Single particle 3D reconstruction for 2D crystal images of membrane proteins

CrossMark

Sebastian Scherer [a], Marcel Arheit [a], Julia Kowal [a], Xiangyan Zeng [b], Henning Stahlberg [a,*]

[a] Center for Cellular Imaging and NanoAnalytics, Biozentrum, University Basel, 4058 Basel, Switzerland
[b] Fort Valley State University, 1005 State University Dr., Fort Valley, GA 31030, USA

ARTICLE INFO

ABSTRACT

In cases where ultra-flat cryo-preparations of well-ordered two-dimensional (2D) crystals are available, electron crystallography is a powerful method for the determination of the high-resolution structures of membrane and soluble proteins. However, crystal unbending and Fourier-filtering methods in electron crystallography three-dimensional (3D) image processing are generally limited in their performance for 2D crystals that are badly ordered or non-flat. Here we present a single particle image processing approach, which is implemented as an extension of the 2D crystallographic pipeline realized in the *2dx* software package, for the determination of high-resolution 3D structures of membrane proteins. The algorithm presented, addresses the low single-to-noise ratio (SNR) of 2D crystal images by exploiting neighborhood correlation between adjacent proteins in the 2D crystal. Compared with conventional single particle processing for randomly oriented particles, the computational costs are greatly reduced due to the crystal-induced limited search space, which allows a much finer search space compared to classical single particle processing. To reduce the considerable computational costs, our software features a hybrid parallelization scheme for multi-CPU clusters and computer with high-end graphic processing units (GPUs). We successfully apply the new refinement method to the structure of the potassium channel MloK1. The calculated 3D reconstruction shows more structural details and contains less noise than the map obtained by conventional Fourier-filtering based processing of the same 2D crystal images.

1. Introduction

The low signal-to-noise ratio (SNR) obtained when unstained proteins embedded in a layer of amorphous ice are imaged by cryo-transmission electron microscopy (cryo-EM) hinders their structural analysis. Both the localization and orientation of smaller molecules and complexes are difficult to define precisely, and the use of averaging methods to enhance the signal is correspondingly difficult. The situation can be improved for membrane proteins as they can be incorporated in a lipid bilayer and crystallized in two dimensions (2D). This strongly restricts the orientational freedom of the macromolecule in the crystal lattice. The regular arrangement lends itself to averaging techniques and, of course, to electron crystallography.

Perfect 2D crystals are flat, distortion-free and, unfortunately, almost impossible to obtain. Rather, deviations have to be recognized and corrected for, as do possible imaging effects. A series of image processing and electron crystallography software packages tackle this problem (reviewed in (Arheit et al., 2013b)), the most famous being the MRC programs (Crowther et al., 1996). These have allowed atomic models to be determined for different membrane proteins. The basic procedure used to process crystal images is comprised of six steps: (i) lattice determination, (ii) determination of the tilt geometry, (iii) lattice correction (unbending), (iv) contrast transfer function (CTF) correction, (v) data averaging for each image, and (vi) merging and lattice line fitting of the data from several images in three dimensions (3D). The *2dx* software (Gipson et al., 2007) implements a user-friendly graphical interface to the MRC programs, and adds several additional modules for automation and additional algorithms. *2dx* offers user guidance

and facilitates the project management. The conventional crystallographic image processing approach addresses locally disordered crystals by *crystal unbending*, where small image patches (typically 25 × 25 pixels) are shifted in the image plane. In the present study we show that the mathematical model behind unbending fails if crystals are disordered and tilted, or are non-flat. In order to overcome these limitations, we introduce a new refinement process that resolves local crystallographic disorder in 3D.

Cryo-electron microscopy (*cryo-EM*) *single particle analysis* (Frank, 1975) is a technique used to determine the structure of proteins without any crystallization. Randomly oriented single particles are imaged under the electron microscope leading to different views of the protein under investigation. The orientation of the recorded projections is initially unknown. The approximate orientation of each particle is determined by comparing the particles with a set of trial projections obtained from an initial model for instance obtained by *random-conical tilt reconstruction* (Radermacher, 1988). Back-projecting all particles into a 3D volume based on their updated orientations leads to an improved 3D model. This iterative process is repeated until the 3D model no longer changes. A broad variety of single particle reconstruction software packages is available today, for instance: EMAN (Ludtke et al., 1999), BSoft (Heymann, 2001), Xmipp (Sorzano et al., 2004), Sparx (Hohn et al., 2007), Frealign (Grigorieff, 2007) Relion (Scheres, 2012), or Simple (Elmlund and Elmlund, 2012).

Here we present a single particle-based reconstruction method for 2D crystals that is implemented as an optional extension of the *2dx* software. The approach applies the principle of single particle processing to 2D crystals, exploiting neighborhood correlation between adjacent proteins in the crystal in order to overcome the limitations originating from the low SNR. High performance computing hardware is used to handle the computational costs of the new procedure. The major differences of the new method to classical single particle analysis are the new local averaging step, the significantly finer angular search space sampled during the refinement and a new crystal-based particle selection procedure validating the alignment parameters of each particle.

2. Approach

In the MRC programs, the mathematical model behind local 2D correction of crystal disorders by in-plane shifting of small crystalline patches, does not cover all the possible crystal imperfections present in a 2D crystal. 3D rotational disorder is not addressed at all. The simplification of the unbending approach is partially valid for non-tilted specimens but not for tilted specimens (Fig. 1). 3D disorder is parameterized by three Euler angles and two in-plane translational parameters, and must be considered for each protein of a 2D crystal individually to fully exploit the data and achieve reconstructions with the best possible resolution.

Below we introduce a single particle-based reconstruction method for 2D crystals that refines the orientation locally for each particle, while using the correlations of neighboring particles in the crystal to overcome limitations imposed by the low SNR of cryo-EM images. The new processing approach (Fig. 2) is a refinement of the conventional crystallographic image processing pipeline described by (Arheit et al., 2013a,b) (Fig. 2A), and is implemented as an option in the open source *2dx* (version 3.5.0) software. This new procedure can be applied to any previously recorded 3D cryo-electron crystallography dataset.

The initial stages of the classical approach (Fig. 2A) include determination of image defocus and astigmatism with the program *CTFFIND3* (Mindell and Grigorieff, 2003), lattice definition, and correction for translational disorder in the crystal, i.e., image unbending. The latter involves cross-correlation of the raw image with an iteratively improved reference image to define the position of individual unit-cells using the MRC program *Quadserch*, and translational adjustment of small crystal patches using the MRC program *CCUnbend*. Finally, the structural data from all unit cells of one crystal image are combined into one resulting unit cell image at a much higher SNR. The ensemble of data from multiple crystals at different specimen tilts is merged into one 3D reconstruction (MRC programs *Origtilt, Latline, and others*).

Our new refinement procedure (Fig. 2B and C) uses the results obtained by the classical method as a starting point, i.e., unit cell positions, CTF parameters of each image and the final 3D model,
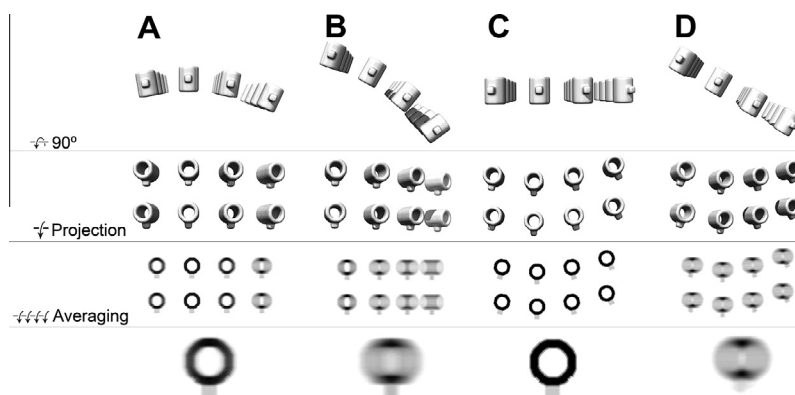

Fig.1. Effect of sample flatness and local disorder on resolution in crystallographic processing. Artificial crystals composed of multiple copies of a tube with a protrusion are considered. From top to bottom: perspective side-view of the crystal; top-view of the crystal (only two protein rows shown for the sake of clarity); the orthographic projection recorded by a transmission electron microscope; the outcome of an optimal crystal unbending procedure as could be achieved by the classical crystallographic processing. (A) Nominally non-tilted, yet locally bent 2D crystal with perfectly in-plane aligned particles. Although the local disorder cannot be retrieved perfectly by unbending the crystal, the merged projection map (bottom row) resembles the true projection of the structure with little resolution loss. (B) The crystal in (A) tilted by 30°. Due to the membrane curvature, the regularity of the crystal is lost in the recorded projection. As a result, high-resolution spots perpendicular to the tilt axis in the Fourier transform of the crystal disappear, which strongly limits the resolution of the projection map (bottom row). Thus the presence of ultra-flat preparations is required when imaging tilted 2D crystals and processing them with the classical lattice unbending approach. (C) Non-tilted perfectly flat crystal with rotational disorder. This is the situation in which the classical unbending procedure performs optimally. Note that even the rotational disorder can be corrected by using an optimal patch size in the unbending step. (D) The perfectly flat crystal in (C) is tilted by 30°. Here, the classical unbending-based processing cannot deal with this situation as the projection image contains multiple different views of the proteins in the crystal. The classical processing fails to align the difference projections from slightly different view points. Taken together, the unbending-based processing performs well on non-tilted crystals, but encounters serious limitations for tilted crystals, even if they are perfectly flat but rotationally disordered.
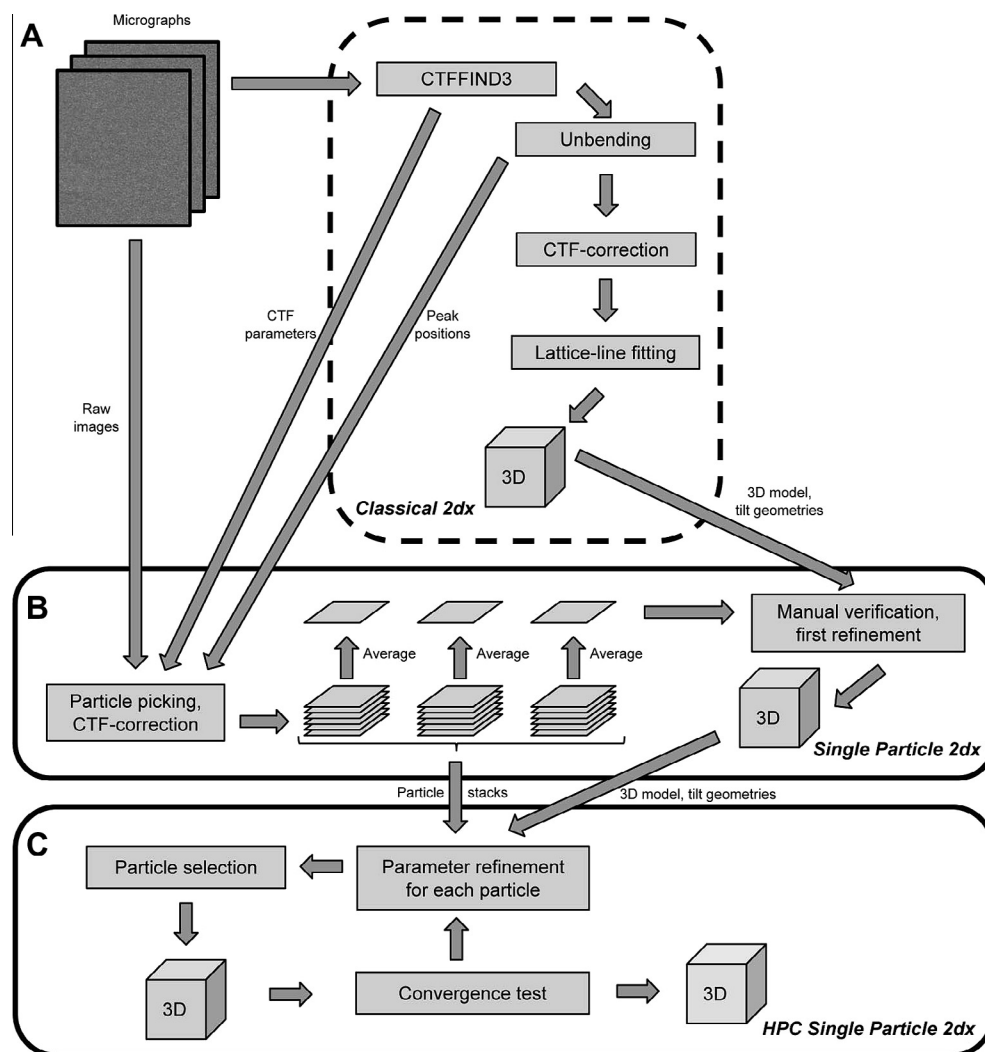
Fig.2. Single particle refinement for 2D crystals. The presented approach refines classically processed cryo-electron crystallographic datasets. (A) All micrographs are processed and merged classically with *2dx* (Gipson et al., 2007). The unit-cell positions determined while unbending the crystals, the defocus values at the center of each micrograph and the final MRC-reconstruction are then forwarded to the new single particle module of *2dx*. (B) The particles are cut out from the original micrographs based on the unit cell positions determined previously with the MRC program *Quadserch* (Crowther et al., 1996). Dependent on their position, the tilt geometry and the central defocus parameters of the crystal from which the particles were originally picked, they are CTF-corrected and finally stored in per-crystal particle stacks. In order to verify and correct the initial tilt geometry of the particles, one averaged particle is calculated for each crystal. The orientation of this average with a much higher SNR is refined either manually by means of a new interactive tool or automatically by an iterative single particle reconstruction procedure. (C) Refining the local tilt geometry variations, is computationally demanding. Therefore, the particle stacks, the refined initial model and the configuration file are uploaded to a high-performance computing cluster.

to improve the resolution of the 3D reconstruction. We propose a new interactive and a new automatic procedure to refine the classically determined tilt geometry of entire 2D crystals (Fig. 2B). Subsequently the 3D orientation of each protein of each crystal is iteratively optimized using our new local refinement tool (Fig. 2C), taking the correlation between neighboring particles into account.

2.1. Particle localization and picking from 2D crystals

The picking of single particles within *2dx* as originally introduced by Zeng et al. (2007) was expanded by an additional step that reduces the number of false positive particles in the background based on estimates of the local particle density of the crystalline structure.

The new particle-picking tool of the *2dx* single-particle processing module (Fig. 2B), uses the unit-cell positions determined by the MRC program *Quadserch* to extract patches centered on the cross-correlation peaks in the cross-correlation profile generated by the

2dx-script *UnbendII*, which indicates the unit-cell position on the raw micrograph. The screw axis symmetry present in a larger number of 2D membrane protein crystals requires inverting the initial tiltgeometry for all protein particles picked from the oppositely facing population of particles. One example is shown in Supplementary Fig. 1. In the current implementation, in case of a screw axis present in the crystal symmetry, particles picked from the two differently oriented conformations are processed independently, and thus no correlation between these two subsets is generated during the processing. Due to this initial tilt geometry inversion of one sub-population of particles in such a case, the remaining symmetry enforcement such as 4-fold symmetrization during the reconstruction step is straight-forward. The advantage of using *Quadserch* via *2dx_image* is that *2dx* allows interactive verification and refinement of the unbending parameters (e.g., reference position and Fourier diffraction spots contributing to the reference) in order to improve the reliability of the cross-correlation profile.

As some false positives in non-crystalline regions are nevertheless unavoidable, the new particle-picking tool uses a particle

density-based selection method to ensure that erroneously selected particles corresponding to false correlation peaks in the background are rejected. In a first step the center of mass (CM) of the crystal is calculated. The n ($n \approx 40$) particles closest to the CM are considered as the *central region*, and a density is calculated for this region based on the number of particles per surface. A region is then defined around each particle based on the n nearest neighbors of the particle, and the unit density of the region is calculated (local density). Particles with a local density significantly deviating from the density of the densely packed central region around the CM of the crystals are rejected, as they most likely correspond to an erroneous noise induced peak rather than a particle induced peak.

2.2. Local CTF-correction

Particles that passed the density-based selection procedure are still affected by unfavorable CTF effects, which depend on (i) the sample tilt of the entire crystal and (ii) the specific position of each particle in the image. Using basic trigonometry, the local defocus and therefore its associated CTF parameters for each particle are calculated from the CTF values at the center of the image determined by *CTFFIND3* (Mindell and Grigorieff, 2003) and manually verified within *2dx*. The particle defocus values were not further refined during the following processing. A detailed explanation of the local CTF correction can be found in the Supplementary material of this article. Finally, the CTF is corrected by phase flipping and the corrected particles are stored in per-crystal particle stacks.

The particle position in the *z*-direction changes if the orientation of regions of the crystal is changed because the 2D crystal is not flat (Fig. 1A and B). However defocus values are not refined in the current implementation. Following (Zhang and Zhou, 2011) the resolution limit imposed by a defocus determination that is inaccurate by 124 nm at 300 kV acceleration voltage is at 7 Å. This resolution limit is beyond the resolution of the test dataset used in the results section. Further, by basic trigonometry and the values reported by Zhang and Zhou, it follows that changing the global tilt angle of an entire non-tilted 1 μm sized crystal by up to 14° during the refinement does not limit the resolution. In the here processed dataset, we have not observed an angular change in this range. We therefore chose to not refine the defocus values of individual particles in our current implementation.

2.3. Initial tilt geometry: determination, verification and refinement

The tilt geometries of entire crystals determined by classical crystallographic processing may still be slightly off (up to 5°). Thus, we have implemented a new user interface (Fig. 3) that allows manual verification and correction of the tilt geometries and in-plane particle positions of an entire dataset, and the use of an automatic routine that optimizes the crystal orientations iteratively based on the results obtained by the classical method.

For this verification, all particles of one crystal are averaged in real-space in order to generate one average with a higher SNR for each crystal, i.e., a 'per-crystal' average. The orientations of the crystals serve as the initial orientation of the averages. The per-crystal averages are then compared with the images obtained by projecting the final 3D model generated using *2dx* in the respective nominal directions. The new user-interface (Fig. 3) enables interactive refinement of the initial orientations determined by conventional processing. Particles from highly inconsistent crystals or from crystals with insufficient structural information can then be removed from the dataset.

Instead of manually refining the dataset, one can run an iterative self-consistent single particle refinement (Frank, 1975) of the per-crystal averages. A set of trial projections of the initial volume around the nominal orientation of each per-crystal average is calculated, and correlated with the per-crystal averages. The orientation of the projection with the highest correlation value is then assigned to the average and, thus, to all particles contributing to this average. Subsequently, the 3D reconstruction is updated by back-projecting the averages into a 3D volume. The described iterative procedure is repeated until the refinement converges, i.e., the orientation parameters of the particles change less than a given threshold. In practice we measured the average angular change per average and stopped the refinement once this average change was below a user-defined threshold of 0.2°. Note that the proposed refinement of the initial orientations resembles a classical real-space single particle refinement of the per-crystal averaged particles. The computational efficiency is significantly improved in case of an approximately flat 2D crystal, which then allows the angular search space to be restricted to ±5°.

The reliability of the in-plane alignment of the particles has a large effect on the quality of the final reconstruction. A better 3D model will lead to a more precise in-plane alignment of the individual particles. After the first round of initial model refinement,
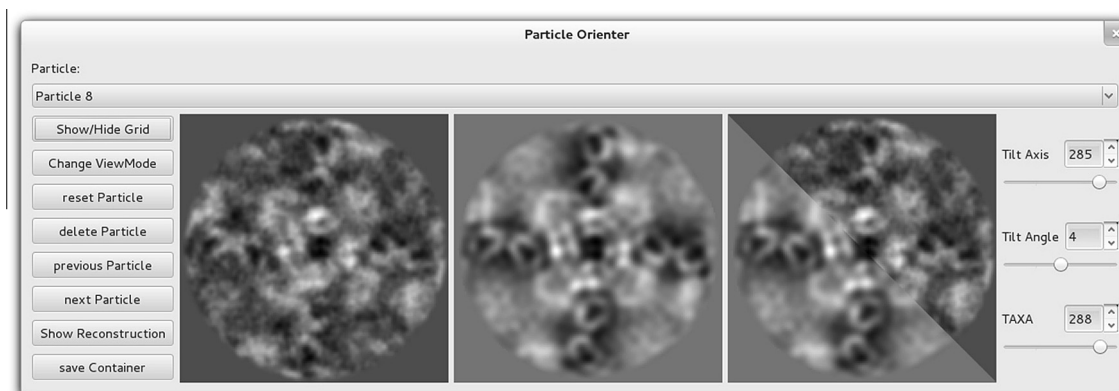

Fig.3. Interactive graphical user interface to verify and refine the particle orientation. The goal is to find the three Euler angles that lead to the projection of the 3D reconstruction (central window) that matches best the per-crystal average (left window). Changing the sliders on the right immediately updates the projection shown in the central window. For a more convenient comparison between projection and average, the right window is used to toggle between the two view modes, which are shown side by side. An updated starting model can be achieved by back-projecting the averages along their refined orientations. This refined model can be inspected in 3D by means of UCSF Chimera (Pettersen et al., 2004) launched directly from the new interface.

our approach optimizes the two in-plane shift parameters of all individual particles. In order to do this, each particle image is cross-correlated with the projection of the initial 3D model into its nominal direction. The vector from the center of the particle to the maximal cross-correlation value is then remembered as an additional shift value for the particle. Note that the described shift refinement resembles synthetic unbending in *2dx*, where projections of the 3D model are used as a reference for the translational unit-cell position correction (Kunji et al., 2000). More reliably in-plane aligned particles produce better per-image averages, which further improves the initial tilt geometry determination.

Dependent on the specific dataset, it can be difficult to refine the initial tilt geometries uniquely based on averages showing one isolated protein. Using image patches that show the central protein surrounded by its crystalline neighbors, generally makes the orientation determination more reliable as it increases the amount of information that can be used for the refinement. In order to stay computationally efficient, we bin the extracted patches once. After the global tilt geometries of all crystals have been optimized based on the binned patches, the refined orientations are applied to the unbinned particle stacks picked from the same dataset.

2.4. Locally averaged single particle orientation parameter refinement

The major aim of the present study is to resolve the locally varying tilt geometry caused by smoothly bent membrane protein crystals. This is achieved by means of an adapted single particle algorithm. The new procedure (i) overcomes the limitations caused by the low SNR and (ii) reduces the computational cost significantly by exploiting the fact that all particles originate from ordered 2D crystals.

The aim of conventional single particle processing (Frank, 1975) is to find the optimal set of alignment parameters $\left(\hat{\theta}, \hat{\psi}, \hat{\phi}, \hat{s}_x, \hat{s}_y\right)_i$, for each particle that maximizes the following expression:

$$\left(\hat{\theta}, \hat{\psi}, \hat{\phi}, \hat{s}_x, \hat{s}_y\right)_i = \text{argmax}_{(\theta, \psi, \phi, s_x, s_y) \in \mathcal{T}} \{sim(P_i, \mathcal{P}(A, \theta, \psi, \phi))\} \qquad (1)$$

where $sim(\cdot, \cdot)$ measures the similarity (i.e., cross-correlation) between the particle P_i and a reference projection $\mathcal{P}(A, \theta, \psi, \phi)$ that depends on the 3D reference A and the three Euler angles (θ, ψ, ϕ) describing the orientation of the particle in 3D. Once the best matching orientation has been determined for each particle, a more accurate 3D structure is obtained from them using back-projection along the updated orientations. The updated structure serves as a reference for the subsequent iteration step. This self-consistent optimization is continued until convergence is reached, i.e., until (i) angular changes are smaller than a certain threshold or (ii) the structure A changes less than a given threshold from one iteration to the next one. Similar to the initial model refinement the iteration is considered to be converged, if the average angular change over all particles is below 0.2° from one iteration to the next. In practice, the optimization of Eq. (1) is carried out over a finite set of trial projection $\mathcal{T}$ uniformly covering the entire angular search space. Note that the optimization also includes an in-plane shift (s_x, s_y) alignment of the particle P_i. As this optimization can be done efficiently by determining the offset between the cross-correlation peak and the center of particle P_i, the shift parameters are not separately mentioned in the optimization target function in Eq. (1).

Both, the significantly lower SNR and the correlation between neighboring particles, are essential differences between classical single particle approaches and our new application. The application of a single particle refinement-based version of Eq. (1) to individual particles picked from cryo-EM images of unstained 2D membrane protein crystals was notsuccessful in our hands, likely due to the too low SNR induced by the small electron dose used

on imaging, and by disturbances from overlapping projections of neighboring particles in images of tilted 2D crystals (data not shown). Conventional single particle processing of randomly oriented particles can increase the SNR by refining the orientations of class averages instead of the individual particles. Similarly, here we exploit the correlation between neighboring particles in order to increase the SNR. Orientations of two neighboring particles will be very similar, while those of distant particles may deviate due to lattice bends and the membrane curvature that our approach aims to resolve (Fig. 1). Accordingly, the orientations of close and distant particles should not be treated as if they are directly coupled. Instead of using individual particles for the alignment step we use locally averaged, radially Gaussian weighted particle averages to determine the angular orientation of each particle. Therefore, we replace the objective function of the underlying optimization problem Eq. (1) by the following expression:

$$\left(\hat{\theta}, \hat{\psi}, \hat{\phi}, \hat{s}_x, \hat{s}_y\right)_i = \text{argmax}_{(\theta, \psi, \phi, s_x, s_y) \in \mathcal{T}_i} \left\{ sim\left(P_i + \sum_{j \in \langle \mathcal{N}_i \rangle} \lambda_j P_j, \mathcal{P}(A, \theta, \psi, \phi)\right) \right\} \qquad (2)$$

where the similarity measure, $sim(\cdot, \cdot)$ is the same as in Eq. (1). The major novelty is that the similarity is now calculated between the projection and the weighted sum of the central particle P_i and particles P_j belonging to the neighborhood $\langle \mathcal{N}_i \rangle$ associated with particle P_i. Note that the projection trial space $\mathcal{T}_i$ depends on the initial orientation of the crystal and thus has an additional subscript index i. Typically, we use a dozen neighbors and a sampling cone of a few degrees. In this way, the computational complexity can be massively reduced and a finer search space can be sampled. The weights λ_j are following a Gaussian distribution, where the particles closer to the central particle contribute more to the average:

$$\lambda_j = \exp\left(-\frac{1}{2}\left(\frac{R_j}{\sigma}\right)^2\right) \qquad (3)$$

where R_j equals the distance between the central particle P_i and the adjacent particle P_j. The parameter σ is used to tune the weight. Our current implementation adjusts σ so that the particle furthest from the center contributes ten times less to the average than the central particle. Once the locally summed particles are aligned, the original central particles are back-projected into a 3D volume in order to generate a new reference for the following iteration step. The procedure is repeated until convergence is reached.

2.5. On the fly likelihood-based particle selection

Due to crystal defects or fallacious noise peaks, some particles of each crystal are incorrectly aligned and thus should not be used for the reconstruction step. Therefore, we introduce a particle selection method that uses the iteratively refined alignment parameters of each particle together with the neighborhood correlation stemming from the crystalline arrangement to judge the plausibility of each particle image.

Based on the alignment we define a 'fingerprint vector' for particle P_i:

$$\vec{x}_i = (\theta_i, \psi_i, \phi_i, (s_x)_i, (s_y)_i, cc_i)^T \qquad (4)$$

consisting of the three Euler angles (θ_i, ψ_i, ϕ_i), the two shift parameters $(s_x)_i, (s_y)_i$ and the corresponding highest cross-correlation value cc_i determined during the refinement process. As suggested by Supplementary Fig. 2 the components of the particle fingerprint vectors follow a 6-fold multivariate normal distribution. Based on this assumption the log-likelihood function $logf(\vec{x}_i|\mu_{c(i)}, \sum_{c(i)})$, which describes the likelihood (or 'probability') of finding a single observation $\vec{x}_i$ given the estimated mean vector $\mu_{c(i)}$ and covariance

matrix $\Sigma_{c(i)} (with \Sigma_{mn} = (x_m - \mu_m)(x_n - \mu_n))$, is given by Anderson and Olkin (1985):

$$\log f\left(\vec{x}_i \mid \mu_{c(i)}, \Sigma_{c(i)}\right) = -\frac{1}{2} \log(|\Sigma_{c(i)}|)$$
$$-\frac{1}{2}\left(\vec{x}_i - \mu_{c(i)}\right)^T \Sigma_{c(i)}^{-1}\left(\vec{x}_i - \mu_{c(i)}\right) \qquad (5)$$

Here, $c(i)$ indicates that the mean vector and the covariance matrix depend on the crystal from which the particle P_i was originally picked. $|\Sigma_{c(i)}|$ represents the determinant of the covariance matrix. Note that a normalized particle finger print vector is used for the likelihood estimation. The mean vector and covariance matrix are estimated separately for each crystal. Particles with a log-likelihood value below a certain threshold are not used for the reconstruction step in order to improve the quality of the structure.

While the parameter refinement process optimizes different parameters serially, the particle fingerprint vector evaluation considers all parameters associated with the alignment simultaneously. This is important as it allows the correlations between the alignment parameters to be taken into account. This is not possible when the individual alignment parameters are verified sequentially, in which case the sequential thresholding and assessment will most likely fail if, e.g., the in-plane orientation worked out fine but the angular assignment failed. Our new method conceptually outperforms approaches that serially judge the plausibility of individual alignment variables and do not account for their correlations.

3. General implementation details

3.1. Software design and 2dx integration

The presented single particle extension for *2dx* has been implemented in version 3.5.0 of the open source software package, which is available for download at www.2dx.org. Following *2dx*'s philosophy, we have extended the set of available image processing scripts by an additional script section dedicated to single particle processing (Supplementary Fig. 3). These new single particle scripts are used to launch the different image processing programs of the single particle backend, which is implemented in c++. The new manual angular refinement and verification user interface (Fig. 3) is natively integrated into *2dx*'s frontend realized with Qt4 (www.qt-project.org). Pieces of code of the new single particle library use functions implemented in *EMAN2* (Tang et al., 2007), i.e., projection, back-projection and symmetry operations. Therefore, the new executables are linked with the EMAN2 backend library *libEM2.a*. The new package makes intensive use of *libboost*, a collection of high quality c++ utility libraries (www.boost.org), for low-level subroutines such as parsing files or multidimensional or sophisticated data structures.

3.2. High performance computing

Automatic image acquisition procedures, such as *Leginon* (Carragher et al., 2000), in combination with the latest generation of direct electron detectors (Bai et al., 2013; Campbell et al., 2012; Li et al., 2013) that can record movies of dose-fractionated image series instead of single images, will increase the size of future datasets tremendously. In order to process these future datasets in a reasonable amount of time, supporting high-performance computing hardware is indispensable for our new software.

Due to the large number of proteins present in each imaged crystal (up to 3000 particles/crystal, 131 Å × 131 Å unit-cell size, 2 proteins/unit-cell, up to 3 μm crystal side-length) the memory consumption of the per-protein refinement step is in most real world datasets higher than the amount of available memory on a regular desktop computer. Thus we perform the conventional MRC-based processing, particle picking and initial model refinement on a local machine, while running the intensive per-protein refinement on a remote high performance cluster computer that should have significantly more memory and computing power. All desktop-based pre-processing steps (Section 2.1–3) are accelerated by means of a shared memory parallelization scheme based on OpenMP (Dagum and Menon, 1998) and follow a highly adapted data-flow mechanism that minimizes the memory requirements by sequentially loading the particle stacks to fulfill the memory boundaries of traditional workstations. Hence the initial pre-processing can be performed on powerful workstations within a reasonable time.

Today's super computers consist of a large number of multi-core processors that are connected by a hierarchical high-bandwidth network. In order to use these machines as efficiently as possible, the per-protein local refinement tool features a hybrid parallelization approach (Jost et al., 2003). Consequently, the set of particles to be refined is distributed equally over a large number of multi-core processors (here called compute nodes) that communicate through the message-passing interface (MPI). Each compute node additionally stores a copy of the actual 3D reconstruction,
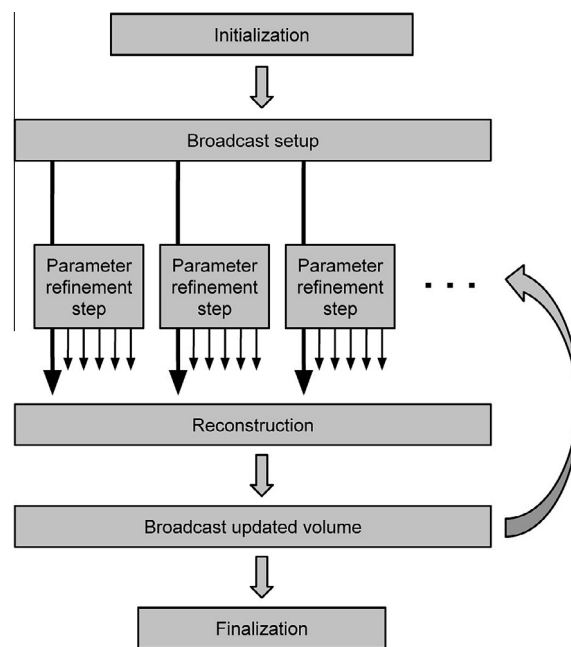

Fig.4. Hybrid parallelization approach. The application is launched on a large number of multi-core compute nodes connected by a high-bandwidth network. Each of the node runs only one MPI rank on one of its cores and keeps the rest of the cores initially idle. In order to setup the refinement, one of the MPI nodes (master node) decides on the optimal particle distribution over all available nodes and broadcasts the load-balancing scheme to all MPI ranks. Each node of the cluster now loads the required particles from the storage system attached to the high performance cluster. A projection gallery is calculated locally on each node based on the orientations of the particles assigned to the node. As there is only one MPI job running per multi-core node all of the per-node computational tasks can be performed in parallel by means of a shared memory parallelism scheme. After the neighborhood-dependent local orientation refinement, all particles stored on a node are back-projected into a 3D volume on each node. In order to update the global reconstruction, all local reconstructions are sent to the master MPI rank, which assembles the global reconstruction. This updated 3D model is subsequently broadcast to all nodes in the system and serves as reference for the next iteration step. Once convergence of the refinement procedure is reached, the optimal orientations are stored to the file system and some statistical parameters are stored for a later quality analysis of the refinement, e.g., the cross-correlation values or assigned angles.

which serves as reference for the alignment described in Section 2.4. The parallelization approach is visualized in Fig. 4.

As particles originating from the same crystal share a considerable number of the required trial projections, pre-calculating and storing the projections decreases the computational complexity of the later cross-correlation maximization significantly compared to an on-the-fly projection calculation scheme. Thus, each compute node uses the orientation of the particles and the current approximation of the reconstruction to pre-calculate the projection gallery required for the subsequent alignment step (Section 2.4). After disabling inconsistent particles based on the procedure detailed in Section 2.5 all particles stored on a compute node are back-projected into a local 3D volume. As the updated global 3D structure depends on all particles spread over a large number of nodes, all local reconstructions are gathered on the master MPI-node and finally merged by weighted Fourier averaging into one volume. Our approach uses Fourier-placing on each node and synchronizes the cumulated interpolation weights towards the master MPI-node to account for non-uniformly sampled voxels. The central slice theorem, storing the Fourier space representation of all objects, and the absence of any per-particle weights, allow us to distribute the reconstruction step over multiple compute nodes. The updated structure is used as a new reference model for the next alignment step and thus is broadcasted to all nodes of the cluster.

The projection and cross-correlation steps per node are accelerated by an OpenMP shared memory parallelization scheme. To achieve this we encapsulated the EMAN2 projection method so that multiple projections can be calculated simultaneously in a thread-safe way, i.e., multi threads can execute a function simultaneously without influencing each other. All together, our hybrid parallelization approach maximizes the performance on homogeneous clusters by using an optimized parallelization paradigm on each level of parallelism.

3.3. GPGPU accelerated projection method

The performance bottleneck of the initial model refinement (Section 2.3) performed on a desktop computer is the calculation of the projection gallery rather than the alignment of the per-crystal averaged particles. We successfully accelerated the projection method by using a general-purpose graphic processing unit (GPGPU) running the c++ language extension CUDA provided by Nvidia. The use of graphic processors significantly accelerates image processing algorithms related to cryo-EM (Castano-Diez et al., 2008). Our implementation makes efficient use of the texture memory and can be used simultaneously by multiple host threads in a thread-safe manner. By means of the hypercube technology introduced in CUDA-5.0, multiple device kernels can run simultaneously on the same graphic processing unit (GPU) by utilizing
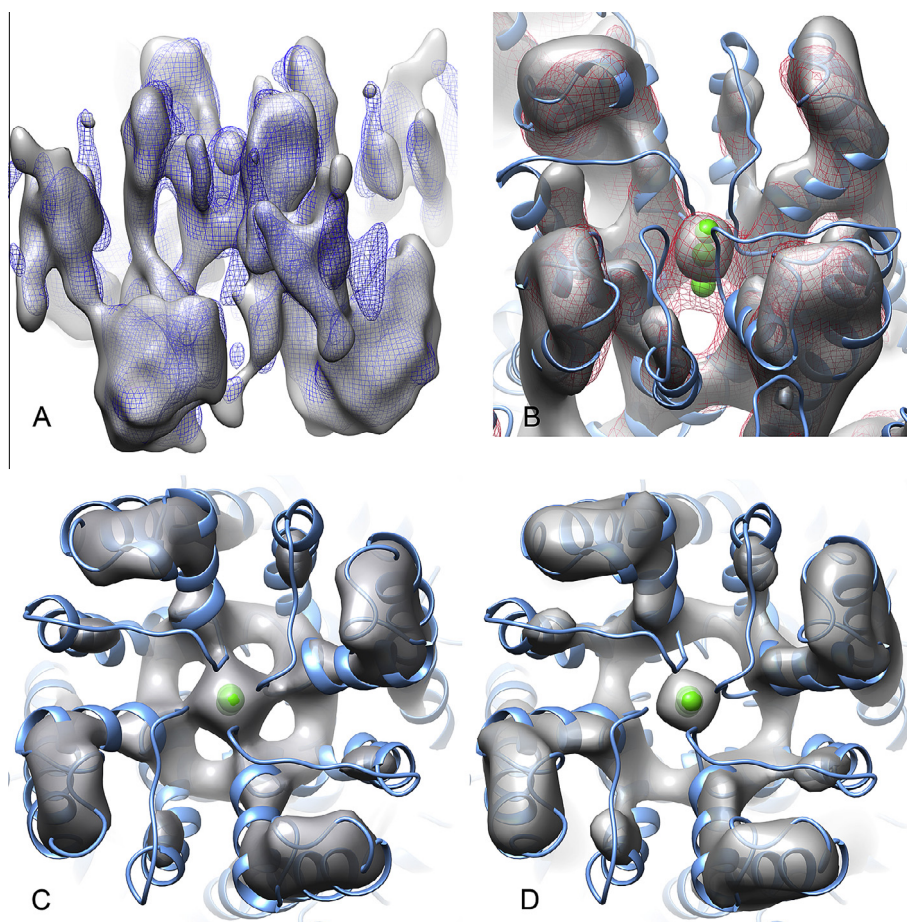

Fig.5. Results obtained with the 3D local tilt geometry refinement procedure on a cryo-EM data set. (A) Reconstruction obtained by classical unbending processing within *2dx* (mesh, blue) compared to the locally refined structure obtained using our new algorithm (solid, grey). (B) Closer look at the pore region of the membrane protein where the quality difference is largest. The atomic model shown was obtained by flexible molecular dynamic fitting using the cryo-EM map generated by the classical unbending approach as an additional constraint. The locally refined structure (solid, grey) resolves the helices much better than the conventionally generated map (mesh, red). (C and D) Comparison of the top-views of the pore region obtained by unbending (C) and the new procedure (D). The iso-contour levels were set so that the long perpendicular helices of the pore region have the same dimensions in all maps shown.

multiple execution streams (implemented in hardware), which once again boosts the projection calculation. Our software additionally supports multi-GPU systems in order to distribute the work over several hardware accelerators. GPU-support on high-performance computing clusters will be realized in a future version.

4. Results and discussion

4.1. MloK1 – an experimental test dataset

We recently determined two different confirmations of a ligand-modulated bacterial potassium channel MloK1 (Kowal et al., 2013). This tetrameric ion channel features cyclic nucleotide binding domains (CNBDs) and putative voltage sensor domains (VSDs) on each of its monomers. In the context of the study presented here we used images of the ligand-bound MloK1 2D crystals to benchmark the new refinement procedure. The test dataset consisted of 73 low dose images recorded on a Philips CM200FEG cryo-transmission electron microscope, operated at an acceleration voltage of 200 kV and a nominal magnification of 50,000×, using a defocus range of 500–2500 nm. The crystal samples were vitrified in the holes of holey carbon film. Images were recorded on photographic films (Kodak SO163), which were digitalized using a drum scanner with a step size of 5 μm, resulting in a pixel size of 1 Å at the specimen level and in a crystallographic unit-cell size of 131 Å. All crystal images were processed and merged into one 3D structure (Fig. 5A, mesh, blue) with *2dx*, yielding a final resolution at 7 Å resolution in *xy*-direction and 12 Å orthogonal to the imaging plane (Kowal et al., 2013).

4.2. Experimental procedure: single particle refinement of a test dataset

In a first step we extracted 1000 patches (300 × 300 pixels) from each crystal. The patches were CTF-corrected, binned by a factor two (resulting in patches of 150 × 150 pixels) and averaged in order to generate per-crystal average patches. Each patch contained the central protein tetramer and four neighboring tetrameric proteins. The presence of these neighbors simplified the orientation determination. The orientation of the per-crystal averages was manually refined against the initial 3D structure by means of the new interactive particle alignment tool (Fig. 3).

Based on the positions of the unit-cells determined while unbending the crystals with *2dx*, we extracted 244,604 un-binned particle images (150 × 150 pixels) from the micrographs. Each particle image was masked so that it contained one unique central protein tetramer. Subsequently we removed 34,066 false positive particles selected in the background, by means of the new density based particle selection tool (Section 2.1). All subsequent refinement steps were performed with the remaining 210,538 individual membrane proteins from 73 crystals. The automatic iterative initial orientation refinement described in Section 2.3 was executed within a ±5° range until the model did not significantly change further. All of the automatically determined orientations were verified manually before any further refinement was performed. After the first initial model refinement, we optimized the in-plane position of each particle within the 21 × 21 pixel window by allowing individual particles to shift by maximally 10 pixels in any direction and recalculated the per-crystal particle averages. This was done by applying the local averaging scheme described in Section 2.4 for the in-plane alignment, i.e., instead of individual noisy particles the weighted sums of nearby particles were used to determine the optimal in-plane correction. Although we propose to use the reconstruction obtained by classical processing as starting model, our software can deal with any starting model provided as MRC-format file. Using a large in-plane shift search space allows the procedure to correct for translational offsets between experimental data and the starting model obtained by another procedure. After the shift correction another round of iterative initial model refinement was applied to the updated per-crystal averages.

The local tilt geometry refinement for each particle was performed in two consecutive refinement rounds. In the first phase of the coarse refinement, the local orientations were fine-tuned within a ±12° cone in which the angular distance between the individual trial direction was maximally 3°. In the second phase, the in-plane rotation (orientation of the crystal on the grid) of each particle was optimized over a ±3° range in 1° steps. For this initial coarse refinement, information in the resolution range from 12–75 Å was used for the cross-correlation calculation. Averaging the central particle with the 16 nearest neighboring particles increased the SNR enough for a consistent orientation determination to be made, i.e., the orientations determined for neighboring regions of the crystal were in a reasonable range (a few degrees). In our hands the refinement characterized by the resulting parameter set showed a remarkably fast convergence. After the third iteration the 3D structure did not visibly change anymore, consequently we limited the number of coarse iteration steps to three. The subsequent fine-grained second refinement was based on the optimal alignment parameters found at the coarser refinement level. In order to resolve finer local variations, we limited the trial projection space to a ±3° cone with 1° separated trial directions. The in-plane rotation was optimized over ±1.5° in 0.5° steps. By reducing the number of neighboring particles contributing to the average used for the alignment to 8, we allowed larger local fluctuations. The second refinement phase used information in the resolution range 10–50 Å and the resolution of the iteratively refined structure was limited to 10 Å.

In order to improve the convergence and stability of the algorithm, we linearly combine the $(n-1)$ reconstruction with the new back-projection based on the refined parameters determined in step (n) to obtain the new reconstruction used as the reference in step $(n+1)$. The most successful refinements were achieved when 30% of the $(n-1)$ structure was retained. The likelihood-based particle selection (Section 2.5) was used to select the 85% most consistent particles per crystal for the back-projection step. Convergence of the second iteration procedure was observed after 5 rounds. A comparison of the structure obtained by classical processing and the new single particle-based reconstruction is shown in Fig. 5A. The probability distributions of the refined parameters of the crystals (three angles and two shift values) show approximately Gaussian distributions around the nominal parameters of the crystals (Supplementary Fig. 2A–E), suggesting that our method resolves the true variations in membrane flatness of the 2D crystals in 3D.

4.3. Resolution measure, prevention of overfitting

From the Fourier shell correlation (FSC) shown in Fig. 6 we conclude (0.5 criterion) that the final reconstruction obtained by means of the presented algorithm is 8.3 Å. This one-dimensional resolution limit is an average value for the anisotropic resolution of the reconstruction due to the missing cone and the beam induced resolution loss for highly tilted samples. A direct comparison between the resolution achieved by the new tool and the resolution of the conventionally generated reconstruction is not possible as there is no universal definition of resolution valid for both crystallography and single particle reconstructions. In electron crystallography the resolution is deduced from the frequency at which the phase residual drops below a commonly accepted value (Unger, 2013), whereas the resolution of a single particle
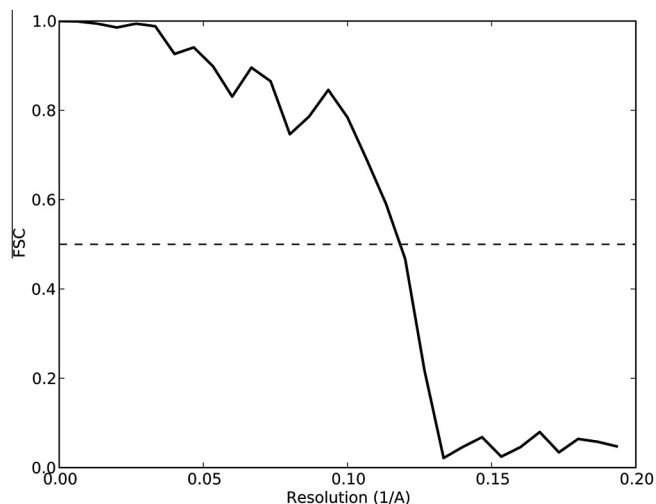
Fig.6. Fourier shell correlation of the final single particle based refined 3D structure indicating a resolution of 8.3 Å (0.5 FSC-criterion, pixel size 1 Å).

reconstruction is commonly calculated by splitting the dataset into two halves from which two independent reconstructions are generated. The investigation of the frequency-dependent decay of the correlation between these two reconstructions allows us to report a unique number as the resolution of the reconstruction (Penczek, 2010). Instead of trying to conceal the two different resolution determination methods, we here estimate the improvement gained by the new refinement procedure by visual comparison of the obtained reconstructions (Fig. 5).

A major issue of commonly used single particle reconstruction approaches is their tendency to overfit the data due to the alignment of noise instead of the true structure (Stewart and Grigorieff, 2004). The choice of appropriate low-pass filter parameters can partly avoid overfitting. However, the careful determination of suitable constraints requires a lot of time even for experienced image processing specialists, and novices generally fail. A more reliable approach is based on the calculation of a so-called gold standard FSC (Scheres and Chen, 2012), where the entire dataset is split before the first round of refinement. The iterative optimization is then performed on the two halves independently. After each iteration, the FSC curve between the two independent reconstructions is used to construct an optimal filter without any user-defined parameters.

Both, the correlation between neighboring particles used during our new alignment procedure and the contribution of all crystals to the starting model, violate the independency assumption on which the gold-standard method is based. Therefore, the realization of the gold-standard procedure is not straight forward for our application. We prevented overfitting by carefully applying conventional low-pass filters with user-defined parameters. Nevertheless, the gold-standard procedure will be implemented in future versions of the presented algorithm in one or another way. For instance, exploiting only a sub-group of the crystallographic symmetry present in most of the datasets could be used to cross-validate the reconstruction and its resolution.

4.4. Structural improvements

The crystals of the presented test dataset were relatively well-ordered and flat, which is favorable for classical MRC processing. Nevertheless, the new refinement procedure improves the resolution of the final reconstruction and shows more structural details (Fig. 5), although some protein backbone loops are not resolved,

likely due to the anisotropic resolution and the basic CTF-correction scheme employed here. From Fig. 5A, where the classically MRC-based merged structure (mesh, blue) and the new local refinement based single particle reconstruction are shown (solid, grey), we can conclude that the algorithm presented here resolves densities that cannot be seen in the reconstruction generated by classical MRC processing. The most significant differences are in the pore region of the membrane protein (Fig. 5B–D). The atomic model shown for comparison was devised based on X-ray structures from parts of the molecule, which were subjected to flexible molecular dynamic fitting (Schröder et al., 2007) to the cryo-EM reconstruction generated by classical MRC-based methods. Although this procedure introduced a bias to the atomistic model from the MRC-based reconstruction, the resulting atomic model matches best to the structure generated using our new single-particle procedure. For instance the new single-particle reconstruction resolves better the isolated helix S6 next to the central channel, as well as the densities for trapped ions in the central axis of the channel in the selectivity filter region (Fig. 5C and D). These and further manual observations suggest that our new refinement procedure outperforms the conventional MRC-based processing on this dataset.

An adequate choice of the iso-contour levels used to display cryo-EM reconstructions is crucial for a reliable comparison between two different structures. The correct way is to select the thresholds so that the shown volumes equal the molecular mass of the protein divided by the average protein density. Due to the missing cone and strongly anisotropic resolution commonly affecting 2D crystallography datasets, such an approach fails in the present situation. Thus, we selected the thresholds so that the long helices of the pore region of the protein perpendicular to the membrane have the same dimensions in both volumes.

4.5. Performance evaluation of the likelihood-based particle selection approach

Fig. 7 demonstrates the efficient skipping procedure developed to avoid the use of particles that are inconsistently aligned due to imaging artifacts or crystal defects. For instance the tilt angle distribution of the particles (black circles) of the crystal region shown in Fig. 7 shows regions of strong inconsistency (Fig. 7A; black ellipse). Based on the procedure described in Section 2.5 we calculate the log-likelihood to find each particle given the mean vector and the covariance matrix of the particle fingerprints of the entire crystal. The resulting log-likelihood values indicating the consistency of the particle alignment parameters are shown in Fig. 7B. The 5–15% of the particles with the lowest consistency values were not used for the subsequent reconstruction (Fig. 7C), which led to a further improvement of the reconstruction.

4.6. Computing performance of the refinement program

The processing presented here was done on a desktop machine featuring two Intel Xeon E5-2660 3.0 GHz 8-core CPUs, 64 Gb memory and two Nvidia GeForce GTX Titan with 2688 CUDA cores and 6 Gb DDR5 RAM each. Solid-state hard drives were used for faster writing and loading of the particle stacks.

As mentioned in Section 3.3 the bottleneck of the initial model refinement is the calculation of the projection gallery. Table 1 shows the computing time required to generate the projection gallery used for one initial model refinement step applied to the presented test dataset. Because the GPU-implementation efficiently uses all available host threads and all available hardware execution and data streams of the GPUs, the initial model refinement is now an interactive processing task that is 40 times faster than the fastest CPU implementation of the same task (Table 1). The calculation
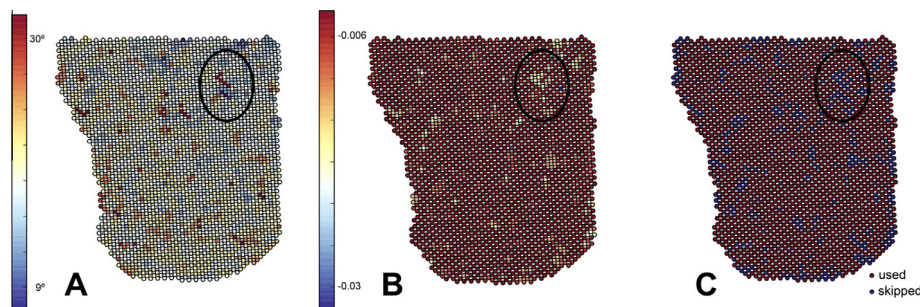
Fig.7. Likelihood-based particle selection within one 2D crystal. (A) Tilt angle distribution within the crystal. The area marked with a black ellipse shows inconsistently aligned particles as the tilt geometry varies more than 10° within the distance of two particles. (B) The likelihood values of the particles. Based on the particle finger print vector Eq. (5), we can calculate the likelihood to find each particle given the mean vector and covariance matrix of all particle finger print vectors of the crystal. (C) Retained and rejected particles. Particles with the lowest likelihood values (also called consistency) should not contribute to the reconstruction as their alignment most likely failed. In the present example 10% of the particles were removed (blue particles) whereas the remaining 90% of the particles were used for the back-projection step.

Table 1
Performance measure of the calculation of 54,166 projections, each 150×150 pixels. Speedup_1 states how many times faster the more sophisticated implementation is compared to the single threaded basic implementation, whereas Speedup_2 compares the GPGPU realizations with the fastest CPU implementation running on all available compute cores.

	Wall-clock time (s)	Speedup_1	Speedup_2
Single core ($1 \times$ Intel Xeon)	4943	–	–
16 cores ($2 \times$ Intel Xeon)	313	15.8	–
Single GPU, 1 CPU thread	128	38.6	2.4
Single GPU, 16 CPU cores	13.3	372	23.5
Dual GPU, 16 CPU cores	7.8	634	40.1

of one single projection in the 'full parallel GPU-setup' took 0.14 ms on average.

One iteration step of the first phase (coarse refinement level, Section 4.2) of the local tilt geometry refinement of the 210,538 particles with a box size of 150×150 pixels took about 100 min wall-clock time using the machine mentioned above. During the second phase of refinement, the cone used for the trial projection calculation depends on the previously optimized local orientation of the particles. Additionally, due to the smaller angular step size much more projections have to be computed. Thus the required projection gallery no longer fits into the memory of our machine. Instead of caching the projections during the second refinement phase, they are calculated on the fly. Due to the GPGPU acceleration of the projection calculation the fine-grained local refinement can be done within a reasonable time, resulting in 300 min wall-clock time for one step of the second refinement phase.

The software was successfully installed and run on several high-performance computing clusters. Refinements using up to 512 processor cores we performed. The hybrid parallelization approach was benchmarked on up to 128 cores and showed a fairly good scaling behavior (Supplementary Fig. 4).

5. Conclusions

The implementation of the described procedure in the *2dx* package and as a standalone high-performance computing cluster program, allows variations in the local tilt geometry of a 2D membrane protein crystal to be taken into account in order to increase the resolution and the amount of structural details visible in the final 3D structure. The concept of local averaging (central and neighboring particles) makes it possible to iteratively refine further imaging parameters, such as local magnification changes, deviations of the defocus from the pre-calculated values, or beam-tilt induced phase distortions. Up to now, the low SNR of low dose cryo-EM images has prohibited the successful fitting of

these imaging parameters based on projections of single membrane proteins picked from a crystal.

The high-performance computing implementation of the here-presented software prepares it for the processing of large datasets recorded automatically on next generation direct electron detectors. Instead of recording one image with the full electron dose, these detectors can be used to distribute the dose over multiple images (frames) in order to track beam-induced distortions. As recently shown (Bai et al., 2013; Campbell et al., 2012), beam-induced motion of individual particles can be resolved by means of processing individual dose-fractionated sub-frames from a direct electron detector operated in "movie mode". A future extension of the here presented approach to resolve beam-induced motion of the crystal lattice should further improve the reconstruction quality.

The here presented alignment procedure is based on cross-correlation maximization, which is prone to noise bias. Several recent studies demonstrated superior results (in both resolution and reliability) obtained by gold-standard maximum likelihood approaches compared to classical cross-correlation maximization (Lyumkis et al., 2013; Scheres, 2012). A maximum likelihood approach that considers the correlation between neighboring particles as prior knowledge due to the presence of crystals is therefore expected to also further improve the single particle reconstruction introduced here.

Although our new algorithm weakens the assumptions made about sample flatness, it still requires a successful pre-processing based on the classical MRC-based 2D crystallography pipeline. Thus, the presented procedure is a refinement tool rather than an independent reconstruction pipeline. As the entire conventional processing pipeline is based on the assumption of perfectly flat and well-ordered 2D crystals, growing them is still crucial. Large well-ordered mono-layered 2D crystals are extremely difficult and time consuming to obtain. In early crystallization trials one often observes 'polycrystals' that at least locally, within smaller patches, show some minimal crystallinity, e.g., of 10×10 crystal unit cells. In combination with the processing procedure presented here, a tool that picks particles from polycrystals would allow such badly ordered 2D membrane protein crystals to be processed. It would then be possible to determine the structure of membrane proteins whose 2D crystallization does not fulfill the stringent order and flatness requirements we know today.

All the here-described procedures are implemented in *2dx*-3.5.0, which will be available on www.2dx.org. The entire set of tilted images of MloK1 used to benchmark the software, together with the *2dx* processing parameters is available at the EBI at http://www.ebi.ac.uk/~ardan/aspera/em-aspera-demo.html under accession number 10006. 2dx/MRC-derived maps of MloK1 are available at the EMDB at accession codes EMD-2526 and EMD-2527. The fitted molecular models are available at the PDB under accession codes 4CHV and 4CHW.

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.2013.12.011.

References

Anderson, T.W., Olkin, I., 1985. Maximum-likelihood estimation of the parameters of a multivariate normal distribution. Linear Algebra Appl. 70, 147–171.

Arheit, M., Castaño-Díez, D., Thierry, R., Abeyrathne, P., Gipson, B.R., Stahlberg, H., 2013a. Merging of image data in electron crystallography. Electron Crystallography of Soluble and Membrane Proteins. Springer, p. 195–209.

Arheit, M., Castaño-Díez, D., Thierry, R., Gipson, B.R., Zeng, X., Stahlberg, H., 2013b. Image processing of 2D crystal images. Electron Crystallography of Soluble and Membrane Proteins. Springer, p. 171–194.

Bai, X.C., Fernandez, I.S., McMullan, G., Scheres, S.H., 2013. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. eLife 2, e00461.

Campbell, M.G., Cheng, A., Brilot, A.F., Moeller, A., Lyumkis, D., Veesler, D., Pan, J., Harrison, S.C., Potter, C.S., Carragher, B., Grigorieff, N., 2012. Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. Structure 20, 1823–1828.

Carragher, B., Kisseberth, N., Kriegman, D., Milligan, R.A., Potter, C.S., Pulokas, J., Reilein, A., 2000. Leginon: an automated system for acquisition of images from vitreous ice specimens. J. Struct. Biol. 132, 33–45.

Castano-Diez, D., Moser, D., Schoenegger, A., Pruggnaller, S., Frangakis, A.S., 2008. Performance evaluation of image processing algorithms on the GPU. J. Struct. Biol. 164, 153–160.

Crowther, R.A., Henderson, R., Smith, J.M., 1996. MRC image processing programs. J. Struct. Biol. 116, 9–16.

Dagum, L., Menon, R., 1998. OpenMP: an industry standard API for shared-memory programming. IEEE Comput. Sci. Eng. 5, 46–55.

Elmlund, D., Elmlund, H., 2012. SIMPLE: software for ab initio reconstruction of heterogeneous single-particles. J. Struct. Biol. 180, 420–427.

Frank, J., 1975. Averaging of low exposure electron micrographs of non-periodic objects. Ultramicroscopy 1, 159–162.

Gipson, B., Zeng, X., Zhang, Z.Y., Stahlberg, H., 2007. *2dx* – user-friendly image processing for 2D crystals. J. Struct. Biol. 157, 64–72.

Grigorieff, N., 2007. FREALIGN: high-resolution refinement of single particle structures. J. Struct. Biol. 157, 117–125.

Heymann, J.B., 2001. Bsoft: image and molecular processing in electron microscopy. J. Struct. Biol. 133, 156–169.

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.

Jost, G., Jin, H., Mey, D., Hatay, F.F., 2003. Comparing the OpenMP, MPI, and hybrid programming paradigms on an SMP cluster. Proc. EWOMP 3, 2003.

Kowal, J., Chami, M., Baumgartner, P., Arheit, M., Chiu, P.-L., Schröder, G.F., Nimigean, C.M., Stahlberg, H., 2013. Ligand-induced structural changes in the cyclic nucleotide-modulated potassium channel MloK1. Nat. Commun. (under review).

Kunji, E.R., von Gronau, S., Oesterhelt, D., Henderson, R., 2000. The three-dimensional structure of halorhodopsin to 5 Å by electron crystallography: a new unbending procedure for two-dimensional crystals by using a global reference structure. Proc. Natl. Acad. Sci. USA 97, 4637–4642.

Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., Cheng, Y., 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat. Methods 10, 584–590.

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

Lyumkis, D., Brilot, A.F., Theobald, D.L., Grigorieff, N., 2013. Likelihood-based classification of cryo-EM images using FREALIGN. J. Struct. Biol..

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

Penczek, P.A., 2010. Resolution measures in molecular electron microscopy. Methods Enzymol. 482, 73–100.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Radermacher, M., 1988. Three-dimensional reconstruction of single particles from random and nonrandom tilt series. J. Electron Microsc. Tech. 9, 359–394.

Scheres, S.H., 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–530.

Scheres, S.H., Chen, S., 2012. Prevention of overfitting in cryo-EM structure determination. Nat. Methods 9, 853–854.

Schröder, G.F., Brunger, A.T., Levitt, M., 2007. Combining efficient conformational sampling with a deformable elastic network model facilitates structure refinement at low resolution. Structure 15, 1630–1641.

Sorzano, C.O., Marabini, R., Velazquez-Muriel, J., Bilbao-Castro, J.R., Scheres, S.H., Carazo, J.M., Pascual-Montano, A., 2004. XMIPP: a new generation of an open-source image processing package for electron microscopy. J. Struct. Biol. 148, 194–204.

Stewart, A., Grigorieff, N., 2004. Noise bias in the refinement of structures derived from single particles. Ultramicroscopy 102, 67–84.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., Ludtke, S.J., 2007. EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. 157, 38–46.

Unger, V.M., 2013. Evaluation of electron crystallographic data from images of two-dimensional crystals. Methods Mol. Biol. 955, 211–227.

Zeng, X., Stahlberg, H., Grigorieff, N., 2007. A maximum likelihood approach to two-dimensional crystals. J. Struct. Biol. 160, 362–374.

Zhang, X., Zhou, Z.H., 2011. Limiting factors in atomic resolution cryo electron microscopy: no simple tricks. J. Struct. Biol. 175, 253–263.