Probing impact of active site residue mutations on stability and activity of *Neisseria polysaccharea* amylosucrase

David Daudé,^{1,2,3} Christopher M. Topham,^{1,2,3} Magali Remaud-Siméon,^{1,2,3*} and Isabelle André^{1,2,3*}

¹Université de Toulouse; INSA, UPS, INP, LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France

²CNRS, UMR5504, F-31400 Toulouse, France

³INRA, UMR792, Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France

Received 19 July 2013; Accepted 10 September 2013

DOI: 10.1002/pro.2375

Published online 1 October 2013 proteinscience.org

Abstract: The amylosucrase from *Neisseria polysaccharea* is a transglucosidase from the GH13 family of glycoside-hydrolases that naturally catalyzes the synthesis of α -glucans from the widely available donor sucrose. Interestingly, natural molecular evolution has modeled a dense hydrogen bond network at subsite –1 responsible for the specific recognition of sucrose and conversely, it has loosened interactions at the subsite +1 creating a highly promiscuous subsite +1. The residues forming these subsites are considered to be likely involved in the activity as well as the overall stability of the enzyme. To assess their role, a structure-based approach was followed to reshape the subsite –1. A strategy based on stability change predictions, using the FoldX algorithm, was considered to identify the best candidates for site-directed mutagenesis and guide the construction of a small targeted library. A miniaturized purification protocol was developed and both mutant stability and substrate promiscuity were explored. A range of 8°C between extreme melting temperature values was observed and some variants were able to synthesize series of oligosaccharides with distributions differing from that of the parental enzyme. The crucial role of subsite –1 was thus highlighted and the biocatalysts generated can now be considered as starting points for further engineering purposes.

Keywords: amylosucrase; GH 13 family; enzyme engineering; enzyme thermostability; structure-based approach; DSF

Introduction

Stability of enzymes has been widely studied during the past decades and in particular its role in enzyme

evolution. Stability changes appear to be involved in protein fitness for either increasing their robustness or altering their fold.^{1–3} In particular, molecular engineering has been extensively considered for optimizing the thermodynamic properties of enzymes.^{4–6} Mutations associated with the apparition of new functions are often detrimental to stability and may lead to poorly functional proteins.⁷ Conversely, mutations enhancing protein stability might cause, above a certain threshold, excess of stability which

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: French Ministry of Research.

*Correspondence to: Isabelle André, Université de Toulouse; INSA, UPS, I NP, LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France. E-mail: isabelle.andre@insa-toulouse.fr; Magali Remaud-Siméon, E-mail: magali.remaud@insa-toulouse.fr

can impair their further evolvability. Nonetheless, mutations or mechanisms that moderately increase enzyme robustness are of prime interest as they could, by buffering damaging effects, increase protein tolerance to additional structural modifications.⁸ Moreover, the localization of mutations also appears critical as buried residue substitutions can have more drastic effects on protein folding and activity than mutations occurring at surface residues whose effects can be more subtle and often considered neutral.⁹ Different strategies have been successfully applied to increase the stability of enzymes based on the introduction of mutations enabling to go back to stabilizing ancestors or consensus mutations,^{10–14} iteratively targeting high B-factor residues to decrease protein flexibility,^{15,16} enabling compensatory effect,¹⁷ or favoring cooperative effects by performing double-mutation cycles.¹⁸ Structure-based strategies for enhancing enzyme stability usually aim at improving packing of the hydrophobic core, extending network of salt bridges and hydrogen bonding interactions that will favourably enhance the enthalpic contribution or reduce the entropy of the unfolded state through salt or disulphide bridges, glycine substitutions, and introduction of prolines in turns.^{19,20} Nevertheless, the mechanisms that relate mutations, stability, fitness and evolvability remain poorly understood. Both experimental and computational tools have been developed over the years to help their elucidation. Among experimental techniques available to investigate protein stability, Differential Scanning Fluorimetry (DSF) has been used to easily and rapidly screen determinants of protein stability and to search for stable enzymes.^{21,22} On a theoretical level, predictive algorithms such as FoldX have been validated to provide $\Delta\Delta G$ estimations with increasing accuracy (i.e. discrepancy calculations in free energy of unfolding between a reference protein and its variants) that can be used to anticipate the effect of mutations on protein stability.^{23,24} Combined with a fast stability evaluation of experimentally generated variants, such predictive algorithms may offer opportune ways to sound out positions of interest in order to explore queer evolution pathways and isolate variants with improved properties. Using this approach, we investigated here the impact of amino acid mutations in the deep and buried active site of the amylosucrase from *Neisseria polysaccharea* (*NpAS*), an α -transglycosylase from the Glycoside-Hydrolase (GH) family 13 according to the Carbohydrate-Active enZYme classification,²⁵ onto its stability and activity.²⁶

NpAS has been described to catalyze the synthesis of an amylose-like polymer using as donor sub-

strate, the widespread sucrose (α -D-glucopyranosyl-1,2- β -D-fructofuranoside) which binds in subsites named -1 and $+1$, located at the bottom of the deep and narrow catalytic pocket. In particular, it has been shown that *NpAS* is highly specific for its natural substrate and that subsite -1 (according to GH nomenclature)²⁷ plays a major role in the recognition of the sucrose glucosyl moiety through a highly efficient hydrogen bonding interaction network (Fig. 1).^{28,29}

In addition, a recent report revealed that wild-type *NpAS* is incapable of using sucrose analogs modified on their glucosyl ring as donor substrate.²⁹ This specificity is so pronounced that simple epimerization of one hydroxyl group of sucrose leads to a total loss of enzymatic activity. Molecular evolution has tailored a highly specific sucrose active site whose residues are highly conserved among the entire GH13_4 subfamily of the CAZy classification containing amylosucrases and sucrose-hydrolases. Substituting amino acid residues constituting subsite -1 has revealed the extreme sensitivity of these positions and their detrimental impact on activity.²⁸ These residues from subsite -1 were thus also assumed to be involved in both sucrose recognition and stabilization of the enzyme overall fold.^{30,31} Nevertheless, deciphering inter-relationships between structure, activity, and stability of *NpAS* and extend further its promiscuity toward artificial donor substrates to enlarge accessible glycodiversity is of prime interest, as first introduced by Buchholz and co-workers.^{32–35} In continuation of our efforts, computational and experimental investigations were coupled in a global approach to probe the effect of amino acid mutations at stringent positions of subsite -1 , on the stability, substrate and product specificity as well as activity of *NpAS*. Altogether these data contribute to emphasize the essential role of subsite -1 for protein fitness and robustness. We also developed a simple screening procedure to explore the stability landscape of GST-tagged amylosucrases that allowed the isolation of several *NpAS* active single variants displaying enhanced stability or producing soluble gluco-oligosaccharides of varying and controlled chain lengths compared with the parental wild-type enzyme.

Results and Discussion

Stability prediction of mutants targeting subsite -1

We have recently demonstrated that *Neisseria polysaccharea* amylosucrase is highly specific for its natural substrate, sucrose, and that such specificity principally stems from the high affinity of subsite -1 for the glucosyl moiety of sucrose.²⁹

A sequence analysis of 66 sequences from the subfamily GH13_4, which contains amylosucrases

Figure 1. *NpAS* active-site residues. Amino acids from subsites -1 and $+1$ are emphasized in blue and catalytic residues are in red. (A) 2D projection of E328Q inactive variant of *NpAS* (pdb: 1JGI) in complex with sucrose (B) 3D representation of wild-type *NpAS* residues involved in substrate recognition (pdb: 1G5A) with sucrose (extracted from pdb: 1JGI). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

and sucrose-hydrolases,³⁶ showed that most residues were globally conserved and that residues of subsite -1 are highly or strictly conserved [Fig. 2(A,D)], what underlines their essential role in both fold and activity of this class of enzymes. In comparison, the analysis of the entire GH13 characterized sequences shows a poor global residue conservation. This variability probably arises from the polyspecificities gathered within the GH13 family and the variety of producing organisms.^{36,37} Nonetheless, the catalytic triad residues from subsite -1 are strictly invariant together with two histidines and one arginine residues from this same subsite which are also highly conserved in the alignment [Fig. 2(B,E)]. In complement, signature regions of high sequence similarity were previously identified.^{30,38–40} Furthermore, analysis of crystallographic B-factor values of *NpAS* showed that residues from subsite -1 display low B-factor values compared to the other residues of the protein, suggesting that they could contribute to the stabilization of the overall fold of *NpAS* [Fig. 2(C)].

To evaluate the contribution of each amino acid forming the subsite -1 to the overall protein thermodynamic stability, we used the FoldX algorithm.^{24,41} *In silico* mutagenesis of seven selected residues of the subsite -1 involved in substrate recognition (D144, Y147, H187, F250, R284, H392, and R509), though not strictly essential to catalysis, was applied in order to substitute individually each residue by the 19 possible amino acids. For each single mutant, the corresponding free energy of unfolding was calculated with respect to the wild-type value to give $\Delta\Delta G$ (kcal mol⁻¹) ($\Delta G_{WT} - \Delta G_{variant} = \Delta\Delta G$).

The changes in thermodynamic stability are represented in Figure 3 (and Supporting Information Fig. 1). Interestingly, two histidine residues appeared relatively tolerant to substitutions, namely H187 and H392, which have been postulated to play a role in transition state stabilization and in enzyme reaction specificity.⁴² Indeed, two H187 variants and nine H392 variants showed lower free energy values than that calculated for the wild-type enzyme. Conversely, introduction of mutations at D144, Y147, F250, R284, and R509 positions led to equivalent or impaired stability compared with the wild-type *NpAS*.

Saturation mutagenesis of amino acid residues from subsite -1 and their analysis

Based on stability prediction results and to evaluate the impact of amino acid substitutions at subsite -1 , 14 variants that were characterized in a prior study, although predicted with lower T_m values than the wild-type *NpAS*, were investigated together with five available variants targeting positions located in the vicinity of the subsite.²⁸ These additional variants were namely D144A, D144E, D144I, Y147A, Y147F, Y147N, R284D, R284H, R284K, R284V, F250A, F250N, F250Y, R446A, R446E, R446F, D507A, D507I, and R509Q. Two additional amino acid residues from subsite -1 that were also predicted to impact significantly enzyme stability, namely histidines H187 and H392, were targeted by saturation mutagenesis to create two libraries of 19 single variants. Altogether, 57 variants were considered.

Figure 2. Analysis of *NpAS* residues. (A) and (B): Colored representation of residue conservation in GH13_4 subfamily and GH13 family respectively, based on Shannon Entropy calculations projected onto the three-dimensional structure of *NpAS* (PDB code: 1G5A). Conserved residues are colored in blue and variable residues are in red. Residues forming subsite -1 are represented in surface. (C): *NpAS* structure (PDB code: 1G5A) colored with respect to B-factor values of the amino acid residues. Residues were gradually colored from blue to red, with those having low B-factors in blue and those with high B-factors in red. (D) and (E): sequence logos of residue conservation within GH13_4 subfamily and GH13 family, respectively. Residues from subsite -1 and +1 are emphasized in blue and catalytic residues are shown in red. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

To experimentally assay the stability of these *NpAS* variants, we undertook the development of a miniaturized purification screening procedure to avoid any noise background due to the host cell enzymatic machinery. We successfully adapted the protocol developed by Emond et al.⁴³ to enable the high-throughput

purification of the crude enzymatic extracts by using pre-packed disposable 96-well filter plates of glutathione S-transferase (GST) tagged proteins. An overview of the general protocol is shown in Figure 4.

All 57 amylosucrase variants-GST fusion proteins were purified to homogeneity (Supporting

Figure 3. $\Delta\Delta G$ (kcal mol⁻¹) values calculated using the FoldX predictive algorithm for single mutants of *NpAS* focused on seven selected positions from subsite -1. A color code from red to green was used to represent detrimental and beneficial substitutions respectively. As the error margin of FoldX is approximately 0.5 kcal mol⁻¹ changes in that range were not considered. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 4. High-throughput screening procedure used to explore the thermodynamic effects of *NpAS* mutations. A: Identification of target amino acid residues for mutagenesis and library construction. B: Miniaturized purification procedure and DSF assay. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 5. Analysis of purified single variants of *NpAS*. A: Experimental T_m values determined by Differential Scanning Fluorometry. Variants colored in green are more stable than wild-type enzyme ($>T_{m,WT} + 0.6^\circ\text{C}$), variants in red are less stable than wild-type ($<T_{m,WT} - 0.6^\circ\text{C}$) and variants in gray are equivalent to the wild-type enzyme ($=T_{m,WT} \pm 0.6^\circ\text{C}$). B: Relative activities of variants towards sucrose (gray) and *p*-nitrophenyl- α -D-glucopyranoside (blue) normalized with respect to the wild-type value. The measurements were performed in triplicates except for the wild-type for which number of single measurements is specified between brackets. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Information Fig. 1). With the exception of H187E, H187P, H392D, and H392R, all remaining mutants were considered suitable in terms of both purity and concentration to perform DSF experiments.

Experimental stability and activity assays towards glycosyl donors

Differential Scanning Fluorometry experiments were carried out to experimentally measure the T_m values of *NpAS* mutants and thus, evaluate their stability. In overall, the experimentally determined T_m values of the variants are consistent with computational predictions of $\Delta\Delta G$ values as they show that amino acid substitutions of residue H392 are mainly beneficial for protein stability (up to 2°C increase) whereas most mutations of H187 residue decrease significantly (down to 6°C) the T_m values of the variants

[Fig. 5(A)]. The screening method used herein appears thus appropriate to explore the thermodynamic stability effects of amino acid mutations and to identify amino acid residues involved in *NpAS* stability. Besides, results obtained for H187 and H392 residues, experimental data on the seven additional residues considered in this study also appeared in agreement with the stability predictions. Only the stability effects of R284 substitutions were found globally higher than expected from $\Delta\Delta G$ predictions. The activity of the variants was subsequently assayed to evaluate whether the variants were still active on the natural substrate sucrose [1] or on a previously identified alternate substrate, namely the *p*-nitrophenyl- α -D-glucopyranoside [2] [Fig. 5(B)]. Interestingly, about half of the variants still exhibited an activity towards sucrose, underlining a certain tolerance to single mutations at

Figure 6. Chain length distribution analysis of gluco-oligosaccharides synthesized using crude enzymatic extracts and sucrose as glucosyl donor (146 mM) analyzed by HPAEC-PAD. A: Products formed by seven variants were investigated compared to wild-type *NpAS* enzyme (WT). B: Difference plots between H392P and WT profiles (purple) as well as Y147F and WT profiles (red) revealed original product distributions. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

positions 187 and 392. The most active variants on sucrose also appeared to be more stable than the wild-type enzyme. Moreover, most variants displaying a lower T_m value were found to be poorly active or inactive on sucrose. The promiscuous activity of the H392 variants on *p*-nitrophenyl- α -D-glucopyranoside was found to be drastically affected compared to the activity on sucrose whereas the promiscuous activity of H187 variants decreased in the same order of magnitude than their activity on sucrose. Thus, by increasing the stability of the enzyme, the specificity towards the natural substrate sucrose was enhanced at the cost of a side activity. Conversely, a diminution of the stability resulted either in a decrease of the overall enzymatic activity or in the full loss of sucrose activity in favor of a side activity towards *p*-nitrophenyl- α -D-glucopyranoside as observed for H187L or H187Q mutants. Altogether, these results highlight that stability is closely related with both activity and substrate specificity changes. Role of equivalent residues to these histidines was previously investigated by mutagenesis in other GH13 family members.^{42,44–46} In cyclodextrin glucanotransferase (CGTase) from alkalophilic *Bacillus* sp. 1011, these residues were shown to be involved in pH dependence and transition state

stabilization,⁴² whereas in CGTase from *Thermoaerobacterium thermosulfurigenes*, their mutation into alanine changed the reaction specificity.⁴⁵

Activity results show that some variants are still able to hydrolyze both substrates. To verify if *NpAS* variants also retained their transglucosylation activities, we checked the production of gluco-oligosaccharides using sucrose as sole glucosyl donor. The analysis of the final reaction media involving crude enzymatic extracts of all 57 mutants were analyzed by HPAEC-PAD. Seven variants (H187D, H392A, H392G, H392P, H392Q, H392S, and Y147F) showed an interesting profile of products compared to wild-type. Interestingly, all seven variants were still able to catalyze the synthesis of oligosaccharides of various polymerization degrees (DP) [Fig. 6(A)]. Two of them showed glucosylated product distributions significantly altered compared with the parental wild-type enzyme. The production of short-chain oligosaccharides (DP 7 to 10) was found enhanced using variant H392P whereas variant Y147F was found to be relevant for the synthesis of soluble gluco-oligosaccharides (<DP25) [Fig. 6(B)]. Altogether these results show that stability, activity and product formation may be altered when engineering core residues of *NpAS* subsite –1.

Figure 7. Putative glycosyl donors investigated with *NpAS*: sucrose (1), *p*-nitrophenyl- α -D-glucopyranoside (2), α -D-galactopyranosyl-1,2- β -D-fructofuranoside (3), α -D-xylopyranosyl-1,2- β -D-fructofuranoside (4), *p*-nitrophenyl- α -D-galactopyranoside (5), *p*-nitrophenyl- α -D-mannopyranoside (6), α -D-galactopyranosyl-1,6- α -D-glucopyranoside-1,2- β -D-fructofuranoside (7) and 1,6-dichloro-1,6-dideoxy- β -D-fructofuranosyl-4-chloro-4-deoxy- α -D-galactopyranoside (8). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Because subsite –1 is known to be responsible for the high affinity for sucrose of *NpAS*, we investigated the capacity of the variants to recognize non-natural substrates. We thus assayed the 57 variants towards 6 sucrose derivatives in the quest for new transglycosylation activity. The substrates retained were : α -D-galactopyranosyl-1,2- β -D-fructofuranoside [3], α -D-xylopyranosyl-1,2- β -D-fructofuranoside [4], *p*-nitrophenyl- α -D-galactopyranoside [5], *p*-nitrophenyl- α -D-mannopyranoside [6], α -D-galactopyranosyl-1,6- α -D-glucopyranoside-1,2- β -D-fructofuranoside (Raffinose) [7] and 1,6-dichloro-1,6-dideoxy- β -D-fructofuranosyl-4-chloro-4-deoxy- α -D-galactopyranoside (Sucralose) [8]. These molecules were selected in priority for their high structural similarity with sucrose (Fig. 7).⁴⁷ However, none of the tested molecule was used as substrate by any of the variants, from the less stable to the most robust. As a consequence, though impacting significantly the stability of enzymes, and the formation of products from sole sucrose, single mutations introduced at subsite –1 were not sufficient to enable a productive binding of sucrose derivatives into the active site and larger structural modifications of the active site would be needed and probably necessitating combinations of mutations of the first and second shell of amino acid residues around subsite –1.

Material and Methods

Enzyme stability prediction

Computation of stability effects was performed by calculating the $\Delta\Delta G$ value of each mutant with

respect to the wild-type enzyme using the FoldX software integrated in Yasara package.^{24,41} The structure of *Neisseria polysaccharea* amylosucrase (PDB code: 1G5A) was first optimized using the repair function of FoldX. The single mutants were generated using the mutation tool of FoldX and the thermodynamic stability changes [$\Delta\Delta G = \Delta G_{WT} - \Delta G_{variant}$ (kcal mol^{–1})] were calculated. Number of runs, temperature, pH, ionic strength, and van der Waals design were set to 3, 303 K, 7.0, 5% and 2, respectively, and neighboring amino acid residues were allowed to move upon mutation of selected position.

Bacterial strains, plasmids, and chemicals

Plasmid pGST-AS derived from the pGEX-6P-3 vector (GE Healthcare Biosciences) and containing the gene encoding for the amylosucrase from *Neisseria polysaccharea* amylosucrase (*NpAS*) was used as template for the construction of saturation mutagenesis libraries. *Escherichia coli* JM109 was used as host strain for the plasmid transformation and gene expression. Sucrose was purchased from Merck KGaA (Darmstadt, Germany). Sucralose, Raffinose, *p*-nitrophenyl- α -D-glucopyranoside, Glutathione reduced, dithiothreitol and lysozyme were purchased from Sigma-Aldrich (Saint-Louis, MO). The *p*-nitrophenyl- α -D-galactopyranoside and *p*-nitrophenyl- α -D-mannopyranoside were purchased from Carbo-synth (Newbury, Berkshire, UK). The α -D-galactopyranosyl-1,2- β -D-fructofuranoside and the α -D-xylopyranosyl-1,2- β -D-fructofuranoside were

enzymatically synthesized and characterized at LISBP, Université de Toulouse, France.²⁹ Ampicillin, isopropyl β -D-thiogalactopyranoside (IPTG) and deoxyribonuclease I were purchased from Euromedex (Souffelweyersheim, France). *Dpn*I restriction enzyme was purchased from New England Biolabs (Beverly, MA). Oligonucleotides were synthesized by Eurogentec (Liege, Belgium).

Engineering of amylosucrase

Saturation mutagenesis directed on residues 187 and 392 was carried out using NNS degenerated primers (H187NNS direct: ATC-TTC-AAC-NNS-ACC-TCC-AAC-GAA-CAC-G, H187NNS reverse: CGT-GTT-CGT-TGG-AGG-TSN-NGT-TGA-AGA-T, H392NNS direct: GTC-AAC-TAC-GTC-CGC-AGC-NNS-GAC-GAC-AT and H392NNS reverse: ATG-TCG-TCS-NNG-CTG-CGG-ACG-TAG-TTG-AC) and pGST-AS as a vector template. PCR amplifications were performed using 0.625 U of Advantage HD polymerase (Clontech, CA) for 30 cycles (98°C, 10 s; 55°C, 5 s; 72°C, 7 min). DNA was digested with *Dpn*I to remove methylated parental template. *E. coli* JM109 competent cells were then transformed with the plasmid pool and plated on LB agar supplemented with 100 μ g mL⁻¹ of ampicillin. Ninety-six transformants per library were collected. The corresponding plasmids were extracted and the amylosucrase variants encoding genes were sequenced (GATC Biotech, Germany). Missing variants were specifically constructed using the same procedure as described before. Primers ATC-TTC-AAC-GAG-ACC-TCC-AAC-GAA-CAC-G and CGT-GTT-CGT-TGG-AGG-TCT-CGT-TGA-AGA-T were used for H187E construction. Primers GTC-AAC-TAC-GTC-CGC-AGC-ATG-GAC-GAC-AT and ATG-TCG-TCC-ATG-CTG-CGG-ACG-TAG-TTG-AC were used for H392M construction.

Purification screening protocol

This procedure was established and adapted from manufacturer recommendations to enable the purification of single variants generated in microplate format. Storage microplates containing the variants were thawed and replicated to inoculate starter culture in 96-well microplates containing in each well, 200 μ L of LB medium supplemented with 100 μ g mL⁻¹ of ampicillin. After growth for 12 h at 30°C (700 rpm), plates were duplicated four times into 96-Deep Well plates containing in each well, 1 mL LB medium supplemented with 100 μ g mL⁻¹ of ampicillin. Cultures were then grown for 24 h at 23°C (700 rpm) and IPTG was added (to reach a final concentration of 1 mM) to induce Glutathion S-transferase-Amylosucrase variant fusion protein expression. After 6 to 8 h, plates were centrifuged (20 min, 3000g, 4°C) and the supernatant was removed. The cell pellet was resuspended in 200 μ L of Phosphate Buffered Saline (PBS) 2X (pH 7.3, containing lyso-

zyme 0.5 mg mL⁻¹ and 1.5 U of deoxyribonuclease I) followed by freezing at -80°C for 8 to 12 h. After thawing at room temperature, the plates were centrifuged (20 min, 3000g, 4°C) and 150 μ L of the four supernatants were transferred to a single GST Mutli-trap 4B (GE Healthcare Biosciences) 96-well plate (600 μ L of supernatant per well). Unbound samples were removed by centrifugation (4 min, 100g, 4°C). 600 μ L of PBS were applied in each well to wash bound samples, this operation was repeated three times. The final elution was performed by adding 200 μ L of Tris buffer (50 mM, pH 7.0) containing 1 mM of dithiothreitol and 15 mM of glutathione reduced. The purity of the proteins was then monitored by sodium dodecyl sulphate-PAGE and protein concentration was determined with a NanoDrop ND-1000 spectrophotometer.

Differential scanning fluorometry (DSF)

The melting points (T_m) of amylosucrase variants were assayed by Differential Scanning Fluorometry (DSF). A mix of previously described purified enzyme (19 μ L), and Sypro-orange (1 μ L, 5X final) (Invitrogen, Paisley, UK) was incubated using a temperature gradient from 25 to 80°C with a 0.3°C increment. The thermal transition was monitored with a RTQ-PCR CFX96 Real-Time System (Biorad, Marnes-la-Coquette, France). T_m was determined by the inflexion point of the curve RFU=f(t), with RFU standing for Relative Fluorescence Unit.

Differential Scanning Fluorometry experiments were run at the Integrated Screening Platform of Toulouse (PICT) located at the Institute of Pharmacology and Structural Biology (Toulouse, France).

Activity assay

Activities towards sucrose were determined by transferring 50 μ L of the purified enzymes into a microplate with 50 μ L of sucrose (final concentration 146 mM) and after incubating for 2 h at 30°C. The reaction was stopped by adding 100 μ L of DNS followed by heating at 95°C for 10 min and the concentrations of released reducing sugars were determined by measuring the absorbance at 540 nm with a Sunrise spectrophotometer. Relative activities were calculated respect to protein concentration and were normalized to the wild-type value.

Activities towards *p*-nitrophenyl- α -D-glucopyranoside were determined by transferring 50 μ L of the purified enzymes into a microplate with 50 μ L of *p*-nitrophenyl- α -D-glucopyranoside (final concentration 15 mM) after incubating for 15 h at 30°C. The conversion of *p*-nitrophenyl- α -D-glucopyranoside was determined by measuring the absorbance of *p*-nitrophenoxide released at 405 nm with a Sunrise spectrophotometer (ϵ l = $3.9 \times 10^3 \times M^{-1} \text{ cm}^{-1}$). Relative activities were calculated respect to protein concentration and were normalized to wild-type value.

The high-throughput screening work was done using equipment of the ICEO facility dedicated to the engineering and screening of new and original enzymes located at the Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (Toulouse, France).

Analysis of chain length distribution of gluco-oligosaccharides

100 μ L of crude enzymatic extracts were incubated with sucrose (final concentration 146 mM) for 24 h at 30°C. Aliquots were collected and diluted with NaOH (final concentration 1M) to avoid any product precipitation. They were subsequently analyzed by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD, Dionex ICS-2500) using a CarboPac PA-100 column (4 x 250 mm). A gradient of sodium acetate in 150 mM NaOH was applied as follows: 0–30 min, 0–300 mM; 30–123 min, 300–450 mM; 123–128 min, 450 mM.

Glycosyl donor assays

The α -D-galactopyranosyl-1,2- β -D-fructofuranoside, α -D-xylopyranosyl-1,2- β -D-fructofuranoside, *p*-nitrophenyl- α -D-glucopyranoside, *p*-nitrophenyl- α -D-galactopyranoside, *p*-nitrophenyl- α -D-mannopyranoside, α -D-galactopyranosyl-1,4- α -D-glucopyranoside-1,2- β -D-fructofuranoside (Raffinose) and 1,6-dichloro-1,6-dideoxy- β -D-fructofuranosyl-4-chloro-4-deoxy- α -D-galactopyranoside (Sucralose) were assayed as potential donor substrates. 50 μ L of previously described purified variants were added to 50 μ L of substrate to reach a concentration of 150 mM (or 15 mM for *p*NP-glycosides). The enzymatic reactions were incubated at 30°C for 24 h under agitation (450 rpm). After 24 h, the enzymes were inactivated at 95°C for 5 min and removed by centrifugation. The supernatants were transferred to 96-well microplate and the reaction media were analyzed by either High Pressure Liquid Chromatography (HPLC) using a Dionex P 680 series pump, a Shodex RI 101 series refractometer and an autosampler HTC PAL with K⁺ column (Biorad, Aminex HPX-87K, 300 $\times$ 7.8 mm) or by measuring the absorbance at 405 nm using a Sunrise microplate reader for *p*NP-glycosides.

Sequence analysis

66 sequences from GH13_4 subfamily and 683 sequences from GH13 characterized family were retrieved from the Uniprot database (www.uniprot.org) with the identifiers collected on the CAZy website (www.cazy.org). Multiple sequence alignments were performed using Clustal Omega⁴⁸ and were further used for profile-profile alignment with MUSCLE v3.7.⁴⁹

Sequence logos were generated using WebLogo 3 application (<http://weblogo.berkeley.edu/>).^{50,51} Position-dependent amino acid residue variation in

multiple sequence alignment data was analysed using the Shannon information entropy measure (H_X), calculated using in house developed SEQUESTER software. The Shannon entropy (H_X) at residue alignment position (X), corrected for the normalized frequency of residue type occurrence, is computed as:

$$H_X = \frac{\sum_{i=1}^{20} p_{(i|X)} \ln p_{(i|X)}}{\sum_{i=1}^{20} p_{(i)} \ln p_{(i)}},$$

where $p_{(i|X)}$ is the conditional probability of residue type (i) occurrence at alignment position (X), and $p_{(i)}$ is the normalized probability of residue type (i) occurrence at any position. To minimize sampling bias, normalized residue type probability values were taken as those documented by Ranganathan and co-workers,^{52,53} garnered from sequence data for all natural proteins. Values of H_X lie in the range zero to one: a zero value corresponds to a fully conserved residue position, and a value of unity represents a distribution in which each residue type has an equal chance of occurrence. SEQUESTER appends Shannon entropies at aligned residue positions in a chosen reference protein structure to an atomic coordinate data file for convenient 3D visual display.

Molecular graphics images were produced using PyMOL software (Schrödinger, LLC) and 2D active-site projection was performed with LigPlot⁺ program.⁵⁴

Conclusions

In this study, we investigated the impact of single core residue substitutions on stability and activity changes of *NpAS*, a glycoside-hydrolase from GH13 family. We developed a simple procedure to explore the thermodynamic stability of GST-fused amylosucrases. We showed that this protocol could be used to identify key residues involved in enzyme robustness and to generate more stable variants. Nine variants displayed enhanced thermostability compared to the wild-type *NpAS* (T_m values increased by up to 2°C, the mutant of larger T_m being H392P). We demonstrate that active site residues which are known to be involved in the catalysis of *NpAS* can also play a role in its overall stability. Both natural and promiscuous substrates were further investigated and several variants kept their hydrolytic activity toward either sucrose or *p*-nitrophenyl- α -D-glucopyranoside upon mutations even if this latter was globally highly affected by mutations. HPAEC-PAD analyses of reaction products using sole sucrose revealed that several variants retained their transglucosylation activity and were still able to catalyze the synthesis of maltooligosaccharides. In particular, two mutants H392P and Y147F displayed original

and controlled product distributions compared to the wild-type parental *NpAS* being more efficient for synthesizing soluble oligosaccharides. Amino acid residue substitutions in *NpAS* buried active site appeared to have an impact on protein fitness as they were found to affect stability, activity, substrate promiscuity and product distribution. Moreover, this study underlines that highly (or even strictly) conserved residues of GH13 enzymes can be replaced while maintaining an activity on their natural substrate. Such engineering strategies may be useful to develop biocatalysts with enhanced properties for biotechnological applications. No new transglycosylation activity was detected when assaying alternative substrate derivatives, however stable variants identified might be used as benchmarks for future protein engineering experiments to further enhance the catalytic properties of this enzyme.

Acknowledgments

The authors are grateful to S. Tranier in using the PICT facility for DSF experiments, N. Monties for her help in the product analysis by HPAEC-PAD as well as S. Bozonnet and S. Pizzut-Serin in using the ICEO facility for enzyme screening. The authors would like to thank A.-M. Besançon for technical assistance.

References

- Bloom JD, Silberg JJ, Wilke CO, Drummond DA, Adami C, Arnold FH (2005) Thermodynamic prediction of protein neutrality. *Proc Natl Acad Sci USA* 102:606–611.
- Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. *Proc Natl Acad Sci USA* 103:5869–5874.
- Besenmatter W, Kast P, Hilvert D (2007) Relative tolerance of mesostable and thermostable protein homologs to extensive mutation. *Proteins* 66:500–506.
- Lehmann M, Wyss M (2001) Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. *Curr Opin Biotech* 12:371–375.
- Eijsink VGH, Bjork A, Gaseidnes S, Sirevag R, Synstad B, van den Burg B, Vriend G (2004) Rational engineering of enzyme stability. *J Biotechnol* 113:105–120.
- Eijsink VGH, Gaseidnes S, Borchert TV, van den Burg B (2005) Directed evolution of enzyme stability. *Biomol Eng* 22:21–30.
- Bershtein S, Segal M, Bekerman R, Tokuriki N, Tawfik DS (2006) Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature* 444:929–932.
- Tokuriki N, Tawfik DS (2009) Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature* 459:668–U671.
- Tokuriki N, Stricher F, Schymkowitz J, Serrano L, Tawfik DS (2007) The stability effects of protein mutations appear to be universally distributed. *J Mol Biol* 369:1318–1332.
- Lehmann M, Pasamontes L, Lassen SF, Wyss M (2000) The consensus concept for thermostability engineering of proteins. *BBA-Protein Struct M* 1543:408–415.
- Bershtein S, Goldin K, Tawfik DS (2008) Intense neutral drifts yield robust and evolvable consensus proteins. *J Mol Biol* 379:1029–1044.
- Gaucher EA, Govindarajan S, Ganesh OK (2008) Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* 451:704–U702.
- Alcolombri U, Elias M, Tawfik DS (2011) Directed evolution of sulfotransferases and paraoxonases by ancestral libraries. *J Mol Biol* 411:837–853.
- Huang RQ, Hippauf F, Rohrbeck D, Haustein M, Wenke K, Feike J, Sorrelle N, Piechulla B, Barkman TJ (2012) Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates. *Proc Natl Acad Sci USA* 109:2966–2971.
- Reetz MT, Carballeira J, Vogel A (2006) Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angew Chem Int Edit* 45:7745–7751.
- Jochens H, Aerts D, Bornscheuer UT (2010) Thermostabilization of an esterase by alignment-guided focussed directed evolution. *Protein Eng Des Sel* 23:903–909.
- Wang XJ, Minasov G, Shoichet BK (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J Mol Biol* 320:85–95.
- Horovitz A (1996) Double-mutant cycles: a powerful tool for analyzing protein structure and function. *Fold Des* 1:R121–R126.
- Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* 65:1–43.
- Eijsink VG, Bjork A, Gaseidnes S, Sirevag R, Synstad B, van den Burg B, Vriend G (2004) Rational engineering of enzyme stability. *J Biotechnol* 113:105–120.
- Niesen FH, Berglund H, Vedadi M (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protoc* 2:2212–2221.
- Uniewicz KA, Ori A, Xu RY, Ahmed Y, Wilkinson MC, Fernig DG, Yates EA (2010) Differential scanning fluorimetry measurement of protein stability changes upon binding to glycosaminoglycans: a screening test for binding specificity. *Anal Chem* 82:3796–3802.
- Guerois R, Nielsen JE, Serrano L (2002) Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *J Mol Biol* 320:369–387.
- Van Durme J, Delgado J, Stricher F, Serrano L, Schymkowitz J, Rousseau F (2011) A graphical interface for the FoldX forcefield. *Bioinformatics* 27:1711–1712.
- Coutinho PM, Henrissat B (1999) Carbohydrate-active enzymes: an integrated database approach. *Recent Adv Carbohydr Bioengin*:3–12.
- Monsan P, Remaud-Simeon M, Andre I (2010) Transglucosidases as efficient tools for oligosaccharide and glucoconjugate synthesis. *Curr Opin Microbiol* 13:293–300.
- Davies GJ, Wilson KS, Henrissat B (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J* 321:557–559.
- Albenne C, Potocki de Montalk G, Monsan P, Skov L, Mirza O, Gajhede M, Remaud-Simeon M (2002) Site-directed mutagenesis of key amino acids in the active site of amylosucrase from *Neisseria polysaccharia*. *Biologia* 57:119–128.
- Daudé D, Champion E, Morel S, Guieysse D, Remaud-Siméon M, André I (2013) Probing substrate

- promiscuities of amylsucrase from *Neisseria polysaccharaea*. *ChemCatChem* 5:2288–2295.
30. Janecek S (2002) How many conserved sequence regions are there in the alpha-amylase family? *Biologia* 57:29–41.
  31. van der Veen BA, Skov LK, Potocki-Veronese G, Gajhede M, Monsan P, Remaud-Simeon M (2006) Increased amylsucrase activity and specificity, and identification of regions important for activity, specificity and stability through molecular evolution. *FEBS J* 273:673–681.
  32. Baciú LE, Jordening HJ, Seibel J, Buchholz K (2005) Investigations of the transfructosylation reaction by fructosyltransferase from *B-subtilis* NCIMB 11871 for the synthesis of the sucrose analogue galactosyl-fructoside. *J Biotechnol* 116:347–357.
  33. Seibel J, Moraru R, Gotze S (2005) Biocatalytic and chemical investigations in the synthesis of sucrose analogues. *Tetrahedron* 61:7081–7086.
  34. Seibel J, Moraru R, Gotze S, Buchholz K, Na'amnieh S, Pawlowski A, Hecht HJ (2006) Synthesis of sucrose analogues and the mechanism of action of *Bacillus subtilis* fructosyltransferase (levansucrase). *Carbohydr Res* 341:2335–2349.
  35. Zuccaro A, Gotze S, Kneip S, Dersch P, Seibel J (2008) Tailor-made fructooligosaccharides by a combination of substrate and genetic engineering. *Chembiochem* 9: 143–149.
  36. Stam MR, Danchin EGJ, Rancurel C, Coutinho PM, Henrissat B (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. *Protein Eng Des Sel* 19:555–562.
  37. Svensson B, Jensen MT, Mori H, Bak-Jensen KS, Bonsager B, Nielsen PK, Kramhoft B, Praetorius-Ibba M, Nohr J, Juge N (2002) Fascinating facets of function and structure of amylolytic enzymes of glycoside hydrolase family 13. *Biologia* 57:5–19.
  38. MacGregor EA, Janecek S, Svensson B (2001) Relationship of sequence and structure to specificity in the alpha-amylase family of enzymes. *BBA-Protein Struct* M 1546:1–20.
  39. Kumar V (2010) Identification of the conserved spatial position of key active-site atoms in glycoside hydrolase 13 family members. *Carbohydr Res* 345:1564–1569.
  40. Kumar V (2010) Analysis of the key active subsites of glycoside hydrolase 13 family members. *Carbohydr Res* 345:893–898.
  41. Krieger E, Koraimann G, Vriend G (2002) Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins* 47: 393–402.
  42. Nakamura A, Haga K, Yamane K (1993) Three histidine-residues in the active-center of cyclodextrin glucanotransferase from alkalophilic *Bacillus* Sp-1011—effects of the replacement on pH-dependence and transition-state stabilization. *Biochemistry* 32: 6624–6631.
  43. Emond S, Potocki-Veronese G, Mondon P, Bouayadi K, Kharrat H, Monsan P, Remaud-Simeon M (2007) Optimized and automated protocols for high-throughput screening of amylsucrase libraries. *J Biomol Screen* 12:715–723.
  44. Sogaard M, Kadziola A, Haser R, Svensson B (1993) Site-directed mutagenesis of histidine-93, aspartic acid-180, glutamic acid-205, histidine-290, and aspartic acid-291 at the active-site and tryptophan-279 at the raw starch binding-site in barley alpha-amylase 1. *J Biol Chem* 268:22480–22484.
  45. Godany A, Majzlova K, Horvathova V, Vidova B, Janecek S (2010) Tyrosine 39 of GH13 alpha-amylase from *Thermococcus hydrothermalis* contributes to its thermostability. *Biologia* 65:408–415.
  46. Yang HQ, Liu L, Shin HD, Chen RR, Li JH, Du GC, Chen J (2013) Structure-based engineering of histidine residues in the catalytic domain of alpha-amylase from *Bacillus subtilis* for improved protein stability and catalytic efficiency under acidic conditions. *J Biotechnol* 164:59–66.
  47. Daudé D, Remaud-Simeon M, André I (2012) Sucrose analogs: an attractive (bio)source for glycodiversification. *Nat Prod Rep* 29:945–960.
  48. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li WZ, Lopez R, McWilliam H, Remmert M, Soding J, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7.
  49. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
  50. Schneider TD, Stephens RM (1990) Sequence logos—a new way to display consensus sequences. *Nucleic Acids Res* 18:6097–6100.
  51. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14: 1188–1190.
  52. Suel GM, Lockless SW, Wall MA, Ranganathan R (2003) Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat Struct Biol* 10:59–69.
  53. Rivoire O, Ranganathan R (2008) Note 103: A summary of SCA calculations. Available at: systems.swmed.edu/Labs/rr/Note103_files/Note103_v4.pdf.
  54. Laskowski RA, Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* 51:2778–2786.