JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Glycoprotein Labeling Using Engineered Variants of Galactose Oxidase Obtained by Directed Evolution

Julie B. Rannes,[†] Avgousta Ioannou,[†] Simon C. Willies,[†] Gideon Grogan,[‡] Carsten Behrens,[§] Sabine L. Flitsch,[†] and Nicholas J. Turner*[,†]

[†]School of Chemistry, University of Manchester, Manchester Interdisciplinary Biocentre, Manchester, M1 7DN, U.K.

[‡]Department of Chemistry, University of York, York, YO10 5YW, U.K.

[§]Department of Protein and Peptide Chemistry, Novo Nordisk A/S, Maaløv, Denmark

Supporting Information

ABSTRACT: A directed evolution approach has been used for the generation of variants of galactose oxidase (GOase) that can selectively oxidize glycans on glycoproteins. The aldehyde function introduced on the glycans D-mannose (Man) and D-*N*-acetyl glucosamine (GlcNAc) by the enzyme variants could then be used to label the glycoproteins and also whole cells that display mannosides on their surface.

Selective bioengineering of the glycans on glycoproteins and other glycoconjugates represents a powerful method for generating biopharmaceuticals with novel pharmacokinetic and pharmacodynamic properties.[1] Particularly attractive tools for such bioengineering methods are carbohydrate modifying enzymes such as glycosidases, glycosyltransferases, and oxidases, which can be highly selective.[2] However, the inherent limitations in the range of substrates accepted by these enzymes highlights the need for carbohydrate-modifying enzymes with broader substrate ranges, and in this respect directed evolution offers an attractive strategy for engineering enzymes with altered substrate specificity.[3] Chemoenzymatic labeling of glycoproteins often relies upon the enzymatic introduction of an aldehyde group into a sugar residue, thereby generating an orthogonal functionality that can be used for site-specific labeling of proteins (Figure 1).[4] Subsequent modification of the aldehyde with functionalized aminooxy probes enables the site-specific labeling of glycoproteins by the formation of a stable oxime product.[5]

Currently, the only enzyme that can introduce such an aldehyde by sugar oxidation is galactose oxidase (GOase; EC.1.1.3.9) which is highly selective for galactose and talose but will not oxidize other sugars commonly found on glycoproteins.[6] GOase is a copper dependent alcohol oxidase, isolated from *Fusarium* sp.,[7] and oxidizes galactose (Gal) residues as either monosaccharides or glycoconjugates that contain galactose at the nonreducing end.[8] GOase has been used in various applications including biosensors,[9] chemical synthesis,[10] and analytical detection of galactose residues present in glycans and glycoconjugates.[11] Several variants of GOase with activity toward secondary alcohols, D-glucose (Glc), and D-fructose (Fru) have previously been identified, but their activity against other glycoconjugates has not been explored.[12−14] Thus GOase represents

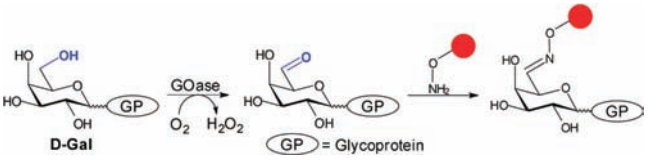

Figure 1. Chemoenzymatic labeling approach.

an attractive target for engineering to accept new carbohydrate substrates that can then be modified in synthetically useful ways.

To broaden the substrate range of GOase, saturation mutagenesis libraries were generated based upon the principle of Combinatorial Active Site Testing (CASTing) reported by Reetz et al.[15] Using this approach it is possible to investigate potential interactions between neighboring residues in the active site of enzymes that are perhaps important in determining substrate specificity. Thereafter, beneficial residues can be recombined for further improvements by Iterative Saturation Mutagenesis (ISM).[16] From initial studies involving the docking of Gal into the active site of GOase, five residues (F194, W290, R330, Q406, and F464) were identified as potentially important for the binding of Gal to the active site of GOase (Figure 2a, underlined residues).[12]

Eight saturation libraries (A−H), based upon randomization at these sites, were constructed using the M_1 (library H) and M_3 (library A−G) sequences as parents[13] and expressed in *E. coli* (Figure 2a). The M_1 sequence contains six mutations compared to the wild-type enzyme (S10P, M70 V, P136, G195E, V494A, N535D) and was chosen due to its improved levels of expression of a more active and stable enzyme in E. *coli* without any change in substrate range.[17] The M_3 sequence contains three additional mutations (W290F, R330K, Q406T) compared to the M_1 sequence and was selected on the basis of its improved activity toward Glc.[14] Approximately 5000 clones from each library were screened using our previously developed colorimetric solid-phase assay,[13] and one variant (H_1), with activity toward Man, was identified from library H. In addition, several variants were isolated with activity toward GlcNAc (E_1, E_2, F_1, F_2) from libraries E and F. The observed changes and positions of key residues important for the altered substrate specificity are shown in Figure 2b. It is noteworthy that variants F_1 and F_2 both possess

Received: March 7, 2011
Published: April 28, 2011

dx.doi.org/10.1021/ja2018477 | *J. Am. Chem. Soc.* 2011, 133, 8436–8439

two mutations relative to the parent and thus represent examples in which CASTing offers advantages compared to multiple single site NNK libraries.

The specific activity of the purified variant GOase enzymes toward both GlcNAc and Man at the 50 mM substrate concentration was measured in order to determine the improvements in activity compared to the parent (M_1 and M_3) and wild type enzymes (Supporting Information (SI), Table S1). The H_1 variant showed the best activity toward Man with a more than 120-fold improvement in activity compared to the wild type enzyme. The H_1 variant carries a single mutation (R330K) which has previously been identified to be important for the activity of GOase toward D-fructose.[12] In contrast, the best variant with activity toward GlcNAc was F_2 with a more than 136-fold improvement in activity compared to the wild type enzyme. The F_2 variant contains the R330K mutation plus additional changes at W290F, Q406E, and Y405F. Additionally, several new variants with activity toward Man and GlcNAc were identified from three of the eight double saturation mutagenesis libraries screened.

Based upon the model of Gal at the active-site of wild type GOase,[12] the amino acid residues responsible for the change in substrate specificity, (P463, Y405, Q406, R330) appear to be in close proximity to either the 2- or 4-OH groups of Gal (Figure 2b). This observation may explain the change in substrate specificity of the variants, as Man and GlcNAc are respectively C2/C4- and C4-epimers of Gal. Further insight into the observed change in substrate specificity was obtained from the crystal structure of the E_1 variant solved at 2.19 Å (SI, Table S2). No significant changes in the overall structure, and only subtle changes in the active site, were observed in comparison to the crystal structure of the wild type enzyme (SI, Figure S1).

This observation suggests that the mutations do not lead to a significant conformational change of the enzyme but that it is the subtle combination of mutations in the active site that is important for the changed substrate specificity. The precise combination of mutations needed to engender new substrate specificity in GOase cannot easily be predicted but can be revealed *via* an approach based upon mutagenesis coupled with screening. Further evidence for the subtle nature of the mutations can be seen from an analysis of the substrate specificity of some of the variants (Table 1).

While the H_1 variant possesses a relatively narrow substrate range, largely favoring Man, the E and F variants display activity toward a broader range of substrates. For these four variants, the improvement in activity toward GlcNAc was accompanied by an increase in the rate of oxidation of both Man and Glc, indicating that these variants possess decreased substrate specificity relative to the wild type enzyme. The activity toward Glc was improved even compared to the parent M_3 variant, which was previously shown to have a 100-fold improvement in activity toward Glc compared to the M_1 variant.[14] Furthermore, all of the variants analyzed retained their activity toward Gal, although at a reduced conversion rate compared to the wild type enzyme. As expected, the catalytic efficiency of the H_1 variant toward Man (k_{cat}/K_m = 35 ± 3 M^{-1} s^{-1}) was reduced compared to the catalytic efficiency of the parent M_1 variant toward Gal (k_{cat}/K_m = 26,964 ± 744 M^{-1} s^{-1}). This decrease arises from a reduced catalytic turnover of the enzyme (~100-fold) and a slight

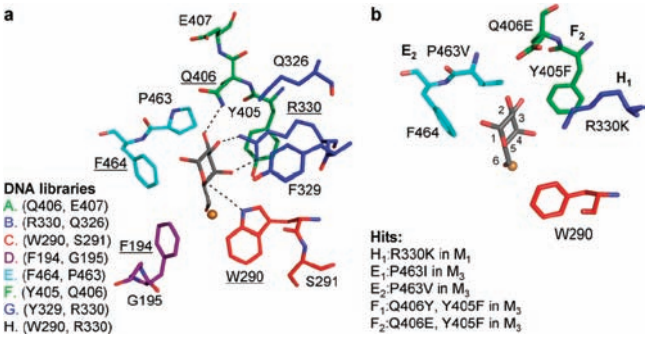

Figure 2. Double saturation mutagenesis libraries (a) and identified hits (b).

Table 1. Substrate Specificity of Parent and Engineered Variants

Conversion rates (μM/min)[a]	WT S2, G6	M_1 S20P M70V P136 G195E V494A N535D in WT	M_3 W290F R330K Q406T in M_1	H_1 R330K in M_1	E_1 P463I in M_3	E_2 P463V in M_3	F_1 Y405F Q406Y in M_3	F_2 Y405F Q406E in M_3
Glc	0.0	0.1	0.4	0.1	0.5	0.8	1.4	1.6
Glc-α-OMe	0.0	0.0	0.2	0.0	0.2	0.3	0.4	0.7
Glc-β-OMe	0.0	0.0	0.2	0.0	0.4	0.6	1.1	1.4
GlcNAc	0.0	0.0	0.1	0.0	0.5	0.9	1.1	1.9
GlcNAc-α-OMe	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
GlcNAc-β-OMe	0.0	0.0	0.0	0.0	0.2	0.6	0.3	1.5
Man	0.0	0.1	0.1	1.7	0.1	0.4	1.1	1.1
Man-α-OMe	0.0	0.1	0.0	0.5	0.0	0.1	0.3	0.5
Gal	222.5*	220.9*	3.4	12.3	4.7	1.4	4.2	3.1
Gal-α-OMe	210.4*	210.5*	4.2	14.7	5.2	1.5	4.2	3.4
Gal-β-OMe	199.9*	199.7*	3.4	16.6	3.9	0.9	3.3	2.9
Tal	127.4*	127.4*	6.1	48.6	13.8	1.9	9.9	6.4
NANA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

0.0-0.1 | 0.2-0.4 | 0.5 - 2.0 | 2.1-10.0 | 10.1-50 | > 50

[a] The conversion rates were determined at 30 °C in phosphate buffer at pH 7.0 using 50 mM of substrate and 12.5 nM enzyme. The reported values represent the mean of three independent measurements performed in triplet. * These values were obtained using 5-fold less enzyme.

dx.doi.org/10.1021/ja2018477 |J. Am. Chem. Soc. 2011, 133, 8436–8439

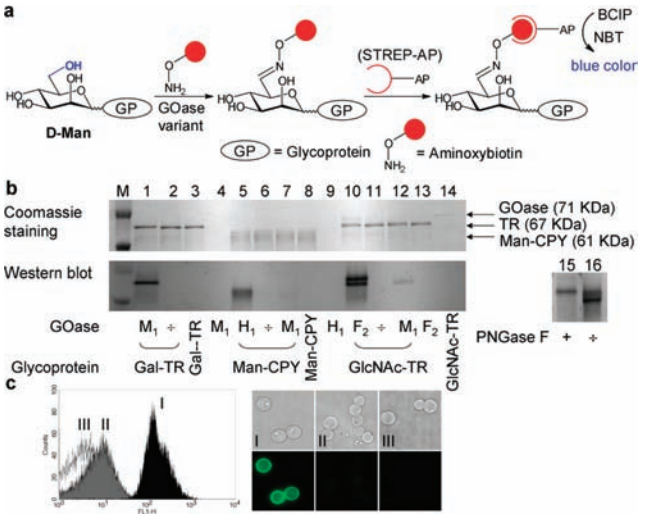

Figure 3. Chemoenzymatic labeling of glycoproteins. (a) Schematic overview of the labeling and detection of glycoproteins terminating in Gal (asialo-transferrin = Gal-TR), Man (mannosylated carboxypeptidase Y = Man-CPY), or GlcNAc (asialo-agalacto-transferrin = GlcNAc-TR). STREP-AP = streptavidin conjugated alkaline phosphatase; BCIP = 5-bromo-4-chloro-3-indolyl phosphate; NBT = nitro blue tetrazolium. (b) Coomassie stained SDS-PAGE gel and Western blot of the labeling reactions with M_1 (lane 1), H_1 (lane 5), and F_2 (lane 10) and their respective glycoproteins. Further shown are control reactions without enzyme (lanes 2, 6, and 11), with M_1 (lanes 7 and 12), with the individual glycoproteins (lanes 3, 8, and 13), and the GOase variants (lanes 4, 9, and 14). Reaction mixtures of biotin-labeled GlcNAc-TR treated with and without PNGase F are shown in lanes 15 and 16, respectively. M - protein marker (top band is 70 KDa and bottom band is 55 KDa). (c) FACS histogram and fluorescence microscopy images of *Pichia pastoris* cells chemoenzymatically labeled with Alexa Fluor 488 using GOase H_1 (I), no enzyme (II), and nontreated cells (III).

increase in the substrate affinity (~7-fold). To search for variants with improved catalytic efficiency, we combined the mutations from the E and F variants according to the principles of Iterative Saturation Mutagenesis. However, none of the combined variants showed improved activity compared to the parent enzymes (data not shown). Additional improvements might be obtained by mutation of amino acids more remote from the active site of the enzyme and could be probed by random mutagenesis techniques. Nevertheless, the observed activities of the H_1 and F_2 variants toward α-Man and β-GlcNAc anomers respectively, both of which are present in *N*-linked mammalian glycoproteins,[18] suggested that these variants were promising candidates for the labeling of glycans and glycoproteins.

Initially, the H_1 and F_2 variants were examined as catalysts for the oxidation of Man and GlcNAc residues respectively. Analysis by ^{13}C NMR spectroscopy revealed that the H_1 and F_2 variants specifically oxidize the C6-OH group of α-D-methyl mannoside (Man-α-OMe) and β-D-methyl *N*-acetylglucosaminide (GlcNAc-β-OMe), respectively (SI, Figures S2 and S3) with >95% conversion in both cases and no evidence of loss of selectivity for the C6 position.

Next, chemoenzymatic labeling of specific target glycoproteins was examined using the variants H_1 and F_2 (Figure 3a). Following treatment of the glycoprotein with the variant GOase, aminoxy-biotin was added and the product examined by gel electrophoresis. In agreement with previous reports,[19] labeling of asialo-transferrin (terminating in Gal residues) was achieved using the M_1 variant which possesses good activity toward Gal. From the Western blot of the reaction mixture, biotinylation was clearly observed for the reaction mixture containing M_1 and asialo-transferrin when compared to the control reaction containing no enzyme (Figure 3b, lane 1 compared to lane 2). Thereafter, the H_1 and F_2 variants were applied in a similar manner for labeling studies. Clear biotinylation of the glycoproteins carrying Man (high mannose carboxypeptidase Y) and GlcNAc (asialo-agalacto-transferrin) was observed using the H_1 and F_2 variants respectively, compared to reaction mixtures without enzyme (Figure 3b, lane 5 compared to lane 6 and lane 10 compared to lane 11). Surprisingly, biotinylation of the F_2 variant itself was also detected during the experiment with asialo-agalacto-transferrin (Figure 3b, lane 10, upper band). This unspecific biotinylation was observed for all of the GOase variants when higher concentrations of GOases were used and seems to be specific for GOase, as other nonglycosylated proteins were not labeled in a similar process (not shown). In addition, we found that the M_1 variant was unable to mediate biotinylation of the mannosylated carboxypeptidase Y whereas a slight biotinylation of asialo-agalacto-transferrin was observed with M_1 (Figure 3b, lanes 7 and 12). We anticipate that the low intensity band observed in Figure 3b, lane 12, is due to the labeling of residual Gal residues that remain from the preparation of the asialo-agalacto-transferrin sample (<2%). Further, the treatment of biotin labeled GlcNAc-TR with PNGase F led to the removal of biotin as shown in lane 15, Figure 3b which indicates that the biotinylation occurs specifically at the glycan moieties. A glycan specific reaction was also demonstrated on RNase B (SI, Figures S4 and S5). Commercial RNase B containing highmannose glycoforms ($GlcNAc_2Man_n$, $n = 5-9$) was treated with a mannose specific H_1 mutant and methoxylamine under one-pot conditions. Reaction mixtures were analyzed by LC-MSD-TOF, and oxymation reactions were observed on RNase-$GlcNAc_2Man_5$ and $GlcNAc_2Man_6$ glycoforms (+26.8 Da). No oxymation reaction was observed using wild type GOase on RNase B under similar conditions.

Finally, we have demonstrated that this methodology can also be applied to the labeling of cells. *Pichia pastoris* cells, which express mannosylated glycoproteins on their surface, were successfully labeled with the fluorescent reagent aminooxy-Alexa Fluor 488 using the GOase H_1 variant (Figure 3c, I), compared to cells treated with the fluorescent reagent but no enzyme (II) and nontreated cells (III).

The method described above for labeling glycoproteins, including those displayed *in situ* on the surface of cells, represents a general approach in that the aldehyde group that is introduced should allow chemical conjugation with a wide variety of aminooxy functionalized probes such as fluorophores, small molecules, and affinity ligands. We believe that this chemoenzymatic labeling approach represents a simple and generally applicable tool for the site-specific oxidation and labeling of previously unacessible glycan motifs in glycoproteins which allows not only the analysis and detection of glycoproteins but also the design and improvement of glycan based therapeutics. However, the GOase variants described herein possess a broader substrate range compared to the wild type enzyme and hence are not suitable for selective oxidation of mixed glycans. Current efforts are directed toward engineering specificity into these variants to provide reagents for diagnostic applications where selective labeling of glycan mixtures is required.

■ ASSOCIATED CONTENT

Supporting Information. Detailed experimental procedures, specific activities of parent and engineered variants, crystal

structures, NMR spectra, LC-MSD-TOF spectra, and X-ray refinement statistics. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
nicholas.turner@manchester.ac.uk

ACKNOWLEDGMENT

We thank Franck Escalettes for the expression plasmids of GOase (M_1 and M_3), Rie Kristine Schjeltved for the preparation and purification of the transferrin glycoproteins, Kate Atkin for preliminary crystallographic work on GOase, and Henning Stennicke for the donation of high-mannose carboxypeptidase Y. This work was supported by grants from the Novo Nordisk Training and Research programme (TRAP), the Biotechnology and Biological Sciences Research Council (BBSRC), the Engineering and Physical Sciences Research Council (EPSRC), and the Royal Society (Wolfson Merit award to S.L.F.).

REFERENCES

(1) (a) Prescher, J. A.; Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13–21. (b) Seeberger, P. H. *Nat. Chem. Biol.* **2009**, *5*, 368–372. (c) Doores, K. J.; Gamblin, D. P.; Davis, B. G. *Chemistry* **2006**, *12*, 656–665.

(2) (a) Daines, A. M.; Maltman, B. A.; Flitsch, S. L. *Curr. Opin. Chem. Biol.* **2004**, *8*, 106–113.(b) Shaikh, F. A.; Withers, S. G. *Biochem. Cell Biol.* **2008**, *86*, 169-177 (2008). (c) Williams, G. J.; Gantt, R. W.; Thorson, J. S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 556–564.

(3) (a) Turner, N. J. *Nat. Chem. Biol.* **2009**, *5*, 1–7. (b) Luetz, S.; Giver, L.; Lalonde, J. *Biotechnol. Bioeng.* **2008**, *101*, 647–653.

(4) (a) Wu, P.; Shui, W.; Carlson, B. L.; Hu, N.; Rabuka, D.; Lee, J.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3000–3005. (b) Prescher, J. A.; Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13–21. (c) Carrico, I. S. *Chem. Soc. Rev.* **2008**, *37*, 1423–1431.

(5) Kalia, J.; Raines, R. T. *Angew. Chem., Int. Ed.* **2008**, *47*, 7523–7526.

(6) van Hellemond, E. W.; Leferink, N. G.; Heuts, D. P.; Fraaije, M. W.; van Berkel, W. J. *Adv. Appl. Microbiol.* **2006**, *60*, 17–54.

(7) (a) Whittaker, J. W. *Chem. Rev.* **2003**, *103*, 2347–2363. (b) Firbank, S. J.; Rogers, M.; Hurtado-Guerrero, R.; Dooley, D. M.; Halcrow, M. A.; Phillips, S. E.; Knowles, P. F.; McPherson, M. J. *Biochem. Soc. Trans.* **2003**, *31*, 506–509.

(8) (a) Avigad, G. *Arch. Biochem. Biophys.* **1985**, *239*, 531–537. (b) Schlegel, R. A.; Gerbeck, C. M.; Montgomery, R. *Carbohydr. Res.* **1986**, *7*, 193–199.

(9) (a) Jeon, S. I.; Hong, J. W.; Yoon, H. C. *Biotechnol. Lett.* **2006**, *28*, 1401–1408. (b) Sharma, S. K.; Kumar, A.; Chaudhary, R.; Suman; Pundir, C. S.; Sehgal, N. *Artif. Cells Blood Substit. Immobil. Biotechnol.* **2007**, *35*, 421–430.

(10) (a) Bülter, T.; Schumacher, T.; Namdjou, D. J.; Gutiérrez Gallego, R.; Clausen, H.; Elling, L. *ChemBioChem* **2001**, *2*, 884–894. (b) Rodriguez, E. C.; Winans, K. A.; King, D. S.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1997**, *119*, 9905–9906.

(11) (a) Shimaoka, H.; Kuramoto, H.; Furukawa, J.; Miura, Y.; Kurogochi, M.; Kita, Y.; Hinou, H.; Shinohara, Y.; Nishimura, S. *Chemistry* **2007**, *13*, 1664–1673. (b) van Wijk, A.; Siebum, A.; Schoevaart, R.; Kieboom, T. *Carbohydr. Res.* **2006**, *341*, 2921–2926.

(12) Deacon, S. E.; Mahmoud, K.; Spooner, R. K.; Firbank, S. J.; Knowles, P. F.; Phillips, S. E.; McPherson, M. J. *ChemBioChem* **2004**, *5*, 972–979.

(13) Escalettes, F.; Turner, N. J. *ChemBioChem* **2008**, *9*, 857–860.

(14) Sun, L.; Bulter, T.; Alcalde, M.; Petrounia, I. P.; Arnold, F. H. *ChemBioChem* **2002**, *3*, 781–783.

(15) Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A. *Angew. Chem., Int. Ed.* **2005**, *44*, 4192–4196.

(16) Reetz, M. T.; Kahakeaw, D.; Sanchis, J. *Mol. Biosyst.* **2009**, *5*, 115–122.

(17) Sun, L; Petrounia, I. P.; Yagasaki, M.; Bandara, G.; Arnold, F. H. *Protein Eng.* **2001**, *14*, 699–704.

(18) Rich, J. R.; Withers, S. G. *Nat. Chem. Biol.* **2009**, *5*, 206–215.

(19) Wilchek, M.; Spiegel, S.; Spiegel, Y. *Biochem. Biophys. Res. Commun.* **1980**, *92*, 1215–1222.