Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 160 (2007) 362–374
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

# A maximum likelihood approach to two-dimensional crystals

Xiangyan Zeng [a], Henning Stahlberg [a], Nikolaus Grigorieff [b,*]

[a] Molecular & Cellular Biology, University of California at Davis, 1 Shields Avenue, Davis, CA 95616, USA
[b] Howard Hughes Medical Institute, Brandeis University-MS029, 415 South Street, Waltham, MA 02454-9110, USA

Received 4 July 2007; received in revised form 5 September 2007; accepted 14 September 2007
Available online 25 September 2007

**Abstract**

Maximum likelihood (ML) processing of transmission electron microscopy images of protein particles can produce reconstructions of superior resolution due to a reduced reference bias. We have investigated a ML processing approach to images centered on the unit cells of two-dimensional (2D) crystal images. The implemented software makes use of the predictive lattice node tracking in the MRC software, which is used to window particle stacks. These are then noise-whitened and subjected to ML processing. Resulting ML maps are translated into amplitudes and phases for further processing within the 2dx software package. Compared with ML processing for randomly oriented single particles, the required computational costs are greatly reduced as the 2D crystals restrict the parameter search space. The software was applied to images of negatively stained or frozen hydrated 2D crystals of different crystal order. We find that the ML algorithm is not free from reference bias, even though its sensitivity to noise correlation is lower than for pure cross-correlation alignment. Compared with crystallographic processing, the newly developed software yields better resolution for 2D crystal images of lower crystal quality, and it performs equally well for well-ordered crystal images.


*Keywords:* Maximum likelihood; Electron crystallography; Protein structure; Single particle; 2dx

## 1. Introduction

Electron crystallography of two-dimensional (2D) crystals is a commonly used technique to obtain high-resolution three-dimensional (3D) structures of proteins (for a recent review, see Renault et al., 2006). The technique was developed mainly by Henderson and Unwin (1975) and one of its first applications led to an atomic model of bacteriorhodopsin (Henderson et al., 1990). Many other structures have been solved at a resolution that allowed an interpretation with an atomic model. The data collected from 2D crystals come in two flavors: images and electron diffraction pattern. Electron diffraction data provide intensities of diffraction spots, much like X-ray diffraction of 3D crystals. These can be measured, and their square root gives the amplitude components of the protein structure. The images of the 2D crystals can be Fourier transformed, and both, amplitudes and phases of the calculated diffraction spots can then be measured by Fourier extraction. This is in contrast to the situation in X-ray crystallography, where phases are not directly observable. Therefore, to collect a complete data set from 2D crystals by electron crystallography, both electron diffraction and imaging are usually performed.

Images also offer another advantage over X-ray crystallography. If disorder is present in a crystal and limits the resolution in a diffraction pattern, the crystal distortions in an image of the crystal can be corrected computationally. This procedure is commonly referred to as "unbending" and significantly improves the resolution attainable by Fourier extraction from 2D crystal images. For the unbending process, a reference image containing only a small number of unit cells is generated either from a filtered version of the image itself, or by projecting an already existing 3D reference structure using the program MAKE-

* Corresponding author. Fax: +1 781 736 2419.
*E-mail address:* niko@brandeis.edu (N. Grigorieff).

1047-8477/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2007.09.013

TRAN (Kunji et al., 2000). A cross-correlation map is then calculated between the reference and the image to determine the location of each unit cell in the crystal. Using the autocorrelation function of the reference, the correlation map can be searched for peaks and unit cell locations are recorded. The MRC program QUADSEARCH (Crowther et al., 1996) performs such a search and exploits the a-priori knowledge of the approximate location of the peaks. It uses an iterative refinement process to predict the peak location in the cross-correlation map, and then searches for the actual local maximum in that map within a limited radius of the predicted location. A coordinate list of the identified lattice nodes is then generated, which is used by the MRC program CCUNBEND to correct the crystal distortions in the image by one of two methods: a better ordered crystal image is generated by either smoothly warping the image so that the unit cells fall onto a perfect lattice, or by creating a discrete montage of unit cells placed at the crystallographically determined grid points. The Fourier transform of the unbent image is then evaluated to obtain values for the amplitudes and phases of the crystal projection map in that image.

The unbending procedure described above has been applied very successfully in many cases and is in common use, but it also has some limitations. First, in its current implementation, it cannot correct for in-plane rotational disorder. Second, the use of correlation functions for alignment is prone to error when the signal is weak, due to the increased chance of mistaking a noise peak for the correlation peak produced by the signal. The signal in the correlation map depends on the size of the reference area used to find the unit cell locations. A larger reference area will produce stronger signal peaks but will also be less sensitive to short-range disorder in the crystal. To detect and remove short-range disorder, a smaller reference area containing one unit cell or even a single molecule would be desirable.

In the case of strong irregularities in a badly ordered 2D crystal, or variations among the unit cells due to sample heterogeneity, single-particle image processing can be applied to the recorded 2D (pseudo-)crystal images. The goal of the single particle processing is similar to the 2D crystal unbending procedure. In both cases alignment and averaging of individual molecules or their assemblies is done to enhance the signal. Sass et al. (1989) have combined phases from correlation averaging (CA) with amplitudes from electron diffraction, obtaining a 3.5 Å projection map of porin. Schultz et al. (1993) imaged negatively stained 2D monolayer crystals by tomographic single-axis tilt series, calculated a 3D reconstruction of the unit cell via single particle methods, and combined two molecules from the unit cell with non-crystallographic symmetry to fill the missing cone. Sherman et al. (1998) applied multivariate statistical analysis to 2D crystal images. Stahlberg et al. (1998) used single particle methods to detect and correct for the non-crystallographic orientation of the photosynthetic reaction center within well-ordered 2D crystal images of the surrounding light-harvesting-complex I. Tahara et al. (2000) could significantly improve the resolution of a Na+/K+-ATPase projection map by applying single particle processing methods to 2D crystal images and allowing for sample heterogeneity.

Here, we apply a maximum likelihood (ML) approach to the single-particle processing of 2D crystal images. The ML processing was introduced for the processing of images of non-crystalline material (single particles), and was shown to have superior performance at low signal-to-noise ratios (SNR, variance ratio of signal over noise) compared with correlation-based alignment (Sigworth, 1998). Combining the single particle processing with an ML approach can therefore lead to an improvement over the currently used unbending process. We utilize a whitening filter to make the ML method applicable to the real data which have nonwhite noise. We discuss here the application of ML to 2D crystals and compare its performance with the traditional correlation-based unbending. Furthermore, we show how the contrast transfer function (CTF) of the electron microscope can be included in the processing.

## 2. Theory

### 2.1. Maximum likelihood for 2D crystals

The application of the ML approach to single particle images has been described previously and employs the iterative expectation maximization algorithm to maximize the likelihood function (Dempster et al., 1977). Since the images we will process as single particles will be centered on unit cells excised from a 2D crystal, it is reasonable to limit the alignment of each particle to an in-plane angle and translation. Individual unit cells may also suffer from out-of-plane tilts, for example due to undulations in the crystal (cryo-crinkling, Vonck, 2000). To perform a full 3D alignment, two additional angles would have to be considered. The computational load associated with the expectation maximization algorithm would be very high in this case and some approximations have to be made (Scheres et al., 2005). However, application of the ML approach to experimental data from 2D crystals (see below), shows that very good results can also be obtained when limiting the alignment to in-plane transformations. We will, therefore, limit our discussion to this case and follow Sigworth's implementation of the maximization algorithm (Sigworth, 1998). Briefly, we assume that we have a set of $N$ images $\mathbf{X} = \{X_i;\ i = 1, \ldots, N\}$ and corresponding transformation parameters $\Phi = \{\phi_i;\ i = 1, \ldots, N\}$ describing how these images are related to the underlying structure $A$. We further assume that the noise in each image follows approximately a Gaussian distribution and is uncorrelated. This assumption will be discussed further below when considering experimentally observed data. We can write for each image $i$

$$X_i = A(-\phi_i) + \sigma R_i, \tag{1}$$