J Bioenerg Biomembr (2012) 44:341–350
DOI 10.1007/s10863-012-9442-3

Solution structure of subunit *a*, $a_{104\text{-}363}$, of the *Saccharomyces cerevisiae* V-ATPase and the importance of its C-terminus in structure formation

Phat Vinh Dip · Wuan Geok Saw · Manfred Roessle · Vladimir Marshansky · Gerhard Grüber

Received: 9 March 2012 / Accepted: 19 April 2012 / Published online: 5 May 2012

Abstract The 95 kDa subunit *a* of eukaryotic V-ATPases consists of a C-terminal, ion-translocating part and an N-terminal cytosolic domain. The latter's N-terminal domain (~40 kDa) is described to bind in an acidification-dependent manner with cytohesin-2 (ARNO), giving the V-ATPase the putative function as pH-sensing receptor. Recently, the solution structure of the very N-terminal segment of the cytosolic N-terminal domain has been solved. Here we produced the N-terminal truncated form $SCa_{104\text{-}363}$ of the N-terminal domain ($SCa_{1\text{-}363}$) of the *Saccharomyces cerevisiae* V-ATPase and determined its low resolution solution structure, derived from SAXS data. $SCa_{104\text{-}363}$ shows an extended S-like conformation with a width of about 3.88 nm and a length of 11.4 nm. The structure has been superimposed into the 3D reconstruction of the related A_1A_O ATP synthase from *Pyrococcus furiosus*, revealing that the $SCa_{104\text{-}363}$ fits well into the density of the collar structure of the enzyme complex. To understand the importance of the C-terminus of the protein $SCa_{1\text{-}363}$, and to determine the localization of the N- and C-termini in $SCa_{104\text{-}363}$, the C-terminal truncated form $SCa_{106\text{-}324}$ was produced and analyzed by SAXS. Comparison of the $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ shapes showed that the additional loop region in $SCa_{104\text{-}363}$ consists of the C-terminal residues. Whereas $SCa_{104\text{-}363}$ is monomeric in solution, $SCa_{106\text{-}324}$ forms a dimer, indicating the importance of the very C-terminus in structure formation. Finally, the solution structure of $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ will be discussed in terms of the topological arrangement of subunit *a* and cytoheisn-2 in V-ATPases.

Keywords Vacuolar ATPase · V_1V_O ATPase · V_1 ATPase · Vph1p · Subunit *a* · Cytohesin-2 · Small angle X-ray scattering

Abbreviations

HSQC	Heteronuclear single quantum coherence
NMR	Nuclear magnetic resonance
RMSD	Root mean square distance
SAXS	Small angle X-ray scattering

Introduction

The vacuolar ATPase (V_1V_O ATPase) has been described as being nature's most versatile H^+-pump (Nishi and Forgac 2002), which is common to all eukaryotic cells. This enzyme is found on the intracellular organelles and also on the plasma membrane of specialized cell types. ATP hydrolysis by the V-ATPases drives the transport of protons into the lumen of organelles, which this macromolecular complex decorates. The V_1V_O ATPase is composed of a water-soluble V_1 sector,

Phat Vinh Dip and Saw Wuan Geok are authors contributed equally to this article.

P. V. Dip · W. G. Saw · G. Grüber (✉)
School of Biological Sciences, Nanyang Technological University,
60 Nanyang Drive,
Singapore 637551, Republic of Singapore
e-mail: ggrueber@ntu.edu.sg

M. Roessle
European Molecular Biology Laboratory, Hamburg Outstation,
EMBL c/o DESY,
22603 Hamburg, Germany

V. Marshansky
Center for Systems Biology, Program in Membrane Biology and
Division of Nephrology, Simches Research Center,
Massachusetts General Hospital and Department of Medicine,
Harvard Medical School,
Boston, MA 02114, USA

in which ATP hydrolysis takes place, and an integral membrane subcomplex, V_O (Grüber et al. 2001). 3D reconstructions of single particle analysis revealed that the V-ATPase complex consists of an A_3B_3 headpiece (Radermacher et al. 2001), which is structurally and functionally connected via peripheral stalks and a central stalk to the membrane-embedded V_O sector and/or the collar domain, which is oriented horizontal to the membrane-embedded sector (Venzke et al. 2005; Wilkens and Forgac 2001).

The integral V_O sector of *Saccharomyces cerevisiae* V-ATPase contains five different subunits in a stoichiometry of a_1:d_1:$c_{4\text{-}5}$:c'_1:c''_1 (Nishi and Forgac 2002). The V_O complex can be subdivided into two parts that rotate relative to each other, the peripheral stalk and the proton-translocating ring. The stator of the V_O part is proposed to consist of subunit *a* (Venzke et al. 2005; Wilkens and Forgac 2001) and *d* (Thaker et al. 2007). The subunits $c_{4\text{-}5}$:c'_1:c''_1 form a ring and each subunit has multiple transmembrane domains, which are termed proteolipids because of their hydrophobic nature (Clare et al. 2006). This ring forms the proton-conducting unit together with subunit *a*. Besides H^+-pumping, subunit *a* of eukaryotic V-ATPases are proposed to be directly involved in the assembly and organelle-specific targeting of V-ATPases (Grüber and Marshansky 2008; Marshansky and Futai 2008). Moreover, subunit *a* has been described as a pH sensor element of V_1V_O ATPases, by interacting with cytohesin-2 (Hurtado-Lorenzo et al. 2006). Cytohesin-2 is a guanine nucleotide exchange factor (GEF) that is responsible for GTP/GDP exchange and activation of Arp-family GTP-binding proteins, including ADP-ribosylation factor 6 (Arf 6) (Franco et al. 1998). Both cytohesin-2 and Arf proteins are essential for various signaling pathways such as cytoskeleton organization, maintenance of Golgi structure and function, synaptic transmission, epithelial cell migration, membrane recycling, exocytosis regulation in the neuroendocrine cells (Marshansky 2007). Subunit *a* is an essential component of this V-ATPase endosomal sensory machinery to control the protein targeting and degradation in the early endosomes in a collective response with Arf6 and its cognate GEF cytohesin-2 (Hurtado-Lorenzo et al. 2006). Loss of endosomal acidification led to the failure of carrier vesicles formation to transport the cargo between early and late endosomes, resulting in the accumulation at early endosomes and inhibition of endocytosis (Hurtado-Lorenzo et al. 2006). Peptide-mapping (Merkulova et al. 2010) of the mammalian isoform subunit *a2* has shown that the very N- and the C-terminal segments of the cytosolic domain of subunit *a* interact with cotohesin-2 (Merkulova et al. 2010). Most recently, the NMR structure of the N-terminal segment has been solved and a defined binding epitope composed of the conserved residues F_5, E_8, M_{10} and Q_{14} at the N-terminal segment has been determined (Hosokawa et al. 2012). NMR studies with the entire soluble region of recombinant subunit *a* (Vph1p) from *Saccharomyces cerevisiae*, called $SCa_{1\text{-}388}$, showed its interaction with the catalytic Sec7 domain of cytohesin-2, highlighting that binding and signaling between V-ATPases and cytohesins is evolutionary conserved among eukaryotes (Hosokawa et al. 2012).

Subunit *a* is divided into a C-terminal, membrane-integrated part and an N-terminal cytosolic domain (~40 kDa) located on the cytoplasmic side. The crystallographic structure of the N-terminal domain, including residues 1–301 ($MRa_{1\text{-}301}$), of the evolutionary related archaea-type A_1A_O ATP synthase has been determined recently (Srinivasan et al. 2011). $MRa_{1\text{-}301}$ has been assigned to the density of the collar domain, oriented parallel with the cytoplasmic surface of the membrane and proposed to be in close proximity to the N-termini of the heterodimers of the peripheral stalk subunits E and G (Srinivasan et al. 2011).

Because of the central role of the cytosolic N-terminal domain of subunit *a* in eukaryotic V-ATPases, insights into its structural features are essential. So far only the very N-terminal segment of the mammalian N-terminal domain of subunit *a* of eukarytotic V-ATPases has been structurally characterized (Hosokawa et al. 2012). Here we describe the production, purification and low resolution solution structure of the truncated forms of the cytosolic N-terminal domain of subunit *a* from the *S. cerevisiae* V-ATPase, which are called $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ and comprise the residues 104–363 and 106–324, respectively. These constructs provided a unique opportunity to use a subtractive approach of the C-terminal truncated form $SCa_{106\text{-}324}$ to understand the contribution of the C-terminal segment to the overall structure of the N-terminal domain of subunit *a,* its orientation inside the enzyme complex as well as its interaction with the Sec7 domain of cytohesin-2.

Material and methods

Biochemicals

Pfu DNA Polymerase and Ni^{2+}-NTA-chromatography resin were obtained from Qiagen (Hilden, Germany); restriction enzymes were purchased from MBI Fermentas (St. Leon-Rot, Germany). Chemicals for gel electrophoresis and trypsin used for in-gel digestion were purchased from Serva (Heidelberg, Germany) and Promega (Madison, WI, USA), respectively. ($^{15}NH_4$) Cl and (^{13}C) glucose were purchased from Cambridge Isotope Laboratories (Andover, U.S.A.). All other chemicals were at least of analytical grade and obtained from BIOMOL (Hamburg, Germany), Merck (Darmstadt, Germany), Sigma (Deisenhofen, Germany) or Serva (Heidelberg, Germany).

Purification of $a_{104\text{-}363}$ and $a_{106\text{-}324}$ of *S. cerevisiae* V_1V_O ATPase

To amplify the *S. cerevisiae* VPH1 (subunit *a*) coding region for $SCa_{104\text{-}363}$, oligonucleotide primers 5′- CGTTCCAC*C CATGG*GTTCAGTGATAGATGATTATGTCCGG-3′ (forward primer) and 5′- TTGGATAAT*GAGCTC*TTAATCAA TACCCAATCTTGCGATCATTTC-3′ (reverse primer), incorporating *NcoI* and *SacI* restriction sites, respectively, were designed. The *S. cerevisiae* genomic DNA was used as the template for the polymerase chain reaction (PCR). The PCR product was ligated into the pET9d-His_3 vector (Grüber et al. 2002), afterwards transformed into *E. coli* cells (strain BL21 (DE3)), which were grown on 30 μg/ml kanamycin-containing Luria-Bertoni (LB) agar-plates. To express His_3-$a_{104\text{-}363}$, liquid cultures were shaken in LB medium containing kanamycin (30 μg/ml) for about 6 h at 30 °C until an optical density OD_{600} of 0.6–0.7 was reached. To induce expression of His_3-$a_{104\text{-}363}$, the cultures were supplemented with isopropyl (thio)-β-D-galactoside (IPTG) to a final concentration of 1 mM. Following incubation for another 16 h at 20 °C, the cells were harvested at 8 500 × *g* for 12 min, 4 °C. Subsequently, they were lysed on ice by sonication for 3×1 min in Buffer A (50 mM Tris/HCl, pH 8.5, 200 mM NaCl, 1 mM DTT, 2 mM PMSF and 2 mM Pefabloc^{SC} (BIOMOL)). Precipitated material was separated by centrifugation at 10 000 × *g* for 35 min. The supernatant was filtered (0.45 μm; Millipore) and passed over a 2 ml Ni^{2+}-NTA resin column to isolate $SCa_{104\text{-}363}$, according to Grüber et al. (2002). The His-tagged protein was allowed to bind to the matrix for 1.5 h at 4 °C and eluted with an imidazole-gradient (25–250 mM) in Buffer A. Fractions containing His_3-subunit $SCa_{104\text{-}363}$ were identified by SDS-PAGE (Laemmli 1970), pooled and concentrated as required using Centricon YM-3 (3 kDa molecular mass cut off) spin concentrators (Millipore). Imidazole was removed by gel filtration chromatography using a Superdex 75 HR 10/30 column (GE Healthcare) and a buffer of 50 mM Tris/HCl, pH 8.5, 200 mM NaCl and 10 mM EDTA.

In order to obtain a further C-terminal truncated form of $SCa_{104\text{-}363}$, $SCa_{106\text{-}324}$ has been produced. Primers 5′- CAC CAA*CCATGG*TAGTGATAGATGATTATGTC-3′ (forward primer) and 5′- ATCATA*GAGCTC*TTAGTTCAAAATTT CAAAAATCG-3′ (reverse primer) were designed, subsequently cloned and purified as described above. The purity and homogeneity of all protein samples were analyzed by SDS-PAGE (Laemmli 1970). SDS-gels were stained with Coomassie Brilliant Blue G250. Protein concentrations were determined by the bicinchoninic acid assay (BCA; Pierce, Rockford, IL., USA). We have also used the corresponding extinction coefficients for protein concentration, resulting in similar values.

Purification of the Sec7 domain of cytohesin-2

The over-expression was performed in BL21 (DE3) cells and purified using a three-step protocol. The first step involved enrichment of His-tagged Sec7 recombinant protein by specific binding to Ni^{2+}-NTA matrix. In order to cleave the N-terminal GST-tag from the Sec7-domain, eluted protein in the imidazole concentration of 75–200 mM were digested overnight with Precision Proteases (GE Healthcare). The second step employed affinity chromatography using GSTrap^{TM} column (GE Healthcare). GST binds to the GSTrap^{TM} column (GE Healthcare), whereas the Sec7 domain remains in the flow through. The final step was performed using a Superdex 75 HR 10/30 column (GE Healthcare). All purification steps were done in a buffer of 50 mM Hepes and 300 mM NaCl, pH 7.5.

Circular dichroism spectroscopy

Steady state CD spectra were measured in the far UV-light (190–260 nm) using a *Chirascan* spectropolarimeter (Applied Photophysics). Spectra were collected in a 60 μl quartz cell (Hellma) with a path length of 0.1 mm, at 20 °C and a step resolution of 1 nm. The readings were average of 2 s at each wavelength and the recorded ellipticity values were the average of three determinations for each sample. CD spectroscopy of subunit $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ (2 mg/ml) was performed in a buffer of 50 mM Tris/HCl (pH 8.5), 200 mM NaCl and 10 mM EDTA. The spectrum for the buffer was subtracted from the spectrum of the protein. CD values were converted to mean residue ellipticity (Θ) in units of degree cm^2 $dmol^{-1}$ using the software *Chirascan Version 1.2*, Applied Photophysics. This baseline corrected spectrum was used as input for computer methods to obtain predictions of secondary structure. The CD spectra were analyzed as described recently (Thaker et al. 2007).

NMR titration experiments of $SCa_{104\text{-}363}$ and Sec7 domain of cytohesin-2

Interaction studies were performed between $SCa_{104\text{-}363}$ and the Sec7 domain of ARNO using NMR sectroscopy. Both proteins have been buffer exchanged to 25 mM phosphate buffer 300 mM NaCl, pH 6.5 prior in presence of 10 % D_2O (v/v). ^1H–^{15}N heteronuclear single quantum coherence (HSQC) spectrum of residues R_{61} to R_{252} corresponding to Sec7 of cytohesin-2 was used as starting point. Ligand protein $SCa_{104\text{-}363}$ was added at increasing amounts. Respective ^1H–^{15}N HSQC spectra were then recorded. A constant amount (0.2 μM) of Sec7 was used, followed by adding unlabeled $SCa_{104\text{-}363}$ as binding partner at increasing amounts to a molar ratio of 1:1–1:2. Changes in chemical shift were then monitored in HSQC spectrum. Experiments

were performed on Bruker Avance 600 machine using Topspin for acquisition and processing of spectra. Respective spectra were overlapped to monitor chemical shift changes, further analysis were done in SPARKY (Goddard and Kneller 1997).

X-ray scattering experiments and data analysis of $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$

Small angle X-ray scattering (SAXS) data for $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ were collected by following standard procedures on the X33 SAXS camera (Boulin et al. 1986; Roessle et al. 2007) of the EMBL Hamburg located on a bending magnet (sector D) on the storage ring DORIS III of the Deutsches Elektronen Synchrotron (DESY). A photon counting Pilatus 1 M pixel detector (67×420 mm^2) was used at a sample—detector distance of 2.4 m covering the range of momentum transfer $0.1<s<4.5\ nm^{-1}$ ($s=4p\ \sin(q)/l$, where q is the scattering angle and l=0.15 nm is the X-ray wavelength). The s-axis was calibrated by the scattering pattern of Silver-behenate salt (d-spacing 5.84 nm). The scattering from the buffer alone was measured before and after each sample measurement and the average of the scattering before and after each sample is used for background subtraction. A range of protein concentrations (1.5 to 6.4 mg/ml) was measured for both yeast constructs $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ to assess and remove any concentration-dependant inter-particle effects. $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ have been measured in buffer containing 50 mM Tris/HCl (pH 8.5) and 200 mM NaCl, 10 mM EDTA and 1 mM DTT. The protein as well as the buffer samples have been injected automatically using the sample-changing robot for solution scattering experiments at the SAXS station X33 (Round et al. 2008). All the data processing steps were performed automatically using the program package PRIMUS (Svergun 1993). The forward scattering $I(0)$ and the radius of gyration R_g were evaluated using the Guinier approximation (Guinier and Fournet 1955) assuming that for spherical particles at very small angles ($s< 1.3/R_g$) the intensity is represented by $I(s) = I(0)\exp\left(-(sR_g)^2/3\right)$. These parameters were also computed from the entire scattering patterns using the indirect transform package GNOM (Svergun et al. 2001), which also provide the distance distribution function $\rho(r)$ of the particle as defined:

$$\rho(r) = 2\pi \int I(s) sr \sin(sr)\, ds$$

The molecular mass of both proteins were calculated by comparison with the forward scattering from the reference solution of bovine serum albumin (BSA). From this procedure a relative calibration factor for the molecular mass (MM) can be calculated using the known molecular mass of BSA (66.4 kDa) and the concentration of the reference solution by applying

$$MM_p = I(0)_p/c_p \times \frac{MM_{st}}{I(0)_{st}/c_{st}}$$

where $I(0)_p$, $I(0)_{st}$ are the scattering intensities at zero angle of the studied and the BSA standard protein, respectively, MM_p, MM_{st} are the corresponding molecular masses and c_p, c_{st} are the concentrations. Errors have been calculated from the upper and the lower I(0) error limit estimated by the Guinier approximation.

Low-resolution models of $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ were built by the program GASBOR (Svergun et al. 2001), which represents the protein as an assembly of dummy atoms forming a chain-compatible model inside a search volume defined by a sphere of the diameter D_{max}. The spatial positions of these dummy atoms are approximately corresponding Cα atoms in the protein structure. The number of residues is equal to that in the protein. Starting from a random model, GASBOR employs simulated annealing to build a scattering equivalent model fitting the experimental data $I_{exp}(s)$ to minimize discrepancy:

$$\chi^2 = \frac{1}{N-1}\sum_j \left[\frac{I_{\exp}(s_j) - cI_{calc}(s_j)}{\sigma(s_j)}\right]^2$$

where N is the number of experimental points, c is a scaling factor and $I_{calc}(s_j)$ and $\sigma(s_j)$ are the calculated intensity and the experimental error at the momentum transfer s_j, respectively. In order to compare the solution structure of $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ with the atomic structure of the N-terminal domain of subunit a ($MRa_{1\text{-}302}$) of the A_1A_O ATP synthase from *M. ruber*, the high resolution model (PDB entry 3rrk; Srinivasan et al. 2011) has been aligned using SUBCOMB (Svergun et al. 2001). This program aligns all possible pairs of models and arranges the smallest average discrepancy among the models.

Results and discussion

Production, purification and secondary structure content of $SCa_{104\text{-}363}$

Most recently the N-termial segment of subunit a of the eukaryotic V-ATPase has been determined by NMR-spectroscopy and its interaction with the Sec7 domain of cytohesin-2 has been mapped (Fig. 1; Merkulova et al. 2010; Hosokawa et al. 2012). To gain more structural insight into the remaining N-terminal domain of subunit a, the construct $a_{104\text{-}363}$ ($SCa_{104\text{-}363}$), including the amino acid residues 104 to 363 of the *S. cerevisiae* V_1V_O ATPase, has been constructed. Induction of His-tagged protein production under

Fig. 1 Amino acid alignment of the N-terminal segment of subunit *a* of *M. musculus* and *S. cerevisiae* using Jalview 271 (Waterhouse et al. 2009). The secondary structure on top of the alignment is based on the NMR structure of the known binding sites to the Sec7 domain (Hosokawa et al. 2012). The mapped binding regions of subunit *a* to the Sec7 domain are shown as red lines (Merkulova et al. 2010). The amino acids of $SCa_{104\text{-}363}$ and $SCa_{106\text{-}324}$ are highlighted by (− −) and the last amino acid residue of $SCa_{106\text{-}324}$ is marked by * *S.c.* 324

the conditions specified resulted in a soluble 30 kDa protein, representing $a_{104\text{-}363}$ of the *S. cerevisiae* V_1V_O ATPase. Ni^{2+}-NTA affinity chromatography has been used in the first purification step. Recombinant $SCa_{104\text{-}363}$ was eluted by an imidazole-gradient (25–250 mM) in buffer consisting of 50 mM Tris/HCl (pH 8.5), 200 mM NaCl, 10 mM EDTA and 2 mM PMSF. Protein eluting at 75 to 125 mM imidazole was collected and subsequently applied to a Superdex 75 column (Fig. 2a, *lane 1*) in order to isolate a pure and monodispersed protein. Analysis of the isolated protein by SDS-PAGE revealed the high purity of the protein (Fig. 2a, *lane 2*).

The secondary structure of recombinant $SCa_{104\text{-}363}$ was determined from circular dichroism spectra, measured between 190 and 260 nm (Fig. 3, solid spectra). The maximum at 192 nm and the minima at 208 and 222 nm indicate the presence of α-helical structures in the protein. The average secondary structure content was 81 % α-helix, 8 % β-sheet and 11 % random coil. This result is consistent with secondary structure predictions based on $SCa_{104\text{-}363}$ amino-acid sequence. The molar ellipticity values at 208 nm and at 222 nm are in a ratio of 0.96, indicating that many of the helices in residues in $SCa_{104\text{-}363}$ are in close neighborhood, since non-interacting helices typically give ratios of around 0.8.

Low resolution structure of V-ATPase $SCa_{104\text{-}363}$ in solution

The high purity of recombinant $SCa_{104\text{-}363}$ enabled to perform small-angle X-ray scattering experiments, to verify the proper three dimensional folding and to determine the first low resolution structure of the major part of the cytosolic N-terminal domain of subunit *a* of an eukaryotic V-ATPase. SAXS patterns of $SCa_{104\text{-}363}$ were recorded as described in Material and Methods to yield the final composite scattering curve shown in Fig. 4a, which indicates a monodispersed protein in solution as shown also by size exclusion chromatography. Inspection of the low angle of the Guinier plots reveals a good data quality and no protein aggregation. $SCa_{104\text{-}363}$ has a radius of gyration (R_g) of 3.88 nm and a maximum dimension D_{max} of 11.4 nm (Fig. 4b). Comparison

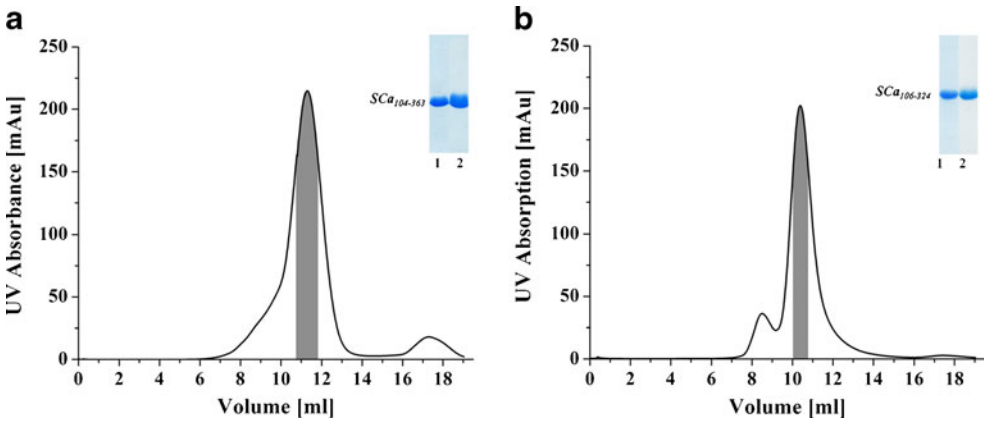

Fig. 2 Purification of $SCa_{104\text{-}363}$ (**a**) and $SCa_{106\text{-}324}$ (**b**). Following purification on Ni^{2+}-NTA resin, the proteins were applied onto a Superdex 75 column using buffer (50 mM Tris/HCl, pH 8.5, 200 mM NaCl, 10 mM EDTA) at a flow rate of 0.5 ml/min. Insert in figures showing SDS gels after purification on Ni^{2+}-NTA resin (*lane 1*), indicated fraction from elution peaks after Superdex 75 (*lane 2*) of $SCa_{104\text{-}363}$ (**a**) and $SCa_{106\text{-}324}$ (**b**), respectively

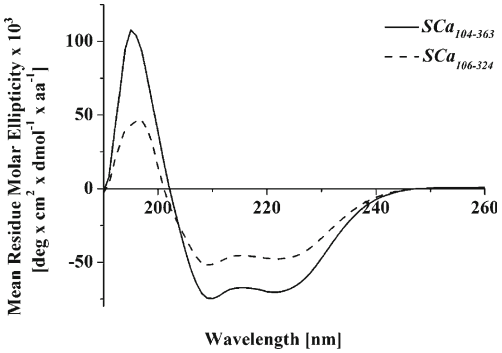

Fig. 3 Far UV-CD spectrum of $SCa_{104\text{-}363}$ (—) and the truncated $SCa_{106\text{-}324}$ (– –)

of the forward scattering of $SCa_{104\text{-}363}$ with the values obtained from a reference solution of bovine serum albumin, (BSA; 66.4±2 kDa) yields a molecular mass of 30±2 kDa, in agreement with the results of the gel filtration chromatography, indicating that $SCa_{104\text{-}363}$ is monomeric at the concentrations used. The distance distribution function $\rho(r)$ shows a bell-shaped function, with a maximum in $p(r)$ at 3.6 nm (Fig. 4b), whereas the slight shoulder arise from 8.4 nm to 12 nm, indicating to an extended formation of $a_{104\text{-}363}$.

The solution structure of $a_{104\text{-}363}$ was restored *ab initio* from the scattering patterns, shown in Fig. 4a. The obtained shape for the protein yields a good fit to the experimental data in the entire scattering range. The corresponding fit, shown in Fig. 4a, has a discrepancies of χ^2=1.08. The protein appears to be an elongated molecule with an S-like shape (Fig. 5a). The low resolution structure has dimensions of about 11.4 nm in length and 3.9 nm in width.

Most recently, the crystallographic structure of the cytosolic N-terminal segment of subunit *a*, $a_{1\text{-}301}$ ($MRa_{1\text{-}301}$), of the related archaea type A_1A_O ATP synthase from *Meiothermus ruber* has been determined (Srinivasan et al. 2011). Superimposing the homologoue region $a_{80\text{-}268}$ ($MRa_{80\text{-}268}$) of the *M. ruber* structure into the low resolution solution structure of $SCa_{104\text{-}363}$ reveals, that $MRa_{80\text{-}268}$ accommodates well in the shape of $SCa_{104\text{-}363}$ with an r.m.s. deviation of 1.8 Å (Fig. 5a). In addition, the entire crystal structure of $MRa_{1\text{-}301}$ fits well into the shape of $SCa_{104\text{-}363}$ with an r.m.s.d. of 1.66 Å, indicating that the truncation of 104

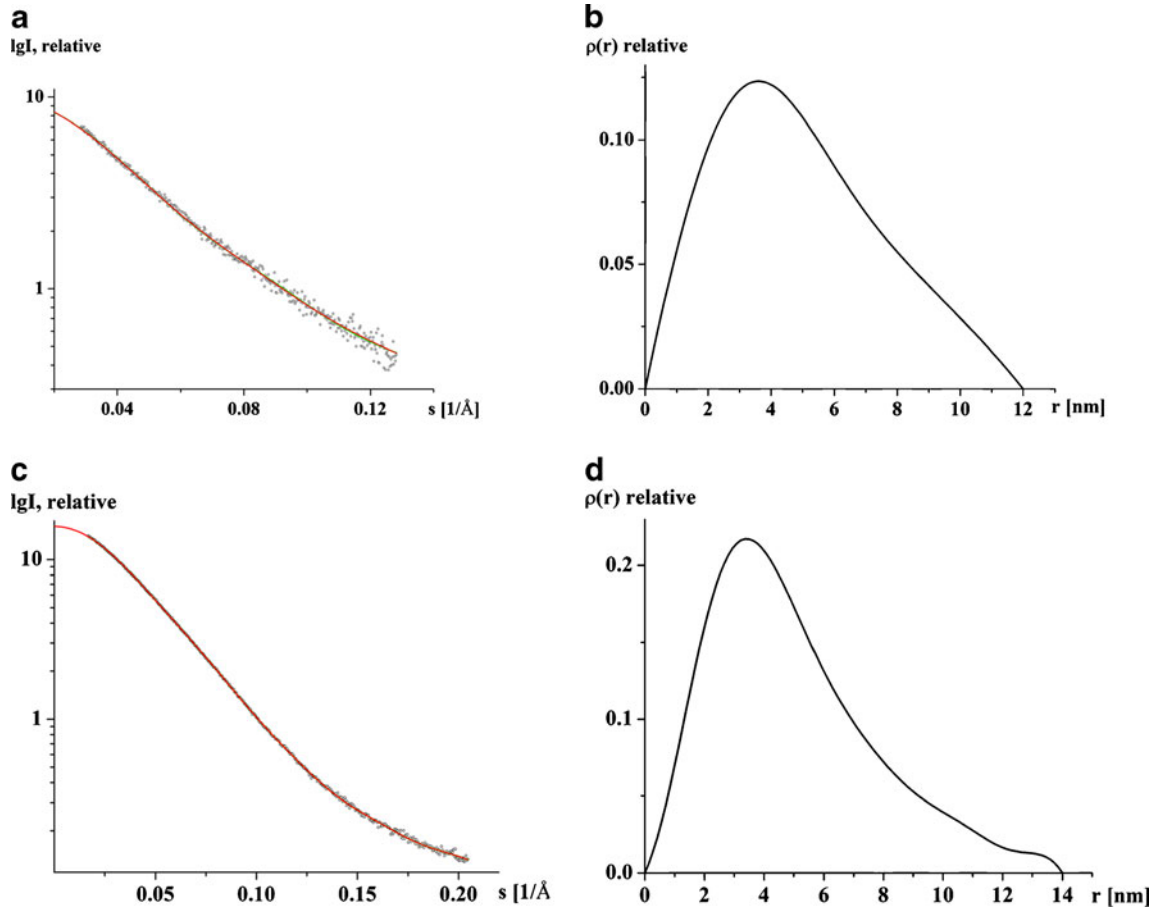

Fig. 4 **a** Experimental scattering curves (*Black Circle*) and the fitting curves (—; *green*: experimental, *red*: calculated from *ab initio* model) of $SCa_{104\text{-}363}$ and (**c**) $SCa_{106\text{-}324}$. Distance distribution functions of $SCa_{104\text{-}363}$ (**b**) and $SCa_{106\text{-}324}$ (**d**), respectively

amino acids in $SCa_{104\text{-}363}$ does not alter the overall dimension and shape of the protein (Fig. 5a). The structural comparisons

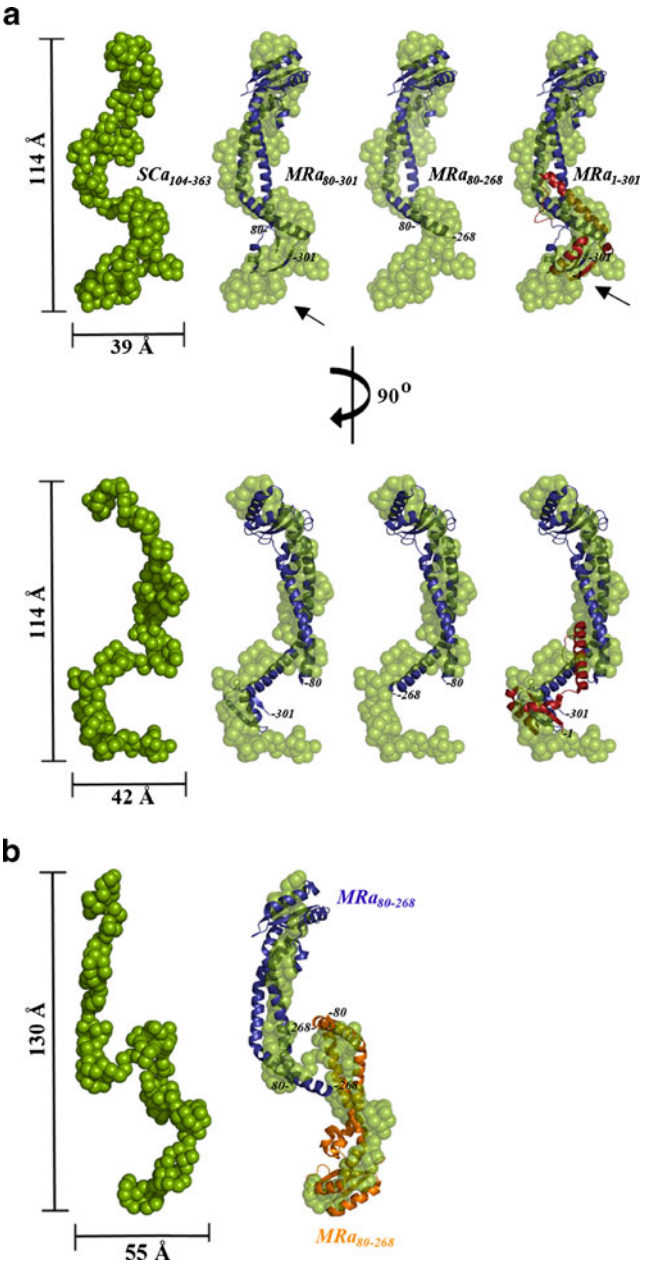

Fig. 5 **a** Low resolution structures of $SCa_{104\text{-}363}$ in green spheres and its 90° view in the first column. Superposition of $SCa_{104\text{-}363}$ with subunit *a* homologous model of the A_1A_O ATP synthase, $MRa_{80\text{-}301}$, $MRa_{80\text{-}268}$, and $MRa_{1\text{-}301}$ (column 2, 3, and 4) (Srinivasan et al. 2011), respectively. As highlighted by an arrow, the comparison of $SCa_{104\text{-}363}$ (column 2) and $MRa_{1\text{-}301}$ indicates that the C-terminal loop of $MRa_{80\text{-}301}$ is located in the lower part of the $SCa_{104\text{-}363}$ shape and that this extended segment of $SCa_{104\text{-}363}$ forms the very C-terminus of $SCa_{104\text{-}363}$. The red colored subunit $MRa_{1\text{-}301}$ segment (column 4) harbors the first 78 amino acids, which are not included in $SCa_{104\text{-}363}$. The comparison of $SCa_{104\text{-}363}$ and $MRa_{1\text{-}301}$ reveals that N- (*red*) and C-terminus (*blue*) are in proximity to each other, which is highlighted by an arrow (column 4). **b** C-terminal truncated $SCa_{106\text{-}324}$ in column 1 superimposed with $MRa_{80\text{-}268}$ (column 2)

$SCa_{104\text{-}363}$ with $MRa_{80\text{-}301}$ as well as with $MRa_{1\text{-}301}$ enable to allocate the remaining C-terminal 62 amino acids of $SCa_{104\text{-}363}$. The superposition shows that this C-terminal region can be assigned to the loop domain at the bottom of $SCa_{104\text{-}363}$. Since the N-terminal domain of subunit *a* has been described to form the so-called collar domain together with subunit C above the Vo-sector (Srinivasan et al. 2011), the low resolution structure of $SCa_{104\text{-}363}$ was superimposed into the 23 Å resolution EM density of the related *P. furiosus* A-ATP synthase and fits well into this density (Fig. 9a). This arrangement shows that the presented C-terminal region of 62 amino acids of $SCa_{104\text{-}363}$ forms a part of the unresolved hinge region, which connects the N-terminal domain of subunit *a*, which is oriented parallel with the cytoplasmic surface of the membrane and in close proximity to the peripheral stalks, with the C-terminal and ion-translocating domain in the membrane.

NMR titration $SCa_{104\text{-}363}$ with Sec7 domain of cytohesin-2

Previously, six peptides ($a2_{1\text{-}17}$, $a2_{35\text{-}49}$, $a2_{195\text{-}214}$, $a2_{215\text{-}230}$, $a2_{313\text{-}331}$ and $a2_{386\text{-}402}$) of the mouse *a*2 isoform have been identified to interact with cytohesin-2 and the two peptides $a2_{1\text{-}17}$ and $a2_{35\text{-}49}$ have been shown to bind most strongly to the catalytic Sec7 domain of cytohesin-2 (Merkulova et al. 2010). NMR titration experiments have been performed to analyze, whether $SCa_{104\text{-}363}$ is still capable to bind to the Sec7 domain of cytohesin-2. All NMR experiments were carried out on a Bruker DRX 600 MHz spectrometer equipped with a cryoprobe using Topspin for acquisition and processing of spectra. Respective spectra were overlapped to monitor chemical shift changes, further analysis were done in SPARKY (Goddard and Kneller 1997). All proteins were buffer exchanged in 25 mM phosphate buffer, pH 6.5 and 300 mM NaCl. $^{1}H–^{15}N$ heteronuclear single quantum coherence spectrum of Sec7 domain (residues R_{61} to R_{252}) of cytohesin-2 was used as starting point, which reveals that Sec7 domain is well folded (Fig. 6, red peaks). The $^{1}H–^{15}N$ HSQC spectra were recorded at 298 K with a fixed concentration of 0.2 mM of Sec7 domain, and $a_{104\text{-}363}$ was titrated with 0.4 mM (1:2) to Sec7. This titration experiments revealed four main chemical shift differences, which are highlighted in Fig. 6 (boxes A–D), indicate binding of $SCa_{104\text{-}363}$ to the Sec7 domain. In comparison, recent NMR titration experiments with Sec7 and $SCa_{1\text{-}388}$ of the yeast V-ATPase showed a more and wide chemical shift, highlighting the strong binding of the $a_{1\text{-}388}$ mainly via its very N-terminal sequence formed by the $a2_{1\text{-}17}$ and $a2_{35\text{-}49}$ (Hosokawa et al. 2012). These results presented here indicate that $SCa_{104\text{-}363}$ is not only properly folded but also functional in its interaction with cytohesin-2. Whether the $SCa_{104\text{-}363}$ and Sec 7 binding occurs via the homologoues peptides of $a2_{195\text{-}214}$, $a2_{215\text{-}230}$, $a2_{313\text{-}331}$ and/or $a2_{386\text{-}402}$ has to be shown in future studies.

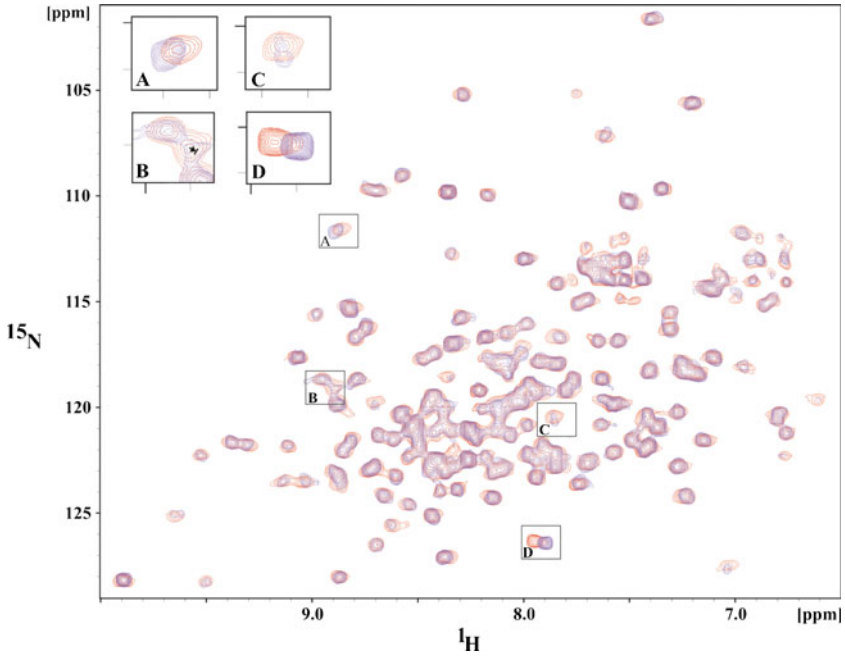

Fig. 6 NMR titration spectra of Sec7 domain of ARNO and $SCa_{104\text{-}363}$. **a** Overlay of 2D ^{1}H–^{15}N HSQC spectrum of Sec7 domain of ARNO (*red*) and after adding of $SCa_{104\text{-}363}$ (1:2, purple) in 25 mM sodium phosphate buffer (pH 6.5), 300 mM NaCl at 298 K. Sec7/$SCa_{104\text{-}363}$ titration indicates chemical shifts of four peaks, indicated as **a**–**d** and highlighted in the section

Production and secondary structure of the truncated form $SCa_{106\text{-}324}$

Previous pull down assays revealed, that besides the very N-terminal sequence of subunit *a*, the C-terminal sequence $a2N_{313\text{-}331}$ of the soluble domain of subunit *a* interacts with cytohesin-2 (Merkulova et al. 2010). Furthermore, as described above the very C-terminal residues (last 62 amino acids) of $SCa_{104\text{-}363}$ forms partially the hinge, connecting the N-terminal domain of subunit *a* with the C-terminal ion-pumping domain. In order to study whether this very C-terminus is important for structure formation, the construct $SCa_{106\text{-}324}$ has been generated. The expressed protein was purified by metal chelate affinity chromatography (Fig. 2b, *lane1*) and gel-filtration (Superdex 75 column) (Fig. 2b, *lane2*). According to the elution profile, $SCa_{106\text{-}324}$ eluted earlier than the larger construct $SCa_{104\text{-}363}$, indicating that the protein is either of higher molecular weight and/or of larger hydrodynamic volume. To confirm a proper folding of $SCa_{106\text{-}324}$, CD spectroscopy has been performed, showing a secondary structure composition of 78 % α-helix, 6.0 % β-sheet and 16 % random coil (Fig. 3).

Shape and domain structure of $SCa_{106\text{-}324}$

$SCa_{106\text{-}324}$ was further investigated by SAXS to determine the shape of the protein. The experimental solution scattering curve of the $SCa_{106\text{-}324}$ protein is presented in Fig. 4c. The obtained shape for $SCa_{106\text{-}324}$ yields a good fit to the experimental data in the entire scattering range with a discrepancies of χ^2=1.25 (Fig. 4c). Compared with the SAXS-data of $SCa_{104\text{-}363}$, the radius of gyration of $SCa_{106\text{-}324}$ increased slightly 0.10 nm and the maximum dimension of the $SCa_{106\text{-}324}$ protein increased to 14.0 nm (Fig. 4d). Comparison of the forward scattering with the values obtained for BSA yields a molecular mass of 50±2 kDa, which confirm the dimerization of the truncated form of $SCa_{106\text{-}324}$. Qualitative analysis of the distance distribution function *ρ(r)* reveals that $a_{106\text{-}324}$ consists of a bell-shaped function, with a principal maximum in the *p(r)* at short distances around 3.2 nm (Fig. 4d) and separated protuberance domain giving rise to a shoulder from 8.4 nm to 14 nm.

The low resolution shape of $SCa_{106\text{-}324}$ determined *ab initio* is shown in Fig. 5b. Like the $SCa_{104\text{-}363}$, $SCa_{106\text{-}324}$ has an elongated S-like shape, with a dimension of about 13.7 nm in length and 3.8 nm in width. The comparison of both shapes reveals that the truncation of 39 amino acids at

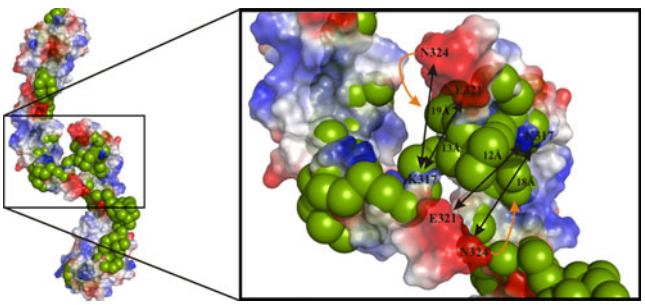

Fig. 7 **a** Superposition of $SCa_{106\text{-}324}$ into the molecular surface with the electrostatic potential of the subunit $MRa_{80\text{-}268}$ of the A_1A_O ATP synthase (Srinivasan et al. 2011). Red and blue areas are negatively and positively charged areas, respectively, calculated with the program Pymol (DeLano 2002). The arrow indicates charged residues in the C-terminus of $MRa_{80\text{-}268}$, which might form an interaction with the C-terminus of each other

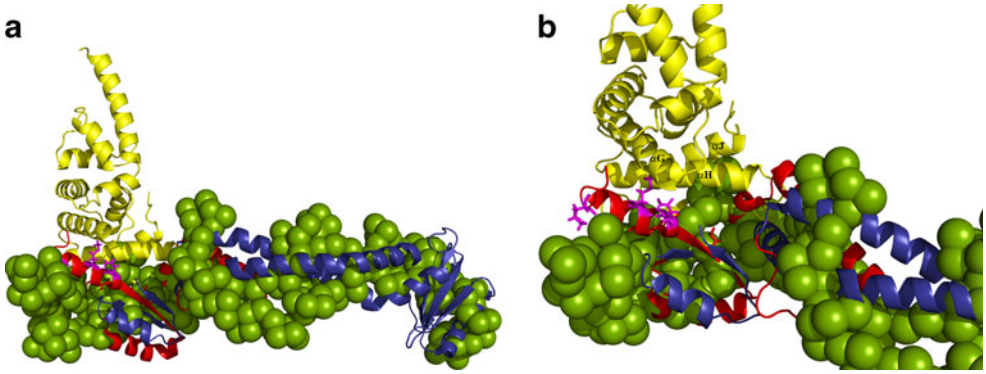

Fig. 8 **a** Model of interaction of the proximal lobe region of $SCa_{104\text{-}363}$ (*green*) with the groove region of the Sec7 domain (Cherfils et al. 1998; *yellow*) of ARNO. The $SCa_{104\text{-}363}$ solution structure and the $MRa_{1\text{-}301}$ crystal structure (the N- and C-terminus are shown in *red* and *blue*, respectively; Srinivasan et al. 2011) were docked with the already characterized binding site of the N-terminal segment $a2_{1\text{-}17}$, whose interacting amino acid residues are labeled in magenta (Hosokawa et al. 2012). **b** The solution structure of $SCa_{104\text{-}363}$ (*green*) accommodates well in the groove between αG, αH and αJ of the Sec7 domain (*yellow*)

the C-terminus of $SCa_{104\text{-}363}$ causes a dimer formation of $SCa_{106\text{-}324}$ (Fig. 5b), which is reflected by the increased D_{max} value (Fig. 4d) and confirms the elution profile of the gel filtration experiments described above. Based on the amino acid alignment of the secondary structure prediction of subunit *a* from yeast with the crystal structure of $MRa_{1\text{-}301}$, $MRa_{80\text{-}268}$ is the homologous segment to $SCa_{106\text{-}324}$. As shown in Fig. 5b, two molecules of $MRa_{80\text{-}268}$ accommodate well in the shape of $SCa_{106\text{-}324}$, where both C-termini are arranged in close proximity. The electrostatic surface potential of $MRa_{80\text{-}268}$ reveals a strong charged surface caused by the C-terminal residues A268 and D266, whose homologue residues are E321 and N324 in $SCa_{106\text{-}324}$, and which are proposed to interact with the positive residue K317 of the second molecule (Fig. 7). In addition, $MRa_{80\text{-}268}$ can also be superimposed into the shape of $SCa_{104\text{-}363}$ (Fig. 5a), showing that the loop of $MRa_{1\text{-}301}$, formed by the residues 268–302, can be assigned to the C-terminal region of $SCa_{104\text{-}363}$ shape.

Interaction of $SCa_{104\text{-}363}$ and Sec7 and its mechanistic implications

This C-terminal region and the very N-terminal segment of the cytosolic N-terminal domain of subunit *a* are in close proximity (Fig. 5a) and both termini are interacting with cytohesin-2. In addition, the four residues of the first

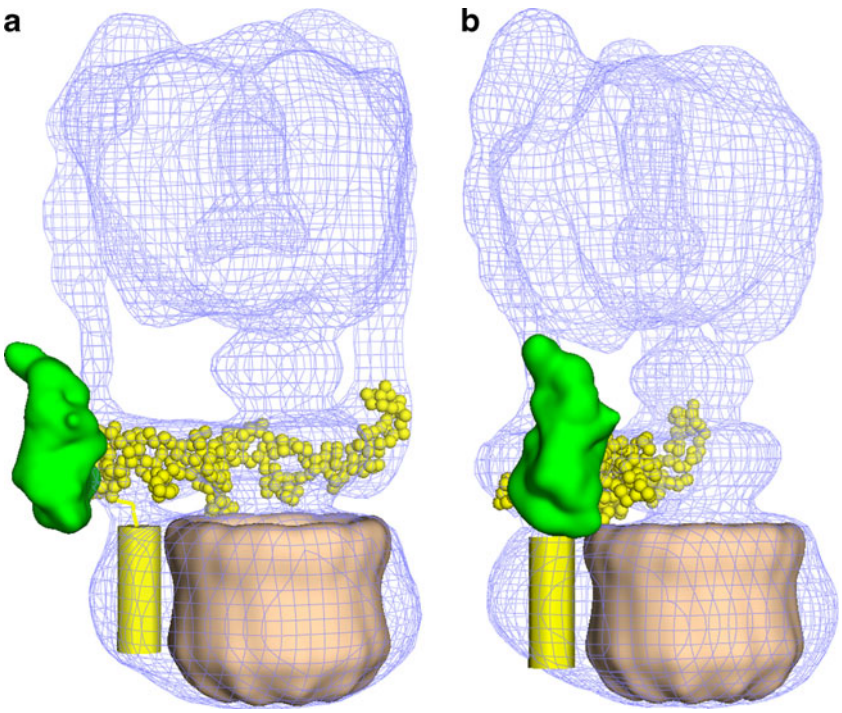

Fig. 9 **a**–**b** The 3D reconstruction EM map of the archaeon *P. furiosus* enzyme (Vonck et al. 2009; EM Data Base ID: EMD-1542) in two different orientations. The structure of $SCa_{106\text{-}324}$ (*yellow*) and the *c*-ring (*wheat*) from PDB 2BL2 (Murata et al. 2005) were used for the fitting. Sec7 (*green*) has been accommodated based on the interactions described in Fig. 8. The models propose that ion-translocation in the interface of the *c*-ring and the C-terminal membrane-embedded domain of subunit *a* (*yellow cylinder*) will be affected by the binding of the Sec7 domain to the N-terminal domain of subunit *a*

seventeen amino acids peptide form the epitope of subunit *a*, which interacts with Sec7, have been determined recently by NMR spectroscopy (Hosokawa et al. 2012). Based on these, a structural model of the $SCa_{104\text{-}363}$, $MRa_{1\text{-}301}$ crystal structure with the characterized binding site $a2_{1\text{-}17}$ and the Sec7 domain is shown in Fig. 8a, demonstrating that the Sec7 binding occurs at the proximal lobe region with no steric hindrance and clashes. The solution structure of $SCa_{104\text{-}363}$ accommodates in the groove between αG, αH, αI and αJ of the Sec7 domain (Cherfils et al. 1998). A detailed view of the model of interaction of the proximal lobe region of $SCa_{104\text{-}363}$ with the groove region of the Sec7 domain is presented in Fig. 8b, revealing that the helices αH and αJ are in proximity to $SCa_{104\text{-}363}$, which explains their interaction in the NMR titration experiments. As shown in Fig. 9a, the low resolution structure of $SCa_{104\text{-}363}$ accommodates well into EM density of the related *P. furiosus* A-ATP synthase, where it forms a part of the so-called collar domain. Since Sec7 does bind $SCa_{104\text{-}363}$ via its groove between αG, αH, αI and αJ the arrangement shown in Fig. 9b reveals that the interaction side of Sec7 and the N-terminal domain of subunit *a* is close to the hinge region, which connects the soluble N-terminal- with the membrane-embedded C-terminal domain, which is involved in proton-pumping. Taken together, the crosstalk of Sec7 of ARNO and the N-terminal domain of subunit *a* may cause an alteration of the hinge region of subunit *a*, which will be transferred to alterations in the C-terminal domain, with the consequence of changes in proton-pumping.

Acknowledgements We thank Dr. S.M. Malathy Sony for her support in SAXS data analysis and art. Phat Vinh Dip is grateful to receive the Singapore International Graduate Award (SINGA). This research was supported by A*STAR BMRC (09/1/22/19/609).

References

Boulin C, Kempf R, Koch MHJ, McLaughlin SM (1986) Nucl Instrum Meth A 249:399–407

Cherfils J, Menetrey J, Mathieu M, Le Bras G, Robineau S, Beraud-Dufour S, Antonny B, Chardin P (1998) Nature 392:101–105

Clare DK, Orlova EV, Finbow MA, Harrison MA, Findlay JB, Saibil HR (2006) Structure 14:1149–1156

DeLano W (2002) DeLano Scientific

Franco M, Boretto J, Robineau S, Monier S, Goud B, Chardin P, Chavrier P (1998) Proc Natl Acad Sci USA 95:9926–9931

Goddard TD, Kneller DG (1997) SPARKY 3, University of California, San Francisco, CA

Grüber G, Marshansky V (2008) BioEssays 30:1–14

Grüber G, Wieczorek H, Harvey WR, Müller V (2001) J Exp Biol 204:2597–2605

Grüber G, Godovac-Zimmermann J, Link TA, Coskun Ü, Rizzo VF, Betz C, Bailer SM (2002) Biochem Biophys Res Commun 298:383–391

Guinier A, Fournet G (1955) Small-angle scattering of x-rays. Wiley, New York

Hosokawa, Merkulova M, Bakulina A, Zhuang Z, Khatri A, Dip PV, Jian X, Randazzo PA, Grüber G, Ausiello DA, Marshansky V (2012) Nat Cell Biol: submitted

Hurtado-Lorenzo A, Skinner M, Annan JE, Futai M, Sun-Wada G-H, Bourgoin S, Casanova J, Wildeman A, Bechoua S, Ausiello DA, Brown D, Marshansky V (2006) Nat Cell Biol 8:124–136

Laemmli UK (1970) Nature 227:680–685

Marshansky V (2007) Biochem Soc Trans 35:1092–1099

Marshansky V, Futai M (2008) Curr Opin Cell Biol 20:415–426

Merkulova M, Bakulina A, Thaker YR, Grüber G, Marshansky V (2010) Biochim Biophys Acta 1797:1398–1409

Murata T, Yamato I, Kakinuma Y, Leslie AG, Walker JE (2005) Science 308:654–659

Nishi T, Forgac M (2002) Nat Mol Cell Biol 31:94–103

Radermacher M, Ruiz T, Wieczorek H, Grüber G (2001) J Struct Biol 135:26–37

Roessle MW, Klaering R, Ristau U, Robrahn B, Jahn D, Gehrmann T, Konarev P, Round A, Fiedler S, Hermes C, Svergun D (2007) J Appl Crystallogr 40:190–194

Round AR, Franke D, Moritz S, Huchler R, Fritsche M, Malthan D, Klaering R, Svergun DI, Roessle M (2008) J Appl Crystallogr 41:913–917

Srinivasan S, Vyas NK, Baker ML, Quiocho FA (2011) J Mol Biol 412:14–21

Svergun DI (1993) J Appl Crystallogr 26:258–267

Svergun DI, Petoukhov MV, Koch MHJ (2001) Biophys J 80:2946–2953

Thaker YR, Roessle M, Grüber G (2007) J Bioenerg Biomembr 39:275–289

Venzke D, Domgall I, Kocher T, Fethiere J, Fischer S, Böttcher B (2005) J Mol Biol 349:659–669

Vonck J, Pisa KY, Morgner N, Brutschy B, Müller V (2009) J Biol Chem 284:10110–10119

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Bioinformatics 25:1189–1191

Wilkens S, Forgac M (2001) J Biol Chem 276:44064–44068