Novel Convergence-Oriented Approach for Evaluation and Optimization of Workflow in Single Particle 2D Averaging of Electron Microscope Images

著者	守屋 俊夫
year	2013
その他のタイトル	電子顕微鏡像の単粒子二次元平均化におけるワークフローの評価と最適化のための収束思考アプローチ
学位授与大学	筑波大学 (University of Tsukuba)
学位授与年度	2013
報告番号	12102甲第6693号
URL	http://hdl.handle.net/2241/00121784

博士（神経科学）

学位論文

Novel Convergence-Oriented Approach for Evaluation and Optimization of Workflow in Single Particle 2D Averaging of Electron Microscope Images

電子顕微鏡像の単粒子二次元平均化におけるワークフローの評価と最適化のための収束思考アプローチ

平成２５年度

守屋　俊夫

筑波大学大学院人間総合科学研究科感性認知脳科学専攻

Table of Contents

General Introduction

Membrane proteins such as ion channels and receptors play major roles in the activity of the brain and nervous system, especially by initiating and conducting action potentials. They change their structures, and quite often oligomerize to higher-order structures to exert their full functions. Specifically, ion channels change their conformations among multiple states to regulate their target ion permeabilities. Therefore, determination and comparison of their 3D structures at near-atomic resolution provide valuable information for understanding brain function at the molecular level, and also for the development of medicines.

Several structure determination methods are currently available, including crystallography, Nuclear Magnetic Resonance (NMR), and Single Particle Analysis (SPA). The choice of method is mainly determined by the resolution required, the molecular mass, and ease of crystallization of the research target. Crystallography analyzes the diffraction patterns from protein crystals, mainly using X-ray or electron beam. It can achieve atomic-level resolution (1-3 Å) and is mainly used for relatively small proteins. However, some proteins are difficult to crystallize, and it is usually difficult to determine high-resolution 3D structure from the poorly ordered crystals. NMR is a spectroscopic technique which measures transitions between spin states of atomic nucleus under an electromagnetic field. Although it can also achieve atomic-level resolution, it has a practical limit of molecular mass up to 20-80 kDa (Skrisovska et al., 2010). SPA is a combination of Electron Microscopy (EM) and computer image processing. The Signal-to-Noise Ratio (SNR) of each individual EM image of the target protein is usually very low, mainly due to the reduced electron dose during the recording necessary to avoid damage to the specimen. The dose is commonly limited to between 15 and 25 electrons/Å^2 as integrated flux densities of electrons (Frank, 2009). However, classifying particles into groups numbering in the hundreds or thousands, and averaging them, can improve the contrast of the target proteins. Computation can also determine the Euler angle of each average and can reconstruct the original 3D structure. SPA does not require crystallization and can theoretically be applied to macromolecules having a mass greater than 100 kDa in cryo-EM (Frank, 2006; Henderson, 1995). Ion channels frequently form higher-order structure by multimeric assembly and are difficult to crystallize. Therefore, SPA is suitable for structural analysis of an entire complex of ion channels in a close-to-native state.

Two major considerations in developing SPA are resolution and analysis time. The current resolution of SPA is 4 to 10 Å (near-atomic resolution) using cryo-EM (Cong et al., 2010; Ludtke et al., 2008; Zhang et al., 2010), which is still lower than that of crystallography and NMR. A notable obstacle to achieving high-resolution with SPA has been the extremely low SNR of electron micrographs. Great improvements in EM have mitigated this without increasing the electron dose. Zero-loss imaging with energy filter (Langmore and Smith, 1992; Schroder et al., 1990; Zhu et al., 1997) has been used to improve SNR in recent high-resolution analysis (Cong et al., 2010; Ludtke et al., 2008; Zhang et al., 2010), which dramatically improves particle recognition from background noise by both human eyes and computers. The development of high-resolution CCD cameras and automatic EM recording systems as well as improvements in computational ability now enable handling of huge volumes of data and accelerate high-resolution analysis.

The other important goal in SPA is reduction of analysis time, which should be achieved simultaneously with improvement of resolution. To understand the conformational change of the target protein (especially ion channels), high-resolution structures of different conformations should be obtained in a realistic period, although SPA is a multi-step image processing.

In typical SPA, there are four major stages: (1) particle pickup, (2) initial 2D averaging (reference-free), (3) initial 3D reconstruction (reference-free), and (4) 3D refinement by projection matching with 3D angle reassignment. After obtaining digitized micrographs of highly purified protein, the first stage is particle pickup: creating a particle library by selecting the regions containing a single particle view. This particle library is corrected for the contrast transfer function (CTF) and is band-pass filtered. In the second stage, class averages are calculated from the particle library to improve SNR of particle views. This stage is repeated until convergence. In the third stage, the initial 3D structure is reconstructed from the class averages, using their estimated 3D projection angles. The fourth stage improves the quality of the 2D averages, their 3D angle estimations, and 3D structure, using projections of the 3D structure from a previous cycle as a reference. This stage is repeated until a stable 3D structure is obtained. Many of these stages consist of several algorithms and are iterative processes; sequences of stages are often repeated multiple times. To reduce the analysis time for SPA, these cycling processes should be effectively operated and automated.

Several software packages have been developed and distributed in the SPA field for reducing analysis time as well as improving resolution. Examples include IMAGIC (van Heel et al., 1996; van Heel et al., 2011), SPIDER (Frank et al., 1996; Shaikh et al., 2008), EMAN (Ludtke et al., 1999; Tang et al., 2007), XMIPP (Marabini et al., 1996; Scheres et al., 2008; Sorzano et al., 2004), and FREALIGN (Grigorieff, 2007). These packages have greatly facilitated the progress and dissemination of this structure determination method. However, a single package is not usually sufficient for the analysis, due to the differences in theoretical approaches of analyses and in the target proteins for their software development. Researchers must often develop an original workflow integrating basic algorithms for each target protein. A workflow is an ordered set of basic algorithms with iterative steps and is likely to be extremely complex. In this workflow optimization task, the comparative evaluation of different algorithm integrations and different parameter settings has often been aided by human image recognition and experience. The necessity of these human interventions is a large part of the long analysis period and hinders automation in these software packages. Particularly, the qualitative evaluation by a human is problematic since this prevents further automation.

To further reduce analysis time in SPA and improve automation of analysis, I carefully examined two main qualitative evaluations during the optimization of SPA workflows: visual inspection of input/output images and convergence evaluation of the iterative steps. Particle images, 2D average images, and projected images from resultant 3D structures are visually inspected to assess the quality of outputs, such as their consistencies and resolutions. Many objective quantitative criteria to replace this qualitative inspection have been already proposed in the SPA field. Examples are a similarity measure using the Cross-Correlation Coefficient (CCC) and a resolution measure using Fourier Shell Correlation / Fourier Ring Correlation (FSC, Harauz and van Heel, 1986; FRC, Saxton and Baumeister, 1982). However, according to our experience from previous analyses, it has been difficult to rate the differences between optimizations by individually checking the resultant values of any single criterion. The value differences of a criterion have been often too subtle to be conclusive and the ratings of different criteria have often conflicted.

The other main qualitative evaluation is for the convergence of iterative steps. Several SPA studies have used a graph of resolution history, such as FSC (Cong and Ludtke, 2010; Frank, 2006; Ludtke et al., 2001) and the Differential Phase Residual (DPR, Frank et al., 1981). However, quantitative criteria have not been established for the final decisions of whether workflows converged or not, even though clear differences have been observed in convergence trends of criterion values before and after workflow optimizations in previous analyses.

Analyzing qualitative evaluation by humans along with our previous experience, three observations emerge: (1) A quantitative criteria of workflow convergence has not been established, (2) The difference between different workflow optimizations has not been statistically evaluated, and (3) Humans have intervened in comprehensive evaluation of multiple criteria. Solving these problems should allow the evaluation and optimization of the SPA workflow without depending on human interpretation and experience. Software can set parameter values systematically and should be able to automatically select the most appropriate optimized state of a workflow using a quantitative evaluation.

Based on this background, I propose the Integrated Convergence-Evaluation Oriented System (IC-EOS). The IC-EOS is an evaluation system, oriented to the convergence of criterion values, for SPA workflow optimization. To enhance the reliability of the decision, the system integrates the convergence evaluations of multiple criteria. Inspired by previous works (Cong and Ludtke, 2010; Frank, 2006; Frank et al., 1981; Ludtke et al., 2001; Sigworth, 1998), it visualizes histories of criterion values. Furthermore, curve fitting is applied to history graphs and provides convergence measurements: fluctuation amplitude and improvement width. These measurements are used to evaluate convergence. The computational selection vote system for most-appropriate condition sets is also developed.

The IC-EOS improves SPA accuracy and minimizes calculation time by reducing trial-and-error steps and by supporting building of complex workflows. This research should open the path to the automatic determination of the structures of ion channels and other important proteins at high-resolution.

Part I: Theoretical background for IC-EOS development

One recent emphasis in SPA methodology has been on the optimization of workflow. Although software packages for SPA provide most of the basic algorithms, the combination and order of algorithms must be frequently rearranged. In addition, multiple searches for adequate parameters are usually necessary. An evaluation step follows to assess each applied optimization. The optimization and evaluation are repeated until a good balance among multiple criteria is reached. The balance is a key to achieving high-resolution, and is specific to the nature of the target protein as well as the type and quality of micrographs. This optimization task often takes several months or even years, especially because the evaluation has been rather qualitative than quantitative. Scoring the workflow optimization is necessary for a more objective evaluation and for automation of SPA, and it is desirable to quantify criteria of these qualitative evaluations. Therefore, to reduce the analysis time of SPA by the automation, the three observations stated in the General Introduction must be addressed. This evaluation system for SPA workflow optimization, named IC-EOS, was developed by focusing on the convergence of multiple quantitative criteria.

1. Establishing a quantitative decision method of workflow convergence

The first agenda is to establish a quantitative method to decide the convergence of a criterion value. For this purpose, a mathematical model of expected convergence trend is fitted to an iteration history curve of a quantitative criterion, and the trend is quantified. Using obtained trend measures, it was hypothesized that it would be possible to make an algorithm of the convergence decision.

The convergence trend of a criterion value was expected to be either decreasing or increasing with saturation, or constant. A good mathematical model for decreasing/increasing trends with saturation is exponential decay. Let y be the value of a criterion and x be the iteration count. Exponential decay is defined by:

$$y = s + we^{-rx} \quad r > 0 \text{ (Decreasing form)}$$

$$y = s + w(1 - e^{-rx}) \quad r > 0 \text{ (Increasing form)}$$

, where s is a value offset (i.e. a saturation value in the decreasing form or an initial value in the increasing form), w is the decay width, and r indicates the decay rate. The parameters of this model can explain some properties of the convergence trend. Using exponential decay as the mathematical model, the "curve fitting" by the least-squares method (minimizing the sum of squared residuals) is applied to the observed history curve of the criterion value in the convergence decision (Fig. 1).

2. Statistical evaluation of workflow optimization difference: fluctuation measurement

The second agenda is to evaluate the difference of workflow optimizations using a statistical method. It was hypothesized that the stability of convergence trend, quantified by curve fitting, should show a statistical difference.

Multiple history curves of a criterion measure are obtained by varying conditions systematically. Using the curve fitting to each history curve, a Root-Mean-Square Deviation (RMSD) is calculated. Since RMSD indicates the average residual of the history curve, the fluctuation measurement, defined by this RMSD, is used as a convergence measurement to quantify the stability of the convergence trend. The small value of fluctuation indicates that the history curve is visually smooth and the criterion value is stable. That is, the condition producing “the most-stable” trend is indicated by the “minimum” fluctuation value of all conditions, and the condition resulting in “the most-unstable” trend is indicated by the “maximum”. The stability of each condition is measured by its statistical difference from these minimum and maximum fluctuation values. Two fluctuations are defined to be significantly-different if they have different variances of residual distributions. Accordingly, the F-test is used, and its test statistic is defined by the ratio between the squared residual variances of two different conditions of a single criterion measure. The larger variance is the numerator. The test statistic is assumed to follow the F-distribution with N1 - 1 and N2 - 1 degrees of freedom under the null hypothesis (i.e. two distributions have the same variance), where N1 and N2 are the sample sizes of the numerator and the denominator (i.e. 30 cycles). The significant level was set to 0.001 with the two-tailed test.

The value difference of a criterion is also considered statistically. I hypothesized that the fluctuation value is the error of a criterion value. Two criterion values are defined to be significantly-different if their difference is larger than 2.58 times the sum of their fluctuations (i.e. the likelihood of their fluctuations causing this size difference is 1%). Otherwise, their fluctuation error range overlaps each other and they are considered to be statistically the same value.

3. Algorithm of comprehensive decision: base measurements and best condition election

The third agenda is to make an algorithm of comprehensive decision. Introducing the computer voting method using multiple criteria, it was hypothesized that the best condition set should obtain the largest number of quantitative criteria showing the most-stable expected trend. This should allow a highly reliable comprehensive decision.

Currently, visual inspection of input/output data is done many times at different stages of the analysis, along with evaluation using quantitative criteria. The most frequently used criterion in SPA is CCC, which measures similarity between a reference image (i.e. a class average or reprojection) and the particle image from EM. Another criterion is resolution. The commonly-used definitions for resolution are the Q-factor (Kessel et al., 1985; van Heel and Hollenberg, 1980), the S-image (Sass et al., 1989), DPR (Frank et al., 1981), FSC (Harauz and van Heel, 1986)/FRC (Saxton and Baumeister, 1982), and the Spectral SNR (SSNR, Unser et al., 1987). Of these, the most widely used are FSC and FRC, because of ease of calculation and versatility, as well as their relationship to the SSNR (Penczek, 2010).

Accordingly, the following five base measurements were selected for quantitative tracking of convergence trends: (1) FRC/FSC-resolution of a 2D class average or a 3D reference structure (in Å/cycle), (2) Total Intra-Class variance (TIC-variance) of a class average (normalized), (3) CCC-similarity of a class average and the corresponding constituent particle (normalized), (4) shift in 2D alignment (in pixels), and (5) rotation in 2D alignment (in degrees). FRC or FSC with a 0.5 CCC criterion was chosen as a resolution measurement. By definition, these indicate how well the fine features match among the constituent particles of a 2D average or a 3D structure. They also measure self-consistency in frequency space, that is, how closely the particles in a class resemble each other. Together, the TIC-variance of class averages can measure self-consistency in real space. This value should also reflect changes in signal and noise variance in particle images. For stability of class constituents, each particle-average pair should show the highest correlation among all other possible combinations at the converged state. Creating a quantitative measurement for the cessation of shift and rotation can be directly associated with the shifted-distance and rotated-angle of particles by 2D alignment. Each of these five values has been individually used as a quantitative criterion or “energy function” (objective function) measuring the achievement level of the workflow objective. IC-EOS uses all of them together, as base measurements for the convergence trend. Mean and Standard Deviation (SD) of each base measurement are calculated at each iteration cycle, and their histories are recorded.

The best condition election by voting method consists of four procedures: (1) convergence sub-vote, (2) fluctuation sub-vote, (3) best condition total vote, and (4) mean-value inspection.

The convergence sub-vote ranks conditions with the number of base measurements showing the expected trend, using all “converged” or “not-converged” decisions. This vote is divided into two categories; for base measurements with “desirable convergences” and for those with “undesirable convergences”. As described in subsection 2.1 of Part II, the preliminary inspection of all obtained data revealed that, when the appropriate condition set was used, values of some base measurements should be “not-converged”. This is because the first cycle of iteration steps was sufficient for these measurements to converge, and their values were constant during the later iteration cycles. Therefore, their convergences are regarded as “undesirable”. The vote simply counts up the number of “converged” marks for the desirable category (A1) and “not-converged” marks for the undesirable category (A2). A larger sub-vote sum (A1 + A2) is better.

The fluctuation sub-vote counts the number of base measurements which obtained the most-stable-trend and rates each condition, using the statistical difference of fluctuation values. For a given base measurement, the procedure evaluates whether the convergence stability is significantly-not-different from the condition of most-stable convergence (i.e. the minimum fluctuation). If this is true, a score is given to the base measurement. The procedure also evaluates whether the convergence stability is significantly-different from the condition of the most-unstable convergence, and again a score is given. A higher score is better for the minimum vote (B1), and lower is better for the maximum vote (B2). A larger sub-vote difference (B1 - B2) is better.

The best-condition total vote is the sub-vote sum plus the sub-vote difference ((A1 + A2) + (B1 - B2)). The condition with the largest number of votes is elected. This ensures that the elected condition has obtained the largest number of base measurements showing the most-stable expected trends.

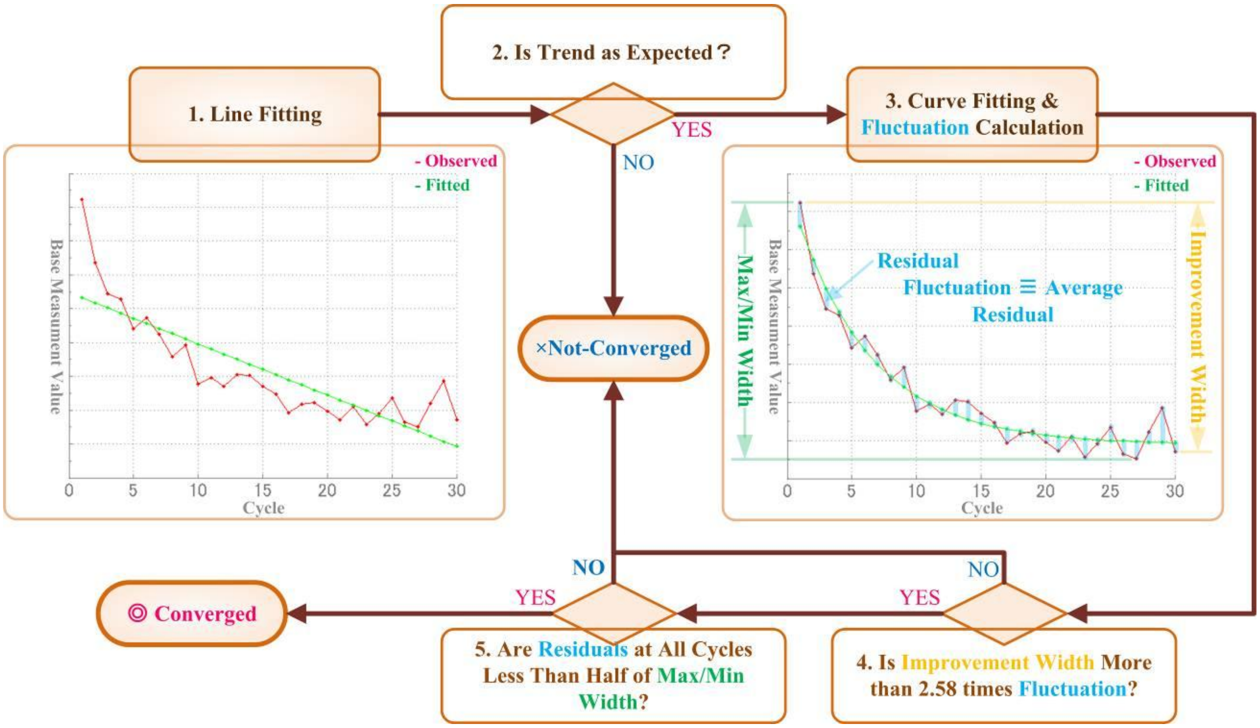
In the mean value inspection, the mean values are evaluated relative to the best and worst values, and used to confirm whether or not each mean value of the elected condition is optimal among the experimental conditions. The statistically-optimal value is defined to be both not-significantly-different from the best mean value and significantly-different from the worst. For a given base measurement history, the procedure checks if the mean value of each condition is statistically optimal.

4. Construction of IC-EOS

IC-EOS is constructed as an external extension tool for existing software packages. The system consists of three components: (1) extraction of five base measurement values from all iteration cycles of the workflow after the completion of 2D averaging, (2) calculation of their statistics, and (3) convergence evaluation. The modules for value extraction and statistic calculation were implemented using C++ and DOS command batch files. For convergence evaluation, including curve fitting and various other statistical analyses, Microsoft Office Excel 2007 (Microsoft Corporation, Redmond, WA) was used.

Figure 1. Convergence evaluation by IC-EOS

Flowchart of convergence evaluation by IC-EOS curve fitting. The evaluation consists of the following process: (1) Line fitting to the base measurement history curve. (2) Determining whether the slope of the fitted line matches the expected increase/decrease trend of the history curve. If not, the history curve is determined to be “not-converged”. (3) Curve fitting of the expected mathematical model (i.e. exponential decay) to the history curve by minimizing the sum of the squared residuals (i.e. deviation between fitted and observed values) at each cycle. At the same time, fluctuation (defined by RMSD) is calculated, which indicates the average residual of the history curve. (4) Determining whether the improvement width (i.e. the difference between values at first and last cycles) is larger than 2.58 times fluctuation (i.e. with a probability of 99%, fluctuation values are expected within this range and the improvement width is not caused by fluctuation). If not, the history curve is determined to be “not-converged”. (5) Determining whether the residual at each cycle is less than 50% of the maximum and minimum value difference (i.e. min/max width), to ensure that unexpected sudden value changes did not happen during the iteration cycles. If not, the history curve is determined to be “not-converged”. A workflow for a condition set is determined to be “converged” only if the history curve passes all these criteria. Statistical analysis of the fluctuations and residuals was made in the best condition election by IC-EOS (defined in section 3 of Part I). This evaluation procedure was applied to each history curve. For a condition set, a total of 10 history curves (two statistic values of five base measurements) were obtained, as seen in Figure 5. Therefore, a total of ten fluctuations and ten convergence decisions were obtained for each condition.


1. Line Fitting
2. Is Trend as Expected？
YES
NO
3. Curve Fitting & Fluctuation Calculation
- Observed
- Fitted
Base Measument Value
Cycle
0
5
10
15
20
25
30
×Not-Converged
- Observed
- Fitted
Max/Min Width
Base Measument Value
Residual
Fluctuation ≡ Average Residual
Improvement Width
Cycle
0
5
10
15
20
25
30
NO
NO
YES
YES
◎ Converged
5. Are Residuals at All Cycles Less Than Half of Max/Min Width?
4. Is Improvement Width More than 2.58 times Fluctuation?

Part II: Case study examination of IC-EOS using NaChBac

1. Introduction

To examine the validity of the convergence-oriented approach for evaluation and optimization of SPA workflow described in Part I, a case study was conducted by applying IC-EOS to the 2D averaging of a negatively stained NaChBac (bacterial Voltage-Gated Sodium (Na_V) channel) from *Bacillus haloduran*. In Part V, the 3D structure of NaChBac was also determined using SPA (see subsection 2.1 and 2.2 of Part V for the preparation of the particle library).

The particle images for SPA are taken with a transmission electron microscope after the purification of proteins. The most commonly used specimen preparations are negative staining and ice-embedding. Negative staining with a heavy metal salt such as uranyl acetate produces high contrast, and protects the protein from deformation under vacuum. Although some particles show deformation or staining variations, the high contrast is still advantageous for the throughput of analysis. For cryo-EM, the purified proteins in solution are rapidly plunged into liquid ethane cooled by liquid nitrogen, whereupon the thin water film vitrifies and embeds the samples. This makes it possible to obtain particle images in a close-to-native state under fully hydrated conditions. However, the density of proteins is only slightly higher than that of vitreous ice, and the contrast between the unstained sample and the surrounding ice is quite low. Because of the extremely low SNR in the recorded particle images, analysis requires a proportionally large number of particle image data and of alignment/classification calculation repetitions, which prolong analysis time. Therefore, negatively stained samples were used in the first examination of IC-EOS.

Addressing 2D averaging is especially important because it is the first critical workflow with iterative steps for 3D reconstruction. A key requirement of SPA is to obtain a high-quality initial structure. From our experience, the quality of the initial structure strongly influences the stability and plausibility of the later stages, as well as the resolution of the final 3D structure. To achieve this objective, the accuracy of 3D projection angle estimation is important. In SPA, the projection directions of 2D particle images must be estimated from these images themselves, since particles are expected to be oriented randomly on an EM sample holder. This estimation process distinguishes SPA from other 3D reconstruction methods such as tomography, where the projection angle of each detected 2D image is known in advance. The estimation accuracy is mainly influenced not only by the 3D angle estimation algorithm but also by the quality of class averages resulting from the iterative alignment/classification of 2D averaging. Therefore, the iteration steps of 2D averaging can be considered to establish the basis for analysis quality.

Several convergence criteria for 2D averaging have been proposed in previous research: (1) resolution improvement of class averages should reach saturation (Frank, 2006; Frank et al., 1981), (2) no particle should migrate between different classes (van Heel and Stoffler-Meilicke, 1985), (3) refinement should decrease total image variance (van Heel and Stoffler-Meilicke, 1985), and (4) most particle images should cease to shift or rotate (Penczek et al., 1992; van Heel and Stoffler-Meilicke, 1985). Since some of these criteria are somewhat qualitative, they need to be translated into quantitative measurements to automatically evaluate convergence of 2D averaging. Ideally, all of these criteria are satisfied when the workflow reaches a converged state. Therefore, the reliability of workflow outputs will be much improved if all criteria are integrated. The proposal here of using five base measurements fulfills all statements.

The optimization of 2D averaging was started from the widely-used workflow described in the IMAGIC-5 manual (IMAGIC-5, 2011; van Heel et al., 1996). This workflow consists of five main steps (Fig. 2a-A): (1) coarse centering of particle views to their total average (Dube et al., 1993), (2) initial classification of coarsely centered particle views (Dube et al., 1993), (3) mutual alignment of class averages, (4) Multi-Reference Alignment (MRA) of particle views relative to class averages, and (5) classification of aligned particle views. The third through fifth steps were repeated for a fixed number of cycles, using the class averages yielded at the fifth step as new references for the subsequent iteration.

2. Materials and Methods

2.1. Expected trend of five base measurements for 2D averaging

To determine the 2D resolution within a class, the class average obtained from 2D averaging was split into two independent sub-averages, which were compared by the FRC method (van Heel and Stoffler-Meilicke, 1985; van Heel et al., 1982) with a 0.5 CCC criterion (van Heel and Schatz, 2005).

The expected convergence trend of mean values for each base measurement is as follows: (1) a decreasing trend with gradual saturation, for FRC-resolution and TIC-variance as well as shift-distance, (2) an increasing trend with saturation for CCC-similarity, and (3) a constant trend for rotation-angle. Mean value of rotation-angles is constant, since the positive and negative angles are expected to cancel each other out. The SD values of all base measurements are expected to show a decreasing trend with saturation. However, contrary to this original hypothesis, the preliminary inspection showed that the convergences are “undesirable” for the SD of FRC-resolution and the SD of TIC-variance. These values converge at the first cycle of 2D averaging and are constant during the iteration cycles.

For mean value inspection, only the FRC-resolution and CCC-similarity mean values of the last cycle are used to simplify the inspection. The minimum mean value is the best for FRC-resolution, while the maximum is the best for CCC-similarity.

2.2 Particle-discard and classification variations in 2D averaging

Two independent variables were chosen: discard ratio and Class-To-Particle ratio (CTP ratio). The trends of base measurement histories were then monitored and the workflow was evaluated. The discard ratio was defined as the ratio of discarded image counts to the total image counts in the library; the discarded images included those of background noise without any particle, aggregated particles, fragmented particles, deformed particles, and particles with variations. The CTP ratio (CTP) was defined as the square root of the ratio of class counts (k) per particle counts in a library (N):

$$CTP = \sqrt{\frac{k}{N}}$$

The number inside the square root is the inverse of the average number of particles in each class. CTP ratio indicates the expected reduction of the noise level by 2D averaging (Rosenthal and Henderson, 2003).

To achieve the highest possible resolution from a given dataset, both discard and CTP ratios must be appropriate. When the number of classification groups is too large, the number of particles in a class will be small. This limits the reduction effect on noise level by 2D averaging. Inversely, if the number of classes is too small, particles of different appearances (or projection angles) are mixed into a single class. This blurs the average image of this class. Even though the appropriate class is used, an unsatisfactory particle selection causes the inclusion of non-true-particle images into some classes. This also blurs class averages. Thus, an inappropriate discard ratio or CTP ratio reduces the resolution of class averages.

3. Results

3.1. IC-EOS evaluation of 2D averaging workflow before optimization

The convergence of unoptimized 2D averaging was first evaluated using the proposed criteria (Fig. 1, Fig. 2a-A). The top four rows in Figure 3a summarize the 2D averaging with discard ratio variation; the bottom four rows are of CTP ratio variation. For the discard ratio variation, the numbers of particles were 10947, 8713, 6848, and 5495 (each a reduction of 20%), while the number of classes was kept constant at 200 (Fig. 3a, Table 1a). The selection was based on the CCC of particle images relative to the total average. For the CTP ratio variation, the numbers of classes were 100, 200, 300, and 400, while the number of particles in the library was kept constant at 5495 (Fig. 3a, Table 1b), where visual inspection did not reveal any obvious non-true particles. The iterative steps of workflows were repeated for 30 cycles.

Most of the base measurements indicated “not-converged” (Fig. 3a, blue crosses). The exception was the TIC-variance mean value, which got “converged” marks (red circles) in all condition sets except in that of 10947 particles (200 classes). The condition set with 100 classes (5495 particles, indicated by an orange outline) got “converged” marks in several measurements including FRC-resolution, indicating the most stable condition set. The visual trends of “converged” and “not-converged” of each base measurement (Fig. 3b) fitted well with the quantitative results. The resultant class averages at the final 30th cycle of this condition set compared well with raw particle images in each class (Fig. 4), indicating that our 2D averaging was consistent with the dataset. Although assessment of convergence achievement could be introduced to the unoptimized workflow, best-condition search was complicated due to the workflow's sensitivity to condition variation. Therefore, I proceeded to optimization of the workflow.

3.1.1. Investigation of condition-sensitivity

To investigate the condition-sensitivity of an unoptimized workflow, it was informative to check the mean value trend of CCC-similarity. History curves of a "not-converged" condition set (5495 particles and 300 classes) show that the mean value of CCC-similarity had a stable trend of increasing to convergence from the 1st to 8th cycle, but the value suddenly dropped off at the 9th cycle (Fig. 3c). The value recovered steadily up to the 15th cycle, but again dropped drastically at the 16th cycle and even more at the 17th cycle. This up-and-down repeated through whole iterations. At the same cycles where CCC mean value dropped drastically, shift-distance mean value also increased (Fig. 3c, highlighted with blue lines), indicating a large movement of most particles. As indicated by the red cross marks in Figure 3d, the orientation of the total average clearly changed from the 16th to the 18th cycle, which showed the most drastic movements of CCC-similarity mean value. In addition, the orientations of class averages were not always consistent, especially in the 17th cycle. From this, it appears that the 17th cycle was a transitional state of the particle library from the well- to poorly-aligned state. These results indicate that the particle orientation and position in each image square tend to change dramatically through whole iterations. These significant changes in appearance, orientation, or position of particle averages between any consecutive cycles are undesirable for alignment and classification algorithms.

3.2. Optimization of 2D averaging by particle uprighting

The workflow was modified to achieve more robust convergence by introducing the "particle uprighting" algorithm (Fig. 2b). This algorithm was designed to obtain the maximum consistency of particle orientation and position through whole iteration cycles, by forcing all particle views to be upright and centered at every cycle. OpenCV v2.3 (OpenCV, 2011) was used for the implementation of the algorithm. The modified workflow (Fig. 2a-B) is comprised of seven main steps: (1) coarse centering of particle images to their total average, (2) initial classification of coarsely centered particle images, (3) uprighting of class averages, (4) uprighting of the total average, (5) mutual alignment of class averages with the upright total average as a base reference (to which all class averages were aligned before the mutual alignment), (6) MRA of particle images using the aligned class averages as references, and (7) classification of the particle images. The fourth through seventh steps were repeated for 30 cycles, using the produced class averages as new references for the subsequent iteration. The introduction of a particle uprighting algorithm was examined using two types of condition variations: CTP ratio variation with a constant discard ratio (Table 2a), and discard ratio variation with a constant CTP ratio (Table 2b).

3.2.1. Number of classes: CTP ratio variation

Introducing the particle uprighting algorithm dramatically enhanced convergence in the CTP ratio variation experiment. The histories of the five base measurements were obtained first (Fig. 5). Fluctuation curves (Fig. 6a) and value curves (Fig. 6b), generated from these histories, represent the changes in each base measurement with the condition variation. These histories were then used as input data for the best condition election of IC-EOS (Fig. 6c). Among the mean-fluctuation curves of the base measurements relative to the CTP ratio (Fig. 6a), the FRC-resolution presented an increasing trend (less stability), while those of the others showed decreasing trends (more stability). All the curves had plateaus at higher CTP ratios. FRC-resolution and CCC-similarity also had plateaus at lower CTP ratios, resembling a sigmoid function. The trends of SD-fluctuations did not have any clear plateaus compared with those of mean fluctuations. The SD-fluctuation curves of FRC-resolution, CCC-similarity, and TIC-variance showed an increasing trend (less stability), while those of rotation-angle and shift-distance showed a decreasing trend (more stability).

Value curves of base measurements depending on CTP ratio (Fig. 6b) showed a variety of trends; the width between the first cycle (orange and aqua curves) and last cycle (red and blue curves) changed relative to CTP ratio variation, for several of the base measurements. As CTP ratio increased, FRC-resolution mean value increased almost linearly, indicating linear degradation. CCC-similarity mean value also showed an almost linear increase, but this trend indicates linear improvement. That is, mean values of FRC-resolution and CCC-similarity were in a tradeoff relationship. TIC-variance mean value decreased almost linearly (see subsection 4.7 of Part II about the interpretation of TIC-variance).

The result of the best condition election shows all base measurements, except FRC-resolution SD and TIC-variance SD, converged in most of the condition sets (Fig. 6c). These SDs converged only at the end points of the experimental condition sets, where other base measurements got "not-converged" marks. The convergence sub-vote section (Fig. 6c-1) indicated that most conditions showed desirable trends (red mark). A few undesirable trends (black mark) appeared at both ends of the CTP ratios (25 and 400 classes). Larger counts of convergence sub-vote sums (Fig. 6c-1), fluctuation sub-vote differences (Fig. 6c-2), and total votes (Fig. 6c-3) concentrated at around 150 classes (0.165 CTP ratio), which were selected by the best condition election (highlighted with yellow). This shows that the condition with 150 classes converged most stably. However, the mean value inspection (Fig. 6c-4) indicated that the condition for statistically-optimal FRC mean value was only at 25 classes while that for statistically-optimal CCC was only at 400, reflecting their tradeoff relationship. Furthermore, both mean values of the selected 150 classes were significantly-different from these statistically-optimal values. All together, FRC and CCC mean values of the most stable condition were located near the average levels of those of all CTP ratios.

3.2.2. Particle selection: discard ratio variation

The particle uprighting algorithm was also examined in relation to discard ratio variation (Fig. 7). Trends of the fluctuation curves of base measurements relative to the discard ratio (Fig. 7a) were not as clear as those for CTP ratio variation. In particular, no clear plateaus were found except on the curve of CCC-similarity mean fluctuation. This curve had a plateau at lower discard ratios and increased almost linearly on the higher end, reflecting unstable CCC-similarity when the discard ratio was high. As the discard ratio increased, FRC-resolution mean value decreased (improved) to the saturation point, while CCC-similarity mean value increased (improved) (Fig. 7b). In these curves, the increasing or decreasing rate changed at higher discard ratios. That is, FRC-resolution and CCC-similarity improved together when the discard ratio increased, at a cost of unstable CCC-similarity. TIC-variance mean value showed a similar trend to the CCC-similarity mean value.

The best condition election of the discard ratio variation experiments (Fig. 7c) showed similar results to those of CTP variation experiments (Fig. 6c). All base measurements, except FRC-resolution SD and TIC-variance SD, converged (Fig. 7c). Notably, TIC-variance SD got "converged" marks at lower discard ratios (Fig. 7c-1, green box). Visual inspection confirmed that these libraries contained a relatively high ratio of non-true-particle images. Higher counts of convergence sub-vote sums (Fig. 7c-1), fluctuation sub-vote differences (Fig. 7c-2), and total votes (Fig. 7c-3) concentrated at around 6848 and 5495 particles. The best condition election selected 5495 particles (0.50 discard ratio; highlighted with yellow). The mean value inspection (Fig. 7c-4) also indicated that this condition achieved the statistically-optimal values of both FRC and CCC. These statistically-optimal values overlapped at 5495, 4373, and 3272 particles (orange box) and they were not-significantly different from values of the most stable condition.

3.3. Further optimization of 2D averaging by intra-class alignment

To further improve 2D averaging, the "intra-class alignment" algorithm was applied (Fig. 2a-C). Particles in each class were aligned to the class average after image classification at the seventh step of the workflow. This alignment was repeated five times in each iteration. Refinement by intra-class alignment was inspired by the work of Ludtke et al. (Ludtke et al., 2004), who used this algorithm in 3D refinement by projection matching to reduce the strong bias of the current volume. Its application to the 2D averaging was expected to reduce the bias of the 2D references. This modification also allowed us to use finer translational and rotational alignment without a huge increase in computation time, since the number of image alignments is much smaller than that using MRA.

Workflows with and without intra-class alignment were compared by IC-EOS using the same condition sets (Table 2a and 2b). Additional intra-class alignment made 2D averaging more condition-sensitive, while it improved FRC-resolution. A few "not-converged" marks were obtained in almost all base measurements. This sensitivity became more apparent in relation to CTP ratio variation; CCC-similarity, rotation-angle, and shift-distance were affected most (Fig. 8c-1). For discard ratio, the condition set with 5495 particles and 100 classes was deselected by the vote, and the CCC-similarity mean value of this condition set was "not-converged" (Fig. 9c-1). Comparing the fluctuation curves between with- and without-intra-class alignment cases, the amplitudes of fluctuations were not so different (Fig. 8a and 9a, solid and dotted curves). However, there were differences in the values of the first cycle (Fig. 8b and 9b, orange and aqua curves) and last cycle (red and blue curves). For example, CCC-similarity SD value curves of the first cycle in Figure 8b show that the values with intra-class alignment (solid aqua curve) were lower than the ones without this algorithm (dotted aqua curve) at all CTP ratios. On the other hand, the value curves of the last cycle (solid and dotted blue curves) were not so different. Consequently, the improvement width became narrower with this algorithm. The other measurements of this optimized workflow also showed narrowing of the improvement width, especially at the "not-converged" condition sets. However, intra-class alignment achieved better FRC-resolution mean value and did not strongly affect its convergence (Fig. 8c-1 and 9c-1). The improvement width of this mean value was almost the same because both the first cycle values (Fig. 8b and 9b, solid relative to dotted orange curve) and last cycle values (solid relative to dotted red curve) shifted lower in all condition sets.

The best CTP ratio was 0.165 (150 classes; Fig. 8c, highlighted with yellow) and the best discard ratio was 0.29 (7817 particles; Fig. 9c, highlighted with yellow), when intra-class alignment was included. The elected CTP ratio was consistent with the workflow without intra-class alignment (Fig. 6c), while the elected discard ratio was different (Fig. 7c). Comparing the total votes between Figure 6c and 8c, the convergence-stability of the 100 class set (0.135 CTP ratio) decreased more than that of the 150 class set, as the condition-sensitivity of the workflow was increased. This observation indicated that the CTP ratio of 0.135 was not adequate for the algorithm with intra-class alignment. Furthermore, the high sensitivity to the CTP ratio made it difficult to find an adequate discard ratio. Accordingly, the 0.5 discard ratio (5495 particles) was concluded to be the most desirable for the present library. Notably, the mean value inspection (Fig. 9c-4) also supported the 5495 particle dataset as a better choice. The statistically-optimal mean values of FRC and CCC overlapped only at 5495 particles (orange box), whose FRC and CCC mean values were not-significantly different from the selected 7817 particles. For the CTP ratio variation, a similar result was obtained (Fig. 8c) compared to without intra-class alignment (Fig. 6c).

4. Discussion

4.1. IC-EOS

It is known that extremely noisy input micrographs can yield an artificially high-resolution with increased misalignments to noise (Penczek, 2010; Stewart and Grigorieff, 2004). To quantify the likelihood of including this artifact into the resultant class averages, not only the resolution but also four other energy functions of 2D averaging and their convergence stabilities were used. The "majority voting" approach softened the noise influence of each individual base measurement and effectively reduced the number of search trials in selecting the best condition set. To organize the findings from the IC-EOS experiment, a decision scheme for workflow optimization was devised (Fig. 10), consisting of three major components: (1) properties of the input/output datasets, (2) workflow, and (3) evaluation system. Among them, properties of datasets and workflow have been well established in SPA, but a quantitative evaluation system has not yet been in wide use. The construction of an evaluation system having a set of quantitative measurements and criteria was the aim for IC-EOS.

4.2. Detection of convergence inhibitor factors

As shown by the present study, IC-EOS can detect three "convergence-inhibitor factors": (1) data, (2) conditions, and (3) processes. First, the non-true-particle images were inhibitor data, which did not match the pre-conditions of the workflow. IC-EOS successfully determined the most appropriate value for the particle discard ratio. Second, an inappropriate number of classes (i.e. particle 3D orientations) was an inhibitor condition, which did not fit the nature of the dataset. The adequate CTP ratio was successfully determined with IC-EOS. Third, the sudden overall shift due to mutual alignments of class averages was the inhibitor process for NaChBac. The orientational constraint by the particle uprighting algorithm was effective in enhancing the convergence of the 2D averaging. The investigation of mismatches between the pre-conditions of the process and the nature of the dataset led to finding the cause of the condition sensitivity of the workflow (Fig. 3c and 3d, and subsection 3.1.1 of Part II). Thus, the system can aid in finding inhibitor processes for the optimization of workflow.

4.3. Optimizations of 2D averaging

The particle uprighting algorithm successfully enhanced the convergence of base measurements by preventing any drastic change in orientation through whole cycles. During this first optimization, three issues were addressed; (1) the appearance of the total average should be stable through all iterations, (2) its orientation and position should also be stable, and (3) the alignment update should occur only between most-similar class average pairs. To maintain a stable appearance, all class averages were aligned relative to the total average before their mutual alignment. The total average should be stable during iterations, since it is the most global statistical feature of the particle library (Penczek et al., 1992). The CCC between the total average and each class average was used to upright particles at the center of the image square. This CCC was also used to ensure that alignment updates occur with only most-similar class average pairs. However, this uprighting algorithm was designed for particles whose projections have a wide variety in shape and size. Caution has to be taken when applying this algorithm to any particle which does not fit this assumption, such as round particles, including icosahedral viruses.

In the second optimization, the addition of intra-class alignment increased the condition-sensitivity of the workflow convergence and was not helpful for finding the adequate parameter range. However, resolutions of the averages were improved. Therefore, this algorithm should be included after narrowing the condition range but not in the condition-search stage.

4.4. Selection of constant ratios

Since the dependence relationship between the CTP ratio and the discard ratio was not clear, independent experiments were designed for variations of these two ratios in the NaChBac study, and constants were chosen based on experience and visual inspection following our previous method (Venturi et al., 2011). For the discard ratio variation, a constant CTP ratio of 0.135 was chosen as an effective number with the size of the present library. The result showed that this CTP ratio was not substantially different from the most appropriate CTP ratio of 0.165 elected in the CTP ratio variation (Fig. 6c, highlighted with yellow), and supported the validity of our empirical method for deciding the number of classes. The library of the 0.50 discard ratio was chosen in the CTP ratio variation, since no non-true-particle images were found in this library by visual inspection. The best condition election with discard ratio variation selected the same library (Fig. 7c, highlighted with yellow), showing consistency with our visual inspection method.

4.5. FRC-resolution relative to discard and CTP ratios

Among the five base measurements used in the present study, FRC-resolution is the most well-balanced because it reflects two important aspects of class averages: self-consistency and SNR (Penczek, 2010). The particle discarding experiment justifies the particle selection. Highly flexible domains of ion channels and the variation of the attached lipids to the hydrophobic transmembrane helices can cause image variation (Maruyama et al., 2007). The increase in the discard ratio may reduce such variations, resulting in the improvement of resolution.

The saturation point offers key information for finding the optimal condition. The FRC mean value was saturated at high discard ratios while it was linearly improved by decreasing CTP ratio. Due to the limited conditions in the present study, this improvement was not bounded by any clear saturation. By contrast, its mean fluctuation curve showed clear plateaus. Therefore, fluctuation assessment is useful in balancing between optimal achievement level and convergence stability.

4.6. CCC-similarity tradeoff relationship with resolution

CCC-similarity and FRC-resolution mean values were in a tradeoff relationship when the CTP ratio was varied. This relationship made it difficult to find the most-plausible CTP ratio using the mean value of FRC-resolution alone. The fluctuations of base measurements were again useful in this matter, because these fluctuations indicated a CTP ratio for the most stable convergence. On the other hand, CCC-similarity and FRC-resolution mean values improved together when the discard ratio was increased. However, the CCC-similarity mean value became increasingly unstable, judging from its mean fluctuation curve. Therefore, a higher discard ratio itself does not mean closer-to-optimal achievement levels. Since multiple objectives interact with each other in relationship with various parameters, our approach of considering 2D averaging as a "complex system" allows IC-EOS to select the most appropriate condition set.

4.7. TIC-variance interpretation

With all the workflows and conditions used in the present study, TIC-variance mean value was the most stable base measurement. However, the interpretation of TIC-variance mean value is not straightforward, since TIC-variance is the sum of signal variance and noise variance from the "additive noise" model (Frank et al., 1981). The FRC mean value reflecting the SNR of class averages should be considered alongside the TIC-variance mean value. The most widely used definition of SNR in SPA is the signal-variance per noise-variance ratio (Frank, 2006; Frank et al., 1981), which can be improved either by relatively reducing noise variance or relatively increasing signal variance. Therefore, the information from other base measurements, such as resolution or SNR, is necessary to understand the trend of the TIC-variance mean value in detail.

5. Conclusion

The case study using NaChBac demonstrated that the numerical computation method of IC-EOS calculating real data is informative to find the characteristics and behaviors of workflows for 2D averaging. Due to the complicated process of image formation in EM for SPA, establishing a noise generation model for input particle images has been challenging for analytical methods to predict the outputs of SPA computations. These results revealed that a characteristic of the workflow with particle uprighting is the ability to stabilize the convergence of base measurements more effectively than does the workflow without this algorithm or with the addition of intra-class alignment. IC-EOS led to this finding by enabling the rating of workflow optimization. The tradeoff behavior of the 2D averaging workflows between resolution and CCC became apparent during the search for the proper number of classes. Even with this complication, IC-EOS still could determine the best number of classes where these two base measurements are well balanced. IC-EOS also showed the behavior of the 2D averaging workflows during particle selection; resolution and CCC improved together with stricter selections, although saturation was possible.

Table 1. Particle selection and class counts

a. Discard ratio variation with constant class counts

Particles	Classes	Members	CTP Ratio	Discard Ratio
10947	200	54.74	0.135	0.00
8713	200	43.57	0.152	0.20
6848	200	34.24	0.171	0.37
5495	200	27.48	0.191	0.50

b. CTP ratio variation with constant particle counts

Particles	Classes	Members	CTP Ratio	Discard Ratio
5495	100	54.95	0.135	0.50
5495	200	27.48	0.191	0.50
5495	300	18.32	0.234	0.50
5495	400	13.74	0.270	0.50

Table 2. 2D Averaging conditions for 1st and 2nd optimizations

a. CTP ratio variation with constant discard ratio

Particles	Classes	Members	CTP Ratio	Discard Ratio
5495	25	219.80	0.067	0.50
5495	50	109.90	0.095	0.50
5495	100	54.95	0.135	0.50
5495	150	36.63	0.165	0.50
5495	200	27.48	0.191	0.50
5495	300	18.32	0.234	0.50
5495	400	13.74	0.270	0.50

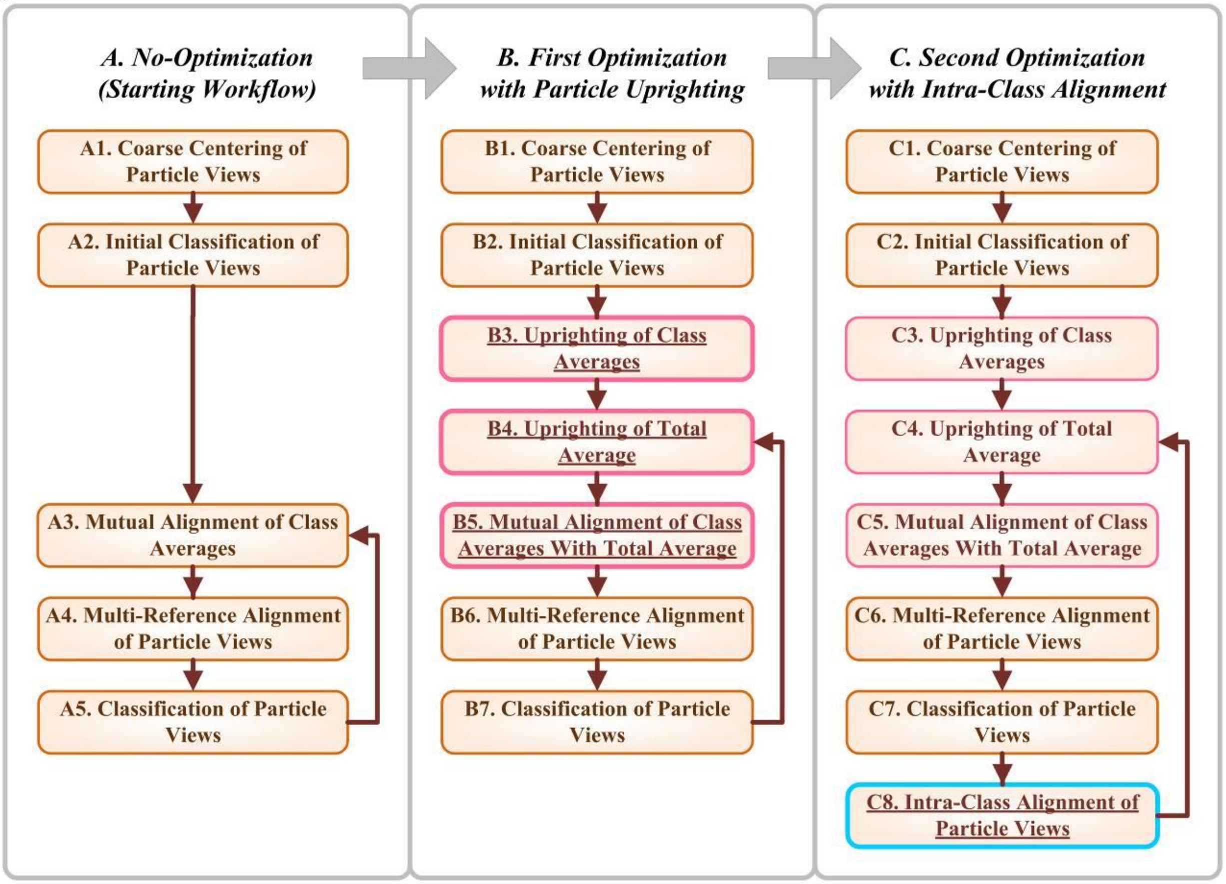
b. Discard ratio variation with constant CTP ratio (variable class counts)

Particles	Classes	Members	CTP Ratio	Discard Ratio
10947	200	54.74	0.135	0.00
9843	180	54.68	0.135	0.10
8713	159	54.80	0.135	0.20
7817	143	54.66	0.135	0.29
6848	125	54.78	0.135	0.37
5495	100	54.95	0.135	0.50
4373	80	54.66	0.135	0.60
3272	60	54.53	0.135	0.70

Figure 2. Optimization of 2D averaging workflow

(a) Three workflows used to test the optimization effect on 2D averaging. The left column shows the starting workflow to be optimized. The first approach (middle column) introduced the "particle uprighting algorithm" (Fig. 2b), expecting to stabilize the orientation and position of aligned particle views through iterations by forcing particle views to be upright and centered at each iteration. This aimed to achieve more robust convergence in 2D averaging. Additions (step B3 and B4) and modifications (from step A3 to B5) of steps are indicated by underlining. The second optimization (right column) further using "intra-class alignment" was expected to reduce the bias due to references used for MRA, enabling finer translational and rotational alignment. The additional step is underlined (step C8). (b) Flow chart of the particle uprighting introduced in the first optimization. This algorithm starts from an average, such as a class average or total average of particles. (1) First, create a binary version of the average. (2) Find the contour of the binary image using the "findContours" function of OpenCV (Open Source Computer Vision Library). (3) Fit an ellipse to the contour by the least-squares method using the "fitEllipse" function of OpenCV. (4) The image contour and the fitted ellipse are uprighted so that the major and minor axes are vertical and horizontal, respectively, the intersection of axes is centered in the image square, and the area surrounded by the contour below the horizontal axis is larger than that above. (5) The third and fourth steps are repeated five times or until the contour is stabilized. (6) Finally, the input image is rotationally and translationally aligned using the resultant values of alignment for the image contour.

a


A. No-Optimization (Starting Workflow)
A1. Coarse Centering of Particle Views
A2. Initial Classification of Particle Views
A3. Mutual Alignment of Class Averages
A4. Multi-Reference Alignment of Particle Views
A5. Classification of Particle Views
B. First Optimization with Particle Uprighting
B1. Coarse Centering of Particle Views
B2. Initial Classification of Particle Views
B3. Uprighting of Class Averages
B4. Uprighting of Total Average
B5. Mutual Alignment of Class Averages With Total Average
B6. Multi-Reference Alignment of Particle Views
B7. Classification of Particle Views
C. Second Optimization with Intra-Class Alignment
C1. Coarse Centering of Particle Views
C2. Initial Classification of Particle Views
C3. Uprighting of Class Averages
C4. Uprighting of Total Average
C5. Mutual Alignment of Class Averages With Total Average
C6. Multi-Reference Alignment of Particle Views
C7. Classification of Particle Views
C8. Intra-Class Alignment of Particle Views

b

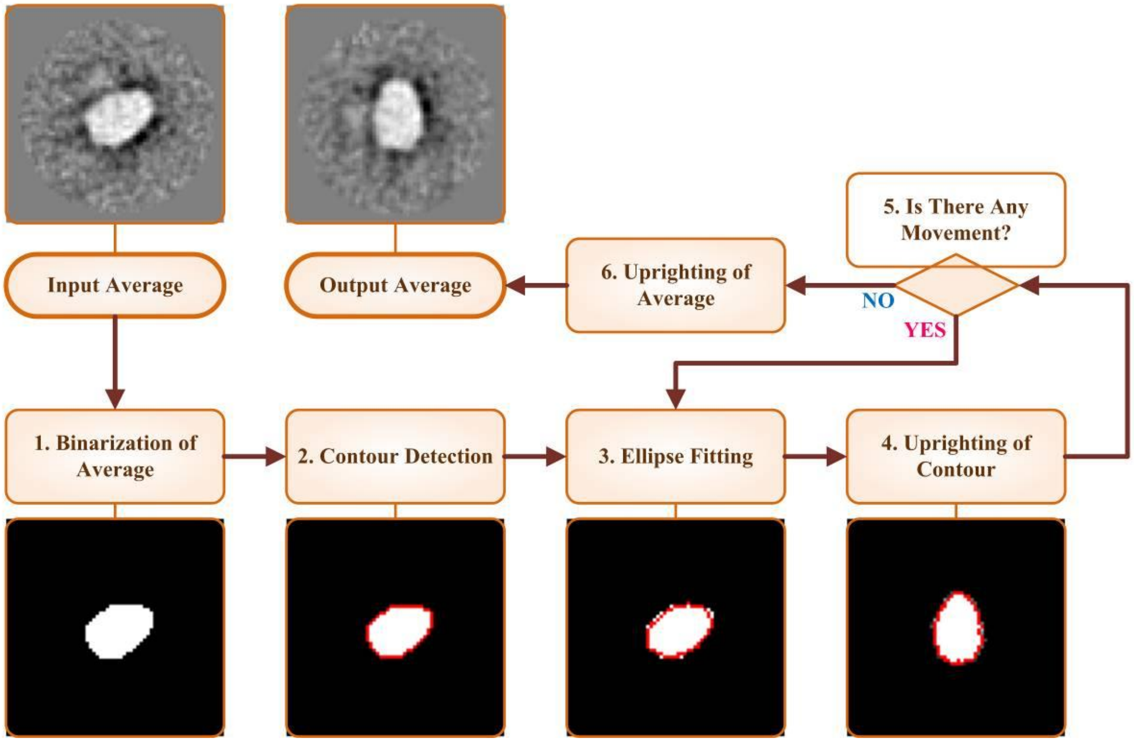
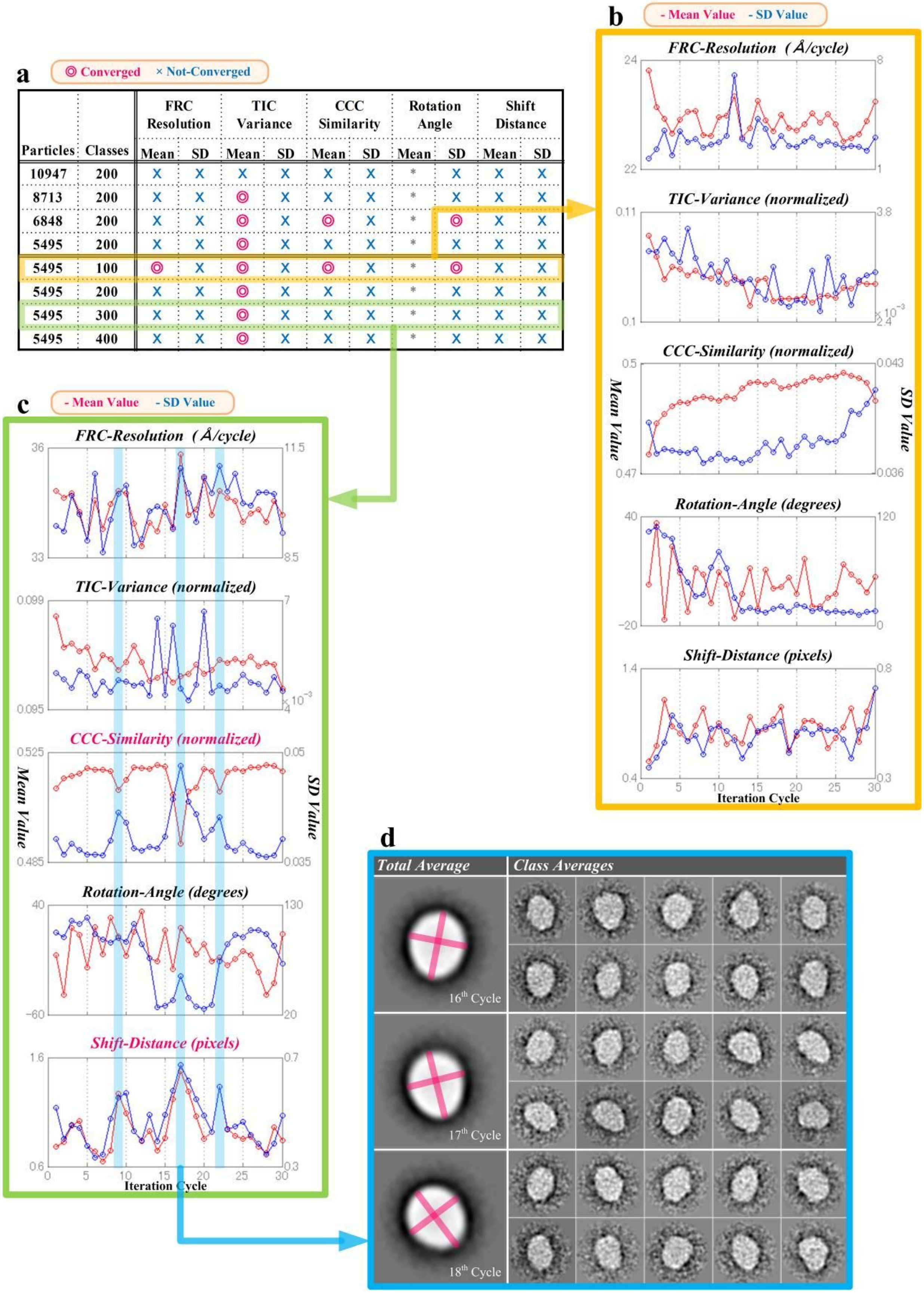

Input Average
Output Average
6. Uprighting of Average
5. Is There Any Movement?
NO
YES
1. Binarization of Average
2. Contour Detection
3. Ellipse Fitting
4. Uprighting of Contour

Figure 3. Workflow without optimization

Using NaChBac particles. (a) Convergence evaluation using the procedure defined in Figure 1. Red circles represent "converged", blue crosses "not-converged". All base measurements, except TIC-variance mean value, did not converge for most of the condition sets. The condition set of 5495 particles classified into 100 classes (indicated by orange outlines) was the most stable because several base measurements, including FRC-resolution, were "converged". (b) History curves of mean values (red line) and SD values (blue line) for the condition set of 5495 particles with 100 classes. The decision of "converged" or "not-converged" by IC-EOS was identical to the visual impression. (c) History curves of the 5495 particle dataset with 300 classes are presented as examples of "not-converged" (Fig. 3a, green box). At cycle numbers 9, 17, and 22 (highlighted with vertical blue lines), mean value of CCC-similarity dropped drastically. Mean value of shift-distance increased at the corresponding cycles, reflecting instability in the overall orientation and position of particle views. (d) Total averages and class averages at cycle numbers 16 to 18, during which mean value of CCC-similarity dropped. The orientations of the total average and class averages were not consistent through the cycles. These inconsistencies can interrupt the convergence of this workflow.



◎ Converged × Not-Converged

		FRC Resolution		TIC Variance		CCC Similarity		Rotation Angle		Shift Distance	
Particles	Classes	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
10947	200	×	×	×	×	×	×	*	×	×	×
8713	200	×	×	◎	×	×	×	*	×	×	×
6848	200	×	×	◎	×	◎	×	*	◎	×	×
5495	200	×	×	◎	×	×	×	*	×	×	×
5495	100	◎	×	◎	×	◎	×	*	◎	×	×
5495	200	×	×	◎	×	×	×	*	×	×	×
5495	300	×	×	◎	×	×	×	*	×	×	×
5495	400	×	×	◎	×	×	×	*	×	×	×

Figure 4. 2D averages of NaChBac created by the unoptimized workflow

Final 2D averages of the NaChBac particles using the unoptimized workflow with the most stable condition set (5495 particles and 100 classes). Four class averages are shown in the top row, with the constituent raw particles below. The constituent particles were highly consistent in size and shape with the 2D averages, indicating successful 2D averaging. However, it was not easy to find a convergence condition for this workflow because of its high sensitivity to condition variation.

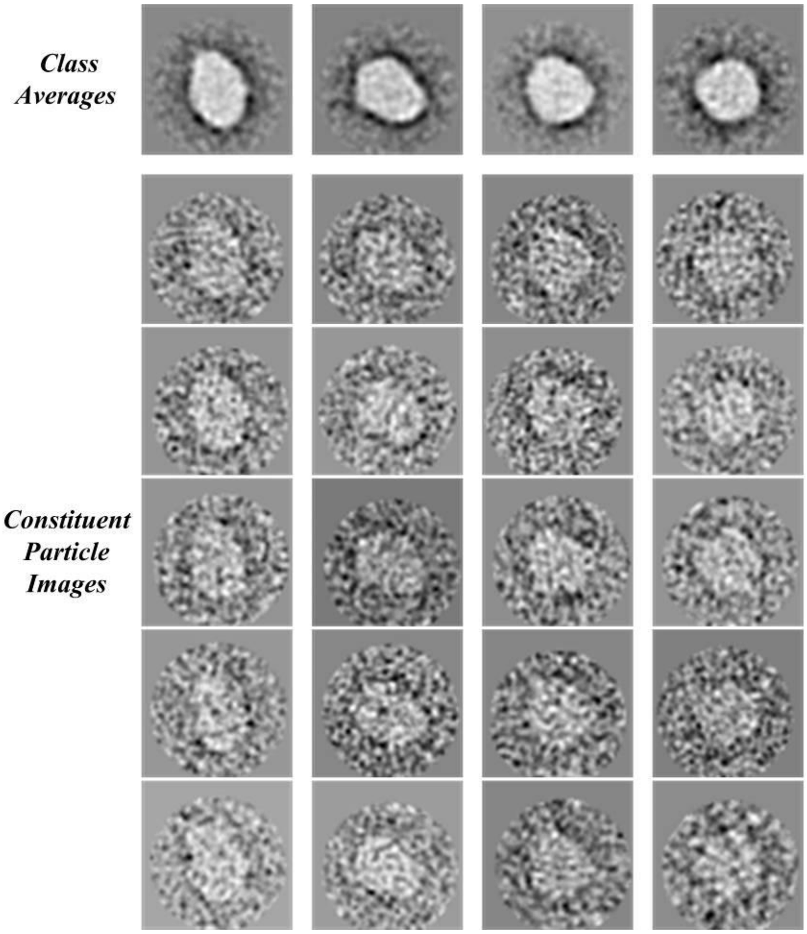
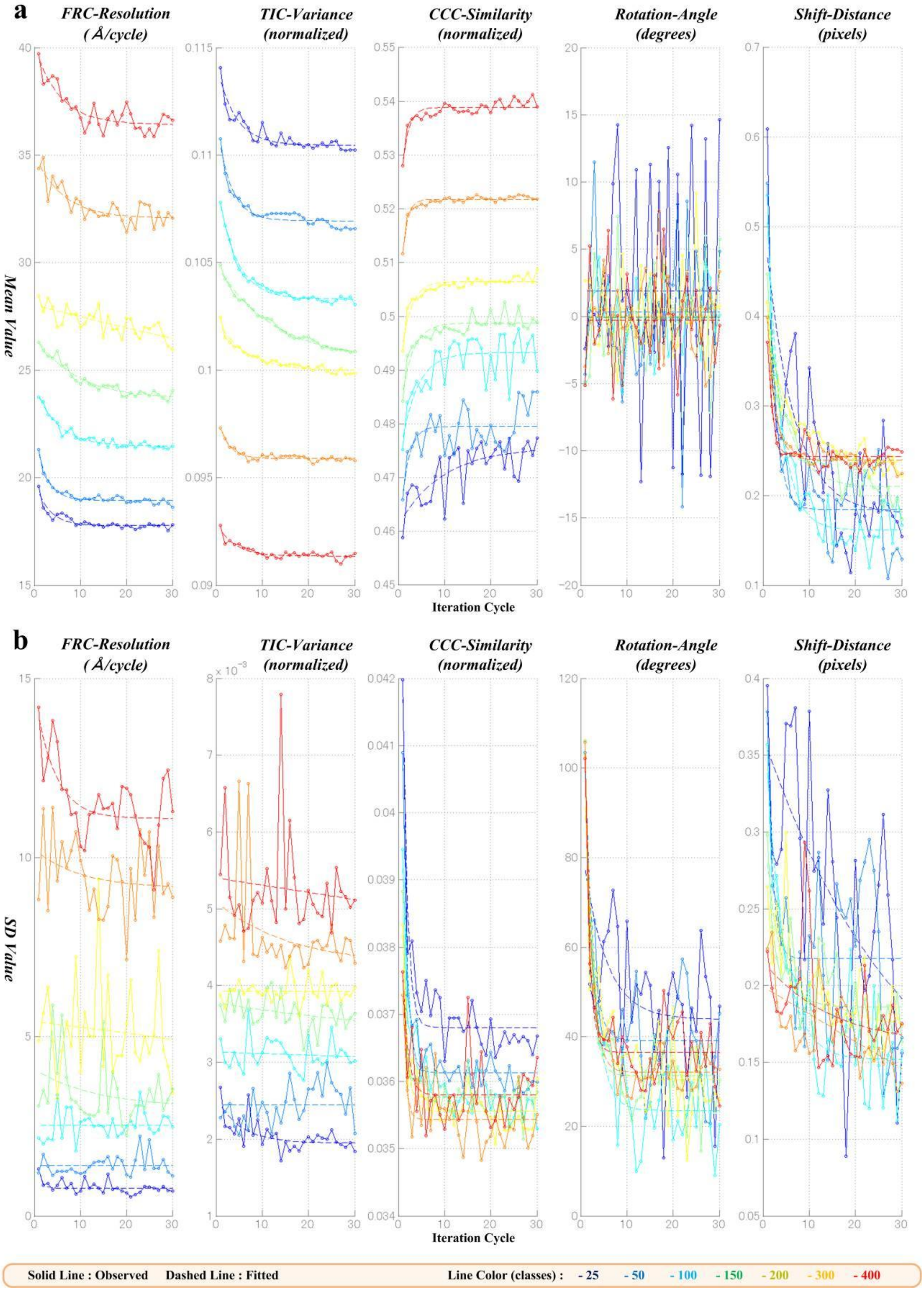

Class Averages
Constituent Particle Images

Figure 5. Five base measurements during the workflow with particle-uprighting and CTP ratio variation in the NaChBac study

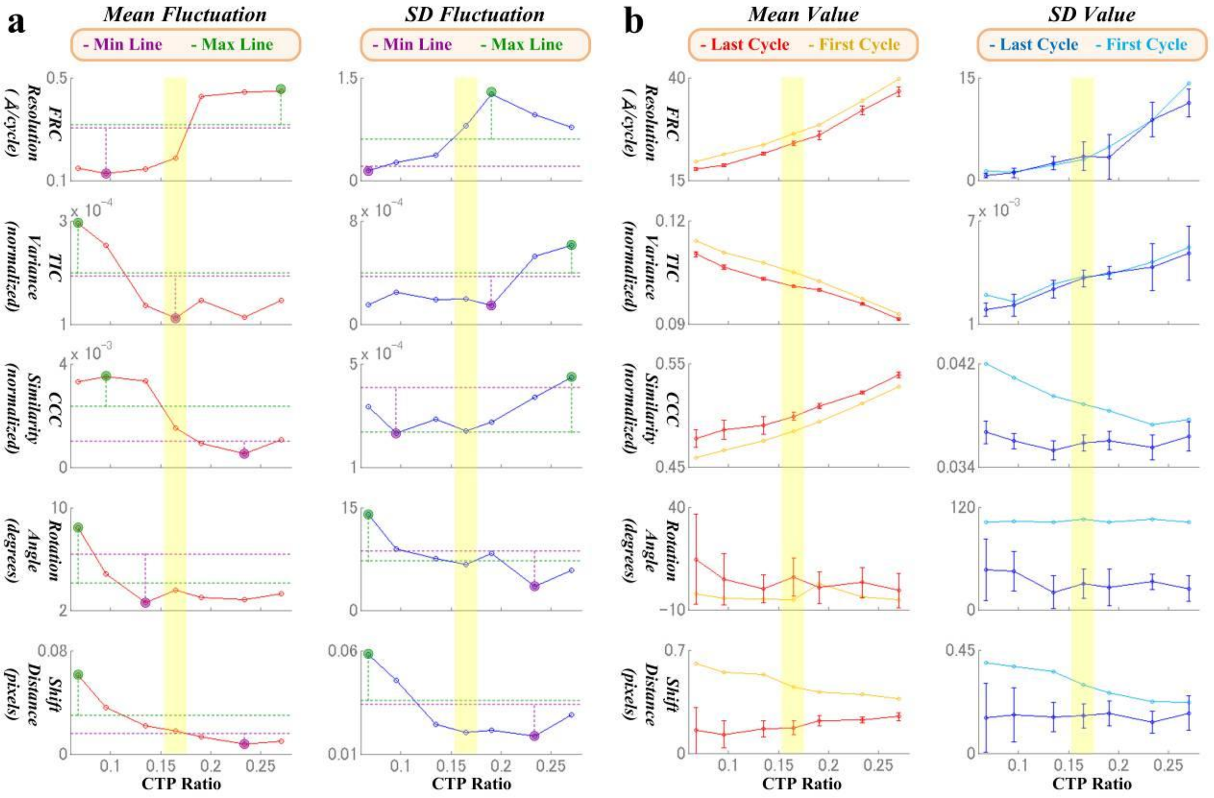
In each graph, a curve corresponds to a history of a base measurement using a particular CTP ratio, which increases as the number of classes increases from 25 (blue) to 400 (red). (a) History curves of mean values, showing convergence trends. CTP ratio influenced the mean values; fluctuation of the value curves helped to visually measure the convergence stability. For example, FRC-resolution increased (degraded or got worse) as the CTP ratio increased. At the same time, its fluctuation amplitude also increased (i.e. curve became rougher), while CCC-similarity increased (improved) and its fluctuation amplitude decreased (i.e. curve became smoother). (b) History curves of SD values. These curves indicate relatively large fluctuations and less clear convergence compared with those of mean values



a
FRC-Resolution (Å/cycle)
TIC-Variance (normalized)
CCC-Similarity (normalized)
Rotation-Angle (degrees)
Shift-Distance (pixels)
Mean Value
Iteration Cycle
0
10
20
30
b
FRC-Resolution (Å/cycle)
TIC-Variance (normalized)
CCC-Similarity (normalized)
Rotation-Angle (degrees)
Shift-Distance (pixels)
SD Value
× 10⁻³
Iteration Cycle
Solid Line : Observed
Dashed Line : Fitted
Line Color (classes) :
- 25
- 50
- 100
- 150
- 200
- 300
- 400

Figure 6. Relationship between CTP ratio and five base measurements in 2D averaging with particle uprighting

(a) Fluctuations of base measurements relative to CTP ratio. The minimum (purple dotted line) and maximum lines (green dotted line) show the ranges of significantly-not-different conditions from the minimum and maximum fluctuations, respectively. In each figure, an increasing or decreasing trend with a plateau was exhibited. Mean fluctuations of FRC and CCC were somewhat sigmoidal. In the fluctuation sub-vote difference (Fig. 6c, 2nd section), a condition with a point lower than the minimum line got a positive vote, while one higher than the maximum line got a negative vote. (b) Means and SDs of base measurements relative to CTP ratio. The changes in the values of first and last cycles represent a variety of trends and improvement widths relative to CTP ratio variation. The mean values of FRC and CCC were in a tradeoff relationship. (c) The result of the best condition election. The condition of 150 classes was selected as the most plausible (0.165 CTP ratio; 3rd section, highlighted with yellow). Convergence sub-vote sums (1st section), fluctuation sub-vote differences (2nd section), and total votes (3rd section) were high around this CTP ratio. The mean value inspection indicated that the optimal FRC mean value was only at 25 classes (0.067 CTP ratio) and the optimal CCC was only at 400 (0.270 CTP ratio), reflecting the tradeoff relationship, and both values of the selected 150 classes were significantly-different from these optimal values (4th section).



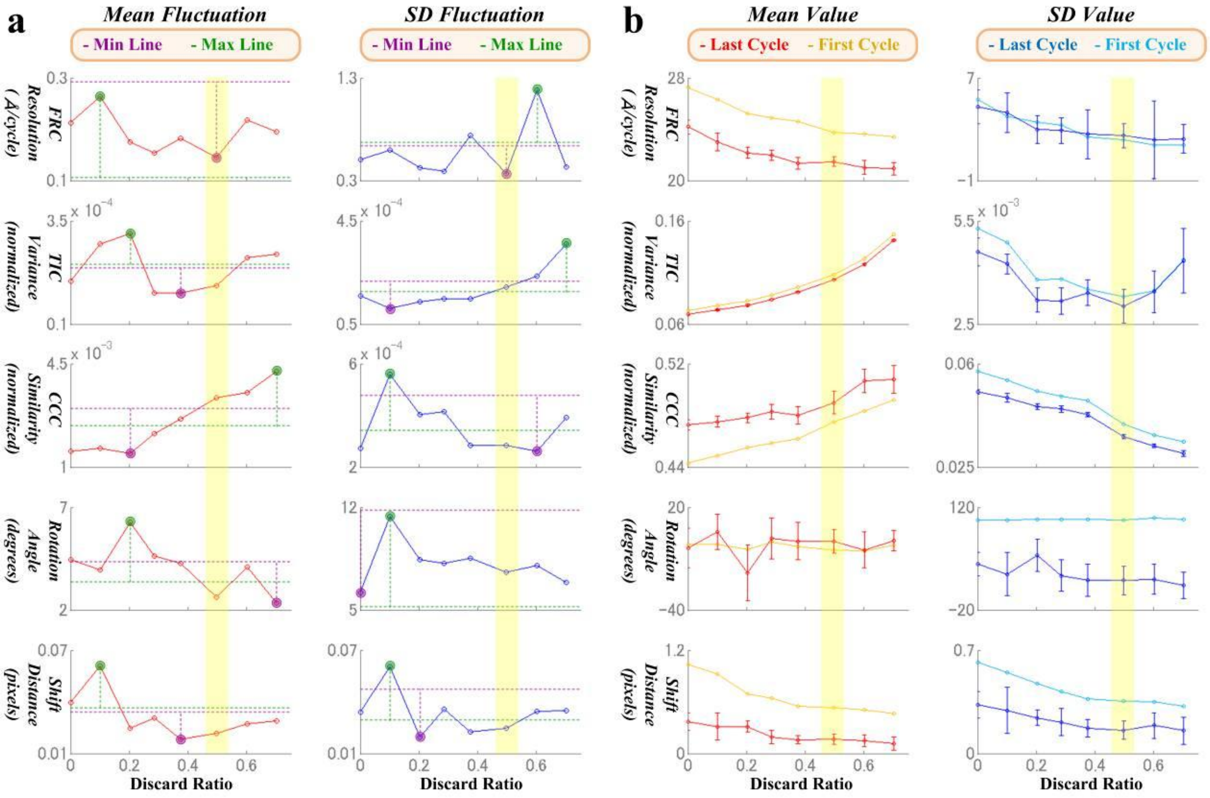
c

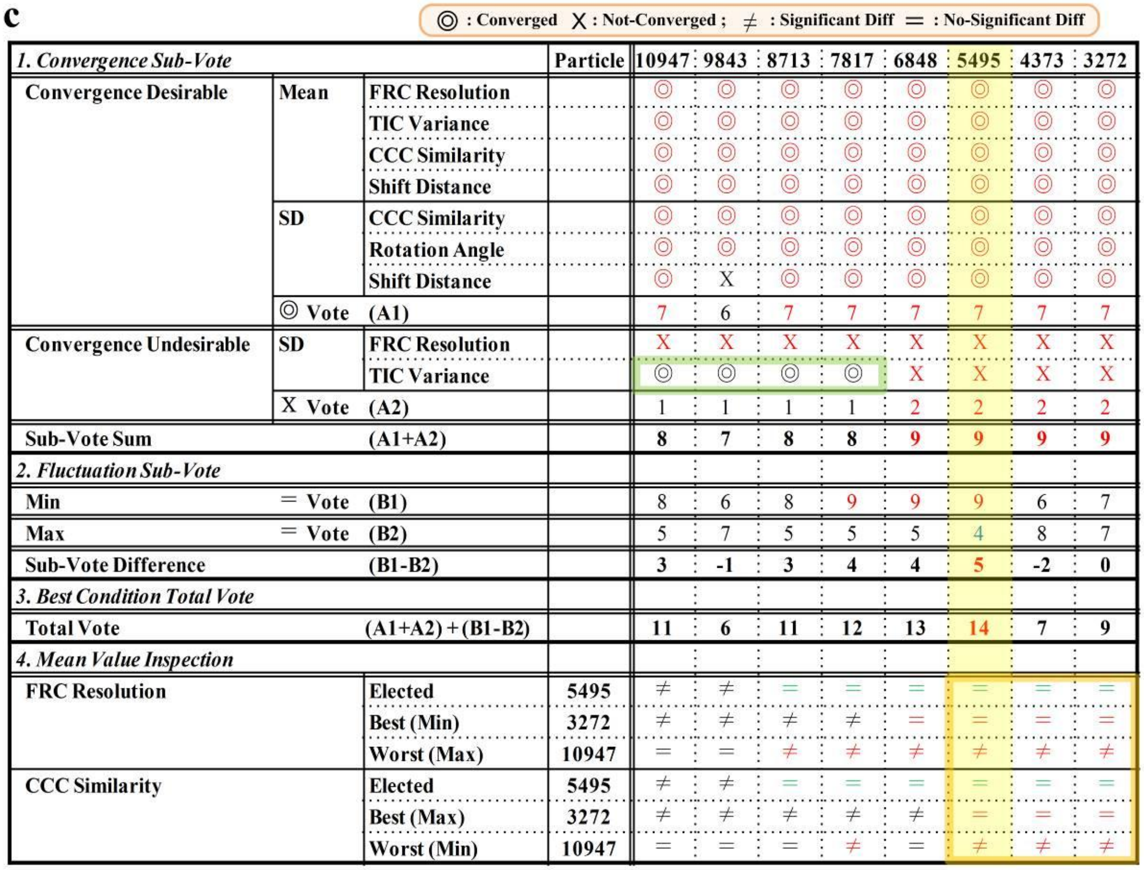
◎ : Converged X : Not-Converged ; ≠ : Significant Diff = : No-Significant Diff

1. Convergence Sub-Vote			Class	25	50	100	150	200	300	400
Convergence Desirable	Mean	FRC Resolution		◎	◎	◎	◎	◎	◎	◎
		TIC Variance		◎	◎	◎	◎	◎	◎	◎
		CCC Similarity		◎	◎	◎	◎	◎	◎	◎
		Shift Distance		◎	◎	◎	◎	◎	◎	◎
	SD	CCC Similarity		◎	◎	◎	◎	◎	◎	X
		Rotation Angle		X	◎	◎	◎	◎	◎	◎
		Shift Distance		◎	◎	◎	◎	X	◎	X
	◎ Vote	(A1)		6	7	7	7	6	7	5
Convergence Undesirable	SD	FRC Resolution		X	X	X	X	X	X	◎
		TIC Variance		◎	X	X	X	X	X	X
	X Vote	(A2)		1	2	2	2	2	2	1
Sub-Vote Sum		(A1+A2)		**7**	**9**	**9**	**9**	**8**	**9**	**6**
2. Fluctuation Sub-Vote										
Min	= Vote	(B1)		4	4	7	7	8	7	5
Max	= Vote	(B2)		7	6	3	2	4	4	4
Sub-Vote Difference		(B1-B2)		**-3**	**-2**	**4**	**5**	**4**	**3**	**1**
3. Best Condition Total Vote										
Total Vote		(A1+A2) + (B1-B2)		**4**	**7**	**13**	**14**	**12**	**12**	**7**
4. Mean Value Inspection										
FRC Resolution		Elected	150	≠	≠	≠	=	≠	≠	≠
		Best (Min)	25	=	≠	≠	≠	≠	≠	≠
		Worst (Max)	400	≠	≠	≠	≠	≠	≠	=
CCC Similarity		Elected	150	≠	=	=	=	≠	≠	≠
		Best (Max)	400	≠	≠	≠	≠	≠	≠	=
		Worst (Min)	25	=	=	=	≠	≠	≠	≠

Figure 7. Particle uprighting workflow with discard ratio variation in NaChBac study

(a) Relationship between base measurement fluctuations and discard ratio variation. The increasing/decreasing trends were less clear than those between fluctuations and CTP ratio variation (Fig. 6a). Plateaus were not clear except for mean fluctuation of CCC-similarity, reflecting that CCC became unstable at higher discard ratios. (b) Means and SDs of base measurements relative to the discard ratio. As discard ratio increased, mean values of FRC and CCC improved together, at a cost of CCC instability. (c) The best condition election selected 5495 particles for the best convergence (0.50 discard ratio; 3rd section, highlighted with yellow). Counts of convergence sub-vote sums (1st section), fluctuation sub-vote differences (2nd section), and total votes (3rd section) were higher around 6848 (0.37 discard ratio) and 5495 particles. According to the mean value inspection, the optimal values for FRC and CCC overlapped at 3272 and 4373 particles (0.7 and 0.6 discard ratio) as well as at the selected 5495 particles (4th section, orange box).





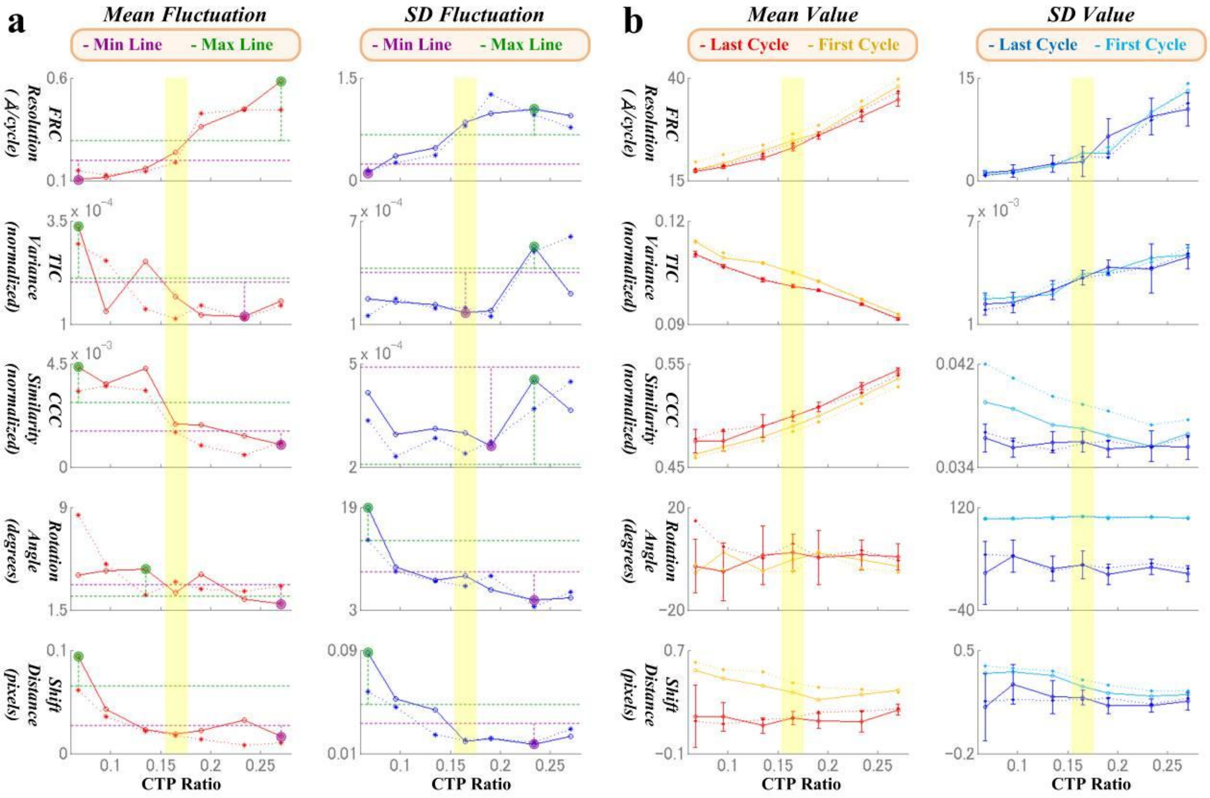
c

◎ : Converged X : Not-Converged ; ≠ : Significant Diff = : No-Significant Diff

1. Convergence Sub-Vote			Particle	10947	9843	8713	7817	6848	5495	4373	3272
Convergence Desirable	Mean	FRC Resolution		◎	◎	◎	◎	◎	◎	◎	◎
		TIC Variance		◎	◎	◎	◎	◎	◎	◎	◎
		CCC Similarity		◎	◎	◎	◎	◎	◎	◎	◎
		Shift Distance		◎	◎	◎	◎	◎	◎	◎	◎
	SD	CCC Similarity		◎	◎	◎	◎	◎	◎	◎	◎
		Rotation Angle		◎	◎	◎	◎	◎	◎	◎	◎
		Shift Distance		◎	X	◎	◎	◎	◎	◎	◎
	◎ Vote	(A1)		7	6	7	7	7	7	7	7
Convergence Undesirable	SD	FRC Resolution		X	X	X	X	X	X	X	X
		TIC Variance		◎	◎	◎	◎	X	X	X	X
	X Vote	(A2)		1	1	1	1	2	2	2	2
Sub-Vote Sum		(A1+A2)		8	7	8	8	9	9	9	9
2. Fluctuation Sub-Vote											
Min	= Vote	(B1)		8	6	8	9	9	9	6	7
Max	= Vote	(B2)		5	7	5	5	5	4	8	7
Sub-Vote Difference		(B1-B2)		3	-1	3	4	4	5	-2	0
3. Best Condition Total Vote											
Total Vote		(A1+A2) + (B1-B2)		11	6	11	12	13	14	7	9
4. Mean Value Inspection											
FRC Resolution		Elected	5495	≠	≠	=	=	=	=	=	=
		Best (Min)	3272	≠	≠	≠	≠	=	=	=	=
		Worst (Max)	10947	=	=	≠	≠	≠	≠	≠	≠
CCC Similarity		Elected	5495	≠	≠	=	=	=	=	=	=
		Best (Max)	3272	≠	≠	≠	≠	≠	=	=	=
		Worst (Min)	10947	=	=	=	≠	=	≠	≠	≠

Figure 8. Intra-class alignment workflow related to CTP ratio variations in NaChBac study

Comparative evaluation between the primary (without intra-class alignment; Fig. 6) and the secondary (with intra-class alignment) optimizations. (a) Fluctuations of base measurements. There were slight differences in the fluctuations with (solid line) and without (dotted line) intra-class alignment at various CTP ratios. (b) Means and SDs of base measurements in the primary (dotted line) and secondary (solid line) optimizations. In general, the improvement widths were narrowed by intra-class alignment. At all CTP ratios except 200 classes (0.191 CTP ratio), the FRC mean values with intra-class alignment were lower (better) than those without intra-class alignment. (c) The best condition election selected 150 classes (0.165 CTP ratio; 3rd section, highlighted with yellow), consistent with the result without intra-class alignment. The convergence sub-vote (1st section) exhibited more undesirable trends than that without intra-class alignment, indicating that intra-class alignment increased condition-sensitivity, as a cost of better FRC-resolution. The result of mean value inspection (4th section) was similar to that without intra-class alignment. This result shows that IC-EOS is effective for revealing differences in the characteristics and behaviors of these two workflows for 2D averaging.



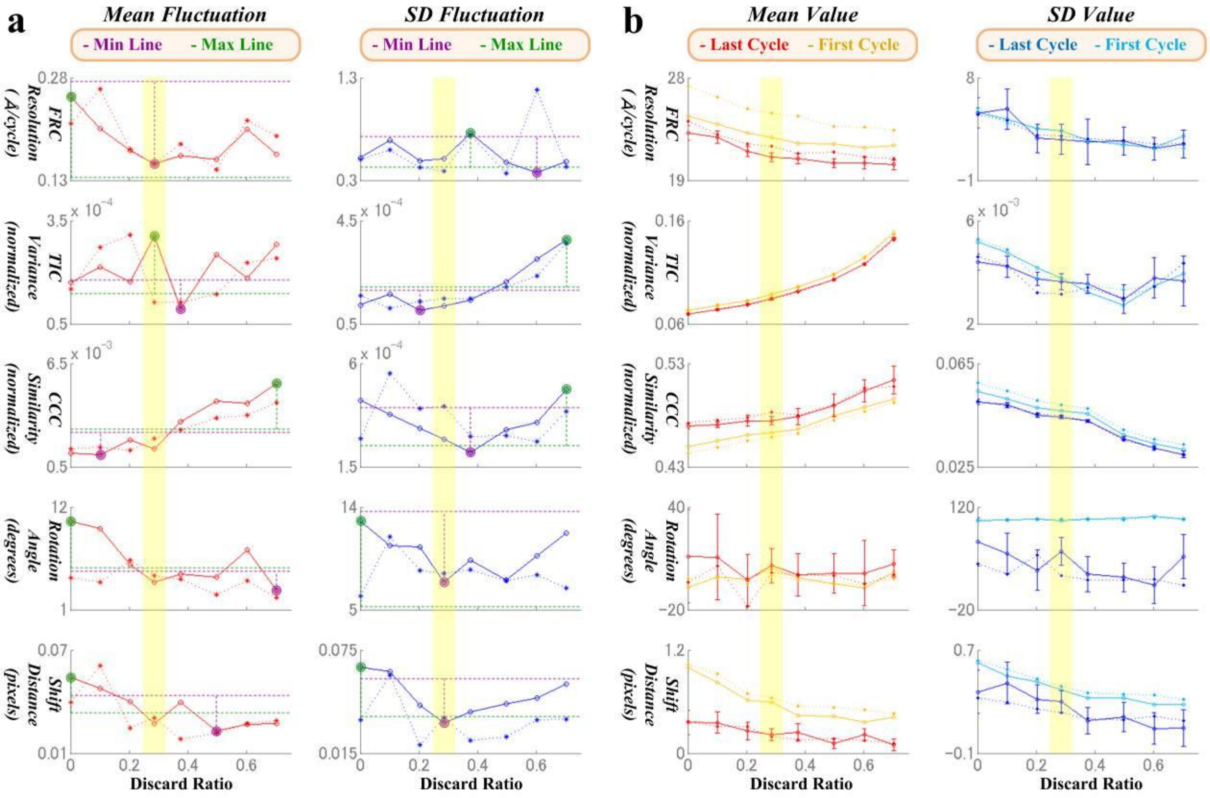
c

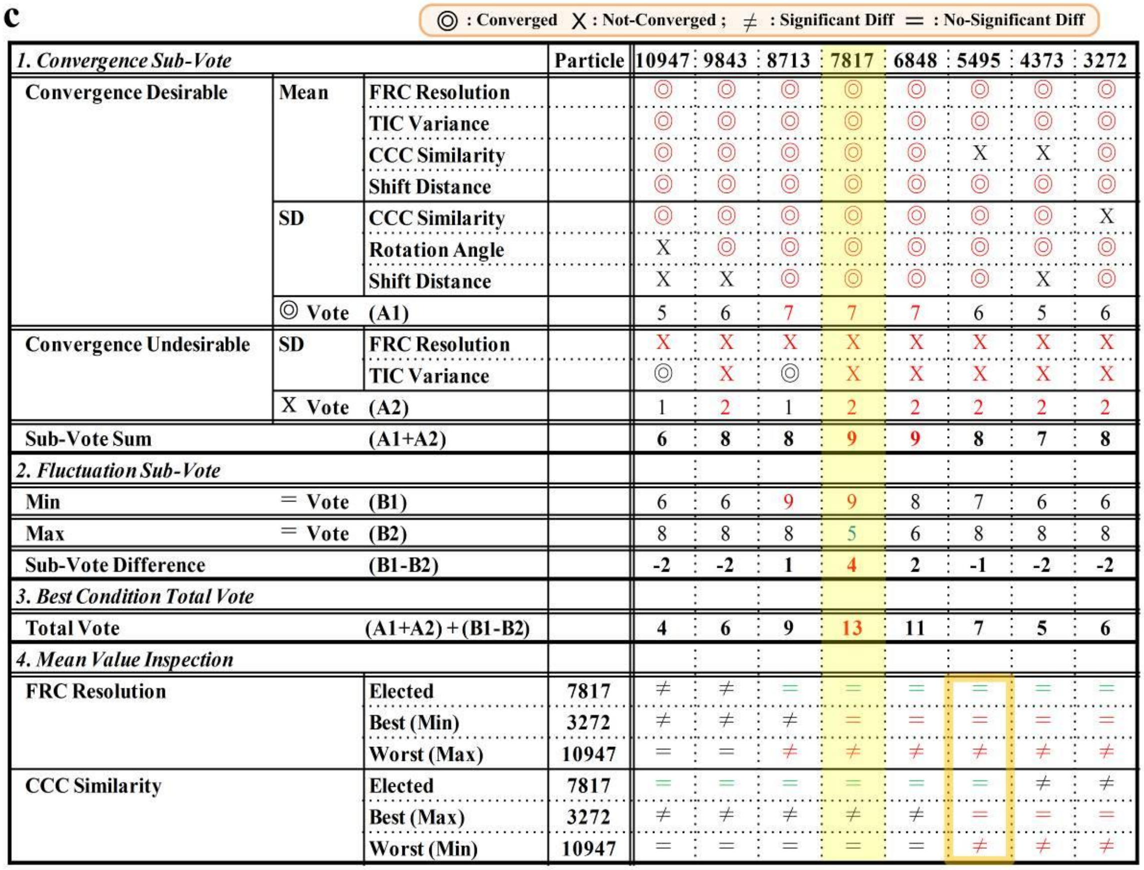
◎ : Converged X : Not-Converged ; ≠ : Significant Diff = : No-Significant Diff

1. Convergence Sub-Vote			Class	25	50	100	150	200	300	400
Convergence Desirable	**Mean**	**FRC Resolution**		◎	◎	◎	◎	X	◎	◎
		TIC Variance		◎	◎	◎	◎	◎	◎	◎
		CCC Similarity		X	X	X	◎	◎	◎	◎
		Shift Distance		X	◎	◎	◎	◎	X	◎
	SD	**CCC Similarity**		◎	◎	◎	◎	◎	X	X
		Rotation Angle		X	◎	◎	◎	◎	◎	◎
		Shift Distance		X	X	◎	◎	X	◎	X
	◎ **Vote**	**(A1)**		3	5	6	7	5	5	5
Convergence Undesirable	**SD**	**FRC Resolution**		X	X	X	X	X	X	X
		TIC Variance		X	X	X	X	X	X	X
	X **Vote**	**(A2)**		2	2	2	2	2	2	2
Sub-Vote Sum		**(A1+A2)**		**5**	**7**	**8**	**9**	**7**	**7**	**7**
2. Fluctuation Sub-Vote										
Min	= **Vote**	**(B1)**		4	4	5	7	6	6	8
Max	= **Vote**	**(B2)**		7	4	4	3	4	4	3
Sub-Vote Difference		**(B1-B2)**		**-3**	**0**	**1**	**4**	**2**	**2**	**5**
3. Best Condition Total Vote										
Total Vote		**(A1+A2) + (B1-B2)**		**2**	**7**	**9**	**13**	**9**	**9**	**12**
4. Mean Value Inspection										
FRC Resolution		**Elected**	**150**	≠	≠	≠	=	≠	≠	≠
		Best (Min)	**25**	=	≠	≠	≠	≠	≠	≠
		Worst (Max)	**400**	≠	≠	≠	≠	≠	≠	=
CCC Similarity		**Elected**	**150**	≠	≠	=	=	=	≠	≠
		Best (Max)	**400**	≠	≠	≠	≠	≠	≠	=
		Worst (Min)	**50**	=	=	=	≠	≠	≠	≠

Figure 9. Intra-class alignment with discard ratio variation in NaChBac study

Differences of the secondary optimization from the primary (Fig. 7). (a) Fluctuations of base measurements with (solid line) and without (dotted line) intra-class alignment. Again, the fluctuations in relationship with discard ratio variation were similar regardless of this algorithm. (b) Means and SDs of base measurements with (solid line) and without (dotted line) intra-class alignment. In general, the improvement widths in the particle selection experiment were narrowed by intra-class alignment. FRC mean value with intra-class alignment was lower (better) than that without intra-class alignment. (c) The convergence vote sub-table with intra-class alignment (1st section) showed more undesirable trends as compared to the sub-table of without intra-class alignment; indicating that intra-class alignment increased condition-sensitivity, as a cost of better FRC. The narrowing of the improvement width might have caused this sensitivity, since one criterion of the convergence evaluation states that the improvement width must be significantly larger than the fluctuation. The best condition election selected 7817 particles (0.29 discard ratio; 3rd section, highlighted with yellow), which was not consistent with the result without intra-class alignment. However, in the mean value inspection (4th section), the optimal values for FRC and CCC overlapped only at 5495 particles (0.50 discard ratio; orange box), whose FRC and CCC mean values were not-significantly different from the selected 7817 particles.



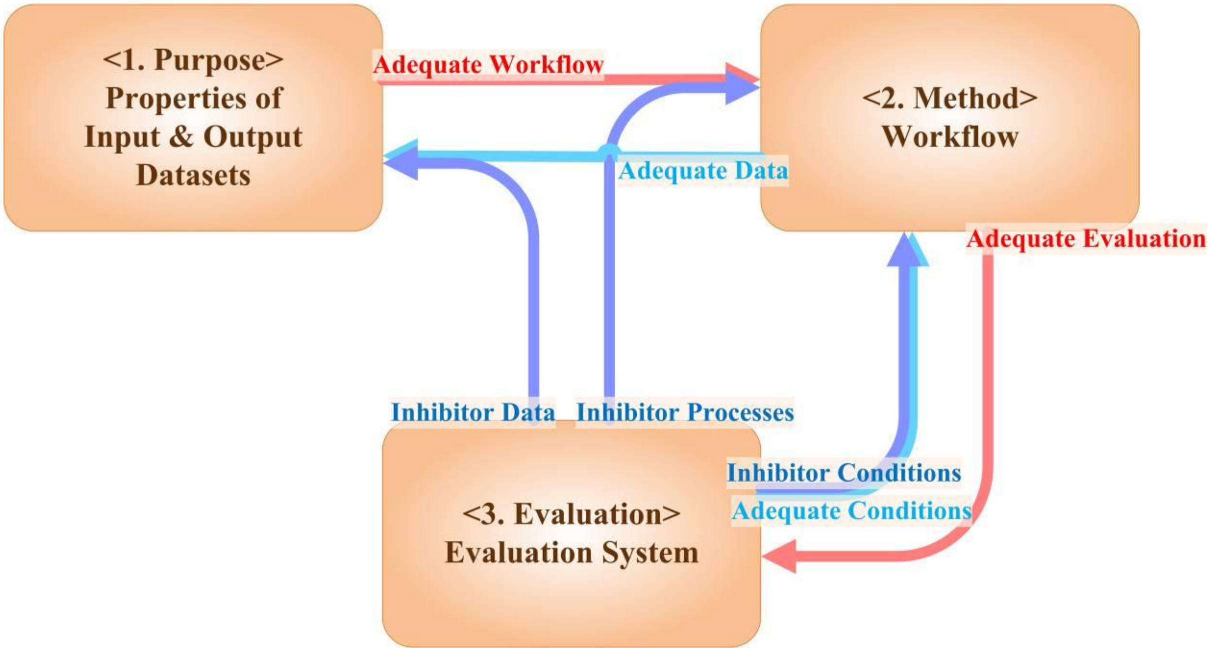


◎ : Converged X : Not-Converged ; ≠ : Significant Diff = : No-Significant Diff

1. Convergence Sub-Vote			Particle	10947	9843	8713	7817	6848	5495	4373	3272
Convergence Desirable	**Mean**	**FRC Resolution**		◎	◎	◎	◎	◎	◎	◎	◎
		TIC Variance		◎	◎	◎	◎	◎	◎	◎	◎
		CCC Similarity		◎	◎	◎	◎	◎	X	X	◎
		Shift Distance		◎	◎	◎	◎	◎	◎	◎	◎
	SD	**CCC Similarity**		◎	◎	◎	◎	◎	◎	◎	X
		Rotation Angle		X	◎	◎	◎	◎	◎	◎	◎
		Shift Distance		X	X	◎	◎	◎	◎	X	◎
	◎ **Vote**	**(A1)**		5	6	7	7	7	6	5	6
Convergence Undesirable	**SD**	**FRC Resolution**		X	X	X	X	X	X	X	X
		TIC Variance		◎	X	◎	X	X	X	X	X
	X **Vote**	**(A2)**		1	2	1	2	2	2	2	2
Sub-Vote Sum		**(A1+A2)**		**6**	**8**	**8**	**9**	**9**	**8**	**7**	**8**
2. Fluctuation Sub-Vote											
Min	= **Vote**	**(B1)**		6	6	9	9	8	7	6	6
Max	= **Vote**	**(B2)**		8	8	8	5	6	8	8	8
Sub-Vote Difference		**(B1-B2)**		**-2**	**-2**	**1**	**4**	**2**	**-1**	**-2**	**-2**
3. Best Condition Total Vote											
Total Vote		**(A1+A2) + (B1-B2)**		**4**	**6**	**9**	**13**	**11**	**7**	**5**	**6**
4. Mean Value Inspection											
FRC Resolution		**Elected**	**7817**	≠	≠	=	=	=	=	=	=
		Best (Min)	**3272**	≠	≠	≠	=	=	=	=	=
		Worst (Max)	**10947**	=	=	≠	≠	≠	≠	≠	≠
CCC Similarity		**Elected**	**7817**	=	=	=	=	=	=	≠	≠
		Best (Max)	**3272**	≠	≠	≠	≠	≠	=	=	=
		Worst (Min)	**10947**	=	=	=	=	=	≠	≠	≠

Figure 10. IC-EOS decision scheme for workflow optimization

The IC-EOS decision scheme for workflow optimization. Each component requires another to be adequate to its precondition (red and cyan arrows). The violations of these preconditions are considered as convergence-inhibitor factors (purple arrow): data, conditions, and processes. The construction of an evaluation system having a set of quantitative measurements and criteria was the aim for IC-EOS.


<1. Purpose>
Properties of
Input & Output
Datasets
Adequate Workflow
<2. Method>
Workflow
Adequate Data
Adequate Evaluation
Inhibitor Data
Inhibitor Processes
Inhibitor Conditions
Adequate Conditions
<3. Evaluation>
Evaluation System

Part III: Simulation study of IC-EOS using GroEL

1. Introduction

In Part II, the case study showed that IC-EOS could determine the best number of classes and selected particle images for the NaChBac dataset. However, an investigation of selection accuracy was still necessary. IC-EOS also revealed trends of fluctuations, means, and SDs of five base measurements relative to the variation of CTP ratio and of discard ratio. Under the assumption that clearer trends of these convergence measurements could be observed and that IC-EOS would select a correct condition, the 2D averaging of simulated data was conducted using the workflow of the first optimization with particle uprighting (Fig. 2a-B). Projection images were artificially generated from an atomic-coordinate data of the GroEL chaperon protein, which was obtained from the Protein Data Bank (PDB) web site (PDB, 2013).

2. Materials and Methods

First, 82 equally distributed projection images were generated from the crystal structure of the GroEL protein (wild-type, apo-GroEL; PDB code: 1XCK, Bartolucci et al., 2005), assuming D7 (7-fold x 2) symmetry (Fig. 11, the leftmost column). Then, 4100 particle images were obtained by creating 50 copies from each projection with random rotations (uniform distribution) and shifts (Gaussian distribution with zero-mean and 4 pixel SD). The dimensions of projection images were 80 x 80 pixels with 4 Å/pixel. The SNR (ratio of signal variance per noise variance) of particle images was set to 0.05 (=1049/20982) by adding zero-mean Gaussian noise (Fig. 11, the middle three columns). To simulate the discard ratio variation, pure noise images of zero-mean Gaussian with a variance of 20982 were added to the library of 4100 particles. To mimic the case of discarding the true-particle images from the library, the number of particles was also reduced from the original 4100 particles by excluding all copied particles of certain projections, based on CCC of projection images relative to their total averages. It was expected that libraries of 4100 particles or fewer would result in almost the same resolution, similarity, and stability. A band pass filter was applied to all generated libraries, with a 288.0 Å low-frequency cutoff and a 10.0 Å high-frequency cutoff at each cycle.

3. Results

The selection accuracy of the discard and CTP ratios was evaluated by simulation analysis using projections created from the crystal structure of GroEL. The condition sets are shown in Tables 3a and 3b. The iterative steps of the workflow were repeated for 30 cycles. The starting dataset of 4100 particles and 82 classes was defined as the correct answer with a zero discard ratio. The negative values of the discard ratio correspond to the ratios of additional noise images in the library, and the positive values correspond to the ratios of discarded true-particle images from the correct answer.

The generated 2D averages fitted well with the original projections (Fig. 11). The best condition election correctly selected 82 classes (0.141 CTP ratio; highlighted with yellow) in CTP ratio variation (Fig. 12c). This result supports that my proposed method is suitable for finding the correct CTP ratio for 2D averaging. The tradeoff relationship between mean values of FRC-resolution and CCC-similarity was observed again (Fig. 12b). The mean and SD fluctuations of

FRC showed clear trends in that the CTP ratios around the correct answer were the most stable (Fig. 12a).

On the other hand, the election selected a library of 3700 particles (0.1 discard ratio; highlighted with yellow) for discard ratio variation (Fig. 13c). This number was 10% fewer than the starting 4100 particles (0.0 discard ratio). However, mean value inspection (Fig. 13c-4) revealed that obtained mean values of FRC and CCC from libraries of 3700, 3900, and 4100 particles (orange box) were not-significantly-different, in agreement with our expectation (see section 2 of Part III).

4. Discussion

The GroEL simulation of discard ratio variation showed that obtained mean values of FRC and CCC from the libraries around the correct answer were not-significantly different according to mean value inspection (Fig. 13c-4, orange box). Assuming that a wider variation of projection angles is more desirable (Frank, 2006), a larger number of particles should be chosen in the discard ratio experiment. The application of this criterion would have resulted in selecting the correct answer of 4100 particles in the simulation. This observation also suggests that the 5495 particle library of NaChBac would have been selected in both with- and without-intra-class-alignment experiments (Fig. 7c-4 and 9c-4, orange box) if the mean value inspection was included in the decision criteria of the best condition. Thus, this simulation result suggests that the mean values of FRC and CCC are also useful for finding the appropriate discard ratio.

5. Conclusion

The result of this GroEL simulation study shows that IC-EOS can successfully select the correct number of classes and of particle images in the simulated library, validating the best-condition-selection method of IC-EOS. Fluctuations, means, and SDs of base measurements relative to these condition variations showed clearer trends than those in the case study using NaChBac. The tradeoff behavior between resolution and CCC was again observed in 2D averaging upon searching for the correct class counts. Noticeable saturation of resolution and CCC improvements at higher numbers of discarded images was observed during particle selection.

For the SPA structure determination of multiple ion channels and their multiple conformational variants, the automation of workflow optimization procedures is crucial for achieving high-resolution. Recent trends in SPA indicate that workflow optimization is increasingly important as individual algorithms improve. The analysis of various proteins with various data styles requires balancing workflow arrangements and their appropriate conditions, using algorithms provided by multiple SPA software packages. Since different proteins particles have different natures, it seems to be difficult, or at least not efficient, to use the same workflow for all proteins. The procedure for designing and constructing a highly optimized workflow should make SPA more accurate, efficient, and accessible to a much broader range of scientists.

Table 3. 2D Averaging conditions for simulation

a. CTP ratio variation with constant discard ratio

Particles	Classes	Members	CTP Ratio	Discard Ratio
4100	25	164.00	0.078	0.00
4100	50	82.00	0.110	0.00
4100	75	54.67	0.135	0.00
4100	82	50.00	0.141	0.00
4100	100	41.00	0.156	0.00
4100	125	32.80	0.175	0.00
4100	150	27.33	0.191	0.00
4100	175	23.43	0.207	0.00
4100	200	20.50	0.221	0.00

b. Discard ratio variation with constant CTP ratio (variable class counts)

Particles	Classes	Members	CTP Ratio	Discard Ratio
5740	115	49.91	0.142	-0.40
4920	98	50.20	0.141	-0.20
4510	90	50.11	0.141	-0.10
4305	86	50.06	0.141	-0.05
4100	82	50.00	0.141	0.00
3900	78	50.00	0.141	0.05
3700	74	50.00	0.141	0.10
3300	66	50.00	0.141	0.20
2450	49	50.00	0.141	0.40

Figure 11. GroEL simulation study

Simulation study using GroEL model projections. Leftmost column: 82 original projections generated from the GroEL crystal structure (PDB code: 1XCK, Bartolucci et al., 2005). Middle three columns: 4100 particle library of 0.05 SNR (the variance ratio). Fifty copies were generated from each original projection with random rotations and shifts. Noise was then added to the copies. Rightmost column: Final 2D averages using particle uprighting workflow with optimized conditions (4100 particles and 82 classes). The particles compared well with the original projections.

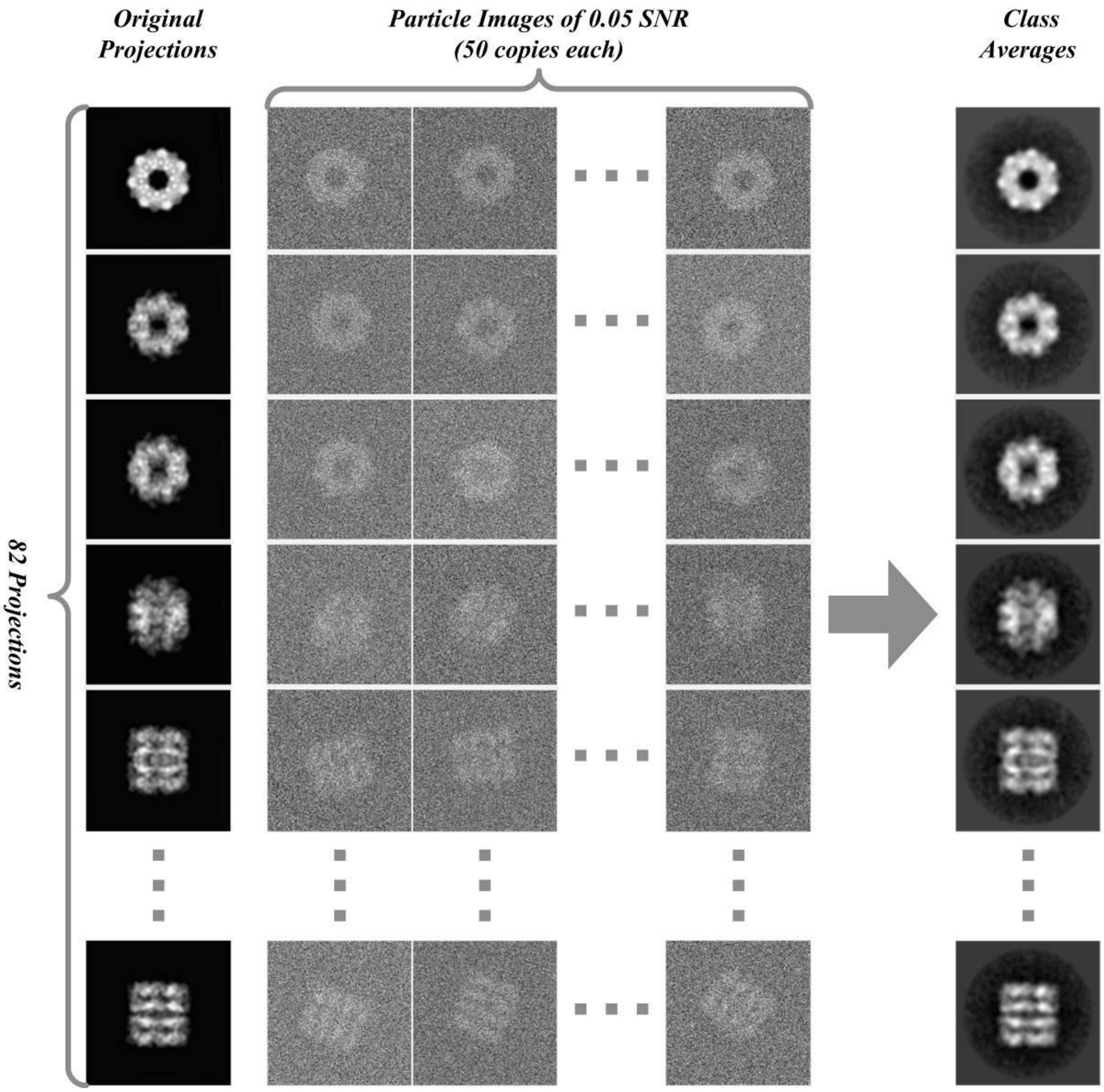
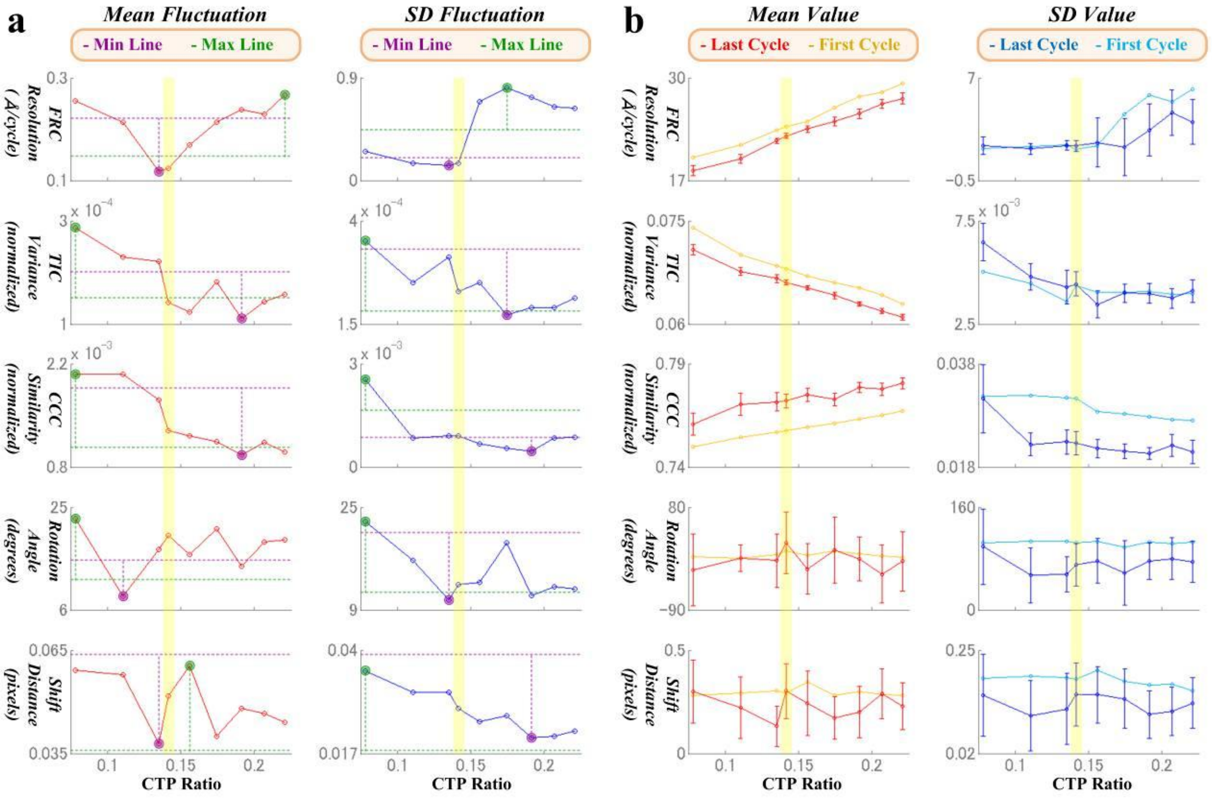

Original Projections
Particle Images of 0.05 SNR (50 copies each)
Class Averages
82 Projections

Figure 12. GroEL simulation related to CTP ratio variations using the workflow with particle uprighting

(a) Fluctuations of base measurements relative to CTP ratio. Mean and SD fluctuations of FRC-resolution showed clear trends, indicating that 75 and 82 classes (0.135 and 0.141 CTP ratios) were the most stable conditions. This coincided with the original 82 classes. Fluctuations of CCC-similarity also had clear plateaus at higher CTP ratios but suggested that the most stable condition was around 150 classes (0.191 CTP ratio). (b) Means and SDs of base measurements relative to CTP ratio. The tradeoff relationship between FRC and CCC were again obtained with this simulation. The improvement width of FRC SD value showed an obvious change at 82 classes and became wider at higher CTP ratios. (c) The best condition election correctly selected 82 classes (3rd section, highlighted with yellow). The convergence sub-vote (1st section) exhibited poor convergences of shift-distance, strongly indicating that this base measurement converged at the first cycle of iterations with this dataset. The concentrations of larger vote counts (1st to 3rd sections) were gained around this selected CTP ratio. The mean value inspection shows a similar result (4th section) compared to the case study of NaChBac (Fig. 6c).



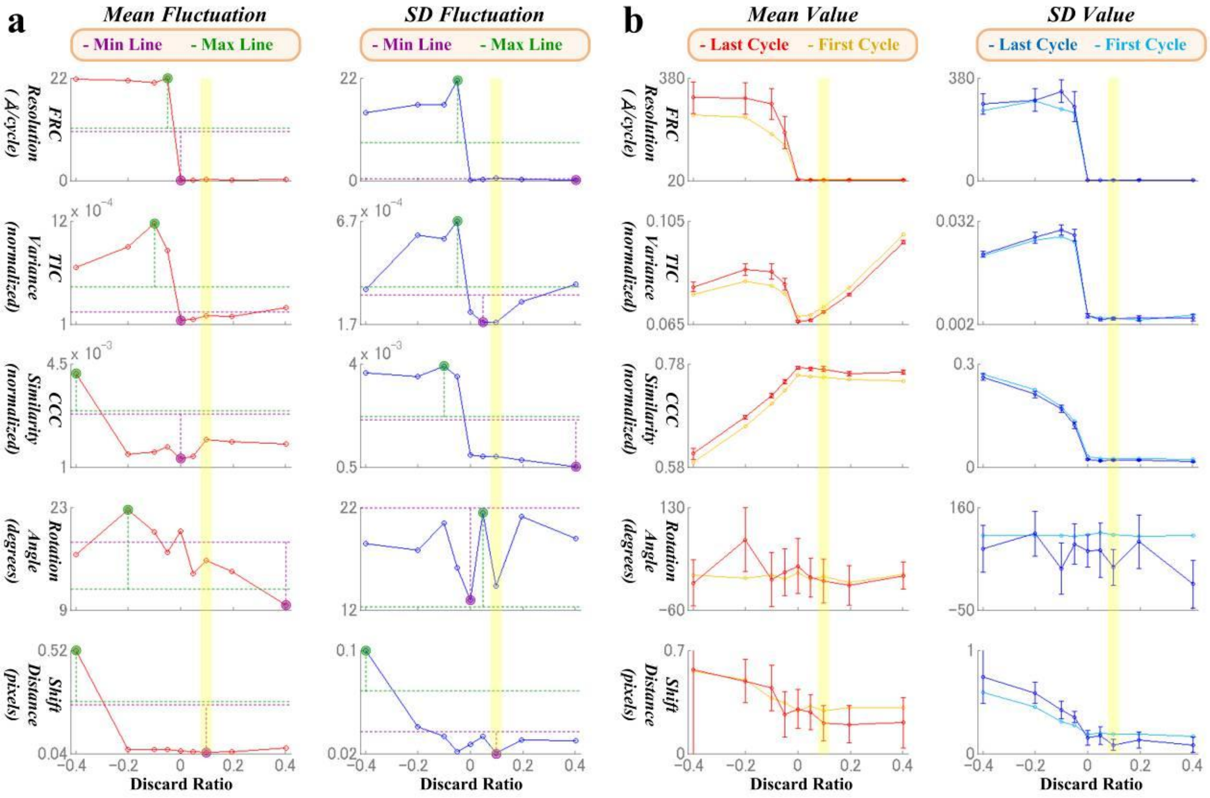
c

◎ : Converged X : Not-Converged ; ≠ : Significant Diff = : No-Significant Diff

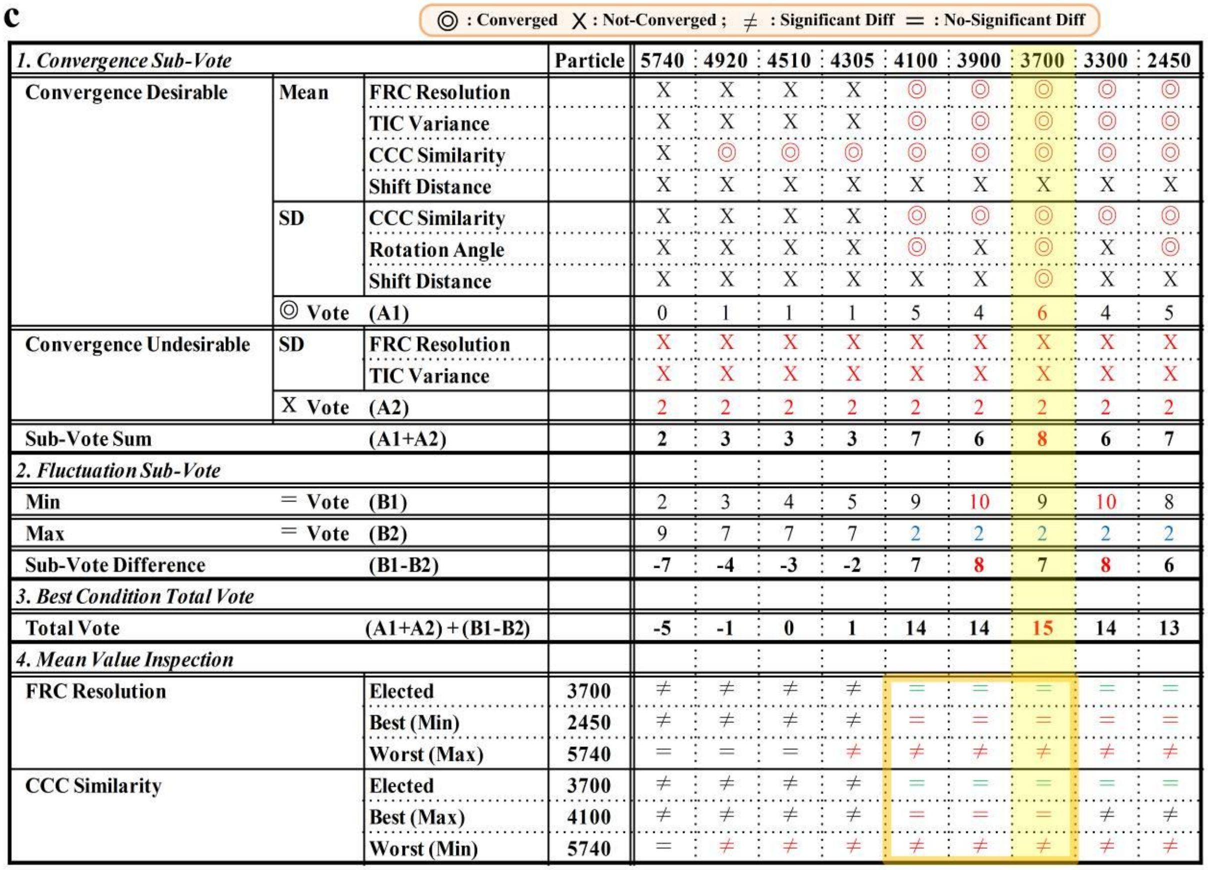
1. Convergence Sub-Vote			Class	25	50	75	82	100	125	150	175	200
Convergence Desirable	Mean	FRC Resolution		◎	◎	◎	◎	◎	◎	◎	◎	◎
		TIC Variance		◎	◎	◎	◎	◎	◎	◎	◎	◎
		CCC Similarity		◎	◎	◎	◎	◎	◎	◎	◎	◎
		Shift Distance		X	X	X	X	X	X	X	X	X
	SD	CCC Similarity		X	◎	◎	◎	◎	◎	◎	◎	◎
		Rotation Angle		X	◎	◎	◎	X	X	◎	X	X
		Shift Distance		X	X	X	X	X	X	X	X	X
	◎ Vote	(A1)		3	5	5	5	4	4	5	4	4
Convergence Undesirable	SD	FRC Resolution		X	X	X	X	X	X	◎	X	◎
		TIC Variance		X	X	X	X	X	X	X	X	X
	X Vote	(A2)		2	2	2	2	2	2	1	2	1
Sub-Vote Sum		(A1+A2)		**5**	**7**	**7**	**7**	**6**	**6**	**6**	**6**	**5**
2. Fluctuation Sub-Vote												
Min	= Vote	(B1)		2	8	7	8	8	8	8	7	6
Max	= Vote	(B2)		9	7	6	6	8	8	6	8	8
Sub-Vote Difference		(B1-B2)		**-7**	**1**	**1**	**2**	**0**	**0**	**2**	**-1**	**-2**
3. Best Condition Total Vote												
Total Vote		(A1+A2) + (B1-B2)		**-2**	**8**	**8**	**9**	**6**	**6**	**8**	**5**	**3**
4. Mean Value Inspection												
FRC Resolution		Elected	82	≠	≠	≠	=	≠	≠	≠	≠	≠
		Best (Min)	25	=	≠	≠	≠	≠	≠	≠	≠	≠
		Worst (Max)	200	≠	≠	≠	≠	≠	≠	≠	=	=
CCC Similarity		Elected	82	≠	=	=	=	=	=	≠	=	≠
		Best (Max)	200	≠	≠	≠	≠	=	≠	=	=	=
		Worst (Min)	25	=	=	≠	≠	≠	≠	≠	≠	≠

Figure 13. GroEL simulation using particle uprighting workflow with discard ratio variation

(a) Fluctuations of base measurements relative to discard ratio. Mean and SD fluctuations of FRC-resolution indicated that the conditions of 4100 particles (0.0 discard ratio) and fewer (positive discard ratios) were stable. Those of TIC-variance and SD fluctuation of CCC-similarity showed similar tendencies. (b) Means and SDs of base measurements relative to the discard ratio. The mean and SD values of FRC were obviously different between negative and non-negative discard ratios. All values of FRC and CCC had plateaus at non-negative discard ratios. The SD value of TIC-variance showed the same trend while its mean value had a trough at the correct 4100 particles. (c) The best condition election selected 3700 particles for the best convergence (0.10 discard ratio; 3rd section, highlighted with yellow), which was 10% fewer than the correct answer. The convergence sub-vote (1st section) again showed poor convergences of shift-distance, suggesting that this measurement of the simulated dataset converged at the first iteration cycle. Counts of convergence sub-vote sums (1st section), fluctuation sub-vote differences (2nd section), and total votes (3rd section) were obviously higher at the non-negative discard ratio side. According to the mean value, the optimal values for FRC and CCC overlapped at 3700 and 3900 particles (0.05 discard ratio) as well as at the correct 4100 particles (4th section, orange box).



c



◎ : Converged X : Not-Converged ; ≠ : Significant Diff ═ : No-Significant Diff

1. Convergence Sub-Vote			Particle	5740	4920	4510	4305	4100	3900	3700	3300	2450
Convergence Desirable	**Mean**	**FRC Resolution**		X	X	X	X	◎	◎	◎	◎	◎
		TIC Variance		X	X	X	X	◎	◎	◎	◎	◎
		CCC Similarity		X	◎	◎	◎	◎	◎	◎	◎	◎
		Shift Distance		X	X	X	X	X	X	X	X	X
	SD	**CCC Similarity**		X	X	X	X	◎	◎	◎	◎	◎
		Rotation Angle		X	X	X	X	◎	X	◎	X	◎
		Shift Distance		X	X	X	X	X	X	◎	X	X
	◎ **Vote**	**(A1)**		0	1	1	1	5	4	6	4	5
Convergence Undesirable	**SD**	**FRC Resolution**		X	X	X	X	X	X	X	X	X
		TIC Variance		X	X	X	X	X	X	X	X	X
	X **Vote**	**(A2)**		2	2	2	2	2	2	2	2	2
Sub-Vote Sum		**(A1+A2)**		**2**	**3**	**3**	**3**	**7**	**6**	**8**	**6**	**7**
2. Fluctuation Sub-Vote												
Min	═ **Vote**	**(B1)**		2	3	4	5	9	10	9	10	8
Max	═ **Vote**	**(B2)**		9	7	7	7	2	2	2	2	2
Sub-Vote Difference		**(B1-B2)**		**-7**	**-4**	**-3**	**-2**	**7**	**8**	**7**	**8**	**6**
3. Best Condition Total Vote												
Total Vote		**(A1+A2) + (B1-B2)**		**-5**	**-1**	**0**	**1**	**14**	**14**	**15**	**14**	**13**
4. Mean Value Inspection												
FRC Resolution		**Elected**	**3700**	≠	≠	≠	≠	═	═	═	═	═
		Best (Min)	**2450**	≠	≠	≠	≠	═	═	═	═	═
		Worst (Max)	**5740**	═	═	═	≠	≠	≠	≠	≠	≠
CCC Similarity		**Elected**	**3700**	≠	≠	≠	≠	═	═	═	═	═
		Best (Max)	**4100**	≠	≠	≠	≠	═	═	═	≠	≠
		Worst (Min)	**5740**	═	≠	≠	≠	≠	≠	≠	≠	≠

Part IV: SPA structure determination of MG23

1. Introduction

In parallel to the work on IC-EOS, I also conducted the SPA structure determinations of two negatively-stained proteins: a cation-conducting channel (Mitsugumin 23 or MG23) and NaChBac. Over the past few decades, an enormous number of 3D structural studies revealed that the remarkable diversity of shapes and sizes of proteins is a reflection of the specificity with which they interact, often in a lock-and-key fashion. The components of these nano-sized "work machines" show mechanical movements such as rotation, gliding, and lever motions. SPA using negatively stained proteins has been used in a significant number of these studies. Here, I followed the procedure of workflow optimization used in our previous SPA studies (e.g. Maruyama et al., 2007; Mio et al., 2007; Yazawa et al., 2007). This led to conclusions about the qualitative criteria used by human evaluation described in the General Introduction, leading to the development of IC-EOS. The structure determination of MG23 is described in this Part IV, and that of NaChBac is in Part V.

MG23 was recently identified as a Sarcoplasmic Reticulum (SR) protein, and was expected to reveal the molecular basis of Ca^{2+} stores (Nishi et al., 1998). This is a 23 kDa transmembrane protein localized to the Endoplasmic Reticulum (ER), SR, and nuclear membranes in a wide variety of cells. The ER/SR is a multifunctional organelle responsible for important cellular processes, including protein maturation, lipid metabolism, Ca^{2+} signaling, and stress response. An important function of the ER/SR is an intracellular Ca^{2+} store, and its Ca^{2+} release controls physiological functions such as muscle contraction, secretion, metabolism, and transcription.

The ubiquitous distribution of MG23 implies its participation in a fundamental function of intracellular membrane systems. Work by our co-research group showed that MG23 behaves as a voltage-dependent, cation-conducting channel, permeable to both K^+ and Ca^{2+} (Venturi et al., 2011). One feature of MG23 gating is that multiple channels always appear to be gating together in the bilayer. In EM observations, most MG23 particles are round-shaped or bowl-shaped, which represent vertical or lateral views, respectively. 3D reconstruction revealed that MG23 forms a large bowl-shaped complex equipped with a putative central pore. However, we realized that the MG23 preparations always contained considerable numbers of asymmetrical particles that were relatively smaller than the bowl-shaped particle. This suggested that two states of multi-metric assemblies existed among MG23 particles. To confirm this observation, I additionally reconstructed the 3D structure of the smaller-sized particles.

2. Materials and Methods

2.1. Affinity purification and transmission electron microscopy of recombinant MG23

I started the 3D reconstruction of MG23 particles from micrographs taken at our laboratory. Purified MG23 particles were kindly provided by Professor H. Takeshima (Graduate School of Pharmaceutical Sciences, Kyoto University, Japan). For production of recombinant MG23 using a methylotrophic yeast system (Invitrogen), a His tag sequence (Hisx6) was inserted into rabbit MG23 cDNA at a site immediately downstream of the N-terminal signal sequence (Venturi et al., 2011). The total microsome fraction was prepared from yeast cells expressing His-tagged MG23, and the recombinant protein was solubilized with NP-40 or n-dodecyl β-D-maltoside (DDM) and purified using combined Ni and mAb-C affinity chromatography. MG23 particles were further separated by density gradient centrifugation. The 5-20% sucrose liner gradient containing the final elution buffer without the peptides was prepared and centrifuged for ~10 hrs in a Beckman SW41 rotor at 220,000xg. MG23-enriched fractions were detected by Western blotting. MG23 particles were applied to thin carbon films, negatively stained with 2% uranyl acetate solution, and imaged using a transmission electron microscope (JEM-100CX; JEOL, Tokyo, Japan) at x39,024 magnification with a 100 kV acceleration voltage. Images were recorded on SO-163 films (Eastman Kodak, Rochester, NY) and digitized with a Scitex Leafscan 45 scanner (Leaf systems Inc., Westborough, MA) at a pixel size of 0.256 nm (the specimen level).

2.2. 3D reconstruction of the MG23 particle

Particle 3D reconstruction was performed essentially as described in previous works (e.g. Maruyama et al., 2007; Mio et al., 2007; Yazawa et al., 2007). MG23 projections were primarily selected in 160 x 160 pixel subframes (corresponding to 41 x 41 nm) using the auto-accumulation method with simulated annealing (Ogura and Sato, 2004b). Selected particles were rotated at 10-degree increments to generate a set of training images for the three-layer pyramidal-type neural network (Ogura and Sato, 2001; Ogura and Sato, 2004a). The trained neural network selected additional particles.

For the 3D reconstruction, 1,707 particle images of the small-sized group were used. The particle library was processed in the following steps. For the initial 2D averaging, particles in each library were aligned rotationally and translationally (Frank, 2006; Harauz and van Heel, 1986; van Heel et al., 2000) relative to each other using the reference free method (Ogura and Sato, 2004a). Aligned images were classified into120 clusters using Multivariate Statistical Analysis (MSA, Bretaudiere and Frank, 1986). The resulting class averages were used as new references and the cycle from alignment to classification was repeated until convergence. Because no symmetry was observed in the averaged images or the raw images, the smaller-sized molecule was reconstructed assuming no intramolecular symmetry. An Euler angle was assigned to each average image, using Echo-correlated 3D reconstruction with simulated annealing (Echo-correlated method, Ogura and Sato, 2006). Using the exact-filter back-projection algorithm (Harauz and van Heel, 1986), a primary 3D model was reconstructed from averages with their assigned Euler angles, and used as the reference structure for the next step. To optimize 3D reconstruction using the projection matching method (Penczek et al., 1994), reprojection images were created from the 3D model. Using these as references, the particles in the library were rotationally and translationally aligned by the MRA method (Frank, 2006; Harauz and van Heel, 1986; van Heel et al., 2000). The aligned particles were further classified using the same classification algorithm as above. For each average image, an Euler angle of the reprojection image was assigned so as to show the best correlation to the average. These steps were repeated until the 3D model became stable. During these iterations, the Echo-correlated method was used to optimize the Euler angles of class average images. To assess the resolution of the resultant 3D density map, the data was divided into odd and even subsets. Using the subsets, two independent 3D reconstructions were computed without masking. These two 3D density maps were compared by FSC (Harauz and van Heel, 1986) at the threshold of 0.5, using IMAGIC-5 (van Heel et al., 1996).

3. Results

In highly-purified MG23 preparations, predominant particles of the large-sized group were about 17-19 nm in length and 14-16 nm in width, and appeared to have 6-fold symmetry. The minor particles had a smaller size (about 13-17 nm in length and 6-8 nm in width) and asymmetrical structures, which appeared to be a component of the predominant particle. For this small-sized group, the final reconstruction included 1,699 particle images (99% of the selected images). The surface representation demonstrates that this structure was roughly crescent-shaped, 14 nm in height, 11 nm in side length, and 6 nm in narrow side length (Fig. 14b). Representative raw images are presented (Fig. 14c, first row), with their corresponding class averages (Fig. 14c, second row) and with surface representations and reprojections (Fig. 14c, third and fourth rows). The almost random orientations of the particles on the grid surface are also shown by a plot of the Euler angles of the 117 adopted class averages (Fig. 15a). According to the FSC function, the resolution limit was 3.6 nm by the correlation coefficient of 0.5 criterion (Fig. 15b).

4. Discussion

EM observations detected a large symmetrical particle as the predominant component and a small asymmetric assembly as the second major component in highly purified MG23 preparations. Because the MG23 preparations after affinity purification and size fractionation using sucrose gradient centrifugation were shown to be highly purified with SDS-PAGE, the concomitance of the large and small particles likely suggests that MG23 continuously disassembles and reassembles. These transitions in channel building may underlie the unusual gating behavior of MG23, and allow rapid cationic flux across intracellular membrane systems (Venturi et al., 2011).

It was considered that the smaller particles constitute a partial domain of the large bowl-shaped assembly.

However, it is unlikely that these larger and smaller structures observed using EM represent hexameric and monomeric MG23, respectively. This is because the MG23 monomer at 23 kDa, composed of 210 amino acids, is too small for the EM observation (Nishi et al., 1998; Venturi et al., 2011). It was therefore proposed that the crescent-shaped particle is a hexameric minimum structural subunit of MG23, and that multiple subunits subsequently constitute a larger assembly such as the bowl shaped structure. With the available data, it was not possible to determine the exact number of crescent-shaped subunits that would make up the full particle, but biochemical and structural analysis points to a number composed of two to six subunits. Fitting of more than six crescent-shaped subunits to the bowl-shaped structure seemed impossible. A docking of six crescent-shaped subunits is presented as the most plausible model (Fig. 14d). The 3D map of the crescent-shaped subunit was contoured at an isosurface containing a volume corresponding to 138 kDa (23 kDa × 6), and the bowl-shaped particle at 828 kDa (138 kDa × 6), assuming a hexagonal assembly of the crescent-shaped subunit. However, this simulation could not exclude other possibilities of subunit construction in the bowl-shaped particle.

Figure 14. 3D reconstruction of MG23 particles in the smaller-sized group

(a) Raw EM images of recombinant MG23 particles. After adsorption to the glow-discharged carbon film, negatively stained samples were imaged by EM. (b) Surface representations of the small-sized MG23 particle group structure from top (left) and oblique (right) views. Dimensions of the smaller-sized crescent-shaped molecule were 14 nm (height), 11 nm (wide side length), and 6 nm (narrow side length). The multiple subunits of this structure were considered to subsequently constitute a larger assembly (Fig. 14d). (c) First row, representative raw images; second row, corresponding class averages; third row, surface representations; and fourth row, projections of the 3D reconstruction. Consistency among data sets was very high in size, shape, and inner structure, indicating successful 3D reconstruction from the original images of recombinant MG23 particles. (d) Proposed configuration of the six-subunit assembly. In this model, the large bowl-shaped structure (full-MG23 particle) was fitted with the six crescent-shaped subunits (putative hexameric subunit). This configuration assumes that each subunit is located longitudinally at 60-degree intervals (left panels). Volume of the full-MG23 was contoured at 828 kDa (yellow mass with red outline) assuming a hexagonal assembly of the 138 kDa (23 kDa x 6) hexameric subunit, and superimposed to the subunit assembly (right panels). The outlines from the top resemble each other, but in the side view, the crescent-shaped subunits are somewhat longer and narrower than would ideally fit the fully assembled MG23 particle. The discrepancy in fitting may suggest the conformational changes of each subunit during assembling into full-MG23 particles. Scale bars are 10 nm.

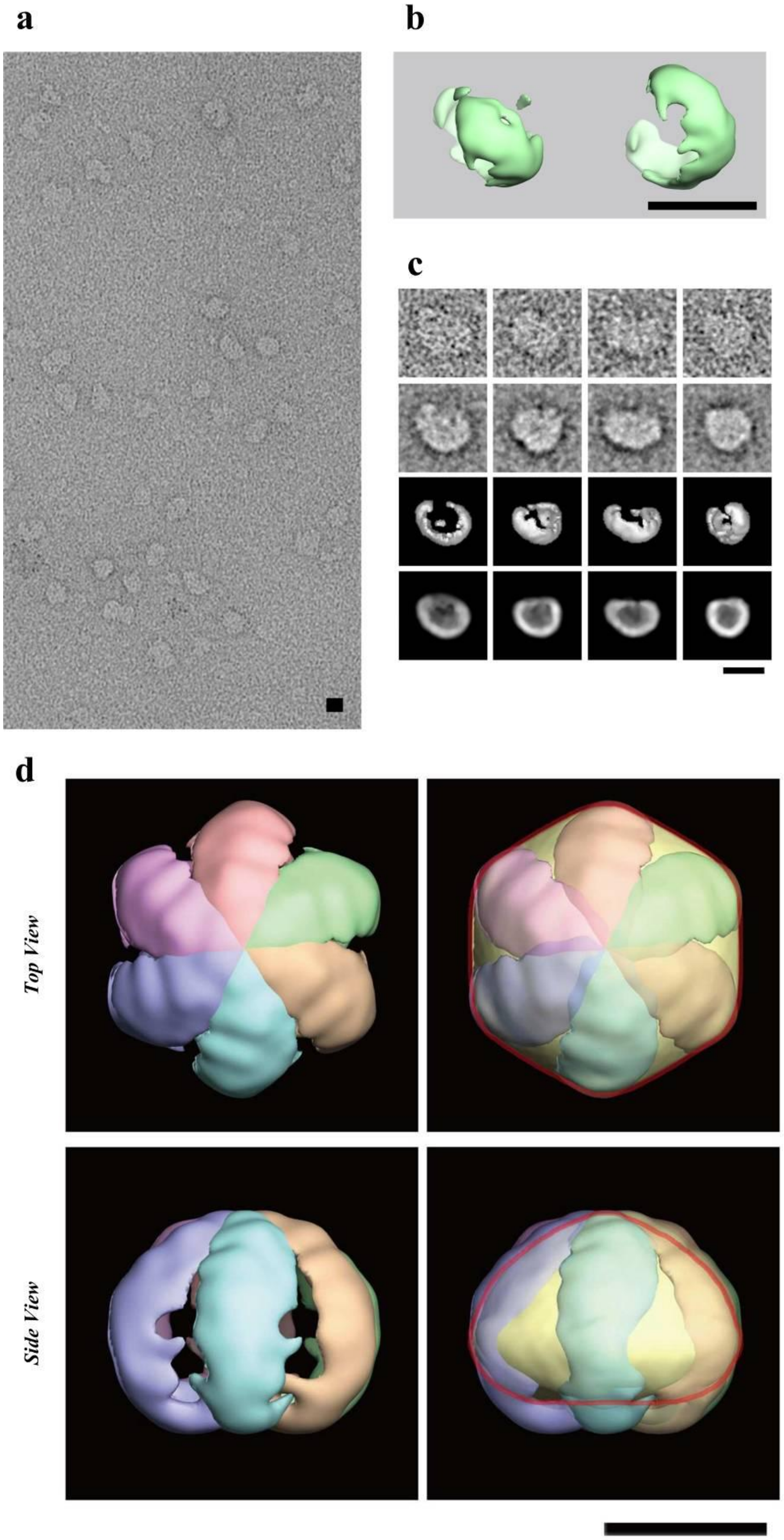
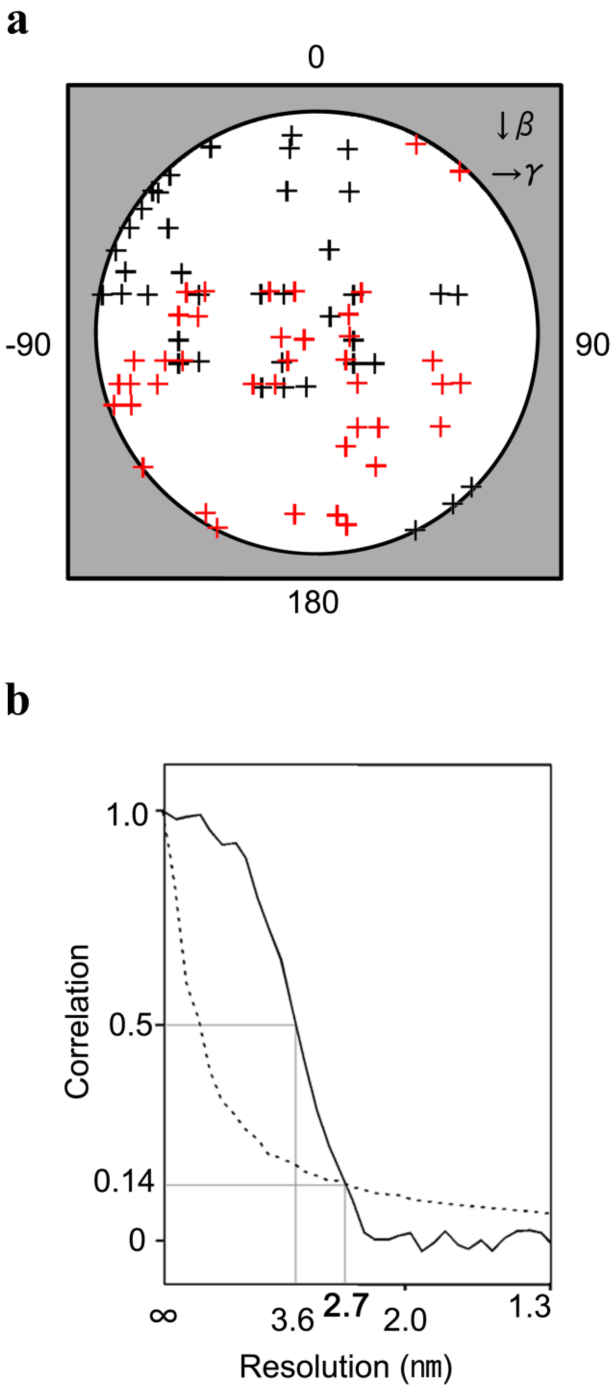

a
b
c
d
Top View
Side View

Figure 15. Euler angle distribution and FSC function of the crescent-shape MG23 particles

(a) Surface projection of Euler angles of the 117 adopted class averages shows that the MG23 particles in the small-sized group were almost randomly oriented on the grid surface. Red crosses indicate their positions were on the far side of the globe. (b) According to the FSC function (solid line), the resolution limit of the particle was shown to be 3.6 nm by the correlation coefficient 0.5 criterion. The noise curve of 3σ (broken line) is also shown as a comparison.


a
0
↓β
→γ
-90
90
180
b
1.0
0.5
0.14
0
Correlation
∞
3.6
2.7
2.0
1.3
Resolution (nm)

Part V: SPA structure determination of NaChBac

1. Introduction

NaChBac from *Bacillus halodurans* is the first functionally characterized bacterial Na_V channel (Ren et al., 2001), and shares key physiological properties with its mammalian homologues, including voltage-dependent activation, inactivation, and pharmacological sensitivity (Mio et al., 2010). Experiments suggest that NaChBac assembles as a homotetramer to exert its channel function. Each subunit has 6 transmembrane segments, and the 24-transmembrane-type mammalian Na_V channel is supposed to be evolved from such a protein by repeated gene duplication. The molecular weight estimated from the amino acid sequence is 33 kDa for the monomer and 134 kDa for the tetramer (Mio et al., 2010). The study of Na_V channels facilitates potential clinical applications. Their mutations underlie several clinical disorders, such as inherited epilepsy, migraine, periodic paralysis, cardiac arrhythmia, and chronic pain syndromes (Ryan and Ptacek, 2010). Na_V channels are molecular targets of drugs used in local anesthesia and in the treatment of genetic and sporadic Na_V channelopthies in the brain, skeletal muscles and heart (Mantegazza et al., 2010). The various channel properties and structural simplicity of bacterial Na_V channels are suitable for a model system of structure-function analysis (Irie et al., 2012; Koishi et al., 2004; Ren et al., 2001; Zhang et al., 2012).

2. Materials and Methods

2.1. Expression, purification, and transmission electron microscopy of NaChBac

The entire sequence of NaChBac was subcloned with an N-terminal FLAG tag to the pCI vector (Promega). The NaChBac protein was expressed in HEK293 cells, solubilized using 20 mM n-dodecyl β-D-maltoside and purified with anti-FLAG affinity gel (Sigma) followed by Superdex 200 size-exclusion chromatography (GE Healthcare) (Mio et al., 2010). Purified proteins were adsorbed to thin carbon films rendered hydrophilic by glow discharge in low-pressure air and supported by copper mesh grids. Samples were washed with 10 drops of double-distilled water, negatively stained with 2% uranyl acetate solution for 30 seconds twice, blotted, and dried in air. Micrographs of negatively stained particles were taken by a JEOL 100CX transmission electron microscope (JEOL, Tokyo, Japan) at x52,100 magnification with 100 kV acceleration voltage. Images were recorded on SO163 film (Eastman Kodak, Rochester, NY) developed with a D19 developer (Kodak) and digitized with a Scitex Leafscan 45 scanner (Leaf system Inc., Westborough, MA) at a pixel size of 1.921 Å (specimen level).

2.2. Automated particle selection and image preprocessing in the NaChBac study

The NaChBac particle selection procedure was same as the MG23 study, except that the initial library of 657 particles was manually divided into subframes with dimensions of 128 x 128 pixels, without using the auto-accumulation method (Ogura and Sato, 2004b). A library of 10,947 images was created from 40 micrographs. A band pass filter was applied to the particle library; the low-frequency cutoff was 221.3 Å and the high-frequency cutoff was 9.6 Å at each cycle. The size of the particle images was reduced to 64 x 64 pixels (a pixel size of 3.842 Å at the specimen level) for the optimization and evaluation of 2D averaging by IC-EOS in Part II.

2.3. 3D Reconstruction of the NaChBac particle

The 3D reconstruction method described in the MG23 study was used again for the NaChBac particles, with several minor modifications. For the initial 2D averaging using “alignment-by-classification” (Dube et al., 1993), the number of clusters for MSA classification was 200, and the iteration steps were repeated for 30 cycles. After the completion of this stage, the averages were examined to test the presumption of symmetry by MSA symmetry analysis (Costa et al., 2006; Dube et al., 1993). Among the putative top images, a set of particles and their averages showed highly 4-fold-symmetrical features. This result is consistent with the homo-tetrameric subunit stoichiometry of NaChBac. Therefore, 4-fold symmetry was imposed in the following computation. In the projection matching, 188 images were generated by the 3D forward projection, the particles were classified according to their assigned Euler angles instead of MSA classification (Bretaudiere and Frank, 1986), and the particles within each class were further aligned relative to the class average. To reduce the bias of the current 3D density map, intra-class alignment was repeated for a fixed number of times (3-5 times) (Ludtke et al., 2004). The iteration steps of projection matching were repeated until the 3D model became stable (15-30 cycles).

3. Results

Most NaChBac particles were bullet- or square-shaped with round corners (Fig. 16a). The variation in shape was interpreted to reflect different orientations of the same molecule on the grid. The square-shaped particles seemed to imply top views of the tetrameric form (perpendicular to the membrane); the bullet-shapes would be side views. Putative side views clearly showed asymmetry with respect to the membrane; one end was wider and the other narrower.

The final 3D reconstruction of the NaChBac molecule included 5,348 particles, 48.9% of all the picked-up images. For surface representation, the 3D map was contoured at an isosurface containing a volume corresponding to 241 kDa: 180 % of the tetrameric NaChBac mass (33 kDa * 4 = 134 kDa, Mio et al., 2010) calculated from the amino acid composition. The surface representation depicts a squared bell-shaped molecule with a pointed bottom tip and a boat-tail top (Fig. 16b). Viewed from the top, NaChBac was a square with a side length of 95 Å and a diagonal length of 115 Å. The height estimated from the side views was 144 Å. Representative raw images are presented (Fig. 16c, first row), with their corresponding class averages (Fig. 16c, second row) and with surface representations and reprojections (Fig. 16c, third and fourth rows). Reprojections from the final density map were consistent with raw images and class averages, reflecting successful reconstruction from the original particle images. A plot of the Euler angles of particles assigned into188 equally distributed angles showed the nearly random orientation of NaChBac on the grid surface (Fig. 17a). According to the FSC function, the resolution limit for the reconstruction was 14.7 Å by the correlation coefficient of 0.5 criterion (Fig. 17b).

4. Discussion

Recent publications of X-ray crystal structures of bacterial Na_V channels have begun to reveal the kinetics of their central ion-conducting pore modules (extracellular funnel, selectivity filter, central cavity, and intracellar activation gate) and of voltage-sensing domains (NavAb from *Arcobacter butzleri*, Payandeh et al., 2011; Payandeh et al., 2012; NavRh, Zhang et al., 2012).

The structure of their crytosolic C-terminal region was not determined, although it was located in the crystal lattice (Payandeh et al., 2011). The crystal structure of the C-terminal region of NavSuIP from *Sulfitobacter pontiacu* was independently published (Irie et al., 2012).

In the preliminary structure determination of NaChBac, the obtained dimensions were highly consistent with the crystal structures of NavAb (PDB code: 3RVY, Payandeh et al., 2011) and NavSulP (PDB code: 3VOU, Irie et al., 2012), lending plausibility to the reconstruction. However, the cytosolic C-terminal coiled-coil structure, which is a conserved amino acid sequence, was not clearly resolved in the density map. Variation of the particles is the likely explanation, blurring structural details in the density map and lowering its resolution. The variations could be caused by domain flexibility of the whole ion channel or by negative staining. In SPA, crystal packing can be avoided. However, deformation and uneven staining of particles can occur in negative staining, and ion channels have highly flexible domains, which may change conformation between the open and closed states (Maruyama et al., 2007). Therefore, it was necessary to classify these 3D variations in the particle library, and IC-EOS was developed as a solution to this problem.

Figure 16. 3D reconstruction of NaChBac particles

(a) Raw EM images of NaChBac particles. (b) Surface representations of the reconstructed density map from top, bottom, and two side views. Dimensions were 144 Å (height), 95 Å (side length), and 115 Å (diagonal length at the widest region). (c) First row, representative constituent particle images; second row, corresponding class averages; third row, surface representations of the 3D reconstruction; and fourth row, projections of the 3D reconstruction. High consistency among data sets indicates successful 3D reconstruction from the original images of NaChBac particles. Scale bars are 100 Å.

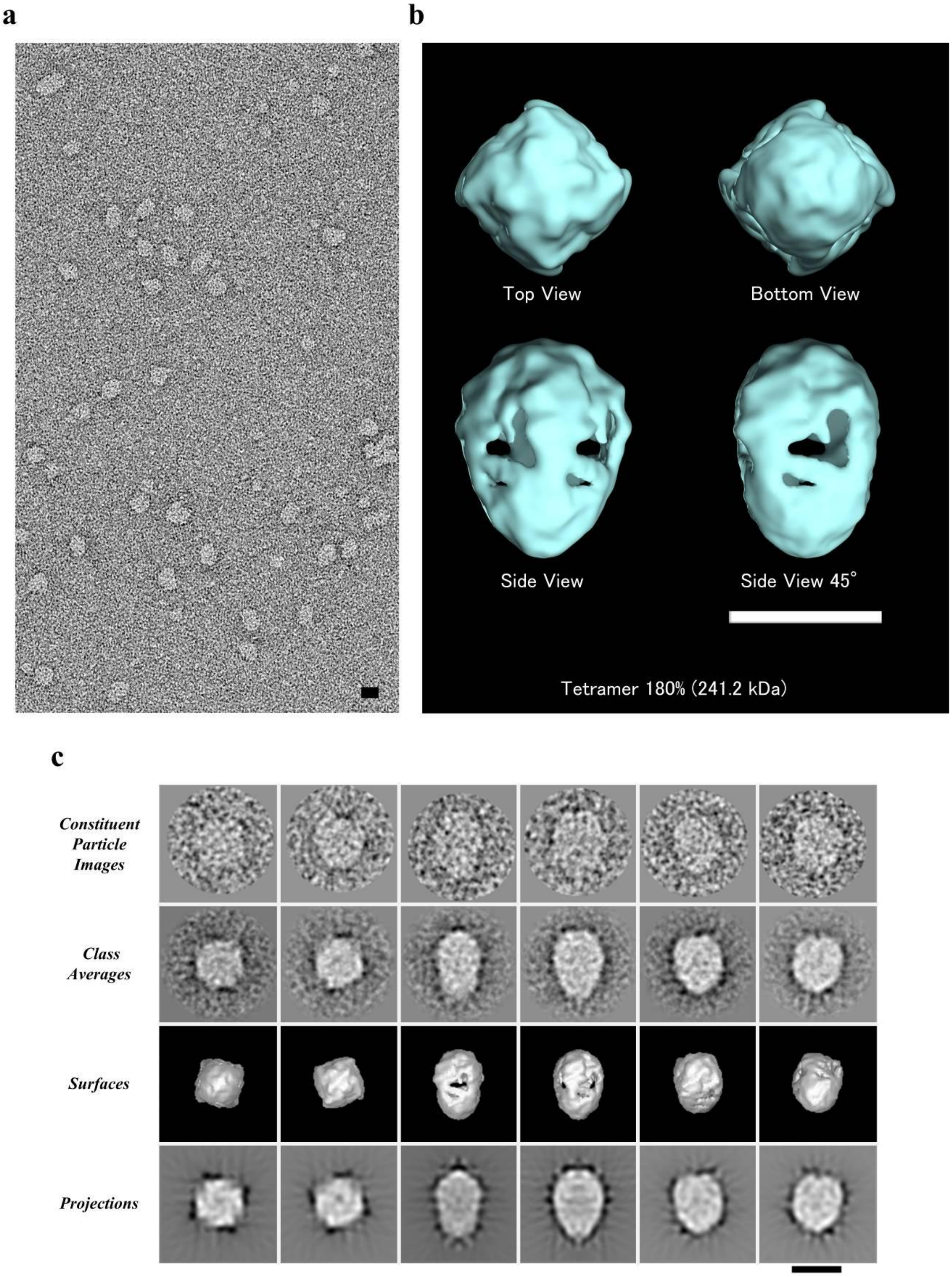
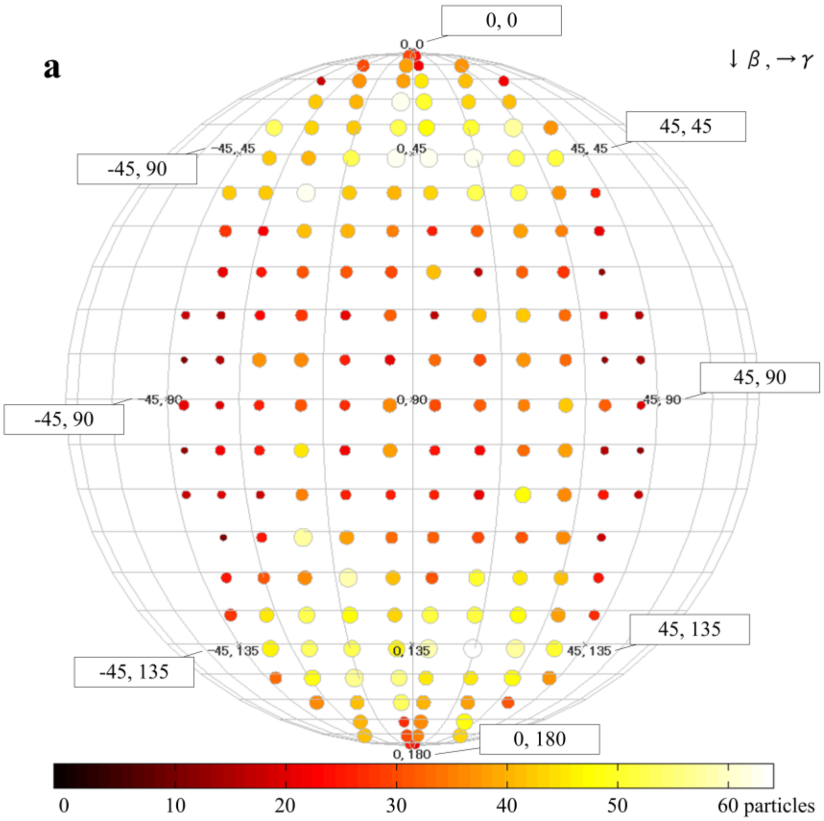
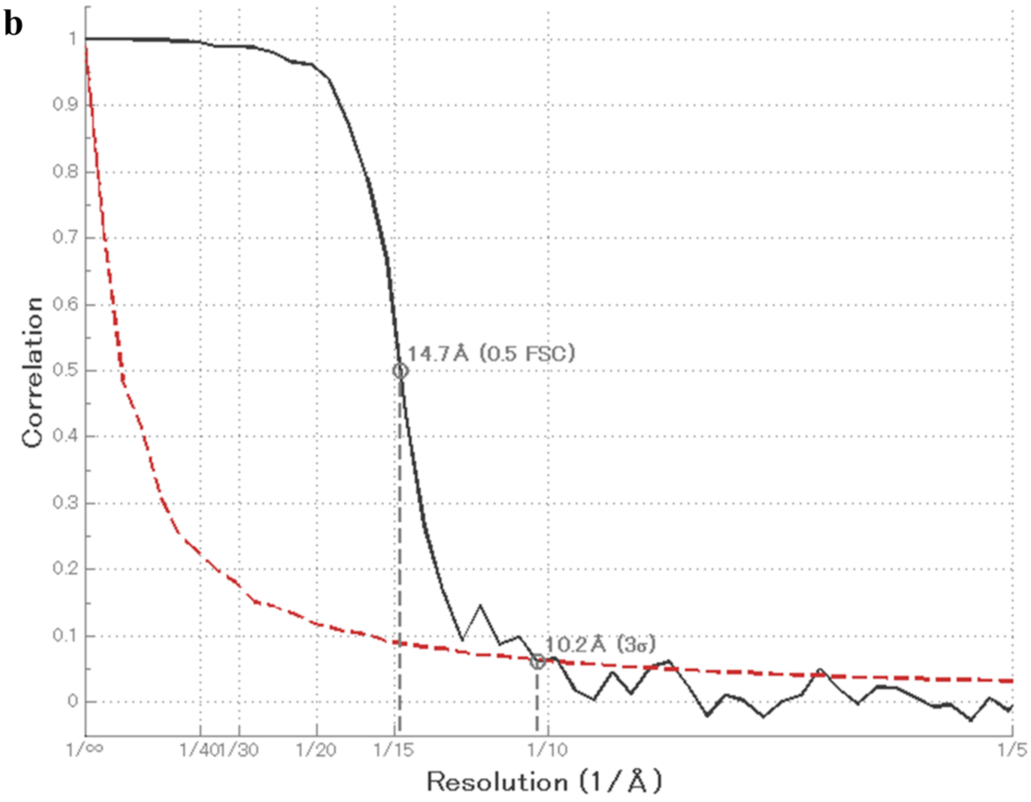

a
b
Top View
Bottom View
Side View
Side View 45°
Tetramer 180% (241.2 kDa)
c
Constituent Particle Images
Class Averages
Surfaces
Projections

Figure 17. Euler angle distribution and FSC function of the NaChBac reconstruction

(a) Surface projection of Euler angles of 5,348 particles assigned into 188 equally distributed angles shows that particles were almost randomly oriented on the grid surface. The size and color of each circle indicate the number of particles (indicated by the color bar at the bottom) assigned to the corresponding Euler angle coordinates. (b) According to the FSC function (solid black line), the resolution limit of the NaChBac particle was 14.7 Å by the correlation coefficient 0.5 criterion. The noise curve of 3σ (broken red line) is also shown as a comparison.


a
0, 0
↓ β , → γ
45, 45
-45, 90
-45, 45
0, 45
45, 90
-45, 90
0, 90
45, 135
-45, 135
0, 135
0, 180
0
10
20
30
40
50
60 particles


b
Correlation
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
14.7Å (0.5 FSC)
10.2Å (3σ)
1/∞
1/40
1/30
1/20
1/15
1/10
1/5
Resolution (1/Å)

General Conclusion

The brain, like other biological organs, has structural hierarchy, from the whole brain to nervous tissue, neurons, organelles, macromolecular assemblies, and macromolecules. EM covers the size range of smaller objects where the resolution of conventional optical microscopy is not sufficient due to the diffraction limit of light. As shown by detailed structural studies of cells, organelles, and viruses, the current achievements of biology as well as brain science would be impossible without EM.

Until recently, preponderance of EM studies in biology had been concluded with qualitative descriptions of sample images, except for the quantitative data related to distances, sizes, or numbers of particles in the images. Over the past few decades, other types of measurements also attempted to form an accurate 3D representation of biological objects. These attempts have been steadily spreading in the EM field, and a group of researchers recently organized the field of Three-Dimensional Electron Microscopy (3DEM). The reconstruction of 3D structure from EM images reveals not only shape but also interior density variations, since transmission images reflect the density of a biological specimen, much like X-ray imaging.

Macromolecular assemblies are in the size range of 50-500 nm, suitable for EM observation but beyond the resolution of conventional optical microscopy (Frank, 2006). Particles of a macromolecular assembly have identical structures, by functional necessity, and exhibit identical views in the EM when placed on the sample holder in the same orientation. Averaging methods for these particle images can reduce the noise level. This allows lower electron dosages, which can help prevent radiation damage. These issues were the impetus for the development of SPA.

Currently, SPA is not used as widely as crystallography for the structure study of proteins, mainly because of the resolution. However, the likelihood of successful crystallization is still unpredictable for a new protein, though sophisticated methods for inducing crystallization are available now. In addition, crystal packing may assume only a small range of the possible physiologically relevant conformations of proteins. The large size of the complexes formed in multiple associations and their likely dynamic variability in the functional context takes most of these complexes out of the reach of crystallography, and leaves SPA as one of the few techniques of structural investigation (Abbott, 2002), especially for ion channels.

In recent years, it has become possible to associate the 3D density distribution of a protein reconstructed by SPA quantitatively to the relevant atomic models from crystallography (e.g. Ludtke et al., 2004; Ludtke et al., 2008). This provides a way of resolving very large structures to atomic-level resolution. Components of a macromolecular assembly can often be induced to form crystals suitable for crystallography, but it is not always true for the entire assembly because of its size or inherent flexibility. On the other hand, it is frequently possible to obtain the entire structure of such an assembly at near-atomic resolution using SPA, while leaving the molecular interactions unconstrained. By matching the atomic model of the component determined by crystallography to the SPA structure of the entire macromolecule, high positional accuracies beyond the SPA resolution can be obtained, and atomic resolution within the large structure can be reached, or at least approximated. Similarly, electron tomography can be undertaken for objects whose structures show considerable individual variability. This method is suitable for the size ranges from bacteria to organelles, and the whole cell. There are now numerous examples for spherical viruses whose structure has been explored by SPA using this combinational approach. At least in principle, atomic-level visualizing and modeling have become possible for a cell and a primitive organism in their entirety, and the same can be said for a neuron in the hierarchy of the brain structure.

Thus, it is desirable to disseminate SPA and make structure analysis a routine procedure in relevant biological research. For this, an easy-to-use SPA system for biologists should be developed, and the system should ultimately be fully automated. The system should require less experience of analysis and less knowledge of multiple disciplines to master. To make such a system, the "know-how" of SPA practice should be integrated into a software system and should be hidden from users as a black box. With a carefully designed user interface and an automated search for the best parameter settings, users will be free from figuring out unnecessary parameter setting for their analysis purpose. Three points of IC-EOS reduce user labor in SPA workflow optimization and make automation possible; (1) automating the convergence decision of a workflow, (2) showing statistical decisions that rate workflow optimization differences, and (3) automating the selection of the workflow in the most-well-balanced optimization state, by comprehensive decisions using the voting method with convergence evaluation and fluctuation. This establishment of easy-to-understand, quantitative criteria for the convergence evaluation of workflows is a step toward the full automation of SPA.

However, as long as analysis depends on human image recognition and experience, it is not possible to make the system fully automated. In the present study, I demonstrate that the use of numerical computation methods is effective for assessing the behavior and characteristics of workflow for 2D averaging. The statistics and the fluctuations of the five base measurements, as energy functions of the workflow, allowed software to evaluate results without the aid of any human.

My future research will attempt further automation of SPA by applying IC-EOS to the refinement of 3D structures. This is the next important iterative workflow after 2D averaging. The basic principle of IC-EOS and its five base measurements are not limited to 2D averaging. The only necessary modification is to convert these measurements to 3D analogues, for example introducing FSC instead of FRC as a resolution measurement. The assessment of IC-EOS effectiveness in 3D refinement should provide additional evidence that the quality of the initial 2D averages strongly influences the resolution of the final 3D structure.

Cryo-EM studies with IC-EOS should follow to further improve the structural resolution of ion channels. In this thesis, a negatively stained sample was used as the first attempt to assess the effectiveness of IC-EOS, because its relatively high contrast makes the duration of analysis shorter and the evaluation of results easier. Negative staining is still used in high-resolution EM of macromolecules as an important first step in identifying characteristic views and assessing if a molecule is suitable for SPA (Frank, 2006). Therefore, it is still important to know the detailed behavior and characteristics of 2D averaging with negatively stained data relative to condition variations. However, deformation and varied staining of particles are inevitable and introduce additional 3D variations. These variations make the interpretation of results more complicated, and limit achievable resolution. On the other hand, cryo-EM preserves a close-to-native state of fully hydrated proteins. Therefore, even with the extremely low contrast of the particle images from cryo-EM, it is possible to reach a much higher resolution than with negative staining method. Because of these differences, I expect that the behaviors of SPA workflows with cryo-EM images will be different from those with negatively stained EM images.

The combination of multiple 3D structure determination methods can also be applied to obtain the atomic resolution of ion channels in multiple conformations. With these atomic models, their mechanical and electrostatic actions can be explained at a molecular level, and they can be understood as nano-machines. These findings fill in details of signaling mechanisms in cells and further the intercellular signal transaction of neural networks. This will facilitate a bottom-up approach to brain research, and SPA occupies a critical position in this combinational approach.

Through the further expansion of IC-EOS, I aim ultimately to develop a fully-automated SPA system for the high-resolution structure determination of ion channels; a system which does not rely on the experience and vast, diverse knowledge of the analyst. Using this system, I hope to advance SPA structure determination toward atomic-level resolution, to determine a larger number of ion channels, and finally to reveal brain mechanisms from the molecular level.

Acknowledgments

I would like to thank Prof. Chikara Sato, who opened a door of opportunity to SPA research and to this Ph. D. work in his laboratory, provided the facilities, and gave valuable advice for publishing journal articles and for writing this dissertation. I also particularly thank Assoc. Prof. Kazuhiro Mio for spending a significant amount of his valuable time in refining the manuscripts for publication and in various academic processes. Special thanks to the members of Prof. Sato's Lab for their support and helpful discussions. I am very thankful to the faculty and to Ms. Yumiko Sakai of the University of Tsukuba for giving advice, support, and help in academic work. I am deeply grateful to Prof. Munetaka Nakata, Tokyo University of Agriculture and Technology, and to Mr. Yusuke Muraoka, CIS Corporation, for support during a difficult time in my life and for guidance for coming back to academic life. My heartfelt appreciation goes to my daughter Alisa, father Junichiro, mother Kuniko, and brother Eiichi for patient support and hearty encouragement. My deepest gratitude goes to my wife Anne for editing all of my English academic writing, and for patient support, hearty encouragement, and comfort. I would like to express deep honor and thankfulness to my late sister Reiko Kuroki, who gave me great encouragement, support, and comfort in the midst of her own courageous fight against cancer.

References

Abbott, A., 2002. The society of proteins. Nature 417, 894-6.

Bartolucci, C., D. Lamba, S. Grazulis, E. Manakova, H. Heumann, 2005. Crystal structure of wild-type chaperonin GroEL. J Mol Biol 354, 940-51.

Bretaudiere, J.P., J. Frank, 1986. Reconstitution of molecule images analysed by correspondence analysis: a tool for structural interpretation. J Microsc 144, 1-14.

Cong, Y., S.J. Ludtke, 2010. Single particle analysis at high resolution. Methods Enzymol 482, 211-35.

Cong, Y., M.L. Baker, J. Jakana, D. Woolford, E.J. Miller, S. Reissmann, R.N. Kumar, A.M. Redding-Johanson, T.S. Batth, A. Mukhopadhyay, S.J. Ludtke, J. Frydman, W. Chiu, 2010. 4.0-A resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement. Proc Natl Acad Sci U S A 107, 4967-72.

Costa, A., T. Pape, M. van Heel, P. Brick, A. Patwardhan, S. Onesti, 2006. Structural studies of the archaeal MCM complex in different functional states. J Struct Biol 156, 210-9.

Dube, P., P. Tavares, R. Lurz, M. van Heel, 1993. The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry. EMBO J 12, 1303-9.

Frank, J., 2006. Three-dimensional Electron Microscopy of Macromolecular Assembles: Visualization of Biological Molecules in Their Native State. Oxford University Press, New York.

Frank, J., 2009. Single-particle reconstruction of biological macromolecules in electron microscopy--30 years. Q Rev Biophys 42, 139-58.

Frank, J., A. Verschoor, M. Boublik, 1981. Computer averaging of electron micrographs of 40S ribosomal subunits. Science 214, 1353-5.

Frank, J., M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj, A. Leith, 1996. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. J Struct Biol 116, 190-9.

Grigorieff, N., 2007. FREALIGN: high-resolution refinement of single particle structures. J Struct Biol 157, 117-25.

Harauz, G., M. van Heel, 1986. Exact filters for general geometry three dimensional reconstruction. Optik 73, 146-156.

Henderson, R., 1995. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. Q Rev Biophys 28, 171-93.

IMAGIC-5, 2011. Single particles in IMAGIC. URL http://www.imagescience.de/manuals/smi.pdf.

Irie, K., T. Shimomura, Y. Fujiyoshi, 2012. The C-terminal helical bundle of the tetrameric prokaryotic sodium channel accelerates the inactivation rate. Nat Commun 3, 793.

Kessel, M., M. Radermacher, J. Frank, 1985. The structure of the stalk surface layer of a brine pond microorganism: correlation averaging applied to a double layered lattice structure. J Microsc 139, 63-74.

Koishi, R., H. Xu, D. Ren, B. Navarro, B.W. Spiller, Q. Shi, D.E. Clapham, 2004. A superfamily of voltage-gated sodium channels in bacteria. J Biol Chem 279, 9532-8.

Langmore, J.P., M.F. Smith, 1992. Quantitative energy-filtered electron microscopy of biological molecules in ice. Ultramicroscopy 46, 349-73.

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

Ludtke, S.J., J. Jakana, J.L. Song, D.T. Chuang, W. Chiu, 2001. A 11.5 A single particle reconstruction of GroEL using EMAN. J Mol Biol 314, 253-62.

Ludtke, S.J., D.H. Chen, J.L. Song, D.T. Chuang, W. Chiu, 2004. Seeing GroEL at 6 A resolution by single particle electron cryomicroscopy. Structure 12, 1129-36.

Ludtke, S.J., M.L. Baker, D.H. Chen, J.L. Song, D.T. Chuang, W. Chiu, 2008. De novo backbone trace of GroEL from single particle electron cryomicroscopy. Structure 16, 441-8.

Mantegazza, M., G. Curia, G. Biagini, D.S. Ragsdale, M. Avoli, 2010. Voltage-gated sodium channels as therapeutic targets in epilepsy and other neurological disorders. Lancet Neurol 9, 413-24.

Marabini, R., I.M. Masegosa, M.C. San Martin, S. Marco, J.J. Fernandez, L.G. de la Fraga, C. Vaquerizo, J.M. Carazo, 1996. Xmipp: an image processing package for electron microscopy. J Struct Biol 116, 237-40.

Maruyama, Y., T. Ogura, K. Mio, S. Kiyonaka, K. Kato, Y. Mori, C. Sato, 2007. Three-dimensional reconstruction using transmission electron microscopy reveals a swollen, bell-shaped structure of transient receptor potential melastatin type 2 cation channel. J Biol Chem 282, 36961-70.

Mio, K., M. Mio, F. Arisaka, M. Sato, C. Sato, 2010. The C-terminal coiled-coil of the bacterial voltage-gated sodium channel NaChBac is not essential for tetramer formation, but stabilizes subunit-to-subunit interactions. Prog Biophys Mol Biol 103, 111-21.

Mio, K., T. Ogura, S. Kiyonaka, Y. Hiroaki, Y. Tanimura, Y. Fujiyoshi, Y. Mori, C. Sato, 2007. The TRPC3 channel has a large internal chamber surrounded by signal sensing antennas. J Mol Biol 367, 373-83.

Nishi, M., S. Komazaki, M. Iino, K. Kangawa, H. Takeshima, 1998. Mitsugumin 23, a novel transmembrane protein on endoplasmic reticulum and nuclear membranes. FEBS Lett 432, 191-6.

Ogura, T., C. Sato, 2001. An automatic particle pickup method using a neural network applicable to low-contrast electron micrographs. J Struct Biol 136, 227-38.

Ogura, T., C. Sato, 2004a. Automatic particle pickup method using a neural network has high accuracy by applying an initial weight derived from eigenimages: a new reference free method for single-particle analysis. J Struct Biol 145, 63-75.

Ogura, T., C. Sato, 2004b. Auto-accumulation method using simulated annealing enables fully automatic particle pickup completely free from a matching template or learning data. J Struct Biol 146, 344-58.

Ogura, T., C. Sato, 2006. A fully automatic 3D reconstruction method using simulated annealing enables accurate posterioric angular assignment of protein projections. J Struct Biol 156, 371-86.

OpenCV, 2011. Open source computer vision library. URL http://sourceforge.net/projects/opencvlibrary/.

Payandeh, J., T. Scheuer, N. Zheng, W.A. Catterall, 2011. The crystal structure of a voltage-gated sodium channel. Nature 475, 353-8.

Payandeh, J., T.M. Gamal El-Din, T. Scheuer, N. Zheng, W.A. Catterall, 2012. Crystal structure of a voltage-gated sodium channel in two potentially inactivated states. Nature 486, 135-9.

PDB, 2013. RCSB Protein Data Bank - RCSB PDB. URL http://www.rcsb.org/pdb/home/home.do.

Penczek, P., M. Radermacher, J. Frank, 1992. Three-dimensional reconstruction of single particles embedded in ice. Ultramicroscopy 40, 33-53.

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

Penczek, P.A., R.A. Grassucci, J. Frank, 1994. The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. Ultramicroscopy 53, 251-70.

Ren, D., B. Navarro, H. Xu, L. Yue, Q. Shi, D.E. Clapham, 2001. A prokaryotic voltage-gated sodium channel. Science 294, 2372-5.

Rosenthal, P.B., R. Henderson, 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J Mol Biol 333, 721-45.

Ryan, D.P., L.J. Ptacek, 2010. Episodic neurological channelopathies. Neuron 68, 282-92.

Sass, H.J., G. Buldt, E. Beckmann, F. Zemlin, M. van Heel, E. Zeitler, J.P. Rosenbusch, D.L. Dorset, A. Massalski, 1989. Densely packed beta-structure at the protein-lipid interface of porin is revealed by high-resolution cryo-electron microscopy. J Mol Biol 209, 171-5.

Saxton, W.O., W. Baumeister, 1982. The correlation averaging of a regularly arranged bacterial cell envelope protein. J Microsc 127, 127-38.

Scheres, S.H., R. Nunez-Ramirez, C.O. Sorzano, J.M. Carazo, R. Marabini, 2008. Image processing for electron microscopy single-particle analysis using XMIPP. Nat Protoc 3, 977-90.

Schroder, R.R., W. Hofmann, J.-F. Menetret, 1990. Zero-loss energy filtering as improved imaging mode in cryoelectronmicroscopy of frozen-hydrated specimens. J Struct Biol 105, 28-34.

Shaikh, T.R., H. Gao, W.T. Baxter, F.J. Asturias, N. Boisset, A. Leith, J. Frank, 2008. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. Nat Protoc 3, 1941-74.

Sigworth, F.J., 1998. A maximum-likelihood approach to single-particle image refinement. J Struct Biol 122, 328-39.

Skrisovska, L., M. Schubert, F.H. Allain, 2010. Recent advances in segmental isotope labeling of proteins: NMR applications to large proteins and glycoproteins. J Biomol NMR 46, 51-65.

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

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

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

Unser, M., B.L. Trus, A.C. Steven, 1987. A new resolution criterion based on spectral signal-to-noise ratios. Ultramicroscopy 23, 39-51.

van Heel, M., J. Hollenberg, 1980. On the stretching of distorted images of two-dimensional crystals, p. 256-260, in: Baumeister, W, (Ed.), Electron Microscopy at Molecular Dimensions, Springer-Verlag, Berlin/New York., pp. 256-260.

van Heel, M., M. Stoffler-Meilicke, 1985. Characteristic views of E. coli and B. stearothermophilus 30S ribosomal subunits in the electron microscope. EMBO J 4, 2389-95.

van Heel, M., M. Schatz, 2005. Fourier shell correlation threshold criteria. J Struct Biol 151, 250-62.

van Heel, M., W. Keegstra, W. Schutter, E.F.J. Van Bruggen, 1982. Arthropod hemocyanin structures studied by image analysis, p. 69-73, in: Wood, E J, (Ed.), Life Chemistry

Reports Supplement 1, The Structure and Function of Invertebrate Respiratory Proteins, EMBO workshop, Leeds, Vol. , pp. 69-73.

van Heel, M., G. Harauz, E.V. Orlova, R. Schmidt, M. Schatz, 1996. A new generation of the IMAGIC image processing system. J Struct Biol 116, 17-24.

van Heel, M., R. Portugal, A. Rohou, C. Linnemayr, C. Bebeacua, R. Schmidt, T. Grant, M. Schatz, 2011. Four-dimensional cryo electron microscopy at quasi atomic resolution: "IMAGIC 4D", p. 624-8, in: Arnold, E, et al., Eds.), Crystallography of Biological Macromolecules, John Wiley and Sons Ltd, New York, Vol. International Tables for Crystallography Volume F, pp. 624-8.

van Heel, M., B. Gowen, R. Matadeen, E.V. Orlova, R. Finn, T. Pape, D. Cohen, H. Stark, R. Schmidt, M. Schatz, A. Patwardhan, 2000. Single-particle electron cryo-microscopy: towards atomic resolution. Q Rev Biophys 33, 307-69.

Venturi, E., K. Mio, M. Nishi, T. Ogura, T. Moriya, S.J. Pitt, K. Okuda, S. Kakizawa, R. Sitsapesan, C. Sato, H. Takeshima, 2011. Mitsugumin 23 forms a massive bowl-shaped assembly and cation-conducting channel. Biochemistry 50, 2623-32.

Yazawa, M., C. Ferrante, J. Feng, K. Mio, T. Ogura, M. Zhang, P.H. Lin, Z. Pan, S. Komazaki, K. Kato, M. Nishi, X. Zhao, N. Weisleder, C. Sato, J. Ma, H. Takeshima, 2007. TRIC channels are essential for Ca2+ handling in intracellular stores. Nature 448, 78-82.

Zhang, J., M.L. Baker, G.F. Schroder, N.R. Douglas, S. Reissmann, J. Jakana, M. Dougherty, C.J. Fu, M. Levitt, S.J. Ludtke, J. Frydman, W. Chiu, 2010. Mechanism of folding chamber closure in a group II chaperonin. Nature 463, 379-83.

Zhang, X., W. Ren, P. DeCaen, C. Yan, X. Tao, L. Tang, J. Wang, K. Hasegawa, T. Kumasaka, J. He, J. Wang, D.E. Clapham, N. Yan, 2012. Crystal structure of an orthologue of the NaChBac voltage-gated sodium channel. Nature 485, 415-54.

Zhu, J., P.A. Penczek, R. Schroder, J. Frank, 1997. Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70S Escherichia coli ribosome. J Struct Biol 118, 197-219.

Abbreviations

2D: Two-Dimensional

3D: Three-Dimensional

NMR: Nuclear Magnetic Resonance

SPA: Single Particle Analysis

EM: Electron Microscopy

SNR: Signal-to-Noise Ratio

CCC: Cross-Correlation Coefficient

FSC: Fourier Shell Correlation

FRC: Fourier Ring Correlation

IC-EOS: Integrated Convergence-Evaluation Oriented System

RMSD: Root-Mean-Square Deviation

TIC-variance: Total Intra-Class variance

SD: Standard Deviation

Na_V channel: Voltage-gated sodium channel

NaChBac: A bacteria-derived Na_V channel

MRA: Multi-Reference Alignment

CTP ratio: Class-To-Particle ratio

GroEL: A chaperon protein

PDB: Protein Data Bank

MG23: Mitsugumin 23; a voltage-dependent, cation-conducting channel

SR: Sarcoplasmic Reticulum

ER: Endoplasmic Reticulum

MSA: Multivariate Statistical Analysis

Echo-correlated method: Echo-Correlated 3D Reconstruction with simulated annealing