Journal of Bioenergetics and Biomembranes, Vol. 35, No. 4, August 2003 (© 2003)

Structure and Function of the Vacuolar H^+-ATPase: Moving From Low-Resolution Models to High-Resolution Structures

Michael Harrison,[1,3] Lyndsey Durose,[1] Chun Feng Song,[2] Elizabeth Barratt,[1] John Trinick,[2] Richard Jones,[1] and John B. C. Findlay[1]

In the absence of a high-resolution structure for the vacuolar H^+-ATPase, a number of approaches can yield valuable information about structure/function relationships in the enzyme. Electron microscopy can provide not only a representation of the overall architecture of the complex, but also a low-resolution map onto which structures solved for individually expressed subunits can be fitted. Here we review the possibilities for electron microscopy of the *Saccharomyces* V-ATPase and examine the suitability of V-ATPase subunits for expression in high yield prokaryotic systems, a key step towards high-resolution structural studies. We also review the role of experimentally-derived structural models in understanding structure/function relationships in the V-ATPase, with particular reference to the complex of proton-translocating 16 kDa proteolipids in the membrane domain of the V-ATPase. This model in turn makes testable predictions about the sites of binding of bafilomycins and the functional interactions between the proteolipid and the single-copy membrane subunit Vph1p, with implications for the constitution of the proton translocation pathway.

KEY WORDS: Vacuolar membrane; V-ATPase; expression; circular dichroism; electron microscopy; modelling.

INTRODUCTION

Both the vacuolar H^+-ATPase and F_1F_0-ATPase belong to a class of bidomain ATPases, with a soluble catalytic domain coupled physically and mechanistically to a proton-translocating membrane domain (Nishi and Forgac, 2002; Stevens and Forgac, 1997). The evolutionary relationship between the two enzymes extends to clear similarity between some component subunits at the level of primary structure (Bowman *et al.*, 1988). However, unlike the F-ATPase, relatively few high-resolution structural data are available for the V-ATPase. It does seem increasingly clear that the vacuolar H^+-ATPase shares the same basic molecular architecture as the F_1F_0-ATPase, and the two enzymes operate via fundamentally similar rotational mechanisms (Imamura *et al.*, 2003; Yasuda *et al.*, 1998). To some extent therefore the F-ATPase represents a reasonable template on which to model core components of the V-ATPase molecular motor. Modelling, as part of an iterative process supported by rigorous experimental testing, can provide credible representations of structures which can explain many functional or mechanistic properties of the protein. However, this approach may be too simplistic and low resolution, and certainly cannot offer an explanation at the structural level for many of the features of the V-ATPase which are unique. For example, high-resolution structures would contribute hugely to our understanding of the subtle differences in functionality conferred by the presence of different subunits or isoforms. In addition, structural data would provide an insight into any dynamic changes that occur as a consequence of the V-ATPase interfacing with physiological control systems or interacting with cytoskeletal elements (Holliday *et al.*, 1999; Kane, 1995).

[1] School of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom.
[2] School of Biomedical Sciences, University of Leeds, Leeds, United Kingdom.
[3] To whom correspondence should be addressed; e-mail: m.a.harrison@leeds.ac.uk.

0145-479X/03/0800-0337/0

Of course, the problems associated with generating high-resolution structures of multi-subunit membrane protein complexes are far from trivial, although not insurmountable. In advance of a complete high-resolution description of the V-ATPase, a number of approaches either have the potential (or in some cases have already proved) to be valuable in understanding structure/function relationships. Examination of the folds of individual subunit species, following structural analysis of recombinantly expressed single polypeptides, can still reveal clues about the mechanism of protein–protein interactions even though the polypeptides are studied in isolation. As a key example, the identification of a β-adaptin-like fold in the crystallographic structure of subunit *H* (Sagermann *et al.*, 2001) (the *VMA13* gene product in yeast) has revealed a mechanism linking the V-ATPase to the endocytic machinery (Geyer *et al.*, 2002a,b). Electron microscopy, while inherently low-resolution, can also provide a detailed map of subunit locations within a complex and has already been applied with success to a number of V-ATPase species (Domgall *et al.*, 2002; Ubbink-Kok *et al.*, 2000; Wilkens *et al.*, 1999; Wilkens and Forgac, 2001). These studies have contributed significantly to structural and mechanistic models for the V-ATPase rotary motor. In the absence of real structures, molecular modelling can provide useful working structures which can in turn be used as test-beds for experimental design.

The aim of this review is provide a synopsis of our work that has focused on modelling components of the V-ATPase, in particular the membrane domain and its assemblage of proteolipid polypeptides. It will extend to examining the possibilities for electron microscopic analysis of the enzyme, and will review work to date on expression, purification, and structural analysis of individual subunits.

ARCHITECTURE OF THE *SACCHAROMYCES* V-ATPase

Advances in electron microscopy instrumentation and image analysis now make it feasible to obtain images of multiprotein complexes in which individual structural elements are resolved. In fact, both soluble (Klaholz *et al.*, 2003) and membrane (Liu *et al.*, 2002; Orlova *et al.*, 2003) proteins have been imaged by single particle averaging at resolutions in the 14–18 Å range. Using cryo-EM resolution to visualize individual secondary structure elements is possible, and rapid freezing methods have been designed which can trap individual conformational states of proteins (Walker *et al*., 1999). These methods effectively allow mapping of the sequence of dynamic transitions that occur during the catalytic cycle of the complex. In advance of high-resolution structural data, the V-ATPase represents an excellent target for this type of approach.

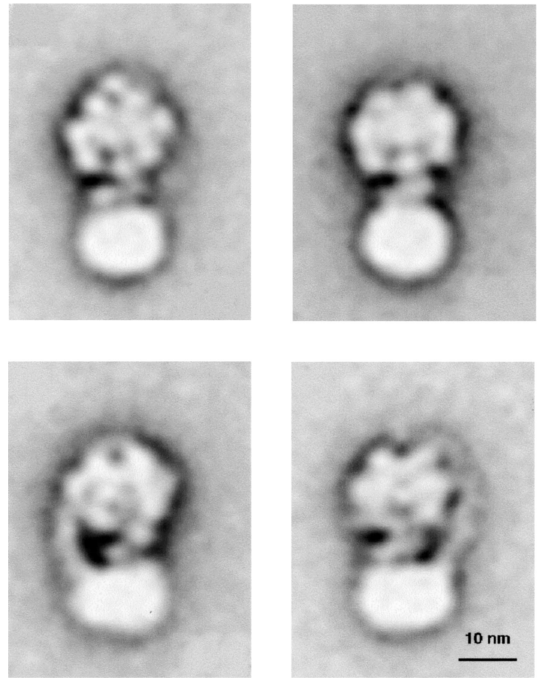

Fig. 1. Electron microscopy and single-particle averaging of the *Saccharomyces* V-ATPase. V-ATPase was prepared from yeast cells expressing a His6-tagged form of Vma3p by solubilization of vacuolar membranes with dodecyl maltoside followed by centrifugation on glycerol gradients, as described in Harrison *et al.* (2000). The panels show image averages of the negatively-stained V-ATPase molecules after classification into homogeneous groups. The image averages show two domains, V_1 and V_0 (lower). Central and peripheral stalks connect V_1 and V_0. The peripheral stalk is most evident in the bottom panels. 5535 molecules were brought into alignment and classified into 60 groups using SPIDER software. The individual averages contain about 100 images.

We have imaged the whole V-ATPase from *Saccharomyces* using electron microscopy with negatively-stained specimens (Fig. 1). Four classes of image averages show a complex with the classical bi-domain structure, approximately 28 nm in the long axis (corresponding to a plane perpendicular to that of the membrane). The soluble domain, with a maximum diameter of 15 nm, is linked to the membrane domain via central and peripheral stalks of 5 and 20 nm, respectively. These structures correspond to those previously observed for both primitive bacterial (Boekema *et al.*, 1997) and higher plant (Domgall *et al.*, 2002; Wilkens *et al.*, 1999) enzymes. The soluble domain appears asymmetrical, with additional densities likely to represent single-copy soluble subunits of (as yet) unknown identities. The membrane domain itself is resolved as a structure with maximum approximate dimensions of 14 nm in a plane parallel to the membrane

by 10 nm (plane perpendicular to the membrane). The peripheral stalk links directly to the membrane domain.

Unambiguous assignment of the locations of subunits remains problematical given the level of resolution. However, it seems clear that the central stalk of the complex comprises Vma8p (subunit *D*) and Vma7p (subunit *F*), since these subunits in combination with the Vma1p/Vma2p hexamer comprise the core V1 complex (Imamura *et al.*, 2003; Tomashek *et al.*, 1997). Additional peripheral structures may comprise Vma13p (subunit *H*), Vma4p (subunit *E*), the soluble domain of Vph1p, or Vma10p (Domgall *et al.*, 2002; Landolt-Marticorena *et al.*, 2000). Vma10p shows partial sequence similarity to subunit *b* of the F-ATPase (Hunt and Bowman, 1997), and by analogy is likely therefore to contribute in some way to the peripheral stalk of the complex, the "stator" element of the V-ATPase rotary motor (Domgall *et al.*, 2002; Ubbink-Kok *et al.*, 2000). However, it is interesting to note that even as a single-extended α-helix, Vma10p would not be long enough to form the peripheral "stator" stalk of the V-ATPase. This structure is likely to also contain Vma4p (subunit *E*), which is reported to make contacts with Vma2p, Vma10p, and the membrane domain (see Domgall *et al.*, 2002, and references therein). The position of Vma13p has been fitted onto a 3D reconstruction of a plant V-ATPase generated by electron microscopy (Domgall *et al.*, 2002), and is proposed to form a peripheral link between the V_1 and membrane domains (Domgall *et al.*, 2002; Landolt-Marticorena *et al.*, 2000). Consistent with this observation, we have previously observed interaction between the 16 kDa proteolipid of the membrane domain and a polypeptide of approximately 54 kDa in isolates of the yeast V-ATPase, after reaction with photoactivatable cross-linkers (Powell, 1999). Its should be stressed that the positions of many of the single-copy subunits of the V-ATPase remain uncertain.

The inherent capability of the yeast V-ATPase to physically uncouple the soluble and membrane domains in response to certain physiological conditions (Doherty and Kane, 1993) may mean that this species of enzyme is fundamentally unstable, limiting its suitability for higher resolution electron microscopy. However, the ease with which the yeast system can be manipulated genetically, allowing expression and incorporation of chimeric or immunologically-tagged subunits, does offer a major advantage: differential imaging of complexes even at relatively low resolution in negative stain can be used to locate the positions of specific subunits within the ATPase. A number of membrane and soluble subunits tolerate addition of immunological tags (Hirata *et al.*, 1997; Nelson *et al.*, 1994), and Vph1p, Stv1p, and Vma2p can be incorporated into assembled V-ATPases as N-terminal GFP fusion proteins (S. Ball, T. Outiero, M. Smithson and M. Harrison, unpublished data). Inclusion of tagged or fused proteins, in combination with labelling by antibody fragments, may offer opportunities for definitive localization of subunits within the V-ATPase complex.

MODELLING COMPONENTS OF THE MEMBRANE DOMAIN

The yeast V_0 domain basically comprises six copies of proteolipid, one copy of the soluble but tightly-bound Vma6p subunit and one copy of the 100 kDa membrane protein Vph1p. The complement of proteolipids includes copies of Vma3p, Vma11p, and Vma16p in an uncertain (perhaps variable) stoichiometry (Hirata *et al.*, 1997). We have focused primarily on modelling the structure of this proteolipid assembly, founded on a few basic premises: the proteolipid assembles as a hexamer (Holzenburg *et al.*, 1993), and each proteolipid monomer is a four-helical bundle (Holzenburg *et al.*, 1993). The monomer constitutes a tandem repeat of the two-helix subunit *c* of F-ATPase (Mandel *et al.*, 1988), for which there is an NMR-derived solution structure (Girvin *et al.*, 1998). Construction of the model has relied on converging a number of experimental approaches that measure lipid accessibility, or directly examine protein–protein contacts within the complex.

Assessing the reactivity of mutagenically introduced thiol side chains to differentially soluble reagents can provide a reasonably accurate map of aqueous phase or lipid phase accessibility. Modification of sites on the 16 kDa proteolipid by water-soluble fluorescein maleimide, conveniently assayed from mass shift of the solvent-extracted polypeptide on SDS-PAGE, showed cysteines on one face of Helix 1 to be freely available to the probe. In contrast, the reagent did not react with any other transmembrane sites (Jones *et al.*, 1995). Differential reactivity towards hydrophobic pyrenyl maleimide of cysteines introduced into Helices 2, 3, and 4 identified discrete lipid-accessible sites (Harrison *et al.*, 1999). This approach was extended further to include labelling of both the native glutamate residue of Helix 4, and several glutamate mutant proteolipids: Modification of specific sites by the fluorescent DCCD analogue pyrenylcyclohexyl carbodiimide confirmed that particular defined faces of Helices 2, 3, and 4 (but not 1) were oriented towards the lipid phase (Harrison *et al.*, 2000). Modification of carboxyl-containing residues by lipid soluble carbodiimides is itself an indirect indicator of lipid phase exposure, since these compounds react only with protonated side chains and form a stable end product only if water is completely excluded.

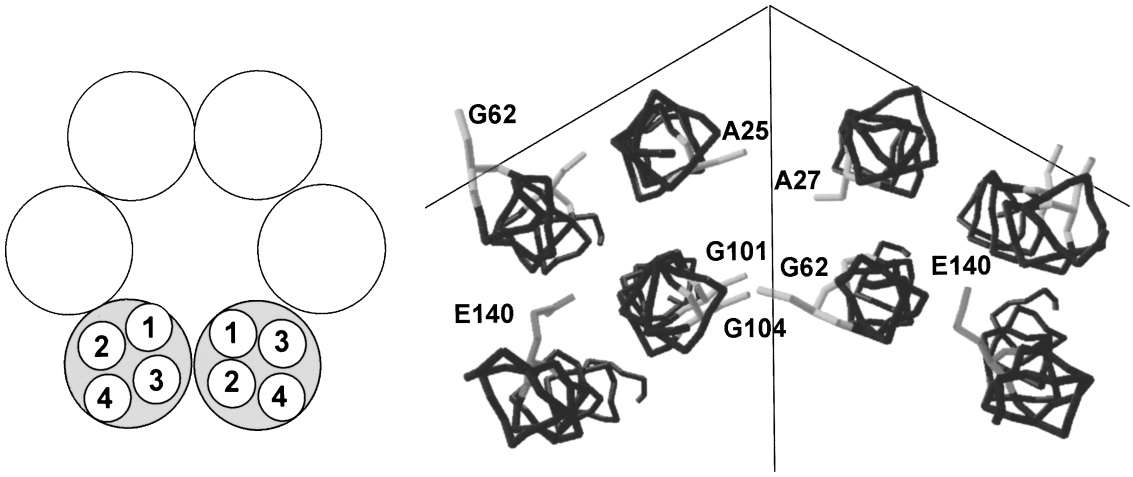

Fig. 2. Experimentally derived model of the 16 kDa proteolipid. The 16 kDa proteolipid is proposed to assemble as a hexamer of four helical bundles surrounding a central pore (left). A dimer of 16 kDa proteolipids was modelled (right) on the NMR-derived structure of the related 8 kDa subunit *c* proteolipid from F-ATPase (Girvin *et al.*, 1998) using the programme Sybyl 6.7 (Tripos Associates). Helices were positioned in accordance with experimentally observed disulphide crosslinks and lipid accessibility measurements. The derived position of the DCCD-reactive glutamate (E140) is consistent with accessibility to the lipid phase. Sites which can be mutated to confer bafilomycin resistance would be positioned on Helix 4 on the face opposite Glu140. This region also contains sites predicted to interact with the region of Vph1p containing Arg735.

Site-directed cross-linking is a powerful method that can provide information about helical packing arrangements in transmembrane proteins (Falke and Koshland, 1987; Jiang and Fillingame, 1998; Pakula and Simon, 1992; Sedgwick *et al.*, 1997). Cysteine residues are introduced mutagenically, and disulphide cross-links indicative of thiol side chain proximity are induced either by incubation under oxidizing conditions or by introduction of homobifunctional crosslinking reagents. We have applied this approach to the 16 kDa proteolipid, constructing an extensive panel of mutant polypeptides containing pairs of cysteine residues (Harrison *et al.*, 1999). A number of intermolecular cross-links resulted in the formation of covalently linked proteolipid dimers. These included cysteines at positions on Helix 1 predicted to be at the lipid/aqueous phase interface (Jones *et al.*, 1995), and at positions on opposite faces of Helix 1 (Fig. 2, residues Ala25 and Ala27) (Harrison *et al.*, 1999). Taken in combination with the observations from the fluorescein labelling studies, these data consistently indicate that Helix 1 from each 16 kDa proteolipid subunit is situated at the center of the V_0 complex, and that these helices interact with one another to form the lining of a large pore which is at least partially accessible to water. Introduction of paired cysteine residues into Helix 2 and Helix 3 of the proteolipid (Fig. 2: Gly101/104 and Gly62) resulted in very rapid and almost irreversible formation of cross-links between adjacent polypeptides (Harrison *et al.*, 1999). These helices are therefore proposed to be adjacent to each other in the proteolipid complex. This proximity is accommodated within the model, in which Helix 1–Helix 1 and Helix 2–Helix 3 junctions form the principal intermolecular contacts within the proteolipid hexamer (Fig. 2). A similar site-directed cross-linking strategy in F_0 resulted in the formation of oligomeric species as large as dodecamers (Jones and Fillingame, 1998). However, in the case of the 16 kDa proteolipid, dimers predominated and no higher order oligomers were observed. There may therefore be some nonequivalence in subunit packing within the V_0 proteolipid complex, leading to pseudo-threefold symmetry (in the form of a "trimer of dimers") rather than simple sixfold symmetry.

Fluorimetric methods can be used not only to assess the general disposition of the individual residues with respect to the lipid phase, but can also provide a quantitative measure of the distance of the fluorescently labelled site from the bilayer center. Using this strategy, it is possible to produce not only broad constraints for helical orientation, but also more precise definitions for the depths of individual helices in the membrane. In the case of the 16 kDa proteolipid, we have used an approach in which the Helix 4 glutamate and a glutamate substitution at position 107 in Helix 3 are modified by pyrenylcyclohexyl carbodiimide. Affinity purification and subsequent reconstitution of pyrene-labelled V_0 complexes show accessibility of the covalently attached fluorophore exclusively to lipid-soluble quenchers (Harrison *et al.*, 2000). This study was extended to the

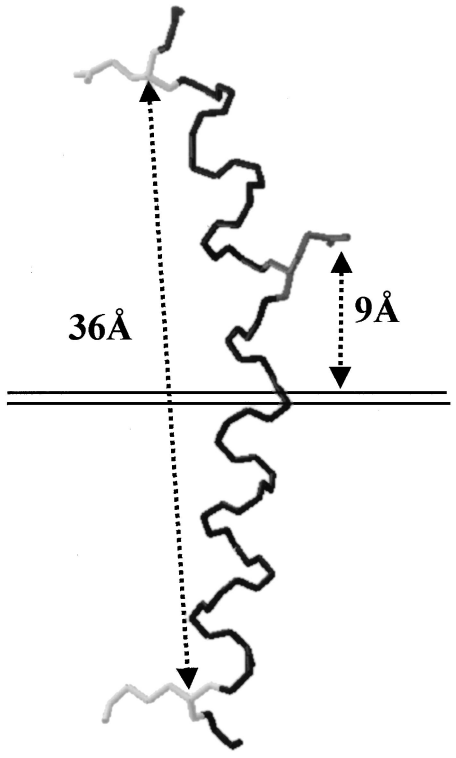

Fig. 3. The depth in the membrane of the DCCD-reactive glutamate of the 16 kDa proteolipid. Quenching of fluorescence of a pyrene adduct bound to Glu140 by positionally spin-labelled lipids indicates that the fluorophore is (on average) 9 Å from the center of the bilayer. Examination of the sequence of Helix 4 of the 16 kDa proteolipid suggests that it is bounded at its extremites by an arginine residue at the N-terminal end (top of the figure) and by a lysine at the C-terminus. This would place Glu140 in the cytoplasmic leaflet of the bilayer.

measurement of fluorescence quenching after reconstitution in the presence of lipids carrying quencher species at defined sites in the acyl chain. Data analysis using the parallax method (Chattopadhyay and London, 1987) permits the distance of the pyrene fluorophore from the center of the bilayer to be calculated. In the case of both Helix 4 (Glu140) and Helix 3 (Ala107 → Glu) sites, this distance was calculated to be approximately 9 Å (Harrison *et al.*, 2000) (Fig. 3). Very similar distances have been calculated for reconstituted V_0 containing PCD-modified Vma3p (M. Harrison, unpublished data). Measurement by ESR methods of spin–spin interactions between the paramagnetic DCCD analogue *N*-(2,2,6,6-tetramethyl piperidineoxy)-*N′*-cyclohexylcarbodiimide attached to Glu140 of the *Nephrops* protein and spin-labelled stearic acids also indicated that the glutamic acid was about 10 Å from the hydrophobic/hydrophilic boundary at the bilayer surface (Pali *et al.*, 1999). These data provide spatial constraints for the model in a plane perpendicular to the membrane plane, and point to a lipid-accessible protonated residue located approximately midway into the cytoplasmic half of the lipid bilayer (Fig. 3) and exposed at the Helix 2–Helix 4 interface (Fig. 2).

The structural model illustrated in Fig. 2 therefore represents a synthesis of mutually supporting data obtained via independent experimental approaches. The orientation of each helix, constrained in accord with the observed disulphide cross-links, is fully consistent with the pattern of lipid accessibility obtained from the labelling experiments. On the basis of common architecture and evolutionary relationship, one could propose an alternative model for the 16 kDa assembly that matches more closely the model of F_0 (PDB entry 1C17). In such an alternative model, the tandemly repeated two-helix elements of the 16 kDa proteolipid would occupy the same positions as the helical hairpins of F_0 subunit *c*, resulting in concentric rings of helices. An inner ring would comprise alternating Helices 1 and 3, the outer ring alternating Helices 2 and 4. However, this model would not accommodate the observed intermolecular cross-links, nor would it allow lipid accessibility to sites on Helix 3.

A question remains however: Is a model of the 16 kDa proteolipid complex which is based on data collected for a heterologously expressed arthropod protein relevant to the yeast V-ATPase? Clearly, the model does not accommodate the existence of Vma11p and Vma16p proteolipid isoforms present in the *Saccharomyces* V_0 complex (Hirata *et al.*, 1997), nor is the sequence of the *Nephrops* proteolipid identical to that of any of the native yeast polypeptides. However, a number of arguments can be presented in mitigation. Firstly, all the eukaryotic proteolipids are very highly conserved, and the *Nephrops* protein is functionally interchangeable with Vma3p in the yeast system (Harrison *et al.*, 1994; Holzenburg *et al.*, 1993). Secondly, the *Nephrops* protein interacts with both Vma11p (T. Outiero and M. Harrison, unpublished data) and Vma16p *in vivo* (Gibson *et al.*, 2002). It is reasonable therefore to propose the model, at least at the level of the dimeric proteolipid unit, is a good representation of the basic 16 kDa proteolipid structural framework.

A number of interesting features are evident from the model of the 16 kDa proteolipid. Firstly, the position of the crucial glutamate residue makes it available for direct interaction with other elements of the proton translocation pathway presumed to be carried by Vph1p (Grabe *et al.*, 2000; Kawasaki-Nishi *et al.*, 2001; Leng *et al.*, 1999). Secondly, sites on Helix 4 which, when mutated, confer resistance to bafilomycin (Bowman and Bowman, 2002), are also located at lipid-exposed sites at the external surface of the complex (see Fig. 2). Resistance infers a direct interaction between these sites and

the lipid-soluble macrolide antibiotic, although indirect structural effects cannot be discounted. Another interesting structural feature is found on Helix 3: the region which appears from cross-linking studies to be involved in intermolecular contacts contains a helical face in which glycine residues are particularly abundant (Fig. 2). This region is also acutely sensitive to mutagenic changes, resulting in a profound loss of functionality (Harrison *et al.*, 1999; Noumi *et al.*, 1991). A series of six glycine residues, which are highly conserved within 16 kDa proteolipids, occur with broadly helical periodicity along the face of Helix 3 which is proposed to form the interface with Helix 2 of the adjacent polypeptide. Recent work examining structural features that drive dimerization of integral membrane proteins has identified the motif GxxxG as a key factor (Ubarretxena-Belandia and Engelman, 2001). Helix 3 of the 16 kDa proteolipid contains a multiple form of this motif (GxxxGxxGxxxGxxxGxxG) (see Fig. 4).

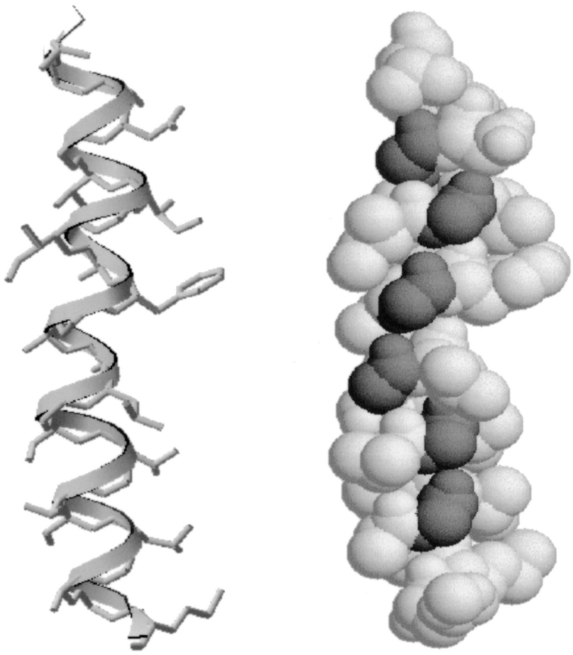

Fig. 4. Helix 3 of the 16 kDa proteolipid contains a glycine repeat motif involved in helix–helix interactions. The Helix 3 region of the 16 kDa proteolipid, corresponding to the second subunit *c*-derived tandem repeat element, contains a glycine repeat motif with α-helical periodicity. The modelled helix is shown in ribbon representation with side chains shown as stick representations, N-terminal residue lowermost (left). The glycine-enriched face is towards the reader. The glycine repeat motif gives rise to a groove on Helix 3 (shown in space-filling representation, right), presumed to allow close intermolecular packing and weak hydrogen bonding between the glycine Cα proton and oxygen of the neighboring chain. The helical hairpin was modelled using Sybyl 6.7 (Tripos Associates) and visualized using Rasmol.

The presence of this motif is proposed to allow multiple weak hydrogen bonds between the Cα protons of the glycine residues and oxygens of the neighboring helix (Senes *et al.*, 2001). The precise role of this structural feature within the 16 kDa proteolipid remains uncertain, but it does seem to represent an example of a more universally abundant dimerization motif and is likely to be involved in facilitating or maintaining highly stable helical interactions within the V_0 complex. The localization of this feature to a region of protein–protein contact by independent experimental approaches in itself provides indirect support for the credibility of the structural model.

The single-copy membrane subunit Vph1p also contributes to the translocation of protons (Kawasaki-Nishi *et al.*, 2001; Leng *et al.*, 1996), and by analogy with the F-ATPase therefore represents the functional homologue of subunit *a*. Both contain essential arginines in transmembrane positions (Arg735 in Vph1p, Arg210 in subunit *a*) (Kawasaki-Nishi *et al.*, 2001; Vik and Antonio, 1994). The application of predictive methods to the question of Vph1p topology has not provided a definitive answer. However, it is likely that Vph1p contains an even number of helices, since both termini are on the cytoplasmic side of the membrane. The soluble N-terminal domain makes contact with V_1 components (Landolt-Marticorena *et al.*, 2000), and the C-terminus may interact with physiological control factors found in the cytoplasm (Ya *et al.*, in press). Eight transmembrane helices is therefore a reasonable estimate. What little sequence similarity there is between Vph1p and subunit *a* is localized to *a* helical region containing the essential Arg735, with the geometry of residues implicated in proton movement proposed to be similar to that in subunit *a* (Kawasaki-Nishi *et al.*, 2001). Molecular modelling of the helical region containing Arg735 and the corresponding section of subunit *a* suggest that the two regions share very similar electrostatics, H-bonding potential and lipophilicity (P. Meek and M. Harrison, unpublished data). It is possible therefore that at a very local level, Vph1p is also structurally similar to subunit *a*. Site-directed cross-linking of the helix of subunit *a* that contains Arg210 makes contact with the subunit-c proteolipid (Jiang and Fillingame, 1998). By extrapolation, one can make predictions about the interaction between Vph1p and the 16 kDa proteolipid. According to the model in Fig. 2, these sites would be on the face of Helix 4 opposite Glu140. These predictions are currently being tested.

TOWARDS STRUCTURES OF SOLUBLE V-ATPase SUBUNITS

The primary requirement for structural studies is the generation of relatively large amounts of highly purified

protein in its native fold. Secondarily, the protein must be stable at high concentrations, and for most applications should be homogeneous and remain monodisperse. Expression in *E. coli* offers the potential for highest yield, but is usually unsuitable for eukaryotic proteins, resulting in insoluble or unfolded polypeptide. Although a number of strategies such as heat-shock (Chen *et al.*, 2002) or osmotic stress (Bhandari and Gowrishankar, 1997) have been exploited to increase yields of eukaryotic proteins, target proteins often remain intractable. The problem of insolubility seems to be particularly exaggerated when attempting to express eukaryotic polypeptides which are components of multiprotein complexes. Aggregation of these polypeptides in the prokaryotic cell may be driven by the same hydrophobic interactions which mediate some specific subunit–subunit interactions in the native system. These reservations aside, the potential for high yields make *E. coli* the system of first recourse.

We have now turned out attention to structural analyses of the single-copy subunits of the *Saccharomyces* V-ATPase. In order to generate material for these studies, we have initially screened Vma7p, Vma4p, Vma8p, Vma5p, Vma6p, and Vma10p for their capacity to be manufactured in *E. coli*, using a variety of fusion constructs and expression conditions. The criteria used to assess suitability are solubility, yield and adoption of a stable fold. In our hands, Vma4p and Vma8p produce only insoluble protein in *E. coli*, irrespective of bacterial strain, fusion partner or induction conditions. Expression of soluble, folded Vma4p has however been reported using an osmoregulated expression system which suppresses formation of insoluble inclusion bodies (Grüber *et al.*, 2002). Vma6p, expressed as a fusion protein with glutathione-S-transferase, produces low yields of soluble protein. However, the smaller subunits Vma7p (118 residues) and Vma10p (114 residues), along with Vma5p (392 residues), appear particularly amenable to expression in *E. coli*, and are produced with high efficiency as soluble, monomeric GST fusion proteins. We have previously reported the expression, purification, and spectroscopic analysis of Vma7p (Jones *et al.*, 2001). Vma5p can be purified to homogeneity (assessed by mass spectrometry, SDS-PAGE and N- and C-terminal amino acid sequencing) via the three-step procedure described for Vma7p in Jones *et al.* (2001), yielding 7–8 mg purified protein per liter culture, comparable to the 5 mg/L achievable for Vma7p. Yields of Vma10p are somewhat lower (1–2 mg/L).

Analysis of secondary structure of the purified proteins by circular dichroism spectroscopy in the range 190–260 nm provides a rigorous measure of α-helical content, but a much less reliable measure of β-sheet (Johnson, 1990). CD spectroscopy indicates that all of the *VMA* gene

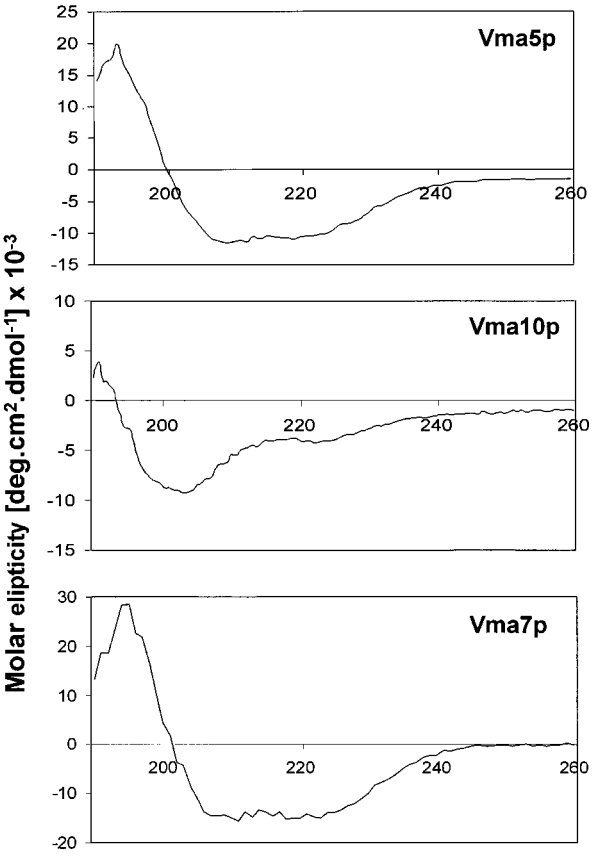

Fig. 5. UV circular dichroism spectra of *Saccharomyces* V-ATPase subunits. Spectra were recorded for Vma5p, Vma10p, and Vma7p after expression in *E. coli* and purification as GST fusion proteins at concentrations of 80, 170, and 150 μg/mL, respectively. Spectra were analyzed using a variety of programs (see Jones *et al.*, 2001, for details) and the secondary structure contents averaged (see Table I).

products analyzed display spectra characteristic of mixed α-helical/β-sheet structures (Fig. 5 and Table I). Vma5p shows a particularly high helical content (50%), whereas the recombinant Vma7p and Vma4p (Grüber *et al.*, 2002) species appear to contain a higher proportion of β-sheet (Table I). The mass and secondary structure of Vma7p, along with its location in the "core" V_1 complex (Imamura *et al.*, 2003; Tomashek *et al.*, 1996), strongly suggest that this subunit represents a structural and functional homologue of the F-ATPase ε-subunit within the V-ATPase (Jones *et al.*, 2001).

CD spectra for Vma10p (Fig. 5) were very noisy at wavelengths below 190 nm, and data were analyzed for the 200–260 nm range only. A small proportion (13%) of α-helix was calculated, the remainder being a combination of β-structure and random coil, between which it was not possible to reliably discriminate. In any case, the secondary structure of heterologously

Table I. Secondary Structure Analysis of Recombinant V-ATPase Subunits by Circular Dichroism Spectroscopy

Subunit	Expression system	Secondary structure (%) α-helix	β-sheet	Other	Reference
Vma5p[a]	*E. coli*, GST fusion	50	21	29	This work
Vma10p[b]	*E. coli*, GST fusion	13	—	84	This work
Vma7p[c]	*E. coli*, GST fusion	30	37	33	Jones *et al.*, 2001
Vma4p	*E. coli*, His-tag	32	23	45	Grüber *et al.*, 2002

[a] Spectra analyzed using Contin (Provencher, 1982).
[b] Spectra (200–260 nm) analyzed using Contin.
[c] Average of outputs from Contin, Lincomb, K2D, G&F, and Jasco data analysis programmes (see Jones *et al.*, 2001, for details).

expressed Vma10p presented here is not consistent with previously reported predictions (Charsky *et al.*, 2000; Hunt and Bowman, 1997). These predictions, supported by localized similarities to the sequence of subunit *b* of the F-ATPase membrane domain, suggest an extended α-helical structure which forms part of the "stator" of the V-ATPase rotary motor (Boekema *et al.*, 1997; Charsky *et al.*, 2000; Domgall *et al.*, 2002; Hunt and Bowman, 1997). Indeed, our own predictions based on molecular modelling using the fold recognition programme "Threader 2" (Jones *et al.*, 1999) also suggest an almost entirely helical structure. "Threader 2" provides a statistically rigorous prediction of fold, comparing segments of test sequence to a library of structures. Scores generated by the programme ("*Z*-scores") are qualitative measures of the quality of matches between folds in the structure library and the sequence under test. In the case of Vma7p, very highly significant *Z*-scores (6.7–4.1) were obtained which predicted α-β-α sandwich architecture with a secondary structure content which correlated with the spectroscopic observations (Jones *et al.*, 2001). In the case of Vma10p, "*Z*-scores" were lower (in the range 2.6–2.9) but still significant, and the polypeptide mapped onto the α-helical regions of the four-helical bundle fold of the *Tar* receptor soluble domain (PDB entry 2asr). This inconsistency between predicted and observed structural characteristics points to the possibility that the heterologously-expressed Vma10p does not adopt its native (helical) fold, even though soluble, monomeric protein is produced. Adoption of an extended helical configuration may require interaction with other subunits of the V-ATPase complex.

FUTURE DIRECTIONS

The ability to generate large amounts of soluble V-ATPase subunits in a stable, monodisperse state is a key step towards solving structures. Analysis of isotopically-labelled Vma7p using heteronuclear NMR methods is now at an advanced state in our laboratory. Modern NMR instrumentation makes it feasible to examine proteins as large as Vma5p using the same methods, in parallel with "traditional" X-ray crystallization approaches. It seems highly likely therefore that over the next few years an increasing number of V-ATPase subunit (and subcomplex) structures will emerge. To complement this, an electron microscopy approach can provide the 3D map into which these structures can be fitted, increasing our understanding of the static and dynamic interactions made by these subunits within the V-ATPase.

ACKNOWLEDGMENTS

This work was supported by United Kingdom BBSRC (project grants to JBCF, grant made available under the MASIF programme to JBCF/MAH/JT), The Wellcome Trust (career development fellowship to MAH and grants for purchase of CD equipment) and the European Commission Framework V programme (QLG1-2000-61801.

REFERENCES

Bhandari, P., and Gowrishankar, J. (2003). *J. Bacteriol.* **179**, 4403–4406.
Boekema, E. J., Ubbink-Kok, T., Lolkema, J. S., Brisson, A., and Konings, W. N. (1997). *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14291–14293.
Bowman, B. J., and Bowman, E. J. (2002). *J. Biol. Chem.* **277**, 3965–3972.
Bowman, E. J., Tenney, K., and Bowman, B. J. (1988). *J. Biol. Chem.* **263**, 13994–14001.
Charsky, C. M. H., Schumann, N. J., and Kane, P. M. (2000). *J. Biol. Chem.* **275**, 37232–37239.
Chattopadhyay, A., and London, E. (1987). *Biochemistry* **26**, 39–45.
Chen, J. Q., Acton, T. B., Basu, S. K., Montelione, G. T., and Inouye, M. (2002). *J. Mol. Microbiol. Biotech.* **4**, 519–524.
Doherty, R. D., and Kane, P. M. (1993). *J. Biol. Chem.* **268**, 16845–16851.

Domgall, I., Venzke, D., Luttge, U., Ratajczak, R., and Böttcher, B. (2002). *J. Biol. Chem.* **277**, 13115–13121.
Falke, J. J., and Koshland, D. E. (1987). *Science* **237**, 1596–1600.
Geyer, M., Fackler, O. T., and Peterlin, B. M. (2002a). *Mol. Biol. Cell* **13**, 2045–2056.
Geyer, M., Yu, H., Mandic, R., Linnemann, T., Zheng, Y. H., Fackler, O. T., and Peterlin, B. M. (2002b). *J. Biol. Chem.* **277**, 28521–28529.
Gibson, L. C., Cadwallader, G., and Finbow, M. E. (2002). *Biochem. J.* **366**, 911–919.
Girvin, M. E., Rastogi, V. K., Abildgaard, F., Markley, J. L., and Fillingame, R. H. (1998). *Biochemistry* **37**, 8817–8824.
Grabe, M., Wang, H., and Oster, G. (2000). *Biophys. J.* **78**, 2798–2813.
Grüber, G., Godovac-Zimmermann, J., Link, T. A., Coskun, U., Rizzo, V. F., Betz, C., and Bailer, S. M. (2002). *Biochem. Biophys. Res. Commun.* **298,** 383–391.
Harrison, M. A., Jones, P. C., Kim, Y.-I., Finbow, M. E., and Findlay, J. B. C. (1994). *Eur. J. Biochem.* **221**, 111–120.
Harrison, M. A., Murray, J., Powell, B., Kim, Y.-I., Finbow, M. E., and Findlay, J. B. C. (1999). *J. Biol. Chem.* **274**, 25461–25470.
Harrison, M. A., Powell, B., Finbow, M. E., and Findlay, J. B. C. (2000). *Biochemistry* **39**, 7531–7537.
Hirata, R., Graham, L. A., Takatsuki, A., Stevens, T. H., and Anraku, Y. (1997). *J. Biol. Chem.* **272**, 4795–4803.
Holliday, L. S., Gluck, S. L., and Lee, B. S. (1999). *J. Bone Miner. Res.* **14**, SU195.
Holzenburg, A., Jones, P. C., Franklin, T., Pali, T., Heimburg, T., Marsh, D., Findlay, J. B. C., and Finbow, M. E. (1993). *Eur. J. Biochem.* **213**, 21–30.
Hunt, I. E., and Bowman, B. J. (1997). *J. Bioenerg. Biomemb.* **29**, 533–540.
Imamura, H., Nakano, M., Noji, H., Muneyuki, E., Ohkuma, S., Yoshida, M., and Yokoyama, K. (2003). *Proc. Natl. Acad. Sci. U.S.A.* **100**, 2312–2315.
Jiang, W., and Fillingame, R. H. (1998). *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6607–6612.
Johnson, W. C., Jr. (1990). *Proteins: Struct. Funct. Genet.* **7**, 205–214
Jones, D. T., Tress, M., Bryson, K., and Hadley, C. (1999). *Proteins: Struct. Funct. Genet.* **Suppl. 3**, 104–111.
Jones, P. C., and Fillingame, R. H. (1998). *J. Biol. Chem.* **273**, 29701–29705.
Jones, P. C., Harrison, M. A., Kim, Y.-I., Finbow, M. E., and Findlay, J. B. C. (1995). *Biochem. J.* **312**, 739–747.
Jones, R. P. O., Hunt, I. E., Jaeger, J., Ward, A., O'Reilly, J., Barratt, E., Findlay, J. B. C., and Harrison, M. A. (2001). *Mol. Memb. Biol.* **18**, 283–290.
Kane, P. M. (1995). *J. Biol. Chem.* **270**, 17025–17032.
Kawasaki-Nishi, S., Nishi, T., and Forgac, M. (2001). *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12397–12402.
Klaholz, B. P., Pape, T., Zavialov, A. V., Myasnikov, A. G., Orlova, E. V., Vestergaard, B., Ehrenberg, M., and van Heel, M. (2003). *Nature* **421**, 90–94.
Landolt-Marticorena, C., Williams, K. M., Correa, J., Chen, W., and Manolson, M. F. (2000). *J. Biol. Chem.* **275**, 15449–15457.
Leng, X. H., Manolson, M. F., Liu, Q., and Forgac, M. (1996). *J. Biol. Chem.* **271**, 22487–22493.
Leng, X. H., Nishi, T., and Forgac, M. (1999). *J. Biol. Chem.* **274**, 14655–14661.
Liu, Z., Zhang, J., Li, P., Chen, S. R. W., and Wagenknecht, T. (2002). *J. Biol. Chem.* **277**, 46712–46719.
Mandel, M., Moriyama, Y., Hulmes, J . D., Pan, Y.-C. E., Nelson, H., and Nelson, N. (1988). *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5521–5524.
Nelson, H., Mandiyan, S., and Nelson, N. (1994). *J. Biol. Chem.* **269**, 24150–24155.
Nishi, T., and Forgac, M. (2002). *Nat. Rev. Mol. Cell Biol.* **3**, 94–103.
Noumi, T., Beltran, C., Nelson, H., and Nelson, N. (1991). *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1938–1942.
Orlova, E. V., Papakosta, M., Booy, F. P., van Heel, M., and Dolly, J. O. (2003). *J. Mol. Biol.* **326**, 1005–1012.
Pakula, A. A., and Simon, M. I. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4144–4148.
Pali, T., Finbow, M. E., and Marsh, D. (1999). *Biochemistry* **38**, 14311–14319.
Powell, B. (1999). PhD Thesis, University of Leeds, Leeds, UK.
Provencher, S. W. (1982). *Comp. Phys. Commun.* **27**, 229–242.
Sagermann, M., Stevens, T. H., and Matthews, B. W. (2001). *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7134–7139.
Sedgwick, E. G., Meuller, J., Hou, C., Rydstrom, J., and Bragg, P. D. (1997). *Biochemistry* **36**, 15285–15293.
Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001). *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9056–9061.
Stevens, T. H., and Forgac, M. (1997). *Annu. Rev. Cell Dev. Biol.* **13**, 779–808.
Tomashek, J. J., Garrison, B. S., and Klionsky, D. J. (1997). *J. Biol. Chem.* **272**, 16618–16623.
Tomashek, J. J., Sonnenburg, J. L., Artimovich, J. M., and Klionsky, D. J. (1996). *J. Biol. Chem.* **271**, 10397–10404.
Ubarretxena-Belandia, I., and Engelman, D. M. (2001). *Curr. Opin. Struct. Biol.* **11**, 370–376.
Ubbink-Kok, T., Boekema, E. J., van Breemen, J. F. L., Brisson, A., Konings, W. N., and Lolkema, J. S. (2000). *J. Mol. Biol.* **296**, 311–321.
Vik, S. B., and Antonio, B. J. (1994). *J. Biol. Chem.* **269**, 30364–30369.
Walker, M., Zhang, X. Z., Jiang, W., Trinick, J., and White, H. D. (1999). *Proc. Natl. Acad. Sci. USA* **96**, 465–470.
Wilkens, S., and Forgac, M. (2001). *J. Biol. Chem.* **276**, 44064–44068.
Wilkens, S., Vasilyeva, E., and Forgac, M. (1999). *J. Biol. Chem.* **274**, 31804–31810.
Ya, S., Aiwu, Z., Al-Lamki, R. S., and Karet, F. E. (2003). *J. Biol. Chem.* **278**, 20013–20018.
Yasuda, R., Noji, H., Kinosita, K., and Yoshida, M. (1998). *Cell* **93**, 1117–1124.