Applying Pairwise Combinations of Amino Acid Mutations for Sorting Out Highly Efficient Glucosylation Tools for Chemo-Enzymatic Synthesis of Bacterial Oligosaccharides

Elise Champion,[†,‡,§,○] Frédéric Guérin,[†,‡,§,∥,⊥,○] Claire Moulis,[†,‡,§] Sophie Barbe,[†,‡,§] Thu Hoai Tran,[†,‡,§,∥,⊥] Sandrine Morel,[†,‡,§] Karine Descroix,[#,∇] Pierre Monsan,[†,‡,§] Lionel Mourey,[∥,⊥] Laurence A. Mulard,[#,∇] Samuel Tranier,[∥,⊥] Magali Remaud-Siméon,[†,‡,§] and Isabelle André*[,†,‡,§]

[†]Université de Toulouse; INSA,UPS,INP; LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France

[‡]CNRS, UMR5504, F-31400 Toulouse, France

[§]INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France

[∥]Département de Biologie Structurale et Biophysique, 205 Route de Narbonne, CNRS, IPBS (Institut de Pharmacologie et de Biologie Structurale), BP 64182, F-31077 Toulouse, France
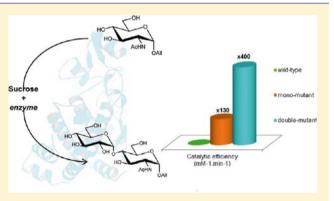
[⊥]Université de Toulouse, UPS, IPBS, F-31077 Toulouse, France

[#]Institut Pasteur, Unité de Chimie des Biomolécules, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France

[∇]CNRS UMR3523, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France

S Supporting Information

ABSTRACT: Iterative saturation mutagenesis and combinatorial active site saturation focused on vicinal amino acids were used to alter the acceptor specificity of amylosucrase from *Neisseria polysaccharea*, a sucrose-utilizing α-transglucosidase, and sort out improved variants. From the screening of three semirational sublibraries accounting in total for 20 000 variants, we report here the isolation of three double mutants of *N. polysaccharea* amylosucrase displaying a spectacular specificity enhancement toward both sucrose, the donor substrate, and the allyl 2-acetamido-2-deoxy-α-D-glucopyranoside acceptor as compared to the wild-type enzyme. Such levels of activity improvement have never been reported before for this class of carbohydrate-active enzymes. X-ray structure of the best performing enzymes supported by molecular dynamics simulations showed local rigidity of the −1 subsite as well as flexibility of loops involved in active site topology, which both account for the enhanced catalytic performances of the mutants. The study well illustrates the importance of taking into account the local conformation of catalytic residues as well as protein dynamics during the catalytic process, when designing enzyme libraries.

INTRODUCTION

Shigella, the causal agent of shigellosis, or bacillary dysentery, represents a major burden worldwide.[1] In countries where shigellosis is endemic, at least 14 distinct *Shigella flexneri* serotypes have been isolated in addition to *Shigella sonnei*, emphasizing the need for a vaccine with broad coverage.[2] The bacterial surface polysaccharides are seen as major targets of the host's protective immunity against infection, and thus as key components of *Shigella* vaccines. This observation led to the development of, among others, a number of *S. flexneri* carbohydrate conjugate vaccine candidates, either LPS-based,[3,4] or synthetic oligosaccharide-based.[5−7] In view of the need for a multivalent *Shigella* vaccine, the development of the latter approach requires a straightforward access to well-defined O-antigen (O-Ag) fragments of the prevalent *S. flexneri*

serotypes. Interestingly, most *S. flexneri* surface polysaccharides share the same linear backbone, but differ in terms of α-D-glucopyranosyl and/or *O*-acetyl decorations.[2,8] As an extension to the total chemical synthesis,[5] such structural diversity of *S. flexneri* O-Ags has led us to investigate, in recent years, the construction of a synthetic toolbox,[9−11] involving both chemical and enzymatic processes, to access chemically defined oligosaccharide fragments of selected *S. flexneri* O-Ags. Despite accomplished advancements, there is still a crucial need for appropriate biocatalysts with requisite efficiency, specificity, and stability to facilitate critical steps in chemical synthesis. Natural availability of "ideal enzymes" able to perform the desired
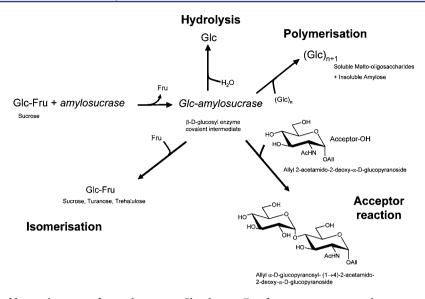
Figure 1. Reactions catalyzed by amylosucrases from sole sucrose. Glc, glucose; Fru, fructose; sucrose, α-D-glucopyranosyl-1,2-β-D-fructofuranoside; turanose, α-D-glucopyranosyl-1,3-β-D-fructose; trehalulose, α-D-glucopyranosyl-1,1-β-D-fructose.

transformation is often a limiting factor. However, with the progress in biomolecular and bioinformatics technologies, protein engineering techniques have contributed to extend the repertoire of available biocatalysts that could be useful in organic synthesis.[12,13]

As a direct continuation of the strategy earlier reported,[10] the current work aims at developing improved chemo-enzymatic routes to fragments of some *S. flexneri* O-Ags and overcome the limitations of the chemical 1,2-*cis* glucosylation developed so far. The use of α-transglucosidases, named glucansucrases,[14] was favored over Leloir-type glucosyltransferases utilizing nucleotide-activated sugar substrates as glucosyl unit donor. Glucansucrases, which belong to glycoside−hydrolase (GH) families 13 and 70,[15,16] naturally proceed by successive transfers of α-D-glucopyranosyl units from cheap and abundant sucrose to the synthesis of α-glucans linked through distinct osidic linkages depending on the enzyme regiospecificity. One major advantage of these enzymes is their natural promiscuity toward a variety of non-natural acceptors that opens access to gluco-derivatives.[11] Previous work already enabled the design of α-transglucosidases with altered specificity toward acceptor substrates. These modified enzymes were used for the synthesis of glucosylated building blocks properly protected for their subsequent chemical conversion into fragments of the *S. flexneri* 3a and 1b O-Ags. In particular, focusing on *Neisseria polysaccharea* amylosucrase (NpAS), which belongs to the GH13 family (Figure 1), a structure-based engineering approach enabled us to identify, from a small library of only 133 single mutants, new catalysts such as the F290K mutant that gained in specificity toward a non-natural acceptor, the allyl 2-acetamido-2-deoxy-α-D-glucopyranoside (**D′**) of interest in the context of *S. flexneri* 1b. However, the observed improvement of enzyme specificity was also accompanied by a decrease of catalytic efficiency as compared to that of the wild-type enzyme toward its natural substrate.[10]

Advances in enzyme engineering have demonstrated the importance of the quality of the variant library and, in particular, the efficiency of iterative saturation mutagenesis and combinatorial active site mutagenesis[17,18] to change enzyme specificity and sort out variants with desired properties from libraries of reasonable size. Here, we have investigated the

value of such an approach for carbohydrate-active enzyme engineering. In particular, we focused part of our library design on combinatorial tandem mutations targeting adjacent amino acids. Indeed, simultaneous mutations of vicinal amino acid residues are not frequently observed in nature. In contrast, they can be easily engineered *in vitro* once key positions are identified and could be of interest for a fine-tuning of unnatural substrates. The ultimate goal of this article remains the enzymatic synthesis of allyl α-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-α-D-glucopyranoside (**ED′**), a lightly protected disaccharide not found in nature, but of interest in a chemo-enzymatic synthesis of *S. flexneri* 1b and 1a oligosaccharides.[10] Key positions for **D′** recognition had been previously identified from a single mutant library targeting seven positions.[10] Here, we examined the effect of the close environment of these amino acids on substrate specificity and enzyme efficiency. To achieve this, combinatorial libraries of tandem mutations were generated and screened. Remarkable improvement of mutant properties was obtained from only one round of screening. X-ray crystallography was used to determine the structures of the best performing enzymes, while molecular dynamics simulations provided insight into their molecular motions. Together, these experimental techniques supplied new insights into structural and dynamic determinants involved in the enhanced catalytic performances of these mutants. The efficiency of the engineering approach used to rapidly sort out improved mutants and enrich our glucosylation toolbox with efficient catalysts is discussed.

■ EXPERIMENTAL METHODS

Bacterial Strains, Plasmids, and Chemicals. Plasmid pGST-AS G537D, derived from the pGEX-6P-3 (GE Healthcare Biosciences) and containing the *N. polysaccharea* amylosucrase encoding gene,[19] was used for the amylosucrase libraries construction. *E. coli* TOP 10 electrocompetent cells (Invitrogen, Carlsbad, U.S.) were used as host for each plasmid library transformation, gene expression, and production of the selected mutants. Fusion DNA-polymerase was purchased from Finnzymes (Espoo, Finland), and *Dpn*I restriction enzyme was from New England Biolabs (Beverly, MA). Oligonucleotides were synthesized by Eurogenetec (Liege, Belgium). DNA extraction (QIASpin) and purification (QIAQuick) columns were purchased from Qiagen (Chatsworth, CA). DNA sequencing was

performed by Beckman Coulter Genomics (Grenoble, France). All positive clones for **D** glucosylation were sequenced on the mutated region (~600 bp) using the primer pGEX_int: CCAACGAACAC-GAATGGGC.

Ampicillin, lysozyme, and isopropyl β-D-thiogalactopyranoside (IPTG) were purchased from Euromedex (Souffelweyersheim, France); Bromothymol Blue, sodium salt, sucrose, and *N*-acetyl-D-glucosamine (D-Glc*p*NAc) were from Sigma-Aldrich (St Louis, MO). **D′** was synthesized using known protocols.[20] The reference disaccharide α-D-Glc*p*-(1→4)-D-Glc*p*NAc (**ED**) was enzymatically synthesized and characterized at LISBP, Université de Toulouse, France.[10]

Construction, Expression, and Screening of Mutant Libraries. Amylosucrase libraries focused on positions 289–290 or 228–229 (libraries 2 and 3) were constructed by site-saturation mutagenesis using pGST-AS G537D as vector template. It was previously checked that G537D mutation had no impact on the native enzyme catalytic properties (data not shown). Three partial overlapping primer pairs surrounding double codons were designed. Each of these codons was replaced with degenerate NNS or NNW sequences, where N = A, C, G, or T; S = C or G; and W = A or T (Table S1, Supporting Information). Such degenerate primers were designed to generate 32 codons encoding the 20 possible amino acids. PCR amplifications were carried out on the whole plasmid with 1 U of Phusion DNA-polymerase for 30 cycles (98 °C, 10 s; 75 °C, 20 s; 57 °C, 15 s; 72 °C, 5 min). The DNA was digested with *Dpn*I to eliminate methylated parental template and purified using Qiaquick spin column, following manufacturer's recommendations. *E. coli* TOP10 was transformed by electroporation with 4 μL of each plasmid library using standard procedures.

Library **1** combining mutations at positions 228 and 290 was constructed by saturation mutagenesis using a two-step procedure described in previous work.[21]

After *E. coli* transformation, clones producing active amylosucrases were detected using a pH-based high-throughput screening assay on solid LB medium containing sucrose.[21] Freshly transformed cells were plated on membranes (Durapore membrane filters, 0.22 μm GV-Millipore, Ireland) previously soaked in physiological water and placed onto square plates (Corning, U.S.) containing LB agar supplemented with ampicillin (100 μg mL^{-1}) and glycerol (1% w/v). After overnight growth at 37 °C, each membrane was transferred onto another square plate containing inducing medium (LB agar, ampicillin, and IPTG (1 mM)), sucrose (50 g L^{-1}), and stained blue by addition of 50 mM Tris-HCl pH 7.5 with Bromothymol blue indicator (0.1 g L^{-1} in 1% ethanol). After 24 h of incubation at 30 °C, active clones (green and yellow) were picked and cultured in 96-well microplates containing LB medium, ampicillin, and glycerol (12% w/v), before storage (−80 °C) or further investigation. The high-throughput screening work was carried out at the Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (Toulouse, France) with the equipment of the ICEO facility, devoted to the engineering and screening of new and original enzymes.

Variants able to glucosylate D-Glc*p*NAc (**D**) were screened using the HPLC procedure previously described.[10,21] Briefly, microplates were duplicated into 96-Deep Well plates containing LB medium supplemented with ampicillin and IPTG (1 mM) to induce GST-AS expression. Cultures were then grown for 24 h at 30 °C under agitation (200 rpm), centrifuged (20 min, 3700*g*, 4 °C), and the supernatant was removed. The cell pellet was resuspended in 200 μL of lysozyme (0.5 mg mL^{-1}), followed by freezing at −80 °C for 8−12 h. After thawing at room temperature, 100 μL of sucrose and 100 μL of acceptor (each at a final concentration of 73 mM) were added to each well. Enzymatic reaction was incubated at 30 °C during 24 h under agitation. The Deep Well plates were then centrifuged (20 min, 3700*g*, 4 °C), and 300 μL of the supernatant was transferred to a filter microplate (glass fiber membrane, PS, 0.25 mm pore, Corning, U.S.) to be clarified. Supernatant filtration was carried out by centrifugation of the filter microplate (5 min, 2000*g*, 4 °C) into a microplate for HPLC screening. Acceptor reaction products were analyzed using a C18-AQ column (Bischoff C18, 125 × 4 mm, 3 μm) kept at room

temperature and eluted with 0.6 mL min^{-1} of ultrapure water to detect **ED** formation (analysis time: 7 min). Complementary HPLC analyses with a Biorad HPX-87K Carbohydrate Analysis column (maintained at 65 °C, and eluted at a flow rate of 0.6 mL min^{-1} with ultrapure water) allowed one to measure sucrose consumption by RI detection and concomitant **ED** formation by UV$_{=220 nm}$ detection. The mutant ability to synthesize the desired disaccharide **ED** was estimated as follows: % glucosyl units transferred onto acceptor derivatives = [glucosyl units transferred onto acceptor derivatives]/[glucosyl units transferred from initial sucrose].

Production, Purification, and Characterization of Improved Variants. Prior to production and purification of double mutants, their entire genes were sequenced by Cogenics (Grenoble, France) showing no other mutations than those at positions 289 and 290. The A289P–F290C, A289P–F290I, and A289P–F290L double mutants were produced and purified to the GST-AS fusion protein stage (96 kDa) as previously described.[19] The protein concentration was estimated using a Nanodrop ND-1000 spectrophotometer. Specific activities were determined at 30 °C in 50 mM Tris-HCl buffer, pH 7.0 using 250 mM sucrose alone, or with 250 mM sucrose supplemented with 250 mM of **D′** acceptor. The concentration of reducing sugars was determined using the dinitrosalicylic acid assay.[22] One unit of amylosucrase activity corresponds to the amount of enzyme that catalyzes the release of 1 μmol of reducing sugars per minute in the assay conditions.

Comparison of Products Synthesized by Wild-type Amylosucrase and Variants. Reactions were performed at 30 °C, in the presence of 250 mM sucrose alone or supplemented with 250 mM **D′** acceptor and 1 U/mL of wild-type NpAS or variants. The reactions were stopped after 24 h by heating at 95 °C for 5 min. The soluble part of the reaction mixture was submitted to HPAEC-PAD (high-performance anion-exchange chromatography with pulsed ampero-metric detection) analysis. To quantify the concentration of monosaccharides and disaccharides, the soluble fraction was diluted in water and separated on a 4 × 250 mm Dionex Carbo-pack PA100 column. A gradient of sodium acetate (from 6 to 300 mM in 28 min) in 150 mM NaOH was applied at 1 mL min^{-1} flow rate. Detection was performed using a Dionex ED40 module with a gold working electrode and an Ag/AgCl pH reference. Note that acceptor **D′**, and its derivatives, were not detected by pulsed amperometric detection system. Therefore, **D′** and their glucosylation products were quantified by HPLC with a Biorad HPX-87K column (see above). Concentration of sucrose, glucose, fructose, turanose, trehalulose, and maltose was determined by HPAEC-PAD.

In parallel, the reaction mixture containing soluble and insoluble malto-oligosaccharides was solubilized in 1 M aqueous KOH at a final total sugar concentration of 10 g L^{-1} and analyzed by HPAEC-PAD using a Dionex Carbo-Pack PA100 column at 30 °C. Mobile phase (150 mM aq NaOH) was set at 1 mL min^{-1} flow rate with a sodium acetate gradient (6−500 mM over 120 min).

Determination of Kinetic Parameters. Catalytic efficiency (k_{cat}/K_m) of wild-type NpAS and variants was determined with (i) sucrose alone and with (ii) both sucrose and acceptor **D′** as substrates. Enzyme assays were carried out in a total volume of 2 mL containing, respectively, (i) 0.073 mg of A289P–F290C, 0.153 mg of A289P–F290I, or 0.092 mg of A289P–F290L in the presence of sucrose concentration varying from 0 to 500 mM or (ii) 0.037 mg of A289P–F290C, 0.076 mg of A289P–F290I, and 0.046 mg of A289P–F290L in the presence of 250 mM sucrose and **D′** concentrations varying from 0 to 250 mM. At regular time intervals, aliquots (200 μL) were removed, heated (95 °C, 2 min), and centrifuged (18 000*g*, 5 min) and analyzed by HPAEC-PAD to monitor fructose release (see section above). The initial rate of sucrose consumption, corresponding to the initial rate of fructose release, was expressed in micromole of fructose released per minute and per gram of enzyme. For reaction with sucrose only, the kinetic parameters k_{cat} and K_m were determined using the Eadie–Hofstee plot. For reaction with sucrose and acceptor, the initial rate of formation of the desired disaccharide (**ED′**) concentration was determined by HPLC with a Biorad HPX-87K column. Saturation was not achieved for the double mutants, and the efficiency k_{cat}/K_m

Table 1. Data Collection and Refinement Statistics

	A289P−F290L (4FLR)	A289P−F290C (4FLO)	A289P−290I (4FLQ)	F290K−E328Q sucrose (4FLS)
Data Collection				
space group	$P2_12_12$			
a, b, c (Å)	96.0, 116.3, 60.2	96.0, 116.0, 60.5	96.5, 115.8, 60.8	95.1, 114.7, 54.6
α, β, γ (deg)	90.0, 90.0, 90.0			
resolution (Å)	38.78−2.40 (2.53−2.40)*	38.66−2.20 (2.32−2.20)	37.06−2.50 (2.64−2.50)	57.35−2.30 (2.42−2.30)
R_{sym}	0.098 (0.337)	0.101 (0.306)	0.096 (0.310)	0.142 (0.418)
$I/\sigma I$	6.8 (2.0)	7.2 (2.5)	7.1 (2.4)	5.1 (1.9)
completeness (%)	97.3 (94.8)	99.8 (99.1)	98.1 (99.4)	98.8 (96.7)
redundancy	3.0 (2.9)	4.7 (4.0)	2.4 (2.4)	3.3 (3.2)
Nb of molecule/au	1			
Matthews coefficient (Å³/Da)	2.3	2.3	2.3	2.3
Refinement				
resolution (Å)	12.00−2.40	14.00−2.20	12.00−2.50	12.00−2.30
no. of unique reflections	26 360 (3664)	35 028 (5007)	23 776 (3464)	26 993 (3768)
R_{work}/R_{free} (%)	18.2/23.55	16.4/21.6	17.3/21.9	17.6/22.0
total nb of atoms	5427	5679	5419	5367
Nb of protein residues	628	628	628	628
Nb of ligand molecules	1 Tris, 18 glycerols, 1 PEG	1 Tris, 3 glycerols, 1 PEG	1 Tris, 17 glycerols, 2 PEG	2 sucroses, 1 glycerol, 1 Cl⁻
Nb of water molecules	258	571	335	283
B-Factors (Å²)				
protein	22.3	16.0	22.6	17.5
ligand	33.9	29.1	35.3	18.5
water	23.1	23.1	25.4	18.8
rmsd				
bond lengths (Å)	0.006	0.007	0.006	0.006
bond angles (deg)	1.0	1.1	1.1	1.1

(D') was calculated by linear regression of the initial rate of formation of desired disaccharide (ED') versus substrate concentration.

Crystallization. A289P−F290L, A289P−F290C, and A289P−F290I mutants were crystallized using conditions previously described by Skov et al.[23,24] Crystals of F290K−E328Q mutant were obtained using the same conditions described above and in the presence of 20 mM sucrose.

Data Collection and Structure Determination. X-ray experiments were carried out at 100 K. Prior to flash cooling, native crystals of NpAS mutants were soaked for a few seconds in the reservoir solution supplemented with 20% (v/v) glycerol to avoid ice formation, except for the cocrystal of F290K−E328Q−sucrose, which was intrinsically cryo-protected by sucrose. A289P−F290L, A289−F290C, and A289P−F290I diffraction data sets were collected to a maximum resolution of 2.40, 2.20, and 2.50 Å, respectively, on beamlines ID14-1 (for A289P−F290C and A289P−F290I mutants) and ID14-2 (for A289P−F290L mutant) at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). The F290K−E328Q−sucrose complex data set was collected to 2.30 Å on the ESRF beamline ID23-1. Diffracted intensities were integrated using iMOSFLM[25] and scaled with SCALA[26] from the CCP4 software suite,[27,28] and 10% (for A289P−F290L, A289P−F290C, and A289P−F290I mutants) or 5% (for F290K−E328Q mutant in complex with sucrose) of the scaled amplitudes was randomly selected and excluded from the refinement procedure. Crystals of all amylosucrase mutants in complex or not with sucrose belong to orthorhombic space group $P2_12_12$ with one molecule per asymmetric unit giving a Matthews coefficient of 2.3 Å³/Da. Data collection statistics are given in Table 1. Structures of amylosucrase mutants were straight refined from their native structures (PDB code: 1G5A)[23,24] using refmac5.[29]

Building and Refinement. Structure refinement was performed using refmac5 from the CCP4 GUI.[27,28] Models were manually reconstructed in SigmaA weighted electron density maps using COOT.[30] Water molecules were automatically assigned, and ligand molecules were manually fitted onto residual maps. The final models of all mutants contain 628 residues out of the 632 theoretical residues

with 4 missing residues at the N-terminal extremity. Refinement statistics are given in Table 1.

Coordinates. Coordinates have been deposited at the protein data bank (PDB codes: 4FLR, 4FLO, 4FLQ, and 4FLS for A289P−F290L, A289P−F290C, A289P−F290I mutants, and F290K−E328Q−sucrose complex, respectively).

Differential Scanning Fluorimetry. The melting point (T_m) of amylosucrase mutants was assayed by differential scanning fluorimetry (DSF). A mix of enzyme (2 μM), Sypro-orange (5 X) (Invitrogen, Paisley, UK), and 50 mM Tris, pH 7.0, 150 mM NaCl, 1 mM DTT, 1 mM EDTA in the absence or in the presence of sucrose 100, 300, or 600 mM was incubated using a temperature gradient from 20 to 80 °C with a 0.3 °C increment. The thermal transition was monitored using a Q-PCR CFX96 Real-Time System (Biorad, Marnes-la-Coquette, France). T_m was given by the inflection point of the curve RFU = $f(T)$, with RFU standing for relative fluorescence unit.

Molecular Dynamics Simulations. All MD simulations were carried out using the AMBER 9 suite of programs.[31] The molecular all-atom ff03[32,33] and carbohydrate GLYCAM06[34] force fields were used for the proteins and the sucrose, respectively. The starting models were derived from the high-resolution crystal structures of NpAS (PDB code: 1G5A)[23,24] and its mutants in *apo*-form (PDB codes: 4FLO, 4FLQ, 4FLR) or in complex with the sucrose substrate (PDB code: 4FLS). To obtain a neutral charge of the simulated systems, a number of counterions were included. Each enzyme or enzyme/sucrose complex together with the counterions was solvated with TIP3P water molecules, using the rectangular parallelepiped box with a minimum distance of 0.12 nm between the solute and the simulation box edge. Preparation of simulations consisted of initial energy minimization steps (steepest descent and conjugate gradient methods). The minimization steps were then followed by a slow heating to 303 K under constant volume over a period of 100 ps. At the final required temperature (303 K), the system was equilibrated under constant volume condition over 10 ps, and then it was turned on constant pressure (1 bar) condition over 90 ps. Atomic positions of the protein backbone were first restrained using a harmonic potential during the minimization schedule. The force constant was then

progressively diminished until a final unrestrained minimization step. Harmonic constraints applied on sucrose were maintained until the end of the heating phase, and thereafter they were gradually removed along the equilibration in the isothermal−isobaric ensemble. The final production phase of simulations was then carried out for a total of 20 ns at constant temperature (303 K) and pressure (1 bar) conditions. The temperature and pressure were controlled using Langevin[35] thermostat and Berendsen barostat[36] with a collision frequency of 2 ps^{-1} and pressure relaxation time of 2 ps. Long-range electrostatic forces were handled by using the particle-mesh Ewald method.[37] The time step of the simulations was 2.0 fs, and the SHAKE algorithm was used to constrain the lengths of all chemical bonds involving hydrogen atoms to their equilibrium values.[38] The resulting trajectories were analyzed using the Ptraj module of the AMBER 9 package. The RMSD was calculated for the protein backbone atoms using least-squares fitting. Atomic positional fluctuations (Δr_i^2) of protein backbone were calculated. A mass-weighted average value was then calculated for each residue. These parameters are related to the *B*-factors through the following relationship:

$$B_i = \frac{8\pi^2}{3}\left\langle \Delta r_i^2 \right\rangle$$

The simulated *B*-factors were calculated using the coordinates of the 20 ns trajectories.

RESULTS AND DISCUSSION

Generation of Libraries Harboring Tandem Mutations. Three saturation libraries of double mutants were generated. Positions 228, 289, and 290, which were previously identified as being favorable for **D′** recognition, were selected.[10] In libraries 1 and 2, the second round of randomization was focused on positions 228/290 and 289−290, respectively. In the 3D structure of NpAS, key positions 228 and 290 were found spatially close and thus likely to influence each other. We also opted for the construction of a library focused onto amino acids 228 and 229 to target two other adjacent positions in the active site. Position 229 was preferred to position 227 due to the presence of a bulky and hydrophobic phenylalanine at position 229 that is likely to disturb **D′** accommodation. We used a NNS or NNW codon degeneracy to generate the 20 possible amino acids at each position. The three libraries were first screened using a previously developed colorimetric pH-based assay on solid medium allowing rapid determination of library size and sucrose-active clone ratio.[21] Altogether, 1140 clones were found able to utilize sucrose as a glucosyl donor, which corresponds to an active clone ratio comprised between 2% and 29%, depending on the sublibrary (Table 2). Notably, with 29% of positive clones, library 2 combined two positions that are quite tolerant to mutations.

Among active clones, 288 clones were considered as very active due to a strong pH change that induced an intense color shift from blue to yellow, whereas 852 were found moderately active (green color). Overall, the screening efforts based on

sucrose utilization allowed downsizing by ∼94% the number of mutants to be tested for their ability to glucosylate commercial *N*-acetyl-glucosamine (D-Glc*p*NAc, **D**), the detection of glucosylation product α-D-Glc*p*-(1→4)-D-Glc*p*NAc (**ED**) being carried out by HPLC. Using this procedure, 30 mutants were found to synthesize the **ED** compound with a glucosylation yield above 20%. Among them, 10 single mutants had already been characterized.[10] Therefore, we focused our work on the 20 remaining active amylosucrase double mutants that are listed in Table 3, among which only two were found

Table 3. Improved Amylosucrase Mutants Identified in the Screening for the α-D-Glc*p*-(1→4)-D-Glc*p*NAc (ED) Synthesis

library	amino acid mutation(s)	mutation codon(s)	occurrence	glucosylation yield to synthesize **ED** (%)
wild-type NpAS			1	2
1	I228V−F290V	GTC GTT	2	90
1	I228V−F290K	GTC AAA	1	89
1	I228A−F290H	GCC CAT	1	88
1	I228T−F290K	ACC AAA	1	72
1	I228V−F290H	GTC CAT	1	61
1	I228T−F290H	ACC CAT	1	56
1	I228K−F290W	AAG TGG	1	50
1	I228K−F290L	AAG CTT	1	47
1	I228V−F290R	GTC CGT	1	25
1	I228P−F290C	CCC TGT	1	22
2	A289P−F290L	CCC TTA	1	100
2	A289P−F290C	CCG TGT	2	100
2	A289P−F290I	CCC ATT	1	100
2	A289T−F290H	ACG CAT	1	40
2	A289H−F290S	CAC TCT	1	40
3	I228C−F229L	TGC CTC	1	78
3	I228V−F229L	GTG TTG	1	56
3	I228L−F229L	TTG CTC	1	35
3	I228V−F229M	GTG ATG	1	29
3	I228M−F229M	ATG ATG	1	28

twice (I228V−F290V and A289P−F290C). As compared to the corresponding single mutants,[10] most of the double mutants produced higher yields of **ED** disaccharide, thus demonstrating the relevance of iterative saturation mutagenesis strategy. Moreover, results obtained from libraries 2 and 3 also show that the double-mutation cycle approach focusing onto vicinal residues is efficient to rapidly identify beneficial coupling interactions that might correspond to optimal compensations of structural rearrangements and isolate superior mutants as compared to single ones. On the basis of our screening criteria, the most promising double mutants (A289P−F290L, A289P−F290C, and A289P−F290I) emerged from library 2.

Biochemical Characterization of Improved Amylosucrase Variants. The three most active double mutants A289P−F290L, A289P−F290C, and A289P−F290I were produced and purified to homogeneity to evaluate their steady-state kinetic properties and the product reaction profile obtained with and without the presence of **D′**. From sole sucrose (250 mM), all three double mutants synthesized malto-oligosaccharides (up to DP 20) in the same way as the wild-type enzyme and the previously characterized F290K single mutant (Figure 2A).[10]

Table 2. Results of the pH Screening of Amylosucrase Sublibraries for Sucrose Utilization

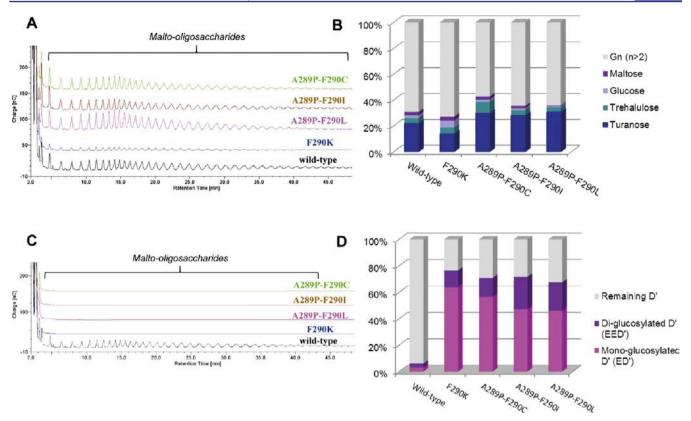|	library 1 228/290	library 2 289−290	library 3 228−229
number of clones	∼8000	∼2000	∼8000
number of isolated active clones (%)	384 (5%)	576 (29%)	180 (2%)
number of active clones (yellow)	96	96	96
number of moderately active clones (green)	288	480	84

Figure 2. Comparison of malto-oligosaccharide synthesis by wild-type amylosucrase, F290K, A289P−F290C, A289P−F290I, and A289P−F290L mutants from 250 mM sucrose supplemented or not with 250 mM allyl 2-acetamido-2-deoxy-α-D-glucopyranoside ($\mathbf{D'}$) and 1 U/mL of enzyme. (A) Superposition of the HPAEC-PAD profiles obtained at the end of the reaction (t_f = 24 h) from sole sucrose. (B) Yields of glucosyl units incorporated into the various products synthesized in the total reaction medium from sole sucrose. (C) Superposition of the HPAEC-PAD profiles obtained at the end of the reaction (t_f = 24 h) from sucrose supplemented with $\mathbf{D'}$. (D) Yields of mono- and diglucosylated $\mathbf{D'}$, $\mathbf{ED'}$ and $\mathbf{EED'}$, respectively, from sucrose supplemented with $\mathbf{D'}$. At final time, >90% of sucrose was consumed.

Table 4. Kinetic Parameters Determination of Wild-type NpAS and Improved Mutants for Sucrose Donor[a]

	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})	$V_{250\ mM}$[d] (μmol/min/g)	behavior (at high concentration)
wild-type NpAS[b]	1.3	50.2	0.0261	900	saturation
F290K[c]	n.d.	n.d.	0.0006	110	linear
A289P−F290L	2.1	10.1	0.2094	1717	saturation
A289P−F290C	2.8	18.9	0.1502	2186	saturation
A289P−F290I	1.1	17.2	0.0655	946	saturation

[a]Values are the mean of three independent measurements; error range is within 5−10%. No acceptor was added. n.d.: not determined. [b]Data from ref 46. [c]Data from ref 10. [d]Initial rate of sucrose consumption was determined at sucrose concentration of 250 mM.

In comparison to these enzymes, double mutants also produced a higher amount of sucrose isomers (trehalulose and mainly turanose) (Figure 2B). When $\mathbf{D'}$ was added to the reaction mixture, double mutants exhibited a strong propensity to catalyze $\mathbf{D'}$ glucosylation, similar to that previously described for the F290K mutant. Indeed, malto-oligosaccharide production was fully suppressed (Figure 2C), more than 90% of the glucosyl residues coming from sucrose being transferred onto $\mathbf{D'}$ to give the monoglucosylated product $\mathbf{ED'}$ and a diglucosylated molecule, the α-D-Glc$p$-(1→4)-α-D-Glc$p$-(1→4)-α-D-Glc$p$NAc-OAll ($\mathbf{EED'}$), as characterized on the basis of MS and NMR analysis (Figure 2D). Notably, A289P−F290L and A289P−F290I double mutants produced a higher amount of trisaccharide $\mathbf{EED'}$ than did A289P−F290C mutant for which the profile of products is close to that observed for the F290K single mutant.

To determine the effect of amino acid substitution on catalytic properties, we performed steady-state kinetics. Upon varying sucrose concentration from 0 to 250 mM, all three double mutants showed a standard saturation kinetic behavior, whereas the previously identified F290K mutant was unable to be saturated. As shown in Table 4, the catalytic efficiency values for all three double mutants are increased by 100−350-fold in comparison to that of the F290K mutant and are higher than that of wild-type NpAS by 2.5−8-fold. The K_m values for sucrose diminished for all three double mutants. The k_{cat} values also increased for the A289P−F290C and A289P−F290L double mutants, whereas that of A289P−F290I remained close to the value of the wild-type.

The kinetic parameters were also determined from reactions carried out in the presence of a fixed concentration of sucrose donor (250 mM) and varying $\mathbf{D'}$ acceptor concentrations. The catalytic efficiency values of the double mutants were increased

18682

dx.doi.org/10.1021/ja306845b | *J. Am. Chem. Soc.* 2012, 134, 18677−18688

by up to 395-fold by comparison to wild-type NpAS and by 3-fold with respect to the F290K single mutant (Table 5). Such

Table 5. Kinetic Parameters Determination for Wild-type NpAS and Improved Mutants for D′ Acceptor Glucosylation Using a Fixed 250 mM Sucrose Concentration[a]

	$k_{cat}/K_{m\ acc}$ (s^{-1} mM^{-1})	$V_{250\ mM}$[c] ($\mu mol/min/g$)
wild-type NpAS[b]	0.002	290
F290K[b,d]	0.265	6690
A289P−F290L	0.790	67 300
A289P−F290C	0.690	50 000
A289P−F290I	0.700	51 400

[a]Values are the mean of three independent measurements; error range is within 5−10%. [b]Data from ref 10. [c]Initial rate of acceptor consumption was determined at a concentration of 250 mM for both sucrose and D′. [d]Apparent second-order rate constant.

an improvement in catalytic properties toward unnatural substrates has never been reported before for this class of enzymes. This shows that NpAS mutants reported here are well adapted for the glucosylation of unnatural D′ monosaccharide. Moreover, the double mutants are more efficient for D′ glucosylation than the wild-type enzyme acting on its natural substrate. This demonstrates that changing vicinal positions was extremely valuable to isolate remarkably improved mutants in the second round of our iterative enzyme optimization process.

Impact of Mutations and Ligand on Protein Stability. The thermal stability of amylosucrases was assessed using a fluorescence-based approach, the differential scanning fluorimetry (DSF), in the absence and in the presence of a bound ligand. Purified proteins were subjected to gradually increasing temperatures to determine the melting temperature (T_m). The denaturation process was monitored by measuring the increase in the fluorescence of a specific, environmentally sensitive, fluorescence dye (SYPRO orange).[39]

The average T_m values determined for the wild-type NpAS and the A289P−F290L, A289P−F290C, and A289P−F290I mutants are listed in Table 6. Measurements were highly reproducible (Figure S1A, Supporting Information). For comparison purposes, the T_m of the earlier reported F290K mutant[10] as well as that of A289P single mutant were measured. We also determined the T_m values in the presence of sucrose substrate at varying concentrations. All enzymes under investigation were first inactivated by introduction of the E328Q mutation to prevent enzymatic cleavage of sucrose. The low impact of E328Q mutation introduction on the T_m value of the enzyme was checked for all mutants (Table 6). The T_m values for E328Q-AS in the presence of 100, 300, and 600 mM sucrose are provided in Table 6 (Figure S1B, Supporting Information).

From these data, we can conclude that all mutations clearly diminished the thermal stability of the enzymes. Comparison between T_m values of the single mutants at positions 289 and 290 also indicates that the mutation at position 290 has a more drastic impact. Replacement of Phe290 at this position is also likely to be responsible for the improved recognition of D′. We have seen that these three double mutants were all more efficient than the wild-type enzyme for catalyzing the natural reaction or the unnatural glycosylation of D′. The increase of catalytic efficiency values could well be due to their highest flexibility (correlated with lower T_m values) and their increased affinity for sucrose (as indicated by lower $K_{m,suc}$, Table 4). Also, the replacement of Ala289 by a proline residue contributes to improve the catalytic efficiency of the double mutants. Indeed, F290K single mutant is well designed for the glucosylation of the D′ acceptor. However, the initial velocity ($V_{250\ mM}$) for transglucosylation is 10-fold lower than that of the double mutants. In addition, this mutant is way less efficient than the double mutants when acting on sucrose alone. Similarly, sucrose consumption by F290C, F290I, and F290L single mutants was not detectable in the absence of D′ (data not shown). Consequently, the replacement of Ala289 by a proline residue appears to restore a good recognition of sucrose substrate without affecting D′ glucosylation, thereby contributing significantly to the improved catalytic properties of the double mutants. To support these observations by molecular interpretation, X-ray crystallographic analyses were performed.

Structural Insight on Improved Acceptor Specificity. The double mutants were crystallized in conditions very close to those used to crystallize wild-type NpAS.[23,24] The structures of A289P−F290L, A289P−F290C, and A289P−F290I mutants were solved at 2.40, 2.20, and 2.50 Å, respectively (Table 1). Our attempts to crystallize the F290K mutant failed, but inactive F290K−E328Q mutant (modified at amino acid position playing the role of general acid/base) was successfully cocrystallized in the presence of sucrose and its structure solved at 2.30 Å. For each mutant, the electron density map over the mutated residues is quite well-defined and allowed an unambiguous reconstruction (Figure 3B).

All mutants show a tertiary structure highly similar to that observed for wild-type NpAS with an overall rmsd between the C_α atoms of 0.2 Å (Figure 3A). The amylosucrases are organized in five domains, three of which are characteristic of the GH family 13: the $(\beta/\alpha)_8$-barrel folded catalytic domain A, the domain B, which is an extension of domain A between β-strand 3 and α-helix 3, and the C-terminal Greek key. Beside these common features, the three-dimensional architecture also contains two additional domains, the N-terminal α-helical domain and a B′-domain between β-strand 7 and α-helix 7 of the catalytic barrel. Positions 289 and 290 are carried by loop 4 (residues 285−303) of the catalytic barrel and closely

Table 6. Average T_m of Amylosucrases Determined by Differential Scanning Fluorimetry

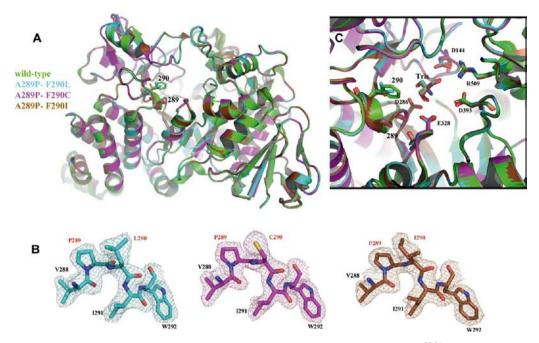|	T_m (°C) of active enzymes	T_m (°C) for inactive enzymes (containing the E328Q mutation)			
	0 mM sucrose	0 mM sucrose	100 mM sucrose	300 mM sucrose	600 mM sucrose
wild-type NpAS	50.4 ± 0.2	49.0 ± 0.2	49.8 ± 0.2	50.3 ± 0.2	52.2 ± 0.4
F290K	45.1 ± 0.2	44.3 ± 0.1	45.5 ± 0.3	45.6 ± 0.2	46.6 ± 0.4
A289P	47.7 ± 0.1	n.d.	n.d.	n.d.	n.d.
A289P−F290L	45.5 ± 0.2	45.4 ± 0.6	47.0 ± 0.8	47.0 ± 1.2	48.8 ± 0.6
A289P−F290C	45.7 ± 0.2	45.6 ± 0.3	46.6 ± 0.2	47.4 ± 0.6	48.6 ± 0.8
A289P−F290I	44.5 ± 0.2	44.2 ± 0.3	44.8 ± 0.5	46.0 ± 0.7	45.9 ± 0.8

Figure 3. Comparison of the three-dimensional architecture of wild-type NpAS in green (PDB code: 1G5A)[23,24] and NpAS mutants: A289P−F290L in cyan (PDB code: 4FLR), A289P−F290C in magenta (PDB code: 4FLO), and A289P−F290I in brown (PDB code: 4FLQ). (A) Overall organization with mutated residues in stick. (B) Zoom of the mutated regions with a SigmaA weighted $2F_o − F_c$ electron density map contoured at 1.0 σ around the residues. (C) Superimposition of active site with mutated residues in positions 289 and 290; catalytic residues D286, E328, D393; salt bridge residues D144, R509; and the Tris ligand.
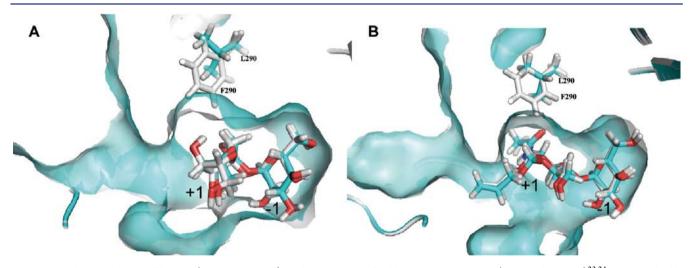

Figure 4. Molecular docking of ligands (sucrose and **ED′**) in the active site of wild-type NpAS in white (PDB code: 1G5A)[23,24] and one double mutant chosen for illustration, A289P−F290L colored in cyan (PDB code: 4FLR). Solvent-accessible surfaces were represented with the same color code to illustrate the volume modification of the active site pocket at subsites +1 and −1. A zoom of the active site is shown. (A) Docking of sucrose in the active site of amylosucrases. (B) Docking of **ED′** in the active site of amylosucrases.

interacting with loop 3 (residues 224−236). Loops 3 and 4 together with loop 7 contribute to the definition of the active site topology.

As was observed in the case of wild-type NpAS, the double mutant structures revealed the presence of a Tris molecule in the active site (Figure 3C). We believe that the presence of a bound ligand might stabilize the loops covering the catalytic site leading to a "closed" conformation of the enzyme. The F290K−E328Q mutant in complex with sucrose revealed two binding sites of sucrose, which are unambiguously defined in the electron density map. These sites were previously described as SB1 for NpAS active site (PDB code: 1JGI)[41] and SB2 at the surface of domain B (PDB code: 1MW1)[40] (Figure S2A,

Supporting Information). SB1 was compared to its equivalent in the inactive E328Q NpAS−sucrose complex (Figure S2B, Supporting Information). Except for a slight reorientation of the Q328 side chain, no significant difference was observed in the sucrose binding mode (a listing of interactions is provided in Table S2, Supporting Information), which involved interactions with the catalytic residues (D286, E328Q, D393), the R509 and D144 amino acids involved in a salt bridge that delineates the bottom of the catalytic pocket, the stabilizing histidines (H187 and H392), which are highly conserved among GH13 family, and other residues (Y147, E284, D394, and R446). Interestingly, close inspection of the electron density map revealed a strong positive residual density located

around a water molecule with a very low B-factor value (peak height of 5σ above the water molecule density) and near a hydroxyl group of the fructose moiety (O1'). Distances between this positive residual density and O1', wat817, wat1075, wat1334, the I330 main chain, and the K290 side chain are 3.1, 3.1, 3.4, 3.0, 3.1, and 4.1 Å, respectively. Such hexa-coordination together with the vicinity of K290 might be compatible with the presence of a chloride ion instead of a water molecule at that position (Figure S3, Supporting Information). Of note, a water molecule (wat1273) has been observed at the same position in the E328Q–sucrose complex (PDB code: 1JGI).[41] Introduction of a positively charged residue at position 290 (F290K mutation) in the amylosucrase active site might thus contribute to stabilize the binding of a Cl$^-$ ion. The chloride ion is known to be an activator of some α-amylases.[42] In α-amylases, chloride is found at an equivalent subsite as in the F290K mutant but not at the same location, close to an arginine residue, which corresponds to Y388 in NpAS.

Regarding the SB2 sucrose-binding site, no significant differences were observed between F290K–E328Q–sucrose and wild-type NpAS–sucrose (PDB code: 1MW1)[40] complexes (Figure S2C, Supporting Information). In both cases, interactions with sucrose involve residues D231, Q437, Y438, and S508 as summarized in Table S2, Supporting Information.

In all mutants, amino acid substitutions at positions 289 and 290 by less hindering residues neither modified side-chain orientation of neighboring residues nor induced distortion of the polypeptide main chain (Figure 3C). In addition, mutations of Phe290 by Leu, Ile, Cys, or Lys led to a slight increase of the active site molecular surface that may facilitate accommodation of the N-acetyl group of the **D'** acceptor. The replacement of Ala289 by Pro did not drastically modify the active site topology. However, it had a major impact on sucrose consumption rate as compared to that observed for single mutant F290K (see table 4). To explore further the effect of mutations, we carried out structural studies. As all attempts to cocrystallize the double mutants with sucrose, the **D'** acceptor, or products were unsuccessful, we performed molecular docking of sucrose and **ED'** in the active site of the double mutants. As shown in Figure 4A, sucrose is bound tightly in the −1 and +1 subsites defined according to the glycoside–hydrolase subsite numbering.[43] While the −1 subsite accommodates the glucosyl unit transferred during the reaction, the +1 subsite plays a dual role: it is occupied by the fructosyl unit from sucrose during the first step of the reaction, and it subsequently receives the acceptor molecule (**D'**) during the second step. Amino acid side chains at positions 289 and 290 are not directly involved in interactions with sucrose. However, they appear to be determinant in recognition of the N-acetyl group in **D'**. Mutations at positions 289 and 290 led to the uncluttering of the active site and the creation of a small lateral pocket, which favors accommodation of the N-acetyl group of **D'** as compared to wild-type NpAS (Figure 4B). Indeed, given that position 290 is able to tolerate amino acids as variable as Lys, Leu, Cys, and Ile, with an increase in **D'** glucosylation rate, one could assume that electrostatic contribution alone is not at the origin of the improvement. In comparison to the F290K mutant, the introduction of a proline residue at position 289 is likely to locally increase rigidity of loop 4 and alter dynamics of amylosucrase while modifying the active site topology. This hypothesis was investigated further by molecular dynamics simulations.

Role of Mutations in Amylosucrase Flexibility. The wild-type, single mutant F290K, and double mutants (A289P–F290L, A289P–F290C, and A289P–F290I) of NpAS were subjected to molecular dynamics (MD) simulations (20 ns) in explicit water under constant temperature and pressure conditions (Figure S4A, Supporting Information). The profiles obtained for wild-type NpAS and double mutants indicate that the backbone rmsd gradually increased within the first 5 ns of the simulation and then remained stable until the end of the simulation around an average value of 1.7 Å. Simulations of substrate-free and sucrose-bound F290K amylosucrase were also compared. Whereas a significant increase of the C$_\alpha$ rmsd is observed between 5 and 15 ns for the unliganded form, the profile corresponding to the liganded form exhibits relatively moderate fluctuations (rmsd ~1.3 Å), with no major conformational change as compared to wild-type NpAS (Figure S4B, Supporting Information).

In addition, the bound form of F290K had lower backbone rmsd values than the free form, suggesting that the stability of F290K structure is strongly enhanced upon sucrose binding in agreement with T_m values (Table 6). Of note, no dissociation of the sucrose from amylosucrase was observed within the 20 ns of the simulation.

For all unliganded amylosucrases, analysis of simulated B-factors during the course of MD simulation showed a particularly high mobility of loop 3 (residues 224−236), loop 4 (residues 285−303), loop 7 (residues 433−449), and loop 8 (residues 487−532) (Figure S4B, Supporting Information). The amplitude of the conformational changes varied depending on the enzyme. These loops play a key role in conferring the active site a pocket topology in the X-ray structure.[23,24] While loops 2 and 8 close up the bottom of the pocket through a salt bridge formed by D144 and R509, loops 3, 7, and 8 cover up the oligosaccharide binding site, shielding it from solvent, and thus promoting transglucosylation reaction over sucrose hydrolysis. Overall, MD simulations indicate that domains B (residues 180−260) and B' (residues 396−460) exhibit the highest B-factors (Figure S4B, Supporting Information), although some differences are observed depending on mutants. Introduction of a proline and leucine or lysine amino acid at positions 289 and 290, respectively, led to a drastic increase in flexibility of loops 3, 4, and 7 as compared to parental NpAS. A less pronounced effect on the flexibility of loops 3, 4, and 7 was observed when inserting mutations A289P–F290C and A289P–F290I. Of all mutants in unliganded form, F290K is the one for which domains B and domain B' were found to be of the highest flexibility. The rigidity that results from the introduction of a proline residue at position 289 may enable catalytic triad residues to be maintained in productive conformation despite the proximity of positions 289 and 290 with nucleophile D286, which are all carried by flexible loop 4. In addition, the A289P mutation enabled one to reinforce sucrose recognition as shown by the lowest K_m for sucrose in the case of the double mutants (Table 4).

Moreover, an opening of the loops was observed for all amylosucrase mutants, exposing the oligosaccharide binding subsites and reshaping the active site into a cleft topology. Interestingly, A289P–F290L mutant and wild-type NpAS also showed a disruption of the salt bridge formed between D144 and R509 after 12 ns of the simulation, displaying putative negative subsites observed in α-amylases, which are not accessible in the amylosucrase X-ray structure. Simultaneously, loops 3 and 7 moved apart by ~12 Å (Figure 5), emphasizing
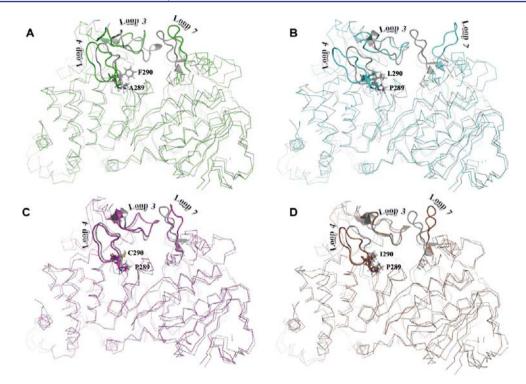
Figure 5. Superimposition of wild-type and double mutant NpAS structures at the beginning (0 ns, in gray) and the extreme conformation identified along 20 ns MD simulation. (A) Wild-type NpAS (green). (B) A289P−F290L mutant (cyan). (C) A289P−F290C mutant (magenta). (D) A289P−F290I mutant (brown).

the significant flexibility of amylosucrases in the absence of ligand. For comparison purpose, simulations were also run with the F290K mutant in complex with sucrose (Figure S4B,C, Supporting Information). When bound to sucrose, the overall flexibility of the enzyme is significantly reduced. Interestingly, external loops 3 and 7 appear strongly stabilized in the presence of sucrose, suggesting a "rigidification" of these loops composing the "roof" of the active site pocket (Figure S5, Supporting Information). On the other hand, only minor influence of sucrose binding was detected on the dynamics of the internal core. Taken together, these results suggest that the amylosucrase is more compact and stable when sucrose substrate is bound at the active site. This could explain why X-ray structures of amylosucrase from *N. polysaccharea* were always solved in "closed" conformation with a ligand molecule (Tris, glucosyl moiety, sucrose, or maltoheptaose) bound in the active site.

The hypothesis that we propose is that only in the presence of a bound molecule in the catalytic site, loops 3, 4, 7, and 8 of the $(\beta/\alpha)_8$-barrel adopt a "closed" conformation conferring a pocket topology to the active site and favoring intermolecular interactions with the ligand. By promoting sucrose and **D**′ recognition, the introduction of mutations at positions 289 and 290 from loop 4 is likely to alter significantly the conformational rearrangements involved in the "closure" of amylosucrase to adopt a catalytically productive state.

Introduction of a short side chain at position 290 led to an increased flexibility of domain B. This flexibility is reduced upon sucrose binding that involves a network of stabilizing interactions with domains B and B′. Such destabilization of the enzyme is suspected to be more important for the F290K mutant than for the double mutants that possess a proline residue at position 289, which locally rigidifies the structure. This could explain why, despite our efforts, we were unable so

far to crystallize the F290K mutant in its *apo* form. Such molecular fluctuations also led to less stable folds of double mutants as compared to wild-type NpAS, in accordance with the decrease of up to 6 °C in the thermal stability of the mutants.

■ CONCLUSIONS

By recombining pairwise mutations at amino acid positions delineating the active site cavity of amylosucrase, we improved by up to 8-fold and 400-fold the catalytic efficiency of the best performing mutants toward sucrose and allyl 2-acetamido-2-deoxy-α-D-glucopyranoside, known as the **D**′ non-natural acceptor. Such levels of activity enhancement have never been reported before for this class of enzymes. These results emphasize the high malleability of α-transglucosidases. They also demonstrate the importance of the experimental design during the construction of the combinatorial library, and especially the importance of exploring in depth the active site topology, in particular by generating vicinal mutations when positive mutants have been targeted in the first run of active site plasticity exploration. This study also demonstrates how important is the understanding of interrelationships between protein flexibility, function, and stability to improve the design of biocatalysts. Indeed, results presented herein suggest that introduction of mutations at positions 289 and 290, which are located at the rim of the substrate pocket, only induced minor structural changes in the amylosucrase 3D structure. More significant modifications induced by mutations were observed in MD simulations that revealed an alteration of the flexibility of some key structural regions that control active site topology, in particular loops 3, 4, and 7. Such conformational changes might thus favor substrate recognition by the enzyme, when in its catalytically active state, as shown by the enhanced catalytic

efficiency of amylosucrase double mutants for **D′** glucosylation. Although the wealth of biochemical and structural data provides invaluable insight into the detailed aspects of the dynamic processes and their impact on the enzymatic mechanisms of amylosucrases, i.e., substrate binding, substrate-induced conformational changes, and control of the transglucosylation/hydrolysis ratio, product release remain unstudied. Prior molecular modeling studies suggested that loop 7 of amylosucrase could play a "conformational gating" role.[44] By opening up, this loop induces an enlargement of the active site cleft that may channel a second sucrose molecule toward the catalytic site and also assist elongation of glycogen branches.[45] The information provided herein about the molecular motion/dynamics of amylosucrase, in its unliganded form or bound to substrate, will thus help to better understand the amylosucrase mechanism and enhanced specificity of mutants.

ASSOCIATED CONTENT

Supporting Information

Supplementary tables and figures showing details of library construction, structure analysis, and MD simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
isabelle.andre@insa-toulouse.fr

Author Contributions
○These authors contributed equally.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Niyogi, S. K. *J. Microbiol.* **2005**, *43*, 133−43.
(2) Foster, R. A.; Carlin, N. I.; Majcher, M.; Tabor, H.; Ng, L. K.; Widmalm, G. *Carbohydr. Res.* **2011**, *346*, 872−6.
(3) Passwell, J. H.; Harlev, E.; Ashkenazi, S.; Chu, C.; Miron, D.; Ramon, R.; Farzan, N.; Shiloach, J.; Bryla, D. A.; Majadly, F.; Roberson, R.; Robbins, J. B.; Schneerson, R. *Infect. Immun.* **2001**, *69*, 1351−7.
(4) Passwell, J. H.; Ashkenzi, S.; Banet-Levi, Y.; Ramon-Saraf, R.; Farzam, N.; Lerner-Geva, L.; Even-Nir, H.; Yerushalmi, B.; Chu, C.; Shiloach, J.; Robbins, J. B.; Schneerson, R. *Vaccine* **2010**, *28*, 2231−5.
(5) Belot, F.; Guerreiro, C.; Baleux, F.; Mulard, L. A. *Chemistry* **2005**, *11*, 1625−35.
(6) Phalipon, A.; Tanguy, M.; Grandjean, C.; Guerreiro, C.; Belot, F.; Cohen, D.; Sansonetti, P. J.; Mulard, L. A. *J. Immunol.* **2009**, *182*, 2241−7.
(7) Said Hassane, F.; Phalipon, A.; Tanguy, M.; Guerreiro, C.; Belot, F.; Frisch, B.; Mulard, L. A.; Schuber, F. *Vaccine* **2009**, *27*, 5419−26.
(8) Lindberg, A. A.; Karnell, A.; Weintraub, A. *Rev. Infect. Dis.* **1991**, *13*, S279−84.
(9) Champion, E.; André, I.; Mulard, L. A.; Monsan, P.; Remaud-Siméon, M.; Morel, S. *J. Carbohydr. Chem.* **2008**, *28*, 142−160.
(10) Champion, E.; André, I.; Moulis, C.; Boutet, J.; Descroix, K.; Morel, S.; Monsan, P.; Mulard, L. A.; Remaud-Siméon, M. *J. Am. Chem. Soc.* **2009**, *131*, 7379−89.
(11) André, I.; Potocki-Véronèse, G.; More, S.; Monsan, P.; Remaud-Siméon, M. *Top. Curr. Chem.* **2010**, *294*, 25−48.
(12) Neylon, C. *Nucleic Acids Res.* **2004**, *32*, 1448−59.
(13) Yuan, L.; Kurek, I.; English, J.; Keenan, R. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 373−92.
(14) Monchois, V.; Willemot, R. M.; Monsan, P. *FEMS Microbiol. Rev.* **1999**, *23*, 131−51.
(15) Henrissat, B.; Davies, G. *Curr. Opin. Struct. Biol.* **1997**, *7*, 637−44.
(16) Coutinho, P. M.; Henrissat, B. Carbohydrate-active enzymes: an integrated database approach. In *Recent Advances in Carbohydrate Bioengineering*; Gilbert, H. J., Davies, G., Henrissat, B., Svensson, B., Eds.; The Royal Society of Chemistry: Cambridge, 1999; pp 3−12.
(17) Reetz, M. T.; Wang, L. W.; Bocola, M. *Angew. Chem., Int. Ed.* **2006**, *45*, 1236−41.
(18) Reetz, M. T.; Carballeira, J. D. *Nat. Protoc.* **2007**, *2*, 891−903.
(19) De Montalk, G. P.; Remaud-Siméon, M.; Willemot, R. M.; Planchot, V.; Monsan, P. *J. Bacteriol.* **1999**, *181*, 375−81.
(20) Ghosh, M.; Dulina, R. G.; Kakarla, R.; Sofia, M. J. *J. Org. Chem.* **2000**, *65*, 8387−90.
(21) Champion, E.; Moulis, C.; Morel, S.; Mulard, L. A.; Monsan, P.; Remaud-Siméon, M.; André, I. *ChemCatChem* **2010**, *2*, 969−975.
(22) Sumner, J.; Howell, S. A. *J. Biol. Chem.* **1935**, *108*, 51−54.
(23) Skov, L. K.; Mirza, O.; Henriksen, A.; Potocki de Montalk, G.; Remaud-Siméon, M.; Sarçabal, P.; Willemot, R. M.; Monsan, P.; Gajhede, M. *Acta Crystallogr., Sect. D* **2000**, *56*, 203−5.
(24) Skov, L. K.; Mirza, O.; Henriksen, A.; De Montalk, G. P.; Remaud-Siméon, M.; Sarçabal, P.; Willemot, R. M.; Monsan, P.; Gajhede, M. *J. Biol. Chem.* **2001**, *276*, 25273−8.
(25) Battye, T. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. *Acta Crystallogr., Sect. D* **2011**, *67*, 271−81.
(26) Evans, P. *Acta Crystallogr., Sect. D* **2006**, *62*, 72−82.
(27) Collaborative Computational Project.. *Acta Crystallogr., Sect. D* **1994**, *50*, 760−3.
(28) Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E. *Acta Crystallogr., Sect. D* **2003**, *59*, 1131−7.
(29) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr., Sect. D* **1997**, *53*, 240−255.
(30) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D* **2004**, *60*, 2126−32.
(31) Case, D. A.; Darden, T. E.; Cheatham, I. T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *AMBER 9*; University of California: San Francisco, CA, 2006.
(32) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(33) Lee, M. C.; Duan, Y. *Proteins* **2004**, *55*, 620−34.
(34) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622−55.
(35) Pastor, R. W.; Brooks, B. R.; Szabo, A. *Mol. Phys.* **1988**, *65*, 1409−1419.
(36) Berendsen, H. J.; Postma, J. P.; Van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(37) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(38) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

(39) Niesen, F. H.; Berglund, H.; Vedadi, M. *Nat. Protoc.* **2007**, *2*, 2212−21.

(40) Skov, L. K.; Mirza, O.; Sprogoe, D.; Dar, I.; Remaud-Siméon, M.; Albenne, C.; Monsan, P.; Gajhede, M. *J. Biol. Chem.* **2002**, *277*, 47741−7.

(41) Mirza, O.; Skov, L. K.; Remaud-Siméon, M.; Potocki de Montalk, G.; Albenne, C.; Monsan, P.; Gajhede, M. *Biochemistry* **2001**, *40*, 9032−9.

(42) Aghajari, N.; Feller, G.; Gerday, C.; Haser, R. *Protein Sci.* **2002**, *11*, 1435−41.

(43) Davies, G. J.; Wilson, K. S.; Henrissat, B. *Biochem. J.* **1997**, *321*, 557−9.

(44) Cortés, J.; Siméon, T.; Remaud-Siméon, M.; Tran, V. *J. Comput. Chem.* **2004**, *25*, 956−967.

(45) Albenne, C.; Skov, L.; Tran, V.; Gajhede, M.; Monsan, P.; Remaud-Simeon, M.; André-Leroux, G. *Proteins* **2007**, *66*, 118−26.

(46) Potocki de Montalk, G.; Remaud-Siméon, M.; Willemot, R. M.; Sarcabal, P.; Planchot, V.; Monsan, P. *FEBS Lett.* **2000**, *471*, 219−23.