Engineering Bifidobacterium longum Endo--N-acetylgalactosaminidase for Neu5Ac2-3Ga11-3Ga1NAc reactivity on Fetuin

Hansen, Dennis K.; Hansen, Anders Lønstrup; Koivisto, Johanna M.; Shuoker, Bashar; Abou Hachem, Maher; Winther, Jakob R.; Willemoës, Martin

Published in:
Archives of Biochemistry and Biophysics

DOI:
10.1016/j.abb.2022.109280

Publication date:
2022

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Hansen, D. K., Hansen, A. L., Koivisto, J. M., Shuoker, B., Abou Hachem, M., Winther, J. R., & Willemoës, M. (2022). Engineering Bifidobacterium longum Endo--N-acetylgalactosaminidase for Neu5Ac2-3Ga11-3Ga1NAc reactivity on Fetuin. Archives of Biochemistry and Biophysics, 725, [109280]. https://doi.org/10.1016/j.abb.2022.109280
Engineering *Bifidobacterium longum* Endo-α-N-acetylgalactosaminidase for Neu5Acα2-3Galβ1-3GalNAc reactivity on Fetuin

Dennis K. Hansen a,1, Anders Lønstrup Hansen a,1, Johanna M. Koivisto a, Bashar Shuoker b,c, Maher Abou Hachem b, Jakob R. Winther b, Martin Willemoes a,b

a The Linderstrøm-Lang Centre for Protein Science, Section for Biomolecular Sciences, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200, Copenhagen, Denmark
b Section for Protein Chemistry and Enzyme Technology, Department of Biotechnology and Biomedicine, Technical University of Denmark, Building 224, DK-2800, Kgs. Lyngby, Denmark
c Biotechnology, Department of Chemistry (KILU), Lund University, Post Office Box 124, 221 00, Lund, Sweden

**ARTICLE INFO**

**Keywords:**
Substrate specificity
Enzyme design
GH101
Molecular docking
O-glycanase
Protein engineering

**ABSTRACT**

Endo-α-N-acetylgalactosaminidase from *Bifidobacterium longum* (EngBF) belongs to the glycoside hydrolase family GH101 and has a strict preference towards the mucin type glycan, Galp1-3GalNAc, which is O-linked to serine or threonine residues on glycopeptides and proteins. While other enzymes of the GH101 family exhibit a wider substrate spectrum, no GH101 member has until recently been reported to process the Neu5Acα2-3Galβ1-3GalNAc. However, work published by others (ACS Chem Biol 2021, 16, 2004–2015) during the preparation of the present manuscript demonstrated that the enzymes from several bacteria are able to hydrolyze this glycan from the fluorophore, methylumbelliferyl. Based on molecular docking using the EngBF homolog, EngSP from *Streptococcus pneumoniae*, substitution of active site amino acid residues with the potential to allow for accommodation of Neu5Acα2-3Galβ1-3GalNAc were identified. Based on this analysis, the mutant EngBF variants W750A, Q894A, K1199A, E1294A and D1295A were prepared and tested.

1. Introduction

Endo-α-N-acetylgalactosaminidase from *Bifidobacterium longum* (EngBF) has been shown to be quite specific for the O-linked Galp1-3GalNAc glycan [1–3] (Fig. 1A) as have other members of the CAZy GH101 family [4]. However, to a varying extent some enzyme family members do release alternative glycans including GalNAc, Galp1-3 (GlCNacp1-6)GalNAc, GlCNacβ1-3GalNAc from para-nitrophenol in synthetic substrates [1–5]. The main species of O-glycosylation on bovine fetuin is the sialylated glycan, Neu5Acα2-3Galβ1-3GalNAc (Fig. 1B), linked to serine or threonine residues [5]. Until recently, no GH101 family member has been shown to release glycans from fetuin without prior sialidase treatment to convert these to Galp1-3GalNAc [1–3]. However, during preparation for the present manuscript, activity towards the extended glycan, Neu5Acα2-3Galβ1-3GalNAc, was demonstrated for four endo-α-N-acetylgalactosaminidases: EngSP (*Streptococcus pneumoniae*), EngCP (*Clostridium perfringens*), EngSN (*Streptomyces natalensis*), and EngCS (*Collinsella stercoris*) utilizing a

**Abbreviations:** PmST1, *Pasteurella multocida* strain P-1059 Sialyltransferase; Nanl, Nanl sialidase from *Clostridium perfringens*; Asialo-fetuin, fetuin treated with Nanl; DMAB, para-dimethylaminobenzaldehyde; EngBF, EngSP; EngCP and EngSN and EngCS, endo-α-N-acetylgalactosaminidase from *Bifidobacterium longum*; *Streptococcus pneumoniae*, *Clostridium perfringens*; *Streptomyces natalensis* and *Collinsella stercoris*, respectively; Gal, galactose; GalNAc, N-acetylgalactosamine; Neu5Ac, N-acetylneuraminic acid; GlcNAc, N-acetylgalcosamine; TLC, thin-layer chromatography; HPEAC-PAD, High-performance Anion Exchange Chromatography coupled to Pulsed Amperometric Detection.

* Corresponding author.
E-mail address: willemoes@bio.ku.dk (M. Willemoes).
1 These authors contributed equally.

https://doi.org/10.1016/j.abb.2022.109280
Received 16 February 2022; Received in revised form 13 April 2022; Accepted 4 May 2022
Available online 21 May 2022
0003-9861/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
and the EngSP variants, Q868T and Q868G, were also able to process enzyme [6]. PNGase, which is responsible for releasing the branched and matured resolved from complexes with both substrate and product that also to e.g. peptides or peptide synthesis precursors. such an enzyme could facilitate transglycosylation by these O-glycans on removal [13]. Not only for analytical purposes could an enzyme that can is the last to follow several enzymatic steps in mucin like O-glycan times referred to as O-glycanase despite that the activity of this enzyme lidase activity of NanI and PmST1 sialidase converts Neu5Ac bond.}

synthetic substrate [6]. In addition, EngSN and EngSP wild-type enzyme and the EngSP variants, Q868T and Q868G, were also able to process fetuin with the EngSP variants showing up to four-fold increased release of the Neu5Ac02-3Gal1-3GalNAc glycan, compared to wild-type enzyme [6].

Crystal structures of EngBF [7], as well as the enzyme EngSP from two Streptococcus pneumoniae strains, R6 and TIGR4 [9], have been reported. The EngSP(TIGR4) structure [9] confirmed interactions with the substrate that were previously inferred from docking experiments [7,8,10]. In addition, detailed substrate interactions of EngSP were resolved from complexes with both substrate and product that also confirmed the role of a dual tryptophan lid closing over the active site, as proposed previously [7].

N-glycans are removed by a single enzyme, peptide:N-glycosidase or PNGase, which is responsible for releasing the branched and matured glycan [11,12]. In an analogy for O-glycans and their release from glycoproteins and -peptides, endo-x-N-acetylgalactosaminidase is sometimes referred to as O-glycanase despite that the activity of this enzyme is the last to follow several enzymatic steps in mucin like O-glycan removal [13]. Not only for analytical purposes could an enzyme that can catalyze the release of full-size O-glycans seem attractive, but also if such an enzyme could facilitate transglycosylation by these O-glycans on to e.g. peptides or peptide synthesis precursors.

The present work describes the successful mutational analysis of EngBF to generate an enzyme releasing trisaccharide Neu5Ac02-3Gal1-3GalNAc from fetuin based on structural models obtained by molecular docking of this extended glycan into the wild-type enzyme and a small set of in silico generated variants having single or double side-chain replacements. The template for the docking experiments was the crystal structures of the EngSP enzyme that allow for comparison of the glycan position in the docked complexes with that of the known crystal structure of the Galβ1-3GalNAc complex.

2. Experimental procedures

Materials. GalNAc for standard solutions and DMAB for colorimetric detection of GalNAc and for use in the enzyme assay were purchased from Carl Roth (Germany). Fetuin from fetal bovine serum and all other chemicals were from Sigma Aldrich (United States). Standards used for peak identification in product characterization by anion chromatography of saccharides released from fetuin were: Gal (Merck, Germany), GalNAc, Neu5Ac and Galβ1-3GalNAc (Carbosynth, Great Britain).

Expression plasmids, protein synthesis and purification. Expression plasmids for synthesis of the mutant variants of EngBF were prepared by the two-reaction site directed mutagenesis procedure [14] as described previously [15] using the following oligo-deoxyribonucleotides: EngBF E1294K Fw: CGGAGGGCGTTAAAGGCAAACCGCATCC.

EngBF E1294K Rv: GATGGGCTGTTGCTTAAACCGCTTCCG.

EngBF E1294A Fw: GTTCTGGAAGGGCCGTGCGCAACCCGATCTCC.

EngBF E1294A Rv: GATGGGCTGTTGCTTAAACCGCTTCCGAAACC, EngBF E1294H Fw: CGGAGGGCGTTAAAGGCAAACCGCATCC, EngBF E1294H Rv: GATGGGCTGTTGCTTAAACCGCTTCCGAAACC, EngBF E1294F M Fw: CGGAGGGCGTTAAAGGCAAACCGCATCC, and EngBF E1294F M Rv: GTGAGGGCTGTTAAGGCAAACCGCATCC.

Protein synthesis in the E. coli Rosetta (DE3) strain transformed with plasmids encoding wild-type EngBF or variants and subsequent protein purification was performed as previously described [15].

A codon optimized reading frame encoding PmST1 sialyltransferase in a truncated form lacking residues 2–25 [16] was ordered from Thermofisher (United States) for recombinant expression in E. coli. Sites for restriction endonucleases were added; Ndel, that overlap the start codon, and BamHI that immediately follows the last codon of the reading frame. The above Ndel/BamHI fragment was transferred to a pET11a derived plasmid encoding a six histidine residue tag added to the 3′-end of a reading frame inserted between the Ndel and BamHI sites of this vector. The resulting plasmid, pDKH13, was sequenced to verify the coding region of the insert. Protein synthesis and purification of PmST1 using Ni-NTA chromatography was performed as previously described for EngBF [15]. Nanl naldiase was produced and purified as previously described [15].

Molecular ligand docking. Ligand dockings were performed with Glide [17–20] accessed from in the Maestro interface [21]. Prior to Glide docking, the EngSP(TIGR4) [9] E796Q variant crystal structure (PDB ID: 5A95) containing the Galβ1-3GalNAc ligand was prepared for docking by optimizing hydrogen bond networks and subsequent energy minimization with charges calculated for pH 6. A Glide grid was calculated from the prepared structure [20].

The ligand Neu5Ac02-3Galβ1-3GalNAc was built using the carbohydrate builder at Glycam.org [22] and prepared for docking in the Maestro LigPrep wizard with charges calculated for pH 6. Ligand docking used the Glide SP algorithm [19,20] with flexible ligand sampling, in which hydrogen bond formation is rewarded by optimizing hydrogen bond networks and subsequent energy minimization with charges calculated for pH 6. A Glide grid was calculated from the prepared structure [20].

The RMSD values between the co-crystallized ligand, Galβ1-3GalNAc, and this moiety of the docked trisaccharide Neu5Ac02-3Galβ1-3GalNAc in the poses were calculated and plotted against the Glide Energy score [18,19].

The identity of the glycan released from fetuin by the EngBF variants was analyzed and established in two ways. The first method using HPEC-PAD was performed as follows, 1 mL of 10 mg/mL fetuin was incubated with 1.5 μM of the E1294K variant at 37 °C for 72 h in 50 mM Acetic acid, 100 mM NaCl, pH 4.0. After incubation, the released saccharides were isolated from enzyme and fetuin, by filtering through an Amicon Ultra-2 0.5 mL 3 K MWCO centrifugal filter (Merck, Germany). The pH of 250 μL of the filtrate was adjusted to 5.0 by addition of HCl, and subsequently added 2 μL of 200 mM CMP and 36 μg of purified PmST1 sialyltransferase and incubated for 2 h at 37 °C. PmST1 was then removed by filtering as above. The filtrate of fetuin reacted with E1294K and the sample additionally treated with PmST1 as well as the standard markers to be used for anion chromatography, were all diluted to 50 μM in 10 mM NaOH
based on the concentration determined by the Morgan Elson colorimetric assay [24,25] for the glycans (see above) or the known stock concentration for small-molecule markers. The samples and standards were then analyzed on a ICS-3000 (Dionex) chromatography system equipped with a CarboPac PA200 column (Dionex, ThermoFisher Scientific) and a gold electrode. The samples were eluted steadily using a mobile phase (0.35 mL/min) of 100 mM NaOH and a stepwise gradient of sodium acetate (from 0 to 25 min, 0–75 mM; 25–30 min, 75–400 mM). Peaks were identified using QuickPeaks in the OriginPro 9.1 suite, which was also used for visualization.

The second method using TLC of products from incubation of fetuin and asialo-fetuin with E1294K was performed as follows: Fetuin (10 mg/mL) was incubated with E1294K for 24 h at 37 °C. Upon completion, a fraction of this incubation was added Nanl sialidase, prepared as described above, to a final concentration of 2.6 μg/mL and incubated for 24 h to remove the Neu5Ac moiety from the released fetuin glycans. The saccharides from samples of enzymatically treated fetuin were isolated from protein by filtering through an Amicon Ultra-0.5 mL 3 K MWCO centrifugal filter (Merck, Germany). The filtrate was loaded on a Silica gel 60 (Merck, Germany) stationary phase by pipetting 5 times 1 mL and allowing the spot to dry between loadings. The chromatogram was visualized by submersion in a solution composed of 2 mL aniline, 10 mL phosphoric acid, 2 g diphenylamine, and acetone to a final volume of 100 mL, followed by heating at 100 °C for 10–15 min.

**Steady-state kinetic analysis.** Fetuin substrate was prepared as follows: 50 mg/mL of fetuin dissolved in 50 mM sodium acetate, 250 mM NaCl, pH 6.0 was dialyzed against the same buffer and centrifuged to recover the soluble fraction resulting in a solution with a concentration ranging from 30 to 35 mg/mL. For preparation of asialo-fetuin, the fetuin solution from above was added Nanl sialidase to a final concentration of 2.6 μg/mL and incubated for 24 h at 37 °C, after which no further increase in Galβ1-3GalNAc release upon treatment with EngBF was detected. Subsequently, the asialo-fetuin preparation was dialyzed against several changes of 50 mM sodium acetate, 50 mM NaCl, pH 4.5.

Assays of glycan hydrolysis on fetuin and asialo-fetuin-based on quantification of released GalNAc residues determined with the Morgan-Elson colorimetric assay and obtained from end points fitted to the progress curves of substrate turnover (see Supplementary Materials).

The second method using TLC of products from incubation of fetuin and asialo-fetuin based on quantification of released Neu5Ac residues determined with the Morgan-Elson colorimetric assay and obtained from end points fitted to the progress curves of substrate turnover (see Supplementary Materials).

### Table 1

| Enzyme       | k_{cat} (s^{-1}) | k_{m} (μM) | k_{cat}/K_{m} (mM^{-1} s^{-1}) | Substrate range (μM) |
|--------------|------------------|------------|--------------------------------|-----------------------|
| Wild-type    | 111 ± 5          | 163 ± 21   | 679 ± 64                       | 800–35                |
| E1294K       | 108 ± 6          | 128 ± 23   | 846 ± 109                      | 700–30                |
| E1294A       | 118 ± 3          | 133 ± 11   | 883 ± 52                       | 970–35                |
| E1294 M      | 151 ± 7          | 218 ± 28   | 693 ± 63                       | 870–35                |
| E1294H       | 119 ± 3          | 178 ± 14   | 669 ± 36                       | 910–35                |
| E1294K:Q894A | 119 ± 8          | 477 ± 79   | 250 ± 26                       | 1400–30               |
| E1294K:K1199A| 109 ± 3          | 201 ± 16   | 539 ± 32                       | 1130–30               |

| Enzyme       | pH    | k_{cat}/K_{m} (mM^{-1} s^{-1}) | Substrate range (μM) |
|--------------|-------|-------------------------------|-----------------------|
| Wild-type    | 4.5   | –                             | NA                    |
| E1294K       | –     | 0.095 ± 0.001                 | 1120–160              |
| E1294A       | –     | 0.030 ± 0.001                 | 1350–160              |
| E1294 M      | –     | 0.0176 ± 0.0008               | 850–110               |
| E1294H       | –     | 0.26 ± 0.01                   | 710–150               |
| E1294K:Q894A | –     | 0.14 ± 0.005                 | 1240–140              |
| E1294K:K1199A| –     | 0.031 ± 0.001                | 1200–200              |
| Wild-type    | 6.0   | NA                            |                       |
| E1294K       | –     | 1.14 ± 0.03                  | 490–10                |
| E1294A       | –     | 1.47 ± 0.06                  | 570–20                |
| E1294 M      | –     | 0.78 ± 0.06                  | 470–60                |
| E1294H       | –     | 1.26 ± 0.09                  | 370–20                |
| E1294K:Q894A | –     | 4.4 ± 0.13                   | 350–30                |
| E1294K:K1199A| –     | 0.76 ± 0.04                  | 350–20                |

### Table 2

| Enzyme       | pH    | k_{cat}/K_{m} (mM^{-1} s^{-1}) | Substrate range (μM) |
|--------------|-------|-------------------------------|-----------------------|
| Wild-type    | 4.5   | –                             | NA                    |
| E1294K       | –     | 0.095 ± 0.001                 | 1120–160              |
| E1294A       | –     | 0.030 ± 0.001                 | 1350–160              |
| E1294 M      | –     | 0.0176 ± 0.0008               | 850–110               |
| E1294H       | –     | 0.26 ± 0.01                   | 710–150               |
| E1294K:Q894A | –     | 0.14 ± 0.005                 | 1240–140              |
| E1294K:K1199A| –     | 0.031 ± 0.001                | 1200–200              |
| Wild-type    | 6.0   | NA                            |                       |
| E1294K       | –     | 1.14 ± 0.03                  | 490–10                |
| E1294A       | –     | 1.47 ± 0.06                  | 570–20                |
| E1294 M      | –     | 0.78 ± 0.06                  | 470–60                |
| E1294H       | –     | 1.26 ± 0.09                  | 370–20                |
| E1294K:Q894A | –     | 4.4 ± 0.13                   | 350–30                |
| E1294K:K1199A| –     | 0.76 ± 0.04                  | 350–20                |

a Experiments were performed as described in Experimental procedures.

b NA, no detectable activity.

c The concentration range of the substrate calculated as released GalNAc residues determined with the Morgan-Elson colorimetric assay and obtained from end points fitted to the progress curves of substrate turnover (see Supplementary Materials).

d k_{cat}/K_{m} was derived as the slope from a linear fit of catalytic rate versus substrate concentration.

mM NaCl). Assay time and enzyme concentrations were adjusted accordingly to obtain full progress-curves as described previously [15]. Fits to obtain initial velocities and substrate concentration from progress curves, as well as steady-state kinetic parameters and their reported errors from the derived saturation curves, were performed as previously described [15].

### 3. Results and discussion

**Molecular ligand docking in the endo-α-N-acetylgalactosaminidase active site.** To probe the accommodation of the Neu5Aco2-3Galβ1-3GalNAc glycan in the endo-α-N-acetylgalactosaminidase active site, molecular dockings with this ligand were performed using the software Glide as described in Experimental procedures. The high-resolution structure of the EngBF homolog, EngSP, in complex with Galβ1-3GalNAc (PDB ID: 5A59) was used as a template for the docking. For each docking experiment the RMSD obtained from comparing the position of the docked Galβ1-3GalNAc moiety of the extended glycan to that of the co-crystallized disaccharide Galβ1-3GalNAc in the EngSP structure were plotted against the Glide energy score (Fig. 2A).

Intriguingly, the extended glycan could readily be docked into the active site to position itself close to the subsites of the Galβ1-3GalNAc glycan. Analyzing the best scoring poses for the extended glycan, demonstrated a marked favorable decrease in the Glide energy score concomitant with lowered RMSD values. In the docked complexes, the galactose G6-hydroxyl in the extended glycan was either within hydrogen bonding...
distance of the Asp-658 carboxylate, common to the crystal structure of the EngSP Galβ1-3GalNAc complex, or rotated towards the Asp-1254 carboxylate (Fig. 2B). Inspection of the best poses, judged by lowest values of Glide energy score and RMSD, showed that the Neu5Ac C1 carboxylate was either pointing towards the enzyme surface (Fig. 2C) or the solvent (Fig. 2D). Not surprisingly, the docking experiments with the Neu5Acα2-3Galβ1-3GalNAc ligand in the EngSP active site show that binding of this glycan is restricted by accommodation of the Neu5Ac moiety. And perhaps more important, the docking experiments indicate that the Neu5Ac residue exerts restrictions on correctly locating the galactose residue to its subsite as defined in the binding of the Galβ1-3GalNAc glycan. Thus, a simple metric for a rational substitution of side chains in the active site of the enzyme that facilitate docking of Neu5Acα2-3Galβ1-3GalNAc, seemed to be the aim for a further concomitant decrease in the Glide energy score and RMSD.

To provide some guidance for the engineering of the EngBF active site, we tested both wild-type enzyme and different EngBF variants with alanine substituted amino acid residues for the ability to cleave the Neu5Acα2-3Galβ1-3GalNAc glycan of fetuin. Most variants are previously described [7,15] and form part of the active site lid, W750A, or comprise the binding site for the galactose moiety, Q894A, K1199A and D1295A (Fig. 3A). To complement these, we also constructed the E1294A variant, as Glu-1294 is also potentially blocking an extended binding site for Neu5Acα2-3Galβ1-3GalNAc (Fig. 3A). Despite recent findings with endo-α-N-acetylgalactosaminidases from other sources [6], no activity on fetuin was detected for the wild-type EngBF or the variants W750A, Q894A, K1199A, D1295A under the reaction conditions and assay method used in the present work. Nevertheless, the side-chain substitutions at the positions corresponding to Gln-894 and Lys-1199 in EngBF were recently shown to improve the activity of EngSP.
Fig. 4. Molecular docking of Neu5Aco2-3Galβ1-3GalNAc in EngSP(TIGR4) (PDB ID: 5A59) and variants substituted at the position of Glu-1253 (Glu-1294 in EngBF). A) The Glide energy score plotted as a function of the RMSD calculated for the superimposition of the Galβ1-3GalNAc moiety of the docked glycan, Neu5Aco2-3Galβ1-3GalNAc, and that of Galβ1-3GalNAc co-crystallized with the EngSP enzyme. The inset shows all the data using an extended x-axis. The data points in the square of the main plot represents the poses with lowest RMSD and Glide energy score for the model EngSP E1253K variant (homologous to the E1294K variant of EngBF). B) Superimposition of the Galβ1-3GalNAc co-crystallized with the EngSP enzyme (dark grey) and the best poses (green) from dockings of Neu5Aco2-3Galβ1-3GalNAc into the model EngSP E1253K structure. Side chains interacting with the Neu5Ac moiety are displayed as sticks (dark grey). The side-chain numbering is according to EngSP and numbering for the corresponding residues in EngBF is shown in parentheses. The figure was created using PyMol (Shr¨odinger Inc.). C) The Glide energy score plotted as a function of the RMSD calculated for the superimposition of the Galβ1-3GalNAc moiety of the docked glycan, Neu5Aco2-3Galβ1-3GalNAc, and that of Galβ1-3GalNAc co-crystallized with the EngSP enzyme. The inset shows all the data using an extended x-axis. The main plot shows the poses with lowest RMSD and Glide energy score for the model EngSP E1253K variant and the double variants, E1253K:Q868A and E1253K:K1156A (homologous to the E1294K, E1294K:Q894A and E1294K:K1199 variants of EngBF). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
on fetuin up to four-fold [6]. EngBF is almost identical to EngSP in the region of the active site (Fig. 3B) suggesting that our assay is not sensitive enough to detect activity at the level reported for EngSP under the reaction conditions that we have chosen. However, the E1294A variant was found to release the Neu5Ac moiety of the Neu5Ac-β-1-3GalNAc ligand when fetuin was used as a substrate. This observation was followed up by analysis of several in silico side-chain substitutions at the homologous position, Glu-1253, in the EngSP structure used for the docking experiments. Here, the E1253K variant was found to produce docking poses with the Neu5Ac-β-1-3GalNAc glycan having a markedly lower RMSD and Glide energy score compared to wild type enzyme and that of other side-chain substitutions at this position (Fig. 4A). Dissection of the best poses from docking of the extended glycan in the E1253K variant (Fig. 4B) shows the hydroxyl groups of C5 and C6 of the galactose moiety to be within hydrogen bond distance of the Asp-658 and Asp-1254 carboxylates. Thus, the interactions of the wild-type enzyme and the Galβ1-3GalNAc ligand found in the crystal structure of the EngSP complex and those of this moiety of the Neu5Ac-β-1-3GalNAc glycan in dockings performed on the E1253K variant, are in better agreement than what is found with the latter glycan in the dockings with the wild-type EngSP enzyme. In addition, the positioning of the Neu5Ac moiety is identical for all the dockings in this group of best poses for the E1253K variant with a solvent exposed positioning of the Neu5Ac moiety at a position later than the Neu5Ac monomer (Fig. 5A). Supporting this we set out to identify the chemical composition of the glycan released by the EngBF variants, exemplified by E1294K. The product released by E1294K upon incubation with fetuin was analyzed by HPEAC-PAD as described in Experimental procedures, before and after α2-3-specific Neu5Ac removal by PmST1 sialyltransferase. The chromatogram shows glycan release after incubation in the absence or presence of E1294K. In 3), the E1294K treated sample from 2) was subsequently incubated with Nan1 sialidase to release Neu5Ac. The composition of the released glycans and carbohydrates are indicated above the chromatogram. Experiments were performed as described in Experimental procedures. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Product analysis. The finding that the EngBF variants substituted at Glu-1294 were able to produce a detectable product from incubation with fetuin, in contrast to wild-type enzyme, gave a strong indication that these variants were able to process the extended glycan, Neu5Ac-α-2-3Galβ1-3GalNAc. To substantiate this we set out to identify the chemical composition of the glycan released by the EngBF variants, exemplified by E1294K.

The product released by E1294K upon incubation with fetuin was analyzed by HPEAC-PAD as described in Experimental procedures, before and after α2-3-specific Neu5Ac removal by PmST1 sialyltransferase in the presence of CMP [16]. Although the Galβ1-3GalNAc entity was fully hydrolyzed to galactose and GalNAc during the HPEAC-PAD analysis, as this is performed in 100 mM NaOH, distinct peaks that are likely to represent the full trisaccharide Neu5Ac-α-2-3Galβ1-3GalNAc or the partial hydrolysate Neu5Ac-α-2-3Gal eluted at a position later than the Neu5Ac monomer (Fig. 5A). Supporting this attribution of the late eluting peaks, was the disappearance of these peaks after treatment of the released glycan with PmST1 sialyltransferase and instead the chromatogram showed increased amounts of Galβ1-3GalNAc hydrolysate (Gal + GalNAc) as well as free Neu5Ac (Fig. 5A).
We also performed a TLC analysis (Fig. 5B) of product from incubation of fetuin with E1294K. The application of standards and the treatment of the product with NanI sialidase indicate that the released glycan is Neu5Acα2-3Galβ1-3GalNAc.

Progress curves of fetuin processed by E1294K showed a maximal release of two glycans per fetuin molecule (Fig. 6A), compared to 2-2.5 glycans obtained with wild-type EngBF or E1294K upon incubation with asialo-fetuin (Fig. 6B). This stoichiometry of Neu5Acα2-3Galβ1-3GalNAc glycans, or Galβ1-3GalNAc after NanI sialidase treatment, per fetuin molecule is in agreement with an average of 2-2.5 O-glycosylations per fetuin molecule with the majority being the Neu5Acα2-3Galβ1-3GalNAc-Ser/Thr as previously determined [5].

Steady-state kinetic analysis of wild-type EngBF and variants using asialo-fetuin or fetuin as a substrate. The EngBF variants altered at position Glu-1294 were subjected to a steady-state kinetic analysis on asialo-fetuin to evaluate the impact of the substituted side chains with respect to the native activity of EngBF of processing Galβ1-3GalNAc O-glycosylations (Table 1). The steady-state kinetic parameters for all variants were close to those of wild-type EngBF (Table 1) and Glu-1294 appears not in any way critical for activity. In addition, this position seems functionally very tolerant to side-chain substitutions, both in terms of size and charge. With fetuin as a substrate, all the EngBF variants altered at Glu-1294 were active, but showed a linear increase in activity with increasing fetuin concentrations and no sign of substrate saturation at up to 30 mg/mL fetuin. This only allowed for the determination of kcat/KM both at pH 4.5 and pH 6.0 (Table 2). It was also clear that all the variants were far better catalysts at pH 6.0 than at pH 4.5 - up to about 50-fold improved for the E1294A variant (Table 2).

Regardless of the pH of the incubation, no significant activity of the wild-type EngBF enzyme was detected (Table 2). Inspired by the docking experiments and the previously published work on EngSP [6], we also investigated if the double substituted variants, E1294K-Q894A and E1294K-K1199A, showed any improved activity on fetuin over that of E1294K alone. Such an improvement, although limited to 1.5 to 4-fold, was observed for the E1294K-Q894A variant both at pH 4.5 and pH 6.0 (Table 2).

Conclusion. In the present work, we have shown that abolishing the negative side chain on a single position in EngBF, Glu-1294, promotes the accommodation of the larger and negatively charged glycan, Neu5Acα2-3Galβ1-3GalNAc, irrespective of the size of the substituted side chains. However, the full potential of engineering EngBF to accept this extended glycan as a substrate, is most likely not fully exhausted with the variants described in the present work. The steady-state kinetic analysis with fetuin as a substrate, where no saturation was observed at the concentrations of fetuin available, indicates that much can be gained by improving the binding of Neu5Acα2-3Galβ1-3GalNAc to the active site of EngBF. We speculate that if 1) positioning of the extended glycan in the binding site of EngBF can be achieved with the Galβ1-3GalNAc moiety exactly overlapping that of the native substrate, and 2) that the structural and chemical environment in the active site where the cleavage reaction takes place is fully preserved, then a similar kcat for the cleavage of both glycans may be expected. This is important, because if the problem of cleaving glycans extended beyond the Galβ1-3GalNAc moiety is reduced to making space for the glycan to position itself to overlap the native substrate, the task is by far more simple than having to manipulate the catalytic machinery. If we were to assume a similar kcat for release of glycans by the EngBF variants active on both asialo-fetuin and fetuin, the values of kcat/KM for Neu5Acα2-3Galβ1-3GalNAc (Table 2) indicate a KM for fetuin of the EngBF variants to be in the molar range at pH 4.5 and in the millimolar range at pH 6.0. Also, it may be that the greatly improved catalytic efficiency displayed by the variants at pH 6.0 compared to pH 4.5, points to a titration of the Neu5Acα2-3Galβ1-3GalNAc glycan resulting in a favorable uniform charge distribution at the higher pH. In the future, docking experiments with Neu5Acα2-3Galβ1-3GalNAc in the active site of EngBF, evaluated as here by superimposition of the glycan moiety and the native substrate, could be combined with e.g. possible neutralization of electrostatic repulsion (Fig. 7A and B). This approach may pave the way for engineering endo-α-N-acetylgalactosaminidase variants with much improved activity towards this extended and negatively charged glycan. As for now, the enzyme variants described here may serve as a tool for identification of O-glycosylations and for determination of stoichiometry of the Neu5Acα2-3Galβ1-3GalNAc glycan in glycoproteins, as suggested previously [6]. But perhaps more interestingly, is the further development of a variant with similar activity levels on Neu5Acα2-3Galβ1-3GalNAc as for wild-type enzyme on Galβ1-3GalNAc, which may find its use in applications involving transglycosylation, which the wild-type endo-α-N-acetylgalactosaminidase performs readily [2,23,26]. Such EngBF variants could show to be important tools in the synthesis of O-linked glycoconjugates.

**Funding sources**

This work was supported by the Danish Dairy Research Foundation, the Villum Fonden and the Independent Research Fund Denmark (9041-00126B) to MW. Additional funding was from a PhD scholarship to BS from the Ministry of Higher Education and Scientific Research of Iraq.

**Acknowledgement**

The authors wish to acknowledge the expert technical assistance by Aida Curovic.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jab.2022.109280.
References

[1] H. Ashida, R. Maki, H. Uzawa, Y. Tani, M. Kiyohara, M. Fujita, A. Imamura, H. Ishida, M. Kiso, K. Yamamoto, Characterization of two different endo-alpha-N-acetylgalactosaminidases from probiotic and pathogenic enterobacteria, Bifidobacterium longum and Clostridium perfringens, Glycobiology 18 (2008) 727–734.

[2] D. Koutsioulis, D. Landry, E.P. Guthrie, Novel endo-alpha-N-acetylgalactosaminidases with broader substrate specificity, Glycobiology 18 (2008) 799–805.

[3] K. Fujita, F. Oura, N. Nagamine, T. Katayama, J. Hiratake, K. Sakata, H. Kumagai, H. Ashida, R. Maki, H. Ozawa, Y. Tani, M. Fujita, A. Imamura, D. Koutsioulis, D. Landry, E.P. Guthrie, Novel endo-alpha-N-acetylgalactosaminidases with broader substrate specificity, Glycobiology 18 (2008) 727–734.

[4] T. Takahashi, M. Ikegami-Kawai, R. Okuda, K. Suzuki, A fluorimetric Morgan-Elson assay method for hyaluronidase activity, Anal. Biochem. 322 (2003) 257–263.

[5] L.M. Willis, R. Zhang, A. Reid, S.G. Withers, W.W. Wakarchuk, Mechanistic investigation of the endo-alpha-N-acetylgalactosaminidase from Streptococcus pneumoniae R6, Biochemistry 48 (2009) 10334–10341.

[6] D. Koutsioulis, D. Landry, E.P. Guthrie, Novel endo-alpha-N-acetylgalactosaminidases with broader substrate specificity, Glycobiology 18 (2008) 727–734.

[7] H. Yu, H. Chokhawala, R. Karpel, H. Yu, B. Wu, J. Zhang, Y. Zhang, Q. Jia, X. Chen, A multifunctional Pasteurella multocida sialyltransferase: a powerful tool for the synthesis of sialo-oligosaccharides libraries, J. Am. Chem. Soc. 127 (2005) 17618–17619.

[8] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.K. Klicic, D.T. Mainz, M. P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, J. Med. Chem. 47 (2004) 1739–1749.

[9] T.A. Halgren, B.R. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, J. L. Banks, Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, J. Med. Chem. 47 (2004) 1750–1759.

[10] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, J. Med. Chem. 49 (2006) 6177–6196.

[11] Schrödinger Release 2016-3, Glide, Schrödinger, LLC, New York, NY, 2018.

[12] Schrödinger Release 2016-3, Maestro, Schrödinger, LLC, New York, NY, 2018.

[13] Woods Group, GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA, 2005-2017. http://glycam.org.

[14] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L.E. Felberg, D.H. Brookes, R.A. Friesner, R.B. Murphy, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, J. Med. Chem. 47 (2004) 1739–1749.

[15] J.L. Banks, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, J. Med. Chem. 49 (2006) 6177–6196.

[16] Enrichment factors in database screening, J. Med. Chem. 47 (2004) 1750–1759.

[17] T.A. Halgren, B.R. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, J. L. Banks, Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, J. Med. Chem. 47 (2004) 1750–1759.

[18] T.A. Halgren, B.R. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, J. L. Banks, Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, J. Med. Chem. 47 (2004) 1750–1759.

[19] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, J. Med. Chem. 49 (2006) 6177–6196.

[20] P. Kötter, D. Obsthein, M. Windwarder, F. Altmann, Site-specific analysis of the O-glycosylation of bovine fetuin by electron-transfer dissociation mass spectrometry, J. Proteonomics 108 (2014) D490–D495.

[21] M. Windwarder, F. Altmann, Site-specific analysis of the O-glycosylation of bovine fetuin by electron-transfer dissociation mass spectrometry, J. Proteonomics 108 (2014) D490–D495.

[22] J.F. Wardman, P. Rahfeld, F. Liu, C. Morgan-Lang, L. Sim, S.J. Hallam, S. G. Withers, Discovery and development of promiscuous O-glycan hydrolases for removal of intact Sialyl T-antigen. ACS Chem. Biol. 16 (2021) 2004–2015.

[23] R. Suzuki, T. Katayama, M. Kinosoku, H. Kumagai, T. Wakagi, H. Shoun, H. Ashida, K. Yamamoto, S. Fushinobu, Crystallographic and mutational analyses of substrate recognition of endo-alpha-N-acetylgalactosaminidase from Bifidobacterium longum, J. Biochem. 146 (2009) 389–398.

[24] L.M. Willis, R. Zhang, A. Reid, S.G. Withers, W.W. Wakarchuk, Mechanistic investigation of the endo-alpha-N-acetylgalactosaminidase from Streptococcus pneumoniae R6, Biochemistry 48 (2009) 10334–10341.

[25] K.J. Gregg, M.D. Suits, L. Deng, D.J. Vocadlo, A.B. Boraston, Structural analysis of the family 101 glycosidase complex with carbohydrates reveals insights into its mechanism, J. Biol. Chem. 290 (2015) 25657–25669.

[26] I. Brockhausen, P. Stanley, O-GalNAc glycans, in: A. Varki, R.D. Cummings, J. D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer, J. H. Prestegard, R.L. Schaara, P.H. Seeberger (Eds.), Essentials of Glycobiology (Rd, Cold Spring Harbor, NY), 2015, pp. 99–111.

[27] D.K. Hansen et al. Archives of Biochemistry and Biophysics 725 (2022) 109280