

Expanding the Thioglycoligase Strategy to the Synthesis of α -Linked Thioglycosides Allows Structural Investigation of the Parent Enzyme/ Substrate Complex

Young-Wan Kim,[†] Andrew L. Lovering,[‡] Hongming Chen,[†] Terrence Kantner,^{†,‡} Lawrence P. McIntosh,[‡] Natalie C. J. Strynadka,[‡] and Stephen G. Withers^{*,†}

Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, British Columbia V6T 1Z1, Canada, and Department of Biochemistry and Molecular Biology, University of British Columbia, 2146 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3, Canada

Received November 21, 2005; E-mail: withers@chem.ubc.ca

The α -retaining glycosidases of family 31 include pharmaceutically important human enzymes, such as the human lysosomal α -glucosidase (GAA), endoplasmic reticulum glucosidase II, and sucrase isomaltase (SUIS). The first structural analysis of a family 31 enzyme was that on the α -xylosidase (EC 3.2.1.) from *E. coli* (YicI), thereby providing a useful structural framework for the understanding of the enzymes from higher order organisms.¹ However, delineation of the substrate specificities and mechanism requires structural information on the enzyme–substrate complex. Nonhydrolyzable substrate analogues, such as thioglycosides, are useful probes for such studies due to their resistance to chemical or enzymatic hydrolysis and their similar conformation to that of their natural O-linked counterparts.² Thioglycoligases are recently developed mutants of retaining glycosidases in which the acid/base carboxylic acid residue has been replaced by an inert amino acid. These mutants catalyze the formation of *S*-glycosidically linked oligosaccharides by coupling of acceptors bearing a suitably positioned nucleophilic thiol group with donors having a good leaving group, such as dinitrophenol or fluoride (Figure 1).³ However, no α -thioglycoligases have yet been reported, raising the question of whether such mutants will function as described. This report describes the generation of the first α -thioglycoligases, the use of one such mutant to form an inhibitor of the wild-type enzyme, and elucidation of the three-dimensional structure of a YicI/thiosugar complex.

Two family 31 glycosidases, an α -xylosidase from *E. coli* (YicI) and an α -glucosidase from *Sulfolobus solfataricus* (MaIA),⁴ were selected for investigation as potential α -thioglycoligases. Previous kinetic studies have revealed a preference of YicI for the α -(1,6)-xylosidic linkage over the α -(1,4)-glucosidic linkage.⁵ By contrast, α -glucosidases, such as MaIA, prefer α -(1,4)-glucosides.

The acid/base catalysts of YicI (Asp482) and MaIA (Asp416) were mutated to alanine using the mega primer method, and the variants were purified by Ni–NTA affinity chromatography (see Supporting Information). The thioglycoligase reactions catalyzed by YicI D482A and MaIA D416A were first investigated using pNP 4-deoxy-4-thio- β -D-glucopyranoside⁶ as acceptor and an equimolar amount of the glycosyl donor, α -D-xylopyranosyl fluoride for YicI D482A, and α -D-glucopyranosyl fluoride for MaIA D416A. YicI D482A catalyzed the thioglycoligase reaction more efficiently than did MaIA D416A, the thio-linked disaccharides, **1** and **2**, being formed in 79 and 48% yields, respectively. Thioglycoligase reactions using pNP 4-deoxy-4-thio- β -D-xylopyranoside as acceptor were less efficient than those with the gluco-version, MaIA D416A not accepting 4-thioxylose at all, while YicI D482A yielded **3** in

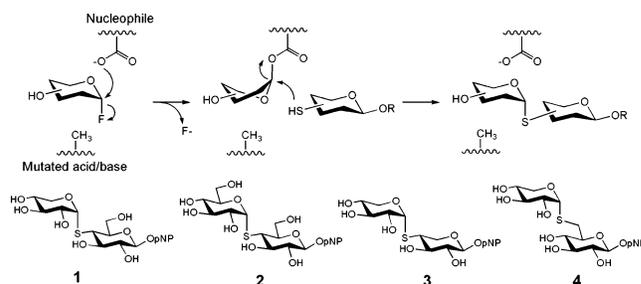


Figure 1. Mechanism of α -thioglycoligases and structures of products.

41% yield. Analysis of the isolated products after acetylation by ¹H and ¹³C NMR and mass spectrometry confirms these structures.

Given the preference of YicI for α -(1,6)-linkages, our interest moved to the synthesis of α -1,6-thioglycosidic linkages using these two thioglycoligases. Using pNP 6-deoxy-6-thio- β -D-glucopyranoside as acceptor, YicI D482A generated a single product **4** in 86% yield. MaIA D416A, consistent with the specificity of its parent enzyme, gave no product.

Next, the thiosaccharides, **1**, **3**, and **4**, were tested as potential inhibitors for YicI. Compound **4**, which differs from isoprimeverose (xylopyranosyl- α -(1,6)-glucopyranose), the natural substrate of YicI, only in the nature of the intersugar linkage where a sulfur atom replaces an oxygen atom, acted as a potent competitive inhibitor with a *K_i* value of 2 μ M. α -(1,4)-Linked thiosugars, **1** and **3**, also inhibited YicI well with *K_i* values of 27 and 14 μ M, respectively.

The three-dimensional structure of YicI in complex with **4** was determined by X-ray crystallography at 1.95 Å resolution. Six active sites in the hexameric quaternary structure of YicI are fully occupied by **4** (see Supporting Information). A seventh molecule of **4** is found between the N-terminal end of the $\alpha_8\beta_8$ barrel and the C-terminal domain of monomer F. This may be an artifact of crystal packing as it is only observed for the one monomer, but it is intriguing to speculate that carbohydrate-binding motifs may exist outside of the active site.¹ The electron density is clear for the ligands, with xyloside moieties in the –1 subsites being in the ⁴C₁ chair conformation, consistent with the conformational itinerary (⁴C₁ → ⁴H₃ → ¹S₃) previously proposed.¹ The Michaelis complex structure reveals the +1 subsite and suggests the role of the N-terminus of YicI (N-domain)¹ in the formation of the aglycon sites (Figure 2A). The +1 subsite of each monomer in the hexameric unit consists of four residues from its own $\alpha_8\beta_8$ barrel (Trp380, Phe417, Arg466, and Phe515) and three residues from the N-domains of three different monomers: Asp185 from its own monomer (black in Figure 2A), Asp49 from an adjacent monomer (orange), and Trp8 from a second distinct monomer (pink). Importantly, the hexameric structure has also been demonstrated in solution. Although Asp185,

[†] Department of Chemistry.

[‡] Department of Biochemistry and Molecular Biology.

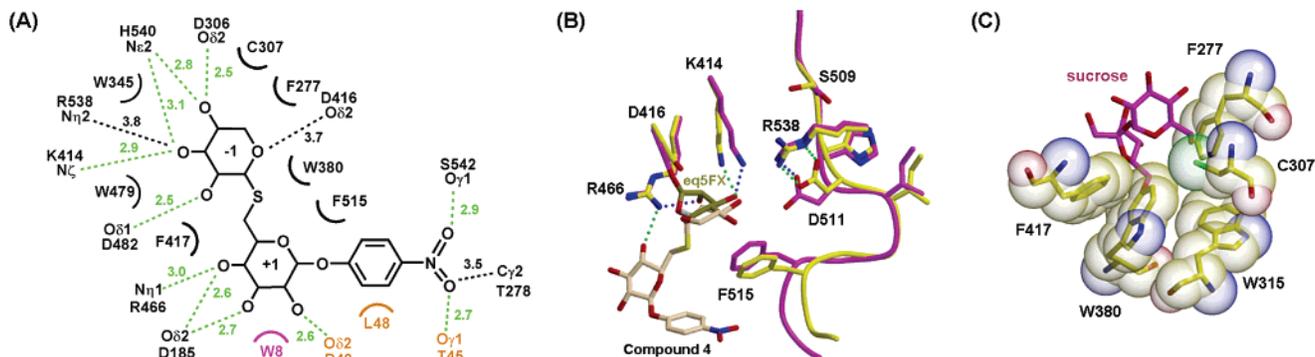


Figure 2. (A) Interactions of YicI with compound 4. Green dotted lines and numbers represent H-bonds and their lengths. Black dotted lines and numbers represent distances between 4 and residues within 4 Å. (B) Superposition of the Michaelis complex (2F2H) and the 5-fluoroxalosyl (eq5FX) enzyme intermediate (1XSK). Side chains for the residues are depicted in bond style coloring with carbon atoms in purple (1XSK) or yellow (2F2H). Hydrogen bonds in 1XSK and 2F2H are represented as blue and green dotted lines, respectively. (C) Modeled structure of YicI with sucrose.

Trp380, and Arg466 are conserved in most family 31 glycosidases, Trp8 and Asp49 are conserved only in α -xylosidases that are highly homologous to YicI. Consequently, the N-domain-mediated hexameric structure of YicI constructs an efficient +1 subsite for isoprimeverose, resulting in substrate recognition and transition state stabilization. Indeed, this structural evidence is consistent with previous kinetic studies that reveal lower K_m and higher k_{cat} values of YicI (0.55 mM and 51.3 s^{-1})⁵ for isoprimeverose than those of α -xylosidase from *S. solfataricus* (28.9 mM and 31.0 s^{-1}),⁷ which lacks the residues corresponding to Trp8 and Asp49 of YicI.

Superposition with the previously determined structure of the covalent intermediate (1XSK)¹ shows the reorientation of several residues and loops in the -1 subsite. Interestingly, a series of mutations of GAA that cause Pompe's disease, G643R (Ser509 in YicI), D645E/H/N (Asp511), C647W (Gly513), and G648S (Gly 514),⁸ are located on one of the loops that shifted upon binding of 4 (Figure 2B). In addition, Asp511 of YicI interacts with Arg538, which corresponds to Arg672 of GAA, another severe mutation site for Pompe's disease.⁸ Therefore, the mutations mentioned above make GAA less functional by disturbing the loop structure required for interactions with substrates. The hydrogen bond between Arg466 and O-2 of the 5-fluoroxalosyl ring (eq5FX) covalently linked with YicI in the intermediate¹ is interesting since Arg466 interacts with O-4 of the glucose in the +1 subsite in the Michaelis complex (Figure 2B). Therefore, Arg466 appears to stabilize the partial negative charge that accumulates on O-2 of the sugar at the transition state, much as has been proposed for Arg227 and His327 in cyclodextrin glucanotransferase.⁹ Indeed, the mutation R600C in GAA (R466 in YicI) results in a 99% loss of activity and confers Pompe's disease.^{8b}

Useful insights into the different aglycon specificities of other family 31 glycosidases are obtained through substrate overlays. Using the xylose bound in the -1 subsite as the basis for superposition with the glucose ring of sucrose, it is clear that O-6' of the fructose in the +1 subsite would clash with Trp380 (Figure 2C). However, both GAA and SUI5 are expected to have space around O-6' of the fructose because Trp380 has been substituted with Val, and Trp315 of YicI (4.4 Å away from O-6' of the fructose) has no counterpart.¹ This information may be valuable in the design of structure-based inhibitors for chemical chaperone therapy¹⁰ of Pompe's disease.

In conclusion, the generation of α -thioglycosides has been successfully achieved, for the first time, using the thioglycosylase strategy. Using the product of one of these thioglycosylases as a nonhydrolyzable substrate analogue, the three-dimensional structure of the Michaelis complex of its parent enzyme has been determined, providing valuable structural insights into the specificity and mechanism of family 31 α -glycosidases and the molecular basis of an associated genetic disease. This general strategy could be applied to other glycosidases.

Acknowledgment. We thank the Natural Sciences and Engineering Research Council of Canada, the Protein Engineering Network of Centres of Excellence of Canada, the Michael Smith Foundation for Health Research (to A.L. and Y.K.), and Commonwealth Scholarship (to T.K.) for financial support. We also thank Dr. Tony Warren for helpful discussions, and Dr. Michael Jahn for kindly providing 6-thio pNP glucoside. Coordinates and structure factors for the YicI:4 complex have been deposited in the Protein Data Bank with accession code 2F2H.

Supporting Information Available: Experimental procedures, NMR data, crystallization, data collection, additional structure figures, and complete refs 7 and 8. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- Lovering, A. L.; Lee, S. S.; Kim, Y. W.; Withers, S. G.; Strynadka, N. C. *J. Biol. Chem.* **2005**, *280*, 2102–2125.
- (a) Driguez, H. *Top. Curr. Chem.* **1997**, *187*, 85–116. (b) Aguilera, B.; Fernandez-Mayoralas, A. *J. Org. Chem.* **1998**, *63*, 2719–2723. (c) Sandgren, M.; Berglund, G. I.; Shaw, A.; Stahlberg, J.; Kenne, L.; Desmet, T.; Mitchinson, C. *J. Mol. Biol.* **2004**, *342*, 1505–1517.
- Jahn, M.; Marles, J.; Warren, R. A. J.; Withers, S. G. *Angew. Chem., Int. Ed.* **2003**, *42*, 352–354.
- Rolfmeier, M.; Haseltine, C.; Bini, E.; Clark, A.; Blum, P. *J. Bacteriol.* **1998**, *180*, 1287–1295.
- Okuyama, M.; Mori, H.; Chiba, S.; Kimura, A. *Protein Expr. Purif.* **2004**, *37*, 170–179.
- p*-Nitrophenyl (pNP) glycosides are preferred as acceptors since they bind more tightly than the free sugars, and products are more easily monitored by TLC and HPLC.
- Moracci, M.; et al. *J. Biol. Chem.* **2000**, *275*, 22082–22089.
- (a) Huie, M. L.; et al. *Biochem. Biophys. Res. Commun.* **1998**, *244*, 921–927. (b) Hermans, M. M.; et al. *J. Hum. Mutat.* **2004**, *23*, 47–56.
- Uitendhaag, J. C.; Mosi, R.; Kalk, K. H.; van der Veen, B. A.; Dijkhuizen, L.; Withers, S. G.; Dijkstra, B. W. *Nat. Struct. Biol.* **1999**, *6*, 432–436.
- Cohen, F. E.; Kelly, J. W. *Nature* **2003**, *42*, 905–909.

JA057904A