Single particle electron microscopy in combination with mass spectrometry to investigate novel complexes of membrane proteins

Ana A. Artenia, Marc Nowaczyk, Julia Lax, Roman Kouřil, Matthias Rögner, Egbert J. Boekema

Abstract

Large data sets of molecular projections of the membrane proteins Photosystem I and Photosystem II from cyanobacteria were analyzed by single particle electron microscopy (EM). Analysis resulted in the averaging of 2D projections from the purified complexes but also in the simultaneous detection and averaging of 2D projections from large contaminating complexes, which were present in frequencies as low as 0.1%. Among them T-shaped and L-shaped contaminants were found. The L-shaped particles could be assigned to Complex I just from the shape, although no Complex I from a cyanobacterium has been structurally characterized. A systematic comparison by single particle EM and mass spectrometry of two differently purified Photosystem II complexes resulted in the assignment of PsbZ, a small peripheral subunit of 6.8 kDa, within the structure. Together these data suggest that screening for membrane protein structures by single particle EM and mass spectrometry may be a new approach to find novel structures of such proteins. We propose here a scheme for searching for novel membrane protein structures in specific types of membranes. In this approach single particle EM and mass spectrometry, after pre-fractionation using one- or multidimensional protein separation techniques, are applied to characterize all its larger components.

Keywords: Electron microscopy; Mass spectrometry; Photosystem I; Photosystem II; Complex I; Thermosynechococcus elongatus

1. Introduction

A proteomics analysis of transmembrane proteins is often focussed on the discovery of novel protein components or on the study of modifications by phosphorylation and so on. Such a systematic analysis is based on one- or multidimensional protein/peptide separation depending on the complexity of the sample followed by extensive (tandem) mass spectrometric analysis (Staub et al., 2003). In comparison with sequence tags and genomic databases, the presence of fully novel or multiple, closely related proteins can be demonstrated. Studies to test this approach have been carried out on several types of membranes, such as mitochondrial membranes from the yeast Saccharomyces cerevisiae (Everberg et al., 2004). In this paper, we will focus on photosynthetic membrane proteins from oxygen evolving organisms. As an illustrative example for a proteomics analysis, Hippler and co-workers were able to reveal nine Lhca antenna proteins associated to Photosystem I (PSI) from the green alga Chlamydomonas reinhardtii. Three of them were identified for the first time (Staub et al., 2003). These proteins have a mass around 25 kDa and are functionally related to the light-harvesting complex II (LHCII) antenna proteins of Photosystem II (PSII), which will be discussed below.
Because of the small dimensions of proteins with a mass below 50 kDa and the low signal-to-noise ratio of transmission EM images of biomacromolecules, single particle EM analysis is not (yet) able to reveal much detail in such proteins, although it is our experience that individual trimers of LHCII (75 kDa) can already be recognized, even without averaging, by their typical shape. But the potential of single particle EM is certainly on the larger complexes and their association into supercomplexes.

Over the last 20 years, single particle analysis of electron microscopy (EM) projections has become a well-established technique to obtain structural information about large biomacromolecules at a resolution of 10–20 Å (Frank, 2002; van Heel et al., 2000). In some favorable cases a much higher resolution has been achieved on ice-embedded virus particles or other complexes with a high symmetry (Ludtke et al., 2004; van Heel et al., 2000). Single particle averaging is simple and attractive if applied to negatively stained specimens with a mass between about 300 and 2000 kDa, since one can process several thousands of projections within a few days, yielding 2D projection maps of at least 20 Å resolution. The statistical analysis and classification procedures used in single particle analysis are often effective in the sorting of different projection views originating from different conformations or subunit compositions.

During the analysis of a large set of 16,600 projections of PS2 complexes, purified from chloroplast grana membranes from spinach, we noticed a strong heterogeneity of particles due to a variable binding of peripheral subunits around the PS2 core complex (Boekema et al., 1999). As a result, 10 protein complexes with a different subunit composition or arrangement were resolved. However, in this analysis which consisted of several cycles of alignments, multivariate statistical analysis and classification about 25% of the projections could not be assigned to any of these PS2 particles (Boekema et al., 1999). One triangular shaped protein contaminant with a diameter of about 200 Å was present in about 200 copies. By averaging these projections, the contaminant could be designated as a multimer of seven copies of trimeric LHCII (Dekker et al., 1999). This assignment was possible for two reasons. First, a high-resolution structure of LHCII trimers was solved by electron microscopy (Kühlbrandt et al., 1994). Even at the 20 Å level the strong similarity between the contaminant and the LHCII structure was apparent. Second, it is known that single LHCII trimers which do not bind to PS2 are present in variable, but appreciable numbers within the thylakoid membrane, from which PS2 complexes were purified. However, the discovery that these trimers would be present as supercomplexes with seven LHCII trimers was a surprise.

These results on a rather heterogeneous set of EM projections indicate that single particle EM analysis has the potential to detect specific unknown large complexes in large data sets and to obtain reasonably well defined 2D maps, even if such contaminants are present in small numbers. Usually we will not be able to assign the structures of the contaminants to specific proteins. Nevertheless, in several studies we could detect remarkably large contaminants, such as a T-shaped particle that co-purified with trimeric PS1 from the cyanobacterium Gloeobacter violaceus (Mangels et al., 2002). These studies also indicated that it would be likely that by checking “rather pure” membrane protein complexes for contaminants more of such novel structures could be found if averaging of several dozens to hundreds of projections is possible.

In a next step of reasoning, we consider here that if it were possible to assign such novel structures to polypeptides in a rather systematic way the technique of single particle analysis could get a novel useful application in the expanding field of proteomics. Here we would like to propose a method which combines single particle EM and protein mass spectrometry. Its strategy is specially meant for membrane proteins, which are harder to characterize by a proteomics analysis than water-soluble proteins and aims to discover novel and larger complexes rather than the single smaller proteins.

To test the possibilities of combining single particle EM with protein mass spectrometry, we describe the comparison of two batches of PS2 from the cyanobacterium Thermosynechococcus elongatus. PS2 is a multimeric membrane protein complex which catalyzes the light-driven oxidation of water (Hankamer et al., 1997). It is a dimer with over 25 subunits (PsbA–PsbZ and others) per monomer depending on the species. About half of them are small intrinsic membrane proteins with a molecular weight below 10 kDa, which extremely complicates the mass spectrometry analysis. However, in a detailed characterization of our His-tagged PS2 preparation almost all of the small subunits could recently be assigned in a rapid and gel-free MALDI-TOF approach (Nowaczyk et al., unpublished results). In contrast to WT PS2 which was purified by a standard HPLC procedure (Kuhl et al., 2000), the His-tagged PS2 complex shows the presence of an additional mass which can be assigned to PsbZ. Single particle EM analysis can reveal the presence of PsbZ in the His-tag preparation and localize it at the periphery of the PS2 complex.

These and other results on proteins contaminating the photosynthetic membrane proteins PS1 and PS2 will be discussed in the framework of a proposal for a combination of the single particle EM technique with a mass spectrometry analysis.
2. Materials and methods

2.1. Purification of WT- and His-tagged photosystem II

PS2 complexes from *Thermosynechococcus elongatus* were prepared by hydrophobic interaction chromatography followed by ionic exchange chromatography as described in Kuhl et al. (1999), with the following modification: after solubilization of the thylakoid membranes by dodecyl-β-D-maltoside (β-DM) and centrifugation, the supernatant was loaded onto a 27 ml sucrose gradient consisting of 14% sucrose (w/w) in buffer B (20 mM Mes, pH 6.5, 10 mM CaCl₂, 10 mM MgCl₂, 0.5 M mannitol, and 0.03% β-DM); this gradient was layered on top of an 80% sucrose cushion in the same buffer. After overnight centrifugation (18 h, 4 °C, 83.000 g), the core centers—visible as a green band in the upper layer—were collected with a syringe and prepared for the first chromatographic step.

For the preparation of His-tagged PS2 complexes, thylakoid membranes were solubilized in 20 mM Mes, pH 6.5, 10 mM CaCl₂, 10 mM MgCl₂, 200 mM (NH₄)₂SO₄, 1.2% β-DM, and 0.5% Na-cholate at a Chl concentration of 1 mg/ml. After centrifugation at 45000 g for 90 min at 4 °C, the supernatant was loaded onto a chelating Sepharose fast flow column (Pharmacia) which had been equilibrated with starting buffer (50 mM Mes, pH 6.5, 10 mM CaCl₂, 10 mM MgCl₂, 300 mM NaCl, 10% glycerol, 0.02% β-DM, and 1 mM imidazole). The column was then flushed with 4 volumes of equilibration buffer at a flow rate of 2 ml min⁻¹. PS2 core complexes were eluted by a linear gradient from 1 to 200 mM imidazole and dialyzed overnight against buffer B. To separate PS2 monomers and dimers, the resulting protein solution was loaded onto a second column (IEC) as in (Kuhl et al., 2000).

2.2. Mass spectrometry

For the matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) analysis, PS2 preparations (adjusted 1 mg/ml Chl) were mixed with a saturated solution of α-cyano-4-hydroxy cinnamic acid at a ratio of 2:5 and spread on a MALDI-target (0.4 μg total Chl). After calibration with insulin, protein masses between 6000 and 7000 (m/z) were determined in linear mode with a MALDI-TOF mass spectrometer (Autoflex, Bruker Daltonics).

2.3. Electron microscopy and image analysis

Negatively stained specimens were prepared with 2% uranyl acetate on glow-discharged carbon-coated copper grids. From both types of purified PS2 particles, images were recorded with a Gatan 4K slow-scan CCD camera on a Philips CM20FEG electron microscope. 2000 x 2000 pixel images were recorded at 66850× magnification with a pixel size of 15 μm and a binning factor of 2, corresponding to a size of 4.49 Å at the specimen level. The data sets were analyzed with Groningen image processing (Grip) software on a PC cluster. Resolution in averaged single particle 2D projections was measured according to van Heel (1987).

Averaging of contaminants present at frequencies well below about 2% was carried out as follows. First, all picked projections were aligned to a reference with features of a potentially interesting contaminant (either one single noisy projection or a projection map from another data set). The projections with the highest correlation coefficient were copied out (usually up to 1% of the complete set) and this subset was submitted to statistical analysis.

3. Results

3.1. PS2 EM analysis

Two sets of about 15,000 projections of dimeric PS2 complexes from the cyanobacterium *Thermosynechococcus elongatus* were analyzed. After several cycles of multireference alignments, statistical analysis and classification the best top-view projections from each set were averaged. The non-symmetrized 2D projections were obtained with a resolution of 14Å (Figs. 1A and B). By comparing the two dimers, it appears that the monomers of each are highly similar, even without imposing any 2-fold symmetry, except for one site at the outer tips outlined in the difference image (Fig. 1C). This tip is larger in the particles with a His-tag. Comparison of our EM data with the 3.7 Å X-ray structure from PS2 (Ferreira et al., 2004) indicates that two membrane-spanning α-helices at the tip of the projection are responsible for this difference (Fig. 1D). These helices were assigned by Ferreira et al. to subunit PsbZ in their structural model of the PS2 complex.

3.2. PS2 analysis by MALDI-TOF

The 1D SDS-PAGE pattern showed no difference between WT- and His-tagged PS2 (data not shown); however, the resolution in the region below 10 kDa is very weak, so that different protein bands could be shielded by overlay. For this reason we developed a technique to apply PS2 complexes directly to MALDI-TOF analysis to determine the entire mass of the small subunits (Nowaczyk et al., unpublished results). With this method we could assign a single mass in the range of 7000–8000 m/z to subunit PsbZ in the sample of the His-tagged PS2 complex (Fig. 2A); this mass is not present in the WT PS2 preparation (Fig. 2B).
3.3. L-shaped contaminant of PS2

In the PS2 preparations studied above one contaminant appeared. L-shaped particles were present with an abundance of 2 and 10% in the batches of PS2 with the His-tag and the ones without, respectively. Image analysis indicated that this L-shaped particle appeared in two dominant views (Figs. 3A and B). The longer horizontal “leg” of the particle does not show much internal detail, which is in general a typical feature for negatively stained membrane proteins in side-view position. Therefore this part is probably a membrane-embedded protein part. Because projections in both views have the same overall dimensions, they likely come from the same particle and thus only have a different handedness in respect to the carbon support film.

3.4. T-shaped contaminant of PS1

Another large contaminant of photosynthetic membrane proteins seen in other samples is a T-shaped particle which copurifies with PS1 trimers purified from *Gloeobacter violaceus* (Mangels et al., 2002). It has an overall length of 190 Å, consisting of a bilobed scaffold with a width of 130 Å and a height of 65 Å, from which a 120 Å long stem extends (Fig. 4B). In comparison to trimeric PS1 (Fig. 4A) it appeared with a frequency of about 20–30% in the batch which was used for EM analysis. A search in data sets from other PS1 particles (see Section 2) showed that it was also present with 70 copies among projections from the cyanobacterium *Synechococcus PCC7942* (Boekema et al., 2001), corresponding with a frequency of only 0.1%. After inspection of original micrographs from which these projections were picked the same frequency was found. Another more spherical contaminant is present in different views in PS1 data sets from *Synechococcus PCC7942* (Figs. 4D–E) (Boekema et al., 2001). It is also found in preparations of another cyanobacterium (*Synechocystis PCC 6803*, Yeremenko et al., 2004).
4. Discussion

Before discussing the role which single particle EM could play in combination with proteomics, we will first discuss our non-systematic observations on contaminants, presented above. For a number of reasons it is possible to assign the L-shaped projection views to molecules of Complex I in side-view position. First, to our knowledge Complex I is the only known major membrane complex which is L-shaped. Second, just as in our data set it appears in all negatively stained single particle EM specimens only in side-views. Third, Complex I sometimes co-purifies with PS2 (Berger et al., 1991). However, the assignment is tentative because no cyanobacterial Complex I has been analyzed by single particle EM. Interestingly, the upper leg of the complex is appreciably shorter (Figs. 3A and B) than found for Complex I from other species (Yagi and Matsuno-Yagi, 2003). This makes sense because three hydrophilic subunits which are responsible for the NADH dehydrogenase activity and which are considered to form the upper unit of the vertical leg (red mass in Fig. 3C) are lacking in cyanobacteria (Friedrich et al., 1995). These subunits have a mass of about 150 kDa in total and the lack of this amount of mass seems in accordance with the much smaller vertical leg observed in our projections (Fig. 3C).

The T-shaped particle is more difficult to assign in the absence of a specific biochemical characterization. Most remarkably, however, it is about 300 times more abundant in *Gloeobacter violaceus* than in other cyanobacteria. In contrast to *Synechocystis* and *Synechococcus* species this cyanobacterium lacks a special thylakoid membrane in which PS1, PS2, and other photosynthetic membrane proteins are located (Rippka et al., 1974). Therefore, PS1 is a component of the cytoplasmic membrane, which is the only intracellular membrane in *Gloeobacter*. If the T-shaped particle is always restricted to the cytoplasmic membrane this could give a possible explanation for its low frequency in the PS1 batch of *Synechococcus* PCC7942. Due to its long stem, it could be that this particle is involved in protein secretion. The rather spherical contaminant of unknown origin (Figs. 4D–F) is likely a water-soluble protein, given its excellent staining profile.

4.1. Assignment of a PS2 subunit

The analysis of two different PS2 preparations by single particle analysis and mass spectrometry revealed that the PsbZ subunit was present in His-tagged PS2 complexes but missing in the WT preparation. The only difference between the two procedures is the first HPLC step. One could speculate that affinity chromatography is more gentle compared to the HIC-column step of the WT preparation but studies with different preparations showed that the amount of PsbZ varies between different batches and with more sensitive techniques traces of PsbZ could also be detected in the WT preparation (data not shown). Nevertheless, we could show in the present study that the combination of single particle analysis and mass spectrometry is a powerful tool to assign a single subunit in the 2D projections. This observation encourages us to propose a scheme for a more systemic search for protein complexes by single particle EM and proteomics tools, especially mass spectrometry.

4.2. Proposed scheme for EM plus proteomics analysis

The following strategy specially focuses on a characterization of membrane proteins by proteomics. The first step is to detect novel 2D structures present in a specific
protein fraction of a detergent-solubilized membrane by single particle EM. It could be a specialized type of membrane, like the thylakoid membrane from photosynthetic organisms, but also a rather non-specific cytoplasmic membrane. The idea is to study all the larger proteins and the easiest way to obtain them would be a one-step size-exclusion chromatography purification or dialysis with a membrane with a cut-off of 100kDa to remove the abundant smaller proteins and fragments. The second step would be to obtain a batch from this membrane in which single particle EM indicates that the complex of interest is more abundant or largely absent. This could be achieved by applying further different partial purification procedures, like sucrose gradient centrifugation or dialysis with membranes with a different cut-off. Another possibility would be to change the detergent used for solubilization. We also like to emphasize here that partial solubilization often yield larger complexes (Boekema et al., 1999). Third, if differences in the frequencies of observed novel structures complexes are detected by single particle EM, SDS–PAGE patterns from the various purified membrane protein batches should be inspected for major differences. Thereafter, mass spectrometry (MALDI-TOF and other techniques) should be used to assign protein bands with a variable intensity. Since the genomes of many organisms have already been sequenced an unambiguous assignment could be achieved in most cases. If necessary additional de novo synthesis could be of help.

Further studies could also be undertaken by subsequent purification of the novel structures to homogeneity. However, this is not the prime objective of our proposal. Moreover, purification to homogeneity of large membrane proteins can be tedious and at the risk of the loss of protein components. Actually, the first indication for the heterogeneity of PS2 due to loosely attached peripheral membrane-bound subunits was obtained by the analysis of single particles from disrupted thylakoid membranes without any purification at all (Boekema et al., 1998). The advantage of EM over many other biophysical methods is certainly the detection of single molecules. Single particle EM should take full advantage out of this. If so, we can expect that by application of our proposed scheme interesting complexes or transient supercomplexes in many membranes will be discovered.

5. Conclusions

1. Membrane protein contaminants can still be detected and averaged from mixtures of membrane proteins by single particle analysis EM if present in frequencies of 0.1%. To get statistical meaningful projections this requires data sets of at least about 20,000 projections. Analysis of such numbers can be performed by a state-of-the-art PC cluster with about 16 processors. This is necessary for performing the most time consuming step in single particle EM, which is the statistical analysis.

2. Mass differences between membrane proteins in the order of 3% are still detectable, as shown for the PsbZ subunit of PS2.
3. Clear differences in frequencies between structures seen by single particle EM can be correlated to clear differences in intensities of protein bands on SDS–PAGE. Since basically a protein band can be assigned by mass spectrometry as long as it is visible with Coomassie staining, we suggest that assignment by mass spectrometry of novel structures seen by single particle EM should be possible.

4. We propose the analysis of the complete set of larger integral membrane proteins from any kind of a specific membrane after detergent disruption in a systematic way by combining single particle EM and proteomics analysis, in particular mass spectrometry.

Acknowledgments

We acknowledge R. Breuckmann for the help with the mass spectrometry measurements and we are grateful to the Protein center of the Ruhr-University Bochum for access to its central facilities. We acknowledge a grant of the Dutch science foundation NWO-ALW (Programme From Molecule to Cell) awarded to E.J.B. and support from the German Research Council DFG (SFB480 and European Graduate College 795) awarded to M.R.

References


