Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

# Structural biology of starch-degrading enzymes and their regulation

Marie Sofie Møller[1,2] and Birte Svensson[1]

Starch is a major energy source for all domains of life. Recent advances in structures of starch-degrading enzymes encompass the substrate complex of starch debranching enzyme, the function of surface binding sites in plant isoamylase, details on individual steps in the mechanism of plant disproportionating enzyme and a self-stabilised conformation of amylose accommodated in the active site of plant α-glucosidase. Important inhibitor complexes include a flavonol glycoside, montbretin A, binding at the active site of human pancreatic α-amylase and barley limit dextrinase inhibitor binding to the debranching enzyme, limit dextrinase using a new binding mode for cereal protein inhibitors.

**Addresses**
[1] Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
[2] Center for Molecular Protein Science, Department of Chemistry, Lund University, 221 00 Lund, Sweden

Corresponding author: Svensson, Birte (bis@bio.dtu.dk)

**Current Opinion in Structural Biology** 2016, **40**:33–42

This review comes from a themed issue on **Carbobydrate–protein interactions and glycosylation**

Edited by **Nagasuma Chandra** and **Harry Gilbert**

http://dx.doi.org/10.1016/j.sbi.2016.07.006



## Introduction

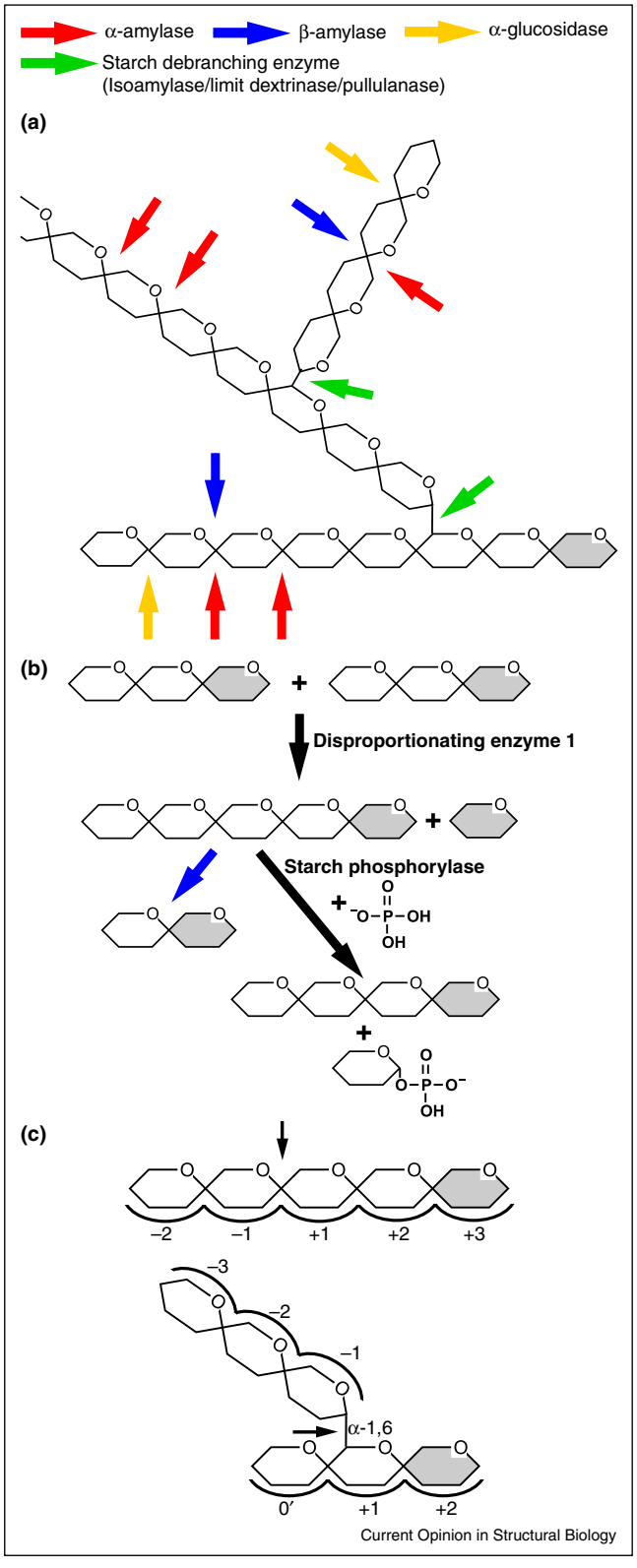
Starch is the most abundant storage polyglucan in nature serving as energy source for all domains of life. Accordingly starch degrading enzymes occur ubiquitously in plants, other eukaryotes, bacteria and archaea (see the carbohydrate-active enzymes database, CAZy; [1]). During the biosynthesis starch is deposited as supramolecular, semicrystalline granules basically consisting of two α-glucans, the branched amylopectin and the essentially linear amylose (see [2] for a recent review on starch). In plants starch can be either a transitory energy and building block supply synthesised and degraded in the diurnal cycle in leaves or a long-term storage residing in grains and tubers.

Starch is a simple polymer of α-1,4-linked and α-1,6-linked glucose residues, yet a multiplicity of enzyme activities is required for degradation of the complex starch granules as well as amylose and amylopectin alone (Figure 1). Starch-degrading enzymes include glycoside hydrolases (GHs), transglycosidases, glycosyl transferases (GTs) (phosphorylases), lyases, phosphatases and lytic polysaccharide monooxygenases (LPMOs). Here emphasis is on enzymes naturally involved in starch degradation, thus bacterial intracellular enzymes acting on glycogen are excluded although they degrade starch. Focus is on structures of new enzyme-types, new structures for previously characterized enzymes, and enzyme complexes with substrates, substrate analogues, natural products, and proteinaceous inhibitors. Starch-degrading enzymes according to CAZy [1] are found in GH families 3, 13, 14, 15, 31, 57, 119, eventually 126, auxiliary activities (AA) family 13 and GT family 35. α-Amylases, α-glucosidases, limit dextrinases/pullulanases, isoamylases, lytic polysaccharide monooxygenases, disproportionating enzymes, and starch phosphorylases will be explicitly addressed (see Table 1 for a complete overview). α-Glucan phosphatase is reviewed elsewhere in this issue [3].

The mentioned enzymes differ in catalytic domain fold, catalytic mechanism, and modular architecture encompassing auxiliary domains, primarily functionally and structurally discrete non-catalytic carbohydrate binding modules (CBMs) (Figure 2). Enzymes of the same family, for example, GH13, can possess differently shaped active sites; (i) a single shallow cleft spanning from a few to more than 10 glucosyl residues accommodating subsites for linear substrates in endo-acting α-amylases; (ii) a pocket shaped binding site extending in exo-glucanases and glucosidases from a catalytic site in a deep cavity and composed of two to several subsites; or (iii) in debranching enzymes two interconnected shallow clefts running in parallel.

Many of these multi-domain enzymes also contain dedicated carbohydrate binding sites that are situated outside of the active site area as seen for example on CBMs. Starch binding CBMs are organised in 12 families (CBM 20, 21, 25, 26, 34, 41, 45, 48, 53, 58, 68, 69; [1]). *In silico* 'domain-walking' using a predicted CBM may disclose starch-active catalytic domains in full length sequences as in case of LPMOs in family AA13 [4•] for which a crystal structure was recently solved [5••]. Carbohydrate binding areas are also identified in form of surface surface binding sites (SBSs) situated on catalytic or intimately associated domains (Figure 2). Identification and functional characterisation of SBSs typically start with its observation in an

**Figure 1**




Schematic overview of the activity of the starch-active enzymes reviewed. The glucose units (hexagons) of the linear maltooligosaccharides are connected by α-1,4-glucosidic linkages, while the branch points are α-1,6-linkages. The reducing ends are indicated by a grey hexagon. **(a)** Starch degradation in general and in particular in the germinating seed of cereals (the only starch debranching enzyme active is limit dextrinase). **(b)** Action of the disproportionating enzyme 1 and associated enzymes during degradation of transitory starch. **(c)** Schematic illustration of general subsite nomenclature for α-1,4-active enzymes (top) and debranching enzymes (bottom), where each subsite interacts with one glucosyl unit and the number of subsites can vary
*Source*: adapted from [66,67].

enzyme:carbohydrate complex [6]. SBSs are difficult to predict even though several starch-active enzymes contain SBSs with validated roles in activity [7]. Recently, SBSs were also characterised in starch synthase I [8,9] and in glycogen synthase [10,11], where they ensure enzyme proximity to the α-glucan molecule during its elongation. SBSs have also been identified in α-glucan phosphatases (reviewed in [3]).

## α-Amylases

According to CAZy, α-amylase (EC 3.2.1.1), probably the best studied amylolytic enzyme, is found in GH13, GH57, GH119, and GH126 (α-amylase activity not fully confirmed) [1,12]. Eleven GH13 subfamilies contain α-amylases: GH13_1 (fungi), GH13_5 (bacterial liquefying enzymes), GH13_6 (plants), GH13_7 (archaea), GH13_15 (insects), GH13_24 (animals), GH13_27 (proteobacteria), GH13_28 (bacterial saccharifying enzymes), GH13_32 (bacteria), GH13_36 (intermediary α-amylase group evolutionary found between oligo-1,6-glucosidases and neopullulanases), GH13_37 (marine bacteria), and GH13_41 (starch degrading enzymes with both an α-amylase and a pullulanase domain [13]) [12,14]. Twenty-three different α-amylases are structure-determined including enzymes from GH13_5, GH13_6, and GH57 published recently [1,15,16•,17•]. The industrial GH13_5 *Bacillus* and *Geobacillus* α-amylases (known as 'Termamyl'-like α-amylases after the trade name of a *Bacillus licheniformis* enzyme) are represented by several structures [18–24], the newest being of an engineered variant of *Geobacillus stearothermophilus* GH13_5 α-amylase (PDB: 4UZU; parent PDB: 1HVX [20]) showing a tighter turning loop due to a two-residue truncation in the three-domain fold stabilised by $Ca^{2+}$ and the $Ca^{2+}$–$Na^{+}$–$Ca^{2+}$ triad of Termamyl-like α-amylases and a sodium ion mimicking the substrate transition-state charge (Figure 2a) [15].

The GH13_6 α-amylases are represented by structures of SBS-containing barley (*Hordeum vulgare*) isozymes 1 [25–28] and 2 [29,30], and the *N*-glycosylated rice (*Oryza sativa*) α-amylase AmyI-1 (PDB: 3WN6) that acts at an early stage of seed germination as well as in leaves and has tartaric acid from the crystallisation binding at an SBS conserved in barley α-amylase (Figure 2b) [16•]. Mutations of this SBS impeded AmyI-1 plastid targeting [31].