Auto-accumulation method using simulated annealing enables fully automatic particle pickup completely free from a matching template or learning data

Toshihiko Ogura and Chikara Sato *

Neuroscience Research Institute and Biological Information Research Center (BIRC), National Institute of Advanced Industrial Science and Technology (AIST), Umezono 1-1-4, Tsukuba, Ibaraki 305-8568, Japan

Received 1 November 2003, and in revised form 5 January 2004

Abstract

Single-particle analysis is a 3-D structure determining method using electron microscopy (EM). In this method, a large number of projections is required to create 3-D reconstruction. In order to enable completely automatic pickup without a matching template or a training data set, we established a brand-new method in which the frames to pickup particles are randomly shifted and rotated over the electron micrograph and, using the total average image of the framed images as an index, each frame reaches a particle. In this process, shifts are selected to increase the contrast of the average. By iterated shifts and further selection of the shifts, the frames are induced to shift so as to surround particles. In this algorithm, hundreds of frames are initially distributed randomly over the electron micrograph in which multi-particle images are dispersed. Starting with these frames, one of them is selected and shifted randomly, and acceptance or non-acceptance of its new position is judged using the simulated annealing (SA) method in which the contrast score of the total average image is adopted as an index. After iteration of this process, the position of each frame converges so as to surround a particle and the framed images are picked up. This method is the first unsupervised fully automatic particle picking method which is applicable to EM of various kinds of proteins, especially to low-contrasted cryo-EM protein images.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Particle picking; Simulated annealing; Single-particle analysis; Pattern recognition; Cryo-electron microscopy

1. Introduction

In single-particle analysis, the signal-to-noise ratio of the electron microscope (EM) images of protein are improved by averaging the projections of similarly oriented molecules, and furthermore, the Euler angles of the averages are determined to construct a three-dimensional (3-D) structure of a protein (Frank, 2002; Sali et al., 2003; van Heel et al., 2000). Because it does not require a crystal, it is advantageous especially for membrane proteins whose crystals are difficult to obtain (Radermacher et al., 1994; Sato et al., 2001; Serysheva et al., 1995). Recently, by using more than 10,000 single-particle images from cryo-EM, the resolution of such analysis has reached a level better than 10 Å even for asymmetric molecules (Matadeen et al., 1999; van Heel et al., 2000). This method is now required to analyze smaller proteins with higher resolution, which requires a large number of particle projections. In the near future, the requirement for the number of projections will reach more than 100,000 for asymmetric protein (Glaeser, 1999; Henderson, 1995). Clearly, this is beyond the ability of manual labor.

Several automatic pickup methods have been devised (Nicholson and Glaeser, 2001). Template matching cross-correlation methods have been mainly used to select very large proteins or viruses since they require reliable templates of projections (Frank and Wagenknecht, 1984;
2. Materials and methods

2.1. Purification of the sodium channels and electron microscopy

The extraction of voltage sensitive sodium channels from the electric organ of *Electrophorus electricus* eels and their purification have been described previously (Sato et al., 1998, 2001). Apoferritin, a soluble protein with a molecular mass of 450 kDa, was kindly provided by Dr. Ichiro Yamashita (Advanced Technology Research Laboratory, Matsushita Electric Ind., Kyoto, Japan). The unstained cryo-EM images of sodium channels and apoferritins were recorded using a JEM3000SFF and a JEM3000EFC electron microscope, respectively, at an acceleration voltage of 300 kV (Fujiiyoshi, 1998). The negatively stained images were recorded using a Hitachi H7000 at an acceleration voltage of 100 kV (Sato et al., 1998). The micrographs were digitized with a Scitex Leaf Scan 45 scanner at pixel sizes of 2.83 and 6.25 Å for cryo and negatively stained images, respectively, at the specimen level. The applied underfocus values for cryo samples ranged from 3.7 to 4.6 μm.

2.2. Image processing

A high-pass filter was applied to the cryo-EM images, the parameters of the filter being set at 195 × 195 pixels for filter size and at 0.005 for cut-off frequency. The average of the pixel intensities in each image was adjusted to 128, which is the median value of 8-bit densities. The sizes of sodium channel and apoferritin images were reduced by half and a quarter, respectively.

To construct a model EM image, 1000 various projections from the randomly oriented sodium channel 3-D model (Sato et al., 2001) were dispersed in wide plane area of 2000 × 5000 pixels. To this image was added various Gaussian noise whose pixel-intensity parameter (σp) was set at a standard deviation (SD) between 5 and 40. The SD of the pixel intensities of the original projection image was almost 4 which is less than those of the noise.

2.3. Initial setting and shifts of frames to search for particles

In the auto-accumulation pickup method, frames which excise subimages to pick up particles are initially set randomly in the whole EM image area. The size of each frame is 64 × 64 pixels, and the total number of frames is set so as to be slightly smaller than the number of particles in the EM image. Each frame has its own coordinate and rotational angle: initially the coordinate (x, y) was set at random values without overlapping with another frame and the angle was set at 0° (Fig. 1A).

In the following algorithm, the frames are shifted randomly and the shifts are further selected. First, one of the frames is selected and shifted randomly: the distance of the shift (Δx, Δy) is randomly extracted from the 2-D normal distribution whose parameter is set at a SD of σx = 7. The frame is further rotated at an angle (Δθ degrees) randomly extracted from the normal distribution of σθ = 30 (Fig. 1B, top left). Second, images are excised from all the frames to calculate the total average image. From the average image, the contrast score is calculated as detailed in Section 2.4, which indicates the degree of overlapping of particle images when all the excised images are stacked. This contrast score is utilized to judge whether to accept the shift or not by the simulated annealing method. After this judgement, the frame shift is accepted or the frame is moved back to the previous position. Third, a new frame is selected, shifted and further rotated again, and the shift is similarly judged by the SA method. The selection of a frame and its shift and judgement are repeated until the number of accepts, Ite2, reaches Ite max 2 or the number of rejects, Ite1, reaches Ite max 1 (Fig. 2). Ite max 1 and Ite max 2 are set at 100 and 20 times the total number of frames, respectively. This whole cycle is repeated for a prefixed number of cycles, C max. In the
annealing iterations, shifts are mostly adopted at the beginning and only the shift toward particle is accepted later by the judgement using the SA algorithm (Figs. 1 and 2).

To avoid the overlapping of frames on the EM map, the following additional rules are also applied just after each shift before the judgement. If the shifted frame overlaps another frame, the shift is cancelled and a new frame is selected. Further, to avoid the stacking of frames at the edge of the entire EM, the following is also applied. If the shifted frame touches or crosses over the edge, the frame is further shifted 100 pixels in the direction perpendicular to the touched edge, toward the center. Thereafter, this shift is judged similarly.

2.4. Calculation of the contrast score which indicates the framing accuracy

From the total average image, the amplitude image of the frequency spectrum is calculated by Fourier-transform. Both the real average image and its amplitude spectrum are used to calculate the contrast score, which is an index to evaluate framing accuracy, i.e., true positive rate, by Eqs. (1)-(6). The score is the sum of both the terms, \( L_{\text{real}} \) and \( L_{\text{freq}} \), in real and Fourier space, respectively.

\[
L = L_{\text{real}} + L_{\text{freq}}.
\]

The left term \( L_{\text{real}} \) is calculated from the real average image in the following steps. First of all, to create the total average image, the framed images are excised and rotated to reverse the sum of their past rotations (Fig. 1B, second row). The rotated images are further averaged to create the total average image, \( d(x, y) \). The average value \( d_{\text{ave}} \) of its pixel intensities is calculated as in Eq. (2): \( x \) and \( y \) represent the horizontal and vertical coordinates of a pixel in the total average image, respectively, whereas \( N \) and \( M \) are the width and height of the frame, respectively. In this case, both \( N \) and \( M \) are 64 pixels. The average value is subtracted from each pixel intensity of the average image \( d(x, y) \), and the calculated image is further masked by a circle \( F_c \) in order to mask its outside area without particle, as in Eq. (3). The diameter of this mask is set according to the size of target molecule: 58 pixels for both sodium channel cryo-EM and model projections, and 45 pixels for both apoferritin cryo-EM and negatively stained sodium channel. Furthermore, in order to avoid the influence of the high frequency which is caused by the highly contrasted edge of this mask, the mask is shaded by the Gaussian averaging kernel, the parameters of which were set to a filter size of 9 \( \times \) 9 pixels and to a SD of 3, in advance. After the masking, each pixel intensity is squared and further summed using the Michaelis–Menten equation, whose parameter is set at a half-saturation constant of \( H \), to yield the \( L_{\text{real}} \) as in Eq. (4). In these experiments, \( H \) was set at 15 for the sodium channel cryo-EM and otherwise to 20. \( P_i \) was set at 128, which is the median value of 8-bit densities. This function is introduced to avoid the effect caused by very high (or low) pixel intensities which are observed locally in the total average, because it suppresses the value of the term calculated from such high (or low) pixel intensity in the right side of Eq. (4). The pixel with such intensities is created very locally and isolatedly, mainly by overlapping of the noise peaks.

\[
d_{\text{ave}} = \frac{1}{N \cdot M} \sum_{x}^{N} \sum_{y}^{M} d(x, y),
\]

2.4. Calculation of the contrast score which indicates the framing accuracy

From the total average image, the amplitude image of the frequency spectrum is calculated by Fourier-transform. Both the real average image and its amplitude spectrum are used to calculate the contrast score, which is an index to evaluate framing accuracy, i.e., true positive rate, by Eqs. (1)-(6). The score is the sum of both the terms, \( L_{\text{real}} \) and \( L_{\text{freq}} \), in real and Fourier space, respectively.

\[
L = L_{\text{real}} + L_{\text{freq}}.
\]

The left term \( L_{\text{real}} \) is calculated from the real average image in the following steps. First of all, to create the total average image, the framed images are excised and rotated to reverse the sum of their past rotations (Fig. 1B, second row). The rotated images are further averaged to create the total average image, \( d(x, y) \). The average value \( d_{\text{ave}} \) of its pixel intensities is calculated as in Eq. (2): \( x \) and \( y \) represent the horizontal and vertical coordinates of a pixel in the total average image, respectively, whereas \( N \) and \( M \) are the width and height of the frame, respectively. In this case, both \( N \) and \( M \) are 64 pixels. The average value is subtracted from each pixel intensity of the average image \( d(x, y) \), and the calculated image is further masked by a circle \( F_c \) in order to mask its outside area without particle, as in Eq. (3). The diameter of this mask is set according to the size of target molecule: 58 pixels for both sodium channel cryo-EM and model projections, and 45 pixels for both apoferritin cryo-EM and negatively stained sodium channel. Furthermore, in order to avoid the influence of the high frequency which is caused by the highly contrasted edge of this mask, the mask is shaded by the Gaussian averaging kernel, the parameters of which were set to a filter size of 9 \( \times \) 9 pixels and to a SD of 3, in advance. After the masking, each pixel intensity is squared and further summed using the Michaelis–Menten equation, whose parameter is set at a half-saturation constant of \( H \), to yield the \( L_{\text{real}} \) as in Eq. (4). In these experiments, \( H \) was set at 15 for the sodium channel cryo-EM and otherwise to 20. \( P_i \) was set at 128, which is the median value of 8-bit densities. This function is introduced to avoid the effect caused by very high (or low) pixel intensities which are observed locally in the total average, because it suppresses the value of the term calculated from such high (or low) pixel intensity in the right side of Eq. (4). The pixel with such intensities is created very locally and isolatedly, mainly by overlapping of the noise peaks.

\[
d_{\text{ave}} = \frac{1}{N \cdot M} \sum_{x}^{N} \sum_{y}^{M} d(x, y),
\]

**Figure 1**

**A** Initial setup

**B** Annealing process with random shift

- Cut image
- Average
- Mask($F_i$)
- FFT
- Mask($F_i$)

**C** After SA pickup

**Figure 6**

**A** Pickup accuracy (%) vs. Cycle and Temperature

**B** Final pickup accuracy (%) vs. Cycle

**C**

- $C_{max}=100$
- $C_{max}=300$
- $C_{max}=500$
- $C_{max}=1000$
The term in Fourier space, $L_{\text{freq}}$, is calculated as follows. By the absolute Fourier transform of the masked average image $(d'_{(x,y)})$, the amplitude spectrum image $|D'_{(u,v)}|$ is calculated. In the spectrum image, the peripheral pixels are further masked by a circle ($F'_{(u,v)}$) to eliminate the high frequency caused mainly by the noise: $u$ and $v$ represent the horizontal and vertical coordinates of a pixel in the spectrum image, respectively, whereas $N$ and $M$ are the width and height of the image, respectively, in Eq. (5). In this case, both $N$ and $M$ are 64 pixels. Since the coordinate $(u_m, v_m)$ represents the center of the amplitude spectrum image in Eq. (6), the center of

\begin{align}
d'_{(x,y)} &= \{d_{(x,y)} - d_{\text{ave}}\} \cdot F_{(x,y)}, \\
L_{\text{real}} &= P_d \cdot \sum_x N \sum_y M d'_{(x,y)}^2 + H^2.
\end{align}
the mask is located in the center of the spectrum image. Therefore, the radius of the mask, \( r_i \), determines the frequency cut-off level. In all the EM presented here, \( r_i \) is set at 6.4 pixels which corresponds to a normalized value of 0.2. By summing all the pixel intensities within the mask in the amplitude spectrum image, the term in Fourier space, \( L_{\text{freq}} \), is calculated as in Eqs. (5) and (6).

\[
L_{\text{freq}} = \sum_{u} \sum_{v} |D(u,v)| \cdot F(u,v),
\]

(5)

\[
F(u,v) = \begin{cases} 
1 & \text{if } r_i > \sqrt{(u-u_m)^2 + (v-v_m)^2} > 0, \\
0 & \text{otherwise.}
\end{cases}
\]

(6)

2.5. SA algorithm to judge whether to accept the shift

The SA method is adopted to rescue the frame from the local minimum which is created where the noise image is accidentally similar to the total average image. The judgment by the SA algorithm consists of several steps (Fig. 2). The contrast score after the shift, \( L_{\text{new}} \), is compared with the score before the shift, \( L_{\text{old}} \). If the shift results in an increase in the score, it is accepted unconditionally. If the shift results in a reduction of the score or in an equal to the score, the value of \( P(\Delta L) \) is calculated based on the following equations to judge whether to accept the shift or not:

\[
\Delta L = L_{\text{old}} - L_{\text{new}},
\]

(7)

\[
P(\Delta L) = \exp \left( -\frac{\Delta L}{T} \right).
\]

(8)

Here, \( \Delta L \) is the change in the total score and \( T \) is the current temperature. A random number \( R \) between 0 and 1 is then generated and compared with \( P(\Delta L) \). The change is accepted if the random number \( R \) is less than \( P(\Delta L) \).

For the annealing, we start at a high temperature of \( T_0 = 200 \) or 500 for model projections or EM images, respectively, then decrease the temperature exponentially by Eq. (9). The time constant \( \tau \) is calculated as in Eq. (10):

\[
T = T_0 \exp \left( -\frac{C}{\tau} \right),
\]

(9)

\[
\tau = -\frac{C_{\text{max}}}{\log \left( T_0 / T_0 \right)}.
\]

(10)

Here, \( T_0, T_0, C \), and \( C_{\text{max}} \) are the initial temperature, the final temperature, the current number of cycles, and the maximum number of cycles, respectively. The temperature is decreased at the end of each cycle.

2.6. Image analysis system

All the calculations in the auto-accumulation method were performed with the image-processing toolbox of Matlab Version 6 (MathWorks) on a personal computer (Pentium 4: 3 GHz, 2 Gbytes RAM) running Windows 2000. Every system was programmed using the Matlab script M-files.

3. Results

3.1. Fully automatic pickup algorithm by the auto-accumulation method

A schematic outline of our proposed pickup method, an auto-accumulation method combined with the SA algorithm, is shown in Fig. 1, and a flow diagram is shown in Fig. 2. Initially, square frames to excise sub-images are positioned randomly on a whole EM image (Fig. 1A). One of the frames is selected randomly and its position and rotational angle are also shifted randomly (Fig. 1B, top left). To evaluate the shift and the rotation, a contrast score is calculated to judge whether to accept the shift or not, as follows. The images of all the framed area are excised and rotated to reverse the sum of its past rotations (Fig. 1B, second row). All these images are then averaged to create the total average image. If scarcely any particles exist in the frames, only a very small contrast appears in the average image. If a particle is surrounded by one of the frames, a small contrast appears in the average image. If many particles are surrounded by frames, more contrast appears in the average image. The contrast increases with an increase in the number of particles surrounded by frames. Furthermore, by the overlapping location and orientation of particle at a corresponding position in each frame, more contrast appears in the average. To use this contrast as an index of framing accuracy, the contrast score at the current iteration is calculated from both the total average image and its Fourier transform as described in Materials and methods. The absolute Fourier transform of a image is called the amplitude spectrum image and it shows the frequency distribution in the image: low frequency appears as a pixel near the center and high frequency appears peripherally. By the circular masking of the spectrum image, we can neglect the peripheral pixels due to the high frequencies which are mainly caused by noise. After the masking, the unmasked pixel intensities of the amplitude spectrum image increase as the contrast of the original total average increases. Therefore, the average of these intensities can also be utilized as an indicator of the particle image location at the corresponding position in each frame. From this value and the value from the real image, the contrast score is further calculated as an indicator of the current pickup accuracy: each shift of the frame is evaluated with this score using the SA method. The SA method is utilized here to rescue frames from the pseudo positives which are local minimums and to enhance convergence. If the
shift of the frame results in an increase in the score, it is accepted unconditionally. If the shift results in a reduction of the score or in an equal to the score, the value of \( P(\Delta L) \) is calculated based on the Boltzmann equation from both the score difference and the current temperature, in order to judge whether to accept it or not. A random number \( R \) between 0 and 1 is then generated and compared with \( P(\Delta L) \). The shift is accepted if the random number \( R \) is less than \( P(\Delta L) \). If \( R \) is higher than \( P(\Delta L) \), the frame is shifted back to the previous position. Thereafter, a new frame is selected and its shift, rotation and further judgement are repeated similarly. This process is iterated as described in Section 2. During the annealing cycles, the temperature decreases exponentially. At high temperature, the Boltzmann equation is almost 1 for a wide range of scores and most of the shifts are accepted. Therefore, the frames move relatively freely on the EM image regardless of the score during the early cycles. Furthermore, by this free movement, a frame which is caught by pseudo-positive, e.g., accidentally matched noise pattern with the average, can escape from the position in the following annealing process. When the reduction of the score, \( \Delta L \), is small, the probability of accepting the shift is high. The probability of acceptance decreases as the temperature decreases. As the temperature further decreases, the probability value \( P(\Delta L) \) also decreases gradually toward 0. Therefore, each frame begins to move toward the position where a higher score is acquired (Fig. 1B, top). When a frame is caught by a pseudo-positive, if the temperature is medium, the frame still has the potential to escape from the position because it can move to the surrounding position where the contrast score decreases. As the temperature further decreases, the frame reaches the particle and continues to vibrate there. As the temperature decreases to near 0, every frame successfully surrounds particle and its vibration gradually ceases (Fig. 1C). Consequently, a few hundred particles are automatically picked up by the frames after the annealing cycles.

3.2. Pickup accuracy of the auto-accumulation method when applied to model-projection images immersed in various noise

The pickup ability of the auto-accumulation method combined with the SA method was investigated, first using model images in which projection images were immersed in Gaussian noise whose pixel-intensity parameter was set at a standard deviation (SD) of \( \sigma_n = 10 \). The SD of the pixel intensities of the original projection image was almost 4, smaller than that of the noise. The white and black circles indicate the frames with and without a particle, respectively. To judge whether a frame possesses a particle or not, the distance between its accumulation center and the center of its nearest particle is measured: a frame with a particle is defined as one whose distance is within 10 pixels. (A) Initial frame setup at random positions. Most frames do not possess particles. Both the total numbers of frames and cycles were set at 500. The parameters of the temperature, \( T_0 = 200 \) and \( T_e = 2 \), were set as described in Section 2. (B) Early stage in annealing process. In most cases, the frames are shifted randomly on a model EM in a manner similar to Brownian movement of a molecule. This is because the frame shift in any direction is almost always accepted since the temperature is high. Therefore, most frames do not have a particle yet. (C) Middle stage of annealing process. As the temperature decreases, the frames are gradually caught by particles and scarcely shift away from the particles. Almost half of the frames hold a particle. (D) After annealing. Every frame possesses a particle. The current number of cycles and the temperature are shown above each image.

Fig. 3. Changes in the positions of frames at annealing cycles by the auto-accumulation method when applied to a model EM. The time course of the frame positions is shown from (A) to (D). A part of the whole model EM image is shown. To construct the model image, the projections from the sodium channel 3-D model are immersed in Gaussian noise whose pixel-intensity parameter was set at a standard deviation (SD) of \( \sigma_n = 10 \). The SD of the pixel intensities of the original projection image was almost 4, smaller than that of the noise. The white and black circles indicate the frames with and without a particle, respectively. To judge whether a frame possesses a particle or not, the distance between its accumulation center and the center of its nearest particle is measured: a frame with a particle is defined as one whose distance is within 10 pixels. (A) Initial frame setup at random positions. Most frames do not possess particles. Both the total numbers of frames and cycles were set at 500. The parameters of the temperature, \( T_0 = 200 \) and \( T_e = 2 \), were set as described in Section 2. (B) Early stage in annealing process. In most cases, the frames are shifted randomly on a model EM in a manner similar to Brownian movement of a molecule. This is because the frame shift in any direction is almost always accepted since the temperature is high. Therefore, most frames do not have a particle yet. (C) Middle stage of annealing process. As the temperature decreases, the frames are gradually caught by particles and scarcely shift away from the particles. Almost half of the frames hold a particle. (D) After annealing. Every frame possesses a particle. The current number of cycles and the temperature are shown above each image.
immersed in various levels of noise and further using the EM images of proteins. Fig. 3 shows the process for detecting particles when applied to a model image. In the image, the various projections from a sodium channel 3-D model (Sato et al., 2001) are immersed in Gaussian noise whose parameter was set at a SD of $\sigma_n = 10$. The initial frames are positioned randomly, where most frames do not include particles (Fig. 3A). After 200 annealing cycles, the temperature decreases and some of the frames begin to be caught by particles (Fig. 3B). As the temperature decreases further, more frames are caught by particles (Fig. 3C). By the end of the annealing, every frame contains a particle projection (Fig. 3D).

We further tested the noise resistance of this method using projection images immersed in heavier noise. To the original projection images (Fig. 4A) were added different levels of noise images whose pixel-intensity parameter were set at a SD, $\sigma_n$, between 20 and 40. In these test images, to the naked eye, the projections seem to almost disappear (Figs. 4B–D). After the annealing, almost all the frames contain a particle at the noise level of $\sigma_n = 20$ although the sodium channel looks very faint (Fig. 4B). At a greater noise level, this method still

---

**Fig. 4.** The selection accuracy when applied to a model EM immersed in various levels of noise. A part of the model EM image is shown. (A) Original model image without noise. Various projections from the sodium channel 3-D model are dispersed randomly, in which the SD of its pixel intensities was 4. (B) Frame positions after annealing when applied to the model EM image immersed in Gaussian noise whose pixel-intensity parameter was set at a SD $\sigma_n$ of 20, which is almost fivefold higher than the SD of the original projections. After annealing, most frames have a particle. (C) Frame positions after annealing when applied to an image with Gaussian noise whose SD $\sigma_n$ is 30. At this noise level, projections are hard to recognize by the naked eye. Even with this noise, approximately 80% of the frames surround a particle. (D) Frame positions after annealing when applied to an image with Gaussian noise $\sigma_n = 40$. At this noise level, projections are difficult to recognize. Even at such a noise level, almost half of the frames have a particle. The parameters adopted here were the same as those in Fig. 3. (E) The relationship between the noise level of the EM image and the final pickup accuracy. The pickup accuracy decreases as the noise increases. The accuracy is the average of five independent annealings, and the error bars show variations (standard deviation, SD).
detects the particles; however, its pickup accuracy gradually decreases as the noise increases (Figs. 4C–E). Even at an extremely high noise level of $\sigma_n = 40$, particles which are very difficult to recognize could be detected by this method (Fig. 4D), which reveals its robustness against noise. In order to determine the accuracy of the auto-accumulation method, the centers of the projections in the model EM were marked in advance. After annealing, the distance between the auto-accumulation center of each frame and the center of the nearest particle was measured. If the distance was less than 10 pixels, the frame was categorized as successful. The ratio of such frames to the total frames was defined as the pickup accuracy. The pickup accuracies after annealing when applied to model EMs of various noise levels are shown in Fig. 4E. The accuracy decreases as the noise increases, although the accuracy is approximately 50% even at the highest noise level of $\sigma_n = 40$. Using these model EM image in Figs. 4B–D, the changes in the total average image at annealing cycles are shown in Fig. 5A. Starting from the image of noise, particle features gradually emerge as the temperature decreases (Fig. 5A, Cycle: 150–200). The total average image at the end of the annealing (Fig. 5A, right end) is similar to the average of the sodium channel projections which are evenly projected from the 3-D model (Fig. 5A, top right). The similarity of the total average after annealing to the model projections average depends on the noise level in the model EM images. At a noise level of $\sigma_n = 20$, very clear particle features appeared (Fig. 5A)

![Fig. 5. Changes in the total average of framed images and its pickup accuracy during annealing cycles when applied to model EM images with various levels of noise. (A) The changes in the features of the total average after a specific number of learning cycles shown above. In the applied model EM images, the parameters of noise are set at a SD of $\sigma_n = 20$ (top row), 30 (middle row), and 40 (bottom row) as shown in Figs. 4B–D, respectively. The time course of the total average image is shown from left to right for the number of cycles indicated at the top. Initially, the total average shows only the noise image (first column on the left). After 150 cycles, an obscure particle feature emerges. As the temperature decreases further, the feature becomes clearer with more contrast. After the annealing, the average reflects the particle feature. Especially, when applied to an EM image with a noise level $\sigma_n$ of 20 or 30, the average is similar to the total average of the model projections, as shown in the top right; however, when applied to a model EM with a noise level $\sigma_n$ of 40, a part of the particle feature is missing in the average image (third row, right end). (B) The changes in the contrast score and the temperature at annealing cycles. As the temperature decreases exponentially, the contrast score increases sigmoidally at every noise level in (A). (C) The changes in the pickup accuracy at annealing cycles. The pickup accuracy is defined as the percentage of frames with particles in all frames. For the beginning 100 cycles, the pickup accuracy is approximately 3%. After this latent period, the accuracy also increases sigmoidally like the contrast score. After 300 cycles, the accuracy gradually reaches a plateau. The final pickup accuracy, when applied to the EM images with noise levels $\sigma_n$ of 20, 30, and 40, were 100% (thin line), 80% (thick line), and 45% (gray line), respectively.}
upper row), whereas at a higher noise level, an unclear image emerged which lacked part of the particle (Fig. 5A, bottom row). As the temperature decreases exponentially, the contrast score which reflects the contrast of the total average image increases sigmoidally as cycles: it increases gradually for the first approximately 100 cycles, and then drastically rises for the next 200 cycles (Fig. 5B). The sharp rising phase of the score is between 180 and 270 cycles during which temperature decreases from 38 to 16. The rising rate then gradually decreases and the score reaches a final plateau. The height of the plateau depends on the noise level of the model EM: the plateau becomes higher as the noise increases. The changes in the pickup accuracy also exhibit a sigmoidal curve (Fig. 5C) as is the case with the contrast score (Fig. 5B). In the first 150 cycles, the accuracy remains at just 3% and, thereafter, increases drastically. As the temperature decreases, it reaches a final plateau, the height of which is markedly dependent on the noise level as shown in Fig. 4E.

In the annealing process, each frame escapes from the local minimum by its rather free movement at medium temperature as described in Section 2 and reaches a particle position which is one of the global minimums. This is similar to crystal formation in nature, in which the molecules move to appropriate positions to form a crystal. In crystal formation, the molecular movements are caused by free energy. If the temperature decreases slowly, the free energy also decreases slowly and a bigger crystal tends to form. By this slow annealing, the molecules decrease their free movements gradually and accurately positions themselves to form a crystal. Likewise, in the present method, a similar relationship between the pickup accuracy and temperature change is expected. When the temperature is decreased at various rates, the accuracy is influenced as shown in Fig. 6A, using the model EM image with a noise level $\sigma_n$ of 40. At every rate, the accuracy increases almost sigmoidally and reaches to its final plateau. The final accuracy becomes better as the annealing becomes slower. By slow temperature reduction during 1000 cycles (Fig. 6B), the pickup accuracy is improved by 30–55%. The total average of the selected images after the annealing is shown in Fig. 6C: the images in each row represent the averages of five independent annealings. The temperature decreased from 200 to 2 during the cycles shown at the left of each row. When the temperature decreased quickly ($C_{\text{max}} = 100$), the resulting average image varied and partially resembles a particle (Fig. 6C, first row). In contrast, when the temperature decreased slowly ($C_{\text{max}} = 1000$), the average image converged into a form (Fig. 6C, fourth row) almost identical to the total average of model projections (Fig. 5A, top right). In total, the more slowly the temperature decreases, the better and more identical the projection average.

### 3.3. Pickup accuracy of auto-accumulation method when applied to EM images

Next, we applied the auto-accumulation pickup method to negatively stained EM images and further to cryo-EM images obtained by transmission electron microscopy (TEM). The negatively stained sodium channel has a strong contrast and is easily recognized by the human eye (Sato et al., 1998). After the annealing, most frames are caught by particles as shown in Fig. 7A. In the total average of the excised images, a round white core arises at the early stage, and this gradually develops into a shape which is similar to the average of the sodium channel projections (Fig. 7B). The changes of both
the temperature and the contrast score during annealing cycles are shown in Fig. 7C.

Because cryo-EM images have low contrast and relatively high noise, the recognition of a particle is generally difficult. Here, the present method is applied to a cryo-EM image of apoferritin, a soluble protein with a molecular mass of 450 kDa (Fig. 8). The image is comparatively easy to recognize (Fig. 8A). In this experiment, the number of frames is set to be as many as 250, which is almost 90% of the total number of particles in this area. After annealing, most frames are successfully caught by the apoferritin particles. The apoferritin-image formation in the total average during the annealing cycles is shown in Fig. 8B. After 180 cycles, a faint image like a ring appears and its contrast becomes stronger as the temperature decreases. Its average image is quite similar to a projection of spherical apoferritin (Fig. 8B, right end). The changes of both the temperature and the contrast score during annealing cycles are shown in Fig. 8C.

The auto-accumulation method was also applied to Cryo-EM images of the sodium channel (Fig. 9). The particle with a molecular mass of 200 kDa shows very low contrast and is very hard to recognize (Fig. 9A) similarly to the cases of Figs. 4C and D. This is because

---

**Fig. 8.** Selection accuracy when applied to micrographs of unstained apoferritins recorded by cryo-EM. (A) Typical particle selection map using apoferritin. Again most frames were caught by soluble apoferritin proteins which have a molecular mass of 450 kDa. In this experiment, the number of frames was set at as many as 250, which is 90% of the total number of particles. The other parameters were set as in Fig. 7. The result clearly shows high efficiency of present method in picking up most particles in the electron micrograph. The scale bar represents 200 Å. (B) Changes in the features of the total average image after a specific number of annealing cycles. The number of cycles is shown above each average. After 180 cycles, a ring-like feature appears, which is similar to a particle projection of this spherical molecule (right end). The contrast of this feature increases as the temperature decreases. A projection from the 3-D model (Granier et al., 1997) is shown on the right end. The scale bar represents 100 Å. (C) Changes in the contrast score and the temperature at annealing cycles. Again, as the temperature decreases exponentially, the contrast score increases sigmoidally. Each arrowhead at the top corresponds to the number of cycles at the top of (B).

---

**Fig. 9.** Selection accuracy when applied to micrographs of unstained sodium channels recorded by cryo-EM. (A) Typical selection map using the sodium channel. The contrast of the protein in the cryo-EM image is weak because the protein is solubilized in a buffer which contains 20% glycerol (Sato et al., 2001). In spite of the faint images of membrane protein with a molecular mass of 200 kDa, more than half of the frames are caught by the proteins after annealing. The total number of frames was set at 500, whereas the maximum number of cycles, $C_{\text{max}}$, was 1000. The scale bar represents 200 Å. (B) Changes in the features of the total average image after a specific number of learning cycles. The number of cycles is shown above. After the annealing, a particle feature which is partially similar to the total average of the sodium channel projections (Fig. 5A, top right) appeared, although it was very noisy. It is also comparable to the total average calculated using the model EM image with very heavy noise (Fig. 5A, $\sigma_s = 40$ and Fig. 6C). The scale bar represents 100 Å. (C) The changes in the contrast score and the temperature at annealing cycles. Again, as the temperature decreases, the contrast score increases sigmoidally. Each arrowhead at the top corresponds to the number at the top of (B).
the surrounding buffer contains 20% glycerol and detergent (Sato et al., 2001). However, by using the auto-accumulation method, almost half of the frames are successfully caught by the particles. An image formation in the total average is shown in Fig. 9B. The image similar to the sodium channel projection appeared, although it is very noisy. It also resembles the total average which was created using the model EM with very heavy noise (Fig. 5A, third row, \( \sigma_r = 40 \), and Fig. 6C). The noisy average suggests the misalignments of the particles in the frames. In the annealing, the contrast score increases sevenfold, from \( 2.5 \times 10^4 \) to \( 1.8 \times 10^5 \) (Fig. 9C), reflecting drastic improvement of the positionings of the frames.

### 3.4. Self-accumulation mechanism of the particle images in the total average of framed images

In the annealing process of the auto-accumulation method, both the position and angle of each frame are induced to change toward increasing the contrast of the total average image. Therefore, the image formation process of the total average image is one of the important keys to understand how an projection average is formed by the simple algorithm presented here. An example of image formation in the total average is shown in Fig. 10A, when applied to each of the EM images in Fig. 4B (\( \sigma_r = 20 \)), 7, 8, and 9. First, a contrasted white core arises near the center of the frame (Fig. 10A, left end). The position of the core is not always centered because the position for accumulation in the frame is not fixed in the present algorithm. As the temperature decreases, the contrast of the core increases and its surrounding structure gradually emerges, (Fig. 10A, 0.2–0.3). Furthermore, with regard to this feature, asymmetry clearly appears (Fig. 10A, third row) and the feature grows to a structure similar to the particle projection average. The quality of the final average image is related to its picking accuracy as shown in Figs. 10A and B. When this method is applied to the model data or Cryo-EM images of apoferritin, the total average is clear and similar to the projection average of the target molecule. With such an image of the total average, the pickup accuracy is high. When applied to the cryo-EM image of the sodium channel, the particle feature of the total average image still contains fuzzy surroundings and its pickup accuracy remains at 50% (Figs. 10A and B, fourth row). However, the accuracy of 50% is sufficiently high to be classified by the modified GNG method (Ogura et al., 2003) to extract only the real positive images.

We further applied the auto-accumulation pickup method to the keyhole limpet hemocyanin Cryo-EM image (Orlova et al., 1997) which is available on the Internet. After the annealing, most frames are caught by particles as shown in Fig. 10C with a pickup accuracy of 95%. The total average image created by this method after the annealing is shown in Fig. 10D.

### 4. Discussion

The principal strategy of single-particle analysis is to make a particle emerge by averaging, even if the details of each particle image are too faint to recognize (Frank, 2002; van Heel et al., 2000). Averaging improves the signal-to-noise ratio because noise is different in each particle image. When we have particle projections in the center of all excised images, the total average image must reflect the shape of the molecule. Accordingly, the contrast of the total average image is higher than the noise average image and becomes even higher when the excised images are aligned rotationally. In combination with shifting of the frames for random excision on the EM, this notion is utilized to indicate the frames which successfully surround particles. To evaluate the alignment, another key idea is introduced: Fourier transform, the amplitude spectrum of the average image, is calculated and the sum of its pixels is also utilized to compute the contrast score. When the particles in a whole EM image are abundant enough to be picked up, this strategy was basically proved to be very effective as a completely automatic pickup. However, a simple combination of these methods failed to reach the particles because the average image frequently matched the noise pattern during the shifts accidentally. To overcome this difficulty, we introduced the SA method, the strategy of which can go through those local minimums. Here, the calculation starts from the high temperature, implying high free energy which enables each frame to shift in any direction. By this energy, the trapped frame can escape from such a pseudo-positive and reach a real positive. The SA method has potential for wide utilization in various steps of single-particle analysis such as our application in the modified growing neural gas (GNG) image classification method (Ogura et al., 2003).

This method is similar to real molecular movements in crystal formation, although real molecules are aligned in 3-D space in natural crystal formation, while 2-D images are aligned in piled 2-D frames in this method. To enable this artificial crystallization in computer, only the contrast of its total average is utilized as an index of the right positionings of frames. Therefore, we do not have to know the shape of the molecules in advance. Namely, this program automatically picks up particles only by the parameter settings of the self-gathering environment. There are two categories of parameters, those which determine the contrast score, e.g., the size of frame and the cut-off level of high frequency in the average image and those which determine the temperature change, e.g., the initial and final temperature, and the maximum number of cycles. As with crystal formation
in nature, an appropriate temperature change is important for this computational crystal formation. If the temperature decreases too rapidly, the shifts of each frame are restricted in a small area and its free movement is lost before a particle appears in the moving frame. Accordingly, the final pickup accuracy becomes low. In contrast, if the temperature decreases too slowly, the annealing reaches the convergence with high accuracy although the calculation time becomes long. In the present experiment, the number of cycles was restricted to 500 or 1000. During the period of annealing, the temperature decreases exponentially from 500 to 5. In this condition, the pickup is accurate as shown and the calculation to pick up 500 particles on the typical PC

Fig. 10. Changes in the particle features accumulated in the total average of framed images and its pickup accuracy after the annealing when applied to various kinds of EM images. (A) The changes in the particle features of the total average after a specific number of cycles. The EM images utilized were those in Fig. 4B (n = 20), 7, 8, and 9. The time course of the average is shown from left to right at the relative contrast score indicated at the top. The relative contrast scores are normalized linearly to their maximum value obtained at the end of annealing. At the relative contrast score of 0.1 in the rising phase of the score, each average has just a small core near the center (first column on the left). The core gradually changes to a dome-like structure at a relative score between 0.2 and 0.3. At a higher score between 0.5 and 0.9, it gradually develops a structure like that of the particle. The pixel intensities in each average image are adjusted to fit to the full range of 8-bit gray scale densities. Each scale bar in the right end column represents 100 Å. (B) The pickup accuracy after the annealing. It is averaged from those after five independent annealings and the error bar shows the variation (SD). The similarity between the total average and the particle projection average is highly related to its pickup accuracy. (C) Typical particle selection map using keyhole limpet hemocyanin (Orlova et al., 1997) after the annealing. The parameters were set as in Fig. 7, except the number of frames which was set at 40 for each micrograph. The average accuracy was calculated using five micrographs with a defocus of 1.5 μm. The pickup accuracy was 95 ± 2.5%. The accuracy is the average of five independent annealings. (D) The total average image after the annealing in (C).
can be finished in one day. The average image after annealing is clear and can be utilized as a starting reference in the following alignment in the single-particle analysis. The index of the contrast score can be generally adopted in multi-reference alignment and possibly in other processes in single-particle analysis, as well as in other types of image processing. This auto-accumulation method itself also can be directly used for various kinds of images.

In conclusion, the auto-accumulation method provides a completely automatic pickup tool that enables a detection of a very low-contrasted protein image in a high level of noise. This method does not require any structural information of the molecule or learning data in advance except the approximate dimensions of the particle, and is characterized by simple parameter settings. On the other hand, its pickup speed and accuracy presented here is not as high as the pickup method based on a three-layer pyramidal-type artificial neural network (NN) (Ogura and Sato, 2001, 2004), although it requires training using a few hundred manually selected particles in advance. This necessity of the manual labor would be solved by the auto-accumulation method. Using particle images picked up here, the modified GNG method would classify a target molecule (Ogura et al., 2003). Using selected images as a learning data set, the NN would pick up particles with higher accuracy. The combination of these two methods would provide a completely automatic particle pickup strategy with both very high accuracy and speed.

Recently, single-particle analysis can be utilized to reveal the dynamics of the molecule and even protein complex formations are beginning to be solved, because it does not require a crystal. However, to reconstruct 3-D structure, much labor is required to pick up particles, which is not easy to do using very faint cryo-EM images, and subsequent image analysis also takes a long time. Many of these problems at the beginning of single-particle analysis would be solved by the present method in combination with the NN method (Ogura and Sato, 2001, 2004), which would contribute to the full automation, speed-up and improvement of resolution of single-particle analysis.

Acknowledgments

The authors would like to express their cordial thanks to Dr. Yoshinori Fujiyoshi (Kyoto University, Japan) for cryo-microscopy and discussions, to Dr. Kenji Iwasaki (Osaka University, Japan) for cryo-microscopy, and to Dr. Andreas Engel (MIH, Biozentrum, Basel, Switzerland) for his suggestions. Dr. Ichiro Yamashita (ATRL, Matsushita Electric Ind., Japan) kindly provided the apoferritin. This work was supported by a grant from Japan New Energy and Industrial Technology Development Organization (NEDO), and by grants from the National Institute of Advanced Industrial Science and Technology (AIST).

References


reconstitution used to visualize the skeletal muscle calcium release channel. Nat. Struct. Biol. 2, 18–24.
