Molecular shape of *Lumbricus terrestris* erythrocruorin studied by electron microscopy and image analysis

Egbert J. Boekema and Marin van Heel

*Fritz-Haber-Institut der Max-Planck Gesellschaft, Berlin (F.R.G.)*

(Received 17 March 1988)

Key words: Erythrocruorin structure; Electron microscopy; Image analysis; (*L. terrestris*)

The molecular structure of erythrocruorin (hemoglobin) from *Lumbricus terrestris* has been studied by electron microscopy of negatively stained particles. Over 1000 molecular projections were selected from a number of electron micrographs and were then classified by multivariate statistical image-processing techniques. The two main groups of top and side views were each subdivided into smaller classes with significantly different features. About half of the top-view projections exhibit perfect hexagonal symmetry at the current resolution of about 2.0 nm, while the other top views lack this symmetry, probably as a result of tilting of the molecules relative to the carbon support film. The side views were separated into two 'families', each associated with the two different stable side-view positions the molecules can take. From these narrow stable side-views, the two families of projections are, again, generated by tilting. The symmetry properties of the three non-tilted projections show that *Lumbricus* erythrocruorin has a pointgroup D6 (622) symmetry rather than D3 (32).

Introduction

Many annelids contain giant extracellular erythrocruorins (alternatively called hemoglobins) freely dissolved in their blood, where they function in oxygen transport. Erythrocruorins are characterized by a distinct subunit architecture: the molecules consist roughly of a hexagonal bi-layer with an apparently empty central cavity [1–3]. Moreover, electron microscopical studies have revealed the shapes of erythrocruorins from different species to be very similar (reviewed in Ref. 4). The exact subunit composition of this highly symmetric structure is not yet known. In one proposed model the molecule is assumed to be built of 216 copies of three different polypeptides

[1]. Erythrocruorin of *Lumbricus terrestris* has been crystallized in different crystal forms suitable for X-ray diffraction studies [1,5,6].

Image-averaging techniques using either images of two-dimensional crystalline arrays of the molecule [7] or individual molecular images [8,3,4] have been applied to this molecule before. Averaging techniques are important to enhance the poor signal-to-noise ratio characteristic of unprocessed electron micrographs of biological macromolecules. In our current study, we present the analysis of a much larger number of individual erythrocruorin molecules than before, with therefore a better statistical significance. Moreover, our automatic classification techniques [9,11] allow us to apply the concept of averaging out noise, while still discriminating between subtle systematic differences present between the images in the data set. This technique has already been applied successfully in studies of a number of different ob-
jects [12–14]. Our study was undertaken in the expectation that knowledge of low resolution (> 2 nm) structural information, obtained by our electron microscopical techniques, may trigger advances in the high-resolution structure determination with X-ray diffraction techniques.

Materials and Methods

Sample preparation and electron microscopy

Three-dimensional crystals of erythrocruorin grown from whole blood [1,5] were redissolved in a 10 mM sodium phosphate buffer of pH 7 to yield a highly purified solution directly suited for electron microscopical sample preparation. The crystals (type II and type III [1,5]) were kindly provided by W.E. Royer and W.A. Hendrickson. During the analysis we noticed that the molecules are slightly better preserved if fixed with a 5% glutaraldehyde solution (diluted from a 50% aqueous solution under argon, from Fluka) in the same phosphate buffer. After 2 h, the fixation procedure was stopped by diluting the sample with 20 aliquots of buffer, after which the microscopical grids were prepared with the droplet technique using 1% uranyl acetate as negative stain. The grids were also rinsed with buffer. Electron micrographs were taken with a Philips EM 300 electron microscope at x 70000 magnification. Care was taken to avoid irradiation of the areas to be imaged prior to the actual recording of the micrograph. An estimated specimen irradiation of about 2000–40000 electrons/nm² was used.

Image processing

The electron micrographs were digitized using a Datacopy Model 610F electronic digitizing camera (Datacopy Corporation, Long Beach, U.S.A.) mounted on a standard 6 × 9 cm enlarger equipped with a direct-current driven halogen lamp. The scanning step used was 32 μm, corresponding to a pixel (image element) size of 0.46 nm at the specimen level. All image analysis was carried out within the IMAGIC software system [15] on a VAX 11/780 computer. The molecular images (a total of 1017 individuals) were selected interactively from 41 digitized micrographs, using a raster-scan image display system. About 30% of the selected molecules were glutaraldehyde fixed, the others were from unfixed samples; only top-view images showed some differences, see results). These were then pretreated by band-pass filtering to suppress the very low and the very high spatial frequencies. The very low spatial frequencies are typically associated with effects like stain gradients around the molecules which are not directly structure related. The very high spatial frequencies, on the other hand, normally contain only noise. Subsequently, the individual molecular images were surrounded by a circular mask to cut away unnecessary background, and their density values were normalized within this mask to zero average density and a standard variance value [12].

The pretreated images were aligned relative to each other by a multi-reference procedure [12]: alignment is a prerequisite for the subsequent pattern recognition and averaging procedures. The aligned images then were submitted to an eigenvector–eigenvalue data compression procedure ('correspondence analysis' [9,10,16]). Using a variance-oriented automatic classification scheme [11] operating in the compressed data space resulting from the correspondence analysis, we then partitioned the data set into between 8 and 16 'classes'. We refer to these Multivariate Statistical Analysis techniques as 'MSA'. During the classification about 10–20% of the population of original images was rejected based on their (too) large contribution to the variance of the data set. The rejected individuals are mostly uncharacteristic images which appear rarely (such as damaged particles). The images comprising the various classes are summed together to yield average-class images with enhanced signal-to-noise ratio.

Results

Two distinct sub-populations of molecular projections could readily be recognized: top views (also described as face-on or hexagonal views) and side views [1,3] (Fig. 1). Since these two views differ so clearly, the data set of 1017 images was split into subsets of 421 top views and 596 side views which were each processed independently.

Analysis of top views

The very large erythrocruorin molecule gives
rather clear projections in the electron micrographs (Fig. 2) which are easy to align. The top-view data set were first aligned relative to a 6-fold symmetrized original image, which was well preserved but otherwise chosen at random from the data set. Three further alignments with refined 6-fold symmetrized references were subsequently applied, but no further translational shifts larger than one pixel were observed. The data set were then partitioned into 12 class averages containing 18–50 members each (Fig. 3).

A comparison of the classes given us an impression of the variations present in the data set. Since 6-fold rotational symmetry was imposed on all references, there are six equivalent optimal positions into which the projections can be brought by the alignments. This technique is used to avoid a bias of the data set towards one particular asymmetric reference image [13]. As a consequence, the classification procedure leads to several classes that are very similar and only differ in their rotational orientation. The classes of Fig. 3F–L illustrate this point: they all have an asymmetric property in that two or three 1/6th parts of the structure are ‘faded’ similarly but in different rotational directions. The rotation between images
Fig. 2. A gallery of digitized and pretreated images of top- and side-view projections of erythrocrurin molecules.

3J and 3K is, for instance, 180°; that between 3I and 3J is 60°. The class-sums in Fig. 3A–E show a more regular shape with a close to 6-fold symmetry.

In a further alignment step, a new reference image was used (in addition to the earlier symmetrical one) which was created by summing the classes 3F–I with the blurred feature brought into rotational register. About 250 images showed a higher correlation coefficient in the multi-reference alignment [12] relative to this new reference image, and these were thus aligned accordingly. The data set were then partitioned into four classes (Fig. 4), ignoring the images of the class shown in Fig. 3D (see below). The resolution within class 4A was calculated with the Fourier ring correlation method [17] and was found to be about 1.7 nm. Within this resolution, the projected structure showed little departure from hexagonal symmetry which was therefore imposed. The other classes typically show a resolution of around 2.0 nm.

For a detailed analysis of the more symmetric top views (Fig. 3A–E), these views were partitioned into two classes by a separate MSA pass. The main inter-class difference was found to lie in the size of the images. One (close to Fig. 3A) had an edge-to-edge diameter of 27 nm, the other (close Fig. 3D) 28.3 nm. A second difference was found in the central cavity of the structure: a central mass was clearly visible in the larger image but almost absent in the smaller one. Part of the data set (about 30% of the molecules) have been
fixed with glutaraldehyde, as described in the Materials and Methods section. We observed that from the unfixed sample about 15% of the molecules had a central mass, whereas only 3% of the fixed molecules (100 examined) showed such a central mass.

**Analysis of side views**

The analysis of the side views was accomplished in two iterations of the multi-reference alignment procedure, using the classes obtained from the first MSA classification as references for the next alignment phase. After this last alignment phase, the images were partitioned into 12 classes with the MSA procedures. The final classification is illustrated in Fig. 5 which shows the data set and the classes projected onto the plane spanned by the most significant eigenvectors (two and three; number one points to the center of mass of the data set [11]), i.e., the directions associated with the largest inter-image variance [11] within the data set. The best nine classes (Fig. 5A–I) each contain 21–88 individual molecular images. Three bad (= large internal variance) classes, with respectively 16, 21 and 31 member images, are not shown in Fig. 5, but their positions are indicated on the map. The map indicates that all images and classes belong to two ‘families’ of views associated
with the two possible narrow stable positions of the molecule (Fig. 5).

**Discussion**

The top (hexagonal) views of molecules from a large number of species, as studied by electron microscopy, all show close similarity [4]. One item of debate is, however, the presence of a central mass. Only a few species like *Oenone fulgida* [18], *Nephys incisa* [19] *Nephtys hombergi* [20] and *Ophelia bicornis* [7] were found to have an additional subunit in the usually empty central cavity. Image analysis of *L. terrestris* erythrocrurin hitherto failed to reveal any appreciable mass in the centre of the structure [2,3]. Our results tend to confirm this opinion. Although a small number of hexagonal projections show a central stain-ex-
Fig. 5. Classification of side view projections. The classification map shows 546 projections and their range of image variation. The vertical axis (axis 2) almost coincides with the direction of the tilting behaviour, while the horizontal axis (axis 3) mostly coincides with the separation of the data set into the two principal narrow projection views. The position of the centre of the nine best classes (SA-I) and of three bad classes (asterix *) is indicated. The number of members for classes A–D is, respectively, 98, 21, 20, 34 and for E–I 57, 61, 39, 45 and 39.

cluding pit (Fig. 3D), it is not as substantial as in the erythrocrurins mentioned earlier. Appearance of a pit is accompanied by a widening of the particle diameter of about 1.5 nm. Fixation of samples with glutaraldehyde, however, drastically reduced the number of particles showing a pit. It is therefore likely that hexagonal views with a central pit result from molecules altered during the preparation for electron microscopy. If this point of view is correct, and if the central pit is indeed a negative-stain preparation artefact, no central pits should be observed in cryo-microscopical images of the molecule. With this technique [21], the molecules are embedded in a layer of vitreous ice which is more like the in vivo environment for this biological macromolecule.

The main phenomenon in the analysis of the top views concerns the departure from 6-fold symmetry. This is a gradual process, as can be seen from Fig. 4A to D. We believe this departure to result from tilting of molecules relative to the carbon support film [22]. This interpretation is supported by the changes in the diameter of the classes: in Fig. 4D the molecule is about 2 nm narrower in the vertical direction than in Fig. 4A. Similar subtle tilting phenomena have been found using multivariate statistical analysis in analyzing projections of 50 S [23] and 70 S [14] ribosome.

Fig. 6. Explanation of the appearance of the two narrow side-view projections. The molecules can either be attached in an ‘on-edge’ position (A) or in an ‘on-point’ position (D). B and E are the corresponding projection views. C shows the difference image between B and E. The regions where the difference is positive are bright. F is the sum of B and E.
The analysis of the side views revealed two main families of classes (Fig. 5). In both groups the narrowest classes are at the same time the largest ones in terms of number of members (Fig. 5A and E). These narrow classes can be interpreted as originating from erythrocruorin molecules in two stable positions: an ‘on edge’ position with four 1/12th subunits making contact with the support and an ‘on-point’ position resting mainly on two subunits [1,3]. These positions are illustrated in Fig. 6. As could be detected, the average ‘on-edge’ projection with a length of 28.7 nm (Fig. 6B) is somewhat longer than the ‘on-point’ projection of 26.8 nm (Fig. 6E). This difference in the length of the molecular projections is also illustrated in Fig. 6C, which shows the difference image of images 6B and 6E. Finally, Fig. 6F shows the image sum of 6B and 6E, i.e., the sum of the ‘on-edge’ and ‘on-point’ side views. This sum resembles the average side-view projection obtained when the subtle differences between the two possible side views have not been taken into account [2,24].

The classification map Fig. 5 shows that the different classes form two disjunct ‘families’ of projections which are directly related to the narrow side views of Fig. 5A and 5E. Again we interpreted these classes (Fig. 5B–D and 5F–I) as originating from molecules tilted relative to the stable narrow position on the support film. The classification map suggests that this tilting occurs gradually, giving extremely tilted projections (Fig. 5C, D, H and I) as well as intermediate ones (Fig. 5B and G). The width of the projections with the largest tilt (Fig. 5D and I) are 23.4 and 20.4 nm, respectively; those of the untilted 19.2 and 18.9 nm (Fig. 5A and E). Upon tilting, one of the long edges of the projection becomes somewhat diffuse (Fig. 5B and F–I).

A special point of attention concerns the symmetry of the oligomeric erythrocruorin molecule: there is not yet a consensus regarding the pointgroup symmetry of the molecule, the two possible choices being D6 (622) or D3 (116); for pointgroup definitions see [25,26]). A D6 structure projected along the 6-fold axis in the electron microscope leads to a 6mm projection; when projected along the two different types of 2-fold axis, two different projections with mm symmetry (perpendicular mirror planes) will result. This is generally true only if the molecules are stained homogeneously. On the other hand, a D3 structure leads to an m projection along the 3-fold axis and to only one type of projection along a 2-fold axis which exhibits only 2-fold symmetry. The three principal views we obtained all show almost perfectly the symmetries predicted for the D6 structure even prior to symmetrization. Note that in Fig. 5 the mirror symmetries have not been imposed at all. We thus must conclude that erythrocruorin had a D6 symmetry at the current resolution level of around 2 nm.

Multivariate statistical classification of images as applied to our erythrocruorin data set has yielded a large number of different projections for this highly symmetric (D6) molecule. This makes this molecule perfectly suited for three-dimensional reconstruction from many individual projections [27].

Acknowledgements

We thank Drs. Wayne A. Hendrickson and William E. Royer for providing us the erythrocruorin samples and for fruitful discussions. Professor Elmar Zeitler is acknowledged for his general support of the project, and Matina Gerdsmeyer for her excellent technical assistance.

References