Cryo-Electron Microscopy Structure of Yeast Ty Retrotransposon Virus-Like Particles

KEVIN J. PALMER, WILLEM TICHELAAR, NICOLA MYERS, NIGEL R. BURNS, SARAH J. BUTCHER, ALAN J. KINGSMAN, STEPHEN D. FULLER, and HELEN R. SAIBIL

Department of Crystallography, Birkbeck College, London, and British Biotech Pharmaceuticals Ltd. and Biochemistry Department, Oxford University, Oxford, United Kingdom, and EMBL, Heidelberg D-69012, Germany

Received 3 February 1997/Accepted 4 June 1997

The virus-like particles (VLPs) produced by the yeast retrotransposon Ty1 are functionally related to retroviral cores. These particles are unusual in that they have variable radii. A paired mass-radius analysis of VLPs by scanning transmission electron microscopy showed that many of these particles form an icosahedral T-number series. Three-dimensional reconstruction to 38-Å resolution from cryo-electron micrographs of T=3 and T=4 shells revealed that the single structural protein encoded by the TYA gene assembles into spiky shells from trimeric units.

Retroelements, which include the retroviruses and retrotransposons, are mobile genetic elements which replicate or transpose via an RNA intermediate and a reverse transcription reaction. Retrotransposons have been found in a wide range of eukaryotic systems. They have a genetic organization similar to that of the retroviruses (35), and they carry out part of their replication cycle in the context of a particle that is directly analogous to the retroviral core. The Ty1 element of yeast, and related retroelements, contains two genes that perform functions similar to those of retroviral gag and pol genes. The gag gene encodes the structural proteins of the particle. The pol gene encodes the enzymes needed for replication, namely protease, which cleaves the precursor proteins, reverse transcriptase, and integrase. However, Ty1-like retrotransposons lack a third gene analogous to the env gene of retroviruses. Env encodes the transmembrane and surface glycoproteins which target new virus particles to other host cells. Retrotransposons are therefore restricted to intracellular transposition and are not infectious.

The Ty1 retrotransposon of the yeast Saccharomyces cerevisiae is part of a family of Ty retroelements dispersed throughout the yeast chromosome. Transcription of the Ty element produces the 5.7-kb genomic RNA transcript, which acts as both message and template for reverse transcription (4). This genome is divided into genes TYA and TYB, which are the gag and pol analogs, respectively (10). As is the case with retroviral gag and pol genes, the TYA and TYB genes overlap and TYB is expressed via a ribosomal frameshift event within the overlap region to produce a 190-kDa TYA-TYB fusion protein called p3. The TYB-encoded component of this fusion protein is a polyprotein precursor for the Ty1 enzymes, which show limited but significant sequence homology to their retroviral counterparts (36). The primary translation product of the TYA gene is the 50-kDa structural protein p1, which is slightly basic, binds nucleic acids, and is phosphorylated in vivo (28). p1 self-assembles with or without the p3 enzymes to form immature, spherical virus-like particles (VLPs) in the cytoplasm of the yeast cell (1). Ty protease is necessary for maturation of Ty-VLPs via cleavage of p1 into the major 46-kDa structural protein p2 and of p3 into the functionally active protease. The cleavage that produces p2 is towards the C terminus of p1 (25, 29).

The similarities of the Ty elements to retroviruses have led them to be regarded as a basic model for all retroelements. They can be produced in large quantities by overexpression of the Ty transcription unit (28), which facilitates structural analysis. However, the particles are very variable in size, and this polydispersity severely complicates data collection. We have therefore used a somewhat simpler system in which Ty-VLPs are produced from the overexpression of a truncated TYA gene encoding only the first 381 amino acids of p1. We have shown previously that this truncated protein, designated p1-381, self-assembles into Ty-VLPs which, like the wild-type particles, package their cognate RNA. These capsids are similar to those of the mature or processed state. The use of this truncated form of capsid protein has reduced the polydispersity to some extent, probably because we have eliminated variation due to differences in extent of proteolytic processing of p1, permitting structural analysis (9).

Our previous characterization of p1-381 Ty-VLPs has shown that although these particles possess fivefold symmetry, they also exhibit polydispersity in size, suggesting that they might form a series of quasi-equivalent icosahedral structures (9). In addition, we showed that the average molecular size of the particles is about 14 MDa and that they are porous to macromolecules. In this study, we have established by scanning transmission electron microscopy (STEM) and cryo-electron microscopy (cryo-EM) that Ty-VLPs form T=3, T=4, and possibly larger icosahedral structures, with trimeric clustering of subunits.

MATERIALS AND METHODS

Sample preparation. The particles were purified by gel filtration and pooled into small-, medium-, and large-size fractions, covering broad size ranges, as previously described (8). The Ty-VLPs in the sample consisted solely of the p1-381 structural protein construct. The sample was treated with RNase to remove the packaged RNA, making the mass measurements proportional to the number of subunits, and increasing the contrast of the shell.

Electron microscopy. Cryo-EM images were recorded on a JEOL 2000 EX microscope with a lanthanum boride filament at 100 kV and an Oxford Instruments cryotransfer stage. The Ty-VLPs used for the image processing came from the medium-size fraction of the RNase-treated sample. The cryo-EM images of both the smaller- and larger-size fractions of the Ty samples were not used in the
analysis, as the preparations were of worse quality than that of the medium-size fraction.

Mass and radius measurements for a 500-particle data set were derived from STEM images. The RNase-treated Ty medium-size fraction (8.6 mg/ml) was diluted 100 times with 40 mM ammonium acetate and applied to a 600-mesh copper grid covered with an ~400-Å-thick carbon film, which was previously glow discharged in air for 10 s, and washed with a few droplets of 20 mM ammonium acetate at pH 7 (ammonium carbaminate). After 1 min the specimen was washed with three droplets of the same buffer, and a droplet of tobacco mosaic virus (TMV) suspension was added. The specimen was further washed with seven droplets of the ammonium acetate buffer, blotted, and vitrified by plunging into liquid ethane.

The specimen was transferred to a Vacuum Generators HB5 STEM equipped with a cold stage described by Homo (23), operating at ~150°C; after transfer the specimen was still vitreous. The specimen was freeze-dried and left overnight at ~100°C. It was then warmed up to room temperature in the microscope. Microscopy was carried out at ~100°C, and areas to be photographed were preselected at low magnification (>10,000) in order to reduce mass loss. Images consisting of 1,024 by 1,024 pixels of 10.9 by 10.9 Å were recorded in a dark field at an electron dose of approximately 4 electrons/Å² and at an acceleration voltage of 80 kV.

The mass analysis was carried out according to the method of Freeman and Leonard (17), with TMV as the internal calibration standard. Only those particles that appeared well preserved were analyzed. Measured mass values were converted into number of subunits, based on the 43-kDa molecular mass of the p1-381 protein.

Image processing and three-dimensional reconstruction. Images were digitized either on a Perkin-Elmer flatbed scanner at 9.1 Å/pixel or on a charge-coupled device camera at 8.2 Å/pixel. Image processing was carried out mainly with the Spider and Web image processing system (16) on a VAX workstation. This included the particle selection, thresholding, and density scaling. FORTRAN 77 programs and Simpleplot Mark 2 (Bradford University Software Services, Ltd.) were used for plotting of common-lines phase residuals and radial density profiles, and for radius measurement. The Semper image processing system (Synoptics, Ltd., Cambridge, United Kingdom) was used on a PC (personal computer) for generating two-dimensional contour maps of the mass-density profiles, and for radius measurement. The problem was formulated as a test of significance between two regressions, in order to determine whether fitting the sum of two Gaussian functions to each distribution is a statistically better fit than that with a single Gaussian function (the null hypothesis). The adjustable parameters were the positions, heights, and widths of the Gaussians. A standard F test was used to statistically assess the relative importance of the residuals for the one-Gaussian and two-Gaussian curve fits (11).

Radius measurement. STEM and cryo-EM particles were centered (32) and cross-correlated with a series of models over the size range of Ty-VLPs present in the data sets. The particles were classified by size according to the best match with the model series. Radius measurements were taken from the radial density profiles of each centered image. To minimize the effect of noise, relative radius measurements were taken at the mean height of the radial-density curves between their maximum and minimum values on the outside of the shell. This was performed by using Neville’s method of inverse iterated interpolation (7), which works by fitting a series of n-degree Lagrange polynomials to the data points.

RESULTS

STEM mass-radius analysis indicates an icosahedral T-number series. Figure 1 shows a STEM image of p1-381 Ty-VLPs used in the mass-radius analysis. These particles had their packaged RNA removed by RNase digestion. The STEM mass measurements of 450 p1-381 Ty-VLPs were obtained by integrating particle scattering over background. A histogram of the mass distribution (Fig. 2, right panel) shows two strong peaks corresponding approximately to particles composed of 180 and 240 subunits. These two size classes are compatible, therefore, with T=3 and T=4 quasi-equivalent icosahedra, respectively. Testing the statistical significance of the regression of the sum of two Gaussian curves against a single Gaussian fit to the mass data gave an F value of 10.46. The P of F3,30 > 10.46 is 0.000014, giving >99.99% certainty that the sum-of-two-Gaussians model is a statistically better fit than the single-Gaussian model for the mass distribution.

With the same data set, relative particle radii were measured from radial-density profiles to produce the radius histogram in Fig. 2 (top). The same statistical test was applied to the radius histogram, which gave an F value of 1.99. In this case, the P of
$F_{3,18} > 1.99$ is 0.15. Hence it is only 85\% certain that the sum of two Gaussians is a statistically better fit than the single-Gaussian model. This result is therefore inconclusive, although it is compatible with a two-Gaussian fit. Radii of $T_5^3$ and $T_5^4$ icosahedra differ by only 15\%, making it difficult to resolve the difference in STEM images.

In order to determine the relationship between mass and radius, these data were paired and shown to have a high linear correlation (coefficient $\approx 0.85$). The paired mass-radius data were plotted as a two-dimensional histogram, displayed as a contour map in Fig. 2 (bottom left). This clearly reveals two peaks, which correspond to radii of 166 and 191 Å and $T_3$ and $T_4$ subunit numbers, respectively. The radius ratio (0.869) is consistent with the predicted ratio for related $T_3$ and $T_4$ icosahedra ($\sqrt[3]{3/4} \approx 0.866$). Thus, there are strong statistical indications that the $T_3$ and $T_4$ icosahedral structures are abundant forms of Ty particles.

Three-dimensional structures of $T_3$ and $T_4$ Ty-VLPs. Figure 3 shows a cryo-EM image of Ty-VLPs used for icosahedral reconstruction. A total of 300 images of Ty particles were selected from the set of images. Radius analysis of this particle set indicated that 50 and 80 particles could be identified as belonging to $T_3$ and $T_4$ radius classes, respectively. Three-dimensional reconstructions of these two classes were determined by a combination of common-lines and cross-correlation methods. The surface representations (Fig. 4) of the reconstructions are shown looking down icosahedral two-, three-, and fivefold symmetry axes. Both structures contain spikes projecting from the capsid shells. In the $T_3$ shell there are 60 spikes, arranged around the fivefold axes. The larger, $T_4$ particle has, in addition to the fivefold spikes, blunt spikes on the threefold axes, giving 80 spikes in total. The spikes are slightly skewed relative to the icosahedral vertices.

Comparison of the capsid spikes and evidence for trimer clustering. In order to compare the different spikes in the $T_3$ and $T_4$ particles, sections bisecting the spikes were taken from each of the reconstructions. Only one section was needed for the $T_4$ particle, since the slice bisecting the threefold-position spike also bisected one of the spikes around the fivefold axis. It can be seen from contour plots of these sections (Fig. 5) that varying degrees of distortion of the p1-381 subunits were needed to form the three different kinds of spikes in the two icosahedral structures. The $T_3$ spike is about 66 Å in height, whereas the fivefold $T_4$ spike protrudes about 76 Å from the shell. The threefold $T_4$ spike is blunt and about 65 Å in height. Trimer clustering of the subunits of Ty-VLPs is indicated by the observations that the number of spikes is one-third of the number of subunits in each case and that the fivefold adjacent spikes appear to have a triangular shape. The apparent differences between the threefold spikes and fivefold
spikes may reflect the relatively low resolution (38 Å) of the reconstruction or their different quasi-equivalent environments.

**DISCUSSION**

The T₅₃ and T₅₄ particle reconstructions presented here provide the first structures for a retroviral core-like particle. It was expected that large holes (>30 Å in diameter) would be found penetrating the surfaces of the Ty-VLP structures, in order to explain how RNase A molecules are able to gain access to the interior of the capsids and digest the Ty RNA (9). However, neither of the reconstructions indicates the presence of such holes. It seems likely that the absence of holes in the structures presented here is a function of the relatively low resolution of the reconstructions. A porous structure becomes necessary if one suggests that the capsid has a role in orchestrating reverse transcription and integration. Holes would allow both low- and high-molecular-weight components of these reagents to pass in and out of the structure. This idea is supported by the observations that Ty-VLPs mediate integration in vitro and that some mutations in the TYA gene which do not disrupt particle formation reduce transposition rates by more than 50 times (14). In addition, it is clear from studies with human immunodeficiency virus (HIV) that retroviral reverse transcription and integration go on in the context of capsid-derived complexes (5, 6). An alternative hypothesis is that release of the integration complex occurs via partial disassembly of the particles, rather than through preexisting holes in the capsid.

The functional significance of the T₅₃ and T₅₄ icosahedral structures is not yet known. Moreover, the histograms in Fig. 2 suggest that larger, T₅₇ structures might also be present in the population. Larger structures could not be included in sufficient numbers for structural analysis, because they were less well preserved.

Interestingly, the reverse transcriptase-containing hepatitis B virus core protein also assembles into T₅₃ and T₅₄ structures (13, 24). These capsids package pregenomic RNA and mediate its reverse transcription to double-stranded DNA. Analysis of the structures formed by the capsid protein upon expression in *Escherichia coli* shows that they are a mixture of dimer-clusters T₅₃ and T₅₄ capsids. Quasi-equivalence requires that the T₅₃ structure must contain three, and the T₅₄ shells must contain four, quasi-equivalent conformations of the capsid protein. The additional conformation is found at the threefold positions of the T₅₄ shell, but the organization of the fivefold regions is identical between the two structures.

The Ty particles also show superimposable fivefold regions but there is a rotation, relative to the icosahedral symmetry axes, between the two. In the T₅₄ particle, a fourth conformation of the subunit must also be added to three conformations present in the T₅₃ form. However, these new subunits form a trimer with each other rather than contributing to dimers as in the hepatitis B capsid. The fivefold regions of the T₅₄ form rotate to accommodate the new trimer in a region where the fivefold subunit clusters of the T₅₃ form interact. This may explain the difference in geometry between the fivefold clustered and threefold trimers.

Retroviruses such as HIV form pleomorphic, nonspherical nucleocapsids which do not obey the strict conventions of quasi-equivalence, although there is some evidence for assembly of retroviral cores with local symmetry (21, 31). HIV Gag particles are about 1,500 Å in diameter, and the maximum dimension of the mature, conical nucleocapsid is ~800 Å, substantially larger than the 300- to 400-Å Ty-VLPs analyzed here. The N-terminal domain of the HIV capsid protein p24 has been shown to be a novel, α-helical structure by nuclear magnetic resonance and X-ray crystallography (20, 22, 30). In the crystal structures it forms dimers, but the contacts are different in the two cases, which contain the capsid N-terminal domain in complex with an antibody fragment in one structure (30) and with cyclophilin A in the other (20). The orientation of p24 in

![FIG. 3. A cryo-EM image of the RNase-treated Ty particles, showing good contrast and preservation of particles. The data are examples of those used in the three-dimensional reconstructions of Ty-VLPs. Bar, 450 Å.](http://jvi.asm.org/)
the HIV capsid is suggested by the locations of binding sites and the cleaved N and C termini. However, the structure of the C-terminal domain, which is important in directing capsid assembly (34, 37), has not been determined. The structure of the HIV matrix protein p17, the amino-terminal subprotein of the Gag precursor, has been shown by nuclear magnetic resonance to resemble interferon-γ (26, 27). A crystal structure of simian immunodeficiency virus p17 shows that it forms trimers which are thought to be important in virus assembly (33). In addition, the transmembrane component of the env-encoded protein from Moloney murine leukemia virus forms stable trimers (15) and is thought to interact with matrix within the particle. The latter observations are consistent with a trimer-clustered retroviral core, but it is unclear how this relates to the p24 dimers seen in the crystal structures. The trimer hypothesis is also consistent with the present observations suggesting that Ty-VLP spikes are formed from p1-381 trimers. A trimeric unit is further supported by recent gel filtration data from which stable p2 trimers have been identified in vitro (25a). It seems likely, therefore, that an understanding of the subunit packing of Ty-VLPs may also help to elucidate the assembly properties of the other retroelements.

ACKNOWLEDGMENTS

K.J.P. thanks the BBSRC and British Biotechnology Limited, Oxford, for provision of a CASE studentship during the course of this work.

The TMV used was a gift of D. Peters (Wageningen). We thank K. R. Leonard for providing mass analysis software and J. Kenney, H. AL-Khayat, and D. Bhella for comments on the data analysis and manuscript.

REFERENCES