

Thomas F. Lerch,^a Qing Xie,^a
Heather M. Ongley,^b Joan Hare^b
and Michael S. Chapman^{a*}

^aDepartment of Biochemistry and Molecular Biology, School of Medicine, Oregon Health and Science University, Portland, OR 97239-3098, USA, and ^bInstitute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306-4381, USA

Correspondence e-mail: chapmami@ohsu.edu

Received 4 December 2008

Accepted 5 January 2009

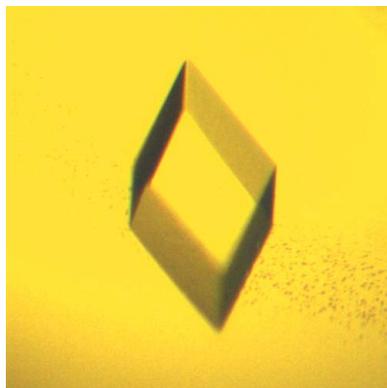
Twinned crystals of adeno-associated virus serotype 3b prove suitable for structural studies

Adeno-associated viruses (AAVs) are leading candidate vectors for gene-therapy applications. The AAV-3b capsid is closely related to the well characterized AAV-2 capsid (87% identity), but sequence and presumably structural differences lead to distinct cell-entry and immune-recognition properties. In an effort to understand these differences and to perhaps harness them, diffraction-quality crystals of purified infectious AAV-3b particles have been grown and several partial diffraction data sets have been recorded. The crystals displayed varying levels of merohedral twinning that in earlier times would have rendered them unsuitable for structure determination, but here is shown to be a tractable complication.

1. Introduction

Adeno-associated viruses (AAVs) are 4.7 kb single-stranded DNA viruses that show potential as gene-therapy vectors (Berns & Giraud, 1996). AAVs require helper viruses (usually adenovirus) for replication, but their infection is not detrimental to cellular viability beyond the effects of the helper virus (Carter, 2006*a*). 12 AAV serotypes have been identified to date that differ in cellular tropism and immunogenic susceptibility (Mori *et al.*, 2004; Schmidt *et al.*, 2008). The most widely characterized serotype is AAV-2 and most gene-therapy vectors used in clinical trials are based upon this serotype (Carter, 2006*b*). There has been much recent interest in exploiting the differences between AAV serotypes (see below) for gene therapy. Serotype 3 is 87% identical to serotype 2 in terms of capsid sequence (Muramatsu *et al.*, 1996). Several closely related isolates of serotype 3 have been characterized virologically and have about six coding differences in the capsid (Rutledge *et al.*, 1998), but no structures are known. AAV-3b is the subject of the current study.

AAV capsids contain 60 subunits of viral protein VP1, VP2 or VP3 which are expressed from overlapping DNA sequences such that they differ only at the N-terminus. VP1, VP2 and VP3 exist in an approximately 1:1:10 ratio within the $T = 1$ icosahedrally symmetric virus capsid. There is no known organization with respect to which of the three types of capsid protein occupies each quasi-symmetrical site in the capsid. VP3 is the dominant capsid protein and in the case of AAV-3b contains 534 amino acids. VP2 has an additional 65 residues at the N-terminus and VP1 a further 137 residues. The molecular determinants of cellular entry reside on the viral capsid surface. Small differences in serotype sequences give rise to diverse mechanisms of cell entry. Despite high sequence conservation among all AAV serotypes (>50% identity), at least three distinct primary cellular receptors are used (Kaludov *et al.*, 2001; Summerford & Samulski, 1998; Schmidt *et al.*, 2008). AAV-1 and AAV-2 share ~83% identity, yet bind different primary receptors (Rabinowitz *et al.*, 2002). AAV-2 and AAV-3b, which share 87% identity, have both been shown to utilize heparan sulfate as the primary receptor (Rabinowitz *et al.*, 2002). However, their binding mechanisms are thought to be distinct as AAV-3b lacks the amino acids most responsible for the binding of AAV-2 to heparin (Kern *et al.*, 2003; Wu *et al.*, 2006; O'Donnell *et al.*,



© 2009 International Union of Crystallography
All rights reserved

2009). It is hoped that structural comparisons of AAV-2 and AAV-3b will facilitate the characterization of key binding interactions and that this information will facilitate the development of gene-therapy vectors targeted to specific cell types.

Immune recognition of AAV limits transduction efficiency, as repeated exposure to one serotype leads to diminishing therapeutic returns. Indeed, 50–80% of the population have antibodies against AAV-2 (Blacklow *et al.*, 1968; Parks *et al.*, 1970; Erles *et al.*, 1999), the template of most vectors currently used in gene-therapy trials. In previous studies, anti-AAV-2 serum was unable to fully neutralize infection by AAV-3b, suggesting that the immunogenic determinants of these two similar capsids are distinct and that AAV-3b could be a useful alternative gene-therapy vector (Rutledge *et al.*, 1998).

Much insight into AAV biology has been gained since the determination of the human AAV-2 structure (Xie *et al.*, 2002; Rabinowitz & Samulski, 2002), which provided a template for engineering viruses with enhanced cellular tropism. Intense interest in the use of AAV as a gene-therapy vector and in the advantages of different serotypes has led to widening structural studies. These have included human and nonhuman AAVs chosen to represent the full diversity of the family (DiMattia *et al.*, 2005; Kaludov *et al.*, 2003; Miller *et al.*, 2006; Nam *et al.*, 2007; Quesada *et al.*, 2007; Xie *et al.*, 2008). In contrast, the current study is of a serotype closely related to AAV-2, which was chosen to help pinpoint which of the differences between serotypes are critical to the functional differences between the human viruses.

2. Results

2.1. Production and purification of AAV-3b

The large-scale production of pure AAV has been an obstacle in crystallographic studies and two strategies to overcome this have emerged. The initial structure of AAV-2 was determined using infectious virions (Xie *et al.*, 2003). AAV production is complicated by the dependence of its replication upon co-infection with a helper virus, which kills the host cells. Subsequent to AAV-2, most structural studies have utilized non-infectious adeno-associated virus-like particles (VLPs) produced using baculovirus expression systems. With AAV-3b, we return to infectious virions, guided by developments for high-yield AAV-2 production (Xie *et al.*, 2004). Functional

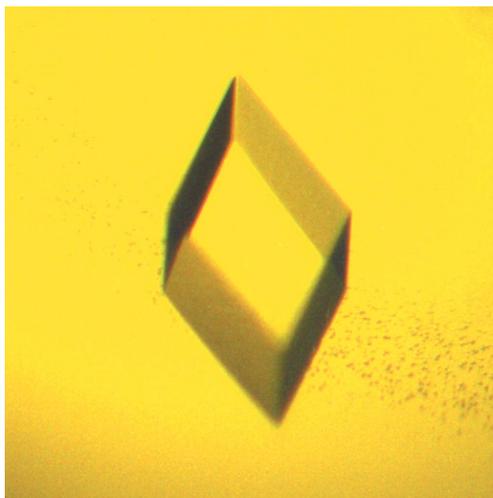


Figure 1
AAV-3b crystals were grown in 3.4% PEG 6000 in HM buffer by vapor diffusion. Corner to corner, the maximum dimension of this crystal is approximately 0.4 mm.

interpretation of the structure will be facilitated by the knowledge that the sample is replication-competent.

To avoid the accumulation of random mutations, each preparation was started from an infectious plasmid clone, pAAV-3b (Rutledge *et al.*, 1998). 10 μg pAAV-3b was transfected into six T225 flasks of HeLa cells grown to $\sim 80\%$ confluency. Adenovirus type 2 (2×10^7 pfu ml^{-1} final concentration) was added 90 min after transfection and the cells were harvested 2–3 d later. A crude cell lysate was prepared by three rounds of freeze-thawing. This cell lysate contained not only AAV-3b but also sufficient adenovirus to provide helper functions and could be used directly as an inoculum for AAV-3b amplification.

Inoculum was added to 1–21 suspension-adapted HeLa cells (0.6×10^6 cells ml^{-1}) in a spinner flask and incubated for 48–72 h at 310 K. The cell suspension was pelleted at 112 400 rcf (25 000 rev min^{-1}) for 3 h using a Beckman SW 28 ultracentrifuge rotor and the cells were subjected to treatment with trypsin (0.125% final concentration) and sodium deoxycholate (0.5% final concentration) for 45 min at 310 K. The lysate was homogenized and AAV-3b was then purified in three consecutive CsCl gradients by the methods previously used for AAV-2 (Xie *et al.*, 2004). The virus was dialyzed against HM buffer (100 mM HEPES, 50 mM MgCl_2 , 0.03% sodium azide pH 7.3) for crystallization trials. Typical AAV-3b yields were $\sim 1\text{--}2$ mg l^{-1} .

The purity of AAV-3b preparations was monitored in several ways. With several serotypes being studied concurrently, restriction digests were used to confirm the identity of the plasmid, taking advantage of a restriction site that was unique to the pAAV-3b plasmid. A multiplex PCR assay (Mitchell *et al.*, 2006) was used to confirm that the amplified virus was of the correct serotype. The homogeneity of DNA-filled capsids (*versus* empty capsids) was confirmed by the absorbance ratio (260 nm/280 nm) of 1.4. LDS-PAGE analysis was performed to determine the purity of viral capsid proteins. Samples

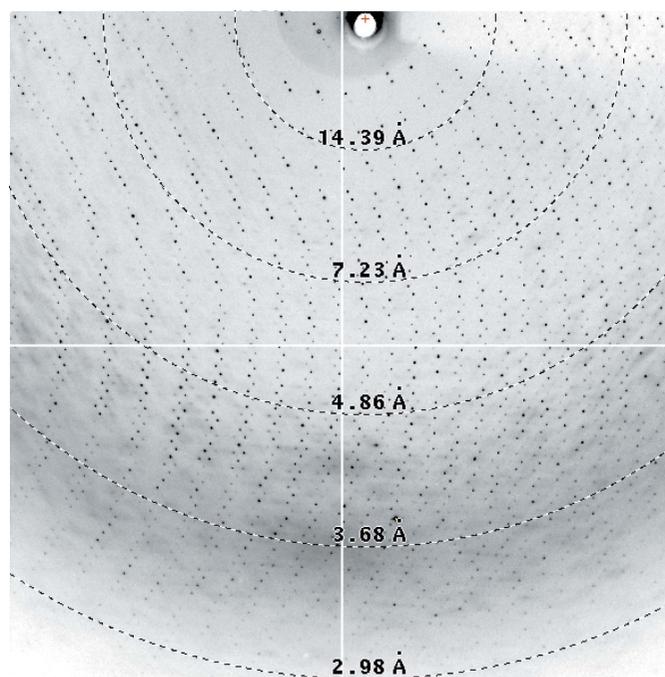


Figure 2
Diffraction image from AAV-3b crystals which diffract X-rays to 3.0 Å resolution. A dual-detector system was used to maximize coverage of the diffraction pattern, while maintaining adequate spot separation, and the diffraction pattern from one of the detectors is shown.

Table 1
 Data-processing statistics.

Values in parentheses are for the highest resolution shell.

	Data set 1	Data set 2
Space group	<i>R</i> 3	<i>R</i> 3
Unit-cell parameters (Å, °)		
Hexagonal	$a = b = 257.8, c = 607.0$	$a = b = 257.7, c = 603.8$
Rhombohedral	$a = 251.2, \alpha = 61.73$	$a = 250.7, \alpha = 61.96$
No. of observations	78243	177517
Unique reflections	76529 (4081)	147388 (3727)
Resolution (Å)	100–3.0 (3.07–3.00)	100–2.6 (2.66–2.60)
$R_{\text{merge}}^{\dagger}$ (%)	6.7 (9.4 \ddagger)	6.6 (42.2 \ddagger)
$\langle I \rangle / \langle \sigma(I) \rangle$	5.55 (1.92)	6.14 (1.11)
Completeness (%)	12.7 (10)	32.1 (12)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the i th observation of a symmetry equivalent of reflection hkl . \ddagger The highest resolution shell has no symmetry-equivalent reflections and therefore R_{merge} was calculated in medium-resolution shells (4.7–4.3 Å for data set 1 and 2.9–2.8 Å for data set 2).

were heated to 343 K in 2% NuPage LDS (Invitrogen) for 10 min and run on 4–12% polyacrylamide gradient gels at various serial dilutions. After staining, VP1, VP2 and VP3 capsid proteins seen in the expected 1:1:10 ratios were used to calibrate neighboring lanes and to demonstrate, using overloaded lanes (~10 µg per well), that the AAV-3b sample purity was >97%.

2.2. Crystallization

AAV-3b crystals (Fig. 1) were grown by the hanging-drop vapor-diffusion method at room temperature. Purified virus (2 µl) at 6.4 mg ml⁻¹ was mixed with 2 µl reservoir solution and equilibrated against 800 µl reservoir solution containing 3.2–3.4% PEG 6000 in HM buffer (100 mM HEPES, 50 mM MgCl₂, 0.03% sodium azide pH 7.3). Unlike previously studied AAV-2 crystals (Xie *et al.*, 2003), crystals grown in the presence of 25% glycerol were of inferior quality and did not produce usable diffraction data.

2.3. Data collection and processing

Diffraction data were collected on the biohazard-level 2 F1 beamline at the Cornell High Energy Synchrotron Source (CHESS) with $\lambda = 0.9186$ Å. Prior to data collection, crystals were soaked for 1 h in a solution containing 30% glycerol and 6% PEG 6000 in HM buffer and then flash-frozen. Diffraction images (Fig. 2) were collected using a side-by-side dual Quantum 4 CCD detector (ADSC, Poway, California, USA) with a 0.3° oscillation angle (Xie *et al.*, 2008). The crystals were exposed for 180 s per image with a crystal-to-detector distance of 580.9 mm. Reflection intensities were indexed, integrated and scaled using the *HKL* v.1.97.9 package (Otwinowski & Minor, 1997).

The diffraction of eight different AAV-3b crystals was evaluated. Many did not diffract beyond 7 Å resolution. Partial data sets were collected from two crystals that diffracted X-rays to 3.0 Å (45 frames collected) and 2.6 Å (80 frames collected) resolution. These data sets were both processed as rhombohedral using mosaicities that were optimized to 0.43° and 0.50°, respectively. Initially, the higher resolution data set appeared to belong to space group *R*32, but this was subsequently corrected (see below). The 3.0 Å data could not be scaled in *R*32, but scaled well (for a virus crystal) in space group *R*3 to an R_{merge} of 6.7% with 12.7% completeness (Table 1).

2.4. Evidence of twinning

There were several indications that the higher resolution data came from a twinned crystal of lower space-group symmetry. The

Table 2
 Twin fraction (α) determined from diffraction data.

	Twin law	R observed	Twin-fraction (α) estimate		
			Britton test	H -test	ML
3.0 Å data set	$(h, -h - k, -l)$	0.431	0.055	0.060	0.048
2.6 Å data set	$(h, -h - k, -l)$	0.108	0.399	0.409	0.478

unit-cell parameters of the two crystals were nearly congruent, yet the crystals had different apparent point-group symmetries. The unit-cell parameters (257 Å) correspond to the expected diameter of a virus particle, so with one virus per unit cell the icosahedral rotation axes are constrained to lie along all crystallographic symmetry elements. A preliminary self-rotation function (see §2.5) for the untwinned 3.0 Å data set confirmed that the viral threefolds were aligned with crystallographic symmetry but that the twofolds were not aligned as expected for *R*32. For the 2.6 Å data set, there were additional strong twofold axes in the ordinary rotation function orthogonal to the rhombohedral axis, as would be expected for space group *R*32, but the other symmetry elements expected of icosahedral symmetry aligned with crystallographic 32 point-group symmetry were not apparent. The ordinary rotation function was initially difficult to interpret owing to its complexity and artifacts from the incompleteness of the data. However, inspection at lower contour levels revealed one set of peaks identical to the untwinned crystal and another related by the aforementioned strong twofolds normal to the rhombohedral axis. Packing constraints limited the contents to one virus per unit cell in the rhombohedral setting, indicating that the twofold additional symmetry seen in scaling and in the rotation function was the result of twinning.

It was with some hesitation that conventional twinning tests were conducted. Traditional tests based on statistical distributions of intensities (Yeates, 1997) are sensitive to the quality of the experimental diffraction measurements. With their large unit cells and consequently diffraction intensities that are 100-fold to 1000-fold weaker, virus data sets are usually of much poorer quality than those encountered in protein crystallography (Fry *et al.*, 1999). The results of several approaches are presented, all of which were calculated using the *PHENIX* program suite (Adams *et al.*, 2002). The L -value (Padilla & Yeates, 2003), as well as the moment of the observed intensity/amplitude distribution, indicated that the 2.6 Å resolution data were twinned. Several twin laws are possible for the *R*3 space group and each was assessed based on merging statistics. R -value analysis of the potential twin laws gives a low merging R value of 0.108 for the twin law $(h, -h - k, -l)$, suggesting that the 2.6 Å resolution data are merohedrally twinned. Estimates of the twin fraction (α) were determined using this twin law and ranged from 0.40 to 0.48 (Table 2). The H -test (Yeates, 1997) measures the distribution of intensities between twin-related reflections as a function of α . Britton plots (Fisher & Sweet, 1980) estimate the twin fraction by minimizing the number of negative intensities after detwinning. The maximum-likelihood (ML) algorithm (Zwart *et al.*, unpublished work) is similar to Britton analyses, but takes experimental errors into account. The consistency of the estimates of the twin fraction (0.40, 0.41 and 0.48) is as much as can be expected from a viral diffraction data set of lower precision and completeness than typical for protein samples. The analysis confirms that the crystals belong to space group *R*3 and that their pseudo-32 symmetry results from near-perfect merohedral twinning.

The same analysis was performed on the 3.0 Å data set, but failed to detect any strong indication of twinning. However, it is possible that the 3.0 Å data set is partially twinned as it is difficult to distin-

guish low twin fractions from nontwinned data. The extent of twinning differs for the crystals used, complicating the merging of data sets. Calculation of the rotation function and initial refinement proceeded using the nontwinned 3.0 Å resolution data alone.

2.5. Rotation function

The orientation of AAV-3b in the crystal was determined from the rotation function calculated using the program *GLRF* (Tong & Rossmann, 1997). The rotation function is compromised by a truncation error resulting from the 13% completeness of the untwinned

data set. Nevertheless, the ordinary self-rotation function at 12 Å resolution shows most of the fivefold, threefold and twofold non-crystallographic symmetry (Fig. 3*a*). The corresponding icosahedrally locked rotation function (Fig. 3*b*) yielded a single solution that was consistent with the relative orientations of the fivefold, threefold and twofold symmetry peaks expected of icosahedral symmetry.

The corresponding 9.3 Å ordinary rotation function for the higher resolution data set was initially more confusing. In retrospect, the dominant twofolds orthogonal to the threefold axis come from the twinning operators. The complexity (numerous peaks) can now be understood in terms of two orientations of the icosahedron superimposed by the twinning operator (Fig. 3*c*).

It is instructive to consider the translation-function problem in two parts. Firstly, there is the positioning of the icosahedral point-group symmetry within the unit cell. Packing considerations (discussed above) require that the NCS point-group center lies on the crystallographic threefold. The position along the threefold can be set arbitrarily. Secondly, there is the placement of the phasing-model subunit within the icosahedral framework. Minor variations have been documented between related viruses (Xie *et al.*, 2003), but the packing of subunits within the icosahedron limits freedom and the adjustments needed can usually be made using rigid-body refinement without calculation of the full translation function.

2.6. Preliminary model building and refinement

The overall strategy would be to use the incomplete untwinned data set for an initial round of phase improvement, map calculation and model building. The more complete twinned data would then be used during atomic refinement.

Rigid-body refinement was first used (with the untwinned crystal) to improve upon the orientation of the icosahedral symmetry coming from the locked rotation function. The icosahedral threefold is constrained to be coincident with the crystallographic threefold, but the atomic model and other icosahedral symmetry operators could rotate about the threefold axis. Refinement using the program *CNS* (Brünger *et al.*, 1998) started with a 20-subunit rigid group from the

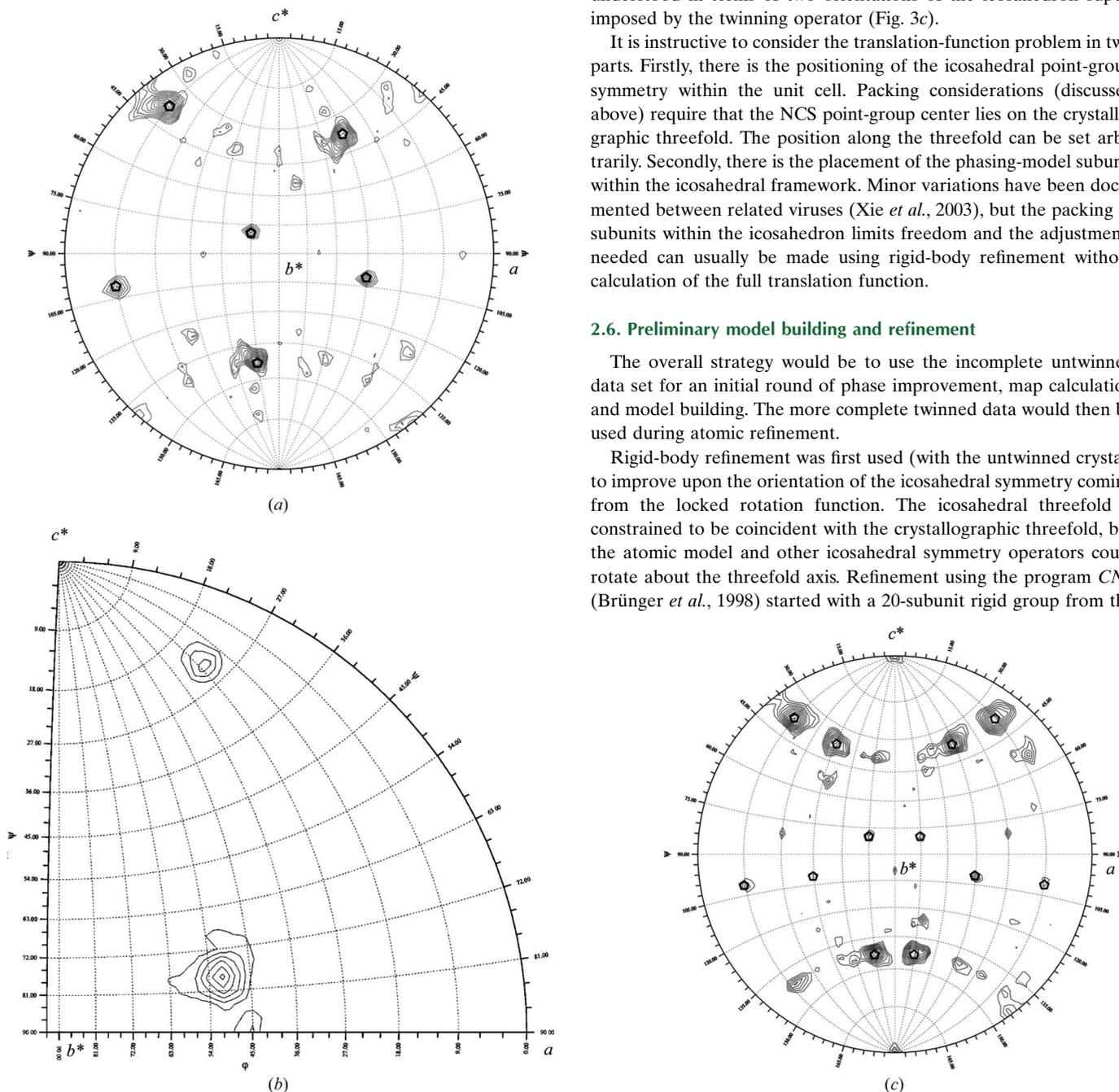


Figure 3

Rotation functions of AAV-3b. Ordinary self-rotation functions for (a) 3.0 Å and (c) 2.6 Å resolution data sets, calculated at 12 and 9.3 Å, respectively. The panels show spherical polar projections oriented with the hexagonal setting of the rhombohedral lattice (a , b^* , c^*) as indicated. (a) The ordinary self-rotation function (RF) at $\kappa = 72^\circ$, calculated with the 13% complete untwinned data set is imperfect, but shows the locations of the six expected icosahedral fivefolds (shown as pentagons). (b) In spite of the truncation errors evident in the ordinary RF, when icosahedral symmetry is imposed as a constraint in a 14 Å locked RF, an unambiguous solution for the icosahedral orientation emerges. (c) The additional symmetry in the 2.6 Å resolution data is a result of twinning. Fivefold peaks are observed from both twin domains and are highlighted as black and gray hexagons.

AAV-2 structure (Xie *et al.*, 2002) corresponding to the contents of the asymmetric unit for AAV-3b. For the 20-subunit group, only the crystallographic symmetry was imposed. The icosahedral NCS was not constrained, so that it was free to rotate about the crystallographic threefold. However, changes to the 20-mer were not significant (less than $0.2 \text{ \AA}/0.2^\circ$), so the icosahedral symmetry operators were left unchanged. The next round of rigid-body refinement allowed the AAV-2 subunit to rotate and translate. A small improvement (about $0.5 \text{ \AA}/0.12^\circ$) brought R/R_{free} from 0.41/0.41 to 0.38/0.39.

Initial phases to 3.75 \AA resolution were obtained from the modified AAV-2 atomic model within the AAV-3b unit cell. They were improved and extended in resolution through NCS averaging using the *RAVE/CCP4* programs (Collaborative Computational Project, Number 4, 1994). Averaging used a mask that initially included all grid points within 4 \AA of any subunit atom and was modified to remove overlap using the program *MAMA* (Kleywegt & Jones, 1994). The mask was generous enough to allow for the expected differences between AAV-3b and AAV-2. Phases were refined at 3.75 \AA resolution and then extended to 3.0 \AA in steps of approximately one-half of the longest reciprocal-lattice unit. Missing reflections were filled based on estimates from the back-transformed averaged map. The resulting map was interpretable, clearly showing differences in the backbone at the site of a sequence insertion and in side chains where AAV-3b and AAV-2 differed. Additionally, a stretch of ten residues was omitted from the AAV-2-based phasing model and its density was recovered in the averaged AAV-3b map, indicating that the completeness of the AAV-3b diffraction data was sufficient to remove model bias in the starting phases.

Manual modifications to the model to incorporate AAV-3b sequence changes were performed using the program *O* (Jones *et al.*, 1991). After refinement of individual atoms by simulated annealing, as well as *B*-factor refinement, the model agreed with the 3.0 \AA resolution data with an R/R_{free} of 0.29/0.30.

Confident that the rotation function and initial phasing and refinement were on track, the AAV-3b model was further refined against the twinned 2.6 \AA resolution data using the program *phenix.refine* with 20-fold NCS restraints (Adams *et al.*, 2002). Without accounting for twinning, R/R_{free} was 0.40/0.43 for data to 2.6 \AA resolution. Use of the twin operator immediately reduced R/R_{free} to 0.30/0.32. Preliminary refinement decreased R/R_{free} to 0.18/0.24, validating the use of this higher resolution data set for further study. Cross-validation data were selected randomly for this preliminary refinement. It is understood that completion of the structure will require re-refinement with resolution-shell test reflections to avoid bias from twinning or noncrystallographic symmetry (Fabiola *et al.*, 2006). Full refinement and rebuilding of the AAV-3b model will be followed by a thorough assessment of the biological significance of capsid-structure difference between AAV-3b and other AAV serotypes.

3. Discussion

Twinning has raised some strategic questions for this project that are relevant to structure determinations of other large and symmetrical (viral) complexes. With improvements in the available methodology, determinations of structures from twinned protein crystals have become much more common in recent years. For virus crystals, the challenges, at first sight, appear to be exacerbated by the weak diffraction and radiation-sensitivity often encountered with very large unit cells (Fry *et al.*, 1999). It is rare that complete data sets can

be obtained from a single crystal. With a variable twinning fraction, merging data from different crystals would only be possible after detwinning. The errors in individual intensities typical of viral data sets, combined with twinning fractions that approach 0.5, do not portend well for accurate deconvolution and structure determination in the conventional manner.

The alternative of persevering in a search for untwinned crystals may also be unsavory. Especially if dependent upon expression in mammalian cells, material for new crystal screens may be in short supply. Isometric capsids are notorious for sometimes crystallizing in multiple polymorphic forms under nominally identical conditions (Wu *et al.*, 1993). Polymorphism and twinning are both manifestations of the near-equivalence in energy of different crystal-packing configurations. Isometric viruses with their high symmetry and nearly spherical shape appear to be particularly prone to this.

Although polymorphic virus crystals have been well documented, there is not an extensive literature reporting twinned virus crystals. This may be because the susceptible forms have been deemed to be unworkable and have gone unreported. The example developed here shows that the presence of high-order noncrystallographic symmetry (NCS) may provide sufficient advantages to offset the data weakness and incompleteness and may support structure determination from sets of variably twinned crystals. Firstly, the presence of NCS redundancy restores a workable ratio between data points and model parameters and has supported numerous virus-structure determinations with completeness in the 10–30% range (Badger *et al.*, 1988). Secondly, the presence of NCS of a known high-order point group allows rotation functions that are complicated by twinning symmetry to be interpreted (Tong & Rossmann, 1990). Thirdly, the wide convergence of phase refinement using high-order NCS has largely obviated the need for experimental phasing of virus structures (Tsao *et al.*, 1992), a hurdle that would be even more daunting with twinning.

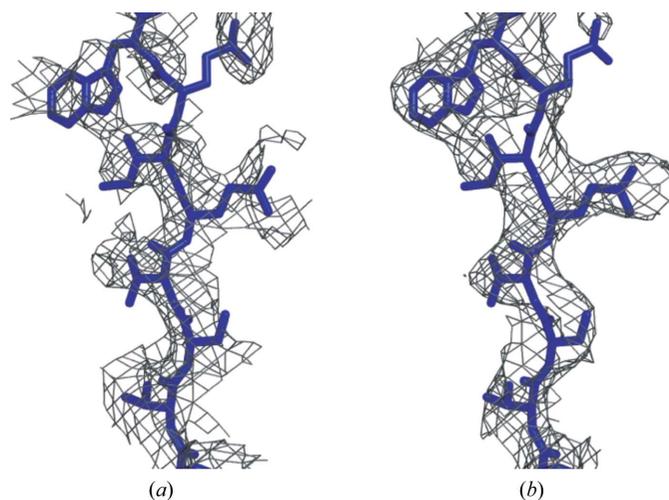


Figure 4 Initial averaged electron-density maps. Maps were created by phase extension from 3.75 to 3.0 \AA , averaging 20 subunits related by NCS. Starting phases were taken from (a) AAV-4 (55% sequence identity to AAV-3B for VP3) or (b) AAV-2 (87% identity). Overlaid is the preliminary model of AAV-3b. The region shown was omitted from the phasing model. Recovery of density for the omitted region shows that the power of 20-fold NCS averaging is sufficient to recover independent phase information, even with the 13% of the data that were from the untwinned crystal. With the AAV-2 phasing model, the density was of sufficient quality to remodel the regions that differed. Remodeling using density from the AAV-4 phasing model would have been more challenging, but might have been possible in several rounds of bootstrapping.

If one wishes to determine a virus capsid structure using twinned data, one can imagine two strategies. The 'more conventional' approach might be to first detwin the experimental data using the intensity statistics, apply initial molecular-replacement phases at low resolution and then refine and extend the phases to the resolution limit (Rossmann, 1995). Here, a more direct approach has been successful. An untwinned crystal has been used as a step in solving the rotation function for a twinned crystal as well as in symmetry averaging and remodeling. This model then sufficed as a starting point for refinement against more complete and higher resolution but twinned data from the other crystal.

How general is the approach? The initial steps of structure determination have been repeated using the structures of AAV-4 (Govindasamy *et al.*, 2006) and canine parvovirus (CPV; Tsao *et al.*, 1991) as phasing models. The sequence identities to the target AAV-3b are 55% and 23%, respectively, *i.e.* much lower than the 87% of the AAV-2 phasing model that was actually used. CPV shares the same core β -barrel fold, but its loops are quite different from those of AAV. AAV-4 and CPV atomic models were used to calculate an AAV-3b map at 3.75 Å resolution and to then extend the phases to 3.0 Å resolution by symmetry averaging. The quality of the phase-refined maps depended on the similarity of the starting phasing model, *i.e.* AAV-2 (87%) > AAV-4 (55%) > CPV (23%). Some features in the AAV-4 phased map can be discerned as unique, but there are also breaks in the density that would make model rebuilding challenging (Fig. 4). Recovery of independent phase information was tested by omitting ten-residue segments from the phasing model. For both the AAV-4 and CPV phasing starts, much but not all of the AAV-3b backbone density could be recovered. All of this is consistent with an understanding that the starting phases for the AAV-4 and CPV models (and their averaging masks) are closer to the limits of convergence and that the power of phase refinement depends on the data completeness and the order of symmetry (Arnold & Rossmann, 1988). The current structure determination was possible using untwinned data as a stepping stone that was only 13% complete, but this was facilitated by the availability of a closely homologous phasing model. A phasing model of lower homology gives an inferior-quality map. It would likely be improved with higher completeness in the untwinned data or higher than 20-fold noncrystallographic symmetry. With 13% complete data and a 55% homologous phasing model, the map shows some of the unique features of AAV-3B, but completion would likely require iterative remodeling as the initial phases for NCS averaging are improved.

For AAV-3b, it appears that a pair of partial data sets, a low-resolution set complementing a twinned higher resolution set, will support at least a preliminary structure determination. In spite of the technical challenges, it looks as if an AAV-3b structure will contribute to the growing understanding of cell entry and immune recognition in the different AAV serotypes (Opie *et al.*, 2003; Lochrie *et al.*, 2006). It is hoped that the insights gained will further accelerate progress in the design of improved AAV-based vectors and that this will contribute to the development of therapies for genetically based diseases (Wu *et al.*, 2000).

The authors would like to thank Thayumanasamy Somasundaram, Weishu Bu and the staff at CHESS who helped with data collection. CHESS is supported by the NSF and NIH/NIGMS *via* NSF award DMR-0225180 and the MacCHESS resource is supported by NIH/NCRR award RR-01646. This research was supported by the National Institutes of Health R01-GM66875 (MSC) and a predoctoral

fellowship from the American Heart Association, Florida and Puerto Rico affiliate 0515201B (HMO).

References

- Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* **D58**, 1948–1954.
- Arnold, E. & Rossmann, M. G. (1988). *Acta Cryst.* **A44**, 270–283.
- Badger, J. *et al.* (1988). *Proc. Natl Acad. Sci. USA*, **85**, 3304–3308.
- Berns, K. I. & Giraud, C. (1996). *Curr. Top. Microbiol. Immunol.* **218**, 1–23.
- Blacklow, N. R., Hoggan, M. D. & Rowe, W. P. (1968). *J. Natl Cancer Inst.* **40**, 319–327.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
- Carter, B. J. (2006a). *Parvoviruses*, edited by J. R. Kerr, S. F. Cotmore, M. E. Bloom, R. M. Linden & C. R. Parrish, pp. 497–498. London: Edward Arnold.
- Carter, B. J. (2006b). *Parvoviruses*, edited by J. R. Kerr, S. F. Cotmore, M. E. Bloom, R. M. Linden & C. R. Parrish, pp. 499–510. London: Edward Arnold.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- DiMattia, M., Govindasamy, L., Levy, H. C., Gurda-Whitaker, B., Kalina, A., Kohlbrenner, E., Chiorini, J. A., McKenna, R., Muzyczka, N., Zolotukhin, S. & Agbandje-McKenna, M. (2005). *Acta Cryst.* **F61**, 917–921.
- Erls, K., Sebkova, P. & Schlehofer, J. R. (1999). *J. Med. Virol.* **59**, 406–411.
- Fabiola, F., Korostelev, A. & Chapman, M. S. (2006). *Acta Cryst.* **D62**, 227–238.
- Fisher, R. G. & Sweet, R. M. (1980). *Acta Cryst.* **A36**, 755–760.
- Fry, E. E., Grimes, J. M. & Stuart, D. I. (1999). *Mol. Biotechnol.* **12**, 13–23.
- Govindasamy, L., Padron, E., McKenna, R., Muzyczka, N., Kaludov, N., Chiorini, J. A. & Agbandje-McKenna, M. (2006). *J. Virol.* **80**, 11556–11570.
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.
- Kaludov, N., Brown, K. E., Walters, R. W., Zabner, J. & Chiorini, J. A. (2001). *J. Virol.* **75**, 6884–6893.
- Kaludov, N., Padron, E., Govindasamy, L., McKenna, R., Chiorini, J. A. & Agbandje-McKenna, M. (2003). *Virology*, **306**, 1–6.
- Kern, A., Schmidt, K., Leder, C., Müller, O. J., Wobus, C. E., Bettinger, K., Von der Lieth, C. W., King, J. A. & Kleinschmidt, J. A. (2003). *J. Virol.* **77**, 11072–11081.
- Kleywegt, G. J. & Jones, T. A. (1994). *Proceedings of the CCP4 Study Weekend. From First Map to Final Model*, edited by S. Bailey, R. Hubbard & D. Waller, pp. 59–66. Warrington: Daresbury Laboratory.
- Lochrie, M. A., Tatsuno, G. P., Christie, B., McDonnell, J. W., Zhou, S., Surosky, R., Pierce, G. F. & Colosi, P. (2006). *J. Virol.* **80**, 821–834.
- Miller, E. B., Gurda-Whitaker, B., Govindasamy, L., McKenna, R., Zolotukhin, S., Muzyczka, N. & Agbandje-McKenna, M. (2006). *Acta Cryst.* **F62**, 1271–1274.
- Mitchell, D. A., O'Donnell, J., Hare, J. T. & Chapman, M. S. (2006). *J. Virol. Methods*, **136**, 277–282.
- Mori, S., Wang, L., Takeuchi, T. & Kanda, T. (2004). *Virology*, **330**, 375–383.
- Muramatsu, S., Mizukami, H., Young, N. S. & Brown, K. E. (1996). *Virology*, **221**, 208–217.
- Nam, H. J., Lane, M. D., Padron, E., Gurda, B., McKenna, R., Kohlbrenner, E., Aslanidi, G., Byrne, B., Muzyczka, N., Zolotukhin, S. & Agbandje-McKenna, M. (2007). *J. Virol.* **81**, 12260–12271.
- O'Donnell *et al.* (2009). In the press.
- Opie, S. R., Warrington, K. H. Jr, Agbandje-McKenna, M., Zolotukhin, S. & Muzyczka, N. (2003). *J. Virol.* **77**, 6995–7006.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* **D59**, 1124–1130.
- Parks, W., Boucher, D., Melnick, J., Taber, L. & Yow, M. (1970). *Infect. Immun.* **2**, 716–722.
- Quesada, O., Gurda, B., Govindasamy, L., McKenna, R., Kohlbrenner, E., Aslanidi, G., Zolotukhin, S., Muzyczka, N. & Agbandje-McKenna, M. (2007). *Acta Cryst.* **F63**, 1073–1076.
- Rabinowitz, J. E., Rolling, F., Li, C., Conrath, H., Xiao, W., Xiao, X. & Samulski, R. J. (2002). *J. Virol.* **76**, 791–801.
- Rabinowitz, J. E. & Samulski, R. J. (2002). *Mol. Ther.* **6**, 443–445.
- Rossmann, M. G. (1995). *Curr. Opin. Struct. Biol.* **5**, 650–655.
- Rutledge, E., Halbert, C. & Russell, D. (1998). *J. Virol.* **72**, 309–319.

- Schmidt, M., Voutetakis, A., Afione, S., Zheng, C., Mandikian, D. & Chiorini, J. A. (2008). *J. Virol.* **82**, 1399–1406.
- Summerford, C. & Samulski, R. J. (1998). *J. Virol.* **72**, 1438–1445.
- Tong, L. & Rossmann, M. G. (1990). *Acta Cryst.* **A46**, 783–792.
- Tong, L. & Rossmann, M. G. (1997). *Methods Enzymol.* **276**, 594–611.
- Tsao, J., Chapman, M. S., Agbandje, M., Keller, W., Smith, K., Wu, H., Luo, M., Smith, T. J., Rossmann, M. G., Compans, R. W. & Parrish, C. (1991). *Science*, **251**, 1456–1464.
- Tsao, J., Chapman, M. S. & Rossmann, M. G. (1992). *Acta Cryst.* **A48**, 293–301.
- Wu, H., Keller, W. & Rossmann, M. G. (1993). *Acta Cryst.* **D49**, 572–579.
- Wu, P., Xiao, W., Conlon, T., Hughes, J., Agbandje-McKenna, M., Ferkol, T., Flotte, T. & Muzyczka, N. (2000). *J. Virol.* **74**, 8635–8647.
- Wu, Z., Asokan, A., Grieger, J. C., Govindasamy, L., Agbandje-McKenna, M. & Samulski, R. J. (2006). *J. Virol.* **80**, 11393–11397.
- Xie, Q., Bu, W., Bhatia, S., Hare, J., Somasundaram, T., Azzi, A. & Chapman, M. S. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 10405–10410.
- Xie, Q., Hare, J., Turnigan, J. & Chapman, M. S. (2004). *J. Virol. Methods*, **122**, 17–27.
- Xie, Q., Ongley, H. M., Hare, J. & Chapman, M. S. (2008). *Acta Cryst.* **F64**, 1074–1078.
- Xie, Q., Somasundaram, T., Bhatia, S., Bu, W. & Chapman, M. S. (2003). *Acta Cryst.* **D59**, 959–970.
- Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.