Journal of Proteomics & Bioinformatics - Open Access
Research Article JPB/Vol.2/August 2009

In Silico Identification of Putative Proton Binding Sites of a Plasma Membrane H+-ATPase Isoform of *Arabidopsis Thaliana*, AHA1

Sunil Kumar[1], Binod Bihari Sahu[2], Niraj Kanti Tripathy[3] and Birendra Prasad Shaw*[2]

[1]Bioinformatics Centre, Institute of Life Sciences, Nalco Square, Bhubaneswar-751023, India
[2]Environmental Biotechnology Laboratory, Institute of Life Sciences, Nalco Square, Bhubaneswar-751023, India
[3]Department of Zoology, Berhampur University, Berhampur, India

*Corresponding author: Birendra Prasad Shaw. Ph D, Environmental Biotechnology Laboratory, Institute of Life Sciences, Nalco Square, Bhubaneswar-751023, India, Tel: +91-674-2302611; Fax: +91-674-2300728; E-mail: bpsils@yahoo.com

Received July 10, 2009; Accepted August 18, 2009; Published August 19, 2009

Citation: Kumar S, Sahu BB, Tripathy NK, Shaw BP (2009) *In Silico* Identification of Putative Proton Binding Sites of a Plasma Membrane H+-ATPase Isoform of *Arabidopsis Thaliana*, AHA1. J Proteomics Bioinform 2: 349-359. doi:10.4172/jpb.1000095

Abstract

The plasma membrane potential and secondary transport systems in all eukaryotes are energized by the activity of P-type ATPase membrane proteins: H+ATPase (the proton pump) in plants and fungi and Na+, K+- ATPase (the sodium-potassium pump) in animals. The overall shape of proton pumps has been revealed by electron microscopy. The crystal structure of AHA2, a plasma membrane H+-ATPase isoform of *Arabidopsis thaliana*, by X-ray crystallography at 3.6 Å was available. The isoform is expressed mainly in root. In the present study homology modeling along with transmembrane topology predictions has been used to build the atomic model of AHA1, another plasma membrane H+-ATPase isoform of *Arabidopsis thaliana* expressing in both root and shoot. AHA2 was used as the template. The homology modeling was done using the MODELLER9v2 software. The model energy was minimized by applying molecular mechanics method. The root mean square deviation (RMSD) for C^{α} atoms between the X-ray and the homology-modeled structures was 0.4 Å. Ten transmembrane helices (TM1-TM10) in the model were identified. The final model obtained by molecular mechanics and dynamics method was assessed by PROCHECK and VERIFY 3D graph, which showed that the final refined model is reliable. Asp^{684} of AHA1 is suggested to act as an essential proton acceptor during proton translocation. The residue may also participate in defining the E_1 proton-binding site. Despite difference in the tissue-specific expression of the two proteins, no remarkable difference in their structure was observed, suggesting that the enzyme isoform evolution may not be linked to its tissue-specific expression.

Keywords: H+-ATPase; Proton binding site; Regulatory domain; Hydronium ion, Proton pump

Abbreviations: PM: Plasma Membrane; EM: Energy Minimization; BLAST: Basic Local Alignment Search Tool

Introduction

PM (plasma membrane)-H+ATPase is a plasma membrane bound proton (H+) pumping ATPase and is member of p-type (E1-E2 type) ATPases representing a superfamily of cation transporters involved in the transport of cations like H+, Na+, K+, Ca2+, Mg2+, etc. at the expense of ATP. Being a member of P-ATPase, the reaction catalysed by it proceed with the formation of a phosphorylated intermediate, unlike that of other ATP hydrolyzing enzymes, V_1V_0-ATPase present in tonoplast and the membranes of the intracellular organelles of eukaryotic cells and F_1F_0 ATPase present in thylakoid and mitochondrial membranes. Like PM-H+-ATPase, V_1V_0-ATPase (or V-ATPase)

is also primarily a H^+ pumping enzyme, the two enzymes differ greatly in their structure, besides being located in different membranes of the cells. Moreover, PM- H^+-ATPase is unique only to the plants and fungi.

The free energy gradient ($\Delta\mu H^+$) produced by plasma membrane H^+-ATPase (PM-H^+-ATPase) provides driving force for secondary active transport for the maintenance of ionic balance across the membrane and uptake of necessary elements or molecules, both charge and neutral, mediated by various transporters like uni-, sym- and antiport carriers (Sze, 1985). These transporters are thus called secondary transporters, as they utilize the energy in the form of proton (H^+ ion) gradient generated by the PM H^+- ATPases. Thus, this membrane enzyme has essential house-keeping functions in the life of plants. It is also involved in special functions such as turgor-dependent processes of growth extension and movements of stomatal guard cells and pulvini (Okazaki, 2002). In most plant species PM-H^+ATPase is encoded by a large gene family. In tomato (*Lycopersicon esculentum* cv. large cherry) seven genes have been identified and genomic DNA blot analysis indicates a family of at least 10 genes (Oufattole et al., 2000). Extensive search of cDNA and genomic libraries from *Nicotiana plumbaginifolia* resulted in identification of nine PM-H^+-ATPase genes (Perez et al., 1992; Moriau et al., 1993; Oufattole et al., 2000), while complete sequencing of *Arabidopsis* genome has uncovered 12 genes (Palmgren et al., 2001) and that of *Oryza sativa* (Japonica) has led to retrieval of 10 *OSA* (*Oryza sativa* PM-H^+ATPase) genes (Arango et al., 2003). Out of the 10 annotated *OSA* genes in *Oryza sativa*, six (*OSA1*, *OSA2*, *OSA3*, *OSA4*, *OSA8* and *OSA9*) have been characterized at cDNA level (Wada et al., 1992; Ookura et al., 1994; Kikuchi et al., 2003). All these isoforms have been cloned from various organs of plants at different stages of development grown under control condition. The existence of multiple isoforms of the enzyme might indicate that some pumps could be redundant. Nonetheless, isoform diversity may also be related to cellular differentiation with individual isoform exhibiting tissue and developmental specific expression and having slightly different biochemical and regulatory properties (Palmgren et al., 2001; Arango et al., 2003).

The overall shape of proton pumps has been revealed by the electron microscopy. The ion translocation proceeds through a Ping-Pong mechanism, cation species being sequentially transferred in opposite directions across the membrane. The opening and closure of the access channels are tightly coupled to conformational changes within the cytoplasmic domains induced by ATP binding, phosphorylation, subsequent dephosphorylation and phosphate release from the enzyme. Therefore, ions are sequentially taken up at ion binding sites from one side of the lipid bilayer, occluded within the transmembrane domain, and then released on the other side of the membrane, in exchange for the second ion species (Kühlbrandt, 2004; Møller et al., 1996; Apell, 2004; Albers, 1967; Post et al., 1969).

The X-ray crystal structure of the sarcoplasmic reticulum Ca^{2+}-ATPase from rabbit skeletal muscle and AHA2 from *Arabidopsis thaliana* are also available (Toyoshima et al., 2000; Toyoshima et al., 2002; Sørensen et al., 2004; Pedersen et al., 2007) and these structures provide critical insights into the structure and mechanism of P-type ATPases. These structures explained the general architecture of the catalytic subunit of P-type ATPase which divided the cytoplasm part of the protein into three domains: (i) A domain (actuator), (ii) P domain (phosphorylation), (iii) N (nucleotide binding) domain and a transmembrane segment of 10 helices (TM1 to TM10) bearing the two embedded Ca^{2+} binding sites. However, despite the existence of a large no. of isoforms of PM-H^+-ATPase with great differences in homologies, a comparative modeling of various isoforms has not been done. This is despite the fact that expressions of many isoforms have been reported to be tissue specific. For example, in *Arabidopsis thaliana* AHA3 expresses in the phloem companion cells (DeWitt et al., 1991)[20], AHA9 in anthers (Houlné et al., 1994) and *AHA6*, *AHA8*, *AHA9* and *AHA12* in pollen grains. Reports are also there for tissue specific expression of PM-H^+ATPase isoforms in other plants, like *Nicotiana plumbaginifolia* (Lefebvre et al., 2005) and *Vicia faba* (Hentzen et al., 1996). The only model available for PM-H^+ATPase is for *Arabidopsis thaliana* AHA2 (Pedersen et al., 2007).

Recently, expression analysis of all the PM-H^+ATPase isoforms of *Arabidopsis thaliana* in root tissue using Affymetrix GeneChip revealed maximum expression of AHA1 and AHA2 in comparison to other isoforms (Dinneny et al., 2008). Palmgren et al., (1999), however, have reported no expression of AHA2 in stem tissue. Besides, a comparative study of the activity of the two enzyme isoforms expressed heterologously it has been observed that the specific activity of AHA2 is approximately double than that of AHA1 (Palmgren et al., (1999). Hence, it was intriguing to know whether there could be difference in the two isoforms of the enzyme at the level of their association with the membrane and substrate affinity/availability. Therefore, in this report we have

modeled AHA1 isoform for a putative proton binding site in the transmembarne region by homology modeling method based on the template of the auto-inhibited AHA2.

Most comparative protein modeling tools adopt rigid body assembly approach which constructs the model from a few core regions and loops obtained from related template protein structures. Another approach adopts segment matching which relies on structurally conserved atoms from template and calculate the co-ordinates of other atoms. The MODELLER uses an approach involving distance geometry or optimization techniques to satisfy spatial restraints obtained from the alignment of target sequence with the template structures. MODELLER allows degree of flexibility and automation which is widely used program in comparative modeling and developed a database of protein models (Modbase). In general, the homology modeling method is based on the assumption that the structure of an unknown protein is similar to known structures of some reference proteins. The MODELLER software employs probability density functions (PDFs) as the spatial restraints rather than energy (Sali et al., 1993; 1994; 1995). The main chain conformation of a given residue in the model is described by restraints that depend upon the residue type, the main chain conformation of equivalent residues in the reference proteins and the local sequence similarity. The PDFs used in restraining the model structure are derived from correlations between structural features in a database of families of homologous proteins aligned on the basis of their 3D structures. These functions are used to restrain Cα- Cα distances, main chain N-O distances, main-chain and side-chain dihedral angles, etc. The 3D model of a protein is obtained by optimization of the molecular pdf such that the model violates the input restraints as little as possible. The molecular pdf is derives as a combination of pdfs restraining individual spatial features of the whole molecule. The optimization procedure is a variable target function method that applies the conjugate gradients algorithm to positions of all non-hydrogen atoms. In our case, the sequence identity between the target and template is 94.2%, and hence the fold (secondary structures arrangement in 3D space well conserved), C-alpha trace, and backbone in the model derived is likely to be highly accurate. Besides, since there is high sequence identity, most of the side chains of the model will have similar side sign torsion angles and any bad steric contacts are removed during energy minimization of the modeled structure. Hence, it was decided to build the model of AHA1 using MODELLER programme. The model built was validated using widely followed Procheck and Verify3D programmes.

Materials and Methods

In the present investigation, the protein sequence of *Arabidopsis thaliana* AHA1 was retrieved from the NCBI database (NP_179486). BLASTP (Altschul et al., 1990) search was performed, against Brook Heaven Protein Data Bank (PDB) (Berman et al., 2000) with default parameters to find the suitable templates for homology modeling. Based on the maximum identity with high score and lower e-value, crystal structure of the plasma membrane proton pump from *Arabidopsis thaliana* at 3.6 Å resolution (PDB code: 3b8c) was used as template for homology modeling. The sequence identity was found to be 94 percent (similarity = 97%, E-Value = 0.0). The ClustalW (http://www.ebi.ac.uk/clustalw) (Thompson et al., 1994) program was used for sequence alignment.

3D Structure Generation

The academic version of MODELER (http//:www.salilab.org/modeler) (Sali et al., 1993) was used for 3D structure generation based on the information obtained from sequence alignment. Out of 20 models generated by MODELLER, the one with the best G-score of PROCHECK (Laskoswki et al., 1993) and with the best VERIFY3D (Luthy et al., 1992) profile was subjected to energy minimization. Using the parameters as a distance-dependent dielectric constant ϵ = 1.0 and non bonding cutoff of 14Å, CHARMm (Brooks et al., 1983), force field and CHARM-all-atom charges, initially a steepest descent algorithm was used to remove close van der Waals contacts, followed by conjugate gradient minimization until the energy showed stability in sequential repetition. All hydrogen atoms were included during the calculation. The energy minimization was started with main chain of the core and then all core side chains were subjected to the same. All calculations were performed by using ACCELRYS DS Modeling 2.0 (Accelrys Inc. San Diego, CA 92121, USA) software suite. During these steps, the quality of the initial model was improved. VERIFY3D (a structure evaluation server) were used to check the residue profiles of the obtained three-dimensional models. The weighted root mean square deviation (RMSD) of the modeled protein was calculated using combinatorial extension (CE) algorithm (Shindyalov et al., 1998). In order to assess the stereo-chemical qualities of the three dimensional models PROCHECK analysis was performed.

Transmembrane Topology Predictions

The transmembrane topology of AHA1 was predicted from the amino acid sequence by averaging the results of four different predictive algorithms: DAS (Cserzo et

	HMMTOP	TMHMM	TMPRED	DAS
M1	64-86	61-83	58-84	96-109
M2	99-115	93-112	93-112	243-262
M3	249-268	240-262	239-261	276-298
M4	281-303	277-299	277-295	624-635
M5	650-669	645-667	618-640	649-684
M6	674-693	671-690	647-666	672-682
M7	712-734	710-732	713-731	712-729
M8	757-774	752-771	753-771	757-770
M9	787-806	783-805	787-804	785-805
M10	817-836	815-832	814-833	819-831

Table 1: AHA1 topological prediction.

	HMMTOP	TMHMM	TMPRED	DAS
M1	50-72	50-72	50-73	
M2	85-102	82-101	82-101	85-97
M3	235-254	229-251	228-250	231-251
M4	267-289	266-288	266-284	265-286
M5	639-655	634-656	607-629	638-672
M6	660-679	660-679	636-655	702-718
M7	698-720	699-721	698-717	747-759
M8	743-760	741-760	742-760	774-793
M9	773-792	772-794	776-793	774-793
M10	807-826	804-826	803-823	

Table 2: 3b8c topological prediction.

al., 1997), HMMTOP (Tusnady et al., 1998), TMHMM (Krogh et al., 2001) and TMPRED (Hofmann et al., 1993) (Table 1). The accuracy of the prediction were assessed by using the same algorithms on 3b8c protein and the results were compared with the topology defined in the 3.6 Å structure (3b8c) (Table 2).

Result and Discussion

Model Building

The structure of the AHA1 protein has not yet determined experimentally (X-ray or NMR), and therefore we built a model following homology modeling protocol. BLASTP search was performed against PDB with default parameters to find the suitable templates for homology modeling. Based on the maximum identity with high score and lower e-value 3b8c (AHA2) was used as the template for homology modeling. Sequence alignment between AHA1 and 3b8c was done using ClustalW program. The sequence – structure alignment used for model building is shown in figure. 1. The alignment is characterized by some insertions and deletions in the loop regions. The first 11 residues and the last 105 residues were deleted because the corresponding residues were not present in the template. Modeling was carried out from 12 to 844 residues followed by a rigorous refinement of the model by means of energy minimization using CHARMm (Brooks et al., 1983) force field. The final stable structure of AHA1 is shown in figure 2.

Protein Structure Validation

To validate the model obtained from the homology modeled AHA1 structure, a Ramachandran plot was drawn and the structure was analyzed by PROCHECK (Laskoswki et al., 1993), a well-known protein structure checking program. It was found that the phi/psi angles of 77.4% residues fell in the most favored regions, 17.5% residues lied in the additional allowed regions 4.0% fell in the generously allowed regions. Only 1.1% of the residues lied in the disallowed conformations (Figure 3). The overall PROCHECK G-factor for the homology modeled structure was -0.35. The score indicates that the modeled structure is acceptable as value is greater than the acceptable value of – 0.50. The structural superimposition of C^{α} trace of the model over the template structure 3b8c (Figure 4) using CE program (http://cl.sdsc.edu/ce.html) gave a root mean square deviation (RMSD) of 0.4 Å (Z-score 8.7) indicating further that the structure of the model was valid.

The amino acid sequence alignment between the template and the final model of AHA1 was generated using CLUSTALW program. The secondary structures was analyzed and compared by the JOY (Mizuguchi et al., 1998) server (protein sequence-structure representation and analysis) and it was found that the secondary structures of the template and final AHA1 model were highly conserved. This also shows that the final model obtained for AHA1 in this study (Figure 5) is highly reliable. However, despite difference in the tissue-specific expression of the two proteins, phylogenetic tree constructed for various H^+ATPase isoforms (Sahu and Shaw, 2009) showed no remarkable difference in their structure, suggesting that the enzyme isoform evolution may not be linked to its tissue-specific expression.

The four clearly defined domains in AHA2, i.e. a transmembrane domain with ten helices, M1 to M10, and the three cytosolic domains comprising of N (nucleotide binding; residues 338-488), P (phosphorylation; residues 308-337 and 489-625) and A (actuator; residues 12-57 and 129-233) are also present in the model derived for AHA1. The N domain remains inserted into the P domain through a hinge (including the conserved sequence motif DPPR) and with bound nucleotide it can move towards the P domain to assemble the catalytic site, at which Asp^{329} will become phosphorylated once every pumping cycle. The A domain, is situated on the top of M2 stimulates dephosphoryla-

```
CLUSTAL 2.0.10 multiple sequence alignment

3B8C            VDLEKIPIEEVFQQLKCSREGLTTQEGEDRIQIFGPNKLEEKKESKLLKFLGFMWNPLSW 60
AHA1            VDLEKIPIEEVFQQLKCTREGLTTQEGEDRIVIFGPNKLEEKKESKILKFLGFMWNPLSW 60
                ******************:************** ***************:**************

3B8C            VMEMAAIMAIALANGDGRPPDWQDFVGIICLLVINSTISFIEENNAGNAAAALMAGLAPK 120
AHA1            VMEAAALMAIALANGDNRPPDWQDFVGIICLLVINSTISFIEENNAGNAAAALMAGLAPK 120
                *** **:**********.********************************************

3B8C            TKVLRDGKWSEQEAAILVPGDIVSIKLGDIIPADARLLEGDPLKVDQSALTGESLPVTKH 180
AHA1            TKVLRDGKWSEQEAAILVPGDIVSIKLGDIIPADARLLEGDPLKVDQSALTGESLPVTKH 180
                ************************************************************

3B8C            PGQEVFSGSTCKQGEIEAVVIATGVHTFFGKAAHLVDSTNQVGHFQKVLTAIGNFCICSI 240
AHA1            PGQEVFSGSTCKQGEIEAVVIATGVHTFFGKAAHLVDSTNQVGHFQKVLTSIGNFCICSI 240
                ****************************************************:*********

3B8C            AIGMVIEIIVMYPIQRRKYRDGIDNLLVLLIGGIPIAMPTVLSVTMAIGSHRLSQQGAIT 300
AHA1            AIGIAIEIVVMYPIQHRKYRDGIDNLLVLLIGGIPIAMPTVLSVTMAIGSHRLSQQGAIT 300
                ***:.***:******:*********************************************

3B8C            KRMTAIEEMAGMDVLCSDKTGTLTLNKLSVDKNLVEVFCKGVEKDQVLLFAAMASRVENQ 360
AHA1            KRMTAIEEMAGMDVLCSDKTGTLTLNKLSVDKNLVEVFCKGVEKDQVLLFAAMASRVENQ 360
                ************************************************************

3B8C            DAIDAAMVGMLADPKEARAGIREVHFLPFNPVDKRTALTYIDGSGNWHRVSKGAPEQILE 420
AHA1            DAIDAAMVGMLADPKEARAGIREVHFLPFNPVDKRTALTYIDSDGNWHRVSKGAPEQILD 420
                *********************************************..***************:

3B8C            LAKASNDLSKKVLSIIDKYAERGLRSLAVARQVVPEKTKESPGAPWEFVGLLPLFDPPRH 480
AHA1            LANARPDLRKKVLSCIDKYAERGLRSLAVARQVVPEKTKESPGGPWEFVGLLPLFDPPRH 480
                **:*  ** ***** ******************************.****************

3B8C            DSAETIRRALNLGVNVKMITGDQLAIGKETGRRLGMGTNMYPSSALLGTHKDANLASIPV 540
AHA1            DSAETIRRALNLGVNVKMITGDQLAIGKETGRRLGMGTNMYPSAALLGTDKDSNIASIPV 540
                ***********************************************:*****.**:*:*****

3B8C            EELIEKADGFAGVFPEHKYEIVKKLQERKHIVGMTGDGVNDAPALKKADIGIAVADATDA 600
AHA1            EELIEKADGFAGVFPEHKYEIVKKLQERKHIVGMTGDGVNDAPALKKADIGIAVADATDA 600
                ************************************************************

3B8C            ARGASDIVLTEPGLSVIISAVLTSRAIFQRMKNYTIYAVSITIRIVFGFMLIALIWEFDF 660
AHA1            ARGASDIVLTEPGLSVIISAVLTSRAIFQRMKNYTIYAVSITIRIVFGFMLIALIWEFDF 660
                ************************************************************

3B8C            SAFMVLIIAILNDGTIMTISKDRVKPSPTPDSWKLKEIFATGVVLGGYQAIMTVIFFWAA 720
AHA1            SAFMVLIIAILNDGTIMTISKDRVKPSPTPDSWKLKEIFATGIVLGGYQAIMSVIFFWAA 720
                *****************************************:**********:*******

3B8C            HKTDFFSDTFGVRSIRDNNHELMGAVYLQVSIISQALIFVTRSRSWSFVERPGALLMIAF 780
AHA1            HKTDFFSDKFGVRSIRDNNDELMGAVYLQVSIISQALIFVTRSRSWSFVERPGALLMIAF 780
                ********.***********.****************************************

3B8C            LIAQLIATLIAVYANWEFAKIRGIGWGWAGVIWLYSIVTYFPLDVFKFAIRYI 833
AHA1            VIAQLVATLIAVYADWTFAKVKGIGWGWAGVIWIYSIVTYFPQDILKFAIRYI 833
                :****:********:* ***::************:******** *::*******
```

Figure 1: Sequence alignment of AHA1 from *Arabidopsis thaliana* with 3b8c from *Arabidopsis thaliana* done using CLUSTALW server that was subsequently submitted to MODELLER. The conserved regions are indicated by '*'.

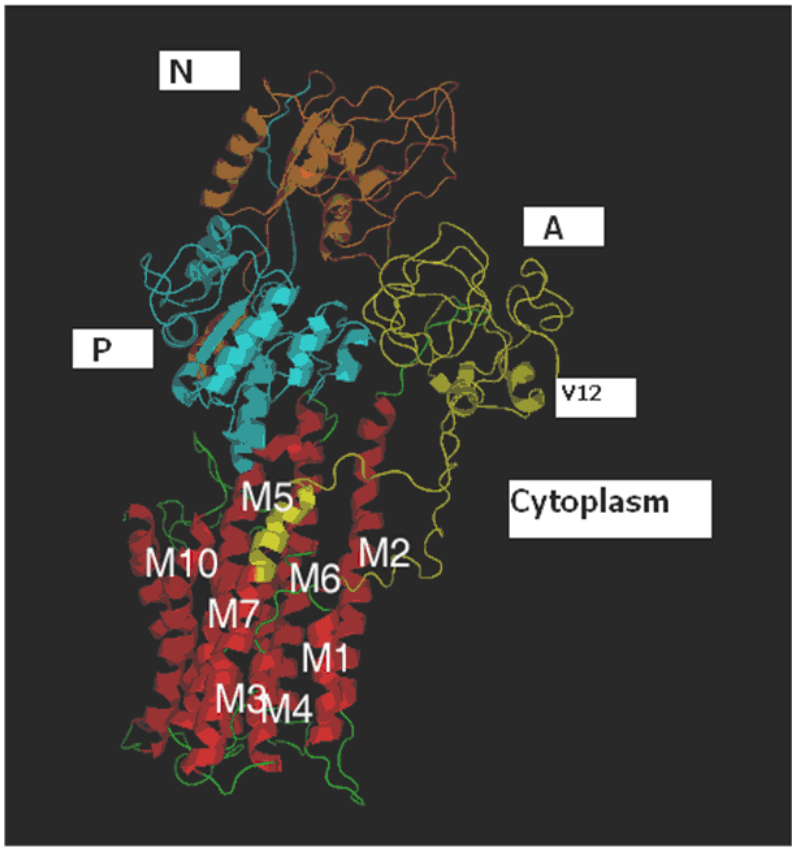

Figure 2: Ribbon representation of modeled AHA1. The α-helices and β -sheets are shown as helices and ribbons, respectively. The rest are shown as loops. The figure was prepared by PyMol.

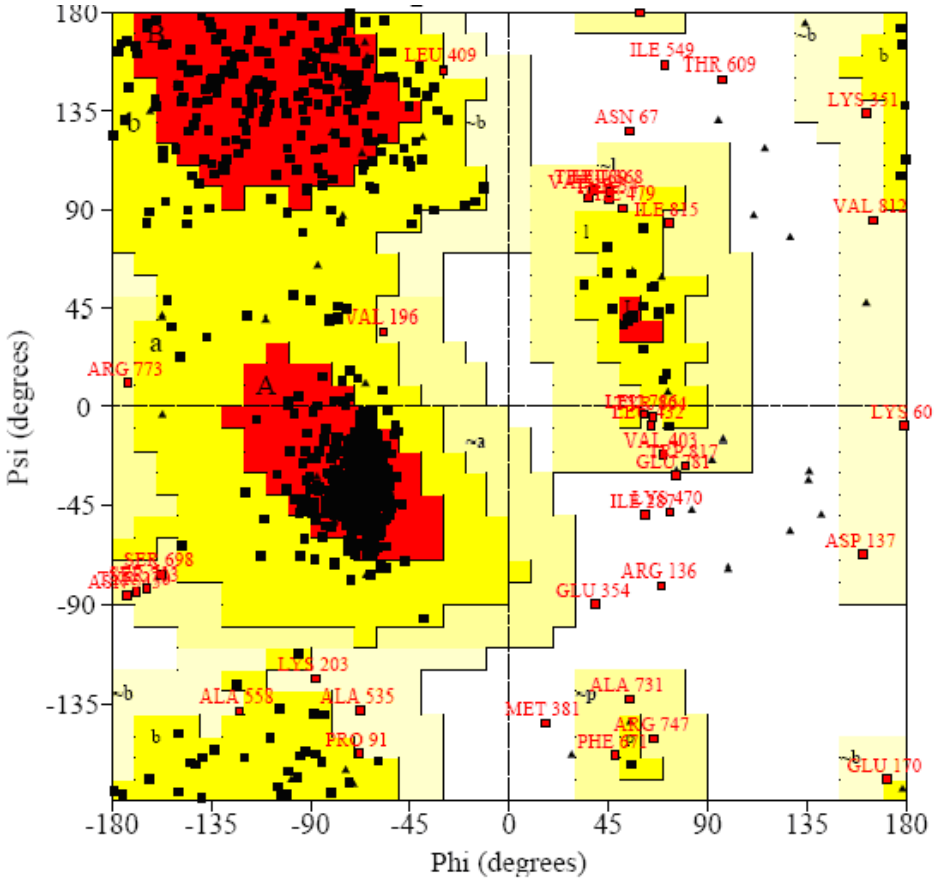

Figure 3: Ramachandran's Map of AHA1 protein. The plot calculation on 3D model of AHA1 protein was calculated with the PROCHECK program.

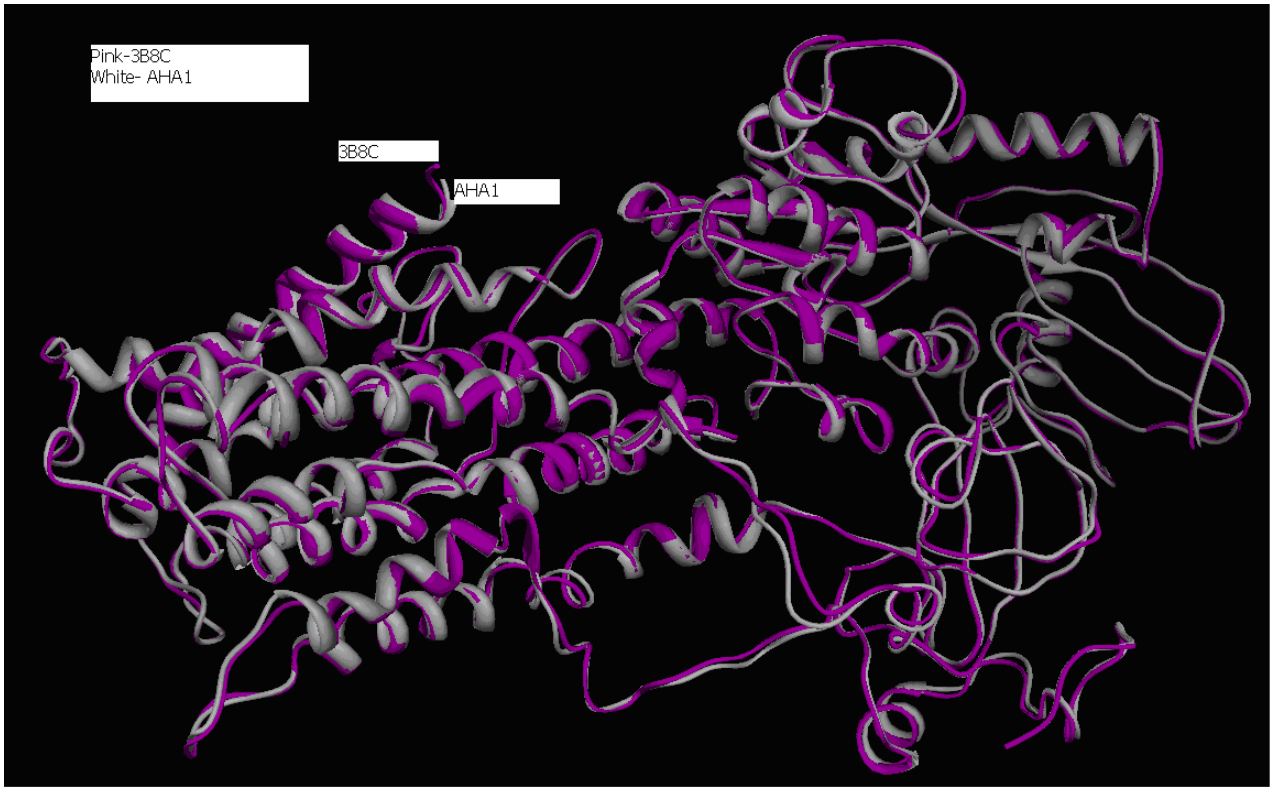

Figure 4: Superimposition of C^{α} trace of AHA1 (represented in white) and 3b8c (represented in pink color).

tion of Asp, which protrudes as a pole out of the membrane, and it is further connected to the M1 and M3 transmembrane segments by extended loops. In the transmembrane domain, the M1 helix shows a prominent 90° kink imposed by a proline residue, Pro^{68}, which is conserved in type III P-type ATPases.

Two ion-binding sites are present in the SR Ca^{2+}-ATPase structure (Toyoshima C et al 2000) but the structure of AHA2 H^{+}-ATPase suggests the presence of only one proton-binding site. The structure of AHA1 also suggests the presence of only one proton-binding site. The residue supposed to contribute to the proton-binding site in AHA1is Asp^{684}. AHA1 seems to be strictly conserved for the corresponding amino acid residues in the AHA2. In fact, this Asp residue appears to be conserved among the P-type ATPases (http://www.patbase.kvl.dk) across the species as it aligns with Asp^{730} of PMA1, Asp^{800} of sercala and Asp^{808} of Na/K.

The side chain of Asp^{684} AHA1 (TM6), as well as the backbone carbonyls of the amino acid residues in the transmembrane helix4 (Ile282, Gly283 and Ile 285) could provide the liganding groups for a high-affinity hydronium ion coordination site in the AHA1 similar to that in AHA2. The mutagenesis experiments support for such a model; it has been observed that after removal of the potential protonation group (D684N, D684A, D684V and D684R) of the Asp684AHA2 side chain there occurs no change in E_1P-E_2P enzyme conformational and proton transport although the ATP hydrolysis continues (Buch-Pedersen et al., 2000; Buch-Pedersen et al., 2003). The mutagenesis experiment also led to the suggestion that the two oxygen of the Asp^{684}AHA1 side chain that participate in the H^{+}- transporting mechanism can be discriminated with respect to their function . This is because the Asp^{684} can be replaced with asparagines with apparently little effect upon initial binding, and removing both the side chains oxygen completely lowers the rate of initial proton binding markedly. Thus, the side chain of the amino acid at position 684 has been ascribed a role in initial proton binding. As reported for AHA2, the Asp^{684}AHA1 side-chain oxygen seem also to influence the rate of initial proton binding to the E_1 proton-binding site.

Effect of Mutations on Molecular Surface Area and Solvent Accessible Surface Area in AHA1 and AHA2 Protein Structures

Molecular surface area (MSA) and solvent accessible surface area (ASA) were calculated for AHA1 and AHA2 using surface racer program (http://apps.phar.umich.edu/tsodikovlab/index_files/Page756.htm). Table 3 shows the changes in MSA and ASA, both polar and charged, from

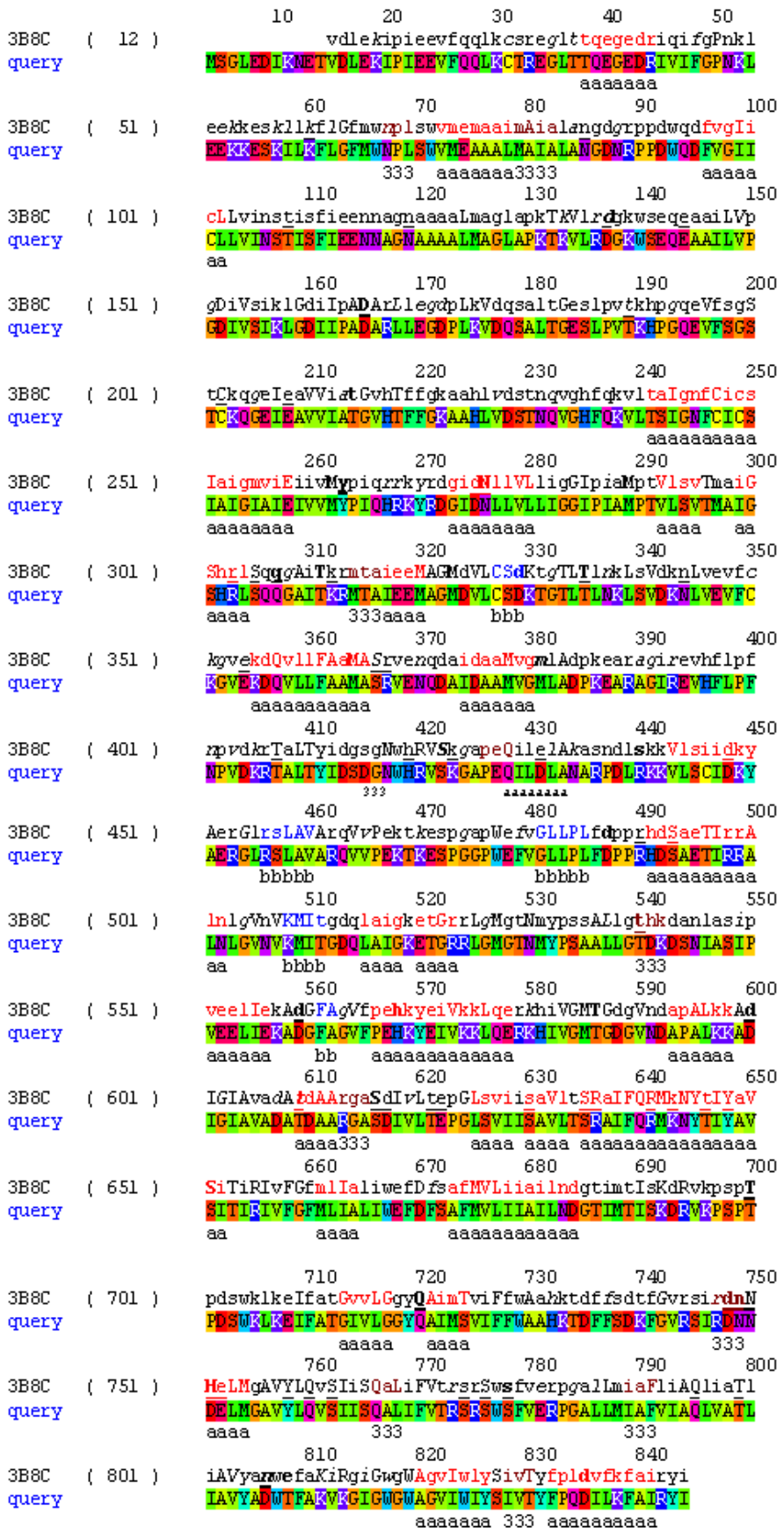

Figure 5: Structure based sequence alignment of template and final structures of the AHA1 using JOY program. The key to the JOY annotation is as follows: lowercase red letter, α-helix; lowercase blue letter, β-strand; lowercase maroon letter, 3_{10}-helix; uppercase letter, solvent-inaccessible residue; lowercase letter, solvent-accessible residue; italic lowercase letter, positive Φ.

	3b8c	AHA1	Change in Area $(A^o)^2$
Total ASA	41012.88	38200.19	2812.69
Total MSA	41216.29	37687.02	3529.27

Area in units: $(A^o)^2$

Table 3: Calculated molecular surface area (MSA) and solvent accessible surface area (ASA).

AHA2 to AHA1 structures due to amino acid mutations. There is a decrease in MSA and ASA from AHA2 (3b8c) to AHA1 due to mutations . However, whether the decrease in the MSA and ASA values could be related to the function of a protein is yet to be addressed. Mutational analysis of the protein could probably provide a suitable answer.

Conclusion

Thus, in this study we built a complete atomic model of *Arabidopsis thaliana* AHA1 plasma membrane H^+-ATPase by homology modeling using the 3D structure of AHA2 from *Arabidopsis thaliana* as a template. The transmembrane topology of the protein was also predicted. The model suggested that a description of the proton translocation pathway through the protein (enzyme) must include a conserved Asp in M6 corresponding to Asp^{684} similar to that in AHA2. This residue is conserved in the P-type H^+ATPase across the species. The model derived matched to AHA2 in all respect, suggesting that the two proteins should not be very much different in its activity or function despite the difference in their tissue-specific expression, which of course in not the case. The difference in the activity of the two isoforms, nevertheless, could be associated with the changing MSA and ASA or possibly other parameters because of change in the amino acids. Further studies may provide conclusive information in this regard, particularly considering the region which was not included in the modeling.

Acknowledgements

The study was supported by Department of Biotechnology grant no.BT/BI/04/058/2002 to Distributed Information Sub Center, Institute of Life Sciences, Bhubaneswar, India. Funds contributed by ILS, Bhubaneswar, to BPS and Senior Research Fellowship to BBS awarded by UGC, New Delhi are kindly acknowledged.

References

1. Albers RW (1967) Biochemical aspects of active transport. Annu Rev Biochem 36: 727-756. » CrossRef » Pubmed » Google Scholar
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) A basic local alignment search tool. J Mol Biol 215: 403-410. » CrossRef » Pubmed » Google Scholar
3. Apell HJ (2004) How do P-type ATPases transport ions? Bioelectrochemistry 63: 149-156. » CrossRef » Pubmed » Google Scholar
4. Arango M, Gevaudant F, Oufattole M, Boutry M (2003) The plasma membrane proton pump ATPase: the significance of gene subfamilies. Planta 216: 355-365. » CrossRef » Pubmed » Google Scholar
5. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000). The Protein Data Bank. Nucleic Acids Research 28: 235-242. » CrossRef » Pubmed » Google Scholar
6. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, et al. (1993) CHARMM: A program for macromolecular energy minimization and dynamics calculations. J Comp Chem 4: 187-217. » CrossRef » Google Scholar
7. Buch-Pedersen MJ, Palmgren MG (2003) Conserved Asp684 in transmembrane segment M6 of the plant plasma membrane H^+-ATPase is a molecular determinant for proton translocation. J Biol Chem 278: 17845-17851. » CrossRef » Pubmed » Google Scholar
8. Buch-Pedersen MJ, Venema K, Serrano R, Palmgren M (2000) Abolishment of proton pumping and accumulation in the E1P conformational state of a plant plasma membrane H^+-ATPase by substitution of a conserved aspartyl residue in transmembrane segment 6. J Biol Chem 275: 39167-39173. » CrossRef » Pubmed » Google Scholar
9. Cserzo M, Wallin E, Simon I, von Heijne G, Elofsson A (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng 10: 673-676. » CrossRef » Pubmed » Google Scholar
10. DeWitt ND, Harper JF, Sussman MR (1991) Evidence for a plasma membrane proton pump in phloem cells of higher plants. Plant J 1: 121-128. » Pubmed » Google Scholar
11. Dinneny JR, Long TA, Wang JY, Jung JW, Mace D, et al. (2008) Cell Identity Mediates the Response of Arabidopsis Roots to Abiotic Stress. Science 320: 942-945. » CrossRef » Pubmed » Google Scholar
12. Hentzen AE, Smart LB, Whimmers LE, Fang HH, Schroeder JI, et al. (1996) Two Plasma Membrane H^+-ATPase Genes Expressed in Guard Cells of *Vicia faba* Are Also Expressed Throughout the Plant. Plant Cell Physiol 37: 650-659. » CrossRef » Pubmed » Google Scholar

13. Hofmann K, Stoffel W (1993) TMbase - A database of membrane spanning protein segments. Biol Chem Hoppe-Seyler 374: 166-171. » CrossRef » Pubmed » Google Scholar

14. Houlné G, Boutry M (1994) Identification of an *Arabidopsis thaliana* gene encoding a plasma membrane H^+-ATPase whose expression is restricted to anther tissues. The Plant Journal 5: 311-317. » CrossRef » Pubmed » Google Scholar

15. Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K, et al. (2003) Collection, Mapping, and Annotation of Over 28,000 cDNA Clones from japonica Rice. Science 301: 376-379. » CrossRef » Pubmed » Google Scholar

16. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305: 567-580. » CrossRef » Pubmed » Google Scholar

17. Kühlbrandt W (2004) Biology, structure and mechanism of P-type ATPases. Nat Rev Mol Cell Biol 5: 282-295. » CrossRef » Pubmed » Google Scholar

18. Laskoswki RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26: 283-291. » CrossRef » Google Scholar

19. Lefebvre B, Arango M, Oufattole M, Crouzet J, Purnelle B, et al. (2005) Identification of a *Nicotiana plumbaginifolia* Plasma Membrane H^+-ATPase Gene Expressed in the Pollen Tube. Plant Molecular Biology 58: 775-787. » CrossRef » Pubmed » Google Scholar

20. Luthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356: 83-85. » CrossRef » Pubmed » Google Scholar

21. Mizuguchi K, Deane CM, Blundell TL, Johnson MS, Overington JP (1998) JOY: protein sequence-structure representation and analysis. Bioinformatics 14: 617-623. » CrossRef » Pubmed » Google Scholar

22. Møller JV, Juul B, Maire M (1996) Structural organization, ion transport, and energy transduction of P-type ATPases, Biochim Biophys Acta 1286: 1-51. » CrossRef » Pubmed » Google Scholar

23. Moriau L, Bogaerts P, Jonniaux JL, Boutry M (1993) Identification and characterization of a second plasma membrane H^+-ATPase gene subfamily in *Nicotiana plumbaginifolia*. Plant Mol Biol 21: 955-963. » CrossRef » Pubmed » Google Scholar

24. Okazaki (2002) Y Blue light inactivates plasma membrane H (+)-ATPase in pulvinar motor cells of *Phaseolus vulgaris* L. Plant Cell Physiol 43: 860-868. » CrossRef » Pubmed » Google Scholar

25. Ookura T, Wada M, Sakakibara Y, Jeong KH, Maruta I, et al. (1994) Identification and Characterization of a Family of Genes for the Plasma Membrane H^+-ATPase of *Oryza sativa* L. Plant Cell Physiol 35: 1251-1256. » CrossRef » Pubmed » Google Scholar

26. Oufattole M, Arango M, Boutry M (2000) Identification and expression of three new *Nicotiana plumbaginifolia* genes which encode isoforms of a plasma-membrane H (+)-ATPase, and one of which is induced by mechanical stress. Planta 210: 715-722. » CrossRef » Pubmed » Google Scholar

27. Palmgren M, Harper J (1999) Pumping with plant P-type ATPases. J Exp Bot 50: 883-893. » CrossRef » Google Scholar

28. Palmgren MG (2001) Plant plasma membrane H^+-ATPases: Powerhouses for Nutrient Uptake. Annu Rev Plant Physiol Plant Mol Biol 52: 817-845. » CrossRef » Pubmed » Google Scholar

29. Pedersen BP, Buch-Pedersen MJ, Preben MJ, Palmgren MG, Nissen P (2007) Crystal structure of the plasma membrane proton pump. Nature 450: 1111-1114. » CrossRef » Pubmed » Google Scholar

30. Pedersen BP, Buch-Pedersen MJ, Morth JP, Palmgren MG, Nissen P (2007) Crystal structure of the plasma membrane proton pump, Nature 450: 1111-1114. » CrossRef » Pubmed » Google Scholar

31. Perez C, Michelet B, Ferrant V, Bogaerts P, Boutry M (1992) Differential expression within a three-gene subfamily encoding a plasma membrane H (+)-ATPase in *Nicotiana plumbaginifolia*. J Biol Chem 267: 1204-1211. » CrossRef » Pubmed » Google Scholar

32. Post RL, Kume S, Tobin T, Orcutt B, Sen AK (1969) Flexibility of an active center in sodiumplus - potassium adenosine triphosphatase. J Gen Physiol 54: 306-326. » CrossRef » Google Scholar

33. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restrains. J Mol Biol 234: 779-815. » CrossRef » Pubmed » Google Scholar

34. Sali A, Matsumoto R, McNeil HP, Karplus M, Stevens RL (1993) Three-dimensional models of four mouse mast cell chymases. Identification of proteoglycan binding regions and protease-specific antigenic epitopes. J Biol Chem 268: 9023-34. » CrossRef » Pubmed » Google Scholar

35. Sali A, Overington JP (1994) Derivation of rules for comparative protein modeling from a database of pro-

tein structure alignments. Protein Sci 3: 1582-96. » CrossRef » Pubmed » Google Scholar

36. Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M (1995) Evaluation of comparative protein modeling by MODELLER. Proteins 23: 318-26. » CrossRef » Pubmed » Google Scholar

37. Shindyalov IN, Bourne PE (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Engineer 11: 739-747. » CrossRef » Pubmed » Google Scholar

38. Sørensen TL, Møller JV, Nissen P (2004) Phosphoryl transfer and calcium ion occlusion in the calcium pump. Science 304: 1672-1675. » CrossRef » Pubmed » Google Scholar

39. Sze H (1985) H^+-translocating ATPase: Advances using membrane vesicles. Annu Rev Plant Physiol 36: 175-208. » CrossRef » Google Scholar

40. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680. » CrossRef » Pubmed » Google Scholar

41. Toyoshima C, Nakasako M, Nomura H, Ogawa H (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolutions. Nature 405: 647-655. » CrossRef » Pubmed » Google Scholar

42. Toyoshima C, Nakasako M, Nomura H, Ogawa H (2000) Crystal structure of the Calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. Nature 405: 647-655. » CrossRef » Pubmed » Google Scholar

43. Toyoshima C, Nomura H (2002) Structural changes in the calcium pump accompanying the dissociation of calcium. Nature 418: 605-611. » CrossRef » Pubmed » Google Scholar

44. Tusnady GE, Simon I (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J Mol Biol 283: 489-506. » CrossRef » Pubmed » Google Scholar

45. Wada M, Takano M, Kasamo K (1992) Nucleotide Sequence of a Complementary DNA Encoding Plasma Membrane H^+-ATPase from Rice (*Oryza sativa* L.). Plant Physiol 99: 794-795. » CrossRef » Pubmed » Google Scholar