Published in final edited form as:

J Struct Biol. 2009 August ; 167(2): 97–105. doi:10.1016/j.jsb.2009.05.002.

Ab initio Reconstruction of Helical Samples with Heterogeneity, Disorder and Coexisting Symmetries

Vincent H. Ramey¹, Hong-Wei Wang², and Eva Nogales^{2,3,*}

¹ Biophysics Graduate Program, UC Berkeley

² Life Science Division, Lawrence Berkeley National Laboratory

³ Department of Molecular and Cell Biology, UC Berkeley/Howard Hughes Medical Institute

Abstract

We describe modifications of the single particle helical reconstruction approach devised for the analysis of a sample that could not be processed with existing methods due to its variable and short range helical order. The added steps of reference-free two-dimensional image classification and alignment, and automated microtubule removal from images, have particular application to proteins or protein complexes that assemble around microtubules. The method was successfully applied to the Dam1 complex, an essential component of the yeast kinetochore that couples replicated chromosomes to spindle microtubules during mitosis. Because of its novel mode of binding, which does not involve a footprint on the microtubule lattice, new steps to deal with the disorder and heterogeneity of the Dam1 complex assembly were required to gain structural information about this complex both routinely and efficiently.

Introduction

The first three-dimensional (3-D) structures of biological macromolecules ever to be obtained by electron microscopy were those of samples with helical symmetry (DeRosier and Klug, 1968). The 3-D structure of an object arranged in a helix can be reconstructed from a single two-dimensional (2-D) projection image, whose Fourier transform is composed of a series of layer lines. These layer lines are Bessel functions of different orders defined by the helical symmetry (DeRosier and Klug, 1968; DeRosier and Moore, 1970; Klug et al., 1958). Application of Fourier-Bessel helical reconstruction methods have led in recent years to close to atomic resolution for a number of important biological macromolecules (Miyazawa et al., 2003; Yonekura et al., 2003). However, these classical helical reconstruction methods are clearly limited by the requirement of robust helical order in the sample to be studied. This condition can sometimes be extremely difficult to attain, if not impossible. In the past decade, new computational methods combining single particle image processing techniques with conventional helical reconstruction have been successfully implemented to address a number of problems in helical assemblies (Egelman, 2000; Egelman, 2007; Holmes et al., 2003; Li et al., 2002; Sachse et al., 2007; Pomfret et al., 2007). For the use of these algorithms, images of several helical filaments are chopped into small fragments, each containing typically a few to tens of turns of the helix. Analogously to single particle analysis of asymmetric molecules, the

*Corresponding author: 708C Stanley Hall, QB3, UC Berkeley, Berkeley, CA 94720-3220; enogales@lbl.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

fragments can be sorted based on their features to reduce heterogeneity. A reconstruction of the experimental filament can be calculated by situating each segment in 3-D space according to its relative orientation and applying the correct helical symmetry through iterative alignment and sorting.

The need to effectively deal with the heterogeneity observed in many helical assemblies was one of the original impetuses for the development of single particle helical methods by Egelman and others, (Bluemke et al., 1988; Egelman, 2000). Our work here builds on these methods to deal with short range helical order on the scale of single turns and competing symmetries during the *ab initio* reconstruction of these assemblies. The most important step of helical reconstruction is to find the correct symmetry for each helical segment in the dataset. This is confirmed by the observation that refinement of helical particles against a featureless cylinder using the close-to-correct symmetry recovers the correct structure (Egelman, 2000; Egelman, 2007). The helical symmetry is normally estimated by examining the power spectrum of the entire helical particle image or summed power spectra of the helical fragments. Such an estimation is, however, rather difficult for highly heterogeneous assemblies, where severe blurriness of the layer lines occurs. The lack of symmetry information can hinder the effectiveness of the iterative single particle analysis in sorting fragments into their homogeneous helical groups, or may prevent the algorithm from converging with the correct solution all together. As a solution to either or both of these problems we have added a reference-free 2-D classification and alignment step that extends the use of the single particle methodology to deal with systems where symmetry determination is non-trivial by other means. This 2-D processing step aims to sort the data into homogeneous classes of helical segments, based on features of structure and local symmetry, but without prior knowledge or manual user intervention. The class averages from these groups can then be used for symmetry estimation, initial 3-D model generation, and projection matching refinement towards higher resolution, following the established single particle methods of Egelman and coworkers. Here we have successfully applied this approach to the Dam1 complex, a kinetochore component required for the robust connection of the mitotic chromosome to the spindle microtubule.

The hetero-decameric Dam1 kinetochore complex has a total weight of 210 kDa. *In vitro* it has been shown to oligomerize into stable rings around microtubules, (Westermann et al, 2005; Miranda et al., 2005), containing about sixteen copies of the complex (Westermann et al, 2006; Wang et al, 2007). Interestingly, the interaction of the ring with the microtubule occurs via flexible elements, likely in both tubulin and the Dam1 complex, (Westermann et al., 2006; Miranda et al., 2007; Wang et al., 2007). The ring geometry and the flexible character of its attachment to the microtubule allows the assembled complexes to act as a force-coupler and transduce the energy released by peeling microtubule protofilaments during depolymerization into a minus-end directed force (Westermann et al., 2006; Asbury et al., 2006; Grishchuk et al., 2008). At saturating concentrations the Dam1 complex forms tightly packed spirals around microtubules, with each turn corresponding very closely to a slightly skewed ring, (or lock-washer). This helical arrangement has the potential to aid structural analysis, by both concentrating the complexes and aligning them. However, the Dam1 spirals formed around microtubules almost never sustain the type of long-range helical order required for conventional helical reconstruction methodology. This is opposed to the well-ordered assembly that is commonly seen for other microtubule-interacting proteins that bind tightly to the microtubule lattice and thus “reproduce” the helical arrangement of tubulin, (e.g. kinesin (Amos and Hirose, 1997; Arnal et al., 1996; Sosa et al., 1997)). Furthermore, because of the flexible character of the attachment of Dam1, as well as its much larger size, the Dam1 assembly has a different, unrelated helical symmetry from that of the underlying microtubule, an interesting property for its biological function that, however, adds technical challenges to using single particle approaches for its study.

The microtubule itself is a well-studied helical structure made of repeating $\alpha\beta$ tubulin heterodimers. The dimers assemble head to tail to form protofilaments, about thirteen of which interact laterally to create a hollow tube approximately 250 Å in diameter. The spiral of Dam1 assemblies around microtubules have an inner diameter of about 320 Å and a helical pitch of about 150 Å (compared to the 40 Å pitch of microtubules). The difference in helical symmetries between the Dam1 complex and the microtubule would make it easy, in principle, to use conventional Fourier-Bessel helical reconstruction methods to study this system, as the two components are separable in Fourier space. In other words, there is no layer line overlap excepting the equator. Thus, the information corresponding to each structure could be extracted independently and each reconstructed with their appropriate symmetries. However, during our cryo-EM studies of Dam1 spirals on microtubules, out of hundreds of filaments examined, only two showed clear helical diffraction to a resolution better than 40 Å (Wang et al., 2007). Most spirals appear ordered by visual inspection but did not display more than one blurred layer line due to the loss of registry between adjacent turns (Fig 1). Ideally we would like to utilize the helical order retained at the level of single turns to avoid this limitation and robustly study this sample. Since conventional Fourier-Bessel methods cannot operate on such short helical stretches, due to extreme layer line broadening and lack of signal, the structural information of the microtubule and Dam1 helix must be separated in a different way and an alternative reconstruction method used to gain structural information.

The approach presented here overcomes these difficulties in studying Dam1 structure by removing the microtubule density from each short segment and then applying 2-D image alignment and classification techniques to sort the assembly turn by turn into homogeneous classes, with no assumptions made about the symmetries present or amount of disorder. Inspection of class averages reveals the number of different symmetries and allows estimation of the helical parameters to use in their reconstruction.

Method

The character of the Dam1 complex assembly around the microtubule as addressed above makes it extremely challenging for studies using either conventional helical methods or single particle approaches. In order to overcome these challenges we have implemented an automated protocol to remove the microtubule structure from the images, which makes it possible to take on a single particle approach using short helical stretches along the spiral. Additionally, we have successfully applied image classification techniques commonly used in single particle methods to address the problem of extreme disorder and symmetry heterogeneity in the Dam1 assembly, opening the door for structural studies of this complex that have been hampered by lack of long-range helical order.

Preprocessing - Removal of microtubule and boxing of single turns

When averaging single particles during 3-DEM reconstruction, as we are doing here by treating a small number of helical turns as a 'single particle', it is commonly assumed that the objects being analyzed are identical biochemically and that each image being averaged corresponds to the same orientation. When mismatched symmetries co-exist in the sample, as for the Dam1 spiral assembled around a microtubule, this is almost never the case, because the specific arrangement between the two assemblies in each segment will only rarely be repeated elsewhere, (i.e. they are two hollow helical tubes that are able to rotate and slide past each other within the scale of individual subunits). Therefore, each image of a Dam1 helical turn has several million Daltons worth of randomly rotated and translated microtubule signal in it. This corrupting signal must be removed to prevent it from biasing the alignment and classification steps of single particle processing and therefore blurring structural features after averaging.

We remove the microtubule signal using an automated approach, in which each small segment, rather than the entire long filament, is processed, thus avoiding the time consuming, manual step of filament straightening. Three main structural features of the underlying microtubule lattice interfere with attempts to align, classify and average Dam1 helices. They are: the projection of the microtubule's cylindrical density in the direction normal to the tube axis; the lateral protofilament pattern of the wall; and the 40 Å axial repeat of tubulin in the lattice. The preprocessing steps taken for removal of this unwanted microtubule signal are outlined in Figure 2. In order to remove the projection density of the microtubule cylinder and the pattern of protofilaments, which for these short segment lengths can be treated as a circular arrangement of vertical cylinders, the raw image of the Dam1-surrounded microtubule, (Fig 2a bottom), is projected along their shared helical axis to generate a one-dimensional (1-D) density profile, (Fig 2a top). Based on the fact that the Dam1 spiral sits at a significant distance from the microtubule wall and therefore their radial density distributions are well separated, we "trimmed" the Dam1 density off from the 1-D profile leaving only the central part as the microtubule density (Fig 2b top). Then a 2-D microtubule mask, (Fig 2b bottom), is generated by backprojecting this truncated 1-D profile along the helical axis. Now the microtubule signal can be removed by simply subtracting this mask from the raw image in real space, resulting in the corrected image shown in Figure 2c. Its corresponding 1-D profile, (Fig 2c top), confirms the removal of this central microtubule density. The signal from the lateral protofilament pattern within the radial range of microtubule is also removed during this process. Power spectra of the raw, masked, and corrected images are shown in Figures 2d, e and f respectively. After masking, the microtubule contribution is removed and the overall shape of the microtubule disappears from the image. This method for microtubule removal is conceptually similar but procedurally different from one presented recently (Narita et al., 2007). In that study it was necessary to remove the microtubule because it interfered with identifying points of sparse decoration on its lattice. They removed the protofilament pattern by zeroing out the entire equator in the Fourier transform of longer stretches of microtubule, removing the moiré pattern and density corresponding to the microtubule wall. Our real-space procedure is equivalent to this Fourier-space operation on the equatorial layer line, but instead of removing the entire layer line we mask only information contained within the boundaries of the microtubule. Parts of the Dam1 cylindrical density (the near and far side of the Dam1 spiral) inside this mask in the image are also removed during this procedure. However, the contribution of these features to the Dam1 spiral's cylindrical density distribution is much smaller than the side views of the walls outside the mask and therefore our method leaves Dam1 relatively unaffected. Thus ample, correct signal remains to effectively align and classify the data. The mild effect of this masking on the alignment is confirmed by the absence of artifacts in the final reconstruction when compared to an independently derived structure. In order to eliminate this bias completely one can also use the angles and shifts found with microtubule masked particles and apply them to the original unmasked data. The inner density corresponding to the microtubule could then be easily removed in the resulting 3-D reconstruction.

The major remaining contaminating signal from the microtubule comes from the axial repeat of tubulin along the protofilament. In Fourier space this is mainly located in the 40 Å layer line, (higher orders were neglected for this low resolution implementation of the process involving a limited number of segments). There is no layer line overlap between the Dam1 and microtubule helices since they have unrelated symmetry. Taking advantage of this independence, the 40 Å microtubule layer line was zeroed out in each short helical segment using a mask width of 3 pixels (Fig 2g). Dam1 helical segments that had been treated to completely remove the microtubule contribution to the image were then boxed to single turns of the helix with a soft rectangle to allow the algorithm to sort heterogeneity turn by turn (Pomfret et al., 2007).

Addressing Heterogeneity With Reference-Free Alignment and Classification

Heterogeneity has been successfully resolved in previous studies using single particle helical methods by comparing segments against 3-D models of different states of the assembly having varying symmetry or conformation (Trachtenberg et al., 2005; Frost et al., 2008; Galkin et al., 2008a; Galkin et al., 2008b; Orlova, A. et al., 2007; Egelman, 2007). Models have been generated in different ways, including: low resolution Fourier-Bessel reconstruction (Frost et al., 2008), application of various symmetries to an 'average' model derived from heterogeneous data (Trachtenberg et al., 2005), and making simulated structures based on the symmetries seen in the averaged power spectra of particles (Wang et al., 2008). By matching each segment to reprojections of these models they can be put into homogeneous groups of known symmetries and reconstructed separately. While we were able to generate a 30 Å resolution reconstruction of the Dam1 complex assembled around microtubules from two rare ordered stretches, we were unable to sort the rest of the data using this model due to the rare and extreme short-ranged order of the sample. Frequently helical segments of Dam1 seemed not to maintain a unique symmetry even for a single turn and this made it difficult to apply the existing methods to create homogeneous groups. The Dam1 complex is a case where the reconstruction of the assembly requires a different approach for sorting heterogeneity.

Our method is to create groups of segments with homogeneous pitch and diameter by aligning and classifying them in 2-D (Fig 3). By making the common assumption that the size of the asymmetric unit along the spiral is constant in each particle, this implies that segments with homogeneous pitch and diameter all represent a single helical symmetry. There are a small number of cases where this is not true, and these degenerate arrangements must be clarified by further experiments (Zhang and Hinshaw, 2001; Chen et al., 2004; Egelman, 2007). Because our sorting is reference-free, it requires no *a priori* knowledge about the number of states or symmetries of the assembly. The classification algorithm will group segments based on their features alone, deriving the distribution of symmetries directly from the data. This is an important point, because existing methods for sorting require a model to be generated for each symmetry in the data. Since Dam1 spirals contain many disordered turns (turns without a constant helical symmetry), we failed to generate models that could sort our data to homogeneity. Reference-free alignment and classification avoids this difficulty. Class averages and their power spectra quickly make it clear which groups are characterized by disorder (and will not be further processed) or have a specific helical symmetry and can be further pursued structurally. This type of iterative, reference-free alignment and classification of images has been used very successfully to reduce heterogeneity and reconstruct samples during single particle analysis of asymmetric samples (van Heel and Frank, 1981; Frank et al., 1996; van Heel et al., 1996; Ludtke et al., 1999). When confronted with a large degree of heterogeneity, as caused by the highly variable symmetries in the Dam1 helix, we must isolate one or a few well ordered turns from a mostly disordered filament of tens of turns. This sparseness of order makes any type of 'ensemble' image processing, i.e. averaged power spectra, impossible. In fact, during our analysis approximately 80% of segments were classified into groups without constant, well defined pitch and diameter, indicating that these segments did not preserve even short range helical order (Fig 3b). This is unsurprising given the large amount of disorder in filaments indicated by the single broad layer line in Figure 1. This initial step, dedicated to finding segments that belong to the same type of helical symmetry, while excluding disorder, is a key step in our single particle helical reconstruction implementation.

During 2-D processing, classification separates helical particles based on several features, which are, in decreasing order of scale: diameter, pitch, out of plane tilt, and azimuthal angle. Changes in these parameters dominate other characteristics and help us separate segments primarily based on these properties. Eigen images generated during Multivariate Statistical Analysis (MSA) classification have been a powerful tool for visualizing and sorting size

differences during single particle processing, (White et al., 2004). This is also the case in our study and can be seen clearly in Figure 4. The major eigen images, (Fig 4a), which contribute most heavily to the separation of images into classes, describe the features of varying pitch/diameter or amount of disorder, but have essentially no features corresponding to the repeating unit. Later eigen images, (Fig 4b), contributing less to the classification, show very strongly the repeat of DamI along the helical lattice and indicate changes in the azimuthal view among segments. Because of this hierarchy of the size of different structural features, we can control the composition of our classes through the choice of eigen images we use for classification (Fig S1, S2), or, when using classification schemes other than MSA, our choice of how many classes to divide our data into. In this study we have mostly used the topology representing networks of Ogura et al. to generate averages, for the practical reason that it allowed easy automation of iterative 2-D processing while still delivering very similar classes to those from MSA (Fig S3). We present the complimentary MSA analysis here to demonstrate that classification based on symmetry can be clearly shown by inspecting eigen images. This result ultimately implies that the sorting techniques used for other single particle samples with heterogeneity (Orlova, E.V. et al., 2000; White et al., 2004) are equally effective for helical samples as well. The progression of separation based on the general properties of diameter, pitch and tilt, and azimuthal angle is robust and reproducible for the case studied here, suggesting that it will be applicable to other helical samples.

3-D Processing – Ab Initio Symmetry Estimation and Refinement

Once a class average of helical segments with homogeneous symmetry has been obtained, it can be used to estimate initial helical parameters for that group, but also, due to helical symmetry, to generate a 3-D reconstruction (Figs 5, S4). Helical symmetry has the special property that all of the information needed for a 3-D reconstruction exists in a projection that is perpendicular to the helical axis, assuming that there are a large number of subunits per turn or helical repeat. Correct helical parameters, the twist ($^{\circ}$) and rise ($\text{\AA}$) per repeating unit that defines the helical arrangement, are essential to accurately determine a structure. The second key step of reconstruction after calculating homogeneous classes is therefore to assign symmetry values to those classes with a visually well-defined pitch and diameter. For any given class, the number of possible symmetries can be narrowed down to only a few by careful inspection of the Fourier transform of class averages and knowledge of the size of the repeating unit.

Interpretation of Fourier transforms or power spectra must be done with caution and symmetry estimation carried out on a case-by-case basis, since 3-D refinement with incorrectly derived symmetry guarantees a structure that is grossly wrong, yet not obviously so, (the chapter by David DeRosier dedicated to helical reconstruction is especially useful (Glaeser et al., 2007)). Still, while indexing has pitfalls, it remains a required and powerful tool for the reconstruction of helical assemblies and has been routinely used in the literature, including analyzing the diffraction of short helical segments, (DeRosier and Moore, 1970; Wang et al., 2007; Orlova et al., 2007; Wang et al., 2008; Frost et al., 2008).

The high amount of signal and homogeneity of symmetry give Fourier transforms of class averages clearer layer line signals than the summed power spectra of unsorted fragments. Analysis of the Fourier transform of averages using conventional helical methods can therefore help define the lattice symmetry by limiting the range of possible values of twist and rise. In cases where a small number of turns per segment are needed to sort extreme heterogeneity, such as for the DamI complex, exact indexing is simply impossible due to layer line broadening and overlap. Indeed for many cases the estimates will not be exact due to the thickness of the protein and the error in measuring layer line peak positions. However, the number of possible symmetries to consider can still be greatly reduced (Fig 3d). Visual inspection of the class

average and the meridional distance of the layer line describing the 150 Å pitch of Dam1 spirals, confirms it is a one-start helix. Layer lines from other helical families are much broader, due to the small box size, and appear as diffuse reflections far from the meridian but that are still recognizable (dashed arrows in Fig 3d). The meridional distance of the lowest order of these reflections is $\sim 1/100 \text{ Å}^{-1}$. Using the equation $n+2\sim 2\pi Rr$, with r being the maximum real-space radius of the helix ($\sim 250 \text{ Å}$ for Dam1) and R the reciprocal space distance of the peak from the meridian ($1/100 \text{ Å}^{-1}$), we estimated the Bessel order of the layer line closest to the equator to be $n = 15.7$. A helical turn with 15.7 subunits and a pitch of 150 Å gives helical parameters of rise = 9.55 Å and twist = 22.9°. This is likely an upper limit of n , since the maximal radius of the particle was used. We cannot rule out smaller values of n , down to $n = 9$, found using the inner radius of the helix. To narrow this range we needed further information about the size and arrangement of the complex when assembled.

Dam1 rings that had fallen off microtubules show 16 fold rotational symmetry, (Wang et al., 2007). The radius and geometry of the ring form of Dam1 are very similar to these helical crystals, lending evidence for a value of n nearer to its upper bound of 15.7. Further evidence for this arrangement comes from the size of the Dam1 complex seen previously by single particle reconstruction, (Wang et al., 2007). It appears as a wide 'T' with a maximum dimension of 120 Å and a protrusion that has been implicated in microtubule binding, and dimerizes head to tail with a spacing of $\sim 110 \text{ Å}$. This estimate of the asymmetric unit size can also be used to calculate helical parameters. Calculation of the path length of the spiral in one turn,

$$L^2 = C^2 + P^2$$

where C is the circumference of a circle with the diameter of the helix (500 Å for Dam1), P is the pitch (150 Å), and L is the path length, leads to a value of $\sim 1,570 \text{ Å}$ for Dam1 spirals of the dominant symmetry. An asymmetric unit length of 110 Å then implies ~ 14.3 subunits per turn and a rise and twist of 10.5 Å and 25.22°. These values agree with the spacing deduced from the ring form of Dam1 and fit near the upper limit of n seen from the power spectrum of the class average. This gives us confidence that we have found roughly the correct symmetry of this assembly. Indeed, as mentioned later, both sets of values lead to the same, correct structure after refinement. Because the inevitable error in the value of parameters used for symmetry estimation, we tested helical parameters corresponding to 18 through 12 subunits per turn of the 150 Å-pitch one start helix.

Classification and alignment generate averages that can be used first for finding symmetry and then to generate an initial model by back-projection of symmetry related copies of the class average. But just as important, the classification procedure groups segments based on their features. Thus, after 2-D processing we have, for each class: estimates of symmetry parameters, an initial model, and a homogeneous group of raw segments. We can therefore follow a projection matching scheme identical to those used in other single particle helical reconstruction approaches, but for each class separately (Egelman, 2000). During each individual refinement, every segment from that class is matched against reprojections of that class reference volume. It is assigned Euler angles, which define its orientation in 3-D space, based on its best matching view, and this information is used to reconstruct a new 3-D model without applying symmetry. Using programs developed by Egelman, we search for the intrinsic symmetry of this volume to gain more accurate helical parameters. These are then applied to segments to create a helically symmetric reconstruction that will more faithfully reflect the data. This model in turn can be used for another cycle of projection matching, and the process is iterated until it converges to a final, stable structure. Underlining the importance of accurate estimation of the helical parameters, we have observed that beginning the refinement for the

Dam1 assembly with an estimate of the twist per subunit that was off more than 2° from the correct 24.6° invariably caused divergence into an incorrect structure, which we could discriminate based on its dissimilarity in projection to the reference-free class average from which it was derived.

It has been shown that a featureful initial model is not required to refine correctly a helical structure, as a cylinder has been used successfully in many cases (Egelman, 2000; Egelman, 2007; Galkin et al., 2006). While this is also true of the Dam1 complex, we observed a shorter convergence time when starting with a model derived from a class average rather than a cylinder. The extra constraint of starting with a model with the correct pitch, diameter, and low resolution features seemed to confine segments close to their correct angles from the beginning. Whether this would make the refinement converge from a wider range of symmetry values was not exhaustively tested and is still an open question.

The Dam1 helix found here contains many views of the repeating unit, even for the short axial lengths used in this study. Other samples, like the archebacterial flagellar filament that has a repeat on the order of 1,000 Å, only provide a limited number of views per short segment (Trachtenberg et al., 2005). To put it another way, the Dam1 helix has a pseudo-repeat that fits within a reasonable box size for processing while a sample like the flagella does not. For this long repeat case classification would be expected to separate distinct views of the same symmetry into different classes, complicating 3-D structure refinement, since needed views would be missing from any single class. This problem could be overcome in two ways. First, during classification of Dam1 segments we found that eigen images describing the pitch and diameter of the filament came first and were separable from the later images that classified based on view, (Fig 4). When only the first 8 eigen images were used in classification, the classes had homogeneous pitch and diameter but contained many azimuthal views, blurring Dam1 along the one start spiral in the average, (Fig S1). Classification of long repeat helices with only a subset of eigen images would therefore be expected to do the same – generate classes of homogeneous symmetry but mixed view that could be used for 3-D reconstruction. A second approach would be to align and classify the power spectra of segments instead of the images themselves. The removal of phase information from the dataset, and therefore information about the view of the filament, apart from small variations in layer lines due to overlap, would cause classification of particles based only on their layer line spacings. Grouping particles into homogeneous power spectrum groups would have the same effect as described previously – homogeneous symmetry with many representative views. Classes generated with either modification to the method would have both the homogeneity of symmetry and range of views necessary for successful 3-D reconstruction.

Results

The application of our algorithm to Dam1 spirals proceeded as shown visually in Figure 5 and schematically in Figure S4. Overlapping segments were selected from digitized cryo-EM micrographs of Dam1 decorated microtubules (see Materials and Methods). The approximate in-plane angle of each segment was estimated based on its orientation during particle picking and used to process each particle. Specifically, this angle was used to project the segment along its axis for microtubule protofilament removal and 40 Å layer line editing. Both of these masks were created with a temporary, vertical version of the segment and then were applied to the raw image to prevent image degradation by interpolation. Each segment was then boxed with a small, soft rectangular mask of fixed size, leaving between one and two turns of the helix per segment. This was done to make the classification as fine grained as possible and thus take advantage of data that was only ordered on the small scale of one or a few turns. After this preprocessing was completed, segments were subjected to automated, iterative 2D alignment

and classification. We used the method of Ogura et al. 2003, which can be considered equivalent to including all eigen images during MSA classification (Fig S3).

One predominant helical symmetry became clear from the class averages and a representative average was selected for further processing. This average came from classification that was fine grained, separating segments into classes based on diameter, pitch, tilt, and azimuthal view. Its parameters were estimated by measuring its pitch, the off meridional distance of layer lines, and the known size of the complex from images of unassembled Dam1 (Wang et al., 2007). This class included 286 segments for a total of ~8500 asymmetric units. A range of symmetries were applied to the class average with values ranging from 20° to 30° twist, or 12 to 18 subunits per turn, and an 8 to 12.5 Å rise selected to preserve the known pitch. Each was then used as a starting model for projection matching refinement using that symmetry as the initial estimate. Only volumes refined with a twist of $24.6^\circ \pm 2^\circ$ (corresponding to the range of 15.9 to 13.5 subunits per helical turn) converged to a common solution whose projection matched both its reference-free class average, and the existing structure obtained by Fourier Bessel reconstruction (Wang et al, 2007). The other parameters caused divergence or led to incorrect volumes whose projections did not match the reference-free class average, as determined by cross-correlation score and visual inspection of their alignment. The final helical parameters of 24.6° and 10.20 Å are very similar to 24.7° and 10.24 Å determined by conventional methods during our previous studies, though different data were used, (Wang et al., 2007). This shows that we found single turns of a similar symmetry that had previously generated multi-turn helical crystals.

The major features in this new, real space reconstruction, are in good agreement with our previous structure from conventional helical reconstruction methods, (Fig 6). The maps superimpose well, with the protrusion from the Dam1 spiral towards the microtubule, the bridge between adjacent turns, and the main body of the spiral, all strongly showing the same morphology. As before, this assembly is a one-start helix made of two anti-parallel head-to-tail arrangements of the complex. To quantitate the similarity of the two structures, a Fourier Shell Correlation (FSC) curve (Fig 6c) of the asymmetric unit and a phase residual over their corresponding layer lines was computed. The 0.5 FSC cutoff at 30 Å coincides with the resolution of both maps, confirming that they are nearly identical over all features. The phase residual over all layer lines within 30 Å was 43°, however, several of the layer lines were barely over the noise level in the power spectrum of the previous helical structure. The phase residual calculated over the strong layer lines in both structures yielded a value of 30°, with 47° for the 30 Å layer line, which indicate good agreement all the way to this resolution.

Discussion

The Dam1 kinetochore complex is essential for proper chromosome segregation in budding yeast. At high concentrations the complex assembles into helical spirals (Westermann et al., 2005; Miranda et al., 2005). Importantly, the Dam1 complex interacts with the microtubule in a novel way and appears not to have a specific footprint on the lattice of the microtubule and instead the interaction is mediated by flexible regions on tubulin and/or the Dam1 complex (Westermann et al., 2006; Wang et al., 2007; Miranda et al., 2007). This property is likely to be an essential feature in its sliding mechanism and therefore on its biological role as a mitotic force coupler. Because of this type of interaction, the Dam1 spiral is not constrained to follow the symmetry of the microtubule. So, while this Dam1 helical arrangement around microtubules has the potential to aid structural analysis, it lacks the long-range helical order commonly seen for microtubule-interacting proteins that bind tightly to the microtubule lattice and are suitable for conventional helical reconstruction methodology. An alternative to traditional helical reconstruction is the use of the type of computational approaches pioneered by Egelman and coworkers that combine helical and single particle methods. However, the

Dam1 complex assembly has added challenges that required modification and extension of those approaches.

The major obstacles to obtain a structure of Dam1 spirals are: the distinct symmetry of the microtubule that confounds alignment of Dam1, and the variability and disorder of the Dam1 assembly itself. In order to overcome the first challenge we implemented an automated removal procedure for the microtubule structure from the experimental images. The ability to automatically remove microtubule signal from short segments will be useful for the study of future complexes found without a footprint on its lattice (e.g. relying on interaction with the flexible C-terminal tails of tubulin (or E-hooks) and/or by using flexible elements of their own for this interaction). To overcome the second challenge concerning Dam1 assembly flexibility, we successfully applied image classification techniques to address the problem of symmetry heterogeneity, identifying short, ordered helical stretches, and separating them into symmetry groups that were amenable to single particle helical reconstruction. By averaging within the groups we obtained homogeneous, high signal class averages. This allowed for a reasonable estimate of helical twist and rise per subunit –by inspection of the class-average Fourier transform, coupled with independent structural knowledge of the complex– and the ability to impose that symmetry on its corresponding class average to generate an initial 3-D model of the assembly.

Implementation of our method yielded a 30 Å reconstruction of the Dam1 complex in very good agreement with our previous reconstruction by traditional Fourier-Bessel methods, (Wang et al., 2007). Importantly, the new structure was obtained using short helical segments that were too disordered for traditional methods and had been discarded in our previous study. Application of this method for future studies concerning mutant complexes, labeled subunits, or interaction with other kinetochore components, will allow us to obtain structures by overcoming the requirement of long-range helical order.

In more general terms, we believe that ours is a broadly applicable approach that has the ability to deal with helical heterogeneity in order to ultimately reach higher resolution from helical macromolecular assemblies. The creation of homogeneous groups prevents the blurring and loss of resolution that accompany averaging images when disorder is present. Besides the case of disorder, polymers can frequently exist in several different well-defined helical states at different points along their length. Microtubules, actin filaments, and Dam1, among others, are examples. To correctly gain structural information about these assemblies, each type of helical symmetry must be treated separately. The helical symmetry parameters can then be estimated for each class and reconstructions done separately.

The application of this classification method to the Dam1 complex allowed for the robust reconstruction of an assembly that had proven difficult to treat by conventional means. It was accomplished using images of helices that did not maintain long range order for more than a few turns and therefore represents an extreme example of the capabilities of this algorithm. It is our hope that this method will be generally applicable, enabling the study of extremely heterogeneous samples and improving the resolution and accuracy of structures from samples with poor helical order.

Materials and Methods

Cryo-EM of Dam1 spirals

Micrographs discarded during our previous studies were used for this method (Wang et al., 2007). Briefly, purified Dam1 complex was incubated with GMPCPP stabilized microtubules overnight at room temperature. The solution was then applied to a holey carbon grid, plunged frozen into liquid ethane and kept at liquid nitrogen temperature during storage and observation.

Images were then taken on a FEI F20 microscope under low-dose conditions on Kodak SO163 films, which were digitized on a Nikon CoolScan 8000 at 2.56 Å/pixel.

Image Processing

Segments having a high degree of overlap were picked with EMAN's Boxer tool (Ludtke et al., 1999). Microtubule removal was performed using the SPIDER software package (Frank et al., 1996). The IMAGIC software package was used to perform the multi-reference alignment and multi-variant statistical analysis (van Heel et al., 1996), and an in-house program using topology representing networks (Ogura et al, 2003) was utilized to classify the images (available for download at <http://cryoem.berkeley.edu/software.shtml>). All 3-D processing was carried out in SPIDER, including application of symmetry to the class average and projection matching refinement to the final model (Egelman, 2007). The SUPRIM image processing package and a modification of MRC modules were used to calculate the phase residual between the current and previous structures (Schroeter and Bretaudiere, 1996; Wang and Nogales, 2005).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Patricia Grob for her useful discussions of helical symmetry, Niko Grigorieff for extremely valuable suggestions and experience, Edward Egelman for generously supplying his programs, and Andres Leschziner for a wealth of useful image processing scripts. We are thankful to Georjana Barnes, David Drubin and Stefan Westermann for our ongoing collaboration on the Dam1 complex studies. This work was supported by training grants from the US National Institutes of Health, as well as by a grant from the National Institute of General Medical Sciences of the US National Institutes of Health and by a Biomedicine chair from the BBVA Foundation to E.N.. E.N. is a Howard Hughes Medical Institute Investigator.

References

- Amos LA, Hirose K. The structure of microtubule-motor complexes. *Curr Biol* 1997;9:4–11.
- Arnal I, Metoz F, DeBonis S, Wade RH. Three-dimensional structures of functional motor proteins on microtubules. *Curr Biol* 1996;6:1265–1270. [PubMed: 8939577]
- Asbury CL, Gestaut DR, Powers AF, Franck AD, Davis TN. The Dam1 kinetochore complex harnesses microtubule dynamics to produce force and movement. *Proc Natl Acad Sci USA* 2006;103:9873–8. [PubMed: 16777964]
- Bluemke DA, Carragher B, Josephs R. The reconstruction of helical particles with variable pitch. *Ultramicroscopy* 1988;26:255–70. [PubMed: 3194997]
- Chen Y-J, Zhang P, Egelman EH, Hinshaw JE. The stalk region of dynamin drives the constriction of dynamin tubes. *Nat Struct Mol Biol* 2004;11:574–5. [PubMed: 15133500]
- DeRoiser DJ, Klug A. Reconstruction of Three Dimensional Structures from Electron Micrographs. *Nature* 1968;217:130–134.
- DeRosier DJ, Moore PB. Reconstruction of 3-dimensional images from electron micrographs of structures with helical symmetry. *J Mol Biol* 1970;52:355. [PubMed: 5485914]
- Egelman EH. A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy* 2000;85:225–34. [PubMed: 11125866]
- Egelman EH. The iterative helical real space reconstruction method: Surmounting the problems posed by real polymers. *J Struct Biol* 2007;157:83–94. [PubMed: 16919474]
- Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* 1996;116:190–9. [PubMed: 8742743]

- Frost A, Perera R, Roux A, Spasov K, Destaing O, Egelman EH, De Camilli P, Unger VM. Structural basis of membrane invagination by F-BAR domains. *Cell* 2008;132:807–17. [PubMed: 18329367]
- Galkin VE, Orlova A, Fatoum A, Walsh MP, Egelman EH. The CH-domain of calponin does not determine the modes of calponin binding to F-actin. *J Mol Biol* 2006;359:478–85. [PubMed: 16626733]
- Galkin VE, Orlova A, Briehner W, Kueh HY, Mitchison TJ, Egelman EH. Coronin-1A stabilizes F-actin by bridging adjacent actin protomers and stapling opposite strands of the actin filament. *J Mol Biol* 2008a;376:607–13. [PubMed: 18177666]
- Galkin VE, Orlova A, Cherepanova O, Lebart MC, Egelman EH. High-resolution cryo-EM structure of the F-actin-fimbrin/plastin ABD2 complex. *Proc Natl Acad Sci USA* 2008b;105:1494–8. [PubMed: 18234857]
- Glaeser, RM.; Downing, K.; DeRosier, D.; Chiu, W.; Frank, J. *Electron Crystallography of Biological Macromolecules*. Oxford University Press; New York, NY: 2007. p. 304
- Grishchuk EL, Spiridonov IS, Volkov VA, Efremov A, Westermann S, Drubin D, Barnes G, Ataullakhanov FI, McIntosh JR. Different assemblies of the DAM1 complex follow shortening microtubules by distinct mechanisms. *Proc Natl Acad Sci USA* 2008;105:6918–23. [PubMed: 18460602]
- Holmes KC, Angert I, Kull FJ, Jahn W, Schröder RR. Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide. *Nature* 2003;425:423–7. [PubMed: 14508495]
- Klug A, Crick FHC, Wyckoff HW. Diffraction of helical structures. *Acta Cryst* 1958;11:199–213.
- Li H, DeRosier DJ, Nicholson WV, Nogales E, Downing KH. Microtubule structure at 8 Å resolution. *Structure* 2002;10:1317–28. [PubMed: 12377118]
- Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 1999;128:82–97. [PubMed: 10600563]
- Mindell JA, Grigorieff N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 2003;142:334–47. [PubMed: 12781660]
- Miranda JLL, King DS, Harrison SC. Protein arms in the kinetochore-microtubule interface of the yeast DASH complex. *Mol Biol Cell* 2007;18:2503–10. [PubMed: 17460120]
- Miranda JLL, De Wulf P, Sorger PK, Harrison SC. The yeast DASH complex forms closed rings on microtubules. *Nat Struct Mol Biol* 2005;12:138–43. [PubMed: 15640796]
- Miyazawa A, Fujiyoshi Y, Unwin N. Structure and gating mechanism of the acetylcholine receptor pore. *Nature* 2003;423:949–55. [PubMed: 12827192]
- Narita A, Mizuno N, Kikkawa M, Maéda Y. Molecular determination by electron microscopy of the dynein-microtubule complex structure. *J Mol Biol* 2007;372:1320–36. [PubMed: 17761194]
- Ogura T, Iwasaki K, Sato C. Topology representing network enables highly accurate classification of protein images taken by cryo electron-microscope without masking. *J Struct Biol* 2003;143:185–200. [PubMed: 14572474]
- Orlova A, Garner EC, Galkin VE, Heuser J, Mullins RD, Egelman EH. The structure of bacterial ParM filaments. *Nat Struct Mol Biol* 2007;14:921–6. [PubMed: 17873883]
- Orlova EV, Rahman MA, Gowen B, Volynski KE, Ashton AC, Manser C, van Heel M, Ushkaryov YA. Structure of alpha-latrotoxin oligomers reveals that divalent cation-dependent tetramers form membrane pores. *Nat Struct Biol* 2000;7:48–53. [PubMed: 10625427]
- Pomfret AJ, Rice WJ, Stokes DL. Application of the iterative helical real-space reconstruction method to large membranous tubular crystals of P-type ATPases. *Journal of Structural Biology* 2007;157:106–16. [PubMed: 16879984]
- Sachse C, Chen J, Coureux PD, Stroupe ME, Fändrich M, Grigorieff N. High-resolution electron microscopy of helical specimens: a fresh look at tobacco mosaic virus. *J Mol Biol* 2007;371:812–35. [PubMed: 17585939]
- Schroeter JP, Bretauiere JP. SUPRIM: easily modified image processing software. *J Struct Biol* 1996;116:131–7. [PubMed: 8742734]
- Sosa HD, Dias DP, Hoenger A, Whittaker M, Wilson-Kubalek E, Sablin E, Fletterick RJ, Vale RD, Milligan RA. A model of the microtubule-Ncd motor protein complex obtained by cryo-electron microscopy and image analysis. *Cell* 1997;90:217–224. [PubMed: 9244296]

- Trachtenberg S, Galkin VE, Egelman EH. Refining the structure of the Halobacterium salinarum flagellar filament using the iterative helical real space reconstruction method: insights into polymorphism. *J Mol Biol* 2005;346:665–76. [PubMed: 15713454]
- van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M. A new generation of the IMAGIC image processing system. *J Struct Biol* 1996;116:17–24. [PubMed: 8742718]
- van Heel M. Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). *Ultramicroscopy* 1984;13:165–83. [PubMed: 6382731]
- van Heel M, Frank J. Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy* 1981;6:187–94. [PubMed: 7268930]
- Wang HW, Ramey VH, Westermann S, Leschziner AE, Welburn JPI, Nakajima Y, Drubin DG, Barnes G, Nogales E. Architecture of the Dam1 kinetochore ring complex and implications for microtubule-driven assembly and force-coupling mechanisms. *Nat Struct Mol Biol* 2007;14:721–6. [PubMed: 17643123]
- Wang HW, Nogales E. An iterative Fourier-Bessel algorithm for reconstruction of helical structures with severe Bessel overlap. *Journal of Structural Biology* 2005;149:65–78. [PubMed: 15629658]
- Wang YA, Yu X, Ng SYM, Jarrell KF, Egelman EH. The Structure of an Archaeal Pilus. *J Mol Biol* 2008;381:456–466. [PubMed: 18602118]
- Westermann S, Avila-Sakar A, Wang HW. Formation of a dynamic kinetochore- microtubule interface through assembly of the Dam1 ring complex. *Mol Cell* 2005;17(2):277–90. [PubMed: 15664196]
- Westermann S, Wang HW, Avila-Sakar A, Drubin DG, Nogales E, Barnes G. The Dam1 kinetochore ring complex moves processively on depolymerizing microtubule ends. *Nature* 2006;440:565–9. [PubMed: 16415853]
- White HE, Saibil HR, Ignatiou A, Orlova E. Recognition and separation of single particles with size variation by statistical analysis of their images. *J Mol Biol* 2004;336:453–60. [PubMed: 14757057]
- Yonekura K, Maki-Yonekura S, Namba K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* 2003;424:643–50. [PubMed: 12904785]
- Zhang P, Hinshaw JE. Three-dimensional reconstruction of dynamin in the constricted state. *Nat Cell Biol* 2001;3:922–6. [PubMed: 11584275]

Figure 1.

a) Micrograph containing several Dam1 decorated microtubules. Scale bar 1000 Å. Inset shows the short ordered helical segment used in our previous studies (Wang et al., 2007). It was one of only a few found after screening hundreds of assembled segments and only sustained order over ~10 turns. Inset scale bar 200 Å. b) CryoEM image of a typical microtubule fully decorated with a Dam1 spiral. Scale bar 200 Å. c) Power spectrum of image in (b). The 40 Å layer line corresponding to the repeat of tubulin in the microtubule is clearly visible (solid white arrow), while only one blurred layer line is visible for the Dam1 spiral (white dashed arrow). This layer line corresponds to the ~100 Å pitch of this variable Dam1 helical assembly. The lack of other layer lines confirms that this assembly does not maintain order past 100 Å over long stretches. Inset shows power spectrum of inset in (a).

Figure 2.

Preprocessing of images to remove the microtubule contribution and to box single turns. a) Raw segment of Dam1 assembled around a microtubule. The 1-D density profile for this image is shown above. b) The central portion of the 1-D density profile in (a) corresponding to the microtubule protofilaments has been projected in 2D to make this mask. Compare 1-D profile above with profile in (a). c) Resulting image and 1-D profile after real space subtraction of image in (b) from image in (a). Both the image and the 1-D profile show the removal of the central microtubule density. d), e) and f) Power spectra of images in (a), (b), and (c) respectively. These illustrate how the masking, which is performed entirely in real space, is equivalent to masking the microtubule's contribution to the equator in Fourier space. g)

Masking of the microtubule's 40 Å layer line in Fourier space (right) removes the signal from the axial tubulin repeat in the image (left). h) Segment from (g) that has had all but one to two turns of Dam1 helix masked out with a soft mask (left). Power spectrum of this small stretch (right) shows that for the size of helical fragment that is required to treat extreme disorder, traditional Fourier space methods of reconstruction fail due to layer line broadening and overlap. The image in (h) is then carried over for 2D processing.

Figure 3.

a) Class averages that represent well-ordered particles of homogeneous symmetry based on a visual assessment of their features and analysis of variance. Averages with sharply defined edges, a clear repeat along the spiral, and a flat 2-D variance were said to be 'well ordered'. b) Averages from the same classification whose features are blurred and represent disordered segments or minority groups such as the small number of particles that have a specific out of plane tilt. c) Class average used to generate the initial model for refinement. While there are clear features along the length of the spiral they cannot be interpreted directly to give helical parameters. d) Power spectrum of the image in (c). Layer lines describing the one start helix (solid white arrow) are strong and clear, while the reflections from its 'supertwist' (area inside white box) are broad and overlapping. This overlap, stemming from the small number of turns in the average, prevents indexing by conventional methods. However, the distance of these reflections from the meridian, (dashed arrows), gives an estimate of the number of subunits

per turn. Furthermore, from the images the helical arrangement appears as a 1-start helix with a 2-fold axial symmetry. Taken together, these observations narrow the range of helical parameters sufficiently to be able to test them by convergence during refinement. All averages in this figure generated by the method of Ogura et al. (Ogura et al., 2003).

Figure 4.

Eigen images generated during reference-free 2D alignment and classification using MSA. Eigen images are ordered with respect to their contribution to sorting (i.e. those with biggest effect first). a) First eight eigen images. These describe large, strong features in the spirals. They all describe the positions, pitch, and diameter of spirals and lack features corresponding to the repeating unit. b) Eigen images 9–18, which mainly contain the repeating pattern of the complex around the spiral. Taken together these 18 eigen images efficiently sort helical segments based on their pitch, diameter and view, into classes characterized by homogeneous symmetry or disorder.

Figure 5.

Visual overview of the reconstruction process. Helical segments are boxed, preprocessed to remove the microtubule, and then aligned and classified to reduce heterogeneity. 2-D averages can then be used to estimate helical symmetry and create an initial model for refinement by single particle methods. Dashed panel on the left compares a central slab from one side of the initial model (right) with the refined structure (left). This area is indicated on the full structures by dashed boxes. The two have the same general morphology but fine features are gained as refinement progresses. This comparison is a quality control step that helps establish that the refinement has progressed smoothly. If the initial and final models do not match, either the refinement has diverged to an incorrect solution or the starting symmetry was far from the correct value. Matching does not guarantee the correct solution, however, and other metrics, like comparing reprojections of the refined model with its class average are necessary to rule out artificial structures.

Figure 6.

Structure generated by this method and its superposition with the structure obtained by the traditional helical Fourier Bessel method (Wang et al., 2007). a) Current structure (blue) modeled around a reconstruction of a 13-protofilament MT (gold). b) FSC curve for the asymmetric unit in the current structure compared to that solved previously. The 0.5 cutoff occurs just above 30 Å, the resolution of both reconstructions. c,d) Comparison with previous structure (gray mesh) in side cutaway and inner surface views. These visually confirm the good agreement of these two structures. In (c) the microtubule is gold, (d) as seen from the microtubule, looking outwards.