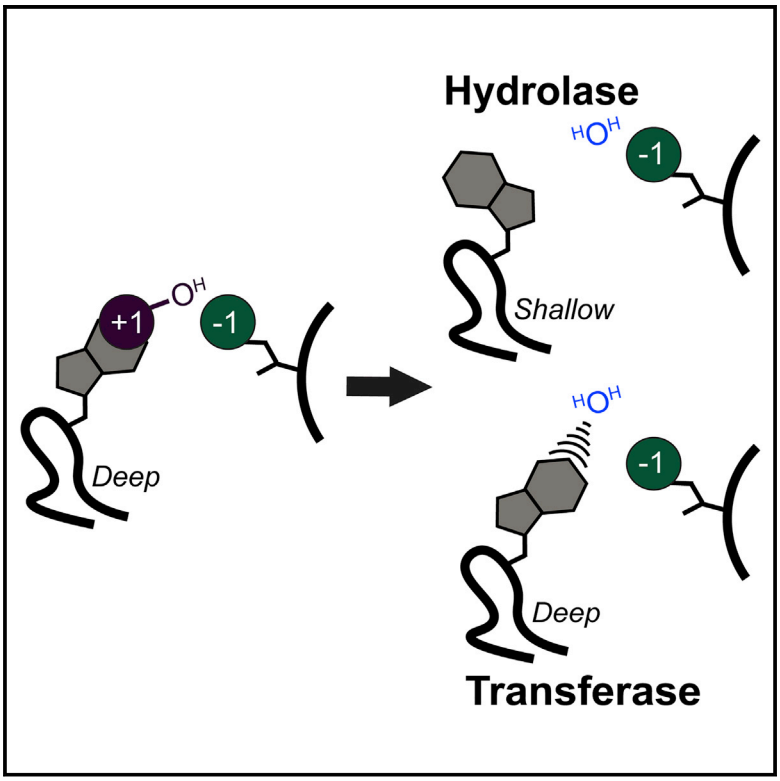
Article

Structure

Transferase Versus Hydrolase: The Role of Conformational Flexibility in Reaction Specificity

Graphical Abstract



Authors

Samuel H. Light, Laty A. Cahoon, Kiran V. Mahasenan, ..., Shahriar Mobashery, Nancy E. Freitag, Wayne F. Anderson

Correspondence

wf-anderson@northwestern.edu

In Brief

Light et al. address the basis of reaction specificity in related enzyme classes. Their findings, that inherent active site flexibility promotes hydrolysis in representative enzymes, have application for the rational design of enzymes with desired transglycosidase activities.

Highlights

- Identification of probable new carbohydrate-binding module family
- Characterization of non-canonical conformational changes that favor hydrolysis
- Description of water-shielding mechanism that favors transferase activity
- Lessons for reengineering enzymes to enhance transglycosylase activity

Accession Numbers

5I0D
5I0E
5I0F
5I0G
5HPO
5HXM

Light et al., 2017, Structure *25*, 295–304
February 7, 2017
http://dx.doi.org/10.1016/j.str.2016.12.007

CellPress

Structure
Article

Transferase Versus Hydrolase: The Role of Conformational Flexibility in Reaction Specificity

Samuel H. Light,[1] Laty A. Cahoon,[2] Kiran V. Mahasenan,[3] Mijoon Lee,[3] Bill Boggess,[3] Andrei S. Halavaty,[1] Shahriar Mobashery,[3] Nancy E. Freitag,[2] and Wayne F. Anderson[1,4,*]
[1]Department of Biochemistry and Molecular Genetics, Center for Structural Genomics of Infectious Diseases, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA
[2]Department of Microbiology and Immunology, University of Illinois at Chicago, Chicago, IL 60612, USA
[3]Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556, USA
[4]Lead Contact
*Correspondence: wf-anderson@northwestern.edu
http://dx.doi.org/10.1016/j.str.2016.12.007

SUMMARY

Active in the aqueous cellular environment where a massive excess of water is perpetually present, enzymes that catalyze the transfer of an electrophile to a non-water nucleophile (transferases) require specific strategies to inhibit mechanistically related hydrolysis reactions. To identify principles that confer transferase versus hydrolase reaction specificity, we exploited two enzymes that use highly similar catalytic apparatuses to catalyze the transglycosylation (a transferase reaction) or hydrolysis of α-1,3-glucan linkages in the cyclic tetrasaccharide cycloalternan (CA). We show that substrate binding to non-catalytic domains and a conformationally stable active site promote CA transglycosylation, whereas a distinct pattern of active site conformational change is associated with CA hydrolysis. These findings defy the classic view of induced-fit conformational change and illustrate a mechanism by which a stable hydrophobic binding site can favor transferase activity and disfavor hydrolysis. Application of these principles could facilitate the rational reengineering of transferases with desired catalytic properties.

INTRODUCTION

One of the most fundamental questions in biochemistry concerns how transferases, the large category of enzymes that catalyze the transfer of a non-water functional group, prevent unwanted hydrolysis (Koshland, 1959). Operating in the aqueous cellular environment where water is present in massive excess over any substrate (55 M water typically compared with nanomolar-millimolar substrate), mechanistically analogous hydrolysis is a persistent threat to overwhelm transferase activity.

In no group of enzymes is this complication more profound than transglycosylases. Not only are these enzymes, which catalyze the transfer of a donor sugar to an acceptor sugar, evolutionarily descendants of glycoside hydrolases, but they also share a common reaction mechanism (Koshland, 1953). In this reaction, a sugar substrate covalently modifies a nucleophilic amino acid (establishing a glycosyl-enzyme intermediate) while the first product is concomitantly released. Next, the glycosyl-enzyme intermediate reacts with a nucleophilic acceptor; with the identity of this acceptor determining whether hydrolysis or transglycosylation results. In glycoside hydrolases, a hydrolytic water acts as acceptor, whereas, in transglycosylases, a hydroxyl group on the second sugar substrate does (Figure 1A).

Considering their glycoside hydrolase origin and conserved reaction mechanism, transglycosylases can be described as hydrolases that have evolved the capacity to: (1) prevent hydrolytic water from binding and (2) promote acceptor sugar binding. Understanding the mechanistic underpinnings of this evolutionary transition would get at the fundamental question of how transferases prevent hydrolysis and have practical implications as well.

Because they catalyze highly specific reactions and use prevalent substrates, transglycosylases have been considered an attractive option for the synthesis of carbohydrates for various research or commercial applications (Cote and Tao, 1990; Edelman, 1956). Unfortunately, the use of transglycosylases for synthetic purposes has been limited by the reality that these enzymes are relatively rare in nature and the fact that the ones that have been characterized act on a limited substrate repertoire. By contrast, hydrolases are extremely common and act on a wide range of substrates. Consequently, the possibility of identifying mutations in hydrolases that shift the balance of hydrolase/transglycosylase activity to favor a desired transglycosylase activity has received considerable research attention. A number of studies have attempted to reengineer hydrolases by introducing mutations in a random or semi-rational manner (reviewed in Bissaro et al., 2015). However, the development of a more predictable reengineering approach, premised on manipulating the fundamental features controlling reaction specificity, is necessary if this strategy is going to achieve a level of practicality prerequisite for widespread use (Bissaro et al., 2015).

We recently showed that the cycloalternan (CA) metabolic pathway is involved in growth of extracellular α-glucans and

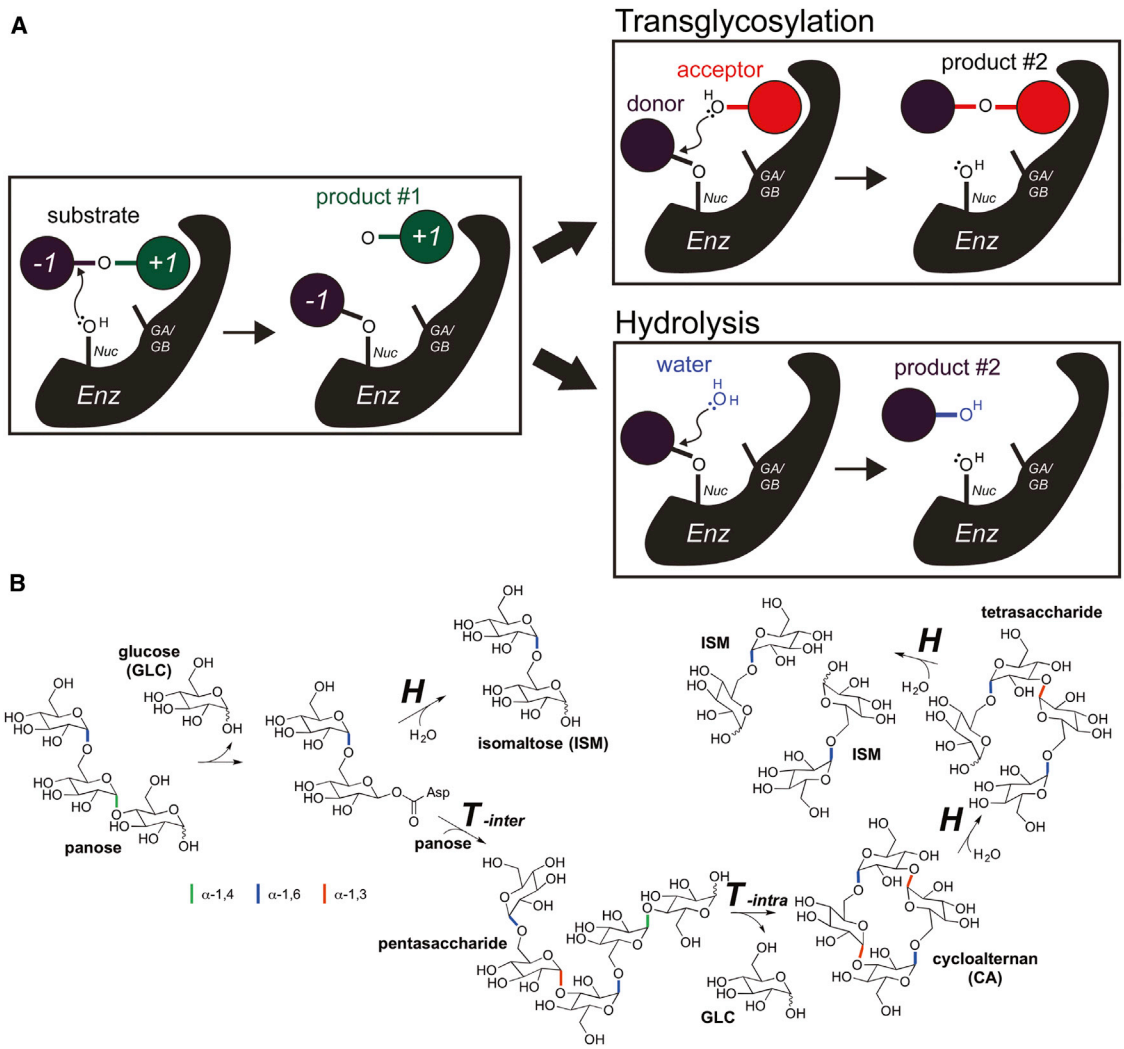


Figure 1. General Transglycosylase/Retaining Glycoside Hydrolase Mechanisms and *Lm*CAFE- and *Tp*CADE-Catalyzed Reactions

(A) Model of key steps in related transglycosylase and retaining glycoside hydrolase reaction mechanisms. Carbohydrate substrates are represented as spheres and the enzyme, with the active site nucleophile (Nuc) and general acid/general base (GA/GB) that transfers protons labeled.

(B) Summary of relevant reactions catalyzed by *Lm*CAFE and *Tp*CADE. Hydrolase (*H*) and intramolecular and intermolecular transglycosylase (*T*) reactions are indicated.

promotes pathogenesis in the foodborne pathogen *Listeria monocytogenes* (Light et al., 2016). In this pathway, secreted enzymes convert environmental α-glucans into a cyclic tetrasaccharide with alternating α-1,3/1,6-glucan linkages called CA. CA is taken up by *L. monocytogenes* and intracellular enzymes catabolize it to glucose. The CA-forming enzyme (CAFE) is an extracellular transglycosylase, that catalyzes the formation of the α-1,3-linkages within CA, while CA-degrading enzyme (CADE), is an intracellular retaining glycoside hydrolase that hydrolyzes these α-1,3-linkages (Figure 1B). We previously determined crystal structures of the *L. monocytogenes* CAFE (*Lm*CAFE) and *Trueperella pyogenes* CADE (*Tp*CADE), which showed that the two enzymes possess nearly identical active sites (Light et al., 2016). TpCADE's shared catalytic scaffold and opposing activities should help eliminate confounding variables and facilitate the identification of features that account for their respective activities. Thus, to clarify the basis of reaction specificity in the two enzyme classes, we initiated a comparative study to identify properties that determine the balance between transglycosylase versus hydrolase activity in the CA enzymes.

RESULTS

Substrate Binding to Non-catalytic Domains Likely Promotes Transglycosylation

*Lm*CAFE and *Tp*CADE contain a core glycoside hydrolase 31 family catalytic domain and four non-catalytic domains that share 31% sequence identity. The most conspicuous difference between the two enzymes is the C-terminal carbohydrate-binding module (CBM) unique to *Lm*CAFE. CBMs are non-catalytic carbohydrate-binding domains contained in enzymes that act

Table 1. Data Collection and Refinement Statistics for *Lm*CAFE

	α-1,4-Glucan	α-1,6-Glucan	CA
PDB:	5HPO	5HXM	5I0D
Data Collection			
Resolution range (Å)	30.00–1.80 (1.83–1.80)[a]	30.00–1.90 (1.93–1.90)	30.00–1.77 (1.80–1.77)
Space group	$C2_1$	$C2_1$	$P2_1$
Unit cell dimensions			
a, b, c (Å)	162.6, 100.1, 73.3	166.2, 102.4, 74.0	74.8, 101.2, 166.4
α, β, γ (°)	90, 105.5, 90	90, 103.8, 90	90, 101.0, 90
Completeness (%)	99.6 (99.2)	99.9 (99.3)	99.7 (100)
No. of reflections	103,966 (5,169)	94,486 (4,690)	233,478 (11,135)
Redundancy	3.8 (3.8)	3.8 (3.7)	3.9 (3.6)
$\langle I/\sigma(I)\rangle$	17.3 (2.7)	20.6 (2.1)	14.0 (2.1)
R_{merge} (%)	7.9 (60.2)	6.6 (60.0)	9.6 (63.8)
Refinement Statistics			
R factor (R_{work}/R_{free}) (%)[b]	13.3/16.4	15.2/19.5	14.4/17.1
No. of atoms			
Protein	8,335	8,343	16,686
Water	1,259	1,025	2,813
Carbohydrate	91	66	413
RMSD			
Bond lengths (Å)	0.009	0.011	0.010
Bond angles (°)	1.55	1.68	1.45
Ramachandran analysis (%)			
Favored regions	97	97	98
Allowed regions	100	100	100
Disallowed regions	0	0	0

RMSD, root-mean-square deviation.

[a]Highest-resolution shell in parentheses.

[b]Definition of R_{work}, R_{free}: $R = \Sigma_{hkl}||F_{obs}| - |F_{calc}|/\Sigma_{hkl}|F_{obs}|$, where *hkl* are the reflection indices used in refinement for R_{work}, and 5% not used in refinement for R_{free}. F_{obs}, and F_{calc} are structure factors deduced from measured intensities or calculated from the model, respectively.

on carbohydrates. A number of CBMs have been shown to boost activity by promoting association with an insoluble substrate, effectively increasing its local concentration (Boraston et al., 2004). We theorized that the CBM domain of *Lm*CAFE might promote transglycosylation in this way.

To determine the binding specificity of *Lm*CAFE's CBM, we separately soaked substrate or product α-1,4- and α-1,6-linked glucans into *Lm*CAFE crystals (Table 1). These studies confirmed that the CBM binds α-1,6-linked glucans (site 1, Figure S1A) and, unexpectedly, revealed three additional non-catalytic binding sites unique to *Lm*CAFE (Figure 2A). Site 2 is positioned in a cleft between the CBM and catalytic domain and recognizes both α-1,4- and α-1,6-linkages (Figure S1B). Sites 3 and 4 are contained on the N-terminal CBM-like domain and specifically recognize α-1,4-linkages (Figures S1C and S1D). One of the sites on the CBM-like domain is also observed to bind α-1,4-glucans in a structurally related domain on the *Thermoactinomyces vulgaris* α-amylase (Abe et al., 2005), strongly implying that this domain represents the founding member of a new CBM family (Figure S1E).

The identified non-catalytic binding sites are distributed across the face of *Lm*CAFE that harbors the active site, and all engage non-reducing oligosaccharide ends. This distribution of non-catalytic binding sites is similar to the distribution of surface binding sites on a number of carbohydrate-active enzymes (Cockburn and Svensson, 2013; Cuyvers et al., 2012). Such a distribution of non-catalytic binding sites can engage a multi-branched particle (such as starch or glycogen) and orient the active site toward exposed branch ends on the particle surface. We reasoned that the presence of non-catalytic binding sites on *Lm*CAFE (a lipoprotein anchored to the cell membrane; Renier et al., 2012) might promote cell adherence to environmental α-glucans. A simple adherence assay was employed to test the ability of an *L. monocytogenes* strain with a disrupted *cafe* gene (*Δcafe*) to bind soluble α-1,3/1,6-linked dextran and insoluble α-1,4/1,6-linked dextrin. Confirming the binding properties of *Lm*CAFE, the mutant exhibited diminished association with both carbohydrates (Figure 2B).

The structural and binding studies suggest that non-catalytic interactions orient the *Lm*CAFE active site toward the surface of branched α-glucan particles. As most α-glucan polymers contain multiple non-reducing chain ends, this non-catalytic association should substantially increase the effective concentration of acceptor sugar substrate. Considering that transglycosylation depends upon acceptor sugar binding outcompeting nucleophilic water, such *Lm*CAFE-specific interactions may

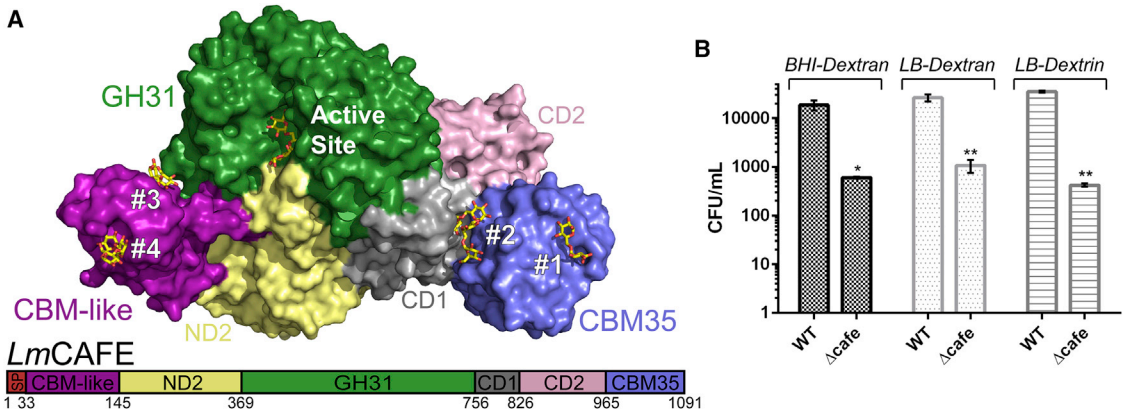


Figure 2. *Lm*CAFE Makes Non-catalytic Interactions that Promote Association with Substrate

(A) Composite of three *LmCAFE* crystal structures shows non-catalytic α-glucan binding sites. Sites 1 and 2 bind α-1,6-linkages and sites 2, 3, and 4 bind α-1,4-linkages. Abbreviations refer to: signal peptide (SP), glycoside hydrolase family 31 catalytic domain (GH31), structural domains (ND2, CD1, and CD2), and carbohydrate binding domains (CBM-like and CBM35).

(B) Wild-type *L. monocytogenes* and a strain with inactivated CAFE (*Δcafe*) were grown to mid-log phase in brain-heart infusion (BHI) broth or Luria broth supplemented with α-glucans before being tested for binding to soluble α-glucan dextran and insoluble dextrin. Representative averages from measurements performed in triplicate are graphed with ±1 SEM error bars. Asterisks signify a statistically significant difference from wild-type (*$p \leq 0.05$ and **$p \leq 0.005$) as assessed by Student's t test.

increase the local concentration of the acceptor sugar substrate and tilt the balance in favor of the intermolecular transglycosylation step in CA synthesis.

The α4-α5 Loop Is Critical for CADE/CAFE Reaction Specificity

We next turned our attention to identifying catalytic features responsible for distinct CAFE and CADE activities. Crystal structures of *Tp*CADE and *Lm*CAFE in complex with CA were determined (Table 2). In both enzymes, CA is sandwiched between D467 and D532 (the amino acid numbering throughout the manuscript corresponds to the residue position in *Tp*CADE), which, based on alignment with homologous carbohydrate-active enzymes, are predicted to function as nucleophile and general acid/general base in catalysis, respectively (Figures 3A and 3B). CA makes very similar interactions with each active site and adopts a nearly identical asymmetric conformation, which differs from previously observed CA conformational states (Figure S2) (Bradbrook et al., 2000; Light et al., 2016). The major feature that distinguishes CA binding to *Lm*CAFE and *Tp*CADE concerns conformational changes in the two enzymes. In *Tp*CADE, a 26-residue insertion between the N-terminal β3 and β4 strands (β3-β4 insertion) of the CBM-like domain is disordered in the unliganded structure, but well resolved in the CA complex. The β3-β4 insertion adopts a trihelical structure and contributes H78 to the active site, which hydrogen bonds with the 6-hydroxyl on the +1 glucose of CA (Figure 3C). In addition to providing H78 to the active site, the β3-β4 insertion abuts the α4-α5 loop and seems to facilitate its shift deeper into the active site. In this deep α4-α5 loop conformation, W430 stacks with the +1 glucose in CA and W429 hydrogen bonds with this sugar (Figure 3D). By contrast, *Lm*CAFE lacks the β3-β4 insertion and the α4-α5 loop assumes a deep conformational state even in the absence of CA (Figure 3E).

Additional crystallographic studies were performed to clarify the functional significance of the observed conformational changes. By soaking CA into wild-type *Tp*CADE crystals, a post-breakage/pre-release glycosyl-enzyme intermediate state was captured (Table 2). In this structure both α-1,3-linkages in CA have been hydrolyzed and the product α-1,6-linked disaccharides remain bound at the active site. One disaccharide is covalently linked to the nucleophilic D467 and the other is positioned non-covalently beside it (Figure 4A). The β3-β4 insertion remains ordered, with H78 continuing to contribute to the active site, but adopts a distinct conformation from the CA complex (Figure S3A). The capture of the post-breakage/pre-release intermediate state likely resulted from the presence of two calcium ions (present in the crystallization condition), which bridge crystal packing contacts made by the β3-β4 insertion and, presumably, stabilize the observed conformation (Figure S3B). By soaking the product disaccharide isomaltose into *Tp*CADE crystals, we also captured a non-covalent post-breakage/post-release complex. In this structure one disaccharide is non-covalently bound at the active site, the β3-β4 insertion is disordered, and the α4-α5 loop assumes its shallow conformation (Figure 4B).

A comparison of the four captured *Tp*CADE reaction states demonstrates that occupancy of the +1 binding site of the substrate represents the key determinant of β3-β4 insertion/α4-α5 loop conformation (Figure 4C). As this observation times the conformational change to just prior to hydrolysis and represents the only identifiable difference from *Lm*CAFE, we hypothesized it played a key role in conferring the hydrolase activity of *Tp*CADE.

To probe the role of the β3-β4 insertion in hydrolysis, we generated an H78A mutant that eliminated the CA-hydrogen bonding side chain from the β3-β4 insertion, and a deletion mutant (Δβ3-β4) that removed residues A69-L87 from the insertion. Hydrolytic activity on the α-1,3-linkages of CA was measured using a coupled glucose oxidase assay and revealed

Table 2. Data Collection and Refinement Statistics for *Tp*CADE

	CA	Intermediate	Isomaltose
PDB:	5I0G	5I0F	5I0E
Data Collection			
Resolution range (Å)	30.00–2.15 (2.19–2.15)[a]	30.00–1.85 (1.88–1.85)	30.00–2.30(2.34–2.30)
Space group	$C2_1$	$C2_1$	$C2_1$
Unit cell dimensions			
a, b, c (Å)	196.4, 103.9, 44.8	195.13, 103.13, 44.09	194.63, 103.4, 44.09
α, β, γ (°)	90, 90.3, 90	90, 92.1, 90	90, 91.3, 90
Completeness (%)	99.7 (100)	100 (100)	98.6 (98.3)
No. of reflections	48,835 (2,376)	74,285 (3,681)	37,721 (1,885)
Redundancy	3.8 (3.4)	3.7 (3.6)	3.3 (2.8)
$\langle I/\sigma(I)\rangle$	10.9 (1.9)	17.3 (2.0)	11.9 (1.5)
R_{merge} (%)	9.4 (52.8)	8.9 (72.1)	7.4 (57.4)
Refinement Statistics			
R factor (R_{work}/R_{free})[b]	17.2/21.0	15.9/19.0	17.4/22.1
No. of atoms			
Protein	5,789	5,783	5,614
Waters	383	609	281
Carbohydrate	44	45	23
RMSD			
Bond lengths (Å)	0.009	0.010	0.010
Bond angles (°)	1.37	1.43	1.40
Ramachandran analysis (%)			
Favored regions	98	99	98
Allowed regions	100	100	100
Disallowed regions	0	0	0

RMSD, root-mean-square deviation.

[a]Highest-resolution shell in parentheses.

[b]Definition of R_{work}, R_{free}: $R = \Sigma_{hkl}||F_{obs}| - |F_{calc}||/\Sigma_{hkl}|F_{obs}|$, where *hkl* are the reflection indices used in refinement for R_{work}, and 5% not used in refinement for R_{free}. F_{obs}, and F_{calc} are structure factors deduced from measured intensities or calculated from the model, respectively.

that both variants had an ~4-fold increase in K_m, with the Δβ3-β4 mutation also causing a ~20-fold reduction in reaction rate (Table 3). These results confirm the involvement of H78 in substrate binding, but demonstrate that the interaction is dispensable for catalysis. The effect of the Δβ3-β4 mutation on reaction rate suggests that interactions with other active site residues (most likely those in the α4-α5 loop) represent the more significant aspect of β3-β4 insertion function. Nevertheless, the rather modest effect of the mutations, coupled with an analysis of putative CADE sequences that revealed that the β3-β4 insertion is not conserved (Figure S4), led us to conclude that the insertion is not a critical determinant of reaction specificity.

We next turned our attention to the role of the α4-α5 loop. A chimeric protein (*Lm*CAFE/CADE-*chimera*) that swapped 93 amino acids in *Lm*CAFE for corresponding residues in the *L. monocytogenes* CADE (replacing the α4-α5 loop and structurally flanking regions) was generated. The relative hydrolase versus transglycosylase activity of *Lm*CAFE/CADE-*chimera* was measured using a liquid chromatography-mass spectrometry (LC-MS)-based assay. Consistent with the α4-α5 loop playing a key role in conferring reaction specificity, *Lm*CAFE/CADE-*chimera* exhibited near exclusive hydrolase activity (Figure 5A, Table 4).

W430 Mobility Can Explain CADE/CAFE Reaction Specificity

Having identified the α4-α5 loop as critical, we next sought to understand the mechanism by which it influences reaction specificity. AMBER14 (Case et al., 2014; Gotz et al., 2012; Salomon-Ferrer et al., 2013) was used to perform 220 ns molecular dynamics (MD) simulations on *Tp*CADE and *Lm*CAFE structures. By tracking the motion of the loop residue W430, a broad range of α4-α5 loop structural fluctuation is evident for unliganded *Tp*CADE, but not for post-breakage/pre-release intermediate state or unliganded *Lm*CAFE (Figure 5B). This pattern of mobility is also suggested by an analysis of electron density (Figure 5C) and temperature factors in the crystal structures (Table S1). Thus, the different *Tp*CADE and *Lm*CAFE α4-α5 loop conformations appear to be associated with distinctions in inherent mobility.

In the equilibrium between transglycosylation and hydrolysis, hydrolysis will be favored when the binding of nucleophilic water at the optimal site is preferred over the binding of acceptor sugar. In this important balance between water and acceptor sugar binding, increased mobility of the α4-α5 loop is poised to affect both sides of the equation in favor of hydrolysis. First,

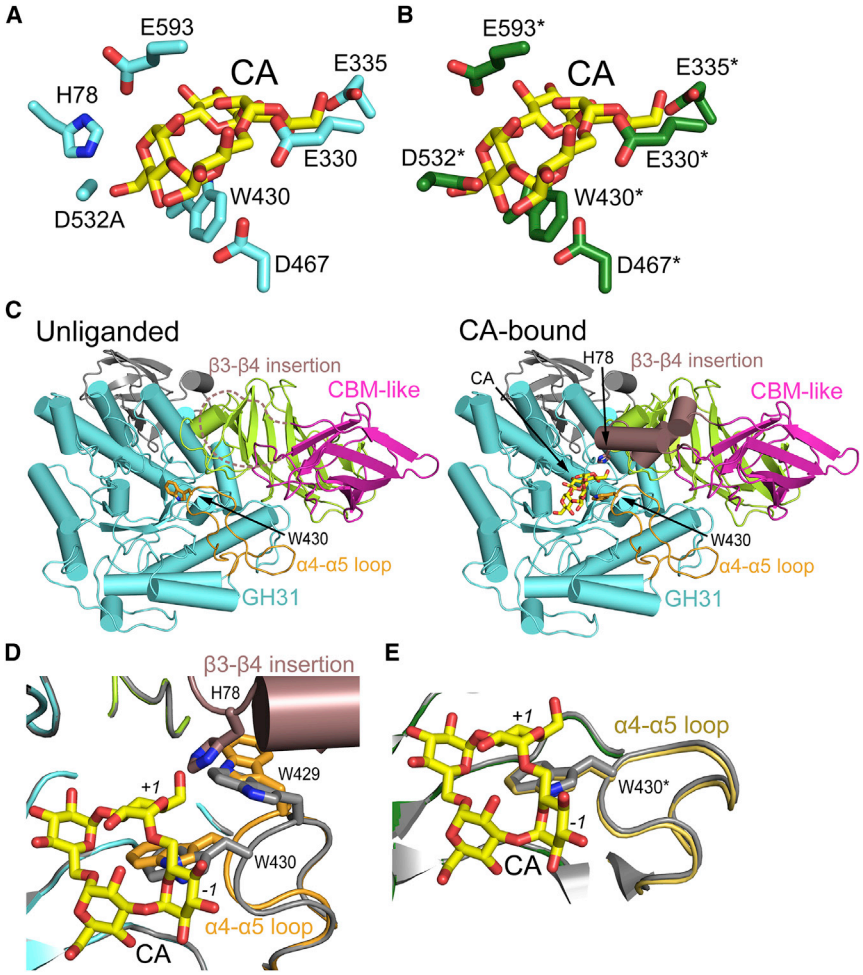


Figure 3. CA Binding Induces Unique Conformational Change in *Tp*CADE Relative to *Lm*CAFE

(A) CA bound to the *Tp*CADE active site. A mutation that replaced the *Tp*CADE general acid/general base aspartate with an alanine (D532A) facilitated capture of this substrate bound state.

(B) CA bound to the *Lm*CAFE active site. For the sake of consistency, residues are labeled (*) based on corresponding positions in *Tp*CADE.

(C) In the unliganded *Tp*CADE structure the β3-β4 insertion is disordered (dashed lines) and the α4-α5 loop adopts a shallow conformation. In the CA-bound structure the insertion is ordered, allowing H78 to hydrogen bond with the +1 glucose of CA, and the α4-α5 loop adopts a deep conformation, where W430 establishes a CH-π-stacking interaction with the +1 CA glucose.

(D) Superposition of unliganded (gray) and CA-bound (colored) *Tp*CADE structures highlights the α4-α5 loop shift from shallow to deep associated with CA binding.

(E) Superposition of unliganded (gray) and CA-bound (colored) *Lm*CAFE structures highlights retention of the deep conformation of the α4-α5 loop.

in *Lm*CAFE, W430* on the α4-α5 loop establishes a critical CH-π-stacking interaction with the +1 acceptor sugar; in *Tp*CADE, conformational change of the α4-α5 loop away from the glycosyl-enzyme linkage takes W430 out of position to stack with the sugar and orient it for nucleophilic attack (Figure 5D). Second, in *Lm*CAFE, the deep conformation of the α4-α5 loop places W430* in close proximity of the glycosyl-enzyme linkage. In this position W430* is poised to act as a hydrophobic shield to sterically disfavor the optimal approach of nucleophilic water (Figure 5E). In *Tp*CADE, the transition of the α4-α5 loop to its shallow conformation takes W430 out of its water-shielding position and, thus, should remove this impediment to hydrolysis.

Since this analysis pinpointed the rigid positioning of the W430* side chain as the key determinant of transglycosylase activity, we generated a *Lm*CAFE W430* to alanine (W430A*) mutant and used the LC-MS assay to measure its relative hydrolase versus transglycosylase activity. The W430A* mutation resulted in >1000× reduction in activity in the assayed conditions, probably in large part due to reduced donor sugar binding affinity. Nevertheless, confirming that the position of the W430 side chain serves a key role in conferring transglycosylase activity; the W430A* variant exhibited near exclusive hydrolase activity (Table 4).

DISCUSSION

We set out to address how transferase/hydrolase specificity is achieved by the virtually identical catalytic apparatuses in *Lm*CAFE and *Tp*CADE. Structural and biochemical studies established that non-catalytic binding sites are situated in such a way as to increase the effective substrate concentration, likely promoting transglycosylation, and that the β3-β4 insertion contributes to hydrolysis. Most importantly, our results suggest that by rigidly retaining a deep conformation, the α4-α5 loop facilitates acceptor sugar binding and sterically disfavors hydrolysis, thereby promoting transglycosylation. By contrast, a dynamic α4-α5 loop seems to promote hydrolysis by allowing for rearrangement to a shallow conformation, which simultaneously withdraws the acceptor sugar binding site and removes the steric impediment to hydrolysis. Based on these findings a model of the specificity-conferring mechanism in *Tp*CADE and *Lm*CAFE is proposed (Figure 6).

Questions pertaining to the basis of transferase/hydrolase reaction specificity have a long history. Consideration of what prevented hexokinase, an enzyme that phosphorylates sugar substrates, from simply hydrolyzing ATP (i.e., phosphorylating water) led Daniel Koshland to propose his famous induced-fit theory of enzyme function (Koshland, 1958, 1959, 1994). His argument, which was later experimentally borne out (Bennett and Steitz, 1978), stated that, in the absence of sugar substrate, hexokinase must adopt an inactive conformation to prevent ATP hydrolysis by ubiquitous water and, thus, the act of substrate binding must trigger activating conformational changes in the enzyme.

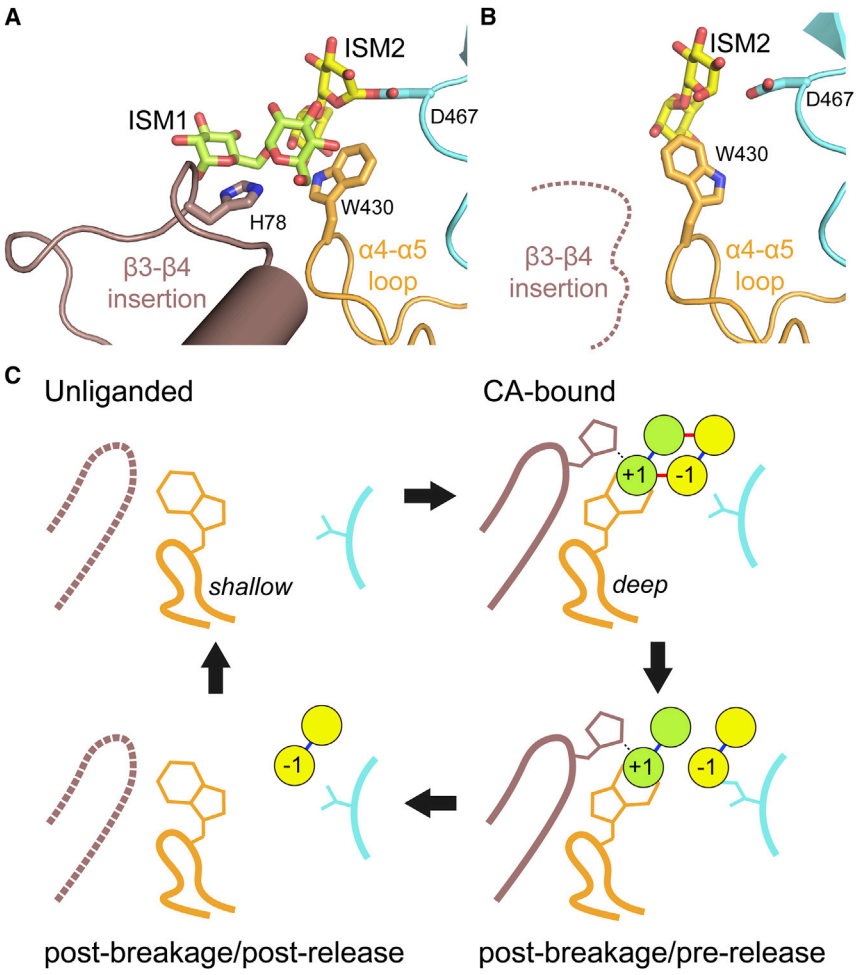


Figure 4. Conformational Change in *Tp*CADE Correlates with Occupancy of the +1 Binding Site

(A) A post-breakage/pre-release covalent intermediate complex was generated by soaking CA into *Tp*CADE crystals. Depicted from a different perspective than in Figure 3, both of the α-1,3-linkages of CA have been hydrolyzed and the two product isomaltose (ISM) disaccharides remain at the active site; ISM1 is non-covalently bound, whereas ISM2 is covalently linked to the nucleophilic D467. The ordered β3-β4 insertion and the deep α4-α5 loop conformation resemble the CA complex.

(B) A post-breakage/post-release complex was generated by soaking isomaltose in *Tp*CADE crystals. In this structure only ISM2 is non-covalently bound and the disordered β3-β4 insertion and the shallow α4-α5 loop conformation resemble the unliganded structure.

(C) Schematic representation of the four captured *Tp*CADE states. Glucoses are represented as spheres and covalent connections as lines. The β3-β4 insertion and the α4-α5 loop both interact with the glucose in the +1 binding site, and a comparison of the different states highlights how their conformation correlates with the occupancy of the +1 site.

Interestingly, *Tp*CADE and *Lm*CAFE appear to reverse Koshland's induced-fit logic. In *Lm*CAFE, the default conformation sterically hinders hydrolysis and, consequently, conformational change is not necessary to prevent this unwanted activity. By contrast, in *Tp*CADE, these same factors necessitate classic induced-fit conformational changes to create a site conducive for the hydrolytic water. Characterized by a default transferase conformation that requires induced-fit conformational change for hydrolysis, the enzymes exemplify the precise opposite mechanism of that stipulated in the classic paradigm. It is thus clear that the role of induced fit in conferring transferase/hydrolysis reaction specificity is more nuanced than has been traditionally appreciated and, depending upon context, can function to either favor or disfavor hydrolysis.

Table 3. Kinetic Characterization of *Tp*CADE Variants

	V_{max} (nM/s)	K_m (mM)
Wild-type	30.5 ± 0.9	2.1 ± 0.2
H78A	30.8 ± 2.1	9.3 ± 1.5
Δβ3-β4	1.5 ± 0.1	7.7 ± 1.1

Errors calculated by fitting of the kinetic data to the Michaelis-Menten equation and expressed as ± SE.

In addition to reframing the question of transferase/hydrolase specificity, the mechanistic insights gained from comparing *Tp*CADE and *Lm*CAFE also have practical implications. In particular, the characterization of a transferase mechanism devoid of induced-fit conformational change is relevant for catalytic reengineering research. Research in this field has sought to generate mutant glycoside hydrolases with synthetically desirable transglycosylation products (reviewed in Bissaro et al., 2015). The identification of principles that favored the evolution of *Lm*CAFE presents new opportunities to guide such design efforts. Specifically, these findings suggest a mutagenesis strategy driven by the goal of introducing an appropriately orientated hydrophobic residue to favor sugar binding at the +1 acceptor site and inhibit hydrolysis. Further study is required to determine best practices and to ascertain how broadly applicable this rational reengineering approach could be.

EXPERIMENTAL PROCEDURES

L. monocytogenes Cell-Adherence Assays

Previously described *L. monocytogenes lmo2446* insertion mutant (*Δcafe*) and control strains were grown in brain-heart infusion broth or Luria broth supplemented with maltodextrin to mid-log phase (OD_{600} = 0.6) (Light et al., 2016). Then 2 × 10^8 bacteria were pelleted and washed with PBS and re-suspended with prewashed dextrin or a mixture of biotinylated dextran (Sigma) + SoftLink Soft Release Avadin Resin (Promega) in PBS. Dextran-bound cells were eluted by the addition of PBS and 5 mM biotin. The dextran elutant and the heterogeneous dextrin mixture were plated and binding was interpreted from the number of resulting colony-forming units.

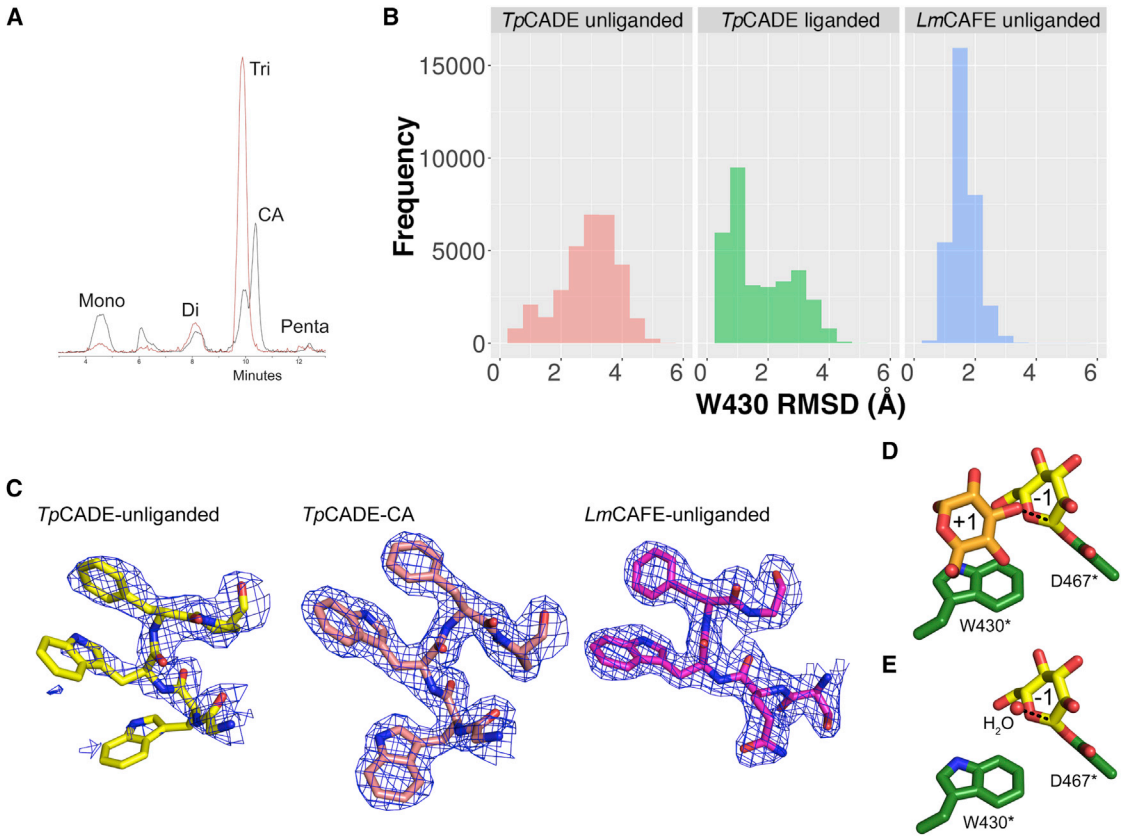


Figure 5. Role of α4-α5 Loop Mobility in Transglycosylase/Hydrolase Reaction Specificity

(A) A representative chromatogram from the LC-MS-based assay showing results from *Lm*CAFE (black) and *Lm*CAFE W528A (red). The oligosaccharide length corresponding to each peak is noted. As summarized in Figure 1, transglycosylation of the substrate panose (tri) produces glucose (mono), pentasaccharide, and CA, while hydrolysis produces glucose (mono) and isomaltose (di).

(B) Histogram plot of root-mean-square deviation (RMSD) fluctuation of the α4-α5 loop residue W430/W430* over 220 ns molecular dynamics simulations.

(C) Electron density ($2F_o - F_c$ map contoured at 1σ) for the α4-α5 loop in identified crystal structures.

(D) Model of the *Lm*CAFE active site with an acceptor sugar in the +1 site and covalently bound donor in the −1 site. The dashed line shows the trajectory of nucleophilic attack. The deep α4-α5 loop conformation positions W430* to interact with and orient the acceptor substrate.

(E) Model of the *Lm*CAFE active site with hydrolytic water and a covalently bound donor in the −1 site. For nucleophilic attack to occur, the static W430* conformation necessitates that the water must travel within an O-C distance of <4 Å of W430*. The dashed line shows the trajectory of nucleophilic attack.

Protein Expression and Purification

Previously described *Lm*CAFE and *Tp*CADE constructs were used for recombinant protein expression (Light et al., 2016). *Lm*CAFE/CADE-*chimera* was generated by digesting the *Lm*CAFE clone with SspI restriction enzyme (New England BioLabs) and introducing a synthetic portion of *Lm*CADE (Integrated DNA Technologies) by Gibson assembly (Gibson et al., 2009), which swapped α4-α5 loop-encompassing *Lm*CAFE residues 486–578 with *Lm*CADE residues 675–768. The *Tp*CADE H78A, *Tp*CADE D532A, *Tp*CADE Δβ3-β4, and *Lm*CAFE W430A* variants were generated using the QuikChange Lightning Kit (Agilent Technologies). Recombinant proteins were expressed and purified in BL21 *Escherichia coli* as described previously (Light et al., 2016).

Table 4. Hydrolase Activity of *Lm*CAFE Variants

Wild-Type	17[a,b]
*Lm*CAFE/CADE-*chimera*[c]	97
W430A*[c]	97

[a]Results represent average of two replicates.
[b]Expressed as percent hydrolase product relative to total products.
[c]To compensate for lower baseline activity reactions were carried out for 10× longer and 100× higher enzyme concentrations.

Crystal Structure Determination

The sitting-drop vapor-diffusion method was used for crystal growth, at a 1:1 ratio of protein to reservoir. *Lm*CAFE was concentrated to 7.3 mg/mL in 500 mM NaCl and 10 mM Tris (pH 8.3) and crystallized in a condition containing 200 mM magnesium formate and 25% (w/v) poly-ethylene glycol (PEG) 3350. To generate the α-1,4-glucan complex, a crystal was incubated in a mother liquor supplemented with 10 mM maltopentaose (Sigma-Aldrich) for ~5 min at ambient temperature. To generate the α-1,6-glucan complex, a crystal was incubated in a mother liquor supplemented with 5 mM panose (Sigma-Aldrich) for ~5 min at ambient temperature. To generate the CA complex, an *Lm*CAFE crystal was incubated in a mother liquor supplemented with 100 mM CA for ~20 min at ambient temperature.

The *Tp*CADE CA complex was obtained after inactivating the enzyme by mutating the putative general acid/general base, D532, to alanine. *Tp*CADE D532A was concentrated to 5.3 mg/mL in 3 mM CA, 250 mM NaCl, and 10 mM Tris (pH 8.3) and crystallized in the JCSG+ Suite (QIAGEN) H7 condition, which contained 200 mM ammonium sulfate, 100 mM Bis-Tris (pH 5.5), and 25% (w/v) PEG 3350. Prior to freezing, the crystal was transferred to a mother liquor supplemented with 10 mM CA. For the *Tp*CADE covalent

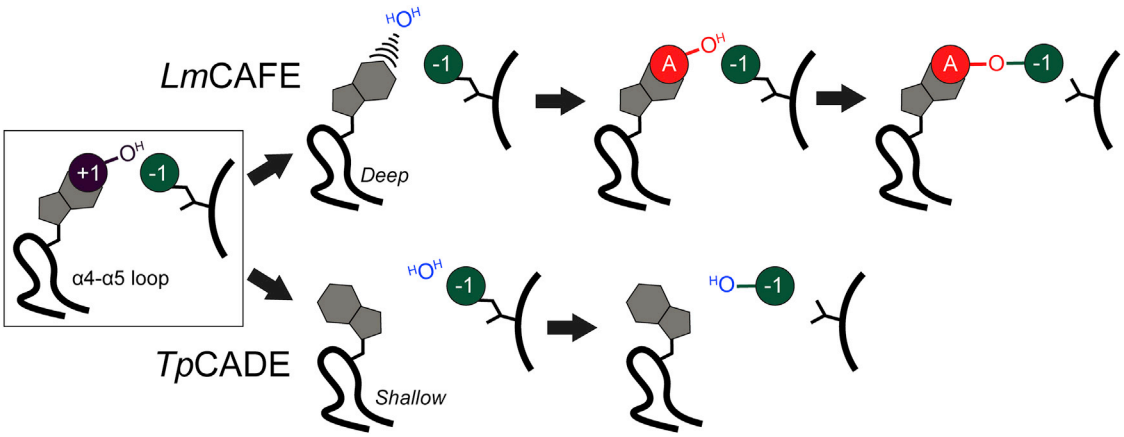


Figure 6. Model of Mechanisms Controlling Reaction Specificity in *Lm*CAFE and *Tp*CADE

The preliminary steps are the same in both reactions. The deep α4-α5 loop conformation promotes substrate binding by positioning W430 to favorably interact with the +1 moiety. The substrate reacts, causing breakage of the +1 to −1 glycoside bond and a glycosyl-enzyme covalent intermediate to form (center box). After product is released from the +1 site, the α4-α5 loop retains its deep conformation in *Lm*CAFE but shifts to a shallow conformation in *Tp*CADE. In *Lm*CAFE, the deep loop conformation positions W430* to provide a surface for the +1 acceptor sugar to bind and sterically disfavor water binding. In *Tp*CADE, the shifted conformation of the loop takes W430 out of the would-be +1 acceptor binding site and removes the steric impediment to hydrolysis. The β3-β4 insertion in *Tp*CADE promotes substrate binding and α4-α5 loop conformational change. For the sake of simplicity, throughout the figure only the reacting +1 and −1 sugars are depicted.

intermediate and isomaltose complexes, crystallization experiments were initiated with 8.0 mg/mL *Tp*CADE in 500 mM NaCl and 10 mM Tris (pH 8.3). Crystals grew in the absence of sugar in the PACT Suite (QIAGEN) B11 condition, which contained 200 mM calcium chloride, 100 mM 2-(N-morpholino)ethanesulfonic acid (pH 6), and 20% (w/v) PEG 6000. To generate the post-breakage/pre-release covalent intermediate complex, a crystal was incubated in a mother liquor supplemented with 25 mM CA for ~1 hr at ambient temperature. To generate the post-breakage/post-release complex a crystal was incubated in a mother liquor supplemented with 25 mM isomaltose for ~5 min at ambient temperature. All crystals were frozen on liquid nitrogen for data collection immediately after incubations.

Diffraction data were collected at 100 K at the Life Sciences-Collaborative Access Team at the Advanced Photon Source, Argonne National Laboratory. Diffraction data were indexed, integrated, and scaled using the HKL3000 program (Otwinowski and Minor, 1997). All structures were phased by molecular replacement in Phaser, using unliganded structures (PDB: 4KMQ and 5F7S) as search models (McCoy et al., 2005). Model building and structure refinement was performed with the Coot and Refmac programs (Emsley and Cowtan, 2004; Murshudov et al., 1997). TLS groups defined using the TLSMD webserver (http://skuld.bmsc.washington.edu/~tlsmd/) were used for refinement (Painter and Merritt, 2006). All structure figures were prepared using the PyMOL Molecular Graphics System, version 1.7.4 (Schrödinger).

*Lm*CAFE and *Tp*CADE Activity Assays

To determine reaction kinetics (Table 3), *Lm*CAFE and *Tp*CADE variants were assayed by measuring glucose liberated from the trisaccharide panose or from CA using the enzyme-coupled glucose oxidase method, as described previously (Light et al., 2016). To measure relative hydrolase/transglycosylase activities (Table 4), *Lm*CAFE was incubated with 20 mM panose and the reaction was allowed to proceed 2 hr before enzymatic heat inactivation. *Lm*CAFE/CADE-*chimera* and *Lm*CAFE W430A exhibited significantly reduced activity and so enzyme concentrations were increased by 100-fold and reaction times by 10-fold. Reaction products were separated by liquid chromatography on an ACQUITY UPLC BEH Amide column (130 Å, 1.7 μm, 2.1 mm × 150 mm, Waters) and their chemical structures were confirmed by comparison of high-resolution mass spectra as well as retention times on liquid chromatography with those of reference controls. The mobile-phase gradient consisted of elution at 0.3 mL/min with a 3-min isocratic, 30% A/70% B, followed by a 10-min linear gradient to 60% A/40% B, a 3-min isocratic, 60% A/40% B, a 1-min linear gradient to 30% A/70% B, and then 30% A/70% B for 6 min (A, water with 0.1% formic acid; B, acetonitrile with 0.1% formic acid). Peak areas from extracted ion chromatograms of corresponding *m/z* values of each product were integrated. Relative hydrolase activity was calculated as the percent hydrolysis product (disaccharide) relative to total reaction products (hydrolase [disaccharide] + transglycosylation [pentasaccharide and CA] products).

MD Simulations

The protein coordinates were prepared with the Maestro program (v10.0, Schrödinger). During the process, bond orders were assigned and hydrogen atoms were added. Further, proper protonation states were assigned for residues using the PROPKA 3.1 program (Olsson et al., 2011). Crystallographic water molecules were retained in the modeling. MD simulations were performed with AMBER14 suite (Case et al., 2014; Gotz et al., 2012; Salomon-Ferrer et al., 2013). Amber ff12SB and GAFF forcefields provided the necessary molecular mechanics parameters. Ligand charges were derived quantum mechanically using RESP methodology at 6-31G* level of theory (Bayly et al., 1993). Solute coordinates were solvated with a truncated octahedron box of the TIP3P water model in such a way that the box edges were at a minimum distance of 10 Å away from the protein surface. The system was neutralized by adding Na^+ ions to the system. Minimization was performed for the solvated water molecule of the complex for a maximum of 2,000 steps and, in a subsequent stage, the whole system was minimized for another 2,000 steps. The system was heated to 300 K and equilibration of the water molecules was performed for 400 ps. In the subsequent equilibration stage of 400 ps, the protein side chains and water molecules were allowed to freely move while positional restraint with a force constant of 100 kcal/mol was applied to the backbone atoms (N, Cα, and C). The restraints were reduced in the succeeding two equilibration stages of 400 ps. Production simulations for 220 ns under NPT condition was performed with a CUDA GPU accelerated PMEMD module implemented in the AMBER14 suite. Bonds to hydrogen atoms were constrained with SHAKE and a 2 fs time step was used in the MD simulation. Berendsen barostat and thermostats were employed for pressure and temperature control. Post-simulation analysis was performed with the cpptraj module implemented in AMBER14 and VMD program (Humphrey et al., 1996; Roe and Cheatham, 2013). Plots were generated with the ggplot2 module in R.

ACCESSION NUMBERS

Atomic coordinates and structure factors have been deposited in the PDB under accession codes PDB: 5I0D, 5I0E, 5I0F, 5I0G, 5HPO, and 5HXM.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.str.2016.12.007.

AUTHOR CONTRIBUTIONS

S.H.L., A.S.H., and W.F.A. were responsible for the molecular biology, as well as biochemical and crystallographic studies. L.A.C. and N.E.F. conducted in vitro pull-down assays. S.H.L., M.L., B.B., and S.M. performed mass spectrometric experiments and analysis. K.V.M. and S.M. provided molecular dynamic characterizations. S.H.L. and W.F.A. wrote the manuscript.

ACKNOWLEDGMENTS

The Center for Structural Genomics of Infectious Diseases has been funded with federal funds from the National Institute of Allergy and Infectious Diseases, NIH, Department of Health and Human Services, under contract nos. HHSN272200700058C and HHSN272201200026C (to W.F.A.). This work was supported by NIH grants R01 AI08324 + AI083241-03S1 (to N.E.F.), F32 AI 115954 (to L.A.C.), and GM61629 (to S.M.). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor under a research program grant (085P1000817). Thanks to George Minasov for experimental assistance and Gregory Côté for providing cycloalternan and helpful discussions.

Received: September 14, 2016
Revised: November 7, 2016
Accepted: December 12, 2016
Published: January 12, 2017

REFERENCES

Abe, A., Yoshida, H., Tonozuka, T., Sakano, Y., and Kamitori, S. (2005). Complexes of *Thermoactinomyces vulgaris* R-47 alpha-amylase 1 and pullulan model oligossacharides provide new insight into the mechanism for recognizing substrates with alpha-(1,6) glycosidic linkages. FEBS J. *272*, 6145–6153.

Bayly, C.I., Cieplak, P., Cornell, W.D., and Kollman, P.A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges – the RESP model. J. Phys. Chem. *97*, 10269–10280.

Bennett, W.S., Jr., and Steitz, T.A. (1978). Glucose-induced conformational change in yeast hexokinase. Proc. Natl. Acad. Sci. USA *75*, 4848–4852.

Bissaro, B., Monsan, P., Faure, R., and O'Donohue, M.J. (2015). Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. Biochem. J. *467*, 17–35.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. *382*, 769–781.

Bradbrook, G.M., Gessler, K., Cote, G.L., Momany, F., Biely, P., Bordet, P., Perez, S., and Imberty, A. (2000). X-ray structure determination and modeling of the cyclic tetrasaccharide cyclo. Carbohydr. Res. *329*, 655–665.

Case, D.A., Babin, V., Berryman, J., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, T.E., III, Darden, T.A., Duke, R.E., Gohlke, H., et al. (2014). AMBER14 (University of California, San Francisco).

Cockburn, D., and Svensson, B. (2013). Surface binding sites in carbohydrate active enzymes: an emerging picture of structural and functional diversity. In Carbohydrate Chemistry, *39* (Royal Society of Chemistry), pp. 204–221.

Cote, G.L., and Tao, B.Y. (1990). Oligosaccharide synthesis by enzymatic transglycosylation. Glycoconj. J. *7*, 145–162.

Cuyvers, S., Dornez, E., Delcour, J.A., and Courtin, C.M. (2012). Occurrence and functional significance of secondary carbohydrate binding sites in glycoside hydrolases. Crit. Rev. Biotechnol. *32*, 93–107.

Edelman, J. (1956). The formation of oligosaccharides by enzymic transglycosylation. Adv. Enzymol. Relat. Sub. Biochem. *17*, 189–232.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gotz, A.W., Williamson, M.J., Xu, D., Poole, D., Le Grand, S., and Walker, R.C. (2012). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. J. Chem. Theor. Comput. *8*, 1542–1555.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38, 27–38.

Koshland, D.E. (1953). Stereochemistry and the mechanism of enzymatic reactions. Biol. Rev. *28*, 416–436.

Koshland, D.E. (1958). Application of a theory of enzyme specificity to protein synthesis. Proc. Natl. Acad. Sci. USA *44*, 98–104.

Koshland, D.E., Jr. (1959). Enzyme flexibility and enzyme action. J. Cell Comp. Physiol. *54*, 245–258.

Koshland, D.E. (1994). The key-lock theory and the induced fit theory. Angew. Chem. Int. Ed. Engl. *33*, 2375–2378.

Light, S.H., Cahoon, L.A., Halavaty, A.S., Freitag, N.E., and Anderson, W.F. (2016). Structure to function of an alpha-glucan metabolic pathway that promotes *Listeria monocytogenes* pathogenesis. Nat. Microbiol. *2*, 16202.

McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr. D Biol. Crystallogr. *61*, 458–464.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Olsson, M.H., Sondergaard, C.R., Rostkowski, M., and Jensen, J.H. (2011). PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. J. Chem. Theor. Comput. *7*, 525–537.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol. *276*, 307–326.

Painter, J., and Merritt, E.A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr. D Biol. Crystallogr. *62*, 439–450.

Renier, S., Micheau, P., Talon, R., Hebraud, M., and Desvaux, M. (2012). Subcellular localization of extracytoplasmic proteins in monoderm bacteria: rational secretomics-based strategy for genomic and proteomic analyses. PLoS One *7*, e42982.

Roe, D.R., and Cheatham, T.E., 3rd (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J. Chem. Theor. Comput. *9*, 3084–3095.

Salomon-Ferrer, R., Gotz, A.W., Poole, D., Le Grand, S., and Walker, R.C. (2013). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. J. Chem. Theor. Comput. *9*, 3878–3888.