Florida State University Libraries

Electronic Theses, Treatises and Dissertations The Graduate School

2017

Segmentation and Structure Determination in Electron Microscopy

Chaity Banerjee Mukherjee

Follow this and additional works at the DigiNole: FSU's Digital Repository. For more information, please contact lib-ir@fsu.edu

FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

SEGMENTATION AND STRUCTURE DETERMINATION IN ELECTRON MICROSCOPY

By

CHAITY BANERJEE MUKHERJEE

A Dissertation submitted to the
Department of Computer Science
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2017

Chaity Banerjee Mukherjee defended this dissertation on July 28, 2017.
The members of the supervisory committee were:

Xiuwen Liu
Professor Co-Directing Dissertation

Kenneth A. Taylor
Professor Co-Directing Dissertation

Adrian Barbu
University Representative

Piyush Kumar
Committee Member

Gary Tyson
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.

This dissertation is dedicated to my parents and husband.

ACKNOWLEDGMENTS

When I started in the Ph.D. program I did not realize that it was a very long marathon where there great many pitfalls waiting to consume the unsuspecting traveler. When I realized the grind that was involved initially I was afraid. But I soon realized that I was fortunate to be under the supervision of two wonderful and experienced professors with whose direction and foresight I could reach my destination without getting into the pitfalls. The fact that I am writing my dissertation today is ample proof of the help and encouragement that I have got from both of them. I was admitted to the Computer Science department under the supervision of my adviser Dr. Xiuwen Liu and with his encouragement I started working in the Institute of Molecular Biophysics under the supervision of Dr. Kenneth Taylor. Without their direction and constant encouragement at every step of the program, I would not have been where I am today. I would like to thank both of them for their professional guidance throughout my Ph.D. work. I have learned a lot, both professionally and personally from both of my advisers and I look up to both of them as my role models, what I want to be going forward in my life.

I would also like to thank Dr. Kenneth Roux, from the Department of Biological Science at FSU for his help and guidance during the first project that I completed as part of my dissertation work. Dr. Roux was kind enough to provide me with data and funding for doing research during my initial years at Florida State University and I would like to express my gratitude to him for having faith in me during that time.

I would like to thank Mrs. Dianne Taylor, who has been like a loving guardian to all of us in Dr. Taylor's lab. Apart from that her expertise in specimen preparation is at the heart of the

success that we have had with all the data that we ever used. I always felt her motherly presence in the lab and would like to thank her for being so nice to me always.

I would like to thank my collaborator Dr. Susan Lowey for providing me with the sample for ACTO-MD, the analysis of which became a part of this dissertation. I am also thankful to Mr. Anthony Worrington of the Taylor Lab for collecting the ACTO-MD data from this sample.

I would also like to take this opportunity to thank past and present members of the Taylor Lab: Dr. Hanspeter Winkler for helping me understand the PROTOMO and I3 software packages and providing the insights for building on the same. Dr. Zong Huang, Dr. Zongjun Hu, Dr. Guiqing Hu, Dr. Aguang Dai, Dr. Zhuan Qin, Dr. Claudia Arakelian for all their help with my daily work. I have been fortunate enough to have had them as my co-workers in the lab and have benefited numerous times from their knowledge about the subject.

I would like to thank Dr. Nan Zhao from Dr. Liu's lab for helping me numerous times with his expertise in understanding of segmentation algorithms. I would also like to thank Dr. Moumita Dutta from Dr, Roux's lab for collecting and sharing the HIV/SIV data with me.

My Ph.D. committee members: Dr. Piyush Kumar, Dr. Gary Tyson and Dr. Adrian Barbu helped me through their constructive criticisms of my work. I would like to express my gratitude to them for their time and patience.

Finally, I would like to thank my husband Dr. Tathagata Mukherjee for his help and encouragement. His being a computer scientist himself helped me a lot in the form of many discussions that we have had on different problems. At the end, I would like to thank my

parents without whom I would never have been born. I thank them for being there for me always and raising me to be what I am today.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The determination of the structure of macro-molecules is one of the first steps in better understanding their functionality. This in turn is useful for understanding how the basic building blocks of life come together to create so many different life forms on this wonderful planet. It also helps us to understand the inner workings of infection causing bacteria and viruses that reek havoc on human civilization without the invention of drugs that can effectively fight the diseases caused by them. Electron microscopy is one of the most effective tools in elucidating and understanding the structure of biological macro-molecules. Through the use of single particle electron microscopy and electron tomography, homogeneous and heterogeneous macro-molecular assemblies have been imaged and studied respectively. In spite of the advances in the implementation of these techniques, still there are problems that are either not well understood or that beg for more automation. This dissertation studies two such problems: one in the realm of tomography and the other in single particle electron microscopy. More precisely, we study the problem of automatic segmentation of heterogeneous macro-molecular structures in 3D volumes obtained from electron tomography. We describe a new learning based method, segmentation by classification that we developed and implemented to address this problem. We report the results of using this algorithm for segmenting the HIV/SIV envelop spikes. For single particle electron microscopy, we study the problem of structure determination of filaments with helical symmetry using filamentous actin in complex with the smooth muscle myosin motor domain, otherwise known as acto-MD. The acto-MD structure provides deeper insights into how muscles work in general and has the potential to impact treatment of hypertension in human physiology.

CHAPTER 1

INTRODUCTION

One of the goals of biology is to be able to understand the structure and interaction of macromolecules, to be able to better understand life at a macromolecular level. One of the most important inventions that revolutionized the study of macromolecular structures is that of the electron microscope [1]. Electron microscopes are used for studying three-dimensional (3D) structures of macromolecular assemblies using two-dimensional (2D) and 3D geometry. The underlying principle of 3D reconstruction from 2D projections is well understood and forms the basis of electron microscopy [2]. Electron microscopy can broadly be categorized as electron tomography and single particle electron microscopy depending on the type of structure under investigation. Electron tomography is method for reconstructing the interior of an object from its projections [3] and is used in those cases where the structures are heterogeneous. In case that they are homogeneous, single particle electron microscopy is used. Single particle electron microscopy has the ability to provide 3D structural information of biological molecules and assemblies by imaging non-crystalline specimens (single particles) [4]. Whatever be the underlying source of the data, tomography or single particle, they involve significant amounts of computational analysis of the images. Many of these problems have been studied in other branches of computer science, like in computer vision and machine learning. However, until very recently, there has not been a significant exchange of ideas between these two disparate communities. This proposal is a step in that direction. We study two problems: the first related to the well-studied problem of segmentation but in the context of electron tomography. The second involves study of a macromolecular structure, the actin-myosin interaction, using single particle 3D reconstruction from electron microscope images. We hope that this effort would be the

beginning of a formal interaction between the two fields with the potential to enrich each other tremendously.

The first problem that we study in this dissertation, is that of the well-studied problem of segmentation in 3D volumes. Segmentation is generally known as the process of feature extraction from a contextual image. However, instead of applying it to point clouds, as is often done in the world of Computer Science [5], we apply it to the volume reconstructed using cryoelectron tomography (cryoET). More precisely, we study the problem of segmentation of Human Immuno Deficiency Virus (HIV)/ Simian Immuno Deficiency Virus (SIV) envelope spikes from electron tomograms of frozen-hydrated virion suspensions. We start with describing the importance of the problem and then briefly describe our proposed approach. Finally, we conclude with a list of items that we will discuss in the final dissertation.

The second problem that we discuss relates to determination of the structure of actin-myosin in in what is known as the nucleotide-free state, also known as the rigor state, using proteins derived from vertebrate smooth muscle. This study involves 3D image reconstruction from electron micrograph of frozen-hydrated filaments using what are broadly known as single particle reconstruction methods followed by analysis of the resulting 3D volume to understand the underlying macromolecular assembly more accurately. As before we first discuss the importance of the problem and then give a description of our approach. We also identify ways in which the process might be improved so that conformational heterogeneity can be satisfactorily dealt with.

1.1 3D Segmentation of Volumetric Data

AIDS, caused by Human Immuno Deficiency Virus (HIV) infection is one of the biggest killers in the history of human civilization [6]. Currently there are no vaccines for this deadly virus. The primary reason for this, is the lack of understanding of the 3D structure of the Human/Simian Immuno Deficiency Virus (HIV/SIV) envelope spikes (Env) and how it is able to escape the bodies ability to generate antibodies. Env is the only virus protein that presents antigens to the human immune response and is thus the key target for neutralization by antibodies that can bind and prevent its entry into target cells. So, understanding the structure of the spikes may eventually help create a vaccine against AIDS. Envelope spikes therefore have been the subject of intense research activity. The structure of a typical virus consists of an inner nucleic acid core surrounded by a protein membrane called the envelope. Env is a glycoprotein, which facilitates entry into the cell via the cell surface receptors CD4 and the chemokine receptors, CCR5 and CXCR4 [7] HIV-1 receptors and cell tropism [8]. CD4 is other-wise known as the cell surface marker for differentiating a class of lymphocytes known as “Helper T-cells”.

In HIV/SIV, the Env spikes comprise two glycoproteins gp120 and gp40, which are cleavage products of a larger protein called gp160. In order to study and ultimately unravel the detailed structure of these spikes, scientists have resorted to cryoET. One of the major problems after creating the tomogram is the segmentation of the Env spikes for further analysis. In this case, segmentation is confined to selecting the Env spikes from the viral membrane for further study. All cryoET volumes, including the HIV/SIV data under study, have very low signal-to-noise ratio (SNR). Moreover, the envelope spikes are distributed across the virus envelope with some tendency to cluster [9]. Therefore, identification of each individual spike manually, needs a great

deal of human intervention and is prone to errors. This work aims to alleviate this problem by proposing a semi-automatic 3D segmentation mechanism using statistical tools.

1.2 Tomography

cryoET has become a powerful tool for revealing 3D structures of macromolecular assemblies. Tomography is a method for reconstructing the interior of a 3D object from its projections. In cryoET, a Transmission Electron Microscope (TEM) is used to collect the 2D projection images in various orientations of a biological specimen frozen in vitrified ice. The specimen is usually rotated about one single axis, called the tilt axis, which is perpendicular to the optical axis of the microscope. In case of dual-axis tomography, the specimen is rotated around two different axes perpendicular to each other. A series of projection images, called the tilt series, are collected for each and every different orientation of the specimen. The whole angular range of the rotation is generally limited to ± 70° and the angular increment for rotation is usually 1°-5°. The projection images are then aligned using one of the two widely used alignment techniques based on either localizing fiducial markers or pattern matching with cross-correlation. Finally, a volume reconstruction is computed from the aligned images using the weighted back projection method [10].

1.2.1 Challenges in Tomography

Despite being the most efficient method of providing a 3-8 nm resolution image of complex biological specimens, cryoET has several inherent challenges. The fundamental reason for the hardships of the cryoET reconstruction results from the fact that most specimens can only tolerate a limited amount of electron dose. This limited electron dose is spread over the entire tilt series, resulting in low SNR in each member of the tilt series, which in turn has a negative effect

on reconstruction quality. The very high radiation sensitivity of the specimen enforces the use of very low electron exposure while recording images, greatly increasing the stochastic noise.

Apart from the problem of poor SNR, there are many other facts that make the interpretation of electron tomograms extremely challenging. The contrast in cryo-tomograms is often low and non-uniform along membranes and fibers, creating vague contours that are difficult to interpret. Moreover, the presence of well-defined point-like objects with strong uniform contrast along with the sample preparation artifacts makes the interpretation of the tomogram even more difficult.

Another most important challenge is data loss due to the "missing wedge". As the angular range of tilting is limited to ± 70°, a significant part of the potential specimen views is not accessible, leaving a missing wedge in the collected data {probably needs a reference}. This problem is most prominent in single-axis tilt series. Because of the missing wedge problem, some parts of the image may have significantly lower contrast than others. Dual-axis tilting can reduce the missing wedge to a missing pyramid, but if the total exposure is limited by radiation damage, the angular increment between views must be increased. Some elaborate tilt angle schemes have been proposed to optimize the dual axis tilting range while minimizing the number of projections needed for a complete reconstruction [11].

1.3 The General Segmentation Problem

Interpretation in cryoET requires segregation of the different features of interest in the tomograms, which includes structures like membranes, filaments, and point-like objects present in the tomogram under study. This process is known as segmentation.

In cryoET, tomographic segmentation is a difficult problem because of the low SNR and the missing wedge. The 3D nature of the data complicates the process further. Commonly used segmentation approaches in medical image processing cannot be applied directly for segmentation in ET because of the SNR issue. Such approaches include the well-known energy minimization based techniques such as active shapes and active contours [12] [13]. The currently available methods for ET segmentation include a variation of the level set method [12] [13] and the immersion-based watershed algorithm [12] [13], which are region-based approaches for segmentation.

In spite of the wide success of these methods, there are some serious limitations. The major limitation of the energy based approaches stems from the fact that the objective functions being optimized are highly non-convex, and hence the segmentation algorithms tend to get stuck at local optima [14]. Moreover, any of the currently available region based segmentation methods involve a significant amount of human intervention at different stages throughout the segmentation process [15].

1.3.1 Segmentation in Electron Tomography

However, the challenges in cryoET data interpretation can be alleviated using an effective segmentation approach. Presence of high amount of white noise suppresses the signal in the tomographic data [16]. Segmentation algorithms should be used in conjunction with a de-noising algorithm such as median filtering [12]. The poor contrast in the cellular tomograms along the membranes and filaments causes a problem for the segmentation algorithms that focus on local intensity properties in images. A common problem when segmenting membranes is the presence of well-defined point-like objects such as ribosome in the tomogram. These have a

strong uniform contrast. Segmentation methods for membrane extraction should use some feature for distinguishing such point-like structures and line-like objects. In addition to all such problems, the low contrast that may result from the missing wedge, can be treated using a proper segmentation algorithm. The goal of my research is to develop a semi-automatic 3D segmentation mechanism that will alleviate the problem of manual segmentation as well as the problem caused by the local optima in some energy based approaches. In general, if the energy function is non-convex, then there will be multiple local minima and any gradient-based algorithm for minimizing this function may get stuck at one of these local minima [17] [12]. Our segmentation approach does not face this problem of multiple local optima using the idea of classification. As mentioned before, high noise and low resolution are two inherent problems in ET. In order to increase the SNR and hence to improve the resolution, the segmented motifs must be well aligned and classified into self-similar groups before averaged. This is an iterative process. In each iteration, the resolution of the class averages and of course the global average improves. In this work, we utilize the idea of "segmentation by classification [18] and have applied the idea to the problem of segmenting HIV/SIV envelope spikes from the ET. SIV/HIV envelope (Env) spikes are the main structure facilitating entry of the virion into the host cell and have been the subject of intense research activity [19] [20] [9]. The spikes are to a first approximation randomly distributed across the virus envelope with some tendency to cluster [9]. Thus, tomographic studies tend to identify envelope spikes manually, a tedious process at best, and one in which human error may bias the selection.

An automatic spike selection method would greatly accelerate research in this area. This problem of automated segmentation is inherently difficult because of all the challenges. Moreover, the molecular structures of the spikes under study are structurally highly heterogeneous [9]. In this

work, we have generated a set of uniformly distributed points that cover the entire surface of the virion at about the radial position of the spike heads. Subvolumes were cut from the tomogram of the virion at the automatically generated positions, aligned translationally, and subjected to Multivariate Data Analysis and classification. Subvolumes that contain spikes near their center were identified using multivariate data analysis.

In the initial cycle of the segmentation process, without any alignment, many class averages lack any spikes because of the poor alignment. After an initial alignment using just translation, against a simplified reference, class averages showing spikes began to appear. After multiple cycles (precisely 8-10 cycles) of a procedure called “alignment by classification” [20], in which only class averages are aligned, spike definition improved and class averages showing only membrane became better defined. Finally, the original set of uniformly distributed data points shows concentration at spike coordinates. The clusters showing pure membranes are separated, and the clusters showing spikes in the average are segmented out for the further study of structure analysis. In essence, our contribution is a semi-automatic approach for 3D segmentation using classification. This method has the possibility of being applied not only for segmenting viral Env spikes, but also segmenting out specific structures from within a cell, for example ribosomes, which is hard to achieve using a manual segmentation approach.

CHAPTER 2

PROCESSING PIPELINES OF ELECTRON TOMOGRAPHY AND SINGLE PARTICLE ELECTRON MICROSCOPY

2.1 Electron Tomography

CcryoET is a technique for revealing the molecular architecture of complex macromolecular structures like viruses, proteins, organelles and cells at a very high resolution of a few nanometers. In cyoET, biological samples are imaged with an electron microscope (EM) and a series of projection images of the 3D specimen are collected at different angles. Before imaging, the biological samples are prepared specially to withstand the conditions inside the EM. Next, the projection images are aligned and combined to yield the 3D reconstruction of the specimen. The 3D image produced after alignment and reconstruction, is known as a tomogram. Many computational steps are involved afterwards to achieve the successful interpretation of the tomogram as well as the structure under study.

In EM, electrons are produced from an electron gun, placed at the top of the microscope and are accelerated by an electrical potential of 100-300 keV. The electron beam travels from gun, through the specimen and to the recording medium at the bottom through a column (the optical path) held at a very high vacuum to minimize the scattering of electrons by residual air. At different positions along the microscope column, electromagnetic lenses are placed to deflect the path of the electrons, which brings them to focus at different positions along the optical path. From the top, the first lens system is called the condenser lens system. This lens system has a set of deflectors, two condenser lenses, called C1 and C2, which control the illumination of the sample, stigmators and an aperture. The purpose of the condenser lens is to take the electron

coming out of the gun and focus and direct them onto the sample. The sample resides within the next lens system, called the objective lens system. The objective lens system has a pair of deflectors, the lens, stigmators and an aperture. The objective lens system produces the magnified image of the sample. This lens is the most important, because it combines the scattered and non-scattered electrons that give the contrast in the image. The magnified image is further magnified by the third lens system, which is called the projector lens system. This lens system has deflectors and several intermediate lenses and at the end a projector lens and stigmators and an aperture. The intermediate lenses are utilized to change the magnification from 50 times to as high as 400,000 times. Finally, the magnified image is sent through a final pair of deflectors on to detectors. Images are then observed in real time on a florescent screen or recorded on a scintillator based CCD camera and viewed in real time on a computer screen and are recorded for further data processing either on a photographic film or a high quality digital camera which can be either a CCD camera coupled to a scintillator or more recently to a direct electron detector (DED) where electrons themselves are detected rather than photons generated by an intervening scintillator.

2.1.1 Electron-sample Interaction and Image Formation

In Transmission Electron Microscopy (TEM), electrons have very high energy, about 100-300 keV, accelerating them to relativistic velocities. While interacting with the sample, electrons undergo a change in direction either because of collisions with atomic nuclei of the specimen or because of electrostatic interactions with the electrons in the electron shells surrounding the nuclei. The following four situations can happen when the electron beam interacts with the sample:

- **Elastic scattering**: Elastic scattering originates when an electron passing closer to the nucleus is more strongly attracted by the positive charge and is therefore deflected through a larger angle (~10^{-2} radians), which is directly proportional to the atomic number of the specimen atom. In this scattering, electrons do not lose their energy, but only change their direction. Elastic scattering generates the high-resolution detail in the image.
- **In-elastic scattering:** Electrostatic interactions and collisions between the beam electrons and the electrons surrounding the atomic nucleus, give rise to inelastic scattering. The deflected electrons are likely to undergo a loss of energy and are deflected through very small angles (~10^{-4} radians), causing almost all of them to pass through the objective aperture. These in-elastically scattered electrons, only contribute in adding noise to the image and also damaging the sample in the process.
- **No interaction:** Beam of electrons, which pass outside the range of the electrostatic field of atomic nuclei and atomic electrons, are not scattered and just go through the molecule without interacting with the specimen. These electrons do not directly contribute to image formation. Rather they interfere with the elastically scattered electrons to produce contrast.
- **Absorption:** To be absorbed, an electron must lose all its energy to the specimen. The portion of the electron being absorbed depends on the overall thickness of the specimen. In TEM, for specimens of normal thickness (<100-200 nm), the portion of the beam absorbed in the specimen is negligible. However, if the electron loses enough energy it is likely to hit one of the fixed apertures further along the optical path, thereby removing it from the other electrons that have passed through the specimen.

The proportions of elastic and in-elastic collisions depend on the accelerating voltage and the nature of the specimen.

2.1.2 Contrast and Image Formation in EM

Contrast in the electron image can arise from both "amplitude" and "phase" effects.

1) **Amplitude contrast**: Amplitude contrast is produced by the loss of amplitude (i.e. electrons) from the beam. It is also called scattering contrast. We can remove the elastically scattered electrons by means of the aperture. The aperture is placed after the objective lens in its back focal plane. The objective aperture allows the un-scattered electrons and electrons scattered up to the radius of the objective aperture to go through but it blocks elastically scattered electrons scattered at high angle and that process generates contrast in the image. Electron opaque object points produce scattering through relatively large angles. Therefore, electrons incident on these points are excluded by the lens aperture and the intensity of images at these points becomes correspondingly low. Conversely, electron transparent regions in the object produces little scattering beyond the lens aperture. The intensity of images of these regions is correspondingly high. Amplitude contrast does not usually provide high-resolution information on proteins or macromolecular complexes because the aperture may block that information.
2) **Phase contrast:** Phase contrast originates from shifts in the relative phases of the portion of the beam that interacts with the sample. Phase contrast in the image arises from differences in phase between scattered and un-scattered rays in different parts of the image and interference between these rays. In this case the contrast transfer function of the objective lens of the microscope is utilized to make the elastically scattered and un-

scattered electrons interfere, generating contrast. Phase contrast provides the high-resolution information.

2.1.3 Resolution and Radiation Damage

A TEM can take images with atomic detail. State of the art TEM, can reach a resolution beyond a single angstrom (0.8 Å) [16]. A very high resolution is achievable if the sample is not radiation sensitive. Unfortunately, biological samples are extremely sensitive to radiation [16]. Biological samples have several other limitations when placed directly into a TEM. Most biological samples live in aqueous solution and as a result, cannot withstand the high vacuum of the microscope. Biological atoms are made of carbon, oxygen and nitrogen and all of them have almost same scattering power, causing very low intrinsic contrast in the sample. The quality of the cryo-tomographic reconstruction is highly correlated with the electron dose accumulated by those tilt series images used to compute the tomogram. As the biological samples are highly radiation sensitive, a total dose of 120 $e^-/Å^2$ is spread over the whole data set which could consist of more than 120 images. As a result, the dose available for single image is below 1 $e^-/Å^2$, which produces a high stochastic noise level in individual images. Therefore, a poor SNR (~0.01) is expected for cryo-tomograms [14].

2.1.4 Cryo Specimen

When in-elastic scattering occurs, the sample gets ionized, generating free radicals that move around the sample breaking the bonds in the molecular assemblies and as a consequence the sample becomes vulnerable enough to “explode” [21]. Because of these issues, biological specimens must be specially prepared prior to imaging. The specimen preparation technique, that ensures the optimal structural preservation relies on rapid freezing of samples, producing what

are called unstained, frozen-hydrated samples. Relatively thin samples (<500 nm), embedded in aqueous solution, are quickly plunged into liquid ethane or liquid propane at temperatures close to that of liquid nitrogen (-196 °C), so that the water molecules in the specimen do not have time to re-organize into a crystal, keeping the ice in vitrified state. Those samples can be examined directly in EM by maintaining the temperature below -170 °C [21]. In this case the sample is hydrated but in a "glassy" state, and can withstand the high vacuum inside the EM.

2.2 Electron Tomography Workflow

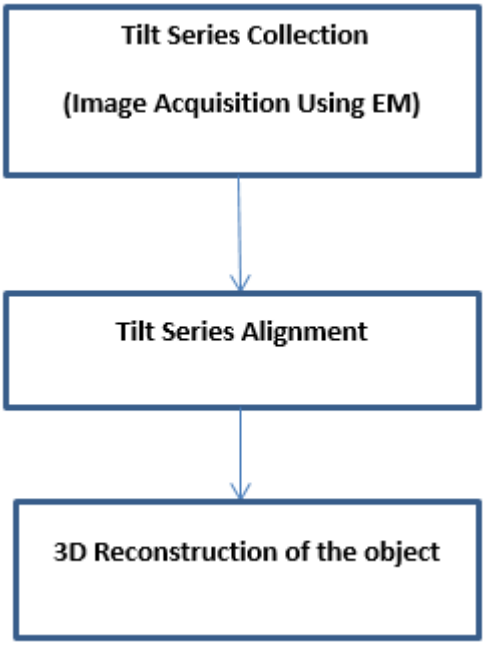

Figure 2.1: Tomographic work flow. Tomographic work flow consists of three major steps – the collection of the projection images from different angles or the tilt series generation, aligning the projection images or the tilt series alignment and reconstruction of the original 3D specimen from the aligned 2D projection images.

Electron tomographic workflow consists of three major steps- 1) tilt series acquisition or collecting 2D projection images of the 3D sample at different viewing directions, 2) Tilt series alignment or registering the 2D projection images to a common coordinate system, 3)

Reconstruction of the original 3D sample from the aligned 2D projection images. Several other steps of the workflow include noise reduction, segmentation, detection and mapping of macromolecular assemblies, 3D subvolume alignment and clustering the subvolumes to reveal the heterogeneous macromolecular structures under study, averaging and validation of the result [14]. Each stage of the workflow involves significant amounts of computation and in reality, most of the steps are computationally challenging because of the low SNR and limited resolution of the tomogram. Each of the major steps will be described in detail below.

2.2.1 Automated Image Acquisition

In essence, ET is the method of 3D reconstruction of a specimen from a series of projection images. In ET, a sample is introduced in the electron microscope and a series of projection images, called the tilt-series, are recorded in a digital camera (Figure 2.2A), by tilting the sample in different angles around a single fixed axis perpendicular to the electron beam. Theoretically, though the angular range for the rotation is ±90°, but in practice, because of some technical limitations of microscope and specimen, a typical acquisition session generates a tilt series of the whole angular range of ±60° or at the most ±70° (Figure 2.3A) at angular increments of 1° - 5°. The image collection is computer automated and the recorded image size is typically 2048 x 2048, 4096 x 4096 or even 8192 x 8192 pixels. Leginon [22], SerialEM [23], TOM [24], and UCSF Tomography [25] are the widely used tilt-series collection softwares

The data acquisition follows a 3D geometry, called tilt geometry (Fig. 2.2). The image coordinate system **(x,y,z)** is fixed with respect to the microscope, with **z** as the optical axis for all the images in the tilt series and **x,y** determined by the recording medium. The electron beam is

2.2.2 Tilt Geometry

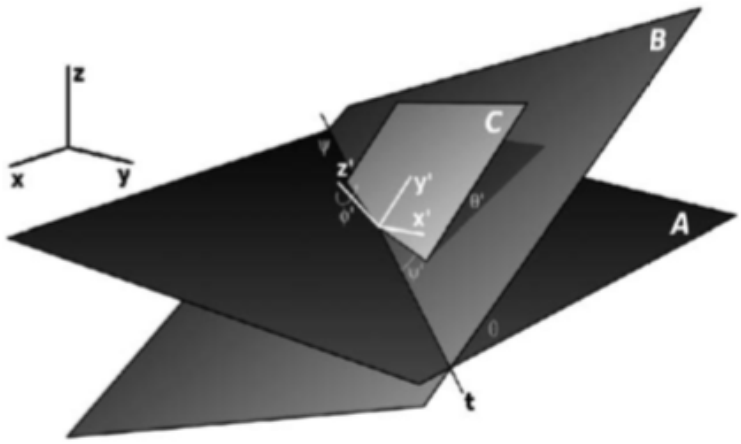

Figure 2.2: Tilt geometry. (x,y,z): coordinate system fixed with respect to the microscope, z is the optical axis, A (the x-y plane) is the image plane. (x',y',z'): coordinate system fixed with respect to the specimen. The transformation from (x,y,z) to (x',y',z') consists of a tilt about the axis t and angle θ, and an additional rotation ($\psi^{'},\theta^{'},\phi^{'}$) which defines the orientation of the specimen (C) with respect to the specimen holder (B). [26]

parallel to the **z** axis, interacts with the sample and after traversing a series of lenses the digitized image is recorded in the **x-y** plane defined as **A** (Fig. 2.2). The tilt axis **t** is assumed to be perpendicular to the optical axis **z**. Hence, the direction of the tilt axis **t** is defined as the azimuthal angle **ψ** with respect to the **x**-axis. The tilt axis for each member in the tilt series is aligned with the tilt axis of the reference image by an in-plane rotation about the optical axis **z**. The specimen holder or the EM grid is represented by plane **B** and is related to the image plane **A** by a rotation **θ** about the tilt axis **t**. The orientation of the specimen with respect to the specimen holder is captured by an additional rotation in the 3D space. Hence, plane **C (x'-y'** plane**)** is related to plane **B** with three Euler angles **($\psi^{'},\theta^{'},\phi^{'}$).** These Euler angles define the departure of the axis of the specimen plane normal from the **z**-axis. In essence, for a single axis tilt series with **n** projection images, the required parameters are one tilt azimuth angle **ψ**, **n** in-

plane rotations about the optical axis **z**, **n** tilt angles $\theta_1,\ldots\ldots,\theta_n$ about **t**, and three Euler angles (ψ',θ',ϕ') for the additional 3D rotation from plane **B** to plane **C** [26].

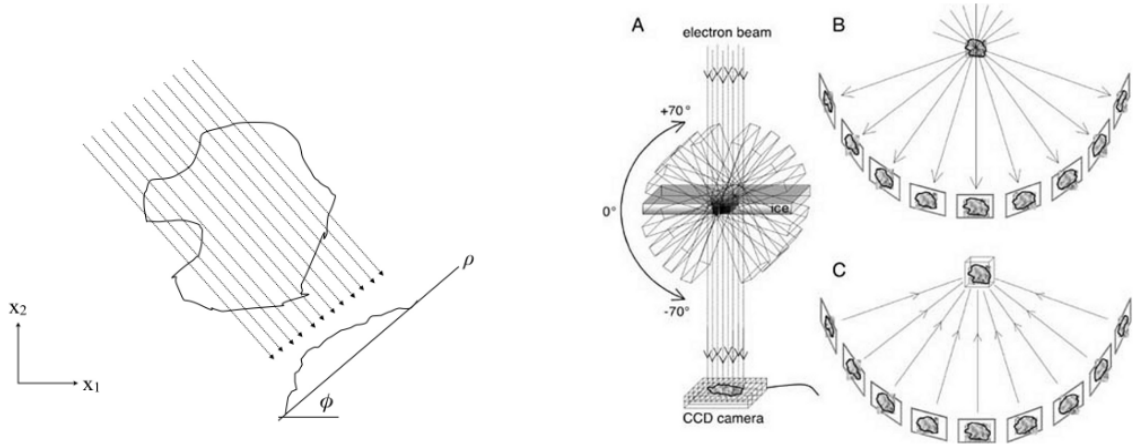

Figure 2.3: Processing pipeline of tomography. Left: Radon transform of a 2D function taken at the projection angle of ∅. Right: A) Transmission electron microscope sample holder rotations. B) Projection of a 3D function, C) Back-projection of 2D Radon transform of a 3D function [124], [125]

2.2.3 Tilt Series Alignment

The term "tilt series alignment" means correction of relative shifts and rotations between the projection images of the tilt series i.e. registering the projection images in the tilt series (Figure 2.3B). The projections are representing the same object from different angles. Hence the projections are similar but not identical having some amount of foreshortening. Moreover, during data acquisition, the imperfection of the mechanical tilt system and the electron optics produce shifts, rotations, magnification changes and other distortion in the image [27]. A large portion of these distortions is compensated during the automated data collection procedure but a more accurate alignment of the tilt series is needed for further processing.

The alignment procedure addresses two different questions: the determination of the direction of the tilt axis and the determination of the x-y positions of the projections relative to a common origin. In other word, the goal of the alignment is to mutually set the images to a common coordinate system correcting the shifts, rotations and other distortions.

Two most widely used alignment techniques are:

1) Least-squares method of alignment using fiducial markers

2) Alignment by cross-correlation.

Marker Based Alignment Technique: In the marker based alignment technique, colloidal gold particles are added to the biological samples and are used as electron-dense fiducial markers. Because of their high contrast, the markers can be easily tracked in the images. The determination of the tilt axis should be done from two or more micrographs, separated by a large tilt angle. Enlargement of these micrographs are made and at least two points are selected which are common to all the micrographs and are as far apart from each other as possible. These points are used for a triangulation [28] and a least square fitting procedure can be used to improve the accuracy. The markers need not be identified in each member of the tilt series if the fitting algorithm can deal with incomplete sets of markers. After the tilt axis has been determined from the two selected micrographs, it can be transferred to the other images of the series, by identifying a line parallel to this axis joining two easily identifiable features. A variant of the marker-based alignment uses specimen features as markers in cases where artificial markers are undesirable. One drawback of the colloidal particles is that they create artifacts that occlude the biological features nearby in the 3D reconstruction due to their extremely high contrast.

Although the positions of the fiducials give information on the origin of the images, it is normally not sufficient for determination of the common origin of a projection series. The problem is that, the origin is relative to the fiducials, which may not be identical to the origin of the specimen itself. Fiducials may move during data acquisition. It is necessary to use information intrinsic to the specimen for alignment, with means of cross correlation function.

Cross Correlation Based Alignment Technique: The cross-correlation techniques are widely used for pattern matching in the electron micrographs to locate common features in the presence of high noise. In the case of electron tomography, projected images represent different views of the 3D structure under study. Any two adjacent projection images are similar but not identical, having different amount of foreshortening and hence cannot be cross correlated directly. To compare two images at different tilt angles, the collected images, which are the orthogonal projections, are stretched in the direction perpendicular to the tilt axis by a factor of 1/cosine (tilt angle), to generate inclined projections [29]**.** Cross correlation based image alignment process involves the following steps in general:

1) Take two projections p_1 and p_2 from the tilt series
2) Stretch the projection images along the direction orthogonal to the tilt axis by the stretching factor 1/cos (tilt angle)
3) Compute Fourier transform of the stretched input projections

   $P_1=F\{p_1\}$

   $P_2=F\{p_2\}$
4) The Fourier transforms of the input images are now high-pass and low-pass filtered.

5) Compute the cross-correlation function (CCF) of the stretched and filtered input image relative to the stretched reference image by multiplying P_1 and complex conjugate of P_2 and then compute inverse transform of the product i.e.

   $CCF=F^{-1}\{F\{p_1\} * F^{*-1}\{p_2\}\}$

6) Compute the CCF peak search iteratively and store the x and y coordinates of the highest cross correlation peak.
7) Calculate the desired shift by un-stretching the x and y coordinates (i.e. by multiplying the y coordinate by cos (tilt angle) assuming x-axis as the tilt axis)
8) The input image is shifted in its frame by the negative of the shift vector.
9) The process continues until all the images in the tilt series are aligned.

Alignment by Cross-correlation in PROTOMO: In PROTOMO, the CCF based alignment is performed in two steps - an initial coarse alignment is followed by area matching using cross-correlation [26]. Area matching is an iterative refinement process in which each iteration includes a refinement of the geometric parameters.

The CCFs are computed according to the algorithm described above. The alignment is based on the fact that, the particle projections separated by a small angular increment are similar to one another. For any two aligned projections, the similarity can be expressed by the size of their correlation coefficient and the factors that contribute to how well the two images will correlate, are specimen thickness and tilt increment.

In the initial coarse alignment, a reference image (usually the 0° tilt image) is selected to define the reference coordinate system and a rectangular region of interest is chosen for area matching. The goal is to find out the identical area in the other images in the tilt series, and to estimate the

required shifts and rotations to achieve the best match. The coordinates of other tilted images in the series are adjusted to fit with this reference coordinate system using translational and rotational shifts. Each image is aligned to its preceding image in the tilt series using cross correlation. The low tilt image is always used as the reference. The alignment is carried out from 0° to the maximum negative tilt ($-\vartheta_{max}$) and then from 0° to the maximum positive tilt (ϑ_{max}).

More precisely, first the selected area in the reference image is padded to the same size of the reference image. The Fourier transform of this new image is computed. Next, the Fourier transform of the image to be aligned is computed. Importantly, the image to be aligned must be stretched and rotated in the proper direction before computing the transform. The complex conjugate of the second transform is computed and the transform of the reference image is multiplied with it. Thereby, the cross correlation is computed in the Fourier domain. Next, an inverse Fourier transform is applied to the computed product, which gives the cross correlation in the real domain. A grid search algorithm is performed to find the highest correlation peak. The shifts are calculated accordingly.

The next step is area matching using cross correlation, which can simply be described as an iterative refinement process. In this step, PROTOMO incorporates a more general approach by introducing a 2D linear transformation matrix, called the distortion matrix to resample the images to the reference coordinate system. This is analogous to the cosine-stretching for a simple conventional cross correlation alignment. The 2D affine transformation for matching the equivalent image areas of two projection images are captured by six parameters, a 2x2 transformation matrix and two origin coordinates. The best match is determined by maximizing the cross-correlation coefficient. A simple grid search algorithm can be used to determine the highest cross correlation peak but it turns out to be computationally very expensive. Hence, a

nonlinear optimization algorithm, namely modified Nelder-Mead algorithm [30] is used to find out the highest CCF peak. The nonlinear peak search algorithm outputs both, the modified transformation matrix and the coordinate of the highest CCF peak, which are the desired shifts.

In PROTOMO, the area-matching step does not use a single image from the tilt series as reference. Instead, the reference image is constructed from the already aligned images. After an image is area matched with the reference image, the aligned images are back projected into a volume using weighted back projection method [31] and then re-projected in the direction of the next image to be aligned. The re-projected image is now used as the new reference image for the next alignment. This process continues until all the images in the tilt series are area matched. Finally, the geometric parameters are simultaneously refined during the iterative process.

2.2.4 Tomographic Reconstruction

The last major step in the Tomography processing pipeline is the 3D reconstruction of the specimen from the set of aligned 2D projection images of the tilt series (Figure 2.3C). The mathematical principle of the tomographic reconstruction is based upon the central slice theorem or the projection theorem which states that ***the Fourier Transform of a 2D projection of a 3D object is a central section of the 3D Fourier Transform of the object***. Therefore, theoretically the 3D Fourier Transform of the specimen can be computed by combining the 2D Fourier transforms from the tilt series and the 3D structure of the specimen can be obtained by an inverse Fourier Transform. This approach is not useful in practice because of the fact that the problem of this approach is related to the non-trivial interpolation of Fourier space.

Traditionally, 3D reconstruction methods have been classified intro two major groups, **Fourier Reconstruction methods and Real-space methods**. In Fourier reconstruction methods, the 3D

Fourier Transform of the specimen is reconstructed from the experimental sample points and the real space distribution of the object is obtained by inverse Fourier transformation. In contrast, direct method demands all calculations in real space. Real-space methods include the convolution back-projection algorithm and iterative algorithms such as Algebraic Reconstruction Technique (ART) [32], Simultaneous Iterative Reconstruction Technique (SIRT) [33], and Iterative Least-squares Technique (ILST) [34].

The standard method for tomographic reconstruction is Weighted back projection (WBP), which can be described as the Fourier approach but working in real space [31].

Back Projection Quantity: The back-projection method assumes that the projection images represent the amount of mass density encountered by imaging rays. In case of single axis tomography, the reconstruction of the 3D volume from 2D projections can be thought of as a series of N_x independent reconstructions of 2D slices from N_y equivalent strips of the projections.

Here each strip is narrow enough to be represented by a one-dimensional set of measurements, and the resulting slice is thought of as a single layer of voxels i.e. without density changes along the direction of slicing. Each micrograph can be thought of as being composed of strips perpendicular to the tilt axis. Each strip is the projection of a slice. Hence the 3D reconstruction problem is reduced to the problem of 2D reconstruction of the density distribution over a slice from a set of 1D strips containing the measured projected densities.

The experimental measurements $p_{i,\vartheta}$ on a given projection under an angle ϑ, is interpreted in terms of summation of voxels θ_j , of the object lying within the projection "ray" $r_{i,\vartheta}$. This ray is a narrow strip whose width corresponds to the size of the projection pixels or a multiple of this. The projection equation, $p_{i,\vartheta} = \sum_{j \epsilon\, r_{i,\vartheta}} \theta_j$ i.e. sum over all voxels lying within the ray, must hold

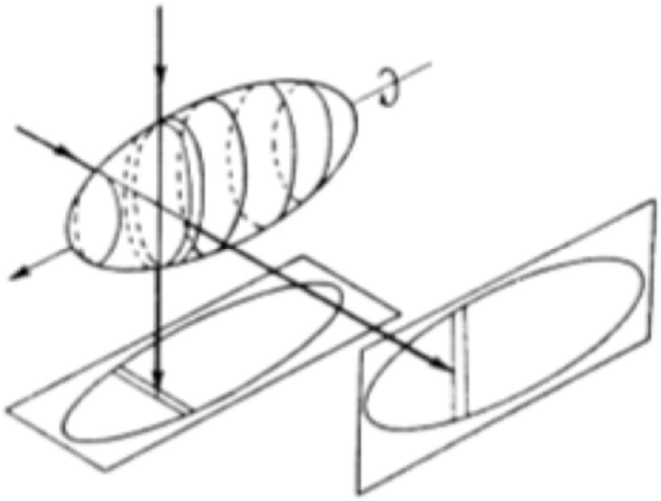

Figure 2.4: Projection geometry relating to a single-axis tilting experiment

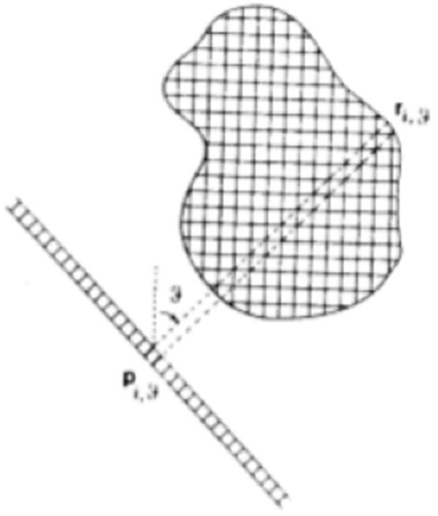

Figure 2.5: Relationship between experimental projection (relating to projection angle θ) and the slice of the reconstructed object.

for all voxels of the projection strip, and for all of the tilt angles ϑ used in the experiment.

The real-space methods are based on the concept of back projection: in essence, a value $p'_{i,\vartheta}$ associated with the projection pixel $p_{i,\vartheta}$ is "smeared out" along the corresponding ray so that each of the voxels falling within the i-th ray receives an equal share. For the iterative techniques, the back-projection quantity is

$$p'_{i,\vartheta} = p_{i,\vartheta} - \sum_{j \epsilon r_{i,\vartheta}} \theta_j \qquad (1)$$

i.e. the difference between the experimental measurement $p_{i,\vartheta}$ and the current ray sum. In each iteration the correction amount i.e. $p'_{i,\vartheta}$ decreases and a distribution of voxels approximating the original object is seen to emerge. For different iterative methods, the remaining error and the way in which the corrections are applied to the 3D density distribution are different.

The convolution back-projection method has higher computational efficiency where the back-projection quantity is

$$p'_{i,\vartheta} = p_{i,\vartheta} \, o \, f_i \qquad (2)$$

where, o is the convolution operation and f_i is the modifying function that has the property of enhancing high-resolution features in the projection. The steps in convolution back-projection method are:

- Convolve each projection with ramp filter
- Compute back-projections
- Compute summation of the back-projection bodies

Simple Back-projection Method: A simple reconstruction method is the simple back-projection method or summation technique shown in Fig: 2.6. In this technique, different projections of an object are smeared out or back projected to form 'back-projection bodies' onto a volume called the back-projection volume. Consequently, back-projection rays from different projection images intersect and reinforce each other at the point where mass is found in the original structure. In other words, to reconstruct the object, all back-projection bodies are summed. Obviously, the reconstruction is better when more projections are used.

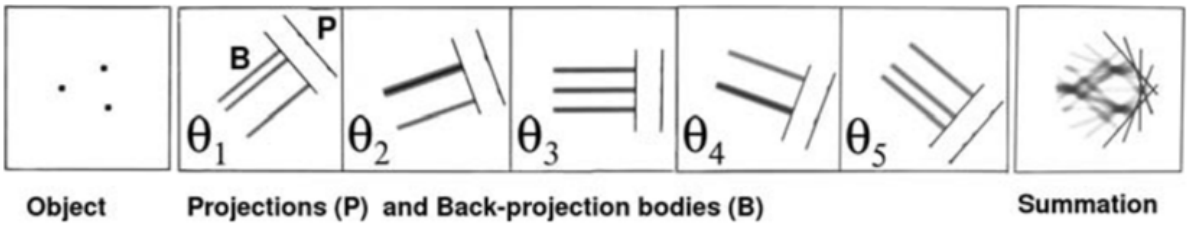

Figure 2.6: Principle of simple back projection method. Back projection bodies are created from five projection P at angles ϑ1, ϑ2, ϑ3, ϑ4 and ϑ5 and the object is reconstructed by addition these back projection bodies [31] .

In practice, the object reconstructed using this technique, is strongly blurred. It can be shown that a simple back projection technique reconstructs the object with a point spread function that overweights the low spatial frequency components.

Drawback of Simple Back-projection: The drawback of the simple back-projection method is the incorrect weighting of the data, i.e. the low spatial frequency domain receives an unduly large share of the sample points. As a result, the back-projection involves an implicit low pass filtering that makes the reconstructed object strongly blurred.

From the given picture(Fig.2.7), it is understandable that Fourier components are concentrated around the center i.e. low frequency components have a higher concentration and the high frequency components are sparsely populated. The only possible way to get the high frequency components, is interpolation of the experimentally sampled Fourier transform onto a Cartesian grid and subsequent inverse Fourier transformation. Unfortunately, the Fourier domain interpolation is non-trivial, since according to the Whittaker-Shannon sampling theorem [35], each of the samples in Fourier space contributes to every point of the Fourier grid [36]**.** Hence a simple bilinear interpolation will not be sufficient for this purpose.

Significantly, a weighting function can compensate this imbalance. The weighting function acts like a high pass filter in the sense that weights down the low frequency Fourier terms to bring them back into balance with the high frequency terms. In PROTOMO, the weighting function is applied to the Fourier transform of the projection images themselves rather than the back-projection body, otherwise referred to as the tomogram. The reason for this comes from the fact that during alignment, references are computed from a back projection of previously aligned images and by weighting the images ahead of time, speeds up the calculation. This reweighing is necessary to restore the correct balance in the reconstruction across all spatial frequencies.

R-weighted Back-projection: To discuss about the R-weighted back-projection method, the concept of point-speared function and transfer function is needed. The point-spread function of an imaging system describes the image of a single point as it results after using a perfect point as the input to the system. If the system is shift-invariant, a property also called isoplanatic in optics, then the system response is independent of the absolute coordinates and depends only on the difference vector ($x - \xi$, $y - \eta$). The point-spread function then can be written as**,**

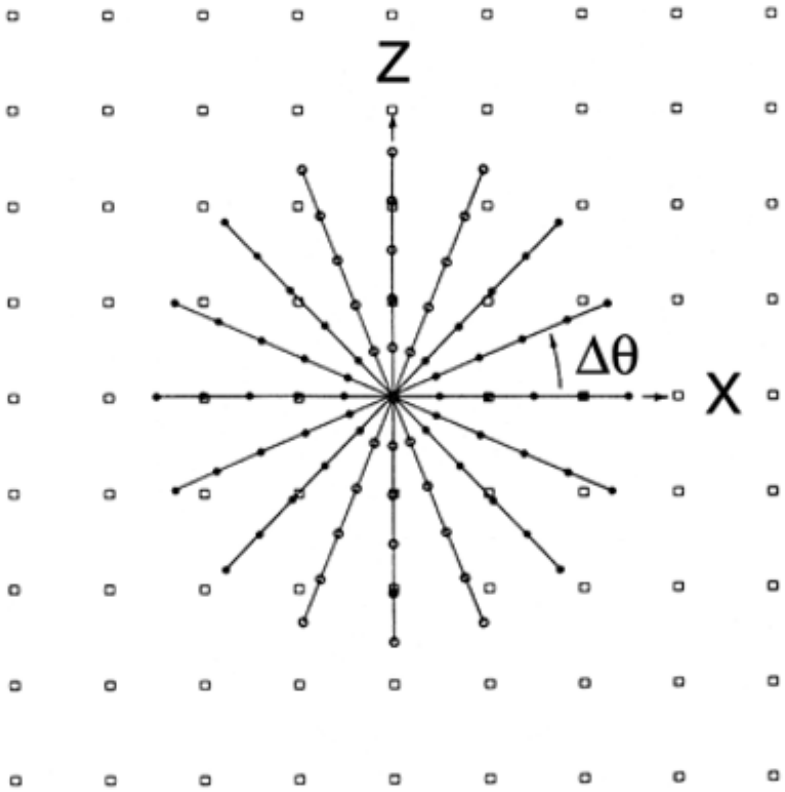

Figure 2.7: Relationship between the polar and Cartesian grid on which Fourier samples are obtained with the Cartesian grid on which data are required for Fourier inversion [3].

h(x, y; ξ , η) = h(x − ξ , y − η) [31]**.**

Let **f** (**x**,**y**,**z**) be a 3D distribution, which is projected under the angles θ_j , φ_j to form a series of projections $p_j\,(x_j, y_j)$. Let $r_j = (x_j, y_j, z_j)$ be the coordinates in the coordinate system of the projection p_j which forms the (x_j, y_j). plane. The geometrical relationship between the object coordinates r = (**x**,**y**,**z**) and $r_j = (x_j, y_j, z_j)$ can be described using the rotation matrices $D_{\theta j}$, $D_{\varphi j}$.

$$r_j = D_{\theta j} \cdot D_{\varphi j} \cdot \mathrm{r} \qquad (3)$$

Now, the rotation matrix about the y-axis by angle θ_j is given by

$$D_{\theta j} = \begin{matrix} cos\theta_j & 0 & -sin\theta_j \\ 0 & 1 & 0 \\ sin\theta_j & 0 & cos\theta_j \end{matrix}$$

And the rottaion matrix about the $z^{'}$ axis by φ_j is

$$D_{\varphi j} = \begin{matrix} cos\varphi_j & sin\varphi_j & 0 \\ -sin\varphi_j & cos\varphi_j & 0 \\ 0 & 0 & 1 \end{matrix}$$

A projection along the direction z_j with angles θ_j , φ_j can be written as

$$p^j = \int x^j \, y^j z^j \qquad (4)$$

Now a back-projection body is formed by convolution of p_j in the (x^j, y^j) plane with the 3D point spread function

$$l_j = \delta(x^j, y^j) \, . \, c\left(z^j\right) \qquad (5)$$

$$\text{with} \quad c\left(z^j\right) = \begin{cases} 1 & for - a \le z^j \le a \\ 0 & otherwise \end{cases} \qquad (6)$$

The convolution conditions for an isoplanatic system are fulfilled within a sphere of diameter $\boldsymbol{a}$ if $\boldsymbol{2 \cdot a}$ is at least twice the object diameter and the projections and the reconstruction volume are sufficiently large to include all back-projection rays.

$$\boldsymbol{p_j^b}\left(x^j, y^j, z^j\right) = \iint \boldsymbol{p_j}\left(x^j, y^j\right). \boldsymbol{l}(\boldsymbol{x^j} - \boldsymbol{x}^{'\,j}, \boldsymbol{y^j} - \boldsymbol{y}^{'\,j}, z^j)\mathrm{d}\boldsymbol{x}^{'\,j}\mathrm{d}\boldsymbol{y}^{'\,j} \qquad (7)$$

$$= \iiint \boldsymbol{f}\,(\boldsymbol{x}^{'\,j}, \boldsymbol{y}^{'\,j}, z^{'\,j})\,\mathrm{d}\boldsymbol{z}^{'\,j}\,\boldsymbol{l}(\boldsymbol{x^j} - \boldsymbol{x}^{'\,j}, \boldsymbol{y^j} - \boldsymbol{y}^{'\,j}, z^j)\,\mathrm{d}\boldsymbol{x}^{'\,j}\mathrm{d}\boldsymbol{y}^{'\,j} \qquad (8)$$

where the $\boldsymbol{j}$'s are the orientation angles.

The back-projection algorithm becomes

$$b(x,y,z) = \sum_j \boldsymbol{p_j^b}\left(x^j, y^j, z^j\right) \quad (9)$$

The point-spread function is found by analyzing how the back-projection algorithm affects a single point in 3D represented by the function

$$q = \delta\,(x, y, z). \quad (10)$$

To find the weighting function for a weighted back-projection for arbitrary geometry, we must first analyze the point-spread function of a simple back-projection in more detail.

The projection of $\boldsymbol{q}$ at angles θ_j , φ_j is

$$\boldsymbol{p_j}\left(x^j, y^j\right) = \delta(x^j, y^j). \quad (11)$$

The back-projection body, using equation (5) and (8), becomes

$$\boldsymbol{p_j^b}\left(x^j, y^j, z^j\right) = \delta(x^j, y^j)\,.\,c\left(z^j\right) \quad (12)$$

and the point back-projected in 3D is found by summation over θ_j , φ_j as

$$b(x,y,z) = \sum_j \delta(x^j, y^j)c(z^j) \quad (13)$$

Thus, b(x,y,z) is the point-spread function of a back-projection calculated from a set of projections taken with arbitrary angles θ_j , φ_j.

The transfer function is the Fourier transform of the point spread function b(x,y,z) and is defined as,

$$\begin{aligned} H(X,Y,Z) &= F[b(x,y,z)] \\ &= F\{\textstyle\sum_j \delta(x^j, y^j)c(z^j)\} \\ &= \textstyle\sum_j F[\delta(x^j, y^j)c(z^j)] \end{aligned} \quad (14)$$

and

$$\begin{aligned} F[\delta\left(x^j, y^j\right)c\left(z^j\right)] &= \textstyle\int_{-\infty}^{+\infty} \iint \delta(x^j, y^j)c(z^j)e^{-2\pi i(x^jX^j + y^jY^j + z^jZ^j)})\, dx^j dy^j dz^j \\ &= 1 \textstyle\int_{-a}^{+a} e^{-2\pi i z^j Z^j} dz^j \end{aligned}$$

$$=\frac{\sin(2\pi a Z^j)}{\pi Z^j}$$

$$=2a\ sinc(2a\pi Z^j) \quad (15)$$

where sinc(**x**) = sin(**x**)/**x**

Given the rotation matrices, Z^j can be expressed in the coordinate system (X,Y,Z) of the object:

$$Z^j = \mathbf{X}\sin\theta_j \cos\varphi_j + \mathbf{Y}\sin\theta_j sin\varphi_j + \mathbf{Z}\cos\theta_j \quad (16)$$

and the transfer function becomes

$$H(X,Y,Z) = \sum_j 2asinc[2a\pi(Xsin\theta_j cos\varphi_j + Ysin\theta_j sin\varphi_j + Zcos\theta_j] \quad (17)$$

Hence, the corresponding weighting function for arbitrary geometry is

$$W_a(X,Y,Z) = \frac{1}{H(X,Y,Z)} \quad (18)$$

$$=\{\sum_j 2asinc[2a\pi(Xsin\theta_j cos\varphi_j + Ysin\theta_j sin\varphi_j + Zcos\theta_j]\}^{-1} \quad (19)$$

Equation (19) is only valid for H≠ 0. The original 3D distribution $o(x,y,z)$ can be recovered from the back-projection $b(x,y,z)$ by multiplying of its Fourier transform $B(X,Y,Z)$ by the weighting function $W_a(X,Y,Z)$, followed by an inverse Fourier transform:

$$o(x,y,z) = F^{-1}[O(X,Y,Z)] = F^{-1}[B(X,Y,Z)W_a(X,Y,Z)]$$

Rebalancing the Fourier coefficients is done with a weighting function, W(**k**), where **k** is the spatial frequency

$$W_a(X,Y,Z) = 1/H(X,Y,Z) \quad (20)$$

$$W_a(X,Y,Z) = \{\sum 2asinc[2\pi a(Xsin\theta_j\ cos\varphi_j + Ysin\theta_j sin\varphi_j + Zcos\theta_j]\}^{-1} \quad (21)$$

This is best done in cylindrical-polar coordinates where Γ is the angular coordinate, θ the tilt increment and Y is the tilt axis

$$H_a(R, \Gamma, Z) = 1/R \quad (22)$$

$$W(R, \Gamma, Z) = \mathrm{R} \quad (23)$$

A weighted back-projection method is a simple back-projection method followed by a deconvolution with the point spread function of the simple back projection algorithm. This deconvolution is done by a division of the Fourier transform of the back-projection by its transfer function, which is essentially the Fourier transform of the points spread function. For single-axis tilting the weighting function is proportional to the radius (R) in Fourier space perpendicular to the tilt axis, hence the term R-weighted back-projection. The multiplication by R can be applied either to the projections or directly to the 3D Fourier transform of the back-projection.

The WBP is widely used in ET mainly because of its computational simplicity. The disadvantage of this method is its sensitivity to the inherent challenges of ET i.e. limited tilt angle, low SNR and poor contrast.

2.2.5 The Missing Wedge

The main challenge in ET is that the reconstruction of an object from a set of its projections. To compute the reconstruction, the Radon transform of the 3D function is measured. Radon transform is a set of parallel line integrals of a density function taken at various projection angles in 2D. In what follows we explain the phenomenon of missing wedge in the context of a 2D function and its 1D radon transform. Whatever we state here can be generalized to 3D functions.

Suppose that we are given an object O and we are taking the projection of the object along a given direction (ρ_j,θ_k). When θ_k is fixed we get one projection denoted by g(ρ, θ_k), by varying ρ over the different 'lines' that cut through the object 'O', in other words by taking the line integral along ρ_j (Fig. 2.8).

Given θ_k , the equation of a line is given by $x\cos\theta_k + ysin\theta_k$=$\rho$. Thus the line along ρ_j is given by

$$x\cos\theta_k + ysin\theta_k = \rho_j \quad (24)$$

Thus, the value of the projection of object 'O' along ρ_j is simply the line integral of 'O' along ρ_j whose value is given by equation (24). This is called "**Radon transform**".

For a 2D object, the object 'O' is denoted by a 2D function $f(x,y)$ and the Radon transform of $f(x,y)$ is given by,

$$g(\rho,\theta_k) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(xcos\theta_k + ysin\theta_k - \rho)\partial x\partial y \quad (25)$$

where δ is the Dirac Delta function.

We note that the Dirac Delta is zero at all places except the origin. Thus, in equation (25), the Dirac Delta enforces the fact that the sum of the densities is taken along lines, $x\cos\theta_k + ysin\theta_k = \rho_j$, for different x and y values, thus giving rise to the line integral In discrete space, we get the Radon transform as,

$$g(\rho,\theta_k) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\delta(xcos\theta_k + ysin\theta_k - \rho) \quad (26)$$

We also note that,

$$g(\rho,\theta+180^\circ) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(xcos(\theta+180^\circ) + ysin(\theta+180^\circ)dxdy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(-xcos\theta - ysin\theta - \rho)dxdy$$

$$= g(-\rho,\theta) \quad (27)$$

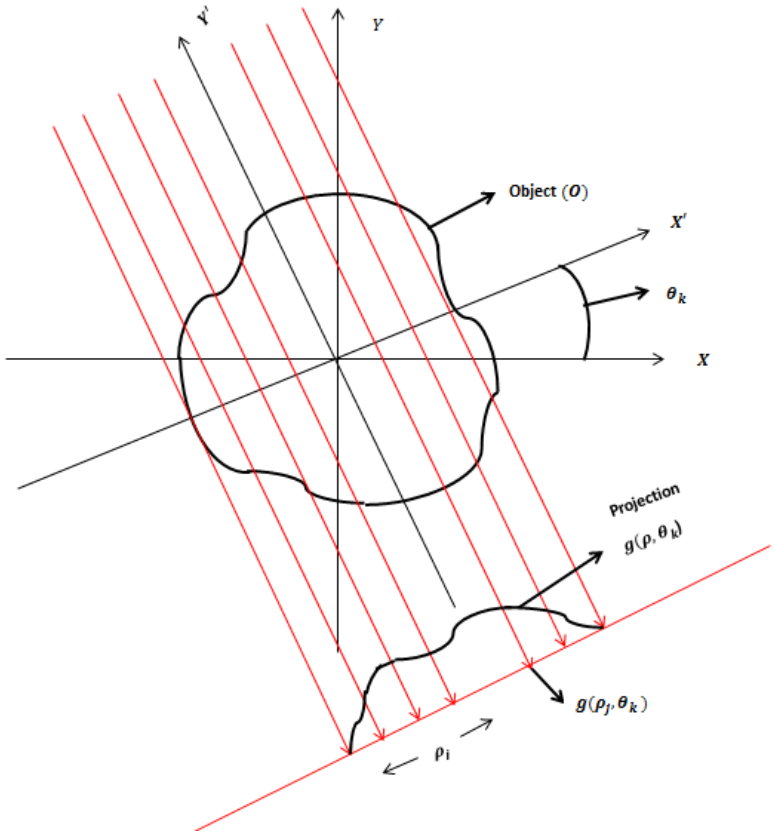

Figure 2.8: The Radon transform of an object (O). The transform is given by $g(\rho, \theta_k)$ where ρ is a vector $(\rho_1, \rho_2 \ldots \ldots \rho_k)$. Each point in the projection is denoted by$g(\rho_j, \theta_k)$ and is nothing but the line integral of the object O along the lineρ_j.

.

Hence to get the projections for $\theta > 180^{\circ}$, we just need to get projections for $\rho < 0$ and thus physically we just need $0 \leq \theta \leq 180^{\circ}$ for the Radon transform. Next, we state the Fourier Slice Theorem that relates the Fourier Transform of the object with the Fourier Transform of the Radon transform.

Fourier Slice Theorem (FST): The FST relates the 1D Fourier transform of the projections $g(\rho, \theta)$ to the 2D Fourier transform of the object $f(x, y)$.

Theorem: *The 1D Fourier transform of* $g(\rho,\theta)$*is nothing but a slice through the origin (along* $x\cos\theta + y\sin\theta = \rho$*) of the 2D Fourier transform of* $f(x,y)$.

Next, we attempt to establish this mathematically.

The 1D Fourier transform of $g(\rho,\theta)$ is given by

$$G(\omega,\theta) = \int_{-\infty}^{+\infty} g(\rho,\theta)e^{-j2\pi\omega\rho}\,\partial\rho \quad (28)$$

We also note that,

$$G(\omega,\theta+180^{\circ}) = G(-w,\theta) \quad (29)$$

Equation (29) relates the Fourier Transform of the slice to the result in (27). It also gives us a way to compute the Fourier transform of the projection $g(\rho,\theta+180^{\circ})$. This is because by definition,

$$G(\omega,\theta+180^{\circ}) = \int_{-\infty}^{+\infty} g(\rho,\theta+180^{\circ})\,e^{-j2\pi\omega\rho}\partial\rho$$

$$= G(-\omega,\theta)$$

Now, from equation (28) we have $G(\omega,\theta) = \int_{-\infty}^{+\infty} g(\rho,\theta)e^{-j2\pi\omega\rho}\,\partial\rho$. Let us put Equation (25) for $g(\rho,\theta)$ i.e. $g(\rho,\theta)= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - \rho)\partial x\partial y$ in Equation (28) and get

$$G(\omega,\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - \rho)\,e^{-j2\pi\omega\rho}\,\partial x\partial y\partial\rho \quad (30)$$

Now putting $\rho = x\cos\theta + y\sin\theta$ in equation (30) we get,

$$\delta(x\cos\theta + y\sin\theta - \rho) = \delta(0) = 1$$

And we can integrate out $\partial\rho$.

Hence,

$$G(\omega,\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,e^{-j2\pi\omega(x\cos\theta+y\sin\theta)}\partial x\partial y \quad (31)$$

$$= F(\omega cos\theta, \omega sin\theta) \quad (32)$$

Equation (32) is the slice of 2D Fourier transform of $f(x,y)$ along the line $xcos\theta + ysin\theta = \rho$, which passes through the origin. Thus equation (32) gives the crux of the Fourier Slice Theorem (FST) i.e. the 1D Fourier transform of $g(\rho,\theta)$ is nothing but a slice through the origin (along $xcos\theta + ysin\theta = \rho$) of the 2D Fourier transform of $f(x,y)$. Hence the FST is proved for 2D and it is essentially the same for 3D as well.

Reconstruction (using back projection): Given the Fourier transform $F(u,v)$ of the 2D object $f(x,y)$, we can get the original object $f(x,y)$ using the inverse Fourier transform as follows:

$$f(x,y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F(u,v)e^{j2\pi(Ux+Vy)}\,\partial U\partial V \quad (33)$$

Now converting to polar coordinates, we get, $U = \omega cos\theta$ and $V = \omega sin\theta \Longrightarrow \partial U\partial V = \omega\partial\omega\partial\theta$

Therefore equation (31) can be written as,

$$f(x,y) = \int_0^{2\pi}\int_0^{\infty} F(\omega cos\theta, \omega sin\theta)e^{j2\pi\omega(xcos\theta+ysin\theta)}\,\omega\partial\omega\partial\theta \quad (34)$$

We note that to get the entire space we take $0 \leq \omega \leq \infty$ and rotate θ from 0 to 2π. Now by the FST we get,

$$F(\omega cos\theta, \omega sin\theta) = G(\omega,\theta)$$

$$\Longrightarrow f(x,y) = \int_0^{2\pi}\int_0^{\infty} G(\omega,\theta)e^{j2\pi\omega(xcos\theta+ysin\theta)}\,\omega\partial\omega\partial\theta \quad (35)$$

Now we also know that,

$$G(\omega,\theta + 180^{\circ}) = G(-\omega,\theta)$$

Using this we get,

$$f(x,y) = \int_0^{\pi}\int_0^{\infty} |\omega| G(\omega,\theta)e^{j2\pi\omega(xcos\theta+ysin\theta)}\partial\omega\partial\theta \quad (36)$$
$$= \int_0^{\pi}[\int_0^{\infty}|\omega| G(\omega,\theta)e^{j2\pi\omega\rho}\partial\omega]\partial\theta$$

Here $\rho = xcos\theta + ysin\theta$

Now the term $\int_0^\infty |\omega| G(\omega,\theta) e^{j2\pi\omega\rho} \partial\omega$ is the 1D inverse Fourier transform of the projections $g(\rho,\theta)$. Here $|\omega|$ is the 1D filtering and hence the name filtered back projection. What we have derived in (13) is what is called the filtered back projection algorithm for reconstruction.

Hence the filtered back projection is done in combination of the following steps:

1) Given the projections $g(\rho,\theta)$ for each fixed θ
2) Compute $G(\omega,\theta)$, which is the 1D Fourier transform of the projections $g(\rho,\theta)$
3) Multiply $G(\omega,\theta)$ by the filter function $|\omega|$
4) Compute inverse Fourier transform of the results from (3)
5) Interchange of sum over θ from 0 to π on the result from (4)

Note that this is the ideal situation where we have slices for all $0 \leq \theta \leq \pi$ or $-\pi/2 \leq \theta \leq \pi/2$. If we have the information for the different angular slices, then the filtered back projection algorithm gives the perfect reconstruction. But, in general, due to the mechanical limitations of the microscope, we get slices from $-70^\circ \leq \theta \leq 70^\circ$. Hence, in Fourier space the missing angular data creates a "bow-tie" pattern. This phenomenon gives rise to the problem of reconstruction when projections for limited tilt angles are available. This is called the problem of "Reconstruction for limited tilt angles".

We are now ready to explain the missing wedge problem for 3D functions. In the Fourier space, the relationship between an object and its projection is referred to as the *central section theorem, the central slice theorem or the Fourier Slice Theorem* says that, the Fourier transformation G of projection g of a 3D object d is the central (i.e., passing though the origin of reciprocal space) 2D

plane cross-section of the 3D Fourier transform of the object D and is perpendicular to the projection vector. This provides the insight as to why the missing wedge problem arises in ET.

Thus, if the projections are such that, their Fourier transforms generate all possible central slices covering the entirety of the 3D Fourier transform of the object, then we have a perfect reconstruction and there is no missing data. In theory, a complete coverage in the Fourier space can be obtained by rotating the sample ±90° about a single axis, called the tilt axis, with equal angular increments. In practice, due to some mechanical limitations of the electron microscopes, the maximum achievable tilt range is ±60° to ±70°. In the Fourier space of the 3D reconstruction, the limited tilt range results in the wedge-shaped region, empty of information, called the "missing wedge" (Fig. 2.9 and Fig: 2.10a). The missing wedge has a significant impact on the resolution of the 3D reconstruction of the specimen, making the resolution of the reconstruction anisotropic i.e. direction dependent [27]. In real space, because of the lack of specimen views in the high tilt angles, the missing wedge produces artifacts in the tomograms, such as blurring the spatial features in the direction parallel to the electron beam and this makes some features look elongated in the beam direction. As a result of the anisotropic resolution, features oriented perpendicular to the tilt axis, tend to fade away from the view causing signification loss of resolution.

The effects of the missing wedge are most clearly demonstrated when following the contours of cell membranes. Because of the missing wedge effect, the cell membrane from the top and the bottom will disappear in the tomogram while the membrane from the sides will be clearly visible.

Minimization of the volume of the missing wedge can alleviate this problem effectively. To minimize the volume, one possible way is to increase the tilt angles of the specimen. Unfortunately, this not achievable due to several factors. Firstly, microscope limitations do not allow tilting more than ~70°. Fortunately, this problem can be partly solved by taking a second dataset after rotating the sample by 90°. The un-sampled volume in the dual axis tomography, called the "missing pyramid" (shown in Fig; 2.10b), is considerably smaller than the "missing wedge". A ±70° tilt range involves that 22% of the information is missing and the use of double-tilt axis acquisition geometry reduces the missing information down to 7% [27].

Secondly, extreme radiation sensitivity of the biological samples limits the number of images from unstained frozen-hydrated samples. To handle this issue, the automated image collection procedure records images under a significantly low electron dose by dividing the maximum tolerable cumulative dose (120 $e^-/Å^2$) over the total number of images. Therefore, an individual image becomes extremely noisy with SNR ~0.01.

2.3 Noise Reduction

As seen from the previous discussion, noise is inevitable in ET and hence a cause for major concern. The best possible way to remove the noise would be to create a model for the noise and apply it to the tomogram. But the noise in a tomogram is usually generated by non-linear combinations of different contributing factors and hence modeling it accurately is quite difficult. This drawback forces researchers to use local techniques, which do not capture the noise model in its entirety.

In general, the noise reduction techniques can be grouped into three categories, namely linear, nonlinear and anisotropic. Linear techniques are used on local averages using Gaussian like

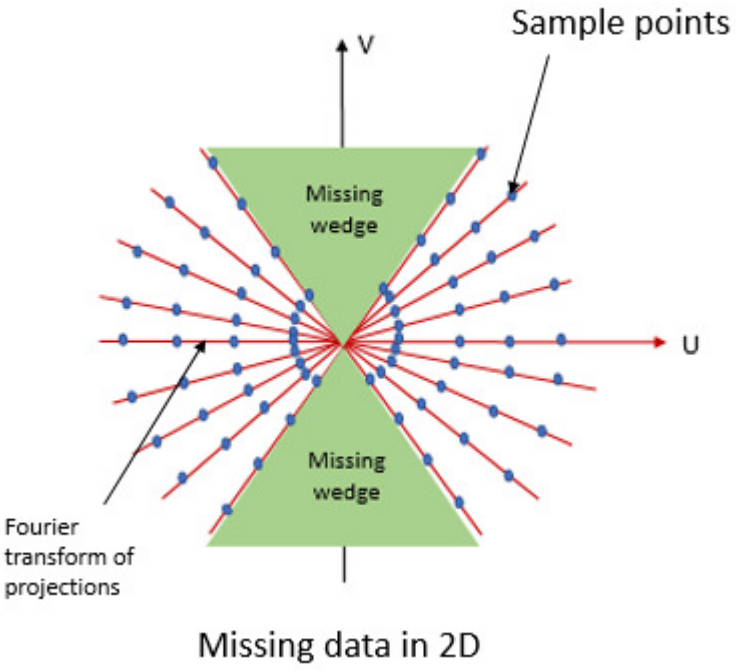

Figure 2.9: Illustration of missing wedge in reconstruction of 2D functions from 1D projection.

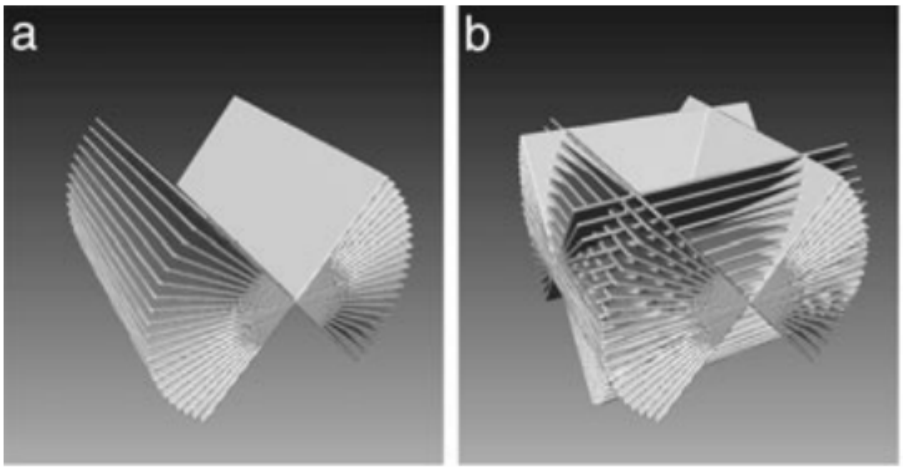

Figure 2.10: a) Missing wedge, b) Missing pyramid [3] .

kernels. Here all the voxels are substituted by a weighted average of the voxels in a neighborhood where the weight functions are given by the Gaussian kernels. In linear techniques, the same kernel is applied for all the voxels. One major drawback of this set of methods is the blurring produced by the use of same kernel across all voxels.

The nonlinear techniques overcome these drawbacks by tuning the kernel parameters to reflect the specific detail of the voxels to which the kernel is being applied. In this way, the filtering is strong in homogeneous areas whereas it is considerably week in areas with high gradient information, which potentially has more detail. One major drawback of this method is that the noise is not completely removed near the edges, which have a higher gradient.

Anisotropic methods overcome this problem by tuning not only the strength of the kernel but also the direction of the filter. Thus, edges are subject to a filtering process that runs parallel to them in the process of cleaning and enhancing the information near the edges.

In the case of tomograms, the first step is usually a linear Gaussian filter, which is used as a preprocessing step to remove noise allowing reliable computation of gradients. Next, more sophisticated techniques, like anisotropic diffusion or median filtering can be used to reduce the noise further. Median filtering is a nonlinear method in which a voxel value is replaced by the median of the values of the nearby voxels. This is an iterative technique and needs to be repeated a number of times (usually 3) to achieve significant reduction of the noise.

However, none of the state-of-the-art denoising techniques work well with ET. This is because both anisotropic diffusion and median filtering increase the contrast in the tomogram but at the same time blurs out the intrinsic detail if the structure under study. As a result, the only remaining option for denoising in ET is through the use of sub-tomogram averaging [20].

2.4 Segmentation

Image segmentation is the process to group pixels into subsets to create non-overlapping partitions in the image, which correspond to meaningful regions or objects. Every pixel in an image is allocated to one of a number of these regions. In a good segmentation, the partitioning of all image pixels is performed in a way, such that the affinity between the pixels belonging to the same region or subset is higher than that between pixels of different regions or subsets. The pixels affinity must be quantified by defining adequate measures based on pixel properties. In essence,

- Pixels in the same region should be homogeneous with respect to some predefined similarity measure such as grayscale value, pixels proximity, texture etc. and form a connected region.
- Pixels belonging to a neighboring region should be as heterogeneous as possible with respect to the multivariate similarity measure and two neighboring regions must be non-overlapping. i.e. the entire image, $R=U_{i=1}^{S} R_i$ and $R_i \cap R_j = \phi$ for $i \neq j$ and each R_i is connected.

One of the primary objectives of image analysis is to be able to identify and analyze objects of interest from a given image. Segmentation is the method by which objects of interest are identified automatically or semi-automatically from a given image. Hence segmentation is the first step that is required to achieve the goals of image analysis. Thus, it is one of the most important steps one encounters in any image analysis application.

CHAPTER 3

RELATED WORKS

In this chapter, we briefly discuss previous work related to two broad aspect of this dissertation. The first relates to the problem of segmenting macromolecular structures from volume data obtained using electron tomographic techniques. The second aspect relates to the determination of the structure of macromolecular assemblies and filaments images using electron tomography and single particle electron microscopy respectively. We start with discussion on the segmentation techniques followed by a discussion on analysis of electron tomographic structures and finally concluding with analysis of helical structures from single particle electron microscopy.

3.1 Segmentation of Volumetric Data

Of all the steps involved in the analysis of electron tomograms, segmentation is particularly challenging partly because of the complexity of biological features and partly due to the poor quality of the micrographs. Some of the challenges for applying state-of-the-art segmentation techniques in ET are discussed below [37]:

1) **Data set variety:** The complex nature of biological specimens comes not only from different biological structures but also from different method of specimen preparation.
2) **Low and non-uniform contrast:** The contrast in cellular tomography is often low and non-uniform along membranes and fibers. Sometimes when the contrast is poor, the structure may be visible in lower magnifications. This creates problem for segmentation methods that rely on local intensity properties.

3) **Low signal-to-noise ratio:** To prevent radiation damage of the biological specimen during data collection limited amount of electron dose is used for image formation. Therefore, the electron tomograms tend to exhibit an extremely high stochastic noise level. In some cases, the segmentation algorithm is applied after using some powerful denoising techniques on the tomograms.
4) **Anisotropic resolution:** Due to the "missing wedge" problem, some parts of the tomogram may have significantly lower contrast than others and the resolution becomes direction dependent and contrast varies across the thickness of the tomogram.
5) **Specimen preparation artifacts:** The specimen preparation method introduces a number of different artifacts such as lack of uniform contrast, discontinuity of membranes etc.
6) **Interfering structures**: This is problem common for segmenting membranes. Some point-like objects, such as ribosomes with a strong uniform contrast, may interfere with the membrane and extend over distances larger than the width of a typical membrane or microtubule.

Developers of segmentation tools for electron tomogram should take care of all the issues, which makes the process inherently challenging.

In practice, image segmentation can broadly be classified into two major categories: The bottom-up approach and the top-down approach. Bottom-up approaches do not consider any prior model of the desired object to start with. Based on the information of the local features, such as gray level values of the pixels, texture, color or edge, the entire image is partitioned into several non-overlapping homogeneous regions. On the other hand, top-down approaches start with a predefined model of the target object and stops when the entire image is segmented into smaller meaningful regions that the operator is looking for. Given an image, the target object is first

localized in the image. After that, it is extracted under guidance of the appearance prior such as texture or shape.

Both approaches, bottom-up and top-down, have some limitations. As the bottom-up approach relies on low-level homogeneity, it often results either in over segmentation or in under segmentation. On the other hand, because of the high intra class variance between the objects being segmented, in terms of object shape and appearance, generating accurate models for the objects are very difficult. Conversely, the shape and appearance of two distinct objects can be very similar in terms of the prior distribution imposed on the appearance. Hence generating accurate models of the desired objects using the top down approaches is quite challenging. Next, we give a brief survey of the different methods that belong to either one of these categories.

3.1.1 Bottom-up Approaches for Segmentation

The bottom-up image segmentation approaches first segment the image into regions and then identify the image regions that correspond to a single object. This method relies mostly on the continuity principle by grouping pixels according to the gray level values or texture uniformity within image regions. In electron tomographic data, the most obvious pixel properties are gray level similarity and proximity.

Several bottom-up segmentation approaches are present in literature, but there is no single method which can be considered good for all types of images, nor are all the techniques applicable for one particular image. In this chapter, different bottom-up segmentation approaches that are used for segmenting ET are presented in brief. Note that, multiple bottom-up techniques can be used in conjunction with others, to solve different segmentation problems.

Histogram-based Thresholding: Thresholding is the simplest and probably the most frequently used technique for image segmentation. In this method, the intensity of the pixels or voxels are compared with a user defined parameter, called the 'threshold' and the pixels or voxels are assigned to the foreground or background according to their intensity above or below the threshold value.

For tomographic data, there is no objective criterion on the choice of the threshold. In single particle technique, where because of comparatively good resolution, a threshold can be set by calculation of the total volume occupied by the protein. Unlike single particle methods, the threshold parameter used in tomography must be set subjectively. This technique was applied to a tomogram of a *Pyrodictium* cell and a binary image was produced [38].

Watershed Segmentation: In grey scale mathematical morphology, the watershed transform is the fundamental image segmentation tool based on an idea inspired by topographic reliefs. The watershed transform can be classified as a region-based segmentation approach. The intuitive idea underlying this method comes from geography: it is that of a landscape or topographic relief which is flooded by water, watersheds being the divide lines of the domains of attraction of rain falling over the region. An alternative approach is to imagine the landscape being immersed in a lake, with holes pierced in local minima. Basins (also called 'catchment basins') will fill up with water starting at these local minima, and, at points where water coming from different basins would meet, dams are built. When the water level has reached the highest peak in the landscape, the process is stopped. As a result, the landscape is partitioned into regions or basins separated by dams, called watershed lines or simply watersheds. Precisely, the basins correspond to the regions of the image being segmented. The algorithm has been used to segment out different 3D electron tomographic volumes.

First the method has been used on simulated and experimental data from smooth-muscle acto-myosin [39]. The watershed algorithm has been applied to the Foot-and-mouth disease virus [40] with a FAB fragment bound.). The FAB fragments are correctly separated from the virus density. The virus capsid is partitioned into its smaller units and the boundaries between the constant and the variable domains of the FAB are clearly visible. Application of this segmentation technique on a raw tomogram of actin cross-linked with aldolase [41] confirms that this method works well even for segmenting the tomograms that are difficult to interpret by eye. Finally, the watershed algorithm was applied to segment out the boundaries of the membrane structures of the Golgi region of a pancreatic beta cell [42].

Normalized Cuts: The idea of a cut based segmentation is to represent the image as a graph, with vertex set equal to the number of pixels in the image. Once this is done, edges are added between the vertices and for every edge a weight is assigned, such that it is proportional to the similarity between the vertices. Once this has been done, the resulting graph encodes the similarity/dissimilarity between the pixels of the image. Now a foreground-background segmentation is obtained by computing the minimum cut (set of edges that partitions the vertex set into two subsets and has the minimum total weight).

The normalized graph cut method was applied to the 2D slice of an electron tomogram of a cell of the archaeon *Pyrodictium abyssi* with its characteristic surface layer, a group of extracellular vesicles, and a fragment of a cannulae [43]. The hierarchical application of this segmentation technique yields the completely segmented *Pyrodictium abyssi* cell with outer membrane, inner membrane, the group of vesicles and the cannulae, segmented from one another and from the ice matrix. This method has been extended to higher dimension and the 3D extension of the method is applied to segment out the complete 3D density map [43].

Active Contour Models & 3D Geodesic Active Contour: A snake as described in [44] is an energy-minimizing spline, whose energy depends on its shape and location within the image. Local minima of this energy then correspond to image properties. Snakes are represented as parametric deformable models. This model is active in the sense that it is always minimizing its energy functional and therefore exhibits dynamic behavior. The basic snake model is a controlled continuity spline under the influence of image forces and external constraint forces. The internal spline forces serve to impose a piece wise smoothness constraint. The image forces push the snake toward salient image forces like lines, edges, and subjective contours. The external constraint forces are responsible for putting the snake near the desired local minimum.

3D geodesic active contour [45] is a technique for detecting object boundaries, based on active contours evolving in time, according to intrinsic geometric measures of the image. The technique is developed based on the relation between active contours and the computation of geodesics, or minimal distance curves.

In [46], a method based on 3D active geodesic contour is reported for the automated segmentation of membranes of HIV virions, which are then used for HIV particle detection. The algorithm finds the boundaries of the object of interest, as global minimal surfaces, in a metric space defined by image features. Then the particles of interest are found by template matching. The segmentation is carried out for individual objects i.e. for individual HIV virions in the tomogram. It must be noted that this method can be used for solving our problem of automatically segmenting HIV/SIV Env pikes. However, the success of this method is dependent upon the availability of a reliable template.

3.1.2 Top-down Approaches to Segmentation or Model-based Segmentation

The complexity of the biological features, the crowded environment, and the inherent low signal-to-noise ratio (SNR), present significant challenges to data-driven methods for segmentation in electron-tomographic reconstructions. Moreover, the subcellular structures that we want to segment out, pose some distinctive geometric properties such as a tubular structure or very thin structure of membrane boundaries. This prior shape knowledge may be obtained from biologists, from statistical analysis of the training shapes, or acquired from user-drawn shapes, and they should be fully exploited to improve the segmentation accuracy and robustness. Next, we discuss the important top-down segmentation methods that have been used in ET segmentation.

Template Matching: In template matching [47] , the template is generally a structure that has been determined at high resolution, usually by an imaging modality other than electron microscopy (most often X-ray crystallography). After filtration of the template, to match the approximate resolution of the electron-tomographic reconstruction, and direction-dependent adjustment, to match the missing angular information, cross correlation is used to find objects matching the template in the 3-D tomographic reconstruction. Preliminary studies indicate that template matching is reasonably effective for identifying large macromolecular complexes [48].

However, these initial efforts have to deal with a significant number of false positives, and have not yet been used to any appreciable extent, to segment 3-D reconstructions of in situ cellular complexes. Other major drawbacks of template matching include difficulty of constructing templates for complex biological structures, and its inability to incorporate variability in the shape of the structures. Template matching is also limited to finding only structures with a strong resemblance to the template; it is not particularly effective at finding an unknown structure.

The template is matched to every location within the tomographic reconstruction. In a brute force template matching [47], to compute the similarity, the cross-correlation is computed between the template and tomogram, at every location within the tomogram, and for all (predefined) angular orientations.

Depending on the feature of interest, simple templates are generated based upon general 3D shapes like cylinders, cubes, spheres, or cuboids. For the relatively low-resolution requirements of the template-matching application, templates like a cylinder for microtubules and cuboids for patches of membrane are generated. For further information on application of template matching to segmentation of ET, readers are directed to [49].

Water-snakes & Shape Driven Water-snakes: One major advantage of the traditional energy-based methods is the ability to easily incorporate prior knowledge into the segmentation process. On the other hand, in traditional watershed methods, though incorporating prior knowledge of the number of objects, and their location is possible, it is not conducive to incorporation of a priori shape information. Hence, by reformulating the watershed segmentation in an energy minimization framework, the water-snake model offers the best of both worlds while avoiding their pitfalls.

The traditional watershed algorithm is implemented via region growing, where seeds are the regional minima of the relief. To avoid over segmentation due to a large number of minima in the original edge evidence function, the watershed line is constructed from a given set of regions called the markers. For more information, please refer to [50].

A 3D model based approach for segmenting biological membranes of mitochondria and bacteria, by incorporating the prior knowledge about the shapes of membranes, was developed by [48] combining the watershed segmentation and the prior shape information.

Maximum Likelihood Estimation: Commonly used approaches of subvolume analysis consist of two steps - an initial alignment step followed by a subsequent classification step. The alignment step is mostly based on either conventional cross correlation or constrained cross correlation [51]. Classification algorithms use a distance metric to measure similarity between sub-tomograms and groups sub-tomograms into classes. An intrinsic limitation of this approach is that, alignment and classification are performed in two subsequent steps and the classification is dependent on alignment. Clearly, poor alignment produces poor classification. Existing strategies of sub-tomogram alignment and classification are less robust when the signal-to-noise ratio of the sub-tomograms is very low.

An alternative 3D alignment approach based on maximum-likelihood optimization is used by [51] for three dimensional subtomogram averaging and classification. This approach, called MLTOMO, is based on a probabilistic data model, which comprises both an estimate of the underlying structure, as well as a formal description of the noise and the distribution of the alignment parameters. The goal of the ML refinement procedure is to find a set of most likely parameter values that describes the experimental data. In MLTOMO, the relative orientations of the subvolumes and their class assignments are treated as hidden variables, and expectation maximization is used to maximize the corresponding marginal likelihood function. The resulting algorithm is robust to high levels of noise and simultaneously tackles the problems of alignment and classification by calculating model parameters as probability-weighted averages over all possible orientation and class assignments. This method has been implemented in the RELION

software package [52] which is used for analyzing structure of both single particle electron micrographs and ET.

3.2 Structure Determination from Electron Tomographic Data

In this section, we describe methods for structure determination of macromolecular assemblies using ET. As the methods used for structure determination depend on the macromolecular assembly being studied, we discuss this in the context of determination of structure of HIV/SIV spikes. The HIV/SIV Env spikes initiate infection by facilitating entry of the virion into the host cells. They are also the sole protein on the surface of the virion accessible to the cells immune system [53]. Hence, understanding their structure will provide insight into host cell infection and may eventually help create effective vaccines against AIDS. Though we discuss some of the more important contribution in this area in Chapter 4, interested readers are hereby referred to [54]

3.3 Structure Determination from Single Particle Electron Microscopic Data

3.3.1 Actin Filament as a Helix

Our second project was structure determination of actin-bound smooth muscle myosin-II (smM) complex in the nucleotide-free state. All muscles contain thin actin-containing filaments and thick myosin-containing filaments. A substantial portion of the cytoplasmic volume of smooth muscle cells is taken up by the molecules myosin and actin. Myosin present in a smooth muscle is primarily myosin II. Muscle thin filaments (diameter 6–10 nm) are a double helix of polymerized actin monomers. The double helix repeats once every 28 monomers if the monomers from both strands are counted. Due to the helical nature of the filament, the molecule

repeats every 14 monomers if the distinction between strands is ignored. The F-actin, can be described as 13 subunits in 6 turns. F-actin can have different helical structures but usually are found between the 13/6 symmetry for vertebrate striated muscle and the 28/13 symmetry for *Lethocerus* flight muscle

Representation of Helix with Cylindrical Coordinates: In Cartesian coordinates, a continuous helix can be defined by a set of three equations, $x=r\cos(2\pi z/P)$, $y=r\sin(2\pi z/P)$, $z=z$; these describe a circle in the x-y plane that gradually rises along the z axis (Fig. 16a). The diagnostic parameters of the helix include the radius (r) and the repeat distance along the z axis, or pitch (P). In cylindrical coordinates, which are the most convenient way to describe a helix, these equations become $r=r$, $\phi=2\pi z/P$, $z=z$. These equations describe a continuous helix, such as the continuous wire path of a spring (Fig. 3.1A), but biological assemblies generally, involve a discontinuous helix (Fig: 3.1B), built with individual building blocks, or subunits, positioned at regular intervals along the helical path.

These assemblies are characterized by the angular and axial interval between the subunits, $\Delta\phi$, and Δz, or alternatively are characterized by the number of subunits per turn of the helix. From these values, the repeat distance (c) can be calculated as $c = u\Delta z = tP$ (The pitch of a helix is the height of one complete helix turn, measured parallel to the axis of the helix). A helical repeat is defined as the distance that a subunit must be translated along the axis to be in register with another subunit. This helical repeat must be the product of an integer multiplied by the axial rise per subunit. Actin filaments have the simplest possible repeat, 13 subunits in 6 turns of the left-handed 1-start helix, with an axial rise per subunit of 27.3 Å. This filament will have a repeat of 13*27.3 Å=355 Å. The angle between adjacent subunits is 360°*6/13=166.15°.

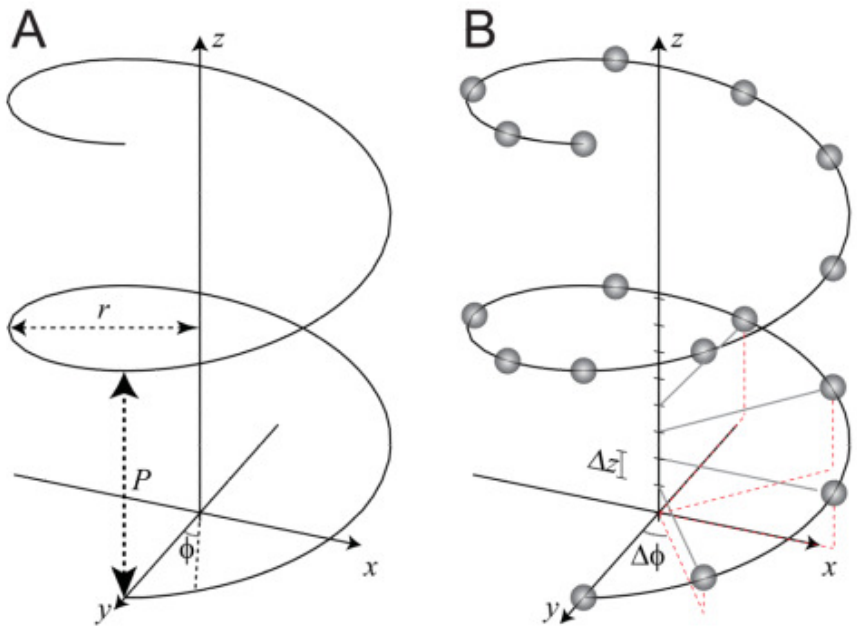

Figure 3.1: Diagrams depicting the geometry of a helix. (A) A continuous helix is characterized by the pitch (P) and the radius (r) adopted by the spiral. Either a Cartesian coordinate system (x,y,z) or cylindrical coordinate system (r,φ,z) can be used. In either case, the z-axis corresponds to the helical axis. (B) Helical assemblies are generally composed of identical subunits arranged along the path of a continuous helix. This requires additional parameters, Δφ and Δz, which describe the incremental translation and rotation between the subunits [55]

3.3.2 Algorithms for Helical Reconstruction

Many macromolecular assemblies in cells are constructed using identical protein subunits arranged with helical symmetry. The presence of helical symmetry is a terrific aid in a structure determination because a single view of the assembly provides multiple different views of the protein subunits. Consequently, fewer micrographs are needed to obtain a structure. In fact, a low resolution 3-D structure can generally be obtained from a single view. In order to determine the 3D structure of helical specimens, the symmetry determination is the critical step of the structure determination procedure. The traditionally used approach for computing a 3-D image of a filament with helical symmetry is the Fourier-Bessel approach [55]. But, in recent time a real

space method, called Iterative Helical Real Space Reconstruction (IHRSR) [56] is the near universally used algorithm for reconstructing helical assemblies.

Fourier-Bessel Method for Helical Reconstruction: The Fourier transform of projection images of helices contains a pattern of structure factor peaks termed layer lines (Fig. 3.2A). The spacing between layer lines is a function of the helical repeat. The layer lines are sharp along the direction of the helical axis, but continuous perpendicular to the helical axis due to the lack of repeats in that direction. The vertical axis of the transform is referred to as the meridian and the horizontal axis termed equator. Diffraction patterns of helices, contain contributions from two separate lattices: first, the lattice of the front or the near side of the helix facing the observer and, second, the lattice of the back or the far side of the helix, which faces away from the observer. Each of these lattices gives rise to a set of corresponding layer lines in the Fourier transform, which are mirror symmetric along the meridian.

The interpretation of the diffraction pattern is accomplished by decomposing the Fourier spectrum into its set of layer lines. This process is also known as indexing the helical diffraction pattern and is the crucial step for de novo structure determination of helical assemblies. Layer lines are found aligned perpendicularly to the meridian and spaced apart by the reciprocal distance of the helical repeat c. Mathematically, the layer lines are described by oscillating Bessel functions $J_n(X)$ with discrete orders n that depend on the helical radius r and the reciprocal radius R from the meridian $J_n(2 * \pi * R * r)$ [55].

The Fourier spectrum reveals that peaks from layer lines of small Bessel order n are located close to the meridian as opposed to large Bessel orders that are spaced further away from the meridian and possess decreasing peak maxima. Each layer line has a horizontal position in the Fourier

spectrum, which is described by the layer line height, h, along the helix axis or the layer line number, l. Once the layer lines have been assigned with their layer line heights, h, and order n, a real-space lattice can be derived from the intersections of the layer line waves, which defines the unit positions of the helical array.

Indexing the diffraction pattern is the key step to determine helical symmetry. If the helical symmetry is incorrect, the structure will be incorrect. First, the layer line heights, h, is measured. Preferably sharp layer lines with low Bessel orders close to the meridian should be examined first and assigned. Second, their meridional distance is measured and the corresponding Bessel order n determined by taking the helical radius, r, obtained from real-space analysis into account $J_n\,(2 * \pi * R * r)$. In addition, the phase difference from opposite sides of a layer line helps to discriminate whether the corresponding Bessel orders, n, are odd or even. The intersections of the corresponding helical waves can be directly converted into the rise/rotation or pitch/unit number parameter convention. Using these parameters, the layer lines selection rule can be determined and used to test whether the predicted layer lines agree with the observed ones and whether the observed ones have been indexed correctly.

The nature of the helical symmetry can better be described using a cylindrical polar coordinate system. Thus, if the density distribution of the object is $\rho(r, \emptyset, z)$ and its Fourier transform is (R, Φ, Z) , where the axis of helical symmetry is the z direction.

The Fourier transform of a structure $\rho(r, \emptyset, z)$, with helical symmetry can be expressed in the form

$$F(R, \Phi, Z) = \sum_n G_{n,l}\,(R)\exp[in\,\left(\Phi + \frac{\pi}{2}\right)] \qquad (37)$$

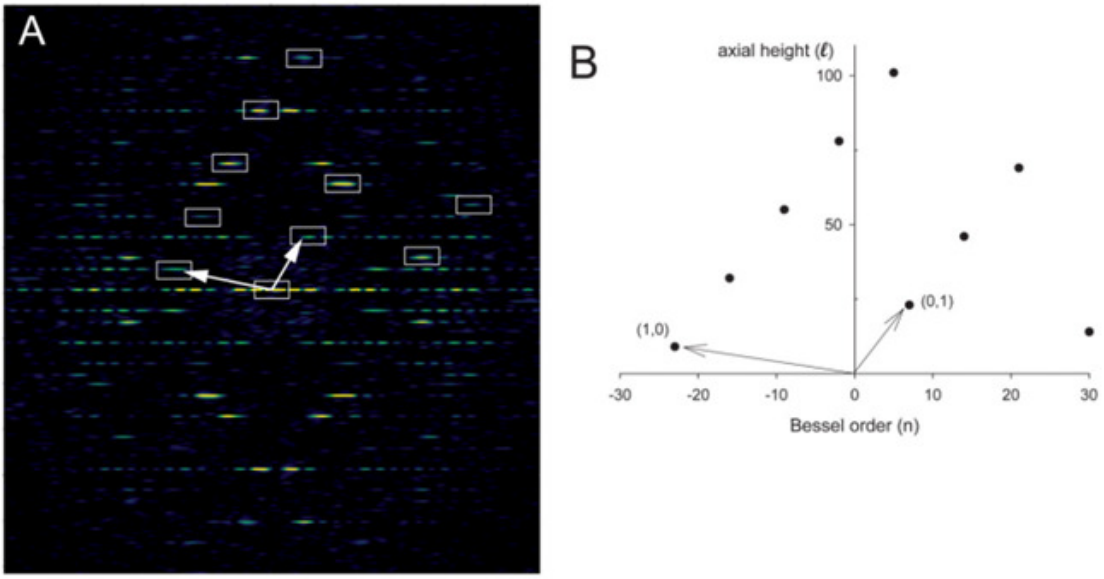

Figure 3.2: A) Fourier transform of a helical assembly. Discrete layer lines that run horizontally across the transform characterize the 2D Fourier transform from a helical tube. Each layer line corresponds to a helical family. The layer line running through the origin is called the equator. The vertical axis is called the meridian and the transform has mirror symmetry across the meridian provided that the layer line does not have contributions from Bessel functions of different orders, referred to as "beating of Bessel functions". The layer lines are mathematically described by oscillating Bessel functions and the start number of each helix (n) determines the order of the Bessel function appearing on that layer line B) Indexing of layer lines in the Fourier transform of a helical assembly. Corresponding plot of Bessel order (n) vs. layer line height (ℓ). Assignment of (1,0) and (0,1) layer lines is arbitrary, but once chosen then all of the other visible layer lines should be a linear combination of these two [55].

The transform $F(R, \Phi, Z)$ is non-zero only for $Z = l/c$ where l is the layer line number and c is the axial repeat of the structure [57]. Repeat distance c satisfies the condition

$$\rho(r, \emptyset, z) = \rho(r, \emptyset, z + c) \quad (38)$$

$G_{n,l}(R)$ is the transform of the layer line data. In real space,

$$g_{n,Z}(\mathrm{R}) = \int G_{n,Z}(R) J_n(2\pi r R) 2\pi dR \quad (39)$$

In real space, the density is computed as,

$$\rho(r, \emptyset, z) = \sum_{-\infty}^{\infty} g_{n,Z}(R) \exp(in\emptyset) \exp(-i2\pi zZ) \quad (40)$$

Problems with Traditional Fourier-Bessel Approach:

If a helical polymer is highly ordered over long distances, is rigid and straight, and diffracts so that only a single Bessel function is present on any layer line within the resolution being studied, then the Fourier-Bessel approach works well. This approach doesn't work for real non-ideal specimens. An alternative, real space approach is Iterative Helical Real Space Reconstruction (IHRSR) method [56]. Problems of Fourier-Bessel methods are manifold:

1) Non-integer symmetries: Fourier-Bessel formalism was based upon crystallographic description symmetry, and was therefore relies on helix having an integer number of subunits in an integer number of turns. For example, the symmetry of F-actin in this formalism can be described as 13/6 i.e. 2.1667 subunits per turn. This problem is quite problematic as small changes in twist leads to catastrophic changes in symmetry [58].
2) Bessel overlap: The problem of Bessel overlap arises where more than one Bessel function occurs on a layer line due to symmetry of the object. This arises in most objects at high resolution, but in some objects (with a small, integral number of subunits in a 'repeat'), it can arise at very low resolution [58].
3) Helical disorder: The disorder that is naturally present in many polymers can be the main limitation in applying Fourier-Bessel analysis. Within a crystal, a space group maintains long range order, while local disorder can exist. Within a helical polymer, there are no forces or factors that maintain long-range order. Thus, all interactions are local, and the variations in these local interactions accumulate to cause liquid disorder, or disorder of second type. For some systems, (such as bacterial flagellar or Tobacco Mosaic Virus), the scale on which a structure is ordered may be long enough that this has no practical consequence. For other

systems, such a F-actin, liquid like cumulative disorder can arise over rather short distances. All helical filaments have some degrees of 'cumulative' disorder. The helical disorder of such filaments precluded using standard Fourier-Bessel reconstruction methods [58].

4) Heterogeneity: The Fourier-Bessel method cannot separate out heterogeneity within and between filaments. The Fourier-Bessel method involves, the imposition of helical symmetry on a single filament, and thus all structural variation is averaged together within this filament. While heterogeneity is a form of "helical disorder", in the sense that strict helical symmetry is absent or broken in both cases, it is helpful to treat structure heterogeneity separately [58].
5) Weakly diffracting specimens: There are many thin helical filaments in biology, (for example bacteriophage and bacterial pili), for which conventional techniques, based upon trying to extract layer lines from individual filaments, fail when the Fourier transform of a single filament is so weak that not even a single layer line is typically seen. Fourier-Bessel method cannot handle weakly diffracting specimens [58].

On the other hand, IHRSR addressed all the above-mentioned problems and overcomes them successfully.

Iterative Helical Real Space Reconstruction (IHRSR): An alternative to Fourier-Bessel reconstruction of helical assemblies, is the Iterative Helical Real-Space Reconstruction (IHRSR) method [56]. The method is based on matching short segments from the image of a helical assembly to a series of projections from a model using the SPIDER software, in a manner analogous to the single particle analysis of isolated macromolecular complexes. The segments are typically much shorter and IHRSR is therefore able to compensate for shorter-range disorder. Furthermore, indexing of the layer lines in the Fourier transform is unnecessary, though knowledge of $\Delta\varphi$ and Δz for the smallest pitch helix is generally required. Specifically, after

using projection matching to determine the relative orientations of all the individual segments along the helical assembly, a 3D structure is generated by back-projection. The helical symmetry of this 3D structure is then determined empirically by examining auto-correlation coefficients after systematically rotating and translating the structure about its helical axis. Once helical parameters ($\Delta\phi$ and Δz) are determined, the structure is symmetrized and used for the next round of alignment and projection matching [55].

The uniqueness about IHRSR approach is that the algorithm determines the screw symmetry, the coupled rotation and axial translation that best fits the reconstructed volume each cycle without the error prone necessity of indexing the Bessel orders. To express this properly, consider a three-dimensional density distribution (the reconstruction) in cylindrical coordinates, $\rho(r, \emptyset, z)$. For a helical object, $\rho(r, \emptyset, z) = \rho(r, \emptyset + \Delta\emptyset, z + \Delta z)$ shows the symmetry between two subunits, where $\nabla\emptyset$ is the rotation between the two subunits, and Δz is the translation along the axis between the two subunits. More generally, $\rho(r, \emptyset, z) = \rho(r, \emptyset + n\Delta\emptyset, z + n\Delta z)$, where n is any integer.

This method for reconstructing helical polymers involved an iterative determination and imposition of helical symmetry upon objects that have been reconstructed without any helical symmetry imposed. A reference volume is used to generate reference projections, where each projection involves a different azimuthal rotation of the reference volume. The actual angular increment between projections depends upon the diameter of the object (D) and the expected resolution (d), and should be $360°*d/(2\pi D)$. Hence, the number of reference projections is $2\pi D/d$. In addition to the helical screw symmetry (an azimuthal rotation coupled with an axial

translation), there can also be a point group symmetry ($C_n$or Dn) present and in this case the number of reference projections is reduced to 2πD/nd.

A multi reference alignment is performed between each projection and the raw images, i.e. the reference projections are cross-correlated against the actual raw image segments. This yielded five parameters for each filament: a correlation coefficient with the most similar reference, an in-plane rotation, an x-shift, a y-shift, and the azimuthal angular orientation of the segment (from the known azimuthal orientation of the reference image that yielded the highest cross-correlation against the raw image). Images that had poor correlation coefficients, or large shifts, were excluded. Back projection was then used on the resulting aligned images, in what is essentially a single-axis tilt series with uniform sampling of all angles and no missing information. This generated a 3D volume in which no assumptions have been made about internal symmetry.

Instead of using multi reference alignment, we have used RELION [52] software for classifying the actual raw segments. In RELION a structure is iteratively refined through a two-step procedure. The first step, which is called Expectation in mathematical terms, has been labeled "Alignment." In this step, computer-generated projections of the structure are compared with the experimental images, resulting in information about the relative orientations of the images .Orientations are not assigned in a discrete manner, but probability distributions over all possible assignments are calculated, and the sharpness of these distributions is determined by the power of the noise in the data. The second step is called Maximization and has been labeled "Smooth reconstruction." In this step, the experimental images are combined with the prior information into a smooth, 3D reconstruction, and updated estimates for the power of the noise and the signal in the data are obtained. The relative contributions of the data and the prior to the reconstruction

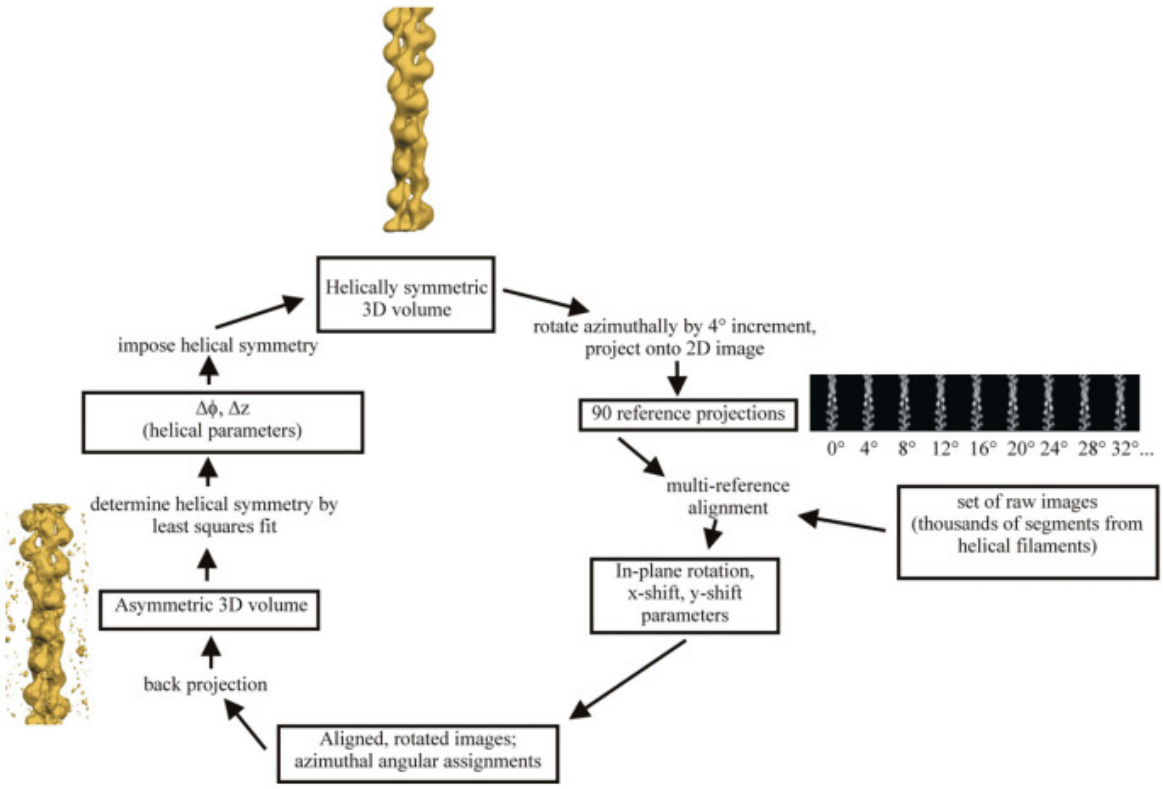

Figure 3.3: A schematic diagram of the IHRSR algorithm [56]. A reference structure (top of diagram) is used to determine the azimuthal orientation, in-plane rotation, and in-plane translation of every image segment. In this example, the reference structure is rotated by 4° increments. The actual azimuthal increment depends upon the diameter of the filament (D), and the expected resolution (d). The reference projections are cross-correlated against the actual image segments. The highest correlation determines the azimuthal orientation of the image in question, as well as providing the in-plane rotation and translation needed to bring it into register with reference projection. A three-dimensional reconstruction is generated by back projection from the aligned images using the azimuthal angular assignment, and this reconstructed volume is used to search for the screw symmetry that minimizes the density deviations. This symmetry is then imposed, generating a new reference volume (top of diagram). The procedure is allowed to cycle until there are no changes in the helical symmetry. The procedure is insensitive to the choice of an initial reference volume, and a solid cylinder can be used quite effectively as an initial reference.

are dictated by Bayes' law and depend on the power of the noise and the power of the signal in the data. The new structure and the updated estimates for the power of the noise and the signal are then used for a subsequent iteration. Iterations are typically stopped after a user-defined number or when the structures do not change further.

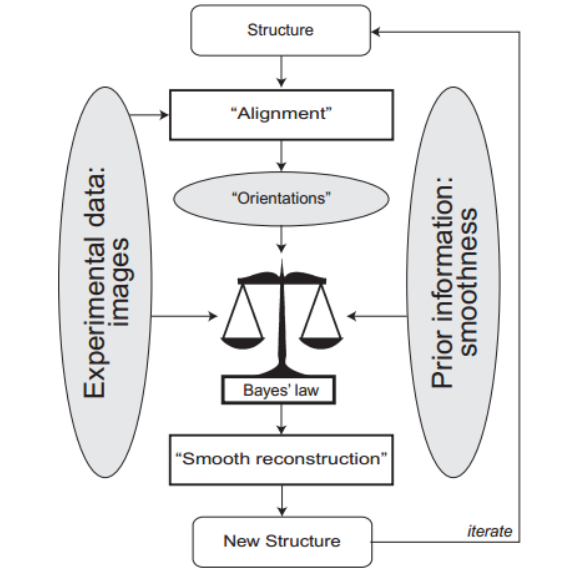

Figure 3.4: A schematic interpretation of the approach of RELION [52].

Determination of helical symmetry in IHRSR: At this point, helical symmetry is imposed upon the central volume, first by determining the two helical parameters, the azimuthal rotation $\Delta\emptyset$ and the axial rise Δz, that related each subunit to its neighbor. The two parameters are not independent (i.e. coupled), hence an initial ''guess'' for Δz is used to determine $\Delta\emptyset$, by calculating the mean-squared deviation ($< d^2 >$) in density between voxels of density and the density at symmetry-related positions in the central volume as $\Delta\emptyset$ was varied [56]. In other words, the central volume is searched for the helical screw operator (the coupled rotation and axial translation) that minimizes the variance between the actual volume and a symmetrized version of the volume. A stochastic gradient descent technique is used to minimize the mean squared deviation. Since the two parameters are coupled, an initial ''guess'' for Δz is used to determine $\Delta\emptyset$, and this value of $\Delta\emptyset$ is then used for a new determination of Δz.The procedure

was then iterated, since the determination of the best value for $\Delta\emptyset$ requires a value for Δz, and vice versa. The test that these values were the best estimates came from the fact that there was no further change in the values of $\Delta\emptyset$ and Δz with further iterations. The parameters $\Delta\emptyset$ and Δz were then imposed upon the 3D volume generated by back-projection, and the resulting perfect helix is then used as a reference for the next cycle.

Structure determination of Actin- Myosin Motor Domain (Acto-MD): In order to understand the mechanism of muscle contraction at the atomic level, it is necessary to understand how myosin binds to actin. Interaction between actin and myosin is one of the key features of contractile events of muscle fibers. Though this has been studied for the last 40 years, not much is known about the structure of the actin-myosin complex. We will discuss the important papers related to the structure of acto-MD in chapter 5 to make it easy for readers to refer back and forth between our work as described in the chapter, and the previous work in this field. However, interested readers are referred to [59] [60] [61] [62] [63] For more recent work about the acto-MD structure intersected red ears are directed to [64] [65] [66] [67].

CHAPTER 4

SEMI-AUTOMATED SEGMENTATION OF SIV SPIKES

The Human Immuno Deficiency Virus (HIV)/ Simian Immuno Deficiency Virus (SIV) envelope (Env) spikes initiate infection by facilitating entry of the virion into the host cells [53]. They are also the sole protein on the surface of the virion accessible to the cells immune system [53]. Hence, understanding their structure will provide insight into host cell infection and may eventually help create effective vaccines against AIDS [9]. To understand the importance of the virus Env spikes we need to understand HIV Replication Cycle first. The HIV/SIV life cycle comprises of the following steps (Fig. 4.1):

1) Fusion of the HIV cell to the host cell surface.

2) HIV RNA, reverse transcriptase, integrase, and other viral proteins enter the host cell.

3) Viral DNA is formed by reverse transcription.

4) Viral DNA is transported across the nucleus and integrates into the host DNA.

5) New viral RNA is used as genomic RNA and to make viral proteins.

6) New viral RNA and proteins move to cell surface to form a new, immature, HIV virus.

7) The virus matures by protease releasing individual HIV proteins.

The spikes are the key parts of a virion that take an important role in the first half of the HIV/SIV life cycle on CD4 T cells. The HIV envelope protein, Env, is comprised of two proteins, gp120 and gp41. In the first step, the Env attaches to the host cell and binds to the CD4 receptor of the helper T cell using gp120; Second step, the conformation of Env changes, and V3 loop is exposed to coreceptor, allowing to coreceptor binding. In the last step, gp41 mediates fusion with

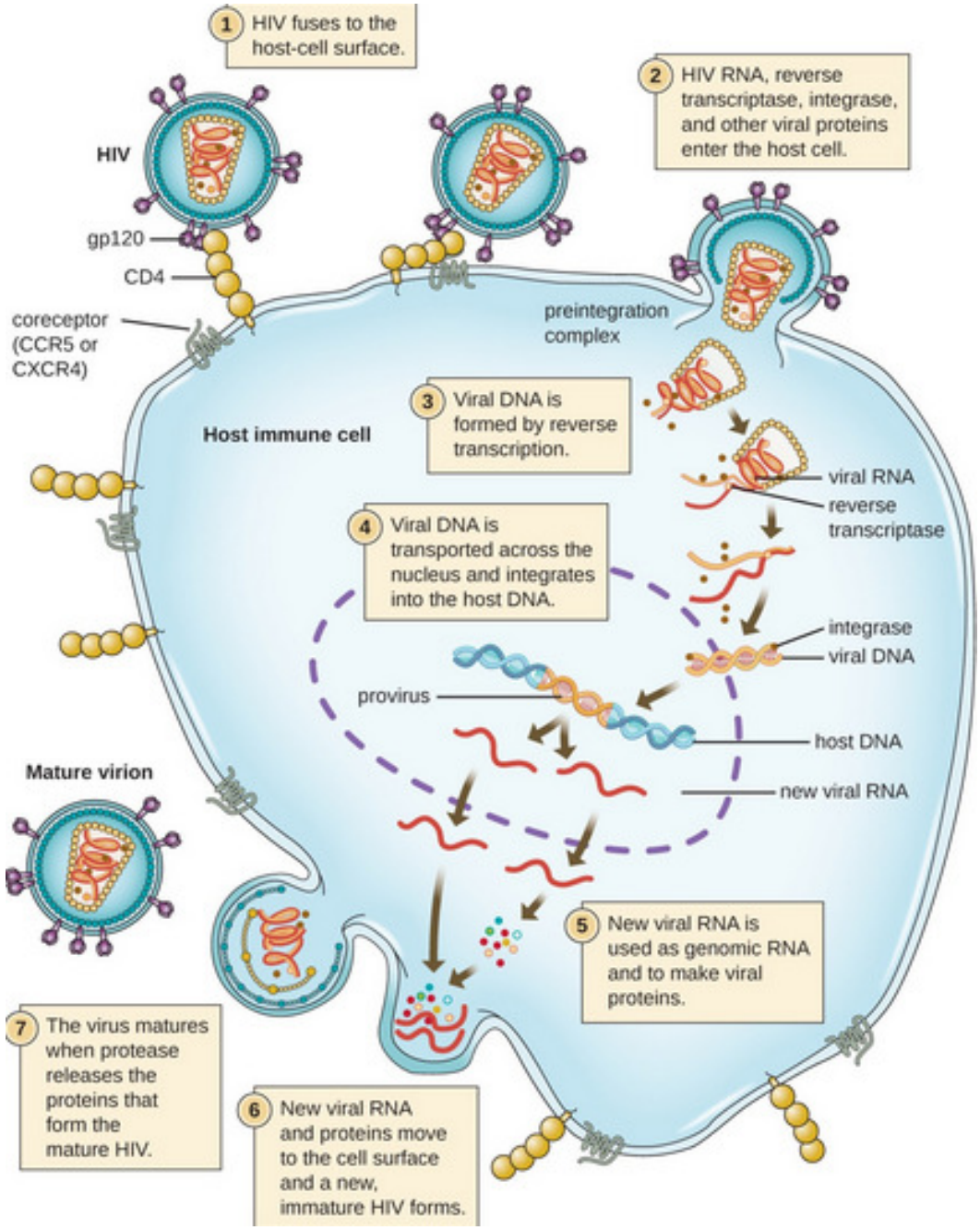

Figure 4.1: Complete HIV replication cycle. 1) Fusion of the HIV cell to the host cell surface. 2) HIV RNA, reverse transcriptase, integrase, and other viral proteins enter the host cell. 3) Viral DNA is formed by reverse transcription. 4) Viral DNA is transported across the nucleus and integrates into the host DNA. 5) New viral RNA is used as genomic RNA and to make viral proteins. 6) New viral RNA and proteins move to cell surface to form a new, immature, HIV virus. 7) The virus matures by protease releasing individual HIV proteins.

the target cell membrane allowing the viral genome to enter the cell. The fusion peptide of gp41 inserts into target membrane and then six-helix bundle forms. Then, membrane fusion completes (Fig: 4.2). Hence, spikes are the key portion of the virus body that initiates the infection in the healthy cell and they are the only location where an antibody can bind with to prevent the infection. Hence, studies of the structure of the spikes are a main emphasis in AIDS research.

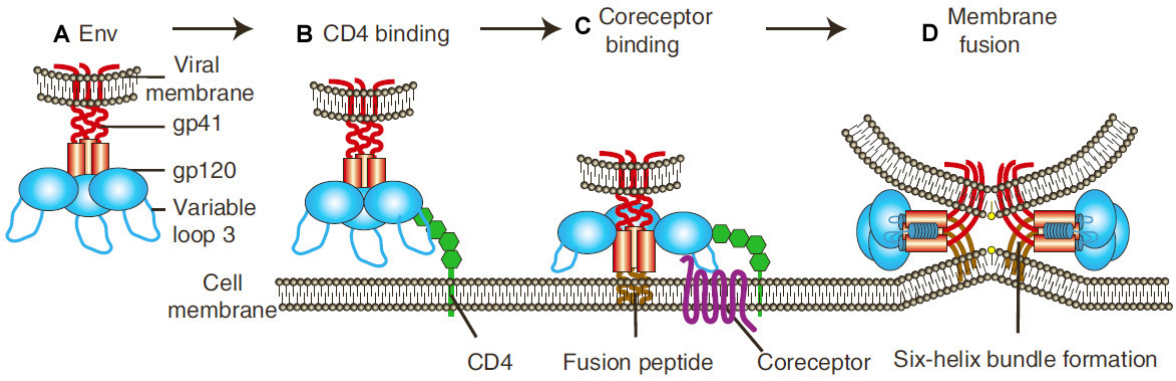

Figure 4.2: The role of spikes in HIV/SIV infection. (A) The sketch of HIV Env. Env is comprised of gp120 and gp41. (B) First step, Env attaches to the host cell and bind CD4. (C) Second step, the conformation of Env changes, and V3 loop is exposure to coreceptor, allowing to coreceptor binding. (D) Last step, membrane fusion is initiated. The fusion peptide of gp41 inserts into target membrane and then six-helix bundle forms. Then, membrane fusion completes. Copied from [130].

A mutant form of SIV virions produced by truncating the small cytoplasmic domain has 80 to 90 envelope spikes per virion (Fig. 4.3), whereas unmodified HIV/SIV virions possess only 8-9 Env spikes [9]. HIV Env spikes have some tendency to cluster [9], whereas truncated SIV Env spikes are more randomly distributed. Thus, any cryoET study of the spike structure becomes very tedious if spikes are selected by hand for further processing. Consequently, automating the process of spike selection is very important for determining the spike structure in situ.

An automatic spike selection method would greatly accelerate research in this area. Automating this process is difficult for four reasons – (1) the automatic process should identify all the spikes

present, (2) it must identify each spike only once, (3) must account for the fact that the virions are not of fixed size or shape, and (4) must account for the possibility that the spikes are heterogeneous in structure. Below, we first describe our segmentation method in detail and then discuss the results of using the same for localizing SIV envelope spikes.

4.1 Introduction

CryoET is a powerful imaging technology for revealing the 3D representation of cellular details to an extent that individual macromolecular assemblies are visualized in the pristine cellular environment. Tomograms from frozen cells contain an overwhelming amount of structural detail. Consequently, a segmentation and/or localization step is inevitably required for the interpretation of an electron density map, in order to identify structures of interest, among numerous other objects that provide the context. To interpret an electron density map, it is absolutely necessary to segregate the map into several constituent parts such as, membrane compartments, filamentous structures, and clusters of associated macromolecules like ribosomes or even virus spikes protruding from the viral membrane. Several techniques are known for tomographic segmentation [43] [41]. The choice of the segmentation method is dictated by the type of structure being segmented and also on the quality of the data. For the segmentation of heterogeneous structures, one of the most common segmentation technique employed is that of template matching or a variation thereof. However, this and related methods can address the problem of segmentation of macromolecular assemblies only if proper priors, in the form of templates, describing these assemblies are available [47]. However, it is usually difficult to get good templates for heterogeneous structures like the SIV/HIV virus envelope spikes. In such cases it is not possible to use template matching and related techniques for the segmentation.

Robust segmentation or localization algorithms for cryoET must overcome some inherent challenges. First, the signal-to-noise ratio of the tomogram is low due to the limited number of electrons that the specimen can tolerate as well as the low contrast inherent in unstained biological specimens [16]. Second, due to the "missing wedge" problem, identical structures are represented differently because the resolution becomes direction dependent and the contrast varies across the thickness of the tomogram [16], [68]. Moreover, cellular tomograms are full of large structural features and macromolecular assemblies. Identification of the different macromolecules from the density map is an inevitable step towards analyzing their structure as well as quantifying their presence. One way of identifying these macromolecules is manual selection. Although manual segmentation is considered reliable, it is also a laborious and time-consuming process. Hence automatic localization techniques are important for macromolecular structure analysis. The most commonly used method of localization is based on a pattern recognition approach using cross-correlation, referred to generally as *template matching*.

Template matching identifies and locates an *a priori* known structure [69] in a cellular tomogram. Template matching is the technique of detection and identification of macromolecular assemblies based on the structural signature [70]. A template can be generated provided a high or medium resolution structure of the macromolecule of interest is available. The templates are then scaled to the voxel size of the target volume, low pass filtered appropriately [71] and then cross correlated with the subvolumes of the same size of the target structure. Based on the value of the cross-correlation thus computed, the structures of interest are localized.

Template matching, in spite of being one of the most common techniques for identification and localization of densely populated macromolecular assemblies, has several drawbacks. Template matching assumes the shape of the macromolecule of interest is known *a priori* and therefore a

template is available. The pattern recognition problem is then reduced to that of finding several occurrences of this template in the density map taking into account the fact that the map suffers from a low signal-to-noise ratio and anisotropic resolution because of the "missing wedge" [71]. In addition, because the approach is driven by a known template, it is ill suited for discovering novel structures within these complex biological volumes.

Motivated by the urge to overcome these shortcomings, we propose a novel localization technique for identification and localization of macromolecular assemblies that does not depend on knowledge of the target object. Our method is based on the idea of "segmentation by classification". "Segmentation by classification" is not a segmentation technique but a localization technique where structures, both known and unknown, within a region of interest are localized by iteratively grouping them into several different classes.

Here we demonstrate our algorithm by using it to localize Env spikes in SIV (Fig.4.3). The HIV/SIV Env spikes initiate infection by facilitating entry of the virion into the host cells [53]. They are also the sole protein on the surface of the virion accessible to the cells immune system [53]. Hence, understanding their structure will provide insight into host cell infection and may eventually help create effective vaccines against AIDS [9]. A mutant form of SIV virion produced by truncating the small cytoplasmic domain has 80 to 90 envelope spikes per virion, whereas unmodified HIV virions possess only 8-9 Env spikes [9]. Thus, any cryoET study of the spike structure becomes very tedious if spikes are selected by hand for further processing. Consequently, automating the process of spike selection is very important for determining the spike structure in situ. An automatic spike selection method would greatly accelerate research in this area.

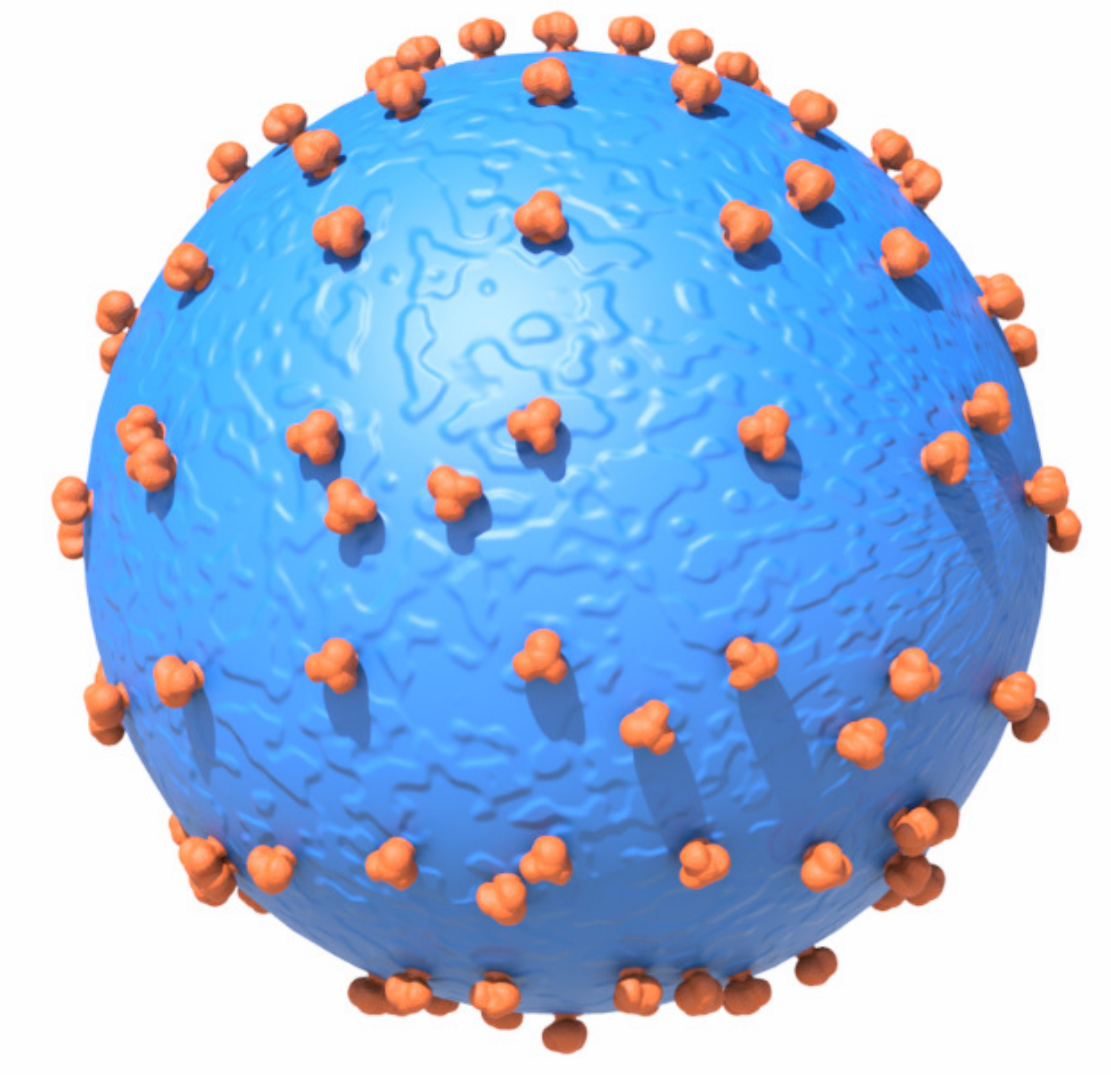

Figure 4.3: A model image of SIV virion (Blue) with spikes (Orange) randomly distributed across the virion surface.

Automating this process is difficult for four reasons – (1) the automatic process should identify all the spikes present, (2) it must identify each spike only once, (3) must account for the fact that the virions are not of fixed size or shape, and (4) must account for the possibility that the spikes are heterogeneous in structure. Below, we first describe our segmentation method in detail and then discuss the results of using the same for localizing SIV envelope spikes.

4.2 Materials and Methods

4.2.1 Virus Sample Preparation

Virus sample: The AIDS Vaccine Program (SAIC Frederick, National Cancer Institute [NCI], Frederick, MD) supplied the highly purified aldrithiol-2-treated virus: SIV 239/251 TAIL/SUPT1-CCR5 CL.30, lot P3978 (SIV short-tailed). With AT-2 treatment the infectivity of viruses was totally eliminated while preserving the Env structure and functions [72] [73]. The production and purification procedures of the viruses are described elsewhere [72]. MAb KT11Fab was generously supplied from Peter Kwong's lab, NIH Vaccine Research Center. Unconjugated 10 nm colloidal gold nanoparticles were purchased from BBI Solutions Cardiff, UK.

Sample Preparation: The purified viral suspensions of 10 µl (~1.5-2 mg/ml total protein) were incubated at room temperature (20°C-25°C) for 30 mins in the presence of the Fab fragment of the KT11 antibody. Ligands were added at a concentration corresponding to an estimated tenfold molar ratio with Env trimers. Samples were then mixed with 10-nm colloidal gold (used for better tracking during tilt series collection) and 3.5 µl sample placed on a 200 mesh R2/1 Quantifoil grid (Quantifoil Micro Tools GmbH, Jena, Germany). Excess liquid was blotted with filter paper from both sides of the grid to form a thin layer of buffer which was then rapidly frozen by plunging the grid in a liquid/solid ethane slush (about -180 °C) and this procedure was done using a FEI Vitrobot Mk IV (FEI, Hillsboro, OR) in conditions of 100% humidity at 4°C. The grids were either transferred to a cryo-holder for EM examination or stored in liquid nitrogen for later use.

4.2.2 Tomographic Data Collection:

SIV-short tailed frozen virus specimens were examined at liquid nitrogen temperatures under low dose conditions using an FEI Titan Krios cryo-electron microscope (FEI, Hillsboro, OR) equipped with a Gatan Tridiem 863 UHS imaging filter with 2k X 2k CCD camera operated in the zero-energy-loss mode with a slit width of 20 eV. The microscope was operated at 300 kV and a magnification of x26,000, resulting in an effective pixel size of 5.4Å. Tilt series were collected at a defocus of -6 μm (underfocus) with a cumulative dose of ~100 e^-/Å. The angular range of the tilt series was from -65° to +65° consisted of 131 images recorded at a fixed tilt increments of 1° using FEI tomography software in automatic batch mode.

4.2.3 Tilt Series Alignment

The tilt series were aligned with the "PROTOMO" software package [18] using marker free alignment and the final maps were computed with the general weighted back projection algorithm [31] that is intrinsic within PROTOMO. Of the 55 tilt series, only 3 were considered of sufficient quality for further analysis.

4.2.4 Subvolume Processing

We used the I3 software for the subvolume processing [18]. More specifically, we used I3 as a black box for computing the initial orientation of the subvolumes using the position of the generated points and the center of the virion. Multivariate data analysis (MDA) package from I3 was used for generating the factors useful for classifying the subvolumes. Finally, we also used the alignment and classification functionalities provided by I3 for aligning the subvolumes and classifying them into different classes**.** All the data analysis reported in this study were carried

out on a desktop computer running Linux Mint on an 8 core AMD Fx processor having 16 GB RAM.

4.3 Approach

In this study, we have developed a semi-automated spike selection technique to segment out the virus spikes from the surface of SIVs. The segmentation approach consists of several steps (Fig. 4.4) – starting from generating a densely populated point cage surrounding the virion surface, classifying the raw subvolumes, aligning the cluster averages against a featureless reference and finally separating the raw subvolumes from the clusters having spikes and those clusters having no spikes at all.

4.3.1 Point Cage Generation

In the tomograms under study, the underlying shape of most virions was either ellipsoidal (Fig. 4.5A) or spheroidal (Fig. 4.5D) but some virions had an irregular shape (Fig. 4.5B and 4.5C). For this study, we used both regular and irregular shaped virions and selected particles using our segmentation technique described in this paper. Spike selection began by identifying four points ~90° apart at axial positions approximating the virion equator. The depth of each virion was estimated by selecting two additional points in the polar (top and bottom) slices respectively. Since the membrane is not visible at the top and bottom, the spikes themselves guide this process. Identification of these six positions was the only manual operation involved in this segmentation process. The selected positions were used to determine the axes and center of each ellipsoidal shaped virion**.** Using the estimated values, we generate a set of uniformly distributed points covering the entire surface (Fig. 4.5) of the virion at approximately the radial position of the "spike heads." The generated points are initially separated by a distance equivalent to the

"spike head radius" to ensure that the generated point cage was dense enough to select all the spikes at least once. Individual spike heads have a diameter of ~11 nm [9] [74]. For an unbinned tomogram with pixel size 0.54 nm, the spike head diameter was 20 pixels, and for a once binned tomogram the spike head diameter is 10 pixels with pixel size 1.08 nm. Coordinates within the point cage were spaced 5 pixels apart in both latitude and longitude. To examine the tradeoff between spike redundancies and identification accuracy, the subvolume size and point spacing can be parametrized so that points can be generated at any density desired.

4.3.2 Segmentation by Classification

In the SIV short-tailed virus, each virion contains ~70-80 spikes. The point cage algorithm automatically generated ~900 points per virion, which is approximately ten times the actual number of spikes present in a virion. Each virion was analyzed separately. From a once binned tomogram of the SIV virions, subvolumes of size 32x32x32 were cut centered on each of the positions in the point cage. For the spheroidal shaped SIV virions, each point in the point cage and the center of the virion were used to generate an initial transformation [18] to align all of the selected subvolumes along a radial line connecting the virion center and the individual cage points. For the ellipsoidal shaped SIV virions, the orientation of the subvolumes, i.e. the radial vector, was initially determined using the centroid of the ellipsoidal surface and the transformation applied to orient the spike axis along the z axis of the subvolume. After applying the initial orientation, the subvolumes were aligned by translation only and subjected to multivariate data analysis (MDA) and hierarchical ascendant classification (HAC). For aligning and classifying subvolumes, we followed the strategy called "alignment by classification" [18] [75]. With this strategy, different orientation of the set of objects are identified by classification and the class averages were aligned with respect to each other instead of the raw subvolumes.

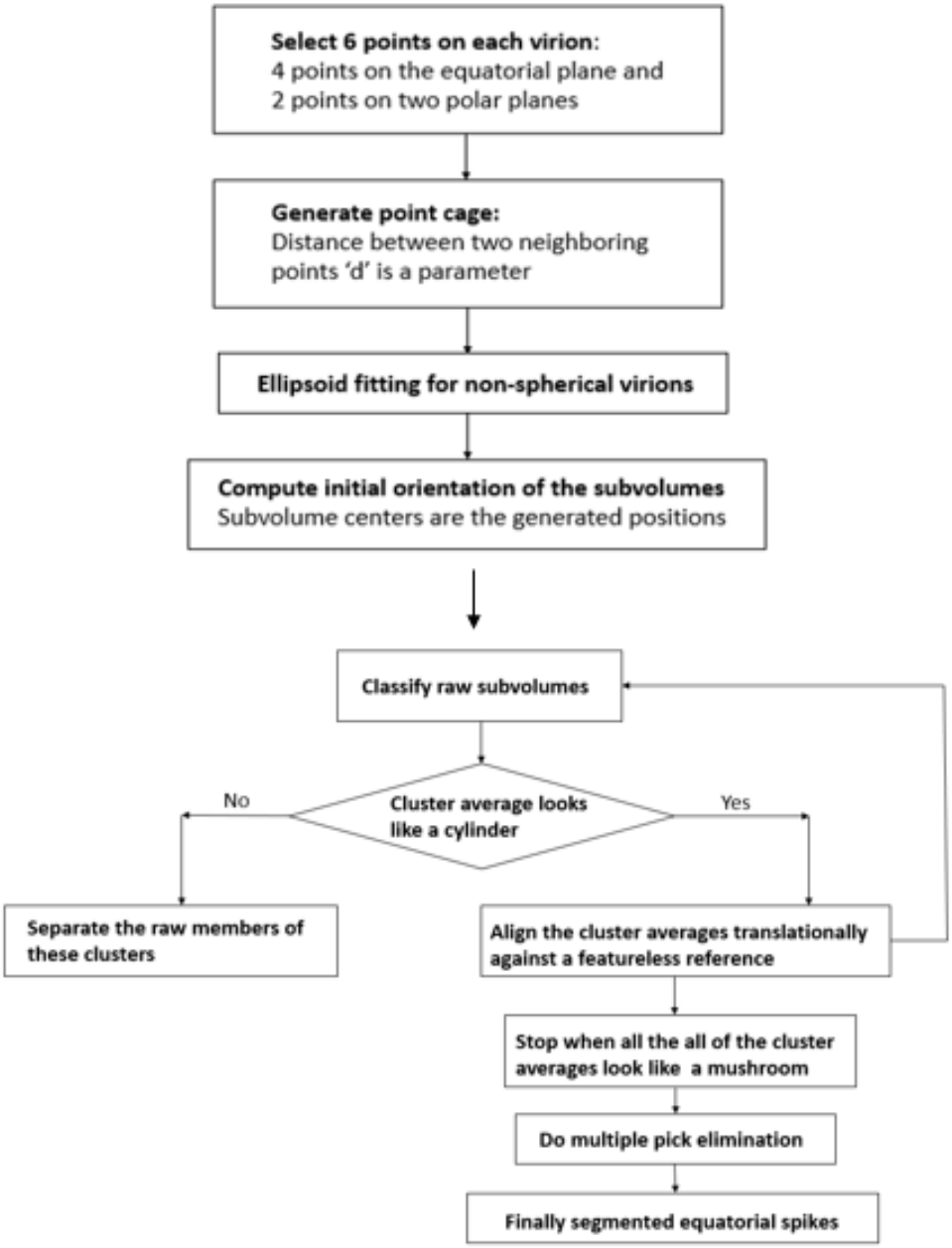

Figure 4.4: A schematic interpretation of the workflow of "segmentation by classification" for segmenting out equatorial SIV spikes. In the first step, a point cage is generated using 6 manually selected positions from individual virions. Subvolumes are cut centering each of the generated positions. In the second step, raw subvolumes are subjected to classification and the cluster averages are evaluated. Clusters with a "mushroom" shaped average are aligned with a featureless reference only translationally. Clusters having averages with just a naked membrane are separated and kept aside to generate averages for intervening membrane. The process is repeated for several cycles until the method converges.

The initial alignment of the class averages only involved translational alignment with respect to an external featureless reference and no rotational alignment was involved. The resulting alignment transformations were then applied to the raw subvolumes, and a new cycle of the iterative procedure is started.

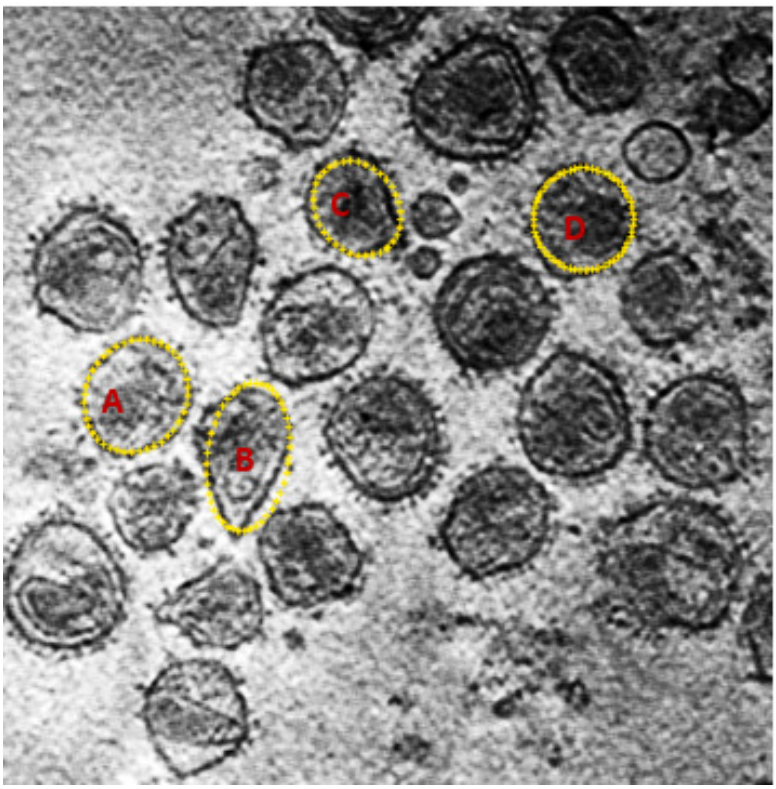

Figure 4.5: One slice of the denoised tomogram showing automatically generated points covering the virion surfaces. Denoising has been done using median filtering to enhance the contrast. Some of the virions are spherical in shape and the distribution of the generated points over a single slice looks like a circle. For the ellipsoid shaped virions, distribution of generated points over a single slice looks like an ellipse. For the irregular shaped virions, the underlying shape is estimated with an ellipse and the initial generated points try to capture the curvature of the membrane as correctly as possible. Subvolumes are generated centering each of these positions and are classified into clusters of any number of user's choice. During alignment, cluster averages are allowed to move relative to the reference in all the three directions by an amount sufficient enough, so that points initially positioning away from the real membrane of the irregular shaped virions, approach the membrane and capture its actual curvature within few iterations.

The signal-to-noise ratio of the SIV tomograms was very low resulting in low contrast between the protein and the background. To get better contrast of the virion contours, as well as the spike definitions, the tomograms were denoised using median filtering [76]. Median filtering improved the contrast by sacrificing the structural details. In this particular study, contrast enhancement initially helped the user to identify the spike and non-spike clusters. In the initial cycle, prior to any alignment of the subvolumes, the raw subvolumes from the denoised data were subjected to classification using the hierarchical ascendant algorithm (HAC) [77]. A rectangular shaped binary mask sufficiently large to contain all of the spike head and a small amount of membrane

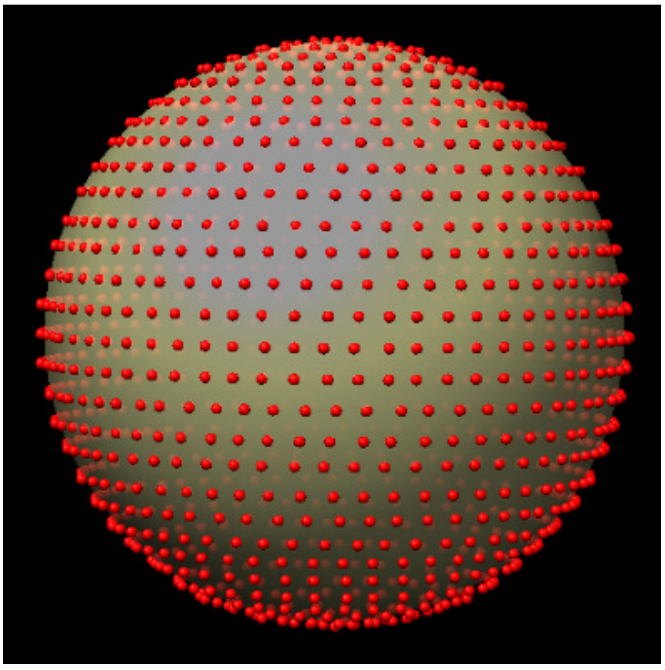

Figure 4.6: Plot of the generated positions. Points are arranged on a regular lattice with a separation of 5 pixels along latitudinal circles and the circles of equidistant points are stacked along longitudinal axis generating the sphere of regularly spaced positions

was used for classification. MDA was required for the pattern classification of the raw subvolumes. MDA is an iterative process and is carried out in each of the successive alignment iterations. In the initial cycle, because of oversampling of the generated positions, approximately,

22,394 raw subvolumes from tomogram 1 were classified into 50 and 100 clusters. Each cluster consists of a large number of raw subvolumes. More precisely, the expected number of raw subvolumes in each class is 448 with a standard deviation of 62, the mean and standard deviations being calculated over 50 clusters. For 100 clusters, the expected number of raw subvolumes is 224 with a standard deviation of 40. In the initial cycle, without any alignment, the raw subvolumes were subjected to classification. Only translational alignment of the cluster averages was carried out for centering up the off centered cluster averages. Only the cluster averages and not the raw subvolumes were aligned with respect to the external reference to avoid reference bias. A featureless reference image (side view Fig: 4.7A & top view Fig: 4.7B) was generated by a combination of both Gaussian sphere and Gaussian cylinder. Combination of two soft Gaussian spheres of correct radii were used to generate the featureless membrane and a soft Gaussian cylinder of correct width and height was added to the center of the membrane to give the reference image a shape of a mushroom which resembles the structure of viral spike protruding out the viral membrane. The alignment mask for aligning the cluster averages with the mushroom shaped reference was a rectangular mask with soft edge including the spike and a small portion of the membrane (Fig: 4.7F). Except for the initial six cycles, the classification mask, on the other hand, was a binary mask including only the spike and not the membrane (Fig: 4.7D) to avoid the bias that separates spikes in the equatorial region (protruding out the sides of the virions) having a prominent membrane density from those of the spikes in the polar region (extends out from the top) where the membrane density is very poor or the membrane is not visible at all. For the later cycles, a more sophisticated classification mask (Fig. 4.7E) was used.

In the initial cycle, with only classifying the raw subvolumes without alignment, most of the class averages lacked any spikes because of the poor alignment (Fig. 4.8A). After an initial

alignment using just translation, against a simplified reference (Fig. 4.7A), class averages showing spikes began to appear (Fig. 4.8B). After four cycles of "alignment by classification", spike definition improved and class averages showing only membrane became better defined (Fig. 4.8C). The original set of uniformly distributed data points showed concentration at spike coordinates and some migrated away from the correct radius. To improve accuracy in identifying spikes, we varied the alignment and classification masks. Subvolumes that contained spikes near their center were identified using MDA.

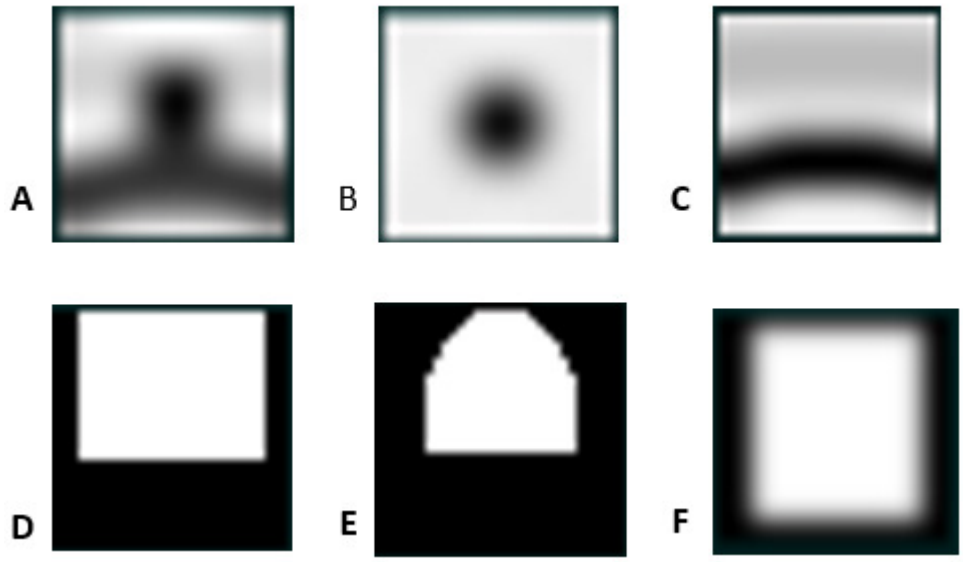

Figure 4.7: The reference image was used for alignment and different masks were used for classification and alignment. A) Side view of the mushroom shaped reference image used for translational alignment. The reference image was created by combining two Gaussian spheres of correct diameter with a Gaussian cylinder. B) Top view of the mushroom shaped reference that looks like a blob when looked down the z-axis. C) Side view of the reference image containing only membrane used for polar alignment. Two Gaussian spheres were used to make the reference membrane of proper width. D) Rectangular shaped classification mask excluding the membrane. E) A sophisticated classification mask excluding the membrane. F) Rectangular alignment mask with soft edge. Alignment mask is big enough to include membrane.

The cluster averages were band-pass filtered and aligned translationally with respect to the featureless reference. Translational alignment was achieved by locating the position of the correlation peak. For the first ten cycles, each class average was allowed to move sufficiently in all three directions to find the best correlation value, and then when the membrane became well aligned, their movement was restricted only in the radial direction for the later cycles. After the first alignment cycle, the cluster averages showed mushroom-like density centered up for very few classes and a significant number of classes lacked any spike density in the middle (Fig: 4.8B), showing pure membrane as the cluster average. The output of an alignment cycle contains the changes in orientation obtained for each of the class averages. These changes are stored and subsequently combined with the parameters already stored for the members of each class. This resulted in a new set of parameters stored again for use in the next cycle. After the first four cycles of the described strategy, the cluster averages showed two significant shapes, mushroom like density in the middle and pure membrane without having any density in the upper part of the membrane. At this point, a manual inspection of the shapes of the cluster averages are performed and, classes with pure membranes were separated out and saved for generating averages for intervening membranes. The class averages having a mushroom shape, were subjected to the translational alignment with respect to the featureless reference. This process was then repeated for another six cycles with the median filtered map until all the cluster averages showed clear spike densities (Fig: 4.8C).

As a natural consequence of the median filtering, the intrinsic details of the spikes structures were blurred significantly. Consequently, even after 10 iterations with the median filtered tomogram, the spike density in each class average improved but remained featureless. At this point, the unfiltered (un-denoised) tomograms were used instead of the median filtered ones and

the same technique was repeated on the unfiltered maps for 20 more cycles (Fig: 4.8D & 4.8E). At the end of each iteration the non-spike classes were separated out from those showing spikes and kept aside for future analysis.

Examination of the class averages, shows that after translational alignment, the angular alignment of the membrane, and consequently the spike, is poor. To correct this, we carry out two cycles of polar alignment utilizing only the first two Euler angles against a synthetic reference having the shape of a curved membrane (Fig. 4.7C). The cross-correlation peaks were restricted to radial movements of only a few pixels. The polar alignment corrects the radial distortion of the subvolumes. During the polar alignment, the classification mask only includes the membrane part of the subvolume and the alignment reference was generated using two Gaussian spheres and with apodized edges.

To get the better-defined spike density, five to eight (depending on the quality of the tomogram) multi-reference spin alignment cycles were carried out and the spin alignment cycles clearly improved the segmentation as the raw subvolumes were nicely classified into spike and non-spike classes after alignment about the polar i.e. the z-axis. It must be noted that this spin alignment is distinctly different from the spin alignment carried out later for revealing the structure of the segmented spikes. This was coarser and was required in order to reveal the classes distinctively and was a result of the quality of the data used for the segmentation. In theory, the segmentation step should require only translational alignment. The spin alignment being applied only in special cases where the translational alignment fails to accurately define the classes because of the noise distribution of the data.

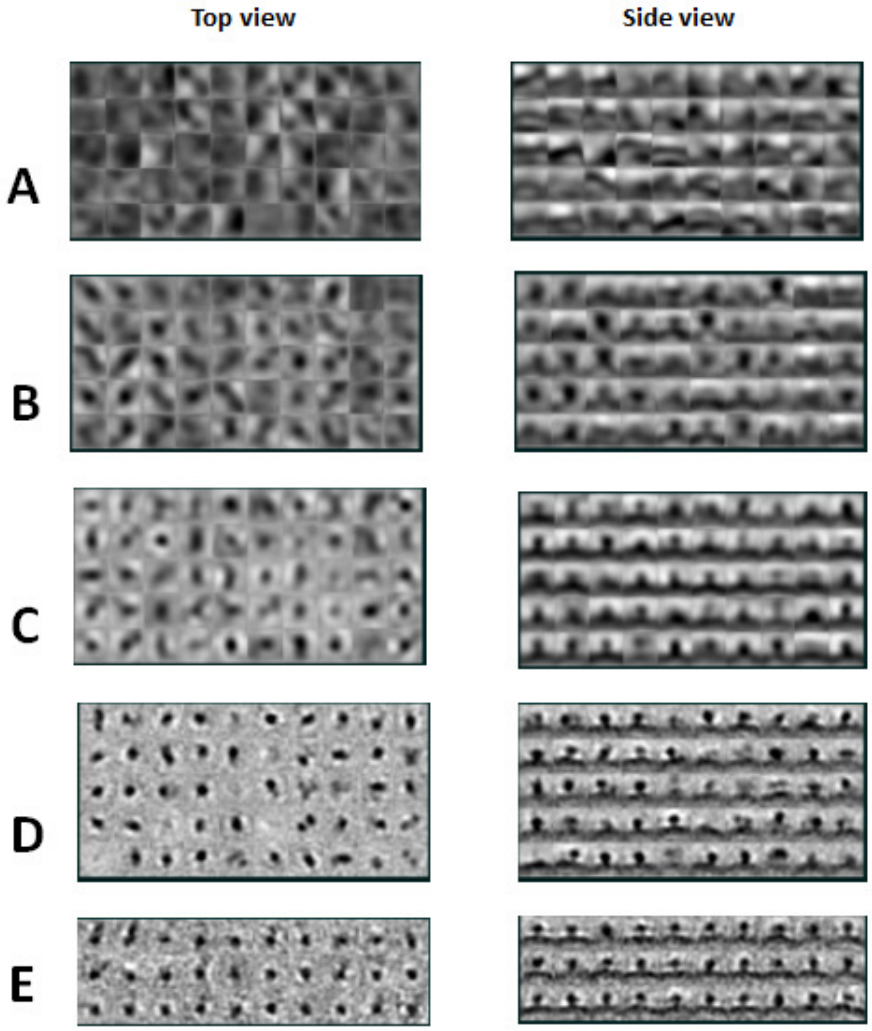

Figure 4.8: Class averages showing the "segmentation by classification" approach. A) Class averages of cycle 0 showing only blurry membranes when the raw subvolumes from a median filtered tomogram were classified into 50 classes. Because of the median filtering, all the intrinsic detail of the membrane structure was blurred and due to lack of alignment spikes were poorly resolved. B) After translational alignment of the class averages from cycle 0 with respect to the featureless mushroom-shaped reference, the raw subvolumes were reclassified. The spike density improved significantly with some classes showing spike-like density and other averages showing naked membrane. C) Class averages of cycle 9. From this point on unfiltered maps were used for better definition of the cluster averages. D) In the first cycle using the unfiltered map, the cluster average showed improved clarity in the spike density; most of the classes showed spike-like averages and very few classes showed pure membrane. Each spike class average had ~230 members; classes showing only membranes had ~400 members. The top view of the non-spike classes showed no significant spike density and the side view showed pure membrane. After manual inspection, non-spike classes were separated from the spike-like classes. E) Class averages for cycle 30 showed spikes in all the 30 classes. In this cycle, the average number of members in each class was 56 with a standard deviation of 12 and the segmentation process was considered converged.

4.3.3 Identification of the Polar Spikes

The method described above captured only the equatorial spikes. To capture the polar spikes, we did a distance analysis of the already segmented equatorial spikes. We measured the radial movement of each of the segmented spike from its original generated position (in the initial cycle). We computed the mean (μ) and standard deviation (σ) of the radial movements of the equatorial spikes. From the originally generated positions, we separated out all the polar positions by simply identifying a point as a polar position if the point has a latitude greater or equal to 45° with respect to the center. Now, each of these identified polar positions may or may not contain spikes or even they may contain a spike only partially. In either of the cases, the membrane is not visible because of the missing wedge but the spike head will be clearly visible. In the initial cycle, we classified the raw subvolumes and the class averages were translationally aligned with respect to a featureless cylinder having no membrane. For the cross-correlation peak search, the subvolumes were allowed to move 9 pixels each in x and y direction and most importantly move independently with 3 different distances along the radial direction: $\mu \pm \sigma$, $\mu \pm 2\sigma$ and $\mu \pm 3\sigma$ respectively (Fig: 4.10). For tomogram1, after continuing two cycles with radial movement of $\mu \pm 2\sigma$, the classification factors started showing clear features indicating polar spikes and classification selecting those factors clearly showed classes having some blob indicating some object (may or may not be spikes) protruding out from the membrane and classes that don't contain anything except a smeared image of the membrane. We continued this analysis for tomogram 2 and found that $\mu \pm \sigma$ amount of radial movement gave the best result in terms of identifying polar spikes. For both tomograms, we separated out the classes having blobs as the average and continued multireference spin alignment to get the shape of the spikes as well as separate out any non-spike element that might be sticking out radially from the invisible

membrane. After continuing 10 spin alignment cycles with a binned tomogram and continuing another 10 spin alignment cycles with an unbinned tomogram, we could capture polar spikes with an excellent accuracy (Fig: 4.11).

In the initial cycles, to get the best cross-correlation peak, the subvolumes were allowed to move in all the three directions. As a result, some of the subvolumes moved radially further away from their initial position while cross-correlated with the reference. Migration along the radial direction, eventually, captured the virion contour correctly over the first 6 to 8 iterations. For the irregular shaped virions, the generated point cages were just a rough estimation of the underlying shapes. Within first 10 iterations, the underlying shapes of these virions were captured almost accurately (Fig: 4.12). A high amount of radial movement caused some subvolumes to move either inside or outside the virion envelope, leaving some outliers. Outlier detection and hence elimination is an essential part of this method.

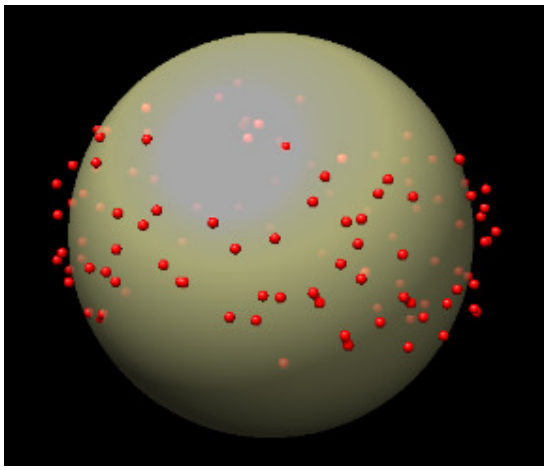

Figure 4.9: 3D plot of spike distribution on one virion after 30 cycles of alignment and classification. After several cycles, due to subvolume alignment the initially generated spherical or ellipsoidal point cages change in shape and eventually capture actual contours of the virions. Hence, initial regular lattice of generated positions become a distorted point cage over the cycles and the remaining points represent the center of those subvolumes that contain spikes.

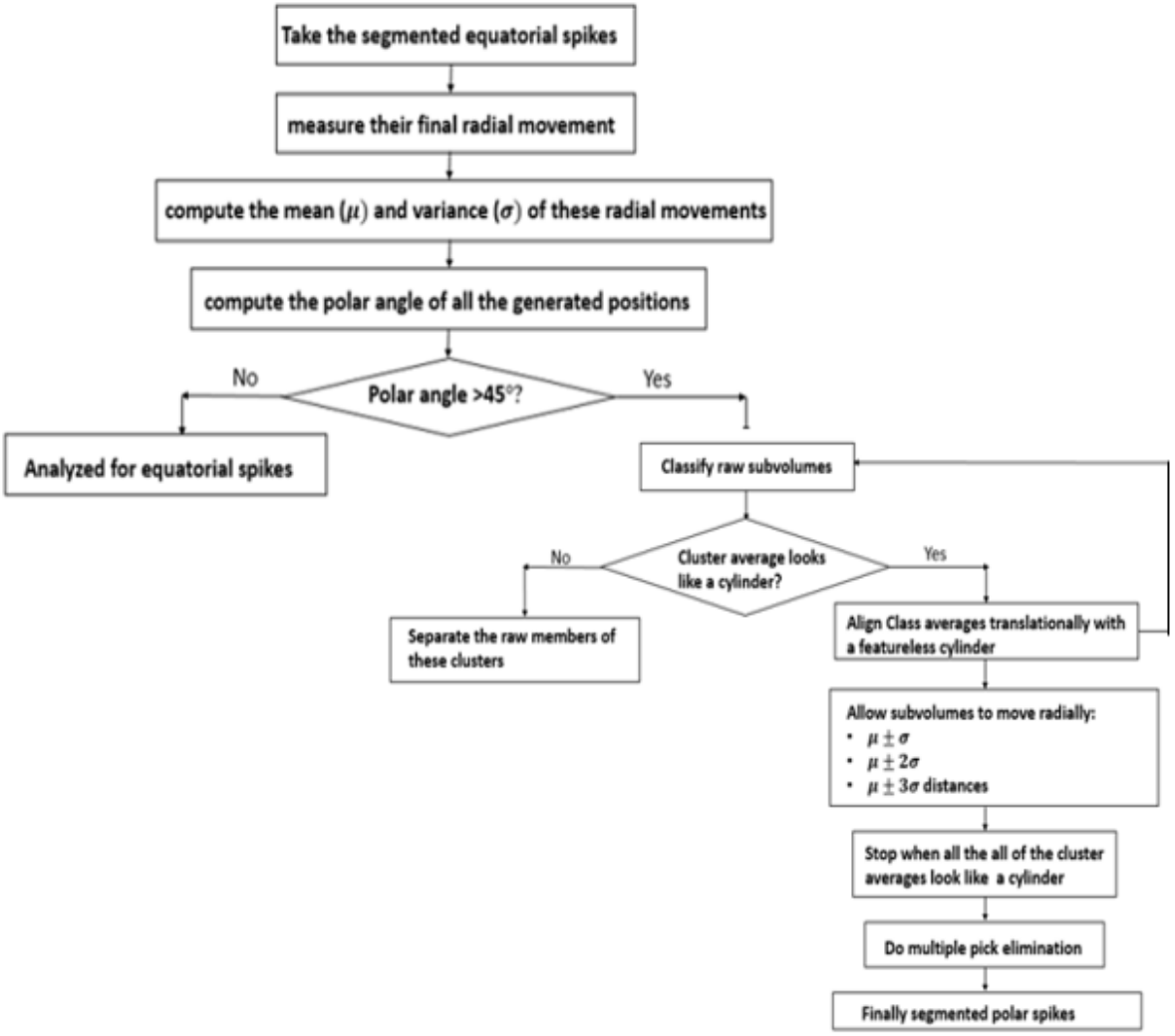

Figure 4.10: A schematic interpretation of the workflow of "segmentation by classification" for segmenting out polar SIV spikes. In the first step, the radial movement of the finally segmented equatorial spikes are measured and their mean (μ) and standard deviation (σ) are computed. In the second step, polar subvolumes are identified. For the polar subvolumes, classification is done with the raw subvolumes and the class averages are translationally aligned with a featureless Gaussian cylinder. The subvolumes are allowed to move radially in three different distances ($\mu \pm \sigma$, $\mu \pm 2\sigma$ and $\mu \pm 3\sigma$) to find the best cross-correlation peak. The best distance is varied for different tomograms. The algorithm is stopped when class averages for all the classes showed a spike in the middle. Multiple pick elimination is performed to eliminate the redundant spikes.

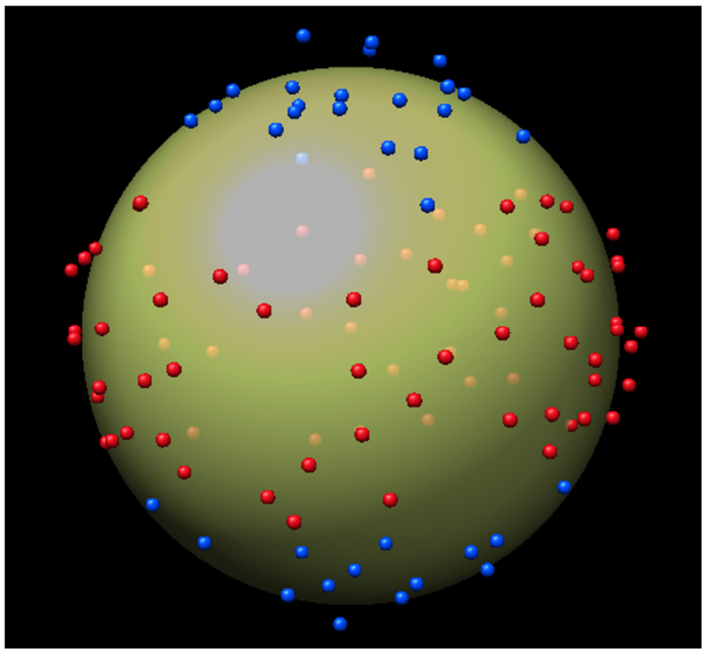

Figure 4.11: 3D plot of polar and equatorial spikes shown for one single virion, segmented using "segmentation by classification". Polar spikes are shown in blue and the equatorial spikes are shown in red. As the spikes are distributed over the entire surface of the virion, hence spikes, that were identified on the other side of the membrane, are visible (in defused red and blue color) inside of the translucent envelope.

In every iteration, clusters showing pure membrane as the average are separated resulting in a significant decrease in the number of raw subvolumes and leaving only those subvolumes that were centered on real spikes. The segmentation progress is depicted in Fig: 4.15. Finally, after ~30 iterations of segmentation by classification, almost 90% of the resultant subvolumes contained spikes in the center. Multiple selection of the same spike is an obvious outcome of the oversampling. Consequently, a multiple pick elimination algorithm was developed to ensure that each spike is selected only once and was performed at the end of the final cycle of the segmentation process.

4.3.4 Post Segmentation Subvolume Analysis

After segmentation using the once binned tomograms, the segmented subvolumes were further analyzed using the unbinned tomogram and the subvolume size was increased to 64x64x64 pixels. In each individual unbinned tomogram, another 15 cycles of multi-reference spin alignment was performed with the angular search range of ±180°. We varied the angular step from coarse (10°) for the first 8 cycles to fine (5°) for the next 7 cycles. Eventually, these spin cycles revealed the trimeric structure of the virus spikes.

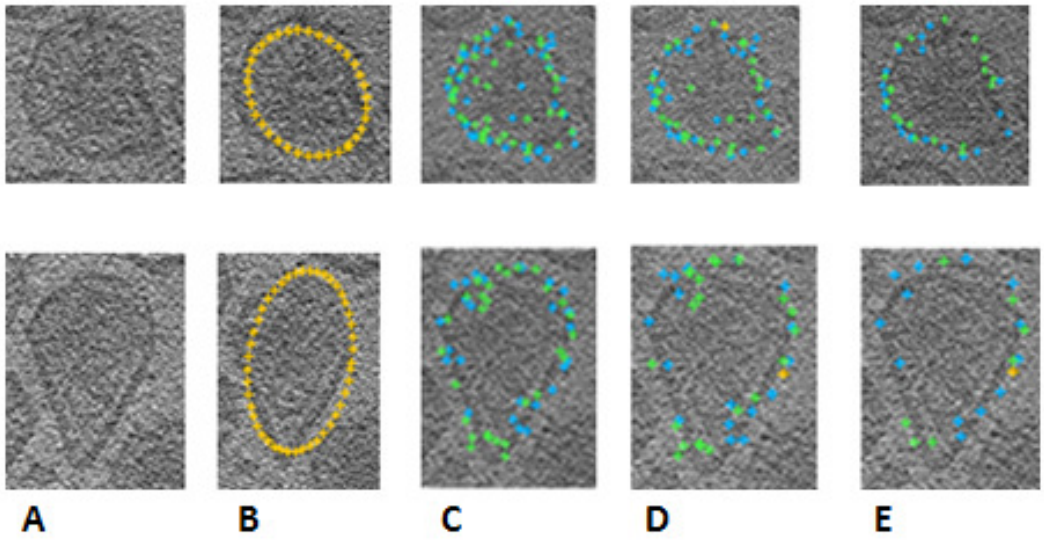

Figure 4.12: Two different virions in both upper and lower panels show the automatic membrane tracking on a particular image plane over the cycles. A) The virions are irregular in shape and B) initially generated point cages were ellipsoidal for both of them. C &D) Intermediate cycles E) After 10 cycles the contours are captured almost perfectly. Blue, yellow and green are showing above, current and below z levels when panned through the tomogram. The point density is reduced as the subvolumes lacking spikes are eliminated over the cycles.

Multi-reference spin alignment for 10 more cycles were carried out on the segmented subvolumes from each of the individual tomograms. At this point in the processing, the trimeric shape of the spikes began to show but was not clearly visible because of the fact that each

individual tomogram did not have enough particles to reveal the structure. Hence, we combined the segmented subvolumes from tomogram-1 and tomogram-2 for further processing to obtain the spike structure. Subvolumes from tomogram 3 were not included for further processing because the subtomogram average of the segmented subvolumes from the unbinned map of tomogram-3 became very noisy with poor resolution. Hence, the segmented subvolumes only from tomogram-1 and tomogram-2 were combined and the raw subvolumes were classified into 10 and 20 clusters (Fig: 4.13A), respectively. The class averages were polar aligned within an angular range of ±45° with angular intervals 5°. One polar alignment was sufficient to get good alignment of the membranes but to reveal the intrinsic details of the spikes, several iterations of spin alignment with finer angular steps were necessary. The cluster averages were aligned with each other by multi-reference spin alignment using a grid search about the polar (i.e. z-axis) and were classified into 20 clusters. During spin alignment, the cross correlation peak search was allowed to move only along the *x* and *y* directions and restricted from moving radially. After 20 spin alignment cycles, though most of the cluster averages were clearly trimeric in shape when viewed along the *z*-axis, some lacked any clear shape (Fig: 4.13B)). At this point, multi reference spin alignment was performed using a subset of cluster averages showing good rotational symmetry. The majority of cluster averages after 20 more spin cycles had a clear trimeric shape from top view (Fig: 4.13C & 4.13D). Very few cluster averages, lacked a trimeric shape after spin cycles and were identified as the outliers and were excluded from further spike analysis. However, this subset of "spikes" could be either cellular membrane spikes or incompletely formed Env spikes. These outliers are separated from the subvolumes that are showing clear rotational symmetry in the spike structure and they are classified again to observe the structural variability amongst them.

4.4 Experimental Results

4.4.1 Identification and Localization of the Envelope Spikes

Our approach is designed to identify spikes more accurately than manual picking while at the same time not biasing the selection toward a particular spike structure. We tested the procedure using three different tomograms of an unstained frozen-hydrated suspension containing an average of 20 individual SIV virions per tomogram (Fig. 4.14). The automatically generated point cage totaled 22,394 coordinates for 23 virions in tomogram 1, 16,886 for 19 virions in tomogram 2 and 18,587 for 22 virions in the tomogram 3. Points were generated at spacing of 5 pixels along the circumference of the horizontal circles (latitudes) which are in turn stacked up along the entire volume of the virion at a separation of 5 pixels (Fig. 4.6). As we move from the equatorial plane to the polar planes of the virion sphere, the circumference of the latitudes decreases, as so does the number of points on each latitude.

At the beginning, before any alignment, many class averages lacked any spikes because the initial coordinates were not obtained with reference to the spikes themselves (Fig. 4.8A). Many coordinates were centered on the intervening membrane. After an initial translational alignment of the class averages was done against a simplified (featureless) mushroom shaped reference, class averages showing spikes began to appear (Fig. 4.8B). After multiple cycles of alignment by classification, spike definition improved and class averages showing only membrane became better defined (Fig. 4.8C). After few alignment iterations, some points in the original set of uniformly distributed data points migrated away from their initial position and concentrated around the actual spikes. The class averages were observed from two different views, the side view and the top view. In the earlier cycles, because of lack of intrinsic details, a class average

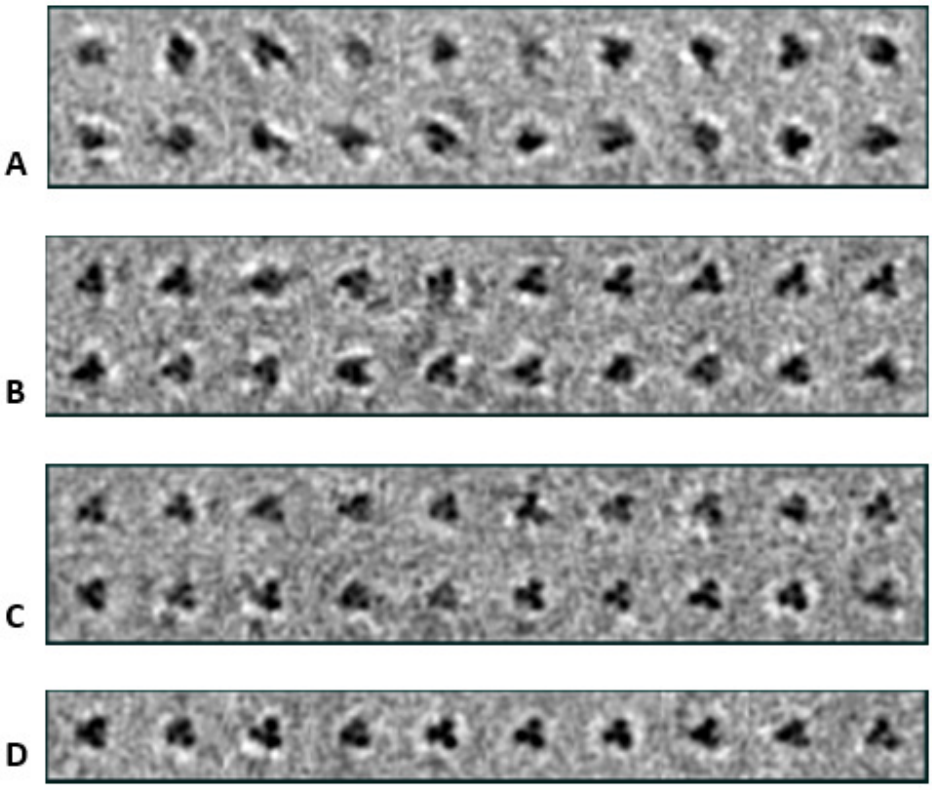

Figure 4.13: Class averages of the equatorial spikes showing 3-fold symmetry. A) Class averages from 20 classes from the first spin cycle after combining the segmented equatorial subvolumes from two tomograms. Class membership is ~150 on average. Some of the classes reveal the trimeric structure but some do not. B) Class averages after 20 more multi-reference spin cycles. Most of the class averages show better 3-fold symmetry but a few of them do not show a clear trimeric shape. C) Continued 20 more Spin alignment cycles against only a subset (those showing good 3-fold symmetry) of the class averages as a reference and discarding the outliers, almost all of the 20 cluster averages show good 3-fold symmetry. Class membership at this point is ~60 spikes per class. D) Class averages of 10 classes are showing better density as the class averages are ~120 per class.

showing spikes should look like a mushroom from the side and a circular blob from the top view. For classes lacking spike concentration, the class averages are expected to show only membrane from the side view and no blobs in the center were expected from the top view. The montage of the class averages was mainly examined from both the views and classes lacking spikes were separated. These “membrane” classes would be used later in reassembly of the virus envelope from class averages of the separate entities. This is an iterative process.

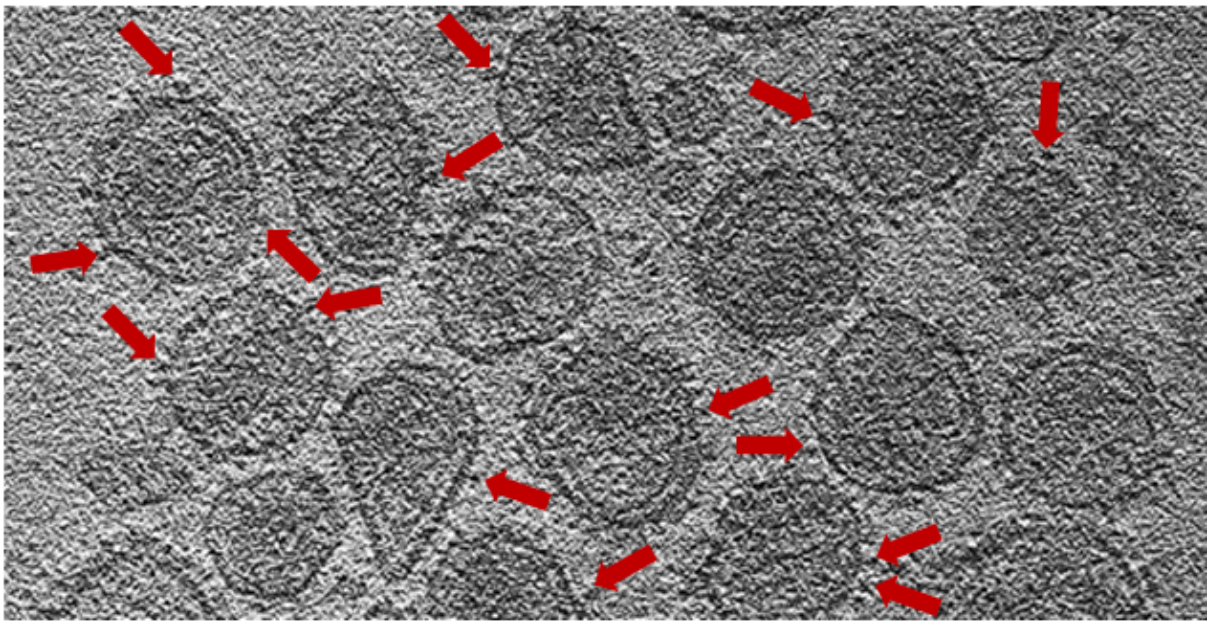

Figure 4.14: One slice from the tomogram-1 of SIV data. Red arrows are showing some of the possible spikes radially protruding from the virus membrane. The underlying shape of individual virions are either spherical or ellipsoidal. Some of the virions have no regular shape.

For all the three tomograms, the first few (4/5) cycles did not show a clear enough result to decide about whether a particular class is a spike or not. As the alignment improved, the spike definition improved as well. Hence, we could separate out classes where the class average looks like a spike from the classes where class average is just naked membrane showing no spikes at all.

The automatically generated points, at the very beginning, were dense enough to ensure that no spikes were missed. Consequently, spike redundancy is an obvious outcome of the oversampling. Spike redundancies were identified by the proximity of centers of the subvolumes on a virion by virion basis at two stages of the analysis - after translational alignment and at the end of the spin alignment cycles. Redundant spikes were eliminated rather than separated. Redundancy

elimination played an important role in downsizing the subvolume numbers and hence, in the segmentation of the spikes.

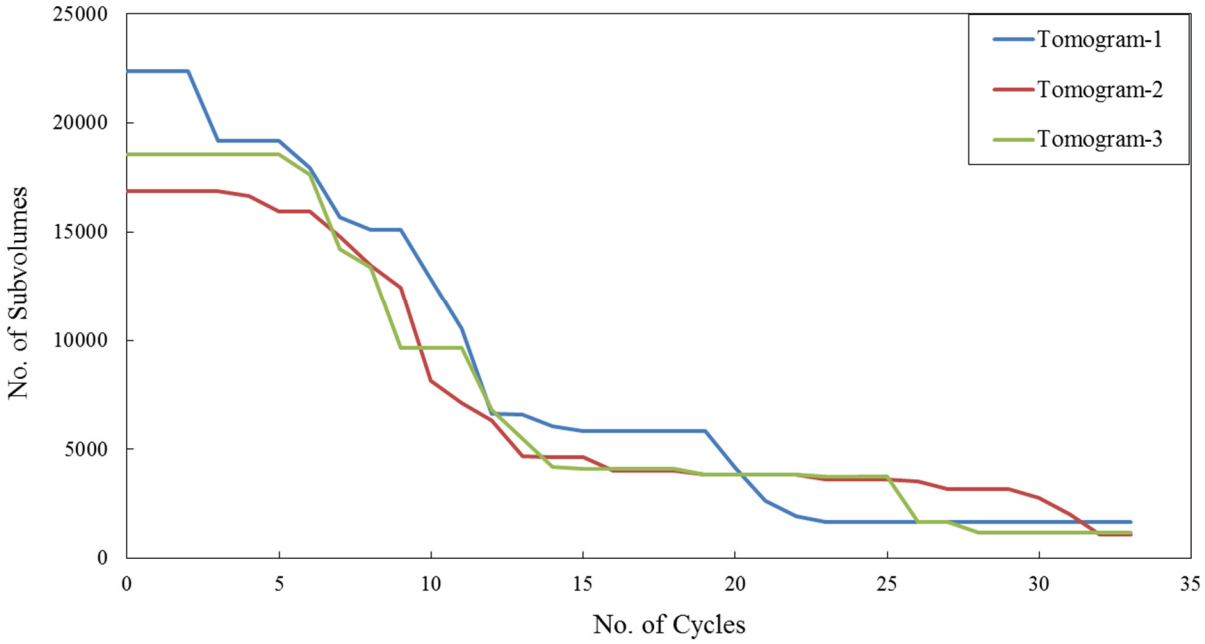

Figure 4.15: The Graphical representation of the "segmentation by classification" method for capturing the equatorial spikes. The method had been tested for three different tomograms containing ~20 virions per tomogram. The graph shows the reduction of the number of subvolumes as a function of cycles. In tomogram-1 initial number of generated positions were 22,394, for tomogram-2 total number of generated positions were 16,886 and for tomogram-3 it was 18,587. For each of the tomograms first 10 translational alignment cycles were performed on denoised tomogram (denoising were done using median filtering) where the class average was aligned with respect to a featureless reference. For further analysis, unfiltered tomograms were used. For next 15 cycles, the class averages were subjected to a translational alignment with the same featureless reference. For next 5 (or 7) cycles a multi-reference spin alignment were carried on improving the spike density in the class averages. At the end of cycle 30, the number of subvolumes were reduced to 1671 for tomogram-1, 1084 for tomogram-2 and 1162 for tomogram-3. The initial cycle took ~4 hours to complete because of the high number of subvolumes caused by the oversampling. Completion time for next 5 cycles were gradually decreased and after cycle 6 each iteration completed in less than an hour.

We cross validated the final segmented spike positions captured by this semi-automated segmentation technique separately, once only for the equatorial spikes and then for both equilateral and polar spikes. We compared the outcome of "segmentation by classification" with

manually picked equilateral positions from each of the three tomograms. For each manually identified position, we calculated the number of automatically selected spike locations within 16 pixels proximity. Our method missed only ~4% of the manually selected spike positions (Table 1). Further investigation showed that 98% of the missed spikes belonged to the irregular shaped virions. Our method also selected some extra spike locations, which were not captured by manual selection.

To cross validate the result of the semi-automated algorithm in combined equatorial and polar locations, we compared the outcome of "segmentation by classification" with manually picked positions from each of the two good quality tomograms from the aforementioned three tomograms. Again, for each manually identified position, we calculated the number of automatically selected spike locations within 16 pixels proximity. Our method missed only ~15% of the manually selected spike positions (Table 2). Our method also selected some extra locations, which were not captured by manual selection. Further investigation, proved that a significant percent, of these extra picks are real spikes which are in a distorted shape. Hence, these extra picks cannot be referred to as false positives.

After significant downsizing of the generated point cage in each individual tomogram, we combined the segmented subvolumes from tomogram 1 and 2 for further processing to obtain the spike structure. The total number of subvolumes combined from both the tomograms was 3230. Though at this point, the spin alignment cycles on individual unbinned maps started showing the 3-fold symmetry but the intrinsic details were not satisfactory probably because the total number of spikes in each individual tomogram were not large enough and as a result the number of members in each class were not sufficient to improve the signal-to-noise ratio to a certain level to reveal the intrinsic details of the trimeric structure.

Figure 4.16: The Graphical representation of the "segmentation by classification" method for capturing the polar spikes. For the first 3 cycles the number of sub-volumes did not reduce. In the fourth cycle for the tomograms a significant drop in the number of subvolumes occurred. Finally, the method converged in cycle 38.

After a few iterations of multi-reference spin alignment cycles with the combined data, the class average started revealing the inherent 3-fold symmetry of the spike structure. In each refinement cycle ~50% of the class averages showed clear trimetric shape from the top view. These class averages were used as the reference for the next iteration and the rest of the classes were aligned against these selected references.

Table 4.1: Segmentation accuracy for equatorial spikes

	Number of manually picked spikes	Number manually selected spike positions missed by the method (false negatives)	Number of extra spike positions localized by the method	Number of actual spikes captured by the method	% of the actual spikes captured by the method
Tomogram1	1231	29	440	1202	97.6%
Tomogram2	579	21	505	558	96.4%
Tomogram3	392	20	770	372	94.5%

Table 4.2: Segmentation accuracy for combined polar and equatorial spikes

	Number of manually picked spikes	Number manually selected spike positions missed by the method (false negatives)	Number of extra spike positions localized by the method	Number of actual spikes captured by the method	% of the actual spikes captured by the method
Tomogram1	1342	201	969	1141	85%
Tomogram2	757	116	848	641	84.7%

The process of alignment against selected references improved the quality of the trimetric shaped class averages. In addition, a few classes having an average with an unexplained shape became more prominent. These classes were separated from the classes showing spikes and kept aside for future analysis. Final redundancy elimination was done at this stage and for each subvolume. In this process, any neighboring subvolume within 10 nm proximity was eliminated. This confirmed that each spike had one and only one copy in the database.

4.4.2 Identification of The KT11 Antibody

The virus specimens used here, had been treated with the monoclonal antibody KT11. If the antibody had labeled a large fraction of the available gp120, it would have been evident in the segmented class averages. Since it was not visible, the labeled fraction of gp120 must have been low. Therefore, evidence of antibody (KT11) attach to the trimetric shaped Env spike head required a more refined classification procedure. We utilized a procedure tried several times earlier, which involved triplicating (making two additional copies) of the raw Env subvolumes and rotating one set 120˚ and the other 240˚ respectively about the spike axis [78] [79]. We then classified the subvolumes using a wedge-shaped mask that covers only a single spike arm. This improved the quality and 4 out of 10 of the class averages (Fig: 4.17A starting from the left and assuming the class numbering starts from 0, four classes, class 1, 3, 6 and 9 showed evidence of antibody) in the form of extra density of varying shape attached with one arm. The surface rendering of the class average from class 9 from both the top and the side view emphasizes the evidence of the antibody (Fig: 4.18 A & B).

In order to eliminate the possibility that the antibody could simply be a neighboring spike arm, we did a proximity analysis. From the triplicated data set, we did the analysis on each set (i.e. each of the three rotational sets) separately. As a first step of this analysis, centers of three arms of the global average were manually selected. Then centers of each spike arm in each of the raw subvolumes were computed by extrapolating the selected arm centers of the global average on the raw subvolumes. After computing the arm centers, we first identified the proximal spikes and then investigated further to identify the proximal arms. Towards this, we first identified whether the proximal spikes belong to the same virion or from two neighboring virions.

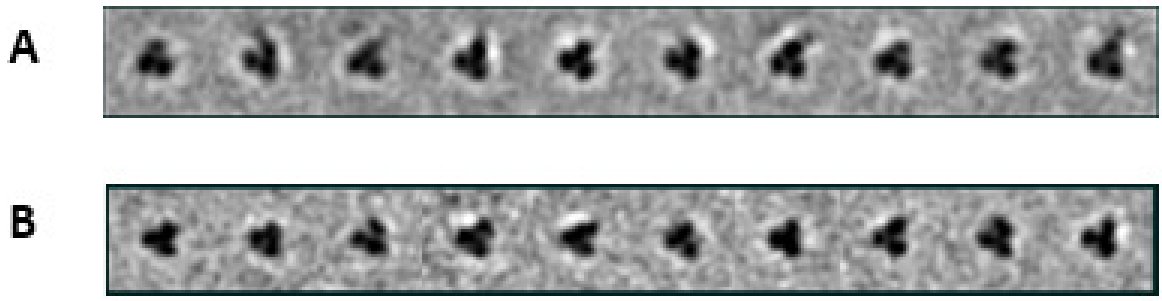

Figure 4.17: A) Top view of the class averages at the final cycle. Class 2,6 and 9 are showing antibody attached to one of the arms. B) Class averages after proximity analysis shows no extra density that could be identified as antibody KT11.

For both the cases, we identified the proximal arms among the three arms of the spikes and eliminated those spikes from further analysis and continued the same identification and elimination process from all the three rotational sets. After reclassifying the non-proximal spikes, we could not resolve the extended density (Fig: 4.17B) attached to one of the spike arms and because of this poor occupancy rate we could not conclude anything about the presence or position of the antibody KT11.

4.4.3 Analysis of Combined Polar and Equatorial Positions

After completing the semi-automated algorithm for the polar positions separately, we ended-up having few polar spike positions for both the tomograms. At this point, we combined the newly segmented polar positions (Fig. 4.11 blue dots) with the existing equatorial positions (Fig. 4.11 red dots), that we segmented out by analyzing only the equatorial positions. We continued few classification cycles with multireference alignment with these combined positions. Surprisingly, the class averages showed two distinct shapes – 1) T-shaped averaged with 2) trimeric-shaped averaged (Fig.4.19). At this point, we separated out the class members of the T-shaped classes from the members of the trimeric-shaped classes and analyzed them separately.

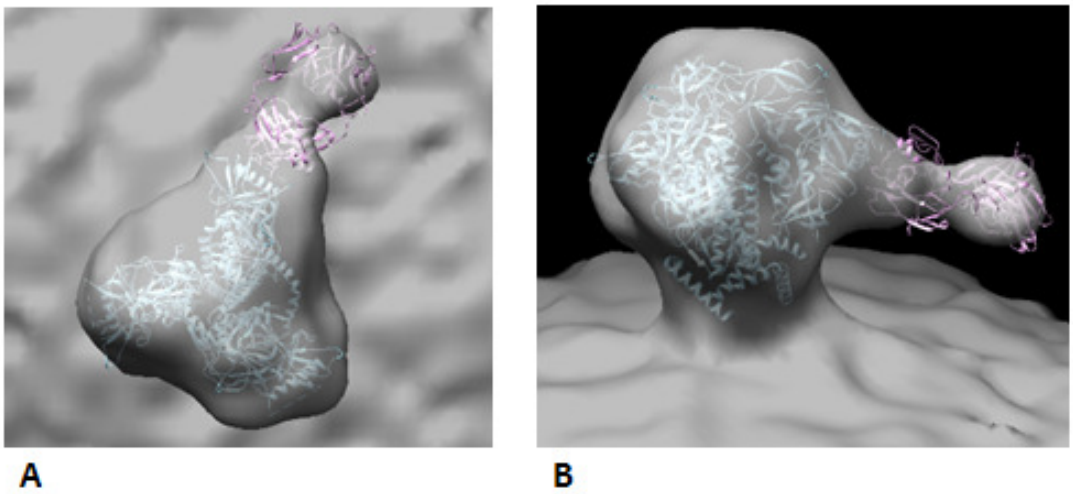

Figure 4.18: Evidence of extra density(KT11). A) Top view and B) side view of the class average of class 9 from figure 10. Clearly the class average shows evidence of extra density, which can be identified as antibody KT11.

In next few cycles, we separately classified the T-shaped subvolumes and aligned then with the global average. Similarly, the trimeric subvolumes were first classified and the class averages were subjected to multi-reference alignment. We performed the symmetrization on the trimeric-shaped subvolumes i.e. we triplicated the dataset by rotating each raw subvolume 120˚ and 240˚ respectively about the spike axis and then classified the subvolumes using a wedge-shaped mask that covers only a single spike arm (Fig. 4.20).

4.5 Discussion

For localizing macromolecular assemblies one of the most common methods used is template matching [47] [70] [80]. Template matching is a model-based approach of segmentation where templates derived from the high-resolution structure of the molecule under scrutiny are used to search the reconstructed volume to localize the object of interest. It must be noted that template matching can also be computationally intensive when applied in a “brute force” approach, for the

reason that the orientation of the particles will be random and, consequently, the whole angular range has to be scanned by rotating the templates and calculating the cross-correlation coefficient for all independent combinations of Eulerian angles. Another limitation of template matching is that it needs a proper template against which the search has to be implemented. But there are situations where such a template may not be available and the data may be heterogeneous which in turn reduces the efficiency of template matching. Thus, in such cases it is advisable to have methods for localizing macromolecular structures that do not depend on the existence of template. The localization of HIV/SIV envelope spikes is one such problem, the data is noisy with poor SNR and the structure is highly heterogeneous. As an alternative to template matching, our method uses classification to localize the macromolecular structures. The advantage of this approach is that it does not depend on a particular template and hence can localize highly heterogeneous macromolecular structures in tomograms even with very low SNR. Our method is not tied to the structure of a given template and hence can localize and find new structures/particles that may be present in the tomogram. This in turn avoids the bias introduced by the use of a template. It also reduces the chance of having false positives in situations where the noise has a distribution similar to that of the template. One of the limitations of our method is that initial cycles are computationally intensive because of oversampling and hence may take more time compared to a method based on the template matching. In the initial cycle, with almost 10 times over sampling, the multivariate data analysis and classification took about three hours on the average to finish and this was the longest part of the process. To speed up the initial cycles, we aligned the cluster averages only translationally with the reference and the alignment took about 30 minutes to complete. The initial spin alignment cycles were little longer (about an

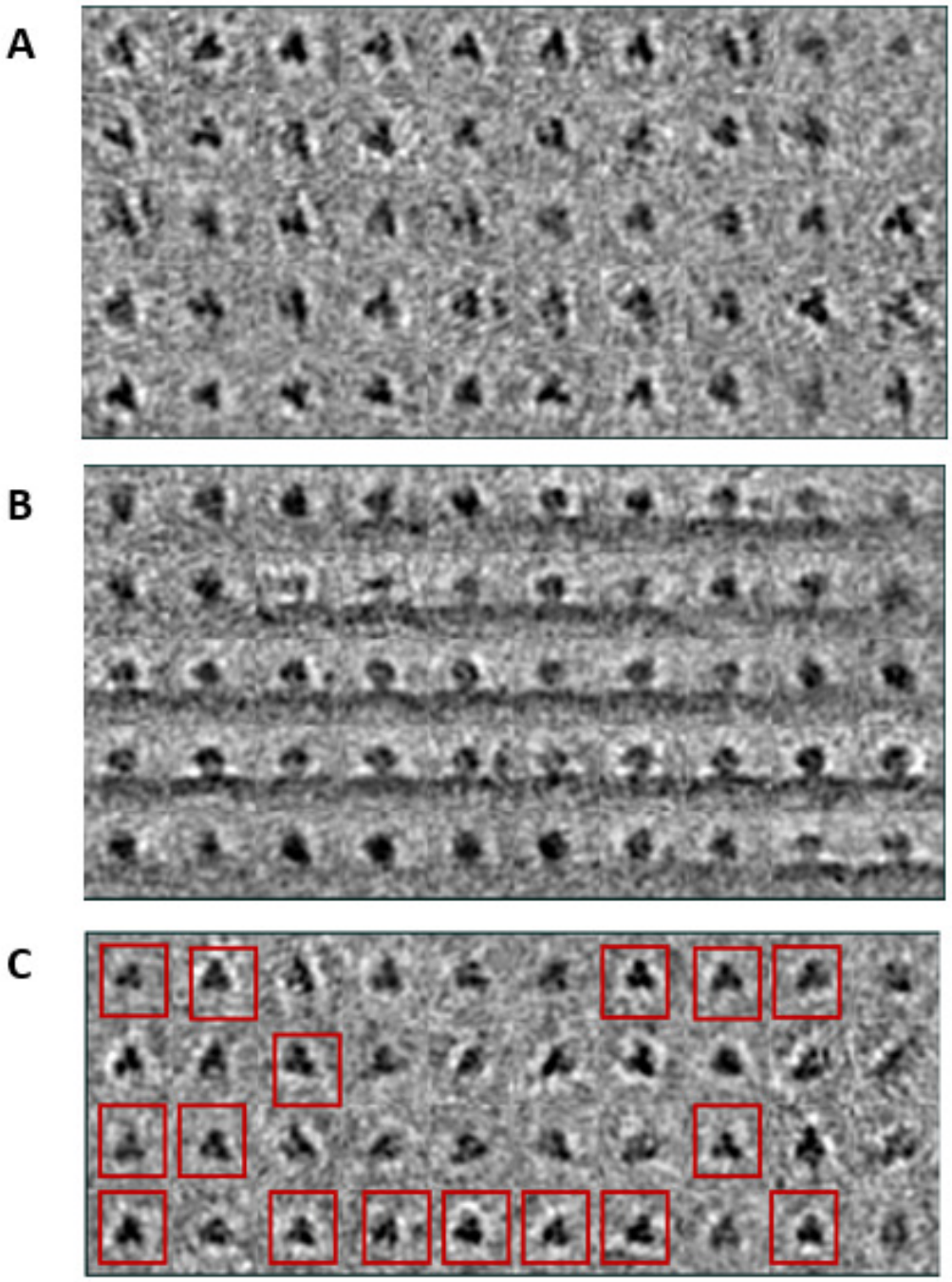

Figure 4.19 A) Top view of 50 classes with combined polar and equatorial spikes and B) side view of the class averages. Some class averages are showing trimeric shape while others are showing a clear T-shape. C) Ref boxes are showing the trimeric classes. Those class members were separated out from the T-shaped classes and analyzed separately.

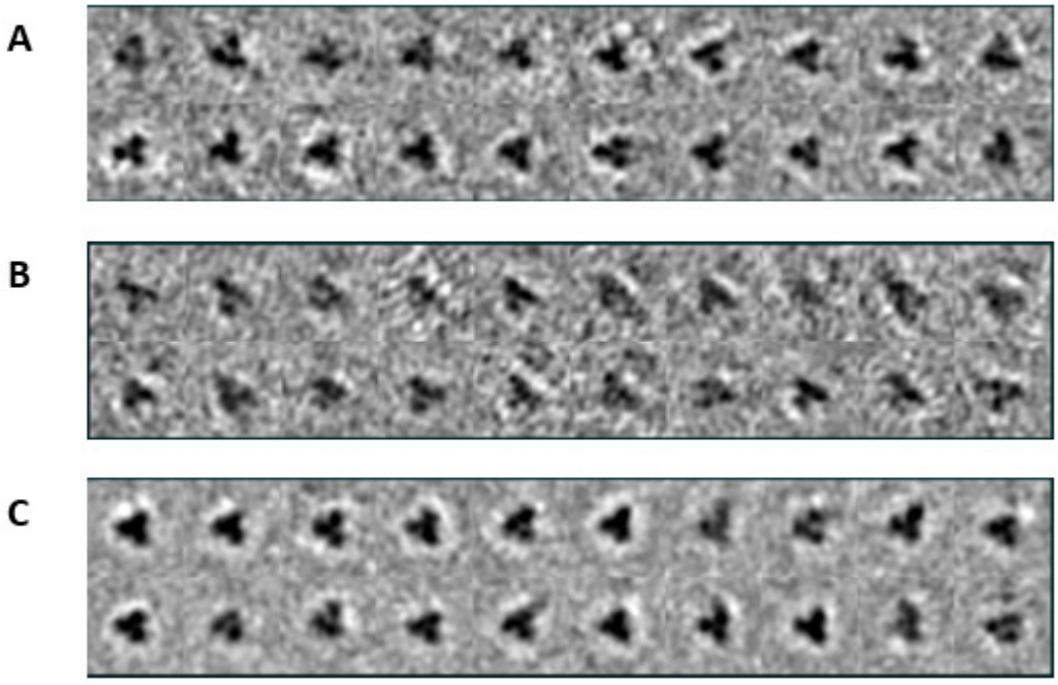

Figure 4.20: Class averages showing the T-shaped and trimeric spikes. A) Top view of 20 classes of only trimeric-shaped spikes and B) Top view of 20 classes of only T-shaped spikes and C) Class averages of trimeric-shaped spikes after symmetrization.

hour) but for the later spin cycles, with gradually decreasing number of subvolumes it took only 30 to 45 minutes to complete the full iteration i.e. MDA, classification and alignment.

The goal of our research is to develop a semi-automatic 3D segmentation mechanism that could be used for any membranous specimen with embedded proteins extending above the membrane. Because of the fact, that HIV/SIV tomograms under study are highly noisy having very low signal-to-noise ratio (SNR), and the envelope spikes are distributed all over across the virus envelope, identification of each individual spike manually, needs huge human intervention and is prone to errors. Our automated segmentation method greatly alleviates this problem substituting computer time for human time.

We analyzed three experimental data sets of frozen-hydrated SIV samples and carried out our algorithm independently in all of them to test and validate our method. The shapes of the virions varied from perfectly spheroidal and ellipsoidal to extremely irregular shapes. We started from an automatically generated point cage either of spheroidal or ellipsoidal shapes, depending on the measured radii of each virion. Generated point cages were the initial estimation of the underlying 3D contours of the virions under study. For the irregularly shaped virions, the initially generated point cage was a rough estimation of the actual shape. Surprisingly, our method could capture the accurate shapes of the virion contours quite efficiently within first few iterations. The spikes, protruding outside the viral membrane, were identified by the method. Interestingly, the spikes located at the top and bottom Z-levels of the tomogram, which are difficult to identify by manual inspection, were identified by this method. Because of the anisotropic resolution as a consequence of the missing wedge, the membrane of the virus envelope is either undefined or poorly defined blurriness in the polar spikes and making polar spikes difficult to identify by manual eye inspection. Our approach was able to identify both equatorial as well as polar spikes with significant accuracy within a reasonable amount of time and very little human intervention. Using manual spike picking as "truth", our results show that only a small number of spikes failed to be identified (false negative rate of 4%), which is a very important result in the analysis of the structure.

We believe the method would be useful for other enveloped viruses or cellular plasma membranes with large molecules extending out the membrane. We have yet to apply this method to HIV virions which pose a greater challenge because the spike density is ~10% of the mutant SIV spike density. This method has a very good potential for whole cell segmentation, where manual segmentation will take several years to finish. Our approach does not depend on the

template of the object that is being looked for. Hence this method could be of great use in whole cellular segmentation, the biggest challenge in the world of biological segmentation.

CHAPTER 5

THE ACTIN-MYOSIN INTERFACE (ACTO-MD) STRUCTURE DETERMINATION

The interaction between myosin heads and F-actin filament is the key feature for force generation in muscle. In order to understand the mechanism of muscle contraction and cargo movement along actin filaments in the cytoplasm at the atomic level, it is necessary to understand how myosin heads bind F-actin during the various steps of the ATPase cycle.

Myosin (Fig. 5.1) is composed of paired molecular trimers, the heavy chain plus the essential and regulatory light chains. The C-terminal part of the paired heavy chain forms a coiled-coil rod; the remainder of the heavy chains and the two light chains form two globular heads (called Myosin cross-bridges), each of which can independently bind the thin (actin) filament.

The binding of myosin to actin weak in some steps of the catalytic cycle and strong in other steps, i.e. the ones where tension is generated. In the absence of nucleotide, the myosin crossbridge binds tightly to the actin filament to form the ‘‘strong’’ or ‘‘rigor’’ complex. The binding of ATP to the myosin crossbridge rapidly dissociates the actomyosin complex [62]. Myosin then hydrolyzes ATP and forms a stable myosin-products complex (M•ADP•Pi). This reaction primes the crossbridge, which then reattaches to a neighboring actin site. Binding to actin causes a crossbridge to change its shape so as to move the actin approximately 10 nm. This phenomenon is called the ‘‘powerstroke’’ or ‘‘working stroke” (Fig 5.1). Binding to actin first releases the phosphate from the crossbridge and at the end of the powerstroke, ADP is released, which allows a new ATP molecule to bind to the myosin. ATP binding produces “weak binding” to F-actin, hence the myosin head detaches. In the absence of ATP the binding is “strong.”

Hence right after the ATP binding, a rapid release of the crossbridge from the actin filament follows and the cycle starts again (Fig. 5.2). This crossbridge cycle was first proposed by Lymn and Taylor (1971). In the absence of ATP (rigor mortis), the crossbridge binds tightly to actin in the end of powerstroke conformation. The mechanism of this interaction and its control by ATP is central to an understanding of muscle contraction. Thus, we need to know the structure of the strong-binding state of myosin to actin in atomic detail.

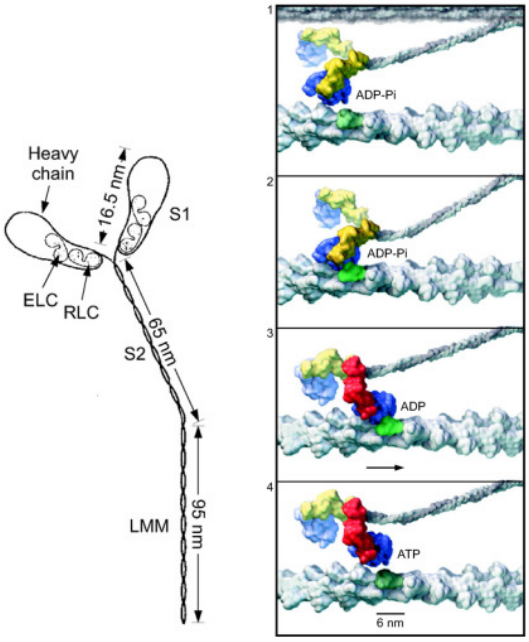

Figure 5.1: Myosin structure. a) Myosin is composed of three paired molecules, the heavy chain and the essential and regulatory light chains. Part of the heavy chains form a coiled-coil tail; the remainder of the heavy chains and the two light chains form two globular heads, each of which can independently bind the thin (actin) filament. b) Acto-myosin power stroke. The acto-myosin power stroke. (1) A myosin head with bound ADP-Pi approaches an actin-binding site. (2) The head becomes strongly bound. (3) During this step the Pi disassociates and the head rotates about a hinge, and the actin filament is displaced. (4) The ADP also disassociates, ATP binds to the myosin head, and the head dissociates from the actin filament, thus allowing the cycle to repeat. Blue is head catalytic core; yellow and red are, respectively, the pre- and post-stroke lever arm of the head. [127]

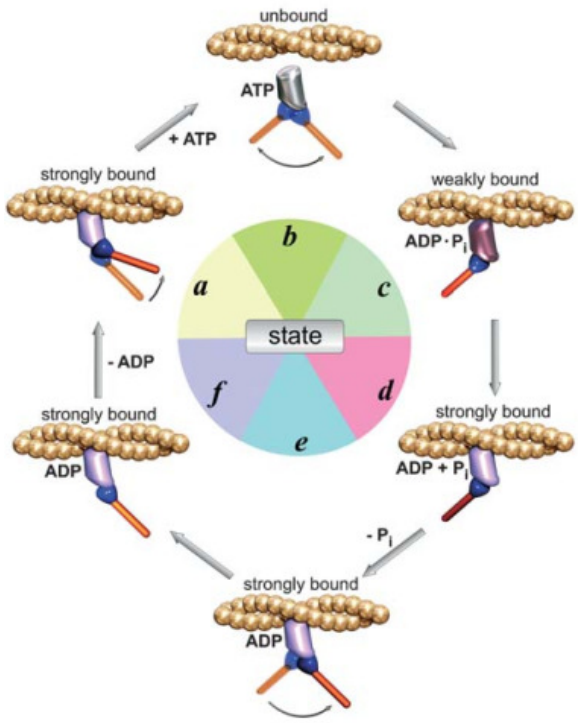

Figure 5.2: A minimal mechanochemical scheme for the acto-myosin cross-bridge cycle [62]**.** Starting from the rigor complex (state a), ATP binds to the ATP binding site and causes rapid dissociatiation of the complex and the lever arm is reprimed to the pre-power-stroke position (state b). This is followed by hydrolysis. The M·D·Pi complex rebinds to actin, initially weakly (state c) and then strongly (state d). Binding to actin induces the dissociation of Pi and the power stroke (state e). The completion of the tail swing (state f) is followed by ADP release to return to the rigor-like complex (state a); in some myosins ADP dissociation is associated with a further displacement of the lever arm. Actin monomers are shown as golden spheres. The motor domain is coloured metallic grey for the free form, purple for the weakly bound form and violet for the strongly bound form. The converter is shown in blue and the lever arm in orange [62].

5.1 The Structure of the Actin-Smooth Muscle Myosin Motor Domain Complex in the Rigor State

5.1.1 Introduction

Myosins form a family of motor proteins currently comprising over 30 identified classes [81] that function within cells to move different types of cargo along F-actin while converting the energy of ATP into work. The myosin-II class, the only filament forming class, plays a central

role in muscle contraction where binding to F-actin, accelerates its ATP hydrolysis rate to produce filament sliding and sarcomere shortening [82].

Muscle myosins are primarily of class II and consist of a pair of heavy chains (HC) each of which has a pair of bound light chains, the regulatory light chain (RLC) and the essential light chain (ELC). In smooth muscle myosin-II (smM-II), the first ~850 HC residues constitute the head which is folded into a globular motor domain (MD) containing the catalytic and actin binding properties, followed in turn by a small folded domain called the converter and a long α-helix to which the light chains bind [83]. This light chain-binding domain or LCD constitutes the lever arm, the motion of which causes movement of the cargo, in this case the thick filament, relative to the actin filament to produce sarcomere shortening. Following the head is a long α-helix in the form of an α-helical coiled-coil, the first ~1/3 of which, the S2 domain, causes the HC to form a dimer and the rest forms the thick filament backbone. A recent 6 Å cryoEM 3-D reconstruction of the thick filaments from the flight muscles of the large waterbug *Lethocerus indicus* has revealed the details of myosin II rods within the backbone in unprecedented detail [84].

Historically, the myosin head has been described as comprising three major proteolytically derived domains that are named after their respective molecular weights, the N-terminal 25-, the 50-, and the C-terminal 20-kDa segments of the heavy chain [85]. The crystal structure of myosin-II from vertebrate skeletal muscle showed that the 25-kDa domain contains an SH3 motif and that the 50-kDa domain is separated into lower and upper domains by a distinct cleft [86]. The 20-kDa domain begins with a helix associated with the lower 50-kDa domain and includes the converter and the HC component of the lever arm. Situated at the center of these four domains is a seven-stranded β-sheet that connects them via a couple of loops and helices.

The actin-based motility of myosin consists of repetitive kinetic cycles in which myosin produces a high rate of ATP hydrolysis when interacting with actin. As a product-inhibited ATPase [83], the actin-induced conformational changes enable myosin to release the hydrolysis products, rebind ATP and continue the cycle. This mechanism was initially elucidated using muscle myosin and evolved into the Lymn-Taylor kinetic model [87]. Several classes of non-muscle myosins work in a similar way, but with modifications of the different rate constants producing different functional adaptations [67]. Thus, this cyclic myosin-actin interaction has become a general ATP hydrolysis mechanism of myosin. During this cycle, the myosin head bridges the separation between thick and thin filaments, initially weakly, followed by conformational changes, some linked to product release, that alter both the position of the lever arm and the affinity for actin. The process results in force generation, the so-called power-stroke, and is now referred to as the swinging lever-arm hypothesis [88].

In the myosin catalytic cycle, not only are large conformational changes observed in the major subdomains, but also subtle changes occur in the connectors, which might have close associations with the accelerated rate of ATP hydrolysis and the swinging of lever arm. It has been reported, initially in myosin V [89] and later in myosin II [90], that the seven-stranded β-sheet is more twisted in the nucleotide-free state than in the actin-detached transition states. Both of these structures differ from other nucleotide free crystal structures in having a closed actin-binding cleft that is associated with the rigor complex of acto-myosin. In the transition state, a structure called the relay helix is kinked but it is straight in the rigor state [88]. The actin binding cleft is open in the weak binding states [91], or half-closed in the prepower-stroke state [92], or entirely closed in the rigor state when attached to actin [64] [93]. The positions of key structures

at the catalytic site, switch 1, switch II, and P-loop, highly depend on the biochemical state of myosin [94].

Knowledge about the structure of myosin in different catalytic steps has been accumulated gradually from the crystal structures of myosin subfragments dissociated from F-actin combined with spectral analysis and electron microscopy of actin filaments decorated with myosin subfragments [95]; [96]; [97]; [98]. To date, structures of three actin-detached catalytic intermediates of myo-II have been solved; the prepower-stroke, sometimes called the transition state [91]; [96], post-rigor [99], and a so-far unique ADP-bound state with an unusual lever arm position [100]. Previously structures of F-actin-myosin-II complexes structures were determined from cryoEM only to medium resolution [93]; [86]; [101]; [102]. Recently, 3-D images of the actin-myosin complex from a non-muscle class I and class II myosin have been reported at near atomic resolution [64]; [65]. The least well-characterized step is the transition between the weakly attached, prepower-stroke and the strongly bound power-stroke despite it being the most critical step of the Lymn-Taylor cycle.

Here we report a sub-nm actin-bound smooth muscle myosin-II motor domain (smMD) complex in the nucleotide-free state at an average resolution ~6 Å using iterative helical real space reconstruction (IHRSR). The reconstruction is very similar to the structure of the non-muscle myosin II class bound to actin in most essentials, particularly in the actin subunit structure, the actin-myosin interface as well as the transducer β-sheet. The nucleotide free myosin-V crystal structure was also an excellent fit to the density map even though that structure was in an actin-free state. Conversely, the myosin-V transition state crystal structure was, as might have been predicted, a poor fit. When our density map and atomic model are compared with the recent 5.2

Å structure of nucleotide F-actin decorated with free myosin V, large differences are seen in the N-terminal 25 kDa domain and upper 50 kDa domains.

5.2 Experimental Procedures

5.2.1 Specimen Preparation

The smooth motor domain consisted of residues 1-Leu790, followed by a FLAG-tag to facilitate affinity purification. Sf9 cells were infected with recombinant baculovirus encoding for the heavy chain, harvested 72 hours later, and purified by FLAG-affinity chromatography (Sigma-Aldrich) essentially as described in [103].

Actin was prepared from rabbit muscle acetone powder [104] with the modification that the chromatography step was done on a Superdex 200 column. Actin was stored as G-actin in a -80°C freezer, thawed as needed. It was then polymerized to 1.5 mg/ml F- actin (with 10mM Imidazole, 10 mM KCl, 2 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, pH 7.4) for 1 hour and diluted to 0.1 mg/ml just before use (with 10 mM Imidazole, 10 mM NaCl, 0.5 mM $MgCl_2$, 0.5 mM DTT, pH 7.4).

Specimens were made for cryo-EM by applying 4 µl of actin to the grid bar side of a 2/1 µm Quantifoil grid for 1 minute, rinsing with MD dilution buffer and applying 3 µl MD for ~5 minutes. Some grids were prepared in a 3° C cold room by manually blotting for 3-4 seconds and followed by plunging into liquid ethane. Other grids were frozen at the University of Vermont in a Gatan CP-3 freezing device operated at 100% relative humidity at room temperature.

5.2.2 Data Collection and Preliminary Analysis

Approximately 4,000 low dose images were collected automatically using the Leginon software package [105] on a Titan Krios electron microscope (FEI, Hillsboro, OR) equipped with a field emission gun and operated at 300 keV. Images were recorded with a DE-20 direct electron detector. The defocus mean and standard deviation was 3.6 ± 0.7 µm under focus; the pixel size was 0.9861Å, as calibrated by FEI. Each micrograph consisted of a 43-frame movie, with a total dose of 60 e^- /Å^2.

We used the Appion software package [106] to manage the data, perform damage compensated motion correction, CTF determination, and particle picking. The damage compensated motion correction process [107] was used to correct for beam induced specimen motion and accumulated electron dose. Defocus was first searched using ACE [108] and then refined by CTFFIND3 [109]. The filaments were manually selected, divided into 384 x 384 pixel boxes, then extracted and normalized using the DoG picker utility within Appion [106]. Each "particle" consisted of a filament segment masked to a length of 210 Å, or slightly more than 7 actin subunits of 27.6 Å separation. Adjacent filament segments overlapped by ~6 repeats (~84% overlap). A total of 346,395 filament segments were selected from 1,417 of the best micrographs. Appion software [106] was used to create a metadata (.star file) having all the positional, orientation and defocus information of the segments and was supplied to RELION [52] version 1.3 for classification. Once the classification is done, we could identify the particles that are then subjected to further processing to elicit the helical structure of the filaments.

5.2.3 Three-Dimensional Reconstruction

Unfortunately, we found that RELION version 1.3 does not incorporate algorithms for helical reconstruction. In order to determine the helical symmetry parameters, we used a specific version of RELION, named as Relion_helix_1.2 implemented by Z. Hong Zhou's group [110], which included the Iterative Helical Real Space Reconstruction (IHRSR) package [58] with RELION version 1.2.

A small set of particles was subjected to 2D classification to eliminate bad particles but unfortunately, none of the class average showed obvious bad particle assemblies. Hence, hierarchical 3D classification was carried out in order to identify "shiny particles" for further analysis.

To reduce the computational burden, the particle stack was divided into two halves and 3D classification was performed on each of them separately. The 8 Å cryo-EM map of the rigor (nucleotide-free) actin-tropomyosin-myosin complex (EMD-1987) [64] was low-pass filtered to 100Å and used as the initial model for 3D refinement of both the particle sets. To increase the speed of the process of selecting good particles by interactive 3D classification, each stack was first binned by a factor of 4. The first particle stack, consisting of 154,559 particles, was subjected to 3D classification from which 82,661 good particles were selected. From the second stack, consisting of 191,836 initial particles, only 104,221 good particles were identified. Then the 186,882 selected good particles were combined, binned by a factor of 2 and subjected to 25 more classification iterations. Particles were randomly divided into four classes and four reconstructions calculated. A set of projections were generated for each reconstruction, and used to reassign each of the 186,882 particles to one of the four groups according to which projection

it most closely resembled. After 25 cycles of 3D classification, the final good-looking classes were selected and combined leaving 101,976 segments for “shiny particle” data analysis.

“Shiny particle” selection utilized 25 additional hierarchical 3D classifications of the 101,976 good particles. This time, one of the best-looking class averages from the final iteration of the previous classifications was chosen as the reference image and filtered to 20 Å. The 101,976 particles were divided into four classes containing 26,683, 17,698, 40,847 and 16,748 filament segments. All of the four class averages appeared good and 3D auto refinement was carried out on each of the classes separately. This process revealed an acto-MD density with an estimated resolution of ~7 Å for all four classes. Each of the four reconstructions were compared with one another in Chimera and the three most homogeneous classes (1-3) were combined to produce the final reconstruction. A combined total of ~85,000 particles were subjected to 3D auto-refinement. One of the good reconstructions from the previous individual auto-refinement scheme was low-pass–filtered to 60 Å and used as the starting model for the final reconstruction. The 3D auto-refinement converged in 24 cycles.

The resolution based on the gold standard FSC (0.143 criterion) [111] showed an average resolution of ~6 Å for the final acto-MD electron density map. The temperature factors are calculated using EM-BFACTOR [112] and the F-actin-MD map was sharpened using a temperature factor of -390.86 Å. Local resolution of the full reconstructed volume computed using Resmap [113] revealed a resolution gradient of ~4.0 Å in the actin core region, ~5 Å in the central part of the map and ~6.5 Å at the outer myosin domains.

5.2.4 Atomic Model Fitting

The starting models of myosin used for homology modeling came from the crystal structure of myosin II subfragment 1 from the adductor muscle of the scallop *Argopecten irradians*, (PDB 1DFK) [100] and the crystal structure of the myosin II motor domain from the slime mold, *Dictyostelium discoideum*, (PDB 1FMV) [114]. First, a homology model of the chicken smooth muscle myosin sequence was built by MODELLER [115] using both PDBs as input. Because several large loops, such as loop 2, are not determined in all the myosin head S1 atomic models, those loops were deleted after homology modeling, in order to avoid clashes in the real space flexible fitting. The real space flexible fitting was performed using Relax in Rosseta [116] at a resolution of 5.5 Å. Because the converter domain has a large conformational change, the SH3 domain, which is located near the converter domain, was not fit well into the density by the real space flexible fitting. So, after the converter domain was fit, the SH3 domain was manually fit into the density and another real space flexible fitting of myosin was performed.

The starting model of actin was taken from the actin-tropomyosin filament structure (PDB 3J8A). The actin species in PDB 3J8A is of skeletal muscle α-actin from the rabbit *Oryctolagus cuniculus*, which is same as our sample. The α-actin atomic model was four residues short of the actual C- terminus. When we found that the γ-actin atomic model [65] fit our actin density quite well, we built in the remaining four residues based on their placement in the γ-actin structure. As of this writing, the subsequent flexible fitting has not been completed. The real space flexible fitting was performed using Relax in Rosseta at a resolution 4.0 Å. In the above fitting, only one myosin MD and one actin subunit are considered. Next, the actin-MD combination was refined using rosetta_scripts in Rosseta [116] with asymm_refine.xml. Next, contact between two

successive actin-MD subunits was refined using rosetta_scripts with symm_refine.xml. Finally, ADP was fit into density map with the fixed actin main chain using Relax.

5.3 Results and Discussion

Motion corrected images recorded on the DE-20 showed F-actin with a high degree of saturation of actin subunits with MD, which are individually resolved (Fig. 5.3). A significant fraction of the filaments appeared bundled and thus only a small fraction of filaments were suitable for further analysis. In addition to the heavily decorated F-actin, a significant fraction of filaments was completely undecorated, a phenomenon typical of F-actin decorated with rigor myosin heads [64].

The reconstruction procedures to find the best-preserved segments of decorated actin eliminated 75% of the segments. The remaining 25% of segments produced a density map with variable resolution that depends roughly on the distance from the helical axis (Fig. 5.4A). The local resolution computed with RESMAP shows regions with ~3.5 Å but these occur mostly near the actin filament. On the MD, the local resolution is mostly in the 5-7 Å range (Fig. 5.4B), being best near the actin and worst near the converter and SH3 domains. Some regions within the smMD show right handed α-helices and their bulky side chains rather than cylindrical shapes in those places where α-helices are expected consistent with a resolution of 4.5-5 Å (Fig. 5.4C). Except around the actin subunits, other places in the reconstruction do not show density corresponding to amino acid side chains with clarity.

We fit an atomic model to the density using Rosseta but starting with a homology model based on the prepower stroke transition state of scallop adductor muscle myosin II. Despite the rather large difference in conformation between the starting model and the final fitted model, the

agreement is quite striking with the atomic model from the higher resolution reconstruction of F-actin decorated with vertebrate non-muscle myosin II motor domain (nmMD) [65]. The structures of the actin subunit are nearly identical (Fig. 5.5A). At the C-terminus, our reconstruction indicated a helical arrangement of the last four residues that were missing in the rabbit α-actin that we started the fitting with (Fig. 5.5B). The non-muscle γ-actin atomic model [65] is a very good fit to this feature so we manually built the last four residues into the density using the γ-actin model as a guide. At the N-terminus of our actin structure, density extends only as far as T5. Two other minor departures occur at G168 (G167 for γ-actin) where the γ-actin chain falls out of the density envelope slightly and T324 (T323 for γ- actin) where both models appear to fit the density equally well. We conclude that at our resolution, the α-actin used in the present study and the γ-actin used for non-muscle myosin II are nearly indistinguishable.

Our density map where F-actin is located also shows strong density in the ADP binding pocket (Fig. 5.5C). We therefore fit ADP into the density. The conformation obtained is similar but not identical to that obtained for γ-actin [65]. The differences are probably not significant at our resolution.

For the smMD, the atomic model begins at residue D2 but at the optimal contour threshold, the model does not enter defined density until N22. The topology of the SH3 domain is the same as the nmMD, but the chains are displaced by about half the spacing between β-strands (Fig. 5.6A,B). The converter domains, the other feature at high radius of interest, have a similar but not identical topology and do not overlap. This could be influenced by the fact that the nmMD has 7 turns of the lever arm α-helix, which are missing in the smMD construct (Fig. 5.10B). The most noticeable difference in the converter occurs between residues E735 and G749. This

segment of polypeptide chain is extended and without secondary structure and differs significantly in the path that the two chains follow.

In the region of the transducer β-sheet, the two structures are virtually superimposable as are most of the α-helices in that neighborhood (Fig. 5.6B, C). The helix from R507-W546 is tilted out of alignment at its beginning but otherwise the differences in this area seem insignificant and the alignment seems quite good.

At the actin-myosin interface, the similarity is more striking than the differences (Fig. 5.6E, F). The most noticeable difference occurs in the loop from R530-G536. The corresponding loop in

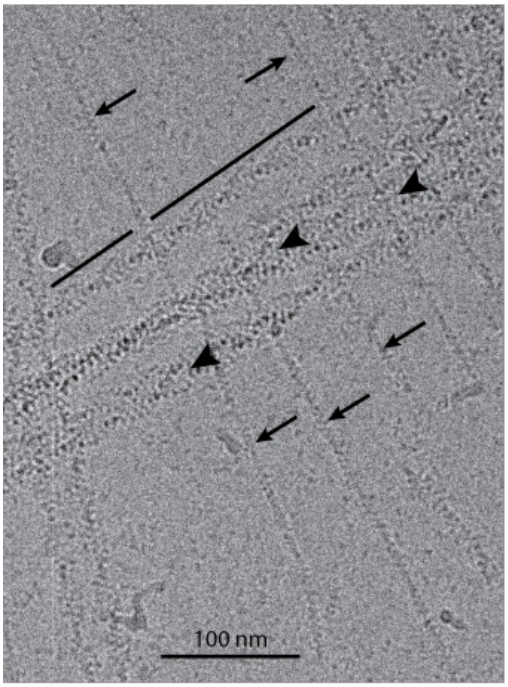

Figure 5.3: Electron micrograph of F-actin decorated with the smooth muscle myosin motor domain. Segments were taken only from the filament region marked by the line. Arrow heads point to bundled filaments. Arrows point to actin filaments completely undecorated with myosin heads.

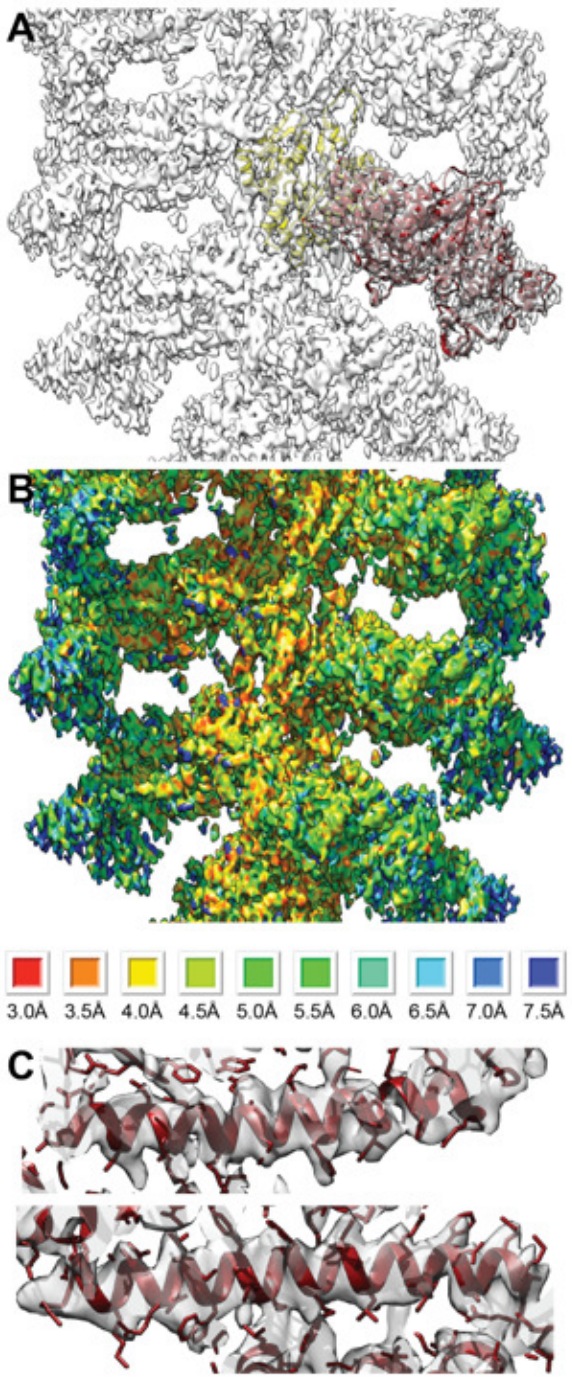

Figure 5.4: Overview and resolution of the reconstruction of F-actin decorated with smMD. (A) Overview showing the atomic models of the actin subunit (yellow) and the smMD (dark red). (B) RESMAP image of the reconstruction. Resolution is clearly highest close to the filament axis and lowest at the high radius where the converter and SH3 domain are positioned. RESMAP color ranges are shown at the bottom. (C) Images of a pair of long α-helices from the acto-smMD reconstruction (purple). Large side chains are clearly visible. Top panel helix comprises residues 477-506; bottom panel comprises residues 420-450.

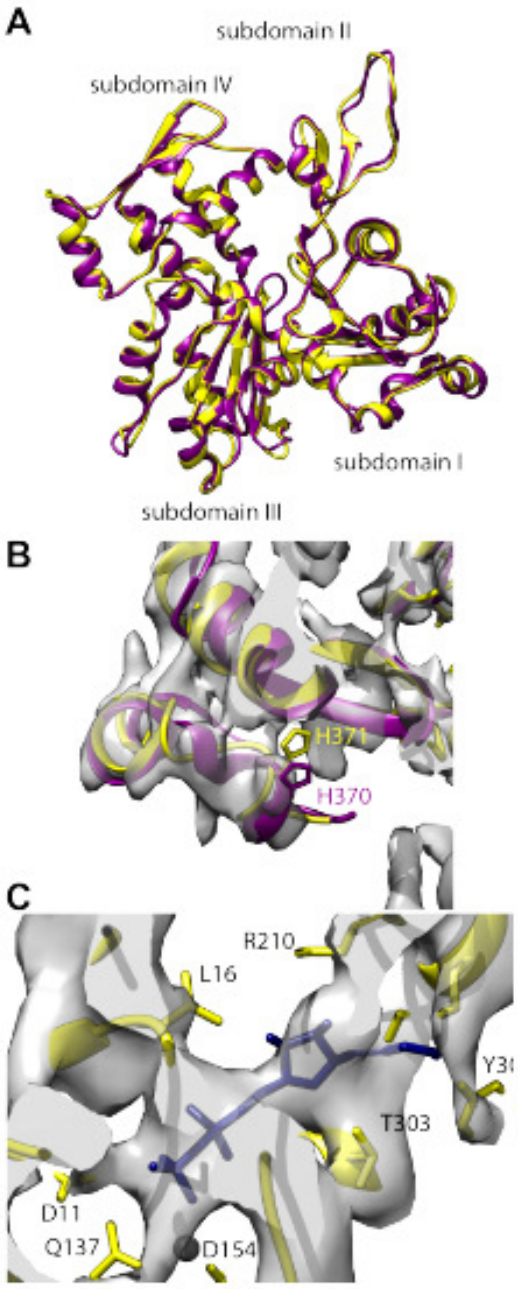

Figure 5.5: Comparison of vertebrate non-muscle and rabbit skeletal muscle actin subunits. (A) Overlay of the fitted actin subunits. Present reconstruction is colored yellow and the actin subunit from vertebrate non-muscle actin subunit colored dark magenta. (B) Region near the C-terminus. The initial rabbit actin subunit did not include four residues at the C-terminus. These were added in later but have not been energy minimized. The C-termini of the non-muscle γ-actin subunit fits the density very well indicating that after refinement, the rabbit muscle α-actin C-terminus will likely be very similar. (C) Region near the ADP binding site. Substantial density is present where the nucleotide binds. The black sphere is a magnesium ion for which clear density is not visible. Its presence provides a useful landmark.

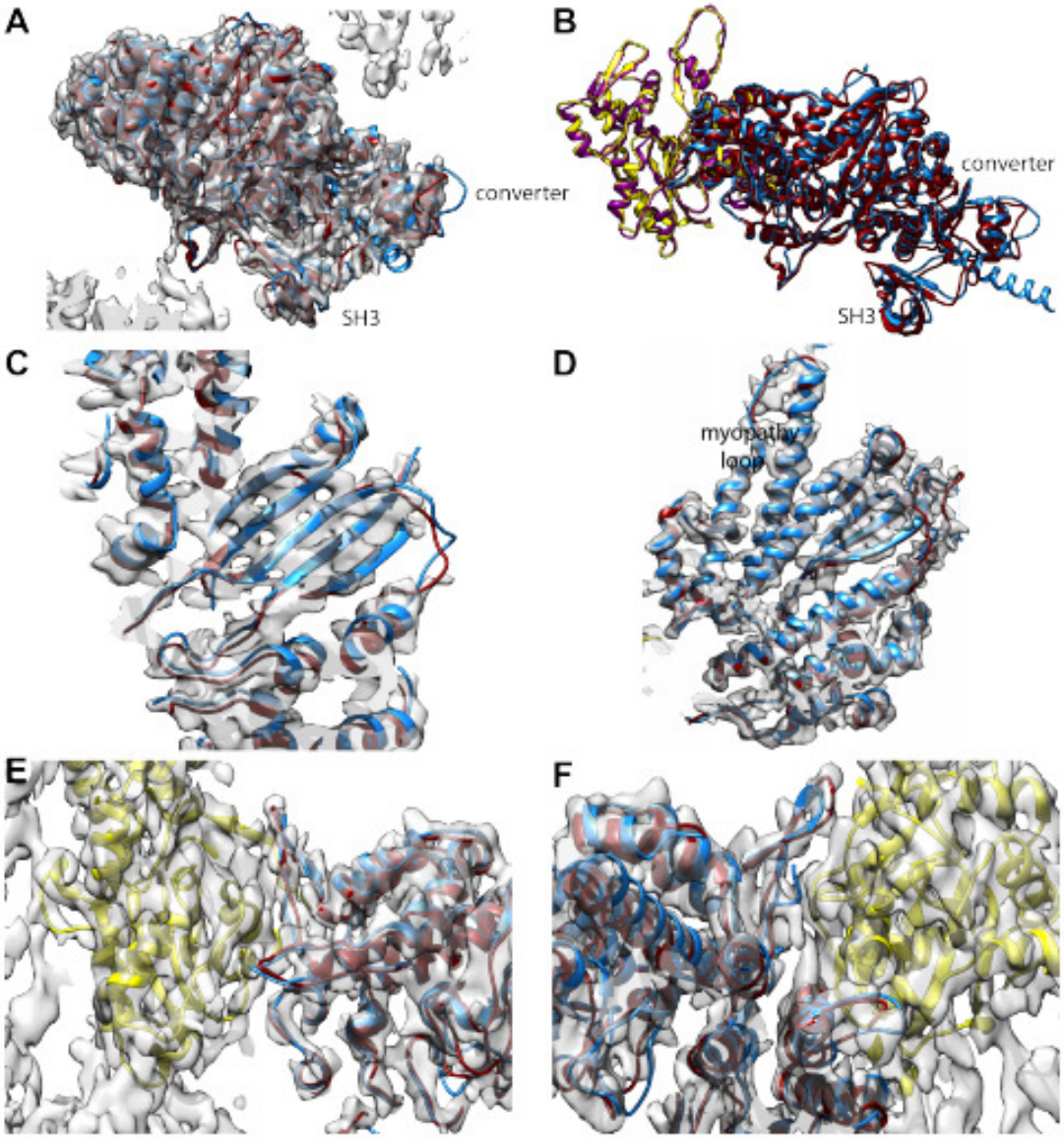

Figure 5.6: Comparison of vertebrate non-muscle and smooth muscle motor domains when bound to F-actin. Vertebrate smooth muscle myosin MD is colored dark red; the vertebrate non-muscle MD is colored dodger blue, rabbit muscle α-actin is yellow, non-muscle γ-actin is colored dark magenta. (A) View showing the relative difference between converter and SH3 domains plus the reconstruction envelope. (B) Slightly different view from (A) showing the actin subunits and the converter in better profile without the map. (C). View showing five strands of the transducer β-sheet with both the smooth muscle MD (dark red) and non-muscle MD (dodger blue). (D) Similar view as (C) but showing four major α-helices, which align well to the smooth muscle MD density and atomic model. (D) Actin-myosin interface from the "front". (E) Actin-myosin interface from the back. In all these views, the vertebrate smooth muscle acto-MD and the vertebrate non-muscle acto-MD atomic models are nearly superimposable at this resolution.

nmMD is R543-G549. In fact, the fit of our smMD in this region is not good, but is closer than is the nmMD. Density corresponding to loop 2 is not visible in our reconstruction, nor is it visible in the nmMD structure. Other than this, the topology in the actin-myosin interface is very similar.

We also compared our structure with the recently published crystal structures of the F-actin-myosin-V MD structure nucleotide free and with ADP strongly bound (PDB 4ZG4) [66]. Here the differences were much greater. Because there is little sequence homology between the smMD and the myosin-V MD, we fit both myosin-V MD atomic coordinates to our map as a rigid body using the *fitinmap* utility of Chimera [117]. The initial fit done this way for nucleotide-free myosin-V MD was entirely satisfactory and could not be visually improve by manual adjustment. The nucleotide-free myosin-V coordinates overlapped the smMD atomic model quite well and fell almost entirely within the density map envelope (Fig. 5.7A-E).

The fit using the ADP bound myosin-V MD was poor (data not shown). When fit as a rigid body into the density, many features were displaced out of the density and poorly aligned with the smMD coordinates. Although the topology of both MD atomic models is similar, most features are displaced or otherwise modified almost nothing overlapped exactly.

We also compared our reconstruction with the recent structure of rabbit striated muscle α-actin decorated with subfragment 1 of rabbit skeletal muscle myosin II [118], hereafter referred to as the skMD since only the motor domain can be compared with the smMD. Here we found significant differences. We tried three alignment methods. The first used the actin subunit coordinates to drive the alignment using the *Matchmaker* utility of Chimera. Done this way, the actin subunit structures are superimposable and matched well and with them the helices that

make up the lower 50 kDa domain that are bound to actin However, nothing else matched well except for those features, like the myopathy loop, that contact actin. We also fit the actin-skMD coordinates as a single rigid body into the acto-smMD density map. This fit was slightly different giving some improvement to the parts that fit poorly using the actin subunit as the alignment driver, i.e. the upper 50 kDa domain, but at the expense of those features that fit well, i.e. the lower 50 kDa domain. The third fit used only the skMD coordinates. This result was also slightly different giving again a small improvement to the upper 50 kDa domain at the expense of the lower 50 kDa domain. We preferred the fit using the actin subunit coordinates as the alignment driver because the differences are more easily visualized but the other alignments did not eliminate these differences; it only reduced them. The following discussion is based on the first method.

When the smMD and the skMD are compared within the density map, two things stand out. First, the so-called loop 3 feature of the skMD is located completely outside of density (Fig. 5.8A). This is the only part of the lower 50 kDa domain that does not match well. Note that the myosin-V MD structure was a good fit at this loop (Fig. 5.7C,D). Second, many of the upper 50 kDa domain helices are positioned outside of density (Fig. 6A) in a way that places them further from the axis of the F-actin. When the map is left out and the models viewed from the other side to show the lower 50 kDa domain, it matches well but the upper 50 kDa domain does not (Fig. 5.8B). We also looked at the position of the transducer β-sheet, which revealed significant differences. Generally, the length of the peptides that conformed to the β-sheet conformation were longer in the smMD atomic model than the skMD model with the first three strands displaced to the right (Fig. 5.8C). The transducer β-sheet is also twisted differently (Fig. 5.8D). When the map is superimposed at the same time, the first four strands are not positioned within

the density map (Fig. 5.8E). When the two atomic models are viewed looking through the actin binding cleft, the general impression is that with the exception of the lower 50 kDa helices, which fit very well, the upper 50 kDa and N-terminal 25 kDa domains are shifted outwards of the smMD atomic model (Fig. 5.8F). This gives the impression that the actin-binding cleft is more open in the skMD structure than in the smMD structure.

We find that the structural differences between the smMD, the nmMD and the myosin-V MD when bound to actin are small whereas the differences between the smMD and the skMD when bound to actin are large. Neither our reconstruction of acto-smMD nor that of acto-skMD has sufficient resolution to make a detailed comparison at the level of amino acid side chains. The most obvious difference at the current resolution lies in the size of the actin binding cleft, reflected in the displacement outward of the upper 50 kDa domain, and the position of the transducer β-sheet, which likely correlate with the properties of the two-myosin species. According to the Protein Data Bank, the myosin species used for the acto-skMD reconstruction is rabbit extraocular muscle, which is a super-fast muscle characterized by rapid shortening and comparatively low tension [119]. On the other hand, vertebrate smooth muscle is a slow muscle capable of sustained, high-tension contractions. Myosin-V also has a comparatively slow actin-activated ATPase with a slow rate of ADP release [120].

The kinetics of the actin-activated ATPase of four muscle myosins from chicken, fast and slow skeletal, cardiac and smooth, were compared in a single study [121]. Measurable differences were found in the rates of ATP induced dissociation from actin, the rate of reassociation with actin after ATP cleavage and rates of release of ADP when actin bound. The rate of ATP induced dissociation from actin at 20°C was slowest for smooth and too rapid to be measurable for fast

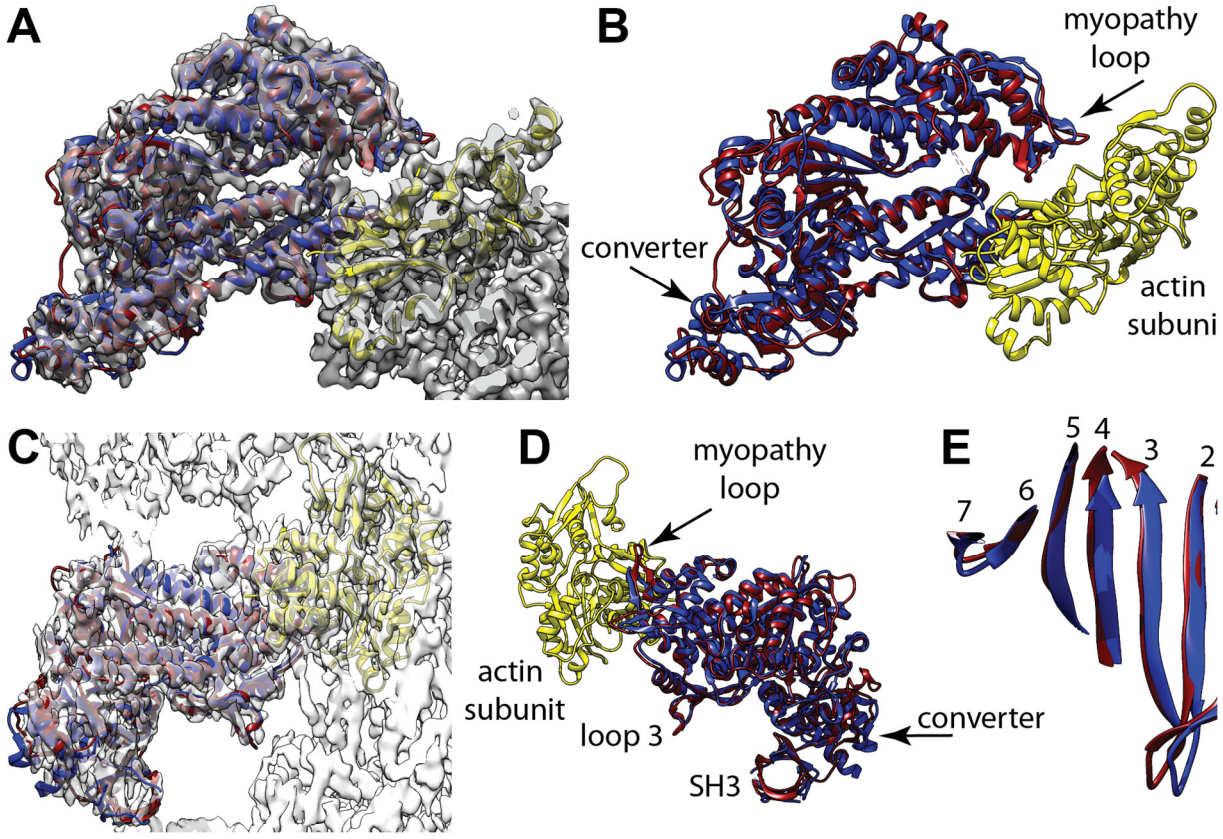

Figure 5.7: Comparison of the actin bound smMD with the nucleotide-free myosin-V MD crystal structure (PDB 1OE9), which has been aligned to the reconstruction using Chimera's *fitinmap* utility. Coloring scheme has the actin subunit yellow, the smMD dark magenta, and the myosin-V MD blue. (A) View down the actin binding cleft showing the excellent fit of the myosin-V crystal structure even though not bound to actin. The myosin-V converter domain has a very similar position and orientation as the smMD converter. (B) Same view direction as panel A but with the map removed to shown the excellent alignment of the myosin-V helices with the corresponding smMD helices and loops. (C) View perpendicular to the helix axis showing the fit of the myosin-V coordinates within the acto-smMD reconstruction. (D) View from the opposite side without the map showing alignment of the myopathy loop, loop 3 and the SH3 domains. (E) View showing the excellent alignment of the 7-stranded transducer β-sheets.

skeletal myosin; at 3°C there was a 4-fold difference. Marston and Taylor concluded that ATP must induce a conformational change in myosin which we now know involves opening of the actin binding cleft [122]. The more open actin-binding cleft in the acto-skMD reconstruction, which may be interpreted as partially along the opening pathway, may offer an explanation for the difference in this rate. Since we only observe a single, nucleotide-free, actin-bound state, the

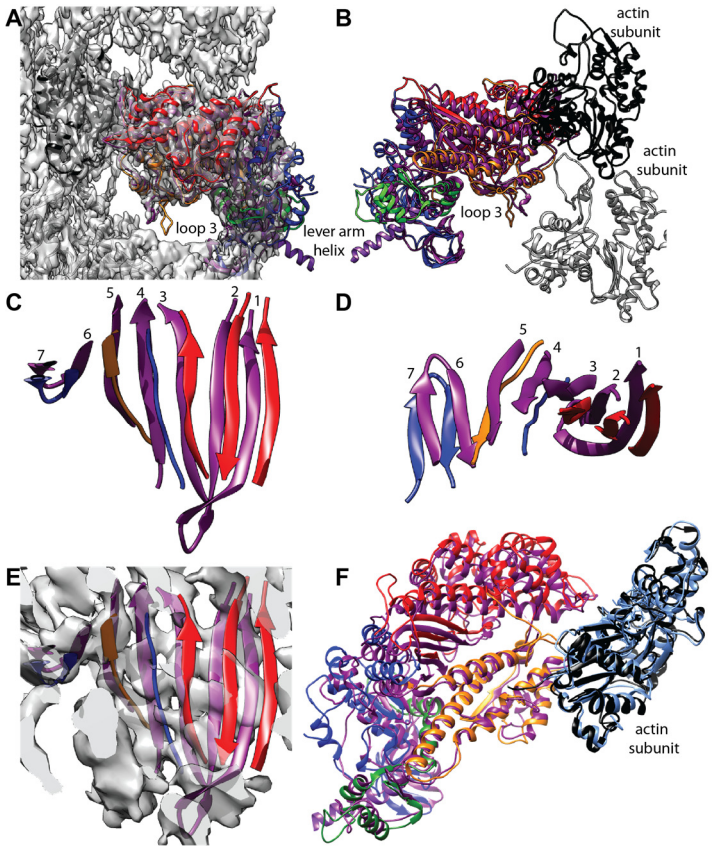

Figure 5.8: Comparison of acto-smMD with acto-skMD. (A) The two reconstructions shown with the acto-smMD reconstruction. The acto-skMD was aligned to the acto-smMD using the coordinates of the actin subunit, which is colored black and comes from the acto-skMD atomic model (PDB 5H53). There is little difference between the two actin subunit atomic models from the two reconstructions. The smMD atomic model is colored purple. The acto-skMD atomic model is colored according to the MD subdomains, which are N-terminal 25 kDa domain (blue), upper 50 kDa domain (red), lower 50 kDa domain (orange), converter domain (green) and the lever arm (magenta). Note that the smMD does not have the lever arm helix. (A) Both atomic models shown within the reconstruction envelope. Many features of the skMD atomic model fall outside of the density envelope of the smMD. The most obvious difference is the position of loop 3 (skeletal residues K567–F579), which falls clearly outside the reconstruction envelope. (B) The atomic models of the skMD and the smMD shown with a pair of actin subunits, one black, the other gray. This view from the opposite direction from that of panel A. (C) Comparison of the transducer β-sheet with the smooth muscle structure shown in purple and the skeletal muscle sheet colored according to subdomain origin. Since the sheet itself is curved, the displacements for strands 1 and 2 are the most obvious. This view direction is from outside the MD looking in towards the actin-binding cleft. The relative displacement has the skeletal β-sheet to the side and on the outside the smooth β-sheet (Roughly looking from the top of panel F towards the bottom). (D) View looking down from the top of panel C. (E) Same view direction as panel C but with the reconstruction envelope showing. Note that the skMD β-sheet mostly falls outside of the corresponding density envelope. (F) View looking down the actin binding cleft showing the actin subunit atomic models from the two reconstructions as well as their MD atomic models. The actin atomic model from the acto-smMD reconstruction is shown in sky blue. Note that the lower 50 kDa domains overlap well, whereas the upper 50 kDa domains overlap poorly. The 25 kDa domains also overlap poorly.

reconstructions cannot offer an explanation for the rates of re-association with actin following ATP cleavage or the differences in ADP release. However, we do point out that in smooth muscle myosin, the ADP release rate is about 20 times slower for smooth compared to fast skeletal muscle myosin [121] and causes a further, 35 Å displacement of the end of the myosin lever arm toward rigor [123], which may affect the rate of ADP release.

CHAPTER 6

CONCLUSION

In this dissertation we have focused broadly on two aspects of electron microscopy – segmentation in electron tomography and structure determination from single particle electron microscopy. We demonstrated the results of our segmentation methods using HIV/SIV data and used Acto-MD data for structure determination from single particle reconstructions. However, as the field of electron microscopy grows every day the challenges are far from over. In fact, even in the limited context of segmentation in electron tomography there are several challenges due to the heterogeneity of the particles, existence of missing wedge and extremely poor SNR. In case of single particle, even though people have been able to get atomic resolutions, some of the methods used have very large time complexity and still some of the solutions are not optimal. For example, in problems like helical reconstruction, the methods primarily employed, can get stuck at local optima. However, getting the optimal solution is expensive. Some of these challenges would be the focus of my work going forward.

In summary, for this dissertation, we have developed a novel and reliable approach for semi-automatic selection o of HIV/SIV Env spikes using cryoET. In our second project, we determined the structure of actin-smooth muscle myosin motor domain complex in the rigor state and revealed the structurer at ~6 Å resolution using single particle data. As a background survey, we performed an in-depth study of segmentation methods that had been already applied in the field of electron tomography. A review article containing the different top-down and bottom-up segmentation methods in cryoET is in preparation and will be a good contribution to future researchers interested in working with segmentation in cryo-ET. Also, this dissertation contains

an elaborate description of processing pipeline of electron tomography. We have submitted a book chapter named Electron Tomography of Biological Specimens that includes the contents of chapter 2 of this dissertation and also some parts of other chapters.

As a future work, we are planning to concentrate on application of deep learning techniques for automatic feature extraction for classification in cryoET. Deep Learning has been used for segmentation and object detection extensively, but till now it has not been used in cryoET or Single particle electron microscopy. Due to the success of deep learning methods elsewhere it would be interesting to apply such techniques in cryoET and single particle electron microscopy.

For the structure determination of Acto-MD we achieved near atomic resolution (~3.5 Å) for the actin-part and could resolve the myosin motor domain part to a resolution of ~6 Å. For symmetry determination we used IHRSR [56]. This method searches for two parameters, helical rise and an in-plane rotation. We noticed that instead of searching a joint parameter space, this method searches on one direction while keeping the other fixed and vice versa. This one-directional search gives a sub-optimal result. The accuracy can be improved by incorporating a grid search.

Lastly, our semi-automated segmentation algorithm works with significant accuracy. Although, our experimental data showed that the method had only ~4% false negative rate while capturing equatorial spikes and ~15% false negative rate while capturing combined (polar and equatorial) spikes even with the presence of very high amount of uncertainty, we still keen to explore a mathematical analysis for segmentation by classification method. We have used this method for ribosomes and chromatin and we would like to continue that line of work and show the efficacy of the method for any such data.

APPENDIX

SOURCE CODE FOR SUBROUTINES

In this appendix we provide code snippets for the most important sub-routines that were used for this dissertation. I am distributing these sub-routines under the assumption that anyone using it would refer to this dissertation and in turn make his/her code that uses parts of this sub-routine available to researchers for free. I forbid anyone from using this code or parts thereof in any commercial enterprise and/or for making profits.

Due to space constrains we omit utility functions for using these subroutines. Anyone interested in using these packages for research purposes can request a copy by emailing to cb10u@my.fsu.edu or taylor@bio.fsu.edu. The author would expect a citation of this dissertation for any such use.

Code for calculating the orientation angle of the virions:

```
% the first radial vector
r1=sqrt(((x1-x0).^2)+((y1-y0).^2)+((z1-z0).^2));
u_vec1=[];
u_vec1=[u_vec1; (x1-x0)./r1 (y1-y0)./r1 (z1-z0)./r1];
% the second radial vector
%r2=sqrt(((x2-x0).^2)+((y2-y0).^2)+((z2-z0).^2));
r2=sqrt(((x1-x2).^2)+((y1-y2).^2)+((z1-z2).^2));
u_vec2=[];
%u_vec2=[u_vec2; (x2-x0)./r2 (y2-y0)./r2 (z2-z0)./r2];
u_vec2=[u_vec2; (x1-x2)./r2 (y1-y2)./r2 (z1-z2)./r2];
% angle between two radial vectors
%dot_prod=((x1-x0)*(x2-x0))+((y1-y0)*(y2-y0))+((z1-z0)*(z2-z0));
dot_prod=((x1-x0)*(x1-x2))+((y1-y0)*(y1-y2))+((z1-z0)*(z1-z2));
%norm1=r1.^2;
%norm2=r2.^2;
norm1=r1;
norm2=r2;
theta=acos((dot_prod)/(norm1*norm2))
theta=pi/2-theta
```

```
function theta = angle_calc_all(V)
    format longE
    % position of the 12 picked points
    %V=load('picked_12_points_new.pos');
    %V=load('vir-2.pos');
    %V=load('ZJ-points.pos');
    %V=load('V_indv.txt')
%     nx=size(V,1);
%     k=nx/4;
%     for i=1:k
%     j=i-1;
%     x0=V(4*j+1,1);
%     y0=V(4*j+1,2);
%     z0=V(4*j+1,3);
%     x1=V(4*j+2,1);
%     y1=V(4*j+2,2);
%     z1=V(4*j+2,3);
%     x2=x1;
%     y2=y1+200;
%     z2=z1;
    n = size(V,1);
    x0=V(1,1);
    y0=V(1,2);
    z0=V(1,3);
    x1=V(2,1);
    y1=V(2,2);
    z1=V(2,3);
    x2=x1;
    y2=y1+200;
    z2=z1;
    % the first radial vector
    r1=sqrt(((x1-x0).^2)+((y1-y0).^2)+((z1-z0).^2));
    u_vec1=[];
    u_vec1=[u_vec1; (x1-x0)./r1 (y1-y0)./r1 (z1-z0)./r1];
    % the second radial vector
    %r2=sqrt(((x2-x0).^2)+((y2-y0).^2)+((z2-z0).^2));
    r2=sqrt(((x1-x2).^2)+((y1-y2).^2)+((z1-z2).^2));
    u_vec2=[];
    %u_vec2=[u_vec2; (x2-x0)./r2 (y2-y0)./r2 (z2-z0)./r2];
    u_vec2=[u_vec2; (x1-x2)./r2 (y1-y2)./r2 (z1-z2)./r2];
    % angle between two radial vectors
    %dot_prod=((x1-x0)*(x2-x0))+((y1-y0)*(y2-y0))+((z1-z0)*(z2-z0));
    dot_prod=((x1-x0)*(x1-x2))+((y1-y0)*(y1-y2))+((z1-z0)*(z1-z2));
    %norm1=r1.^2;
    %norm2=r2.^2;
    norm1=r1;
    norm2=r2;
    theta=acos((dot_prod)/(norm1*norm2))
    theta=pi/2-theta
    %end
%     cos(theta);
end
```

Code for generating points on the sphere:

```
for i=1:N
    theta=(i*d/R);
    R_new=R*(cos(theta));
    Z=-R*(sin(theta));
    n=round((2*pi*R_new/d));
    t=[0:n-1]';
    A=R_new*cos(2*pi*t/n);
    B=R_new*sin(2*pi*t/n);
    plot(R_new*cos(2*pi*t/n),R_new*sin(2*t*pi/n),'o')
    C=horzcat(A,B);
    D=repmat(Z,n,1);
    I=horzcat(C,D);
    %F=[0 0 0];
    %F=[xc yc zc];
    G=repmat(F,n,1);
    I=I+G;
    E=vertcat(E,I);
```

Code for generating points on the ellipse:

```
function generate_ellipse(argsmat,base_name,suffix)
% this script will generate points that are 'd' distance apart over a
% Ellipsoid shaped virion
%-------------------------------------------------------------------------------------
--------------------------
% unwrap the arguments here
% semi major
a= argsmat(1)
% semi minor
b = argsmat(2)
% height
c = argsmat(3);
% gap
d= argsmat(4);
% tomo num
l = argsmat(5);
j=num2str(l);
% angle
theta = argsmat(6);
% xc,yc,zc,X,Y,Z
F=[argsmat(7) argsmat(8) argsmat(9)];
X = argsmat(10);
Y = argsmat(11);
h1 = argsmat(12);
h2 = argsmat(13);
xx=max(h1,h2);
yy=min(h1,h2);
%.....................................................................................
.........................
Pts_1=[];
%-------------------------------------------------------------------------------------
--------------------------
ptmat=[];
%c=Height(l);
    %S=(pi*(3/2*(c+a)+sqrt(c*a)))/4
    %Ramanujan's approximation for the perimenetr of the ellipse
    S=(pi*(3*(c+a)-sqrt((3*c+a)*(c+3*a)))./4;
    %no of ellipses in one half
    N=floor(S/d);
    % Call ellipse point generation code
    Pts = calculateEllipse(a,b,theta);
    %Pts = calculateEllipse(b,a,theta);
    nx=size(Pts,1);
    Z=0;
    D=repmat(Z,nx,1);
    E=horzcat(Pts,D);
    G=repmat(F,nx,1);
    E=E+G;
    %dlmwrite('ellipse_0.pos',E,' ');
    %perometer of the ellipses
    % P_a = pi*(3/2*(c+a)+sqrt(c*a));
    % P_b = pi*(3/2*(c+b)+sqrt(c*b));
    %Ramanujan's approximation
    P_a=pi*(3*(c+a)-sqrt((3*c+a)*(c+3*a)));
    P_b=pi*(3*(c+b)-sqrt((3*c+b)*(c+3*b)));

%-------------------------------------------------------------------------------------
---------------------------
    % points on upper ellipses
```

Code for elimination of multiple picks:

```
% this code deletes the points that are within a circle of radius 10 from a point.
So this points gives the reduced data sets
V=load('points.pos');
%change_counter = 0;
%V
final_deleted_pts_index = [];
main_data_matrix = V;
dataMatrix=V;
queryMatrix=V;
x=size(queryMatrix,1)
%=size(dataMatrix,1);
final_points=[];
deleted_points = [];
%neighborIds = zeros(size(queryMatrix,1),k);
%neighborDistances = neighborIds;
numDataVectors = size(dataMatrix,1);
numQueryVectors = size(queryMatrix,1);
dist =sqrt(sum((repmat(queryMatrix(1,:),numDataVectors,1)-dataMatrix).^2,2));
old_row_count= size(dist,1);
cnt = 0;
while(numDataVectors>0)
    i=1;
%for i=1:numQueryVectors,
    change_counter = 0;
    dist =sqrt(sum((repmat(queryMatrix(1,:),numDataVectors,1)-dataMatrix).^2,2));
    %sort(dist,'ascend')
    n_dist = size(dist,1);
    for j=1:n_dist,
        cnt = cnt + 1;
        % set condition for distance
        if (dist(j) > 0 && dist(j) <= 5)
            cnt;
            j_new=j-change_counter;
            % resize the two matrices and continue
            %fprintf('*** Deleted points are ***');
            %dataMatrix(j_new,:)
            %deleted_points = [deleted_points;queryMatrix(1,:)];
            deleted_points = [deleted_points;dataMatrix(j_new,:)];
            f_dataMatrixnew = dataMatrix(1:j_new-1,:);
            s_dataMatrixnew = dataMatrix(j_new+1:numDataVectors,:);
            dataMatrix = vertcat(f_dataMatrixnew,s_dataMatrixnew);
            f_queryMatrixnew = queryMatrix(1:j_new-1,:);
            s_queryMatrixnew = queryMatrix(j_new+1:numDataVectors,:);
            queryMatrix = vertcat(f_queryMatrixnew,s_queryMatrixnew);
            numDataVectors = size(dataMatrix,1);
            numQueryVectors = size(queryMatrix,1);
            change_counter = change_counter+1;
        end
   end
  %i_f_dataMatrixnew=dataMatrix(1:i-1,:);
  %i_s_dataMatrixnew=dataMatrix(i+1:numDataVectors,:);
  %dataMatrix=vertcat(i_f_dataMatrixnew,i_s_dataMatrixnew)
  %fprintf('*** Deleted points are ***');
  %deleted_points
  newdataMatrix=dataMatrix(1,:);
  final_points = [final_points;newdataMatrix];
  dataMatrix=dataMatrix(2:numDataVectors,:);
  %i_f_queryMatrixnew=queryMatrix(1:i-1,:);
  %i_s_queryMatrixnew=queryMatrix(i+1:numQueryVectors,:);
```

Code for aligning class averages with the selected class averages:

```
ls
#!/bin/bash


# This script will align teh class averages with respect to some selected class
#averages

#number of cycle
i=7


#the class number for which you want to do the alignment
c=30

#array of the selected class averages
array_selected=(00 01 07 08 09 11 13 14 15 16 24);
#array of the rest of the class averages
array_rest=(02 03 04 05 06 10 12 17 18 19 20 21 22 23 25 26 27 28 29);


#number of elements on the array_slected
tLen_s=${#array_selected[@]}
echo ${tLen_s}
#number of elements in the array_rest
tLen_r=${#array_rest[@]}
echo ${tLen_r}



#.....................................................
#check if i is a one digit number or a two digit number

if [${i} -le 9]
then
     i=0$i
else
     i=$i
fi

echo $i
mkdir siv-0${i}-sel


#copying the class averages of the selected classes to the -sel directory
for l in "${array_selected[@]}"
do
cp siv-0${i}-class-0${c}/0${c}-0${l}.avg siv-0${i}-sel
done

cd ..
i3mraselect-copy.sh ${i}

cd cycle-0${i}
mv siv-0${i}-sel siv-0${i}-sel.original
mv siv-0${i}-sel.i3d siv-0${i}-sel.i3d.original
mkdir siv-0${i}-sel


#copying the class averages of rest of the classes to the -sel directory
```

Code to figure out the relationship of each segment of filament and the micrograph

```
# This script will relate the segments with its original image name

def find_micrograph(file1,file2):
    """
    The first file is the required box file and the second one
    is the target star file.
    :param file1:
    :param file2:
    :return:
    """
    outfile = open('output_check.star','w')
    file1_h = open(file1,'r')
    file2_h = open(file2,'r')
    data1 = []
    data2 = []
    for line in file1_h:
        line = str(line).strip('\n').split(' ')
        data1.append(line)
    for line in file2_h:
        line = str(line).strip('\n').split(' ')
        data2.append(line)
    sum = 0.0
    for i in xrange(0, len(data1)):
        dt1 = data1[i]
        sum += int(dt1[0])
    print " Total number of rows in box file: ",sum
    val = 0
    for i in xrange(0,len(data1)):
        dt1 = data1[i]
        print "Starting from: ",val
        print "Number of micrographs:",dt1[0]
        ptr = 0
        for j in xrange(0, int(dt1[0])):
            ptr = val+j
            dt2 = data2[ptr]
            a_str = 'mic'+str(i+1)
            dt2[8] = a_str
            #dt2.append(a_str)
            dt2.append(dt1[1])
            dt2str = ' '.join(dt2)
            outfile.write(dt2str)
            outfile.write('\n')
        val = ptr+1
        print "Next start: ",val
    outfile.close()

if __name__ == '__main__':
    file1 = 'box.txt'
    file2 = 'start.star'
    find_micrograph(file1,file2)
```

Code to figure out if the segments are coming from same filament:

```
% Script to check whether the points in one box files fall on same filament
% or not
V=load('14may08c_8a1_00038gr_00012sq_00007hl_00003ex.box');
nx=size(V,1);
U=[];
for i=1:nx;
    U=[U; V(i,1) V(i,2)];
end
m=[];
for j=1:(nx-1);
    m=[m;(U(j+1,2)-U(j,2))/(U(j+1,1)-U(j,2))];
end
EPS=0.002;
X=[];
Eps=[];
J=1;
for i=1:nx-2
    eps=abs(m(i+1)-m(i));
    Eps=[Eps;eps];
end
for k=1:2
    X=[X;U(k,1) U(k,2) J];
end
    for l=1:nx-2
     if ((Eps(l)<EPS) || (Eps(l-1)>EPS))
       X=[X;[U(l+2,1) U(l+2,2) J]];
     else
         J=J+1;
         X=[X;[U(l+2,1) U(l+2,2) J]];
     end
    end

dlmwrite('14may08c_8a1_00038gr_00012sq_00007hl_00003ex.box.fil.track.txt',X,' ');
```

REFERENCES

[1] K. Murata and M. Wolf, "Cryo-electron microscopy for structural analysis of dynamic biological macromolecules," *Biochimica et Biophysica Acta (BBA)-General Subjects,* 2017.

[2] S. Jonić, "Computational methods for analyzing conformational variability of macromolecular complexes from cryo-electron microscopy images," *Current opinion in structural biology,* vol. 43, pp. 114--121, 2017.

[3] J. Frank, Electron tomography: methods for three-dimensional visualization of structures in the cell, Springer, 2008.

[4] Y. Cheng, N. Grigorieff, P. A. Penczek and T. Walz, "A primer to single-particle cryo-electron microscopy," *Cell,* vol. 161, no. 3, pp. 438--449, 2015.

[5] R. B. Rusu and S. Cousins, "3d is here: Point cloud library (pcl)," in *Robotics and automation (ICRA), 2011 IEEE International Conference on*, 2011.

[6] International AIDS Society Scientific Working Group on HIV Cure and others, "Towards an HIV cure: a global scientific strategy," *Nature reviews. Immunology,* vol. 12, no. 8, p. 607, 2012.

[7] P. R. Clapham and Á McKnight, "HIV-1 receptors and cell tropism," *British medical bulletin,* vol. 58, no. 1, pp. 43--59, 2001.

[8] P. R. Clapham and Á McKnight, "HIV-1 receptors and cell tropism," *British medical bulletin,* vol. 58, no. 1, pp. 43--59, 2001.

[9] P. Zhu, J. Liu, J. Bess, E. Chertova, J. D. Lifson, H. Grisé, G. A. Ofek, K. A. Taylor and K. H. Roux, "Distribution and three-dimensional structure of AIDS virus envelope spikes," *Nature,* vol. 441, no. 7095, p. 847, 2006.

[10] M. Radermacher, "Weighted Back-projection Methods," in *Electron tomography: methods for three-dimensional visualization of structures in the cell*, NY, Springer, 2007, pp. 245-273.

[11] P. Penczek and J. Frank, "Resolution in Electron Tomography," in *Electron tomography: methods for three-dimensional visualization of structures in the cell*, NY, Springer, 2007, pp. 307-330.

[12] S. J. Prince, Computer vision: models, learning, and inference, Cambridge University Press, 2012.

[13] M. Sonka, V. Hlavac and R. Boyle, Image processing, analysis, and machine vision, Cengage Learning, 2014.

[14] N. Volkmann, "methods for segmentation and interpretation of electron tomographic reconstruction," in *Methods in Enzymology*, vol. 483, Elsevier, 2010, pp. 31-45.

[15] D. L. Pham, C. Xu and J. L. Prince, "Current methods in medical image segmentation," *Annual review of biomedical engineering,* vol. 2, no. 1, pp. 315--337, 2000.

[16] N. Volkmann, "Chapter two-methods for segmentation and interpretation of electron tomographic reconstructions," *Methods in enzymology,* vol. 483, pp. 31--46, 2010.

[17] I. Griva, S. G. Nash and A. Sofer, Linear and nonlinear optimization, Siam, 2009.

[18] H. Winkler, P. Zhu, J. Liu, F. Ye, K. H. Roux and K. A. Taylor, "Tomographic subvolume alignment and subvolume classification applied to myosin V and SIV envelope spikes," *Journal of structural biology,* 2009.

[19] W. Hanspeter and K. A. Taylor, "Accurate marker-free alignment with simultaneous geometry determination and reconstruction of tilt series in electron tomography," *Ultramicroscopy,* vol. 106, no. 3, pp. 240--254, 2006.

[20] H. Winkler, P. Zhu, J. Liu, F. Ye, K. Roux and K. Taylor, "Tomograhic subvolume alignment and subvolume classification aplied to myosin V and SIV envelope spikes," *Journal of Structural Biology,* vol. 165, pp. 64-77, 2009.

[21] J. Dubochet, M. Adrian, J.-J. Chang, J.-C. Homo, J. Lepault, A. W. McDowall and P. Schultz, "Cryo-electron microscopy of vitrified specimens," *Quarterly reviews of biophysics,* vol. 21, no. 2, pp. 129--228, 1988.

[22] C. S. Potter, H. Chu, B. Frey, C. Green, N. Kisseberth, T. Madden, K. Miller, K. Nahrstedt, J. Pulokas, A. Reilein and others, "Leginon: a system for fully automated acquisition of 1000 electron micrographs a day," *Ultramicroscopy,* vol. 77, no. 3, pp. 153--161, 1999.

[23] D. N. Mastronarde, "SerialEM: a program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position," *Microscopy and Microanalysis,* vol. 9, no. S02, pp. 1182--1183, 2003.

[24] S. Nickell, F. Förster, A. Linaroudis, W. Del Net, F. Beck, R. Hegerl, W. Baumeister and J. Plitzko, "TOM software toolbox: acquisition and analysis for electron tomography," *Journal of structural biology,* vol. 149, no. 3, pp. 227--234, 2005.

[25] S. Q. Zheng, B. Keszthelyi, E. Branlund, J. M. Lyle, M. B. Braunfeld, J. W. Sedat and D. A. Agard, "UCSF tomography: an integrated software suite for real-time electron microscopic tomographic data collection, alignment, and reconstruction," *Journal of structural biology,* vol. 157, no. 1, pp. 138--147, 2007.

[26] H. Winkler and K. Taylor, "Accurate marker-free alignmnet with simultaneous geometry determination and reconstruction of tilt series in electron toography.," *Ultramicroscopy,* pp. 240-254, 2006.

[27] J.-J. Fernandez, "Computational methods of electron tomography," *Micron 43,* pp. 1010-1030, 2012.

[28] J. Berringman, R. Bryan, R. Freeman and K. R. Leonard, "Methods for specimen thickness determination in electron microscopy," *Ultramicroscopy 13,* pp. 351-364, 1984.

[29] M. Guckenberger, "Determination of a common origin in the micrographs of tilt series in three dimemsional electron microscopy," *Ultramicroscopy 9,* pp. 167-174, 1982.

[30] J. Nash, Compact Numerical Methods for Computers: Linear Algebra and Function Minimization, Bristol: Adam Hilger Ltd, 1979.

[31] M. Radermacher, "Weighted back-projection methods," in *Electron Tomography: Three-dimensional Imaging with the Transmission Electron Microscope*, New York, Plenum Press, 1992.

[32] R. Gordon, R. Bender and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," *Journal of theoretical Biology,* vol. 29, no. 7, pp. 471IN1477--476IN2481, 1970.

[33] J. Trampert and J.-J. Leveque, "Simultaneous iterative reconstruction technique: physical interpretation based on the generalized least squares solution," *Journal of Geophysical Research: Solid Earth,* vol. 95, no. B8, pp. 12553--12559, 1990.

[34] J. G. Colsher, "Iterative three-dimensional image reconstruction from tomographic projections," *Computer Graphics and Image Processing,* vol. 6, no. 6, pp. 513--537, 1977.

[35] C. Shannon, "Communication in the presence of noise," in *Proceedings of the IRE 37*, 1949.

[36] R. Crowther, D. J. DeRosier and A. Klug, "The reconstruction of a three-dimensional structure from projections and its application to electron microscopy," London, 1970.

[37] K. Sandberg, "Methods for image segmentation in celluar tomography," *DS in cell Biology,* vol. 79, pp. 769-798, 2007.

[38] A. Frangakis and R. Hegerl, "Segmentation of Three-dimensional Electron Tomographic Images," in *Electron Tomography*, 2010, pp. 353-370.

[39] N. Volkmann, D. Hanein, G. Ouyang, K. M. Trybus, D. J. DeRosier and S. Lowey, "Evidence for cleft closure in actomyosin upon ADP release," 2000.

[40] E. A. Hewat, N. Verdaguer, I. Fita, W. Blakemore, S. Brookes, A. King, J. Newman, E. Domingo, M. G. Mateu and D. I. Stuart, "Structure of the complex of an Fab fragment of a neutralizing antibody with foot-and-mouth disease virus: positioning of a highly mobile antigenic loop," 1997.

[41] N. Volkmann, "A novel three-dimensional variant of the watershed transform for segmentation of electron density maps," *Journal of structural biology,* vol. 138, no. 1, pp. 123--129, 2002.

[42] B. J. Marsh, D. N. Mastronarde, K. F. Buttle, K. E. Howell and J. R. McIntosh, "Organellar relationships in the Golgi region of the pancreatic beta cell line, HIT-T15, visualized by high resolution electron tomography," *Proceedings of the National Academy of Sciences,* vol. 98, no. 1, pp. 2399--2406, 2001.

[43] A. S. Frangakis and R. Hegerl, "Segmentation of two-and three-dimensional data from electron microscopy using eigenvector analysis," *Journal of Structural Biology,* vol. 138, no. 1, pp. 105--113, 2002.

[44] M. Kass, A. Witkin and D. Terzopoulos, "Snake: Active contour Models," *International Journal of Computer Vsion,* 1988.

[45] V. Caselles, R. Kimmel and G. Sapiro, "Geodesic active contours," *Computer Vision, 1995. Proceedings., Fifth International Conference on,* pp. 694--699, 1995.

[46] A. Betesaghi, G. Sapiro and S. Subramaniam, "An energy -based three dimensional segmenattion approach for the quantitative interpretation of electron tomograms," *IEEE Trans Image Process,* 2005.

[47] J. Böhm, A. S. Frangakis, R. Hegerl, S. Nickell, D. Typke and W. Baumeister, "Toward detecting and identifying macromolecules in a cellular context: template matching applied to electron tomograms," *Proceedings of the National Academy of Sciences,* vol. 97, no. 26, pp. 14245--14250, 2000.

[48] N. Hien and J. Qiang, "Shape-driven three dimentionl watersnake segmentation of biological membranes in electron tomography," *IEEE transactions on medical imaging,* vol. 27, 2008.

[49] M. Lebbink, J. C. W. Geerts and J. A. Koster, "Template matching as a tool for annottaion of tomograms of stained biological structures," *journal of Stuctural Biology,* 2007.

[50] N. T. Hieu, W. Marcel and V. Boomgaard, "Watersnakes:energy-driven watershed segmentation," *IEEE transaction on pattern analysis and machine intellegence,* 2003.

[51] M. Stolken, F. Beck, T. Haller, R. Hegerl, I. Gutsche, J. Crazo, W. Baumeister, S. Scheres and S. Nickell, "Maximum likelihood based classification of electron tomographic data," *Journal of Structural biology,* 2011.

[52] S. H. Scheres, "RELION: implementation of a Bayesian approach to cryo-EM structure determination," *Journal of structural biology,* vol. 180, no. 3, pp. 519--530, 2012.

[53] K. H. Roux and K. A. Taylor, "AIDS virus envelope spike structure," *Current opinion in structural biology,* vol. 17, no. 2, pp. 244--252, 2007.

[54] H. B. Gristick, L. von Boehmer, A. P. West Jr, M. Schamber, A. Gazumyan, J. Golijanin, M. S. Seaman, G. Fätkenheuer, F. Klein, M. C. Nussenzweig and others, "Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site," *Nature structural & molecular biology,* vol. 23, no. 10, p. 906, 2016.

[55] R. Diaz, J. R. William and D. L. Stokes, "Chapter Five-Fourier--Bessel Reconstruction of Helical Assemblies," *Methods in enzymology,* vol. 482, pp. 131--165, 2010.

[56] E. H. Egelman, "A robust algorithm for the reconstruction of helical filaments using single-particle methods," *Ultramicroscopy,* vol. 85, no. 4, pp. 225--234, 2000.

[57] D. DeRosier and P. Moore, "Reconstruction of three-dimensional images from electron micrographs of structures with helical symmetry," *Journal of molecular biology,* vol. 2, no. 355--369, p. 52, 1970.

[58] E. H. Egelman, "The iterative helical real space reconstruction method: surmounting the problems posed by real polymers," *Journal of structural biology,* vol. 157, no. 1, pp. 83--94, 2007.

[59] D. A. Winkelmann, T. S. Baker and I. Rayment, "Three-dimensional structure of myosin subfragment-1 from electron microscopy of sectioned crystals.," *The Journal of cell biology,* vol. 114, pp. 701--713, 1991.

[60] I. Rayment, W. R. Rypniewski, K. Schmidt-Bäse, D. A. Winkelmann, G. Wesenberg and H. M. Holden, "Myosin Subfragment-1: A Molecular Motor," *Science,* vol. 261, p. 2, 1993.

[61] I. Rayment, H. M. Holden, M. Whittaker, C. B. Yohn, M. Lorenz, K. C. Holmes and R. A. Milligan, "Structure of the actin-myosin complex and its implications for muscle contraction," *SCIENCE-NEW YORK THEN WASHINGTON,* vol. 261, pp. 58--58, 1993.

[62] M. A. Geeves, R. Fedorov and D. J. Manstein, "Molecular mechanism of actomyosin-based motility," *Cellular and Molecular Life Sciences CMLS,* vol. 62, no. 13, pp. 1462--1477, 2005.

[63] M. Lorenz and K. C. Holmes, "The actin-myosin interface," *Proceedings of the National Academy of Sciences,* vol. 107, no. 28, pp. 12529--12534, 2010.

[64] E. Behrmann, M. Müller, P. A. Penczek, H. G. Mannherz, D. J. Manstein and S. Raunser, "Structure of the rigor actin-tropomyosin-myosin complex," *Cell,* vol. 150, no. 2, pp. 327--338, 2012.

[65] J. von der Ecken, S. M. Heissler, S. Pathan-Chhatbar, D. J. Manstein and S. Raunser, "Cryo-EM structure of a human cytoplasmic actomyosin complex at near-atomic resolution," *Nature,* vol. 534, no. 7609, pp. 724--728, 2016.

[66] S. F. Wulf, V. Ropars, S. Fujita-Becker, M. Oster, G. Hofhaus, L. G. Trabuco, O. Pylypenko, H. L. Sweeney, A. M. Houdusse and R. R. Schröder, "Force-producing ADP state of myosin bound to actin," *Proceedings of the National Academy of Sciences,* vol. 113, no. 13, pp. E1844--E1852, 2016.

[67] A. Houdusse and H. L. Sweeney, "How myosin generates force on actin filaments," *Trends in biochemical sciences,* vol. 41, no. 12, pp. 989--997, 2016.

[68] J.-J. Fernandez, "Computational methods for electron tomography," *Micron,* vol. 43, no. 10, pp. 1010--1030, 2012.

[69] S. Nickell, C. Kofler, A. P. Leis and W. Baumeister, "A visual approach to proteomics," *Nature reviews Molecular cell biology,* vol. 7, no. 3, pp. 225--230, 2006.

[70] A. S. Frangakis, J. Böhm, F. Förster, S. Nickell, D. Nicastro, D. Typke, R. Hegerl and W. Baumeister, "Identification of macromolecular complexes in cryoelectron tomograms of phantom cells," *Proceedings of the National Academy of Sciences,* vol. 99, no. 22, pp. 14153--14158, 2002.

[71] C. Best, S. Nickell and W. Baumeister, "Localization of protein complexes by pattern recognition," *Methods in cell biology,* vol. 79, pp. 615--638, 2007.

[72] E. Chertova, B. J. Crise, D. R. Morcock, J. Bess, L. E. Henderson and J. D. Lifson, "Sites, mechanism of action and lack of reversibility of primate lentivirus inactivation by preferential covalent modification of virion internal proteins," *Current molecular medicine,* vol. 3, pp. 265--272, 2003.

[73] J. Rossio, M. Esser, K. Suryanarayana, D. Schneider, J. Bess, G. Vasquez, T. Wiltrout, E. Chertova, M. Grimes and Q. Sattentau, "Inactivation of human immunodeficiency virus type 1 infectivity with preservation of conformational and functional integrity of virion surface proteins.," *Journal of virology,* vol. 72, pp. 7992--8001, 1998.

[74] P. Zhu, H. Winkler, E. Chertova, K. A. Taylor and K. H. Roux, "Cryoelectron tomography of HIV-1 envelope spikes: further evidence for tripod-like legs," *PLoS pathogens,* vol. 4, no. 11, p. e1000203, 2008.

[75] P. Dube, P. Tavares, R. Lurz and M. Van Heel, "The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry.," *The EMBO journal,* vol. 12, no. 4, p. 1303, 1993.

[76] T. S. Huang, G. J. Yang and G. Y. Tang, "A fast two-dimensional median filtering algorithm," *Acoustics, Speech and Signal Processing, IEEE Transactions on,* vol. 27, pp. 13-18, 1979.

[77] T. Pun, D. Hochstrasser, R. D. Appel, M. Funk, V. Villars-Augsburger and C. Pellegrini, "Computerized classification of two-dimensional gel electrophoretograms by correspondence analysis and ascendant hierarchical clustering.," *Applied and theoretical electrophoresis: the official journal of the International Electrophoresis Society,* vol. 1, pp. 3-9, 1987.

[78] G. Hu, J. Liu, K. A. Taylor and K. H. Roux, "Structural comparison of HIV-1 envelope spikes with and without the V1/V2 loop," *Journal of virology,* vol. 85, no. 6, pp. 2741--2750, 2011.

[79] M. Dutta, J. Liu, K. H. Roux and K. A. Taylor, "Visualization of retroviral envelope spikes in complex with the V3 loop antibody 447-52D on intact viruses by cryo-electron tomography," *Journal of virology,* vol. 88, no. 21, pp. 12265--12275, 2014.

[80] S. H. Scheres, "Semi-automated selection of cryo-EM particles in RELION-1.3," *Journal of structural biology,* vol. 189, no. 2, pp. 114--122, 2015.

[81] F. Odronitz and M. Kollmar, "Drawing the tree of eukaryotic life based on the analysis of 2,269 manually annotated myosins from 328 species," *Genome biology,* vol. 8, no. 9, p. R196, 2007.

[82] H. L. Sweeney and A. Houdusse, "Structural and functional insights into the myosin motor mechanism," *Annual review of biophysics,* vol. 39, pp. 539--557, 2010.

[83] K. C. Holmes and M. A. Geeves, "The structural basis of muscle contraction," *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* vol. 355, no. 1396, pp. 419--431, 2000.

[84] Z. Hu, D. W. Taylor, M. K. Reedy, R. J. Edwards and K. A. Taylor, "Structure of myosin filaments from relaxed Lethocerus flight muscle by cryo-EM at 6 A resolution," *Science advances,* vol. 2, no. 9, p. e1600058, 2016.

[85] D. Mornet, R. Bertrand, P. Pantel, E. Audemard and R. Kassab, "Proteolytic approach to structure and function of actin recognition site in myosin heads," *Biochemistry,* vol. 20, no. 8, pp. 2110--2120, 1981.

[86] I. Rayment, W. R. Rypniewski, K. Schmidt-Bäse, R. Smith, D. R. Tomchick, M. M. Benning, D. A. Winkelmann, G. Wesenberg and H. M. Holden, "Three-dimensional structure of myosin subfragment-1: a molecular motor," *Science,* pp. 50--58, 1993.

[87] R. Lymn and E. W. Taylor, "Mechanism of adenosine triphosphate hydrolysis by actomyosin," *Biochemistry,* vol. 10, no. 25, pp. 4617--4624, 1971.

[88] M. A. Geeves and K. C. Holmes, "The molecular mechanism of muscle contraction," *Advances in protein chemistry,* vol. 71, pp. 161--193, 2005.

[89] P.-D. Coureux, A. L. Wells, J. Ménétrey, C. M. Yengo and others, "A structural state of the myosin V motor without bound nucleotide," *Nature,* vol. 425, no. 6956, p. 419, 2003.

[90] Y. Yang, S. Gourinath, M. Kovács, L. Nyitray, R. Reutzel, D. M. Himmel, E. O'Neall-Hennessey, L. Reshetnikova, A. G. Szent-Györgyi, J. H. Brown and others, "Rigor-like structures from muscle myosins reveal key mechanical elements in the transduction pathways of this allosteric motor.," *Structure,* vol. 15, no. 5, pp. 553--564, 2007.

[91] R. Dominguez, Y. Freyzon, K. M. Trybus and C. Cohen, "Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: visualization of the preâ≠“power stroke state.," *Cell,* vol. 94, no. 5, pp. 559--571, 1998.

[92] C. A. Smith and I. Rayment, "X-ray structure of the magnesium (II) · ADP· vanadate complex of the Dictyostelium discoideum myosin motor domain to 1.9 Å resolution," *Biochemistry,* vol. 35, no. 17, pp. 5404--5417, 1996.

[93] K. C. Holmes, I. Angert, F. J. Kull, W. Jahn and R. R. Schroder, "Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide.," *Nature,* vol. 6956, no. 423, p. 425, 2003.

[94] P.-D. Coureux, H. L. Sweeney and A. Houdusse, "Three myosin V structures delineate essential features of chemoâ≠≈mechanical transduction.," *The EMBO journal,* vol. 23, no. 23, pp. 4527--4537, 2004.

[95] A. Criddle, M. Geeves and T. Jeffries, "The use of actin labelled with N-(1-pyrenyl) iodoacetamide to study the interaction of actin with myosin subfragments and troponin/tropomyosin," *Biochemical Journal,* vol. 232, no. 2, pp. 343--349, 1985.

[96] S. Gourinath, D. M. Himmel, J. H. Brown, L. Reshetnikova, A. G. Szent-Györgyi and C. Cohen, "Crystal structure of scallop myosin S1 in the pre-power stroke state to 2.6 Ã… resolution: flexibility and function in the head.," *Structure,* vol. 11, no. 12, pp. 1621--1627, 2003.

[97] D. M. Himmel, S. Gourinath, L. Reshetnikova, Y. Shen, A. G. Szent-Györgyi and C. Cohen, "Crystallographic findings on the internally uncoupled and near-rigor states of myosin: further insights into the mechanics of the motor.," *Proceedings of the National Academy of Sciences,* vol. 99, no. 20, pp. 12645--12650, 2002.

[98] R. Milligan and P. Flicker, "Structural relationships of actin, myosin, and tropomyosin revealed by cryo-electron microscopy.," *The Journal of Cell Biology,* vol. 105, no. 1, pp. 29--39, 1987.

[99] J. Ménétrey, P. Llinas, J. Cicolari, G. Squires, X. Liu, A. Li, H. L. Sweeney and A. Houdusse, "The post-rigor structure of myosin VI and implications for the recovery stroke.," *The EMBO journal,* vol. 27, no. 1, pp. 244--252, 2008.

[100] A. Houdusse, A. G. Szent-Györgyi and C. Cohen, "Three conformational states of scallop myosin S1," *Proceedings of the National Academy of Sciences,* vol. 97, no. 21, pp. 11238--11243, 2000.

[101] N. Volkmann, G. Ouyang, K. M. Trybus, D. J. DeRosier, S. Lowey and D. Hanein, "Myosin isoforms show unique conformations in the actin-bound state," *Proceedings of the National Academy of Sciences,* vol. 100, no. 6, pp. 3227--3232, 2003.

[102] N. Volkmann, H. Liu, L. Hazelwood, E. B. Krementsova, S. Lowey, K. M. Trybus and D. Hanein, "The structural basis of myosin V processive movement as revealed by electron cryomicroscopy.," *Molecular cell,* vol. 19, no. 5, pp. 595--605, 2005.

[103] K. M. Trybus, "Biochemical studies of myosin," *Methods,* vol. 22, no. 4, pp. 327--335, 2000.

[104] J. D. Pardee and J. A. Spudich, "Purification of muscle actin.," *Methods in cell biology,* vol. 24, pp. 271--289, 1982.

[105] C. Suloway, J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C. S. Potter and B. Carragher, "Automated molecular microscopy: the new Leginon system.," *Journal of structural biology,* vol. 151, no. 1, pp. 41--60, 2005.

[106] G. C. Lander, S. M. Stagg, N. R. Voss, A. Cheng, D. Fellmann, J. Pulokas, C. Yoshioka, C. Irving, A. Mulder, P.-W. Lau and others, "Appion: an integrated, database-driven pipeline to facilitate EM image processing.," *Journal of structural biology,* vol. 166, no. 1, pp. 95--102, 2009.

[107] Z. Wang, C. F. Hryc, B. Bammes, P. V. Afonine, J. Jakana, D.-H. Chen, X. Liu, M. L. Baker, C. Kao, S. J. Ludtke and others, "An atomic model of brome mosaic virus using direct electron detection and real-space optimization," *Nature communications,* vol. 5, 2014.

[108] S. P. Mallick, B. Carragher, C. S. Potter and D. J. Kriegman, "ACE: automated CTF estimation.," *Ultramicroscopy,* vol. 104, no. 1, pp. 8--29, 2005.

[109] J. A. Mindell and N. Grigorieff, "Accurate determination of local defocus and specimen tilt in electron microscopy.," *Journal of structural biology,* vol. 142, no. 3, pp. 334--347, 2003.

[110] D. L. Clemens, P. Ge, B.-Y. Lee, M. A. Horwitz and Z. H. Zhou, "Atomic structure of T6SS reveals interlaced array essential to function.," *Cedll,* vol. 160, no. 5, pp. 940--951, 2015.

[111] S. H. cheres and S. Chen, "Prevention of overfitting in cryo-EM structure determination," *Nature methods,* vol. 9, no. 9, pp. 853--854, 2009.

[112] J. Fernandez, D. Luque, J. Caston and J. Carrascosa, "Sharpening high resolution information in single particle electron cryomicroscopy," *Journal of structural biology,* vol. 164, no. 1, pp. 170--175, 2008.

[113] A. Kucukelbir, F. J. Sigworth and H. D. Tagare, "Quantifying the local resolution of cryo-EM density maps.," *Nature methods,* vol. 11, no. 1, pp. 63--65, 2014.

[114] C. B. Bauer, H. M. Holden, J. B. Thoden, R. Smith and I. Rayment, "X-ray structures of the apo and MgATP-bound states of Dictyostelium discoideum myosin motor domain.," *Journal of Biological Chemistry,* vol. 275, no. 49, pp. 38494--38499, 2000.

[115] A. Fiser and A. Šali, "Modeller: generation and refinement of homology-based protein structure models," *Methods in enzymology,* vol. 374, pp. 461--491, 2003.

[116] F. DiMaio, Y. Song, X. Li, M. J. Brunner, C. Xu, V. Conticello, E. Egelman, T. C. Marlovits, Y. Cheng and D. Baker, "Atomic-accuracy models from 4.5-A cryo-electron microscopy data with density-guided iterative local refinement.," *Nature methods,* vol. 12, no. 4, pp. 361--365, 2015.

[117] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, "UCSF Chimera—a visualization system for exploratory research and analysis," *Journal of computational chemistry,* vol. 25, no. 13, pp. 1605--1612, 2004.

[118] T. Fujii and K. Namba, "Structure of actomyosin rigour complex at 5.2 Å resolution and insights into the ATPase cycle mechanism," *Nature communications,* vol. 8, 2017.

[119] M. M. Briggs and F. Schachat, "Early specialization of the superfast myosin in extraocular and laryngeal muscles.," *Journal of Experimental Biology,* vol. 203, no. 16, pp. 2485--2494, 2000.

[120] F. Wang, L. Chen, O. Arcucci, E. V. Harvey, B. Bowers, Y. Xu, J. A. Hammer and J. R. Sellers, "Effect of ADP and ionic strength on the kinetic and motile properties of recombinant mouse myosin V," *Journal of Biological Chemistry,* vol. 275, no. 6, pp. 4329--4335, 2000.

[121] S. Marston and E. W. Taylor, "Comparison of the myosin and actomyosin ATPase mechanisms of the four types of vertebrate muscles.," *Journal of molecular biology,* vol. 139, no. 4, pp. 573--600, 1980.

[122] K. C. Holmes, R. R. Schröder, H. Sweeney and A. Houdusse, "The structure of the rigor complex and its implications for the power stroke," *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 359, no. 1452, p. 1819, 2004.

[123] M. Whittaker, E. M. Wilson-Kubalek, J. E. Smith, L. Faust and others, "A 35-angstrom movement of smooth muscle myosin on ADP release," *Nature,* vol. 378, no. 6558, p. 748, 1995.

[124] A. C. Steven and U. Aebi, "The next ice age: cryo-electron tomography of intact cells," 2003.

[125] K. Song, "Partially Observed Tomographic Reconstruction Alignment Using Matrix Norm Minimization.," 2014.

[126] J. Böhm, A. S. Frangakis, R. Hegerl, S. Nickell, D. Typke and W. Baumeister, "Toward detecting and identifying macromolecules in a cellular context: template matching applied to electron tomograms," *Proceedings of the National Academy of Sciences,* vol. 97, no. 26, pp. 14245--14250, 2000.

[127] S. L. Hooper, K. H. Hobbs and J. B. Thuma, "Invertebrate muscles: thin and thick filament structure; molecular basis of contraction and its regulation, catch and asynchronous muscle," *Progress in neurobiology,* vol. 86, no. 2, pp. 72--127, 2008.

[128] N. Volkman, "A novel three dimensional variant of the watershed transform for segmentation of electron density maps.," *Journal of structural biology,* 2002.

[129] I. Rayment and H. M. Holden, "Myosin subfragment-1: structure and function of a molecular motor," *Current Opinion in Structural Biology,* vol. 3, no. 6, pp. 944--952, 1993.

[130] C. B. Wilen, J. C. Tilton and R. W. Doms, "HIV: cell binding and entry," *Cold Spring Harbor perspectives in medicine,* vol. 2, no. 8, p. a006866, 2012.

[131] K. A. Taylor and D. W. Taylor, "Formation of 2-D paracrystals of F-actin on phospholipid layers mixed with quarternary ammonium surfactants," *Journal of structural biology,* vol. 108, no. 2, pp. 140--147, 1992.

[132] G. Hu, J. Liu, K. H. Roux and T. K. A., "Structure of Simian Immunodeficiency Virus Envelope Spikes bound with CD4 and Monoclonal Antibody 36D5," *Journal of Virology,* pp. JVI--00134, 2017.

BIOGRAPHICAL SKETCH

Ms Chaity Banerjee Mukherjee is a final year Ph.D student at Florida State University working jointly with the Departments of Computer Science and Molecular Biophysics. Her research interests include unsupervised methods for segmentation in electron tomographic images, helical reconstruction in single particle electron microscopy and unsupervised learning methods in general. She hold a Master of Science degree in Computer Science from the Department of Computer Science at Florida State University, a Master's degree in Computer Applications and a Bachelor's degree in Statistics with Honors from India.

Prior to joining Florida State University for her graduate work, she was an assistant professor at Bengal Institute of Technology (BIT) in India. She has also taught extensively at different colleges in India at the undergraduate level in Computer Science before joining BIT as an assistant professor. She has also worked as a research intern at the Indian Statistical Institute in India, working on genetic algorithms with elitist models and their applications to set estimation.

At Florida State University she worked with Dr. Xiuwen Liu from the Computer Science department and Dr. Kenneth Taylor from the department of Molecular Biophysics. She has also collaborated with other members of Dr. Liu's lab and Dr. Taylor's lab.

In her spare time, she enjoys reading books both in her mother tongue and English. She is a avid lover of nature and likes to relax in the lap of mother nature whenever possible. She plans to stay in the academic in the future and hopes to make lasting contributions in her chosen field of research going forward