

Journal of Molecular Catalysis B: Enzymatic 11 (2001) 155–163

www.elsevier.com/locate/molcatb

Glycosynthases: new enzymes for oligosaccharide synthesis

Marco Moracci^{a,*}, Antonio Trincone ^b, Mosé Rossi^{a,c}

^a Institute of Protein Biochemistry and Enzymology-CNR, Via Marconi 10, 80125 Naples, Italy
^b Istituto per la Chimica di Molecole di Interesse Biologico-CNR, Via Toiano 2, Arco Felice, 80072 Naples, Italy
^c Dipartime

Abstract

The mutation of putative acid/base and nucleophile of the active sites of *retaining* glycosyl hydrolases, together with kinetic analysis of the mutants, and stereochemical identification of products lead to useful information for the understanding of the reaction mechanism of these enzymes. This was the preliminary and fundamental step toward the preparation of new enzymatic activities called *glycosynthases*. Direct exploitation of this information has been possible, leading to the design of four new enzymes for oligosaccharides synthesis. The interest for these biocatalysts rises from the fact that the yield of the reaction can be increased and selectivity can be interpreted as key characteristic of the transfer reaction instead of a balance of hydrolytic and transferring pathways followed either by substrates and products. These new biocatalysts possess different specificities and are promising and useful tools in the construction of oligosaccharide molecules of great biological interest. This short review focused the attention on different glycosynthases obtained from four glycosyl hydrolases highlighting on the preparation and development of these new enzymes. \heartsuit 2001 Elsevier Science B.V. All rights reserved.

Keywords: Gycosyl hydrolase; Thermostable enzymes; Archaea; Oligosaccharides; Glycosynthase

1. Introduction

It has been known that carbohydrates can serve as structural components and energy source of the cell.

More interestingly, however, these biomolecules are involved in a variety of molecular recognition processes in intercellular communication and signal transduction such as cell adhesion, differentiation, development, regulation, etc. $[1]$. For these reasons, great interest arose on carbohydrate-based pharmaceuticals and on the development of techniques for the analysis and synthesis of oligosaccharides. However, despite the recent advances $[2]$, the synthesis for large-scale production of complex carbohydrates still cannot be easily performed by the classical chemical procedures which require long protection– deprotection steps for selectivity control and give low final yields. Hence, the enzyme-catalyzed synthesis of oligosaccharides represents, nowadays, an

Abbreviations: NMR, nuclear magnetic resonance; HPLC, high pressure liquid chromatography; Abg, *Agrobacterium faecalis* β -glucosidase; Cex, *Cellulomonas fimi* exoglucanase/ xylanase; Ssß-gly, Sulfolobus solfataricus ß-glycosidase; 2-/4- NP , $2-/4$ -nitro-phenyl-; $2,4-DNP$ -, $2,4$ -dinitrophenyl-; $Glc/Gal/Fuc/Xyl/GluNAc$, glucoside/galactoside/ fucoside/xyloside/Nacetylglucosamine; α -F-Glc/Gal, α -fluoroglucoside/galactoside
 $*$ Corresponding author. Tel.: $+39-081-7257246$; fax: $+39-$

^{081-2396525.}

E-mail address: moracci@dafne.ibpe.na.cnr.it (M. Moracci).

interesting alternative to the classical chemical methods allowing the control of both the regioselectivity and the stereochemistry of bond formation. The enzymatic approach involves mainly two classes of carbohydrate modifying enzymes: glycosyl transferases and glycosyl hydrolases. Glycosyl transferases have been used for oligosaccharide synthesis on a large scale $[3]$, but their poor availability and the high cost of their substrates limited their exploitation. In contrast, glycosyl hydrolases, which are present in all living organisms and allow the use of relatively inexpensive substrates, represent an alternative choice used in transglycosylation or reverse hydrolysis mode. This can be done by modifying the reaction conditions but the products of the reaction are always a substrate for the enzyme, resulting in poor yields unless the equilibrium is displaced in different ways. Recently, an alternative approach to oligosaccharide synthesis, involving a mutant glycosyl hydrolase, was proposed by Withers et al. [4] which lead to a new class of biocatalysts: the *glycosynthases*. These enzymes efficiently synthesize oligosaccharides, but do not hydrolyze them. Moreover, the studies about glycosyl hydrolases reactivation have been of great importance in defining the mechanistic aspects of their functions.

This review highlights the most recent work on the preparation and development of different *glycosynthases*, with particular focus on a thermophilic representative obtained from the b-glycosidase from the archeon *Sulfolobus solfataricus*.

2. The reaction mechanisms of glycosyl hydrolases

Glycosyl hydrolases follow two distinct mechanisms that differ for the configuration of the anomeric carbon in the product of the reaction catalyzed. The reaction mechanisms are termed *retaining* or *inverting* if the enzymatic cleavage of the glycosidic bond liberates a sugar hemiacetal with the same or opposite anomeric configuration of the glycosidic substrate, respectively. The two mechanisms are depicted in Fig. 1: in *inverting* enzymes (Fig. 1A) one residue in the active site acts as a general acid and the other as a general base and the reaction proceeds

Fig. 1. Reaction mechanism of glycosyl hydrolases. (A) *Inverting* enzymes. (B) *Retaining* enzymes.

with a direct displacement of the leaving group by water. The catalytic mechanism of *retaining* enzymes (Fig. $1B$) was proposed for the first time by Koshland in 1953 $[5]$ and is still largely accepted nowadays. These enzymes follow a two-step reaction with a double displacement and the formation of a covalent glycosyl intermediate. The two carboxylic residues in the active site act as a general acid/base and as the nucleophile of the reaction, respectively. In the first step, the concerted action of the general acid and of the nucleophile residues leads to the departure of the aglycon group and to the formation of a glycosyl ester intermediate. In the second step, a water molecule partially deprotonated by the conjugate base of the catalytic acid attacks the anomeric carbon and cleaves the glycosyl ester intermediate

leading to the overall retention of the anomeric configuration of the substrate. When, instead of water, different acceptors intercept the reactive glycosyl-enzyme intermediate, *retaining* enzymes work in transglycosylation mode.

Despite the differences, the two mechanisms show significant similarities: both classes of enzymes employ a pair of carboxylic acids at the active site and both mechanisms operate via transition states with substantial oxocarbenium ion character. However, in *retainers*, as described by a number of excellent reviews, strong evidences support the hypothesis that a covalent intermediate is formed $[6-8]$. Structural data on both class of enzymes increased sharply in recent years revealing that *inverting* enzymes have an active-site structure similar to that of *retaining* enzymes. However, closer inspection revealed that the distance between the catalytic residues in inverters is significantly greater than in retainers (9.2) versus 5.0 Å). This characteristic has been observed almost invariantly and allowed to diagnosticate the mechanism followed by the enzyme. Recent insights into inhibition, structure, and mechanism of *retaining* glycosyl hydrolases lead to a new hypothesis about the location of catalytic acid/base and protonation of substrate. In this respect, Heightman and Vasella [9] classified these enzymes according to *anti* or *sin* protonation trajectory, finding an unsurprising correlation with glycosidase families classified according to amino acid sequences (see below). Amino acids important in catalysis can be clearly identified through specific labeling and kinetic anal-

ysis of site specific mutants. These studies, and the NMR analysis of the reaction products labeled in the anomeric carbon of the substrate, provide direct evidences of the mechanisms followed by these enzymes.

3. Identification of the glycosyl hydrolases residues involved in catalysis

The identification of the reaction mechanism, of the residues involved in catalysis and of the overall folding of any new glycosyl hydrolase available, is nowadays greatly improved. Several excellent reviews of the techniques used to this aim exist in literature $[7,8]$; we will describe here only a few examples.

Glycosyl hydrolases are classified into families on the basis of their amino acid sequence $[10,11]$. This classification included at present a huge number of enzymes organized in 77 families and 10 clans or superfamilies, sharing three-dimensional structure and reaction mechanisms. Only 10 enzymes are left unclassified at this time. This classification, continuously updated and available on the web $\frac{http://}{ }$ afmb.cnrs-mrs.fr/ \sim pedro/CAZY/db.html), give a valuable contribution to the general understanding of glycosidases. Together with the analysis of structural data, and even in their absence, amino acids important in catalysis can be identified with several techniques. Labeling with carboxyl-specific reagents, such as epoxides and carbodiimides, was found to be a successful approach for the identification of active site residues in both *inverting* and *retaining* enzymes $[12-14]$. However, this method does not reliably suggest the specific role of each residue, but has to be combined with amino acid sequence alignments, crystal structure inspection and/or kinetic analysis of site-specific mutants (see below). In contrast, activated 2-deoxy-2-fluoro-glycoside inhibitors allowed the formation of stabilized glycosyl-enzyme complexes through the specific binding to the nucleophile of the reaction. In combination with mass spectrometry, this technique allowed the rapid identification and sequencing of labeled peptides even without the need of radiolabels $[15]$. The use of these

mechanism-based inhibitors is nowadays the most powerful method for the direct identification of the carboxyl group acting as the nucleophile of the reaction. With this approach, the sweet almond β glucosidase was assigned to glycosyl hydrolase Family 1 even in the absence of the amino acid sequence of the enzyme $[16]$. In contrast, no reliable chemical method for the identification of the acid catalyst is currently available. Hence, site-directed mutagenesis and kinetic analysis of the mutants remains the approach most used for the definition of the role played by the active site carboxylic groups. Several examples of these methodologies exist in literature and have been reviewed recently $[7]$; we will briefly describe here the identification of the nucleophile and the acid/base catalysts in *retaining* enzymes by mutagenesis and on the use of exogenous nucleophiles.

In *retaining* enzymes, replacement of the active site nucleophile with a non-nucleophilic residue completely inactivated the enzyme since the new amino acid side chain cannot form the glycosyl enzyme intermediate. Whereas, if the acid/base is removed, some activity on activated substrates can be found. In this case, the first and the second step of the reaction, glycosylation and deglycosylation, respectively, are both slowed down, but, in the presence of substrates with good leaving group, the first step (glycosylation) is only marginally affected by mutation. Hence, the acid/base catalyst can be identified among other carboxyl groups of the enzyme through the analysis of the kinetic parameters of the

Fig. 2. Molecular mechanism of reactivation of mutated glycosyl hydrolases in the presence of high concentrations of NaN₃. The residues (Ala or Gly) replacing the acid/base or the nucleophile are indicated by an X. (A) Reactivation of enzymes mutated in the residue acting as the acid/base. (B) Reactivation of enzymes mutated in the residue acting as the nucleophile.

mutants, obtained on various substrates with leaving groups of different ability. The cavity created by mutation could allow the accommodation of a small anion at the β -face of the substrates adjacent to the anomeric center (Fig. 2A) and the attack of his exogenous nucleophile to the intermediate producing β -product [17]. An increase in the reaction rate of the second step (deglycosilation) will be observed in this case, whereas catalysis on substrates with poorer leaving groups will be reactivated to a lower extent.

4. Reactivation of mutated glycosyl hydrolases

When the active site nucleophile is replaced with a non-nucleophilic amino acid of reduced side chain, the mutant resulted completely inactive since it cannot form the glycosyl-enzyme intermediate. However, the rest of the active site is intact and the small cavity created upon mutation can accommodate a small anion exactly as described above for the mutated acid/base catalyst. In the presence of a substrate with good leaving group ability the activity can be restored. This approach was developed for the first time by Withers et al. [18] on the *retaining* β-glucosidase from *Agrobacterium faecalis* (Abg). When the active site nucleophile was changed into alanine (Glu358Ala) the activity of the mutant was 10⁷ -fold lower than that of the wild type, indicating that the direct attack of water was very inefficient. However, the addition of high concentrations of sodium azide or sodium formate accelerated catalysis 10⁵ -fold. In the case of azide, the reaction proceeded through the direct attack of the exogenous nucleophile to the anomeric center of the substrate from the α side of the pyranose ring (Fig. 2B), leading to α -glucosyl azide as product identified by ¹H NMR. The activity of the azide-restored mutant followed an *inverting* mechanism when compared to the wild type. With the same approach, the reaction mechanism of three other enzymes, the exoglucanase/xylanase from *Cellulomonas fimi* (Cex), the 1,3-1,4-b-glucanase from *Bacillus licheniformis* and the β-glycosidase from *S. solfataricus* (Ssβ-gly) was changed from *retaining* to *inverting* [17,19,20]. In the case of mutant Ss β -gly, direct α -glucosyl azide

formation has been followed in a reaction using anomerically ${}^{13}C$ labeled substrate [20]. These results demonstrated that this methodology could be of general applicability, involving both exo- and endoenzymes from diverse sources and with different functions and specificities. Interestingly enough, with a complementary approach, Kuroki et al. [21] were able to convert the T4 lysozyme from *inverting* glycosyl hydrolase to a *retaining*.

The degree of reactivation obtained in the presence of external nucleophiles varied significantly if compared to the wild type: in Table 1 the kinetic parameters for mutated glycosyl hydrolases reactivated with exogenous nucleophiles are reported. In all cases the only two anions that revealed a relevant effect were azide and formate. Kinetic constants obtained at the optimal conditions of pH, temperature, etc., greatly differ among the enzymes described, however, for a simple comparison, only the values obtained on 2,4-dinitro-phenyl-glycoside substrates are reported in Table 1. The comparison among Glu into Ala mutants revealed that the highest rescue of activity is obtained with the modified $Ss\beta$ -gly, presumably for the general increased resistance of the thermophilic enzymes to the high ionic strength. Remarkably, this rescue is much more evident in mutant with Gly as substituting amino acid, rendering it active also on the more stable 2-NP-Glc. This suggests that the size of the substituting residue can play a relevant role.

Although azide and formate restored the activities of all enzymes, their functioning is different at the molecular level. In the presence of azide the product formed in the reaction $(\alpha$ -glycosyl azide) possesses an anomeric configuration opposed to that of the substrate, leading to the consideration that an *inverting* enzyme has been generated by mutation $[17–20]$. In the case of $Ss\beta$ -gly Glu387Gly mutant, the anomeric configuration of the product of the reaction in the presence of sodium formate is retained as demonstrated by observing the formation of the 2,4 dinitrophenyl β -D-laminaribioside product [20]. The stereochemistry of the new glycosidic bond implies a two step mechanism (see below) in which the external nucleophile assists the leaving of the phenolate group and then stabilizes the glucosyl unit until it was transferred to the acceptor. This biomimetic role of formate has been described as assistant nucle-

ophile [20] with formation of the intermediate, which has also been observed by 1 H NMR [17].

5. *Glycosynthases* **in the oligosaccharide synthesis**

The two main approaches for the glycosyl hydrolases-catalyzed synthesis of glycosidic bond are: direct reversal of hydrolysis and the trapping of glycosyl enzyme intermediate by an acceptor other than water. Both methodologies were described and excellent reviews are present in literature in this field $[22–24]$. The product of the reaction is still substrate for the enzyme and subject to hydrolysis, hence reaction conditions must be carefully controlled and reaction must be closely monitored to increase final yields. In this case, the overall stereochemistry of the reaction depends on the relative activity of the substrate and the different products formed.

Owing to the possibility of restoring the activity of mutated glycosyl hydrolases by external nucleophiles, these biocatalysts become interesting enzymes for the synthesis. In fact, since they are lacking catalytic nucleophile, they act only on activated substrates and can efficiently synthesize new glycosidic bonds without hydrolyzing the products formed. For these reasons they were called *glycosynthases*.

The identification of α -cellobiosyl fluoride recognized as a coupling product of the reactivated Abg

Glu358Ala mutant $[18]$ was the starting point for the development of *glycosynthases*. These novel biocatalysts operate by the two mechanisms described in Fig. 3: the two pathways differ about the origin of glycosyl intermediate. In the first case (*inverting glycosynthases*) the donor substrate mimics the glycosyl intermediate containing an α bond whilst in the second (*retaining glycosynthases*) the intermediate is formed from an activated β substrate by the enzyme.

The *glycosynthases* differ in the selectivity of the reaction catalyzed and in the use of different donors. After the first preliminary report, Withers et al. $[4]$ described the synthesis of different oligosaccharides using the Abg mutant and α -F-Glc or -Gal as donors and different aryl glycosides as acceptors. Mesophilic *glycosynthases* are used in the synthesis only by pathway 1 in Fig. 3, by using α -F-monosaccharides (Gal, Glc) and α -laminaribiosyl fluoride [4,25]. The thermophilic *glycosynthase* is able to act on a wide range of donors of 2- and 2,4-dinitrophenyl-glycosides (β -Glc, β -Fuc, β -Gal and β -Xyl) through pathway 2 and α -F-Glc using pathway 1 [20,26].

Yields and selectivity are key features of these reactions: in the case of Abg *glycosynthase*, the yield ranged from 64% to 92% and β -1,4 glycosidic bonds were always formed, but not in the case of 4-NP-Xyl $(\beta-1,3)$. Differences in the composition of products were reported depending upon the usage of glucosyl and galactosyl fluoride as donors. In the

Fig. 3. Pathways followed by *glycosynthases* in the oligosaccharide synthesis. In *inverting glycosynthases* (pathway 1) the glycosyl fluoride substrate mimics the glycosyl intermediate. In *retaining glycosynthases* (pathway 2) the "Substrate" is an activated β -glycoside and the intermediate is formed by the enzyme upon the glycosylation of formate. The following deglycosylation step is common to the two pathways and the carbohydrate molecule is transferred to the acceptor (R).

first case tri- and tetrasaccharides were observed, while the reaction stops disaccharide formation in the latter, showing the inhibitory effect of the axial OH group for sequential transferring reactions.

Selectivity of formation versus β -1,4 glycosidic link has also been observed with *glycosynthase* enzyme obtained from the endo-glucanase from *B. licheniformis* [25]. In this case α -laminaribiosyl fluoride is the donor used with different acceptors such as: methylumbelliferyl mono- and disaccharides and 4-NP-Gal and 4-NP-GluNAc. The products formed were laminaribiosyl derivatives of the acceptors; the yield in the best case reported is 90% as calculated by HPLC. It is worth noting that, in this case, the transfer involved directly a disaccharide unit. This finding demonstrates that an endoglucanase can be transformed into an *endoglucansynthase* able to transfer preformed sequences of carbohydrate units. More recently, another mutated endoglucanase, from *Humicola isolens*, was used in the synthesis of natural or modified cellooligosaccharides and for the synthesis of different fluorescent oligosaccharides as potential substrate analogues for cellulases of different families [27].

The thermophilic representative is the only *glycosynthase* described so far that is able to function with both pathways in Fig. 3 either as *inverting* or *retaining glycosynthase*. The enzyme can be used with 2,4-DNP-Glc and 2-NP-Glc in presence of formate leading to the synthesis of β -1,3-disaccharides and tri- and tetrasaccharides containing both β -1,3and β -1,6-bonds with high regioselectivity and yields $(80-90\%)$ [26]. By use of α -F-Glc as donor, this enzyme acts as *inverting glycosynthase*, allowing the design of new products by supporting the reaction with different acceptors such as simple alcohols or various pyranosidic acceptors. Moreover, the finetuning of the donor/acceptor ratios could improve yields and selectivity. Using this approach single products were obtained with high selectivity and over 80% yield [26]. These studies demonstrated that both approaches are possible with this thermophilic *glycosynthase* widening the potentiality of these novel biocatalysts and encouraging their exploitation in the synthesis of target oligosaccharides. In fact, the branching functionalization represents a unique characteristic of Ss_{B-gly} *glycosynthase*, which is of applicative interest in the pharmaceutical and food fields, but can also be used for the synthesis of new substrates and inhibitors for glycosidases.

Acknowledgements

We thank the coworkers, whose names are listed in the references, for their contributions to this research. This work was supported by EC project: "Biotechnology of extremophiles" contract no. BIO2-CT93-0274 and by the EC FAIR project: "Enzymatic Lactose Valorization" contract no. 1048/96. We thank S. Zambardino, V. Mirra, E. Pagnotta, G. Imperato and O. Piedimonte for the technical assistance.

References

- [1] A. Varki, Glycobiology 3 (1993) 97.
- [2] G.J. Boons, Tetrahedron 52 (1996) 1095.
- [3] J.G.I. Nilsson, TIBTECH 6 (1988) 256.
- [4] L.F. Mackenzie, Q. Wang, R.A.J. Warren, S.G. Withers, J. Am. Chem. Soc. 120 (1998) 5883.
- [5] D.E. Koshland, Biol. Rev. 28 (1953) 416.
- [6] M.L. Sinnot, Chem. Rev. 90 (1990) 1171.
- [7] J.D. McCarter, S.G. Withers, Curr. Opin. Struct. Biol. 4 (1994) 885.
- [8] A. White, D.R. Rose, Curr. Opin. Struct. Biol. 7 (1997) 645.
- [9] T.D. Heightman, A.T. Vasella, Angew. Chem., Int. Ed. 38 (1999) 750.
- [10] G. Davies, B. Henrissat, Structure 3 (1995) 853.
- [11] B. Henrissat, A. Bairoch, Biochem. J. 316 (1996) 695.
- 12] H. Toda, Y. Nitta, S. Asanami, J. Kim, F. Sakiyama, Eur. J. Biochem. 216 (1993) 25.
- [13] M.R. Bray, A.J. Clarke, Eur. J. Biochem. 219 (1994) 82.
- [14] F. Febbraio, R. Barone, S. D'Auria, M. Rossi, R. Nucci, G. Piccialli, L. De Napoli, S. Orrù, P. Pucci, Biochemistry 36 (1997) 3068.
- [15] S. Miao, J.D. McCarter, M. Grace, G. Grabowski, R. Aebersold, S.G. Withers, J. Biol. Chem. 269 (1994) 10975.
- [16] S. He, S.G. Withers, J. Biol. Chem. 272 (1997) 24864.
- [17] J.L. Viladot, E. de Ramon, O. Durany, A. Planas, Biochemistry 37 (1998) 11332.
- [18] Q. Wang, R.W. Graham, D. Trimbur, R.A.J. Warren, S.G. Withers, J. Am. Chem. Soc. 116 (1994) 11594.
- [19] A. MacLeod, D. Tull, K. Rupitz, R.A.J. Warren, S.G. Withers, Biochemistry 35 (1996) 13165.
- [20] M. Moracci, A. Trincone, G. Perugino, M. Ciaramella, M. Rossi, Biochemistry 37 (1998) 17262.
- [21] R. Kuroki, L.H. Weaver, B.W. Matthews, Nat. Struct. Biol. 2 (1995) 1007.
- [22] C.-H. Wong, R.J. Halcomb, Y. Ichikawa, T. Kajimoto, Angew. Chem., Int. Ed. Engl. 34 (1995) 521.
- [23] H.J.M. Gijsen, L. Qiao, W. Fitz, C.-H. Wong, Chem. Rev. 96 (1996) 443.
- [24] G.M. Watt, P.A.S. Lowden, S.L. Flitsch, Curr. Opin. Struct. Biol. 7 (1997) 652.
- [25] C. Malet, A. Planas, FEBS Lett. 440 (1998) 208.
- [26] A. Trincone, G. Perugino, M. Rossi, M. Moracci, Bioorg. Med. Chem. Lett. 10 (2000) 365.
- [27] V. Boyer, S. Fort, L. Christiansen, M. Schulein, S. Cottaz, H. Driguez, 3rd Carbohydrate Bioengineering Meeting 11–14 April Poster 1.2(1999).