BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Structural and Functional Characterization of an Ancient Bacterial Transglutaminase Sheds Light on the Minimal Requirements for Protein Cross-Linking

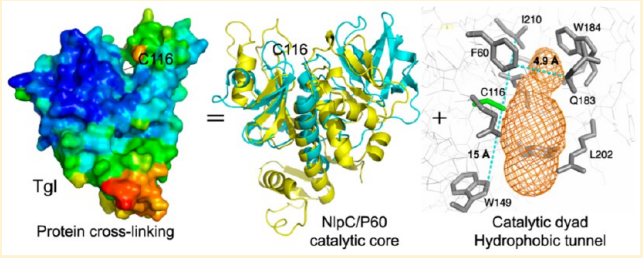
Catarina G. Fernandes,[†,⊥] Diana Plácido,[‡] Diana Lousa,[§] José A. Brito,[‡] Anabela Isidro,[†,#] Cláudio M. Soares,[§] Jan Pohl,[∥] Maria A. Carrondo,[‡] Margarida Archer,[*,‡] and Adriano O. Henriques[*,†]

[†]Microbial Development Group, [‡]Membrane Protein Crystallography, and [§]Protein Modelling Group of Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, ITQB-UNL, 2780-157 Oeiras, Portugal

[∥]Biotechnology Branch, Centers for Disease Control and Prevention, Atlanta, Georgia 30333, United States

Supporting Information

ABSTRACT: Transglutaminases are best known for their ability to catalyze protein cross-linking reactions that impart chemical and physical resilience to cellular structures. Here, we report the crystal structure and characterization of Tgl, a transglutaminase from the bacterium *Bacillus subtilis*. Tgl is produced during sporulation and cross-links the surface of the highly resilient spore. Tgl-like proteins are found only in spore-forming bacteria of the *Bacillus* and *Clostridia* classes, indicating an ancient origin. Tgl is a single-domain protein, produced in active form, and the smallest transglutaminase characterized to date. We show that Tgl is structurally similar to bacterial cell wall endopeptidases and has an NlpC/P60 catalytic core, thought to represent the ancestral unit of the cysteine protease fold. We show that Tgl functions through a unique partially redundant catalytic dyad formed by Cys116 and Glu187 or Glu115. Strikingly, the catalytic Cys is insulated within a hydrophobic tunnel that traverses the molecule from side to side. The lack of similarity of Tgl to other transglutaminases together with its small size suggests that an NlpC/P60 catalytic core and insulation of the active site during catalysis may be essential requirements for protein cross-linking.



Transglutaminases (TGases) catalyze transamidation reactions and are best known for their ability to cross-link proteins. Protein cross-linking results from the formation of stable and protease resistant ε-(γ-glutamyl)lysyl isopeptide bonds between protein-bound glutaminyl and lysyl residues, the Q and K substrates, respectively.[1,2] Protein cross-linking is central to a variety of important processes, including clotting phenomena, cell and tissue morphogenesis, and development.[1,2] Deregulation of TGase activity in humans contributes to a variety of severe pathologies, including progressive fibrosis, neurodegenerative, autoimmune, and infectious diseases, and illnesses related to malformations of the skin epidermis.[3,4]

The animal TGases, of which the human enzymes have been the most extensively studied, share a similar active site arrangement and acylation−deacylation pathway with the papain-like cysteine proteases[5] (Figure 1). In general, TGases employ a charge-relay Cys-His-Asp catalytic triad with the reaction involving formation of a covalent acyl−enzyme intermediate (Figure 1B, step 3); during the reaction, a thiolate/imidazolium ion pair is formed between the catalytic Cys and His residues, the latter acting both as a proton donor and as an acceptor during catalysis (steps 2 and 4, in Figure 1B). A hydrogen bond between the Asp and His residues facilitates the correct orientation of the basic amino acid[5−7] (Figure 1B). The catalytic core of TGases appears to derive from the minimal ancestral structural unit of the thiol-protease fold, whose closest current representative is found in the NlpC/P60 domain, which adds to the mechanistic similarity.[8,9] This domain takes its name from the bacterial NlpC/P60 cell wall endopeptidases, which show a catalytic triad composed of a Cys, His, and a third polar residue.[9,10] Because of the mechanistic and structural similarities, it has been suggested that TGases evolved from an ancestral, papain-like protease.[8,11]

Reflecting their multiple functions and interactions *in vivo* and the need for tight control, the animal-like TGases are structurally complex and have intricate activation mechanisms.[3,4,12] The animal-like TGases are activated by calcium and produced in zymogenic form, bound by inhibitory subunits, and/or negatively modulated by GTP/GDP or ATP.[2,12,13] Human TGase 2, for instance, a 80 kDa enzyme, adopts a compact, transamidase-inactive conformation in the absence of Ca^{2+}.[14] Activation of TGase 2 involves a large conformational rearrangement that exposes a hydrophobic tunnel leading to the active site.[15] Concealment of the tunnel in the inactive form

Received: June 14, 2015
Revised: August 30, 2015
Published: August 31, 2015

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723−5734

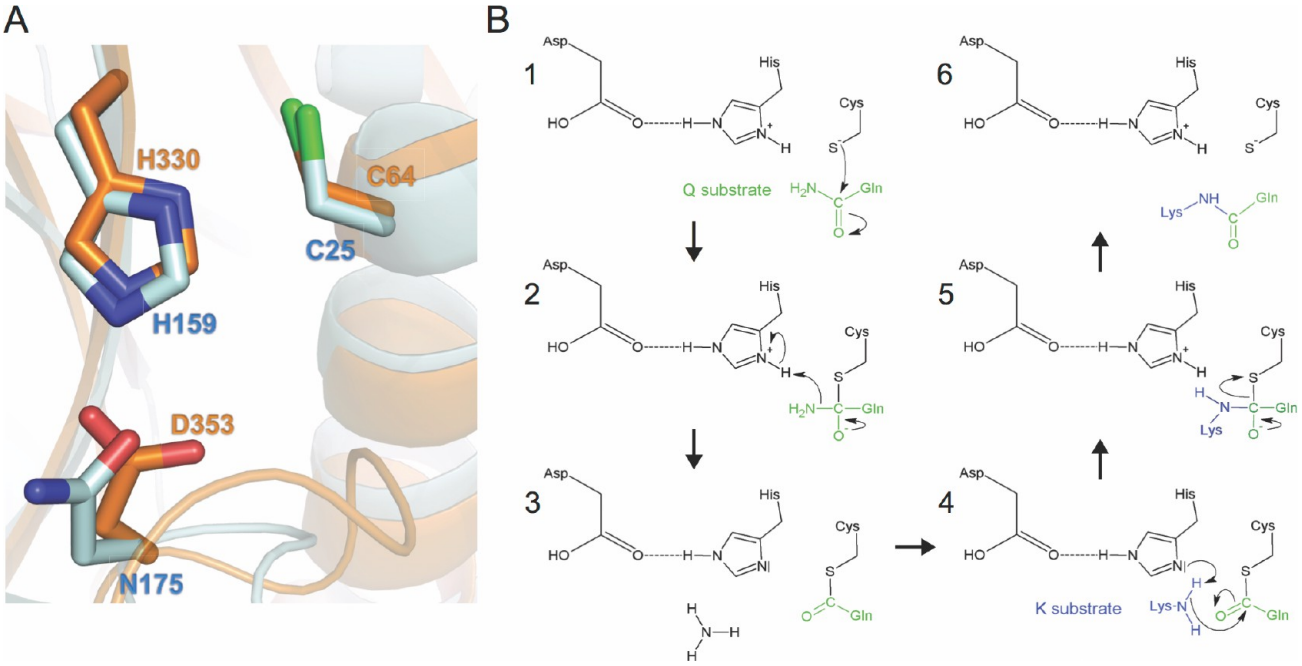


Figure 1. Proposed reaction mechanism for the animal TGases. (A) Superimposition of the catalytic residues of human TGase 3 (PDB entry 1L9M, orange) and papain (PDB entry 1PPN, blue). Note that only Asn175 of the dual Cys26-His159-Asp158 or Cys26-His159-Asn175 catalytic triad of papain is represented.[56] (B) Reaction mechanism of TGases proposed on the basis of the similarity of their catalytic core to that of papain-like proteases. Gln and Lys refer to protein-bound Gln and Lys residues, respectively. The reaction mechanism involves six different steps. (1) The thiolate ion nucleophilically attacks the Gln side chain (the acyl donor or Q substrate), leading to the formation of an oxyanion intermediate (2). Then, the regeneration of the carbonyl group of the tetrahedral intermediate leads to the release of ammonia and to the formation of the acyl−enzyme intermediate (3); the catalytic His removes a proton from the primary amine of the Lys residue (4), leading to the nucleophilic attack of the Lys residue on the carbonyl group of the acyl−enzyme intermediate formed in step 3 and to the formation of a second tetrahedral intermediate (5). As the oxyanion intermediate regenerates the carbonyl group, the thioester bond is cleaved, regenerating the initial catalytic core and releasing the final, cross-linked, protein product (6).

of the enzyme, as well as the presence of hydrophobic residues at its entry, is thought to reduce the potential for deamidation (hydrolytic) reactions.[15] A tunnel is also seen in the calcium-activated forms of human TGase 3 and Factor XIIIa.[16,17] In all cases, the active site is inaccessible from the outside in the inactive forms of the enzymes.[14,16,18] Although an NlpC/P60 catalytic core together with the insulation of the active site from water may be essential characteristics of TGases, the structural complexity of the animal enzymes makes this inference difficult.[19]

Proteins related to the classical animal TGases, carrying an NlpC/P60 core, were detected in the genomes of several other groups of organisms, including bacteria.[11] However, the NlpC/P60 domain defines a superfamily of proteins with distinct although related enzymatic activities.[8] Thus, while one of the putative TGases, TgpA, a large multidomain protein from *Pseudomonas aeurginosa*, indeed proved to be a TGase,[20] several phage-associated proteins were shown to be proteases,[21,22] and a protein from *Bordetella bronchiseptica* turned out to be involved in lipopolysaccharide maturation.[23] Conversely, not all proteins with TGase activity can be identified through sequence similarity as related to the animal-like TGases.[24,25] GP42, for example, a large (80 kDa), multidomain, Ca^{2+}-dependent TGase from a plant pathogen oomycete, shows no sequence similarity to the classical TGases.[26] Remarkably, GP42 also has an NlpC/P60 core and a Cys-His-Asp triad and seems to be related to a group of cysteine proteases from marine bacteria.[26] The possible convergent evolution of GP42 raises the possibility that an NlpC/P60 core is an essential structural requirement for TGase activity.[26]

Two TGases of bacterial origin, which were also not detected by sequence similarity as being related to the animal-type enzymes,[11] are much smaller in size. One, the MTG protein from *Streptomyces mobaraensis*, is the only bacterial TGase structurally characterized so far and appears to function through a Cys-Asp dyad.[24,25] However, it was not reported as having an NlpC/P60 core.[24,25] While MTG is a 38 kDa, single-domain, Ca^{2+}-independent enzyme, it is still produced as a zymogen and exhibits a complex activation mechanism.[27−29] The second is Tgl, found in a deeply rooted group of spore-forming microorganisms, which includes some *Bacillus* and *Clostridium* species, suggesting an ancient origin.[30,31] The 28 kDa Tgl protein from *Bacillus subtilis* has a well-established role in cross-linking a multiprotein surface structure known as the coat that encases and protects the spore.[32] Importantly, Tgl is produced in its active form, and no factor (Ca^{2+}, GTP, or other) is known to control its activity.[33]

Here, we present the crystal structure of Tgl, the smallest TGase characterized to date, and show that the enzyme has the NlpC/P60 fold at its catalytic core and a unique, partially redundant, catalytic dyad. Moreover, the catalytic Cys116 is insulated within a hydrophobic tunnel that traverses the molecule from side to side. Tgl is structurally related to a group of bacterial cell wall endopeptidases. Together with the finding that MTG also possesses an NlpC/P60 core, the structural and functional characterization of Tgl herein reported supports the idea that, irrespective of their evolutionary history,

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723−5734

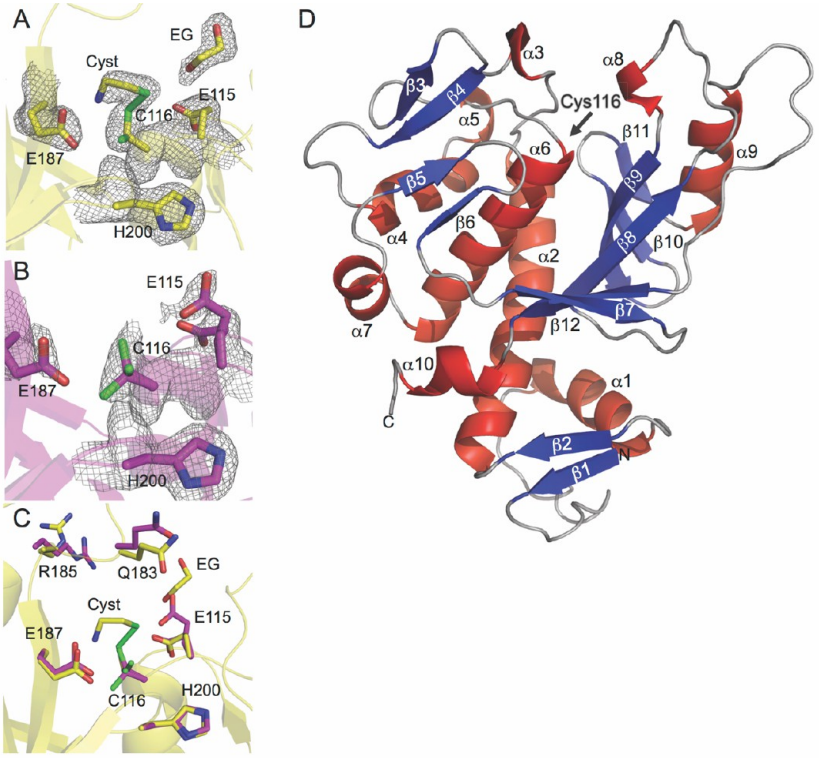


Figure 2. Active site and overall structure of Tgl. The $2F_o - F_c$ electron density maps contoured at 1.2σ around relevant residues (Glu115, Cys116, Glu187, and His200) for monomers A (A) and B (B) of the Tgl:cysteamine complex (see also Table S3). Monomer A shows a cysteamine (Cyst) covalently bound to the Sγ atom of conformer B of Cys116 and a molecule of ethylene glycol (EG) near the side chain of Glu115. In monomer B, the side chain of Glu115 displays two conformations and Cys116 was modeled in three different conformations: red for O, blue for N, and green for S and yellow for C for monomer A and magenta for C for monomer B (the same color code is followed throughout the text for Tgl). (C) Superposition of monomers A and B. Note that Gln183 and Arg185 show slightly different conformations in both monomers. (D) Overall fold of Tgl. Secondary structural elements are indicated as well as the position of Cys116 (black arrow).

all TGases have evolved from cysteine proteases. Furthermore, the structure of Tgl offers insight into the minimal structural requirements for protein cross-linking by TGases.

EXPERIMENTAL PROCEDURES

Protein Production and Purification. Point mutations were introduced into the *tgl* gene by polymerase chain reaction, creating pET30a(+) (Novagen) derivatives for the expression of Tgl, the wild type or mutant forms. GST-SpoVID$_{202}$ was expressed from pTC49.[34] All plasmids were introduced in *Escherichia coli* BL21(DE3), and the proteins were overproduced by a modified autoinduction method and the proteins purified by affinity chromatography. For activity assays, the purified proteins were dialyzed overnight against 0.1 M Tris-HCl (pH 8.0) using a 10 kDa cutoff membrane (SnakeSkin, from Pierce). For details on plasmid constructions, protein expression, and purification, see the Supporting Information. All primer sequences, plasmids, and strains used in this study are listed in Tables S1 and S2.

Crystallization, X-ray Data Collection, Phasing, and Refinement. Tgl was crystallized as previously described.[33] It is noteworthy that no Tgl crystals were formed at pH values higher than 4.5−5. Transfer of crystals to higher-pH solutions led to their degradation, even with crystals cross-linked with glutaraldehyde. The Tgl structure was determined by single isomorphous replacement with anomalous scattering (SIRAS), using the Tgl:Gd derivative (Table S3). An initial model, comprising 501 residues in eight chains, which belong to two molecules in the asymmetric unit, was built and improved. The refined structure of the Tgl:cysteamine complex was obtained from crystals soaked in a cystamine [2,2′-dithiobis-(ethylamine)] solution.

The stereochemistry of the refined model was assessed with Procheck/MolProbity[35,36] (Table S3). Structural representations were rendered with Pymol.[37] For details, see the Supporting Information.

Dansylcadaverine Labeling Assays. Amine incorporation activity was assessed by monitoring labeling of bovine serum albumin (BSA) (New England Biolabs) by dansylcadaverine (dansyl-cd, Fluka).[38] Reaction mixtures containing Tgl (16 μM), BSA (60 μM), and dansyl-cd (0.5 mM) in 0.1 M Tris-HCl (pH 8.0) were incubated at 50 °C and samples withdrawn over time and resolved on a 10% sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) gel. Labeling of BSA by dansyl-cd was recorded on a UV transilluminator (Chemidoc XRS, Bio-Rad) and quantified using ImageJ 1.37v.[39] A sample collected following reaction with wild-type Tgl for 120 min was loaded onto each gel for the normalization of the fluorescence signal among different assays. All Tgl forms were purified and assayed independently at least three times.

Cross-Linking Assays. The wild type or mutant forms of Tgl (16 μM) were incubated with GST-SpoVID$_{202}$ (60 μM) at 50 °C in 0.1 M Tris-HCl (pH 8.0) and samples collected at different times for analysis on 10% SDS−PAGE gels. As controls, GST-SpoVID$_{202}$ and Tgl were incubated alone, and samples were taken at the beginning and end of the assays. All

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723−5734

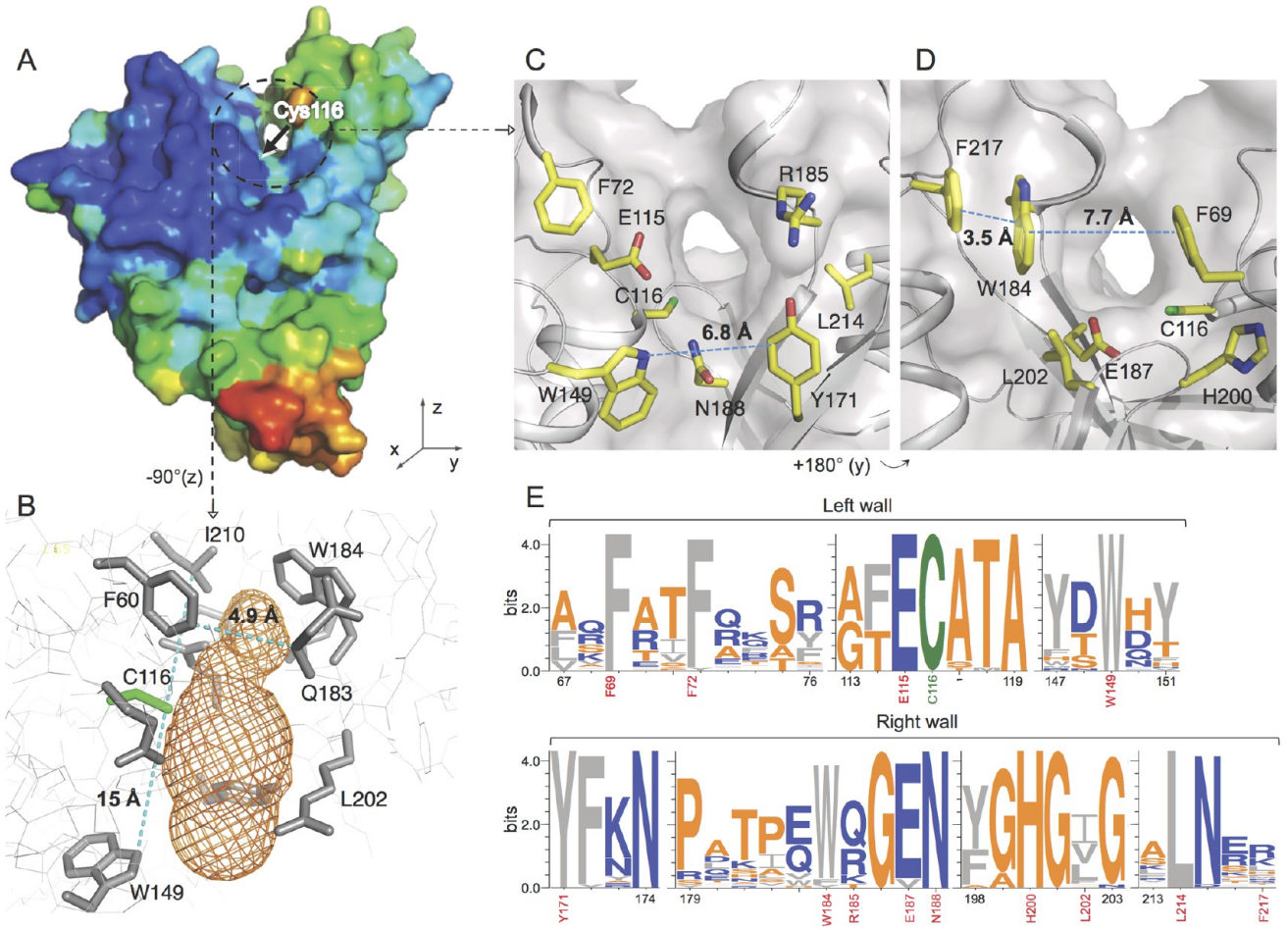


Figure 3. Tgl tunnel. (A) Molecular surface representation of Tgl (same orientation as in Figure 2D), color-coded according to the thermal displacement parameters (*B* factors; from blue, lower, to red, higher). The region around the Tgl tunnel is encircled. (B) Mesh representation of the cavity in the active site region, viewed as a cross section from the top of the Tgl molecule (residues forming the ceiling of the tunnel are not represented for the sake of clarity). (C) Close-up of the front entry of the hydrophobic tunnel. (D) Close-up of the back entry of the tunnel (180° rotation around the *y* axis, relative to panel C). In panels B–D, the distances between relevant residues, shown in stick representation, are indicated. (E) Sequence Logo[47] of the blocks of residues that form the left and right walls of the tunnel (numbered according to the Tgl sequence), for 40 selected Tgl orthologues (see Figure S1). Residues, in singe-letter notation, are color-coded as follows: M, L, I, V, F, and W in gray; A, G, T, and P in orange; E, D, R, K, Q, and N in blue; C in green. The residues represented in panels C and D are colored red.

forms of Tgl were assayed independently a minimum of three times.

Mass Spectrometry. The high-molecular weight species detected upon incubation of GST-SpoVID$_{202}$ with Tgl^{wt}, Tgl^{E187A}, and Tgl^{H200A} were analyzed by mass spectrometry and by N-terminal sequence (for the case of assays with wild-type Tgl), where GST-SpoVID$_{202}$ and Tgl (wild type or mutant forms) were positively identified (for details, see the Supporting Information).

Molecular Dynamics Simulations. Molecular dynamics (MD) simulations were performed with the GROMACS package,[40] version 4.0.4,[41] using the 53A6 force field.[42] Details of the methodology used for the calculations of the protonation states and for the MD simulations are given in the Supporting Information.

Accession Numbers. The atomic coordinates and the structure factor amplitudes of the Tgl crystal structure were deposited in the PDB as entries 4P8I (apo form) and 4PA5 (Tgl:cysteamine complex).

RESULTS

Structure of Tgl. The X-ray structures of Tgl were obtained from crystals soaked with cystamine, a TGase inhibitor.[43] Two X-ray structures of Tgl, in apo form and bound to cysteamine (the reduced form of cystamine), were refined to 1.85 and 1.86 Å with *R* values of 15.8 and 16.6%, respectively, and R_{free} values of 20.8 and 21.2%, respectively (relevant statistics are listed in Table S3). The apo and Tgl:cysteamine structures form dimers, which are very similar [root-mean-square deviation (rmsd) of 0.29 Å for all Cα atoms]. For the sake of simplicity, only the Tgl:cysteamine structure will be described in detail. The Tgl:cysteamine complex comprises 490 amino acid residues, a cysteamine in chain A [absent in chain B (Figure 2A–C)], four PEG fragments, two sulfate ions, one citrate, one glycerol, one zinc ion, and 204 water molecules. The Ramachandran plot shows 95.2% of the residues in the favored region and only 0.8% in the disallowed region.[44] The electron density maps are generally of good quality, except for three disordered regions: residues 1–20 in both chains (Asp13–Glu15 not included in the model) and residues 180–184 and 219–224 in chain B. Each molecule has approximate dimensions of 60 Å × 50 Å ×

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723–5734

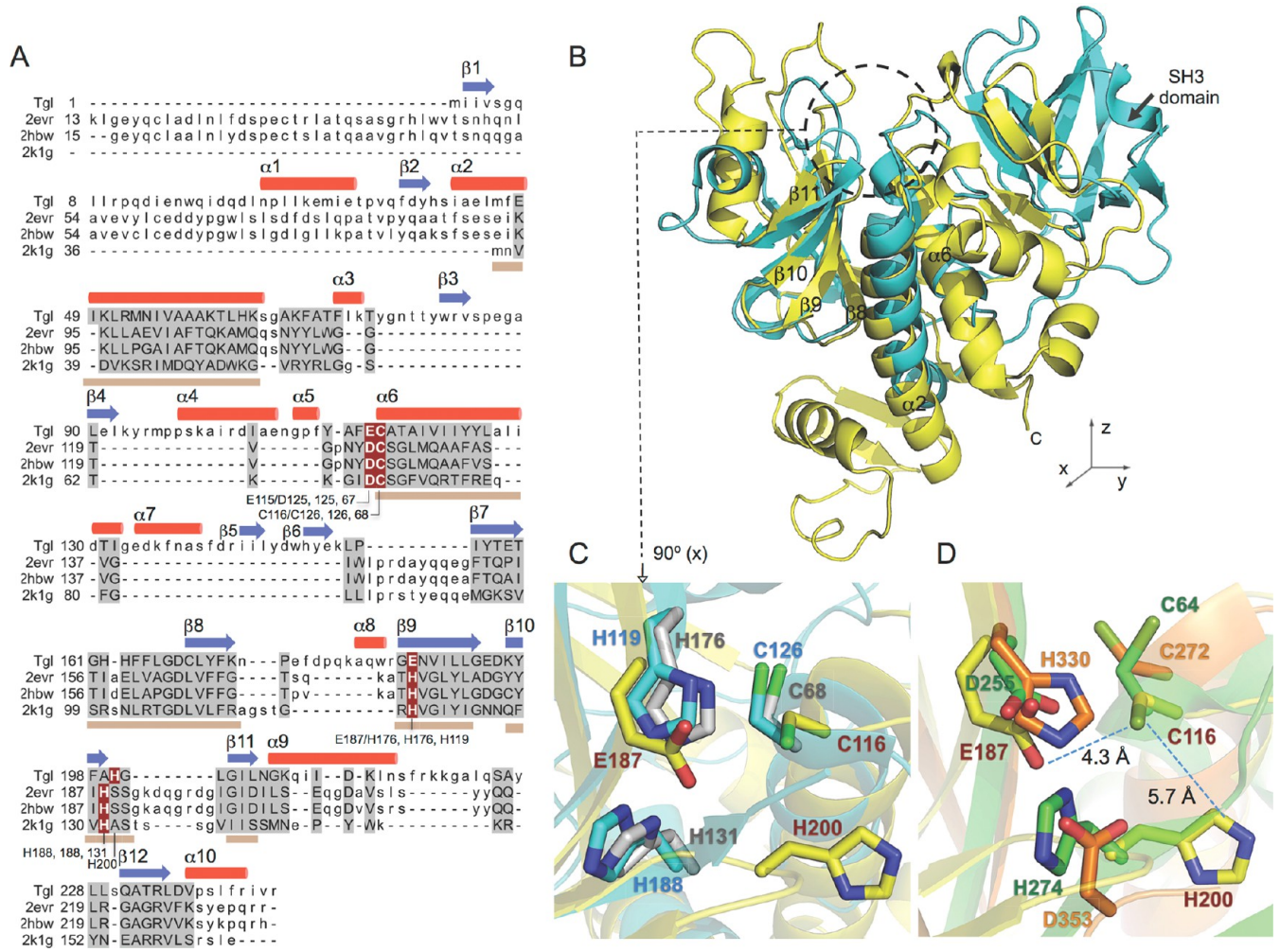


Figure 4. Catalytic core of Tgl has an NlpC/P60-like fold. (A) Structure-based amino acid sequence alignment of Tgl, NpPCP (PDB entry 2EVR), AvPCP (PDB entry 2HBW), and Spr (PDB entry 2KIG), generated with DALI.[48] Helices (red) and β-strands (blue) refer to the Tgl structure. Residues aligned in all four structures are shown in capital letters and highlighted in gray. The secondary structure elements corresponding to the NlpC/P60 core are marked with brown lines below the alignment. Actual or presumptive catalytic residues are colored white against a dark red background. (B) Structural superposition of Tgl (yellow) with NpPCP (cyan). NpPCP (as well as AvPCP) contains an N-terminal SH3-like domain (as indicated), which is absent in Tgl. The figure highlights the conserved core of the NlpC/P60 domain (strands β8–β11 in Tgl, forming a β-sheet packed against α2 and α6; only the clearly visible elements are labeled). The active site region is circled. Superimposition of the putative catalytic residues of Tgl (brown numbering) with (C) NpPCP (cyan) and Spr (gray) and (D) TGase 3 (PDB entry 1L9M, orange) and MTG (PDB entry 1IU4, green). For the animal-type TGases (e.g., TGase 3), the arrangement of the catalytic residues (Cys-His-Asp) is very similar. The inversion of the His and Asp/Glu residues of MTG/Tgl when compared to those of animal TGases is illustrated. MTG shows a catalytic Cys-Asp dyad instead of the typical triad (see also Figure S2).

30 Å and comprises three β-sheets (two-, four-, and six-β-stranded) and 10 helices, of which three are 3_{10}-helices (Figure 2D).

Cys116, essential for the activity of Tgl both *in vivo* and *in vitro*,[45,46] is located at the N-terminus of a long helix, α6 (Figure 2D). In monomer A, the side chain of Cys116 ($Cys116^{A}$) is split into two conformations and additional electron density consistently appeared around the Sγ atom. This extra density was modeled as cysteamine ($-S-CH_2-CH_2-NH_2$) covalently bound to conformer B and was refined to an occupancy of 88% (Figure 2A). For human TGase 2, under reducing conditions, cystamine ($NH_2-CH_2-CH_2-S-S-CH_2-CH_2-NH_2$) is readily reduced to two cysteamine ($NH_2-CH_2-CH_2-SH$) molecules that then act as a K donor substrate, competitively inhibiting the enzyme.[43] In contrast, the structure deriving from inhibition of Tgl by cystamine is suggestive of a disulfide-exchange reaction with the catalytic Cys116, and evidence of an alternative mechanism of inhibition that may occur with many different TGases.

In monomer B, $Cys116^{B}$ was modeled in a triple conformation (Figure 2B) with two nearby water molecules (not shown). These water molecules are within H-bonding distances to $Glu187^{B}$ and to one of two alternate conformations of the $Glu115^{B}$ side-chain. In monomer A, the second conformation of Glu115 is not present and a molecule of ethylene glycol occupies the corresponding space (Figure 2A). In all, the results suggest high flexibility around Cys116 and Glu115.

The Hydrophobic Tunnel of Tgl and Substrate Accessibility. Strikingly, the Tgl structure shows a tunnel near the surface that crosses the molecule from side to side (Figure 3A) and is not apparent in monomer A because of the covalent binding of cysteamine to Cys116 (not shown). The tunnel is ~15 Å long, with its narrowest point (4.9 Å)

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723–5734

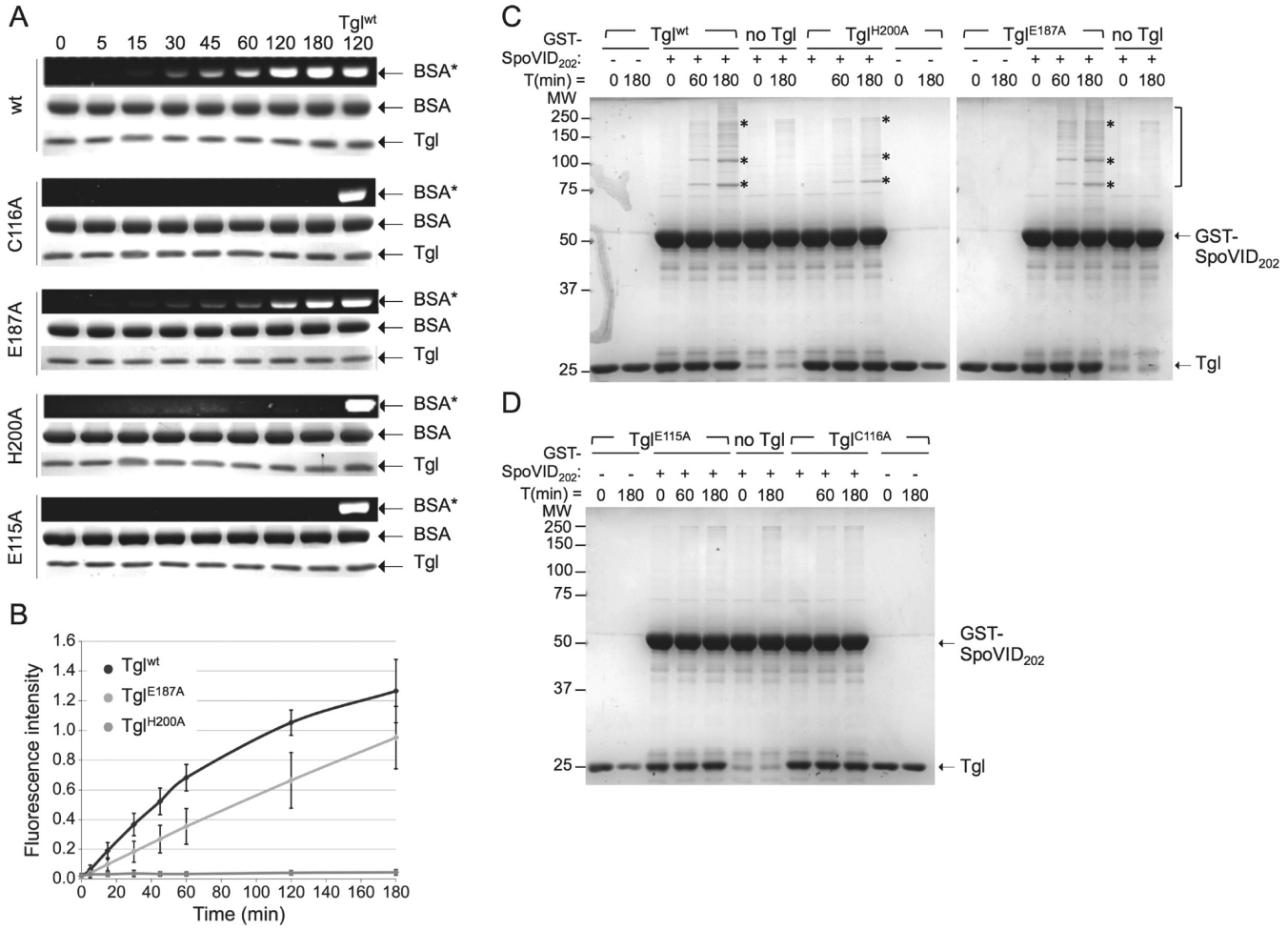


Figure 5. Activity assays with wild-type and mutant forms of Tgl. (A) Enzyme activity was monitored for Tgl^{wt} and for the Tgl^{C116A}, Tgl^{E187A}, Tgl^{H200A}, and Tgl^{E115A} forms by the incorporation of fluorescent dansylcadaverine into BSA, as detected under UV light (BSA*) after SDS−PAGE of the reaction products. The top rows of each panel represent samples taken at the indicated reaction times (in minutes). The middle and bottom rows show the regions of the same gel, stained with Coomassie. The fluorescence signal was normalized using the level of BSA* detected after incubation with Tgl^{wt} for 120 min. (B) The normalized fluorescence was plotted for the various forms of the enzyme, except for Tgl^{C116A} and Tgl^{E115A} for which no labeled BSA could be detected; represented are the mean and standard deviation values for each time point. (C and D) Cross-linking assays of GST-$SpoVID_{202}$ with Tgl^{wt}, Tgl^{H200A}, or Tgl^{E187A} (C) and Tgl^{E115A} or Tgl^{C116A} (D). GST-$SpoVID_{202}$ was incubated with the different forms of Tgl for the indicated times (in minutes), after which the reaction mixture was resolved by SDS−PAGE and the gels were stained with Coomassie. The parentheses indicate the region of the gel where cross-linked products are found (absent from control reactions where GST-$SpoVID_{202}$, Tgl^{wt}, or the mutant enzymes were incubated alone). The bands labeled with an asterisk were subjected to mass spectrometry and N-terminal sequence analysis. The slight decrease in the level of Tgl following incubation for 180 min, especially in the case of Tgl^{H200A} or Tgl^{E115A}, is likely due to precipitation as it proved to be independent of enzymatic activity and was not seen in all the replicas of the assays.

approximately at the center (Figure 3B). The front and back entrances are 6.8 and 7.7 Å across, respectively (Figure 3C,D). The front side of the tunnel includes residues Phe72, Trp149, Tyr171, and Leu214 (Figure 3C) and shows a network of solvent molecules, including a PEG fragment, while the backside has fewer water molecules (not shown). A conspicuous feature of the backside is the close stacking of Trp184 with Phe217 (the shortest distance is ~3.5 Å) (Figure 3D). Together with Phe69, these aromatic residues are in a parallel packing and create a hydrophobic environment to which Leu202 also contributes (Figure 3D). With the exception of Phe217, all of the aromatic residues at both tunnel entrances are highly conserved among Tgl orthologues as shown by a sequence Logo[47] (Figure 3E and Figure S1) and may contribute to the insulation of the active site from water.

The left side wall of the channel, as viewed from its front side entrance, includes Phe69, Phe72, Glu115, the catalytic Cys116, Ala117, Asp148, and Trp149 (Figure 3C,D). The right side comprises Tyr171, Gln183−Asn188, His200, and Gly201 (Figure 3C,D), most of which are highly conserved (Figure 3E and Figure S1). Some right wall residues show high thermal displacement parameters (*B* factors), in particular Arg185, suggesting flexibility (Figures 2C and 3A). Flexibility in this region of the protein may be important for displacement of the ceiling of the tunnel, so that following catalysis, the cross-linked protein product may be released and the enzyme regenerated.[15,19]

The Catalytic Core of Tgl Adopts the NlpC/P60 Papain-like Fold. Proteins related to Tgl at the primary structure level can be found in only spore-forming bacteria of the *Bacillus* and *Clostridium* genera, and closely related organisms.[30,31] However, a DALI search[48] reveals structural similarity between Tgl and several proteins having in common the presence of an NlpC/P60 core. The highest hits were with

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723−5734

the soluble domain (residues 37−162) of the Spr protein from *E. coli*[10] and two homologous cyanobacterial proteins, NpPCP from *Nostoc punctiforme* and AvPCP from *Anabaena variabilis*[9] (Figure 4A,B). All three proteins are peptidoglycan cysteine endopeptidases (PCPs) that exhibit *Z* scores around 7 and rmsds between 2.3 and 2.7 Å for the 188 (for NpPCP and AvPCP) or 126 (for Spr) residues aligned with Tgl. Structural elements shared by Tgl and the PCPs include helices α2 and α6, which are longer in Tgl, and four β-strands (β8−β11) of an antiparallel β-sheet element that is packed against α6 (Figure 4A,B). A long α-helix with the catalytic Cys residue at its N-terminus (Cys116 in Tgl), against which is packed a three-stranded antiparallel β-sheet that contributes with additional catalytic residues, typifies the core NlpC/P60 fold.[8] The core of the NlpC/P60 domain is thought to represent the ancestral, minimal fold of the papain-like cysteine proteases and TGases.[8] In the papain-like and NlpC/P60 proteases, the core β-sheet element is part of a five-stranded β-barrel, while in animal TGases, this β-barrel consists of eight strands.[8−10] The catalytic core of Tgl thus adopts the α+β fold of the core NlpC/P60 domain, with the β-barrel element consisting of six strands [β7−β12 (Figures 2D and 4A,B)].

Active Site Architecture. NpPCP and AvPCP, and presumably also Spr, use a Cys-His-His catalytic triad to cleave the peptidoglycan stem peptide.[9,10] A closer inspection of the superimposed structures of Tgl, Spr, and NpPCP in the vicinity of the catalytic Cys shows that the position occupied by one of the catalytic His residues in the cell wall endopeptidases corresponds to Glu187 in Tgl, while the second His is close to His200 (Figure 4C). Both His200 (on a loop between strands β10 and β11) and Glu187 (on β9) are located within the six-stranded β-sheet of the NlpC/P60 core of Tgl (Figures 2D and 4A,B). Thus, the position of His200 and Glu187 in the structure seems to suggest that both are catalytic residues (but see below).

We also superimposed the region around Cys116 of Tgl with the active site of MTG, the only other current structurally characterized microbial TGase,[24,25] which, as we show, also possesses an NlpC/P60 core (Figure S2), and human TGase 3, an archetypal animal TGase. In TGases, catalysis involves a Cys-His-Asp triad (for the animal enzymes) or a Cys-Asp dyad (for MTG) (Figure 1B). The relative position of the catalytic residues in MTG and TGase 3 results from a permutation, such that the His and Asp residues in MTG spatially correspond to the Asp and His residues of TGase 3, respectively (Figure 4D).[24] This is also a functional permutation as Asp255 of MTG plays the role of the catalytic His in the animal-like TGases while His274 of MTG is not essential for catalysis.[24,49] The superimposition of the catalytic Cys in Tgl (Cys116), MTG (Cys64), and TGase 3 (Cys272) places His200 of Tgl in the same relative position of His274 in MTG, raising the possibility that His200 is also not essential for catalysis. Furthermore, it also suggests that Glu187 in Tgl could be a catalytic residue: not only does it occupy the position equivalent to His330 in TGase 3 or Asp255 in MTG, but the nearest side-chain atoms of Glu187 are close (~4.3 Å) to the Sγ atom of the catalytic Cys116 residue (Figure 4D).

Probing the Active Site of Tgl. To define the catalytic residues of Tgl, forms of the enzyme bearing single Ala substitutions of residues Cys116 (Tgl^{C116A}), Glu187 (Tgl^{E187A}), and His200 (Tgl^{H200A}) were purified and tested in parallel with the wild-type enzyme (Tgl^{wt}), for activity. In addition to protein cross-linking, a second transamidation reaction catalyzed by TGases involves a protein substrate and a primary amine present in simple molecules or in naturally occurring polyamines, which results in amine incorporation.[1,2] In previous work, we have shown that BSA serves as a substrate for Tgl cross-linking, and we reasoned that it could also be used in amine incorporation reactions.[33] We first monitored the incorporation of the fluorescent K donor substrate dansylcadaverine (dansyl-cd) into bovine serum albumin (BSA), the Q substrate. Tgl^{wt} efficiently catalyzed the transfer of dansyl-cd to BSA, resulting in fluorescent BSA (Figure 5A), with the level of labeling increasing linearly with time, during the first 45 min of the assay (Figure 5B). Importantly, cross-linking of BSA, auto-cross-linking of Tgl, or labeling of Tgl by dansyl-cd was not significant during the time of the assay (not shown). In confirmation and extension of earlier results,[45,46] the C116A substitution resulted in a loss of enzymatic activity (Figure 5A). In contrast, labeling of BSA could be detected for Tgl^{H200A} (albeit at low levels) and Tgl^{E187A}, the latter showing ~60% of the wild-type enzyme activity (Figure 5A,B).

Second, we assessed the ability of the various forms of Tgl to cross-link a glutathione *S*-transferase (GST) fusion protein. GST (~30 kDa) can serve as a substrate for TGases.[38,50] However, in screening a collection of proteins available in the laboratory, we found that a fusion of GST to the first 202 amino acids of coat protein SpoVID[34] was cross-linked by Tgl. The use of GST-$SpoVID_{202}$ (50 kDa) increased the mass difference to Tgl (28 kDa) and allowed the enzyme and substrate to be unambiguously identified and their levels to be monitored on a Coomassie-stained gel. Incubation of Tgl^{wt}, Tgl^{E187A}, or Tgl^{H200A} with GST-$SpoVID_{202}$ resulted in the formation of high-molecular weight species over time (indicated by the parentheses in Figure 5C), although at reduced levels for Tgl^{H200A}. In contrast, these species were not detected upon incubation of GST-$SpoVID_{202}$ with Tgl^{C116A} or when Tgl or GST-$SpoVID_{202}$ were incubated alone (Figure 5C,D). The more defined 75, 100, and 250 kDa bands detected after incubation with Tgl^{wt}, Tgl^{E187A}, and Tgl^{H200A} for 180 min (Figure 5C, asterisks) were shown by mass spectrometry to contain sequences of both Tgl and GST-$SpoVID_{202}$. While Tgl alone shows little auto-cross-linking activity under the conditions of the assay (Figure 5C, lanes with no GST-$SpoVID_{202}$), these results suggest that Tgl^{wt}, Tgl^{E187A}, or Tgl^{H200A} catalyzes the cross-linking to GST-$SpoVID_{202}$. In any event, formation of the 75, 100, and 250 kDa species is the result of Tgl activity. Thus, both the dansyl-cd incorporation and cross-linking assays show that unlike Cys116, Glu187 and His200 are not essential for catalysis.

Interactions among Residues near the Catalytic Cys. Our activity assays suggest that the most likely partner and proton acceptor for Cys116 is His200. However, this seems to be in contradiction with the apparent orientation of His200 in the crystal structure. As shown in Figure 4D, the His200 imidazole ring is not only too far from Sγ of Cys116 to allow proton abstraction (the shortest distance is ~5.7 Å), as its side chain is also oriented away from the active site Cys (Nδ1 is H-bonded to a water molecule). Because the X-ray structures of Tgl were obtained at pH 4.0, outside the range for optimal enzyme activity,[51] we reasoned that our crystallization conditions may have caused artifacts in the relative orientation of critical side chains in the active site region, specifically those of Glu187 and His200. Because we could not obtain crystals at a higher pH (see Experimental Procedures), molecular dynamics (MD) simulations were conducted at pH 7 (the

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723−5734

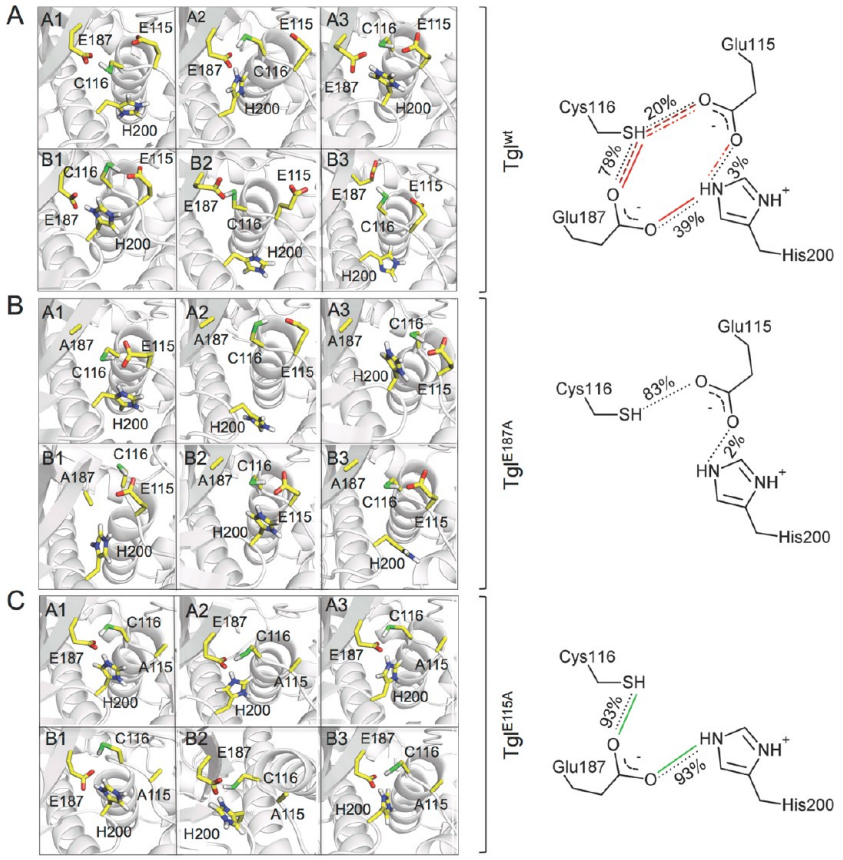


Figure 6. Behavior of Tgl^{wt}, Tgl^{E187A}, and Tgl^{E115A} in MD simulations. Two molecules were observed in the final model of the X-ray structure, and both were simulated, named monomers A and B, and three simulations conducted for each monomer (see also Figure S3). Representation of the MD simulations for (A) Tgl^{wt}, (B) Tgl^{E187A}, and (C) Tgl^{E115A}. In the left-hand side of each panel is represented the average structure of the active site of Tgl during the last 5 ns of simulation. The right side of each panel shows a schematic representation of the interactions (black dotted lines) formed among the residues studied and their frequency (numbers above dotted lines) and correspondent negative (red lines) or positive (green lines) correlations. For the sake of simplicity, the interaction between His200 and Glu187/Glu115 is depicted as occurring through the Nϵ2 proton of His200; however, the interaction was also detected via the Nδ1 proton of His200, as seen on the average structure on the left side (see also Tables S4 and S5). From the average structure of the active site of Tgl, it is clear that His200 does not interact with Cys116. Furthermore, in (A) Tgl^{wt} and (B) Tgl^{E187A}, the simultaneous interaction of Glu187 or Glu115 with both Cys116 and His200 is unfavorable unlike what is seen with Tgl^{E115A} (C).

MD simulations are described in detail in the Supporting Information). Our results indicate that the orientation of His200 detected in the crystal structure is a consequence of the low-pH crystallization conditions and that an increase in pH can change the orientation of His200 toward the active site pocket.

To gain further insight into the contribution of His200 and Glu187 to the activity of Tgl, we analyzed the hydrogen bonding interactions among the Cys116, Glu187, and His200 residues, as well as the correlations between all pairs of hydrogen bonds at pH 7.0 (Figure 6A; see also Tables S4 and S5). The results suggest that Cys116 and His200 do not interact through hydrogen bonding (Figure 6A), and therefore, His200 probably does not act as the proton acceptor for Cys116 (step 1 in Figure 1B). However, Glu187 forms a hydrogen bond with Cys116 with a high frequency (of ~78%) (Figure 6A). Furthermore, while Glu187 may interact with His200, the simultaneous interaction of Glu187 with both His200 and Cys116 appears to be unfavorable, as the Cys116-Glu187 and His200-Glu187 pairs show a negative correlation (Figure 6A). Therefore, and in spite of its presumed orientation toward the active site at pH 7.0, as suggested by our MD analysis (above), His200 does not appear to be a catalytic residue. This is consistent with the observation that Tgl^{H200A} still retains activity in both amine incorporation and cross-linking assays (Figure 5A–C). In contrast, Glu187 appears to be a catalytic residue serving the role of the proton acceptor for Cys116. However, Tgl^{E187A} still retains considerable activity (Figure 5A–C), raising the possibility that another residue, near the catalytic Cys116, can compensate for the absence of Glu187.

A New Catalytic Residue. Our MD analysis of the interplay between Cys116 and its neighboring residues suggests that a second Glu residue, Glu115, may form a transient hydrogen bond with Cys116, with a frequency of ~20% (Figure 6A). As such, at pH 7, Glu115 is most likely deprotonated and may function as the proton acceptor for Cys116 (step 1 in Figure 1B). Furthermore, we found a strong negative correlation between the Cys116-Glu187 and the Cys116-Glu115 pairs, suggesting that the two acids compete for the catalytic Cys (Figure 6A). This could explain why Tgl^{E187A} still shows enzymatic activity (Figure 5A–C). His200 and Glu115 also appear to form a hydrogen bond, but the simulations suggest that this is a rare event (frequency of 3%), especially when compared to formation of the His200–Glu187 salt bridge (Figure 6A; see also above). In addition, the simulations also suggest a negative correlation between the Cys116-Glu115 and His200-Glu115 pairs (Figure 6A). We infer that, as was also

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723–5734

seen for Glu187, Glu115 does not tend to simultaneously interact with both Cys116 and His200. Finally, we also found strong positive correlations between the His200-Glu187 and Cys116-Glu115 interactions (Figure 6A). Thus, the interaction between His200 and Glu187 weakens the Glu187-Cys116 interaction, so that Cys116 turns to the free Glu115. Similarly, the Cys116-Glu187 and His200-Glu115 pairs are positively correlated (Table S5). It follows that the His200-Glu115 interaction directs the catalytic Cys116 to the alternative Glu187. While strengthening the view that His200 is not a catalytic residue, these observations suggest that Tgl uses a catalytic dyad formed by either Cys116 and Glu187 or Cys116 and Glu115.

A Partially Redundant Catalytic Dyad: Nonreciprocal Substitution of Glu187 with Glu115. The results presented in the preceding section suggested that Tgl has a catalytic dyad. This dyad is however atypical, as the two acidic residues, Glu187 and Glu115, appear to be redundant. If so, then an E115A substitution would not be expected to eliminate enzyme activity. To test this prediction, purified Tgl^{E115A} was tested for activity. Contrary to our expectations, however, the single E115A substitution rendered Tgl unable to detectably label BSA (Figure 5A) or to cross-link GST-$SpoVID_{202}$ (Figure 5D). Thus, while Glu115 seems to replace Glu187, the reverse is not observed. In an attempt to understand this effect, we conducted MD simulations using *in silico* mutants (Figure 6B,C and Figure S3D,E). The results show that in Tgl^{E187A} the hydrogen bond between Cys116 and Glu115 becomes very likely, occurring with a frequency of ∼80%, i.e., 4 times higher than in the wild-type enzyme (Figure 6B). This is in line with the idea that in the absence of the Glu187 side chain, Cys116 readily interacts with Glu115. Also, in Tgl^{E187A}, the frequency of the interaction between His200 and Glu115 is low [∼2% (Figure 6B and Figure S3D)], and in fact, the average structures of the last 5 ns of the simulations with Tgl^{E187A} show that His200 does not tend to interact with any other active site residue (Figure 6B and Table S4). Likewise, the simulations with the Tgl^{E115A} mutant (Figure 6C and Figure S3E) indicate that Cys116 tends to interact with the remaining acidic residue, Glu187, with the frequency of hydrogen bonding between Cys116 and Glu187 increasing relative to that of the wild type (from 78.3 to 93.2%). Importantly, however, and unlike the results for the wild-type enzyme or Tgl^{E187A}, the absence of Glu115 additionally leads to the frequent interaction between His200 and Glu187 [frequency of 93% (Figure 6C and Figure S3E)]. Thus, under these conditions, the Glu187-His200 interaction is stable. The once negative correlation between the Cys116-Glu187 and His200-Glu187 pairs is now positive, and both Cys116 and His200 appear to interact simultaneously with Glu187, which acts as a proton acceptor for both (Figure 6C). Considering that in Tgl^{E115A}, Glu187 is the proton acceptor for Cys116, then the simultaneous interaction with His200 would weaken its ability to abstract a proton and to produce a thiolate ion (step 1 in Figure 1B). The result would be the decreased reactivity of Cys116. This possibly explains the strongly reduced activity of Tgl^{E115A} and the partial redundancy of the Tgl Cys116-Glu187 or Cys116-Glu115 catalytic dyad.

DISCUSSION

Here, we present the crystal structure and functional characterization of Tgl, a 28 kDa sporulation-specific bacterial TGase. Sporulation emerged at the basis of the Firmicutes phylum some 3 billion years ago.[30] A much larger TGase (YebA, with a predicted molecular weight of 75 kDa), related to the animal-type enzymes, is also detected in *B. subtilis*.[11] Thus, it seems possible that Tgl emerged as a sporulation-specific TGase dedicated to the cross-linking of the spore surface.[46,52−54] Although the evolutionary history of Tgl is presently unknown, its structural similarity to a group of bacterial cell wall endopepetidases suggests that regardless of their origin, all TGases evolved from cysteine proteases. Moreover, it shows that the adoption of the NlpC/P60 core by TGases is ancient.

Another finding of this study is that the catalytic center of Tgl is located within a 15 Å long tunnel that crosses the molecule from side to side, with both entrances surrounded by hydrophobic residues (Figure 3B−D). This tunnel is a second feature that Tgl shares with the active form of animal TGases. The Tgl tunnel appears to be wide enough (∼6 Å) to accommodate a peptide-chain backbone bearing a reactive Q or K residue. Presumably, this tunnel contributes to waterproofing the active site favoring transamidation over deamidation reactions. The insulation of the catalytic Cys within a hydrophobic tunnel that channels the Q and K substrates to the active site has been proposed as an essential structural requirement for protein cross-linking by TGases.[19] This requirement is fulfilled in the active forms of human TGase 2 and 3 and Factor XIIIa[15−17] and may be a conserved feature of the animal-type enzymes.[4] The discovery of a tunnel in Tgl suggests that the adoption of this structural feature was also an early event. Strikingly, in MTG and GP42, the active site is close to the surface of the molecule.[24,26] However, the prosequence that covers the active site cleft of MTG positions two Tyr residues on top of the catalytic residues,[25] and in GP42, several aromatic residues are located above the catalytic Cys.[26] It is tempting to suggest that binding of MTG or GP42 substrates may exclude water from their active sites during catalysis. In any event, the animal-like enzymes,[8] as well as GP42,[26] MTG (see Figure S2), and the smaller sized Tgl, all show an NlpC/P60 core. This suggests that an NlpC/P60 core, together with a mechanism for active site insulation, may be universal requirements for protein cross-linking.

Both entries of the tunnel are most likely docking sites for the Q and K substrates of Tgl. Interestingly while many of the residues at both entrances of the tunnel are conserved among Tgl orthologues, some stretches are highly variable (Figure 3E). It is possible that this variability represents specificity toward different substrates. We do not presently know which sides of the tunnel are engaged by the Q and K substrates. However, waterproofing of the K acceptor side is important to reduce the potential for hydrolytic reactions, as shown for human TGase 2.[15] Thus, the K acceptor side of Tgl may correspond to the more hydrophobic "back" entrance of the tunnel (Figure 3E). His200, at the "back" entrance of the Tgl tunnel, does not seem to be a catalytic residue (see above), as the simultaneous interactions among Cys116, Glu187, and His200 are unfavorable. Rather, His200 may be involved in substrate recognition, contributing in that way to the overall activity of the enzyme. However, the initial deprotonation of the acyl acceptor (K) substrate, which in solution is likely to exist predominantly in its protonated form, has been proposed to be an important step in the reaction mechanism of TGases.[17,55] In Factor XIII and in human TGase 2, His residues outside the catalytic triad appear to facilitate the nucleophilic attack by the K substrate on the acyl−enzyme intermediate.[17,55] It is tempting to suggest that His200 plays a similar role, perhaps coupling binding of the K substrate to its deprotonation. An Ala

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723−5734

substitution of His274 in MTG, which occupies the same relative position of His200 in Tgl (Figure 4D), also reduces but does not eliminate the activity of the enzyme.[24,49] In MTG, His274 was proposed to be involved in substrate interactions.[49] However, it could also be involved in the initial deprotonation of the K substrate, as suggested above for Tgl.

A third significant finding is that Tgl has a distinctive catalytic center. While Glu187 appears to be the primary proton acceptor for Cys116, a second acidic residue, Glu115, in the close vicinity of Cys116, can substitute for Glu187, at least when the latter is substituted with an Ala. The reciprocal substitution (E115A) abrogates enzyme activity, possibly because in the mutant, as suggested by the MD simulations, Glu187 interacts with equal frequency with both Cys116 and His200, reducing its capacity to deprotonate Cys116. While Glu115 (but not Glu187) is strictly conserved among Tgl orthologues (Figure 3E), further studies are needed to elucidate its role in catalysis. Both the MD simulations and the activity assays converge to indicate that Tgl has a catalytic dyad. Why the two bacterial TGases, MTG and Tgl, employ a catalytic dyad[24] whereas the animal-type enzymes use a triad is presently unclear. However, it has been suggested that in animal-like TGases in addition to the classical Cys-His-Asp triad, a Cys-His catalytic dyad may also be involved in catalysis.[17] Interestingly, papain also seems to function through a dual catalytic Cys26-His159-Asp158/Asn175 triad.[56] The dual partial redundant catalytic dyad of Tgl (Cys116−Glu187 or Glu115) implies that catalysis may take place from one or two distinct steric positions. Perhaps in Tgl different physiological substrates are handled in distinct ways. Variations in the spatial arrangement of the catalytic residues, such as the permutations seen in MTG and Tgl (Figure 3), may in fact reflect enzyme−substrate interactions. An extreme example may be found in GP42, in which the His of the Cys-His-Asp triad is located in the same helix, part of the NlpC/P60 core, that carries the catalytic Cys, while the β-barrel element contributes with aromatic residues possibly involved in insulation of the catalytic Cys (as mentioned above) and substrate binding. Rearrangement of the triad was proposed to have occurred during evolution so that aromatic residues required for substrate specificity could be accommodated.[26]

Substrate interactions may also induce important functional rearrangements of the active site region. Our MD simulations also show that unlike what has been proposed for the animal TGases, where the catalytic Cys appears as a thiolate ion in the initial step of the reaction mechanism (Figure 1B, step 1), in Tgl the Cys appears to be protonated even at high pH values (Supporting Information). It is thus possible that in Tgl binding of a substrate will induce modifications in the pK_a of the catalytic Cys and the formation of the thiolate ion will be concomitant with the initial nucleophilic attack on the Q substrate and the transfer of the proton to Glu187 or Glu115.

TGases catalyze a variety of reactions of which protein cross-linking is the most notorious. They play important roles in cell and tissue morphogenesis and participate in a variety of processes, from clotting phenomena to development. Proteins homologous to the animal-type TGases, of which the human enzymes are the most extensively studied, show a broad phylogenetic distribution. They are structurally and mechanistically related to the papain-like cysteine proteases and, accordingly, show an NlpC/P60-type catalytic core, though to represent the minimal structural unit of the thiol protease fold. The TGases are large multidomain enzymes, with complex activation mechanisms, making the identification of the essential requirements for protein cross-linking difficult. The bacterial TGase, Tgl, a single-domain enzyme, is the smallest TGase characterized to date. It most likely emerged at the basis of the Firmicutes phylum, some 3 billion years ago, as an enzyme dedicated to the cross-linking of the spore coat. Tgl is structurally related to a family of bacterial cell wall endopeptidases and has an NlpC/P60 core and a hydrophobic tunnel that insulates the active site. These features may thus be essential requirements for protein cross-linking. Moreover, it has a partially redundant catalytic dyad, reminiscent of papain. The properties of Tgl lend support to the notion that all TGases have evolved from ancestral thiol proteases.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.5b00661.

Additional experimental procedures, detailed description of the molecular dynamics simulations methods and results, ClustalW alignment of Tgl orthologues, evidence of the NlpC/P60 core in MTG, data collection, phasing, and refinement statistics of the Tgl structure, and interaction frequency and correlation analysis among active site residues (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*Phone: (+351) 214469747. E-mail: archer@itqb.unl.pt.
*Phone: (+351) 214411277. E-mail: aoh@itqb.unl.pt.

Present Addresses

⊥C.G.F.: Department of Microbiology and Immunology, Emory University School of Medicine, Atlanta, GA 30322.
#A.I.: Fundação para a Ciência e Tecnologia, Av. D. Carlos I, 126, 1249-074 Lisboa, Portugal.

Author Contributions

D.P. collected the diffraction data. D.P., J.A.B., and M.A. analyzed the diffraction data and determined the structure. C.G.F. and A.I. performed mutagenic studies. C.G.F. conducted the functional characterization of Tgl and its mutants. D.L. and C.M.S. performed the molecular dynamics studies. C.G.F., D.P., M.A., A.O.H., and M.A.C. conceived the experimental design and interpreted the data. C.G.F., D.P., D.L., J.A.B., C.M.S., M.A., A.O.H., and M.A.C. prepared the manuscript.

Author Contributions

C.G.F. and D.P. contributed equally to this work.

Funding

This work was funded by "Fundação para a Ciência e a Tecnologia" (FCT) trough Ph.D. (SFRH/BD/43200/2008, SFRH/BD/14384/2003, and SFRH/BD/28269/2006) and postdoctoral (SFRH/BPD/8967/2002 and SFRH/BPD/79224/2011) fellowships and FCT Grants (Pest-E/EQB/LA0004/2011, POCI/BIA-BCM/60855/2004, POCTI/BCI/48647/2002, and PTDC/BIA-PRO/118535/2010). Support from the European Commission under the sixth Framework Programme through the Key Action "Strengthening the European Research Area Programme" (Contract n° RII3-CT-2004-506008) was provided to collect X-ray data at EMBL/DESY (Hamburg, Germany) and SLS (Villigen, Switzerland).

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723−5734

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Pedro Matias (ITQB) for his help with data collection. We also acknowledge support from the European Synchrotron Radiation Facility (ESRF, Grenoble, France).

■ ABBREVIATIONS

TGase, transglutaminase; PCP, peptidoglycan cysteine endopeptidase; GST, glutathione *S*-transferase; MD, molecular dynamics; PDB, Protein Data Bank.

■ REFERENCES

(1) Griffin, M., Casadio, R., and Bergamini, C. M. (2002) Transglutaminases: nature's biological glues. *Biochem. J. 368*, 377−396.
(2) Lorand, L., and Graham, R. M. (2003) Transglutaminases: crosslinking enzymes with pleiotropic functions. *Nat. Rev. Mol. Cell Biol. 4*, 140−156.
(3) Kim, S. Y., Jeitner, T. M., and Steinert, P. M. (2002) Transglutaminases in disease. *Neurochem. Int. 40*, 85−103.
(4) Iismaa, S. E., Mearns, B. M., Lorand, L., and Graham, R. M. (2009) Transglutaminases and disease: lessons from genetically engineered mouse models and inherited disorders. *Physiol. Rev. 89*, 991−1023.
(5) Pedersen, L. C., Yee, V. C., Bishop, P. D., Le Trong, I., Teller, D. C., and Stenkamp, R. E. (1994) Transglutaminase factor XIII uses proteinase-like catalytic triad to crosslink macromolecules. *Protein Sci. 3*, 1131−1135.
(6) Ahvazi, B., Boeshans, K. M., Idler, W., Baxa, U., and Steinert, P. M. (2003) Roles of calcium ions in the activation and activity of the transglutaminase 3 enzyme. *J. Biol. Chem. 278*, 23834−23841.
(7) Boeshans, K. M., Mueser, T. C., and Ahvazi, B. (2006) A three-dimensional model of the human transglutaminase 1: insights into the understanding of lamellar ichthyosis. *J. Mol. Model. 13*, 233−246.
(8) Anantharaman, V., and Aravind, L. (2003) Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes. *Genome biology 4*, R11.
(9) Xu, Q., Sudek, S., McMullan, D., Miller, M. D., Geierstanger, B., Jones, D. H., Krishna, S. S., Spraggon, G., Bursalay, B., Abdubek, P., Acosta, C., Ambing, E., Astakhova, T., Axelrod, H. L., Carlton, D., Caruthers, J., Chiu, H. J., Clayton, T., Deller, M. C., Duan, L., Elias, Y., Elsliger, M. A., Feuerhelm, J., Grzechnik, S. K., Hale, J., Won Han, G., Haugen, J., Jaroszewski, L., Jin, K. K., Klock, H. E., Knuth, M. W., Kozbial, P., Kumar, A., Marciano, D., Morse, A. T., Nigoghossian, E., Okach, L., Oommachen, S., Paulsen, J., Reyes, R., Rife, C. L., Trout, C. V., van den Bedem, H., Weekes, D., White, A., Wolf, G., Zubieta, C., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A., and Wilson, I. A. (2009) Structural basis of murein peptide specificity of a gamma-D-glutamyl-l-diamino acid endopeptidase. *Structure 17*, 303−313.
(10) Aramini, J. M., Rossi, P., Huang, Y. J., Zhao, L., Jiang, M., Maglaqui, M., Xiao, R., Locke, J., Nair, R., Rost, B., Acton, T. B., Inouye, M., and Montelione, G. T. (2008) Solution NMR structure of the NlpC/P60 domain of lipoprotein Spr from *Escherichia coli*: structural evidence for a novel cysteine peptidase catalytic triad. *Biochemistry 47*, 9715−9717.
(11) Makarova, K. S., Aravind, L., and Koonin, E. V. (1999) A superfamily of archaeal, bacterial, and eukaryotic proteins homologous to animal transglutaminases. *Protein Sci. 8*, 1714−1719.
(12) Gundemir, S., Colak, G., Tucholski, J., and Johnson, G. V. W. (2012) Transglutaminase 2: A molecular Swiss army knife. *Biochim. Biophys. Acta, Mol. Cell Res. 1823*, 406−419.
(13) Klöck, C., and Khosla, C. (2012) Regulation of the activities of the mammalian transglutaminase family of enzymes. *Protein Sci. 21*, 1781−1791.
(14) Liu, S., Cerione, R. A., and Clardy, J. (2002) Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity. *Proc. Natl. Acad. Sci. U. S. A. 99*, 2743−2747.
(15) Pinkas, D. M., Strop, P., Brunger, A. T., and Khosla, C. (2007) Transglutaminase 2 undergoes a large conformational change upon activation. *PLoS Biol. 5*, e327.
(16) Ahvazi, B., Kim, H. C., Kee, S. H., Nemes, Z., and Steinert, P. M. (2002) Three-dimensional structure of the human transglutaminase 3 enzyme: binding of calcium ions changes structure for activation. *EMBO J. 21*, 2055−2067.
(17) Stieler, M., Weber, J., Hils, M., Kolb, P., Heine, A., Büchold, C., Pasternack, R., and Klebe, G. (2013) Structure of Active Coagulation Factor XIII Triggered by Calcium Binding: Basis for the Design of Next-Generation Anticoagulants. *Angew. Chem., Int. Ed. 52*, 11930−11934.
(18) Yee, V. C., Pedersen, L. C., Le Trong, I., Bishop, P. D., Stenkamp, R. E., and Teller, D. C. (1994) Three-dimensional structure of a transglutaminase: human blood coagulation factor XIII. *Proc. Natl. Acad. Sci. U. S. A. 91*, 7296−7300.
(19) Nemes, Z., Petrovski, G., Csosz, E., and Fesus, L. (2005) Structure-function relationships of transglutaminases–a contemporary view. *Progress in experimental tumor research. Fortschritte der experimentellen Tumorforschung 38*, 19−36.
(20) Milani, A., Vecchietti, D., Rusmini, R., and Bertoni, G. (2012) TgpA, a protein with a eukaryotic-like transglutaminase domain, plays a critical role in the viability of *Pseudomonas aeruginosa*. *PLoS One 7*, e50323.
(21) Luo, Y., Pfister, P., Leisinger, T., and Wasserfallen, A. (2002) Pseudomurein endoisopeptidases PeiW and PeiP, two moderately related members of a novel family of proteases produced in *Methanothermobacter* strains. *FEMS Microbiol. Lett. 208*, 47−51.
(22) Pfister, P., Wasserfallen, A., Stettler, R., and Leisinger, T. (1998) Molecular analysis of Methanobacterium phage psiM2. *Mol. Microbiol. 30*, 233−244.
(23) King, J. D., Vinogradov, E., Preston, A., Li, J., and Maskell, D. J. (2008) Post-assembly modification of *Bordetella bronchiseptica* O polysaccharide by a novel periplasmic enzyme encoded by *wbmE*. *J. Biol. Chem. 284*, 1474−1483.
(24) Kashiwagi, T., Yokoyama, K., Ishikawa, K., Ono, K., Ejima, D., Matsui, H., and Suzuki, E. (2002) Crystal structure of microbial transglutaminase from *Streptoverticillium mobaraense*. *J. Biol. Chem. 277*, 44252−44260.
(25) Yang, M. T., Chang, C. H., Wang, J. M., Wu, T. K., Wang, Y. K., Chang, C. Y., and Li, T. T. (2011) Crystal Structure and Inhibition Studies of Transglutaminase from *Streptomyces mobaraense*. *J. Biol. Chem. 286*, 7301−7307.
(26) Reiss, K., Kirchner, E., Gijzen, M., Zocher, G., Loffelhardt, B., Nurnberger, T., Stehle, T., and Brunner, F. (2011) Structural and Phylogenetic Analyses of the GP42 Transglutaminase from *Phytophthora sojae* Reveal an Evolutionary Relationship between Oomycetes and Marine *Vibrio* Bacteria. *J. Biol. Chem. 286*, 42585−42593.
(27) Pasternack, R., Dorsch, S., Otterbach, J. T., Robenek, I. R., Wolf, S., and Fuchsbauer, H. L. (1998) Bacterial pro-transglutaminase from *Streptoverticillium mobaraense*-purification, characterisation and sequence of the zymogen. *Eur. J. Biochem. 257*, 570−576.
(28) Zotzel, J., Keller, P., and Fuchsbauer, H. L. (2003) Transglutaminase from *Streptomyces mobaraensis* is activated by an endogenous metalloprotease. *Eur. J. Biochem. 270*, 3214−3222.
(29) Zotzel, J., Pasternack, R., Pelzer, C., Ziegert, D., Mainusch, M., and Fuchsbauer, H. L. (2003) Activated transglutaminase from *Streptomyces mobaraensis* is processed by a tripeptidyl aminopeptidase in the final step. *Eur. J. Biochem. 270*, 4149−4155.
(30) Abecasis, A. B., Serrano, M., Alves, R., Quintais, L., Pereira-Leal, J. B., and Henriques, A. O. (2013) A genomic signature and the identification of new sporulation genes. *Journal of bacteriology 195*, 2101−2115.
(31) Galperin, M. Y., Mekhedov, S. L., Puigbo, P., Smirnov, S., Wolf, Y. I., and Rigden, D. J. (2012) Genomic determinants of sporulation in

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723−5734

Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. *Environ. Microbiol. 14*, 2870−2890.
(32) Henriques, A. O., and Moran, C. P., Jr. (2007) Structure, assembly, and function of the spore surface layers. *Annu. Rev. Microbiol. 61*, 555−588.
(33) Placido, D., Fernandes, C. G., Isidro, A., Carrondo, M. A., Henriques, A. O., and Archer, M. (2008) Auto-induction and purification of a *Bacillus subtilis* transglutaminase (Tgl) and its preliminary crystallographic characterization. *Protein Expression Purif. 59*, 1−8.
(34) Costa, T., Isidro, A. L., Moran, C. P., Jr., and Henriques, A. O. (2006) Interaction between coat morphogenetic proteins SafA and SpoVID. *Journal of bacteriology 188*, 7731−7741.
(35) Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 12−21.
(36) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283−291.
(37) DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, Palo Alto, CA.
(38) Sugimura, Y., Hosono, M., Wada, F., Yoshimura, T., Maki, M., and Hitomi, K. (2006) Screening for the preferred substrate sequence of transglutaminase using a phage-displayed peptide library: identification of peptide substrates for TGase 2 and Factor XIIIa. *J. Biol. Chem. 281*, 17699−17706.
(39) Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods 9*, 671−675.
(40) Berendsen, H. J. C., van der Spoel, D., and van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun. 91*, 43−56.
(41) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput. 4*, 435−447.
(42) Oostenbrink, C., Villa, A., Mark, A. E., and van Gunsteren, W. F. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem. 25*, 1656−1676.
(43) Jeitner, T. M., Delikatny, E. J., Ahlqvist, J., Capper, H., and Cooper, A. J. (2005) Mechanism for the inhibition of transglutaminase 2 by cystamine. *Biochem. Pharmacol. 69*, 961−970.
(44) Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins: Struct., Funct., Genet. 50*, 437−450.
(45) Kobayashi, K., Hashiguchi, K., Yokozeki, K., and Yamanaka, S. (1998) Molecular cloning of the transglutaminase gene from *Bacillus subtilis* and its expression in *Escherichia coli*. *Biosci., Biotechnol., Biochem. 62*, 1109−1114.
(46) Zilhao, R., Isticato, R., Martins, L. O., Steil, L., Volker, U., Ricca, E., Moran, C. P., Jr., and Henriques, A. O. (2005) Assembly and function of a spore coat-associated transglutaminase of *Bacillus subtilis*. *Journal of bacteriology 187*, 7753−7764.
(47) Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res. 14*, 1188−1190.
(48) Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol. 233*, 123−138.
(49) Tagami, U., Shimba, N., Nakamura, M., Yokoyama, K. i., Suzuki, E. i., and Hirokawa, T. (2009) Substrate specificity of microbial transglutaminase as revealed by three-dimensional docking simulation and mutagenesis. *Protein Eng., Des. Sel. 22*, 747−752.
(50) Piredda, L., Farrace, M. G., Lo Bello, M., Malorni, W., Melino, G., Petruzzelli, R., and Piacentini, M. (1999) Identification of 'tissue' transglutaminase binding proteins in neural cells committed to apoptosis. *Faseb J. 13*, 355−364.
(51) Kobayashi, K., Suzuki, S. I., Izawa, Y., Yokozeki, K., Miwa, K., and Yamanaka, S. (1998) Transglutaminase in sporulating cells of *Bacillus subtilis*. *J. Gen. Appl. Microbiol. 44*, 85−91.
(52) Monroe, A., and Setlow, P. (2006) Localization of the transglutaminase cross-linking sites in the *Bacillus subtilis* spore coat protein GerQ. *Journal of bacteriology 188*, 7609−7616.
(53) Ragkousi, K., and Setlow, P. (2004) Transglutaminase-mediated cross-linking of GerQ in the coats of *Bacillus subtilis* spores. *Journal of bacteriology 186*, 5567−5575.
(54) Kuwana, R., Okuda, N., Takamatsu, H., and Watabe, K. (2006) Modification of GerQ reveals a functional relationship between Tgl and YabG in the coat of *Bacillus subtilis* spores. *J. Biochem. 139*, 887−901.
(55) Keillor, J. W., Clouthier, C. M., Apperley, K. Y. P., Akbar, A., and Mulani, A. (2014) Acyl transfer mechanisms of tissue transglutaminase. *Bioorg. Chem. 57*, 186−197.
(56) Wang, J., Xiang, Y. F., and Lim, C. (1994) The double catalytic triad, Cys25-His159-Asp158 and Cys25-His159-Asn175, in papain catalysis: role of Asp158 and Asn175. *Protein Eng., Des. Sel. 7*, 75−82.

DOI: 10.1021/acs.biochem.5b00661
Biochemistry 2015, 54, 5723−5734