10 mg/ml Ratio 1:1	2nd Batch 10 mg/ml Ratio 1:1	2nd Batch 10 mg/ml Ratio 1:1
B	1st Batch 11 mg/ml Ratio 2:1	1st Batch 11 mg/ml Ratio 2:1	1st Batch 11 mg/ml Ratio 2:1	2nd Batch 10 mg/ml Ratio 2:1	2nd Batch 10 mg/ml Ratio 2:1	2nd Batch 10 mg/ml Ratio 2:1
C	1st Batch 13 mg/ml Ratio 1:1	1st Batch 13 mg/ml Ratio 1:1	1st Batch 13 mg/ml Ratio 1:1	2nd Batch 13 mg/ml Ratio 1:1	2nd Batch 13 mg/ml Ratio 1:1	2nd Batch 13 mg/ml Ratio 1:1
D	1st Batch 13 mg/ml Ratio 2:1	1st Batch 13 mg/ml Ratio 2:1	1st Batch 13 mg/ml Ratio 2:1	2nd Batch 13 mg/ml Ratio 2:1	2nd Batch 13 mg/ml Ratio 2:1	2nd Batch 13 mg/ml Ratio 2:1

B

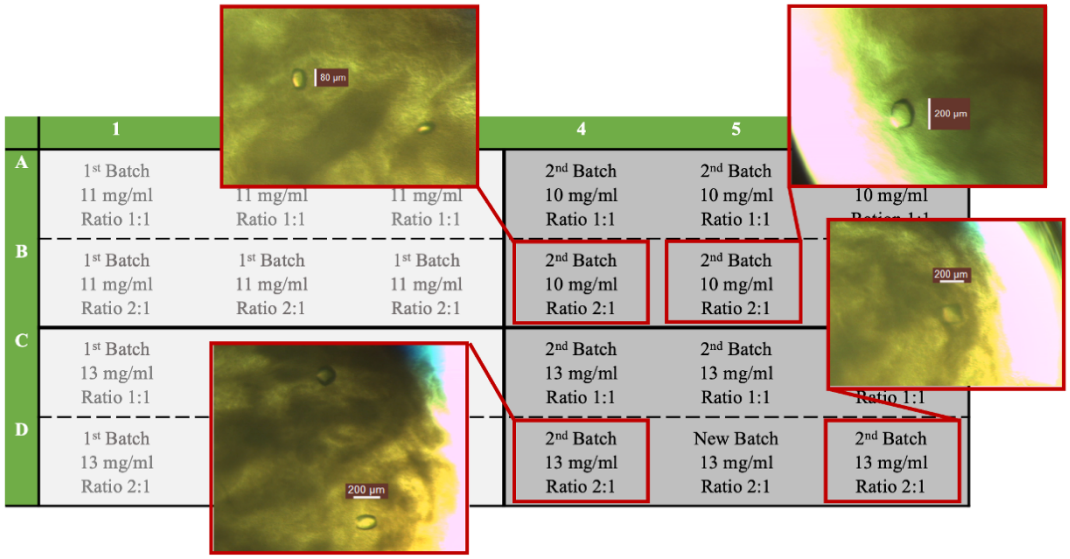


Figure 3.13. Images of different batches of EV-D68 2A^{C107A} crystallization design and results. A) crystallization condition design for two batches of samples. 1st and 2nd indicate the batch of protein, 10 mg/ml, 11mg/ml, and 13 mg/ml are the concentrations of corresponding protein sample, and the ratio is protein to precipitant.

9. Overall Structure of EV-D68 2A ^{pro}, EV-D68 2A ^{C107A,} and EV-D68 2A ^{N84T}

An EV-D68 2A^{C107A} mutant in which the active site Cys107 was mutated to Ala was crystallized at pH 6.5, the optimal pH. The crystal structure was resolved to 2.5 Å with

well-defined electron density from residues 4 to 140. There are six protein molecules (chains A to F) in the asymmetric unit. The overall structure of all six molecules is very similar. The numbers of residues built for different chains are slightly different from 136 residues to 142 residues. Therefore, the structure analysis is limited to molecule E because it has the most residues (142 residues) and also has the best geometry display,s so it is representative of all the molecules present in the asymmetric unit. Same as the $2A^{C107A}$, EV-D28 $2A^{N84T}$ is a hexamer in the unite cell and was resolved at 2.6 Å. The six chains of $2A^{N84T}$ have an average length of 142 residues and are highly structural similar. So here, we use chain E to perform structural analysis. As of EV-D28 $2A^{pro}$, the resolution of the structure is slightly lower than the previous two, which is 3.5 Å, with 6 monomeric chains. The chain B of $2A^{pro}$ was used to perform analysis due to its best geometry displays.

EV-D28 2A consists of two domains: An N-terminal domain comprising a four-stranded sheet (bI2, cI, eI2, and fI) and a C-terminal domain up of a six-stranded ß-barrel (aII, bII, cII2, dII, eII, and fII) with a tightly bound zinc atom. The N-terminal domain is linked to the C-terminal domain by a long inter-domain loop (residues 43-56). Within the N-terminal domain, a helical turn (residues Ile22-Ser27) connects cI to eI2. In the C-terminal domain, an antiparallel ß-hairpin formed by the bII2 and cII1 strands is located next to the six-stranded ß-barrel. (Figure 3.2a) Furthermore, this ß-hairpin also makes close contact with residues from the N-terminal domain. Three highly conserved residues, His18, Asp36, and Cys107 (mutate to Ala in this structure), were found in EV-D68 $2A^{pro}$, and from all enteroviruses (Figure 3.2c). Besides, four conserved residues,

Cys53, Cys55, Cys113, and His115, make up the zinc ion-binding site (Figure 3.14c). The structure of EV-D68 2A^{C107A}, EV-D28 2A^{N84T}, EV-D28 2A^{pro} are similar (Figure 3.2b and 3.2c).

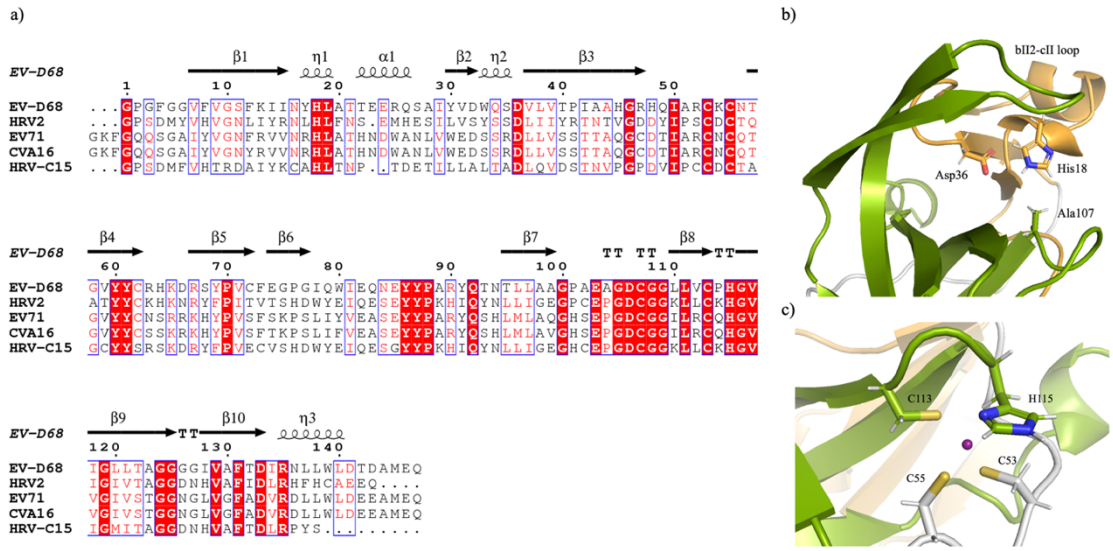


Figure 3.14 Analysis of the EV-D68 2A^{pro} a) The amino-acid sequences of 2A^{pro} from human Enterovirus D68 (EV-D68), human rhinovirus serotype 2 (HRV2-2A), human enterovirus 71 (EV71), coxsackievirus A16 (CVA16) and human rhinovirus C15 (HRV-C15) were aligned using Clustal Omega. (Madeira et al., 2019), and the result of the alignment was graphically displayed using ESPript (Gouet et al., 1999). Secondary-structure elements are indicated according to the structure of 2A^{pro} from EV-D68. b) The catalytic triad of EV-D68 2A^{C107A} is located beneath the bII2–cII1 loop. The active-site residues are shown as sticks. The active site Cys107 was mutated to Ala. c) A tetrahedral coordination site for the zinc ion in the EV-D68 2A^{C107A} structure.

3.5 Discussion

1. Comparison with other enterovirus $2A^{C107A}$ structures

The overall structure of EV-D68 $2A^{\mathbf{C107A}}$ shares a high degree of structural similarity with the $2A^{\mathbf{pro}}$ proteins (Figure 3.15a), as well as sequence identities with EV71 (49%), HRV-C15 (40%), and CVA16 (49%) (NCBI standard protein BLAST). The structure of EV-D68 $2A^{\mathbf{C107A}}$ could be superimposed onto the structure of EV71 $2A^{\mathbf{C107A}}$ with an overall r.m.s.d. (root-mean-square deviation of atomic positions) of 1.035 Å (Figure 3.15b). Similarly, an overall r.m.s.d. of 0.955 Å was observed when the structure of EV-D68 $2A^{C107A}$ was overlaid onto the structure of $2A^{\mathbf{pro}}$ from CVA16 (Figure 3.15c), which is the main factor related to human hand-foot-and-mouth disease (HFMD). The N-terminal domain of EV-D68 $2A^{\mathbf{C107A}}$ is very similar to that observed in EV71 $2A^{\mathbf{pro}}$, except for some minor differences in the flexible loops' orientations (Figure 3.15b). Compared to EV71 $2A^{\mathbf{pro}}$, the EV-D68 $2A^{C107A}$ possesses a longer bII1 and bII2 ß-strand. The a-helix connecting cI to eI2 in EV-D68 $2A^{\mathbf{C107A}}$ move outwards compared to that in the EV71 $2A^{pro}$ structure. The active site and the Zn^{2+} binding site are highly conserved between the two structures (Figure 3.15e).

The N-terminal domain of EV-D68 $2A^{\mathbf{C107A}}$ differs more than the C-terminal domain compared with the structure of CVA16 $2A^{\mathbf{C107A}}$ (Figure 3.15c). The orientation of C-terminal loops presents a noticeable difference between these two structures. Compared to CVA16 $2A^{\mathbf{pro}}$, all four ß-strands in N-terminal in EV-D68 $2A^{C107A}$ are shorter than those in CVA16 $2A^{\mathbf{pro}}$. Regarding the C-terminal structure, EV-D68 $2A^{C107A}$ presents a

longer bII1 ß-strand, a longer bII2 strand, and a longer dII ß-strand. Except for connecting by loop, there is an extra a-helix between cII and dII.

The overall r.m.s.d. between the crystal structure of $2A^{pro}$ from EV-D68 and HRV-C15 is 1.217 Å, which is higher than the r.m.s.d. across the crystal structures of EV-D68 $2A^{\mathbf{C107A}}$ and CVA16 $2A^{pro}$ (Figure 3.15d). When compared with the HRV-C15 $2A^{pro}$ sequence, the N-terminal of EV-D68 $2A^{C107A}$ processes two additional residues (Glu23 and Glu24). This region forms an a-helical structure in EV-D68. A loop occupies the corresponding region in HRV-C15 $2A^{pro}$. The C-terminal of EV-D68 $2A^{C107A}$ also differs when compared to the structure of HRV-C15 $2A^{\mathbf{pro}}$. Notably, the bII ß-strand is super long in HRV-C15 $2A^{pro}$. In contrast, in the EV-D68 $2A^{C107A}$ structure, a longer bII1 ß-strand is connected to a shorter bII2 ß-strand by a short loop.

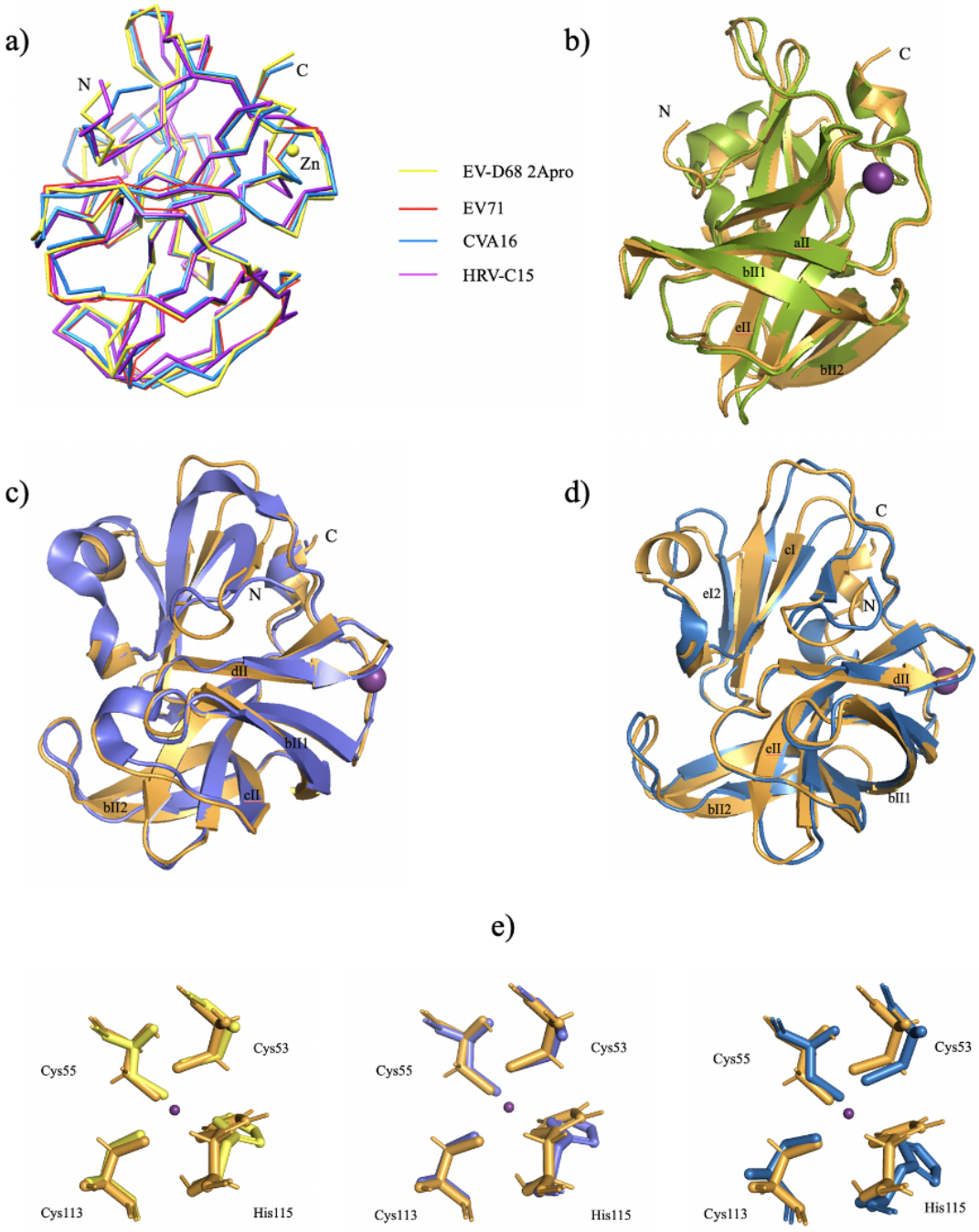


Figure 3.15. Comparison of the structure of EV-D68 2A^{pro} with homologous structures.

a) Superimposition of 2A^{pro} homologs. Structures are shown as ribbons.

b) Superimposition of the structure of EV-D68 2A^{pro} (shown in orange) with the structure

of EV71 $2A^{pro}$ (shown in yellow). c) Superimposition of the structure of EV-D68 $2A^{pro}$ (shown in orange) with the structure of CVA16 $2A^{pro}$ (shown in purple). d) Superimposition of the structure of EV-D68 $2A^{pro}$ (shown in orange) with the structure of HRV-C15 $2A^{pro}$ (shown in blue). e) Superimposition of the Zn binding sites of EV-D68 $2A^{pro}$ (shown in orange) with corresponding sites in the structure of EV71 $2A^{pro}$ (shown in yellow), CVA16 $2A^{pro}$ (shown in purple), and HRV-C15 $2A^{pro}$ (shown in blue).

2. Open conformation binding cleft

The surface rendering of EV-D68 $2A^{C017A}$ presents an open cleft, with the active site being accessible (Figure 3.16a). The cleft's width is about 12 Å, which makes the active site accessible for substrates. The confirmation of the active site of EV-D68 $2A^{C107A}$ is similar to the open conformation of $2A^{pro}$ observed in the structure of EV71 (Figure 3.16b) and HRV-C15 (Figure 3.16d), where both of the structure retains the width of the cleft about 12 Å. However, the active site His 18 is easier to approach in HRV-C15 $2A^{pro}$ compared to EV-D68, which may indicate a higher dynamic for a different binding substrate. The open cleft conformation promises the interaction between the active site on EV-D68 $2A^{C107A}$ and different substrates.

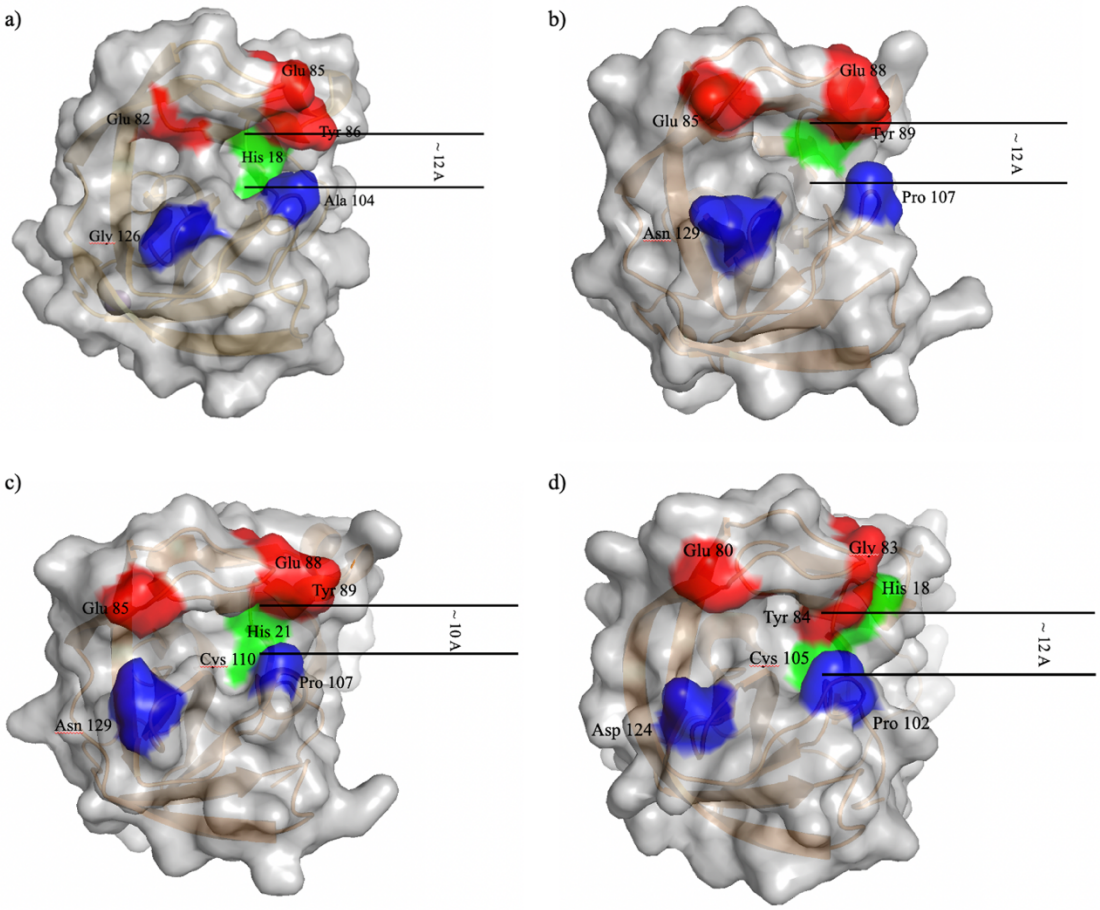


Figure 3.16. Open cleft conformation for substrate binding. a) Surface rendering of EV-D68 2A^{pro} shows an open cleft, with the active site being accessible. The width of the cleft is about 12 Å and is mainly estimated from a distance between Glu82 and Gly126 and the distance between Tyr86 and Ala104. b) Surface rendering of EV71 2A^{pro} shows an open cleft, with the active site being accessible. The width of the cleft is about 12 Å and is mainly estimated from a distance between Glu85 and Asn129 and the distance between Tyr89 and Pro107. c) Surface rendering of CVA16 2A^{pro} shows an open cleft, with the active site being accessible. The width of the cleft is about 10 Å and is mainly estimated from a distance between Glu85 and Asn129 and the distance between Tyr89 and Pro107. c) Surface rendering of HRV-C15 2A^{pro} shows an open cleft, with the active

site being accessible. The width of the cleft is about 12 Å and is mainly estimated from a distance between Glu80 and Asp124 and the distance between Tyr84 and Pro102.

CHAPTER 4

Structural Study Interactions of the Leukocyte Integrin $\alpha_M\beta_2$ with Cationic Ligand

4.1 Abstract

Integrins are important adhesion receptors that are associated with various vital cellular functions. Integrin $\alpha_M\beta_2$ is an integrin that involves in immunological processes via cell extravasation, phagocytosis, and immune synapse formation, so it is necessary to understand the molecular ligand-binding mechanism and activation mechanism of $\alpha_M\beta_2$. Pleiotrophin was proved to be a cationic ligand to $\alpha_M\beta_2$ via the specific binding sites involving the I domain on $\alpha_M\beta_2$. The preliminary crystallization condition and initial structural model of $\alpha_M\beta_2$-I domain in complex with pleiotrophin have been reported in this chapter, shedding light on the binding sites and guiding the crystallization optimization to get a complete and high-resolution structure in the future.

4.2 Introduction

Integrins are cell surface adhesion receptors composed of two non-covalently associated αβ heterodimers that play vital roles in lots of cellular functions, such as cell differentiation, cell adhesion, cell migration, and immune response (Hynes, 1992). Both integrin αβ subunit formed by a large multicomponent extracellular domain, a single membrane-spanning helix, and a short unstructured cytoplasmic tail (Shen, 2020). The major ligand-binding site of integrin was believed to be the α I-domain or von Willebrand factor A-type domain, which exist on half of the integrin α-subunits extracellular region as an inserted domain (KAMATA & TAKADA, 1994; MICHISHITA et al., 1993;

TUCKWELL et al., 1995). The homologous β I-domain will be the alternative ligand-binding site if no α I-domain exists in an integrin (Mould et al., 1997). Thus, it's necessary to understand the interaction between the integrin and ligand to reveal the mechanism of integrin activity.

Integrin $\alpha_M\beta_2$ is an integrin with the α I-domain, expressed on leukocytes like the myeloid, natural killer, and T cells (Shen, 2020). $\alpha_M\beta_2$ is associated with many immunological functions, including cell extravasation, phagocytosis, and immune synapse formation (Kinashi, 2007), so it's very important to study the molecular ligand-binding mechanism and activation mechanism of $\alpha_M\beta_2$. As mentioned before, the αM I-domain is related to integrin activation (Shen, 2020). The previously solved structure of αM I-domain shows that there are 6 β sheets surrounded by 7 α-helixes linked by flexible loops, which formed a classic dinucleotide binding or Rossmann fold (J. O. Lee et al., 1995). The ligand-binding site named metal ion-dependent adhesion site (MIDAS) is lying on the top face of the α_M I-domain ($\alpha_M I_d$).

Pleiotrophin (PTN) is a glycosaminoglycan-binding cytokine and growth factor with potent mitogenic and angiogenic activities (Shen, 2020). PTN has been proved to be associated with several significant physiological activities, such as injured tissue repair and regeneration (Perez-Pinera et al., 2008), maintenance of hematopoietic stem cells (Himburg et al., 2012), inflammation, and leukocyte recruitment (Ochiai et al., 2004; Yokoi et al., 2012). There are several proteins that have been studied as the receptor of PTN, such as the heparan sulfate proteoglycan (HSPG) N-syndecan and the chondroitin

sulfate proteoglycan receptor-type protein tyrosine phosphatase ζ (PTPRZ) (Maeda et al., 1996; Meng et al., 2000; Raulo et al., 1994). It's recently been discovered that $\alpha_M\beta_2$ is a receptor of PTN (Shen, 2020). The interaction between $\alpha_M\beta_2$ and PTN was initially accessed by the ability of Mac- 1-HEK293 to adhere to immobilized PTN, where demonstrates the fact that both HSPG and Mac-1 can act as receptors for PTN and confirmed by solution NMR spectroscopy (Shen, 2020). Additionally, the IC-21 cells adhesion assay was used to investigate the interaction of PTN-$\alpha_M\beta_2$ in immune cells. The little effect of heparin on the adhesion of IC-21 suggests that the surface of IC-21 cells largely contributes to the interaction with PTN (Shen, 2020). The binding domain on $\alpha_M\beta_2$ responsible for the interaction was determined to be the I-domain by using the biolayer interferometry (BLI). The binding affinity of PTN of soluble active $\alpha_M I_d$ is 1.2 ± 0.2 μM, higher than that of inactive $\alpha_M I_d$ (Shen, 2020). The presence of EDTA merely decreased the binding of active $\alpha_M I_d$, suggesting that the interaction is independent of the presence of Mg^{2+} (Shen, 2020). This study also revealed that the thrombospondin type-1 repeat (TSR) domains of PTN are involved in binding integrin $\alpha_M I_d$ (Shen, 2020). Here we report the preliminary crystallization condition of the PTN-$\alpha_M I_d$ complex and propose an initial structure model with a 4.0 Å resolution that provides a closer understanding of PTN-$\alpha_M I_d$ interaction and shed light on solving this complex structure in the future.

4.3 Materials and Methods

1. Materials and Instruments

Commercialized crystallization screens PEG/Ion HT (HR2-139), Index HT (HR2-134), Crystal Screen HT (HR2-130), and Natrix HT (HR2-131) were purchased from Hampton

Research. High-throughput crystallization robot Crystal Gryphon was purchased from Art Robbins Instruments. Crystallization plate storage and inspection robot Rigaku was purchased by Petra Fromme Group and was shared with the whole entire Biodesign center Center of Applied Structure Discovery (BCASD).

Protein samples were overexpressed and purified by Di Shen and Thi Thanh Hoa from Dr. Xu Wang's lab at Arizona State University.

2. First-round of crystallization screening

$\alpha_M I_d$ -PTN complex sample was obtained in a buffer solution (20 mM HEPES pH 7.0, 300 mM NaCl, 1 mM $MgCl_2$), on ice, at 7.5 mg/ml. The protein solution was centrifuged at 16,000*g* for 10 min before setting up crystallization. It was sitting on ice if possible since it's easy to crash out at 20 °C. 6 commercialization screens were tested (Table 4.1).

Table 4.1

First-round $a_M I_d$-PTN crystallization screening

Sample	Concentration	Protein: Precipitant	Screen
$\alpha_M I_d$-PTN	7.5 mg/ml	2:1 in subwell1	HR2-139 PEG/ION HT
			HR2-134 Index HT
		1:1 in subwell2	HR2-131 Natrix HT
			HR2-137 MembFac HT
			HR2-248 Grid Screen Salt HT
			HR2-130 Crystal Screen HT

For each well of the plate, 200 nL protein and 100 nL precipitant solution were dispensed in subwell1, and 100 nL protein and 100 nL precipitant solution were dispensed in subwell2. All screens were incubated in 12 °C storage, and images were taken by Rigaku.

3. Second-round of crystallization screening

$\alpha_M I_d$-PTN complex sample was obtained in a buffer solution (20 mM HEPES pH 7.0, 130 mM NaCl, 1 mM $MgCl_2$), on ice, at 10.0 mg/ml. The protein solution was centrifuged at 16,000*g* for 10 min before setting up crystallization. It was sitting on ice if possible since it's easy to crash out at 20 °C. 6 commercialization screens were tested (Table 4.2).

Table 4.2

First-round $a_M I_d$-PTN crystallization screening

Sample	Concentration	Protein: Precipitant	Screen
$\alpha_M I_d$-PTN	10.0 mg/ml	2:1 in subwell1	HR2-139 PEG/ION HT
		1:1 in subwell2	HR2-134 Index HT
			HR2-131 Natrix HT
			HR2-248 Grid Screen Salt HT
			HR2-130 Crystal Screen HT

For each well of the plate, 200 nL protein and 100 nL precipitant solution were dispensed in subwell1, and 100 nL protein and 100 nL precipitant solution were dispensed in subwell2. All screens were incubated in 4 °C storage, and plates were checked manually.

4. Third-round of crystallization screening

$\alpha_M I_d$-PTN complex sample was obtained in a buffer solution (20 mM HEPES pH 7.0, 130 mM NaCl, 1 mM $MgCl_2$), on ice, at 10.0 mg/ml. The protein solution was diluted with the same volume of buffer and then was centrifuged at 16,000*g* for 10 min before setting up crystallization. It was sitting on ice if possible since it's easy to crash out at 20 °C. 5 commercialization screens were tested (Table 4.3).

Table 4.3

Third-round $\alpha_M I_d$ -PTN crystallization screening

Sample	Concentration	Protein: Precipitant	Screen
$\alpha_M I_d$-PTN	10.0 mg/ml	2:1 in subwell1	HR2-139 PEG/ION HT
		1:1 in subwell2	HR2-134 Index HT
			HR2-131 Natrix HT
			HR2-137 MembFac HT
			HR2-248 Grid Screen Salt HT
			HR2-130 Crystal Screen HT

For each well of the plate, 200 nL protein and 100 nL precipitant solution were dispensed in subwell1, and 100 nL protein and 100 nL precipitant solution were dispensed in subwell2. All screens were incubated in 12 °C storage and images were taken by Rigaku.

5. First-round of crystallization optimization

$\alpha_M I_d$-PTN complex sample was obtained in a buffer solution (20 mM HEPES pH 7.0, 300 mM NaCl, 1 mM $MgCl_2$), on ice, at 10.0 mg/ml. The protein sample was divided into two aliquots. The first aliquot was diluted with the same volume of buffer, which kept the 300 mM NaCl. The second aliquot was diluted with the same volume of the buffer without NaCl, which made the concentration of NaCl drop to 150 mM. Then samples were centrifuged at 16,000*g* for 10 min before setting up crystallization. Three optimization screens (Table 4.4) were designed based on the previous crystallization condition, which contained 0.1 M Potassium chloride, 0.025 M Magnesium chloride hexahydrate, 0.05 M Sodium cacodylate trihydrate pH 6.0, and 15% v/v 2-Propanol.

Table 4.4

Optimization screens for first-round $\alpha_M I_d$-PTN crystallization optimization

Sample	Concentration	Protein: Precipitant	Screen
$\alpha_M I_d$-PTN (300 mM NaCl)	5.0 mg/ml	2:1 in subwell1 1:1 in subwell2	OPT-aM-1 OPT-aM-2 OPT-aM-3
$\alpha_M I_d$-PTN (150 mM NaCl)	5.0 mg/ml	2:1 in subwell1 1:1 in subwell2	OPT-aM-1 OPT-aM-2 OPT-aM-3

The detailed composition of each screen was deployed as follows (Figure 4.1 and 4.2). For OPT-aM-2, it was used as additive optimization; thus, each well contained the

original condition (except for spermine) plus 10 µl additives from the Additive Screen (Hampton Research HR2-428). All screens were incubated in 12 °C storage and images were taken by Rigaku.

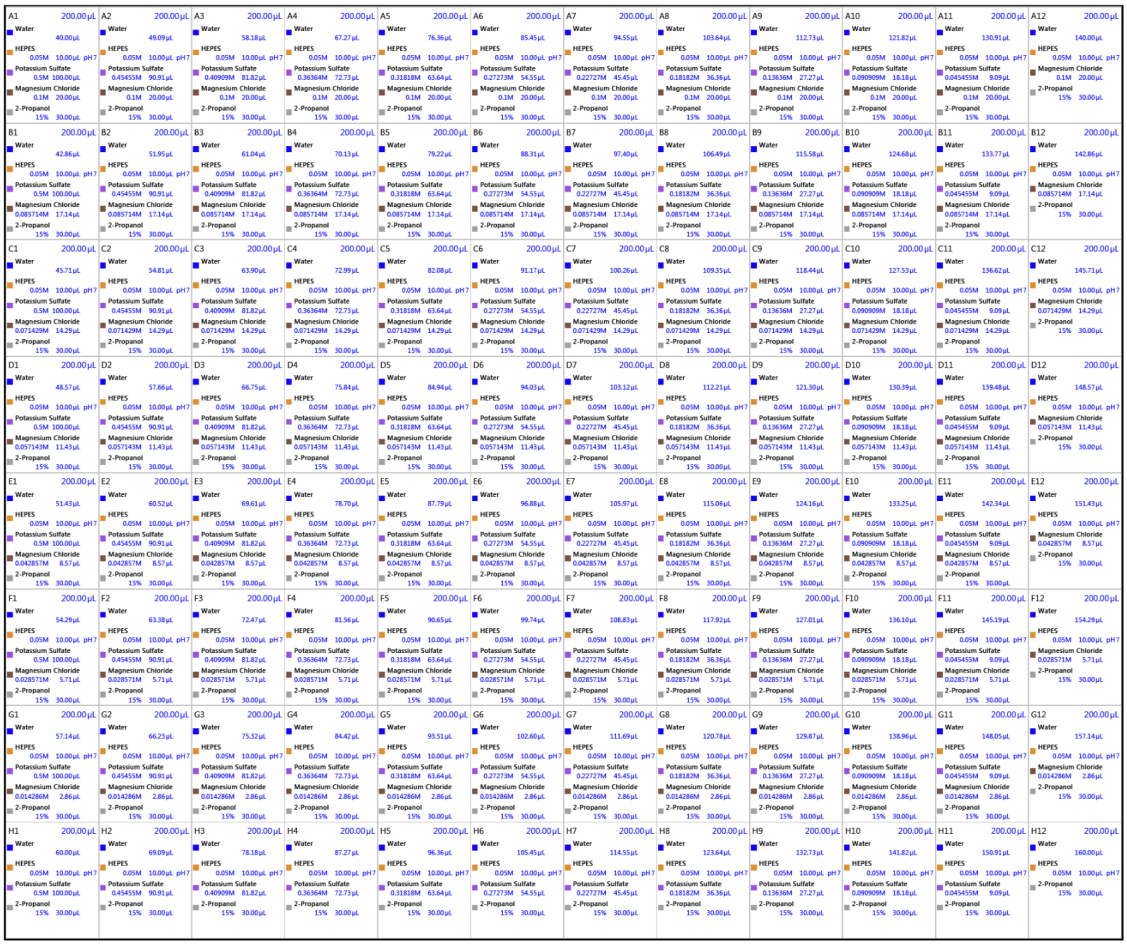


Figure 4.1. OPT-αM-1. Crystallization condition details of each well on OPT-aM-1. This screen was used to optimize the concentration of potassium chloride (0-0.5 M) and magnesium chloride (0-0.1 M). The concentration of sodium cacodylate pH6.0 was constant in each well of 0.05 M, and the concentration of 2-propanol was constant in each well of 15% v/v. The image was created by the screen-making robot Scorpion. The

“potassium sulfate” on the image represents “potassium chloride” because “potassium chloride” was not included in the robot database.

Figure 4.2. OPT-αM-3. Crystallization condition details of each well on OPT-aM-3. This screen was used to optimize the concentration of 2-propanol (0-30% v/v.). The concentration of sodium cacodylate pH 6.0 was constant in each well of 0.05 M and the same as potassium chloride of 0.1 M and magnesium chloride of 0.025 M. The image was created by screen making robot Scorpion. The “potassium sulfate” on the image

represents “potassium chloride” because “potassium chloride” was not included in the robot database.

6. Second-round of crystallization optimization

$\alpha_M I_d$-PTN complex sample was obtained in a buffer solution (20 mM HEPES pH 7.0, 300 mM NaCl, 1 mM $MgCl_2$), on ice, at 10.0 mg/ml. The protein sample was diluted with the same volume of the buffer without NaCl, which made the concentration of NaCl drop to 150 mM. Then the sample was centrifuged at 16,000*g* for 10 min before setting up crystallization. Five optimization screens (Table 4.5) were designed based on the previous crystallization condition of the first round of optimization. The detailed crystallization conditions were described as follows (Figure 4.3-4.7)

Table 4.5

Optimization screens for second-round aMId-PTN crystallization optimization

Sample	Concentration	Protein: Precipitant	Screen
$\alpha_M I_d$-PTN	5.0 mg/ml	2:1 in subwell1 1:1 in subwell2	OPT-aM-4 OPT-aM-5 OPT-aM-6 OPT-aM-7 OPT-aM-8

		1	2	3	4	5	6	7	8	9	10	11	12
A	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine tetrahydrochloride	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	140	137.5	135	132.5	130	127.5	130	127.5	125	122.5	120	117.5
B	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine tetrahydrochloride	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
C	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine tetrahydrochloride	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
D	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine tetrahydrochloride	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
E	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine tetrahydrochloride	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
F	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine tetrahydrochloride	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
G	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine tetrahydrochloride	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
H	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Spermine tetrahydrochloride	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	120	117.5	115	112.5	110	107.5	110	107.5	105	102.5	100	97.5

Figure 4.3. OPT-αM-4. Crystallization condition details (in μl) of each well on OPT-aM-4. This screen was used to optimize the concentration of sodium cacodylate pH 6.0 (0.0125 M-0.075 M), potassium chloride (0.025 M-0.1 M), and spermine tetrahydrochloride (0.005 M-0.0125 M).

		1	2	3	4	5	6	7	8	9	10	11	12
A	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Sarcosine	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	140	137.5	135	132.5	130	127.5	130	127.5	125	122.5	120	117.5
B	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Sarcosine	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
C	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Sarcosine	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
D	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Sarcosine	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
E	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Sarcosine	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
F	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Sarcosine	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
G	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Sarcosine	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
H	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Sarcosine	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	120	117.5	115	112.5	110	107.5	110	107.5	105	102.5	100	97.5

Figure 4.4. OPT-αM-5. Crystallization condition details (in µl) of each well on OPT-aM-5. This screen was used to optimize the concentration of sodium cacodylate pH 6.0 (0.0125 M-0.075 M), potassium chloride (0.025 M-0.1 M), and sarcosine (0.005 M-0.0125 M).

		1	2	3	4	5	6	7	8	9	10	11	12
A	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Betane hydrochloride	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	140	137.5	135	132.5	130	127.5	130	127.5	125	122.5	120	117.5
B	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Betane hydrochloride	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
C	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Betane hydrochloride	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
D	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Betane hydrochloride	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
E	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Betane hydrochloride	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
F	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Betane hydrochloride	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
G	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Betane hydrochloride	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
H	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1M Betane hydrochloride	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	120	117.5	115	112.5	110	107.5	110	107.5	105	102.5	100	97.5

Figure 4.5. OPT-αM-6. Crystallization condition details (in µl) of each well on OPT-aM-6. This screen was used to optimize the concentration of sodium cacodylate pH 6.0 (0.0125 M-0.075 M), potassium chloride (0.025 M-0.1 M), and betane hydrochloride (0.005 M-0.0125 M).

		1	2	3	4	5	6	7	8	9	10	11	12
A	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	30% v/v Methanol	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	140	137.5	135	132.5	130	127.5	130	127.5	125	122.5	120	117.5
B	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	30% v/v Methanol	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
C	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	30% v/v Methanol	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
D	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	30% v/v Methanol	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
E	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	30% v/v Methanol	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
F	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	30% v/v Methanol	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
G	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	30% v/v Methanol	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
H	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	30% v/v Methanol	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	120	117.5	115	112.5	110	107.5	110	107.5	105	102.5	100	97.5

Figure 4.6. OPT-αM-7. Crystallization condition details (in µl) of each well on OPT-aM-7. This screen was used to optimize the concentration of sodium cacodylate pH 6.0 (0.0125 M-0.075 M), potassium chloride (0.025 M-0.1 M) and methanol (1.5%-3.75% v/v).

		1	2	3	4	5	6	7	8	9	10	11	12
A	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1 M Spermidine	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	140	137.5	135	132.5	130	127.5	130	127.5	125	122.5	120	117.5
B	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1 M Spermidine	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
C	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1 M Spermidine	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
D	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1 M Spermidine	10	10	10	10	10	10	20	20	20	20	20	20
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
E	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	5	5	5	5	5	5	5	5	5	5	5	5
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1 M Spermidine	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	135	132.5	130	127.5	125	122.5	125	122.5	120	117.5	115	112.5
F	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	10	10	10	10	10	10	10	10	10	10	10	10
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1 M Spermidine	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	130	127.5	125	122.5	120	117.5	120	117.5	115	112.5	110	107.5
G	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	15	15	15	15	15	15	15	15	15	15	15	15
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1 M Spermidine	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	125	122.5	120	117.5	115	112.5	115	112.5	110	107.5	105	102.5
H	1M Sodium cacodylate pH 6.0	2.5	5	7.5	10	12.5	15	2.5	5	7.5	10	12.5	15
	1M KCl	20	20	20	20	20	20	20	20	20	20	20	20
	1M MgCl2	5	5	5	5	5	5	5	5	5	5	5	5
	80% 2-propanol	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5	37.5
	0.1 M Spermidine	15	15	15	15	15	15	25	25	25	25	25	25
	dH2O	120	117.5	115	112.5	110	107.5	110	107.5	105	102.5	100	97.5

Figure 4.7. OPT-αM-8. Crystallization condition details (in μl) of each well on OPT-aM-8. This screen was used to optimize the concentration of sodium cacodylate pH 6.0 (0.0125 M-0.075 M), potassium chloride (0.025 M-0.1 M), and methanol (0.005 M-0.0125 M).

7. Data collection and initial model building

The crystals of $\alpha_M I_d$-PTN were transferred into a mother liquor solution containing 40% of glycerol as a cryoprotectant and were incubated for 10 min under 20 ºC before flash freezing in liquid nitrogen for data collection. X-ray diffraction data were collected

remotely at the Argonne Advanced Photon Source (GM/CA) beamline 23-ID-B on a Dectris Pilatus3-6M detector.

Single crystal diffraction data were processed by the JBluIce data processing pipeline named GMCAproc, and the space group was determined to belong to $P12_11$. Then, the phase of aMId-PTN was determined using molecular replacement (MR) by the program Phaser-MR on the Phenix (Adams et al., 2010). The crystal structure of the human Integrin CR3 I-domain (PDB code: 1IDO) and human pleiotrophin (PDB code: 2N6F, truncated as C-terminal-domain) were used as the search model. The model was further refined by multiple rounds of phenix.refine, followed by Real-space refinement in the Phenix suite (highest resolution was adjusted to 4.0 Å during refinement).

4.4 Results

1. First-round crystallization screening

Protein crashed out or precipitated in most of the wells (Figure 4.8), which is consistent with the fact that $\alpha_M I_d$-PTN is temperature sensitive and easy to crash out. No promising crystal was observed.

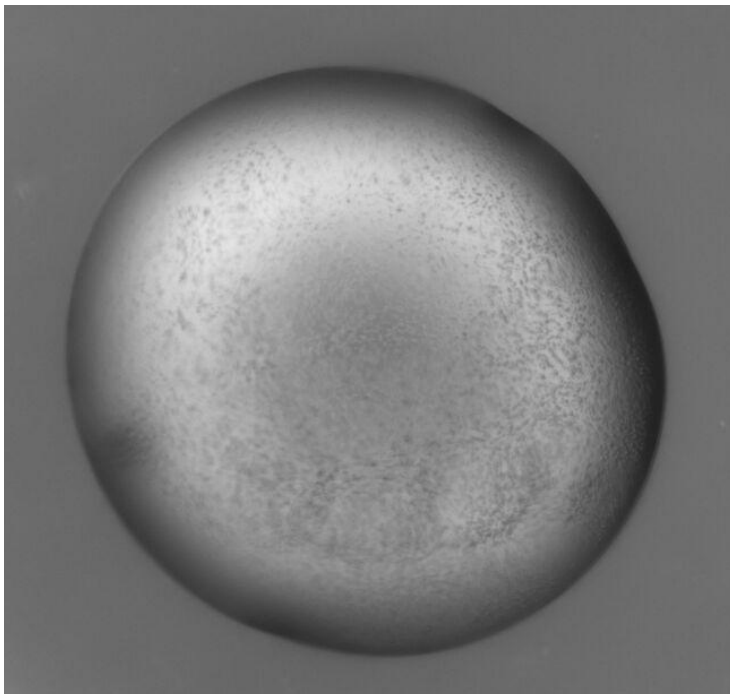


Figure 4.8. Sample brightfield image of first-round crystallization screening.

2. Second-round crystallization screening

Similar observations of heavy protein precipitation in most of the wells. Due to the instability of protein and ease of precipitation at high concentrations, lower protein concentrations would be tested in the next round.

3. Third-round crystallization screening

Crystals were observed (Figure 4.9A) in the condition with 0.1 M Potassium chloride, 0.025 M Magnesium chloride hexahydrate, 0.05 M Sodium cacodylate trihydrate pH 6.0, and 15% v/v 2-Propanol. Crystals diffracted up to 4 Å.

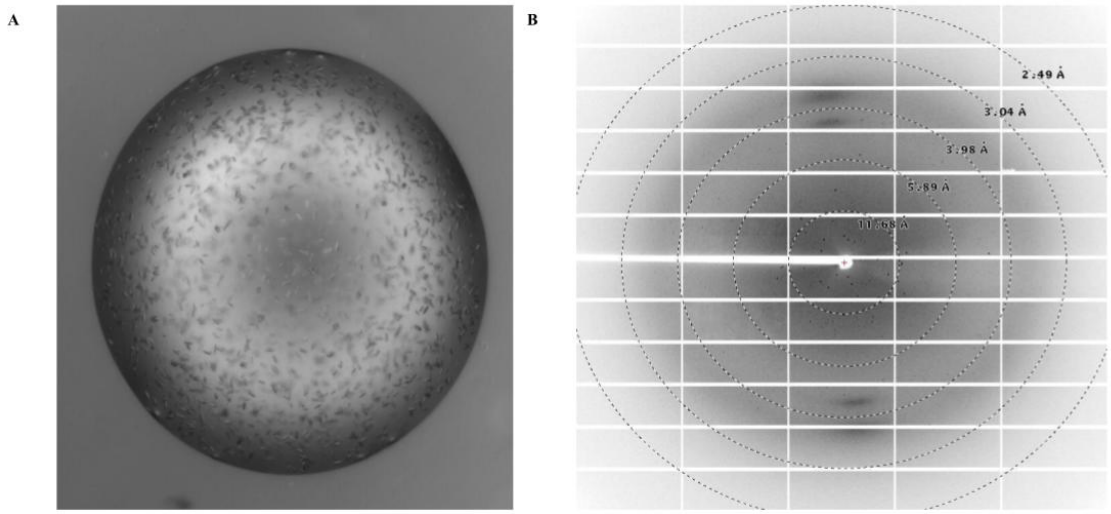


Figure 4.9. Result of third-round crystallization screening. A. Brightfield image of crystals for the first-round crystallization screening. B. Diffraction image of single crystals.

4. First-round crystallization optimization

Crystals were observed in this round of crystallization optimization, and bigger crystals were found in five conditions with additives as well as 150 mM NaCl. Besides composition of original crystallization condition, condition A contained 0.005 M Spermine tetrahydrochloride; condition B contained 0.005 M Sarcosine; condition C contained 0.1 M Betaine hydrochloride; condition D contained 0.005 M Spermidine; condition E contained 1.5% v/v Methanol.

Crystals were harvested and sent for diffraction experiments. This batch of crystals diffracted better compared to the last batch with more diffraction spots and higher resolution up to 3.5 Å.

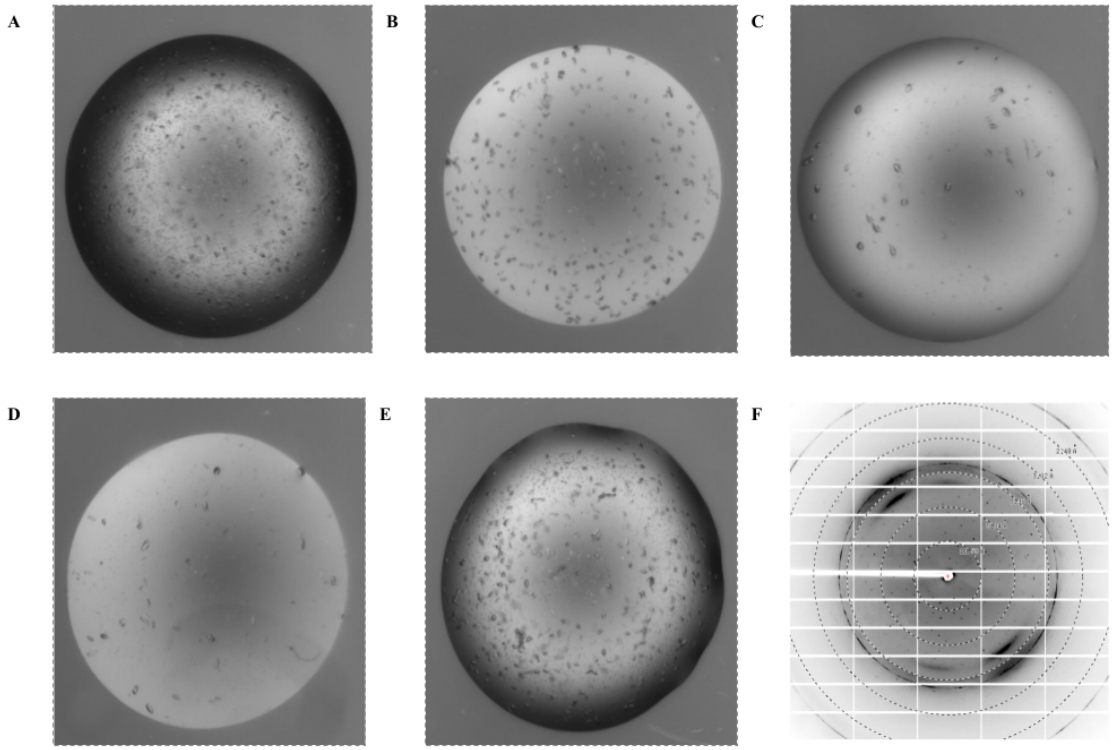


Figure 4.10. Result of first-round crystallization optimization. A-E shows the brightfield crystal images found in different conditions. F shows the diffraction pattern of crystals.

5. Second-round crystallization optimization

Biggest crystals were observed in three conditions: A. (OPT-aM-4) 0.0125 M Sodium Cacodylate pH6.0, 0.05 M Potassium Chloride, 0.025 M Magnesium Chloride, 15% v/v 2-propanol and 0.001 M Spermine tetrahydrochloride (Figure 4.11A). B. (OPT-aM-5) 0.025 M Sodium Cacodylate pH6.0, 0.05 M Potassium Chloride, 0.025 M Magnesium Chloride, 15% v/v 2-propanol and 0.005 M Sarcosine (Figure4.11B). C. (OPT-aM-5) 0.025 M Sodium Cacodylate pH6.0, 0.1 M Potassium Chloride, 0.025 M Magnesium Chloride, 15% v/v 2-propanol and 0.0075 M Sarcosine (Figure 4.11C).

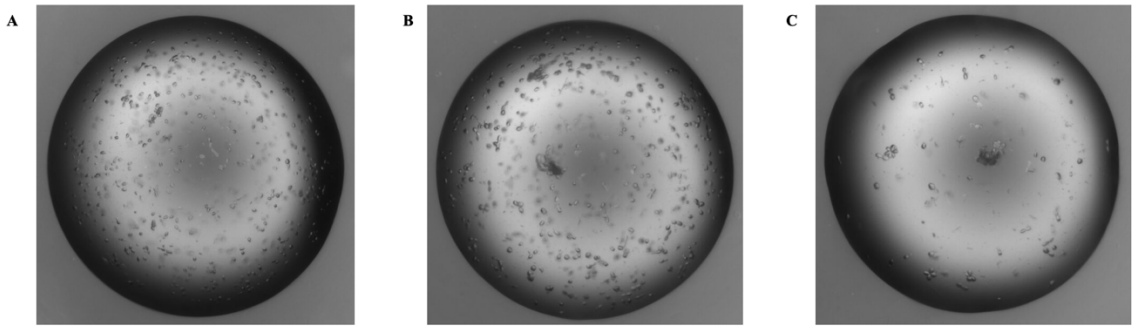


Figure 4.11. Result of second-round crystallization optimization. A-C show the brightfield crystal images found in different conditions.

6. $\alpha_M I_d$-PTN complex structural model

The complex of $\alpha_M I_d$-PTN was crystallized at pH 6.0, the optimal pH. The structural model was built to 4.0 Å from residues 132 to 315 of $\alpha_M I_d$ and residues 66-110 of PTN. There are two $\alpha_M I_d$ molecules (chains A and B) and one PTN molecule (chain C) in the asymmetric unit (Figure 4.12A). The overall structures of the two $\alpha_M I_d$ molecules are very similar. There is no obvious electron density for the C-terminal flexible loop region, so the model was only built up to residue 110.

The C-terminal TSR domain of PTN poses closely to the MIDAS domain on the $\alpha_M I_d$ (Figure 4.12B), which is consistent with the hypothesis that the TSR and MIDAS are the major binding sites. No metal ion was observed in this structural model.

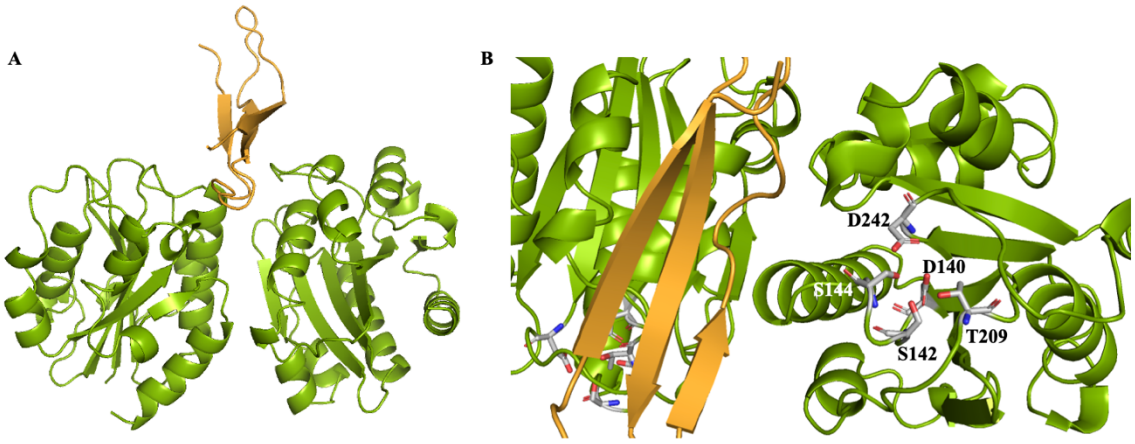


Figure 4.12. Structural model of the $\alpha_M I_d$-PTN complex. A shows the overall structural model. B shows the structural detail with a close look at the MIDAS domain where the sticks highlight the metal coordinating residues. The chains colored with green are the $\alpha_M I_d$ molecules, and the chain colored with orange is the PTN molecule.

4.5 Discussion

Although PTN was proved to be a novel ligand of integrin $\alpha_M I_d$and, the hypothesized major binding sites are the TRS of PTN and the MIDAS on $\alpha_M I_d$. However, due to the limitation in the active αM I-domain's spectral quality, not all the peaks for the active $\alpha_M I_d$ can be assigned in the HSQC spectrum (Shen, 2020). Thus, to discover the interaction between PTN and $\alpha_M \beta_2$ I-domain, a high-resolution structure of the $\alpha_M I_d$-PTN complex is necessary.

The crystallization of the $\alpha_M I_d$-PTN complex was difficult due to the instability of active state $\alpha_M I_d$ and the ease of protein precipitate under room temperature. In many cases of

soluble protein crystallization, higher protein concentration gives the better possibility to grow crystals and yield bigger crystals under the same condition. However, it's not the case for $\alpha_M I_d$-PTN, where 10 mg/ml of protein failed to grow any crystals and precipitated heavily. After many rounds of screening and optimization, the optimal crystallization concentration was determined to be 5 mg/ml of $\alpha_M I_d$ with the same molar amount of PTN added prior to crystallization. Also, since the instability of the protein, the sample needs to be kept on ice, if possible, when set up crystallization to reduce the protein precipitation, and the crystallization plates need to be stored at 12 °C for the same reason. The current optimal pH of crystallization is 6.0 with 0.0125 M Sodium Cacodylate, 0.05 M Potassium Chloride, 0.025 M Magnesium Chloride, 15% v/v 2-propanol. Two different additives were proved to facilitate the crystal growth, which is spermine tetrahydrochloride (0.001M) and sarcosine (0.005 M and 0.0075 M).

Although the size of $\alpha_M I_d$-PTN complex crystals has been increased with optimization, most of the crystals are still smaller than 20 μm which is not very suitable for crystal harvest. Besides that, problems such as double lattice and fast crystal diminish persisted. To solve these issues, there are two alternatives to try. Firstly, we can continue to optimize crystallization conditions, including pH, composition concentration, or discover more additives, which may lead to larger crystals with a size of 50 μm, but not guaranteed. Larger single crystals will be easier to harvest, which may avoid the double lattice issue. Additionally, larger crystals will survive better under radiation damage. The other method is to try current crystals with serial femtosecond crystallography (SFX). SFX allows thousands of microcrystals to be injected into the synchronized X-ray bunch

in random orientation so that enough patterns of various orientations can be collected to produce a complete data set for structure determination. The power of the radiation is high enough to destroy the crystals immediately in every encounter. The pulse duration enables to catch the diffraction patterns before the influence of radiation damage caused by scattered photons. Thus, XFELs allow better collections of high-resolution diffraction from small crystals compared to synchrotron radiation. The nature of "diffraction before destruction" also eliminates the need for cryocooling, which enables collections under room temperature that better mimic in vivo environments. The lipidic cubic phase (LCP) can be used as the delivery medium because its high viscosity can decrease the flow rate with lower sample waste to achieve a higher efficiency of crystal preparation. To prepare a sample suitable for SFX, crystallization conditions should be tested with the lipid to make sure any of the compositions won't interfere with the correct phase formation. Additionally, the crystallization condition needs to be optimized in a larger volume. At least ten to twenty microliters protein samples will be needed to prepare an LCP sample, but under current conditions, the crystallization volume is only 100 nl for each well, which requires a collection of at least 100 to 200 wells. Optimizing the crystallization volume up to 5 μl or trying batch crystallization will make this process easier.

With all mentioned barriers addressed, it's promising to solve the complex structure of $\alpha_M I_d$-PTN, which will reveal the mechanism of how PTN interacts with the active state of $\alpha_M I_d$.

CHAPTER 5

SUMMARY AND OUTLOOK

This thesis is focused on highlighting the significance of protein structures to structural based drug discovery. As the most used structure determination method, the extensive automation of third-generation synchrotron radiation has allowed high-throughput crystallography (Abola et al., 2000; Beteva et al., 2006; Cohen et al., 2002; Muchmore et al., 2000). X-ray crystallography has been upgraded and innovated to overcome major limitations. One of the limiting factors is the burden of handling and mounting crystals, which has been alleviated by *in situ* diffractions (Bingel-Erlenmeyer et al., 2011; Gelin et al., 2015; Jacquamet et al., 2004; le Maire et al., 2011) and automated crystal harvesting (Cipriani et al., 2012; Deller & Rupp, 2014; Yin et al., 2014). Additionally, growing big crystals were always a huge challenge for structural biologists until the application of micro-focused beams (Smith et al., 2012), which makes data collection possible on microcrystals (10-50 μm) or imperfect crystals. Radiation damages, however, were also a remaining issue, though it's been alleviated by cryo-cooling (Garman, 2010; Garman & Owen, 2006). X-ray free-electron laser (XFEL) enabled serial femtosecond crystallography (SFX) eliminated the radiation damage to the largest extent by the application of "diffract before destroy" (Chapman et al., 2011). The SFX allows microcrystals to be delivered through a viscous medium to an injection device and shot only once by a femtosecond-paused X-ray beam, leading to room-temperature, radiation-damage-free structures.

Despite the development of X-ray crystallography as a dispensable tool for structural based drug discovery, there are still sources of failure in generating crystals or structures of target complexes. The first and most obvious one is the high attrition rate in early experimental steps, including protein production, purification, and crystallization. In this case, computational tools might be a potential helper in selecting protein targets, designing protein constructs, and suggesting crystallization conditions based on protein sequences. Even though scientists managed to get the crystals of target the protein, it might still be difficult to get the crystals of the desired complex, just like the case of EV-D68 2A protease mentioned in this thesis. Additionally, the mixture of free protein and protein in complex in the crystal may also cause poor electron density, which makes it hard to fit the ligand into the structure. Thus, it is critical to use the right method to obtain the co-crystals and deicide what solvent to use to help the ligand solubilize and be compatible with the native state of the protein.

The first structure of EV-D68 2A protease was reported in this thesis, as well as the structure of its two mutants, $2A^{N84T}$ and $2A^{C107A}$. The $2A^{pro}$ was proved to be the popular drug target for EV-D68 infection associated with acute flaccid myelitis, which arose concerns on global health significantly, yet no treatment is accessible. The crystallization conditions of the $2A^{pro}$ and two mutants are quite similar due to their high similarity, and all of them adopt a structure of an N-terminal domain comprising a four-stranded sheet and a C-terminal domain made up of a six-stranded ß-barrel with a tightly bound zinc atom. The structure of $2A^{pro}$ is like the structures of other enterovirus 2A with a sub-one angstrom r.m.s.d. All three solved structures were crystallized in apo, and no condition

was found so far to grow complex crystals with a ligand. The reason might be the adding of ligand significantly changed the protein conformation; thus, original crystal packing would not fit anymore, so more crystallization condition screening is needed to dig out the optimal condition for complex crystals. The second route, other than co-crystallization, is soaking, where the apo crystals can be incubated with crystallization precipitant containing an appropriate amount of ligand. The potential pitfall of soaking is the original crystallization condition may not be compatible with the ligand, meaning the affinity of the ligand could be decreased. To increase the possibility of obtaining a complex structure, we will continue to test both co-crystallization and soaking in parallel.

The preliminary crystallization condition of the $\alpha_M I_d$-I domain in complex with Pleiotrophin was reported in this thesis, as well as the initial structural model. The journey of obtaining this complex crystal was long and difficult. Unlike other common soluble proteins, this complex crystalized better under lower concentration than higher and required some extra caution to keep in 12°C to reduce the possibility of precipitation. After rounds of crystallization optimization, the size of crystals (10 μm or even smaller) is still not satisfying the need to harvest the single crystal for data collection, and even diffractable crystals failed to tolerate radiation damages and diminished fast. Even though with an incomplete dataset, the initial structural model with built after many tries of figuring out the right data processing strategy. The space group of the $\alpha_M I_d$-PTN complex was determined to belong to $P12_11$. The initial model is composed of two monomers of $\alpha_M I_d$ with a single molecule of PTN lying on top. The C-terminal TSR domain of PTN poses closely to the MIDAS domain on the $\alpha_M I_d$, which is consistent with the hypothesis

that the TSR and MIDAS are the major binding sites(Shen, 2020). Further analysis of a complete and high-resolution structure of the $\alpha_M I_d$-PTN complex is essential to uncover the mechanism of how PTN interacts with the $\alpha_M I_d$. With the preliminary crystallization condition and the initial model, it's only a matter of time to obtain the atomic structure of the complex.

REFERENCE

Abola, E., Kuhn, P., Earnest, T., & Stevens, R. C. (2000). Automation of X-ray crystallography [Article]. *Nature Structural Biology*, *7*(11s), 973–977. https://doi.org/10.1038/80754

Adams, P. D., Afonine, P. V, Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution [Article]. *Acta Crystallographica. Section D, Biological Crystallography.*, *66*(2), 213–221. https://doi.org/10.1107/S0907444909052925

Audet, M., White, K. L., Breton, B., Zarzycka, B., Han, G. W., Lu, Y., Gati, C., Batyuk, A., Popov, P., Velasquez, J., Manahan, D., Hu, H., Weierstall, U., Liu, W., Shui, W., Katritch, V., Cherezov, V., Hanson, M. A., & Stevens, R. C. (2019). Crystal structure of misoprostol bound to the labor inducer prostaglandin E2 receptor [Article]. *Nature Chemical Biology*, *15*(1), 11–17. https://doi.org/10.1038/s41589-018-0160-y

Ayyer, K., Yefanov, O. M., Oberthür, D., Roy-Chowdhury, S., Galli, L., Mariani, V., Basu, S., Coe, J., Conrad, C. E., Fromme, R., Schaffer, A., Dörner, K., James, D., Kupitz, C., Metz, M., Nelson, G., Xavier, P. L., Beyerlein, K. R., Schmidt, M., … Chapman, H. N. (2016). Macromolecular diffractive imaging using imperfect crystals [Article]. *Nature (London)*, *530*(7589), 202–206. https://doi.org/10.1038/nature16949

Barends, T. R. M., Foucar, L., Botha, S., Doak, R. B., Shoeman, R. L., Nass, K., Koglin, J. E., Williams, G. J., Boutet, S., Messerschmidt, M., & Schlichting, I. (2014). De novo protein crystal structure determination from X-ray free-electron laser data [Article]. *Nature (London)*, *505*(7482), 244–247. https://doi.org/10.1038/nature12773

Batool, M., Ahmad, B., & Choi, S. (2019). A structure-based drug discovery paradigm. *International Journal of Molecular Sciences*, *20*(11). https://doi.org/10.3390/ijms20112783

Batyuk, A., Galli, L., Ishchenko, A., Han, G. W., Gati, C., Popov, P. A., Lee, M.-Y., Stauch, B., White, T. A., Barty, A., Aquila, A., Hunter, M. S., Liang, M., Boutet, S., Pu, M., Liu, Z. -j., Nelson, G., James, D., Li, C., … Cherezov, V. (2016). Native phasing of x-ray free-electron laser data for a G protein-coupled receptor. *Science Advances*, *2*(9), e1600292–e1600292. https://doi.org/10.1126/sciadv.1600292

Baxter, N. J., Roetzer, A., Liebig, H.-D., Sedelnikova, S. E., Hounslow, A. M., Skern, T., & Waltho, J. P. (2006). Structure and Dynamics of Coxsackievirus B4 2A Proteinase, an Enyzme Involved in the Etiology of Heart Disease [Article]. *Journal of Virology*, *80*(3),

1451–1462. https://doi.org/10.1128/jvi.80.3.1451-1462.2006

Berman, H., Henrick, K., & Nakamura, H. (2003). Announcing the worldwide protein Data Bank [Article]. *Nature Structural Biology*, *10*(12), 980–980. https://doi.org/10.1038/nsb1203-980

BERNAL, J. D., & CROWFOOT, D. (1934). X-Ray Photographs of Crystalline Pepsin. *Nature*, *133*(3369), 794–795. https://doi.org/10.1038/133794b0

Bernstein, F. C., Koetzle, T. F., Williams, G. J. ., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1978). The protein data bank: A computer-based archival file for macromolecular structures [Article]. *Archives of Biochemistry and Biophysics*, *185*(2), 584–591. https://doi.org/10.1016/0003-9861(78)90204-7

Beteva, A., Cipriani, F., Cusack, S., Delageniere, S., Gabadinho, J., Gordon, E. J., Guijarro, M., Hall, D. R., Larsen, S., Launer, L., Lavault, C. B., Leonard, G. A., Mairs, T., McCarthy, A., McCarthy, J., Meyer, J., Mitchell, E., Monaco, S., Nurizzo, D., … McSweeney, S. M. (2006). High-throughput sample handling and data collection at synchrotrons: embedding the ESRF into the high-throughput gene-to-structure pipeline [Article]. *Acta Crystallographica. Section D, Biological Crystallography.*, *62*(10), 1162–1169. https://doi.org/10.1107/S0907444906032859

Bingel-Erlenmeyer, R., Olieric, V., Grimshaw, J. P. A., Gabadinho, J., Wang, X., Ebner, S. G., Isenegger, A., Schneider, R., Schneider, J., Glettig, W., Pradervand, C., Panepucci, E. H., Tomizaki, T., Wang, M., & Schulze-Briese, C. (2011). SLS Crystallization Platform at Beamline X06DA□A Fully Automated Pipeline Enabling in Situ X-ray Diffraction Screening [Article]. *Crystal Growth & Design*, *11*(4), 916–923. https://doi.org/10.1021/cg101375j

Bublitz, M., Nass, K., Drachmann, N. D., Markvardsen, A. J., Gutmann, M. J., Barends, T. R. M., Mattle, D., Shoeman, R. L., Doak, R. B., Boutet, S., Messerschmidt, M., Seibert, M. M., Williams, G. J., Foucar, L., Reinhard, L., Sitsel, O., Gregersen, J. L., Clausen, J. D., Boesen, T., … Schlichting, I. (2015). Structural studies of P-type ATPase-ligand complexes using an X-ray free-electron laser [Article]. *IUCrJ*, *2*(Pt 4), 409–420. https://doi.org/10.1107/S2052252515008969

Caffrey, M. (2003). Membrane protein crystallization. *Journal of Structural Biology*, *142*(1), 108–132. https://doi.org/10.1016/s1047-8477(03)00043-1

Cai, Q., Yameen, M., Liu, W., Gao, Z., Li, Y., Peng, X., Cai, Y., Wu, C., Zheng, Q., Li, J., & Lin, T. (2013). Conformational Plasticity of the 2A Proteinase from Enterovirus 71 [Article]. *Journal of Virology*, *87*(13), 7348–7356. https://doi.org/10.1128/jvi.03541-12

Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter,

M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V, Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., … Spence, J. C. H. (2011). Femtosecond X-ray protein nanocrystallography [Article]. *Nature (London)*, *470*(7332), 73–78. https://doi.org/10.1038/nature09750

Chen, V. B., Davis, I. W., & Richardson, D. C. (2009). KING (Kinemage, Next Generation): A versatile interactive molecular and scientific visualization program [Article]. *Protein Science*, *18*(11), 2403–2409. https://doi.org/10.1002/pro.250

Cheng, T., Li, Q., Zhou, Z., Wang, Y., & Bryant, S. H. (2012). Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review [Article]. *The AAPS Journal*, *14*(1), 133–141. https://doi.org/10.1208/s12248-012-9322-0

Cherezov, V. (2012). Lipidic Cubic Phase Technologies for Membrane Protein Structural Studies [Article]. *Current Opinion in Structural Biology*, *21*(4), 559–566. https://doi.org/10.1016/j.sbi.2011.06.007.Lipidic

Cherezov, V., Hanson, M. A., Griffith, M. T., Hilgart, M. C., Sanishvili, R., Nagarajan, V., Stepanov, S., Fischetti, R. F., Kuhn, P., & Stevens, R. C. (2009). Rastering strategy for screening and centring of microcrystal samples of human membrane proteins with a sub-10 microm size X-ray synchrotron beam. [Article]. *Journal of the Royal Society, Interface / the Royal Society*, *6 Suppl 5*(June), S587-97. https://doi.org/10.1098/rsif.2009.0142.focus

Cipriani, F., Röwer, M., Landret, C., Zander, U., Felisaz, F., & Márquez, J. A. (2012). CrystalDirect: a new method for automated crystal harvesting based on laser-induced photoablation of thin films. *Acta Crystallographica. Section D, Biological Crystallography*, *68*(Pt 10), 1393–1399. https://doi.org/10.1107/S0907444912031459

Claesson, E., Wahlgren, W. Y., Takala, H., Pandey, S., Castillon, L., Kuznetsova, V., Henry, L., Panman, M., Carrillo, M., Kübel, J., Nanekar, R., Isaksson, L., Nimmrich, A., Cellini, A., Morozov, D., Maj, M., Kurttila, M., Bosman, R., Nango, E., … Westenhoff, S. (2020). The primary structural photoresponse of phytochrome proteins captured by a femtosecond x-ray laser [Article]. *ELife*, *9*. https://doi.org/10.7554/eLife.53514

Clark, G. L., & Corrigan, K. E. (1932). The Crystal Structure of Insulin. *Phys. Rev.*, *40*(4), 639. https://doi.org/10.1103/PhysRev.40.639

Cohen, A. E., Ellis, P. J., Miller, M. D., Deacon, A. M., & Phizackerley, R. P. (2002). An automated system to mount cryo-cooled protein crystals on a synchrotron beamline, using compact sample cassettes and a small-scale robot [Article]. *Journal of Applied Crystallography*, *35*(6), 720–726. https://doi.org/10.1107/S0021889802016709

Cui, S., Wang, J., Fan, T., Qin, B., Guo, L., Lei, X., Wang, J., Wang, M., & Jin, Q. (2011). Crystal Structure of Human Enterovirus 71 3C Protease [Article]. *Journal of*

Molecular Biology, *408*(3), 449–461. https://doi.org/10.1016/j.jmb.2011.03.007

de Graaf, C., Song, G., Cao, C., Zhao, Q., Wang, M. W., Wu, B., & Stevens, R. C. (2017). Extending the Structural View of Class B GPCRs. *Trends in Biochemical Sciences*, *42*(12), 946–960. https://doi.org/10.1016/j.tibs.2017.10.003

de la Mora, E., Coquelle, N., Bury, C. S., Rosenthal, M., Holton, J. M., Carmichael, I., Garman, E. F., Burghammer, M., Colletier, J.-P., & Weik, M. (2020). Radiation damage and dose limits in serial synchrotron crystallography at cryo- and room temperatures. *Proceedings of the National Academy of Sciences*, *117*(8), 4142 LP – 4151. https://doi.org/10.1073/pnas.1821522117

Deller, M. C., & Rupp, B. (2014). Approaches to automated protein crystal harvesting. *Acta Crystallographica. Section F, Structural Biology Communications*, *70*(Pt 2), 133–155. https://doi.org/10.1107/S2053230X14000387

DePonte, D. P., Weierstall, U., Schmidt, K., Warner, J., Starodub, D., Spence, J. C. H., & Doak, R. B. (2008). Gas dynamic virtual nozzle for generation of microscopic droplet streams [Article]. *Journal of Physics. D, Applied Physics*, *41*(19), 195505. https://doi.org/10.1088/0022-3727/41/19/195505

Dismukes, G. ., Klimov, V. ., Baranov, S. ., Kozlov, Y. ., DasGupta, J., & Tyryshkin, A. (2001). The Origin of Atmospheric Oxygen on Earth: The Innovation of Oxygenic Photosynthesis [Article]. *Proceedings of the National Academy of Sciences - PNAS*, *98*(5), 2170–2175. https://doi.org/10.1073/pnas.061514798

Dods, R., Båth, P., Arnlund, D., Beyerlein, K. R., Nelson, G., Liang, M., Harimoorthy, R., Berntsen, P., Malmerberg, E., Johansson, L., Andersson, R., Bosman, R., Carbajo, S., Claesson, E., Conrad, C. E., Dahl, P., Hammarin, G., Hunter, M. S., Li, C., … Neutze, R. (2017). From Macrocrystals to Microcrystals: A Strategy for Membrane Protein Serial Crystallography [Article]. *Structure (London)*, *25*(9), 1461-1468.e2. https://doi.org/10.1016/j.str.2017.07.002

Dods, R., Båth, P., Morozov, D., Gagnér, V. A., Arnlund, D., Luk, H. L., Kübel, J., Maj, M., Vallejos, A., Wickstrand, C., Bosman, R., Beyerlein, K. R., Nelson, G., Liang, M., Milathianaki, D., Robinson, J., Harimoorthy, R., Berntsen, P., Malmerberg, E., … Neutze, R. (2021). Ultrafast structural changes within a photosynthetic reaction centre [Article]. *Nature (London)*, *589*(7841), 310–314. https://doi.org/10.1038/s41586-020-3000-7

Evans, G., Axford, D., Waterman, D., & Owen, R. L. (2011). Macromolecular microcrystallography. *Crystallography Reviews*, *17*(2), 105–142. https://doi.org/10.1080/0889311X.2010.527964

Fang, Y. (2012). Ligand-receptor interaction platforms and their applications for drug

discovery. *Expert Opinion on Drug Discovery*, *7*(10), 969–988. https://doi.org/10.1517/17460441.2012.715631

Fenalti, G., Zatsepin, N. A., Betti, C., Giguere, P., Han, G. W., Ishchenko, A., Liu, W., Guillemyn, K., Zhang, H., James, D., Wang, D., Weierstall, U., Spence, J. C. ., Boutet, S., Messerschmidt, M., Williams, G. J., Gati, C., Yefanov, O. M., White, T. A., … Cherezov, V. (2015). Structural basis for bifunctional peptide recognition at human δ-opioid receptor [Article]. *Nature Structural & Molecular Biology*, *22*(3), 265–268. https://doi.org/10.1038/nsmb.2965

Feng, Q., Langereis, M. A., Lork, M., Nguyen, M., Hato, S. V, Lanke, K., Emdad, L., Bhoopathi, P., Fisher, P. B., Lloyd, R. E., & van Kuppeveld, F. J. M. (2014). Enterovirus 2Apro targets MDA5 and MAVS in infected cells [Article]. *Journal of Virology*, *88*(6), 3369.

Fraser, J. S., Van Den Bedem, H., Samelson, A. J., Lang, P. T., Holton, J. M., Echols, N., & Alber, T. (2011). Accessing Protein conformational ensembles using room-temperature X-ray crystallography [Article]. *Proceedings of the National Academy of Sciences - PNAS*, *108*(39), 16247–16252. https://doi.org/10.1073/pnas.1111325108

Fromme, R., Ishchenko, A., Metz, M., Chowdhury, S. R., Basu, S., Boutet, S., Fromme, P., White, T. A., Barty, A., Spence, J. C. H., Weierstall, U., Liu, W., & Cherezov, V. (2015). Serial femtosecond crystallography of soluble proteins in lipidic cubic phase [Article]. *IUCrJ*, *2*(Pt 5), 545–551. https://doi.org/10.1107/S2052252515013160

Garman, E. F. (2010). Radiation damage in macromolecular crystallography: what is it and why should we care? [Article]. *Acta Crystallographica. Section D, Biological Crystallography.*, *66*(4), 339–351. https://doi.org/10.1107/S0907444910008656

Garman, E. F., & Owen, R. L. (2006). Cryocooling and radiation damage in macromolecular crystallography [Article]. *Acta Crystallographica. Section D, Biological Crystallography.*, *62*(1), 32–47. https://doi.org/10.1107/S0907444905034207

Gelin, M., Delfosse, V., Allemand, F., Hoh, F., Sallaz-Damaz, Y., Pirocchi, M., Bourguet, W., Ferrer, J. L., Labesse, G., & Guichou, J. F. (2015). Combining “dry” co-crystallization and in situ diffraction to facilitate ligand screening by X-ray crystallography. *Acta Crystallographica. Section D, Biological Crystallography*, *71*(Pt 8), 1777–1787. https://doi.org/10.1107/S1399004715010342

Gisriel, C., Coe, J., Letrun, R., Yefanov, O. M., Luna-Chavez, C., Stander, N. E., Lisova, S., Mariani, V., Kuhn, M., Aplin, S., Grant, T. D., Dörner, K., Sato, T., Echelmeier, A., Cruz Villarreal, J., Hunter, M. S., Wiedorn, M. O., Knoska, J., Mazalova, V., … Zatsepin, N. A. (2019). Membrane protein megahertz crystallography at the European XFEL [Article]. *Nature Communications*, *10*(1), 5021–5021. https://doi.org/10.1038/s41467-019-12955-3

Gorel, A., Motomura, K., Fukuzawa, H., Doak, R. B., Grünbein, M. L., Hilpert, M., Inoue, I., Kloos, M., Kovácsová, G., Nango, E., Nass, K., Roome, C. M., Shoeman, R. L., Tanaka, R., Tono, K., Joti, Y., Yabashi, M., Iwata, S., Foucar, L., … Schlichting, I. (2017). Multi-wavelength anomalous diffraction de novo phasing using a two-colour X-ray free-electron laser with wide tunability [Article]. *Nature Communications*, *8*(1), 1170–1178. https://doi.org/10.1038/s41467-017-00754-7

Hendrickson, W. A. (n.d.). *Anomalous Diffraction in Crystallographic Phase Evaluation*. https://doi.org/10.1017/S0033583514000018

Himburg, H. A., Harris, J. R., Ito, T., Daher, P., Russell, J. L., Quarmyne, M., Doan, P. L., Helms, K., Nakamura, M., Fixsen, E., Herradon, G., Reya, T., Chao, N. J., Harroch, S., & Chute, J. P. (2012). Pleiotrophin Regulates the Retention and Self-Renewal of Hematopoietic Stem Cells in the Bone Marrow Vascular Niche [Article]. *Cell Reports (Cambridge)*, *2*(6), 1774–1774. https://doi.org/10.1016/j.celrep.2012.11.005

Hol, W. G. J. (1987). Protein crystallography, computer graphics and drug design [Article]. *Pure and Applied Chemistry*, *59*(3), 431–436. https://doi.org/10.1351/pac198759030431

Hynes, R. O. (1992). Integrins: Versatility, modulation, and signaling in cell adhesion [Article]. *Cell*, *69*(1), 11–25. https://doi.org/10.1016/0092-8674(92)90115-S

Ibrahim, M., Fransson, T., Chatterjee, R., Cheah, M. H., Hussein, R., Lassalle, L., Sutherlin, K. D., Young, I. D., Fuller, F. D., Gul, S., Kim, I.-S., Simon, P. S., de Lichtenberg, C., Chernev, P., Bogacz, I., Pham, C. C., Orville, A. M., Saichek, N., Northen, T., … Yano, J. (2020). Untangling the sequence of events during the S 2 -> S 3 transition in photosystem II and implications for the water oxidation mechanism [Document]. In *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 117, Issue 23). https://doi.org/10.1073/pnas.2000529117

Ihara, K., Hato, M., Nakane, T., Yamashita, K., Kimura-Someya, T., Hosaka, T., Ishizuka-Katsura, Y., Tanaka, R., Tanaka, T., Sugahara, M., Hirata, K., Yamamoto, M., Nureki, O., Tono, K., Nango, E., Iwata, S., & Shirouzu, M. (2020). Isoprenoid-chained lipid EROCOC17+4: a new matrix for membrane protein crystallization and a crystal delivery medium in serial femtosecond crystallography [Article]. *Scientific Reports*, *10*(1), 19305–19305. https://doi.org/10.1038/s41598-020-76277-x

Im, D., Inoue, A., Fujiwara, T., Nakane, T., Yamanaka, Y., Uemura, T., Mori, C., Shiimura, Y., Kimura, K. T., Asada, H., Nomura, N., Tanaka, T., Yamashita, A., Nango, E., Tono, K., Kadji, F. M. N., Aoki, J., Iwata, S., & Shimamura, T. (2020). Structure of the dopamine D2 receptor in complex with the antipsychotic drug spiperone [Article]. *Nature Communications*, *11*(1), 6442–6442. https://doi.org/10.1038/s41467-020-20221-0

Ishchenko, A., Cherezov, V., Liu, W., Ishchenko, A., & Cherezov, V. (2014). Preparation of microcrystals in lipidic cubic phase for serial femtosecond crystallography. *Nature Protocols*, *9*(9), 2123–2134. https://doi.org/10.1038/nprot.2014.141

Ishchenko, A., Stauch, B., Han, G. W., Batyuk, A., Shiriaeva, A., Li, C., Zatsepin, N., Weierstall, U., Liu, W., Nango, E., Nakane, T., Tanaka, R., Tono, K., Joti, Y., Iwata, S., Moraes, I., Gati, C., & Cherezov, V. (2019). Toward G protein-coupled receptor structure-based drug design using X-ray lasers [Article]. *IUCrJ*, *6*(Pt 6), 1106–1119. https://doi.org/10.1107/S2052252519013137

Ishigami, I., Lewis-Ballester, A., Echelmeier, A., Brehm, G., Zatsepin, N. A., Grant, T. D., Coe, J. D., Lisova, S., Nelson, G., Zhang, S., Dobson, Z. F., Boutet, S., Sierra, R. G., Batyuk, A., Fromme, P., Fromme, R., Spence, J. C. H., Ros, A., Yeh, S.-R., & Rousseau, D. L. (2019). Snapshot of an oxygen intermediate in the catalytic reaction of cytochrome c oxidase. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(9), 3572–3577. https://doi.org/10.1073/pnas.1814526116

Ishigami, I., Zatsepin, N. A., Hikita, M., Conrad, C. E., Nelson, G., Coe, J. D., Basu, S., Grant, T. D., Seaberg, M. H., Sierra, R. G., Hunter, M. S., Fromme, P., Fromme, R., Yeh, S.-R., & Rousseau, D. L. (2017). Crystal structure of CO-bound cytochrome c oxidase determined by serial femtosecond X-ray crystallography at room temperature. *Proceedings of the National Academy of Sciences*, *114*(30), 8011 LP – 8016. https://doi.org/10.1073/pnas.1705628114

Jacquamet, L., Ohana, J., Joly, J., Borel, F., Pirocchi, M., Charrault, P., Bertoni, A., Israel-Gouy, P., Carpentier, P., Kozielski, F., Blot, D., & Ferrer, J.-L. (2004). Automated Analysis of Vapor Diffusion Crystallization Drops with an X-Ray Beam [Article]. *Structure (London)*, *12*(7), 1219–1225. https://doi.org/10.1016/j.str.2004.04.019

Jaeger, K., Dworkowski, F., Nogly, P., Milne, C., Wang, M., & Standfuss, J. (2016). Serial Millisecond Crystallography of Membrane Proteins [Article]. *The Next Generation in Membrane Protein Structure Determination*, *922*, 137–149. https://doi.org/10.1007/978-3-319-35072-1_10

Johansson, L. C., Arnlund, D., White, T. A., Katona, G., DePonte, D. P., Weierstall, U., Doak, R. B., Shoeman, R. L., Lomb, L., Malmerberg, E., Davidsson, J., Nass, K., Liang, M., Andreasson, J., Aquila, A., Bajt, S., Barthelmess, M., Barty, A., Bogan, M. J., … Neutze, R. (2012). Lipidic phase membrane protein serial femtosecond crystallography [Article]. *Nature Methods*, *9*(3). https://doi.org/10.1038/nmeth.1867

Johansson, L. C., Stauch, B., McCorvy, J. D., Han, G. W., Patel, N., Huang, X.-P., Batyuk, A., Gati, C., Slocum, S. T., Li, C., Grandner, J. M., Hao, S., Olsen, R. H. J., Tribo, A. R., Zaare, S., Zhu, L., Zatsepin, N. A., Weierstall, U., Yous, S., … Cherezov, V. (2019). XFEL structures of the human MT2 melatonin receptor reveal the basis of subtype selectivity [Article]. *Nature (London)*, *569*(7755), 289–292.

https://doi.org/10.1038/s41586-019-1144-0

Juers, D. H., & Matthews, B. W. (2004). Cryo-cooling in macromolecular crystallography: advantages, disadvantages and optimization [Article]. *Quarterly Reviews of Biophysics*, *37*(2), 105–119. https://doi.org/10.1017/S0033583504004007

Kabsch, W. (2010). XDS. *Acta Crystallographica. Section D, Biological Crystallography*, *66*(Pt 2), 125–132. https://doi.org/10.1107/S0907444909047337

KAMATA, T., & TAKADA, Y. (1994). Direct binding of collagen to the I domain of integrin alpha 2 beta 1 (VLA-2, CD49b/CD29) in a divalent cation-independent manner [Article]. *The Journal of Biological Chemistry*, *269*(42), 26006–26010. https://doi.org/10.1016/S0021-9258(18)47151-7

Kang, Y., Zhou, X. E., Gao, X., He, Y., Liu, W., Ishchenko, A., Barty, A., White, T. A., Yefanov, O., Han, G. W., Xu, Q., de Waal, P. W., Ke, J., Tan, M. H. E., Zhang, C., Moeller, A., West, G. M., Pascal, B. D., Van Eps, N., … Xu, H. E. (2015). Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser [Article]. *Nature (London)*, *523*(7562).

Keedy, D. A., van den Bedem, H., Sivak, D. A., Petsko, G. A., Ringe, D., Wilson, M. A., & Fraser, J. S. (2014). Crystal Cryocooling Distorts Conformational Heterogeneity in a Model Michaelis Complex of DHFR [Article]. *Structure (London)*, *22*(6), 899–910. https://doi.org/10.1016/j.str.2014.04.016

Kern, J., Chatterjee, R., Young, I. D., Fuller, F. D., Lassalle, L., Ibrahim, M., Gul, S., Fransson, T., Brewster, A. S., Alonso-Mori, R., Hussein, R., Zhang, M., Douthit, L., de Lichtenberg, C., Cheah, M. H., Shevela, D., Wersig, J., Seuffert, I., Sokaras, D., … Yachandra, V. K. (2018). Structures of the intermediates of Kok's photosynthetic water oxidation clock [Article]. *Nature (London)*, *563*(7731), 421–425. https://doi.org/10.1038/s41586-018-0681-2

Kinashi, T. (2007). Integrin Regulation of Lymphocyte Trafficking: Lessons from Structural and Signaling Studies [Article]. *Advances in Immunology*, *93*, 185–227. https://doi.org/10.1016/S0065-2776(06)93005-3

Krissinel, E., Uski, V., Lebedev, A., Winn, M., & Ballard, C. (2018). Distributed computing for macromolecular crystallography. *Acta Crystallographica. Section D, Structural Biology*, *74*(Pt 2), 143–151. https://doi.org/10.1107/S2059798317014565

Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N. A., Rendek, K. N., Hunter, M. S., Shoeman, R. L., White, T. A., Wang, D., James, D., Yang, J.-H., Cobb, D. E., Reeder, B., Sierra, R. G., Liu, H., Barty, A., Aquila, A. L., Deponte, D., … Fromme, P. (2014). Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser [Article]. *Nature (London)*, *513*(7517). https://doi.org/10.1038/nature13453

Lavecchia, A, & Di Giovanni, C. (2013). Virtual screening strategies in drug discovery: a critical review. *Current Medicinal Chemistry*, *20*(23), 2839–2860. https://doi.org/10.2174/09298673113209990001

Lavecchia, Antonio, & Cerchia, C. (2016). In silico methods to address polypharmacology: current status, applications and future perspectives [Article]. *Drug Discovery Today*, *21*(2), 288–298. https://doi.org/10.1016/j.drudis.2015.12.007

le Maire, A., Gelin, M., Pochet, S., Hoh, F., Pirocchi, M., Guichou, J.-F., Ferrer, J.-L., & Labesse, G. (2011). In-plate protein crystallization, in situ ligand soaking and X-ray diffraction [Article]. *Acta Crystallographica. Section D, Biological Crystallography.*, *67*(9), 747–755. https://doi.org/10.1107/S0907444911023249

Lee, J. O., Rieu, P., Arnaout, M. A., & Liddington, R. (1995). Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18) [Article]. *Cell*, *80*(4), 631–638.

Lee, M.-Y., Geiger, J., Ishchenko, A., Han, G. W., Barty, A., White, T. A., Gati, C., Batyuk, A., Hunter, M. S., Aquila, A., Boutet, S., Weierstall, U., Cherezovb, V., & Liua, W. (2020). Harnessing the power of an X-ray laser for serial crystallography of membrane proteins crystallized in lipidic cubic phase [Article]. *IUCrJ*, *7*(Pt 6), 976–984. https://doi.org/10.1107/S2052252520012701

Lee, W., Watters, K. E., Troupis, A. T., Reinen, N. M., Suchy, F. P., Moyer, K. L., Frederick, R. O., Tonelli, M., Aceti, D. J., Palmenberg, A. C., & Markley, J. L. (2014). Solution structure of the 2A protease from a common cold agent, human rhinovirus C2, strain W12 [Article]. *PloS One*, *9*(6), e97198–e97198. https://doi.org/10.1371/journal.pone.0097198

Lei, X., Sun, Z., Liu, X., Jin, Q., He, B., & Wang, J. (2011). Cleavage of the Adaptor Protein TRIF by Enterovirus 71 3C Inhibits Antiviral Responses Mediated by Toll-Like Receptor 3 [Article]. *Journal of Virology*, *85*(17), 8811–8818. https://doi.org/10.1128/JVI.00447-11

Li, D., Stansfeld, P. J., Sansom, M. S. P., Keogh, A., Vogeley, L., Howe, N., Lyons, J. A., Aragao, D., Fromme, P., Fromme, R., Basu, S., Grotjohann, I., Kupitz, C., Rendek, K., Weierstall, U., Zatsepin, N. A., Cherezov, V., Liu, W., Bandaru, S., … Caffrey, M. (2015). Ternary structure reveals mechanism of a membrane diacylglycerol kinase [Article]. *Nature Communications*, *6*(1), 10140–10140. https://doi.org/10.1038/ncomms10140

Lieske, J., Cerv, M., Kreida, S., Komadina, D., Fischer, J., Barthelmess, M., Fischer, P., Pakendorf, T., Yefanov, O., Mariani, V., Seine, T., Ross, B. H., Crosas, E., Lorbeer, O., Burkhardt, A., Lane, T. J., Guenther, S., Bergtholdt, J., Schoen, S., … Meents, A. (2019).

On-chip crystallization for serial crystallography experiments and on-chip ligand-binding studies [Article]. *IUCrJ*, *6*(4), 714–728. https://doi.org/10.1107/S2052252519007395

Ling, H., Yang, P., Hou, H., & Sun, Y. (2018). Structural view of the 2A protease from human rhinovirus C15 [Article]. *Acta Crystallographica Section F: Structural Biology Communications*, *74*(4), 255–261. https://doi.org/10.1107/S2053230X18003382

Lionta, E., Spyrou, G., Vassilatis, D. K., & Cournia, Z. (2014). Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Current Topics in Medicinal Chemistry*, *14*(16), 1923–1938. https://doi.org/10.2174/1568026614666140929124445

Liu, Q., Dahmane, T., Zhang, Z., Assur, Z., Brasch, J., Shapiro, L., Mancia, F., & Hendrickson, W. A. (2012). Structures from Anomalous Diffraction of Native Biological Macromolecules [Article]. *Science (American Association for the Advancement of Science)*, *336*(6084), 1033–1037. https://doi.org/10.1126/science.1218753

Liu, W., Wacker, D., Gati, C., Han, G. W., James, D., Wang, D., Nelson, G., Weierstall, U., Katritch, V., Barty, A., Zatsepin, N. A., Li, D., Messerschmidt, M., Boutet, S., Williams, G. J., Koglin, J. E., Seibert, M. M., Wang, C., Shah, S. T. A., … Cherezov, V. (2013). Serial Femtosecond Crystallography of G Protein-Coupled Receptors [Article]. *Science (American Association for the Advancement of Science)*, *342*(6165), 1521–1524. https://doi.org/10.1126/science.1244142

Liu, W., Wacker, D., Wang, C., Abola, E., & Cherezov, V. (2014). Femtosecond crystallography of membrane proteins in the lipidic cubic phase. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1647), 20130314. https://doi.org/10.1098/rstb.2013.0314

Liu, W., Wacker, D., Wang, C., Abola, E., Cherezov, V., Liu, W., Wacker, D., Wang, C., Abola, E., & Cherezov, V. (2014). Femtosecond crystallography of membrane proteins in the lipidic cubic phase. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1647), 20130314. https://doi.org/10.1098/rstb.2013.0314

Luginina, A., Gusach, A., Marin, E., Mishin, A., Brouillette, R., Popov, P., Shiriaeva, A., Besserer-Offroy, É., Longpré, J.-M., Lyapina, E., Ishchenko, A., Patel, N., Polovinkin, V., Safronova, N., Bogorodskiy, A., Edelweiss, E., Hu, H., Weierstall, U., Liu, W., … Cherezov, V. (2019). Structure-based mechanism of cysteinyl leukotriene receptor inhibition by antiasthmatic drugs [Article]. *Science Advances*, *5*(10), eaax2518–eaax2518. https://doi.org/10.1126/sciadv.aax2518

Maeda, N., Nishiwaki, T., Shintani, T., Hamanaka, H., & Noda, M. (1996). 6B4 proteoglycan/phosphacan, an extracellular variant of receptor-like protein-tyrosine phosphatase zeta/RPTPbeta, binds pleiotrophin/heparin-binding growth-associated molecule (HB-GAM) [Article]. *The Journal of Biological Chemistry*, *271*(35), 21446–

21452.

Malcolm, B. A., Allaire, M., James, M. N. G., & Chernaia, M. M. (1994). Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases [Article]. *Nature (London)*, *369*(6475), 72–76. https://doi.org/10.1038/369072a0

Martin-Garcia, J. M., Conrad, C. E., Nelson, G., Stander, N., Zatsepin, N. A., Zook, J., Zhu, L., Geiger, J., Chun, E., Kissick, D., Hilgart, M. C., Ogata, C., Ishchenko, A., Nagaratnam, N., Roy-Chowdhury, S., Coe, J., Subramanian, G., Schaffer, A., James, D., … Liu, W. (2017). *Serial millisecond crystallography of membrane and soluble protein microcrystals using synchrotron radiation*. *4*, 439–454. https://doi.org/10.1107/S205225251700570X

Martin-Garcia, J. M., Zhu, L., Mendez, D., Lee, M.-Y., Chun, E., Li, C., Hu, H., Subramanian, G., Kissick, D., Ogata, C., Henning, R., Ishchenko, A., Dobson, Z., Zhang, S., Weierstall, U., Spence, J. C. H., Fromme, P., Zatsepin, N. A., Fischetti, R. F., … Liu, W. (2019). High-viscosity injector-based pink-beam serial crystallography of microcrystals at a synchrotron radiation source [Article]. *IUCrJ*, *6*(Pt 3), 412–425. https://doi.org/10.1107/S205225251900263X

Matthews, D. A., Smith, W. W., Ferre, R. A., Condon, B., Budahazi, G., Slsson, W., Villafranca, J. ., Janson, C. A., McElroy, H. ., Gribskov, C. ., & Worland, S. (1994). Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein [Article]. *Cell*, *77*(5), 761–771. https://doi.org/10.1016/0092-8674(94)90059-0

Maveyraud, L., & Mourey, L. (2020). Protein X-ray crystallography and drug discovery. *Molecules*, *25*(5). https://doi.org/10.3390/molecules25051030

Meents, A., Wiedorn, M. ., Srajer, V., Henning, R., Sarrou, I., Bergtholdt, J., Barthelmess, M., Reinke, P. Y. ., Dierksmeyer, D., Tolstikova, A., Schaible, S., Messerschmidt, M., Ogata, C. ., Kissick, D. ., Taft, M. ., Manstein, D. ., Lieske, J., Oberthuer, D., Fischetti, R. ., & Chapman, H. . (2017). Pink-beam serial crystallography [Article]. *Nature Communications*, *8*(1), 1281–12. https://doi.org/10.1038/s41467-017-01417-3

Meng, K., Rodriguez-Peña, A., Dimitrov, T., Chen, W., Yamin, M., Noda, M., & Deuel, T. F. (2000). Pleiotrophin Signals Increased Tyrosine Phosphorylation of β -Catenin through Inactivation of the Intrinsic Catalytic Activity of the Receptor-Type Protein Tyrosine Phosphatase β /ζ [Article]. *Proceedings of the National Academy of Sciences - PNAS*, *97*(6), 2603–2608. https://doi.org/10.1073/pnas.020487997

Messacar, K., Asturias, E. J., Hixon, A. M., Van Leer-Buter, C., Niesters, H. G. M., Tyler, K. L., Abzug, M. J., & Dominguez, S. R. (2018). Enterovirus D68 and acute flaccid myelitis—evaluating the evidence for causality [Article]. *The Lancet Infectious*

Diseases, *18*(8), e239–e247. https://doi.org/10.1016/S1473-3099(18)30094-X

MICHISHITA, M., VIDEM, & ARNAOUT, M. (1993). A NOVEL DIVALENT CATION-BINDING SITE IN THE A-DOMAIN OF THE BETA-2-INTEGRIN-CR3 (CD11B/CD18) IS ESSENTIAL FOR LIGAND-BINDING [Article]. *Cell*, *72*(6), 857–867. https://doi.org/10.1016/0092-8674(93)90575-B

Midgley, C. M., Watson, J. T., Nix, W. A., Curns, A. T., Rogers, S. L., Brown, B. A., Conover, C., Dominguez, S. R., Feikin, D. R., Gray, S., Hassan, F., Hoferka, S., Jackson, M. A., Johnson, D., Leshem, E., Miller, L., Nichols, J. B., Nyquist, A.-C., Obringer, E., … Gerber, S. I. (2015). Severe respiratory illness associated with a nationwide outbreak of enterovirus D68 in the USA (2014): a descriptive epidemiological investigation. *The Lancet. Respiratory Medicine*, *3*(11), 879–887. https://doi.org/10.1016/S2213-2600(15)00335-5

Misquitta, Y., Cherezov, V., Havas, F., Patterson, S., Mohan, J. M., Wells, A. J., Hart, D. J., & Caffrey, M. (2004). *Rational design of lipid for membrane protein crystallization* [Article]. *148*(2), 169–175. https://doi.org/10.1016/j.jsb.2004.06.008

Moffat, K. (1998). Ultrafast time-resolved crystallography [Article]. *Nature Structural Biology*, *5*(8), 641–643. https://doi.org/10.1038/1333

Moffat, Keith, Bilderback, D., Schildkamp, W., & Volz, K. (1986). Laue diffraction from biological samples [Article]. *Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment*, *246*(1), 627–635. https://doi.org/10.1016/0168-9002(86)90164-6

Moffat, Keith, Szebenyi, D., & Bilderback, D. (1984). X-ray Laue Diffraction from Protein Crystals [Article]. *Science (American Association for the Advancement of Science)*, *223*(4643), 1423–1425. https://doi.org/10.1126/science.223.4643.1423

Mosimann, S. C., Cherney, M. M., Sia, S., Plotch, S., & James, M. N. G. (1997). Refined X-ray Crystallographic Structure of the Poliovirus 3C Gene Product [Article]. *Journal of Molecular Biology*, *273*(5), 1032–1047. https://doi.org/10.1006/jmbi.1997.1306

Mould, A. P., Askari, J. A., Aota, S. i, Yamada, K. M., Irie, A., Takada, Y., Mardon, H. J., & Humphries, M. J. (1997). Defining the topology of integrin alpha5beta1-fibronectin interactions using inhibitory anti-alpha5 and anti-beta1 monoclonal antibodies. Evidence that the synergy sequence of fibronectin is recognized by the amino-terminal repeats of the alpha5 subunit [Article]. *The Journal of Biological Chemistry*, *272*(28), 17283–17292.

Muchmore, S. W., Olson, J., Jones, R., Pan, J., Blum, M., Greer, J., Merrick, S. M., Magdalinos, P., & Nienaber, V. L. (2000). Automated Crystal Mounting and Data Collection for Protein Crystallography [Article]. *Structure (London)*, *8*(12), R243–R246.

https://doi.org/10.1016/S0969-2126(00)00535-9

Musharrafieh, R., Zhang, J., Hu, Y., Diesing, J. M., Marty, M. T., & Wang, J. (2019). Validating Enterovirus D68-2Apro as an Antiviral Drug Target and the Discovery of Telaprevir as a Potent D68-2Apro Inhibitor. *J Virol*, *93*(7), 1–16.

Nakane, T., Hanashima, S., Suzuki, M., Saiki, H., Hayashi, T., Kakinouchi, K., Sugiyama, S., Kawatake, S., Matsuoka, S., Matsumori, N., Nango, E., Kobayashi, J., Shimamura, T., Kimura, K., Mori, C., Kunishima, N., Sugahara, M., Takakyu, Y., Inoue, S., … Mizohata, E. (2016). Membrane protein structure determination by SAD, SIR, or SIRAS phasing in serial femtosecond crystallography using an iododetergent. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(46), 13039–13044. https://doi.org/10.1073/pnas.1602531113

Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., … Iwata, S. (2016). A three-dimensional movie of structural changes in bacteriorhodopsin [Article]. *Science (American Association for the Advancement of Science)*, *354*(6319), 1552–1557. https://doi.org/10.1126/science.aah3497

Nass, K. (2019). Radiation damage in protein crystallography at X-ray free-electron lasers [Article]. *Acta Crystallographica. Section D, Structural Biology*, *75*(2), 211–218. https://doi.org/10.1107/S2059798319000317

Nass Kovacs, G., Colletier, J.-P., Grünbein, M. L., Yang, Y., Stensitzki, T., Batyuk, A., Carbajo, S., Doak, R. B., Ehrenberg, D., Foucar, L., Gasper, R., Gorel, A., Hilpert, M., Kloos, M., Koglin, J. E., Reinstein, J., Roome, C. M., Schlesinger, R., Seaberg, M., … Schlichting, I. (2019). Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin [Article]. *Nature Communications*, *10*(1), 3177–17. https://doi.org/10.1038/s41467-019-10758-0

Nogly, P., James, D., Wang, D., White, T. A., Zatsepin, N., Shilova, A., Nelson, G., Liu, H., Johansson, L., Heymann, M., Jaeger, K., Metz, M., Wickstrand, C., Wu, W., Båth, P., Berntsen, P., Oberthuer, D., Panneels, V., Cherezov, V., … Weierstall, U. (2015). Lipidic cubic phase serial millisecond crystallography using synchrotron radiation [Article]. *IUCrJ*, *2*(Pt 2), 168–176. https://doi.org/10.1107/S2052252514026487

Nogly, P., Weinert, T., James, D., Carbajo, S., Ozerov, D., Furrer, A., Gashi, D., Borin, V., Skopintsev, P., Jaeger, K., Nass, K., Båth, P., Bosman, R., Koglin, J., Seaberg, M., Lane, T., Kekilli, D., Brünle, S., Tanaka, T., … Standfuss, J. (2018). Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser [Article]. *Science (American Association for the Advancement of Science)*, *361*(6398), eaat0094. https://doi.org/10.1126/science.aat0094

Oberste, M. S., Maher, K., Schnurr, D., Flemister, M. R., Lovchik, J. C., Peters, H., Sessions, W., Kirk, C., Chatterjee, N., Fuller, S., Hanauer, J. M., & Pallansch, M. A. (2004). Enterovirus 68 is associated with respiratory illness and shares biological features with both the enteroviruses and the rhinoviruses [Article]. *Journal of General Virology*, *85*(9), 2577–2584. https://doi.org/10.1099/vir.0.79925-0

Ochiai, K., Muramatsu, H., Yamamoto, S., Ando, H., & Muramatsu, T. (2004). The role of midkine and pleiotrophin in liver regeneration [Article]. *Liver International*, *24*(5), 484–491. https://doi.org/10.1111/j.1478-3231.2004.0990.x

Opanda, S. M., Wamunyokoli, F., Khamadi, S., Coldren, R., & Bulimo, W. D. (2014). Genetic diversity of human enterovirus 68 strains isolated in Kenya using the hypervariable 3′- end of VP1 gene [Article]. *PloS One*, *9*(7), e102866–e102866. https://doi.org/10.1371/journal.pone.0102866

Pandey, S., Bean, R., Sato, T., Poudyal, I., Bielecki, J., Cruz Villarreal, J., Yefanov, O., Mariani, V., White, T. A., Kupitz, C., Hunter, M., Abdellatif, M. H., Bajt, S., Bondar, V., Echelmeier, A., Doppler, D., Emons, M., Frank, M., Fromme, R., … Schmidt, M. (2020). Time-resolved serial femtosecond crystallography at the European XFEL [Article]. *Nature Methods*, *17*(1), 73–78. https://doi.org/10.1038/s41592-019-0628-z

Pellegrini, C. (2012). The history of X-ray free-electron lasers [Article]. *European Physical Journal H*, *37*(5), 659–708. https://doi.org/10.1140/epjh/e2012-20064-5

Perez-Pinera, P., Berenson, J. R., & Deuel, T. F. (2008). Pleiotrophin, a multifunctional angiogenic factor: mechanisms and pathways in normal and pathological angiogenesis [Article]. *Current Opinion in Hematology*, *15*(3), 210–214. https://doi.org/10.1097/MOH.0b013e3282fdc69e

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera-A visualization system for exploratory research and analysis [Article]. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Poelman, R., Schuffenecker, I., Van Leer-Buter, C., Josset, L., Niesters, H. G. ., & Lina, B. (2015). European surveillance for enterovirus D68 during the emerging North-American outbreak in 2014 [Article]. *Journal of Clinical Virology*, *71*, 1–9. https://doi.org/10.1016/j.jcv.2015.07.296

Raulo, E., Chernousov, M. ., Carey, D. ., Nolo, R., & Rauvala, H. (1994). Isolation of a neuronal cell surface receptor of heparin binding growth-associated molecule (HB-GAM). Identification as N-syndecan (syndecan-3) [Article]. *The Journal of Biological Chemistry*, *269*(17), 12999–13004. https://doi.org/10.1016/S0021-9258(18)99975-8

Renaud, J. (2020). The Evolving Role of Structural Biology in Drug Discovery. *Structural Biology in Drug Discovery*, 1–22. https://doi.org/10.1002/9781118681121.ch1

Schieble, J. H., Fox, V. L., & Lennette, E. H. (1967). A probable new human picornavirus associated with respiratory diseases. *American Journal of Epidemiology*, *85*(2), 297–310. https://doi.org/10.1093/oxfordjournals.aje.a120693

Shen, D. (2020). *Functional and Structural Studies on Interactions of the Leukocyte Integrin αMβ2 with Cationic Ligands* [Dissertation]. Thesis (Ph.D.)--Arizona State University, 2020.

Skopintsev, P., Ehrenberg, D., Weinert, T., James, D., Kar, R. K., Johnson, P. J. M., Ozerov, D., Furrer, A., Martiel, I., Dworkowski, F., Nass, K., Knopp, G., Cirelli, C., Arrell, C., Gashi, D., Mous, S., Wranik, M., Gruhl, T., Kekilli, D., … Standfuss, J. (2020). Femtosecond-to-millisecond structural changes in a light-driven sodium pump [Article]. *Nature (London)*, *583*(7815), 314–318. https://doi.org/10.1038/s41586-020-2307-8

Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., & Bricogne, G. (2012). Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER [Article]. *Acta Crystallographica. Section D, Biological Crystallography.*, *68*(4), 368–380. https://doi.org/10.1107/S0907444911056058

Smith, J. L., Fischetti, R. F., & Yamamoto, M. (2012). Micro-crystallography comes of age [Article]. *Current Opinion in Structural Biology*, *22*(5), 602–612. https://doi.org/10.1016/j.sbi.2012.09.001

Song, C. M., Lim, S. J., & Tong, J. C. (2009). Recent advances in computer-aided drug design [Article]. *Briefings in Bioinformatics*, *10*(5), 579–591. https://doi.org/10.1093/bib/bbp023

Spence, J. C. H., Weierstall, U., & Chapman, H. N. (2012). X-ray lasers for structural and dynamic biology [Article]. *Reports on Progress in Physics*, *75*(10), 102601–102601. https://doi.org/10.1088/0034-4885/75/10/102601

Šrajer, V., Teng, T.-Y., Ursby, T., Pradervand, C., Ren, Z., Adachi, S.-I., Schildkamp, W., Bourgeois, D., Wulff, M., & Moffat, K. (1996). Photolysis of the Carbon Monoxide Complex of Myoglobin: Nanosecond Time- Resolved Crystallography [Article]. *Science (American Association for the Advancement of Science)*, *274*(5293), 1726–1729. https://doi.org/10.1126/science.274.5293.1726

Stauch, B., & Cherezov, V. (2018). Serial Femtosecond Crystallography of G Protein-Coupled Receptors [Article]. *Annual Review of Biophysics*, *47*(1), 377–397. https://doi.org/10.1146/annurev-biophys-070317-033239

Stauch, B., Johansson, L. C., McCorvy, J. D., Patel, N., Han, G. W., Huang, X.-P., Gati, C., Batyuk, A., Slocum, S. T., Ishchenko, A., Brehm, W., White, T. A., Michaelian, N., Madsen, C., Zhu, L., Grant, T. D., Grandner, J. M., Shiriaeva, A., Olsen, R. H. J., … Cherezov, V. (2019). Structural basis of ligand recognition at the human MT1 melatonin receptor [Article]. *Nature (London)*, *569*(7755), 284–288. https://doi.org/10.1038/s41586-019-1141-3

Stellato, F., Oberthür, D., Liang, M., Bean, R., Gati, C., Yefanov, O., Barty, A., Burkhardt, A., Fischer, P., Galli, L., Kirian, R. A., Meyer, J., Panneerselvam, S., Yoon, C. H., Chervinskii, F., Speller, E., White, T. A., Betzel, C., Meents, A., & Chapman, H. N. (2014). Room-temperature macromolecular serial crystallography using synchrotron radiation [Article]. *IUCrJ*, *1*(Pt 4), 204–212. https://doi.org/10.1107/S2052252514010070

Stempniak, M., Hostomska, Z., Nodes, B. R., & Hostomsky, Z. (1997). The NS3 proteinase domain of hepatitis C virus is a zinc-containing enzyme [Article]. *Journal of Virology*, *71*(4), 2881–2886. https://doi.org/10.1128/jvi.71.4.2881-2886.1997

STRANDBERG, B. E., DAVIES, D. R., SHORE, V. C., KENDREW, J. C., DICKERSON, R. E., HART, R. G., & PHILLIPS, D. C. (1960). Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 . Resolution [Article]. *Nature (London)*, *185*(4711), 422–427. https://doi.org/10.1038/185422a0

Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H., & Shen, J.-R. (2015). Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses [Article]. *Nature (London)*, *517*(7532), 99–103. https://doi.org/10.1038/nature13991

Suga, M., Akita, F., Sugahara, M., Kubo, M., Nakajima, Y., Nakane, T., Yamashita, K., Umena, Y., Nakabayashi, M., Yamane, T., Nakano, T., Suzuki, M., Masuda, T., Inoue, S., Kimura, T., Nomura, T., Yonekura, S., Yu, L.-J., Sakamoto, T., … Shen, J.-R. (2017). Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL [Article]. *Nature (London)*, *543*(7643), 131–135. https://doi.org/10.1038/nature21400

Suga, M., Akita, F., Yamashita, K., Nakajima, Y., Ueno, G., Li, H., Yamane, T., Hirata, K., Umena, Y., Yonekura, S., Yu, L.-J., Murakami, H., Nomura, T., Kimura, T., Kubo, M., Baba, S., Kumasaka, T., Tono, K., Yabashi, M., … Shen, J.-R. (2019). An oxyl/oxo mechanism for oxygen-oxygen coupling in PSII revealed by an x-ray free-electron laser [Article]. *Science (American Association for the Advancement of Science)*, *366*(6463), 334–338. https://doi.org/10.1126/science.aax6998

Sun, J., Hu, X.-Y. Y., & Yu, X.-F. F. (2019). Current understanding of human enterovirus D68 [Article]. *Viruses*, *11*(6), 490. https://doi.org/10.3390/v11060490

Sun, Y., Wang, X., Yuan, S., Dang, M., Li, X., Zhang, X. C., & Rao, Z. (2013). An open conformation determined by a structural switch for 2A protease from coxsackievirus A16 [Article]. *Protein and Cell*, *4*(10), 782–792. https://doi.org/10.1007/s13238-013-3914-z

Suno, R., Kimura, K. T., Nakane, T., Yamashita, K., Wang, J., Fujiwara, T., Yamanaka, Y., Im, D., Horita, S., Tsujimoto, H., Tawaramoto, M. S., Hirokawa, T., Nango, E., Tono, K., Kameshima, T., Hatsui, T., Joti, Y., Yabashi, M., Shimamoto, K., … Kobayashi, T. (2018). Crystal Structures of Human Orexin 2 Receptor Bound to the Subtype-Selective Antagonist EMPA [Article]. *Structure (London)*, *26*(1), 7-19.e5. https://doi.org/10.1016/j.str.2017.11.005

Tan, J., George, S., Kusov, Y., Perbandt, M., Anemüller, S., Mesters, J. R., Norder, H., Coutard, B., Lacroix, C., Leyssen, P., Neyts, J., & Hilgenfeld, R. (2013). 3C Protease of Enterovirus 68: Structure-Based Design of Michael Acceptor Inhibitors and Their Broad-Spectrum Antiviral Effects against Picornaviruses [Article]. *Journal of Virology*, *87*(8), 4339–4351. https://doi.org/10.1128/JVI.01123-12

Thomaston, J. L., Woldeyes, R. A., Nakane, T., Yamashita, A., Tanaka, T., Koiwai, K., Brewster, A. S., Barad, B. A., Chen, Y., Lemmin, T., Uervirojnangkoorn, M., Arima, T., Kobayashi, J., Masuda, T., Suzuki, M., Sugahara, M., Sauter, N. K., Tanaka, R., Nureki, O., … DeGrado, W. F. (2017). XFEL structures of the influenza M2 proton channel [Article]. *Proceedings of the National Academy of Sciences - PNAS*, *114*(51), 13357–13362. https://doi.org/10.1073/pnas.1705624114

Torii, H., Sueki, H., Kumada, H., Sakurai, Y., Aoki, K., Yamada, I., & OHtsuki, M. (2013). Dermatological side-effects of telaprevir-based triple therapy for chronic hepatitis C in phase III trials in Japan [Article]. *Journal of Dermatology*, *40*(8), 587–595. https://doi.org/10.1111/1346-8138.12199

Tosha, T., Nomura, T., Nishida, T., Saeki, N., Okubayashi, K., Yamagiwa, R., Sugahara, M., Nakane, T., Yamashita, K., Hirata, K., Ueno, G., Kimura, T., Hisano, T., Muramoto, K., Sawai, H., Takeda, H., Mizohata, E., Yamashita, A., Kanematsu, Y., … Kubo, M. (2017). Capturing an initial intermediate during the P450nor enzymatic reaction using time-resolved XFEL crystallography and caged-substrate [Article]. *Nature Communications*, *8*(1), 1585–1589. https://doi.org/10.1038/s41467-017-01702-1

TUCKWELL, D., CALDERWOOD, D., GREEN, L., & HUMPHRIES, M. (1995). INTEGRIN ALPHA-2 I-DOMAIN IS A BINDING-SITE FOR COLLAGENS [Article]. *Journal of Cell Science*, *108*, 1629–1637.

Ulferts, R., de Boer, S. M., van der Linden, L., Bauer, L., Lyoo, H. R., Maté, M. J., Lichière, J., Canard, B., Lelieveld, D., Omta, W., Egan, D., Coutard, B., & van Kuppeveld, F. J. M. (2016). Screening of a Library of FDA-Approved Drugs Identifies Several Enterovirus Replication Inhibitors That Target Viral Protein 2C [Article].

Antimicrobial Agents and Chemotherapy, *60*(5), 2627–2638. https://doi.org/10.1128/AAC.02182-15

Visser, L. J., Langereis, M. A., Rabouw, H. H., Wahedi, M., Muntjewerff, E. M., de Groot, R. J., & van Kuppeveld, F. J. M. (2019). Essential role of enterovirus 2A protease in counteracting stress granule formation and the induction of type I interferon [Article]. *Journal of Virology*, *93*(10). https://doi.org/10.1128/JVI.00222-19

Wacker, D., Wang, C., Katritch, V., Han, G. W., Huang, X.-P., Vardy, E., McCorvy, J. D., Jiang, Y., Chu, M., Siu, F. Y., Liu, W., Xu, H. E., Cherezov, V., Roth, B. L., & Stevens, R. C. (2013). Structural Features for Functional Selectivity at Serotonin Receptors [Article]. *Science (American Association for the Advancement of Science)*, *340*(6132), 615–619. https://doi.org/10.1126/science.1232808

Wadsten, P., Wöhri, A. B., Snijder, A., Katona, G., Gardiner, A. T., Cogdell, R. J., Neutze, R., & Engström, S. (2006). Lipidic Sponge Phase Crystallization of Membrane Proteins [Article]. *Journal of Molecular Biology*, *364*(1), 44–53. https://doi.org/10.1016/j.jmb.2006.06.043

Wang, C., Wu, H., Katritch, V., Han, G. W., Huang, X.-P., Liu, W., Siu, F. Y., Roth, B. L., Cherezov, V., & Stevens, R. C. (2013). Structure of the human smoothened receptor bound to an antitumour agent [Article]. *Nature (London)*, *497*(7449), 338–343. https://doi.org/10.1038/nature12167

Weierstall, U., James, D., Wang, C., White, T. A., Wang, D., Liu, W., Spence, J. C., Bruce Doak, R., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N. A., Liu, H., Basu, S., Wacker, D., Won Han, G., Katritch, V., … Cherezov, V. (2014). Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nat Commun*, *5*, 3309. https://doi.org/10.1038/ncomms4309

Westenhoff, S., Nazarenko, E., Malmerberg, E., Davidsson, J., Katona, G., & Neutze, R. (2010). Time-resolved structural studies of protein reaction dynamics: a smorgasbord of X-ray approaches. *Acta Crystallographica. Section A, Foundations of Crystallography*, *66*(Pt 2), 207–219. https://doi.org/10.1107/S0108767309054361

Woodhouse, J., Nass Kovacs, G., Coquelle, N., Uriarte, L. M., Adam, V., Barends, T. R. M., Byrdin, M., de la Mora, E., Bruce Doak, R., Feliks, M., Field, M., Fieschi, F., Guillon, V., Jakobs, S., Joti, Y., Macheboeuf, P., Motomura, K., Nass, K., Owada, S., … Weik, M. (2020). Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy [Article]. *Nature Communications*, *11*(1), 741–741. https://doi.org/10.1038/s41467-020-14537-0

Xu, D., Shen, B., Xu, J., & Liang, Z. (2020). XFEL beamline design for vacuum birefringence experiment [Article]. *Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment*, *982*,

164553. https://doi.org/10.1016/j.nima.2020.164553

Yanmei Hu, Rami Musharrafieh, Madeleine Zheng, J. W. (2020). Enterovirus D68 antivirals: past, present and future. *Physiology & Behavior*, *176*(3), 139–148. https://doi.org/10.1021/acsinfecdis.0c00120.Enterovirus

Yin, X., Scalia, A., Leroy, L., Cuttitta, C. M., Polizzo, G. M., Ericson, D. L., Roessler, C. G., Campos, O., Ma, M. Y., Agarwal, R., Jackimowicz, R., Allaire, M., Orville, A. M., Sweet, R. M., & Soares, A. S. (2014). Hitting the target: fragment screening with acoustic in situ co-crystallization of proteins plus fragment libraries on pin-mounted data-collection micromeshes. *Acta Crystallographica. Section D, Biological Crystallography*, *70*(Pt 5), 1177–1189. https://doi.org/10.1107/S1399004713034603

Yokoi, H., Kasahara, M., Mori, K., Ogawa, Y., Kuwabara, T., Imamaki, H., Kawanishi, T., Koga, K., Ishii, A., Kato, Y., Mori, K. P., Toda, N., Ohno, S., Muramatsu, H., Muramatsu, T., Sugawara, A., Mukoyama, M., & Nakao, K. (2012). Pleiotrophin triggers inflammation and increased peritoneal permeability leading to peritoneal fibrosis [Article]. *Kidney International*, *81*(2), 160–169. https://doi.org/10.1038/ki.2011.305

Young, I. D., Ibrahim, M., Chatterjee, R., Gul, S., Fuller, F. D., Koroidov, S., Brewster, A. S., Tran, R., Alonso-Mori, R., Kroll, T., Michels-Clark, T., Laksmono, H., Sierra, R. G., Stan, C. A., Hussein, R., Zhang, M., Douthit, L., Kubin, M., De Lichtenberg, C., … Yano, J. (2016). Structure of photosystem II and substrate binding at room temperature [Article]. *Nature (London)*, *540*(7633), 453–457. https://doi.org/10.1038/nature20161

Zhang, Haitao, Han, G. W., Batyuk, A., Ishchenko, A., White, K. L., Patel, N., Sadybekov, A., Zamlynny, B., Rudd, M. T., Hollenstein, K., Tolstikova, A., White, T. A., Hunter, M. S., Weierstall, U., Liu, W., Babaoglu, K., Moore, E. L., Katz, R. D., Shipman, J. M., … Cherezov, V. (2017). Structural basis for selectivity and diversity in angiotensin II receptors [Article]. *Nature (London)*, *544*(7650), 327–332. https://doi.org/10.1038/nature22035

Zhang, Haitao, Unal, H., Gati, C., Han, G. W., Liu, W., Zatsepin, N. A., James, D., Wang, D., Nelson, G., Weierstall, U., Sawaya, M. R., Xu, Q., Messerschmidt, M., Williams, G. J., Boutet, S., Yefanov, O. M., White, T. A., Wang, C., Ishchenko, A., … Cherezov, V. (2015). Structure of the Angiotensin Receptor Revealed by Serial Femtosecond Crystallography [Article]. *Cell*, *161*(4).

Zhang, Haonan, Qiao, A., Yang, D., Yang, L., Dai, A., de Graaf, C., Reedtz-Runge, S., Dharmarajan, V., Zhang, H., Han, G. W., Grant, T. D., Sierra, R. G., Weierstall, U., Nelson, G., Liu, W., Wu, Y., Ma, L., Cai, X., Lin, G., … Wu, B. (2017). Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature*, *546*(7657), 259–264. https://doi.org/10.1038/nature22363

Zhang, X., Zhao, F., Wu, Y., Yang, J., Han, G. W., Zhao, S., Ishchenko, A., Ye, L., Lin,

X., Ding, K., Dharmarajan, V., Griffin, P. R., Gati, C., Nelson, G., Hunter, M. S., Hanson, M. A., Cherezov, V., Stevens, R. C., Tan, W., … Xu, F. (2017). Crystal structure of a multi-domain human smoothened receptor in complex with a super stabilizing ligand. *Nature Communications*, *8*(May), 15383. https://doi.org/10.1038/ncomms15383

Zhu, L., Chen, X., Abola, E. E., Jing, L., & Liu, W. (2020). Serial Crystallography for Structure-Based Drug Discovery [Article]. *Trends in Pharmacological Sciences (Regular Ed.)*, *41*(11), 830–839. https://doi.org/10.1016/j.tips.2020.08.009

APPENDIX A

[REFERENCFE LIST OF TABLE 2.2]

Year	Protein	PDB ID	Reference
2013	5-HT_{2B}	4NC3	(W. Liu et al., 2013)
2014	δ-OR	4RWD	(Fenalti et al., 2015)
2014	SMO	4O9R	(Weierstall et al., 2014)
2015	AT_1R	4YAY	(Haitao Zhang et al., 2015)
2015	rhodopsin	4ZWJ	(Kang et al., 2015)
2016	OX_2R	5WS3	(Suno et al., 2018)
2017	GCGR	5XEZ	(Haonan Zhang et al., 2017)
2017	AT_2R	5UNG, 5UNF	(Haitao Zhang et al., 2017)
2017	SMO	5V56,	(X. Zhang et al., 2017)
2018	MT_2	6ME7, 6ME6, 6ME9, 6ME8	(Johansson et al., 2019)
2018	MT1	6ME3, 6ME2, 6ME5, 6ME4	(Stauch et al., 2019)
2018	EP3	6M9T	(Audet et al., 2019)
2019	MT1	6PS8	(Ishchenko et al., 2019)
2019	$CysLT_1R$	6RZ5	(Luginina et al., 2019)
2019	β_2AR	6PS1, 6PS4, 6PS3, 6PS6, 6PS0, 6PRZ, 6PS5	(Ishchenko et al., 2019)
2019	$A_{2A}AR$	6PS7	(Ishchenko et al., 2019)
2020	D_2R	7DFP	(Im et al., 2020)

Year	Protein	PDB ID	Reference
2014	PS II	4PBU	(Kupitz et al., 2014)
2016	PS II	5TIS, 5KAF, 5TIS	(Young et al., 2016)
2016	PS II	5KAI	(Young et al., 2016)
2016	PS II	5GTH, 5WS6, 5WS5, 5GTI	(Suga et al., 2017)
2016	bR	5B6Z, 5B6Y, 5B6V, 5B6X, 5B6W, 5H2P, 5H2O, 5H2H, 5H2J, 5H2I, 5H2L, 5H2K, 5H2N, 5H2M	(Nango et al., 2016)
2018	bR	6GAH, 6GAA, 6GAC, 6GAB, 6GAE, 6GAD, 6GAG, 6GAF, 6GA9, 6GA8, 6GA1, 6GA3, 6GA2, 6GA5, 6GA4, 6GA7, 6GA6, 6RMK	(Nass Kovacs et al., 2019)
2018	bR	6G7I, 6G7H, 6G7K, 6G7J, 6G7L	(Nogly et al., 2018)
2019	KR2	6TK6, 6TK5, 6TK7, 6TK2, 6TK1, 6TK4, 6TK3	(Skopintsev et al., 2020)
2020	PS II	6W1O, 6W1Q, 6W1P, 6W1R, 6W1U, 6W1T, 6W1V	(Ibrahim et al., 2020)
2020	*Bv*RC	6ZIA, 6ZID, 6ZI6, 6ZI5, 6ZI9, 6ZI4	(Dods et al., 2021)

Year	Protein	PDB ID	Reference
2014	DgkA	4UYO	(Li et al., 2015)
2017	bCcO	5W97	(Ishigami et al., 2017)
2019	bCcO	6NKN, 6NMP, 6NMF	(Ishigami et al., 2019)
2017	M_2	5TTC	(Thomaston et al., 2017)
2017	M_3	5JOO	(Thomaston et al., 2017)
2017	M_4	5UM1	(Thomaston et al., 2017)
2014	PS II	4UB6, 4UB8	(Suga et al., 2015)
2019	PS II	6JLK, 6JLJ, 6JLM, 6JLL, 6JLP, 6JLO, 6JLN	(Suga et al., 2019)
2015	PS II	5E7C, 5E79	(Ayyer et al., 2016)
2015	SR Ca^{2+}-ATPase	4XOU	(Bublitz et al., 2015)
2016	bR	5B34, 5B35	(Nakane et al., 2016)
2017	RC_{vir}	5NJ4, 5O4C	(Dods et al., 2017)
2019	hAQP2	6QF5	(Lieske et al., 2019)
2019	PS I	6PGK	(Gisriel et al., 2019)
2020	$A_{2A}AR$	6WQA	(M.-Y. Lee et al., 2020)
2020	$A_{2A}AR$	6LPK, 6LPJ	(Ihara et al., 2020)

APPENDIX B

[PERMISSION OF TABLE 1.1]

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Oct 11, 2021

This Agreement between Miss. Chang Liu ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	5166170699064
License date	Oct 11, 2021
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Wiley Books
Licensed Content Title	The Evolving Role of Structural Biology in Drug Discovery
Licensed Content Author	Jean-Paul Renaud
Licensed Content Date	Jan 10, 2020
Licensed Content Pages	21
Type of use	Dissertation/Thesis
Requestor type	University/Academic

APPENDIX C

[PERMISSION OF FIGURE 1.2]

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Oct 11, 2021

This Agreement between Miss. Chang Liu ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	5166170699064
License date	Oct 11, 2021
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Wiley Books
Licensed Content Title	The Evolving Role of Structural Biology in Drug Discovery
Licensed Content Author	Jean−Paul Renaud
Licensed Content Date	Jan 10, 2020
Licensed Content Pages	21
Type of use	Dissertation/Thesis
Requestor type	University/Academic

Format	Print and electronic
Portion	Figure/table
Number of figures/tables	3
Will you be translating?	No
Title	Structural Based Drug Discovery: The Significance of Protein Structure
Institution name	Arizona State University
Expected presentation date	Oct 2021
Portions	Table 1.1, Table 1.2, Figure 1.1
Requestor Location	Miss. Chang Liu 2396 w ivanhoe st CHANDLER, AZ 85224 United States Attn: Miss. Chang Liu
Publisher Tax ID	EU826007151