

Assessment of electron crystallographic data obtained from two-dimensional crystals of biological specimens

Vinzenz M. Unger

Department of Molecular Biophysics and
Biochemistry, Yale University, 333 Cedar Street,
New Haven, Connecticut 06510, USA

Correspondence e-mail:
vinzenz.unger@yale.edu

Received 24 March 2000

Accepted 9 August 2000

Over the past few years, an increasing number of electron crystallographic studies using two-dimensional crystals have shed light on the structure of biologically important macromolecules. Steady progress in the development of specimen-preparation techniques and image-processing tools enable researchers to achieve resolutions in the range of 5–10 Å almost routinely. However, reaching near-atomic resolution remains a formidable task that is likely to require several years. Without doubt, this process will become far less time-consuming as methods are improved further. However, the immediate future is more likely to be dominated by structures solved to an intermediate level of resolution. Since the reliability of such structures is more difficult to assess than that of density maps at near-atomic resolution and as the popularity of electron cryo-microscopy increases it becomes more important to define standardized criteria for the evaluation of electron crystallographic data. This article discusses some of the relevant issues with the aim of stimulating further discussion about the assessment and presentation of electron crystallographic data.

1. Introduction

In 1990, Henderson and coworkers reported the 3.5 Å structure of the light-driven proton pump bacteriorhodopsin (Henderson *et al.*, 1990). This breakthrough summarized two decades of intense effort spent on the development of methods that have established electron crystallography as a new tool for analyzing the macromolecular structure of biological specimens at near-atomic resolution. Further validation for the suitability of this approach emerged when in 1994 Kühlbrandt and colleagues derived the 3.4 Å structure of another membrane protein, the plant light-harvesting complex II (Kühlbrandt *et al.*, 1994). To add to the excitement, the 4 Å tubulin structure solved recently by Nogales *et al.* (1998) demonstrated that electron microscopic high-resolution studies are not limited to two-dimensional (2D) crystals of integral membrane proteins. Nevertheless, to date electron crystallography has found its major application in the structure analysis of integral membrane proteins, for which well ordered three-dimensional (3D) crystals cannot be easily obtained. Projection density maps at 3.5–4 Å resolution of aquaporin-1 (Jap & Li, 1995), microsomal glutathione transferase (Schmidt-Krey *et al.*, 1999) and the bacterial Na⁺/H⁺-antiporter NhaA (Williams *et al.*, 1999) indicate that the generation of high-resolution data will become more common in the future. Furthermore, a number of 3D maps at intermediate resolutions [aquaporin-1 (Cheng *et al.*, 1997; Li *et al.*, 1997; Walz *et al.*, 1997), rhodopsin (Unger, Hargrave *et al.*

1997), H⁺-ATPase (Auer *et al.*, 1998), photosystem II (Rhee *et al.*, 1998), gap-junction channels (Unger *et al.*, 1999), NhaA (Williams, 2000)] provided the first visualization of the overall fold of their respective protein families. Although at modest resolution, these structures provide important constraints for model building (see, for example, Baldwin *et al.*, 1997) and biochemical structure–function studies. As the field of electron crystallography grows and new techniques are being developed extending its use to soluble proteins through crystallization on lipid monolayers and lipid nanotubes (Uzgiris & Kornberg, 1983; Kubalek *et al.*, 1994; Wilson-Kubalek *et al.*, 1998; Asturias & Kornberg, 1999), new standards will emerge defining the scopes and publication policies of electron crystallographic studies. This paper is not a review on 2D crystallization techniques, nor of the structures that have been published. The scope of this contribution is to look at how results in our field have been presented in the past. This may serve as a starting point for further discussions defining standardized criteria for the evaluation of electron crystallographic work, similar to those that exist in X-ray crystallography. Identifying such standards may enhance the transition of electron crystallography from being an exotic application to becoming a routine tool in the area of structural biology.

2. Discussion

2.1. Some general remarks

A decade ago, the capability to perform cryo-EM was limited to a few well endowed laboratories. However, as the potential of cryo-EM gains more recognition, new facilities have been established at an accelerated pace during the past three years. This rapid growth demonstrates that cryo-EM is at the cutting edge of biomedical research. However, fast growth also holds the danger that investigators do not take sufficient time to familiarize themselves with required background knowledge. This can become a problem, for example in the analysis of low-dose micrographs of 2D crystals, where the relevant software is largely without annotations. In the hands of inexperienced users, this may involuntarily result in errors that are not easily detected once a ‘completed’ data set is presented. Furthermore, to date the evaluation of data is largely subject to personal preferences. To give just two examples: any 3D reconstruction from 2D crystals requires data from tilted specimens because each individual image only provides projection data for the particular view dictated by the tilt geometry. Hence, electron crystallographic data will always be anisotropic owing to the limited tilt capabilities of the cryo-specimen holders. This raises the question: what resolutions/tilts are required for the generation of a reliable representation of one’s specimen? Based on discussions with colleagues, views on this issue are quite diverse. Similarly, the electron microscope can be used for collection of electron-diffraction patterns as well as images. However, electron-diffraction data may not be available in cases where 2D crystals are small or poorly ordered. In these cases, structure-factor amplitudes and

phases are obtained from the images. In contrast to the phases, however, image-derived amplitudes are strongly influenced by imperfections in the imaging process. The resulting fall-off of the amplitudes is very pronounced, is resolution dependent and is commonly overcome by the use of inverse *B* factors that are large compared with those used in X-ray crystallography. This raises the question of what magnitude of image sharpening is permissible – an issue that is less disputed amongst electron crystallographers but frequently raises concerns from reviewers who are more familiar with X-ray crystallographic data. The following paragraphs will provide a short discussion of some of these issues. The discussion will be focused on intermediate resolution data, *i.e.* those in the 5–10 Å region, where quality assessment is the most difficult.

2.2. Evaluation of projection data

The first step in solving a structure from 2D crystals is the generation of an end-on view of the molecule by recording data from untilted crystals. A number of such projection structures have been published stand-alone (for instance, Schertler *et al.*, 1993; Karrasch *et al.*, 1995; Mitra *et al.*, 1995; Jap & Li, 1995; Walz *et al.*, 1995; Rhee *et al.*, 1997; Unger, Kumar *et al.*, 1997; Schmidt-Krey *et al.*, 1999; Williams *et al.*, 1999; Koning *et al.*, 1999). As indicated in Table 1, a projection structure allows conclusions about the in-plane dimensions of the protein (complex), any inherent symmetry within the molecule (complex), its oligomeric state and formal crystallographic packing arrangement. Other insights derived from such maps, especially referring to the type and number of secondary-structure elements, are mostly speculative. The sole exceptions are α -helices oriented (almost) perpendicular to the plane of the untilted specimen. If not obscured by other parts of the molecule, such α -helices will appear as individual round density features ~ 10 Å in diameter and 10–12 Å apart from any adjacent density features. This characteristic appearance has allowed the identification of α -helices in the projection maps of a number of integral membrane proteins. If more than one α -helix can be observed in projection, then a comparison may be possible with other proteins of similar topology. For instance, the projection density maps of the G-protein-coupled receptor rhodopsin and the light-driven proton pump bacteriorhodopsin (Schertler *et al.*, 1993) revealed that the seven transmembrane α -helices are packed differently in each of these proteins. In exceptional cases, it might even be possible to analyze conformational changes of a protein (complex) or to locate certain parts of a protein by introduction of heavy-atom labels or binding of antibody fragments. Such data will continue to be of significant interest, especially for membrane proteins, where the structures are more restricted owing to the constraints imposed by the lipid bilayer. This raises the question of how to present projection data so that their quality can easily be appreciated.

2.2.1. Phase statistics. Of all the items presented upon publication of a projection density map, phase statistics provide the best evaluation criterion for the quality of the image data, because the phase errors directly reflect how well

Table 1
Validity of conclusions drawn from projection density maps.

Conclusions	Validity
<i>x,y</i> dimensions of molecule	Yes
Inherent symmetry	Yes
Crystallographic symmetry	Yes, program <i>ALLSPACE</i> for cryo-data. Note that <i>ALLSPACE</i> may not provide the correct answer for negatively stained specimens.
Oligomeric state of protein	Yes. However, physiological significance (if any) has to be demonstrated by other means.
Resolution	Yes, can be proven by phase statistics.
Type of secondary structure	No, except for α -helices oriented (almost) perpendicular to direction of view and not obscured by large loops/ectodomains/other parts of molecule.
Number of secondary-structure elements	No, except for (transmembrane) α -helices that can be identified unambiguously.
Detection of heavy-atom labels	Possible if the resolution is sufficiently high.
Detection of bound antibody (fragment)	Possible if the occupancy is high enough. May break crystallographic symmetries that possess twofold screw-axes.
Comparison with proteins of assumed identical/similar topology	Sometimes possible if the overall mass of the two proteins is similar and no large ectodomains exist or are disordered. If an X-ray structure is known for one of the proteins, relevant projections may be calculated from the atomic coordinates.
Detection of conformational changes	Possible, but very difficult to interpret. Preferably, such an analysis will use electron diffraction data. True vector-difference maps should be presented if the changes are large or globally affect the unit cell (Lindahl & Henderson, 1997). Methods and data should be presented in detail (as a supplement if necessary) to allow evaluation, especially if amplitudes are not from electron diffraction. If differences are interpreted, then biochemical data should be provided or already be published elsewhere to support the argument.

structural details agree in independent micrographs. There are at least three approaches by which phase statistics have been presented in the past.

(1) Using image data, an amplitude-weighted phase error is calculated for each reflection (*AVRGAMPHS* program; Crowther *et al.*, 1996) and converted to a figure of merit (FOM) based on the combined phase probability distribution (Henderson *et al.*, 1986). In plane groups with no or only few phase constraints, this FOM is the only measure to represent the reliability of the data.

(2) For plane groups with a twofold axis along *z*, all phases adopt values of 0 or 180° in projection. In this case, phase quality can be expressed in terms of a 'centric phase error', *i.e.* the deviation of the averaged experimental phases from their theoretical values. This representation has become very popular. Usually, the 'centric error' is tabulated for groups of reflections after binning into several resolution shells.

(3) A third method for the presentation of the phase data combines the first and second approach, *i.e.* the error obtained for the averaged phase angle of a reflection is weighted by an additional term accounting for the deviation of the average phase from its theoretical target value.

Using these approaches interchangeably, a potential point of confusion is that their criteria for 'randomness' are different. Random data will give a 90° error for the first and third approach, while a mean error of 45° indicates random data in a pure comparison against 0/180°. It is therefore important to state clearly what is being reported.

While each of the three approaches outlined above is valid, approaches (1) and (2) may not reveal all aspects of the data structure for a given specimen. For instance, the FOM obtained by data averaging (approach 1) depends mostly on the signal-to-noise ratio of individual measurements and the number of independent observations. This should also apply to 'centric phase' errors (approach 2), assuming that the structure is truly centrosymmetric and that no systematic errors are present. Fig. 1 illustrates that in practice these assumptions do not necessarily apply. Image data from crystals of two very different specimens, the bacterial Na⁺/H⁺ antiporter NhaA (Williams *et al.*, 1999) and a truncated gap-junction channel (Unger, Kumar *et al.*, 1997) were sequentially combined. As expected, the overall mean phase error obtained

by approach (1) steadily decreases with an increase in the number of independent observations. In contrast, increasing the size of the data set effects only a small improvement for the overall mean error of the centric comparison. Such a behavior of the centric phase errors can only be explained if the projection is not truly centrosymmetric or if systematic errors are present, or both. Consequently, neither approach (1) nor approach (2) provides a satisfactory description of the data in these cases. A possible solution is the presentation of a 'combined phase error' (approach 3). In praxis this involves an adjustment of the FOM (based on approach 1) by the centric phase error (approach 2) according to $FOM_{\text{adjusted}} = (FOM \cdot \cos \alpha_{\text{av}})$. The program *PLOTALL* can then be used to plot the 'combined phase error' for each reflection (Fig. 2). Note that the program *SCALIMAMP* (Schertler *et al.*, 1993) or its derivative *FOMSTATS* (Yeager *et al.*, 1999) automatically generate the adjusted FOM values (FOM_{adjusted}) whenever centrosymmetric phase constraints are applied. Use of these FOM_{adjusted} values in the Fourier inversion results in a 'double weighting' of the Fourier components. The rationale for this approach is that it puts the highest weight on the most reliable data, *i.e.* well defined reflections (FOM close to 1) that are in good agreement with the centrosymmetric constraint ($\cos \alpha_{\text{av}}$ close to 1). At the same time, this approach efficiently reduces the contribution of reflections that are ill-defined (FOM is small) and/or are affected by some sort of systematic error ($\cos \alpha_{\text{av}}$ is small). Consequently, showing the 'combined phase error' for

each reflection as illustrated in Fig. 2(b) provides a comprehensive description of the quality of the phase data and their conformity with applicable constraints.

2.2.2. Symmetry. In most cases, crystallographic symmetry is enforced in the final calculation of the projection density map. Whenever this occurs, the validity of the symmetry averaging should be demonstrated. For unstained specimens, this can be performed most easily by analyzing phase relationships in the data from untilted crystals (*ALLSPACE* program; Valpuesta *et al.*, 1994). Nevertheless, any symmetry so determined should be enforced only if the phase relationships are obeyed to the highest resolution chosen for density-map calculations. Otherwise, the map should be presented without symmetry averaging, *i.e.* in plane group $p1$. This strategy seems reasonable mostly for two reasons: inaccuracies in the correction of phase data for the contrast-transfer function and specimen tilt or both can result in a loss of the required phase relationships. Especially in the case of thick specimens ($>100 \text{ \AA}$), involuntary specimen tilt may invalidate the enforcement of symmetry. For instance, crystal tilts of as little as $2\text{--}3^\circ$ were sufficient to completely abolish $p6$ symmetry in the images of untilted truncated gap-junction channels (Unger, Kumar *et al.*, 1997). In this case, a specimen thickness of $\sim 150 \text{ \AA}$ resulted in partial loss of symmetry even

at 0.5° crystal tilt. Such circumstances require the tilt geometry to be accounted for in order to generate true projection data. For thinner specimens the effect is less pronounced, yet care should be taken whenever *ALLSPACE* fails to detect symmetry. Besides of a true lack of symmetry, the latter may simply be obscured.

2.2.3. Use of inverse B factors. In the past, crystalline areas of the majority of specimens were too small (*i.e.* less than $\sim 0.5 \times 0.5 \mu\text{m}$) to allow the collection of electron diffraction data. Nevertheless, density maps could be calculated because the images provided both the phase and amplitude information. However, in contrast to electron-diffraction data, image-derived amplitudes are less accurate owing to the influence of the contrast-transfer function and other factors that degrade image contrast, such as specimen charging or movement. Consequently, image-derived amplitudes show a pronounced fall-off at intermediate and high resolutions (see Henderson, 1992), causing a diminution of fine structural detail in the calculated density maps. In principle, this effect can be overcome by applying an inverse B factor ('temperature factor') which will restore power to the higher resolution terms. However, this procedure will amplify noise with the same fidelity. Therefore, image sharpening should only be used if the phases of the higher resolution reflections are well deter-

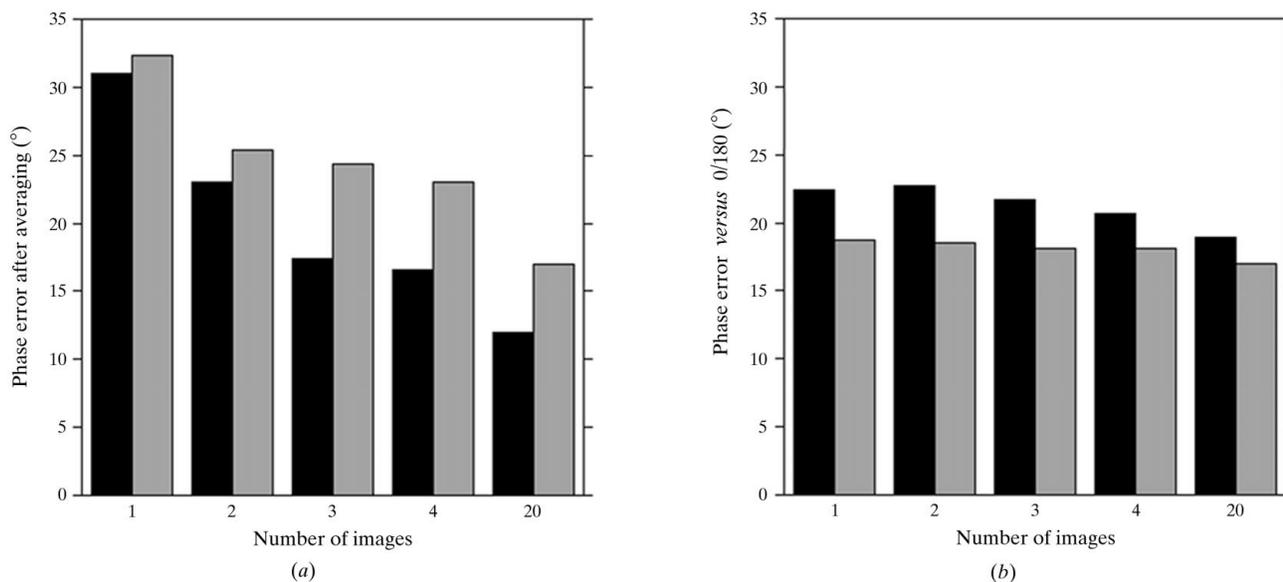


Figure 1

Characterization of phase errors for projection data. Phase errors are the best tool for the evaluation of data quality in electron crystallographic work. Shown in this figure is how different types of overall mean phase errors of projection data behave as a function of the number of images used. The errors were obtained by sequential averaging of data from 1, 2, 3, 4 and 20 images of a gap-junction channel (black columns; Unger, Kumar *et al.*, 1997) and the bacterial Na⁺/H⁺-antiporter NhaA (grey columns; Williams *et al.*, 1999). Data from the first three images were of truly untilted crystals. In all other cases, only those measurements within 0.0016 \AA^{-1} (gap junction) and 0.005 \AA^{-1} (NhaA) along the direction of z^* were included in the averaging. Errors shown in (a) correspond to the overall figure of merit (FOM) calculated by *AVRGAMPHS*. An error of 90° would be obtained for random data. The errors shown in (b) represent the mean centric phase error. This error reflects the average deviation of the average phases from their theoretical target values ($0/180^\circ$) imposed by the $p6$ (gap junctions) and $p22_12_1$ (NhaA) symmetries of the two specimens. In this case, random data would result in a 45° deviation overall. The total number of independent measurements to 6 \AA resolution was 179, 353, 516, 596 and 1266 in case of the gap junction, describing 76 unique reflections. In the case of NhaA, 368, 713, 936, 1152 and 1704 individual measurements were averaged to yield phases for 204 unique reflections. The smaller redundancy in the data and the lower symmetry in the case of NhaA explain the slightly higher phase errors shown in (a) compared with those obtained for the gap-junction specimen. However, the overall centric phase errors are very similar in both cases and do not improve much as data are added. This unexpected behavior is related to the 3D molecular transforms, which in both cases cause systematic errors in the determination of the projection data. The conclusion to be drawn here is that the use of a large and highly redundant data set will provide highly significant averages for the phase values. However, significance of the averages and their conformity with applicable phase constraints are not necessarily related to each other.

mined and weighted according to their reliability. If this applies, then the use of an inverse B factor can be very useful. This is illustrated in Figs. 3(a) and 3(b) showing a side-by-side comparison of projection density maps obtained for a truncated gap-junction channel. The significant increase in structural detail upon sharpening frequently raises concerns about potential artifacts. Accompanying this is the question of which B factors are permissible and how one can determine what B factor will be best for any given data set. Based on what has been used in the past, B factors of -300 to -600 \AA^2 seem required/sufficient to restore the appropriate weight to image-derived amplitudes in the $5\text{--}10 \text{ \AA}$ region. To put these numbers in perspective, a third panel has been added in Fig. 3(c). This part of the figure shows the same projection based on data where all amplitudes had been set to unity and the cosine of the 'combined phase error' (see Fig. 2b) was used as a weight in the Fourier summation. This treatment abolished the 110-fold difference in amplitude that separated the strongest reflection in the data set from the average amplitude for reflections between 6 and 7.5 \AA . In comparison, a maximal boost of ~ 11 -fold (B factor of -350 \AA^2) was applied for the calculation of the map shown in Fig. 3(b). Undoubtedly, the use of unity amplitudes (Fig. 3c) strongly emphasized the main features of the projected channel structure and introduced additional strong density features close to the twofold and

threefold axes. Note, however, that the overall appearances of the maps in Figs. 3(b) and 3(c) were still remarkably similar despite the large differences in amplitudes used for their calculation. Interestingly, not all of the additional density features in Fig. 3(c) were artifacts. The 3D reconstruction (Unger *et al.*, 1999) reveals that the additional densities around the threefold axes correspond to part of one of the transmembrane α -helices. Without doubt, using unity amplitudes instead of the real amplitudes is an extreme. Nevertheless, the fact that most of the information in the density maps remains meaningful regardless of how the amplitudes are treated emphasizes that the choice of B factor is arbitrary and need not be finely tuned. Furthermore, the danger of introducing gross distortions by using B factors is small as long as the phases are well determined and properly weighted. However, caution is warranted when scaling data in the resolution range between 5 and 7 \AA because the molecular transform may be intrinsically weak in this region. Two examples where this is observed are bacteriorhodopsin and aquaporins. Mistaking the characteristic feature of the molecular transform for an imaging artifact and attempting an inappropriate boost of these amplitudes would not be beneficial. Even more caution seems appropriate as near-atomic resolution is approached because the degradation of image data may no longer follow a simple exponential. Luckily, in

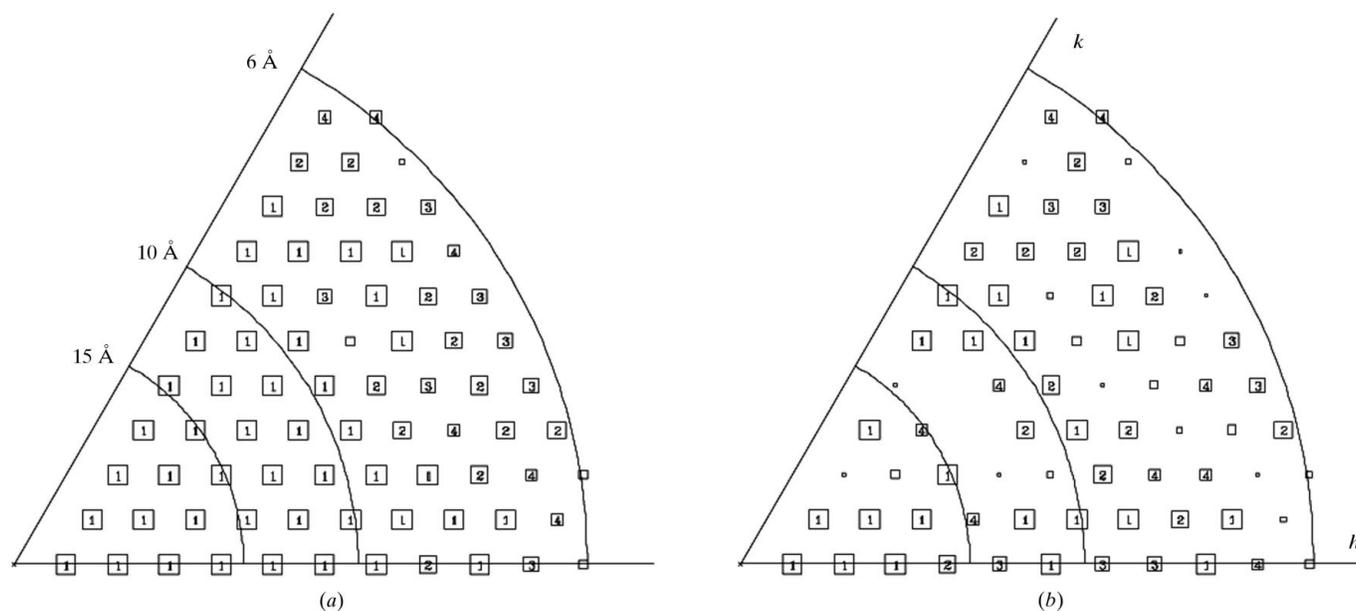


Figure 2

Combined phase errors as a measure of projection data quality. As shown in Fig. 1, neither the phase error obtained from averaging of the data nor the deviation of the average phase angles from applicable theoretical target values provide a straightforward measure of data quality in plane groups with twofold symmetry along the z axis. In these cases, the presentation of 'combined phase errors' may be the best solution to provide a complete account of data quality. Data from the gap-junction specimen were used to compile Fig. 2. Shown in (a) is the FOM for each reflection obtained by data averaging using the program *AVRGAMPHS*. The corresponding 'combined phase error' for each reflection is presented in (b). The 'combined phase error' of a reflection is defined as 'combined error' ($^{\circ}$) = $\cos^{-1}(\text{FOM} \cdot \cos \alpha_{\text{av}})$, where FOM is the figure of merit (ranging from 0 to 1) and α_{av} is the average phase angle after data averaging. Both FOM and α_{av} are calculated by *AVRGAMPHS*. According to its definition, the 'combined phase error' reflects both the statistical significance of the averaged phase angle of a reflection and its deviation from the applicable target phase value. Furthermore, the term (FOM $\cdot \cos \alpha_{\text{av}}$) normally serves as a weight for the structure factors if centrosymmetric phase constraints are enforced in the Fourier summation. This procedure minimizes the contribution of unreliable reflections. For plotting purposes, increasing errors are represented by decreasing box sizes. Numbers inside the boxes are assigned to phase errors as follows, 1 < 8° , 2 < 14° , 3 < 20° , 4 < 30° . No numbers are written out for errors encoded by the small box sizes, 5 < 40° , 6 < 50° , 7 < 70° and 8 < 90° . In both representations, an error of 90° indicates that the phase of the reflection is not significant.

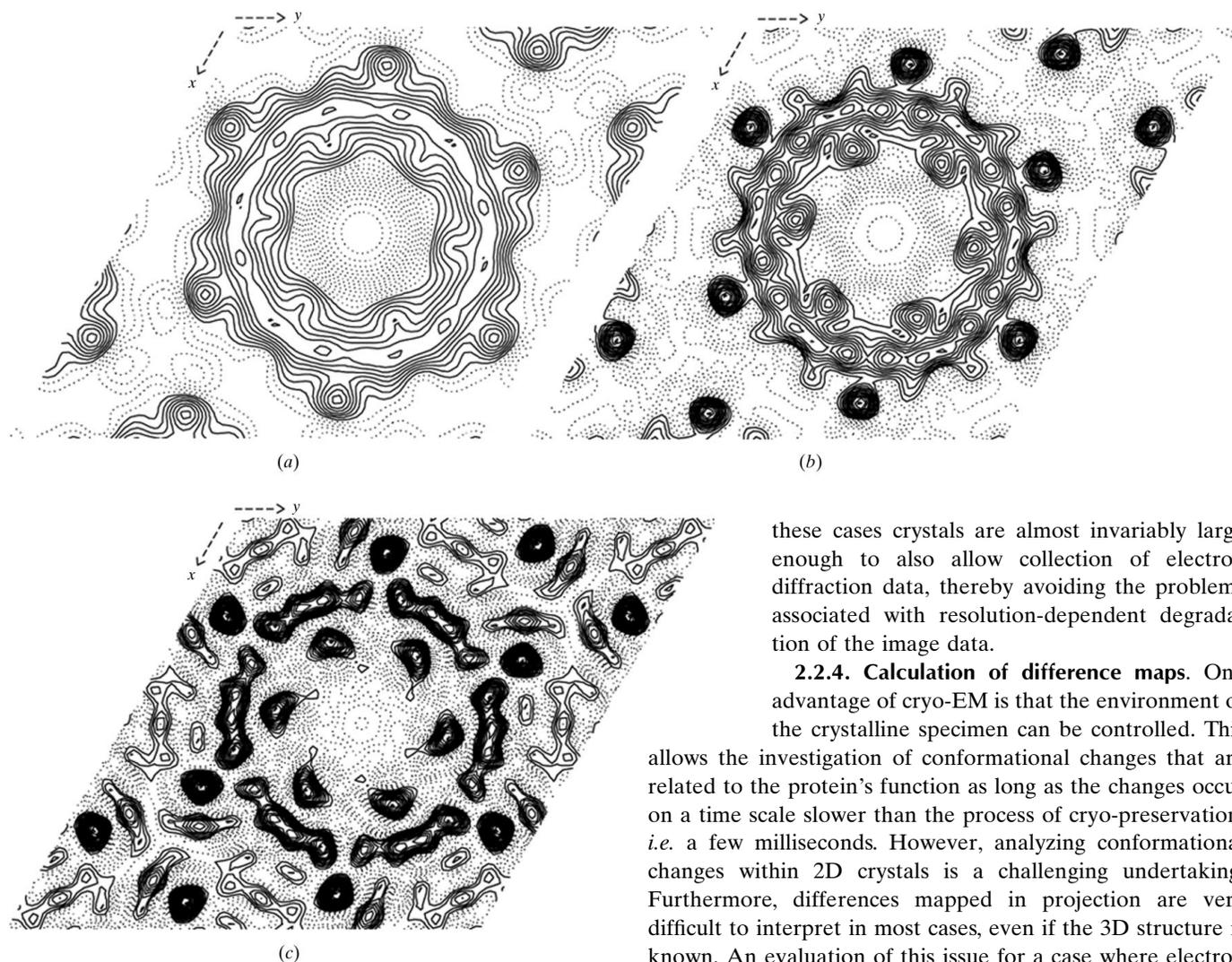


Figure 3

Effect of inverse B factors on projection density maps. Inverse B factors are frequently used in electron crystallography to overcome the resolution-dependent fall-off of image-derived amplitudes in cases where no electron diffraction data are available. Three projection density maps of a single truncated gap-junction channel are shown. Each map was calculated to a nominal resolution of 6 Å. Inverse B factors of $B = 0$ and -350 Å^2 were used in (a) and (b), respectively. The map shown in (c) was calculated after setting all amplitudes to unity and using $(\text{FOM} \cdot \cos \alpha_{av})$ (see Fig. 2b) as a weight in the Fourier inversion. All three maps are contoured in increments of 0.25σ . Solid lines are drawn for contour levels above the mean. Density maps were generated using the CCP4 program package (Collaborative Computational Project, Number 4, 1994). The 'sharpening-factor' (sf) is obtained as $\text{sf} = \exp[-B/(4d^2)]$, where B is the B factor and d is the resolution in Å. For instance, $B = -350 \text{ Å}^2$ increases the amplitude of a reflection at 6 Å by ~ 11 -fold, provided that its FOM is 1. For $\text{FOM} < 1$, the boost will be reduced to $(\text{FOM} \cdot \text{sf})$. Similarly, applying $(\text{FOM} \cdot \cos \alpha_{av})$ (see Fig. 2b) as a weight in plane groups with twofold constraint ensures that the full amount of sharpening is applied only to the most reliable reflections. In this case, even amplitudes as incorrect as those used for the map shown in (c) will result in a 'density distribution' that still describes the main features of the projected structure correctly. In case of the gap-junction channel, these features are two groups of six roughly circular and rotationally staggered densities that are separated by a continuous band of density.

these cases crystals are almost invariably large enough to also allow collection of electron diffraction data, thereby avoiding the problems associated with resolution-dependent degradation of the image data.

2.2.4. Calculation of difference maps. One advantage of cryo-EM is that the environment of the crystalline specimen can be controlled. This allows the investigation of conformational changes that are related to the protein's function as long as the changes occur on a time scale slower than the process of cryo-preservation, *i.e.* a few milliseconds. However, analyzing conformational changes within 2D crystals is a challenging undertaking. Furthermore, differences mapped in projection are very difficult to interpret in most cases, even if the 3D structure is known. An evaluation of this issue for a case where electron diffraction data were available can be found in Lindahl & Henderson (1997). Their paper draws the conclusion that a consideration of changes in either the amplitudes or phases alone is sufficient for the determination of localized changes in the structure. However, if the structural changes are spread over the entire unit cell then only a true vector difference will unambiguously describe the differences between the two structures. In this case, a full set of amplitudes and phases is required for each of the states that are compared. Things are complicated further if the projection is centrosymmetric, because in this case phase changes cannot be subtle but always take the form of a swing through 180° . Unless the structural changes are very large, this may result in a situation where phase changes are observed for weak reflections only. Since those reflections are also noisier than strong reflections, it becomes very difficult to determine whether the observed changes are meaningful. If in addition image-derived amplitude data are used for the calculation of difference maps, then it seems increasingly doubtful that any observed differences in the density distribution are significant. How such projection difference maps can be presented will need further discussion. The calculation of independent difference maps based on 'amplitudes only', 'phases only' and a true vector difference

may provide a possible way out of this dilemma in cases where all data are obtained from the images. If the full vector difference combines contributions that are clearly visible in both the 'amplitude-only' and 'phase-only' difference maps, then even image data alone may be sufficient to describe the differences correctly. However, if the full vector difference represents only changes in either the amplitudes or the phases, then the result is probably not significant.

2.3. 3D data

2.3.1. 3D density maps at intermediate resolutions – why are they useful? The atomic models of bacteriorhodopsin, light-harvesting complex II and tubulin demonstrate that high-resolution density maps can be generated by electron crystallography. Nevertheless, large 2D crystals that are ordered to near atomic resolution are difficult to obtain. In contrast to X-ray crystallography, however, even small and moderately ordered 2D arrays can provide useful data in the 5–10 Å region. This is possible for two reasons. Firstly, translational lattice disorder can be corrected computationally and, secondly, image-derived phases and amplitudes are sufficient to obtain interpretable density maps at intermediate resolutions. Especially in the case of membrane proteins, such maps continue to be useful because the structures are generally less complex than those of soluble proteins. For instance, the intermediate resolution map of a typical membrane protein will identify the following.

(i) The type of secondary-structure elements: α -helices will be resolved whereas β -structure will not.

(ii) The packing arrangement of α -helices: this has been shown in a number of cases [aquaporin-1 (Cheng *et al.*, 1997; Li *et al.*, 1997; Walz *et al.*, 1997), frog rhodopsin (Unger, Hargrave *et al.*, 1997), the sarcoplasmic Ca^{2+} and plasma membrane H^+ -ATPases (Zhang *et al.*, 1998; Auer *et al.*, 1998), gap-junction intercellular channels (Unger *et al.*, 1999) and the bacterial Na^+/H^+ -antiporter (Williams, 2000)].

Furthermore, intermediate resolution density maps may be useful in the following ways.

(i) To detect conformational changes that are associated with protein function. For instance, at 9 Å resolution, Unwin (1995) visualized changes in the structure of the nicotinic acetylcholine receptor that occur during opening and closing of the channel part of the receptor. Similarly, large differences in the orientation of the ectodomains in the structures of the sarcoplasmic Ca^{2+} and plasma membrane H^+ -ATPases were proposed to represent the closed and open states of P-type ATPases (Stokes *et al.*, 1999).

(ii) To help in phasing data from 3D crystals.

(iii) To improve crystallinity of specimens: knowing the thickness and surface relief of the molecule may help to identify detergents and additives that are suitable to improve crystal quality.

Finally, intermediate resolution maps also provide experimental constraints for the following.

(i) Hybrid crystallography: resolving details within large globular domains is difficult to achieve by cryo-EM because

these regions often show more flexibility and the orientation of secondary-structure elements is not constrained. However, the cryo-EM reconstruction will provide a molecular envelope suitable for the docking of high-resolution X-ray structures into the cryo-EM density map of the 2D crystal or macromolecular assembly. An example for this approach is the high-resolution model of microtubules (Nogales *et al.*, 1999).

(ii) Model-building studies: for instance, Baldwin used a 7 Å density map of frog rhodopsin and constraints obtained from an extensive analysis of the amino-acid sequences for the construction of the first realistic C^α template for G-protein coupled receptors (GPCR). The Baldwin model (Baldwin *et al.*, 1997) was the first to account for the significant differences in helix packing that separate GPCR from bacteriorhodopsin, which until then had widely been used as a template for GPCR modeling.

None of the above replaces the rich detail of a high-resolution structure. However, intermediate resolution models do represent a milestone in the analysis of a new specimen and hence will continue to provide a useful template until high-resolution data become available.

2.3.2. Specimen tilts. As a continuation of the previous section, the question arises as to what extent a structure needs to be determined in order to be useful for any of above applications. The answer is that this will largely depend on the actual structure and the questions being asked. For instance, assume that the transmembrane domains of an integral membrane protein are all α -helical. Furthermore, let the highest inclination angle of any of these helices with respect to the membrane plane be 30°. In this case, an in-plane resolution of better than 10 Å and specimen tilts in the range of 25–30° will suffice if the main intention is to demonstrate α -helical character of the transmembrane domains and to visualize their packing. This is illustrated in Fig. 4. Identical cross-sections parallel to the membrane plane are shown from one of the two membrane-spanning parts of a gap-junction channel. While only 5° of specimen tilt (Fig. 4*a*) did not resolve either of the two more highly tilted helices, the general 'story' emerged at tilt angles as little as 15°. Finally, at 30° of specimen tilt all transmembrane α -helices were clearly resolved. However, the exact length of the α -helices remained unknown owing to the elongation of the entire structure caused by the large 'missing cone' of data. The corresponding low vertical resolution makes model building difficult because such an exercise largely depends on accurate density constraints. To facilitate efficient modeling would require resolutions of ~4.5–5 Å (xy plane) and 6–7 Å (along z). Approaching these target values, the α -helical repeat becomes visible allowing unambiguous placement of canonical α -helices into the density map. Yet, achieving vertical resolutions significantly better than 10 Å requires good data from specimens tilted to more than 40°, which is not easily accomplished. Such high tilts are also required for the visualization of ordered connecting loops or α -helices oriented parallel to the membrane plane. Similarly, large globular ectodomains of membrane proteins as well as 2D crystals of soluble proteins are more demanding. In these cases, both α - and β -secondary structure may be present that

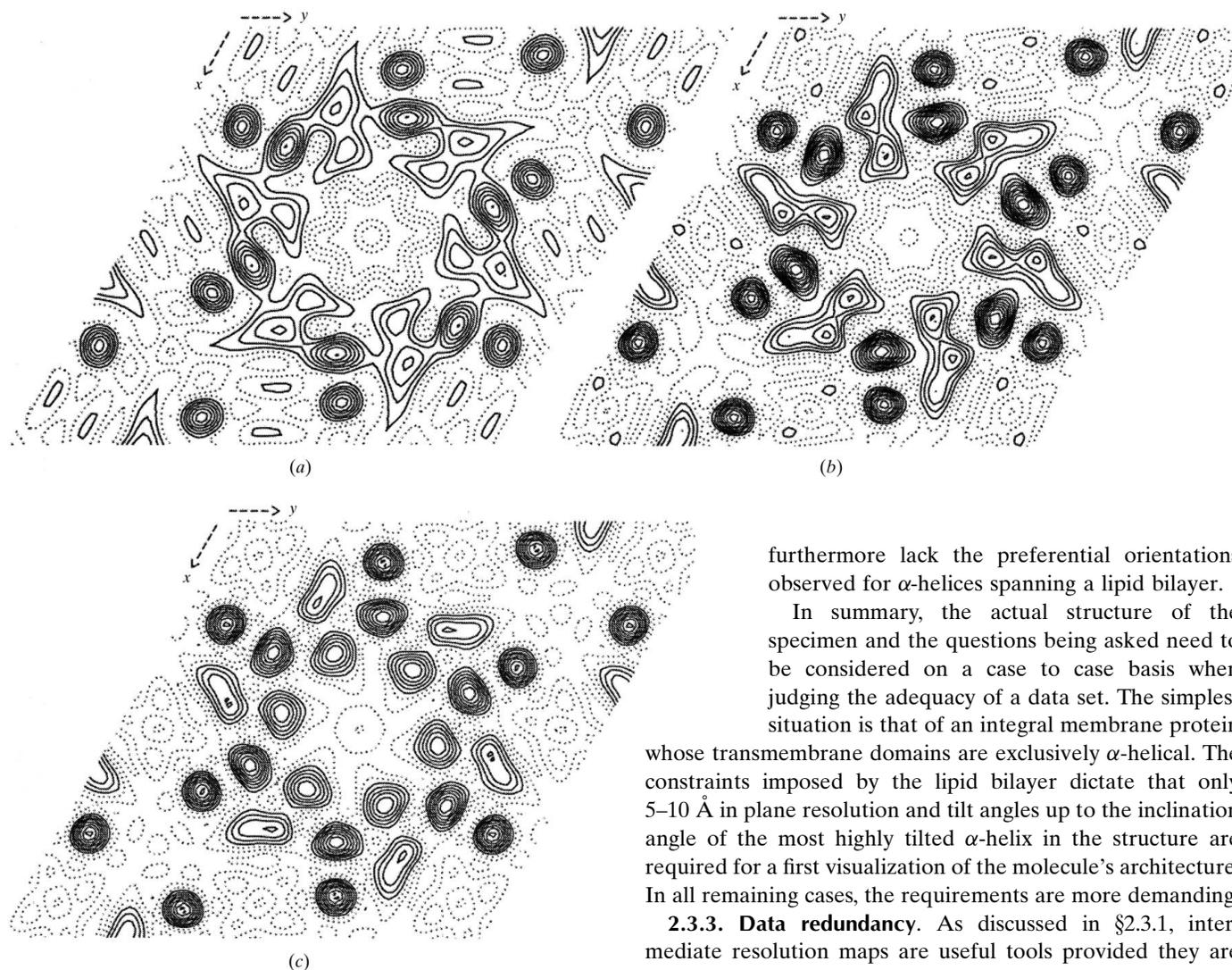


Figure 4

Impact of specimen tilt on definition of α -helices. Gap-junction channels directly connect two adjacent cells. Each of the channel's 12 subunits adopts an α -helical fold within its membrane-spanning part (Unger *et al.*, 1999). The α -helices can be resolved as soon as the data sample views of the molecule at tilt angles equal to the inclination angle of the most highly tilted α -helix in the structure. To illustrate this fact, three 3D density maps were calculated from data corresponding to maximum specimen tilts of (a) 5°, (b) 15° and (c) 30°, respectively. An inverse B factor of -350 \AA^2 was applied to correct the resolution-dependent fall-off of the image-derived amplitudes. Shown are identical cross-sections from each of these maps. The sections are contoured in steps of σ with solid lines representing contour levels above the mean. Even a specimen tilt of only 5° suffices to resolve two of the α -helices. However, the more highly tilted α -helices remain ill-defined. Inclusion of data from crystal tilts up to 15° completely resolves the two α -helices that are oriented almost perpendicular to the membrane plane. Furthermore, the two more highly tilted α -helices start being recognisable. At 30° of specimen tilt, the vertical resolution is sufficient to resolve all of the α -helices. Nevertheless, the cross-section through the most highly tilted α -helix is still elliptical in shape, indicating that its tilt is very close to the maximum specimen tilt covered by the data. Addition of data from even more highly tilted crystals would improve the definition of each of the α -helices. However, the main conclusion drawn from this section of the 3D density-map will remain unchanged, *i.e.* each of the six connexin subunits has four α -helical transmembrane domains.

furthermore lack the preferential orientations observed for α -helices spanning a lipid bilayer.

In summary, the actual structure of the specimen and the questions being asked need to be considered on a case to case basis when judging the adequacy of a data set. The simplest situation is that of an integral membrane protein

whose transmembrane domains are exclusively α -helical. The constraints imposed by the lipid bilayer dictate that only 5–10 Å in plane resolution and tilt angles up to the inclination angle of the most highly tilted α -helix in the structure are required for a first visualization of the molecule's architecture. In all remaining cases, the requirements are more demanding.

2.3.3. Data redundancy. As discussed in §2.3.1, intermediate resolution maps are useful tools provided they are well defined. Achieving this goal requires data redundancy to support the iterative refinement of image parameters such as astigmatism and tilt geometry. In turn, these refinements improve the quality of the fitted lattice lines giving more reliable phase estimates for the structure factors. In practice, a fivefold to tenfold redundancy provides a solid base for data refinement and structure calculation. However, exactly how many images are needed will depend on their quality, the thickness of the specimen and the crystallographic symmetry. Generally, fewer images are required for specimens that exhibit high degrees of symmetry. Still, even in these cases a sufficiently large number of independent measurements can only be obtained by inclusion of data from several images per degree of specimen tilt. This applies particularly for tilt angles $>30^\circ$, where the volume of reciprocal space increases rapidly while data quality from individual images tends to go exactly the opposite way. An annoying habit in the presentation of data is the listing of a nominal maximum tilt used during data collection without an evaluation of how well the corresponding part of reciprocal space has been sampled. Certainly, it looks appealing and possibly more acceptable to reviewers if images were recorded at $>45^\circ$ of tilt. Yet, what good does this do if in reality the resolution in those images was either not

isotropic (*i.e.* less good in the direction perpendicular to the tilt axis) or the actual number of measurements was insufficient to support a reliable fit of the lattice lines? Therefore, authors should provide information about the ‘effective specimen tilt’ for which reciprocal space has been sampled efficiently. As with other issues in electron crystallography, the definition for ‘efficiently’ remains subject to further debate. Until settled, authors should provide an evaluation of the completeness of their fitted data. For instance, a comparison of the number of fitted structure factors having phase errors of $<60^\circ$ with the number of theoretically possible structure factors to a given resolution and set of maximum specimen tilt angles would indicate where the sampling of reciprocal space begins to break down. Furthermore, a final evaluation of the effective resolution limits of the reconstruction is necessary to establish what conclusions can be drawn without over-interpreting the map.

2.3.4. Assessment of resolution and treatment of constrained lattice lines. As for projection density maps, average phase errors obtained from the lattice-line fit provide the best measure for data quality. In my experience, the appearance of an intermediate resolution density map settles once the overall weighted phase error calculated by *LATLINE* (Agard, 1983) approaches 20° . The likely causes for higher overall phase errors (especially in excess of 30°) are either sparse data distribution or insufficiently refined image parameters or both. In these cases, a listing of the phase errors in resolution bins is desirable. A major disadvantage of intermediate resolution density maps lies in the lack of a clear criterion defining their effective resolution. This applies in particular to the vertical resolution, which owing to the ‘missing cone’ of data will be lower than the in-plane resolution. For instance, at an in-plane resolution of 6 \AA , resolution along the z axis can vary anywhere between ~ 7 and 25 \AA depending on the ‘nominal’ and ‘effective’ maximum tilts of the specimen (see §2.3.3). How to assess vertical resolution in these cases is still subject to discussion. A possible solution consists of the analysis of the point-spread function of the experimental data. In the past, this function has been calculated by setting all phases to 0° , all amplitudes to unity and by using the experimental FOMs (calculated by *LATLINE*) in the Fourier inversion. If the data are isotropic and perfect this procedure results in a sphere centered about the origin. However, missing and/or imperfect data cause a distortion of the sphere (Henderson *et al.*, 1990; Unger & Schertler, 1995). The distortion can be analyzed along the principal planes of the reconstruction, giving a quantifiable measure of the effective resolution. The advantage of this approach is that one can easily simulate ‘perfect’ data for any resolution/maximum tilt cutoffs. This simulation provides an estimate for the best possible resolution along all directions and thereby serves as a useful comparison to the result obtained for the experimental data. An example for such a comparison is shown in Fig. 5. A disadvantage of the traditional method of calculating the point spread (*i.e.* setting all amplitudes to unity) lies in the disregard of experimental amplitude data. Setting all amplitudes to unity *de facto* mimics the use of a very

significant B factor (see also Fig. 3c). Under these circumstances, poor or missing FOM values will have almost no impact on the point-spread function, which as a consequence becomes largely a measure for the presence/absence of structure factors. This can result in an overestimation of the vertical resolution, especially for two-sided plane groups that impose phase constraints on certain lattice lines. The latter cannot be emphasized enough because the lattice-line fitting program *LATLINE* assigns $\text{FOM} = 100$ (= perfect measurement) to structure factors with applicable phase constraints,

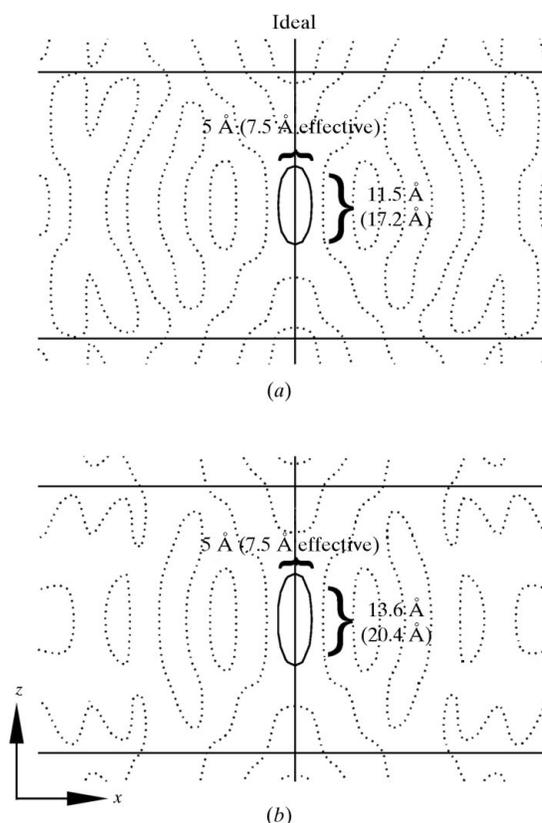


Figure 5

Assessment of resolution by means of the point-spread function. The effective resolutions of a reconstruction can be evaluated by comparison of the point-spread functions calculated for a simulated perfect data set and the experimental data (Unger & Schertler, 1995). Using 7.5 \AA in plane resolution and a maximum specimen tilt of 30° , the zx sections for (a) the simulated perfect data set and (b) the experimental data are shown. A grid with 40 \AA spacing in x and z was superimposed on the sections to provide an internal scale. The elongation of the spheres along z , which is characteristic for the presence of a ‘missing cone’ of data, can clearly be seen. The spheres were scaled to the same maximum density and sections were contoured with a single positive contour set at half height of the peak to satisfy the Rayleigh criterion of resolution. At this height, the peak measures 5 \AA across in the direction of x . A conversion factor of 1.5 correlates the nominal resolution of 7.5 \AA with the width at half height (*i.e.* nominal resolution/width at half height = $7.5 \text{ \AA}/5 \text{ \AA}$). This conversion factor can be used for the determination of the effective resolution from the half-width of the point-spread functions. Accordingly, 17.2 \AA is the best possible vertical resolution (*i.e.* along z) at 7.5 \AA in-plane resolution and a 30° maximum tilt for the specimen (= $1.5 \times 11.5 \text{ \AA}$). Applied to the experimental data (*i.e.* set all amplitudes to unity, set all phases to 0° and use FOMs obtained by *LATLINE* as weight in Fourier inversion), an effective vertical resolution of 20.4 \AA is obtained.

regardless of the quality and distribution of the actual data. Hence, *LATLINE* will generate apparently perfect data points along constrained lattice lines even if the measurements are too sparse and/or noisy to validate the fit. Inclusion of these points not only causes an overly optimistic estimate of the vertical resolution but may also introduce distortions in the 3D density map. Two steps are necessary to avoid these problems. First, the data of all lattice lines (constrained and unconstrained) need to be manually truncated to exclude regions where measurements are distributed too sparsely for a meaningful curve fit. Second, measurements for constrained lattice lines need to be isolated and then fitted in plane group $p1$. If phase constraints are enforced, then the FOMs of any constrained structure factor should be adjusted to reflect the combined phase error, as was discussed for projection data. Note, however, that phases are constrained to $\pm 90^\circ$ along certain lattice lines in some two-sided plane groups. In these cases, the FOM obtained from the fitting should be multiplied with the sine of the phase angle instead of the cosine because phase angles of 0 or 180° are 'random' in this case. Independently of these suggestions, the question remains whether the use of unity amplitudes is commendable for the calculation of the point-spread function. For future use of this approach, it might be preferable to retain the experimental amplitudes and FOMs, set all phases to 0° and to apply the same inverse B factor for calculation of the point spread as is used for the calculation of the final 3D density map. Applied to the example shown in Fig. 5, retaining the experimental amplitudes and using a B factor of -350 \AA^2 (see Fig. 4) results in an estimate of $\sim 27 \text{ \AA}$ for the vertical resolution. This compares 'unfavorably' to the $\sim 20 \text{ \AA}$ obtained when using unity amplitudes (Fig. 5*b*). Stemming from the same set of original data, this discrepancy in estimates obtained for the vertical resolution emphasizes why this issue requires further discussion. Until firm guidelines have been established, it seems appropriate to clearly state how the resolution limits of an intermediate resolution 3D map have been estimated.

2.4. Use of negatively stained specimen

Negative staining is a very valuable approach for screening purposes and initial crystal evaluation. However, the information obtainable from negatively stained specimens is limited to about 12–15 \AA resolution and no internal molecular detail can be visualized. Therefore, priority should be given to data collection from unstained specimens wherever possible. However, extremely low quantities of sample as well as highly disordered, small and/or low yields of 2D crystals all may justify an in-depth analysis of a negatively stained specimen. In these cases, one has to be aware of the potential artifacts that are commonly associated with work in negative stain. For instance, *ALLSPACE* results obtained for negatively stained crystals should always be treated with caution. In contrast to unstained specimen, data from negatively stained samples are dominated by the exclusion pattern of the stain molecules. Uneven staining of the two sides of the crystal and a small number of phase comparisons (owing to the inherently low

resolution) may suggest symmetries that are higher or lower than the true symmetry of the molecules on the lattice. Similarly, biological molecules are prone to shrinkage and deformation if dried down in negative stain. Consequently, one would expect these factors to be discussed along with the results of any reconstruction. Another aspect that is often overlooked is the potential influence of the stain's pH on the specimen. For instance, the acidic pH of uranyl acetate may cause protein denaturation or induce disordering of the lattice. Hence, comparison with results obtained in other negative stains such as phosphotungstic acid (titrated to pH 7) would be useful for an assessment of the significance of the reconstruction. Lastly, there have been incidences where negatively stained specimens were reported to diffract to better than 10 \AA . This is possible, especially if the crystals consist of multiple layers that are stacked in register. However, the higher resolution data require corrections for the reversal of contrast that occurs past the resolution limit of the stain. If no such correction is implemented then inclusion of any data past the resolution limit of the stain (*i.e.* 12–15 \AA) will render the final density map largely uninterpretable.

2.5. Submittal of data to the PDB

Guidelines for the submission of electron-crystallographic data are currently being developed (H. M. Berman, personal communication). In cases of structures determined to intermediate resolutions, deposition of the structure factors and basic crystallographic information [unit cell, symmetry, resolution (nominal), suitable B factor for map calculation (where applicable)] might be preferable over deposition of an actual density map. In either case, the following information could also be included to allow users an independent assessment of the reliability of the data.

- (i) Software package used.
- (ii) Were images recorded with flood-beam or spot-scan?
- (iii) Was dynamic focusing used to record images of tilted crystals?
- (iv) Are the amplitudes of the structure factor based on electron diffraction, or were they derived from the images?
- (v) What was the highest 'effective specimen tilt angle' covered by the data?
- (vi) What was the average redundancy in the data set before fitting of the lattice lines?
- (vii) What was the weighted overall phase error obtained from the fitting of the lattice lines?
- (viii) What is the effective vertical resolution of the reconstruction? How was this vertical resolution determined?
- (ix) Were the images reprocessed using an initial 3D model as reference? If so, how did the resolution improve?

I am grateful to Drs K. A. Williams and M. J. Yeager for permission to use NhaA and α_1 -Cx263T data sets for generation of the figures. I also thank C. Breyton, M. Lindahl and J. Vonck at the Max-Planck-Institute for Biophysics,

Frankfurt for comments on the manuscript and discussions on the issues covered by this contribution.

References

- Agard, D. A. (1983). *J. Mol. Biol.* **167**(4), 849–852.
- Asturias, F. J. & Kornberg, R. D. (1999). *J. Biol. Chem.* **274**(11), 6813–6816.
- Auer, M., Scarborough, G. A. & Kühlbrandt, W. (1998). *Nature (London)*, **392**, 840–843.
- Baldwin, J. M., Schertler, G. F. X. & Unger, V. M. (1997). *J. Mol. Biol.* **272**, 144–164.
- Cheng, A., van Hoek, A. N., Yeager, M., Verkman, A. S. & Mitra, A. K. (1997). *Nature (London)*, **387**, 627–630.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Crowther, R. A., Henderson, R. & Smith, J. M. (1996). *J. Struct. Biol.* **116**, 9–16.
- Henderson, R. (1992). *Ultramicroscopy*, **46**(1–4), 1–18.
- Henderson, R., Baldwin, J. M., Céska, T. A., Zemlin, F., Beckman, E. & Downing, K. H. (1990). *J. Mol. Biol.* **213**(4), 899–929.
- Henderson, R., Baldwin, J. M., Downing, K. H., Lepault, J. & Zemlin, F. (1986). *Ultramicroscopy*, **19**, 147–178.
- Jap, B. K. & Li, H. (1995). *J. Mol. Biol.* **251**, 413–420.
- Karrasch, S., Bullough, P. A. & Gosh, R. (1995). *EMBO J.* **14**, 631–638.
- Koning, R. I., Keegstra, W., Oostergetel, G. T., Schuurman-Wolters, G., Robillard, G. T. & Brisson, A. (1999). *J. Mol. Biol.* **287**, 845–851.
- Kubalek, E. W., Le Grice, S. F. & Brown, P. O. (1994). *J. Struct. Biol.* **113**(2), 117–123.
- Kühlbrandt, W., Wang, D. N. & Fujiyoshi, Y. (1994). *Nature (London)*, **367**, 614–621.
- Li, H., Lee, S. & Jap, B. K. (1997). *Nature Struct. Biol.* **4**, 263–265.
- Lindahl, M. & Henderson, R. (1997). *Ultramicroscopy*, **70**, 95–106.
- Mitra, A. K., van Hoek, A. N., Wiener, M. C., Verkman, A. S. & Yeager, M. (1995). *Nature Struct. Biol.* **2**, 726–729.
- Nogales, E., Whittaker, M., Milligan, R. A. & Downing, K. H. (1999). *Cell*, **96**(1), 79–88.
- Nogales, E., Wolf, S. G. & Downing, K. H. (1998). *Nature (London)*, **391**, 199–203.
- Rhee, K. H., Morris, E. P., Barber, J. & Kühlbrandt, W. (1998). *Nature (London)*, **396**, 283–286.
- Rhee, K. H., Morris, E. P., Zheleva, D., Hankamer, B., Kühlbrandt, W. & Barber, J. (1997). *Nature (London)*, **389**, 522–526.
- Schertler, G. F., Villa, C. & Henderson, R. (1993). *Nature (London)*, **362**, 770–772.
- Schmidt-Krey, I., Murata, K., Hirai, T., Cheng, Y., Morgenstern, R., Fujiyoshi, Y. & Hebert, H. (1999). *J. Mol. Biol.* **288**, 243–253.
- Stokes, D. L., Auer, M., Zhang, P. & Kühlbrandt, W. (1999). *Curr. Biol.* **9**, 672–679.
- Unger, V. M., Hargrave, P. A., Baldwin, J. M. & Schertler, G. F. X. (1997). *Nature (London)*, **389**, 203–206.
- Unger, V. M., Kumar, N. M., Gilula, N. B. & Yeager, M. J. (1997). *Nature Struct. Biol.* **4**(1), 39–43.
- Unger, V. M., Kumar, N. M., Gilula, N. B. & Yeager, M. (1999). *Science*, **283**, 1176–1180.
- Unger, V. M. & Schertler, G. F. (1995). *Biophys J.* **68**(5), 1776–1786.
- Unwin, N. (1995). *Nature (London)*, **373**, 37–43.
- Uzgiris, E. E. & Kornberg, R. D. (1983). *Nature (London)*, **301**, 125–129.
- Valpuesta, J. M., Carrascosa, J. L. & Henderson, R. (1994). *J. Mol. Biol.* **240**, 281–287.
- Walz, T., Hirai, T., Murata, K., Heymann, J. B., Mitsuoka, K., Fujiyoshi, Y., Smith, B. L., Agre, P. & Engel, A. (1997). *Nature (London)*, **387**, 624–627.
- Walz, T., Typke, D., Smith, B. L., Agre, P. & Engel, A. (1995). *Nature Struct. Biol.* **2**(9), 730–732.
- Williams, K. A. (2000). *Nature (London)*, **403**, 112–115.
- Williams, K. A., Geldermacher-Kaufer, U., Padan, E., Schuldiner, S. & Kühlbrandt, W. (1999). *EMBO J.* **18**, 3558–3563.
- Wilson-Kubalek, E. M., Brown, R. E., Celia, H. & Milligan, R. A. (1998). *Proc. Natl Acad. Sci. USA*, **95**(14), 8040–8045.
- Yeager, M., Unger, V. M. & Mitra, A. (1999). *Methods Enzymol.* **294**, 135–180.
- Zhang, P., Toyoshima, C., Yonekura, K., Green, N. M. & Stokes, D. L. (1998). *Nature (London)*, **392**, 835–839.