J|A|C|S
ARTICLES
Published on Web 11/02/2004

Directed Evolution of a Glycosynthase via Chemical Complementation

Hening Lin,† Haiyan Tao, and Virginia W. Cornish*

Contribution from the Department of Chemistry, Columbia University, New York, New York 10027

Received June 24, 2004; E-mail: vc114@columbia.edu

Abstract: Recently, we reported a general assay for enzyme catalysis based on the yeast three-hybrid assay, Chemical Complementation, which is intended to expand the range of chemical reactions to which directed evolution can be applied. Here, Chemical Complementation was applied to a glycosynthase derived from a retaining glycosidase, an important class of enzymes for carbohydrate synthesis. Using the yeast three-hybrid assay, the glycosynthase activity of the E197A mutant of the Cel7B from *Humicola insolens* was linked to transcription of a *LEU2* reporter gene, making cell growth dependent on glycosynthase activity in the absence of leucine. Then the *LEU2* selection was used to isolate the most active glycosynthase from a Glu197 saturation library, yielding an E197S Cel7B variant with a 5-fold increase in glycosynthase activity. These results not only establish Chemical Complementation as a platform for the directed evolution of glycosynthases, but also show the generality of this approach and the ease with which it can be applied to new chemical reactions.

Introduction

Directed evolution has been used to improve the activity of an existing enzyme and, more recently, even to modify enzyme substrate specificity.[1,2] Directed evolution involves generating large numbers of protein variants and then assaying these variants en masse for the desired function. While potentially a powerful approach for generating enzymes for chemical synthesis, commercial products, biomedical research, and even therapeutics, directed evolution in practice is limited to reactions that are inherently screenable or selectable, such as reactions where the product is fluorescent or an essential metabolite. Thus, several laboratories have sought to develop high-throughput assays for enzyme catalysis that are general and can be applied to a broad range of chemical reactions.[3−9] Our laboratory recently reported such an assay, which uses the yeast three-hybrid assay to link enzyme catalysis to reporter gene transcription in vivo.[10,11] We called this assay "Chemical Complementation" because it is like a complementation assay in that enzyme activity is required for the cellular phenotype, but the chemistry being complemented is controlled by the chemical linker between two small-molecule ligands and thus can be readily varied. In our initial report, we showed the feasibility of this approach using a β-lactamase enzyme. Here, we further show the generality of this approach, applying Chemical Complementation to the directed evolution of a glycosynthase, which can be used for oligosaccharide synthesis (Figure 1).

Despite their fundamental role in biological processes and potential use as therapeutics, it still remains difficult to synthesize carbohydrates. In the past two decades, there has been tremendous progress in the chemical synthesis of complex carbohydrates.[12−20] However, chemical synthesis is still limited by the need for differentially protected intermediates and reactant-dependent coupling yields and stereo-control. Enzymes, with their control of both regio- and stereochemistry, provide an obvious alternative to traditional small molecule chemistry for the synthesis of oligosacchrides. Glycosyltransferases, the natural enzymes responsible for the synthesis of oligo- and polysacchrides, and glycosidases, the enzymes that normally

† Hening Lin's current address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115.

(1) Joo, H.; Lin, Z.; Arnold, F. H. *Nature* **1999**, *399*, 670−673.
(2) Yano, T.; Oue, S.; Kagamiyama, H. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5511−5515.
(3) Pedersen, H.; Holder, S.; Sutherlin, D. P.; Schwitter, U.; King, D. S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 10523−10528.
(4) Xia, G.; Chen, L.; Sera, T.; Fa, M.; Schultz, P. G.; Romesberg, F. E. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6597−6602.
(5) Olsen, M. J.; Stephens, D.; Griffiths, D.; Daugherty, P.; Georgiou, G.; Iverson, B. L. *Nat. Biotechnol.* **2000**, *18*, 1071−1074.
(6) Tawfik, D. S.; Griffiths, A. D. *Nat. Biotechnol.* **1998**, *16*, 652−656.
(7) Ghadessy, F. J.; Ong, J. L.; Holliger, P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4552−4557.
(8) Griffiths, A. D.; Tawfik, D. S. *Embo J.* **2003**, *22*, 24−35.
(9) Firestine, S. M.; Salinas, F.; Nixon, A. E.; Baker, S. J.; Benkovic, S. J. *Nat. Biotechnol.* **2000**, *18*, 544−547.
(10) Baker, K.; Bleczinski, C.; Lin, H.; Salazar-Jimenez, G.; Sengupta, D.; Krane, S.; Cornish, V. W. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16537−16542.
(11) Michnick, S. W.; Valois, F. X. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16513−16515.
(12) Toshima, K.; Tatstusa, K. *Chem. Rev.* **1993**, *93*, 1503−1531.
(13) Fraser-Reid, B.; Konradsson, P.; Mootoo, D. R.; Udodong, U. *J. Chem. Soc. Chem. Commun.* **1988**, *12*, 823−825.
(14) Fraser-Reid, B.; Udodong, U. E.; Wu, Z.; Ottosson, H.; Merritt, J. R.; Rao, C. S.; Roberts, C.; Madsen, R. *Synlett* **1992**, 927−942.
(15) Schmidt, R. R.; Michel, J. *Angew. Chem., Int. Ed. Engl.* **1980**, *19*, 731−732.
(16) Schmidt, R. R. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 212−135.
(17) Seeberger, P. H.; Bilodeau, M. T.; Danishefsky, S. J. *Aldrichim. Acta* **1997**, *30*, 75−92.
(18) Hashimoto, S.; Honda, T.; Ikegami, S. *J. Chem. Soc. Chem. Commun.* **1989**, 685−687.
(19) Plante, O. J.; Andrade, R. B.; Seeberger, P. H. *Org. Lett.* **1999**, *1*, 211−214.
(20) Paulsen, H. *Angew. Chem., Int. Ed. Engl.* **1982**, *21*, 155−173.

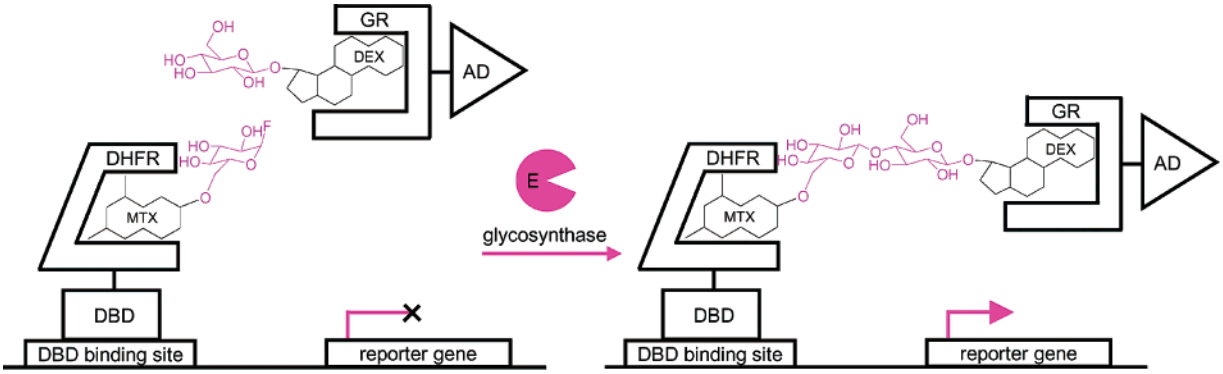

Figure 1. Chemical Complementation provides a high-throughput assay for glycosynthase activity. Chemical Complementation detects enzyme catalysis of bond formation or cleavage reactions based on covalent coupling of two small molecule ligands. The heterodimeric small molecule reconstitutes a transcriptional activator, turning on the transcription of a downstream reporter gene. Here, a dexamethasone (Dex)-methotrexate (Mtx) yeast three-hybrid system is used. Glycosynthase activity is detected as formation of a glycosidic linkage between a Mtx α-fluoride donor and a Dex acceptor.

hydrolyze carbohydrates, have both been used for carbohydrate synthesis.[21−25] The use of glycosyltransferases, however, is limited by the need for nucleotide diphosphate glycosyl donors, which are expensive and difficult to synthesize. Oligosaccharide synthesis using glycosidases, not surprisingly, suffers from low yields since the enzyme also catalyzes the hydrolysis of the desired product. Thus, alternative methods are being sought for enzyme-catalyzed carbohydrate synthesis.

Recently, Withers and co-workers demonstrated that retaining glycosidases can be engineered to glycosynthases simply by mutating the nucleophilic Glu residue at the base of the active site to a small hydrophobic residue and using an α-glycosyl fluoride as the donor substrate.[26] This strategy is based on extensive characterization of the mechanism of retaining glycosidases. Retaining glycosidases have two acidic residues (Glu or Asp) that flank the glycosidic bond being hydrolyzed. One acidic residue serves as a nucleophile, forming a glycosyl-enzyme intermediate with the substrate. While the other acts as a general acid/base catalyst both for the formation and subsequent hydrolysis of the glycosyl-enzyme intermediate. Mutation of the active site nucleophile to a small, hydrophobic residue both accommodates the glycosyl fluoride donor and inactivates the hydrolytic activity of the enzyme, allowing the reaction to proceed in the reverse direction (Figure 2). This approach was first demonstrated using the *Agrobacterium sp.* β-glucosidase/galactosidase (Abg).[26] The active site nucleophile Glu358 was mutated to Ala. This Abg:E358A variant was shown to accept both galactosyl fluoride and glucosyl fluoride as donors to form glycosidic bonds with several mono- and di-saccharides. This result opened a new route for carbohydrate synthesis, and already several retaining glycosidases have been successfully converted to glycosynthases using this strategy.[27−33] Directed

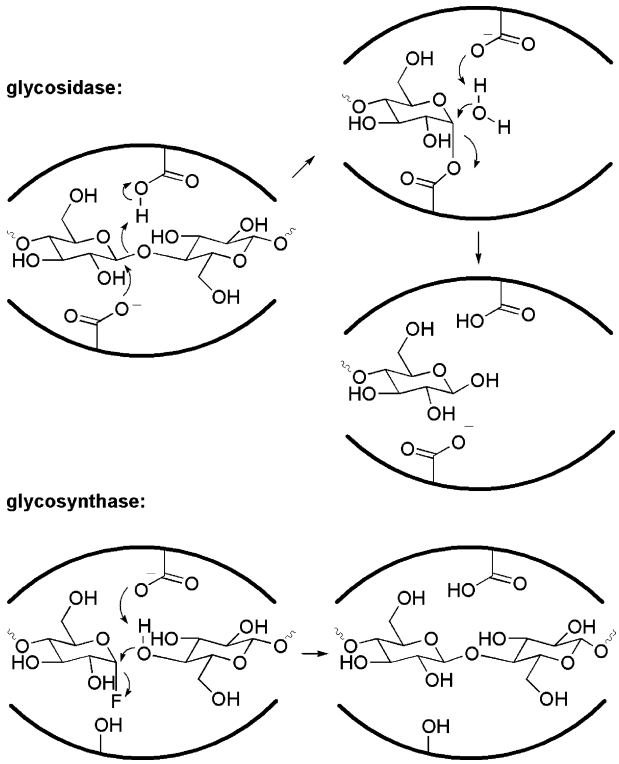

Figure 2. Single point mutation converts a retaining glycosidase to a glycosynthase. Retaining glycosidases have two acidic residues that flank the glycosidic bond to be hydrolyzed, one acting as a general acid/base catalyst and the other acting as a nucleophile, which attacks the anomeric position forming a glycosyl-enzyme intermediate. If the nucleophilic residue is mutated to a small hydrophobic residue, such as Ser or Ala, the protein can no longer hydrolyze the glycosidic bond. This enzyme variant can, however, catalyze the formation of a glycosidic bond between an α-glycosyl fluoride donor and a glycosyl acceptor.

evolution would offer an obvious route to improve the activity and alter the substrate selectivity of these enzymes, except that there is no intrinsic way to screen or select for glycosynthase activity. Mayer et al. developed a coupled enzyme assay using an endo-cellullase that can be used to screen for glycosynthase

(21) Flitsch, S. L. *Curr. Opin. Chem. Biol.* **2000**, *4*, 619−625.
(22) Koeller, K. M.; Wong, C. H. *Chem. Rev.* **2000**, *100*, 4465−4493.
(23) Koeller, K. M.; Wong, C. H. *Nature* **2001**, *409*, 232−240.
(24) Gijsen, H. J. M.; Qiao, L.; Fitz, W.; Wong, C. H. *Chem. Rev.* **1996**, *96*, 6, 443−473.
(25) Sears, P.; Wong, C. H. *Science* **2001**, *291*, 2344−2350.
(26) Mackenzie, L. F.; Wang, Q.; Warren, R. A. J.; Withers, S. G. *J. Am. Chem. Soc.* **1998**, *120*, 5583−5584.
(27) Davies, G. J.; Charnock, S. J.; Henrissat, B. *Trends Glycosci. Glycotechnol.* **2001**, *13*, 105−120.
(28) Faijes, M.; Fairweather, J. K.; Driguez, H.; Planas, A. *Chem.-Eur. J.* **2001**, *7*, 4651−4655.
(29) Fairweather, J. K.; Stick, R. V.; Withers, S. G. *Aust. J. Chem.* **2000**, *53*, 913−916.
(30) Mayer, C.; Zechel, D. L.; Reid, S. P.; Warren, R. A. J.; Withers, S. G. *Febs Lett.* **2000**, *466*, 40−44.
(31) Trincone, A.; Perugino, G.; Rossi, M.; Moracci, M. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 365−368.
(32) Nashiru, O.; Zechel, D. L.; Stoll, D.; Mohammadzadeh, T.; Warren, R. A. J.; Withers, S. G. *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 417−420.
(33) Fort, S.; Boyer, V.; Greffe, L.; Davies, G.; Moroz, O.; Christiansen, L.; Schulein, M.; Cottaz, S.; Driguez, H. *J. Am. Chem. Soc.* **2000**, *122*, 5429−5437.

mutants with improved activity.[34] This screen, however, can only be used for glycosynthases that synthesize products that are substrates of the endo-cellulase. Screens only allow relatively small libraries, ca. 10^4 variants, to be assayed.

Here, using Chemical Complementation, we have developed a *LEU2* growth selection for glycosynthase activity. The method was developed using the E197A mutant of Cel7B from *Humicola insolens*.[33] Cel7B is an endoglucanase that catalyzes the hydrolysis of β-1,4-linked glucosidic bonds in cellulose with retention of stereochemistry at the anomeric center.[35] The E197A mutant has been shown to be a glycosynthase.[33,36] Here, we present results showing that Chemical Complementation can read-out Cel7B:E197A glycosynthase activity. Further, the Chemical Complementation *LEU2* selection is used for the directed evolution of the Cel7B glycosynthase with a Glu197 saturation library.

Results and Discussion

Selection Scheme. Chemical Complementation detects enzyme catalysis of bond formation or cleavage reactions based on covalent coupling of two small molecule ligands in vivo (Figure 1). The heterodimeric small molecule reconstitutes a transcriptional activator, turning on the transcription of a downstream reporter gene. Bond formation is detected as activation of an essential reporter gene; bond cleavage, repression of a toxic reporter gene. The assay is high-throughput because it can be run as a growth selection where only cells containing functional enzyme survive. The assay can be readily extended to new chemistry simply by synthesizing small molecule heterodimers with different chemical linkers as the enzyme substrates.

We envisioned that Chemical Complementation could detect Cel7B:E197A glycosynthase activity as formation of a bond between a methotrexate-disaccharide-fluoride donor (Mtx-Lac-F) and a dexamethasone-disaccharide acceptor (Dex-Cel) in a Dex-Mtx yeast three-hybrid system (Figure 1).[37] In this system, a heterodimeric Dex-Mtx small molecule dimerizes the hormone-binding domain of the glucocorticoid receptor (GR), which binds to Dex, and dihydrofolate reductase (DHFR), which binds to Mtx. DHFR is fused to a DNA-binding domain (DBD), and GR is fused to a transcription activation domain (AD), such that Dex-Mtx effectively reconstitutes the transcriptional activator (DBD-AD) and increases transcription of a downstream reporter gene. Here, LexA is used as the DBD, and B42 as the AD. Both the DBD-DHFR and AD-GR fusion proteins are expressed from a GAL1 promoter. The DBD-DHFR gene is integrated into the chromosome at the *ade4* locus, and the AD-GR gene is on a 2μ plasmid. This combination stabilizes the transcription read-out without significantly decreasing transcription activation.[38] The reporter gene is *LEU2*, which allows for a growth selection in the absence of leucine. The *LEU2* gene is integrated into the chromosome under the control of 6 tandem LexA operators.[39] The Cel7B gene is expressed from a 2μ plasmid under control of a repressible MET promoter.

Dex-Cel and Mtx-Lac-F substrates. The carbohydrate substrates for the Chemical Complementation selection must incorporate the Dex and Mtx ligands, yet still be efficient substrates for Cel7B glycosynthase variants. The structure of the two substrates designed for use in this study, Dex-Cel and Mtx-Lac-F, are shown in Figure 3. The design was based on the high-resolution X-ray structure of the *H. insolens* Cel7B,[40] the X-ray structure of a homologous cellulase *Fusarium oxysporum* (*F. oxysporum*) Cel7B complexed with a nonhydrolyzable substrate analogue thio-DP5,[41] and in vitro characterization of the substrate specificity of the *H. insolens* Cel7B glycosidase and Cel7B:E197A glycosynthase.[33,35] Cel7B from *H. insolens* and Cel7B from *F. oxysporum* share more than 50% amino acid identity, and their three-dimensional structures are very similar. Both the kinetic characterization and high-resolution structures suggest that there are five subsites in the active site that accommodate five glucose units. Four of the five subsites (-2, -1, $+1$, $+2$) contribute most of the binding energy, with the fifth ($+3$) contributing only slightly. We therefore decided to use two disaccharide compounds as the substrates for the Cel7B:E197A glycosynthase. Since Cel7B protein's natural substrate is β-1,4-linked glucose polymer, the two disaccharide substrates should be or be very close to cellobiose. For the acceptor compound, we used cellobiose with Dex attached at the anomeric position. Attachment of Dex at the anomeric position should not disrupt Cel7B binding since attachment of another sugar unit at this position is tolerated. However, for the glycosyl fluoride donor, cellobiosyl fluoride cannot be used because it can also act as an acceptor and therefore self-polymerize. Thus, we chose to use an epimer of cellobiose fluoride, lactosyl fluoride, which differs only in the stereochemistry at the 4′ position. As an artifact of the synthetic protocol, the 4′ hydroxy was instead replaced by a bromide, also a suitable substrate here. Mtx was installed at the 6′ position of the galactose unit to facilitate the chemical synthesis and because the crystal structure suggests that this position is exposed and, therefore, that addition of Mtx should not interfere with substrate binding.

The retro-synthetic analysis of the two substrates is shown in Figure 3. Dex-Cel was prepared from the commercially available cellobiose octa-acetate. The anomeric O-acetate was converted to the bromide, and then the glycosidation reaction between cellobiosyl bromide and 6-bromo-1-hexanol was carried out with catalysis by mercury salts. A thiol derivative of Dex could then be introduced by a simple substitution reaction. After deprotection with sodium methoxide, 18 mg of Dex-Cel was obtained in 4 steps in 27% overall yield. Mtx-Lac-F was assembled from several smaller building blocks. The α-methyl ester of Mtx was obtained in 3 steps. An amino thiol linker was synthesized from 6-aminohexanoic acid in 3 steps. Next, the Mtx α-methyl ester was coupled to a protected lactose derivative via the amino thiol linker. Transient protection of the 4′,6′-hydroxyl groups of the galactose unit of lactose as the

(34) Mayer, C.; Jakeman, D. L.; Mah, M.; Karjala, G.; Gal, L.; Warren, R. A.; Withers, S. G. *Chem. Biol.* **2001**, *8*, 437−443.
(35) Schou, C.; Rasmussen, G.; Kaltoft, M. B.; Henrissat, B.; Schulein, M. *Eur. J. Biochem.* **1993**, *217*, 947−953.
(36) Ducros, V. M.-A.; Tarling, C. A.; Zechel, D. L.; Brzozowski, A. M.; Frandsen, T. P.; Ossowski, I. v.; Schulein, M.; Withers, S. G.; Davies, G. J. *Chem. Biol.* **2003**, *10*, 619−628.
(37) Lin, H.; Abida, W. M.; Sauser, R. T.; Cornish, V. W. *J. Am. Chem. Soc.* **2000**, *122*, 4247−4248.
(38) Baker, K.; Sengupta, D.; Salazar-Jimenez, G.; Cornish, V. W. *Anal. Biochem.* **2003**, *315*, 134−137.
(39) Licitra, E. J.; Liu, J. O. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12817−12821.
(40) Mackenzie, L. F.; Sulzenbacher, G.; Divne, C.; Jones, T. A.; Woldike, H. F.; Schulein, M.; Withers, S. G.; Davies, G. J. *Biochem. J.* **1998**, *335*, 409−416.
(41) Sulzenbacher, G.; Driguez, H.; Henrissat, B.; Schulein, M.; Davies, G. J. *Biochemistry* **1996**, *35*, 15280−15287.

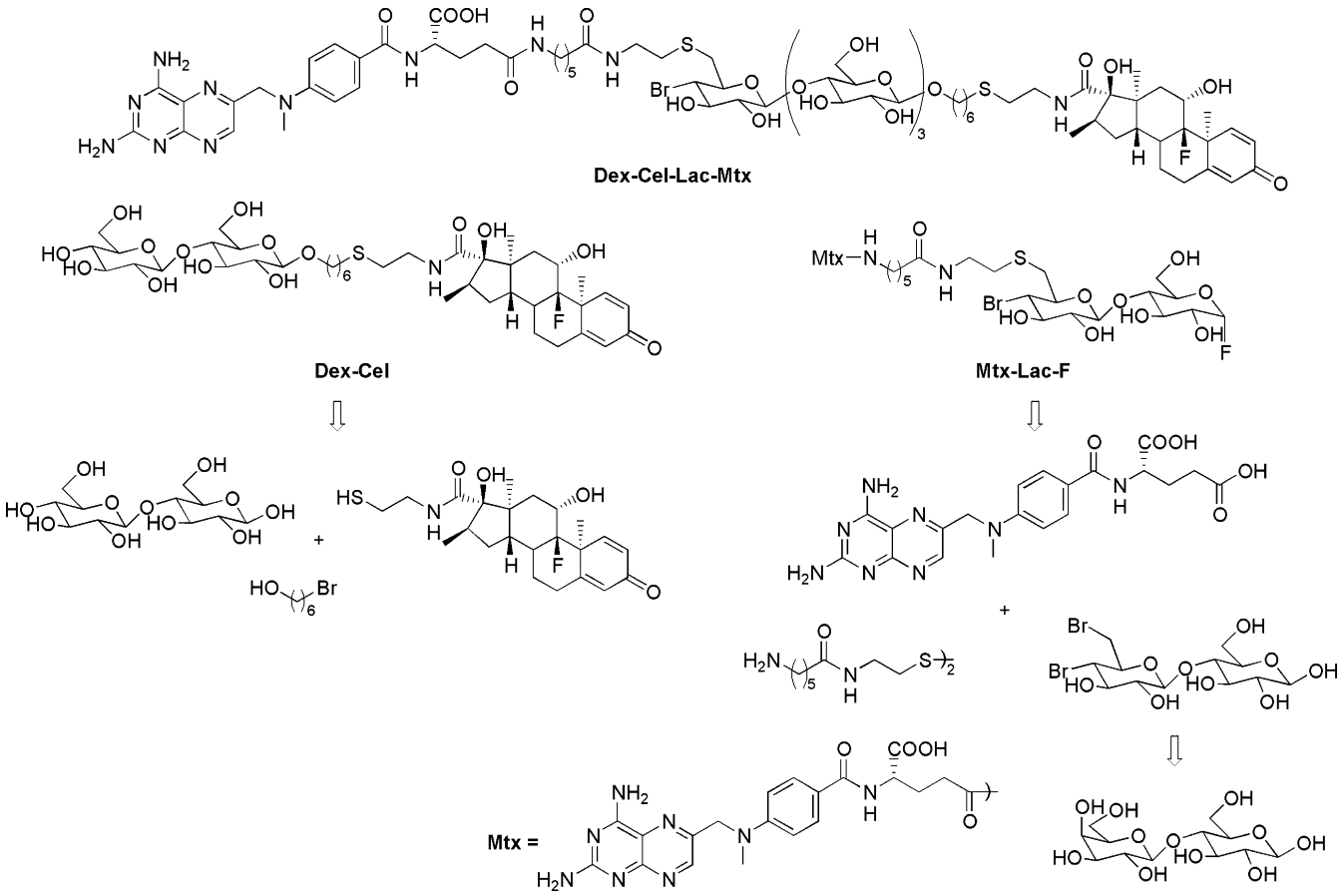

Figure 3. Retro-synthetic analysis of the Dex-Cel and Mtx-Lac-F substrates.

4-methoxyl benzylidene acetal allows Mtx to be introduced selectively at the 6′ position. The Mtx-Lac compound was then converted to the α-fluoride using a 7:3 hydrogen fluoride-pyridine. The α-fluoride is the thermodynamic product. After global deprotection by lithium hydroxide, 110 mg of Mtx-Lac-F was synthesized in 9 steps in 7% overall yield when calculated from the longest route (from L-glutamic acid 5-*tert*-butyl 1-methyl ester). The detailed synthesis of the two substrates will be reported in a separate publication (H. Tao, P. P. Peralta-Yahya, H. Lin and V. W. Cornish, unpublished results).

An in vitro assay was developed to ensure that Dex-Cel and Mtx-Lac-F are in fact substrates for the Cel7B:E197A glycosynthase. While Cel7B:E197A had been shown to be a glycosynthase based on product synthesis,[33] there was no reported kinetic characterization of this enzyme at the time of this study.[36] Thus, we developed a qualitative assay based on the disappearance of the anomeric proton of Mtx-Lac-F coupled to F as measured by ^{1}H NMR and high-pressure liquid chromatography (HPLC). Because Cel7B is known to be difficult to express using standard methods, the Cel7B:E197A enzyme was provided by Novozymes using a proprietary protocol for overexpression in *Aspergillus oryzae*. Lactosyl fluoride and benzyl cellobioside, known substrates of Cel7B:E197A were used as a control. In a reaction containing 8.6 mg (10 mM) Dex-Cel, 9.5 mg (10 mM) Mtx-Lac-F and 1.2 mg (30 μM) Cel7B:E197A in 0.1 M sodium phosphate buffer pH 7 with 13% (v/v) DMSO at 37 °C, Dex-Cel-Lac-Mtx was produced in 25% yield after 24 h (Figure 3). The Dex-Cel-Lac-Mtx product was purified by HPLC, and its structure was confirmed by ^{1}H NMR and mass spectrometry. By comparison, the reaction with lactosyl fluoride and benzyl cellobioside under identical conditions gave the desired product in >90% yield based on HPLC after 12 h. The in vitro synthesis of the Dex-Cel-Lac-Mtx product further allowed us to test to ensure that this product is an efficient chemical dimerizer in the yeast three-hybrid assay. In a standard *LEU2* growth assay, cells incubated with 10 μM Dex-Cel-Lac-Mtx showed a significant growth advantage compared with no Dex-Cel-Lac-Mtx, and the growth depended on the concentration of Dex-Cel-Lac-Mtx in the media (Figure 4).

Chemical Complementation Links Glycosynthase Activity to *LEU2* Transcription in Vivo. Glycosynthase activity could be detected in vivo using Chemical Complementation. Using standard conditions for a *LEU2* growth assay, we showed that expression of Cel7B:E197A in the presence of Dex-Cel and Mtx-Lac-F conferred a growth advantage to the yeast three-hybrid selection strain V1019Y (Figure 5), presumably because the Cel7B:E197A glycosynthase catalyzed the synthesis of Dex-Cel-Lac-Mtx. Several control experiments were carried out to confirm that the growth advantage is indeed caused by the catalytic activity of Cel7B:E197A. First, we showed that transcription activation required both the donor and the acceptor substrates. The yeast three-hybrid selection strain expressing Cel7B:E197A was grown in the presence of no small molecule, 10 μM Dex-Cel alone, 10 μM Mtx-Lac-F alone, or 10 μM Dex-Cel and 10 μM Mtx-Lac-F in media lacking the appropriate auxotrophs and leucine. Activation of the *LEU2* reporter gene results in an increase in cell growth and hence in the OD_{600} for the cell culture. Cell growth was at background levels for cells grown with no small molecule or with only Dex-Cel and near background levels with only Mtx-Lac-F. Cells grown in the

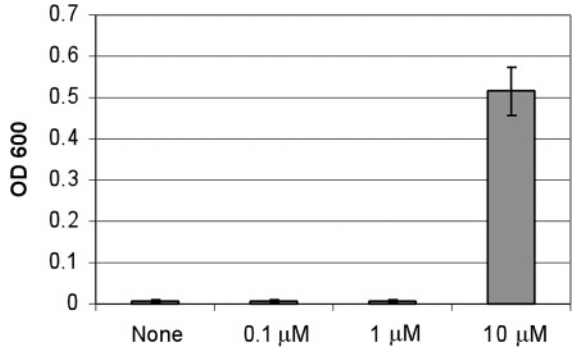

Figure 4. Dex-Cel-Lac-Mtx product activates *LEU2* reporter gene transcription in the yeast three-hybrid assay. Yeast cells expressing the DBD-DHFR and AD-GR fusion proteins and containing a *LEU2* reporter gene were grown in selective media lacking the appropriate auxotrophs and leucine with various concentrations of Dex-Cel-Lac-Mtx for 3 days. Activation of the *LEU2* reporter gene results in an increase in cell growth in the absence of leucine, which is measured as an increase in the OD_{600}. The levels of transcription activation (OD_{600}) after 6 days with 10 μM Dex-Cel-Lac-Mtx are comparable to those with 1 μM Dex-Mtx, shown in Figure 5. The growth assays were carried out in triplicate, and the error bars correspond to the standard deviation from the mean.

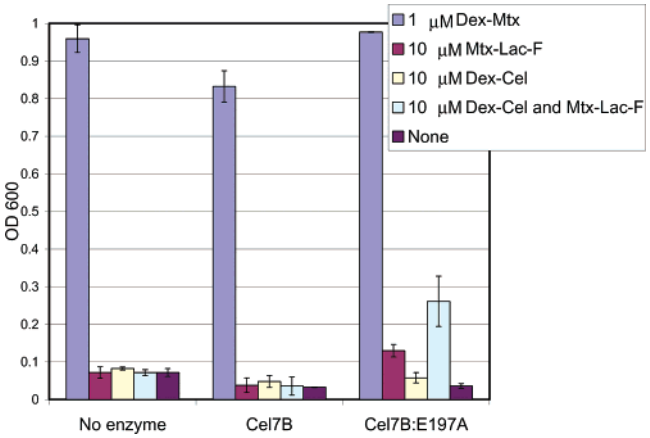

Figure 5. Chemical Complementation links Cel7B:E197A glycosynthase activity to *LEU2* reporter gene transcription in vivo. Yeast cells containing the DBD-DHFR and AD-GR fusion proteins and the *LEU2* reporter gene, expressing either no enzyme, Cel7B glycosidase, or Cel7B:E197A glycosynthase were grown in selective medium with or without the Dex-Cel and Mtx-Lac-F substrates. Dex-Mtx, a small molecule dimerizer with a methylene linker, was used as a positive control. Cells expressing no enzyme or Cel7B glycosidase were used as negative controls. Activation of the *LEU2* reporter gene and hence cell growth in the absence of leucine is reported here as the OD_{600} after 6 days of growth at 30 °C. The growth assays were carried out in triplicate, and the error bars correspond to the standard deviation from the mean.

presence of both Dex-Cel and Mtx-Lac-F showed a clear growth advantage (Figure 5). Next, we showed that transcription activation is dependent on the catalytic activity of Cel7B:E197A. Cells expressing the Cel7B glycosidase, which should be able to bind the two substrates, but should be much less efficient at product synthesis, showed no growth advantage in the presence of Dex-Cel and Mtx-Lac-F. Furthermore, cells expressing Cel7B were indistinguishable from cells expressing no enzyme. Together, these data suggest that the Cel7B:E197A glycosynthase activity can be detected using Chemical Complementation.

Further, the growth advantage conferred by the Cel7B:E197A glycosynthase was used to select the glycosynthase from a pool of inactive variants. A mock library containing 100:1 Cel7B glycosidase to Cel7B:E197A glycosynthase was used to transform the yeast three-hybrid selection strain V1019Y. Here, the Cel7B glycosidase is used as the "inactive" control. First, the transformed cells were plated under nonselective conditions to determine the transformation efficiency and 10 random colonies were analyzed to establish the integrity of the library. The transformation efficiency was 10^5 transformants per μg of DNA. All 10 colonies were found to express Cel7B, as would be expected from the 100:1 library ratio, based on colony PCR and restriction digestion. Then the library was plated on selective media lacking the appropriate auxotrophs and leucine and containing the two substrates Dex-Cel and Mtx-Lac-F and grown at 30 °C. After 6 days, the 10 largest colonies were picked and analyzed by colony PCR and restriction digestion. Of these 10 colonies, 8 contained the Cel7B:E197A glycoynthase, and 2 contained the Cel7B glycosidase, which corresponds to an enrichment of 400-fold after a single round of selection.

Directed Evolution of Cel7B. Finally, the Chemical Complementation *LEU2* selection was used for the directed evolution of Cel7B with a Glu197 saturation library. Position Glu197 was randomized to all 20 amino acids using cassette mutagenesis with a NNK codon where N is all four bases and K is either T or G. This library was transformed into the yeast three-hybrid selection strain V1019Y and first plated under nonselective conditions. The transformation efficiency was 10^6 colonies per μg DNA, more than sufficient to cover the library size. These colonies were combined and plated on selective media lacking the appropriate auxotrophs and leucine and containing the Dex-Cel and Mtx-Lac-F substrates. After 10 days incubation at 30 °C, the 96 largest colonies were picked from the selection plate. To eliminate false positives, a secondary screen was carried out by monitoring the growth of the selected mutants under selective conditions with and without the two substrates in a 96-well plate. Only those colonies that showed better growth with the two substrates were considered true positives and subjected to further characterization. Of the 96 colonies picked, the first 8 were all active, and totally 35 were active in the secondary screen. The Cel7B genes from those 35 positives were PCR amplified and sequenced. One colony that did not show small molecule-dependent growth was also picked for sequencing and thus served as a negative control.

Three mutants, Ala (10), Gly (11), and Ser (7), occurred most frequently and represent 80% of the positive clones. In addition, the Leu mutant was isolated twice and the Pro, Asp, and Thr mutants were all isolated once. Of the 10 Ala mutants isolated, three of them also had an additional N196D mutation, which was caused by a single base change from A to G. Sequencing of total DNA from the original library ruled out the possibility that this mutation was encoded at a high level by error in the E197X cassette library. The negative control chosen for sequencing turned out to be Arg, which is unlikely to be an active glycosynthase based on our understanding of the mechanism of glycosidases and glycosynthases. To confirm that these mutants are indeed glycosynthases, plasmids for individual mutants were isolated, transformed back into the yeast three-hybrid selection strain, and re-subjected to the secondary screen with or without small molecules. All of these variants again showed a growth advantage with the Dex-Cel and Mtx-Lac-F substrates, and the Cel7B:N196D/E197A variant grew faster than the Cel7B:E197A variant with the two substrates.

Characterization of Evolved Variants. The Cel7B:E197A, Cel7B:E197S, and Cel7B:N196D/E197A variants were all then

Table 1. Glycosynthase Activities and Protein Purification Yields for Cel7B Variants[a]

	E197A	E197S	N196D/E197A
specific activity: mol [F]/(min^{-1}•mol [E_0])	8 ± 2	40 ± 5	7 ± 1
protein purification yield: nmol/L	6.1	4.6	7.3

[a] Glycosynthase activity for tetrasaccharide synthesis from α-lactosyl fluoride and *p*-nitrophenyl β-cellobioside (PNPC) was measured for the *Humicola insolens* Cel7B variants in sodium phosphate buffer, pH 7.0, at room temperature. Specific activities were determined by measuring the fluoride ion release rate by a fluoride ion selective electrode. The protein purification yields are the yield of purified protein as determined by western analysis from total cell culture.

overexpressed and purified to allow determination of their glycosynthase activities. The Cel7B protein variants were overexpressed from the MET promoter using the same 2μ expression plasmid as for the Chemical Complementation assay. Sequencing showed the Cel7B:E197S plasmid to have four additional mutations, S-L17P (in the *S*ignal peptide), I98V, E150G, and T267A, and the Cel7B:N196D/E197A variant to have one additional mutation, F16L. The proteins were purified with these additional mutations, which may also contribute to the improved properties of the proteins, although all of these mutations are distant from the enzyme active site. The protein variants were purified via their 6-His tags using HisTrap kit (Amersham Biosciences). All three proteins were shown to be pure, except for a common lower MW contaminant, by SDS-PAGE and staining with Commassie Blue. The Cel7B protein was further identified by comigration with authentic Cel7B:E197A protein purified from *Aspergillus oryzae* by Novozymes (vide infra) and Western analysis with an anti-His tag antibody. On the basis of this comigration, these Cel7B variants purified from *S. cerevisiae* are believed to both have their signal peptide processed and to be glycosylated. Efforts to confirm whether signal peptide processing had occurred by *N*-terminal sequencing were inconclusive. The concentrations of the Cel7B protein variants were determined by comparison to a 6-His tagged protein of known concentration by a standard curve derived from Western analysis with an anti-His tag antibody.

The glycosynthase activities of these three variants were measured based on the rate of expulsion of fluoride ion from the α-fluoro donor substrate (Table 1).[36] For the substrates, α-lactosyl flouride (Lac-F) and *p*-nitrophenyl-β-cellobioside (PNPC) were used. These are two of the substrates used in the literature to characterize cellulase glycosynthase variants, and these two substrates are structurally similar to Mtx-Lac-F and Dex-Cel. The rate of fluoride ion expulsion was determined essentially as reported for other glycosynthase enzymes. The specific activity reported here is with 10 mM Lac-F, 14.3 mM PNPC, and 0.86−1.4 μM enzyme in 37.5 mM sodium phosphate, pH 7. The original glycosynthase variant Cel7B:E197A has a specific activity of 8 ± 2 mol product/min/mol enzyme. The Cel7B:E197S variant isolated from this selection shows a significant improvement in activity, with a specific activity of 40 ± 5 mol/min/mol. The Cel7B:N196D/E197A variant, somewhat surprisingly, is within experimental error of Cel7B:E197A, with a specific activity of 7 ± 1 mol/min/mol.

Significantly, the Chemical Complementation *LEU2* selection has given a 5-fold improvement in the glycosynthase activity of Cel7B:E197A via directed evolution. While mutational studies on other retaining glycosidases have shown that mutation of the Glu nucleophilic catalyst at the base of the binding pocket can improve the glycosynthase activity of these enzymes, the effect of individual mutations differs for retaining glycosidases from different species.[42−45] For example, the glycosynthase activity (k_{cat}/K_M) of the E231A and E231S variants of the Barley (1,3)-β-D-glucan endohydrolase are within experimental error of one another.[43] Concurrent with this work, characterization of the *H. insolens* Cel7B:E197A and Cel7B:E197S variants has been reported recently by Ducros and Davies and co-workers.[36] While the proteins were purified from different sources, the specific activities reported here are in good agreement with this report. The activity of the Cel7B:E197A variant is within experimental error with this report, and the Cel7B:E197S variant is slightly less active, which may also reflect the additional mutations in the Cel7B:E197S gene derived from the selection. Despite the fact that the Cel7B:N196D/E197A variant arose frequently even though the 196 mutation was not encoded in the library, the specific activity of this variant is indistinguishable from that of Cel7B:E197A. While we cannot rule out that this mutation specifically increases the activity with the Mtx-Lac-F and Dex-Cel substrates, this explanation seems unlikely given the structural similarity of Mtx-Lac-F and Dex-Cel to Lac-F and PNPC in the vicinity of the mutation. Finally, the finding that the Ala, Ser, and Gly variants arise most frequently from the growth selection provides experimental support for the idea that these mutations would give rise to the most active glycosynthase variants.

In addition to the catalytic activity, the in vivo expression levels of an enzyme limit its practical utility on a preparative scale. Thus, the expression levels of the Cel7B:E197A, Cel7B:E197S, and Cel7B:N196D/E197A variants in *S. cerevisiae* were compared. Interestingly, Western analysis of the crude cell lysates showed that the Cel7B:E197S protein had impaired expression, while the Cel7B:N196D/E197A protein had improved expression levels (data not shown). This observation was confirmed by the protein expression yields. While 6.1 nmol of protein was isolated from 1 L of cell culture for the original Cel7B:E197A glycosynthase, only 4.6 nmol/L of protein was isolated for the Cel7B:E197S variant, and 7.3 nmol/L was isolated for the Cel7B:N196D/E197A variant (Table 1). The impaired expression of the Cel7B:E197S protein could explain why this more active variant was not isolated more frequently than the E197A variant from the Chemical Complementation *LEU2* selection. While several factors ranging from codon usage to protein folding to the additional mutations in these mutant genes could be responsible for a decrease in protein expression levels, it is interesting to speculate that the N196D charge compensation mutation may effect the folding and/or stability of the Glu197 variant, rather than its catalytic activity. These results suggest the most useful variants may be those that balance in vitro activity with in vivo expression. The active site residues of retaining glycosidases may have been fine-tuned through evolution for the stabilization of two acidic residues juxtaposed across the protein active site. Thus, in addition to modifying the Glu nucleophile at the base of the binding pocket,

(42) Mayer, C.; Gal, L.; Warren, R. A. J.; Withers, S. G. *Faseb J.* **1999**, *13*, A1444−A1444.
(43) Hrmova, M.; Imai, T.; Rutten, S. J.; Fairweather, J. K.; Pelosi, L.; Bulone, V.; Driguez, H.; Fincher, G. B. *J. Biol. Chem.* **2002**, *277*, 30102−30111.
(44) Malet, C.; Planas, A. *Febs Lett.* **1998**, *440*, 208−212.
(45) Jakeman, D.; Withers, S. *Can. J. Chem.* **2002**, *80*, 866−870.

the surrounding active site residues may need to be readjusted to accommodate this mutation in order to convert a retaining glycosidase to a glycosynthase.

Chemical Complementation. Here, Chemical Complementation has been successfully applied to a new chemical reaction, glycoside bond formation. As hoped, no further modifications were needed to the yeast three-hybrid system. All that was required to detect glycosynthase activity was to add the Dex and Mtx saccharine substrates. This result shows the generality of Chemical Complementation, and the ease with which it can be applied to new chemical reactions. In addition, it shows that Chemical Complementation can detect not only bond cleavage, but also bond formation reactions. With a bond formation reaction, enzyme activity could be linked to transcription of an essential gene, allowing Chemical Complementation to be run as a growth selection. The size of the E197 saturation library selected here was quite small, with only 32 members at the DNA level. The transformation efficiency of *S. cerevisiae* will allow much larger libraries, on the order of 10^5-10^7. For the saturation library here the transformation efficiency was in fact 10^6 colonies per μg of DNA, showing that the yeast three-hybrid selection strain is not impaired in its transformation efficiency. For yet larger library sizes, the system will have to be moved to *E. coli*, and we have made efforts in that direction.[46] Rather than library size, the challenge for more sophisticated libraries likely will be discriminating variants with a broad range of activities, as in any high-throughput assay. Currently, we are using the Chemical Complementation *LEU2* selection to increase the activity and expression levels of the glycosynthase and to modify its substrate specificity.

Experimental Section

General Methods. Restriction enzymes, Vent DNA polymerase and T4 DNA ligase were purchased from New England Biolabs. The dNTPs used in the Polymerase Chain Reaction (PCR) were purchased from Pharmacia Biotech. Oligonucleotides were purchased from The Great American Gene Company (www.geneco.com). Single stranded DNA (DNA sodium salt type III, salmon testes, cat# D1626) used for yeast transformation was purchased from Sigma. The bacto-agar, bacto-peptone and bacto-yeast extract were purchased from DIFCO. Except otherwise noted, all chemicals were purchased from Aldrich or Sigma. Dex-Cel (15 mM), Mtx-Lac-F (15 mM) and Dec-Cel-Lac-Mtx (2 mM) were dissolved in DMF and stored at −80 °C. All aqueous solutions were made with distilled water prepared from a Mili-Q Water System. For PCR, a MJ Research PTC-200 Pellier Thermal Cycler was employed. The transformation of *E. coli* was carried out by electroporation using a Bio-Rad *E. coli* Pulser. Restriction digests were carried out as recommended by New England Biolabs. All yeast techniques including the preparation of yeast media were carried out following standard protocols.[47] All other standard molecular biology techniques were carried out essentially as described. [48,49]

Chemical Synthesis of Dex-Cel and Mtx-Lac-F. The detailed procedure for the synthesis of Dec-Cel and Mtx-Lac-F will be reported in a separate publication. Dex-Cel: ^{1}H NMR (400 MHz, CD_3OD) δ 7.45 (d, J = 10.1 Hz, 1), 6.32 (dd, J = 1.9, 10.1 Hz, 1), 6.10 (s, 1), 4.43 (d, J = 7.8 Hz, 1), 4.30 (d, J = 7.8 Hz, 1), 4.26 (d, J = 11.2 Hz, 1), 3.90 (m, 4), 3.68 (dd, J = 5.4, 11.8 Hz, 1), 3.63−3.20 (m, 14), 3.14 (m, 1), 2.74 (m, 1), 2.66 (m, 2), 2.60 (t, J = 7.3 Hz, 2), 2.57−2.36 (m, 2), 2.22 (m, 2), 1.90 (m, 1), 1.77 (q, J = 11.5 Hz, 1), 1.7−1.4 (m, 14), 1.23 (m, 1), 1.11 (s, 3), 0.92 (d, J = 7.4 Hz, 3);^{13}C NMR (100 MHz, CD_3OD) δ 187.81, 174.41, 170.05, 155.00, 128.68, 124.01, 103.60, 103.19, 102.54, 100.79, 87.22, 79.77, 77.18, 76.92, 75.51, 74.00, 73.93, 72.36, 71.99, 70.44, 69.91, 61.54, 60.97, 49.64, 49.42, 44.04, 39.23, 36.03, 35.49, 35.08, 34.88, 32.66, 31.68, 31.49, 29.99, 29.92, 28.90, 28.09, 25.95, 22.94, 17.06, 14.53. HRMS m/z 862.4033 $(M+H)^+$, calculated 862.4059; Mtx-Lac-F: ^{1}H NMR (400 MHz, 4:1 CD_3OD:DMSO-d_6) δ 8.73 (s, 1), 7.84 (d, J = 8.9 Hz, 2), 6.93 (d, J = 9.0 Hz, 2), 5.60 (dd, J = 2.8, 53.7 Hz, 1), 4.98 (s, 2), 4.6−4.5 (m, 2), 4.0−3.45 (m, 10), 3.30−3.15 (m, 5), 3.92 (dd, J = 5.6, 14.4 Hz, 1), 2.87−2.70 (m, 2), 2.5−2.0 (m, 7), 1.7−1.3 (m, 6); ^{13}C NMR (100 MHz, $CDCl_3$) δ 174.53, 173.46, 168.23, 163.62, 151.93, 151.73, 149.25, 129.40, 121.78, 111.80, 109.10, 106.87, 104.09, 103.48, 78.81, 77.06, 76.29, 75.38, 73.74, 72.65, 72.31, 72.06, 71.87, 71.79, 66.52, 60.25, 56.29, 55.85, 55.22, 53.25, 52.10, 49.19, 36.27, 34.81, 32.93, 32.49, 30.07, 29.44, 27.59, 26.88, 25.96. HRMS m/z 1015.3007 $(M+H)^+$, calculated 1015.2995. During the final deprotection step for Mtx-Lac-F, a side product, which is believed to be a keto derivative of the galactose unit resulting from an elimination reaction, was generated (HRMS m/z 935.3732, calculated 935.3733). The *LEU2* growth assays and selection experiments were carried out with a 1:1 mixture of Mtx-Lac-F to this side product. The *LEU2* growth assays were confirmed using Mtx-Lac-F of >90% purity, prepared by preparative HPLC using a linear gradient of 100% water to 100% acetonitrile over 85 min and a C_{18} column. Cel7B:E197A catalyzed synthesis of Dex-Cel-Lac-Mtx from Dex-Cel and Mtx-Lac-F showed the side product was not a substrate for the Cel7B:E197A glycosynthase.

Enzymatic synthesis of Dex-Cel-Lac-Mtx. Following a published procedure,[33] Dex-Cel (8.6 mg, 0.010 mmol) and Mtx-Lac-F (9.5 mg, 0.010 mmol) were dissolved in 0.1 M pH 7 sodium phosphate buffer (1.0 mL) and DMSO (75 μL). The Cel7B:E197A enzyme (1.2 mg, provided by Novozymes) was added to the solution, and the reaction was incubated at 37 °C for 2 days. The product was purified by HPLC with a gradient of 100% water to 100% acetonitrile over 85 min on a C_{18} column (retention time = 56 min). The product was obtained in 25% yield: ^{1}H NMR (400 MHz, 5:1 DMSO-d_6:CD_3OD) δ 8.57 (s, 1), 8.29 (d, J = 7.3 Hz, 1, amide proton), 7.89 (t, J = 5.2 Hz, 1, amide proton), 7.81 (t, J = 5.9 Hz, 1, amide proton), 7.73 (d, J = 9.0 Hz, 2), 7.57 (t, J = 5.8 Hz, 1, amide proton), 7.30 (d, J = 10.1 Hz, 1), 6.82 (d, J = 9.0 Hz, 2), 6.21 (dd, J = 1.7 and 10.1 Hz, 1), 5.99 (s, 1), 4.80 (s, 2), 4.45 (d, J = 8.1 Hz, 1), 4.40−0.75 (m, m). MS m/z 1855.9 $(M+H)^+$, 928.5 $(M+2H)^{2+}$.

Construction of the Enzyme Expression Vector. Plasmid pHL1262 encoding Cel7B under control of the MET promoter, and plasmid pHL1263 encoding Cel7B:E197A were prepared from plasmid pHW704 and pCE10E197A, respectively, provided by Novozymes. A 1.3 kb Spe I to Sma I fragment encoding Cel7B or Cel7B:E197A were obtained from PCR using primers VWC9845′Cel7BHg422, 5′GCA TAC GTC *ACT AGT* ATG GCT CGC GGT ACC GCT CT3′ and VWC9863′Cel7BHg, 5′GCA TAC GTC *CCC GGG* TTA ATG GTG ATG GTG ATG GTG CTG AAC CTC CTG GTA GGT C3′. Taq polymerase was used for the PCR reaction, and the PCR cycles were as follows: step1, 94 °C, 5 min; step 2, 94 °C, 0.5 min, 50 °C, 0.5 min, 72 °C, 2 min, 25 cyles; Step 3, 72 °C, 10 min. The PCR product and plasmid p426MET25[50] (from ATCC) were double digested with Spe I and Sma I at room temperature for 4 h and then 37 °C for 4 h. The Cel7B and Cel7B:E197A coding regions were confirmed by DNA sequencing using the primers VWC1051Met25S, 5′CGT GTA ATA CAG GGT CGT C3′, VWC1052CYC1S, 5′GGG ACC TAG ACT TCA

(46) Althoff, E. A.; Cornish, V. W. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2327−2330.
(47) Adams, A.; Kaiser, C.; Gottschling, E. D.; Stearns, T. *Methods in Yeast Genetics: A Laboratory Course Manual*; Cold Spring Harbor Laboratory Press: Plainview, 1998.
(48) Ausubel, F.; Brent, R.; Kingston, R. E.; Moore, D. D.; Seidman, J. G.; Smith, J. A.; Struhl., K. *Current Protocols in Molecular Biology*; John Wiley & Sons: New York, 1995; Vol. 3.
(49) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Plainview, 1989.
(50) Mumberg, D.; Muller, R.; Funk, M. *Nucleic Acids Res*. **1994**, *22*, 5767−5768.

GGT TG3′, Cel7BForwardPrimer (provided by Novozymes), 5′GCT CTG TAC CTG TCC GAG AT3′.

Yeast Transformation. The standard LiOAc protocol was modified to improve the library transformation efficiency for V1019Y strain. A single colony from streak plate was picked and incubated in 10 mL liquid SC medium containing 2%glucose,lacking histidine and tryptophan. Then the culture was diluted with 100 mL YPD media and incubated at 30 °C for about 5 h (cell OD_{600} between 0.25 and 0.35). The cells were pelletted at room temperature using a Sorvail RT7 Plus centrifuge (2000 rpm, 3 min), washed with sterile distillated water, pelletted again at room temperature (2000 rpm, 3 min). The cell pellet was resuspended with 750 μL of 0.1 M LiOAc, incubated at 30 °C for 10 min, and then distributed to several eppendorf tubes (cells obtained from 100 mL YPD media can be used for 5−10 transformations). The cells were pelletted at room temperature using a Eppendorf 5417C centrifuge (14 000 rpm, 10 s). The liquid was carefully pulled out using a pipet. To the cell pellet were added the following reagents in the order indicated: (1) 170 μL 50% PEG; (2) 20 μL 1M LiOAc; (3) 10 μL of 10 mg/mL ssDNA; (4) 1−2 μL of plasmid DNA. The transformation mixture was mixed by vortexing until no cell pellet was seen, and then incubated at 30 °C for 30 min, then 42 °C for 30 min. After incubation at 42 °C, the transformation mixture was plated directly onto selective plate without pelletting cells and resuspending them in water.

Construction of Yeast Three-Hybrid Strain. The *URA3* gene of yeast strain V1016Y (similar to V784Y but with pMW112 integrated into ura3 locus only once) was replaced with the *ADE4* gene by integration.[38] The resulting strain V1019Y was the yeast three-hybrid strain used for all of the experiments in this study.

***LEU2* Growth Assay.** Strain V1019Y was transformed with either p426MET25, pHL1262, or pHL1263 using lithium acetate under standard conditions resulting in strain HL1259, HL1260, and HL1261, respectively. For each strain, at least four independent colonies were analyzed. The cells were grown in 100 μL of liquid SC medium containing 2% glucose and lacking histidine, uracil, and tryptophan in 96-well plates. After incubation in a 30 °C shaker for 24 h, 2 μL of the culture was transferred to another 96-well plate containing 100 μL of liquid SC medium containing 2% galactose, 2% raffinose, and lacking histidine, uracil, tryptophan, methionine, and leucine, with or without the small molecules Dex-Cel and Mtx-Lac-F as indicated. The plate was then incubated in a 30 °C shaker for up to 10 days. The OD_{600} of the cell culture was taken starting from day four and was measured once a day.

Mock Library Selection. A mixture of 100:1 plasmid pHL1262 to pHL1263 was transformed into V1019Y. The transformation mixture was plated onto an SC plate containing 2% glucose, lacking histidine, uracil, and tryptophan, and incubated at 30 °C for 2−3 days. About 10^4 colonies formed on the plate. Then 5 mL of sterile distilled water was added to the plate, and the plate was gently shaken by hand until all of the colonies were resuspended in the water. The cell suspension was then collected into a 15 mL Falcon tube with a sterile pipet. The plate was washed once with 5 mL of sterile distilled water to collect the remaining cells on the plate. The cell suspension was centrifuged at room temperature using a Sorvall RT7 Plus centrifuge (2000 rpm, 5 min), and the cell pellet was resuspended in 20% glycerol to an OD_{600} = 10 (measured by 10x dilution). Finally, 100 μL of this cell suspension was plated onto two SC plates containing 2% galactose, 2% raffinose, and lacking histidine, uracil, tryptophan, methionine, and leucine, with 40 μM Dex-Cel and 10 μM Mtx-Lac-F. After incubation at 30 °C for 6 days, the 5 biggest colonies were picked from each plate for colony PCR using Taq polymerase with primers VWC9845′Cel7BHg422 and VWC9863′Cel7BHg. The PCR program used are as follows: step1, 94 °C, 5 min; step 2, 94 °C, 0.5 min, 50 °C, 0.5 min, 72 °C, 2 min, 30 cycles; step 3, 72 °C, 10 min. Colony PCR was done on a 30 μL scale-10 μL was used to check the colony PCR by DNA gel electrophoresis, and the remaining 20 μL was digested with Noc I for 3 h and also checked by DNA gel electrophoresis. Since only Cel7B:E197A is digested by Nco I, which colony encodes Cel7B and which colony encodes Cel7B:E197A can be easily determined.

E197X Library Construction and Selection. The E197X library was constructed by overlap extension PCR. First, two 50-μL PCR reactions were done using the Cel7B gene as the template. The primers for the two PCR reactions are VWC9845′Cel7BHg422 and VWC1193p3Cel7BHI197 (5′GCA GCA CGA GCC CTT GCC3′), and VWC9863′Cel7BHg and VWC1236p5Cel7B197 (5′GGC AAG GGC TCG TGC TGC AAC *NNK* ATG GAT ATC TGG GAG GCC AAC TC3′), respectively. The two PCR reactions gave two products of about 600 bp which were gel purified and eluted with 50 μL of buffer EB (Qiagen). Then another 100 μL PCR reaction was done with VWC9845′Cel7BHg422 and VWC9863′Cel7BHg as the primers and 1 μL of the two PCR products as the template. The PCR gave a product of about 1300 bp which was PCR purified and eluted with 35 μL sterile distilled water. The purified PCR product was subjected to a double digest with Spe I and Sma I as described earlier and then ligated into the similarly digested p426MET25 plasmid in a 20 μL scale ligation reaction. After ligation, the ligase was inactivated by heating at 65 °C for 10 min. The DNA in the ligation mixture was then precipitated using Pellet Paint (Novagen) and redissolved in 5 μL of buffer EB and transformed to TG1 competent cells in two electroporations. Each electroporation was diluted with 2 mL LB medium, and the combined 4 mL culture was incubated at 37 °C for 1 h. Then 2 μL was taken out, diluted with 100 μL sterile distilled water, and plated onto LB/Amp plate. By counting the colonies that formed, the plasmid library size was estimated to be around 10^7. The rest of the 4 mL culture was diluted with 50 mL LB with 100 μg/mL Amp, incubated at 37 °C overnight and then subjected to plasmid maxiprep using the High-speed Maxiprep Kit (Qiagen). The DNA was eluted with 1 mL buffer EB.

To transform V1019Y, 1 to 2 μL of the maxiprep DNA was used following the transformation protocol described above. Selection was carried out essentially as for the mock library selection except that only 10 μM Dex-Cel was used. In total, the 96 largest colonies were picked and transferred to 100 μL media in 96 well plates. The cells were allowed to grow at 30 °C for 24 h and then 2 μL of the cell culture was transferred to 100 μL liquid SC medium containing 2% galactose, 2% raffinose, and lacking histidine, uracil, tryptophan, methionine, and leucine, with or without 10 μM Dex-Cel and 10 μM Mtx-Lac-F. The growth of the cells was monitored by measuring the OD_{600}. Only cells that grew faster with both small molecules were chosen for colony PCR and sequencing. The primers for colony PCR were VWC9845′Cel7BHg422 and VWC9863′Cel7BHg. The primer used for sequencing was Cel7BForwardPrimer.

Enzyme Purification. The glycosynthase Cel7B:E197A, Cel7B:E197S or Cel7B:N196D:E197A was purified from the yeast strain V1019Y carrying plasmid HL1261, HL1286, or HL1288, respectively. The cells were grown in 100 mL liquid SC medium containing 2% glucose and lacking uracil and methionine to saturation. Then the cells were diluted into 1 L liquid SC medium containing 2% glucose and lacking uracil and methionine and allowed to grow to an OD_{600} of 1. The cells were harvested by centrifugation and washed by protein extraction buffer, which contains 42.5 mM sodium phosphate buffer (pH = 7.5), 255 mM NaCl, 8.5% glycerol, 0.007% β-mercaptoethanol, 1% protease inhibitor and 2.5% 40 mM PMSF in an 2-propanol solution. The cell pellet was stored at −80 °C and was thawed at 4 °C for purification. The cells were lysed using acid-washed glass beads (vortex 7 min, 4 °C, protein extraction buffer). The cell crude extract was centrifuged at 14 000 g for 10 min, and the soluble portion was then purified under standard nondenaturing conditions using a HisTrap Kit according to the manufacturer's protocol (Amersham Bioscience). For purification of Cel7B:E197S and Cel7B:N196D/E197A, the imidazole gradient was from 10 mM to 100 mM over 20 min and then from 100 mM to 500 mM over 40 min. For purification of Cel7B:E197A, the imidazole gradient was from 10 mM to 500 mM over 60

min. The fractions containing glycosynthase protein were concentrated to 3.8 mL using a Millipore concentrator (30 000 MWCO, 15 mL). The enzyme solutions were stored at 4 °C and used directly for kinetic experiments and Western blots. The purity of each protein was judged based on Coomassie staining of a sodium dodecyl sulfate (SDS) polyacrimide gel electrophoresis (PAGE), and protein concentrations were determined by western blot by comparison to a His-tagged protein of known concentration using a standard curve.

Enzyme Characterization. The glycosynthase specific activity was measured by monitoring fluoride ion release using α-lactosyl fluoride (Lac-F) as the donor and *p*-nitrophenyl *β*-cellobioside (PNPC) as the acceptor. The activity measurement is essentially as reported.[36] Activity was assayed in falcon 15 mL tubes containing 500 μL of enzyme solution, 70 μL of 100 mM Lac-F in 150 mM sodium phosphate buffer (pH = 7), 25 μL of 400 mM PNPC in 40% (v/v) DMSO aqueous solution and 105 μL 150 mM sodium phosphate buffer (pH = 7). To measure the E197A standard specific activity, 3 μL of E197A standard from Novozymes was added to 497 μL of 230 mM imidazole sodium phosphate buffer (from HisTrap kit).

Conclusion

In conclusion, we have shown that Chemical Complementation can provide a high-throughput selection for glycosynthases, an important class of enzymes for carbohydrate synthesis. Further, this selection has been used to isolate active-site variants of the Cel7B glycosynthase from *H. insolens* with improved catalytic activity and in vivo expression levels. These results not only establish Chemical Complementation as a platform for the directed evolution of glycosynthases, but also show the generality of this approach and the ease with which it can be applied to new chemical reactions. Currently, the Chemical Complementation *LEU2* selection is being used to improve the expression of Cel7B glycosynthase in *S. cerevisiae* and further increase the activity of the Cel7B glycosynthase to make use of this enzyme for chemical synthesis practical on a preparative scale. In the longer term, directed evolution will be used to modify the substrate specificity of the glycosynthase for the synthesis of diverse oligosaccharides. We anticipate that Chemical Complementation will also find broad application in the directed evolution of other catalytic activities, particularly for reactions that are neither naturally selectable nor screenable.

Acknowledgment. The authors would like to thank Dr. J. Cherry and Dr. C. Ward from Novozymes for providing the Cel7B genes and the Cel7B:E197A glycosynthase and for sequencing the selected mutants. We also would like to thank Prof. D. Kahne, Prof. S. Walker and the members of their labs for help with the carbohydrate chemistry.We are grateful for financial support for this work from the National Science Foundation (99-84928). V.W.C. is a Sloan Fundation Fellow. H.L. was a Guthikonda Fellow (2001).

Supporting Information Available: The ^{1}H NMR spectrums of Mtx-Lac-F, Dex-Cel, and Mtx-Lac-Cel-Dex, the Coomassie stained SDS-PAGE gel of purified *Humicola insolens* Cel7B variants, the specific activity measurement for Cel7B Variants and the western blots to determine the protein concentration of Cel7B variants. This material is available free of charge via the Internet at http://pubs.acs.org.

JA046238V