Journal of Structural Biology 186 (2014) 234–244

Contents lists available at ScienceDirect

Journal of Structural Biology

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

# Frealix: Model-based refinement of helical filament structures from electron micrographs

Alexis Rohou, Nikolaus Grigorieff*

*Department of Biochemistry, Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454, USA*
*Janelia Farm Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, VA 20147, USA*

ARTICLE INFO

*Article history:*
Received 9 January 2014
Received in revised form 12 March 2014
Accepted 13 March 2014
Available online 20 March 2014

*Keywords:*
Amyloid
Structure
Cryo-EM
Curvature
Filament
Helical

ABSTRACT

The structures of many helical protein filaments can be derived from electron micrographs of their suspensions in thin films of vitrified aqueous solutions. The most successful and generally-applicable approach treats short segments of these filaments as independent "single particles", yielding near-atomic resolution for rigid and well-ordered filaments. The single-particle approach can also accommodate filament deformations, yielding sub-nanometer resolution for more flexible filaments. However, in the case of thin and flexible filaments, such as some amyloid-β (Aβ) fibrils, the single-particle approach may fail because helical segments can be curved or otherwise distorted and their alignment can be inaccurate due to low contrast in the micrographs. We developed new software called Frealix that allows the use of arbitrarily short filament segments during alignment to approximate even high curvatures. All segments in a filament are aligned simultaneously with constraints that ensure that they connect to each other in space to form a continuous helical structure. In this paper, we describe the algorithm and benchmark it against datasets of Aβ(1–40) fibrils and tobacco mosaic virus (TMV), both analyzed in earlier work. In the case of TMV, our algorithm achieves similar results to single-particle analysis. In the case of Aβ(1–40) fibrils, we match the previously-obtained resolution but we are also able to obtain reliable alignments and ~8-Å reconstructions from curved filaments. Our algorithm also offers a detailed characterization of filament deformations in three dimensions and enables a critical evaluation of the worm-like chain model for biological filaments.



## 1. Introduction

Electron micrographs of thin (tens of nanometers) films of vitrified, dilute suspensions of biological macromolecules and their assemblies can be analyzed to deduce their three-dimensional (3D) structure (Frank, 2006). Cases in which proteins form filamentous assemblies are particularly well suited to structure determination, since a single image of a helical filament can yield a full tomographic series of projections through the helical protomer. Helical filaments permitted some of the earliest examples of structure determination by electron microscopy (De Rosier and Klug, 1968) and continue to be studied in many different contexts (DeRosier, 2007).

In recent years, most investigators have been using variations on iterative algorithms originally developed for the study of so-called single particles. When adapted to helical structure determination, these methods treat short (tens of nanometers) segments of imaged filaments as independent projections of the unknown 3D structure to be determined. The length of those segments is chosen carefully. They must be long enough to allow for reliable and accurate alignment against projections calculated from the current 3D reconstruction yet, because no helical filament is perfectly rigid, short enough to approximate the canonical helical assembly depicted by the 3D reconstruction (Bluemke et al., 1988).

The single-particle approach may be less successful when applied to a class of protein filaments whose mass-per-length and persistence lengths (with regards to bending, torsion and/or stretching) are too low to be amenable to such image analysis, given a set of optical and detection conditions. Such filaments are so deformable that any given image segment would be unlikely to satisfy both length requirements. Luckily, many filaments of biological importance do not fall under this regime. Even filamentous

*Abbreviations:* WLC, worm-like chain; Aβ, amyloid beta.

* Corresponding author at: Janelia Farm Research Campus, 19700 Helix Drive, Ashburn, VA 20147, USA.
*E-mail address:* niko@grigorieff.org (N. Grigorieff).

http://dx.doi.org/10.1016/j.jsb.2014.03.012
1047-8477/

actin (f-actin) appears to be (or can become) rigid enough when imaged in cryo-EM experiments to yield sub-nanometer resolution (Fujii et al., 2010; Galkin et al., 2012).

Fibrils formed by amyloid-β (Aβ) peptides, which are implicated in Alzheimer's disease, present an interesting intermediate case, whence sub-nanometer reconstructions have been obtained, but with some difficulty. A strong meridional reflection at ∼1/4.8 Å$^{-1}$ can be seen in averaged power spectra computed from existing micrographs of Aβ fibrils (Sachse et al., 2008), suggesting that a high degree of axial order is preserved in those specimens and that the images should therefore be of sufficient quality for higher-resolution reconstructions to be attainable. However, processing micrographs of unstained Aβ fibrils is challenging because they are essentially featureless in the axial direction at resolutions >4.8 Å and because the signal-to-noise ratio (SNR) in micrographs recorded onto film is insufficient to give reliable alignments (Sachse, 2007; Sachse et al., 2008). Many Aβ fibril morphologies resemble flat nanoscopic ribbons with slow twists on the order of 1 °/nm (Meinhardt et al., 2009), so that their projection images periodically exhibit thin (∼4–7 nm) high-contrast crossovers with projected molecular weights of ∼5–8 kDa nm$^{-2}$ and wide (∼10–20 nm) low-contrast regions with projected molecular weights of ∼2.3–3.4 kDa nm$^{-2}$ (Schmidt et al., 2009), which is comparable with f-actin (∼2.8 kDa nm$^{-2}$, width ∼6–10 nm).

Two avenues to improving the resolution of Aβ fibril structures therefore present themselves: improving the SNR of micrographs and/or improving the image analysis algorithms to make them even more robust to low SNR and filament deformations. Here, we explore the second avenue and attempt to make iterative real-space refinement of Aβ fibrils even more robust to low SNR and fibril deformations.

In previous work, knowledge about the connectivity of image segments coming from the same filaments and their geometric relationships due to helical symmetry was used as a criterion for validation of segment alignments and for the *a posteriori* selection of segments. For example, when analyzing micrographs of TMV, Sachse et al. (2007) discarded those segments for which either the assigned polarity contradicted that of other segments from the same filament or the shifts perpendicular to the helical axis were greater than ∼10 Å. Similar *a posteriori* exclusion of segments is employed by the commonly-used method developed by Egelman (2000). In our approach we tested whether this type of criterion could also be used as a *prior* during the iterative real-space processing of filament segments to improve the overall quality of their alignments. In particular, we were interested in whether it would be possible to reliably "align" filaments with high curvature and/or low contrast.

To help answer these questions, we developed Frealix, a software tool that introduces "full filament" restraints so that helical deformations can be tracked accurately using arbitrarily short linear segments, which are not treated independently from each other.

## 2. Theory

### 2.1. Frealix

Frealix is a program for the analysis of electron micrographs of helical filaments. Its inputs are micrographs, filament coordinates, estimated helical parameters and a preexisting 3D reconstruction. Its outputs are a 3D reconstruction, refined coordinates and refined helical parameters.

Internally, each filament is represented as an assembly of (rigid-body) subunits positioned along a helix which has a space curve as its axis. The space curve and helical parameters are refined iteratively by maximizing a function which compares the experimental (noisy) image of the filament to projections of the current reconstruction as predicted by its model. The scoring function also integrates restraints derived from mechanical considerations when modeling filaments.

Below, we describe the parametrization of our model for helical filaments (Section 2.2), the function used to "score" sets of parameter values given a model and a micrograph (Section 2.3), maximization strategies we use during refinement (Section 2.4) and the 3D reconstruction protocol (Section 2.6).

### 2.2. Modeling helical filaments

The simplest model of a straight filament without distortions can be described by two parameters: the rise ($\Delta_t$) and twist ($\Delta_\varphi$) per helical subunit (Fig. 1A). If we let the helical axis coincide with the $Z$ axis and position the first asymmetric unit on the $Z = 0.0$ plane, the $Z$ position of the $i$th asymmetric unit is $Z = (i-1)\Delta_t$ and its $(X,Y)$ coordinates are obtained from the $(X,Y)$ coordinate of the first unit by $i-1$ rotations of $\Delta_\varphi$ around $Z$. All the asymmetric units lie on a continuous helix which has a characteristic pitch, the distance along its axis over which a revolution ($2\pi$) is completed. One can define $t$, the distance along the helical axis: in this simple case $t = Z$.

A more generalized description of observable filaments needs to account for their elasticity with regards to bending, torsion and stretching. To achieve this in the simplest possible way, we chose to describe the axis of a filament as a space curve $\mathbf{r}$ defined by 3 cubic spline functions $x(t)$, $y(t)$ and $z(t)$ ($t$, as before, is the arc length along the axis), which interpolate a set of $n$ waypoints defined by $(x_i,y_i,z_i)$ coordinates, where $i = 1, \ldots, n$.

At waypoint $i$, the $\theta_i$ (out-of-plane) and $\psi_i$ (in-plane) Euler angles are related to the curve's tangent vector (Fig. 1B) and thus its derivatives $x'(t_i)$, $y'(t_i)$ and $z'(t_i)$, and can be used as constraints when solving the splines:

$$\begin{aligned} x'(t_i) &= \sin\theta_i \cos\psi_i \\ y'(t_i) &= \sin\theta_i \sin\psi_i \\ z'(t_i) &= \cos\theta_i \end{aligned} \tag{1}$$

where $t_i$ is the arc length from the filament's first waypoint to waypoint $i$. Interpolating cubic splines which are thus constrained by their tangents are sometimes called Hermite splines (Knott, 2000 p. 66). Conversely, Euler angles at any point along the filament axis can be computed from the local tangent vector:

$$\begin{aligned} \psi(t) &= \tan^{-1}\left(\frac{y'(t)}{x'(t)}\right) \\ \theta(t) &= -\tan^{-1}\left(\frac{y'(t)}{z'(t)\sin\psi(t)}\right) \text{or} \\ \theta(t) &= -\tan^{-1}\left(\frac{-x'(t)}{z'(t)\cos\psi(t)}\right). \end{aligned} \tag{2}$$

Arbitrary bending deformations of the helical axis can be accurately described by these three spline functions, given a sufficient number of waypoints.

We also define two additional parameters per waypoint – rotation around the helical axis ($\varphi_i$) and helical subunit number ($h_i$) – and use natural cubic splines to interpolate values for these parameters. This allows us to define at every point along the filament the local helical parameters $\Delta_\varphi$ and $\Delta_t$, the axial twist and rise which relate a subunit to its neighbors:

$$\begin{aligned} \Delta_t(t) &= \frac{1}{h'(t)} \\ \Delta_\phi(t) &= \frac{\phi\prime(t)}{h'(t)}. \end{aligned} \tag{3}$$

Each waypoint thus contributes 7 parameters to the description of a filament: $x$, $y$, $z$, $\theta$ and $\psi$ describe the trajectory of the helical axis