Three dimensional structure prediction and ligand-protein interaction study of expansin protein ATEXPA23 from *Arabidopsis thaliana* L.

Anamika Basu[1], Anasua Sarkar[2]*, Ujjwal Maulik[2] & Piyali Basak[3]

[1]Department of Biochemistry, Gurudas College, Kolkata -700 054, West Bengal, India

[2]Department of Computer Science and Engineering; [3]School of Bioscience and Engineering,
Jadavpur University, Kolkata -700 032, West Bengal, India

Arabidopsis thaliana L. is a small flowering plant that is widely used as a model organism in plant biology. In the present study, we study the peripheral membrane protein ATEXPA23 from *Arabidopsis thaliana* L. using homology modelling and molecular docking. The microarray analysis shows expression of ATEXPA23 (AT5G39280) protein, which leads to loosening and extension of plant cell walls. This protein is differentially expressed during different stages of plant embryogenesis. It contains one expansin-like CBD domain and one expansin-like EG45 domain. ATEXPA23 belongs to the expansin family in expansin a subfamily. The 3D model after refinement is used to explore the xyloglucan binding characteristics of ATEXPA23 using AutoDock. The docking analysis shows that the surface exposed aromatic amino acid residues Phe 193 and Phe 265 interact with ligand xyloglucan through CH-л interaction. The binding energy values of docking reflect a stable conformation of the docked complex. The interaction of expansin protein with carbohydrate xyloglucan, present in hemicellulose structures of plant cell wall, is thoroughly analysed with cellotetrose and xyloglucan heptasaccharide using electrostatic potential calculation. This CH-л non-covalent interaction predominates on the cellulose-xyloglucan interaction in plant cell wall during cell growth.

Keywords: ATEXPA23, CH-л interaction, Molecular docking, Molecular modelling

Expansins are typically 250-275 amino acids long, pH-dependent wall-loosening proteins[1] required for cell expansion[2] in many developmental processes *e.g.* plant embryogenesis, leaf initiation fruit softening[3], xylem formation[4], abscission[5], seed germination[6], root colonization[7], internodal elongation[8], root elongation[9], root hair initiation[10] and the penetration of pollen tubes *etc*. Expansins are peripheral membrane proteins (UniProt кB - Q9FL79 (EXP23_ARATH)) that have an affinity for the membrane because they bind either another membrane protein or lipid head groups. Peripheral membrane proteins do not integrate into the hydrophobic core layer. Expansin genes, first isolated from cucumber (*Cucumis sativus* L.)[11] have now been identified in many plant species *e.g. Arabidopsis*[10], tomato, oat, maize[12], rice[8], soybean[9] *etc*. Among 35 expansin genes in *Arabidopsis thaliana* L., 26 genes code for α-expansins (EXPA), 5 genes code for β-expansins (EXPB) whereas 3 genes code for α-expansin like-proteins (EXLA) and the rest one gene codes for β-expansin like-protein (EXLB)[13,14].

The mature expansins are of 25-27 kDa, consisting of two domains[13].

i. An amino-terminal domain of ~120 amino acid residues with structural and sequence similarity to family-45 endoglucanases and hosts a series of conserved cysteines and a His-Phe-Asp (HFD) motif that makes up part of the catalytic site of family-45 endoglucanases.

ii. A carboxy-terminal domain of ~98 amino acid residues that is hypothesized to function as a polysaccharide-binding domain (this is not experimentally established), based on conserved aromatic and polar residues on the surface of the protein.

In the plant cell wall, the cellulose microfibrils are linked via hemi cellulosic tethers to form the cellulose-hemicellulose network, which is embedded in the pectin matrix. Hemicellulose is a complex polysaccharide matrix composed of different residues branched in three kinds of backbones, named xylan, xyloglucan (XyG) and mannan. Xyloglucans have a main β-D-(1→4)-glucan backbone (denoted as G) generally branched with α(1→6)-linked D-xylopyranosyl (denoted as X) or β-D-galactopyranosyl (1→2)-D-xylopyranosyl

residues (denoted as L). The structure and molecular distribution of these XyG side chains varies in different plant tissues and species. As stated earlier, expansins disrupt the cellulose-hemicellulose association transiently, allowing slippage or movement of cell wall polymers[15]. In the year 2000, Cosgrove DJ[1] proposed that at low pH, plant cell wall extends faster. This event is known as 'acid growth'. Cellulose microfibrils exhibited few cellulose cross-peaks with the main hemicellulose, xyloglucan, suggesting limited entrapment in the microfibrils rather than extensive surface coating. The entrapped xyloglucan has motional properties that are intermediate between the rigid cellulose and the dynamic pectins[16]. But the molecular mechanism by which expansin loosens the cellulosic network within the cell wall is not yet established. α-Expansins (EXPA) may promote such movement by inducing local dissociation and slippage of xyloglucans on the surface of the cellulose and the optimum pH for cell wall extension for α-expansins is pH 4 whereas β-expansins (EXPB) work on a different glycan, perhaps xylan, for similar effect with pH optimum in between 5 and 6 (for maize pollen)[13].

By analysing the results of microarray datasets for different stages of embryogenesis in *Arabidopsis*, it has been observed that certain expansin genes are over expressed in specific tissues during their growth. In the early stage of embryogenesis, the zygote elongates 3 times before the cell division as a result of the extension of the cells along the side walls. In preglobular stage of embryogenesis. α-*expansin* proteins *e.g.* ATEXPA21 and ATEXPA23 are expressed in higher amounts (results are not shown). ATEXPA21 functions through plant-type cell wall modification involved in multidimensional cell growth as well as unidimensional cell growth. But ATEXPA23 involves in unidimensional cell growth only. ATEXPA23 is expressed significantly only during globular (specifically preglobular) stage of embryogenesis as shown in the results from *Arabidopsis* eFP Browser[17].

Computational approaches are frequently used for the sequence analysis and functional characterization of proteins[18,19]. Various structural and physicochemical properties of proteins can be illustrated by using computational tools when the crystal structure of interested protein is not available[20]. Although precise and accurate structure of proteins can be guaranteed by experimental methods but the disadvantage is that the wet laboratory based methods are time consuming and large amount of purified protein is required for this purpose.

Electrostatic interactions play important role in biological process by controlling protein- ligand interactions. Many proteins *e.g.* trypsin, acetylcholinesterase[21], bacterial expansin[22] are analyzed with their electrostatic potentials along with their functions.

We have used computational methods as an excellent and cost-effective alternative for analyzing structure and function of 259 amino acid long protein ATEXPA23. By using computational biology methods, we predict the structure of the protein and its interactions the ligand.

Though there are innumerable reports on the involvement of expansin proteins in various developmental processes in plants, the crystal structure of the *Arabidopsis* protein is not available yet. Not only 3D structures but also the role of these proteins during cell elongation must be clarified. Exact mechanism of interaction of expansin proteins with different carbohydrate moieties is not completely explained till now. In this context, structural aspects of ATEXPA23 and studies on interaction with plant cell wall carbohydrate molecules *e.g.* xyloglucan, cellotetrose molecules have been carried out in our present work. The results obtained might be helpful to identify the role of expansin in cell elongation. The protein is subjected to several online and desktop based bioinformatics tools to study physicochemical properties. The 3D structure of the protein is established using homology modelling approach. The predicted structure is used for docking xyloglucan and identify the ligand-protein interaction sites. The study reveals amino acids present in the ligand binding site.

Materials and Methods
3D structure prediction

For a monomeric protein, tertiary structure prediction is of vital importance for function prediction. Three dimensional structure prediction can be done by using different methods *e.g.* homology modelling, fold recognition/threading and De-novo/ Abs-initio methods. Among several web servers for structure prediction, we use SWISS-Model[23].

Homology modelling

Homology modelling of ATEXPA23 was carried out by SWISS-MODEL Version 8.05[23] by using NCBI Reference Sequence: NP_198744.1. The Following steps were followed:

Template Search

Template search with Blast and HHBlits was performed against the SWISS-MODEL template library (SMTL, last update: 2016-01-06, last included PDB release: 2015-12-31). The target sequence was searched with BLAST[24] against the primary amino acid sequence contained in the SMTL. A total of 30 templates were found. An initial HHblits profile was built using the procedure outlined in[25] followed by 1 iteration of HHblits against NR20. The obtained profile was then searched against all profiles of SMTL. A total of 26 templates were found.

Template Selection

For each identified template, the template's quality was predicted from features of the target-template alignment. The templates with the high quality were then been selected for model building.

Model Building

Models were built based on the target-template alignment using Pro mod-II. Coordinates which were conserved between the target and the template were copied from the template to the model. Insertions and deletions were remodeled using a fragment library. Side chains were then rebuilt. Finally, the geometry of the resulting model was regularized by using a force field. When loop modelling with ProMod-II[26] did not give satisfactory results, an alternative model was built using MODELLER[27].

Model Quality Estimation

The global and per-residue model quality was assessed using the QMEAN scoring function[28]. For improved performance, weights of the individual QMEAN terms were trained specifically for SWISS-MODEL. The QMEAN score is a composite score consisting of a linear combination of 6 terms. The pseudo-energies of the contributing terms are given below together with their Z-scores with respect to scores obtained for high-resolution experimental structures of similar size solved by X-ray crystallography.

Secondary structure analysis of the modelled protein

Secondary structure analyses of the query protein were performed by PSIPRED Server[29]. The physicochemical parameters of the protein sequence that included amino-acid and atomic compositions, molecular weight and isoelectric point (ρI) were computed by FFPred v2.0 (Eukaryotic Function Prediction) obtained from The PSIPRED Protein Sequence Analysis Workbench[30].

The sequence was then submitted to PDBsum[31], a pictorial database that provides an at-a-glance overview of the contents of each 3D structure deposited in the Protein Data Bank (PDB) and to evaluate the detailed topology and identify the clefts in the protein.

Protein- ligand Molecular docking

The 3D structure of the ligand molecule was obtained from PubChem database (CID: 11147508). Docking of ATEXPA23 with ligand xyloglucan was carried with AutoDock 4.2.3. The most favorable docking strucure with least binding energy was visualized by the USCF Chimera software[32]. The MMFF94 force field[33] was used for energy minimization of ligand molecule xyloglucan. Gasteiger partial charges were added to the ligand atoms. Non-polar hydrogen atoms were merged, and rotatable bonds were defined. Affinity (grid) maps of 60X60X60 Å grid points and 1.5Å spacing were generated using the Autogrid program[34]. Each docking experiment was derived from 20 different runs that are set to terminate after a maximum of 250000 energy evaluations. The population size was set to 50.

Electrostatic potential map generation

For both protein ATEXPA23 and protein- ligand binding structure, electrostatic potential energy was calculated by using Adaptive Poisson-Boltzmann Solver (APBS) system[35] by solving linear Poisson Boltzmann equation.

Results

Results for homology modelling and model quality estimation
Results for template search and template selection

Of all the PDB templates retrieved by HHblits and BLAST search from SWISS MODEL server, 2HCZ1.A template for amino acid range 31-259 showed the best sequence similarity with highest coverage with NCBI Reference Sequence: NP_198744.1 with β–expansin 1a protein from *Zea maize* (Table 1).

Model building and model quality estimation

The structure predicted by SWISS-MODEL Version 8.05 with the α helices and β pleated sheets visualized by Jmol 11.6.1 and its (blue) superimposition with the template (red) was illustrated in (Fig. 2). The modelled structure had GMQE score of 0.57 and QMEAN4 of −10.17 (Table 2).

Superimposition of C_α backbone structure of query protein ATEXPA23 with its template 2hcz.1. A using

Table 1 — Summary of the results obtained by SWISS-Model using as templates the homologous structures listed

Model no.	Template	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Range	Coverage	Description	GMQE score	QMEAN4 score
1	2hcz.1.A	29.95	monomer	HHblits	X -ray	2.75Å	0.36	31-259	0.84	β-expansin 1a	0.57	−10.17
2	2hcz.1.A	33.49	monomer	BLAST	X -ray	2.75Å	0.37	35-259	0.82	β-expansin 1a	0.56	−9.38
3	1n10.2.A	30.84	monomer	HHblits	X -ray	2.90Å	0.36	31-255	0.83	Pollen allergen Phl p 1	0.59	−5.68

Table 2 — QMEAN score for ATEXPA23 modelled protein

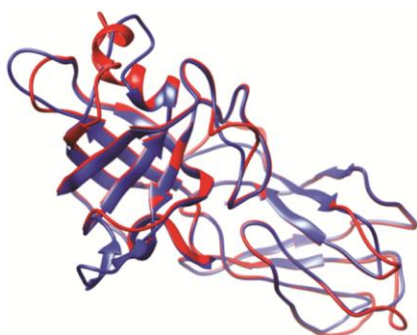| Parameters	Values for predicted model
C_β interaction energy	−4.07
All-atom pairwise energy:	−4.46
Solvation energy:	−6.75
Torsion angle energy:	−7.15
Total QMEAN-score:	−10.17

Fig. 1 — Structural superposition of the template 2hcz 1.A and modelled ATEXPA23 structure

UCSF Chimera[32] (Fig. 1) suggested that structural alignment was good. In both the proteins, the two domains namely family-45 endoglucanase-like domain and Cellulose-binding-like domain obtained from Prosite motif search library are found distinguishably.

2hcz.1.A (β-expansin 1a) (EXPB1_MAIZE) was identified as the template with sequence identity 29.95 with EXP23_ARATH (Fig. 2). The residues Gly 74, 77,104, 206, Ala 75, Cys 76, 106 Tyr 107, His 150, Phe 151, Asp 152, 207, Leu 153, Ser 154, Ile 208, Leu 210, and Met 211 were found to be conserved in template and query sequences in patches of primary sequences where four or more amino acids were similar in two sequences. Five amino acid residues, Gly 206, Asp 207, Ile 208, Leu 210 and Met 211 are present at a stretch in the cellulose binding domain of both the proteins.

Seconadry structure analysis of predicted model

FFPred v2.0 (Eukaryotic Function Prediction)[30] was used for primary sequence analyses of ATEXPA23. Examination of the predicted protein model (PM) using PSIPRED[29] reveals 5 β sheets, 5 hairpins, 3 psi loops, 7 β bulges, 16 strands, 3 helices, 27 β turns, 5 γ turns and 3 disulphide bonds .

Amino acid compositions of ATEXPA23 shows that aliphatic, hydrophobic amino acids *e.g.* gly and valine are present in higher amount whereas amino acids containing charged side chains *e.g.* glutamate, arginine and histidine are present in very less amounts.To check the reliability of 3D model, Ramachandran plot corresponding to the predicted model (PM) was generated using website PDBSum[31]/PROCHECK[36]. Validation shows that 66.7% of the residues were in the most favoured regions of Ramachandran plot, followed by 28.5% in additional allowed regions, 3.8% in generously allowed regions and 1.1% in disallowed regions.

Docking studies

Molecular docking method was applied for prediction of of the binding site of ATEXPA23 for various ligands *e.g.* cellotetrose (CID 91972060), xyloglucan oligosaccharide (CID 54758601), xyloglucan (CID 529401) and xyliglucan heptasaccharide (CID 4463046). Docking calculations were carried out using Docking Server[37]. Gasteiger partial charges were added to the four ligand atoms. Non-polar hydrogen atoms were merged, and rotatable bonds were set as 10. Docking calculations were carried out on every protein ligand model. Essential hydrogen atoms, Kollman united atom type charges, and solvation parameters were added with the aid of AutoDock tools [34]. Affinity (grid) maps of 60X60X60 Å grid points and 1Å spacing were generated using the Autogrid program[34]. Auto Dock parameter set-and distance-dependent dielectric functions were used in the calculation of the van der Waals and the electrostatic terms, respectively.

Docking simulations were performed using the Lamarckian genetic algorithm (LGA) and the Solis & Wets local search method[38]. Initial position, orientation, and torsions of the ligand molecules were set randomly. Highest number of rotatable torsions are released during each docking. Each docking result was derived from 10 different runs that were set to

```
CLUSTAL O(1.2.1) multiple sequence alignment
sp|Q9FL79|EXP23_ARATH    MNLLGKMIYVEGFMMIMATLLVSMSYGHRAMINDVAEAPVFDDVVSPNGLDSSWYDARAT   60
sp|P58738|EXPB1_MAIZE    MGSLA------NNIMVVGAVLAALVAGGSCGPPKVPP-----GPNITTNYNGKWLTARAT   49
                         *   *.         :*::.::*.::  *   .  .*       .  ::..*  ****

sp|Q9FL79|EXP23_ARATH    FYGDIHGG--ETQQGACGYGDLFKQGYGLETAALSTALFNEGYTCGACYQIMCVNDPQWC  118
sp|P58738|EXPB1_MAIZE    WYGQPNGAGAPDNGGACGIKNVNLPPYSGMTACGNVPIFKDGKGCGSCYEVRCKEKPECS  109
                         :**: .*.    : **** ::     .. :**:*     **:**:: * :.*: .

sp|Q9FL79|EXP23_ARATH    L-PGSVKITATNFCPPDYSKTEGVWCNPPQKHFDLSLPMFLK-------IAQYKAGVVPV  170
sp|P58738|EXPB1_MAIZE    GNPVTVYIITDMNYEP-----------IAPYHFDLSGKAFGSLAKPGLNDKIRHCGIMDV  157
                         * :* ** *:  *            ***** *  .        :.*:: *

sp|Q9FL79|EXP23_ARATH    KYRRISCARTGGVK----FETKGNPYFLMILPYNVGGAGDIKLMQVKGD-KTGWITMQKN  225
sp|P58738|EXPB1_MAIZE    EFRRVRCKYPAGQKIVFHIEKGCNPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPMKLS  217
                         ::**: *   .* *    :*.  ** :*  :*   *.  *** **::: ..: *   *: .

sp|Q9FL79|EXP23_ARATH    WGQNWTTGVNLTGQ-GISFRVTTSDGVTKDFNNVMPNNWGFGQTFDGKINF-  275
sp|P58738|EXPB1_MAIZE    WGAIWRMDTAKALKGPFSIRLTSESGKKVIAKDVIPANWRPDAVYTSNVQFY  269
                         **   .  .  : : :*:*:*:..*  .  ::*:* **  .: .:::*
```

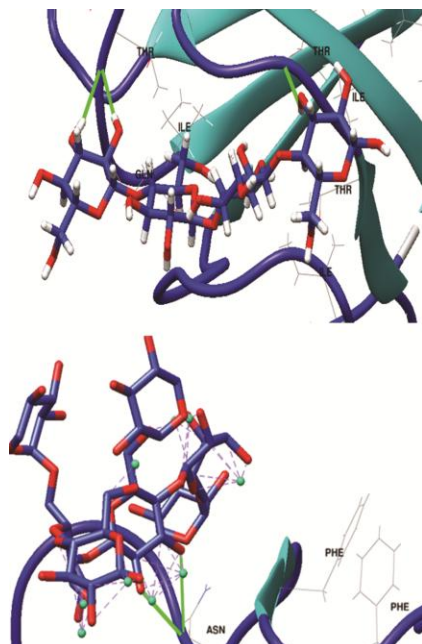Fig. 2 — Amino acid sequence of Alignment of ATEXPA23 and the template 2hcz.1.A used for modelling

Fig. 3 — A Docking structures of cellotetrose and xyloglucan heptasaccharide with ATEXPA23

terminate after a maximum of 2500 energy evaluations. The population size was set to 50. During the search, a translational step of 0.2 Å, and quaternion and torsion steps of 5 were applied. Results of docking of ATEXPA23 with the four ligands were shown in (Table 3).

Protein –ligand binding for two ligands *e.g.* cellotetrose and xyloglucan heptasaccharide were visualized using UCSF Chimera[32] as shown in following (Fig. 3).

Docking structures of cellotetrose and xyloglucan heptasaccharide with ATEXPA23

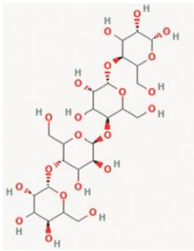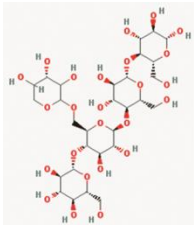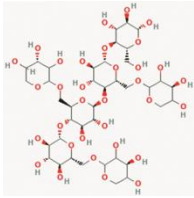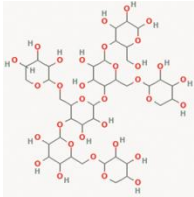In cellotetrose and expansin protein complex, the interacting amino acids are Thr 232, Thr 231, Gln 223, Ile 220, and Thr 221. Cellotetrose forms H bonds with Thr 232 and Thr 231 with bond lengths of 3.08 and 3.12 Å respectively, whereas Gln 223 forms a polar interaction and Ile 220 shows hydrophobic interaction with ATEXPA23. Specifically, O5 atom and O21 atom of cellotetrose interacts with CB, CG2, OG1 of Thr 232, respectively, through hydrogen bonds. Similarly, polar interactions occur between O18 and H39 with NE2 atom of Gln 223, and H14 and O8 with OG1 atom of Thr 231. Hydrogen bonding is also observed between H42 and OG1 of Thr 232.

Two types of interactions are observed during binding of xyloglucan heptasaccharide with expansin protein ATEXPA23. First one is a polar interaction with Asn 190 and cation-pi interactions with Phe 265 and Phe 193. The second type of interaction is hydrophobic and involves C9, C10, C11 and C12 atoms of xyloglucan heptasaccharide and CD2, CE2 atoms of Phe 265 with distances 3.26, 3.60 3.86, and 3.64 Å, respectively. Similarly, polar interactions take place between H8, O3 and H6 of xyloglucan heptasaccharide with ND2 atom of Asn 190 with distances 3.55, 3.63 and 3.64 Å, respectively. Similarly, polar interactions are observed between H8, O3 and H6 of ligand with ND2 atom of Asn190 with distances 3.55, 3.63 and 3.81 Å, respectively. Most importantly cation-pi interactions occur among H8 and H3 atoms of this ligand with Phe 193 and Phe 265 with distances of 3.32 and 3.39 Å, respectively.

Electrostatic potential calculation

Electrostatic potential calculation for protein ATEXPA23 shows that in front part only amino acid residues namely Thr 112, Cys 90, Ala 75, Ser 160 exhibit electrostatic potential of -5 KT/e (Fig. 4). A few patches on same part of protein are with +5 KT/e electrostatic potential *e.g.* surface around Asn 97, Trp 38, Arg 42, Arg 157, Gly 25 and

Table 3 — Molecular docking results for different ligands

Comp no.	Pubchem ID no. and 2D structure	Name of ligand	Binding energy (Kcal/ mol)	Internal energy (Kcal/ mol)	Intermolecular energy (Kcal/mol)	Torsional energy (Kcal/mol)	Electrostatic energy(Kcal/mol)
1	CID91972060	Cellotetrose	−2.27	−8.63	−2.51	7.16	0.06
2	CID54758601	Xyloglucan oligosaccharide	1.33	−6.03	−1.92	8.35	0
3	CID52940180	Xyloglucan	−2.91	−9.92	−3.48	8.35	0
4	CID44630346	Xyloglucan heptasaccharide	−0.14	−9.52	−3.33	8.05	0

Thr 231. The white region of the molecule in images indicates the electronegativity difference in that region are not very significant. Surfaces around Phe 177, Pro 175, Tyr 176, and Val 218 show white colour in front part of the protein. Surprisingly in back part no red colouration is shown (image not shown). Some white and blue patches are shown. Areas of low electrostatic potential in images are characterized by an abundance of electrons shown in red.

After binding with xyloglucan, electrostatic potential on the surface of ATEXPA23 protein changes totally (Fig. 5). Carbohydrate binding domain (CBD) turns red with -10 KT/e electrostatic potential whereas family-45 endoglucanases domain remains less electronegative. After binding with xyloglucan molecular electrostatic potential surface of CBD domain changes remarkably from white to dark red. So, this region becomes highly electronegative from its neutral character on its surface.
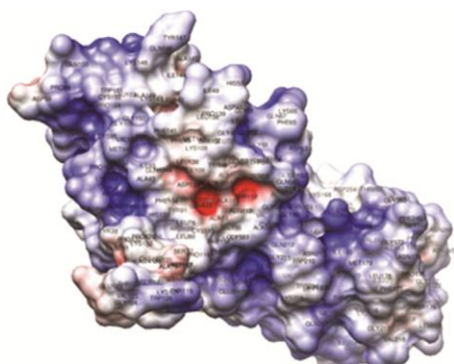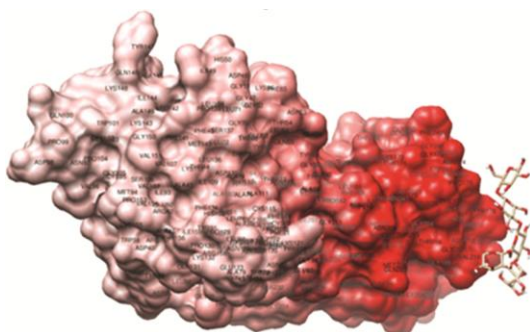
Fig. 4 — Electrostatic potential map of ATEXPA23

Fig. 5 — Electrostatic potential map of ATEXPA23 after binding with xyloglucan

Discussion

Cosgrove in 2016[39] compared the two models of primary cell wall structure of plant cell wall structure of plant cells. In the first model, known as the tethered network model, xyloglucan molecule binds to cellulose microfibril to form a load- bearing molecular network. Similarly, in the second one, the biomechanical hotspot model depicts cellulose-cellulose junctions in primary cell wall, are linked by xyloglucan-cellulose amalgam. For both models, cellulose-xyloglucan interaction is common in primary cell wall structure. Protein expansin plays an important role in cell wall expansion by slippage or rearrangement of matrix polymers. Structural difference between cellulose and hemicelluloses is mainly due to substituted xylose residues in xyloglucan. From molecular docking studies of carbohydrates with expansin protein it can be concluded that all interacting amino acids *e.g.* Thr 232, Thr 231, Glu 223, Ile 220, Thr 221 for cellotetrose and Asn 190, Phe 193, Phe265 for xyloglucan, are present in carbohydrate -binding domain of ATEXPA23 protein. But the cation-pi interaction occurs during xyloglucan heptasaccharide binding with ATEXPA23 which is completely absent in cellotetrose binding.

Here expansin protein is non-covalently bound with carbohydrates. Specifically, C-H bonds of carbohydrate interact preferably with aromatic amino acid residues of ATEXPA23. Electropositive C-H bonds involve in CH-л interactions with electron-rich aromatic groups of Phe 193 and Phe 265, both amino acids are present in loop structure of expansin protein. This weak CH- л noncovalent interaction depends on the type of carbohydrate residue and their position within carbohydrate –binding modules of the protein[40]. Molecular docking studies clearly show that expansin protein preferably bind with xyloglucan moiety, compared to cellulose. Electrostatic potential study indicates in presence of xyloglucan, electrostatic potential of carbohydrate binding domain changes completely from neutral to highly electronegative. This change also confirms the CH-л noncovalent interaction between expansin and xyloglucan. Probably this type of non-covalent interaction predominates on the cellulose-xyloglucan interaction in plant cell wall during cell growth. Thus, expansin protein interfere the cellulose- xyloglucan network and helps in slippage of xyloglucan on the surface of the cellulose during cell wall extension.

References

1   Cosgrove DJ, Loosening of plant cell walls by expansins. *Nature*, 407 (2000) 321.
2   McQueen-Mason SJ, Durachko DM & Cosgrove DJ, Two endogenous proteins that induce cell wall extension in plants. *Plant Cell*, 4 (1992) 1425.
3   Hiroko H, Ito A, Moriguchi T & Kashimura Y, Identification of a new expansin gene closely associated with peach fruit softening. *Postharvest Biol Technol*, 29 (2003) 1.
4   Gray-Mitsumune M, Mellerowicz EJ, Abe H, Schrader J, Winzéll A, Sterky F, Blomqvist K, McQueen-Mason SJ, Teeri TT & Sundberg B, Expansins abundant in secondary xylem belong to subgroup A of the α-expansin gene family. *Plant Physiol*, 135 (2004) 1552.
5   Cho HT & Cosgrove DJ, Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana* L. *Proc Natl Acad Sci U S A*, 97 (2000) 97836.
6   Feng C, Dahal P & Bradford KJ, Two tomato expansin genes show divergent expression and localization in embryos during seed development and germination. *Plant Physiol*, 127 (2001) 928.
7   Frédéric K, Amoroso A, Herman R, Sauvage E, Petrella S, Filée P, Charlier P, Joris B, Tabuchi A, Nikolaidis N & Cosgrove DJ, Crystal structure and activity of Bacillus subtilis YoaJ (EXLX1), a bacterial expansin that promotes root colonization. *Proc Natl Acad Sci U S A*, 105 (44) (2008) 168768.
8   Yi L & Kende H, Expression of β-expansins is correlated with internodal elongation in deep water rice. *Plant Physiol*, 127 (2001) 6459.

9   Lee DK, Ahn JH, Song SK, Choi YD & Lee JS, Expression of an expansin gene is correlated with root elongation in soybean. *Plant Physiol*, 131 (2003) 985.

10   Cho HT & Cosgrove DJ, Regulation of root hair initiation and expansin gene expression in *Arabidopsis. Plant Cell,* 14 (2002) 3237.

11   Tatyana SY, Shi J, Durachko DM, Guiltinan MJ, McQueen-Mason SJ, Shieh M &Cosgrove DJ, Molecular cloning and sequence analysis of expansins--a highly conserved, multigene family of proteins that mediate cell wall extension in plants. *Proc Natl Acad Sci U S A*, 92 (1995) 9245.

12   Yajun W, Meeley RB & Cosgrove DJ, Analysis and expression of the α-expansin and β-expansin gene families in maize. *Plant Physiol*, 126 (2001) 22213.

13   Javier S & Cosgrove DJ, The expansin superfamily. *Genome Biol*, 6 (2005) 242.

14   Kende H, Bradford K, Brummell D, Cho HT, Cosgrove DJ, Fleming A, Gehring C, McQueen-Mason SJ, Rose J, Voesenek LA, Nomenclature for members of the expansin superfamily of genes and proteins. *Plant Mol Biol*, 55 (2004) 311.

15   Cosgrove DJ, Assembly and enlargement of the primary cell wall in plants. *Annu Rev Cell Dev Biol*, 13 (1997) 171.

16   Dick-Pérez, M, Zhang Y, Hayes J, Salazar A, Zabotina OA, & Hong M, Structure and interactions of plant cell-wall polysaccharides by two-and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry*, 50 (2011) 989.

17   Debbie W, Vinegar B, Nahal H, Ammar R, Wilson GV & Provart NJ, An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS one*, 2 (2007) e718.

18   Yata VK, Mahajan S, Thapa A, Ahmed S, Biswas AD, Sanjeev A & Mattaparthi VSK, *In silico* methods reconfirm CDK2 as a potential molecular target of 5-flourouracil. *Indian J Biochem Biophys*, 53 (2016) 199.

19   Gupta M, Singh VK & Gupta S, Molecular modeling and virtual docking studies on D1 protein of Phalaris minor biotypes with isoproturon. *Indian J Biochem Biophys*, 53 (2016) 39.

20   Sudha S & Bandopadhyay R, "*In silico* analysis of the structure and interaction of COP1 protein of *Arabidopsis thaliana* L. *Indian J Biochem Biophys*, 51 (2014) 343.

21   Hildebrandt A, Blossey R, Rjasanow S, Kohlbacher O & Lenhof HP, Electrostatic potentials of proteins in water: a structured continuum approach. *Bioinformatics*, 23 (2007) e99.

22   Pastor N, Dávila S, Pérez-Rueda E, Segovia L & Martínez-Anaya C, Electrostatic analysis of bacterial expansins. *Proteins*, 83 (2015) 215.

23   Konstantin A, Bordoli L, Kopp J & Schwede T, The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22 (2006) 195.

24   Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipman, DJ, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25 (1997) 3389.

25   Remmert M, Biegert A, Hauser A & Söding J, HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods*, 9 (2012) 173.

26   Guex N & Peitsch MC, SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis*, 18 (1997) 2714.

27   Šali A & Blundell TL, Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*, 234 (1993) 779.

28   Pascal B, Biasini M & Schwede T, Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, 27 (2010) 343.

29   McGuffin LJ, Bryson K & Jones DT, The PSIPRED protein structure prediction server. *Bioinformatics*, 16 (2000) 40430.

30   Buchan DWA, Minneci F, Nugent TC, Bryson K & Jones DT, Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res*, 41 (2013) W34931.

31   Laskowski RA, PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res*, 29 (2001) 221.

32   Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE, UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*, 25 (2004) 1605.

33   Halgren TA, Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comput Chem*, 17 (1996) 49034.

34   Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS & Olson AJ, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, 30 (2009) 2785.

35   Konecny R, Baker NA & McCammon JA, iAPBS: a programming interface to the adaptive Poisson–Boltzmann solver. *Comput Sci Discov*, 5 (2012) 015005.

36   Laskowski RA, MacArthur MW & Thornton JM, PROCHECK: validation of protein structure coordinates. International tables of crystallography, volume F. *Crystallograp Biol Macromol* (2001) 722.

37   Bikadi Z & Hazai E, Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock. *J Cheminform*, 1 (2009) 15.

38   Solis FJ & Wets RJB, Minimization by random search techniques. *Math Opera Res*, 6 (1981) 19.

39   Cosgrove DJ, Catalysts of plant cell wall loosening. *F1000Res*, 5 (2016) 40.

40   Hudson KL, Bartlett GJ, Diehl RC, Agirre J, Gallagher T, Kiessling LL & Woolfson DN, Carbohydrate–aromatic interactions in proteins. *J Am Chem Soc*, 137 (2015) 15152.