

Regular Article

# Homology modeling and *in silico* characterization of synaptotagmin 1 (SYT1) protein from *Arabidopsis thaliana* (L.) Heynh.

Ashutosh Mukherjee

Department of Botany, Dinabandhu Mahavidyalaya, Bongaon, North 24 Parganas - 743235, West Bengal, India

Synaptotagmins are a group of C2 domain containing proteins which play important roles in vesicle trafficking in synaptic vesicles of animals. With the advent of plant genome sequencing, many synaptotagmins are also discovered from plants and they have been found to perform many crucial physiological roles in plants. The model plant *Arabidopsis thaliana* (L.) Heynh. contains five synaptotagmins of which SYT1 is responsible for many important functions including maintenance of plasma membrane integrity and cell viability. In this study, the three dimensional structure of this protein has been developed *in silico* by homology modeling method to collect knowledge about its structure-function relationship. Additionally, the interaction of calcium ion with this protein was studied. The structures of C2 domains are somewhat different from the known animal synaptotagmins. The protein contains many phosphorylation sites which indicate that SYT1 is part of one or more signaling cascades.

**Keywords:** C2-domain, disorder, flexibility, motif, ligand, topology

## Introduction

Synaptotagmins are a group of membrane trafficking proteins containing an N-terminal transmembrane region (TMR), a linker of variable length and two tandemly arranged C-terminal C2 domains (Craxton, 2004), called C2A and C2B, respectively. C2 domain is a calcium ion-binding domain; about 130-145 amino acids long and are found in many membrane-associated signaling proteins in a large number of organisms (Nalefski and Falke, 1996). Synaptotagmins are calcium sensors and control the exocytosis and endocytosis of synaptic vesicles in animals (Lewis and Lazarowitz, 2010). Although these were previously considered to be exclusive to animals (Lewis and Lazarowitz, 2010), the comprehensive public sequence database resources show that they are also present in

plants. From the sequenced plant genomes, many synaptotagmin genes have been identified by several computational procedures (Craxton, 2004).

The model plant *Arabidopsis thaliana* (L.) Heynh. contains 5 synaptotagmin genes in its genome. Of these, the gene AT2G20990 was named as SYT1 (synaptotagmin 1, also called SYTA) as this gene codes for a protein showing similar structure to animal synaptotagmin (Craxton, 2004). This gene has three transcripts as shown by the TAIR (The Arabidopsis Information Resource) database (<http://www.arabidopsis.org/index.jsp>). However, the gene model AT2G20990.1 encodes a 541 amino acid long protein whose function in many physiological processes has been vastly studied (Yamazaki *et al.*, 2008; Schapire *et*

*al.*, 2008; Yamazaki *et al.*, 2010). This is a plasma membrane localized protein, expressed in all tissues and found to play many vital physiological roles including the maintenance of plasma membrane integrity and cell viability (Schapire *et al.*, 2008), calcium dependent freezing tolerance (Yamazaki *et al.*, 2008) and regulation of endocytosis and inhibition of the cell-to-cell movement of viral particles (Lewis and Lazarowitz, 2010). It has also been shown that loss of function mutation of SYT1 causes hypersensitivity to NaCl which causes electrolyte leakage from the plasma membrane (Schapire *et al.*, 2008). All these studies suggested that SYT1 performs crucial roles in maintaining plasma membrane integrity and helps to overcome stress situations.

Biological function of a protein is the manifestation of its tertiary structure and thus, information of the structure of the protein is a prerequisite for understanding its detailed functional aspects (Paital *et al.*, 2011). As no 3D structure of SYT1 is known, in the present investigation, effort was made to generate a 3D structure of the SYT1 protein by homology modeling. Additionally, its physicochemical, structural and functional aspects were analyzed with the help of several bioinformatics tools.

### Materials and Methods

The *Arabidopsis thaliana* (L.) Heynh. SYT1 protein sequence (GenBank accession NP\_565495; GI: 18399541) was downloaded from the NCBI Refseq (Pruitt *et al.*, 2007) database (<http://www.ncbi.nlm.nih.gov/projects/RefSeq/>). Three dimensional structure of this protein was not yet available in the Protein Data bank *i.e.* PDB (Berman *et al.*, 2000). This sequence was further utilized for further characterization and structure prediction.

### Physico-chemical analysis

The computations of various physical and chemical parameters, such as molecular weight, amino acid composition,

isoelectric point (pI), total number of negatively and positively charged residues, extinction coefficient, instability index, aliphatic index and Grand Average of Hydropathy (GRAVY), were done using ProtParam tool (Gasteiger *et al.*, 2005) available at <http://us.expasy.org/tools/protparam.html>.

### Structural and functional characterization

Secondary structure prediction was carried out with SOPMA (Geourjon and Deléage, 1995) server ([http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_sopma.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)). The transmembrane regions were predicted with TMHMM (Krogh *et al.*, 2001) server 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). The output of TMHMM was displayed using the TOPO2 (Johns, 2005) transmembrane protein display server ([www.sacs.ucsf.edu/TOPO2/](http://www.sacs.ucsf.edu/TOPO2/)). Subcellular localization was predicted using TargetP (Emanuelsson *et al.*, 2000) 1.1 server (<http://www.cbs.dtu.dk/services/TargetP/abstract.php>). Protein disorder was predicted using Disopred (Ward *et al.*, 2004) (<http://bioinf.cs.ucl.ac.uk/disopred/>) server. Motif Scan (Pagni *et al.*, 2007; Sigrist *et al.*, 2010) server ([http://myhits.isb-sib.ch/cgi-bin/motif\\_scan](http://myhits.isb-sib.ch/cgi-bin/motif_scan)) was also scanned with the default parameters for the presence of any signature domain.

### Homology modeling and model evaluation

HHPred (Söding *et al.*, 2005), BLASTP (Altschul *et al.*, 1990) and PSI-BLAST (Altschul *et al.*, 1997) were employed to find suitable templates for SYT1. I-TASSER (Zhang, 2007; Roy *et al.*, 2010), the iterative threading assembly refinement server (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>), was chosen to generate the homology model because it is automated and easy to use, its algorithm incorporates multiple templates and it has a high degree of accuracy based on blind CASP experiments (Roy *et al.*, 2010). Rather than specifying one single template for homology modeling, I-

TASSER was allowed to incorporate multiple templates since it is recommended that multiple templates should be used in order to avoid biasing the model toward one protein or one set of side chain conformations (Ginalski, 2006; Rhodes, 2006). I-TASSER generated five predicted structures for the protein of with the model with highest C-score was chosen for further analysis. After modeling, the validation of the modelled structure was also carried out using Protein Structure Validation Suite (PSVS) tool (Bhattacharya *et al.*, 2007) available at [http://psvs-1\\_4-dev.nesg.org/](http://psvs-1_4-dev.nesg.org/). Within PSVS, the models were analyzed by PROCHECK (Laskowski *et al.*, 1993), RAMPAGE and Molprobit (Lovell *et al.*, 2003). 3D structures of the proteins and protein-calcium complex were visualized with Chimera (Pettersen *et al.*, 2004). For an

at-a-glance overview of the modeled protein, PDBsum (Laskowski, 2009) web server (<http://www.ebi.ac.uk/pdbsum/>) was used. To analyze protein flexibility, Flexserv (Camps *et al.*, 2009) server (<http://mmb.pcb.ub.es/FlexServ/>) was employed. To predict the interacting surfaces of SYT1 protein, POLYVIEW-3D (Porollo and Meller, 2007) server (<http://polyview.cchmc.org/polyview3d.html>) was employed.

## Results and Discussion

### Physico-chemical and structural features

The physico-chemical properties of the SYT1 from *Arabidopsis* were predicted using ProtParam (<http://expasy.org/cgi-bin/protparam>) and the results are shown in table 1.

**Table 1: ProtParam table showing different physicochemical properties of SYT1 protein**

Parameters	Value	Explanation
pI	7.18	Indicates that the protein is slightly basic.
Total number of negatively charged residues (Asp + Glu)	73	Total number of negatively charged residues is equal to the total number of positively charged residues. This indicates that the protein have extracellular portions.
Total number of positively charged residues (Arg + Lys)	73	
The instability index (II)	34.24	This classifies the protein as stable.
Aliphatic index	91.98	Indicates that this globular protein is thermostable.
Grand average of hydropathicity (GRAVY)	-0.256	A negative GRAVY score indicates that the protein is hydrophilic.

The most frequent amino acids present in the sequence were found to be both leucine and lysine (52 residues each, 9.6%) and the least frequent was cystine (4 residues, 0.7%). Leucine has strong helical-forming power (Chou and Fasman, 1973). Thus, leucine may help the formation of helices in this protein. Lysine, on the other hand, can be modified by acetylation, methylation, ubiquitination, sumoylation, neddylation, biotinylation, pupylation and carboxylation which have a propensity to modify protein function (Sadoul *et al.*, 2008). Thus, in SYT1 of *A. thaliana*, leucine and lysine appear to contribute in the structural and functional aspects, respectively. Cysteines are over-represented in nonflexible regions of the

proteins to form disulfide bridges (Schlessinger and Rost, 2005). As the SYT1 protein has only four cysteines, it appears to have flexible regions. The total number of negatively charged residues (Asp + Glu) and the total number of positively charged residues (Arg + Lys) were equal (73). This indicates that the protein has an extracellular portion because intracellular proteins have higher fraction of negatively charged residues (Cedano *et al.*, 1997). The intended isoelectric point (pI) is helpful for the fact that at isoelectric point, there is least solubility and the mobility in an electric field is zero. Isoelectric point (pI) is that pH at which the surface of a protein is covered with charge although net charge of

the protein is zero. The calculated isoelectric point (pI) was computed to be 7.18 which indicate that the protein is slightly basic. Schwartz *et al.* (2001) reported that integral membrane proteins usually have high pI values. The slightly basic nature of this protein, thus, indicates that the protein is membrane attached. It also appeared that lysine (which is one of the two most frequent amino acids in this protein) contributes the basic nature of SYT1 of this study. The high aliphatic index (91.98) indicates that this protein is stable for a broad range of temperature range. The instability index (34.24) also provides the evidence that the protein is stable. The Grand Average of Hydropathicity (GRAVY) value was negative (-0.256) which indicates better interaction of this protein with the surrounding water molecules.

Table 2 presents the results of secondary structure prediction analysis by SOPMA from which it is clear that random coil is predominant (42.33%), followed by alpha helix (31.24%) and extended strand (23.48%). SOPMA also predicted the presence of Beta turn (2.96%).

**Table 2: Secondary structure of SYT1 as predicted by SOPMA**

Parameters	Number of amino acids	Percentage of amino acids
Alpha helix (Hh)	169	31.24
3 <sub>10</sub> helix (Gg)	0	0.00
Pi helix (Ii)	0	0.00
Beta bridge (Bb)	0	0.00
Extended strand (Ee)	127	23.48
Beta turn (Tt)	16	2.96
Bend region (Ss)	0	0.00
Random coil (Cc)	229	42.33
Ambiguous states	0	0.00
Other states	0	0.00

Presence of random coils or more specifically coils adds flexibility to the protein structure (Buxbaum, 2007). Flexibility is very important for protein

function and the changes in protein structure can be remarkable during interactions with its binding partners (Teillum *et al.*, 2009). Additionally, Sharmin *et al.* (2011) reported that random coils have some relationships with dehydration response. Qu *et al.* (1998) reported that random coils are contracted by osmolytes which leads to decrease of conformational entropy and subsequent stabilization of protein folding. Considering all these, it appears that random coil regions of SYT1 protein from *A. thaliana* play a pivotal role in interaction with other proteins which is essential to combat stress situations.

### Domains and motifs

The Motif scan tool identified several phosphorylation sites (table 3). These include one cAMP - and cGMP-dependent protein kinase phosphorylation site, six Caesin kinase II phosphorylation sites, six N-myristoylation sites, six Protein kinase C phosphorylation sites and one tyrosine kinase phosphorylation site. Previously, Hilfiker *et al.* (1999) showed in animal systems that synaptotagmins were phosphorylated by Ca<sup>2+</sup>/calmodulin-dependent protein kinase II (CaMKII), protein kinase C (PKC), and casein kinase II. They also suggested that the regulation of neurotransmitter release in animals was possibly contributed by the phosphorylation of synaptotagmin I by CaMKII and PKC. The presence of many phosphorylation sites in SYT1 of present study indicates that SYT1 is heavily regulated.

N-terminal myristoylation plays a crucial function in signal transduction and membrane targeting in plants in responses to environmental stress conditions (Podell and Gribskov, 2004). These myristoylation events influence the conformational stability of a protein as well as their interacting capability with membranes or the hydrophobic domains of other proteins (Zheng *et al.*, 1993; Resh 1999; Olsen and Kaarsholm, 2000). If an attached myristic acid (C<sub>14</sub> saturated fatty acid) is placed on protein's surface, it can tether the modified

protein to any membrane system of the cell which ultimately increases the chance to interact with other proteins (Podell and Gribskov, 2004). Two myristoylation sites were positioned in tandem within the transmembrane region (Gly14-Leu19 and Gly20-Tyr25) which indicates its role in the interaction with the membrane. All other myristoylation sites form some part of the surface of the SYT1 protein indicating that they help to interact with other proteins. Previously, Kang *et al.* (2004) showed the presence of palmitoylation sites in animal synaptotagmin I and palmitoylation *i.e.* attachment of palmitic acid ( $C_{16}$  saturated fatty acid) is important to sort synaptotagmin I to the intracellular vesicular pool at presynaptic contacts. From these observations, it is clear that protein modification involving addition of hydrophobic groups (myristic acid and palmitic acid) is important for proper functioning of synaptotagmin I. It is also noteworthy that the cAMP - and cGMP-dependent protein kinase phosphorylation

site, the tyrosine kinase phosphorylation site as well as two N-myristoylation sites and four Protein kinase C phosphorylation sites contain lysine residues indicating the role of lysine in SYT1 protein function.

### Protein disorder

The total intrinsic disorder as detected by PrDOS shows total disordered amino acid residues were 23 (4.25%). However, they were spread over the protein in 3 regions (figure 1). The longest disordered region was spread from Met398 to Glu413. However, the C2 domains were not disordered. This shows that the C2 domains are structured which may be important for binding with calcium ion. Disordered regions differ considerably from the well-structured regions of a protein (Wright and Dyson, 1999). These disordered regions play significant roles in protein interaction (Patil and Nakamura, 2006). From these results, it seems that these disordered regions interact with other proteins with novel properties.

Table 3: Motifs identified with Motif Scan

Motif information	No. of sites	Amino acid residues
cAMP- and cGMP-dependent protein kinase phosphorylation site	1	297-300
Casein kinase II phosphorylation site	6	43-46, 198-201, 317-320, 329-332, 394-397, 428-431
N-myristoylation site	6	14-19, 20-25, 162-167, 208-213, 370-375, 414-419
Protein kinase C phosphorylation site	6	92-94, 181-183, 296-298, 300-302, 317-319, 489-491
Tyrosine kinase phosphorylation site	1	383-390

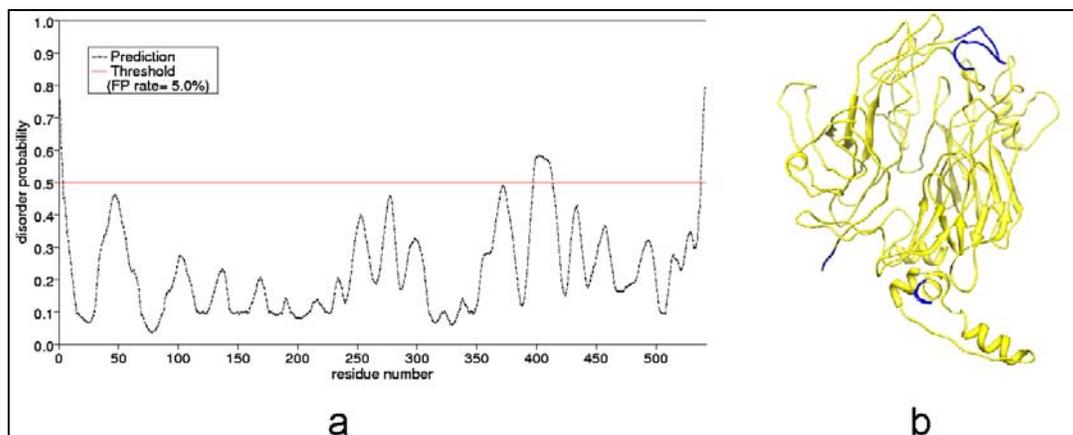
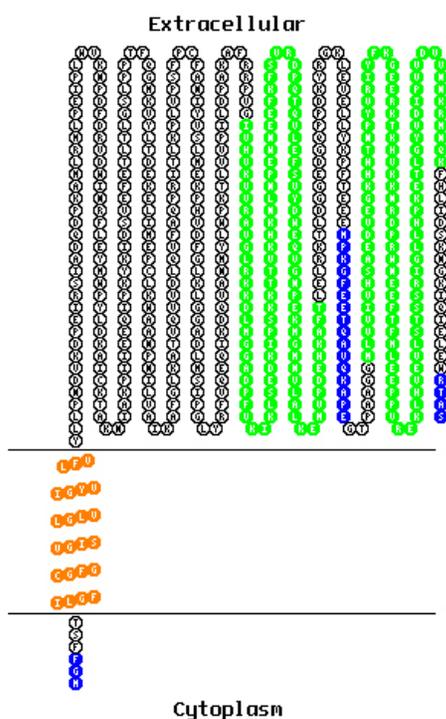


Figure 1: Disordered regions of the SYT1 protein. a. schematic diagram of disorder over the entire sequence; b. Disordered portions of the protein in the three dimensional structure. Disordered regions are shown in blue.

### Localization of the protein

TargetP score (0.989) was the highest for SP (a signal peptide) and the Loc (prediction of localization) indicated that the protein is guided through a secretory pathway to the membrane. TargetP also predicted that the protein has a 24 amino acid long signal peptide. TMHMM predicted that the protein has a transmembrane domain (residue 7-29). The pI value (table 1) also indicates the protein is membrane localized. All these indicate that the protein is localized in the plasma membrane by a short transmembrane domain. The topology of the protein as generated by TOPO2 server is shown in figure 2.



**Figure 2: Topological structure of SYT1 as drawn by TOPO2 tool. Transmembrane region is shown in orange, C2 domains are shown in green and disordered regions are shown in blue.**

### Homology modeling and model evaluation

Preliminary template search with HHpred and BLAST (both BLASTP and PHI-BLAST) against PDB database resulted several templates aligned with the C-terminal

region of SYT1, but not the N-terminal region. Thus, I-TASSER tool was employed for model generation as it incorporates multiple templates for model generation. I-TASSER generated several models almost spanning the whole length of the protein as it used multiple templates.

The models generated by I-TASSER were considered for further quality assessment studies. Among the several models generated by I-TASSER, 'Model 1' possessed the highest C-score of -1.54 and thus, was selected for further analyses. C-score is a confidence score provided by the I-TASSER server which estimates the quality of the predicted structure. C-score is defined based on the quality of the threading alignments and the convergence of the I-TASSER's structural assembly refinement simulations. The model showed TM-scores (template modeling scores) ranged from 0.500-0.833, indicating largely correct topology for the modeled structure. The normalized Z-score for each alignment (between the target and a given template) indicates the significance of the alignment compared to the average. Z-score greater than 1 signifies confident alignment. The model of SYT1 showed Z-scores ranging from 1.20 - 4.59, indicating the confidence of the alignments between the target and templates.

Besides the quality assessment measures taken by I-TASSER, other assessments were also performed. The stereochemical quality of the modeled protein was analyzed by PROCHECK and RAMPAGE. It was found by PROCHECK that the phi/psi angles of 72.7% residues fell in the most favored regions. Only 4.3% of the residues fell in disallowed region. RAMPAGE, however, showed that 80.3% residues fell in favoured region (figure 3). The stereochemical quality of the structure was also evaluated with the help of Molprobit by calculating phi and psi torsion angles, backbone bond lengths and backbone bond angles. Molprobit provides a clashscore as a result of an all-atom contact analysis, performed

after adding the hydrogen atoms to the structure. When non-donor acceptor atoms overlap by more than 0.4Å, at least one of the two atoms must be modeled incorrectly. A clash at this location is noted and incorporated into the clashscore. Clash

score is simply the number of clashes per 1000 atoms. The modeled SYT1 showed a clashscore of 134.65 which is quite low. These results showed that the modeled structure was of good quality.

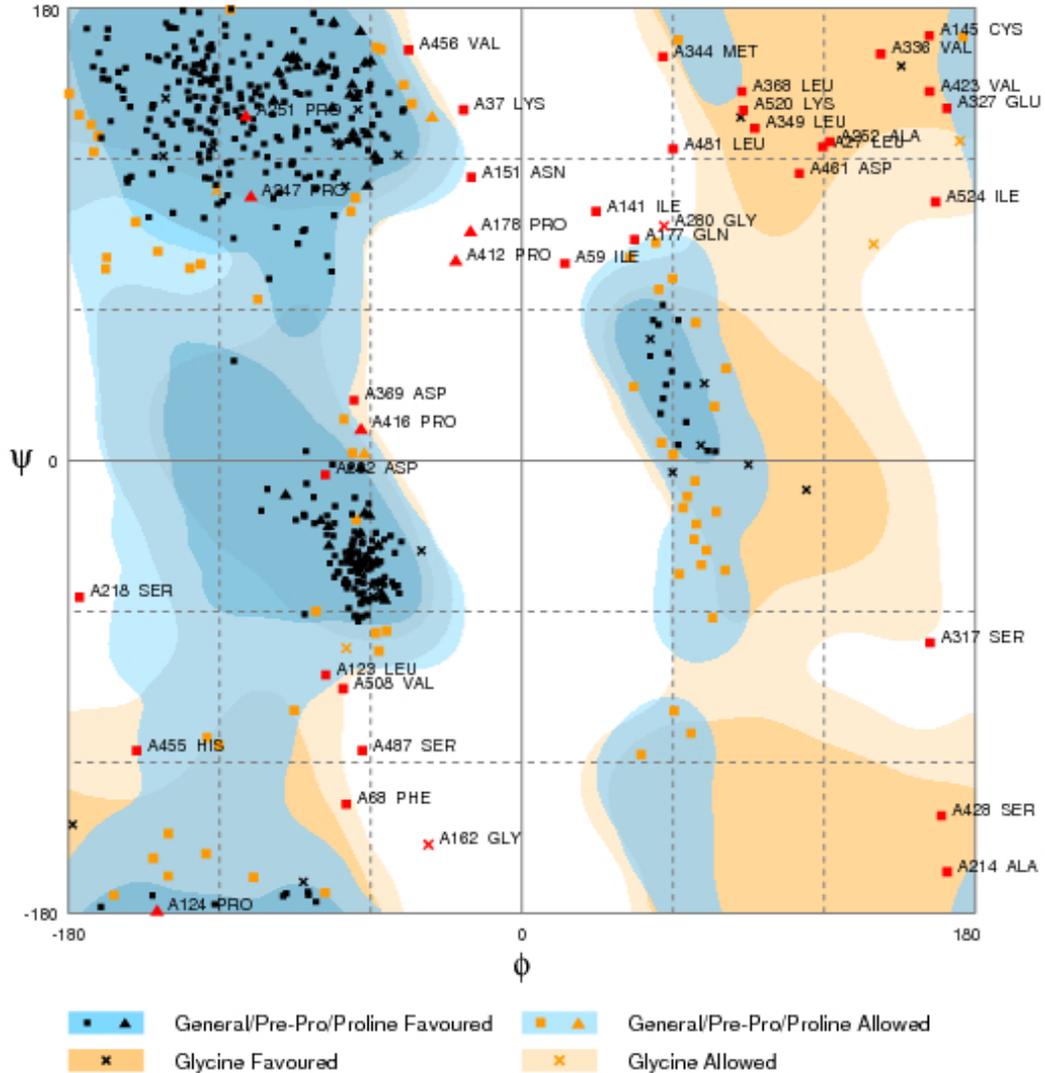


Figure 3: Ramachandran plot of SYT1 homology model as shown by RAMPAGE

**Tertiary structure of SYT1 protein**

The tertiary structure of SYT1 is shown in figure 4. This showed the presence of 7 helices and 19 strands (which formed 7 sheets) and 12 beta hairpins. The topology (figure 5) showed that the N-terminal part is primarily consisted of alpha helices, while the C-terminal portion was made primarily of beta sheets. The C2A domain contains three beta sheets and one alpha

helix and the C2B domain contains four beta sheets and one alpha helix. TMHMM predicted residues 7-29 to be transmembrane. The region is mainly comprised of alpha helix as expected for a transmembrane region. The C2 domain is a Ca<sup>2+</sup>-binding protein structural domain which is approximately 130-145 amino acids long and found in various membrane-associated signaling proteins in a large

number of organisms (Nalefski and Falke, 1996). These proteins mediate a broad array of crucial intracellular processes including membrane trafficking, the generation of lipid second messengers, activation of GTPases and the control of protein phosphorylation. Although the function of the C2 domain in several proteins is not known, it is considered that  $Ca^{2+}$  neutralizes negatively charged residues in the loop regions of the C2 domain and permits its interaction with phospholipids in the membrane which leads to trafficking (Rizo and Sudhof, 1998).

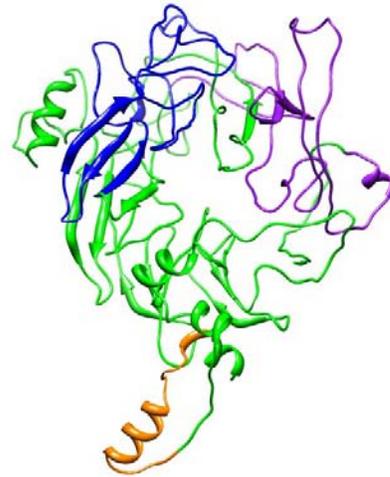


Figure 4: Three-dimensional model of SYT1. C2A and C2B are shown in blue and purple respectively. Transmembrane region is shown in orange.

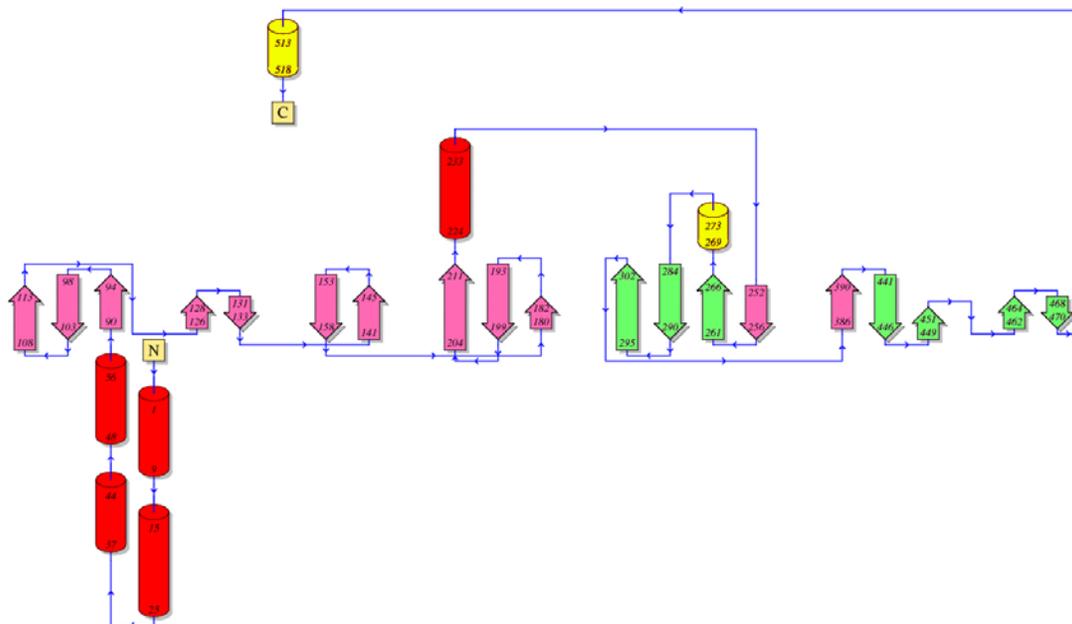
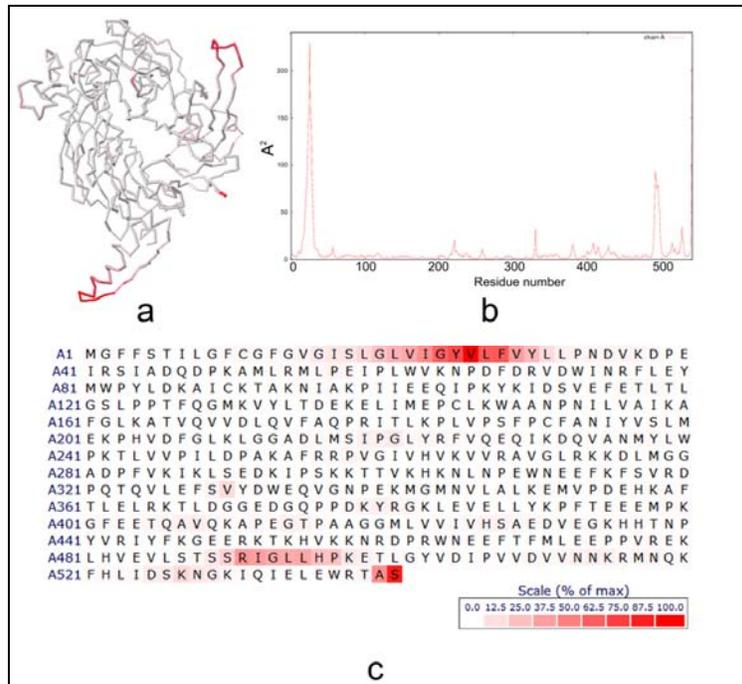


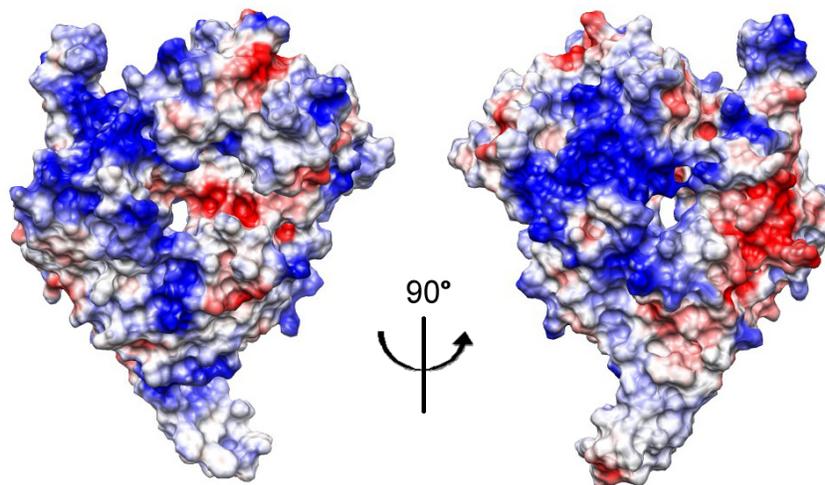
Figure 5: Topology of three-dimensional model of SYT1 as predicted by PDBsum. Helices and sheets are shown in red and pink, respectively. Helices and sheets which are part of C2 domains are shown in yellow and green, respectively.

Investigation of molecular flexibility by means of coarse-grained protein dynamics (Camps *et al.*, 2009) identified three major regions as flexible. A graphical representation of the B factor, which reflects spatial uncertainty, for different residues of the SYT1 is shown in figure 6. The extreme C-terminal residues are flexible. Part of the transmembrane region and the loop region

of C2B region was found to be flexible. Flexibility of the transmembrane region is expected as the bulk portion of the protein is outside the membrane and it moves around the transmembrane region. The loop region of the C2B region do not form any rigid structure in the protein, flexibility is also expected for this region.



**Figure 6: Flexibility of SYT1. a.** Flexibility to rigidity is shown in a gradient of red to white in the 3D model; **b.** Flexibility is indicated by peaks; **c.** Flexibility is indicated in a red white gradient over the entire sequence.



**Figure 7: Surface electrostatic potential of SYT1.** Red portions are electronegative and blue portions are electropositive. White portions are neutral.

The distribution of electrostatic potentials (figure 7) showed that the transmembrane region is neutral, which is expected as this region interacts with lipid bilayer. However, the C2A and C2B regions were primarily positively charged. It is also notable that the highly flexible region of the

protein had either positive or negative electrostatic potentials. The presence of charged residues in the loop regions of high flexibility in the C2B strongly suggests their involvement in dynamic charge-mediated interactions with other molecules. The prediction of interacting surfaces predicted

by POLYVIEW3D also showed that the positively charged loop region of C2B is highly interacting (figure 8). POLYVIEW3D predicted several interacting residues of which 18 and 19 residues fall in C2A and C2B regions, respectively. This shows that C2 domains are important for interacting with other macromolecules. The transmembrane region is largely interacting as it is placed into the lipid bilayer.

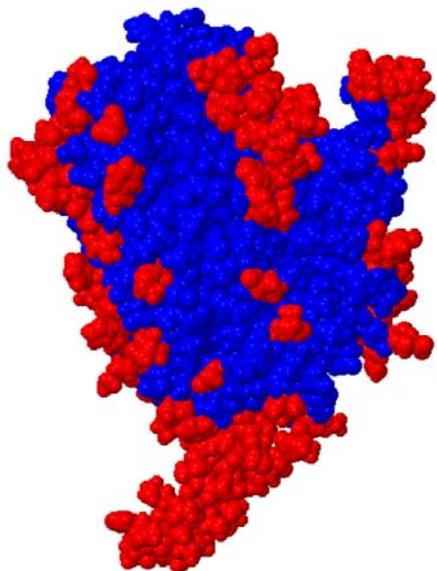


Figure 8: Interacting surface areas of SYT1. Interacting residues are shown in red.

#### Interaction with the $\text{Ca}^{2+}$ ion

The results of I-TASSER indicate that the calcium ion binding sites are Asp282, Lys286, Thr299, Val301 and His303. These residues were identified for docking on the basis of the structural alignment of 3HXA (Crystal structure of *Pseudomonas aeruginosa* PilY1 C-terminal domain). The  $\text{Ca}^{2+}$  was docked with modeled SYT1 protein by I-TASSER server. The Ligplot (Wallace *et al.*, 1996) of the complex as shown by PDBSum showed that the calcium ion interacts with Phe284, Thr300 and Lys302.

The three dimensional positioning of the amino acids around the calcium ion and the Ligplot are shown in the figure 9. The interaction between the protein and the calcium ion proposed in this study are useful for understanding the potential mechanism of protein and substrate binding. It is widely recognized that hydrogen bonds play an important role for the functioning of biological macromolecules. The binding mode of  $\text{Ca}^{2+}$  as predicted by I-TASSER indicates that Phe284, Thr300 and Lys302 are important for interaction with  $\text{Ca}^{2+}$ . As these three amino acids interact with the calcium ion, others are responsible for the formation of binding pocket of the protein.

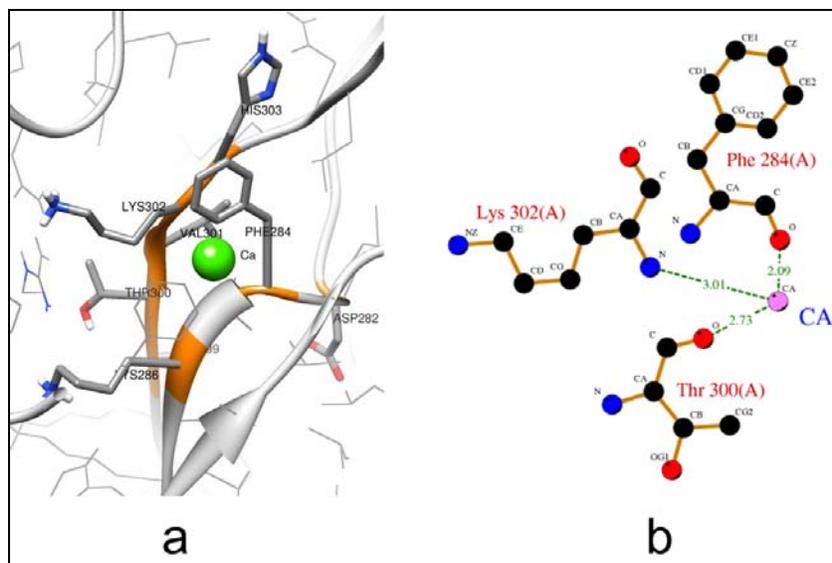


Figure 9: Interaction of Calcium ion with the SYT1 protein. a. Three-dimensional orientation of active site residues surrounding  $\text{Ca}^{2+}$  ion; b. Ligplot of SYT1 complexed with  $\text{Ca}^{2+}$ .

Thus, the present study showed that

- a. the SYT1 protein is a membrane localized protein
- b. it is part of one or more signal transduction cascades as it has several phosphorylation sites and its surface is highly interacting
- c. it is highly structured as shown by flexibility and disorder studies
- d. C2A mainly acts by the binding of calcium ion while C2B probably acts by interaction of other proteins.

In conclusion, this study is a starting point intended to guide laboratory experiments related with the biochemical properties and function of SYT1 protein from *Arabidopsis thaliana*. Further characterization is needed for revealing its role in various physiological processes of *Arabidopsis*.

#### Acknowledgement

The facility situated at the Department of Botany, Dinabandhu Mahavidyalaya, Bongaon is gratefully acknowledged.

#### References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215(3): 403-410.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28(1): 235-242.
- Bhattacharya A, Tejero R, Montelione GT. 2007. Evaluating protein structures determined by structural genomics consortia. *Proteins.* 66: 778-795.
- Buxbaum, E. 2007. *Fundamentals of protein structure and function*, Springer Science Business Media, LLC, New York.
- Camps J, Carrillo O, Emperador A, Orellana L, Hospital A, Rueda M, Cicin-Sain D, D'Abramo M, Gelpí JL, Orozco M. 2009. FlexServ: an integrated tool for the analysis of protein flexibility. *Bioinformatics* 25(13): 1709-1710.
- Cedano J, Aloy P, Pérez-Pons JA, Querol E. 1997. Relation between amino acid composition and cellular location of proteins. *J. Mol. Biol.* 266(3): 594-600.
- Craxton M. 2004. Synaptotagmin gene content of the sequenced genomes. *BMC Genomics* 5: 43.
- Chou PY, Fasman GD. 1973. Structural and functional role of leucine residues in proteins. *J. Mol. Biol.* 74: 263-281.
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G. 2000. Predicting Subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300(4): 1005-1016.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, MR., Appel, RD. and Bairoch, A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*, Ed. J.M. Walker, Humana Press, Totowa, New Jersey, USA, pp: 571-607.
- Geourjon C, Deléage G. 1995. SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput. Appl. Biosci.* 11: 681-684.
- Ginalski K. 2006. Comparative modeling for protein structure prediction. *Curr. Opin. Struct. Biol.* 16(2): 172-177.
- Hilfiker S, Pieribone VA, Nordstedt C, Greengard P, Czernik AJ. 1999. Regulation of synaptotagmin I phosphorylation by multiple protein kinases. *J. Neurochem.* 73: 921-932.
- Johns SJ. 2005. TOPO2, Transmembrane protein display software, <http://www.sacs.ucsf.edu/TOPO2/>
- Kang R, Swayze R, Lise MF, Gerrow K, Mullard A, Honer WG, El-Husseini A. 2004. Presynaptic trafficking of synaptotagmin I is regulated by protein palmitoylation. *J. Biol. Chem.* 279(48): 50524-50536.
- 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(3): 567-580.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: a program to check the stereochemistry of protein structures. *J. Appl. Crystallogr.* 26: 283-291.
- Laskowski RA. 2009. PDBsum new things. *Nucleic Acids Res.* 37: D355-D359.
- Lewis JD, Lazarowitz SG. 2010. *Arabidopsis* synaptotagmin SYTA regulates endocytosis and virus movement protein cell-to-cell transport. *Proc. Natl. Acad. Sci. U S A.* 107(6): 2491-2496.
- Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC. 2003. Structure validation by C $\alpha$  geometry:  $\phi$ ,  $\psi$  and C $\beta$  deviation. *Proteins* 50: 437-450.
- Nalefski EA, Falke JJ. 1996. The C2 domain calcium-binding motif: Structural and functional diversity. *Prot. Sci.* 5: 2375-2390.
- Olsen HB, Kaarsholm NC. 2000. Structural effects of protein lipidation as revealed by Lys<sup>B29</sup>-myristoyl, des(B30) insulin. *Biochemistry* 39(39): 11893-11900.
- Pagni M, Ioannidis V, Cerutti L, Zahn-Zabal M, Jongeneel CV, Hau J, Martin O, Kuznetsov D, Falquet L. 2007. MyHits: improvements to an interactive resource for analyzing protein sequences. *Nucleic Acids Res.* 35: W433 - W437.
- Paital B, Kumar S, Farmer R, Tripathy NK, Chainy GBN. 2011. *In silico* Prediction and characterization of 3D structure and binding properties of catalase from the commercially important crab, *Scylla serrata*. *Interdiscip. Sci. Comput. Life Sci.* 3: 110-120.
- Patil A, Nakamura H. 2006. Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Letters.* 580(8): 2041-2045.
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25(13): 1605-1612.
- Podell S, Gribskov M. 2004: Predicting N-terminal myristoylation sites in plant proteins. *BMC Genomics.* 5: 37.
- Porollo A, Meller J. 2007. Versatile Annotation and Publication Quality Visualization of Protein Complexes Using POLYVIEW-3D. *BMC Bioinformatics* 8: 316.
- Pruitt KD, Tatusova T, Maglott TR. 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 35: D61-D65.
- Qu Y, Bolen CL, Bolen DW. 1998. Osmolyte-driven contraction of a random coil protein. *Proc. Natl. Acad. Sci., USA.* 95: 9268-9273.
- Resh MD. 1999. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim. Biophys. Acta* 1451(1): 1-16
- Rhodes, G. 2006. *Crystallography Made Crystal Clear*. 3rd ed., Academic Press, Burlington, MA.
- Rizo J, Sudhof TC. 1998. C2-domains, structure and function of a universal Ca<sup>2+</sup>-binding domain. *J. Biol. Chem.* 273: 15879-15882.
- Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5(4): 725-738.
- Sadoul K, Boyault C, Pabion M, Khochbin S. 2008. Regulation of protein turnover by acetyltransferases and deacetylases. *Biochimie* 90(2): 306-312.
- Schapiro AL, Voigt B, Jasik J, Rosado A, Lopez-Cobollo R, Menzel D, Salinas J, Mancuso S, Valpuesta V, Baluska F, Botella MA. 2008. *Arabidopsis* synaptotagmin 1 is required for the

- maintenance of plasma membrane integrity and cell viability. *Plant Cell* 20: 3374-3388.
- Schlessinger A, Rost B. 2005. Protein Flexibility and Rigidity Predicted From Sequence. *Proteins Struct. Funct. Bioinform.* 61: 115-126.
- Schwartz R, Ting CS, King J. 2001. Whole proteome pI values correlate with subcellular localizations of proteins for organisms within the three domains of life. *Genome Res.* 11(5): 703-709.
- Sharmin S, Moosa MM, Islam MS, Kabir I, Akter A, Khan H. 2011. Identification of a novel dehydration responsive transcript from tossa jute (*Corchorus olitorius* L.). *J Cell Mol Biol* 9(1): 21-29.
- Sigrist CJA, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, Hulo N. 2010. PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res.* 38: D161-D166.
- Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33: W244-W248.
- Teilum K, Olsen JG, Kragelund BB. 2009. Functional aspects of protein flexibility. *Cell Mol Life Sci.* 66(14): 2231-2247.
- Wallace AC, Laskowski RA, Thornton JM. 1996. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* 8: 127-134.
- Ward JJ, McGuffin LJ, Bryson K, Buxton BF, Jones DT. 2004. The DISOPRED server for the prediction of protein disorder. *Bioinformatics* 20: 2138-2139.
- Wright PE, Dyson HJ. 1999. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* 293: 321-331.
- Yamazaki T, Kawamura Y, Minami A, Uemura, M. 2008. Calcium-dependent freezing tolerance in *Arabidopsis* involves membrane resealing via synaptotagmin SYT1. *Plant Cell* 20(12): 3389-3404.
- Yamazaki T, Takata N, Uemura M, Kawamura Y. 2010. *Arabidopsis* synaptotagmin SYT1, a type I signal-anchor protein, requires tandem C2 domains for delivery to the plasma membrane. *J. Biol. Chem.* 285(30): 23165-23176.
- Zhang Y. 2007. Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins* 69(S8): 108-117.
- Zheng J, Knighton DR, Xuong NH, Taylor SS, Sowadski JM, Ten Eyck LF. 1993. Crystal structures of the myristoylated catalytic subunit of cAMP-dependent protein kinase reveal open and closed conformations. *Prot. Sci.* 2(10): 1559-1573.