Chapter 1

Introduction

1.1 Structural Biology

Structural biology is the approach to understanding cell biology through determining the structures of objects found in the cell. These objects range from proteins and molecular machines to organelles. To accommodate the difference in scales of these objects, which span from nanometers to microns, a variety of complementary imaging techniques are used. The imaging techniques, together, determine the structures of molecular machines and cellular structures and provide information about their quantity, distribution, and location. Also, real-time information about processes within cells, sometimes in their native states, can be extracted.

1.2 Structure Determination Techniques

The main techniques used in structural biology are X-ray crystallography (XRC), nuclear magnetic resonance spectroscopy (NMR), light microscopy (LM), computational biology and cryo-electron microscopy (Cryo-EM). These methods work together in a complementary way to reveal information about a variety of structures in different physical conditions.

As of June 2008, XRC has produced by far the largest number of atomic models of proteins as compared to NMR and Cryo-EM according to the Protein Data Bank. XRC works well for proteins that can be crystallized and the structures often reach atomic resolution. The difficulty with this technique is that the crystallization process requires trying numerous conditions of temperature, pH, and buffer concentrations to produce a crystal that diffracts to sufficiently high resolution. These conditions result in structures of the proteins in non-native states. Once such crystals can be grown, X-ray diffraction patterns are then recorded, giving the Fourier amplitudes of the crystal. Next, the phases need to be determined ("phase problem") before the structures can be obtained.

NMR also produces atomic resolution structures but is limited to molecular masses of less than 50 kDa, which includes only the smaller proteins. On rare occasions, larger protein structures may be determined, for example, an 82-kDa enzyme in 2005 (Tugarinov, Choy et al. 2005).

LM allows for real-time imaging of live cells. Traditionally, this technique was limited in resolution by the wavelength of light and thus could not reveal the workings of the cell to higher resolutions. Recently, "super-resolution" techniques have been developed to surpass the diffraction limit as described in a recent review (Hell 2007) and have reached sub 100-nm resolutions (Juette, Gould et al. 2008; Schmidt, Wurm et al. 2008).

Computational biology techniques include comparative structure prediction, where protein structures are predicted using known structures as a reference, and *de novo* predictions in which no assumptions are made about the structures.

1.3 Cryo-Electron Microscopy

Cryo-EM delivers structures that span the resolution and size range between the atomic models provided by XRC or NMR, and the imaging of entire cells by LM. Its advantages are that samples are easily obtained, and when used in conjunction with plunge freezing (Dubochet and Mcdowall 1981) using a Vitrobot (Iancu, Tivol et al. 2006), the proteins or cells can be studied in their near-native state. This is achieved by first having the sample in a buffer which is spread onto a carbon film. The film is then plunged into liquid ethane, which cools the sample quickly enough so that the water in the sample is frozen in vitreous form (Angell 2004). This prevents the crystallization of water, which would damage the sample. The sample is then inserted into the microscope and imaged with electrons, which are scattered and then focused by electron lenses to form an image that is recorded on film or on a digital camera such as a charged-coupled device (CCD) or CMOS detector. An advantage of cryo-EM over XRC is the recording of images instead of just amplitudes. However, cryo-EM samples are limited to a thickness of ~ ½ micron (Lucic, Forster et al. 2005) to prevent multiple scattering of electrons within the cell. Also, the electron beam causes significant damage to the sample and thus the electron dose has to be kept low in order to reduce damage. This low dose results in images with low signal-to-noise ratios (SNRs).

There are several cryo-EM techniques available. Electron crystallography (EC) is used when 2D crystals of proteins, which are one unit cell thick, can be formed. In such situations, near-atomic resolution has been achieved (Henderson, Baldwin et al. 1990).

Similarly, the imaging of helical or tubular crystals also allows for atomic structures to be determined (Unwin 2005).

Electron cryo-tomography (ECT) is a technique which allows for the study of large structures and even entire small cells (Henderson and Jensen 2006). ECT can image the sample to high resolution in its native state, which is not possible with XRC, NMR, and LM. ECT complements LM because cells can be first observed *in vivo* with LM and then plunge-frozen to be imaged by ECT (Briegel, Ding et al. 2008). The ECT technique images cells from various tilt angles along one or more tilt axes. In theory, this technique would allow for a full reconstruction of a cell if the tilt angles ranging from -90° to +90° could be used. In practice, a maximum tilt of about ±65° is used, resulting in an artifact known as the "missing wedge or pyramid" (Iancu, Wright et al. 2005) in reconstructions of the cell. This artifact arises due to a wedge or pyramid of missing information in Fourier space. Another limitation of this technique is that the maximum dose to which the sample can be exposed has to be shared by all images of the tilt series in order to prevent information loss due to structural damage by the beam.

Lastly, Single particle analysis (SPA) is a technique in which many identical copies of a specimen are imaged. The particles in solution are applied to a grid and plunge-frozen. These grids are imaged resulting ideally in random views of these particles from all angles, although certain types of particles have preferred orientations. The images obtained from electron microscopes are noisy due to the low electron dose that can be tolerated by the sample. Fortunately, the information from these views can be averaged to improve the SNR and produce high-resolution reconstructions of particles through Fourier reconstruction techniques (Crowther, Amos et al. 1970).

1.4 Reconstruction Theory

The reconstruction process can be simplified into three main stages (Figure 1-1). First, information about the object to be reconstructed is obtained in the form of raw projection images in various orientations, which are described by Euler angles and determined by the common-line method (Fuller, Butcher et al. 1996) for particles of high symmetry, or by 3D projection matching (Penczek, Grassucci et al. 1994). Secondly, corrected images are produced by the correction of raw images, which removes artifacts that were introduced during the imaging process due to the point spread function (PSF). This process is called contrast transfer function (CTF) correction and is performed by taking the 2D Fourier transform (FT) of a raw image and dividing it by the CTF, which is the FT of the PSF, before taking the inverse FT to get a corrected image. Thirdly, a 3D real-space reconstruction of the object is determined by a reconstruction algorithm.

To a good approximation, corrected images are projections of the object, which are equivalent to the inverse FT of central slices in the 3D FT of the object being reconstructed (Bragg 1929):

$$\begin{aligned} p(x,y) &= \int \rho(x,y,z)dz \\ &= \int \iiint F(X,Y,Z)e^{i2\pi(xX+yY+zZ)}dXdYdZdz \\ &= \iiint F(X,Y,Z)e^{i2\pi(xX+yY)}\delta(Z)dXdYdZ \\ &= \iint F(X,Y,0)e^{i2\pi(xX+yY)}dXdY \end{aligned} \tag{1}$$

where $p(x,y)$ is a projection of the object along the z-axis, $\rho(x,y,z)$ is the density of the object and $F(X,Y,Z)$ is the 3D FT of the object. This derivation can be generalized for projections in all possible directions and is called the projection theorem.

Using the property above, the 3D FT of the object can be determined by adding many central slices with different orientations using Whittaker-Shannon interpolation (Whittaker 1915; Shannon 1949) or by Fourier-Bessel synthesis (Klug, Crick et al. 1958). Once the 3D FT has been sufficiently sampled, the inverse FT can be calculated to give the reconstruction of the object.

1.5 Resolution Measures

When discussing resolution, a high resolution (or spatial frequency) corresponds to the resolvability of features separated by small distances, while a low resolution (or spatial frequency) corresponds to the resolvability of features separated by large distances; Atomic resolution refers to the resolvability of the distances between atoms while near-atomic resolution, which is slightly lower, implies that atomic models can be fit with the help of additional information such as the protein sequence.

In SPA, the quality of a reconstruction is measured in terms of the resolution achieved, which can be measured numerically or visually. Both these methods are subjective and can be manipulated to provide better or worse results by adjusting certain parameters.

The most commonly used numerical resolution measure is the Fourier shell coefficient (FSC) (Harauz and Van Heel 1986). In order to calculate the FSC, a data set consisting of a large number of images is split randomly into two halves. Independent reconstructions of each half of the data set are generated. The two reconstructions are then compared by calculating the value of the FSC at each spatial frequency

$$FSC(|s|) = \frac{\sum_i |F_1^i||F_2^i| Cos(\phi_1^i - \phi_2^i)}{\sqrt{\sum_i |F_1^i|^2 \sum_i |F_2^i|^2}} \tag{2}$$

where i enumerates the set of points found at spatial frequency s in the 3D FTs of the two reconstructions, F_1^i and F_2^i represent the values of the Fourier coefficients for each half of the data set and ϕ_1^i and ϕ_2^i represent their phases.

A variety of factors (van Heel and Schatz 2005) can affect the value of the FSC resolution, such as the number of additional voxels in the reconstruction which are in excess to the object being reconstructed. Changing the size of the volume containing the reconstruction adjusts the amount of additional voxels. Other factors that affect the measured resolution include the types of masks and how sharp these masks are, and most importantly the FSC threshold value, which indicates the maximum resolution of the reconstruction.

The resolution of a reconstruction can be determined visually if the resolution is sufficiently high. This has recently been possible with high-resolution reconstructions of icosahedral virus particles at ~ 4 Å resolution (Jiang, Baker et al. 2008; Yu, Jin et al. 2008; Zhang, Settembre et al. 2008). Table 1-1 gives a list of biological structural features that can be observed at various resolutions. High-resolution details can be enhanced to a certain extent by applying an "inverse" B-factor to the reconstructions, which adjusts the weighting of higher-resolution information by multiplication with the following factor:

$$e^{Bs^2} \tag{3}$$

where *B* is the B-factor and *s* is the spatial frequency.

However, it is important to note the FSC is not affected by the B-factor:

$$\begin{aligned} FSC_B(|s|) &= \frac{\sum_i | F_1^i e^{Bs^2} || F_2^i e^{Bs^2} | Cos(\phi_1^i - \phi_2^i)}{\sqrt{\sum_i | F_1^i e^{Bs^2} |^2 \sum_i | F_2^i e^{Bs^2} |^2}} \\ &= \frac{\sum_i | F_1^i || F_2^i | Cos(\phi_1^i - \phi_2^i)}{\sqrt{\sum_i | F_1^i |^2 \sum_i | F_2^i |^2}} \left(\frac{e^{Bs^2}}{e^{Bs^2}} \right)^2 \\ &= FSC(|s|) \end{aligned} \tag{4}$$

In addition, since the FSC calculation uses half datasets while the visually determined resolution uses the entire dataset, the latter gives a higher measure of resolution.

1.6 Resolution Limitations

There are two sets of resolution limitations involved in the SPA process. The first set consists of instrumentation limitations. These include incoherent beam sources, specimen preservation during the imaging process, and specimen charging by the electron beam, among others. The second set of resolution limitations consist of processing limitations, which include orientation, origin, and defocus determination and lack of computational power. There also exists the depth of field or equivalently the Ewald sphere curvature problem, which can be solved both computationally and instrumentally. Further discussion of these resolution limitations can be found in cryo-EM reviews (Baker, Olson et al. 1999; van Heel, Gowen et al. 2000).

1.7 Instrumentation Progress

Better electron sources and energy filters, more stable cooling stages, and larger, more sensitive CCD cameras have allowed structure determination by cryo-EM to approach near-atomic resolution by improving the recording of higher-resolution information with fewer artifacts and increasing data throughput.

In modern electron microscopes, the electron beam source is a highly coherent field emission electron gun (FEG). The FEG consists of a pointed field emission tip placed near a positive electrode. This causes a strong electric field to form which allows electrons to overcome the work function of the filament (usually tungsten) and be emitted. FEGs are better than previous electron sources, such as the thermionic W or LaB_6 and Schottky ZrO/W guns. They are spatially and temporally more coherent because they produce better point electron sources and are colder, which reduces the thermal energy spread, leading to more monochromatic beams, respectively. The electron beams are focused with improved electron lenses that have lower spherical aberrations than previously. Samples are cooled by liquid nitrogen in ECT (Iancu, Wright et al. 2006) and by liquid helium (Fujiyoshi, Mizusaki et al. 1991) in SPA (van Heel, Gowen et al. 2000) and EC (Hite, Raunser et al. 2007) to reduce beam damage. In addition, energy filters are used to ensure that only elastically scattered electrons are recorded on the CCD. Furthermore, the entire data collection process can be automated (Potter, Chu et al. 1999).

1.8 Progress in Processing Techniques

Although the fundamentals of the reconstruction process are still the same, there now exist several popular software packages that are used in the reconstruction of virus particles by single particle analysis. For example, IMIRS (Liang, Ke et al. 2002) utilizes the Fourier-Bessel synthesis method and was written for Microsoft Windows XP, while EMAN (Ludtke, Baldwin et al. 1999), FREALIGN (Grigorieff 2007) and Bsoft (Heymann 2001) are Cartesian-coordinate, UNIX-based packages which use a variety of interpolations which are approximations of a full 3D Fourier interpolation (Whittaker 1915; Shannon 1949).

Fundamental improvements to the reconstruction process include CTF correction of images and more sophisticated orientation determination algorithms, among others.

Improvements in computer hardware have also allowed for larger reconstructions to be computed because of 64-bit memory addressing and faster CPU speeds.

1.9 Approaching Atomic Resolution by Cryo-EM

With these advances, near-atomic resolution of biological structures was first achieved using EC (Henderson, Baldwin et al. 1990) and then by helical or tubular reconstructions (Unwin 2005). Thus the next technique by Cryo-EM that will approach these high resolutions is SPA. The alignment and orientation determination process, which is not required for EC and helical reconstructions, is non trivial, but using particles with large masses lessens this obstacle. In addition, high physical symmetry allows for fewer particles to be used in the reconstruction process. Thus large icosahedral virus particles are the best candidates for SPA to achieve atomic models.

1.10 Icosahedral Virus Structures

Virus capsids are composed of many identical copies of one or a few different capsid proteins, and as a result, the genetic material of the virus can be smaller and the production of a complete virus capsid quicker (Crick and Watson 1956; Caspar and Klug 1962). This use of identical proteins usually results in capsids of helical symmetry, the best known example being the tobacco mosaic virus (Bloomer, Champness et al. 1978), or icosahedral symmetry, for example, the herpes simplex virus (Zhou, Dougherty et al. 2000). Icosahedral symmetry is the naturally preferred structure for containing the virus genome because it provides the largest volume using the fewest capsid units possible. Each of the 20 triangular faces of the icosahedral structure consists of three asymmetric units. Furthermore, each of these asymmetric units can be composed of a number of either identical or different subunits. The triangulation (T) number (Caspar and Klug 1962) specifies the number of subunits in each asymmetric unit.

Any image of an icosahedral virus particle can be used 60 times in the reconstruction process because icosahedral virus particles possess 60-fold symmetry. Alternatively, only 1/60th of the total information is required to reconstruct a virus particle. The latter approach is more difficult to achieve in reconstruction algorithms but some progress has been made towards it with the Fourier-Bessel reconstruction algorithm (Crowther, Amos et al. 1970) which uses 1/10th of the information by aligning the 5-fold axis along the z-axis and utilizing 2-fold symmetry which results in information being required only between the azimuthal angles of 0° and 36° in a cylindrical coordinate system. Likewise, orientation determination of icosahedral particles is also easier due to the symmetry which allows for the use of the common-line method (Fuller, Butcher et al. 1996), which compares intersections of the 60 central slices from each image to derive the correct orientation.

1.11 Viruses

Virus structures are being intensively researched, as shown by a recent PubMed search for “virus structure”, which yielded over 37,000 hits. An old review of solved icosahedral virus structures listed over 175 reconstructions (Baker, Olson et al. 1999), further underlining the effort being invested.

Viruses consist of genetic material enclosed in capsids, with or without envelopes. A classification scheme was proposed (Baltimore 1971) which separated viruses into classes depending on the type of genetic material contained within the capsids. Viruses infect host cells either by being transported through the cellular membranes, or by injecting their genetic material, in the form of DNA or RNA, into the cell. If viral DNA is introduced into the cell, it is transcribed to produce RNA. The viral RNA is subsequently translated into proteins that form the virus capsid. Despite detailed understanding, there is still much to learn and exploit, for example, targeted viruses can be used to cause cancer cells to kill themselves (Ito, Aoki et al. 2006).

Viruses cause a wide range of diseases, such as AIDS (human immunodeficiency virus), cold sores (herpes virus) and even cancer (papilloma virus) (zur Hausen 2002). Greater understanding of viruses aids us in our attempts to cure or prevent certain diseases, which in turn would allow us to improve or save the lives of millions of people. While reconstructions that achieve a resolution of ~ 3.5 Å allow atomic models to be fit within the density, higher resolutions of ~ 2 Å allow predictions of the behavior and location of the interaction surfaces of virus capsids, which in turn guide drug design in producing drugs that target these surfaces by disrupting the original interaction surface properties, thereby disrupting assembly of capsids.

In addition, the study of viruses as simplified cellular machines continues to improve our understanding of evolution, for example, by understanding that viruses may be agents in horizontal gene transfer. These studies have also improved our knowledge of cell biology.

1.12 Approach Atomic Resolution by Single Particle Analysis

3D reconstructions of virus particles from electron micrographs by Fourier synthesis were first accomplished in 1970 (Crowther, Amos et al. 1970). Since then, reconstruction algorithms have improved and matured, resulting in sub-nanometer resolution in 1997 (Bottcher, Wynne et al. 1997; Conway, Cheng et al. 1997; Trus, Roden et al. 1997).

According to Glaeser (Glaeser 1999), achieving atomic resolution, which requires the determination of orientations from 10^6 images, would require an estimated 10^{23} floating point operations, which would take the world's fastest super computer with a maximum processing power of 1.375 PFlops (June 2008, www.top500.org) over two years to complete.. Fortunately, the 60-fold symmetry of icosahedral viruses reduces that number by nearly two orders of magnitude.

When I first began my thesis work, several factors that limited the resolution of SPA reconstruction had not been addressed. I attempted to address two of these challenges, namely the lack of computing power in reconstruction algorithms and the depth of field or equivalently, the Ewald sphere curvature problem (DeRosier 2000).

The resolutions of SPA reconstructions have improved significantly in the last few years and towards the end of my thesis work in 2008, three structures reached near-atomic resolution (Jiang, Baker et al. 2008; Yu, Jin et al. 2008; Zhang, Settembre et al. 2008).

1.13 Computational Complexity of 3D Reconstruction Algorithm

3D reconstructions are highly computationally and memory intensive. Despite the increasing amounts of memory available, increasing speeds of processors, and the increase in number of cores and processors per computer, the computation requirements are still very high when trying to perform reconstructions of very large viruses to high resolutions.

The basic reconstruction algorithm requires that the 3D FT be held in memory as samples are applied to it, which results in a $O(n^3)$ memory requirement where n is the length of one side of the transform. Due to the large memory requirements, it is necessary that the computer performing the reconstruction possess enough RAM to meet this requirement. Computers lacking the necessary RAM will require swapping of memory, a process that utilizes the hard disk as additional memory. As hard disk access is several orders of magnitude slower than RAM access, the resulting computation would not be completed in a reasonable amount of time. The number and size of images being used in reconstructions are very large when high-resolution reconstructions are required, due to the smaller pixel sizes and the higher sampling of images. In practice, for a reconstruction of a virus particle using 1k x 1k images, the memory requirements would be approximately 16, 20, and 30 GB for EMAN (Ludtke, Baldwin et al. 1999), Bsoft (Heymann 2001), and FREALIGN (Grigorieff 2007), respectively. IMIRS (Liang, Ke et al. 2002), which is highly optimized, would require less than 2GB. Currently, 64-bit systems allow for access of sufficient memory for even the largest of virus particles. Thus, memory requirements are a cost issue, which can be overcome with purchasing of sufficient RAM.

The computation of the basic reconstruction algorithm consists of applying the value of each pixel of the 2D FT of the images to the 3D FT making this a $O(mn^2)$ computation problem where m is the number of images and n is the length of one side of the 2D FT of an image. While the problem is tractable, it does take a significant amount of time for high-resolution structures of large virus particles, once again, due to the larger images used in the reconstruction process. While it may seem that purchasing faster computers can likewise solve the computation problem, it is not a good solution because CPU speeds have already started to plateau. Fortunately, the computation problem is trivially parallelizable for the most part and thus parallel and distributed computation are possible solutions to solve the problem efficiently.

1.14 Parallel Computation

One approach is the parallelization of the reconstruction process, which allows for the utilization of multiple cores or processors on a single computer or supercomputer that has shared memory and fast access to this memory. Parallelization takes advantage of the recent trend by CPU chip manufacturers to increase the number of cores per CPU instead of increasing the speed of the processors. A program that is multi-threaded will be able to process multiple calculations simultaneously and would take advantage of these additional resources. This multi-threaded approach which utilizes shared memory would require only one copy of the 3D FT to be stored in memory while allowing for the computation time to be reduced due to the increased number of threads performing calculations on the various processors or cores without any significant additional memory requirements.

The most significant drawback to this approach is that when too many threads are utilized, a bottleneck of the process occurs in the write access to the large memory holding the 3D FT of the object. The number of memory accesses, due to sample values being applied to the 3D FT, would be $O(mn^2)$ where m is the number of images used in the reconstruction and n is the length of one side of the image. These memory accesses are essentially random in their access pattern of the memory and thus require that the shared memory be locked before changes are made to it to prevent race conditions where changes are inadvertently lost when multiple threads access the same memory location at the same time. This bottleneck is encountered when the rate of samples calculated, which scale linearly with the number of cores, exceeds the rate at which samples are applied to the 3D FT, which is limited to the RAM access rate that is a constant. At this point, additional cores cannot accelerate the reconstruction process any further because the additional threads would spend increasing amounts of time waiting for access to the shared memory.

Implementations of parallel optimizations to the reconstruction algorithms using multi-threading libraries, such as the pthreads library, are described in Chapter 3.

1.15 Distributed Computation

The second available approach to reducing computation time is by distributed computation. This means that individual processes, which are executed on multiple computers with the necessary memory requirements, can take a subset of the data and perform independent reconstructions that are later combined to produce the full reconstruction. It is also possible to execute multiple processes on a single machine with the requisite number of processors and the required multiples of RAM.

Many structural biology laboratories possess mixtures of heterogeneous workstations purchased individually or in small sets for laboratory personnel, which constitutes a wealth of underutilized computation capacity. This is an ideal situation for this using the distributed approach to solve the computational problem.

This untapped resource was previously unworkable because of the effort required of researchers to log in to multiple computers and manually distribute jobs across computers with different operating systems. In addition, custom scripts were needed to submit jobs one after another through the night or weekend and watch for their completion. Lastly, computer usage had to be coordinated with laboratory colleagues so as not to impede their own computation efforts. Despite such efforts on the parts of some researchers, most workstations were still only used to a small fraction of their capacity due to the difficulty of manually managing multiple tasks on multiple workstations.

In 2003, only a few distributed systems were available, including Open PBS from Veridian Systems, Condor (Tannenbaum and Litzkow 1995), and BOINC, the Berkeley Open Infrastructure for Network Computing, which mediates the SETI@home project (Anderson, Cobb et al. 2002). These systems did not meet all our requirements for processing jobs that had extensive read, write, and memory requirements, were computationally intensive; had little or no fault tolerance, needed no changes to source code, and enabled desktop harvesting.

Peach, a distributed computation system, which is described in detail in Chapter 2, was developed in order to meet those requirements and also be simple to use and administer, scalable, secure, robust, and as compatible as possible with the existing hardware and software in structural biology. Essentially, Peach allows for multiple jobs to be submitted to a heterogeneous cluster of computers and utilizes clock cycles of idle computers. This distributed approach requires many powerful computers with sufficient RAM when used in the reconstruction of large virus particles. Furthermore, distributed computation is also applicable to a wide range of tasks in image processing.

The combination of the information, after the independent reconstructions are completed, requires $O(\log(c)\,n^3)$ steps, where c is the number of separate computers used in the reconstruction and n is the length of one side of the reconstruction, which is independent of the total number of images. This combination by binary merging is significantly quicker than in the parallel approach because results have already been accumulated by the individual reconstructions before being combined and can be combined in parallel, i.e., n reconstructions can be merged by $\frac{1}{2}n$ individual processes repeatedly until the final reconstruction is left and thus require only $\log(c)$ stages of combinations. If there is availability of computers with sufficient RAM, then distributed computation is a better solution than the parallel approach because it does not encounter the memory access bottleneck.

1.16 Hybrid Approach

A hybrid approach, using both parallel and distributed approaches together, would be the best solution in the reconstruction of large viruses as it utilizes computing resources maximally by using all available cores on all available computers. This approach is feasible with new implementations (Chapter 3) of the reconstruction algorithms in Bsoft (Heymann 2001) and EMAN (Ludtke, Baldwin et al. 1999) that possess capabilities for both parallel and distributed computation, and which may be used in conjunction with a suitable distributed computation system such as Peach (Chapter 2) or by processing on several multi-core nodes of a supercomputer.

1.17 Depth of Field and Ewald Sphere Curvature

As mentioned above, one of the resolution limitations of SPA of large virus particles is the depth of field problem, or equivalently, the Ewald sphere curvature. The depth of field, which is the distance over which the sample is in focus, is sometimes mistakenly called the depth of focus, which corresponds to the distance over which the recorded image is in focus (Fultz and Howe 2002). The depth of field can be geometrically calculated according to the following formula:

$$D = \tfrac{d}{\alpha} \tag{5}$$

where D is the depth of field, d is the resolution, and α is the aperture angle of the lens. For a typical transmission electron microscope, the aperture angle α is ~ 10^{-3} rad and the resolution d ~ 5 Å giving a depth of field D of ~ 5000 Å or a ½ micron.

The geometric estimate, however, cannot be applied to high-resolution phase contrast information because small defocus changes Δd, on the order of 10^2 Å, affect the image intensity distribution (Reimer 1997). This effect is due to the wave aberration

$$\chi = \tfrac{\pi}{2} C_s \lambda^3 s^4 - \pi \Delta f \, \lambda s^2 \tag{6}$$

where C_s is the spherical aberration, λ is the electron wavelength, Δf is the defocus value, and s is the spatial frequency. We find that $\Delta d \le \tfrac{1}{\lambda s^2}$ when setting the change in the wave aberration to be less than π. For a resolution of 3.8 Å at 300 kV, where λ ~ 0.02 Å and s ~ 0.263 Å^{-1}, Δd is ~ 720 Å which is approximately the diameters of the CPV, $\varepsilon 15$, and DLP capsids.

The defocus gradient and Ewald sphere curvature problems were shown to be equivalent first in 1978 (Amelinckx, Gevers et al. 1978), then in 2000 (DeRosier 2000) and again in 2004 (Wan, Chiu et al. 2004). Further elaboration about their equivalence qualitatively and quantitatively is provided below.

Firstly to understand the situation qualitatively, consider the Ewald sphere in XRC. Reciprocal lattice points have dimensions that are inversely proportional to the size of the crystal. If the crystal thickness in one direction is large, then the dimension of the reciprocal lattice point in that direction becomes small. Likewise, if we have a thin crystal, then the dimension of the reciprocal lattice point in that direction becomes very long and is known as a reciprocal rod or “rel-rod”. The intersections of the Ewald sphere and reciprocal lattice points are where scattering occurs. Take the situation where reciprocal lattice points lie along the XY plane. If the incoming beam is along the Z-axis, then at high resolutions along the plane, there will be reciprocal lattice points which do not intersect the Ewald sphere. If the crystal is thin, then the rel-rods stretch and intersect the Ewald sphere. Alternatively, if instead a higher voltage is used, the Ewald sphere flattens or has a larger radius. In this situation, the reciprocal lattice points also intersect the Ewald sphere without needing to be rel-rods. Thus, in this situation having a crystal thin enough will render the Ewald sphere curvature negligible. Conversely, if a crystal is thick enough, then the Ewald sphere curvature cannot be neglected at high resolution.

The XRC example can be viewed as a simple case of what occurs in electron microscopy (EM). The difference is that in EM, the sample is not crystalline and thus, the Fourier amplitudes vary continuously in all directions, as opposed to discrete reciprocal lattice points in crystallography, and scattering occurs at all points on the Ewald sphere. Analogously, the variations in the FT are quicker with thicker EM samples and slower for thinner EM samples in the direction of the thickness. Thus, when larger virus particles are imaged, the effect of the Ewald sphere is significant and should not be ignored. Alternatively, if a small virus particle is imaged, the variations of the FT are slower, so the Ewald sphere curvature is less significant.

This can also be explained quantitatively. First let us take a point $(X_0,Y_0,Z(X_0,Y_0))$ from the 3D FT of an object of radius R, where $Z(X,Y) \approx \frac{1}{2}\lambda(X^2+Y^2)$. The object can be broken down in real space as a set of thin slabs at different defocus values. During the recording of the image, all the slabs contribute to $(X_0,Y_0,Z(X_0,Y_0))$ but with different defocus values or, equivalently, different phase delays due to the wave aberration χ (equation 6). The difference in the defocus values, Δd, from the center defocus, result in phase delays of $-\pi\Delta d\lambda s^2$ with respect to the defocus value at the center of the object. Thus contributions from the top slab of the object will have an additional phase delay of $\pi R\lambda s^2$, as compared to the slab at the center.

Taking the alternative view, we assume a single defocus for the entire object. Then, the slabs have contributed to $(X_0,Y_0,Z(X_0,Y_0))$ without additional phase delays as all slabs are of the same defocus. However, for each slab to have no additional phase delays, the slabs would have to be located at the center of the object. Since the slabs are physically located away from the center, each slab has a phase shift due to its location, and thus a phase delay according to the Fourier shift theorem, which states that $F(X,Y,Z)$ becomes $F(X,Y,Z)e^{i2\pi zZ}$ when an object is shifted in position by a value z. The separate slabs, with physical shifts z, thus have phase delays of $2\pi zZ \approx 2\pi z(\frac{1}{2}\lambda s^2) = \pi z\lambda s^2$ where Z is due to the curvature of the Ewald sphere. Once again, a phase delay of $\pi R\lambda s^2$ occurs between contributions at the top slab of the object and the center. This phase delay would not have existed if the Ewald sphere curvature were negligible or equivalently when Z is set to 0.

The phase delays are identical in both cases. This indicates that considering the defocus gradient over an object is equivalent to taking into account the Ewald sphere curvature while assuming a single defocus value. Alternatively, if an object possesses an insignificant defocus gradient, then curvature of the Ewald sphere can be ignored.

1.18 Virus Structures Limited by Ewald Sphere Curvature

Various studies have shown that the Ewald sphere curvature is significant for particles ~ 700 Å or greater in diameter, at near-atomic resolution. In 2008, three virus structures of this diameter were reconstructed to near-atomic resolution of ~ 4 Å (Jiang, Baker et al. 2008; Yu, Jin et al. 2008; Zhang, Settembre et al. 2008). According to Jensen and Kornberg's envelope function (Jensen and Kornberg 2000), half of the signal in a conventional reconstruction of such a large virus at 300 kV would be lost due to curvature of the Ewald sphere by ~ 3.5 Å resolution. Likewise, DeRosier's formula (DeRosier 2000) predicts that the curvature problem in this same situation would become significantly limiting by ~ 3.3 Å resolution.

Thus, the Ewald sphere curvature will be most significant for three families of large icosahedral viruses, namely, the adenoviridae, herpesviridae and reoviridae, as their diameters are large enough that the curvature of the Ewald sphere will become significant at near-atomic resolution (Table 1-2). These families are medically important as they are responsible for a large range of diseases; for example, respiratory tract infections, conjunctivitis, hemorrhagic cystitis, and gastroenteritis (Adenoviridae), oral and genital herpes, chickenpox, and shingles (Herpesviridae), and human infantile gastroenteritis (Reoviridae). For instance, the 1250 Å diameter herpes simplex virus (HSV) (Zhou, Dougherty et al. 2000), which is currently present in over 60% of the US population, is responsible for herpes, cowpox, cancer, and many other dangerous diseases.

To overcome the Ewald sphere curvature resolution limit, the paraboloid reconstruction (Prec) algorithm for Cryo-EM, was developed to correct for the effects of the Ewald sphere curvature in the context of 3D reconstructions. Details of the algorithm are discussed in Chapter 3.

1.19 References

Amelinckx, S., R. Gevers, et al. (1978). Diffraction and imaging techniques in material science. Amsterdam ; New York, Elsevier North-Holland.

Anderson, D. P., J. Cobb, et al. (2002). "SETI@home - An experiment in public-resource computing." Communications of the Acm **45**(11): 56−61.

Angell, C. A. (2004). "Amorphous water." Annual Review of Physical Chemistry **55**: 559−583.

Baker, T. S., N. H. Olson, et al. (1999). "Adding the third dimension to virus life cycles: Three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs." Microbiology and Molecular Biology Reviews **63**(4): 862−922.

Baltimore, D. (1971). "Expression of Animal Virus Genomes." Bacteriological Reviews **35**(3): 235−241.

Bloomer, A. C., J. N. Champness, et al. (1978). "Protein Disk of Tobacco Mosaic-Virus at 2.8-a Resolution Showing Interactions within and between Subunits." Nature **276**(5686): 362−368.

Bottcher, B., S. A. Wynne, et al. (1997). "Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy." Nature **386**(6620): 88−91.

Bragg, W. L. (1929). "The determination of parameters in crystal structures by means of fourier series." Proceedings of the Royal Society of London Series A-Containing Papers of a Mathematical and Physical Character **123**(792): 537−559.

Briegel, A., H. J. Ding, et al. (2008). "Location and architecture of the Caulobacter crescentus chemoreceptor array." Molecular Microbiology **69**(1): 30−41.

Caspar, D. L. D. and A. Klug (1962). "Physical Principles in Construction of Regular Viruses." Cold Spring Harbor Symposia on Quantitative Biology **27**: 1−24.

Conway, J. F., N. Cheng, et al. (1997). "Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy." Nature **386**(6620): 91−94.

Crick, F. H. C. and J. D. Watson (1956). "Structure of Small Viruses." Nature **177**(4506): 473−475.

Crowther, R. A., L. A. Amos, et al. (1970). "3 Dimensional Reconstructions of Spherical Viruses by Fourier Synthesis from Electron Micrographs." Nature **226**(5244): 421−425.

DeRosier, D. J. (2000). "Correction of high-resolution data for curvature of the Ewald sphere." Ultramicroscopy **81**(2): 83−98.

Dubochet, J. and A. W. Mcdowall (1981). "Vitrification of Pure Water for Electron-Microscopy." Journal of Microscopy-Oxford **124**(DEC): RP3−RP4.

Fujiyoshi, Y., T. Mizusaki, et al. (1991). "Development of a Superfluid-Helium Stage for High-Resolution Electron-Microscopy." Ultramicroscopy **38**(3−4): 241−251.

Fuller, S. D., S. J. Butcher, et al. (1996). "Three-dimensional reconstruction of icosahedral particles - The uncommon line." Journal of Structural Biology **116**(1): 48−55.

Fultz, B. and J. M. Howe (2002). Transmission electron microscopy and diffractometry of materials. Berlin ; New York, Springer.

Glaeser, R. M. (1999). "Review: Electron crystallography: Present excitement, a nod to the past, anticipating the future." Journal of Structural Biology **128**(1): 3-14.

Grigorieff, N. (2007). "FREALIGN: High-resolution refinement of single particle structures." Journal of Structural Biology **157**(1): 117−125.

Harauz, G. and M. Van Heel (1986). "Exact Filters for General Geometry 3-Dimensional Reconstruction." Optik **73**(4): 146−156.

Hell, S. W. (2007). "Far-field optical nanoscopy." Science **316**(5828): 1153−1158.

Henderson, G. P. and G. J. Jensen (2006). "Three-dimensional structure of Mycoplasma pneumoniae's attachment organelle and a model for its role in gliding motility." Molecular Microbiology **60**(2): 376−385.

Henderson, R., J. M. Baldwin, et al. (1990). "Model for the Structure of Bacteriorhodopsin Based on High-Resolution Electron Cryomicroscopy." Journal of Molecular Biology **213**(4): 899-929.

Henderson, R., J. M. Baldwin, et al. (1990). "Model for the Structure of Bacteriorhodopsin Based on High-Resolution Electron Cryomicroscopy." Journal of Molecular Biology **213**(4): 899−929.

Heymann, J. B. (2001). "Bsoft: Image and molecular processing in electron microscopy." Journal of Structural Biology **133**(2−3): 156−169.

Hite, R. K., S. Raunser, et al. (2007). "Revival of electron crystallography." Current Opinion in Structural Biology **17**(4): 389−395.

Iancu, C. V., W. F. Tivol, et al. (2006). "Electron cryotomography sample preparation using the Vitrobot." Nature Protocols **1**(6): 2813−2819.

Iancu, C. V., E. R. Wright, et al. (2005). "A "flip-flop" rotation stage for routine dual-axis electron cryotomography." Journal of Structural Biology **151**(3): 288−297.

Iancu, C. V., E. R. Wright, et al. (2006). "A comparison of liquid nitrogen and liquid helium as cryogens for electron cryotomography." Journal of Structural Biology **153**(3): 231−240.

Ito, H., H. Aoki, et al. (2006). "Autophagic cell death of malignant glioma cells induced by a conditionally replicating adenovirus." Journal of the National Cancer Institute **98**(9): 625−636.

Jensen, G. J. and R. D. Kornberg (2000). "Defocus-gradient corrected back-projection." Ultramicroscopy **84**(1−2): 57−64.

Jiang, W., M. L. Baker, et al. (2008). "Backbone structure of the infectious epsilon 15 virus capsid revealed by electron cryomicroscopy." Nature **451**(7182): 1130−1134.

Juette, M. F., T. J. Gould, et al. (2008). "Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples." Nature Methods **5**(6): 527−529.

Klug, A., F. H. C. Crick, et al. (1958). "Diffraction by Helical Structures." Acta Crystallographica **11**(3): 199−213.

Liang, Y. Y., E. Y. Ke, et al. (2002). "IMIRS: a high-resolution 3D reconstruction package integrated with a relational image database." Journal of Structural Biology **137**(3): 292−304.

Lucic, V., F. Forster, et al. (2005). "Structural studies by electron tomography: From cells to molecules." Annual Review of Biochemistry **74**: 833−865.

Ludtke, S. J., P. R. Baldwin, et al. (1999). "EMAN: Semiautomated software for high-resolution single-particle reconstructions." Journal of Structural Biology **128**(1): 82−97.

Murray, P. R., K. S. Rosenthal, et al. (2005). Medical microbiology. Philadelphia, Elsevier Mosby.

Penczek, P. A., R. A. Grassucci, et al. (1994). "The Ribosome at Improved Resolution - New Techniques for Merging and Orientation Refinement in 3d Cryoelectron Microscopy of Biological Particles." Ultramicroscopy **53**(3): 251−270.

Potter, C. S., H. Chu, et al. (1999). "Leginon: a system for fully automated acquisition of 1000 electron micrographs a day." Ultramicroscopy **77**(3−4): 153−161.

Reimer, L. (1997). Transmission electron microscopy : physics of image formation and microanalysis. Berlin ; New York, Springer.

Schmidt, R., C. A. Wurm, et al. (2008). "Spherical nanosized focal spot unravels the interior of cells." Nature Methods **5**(6): 539−544.

Shannon, C. E. (1949). "Communication in the Presence of Noise." Proceedings of the Institute of Radio Engineers **37**(1): 10−21.

Tannenbaum, T. and M. Litzkow (1995). "The Condor Distributed-Processing System." Dr Dobbs Journal **20**(2): 40−48.

Trus, B. L., R. B. S. Roden, et al. (1997). "Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 angstrom resolution." Nature Structural Biology **4**(5): 413−420.

Tugarinov, V., W. Y. Choy, et al. (2005). "Solution NMR-derived global fold of a monomeric 82-kDa enzyme." Proceedings of the National Academy of Sciences of the United States of America **102**(3): 622−627.

Unwin, N. (2005). "Refined structure of the nicotinic acetylcholine receptor at 4 angstrom resolution." Journal of Molecular Biology **346**(4): 967−989.

Unwin, N. (2005). "Refined structure of the nicotinic acetylcholine receptor at 4 angstrom resolution." Journal of Molecular Biology **346**(4): 967-989.

van Heel, M., B. Gowen, et al. (2000). "Single-particle electron cryo-microscopy: towards atomic resolution." Quarterly Reviews of Biophysics **33**(4): 307−369.

van Heel, M. and M. Schatz (2005). "Fourier shell correlation threshold criteria." Journal of Structural Biology **151**(3): 250−262.

Wan, Y., W. Chiu, et al. (2004). "Full contrast transfer function correction in 3D cryo-EM reconstruction". IEEE Proceedings of ICCCAS 2004 Chengdu, Sichuan, China.

Whittaker, E. T. (1915). "On the Functions which are Represented by the Expansion of Interpolation Theory." Proceedings of the Royal Society of Edinburgh **35**: 181−194.

Yu, X. K., L. Jin, et al. (2008). "3.88 angstrom structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy." Nature **453**(7193): 415−419.

Zhang, X., E. Settembre, et al. (2008). "Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction." Proceedings of the National Academy of Sciences of the United States of America **105**(6): 1867−1872.

Zhou, Z. H., M. Dougherty, et al. (2000). "Seeing the herpesvirus capsid at 8.5 angstrom." Science **288**(5467): 877−880.

zur Hausen, H. (2002). "Papillomaviruses and cancer: From basic studies to clinical application." Nature Reviews Cancer **2**(5): 342−350.

1.20 Figures and tables

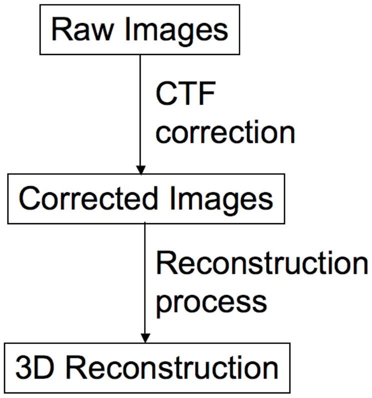

Figure 1-1. **Flow chart of simplified reconstruction process.** The reconstruction process consist of three stages: (1) Raw images from electron micrographs, (2) Corrected images produced by CTF correction of Raw images, (3) 3D real-space reconstruction generated by reconstruction algorithm using corrected images

Biological Structural Features	**Approximate Resolution**
α-helices	~ 7Å
Main chain	~ 4Å
Side chains	~ 3Å
Atomic details	~1–2Å

Table 1-1. **Table of biological structural features observable at different resolutions.** Visual resolution of a reconstruction can be determined by the observation of various structures common to biological samples

Viruses Shown to Infect Humans	**Size (Å)**
Adenoviridae	
Human Adenovirus Serotypes 1–47	700–900
Herpesviridae	
Herpes Simplex Virus Type 1 (HSV-1)	~ 1,500
Herpes Simplex Virus Type 2 (HSV-2)	
Varicella-Zostrer Virus	
Epstein-Barr Virus	
Cytomegalovirus (CMV)	
Human Herpesvirus 6 (Roseola Infantum)	
Human Herpesvirus 7	
Reoviridae	
Reovirus 1, 2, 3	600–800
Colorado Tick Fever Virus	
Rotavirus Groups A, B, C	

Table 1-2. **Table of viruses known to infect humans.** Viruses known to infect humans (Murray, Rosenthal et al. 2005) for which the correction of the curvature of the Ewald sphere will be required to derive atomic models by cryo-EM