Title
Electron microscopy methods to overcome the challenges of structural heterogeneity and preferred orientations in small (sub-500 kDa) macromolecular complexes

Author
Lipscomb, Dawn Michelle

Publication Date
2017

Peer reviewed|Thesis/dissertation

Electron microscopy methods to overcome the challenges of structural heterogeneity and preferred orientations in small (sub-500 kDa) macromolecular complexes

By

Dawn Michelle Lipscomb

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

In

Biophysics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eva Nogales, Chair
Professor Gloria Brar
Professor James Hurley
Professor David Savage

Summer 2017

Abstract

Electron microscopy methods to overcome the challenges of structural heterogeneity and preferred orientations in small (sub-500 kDa) macromolecular complexes

By

Dawn Michelle Lipscomb

Doctor of Philosophy in Biophysics

University of California, Berkeley

Professor Eva Nogales, Chair

While cryo-EM imaging technology and software suites have led to a "resolution revolution", the poor reproducibility and inefficiency offered by current specimen preparation techniques remain challenges largely unexamined and under-prized. Regardless of camera or software proficiency, a poor-quality specimen will always produce an equivalently substandard 3D reconstruction. To further advance the high-resolution EM pipeline and establish quality control for grid specimens, we tested three techniques to assess optimum conditions for vitrification of small (<500kDa), multi-subunit biological macromolecular complexes.

We performed specimen optimization for a 300kDa, five-component, macromolecular complex, PRC2. PRC2 is a key regulator of gene silencing in eukaryotes and mutations in its catalytic subunit, Ezh2, are linked to a number of human cancers and degenerative diseases. To date, no high-resolution structure of the complete complex has been solved. PRC2 presents a number of unique challenges for the structural biologist, including its small size, highly flexible regions, and conformation heterogeneity. A high-resolution structure of PRC2 would provide unparalleled insight into the biochemical mechanisms that mediate gene silencing by methylation of histone tails. For these reasons, PRC2 is an ideal candidate for optimizing cryo-EM grid preparation techniques translatable to similar complexes and for future research.

Previous EM studies of PRC2 revealed a distinct preferred orientation when bound to carbon substrates, necessitating the collection of tens of thousands of images to generate a complete structure containing high-resolution features. To overcome this impediment, we experimented with two other grid specimen preparation techniques. First, we vitrified solution-suspended PRC2 in open holes on carbon mesh support. Despite modifications of chemical and physical buffer and substrate parameters to stabilize the complex, denaturation upon surface binding at the air-water interface and dissociation of the complex during vitrification proved insurmountable. We next tested a method of affinity binding to a streptavidin monolayer substrate covering the holes of a carbon mesh support. We found that chemically biotinylated PRC2 was intact, monodisperse, and bound the grid in random orientations, all factors critical to high-resolution cryo-EM structure determination. Negative stain EM analysis revealed a more even distribution of views for 3D reconstruction. From these results, we conclude that streptavidin monolayer substrates are an option for overcoming the obstacle of preferred

orientations for small complexes and provide an easily reproducible protocol for grid preparation.

Table of Contents

List of Figures

Acronyms

AEBP2	Adipocyte Enhancer-Binding Protein
CCD	Charged Coupled Device
cryo-EM	Cryogenic Electron Microscopy
DDD	Direct Detection Device
DNA	Deoxyribonucleic Acid
DQE	Detective Quantum Efficiency
EED	Embryonic Ectoderm Development
EM	Electron Microscopy
Ezh2	Enhancer of Zeste 2
FFT	Fast Fourier Transform
FSC	Fourier Shell Correlation
GFP	Green Fluorescent Protein
GPU	Graphic Processing Unit
GST	Glutathione S-transferase
H3K27me3	Histone 3 Lysine 27 tri-methylation
HOTAIR	HOX Transcript Antisense RNA

Jarid2	Jumonji, AT Rich Interactive Domain 2
kDa	Kilodalton
keV	Kiloelectron Volt
ncRNA	Non-Coding RNA
PRC2	Polycomb Repressive Complex 2
RELION	Regularised Likelihood Optimisation
RNA	Ribonucleic Acid
SA	Streptavidin
SAM	S-Adenosyl methionine
SAXS	Small Angle X-ray Scattering
SET	Su(var)3-9, Enhancer of zeste and Trithorax
Suz12	Suppressor of Zeste 12
SRM	Stimulatory Recognition Motif
TEM	Transmission Electron Microscopy
VEFS	VRN2-EMF2-FIS2-Su(z)12
WD40	40 amino acid terminating in tryptophan-aspartic acid dipeptide
Xist	X-Inactivate Specific Transcript

Acknowledgements

With a novel in hand, and a dream settling into the recesses of my imagination, I traveled west. Amongst the soft, golden hills, towering turbines dotted the landscape, proudly proclaiming, "Welcome, to California." Filled with ambition and unmarred by the rigidities that chafe the enthusiasms and whimsical fancies of a fresh young mind, I arrived eager to try something novel and exciting, hungry for knowledge and unrelenting in the quest for purpose.

I always imagined life as a beautiful, wonderful, dynamic adventure. I would discover, build, grow, advance, imagine, and eventually become part of a broader, knowledgeable scientific community at one of the world's most renowned Universities. Eva Nogales provided a platform for me to expand academically, intellectually, and personally. I am eternally gratefully for the opportunities afforded to me throughout my graduate studies. I've been part of an amazing team. Simon, Vignesh, and my peers have provided endless enthusiasm and reinforcement, cultivating an environment that allowed growth and individuality as a researcher. Patricia served as a true inspiration, emboldening my curiosities and never tiring of my relentless onslaught of questions. Teresa is certainly the heart of lab, coordinating with all the arterial forces that build a captivating environment and keeping me on a path to success.

There are several people who have truly touched my life and my world would be empty without them. I'm indebted to Mica Smith, who I've known since memories were worth remembering and our tangential growth made us two of the same, strange individuals. Despite my flaws and at times questionable actions, I would not have completed my project without Nicole Haloupek to highlight my greatest attributes and edit out my failings. The endless hours of silence reinforced during long nights at the lab were dampened by the lyrical genius of 2 Chainz reminding me that yeah, "I'm different."

I have become myself; five solid years of growth, endless interpersonal interactions, finding love, conquering fear, and exploring the world. I'm so grateful to have a better half, my life partner and most trusted confidant, Dylan Rosario. Our life is only beginning.

Ultimately, I am dedicating my dissertation to my mother and father who have provided warmth, unwavering love, and unquestionable certainty. I am so proud to have parents who nurtured and encouraged me through each grueling mile, regardless of perceived obstacles, or the outlandishness of my ambitions.

Chapter 1

Introduction

1.1 Cryo-EM: A method for analyzing macro-molecular complexes

Over the past decade, the burgeoning field of cryo-EM has undergone a technological and computational revolution[1,2]. Cryo-EM allows researchers to view snapshots of macro-molecular complexes flash-frozen in buffer solutions, closely approximating biological conditions.[2] The first EM images were recorded on film, which was developed in a dark room, and finally digitized for analysis, making data collection slow and laborious.[1] Data collection soon became automated and faster after the invention of CCD cameras that relied on the conversion of electrons into photons that were then detected by the scintillator.[1] However, the "photon cloud" generated by the conversion of multiple electron events in neighboring areas would be much larger than the pixel size of the camera, introducing noise and signal attenuation at high-resolution.[1] CCD cameras also captured snapshots of particles in a single short exposure. This exposure often resulted in micrographs containing blurry images of particles due to beam induced movement.[1]

New imaging technologies, such as direct electron detectors, allow the collection of multiple image frames over long exposure times (~20s or more). Aligning these frames minimizes blurring caused by movement of the particles and surrounding cryo-environment induced by settling of the sample upon exposure to the electron beam.[1,2] Also, this novel camera technology has heightened DQE, which measures the performance of an imaging system by comparing the signal-to-noise ratio in its measurements. This allows for imaging that preserves higher order structural information. [2]

These detectors were invented in tandem with new image processing software and increasingly reliable algorithms. The combination of direct electron detectors and improved image processing software has made possible the generation of EM density maps of biological complexes at atomic or near-atomic resolution.[1]

Additionally, automated data collection and increasingly powerful microscopes, such as the Titan Krios TEM, have streamlined and popularized cryo-EM as a primary method for

tackling some of the most difficult problems in structural biology.[2] Taken together, these advances have reduced the lower mass limit for EM, allowing generation of atomic models of complexes from ~300kDa to what is theoretically predicted by physics (~150-50kDa).[2]

Despite these advances, the high-resolution EM pipeline is far from optimized [Fig 1.1]. This chapter reviews the procedures involved in high-resolution EM map generation and outlines the obstacles and opportunities for improvement along the path to atomic modeling of biological complexes.

1.2 Strategies for high-resolution structure solving in cryo-EM

1.2.1 Protein purification

All samples exhibit some degree of structural or compositional heterogeneity. Heterogeneity must be minimized as it may prevent a sufficient distribution of views for the class averaging necessary for high-resolution reconstructions.[2,3]

Before cryo-EM specimen preparation, a mostly homogeneous biological sample with minimum contamination is achieved through careful modification of biochemical parameters. Negative-stain EM analysis is used to assess and fine-tune these parameters to verify the suitability of a sample for higher resolution cryo-EM studies.[1,2,3] During this preliminary analysis, samples are preserved through embedment in heavy metal salts such as uranyl acetate and uranyl formate on continuous carbon grids.

Sample heterogeneity falls into two broad categories.[3] The first, compositional heterogeneity, refers to the presence of complexes with sub-stoichiometric components, aggregates, or contaminants. This type of heterogeneity is diminished through biochemical techniques such as affinity tag purification, improving buffer conditions, and/or chemical cross-linking during or after purification, and is assessed using negative-stain EM.[3]

The second type of heterogeneity, conformational heterogeneity, exists due to high flexibility of many biological complexes. Chemical crosslinking can be used to reduce heterogeneity caused by structural variations in macromolecules; however, caution should be exerted when interpreting density maps produced using crosslinking. Complexes may no longer exist in their native state, as crosslinking can introduce artifacts and irreversibly fix flexible regions together, calling into question the physiological relevance of a structure.[3]

1.2.2 Specimen preparation

Preparation and analysis of negatively stained samples is a generally rapid and relatively reliable technique for assessing the degree of sample heterogeneity. Due to the extent to which they scatter electrons, staining with uranium salts creates high contrast, but the grain size of the stain prevents negatively stained samples from surpassing ~15 Å resolution.[1] Further, the stain may cause artifacts by flattening or collapsing the complex being analyzed.[2] To reach higher resolutions, observe secondary structural elements, and, in the best cases, build atomic models, the sample must be vitrified rather than being stained.[3]

Specimen vitrification—that is, the preparation of samples for cryo-EM analysis—is the most reliable technique for generating high-resolution EM density maps. The vitrification

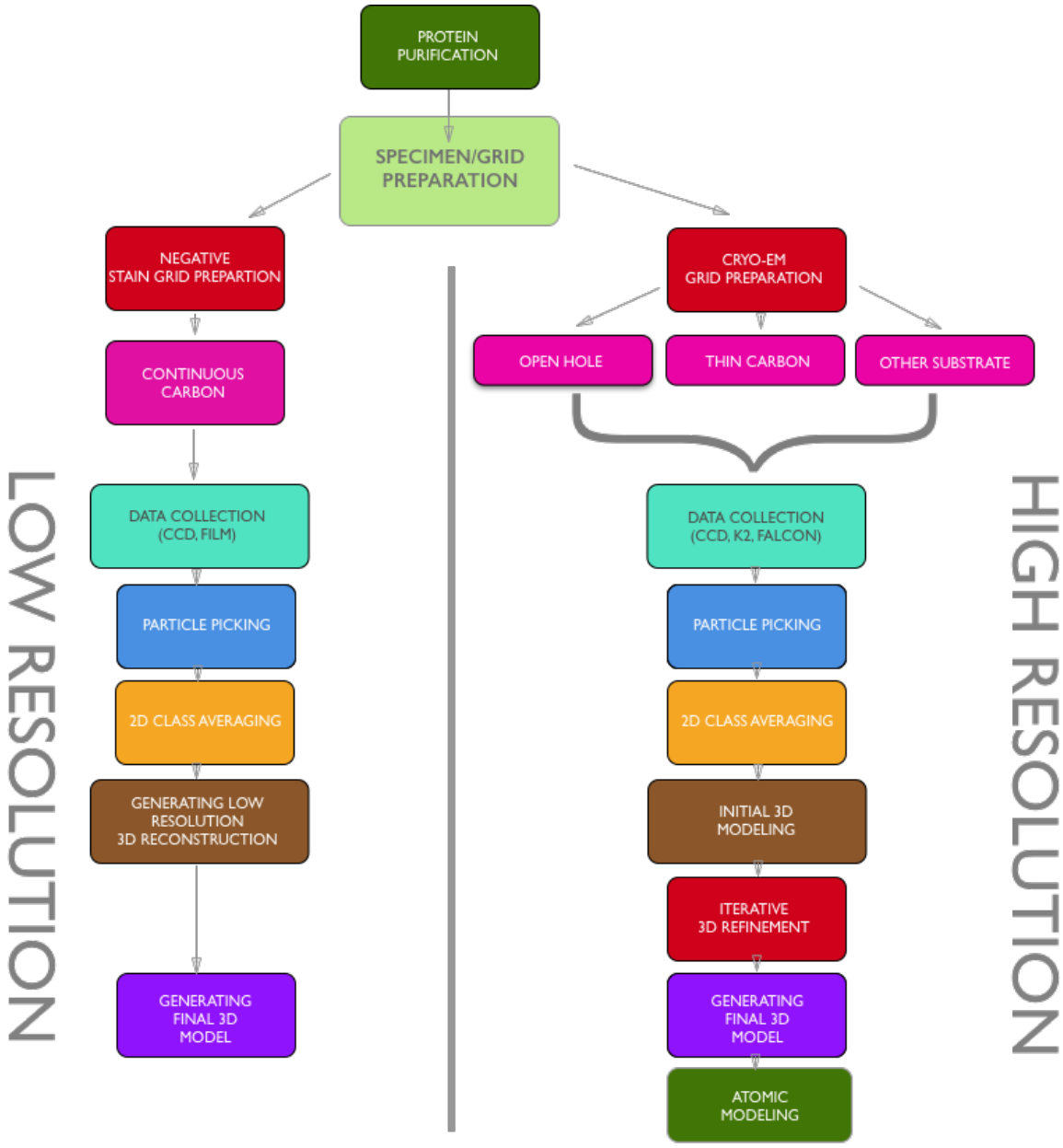


Figure 1.1 Single-particle EM workflow diagram. The left side shows the procedures involved in generating a low-resolution model and testing various sample conditions for implementation in cryo-EM. The right side maps the steps involved in obtaining a high-resolution structure. This dissertation focuses on grid preparation techniques, a critical of step to obtain a high-resolution structure.

process involves plunging a sample of protein complexes suspended in buffer into a pool of liquid ethane [Fig 1.2]. This rapid cooling process prevents the formation of opaque, crystalline ice; instead, it produces vitreous ice, which is relatively transparent to the electron beam, allowing the biological particles to be visualized. This method also preserves the native or near-native structures of protein complexes as near-physiological conditions are maintained, unlike the case in X-ray crystallography, which requires protein crystal growth through alteration of buffer conditions and may change structural conformation.[4]

Cryo-EM specimens are prepared on circular grids of one of three types: open-hole, thin continuous carbon, or affinity substrate [Fig 1.1 & 1.2]. Continuous carbon grids consist of a layer of thin, amorphous carbon that overlays the holes and provides a support film for complex adsorption. However, particles may fall in only one orientation on the grid, limiting the number of views for 3D reconstruction.[5] Ideally, researchers seek conditions that minimize background noise introduced by support films and prevent preferred orientations caused by interactions between the sample and the substrate.[6]

Open-hole grid preparation is a viable option for overcoming both since these grids lack a support film or substrate for protein adsorption, meaning the protein is suspended in a thin film of liquid in the holes.[6] However, this technique presents its own set of challenges. Many complexes are highly sensitive to the denaturing effects of the air-water interface and tend to dissociate or fall apart upon contact with the surface.[6,7] The use of a substrate grid typically serves to protect complexes from encountering one side of the air-water interface, and the adsorption process prevents further contact by providing a surface for protein binding.[6] Finally, the use of an affinity substrate, such as a SA monolayer, not only partially protects complexes from the air-water interface, but may also increase the distribution of views for particularly small, asymmetric particles, by randomly orienting them on the grid through chemical binding.[7,8] This thesis explores various strategies of open-hole and affinity substrate grid preparation to reduce denaturation, heterogeneity, and preferred orientations on continuous carbon substrates.

1.2.3 Data collection and analysis

As previously mentioned, DDD cameras are capable of recording a series of dose-fractionated frames over one long exposure that are then aligned and averaged[1,2]. Image alignment produces enhanced contrast of individual particles otherwise blurred by electron beam-induced motion[1,2]. Accurate alignment of the microscope is critical to ensure optimal usage of enhanced image capturing technologies. Some of these alignment parameters include but are not limited to selection of an appropriate objective aperture size, spot size, electron dose and defocus values, and coma-free alignments to ensure accurate beam spread and coherence.[2,3] Microscope parameters must be considered for each specimen as the size of the complex and materials used for grid preparation may affect the degree of contrast required for frame alignment during movie processing.[3] For instance, smaller complexes require higher defocus values of a few micrometers in order to achieve sufficient contrast.[3] Once aligned, the microscope is ready for semi-automated data collection using a DDD camera.

Image analysis begins by frame alignment and fastidious selection of micrographs with concentric Thon rings that extend to high resolution in all directions in Fourier space.[3,4] Then, individual particles are selected using algorithms or by a user, or some combination of both. This step is crucial to ensure that noise, artifacts, aggregates, and partially dissociated complexes, are not introduced into the final reconstruction. The particles are subsequently boxed out of

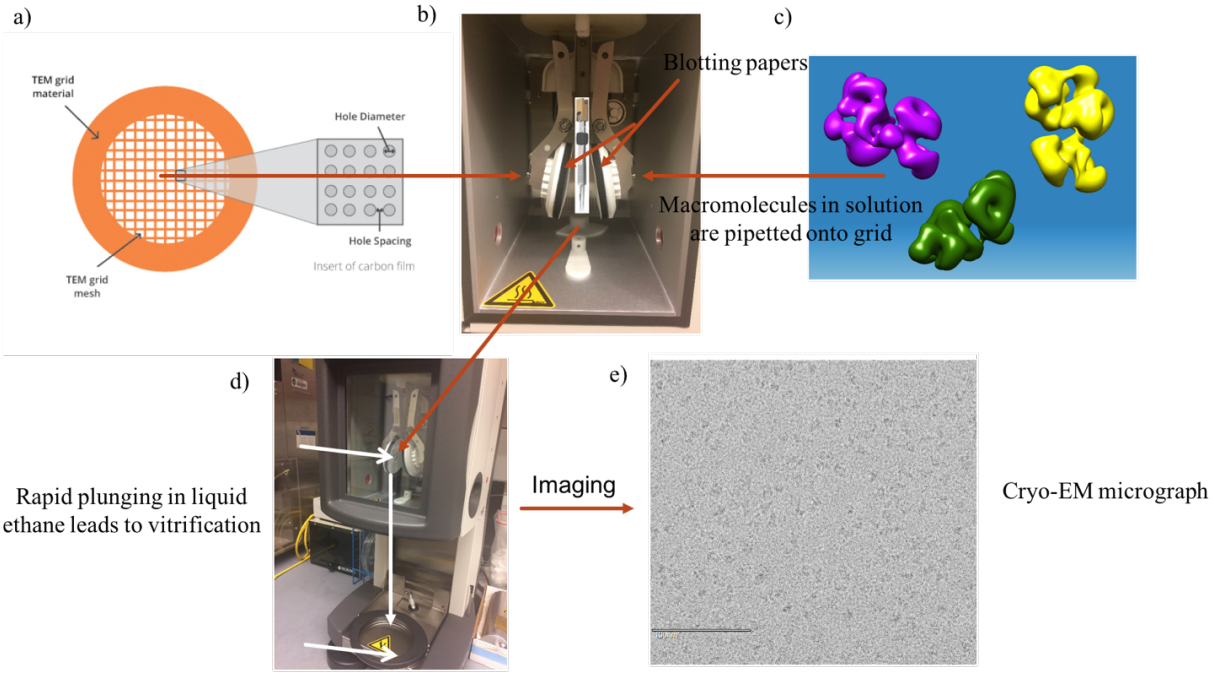


Figure 1.2 Overview of cryo-EM grid preparation. (a) Typical layout of an EM grid. The grids normally consist of a square copper mesh support structure. Each individual square has a series of small open holes of carbon or continuous carbon that are filled with sample solution as shown in (c). These grids are blotted (b) and then plunge-frozen in liquid ethane (d) and imaged using an electron microscope (e).

micrographs and CTF corrected to account for the effect of the imaging system. However, no particle picking method is without flaws. 2D analysis in programs such as EMAN2 or RELION groups and aligns particles into similar projection stacks and weeds out most "junk" particles.[9] These stacks simultaneously indicate whether particles adopt preferred orientations and whether the distribution of views is sufficient for 3D reconstructions.

Another approach to culling junk from datasets is implementing an unsupervised 3D classification software, such as the one used in RELION, designed to sort particles through multiple rounds of 3D reconstruction cycles.[9] Traditional means of data interpretation relied on a series of refinement procedures that required fine-tuning of ad hoc parameters through user input.[9] Conversely, RELION applies a maximum a posteriori (MAP) estimate, or empirical Bayesian approach, to generate a statistical model from the input data with little tuning of parameters by the user.[9] This approach calculates model correctness by probability based on data observed and the assumption that reconstructions of macromolecules are smooth.[9] After completing 3D classification and discarding undesired particles at each step, the remaining data is subjected to iterative 3D refinement, and in some cases, flexible regions of complexes are isolated and refinement is focused on specific regions of interest. Finally, FSC calculations serve as a "gold-standard" for estimating overall and local resolution for data interpretation.[9]

A couple years ago, programs like RELION were simply too computationally expensive for widespread implementation.[1] Today, more affordable supercomputing clusters and GPUs allow researchers to generate reconstructions much more quickly, sometimes within days.[1] While computing challenges no longer pose a barrier, the Bayesian approach still necessitates mostly homogenous, high-contrast images to generate sufficient information for eventual atomic modeling of macromolecules.[9] Given that RELION exclusively uses the data observed and a heavily low-pass filtered starting model (~60-100 Å) as the only input parameters for generating a statistical model, acquiring high-quality images with particles randomly oriented across the field of view is critical for the success of the Bayesian approach.[9] Therefore, optimizing specimen preparation is critical, yet it remains a challenge in cryo-EM studies.[9,10]

1.3 Research rationale and significance

Section 1.2 describes the recent improvements in technological and computational approaches that facilitated the mainstream adoption of cryo-EM for structural biology studies. The previous macromolecular size limitations (~300 Å) have been broken, and EM map depositions with resolutions better than 4 Å have increased exponentially within the last two years.[1,2] However, at each major junction in the EM pipeline [Fig. 1.1], hurdles in specimen preparation continue to hinder researchers from pushing the resolution of particularly difficult specimens.[10] A translatable, reliable grid preparation technique that preserves the native structure of complexes, while simultaneously protecting macromolecules from the denaturing effects of the air-water interface and reducing preferred orientations, would aid in completing the workflow for many asymmetric, small, and conformationally diverse macromolecules.

My research explores the use of continuous carbon, open-hole, and affinity substrate grid preparation techniques. In this thesis, we demonstrate that while substrate-free, open-hole conditions are generally preferred due to higher contrast and random orientation in the thin film of vitrified solution, the data indicates that denaturation at the air-water interface may present too formidable a challenge, necessitating the use of a substrate to reduce protein contact with the liquid surface during vitrification. The use of continuous carbon substrates is simple,

economical, and popular, but may produce preferred orientations.[10] Here, we show SA monolayer affinity substrate grid specimens may overcome these challenges for small (<500kDa), asymmetric, multi-subunit complexes. The experiments outlined in the proceeding chapters specifically pertain to five-component PRC2-AEBP2, a 300kDa methyltransferase chromatin remodeling complex that silences transcription and is essential for eukaryotic life.

Chapter 2

PRC2 as a model complex for EM grid specimen optimization

2.1 PRC2's essentiality for eukaryotic life

Polycomb group (PcG) protein complexes, which define patterns of gene expression by directly interacting with chromatin structure,[11] are essential to transcriptional memory over multiple lineages of cell division. Polycomb repressive complex 2 (PRC2) is evolutionarily conserved across eukaryotes and plays a direct role in gene silencing through post-translational modification of histones.[11,12] From Drosophila to mammals, PRC2 has been shown to silence genes by di and tri-methylating H3K27 (histone 3, lysine 27) of nucleosomes.[12] High levels of methylated H3K27 are associated with chromatin compaction and gene repression, and partial knockdown or full deletion of PRC2 in mice leads to developmental defects or lethality.[13] Despite the importance of PcG proteins, little is known about their structures, the ways they are recruited to binding sites, and the mechanisms by which the epigenetic marks they generate are propagated.

Histone post-translational modifications are essential for transcriptional regulation in eukaryotes, and abnormal patterns of posttranslational modifications are associated with a number of human cancers. Therefore, understanding genome maintenance, including how PcG proteins contribute to it, is indispensable for the development of effective drugs and therapies to combat disease. PRC2 enzymatic activity resides in a SET domain that is functional in the context of a core of four subunits, but the mechanisms through which methyltransferase activity is stimulated or repressed within PRC2 remain largely unknown.

he PRC2 core is composed of four major components: EZH1/2, EED, Suz12, and RBAp46/48.[12] Both Ezh1 and Ezh2 contain the SET catalytic domain responsible for histone lysine methyltransferase (HKMT) activity [Fig. 2.1].[14,15] Ezh2 is generally expressed in proliferating tissues and exhibits much higher levels of HKMT activity than does Ezh1, which maintains chromatin compaction in adult tissues and is less enzymatically active.[14,15] EED is a WD40 repeat protein that binds to histone tails with H3K27me3 marks and allosterically facilitates initiation of methyltransferase activity through direct interaction with Ezh2.[15-20]

Suz12 forms a structural core with Ezh2 and EED and regulates methyltransferase activity through its VEFS domain.[21] RbAp48, another WD40 repeat protein, forms a stable

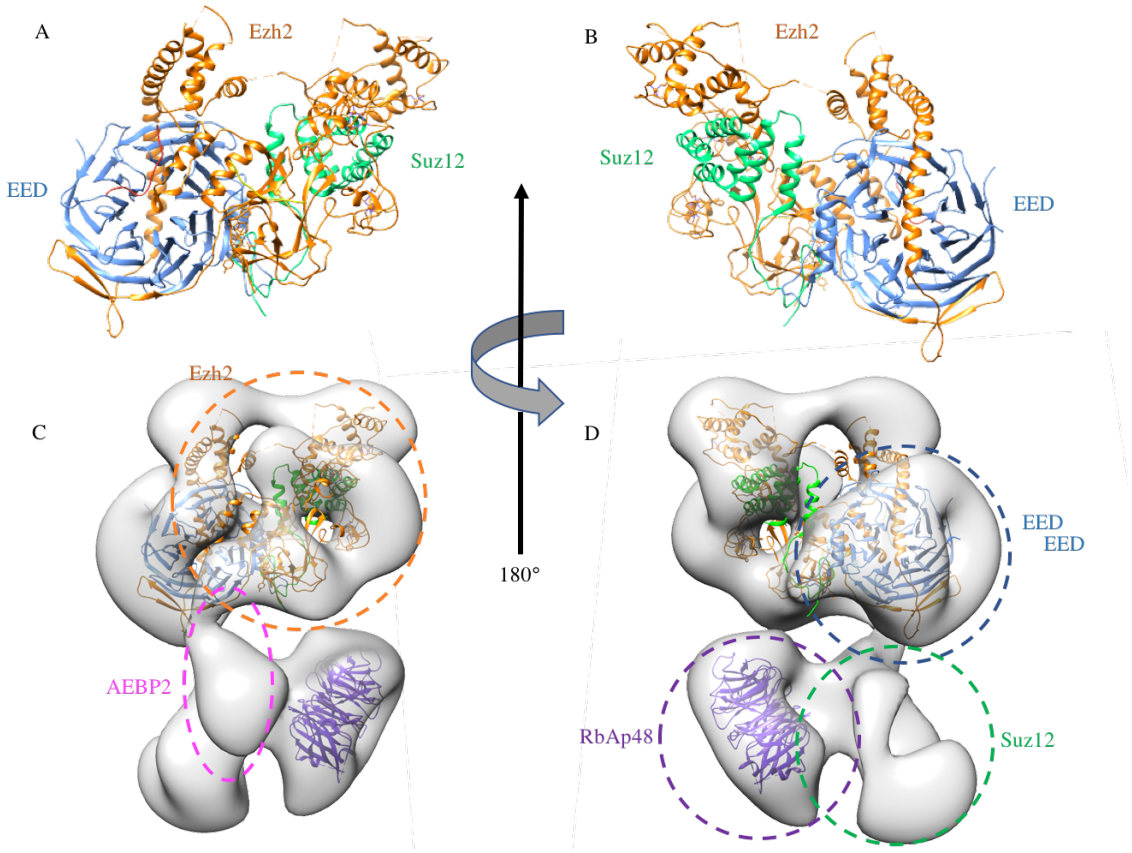


Figure 2.1 EM reconstruction of PRC2 with docked crystal structures. (A) Crystal structure of a human PRC2 core consisting of Ezh2, EED, and the VEFS domain of Suz12. The complex was activated by binding of a H3K27me3-containing peptide to EED and included also a peptide containing the H3K27M mutant bound to the catalytic SET domain of Ezh2. (B) Back view of PRC2. (C) Crystal structure shown in (A) docked into the 21A resolution EM density map of the full human PRC2 (containing Ezh2-EED-Suz12-RbAp48) bound to the cofactor AEBP. The EM structure consists of 2 major lobes stapled by AEBP2. The density regions for Ezh2 and AEBP2 are highlighted in the map. (D) Back view of (C) showing the locations of EED, Suz12, and RbAp48, along with the RbAp48 docked crystal structure. As illustrated by the last two panels, high-resolution information regarding Suz12 and AEBP2 and their interaction with the catalytic core remain to be elucidated. Crystal structures and EM maps were modeled in Chimera (EMDB ID: 3336; PDB: 5HYN; PDB: 3GFC).

complex with Suz12.[23] Finally, other polypeptides and ncRNAs, such as HOTAIR and Xist, have been found to stabilize PRC2 and facilitate its recruitment to nucleosomes.[12,24] A low-resolution (21 Å) EM map of five component PRC2, together with subunit tagging, chemical crosslinking followed by mass spectrometry, and crystal structure docking of the WD40 repeat domains yielded a model of the PRC2-AEBP2 complex that revealed the relative position of the core subunits and the cofactor AEBP2 [Fig. 2.1].[11] Without complete crystal structures or a high-resolution map, the structure and function of individual components of PRC2 in the context of the full complex and their interactions with potential chromatin recruiters leaves many questions regarding the mechanisms of PRC2-mediated gene silencing unanswered.

2.2 Biochemical interactions of PRC2 with chromatin

2.2.1 PRC2 activation via EED

Biochemical analysis showed that an H3K27me3 (amino acids 21-32) peptide binds to EED with nanomolar affinity and increases the enzymatic activity of Ezh2's catalytic SET domain, both *in vitro* and *in vivo*.[19,25] A fivefold increase in the enzymatic activity of the SET domain was detected upon peptide binding to the EED subunit, and mutagenesis of the peptide-EED interface decreased this activity back to baseline.[19,24,25] Increased catalytic turnover in the presence of the H3K27me3 peptide suggests SET domain regulation through an allosteric mechanism.[25] The stimulation of PRC2's methyltransferase activity caused by recognition of existing H3K27me3 marks through binding via the EED subunit in trans suggests a mechanism for methyltransferase propagation to nearby histone tails to disseminate the H3K27me3 epigenetic marker, resulting in gene silencing.[19,25]

2.2.2 PRC2 activation via Suz12

It is well-documented that PRC2 HKMT activity is higher in the presence of dense oligonucleosomes compared to mononucleosomes or free histones; therefore, modifications to neighboring histone tails may modulate gene silencing by PRC2.[22,26] A peptide fragment of unmodified H3 (amino acids 31-42) was shown to robustly stimulate PRC2's catalytic activity through allosteric recognition via the highly conserved VEFS domain of Suz12.[23] This stimulatory response provides a means for PRC2 to methylate histone tails independent of EED subunit recognition of preexisting, nearby H3K27me3 marks, thereby allowing genes to be switched from on to off.[23]

2.2.3 PRC2 inhibition

Several histone modifications have been shown to inhibit PRC2 HKMT activity, including H3K4me3 and H3K36me3, which are typically found near genes that are actively transcribed.[23,26-28] Two major studies investigated the mechanisms behind this inhibitory response.[23,26] HKMT and EMSA assays established that these histone tail modifications did not preclude PRC2's binding to nucleosomes, but PRC2's catalytic activity was suppressed, suggesting allosteric regulation of the Ezh2 SET domain through detection of these inhibitory markers.[23,26] Mass spectrometry and mutagenesis studies showed the binding sites for these

markers likely reside on the VEFS domain of Suz12, but the exact location and mechanism by which they regulate PRC2 function is yet to be elucidated.[23]

2.3 Structural studies of PRC2

2.3.1 Known interactions of Ezh2 with EED, SUZ12, RBAP48, and stabilizing co-factor AEPB2

PRC2 requires the cooperative interaction of five components for heightened catalytic activity.[11,12] In addition to the C-terminal SET domain that catalyzes three sequential methyl-transferase reaction, Ezh2 contains two SANT domains thought to participate in histone binding to mediate catalysis.[29,30,31] In direct contact with Ezh2 is the EED subunit, composed of a WD-40 domain and an unstructured 80-residue N-terminal domain.[17,32] EED recognizes tri-methylated lysine residues on histone H3 in trans, a step essential for HKMT activity.[18,19] Crystallographic studies and GST pull-down assays showed that an Ezh2 N-terminal peptide binds across EED's beta propeller through electrostatic interactions and a H3K27me3 peptide binds an aromatic pocket in the propeller.[17-19] The unstructured EED N-terminus was found to bind within the histone fold domain of H3, indicating it may serve to position the PRC2 complex on the nucleosome.[32]

Suz12 is required for PRC2 assembly *in vitro* and *in vivo*, but its specific role remains elusive.[21,29] Through protein labeling in EM studies as well as crosslinking and mass spectrometry, the highly conserved C-terminal VEFS domain was shown to directly interact with the Ezh2 N-terminal SANT domain, and deletion of this region inhibits PRC2 assembly.[11,12,21,29,33] Biochemical studies have demonstrated Suz12's ability to allosterically inhibit or stimulate the catalytic SET domain through recognition of different histone tail modifications, though the exact mechanism by which this occurs remains unknown.[21,29]

RbAp48, a nonessential component for methyltransferase activity, is a WD40 repeat protein shown to bind the N-terminal potion of Suz12 within a hydrophobic pocket of the beta propeller.[23] GFP and MBP tags were used to determine the location of the subunit in an EM map at 21 Å resolution, producing results in agreement with biochemical studies.[11] Crystallography, competitive binding assays, and fluorescence polarization experiments demonstrated the ability of the RbAp48-Suz12 complex to bind the N-terminus of H3 in the ligand binding site of the WD40 propeller; however, disruption of the RbAp48 component bound to the H3 tail has little to no effect on the catalytic activity of Ezh2.[23] Biochemical evidence and the low-resolution EM map of PRC2 suggest that EED and Suz12-RbAp48 work in a concerted manner to allosterically relay information to the catalytic SET domain, driving activation or repression.[11,23,24]

Additionally, heightened enzymatic function requires AEBP2, a three-zinc finger protein and one of several known cofactors of PRC2.[11,34] Without AEBP2, the complex is highly flexible and tends to fall apart on EM grids. AEBP2 spans PRC2's narrowest region, forming a critical link that stabilizes the Suz12 subunit's contact with Ezh2.[11] Mass spectrometry also identified 60 intra-and intermolecular sites of interaction within the PRC2-AEBP2 complex, including interactions between the catalytic SET domain of Ezh2 and AEBP2, and several strong links between the Suz12 domain and the cofactor.[11] This suggests that AEBP2 plays an important role in relaying information and maintaining the structure between the two domains, though the mechanisms are yet to be determined.

Recently, crystal structures of the three core components that enable and regulate trimethylation, Ezh2, EED, and the VEFS domain of Suz12, along with bound histone peptides and other constituents, were solved.[26,38] The first of these structures was from the included trimeric PRC2 (Ezh2, EED, Suz12) of thermophilic fungus *Chaetomium thermophilum*.[26] This structure contained stimulating peptide H3K27me3 bound to the EED domain and an oncogenic histone variant that precludes methylation of H3 by PRC2, H3K27M, bound to the active site of Ezh2.[26] The structure revealed that PRC2 undergoes significant conformational changes upon activation.[26] The substrate binding induces rotational motion of an autoinhibitory domain, shifting it away from the active site and completing the binding site for SAM, a cofactor that mediates methylation.[26] Furthermore, the H3K27me3 peptide bound to EED stabilizes the SRM site of Ezh2, a helical segment that extends across EED and makes extensive contacts between the peptide and SET domain of Ezh2.[26] This stabilization is thought to activate the catalytic site of PRC2, allowing for trimethylation of H3.[26] The second is a crystal structure of human PRC2, again containing Ezh2, EED, and the VEFS domain of Suz12, in this case in complex with a Jarid2 peptide, a regulatory component whose trimethylation leads to binding of EED and stimulation of PRC2 analogous to allosteric activation via trimethylated H3.[38]

The human trimeric crystal structure can be readily docked into the low-resolution EM map to place it in the context of the full PRC2-AEBP2 [Fig. 2.1]. However, the Suz12 subunit and AEBP2 cofactor of the complex remains largely unsolved. Recently, cryo-EM has become an indispensable tool for studying macromolecular complexes like PRC2, with cameras, instruments, and algorithms approaching what is theoretically predicted by physics, and leading to EM maps reaching resolutions that allowed atomic modeling (<4 Å). This work, exploring new grid preparation techniques, aids in completing the high-resolution EM pipeline, allowing researchers to push the bounds of resolution for PRC2 and other complexes alike.

Chapter 3

Comparison of grid preparation techniques

To survive the vacuum of an electron microscope and bombardment with high-energy electrons, samples for EM analysis must be preserved through embedment in heavy metal salts or by vitrification, as discussed in Chapter 1.[1,2] Embedding biological complexes in salts may distort protein structures, and the grain of the stain prevents high resolution maps from being obtained[1]. This Chapter explores the various methods of cryogenic sample preparation and the challenges faced during the vitrification process. A comparison of continuous carbon substrates and open-hole techniques reveals that although thin carbon may introduce additional noise and preferred orientations, open-hole grid preparation is not a viable option for many samples, so other methods of specimen preparation must be considered.

3.1 Sample instability during cryo-EM preparation

All high-resolution EM density maps are derived from vitrified samples. Biological complexes suspended in buffer that approximates physiological conditions are plunge-frozen in liquid ethane, vitrifying the liquid and capturing snapshots of the molecules in native or near-native conformations. Sample preparation parameters must be determined for each complex of interest.

On a macroscopic scale, deformations and debris in vitrified ice are the first challenges to overcome.[3] Upon cryo-plunging, the liquid contents are flash frozen, producing vitreous ice, which is amorphous and semi-transparent to the electron beam. Warming of the sample causes water molecules to arrange in a geometrically repeating pattern known as crystalline ice that is opaque and precludes imaging.[3] Samples may also be contaminated with ethane or ice contamination from water trapped in the microscope column, undermining image quality.[3] Therefore, at all phases of the cryo-freezing process, care must be exerted.

Using optimal grid materials and buffer conditions is essential for obtaining high-resolution reconstructions.[10] This thesis explores three major types of grid preparation (continuous carbon, open-hole, and affinity substrate) along with investigating how the addition of detergents and chemical modifications all uniquely contribute to specimen quality.

Importantly, physical forces, chemical properties of solutions, and interactions of biological macro-molecules with the air water interface all contribute to the stability of molecules examined using EM.[6,7,10] Although difficult to quantify, Brownian motion and chemical properties such as hydrophobicity cause complexes to bind or interact with the air-water interface.[6, 7] Upon contact with this interface, many proteins irreversibly undergo partial or complete denaturation. [34,35,6,2] Early experiments with ferritin revealed this denaturation can occur spontaneously and result in a thin film or "cover-slip" that floats on the solution's surface.[6] Further experimentation showed intact ferritin readily adsorbed to the denatured protein and could be extracted by lifting the film using filter paper.[6]

Following ferritin experiments, surfactants gained widespread popularity as barriers to the denaturing effects of the air-water interface, particularly for open-hole specimen preparation.[6] The use of detergents such as NP-40, when added to protein solutions at levels lower than the threshold for micelle formation, have been used to protect many biological samples from direct contact with the surface. [34,35,6,2]

However, for some samples, surfactants may do more harm than good. After incubating a protein solution on the grid's surface, excess liquid is removed by pressing filter paper to one or both sides of the grid. The shear forces produced may wick away the surfactant or denatured protein "cover-slip" along with any bound protein.[6,7]. This action may greatly reduce the number of particles present in the holes or remove them altogether. Surfactants also change the physical properties of the solution, such as viscosity, resulting in thicker, less transparent ice.[6]

Rather than blotting away excess liquid following incubation, some protocols call for a brief evaporation period to remove the remaining solution.[6] This recommendation is impractical for the many samples that are particularly sensitive to changes in buffer conditions. During evaporation, chemical properties like salt concentrations and pH are significantly altered, which may compromise the structures of many proteins.[6,2] Modern EM tools, such as Vitrobots, incubate grids in a humidity chamber to prevent evaporation [Fig. 1.2].

Aggregation and dissociation of macro-molecular complexes pose another concern during open-hole specimen preparation. Without the buttressing support layer, the vitrification process may introduce stress in already fragile interactions, causing protein components to disconnect, greatly increasing the degree of heterogeneity. Chemical crosslinking with glutaraldehyde or BS3 may reinforce such interactions, stabilizing complexes for vitrification and analysis [Fig. 3.1(a)].[2] However, this chemical fixation can introduce intermolecular aggregation and crosslinking conditions (crosslinker concentration, crosslinking time, and quenching method) should be optimized for every sample.[2]

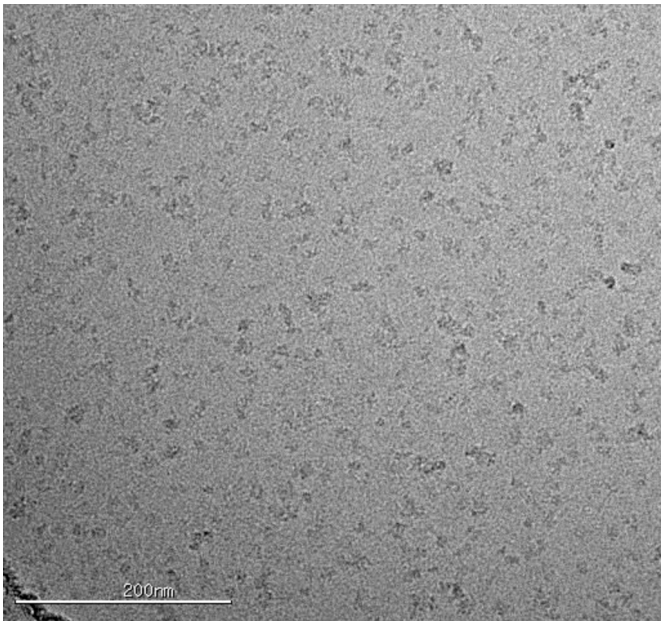
Finally, adsorption to any substrate may introduce preferred orientations or alterations in the physical properties of biological complexes.[2] No method is without faults and benefits must outweigh the risks when determining a proper protocol for specimen preparation.

3.2 Open hole vs. substrate grid preparation in solution

3.2.1 PRC2's preferred orientations on thin carbon substrates

To increase productivity and streamline the high-resolution EM pipeline for small complexes such as PRC2, we pursued a general method for specimen preparation that was easily

a)



b)

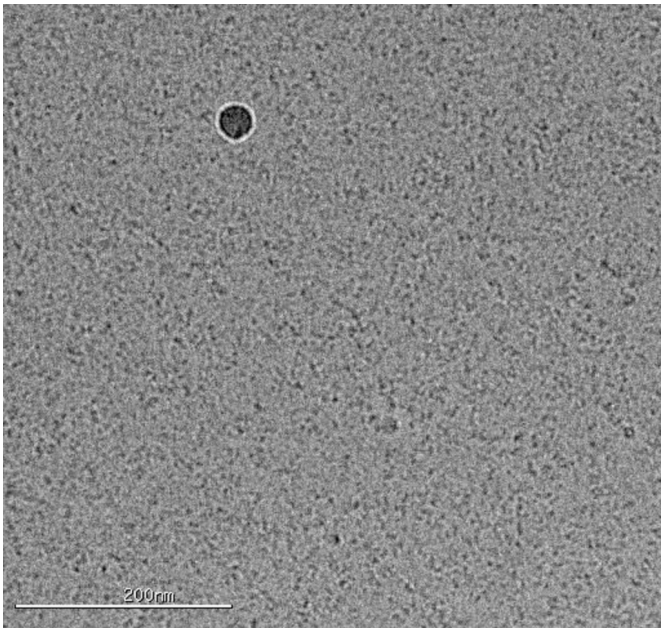


Figure 3.1 Comparison of PRC2 particles on continuous carbon grids and in open-holes. (a) Cryo-EM image of crosslinked PRC2 on thin continuous carbon substrate. (b) Cryo-EM image of PRC2 without continuous carbon or crosslinker illustrates that the complex falls apart or denatures in open-holes.

reproducible and guaranteed a uniform distribution of angular orientations. Additionally, we sought a method that ensured complex stability and homogeneity under cryo-conditions.

PRC2 has a strong tendency to adopt a preferred orientation on continuous carbon grids.[11] To compound the problem, the presence of a carbon substrate introduces background noise into the images, which may compromise higher order structural features given that biological macromolecules produce low contrast in EM because they scatter electrons only slightly more than their vitrified buffer surroundings.[2]

Additionally, determining thickness of carbon films remains mostly guesswork as deposition varies and the protocol is not easily reproducible.[2] Continuous carbon grids are created in-house by evaporating a thin carbon layer onto a freshly cleaved sheet of mica in a vacuum chamber. The carbon sheet is removed by dipping the sheet of mica into a well of distilled water at an angle. The hydrophobic nature of the carbon sheet causes it to float on the surface of the water, which is subsequently drained, allowing the sheet to settle over perforated carbon grids with regularly arrayed empty holes. These continuous carbon grids are normally treated by glow discharging, or cleaning via a mixed gas plasma in a vacuum chamber to render them hydrophilic for sample preparation. The introduced variability in the hydrophilicity of the carbon results in inconsistent ice thickness from grid to grid. [2]

Carbon backing presently provides a simple, well-documented solution to shield proteins from the denaturing effects of contact with the air-water interface, and adsorption on carbon may protect sensitive macromolecules from dissociating into constituent components prior or during freezing [Fig. 3.1(b)].[2,10] However, few controls exist that indicate binding does not adversely affect the overall structural integrity, and introduction of additional noise in images may constitute too formidable a challenge for obtaining high-resolution structures of small complexes.[2 , 7,10]

Given the skewed distribution of views and large variation between one grid preparation and another, an unreasonable number of images taken from ideal carbon substrate cryo-samples would be required to garner sufficient angular orientations and homogeneity to generate a high-resolution structure of PRC2. Additionally, small particles require greater defocus values to achieve distinction from their noisy environment.[2] This application of high defocus decreases spatial coherence, or the parallel illumination of the electron beam, and may further limit resolution.[2] Ultimately, reduced contrast caused by electron scattering by the variable ice and carbon substrate may place a physical limitation on achievable resolution regardless of other contributing parameters.[2]

3.2.2 Open-hole grid preparation: perspectives and challenges

Ideally, a sample consists of monodisperse, intact particles, adopting random orientations in open holes with an ice thickness no greater than the depth of field; ice thicker than this increases the number of scattered electrons and leads to loss of useful signal.[7] However, as discussed in prior sections, physical and chemical properties of the specimen all contribute to sample defects.

We used a Mark IV Vitrobot and Quantifoil 1.2/1.3R mesh holey carbon grids to prepare open-hole specimens for the five-component PRC2 complex [Fig. 3.2]. Samples were applied to grids within the Vitrobot, which was kept at 100% humidity and 24˚C; afterward, grids were blotted and immediately plunge-frozen. We hypothesized that limiting the amount of time the sample was exposed to the air-water interface prior to freezing would decrease contact between the proteins and the interface and minimize denaturation[7]. We exhausted the spectrum of blot

Figure 3.2: Table of open hole experiments

Experiment	Chloroform	No chloroform	Detergent	No Detergent	Glow Discharge	No glow discharge	Cross-linker	No Cross-linker	Results:
1		✓		✓		✓		✓	No particles observed
2		✓		✓		✓	✓		No particles observed
3		✓		✓	✓		✓		Less aggregation/still falling apart
4		✓	✓		✓		✓		Aggregation around rim
5	✓		✓		✓		✓		Aggregation around rim
6		✓		✓	✓			✓	Complex fell apart/denatured
7		✓	✓		✓			✓	Complex fell apart/denatured
8	✓		✓		✓			✓	Complex fell apart/denatured
9		✓	✓			✓		✓	No particles Observed
10	✓		✓			✓		✓	No particles Observed
11	✓		✓			✓	✓		No particles Observed
12	✓			✓		✓		✓	No particles observed
13	✓			✓		✓	✓		No particles observed
14	✓			✓	✓		✓		Less aggregation/still falling apart

force and blot time values for each experimental condition [Fig. 3.2] and concluded that 2s blot time +/- .5s with 10N blot force produced optimal ice thickness in all cases.

The bulk of open-hole experimentation involved combinatorically testing four factors hypothesized to stabilize the complex, ensure random orientations across the field of view, and limit contact with the air-water interface. These factors included addition of a crosslinker to the sample, washing grids with chloroform prior to sample application, use of detergent in the sample buffer, and glow discharging of the grids prior to sample application [Fig. 3.2].

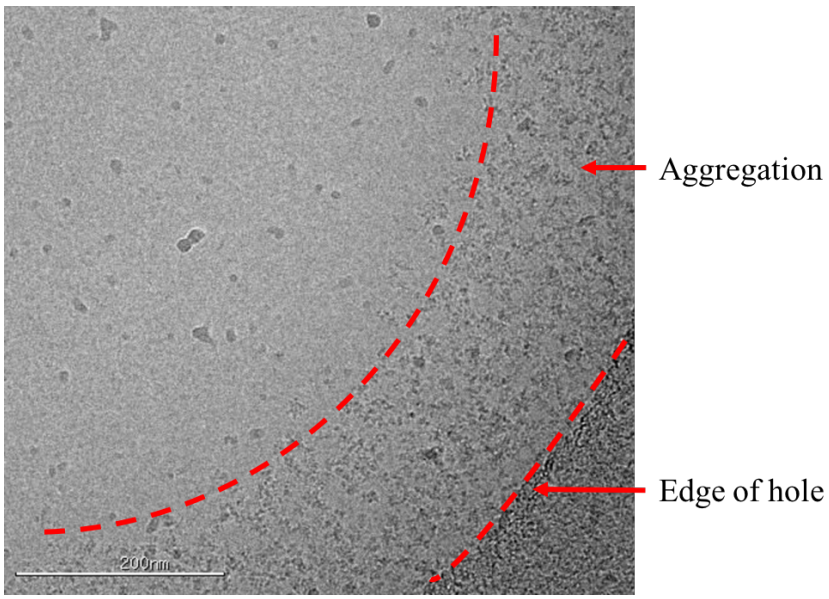
Comparing the sample concentration to the observed particle density in the holes is an indirect way of analyzing what may be physically occurring during grid preparation.[6] For instance, large quantities of protein bound to surrounding carbon but none visible in the holes suggests the protein has high affinity for the carbon, and that the remaining complexes are rapidly denatured upon contact with the air-water interface.[6]

We attempted to overcome aggregation and binding to the rim of the holes [Fig. 3.3(a)] by altering the surface properties of the carbon support. We hypothesized that glow discharging the grid before sample application, also called plasma cleaning, may cause PRC2 to bind with higher affinity along the rim and to the carbon support due to the increased hydrophilicity of the carbon after discharging.[7] Rinsing grids with chloroform prior to sample application was also tested in an attempt to wash away impurities and residual hydrocarbons deposited during the grid manufacturing process.[2,7] Neither one, nor various combinations of both, produced monodisperse, intact complexes in the holes [Fig. 3.3(b)]. We also tested crosslinking complexes prior to vitrification, but subunit dissociation was seen in all cases, indicating the vitrification process may break the fragile, flexible links binding individual subunits together despite the stabilization crosslinking should afford [Fig. 3.4]. Finally, we hypothesized that including the detergent NP-40 in the sample buffer may shield the complex from denaturation at the air-water interface by forming a protective surfactant barrier, and, due to increased viscosity, evenly spread particles across the field of view.[7] Unfortunately, in all instances, the detergent seemed to encourage aggregation or preclude particle migration into open holes [Fig. 3.3 & Fig 3.4].

Given the combinatorial testing of open-hole parameters and resulting negative conclusions, including subunit dissociation, denaturation at the air-water interface, aggregation along the rim, and propensity to stick to the carbon mesh support, we conclude open-hole specimen preparation is not a viable option for overcoming PRC2's preferred orientation problem on carbon substrates [Fig. 3.2].

These experiments demonstrate that PRC2 does not remain intact and may completely denature during vitrification. Therefore, an ideal grid support must shield the complex from the air-water interface, protect particles from physical stresses induced during freezing, while also preventing binding of the complex along a single axis.

a)



b)

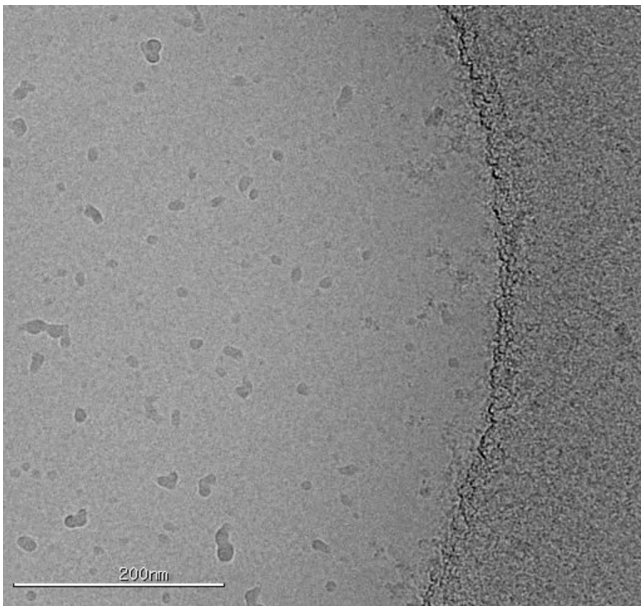
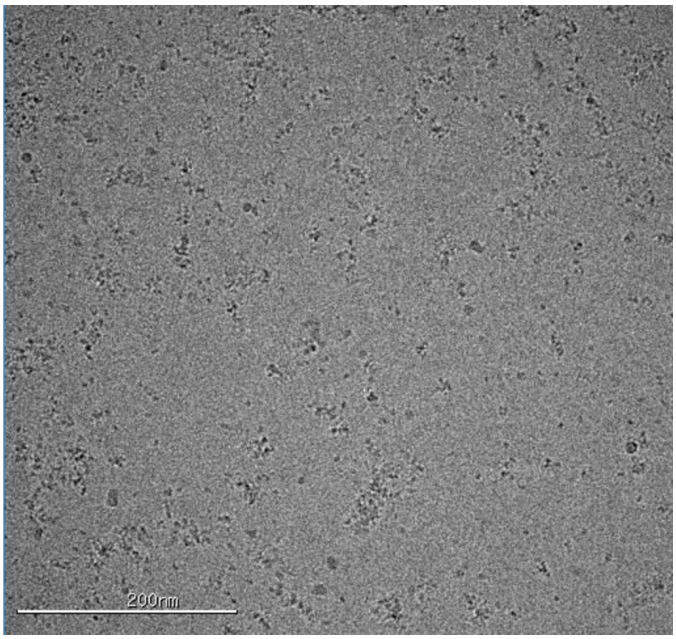


Figure 3.3 Images from two different open-hole grids showing results of using detergents and not glow discharging. a) Image shows sample aggregation along the rim following addition of detergent. b) Image showing failure of sample to spread into holes without glow-discharging grid prior to freezing.

a)



b)

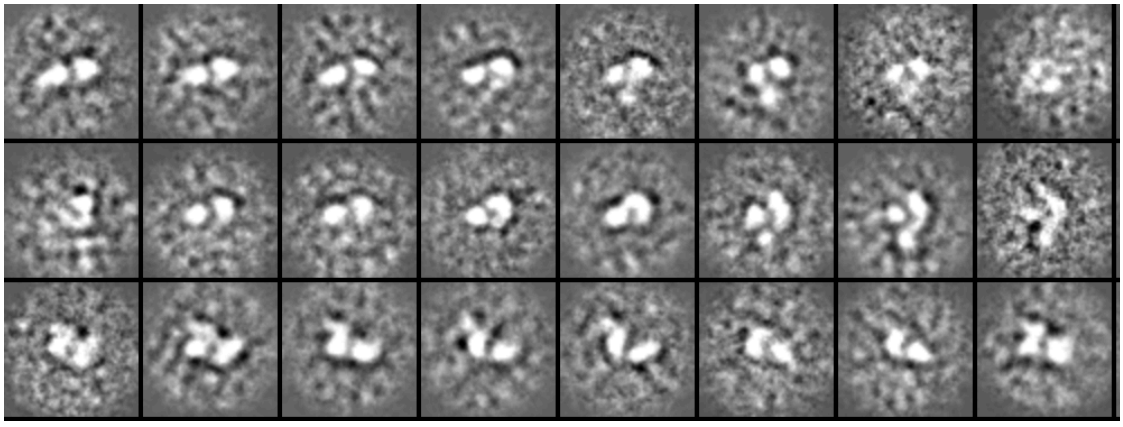


Figure 3.4 Image of crosslinked PRC2 on open-hole grids and 2D class averages. (a) Image shows crosslinked PRC2 on open holes with glow-discharging and no detergent. 9b) 2D class averages of crosslinked PRC2 on open hole grids reveal dissociation of complex subunits.

Chapter 4

Use of streptavidin monolayers as support films for small complexes

4.1 Rationale for choosing SA monolayers

Sample preparation remains one of the greatest challenges facing cryo-EM researchers.[10] Despite recent advances in data collection technology and algorithms to sort sample heterogeneity, grid preparation has persisted as a major impediment to achieving high-resolution reconstructions, particularly of small, asymmetric macromolecules.[10]

Although the breadth of physical and biochemical challenges is by no means catalogued in this thesis, two major factors have been thoroughly documented in the literature, and analyzed during the course of this study. First, the popular method of overlaying a thin film of continuous carbon across the copper mesh support grid may introduce background noise, compromise the native state of proteins when bound, or may introduce preferred orientations in oblong macromolecules containing highly charged localized regions.[6] Second, the technique of using open holes, as was demonstrated in Chapter 3, may cause aggregation, or partial or full denaturation of the complex upon interaction with the air water interface. Brownian motion may cause particles to contact the air-water interface, resulting in the formation of a thin film of denatured protein that serves as a barrier for further denaturation of the remaining particles.[6]

SA monolayers are grown in open hole carbon or gold mesh grids and consist of a biotinylated lipid monolayer with bound tetrametic streptavidin and a thin carbon backing providing additional support to the lipid tails [Fig. 4.1]. These monolayers have been shown to provide a robust structural support to reduce beam induced motion in cryo-EM.[35] Their use with ribosomes has been shown to prevent preferred orientations and produce monodisperse complexes.[6] SA monolayers are also highly uniform and provide reproducible ice thickness.[6]

The small, oblong nature of PRC2 and the tendency of the complex to aggregate and denature on open-hole grids, coupled with the need to collect thousands of micrographs to overcome the problem of preferential views in order to achieve higher resolution, makes SA monolayer grids a promising system for solving PRC2's structure. If fewer micrographs are

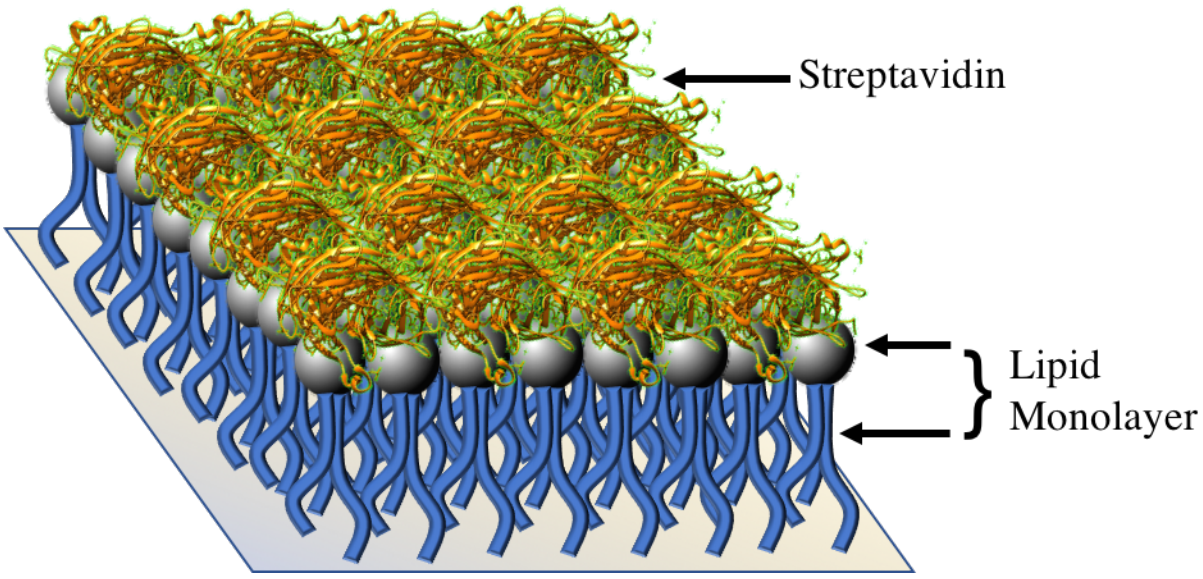


Figure 4.1 Streptavidin monolayer configuration. Situated in the open holes of the carbon mesh, biotinylated lipid heads are bound to a streptavidin monolayer crystal and a thin carbon sheet is cast over the lipid tail. Proteins are attached to the streptavidin monolayer via a biotin tag.

required to supply enough particles in each view, the burdens of massive data collection and heavy computational requirements will be alleviated.

4.2 Recent improvements in the reproducibility of SA monolayer grid preparation

The earliest use of monolayers as an affinity support system involved binding a biotinylated molecule to an SA crystal supported by a lipid monolayer.[36] Biotinylated ferritin was used to demonstrate that particles could be bound in high density to SA crystals. Later, algorithms were developed to subtract the signal contributed by the monolayers.[37] Since then, researchers have made great strides in producing streptavidin monolayer crystals that completely coat EM grids with high reproducibility and fidelity.[5]

Although the use of SA monolayers as EM substrates was first conceived of decades ago, the method for producing them remained too unreliable for widespread implementation.[5] Fourier subtraction of the Bragg diffraction spots necessitates a high degree of regularity and uniformity in the crystal lattice. Because the original protocol for preparing SA monolayers involved separately growing the crystals in wells and then transferring them onto EM grids, physical forces often fractured the fragile crystals, which caused uneven coverage of the crystals across the grid holes. A new protocol involves growing SA crystals directly on EM grids, generating consistent, easily reproducible lattices.[5]

Initially, the short shelf-life of SA monolayer crystals provided a narrow window for experimentation, imposing constraints on their practicality for studying many biological processes.[5] The new protocol involves embedding the monolayers in a protective surface barrier of trehalose, preventing their deterioration and allowing preparation of stock grids up to a month in advance.[5] Also, the use of the Edwards vacuum pump plasma cleaner to evaporate a very thin (~20 Å) layer of carbon on the bottom side of the grid further preserves monolayer stability [Figure 4.1].[5]

Most recently, 70s ribosomal particles were used as a model specimen to demonstrate the ability of SA monolayer support films to produce a high-resolution structure from fewer particles than are required without the monolayers as a result of the more even distribution of orientations on the monolayer grids.[5] In this thesis, I show that SA monolayers provide a more even distribution of orientations for a smaller complex (<500kDa): namely, PRC2.

4.3 Materials and Methods

4.3.1 SA monolayer preparation

We chose Quantifoil UltraAuFoil Holey Gold Films, rather than traditional copper with overlaid carbon, with micrometer-sized circular holes R2/2 for their added stability during irradiation. The gold support minimizes movement during imaging and provides greater reproducibility in the preparation of SA monolayers grids. Approximately two days prior to SA monolayer prep, these grids were rinsed three times in chloroform and then three times in ethanol to remove any impurities deposited during the manufacturing process. The grids were allowed to dry and a layer of carbon (~3 nm) was evaporated across the tops, rendering the grids hydrophilic. This added

layer of carbon sequesters the lipid monolayers into the open holes in the gold mesh during preparation.

On the day of SA monolayer preparation, grids were rinsed with ethanol to remove impurities deposited during the carbon evaporation process and placed carbon side up on filter paper to dry.

The lipids used were 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-(cap biotinyl) purchased from Avanti Polar Lipids at a concentration of 10mg/mL in a solution of chloroform/methanol/water. The lipids were diluted to 1mg/mL prior to use. Streptavidin (catalog number N7021S) was purchased from New England Biolabs at a concentration of approximately 1mg/mL. Both the lipids and streptavidin were stored at -80 ˚C for preservation. For more details on storage and allocation see Han *et al*.[5]

The following describes the protocol developed by Han *et al*.[5] for growing SA crystal monolayers directly on the gold mesh grids: we filled a small Falcon brand petri dish with crystallization buffer consisting of 50nM HEPES (pH 7.5), 0.15M KCl, and 10% trehalose.[5] A small amount of talcum powder was sprinkled along the rim of the dish and allowed to spread across the surface. A drop of castor oil was then applied to the surface of buffer. The oil pushes away the talcum powder and any contamination lingering at the surface of the air-water interface at the center of the dish.[5] The talcum powder spreading indicates the diameter of oil spread. The castor oil is hypothesized to act as a piston that exerts a consistent pressure on the lipid surface.

A Hamilton syringe was sterilized and decontaminated by repeatedly rinsing with chloroform before use. After rinsing, the syringe was used to deposit a small amount (<1 μL) of lipids directly onto the surface of the castor oil. The lipids are visible when viewed from a 25-degree angle from the surface of the well. The quantity of lipids deposited on the surface allows the user to make multiple grids from the same well.

The grids were then individually touched to the surface of the lipids, carbon-side down, and washed in three successive drops of crystallization buffer by gently touching the surface with the lipids to the top of the drop. The washing procedure rinses off any residual lipids on the mesh surrounding the open holes to ensure uniformity in crystal formation.

The grids were placed carbon-side up and 4 μL of streptavidin diluted to 0.5mg/mL was pipetted directly onto the grid. Grids were then allowed to incubate in a humidity chamber for approximately 30 minutes.

After incubation, unbound streptavidin was washed away by gently placing the grid, streptavidin-side down, onto a large droplet (~450 μL) of rinsing buffer consisting of 10mM HEPES (pH 7.5), 50nM KCl, and 10% trehalose for resin embedding. After a couple seconds the grid was lifted straight up from the drop, taking caution not to rupture the delicate, freshly formed crystal. The grid was blotted by touching a piece of filter paper to the side, wicking away any excess buffer and free streptavidin. The grid was then placed streptavidin-side up on a piece of filter paper and left to dry for 30 mins to 1 hr.

After drying, we flipped the grids over to expose the side with the lipid tails. The grids were placed in an Edwards vacuum chamber system and a very thin layer (~40 Å or less) of carbon was deposited on the back side of the grids for added structural stability. The grids were stored in a desiccation chamber to eliminate the possibility of moisture contaminating the structural integrity of the SA crystal.

4.3.2 Random chemical biotinylation of PRC2-AEBP2

Streptavidin homotetramers exhibit an incredibly high affinity for biotin in one of the strongest non-covalent interactions in nature. [5,8] This property, combined with a tolerance to large fluctuations in buffer conditions, makes streptavidin popular for biochemical studies and protein purification.[5,8] Most importantly, once biotinylated proteins bind to streptavidin monolayers, physical forces applied during blotting and plunge-freezing are highly unlikely to disrupt the interactions.[5,8]

We investigated three possible methods for biotinylating PRC2: addition of a biotin-conjugated H3K27me3 peptide, crosslinking using the biotinylated homobifunctional crosslinker CBDPSS-H8/D8, and random chemical biotinylation of lysine residues using Sulfo Chromalink Biotin.

Recently, the Ezh2-EED-Suz12(VEFS) crystal structure with bound H3K27me3 peptide showed an allosteric modification of the catalytic domain, shifting PRC2 into an active state. [38] We initially tested the peptide-biotin conjugate because crystal structure docking into a high-resolution map of the five-component complex would provide model validation and key insights into PRC2's role in gene silencing as described in 2.3.1.

For the peptide-biotin conjugate tests, PRC2 was stored in aliquots of 5 μL at -80˚C in a solution of 25 mM HEPES (pH 7.5), 150mM NaCl, and 10% glycerol, which served as a cryoprotectant. An aliquot was thawed at room temperature prior to chemical modification.

The sample was incubated with a 10:1 ratio of peptide to protein for 30 minutes at room temperature and then crosslinked with 2% BS3 for an additional 30 minutes, then buffer exchanged into 25mM HEPES (pH 7.5) with 150mM NaCl to remove glycerol and excess biotin, which may outcompete protein during monolayer attachment. As illustrated in Chapter 3, crosslinking prevented dissociation of the complex during sample application, blotting, and vitrification. Crosslinking was also necessary to stabilize the flexible regions of PRC2 that were shown to dissociated by physical strain during desalting.

Following buffer exchange, protein concentration was quantified using a spectrophotometer. The sample was diluted to 100 nM and 4uL was incubated for 30 minutes on the SA grid. After following the negative staining protocol described in 4.3.3, we collected a series of micrographs that revealed intact monolayers, but no bound protein. After multiple trials altering incubation times and crosslinker concentrations, we concluded that the H3K27me3 peptide failed to bind with sufficient affinity to allow for complex binding to the grid. Alternatively, crosslinker may preclude peptide binding to the active site.

We then tested crosslinking PRC2 with the biotinylated, homobifunctional crosslinker CBDPSS-H8/D8 in a similar ratio as mentioned in the prior protocol. CBDPSS-H8/D8 crosslinks primary amino groups and we aimed for a 1:1 ratio of biotin to complex, thus limiting the amount of crosslinker available per complex. Therefore, we included an additional15 min incubation with BS3 to reinforce complex stability. Micrographs of the desalted and stained complex revealed large aggregates and contaminants, which persisted despite a number of experimental modifications. We hypothesize that CBDPSS-H8/D8 and BS3 may have worked cooperatively to intramolecularly crosslink PRC2, resulting in the observed aggregation.

Finally, we attempted random chemical biotinylation using Sulfo Chromalink Biotin added in a 10:1 ratio to a fresh PRC2 sample. The solution was crosslinked with BS3 and subsequently desalted. Before and after desalting, we quantified the ratio of biotin to protein by targeting a traceable bis aryl hydrozone chromophore on the reagent. Spectrometry revealed a ratio of 0.9 biotin per complex. Randomly chemically biotinylated PRC2 successfully bound to SA monolayers in a wide array of angular distributions [Fig. 4.2].

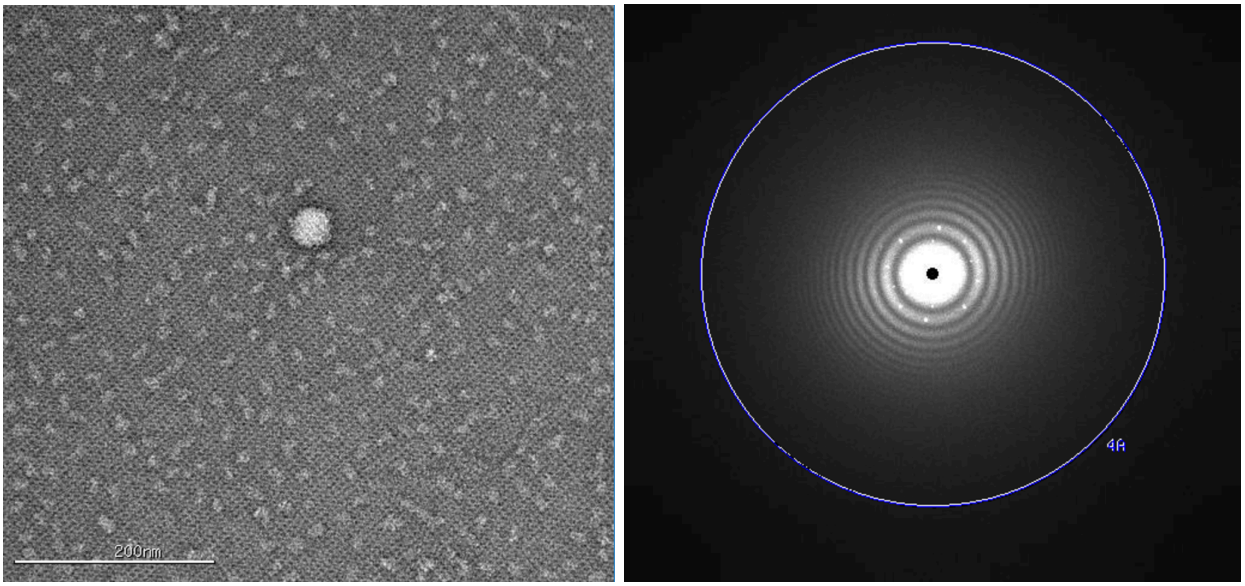


Figure 4.2 Negative stain image of PRC2 bound to streptavidin monolayer and corresponding FFT. Negative stain image of streptavidin monolayers with randomly chemically biotinylated PRC2 bound to the crystal lattice (left) and FFT of micrograph clearly showing the lattice repeats within the Thon rings (right).

4.3.3 Negative stain grid preparation of biotinylated PRC2-AEBP2 bound to SA monolayers

SA monolayer grids were rehydrated by repeated washing with rehydration buffer (25 mM HEPES pH 7.5 with 150 mM NaCl). The grids were gently touched to the top of two 50 µL drops of buffer and lifted straight up so as not to disturb the crystal. A third 100 µL drop was used for the final rehydration step, during which the grid was left floating streptavidin-side down for 10 minutes, supported by surface tension. The entire rehydration procedure was repeated at least once more to ensure all excess trehalose and other contaminants were washed away during the process.

An additional buffer exchange step into the sample's buffer may be added immediately after the rehydration process should the sample in question be easily affected by changes in buffer conditions.[5] However, we found the structural integrity of the crystal not to be impacted by using the buffer for PRC2 rather than that used in the original protocol, which was the rinsing buffer without the additional trehalose.[5]

Biotinylated PRC2 was diluted to an approximate concentration of 100 nM and 4 µL was applied to the SA side of the grid. The contents were gently mixed by pipetting up and down on the surface of the grid. Subsequently, the grid was placed in a hydration chamber and allowed to incubate for 25 minutes. Following incubation, unbound sample was washed away by touching the liquid surface twice to three successive drops of PRC2-AEBP2 buffer.

After washing away unbound sample, the grid was flipped over and excess liquid pipetted off the top. 4 µL of 2% UA was applied to the grid and gently mixed by pipetting. The UA was allowed to incubate for 1.5 minutes. 3uL of UA was pipetted from the surface and replaced with an additional 3 µL of 2% UA and incubated for another 1.5 minutes. The procedure was repeated once more to ensure a uniform distribution of stain across the crystal.

One particularly challenging aspect of negatively staining monolayer grids was the tendency of the grids to dry out after blotting, leaving behind halos of stain around the particles, rendering the sample unsuitable for data collection. We overcame this challenge by leaving behind a thin layer of liquid after blotting to guarantee the stain remained in the holes containing the monolayers. This final step ensured the grids had an even distribution of stain with distinct, monodisperse particles across the field of view.

A control grid without monolayers was simultaneously produced from the same sample of biotinylated PRC2 and negatively stained for comparative imaging.

4.4 Data collection and processing

Micrographs were collected using a T20F Twin-TEM equipped with an FEG operating at 120 keV at a magnification of 80,000X, resulting in a 1.5 Å/pixel size at the level of the detector. Images were recorded in a semi-automated manner using the LEGINON data collection software package with a defocus range of -0.7 to -1.2 µm using a Gatan CCD camera with a pixel size of 15 µm.

Monolayers were assessed for quality by visual inspection of the images and observation of lattice repeats in the FFTs [Fig. 4.2]. Clear, concentric Thon rings with obvious Bragg peaks, or spots in the FFT indicating lattice repeats, served as the criteria for selecting micrographs. Periodic signal contributed by the crystal was removed from the images by Fourier transform and subsequent masking out the diffraction pattern [Fig. 4.3]. The amplitudes of neighboring pixels

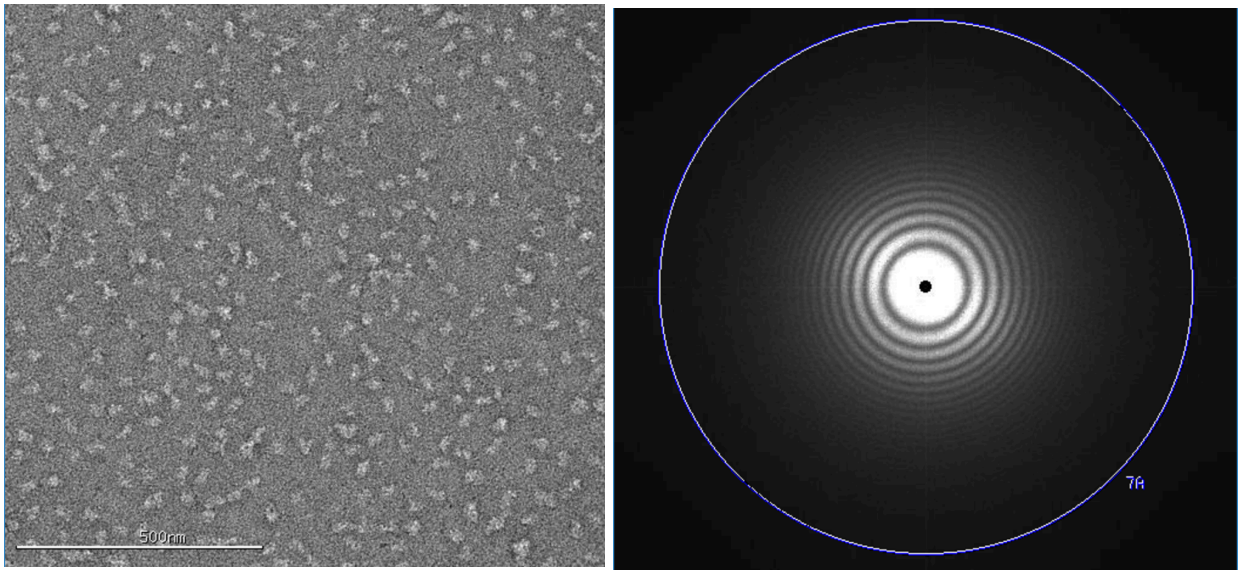


Figure 4.3 Image and FFT following background subtraction of crystal lattice. Image of PRC2-AEBP2 bound to streptavidin monolayer after background subtraction (left) and removal of the crystal diffraction pattern from the FFT (right).

were averaged and these values were assigned to the Bragg peaks. The inverse Fourier transform restored particle information to the micrographs without the introduction of artifacts in the resulting 2D class averages of PRC2[Fig. 4.4].

Images were preprocessed using the RELION 1.4 software package.[9] During preprocessing, particles were picked using a combination of the automated DogPicker software and manual selection. CTF values were calculated in tandem using ACE2 and CTFFind4 and images exhibiting astigmatism, or asymmetry in the FFT, were discarded. Particles were phase-flipped and extracted with a box size of 224x224 pixels then binned by a factor of 2 in RELION.[9] 2D class averages were generated in RELION and junk particles were discarded.

The resulting 2D class averages show an array of orientations for the PRC2, with clear structural features and distinct lobes [Fig. 4.4]. The preferred orientations normally seen when imaging PRC2 in negative stain without SA monolayers were observed in the control [Fig4.5]. When compared to the control grid without monolayers, chemically biotinylated PRC2 exhibited a broader spectrum of views throughout 2D classification. [Fig. 4.4 & 4.5]

Approximately 9,000 selected particles were processed using RELION 3D refinement with a regularization parameter (T) of 3, which specifies the weight of the experimental data versus the prior. After multiple rounds of refinement, a structure at a resolution of 28 Å was generated along with the 3D angular distribution maps, which indicated the particle angular spread across the surface of the unit sphere when compared to the particle distribution [Fig. 4.6].

4.5 SA monolayer support films are a viable option for small complexes

4.5.1 Random biotinylation allowed PRC2 to adopt multiple views on the streptavidin monolayer grids

Our experiments tested two different substrate support grids, SA monolayer and continuous carbon (control), that were produced by incubating the same biotinylated sample of PRC2 and subsequently negatively stained. No additional changes to buffer or staining conditions were made between grids when preparing the specimens. The data show that the angular distribution of biotinylated PRC2 particles across the unit sphere was greater in the SA monolayer reconstruction [Fig. 4.6]. We conclude the biotinylation process does not impact the structure or distribution of views of PRC2; only the binding to the streptavidin monolayer produces a change in orientation.

When compared to the published low-resolution map of PRC2, no apparent artifacts from masking out the streptavidin lattice from the FFTs were generated in the final EM density map. We conclude that SA monolayer support films are a viable option for cryo-EM specimen preparation for PRC2.

Biotinylated PRC2 binds with a more even distribution of orientations on SA monolayer substrates than it does on other substrates. Future researchers may use the protocols and conclusions outlined throughout this thesis for further pursuit of high-resolution cryo-EM density maps of PRC2.

These established and tested protocols may extend to specimen preparation of other small, flexible complexes that remain a challenge for structural biology. This thesis contributes to further expanding and enhancing the methodologies used in high-resolution cryo-EM.

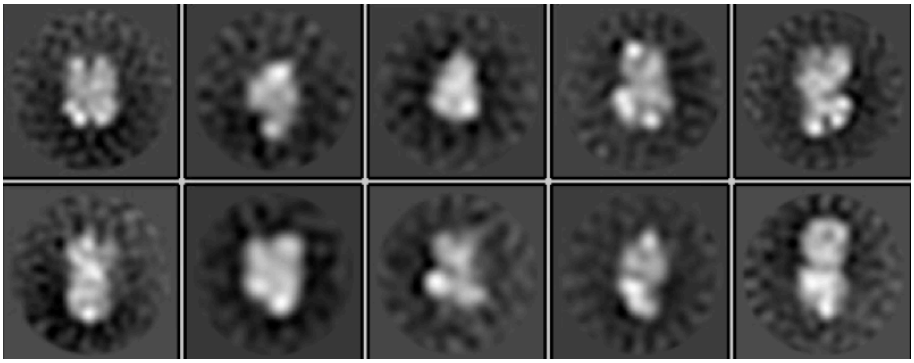


Figure 4.4 2D class averages of PRC2 calculated from SA monolayer substrate grids. 2D class averages of negatively stained PRC2 after chemical biotinylation and tethering to the streptavidin monolayer crystal generated after background subtraction of the crystal lattice. Class averages contained 200-300 particles per class from ~9000 total particles

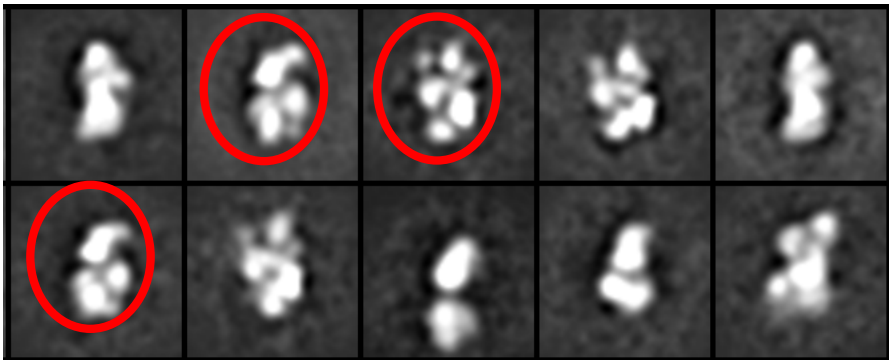


Figure 4.5 2D class averages of PRC2 calculated from carbon substrate grids. 2D class averages of negatively stain, biotinylated PRC2 on continuous carbon (control). Class averages contained 200-300 particles per class from ~9000 total particles. The red circles highlight the dominant view shown throughout the averaging.

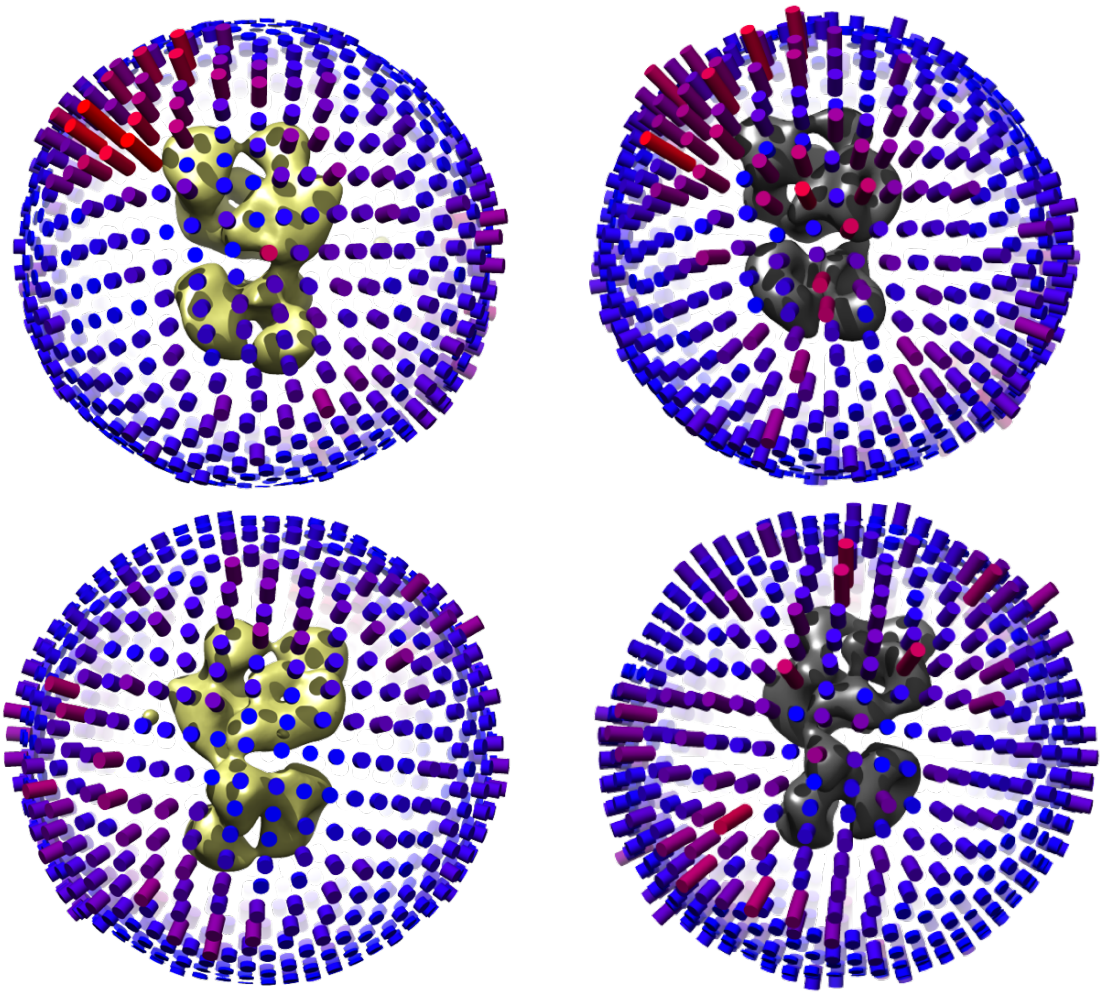


Figure 4.6 Angular distribution plots and low-resolution reconstructions for PRC2 imaged on continuous carbon and SA monolayer substrates. Euler angles distribution plots for different reconstructions of PRC2-AEBP2 samples imaged on continuous carbon (right) and bound to an SA monolayer substrate (left). Each reconstruction included ~9,000 particles. The depth of the rods along with the color are indicative of the number of particles from that view. As can be observed from the distribution on the right, a cluster of long, red rods appear at the top, indicating that a majority of particles on the continuous carbon grid were arranged in that orientation. The rest of the angular distribution plot is mostly short, blue rods, indicating that there were very few other views on the grid. The particle distribution collected from the SA monolayer data on the right shows a more widespread array of light red and purple rods of varying length. This indicates that randomly biotinylated PRC2 has a tendency to bind in multiple orientations on the SA monolayer grids.

Bibliography

1. Nogales, E. (2016). The development of cryo-EM into a mainstream structural biology technique. Nature Methods, 13(1), 24–27.
2. Agard, D., Cheng, Y., Glaeser, R. M., & Subramaniam, S. (2014). Single-particle cryo-electron microscopy (cryo-EM): Progress, challenges, and perspectives for further improvement. In Advances in Imaging and Electron Physics (Vol. 185, pp. 113-137). (Advances in Imaging and Electron Physics; Vol. 185). Academic Press Inc.
3. Rebecca F. Thompson, Matt Walker, C. Alistair Siebert, Stephen P. Muench, Neil A. Ranson, An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology, Methods, Volume 100, 1 May 2016, Pages 3-15, ISSN 1046-2023
4. Cheng Y. Single-particle cryo-EM at crystallographic resolution. Cell. 2015;161(3):450-457.
5. Han B-G, Watson Z, Kang H, et al. Long shelf-life streptavidin support-films suitable for electron microscopy of biological macro-molecules. Journal of structural biology. 2016;195(2):238-244.
6. Robert M. Glaeser,Bong-Gyoon Han. Opinion:hazards faced by macro-molecules when confined to thin aqueous films[J]. Biophysics Reports, 2017, 3(1-3): 1-7.
7. Factors that Influence the Formation and Stability of Thin, Cryo-EM Specimens. Robert M. Glaeser, Bong-Gyoon Han, Roseann Csencsits, Alison Killilea, Arto Pulk, Jamie H.D. Cate. Biophys J. 2016 Feb 23; 110(4): 749–755. Published online 2015 Sep 17.
8. Bong-Gyoon Han, Zoe Watson, Jamie H.D. Cate, Robert M. Glaeser, Monolayer-crystal streptavidin support films provide an internal standard of cryo-EM image quality, Journal of Structural Biology, 2017, ISSN 1047-8477
9. Scheres, 2012. S.H.W. Scheres RELION: implementation of a Bayesian approach to cryo-EM structure determination J. Struct. Biol., 180 (2012), pp. 519-530
10. How good can cryo-EM become? Robert M Glaeser Nature Methods 13, 28–32 (2016)
11. Ciferri, C., Lander, G.C., Maiolica, A., Herzog, F., Aebersold, R., and Nogales, E. (2012). Molecular architecture of human polycomb repressive complex 2. elife 1, e00005.
12. Margueron R, Reinberg D. 2011. The Polycomb complex PRC2 and its mark in life. Nature 469:343–9.

13. O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol 21:4330–6.
14. Margueron R, Li G, Sarma K, Blais A, Zavadil J, Woodcock CL, et al. 2008. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol Cell 32:503–18.
15. Cao R, Zhang Y. (2004a). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr Opin Genet Dev 14:155–64.
16. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, et al. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298:1039–43.
17. Han Z, Xing X, Hu M, Zhang Y, Liu P, and Chai J. (2007). Structural basis of EZH2 recognition by EED. Structure 15:1306–15.
18. Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, et al. 2008. A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol 10:1291–300.
19. Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ III, et al. 2009. Role of the polycomb protein EED in the propagation of repressive histone marks. Nature 461:762–7.
20. Montgomery ND, Yee D, Montgomery SA, Magnuson T. Molecular and functional mapping of EED motifs required for PRC2-dependent histone methylation. J Mol Biol (2007); 374: 1145–57.
21. Cao R, Zhang Y. (2004b). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol Cell 15:57–67.
22. Schmitges FW, Prusty AB, Faty M, Stutzer A, Lingaraju GM, Aiwazian J, et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. Mol Cell 42:330–41.
23. Tan, J. Z., Yan, Y., Wang, X. X., Jiang, Y., & Xu, H. E. (2013). EZH2: biology, disease, and structure-based drug discovery. Acta Pharmacologica Sinica.
24. Van Aller, G. S., Pappalardi, M. B., Ott, H. M., Diaz, E., Brandt, M., Schwartz, B. J., ... & Kruger, R. G. (2013). Long Residence Time Inhibition of EZH2 in Activated Polycomb Repressive Complex ACS chemical biology.
25. Yuan, W., Wu, T., Fu, H., Dai, C., Wu, H., Liu, N., Li, X., Xu, M., Zhang, Z., Niu, T., et al. (2012). Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. Science 337, 971–975.
26. L. Jiao, X. Liu, Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2, Science 350 (2015) aac4383.
27. S. Antonysamy, B. Condon, Z. Druzina, J.B. Bonanno, T. Gheyi, F. Zhang, I. MacEwan, A. Zhang, S. Ashok, L. Rodgers, et al., Structural context of disease associated mutations and putative mechanism of autoinhibition revealed by X-ray crystallographic analysis of the EZH2-SET domain, PLoS One 8 (2013) e84147.
28. O'Meara, M.M., and Simon, J.A. (2012). Inner workings and regulatory inputs that control Polycomb repressive complex 2. Chromosoma 121, 221–234.
29. Boyer LA, Latek RR, and Peterson CL. (2004) The SANT domain: a unique histone-tail-binding module? Na tRev Mol Cell Biol 5:158–63.
30. Zhang, X., & Bruice, T. C. (2008). Enzymatic mechanism and product specificity of SET-domain protein lysine methyltransferases. Proceedings of the National Academy of Sciences, 105(15), 5728-5732.

31. Tie F, Stratton CA, Kurzhals RL, Harte PJ. (2007). The N terminus of Drosophila ESC binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3 lysine 27. Mol Cell Biol 27:2014–26.
32. Margueron R, Reinberg D. (2010). Chromatin structure and the inheritance of epigenetic information. Nat Rev Genet 11:285–96.
33. Kim H, Kang K, Kim J. 2009. AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. Nucleic Acids Res 37:2940–50.
34. Glaeser RM, Han B-G, Csencsits R, Killilea A, Pulk A, Cate JHD. Factors that Influence the Formation and Stability of Thin, Cryo-EM Specimens. Biophysical Journal. 2016;110(4):749-755.
35. Bong-Gyoon Han, Zoe Watson, Jamie H.D. Cate, Robert M. Glaeser, Monolayer-crystal streptavidin support films provide an internal standard of cryo-EM image quality, Journal of Structural Biology, 2017, ISSN 1047-8477,
36. L.G. Wang, P. Ounjai, F.J. Sigworth Streptavidin crystals as nanostructured supports and image-calibration references for cryo-EM data collectionJ. Struct. Biol., 164 (2008), pp. 190-198
37. Wang, L.G., Sigworth, F.J., 2010. Liposomes on a streptavidin crystal: a system to study membrane proteins by cryo-EM p. 147–164. In: Jensen, G.J. (Ed.), Methods in Enzymology, Vol 481: Cryo-Em, Part a – Sample Preparation and Data Collection, Vol. 481, pp. 147–164.
38. N. Justin, Y. Zhang, C. Tarricone, S.R. Martin, S. Chen, E. Underwood, V. De Marco, L.F. Haire, P.A. Walker, D. Reinberg, et al., Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2, Nat. Commun. 7 (2016) 11316.