Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Structural biology of glucan phosphatases from humans to plants

Matthew S Gentry, M Kathryn Brewer and Craig W Vander Kooi

Glucan phosphatases are functionally conserved at the enzymatic level, dephosphorylating glycogen in animals and starch in plants. The human glucan phosphatase laforin is the founding member of the family and it is comprised of a carbohydrate binding module (CBM) domain followed by a dual specificity phosphatase (DSP) domain. Plants encode two glucan phosphatases: Starch EXcess4 (SEX4) and Like Sex Four2 (LSF2). SEX4 contains a DSP domain followed by a CBM domain, while LSF2 contains a DSP domain and lacks a CBM. This review demonstrates how glucan phosphatase function is conserved and highlights how each family member employs a unique mechanism to bind and dephosphorylate glucan substrates.

Address
Department of Molecular and Cellular Biochemistry and Center for Structural Biology, University of Kentucky, Lexington, KY 40536, United States

Corresponding authors: Gentry, Matthew S (matthew.gentry@uky.edu) and Vander Kooi, Craig W (craig.vanderkooi@uky.edu)

Current Opinion in Structural Biology 2016, **40**:62–69

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

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

http://dx.doi.org/10.1016/j.sbi.2016.07.015

Introduction

Animals and plants both utilize glucans comprised of α-1,4 linked glucose chains with α-1,6 branch points as energy storage macromolecules. Glycogen is water-soluble and is utilized by opisthokonts (e.g. animals and fungi) and most prokaryotes, while starch is water-insoluble and is utilized by Archaeplastida (e.g. algae and land plants) and eukaryotes with Archaeplastida-derived plastids. Glycogen and starch each can be enzymatically modified by the addition of phosphate and this phosphoryl group links research on a human epilepsy with plant starch metabolism.

Glycogen and Lafora disease

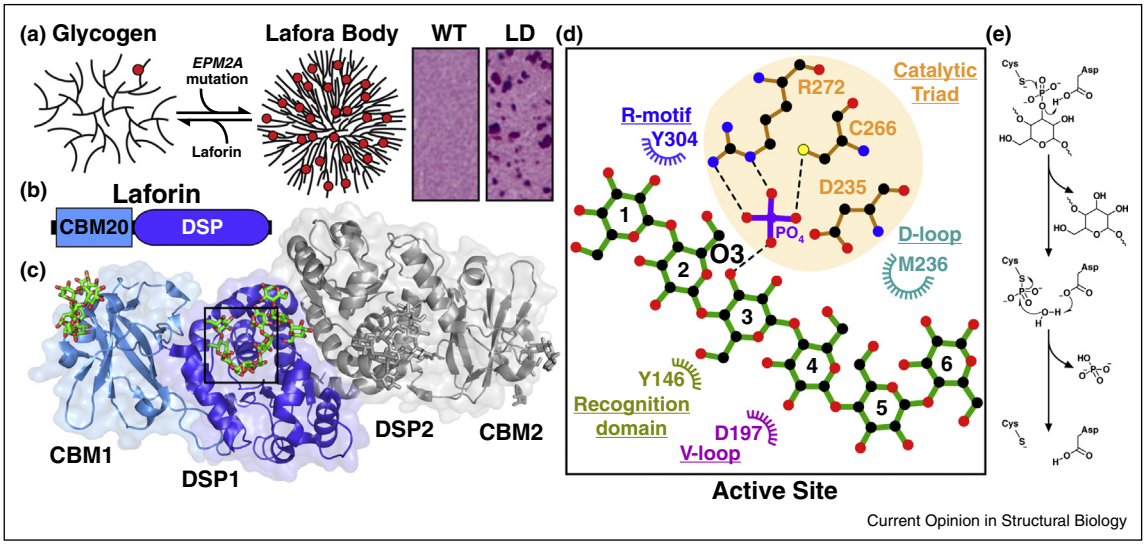
Glycogen, the major mammalian storage carbohydrate, has a critical function in energy metabolism [1]. Glycogen is typically modeled with an average of 13 glucose residues per chain and successive layers of α-1,6 branches up to 12 tiers total (Figure 1a). This unique organization allows cells to store up to ≈55 000 glucose units/molecule in a water-soluble form allowing rapid glucose mobilization. Recent discoveries from the study of Lafora disease have enabled identification of glycogen phosphorylation as an important regulatory mechanism with direct implications for human disease.

In 1911, Gonzalo Lafora, a student of Cajal and Alzheimer, reported cases of teen-age myoclonus epilepsy with dementia and elegantly described 'amyloid bodies' in patient neurons (Figure 1a) [2,3]. While amyloid now is used to refer to proteinaceous deposits, it originally referred to any material that stained in a manner similar to starch. The disease identified by Lafora was later called Lafora disease (LD) and the starch-like inclusions Lafora bodies (LBs).

LD is a fatal, autosomal recessive, neurodegenerative epilepsy. The majority of cases are caused by mutations in the *epilepsy, progressive myoclonus 2A* (*EPM2A*) gene [4]. *EPM2A* encodes the bimodular protein laforin that contains a carbohydrate binding module-family 20 (CBM20) domain followed by a dual specific phosphatase (DSP) domain (Figure 1b) [4,5•,6•]. CBMs are a divergent family of domains that allow proteins to bind an array of different carbohydrates with the CBM20 family being specific for glycogen and starch [7]. DSPs belong to the protein tyrosine phosphatase (PTP) superfamily, a family of ~105 phosphatases that dephosphorylate proteinaceous or non-proteinaceous substrates regulating multiple signaling cascades [8]. Driven by the observation that many of the enzymes involved in glycogen metabolism are regulated by reversible phosphorylation, including the 1992 Nobel Prize for the discovery of reversible phosphorylation of glycogen phosphorylase, it was initially hypothesized that laforin's function was to dephosphorylate an enzyme(s) involved in glycogen metabolism.

Surprisingly, insights from starch metabolism led to the discovery of laforin's function. Austin and coworkers demonstrated that LBs are glucose polymers more similar to starch than glycogen, and identified ~10-fold more phosphate in LBs than in normal glycogen [9,10]. Dixon and colleagues were struck by the realization of the aberrant glycogen metabolism in LD patients, the increased phosphate on LBs, and that the mutated *EPM2A* gene encoded a phosphatase with an unknown substrate.

Figure 1



(a) Glycogen model depicting 4 of the possible 12 tiers. Red spheres represent phosphate moieties. Laforin is a glycogen phosphatase encoded by the *EPM2A* gene. Mutations in *EPM2A* result in the hyperphosphorylation of glycogen that changes the normal branching pattern of glycogen and results in a Lafora body (LB). LBs are inclusions present in neurons of LD patients and stain Periodic acid-Schiff (PAS) positive. The micrographs are PAS stain slices of wildtype (WT) and *Epm2a*−/− (LD) mice hippocampus. Modified from [58] with permission. **(b)** Laforin is composed of an N-terminal carbohydrate binding module (CBM) and a C-terminal dual-specificity phosphatase (DSP) domain. **(c)** The crystal structure of the laforin dimer bound to maltohexaose and phosphate (PDB: 4RKK). The CBM and DSP of one subunit are shown in light blue and dark blue, respectively; the second subunit is shown in grey. The quaternary arrangement of the two subunits results in a CBM–DSP–DSP–CBM architecture. Structure figures were generated using Pymol [59]. **(d)** Ligplot+ generated representation of hydrophobic and hydrophilic protein−glucan interactions at the DSP domain active site [60]. Maltohexaose is shown in green and hydrogen bonds are shown in black. Glucose moieties are numbered from the non-reducing end to the reducing end. Shared glucan phosphatase-family DSP motifs contributing to glucan binding and their corresponding key residues are displayed. Multiple elements of the laforin DSP domain contribute to glucan binding at the active site. Both the recognition domain and the variable loop (V-loop) interact with the convex side of the glucan chain with residues of the R-motif, catalytic triad, and D-loop all interacting with the concave face. **(e)** The reaction mechanism for glucan dephosphorylation is proposed based on the well-characterized mechanism of PTPs [8,61]. The first step of catalysis involves nucleophilic attack by the sulfur atom of the active site Cys, which is believed to exist in thiolate form. The highly conserved Asp functions as a general acid, donating its proton to the substrate leaving group, and a cysteinyl-phosphate intermediate is formed. In the second step of catalysis, the same Asp functions as a general base, accepting a proton from an attacking water molecule, which hydrolyzes the phosphoenzyme intermediate.

They demonstrated that laforin could remove phosphate from glucans and hypothesized that laforin is a glycogen phosphatase [11••]. Subsequent work demonstrated that laforin does remove phosphate from glycogen, that glycogen becomes hyperphosphorylated in the absence of laforin activity, that this hyperphosphorylated glycogen becomes a LB, and that LBs are the pathogenic cause of LD (Figure 1a) [12,13,14••,15,16]. Recent structural studies have elucidated the basis for the function of laforin as a glucan-specific phosphatase, and provided important insights into the molecular basis of LD.

Laforin possesses a cooperative tetramodular glucan-binding platform

We determined the structure of the human laforin dimer bound to maltohexaose and phosphate to 2.4 Å (Figure 1c) [17••]. Maltohexaose is a linear, six-glucose polysaccharide with α-1,4 glycosidic linkages, representative of a linear glycogen chain. Laforin engages a maltohexaose chain within its CBM and DSP domain active site via aromatic and hydrophilic residues.

The laforin CBM and DSP domains form an integrated tertiary structure with >1200 Å^2 of interfacial surface area. This interdomain interface is a hot spot for disease mutations, suggesting domain coupling is critical for laforin function. Further, laforin was analyzed by multiple biophysical techniques to confirm that it is a dimer in solution. Laforin forms an antiparallel dimer with a tetramodular CBM–DSP–DSP–CBM architecture (Figure 1c). The four glucan binding sites are oriented on one face of the dimer. Binding experiments demonstrated that shorter oligosaccharide chains, such as maltohexaose, bind without any observed cooperativity, consistent with the multiple distinct chains bound in the crystal structure. Strikingly, longer chains showed a single cooperative high-affinity binding. Importantly, the disease mutation (laforin-F321S) was shown to disrupt dimerization and abolish cooperative glucan binding. Thus, structural