Structural Studies of Protein Complexes Using Single-particle Cryo-electron Microscopy

by

Chloe Du Truong

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

Approved October 2021 by the
Graduate Supervisory Committee:

Po-Lin Chiu, Chair
Wei Liu
Yuval Mazor

ARIZONA STATE UNIVERSITY

December 2021

ABSTRACT

Cryogenic Electron Microscopy (Cryo-EM) is a method that can be used for studying the structure of biological systems. Biological samples are frozen to cryogenic temperatures and embedded in a vitreous ice when they are imaged by electrons. Due to its ability to preserve biological specimens in near-native conditions, cryo-EM has a significant contribution to the field of structural biology.

Single-particle cryo-EM technique was utilized to investigate the dynamical characteristics of various protein complexes such as the Nogo receptor complex, polymerase ζ (Polζ) in yeast and human integrin αvβ8-pro-TGFβ1-GARP complex. Furthermore, I proposed a new method that can potentially improve the sample preparation for cryo-EM. The Nogo receptor complex was expressed using baculovirus expression system in *sf9* insect cells and isolated for structural studies. Nogo receptor complex was found to have various stoichiometries and interactions between individual proteins. A structural investigation of the yeast apo polymerase ζ holoenzyme was also carried out. The apo Polζ displays a concerted motions associated with expansion of the Polζ DNA-binding channel upon DNA binding. Furthermore, a lysine residue that obstructs the DNA-binding channel in apo Polζ was found and suggested a gating mechanism. In addition, cryo-EM studies of the human integrin αvβ8-pro-TGFβ1-GARP complex was conducted to assess its dynamic interactions. The 2D classifications showed the αvβ8-pro-TGFβ1-GARP complex is highly flexible and required several sample preparation techniques such as crosslinking and graphene oxide coating to improve protein homogeneity on the EM grid. To overcome challenges within the cryo-EM technique such as particle adsorption on air-water interface, I have documented a collaborative work on the development and

application of lipid monolayer sandwich on cryo-EM grid. Cryogenic electron tomography (cryo-ET) along with cryo-EM were used to study the characteristics of lipid monolayer sandwich as a potential protective layer for EM grid. The cryo-ET results demonstrated that the thickness of lipid monolayer is adequate for single-particle cryo-EM processing. Furthermore, there was no appearance of preferred orientations in cryo-EM and cryo-ET images. To establish that this method is actually beneficial, more data must be collected, and high-resolution structures of protein samples must be obtained using this methodology.

To grandma (1931-2021)

ACKNOWLEDGMENTS

This dissertation would not have been possible without the support and guidance that I received from many people.

First and foremost, I would like to express my sincere gratitude to my advisor Dr. Po-Lin Chiu for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. Po-Lin is someone you will instantly like and never forget once you meet him. He is kind, open, honest, funny, and is one of the smartest people I know. Through every experiment I conducted, his insightful feedback pushed me to sharpen my critical thinking and brought my work to a higher level. I truly admire his dedication and relentless love for science; and I hope to become half the scientist he is someday. I would also like to thank the rest of my thesis committee: Dr. Wei Liu, and Dr. Yuval Mazor for many great insights, which incented me to widen my research from various perspectives. Likewise, I would like to thank our cryo-EM facility manager, Dr. Dewight Williams for his technical support and for teaching me everything he knows about microscopy. Yet, I still have many things to learn from him.

Most of my research would not be nearly as exciting as it was without the help of my collaborators. I would like to particularly acknowledge Dr. Albert Lin, Dr. Eugene Nebelitsky, Dr. Nadia Leloup at Morphic Therapeutics for their support on the integrin project. I would also like to thank Dr. Rajiv Kumar, Dr. Georges Mer, Dr. Theodore Craig, and Rachel Serkasevich at Mayo Clinic for their work on polymerase ζ. I am also thankful for all the guidance that Dr. Joseph Wang at Pennsylvania State University has kindly provided me on the lipid monolayer sandwich project.

I would like to thank my fellow lab mates in for all the fruitful discussions, the long hours we were working together for an experiment, and for all the fun boba runs. Special thanks to Purbasha Nandi for being an amazing friend ever since our lab started, to Kayi Chan and Kazi Waheeda for helping me out in the lab, to Ryan Puskar for all the interesting conversations we had. Additionally, I would like to thank Dr. Guanhong Bu and Kyle Swain in Dr. Brent Nannenga's lab for being a team player and always giving great advice to all my problems, whether big or small. You are a bundle of friendly, fun, and hardworking people that I have the honor to get to know and work with. I am truly grateful.

A good support system is important to surviving and staying sane in graduate school. I was lucky to be a part of one that is called "CX buddies," which refer to Dr. Zina Al-Sahouri, Dr. Michael Moran, and myself. They were the ones who dragged me away from thinking about work when I over-stressed myself, whether by playing Stardew Valley or going on a random trip to a furniture store. You guys are my lifelong friends, and I thank you.

I would like to thank all my friends and family: my parents and my little sister for supporting me emotionally throughout my Ph.D. and my life in general. To Vy, although sometimes I feel like I still have to babysit you, you are my forever best friend and the first person I go to for anything. Your maturity triumphs the fact that you are my little sister, for that I have an immense respect for you. I am extremely proud of my little feminist.

The best outcome from these past four and a half years is finding my best friend, travel buddy and partner-in-crime, that is Dr. Thi Ngo. I am very lucky to have met someone who understands my bad day and turn it around with laughter, someone who truly believes in me and always cheers me on. You are my shoulder to lean on, my love, my rock, and I cannot wait to be building a life together with you.

TABLE OF CONTENTS

Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1: INTRODUCTION

1.1 Single-particle Cryo-Electron Microscopy

Cryogenic Electron Microscopy (cryo-EM) is a powerful tool used to study the structures of biological macromolecules. Cryo-EM uses electron to image vitrified biological samples that are close to their native-like conditions[1]. Cryogenic environments are characterized by temperatures below -150°C. Vitreous ice is characterized by its amorphous structure and formed when an object is quickly plunge frozen into a cryogen with a high heat capacity, such as liquid ethane[10]. Proteins embedded in vitreous ice are then imaged under a transmission electron microscopy, and cryo-EM data is collected. Through image processing, the images of the particles can be extracted from micrographs, then averaged into two-dimensional (2D) class averages. A 3D reconstruction map is generated and further refined computationally. A final high-resolution density map of the target molecule can then be reconstructed. Single-particle Cryo-EM has significantly contributed to the field of structural biology due to its applications in the preservation of biological specimens in their near-native conditions.

Recent advances in detector technology and software algorithms have enabled the discoveries of biomolecular structure determination at near atomic resolution[2]. These advances came from the development of direct electron detector device (DDD) cameras[3,4,5]. DDD cameras has a higher detective quantum efficiency (DQE) than any other conventional recording media[6]. A complementary metal-oxide semiconductor (CMOS) technology also makes the technique more prominent to collect dose-fractionated image stacks, typically referred to as movies. Cryo-EM movies undergo a computational

correction of specimen movements, so-called motion correction, which corrects the beam-induced movements of the sample[4,6,7]. These improvements have enabled the reconstruction of a high-quality density map from the raw electron movies. In parallel, increases in computing power enable faster processing speed and allow processing a large dataset for new algorithms to generate more reliable 3D density maps[8,9].

1.2 Specimen preparation for single-particle cryo-EM

Single-particle cryo-EM experiments start with a purified biological specimen, such as proteins, in aqueous solution at a typical concentration between 50 nM and 5 μM[11]. The protein solution is applied to an EM grid, which consists of a holey carbon film supported by a metal frame made of copper or gold. Excess buffer from the sample application step is then blotted away with filter paper, and the grid is immediately plunge frozen in liquid ethane at -180°C, then subsequently cooled to -196°C with liquid nitrogen[10]. Vitreous ice can be formed due to liquid ethane having a higher heat capacity and boiling point (-188 *vs.* -196°C) than liquid nitrogen, which minimizes boiling and enables rapid freezing of the sample[12].

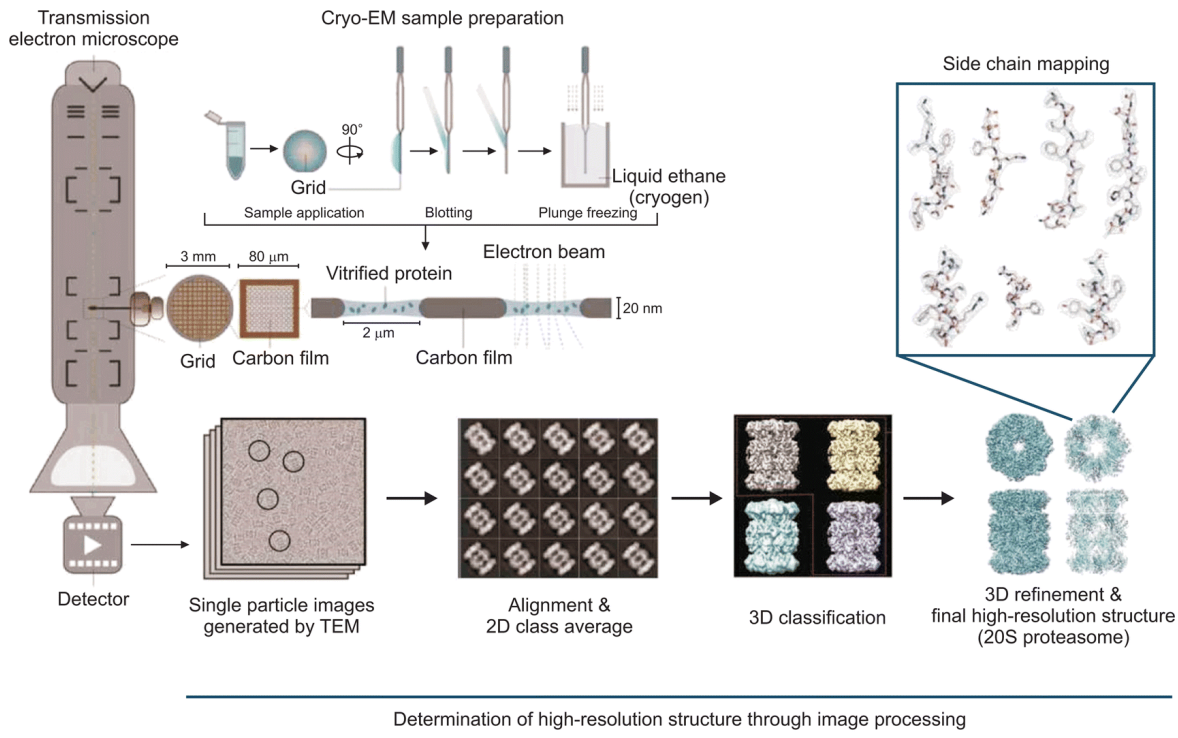

Figure 1.1. General scheme of single-particle cryo-EM workflow. A purified protein sample is applied to a cryo-EM grid and plunge frozen in liquid ethane, then transferred to liquid nitrogen. Particles embedded in vitreous ice are imaged using a cryo-transmission electron microscope. From these micrographs, individual particles are then selected and aligned to calculate two-dimensional class average. Three-dimensional classification and further refinement subsequently calculate the high-resolution cryo-EM structure[10].

To ensure the formation of vitreous ice, the sample must be frozen quickly at a rate of approximately 106°C per second, so that the water in and around the specimen is immobilized in a vitreous state[13]. If freezing occurs too slowly, the specimen is at the risk forming crystalline ice, since water devitrifies at -137°C [13]. This rapid freezing step allows the biological molecules to be preserved in their near-native state without dehydration or ice crystallization. Ice contamination can cause damage to the structural integrity of the specimen, due to crystal formation removing water molecules from the hydration shells of the specimen[14]. Formation of crystalline ice also worsens image quality, as these crystals diffract electrons and dominate the imaged area. Other types of contamination such as

hexagonal ice can also occur during the freezing process, which can be reduced by working in a humidity-controlled environment and minimizing ice contamination in the liquid nitrogen. Furthermore, plunge freezing a specimen can localize the effects of radiation damage sustained in the electron microscope.

An amorphous ice layer of sufficient thickness allows particles to be visualized better under EM. Particles that are both evenly distributed across the field of view and adopt a wide range of orientations are most desirable. To achieve optimal conditions, several parameters need to be adjusted according to the characteristics of a specific protein. A thin ice layer allows for higher contrast between the molecule and its background by minimizing defocus spread, which is caused by different heights of the molecules in the ice layer and can hinder high-resolution structure determination. Two key parameters that can affect ice thickness are blotting time and sample support structure, and both can be optimized based on the protein characteristics to achieve the best results. Blotting time can be precisely controlled using a commercially available plunger such as the Vitrobot (FEI), which enables users to automate and fine tune their plunge freezing conditions. Likewise, the type of support used can be adjusted during the specimen preparation step by selecting grids with the appropriate structural parameters. The EM grids are typically copper, coated in either a continuous or holey carbon film for use in negative-stain EM or cryo-EM, respectively. Holey carbon grids can be either home-made or nanofabricated with fixed hole sizes (e.g., C- Flat^{TM} grids from Electron Microscopy Sciences).

1.3 Single-particle cryo-EM data collection

1.3.1 Anatomy and physiology of a transmission electron microscope

A typical transmission electron microscope contains several essential components including an electron gun, electromagnetic lenses and an image detecting system.

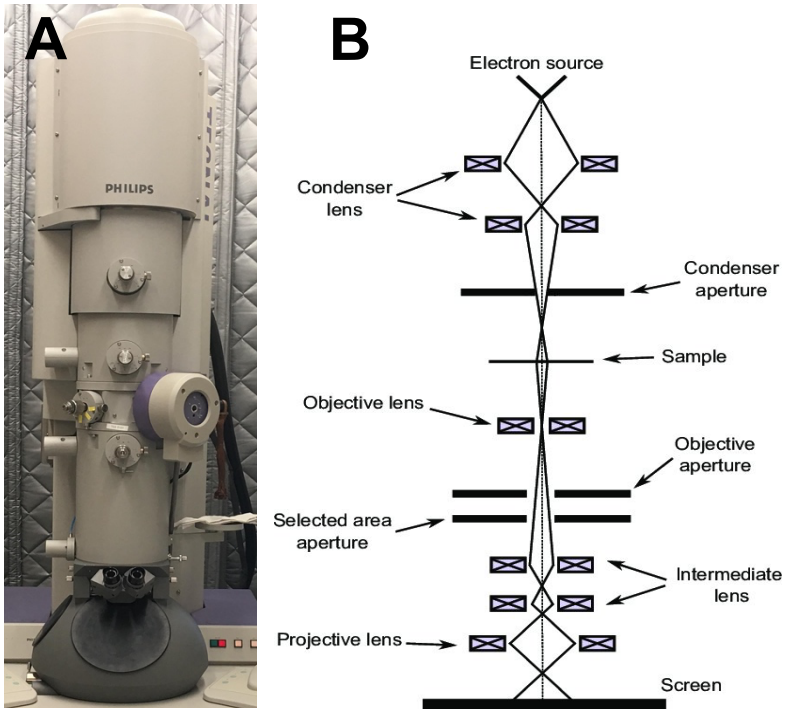

Figure 1.2. Anatomy of a transmission electron microscope (TEM). (A) Photograph of the TEM, FEI TF20 in Eyring Material Science Center, ASU. (B) Schematic illustrations of TEM[15].

The electron sources can be generated from a heated tungsten filament, a LaB6 crystal, or a field emission gun (FEG). FEG emits a constant, near-monochromatic, and high-current electron beam. So far, the FEG is commonly used electron source for cryo-EM. Similar to a light microscope, the electromagnetic lens system used in TEM is composed of condenser, objective, intermediate and projective lenses. In bright-field imaging, the condenser lenses generate the beam that is parallel and illuminated on the specimen. The condenser aperture is placed below the condenser lens to select the low-angle scattered forwarded electrons for a parallel beam. The objective lens forms an initial image and

provides a primary magnification, approximately 20-50X. It is then further magnified by intermediate and projector lenses before the electrons arrive at its detector. Furthermore, adjusting the "focus knob" changes the objective lens current and set the targeted focus values.

Even though TEM is a powerful tool for structural determination, there are some fundamental limitations in the performance of any microscope. These limitations are called spherical, chromatic, and astigmatic aberrations. Spherical aberration relates to the performance of the objective lens. The electrons passing through the edge of a lens are refracted more than those passing through the center. Hence, the electrons that do not reach a common focal point can distort the image and cause points to become blurred disks. Chromatic aberrations are wavelength-dependent artifacts resulted by electrons vary in energy. Electrons that have longer wavelength are focused intensely by the lens, which makes the image formed in a plane closer to the object. Finally, axial astigmatism is a consequence of the objective lens having imperfect axial symmetry. Axial astigmatism can cause circular images of the electron source to deform and become ellipses. Axial astigmatism can be corrected using stigmators that are situated under the objective lens[16].

1.3.2 Image formation in the transmission electron microscope

To form an image using TEM, electrons must first interact with an object present in the beam path. When the electrons pass through the specimen, they can either be scattered without loss of energy (elastic scattering) or be scattered with an exchange of energy between the electron beam and the specimen (inelastic scattering). Elastically scattered electrons change direction without changing the wavelength, which contribute to the image

intensity change. Inelastically scattered electrons increase in wavelength since the electrons lose part of their energy through interactions with the specimen.

Objects that interact with electron beam are typically biological molecules, which usually consist of light atoms such as hydrogen, oxygen, nitrogen and carbon. Thus, biological specimens are regarded as weak-phase objects that will not change the amplitude of the incident electron wave but slightly change the phase of the wave. They are called weak phase objects[17]. The contrast of an image of a biological sample is approximately proportional to the projected Coulomb potential.

A TEM sample is usually described by a potential model where the atomic potentials modify the phase of the wave of the incident electrons as the electrons passes through the material. It is represented by the phase object approximation (POA), given σ as an interaction constant, V_t as the projected potential of the specimen, and r is a two-dimensional vector x and y.

$$q(r) = e^{-i\sigma V_t(r)}$$

For a thin specimen, σV_t is less than unity and the phase change is small. Using the first-degree Taylor polynomial of q(**r**), the POA model can be simplified by the weak phase object approximation (WPOA), which is a linear approximation, given by

$$q(r) \approx 1 + i\sigma V_t(r)$$

Weak-phase-objects (WPO) specimens scatter electrons weakly, so the electron beam passing thorough the thin specimen only undergoes a phase shift and its amplitude is unchanged. So, there is no amplitude contrast when imaging WPO specimens, as electrons pass through different regions there is no intensity difference in images with a microscope that is exactly in focus. As a result, defocus is generally applied to create phase contrast in

an image. Applying defocus leads to a systematic alteration of the image data, and this alteration is denoted as contrast transfer function (CTF). In conclusion, the image formation is described by the action of the contrast transfer function (CTF) H(k). In Fourier space, image contrast can be expressed as

$$I(k) = O(k)H(k)$$

where I(k) represents the Fourier transform of the observed image, O(k) is the Fourier transform of the object, H(k) is the CTF, and k is the spatial frequency. The term k is a direct conversion of the 2D vector x and y, k= (*kx, ky*). Depending on the defocus setting, features of the object can appear enhanced or repressed in the image. This is due to oscillation of the CTF between -1 (negative contrast transfer) and +1 (positive contrast transfer) as the spatial frequency goes from low to high. So, the exact location of the zero crossings where no contrast is transferred, and information is lost is highly dependent on the defocus.

1.3.3 Electron detection and image acquisition

Imaging in electron microscopy is conducted by one of the three main detector technologies: photographic film, charge-coupled devices (CCD) and direct detector devices (DDD)[18]. Before the widespread adoption of digital imaging, TEM images had mostly been recorded with photographic film. This film is coated with thick emulsion on a polyester base, which contains small light-sensitive silver halide crystals[19]. In image recording, the film is exposed to electrons directly within the camera chamber, resulting in an intensity-based exposure pattern in the silver halide that can then be washed away in a chemical development step.

An alternative to photographic film is digital capture using a computer digitizing and archiving camera. The CCD is a two-dimensional digital semiconductor photoelectric conversion device that contains the following components: A scintillator, fiber optics, and a photo sensor. In order for the CCD to detect an electron beam, electrons are converted into photons by a fluorescent material or yttrium aluminium garnet crystal. When the CCD is exposed to light, electron charges are accumulated in a depletion region. These charges are subsequently transferred to a neighboring region, where they can be converted to electric signals[20]. Compared with traditional photographic films, CCD cameras have the advantage of rapid real-time readout, and the image can be reconstructed and displayed digitally on a computer monitor. With this said, the CCD camera does have a major drawback: The signal it produces is the result of two conversions, electron to photon, then photon back to electron. Because of this, the spatial resolution is reduced due to multiple scattering within signal transferring or when the signals arrive at the optical interface[21] (Figure 1.4).

Since DDD has been created, it can be directly exposed to the high energy electron beam and convert electrons into voltage signals. Different from CCD, DDD is a product of signals based on accumulated charges. With a new CMOS design, a DDD allows for direct exposure to the incident electron beam, and significantly improves the signal-to-noise ratio (SNR). The first commercially available CMOS-based DDD camera for cryo-EM was the DE-12 camera from Direct Electron, LP (San Diego, CA)[17]. Other popular DDD cameras are the K2 Summit from Gatan, Inc. (Pleasanton, CA), and Falcon camera from FEI (Hillsboro, OR).

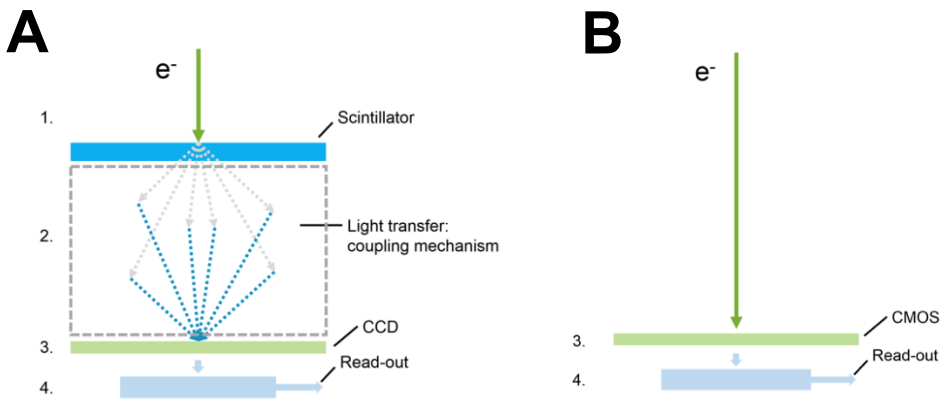

Figure 1.3. Schematic representation of a CCD and CMOS-based DDD for TEM. (A) CCD contains multiple conversion steps (left). Step 1: the scintillator converts electron to signal. Step 2: signal is transferred via coupling mechanism. Step 3: CCD sensor detects signal. Step 4: signal is electronically transferred and read-out to form an image. (B) CMOS-based DDD contains only two steps in TEM image (right). Step 3: sensor directly detects signal. Step 4: signal is electronically transferred and read-out to form an image[22].

1.4 Single-particle 3D reconstruction

Single-particle three-dimensional (3D) reconstruction is accomplished using a 3D electron density map of the molecule. The 3D density map is computed from a set of single-particle images that represent the projections of molecules shown from different orientations. There are several steps required before a final 3D density map can be created: CTF estimation and correction, particle picking and selection, two-dimensional classification, initial three-dimensional reconstruction, and refinement of the reconstructed density map.

1.4.1 CTF estimation and correction

After single-particle data is collected, micrographs are subjected to CTF estimation and correction. Micrographs with minimal ice contaminations are identified and sorted out for further processing. The CTF models a linear image formation process in the transmission electron microscope. The key parameters required for CTF estimation are acceleration voltage, spherical aberration constant, defocus, amount of axial astigmatism, and amplitude

contrast. Acceleration voltage and spherical aberration coefficient can be obtained from the instrument. The defocus value is set during data collection, which is only an estimate. More accurate values for defocus can be obtained through later measurements. CTF determination is carried out using a least-squared fitting method to find the set of parameters that minimize the difference between the calculated CTF model and the observed power spectrum. There are several software packages capable of CTF correction that are currently available, such as SPIDER[23], EMAN[24], CTFFIND4[25] and GCTF[26]. In our studies, CTFFIND4 software was used for CTF estimation and correction. CTFFIND4 is an algorithm that consists of several steps: an amplitude spectrum is computed from the input micrograph, the spectrum's background is then estimated and subtracted from the original spectrum, then the similarity between the theoretical two-dimensional CTF functions and the remaining oscillatory signal is finally evaluated. The parameters for the theoretical CTF can be varied until the similarity is maximized, which results in an estimate of the microscope's defocus and astigmatism parameters[25].

1.4.2 Particle picking and extraction

After micrograph sorting, particles representing the biological molecules of interest are selected and extracted for downstream image analysis. The particle selection is created to distinguish between, and sort particles selected from micrographs that include contaminants or noise, and good particles that represent only the target biological molecule. There are several particle picking software packages that are currently available, including EMAN2[27], DOGPICKER[28], RELION[29], and CryoSPARC[30]. The particle picking step can be done using either a reference-based or reference-free picking algorithm. In the reference-based approach, a reference is generated from a 2D class average of particles

selected from micrographs[28]. In the reference-free approach, characteristics such as shapes and edges of the particles are expressed as a vector. The experimental features calculated from local areas within the micrographs are compared to a corresponding set of predicted features.

In our cryo-EM studies of polymerase ζ (Chapter 3) and integrin (Chapter 4), we used the Manual Picker and Template Picker feature in CryoSPARC[30]. Manual Picker allows us to manually pick the particles interactively from the micrographs. The particles selected in Manual Picker are extracted and classified into 2D classes to generate a reference for automated Template Picker. The Template Picker feature then auto-picks particles from all micrographs using this previously generated template.

1.4.3 Two-dimensional classification

Two-dimensional classification is performed for the following purposes: to eliminate any invalid particles from the particle picking and extraction step, to quantitatively evaluate data quality and show angular distribution of particle views, and to estimate the potential resolution before performing the subsequence three-dimensional (3D) reconstruction steps. The 2D classification often utilizes an un-supervised method known as K-means clustering. K-means groups objects into k groups based on their characteristics. The number k refers to the number of centroids in a dataset. A centroid is a location representing the center of the cluster[31]. For this method, the distance used is usually the quadratic or Euclidean distance. The K-means clustering algorithm has three steps. The first step is initialization, meaning the number of groups k are chosen and k centroids are established in the data space in random. The second step is assignment of objects to the centroids, as each object of the data is assigned to its nearest centroid. The final step is a centroid update, in which

the position of the centroid of each group is modified and a new centroid location is assigned according to the average position of the objects belonging to said group[32]. The major advantage of K-means clustering is its speed and simplicity; however, it is important to decide the value of k, as its final result is dependent upon the initialization step. K-means clustering can also be validated using the Maximum Likelihood[33].

1.4.4 Three-dimensional reconstructions

A three-dimensional (3D) object can be reconstructed from its 2D classifications using several methods, such as a back-projection or Fourier based approach. The underlying principle in 3D reconstruction is the projection theorem. The projection theorem means each projection of an object provides a central section of the 3D Fourier transform of the object (Figure 1.5). According to this theorem, a reconstruction can be obtained with a 3D inverse Fourier transformation of the 3D Fourier domain, which can be filled in by the data obtained from the 2D Fourier sections.

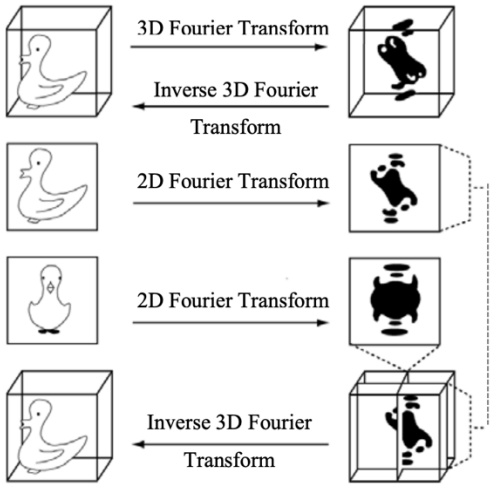

Figure 1.4. Illustration of the projection theorem and its use in 3D reconstruction[34].

3D reconstruction using back-projection, an inverse function of projection, is commonly used to generate a 3D model. For an n-dimensional object that is projected, each projection is an n-1 dimensional sum of its density along the projection axis. For example, all projections of a sphere would be circles while projections of a cube shape object could be squares or any other parallelograms based on the corresponding direction of projection. Back projection inverts these resulting 2D projections, which allows reconstruction of the original object.

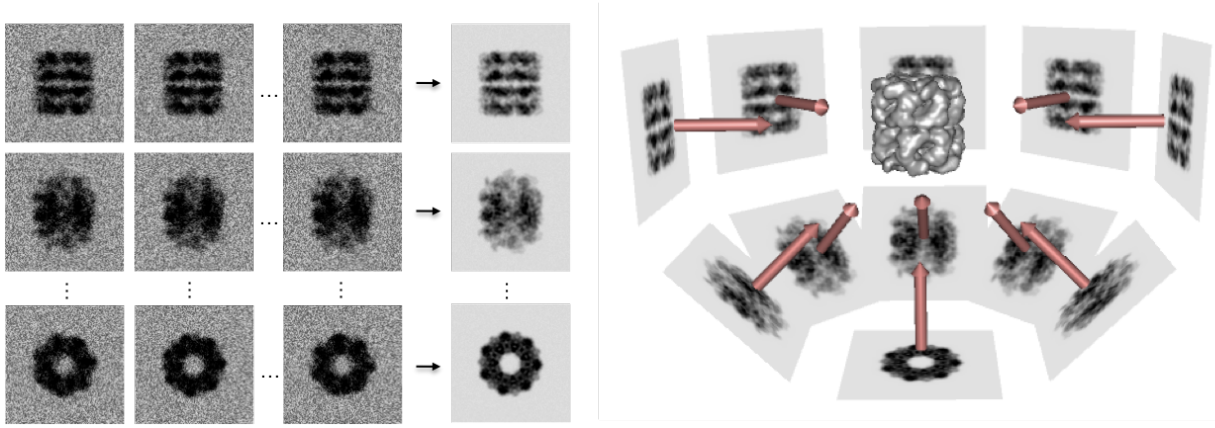

Figure 1.5. Illustration of the back-projection method in 3D reconstruction. Cryo-EM 2D class averages (left) and back-projection (right) of protein. Three clusters are averaged from the images on the same row. The 3D orientation of each of these average images is found. With the resulted orientations, images are back projected to create a 3D density map[35].

Using back-projection methods also means "smearing" each projection across the reconstructed image, which uses a 2D image to create a 3D volume. However, the result of this method yields a blurry object due to a decrease in sampling points as the spatial frequency increases in Fourier space. The Fourier amplitudes at low spatial frequencies are weighted more than the amplitudes at high spatial frequencies. This problem can be solved by reducing the weighting on low frequency components, as well as applying a high-pass filter to each image before back-projection.

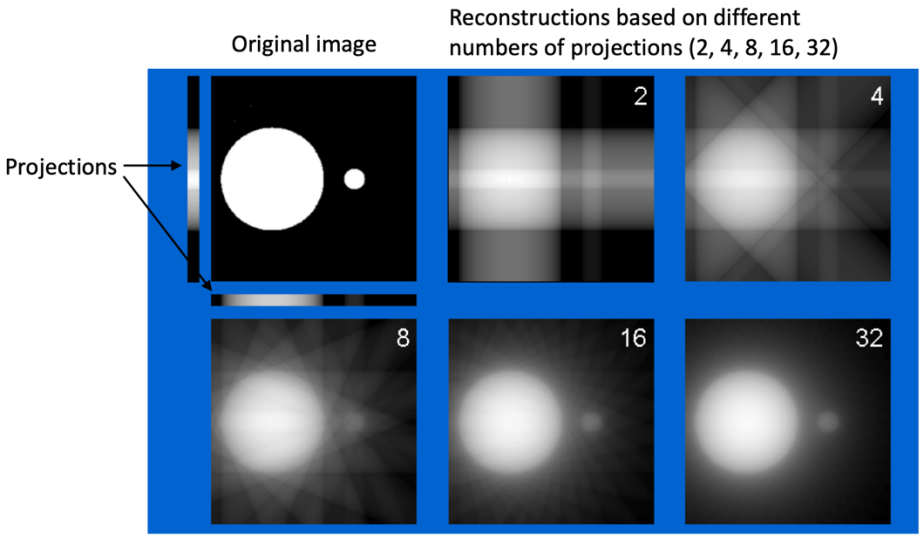

Figure 1.6. The result of back-projection is a blurred version of the original image. Original image is shown in the top left corner, generated from different projections labeled 2, 4, 8, 16, and 32[36].

The Fourier based approach is similar to the approach taken in X-ray crystallography, in which data is collected to fill Fourier space. The inverse Fourier transform generates the map of the object in real space, but interpolation and inversion methods are still required since the information is not complete and must be rebuilt by interpolation between the central sections.

There are several software packages for single-particle 3D reconstruction algorithms such as: SPIDER[23], EMAN[24,27], RELION[29] and CryoSPARC[30]. In this study, RELION and CryoSPARC were used, and CryoSPARC was being used most extensively.

1.4.5 Structural refinement and resolution evaluation

There are a few steps to further refine the initial 3D model. The first step is to find the view angle that best matches the 3D model from each 2D class averages. Second, using these newly estimated view angles, a better 3D model can be reconstructed. However, structure refinement methods are prone to overfitting- The converged model can show features that

do not exist but are miscategorized as such due to reflections of noise in the images. Recently, several methods have been developed to mitigate this issue, for example some software like RELION and CryoSPARC use a Bayesian approach to solve this problem.

After multiple rounds of 3D refinement, a good density map containing high-resolution information of a structure is then generated. Determining the resolution of the final density map is a commonly used method to evaluate the refined results. The term resolution is often used as a numerical indicator to quantify the resolvability of object details reconstructed in the density map. In our cryo-EM study of polymerase ζ (chapter 3), we evaluated the resolution of the protein using gold standard protocol with Fourier Shell Correlation (FSC) = 0.143 criterion. For this method, the data are split into two equivalent subsets by separating even and odd numbered images. Each set of data is reconstructed and refined independently from each other. FSC is expressed as a function of spatial frequency by the following equation:

$$FSC(k) = \frac{\sum_{k,\Delta k} F_1(k)F_2(k)^*}{\sqrt{\sum_{k,\Delta k} |F_1(k)|^2 \sum_{k,\Delta k} |F_2(k)|^2}}$$

In this equation, F1(k) and F2(k) are 3D Fourier transforms of the two structures reconstructed from the two independent subsets. F2(k)* is a complex conjugate of F2(k). The Σk, ΔkF(k) means summing up F(k) over a small range Δk around k. FSC (k) takes a value between -1 and +1 for every spatial frequency k. When the FSC (k) value is close to +1, it means that the two reconstructed structures are agreeable and reliable for the reconstructed structure.

Moreover, detailed structural features can also confirm the resolution indicated by the FSC. For low-resolution map (10 Å or higher), there is less information about the overall shape

of the molecules or the arrangements of different subunits in a protein complex. With a high-resolution map (4 Å or higher), secondary structures such as α-helices and β-strands become identifiable and can be used for polypeptide backbone determination. Likewise, the relative arrangements of subunits in a protein complex can also be indicated.

1.5 Cryo-EM for protein complexes: Nogo receptor complex, Polζ, and integrin

In my studies, I utilized this highly advanced and robust technique to study biological specimen such as protein complexes in their near native environment. The investigation of a protein complex via cryo-EM requires a comprehensive approach going from gene to structure. First, proteins of interest must either be recombinantly expressed and isolated from a cell system or be extracted from their native cells[37]. Chapter 2 describes the process of protein expression and purification, in which I recombinantly expressed the Nogo receptor complex using baculovirus expression system in the *sf9* insect cells. After the proteins are expressed and isolated using different chromatography techniques such as Affinity Chromatography and Size Exclusion Chromatography (SEC), a negative stain EM experiment is performed to assess the protein quality[38]. Once the negative stain EM shows homogenous protein particles, downstream methods such as plunge freezing sample in an EM grid for cryo-EM data collection, and image processing can be carried out to achieve high-resolution protein structures. Chapter 3 showcases an example of solving high-resolution protein structure using cryo-EM. This chapter describes the structural study of biological macromolecule such as the polymerase ζ holoenzyme.

1.6 Recent advances and limitations of Cryo-EM

The revolution of cryo-EM is mainly from the development of novel instruments and image-processing technologies, as well as the introduction of automation in data collection

and analysis. These factors allow for a comprehensive study of complex structures and their functions at the near-atomic level. The recent development of a new generation of direct electron detection cameras over the past decade has offered clear advantages over the traditional photographic film and scintillator-based CCD cameras. DDD cameras provide high quality images and fast readout time, which have greatly increased the efficacy of structural biology[17,18]. Furthermore, newly developed, and user-friendly data processing software such as cryoSPARC allows cryo-EM structures to be determined within a shorter period of time and less requirements for computer power[30]. This advancement in image processing has allowed biologists to investigate highly flexible protein structure such as the transmembrane linker integrin[39]. Although, this can also be a bottleneck for many scientists, since cryo-EM image data reveal heterogeneity in the protein structure and to refine 3D maps to high resolution becomes challenging. Chapter 4 describes our recent investigation of the integrin αvβ8 in complex with its binding partner, pro-TGFβ1-GARP protein, and Fab antibody.

Many advancements in the cryo-EM technique have helped increase the feasibility of high-resolution structural discovery. However, there are some limitations when it comes to sample preparation. In order for a protein structure to be solved using cryo-EM, the protein particles must be homogenous and evenly distributed on the EM grid. Particles that are well distributed in the grid hole, and also embedded in a thin vitreous ice allows for high quality data to be collected automatically on the microscope. In certain cases, proteins exhibit a preferred orientation on the grid, and hinders the high-resolution structure discovery. The preferred orientation problem is resulted from particles within each hole are adsorbed to one of the two air-water interfaces, causing them to adopt a preferred

orientation on grids. Notorious cases of the preferred orientation problem are the ATPase p97 protein[40] and influenza hemagglutinin trimer[41]. In chapter 5, we addressed this problem and developed a new method for minimizing the particle adsorption in air-water interface via lipid monolayer sandwich.

In summary, cryo-EM is an effective technique for identifying the structures of biological macromolecules. Recent advancements in microscope design and imaging technology, along with improved image processing and automation capabilities, increase the effectiveness of cryo-EM techniques. These advancements enhance the speed and scope of automation, as well as the resolutions that can be reached, making this technology capable of determining a wide range of biological structures at high-resolution details.

1.7 References

1. Cheng, Y., Grigorieff, N., Penczek, Pawel A., & Walz, T. (2015). A Primer to Single-Particle Cryo-Electron Microscopy. *Cell*, *161*(3), 438–449. https://doi.org/10.1016/j.cell.2015.03.05

2. Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P. M. G. E., Grigoras, I. T., Malinauskaite, L., Malinauskas, T., Miehling, J., Uchański, T., Yu, L., Karia, D., Pechnikova, E. V., de Jong, E., Keizer, J., Bischoff, M., McCormack, J., Tiemeijer, P., & Hardwick, S. W. (2020). Single-particle cryo-EM at atomic resolution. *Nature*, *587*(7832), 152–156. https://doi.org/10.1038/s41586-020-2829-0

3. Faruqi, A. R., & McMullan, G. (2011). Electronic detectors for electron microscopy. *Quarterly Reviews of Biophysics*, *44*(3), 357–390. https://doi.org/10.1017/s0033583511000035

4. Li, X., Mooney, P., Zheng, S., Booth, C. R., Braunfeld, M. B., Gubbens, S., Agard, D. A., & Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic resolution single-particle cryo-EM. *Nature Methods*, *10*(6), 584–590. https://doi.org/10.1038/nmeth.2472

5. Milazzo, A.-C., Cheng, A., Moeller, A., Lyumkis, D., Jacovetty, E., Polukas, J., Ellisman, M. H., Xuong, N.-H., Carragher, B., & Potter, C. S. (2011). Initial evaluation of a direct detection device detector for single particle cryo-electron microscopy. *Journal of Structural Biology*, *176*(3), 404–408. https://doi.org/10.1016/j.jsb.2011.09.002

6. McMullan, G., Chen, S., Henderson, R., & Faruqi, A. R. (2009). Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy*, *109*(9), 1126–1143. https://doi.org/10.1016/j.ultramic.2009.04.002

7. Bai, X., Fernandez, I. S., McMullan, G., & Scheres, S. H. (2013). Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *ELife*, *2*. https://doi.org/10.7554/elife.00461

8. Campbell, Melody G., Cheng, A., Brilot, Axel F., Moeller, A., Lyumkis, D., Veesler, D., Pan, J., Harrison, Stephen C., Potter, Clinton S., Carragher, B., & Grigorieff, N. (2012). Movies of IceEmbedded Particles Enhance Resolution in Electron Cryo-Microscopy. *Structure*, *20*(11), 1823–1828. https://doi.org/10.1016/j.str.2012.08.026

9. Cheng, Y. (2015). Single-Particle Cryo-EM at Crystallographic Resolution. *Cell*, *161*(3), 450–457. https://doi.org/10.1016/j.cell.2015.03.049

10. Dutta, M. (2018). Recent Advances in Single Particle Cryo-electron Microscopy and Cryo electron Tomography to Determine the Structures of Biological Macromolecules. *Journal of the Indian Institute of Science*, *98*(3), 231–245. https://doi.org/10.1007/s41745-018-0087-z

11. Passmore, L. A., & Russo, C. J. (2016). Specimen Preparation for High-Resolution Cryo-EM. *Methods in Enzymology*, 51–86. https://doi.org/10.1016/bs.mie.2016.04.011

12. Syed, T. H., Hughes, T. J., Marsh, K. N., & May, E. F. (2012). Isobaric Heat Capacity Measurements of Liquid Methane, Ethane, and Propane by Differential Scanning Calorimetry at High Pressures and Low Temperatures. *Journal of Chemical & Engineering Data*, *57*(12), 3573–3580. https://doi.org/10.1021/je300762m

13. Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowall, A. W., & Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics*, *21*(2), 129–228. https://doi.org/10.1017/s0033583500004297

14. Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P., & Ranson, N. A. (2016). An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*, *100*, 3–15. https://doi.org/10.1016/j.ymeth.2016.02.017

15. Marturi, N. (2013). *Vision and visual servoing for nanomanipulation and nanocharacterization in scanning electron microscope* [PhD Thesis].

16. Rodenburg, J. (n.d.). *understand astigmatism in the electron lens*. Www.rodenburg.org. Retrieved September 22, 2021, from http://www.rodenburg.org/guide/t600.html

17. Liao, Y. (2006). *Practical Electron Microscopy and Database*. www.globalsino.com/EM/

18. Bammes, B. E., Rochat, R. H., Jakana, J., Chen, D.-H., & Chiu, W. (2012). Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency. *Journal of Structural Biology*, *177*(3), 589–601. https://doi.org/10.1016/j.jsb.2012.01.008

19. Ayato, H., Mori, N., Miyahara, J., & Oikawa, T. (1990). Application of the Imaging Plate to TEM Observation. *J Electron Microsc (Tokyo)*, *39*(6), 444–448. https://doi.org/10.1093/oxfordjournals.jmicro.a050835

20. Fan, G. Y., & Ellisman, M. H. (2000). Digital imaging in transmission electron microscopy. *Journal of Microscopy*, *200*(1), 1–13. https://doi.org/10.1046/j.1365-2818.2000.00737.x

21. Fan, G. Y., & Ellisman, M. H. (1993). High-sensitivity lens-coupled slow-scan CCD camera for transmission electron microscopy. *Ultramicroscopy*, *52*(1), 21–29. https://doi.org/10.1016/0304-3991(93)90019-t

22. *Imaging | Gatan, Inc.* (n.d.). Www.gatan.com. Retrieved September 22, 2021, from https://www.gatan.com/techniques/imaging

23. Frank, J., Shimkin, B., & Dowse, H. (1981). Spider—A modular software system for electron image processing. *Ultramicroscopy*, *6*(4), 343–357. https://doi.org/10.1016/s0304-3991(81)80236-7

24. Ludtke, S. J., Baldwin, P. R., & Chiu, W. (1999). EMAN: Semiautomated Software for High Resolution Single-Particle Reconstructions. *Journal of Structural Biology*, *128*(1), 82–97. https://doi.org/10.1006/jsbi.1999.4174

25. Rohou, A., & Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, *192*(2), 216–221. https://doi.org/10.1016/j.jsb.2015.08.008

26. Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *Journal of Structural Biology*, *193*(1), 1–12. https://doi.org/10.1016/j.jsb.2015.11.003

27. Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., & Ludtke, S. J. (2007). EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology*, *157*(1), 38–46. https://doi.org/10.1016/j.jsb.2006.05.009

28. Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S., & Carragher, B. (2009). DoG Picker and TiltPicker: Software tools to facilitate particle selection in single particle electron microscopy. *Journal of Structural Biology*, *166*(2), 205–213. https://doi.org/10.1016/j.jsb.2009.01.004

29. Scheres, S. H. W. (2015). Semi-automated selection of cryo-EM particles in RELION-1.3. *Journal of Structural Biology*, *189*(2), 114–122. https://doi.org/10.1016/j.jsb.2014.11.010

30. Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

31. Garbade, M. (2018, September 12). *Understanding K-means Clustering in Machine Learning*. Towards Data Science; Towards Data Science.

https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1

32. Xu, Y., Wu, J., Yin, C.-C., & Mao, Y. (2016). Unsupervised Cryo-EM Data Clustering through Adaptively Constrained K-Means Algorithm. *PLOS ONE*, *11*(12), e0167765. https://doi.org/10.1371/journal.pone.0167765

33. McKay, D. J. C. (2003, October). *Information Theory, Inference and Learning Algorithms | Pattern recognition and machine learning*. Cambridge University Press. http://www.cambridge.org/0521642981

34. Renault, L., & Stahlberg, H. (2007). Three-Dimensional Electron Microscopy of Macro-molecular Assemblies: Visualization of Biological Molecules in Their Native State. By Joachim Frank. Oxford University Press, New York (2006). ISBN 0195182189; softback; 432 pages. *Scanning*, *29*(1), 37–37. https://doi.org/10.1002/sca.20031

35. Pintilie, G. (n.d.). *Greg Pintilie*. People.csail.mit.edu. https://people.csail.mit.edu/gdp/cryoem.html

36. *Back Projection*. (n.d.). Www.impactscan.org. http://www.impactscan.org/slides/impactcourse/basic_principles_of_ct/img12.html

37. Wingfield, P. T. (2015). Overview of the Purification of Recombinant Proteins. *Current Protocols in Protein Science*, 6.1.1–6.1.35. https://doi.org/10.1002/0471140864.ps0601s80

38. Scarff, C. A., Fuller, M. J. G., Thompson, R. F., & Iadaza, M. G. (2018). Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems. *Journal of Visualized Experiments*, *132*. https://doi.org/10.3791/57199

39. Mezu-Ndubuisi, O. J., & Maheshwari, A. (2020). The role of integrins in inflammation and angiogenesis. *Pediatric Research*. https://doi.org/10.1038/s41390-020-01177-9

40. Rouiller, I., Butel, V. M., Latterich, M., Milligan, R. A., & Wilson-Kubalek, E. M. (2000). A Major Conformational Change in p97 AAA ATPase upon ATP Binding. *Molecular Cell*, *6*(6), 1485–1490. https://doi.org/10.1016/s1097-2765(00)00144-1

41. Tan, Y. Z., Baldwin, P. R., Davis, J. H., Williamson, J. R., Potter, C. S., Carragher, B., & Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods*, *14*(8), 793–796. https://doi.org/10.1038/nmeth.434

CHAPTER 2: HETEROLOGOUS EXPRESSION AND PURIFICATION OF RECOMBINANT HUMAN NOGO RECEPTOR ECTODOMAIN COMPLEX FROM SF9 INSECT CELLS

2.1 Introduction

Neuroplasticity regulates the connections between neurons, which affects learning and memory in the brain[1]. Neuronal axons play an important role in connecting neurons and the axonal growth is tightly controlled[2-6]. Myelin, a lipid-rich insulating layer surrounding axons, involves in the regulation of axonal generation and degeneration[1]. Oligodendrocytes are the main component of myelin cells in the adult central nervous system (CNS) and they do not contain sheath[1,13,27]. In addition to the protective support for the nerve axons, oligodendrocytes regulate their metabolic processes[2]. When an axon encounters injury in the CNS, unlike in the peripheral nerve system (PNS) where neuronal damages are readily repaired, the axonal regeneration is significantly inhibited by oligodendrocytes[3]. This inhibition is triggered by myelin-associated inhibitors (MAIs) expressed by oligodendrocytes, including Nogo-A, myelin-associated glycoprotein (MAG) and oligodendrocyte myelin glycoprotein (OMgp)[4]. The MAI interacts with Nogo receptor-1 (NgR1) on the axonal membrane and initiate a signaling transduction, so called Nogo signaling[5]. Nogo signaling inhibits the axonal regeneration and leads to growth cone collapse, thereby modulates neural plasticity[6,7]. Improper regulations on the axonal growth can result in miscommunication between neurons, neurodegeneration, spinal cord injuries, or multiple sclerosis[8].

The molecular mechanism of the Nogo signaling initiation has remained unclear. The binding site of the Nogo-A on the NgR1 has been identified, but those for MAG and OMgp are still undetermined[1,26]. Nogo-A is a membrane protein, which belongs to the reticulon family (RTN4) and is expressed on the surface of the oligodendrocyte[26]. A 66-amino-acid extracellular loop of the Nogo-A, the Nogo-66 domain, binds to NgR1[5]. The ligand-bound NgR1 will then recruit LINGO-1 (leucine-rich repeat and immunoglobin-like domain-containing protein-1) and either p75 neurotrophin receptor ($p75^{NTR}$) or tumor necrosis factor receptor superfamily, member 19 (TROY) to form a membrane receptor complex[9], the so-called Nogo-receptor complex, which then activates the downstream RhoA/ROCK signaling pathway and modulates cytoskeleton structures, causing the growth cone to collapse[10] (Figure 1).

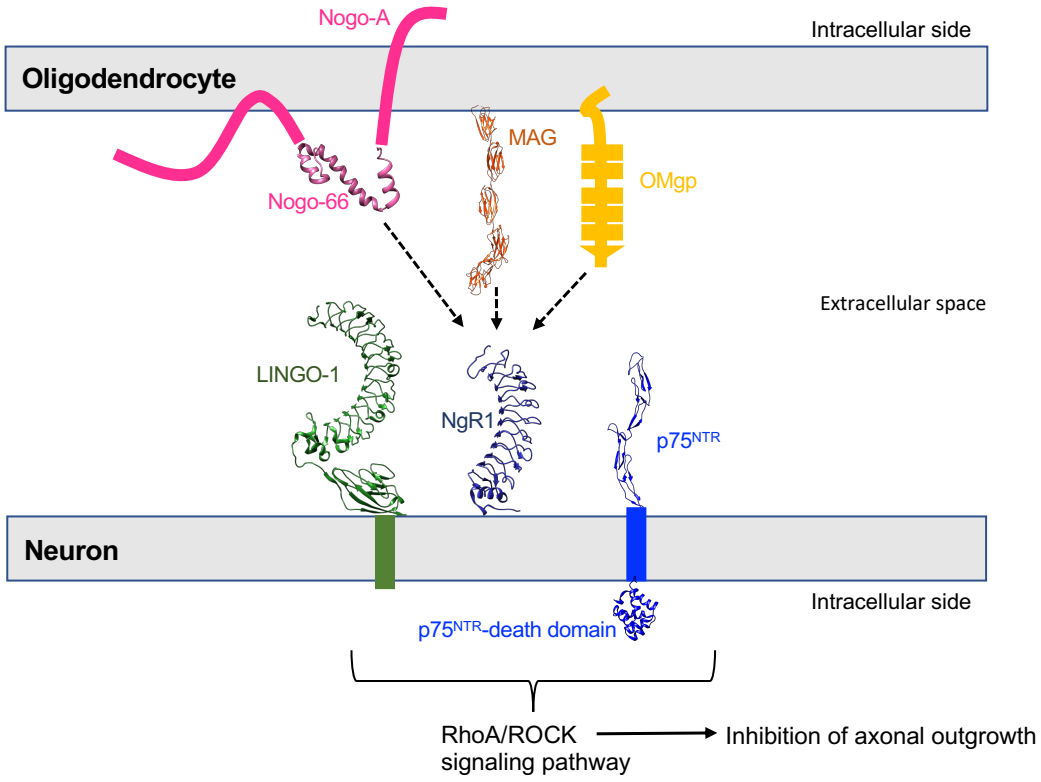

Figure 2.1. Schematic representation of the interactions of myelin-associated inhibitors (MAIs), Nogo-A, MAG and OMgp, NgR1 and its co-receptors. OMgp (yellow), MAG (orange; PDB code: 5LF5[10]) or Nogo-66 domain (pink; PDB code: 2KO2[11]) of the Nogo-A binds to NgR1 receptor (dark blue; PDB code: 1P8T[12]). NgR1 forms a complex with LINGO-1 (green; PDB code: 2ID5[13]) and $p75^{NTR}$ (blue; PDB code: 3BUK[14] and 1NGR[15]) to transduce the signaling to RhoA/ROCK downstream pathway. NgR1 and its co-receptors are expressed in neurons, while the MAI ligands (Nogo-A, MAG and OMgp) are expressed in oligodendrocytes.

The relationship between each component of the Nogo receptor complex is particularly interesting, since the molecular mechanism that drives this complex formation was not fully understood[24, 25]. Therefore, we would like to study how individual component with each other within its complex formation by determining the structure of Nogo receptor complex. In other to do that, we first need to successfully express and purify the Nogo complex for structural determination.

To express individual components within Nogo complex, we used the baculovirus-insect cell expression system to generate the target proteins. Baculovirus-insect cell system is a well-established method for high-quality recombinant protein production[16]. This system has simple post-translational modifications for overexpressing recombinant proteins[17]. Previous studies have shown the overexpression of individual Nogo proteins, such as NgR1[18], $p75^{NTR}$[19], MAG[20], Nogo-66[11] and LINGO-1[13], were feasible via various expression systems, including Baculovirus-insect cell system. To our best knowledge, the in vitro formation of the Nogo-receptor complex has not been reported yet.

In this chapter, we present the results of overexpression and purification of the extracellular domains of Nogo receptor complex (NgR1-LINGO-1-$p75^{NTR}$) from Sf9 insect cell hosts, a clonal isolate of Spodoptera frugiperda Sf21 cells (IPLB-Sf21-AE). Our goal is to assemble the complex and characterize the complex formation in vitro. We expressed

individual proteins and performed the baculoviral co-infection with multiple baculoviruses that carry individual protein genes in the same culture. We were able to identify the Nogo complex from the co-infection experiments. Compared to the overexpression of individual proteins, the protein yield of the co-infected complex was higher than those of the individual components; presumably, the complex formation may stabilize the isolated components. Negative-stain electron microscopy (EM) analysis further confirmed the formation of the complex. Thus, our purification workflow will be suitable for characterizing the protein-protein interactions between neuronal membranes.

2.2 Materials and Methods

2.2.1. Cell culture

Sf9 insect cells, Spodoptera frugiperda cell line IPLB-Sf-21-AE, were purchased from Thermo Fisher Scientific (SKU# 11496-015). The cells were cultured in the purchased Insectagro® Sf9 serum-free medium (Corning 13-410-CV).

2.2.2 Expression vectors

DNA sequences with codon optimization from NgR1 (NP_848663), LINGO-1 (NP_116197), MAG (NP_002352), and Nogo-66 (AAI52556.1) were synthesized with the purification tags, that is, histidine (His)-tag and/or 3X FLAG tag, and the adjacent restriction sites. A detailed description was mentioned in the supplementary materials. The ectodomains of individual gene fragments were then subcloned into pFastBac1 (Thermo Fisher Sciences 10360014). The $p75^{NTR}$-RFP plasmid was requested from Addgene (Watertown, MA; Cat No. 24092). We designed the primers to amplify the DNA fragment of the $p75^{NTR}$ ectodomain and then subcloned into the pFastBac1 vector with the

addition of maltose-binding protein (MBP) and 8X His-tag on the N-terminus (Figure S1). The sequences of all the expression constructs were verified.

2.2.3 Overexpression of individual Nogo proteins in Sf9 cells

The recombinant plasmids of the individual gene (NgR1, LINGO-1, $p75^{NTR}$, MAG and Nogo-66) were transformed into DH10Bac competent cells (Thermo Fisher Sciences 10361012) to generate the bacmid. White colonies in the culture were selected, and the recombinant bacmid was extracted using the procedure provided by Invitrogen (Thermo Fisher Science 10359016). Then, 1 mL Sf9 cells were transfected with 5 μg recombinant bacmid using 8 μL Cellfectin Reagent II (Thermo Fisher Scientific) to produce the recombinant baculovirus. The baculovirus was then amplified in Sf9 cells individually for two more generations (P2 and P3, with the volume increasing from 4 to 30 mL). For protein overexpression, the Sf9 cells were infected with the P3 virus in a logarithmic growth phase (1x106 cells/mL) at a volume density of 20 mL virus per one-liter cell culture. Cells were harvested after 72-hour post-infection, and the overexpressed proteins were verified using SDS-PAGE and Western blotting with an anti-His-tag monoclonal antibody (1:2000 dilution). The co-infection (NgR1-LINGO-1-$p75^{NTR}$-MAG or NgR1-LINGO-1-$p75^{NTR}$-Nogo-66) was carried out by mixing the individually prepared viruses (10 ml of each virus vectors containing the individual component) into 1 L cell culture, which incubated at 27°C for 72 hours.

2.2.4 Immobilized metal chelates affinity and size-exclusion chromatography of NgR1, LINGO-1, and MAG

Cell lysis was carried out using dounce homogenization in Buffer A containing 20 mM HEPES (pH 7.5), 150 mM NaCl, with cOmpleteTM Protease Inhibitor Cocktail (1 tablet

per 50-ml volume, Millipore 11697498001). The final buffer volume was adjusted to a cell density of 10 million cells per 2 mL buffer. The membrane fraction was removed by centrifugation at 45,000×g for 30 minutes at 4oC. The collected supernatant was filtered through a 0.22-μm filter and loaded onto a HisTrap HP column (GE Health Care). The column was first washed with Buffer A for 20 column volumes (CV) with 30 mM imidazole to remove non-specific bindings. Approximately 5 to 7 CV of the Buffer A with 300 mM imidazole was used to elute the protein from the column. The eluted sample was loaded on a Superdex 200 Increase 10/300 GL column (GE Healthcare 28990944) for size-exclusion chromatography (SEC) on a GE ÄKTA pure 25 M1. The peak fractions of target proteins were collected according to the expected elution volumes and their corresponding molecular weights. The cell lysate and eluted sample were analyzed by SDS-PAGE and Western blotting (Figure 2.2 and S2-S5).

2.2.5 Purification of $p75^{NTR}$ proteins

Most of the purification procedures for $p75^{NTR}$ were the same as those for other Nogo proteins, but a different affinity tag on $p75^{NTR}$ for purification was used. We loaded cell lysates onto an MBPTrap HP column (GE Healthcare, SKU-28-9187-80). Lysis and binding buffer were composed of 20 mM HEPES (pH 7.5), 200 mM NaCl, 1 mM Ethylenediaminetetraacetic acid (EDTA), and 1 mM Dithiothreitol (DTT). 10 mM maltose was used for elution. After purification, the eluted sample was loaded onto a Superdex 200 Increase 10/300 GL column for SEC, and the target peak fraction was collected for further analysis (Figure 2G).

2.2.6 Immobilized metal chelate affinity and size-exclusion chromatography of co-infected NgR1-LINGO-1-$p75^{NTR}$-MAG protein complex

Whole cells were collected after co-infection and subjected to cell lysis. The insoluble pellet of the cell lysate was removed by highspeed centrifugation at 45,000Xg at 4°C for 30 minutes. The co-infected proteins were isolated using the Ni-NTA affinity chromatography via a HisTrap HP column (GE Health Care), with Buffer A mentioned previously. The purification procedures were carried out the same as the individual protein purification. The eluate was loaded on a Superdex 200 Increase 10/300 GL column and further purified using the SEC. The peak fractions were collected in the corresponding elution volumes. Samples of the cell lysate and eluates were analyzed by SDS-PAGE and Western blotting (Figure 3).

2.2.7 Antigen-antibody affinity and size-exclusion chromatography of co-infected NgR1-LINGO-1-p75^{NTR}-Nogo-66 protein complex

The insoluble pellet of the cell lysate was removed by highspeed centrifugation at 45,000Xg at 4°C for 45 minutes. The co-infected sample was isolated using the FLAG-affinity chromatography with a column loaded with ANTI-FLAG® M2 Affinity gel (Millipore, SKU# A220). The supernatant was filtered through a 0.22-μm filter and then incubated with the pre-equilibrated 2 mL ANTI-FLAG® M2 Affinity Gel for 1 to 3 hours at 4°C under constant slow agitation. The protein complex-FLAG-resin mixture was then loaded onto a gravity column and washed with the binding buffer composed of 20 mM HEPES (pH 7.5) and 150 mM NaCl. The eluted proteins were collected using the competitive binding of 3X FLAG peptide (Millipore F4799) at the concentration of 0.1 mg/mL for 5-mL elution. The eluate was loaded on a Superose 6 Increase 10/300 GL column and further purified using the SEC. The peak fractions were collected in the

corresponding elution volumes. Samples of the cell lysate and eluates were analyzed by SDS-PAGE and Western blotting (Figure 4).

2.2.8 Polyacrylamide Gel Electrophoresis (PAGE)

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed on 4% stacking gels followed by 12% resolving gels. Protein solutions were mixed with equal volume of 62.5 mM Tris-HCl pH 6.8, 5 mM DTT, 25% glycerol, 0.1% bromophenol blue and 0.2% SDS loading buffer. The protein bands were visualized by 0.1% Coomassie blue R-250.

2.2.9 Negative-stain electron microscopy

We further characterized our complex assemblies using negative stain EM. The sample was negatively stained by 0.75% uranyl formate, followed by the previous method[21]. The electron micrographs of negatively stained complexes were imaged under a Philips CM12 transmission electron microscope (TEM) at an accelerating voltage of 80 keV. The electron scattering signals were recorded using a Gatan model 791 CCD camera with 1,024X1,024 pixels. The image pixel size was determined as 8 Å/pixel at the specimen level.

For single-particle image analysis, 4,016 particles were manually selected from 215 micrographs. The selected particles were subjected to the two-dimensional (2D) image alignment and classification iteratively using RELION software[22]. The class k number was set to 30.

2.3 Results and Discussion

2.3.1 Purification of individual protein NgR1, LINGO-1, $p75^{NTR}$, and MAG

The SEC profile of the NgR1 indicated two dominant peaks, in which the first eluted peak was observed to be aggregates (Figure 2.2A). The Western blot confirmed the identity of NgR1 from the second peak (Figure 2.2B). The molecular weight of the NgR1 is 54 kDa. However, the gel band indicated a molecular weight of the target protein as 90 kDa, which is larger than expected. The previous study suggested that the NgR1 contains two glycosylation sites[18], and the glycosylated form of NgR1 explains the band shift on the gel. The SEC profile of LINGO-1 showed the second most dominant peak, and the sample in the peak fraction was confirmed as LINGO-1 using Western blotting (Figure 2.2C and D). The target protein was shown on two major bands between 100 and 150 kDa. This result is consistent to the previous finding that LINGO-1 has seven glycosylation sites on the extracellular domain and various levels of glycosylation caused the different sizes of band shift on the gel[13]. Our results showed that the insect cell system was able to overexpress the glycosylated LINGO-1 despite aggregations. The purification of $p75^{NTR}$ was conducted and the SEC results were provided (Figure 2.2E). The $p75^{NTR}$ ectodomain has a molecular weight of 45 kDa; the addition of MBP being 42.5 kDa gives the total size of the recombinant MBP-$p75^{NTR}$ to be 87.5 kDa. The Western blot confirmed the identity of glycosylated $p75^{NTR}$ at the expected size (Figure 2.2F).

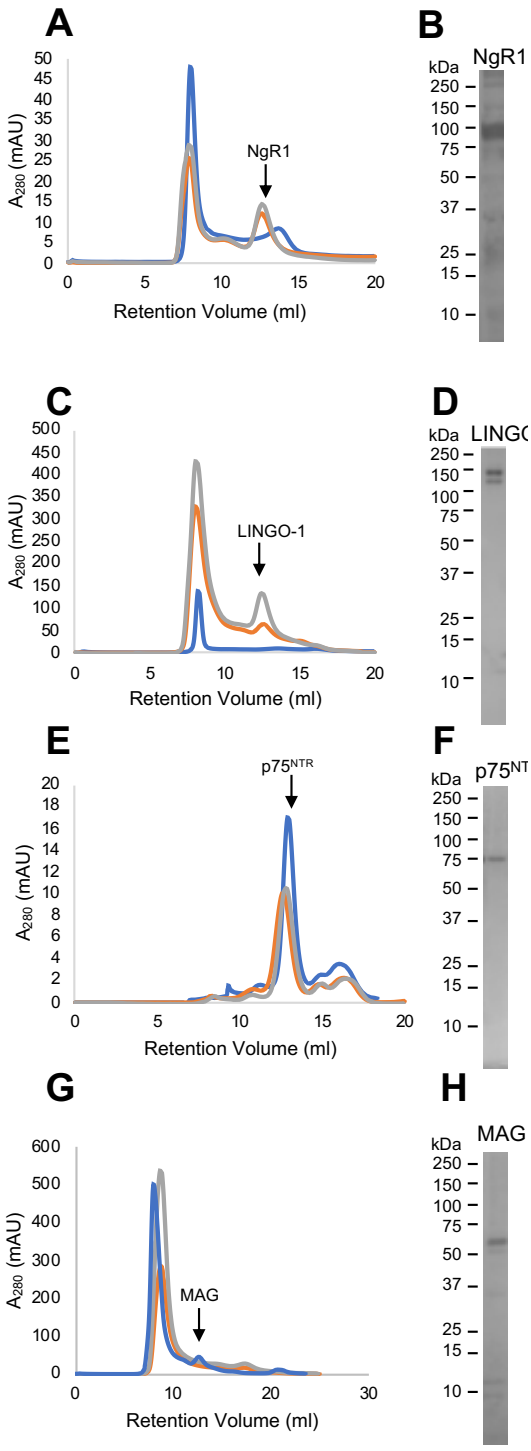

Figure 2.2. Purification of recombinant NgR1, LINGO-1, $p75^{NTR}$ and MAG. Size-exclusion chromatographic profiles of (A) NgR1, (C) LINGO-1, (E) $p75^{NTR}$, and (G) MAG. Western blotting gels of (B) NgR1, (D) LINGO-1, (F) $p75^{NTR}$ and (H) MAG.

According to the purification results, $p75^{NTR}$ was shown to be expressed well with no aggregations. The purification of MAG was carried out, in which the SEC profile of MAG was shown in Figure 2.2G. Multiple trials and optimization of the purification protocol were done. A peak fraction eluted at 12 mL was suggested to contain the dimeric MAG, which is consistent with the previous study that the crystal structure of MAG was shown in a dimeric form and the molecular weight is 120 kDa[23]. The Western blot confirmed the identity of the MAG (Figure 2.2H). Our purification results showed that MAG was overexpressed successfully, and Sf9 cells could produce the protein.

2.3.2 Purification of co-infected Nogo protein complexes

A more stable protein complex with a higher yield was formed when the viruses co-infected the insect hosts. Purification of the Nogo receptor complex with MAG was conducted. The complex was predicted to be approximately 680 kDa, with each component at its corresponding stoichiometry found from previous studies. Figures 2.3A and 2.3B presented the SDS-PAGE and Western blot of the co-infected complex sample. Figure 2.3C showed the SEC profile of the complex. The complex was eluted at 8 ml, which is corresponding to the first peak in Figure 2.3C. Aside from the complex assembled in the first peak, some unbound individual components made up the second peak. It was most likely to be NgR1, $p75^{NTR}$ or MAG since their stoichiometries resulted in a similar molecular weight of approximately 120-160 kDa. The shoulder peak on the left of the second peak most likely to be LINGO-1, because the tetramers were at a higher molecular weight (approximately 240 kDa). Figure 2.3D shows the negative-stain EM analysis of the Nogo receptor complex with its ligand MAG. Along with the purified complex, the resulting SEC and negative stain analysis showed the assembled complex's heterogeneity.

These results entailed that the complex formation was partially stable. Yet, the quantity of the co-infected protein complex was higher than those of individually expressed protein. The overexpressed proteins from co-infection enhance the expression of one another and facilitate the interactions between each other.

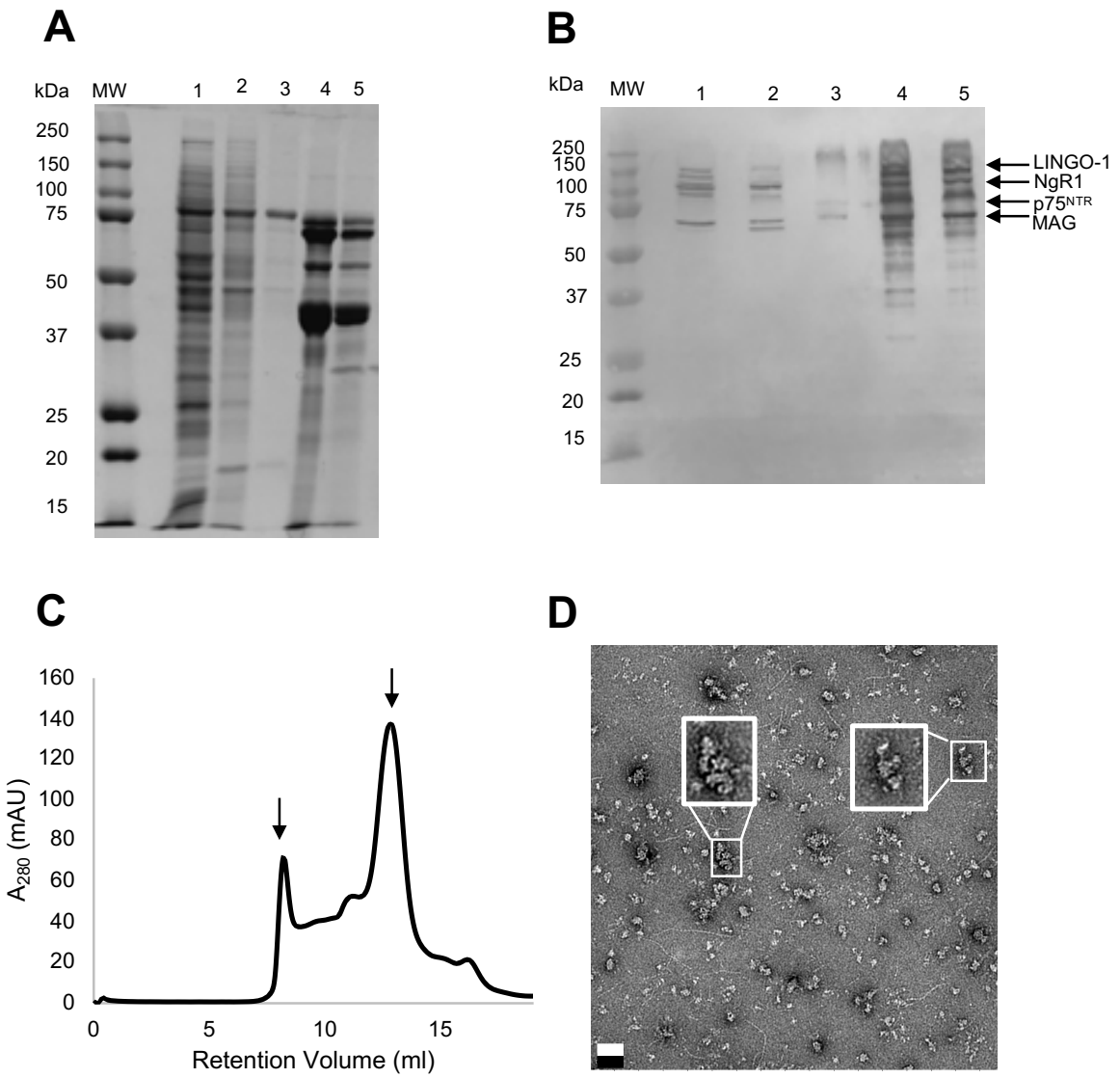

Figure 2.3. Purification of Nogo receptor complex of NgR1, LINGO-1, $p75^{NTR}$ and MAG via co-infection. (A) SDS-PAGE analysis of the Ni-NTA purification of the co-infected complex. Lanes 1, 2, and 3 were the flow-through, first and second washes of the sample. Lanes 4 and 5 were the earlier and later competitive elution fractions using imidazole. Individual components were shown with their molecular weights (MW) on the bands: 120 kDa, 75 kDa, 66 kDa, and 54 kDa. Bands were shifted to a larger MW due to multiple

glycosylation on the receptors. (B) Anti-His Western blotting on the SDS-PAGE gel. (C) Size-exclusion chromatogram of the co-infected protein complex. Arrows indicate the estimated elution position of the assembled complex (left) and unbound proteins (right), respectively. (D) Negative-stain EM analysis of the Nogo-ectodomain complexes. Proteins are shown in white and the background in black. White boxes enlarge the representative particle images. Scale bar indicates 50 nm.

We have also produced another co-infected Nogo complex with another ligand, Nogo-66 peptide. Figures 2.4A and 2.4B presented the SDS-PAGE and Western blot of the purification of the co-infected complex with Nogo-66. Observing from the SEC profile in Figure 2.4C, each peak was better separated, and the individual proteins were isolated along with the bound complex. The complex (594.6 kDa) was eluted at 11 ml and was used for subsequent negative-stain electron microscopic analysis. The 2D class averages of the particle images showed the appearance of the Nogo complex (Figure 2.4D and E). Also, the quantity of the co-infected protein complex was higher than those of individually expressed protein. The protein yield from co-infection was higher because the SDS-PAGE analysis showed more dominant bands of the Nogo proteins. Thus, the overexpression and purification of the Nogo-receptor complex with Nogo-66 were concluded to be feasible for complex formation.

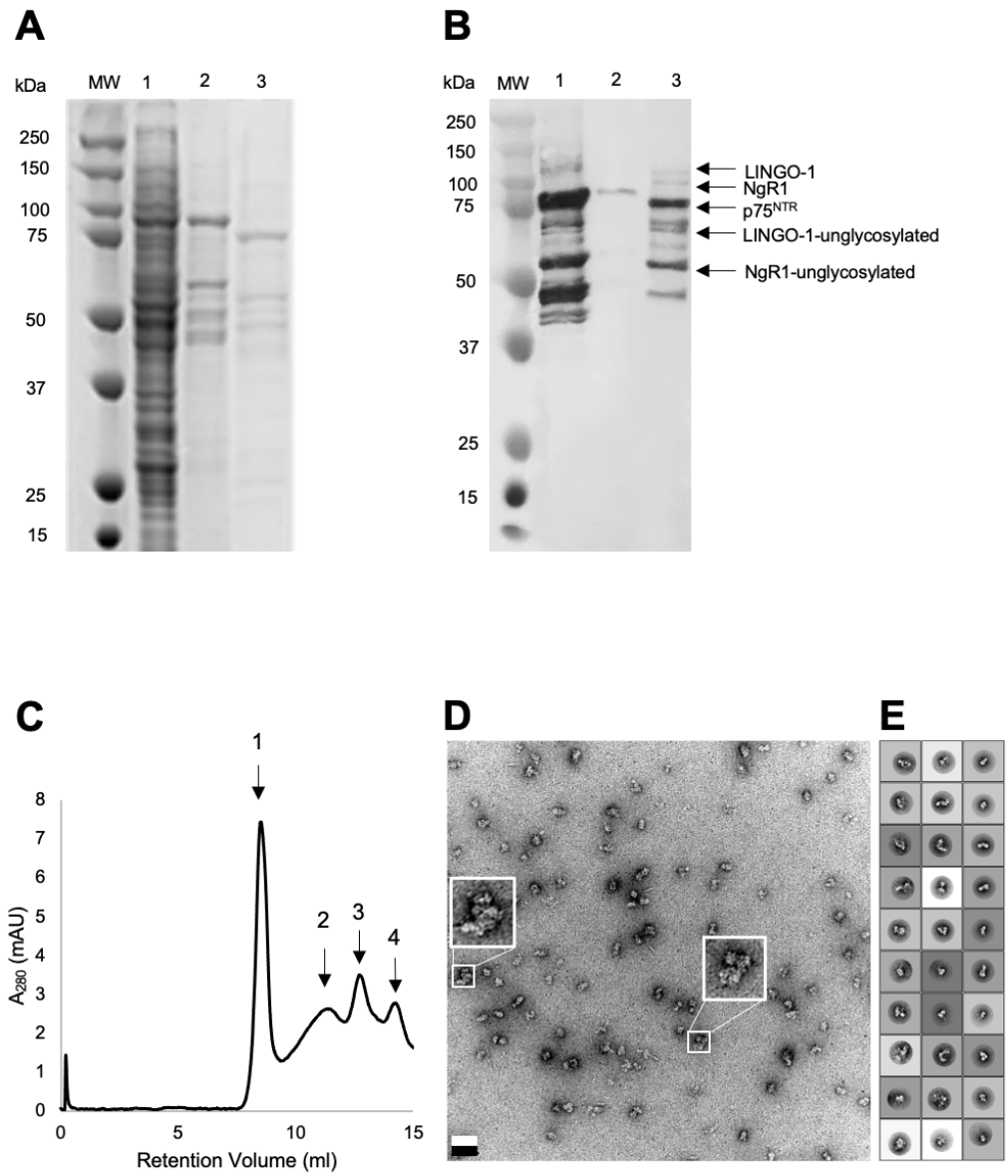

Figure 2.4. Purification of Nogo receptor complex of NgR1, LINGO-1, $p75^{NTR}$ and Nogo-66 via co-infection. (A) SDS-PAGE and Western blot (B) analysis of the FLAG-tag affinity chromatography of the co-infection experiment. Lane 1 was the elution of the FLAG-tag affinity column. Lane 2 was the wash. Lane 3 was the competitive elution, and the bands were shown in the molecular weights (MW) of 130 kDa, 90 kDa, and 75 kDa for LINGO-1, NgR1, and $p75^{NTR}$, respectively. Bands shifted due to multiple glycosylation on the receptor proteins. The bands observed below 75 kDa were unglycosylated LINGO-1 and NgR1, respectively. (C) Size-exclusion chromatogram of the co-infected protein complex. Arrows indicate the estimated elution position of the assembled complex (peak 2) and unbound proteins namely: LINGO-1 (peak 3), NgR1 (peak 4) and $p75^{NTR}$ (peak 4), and peak 1 is aggregation. (D) Negative-stain EM analysis of the Nogo-ectodomain complexes with Nogo-66. Proteins are shown in white and the background in black. White

boxes are representative protein complexes. Scale bar indicates 50 nm. (E) 2D class averages of the Nogo-ectodomain complexes with Nogo-66. Each box length is 35 nm.

2.4 Conclusions

Here we performed the overexpression and purification of the Nogo receptor complex in Sf9 cells and followed by characterization of the expressed proteins by SDS-PAGE and Western blotting. In this study, we reported the production of the native Nogo receptor complex in vivo using a baculovirus expression system in Sf9 cells for the first time. When the complex is co-infected together in the same culture, the chances of protein-protein interaction is observed. The baculovirus expression system enables the glycosylation-rich proteins to produce without structural limitations by having post-translational modification property. Although, the stability of the complex was lower than anticipated. Thus, further studies are needed to refine the expression and purification techniques to improve the stability of Nogo receptor complexes.

2.5 References

1. Schwab, J. M., Tuli, S. K., & Failli, V. (2006). The Nogo receptor complex: confining molecules to molecular mechanisms. *Trends in Molecular Medicine*, *12*(7), 293–297. https://doi.org/10.1016/j.molmed.2006.05.001

2. Chamberlain, K. A., Nanescu, S. E., Psachoulia, K., & Huang, J. K. (2016). Oligodendrocyte regeneration: Its significance in myelin replacement and neuroprotection in multiple sclerosis. *Neuropharmacology*, *110*(Pt B), 633–643. https://doi.org/10.1016/j.neuropharm.2015.10.010

3. Gordon, T., Udina, E., Verge, V. M. K., & de Chaves, E. I. P. (2009). Brief Electrical Stimulation Accelerates Axon Regeneration in the Peripheral Nervous System and Promotes Sensory Axon Regeneration in the Central Nervous System. *Motor Control*, *13*(4), 412–441. https://doi.org/10.1123/mcj.13.4.412

4. Rao, S. N. R., & Pearse, D. D. (2016). Regulating Axonal Responses to Injury: The Intersection between Signaling Pathways Involved in Axon Myelination and The Inhibition of Axon Regeneration. *Frontiers in Molecular Neuroscience*, *9*. https://doi.org/10.3389/fnmol.2016.00033

5. Akbik, F., Cafferty, W. B. J., & Strittmatter, S. M. (2012). Myelin associated inhibitors: A link between injury-induced and experience-dependent plasticity. *Experimental Neurology*, *235*(1), 43–52. https://doi.org/10.1016/j.expneurol.2011.06.006

6. Yiu, G., & He, Z. (2006). Glial inhibition of CNS axon regeneration. *Nature Reviews Neuroscience*, *7*(8), 617–627. https://doi.org/10.1038/nrn1956

7. Silver, J., Schwab, M. E., & Popovich, P. G. (2014). Central Nervous System Regenerative Failure: Role of Oligodendrocytes, Astrocytes, and Microglia. *Cold Spring Harbor Perspectives in Biology*, *7*(3), a020602. https://doi.org/10.1101/cshperspect.a020602

8. McDonald, L., Bandtlow, C., & Reindl, M. (2011). Targeting the Nogo Receptor Complex in Diseases of the Central Nervous System. *Current Medicinal Chemistry*, *18*(2), 234–244. https://doi.org/10.2174/092986711794088326

9. Meabon, J. S., De Laat, R., Ieguchi, K., Wiley, J. C., Hudson, M. P., & Bothwell, M. (2015). LINGO-1 Protein Interacts with the p75 Neurotrophin Receptor in Intracellular Membrane Compartments. *Journal of Biological Chemistry*, *290*(15), 9511–9520. https://doi.org/10.1074/jbc.m114.608018

10. Mi, S., Lee, X., Shao, Z., Thill, G., Ji, B., Relton, J., Levesque, M., Allaire, N., Perrin, S., Sands, B., Crowell, T., Cate, R. L., McCoy, J. M., & Pepinsky, R. B. (2004).

LINGO-1 is a component of the Nogo-66 receptor/p75 signaling complex. *Nature Neuroscience*, *7*(3), 221–228. https://doi.org/10.1038/nn1188

11. Vasudevan, S. V., Schulz, J., Zhou, C., & Cocco, M. J. (2010). Protein folding at the membrane interface, the structure of Nogo-66 requires interactions with a phosphocholine surface. *Proceedings of the National Academy of Sciences*, *107*(15), 6847–6851. https://doi.org/10.1073/pnas.0911817107

12. Barton, W. A. (2003). Structure and axon outgrowth inhibitor binding of the Nogo-66 receptor and related proteins. *The EMBO Journal*, *22*(13), 3291–3302. https://doi.org/10.1093/emboj/cdg325

13. Mosyak, L., Wood, A., Dwyer, B., Buddha, M., Johnson, M., Aulabaugh, A., Zhong, X., Presman, E., Benard, S., Kelleher, K., Wilhelm, J., Stahl, M. L., Kriz, R., Gao, Y., Cao, Z., Ling, H.-P., Pangalos, M. N., Walsh, F. S., & Somers, W. S. (2006). The Structure of the Lingo-1 Ectodomain, a Module Implicated in Central Nervous System Repair Inhibition. *Journal of Biological Chemistry*, *281*(47), 36378–36390. https://doi.org/10.1074/jbc.m607314200

14. Gong, Y., Cao, P., Yu, H., & Jiang, T. (2008). Crystal structure of the neurotrophin-3 and p75NTR symmetrical complex. *Nature*, *454*(7205), 789–793. https://doi.org/10.1038/nature07089

15. Liepinsh, E. (1997). NMR structure of the death domain of the p75 neurotrophin receptor. *The EMBO Journal*, *16*(16), 4999–5005. https://doi.org/10.1093/emboj/16.16.4999

16. Contreras-Gómez, A., Sánchez-Mirón, A., García-Camacho, F., Molina-Grima, E., & Chisti, Y. (2013). Protein production using the baculovirus-insect cell expression system. *Biotechnology Progress*, *30*(1), 1–18. https://doi.org/10.1002/btpr.1842

17. Jarvis, D. L. (2003). Developing baculovirus-insect cell expression systems for humanized recombinant glycoprotein production. *Virology*, *310*(1), 1–7. https://doi.org/10.1016/s0042-6822(03)00120-x

18. He, X. L., Bazan, J. Fernando., McDermott, G., Park, J. B., Wang, K., Tessier-Lavigne, M., He, Z., & Garcia, K. Christopher. (2003). Structure of the Nogo Receptor Ectodomain. *Neuron*, *38*(2), 177–185. https://doi.org/10.1016/s0896-6273(03)00232-0

19. Feng, D., Kim, T., Özkan, E., Light, M., Torkin, R., Teng, K. K., Hempstead, B. L., & Garcia, K. C. (2010). Molecular and Structural Insight into proNGF Engagement of p75NTR and Sortilin. *Journal of Molecular Biology*, *396*(4), 967–984. https://doi.org/10.1016/j.jmb.2009.12.030

20. Johnson, P. W., Attia, J., Richardson, C. D., Roder, J. C., & Dunn, R. J. (1989). Synthesis of soluble myelin-associated glycoprotein in insect and mammalian cells. *Gene*, *77*(2), 287–296. https://doi.org/10.1016/0378-1119(89)90076-0

21. Ohi, M., Li, Y., Cheng, Y., & Walz, T. (2004). Negative staining and image classification — powerful tools in modern electron microscopy. *Biological Procedures Online*, *6*(1), 23–34. https://doi.org/10.1251/bpo70

22. Scheres, S. H. W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

23. Pronker, M. F., Lemstra, S., Snijder, J., Heck, A. J. R., Thies-Weesie, D. M. E., Pasterkamp, R. J., & Janssen, B. J. C. (2016). Structural basis of myelin-associated glycoprotein adhesion and signalling. *Nature Communications*, *7*(1). https://doi.org/10.1038/ncomms13584

24. George, J. M., Jin, H., Woods, W. S., & Clayton, D. F. (1995). Characterization of a novel protein regulated during the critical period for song learning in the zebra finch. *Neuron*, *15*(2), 361–372. https://doi.org/10.1016/0896-6273(95)90040-3

25. Cafferty, W. B. J., & Strittmatter, S. M. (2006). The Nogo-Nogo Receptor Pathway Limits a Spectrum of Adult CNS Axonal Growth. *Journal of Neuroscience*, *26*(47), 12242–12250. https://doi.org/10.1523/jneurosci.3827-06.2006

26. Kempf, A., & Schwab, M. E. (2013). Nogo-A Represses Anatomical and Synaptic Plasticity in the Central Nervous System. *Physiology*, *28*(3), 151–163. https://doi.org/10.1152/physiol.00052.2012

27. Bradl, M., & Lassmann, H. (2009). Oligodendrocytes: biology and pathology. *Acta Neuropathologica*, *119*(1), 37–53. https://doi.org/10.1007/s00401-009-0601-5

CHAPTER 3: CRYO-EM REVEALS CONFORMATIONAL FLEXIBILITY IN APO DNA POLYMERASE ζ

This chapter is a paper published in *Journal of Biological Chemistry*

Du Truong, C., Craig, T. A., Cui, G., Botuyan, M. V., Serkasevich, R. A., Chan, K.-Y., Mer, G., Chiu, P.-L., & Kumar, R. (2021). Cryo-EM reveals conformational flexibility in apo DNA polymerase ζ. *Journal of Biological Chemistry*, *297*(2), 100912. https://doi.org/10.1016/j.jbc.2021.100912

3.1 Introduction

Environmental factors such as chemicals and ultraviolet light as well as metabolic processes cause cellular DNA damage and genomic instability, resulting in DNA lesions that can stall replicative DNA polymerases Polζ or Polζ[1-4]. Eukaryotes and prokaryotes have evolved a mechanism, called translesion synthesis (TLS), that allows the replication machinery to bypass DNA lesions. This process can be mutagenic due to the misincorporation of nucleotides across the lesion site[4-9].

There are three known TLS polymerases in E. coli and 15 in eukaryotes[10]. One of the eukaryotic TLS polymerases, Rev1, serves as a scaffolding protein that recruits other TLS polymerases to replication forks[4,5,11-16]. Rev1 interacts with other TLS polymerases via distinct interfaces in its C-terminal domain[5,12,17]. Rev1 also possesses deoxycytidine monophosphate (dCMP) transferase activity and functions as an insertion TLS polymerase, which incorporates nucleotides (usually deoxycytidine triphosphate or dCTP) opposite damaged and non-damaged guanines[4,5,11,12,18,19].

Among Rev1 partner TLS polymerases, Polζ has been extensively studied both in human and yeast and shown to mediate damage-induced mutagenesis[20-23]. It belongs to the B family of polymerases and is composed of subunits Rev3, Rev7, Pol31 and Pol32[24,25] (Figure 1A). Polζ has lower processivity but higher fidelity than the Y family of polymerases[24-27]. Rev3 is the catalytic subunit of Polζ and can perform its function alone[26]. The accessory subunit Rev7 increases Rev3 activity by at least 20-fold, suggesting an enhancing role in Polζ processivity[26]. Two three-dimensional (3D) structures of S. cerevisiae Polζ were recently determined in the presence of DNA oligomers of different lengths using single-particle cryogenic electron microscopy (cryo-EM)[28]. Although the

cryo-EM density for DNA was not detectable in one of the DNA-Polζ complex structures, the DNA might still influence the structure. Without a true apo Polζ structure, our mechanistic understanding of TLS initiation by Polζ remains incomplete.

In this study, we expressed and purified Polζ from S. cerevisiae and characterized its association with Rev1. We verified that Rev1 binds Polζ with high affinity and functions synergistically with Polζ to extend DNA beyond an abasic lesion. We then used single-particle cryo-EM to visualize the structure of the pentameric apo Polζ holoenzyme comprised of two Rev7 subunits and one subunit each of Rev3, Pol31 and Pol32 arranged around a central canal. By comparison with the previously determined cryo-EM structures of DNA-bound Polζ, we assessed possible conformational changes in Polζ associated with DNA binding to gain mechanistic insights into the initiation phase of Polζ-mediated TLS.

3.2 Materials and Methods

3.2.1 Yeast cells and plasmids

Protease-deficient yeast (*S. cerevisiae*) host PY265 (PY265, mat a, genotype: can 1 his3 leu 2 trp 1 ura 3 pep4::HIS3 GAL nam7Δ::Mx4) and plasmids pBL813-Zeta_opt (for expression of GST-Rev3 and Rev7 under the control of a GAL1-10 promoter with codon-optimized genes) (URA selection), pBL347_p31-32_his (for expression of Pol31 and heptahistidine-Pol32 under the control of a GAL1-10 promotor) (LEU selection), and pBL824-0 Rev1 PPCS (for expression of yeast GST-Rev1 under the control of a GAL1-10 promotor) (URA selection), were gifts from Dr. P.M. Burgers[13] (Washington University, Saint Louis, MO).

3.2.2 Cloning of human Rev7 and SHLD3

The cDNAs of the full-length human Rev7 and various lengths of shieldin-3 (SHLD3) (full-length, 1-29 aa, 28-74 aa, 41-74 aa, 1-83 aa and 28-83 aa) were dually inserted into a pETDuet1-based vector, producing co-expressed Rev7-SHLD3 complexes in which Rev7 has an N-terminal hexahistine tag cleavable by PreScission protease and SHLD3 is untagged. QuikChange (Agilent) was used to introduce a homodimer disrupting R124A single point mutation in Rev7.

3.2.3 Yeast transformation

Plasmids were amplified by chemical transformation of TOP10 cells *E. coli* (Thermo Fisher Life Technologies) with ampicillin selection. Plasmids (pBL813-Zeta_opt, pBL347_p31.32_his and pBL824-0 Rev1 PPCS) were purified using Wizard® Plus DNA purification kits (Promega) and then used for PY265 yeast transformation by electroporation (Becker and Guarente, 1991).

3.2.4 Expression of full-length yeast Polζ

A single colony transformed with pBL813-Zeta_opt and pBL347_p31.32_his plasmids from an SD-Ura-Leu agarose plate (Synthetic Defined (SD) agar plates: 2% glucose, 6.7 g yeast nitrogen base without amino acids, 20 mg each of histidine, arginine, tryptophan, tyrosine, methionine and adenine, 40 mg threonine, 50 mg phenylalanine, 60 mg lysine and 20 g agar in 1 L) was used to inoculate 30 mL of SD-Ura-Leu media (SD components without agar). The culture was grown for 3 days at 30°C and 240 rpm to achieve cell saturation. Once sufficiently dense, a primary SCGL starter culture was created by inoculating 30 mL of SCGL-Ura-Leu media with 500 µL of cells grown in SD media. The SCGL medium contains per liter: 1.7 g of yeast nitrogen base without amino acids and

ammonium sulfate, 5 g ammonium sulfate, 30 mL glycerol, 20 mL lactic acid, 1 g glucose, 20 mg each of adenine, histidine, tryptophan, proline, arginine and methionine, 30 mg each of isoleucine, tyrosine and lysine, 50 mg phenylalanine, and 100 mg each of glutamic acid, aspartic acid, valine, threonine, and serine. Uracil and leucine were omitted to ensure the selective maintenance of plasmids. Prior to autoclaving, the pH of the media was adjusted to 5-6. The primary starter was grown for 2-3 days at 30°C and 240 rpm. Secondary cultures were produced by filling 50 mL conical tubes with 30 mL of SCGL-Ura-Leu media and inoculating each with 1 mL of the primary starter. Secondary cultures were grown for 2-3 days at 30°C and 240 rpm and then used to inoculate SCGL-Ura-Leu media in 2 L baffled Erlenmeyer flasks (50 mL starter per 600 mL media). Growth was continued at 30°C and 240 rpm for 24-26 h to achieve an OD_{660} of 2.8-3.0. At this point, an equal volume of rich YPGLA medium (30 g yeast extract, 60 g peptone, 90 mL glycerol, 60 mL lactic acid and 60 mg adenine per L, pH adjusted to 5-6) was added to each flask. The flasks were gently swirled and then the mixed media distributed equally between two 2 L flasks. Cells were grown for 2-3 h at 30°C and 240 rpm and then induced with 2% solid galactose. Cell growth was continued for 12-16 h. Cells were harvested by centrifugation at 7,000×g for 10 min at 4°C using a swinging bucket rotor in a Lynx 4000 centrifuge. To avoid disruption of the cell pellet, rotor deceleration was set to a low level of 4. Centrifuged cells were washed with ice cold water and centrifuged again. The resulting pellet was transferred to sterile 50 mL conical tubes and immediately frozen using liquid nitrogen for storage at -80°C.

3.2.5 Purification of Polζ

All steps were performed at 4°C unless otherwise stated. Frozen cells in which the four subunits of Polζ were expressed were thawed using room temperature water and then transferred to a bead beater chamber for lysis. To 200 g yeast cells in the chamber, were added 100 mL of lysis buffer (150 mM HEPES (pH 7.8), 900 mM KCl, 90 mM K_2HPO_4/HK_2PO_4, 8% glycerol, 7.5 mM sucrose, 0.15% Tween 20, 0.03% Nonidet P-40, 6 mM DTT, 30 μM pepstatin A, 30 μM leupeptin, 7.5 mM benzamidine and 0.5 mM PMSF) and 150 mL of 0.5 mm glass beads. Cells were lysed by 55 cycles of alternating lysis (30 sec) and cooling (2 min). After lysis, the homogenate was removed from the chamber and transferred to a clean beaker. The glass beads were allowed to settle, the supernatant was collected, and the beads were rinsed three times with lysis buffer to ensure collection of the entire homogenate. Nucleic acids were precipitated from the homogenate by adding 45 mL of 10% Polymin P per L of homogenate. The mixture was stirred for 20 min and then centrifuged at 29,000×g for 60 min to remove cellular debris/insoluble material. The supernatant was collected, and ammonium sulfate was added to a concentration of 0.31 g/mL. The sample was stirred overnight and then centrifuged at 29,000×g for 90 min. The resulting pellet was collected and resuspended in Buffer A1 (50 mM HEPES (pH 7.4), 300 mM KCl, 30 mM K_2HPO_4/KH_2PO_4, 8% glycerol, 2.5 mM sucrose, 0.05% Tween 20, 0.01% Nonidet P-40, 2 mM DTT, 8 μM pepstatin A, 8 μM leupeptin, 2 mM benzamidine and 0.5 mM PMSF) in a total volume of 1 L for 1 h before centrifuging at 29,000×g for 5 h. The resulting soluble material was collected and filtered using a 0.45-μm bottle top filter. The filtered material was passed over a GST-Prep FF 16/10 affinity column (GE Healthcare/Cytiva) at a flow rate of 2 mL/min using a peristaltic pump. For batch chromatography, protein pellet was resuspended with 3 L Buffer A1 for

12-16 h, centrifuged at 29,000×g for 2 h, incubated with 20 mL glutathione-agarose resin (Thermo Fisher Scientific) for 4 h, and then packed into a disposable column. Following application of the supernatant material, the column was washed with 200 mL each of Buffer A2 (30 mM HEPES (pH 7.8), 200 mM KCl, 30 mM K_2HPO_4/KH_2PO_4, 8% glycerol, 2.5 mM sucrose, 0.05% Tween 20, 0.01% Nonidet P-40, 1 mM DTT, 5 mM $MgCl_2$, 1 mM ATP, 2 μM pepstatin A and 0.5 mM PMSF), and Buffer A3 (30 mM HEPES (pH 8.0), 100 mM KCl, 30 mM K_2HPO_4/KH_2PO_4, 8% glycerol, 2.5 mM sucrose, 0.05% Tween 20, 0.01% Nonidet P-40, 1 mM DTT, 2 μM pepstatin A and 0.5 mM PMSF) at a flow rate of 2.5-3.0 mL/min. The Polζ complex was eluted from the resin using Buffer A3 with 50 mM reduced glutathione, at a flow rate of 0.5 mL/min. Three mL fractions were collected and analyzed by SDS-PAGE. The Polζ complex-containing fractions were combined and treated with PreScission protease (Cytiva) to cleave the GST tag from Rev3. The sample was then diluted with equal volume of Buffer E (30 mM HEPES (pH 7.4) 20 mM KCl, 20 mM K_2HPO_4/KH_2PO_4, 5% glycerol, 2.5 mM sucrose, 1 mM DTT and 0.5 mM PMSF) and imidazole was added to a final concentration of 20 mM. The sample was applied to a HisPrep FF 16/10 nickel affinity column (Cytiva) at a flow rate of 3 mL/min using a peristaltic pump. The column was washed with 100 mL of Buffer E followed by 100 mL of Buffer E containing 20 mM imidazole at a flow rate of 3 mL/min. Polζ was eluted from the column with Buffer E containing 200 mM imidazole at a flow rate of 0.5 mL/min. Fractions eluted from the column were analyzed by SDS-PAGE and gels were stained with Coomassie Blue or silver. The four components of the Polζ complex was analyzed by mass spectrometry. When performing bulk chromatography, fractions treated with PreScission protease were diluted 5× with Buffer A3 with 400 mM KCl and 20 mM imidazole. Two

mL of Ni-Sepharose 6 resin (GE Healthcare) were added to the diluted sample, stirred for 2 h, and packed into a disposable column. The resin was washed with 100 mL of Buffer A3 with 400 mM KCl and 20 mM imidazole and the protein were eluted using Buffer A3 with 400 mM KCl and 200 mM imidazole.

3.2.6 Purification of Rev1

Full-length GST-Rev1 was purified from ~180 g of Rev1 expressing yeast cells following the same steps used for the purification of the Polζ complex but without the nickel affinity chromatography step.

3.2.7 Purification of human Rev7 and SHLD3

Various Rev7(R124A)-SHLD3 complexes were co-expressed in BL21(DE3) *E. coli* cells grown at 37°C in LB media to an OD_{600} of ~0.6 and induced with 0.5 mM isopropyl-β-D-thiogalactoside at 15°C for ~16 h. Harvested cells were resuspended in bind buffer, lyzed with a microfluidizer (Avestin Emulsiflex C5), and centrifuged. The resulting supernatant was loaded onto a column with Ni^{2+}-NTA agarose (Qiagen). After extensively washing the column with wash buffer, the complex was eluted with elution buffer. The bind buffer is made up of 50 mM sodium phosphate (pH 7.5) and 300 mM NaCl, while the wash and elution buffers have the bind buffer components with additional 20 and 250 mM imidazole, respectively. The hexahistidine tag on Rev7(R124A) was removed by addition of PreScission protease at 4°C overnight. The complex was further purified by size exclusion chromatography using a HiLoad 16/60 Superdex 75 column (GE Healthcare) and 5 mM HEPES (pH 7.4) and 100 mM NaCl as running buffer. From protein expression results, a

minimal complex of Rev7(R124A) and SHLD3 (41-74 aa) could be formed. This complex was used for structure determination.

3.2.8 Mass spectrometry of Polζ

Gel bands obtained from SDS-PAGE gels were subjected to in-gel trypsin digestion after reduction and carboxymethylation, and the treated and extracted peptides were analyzed by nano-ESI-LC/MS/MS with a Q Exactive mass spectrometer coupled to a Dionex nano-LC system (Thermo Fisher Scientific). The LC system used multi-step linear gradients with solvents A (2% acetonitrile, 0.2% formic acid, in water) and B (80% acetonitrile, 10% isopropyl alcohol, 0.2% formic acid, in water) as follows: 4-5 min, at 5% B; 5- 35 min 5-45 % B ; 35-38 min 45-95% B; 38-42 min 95% B; 42-44 min 95% A-10% B; 44-47 min 10% B; 47-55 min 10-95% B; 55-58 min 95% B; 58-61 min 95-5% B; 61-67 min 5% B. The mass spectrometer had a resolution of 70,000 (at 200 m/z) and used data dependent acquisition, with a full MS1 scan ranging from 350-1800 m/z, then selecting the top 15 ions for MS2 analysis with a dynamic range set to 8 sec. All MS/MS spectra were analyzed using Mascot (version 2.4; Matrix Science), and X! Tandem (www.thegpm.org; version 2013.09.01 is provided in the public domain by the Global Proteome Machine Organization, Manitoba Centre for Proteomics and Systems Biology, Winnipeg, Canada). Each software was set up to search the current SwissProt database, assuming trypsin digestion with up to two miscleavages with a fragment ion tolerance of 10.0 PPM (www.uniprot.org; SwissProt). Oxidation of methionine was set as a variable modification, and carbamidomethylation of cysteine (iodoacetamide derivative) was set as a fixed modification. Proteomics software (Scaffold, ver. 4.11.0; Proteome Software Inc.) was

used to view MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at > 95.0% probability, as specified by the peptide prophet algorithm. Protein identifications were accepted if they could be established at > 95% probability and contain at least two unique peptides. Protein probabilities were assigned by the protein prophet algorithm.

3.2.9 Database searching for Polζ

Tandem mass spectra were extracted using ProteoWizard MsConvert. Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.4.0) and X! Tandem (The GPM, thegpm.org; version X! Tandem Sledgehammer (2013.09.01.1)). Mascot was set up to search the *S. cerevisiae* Swissprot database (downloaded in April 2019, 16,060 entries), assuming the digestion enzyme stricttrypsin. X! Tandem was set up to search the *S. cerevisiae* Swissprot database (downloaded in April 2019, 16,060 entries) also assuming stricttrypsin. Mascot and X! Tandem were searched with a fragment ion mass tolerance of 0.020 Da and a parent ion tolerance of 10.0 ppm. Carbamidomethyl of cysteine was specified in Mascot and X! Tandem as a fixed modification. Glu->pyro-Glu of the N-terminus, ammonia-loss of the N-terminus, Gln->pyro-Glu of the N-terminus and oxidation of methionine were specified in X! Tandem as variable modifications. Oxidation of methionine and acetyl of the N-terminus were specified in Mascot as variable modifications.

3.2.10 Criteria for Polζ identification

Scaffold (version Scaffold_4.11.0, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were

accepted if they could be established at ≥ 95.0% probability by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at ≥ 95.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm[54]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters.

3.2.11 Database searching for Rev1

Tandem mass spectra were extracted using ProteoWizard MsConvert. Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using Mascot (Matrix Science, London UK; version 2.4.0). Mascot was set up to search the Swissprot *S. cerevisiae* database (downloaded in August 2019, 13,582 entries), assuming the digestion enzyme stricttrypsin. Mascot was set up to search with a fragment ion mass tolerance of 0.020 Da and a parent ion tolerance of 10.0 ppm. Carbamidomethyl of cysteine was specified in Mascot as a fixed modification. Oxidation of methionine was specified in Mascot as a variable modification.

3.2.12 Criteria for Rev1 identification

Scaffold (version Scaffold_4.11.0, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at ≥ 95.0% probability. Peptide Probabilities from Mascot (samples (qe1_2020feb21_P20025_upperband) and (qe1_2020feb21_P20025_lowerband)) were assigned by the Scaffold Local FDR algorithm. Peptide Probabilities from Mascot (samples (qe1_2020feb21_P20025_postrunblank) and (qe1_2020feb21_P20025_prerunblank))

were assigned by the Peptide Prophet algorithm[55] with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at ≥ 95.0% probability and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm[54]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters.

3.2.13 Biolayer interferometry and K_D determination

Biolayer interferometry assays were performed on a BLItz biolayer interferometry instrument (ForteBio). All measurements were done at 22ºC in basic kinetic mode and consisted of three main steps. An initial base line reading (30 sec) using 1× kinetic buffer (10× kinetic buffer, #18-1105 Pall ForteBio was diluted to 1× with PBS, #10010-23 Gibco) was performed. This was followed by a step in which association of proteins occurred over 600 sec. A dissociation step of 120 sec was then performed. Prior to use, Ni^{+2}-NTA sensor tips (Pall ForteBio) were equilibrated in 1× kinetic buffer for at least 20 min. Yeast Polζ (19 µg/mL, 328 kDa) (Rev3 cleaved of GST-tag, heptahistidine--tagged Rev7, Pol31 and Pol32) in Buffer E with 50 mM glutathione was diluted with 1× kinetic buffer to 9.5 µg/mL for association with Ni^{+2}-NTA tips on the Blitz. Yeast Rev1 (142 µg/mL, 112 kDa) in Buffer A3 with 50 mM glutathione was diluted 30-fold with 1× kinetic buffer to 4.7 µg Rev1/mL and likewise applied to Ni^{+2}-NTA tips.

3.2.14 Translesion DNA polymerase assay

Extension of a 32-base DNA oligonucleotide primer annealed to a 52-base DNA oligonucleotide containing an abasic site (tetrahydrofuran covalent linked dSpacer in

oligonucleotide sugar backbone, without a purine or pyrimidine base) were prepared to test translesion extension ability of Polζ and Rev1. Below are sequences of the synthetic oligonucleotides (GENEWIZ).

32-mer:

5'-GTTTTCCCAGTCACGACGATGCTCCGGTACTC-3'

52-mer:

5'-

TTCGTATAATGCCTACACT*GAGTACCGGAGCATCGTCGTGACTGGGAAAAC-3'

(* = abasic site: tetrahydrofuran without nucleotide base)

The 32-base primer was first end labeled with ^{32}P using gamma ^{32}P ATP (adenosine 5'-triphosphate, #BLU-002H Perkin Elmer) and T4 polynucleotide kinase (New England Biolabs) at 37°C in T4 polynucleotide kinase buffer (70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$ and 5 mM DTT). After labeling, sample was passed through a Micro Bio-Spin P6 spin column (Bio-Rad) pre-equilibrated with annealing buffer (10 mM Tris-HCl (pH 7.6), 50 mM NaCl and 1 mM EDTA) to remove excess ^{32}P ATP. The ^{32}P labeled 32-base primer and the unlabeled 52-base template were annealed by putting the tube with the mixture in a 500 mL beaker of boiling water and letting the water cool to room temperature for ~4 h. The annealed DNA was stored at -20°C until use.

Ten μL solution of 100 fmol ^{32}P-labeled annealed DNA in 1× translesion DNA extension buffer (25 mM KH_2PO4 (pH 7.0), 5 mM $MgCl_2$, 5 mM DTT, 100 μg/mL BSA and 10% glycerol)[56] were added to each of the mixtures below.

To test TLS DNA polymerase activity of Polζ, mixtures containing 53, 131 or 526 fmol of Polζ in 20 μL Buffer A3 with 25 mM glutathione were prepared. To examine enhancement of Polζ activity by Rev1, a similar 20 μL mixture containing 53 fmol of Polζ and 12 fmol of GST-Rev1 was prepared. Mixtures containing only 12, 30 and 121 fmol of Rev1 in 20 μL were likewise set up to examine the polymerase activity of Rev1. After combining DNA and protein mixtures on ice, DNA synthesis (translesion and extension) was initiated by addition of 0.1 mM deoxynucleotide triphosphates (dNTPs) and incubation at 30°C. The reactions were stopped after 30 min by addition of 10 μL 1× Novex® Hi-Density TBE sample buffer (Thermo Fisher Scientific). Reaction samples were stored at -20°C or combined with equal volume of 80% formamide, 1× TBE. Samples from the latter were heated to 95°C for 6 min, loaded onto 1× TBE 7 M urea 10% acrylamide gels (Thermo Fisher Scientific) pre-warmed to 45°C and with buffer pre-warmed to 50°C, and run at 180 volts constant voltage until loading dye was near bottom of gel (~20 bp). ^{32}P-labeled 32-mer primer and ^{32}P-labeled 52-mer template (150 fmol each in 1× NOVEX Hi-Density TBE sample buffer), prepared similarly as the samples, were also run on gels to serve as molecular weight standards. Undried gels covered with plastic wrap were exposed to X-ray films for various lengths of time (h-days) and developed.

3.2.15 X-ray crystallography of human Rev7 and SHLD3

Crystals of the full-length Rev7(R124A)-SHLD3 (41-74 aa) were obtained by the hanging drop method, putting 1 μL of the protein sample (25 mg/mL in 5 mM HEPES (pH 7.4), 100 mM NaCl and 5 mM DTT) and 1 μL of the reservoir solution for the drop and 0.5 mL for the reservoir solution (0.1 M MES monohydrate (pH 6.5) and 1.4 M $MgSO_4 \cdot 6H_2O$) in the well. Crystals formed within 2 to 4 weeks at 15ºC. The crystals were cryoprotected in 50% PEG400 and quick-frozen in a cryoloop with liquid nitrogen. Diffraction data were collected at the 19-BM beamline at the Advanced Photon Source, Argonne National Laboratory. Diffraction patterns were indexed, integrated, and scaled with HKL2000[57]. The initial phases were obtained by molecular replacement using the coordinates of the Rev7-Rev3 structure (PDB code: 3ABD) as a search model in Phenix[36,58]. The starting model was completed and refined in Coot[59] and Phenix (58) in an iterative manner. The crystals of Rev7 (R124A)-SHLD3 (41-74 aa) complex have a $P3_221$ space group, with one molecule of Rev7(R124A) and one molecule of SHLD3 (41–74 aa) in the complex. One copy of the complex molecule is found in the asymmetric unit. Statistics of the final structure are shown in Table 2. All molecular representations were generated with PyMOL (60) and UCSF Chimera[61].

3.2.16 Negative-stain electron microscopy

Negatively stained samples of 0.01 mg/mL apo Polζ complex was prepared using 0.75% uranyl formate and followed by the previous method[62]. The stained samples were imaged using a Tecnai TF20 TEM at an accelerating voltage of 200 keV with a CCD camera recording at a pixel size of 1.4 Å/pixel at the specimen level. Thirty-nine electron images were collected and imported into Relion (version 3.1-beta-commit-ca101f)[63] for general

image processing. A total of 2,330 particles were manually selected from the electron images, and the 2D class averages with an assigned *k* of 50 were calculated, respectively.

3.2.17 Single-particle cryo-EM data collection

The Polζ complex sample was loaded onto a Superose 6 column (GE Healthcare) for size-exclusion chromatography. The purified peak fraction was used for further cryo-EM imaging. A C-flat 400-mesh holey-carbon coated copper grid (2/1 4C; Protochips) was glow-discharged for 15 sec in a Pelco easiGlow glow-discharge system (Ted Pella). Five μL of 0.1 mg/mL protein sample were then applied on the EM grid. The grid was blotted by a filter paper to remove the excess solution and quickly frozen into liquid ethane. The plunge freezing process was automated using a Vitrobot Mark IV plunge freezer (Thermo Fisher/FEI) at a humidity of 100% with a blotting time of 6 sec.

All the cryo-EM data collections were completed in the Eyring Materials Center (EMC) at Arizona State University (ASU). The grid specimen was imaged using a Thermo Fisher/FEI Titan Krios TEM (Thermo Fisher/FEI) at an accelerating voltage of 300 keV. The electron scattering was recorded by a Gatan Summit K2 DDD camera in super-resolution mode[64]. The nominal magnification was set to 48,780X, corresponding to a pixel size of 1.025 Å/pixel at the specimen level. The defocus was set to vary from -0.8 to -3.0 μm. The camera counted rate was calibrated to 8 e^-/pixel/sec. The exposure time was 6 sec, accumulating to a total dosage of 45.7 $e^-/Å^2$. The dataset was collected in counting mode. The beam-image shift scheme was applied to accelerate the data collection[65]. The procedure of low-dose imaging was automated using SerialEM software (version 3.9) with customized macros[66].

3.2.18 Image processing

Image processing was generally conducted using cryoSPARC (version 3.0)[67]. A total of 11,698 cryo-EM movies was imported into the program for processing. The frame registration and averaging for motion correction were performed using the 'Patch motion correction' and the estimation of the defocus was performed using the 'Patch CTF estimation'. An ensemble of 2,974,553 particles was automatically selected using a neural network and positive-unlabeled learning by the Topaz program (version 0.2.3)[68]. The curation of the particle images was performed using iterative 2D image classification. A total of 1,658,585 particles was selected for an *ab initio* 3D map generation[67]. The two generated volumes were refined further against their individual particle subsets and only one volume showed discernible structural features of a protein. The subgroup of its 3D reconstruction with discernible protein features was carried over for homogeneous refinement and subsequent image processing. A total of 213,120 particle images was selected for further processing. The 3D map was then refined using homogenous and non-uniform refinement procedures in cryoSPARC[67,69]. The final map resolution reached 4.11 Å, estimated using the golden standard Fourier-shell correlation (FSC) method at the cutoff of 0.143[29]. The local resolution was assessed using an FSC-windowed method[70]. The directional FSC of the reconstruction was assessed using 3DFSC program wrapped in cryoSPARC[71].

Further signal subtraction and focused classification were performed to improve the quality of the local densities for Rev3 and Rev7-Pol31-Pol32 using cryoSPARC software[67]. Masks were generated using Segger implemented in UCSF Chimera[72]. The local refinement was focused on the region of interest and the remaining densities were subtracted[30]. The

resolutions of the local densities of the Rev3 and Rev7-Pol31-Pol32 were 3.65 Å and 3.72 Å, respectively. The two improved maps were then combined using the 'phenix.combine_focused_maps' program in Phenix software (version 1.18.2-3874) for subsequent modeling[58].

3.2.19 Molecular modeling

Previous atomic coordinates of the Polζ-DNA complex (for which DNA density was not detected) (PDB code: 6V8P)[28] were used as starting template. These initial coordinates were first docked into the cryo-EM density using the 'Fit in the Map' function in UCSF Chimera software (version 1.14)[61]. The fitted coordinates were manually rebuilt and adjusted using Coot (version 0.9-pre)[59,73]. The rebuilt coordinates were refined against the cryo-EM density using the 'phenix.real_space_refine' program in Phenix software package (version 1.18.2-3874)[58]. The molecular graphic presentation for the final model was made using UCSF Chimera or UCSF ChimeraX (version 0.91)[74].

The structures of the DNA-bound (PDB code: 6V8P and 6V93)[28] and apo Polζ were used to calculate a morph movie by using the Needleman-Wunsch algorithm in UCSF Chimera[61]. The generated movie is shown in supplementary material Movie S1.

3.3 Results and Discussion

3.3.1 Characterization of the Polζ complex and its interaction with Rev1

We purified the Polζ complex from *S. cerevisiae* using GST-affinity and metal-chelation chromatography. For a typical purification, we processed 0.5 to 2.5 kg of yeast cells, yielding approximately 250 μg of purified Polζ per kg of cells. SDS-PAGE of the purified protein showed four bands at 175, 55, 49 and 28 kDa (Figure S1A). Mass spectrometric

analysis indicated that the 175 kDa band was the Rev3 catalytic subunit; the 55 kDa band was the Pol31 subunit; the 49 kDa band was the Pol32 subunit; and the 28 kDa band was the Rev7 processivity subunit (Figure S1B-C). The quality of the protein complex was assessed using negative-stain electron microscopy (EM) (Figure S1D-E). The EM images showed a stable and homogeneous protein complex, and the two-dimensional (2D) class averages showed clear features of Polζ in different views. Thus, the apo Polζ protein complex can be stably formed in the absence of DNA oligomers.

Rev1 was purified using GST-affinity chromatography. SDS-PAGE showed a major band at 115 kDa and a minor band at 100 kDa (Figure S2A-C). Analysis of both bands by mass spectrometry produced a sequence compatible with full-length Rev1 with 90-95% coverage, suggesting that the difference in gel mobility was due to post-translational modifications (Figure S2A-C).

We tested the binding of Rev1 to Polζ using nickel-affinity chromatography and showed that the two proteins co-purified and formed a tight complex (Figure 3.1B). The Pol32 subunit of Polζ has a heptahistidine-tag. Using biolayer interferometry (BLI), we measured a dissociation constant (K_D) of 0.11 ± 0.12 μM (mean ± s.d. of n = 4 independent experiments) for the Rev1-Polζ complex.

Next, we tested the capacity of Polζ and Rev1 to extend DNA beyond an abasic lesion. In the absence of Rev1, low concentration of Polζ (2.65 nM) could not extend DNA but when a small amount of Rev1 (0.6 nM, final concentration) was added, DNA extension was detected (Figure 1C). Rev1 by itself at low concentrations (0.6 and 1.5 nM) could not extend DNA (Figure 3.1C). Therefore, Polζ and Rev1 polymerases have a synergistic

effect on DNA extension. However, at higher concentrations, Rev1 (6 nM) and Polζ (6.6 and 26.3 nM) are each capable of extending DNA without the other polymerase. At elevated concentrations, the two enzymes might have redundant functions or synergize with co-purified endogenous yeast enzymes.

3.3.2 Visualization of the apo Polζ using single-particle cryo-EM

We used single-particle cryo-EM to study the structural organization of bioactive Polζ that included all five subunits (Figure 3.1A). We obtained a consensus 3D cryo-EM reconstruction at 4.11 Å resolution, determined by the golden standard Fourier-shell correlation (FSC) method at a cutoff of 0.143[29] (Figure S3A-B). Local resolution analysis of the 3D density showed an anisotropic resolution distribution, implying flexibility for the apo Polζ enzyme (Figure S3C). To improve the quality of the local densities, we performed signal subtraction and focused refinement procedures on the Rev3 and on Rev7-Pol31-Pol32 subunits, separately[30]. We then generated cryo-EM densities of the two separate systems at higher resolutions (3.65 Å for the Rev3 density and 3.72 Å for the Rev7-Pol31-Pol32 density) (Figure S4). The atomic coordinates for Rev3, Rev7, Pol31 and Pol32 we modeled into the cryo-EM density maps (see Methods) have well-defined rotameric side chain conformations (Figure 3.1D and S5). We can also identify the cryo-EM density of the [4Fe-4S] cluster in Rev3, which is essential to the catalytic activity of Rev3[31]. The cryo-EM structure determination statistics are summarized in Table 1.

The five subunits of apo Polζ organize into a ring-like structure (Figure 3.1D). This arrangement of subunits is the same as that in previous Polζ structures determined in complex with DNA oligomers of different lengths (PDB codes: 6V8P and 6V93)[28]. The subunits Pol31 and Pol32 bind the C-terminal domain of Rev3. In particular, the interaction

of Pol31 is stabilized by the iron-sulfur cluster (4Fe-4S) in Rev3[23]. Different from the aforementioned Polζ cryo-EM structures[28], our structure is purely the apo form of Polζ (Figure S6A). Superposition of the apo and DNA-bound Polζ structures show a concerted rigid-body movement of several Polζ regions associated with DNA binding (Figure S6A).

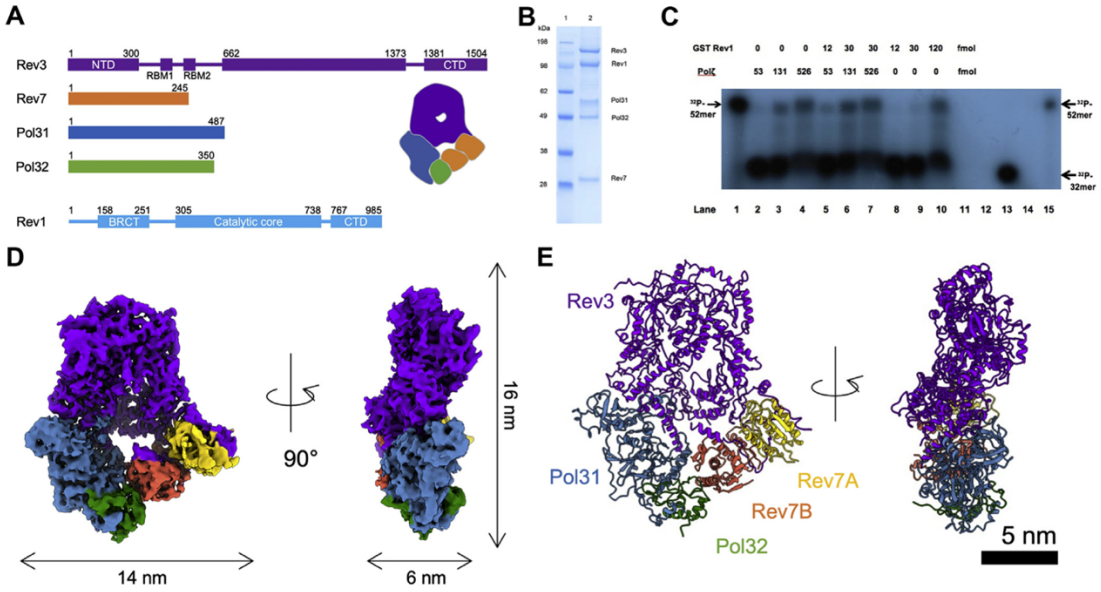

Figure 3.1. Analysis of the interaction between yeast Polζ and Rev1 TLS polymerases and cryo-EM reconstruction of apo Polζ. (A) schematic of the structural organization of yeast Polζ subunits and Rev1. Rev3, Rev7, Pol31, Pol32, and Rev1 are shown in purple, orange, steel blue, and green, respectively. NTD and CTD stand for N-terminal and C-terminal domains, respectively. RBM1 (517–540 aa) and RBM2 (599–623 aa) are Rev7-binding motifs 1 and 2. BRCT stands for BRCA1 C-terminal domain. (B) SDS-PAGE analysis of the coexpression of the Polζ and Rev1 TLS polymerases in yeast. The gel was stained with Coomassie blue. (C) Polζ, Rev1, or a mixture of the two polymerases was used to extend a ^{32}P oligonucleotide 32-mer primer annealed to a 52-mer template strand with an abasic site. Low amounts of either polymerase (lanes 2 and 8) did not extend the primer, but Polζ combined with Rev1 allowed extension past the abasic site (lanes 5 and 6). (D) cryo-EM density of the apo Polζ enzyme complex. E, atomic model of the apo Polζ enzyme complex. The coordinates of Rev3 (purple), Rev7 (gold (Rev7A), and orange (Rev7B)), Pol31 (steel blue) and Pol32 (green) were built along the determined cryo-EM densities. Scale bar indicates 5 nm.

The Pol31 and Pol32 subunits are also essential to Polδ polymerase, and the spatial arrangement of these subunits in apo Polζ is the same as that in Polδ[32,33]. The difference

between Polδ and Polζ is the addition of two Rev7 subunits in Polζ interfacing with Rev3, Pol31 and Pol32 (Figure 3.1D). Rev3 interacts directly with Pol31 and Rev7 but does not contact Pol32 (Figure 3.1D).

3.3.3 Conformational changes in Polζ upon DNA binding

The ring structure of Polζ harbors a narrow central channel for oligonucleotide binding (Figure 3.1D, 3.1E and 3.2). Upon DNA binding, concerted movements of three Rev3 loops (948-960 aa, 1050-1094 aa and 1095-1107 aa) increase the size of the DNA-binding channel opening, from a closed to an open state (Figure 3.2A-B). The loop (1050-1094 aa) near the central channel experiences an up-to-down movement upon DNA binding (Figure 3.2B and S6C) while adjacent flexible loop region (1326-1344 aa) folds into a short α-helix upon DNA binding (Figure S4C). Such changes could be a result of DNA insertion and Polζ processivity (Figure 3.2B and S6C-D). The channel opening sizes of apo Polζ, and Polζ states A and B are 18.2 Å, 20.6 Å and 20.7 Å, respectively (Figure 3.2A). State A represents the short DNA oligomer-bound Polζ (PDB: 6V8P) and state B represents the longer DNA oligomer-bound Polζ (PDB: 6V93). Thus, the channel of Polζ is closed in the absence of DNA. It is possible that the presence of a short DNA oligomer induces structural changes in the Rev3 loops at residues 948-960 and 1095-1107 (Figure 3.2B) and allows the DNA oligomer to initiate a contact with the Rev3 active site, opening up the central channel for DNA processing in translesion synthesis. The channel size change seems to be solely linked to local structural variations of the Rev3 loops, independent of other Polζ subunits. A previous study showed that Rev3 alone has catalytic activity without binding any other subunits[26]. The initiation phase of DNA translesion synthesis may therefore only

require Rev3 and DNA. It is possible that other subunits in the complex, that is, Rev7, Pol31 and Pol32, play a role in regulating the Rev3 activity or interacting with other proteins. To further illustrate the movements associated with the DNA-bound and unbound states of Polζ, we morphed the three cryo-EM structures of apo Polζ and DNA-bound Polζ states A and B together and generated a movie that highlights concerted movements of the local regions (Movie S1).

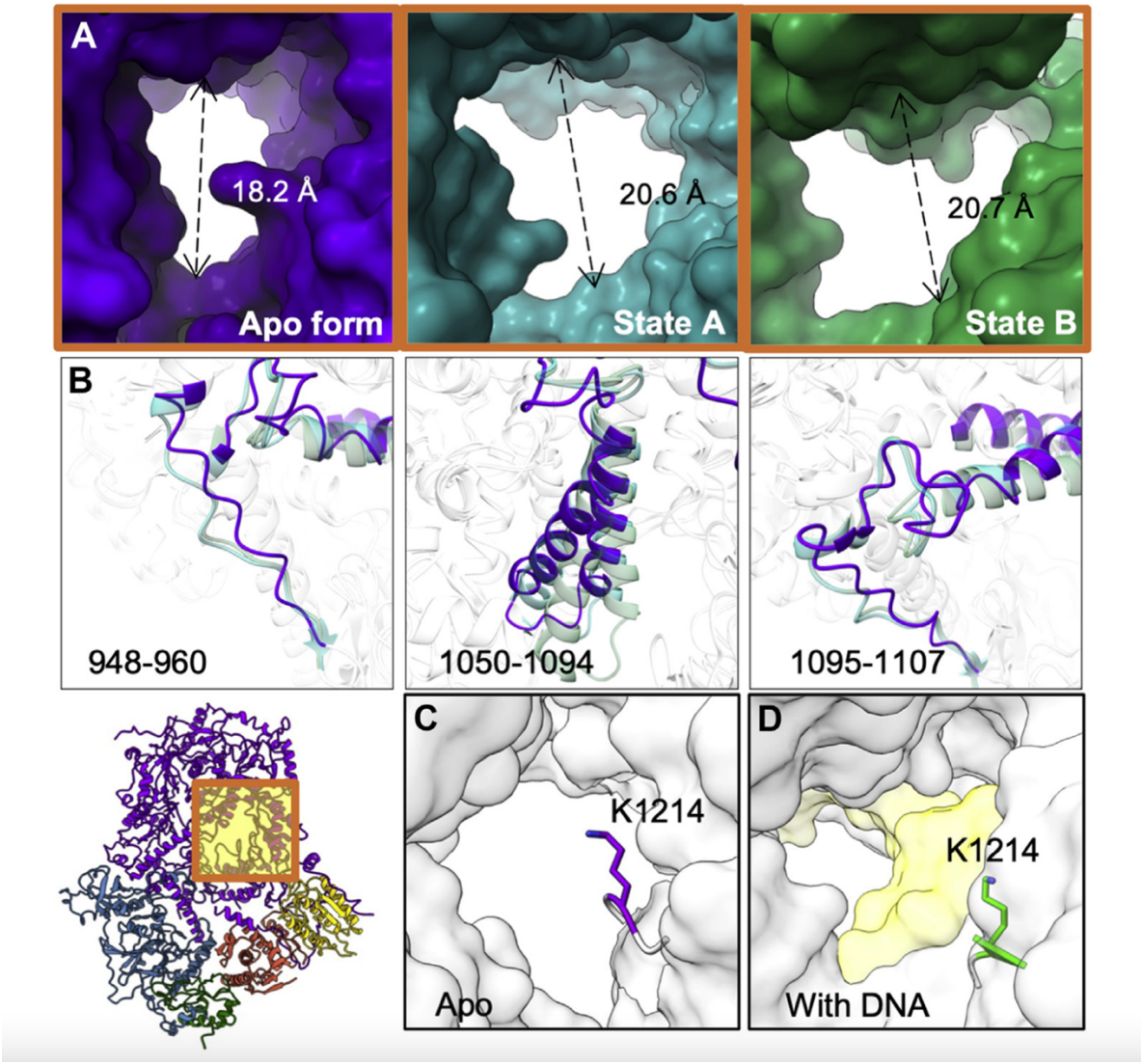

Figure 3.2. Central channel of the Polζ enzyme complex and conformational changes associated with DNA binding. (A) central channel of Rev3 for DNA binding. Surfaces of the apo form and states A and B are shown in purple, cyan and green, respectively. The

sizes of the channel openings were measured as 18.2 Å, 20.6 Å, and 20.7 Å for the apo form and states A and B, respectively. (B) conformational changes of Rev3 loops (948–960 aa, 1050–1094 aa and 1095–1107 aa) upon DNA oligomer binding. The indicated loops of the apo form and DNA-bound Polζ are highlighted in purple and sea green, respectively. (C) rotameric conformation of Rev3 K1214. The side chain of K1214 primary amine points toward the center of the hole in the apo form, whereas it moves away when the Polζ complex binds DNA. (D) side chain of K1214 in DNA-bound Polζ. DNA backbone is colored yellow. The positively charged K1214 side chain points away from the negatively charged sugar-phosphate backbone.

Our apo structure shows that one lysine residue (K1214) of Rev3 protrudes up and points toward the pore where nucleotide insertion occurs (Figure 3.2C). This lysine is conserved in TLS polymerases across different species (Figure S7). When a DNA oligomer is present, K1214 moves away from the pore and leaves space for DNA binding. The positively charged K1214 (PDB: 6V93; state B) is close to but does not interact with the negatively charged DNA polar sugar-phosphate backbone in the minor groove (Figure 3.2D)[28]. This may suggest less friction for DNA binding and Polζ processivity due to removal of electrostatic contacts between K1214 and DNA. Because K1214 does not strongly interact with neighbor residues, it is possible that different conformations of K1214 co-exist, and that single-particle cryo-EM only captured one of the rotameric states. Thus, K1214 may play a role in gating the oligonucleotide passage.

3.3.4 Dimerization of Rev7 subunits in the apo Polζ complex

Polζ has two identical Rev7 accessory subunits, which are important for promoting Polζ catalytic activity and interacting with other proteins that regulate the activity of Polζ[26, 34-38]. The Rev7 homodimer is essential for the stability and function of the Polζ complex[39,40]. Rev7 directly binds Rev1, and this interaction is important for TLS regulation[39,41,42]. In yeast, Rev7 also stimulates the catalytic activity of Rev3[26]. Within the Polζ complex, the two subunits of Rev7, Rev7A and Rev7B, bind the Rev7-binding motifs 1 (RBM1; 17-540

aa) and 2 (RBM2; 599-623 aa) of Rev3, respectively (Figure 3.3A). Additionally, Rev7B bridges Pol31 to Pol32 (6) (Figure 3.3A). These interactions bring together the accessory subunits Rev7, Pol31 and Pol32 and the catalytic subunit Rev3.

Rev7 is a member of the HORMA (Hop1, Rev7 and Mad2 proteins) domain protein family[43]. HORMA proteins share a common core structure composed of three β-strands flanked by three α-helices and a "safety-belt" region, which closes around target proteins. The safety-belt motif also mediates the assembly and disassembly of the HORMA domain protein dimer[44,45]. HORMA domains can adopt an open and a close state[40]. The closed state has two β-strands in the safety belt moving away from its core and wrapping around the HORMA domain, allowing HORMA proteins to bind their partners (Figure 3.3C and 3.3F). Our apo Polζ structure shows that the safety belts of Rev7A and Rev7B adopt a closed conformation in their $Rev3^{RBM1}$- and $Rev3^{RBM2}$-bound states, respectively (Figure 3.3F). The yeast Rev7 dimer also shows a head-to-tail arrangement when bound to the RBM1 and RBM2 motifs of Rev3 (Figure 3.3A-B).

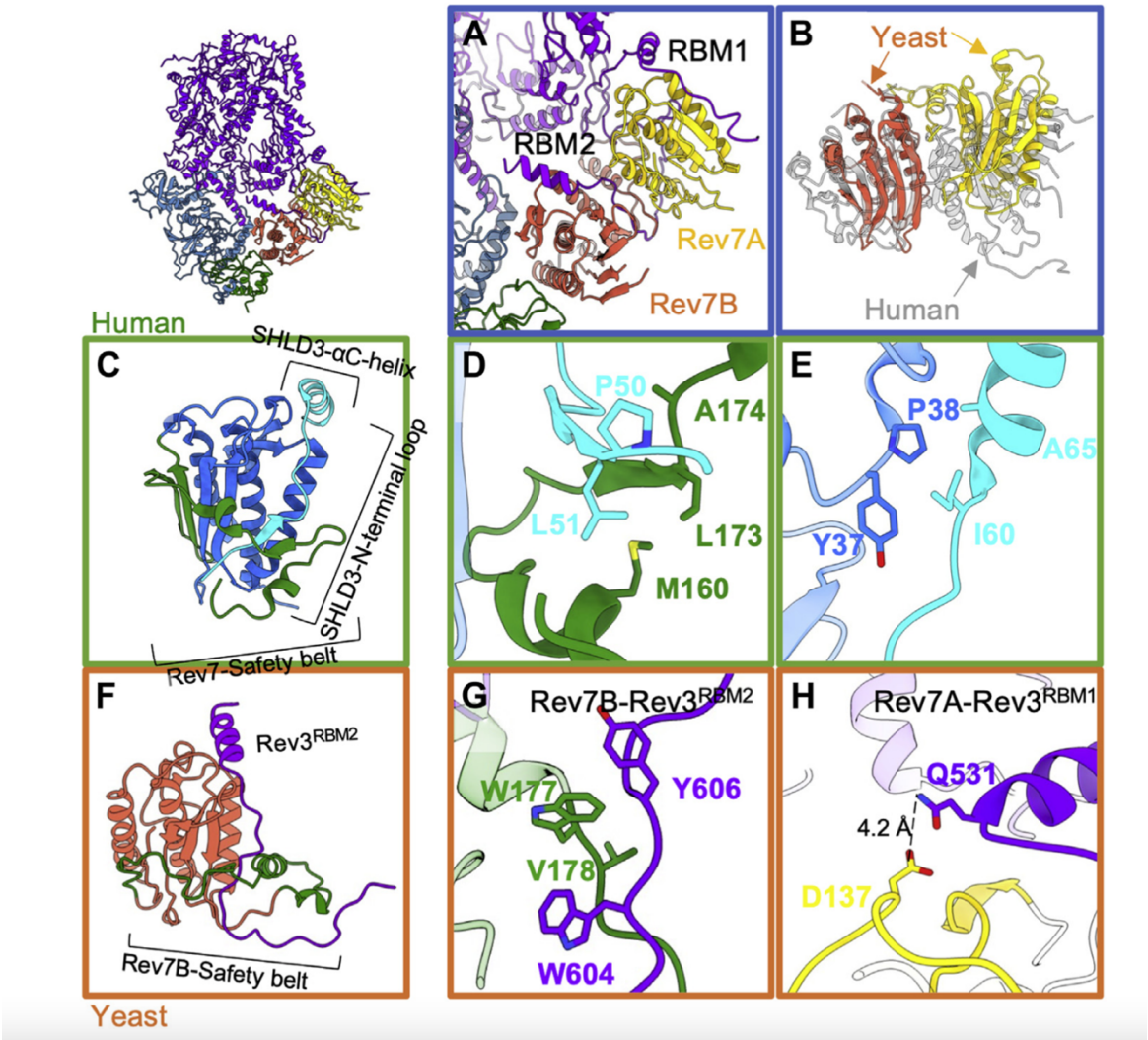

Figure 3.3. Comparison of yeast and human Rev7 dimerization states and binding modes. (A) Rev7 dimer in the yeast apo Polζ complex. Rev7A (yellow) and Rev7B (orange) bind to Rev7-binding motif 1 (RBM1) and 2 (RBM2) of Rev3 (purple), respectively. The RBM2 motif is sandwiched between the two Rev7 protomers, limiting its structural mobility. The RBM1 motif is more accessible than RBM2 suggesting less spatial restriction. (B) overlay of yeast and human Rev7 dimers with yeast Rev7B (orange) and one of the human Rev7 protomers (gray) oriented in the same manner. This comparison highlights the radically different orientations of the two protomers in yeast (orange and yellow) and human (gray) Rev7. (C) crystal structure of human Rev7(R124A) in complex with SHLD3 (41–74 aa). Rev7 core, safety belt region, and SHLD3 are shown in blue, green and cyan, respectively. (D) safety belt region of human Rev7 (M160, L173 and A174) interacts with the N-terminal loop of SHLD3 (P50 and L51). (E) interaction between human Rev7 (Y37 and P38) and SHLD3 (I60 and A65). (F) structure of yeast Rev7B-Rev3^{RBM2} within the apo Polζ complex. Rev7B, safety belt region, and Rev3^{RBM2} are shown in orange, green and purple, respectively. (G) interaction between the safety belt of yeast Rev7B (W177 and V178) and

Rev3^{RBM2} motif (W604 and Y606). H, interaction between yeast Rev7A (D137) and Rev3^{RBM1} (Q531). Rev7A is colored yellow.

3.3.5 Binding modes of yeast and human Rev7

Budding yeast and human Rev7 only share 27% amino acid sequence identity (Figure S8A) but comparison of their 3D structures shows similar folds. The structure of human Polζ is still unavailable, but we examined the crystal structures of human Rev7 in complex with fragments of shieldin-3 (SHLD3) (46-48), including that with a SHLD3 peptide (41-74 aa), which we determined to a resolution of 2.0 Å (Table 2) to compare the binding modes of yeast and human Rev7. SHLD3 and Rev7 are two of the four components of the protein complex shieldin which plays a key regulatory role in DNA double-strand break repair by blocking DNA end resection necessary for homologous recombination (49-51). The small backbone RMSDs with respect to Rev7A-Rev3^{RBM1} (RMSD 1.20 Å) and Rev7B-Rev3^{RBM2} (RMSD 1.08 Å) indicate that the yeast and human Rev7 folds are very similar. In human just like in yeast, Rev7 forms a homodimer, but the orientations of the two Rev7 protomers differ radically in the two systems (Figure 3.3A-B). In addition, the safety belt region of yeast Rev7 is longer than that of human Rev7. In yeast Rev7, the safety belt presents a short helix-turn-helix motif, but the one in human Rev7 is a short helix connected to a loop (Figure S8B). Since the two yeast Rev7 molecules have different organizations in the Polζ complex, we compared each to our human Rev7-SHLD3 crystal structure, which was determined in a Rev7 monomeric state using the R124A mutation (Figure 3.3C). We superimposed the Rev7 structure in human Rev7(R124A)-SHLD3 to the Rev7 structures in yeast Rev7A-Rev3^{RBM1} and Rev7B-Rev3^{RBM2}. The binding modes are very different (Figure 3.3). Rev7B W177 and V178 interact with W604 and Y606 of Rev3^{RBM2} through

hydrophobic interactions or π-stacking contacts (Figure 3.3G), while Rev7A N137 forms a weak hydrogen bond with Q531 of Rev3^{RBM1} (Figure 3.3H). In the crystal structure of Rev7(R124A)-SHLD3, Rev7 binds the SHLD3 peptide at two sites. In one site, SHLD3 N-terminal loop residues P50 and L51 interact mainly via hydrophobic interactions with Rev7 residues A174, M160 and L173 in the safety belt region of Rev7 (Figure 3.3D). In another hydrophobic interface, residues I60 and A65 in the C-terminal α-helix (αC-helix) of SHLD3 interact with Rev7 Y37 and P38 (Figure 3.3E). As can be seen from side-by-side comparison (Figure 3.3B, 3.3C, 3.3F and S8A), yeast and human Rev7 show little similarity in target recognition. Therefore, one may not reliably infer from the yeast Polζ structure how human Rev7 binds other components of the human Polζ complex.

3.4 Conclusions

In this paper, we have determined the structure of the apo form of budding yeast DNA polymerase Polζ, revealing a closed conformation in the absence of DNA. DNA binding to Polζ induces a concerted movement of local structural motifs associated with the opening of the central DNA-binding channel of the polymerase. Moreover, we identified a lysine (K1214) in Rev3 as a putative gating residue for DNA binding. In our apo Polζ structure, the side chain of K1214 protrudes out and obstructs the DNA-binding channel. In the presence of DNA, the side chain of K1214 moves away from the central channel and creates a wider space that allows insertion of incoming DNA. Future molecular dynamics simulations will incorporate this new knowledge to further explore the mechanism of action of yeast Polζ.

The Rev7 subunit of Polζ is an evolutionarily conserved hub protein found in many different protein complexes. For example, human Rev7 is not only a component of Polζ but is also an essential subunit of DNA double-strand break repair protein complex shieldin. Since there is no structure of human Polζ, we compared the molecular interactions of human Rev7 in the context of shieldin to those of Rev7 in yeast Polζ. Although the yeast and human Rev7 proteins display virtually identical folds, their dimeric conformations and how they bind target proteins differ markedly. These differences are surprising and suggest that the oligomeric conformation of Rev7 may be context-dependent if one assumes that the spatial arrangement of subunits is the same in budding yeast and human Polζ. Alternatively, yeast and human Polζ may differ in structure and modes of action. A future challenge will be to determine the structure of human Polζ, which would be of high value for the long-term development of new anti-cancer drugs[15,17,52,53].

3.5 References

1. Baynton, K., & Fuchs, R. P. P. (2000). Lesions in DNA: hurdles for polymerases. *Trends in Biochemical Sciences*, *25*(2), 74–79. https://doi.org/10.1016/s0968-0004(99)01524-8

2. Kunz, B. A., Straffon, A. F. L., & Vonarx, E. J. (2000). DNA damage-induced mutation: tolerance via translesion synthesis. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *451*(1-2), 169–185. https://doi.org/10.1016/s0027-5107(00)00048-8

3. Marians, K. J. (2018). Lesion Bypass and the Reactivation of Stalled Replication Forks. *Annual Review of Biochemistry*, *87*(1), 217–238. https://doi.org/10.1146/annurev-biochem-062917-011921

4. Prakash, S., Johnson, R. E., & Prakash, L. (2005). EUKARYOTIC TRANSLESION SYNTHESIS DNA POLYMERASES: Specificity of Structure and Function. *Annual Review of Biochemistry*, *74*(1), 317–353. https://doi.org/10.1146/annurev.biochem.74.082803.133250

5. Acharya, N., Haracska, L., Johnson, R. E., Unk, I., Prakash, S., & Prakash, L. (2005). Complex Formation of Yeast Rev1 and Rev7 Proteins: a Novel Role for the Polymerase-Associated Domain. *Molecular and Cellular Biology*, *25*(21), 9734–9740. https://doi.org/10.1128/mcb.25.21.9734-9740.2005

6. Gómez-Llorente, Y., Malik, R., Jain, R., Choudhury, J., Johnson, Robert E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., & Aggarwal, Aneel K. (2013). The Architecture of Yeast DNA Polymerase ζ. *Cell Reports*, *5*(1), 79–86. https://doi.org/10.1016/j.celrep.2013.08.046

7. Goodman, M. F. (2002). Error-Prone Repair DNA Polymerases in Prokaryotes and Eukaryotes. *Annual Review of Biochemistry*, *71*(1), 17–50. https://doi.org/10.1146/annurev.biochem.71.083101.124707

8. Leung, W., Baxley, R., Moldovan, G.-L., & Bielinsky, A.-K. (2018). Mechanisms of DNA Damage Tolerance: Post-Translational Regulation of PCNA. *Genes*, *10*(1), 10. https://doi.org/10.3390/genes10010010

9. Goodman, M. F., & Woodgate, R. (2013). Translesion DNA Polymerases. *Cold Spring Harbor Perspectives in Biology*, *5*(10), a010363–a010363. https://doi.org/10.1101/cshperspect.a010363

10. Acharya, N., Johnson, R. E., Prakash, S., & Prakash, L. (2006). Complex Formation with Rev1 Enhances the Proficiency of Saccharomyces cerevisiae DNA Polymerase ζ for Mismatch Extension and for Extension Opposite from DNA

Lesions. *Molecular and Cellular Biology*, *26*(24), 9555–9563. https://doi.org/10.1128/mcb.01671-06

11. Acharya, N., Johnson, R. E., Pages, V., Prakash, L., & Prakash, S. (2009). Yeast Rev1 protein promotes complex formation of DNA polymerase with Pol32 subunit of DNA polymerase. *Proceedings of the National Academy of Sciences*, *106*(24), 9631–9636. https://doi.org/10.1073/pnas.0902175106

12. Makarova, A. V., Stodola, J. L., & Burgers, P. M. (2012). A four-subunit DNA polymerase ζ complex containing Pol δ accessory subunits is essential for PCNA-mediated mutagenesis. *Nucleic Acids Research*, *40*(22), 11618–11626. https://doi.org/10.1093/nar/gks948

13. Makarova, A. V., & Burgers, P. M. (2015). Eukaryotic DNA polymerase ζ. *DNA Repair*, *29*, 47–55. https://doi.org/10.1016/j.dnarep.2015.02.012

14. Yamanaka, K., Chatterjee, N., Hemann, M. T., & Walker, G. C. (2017). Inhibition of mutagenic translesion synthesis: A possible strategy for improving chemotherapy? *PLOS Genetics*, *13*(8), e1006842. https://doi.org/10.1371/journal.pgen.1006842

15. Cui, G., Botuyan, M. V., & Mer, G. (2018). Structural Basis for the Interaction of Mutasome Assembly Factor REV1 with Ubiquitin. *Journal of Molecular Biology*, *430*(14), 2042–2050. https://doi.org/10.1016/j.jmb.2018.05.017

16. Taniguchi, T. (2019). REV1-POL ζ Inhibition and Cancer Therapy. *Molecular Cell*, *75*(3), 419–420. https://doi.org/10.1016/j.molcel.2019.07.012

17. Nair, D. T. (2005). Rev1 Employs a Novel Mechanism of DNA Synthesis Using a Protein Template. *Science*, *309*(5744), 2219–2222. https://doi.org/10.1126/science.1116336

18. Nelson, J. R., Lawrence, C. W., & Hinkle, D. C. (1996). Deoxycytidyl transferase activity of yeast REV1 protein. *Nature*, *382*(6593), 729–731. https://doi.org/10.1038/382729a0

19. Chang, D. J., & Cimprich, K. A. (2009). DNA damage tolerance: when it's OK to make mistakes. *Nature Chemical Biology*, *5*(2), 82–90. https://doi.org/10.1038/nchembio.139

20. Waters, L. S., Minesinger, B. K., Wiltrout, M. E., D'Souza, S., Woodruff, R. V., & Walker, G. C. (2009). Eukaryotic Translesion Polymerases and Their Roles and Regulation in DNA Damage Tolerance. *Microbiology and Molecular Biology Reviews : MMBR*, *73*(1), 134–154. https://doi.org/10.1128/MMBR.00034-08

21. Bezalel-Buch, R., Cheun, Y. K., Roy, U., Schärer, O. D., & Burgers, P. M. (2020). Bypass of DNA interstrand crosslinks by a Rev1–DNA polymerase ζ complex. *Nucleic Acids Research*, *48*(15), 8461–8473. https://doi.org/10.1093/nar/gkaa580

22. Martin, S. K., & Wood, R. D. (2019). DNA polymerase ζ in DNA replication and repair. *Nucleic Acids Research*, *47*(16), 8348–8361. https://doi.org/10.1093/nar/gkz705

23. Lawrence, C. W. (2004). Cellular Functions of DNA Polymerase ζ and Rev1 Protein. *Advances in Protein Chemistry*, 167–203. https://doi.org/10.1016/s0065-3233(04)69006-1

24. Morrison, A., Christensen, R. B., Alley, J., Beck, A. K., Bernstine, E. G., Lemontt, J. F., & Lawrence, C. W. (1989). REV3, a Saccharomyces cerevisiae gene whose function is required for induced mutagenesis, is predicted to encode a nonessential DNA polymerase. *Journal of Bacteriology*, *171*(10), 5659–5667. https://doi.org/10.1128/jb.171.10.5659-5667.1989

25. Nelson, J. R., Lawrence, C. W., & Hinkle, D. C. (1996a). Thymine-Thymine Dimer Bypass by Yeast DNA Polymerase zeta. *Science*, *272*(5268), 1646–1649. https://doi.org/10.1126/science.272.5268.1646

26. Zhong, X., Garg, P., Stith, C. M., McElhinny, S. A. N., Kissling, G. E., Burgers, P. M. J., & Kunkel, T. A. (2006). The fidelity of DNA synthesis by yeast DNA polymerase zeta alone and with accessory proteins. *Nucleic Acids Research*, *34*(17), 4731–4742. https://doi.org/10.1093/nar/gkl465

27. Malik, R., Kopylov, M., Gomez-Llorente, Y., Jain, R., Johnson, R. E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., & Aggarwal, A. K. (2020). Structure and mechanism of B-family DNA polymerase ζ specialized for translesion DNA synthesis. *Nature Structural & Molecular Biology*, *27*(10), 913–924. https://doi.org/10.1038/s41594-020-0476-7

28. Scheres, S. H. W., & Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. *Nature Methods*, *9*(9), 853–854. https://doi.org/10.1038/nmeth.2115

29. Nakane, T., Kimanius, D., Lindahl, E., & Scheres, S. H. (2018). Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *ELife*, *7*. https://doi.org/10.7554/elife.36861

30. Netz, D. J. A., Stith, C. M., Stümpfig, M., Köpf, G., Vogel, D., Genau, H. M., Stodola, J. L., Lill, R., Burgers, P. M. J., & Pierik, A. J. (2011). Eukaryotic DNA

polymerases require an iron-sulfur cluster for the formation of active complexes. *Nature Chemical Biology*, *8*(1), 125–132. https://doi.org/10.1038/nchembio.721

31. Jain, R., Rice, W. J., Malik, R., Johnson, R. E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., & Aggarwal, A. K. (2019). Cryo-EM structure and dynamics of eukaryotic DNA polymerase δ holoenzyme. *Nature Structural & Molecular Biology*, *26*(10), 955–962. https://doi.org/10.1038/s41594-019-0305-z

32. Johnson, R. E., Prakash, L., & Prakash, S. (2012). Pol31 and Pol32 subunits of yeast DNA polymerase are also essential subunits of DNA polymerase. *Proceedings of the National Academy of Sciences*, *109*(31), 12455–12460. https://doi.org/10.1073/pnas.1206052109

33. Brandão, L. N., Ferguson, R., Santoro, I., Jinks-Robertson, S., & Sclafani, R. A. (2014). The Role of Dbf4-Dependent Protein Kinase in DNA Polymerase ζ-Dependent Mutagenesis in Saccharomyces cerevisiae. *Genetics*, *197*(4), 1111–1122. https://doi.org/10.1534/genetics.114.165308

34. Fattah, F. J., Hara, K., Fattah, K. R., Yang, C., Wu, N., Warrington, R., Chen, D. J., Zhou, P., Boothman, D. A., & Yu, H. (2014). The Transcription Factor TFII-I Promotes DNA Translesion Synthesis and Genomic Stability. *PLoS Genetics*, *10*(6), e1004419. https://doi.org/10.1371/journal.pgen.1004419

35. Hara, K., Hashimoto, H., Murakumo, Y., Kobayashi, S., Kogame, T., Unzai, S., Akashi, S., Takeda, S., Shimizu, T., & Sato, M. (2010). Crystal Structure of Human REV7 in Complex with a Human REV3 Fragment and Structural Implication of the Interaction between DNA Polymerase ζ and REV1. *Journal of Biological Chemistry*, *285*(16), 12299–12307. https://doi.org/10.1074/jbc.m109.092403

36. Lemontt J. F. (1971). Mutants of yeast defective in mutation induced by ultraviolet light. *Genetics, 68*(1), 21–33. https://doi.org/10.1093/genetics/68.1.21

37. Bhat, A., Wu, Z., Maher, V. M., McCormick, J. J., & Xiao, W. (2015). Rev7/Mad2B plays a critical role in the assembly of a functional mitotic spindle. *Cell Cycle*, *14*(24), 3929–3938. https://doi.org/10.1080/15384101.2015.1120922

38. Murakumo, Y., Ogura, Y., Ishii, H., Numata, S., Ichihara, M., Croce, C. M., Fishel, R., & Takahashi, M. (2001). Interactions in the Error-prone Postreplication Repair Proteins hREV1, hREV3, and hREV7. *Journal of Biological Chemistry*, *276*(38), 35644–35651. https://doi.org/10.1074/jbc.m102051200

39. Rizzo, A. A., Vassel, F.-M., Chatterjee, N., D'Souza, S., Li, Y., Hao, B., Hemann, M. T., Walker, G. C., & Korzhnev, D. M. (2018). Rev7 dimerization is important for assembly and function of the Rev1/Polζ translesion synthesis complex. *Proceedings of the National Academy of Sciences*, *115*(35), E8191–E8200. https://doi.org/10.1073/pnas.1801149115

40. Guo, C. (2003). Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *The EMBO Journal*, *22*(24), 6621–6630. https://doi.org/10.1093/emboj/cdg626

41. Masuda, Y., Ohmae, M., Masuda, K., & Kamiya, K. (2003). Structure and Enzymatic Properties of a Stable Complex of the Human REV1 and REV7 Proteins. *Journal of Biological Chemistry*, *278*(14), 12356–12360. https://doi.org/10.1074/jbc.m211765200

42. Rosenberg, S. C., & Corbett, K. D. (2015). The multifaceted roles of the HORMA domain in cellular signaling. *The Journal of Cell Biology*, *211*(4), 745–755. https://doi.org/10.1083/jcb.201509076

43. Miniowitz-Shemtov, S., Eytan, E., Kaisari, S., Sitry-Shevah, D., & Hershko, A. (2015). Mode of interaction of TRIP13 AAA-ATPase with the Mad2-binding protein p31comet and with mitotic checkpoint complexes. *Proceedings of the National Academy of Sciences*, *112*(37), 11536–11540. https://doi.org/10.1073/pnas.1515358112

44. Zhang, L., Yang, S.-H., & Sharrocks, A. D. (2007). Rev7/MAD2B Links c-Jun N-Terminal Protein Kinase Pathway Signaling to Activation of the Transcription Factor Elk-1. *Molecular and Cellular Biology*, *27*(8), 2861–2869. https://doi.org/10.1128/mcb.02276-06

45. Dai, Y., Zhang, F., Wang, L., Shan, S., Gong, Z., & Zhou, Z. (2020). Structural basis for shieldin complex subunit 3–mediated recruitment of the checkpoint protein REV7 during DNA double-strand break repair. *Journal of Biological Chemistry*, *295*(1), 250–262. https://doi.org/10.1074/jbc.ra119.011464

46. Liang, L., Feng, J., Zuo, P., Yang, J., Lu, Y., & Yin, Y. (2020). Molecular basis for assembly of the shieldin complex and its implications for NHEJ. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15879-5

47. Xie, W., Wang, S., Wang, J., de la Cruz, M. J., Xu, G., Scaltriti, M., & Patel, D. J. (2021). Molecular mechanisms of assembly and TRIP13-mediated remodeling of the human Shieldin complex. *Proceedings of the National Academy of Sciences*, *118*(8), e2024512118. https://doi.org/10.1073/pnas.2024512118

48. Wang, D., Ma, J., Botuyan, M. V., Cui, G., Yan, Y., Ding, D., Zhou, Y., Krueger, E. W., Pei, J., Wu, X., Wang, L., Pei, H., McNiven, M. A., Ye, D., Mer, G., & Huang, H. (2021). ATM-phosphorylated SPOP contributes to 53BP1 exclusion from chromatin during DNA replication. *Science Advances*, *7*(25), eabd9208. https://doi.org/10.1126/sciadv.abd9208

49. Gupta, R., Somyajit, K., Narita, T., Maskey, E., Stanlie, A., Kremer, M., Typas, D., Lammers, M., Mailand, N., Nussenzweig, A., Lukas, J., & Choudhary, C. (2018). DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell*, *173*(4), 972-988.e23. https://doi.org/10.1016/j.cell.2018.03.050

50. Noordermeer, S. M., Adam, S., Setiaputra, D., Barazas, M., Pettitt, S. J., Ling, A. K., Olivieri, M., Álvarez-Quilón, A., Moatti, N., Zimmermann, M., Annunziato, S., Krastev, D. B., Song, F., Brandsma, I., Frankum, J., Brough, R., Sherker, A., Landry, S., Szilard, R. K., & Munro, M. M. (2018). The shieldin complex mediates 53BP1-dependent DNA repair. *Nature*, *560*(7716), 117–121. https://doi.org/10.1038/s41586-018-0340-7

51. Wojtaszek, J. L., Chatterjee, N., Najeeb, J., Ramos, A., Lee, M., Bian, K., Xue, J. Y., Fenton, B. A., Park, H., Li, D., Hemann, M. T., Hong, J., Walker, G. C., & Zhou, P. (2019). A Small Molecule Targeting Mutagenic Translesion Synthesis Improves Chemotherapy. *Cell*, *178*(1), 152-159.e11. https://doi.org/10.1016/j.cell.2019.05.028

52. Chatterjee, N., Whitman, M. A., Harris, C. A., Min, S. M., Jonas, O., Lien, E. C., Luengo, A., Vander Heiden, M. G., Hong, J., Zhou, P., Hemann, M. T., & Walker, G. C. (2020). REV1 inhibitor JH-RE-06 enhances tumor cell response to chemotherapy by triggering senescence hallmarks. *Proceedings of the National Academy of Sciences*, *117*(46), 28918–28921. https://doi.org/10.1073/pnas.2016064117

53. Nesvizhskii, A. I., Keller, A., Kolker, E., & Aebersold, R. (2003). A Statistical Model for Identifying Proteins by Tandem Mass Spectrometry. *Analytical Chemistry*, *75*(17), 4646–4658. https://doi.org/10.1021/ac0341261

54. Keller, A., Nesvizhskii, A. I., Kolker, E., & Aebersold, R. (2002). Empirical Statistical Model To Estimate the Accuracy of Peptide Identifications Made by MS/MS and Database Search. *Analytical Chemistry*, *74*(20), 5383–5392. https://doi.org/10.1021/ac025747h

55. Guo, D. (2004). Translesion synthesis of acetylaminofluorene-dG adducts by DNA polymerase is stimulated by yeast Rev1 protein. *Nucleic Acids Research*, *32*(3), 1122–1130. https://doi.org/10.1093/nar/gkh279

56. Otwinowski, Z., & Minor, W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, 307–326. https://doi.org/10.1016/s0076-6879(97)76066-x

57. 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. *Acta Crystallographica Section D Biological Crystallography*, *66*(2), 213–221. https://doi.org/10.1107/s0907444909052925

58. Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography*, *60*(12), 2126–2132. https://doi.org/10.1107/s0907444904019158

59. DeLano, W. L. (2002) The PyMOL molecular graphics system. http://www. pymol. org

60. 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. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

61. Ohi, M., Li, Y., Cheng, Y., & Walz, T. (2004). Negative staining and image classification- powerful tools in modern electron microscopy. *Biological Procedures Online*, *6*(1), 23–34. https://doi.org/10.1251/bpo70

62. Scheres, S. H. W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

63. Chiu, P.-L., Li, X., Li, Z., Beckett, B., Brilot, A. F., Grigorieff, N., Agard, D. A., Cheng, Y., & Walz, T. (2015). Evaluation of super-resolution performance of the K2 electron-counting camera using 2D crystals of aquaporin-0. *Journal of Structural Biology*, *192*(2), 163–173. https://doi.org/10.1016/j.jsb.2015.08.015

64. Cheng, A., Eng, E. T., Alink, L., Rice, W. J., Jordan, K. D., Kim, L. Y., Potter, C. S., & Carragher, B. (2018). High resolution single particle cryo-electron microscopy using beam-image shift. *Journal of Structural Biology*, *204*(2), 270–275. https://doi.org/10.1016/j.jsb.2018.07.015

65. Mastronarde, D. N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology*, *152*(1), 36–51. https://doi.org/10.1016/j.jsb.2005.07.007

66. Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

67. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A. J., & Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods*, *16*(11), 1153–1160. https://doi.org/10.1038/s41592-019-0575-8

68. Punjani, A., Zhang, H., & Fleet, D. J. (2020). Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nature Methods*, *17*(12), 1214–1221. https://doi.org/10.1038/s41592-020-00990-8

69. Cardone, G., Heymann, J. B., & Steven, A. C. (2013). One number does not fit all: Mapping local variations in resolution in cryo-EM reconstructions. *Journal of Structural Biology*, *184*(2), 226–236. https://doi.org/10.1016/j.jsb.2013.08.002

70. Tan, Y. Z., Baldwin, P. R., Davis, J. H., Williamson, J. R., Potter, C. S., Carragher, B., & Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods*, *14*(8), 793–796. https://doi.org/10.1038/nmeth.4347

71. Pintilie, G. D., Zhang, J., Goddard, T. D., Chiu, W., & Gossard, D. C. (2010). Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *Journal of Structural Biology*, *170*(3), 427–438. https://doi.org/10.1016/j.jsb.2010.03.007

72. Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development ofCoot. *Acta Crystallographica Section D Biological Crystallography*, *66*(4), 486–501. https://doi.org/10.1107/s0907444910007493

73. Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., & Ferrin, T. E. (2017). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science*, *27*(1), 14–25. https://doi.org/10.1002/pro.3235

CHAPTER 4: CRYO-EM STUDY OF THE INTEGRIN αvβ8 COMPLEXED WITH PRO-TGFB1-GARP

4.1 Introduction

In our body, the right reactions to extracellular, intracellular, and intercellular signals maintain cellular homeostasis in the organism[1]. Imbalances in these signals can cause cellular homeostasis to be disrupted, resulting in alterations in cell mass and/or tissue organization. The secreted, cell surface, and intracellular chemicals that make up the cellular homeostatic machinery keep cellular differentiation and the balance between quiescence and cell cycle entrance in check[2]. Controlling cell proliferation via cell–extracellular matrix interactions, in conjunction with growth-promoting and inhibitory cytokines, is how homeostasis is maintained[3]. TGF-β is a multifunctional cytokine that plays an important role in tissue homeostasis as a growth inhibitory cytokine. TGF-β plays a crucial role in homeostasis, as evidenced by the fact that TGF-β1 deficient animals exhibit epithelial hyperplasias and fatal multiorgan inflammation within weeks after birth[4]. TGF-β1 is typically kept latent or inactive by noncovalent interactions between the bioactive peptide of TGF-β1 with its NH2-terminal propeptide, latency-associated peptide LAP1. As a result, normal TGF-β function is considered to be primarily regulated by its activation from a dormant state, which is a process that is not entirely understood.

Recent data shows that TGF-β may be activated by cell surface molecules or secreted extracellular molecules. Moreover, the TGF-β1 activation has been linked to nonproteolytic processes involving the integrin αvβ8 and glyco-protein A repetitions

predominant (GARP)[5]. In recent discoveries, human regulatory T cells (Tregs) inhibit other T cells by transforming the dormant, inactive form of TGF-β1 (pro-TGF-β1) into active TGF-β1. TGF-β1 activation in Tregs requires GARP, which is a transmembrane protein that binds and presents latent TGF-1 on the surface of Tregs activated by their T cell receptor. However, GARP alone is insufficient since transduction of GARP in non-Treg T cells does not result in active TGF-1 production. Since it is required for T cell, myeloid, and endothelial cell differentiation during development, integrin avb8 is particularly crucial for TGF-β activation[6]. Avb8 is involved in fibroinflammatory processes as well as anti-tumor immunity in the postnatal period[7]. In particular, avb8 exclusively binds to TGF-β effectively, in contrast to the other integrins, which bind to a variety of ligands in addition to TGF-β[8, 9, 10].

The activation of pro-TGF-β by GARP (glycoprotein A repeats predominate) on their cell surface together with avb8 plays a key role in understanding immunosuppression in cancer or chronic infections. As a result, identifying the basis of avb8 specificity for pro-TGF-b and GARP is significant for understanding integrin-mediated TGF-β activation and guiding the creation of therapeutic antibodies and integrin-specific small compounds, which has proven challenging.

To understand the interactions within αvβ8-pro-TGFβ1-GARP, we utilize the single-particle cryo-EM technique and studied this protein complex. In addition, we apply different cryo-EM sample preparation methods to optimize the plunge frozen grid for high-resolution structure determination.

4.2 Materials and Methods

4.2.1 Protein complex formation of αvβ8-pro-TGFβ1

Three proteins of αvβ8, pro-TGFβ1, GARP were produced in Dr. Albert Lin's laboratory in collaboration between ASU and Morphic Therapeutic. Protein solutions were flash frozen and shipped to ASU in 1-4 mg quantity. The binding efficiently of αvβ8-pro-TGFβ1 complex was studied to understand its initial interaction. Integrin αvβ8 was combined with pro-TGFβ1 at a 3:1 molar ratio (αvβ8: pro-TGFβ1) and incubated at room temperature for one hour. Then, the mixture was injected onto a Superdex 200 10/300 GL SEC column to separate the complex from unbound protein by size. The peak fractions of target complex were collected according to the expected elution volumes and their corresponding molecular weights, which is subsequently used for negative stain EM experiment.

4.2.2. On-column crosslinking of αvβ8-pro-TGFβ1-GARP

Prior to the experiment, a Superdex 200 10/300 GL SEC column was equilibrated with 2CV (50 mL) of buffer containing 20 mM HEPES pH 7.3, 150 mM NaCl, 2 mM MnCL2, and 0.1 mM CaCl2. A 4 mL injection loop was attached and washed with deionized water (diH2O) and then the previously mentioned buffer. After that, 4 mL of 0.5% of glutaraldehyde was injected onto the column, at 0.25 mL/min flow rate. A 500 µL injection loop was used to replace the 4 mL injection loop, which it was washed with deionized water, then buffer.

To start the experiment, αvβ8-pro-TGF-β1-GARP protein mixture was combined at a 1: 1.25: 1.25 molar ratio, respectively. The mixture incubated in room temperature for 1 hour before subjected to SEC using a Superose 6 Increase GL 10/300 SEC column. Peak fractions of target complex were collected according to the expected elution volumes and

their corresponding molecular weights, which is subsequently used for negative stain EM experiment.

4.2.3 Complex analysis using negative-stain EM

The negative-stain electron microscopy was strictly followed by the procedures mentioned in Chapter 2.

4.2.4 Preparation of graphene oxide coated EM grid

First, graphene oxide (GO) solution was diluted to 0.2mg/mL with deionized water. The carbon side of a C-flat 400-mesh holey-carbon coated copper grid (2/1 4C; Protochips) was glow-discharged for 30 seconds in a Pelco easiGlow glow-discharge system (Ted Pella). Then, the clean surface of a parafilm was prepared. One drop containing 5 µl of graphene oxide solution and two drops of diH2O were placed onto the parafilm surface. The carbon side of the grid was placed down onto GO solution and incubated for 1 minute. An anti-capillary tweezers was used to pick up the grid and a folded filter paper was used to remove excess solution. The first drop of ddH20 was taken up by the carbon side of the grid, excess diH2O was removed with the filter paper. The second drop of water was done the same. After all excess water was removed from the grid, it was let dry for 5 minutes before usage.

4.2.5 Single-particle cryo-EM data collection

Five µL of 0.1 mg/mL protein sample were then applied on the graphene oxide coated EM grid. The grid was blotted by a filter paper to remove the excess solution and quickly frozen into liquid ethane. The plunge freezing process was automated using a Vitrobot Mark IV plunge freezer (Thermo Fisher/FEI) at a humidity of 100% with a blotting time of 6 sec.

All the cryo-EM data collections were completed in the Eyring Materials Center (EMC) at Arizona State University (ASU). The grid specimens were imaged using a Thermo

Fisher/FEI Titan Krios transmission electron microscope (TEM) (Thermo Fisher Scientific/FEI, Waltham, MA) at an accelerating voltage of 300 keV. The electron scattering was recorded by a Gatan Summit K2 DDD camera in super-resolution mode. The nominal magnification was set to 48,780×, corresponding to a pixel size of 1.04 Å/pixel at the specimen level. The defocus was set to vary from -1.0 to -3.5 μm with a total dosage of a total electron dosage of 60 $e^-/Å^2$. The dataset was collected in counting mode. The beam-image shift scheme was applied to accelerate the data collection.

4.2.6 Image processing

Image processing was generally conducted using cryoSPARC (version 3.0)[11] . A total of 11,415 cryo-EM movies was imported into the program for processing. The frame registration and averaging for motion correction were performed using the 'Patch motion correction' and the estimation of the defocus was performed using the 'Patch CTF estimation'. An ensemble of 1,358,170 particles was automatically selected using a neural network and positive-unlabeled learning by the Topaz program (version 0.2.3)[14]. The curation of the particle images was performed using iterative 2D image classification. A total of 509,015 particles was selected for an *ab initio* 3D map generation[11].

4.3 Results and Discussion

Individual protein αvβ8 and pro-TGF-β1 were combined at the molar ratio of 3:1 (αvβ8: pro-TGF-β1) and incubated at room temperature for one hour. The mixture was then subjected to SEC to separate the complex from unbound protein by size. The peak fraction that was within the correct range of molecular weight was pooled and the sample was used for negative staining EM (Figure 4.1). Negative stain EM allows us to visualize and assess the homogeneity of the complex. According to the negative stain EM analysis, there were

indication of complex formation. However, there were also falloffs and individual unbound proteins.

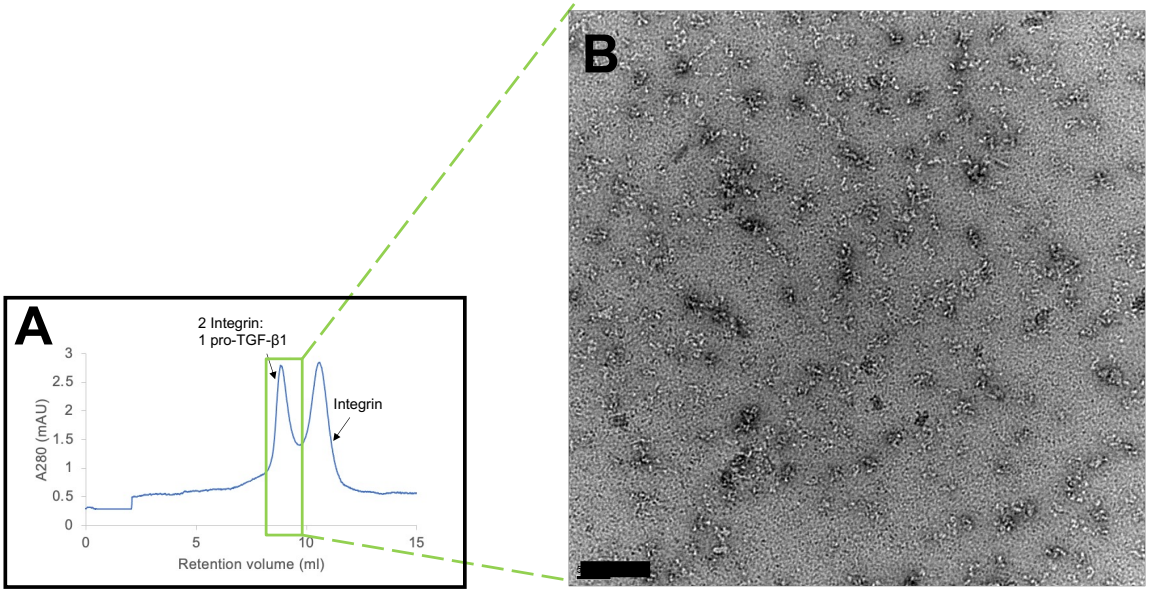

Figure 4.1. αvβ8 and pro-TGF-β1 complex assembly. (A) Size exclusion chromatogram of αvβ8 and pro-TGF-β1 showing the complex is formed with a molar ratio of 2 integrin to 1 pro-TGF-β1. (B) Peak fraction of the complex was used for negative stain EM analysis. Proteins are shown in white and the background in black. Scale bar indicates 50 nm.

To stabilize complex formation of αvβ8-pro-TGFβ1 as an initial assessment, the mixture was subjected to on-column crosslinking using 0.5% glutaraldehyde. Different species of the complex were identified. There were complexes in different molar ration 1:1 (αvβ8: pro-TGFβ1) and 2:1 (αvβ8: pro-TGFβ1) as confirmed by the SEC. Peak fraction at the correct molecular weight range was used for negative stain analysis, which shows a better homogeneity of the proteins. Negative stain data was collected and 2D classification was performed to assess the protein quality. According to the 2D class averages, there were evidence of αvβ8-pro-TGFβ1 complex formation (Figure 4.2).

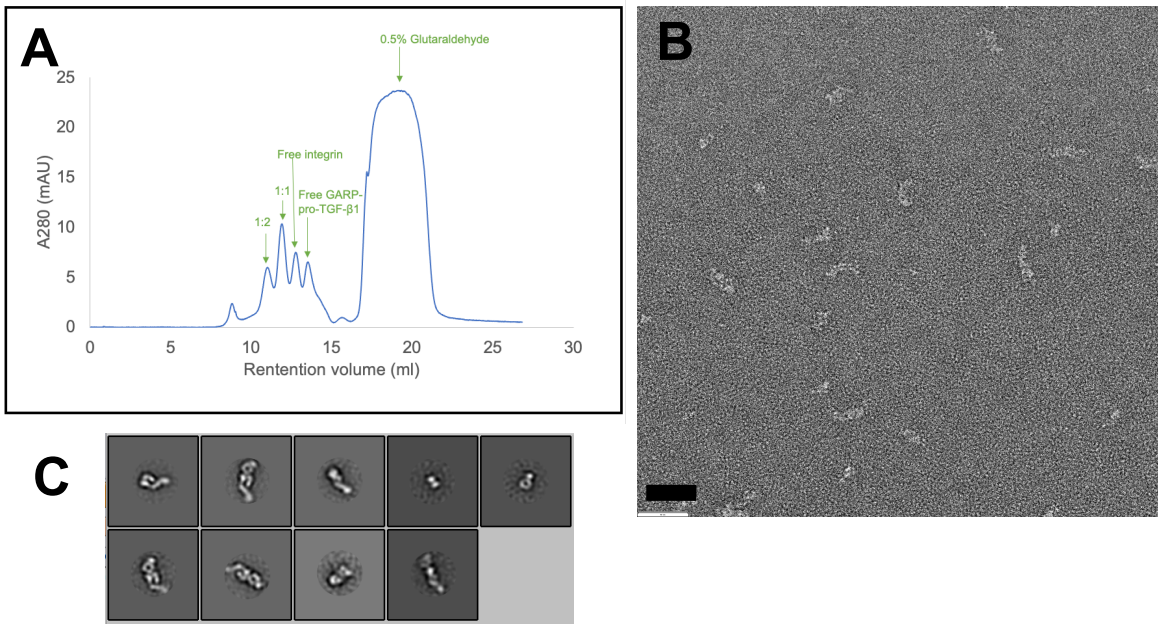

Figure 4.2. Crosslinking analysis of αvβ8 and pro-TGF-β1 complex. (A) Size exclusion chromatogram of αvβ8 and pro-TGF-β1 showing different species of the complex is formed with molar ratio of 2:1 or 1:1 integrin to pro-TGF-β1. (B) Peak fraction of the 1:1 complex was used for negative stain EM analysis. Proteins are shown in white and the background in black. Scale bar indicates 50 nm.

Therefore, the sample was used for plunge freezing and further cryo-EM study. However, proteins tended to cluster in the grid hole and not evenly distributed to improve the homogeneity on the EM grid, a layer of Graphene Oxide was coated onto the carbon support of the EM grid before sample application (Figure 4.3)[12,13]. As a result, cryo-EM data of the complex was collected and 2D classified to assess protein quality and carried on with further image analysis (Figure 4.4). The 2D classes show many different views of the complex, indicating that αvβ8-pro-TGF-β1 is highly flexible.

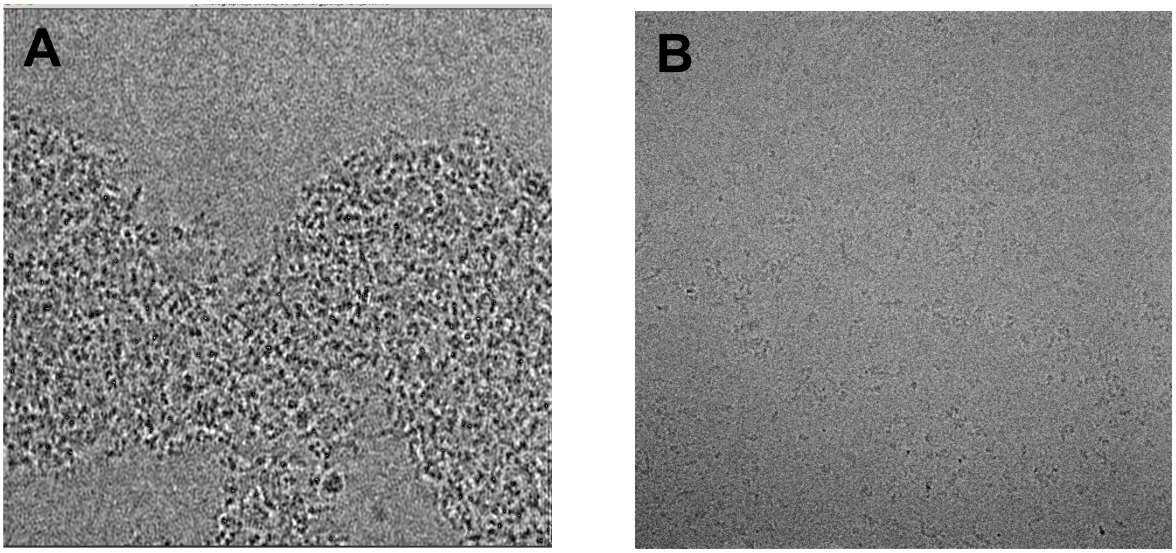

Figure 4.3. Cryo-EM analysis of αvβ8-pro-TGFβ1. (A) Plunge frozen particles appeared in the cryo-EM grid are shown in black with the gray background. Proteins clustered together and were not homogeneous. (B) Cryo-EM grid coated with graphene oxide was used for plunge freezing, which shows an improved homogeneity of particle distribution.

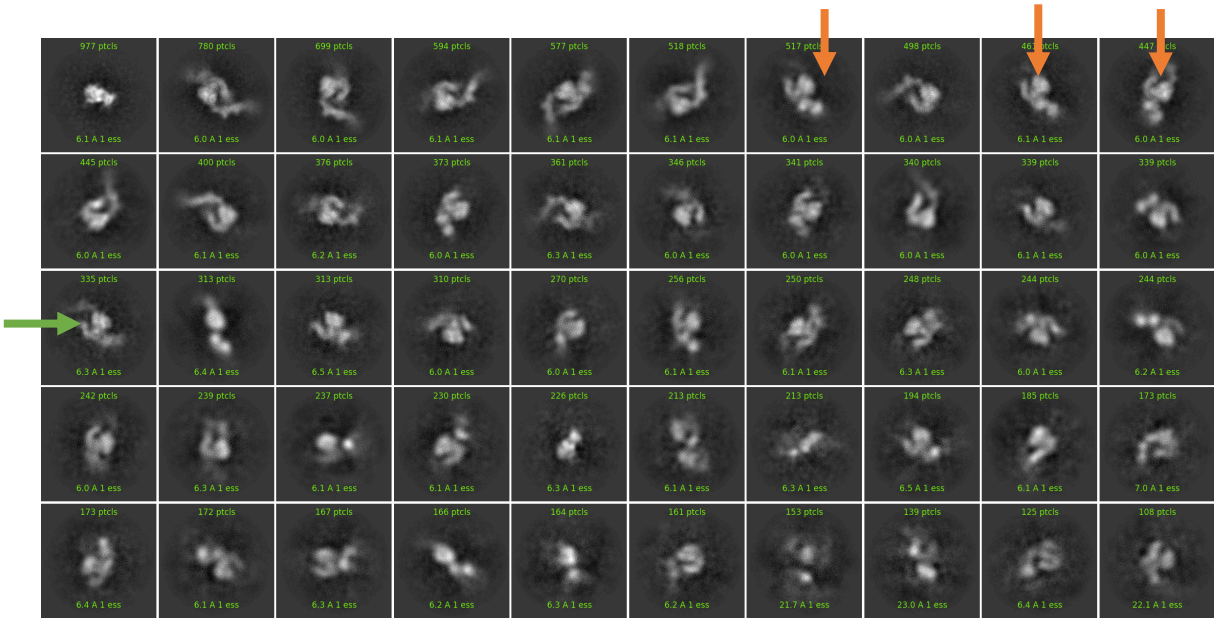

Figure 4.4. Two-dimensional classification of αvβ8-pro-TGFβ1. Different orientations of the complex were represented in orange arrow, and the unbound integrin was represented in green arrow. The 2D classes were performed with k number of 50.

Knowing the αvβ8-pro-TGFβ1 can form a stable complex, GARP protein was then introduced into the arrangement. With the addition of GARP and a Fab antibody that was

designed to bind pro-TGFβ1, a full understanding of the interactions within αvβ8-pro-TGFβ1-GARP can be achieved. Size exclusion chromatography and EM were conducted to study the structure of said complex (Figure 4.5). However, we soon realized that in order to reconstruct and acquire a high-resolution structure of αvβ8-pro-TGFβ1-GARP, numerous particles must be obtained. For more than 11,000 cryo-EM movies collected, only 2,705 particles within the population are αvβ8-pro-TGFβ1-GARP. With previous results showing the complex is highly flexible, obtaining a full density map is highly challenging. These factors resulted to missing fragments of the density map (Figure 4.6) after 3D reconstruction. Only the density of αvβ8 and pro-TGFβ1can be shown, other components are clearly lacking.

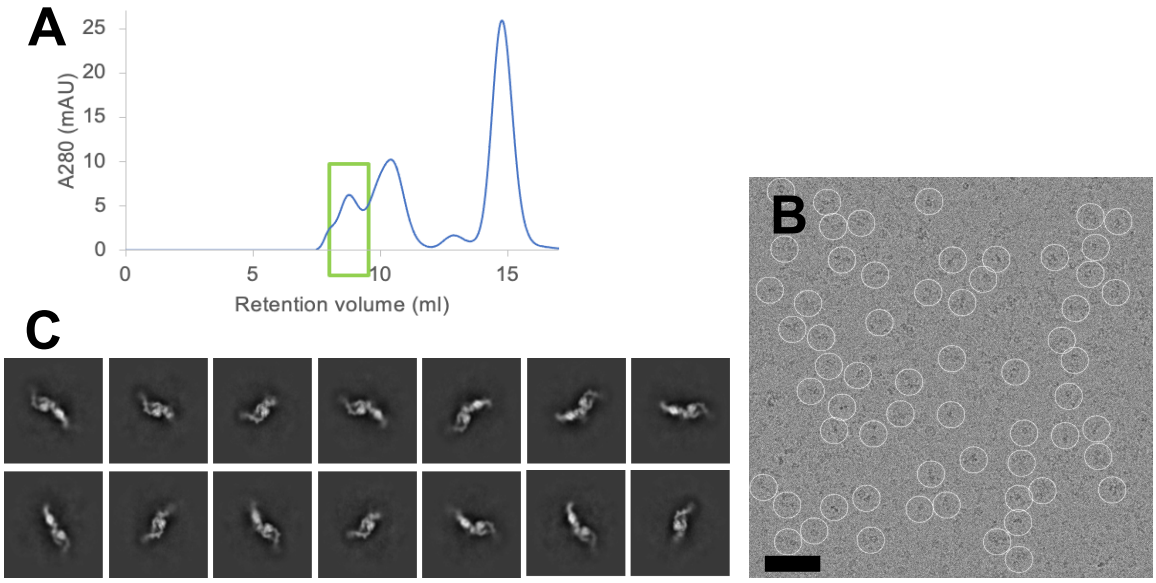

Figure 4.5. Cryo-EM analysis of αvβ8-pro-TGF-β1-GARP complex. (A) Size exclusion chromatogram of αvβ8-pro-TGF-β1-GARP with peak fraction highlighted in green box. (B) Peak fraction of the complex was used for subsequence cryo-EM data collection. Proteins are shown in black and the background in gray. Scale bar indicates 50 nm. (C) 2D classification of the collected data on αvβ8-pro-TGF-β1-GARP, k number is 14.

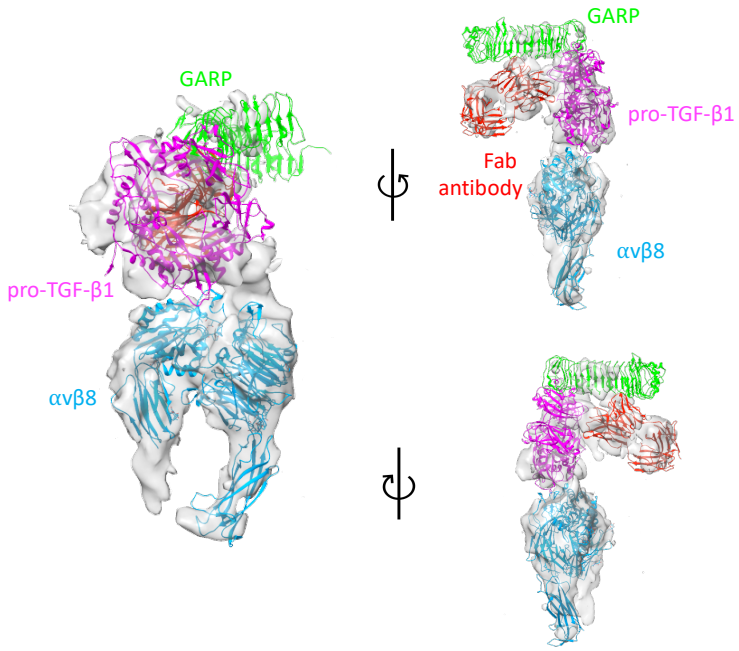

Figure 4.6. Three-dimensional reconstruction of αvβ8-pro-TGF-β1-GARP. Cryo-EM density of αvβ8-pro-TGF-β1-GARP is shown in gray. Structure of GARP shown in bright green, pro-TGF-β1 shown in pink, and Fab antibody shown in red (PDB: 6GFF), integrin αvβ8 is shown in blue (PDB: 6UJA) is fitted into the density map and represented in different angle rotated on the y-axis.

4.4 Conclusions

Understanding the activation and structure of pro-TGF-β1 in association with αvβ8 and GARP are essential for immunotherapy. Here, we performed a series of cryo-EM study showing the interaction of pro-TGF-β1 with αvβ8, as well as together with GARP. The complex of αvβ8-pro-TGF-β1 can be steadily formed using crosslinking chemical such as glutaraldehyde. As the sample being plunge frozen into the cryo-EM grid, image analysis under the microscope showed uneven particle distribution and clustering. One way to overcome this challenge is to coat the EM grid with graphene oxide prior to adding proteins and plunge freezing grid. Throughout our cryo-EM study of the αvβ8-pro-TGF-β1-GARP, we learned that the complex is extremely flexible as it was shown in 2D class averages. With such various views and orientation of the particles, it is inherently difficult to produce

coherent 2D class averages and reconstruct a high-resolution 3D density map for structure study. In order to achieve higher resolution 3D reconstruction, higher particle number must be obtained. To extract more particles, more cryo-EM data collection must be conducted until the cryo-EM density map shows a clear view of the complex with all components well fitted. After that, further 3D refinement can be carried out to reach a high-resolution structure.

4.5 References

1. Potter, V. R. (1974). Probabilistic Aspects of the Human Cybernetic Machine. *Perspectives in Biology and Medicine*, *17*(2), 164–183. https://doi.org/10.1353/pbm.1974.0023

2. LORD, B. I. (1988). Feedback regulators in normal and tumour tissues. *Journal of Cell Science*, *1988*(Supplement_10), 231–242. https://doi.org/10.1242/jcs.1988.supplement_10.16

3. Giancotti, F. G. (1997). Integrin signaling: specificity and control of cell survival and cell cycle progression. *Current Opinion in Cell Biology*, *9*(5), 691–700. https://doi.org/10.1016/s0955-0674(97)80123-8

4. Crawford, S. E., Stellmach, V., Murphy-Ullrich, J. E., Ribeiro, S. M. F., Lawler, J., Hynes, R. O., Boivin, G. P., & Bouck, N. (1998). Thrombospondin-1 Is a Major Activator of TGF-β1 In Vivo. *Cell*, *93*(7), 1159–1170. https://doi.org/10.1016/s0092-8674(00)81460-9

5. Stockis, J., Liénart, S., Colau, D., Collignon, A., Nishimura, S. L., Sheppard, D., Coulie, P. G., & Lucas, S. (2017). Blocking immunosuppression by human Tregs in vivo with antibodies targeting integrin αVβ8. *Proceedings of the National Academy of Sciences*, *114*(47), E10161–E10168. https://doi.org/10.1073/pnas.1710680114

6. Aluwihare, P., Mu, Z., Zhao, Z., Yu, D., Weinreb, P. H., Horan, G. S., Violette, S. M., &Munger, J. S. (2009). Mice that lack activity of αvβ6- and αvβ8-integrins reproduce the abnormalities of Tgfb1- and Tgfb3-null mice. *Journal of Cell Science*, *122*(2), 227–232. https://doi.org/10.1242/jcs.035246

7. Kudo, M., Melton, A. C., Chen, C., Engler, M. B., Huang, K. E., Ren, X., Wang, Y., Bernstein, X., Li, J. T., Atabai, K., Huang, X., & Sheppard, D. (2012). IL-17A produced by αβ T cells drives airway hyper-responsiveness in mice and enhances mouse and human airway smooth muscle contraction. *Nature Medicine*, *18*(4), 547–554. https://doi.org/10.1038/nm.2684

8. Humphries, J. D. (2006). Integrin ligands at a glance. *Journal of Cell Science*, *119*(19), 3901–3903. https://doi.org/10.1242/jcs.03098

9. Mu, D., Cambier, S., Fjellbirkeland, L., Baron, J. L., Munger, J. S., Kawakatsu, H., Sheppard,D., Broaddus, V. C., & Nishimura, S. L. (2002). The integrin αvβ8 mediates epithelial homeostasis through MT1-MMP–dependent activation of TGF-β1. *Journal of Cell Biology*, *157*(3), 493–507. https://doi.org/10.1083/jcb.200109100

10. Ozawa, A., Sato, Y., Imabayashi, T., Uemura, T., Takagi, J., & Sekiguchi, K. (2016). Molecular Basis of the Ligand Binding Specificity of αvβ8 Integrin. *Journal of Biological Chemistry*, *291*(22), 11551–11565. https://doi.org/10.1074/jbc.m116.719138

11. Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

12. Patel, A., Toso, D., Litvak, A., & Nogales, E. (2021). Efficient graphene oxide coating improves cryo-EM sample preparation and data collection from tilted grids. *BioRXiv*. https://doi.org/10.1101/2021.03.08.434344

13. Wang, F., Liu, Y., Yu, Z., Li, S., Feng, S., Cheng, Y., & Agard, D. A. (2020). General and robust covalently linked graphene oxide affinity grids for high-resolution cryo-EM. *Proceedings of the National Academy of Sciences*, *117*(39), 24269–24273. https://doi.org/10.1073/pnas.2009707117

14. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A. J., & Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods*, *16*(11), 1153–1160. https://doi.org/10.1038/s41592-019-0575-8

CHAPTER 5: MINIMIZATION OF PARTICLE ADSORPTION AT AIR-WATER INTERFACE USING LIPID MONOLAYER FOR SINGLE-PARTICLE CRYO-EM

5.1 Introduction

Electron cryogenic microscopy (cryo-EM) is a rapidly advancing technique in structural biology[1]. The principle of cryo-EM has been discussed in Chapter 1, which is about the electron beam interacts with the frozen specimen and travels through a combination of electromagnetic lenses to form an image, which provides high-resolution information for structure determination. However, biological specimens are extremely susceptible to beam radiation, therefore the sample is frozen in vitreous ice to localize structural damage and maintain hydration in the EM vacuum column[2,3]. A three-dimensional (3D) map of the target molecule's electrostatic potential may be rebuilt by combining molecular pictures with their corresponding Euler angles. The quality of 3D protein reconstructions has recently achieved atomic resolution in favorable instances, thanks to recent technological advances in the direct electron detector camera and high-performance computers.

The cryo-EM sample preparation is one of the bottlenecks that continues to restrict resolution[4]. The most popular approach for preparing frozen protein samples for imaging is plunge-freezing vitrification. The typical procedure is to blot out excess buffer from the EM grid using filter paper before immediately immersing the grid in a cryogen, such as liquid ethane or propane, for vitrification. This blotting procedure rapidly raises the surface-to-volume ratio of the sample, resulting in proteins clustering at the air-water interface in many cases[5]. Chapter 4 in this dissertation is a typical example of this problem.

Recent electron tomographic studies have shown the protein particles populated at the air-water interface in the vitreous ice[5]. The air-water interface has a high surface tension that may cause potential protein deformation[6]. Furthermore, particles adsorbed at the air-water interface have a preferential orientation in electron images[7], leading in an under-sampling of particular projection angles needed for 3D reconstruction. As a result, the quality of the single-particle cryo-EM reconstruction suffers.

One typical method for minimizing this air-water interface issue is to add a small quantity of detergent to the buffer, changing the chemical property of the air-water interface and decreasing particle adsorption[8]. However, detergent may affect the protein stability or structure[9]. Also, the type and concentration of the additive detergent must be optimized for an EM grid specimen to evenly disperse the particles. The optimization can be a time and labor-consuming process.

An alternative approach would be applying a supported layer such as a thin, continuous carbon film or graphene oxide flakes, onto the EM grid to reduce the contact of the particles with the air-water interface[10,11]. However, this additional support layer usually reduces the scattering signal and increases background in the image[12]. Furthermore, coverage of the graphene oxide layer on the EM grid is usually small and may occur in multiple layers[13]. In some cases, reducing the time duration between sample spotting and plunging reduced the number of particles adsorbed to the air-water interface[14]. However, there are practical limits to this approach so that it does not eliminate the particle absorbing to the air-water interface.

In this chapter, we present a new method to prepare the grid specimen for single-particle cryo-EM imaging. We used the lipid monolayer as a supporting layer to cover the EM grid

specimen on both opposite sides. Lipid monolayer was used to serve as a support layer for 2D crystals for the electron crystallographic studies[15]. The monolayer has also been used as a purification method for single-particle cryo-EM structure determination[15]. A lipid monolayer is formed on the surface of the aqueous solution by evaporating the volatile solvent, such as chloroform, used for dissolving lipids[16]. The lipid head groups will then interact the water molecules near the surface, leaving the hydrophobic acyl chains in the air (Figure 5.1).

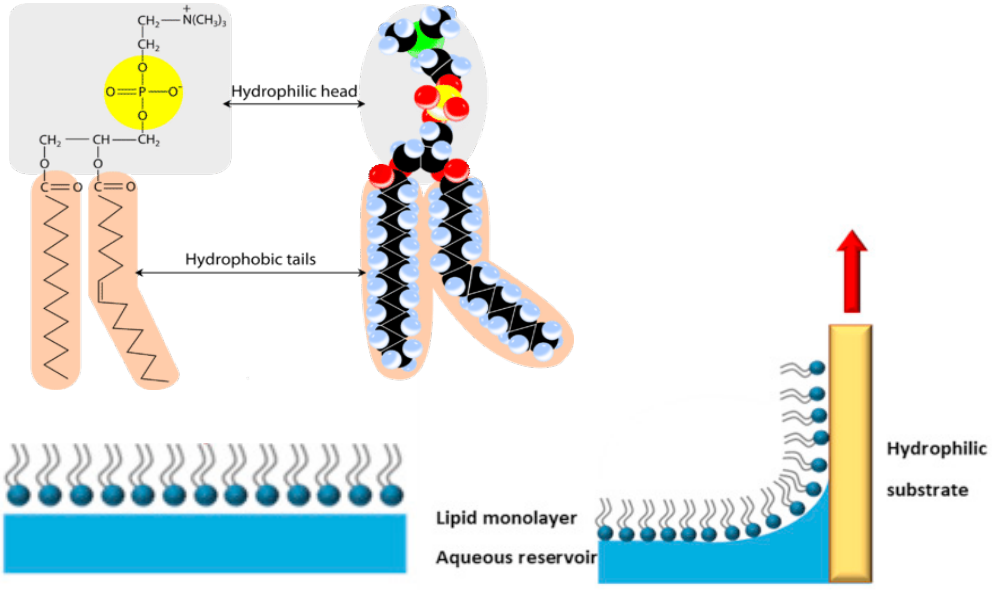

Figure 5.1. Illustration of the phosphor lipid structure[18] and the schematic representation of lipid monolayer formation on an aqueous reservoir[17]. As a hydrophilic substrate such as the carbon supported grid interacts with the lipid layer, the water surface tension produces a pull curve that create a sandwich formation with lipid monolayer being the protective layer outside the grid.

The lipid monolayer is flat and thin, which makes it a good material as a supporting layer. This technique has been used to grow 2D crystals for electron crystallography[19] and provide an affinity binding of target proteins for single-particle cryo-EM, such as

streptavidin and annexin IV[19]. We use lipid monolayers to sandwich the grid specimen and quickly freeze into cryogen for vitrification. The coverage of the monolayer sandwich is wide and stable. Also, the thickness of the lipid monolayer is thin. We used the protein samples of the apoferritin and the human p97 ATPase to test the idea, which can be proposed as another preservation method of the vitreous sample.

5.2 Materials and Methods

5.2.1 Protein samples

Lyophilized apoferritin from horse spleen was purchased from fisherSci (Cat no. ICN10026025). The p97 ATPase and p47 protein samples were provided by Dr. Tsui-Fen Chou laboratory (California Institute Technology, Pasadena, CA).

5.2.2 Sandwiched grid specimen preparation

DMPC (1,2-dimyristoyl-*sn*-glycero-3-phosphocholine) lipids (Avanti Polar Lipids, Alabaster, AL) was used to form lipid monolayer on the buffer surface. The lipids were in powder form and solubilized with 8:1 (v/v) chloroform/methanol. A Teflon plate with manually drilled wells was used as a buffer reservoir to provide enough volume for grid pick-up (Figure 5.2). The protein was added into the buffer to a final concentration of 0.2 mg/ml. The lipid monolayer formation was followed using the previous method[21]. The EM grid was glow-discharged using a Pelco easiGlow glow discharger (Ted Pella, Redding, CA) for 15 seconds on both sides of the grid. Once the grid is completely moved out of the buffer surface, a filter paper was used to blot away the excess solution from the side of the grid. The grid specimen was then quickly frozen into liquid ethane and saved in the dewar until cryo-EM imaging.

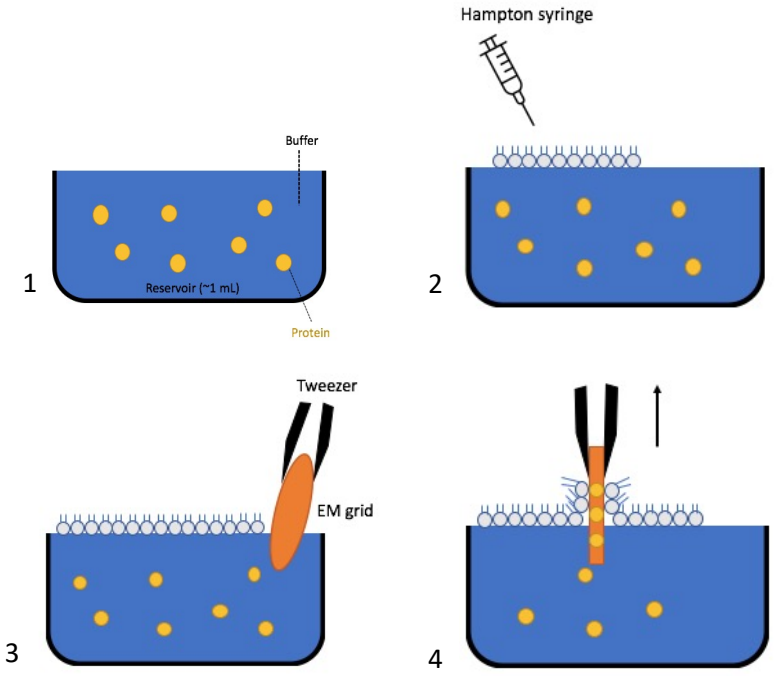

Figure 5.2. Schematic of the experimental procedure of lipid monolayer sandwich. The procedure was explained from step 1 to step 4. Step 1, protein and buffer are added into a reservoir. Step 2, lipid is introduced to the solution surface and allowed to form a lipid monolayer Step 3, A cryo-grid is then introduced into the reservoir, proteins are picked up by the grid and protected by lipid monolayer. Step 4, The grid is removed from solution, excess solution is blotted away, and the grid is quickly frozen into liquid ethane and imaged by cryo-EM.

5.2.3 Monolayer sandwich formation using a perfect loop

A new and improved procedure has been created for lipid monolayer formation on the EM grid. The monolayer is first formed on the buffer reservoir. A perfect loop (EMS, Cat. No 70944), designed to pick up thin section for ultramicrotomy, was used to pick up the lipid monolayer from the buffer reservoir. The monolayer was then transferred onto one side of the glow discharged grid. Protein sample was then added onto the other side of the grid. Excess buffer was removed. Another lipid monolayer was then introduced on that same side of the grid, which now has created a monolayer sandwich. The grid is now ready for EM imaging.

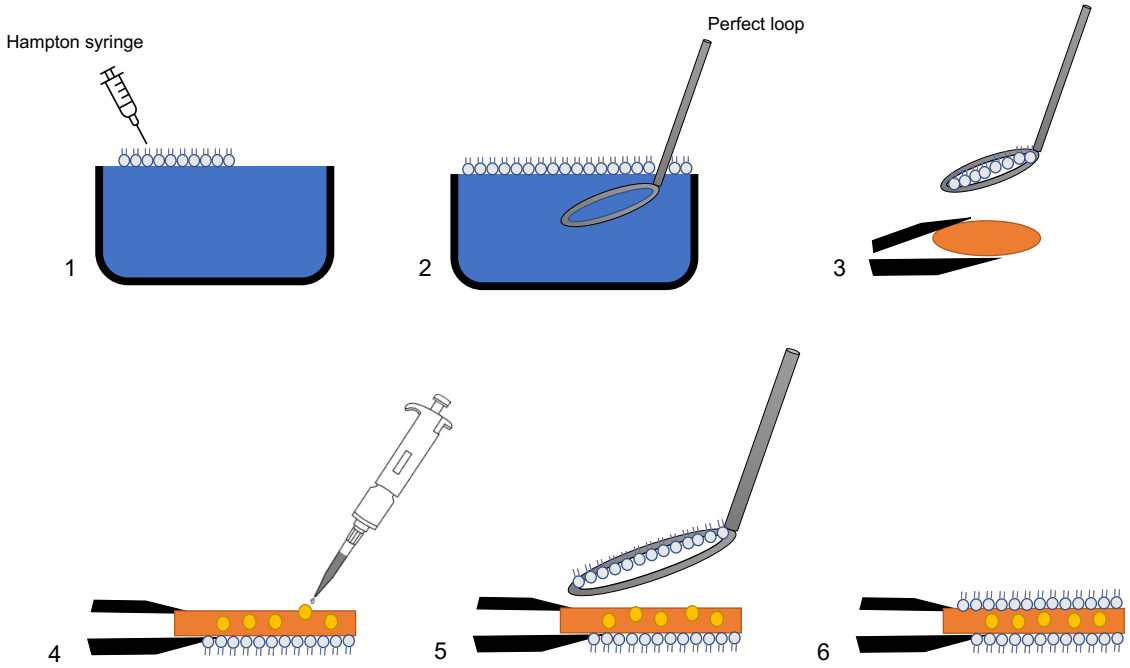

Figure 5.3. Lipid monolayer formation via a perfect loop. The procedure was explained from step 1 to step 6. Step 1: lipids are added on top of a buffer reservoir, forming the lipid monolayer. Step 2: a perfect loop is used to pick up the lipid monolayer. Step 3: the lipid monolayer is deposited onto the EM grid. Step 4: protein sample is added to the other side of the EM grid. Step 5: second layer of lipid is added. Step 6: a lipid monolayer sandwich is formed, the grid is quickly frozen into liquid ethane and imaged by cryo-EM.

5.2.3 Cryo-EM imaging

The grid specimens were imaged using a Thermo Fisher/FEI Titan Krios transmission electron microscope (TEM) (Thermo Fisher Scientific/FEI, Waltham, MA) at an accelerating voltage of 300 keV. The electron scattering was recorded by a Gatan Summit K2 DDD camera in super-resolution mode. The nominal magnification was set to 48,780×, corresponding to a pixel size of 1.04 Å/pixel at the specimen level. The defocus was set to vary from -0.8 to -2.8 µm. The camera counted rate was calibrated to 8.8 e^-/pixel/sec. The exposure time was 8 sec, accumulating to a total dosage of 71 $e^-/Å^2$. The dataset was collected in counting mode. The beam-image shift scheme was applied to accelerate the

data collection. The procedure of low-dose imaging was automated using SerialEM software (version 3.9).

5.2.4 Cryo-ET imaging

The cryo electron tomography data were collected in Pennsylvania State University, in collaboration with Dr. Joseph Wang at the College of Mecidine. Single-axis bidirectional tilt series spanning the -60 degrees to +60 degrees in 2 degrees increments were collected for the monolayer sandwiched sample in a Titan Krios TEM (Thermo Fisher Scientific/FEI, Waltham, MA) at an accelerating voltage of 300 keV with a Gatan BioQuantum K3 DDD camera. The exposure in each image was 0.7 e-/Å2 over 0.5 seconds at a pixel size of 1.7 Å/pixel at the specimen level. The tomographic series of the control dataset was collected using the same parameters. TFS Tomography (version 4) was used to automate the data collection processing. Tilt series alignment, CTF correction, and tomogram reconstruction were performed using IMOD (version 4.8.12)[22]. The final 3D reconstruction was generated using nonlinear anisotropic diffusion to reduce noise[23].

5.2.5 Data processing

Image processing was generally conducted using cryoSPARC (version 3.0)[24]. A total of 302 cryo-EM movies was imported into the program for processing. The frame registration and averaging for motion correction were performed using the 'Patch motion correction' and the estimation of the defocus was performed using the 'Patch CTF estimation'. Approximately 10,516 particles were automatically selected. After multiple rounds of particle curation, 4,941 particles were selected for 2D classification.

5.3 Results and Discussion

To analyze the properties of the lipid monolayers as supporting layers for the proteins, the grid of monolayer sandwiched was first visualized in room temperature EM, then assess it under cryogenic condition. When the EM grid picked up the lipid monolayers, the monolayers covered on both opposite sides will have each head group facing each other, leaving the hydrophobic tails in the air. This coverage can be seen in a darker shade inside the grid square. The broken monolayers can be seen as the brighter area, where the electron beam penetrated through the grid (Figure 5.4).

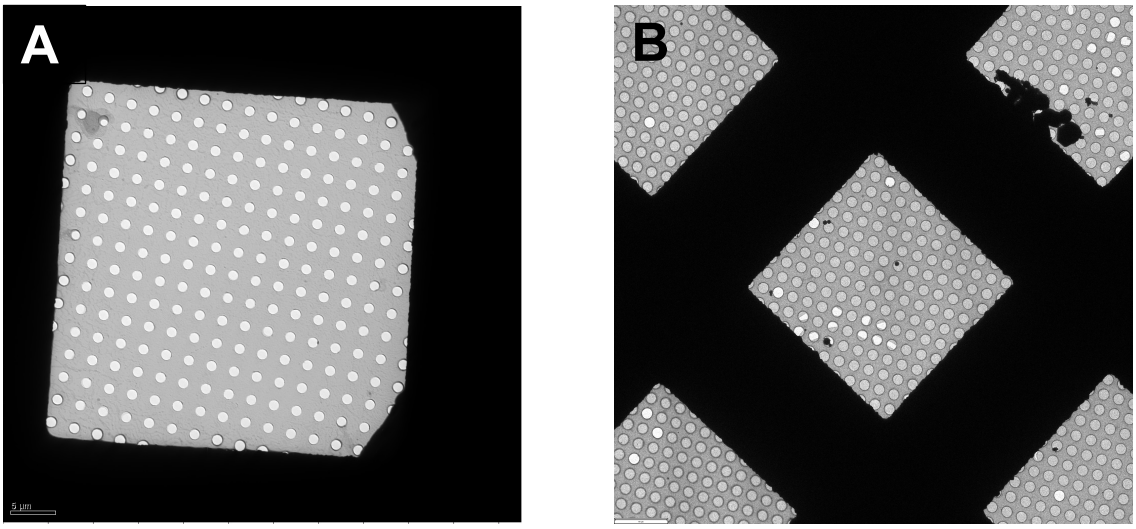

Figure 5.4. Assessment of lipid monolayer coverage on an EM grid. (A) Low magnification image of an EM grid sandwiched in lipid monolayer imaged in room temperature EM. (B) Low magnification image of plunge frozen EM grid sandwiched in lipid monolayer plunge frozen and imaged under cryogenic condition.

Cryo-electron tomography (cryo-ET) of the apoferritin sample in the monolayer sandwich (Figure 5.5) was conducted to analyze the properties of lipid monolayer in vitreous ice. At high tilted angle (60º), the carbon film has ripple like pattern. However, grid holes have nice distributions of particles.

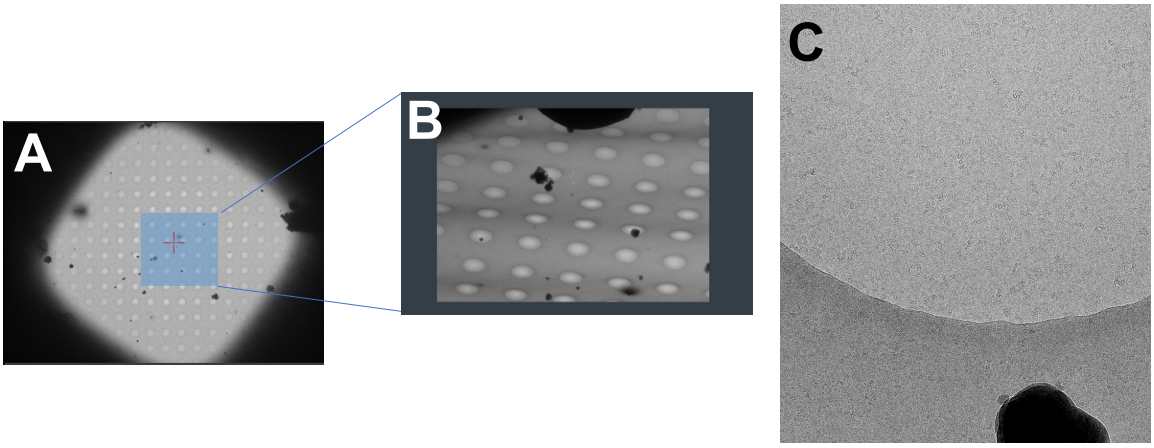

Figure 5.5. Cryo-ET of apoferritin sandwiched in lipid monolayer. (A) Low magnification of apoferritin plunge frozen grid. (B) High-tilted angle (60^{o}) of the grid showing ripple like pattern of the carbon film. (C) High magnification image showing homogenous particle distribution of apoferritin.

The 3D reconstructed tomogram of apoferritin sandwiched in lipid monolayer showed the proteins located at the top and bottom surface of the hole and stay in a thin ice at the center of the hole (Figure 5.6). Furthermore, single-particle cryo-EM data of apoferritin was collected and 2D classification was performed to assess the protein quality, as well as to analyze any difference or difficulty in image processing using this set of data. The cryo image and 2D classes contained high resolution details of apoferritin particles, showcased the feasibility of this method.

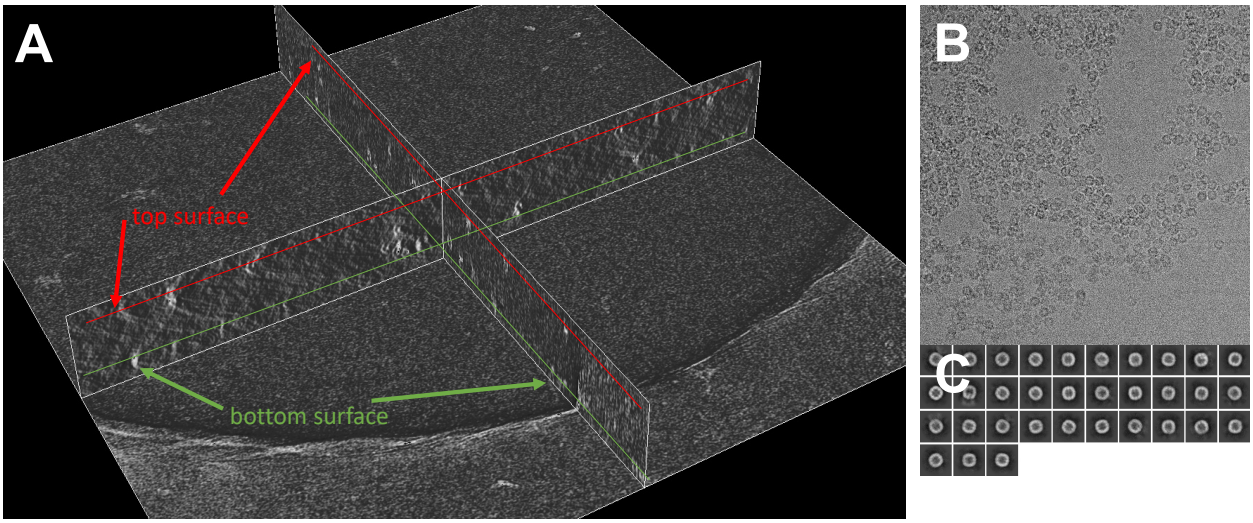

Figure 5.6. Three-dimensional tomogram of apoferritin. (A) 3D reconstruction of apoferritin showing particles appear on top, bottom and middle of the grid hole. (B) Cryo-EM image of apoferritin showing normal distribution of apoferritin in lipid monolayer grid. (C) 2D class averages of cryo-EM data collected, k number is 33.

Furthermore, the cryo-ET of p97 and p47 protein complex was tested to see whether lipid monolayer could solve the preferred orientation problem in cryo-EM technique. The complex of p97 with p47 represented only one preferred view when plunge frozen with conventional plunge freezing technique[20]. Surprisingly, the cryo-ET images showed particles with different orientations in the grid hole (Figure 5.7). This result represented a potential for lipid monolayer method to be a problem solver for currently problem in cryo-EM.

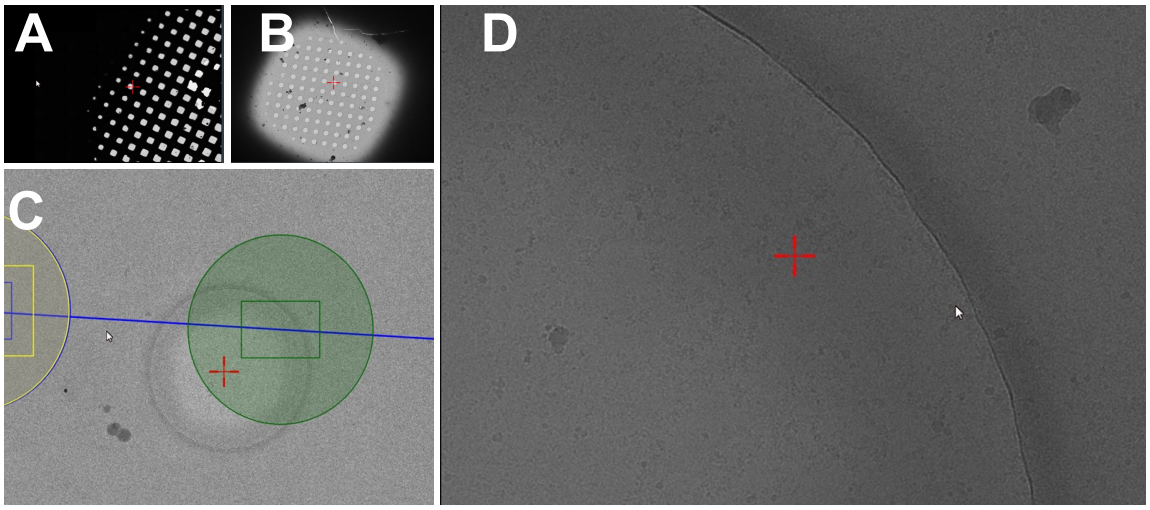

Figure 5.7. Cryo-ET of p97 in complex with p47 protein. (A) Overview of the lipid monolayer plunge frozen grid. (B) Overview of one representative grid square within the frozen grid. (C) View of one representative grid hole in grid square shown in B. The green circle shows the imaging area of the hole. (D) Area imaged within the grid hole showing different orientations of p97-p47.

5.4 Conclusions

The cryo-ET and cryo-EM study of the monolayer sandwiched apoferritin showed a high quality of 2D classification and 3D reconstructed tomogram, showing a good sample preservation between the sandwich without losing significant scattering signals. The cryo-ET data showed that the lipid monolayer is thin enough for regular single-particle cryo-EM processing. Furthermore, the angle distribution of the particle in the cryo-EM and cryo-ET image showed no sign of the preferred orientations. However, more data needed to be collected and high-resolution structure of both apoferritin and p97-p47 complex must be achieved using this technique to prove this method is truly useful. If the structures of both macromolecules were achieved with the implementation of lipid monolayer, this technique can possibly be adapted to use for a routine single-particle cryo-EM sample preparation.

5.5 References

1. Kühlbrandt, W. (2014). The Resolution Revolution. *Science*, *343*(6178), 1443–1444. https://doi.org/10.1126/science.1251652

2. Al-Amoudi, A., Chang, J.-J., Leforestier, A., McDowall, A., Salamin, L. M., Norlén, L. P. O., Richter, K., Blanc, N. S., Studer, D., & Dubochet, J. (2004). Cryo-electron microscopy of vitreous sections. *The EMBO Journal*, *23*(18), 3583–3588. https://doi.org/10.1038/sj.emboj.7600366

3. Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowall, A. W., & Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics*, *21*(2), 129–228. https://doi.org/10.1017/s0033583500004297

4. Lyumkis, D. (2019). Challenges and Opportunities in Cryo-EM Single-Particle Analysis. *Journal of Biological Chemistry*, *294*, jbc.REV118.005602. https://doi.org/10.1074/jbc.REV118.005602

5. D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., & Kühlbrandt, W. (2019). Protein denaturation at the air-water interface and how to prevent it. *ELife*, *8*. https://doi.org/10.7554/elife.42747

6. Glaeser, R. M., & Han, B.-G. (2016). Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophysics Reports*, *3*(1-3), 1–7. https://doi.org/10.1007/s41048-016-0026-3

7. Glaeser, R. M. (1985). Electron Crystallography of Biological Macromolecules. *Annual Review of Physical Chemistry*, *36*(1), 243–275. https://doi.org/10.1146/annurev.pc.36.100185.001331

8. Chen, J., Noble, A. J., Kang, J. Y., & Darst, S. A. (2019). Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *Journal of Structural Biology: X*, *1*, 100005. https://doi.org/10.1016/j.yjsbx.2019.100005

9. Otzen, D. E. (2002). Protein Unfolding in Detergents: Effect of Micelle Structure, Ionic Strength, pH, and Temperature. *Biophysical Journal*, *83*(4), 2219–2230. https://doi.org/10.1016/s0006-3495(02)73982-9

10. Bernal, R. A., & Stock, D. (2004). Three-Dimensional Structure of the Intact Thermus thermophilus H+-ATPase/Synthase by Electron Microscopy. *Structure*, *12*(10), 1789–1798. https://doi.org/10.1016/j.str.2004.07.017

11. Passmore, L. A., & Russo, C. J. (2016). Specimen Preparation for High-Resolution Cryo-EM. *Methods in Enzymology*, *579*, 51–86. https://doi.org/10.1016/bs.mie.2016.04.011

12. Brink, J., Sherman, M. B., Berriman, J., & Chiu, W. (1998). Evaluation of charging on macromolecules in electron cryomicroscopy. *Ultramicroscopy*, *72*(1-2), 41–52. https://doi.org/10.1016/s0304-3991(97)00126-5

13. Wu, K., Wu, D., Zhu, L., & Wu, Y. (2021). Application of Monolayer Graphene and Its Derivative in Cryo-EM Sample Preparation. *International Journal of Molecular Sciences*, *22*(16), 8940. https://doi.org/10.3390/ijms22168940

14. Noble, A. J., Wei, H., Dandey, V. P., Zhang, Z., Tan, Y. Z., Potter, C. S., & Carragher, B. (2018). Reducing effects of particle adsorption to the air–water interface in cryo-EM. *Nature Methods*, *15*(10), 793–795. https://doi.org/10.1038/s41592-018-0139-3

15. Yeager, M., Dryden, K. A., & Ganser-Pornillos, B. K. (2012). Lipid Monolayer and Sparse Matrix Screening for Growing Two-Dimensional Crystals for Electron Crystallography: Methods and Examples. *Methods in Molecular Biology*, 527–537. https://doi.org/10.1007/978-1-62703-176-9_28

16. Kurniawan, J., Ventrici de Souza, J. F., Dang, A. T., Liu, G., & Kuhl, T. L. (2018). Preparation and Characterization of Solid-Supported Lipid Bilayers Formed by Langmuir–Blodgett Deposition: A Tutorial. *Langmuir*, *34*(51), 15622–15639. https://doi.org/10.1021/acs.langmuir.8b03504

17. Andersson, J., & Köper, I. (2019). Biomimetic Membranes. *Comprehensive Nanoscience and Nanotechnology*, 49–64. https://doi.org/10.1016/b978-0-12-803581-8.10447-3

18. *26.9: Phospholipids*. (2016, June 27). Chemistry LibreTexts. https://chem.libretexts.org/Bookshelves/Introductory_Chemistry/Book%3A_Introductory_Chemistry_(CK-12)/26%3A_Biochemistry/26.09%3A_Phospholipids

19. Dietrich, J., & Vénien-Bryan, C. (2005). *Strategies for Two-Dimensional Crystallization of Proteins Using Lipid Monolayers*. https://doi.org/10.1142/p313

20. Rouiller, I., Butel, V. M., Latterich, M., Milligan, R. A., & Wilson-Kubalek, E. M. (2000). A Major Conformational Change in p97 AAA ATPase upon ATP Binding. *Molecular Cell*, *6*(6), 1485–1490. https://doi.org/10.1016/s1097-2765(00)00144-1

21. Peter, B., & Higgins, M. (n.d.). *Monolayer Assay Protocol*. Www2.Mrc-Lmb.cam.ac.uk. Retrieved September 21, 2021, from https://www2.mrc-lmb.cam.ac.uk/groups/hmm/techniqs/mono.html

22. Kremer, J. R., Mastronarde, D. N., & McIntosh, J. Richard. (1996). Computer Visualization of Three-Dimensional Image Data Using IMOD. *Journal of Structural Biology*, *116*(1), 71–76. https://doi.org/10.1006/jsbi.1996.0013

23. Frangakis, A. S., & Hegerl, R. (2001). Noise Reduction in Electron Tomographic Reconstructions Using Nonlinear Anisotropic Diffusion. *Journal of Structural Biology*, *135*(3), 239–250. https://doi.org/10.1006/jsbi.2001.4406

24. Cheng, A., Eng, E. T., Alink, L., Rice, W. J., Jordan, K. D., Kim, L. Y., Potter, C. S., and Carragher, B. (2018) High resolution single particle cryo-electron microscopy using beam-image shift. J. Struct. Biol. 204, 270-275

CHAPTER 6: CONCLUSIONS

During my doctoral research studies, I have focused on the structural characterization of protein complexes using cryo-electron microscopy (EM).

In chapter 2, we used a baculovirus expression method in Sf9 cells to produce the native Nogo receptor complex in vivo for the first time. The possibility of protein-protein interaction is observed when the complex is co-infected in the same culture. Because of the post-translational modification characteristic of the baculovirus expression method, glycosylation-rich proteins may be produced without structural constraints. However, the complex's stability was weaker than expected. As a result, more research is needed to fine-tune the expression and purification procedures in order to increase the stability of Nogo receptor complexes.

The structural investigation of apo polymerase ζ in yeast was successfully conducted in chapter 3. In the absence of DNA, polymerase Polζ reveals a closed conformation. When DNA binds to Polζ, it causes a coordinated movement of local structural motifs involved with the opening of the polymerase's primary DNA-binding channel. Furthermore, we discovered a lysine (K1214) in Rev3 that may act as a gating residue for DNA binding. In the presence of DNA, the side chain of K1214 shifts away from the core channel, creating a larger area for incoming DNA insertion. Furthermore, we learned that even though the folding of the yeast and human Rev7 proteins are almost identical, their dimeric conformations and how they bind target proteins differ significantly. Future research goals will be determining the structure of human Polζ, which will be extremely valuable for the long-term creation of novel anti-cancer medicines.

Chapter 4 is our current study of integrin αvβ8 in complex with pro-TGF-β1-GARP and Fab antibody, which reveals structural flexibility of this macromolecule. With such a wide range of particle orientations, further studies needed to be done on 2D class averages and 3D reconstruction for detailed structure investigation. Likewise, additional cryo-EM data must be collected to extract more particles until the cryo-EM density map gives a clear image of the complex with all components nicely matched. Following that, more 3D refining might be performed to get a high-resolution structure.

In chapter 5, other than structural investigation of protein complexes via cryo-EM, I also worked on lipid monolayer sandwich on an EM grid. This is a technique improvement method developed in the Chiu lab. Our cryo-ET and cryo-EM studies of lipid monolayer sandwiched of apoferritin revealed a high quality of 2D classification and 3D reconstructed tomogram. This indicates good sample preservation between the sandwich without loss of substantial scattering signals. Our first proof-of-principle experiments with the lipid monolayer opens the door of possibility in future implications of the technique as another sample preparation method for cryo-EM study. Once this study is successfully carried out, it can help solving the air-water interface problem that has been a challenge for many cryo-EM scientist over the decades.

REFERENCES

1. Cheng, Y., Grigorieff, N., Penczek, Pawel A., & Walz, T. (2015). A Primer to Single-Particle Cryo-Electron Microscopy. *Cell*, *161*(3), 438–449. https://doi.org/10.1016/j.cell.2015.03.05

2. Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P. M. G. E., Grigoras, I. T., Malinauskaite, L., Malinauskas, T., Miehling, J., Uchański, T., Yu, L., Karia, D., Pechnikova, E. V., de Jong, E., Keizer, J., Bischoff, M., McCormack, J., Tiemeijer, P., & Hardwick, S. W. (2020). Single-particle cryo-EM at atomic resolution. *Nature*, *587*(7832), 152–156. https://doi.org/10.1038/s41586-020-2829-0

3. Faruqi, A. R., & McMullan, G. (2011). Electronic detectors for electron microscopy. *Quarterly Reviews of Biophysics*, *44*(3), 357–390. https://doi.org/10.1017/s0033583511000035

4. Li, X., Mooney, P., Zheng, S., Booth, C. R., Braunfeld, M. B., Gubbens, S., Agard, D. A., & Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic resolution single-particle cryo-EM. *Nature Methods*, *10*(6), 584–590. https://doi.org/10.1038/nmeth.2472

5. Milazzo, A.-C., Cheng, A., Moeller, A., Lyumkis, D., Jacovetty, E., Polukas, J., Ellisman, M. H., Xuong, N.-H., Carragher, B., & Potter, C. S. (2011). Initial evaluation of a direct detection device detector for single particle cryo-electron microscopy. *Journal of Structural Biology*, *176*(3), 404–408. https://doi.org/10.1016/j.jsb.2011.09.002

6. McMullan, G., Chen, S., Henderson, R., & Faruqi, A. R. (2009). Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy*, *109*(9), 1126–1143. https://doi.org/10.1016/j.ultramic.2009.04.002

7. Bai, X., Fernandez, I. S., McMullan, G., & Scheres, S. H. (2013). Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *ELife*, *2*. https://doi.org/10.7554/elife.00461

8. Campbell, Melody G., Cheng, A., Brilot, Axel F., Moeller, A., Lyumkis, D., Veesler, D., Pan, J., Harrison, Stephen C., Potter, Clinton S., Carragher, B., & Grigorieff, N. (2012). Movies of IceEmbedded Particles Enhance Resolution in Electron Cryo-Microscopy. *Structure*, *20*(11), 1823–1828. https://doi.org/10.1016/j.str.2012.08.026

9. Cheng, Y. (2015). Single-Particle Cryo-EM at Crystallographic Resolution. *Cell*, *161*(3), 450–457. https://doi.org/10.1016/j.cell.2015.03.049
10. Dutta, M. (2018). Recent Advances in Single Particle Cryo-electron Microscopy and Cryo electron Tomography to Determine the Structures of Biological Macromolecules. *Journal of the Indian Institute of Science*, *98*(3), 231–245. https://doi.org/10.1007/s41745-018-0087-z

11. Passmore, L. A., & Russo, C. J. (2016). Specimen Preparation for High-Resolution Cryo-EM. *Methods in Enzymology*, 51–86. https://doi.org/10.1016/bs.mie.2016.04.011

12. Syed, T. H., Hughes, T. J., Marsh, K. N., & May, E. F. (2012). Isobaric Heat Capacity Measurements of Liquid Methane, Ethane, and Propane by Differential Scanning Calorimetry at High Pressures and Low Temperatures. *Journal of Chemical & Engineering Data*, *57*(12), 3573–3580. https://doi.org/10.1021/je300762m

13. Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowall, A. W., & Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics*, *21*(2), 129–228. https://doi.org/10.1017/s0033583500004297

14. Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P., & Ranson, N. A. (2016). An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*, *100*, 3–15. https://doi.org/10.1016/j.ymeth.2016.02.017

15. Marturi, N. (2013). *Vision and visual servoing for nanomanipulation and nanocharacterization in scanning electron microscope* [PhD Thesis].

16. Rodenburg, J. (n.d.). *understand astigmatism in the electron lens*. Www.rodenburg.org. Retrieved September 22, 2021, from http://www.rodenburg.org/guide/t600.html

17. Liao, Y. (2006). *Practical Electron Microscopy and Database*. www.globalsino.com/EM/

18. Bammes, B. E., Rochat, R. H., Jakana, J., Chen, D.-H., & Chiu, W. (2012). Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency. *Journal of Structural Biology*, *177*(3), 589–601. https://doi.org/10.1016/j.jsb.2012.01.008

19. Ayato, H., Mori, N., Miyahara, J., & Oikawa, T. (1990). Application of the Imaging Plate to TEM Observation. *J Electron Microsc (Tokyo)*, *39*(6), 444–448. https://doi.org/10.1093/oxfordjournals.jmicro.a050835

20. Fan, G. Y., & Ellisman, M. H. (2000). Digital imaging in transmission electron microscopy. *Journal of Microscopy*, *200*(1), 1–13. https://doi.org/10.1046/j.1365-2818.2000.00737.x

21. Fan, G. Y., & Ellisman, M. H. (1993). High-sensitivity lens-coupled slow-scan CCD camera for transmission electron microscopy. *Ultramicroscopy*, *52*(1), 21–29. https://doi.org/10.1016/0304-3991(93)90019-t

22. *Imaging | Gatan, Inc.* (n.d.). Www.gatan.com. Retrieved September 22, 2021, from https://www.gatan.com/techniques/imaging

23. Frank, J., Shimkin, B., & Dowse, H. (1981). Spider—A modular software system for electron image processing. *Ultramicroscopy*, *6*(4), 343–357. https://doi.org/10.1016/s0304-3991(81)80236-7

24. Ludtke, S. J., Baldwin, P. R., & Chiu, W. (1999). EMAN: Semiautomated Software for High Resolution Single-Particle Reconstructions. *Journal of Structural Biology*, *128*(1), 82–97. https://doi.org/10.1006/jsbi.1999.4174

25. Rohou, A., & Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, *192*(2), 216–221. https://doi.org/10.1016/j.jsb.2015.08.008

26. Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *Journal of Structural Biology*, *193*(1), 1–12. https://doi.org/10.1016/j.jsb.2015.11.003

27. Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., & Ludtke, S. J. (2007).EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology*, *157*(1), 38–46. https://doi.org/10.1016/j.jsb.2006.05.009

28. Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S., & Carragher, B. (2009). DoG Picker and TiltPicker: Software tools to facilitate particle selection in single particle electron microscopy. *Journal of Structural Biology*, *166*(2), 205–213. https://doi.org/10.1016/j.jsb.2009.01.004

29. Scheres, S. H. W. (2015). Semi-automated selection of cryo-EM particles in RELION-1.3. *Journal of Structural Biology*, *189*(2), 114–122. https://doi.org/10.1016/j.jsb.2014.11.010

30. Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

31. Garbade, M. (2018, September 12). *Understanding K-means Clustering in Machine Learning*. Towards Data Science; Towards Data Science.

https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1

32. Xu, Y., Wu, J., Yin, C.-C., & Mao, Y. (2016). Unsupervised Cryo-EM Data Clustering through Adaptively Constrained K-Means Algorithm. *PLOS ONE*, *11*(12), e0167765. https://doi.org/10.1371/journal.pone.0167765

33. McKay, D. J. C. (2003, October). *Information Theory, Inference and Learning Algorithms | Pattern recognition and machine learning*. Cambridge University Press. http://www.cambridge.org/0521642981

34. Renault, L., & Stahlberg, H. (2007). Three-Dimensional Electron Microscopy of Macro-molecular Assemblies: Visualization of Biological Molecules in Their Native State. By Joachim Frank. Oxford University Press, New York (2006). ISBN 0195182189; softback; 432 pages. *Scanning*, *29*(1), 37–37. https://doi.org/10.1002/sca.20031

35. Pintilie, G. (n.d.). *Greg Pintilie*. People.csail.mit.edu. https://people.csail.mit.edu/gdp/cryoem.html

36. *Back Projection*. (n.d.). Www.impactscan.org. http://www.impactscan.org/slides/impactcourse/basic_principles_of_ct/img12.html

37. Wingfield, P. T. (2015). Overview of the Purification of Recombinant Proteins. *Current Protocols in Protein Science*, 6.1.1–6.1.35. https://doi.org/10.1002/0471140864.ps0601s80

38. Scarff, C. A., Fuller, M. J. G., Thompson, R. F., & Iadaza, M. G. (2018). Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems. *Journal of Visualized Experiments*, *132*. https://doi.org/10.3791/57199

39. Mezu-Ndubuisi, O. J., & Maheshwari, A. (2020). The role of integrins in inflammation and angiogenesis. *Pediatric Research*. https://doi.org/10.1038/s41390-020-01177-9

40. Rouiller, I., Butel, V. M., Latterich, M., Milligan, R. A., & Wilson-Kubalek, E. M. (2000). A Major Conformational Change in p97 AAA ATPase upon ATP Binding. *Molecular Cell*, *6*(6), 1485–1490. https://doi.org/10.1016/s1097-2765(00)00144-1

41. Tan, Y. Z., Baldwin, P. R., Davis, J. H., Williamson, J. R., Potter, C. S., Carragher, B., & Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods*, *14*(8), 793–796. https://doi.org/10.1038/nmeth.434

42. Schwab, J. M., Tuli, S. K., & Failli, V. (2006). The Nogo receptor complex: confining molecules to molecular mechanisms. *Trends in Molecular Medicine*, *12*(7), 293–297. https://doi.org/10.1016/j.molmed.2006.05.001

43. Chamberlain, K. A., Nanescu, S. E., Psachoulia, K., & Huang, J. K. (2016). Oligodendrocyte regeneration: Its significance in myelin replacement and neuroprotection in multiple sclerosis. *Neuropharmacology*, *110*(Pt B), 633–643. https://doi.org/10.1016/j.neuropharm.2015.10.010

44. Gordon, T., Udina, E., Verge, V. M. K., & de Chaves, E. I. P. (2009). Brief Electrical Stimulation Accelerates Axon Regeneration in the Peripheral Nervous System and Promotes Sensory Axon Regeneration in the Central Nervous System. *Motor Control*, *13*(4), 412–441. https://doi.org/10.1123/mcj.13.4.412

45. Rao, S. N. R., & Pearse, D. D. (2016). Regulating Axonal Responses to Injury: The Intersection between Signaling Pathways Involved in Axon Myelination and The Inhibition of Axon Regeneration. *Frontiers in Molecular Neuroscience*, *9*. https://doi.org/10.3389/fnmol.2016.00033

46. Akbik, F., Cafferty, W. B. J., & Strittmatter, S. M. (2012). Myelin associated inhibitors: A link between injury-induced and experience-dependent plasticity. *Experimental Neurology*, *235*(1), 43–52. https://doi.org/10.1016/j.expneurol.2011.06.006

47. Yiu, G., & He, Z. (2006). Glial inhibition of CNS axon regeneration. *Nature Reviews Neuroscience*, *7*(8), 617–627. https://doi.org/10.1038/nrn1956

48. Silver, J., Schwab, M. E., & Popovich, P. G. (2014). Central Nervous System Regenerative Failure: Role of Oligodendrocytes, Astrocytes, and Microglia. *Cold Spring Harbor Perspectives in Biology*, *7*(3), a020602. https://doi.org/10.1101/cshperspect.a020602

49. McDonald, L., Bandtlow, C., & Reindl, M. (2011). Targeting the Nogo Receptor Complex in Diseases of the Central Nervous System. *Current Medicinal Chemistry*, *18*(2), 234–244. https://doi.org/10.2174/092986711794088326

50. Meabon, J. S., De Laat, R., Ieguchi, K., Wiley, J. C., Hudson, M. P., & Bothwell, M. (2015). LINGO-1 Protein Interacts with the p75 Neurotrophin Receptor in Intracellular Membrane Compartments. *Journal of Biological Chemistry*, *290*(15), 9511–9520. https://doi.org/10.1074/jbc.m114.608018

51. Mi, S., Lee, X., Shao, Z., Thill, G., Ji, B., Relton, J., Levesque, M., Allaire, N., Perrin, S., Sands, B., Crowell, T., Cate, R. L., McCoy, J. M., & Pepinsky, R. B.

(2004). LINGO-1 is a component of the Nogo-66 receptor/p75 signaling complex. *Nature Neuroscience*, *7*(3), 221–228. https://doi.org/10.1038/nn1188

52. Vasudevan, S. V., Schulz, J., Zhou, C., & Cocco, M. J. (2010). Protein folding at the membrane interface, the structure of Nogo-66 requires interactions with a phosphocholine surface. *Proceedings of the National Academy of Sciences*, *107*(15), 6847–6851. https://doi.org/10.1073/pnas.0911817107

53. Barton, W. A. (2003). Structure and axon outgrowth inhibitor binding of the Nogo-66 receptor and related proteins. *The EMBO Journal*, *22*(13), 3291–3302. https://doi.org/10.1093/emboj/cdg325

54. Mosyak, L., Wood, A., Dwyer, B., Buddha, M., Johnson, M., Aulabaugh, A., Zhong, X., Presman, E., Benard, S., Kelleher, K., Wilhelm, J., Stahl, M. L., Kriz, R., Gao, Y., Cao, Z., Ling, H.-P., Pangalos, M. N., Walsh, F. S., & Somers, W. S. (2006). The Structure of the Lingo-1 Ectodomain, a Module Implicated in Central Nervous System Repair Inhibition. *Journal of Biological Chemistry*, *281*(47), 36378–36390. https://doi.org/10.1074/jbc.m607314200

55. Gong, Y., Cao, P., Yu, H., & Jiang, T. (2008). Crystal structure of the neurotrophin-3 and p75NTR symmetrical complex. *Nature*, *454*(7205), 789–793. https://doi.org/10.1038/nature07089

56. Liepinsh, E. (1997). NMR structure of the death domain of the p75 neurotrophin receptor. *The EMBO Journal*, *16*(16), 4999–5005. https://doi.org/10.1093/emboj/16.16.4999

57. Contreras-Gómez, A., Sánchez-Mirón, A., García-Camacho, F., Molina-Grima, E., & Chisti, Y. (2013). Protein production using the baculovirus-insect cell expression system. *Biotechnology Progress*, *30*(1), 1–18. https://doi.org/10.1002/btpr.1842

58. Jarvis, D. L. (2003). Developing baculovirus-insect cell expression systems for humanized recombinant glycoprotein production. *Virology*, *310*(1), 1–7. https://doi.org/10.1016/s0042-6822(03)00120-x

59. He, X. L., Bazan, J. Fernando., McDermott, G., Park, J. B., Wang, K., Tessier-Lavigne, M., He, Z., & Garcia, K. Christopher. (2003). Structure of the Nogo Receptor Ectodomain. *Neuron*, *38*(2), 177–185. https://doi.org/10.1016/s0896-6273(03)00232-0

60. Feng, D., Kim, T., Özkan, E., Light, M., Torkin, R., Teng, K. K., Hempstead, B. L., & Garcia, K. C. (2010). Molecular and Structural Insight into proNGF Engagement of p75NTR and Sortilin. *Journal of Molecular Biology*, *396*(4), 967–984. https://doi.org/10.1016/j.jmb.2009.12.030

61. Johnson, P. W., Attia, J., Richardson, C. D., Roder, J. C., & Dunn, R. J. (1989). Synthesis of soluble myelin-associated glycoprotein in insect and mammalian cells. *Gene*, *77*(2), 287–296. https://doi.org/10.1016/0378-1119(89)90076-0

62. Ohi, M., Li, Y., Cheng, Y., & Walz, T. (2004). Negative staining and image classification — powerful tools in modern electron microscopy. *Biological Procedures Online*, *6*(1), 23–34. https://doi.org/10.1251/bpo70

63. Scheres, S. H. W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

64. Pronker, M. F., Lemstra, S., Snijder, J., Heck, A. J. R., Thies-Weesie, D. M. E., Pasterkamp, R. J., & Janssen, B. J. C. (2016). Structural basis of myelin-associated glycoprotein adhesion and signalling. *Nature Communications*, *7*(1). https://doi.org/10.1038/ncomms13584

65. George, J. M., Jin, H., Woods, W. S., & Clayton, D. F. (1995). Characterization of a novel protein regulated during the critical period for song learning in the zebra finch. *Neuron*, *15*(2), 361–372. https://doi.org/10.1016/0896-6273(95)90040-3

66. Cafferty, W. B. J., & Strittmatter, S. M. (2006). The Nogo-Nogo Receptor Pathway Limits a Spectrum of Adult CNS Axonal Growth. *Journal of Neuroscience*, *26*(47), 12242–12250. https://doi.org/10.1523/jneurosci.3827-06.2006

67. Kempf, A., & Schwab, M. E. (2013). Nogo-A Represses Anatomical and Synaptic Plasticity in the Central Nervous System. *Physiology*, *28*(3), 151–163. https://doi.org/10.1152/physiol.00052.2012

68. Bradl, M., & Lassmann, H. (2009). Oligodendrocytes: biology and pathology. *Acta Neuropathologica*, *119*(1), 37–53. https://doi.org/10.1007/s00401-009-0601-5

69. Baynton, K., & Fuchs, R. P. P. (2000). Lesions in DNA: hurdles for polymerases. *Trends in Biochemical Sciences*, *25*(2), 74–79. https://doi.org/10.1016/s0968-0004(99)01524-8

70. Kunz, B. A., Straffon, A. F. L., & Vonarx, E. J. (2000). DNA damage-induced mutation: tolerance via translesion synthesis. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *451*(1-2), 169–185. https://doi.org/10.1016/s0027-5107(00)00048-8

71. Marians, K. J. (2018). Lesion Bypass and the Reactivation of Stalled Replication Forks. *Annual Review of Biochemistry*, *87*(1), 217–238. https://doi.org/10.1146/annurev-biochem-062917-011921

72. Prakash, S., Johnson, R. E., & Prakash, L. (2005). EUKARYOTIC TRANSLESION SYNTHESIS DNA POLYMERASES: Specificity of Structure and Function. *Annual Review of Biochemistry*, *74*(1), 317–353. https://doi.org/10.1146/annurev.biochem.74.082803.133250

73. Acharya, N., Haracska, L., Johnson, R. E., Unk, I., Prakash, S., & Prakash, L. (2005). Complex Formation of Yeast Rev1 and Rev7 Proteins: a Novel Role for the Polymerase-Associated Domain. *Molecular and Cellular Biology*, *25*(21), 9734–9740. https://doi.org/10.1128/mcb.25.21.9734-9740.2005

74. Gómez-Llorente, Y., Malik, R., Jain, R., Choudhury, J., Johnson, Robert E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., & Aggarwal, Aneel K. (2013). The Architecture of Yeast DNA Polymerase ζ. *Cell Reports*, *5*(1), 79–86. https://doi.org/10.1016/j.celrep.2013.08.046

75. Goodman, M. F. (2002). Error-Prone Repair DNA Polymerases in Prokaryotes and Eukaryotes. *Annual Review of Biochemistry*, *71*(1), 17–50. https://doi.org/10.1146/annurev.biochem.71.083101.124707

76. Leung, W., Baxley, R., Moldovan, G.-L., & Bielinsky, A.-K. (2018). Mechanisms of DNA Damage Tolerance: Post-Translational Regulation of PCNA. *Genes*, *10*(1), 10. https://doi.org/10.3390/genes10010010

77. Goodman, M. F., & Woodgate, R. (2013). Translesion DNA Polymerases. *Cold Spring Harbor Perspectives in Biology*, *5*(10), a010363–a010363. https://doi.org/10.1101/cshperspect.a010363

78. Acharya, N., Johnson, R. E., Prakash, S., & Prakash, L. (2006). Complex Formation with Rev1 Enhances the Proficiency of Saccharomyces cerevisiae DNA Polymerase ζ for Mismatch Extension and for Extension Opposite from DNA Lesions. *Molecular and Cellular Biology*, *26*(24), 9555–9563. https://doi.org/10.1128/mcb.01671-06

79. Acharya, N., Johnson, R. E., Pages, V., Prakash, L., & Prakash, S. (2009). Yeast Rev1 protein promotes complex formation of DNA polymerase with Pol32 subunit of DNA polymerase. *Proceedings of the National Academy of Sciences*, *106*(24), 9631–9636. https://doi.org/10.1073/pnas.0902175106

80. Makarova, A. V., Stodola, J. L., & Burgers, P. M. (2012). A four-subunit DNA polymerase ζ complex containing Pol δ accessory subunits is essential for PCNA-mediated mutagenesis. *Nucleic Acids Research*, *40*(22), 11618–11626. https://doi.org/10.1093/nar/gks948

81. Makarova, A. V., & Burgers, P. M. (2015). Eukaryotic DNA polymerase ζ. *DNA Repair*, *29*, 47–55. https://doi.org/10.1016/j.dnarep.2015.02.012

82. Yamanaka, K., Chatterjee, N., Hemann, M. T., & Walker, G. C. (2017). Inhibition of mutagenic translesion synthesis: A possible strategy for improving chemotherapy? *PLOS Genetics*, *13*(8), e1006842. https://doi.org/10.1371/journal.pgen.1006842

83. Cui, G., Botuyan, M. V., & Mer, G. (2018). Structural Basis for the Interaction of Mutasome Assembly Factor REV1 with Ubiquitin. *Journal of Molecular Biology*, *430*(14), 2042–2050. https://doi.org/10.1016/j.jmb.2018.05.017

84. Taniguchi, T. (2019). REV1-POL ζ Inhibition and Cancer Therapy. *Molecular Cell*, *75*(3), 419–420. https://doi.org/10.1016/j.molcel.2019.07.012

85. Nair, D. T. (2005). Rev1 Employs a Novel Mechanism of DNA Synthesis Using a Protein Template. *Science*, *309*(5744), 2219–2222. https://doi.org/10.1126/science.1116336

86. Nelson, J. R., Lawrence, C. W., & Hinkle, D. C. (1996). Deoxycytidyl transferase activity of yeast REV1 protein. *Nature*, *382*(6593), 729–731. https://doi.org/10.1038/382729a0

87. Chang, D. J., & Cimprich, K. A. (2009). DNA damage tolerance: when it's OK to make mistakes. *Nature Chemical Biology*, *5*(2), 82–90. https://doi.org/10.1038/nchembio.139

88. Waters, L. S., Minesinger, B. K., Wiltrout, M. E., D'Souza, S., Woodruff, R. V., & Walker, G. C. (2009). Eukaryotic Translesion Polymerases and Their Roles and Regulation in DNA Damage Tolerance. *Microbiology and Molecular Biology Reviews : MMBR*, *73*(1), 134–154. https://doi.org/10.1128/MMBR.00034-08

89. Bezalel-Buch, R., Cheun, Y. K., Roy, U., Schärer, O. D., & Burgers, P. M. (2020). Bypass of DNA interstrand crosslinks by a Rev1–DNA polymerase ζ complex. *Nucleic Acids Research*, *48*(15), 8461–8473. https://doi.org/10.1093/nar/gkaa580

90. Martin, S. K., & Wood, R. D. (2019). DNA polymerase ζ in DNA replication and repair. *Nucleic Acids Research*, *47*(16), 8348–8361. https://doi.org/10.1093/nar/gkz705

91. Lawrence, C. W. (2004). Cellular Functions of DNA Polymerase ζ and Rev1 Protein. *Advances in Protein Chemistry*, 167–203. https://doi.org/10.1016/s0065-3233(04)69006-1

92. Morrison, A., Christensen, R. B., Alley, J., Beck, A. K., Bernstine, E. G., Lemontt, J. F., & Lawrence, C. W. (1989). REV3, a Saccharomyces cerevisiae gene whose

function is required for induced mutagenesis, is predicted to encode a nonessential DNA polymerase. *Journal of Bacteriology*, *171*(10), 5659–5667. https://doi.org/10.1128/jb.171.10.5659-5667.1989

93. Nelson, J. R., Lawrence, C. W., & Hinkle, D. C. (1996a). Thymine-Thymine Dimer Bypass by Yeast DNA Polymerase zeta. *Science*, *272*(5268), 1646–1649. https://doi.org/10.1126/science.272.5268.1646

94. Zhong, X., Garg, P., Stith, C. M., McElhinny, S. A. N., Kissling, G. E., Burgers, P. M. J., & Kunkel, T. A. (2006). The fidelity of DNA synthesis by yeast DNA polymerase zeta alone and with accessory proteins. *Nucleic Acids Research*, *34*(17), 4731–4742. https://doi.org/10.1093/nar/gkl465

95. Malik, R., Kopylov, M., Gomez-Llorente, Y., Jain, R., Johnson, R. E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., & Aggarwal, A. K. (2020). Structure and mechanism of B-family DNA polymerase ζ specialized for translesion DNA synthesis. *Nature Structural & Molecular Biology*, *27*(10), 913–924. https://doi.org/10.1038/s41594-020-0476-7

96. Scheres, S. H. W., & Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. *Nature Methods*, *9*(9), 853–854. https://doi.org/10.1038/nmeth.2115

97. Nakane, T., Kimanius, D., Lindahl, E., & Scheres, S. H. (2018). Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *ELife*, *7*. https://doi.org/10.7554/elife.36861

98. Netz, D. J. A., Stith, C. M., Stümpfig, M., Köpf, G., Vogel, D., Genau, H. M., Stodola, J. L., Lill, R., Burgers, P. M. J., & Pierik, A. J. (2011). Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nature Chemical Biology*, *8*(1), 125–132. https://doi.org/10.1038/nchembio.721

99. Jain, R., Rice, W. J., Malik, R., Johnson, R. E., Prakash, L., Prakash, S., Ubarretxena-Belandia, I., & Aggarwal, A. K. (2019). Cryo-EM structure and dynamics of eukaryotic DNA polymerase δ holoenzyme. *Nature Structural & Molecular Biology*, *26*(10), 955–962. https://doi.org/10.1038/s41594-019-0305-z

100. Johnson, R. E., Prakash, L., & Prakash, S. (2012). Pol31 and Pol32 subunits of yeast DNA polymerase are also essential subunits of DNA polymerase. *Proceedings of the National Academy of Sciences*, *109*(31), 12455–12460. https://doi.org/10.1073/pnas.1206052109

101. Brandão, L. N., Ferguson, R., Santoro, I., Jinks-Robertson, S., & Sclafani, R. A. (2014). The Role of Dbf4-Dependent Protein Kinase in DNA Polymerase ζ-Dependent Mutagenesis in Saccharomyces cerevisiae. *Genetics*, *197*(4), 1111–1122. https://doi.org/10.1534/genetics.114.165308

102. Fattah, F. J., Hara, K., Fattah, K. R., Yang, C., Wu, N., Warrington, R., Chen, D. J., Zhou, P., Boothman, D. A., & Yu, H. (2014). The Transcription Factor TFII-I Promotes DNA Translesion Synthesis and Genomic Stability. *PLoS Genetics*, *10*(6), e1004419. https://doi.org/10.1371/journal.pgen.1004419

103. Hara, K., Hashimoto, H., Murakumo, Y., Kobayashi, S., Kogame, T., Unzai, S., Akashi, S., Takeda, S., Shimizu, T., & Sato, M. (2010). Crystal Structure of Human REV7 in Complex with a Human REV3 Fragment and Structural Implication of the Interaction between DNA Polymerase ζ and REV1. *Journal of Biological Chemistry*, *285*(16), 12299–12307. https://doi.org/10.1074/jbc.m109.092403

104. Lemontt J. F. (1971). Mutants of yeast defective in mutation induced by ultraviolet light. *Genetics, 68*(1), 21–33. https://doi.org/10.1093/genetics/68.1.21

105. Bhat, A., Wu, Z., Maher, V. M., McCormick, J. J., & Xiao, W. (2015). Rev7/Mad2B plays a critical role in the assembly of a functional mitotic spindle. *Cell Cycle*, *14*(24), 3929–3938. https://doi.org/10.1080/15384101.2015.1120922

106. Murakumo, Y., Ogura, Y., Ishii, H., Numata, S., Ichihara, M., Croce, C. M., Fishel, R., & Takahashi, M. (2001). Interactions in the Error-prone Postreplication Repair Proteins hREV1, hREV3, and hREV7. *Journal of Biological Chemistry*, *276*(38), 35644–35651. https://doi.org/10.1074/jbc.m102051200

107. Rizzo, A. A., Vassel, F.-M., Chatterjee, N., D'Souza, S., Li, Y., Hao, B., Hemann, M. T., Walker, G. C., & Korzhnev, D. M. (2018). Rev7 dimerization is important for assembly and function of the Rev1/Polζ translesion synthesis complex. *Proceedings of the National Academy of Sciences*, *115*(35), E8191–E8200. https://doi.org/10.1073/pnas.1801149115

108. Guo, C. (2003). Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *The EMBO Journal*, *22*(24), 6621–6630. https://doi.org/10.1093/emboj/cdg626

109. Masuda, Y., Ohmae, M., Masuda, K., & Kamiya, K. (2003). Structure and Enzymatic Properties of a Stable Complex of the Human REV1 and REV7 Proteins. *Journal of Biological Chemistry*, *278*(14), 12356–12360. https://doi.org/10.1074/jbc.m211765200

110. Rosenberg, S. C., & Corbett, K. D. (2015). The multifaceted roles of the HORMA domain in cellular signaling. *The Journal of Cell Biology*, *211*(4), 745–755. https://doi.org/10.1083/jcb.201509076

111. Miniowitz-Shemtov, S., Eytan, E., Kaisari, S., Sitry-Shevah, D., & Hershko, A. (2015). Mode of interaction of TRIP13 AAA-ATPase with the Mad2-binding protein p31comet and with mitotic checkpoint complexes. *Proceedings of the National Academy of Sciences*, *112*(37), 11536–11540. https://doi.org/10.1073/pnas.1515358112

112. Zhang, L., Yang, S.-H., & Sharrocks, A. D. (2007). Rev7/MAD2B Links c-Jun N-Terminal Protein Kinase Pathway Signaling to Activation of the Transcription Factor Elk-1. *Molecular and Cellular Biology*, *27*(8), 2861–2869. https://doi.org/10.1128/mcb.02276-06

113. Dai, Y., Zhang, F., Wang, L., Shan, S., Gong, Z., & Zhou, Z. (2020). Structural basis for shieldin complex subunit 3–mediated recruitment of the checkpoint protein REV7 during DNA double-strand break repair. *Journal of Biological Chemistry*, *295*(1), 250–262. https://doi.org/10.1074/jbc.ra119.011464

114. Liang, L., Feng, J., Zuo, P., Yang, J., Lu, Y., & Yin, Y. (2020). Molecular basis for assembly of the shieldin complex and its implications for NHEJ. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15879-5

115. Xie, W., Wang, S., Wang, J., de la Cruz, M. J., Xu, G., Scaltriti, M., & Patel, D. J. (2021). Molecular mechanisms of assembly and TRIP13-mediated remodeling of the human Shieldin complex. *Proceedings of the National Academy of Sciences*, *118*(8), e2024512118. https://doi.org/10.1073/pnas.2024512118

116. Wang, D., Ma, J., Botuyan, M. V., Cui, G., Yan, Y., Ding, D., Zhou, Y., Krueger, E. W., Pei, J., Wu, X., Wang, L., Pei, H., McNiven, M. A., Ye, D., Mer, G., & Huang, H. (2021). ATM-phosphorylated SPOP contributes to 53BP1 exclusion from chromatin during DNA replication. *Science Advances*, *7*(25), eabd9208. https://doi.org/10.1126/sciadv.abd9208

117. Gupta, R., Somyajit, K., Narita, T., Maskey, E., Stanlie, A., Kremer, M., Typas, D., Lammers, M., Mailand, N., Nussenzweig, A., Lukas, J., & Choudhary, C. (2018). DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell*, *173*(4), 972-988.e23. https://doi.org/10.1016/j.cell.2018.03.050

118. Noordermeer, S. M., Adam, S., Setiaputra, D., Barazas, M., Pettitt, S. J., Ling, A. K., Olivieri, M., Álvarez-Quilón, A., Moatti, N., Zimmermann, M., Annunziato, S., Krastev, D. B., Song, F., Brandsma, I., Frankum, J., Brough, R., Sherker, A., Landry, S., Szilard, R. K., & Munro, M. M. (2018). The shieldin complex mediates

53BP1-dependent DNA repair. *Nature*, *560*(7716), 117–121. https://doi.org/10.1038/s41586-018-0340-7

119. Wojtaszek, J. L., Chatterjee, N., Najeeb, J., Ramos, A., Lee, M., Bian, K., Xue, J. Y., Fenton, B. A., Park, H., Li, D., Hemann, M. T., Hong, J., Walker, G. C., & Zhou, P. (2019). A Small Molecule Targeting Mutagenic Translesion Synthesis Improves Chemotherapy. *Cell*, *178*(1), 152-159.e11. https://doi.org/10.1016/j.cell.2019.05.028

120. Chatterjee, N., Whitman, M. A., Harris, C. A., Min, S. M., Jonas, O., Lien, E. C., Luengo, A., Vander Heiden, M. G., Hong, J., Zhou, P., Hemann, M. T., & Walker, G. C. (2020). REV1 inhibitor JH-RE-06 enhances tumor cell response to chemotherapy by triggering senescence hallmarks. *Proceedings of the National Academy of Sciences*, *117*(46), 28918–28921. https://doi.org/10.1073/pnas.2016064117

121. Nesvizhskii, A. I., Keller, A., Kolker, E., & Aebersold, R. (2003). A Statistical Model for Identifying Proteins by Tandem Mass Spectrometry. *Analytical Chemistry*, *75*(17), 4646–4658. https://doi.org/10.1021/ac0341261

122. Keller, A., Nesvizhskii, A. I., Kolker, E., & Aebersold, R. (2002). Empirical Statistical Model To Estimate the Accuracy of Peptide Identifications Made by MS/MS and Database Search. *Analytical Chemistry*, *74*(20), 5383–5392. https://doi.org/10.1021/ac025747h

123. Guo, D. (2004). Translesion synthesis of acetylaminofluorene-dG adducts by DNA polymerase is stimulated by yeast Rev1 protein. *Nucleic Acids Research*, *32*(3), 1122–1130. https://doi.org/10.1093/nar/gkh279

124. Otwinowski, Z., & Minor, W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, 307–326. https://doi.org/10.1016/s0076-6879(97)76066-x

125. 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. *Acta Crystallographica Section D Biological Crystallography*, *66*(2), 213–221. https://doi.org/10.1107/s0907444909052925

126. Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography*, *60*(12), 2126–2132. https://doi.org/10.1107/s0907444904019158

127. DeLano, W. L. (2002) The PyMOL molecular graphics system. http://www. pymol. org

128. 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. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

129. Ohi, M., Li, Y., Cheng, Y., & Walz, T. (2004). Negative staining and image classification- powerful tools in modern electron microscopy. *Biological Procedures Online*, *6*(1), 23–34. https://doi.org/10.1251/bpo70

130. Scheres, S. H. W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

131. Chiu, P.-L., Li, X., Li, Z., Beckett, B., Brilot, A. F., Grigorieff, N., Agard, D. A., Cheng, Y., & Walz, T. (2015). Evaluation of super-resolution performance of the K2 electron-counting camera using 2D crystals of aquaporin-0. *Journal of Structural Biology*, *192*(2), 163–173. https://doi.org/10.1016/j.jsb.2015.08.015

132. Cheng, A., Eng, E. T., Alink, L., Rice, W. J., Jordan, K. D., Kim, L. Y., Potter, C. S., & Carragher, B. (2018). High resolution single particle cryo-electron microscopy using beam-image shift. *Journal of Structural Biology*, *204*(2), 270–275. https://doi.org/10.1016/j.jsb.2018.07.015

133. Mastronarde, D. N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology*, *152*(1), 36–51. https://doi.org/10.1016/j.jsb.2005.07.007

134. Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

135. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A. J., & Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods*, *16*(11), 1153–1160. https://doi.org/10.1038/s41592-019-0575-8

136. Punjani, A., Zhang, H., & Fleet, D. J. (2020). Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nature Methods*, *17*(12), 1214–1221. https://doi.org/10.1038/s41592-020-00990-8

137. Cardone, G., Heymann, J. B., & Steven, A. C. (2013). One number does not fit all: Mapping local variations in resolution in cryo-EM reconstructions. *Journal of Structural Biology*, *184*(2), 226–236. https://doi.org/10.1016/j.jsb.2013.08.002

138. Tan, Y. Z., Baldwin, P. R., Davis, J. H., Williamson, J. R., Potter, C. S., Carragher, B., & Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods*, *14*(8), 793–796. https://doi.org/10.1038/nmeth.4347

139. Pintilie, G. D., Zhang, J., Goddard, T. D., Chiu, W., & Gossard, D. C. (2010). Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *Journal of Structural Biology*, *170*(3), 427–438. https://doi.org/10.1016/j.jsb.2010.03.007

140. Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development ofCoot. *Acta Crystallographica Section D Biological Crystallography*, *66*(4), 486–501. https://doi.org/10.1107/s0907444910007493

141. Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., & Ferrin, T. E. (2017). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science*, *27*(1), 14–25. https://doi.org/10.1002/pro.3235

142. Potter, V. R. (1974). Probabilistic Aspects of the Human Cybernetic Machine. *Perspectives in Biology and Medicine*, *17*(2), 164–183. https://doi.org/10.1353/pbm.1974.0023

143. LORD, B. I. (1988). Feedback regulators in normal and tumour tissues. *Journal of Cell Science*, *1988*(Supplement_10), 231–242. https://doi.org/10.1242/jcs.1988.supplement_10.16

144. Giancotti, F. G. (1997). Integrin signaling: specificity and control of cell survival and cell cycle progression. *Current Opinion in Cell Biology*, *9*(5), 691–700. https://doi.org/10.1016/s0955-0674(97)80123-8

145. Crawford, S. E., Stellmach, V., Murphy-Ullrich, J. E., Ribeiro, S. M. F., Lawler, J., Hynes, R. O., Boivin, G. P., & Bouck, N. (1998). Thrombospondin-1 Is a Major Activator of TGF-β1 In Vivo. *Cell*, *93*(7), 1159–1170. https://doi.org/10.1016/s0092-8674(00)81460-9

146. Stockis, J., Liénart, S., Colau, D., Collignon, A., Nishimura, S. L., Sheppard, D., Coulie, P. G., & Lucas, S. (2017). Blocking immunosuppression by human Tregs in vivo with antibodies targeting integrin αVβ8. *Proceedings of the National Academy of Sciences*, *114*(47), E10161–E10168. https://doi.org/10.1073/pnas.1710680114

147. Aluwihare, P., Mu, Z., Zhao, Z., Yu, D., Weinreb, P. H., Horan, G. S., Violette, S. M., &Munger, J. S. (2009). Mice that lack activity of αvβ6- and αvβ8-integrins reproduce the abnormalities of Tgfb1- and Tgfb3-null mice. *Journal of Cell Science*, *122*(2), 227–232. https://doi.org/10.1242/jcs.035246

148. Kudo, M., Melton, A. C., Chen, C., Engler, M. B., Huang, K. E., Ren, X., Wang, Y., Bernstein, X., Li, J. T., Atabai, K., Huang, X., & Sheppard, D. (2012). IL-17A produced by αβ T cells drives airway hyper-responsiveness in mice and enhances mouse and human airway smooth muscle contraction. *Nature Medicine*, *18*(4), 547–554. https://doi.org/10.1038/nm.2684

149. Humphries, J. D. (2006). Integrin ligands at a glance. *Journal of Cell Science*, *119*(19), 3901–3903. https://doi.org/10.1242/jcs.03098

150. Mu, D., Cambier, S., Fjellbirkeland, L., Baron, J. L., Munger, J. S., Kawakatsu, H., Sheppard,D., Broaddus, V. C., & Nishimura, S. L. (2002). The integrin αvβ8 mediates epithelial homeostasis through MT1-MMP–dependent activation of TGF-β1. *Journal of Cell Biology*, *157*(3), 493–507. https://doi.org/10.1083/jcb.200109100

151. Ozawa, A., Sato, Y., Imabayashi, T., Uemura, T., Takagi, J., & Sekiguchi, K. (2016). Molecular Basis of the Ligand Binding Specificity of αvβ8 Integrin. *Journal of Biological Chemistry*, *291*(22), 11551–11565. https://doi.org/10.1074/jbc.m116.719138

152. Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

153. Patel, A., Toso, D., Litvak, A., & Nogales, E. (2021). Efficient graphene oxide coating improves cryo-EM sample preparation and data collection from tilted grids. *BioRXiv*. https://doi.org/10.1101/2021.03.08.434344

154. Wang, F., Liu, Y., Yu, Z., Li, S., Feng, S., Cheng, Y., & Agard, D. A. (2020). General and robust covalently linked graphene oxide affinity grids for high-resolution cryo-EM. *Proceedings of the National Academy of Sciences*, *117*(39), 24269–24273. https://doi.org/10.1073/pnas.2009707117

155. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A. J., & Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature Methods*, *16*(11), 1153–1160. https://doi.org/10.1038/s41592-019-0575-8

156. Kühlbrandt, W. (2014). The Resolution Revolution. *Science*, *343*(6178), 1443–1444. https://doi.org/10.1126/science.1251652

157. Al-Amoudi, A., Chang, J.-J., Leforestier, A., McDowall, A., Salamin, L. M., Norlén, L. P. O., Richter, K., Blanc, N. S., Studer, D., & Dubochet, J. (2004). Cryo-electron microscopy of vitreous sections. *The EMBO Journal*, *23*(18), 3583–3588. https://doi.org/10.1038/sj.emboj.7600366

158. Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowall, A. W., & Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics*, *21*(2), 129–228. https://doi.org/10.1017/s0033583500004297

159. Lyumkis, D. (2019). Challenges and Opportunities in Cryo-EM Single-Particle Analysis. *Journal of Biological Chemistry*, *294*, jbc.REV118.005602. https://doi.org/10.1074/jbc.REV118.005602

160. D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., & Kühlbrandt, W. (2019). Protein denaturation at the air-water interface and how to prevent it. *ELife*, *8*. https://doi.org/10.7554/elife.42747

161. Glaeser, R. M., & Han, B.-G. (2016). Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophysics Reports*, *3*(1-3), 1–7. https://doi.org/10.1007/s41048-016-0026-3

162. Glaeser, R. M. (1985). Electron Crystallography of Biological Macromolecules. *Annual Review of Physical Chemistry*, *36*(1), 243–275. https://doi.org/10.1146/annurev.pc.36.100185.001331

163. Chen, J., Noble, A. J., Kang, J. Y., & Darst, S. A. (2019). Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *Journal of Structural Biology: X*, *1*, 100005. https://doi.org/10.1016/j.yjsbx.2019.100005

164. Otzen, D. E. (2002). Protein Unfolding in Detergents: Effect of Micelle Structure, Ionic Strength, pH, and Temperature. *Biophysical Journal*, *83*(4), 2219–2230. https://doi.org/10.1016/s0006-3495(02)73982-9

165. Bernal, R. A., & Stock, D. (2004). Three-Dimensional Structure of the Intact Thermus thermophilus H+-ATPase/Synthase by Electron Microscopy. *Structure*, *12*(10), 1789–1798. https://doi.org/10.1016/j.str.2004.07.017

166. Passmore, L. A., & Russo, C. J. (2016). Specimen Preparation for High-Resolution Cryo-EM. *Methods in Enzymology*, *579*, 51–86. https://doi.org/10.1016/bs.mie.2016.04.011

167. Brink, J., Sherman, M. B., Berriman, J., & Chiu, W. (1998). Evaluation of charging on macromolecules in electron cryomicroscopy. *Ultramicroscopy*, *72*(1-2), 41–52. https://doi.org/10.1016/s0304-3991(97)00126-5

168. Wu, K., Wu, D., Zhu, L., & Wu, Y. (2021). Application of Monolayer Graphene and Its Derivative in Cryo-EM Sample Preparation. *International Journal of Molecular Sciences*, *22*(16), 8940. https://doi.org/10.3390/ijms22168940

169. Noble, A. J., Wei, H., Dandey, V. P., Zhang, Z., Tan, Y. Z., Potter, C. S., & Carragher, B. (2018). Reducing effects of particle adsorption to the air–water interface in cryo-EM. *Nature Methods*, *15*(10), 793–795. https://doi.org/10.1038/s41592-018-0139-3

170. Yeager, M., Dryden, K. A., & Ganser-Pornillos, B. K. (2012). Lipid Monolayer and Sparse Matrix Screening for Growing Two-Dimensional Crystals for Electron Crystallography: Methods and Examples. *Methods in Molecular Biology*, 527–537. https://doi.org/10.1007/978-1-62703-176-9_28

171. Kurniawan, J., Ventrici de Souza, J. F., Dang, A. T., Liu, G., & Kuhl, T. L. (2018). Preparation and Characterization of Solid-Supported Lipid Bilayers Formed by Langmuir–Blodgett Deposition: A Tutorial. *Langmuir*, *34*(51), 15622–15639. https://doi.org/10.1021/acs.langmuir.8b03504

172. Andersson, J., & Köper, I. (2019). Biomimetic Membranes. *Comprehensive Nanoscience and Nanotechnology*, 49–64. https://doi.org/10.1016/b978-0-12-803581-8.10447-3

173. *26.9: Phospholipids*. (2016, June 27). Chemistry LibreTexts. https://chem.libretexts.org/Bookshelves/Introductory_Chemistry/Book%3A_Introductory_Chemistry_(CK-12)/26%3A_Biochemistry/26.09%3A_Phospholipids

174. Dietrich, J., & Vénien-Bryan, C. (2005). *Strategies for Two-Dimensional Crystallization of Proteins Using Lipid Monolayers*. https://doi.org/10.1142/p313

175. Rouiller, I., Butel, V. M., Latterich, M., Milligan, R. A., & Wilson-Kubalek, E. M. (2000). A Major Conformational Change in p97 AAA ATPase upon ATP Binding. *Molecular Cell*, *6*(6), 1485–1490. https://doi.org/10.1016/s1097-2765(00)00144-1

176. Peter, B., & Higgins, M. (n.d.). *Monolayer Assay Protocol*. Www2.Mrc-Lmb.cam.ac.uk. Retrieved September 21, 2021, from https://www2.mrc-lmb.cam.ac.uk/groups/hmm/techniqs/mono.html

177. Kremer, J. R., Mastronarde, D. N., & McIntosh, J. Richard. (1996). Computer Visualization of Three-Dimensional Image Data Using IMOD. *Journal of Structural Biology*, *116*(1), 71–76. https://doi.org/10.1006/jsbi.1996.0013

178. Frangakis, A. S., & Hegerl, R. (2001). Noise Reduction in Electron Tomographic Reconstructions Using Nonlinear Anisotropic Diffusion. *Journal of Structural Biology*, *135*(3), 239–250. https://doi.org/10.1006/jsbi.2001.4406

179. Cheng, A., Eng, E. T., Alink, L., Rice, W. J., Jordan, K. D., Kim, L. Y., Potter, C. S., and Carragher, B. (2018) High resolution single particle cryo-electron microscopy using beam-image shift. J. Struct. Biol. 204, 270-275

APPENDIX A

LIST OF PUBLICATIONS

LIST OF PUBLICATIONS DURING GRADUATE STUDY

1. Chan, K.-Y., Du Truong, C., Poh, Y.-P., & Chiu, P.-L. (2020). Sample Preparation and Data Collection for Electron Crystallographic Studies on Membrane Structures and Lipid–Protein Interaction. *CryoEM*, 247–265. https://doi.org/10.1007/978-1-0716-0966-8_11

2. Du Truong, C., Craig, T. A., Cui, G., Botuyan, M. V., Serkasevich, R. A., Chan, K.-Y., Mer, G., Chiu, P.-L., & Kumar, R. (2021). Cryo-EM reveals conformational flexibility in apo DNA polymerase ζ. *Journal of Biological Chemistry*, *297*(2), 100912. https://doi.org/10.1016/j.jbc.2021.100912

3. Maqsood, S., Kilpatrick, S. M., Truong, C. D., & Lefler, S. R. (2021). Analysis of Amylase in the Kitchen: An At-Home Biochemistry Experiment for the COVID-19 Pandemic. *Journal of Chemical Education*, *98*(3), 858–865. https://doi.org/10.1021/acs.jchemed.0c01236

4. March, K., Venkatraman, K., Truong, C. D., Williams, D., Chiu, P., & Rez, P. (2021). Protein secondary structure signatures from energy loss spectra recorded in the electron microscope. *Journal of Microscopy*, *282*(3), 215–223. https://doi.org/10.1111/jmi.12995

5. Robertson, K. E., Truong, C. D., Craciunescu, F. M., Yang, J.-H., Chiu, P.-L., Fromme, P., & Hansen, D. T. (2019). Membrane directed expression in Escherichia coli of BBA57 and other virulence factors from the Lyme disease agent Borrelia burgdorferi. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-53830-x

6. Du Truong, C., Williams, D.R., Zhu, M., & Chiu, P.-L. (2021). Sample preparation using a lipid monolayer method for electron crystallographic studies. *Journal of Visualize Experiments*. (revision)

7. Baker, A.T, Boyd, R.J., Sarkar, D., Crespo, A.T., Chan, C.K., Bates, E., Waraich, K., Vant, J., Wilson, E., Du Truong, C., Lipka-Lloyd, M., Fromme, P., Vermaas, J., Williams, D.R., Machiesky, L., Heurich1, M., Nagalo, B.M., Coughlan, L., Umlauf, L., Chiu, P.-L., Rizkallah, P.J., Cohen, T.S., Parker, A.L, Singharoy, A., & Borad, M.J. (2021). ChAdOx1 interacts with CAR and PF4 with implications for thrombosis with thrombocytopenia syndrome. *Science Advances*. (revision)

APPENDIX B

CRYO-EM DATA STATISTIC OF POLYMERASE ζ

Table 1. Statistics of the single-particle cryo-EM structure determination of the apo DNA polymerase ζ complex of *Saccharomyces cerevisiae.*

Protein	Apo DNA polymerase ζ (Polζ) (PDB: 7LXD; EMDB: EMD-23570)
Data collection	
Electron microscope	Thermo Fisher/FEI Titan Krios TEM
Accelerating voltage (kV)	300
Spherical aberration constant (mm)	2.7
Detector camera	Gatan K2 Summit DDD camera
Defocus (μm)	-0.6 - -3.0
Nominal magnification	48,077×
Physical pixel size (Å/pixel)	1.025
Image dose ($e^-/Å^2$)	44.3
Image processing	
Number of movies	11,698
Number of particles selected (initial)	1,658,585
Number of particles used for final 3D density (final)	213,120
Spatial frequency at FSC of 0.143 ($Å^{-1}$)	4.11
Imposed symmetry	C1
Sharpening *b*-factor ($Å^2$)	-162.8
Modeling	
Initial model used (PDB code)	6V8P
Model composition	
Non-hydrogen atoms	17,456
Protein residues	2,093
Ligands	8
B factors ($Å^2$)	
Protein	120.3

Ligands	94.0
RMS deviations	
Bond length (Å)	0.006
Bond angle (°)	1.299
Clash score	9.33
MolProbity score	1.62
Rotamer outlier (%)	0.00
Ramachandran plot (%)	
Disallowed	0.00
Allowed	2.68
Favored	97.32

Table 2. Data collection and refinement statistics for human Rev7(R124A)-SHLD3(41-74aa)

Protein	human Rev7(R124A)-SHLD3 (41-74 aa)
Data collection	
Space group	P32 2 1
Cell dimensions	
a, *b*, *c* (Å)	59.99, 59.99, 132.47
□□□□□□□ (°)	90, 90, 120
Resolution (Å)	29.25-2.00 (2.05-2.00) *
R_{sym} or R_{merge}	0.083 (1.193)
I / □*I*	40.89 (4.11)
Completeness (%)	98.40 (98.89)
Redundancy	22.6 (20.9)
Refinement	
Resolution (Å)	29.25-2.00 (2.05-2.00)
No. reflections	19,167 (1339)
R_{work} / R_{free}	0.22 (0.30)/0.24 (0.32)
No. atoms	2,016
Protein	1,817
Ligand/ion	15
Water	184
B-factors	
Protein	36.17
Ligand/ion	114.28
Water	45.14
R.m.s. deviations	

Bond lengths (Å)	0.004
Bond angles (°)	0.65
Ramachandran plot (%)	
Disallowed	0.00
Allowed	0.44
Favored	99.56

*Values in parentheses are for highest-resolution shell.