The *Toxoplasma* glucan phosphatase TgLaforin utilizes a distinct functional mechanism that can be exploited by therapeutic inhibitors

Robert D. Murphy[1,2,‡], Tiantian Chen[1,‡], Jianping Lin[3,‡], Rongjun He[3], Li Wu[3], Caden R. Pearson[1], Savita Sharma[1], Carl D. Vander Kooi[1], Anthony P. Sinai[2], Zhong-Yin Zhang[3,*], Craig W. Vander Kooi[1,*], and Matthew S. Gentry[1,*]
From the [1]Department of Molecular and Cellular Biochemistry, and [2]Department of Microbiology, Immunology, and Molecular Genetics, College of Medicine, University of Kentucky, Lexington, Kentucky, USA; [3]Departments of Medicinal Chemistry and Molecular Pharmacology and of Chemistry, Purdue Institute for Drug Discovery, Purdue University, West Lafayette, Indiana, USA

Edited by Roger Colbran

Toxoplasma gondii is an intracellular parasite that generates amylopectin granules (AGs), a polysaccharide associated with bradyzoites that define chronic *T. gondii* infection. AGs are postulated to act as an essential energy storage molecule that enable bradyzoite persistence, transmission, and reactivation. Importantly, reactivation can result in the life-threatening symptoms of toxoplasmosis. *T. gondii* encodes glucan dikinase and glucan phosphatase enzymes that are homologous to the plant and animal enzymes involved in reversible glucan phosphorylation and which are required for efficient polysaccharide degradation and utilization. However, the structural determinants that regulate reversible glucan phosphorylation in *T. gondii* are unclear. Herein, we define key functional aspects of the *T. gondii* glucan phosphatase TgLaforin (TGME49_205290). We demonstrate that TgLaforin possesses an atypical split carbohydrate-binding-module domain. AlphaFold2 modeling combined with hydrogen–deuterium exchange mass spectrometry and differential scanning fluorimetry also demonstrate the unique structural dynamics of TgLaforin with regard to glucan binding. Moreover, we show that TgLaforin forms a dual specificity phosphatase domain–mediated dimer. Finally, the distinct properties of the glucan phosphatase catalytic domain were exploited to identify a small molecule inhibitor of TgLaforin catalytic activity. Together, these studies define a distinct mechanism of TgLaforin activity, opening up a new avenue of *T. gondii* bradyzoite biology as a therapeutic target.

Toxoplasma gondii is an obligate, intracellular, protozoan parasite that can infect any warm-blooded animal and chronically infects one-third of humans worldwide (1). Infection with the parasite occurs through several routes, including ingestion of environmental oocysts shed through cat feces, consumption of tissue cysts in raw or undercooked meat, or through vertical transmission from an acutely infected pregnant individual to their fetus (2, 3). Upon consumption, the cyst forms (oocysts or tissue cysts) of the parasite eventually convert into rapidly dividing tachyzoites that disseminate throughout the body of the host, defining the acute stage of infection (4). Under host immune pressure, a small number of tachyzoites evade host defenses by converting into bradyzoites that populate tissue cysts located in the central nervous system or muscle tissue of the host (5, 6). If the host becomes immunocompromised during the course of the life-long chronic infection, spontaneous reactivation of bradyzoites into tachyzoites can result in symptomatic toxoplasmosis that primarily manifests as the life-threatening toxoplasmic encephalitis (7). The current first-line treatment is a combination of pyrimethamine and sulfadiazine that targets folate synthesis in tachyzoites during the acute phase of infection or after reactivation. However, this treatment regime only targets tachyzoites, results in serious side effects, must be taken for long periods of time, and cannot be administered to pregnant individuals (8). Notably this drug combination does not reduce the overall cyst burden. Other drugs targeting the parasite mitochondrion result in a partial clearance of tissue cysts, without achievement of a sterile cure. The potential for amylopectin granule (AG) metabolism as a target for drug intervention provides a unique molecular target for *T. gondii* therapeutics that has not been explored.

A distinguishing feature of *T. gondii* bradyzoites is their accumulation of large glucose-based polymers known as AGs (9). AGs are believed to function as an insoluble form of glucose storage for use in bradyzoite persistence, replication, transmission, and reactivation, analogous to starch utilization during the night in plants (10, 11). Purified AGs were found to be similar to plant amylopectin, the major glucose component of starch, with respect to glucose chain-length and branching frequency, demonstrating that the structure of AGs is more similar to plant starch than to animal and fungal glycogen (9, 12, 13). Unexpectedly, it was demonstrated that AGs are synthesized from UDP-glucose, similar to the sugar nucleotide used in fungal and animal glycogen synthesis, rather than ADP-glucose, which is used by plants (13). Additionally,

T. gondii AGs reside in the cytoplasm, which is also characteristic of fungal and animal glycogen and are visible by transmission electron microscopy (TEM) in encysted bradyzoites but not the rapidly dividing tachyzoites (Fig. 1*A*). The presence of AGs in the cytoplasm is noteworthy given that *T. gondii* retains a plastid remnant organelle termed the apicoplast that has its origins in red alga.

These seemingly disparate findings regarding AGs are consistent with the fact that the *T. gondii* proteome contains a hybrid molecular machinery exhibiting characteristics of plants, fungi, and animals (14, 15). This hybrid machinery is exemplified in *T. gondii*'s putative cycle of reversible glucan phosphorylation (Fig. 1*B*). In plants, the cycle begins with the actions of glucan, water dikinase (GWD) and phosphoglucan, water dikinase (PWD) that use ATP to phosphorylate glucose at the C6- and C3-hydroxyl positions within the glucan, respectively (16, 17). The presence of covalently attached phosphate disrupts the crystalline organization of starch, thus solubilizing the surface glucose chains and enhancing access of multiple amylases that release glucose, maltose, or other oligosaccharides (18–20). To allow processive hydrolysis, the activity of a glucan phosphatase is required to remove the residual phosphate groups and reset the cycle (21–23). The *T. gondii* genome contains sequences that encode putative enzymes with these activities: a glucan, water-dikinase (TgGWD [TGME49_214260]; plant-like), multiple amylases and debranching enzymes, and the glucan phosphatase *T. gondii* laforin (TgLaforin [TGME49_205290]; animal-like), along with all of the enzymes needed in glucan synthesis (Figs. 1*B* and S1*A*) (13, 24–26).

The mechanism regulating the utilization of AGs in the *T. gondii* lifecycle is only beginning to be defined. The critical regulatory role of reversible phosphorylation of both starch and glycogen has been increasingly recognized, leading to potentially unique insights and molecular targets in the *T. gondii* system (27, 28). Notably, glucan phosphatases are a family of carbohydrate-specific enzymes that play a critical role in controlling polyglucan utilization in both plants and animals. Glucan phosphatases are required for efficient starch catabolism as loss of *starch excess4* (SEX4), the prototypical glucan phosphatase found in plants, leads to aberrant starch accumulation and morphology (22, 29). Strikingly, mutations in human laforin (*Hs*Laforin), the glucan phosphatase in humans, results in hyperphosphorylated and aberrantly branched glycogen that is the driver of Lafora disease, a fatal epilepsy and childhood dementia (30–32).

Glucan phosphatases are an enzyme family that dephosphorylate glucans *via* a dual specificity phosphatase (DSP) domain coupled with either a carbohydrate-binding module (CBM) domain or carbohydrate surface binding sites (23, 24). Glucan phosphatases are members of the protein tyrosine phosphatase (PTP) superfamily within the DSP clade (33–35). Previous work has demonstrated that glucan phosphatases employ a wide variety of platforms that allow them to bind and dephosphorylate carbohydrates (36–39). In land plants, SEX4 binds and dephosphorylates carbohydrates by utilizing an integrated binding pocket formed by the CBM/DSP interface (40–42). Like Sex Four-2, LSF2, another plant glucan phosphatase, lacks a traditional CBM and instead binds carbohydrates through surface binding sites within its DSP domain (43). Human laforin is an antiparallel dimer with a spatially separated CBM and DSP that are each capable of binding to carbohydrates independently (44). While they are not orthologs, SEX4 and laforin are functional equivalents as complementation of laforin into Δ*sex4 Arabidopsis thaliana* plants rescues the starch excess phenotype (24).

In this study, we build on our understanding of glucan phosphatases and reversible phosphorylation in *T. gondii* by defining the biophysical and biochemical properties of TgLaforin. We identified an atypical CBM in TgLaforin, defined TgLaforin's oligomerization status, and determined its activity. Moreover, we developed and characterized the first reported TgLaforin inhibitor. These findings demonstrate the value of a detailed examination of the enzymology related to glucan
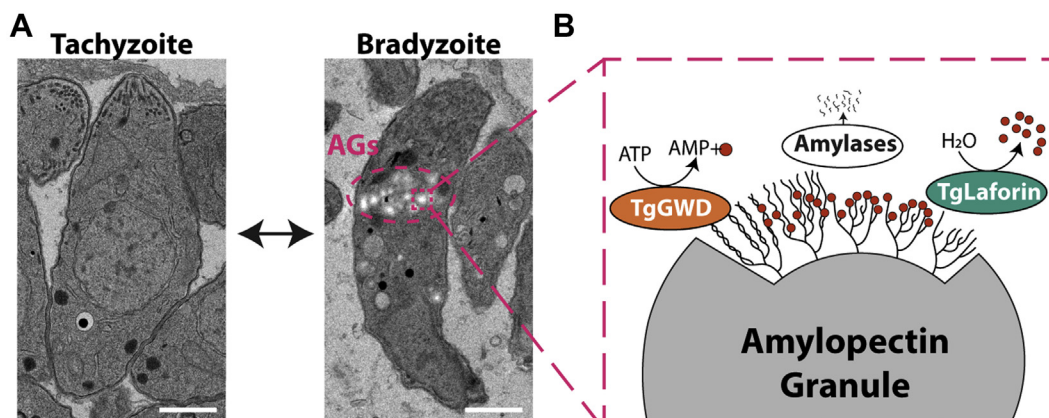

Figure 1. *Toxoplasma gondii* contains a hybrid starch metabolism machinery. *A*, *T. gondii* tachyzoites contain a smaller, highly turned-over glucan in their cytoplasm. Tissue culture induced bradyzoites, on the other hand, contain insoluble AGs that can be visualized using TEM. Scale bar = 1 μm. *B*, schematic of proposed AG reversible phosphorylation in *T. gondii* is representative of the hybrid approach utilized in AG degradation: a GWD kinase-activity (plant-like) opposes a laforin glucan phosphatase (animal-like). The schematic depicts TgGWD utilization of the β-phosphate in ATP for glucan phosphorylation alongside the release of AMP and the γ-phosphate (*red circles* represent free or glucan-bound phosphate). Amylases are then able to efficiently release oligosaccharides (*small black lines*) until a glucan phosphatase (TgLaforin) is required to remove phosphate to reset the cycle and allow for processive glucan hydrolysis. AG, amylopectin granule; GWD, glucan, water-dikinase.

metabolism in *T. gondii* and open new doors for anti-*Toxoplasma* therapeutics.

Results

Biophysical evidence for a split-CBM20 in TgLaforin

T. gondii AG metabolism employs reversible glucan phosphorylation where the plant-like TgGWD dikinase and animal-like TgLaforin phosphatase are utilized (Fig. 1*B*). In addition to this plant-like *versus* animal like axis, TgLaforin is also unique in its domain organization. We previously predicted that the *T. gondii* genome encodes for a glucan phosphatase that is more similar to human laforin than other glucan phosphatases based on its domain orientation with an N-terminal CBM family 20 (CBM20) followed by a DSP domain (24, 25). However, TgLaforin contains 523 amino acids, and the combined length of a typical CBM (90–120 amino acids) and DSP (150 amino acids) is only 240 to 270 amino acids, leaving approximately half of the protein uncharacterized. Bioinformatic analyses suggested that TgLaforin contains two unusual inserts within its CBM20 domain (Fig. 2*A*) (45).

To define the architecture of TgLaforin's CBM, we utilized AlphaFold2 (46) to generate a predicted three-dimensional model of TgLaforin. The model generated by AlphaFold2 predicted that TgLaforin contains a single CBM20 over the span of 325 amino acids. With high confidence (predicted Local Distance Difference Test [pLDDT] score > 70), Alpha-Fold2 modeled a core CBM that consists of β-strands from the three distally located subdomains of the CBM region. The model predicts that the unusual inserts form large unstructured linker regions connecting the core structured CBM regions (Fig. 2*B*). The unstructured portions of the CBM20 modeled with low confidence (pLDDT < 50) (Fig. 2*B*). A canonical DSP domain (reviewed in (33)) was also modeled in TgLaforin containing the characteristic DSP α+β structure with a central β-sheet consisting of five twisted β-strands (Fig. S1*B*). TgLaforin contains the canonical CX_5R motif with C452 serving as the catalytic cysteine (Fig. S1, *B* and *C*), and R458 that coordinates the three nonbridging oxygens of the incoming phosphoryl substrate. This catalytic motif is located between a β-strand and α-helix as expected. DSP domains also contain a critical, upstream aspartate (D421 in TgLaforin) that defines the D-loop and functions as an acid–base catalyst,
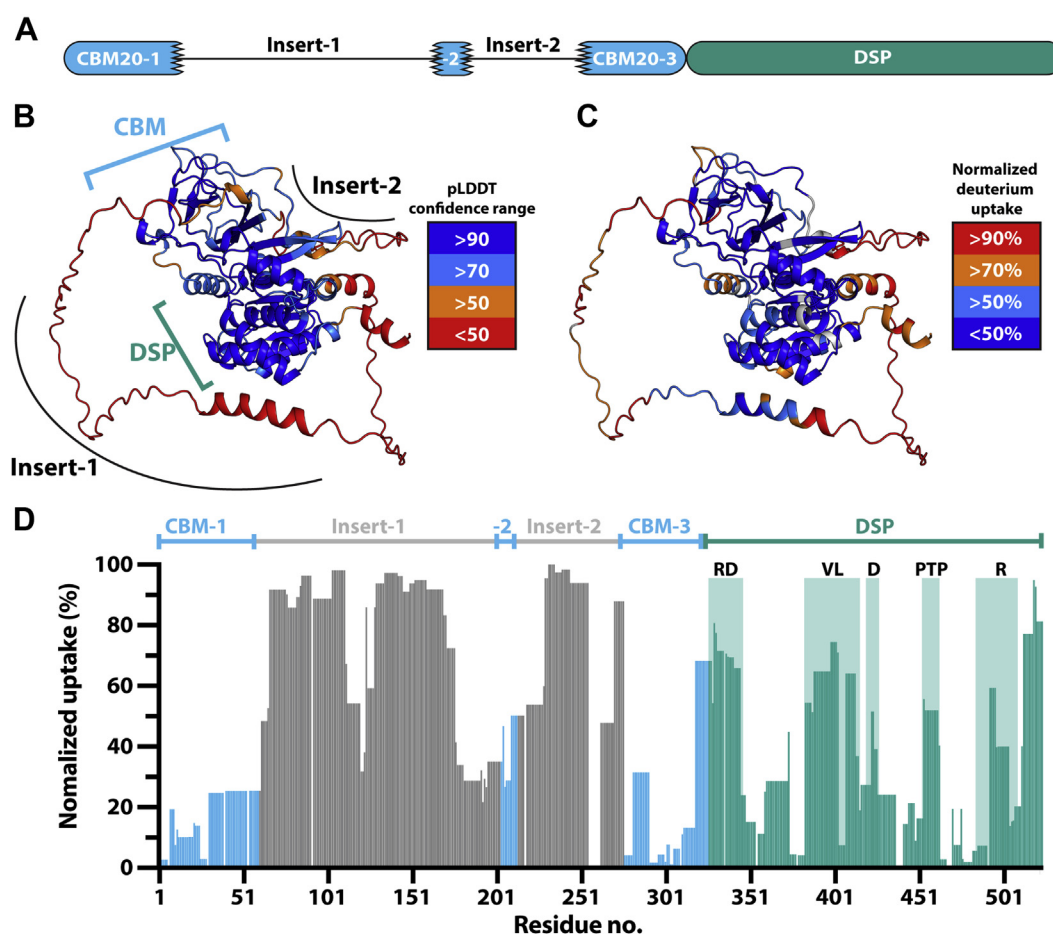

Figure 2. Modeling and HDX data indicate that TgLaforin contains a split-CBM20. *A*, schematic of a previously proposed domain arrangement of TgLaforin (45). *B*, TgLaforin modeled by AlphaFold2 with confidence levels mapped onto the structure. CBM, DSP, and inserts are indicated. *C*, deuterium uptake of TgLaforin after 15 s as determined by HDX-MS mapped onto the AlphaFold2 model. *D*, normalized deuterium uptake by TgLaforin after 15 s deuteration represented in 2D, highlighting uptake in each domain of the protein. Other time-points from 15 s to 15,000 s displayed in Fig. S3. CBM, carbohydrate-binding module; D, D-loop; DSP, dual-specificity phosphatase domain; HDX-MS, hydrogen–deuterium exchange mass spectrometry; PTP, protein tyrosine phosphatase loop; R, R-motif; RD, recognition domain; VL, variable loop.

modeled in TgLaforin as participating in the active site (Fig. S1*B*). The presence and location of these residues place TgLaforin in the PTP superfamily (47).

To validate the AlphaFold2 model of TgLaforin, we expressed and purified full-length TgLaforin in Sf9 insect cells to >95% homogeneity (Fig. S2*A*) and utilized the recombinant protein in hydrogen–deuterium exchange (HDX) mass spectrometry (MS) experiments. This technique has previously been used to illuminate conformational dynamics of proteins by demonstrating areas of high and low solvent accessibility (44, 48). We first determined deuterium uptake across the entire protein from 15 to 15,000 s (Fig. S3) and mapped the results onto the AlphaFold2 model (Fig. 2, *C* and *D*). Critically, the core CBM20 β-sandwich exhibited low uptake of deuterium, while the CBM insert regions that were modeled with low confidence displayed much higher rates of deuterium uptake (Fig. 2, *C* and *D*). These regions with high deuterium exchange are thus more mobile and less structured. Conversely, β-sheets within the core CBM were among the structural elements with the lowest deuterium uptake. Within the DSP domain, core structural elements had low uptake, while elements of the DSP active site, known to undergo conformational changes required for substrate interaction (44, 49), displayed higher uptake. These active site regions included the recognition domain, variable loop, D-loop, PTP-loop, and R-motif (Figs. 2*D* and S3) that have previously been reported as exhibiting higher solvent accessibility in SEX4 and laforin (44, 50).

Integrated CBM is required for carbohydrate binding

To further define the nature of the unique split-CBM domain found in TgLaforin, we aligned the TgLaforin CBM with other glucan phosphatases containing a CBM20 (Fig. 3*A*). The three regions of the TgLaforin CBM20-domain, split over 325 amino acids and interrupted by linker regions, each contain a part of the key consensus amino acids predicted to be critical to glucan binding. However, none of the individual regions contain a complete array of consensus amino acids that would allow it to be an individual CBM domain. Importantly, these three subdomains are predicted to contain the total number of β-strands that are typically identified in a CBM. These β-strands are predicted to come together to form two anti-parallel β-sheets, comprising the canonical β-sandwich fold that is characteristic of CBM20 domains (Fig. 3*B*). Indeed, the core CBM20 in the TgCBM is structurally similar to other CBM20 domains. A comparison of the TgCBM to all structures in the Dali server identified the top three most similar structures as: (1) HsLaforin (4RKK), RMSD = 1.7 (44); (2) glucoamylase from *Hypocrea jecorina* (2VN4), RMSD = 2.2 (51); and (3) glucanotransferase from *Paenibacillus macerans* (4JCL), RMSD = 2.1.

Each subdomain contains one of the three conserved amino acids involved in either aromatic glucose stacking interactions (W29 and W278) or hydrogen bonding with glucose rings (K206) that are each critical to glucan binding (52). Conversely, the linker regions do not contain the key conserved glucan-binding residues. Importantly, the three

critical carbohydrate-binding residues (W29, K206, and W278) are predicted to be organized into a single three-dimensional binding site similar to other CBM20s (Figs. 3*C* and S2*B*). To further test if TgLaforin contains a single integrated CBM, we expressed each of the three subdomains (CBM20-1, CBM20-2, and CBM20-3) separately and in different combinatorial arrangements (Fig. S2*C*). Each of the constructs encoding individual portions of this region were either not expressed, not folded, or aggregated (Fig. S2, *D* and *E*). However, we were able to express and purify the predicted full CBM comprising all three subdomains with the two linkers, TgCBM123 (referred to simply as "TgCBM") (Fig. S2*E*).

We next analyzed TgCBM using size-exclusion chromatography coupled with multiangle light scattering (SEC-MALS) to define its oligomerization state. SEC-MALS analysis demonstrated that TgCBM has a molecular weight (MW) of 36.6 ± 6.8 kDa, compared to the predicted monomeric mass of 35.6 kDa, indicating that the TgCBM is a stable monomer in solution (Fig. 3*D*). To define TgCBM carbohydrate binding, we utilized differential scanning fluorimetry (DSF) (53, 54). In the presence of a carbohydrate, an increase in T_m is associated with increased protein stability due to protein–ligand interaction. We incubated TgCBM with the linear oligosaccharide maltohexaose and observed a robust dose-dependent stabilization of TgCBM with an $K_{d,app}$ = 20 ± 4 mM (Fig. 3*E*).

To further probe the glucan binding of TgCBM, we designed alanine mutants of the three predicted critical carbohydrate-binding residues. K206A and W278A were produced and purified to near homogeneity, while W29A was deleterious to protein folding. We then tested the effect of the K206A and W278A mutations on carbohydrate binding. Strikingly, neither mutant was able to bind to maltohexaose, demonstrating their key function in carbohydrate binding (Fig. 3*E*). If TgLaforin contained two distinct CBMs, then carbohydrate binding should have been only partially reduced by each independent mutant. Taken together, these data demonstrate the structure, dynamics, and carbohydrate binding of the split CBM20 of TgLaforin.

TgLaforin is an antiparallel dimer

Previous work demonstrated that glucan phosphatases utilize unique platforms to bind and dephosphorylate carbohydrates (37). While the plant glucan phosphatases SEX4 and LSF2 are monomeric, dimerization is a critical feature of human laforin activity and stability, and dimerization is mediated through the DSP domain resulting in an antiparallel dimer (44). To determine the glucan binding platform employed by TgLaforin, we analyzed the oligomerization status of full-length TgLaforin and found that TgLaforin eluted as a single species from SEC (Fig. 4*A*). However, like other glucan phosphatases, it was significantly shifted in elution volume due to interactions with the carbohydrate-based SEC matrix and eluted at a higher volume than would be expected when compared to SEC standards (55). We therefore utilized SEC-MALS to determine the oligomeric state of TgLaforin. SEC-MALS demonstrated that TgLaforin
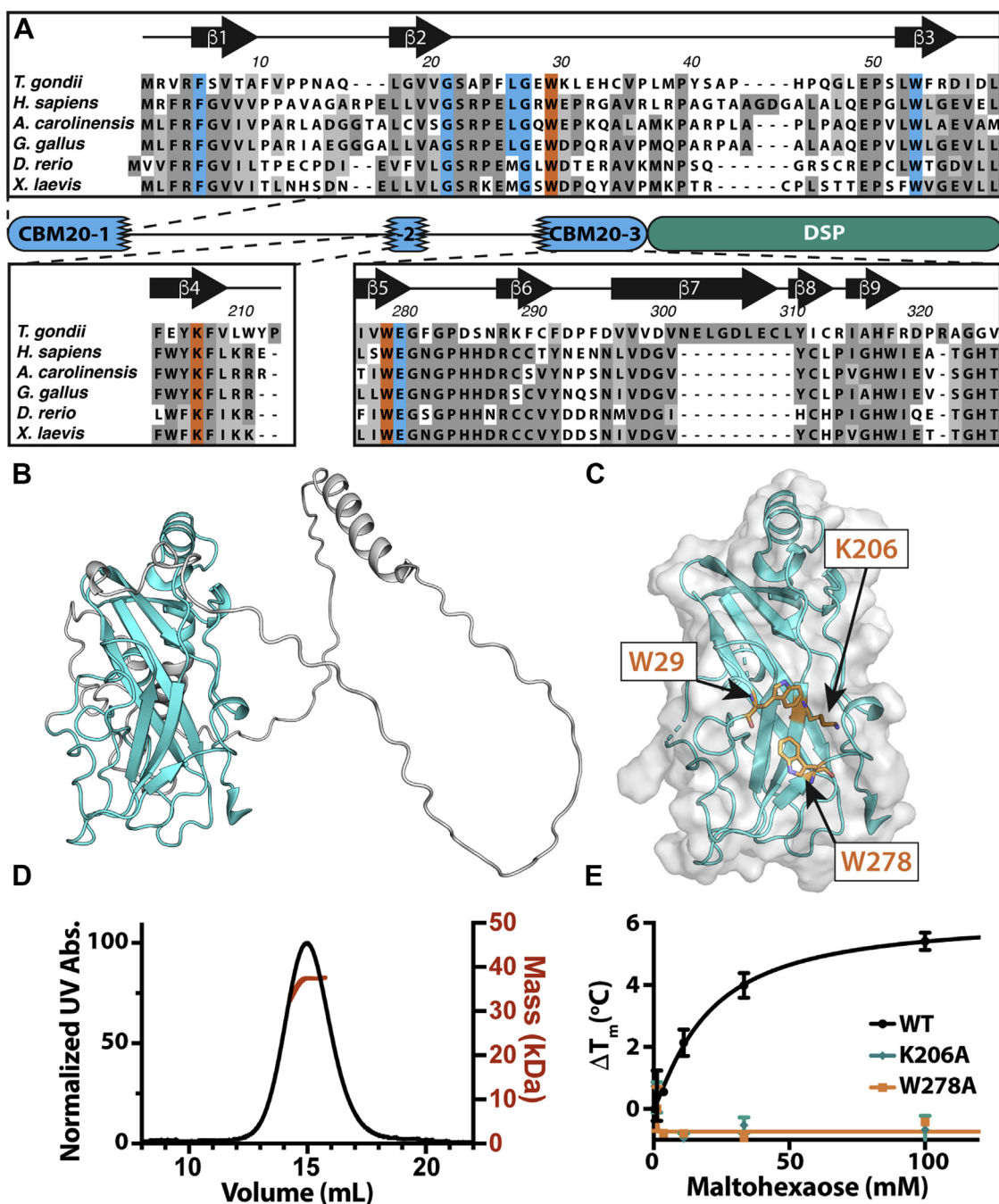
Figure 3. Involvement of complete TgCBM is required to bind carbohydrates. *A*, multiple sequence alignment of laforin orthologue CBMs found across multiple vertebrate species, overlaid with the predicted domain structure of TgLaforin. *Black lines* connecting CBM20 fragments are predicted to be largely unstructured loops. Residues highlighted in *blue* are involved in carbohydrate-binding specificity, and residues in *orange* are consensus residues needed to engage carbohydrates. *Arrows* symbolize β-strands, and numbering corresponds to the TgLaforin sequence. *B*, AlphaFold2 model of TgLaforin's split CBM20 highlighting the core CBM predicted with high confidence in *cyan* and the low confidence insert regions in *gray*. *C*, TgLaforin's CBM core structure is predicted to bring together three putative, canonical residues used in carbohydrate interactions: W29, K206, and W278. *D*, TgCBM SEC profile with MW prediction overlay determines that TgCBM MW is 36.6 ± 6.8 kDa. *E*, TgCBM is stabilized in the presence of increasing concentrations of maltohexaose, while mutagenesis of carbohydrate binding residues results in mutants that no longer exhibit carbohydrate binding. Error bars represent standard deviation of three independent replicates. CBM, carbohydrate-binding module; DSP, dual-specificity phosphatase domain; MW, molecular weight; SEC, size-exclusion chromatography.

has a MW of 136.6 ± 0.6 kDa, twice the predicted size of the 61.5 kDa monomeric protein, indicating that it exists in solution as a stable dimer (Fig. 4*A*). When considered alongside the fact that TgLaforin-CBM is a monomer in solution (Fig. 3*D*), this is highly suggestive that dimerization of TgLaforin is mediated through the DSP domain.

The structure of a homodimer of TgLaforin was next modeled with AlphaFold-Multimer (56), resulting in a high confidence model that supports the exclusively DSP-driven anti-parallel dimer predicted by SEC-MALS experimental data (Fig. 4*B*). The predicted dimer displays an extended 1189 Å^2 interface and high shape complementarity (Sc = 0.71).
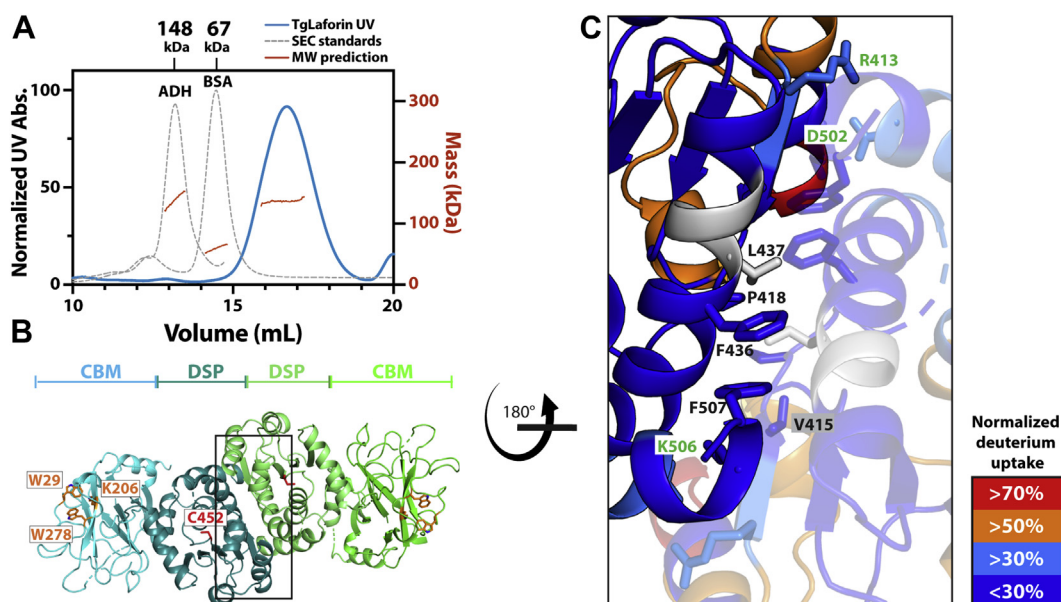
Figure 4. TgLaforin is an antiparallel dimer. *A*, size-exclusion chromatogram (SEC) overlaid with MALS-measured mass (*red*) for two SEC standards (*gray*; MW of standards indicated at elution volume on top) and TgLaforin (*blue*). MALS predicts that TgLaforin is 136.6 ± 0.6 kDa in solution. *B*, AlphaFold2 model of TgLaforin as a dimer. Monomeric subunits each are shown in shades of *blue* and *green*. Residues critical to glucan phosphatase function are shown on the *left subunit*: consensus CBM residues are depicted and identified in *orange*, and the catalytic cysteine is shown in *red*. *C*, close-up view of the dimer interface with each monomeric subunit indicated by differential transparency. Interfacial hydrophobic and hydrophilic residues are shown in *sticks* and labeled in *black* and *green*, respectively. HDX data of 15 s were mapped onto the structure model. Deuterium uptake levels are color coded as indicated in the figure legend, *gray regions* are regions not covered in HDX-MS experiment. CBM, carbohydrate-binding module; DSP, dual-specificity phosphatase domain; HDX-MS, hydrogen–deuterium exchange mass spectrometry; MALS, multiangle light scattering; MW, molecular weight.

Key interfacial residues include a core hydrophobic region and flanking electrostatic/hydrophilic interactions. Residues V415, P418, F436, L437, L441, and F507 form the hydrophobic core in the dimer interface between two DSP domains (Fig. 4C). The outer hydrophilic region contains R413, D502, and K506. This antiparallel dimer is very similar to that observed in human laforin (44). In particular, the hydrophobic residues with high conservation include: TgV415/HsI219, TgL437/HsL251, TgL441/HsL255, and TgF507/HsF321 (Fig. S1C, residues highlighted in *cyan*). Additionally, F507 corresponds to F321 in human laforin, and point mutations at this position have been reported to destabilize HsLaforin in human Lafora disease patients (54).

This dimer model was further analyzed using the HDX–MS data. Importantly, peptides predicted to be at the dimer interface displayed low deuterium uptake compared to solvent accessible surface regions, consistent with a key role in the obligate dimer interface (Figs. 4C and S3). The conserved hydrophobic residues in the hydrophobic core were particularly protected, with exchange ranging from 14% to 27%. This region includes F507 that exhibited an exchange rate of 20%. Flanking hydrophilic residues displayed exchange rates varying from 16% to 40%.

TgLaforin is an active glucan phosphatase

Phosphotyrosine specificity is conferred to PTPs through the ∼40 amino acid recognition domain that forms a deep catalytic pocket. In DSPs, however, the recognition region is much shorter, which results in a shallower active site that allows DSPs to dephosphorylate phospho-serine/threonine residues as well (33). As with other DSP containing proteins, the TgLaforin recognition region is much shorter than a typical PTP, confirming its placement in the DSP family. When compared with other glucan phosphatases, the TgLaforin DSP domain also shares many of the same glucan binding residues with human laforin, but not with plant glucan phosphatases (Fig. S1C; residues highlighted in blue).

These carbohydrate-binding residues identified in the DSP play a role in both specific and selective activity against glucan substrates. To determine TgLaforin activity, we first tested its activity against the artificial substrate para-nitrophenyl phosphate (pNPP). TgLaforin demonstrated robust activity against pNPP that was both dose-dependent and time-dependent (Fig. 5A). Moreover, when the catalytic cysteine (C452) was mutated to a serine (C452S), the activity was completely ablated, consistent with its role as the catalytic nucleophile. Against pNPP, TgLaforin displayed standard Michaelis–Menten kinetics with a K_m of 1.0 ± 0.05 mM and a V_{max} of 1.4 ± 0.02 nmol PO_4/min (Fig. 5B).

To test its activity against a biologically relevant substrate, we used a malachite green-based assay to detect phosphate release from solubilized potato amylopectin. TgLaforin readily dephosphorylated amylopectin in a time-dependent manner, whereas the catalytically inactive C452S control did not (Fig. 5C). Importantly, Vaccinia H1-related phosphatase (VHR), a member of the DSP family lacking a CBM, was unable to release phosphate from amylopectin. Therefore, in
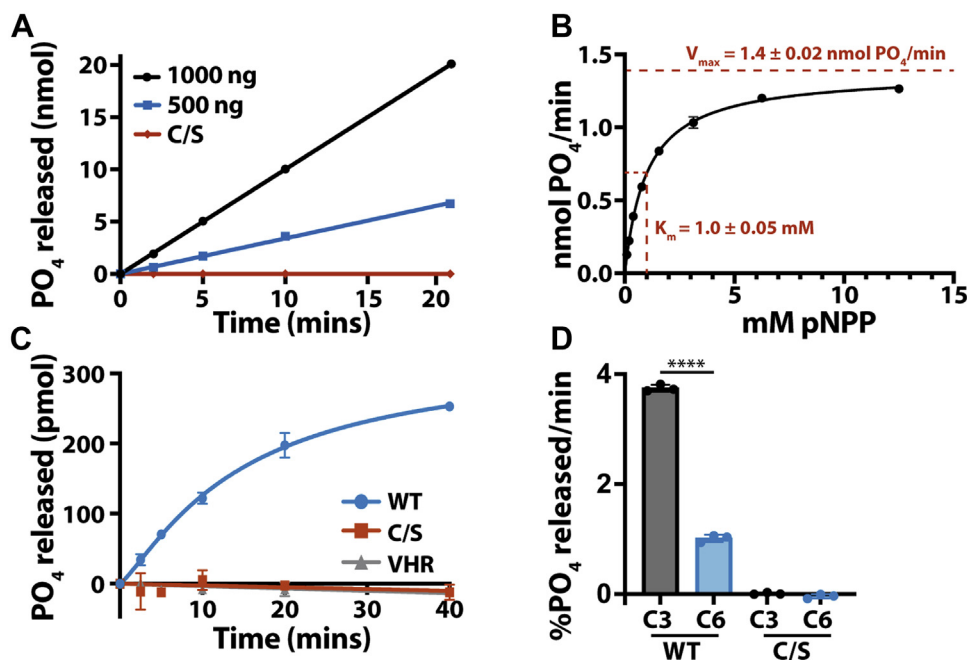
Figure 5. Full-length TgLaforin is an active glucan phosphatase that preferentially dephosphorylates starch at the C3 position. *A*, TgLaforin activity against pNPP is both dose and time-dependent and abrogated when the catalytic cysteine (C452) is mutated to serine (C/S). *B*, the rate of pNPP hydrolysis by TgLaforin as a function of pNPP concentration, demonstrating standard hyperbolic, Michaelis–Menten kinetics. K_m and V_{max} values are the mean of three independent replicates with uncertainties presented as the standard deviation from the mean. *C*, TgLaforin releases phosphate from solubilized potato amylopectin, whereas the C/S mutant and VHR cannot. *D*, TgLaforin preferentially dephosphorylates glucose from insoluble, ^{33}P-radiolabeld *A. thaliana* starch at the C3 position. Statistical comparison of site specificity was done using an unpaired two-tailed *t* test; statistical significance: ****$p < 0.0001$. All graphical error bars represent the standard deviation of three independent replicates. pNPP, para-nitrophenylphosphate; VHR, vaccinia H1-related phosphatase.

addition to providing evidence for its glucan phosphatase activity, these data also demonstrate the ability of TgLaforin to bind carbohydrates and confirm the function of its split CBM. The activity of TgLaforin against amylopectin deviated from Michaelis–Menten kinetics and displayed positive cooperativity toward this substrate with a Hill coefficient of 2.1 ± 0.4, indicative of positive substrate cooperativity (Fig. S4*A*). This result is consistent with the nature of the complex carbohydrate substrate and similar to the recently reported activity of the glucan phosphatase SEX4 (42). Lineweaver–Burk analysis of this data produced a nonlinear plot, confirming the non-Michaelis–Menten kinetics utilized by TgLaforin against amylopectin (Fig. S4*B*).

Glucose within both starch and glycogen is phosphorylated at both the C3- and C6-hydroxyl groups, and glucan phosphatases display varying specificities with regard to the site that they dephosphorylate (39). SEX4 displays a ~2.5-fold preference for dephosphorylating the C6-position (41), human laforin displays a reciprocal ~2.5-fold preference for the C3-position (36), and LSF2 exclusively dephosphorylates the C3-position (43). We tested TgLaforin for selectivity by incubating it with starch differentially radiolabeled at the C3- or C6-position. TgLaforin displayed a significant 3.5-fold preference for the C3-position (Fig. 5*D*). Thus, by activity, TgLaforin more closely resembles animal laforin compared to plant glucan phosphatases. Together, these data demonstrate that TgLaforin is a glucan phosphatase with a preference for C3-phosphate.

Design and synthesis of a TgLaforin inhibitor

Because TgLaforin represents a possible therapeutic target as a critical regulator of AG metabolism in *T. gondii*, we sought to exploit its observed unique features to identify a specific inhibitor. We utilized the sulfophenyl acetic amide (SPAA) platform, which is a novel pTyr-mimetic derived from the FDA-approved drug cefsulodin (57, 58). Based on SPAA, we previously developed potent and specific inhibitors to several PTP enzymes, including SHP2, mPTPA, mPTPB, and LMW-PTP (57–60). These inhibitors were discovered by fragment-based combinatorial chemistry approach where a library of compounds was prepared by reacting SPAA containing cores with a set of carboxylic acids or amines. An initial fragment screen identified the newly designed SPAA containing core L319-21 as the basis for identification of a TgLaforin inhibitor (Fig. S5). L319-21 was designed to incorporate SPAA and the linker 1,2-diaminobenzene 2-((2-aminophenyl)amino)-2-oxoacetic acid. To increase the potency and selectivity of core L319-21 for TgLaforin, we aimed to install molecular diversity in order to capture additional and less conserved interactions outside the pTyr binding cleft (*i.e.*, active site). L319-21 was then reacted with 192 amines that differ in size, charge, and lipophilicity to generate a library of compounds under standard HBTU coupling reactions conditions (reported in the Experimental procedures). These compounds were then subjected to screening against TgLaforin, PTP1B, and VHR at a concentration of 10 μM. L319-21-M49 was the most potent TgLaforin inhibitor, exhibiting >90% at 10 μM but <10% at

10 μM for inhibition against PTP1B and VHR, and was selected for detailed analysis (Fig. 6*A*).

To test the specificity and potency of L319-21-M49, we used the pNPP assay with a TgLaforin and a variety of other phosphatases (Fig. 6*B*). Each phosphatase was incubated with a concentration of pNPP equivalent to the K_m value for each respective enzyme (reported in the Experimental procedures) alongside a range of inhibitor concentrations. Under these conditions, L319-21-M49 possessed an IC$_{50}$ of 1.5 ± 0.3 μM against TgLaforin (Fig. 6*C*). Moreover, when tested against other phosphatases, the inhibitor was remarkably specific to TgLaforin. L319-21-M49 was at least 100-fold selective toward TgLaforin compared to another DSP (VHR), a general PTP (PTP1B), and an alkaline phosphatase (CIP) (Fig. 6*C*). To determine the mechanism by which L319-21-M49 inhibits the activity of TgLaforin against pNPP, a double reciprocal Lineweaver-Burk plot was generated. At L319-21-M49 concentrations of 0 μM, 1.5 μM, and 5 μM, the y-intercepts were 0.73 ± 0.01 μM, 0.86 ± 0.03 μM, and 0.81 ± 0.10 μM, respectively. Importantly, L319-21-M49 did not significantly change the y-intercept (*i.e.*, V$_{max}$) of the data ($p > 0.05$), but did significantly alter the x-intercept (*i.e.*, K_m; $p < 0.0001$) indicating that L319-21-M49 is a competitive inhibitor of TgLaforin (Fig. 6*D*).

Next, L319-21-M49 was tested against the glucan phosphatase activity of TgLaforin. To do this, we incubated TgLaforin with amylopectin in the presence of the inhibitor and detected phosphate release with malachite green.

Importantly, the inhibitor was also effective at preventing TgLaforin activity against soluble amylopectin with an IC$_{50}$ of 5.1 ± 0.4 μM (Fig. 6*E*). Finally, we tested the inhibitor's ability to inhibit phosphate release using the radiolabel assay. L319-21-M49 was also effective in preventing phosphate release from both the C3 and C6 positions with similar IC$_{50}$s of 6.7 ± 0.1 μM and 6.5 ± 1.4 μM, respectively (Fig. 6*F*). Thus, in L319-21-M49, we have identified a first-in-class lead compound for the inhibition TgLaforin.

Discussion

In animals, glycogen breakdown provides substrates for epigenetic modifications, central carbon metabolism, and protein posttranslational modifications (27, 28, 61). In plants, starch characteristically acts as both a transient and long-term storage form of glucose, dictating the level of free sugars and allowing for energy maintenance through dark periods (11). Interestingly, *T. gondii* appears to utilize both plant-like and animal-like systems of glucan storage and degradation making this system an appealing candidate for therapeutic development. Glucan phosphatases are critical regulators of polyglucan metabolism in both plants and animals. Here, we characterized the structure, dynamics, and activity of the unique *T. gondii* glucan phosphatase TgLaforin, demonstrating it possesses a single CBM20 N terminal to its DSP, a dimerization interface *via* the DSP domain, and preference for dephosphorylating the C3-hydroxyl position on glucose.
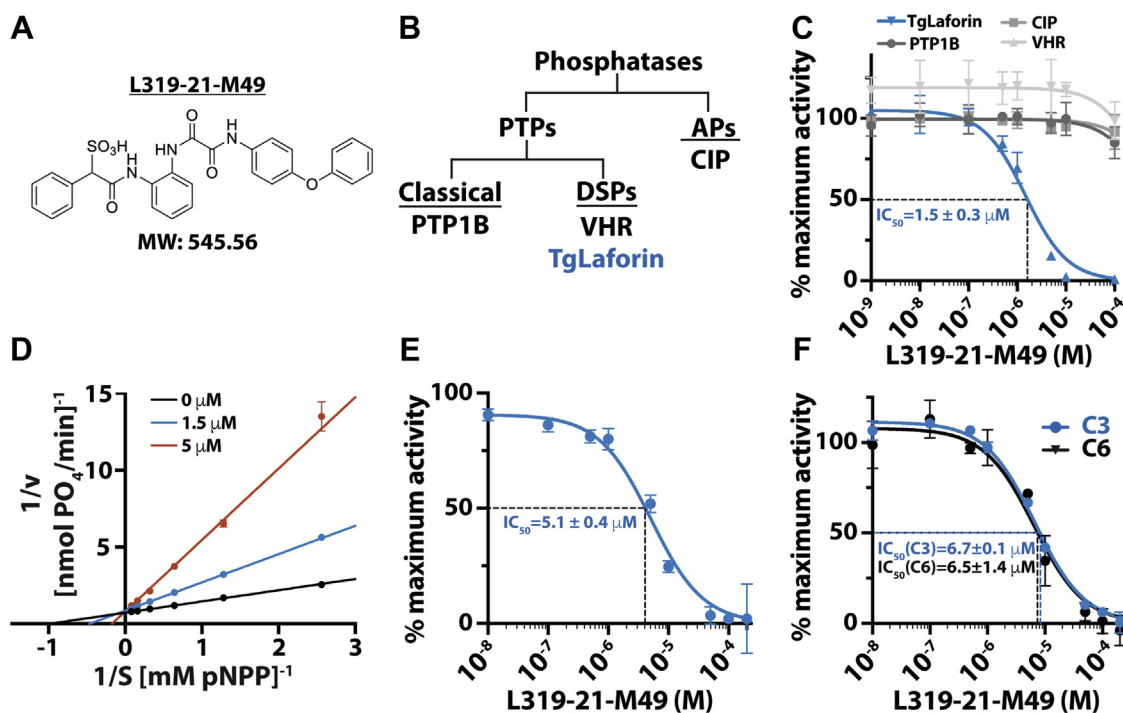

Figure 6. Development and testing of a novel glucan phosphatase inhibitor. *A*, the structure the TgLaforin inhibitor L319-21-M49. *B*, classification tree of the phosphatases used in this study. *C*, L319-M21-M49 inhibits TgLaforin activity against the model phosphatase substrate pNPP with >100-fold specificity over other phosphatases. *D*, Lineweaver-Burk double-reciprocal plot demonstrating that L319-21-M49 is a competitive inhibitor of TgLaforin. Each dataset represents a different inhibitor concentration, indicated in legend. *E*, L319-M21-M49 inhibits TgLaforin activity against soluble amylopectin; *F*, activity against insoluble starch with similar efficiency. All graphical error bars represent the standard deviation of at least three independent replicates, and all IC$_{50}$ values are the mean of three independent replicates with uncertainty presented as standard deviation from the mean. AP, alkaline phosphatase; CIP, calf intestinal phosphatase; DSP, dual-specificity phosphatase; PTP, protein tyrosine phosphatase; VHR, Vaccinia H1-related phosphatase.

Finally, we identified a first-in-class lead compound that inhibits TgLaforin.

To elucidate the split nature of this novel CBM20, we employed the recently developed AlphaFold2 neural network for an initial structural prediction. In agreement with our sequence predictions and site-directed mutagenesis experiments, the computational prediction of AlphaFold2 suggested that TgLaforin's CBM20 folds into a central core β-sandwich with large unstructured loops outside of this core. This orientation connects the β-sheets so that they form the canonical immunoglobulin-like fold. The linker regions were modeled with a very low level of confidence by AlphaFold2, so we utilized HDX to define both core and linker regions of the TgCBM20. Strikingly, HDX demonstrated that regions that were predicted to form the CBM20 core exhibited much lower solvent exchange. Conversely, regions displaying the highest solvent uptake were also the low-confidence regions of the AlphaFold2 model, potentially revealing intrinsically disordered protein-like character of these inserts (62, 63). Low complexity inserts have been previously noted as being unusually abundant in the proteome of *T. gondii* and other related parasitic protozoa, and these regions have been suggested to facilitate protein–protein interactions (64–66). Indeed, HsLaforin has been extensively characterized as possessing many protein binding partners (67), so intrinsically disordered regions in the TgCBM may facilitate such interactions in the *T. gondii*.

We further demonstrated that each homologous region and the aromatic residues within them are critical to the ability of TgLaforin to bind carbohydrates, consistent with the fact that it indeed contains a single CBM20. This combination of protein modeling, mutagenesis, and HDX-MS allowed us to define the first reported split-CBM. We attempted to identify other proteomes containing a split-CBM20 using BLASTp and TgCBM as a query, but we were only able to identify putative split-CBM20 orthologues in highly related *T. gondii* genera that are also cyst forming members of the Phylum Apicomplexa such as *Neospora*, *Besnoitia*, and *Eimeria*.

Therefore, despite its similarities to vertebrate glucan phosphatases, TgLaforin employs a unique starch-binding domain (SBD) in the form of a split-CBM20. Currently, there are 88 CBM families identified in the CAZy database, 15 of which are SBDs (52). The accepted definition of a SBD requires that it must bind to complex polysaccharides and that it be a continuous and discreet domain of about 100 amino acids. However, the discovery and characterization of a *Microbacterium aurum* GH13 α-amylase (MyAmyA) revealed the first *bona fide* SBD containing 300 amino acid residues (CBM74), demonstrating that SBDs can be much longer than the previously expected length (68). In characterizing the CBM74, structural and mutagenesis studies were not performed on the CBM74 itself, and 3D modeling with the Phyre2 server, a previous modeling software, was unable to generate a confident tertiary structure prediction for the domain. Phyre2 predicted 12 β-strands throughout the CBM74 domain, but no continuous sequence between strands approached the length of the linkers in the TgCBM. Here, we describe a CBM20 that

is similar in length to CBM74, but distinct in that its carbohydrate-binding residues and secondary structures are split over approximately 325 amino acids. This split-CBM20 contains discreet regions of homology to other CBM20s, separated by long linker regions with high solvent accessibility.

Recent work has demonstrated that glucan metabolism is not restricted to bradyzoites in *T. gondii* but is critical throughout the asexual lifecycle, playing a role in parasite growth, morphology, and infection (69–74). Tachyzoites have been recognized as containing a much smaller cytoplasmic polyglucan that is not typically visible by electron microscopy but is detectable using a carbohydrate-specific stain such as periodic-acid Schiff (PAS) (69, 73–75). This tachyzoite polyglucan is rapidly turned over as observed in other protozoan glucans during their actively replicating trophozoite stages (69, 75, 76). Upon stage conversion, bradyzoites accumulate large, insoluble AGs that are visible by TEM and, thus, distinct from the glucan in tachyzoites. In other organisms, loss of glucan phosphatases results in increased glucan phosphate content that leads to enlarged granule morphology in plants, and the precipitation of glycogen into amylopectin-like Lafora bodies in animals and human. Posttranslational regulation of TgLaforin's activity in *T. gondii via* its significantly unstructured CBM might provide a mechanism for finely tuned precipitation of the tachyzoite glucan into the larger, crystalline bradyzoite AG.

Of note, Ca^{2+}-dependent protein kinase 2, CDPK2, was identified as a regulator of AG metabolism in tachyzoites and bradyzoites. The ability of CDPK2 to regulate glucan metabolism in tachyzoites was mediated through its CBM20 domain. As such, loss of CDPK2 results in parasites with aberrant cell morphology, excess AG accumulation, and the inability to form tissue cysts in mice (69). Importantly, multiple CDPK2 substrates are involved in glucan turnover including glycogen phosphorylase, α-amylase, debranching enzyme, and TgGWD. These data highlight the critical role reversible phosphorylation might play in *T. gondii* beyond its role in bradyzoites. Additionally, they suggest that enzymes involved in glucan metabolism are potential drug targets throughout the parasite's asexual lifecycle.

Historically, phosphatases have been difficult to specifically target (77). However, recent breakthroughs have resulted in a number of successful compounds being developed (60, 78–83). For TgLaforin, we exploited differences in phosphatase tertiary structure outside of the well-conserved PTP-loop to develop an inhibitor. L319-21-M49 was a selective and competitive inhibitor of TgLaforin with >100-fold specificity compared to VHR, another member of the DSP family. Further work will seek to enhance the specificity of the TgLaforin inhibitor such that it can differentiate among glucan phosphatases. Given the differences in active site aromatic residues among glucan phosphatases, it is predicted that increases in potency will also be associated with increases in selectivity for the SPAA-based inhibitors as has been demonstrated with other phosphatases (58–60). Importantly, L319-21-M49 not only inhibited the generic pNPP activity of TgLaforin, but it also inhibited the ability of TgLaforin to dephosphorylate complex carbohydrates.

Glucose metabolism plays a critical role in central carbon metabolism and cellular signaling from bacteria to humans. Glucose storage and utilization in AGs is beginning to be elucidated in *T. gondii*. Given that a glucose aggregate has now been established in both life cycles that replicate in humans (84), the enzymes that control AG metabolism are potential therapeutic targets, especially the enzymes involved in AG reversible phosphorylation. Importantly, AG metabolism is intricately linked to bradyzoite development (85), making reversible phosphorylation in AGs an intriguing target. Our studies demonstrate the key enzymatic features of TgLaforin as a glucan phosphatase and lay the groundwork for targeting TgLaforin as a means to prevent *T. gondii* bradyzoite reactivation and transmission.

Experimental procedures

Bioinformatic analyses

The amino acid sequences of TgLaforin (TGME49_205290) was obtained from ToxoDB.org, and other glucan phosphatase sequences were obtained using BLASTp (86). Domains were defined using the NCBI's Conserved Domain Database (87). Accession numbers for each protein are listed in Table S1. Amino acid sequences of the glucan phosphatases were aligned using ClustalW and prepared using MacVector 18 (MacVector, Inc). MW calculations of TgCBM and full-length TgLaforin recombinant protein was calculated using the ProtParam tool on the ExPASy server (88).

AlphaFold2 modeling

The three-dimensional models for TgLaforin were obtained by running AlphaFold2 *via* the DeepMind official Colab server using AlphaFold v2.1.0 (https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb) (46, 56). A slightly simplified version of AlphaFold v2.1.0 was hosted on Colab notebook, which supports both monomeric and multimeric structure predictions. For the TgLaforin monomer structure prediction, model confidence was assessed by the pLDDT score (46), whereas a weighed combination of predicted TM-score and interface predicted TM score was used to assess the reliability of the TgLaforin dimer model (56). In both cases, top ranked models were used for subsequent analysis. The interfacial interactions and shape complementarity were analyzed using PDBePISA server (89) and SC software within CCP4 package (90). Structural models were analyzed, and figures were made using PyMOL 2.2.3 (Schrodinger).

Electron microscopy

Type II *T. gondii* strain ME49 tachyzoites were used to infect human foreskin fibroblast monolayers. Tachyzoite samples were fixed with 3% glutaraldehyde after 2 days of growth under standard cell culture conditions (37 °C and 5% CO_2). Bradyzoite samples were generated by differentiating parasites under alkaline stress (pH 8.2) for 6 days and then subsequently fixing with 3% glutaraldehyde. Infected cell monolayers were prepared for transmission electron microscopy as previously described (91) in the University of Kentucky's Imaging Center. Micrographs were collected at the University of Kentucky's Electron Microscopy Center on a Talos F200X TEM (Thermo Fisher Scientific) operated at 200 kV accelerating voltage with a 50 μm objective aperture inserted to enhance contrast using a 16M pixel 4k × 4k CMOS camera (Ceta, Thermo Fisher Scientific).

Cloning and production of TgLaforin baculovirus

The cDNA sequence of TgLaforin was obtained from ToxoDB.org. For expression in Sf9 insect cells, the cDNA sequence of the wildtype and C452S mutant were synthesized and cloned into pFastBac-HTA (Invitrogen) using NdeI/XhoI restriction digest sites (GenScript). The bacmid and baculovirus used in TgLaforin expression were then created using the manufacturer's product specifications (MAN0000414) with several modifications. Briefly, the TgLaforin bacmid was created by transforming pFastBac-HTA containing TgLaforin into DH10Bac cells where transposition of TgLaforin into a bacmid occurred, and the bacmid was then purified from a transformed colony. The baculovirus was produced by transfecting 2 μg of purified bacmid containing TgLaforin into 10^6 adherent Sf9 insect cells in a 6-well plate using Cellfectin II reagent (Invitrogen), and the supernatant containing the initial baculovirus was harvested 72 h later. Serial passage of the baculovirus through insect cells in suspension was used to increase viral titer. This was done by repeatedly infecting Sf9 cells in log phase of growth (1.5–2.5 × 10^6 cells/ml) and harvesting supernatant containing baculovirus 72 h later. Final viral stock infection volume was optimized by infecting insect cells in log phase at various concentrations (1:100–1:2000 v/v) and measuring cell viability and morphology after 72 h using Trypan blue staining (>80% living cells, the majority of which were enlarged). An effective infection ratio was determined to be 1:1000 of virus into log-phase insect cells. Protein production was then initiated using optimized volume of viral stock to infect log-phase insect cells for 72 h. Throughout study, insect cells were cultured at 27 °C rotating at 135 rpm without antibiotics and maintained between 10^6 and 10^7 cells/ml.

Expression and purification of full-length TgLaforin

Infected cells were centrifuged at 500*g* for 10 min to collect cells, washed once with PBS, and pellets were flash frozen and stored at −20 °C until needed. To purify, cells were resuspended in lysis buffer (50 mM Hepes, 100 mM NaCl, 15 mM imidazole, 2 mM dithiothreitol [DTT], 0.1% Triton X-100, pH 8.0). Cells were lysed by sonication, lysate was clarified by centrifugation at 14,000*g* for 30 min at 4 °C, and the supernatant containing protein was harvested. Protein was isolated from the supernatant using immobilized metal affinity chromatography with Ni-agarose beads (Sigma-Aldrich). Protein bound to beads was washed three times with purification buffer (lysis buffer without Triton X-100) and eluted in purification buffer containing 300 mM imidazole. Finally, protein was purified to >95% homogeneity *via* SEC using an AKTA

Pure fitted with a Superdex 200 16/200 column (GE Health-care). Two ml fractions were collected at a rate of 0.5 ml/min, analyzed for purity by Coomassie blue staining after SDS-PAGE, and only fractions containing pure TgLaforin were combined and concentrated to 2 to 3 mg/ml using Amicon centrifugal filters (30k MW cutoff). Final protein concentration was determined using absorbance at 280 nm on a Nanodrop (Thermo Fisher Scientific). Protein was flash-frozen and stored in purification buffer containing 10% glycerol at −80 °C.

Cloning, expression, and purification of TgCBM constructs

TgCBM truncations and point mutations were generated and cloned into pET28b for bacterial expression. Each mutant was produced in BL21-DE3 (NEB) cells grown in 2xYT media at 37 °C, with shaking at 225 rpm. When the A_{600} reached 0.6, protein production was induced with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; Gold Bio) at 16 °C for ∼16 h. Bacteria were pelleted by centrifugation at 5000g for 10 min at 4 °C. Cell pellets were either frozen (as described for insect cells) or resuspended immediately in lysis/purification buffer (20 mM Tris-HCl, 100 mM NaCl, 2 mM DTT, at pH 7.5). Lysate was generated *via* sonication, and protein was purified as described above for full-length constructs.

Differential scanning fluorimetry

Thermal stability of TgCBM constructs in the presence of oligosaccharides was determined as previously described (36, 42). Briefly, DSF was performed using a CFX96 Real-Time PCR system (Bio-Rad) with the FRET channel excitation and emission wavelength set to 450 to 470 nm and 560 to 580 nm, respectively. Melting temperature of each protein (2 μM final concentration) was measured in 40 μl volumes in DSF buffer (50 mM Hepes, 100 mM NaCl, 2 mM DTT, pH 7.5) in the presence of 5X SYPRO Orange (final concentration; Invitrogen). Thermal denaturation curves were obtained by measuring fluorescence intensity in the FRET channel as the temperature was increased from 20 to 95 °C at a rate of 1 °C/min. The melting temperature (T_m) of each protein with and without carbohydrate was obtained by calculating the first derivative of the fluorescence melt curve, fitting the derived curve with a Gaussian peak using Prism 9 (GraphPad) and determining the mean of the peak which corresponded to the T_m (the inflection point of the original melt curve).

Size-exclusion chromatography coupled with multiangle light scattering

SEC-MALS analysis was performed as previously described (55). Proteins were run on an AKTA Pure system using a Superdex 75 Increase 100/300 column (TgCBM) or a Superdex 200 10/300 column (full-length TgLaforin; GE Healthcare), miniDAWN TREOS, and Optilab T-rEX (Wyatt Technologies) arranged in sequence with one another. The columns were first equilibrated with the purification buffer used for each protein (described in expression and purification methods). Then,

500 μl of 2 mg/ml protein was then loaded and run at 0.5 ml/min, and elution profiles were monitored using UV-Vis absorbance. Light scattering data were analyzed on ASTRA software (Wyatt Technologies). The MW of each protein was then determined by analyzing peaks at half height using the refractive index to determine accurate protein concentration. UV and MW data were exported and plotted in Prism 9.

Hydrogen–deuterium exchange mass spectrometry

HDX-MS experiments were performed using a Synapt G2-SX HDMS system (Waters). Samples were processed using automated handling with a HDX PAL liquid handling system (LEAP). Samples were measured in triplicate, with technical replicates randomized using a workflow control software (Chronos). For labeled samples, 3 μl of TgLaforin (1 mg/ml, 20 mM Tris, 100 mM NaCl, 2 mM DTT, pH 7.5) was diluted with 57 μl of labeling buffer (20 mM Tris, 100 mM NaCl, 99% D_2O, pD 7.4). The reactions were incubated at 20 °C for 15 s, 150 s, 1500 s, and 15,000 s. Nondeuterated data were acquired the same way except that samples were diluted with an H_2O-based buffer (20 mM Tris, 100 mM NaCl, H_2O, pH 7.4). Reference deuterated samples were prepared by incubating 3 μl of TgLaforin in 57 μl D_2O buffer (99% D_2O + 1% (v/v) formic acid) for 24 h at room temperature. For all samples, 50 μl of the reaction solution was then rapidly mixed with 50 μl of quench buffer (100 mM phosphate, H_2O, pH 2.5) at 1 °C.

After the samples were quenched, 95 μl of the sample was immediately loaded onto an Acquity M-class UPLC (Waters) to minimize back exchange. UPLC runs were performed using sequential inline pepsin digestion (Waters Enzymate Beh Pepsin column, 2.1 mm × 30 mm) at 15 °C followed by reverse phase purification (Acquity UPLC BEH C18 1.7 μm at 0.2 °C). Sample was loaded onto the column equilibrated with 95% water, 5% acetonitrile, and 0.1% formic acid at a flow rate of 40 μl/min. A 7 min linear gradient (5%–35% acetonitrile) followed by a ramp and 2 min block (85% acetonitrile) was used for separation and directly continuously infused onto a Synapt XS using Ion Mobility (Waters). [Glu1]-Fibrinopeptide B was used as a reference.

Data from nondeuterated samples were used for peptide identification with ProteinLynx Global Server 3.0 (Waters). Peptides with intensity lower than 5000 and/or score below seven were excluded, long peptides (length > 20) were also filtered out. The filtered peptide list and MS data were imported into DynamX 3.0 (Waters) for deuterium uptake calculation using both retention time and mobility matching. Deuterium uptake levels were normalized as recommended (92) using the following equation:

$$D_{norm} = (m - m_{0\%}) / (m_{100\%} - m_{0\%})$$

where D_{norm} is the normalized deuteration level, m is the centroid mass of a given peptide at a specific time point, $m_{0\%}$ and $m_{100\%}$ are the centroid masses of the nondeuterated peptide, and the maximally deuterated peptide, respectively. The normalized deuterium uptake data at 15 s were mapped

onto the AlphaFold predicted structural model using PyMol 2.2.3.

pNPP phosphatase assay

Hydrolysis of pNPP to 4-nitrophenylphosphate (4NP) was performed in 50 μl reactions containing phosphatase buffer (0.1 M sodium acetate, 0.05 M bis-Tris, 0.05 M Tris-HCl, and 2 mM DTT at pH 5.5) at 37 °C. pNPP concentration was equivalent to the K_m value determined for each phosphatase under these reaction conditions containing 50 nM enzyme (unless otherwise indicated). Reactions were terminated by the addition of 200 μl of 0.25 N NaOH, and the production of 4NP was determined by measuring the absorbance at 410 nm using a Synergy HTX Multi-Mode Reader (BioTek). Absorbance was converted to nmol phosphate released (1:1-phosphate:4NP ratio) using a 4NP standard curve.

V_{max} and K_m were calculated by varying pNPP concentrations and measuring absorbance in the linear range of enzyme activity. K_m values are as follows for each enzyme: TgLaforin = 1.0 ± 0.05 mM, PTP1B = 1.0 ± 0.03 mM, VHR = 2.5 ± 0.1 mM, calf intestinal phosphatase = 53.8 ± 4.2 μM, with all uncertainties in these values calculated and presented as standard deviation from the mean of three independent replicates, each consisting of three technical replicates. Parameters were calculated using the Michaelis–Menten nonlinear curve-fitting function in Prism 9 software. The optimal pH (5.5) for TgLaforin was selected by determining specific activity as a function of pH from 5.0 to 9.0 in phosphatase buffer (see above) that maintains a constant ionic strength over this pH range.

Malachite green phosphatase assay

Release of phosphate from potato amylopectin was determined as done previously (23, 24, 42, 93) with the following modifications. Briefly, phosphate release was monitored using the PiColorLock Phosphate Detection Reagent (Novus Biologicals), a malachite green–based assay. For time course assays, 5 nM recombinant protein was incubated with 90 μg solubilized potato amylopectin (Sigma-Aldrich), supplied as a powder; solubilized at a stock concentration of 5 mg/ml using alcohol/alkaline method (also referred to as the "Roach method" in Ref. (93)) in phosphatase buffer in a final volume of 80 μl at pH 6.5. For TgLaforin kinetic characterization, 5 nM TgLaforin was incubated with varying amylopectin concentrations for 10 min. All reactions were terminated by addition of 20 μl (0.25 initial reaction volume) of the PiColorLock Gold solution containing Accelerator in a 100:1 ratio of Gold solution to the accelerator. After 5 min, 8 μl stabilizer solution (0.1 initial reaction volume) was added, and reaction was allowed to develop for 30 min at r.t. before the absorbance of each reaction was measured at 635 nm using a Synergy HTX Multi-Mode Reader (BioTek). Absorbance was converted to pmol P_i release using a P_i absorbance standard curve. Data points are presented as the mean of three independent replicates, each consisting of three technical replicates.

Radiolabeled starch phosphatase assay

Position-dependent phosphate release from ^{33}P-labeled starch was performed as previously described (36, 41–43). The substrate for this assay was generated by phosphorylating phosphate-free *Arabidopsis sex1-3* starch with β[33]P-labeled ATP (Hartmann Analytic) at the C6 and C3 positions sequentially. The starch granules were phosphorylated at the C6 position using purified recombinant StGWD enzyme. Next, the granules were phosphorylated at the C3 position using purified recombinant AtPWD enzyme.

To determine the site-specificity of TgLaforin, TgLaforin (1.3 nM) was incubated separately with 0.6 mg both C6- and C3-^{33}P-labeled starch in phosphatase buffer (pH 6.5) containing 1 mg/ml bovine serum albumin in a final volume of 150 μl. Dephosphorylation was allowed to proceed for 15 min with shaking at room temperature. The reaction was terminated with 50 μl 10% SDS, and starch was pelleted by centrifugation at maximum speed for 5 min. Supernatant was then transferred to 3 ml scintillation fluid. The amount of phosphate released from each starch suspension was quantified by measuring counts per minute (CPM) of the supernatant on a 1900 TR liquid scintillation counter (Packard). CPM measurements were then converted to % phosphate release by dividing CPM in the supernatant (released phosphate) by initial CPM on starch (bound phosphate).

Synthesis and characterization of novel inhibitor of TgLaforin, L319-21-M49

Materials and general procedures

Unless otherwise specified, all reagents were purchased from commercial suppliers and used directly without further purification. Analytical thin layer chromatography (TLC) was performed on 0.25 mm silica gel 60-F_{254}. Column chromatography was performed using KP-SIL silica gel (Biotage), and flash column chromatography was performed on Biotage prepacked columns using the automated flash chromatography system Biotage Isolera One. The ^{1}H and ^{13}C NMR spectra were recorded on a Bruker AVANCE 500 MHz instrument. Chemical shifts for proton magnetic resonance spectra (^{1}H NMR) were quoted in parts per million (ppm) referenced to the appropriate solvent peak or 0.0 ppm for tetramethylsilane. The following abbreviations were used to describe peak splitting patterns when appropriate: br = broad, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublet. Coupling constants, J, were reported in hertz unit (Hz). Chemical shifts for 13C NMR were reported in ppm referenced to the center line at 39.52 of DMSO-d. HPLC purification was carried out on a Waters Delta 600 equipped with a Sunfire Prep C18 OBD column (30 mm/150 mm, 5 μm) with methanol water (both containing 0.1% TFA) as mobile phase (gradient: 50–100% methanol, flow 10 ml/min). Low-resolution mass spectra and purity data were obtained using an Agilent Technologies 6470 series, triple quadrupole LC/MS. High-resolution mass spectra were recorded on an Agilent Mass spectrometer using electrospray ionization-time of flight.

Library synthesis

Reactions were assembled directly on 96-well plates by HBTU peptide coupling conditions. To each well of a 96-well plate L319-21 (20 μl, 20 mM in dimethylformamide [DMF]) was first added, HBTU (20 μl, 20 mM in DMF), HOBt (20 μl, 20 mM in DMF), and DIEA (20 μl, 75 mM in DMF) 5 min later, and various amines (1 μl, 500 mM in DMF) were added to each well. The plate was allowed to stand at room temperature for overnight to give products in stock solutions at 4 mM (assuming products were obtained at 80% yield). Representative wells from each plate were monitored by LC-MS, which indicated that the reactions went well, affording products in good conversions (70–80%). Following screening, the most potent inhibitor of TgLaforin, L319-21-M49, was resynthesized and purified with reversed-phase HPLC for further characterization (Fig. S6A).

Synthesis and characterization of compound L319-21-M49

Treating the starting material 2-phenyl-2-sulfoacetic acid (compound 1) with thionyl chloride afforded phenyl sulfoacetyl chloride (compound 2). Then, compound 3 was synthesized by reacting compound 2 with *o*-Phenylenediamine in the presence of DIEA. Compound 3 upon treatment with methyl chlorooxoacetate yielded compound 4, which was hydrolyzed in 0.5 M KOH (aq.) at room temperature for 1 h to produce key intermediate core L319-21. Final product L-319-M49 was obtained through condensation of core L319-21 with 4-phenoxyaniline in the presence of HOBt, HBTU, and DIEA and later purified by HPLC (Fig. S6A).

To a round-bottom flask was added 2-phenyl-2-sulfoacetic acid (3.0 g, 13.88 mmol) and thionyl chloride (16.51 g, 138.75 mmol), the resulting mixture was stirred at 80 °C overnight. Then, removal of extra thionyl chloride by rotary evaporator gave product phenyl sulfoacetyl chloride as a brown oil which was directly used for the next step without purification.

2-((2-aminophenyl)amino)-2-oxo-1-phenylethanesulfonic acid (3)

To *o*-phenylenediamine (1.50 g, 13.88 mmol) and DIEA (7.41 ml, 41.65 mmol) in DCM (40 ml), phenyl sulfoacetyl chloride (2.98 g, 13.88 mmol) and DCM (20 ml) solution were slowly added. The resulting mixture was stirred at r.t. for 12 h and then was concentrated by rotary evaporator. The mixture was subjected to HPLC purification, and product 2-((2-aminophenyl)amino)-2-oxo-1-phenylethanesulfonic acid was obtained as off-white solid (2.65 g, 62% yield, >95% purity). ^{1}H NMR (500 MHz, DMSO) δ 10.22 (s, 1H), 7.60 (d, J = 6.5 Hz, 2H), 7.32 to 7.24 (m, 5H), 7.24 to 7.17 (m, 2H), 4.90 (s, 1H). LC-MS (ESI): m/z [M + H]$^+$ calcd. For $C_{14}H_{15}N_2O_4S$: 307.08, found: 307.20.

2-((2-(2-methoxy-2-oxoacetamido)phenyl)amino)-2-oxo-1-phenylethanesulfonic acid (4)

To 2-((2-aminophenyl)amino)-2-oxo-1-phenylethane sulfonic acid (500 mg, 1.63 mmol) and DIEA (0.87 ml, 4.90 mmol) in DCM (20 ml), methyl chlorooxoacetate (220 mg, 1.80 mmol) was slowly added. The resulting mixture was stirred at r.t. for 1 h and then was concentrated by rotary evaporator. The mixture was purified by column chromatography eluting with dichloromethane/methanol 10:1 v/v to give 2-((2-(2-methoxy-2-oxoacetamido)phenyl)amino)-2-oxo-1-phenylethanesulfonic acid as a colorless oil (590 mg, 92% yield, >95% purity). ^{1}H NMR (500 MHz, DMSO) δ 10.26 (s, 1H), 10.22 (s, 1H), 7.74 (dd, J = 8.1, 1.3 Hz, 1H), 7.53 (dd, J = 7.6, 1.9 Hz, 2H), 7.46 (dd, J = 7.9, 1.3 Hz, 1H), 7.27 to 7.21 (m, 4H), 7.16 (td, J = 7.7, 1.4 Hz, 1H), 4.69 (s, 1H), 3.80 (s, 3H). LC-MS (ESI): m/z [M – H]$^-$ calcd. For $C_{17}H_{15}N_2O_7S$: 391.06, found: 391.10.

2-oxo-2-((2-(2-phenyl-2-sulfoacetamido)phenyl)amino)acetic acid (L319-21)

To a solution of compound 4 (500 mg, 1.27 mmol) in methanol (20 ml) and H$_2$O (20 ml), KOH (570 mg, 10.18 mmol) was added. The obtained mixture was stirred at r.t. for 1 h. The mixture was brought to 0 °C, carefully acidified with 1 N HCl until pH = 1 to furnish product 2-oxo-2-((2-(2-phenyl-2-sulfoacetamido)phenyl)amino)acetic acid (L319-21). Subsequently, it was subjected to HPLC purification, and product L319-21 was obtained as colorless oil (366 mg, 76% yield, >95% purity). LC-MS (ESI): m/z [M – H]$^-$ calcd. For $C_{16}H_{13}N_2O_7S$: 377.04, found: 377.10.

2-oxo-2-((2-(2-oxo-2-((4-phenoxyphenyl)amino)acetamido)phenyl)amino)-1-phenylethanesulfonic acid (L319-21-M49)

L319-21 (200 mg, 0.53 mmol), HOBt (86 mg, 0.63 mmol), and HBTU (241 mg, 0.63 mmol) were dissolved in dry DMF (20 ml). The mixture was stirred at r.t. for 15 min, 4-phenoxyaniline (98 mg, 0.53 mmol) and DIEA (0.28 ml, 1.59 mmol) were then added, and the resulting mixture was stirred at r.t. over night. DMF was removed by rotary evaporator, the mixture was subjected to HPLC purification, and product 2-oxo-2-((2-(2-oxo-2-((4-phenoxyphenyl)amino)acetamido)phenyl)amino)-1-phenylethanesulfonic acid (L319-21-M49) was obtained as off-white solid (234 mg, 81% yield, >95% purity). ^{1}H NMR (Fig. S6B) (500 MHz, DMSO) δ 10.74 (s, 1H), 10.42 (s, 1H), 10.24 (s, 1H), 7.87 (d, J = 8.9 Hz, 2H), 7.79 (d, J = 8.0 Hz, 1H), 7.53 (d, J = 7.2 Hz, 2H), 7.47 (d, J = 7.8 Hz, 1H), 7.37 (t, J = 7.9 Hz, 2H), 7.28 to 7.16 (m, 5H), 7.11 (t, J = 7.3 Hz, 1H), 7.04 (d, J = 8.9 Hz, 2H), 7.00 (d, J = 7.9 Hz, 2H), 4.68 (s, 1H). ^{13}C NMR (Fig. S6C) (126 MHz, DMSO) δ 166.75 (s), 159.34 (s), 158.12 (s), 157.16 (s), 152.75 (s), 135.14 (s), 133.61 (s), 132.97 (s), 130.02 (s), 129.85 (s), 127.47 (s), 126.86 (s), 126.66 (s), 126.29 (s), 126.21 (s), 124.56 (s), 123.13 (s), 122.08 (s), 119.27 (s), 118.09 (s), 72.01 (s). High-resolution mass spectra (electrospray ionization-time of flight): m/z [M + H]$^+$ calcd. For $C_{28}H_{34}N_3O_7S$: 546.1335, found: 546.1342.

Inhibitor assays

L319-M21-M49 stock solutions were prepared by resuspending the inhibitor in DMSO at a concentration of 20 mM. Various concentrations of inhibitor were prepared by serial

dilution in DMSO such that stock concentrations were 50× the working concentration, and the final concentration of DMSO in each assay was 2% (v/v). The efficacy of L319-M21-M49 against various phosphatases was determined using the assay conditions described above. Enzymes were preincubated with L319-M21-M49 for 5 min on ice before initiating the assay with the addition of substrate. All inhibitor assays were performed within the linear time/concentration range of phosphatase activity and terminated after 10 min. All reported uncertainties in IC$_{50}$ values were calculated and presented as standard deviation from the mean of at least three independent replicates, each consisting of three technical replicates. Statistical analysis of y- and x-intercept values obtained from Lineweaver–Burk plots was done using a one-way ANOVA in Prism 9. Intercepts are presented in text as the mean of three independent replicates with uncertainty presented as the standard deviation from the mean.

Data availability

All data presented are contained within the manuscript.

dual-specificity phosphatase domain; GWD, glucan, water-dikinase; HDX-MS, hydrogen–deuterium exchange mass spectrometry; LSF2, like-SEX4 2; pLDDT, predicted Local Distance Difference Test; pNPP, para-nitrophenylphosphate; PTP, protein tyrosine phosphatase; SEC-MALS, size-exclusion chromatography with multiangle light scattering; SEX4, starch excess4; TEM, transmission electron microscopy.

References

1. Robert-Gangneux, F., and Dardé, M. L. (2012) Epidemiology of and diagnostic strategies for toxoplasmosis. *Clin. Microbiol. Rev.* **25**, 264–296
2. Jones, J. L., and Dubey, J. P. (2012) Foodborne toxoplasmosis. *Clin. Infect. Dis.* **55**, 845–851
3. Torrey, E. F., and Yolken, R. H. (2013) Toxoplasma oocysts as a public health problem. *Trends Parasitol.* **29**, 380–384
4. Derouin, F., and Garin, Y. J. F. (1991) Toxoplasma gondii: blood and tissue kinetics during acute and chronic infections in mice. *Exp. Parasitol.* **73**, 460–468
5. Smith, N. C., Goulart, C., Hayward, J. A., Kupz, A., Miller, C. M., and van Dooren, G. G. (2021) Control of human toxoplasmosis. *Int. J. Parasitol.* **51**, 95–121
6. Montoya, J. G., and Liesenfeld, O. (2004) Toxoplasmosis. *Lancet* **363**, 1965–1976
7. Luft, B. J., and Remington, J. S. (1992) Toxoplasmic encephalitis in AIDS. *Clin. Infect. Dis.* **15**, 211–222
8. Silva, M.d., Teixeira, C., Gomes, P., and Borges, M. (2021) Promising drug targets and compounds with anti-Toxoplasma gondii activity. *Microorganisms* **9**, 1960
9. Coppin, A., Dzierszinski, F., Legrand, S., Mortuaire, M., Ferguson, D., and Tomavo, S. (2003) Developmentally regulated biosynthesis of carbohydrate and storage polysaccharide during differentiation and tissue cyst formation in Toxoplasma gondii. *Biochimie* **85**, 353–361
10. Sinai, A. P., Watts, E. A., Dhara, A., Murphy, R. D., Gentry, M. S., and Patwardhan, A. (2016) Reexamining chronic Toxoplasma gondii infection: surprising activity for a "dormant" parasite. *Curr. Clin. Microbiol. Rep.* **3**, 175–185
11. MacNeill, G. J., Mehrpouyan, S., Minow, M. A., Patterson, J. A., Tetlow, I. J., Emes, M. J., *et al.* (2017) Starch as a source, starch as a sink: the bifunctional role of starch in carbon allocation. *J. Exp. Bot.* **68**, 4433–4453
12. Guérardel, Y., Leleu, D., Coppin, A., Liénard, L., Slomianny, C., Strecker, G., *et al.* (2005) Amylopectin biogenesis and characterization in the protozoan parasite Toxoplasma gondii, the intracellular development of which is restricted in the HepG2 cell line. *Microbes Infect.* **7**, 41–48
13. Coppin, A., Varre, J.-S., Lienard, L., Dauvillee, D., Guerardel, Y., Soyer-Gobillard, M.-O., *et al.* (2005) Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite Toxoplasma gondii argues for a red alga ancestry. *J. Mol. Evol.* **60**, 257–267
14. Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., *et al.* (2012) CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS Biol.* **10**, e1001419
15. Behnke, M. S., Dubey, J. P., and Sibley, L. D. (2016) Genetic mapping of pathogenesis determinants in Toxoplasma gondii. *Annu. Rev. Microbiol.* **70**, 63–81
16. Ritte, G., Lloyd, J. R., Eckermann, N., Rottmann, A., Kossmann, J., and Steup, M. (2002) The starch-related R1 protein is an alpha -glucan, water dikinase. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 7166–7171
17. Kötting, O., Pusch, K., Tiessen, A., Geigenberger, P., Steup, M., and Ritte, G. (2005) Identification of a novel enzyme required for starch metabolism in Arabidopsis leaves. The phosphoglucan, water dikinase. *Plant Physiol.* **137**, 242–252
18. Edner, C., Li, J., Albrecht, T., Mahlow, S., Hejazi, M., Hussain, H., *et al.* (2007) Glucan, water dikinase activity stimulates breakdown of starch granules by plastidial beta-amylases. *Plant Physiol.* **145**, 17–28
19. Hejazi, M., Fettke, J., Haebel, S., Edner, C., Paris, O., Frohberg, C., *et al.* (2008) Glucan, water dikinase phosphorylates crystalline maltodextrins and thereby initiates solubilization. *Plant J.* **55**, 323–334

20. Hansen, P. I., Spraul, M., Dvortsak, P., Larsen, F. H., Blennow, A., Motawia, M. S., *et al.* (2009) Starch phosphorylation–maltosidic restrains upon 3'- and 6'-phosphorylation investigated by chemical synthesis, molecular dynamics and NMR spectroscopy. *Biopolymers* **91**, 179–193

21. Takeda, Y., and Hizukuri, S. (1981) Re-examination of the action of sweet-potato beta-amylase on phosphorylated (1→4)-χ-d-glucan. *Carbohydr. Res.* **89**, 174–178

22. Kötting, O., Santelia, D., Edner, C., Eicke, S., Marthaler, T., Gentry, M. S., *et al.* (2009) STARCH-EXCESS4 is a laforin-like phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. *Plant Cell* **21**, 334–346

23. Worby, C. A., Gentry, M. S., and Dixon, J. E. (2006) Laforin, a dual specificity phosphatase that dephosphorylates complex carbohydrates. *J. Biol. Chem.* **281**, 30412–30418

24. Gentry, M. S., Dowen, R. H., Worby, C. A., Mattoo, S., Ecker, J. R., and Dixon, J. E. (2007) The phosphatase laforin crosses evolutionary boundaries and links carbohydrate metabolism to neuronal disease. *J. Cell Biol.* **178**, 477–488

25. Gentry, M. S., and Pace, R. M. (2009) Conservation of the glucan phosphatase laforin is linked to rates of molecular evolution and the glucan metabolism of the organism. *BMC Evol. Biol.* **9**, 138

26. Kloehn, J., Krishnan, A., Tonkin, C. J., McConville, M. J., and Soldati-Favre, D. (2020) Chapter 10 - metabolic networks and metabolomics. In: Weiss, L. M., Kim, K., eds. *Toxoplasma gondii*, 3rd ed, Academic Press, Cambridge, MA: 451–497

27. Sun, R. C., Dukhande, V. V., Zhou, Z., Young, L. E. A., Emanuelle, S., Brainson, C. F., *et al.* (2019) Nuclear glycogenolysis modulates histone acetylation in human non-small cell lung cancers. *Cell Metab.* **30**, 903–916.e7

28. Sun, R. C., Young, L. E. A., Bruntz, R. C., Markussen, K. H., Zhou, Z., Conroy, L. R., *et al.* (2021) Brain glycogen serves as a critical glucosamine cache required for protein glycosylation. *Cell Metab.* **33**, 1404–1417.e9

29. Zeeman, S. C., Northrop, F., Smith, A. M., and Rees, T. (1998) A starch-accumulating mutant of Arabidopsis thaliana deficient in a chloroplastic starch-hydrolysing enzyme. *Plant J.* **15**, 357–365

30. Minassian, B. A., Lee, J. R., Herbrick, J. A., Huizenga, J., Soder, S., Mungall, A. J., *et al.* (1998) Mutations in a gene encoding a novel protein tyrosine phosphatase cause progressive myoclonus epilepsy. *Nat. Genet.* **20**, 171–174

31. Gentry, M. S., Guinovart, J. J., Minassian, B. A., Roach, P. J., and Serratosa, J. M. (2018) Lafora disease offers a unique window into neuronal glycogen metabolism. *J. Biol. Chem.* **293**, 7117–7125

32. Duran, J., Hervera, A., Markussen, K. H., Varea, O., López-Soldado, I., Sun, R. C., *et al.* (2021) Astrocytic glycogen accumulation drives the pathophysiology of neurodegeneration in Lafora disease. *Brain* **144**, 2349–2360

33. Alonso, A., Rojas, A., Godzik, A., and Mustelin, T. (2004) The dual-specific protein tyrosine phosphatase family. In *Protein Phosphatases*, Springer, Berlin/Heidelberg: 333–358

34. Patterson, K. I., Brummer, T., O'Brien, P. M., and Daly, R. J. (2009) Dual-specificity phosphatases: critical regulators with diverse cellular targets. *Biochem. J.* **418**, 475–489

35. Alonso, A., and Pulido, R. (2016) The extended human PTPome: a growing tyrosine phosphatase family. *FEBS J.* **283**, 1404–1429

36. Meekins, D. A., Raththagala, M., Auger, K. D., Turner, B. D., Santelia, D., Kötting, O., *et al.* (2015) Mechanistic insights into glucan phosphatase activity against polyglucan substrates. *J. Biol. Chem.* **290**, 23361–23370

37. Gentry, M. S., Brewer, M. K., and Vander Kooi, C. W. (2016) Structural biology of glucan phosphatases from humans to plants. *Curr. Opin. Struct. Biol.* **40**, 62–69

38. Emanuelle, S., Brewer, M. K., Meekins, D. A., and Gentry, M. S. (2016) Unique carbohydrate binding platforms employed by the glucan phosphatases. *Cell. Mol. Life Sci.* **73**, 2765–2778

39. Meekins, D. A., Vander Kooi, C. W., and Gentry, M. S. (2016) Structural mechanisms of plant glucan phosphatases in starch metabolism. *FEBS J.* **283**, 2427–2447

40. Vander Kooi, C. W., Taylor, A. O., Pace, R. M., Meekins, D. A., Guo, H. F., Kim, Y., *et al.* (2010) Structural basis for the glucan phosphatase activity of Starch Excess4. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15379–15384

41. Meekins, D. A., Raththagala, M., Husodo, S., White, C. J., Guo, H. F., Kötting, O., *et al.* (2014) Phosphoglucan-bound structure of starch phosphatase Starch Excess4 reveals the mechanism for C6 specificity. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7272–7277

42. Mak, C. A., Weis, K., Henao, T., Kuchtova, A., Chen, T., Sharma, S., *et al.* (2021) Cooperative kinetics of the glucan phosphatase Starch Excess4. *Biochemistry* **60**, 2425–2435

43. Meekins, D. A., Guo, H. F., Husodo, S., Paasch, B. C., Bridges, T. M., Santelia, D., *et al.* (2013) Structure of the Arabidopsis glucan phosphatase like sex four2 reveals a unique mechanism for starch dephosphorylation. *Plant Cell* **25**, 2302–2314

44. Raththagala, M., Brewer, M. K., Parker, M. W., Sherwood, A. R., Wong, B. K., Hsu, S., *et al.* (2015) Structural mechanism of laforin function in glycogen dephosphorylation and lafora disease. *Mol. Cell* **57**, 261–272

45. Kuchtová, A., Gentry, M. S., and Janeček, Š. (2018) The unique evolution of the carbohydrate-binding module CBM20 in laforin. *FEBS Lett.* **592**, 586–598

46. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589

47. Denu, J. M., Stuckey, J. A., Saper, M. A., and Dixon, J. E. (1996) Form and function in protein dephosphorylation. *Cell* **87**, 361–364

48. Simmons, Z. R., Sharma, S., Wayne, J., Li, S., Vander Kooi, C. W., and Gentry, M. S. (2021) Generation and characterization of a laforin nanobody inhibitor. *Clin. Biochem.* **93**, 80–89

49. Xiao, J., Engel, J. L., Zhang, J., Chen, M. J., Manning, G., and Dixon, J. E. (2011) Structural and functional analysis of PTPMT1, a phosphatase required for cardiolipin synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 11860–11865

50. Hsu, S., Kim, Y., Li, S., Durrant, E. S., Pace, R. M., Woods, V. L., Jr., *et al.* (2009) Structural insights into glucan phosphatase dynamics using amide hydrogen-deuterium exchange mass spectrometry. *Biochemistry* **48**, 9891–9902

51. Bott, R., Saldajeno, M., Cuevas, W., Ward, D., Scheffers, M., Aehle, W., *et al.* (2008) Three-dimensional structure of an intact glycoside hydrolase family 15 glucoamylase from Hypocrea jecorina. *Biochemistry* **47**, 5746–5754

52. Janeček, Š., Mareček, F., MacGregor, E. A., and Svensson, B. (2019) Starch-binding domains as CBM families-history, occurrence, structure, function and evolution. *Biotechnol. Adv.* **37**, 107451

53. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221

54. Brewer, M. K., Machio-Castello, M., Viana, R., Wayne, J. L., Kuchtová, A., Simmons, Z. R., *et al.* (2021) An empirical pipeline for personalized diagnosis of Lafora disease mutations. *iScience* **24**, 103276

55. Sharma, S., Vander Kooi, C. D., Gentry, M. S., and Vander Kooi, C. W. (2018) Oligomerization and carbohydrate binding of glucan phosphatases. *Anal. Biochem.* **563**, 51–55

56. [preprint] Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., *et al.* (2021) Protein complex prediction with AlphaFold-multimer. *bioRxiv*. https://doi.org/10.1101/2021.10.04.463034

57. He, R., Yu, Z. H., Zhang, R. Y., Wu, L., Gunawan, A. M., Lane, B. S., *et al.* (2015) Exploring the existing drug space for novel pTyr mimetic and SHP2 inhibitors. *ACS Med. Chem. Lett.* **6**, 782–786

58. He, R., Yu, Z. H., Zhang, R. Y., Wu, L., Gunawan, A. M., and Zhang, Z. Y. (2015) Cefsulodin inspired potent and selective inhibitors of mPTPB, a virulent phosphatase from Mycobacterium tuberculosis. *ACS Med. Chem. Lett.* **6**, 1231–1235

59. Dutta, N. K., He, R., Pinn, M. L., He, Y., Burrows, F., Zhang, Z. Y., *et al.* (2016) Mycobacterial protein tyrosine phosphatases A and B inhibitors augment the bactericidal activity of the standard anti-tuberculosis regimen. *ACS Infect. Dis.* **2**, 231–239

60. He, R., Wang, J., Yu, Z.-H., Zhang, R.-Y., Liu, S., Wu, L., *et al.* (2016) Inhibition of low molecular weight protein tyrosine phosphatase by an induced-fit mechanism. *J. Med. Chem.* **59**, 9094–9106

61. Brewer, M. K., Uittenbogaard, A., Austin, G. L., Segvich, D. M., DePaoli-Roach, A., Roach, P. J., *et al.* (2019) Targeting pathogenic Lafora bodies in Lafora disease using an antibody-enzyme fusion. *Cell Metab.* **30**, 689–705.e6

62. Ruff, K. M., and Pappu, R. V. (2021) AlphaFold and implications for intrinsically disordered proteins. *J. Mol. Biol.* **433**, 167208

63. Strodel, B. (2021) Energy landscapes of protein aggregation and conformation switching in intrinsically disordered proteins. *J. Mol. Biol.* **433**, 167182

64. Feng, Z. P., Zhang, X., Han, P., Arora, N., Anders, R. F., and Norton, R. S. (2006) Abundance of intrinsically unstructured proteins in P. falciparum and other apicomplexan parasite proteomes. *Mol. Biochem. Parasitol.* **150**, 256–267

65. Mohan, A., Sullivan, W. J., Jr., Radivojac, P., Dunker, A. K., and Uversky, V. N. (2008) Intrinsic disorder in pathogenic and non-pathogenic microbes: discovering and analyzing the unfoldomes of early-branching eukaryotes. *Mol. Biosyst.* **4**, 328–340

66. Dixon, S. E., Bhatti, M. M., Uversky, V. N., Dunker, A. K., and Sullivan, W. J., Jr. (2011) Regions of intrinsic disorder help identify a novel nuclear localization signal in Toxoplasma gondii histone acetyltransferase TgGCN5-B. *Mol. Biochem. Parasitol.* **175**, 192–195

67. Gentry, M. S., Romá-Mateo, C., and Sanz, P. (2013) Laforin, a protein with many faces: glucan phosphatase, adapter protein, et alii. *FEBS J.* **280**, 525–537

68. Valk, V., Lammerts van Bueren, A., van der Kaaij, R. M., and Dijkhuizen, L. (2016) Carbohydrate-binding module 74 is a novel starch-binding domain associated with large and multidomain α-amylase enzymes. *FEBS J.* **283**, 2354–2368

69. Uboldi, A. D., McCoy, J. M., Blume, M., Gerlic, M., Ferguson, D. J., Dagley, L. F., *et al.* (2015) Regulation of starch stores by a Ca2+-dependent protein kinase is essential for viable cyst development in Toxoplasma gondii. *Cell Host Microbe* **18**, 670–681

70. Sugi, T., Tu, V., Ma, Y., Tomita, T., and Weiss, L. M. (2017) Toxoplasma gondii requires glycogen phosphorylase for balancing amylopectin storage and for efficient production of brain cysts. *mBio* **8**, e01289-17

71. Cao, X. Z., Wang, J. L., Elsheikha, H. M., Li, T. T., Sun, L. X., Liang, Q. L., *et al.* (2019) Characterization of the role of amylo-alpha-1, 6-glucosidase protein in the infectivity of Toxoplasma gondii. *Front. Cell. Infect. Microbiol.* **9**, 418

72. Xia, N., Ye, S., Liang, X., Chen, P., Zhou, Y., Fang, R., *et al.* (2019) Pyruvate homeostasis as a determinant of parasite growth and metabolic plasticity in Toxoplasma gondii. *mBio* **10**, e00898-19

73. Lyu, C., Yang, X., Yang, J., Hou, L., Zhou, Y., Zhao, J., *et al.* (2021) Role of amylopectin synthesis in Toxoplasma gondii and its implication in vaccine development against toxoplasmosis. *Open Biol.* **11**, 200384

74. Yang, J., Yang, C., Qian, J., Li, F., Zhao, J., and Fang, R. (2020) Toxoplasma gondii α-amylase deletion mutant is a promising vaccine against acute and chronic toxoplasmosis. *Microb. Biotechnol.* **13**, 2057–2069

75. Ralton, J. E., Sernee, M. F., and McConville, M. J. (2021) Evolution and function of carbohydrate reserve biosynthesis in parasitic protists. *Trends Parasitol.* **37**, 988–1001

76. Sernee, M. F., Ralton, J. E., Nero, T. L., Sobala, L. F., Kloehn, J., Vieira-Lara, M. A., *et al.* (2019) A family of dual-activity glycosyltransferase-phosphorylases mediates mannogen turnover and virulence in Leishmania parasites. *Cell Host Microbe* **26**, 385–399.e9

77. Mullard, A. (2018) Phosphatases start shedding their stigma of undruggability. *Nat. Rev. Drug Discov.* **17**, 847–849

78. Krabill, A. D., and Zhang, Z. Y. (2021) Functional interrogation and therapeutic targeting of protein tyrosine phosphatases. *Biochem. Soc. Trans.* **49**, 1723–1734

79. Ruddraraju, K. V., Aggarwal, D., and Zhang, Z. Y. (2020) Therapeutic targeting of protein tyrosine phosphatases from Mycobacterium tuberculosis. *Microorganisms* **9**, 14

80. Ruddraraju, K. V., Aggarwal, D., Niu, C., Baker, E. A., Zhang, R.-Y., Wu, L., *et al.* (2020) Highly potent and selective N-aryl oxamic acid-based inhibitors for Mycobacterium tuberculosis protein tyrosine phosphatase B. *J. Med. Chem.* **63**, 9212–9227

81. Bai, Y., Yu, Z.-H., Liu, S., Zhang, L., Zhang, R.-Y., Zeng, L.-F., *et al.* (2016) Novel anticancer agents based on targeting the trimer interface of the PRL phosphatase. *Cancer Res.* **76**, 4805–4815

82. Zeng, L.-F., Zhang, R.-Y., Yu, Z.-H., Li, S., Wu, L., Gunawan, A. M., *et al.* (2014) Therapeutic potential of targeting the oncogenic SHP2 phosphatase. *J. Med. Chem.* **57**, 6594–6609

83. Ho, W. J., Croessmann, S., Lin, J., Phyo, Z. H., Charmsaz, S., Danilova, L., *et al.* (2021) Systemic inhibition of PTPN22 augments anticancer immunity. *J. Clin. Invest.* **131**, e146950

84. Watts, E., Zhao, Y., Dhara, A., Eller, B., Patwardhan, A., and Sinai, A. P. (2015) Novel approaches reveal that Toxoplasma gondii bradyzoites within tissue cysts are dynamic and replicating entities *in vivo*. *mBio* **6**, e01155-15

85. Sinai, A. P., and Suvorova, E. S. (2020) The RESTRICTION checkpoint: a window of opportunity governing developmental transitions in Toxoplasma gondii. *Curr. Opin. Microbiol.* **58**, 99–105

86. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410

87. Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., *et al.* (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* **48**, D265–d268

88. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., *et al.* (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **112**, 531–552

89. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797

90. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242

91. Dhara, A., de Paula Baptista, R., Kissinger, J. C., Snow, E. C., and Sinai, A. P. (2017) Ablation of an ovarian tumor family deubiquitinase exposes the underlying regulation governing the plasticity of cell cycle progression in Toxoplasma gondii. *mBio* **8**, e01846-17

92. Masson, G. R., Burke, J. E., Ahn, N. G., Anand, G. S., Borchers, C., Brier, S., *et al.* (2019) Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods* **16**, 595–602

93. Sherwood, A. R., Paasch, B. C., Worby, C. A., and Gentry, M. S. (2013) A malachite green-based assay to assess glucan phosphatase activity. *Anal. Biochem.* **435**, 54–56

Robert D. Murphy recently received his PhD from the University of Kentucky in Molecular and Cellular Biochemistry. His research focused on defining the underlying mechanisms of reversible glucan phosphorylation in the parasite *Toxoplasma gondii* using a number of techniques involving both enzymology and cell biology. Bobby will next begin a fellowship in public health with the CDC and can be found on Twitter @Bobby_D_Murphy.

Tiantian Chen is a graduate student in the Department of Molecular and Cellular Biochemistry at the University of Kentucky. His studies focus on understanding the molecular mechanisms of reversible carbohydrate phosphorylation. He specializes in structural biology, protein modelling, and structure-based drug design. He plans to continue fundamental studies to identify and structurally characterize target protein in carbohydrate phosphorylation pathway and utilize these insights to develop potent specific inhibitors.

Jianping Lin is a postdoctoral research associate in the Department of Medicinal Chemistry and Molecular Pharmacology at Purdue University. His current research involves design and synthesis of small-molecule inhibitors that target protein tyrosine phosphatases (PTPs) as chemical probes to interrogate PTP function and/or as novel therapeutics for the treatment of cancer, autoimmune disorders, and neurodegenerative and infectious diseases.

ASBMB

J. Biol. Chem. (2022) 298(7) 102089 **17**