Biochimica et Biophysica Acta 851 (1986) 353–360
Elsevier

BBA 42131

Structure of mitochondrial F_1-ATPase studied by electron microscopy and image processing

Egbert J. Boekema [a], Jan A. Berden [b] and Marin G. van Heel [a]

[a] *Fritz-Haber-Institut der Max-Planck Gesellschaft, Faradayweg 4–6, D-1000 Berlin 33 (Germany) and*
[b] *Laboratory of Biochemistry, University of Amsterdam, Plantage Muidergracht 12, Amsterdam (The Netherlands)*

(Received 7 April 1986)

Key words: ATPase structure; Electron microscopy; Image analysis

The structure of soluble F_1-ATPase (EC 3.6.1.3) has been investigated by computer analysis of individual molecular images extracted from electron micrographs of negatively stained particles. A total of 1241 images was interactively selected from several digitized micrographs and these images were subsequently aligned relative to different reference images. They were then submitted to a multivariate statistical classification procedure. We have focussed our attention on the main 'hexagonal' view which represents some 40% of our population of images. In this view, six masses are located on the outer region of the projection which are associated with the alpha and the beta subunits of the protein. A seventh mass is located close to the centre of the hexagon, but slightly off its exact midpoint. It has the shape of the letter V and its two legs point to two of the outer protein masses, or one alpha-beta subunit pair. The corner of the V has a density as high as those of the large subunits. Possible subunit arrangements and their consequences for the mechanism of ATP synthesis are discussed.

Introduction

ATP synthesis is carried out by a large membrane-bound enzyme (ATP synthetase), which consists of a membrane-integrated 'F_0' part and a water-soluble catalytic 'F_1' component which sticks out of the membrane and seems to be connected to the F_0 part through a 'stalk'. The enzymes from different sources have been found to be very similar (reviewed in Refs. 1, 2 and 3). Bovine heart F_1 ATPase is composed of five different types of subunits: α, β, γ, δ and ϵ, with molecular weights of 55164, 51595, 30141, 15065 and 5652 respectively [4]. One protein molecule contains three copies of the α and β subunits, and one copy of each of the three smaller ones [4,5]. The total molecular weight of bovine heart F_1 ATPase is 371135 [4].

The structure of the soluble part of the ATPase system has been the subject of many investigations, especially since the presence of three α and three β subunits, containing the adenine-nucleotide binding sites, seemed to contradict many data suggestive of a dimeric character of the enzyme. X-ray diffraction studies of crystals resulted in a low-resolution model [6], showing only six regions of approximately equal size. This makes it difficult to unambiguously interpret the location of the seven larger individual subunits (α, β and γ). Wakabayashi et al. [7] showed by Fourier-filtering of images from two-dimensional crystals of ATPase from a thermophilic bacterium that the molecule has a pseudo six-fold and three-fold symmetry. Since, however, their crystals had long-order bends, the Fourier filtering (a tech-

Correspondence address: Fritz-Haber-Institut der Max-Planck Gesellschaft, Faradayweg 4–6, D-1000 Berlin 33, Germany.

nique to suppress noise by averaging neighbouring unit cells) was not optimal.

An arrangement of the larger subunits was proposed on the basis of electron micrographs of single molecule preparations [8,9]. In these studies, the electron images of the individual particles were interpreted visually. However, electron microscopical images are very noisy and can more reliably be interpreted after enhancement of the signal-to-noise ratio by averaging, comparable to the averaging technique for two-dimensional crystals mentioned above. Techniques to average images of single macromolecules, after an alignment procedure in which correlation functions are used extensively, are now well established (see Ref. 10). To circumvent the problem that single molecules have more orientational freedom in electron microscopical preparations than molecules arranged in two-dimensional crystals, multivariate statistical techniques were developed which allow the analysis of mixed populations of images [11–15]. The characteristic views present in the population of images are determined with an automatic classification scheme [13] and the resulting class averages represent noise-free projection images of the molecule under investigation. We have applied these new techniques to a large set of 1241 electron microscopical images of F_1-ATPase.

Materials and Methods

Sample preparation and electron microscopy

F_1-ATPase was isolated from bovine heart mitochondria according to the method of Knowles and Penefsky [16]. The protein was stored in small portions under liquid nitrogen at a concentration of 20 mg/ml in a 20 mM Tris-HCl buffer of pH 7.5, containing 1 M sucrose. Specimens for electron microscopy were prepared by the droplet method or the Valentine method [17], using uranyl acetate as a negative stain. The original protein batch was diluted 60–600-times immediately before preparation with a 10 mM Tris-HCl buffer (pH 7.5) containing 4 mM ATP and 2 mM EDTA. The grids were washed during the staining procedure for several seconds with distilled water to remove disturbing remnants of sucrose. Routine microscopy was performed on a Philips EM 400 electron microscope. Micrographs used for further analysis were taken at 80 000 magnification on the DEEKO 100 electron microscope built at the Fritz-Haber Institut [18]. The total electron dose for each micrograph was about 5500 electrons/nm^2.

Image processing

The micrographs used (see Fig. 1) were digitized with a Datacopy Model 610F electronic digitizing camera (Datacopy Corporation, Long Beach, U.S.A.) mounted under 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 0.4 nm at the specimen level. Image analysis was carried out within the framework of the IMAGIC software system [19] on a VAX 11/780 computer (Digital Equipment Corporation). Particles were selected interactively from the digitized micrographs using a raster-scan image display system. A total of 1241 molecular images were extracted from 10 different micrographs. The single molecular images were pretreated by band-pass filtering to suppress the very low and very high spatial frequencies which normally do not contain useful information [13]. The images were then surrounded by a circular mask to cut away unnecessary background; the average density inside the mask was set to zero and the variance of the densities inside the mask was normalized to an arbitrary value of 100 [13]. A gallery of F_1-ATPase images from this standardized data set is shown in Fig. 2.

The next – essential – step in the analysis is the alignment of the images with respect to each other by rotation and translation in the plane of the micrograph using correlation techniques. Alignment of the data set is a prerequisite for the subsequent (differential) average procedures. The aligned images are then submitted to an eigenvector-eigenvalue data compression procedure ('correspondence analysis' [11–14] which reduces the number of measurements per image from the original $48 \times 48 = 2304$ density values to typically between eight and twenty 'factorial coordinates' [14]. The data compression facilitates the grouping of images that are very similar. Using the variance-oriented automatic classification scheme we then partitioned the data set into between 16 and

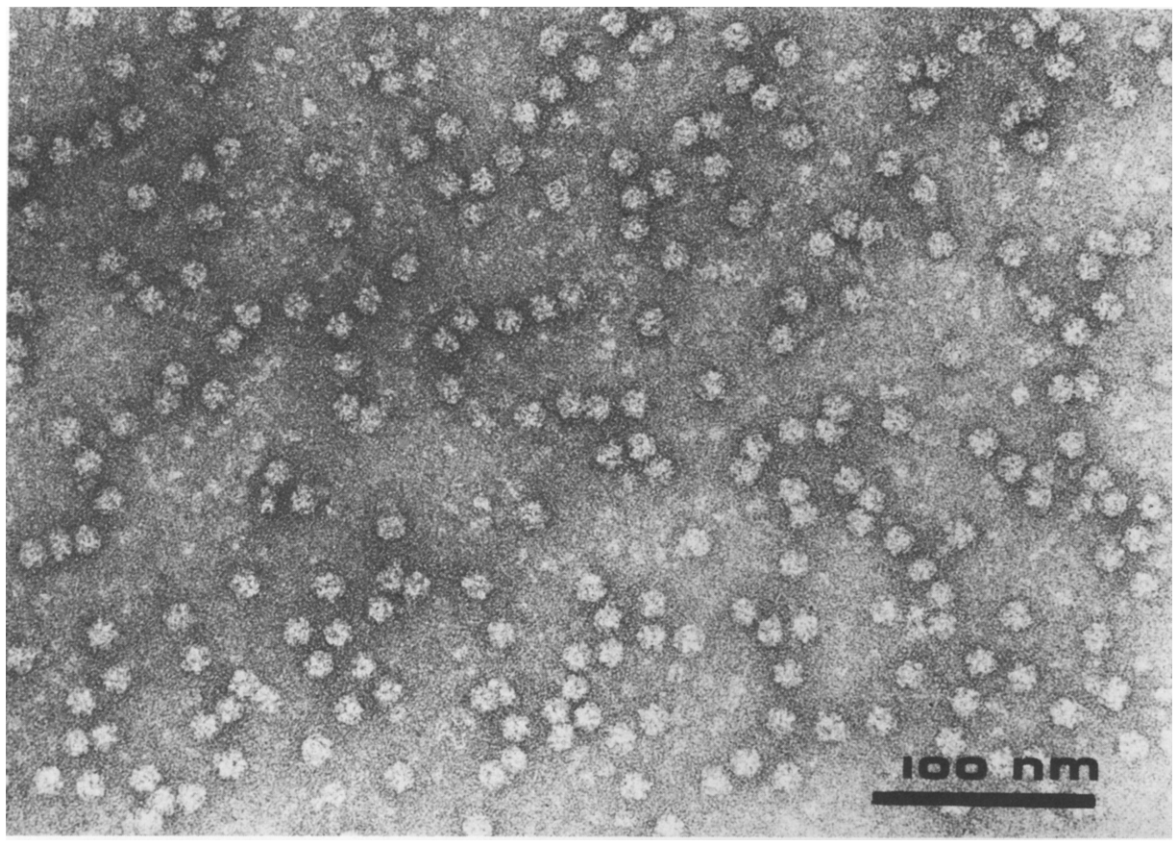

Fig. 1. Part of an original electron micrograph showing F_1-ATPase molecules used for the analysis of single particles. The specimen was negatively stained with a 1% solution of uranyl acetate. The electron optical magnification was 80000; the estimated integrated object dose used was 5500 electrons/nm^2.

60 'classes'. During the classification we normally reject 15% of the population of original images on the basis of their (too) large contribution to the internal variance of the class into which they would otherwise be placed [14]. The rejected individuals are mostly uncharacteristic images which appear rarely. The images comprising the various classes are then summed to give an average image with an enhanced signal-to-noise ratio.

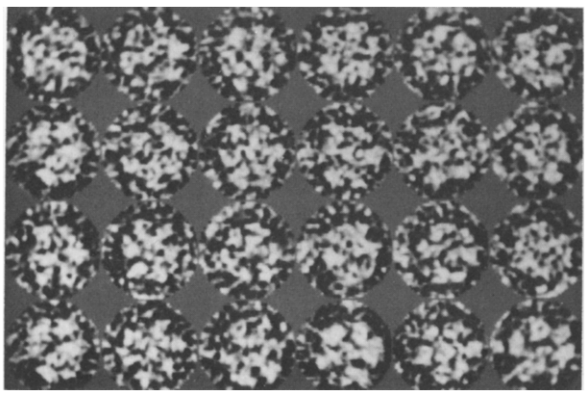

Fig. 2. A gallery of molecular projections taken from digitized images and pretreated by band-pass filtering to suppress the very low and very high spatial frequencies.

A fundamental problem associated with alignments is that we can only align an image if the image and the reference correlate well with each other ('look alike'). In a mixed population of images, such as our set of F_1-ATPase images, this implies that we have to align the data set relative to a number of different reference images in a so called multi-reference alignment procedure [14]. Whereas in the case of the much larger ribosomal subunits ($(1–2.6) \cdot 10^3$ kDa) the multi-reference alignment procedure leads to a consistent result independent of the starting point, i.e., independent of which image is used as a first reference,

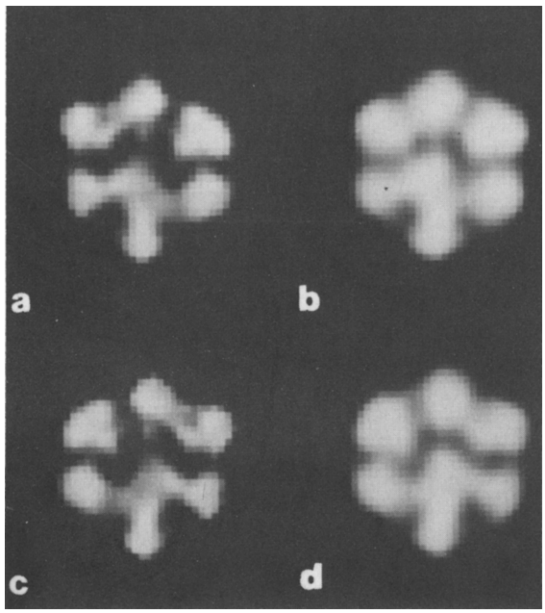

Fig. 3. The influence of the reference on the alignment of the dataset. A reference (3A) made from the summation of 9 aligned hexamers was used to align a subset of 130 images. In 3C it was mirrored and brought into a position equivalent to 3A, but the dataset itself was not mirrored. The summations (3B and 3D) of the aligned 130 images differ (see arrows), although the dataset was in both cases exactly the same.

the rather small F_1-ATPases represent a problem in that respect (Fig. 3). Depending on the starting point of the analysis, slightly different alignments of the data set are obtained and therewith somewhat different partitionings of the images into classes. We have used various starting points for the multi-reference alignments and we will only discuss the main class that is consistently present irrespective of the starting point. By restricting ourselves to the interpretation of the main 'hexagonal' class, we simultaneously restrict ourselves to those images for which the alignment is the least problematic.

Results

Various reference images used as starting point for the analysis are shown in Fig. 4. An image showing a roughly hexagonal projection (Fig. 4A), but otherwise chosen arbitrarily from the set of images, was used a one of the first references. After the first alignment relative to this noisy image, a new reference was generated by summation of the 50 images with the highest correlation coefficients relative to that first reference (Fig.

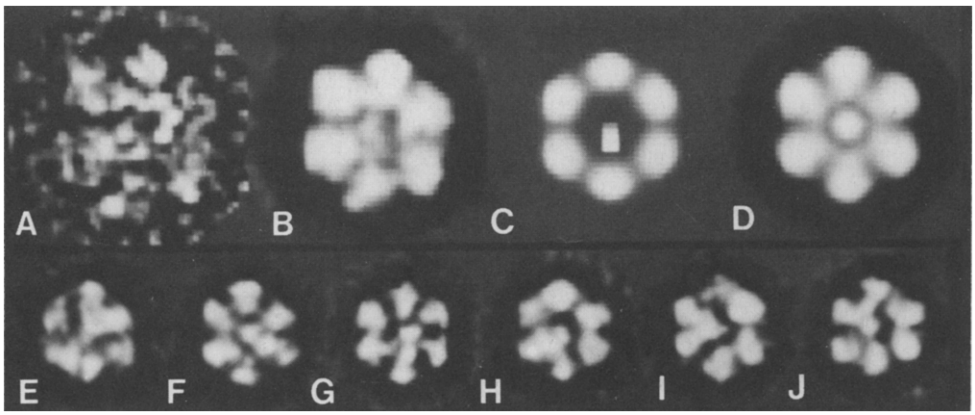

Fig. 4. Different reference images used to align the dataset. A shows an arbitrarily chosen first reference image. Those 50 images that exhibit the highest correlation with respect to the first reference in A were summed to make a second, relatively noise-free reference image shown in B. A six-fold symmetrized reference with an additional off-axis density artificially added to it is shown in C. In D a six-fold symmetrized reference image is shown that was generated by symmetrization of a good hexagonal class from an earlier classification run. When the data set was aligned relative to the perfect six-fold symmetric reference in D, the images would assume (by chance) one out of six possible orientations. We then looked for classes that were very similar, but were rotated over multiples of 60° relative to each other. In this way one can prove an asymmetry independently of the reference. E–J show six almost identical asymmetrical class averages that only differ from each in terms of orientation (by multiples of 60°C).

4B). The analysis was then continued using different class averages from the first classification runs in a multi-reference alignment scheme.

If a single reference image is used to align the data set, one may strongly impose the properties of that reference image onto the data set. For example, the total sum of a set of images aligned relative to a single reference image will normally resemble that reference image. Even if automatic classification schemes are used [14], the subset sums obtained may still merely reflect bias imposed by the alignment; even these objective classification schemes cannot 'look beyond' a given alignment. It is therefore difficult to prove that a molecule is asymmetric by only showing sums of the aligned molecules. The sum may just reflect an accidental asymmetry (noise) of the reference molecule. To circumvent this problem, different images were used as starting point for the analysis and we verified that those led to identical results. We also tried artificially generated reference images like the one shown in Fig. 4C, which was generated with exact six-fold symmetry and a seventh mass placed off the centre of the image in the direction one of the six masses.

Almost unambiguous evidence for departure from symmetry can be provided by using an oversymmetrized reference image for the alignment. When the data set is aligned relative to the perfect six-fold symmetric reference (Fig. 4D), members of an asymmetric class of images will assume (by chance) one out of six possible orientations. What we then look for in the analysis are six classes that are all very similar, but are rotated over multiples of 60° relative to each other. In this way asymmetry in a set of molecular images can be revealed, without imposing the asymmetric properties of one particular reference on the full data set. A disadvantage of the procedure is, however, that the data set size is effectively reduced by a factor of six and that the statistics are thus poorer. We therefore only used this analysis to confirm the correctness of the results of the earlier multi-reference alignment. The main view which resulted from our analysis is an asymmetric hexagonal view as is illustrated by the six best classes (Fig. 4E–J) that resulted after alignment relative to the six-fold symmetric reference.

The better classes that resulted from the non-symmetrized analysis (Fig. 5) show somewhat more detail due to the better statistics: a larger number of images are summed into each class. This asymmetric hexagonal view represents the largest homogeneous subpopulation present in the set of images (Fig. 5A–E). It accounts for about 40% of the molecular images present in our micrographs. Other classes, including molecules in a 'flattened' hexagonal projection (Fig. 5F–H), will be discussed in a future paper; they are likely to represent other positions of the molecule on the support film.

The main hexagonal view of Fig. 6 is a summa-

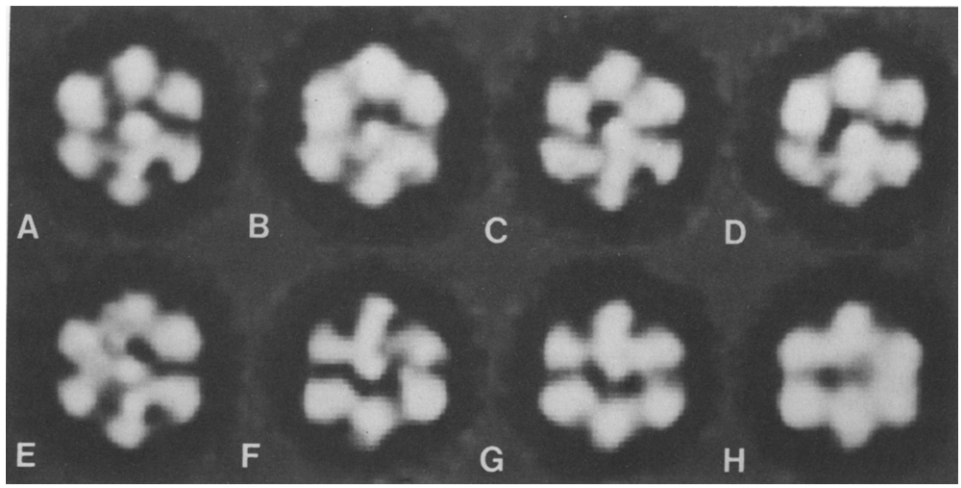

Fig. 5. Molecular projections of F_1-ATPase as determined by multireference alignments and classification of 1241 molecules; the best 8 classes out of a decomposition of the dataset into 16 classes, automatically ignoring 15% of the images.

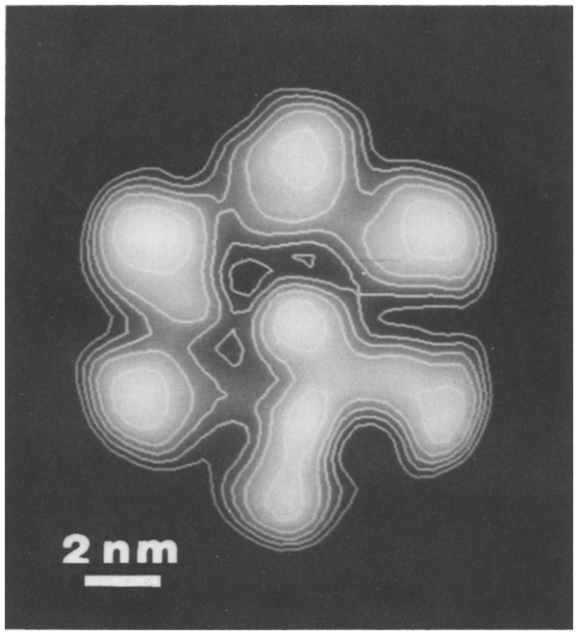

Fig. 6. The final result of the image analysis procedure: a summation of 379 projections from the five main hexagonal classes.

tion of 279 members of the five hexagonal classes of Fig. 5A–E. To determine the reproducible resolution within this view, it was split into two independent sub-averages which were compared with the Fourier ring correlation method [14]. A resolution of 2.0 nm was found. The view shows a close-to-hexagonal outer ring of six densities plus a seventh, asymmetric, density placed close to the centre of the image. This seventh density (or 'stain-excluding region') has the shape of the letter V, and the tip of the 'V' has a high density, comparable to that of the six outer densities. The two legs of the 'V' form connections with two neighbouring members of the outer ring densities. However, this view is merely a projection through the F_1-ATPase structure in a specific direction and asymmetry of a projection does not necessarily reflect asymmetry of the projected object. If, for example, a three-fold rotationally symmetric structure is projected along the three-fold axis, then the projection will also exhibit this three-fold symmetry; however, a tilt of the projection direction of say 20° relative to the rotation axis will break the symmetry at our current resolution level. The question concerning the symmetry of F_1-ATPase is thus not automatically answered by the finding of our asymmetric view.

Discussion

One molecule of F_1-ATPase contains three α and three β subunits, which together account for 86% of the mass of the protein. The α and β subunits have high homology [4], indicating that they fold in three dimensions in a very similar way. The most likely spatial arrangement for these six subunits is a D3 pointgroup arrangement if we consider the six subunits to be equivalent. Since, however, the alphas and the betas are not fully equivalent, the pointgroup symmetry of the arrangement must break down to, most likely, a C3 pointgroup symmetry with one layer of three alphas on top of a layer containing the beta subunits. Such arrangements for the main subunits were proposed on the basis of electron microscopical images by various authors [8,9]. The appearance of the electron images in these studies suggests a staggered rather than an eclipsed arrangement of the two layers of subunits. The C3 arrangement was recently demonstrated directly using antibody labeling techniques [21,22]: angles of 120° between the attachment sites of anti-α antibodies were found consistently in the immuno-electron microscopical experiments.

Our findings are consistent with these C3 models (close to D3 models), in which the large α and β subunits alternate on the outside of the structure. Assuming this C3 arrangement for the large subunits to be correct, we can measure the distances between the members of the six outer densities that are related by the approximate three-fold rotation axis. The distances found in various independent classes are all within around 5% of each other and the distance variations between the five independent classes (Fig. 5A–E) are unsystematic. This indicates that the main view is indeed a view that is within 10° of the three-fold rotation axis. The asymmetry observed in the main view, at the current resolution level, is thus likely to be due to an asymmetry in the F_1-ATPase molecule itself and not to a tilting of the projection direction relative to the approximate three-fold axis of the F_1-ATPase.

Since γ and the other small subunits appear as

single copies in each ATPase molecule, these subunits are likely to be found in asymmetric components of the structure: the central pit and the legs of the V shape. The V shape suggests that most of the interactions of γ and the other single-copy subunits are with two neighbouring α-β subunits. This result directly supports the biochemical evidence of Williams et al. [23] that the γ subunit binds to a single α-β pair.

Upon projection along the three-fold axis a D3 structure yields a three-fold rotationally symmetric view with three mirror axes at angles of 60° from each other. The mirror axes are a consequence of projecting in a direction perpendicular to the two-fold axes. In our case the two-fold symmetry elements are only approximate due to the differences between the α and the β subunits. The connections found between the central subunit and two neighbouring subunits break the three-fold symmetry of the most common 'hexagonal' projection seen in electron micrographs. One of the axes of approximate mirror symmetry remains visible in spite of the presence of the seventh subunit, namely, the axis running through the α-β pair to which the central subunit is thought to be connected. However, even at the current resolution level, this 'remaining' mirror symmetry is significantly not exact (Figs. 5 and 6). This may be due to a genuine structural departure from three-fold symmetry of the F_1-ATPase molecule, to preferential staining effects, or to both.

Interestingly, in a recent X-ray crystallographic study [6,24] only half a molecule was found in the asymmetric unit, related to the other half by a two-fold rotation axis, which could correspond to the 'remaining' two-fold axis discussed above. In their work, the authors discuss the arrangement we propose as one of the possibilities, with however, a different assignment for the α and β subunits. No clear density close to the centre of the protein was found; but since the smaller subunits do not conform to the crystal symmetry they tend to have low X-ray diffraction electron density and could go undetected [24].

The hexagonal view in Fig. 6 represents the largest homogeneous subpopulation present in the set of images. In earlier electron microscopical studies, the 'hexagonal' views were selected purely visually [8,9] or using additional symmetry criteria [20] prior to any analysis. Such procedures may unintentionally bias the data set. In our study the main class is rather the result of the analysis and is determined by objective classification techniques. Moreover, the number of molecular images analysed in our study is by an order of magnitude larger than that used in earlier studies; the statistical significance of our results is correspondingly higher.

In a study by Tsuprun et al. [9] the molecules showed no additional protein density in the centre of the structure. This could be due to the visual selection procedure. In our data set we also found some molecules with hexagonal appearance but without the central density; however, this rare class lacked statistical significance. Absence of the central density could also be due to loss of one or more of the smaller subunits during isolation or preparation. Our results agree better with those of Tiedge et al. [8]. They have placed a smaller protein mass on the inside of the structure. The inner protein density has connections with three subunits on the outside, a symmetrical arrangement which was not confirmed in our work. In a study by Akey et al. [20], only some 30–50 molecules were analysed. The particle selection, the centering of the particles using the $n = 6$ harmonic component and other details of the analysis represent a bias toward the symmetric – in particular the six-fold symmetric – image components. The possibility of asymmetry of the projection was not investigated; the averages shown in their paper miss the important $n = 1$ harmonic component and therefore cannot be compared to our images.

It is difficult to bring our result of an asymmetric projection into harmony with the results of Gresser et al. [25] in which three equivalent catalytic binding sites on the three β subunits were proposed. On the other hand, it was found that that the three β subunits are not equivalent in the enzyme [26,27]. In the *E. coli* enzyme two β subunits can bind to DCCD. The third β subunit is different: it does not bind (easily) to DCCD, but it can bind EDC and crosslink to the ε subunit [26]. For the bovine heart enzyme it was found that only four out of the six nucleotide binding sites are exchangeable and can be labelled with 8-azido-AT(D)P [27]. Our results, however, cannot absolutely prove that the observed asymmetry in

the F_1-ATPase reflects a difference in the catalytic behaviour of the three α-β pairs.

Our results strongly suggest the presence of a seventh structural subunit in F_1-ATPase. This central subunit, which is likely to contain at least the γ subunit, is connected to a pair of neighbouring α-β subunits. The connections found between the central subunit and two neighbouring subunits break the approximate three-fold symmetry of the most common 'hexagonal' projection seen in electron micrographs. Since we have concentrated on this single view, it is not yet clear whether or not the seventh subunit protrudes outside (stalk?) the boundaries of the approximate D3 pointgroup structure formed by the three α and three β subunits. An extended study of an even larger data set, to render statistical significance to the rare non-hexagonal views, combined with three-dimensional reconstruction will be necessary to answer such questions.

Acknowledgments

We thank Mr. K. Weiss and Mr. H. Hollmann for expert technical assistance, and Dr. George Harauz for critical reading of the manuscript and for his contributions to the programming effort. We furthermore thank Professor Elmar Zeitler for continuous support of our work. Computer equipment grants from the Max Planck Society are acknowledged. This work was supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 312).

References

1 Cross, R.L. (1981) Annu. Rev. Biochem. 50, 681–714
2 Amzel, L.M. and Pedersen, P.L. (1983) Annu. Rev. Biochem. 52, 801–824
3 Senior, A.E. and Wise, J.G. (1983) J. Membrane Biol. 73, 105–124
4 Walker, J.E., Fearnley, I.M., Gay, N.J., Gibson, B.W., Northrop, F.D., Powell, S.J., Runswick, M.J., Saraste, M. and Tybulewicz, V.L.J. (1985) J. Mol. Biol. 184, 677–701
5 Todd, R.D., Griesenbeck, T.A. and Douglas, M.G. (1980) J. Biol. Chem. 255, 5461–5467
6 Amzel, L.M., Mckinney, M., Narayanan, P. and Pedersen, P.L. (1982) Proc. Natl. Acad. Sci. USA 79, 5852–5856
7 Wakabayashi, T., Kubota, M., Yoshida, M. and Kagawa, Y. (1977) J. Mol. Biol. 117, 515–519
8 Tiedge, H., Schäfer, G. and Mayer, F. (1983) Eur. J. Biochem. 132, 37–45
9 Tsuprun, V.L., Mesyanzhinova, I.V., Kozlov, I.A. and Orlova, E.V. (1984) FEBS Lett. 167, 285–290
10 Frank, J., Verschoor, A. and Boublik, M. (1981) Science 214, 1353–1355
11 Van Heel, M. and Frank, J. (1981) Ultramicroscopy 6, 187–194
12 Frank, J., Verschoor, A. and Boublik, M. (1982) J. Mol. Biol. 161, 107–137
13 Van Heel, M. (1984) Ultramicroscopy 13, 165–184
14 Van Heel, M. and Stöffler-Meilicke, M. (1985) EMBO J. 4, 2389–2395
15 Bijlholt, M.M.C., Van Heel, M.G. and Van Bruggen, E.F.J. (1982) J. Mol. Biol. 161, 139–153
16 Knowles, A.F. and Penefsky, H.S. (1972) J. Biol. Chem. 245, 1336–1344
17 Valentine, R.C., Shapiro, B.M. and Stadmann, E.R. (1968) Biochemistry 7, 2143–2152
18 Kunath, W., Weiss, K., Sack-Kongehl, H., Kessel, M. and Zeitler, E. (1984) Ultramicroscopy 13, 241–252
19 Van Heel, M.G. and Keegstra, W. (1981) Ultramicroscopy 7, 113–130
20 Akey, C.W., Crepeau, R.H., Dunn, S.D., McCarthy, R.E. and Edelstein, S.J. (1983) EMBO J. 2, 1409–1415
21 Tiedge, H., Lünsdorf, H., Schäfer, G. and Schairer, H.U. (1985) Proc. Natl. Acad. Sci. USA 82, 7874–7878
22 Lünsdorf, H., Ehrig, K., Friedl, P. and Schairer, H.U. (1985) J. Mol. Biol. 173, 131–136
23 Williams, N., Hullihen, J.M. and Pedersen, P.L. (1984) Biochemistry 23, 780–785
24 Pedersen, P.L. and Amzel, L.M. (1985) in Achievements and Perspectives of Mitochondrial Research (Quagliariello, E., ed.), Elsevier, Amsterdam
25 Gresser, M.J., Myers, J.A. and Boyer, P.D. (1982) J. Biol. Chem. 257, 12030–12038
26 Lötscher, H.R. and Capaldi, R.A. (1984) Biochem. Biophys. Res. Commun. 121, 331–339
27 Van Dongen, M.B.M. and Berden, J.A. (1986) Biochim. Biophys. Acta 850, 121–130