SPRING - an image processing package for single-particle based helical reconstruction from electron cryomicrographs.

Ambroise Desfosses, Rodolfo Ciuffa, Irina Gutsche, Carsten Sachse

Journal of Structural Biology 185 (2014) 15–26

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

SPRING – An image processing package for single-particle based helical reconstruction from electron cryomicrographs

Ambroise Desfosses [a,b,c,d], Rodolfo Ciuffa [a], Irina Gutsche [b,c,d], Carsten Sachse [a,*]

[a] EMBL – European Molecular Biology Laboratory, Structural and Computational Biology Unit, Meyerhofstr. 1, 69917 Heidelberg, Germany
[b] Univ. Grenoble Alpes, UVHCI, F-38000 Grenoble, France
[c] CNRS, UVHCI, F-38000 Grenoble, France
[d] Unit for Virus Host-Cell Interactions, Univ. Grenoble Alpes-EMBL-CNRS, 6 rue Jules Horowitz, 38042 Grenoble, France

ARTICLE INFO

Article history:
Received 14 October 2013
Received in revised form 12 November 2013
Accepted 13 November 2013
Available online 21 November 2013

Keywords:
Helical assembly
Electron cryomicroscopy
Software
Image processing
3D reconstruction
Helical symmetry

ABSTRACT

Helical reconstruction from electron cryomicrographs has become a routine technique for macromolecular structure determination of helical assemblies since the first days of Fourier-based three-dimensional image reconstruction. In the past decade, the single-particle technique has had an important impact on the advancement of helical reconstruction. Here, we present the software package SPRING that combines Fourier based symmetry analysis and real-space helical processing into a single workflow. One of the most time-consuming steps in helical reconstruction is the determination of the initial symmetry parameters. First, we propose a class-based helical reconstruction approach that enables the simultaneous exploration and evaluation of many symmetry combinations at low resolution. Second, multiple symmetry solutions can be further assessed and refined by single-particle based helical reconstruction using the correlation of simulated and experimental power spectra. Finally, the 3D structure can be determined to high resolution. In order to validate the procedure, we use the reference specimen Tobacco Mosaic Virus (TMV). After refinement of the helical symmetry, a total of 50,000 asymmetric units from two micrographs are sufficient to reconstruct a subnanometer 3D structure of TMV at 6.4 Å resolution. Furthermore, we introduce the individual programs of the software and discuss enhancements of the helical reconstruction workflow. Thanks to its user-friendly interface and documentation, SPRING can be utilized by the novice as well as the expert user. In addition to the study of well-ordered helical structures, the development of a streamlined workflow for single-particle based helical reconstruction opens new possibilities to analyze specimens that are heterogeneous in symmetries.

1. Introduction

Structure determination of large macromolecular assemblies embedded in vitreous ice using electron microscopy (EM) is becoming increasingly popular as evidenced by the steady increase in the number of structure depositions into the EM databank (EMDB) (Lawson et al., 2011). Depending on the molecular weight and order of the assembly, a series of three-dimensional (3D) structures at near-atomic resolution have become available in the past decade. Pioneering work on highly symmetric structures derived from two-dimensional arrays, helical or icosahedral assemblies (Henderson et al., 1990; Unwin, 2005; Zhang et al., 2008) have demonstrated the potential of electron cryomicroscopy (cryo-EM) based structure determination. Historically, the first 3D reconstructions were computed from electron micrographs of helical assemblies (De Rosier and Klug, 1968). These assemblies have the advantage that a single helix already represents many views of the asymmetric unit whose structure needs to be determined.

Currently, helical assemblies make up ~10% of the determined structures in the entire EMDB. This is due to the fact that only a limited number of proteins form arrays of helical symmetry. Nevertheless, many of these are functional in the helical state and as such, are of fundamental importance to the cell (Moore et al., 1970; Nogales et al., 1999). Several structures that mediate the modulation of membrane shapes have been determined with the protein coat assembled at the membrane in a helical geometry (Frost et al., 2008; Low et al., 2009). In addition, there are examples of helical assemblies that form protein crystals in the context of a tubular membrane (Korkhov et al., 2010; Unwin, 1993). Such assemblies have also been successfully formed by affinity-tagged membrane-associated proteins (Wilson-Kubalek et al., 1998).

* Corresponding author.
E-mail address: carsten.sachse@embl.de (C. Sachse).

1047-8477
http://dx.doi.org/10.1016/j.jsb.2013.11.003

Methods for structure determination of helical assemblies have significantly evolved since the birth of 3D electron microscopy. In the past, the procedure relied on entire and straight filaments, filamentous viruses or tubules that were processed in Fourier space by indexing the helical lattice and extracting the amplitudes and phase from the corresponding layer lines. Multiple helices were averaged and brought to a common phase origin and a 3D reconstruction was computed by Fourier inversion of the structure factors. For a more comprehensive description consult (Stewart, 1988). More recently, adapted Fourier-based techniques and real-space approaches that treat helices as small segments have significantly improved the attainable resolution (Beroukhim and Unwin, 1997; Ge and Zhou, 2011; Sachse et al., 2007; Yonekura et al., 2003). In addition, some helical assemblies deviate from their ideal straight path and can also vary in their helical symmetries because of inherent flexibilities (Fujii et al., 2010; Sachse et al., 2008). In certain cases, Fourier-based helical 3D reconstruction can be complicated by particular symmetries. First, in cases of long helical pitches many layer lines are required to represent the entire helical structure as in the case of amyloid fibrils. Second, several layer lines can interfere on a single reciprocal pixel line and the resulting Bessel overlap makes the assignment of Bessel order impossible. Nevertheless, real-space helical reconstruction can cope with these complications and determine the 3D structures of these helical assemblies (Jiménez et al., 1999; Sachse et al., 2008).

Despite the previous successes of helical structure determination, a simple standardized workflow for 3D helical reconstruction is still lacking. The most widely used approach is the implementation of the iterative helical real-space reconstruction (IHRSR) based on the SPIDER package (Frank et al., 1996) and additional tools for helical symmetry determination and imposition (Egelman, 2000). In the meantime, other packages such as SPARX have adapted the IHRSR algorithm (Behrmann et al., 2012). Moreover, several structures have been determined by extending and modifying the original IHRSR approach significantly with additional SPIDER operations (Sachse et al., 2007). Using a full correction of the contrast-transfer function, alignment restraints and an adapted 3D symmetrization procedure, a series of structures were determined (Bharat et al., 2012; Korkhov et al., 2010; Low et al., 2009; Sachse et al., 2007; Sachse et al., 2008). In order to condense the adapted procedures into a generally usable workflow, we describe here a package for single-particle based helical reconstruction termed SPRING (Single particle reconstruction from images of known geometries). We demonstrate the full functionality of the package by processing a subset of previously published micrographs of Tobacco Mosaic Virus (TMV) (Sachse et al., 2007) (http://grigoriefflab.janelia.org/datadownload). SPRING contains programs that determine the microscope parameters, analyze and classify the segmented helices, explore helical symmetry at low resolution, refine high-resolution symmetry and determine the 3D structure.

2. Overview

SPRING aims to provide a comprehensive workflow for processing electron micrographs of helical specimens from micrographs to 3D structure analysis and interpretation. The workflow has been subdivided into three separate suites of programs: "Springmicrograph", "Spring2d" and "Spring3d" (Table 1). In SPRINGMICROGRAPH, digital micrographs can be analyzed and processed. The extraction and analysis of helical segments is implemented in the second suite, SPRING2D. The third suite of programs, SPRING3D, generates, refines and analyzes 3D structures. The individual programs can be operated from a graphical user interface (GUI) (Fig. 1), from the command line prompt, from command line options or using a simple text file as input parameter file. In all programs, the user can specify three levels of expertise: beginner, intermediate and expert. The beginner level reduces the complexity of the input parameters by using sensible default values. As their familiarity with the processing operations increases the user can choose to add more parameters. In the current implementation of SPRING, a significant effort was invested to streamline analysis and diagnosis of the obtained results in a user-friendly manner. Where possible, either condensed graphical plots are generated or more complex data representations can be browsed interactively (Fig. 2).

The SPRING package is entirely written in object-oriented python and uses EM-related libraries and functions from SPARX and EMAN2 (Hohn et al., 2007). Microscope parameters are determined by CTFFIND and CTFTILT (Mindell and Grigorieff, 2003). In addition, scientific computing tasks are performed by Numpy and Scipy functions (numpy.scipy.org). For parameter storage sqlite3 databases are used and interfaced by SQLAlchemy (www.sqlalchemy.org). Interactive and diagnostic plots were made with the plotting libraries of matplotlib (http://matplotlib.sourceforge.net). SPRING's GUI has been built using PyQT libraries. SPRING is optimized to run in a multi-CPU environment on high-performance computer cluster implemented by Mpi4Py (http://mpi4py.scipy.org).

Python can be used as a scripting language as well as a structured programming language. Both of these features make the usage of isolated existing functions in a new processing context and the easy modification of SPRING possible. In addition, the widespread use of python as a programming language and the excellent interfaces to scientific computing libraries such as Numpy and Scipy are a great advantage for prototyping any numerical computations and thus promoting further development of the package. The python programming language facilitates code structuring and readability and the code is directly documented and available on SPRING's website as a detailed reference (Fig. 1B). Python has become a popular tool to master the scripting and programming tasks in a variety of other EM software packages such as PyTOM and Xmipp (Hrabe et al., 2012; Scheres et al., 2008).

3. Initial analysis of micrographs and segments

The EM operator records electron micrographs in several different ways. Currently, film, CCD cameras and direct detectors are the common sources of EM data. After film has been digitized all types of data are available as images in various formats. SPRING accepts all the formats of micrograph data that EMAN2 currently supports such as standard MRC, IMAGIC, SPIDER, TIF formats. MICEXAM examines the micrographs by analyzing the power spectra tiles to exclude images that suffer from poor information transfer at higher resolutions due to charging or drift (Hohn et al., 2007). MICCTFDETERMINE determines the CTF of the micrographs by interfacing with CTFFIND initially and optionally refines parameters using CTFTILT (Mindell and Grigorieff, 2003). The program captures a reduced output of CTFFIND and CTFTILT and the results are stored in the SPRING database to be retrieved for further processing.

After the selection of high-quality micrographs, helices need to be extracted from the images. For this purpose, the helices are interactively picked using external programs. In the past, EMAN's HELIXBOXER or BOXER with the helix option was used (Ludtke et al., 1999) (EMAN2 has an updated version named E2HELIXBOXER). BSOFT is also capable of picking filaments with significant curvature and recording their helix paths (Heymann and Belnap, 2007). The program SEGMENT from SPRING extracts a complete data set of overlapping segments using the provided coordinates from either EMAN, EMAN2 or BSOFT, applies CTF correction by either phase-flipping or convolving the segments with the determined CTF and stores coordinates and the derived in-plane

Table 1
Summary of individual SPRING programs and associated functionality.

Program	Description
Micrograph analysis (Springmicrograph)	
Micexam	Program to examine micrograph quality by computing a localized power spectrum using EMAN2's e2scaneval.py and an averaged power spectrum from overlapping tiles using SPARX' sx_welch_pw2.py
Micctfdetermine	Program to determine CTF parameters from a set of micrographs using CTFFIND and CTFTILT (Mindell and Grigorieff, 2003)
Segment analysis (Spring2d)	
Segment	Program to extract overlapping segments from micrographs
Segmentexam	Program to examine all of excised in-plane rotated segments and compute their collapsed (1D) and 2D power spectrum and width profile of helices
Segmentclass	Program to classify excised in-plane rotated segments using SPARX's k-means clustering
Segclassexam	Program to examine helix classes to compute their collapsed (1D) and 2D power spectrum and width profile of helices
Segclasslayer	Program to extract amplitudes and phases from desired layer lines of class averages
Seglayer2lattice	Program to simulate helical diffraction pattern and plot helical lattice from a series of indexed layer lines or rise/rotation parameters
Segmentplot	Program to plot parameters from segmented helices
Segment 3D reconstruction (Spring3d)	
Segclassreconstruct	Program to compute 3D reconstruction from a single class average using a range of different helical symmetries
Seggridexplore	Program to interactively explore grid searches according to different criteria
Segmentrefine3d	Program to iteratively refine a 3D structure of helical specimens from segment stacks
Segrefine3plot	Program to plot refinement parameters from segmentrefine3d
Segrefine3dgrid	Program to optimize segmentrefine3d reconstruction by varying refinement parameters systematically on a grid
Segrefine3dinspect	Program to interactively inspect 3D reconstructions from segmentrefine3d

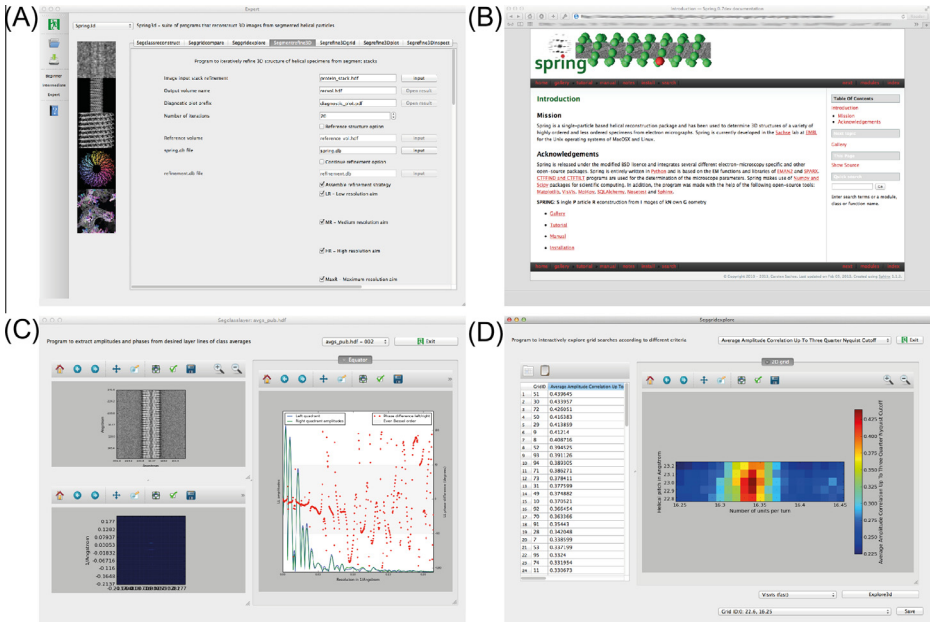

Fig.1. SPRING - a package for single-particle based helical processing. (A) GUI of SPRING to launch individual programs. The GUI can be visualized in beginner, intermediate and expert mode with increasing complexity of the input parameters. (B) Website and documentation for SPRING. All input parameters are listed and described on the website. (C) SEGCLASSLAYER - GUI program to plot left and right layer line profile with corresponding amplitude and phase difference. (D) SEGGRIDEXPLORE - GUI program to interactively visualize grids e.g. for optimization of helical symmetry parameters.

rotation in a database file. For further 2D analysis, the stack of in-plane rotated segments is analyzed by adding their power spectra using SEGMENTEXAM. The presence or absence of layer lines is a good indicator of the degree of order and symmetry of the specimen. In particular, the reduction of the power spectrum into a "collapsed" one-dimensional profile by averaging the spectrum along the layer lines is useful to assess the existence of weak layer lines. Furthermore, width profiles of segments allow a quick characterization of the specimen. For further characterization, SEGMENTCLASS iteratively classifies and aligns the segment stack using a k-means clustering algorithm from SPARX (Hohn et al., 2007) (Fig. 3A–D). Segment assignment to classes is recorded in a database and can be used as a selection criterion at a later stage during the 3D structure refinement. The resulting classes can be further analyzed in real space and in the Fourier domain. SEGCLASSEXAM can extract width and power spectrum from the 2D class averages with significantly improved signal-to-noise ratio when compared with raw segments (Fig. 3D/E). The process of segmenting helices for classification-based averaging using a regular step size bears the risk of introducing artificial frequency components (Fig. 3B/C). The artifacts show high intensities along all expected layer lines up to highest resolution but do not follow the intensity profile of the expected Bessel functions. These artifacts arise as a consequence of using overlapping segments. If a small and constant step size is chosen, the images are overlapping. Thus, the overlapping image information produces a repeat in the average of neighboring segments that will be enhanced during the classification procedure. In order to address this issue, the classification procedure is

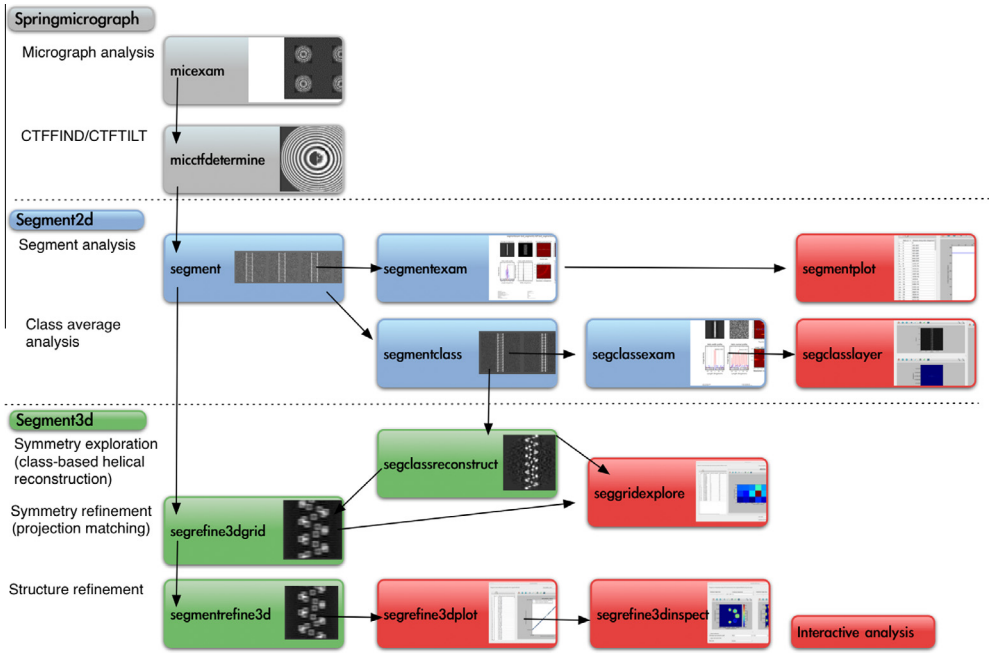

Fig.2. Flow chart of processing helical specimens including micrograph analysis, CTFFIND/CTFTILT, segment and class average analysis. The class average can be used for symmetry exploration. The helical symmetry parameters can be refined by a projection-matching based symmetry grid. Finally, the structure can be refined to highest resolution. Interactive GUI programs (red) help to analyze results of the respective programs.

performed with a different stack that was not segmented by a regular step size but by using a randomized step between 0 and the desired step size. This approach results in power spectra that do not show artificially introduced repeats (Fig. 3E). In addition, SEGCLASSLAYER allows the interactive exploration of layer lines from class averages (Fig. 1C). First, class averages have improved signal-to-noise ratio in comparison with raw images. Second, classification can separate different symmetries from a heterogeneous data set. Third, in comparison with the added power spectra from SEGMENTEXAM (Fig. 3F) class averages have the additional advantage that they contain phase information and can therefore be used to determine whether Bessel orders are odd or even due to the characteristic phase difference between left and right sides of the layer line. In the case of TMV, the inspection of the 1/23 and 1/11.5 1/Å layer lines reveals that they have a Bessel order of 1 and 2 due to the position of the first amplitude maximum and phase difference of left and right layer line profile (Fig. 3G). The assignment of the Bessel order is essential to index the helical lattice, which is a common procedure for determination of helical symmetry of helices (Diaz et al., 2010; Korkhov et al., 2010). Once all helices are segmented the program SEGMENTPLOT visualizes quantities of segments interactively, e.g. coordinates or in-plane angle rotation as a function of the segment assignment to classes.

4. Class-based helical reconstruction and symmetry exploration

The determination of helical symmetry is one of the most critical and time-consuming steps in the structure determination of helical specimens. Awareness of this problem has been raised previously (Yu and Egelman, 2010) and stems from the fact that projections of helices with different helical symmetries can be identical as a set of helical densities at varying radii can give rise to the same projections in two dimensions. For a detailed discussion using TMV as an example refer to (Egelman, 2010). Thus, an inherent problem remains the choice of the initial reference symmetry that severely biases any structure refinement steps because these related symmetries will "converge" to a stable solution but only one of them corresponds to the true symmetry. Currently, the IHRSR approach iteratively refines the helical symmetry but it strongly depends on the choice of the starting parameters. In previous approaches other groups have derived starting estimates of helical symmetry parameters from classical Fourier analysis (Pomfret et al., 2007; Ramey et al., 2009) to further refine the symmetry by the provided IHRSR programs. Alternatively, a motif-based maximum likelihood estimator has been formulated to determine the helical symmetry parameters (Lee et al., 2011). In principle, the symmetry problem could be overcome, if all 3D reconstructions with all possible combinations of helical rise and helical rotation during 3D structure refinement were computed at the highest possible resolution. In the presence of an atomic structure with a characteristic fold, subnanometer resolution should be sufficient to unambiguously place the structure into the EM density reconstructed using the true symmetry parameters and thus verify the imposed helical symmetry. By contrast, an EM density reconstructed with the wrong helical symmetry could not be fitted with the PDB structure. Currently, such an exhaustive search is computationally not feasible in particular with small increments between different symmetries.

In order to address this problem, we propose a class-based helical reconstruction scheme that explores the helical symmetry by reconstructing the 3D structures from a large number of symmetry combinations using a single class average. Since a single view already represents many projections of a helix, a single class average is sufficient to reconstruct a low-resolution initial 3D reconstruction. Subsequently, the class average will be duplicated and multiples of helical rise and rotation will be applied to image duplicates that are subsequently used for the 3D reconstruction. A similar approach has been applied to class averages to generate an initial 3D structure once the symmetry had been determined through indexing the helical lattice (Ramey et al., 2009). The program SEGCLASSRECONSTRUCT performs the class-based helical reconstruction. The underlying hypothesis is that if the correct helical symmetry

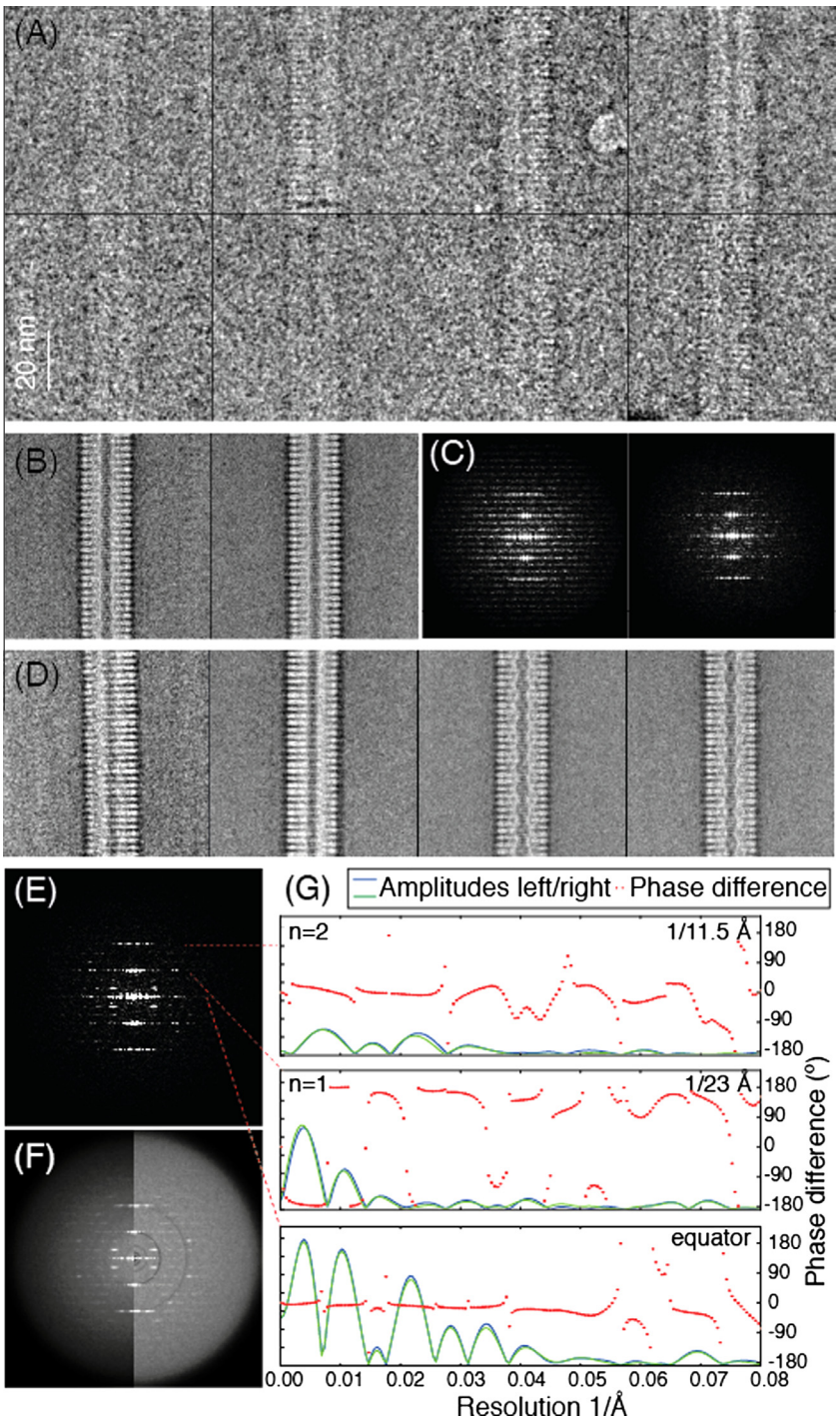

Fig.3. Segment analysis of TMV. (A) Eight TMV segments of 670 Å dimension from cryo-EM images. Segments were rotated in the plane and phase-flipped by the determined contrast-transfer function. (B) Classification results from repeated cycles of k-means clustering and multi-reference alignment using a segmentation step size of 70 Å and 55 Å (left and right panel respectively). (C) Corresponding power spectra computed from (B) give rise to noticeable artifacts of the overlapping segments. (D) Four exemplary classes from a random segmentation step from 0 to 70 Å. (E) The power spectrum reveals layer lines up to 1/11.5 1/Å and is devoid of artifacts from the fourth class of (B). (F) High-resolution power spectrum computed as a sum of power spectra of in-plane rotated segments. Right. Enhanced power spectrum obtained by B-factor compensation. (G) Closer layer-line analysis reveals that 1/23 and 1/11.5 1/Å have an order of 1 and 2 supported by the peak position and the phase difference from layer line profiles (E).

was applied, the reprojection will match the original class average. This assumption should hold true for real-space and Fourier-space comparisons.

In order to quantify this comparison, we have measured amplitude correlation from the power spectra of the class average and reprojections of the 3D model, the cross-correlation of the class average and the reprojections of the 3D model and the variance of the 3D reconstruction from a large number of possible symmetry combinations (Fig. 4A–D). We have applied this strategy to a class average from TMV (Fig. 4F). The symmetry was searched by varying the helical pitch from 22.0 to 24.0 Å in 0.1 Å steps and the number of units per turn from 10.00 to 20.00 in 0.01 steps. First, amplitude correlations were determined by correlating the power spectra of the duplicates of the class average with the power

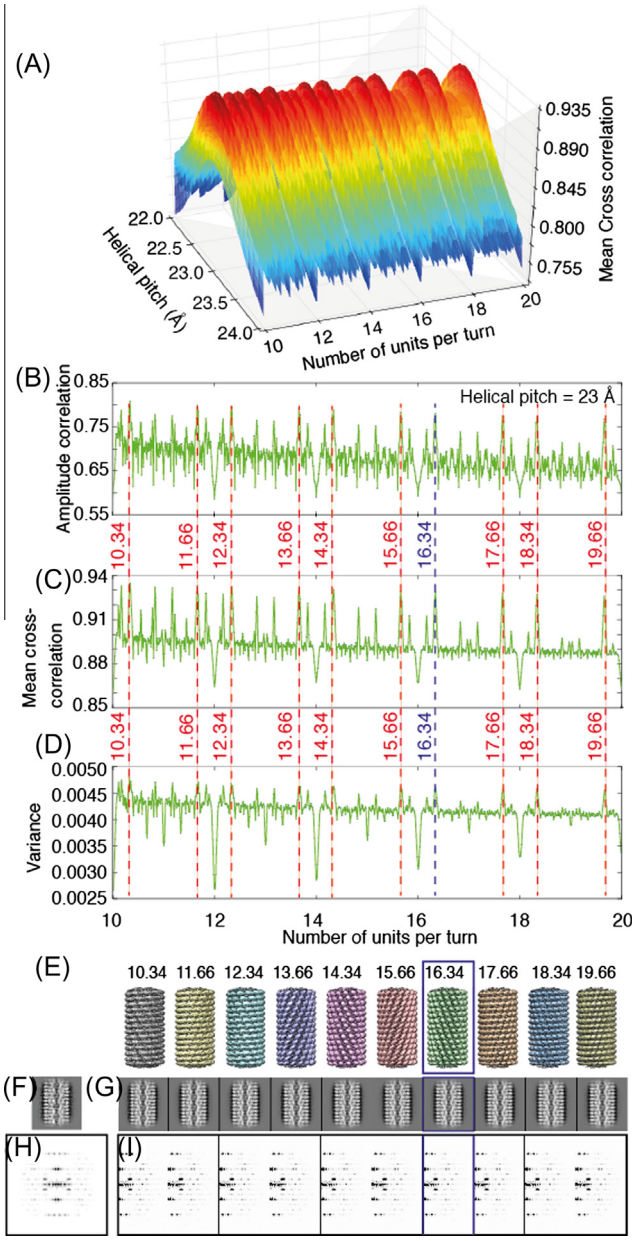

Fig.4. Helical symmetry exploration based on TMV class averages reveals a number of possible solutions at low resolution. (A) 3D surface presentation of mean cross-correlation between class average power spectrum and power spectrum from reprojection of 3D reconstruction from a grid of helical pitch (22–24) and number of units per turn (10–20). (B) Mean amplitude correlation profile at 23.0 Å helical pitch. (C) Mean cross-correlation profile at 23.0 Å helical pitch as a function of number of units per turn indicates several possible solutions for the helical symmetry. (D) Variance of 3D reconstruction. (E) Ten 3D reconstructions with number of units per turn from 10.34 to 19.66 imposed. (F) Class average used to generate 3D reconstructions. (G) Ten reprojections from (E). (H) Power spectrum of class average from (F). (I) Ten half power spectra computed from reprojections (G).

spectra of the corresponding reprojections of the 3D model compared up to 1/7 1/Å resolution and averaged. Second, cross-correlations were computed between the duplicates of the class average and the corresponding reprojections of the 3D model and averaged. Therefore, from here on we refer to these measurements as mean cross-correlation and mean amplitude correlation. The mean cross-correlation grid gave rise to high correlations at a pitch of 23 Å and several maxima ranging from 10.34 to 16.34 number of units per turn (Fig. 4A/C). Similarly, the mean amplitude correlation grid confirmed the observations resulting in a slightly noisier profile when compared with the mean cross-correlation. We noticed by visual inspection that the 3D reconstructions can differ significantly (Fig. 4E), while the 2D reprojections and power spectra of these 3D reconstructions are virtually indistinguishable (Fig. 4F–I). In conclusion, the grids demonstrate that multiple symmetry solutions can be obtained using this method without the need of structure refinement. Therefore, the class-based helical reconstruction can limit the symmetry space to be explored and thus provides good starting estimates for launching a confined symmetry grid search of iterative structure refinements (see section on symmetry refinement below). In some cases, indexing of a helical lattice may help to limit the number of solutions. In order to better access and evaluate the large grid of different symmetry parameters, we developed the interactive GUI-based SEGGRIDEXPLORE program (Fig. 1D) where multiple comparison criteria can be browsed simultaneously and the associated data files such as the simulated power spectrum, the corresponding experimental power spectrum and 3D reconstruction can be immediately opened by clicking on the respective grid point. This option of interactively browsing the results makes the evaluation of these many symmetry combinations easily accessible and user-friendly.

To further test the applicability of the class-based reconstruction approach, we generated test classes from 3D reconstructions of the flagellar hook and Respiratory Syncytial Virus (RSV) nucleocapsid (EMDB 1647/ 1622, respectively) (Fujii et al., 2009; Tawar et al., 2009). In agreement with the TMV analysis, when evaluating symmetry combinations with increasing number of units per turn along the constant pitch, we noticed that the ambiguous symmetry peaks of the mean cross-correlation are not randomly distributed. Instead they repeat every two units per turn and possess a mirror symmetry around integer number of units per turn (Fig. 5B/D). Although not explicitly described, a similar relationship between ambiguous helical symmetries has already been observed (Chen et al., 2004; Egelman, 2010; Wang et al., 2006). In addition to the cited examples, the here-investigated cases all share this ambiguity pattern. It is important to note that all considered cases are one-start helices and cases with additional rotational symmetry further complicate the interpretation of the symmetry pattern. For one-start helices, the relationship can be understood by examining the reciprocal space pattern arising from such symmetric structures. First, all solutions share the same exact repeat distance but differ in their number of units per repeat and thus they will give rise to layer lines at the same reciprocal heights (Fig. S1, Table S1). Second, the orders of the Bessel functions at a given layer line from the ambiguous solutions are either always odd or always even. Thus the parity of the Bessel orders at the same layer lines is conserved (Table 2, Appendix). For all ambiguous solutions, the absolute first peak position of the layer line is identical thus the change in Bessel order is compensated by a radius change of the real-space densities. For a more detailed mathematical treatment of helical ambiguity, refer to the Appendix. Nevertheless, there is only one true helical symmetry among the ambiguous solutions, whose 3D reconstruction can be interpreted with the correct protein structure.

In practice, we noticed several complications of class-based helical reconstruction that will limit the results for particular sets of helical symmetry parameters. First, we prefer to use phase-flipped in-plane rotated images rather than CTF-convolved images because the low-resolution frequency signal is stronger when compared with CTF-convolved images. Second, the helical parameters of rise and rotation dictate the number of images that can be used for the 3D reconstruction. For example, for a helical rise half of the image dimension only two images can be used. This can in principle be compensated for by using larger image segments for classification but is limited by the straightness of the helix. Furthermore,

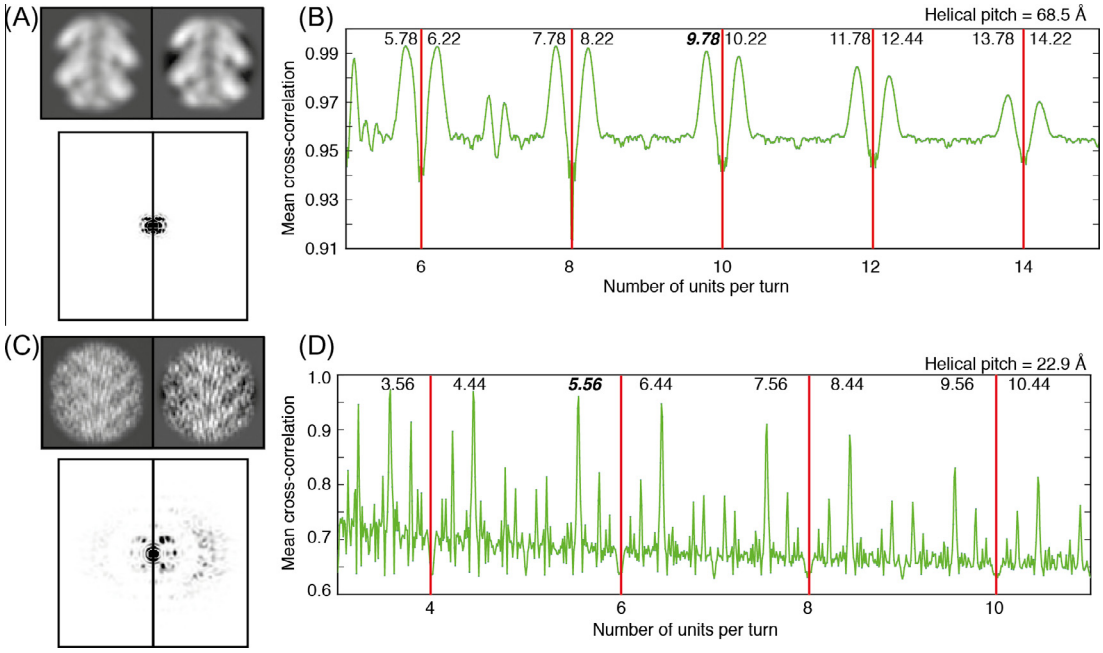

Fig.5. Class-based symmetry determination profiles at constant pitches for RSV and flagellar hook. (A) Side-by-side comparison of RSV class average (left) and reprojection (right) from true symmetry and corresponding side-by-side power spectra. (B) Mean cross-correlation profile of class average with reprojections. (C) Side-by-side display of flagellar hook class average and reprojection from correct symmetry and corresponding power spectra. (D) Mean cross-correlation profile as a function of number of units per turn give rise to a number of symmetry solutions that are ambiguous. The ambiguous solutions are related to each other such that they repeat every two units per turn including a mirror symmetry around integer values.

Table 2
Comparison of Bessel orders of layer lines from ambiguous symmetry solutions (odd and even Bessels orders are highlighted in red and blue).

n odd	*n* even	Number of units per turn					
Layer line position	Layer-line position (1/Å)	13.66	14.34	15.66	16.34	17.66	18.34
1	1/67.6	14	−14	16	−16	18	−18
2	1/34.8	−13	15	−15	17	−17	19
3	1/23.0	1	1	1	1	1	1
4	1/17.2	15	−13	17	−15	19	−17
5	1/13.9	−12	16	−14	18	−16	20
6	1/11.5	2	2	2	2	2	2

helical rotations that have integer divisors of 360° will result in sparse sampling of the azimuthal rotation angle. For example, a helical rotation of 60° will only be inserted six times into the 360° regardless of the helical rise and the number of image duplicates. As a result, the 3D reconstructions contain significant artifacts because of the gaps in the angular coverage. Third, prior knowledge of the out-of-plane tilt angle of the helix is required for a correct 3D reconstruction. If no tilt stage was induced, the out-of-plane angles are usually Gaussian distributed around zero and thus the majority of classes will not show a significant out-of-plane tilt. Inspection of the phase differences using the program SEGCLASSLAYER along the layer lines will indicate whether the class average view is tilted significantly out of plane (Diaz et al., 2010; Stewart, 1988).

5. Projection-matching based symmetry refinement

The symmetry solutions of the class-based symmetry exploration can now be assessed until the correct fold of the protein is identified. This way, we can rule out ambiguous solutions from the symmetry search. Therefore, we performed a 3D structure refinement with the symmetries at a pitch of 23 Å and 13.66, 14.34, 15.66, 16.34, 17.66 and 18.34 subunits per turn (Fig. 6). For this computation, we used a total of 1400 segments from 40 viruses extracted from two micrographs that were taken at 1.5 and 3.5 μm underfocus (Sachse et al., 2007) (http://grigoriefflab.janelia.org/datadownload). As a result, only the 3D reconstruction with 16.34 numbers of subunits per turn imposed is compatible with the known four-helix bundle structure (Fig. 6, PDB code 2OM3, 2TMV) (Namba et al., 1989; Sachse et al., 2007) and the outer densities in particular are best defined in comparison with the other 3D structures. It is important to note that refinement convergence and statistics such as FSC and azimuthal angle distribution do not reveal whether the symmetry is correct or wrong. Some solutions may even be interpreted as a fold containing α-helices but are clearly incompatible with the atomic reference structure. In conclusion, the aim remains to determine an EM-map at sufficient resolution to verify the 3D reconstruction and symmetry in order to recognize a characteristic fold or polypeptide backbone at medium to high resolution (4–10 Å) or an unambiguous fit of the PDB structure at lower resolution (>10 Å).

To further refine the correct symmetry, we performed a symmetry grid search of iterative 3D refinements around 16.34 numbers of subunits per turn for the TMV data set using the program SEGREFINE3DGRID. In comparison with the single-class

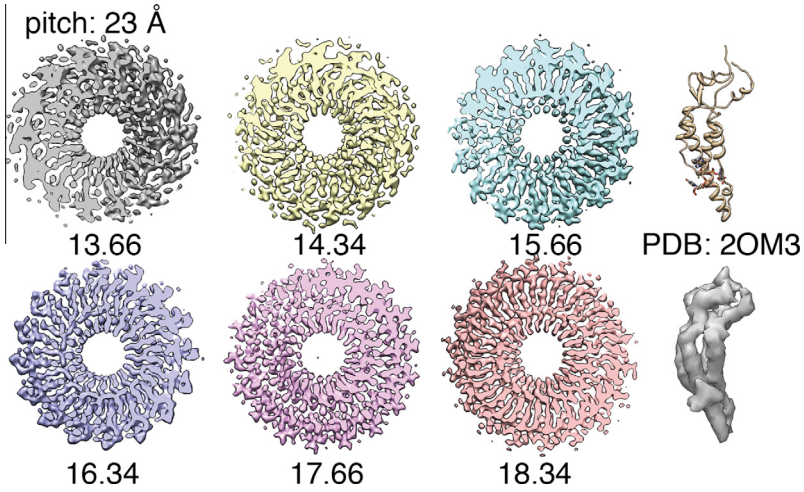

Fig.6. Comparison of high-resolution 3D TMV structures by imposing different ambiguous symmetries. 3D structures were sharpened by a B-factor of -200 1/Å^2 filtered to a resolution of 7 Å. Slab view through helical rod. The reconstructions share the pitch of 23 Å and possess 13.66, 14.34, 15.66, 16.34, 17.66 and 18.34 subunits per turn, respectively. Only the correct reference structure with 16.34 numbers of subunits per turn is compatible with the four-helix bundle atomic structure (PDB code 2OM3).

reconstructions, this approach has the advantage that limited angular sampling can be overcome if many segments contribute to an even angular coverage. As the iterative refinement is a CPU-expensive task, we chose a smaller subset of data and restricted the number of iteration cycles to 4–8.

Using the TMV data set, we performed a refined search of pitches ranging from 22.8 to 23.2 with a step size of 0.1 and number of units per turn ranging from 16.25 to 16.45 with a step size of 0.01 (Fig. 7). First, we evaluated the amplitude correlation between experimental and simulated projection. For each grid point, we averaged all power spectra of in-plane rotated segments that showed no out-of-plane tilt and compared it with the power spectra of the matched reprojections (Fig. 7A/D/E). The advantage of this approach over a mean cross-correlation measure between segments and projections is that no translational alignment information is required to compute and compare high-resolution power spectra of the segments (Fig. 4E). Thus, this approach avoids the known effects of noise bias when comparing references with noisy segments (Stewart and Grigorieff, 2004). In the case of TMV, we can refine the symmetry by comparing the layer lines beyond 1/7.67 (0.13) 1/Å. Here, we have taken the same principal approach as in the cases of helical tubules of TspO and bacterial dynamin (Korkhov et al., 2010; Low et al., 2009). In addition, we have tested a series of different criteria to evaluate and distinguish reconstructions with different symmetries: variance of the 3D reconstruction, number of excluded segments and mean cross-correlation peak that represents the similarity measure of the projection matching procedure (Fig. 7B/C). The obtained pitch of 23.0 Å and number of units per turn of 16.34, i.e. 49.02 subunits for three turns, are in exact agreement with results from fiber diffraction (Stubbs and Makowski, 1982; Stubbs et al., 1977). Finally, the amplitude correlation criterion remains most reliable and it has the advantage that a simple visual inspection can confirm whether the refinement has found a good agreement between experimental and simulated power spectra.

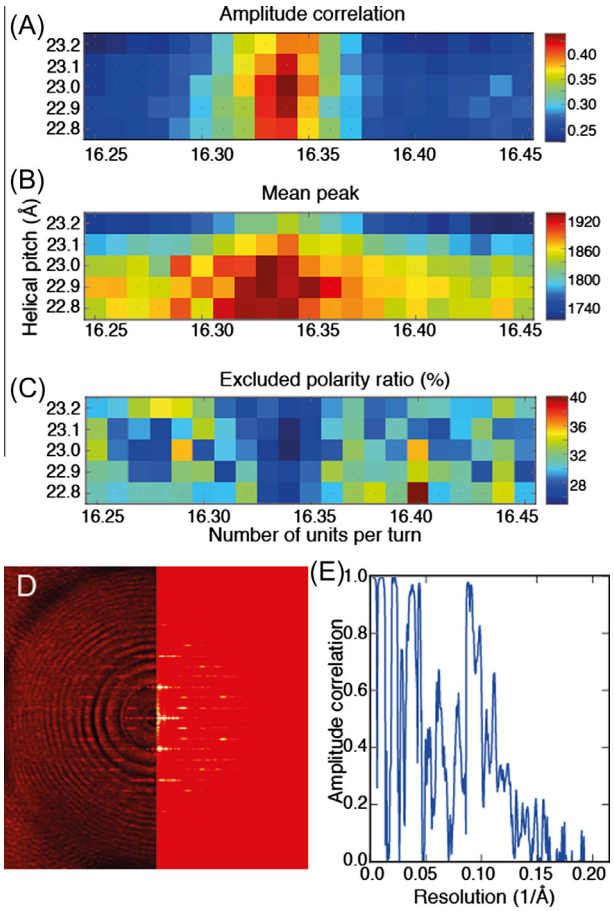

Fig.7. Symmetry refinement of TMV. Grid search of different symmetry combinations from 22.8 to 23.2 pitch and from 16.25 to 16.45 number of units per turn. (A) Amplitude correlation of symmetry grid search. The highest correlation is found at a pitch of 23.0 Å and 16.34 units per turn corresponding to a helical rise of 1.408 Å and 22.03°. (B) The mean cross-correlation peak measure of all matched segments plotted as a function of the changed symmetry parameters. The mean peak shows slight deviation from the determined number of units per turn. (C) A similar convergence can be found by plotting the ratio of excluded segments based on incorrect polarity matches. The minimum of 25 % also at 23.0 Å pitch and 16.34 units per turn. (D) Side-by-side display of highest correlation of experimental and simulated power spectrum. (E) Resolution-dependent amplitude correlation of experimental with simulated power spectrum indicates correlation up to 0.13 1/Å.

6. High-resolution helical reconstruction

Our previous description of single-particle helical based reconstruction for TMV was implemented as a prototype using SPIDER scripts (Sachse et al., 2007). The current article focuses on the bundling and extension of the previous approach into an efficient and

reproducible image processing pipeline with significant optimizations (Fig. 2). In SPRING, we have re-implemented many of the features using the python API of SPARX/EMAN2. Briefly, we would like to summarize the important features of the structure refinement that we have implemented as previously described. For a more detailed discussion of these topics consult our article (Sachse et al., 2007). Once the overlapping segments have been extracted the images are convolved with the corresponding CTF as estimated from the CTFTILT measurement and the position of the segment on the micrograph. The original in-plane rotation of the segments is retained in the images and recorded in a database. This step is critical for correcting astigmatism of the CTF on the segment level and beneficial to avoid multiple interpolation steps when segments are included in the 3D reconstruction. These segments are subjected to projection matching and followed by 3D reconstruction. The polarity assignments of segments within a helix are recorded and excluded if an opposite polarity has been found. Furthermore, alignment restraints are applied to exclude excessive shifts against a mean path of the helix. The 3D reconstruction is computed from the matched segments and symmetry-related views that are not covered by the overlapping segments are included in the 3D reconstruction. The helical symmetry is fixed during the entire run. The resulting 3D reconstruction is divided by a 3D CTF squared with a Wiener filter constant to compensate for amplitude modulations from the specific set of defoci used for image acquisition (Sachse et al., 2007).

In single-particle structure refinement, it is crucial to be aware of the inherent noise bias and the danger of over-fitting noise during the alignment procedure (Scheres and Chen, 2012; Stewart and Grigorieff, 2004) which can result in wrong or over-refined structures. In order to address this issue, we have implemented two precautionary measures to avoid over-fitting. First, we derive a low-pass filter from the square root of the FSC curve of the previous cycle, which is applied during projection matching. This filter will gradually include higher resolution signal as the resolution of the 3D reconstruction improves. Second, we implement a strict low-pass filter cutoff at 11 Å resolution to the alignment procedure which excludes the risk of over-fitting in the resolution range where it is most prominent. In agreement with previous results, we found that the low-resolution frequencies provide sufficient detail for alignment and 3D reconstruction at high resolution (Rosenthal and Henderson, 2003).

High-resolution structure refinement is computationally expensive and requires the use of multiple processors. For this reason, a significant effort was invested in improving the performance of 3D structure refinement. First, seamless parallelization was implemented by Mpi4Py – a python interface to OpenMPI. Moreover, segments are downscaled internally in multiple resolution steps to reduce computational cost. The complete refinement is broken up into appropriate resolution aim ranges that the user is able to select (Table 3). The underlying principle is that a 3D reconstruction at the targeted resolution can be obtained using a pixel size where 3× pixel size is assumed as the practical resolution limit due to interpolation errors. SPRING will determine the required decimation factor to achieve the respective resolution aim. Thus, the segments are low-pass filtered and subsequently decimated and as a result the convergence of the refinement is accelerated.

Table 3
Summary of refinement strategy of TMV structure determination.

Refinement strategy	Low	Medium	High	Maximum
Resolution aim (Å)	>20	20–10	10–5	<5
Decimation factor	8	3	2	–
Pixel size (Å)	8.141	3.489	2.326	1.163
Number of iterations	4	4	4	4

By default, the refinement projects 90 projections around the helical axis and assumes seven possible out-of-plane tilt angles from −12° to 12°. The azimuthal angles do not increase by a constant step but are distributed such that the asymmetric unit is regularly azimuthally sampled and so only unique helical views are tested for projection matching (Ge and Zhou, 2011). The out-of-plane tilt step follows an even distribution between 0 and the maximum tilt angle. The number of angular projections will be further increased by a factor of five for high-resolution refinements. To reduce computational cost further, at a later stage of refinement the number of projections tested for projection matching are reduced such that the search is limited to views around ±20° and later by ±2° around the previous match of the previous cycle. The following 3D reconstruction step tends to be computationally expensive depending on the pixel dimensions of the segment. We, therefore, reconstruct a reduced size volume of helix width dimensions and for the subsequent step of projection helically expand the volume back to the intended dimension required for alignment.

In order to further demonstrate the complete functionality of the package, we performed a structure refinement with a smaller subset of our previous data set used for the prototype of high-resolution reconstruction. With a total of 50,000 asymmetric units extracted from two micrographs we computed a 3D reconstruction at 6.4 Å resolution (Fig. 8A–F). The final 3D structure was interactively B-factor sharpened by −150 1/(Å)^2 and filtered to a resolution cutoff of 5 Å using the program SEGREFINE3DINSPECT to correctly display the EM densities (Rosenthal and Henderson, 2003). The total refinement time of 16 iterations from low to highest resolution took 45 h on 48 Intel Xeon 2.4 GHz processors. Next, we wondered how the resolution of SPRING reconstructions compares with the previously published structure that was computed with SPIDER scripts (Sachse et al., 2007). Consequently, we determined a structure using the previous data set of 200,000 asymmetric units and measured very similar resolutions with cutoffs at 4.8/4.3 Å (FSC 0.5/0.143) when compared with the previously determined 4.7/4.3 Å (EMDB-1316) (Fig. 8F). At closer inspection of the map, the EM densities provide a very similar level of detail for structure interpretation (Fig. 8H/I).

In order to increase the homogeneity of single-particle data sets, classification strategies are routinely employed that rely on principal component analyses (van Heel and Frank, 1981) or cross-correlation based approaches. In helical processing in particular, projection-matching based sorting of varying pitches has been applied to improve structural homogeneity of particular helical symmetries (Wang et al., 2006). In order to conveniently track the membership of segments to micrographs, helices, classes and other microscope-related parameters SPRING relies on an SQL database that records the available data of each segment. If desired, the user can apply segment exclusion criteria that are based on properties before any refinement is done, e.g. class assignment, micrograph, helix, defocus and astigmatism. In addition, segment exclusion can also be based on refinement parameters such as cross-correlation peaks, out-of-plane tilt and excessive x-shifts. For further analysis, refinement parameters can easily be plotted as a function of distance along the helix using the program SEGREFINE3DPLOT. The segment database will simplify processing of large heterogeneous data sets and enable comprehensive selection schemes to obtain 3D structures from homogeneous subsets of the data.

7. Outlook

The development of the here-described processing package SPRING will facilitate many of the steps of single-based particle helical reconstruction. One of the most time-consuming processing steps, however, will remain the determination of the helical

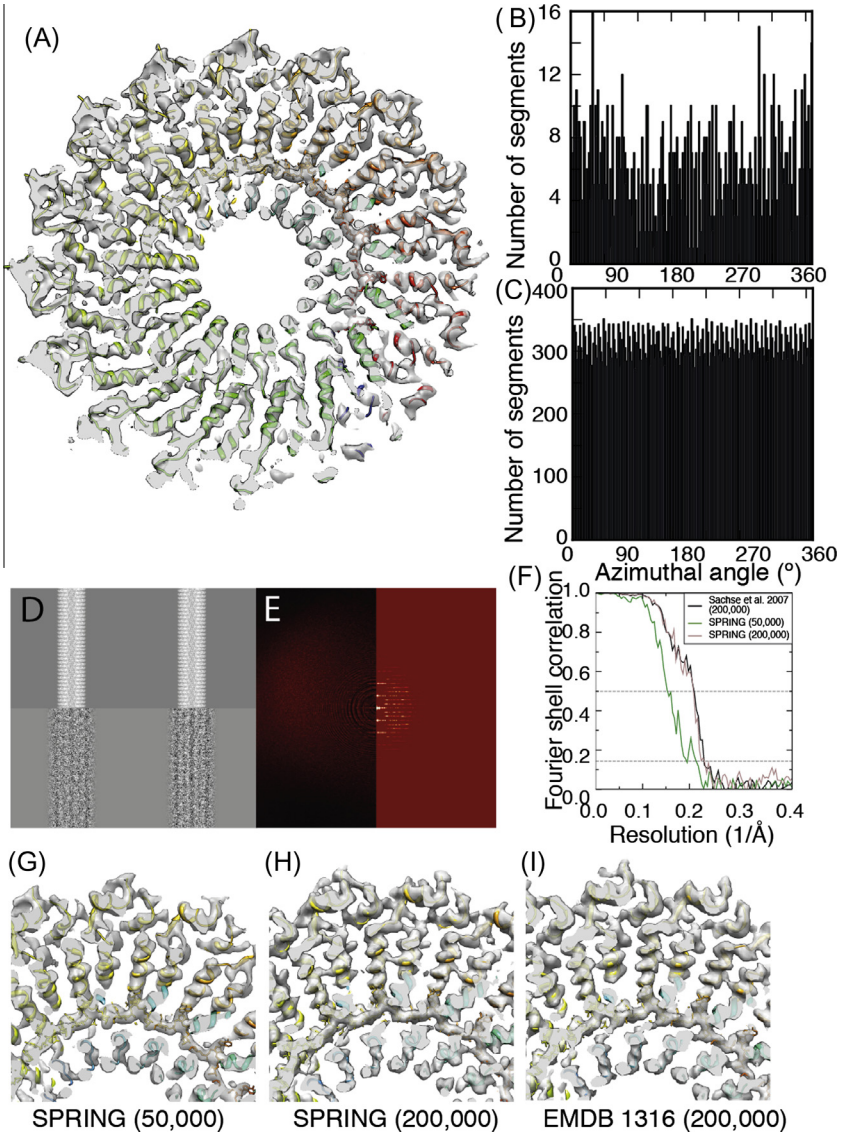

Fig.8. High-resolution helical reconstruction of TMV of 50,000 asymmetric units from two micrographs. (A) Final 3D density map fitted with PDB code 2OM3 at 6.4 Å resolution including a B-factor sharpening of −150 $1/(Å)^2$ using the program SEGREFINE3DINSPECT that allows interactive B-factor application to avoid oversharpening of the structure. (B) Histogram of azimuthal angles per segments in degrees. (C) Histogram of azimuthal angles in degrees after symmetrization. (D) Projections of final structure and corresponding segment. (E) Side-by-side display of B-factor compensated experimental and simulated power spectrum. (F) Using 50,000 asymmetric units from two micrographs the final resolution of the SPRING reconstruction is estimated at 6.4/5.1 Å using the Fourier Shell Correlation criteria at 0.5/0.143. In comparison with the Sachse et al., 2007 data set of 200,000 asymmetric units, SPRING computes a very similar structure with resolutions estimated at 4.8/4.3 Å at the respective FSC cutoffs. (G/H/I) Close-up views of EM densities including fitted PDB (2OM3) structures from SPRING reconstructions using 50,000 and 200,000 asymmetric units (G/H) in comparison with reference structure from Sachse et al., 2007 (EMDB-1316).

symmetry parameters because validation is really only possible once the obtained resolution allows an unambiguous fit of a high-resolution PDB model. This ambiguity in 2D projections may be overcome by additional experimental approaches to evaluate the helical symmetry directly from 3D tomograms as was previously demonstrated on retroviral Mason-Pfizer Monkey virus tubes (Bharat et al., 2012). Those initial lattice parameters can be further refined using SPRING without the need of large-scale exploration and validation of symmetry parameters. In conclusion, SPRING will significantly speed up helical structure determination by merging many processing steps into a single processing workflow. Furthermore, it will enable the user to address more challenging samples that exhibit significant heterogeneity in their helical symmetries or helical order. With the advent of direct electron detectors additional improvements in image quality can be expected and will increase the obtainable resolution of helical specimens with fewer asymmetric units and for example facilitate the processing of data sets of helices with different symmetry families.

The SPRING website (http://www.sachse.embl.de/emspring) provides a tutorial on how to process the two described micrographs of the TMV data set used in this publication.

Acknowledgments

We are grateful to Pawel Penczek for helpful advice on the SPARX package. We appreciate feedback on the manuscript from Tanmay Bharat and Abul Tarafder. In addition, we thank Andres Lindau and Michael Wahlers from the EMBL-IT support for assistance with high-performance computing.

Appendix

Ambiguous solutions observed in the mean cross-correlation patterns are related to each other by the expression shown in Eq. (1). Consider three one-start helices A, B and C of the same pitch P with the number of subunits per turn $N(A)$, $N(B)$ and $N(C)$:

$$\begin{cases} N(A) = Z - \Delta f \\ N(B) = Z + \Delta f \\ N(C) = Z - \Delta f + 2 \end{cases} \quad (1)$$

All numbers N can be written in the form $Z + \Delta f$ where Z is the closest integer to the number of units per turn and Δf the remaining fractional difference to the closest integer ($0 < \Delta f <= 0.5$). Those helices represent ambiguous solutions of helical symmetry that are observed when comparing a 2D projection of helix A, B or C with reprojections of helices A, B and C. In the example of TMV, helix B corresponds to the true symmetry at 16.34 subunits per turn whereas helix A and C are incorrect symmetry at 15.66 and 17.66, respectively. The same ambiguous relationship holds true for a set of artificial helices H1, H2 and H3 that are illustrated for in Figure S1.

Because helices A, B and C share the fractional difference Δf but differ in Z, they will have an exact repeat after the same number of turns t with $t*\Delta f$ being the smallest integer multiple of Δf (Table S1). As the three helices have the same pitch P, their repeat c will occur at the same axial distance $t*P$. Therefore, the Fourier Transform of 2D projections of those helices will be confined to layer lines at same positions ζ:

$$\zeta = \frac{l}{c} \text{Å}^{-1} \quad (2)$$

where l is an integer that represents the l-th layer line. On each layer line, the order of contributing Bessel functions n are restricted by the selection rule

$$\zeta = \frac{n}{P} + \frac{m}{p} = \frac{l}{c} \text{Å}^{-1} \quad (3)$$

that relates the layer lines positions ζ to the pitch P, the rise per subunit p and the integers n and m. By replacing the axial rise p by the pitch over the number of subunits per turn as expressed in (1), and writing the selection rule for the three helices A, B and C, we obtain after simplification:

$$\begin{cases} \zeta(A) = \frac{n(A)+m(A)Z}{P} - \frac{m(A)\Delta f}{P} \\ \zeta(B) = \frac{n(B)+m(B)Z}{P} + \frac{m(B)\Delta f}{P} \\ \zeta(C) = \frac{n(C)+m(C)Z}{P} - \frac{m(C)f}{P} + \frac{2m(C)}{P} \end{cases} \quad (4)$$

If we assume that helices A, B, and C share layer line positions ζ, so that

$$\begin{cases} \zeta(B) = \zeta(A) \\ \zeta(C) = \zeta(A) \end{cases} \quad (5)$$

We can first deduce relationships of values of m between helices A and B and helices A and C by multiplying left and right side of (5) by P, using the expression of ζ given in (5), and further simplification:

$$\begin{cases} n(B) + m(B)Z - n(A) - m(A)Z = -\Delta f(m(A) + m(B)) \\ n(C) + m(C)Z + 2m(C) - n(A) - m(A)Z = \Delta f(m(C) - m(A)) \end{cases} \quad (6)$$

The left part of Eq. (6) is by definition an integer because m, n and Z are confined to integers. In the simplest case, it follows:

$$\begin{cases} m(B) = -m(A) \\ m(C) = m(A) \end{cases} \quad (7)$$

We deduce a relationship between the possible orders of Bessel function for helices A, B and C on each layer line, using Eqs. (4), (5) and (7):

$$\begin{cases} n(B) = n(A) + 2Zm(A) \\ n(C) = n(A) - 2m(A) \end{cases} \quad (8)$$

For any considered layer line, Bessel function orders of helices A, B and C differ from each other by a factor of two. Thus, we conclude that the parity of Bessel function orders is conserved between ambiguous solutions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jsb.2013.11.003.

References

Behrmann, E., Tao, G., Stokes, D.L., Egelman, E.H., Raunser, S., Penczek, P.A., 2012. Real-space processing of helical filaments in SPARX. J. Struct. Biol., 1–12.

Beroukhim, R., Unwin, N., 1997. Distortion correction of tubular crystals: improvements in the acetylcholine receptor structure. Ultramicroscopy 70, 57–81.

Bharat, T.A.M., Davey, N.E., Ulbrich, P., Riches, J.D., de Marco, A., Rumlova, M., Sachse, C., Ruml, T., Briggs, J.A.G., 2012. Structure of the immature retroviral capsid at 8 Å resolution by cryo-electron microscopy. Nature 487, 385–389.

Chen, Y.-J., Zhang, P., Egelman, E.H., Hinshaw, J.E., 2004. The stalk region of dynamin drives the constriction of dynamin tubes. Nat. Struct. Mol. Biol. 11, 574–575.

De Rosier, D., Klug, A., 1968. Reconstruction of three dimensional structures from electron micrographs. Nature 217, 130–134.

Diaz, R., Rice, W.J., Stokes, D.L., 2010. Fourier-Bessel reconstruction of helical assemblies. Methods Enzymol. 482, 131–165.

Egelman, E.H., 2000. A robust algorithm for the reconstruction of helical filaments using single-particle methods. Ultramicroscopy 85, 225–234.

Egelman, E.H., 2010. Reconstruction of helical filaments and tubes. Methods Enzymol. 482, 167–183.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., Leith, A., 1996. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. J. Struct. Biol. 116, 190–199.

Frost, A., Perera, R., Roux, A., Spasov, K., Destaing, O., Egelman, E.H., De Camilli, P., Unger, V.M., 2008. Structural basis of membrane invagination by F-BAR domains. Cell 132, 807–817.

Fujii, T., Iwane, A.H., Yanagida, T., Namba, K., 2010. Direct visualization of secondary structures of F-actin by electron cryomicroscopy. Nature 467, 724–728.

Fujii, T., Kato, T., Namba, K., 2009. Specific arrangement of alpha-helical coiled coils in the core domain of the bacterial flagellar hook for the universal joint function. Structure 17, 1485–1493.

Ge, P., Zhou, Z.H., 2011. Hydrogen-bonding networks and RNA bases revealed by cryo electron microscopy suggest a triggering mechanism for calcium switches. Proc. Natl. Acad. Sci. USA 108, 9637–9642.

Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E., Downing, K.H., 1990. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. J. Mol. Biol. 213, 899–929.

Heymann, J.B., Belnap, D.M., 2007. Bsoft: image processing and molecular modeling for electron microscopy. J. Struct. Biol. 157, 3–18.

Hohn, M., Tang, G., Goodyear, G., Baldwin, P.R., Huang, Z., Penczek, P.A., Yang, C., Glaeser, R.M., Adams, P.D., Ludtke, S.J., 2007. SPARX, a new environment for Cryo-EM image processing. J. Struct. Biol. 157, 47–55.

Hrabe, T., Chen, Y., Pfeffer, S., Cuellar, L.K., Mangold, A.-V., Förster, F., 2012. PyTom: a python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. J. Struct. Biol. 178, 177–188.

Jiménez, J.L., Guijarro, J.I., Orlova, E., Zurdo, J., Dobson, C.M., Sunde, M., Saibil, H.R., 1999. Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. EMBO J. 18, 815–821.

Korkhov, V.M., Sachse, C., Short, J.M., Tate, C.G., 2010. Three-dimensional structure of TspO by electron cryomicroscopy of helical crystals. Structure 18, 677–687.

Lawson, C.L., Baker, M.L., Best, C., Bi, C., Dougherty, M., Feng, P., van Ginkel, G., Devkota, B., Lagerstedt, I., Ludtke, S.J., et al., 2011. EMDataBank.org: unified data resource for CryoEM. Nucleic Acids Res. 39, D456–D464.

Lee, S., Doerschuk, P.C., Johnson, J.E., 2011. Multiclass maximum-likelihood symmetry determination and motif reconstruction of 3-d helical objects from projection images for electron microscopy. IEEE Trans. Image Process. 20, 1962–1976.

Low, H.H., Sachse, C., Amos, L.A., Löwe, J., 2009. Structure of a bacterial dynamin-like protein lipid tube provides a mechanism for assembly and membrane curving. Cell 139, 1342–1352.

Ludtke, S., Baldwin, P., Chiu, W., 1999. EMAN: semiautomated software for high-resolution single-particle reconstructions. J. Struct. Biol. 128, 82–97.

Mindell, J.A., Grigorieff, N., 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. J. Struct. Biol. 142, 334–347.

Moore, P.B., Huxley, H.E., DeRosier, D.J., 1970. Three-dimensional reconstruction of F-actin, thin filaments and decorated thin filaments. J. Mol. Biol. 50, 279–295.
Namba, K., Pattanayek, R., Stubbs, G., 1989. Visualization of protein-nucleic acid interactions in a virus. Refined structure of intact tobacco mosaic virus at 2.9 A resolution by X-ray fiber diffraction. J. Mol. Biol. 208, 307–325.
Nogales, E., Whittaker, M., Milligan, R.A., Downing, K.H., 1999. High-resolution model of the microtubule. Cell 96, 79–88.
Pomfret, A.J., Rice, W.J., Stokes, D.L., 2007. Application of the iterative helical real-space reconstruction method to large membranous tubular crystals of P-type ATPases. J. Struct. Biol. 157, 106–116.
Ramey, V.H., Wang, H.-W., Nogales, E., 2009. Ab initio reconstruction of helical samples with heterogeneity, disorder and coexisting symmetries. J. Struct. Biol. 167, 97–105.
Rosenthal, P.B., Henderson, R., 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721–745.
Sachse, C., Chen, J.Z., Coureux, P.-D., Stroupe, M.E., Fändrich, M., Grigorieff, N., 2007. High-resolution electron microscopy of helical specimens: a fresh look at tobacco mosaic virus. J. Mol. Biol. 371, 812–835.
Sachse, C., Fändrich, M., Grigorieff, N., 2008. Paired beta-sheet structure of an Abeta(1–40) amyloid fibril revealed by electron microscopy. Proc. Natl. Acad. Sci. USA 105, 7462–7466.
Scheres, S.H.W., Chen, S., 2012. Prevention of overfitting in cryo-EM structure determination. Nat. Methods 9, 853–854.
Scheres, S.H.W., Núñez-Ramírez, R., Sorzano, C.O.S., Carazo, J.M., Marabini, R., 2008. Image processing for electron microscopy single-particle analysis using XMIPP. Nat. Protoc. 3, 977–990.
Stewart, A., Grigorieff, N., 2004. Noise bias in the refinement of structures derived from single particles. Ultramicroscopy 102, 67–84.
Stewart, M., 1988. Computer image processing of electron micrographs of biological structures with helical symmetry. J. Electron Microsc. Tech. 9, 325–358.
Stubbs, G., Makowski, L., 1982. Coordinated use of isomorphous replacement and layer-line splitting in the phasing of fiber diffraction data. Acta. Cryst. A 38, 417–425.
Stubbs, G., Warren, S., Holmes, K., 1977. Structure of RNA and RNA binding site in tobacco mosaic virus from 4-Å map calculated from X-ray fibre diagrams. Nature 267, 216–221.
Tawar, R.G., Duquerroy, S., Vonrhein, C., Varela, P.F., Damier-Piolle, L., Castagné, N., MacLellan, K., Bedouelle, H., Bricogne, G., Bhella, D., et al., 2009. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. Science 326, 1279–1283.
Unwin, N., 1993. Nicotinic acetylcholine receptor at 9 A resolution. J. Mol. Biol. 229, 1101–1124.
Unwin, N., 2005. Refined structure of the nicotinic acetylcholine receptor at 4A resolution. J. Mol. Biol. 346, 967–989.
van Heel, M., Frank, J., 1981. Use of multivariate statistics in analysing the images of biological macromolecules. Ultramicroscopy 6, 187–194.
Wang, Y.A., Yu, X., Yip, C., Strynadka, N.C., Egelman, E.H., 2006. Structural polymorphism in bacterial EspA filaments revealed by cryo-EM and an improved approach to helical reconstruction. Structure 14, 1189–1196.
Wilson-Kubalek, E.M., Brown, R.E., Celia, H., Milligan, R.A., 1998. Lipid nanotubes as substrates for helical crystallization of macromolecules. Proc. Natl. Acad. Sci. USA 95, 8040–8045.
Yonekura, K., Maki-Yonekura, S., Namba, K., 2003. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. Nature 424, 643–650.
Yu, X., Egelman, E.H., 2010. Helical filaments of human Dmc1 protein on single-stranded DNA: a cautionary tale. J. Mol. Biol. 401, 544–551.
Zhang, X., Settembre, E., Xu, C., Dormitzer, P.R., Bellamy, R., Harrison, S.C., Grigorieff, N., 2008. Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. Proc. Natl. Acad. Sci. USA 105, 1867–1872.