



The Open Biotechnology Journal

Content list available at: <https://openbiotechnologyjournal.com>



RESEARCH ARTICLE

Functional and Structural Characterization of a Novel Isoamylase from *Ostreococcus tauri* and Role of the N-Terminal Domain

Nicolas Hedín¹, Julieta Barchiesi¹, Diego F. Gomez-Casati^{1,*} and María V. Busi^{1,*}

¹Centro de Estudios Fotosintéticos y Bioquímicos (CEFOTBI), CONICET - Universidad Nacional de Rosario, Suipacha 531, 2000 Rosario, Argentina

Abstract:

Background:

The debranching starch enzymes, isoamylase 1 and 2 are well-conserved enzymes present in almost all the photosynthetic organisms. These enzymes are involved in the crystallization process of starch and are key components which remove misplaced α -1,6 ramifications on the final molecule.

Aim:

In this work, we performed a functional and structural study of a novel isoamylase from *Ostreococcus tauri*.

Methods:

We identified conserved amino acid residues possibly involved in catalysis. We also identified a region at the N-terminal end that resembles a Carbohydrate Binding Domain (CBM), which is more related to the family CBM48, but has no spatial conservation of the residues involved in carbohydrate binding.

Results:

The cloning, expression and biochemical characterization of this N-terminal region confirmed that it binds to polysaccharides, showing greater capacity for binding to amylopectin rather than total starch or amylose.

Conclusion:

This module could be a variant of the CBM48 family or it could be classified within a new CBM family.

Keywords: Carbohydrate Binding Domain (CBM), Debranching starch enzymes, Amino acid residues, Polysaccharides, Biochemical characterization, Carbohydrates.

Article History

Received: November 04, 2019

Revised: November 28, 2019

Accepted: January 04, 2020

1. INTRODUCTION

Carbohydrates are the most abundant molecules on earth; each year photosynthesis fixes approximately 100 billion tons of CO₂ and H₂O into cellulose, starch, sucrose and other sugars. Starch is one of the most important sources of energy in the human diet. It is composed of two polymers, amylose (essentially an α -1,4-glucan chain) and amylopectin (similar to amylose but branched with α -1,6 glycosidic bonds) whose combination gives rise to a highly insoluble starch granule [1, 2]. Amylo-

pectin represents a 75-90% of total starch content, and approximately 5% of its glucose residues are involved in branch bonds. It was reported that the branching pattern is responsible for amorphous lamellae formation, which is important for the semi-crystalline structure of the granule, and along the linear part of amylopectin molecule, it is responsible for the lamellar organization of starch [3]. In contrast, amylose is present in a smaller amount in the starch granule and is believed to be responsible for granule “filling”, increasing its density [4]. The way in which starch accumulates is controlled by the enzymes involved in its synthesis [5]. The frequency and pattern of amylopectin branching, far from being randomly generated, follow a model that leads to the formation of crystalline lamellae [6]. The ADP-glucose pyrophosphorylase,

* Address correspondence to this author at the Centro de Estudios Fotosintéticos y Bioquímicos (CEFOTBI), CONICET - Universidad Nacional de Rosario, Suipacha 531, 2000 Rosario, Argentina; Fax: +54-0341-4370044; Tel: +54-0341-4371955; E-mails: busi@cefobi-conicet.gov.ar, gomezcasati@cefobi-conicet.gov.ar

which synthesizes ADPGlc, is one of the most important enzymes in this process, being the regulatory point of starch synthesis [7 - 9]. Other enzymes involved in starch synthesis are Starch Synthases (SS) and Branching Enzymes (SBE) [7 - 9]. Both Granular Bound Starch Synthases (GBSS) and Soluble Starch Synthases (SSS) catalyze the addition of a glucose moiety from ADPGlc to the non-reducing end of the linear α -1,4-glucan chain. While starch synthases generate new α -1,4 bonds, SBE are responsible for the branching of the linear glucan chain, generating α -1,6 glycosidic bonds [10]. On the other hand, Debranching Enzymes (DBEs) participate in the cleavage of the α -1,6 glycosidic linkages generated by SBEs. The latter group of enzymes can be classified into two types: Isoamylases (ISA) and Pullulanases or Limit dextrinases (LDA). While isoamylase can debranch amylopectin, glycogen and phytoglycogen, pullulanase can only act on pullulan and amylopectin [11]. In *Arabidopsis thaliana*, ISA1 and ISA2 are implicated in amylopectin synthesis, whereas ISA3 and LDA have a major role in starch degradation [1].

Most of the enzymes involved in carbohydrate metabolism comprise a modular organization composed of a catalytic domain and one or more modules named Carbohydrate Binding Domain (CBM) [12]. A CBM is defined as an amino acid sequence with a particular folding configuration capable to interact and bind to polysaccharides [13]. Although many CBMs are commonly found attached to active enzymes, there are several exceptions where they are part of non-catalytical proteins (e.g. PTST from *Arabidopsis thaliana* [14]). CBM can be composed of 30 to 200 amino acids and can be as a sole domain or as a part of a tandem. According to their sequence, structure and substrate binding specificity, CBM can be classified into 86 families [15]. They can serve different purposes, and the reason why they are present on the proteins may vary. They can increase the rate of an enzymatic reaction through a strong physical association with the substrate, or they can also act as an organizing subunit, which helps assembling multienzyme complexes [13]. Among CBMs, a particular class is the Starch Binding Domain (SBD). SBDs are found in archaea, bacteria and eukaryotes and are distributed along twelve families of CBM: 20, 21, 25, 26, 34, 41, 45, 48, 53, 58, 68, & 69 (<http://www.cazy.org/>). These SBD modules have the evolutionary advantage of being able to disrupt the surface of the substrate due to the presence of two different binding sites [16 - 18].

CAZY is a constantly growing database and hence the number of CBM families is in perpetual growth, as data from new sequencing projects are analyzed and novel CBMs are found. In the last decade, many algal genomes have been sequenced, like those from *Micromonas pusilla* [19], *Ostreococcus lucimarinus* [20], *Volvox carterii* [21], *Chlorella variabilis* [22], *Chlamydomonas reinhardtii* [23] and *Ostreococcus tauri* [24, 25].

Recent effort toward understanding starch metabolism has set *O. tauri* on the spotlight [26 - 30]. *O. tauri* is the tiniest free-living photosynthetic eukaryote known to date and belongs to the ancient Prasinophyceae group [31, 32]. This picoalga has a rather compact genome when compared to other photosynthetic organisms, but it has a large set of genes coding

for many proteins involved in starch metabolism. One of the most extreme examples is the case of Starch Synthase III (SSIII) gene: Many plants such as *A. thaliana* have a single copy of this isoform, while *O. tauri* has three different copies that codes for SSIII [27, 28, 33 - 36]. The same is true for Glucan Water Dikinase (GWD) and Phosphoglucan Water Dikinase (PWD); *O. tauri* has 3 GWD and 2 PWD, while *A. thaliana* only has 2 GWD and 1 PWD [37]. Another feature of this organism is the presence of a single mitochondrion and chloroplast, that contains a single starch granule [38, 39]. These traits make *Ostreococcus* an excellent model organism to study starch metabolism and to understand how the fine structure of amylopectin molecule impacts on the final shape of the starch granule through the action of branching and debranching enzymes.

Here, we studied the structural properties of OsttaISA1, an isoamylase coded by Ot14g02550, which is a homolog to ISA1 from *Arabidopsis thaliana*. We cloned and expressed the N-terminal region of OsttaISA1 to evaluate its function as a CBM, and the interactions with different polysaccharides. In addition, we performed a structural model in order to compare the proposed structure and the location of amino acids important for OsttaISA1 function with other related proteins.

2. MATERIALS AND METHODS

2.1. Cloning, Expression and Purification

The part of the gene Ot14g02550 coding for the putative CBM from *O. tauri* (named as OsttaCBM) was cloned from genomic *O. tauri* DNA using primers containing EcoRI and XhoI restriction sites, which are compatible with pRSFDuet-1 expression vector (Novagen, EMD Biosciences Inc, Madison, WI, USA). The procedure was performed using standard molecular biology protocols and the following primers: CBM48I-Fw, ATCGAATTCTCTGGGAGCGACGCGAGTG G (EcoRI site is underlined); CBM48I-Rv, GATCTCGAGT CACTTTGCGTACGGGTCGAGC (XhoI site is underlined).

The plasmid containing the sequence coding for OsttaCBM was transformed into *Escherichia coli* BL21 (DE3) strain (F-*ompT gal dcm lon hsdSB*(rB- mB-) λ (DE3 [*lacI lacUV5-T7 gene 1 ind1 sam7 nin5*])). The cells were grown at 37°C to an OD600 = 0.5, and the expression was induced using IPTG to a final concentration of 0,8 mM, and further incubation at 18°C for 14 h. The cells were harvested by centrifugation at 13000 x g for 10 min at 4°C. The pelleted cells were resuspended in a buffer containing 25 mM Tris-HCl (pH 7.5), disrupted by sonication and centrifuged at 13000 x g for 10 min at 4°C. The supernatant was loaded onto a HiTrap chelating HP column (GE Healthcare BioSciences, Little Chalfont, UK) equilibrated with binding buffer (20 mM Tris-HCl pH 7.4, 500 mM NaCl and 25 mM imidazole). The column was washed twice with 15 volumes of binding buffer, and bound proteins were eluted using a linear gradient of elution buffer (25 mM Tris-HCl pH 7.5, 500 mM NaCl and 25-500 mM imidazole) [33]. The presence of the CBM protein in the eluted fractions was confirmed by SDS-PAGE analysis and immunoblotting. The fractions containing the protein of interest with no contaminations were pooled, concentrated to more than 0.5 mg/ml and desalted using Vivaspin 6 3000 MWCO

concentrators. Aliquots were stored at -20°C until use after the addition of glycerol to a final concentration of 20% (v/v). Under these final conditions, the protein was stable for at least 3 months.

2.2. Gel Electrophoresis and Immunological Studies

Polyacrylamide gel electrophoresis under denaturing conditions (SDS-PAGE) was done using a Bio-Rad Mini Protean system as described by Laemmli [40] and using 17% (v/v) polyacrylamide/bisacrylamide gels. Gels were stained after electrophoresis with Coomassie Brilliant Blue R250 (Sigma-Aldrich, St. Louis, MO, USA) or transferred to nitrocellulose membranes (BioRad, Hercules, CA, USA). Electrobotted membranes were incubated with penta-His antibody (anti-His antibody selector kit, Qiagen). The antigen-antibody complex was visualized with alkaline phosphatase-linked a-rabbit IgG, followed by staining with BCIP and NBT as described previously [41].

Native PAGE was performed as previously described by Bollag and Edelstein [42]. Protein samples were prepared in 20 mM Tris-HCl (pH 6.8) containing 10% (v/v) glycerol and 0.01% (w/v) bromophenol blue and then loaded on the gel, which was pre-run for about 2 h at 4°C (1 h at 100 V, 30 min at 150 V and 1 h at 200 V). Once loaded, the gel was performed at a constant voltage of 200 V at 4°C using a 10% separating gel (10% (w/v) acrylamide-0.22% (w/v) bisacrylamide, 250 mM Tris-HCl, pH 8.8, 6 x 8 cm x 1mm). Electrophoresis buffer contained 25 mM Tris-HCl and 192 mM glycine (pH 8.8). The gel was then stained using Coomassie Blue.

2.3. Polysaccharide-Binding Assay

Purified OsttaCBM (approximately 15 μg) was mixed with amylose (Fluka-10130), amylopectin (Fluka-10118) or starch (Fluka-85649, St. Louis, MO, USA) in 25 mM Tris-HCl (pH 7.5) at a final polysaccharide concentration of 5% (w/v). The assay was carried out at 20°C under orbital mixing for 20 minutes. The polysaccharides were pelleted by centrifugation at 13000 x g during 10 min and the supernatant was removed and boiled in SDS containing buffer (Tris-HCl 50mM pH 6.8, SDS 2% (w/v), Glycerol 6% (v/v), Bromophenol blue 0.004% (w/v) and β -Mercaptoethanol 1% (v/v)). The pellet was further washed using 25 mM Tris-HCl (pH 7.5) centrifugation, and then resuspended in an equal volume to the supernatant using SDS containing buffer and finally boiled [29, 43 - 45]. Samples were analyzed on SDS-PAGE and protein levels were determined by densitometric analysis of the protein bands stained with Coomassie Blue R-250, using the software GelPro analyzer (MC, Bethesda, MD, USA).

2.4. Homology Modeling and Sequence Alignment

Sequence similarity search of OsttaISA1 (Genbank code: XP_003083046, 1-851 amino acid residues) whole amino acid sequences was performed with PSI-BLAST [46] with the default parameters (inclusion threshold 0.005) until convergence, using no redundant databases. The sequences with an E value less than 1.10^{-5} were retrieved and aligned with the program CLUSTALW [47]. Alignment figures were performed using the online program ESPript 3.0 using the

default parameters [48]. Phylogenetic tree was built by maximum likelihood using a bootstrap analysis of 500 replicates in MEGA 10.0.2 version [49].

A 3D structural model was obtained using the @TOME V2.3 program, which includes T.I.T.O. (Tool for Incremental Threading Optimization), Scwrl and Modeller [50]. OsttaISA1 structure was modeled using glycogen debranching enzyme Trex from *Sulfolobus solfataricus* (PDB code 2VR5) [51] and Isoamylase 1 from *Chlamydomonas reinhardtii* (ChlreISA, PDB code 4OKD) [52]. CBM48 was modelled using GlgX from *Escherichia coli* a template (PDB code 2WSK) [53]. Alignment with the template was based on homology and secondary structure. Models were evaluated using Verify-3D [54] and ProSA-Web [55, 56] online programs. Superposition of the structures was performed using the PyMOL v2.1.0 [57] server v 1.0 [57].

3. RESULTS AND DISCUSSION

3.1. Sequence Alignment and Phylogenetic Analysis of OsttaISA1

Using PSI-Blast, we conducted similarity searches with OsttaISA1 as a query. Several results were retrieved with E-values lower than 10^{-5} . According to similarity and domain organization, we were able to determine that this protein is strongly related to isoamylases from plants and algae. This bioinformatic prediction allowed us to classify this protein as a family 1 isoamylases-related protein (named as OsttaISA1). The retrieved amino acid sequences were aligned using ClustalW [47], and analyzed using ESPript online server [48]. A partial alignment is shown in Fig. (1) and catalytic amino acids are indicated using the isoamylase from *Sulfolobus solfataricus* (Trex, PDB code 2VR5) as a template, as the catalytically-important residues have been identified for this organism (numbering refers to OsttaISA1) [51]. It can be observed that Asp363 (Asp466 in OsttaISA1) is conserved not only in OsttaISA1, but also in each analyzed sequence. Other aminoacids that are well conserved are Glu399 (Glu526 in OsttaISA1) and Asp471 (Asp597 on OsttaISA1) Table 1. Numerous studies analyzing the role of these amino acids in the GH13 proteins showed that Asp363 acts as a catalytic nucleophile and would form a covalent intermediate by binding to the glucose moiety at the reducing end of the molecule, while Glu399 is the catalytic acid/base that would protonate the oxygen from the -1 glucose [29, 51, 58]. These aminoacids are involved in the so-called double displacement mechanism, which is proposed for all the reactions involving GH13 enzymes [59]. According to the predictions based on sequence analysis, Asp471 (*S. solfataricus* numeration) would behave as the transition state stabilizer of the reaction. Additionally, in the characterization of isoamylase I (4OKD - ChlreISA1) from *Chlamydomonas reinhardtii* [52] two different carbohydrate Surface Binding Sites (SBS) have been described. These SBS are non-catalytic regions of the protein that interact with sugar rings through stacking interactions with aromatic residues [60]. Both of these sites, SBS1 and SBS2 are believed to recognize the helical structure of crystalline amylopectin [52]. The residues that provide aromatic rings for the stacking interaction are Trp757 according to 4OKD numbering (Trp740 in

OsttaISA1) and Trp767 (Trp750 in OsttaISA1) on site 1, and Trp652 (Trp629 in OsttaISA1), His794 (His778 on OsttaISA1) and Trp856 (Arg840 in OsttaISA1) on site 2. Analyzing complete protein sequence alignment, it can be observed that both Trp from SBS1 are conserved (Trp740 and 750 OsttaISA1 numbering), while SBS2 maintains Trp652 (ChlreISA1) as Trp629 (OsttaISA1) and His794 as His778 (ChlreISA1 and OsttaISA1 numbering, respectively) (Supplementary Fig. 1). The case for the third residue from SBS2, Trp856, is at least unexplored as it is not conserved on OsttaISA1 (Arg840), while the rest of the sequences analyzed in this alignment have an aromatic residue on that position (Table 1). Another feature that is absent on OsttaISA1 but present on ChlreISA1 is the presence of three loops (aa 216 – 226, 466 – 484 and 586 – 604, ChlreISA1 numbering), which are believed to be involved in interface formation on ISA1/ISA2 oligomers on the *Chlamydomonas reinhardtii* enzyme [52]. Furthermore, it remains to be determined if OsttaISA1 homodimerizes tail-to-tail like ChlreISA1 does, as bioinformatics prediction failed to detect the conservation of the majority of the 48 residues, which are responsible for this interaction. Additionally, a set of amino acid residues that are key to this dimerization are Arg667 and Asp848 (ChlreISA1 numbering). It was described that Arg667 generates a salt bridge between its NH1 with the OD1 of Asp848, however, both residues are absent in OsttaISA1 sequence (Table 1 and Supplementary Fig. 1) [51]. From the amino acid sequence alignment, it can be seen that both residues are not present in the Trex sequence either, suggesting that either these enzymes would not have the ability to form oligomers or that other residues would be involved in this mechanism (Fig. 1).

Using the sequences retrieved from BLAST, we performed a maximum likelihood phylogenetic tree (Fig. 2), which demonstrates a clear relatedness of OsttaISA1 to those enzymes from algae. Although the Trex enzyme from *Sulfolobus* is inside the algae cluster and it seems to evolve from a common ancestor to OsttaISA1, the evolutionary difference between this sequence and OsttaISA1 is clearly evident. Because Trex protein has been crystallized and studied in detail, it is an excellent candidate to obtain OsttaISA1 molecular structure using homology modeling.

3.2. Homology Modeling of OsttaISA1

Using the 3D structure of the previously crystallized 4OKD (ChlreISA1) and 2VR5 (Trex) as a template, we were able to generate a homology model for the isoamylase from *Ostreococcus* as described in the methods section. The analysis of the structure proposed using 4OKD as a template was evaluated with Verify 3D program and it showed that 98.0% of the residues had a positive score, and 89.6% of them were higher than 0.2 (average score 0.392). Using 2VR5, 99.8% of the residues were greater than zero, while 91.5% were higher than 0.2, averaging on 0.389. The results from Prosa II were also in accordance with high quality models, with its Z-score being -6.45 for 4OKD and -5.68 for 2VR5 [55, 56]. According to these results, we can conclude that the proposed model for OsttaISA1 is of good quality and accuracy to reach confidence in the molecular structure analysis. As described in the previous paragraph, 2VR5 is a glycogen-debranching enzyme

with dual activities: α -1,4 transferase and α -1,6 glucosidase, while 4OKD belongs to an isoamylase from the green algae *Chlamydomonas reinhardtii*. Although 4OKD is more related to OsttaISA1 from an evolutionary point of view, Trex has been studied in more depth and allowed us to determine the most important catalytic residues.

According to protein sequence analysis and classification from InterPro (detailed in Materials and Methods) the region from amino acid 121 to 218 resembles a carbohydrate-binding domain from family 48 (CBM48). Fig. (3A) shows in magenta the position of this putative CBM from OsttaISA1, and it can be clearly observed that it comprises six β -strands forming the classical CBM β -sandwich folding [27, 29]. The central part of the protein is composed of the Catalytic Domain (CD), which is also conserved among the α -amylase family, and belongs to the family 13 of glycoside hydrolase from the CAZY database [51, 61]. The CD domain consists of eight parallel β -strands, which are surrounded by eight α -helices. Furthermore, Fig. (3A) also shows the position of the most important aminoacids for catalysis, being all spatially and rotationally conserved (numbering is referred to OsttaISA1, 2VR5 yellow residues, OsttaISA1 red residues). The zoom-in shown in Fig. (3B) represents the accommodation of the substrate on the cleft and shows the possible interaction with the catalytic triad. Finally, Fig. (3C) shows the surface of the proposed protein folding. It is interesting to observe that both SBS1 and SBS2 have the ability to interact with oligosaccharides as observed for ChlreISA1. The slight curvature of the binding hints the possibility that those SBSs might recognize the helical structure of amylopectin.

3.3. Polysaccharide Binding Assay of the N-Terminal Region of OsttaISA1

Although CBMs can be encountered in non-active enzymes [14], they are normally associated with active enzymes involved in carbohydrate metabolism, driving the substrate to a more intimate and prolonged association with the catalytic domain [13]. It was also demonstrated that this module can act as a macromolecule complex aggregator and can also disrupt the polysaccharide surface [62, 63].

CBM48 can be found both in branching, debranching, and also in non-amylolytic enzymes (e.g. SEX4 protein, SNF1 beta subunit, AMPK beta subunit). They can be on the N- or C-terminal region of a protein and also in the middle of its amino acid sequence, as is the case for SNF1 complex from *Saccharomyces cerevisiae* [64].

Based on the structure proposed and alignment of the amino acid sequence of the N-terminal end of OsttaISA1 with sequences from other CBMs from family 20, 48 and 53, it was possible to detect amino acid residues that align with those already identified as responsible for polysaccharide binding (Fig. (4) and Supplementary Fig. 1). Fig. (4) shows that W151 and Y199 residues are spatially conserved, but it is interesting to note that the orientation of their aromatic rings differs from those from the amino acids present in the GlgX protein.

Thus, it is necessary to test if this N-terminal domain would be involved in the binding to polysaccharides. To evaluate this possibility, we cloned the sequence coding for the

Table 1. List of residues involved in catalysis and/or binding.

Residue	Function	Conserved Function
Asp363 (Trex)	catalytic nucleophile	YES (Asp466)
Glu399 (Trex)	catalytic acid/base	YES (Glu526)
Asp471 (Trex)	transition state stabilizer	YES (Asp597)
Trp757 (ChlreISA1)	stacking interactions SBS1	YES (Trp740)
Trp767 (ChlreISA1)	stacking interactions SBS1	YES (Trp750)
Trp652 (ChlreISA1)	stacking interactions SBS2	YES (Trp629)
His794 (ChlreISA1)	stacking interactions SBS2	YES (His778)
Trp856 (ChlreISA1)	stacking interactions SBS2	NO (Arg840)
Arg667 (ChlreISA1)	key for dimerization	NO (Ala648)
Asp848 (ChlreISA1)	key for dimerization	NO (Lys832)

Ref: Trex: *Sulfolobus solfataricus* glycogen debranching enzyme; ChlreISA1: *Chlamydomonas reinhardtii* isoamylase; SBS1: carbohydrate surface binding site 1; SBS2: carbohydrate surface binding site 2.

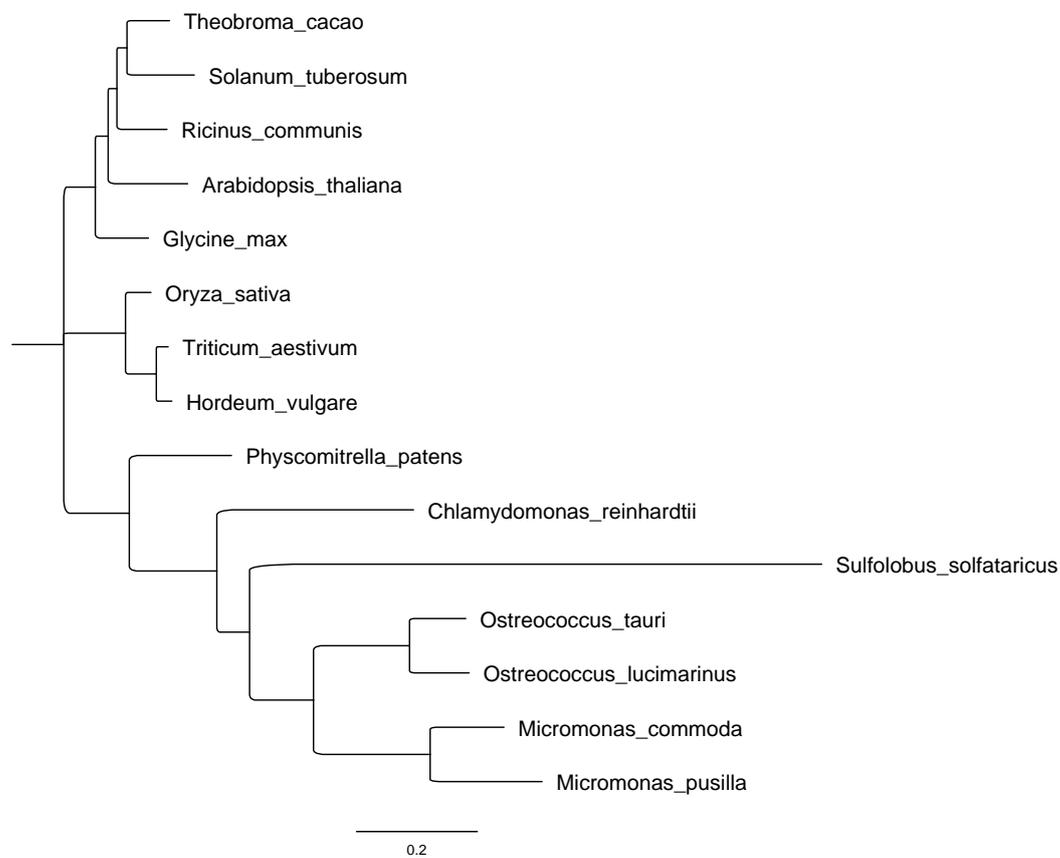


Fig. (2). Phylogenetic tree of ISA enzymes. The tree is based on the alignment of OsttaISA1 with the amino acid sequences obtained with PSI-Blast. The tree was built by Maximum Parsimony method in MEGA X version 10.0.4 (49). Phylogenies were determined by Bootstrap Analysis of 500 replicates in MEGA X. Branch lengths are proportional to distances. Bootstrap values are shown above branches.

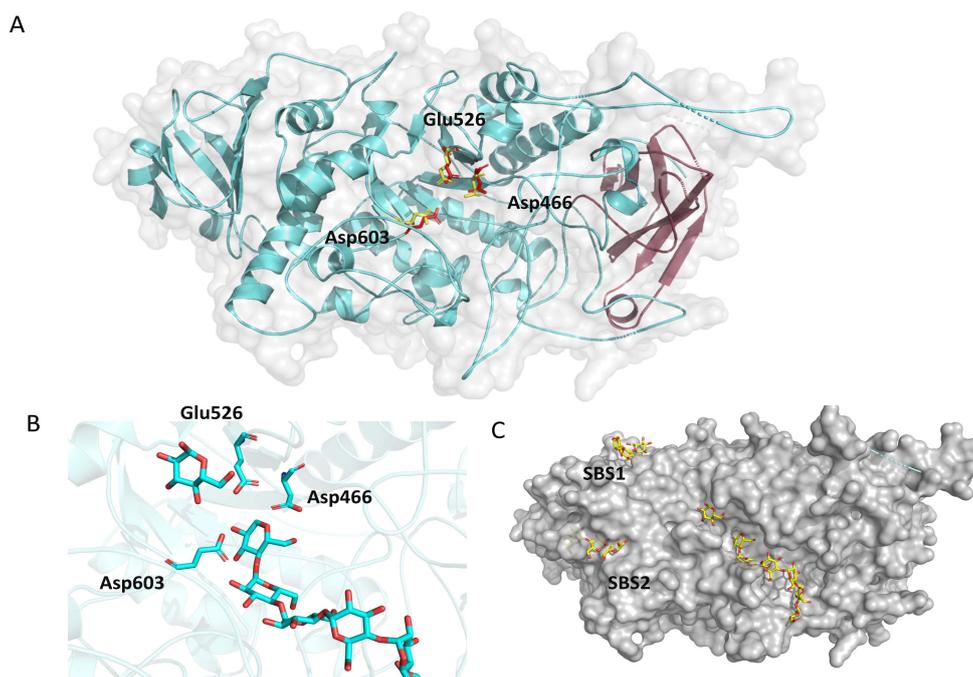


Fig. (3). Homology modeling of OsttaISA1. **A)** 3D structure of OsttaISA1 with the CBM domain represented in magenta and the rest of the polypeptide in cyan. Red residues corresponds to the catalytic core of OsttaISA1, and yellow residues represents the core from Isoamylase 1 from *Chlamydomonas reinhardtii*. Numbering refers to OsttaISA1. **B)** Close-up of the catalytic domain with glucose residues represented. **C)** Surface of OsttaISA1 showing catalytic cleft and SBS1 and SBS2.

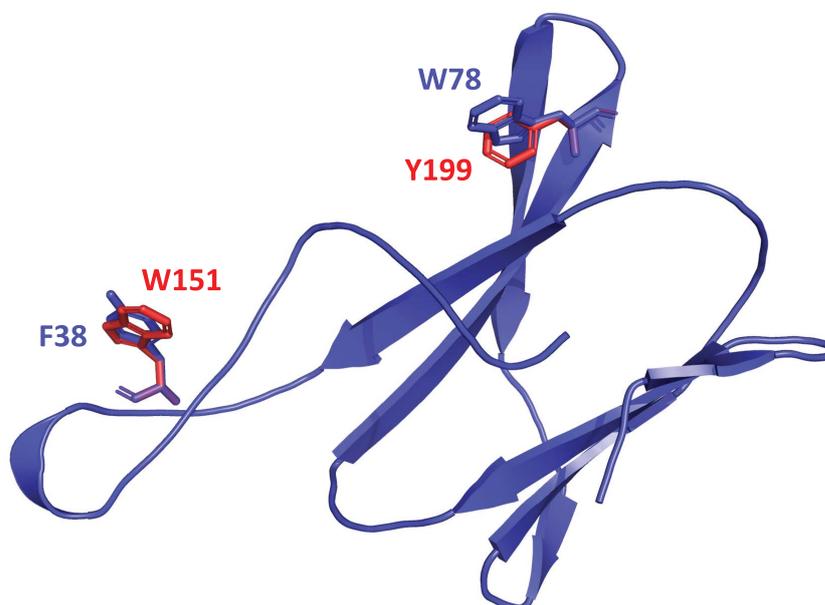


Fig. (4). Proposed structure of CBM48 from *O. tauri* using 2WSK as model. Aminoacids in BLUE are from glycogen debranching enzyme from *Escherichia coli* (2WSK), residues in RED are from *O. tauri*.

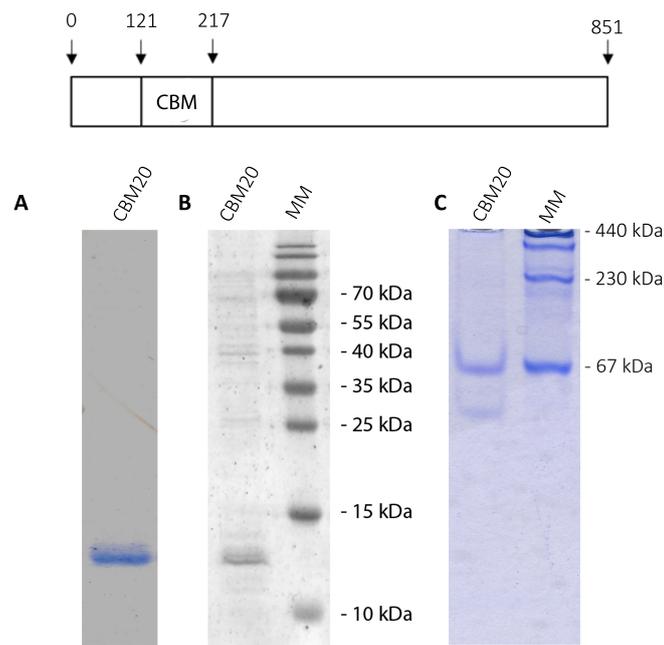


Fig. (5). Schematic representation of the protein studied in this article showing the position of the cloned CBM. The numbers indicates the amino acid positions. (A) SDS-PAGE analysis and (B) immunodetection of the recombinant OsttaCBM protein using anti-His antibody. MM corresponds to PageRuler prestained protein ladder (Thermo Scientific). (C) Native PAGE analysis of the purified OsttaCBM. Numerals on the right of Figures B and C indicate the molecular masses of the standards.

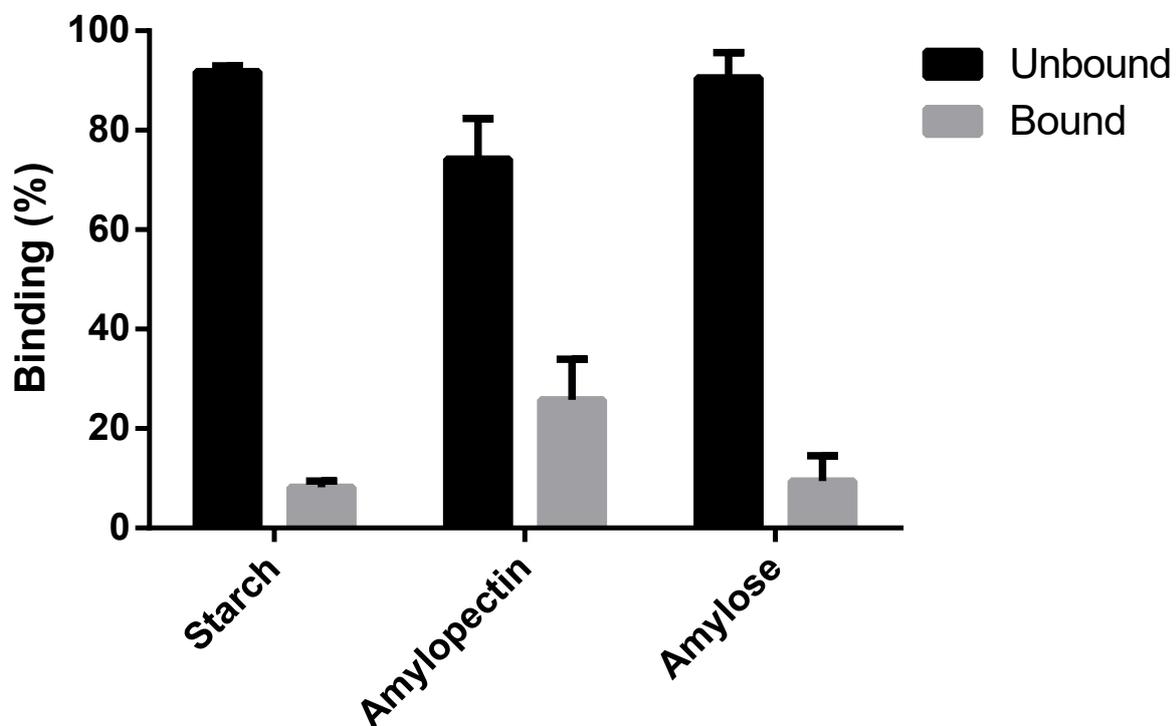


Fig. (6). Percentage of bound or unbound protein to different polysaccharides (starch, amylopectin and amylose) using recombinant OsttaCBM. Black boxes show the amount of free protein, while grey boxes shows the amount of protein that binds to each polysaccharide.

CONCLUSION

OsttaISA1 is classified as a debranching enzyme as its sequence has a very high identity to other debranching enzymes analyzed. According to our alignment and prediction, OsttaISA1 has every amino acid responsible for α -1,6-transferase activity, but lacks a key structural lid region believed to be involved in α -1,4-transferase activity located close to the active site [51]. As previously shown, these lids located close to the active site are key on the latter transferase activity, and in some polypeptides, the lid may be formed through dimerization of the protein [65, 66]. As predicted by our *in silico* analysis, OsttaISA1 lacks specific residues previously reported to be involved in dimers or higher oligomers formation [51, 52]; however, results show that OsttaCBM is able to form oligomers, at least *in vitro*. It is clear that further experiments are needed to probe the possible oligomerization capacity of OsttaISA1 and its functional relevance.

The binding assays give new insight into the function of the N-terminal region of the polypeptide. According to the bioinformatic prediction, this domain, which has a greater capacity for amylopectin binding, could be from the CBM family 48 or a variant of this family, because, as mentioned before, although important amino acids are conserved, there was no spatial co-localization for W151 and Y199 (which are responsible for polysaccharide binding) with those from the CBM48 model. These differences in the binding site could be a consequence of structural discrepancies between amylopectin and glycogen.

It is important to note that debranching enzymes are subdivided into two groups according to substrate specificity and protein sequence: ISA and Limit Dextrinase (LDA) [1]. Both *Arabidopsis thaliana* and *Ostreococcus tauri* have three ISA proteins: ISA1, ISA2 and ISA3, while there is only one LDA. As has been reported for *Arabidopsis*, ISA1 (homolog to OsttaISA1) and ISA2 are responsible for the process of debranching activity during the synthesis of starch, while ISA3 and LDA are involved on starch degradation by cleaving α -1,6 bonds [37]. ISA1 and ISA2 might be involved in the so-called “glucan trimming” model, where they correct the misplaced α -1,6 bonds by cleaving the ones that will not allow amylopectin packing into a highly ordered structure [67]. From this point of view, the result obtained is interesting, where amylopectin is the polysaccharide that showed the highest affinity to OsttaCBM. Taken together, these results might explain part of the mechanism that facilitates the cleavage of certain bonds but not others, generating the crystalline starch granule structure this way. In addition, the characterization of new enzymes capable of conferring differential properties to starch is of high importance to eventually design new strategies to obtain novel starches for different uses in different industries.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

This work was supported by grants from CONICET (PIP 0476), ASaCTeI (IO-2018-00098) and ANPCyT (PICT 2018-01440). NH is a research fellow from ANPCyT. JB, DFGC and MVB are research members from CONICET.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

We are grateful to María Belén Velázquez for excellent technical assistance.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publishers web site along with the published article.

REFERENCES

- [1] Zeeman SC, Kossmann J, Smith AM. Starch: Its metabolism, evolution, and biotechnological modification in plants. *Annu Rev Plant Biol* 2010; 61: 209-34. [<http://dx.doi.org/10.1146/annurev-arplant-042809-112301>] [PMID: 20192737]
- [2] Christiansen C, Abou Hachem M, Janecek S, Viksø-Nielsen A, Blennow A, Svensson B. The carbohydrate-binding module family 20-diversity, structure, and function. *FEBS J* 2009; 276(18): 5006-29. [<http://dx.doi.org/10.1111/j.1742-4658.2009.07221.x>] [PMID: 19682075]
- [3] Nakamura S, Satoh H, Ohtsubo K. Development of formulae for estimating amylose content, amylopectin chain length distribution, and resistant starch content based on the iodine absorption curve of rice starch. *Biosci Biotechnol Biochem* 2015; 79(3): 443-55. [<http://dx.doi.org/10.1080/09168451.2014.978257>] [PMID: 25384364]
- [4] Wattedled F, Dong Y, Dumez S, *et al.* Mutants of *Arabidopsis* lacking a chloroplastic isoamylase accumulate phytoglycogen and an abnormal form of amylopectin. *Plant Physiol* 2005; 138(1): 184-95. [<http://dx.doi.org/10.1104/pp.105.059295>] [PMID: 15849301]
- [5] Wang XY, Ma J, Wang CS, *et al.* Characterization of starch branching enzyme I (SBE I) gene in two *Triticum monococcum* accessions with different starch content. *Starke* 2015; 67(7-8): 663-72. [<http://dx.doi.org/10.1002/star.201500027>]
- [6] Pfister B, Zeeman SC. Formation of starch in plant cells. *Cell Mol Life Sci* 2016; 73(14): 2781-807. [<http://dx.doi.org/10.1007/s00018-016-2250-x>] [PMID: 27166931]
- [7] Smith AM. The biosynthesis of starch granules. *Biomacromolecules* 2001; 2(2): 335-41. [<http://dx.doi.org/10.1021/bm000133c>] [PMID: 11749190]
- [8] Ball SG. Starch metabolism. In: *The Chlamydomonas sourcebook*. second. 2009. [<http://dx.doi.org/10.1016/B978-0-12-370873-1.00009-5>]
- [9] Busi MV, Barchiesi J, Martin M, Gomez-Casati DF. Starch metabolism in green algae. *Starke* 2014; 66(1-2): 28-40. [<http://dx.doi.org/10.1002/star.201200211>]
- [10] Myers AM, Morell MK, James MG, Ball SG. Recent progress toward

- understanding biosynthesis of the amylopectin crystal. *Plant Physiol* 2000; 122(4): 989-97.
[http://dx.doi.org/10.1104/pp.122.4.989] [PMID: 10759494]
- [11] Manners DJ. Recent developments in our understanding of amylopectin structure. *Carbohydr Polym* 1989; 11(2): 87-112.
[http://dx.doi.org/10.1016/0144-8617(89)90018-0]
- [12] Henrissat B, Davies G. Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol* 1997; 7(5): 637-44.
[http://dx.doi.org/10.1016/S0959-440X(97)80072-3] [PMID: 9345621]
- [13] Shoseyov O, Shani Z, Levy I. Carbohydrate binding modules: Biochemical properties and novel applications. *Microbiology and Molecular Biology Reviews* 2006 June; 2006;70(2): 283-95.
- [14] Seung D, Soyk S, Coiro M, Maier BA, Eicke S, Zeeman SC. Protein targeting to starch is required for localising granule-bound starch synthase to starch granules and for normal amylose synthesis in *Arabidopsis*. *PLoS Biol* 2015; 13(2):e1002080
[http://dx.doi.org/10.1371/journal.pbio.1002080] [PMID: 25710501]
- [15] Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. Carbohydrate-binding modules: Fine-tuning polysaccharide recognition. *Biochem J* 2004; 382(Pt 3): 769-81.
[http://dx.doi.org/10.1042/BJ20040892] [PMID: 15214846]
- [16] Penninga D, van der Veen BA, Knechtel RM, et al. The raw starch binding domain of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. *J Biol Chem* 1996; 271(51): 32777-84.
[http://dx.doi.org/10.1074/jbc.271.51.32777] [PMID: 8955113]
- [17] Williamson MP, Le Gal-Coëffet MF, Sorimachi K, Furniss CS, Archer DB, Williamson G. Function of conserved tryptophans in the *Aspergillus niger* glucoamylase 1 starch binding domain. *Biochemistry* 1997; 36(24): 7535-9.
[http://dx.doi.org/10.1021/bi9702896] [PMID: 9200704]
- [18] Giardina T, Gunning AP, Juge N, et al. Both binding sites of the starch-binding domain of *Aspergillus niger* glucoamylase are essential for inducing a conformational change in amylose. *J Mol Biol* 2001; 313(5): 1149-59.
[http://dx.doi.org/10.1006/jmbi.2001.5097] [PMID: 11700070]
- [19] Worden AZ, Lee JH, Mock T, et al. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 2009; 324(5924): 268-72.
[http://dx.doi.org/10.1126/science.1167222] [PMID: 19359590]
- [20] Palenik B, Grimwood J, Aerts A, et al. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc Natl Acad Sci USA* 2007; 104(18): 7705-10.
[http://dx.doi.org/10.1073/pnas.0611046104] [PMID: 17460045]
- [21] Prochnik SE, Umen J, Nedelcu AM, et al. Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science* 2010; 329(5988): 223-6.
[http://dx.doi.org/10.1126/science.1188800] [PMID: 20616280]
- [22] Blanc G, Duncan G, Agarkova I, et al. The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell* 2010; 22(9): 2943-55.
[http://dx.doi.org/10.1105/tpc.110.076406] [PMID: 20852019]
- [23] Merchant SS, Prochnik SE, Vallon O, et al. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 2007; 318(5848): 245-50.
[http://dx.doi.org/10.1126/science.1143609] [PMID: 17932292]
- [24] Derelle E, Ferraz C, Rombauts S, et al. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc Natl Acad Sci USA* 2006; 103(31): 11647-52.
[http://dx.doi.org/10.1073/pnas.0604795103] [PMID: 16868079]
- [25] Blanc-Mathieu R, Verhelst B, Derelle E, et al. An improved genome of the model marine alga *Ostreococcus tauri* unfolds by assessing Illumina de novo assemblies. *BMC Genomics* 2014; 15: 1103.
[http://dx.doi.org/10.1186/1471-2164-15-1103] [PMID: 25494611]
- [26] Kuhn ML, Falaschetti CA, Ballicora MA. *Ostreococcus tauri* ADP-glucose pyrophosphorylase reveals alternative paths for the evolution of subunit roles. *J Biol Chem* 2009; 284(49): 34092-102.
[http://dx.doi.org/10.1074/jbc.M109.037614] [PMID: 19737928]
- [27] Barchiesi J, Hedin N, Gomez-Casati DF, Ballicora MA, Busi MV. Functional demonstrations of starch binding domains present in *Ostreococcus tauri* starch synthases isoforms. *BMC Res Notes* 2015; 8: 613.
[http://dx.doi.org/10.1186/s13104-015-1598-6] [PMID: 26510916]
- [28] Barchiesi J, Hedin N, Iglesias AA, Gomez-Casati DF, Ballicora MA, Busi MV. Identification of a novel starch synthase III from the picoalgae *Ostreococcus tauri*. *Biochimie* 2017; 133: 37-44.
[http://dx.doi.org/10.1016/j.biochi.2016.12.003] [PMID: 28003125]
- [29] Hedin N, Barchiesi J, Gomez-Casati DF, Iglesias AA, Ballicora MA, Busi MV. Identification and characterization of a novel starch branching enzyme from the picoalgae *Ostreococcus tauri*. *Arch Biochem Biophys* 2017; 618: 52-61.
[http://dx.doi.org/10.1016/j.abb.2017.02.005] [PMID: 28235467]
- [30] Carrillo JB, Gomez-Casati DF, Martín M, Busi MV. Identification and analysis of OstaDSP, a phosphoglucan phosphatase from *Ostreococcus tauri*. *PLoS One* 2018; 13(1):e0191621
[http://dx.doi.org/10.1371/journal.pone.0191621] [PMID: 29360855]
- [31] Courties C, Perasso R, Chrétiennot-Dinet MJ, Gouy M, Guillou L, Troussellier M. Phylogenetic analysis and genome size of *Ostreococcus tauri* (Chlorophyta, Prasinophyceae). *J Phycol* 1998; 34(5): 844-9.
[http://dx.doi.org/10.1046/j.1529-8817.1998.340844.x]
- [32] Courties C VA, Troussellier M, Lautier J, et al. Smallest eukaryotic organism 1994.
[http://dx.doi.org/10.1038/370255a0]
- [33] Valdez HA, Busi MV, Wayllace NZ, Parisi G, Ugalde RA, Gomez-Casati DF. Role of the N-terminal starch-binding domains in the kinetic properties of starch synthase iii from *Arabidopsis thaliana*. *Biochemistry* 2008 Mar; 47(9): 3026-2.
- [34] Wayllace NZ, Valdez HA, Ugalde RA, Busi MV, Gomez-Casati DF. The starch-binding capacity of the noncatalytic SBD2 region and the interaction between the N- and C-terminal domains are involved in the modulation of the activity of starch synthase III from *Arabidopsis thaliana*. *FEBS J* 2010; 277(2): 428-40.
[http://dx.doi.org/10.1111/j.1742-4658.2009.07495.x] [PMID: 19968859]
- [35] Busi MV, Palopoli N, Valdez HA, et al. Functional and structural characterization of the catalytic domain of the starch synthase III from *Arabidopsis thaliana*. *Proteins* 2008; 70(1): 31-40.
[http://dx.doi.org/10.1002/prot.21469] [PMID: 17623838]
- [36] Gomez-Casati DF, Martín M, Busi MV. Polysaccharide-synthesizing glycosyltransferases and carbohydrate binding modules: The case of starch synthase III. *Protein Pept Lett* 2013; 20(8): 856-63.
[http://dx.doi.org/10.2174/0929866511320080003] [PMID: 23286550]
- [37] Deschamps P, Moreau H, Worden AZ, Dauvillée D, Ball SG. Early gene duplication within chloroplasts and its correspondence with relocation of starch metabolism to chloroplasts. *Genetics* 2008; 178(4): 2373-87.
[http://dx.doi.org/10.1534/genetics.108.087205] [PMID: 18245855]
- [38] Chrétiennot-Dinet M, Courties C, Vaquer A, et al. A new marine picoeukaryote: *Ostreococcus tauri* gen. et sp. nov. (Chlorophyta, Prasinophyceae). *Phycologia* 1995; 34(4): 285-92.
[http://dx.doi.org/10.2216/i0031-8884-34-4-285.1]
- [39] Ral JP, Derelle E, Ferraz C, et al. Starch division and partitioning. A mechanism for granule propagation and maintenance in the picophytoplanktonic green alga *Ostreococcus tauri*. *Plant Physiol* 2004; 136(2): 3333-40.
[http://dx.doi.org/10.1104/pp.104.044131] [PMID: 15448195]
- [40] Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970; 227(5259): 680-5.
[http://dx.doi.org/10.1038/227680a0] [PMID: 5432063]
- [41] Gomez Casati DF, Sesma JI, Iglesias AA. Structural and kinetic characterization of NADP-dependent, non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase from celery leaves. *Plant science : An int j experi plant biol* 2000 May; 29154(2): 107-5.
[PMID: 10729609]
- [42] Bollag DME, Rozycki SJ, Bollag MD, Rozycki R, Edelstein S. *Proteins methods* 1996.
- [43] Kerk D, Conley TR, Rodriguez FA, Tran HT, Nimick M, Muench DG, et al. A chloroplast-localized dual-specificity protein phosphatase in *Arabidopsis* contains a phylogenetically dispersed and ancient carbohydrate-binding domain, which binds the polysaccharide starch. *The Plant j : For Cell and Mol Biol* 2006 May; 46(3): 400-13.
[http://dx.doi.org/10.1111/j.1365-313X.2006.02704.x]
- [44] Valdez HA, Peralta DA, Wayllace NZ, Grisolia MJ, Gomez-Casati DF, Busi MV. Preferential binding of SBD from *Arabidopsis thaliana* SSIII to polysaccharides: Study of amino acid residues involved. *Starke* 2011; 63(8): 451-60.
[http://dx.doi.org/10.1002/star.201000111]
- [45] Valdez HA, Busi MV, Wayllace NZ, Parisi G, Ugalde RA, Gomez-Casati DF. Role of the N-terminal starch-binding domains in the kinetic properties of starch synthase III from *Arabidopsis thaliana*. *Biochemistry* 2008; 47(9): 3026-32.
[http://dx.doi.org/10.1021/bi702418h] [PMID: 18260645]
- [46] Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and

- PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 1997; 25(17): 3389-402. [http://dx.doi.org/10.1093/nar/25.17.3389] [PMID: 9254694]
- [47] Larkin MA, Blackshields G, Brown NP, *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* 2007; 23(21): 2947-8. [http://dx.doi.org/10.1093/bioinformatics/btm404] [PMID: 17846036]
- [48] Robert X, Gouet P. Deciphering key features in protein structures with the new END script server. *Nucleic Acids Res* 2014; 42 Web Server issue W320-4 [http://dx.doi.org/10.1093/nar/gku316] [PMID: 24753421]
- [49] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018; 35(6): 1547-9. [http://dx.doi.org/10.1093/molbev/msy096] [PMID: 29722887]
- [50] Labesse G, Mornon J. Incremental threading optimization (TITO) to help alignment and modelling of remote homologues. *Bioinformatics* 1998; 14(2): 206-11. [http://dx.doi.org/10.1093/bioinformatics/14.2.206] [PMID: 9545453]
- [51] Woo E-J, Lee S, Cha H, *et al.* Structural insight into the bifunctional mechanism of the glycogen-debranching enzyme TreX from the archaeon *Sulfolobus solfataricus* *J Biol Chem* 2008; 283(42): 28641-8. [http://dx.doi.org/10.1074/jbc.M802560200] [PMID: 18703518]
- [52] Sim L, Beeren SR, Findinier J, *et al.* Crystal structure of the Chlamydomonas starch debranching enzyme isoamylase ISA1 reveals insights into the mechanism of branch trimming and complex assembly. *J Biol Chem* 2014; 289(33): 22991-3003. [http://dx.doi.org/10.1074/jbc.M114.565044] [PMID: 24993830]
- [53] Song HN, Jung TY, Park JT, *et al.* Structural rationale for the short branched substrate specificity of the glycogen debranching enzyme GlgX. *Proteins* 2010; 78(8): 1847-55. [http://dx.doi.org/10.1002/prot.22697] [PMID: 20187119]
- [54] Lüthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. *Nature* 1992; 356(6364): 83-5. [http://dx.doi.org/10.1038/356083a0] [PMID: 1538787]
- [55] Wiederstein M, Sippl MJ. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 2007 Jul; 35 Web Server issue W407-10 [http://dx.doi.org/10.1093/nar/gkm290] [PMID: 17517781]
- [56] Sippl MJ. Recognition of errors in three-dimensional structures of proteins. *Proteins* 1993; 17(4): 355-62. [http://dx.doi.org/10.1002/prot.340170404] [PMID: 8108378]
- [57] Schrodinger LLC. The PyMOL Molecular Graphics System, Version 1.8 2015.
- [58] Davies GJ, Wilson KS, Henriissat B. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J* 1997; 321(2): 557-9. [http://dx.doi.org/10.1042/bj3210557] [PMID: 9020895]
- [59] MacGregor EA, Janeček S, Svensson B. Relationship of sequence and structure to specificity in the α -amylase family of enzymes. *Biochim Biophys Acta* 2001; 1546(1): 1-20. [http://dx.doi.org/10.1016/S0167-4838(00)00302-2] [PMID: 11257505]
- [60] Cuyvers S, Dornez E, Delcour JA, Courtin CM. Occurrence and functional significance of secondary carbohydrate binding sites in glycoside hydrolases. *Crit Rev Biotechnol* 2012; 32(2): 93-107. [http://dx.doi.org/10.3109/07388551.2011.561537] [PMID: 21711082]
- [61] Janeček Š, Gabriško M. Remarkable evolutionary relatedness among the enzymes and proteins from the α -amylase family. *Cell Mol Life Sci* 2016; 73(14): 2707-25. [http://dx.doi.org/10.1007/s00018-016-2246-6] [PMID: 27154042]
- [62] Southall SM, Simpson PJ, Gilbert HJ, Williamson G, Williamson MP. The starch-binding domain from glucoamylase disrupts the structure of starch. *FEBS Lett* 1999; 447(1): 58-60. [http://dx.doi.org/10.1016/S0014-5793(99)00263-X] [PMID: 10218582]
- [63] Tormo J, Lamed R, Chirino AJ, *et al.* Crystal structure of a bacterial family-III cellulose-binding domain: A general mechanism for attachment to cellulose. *EMBO J* 1996; 15(21): 5739-51. [http://dx.doi.org/10.1002/j.1460-2075.1996.tb00960.x] [PMID: 8918451]
- [64] Janeček Š, Svensson B, MacGregor EA. Structural and evolutionary aspects of two families of non-catalytic domains present in starch and glycogen binding proteins from microbes, plants and animals. *Enzyme Microb Technol* 2011; 49(5): 429-40. [http://dx.doi.org/10.1016/j.enzmictec.2011.07.002] [PMID: 22112614]
- [65] Hondoh H, Kuriki T, Matsuura Y. Three-dimensional structure and substrate binding of *Bacillus stearothermophilus* neopullulanase. *J Mol Biol* 2003; 326(1): 177-88. [http://dx.doi.org/10.1016/S0022-2836(02)01402-X] [PMID: 12547200]
- [66] Kim J-S, Cha S-S, Kim H-J, *et al.* Crystal structure of a maltogenic amylase provides insights into a catalytic versatility. *J Biol Chem* 1999; 274(37): 26279-86. [http://dx.doi.org/10.1074/jbc.274.37.26279] [PMID: 10473583]
- [67] Ball SG, Morell MK. From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. *Annu Rev Plant Biol* 2003; 54: 207-33. [http://dx.doi.org/10.1146/annurev.arplant.54.031902.134927] [PMID: 14502990]