Recent Advances in the Oligosaccharide Synthesis Promoted by Catalytically Engineered Glycosidases

Giuseppe Perugino,[a] Beatrice Cobucci-Ponzano,[a] Mosè Rossi,[a,b] Marco Moracci[a,*]

[a] Institute of Protein Biochemistry – Consiglio Nazionale delle Ricerche, Via P. Castellino 111, 80131, Naples, Italy
Fax (+39)-081-6132277, e-mail m.moracci@ibp.cnr.it
[b] Dipartimento di Biologia Strutturale e Funzionale, Università di Napoli "Federico II", Via Cinthia 16, 80126, Naples, Italy

Received: February 2, 2005; Accepted: March 31, 2005

Abstract: In oligosaccharide synthesis the chemo-enzymatic approach is often chosen to obtain the desired glycosidic bond by exploiting the exquisite substrate specificity and stereoselectivity of glycosidases and glycosyltransferases. In recent years the invention of new enzymatic activities, named glycosynthases, created a third route for the enzymatic synthesis of oligosaccharides. Glycosynthases are mutant glycosidases, devoid of hydrolytic activity, that produce oligosaccharides with yields >80%. The glycosynthase technology is now well established and was greatly improved recently. Several groups experimentally proved that this methodology can be applied to several glycosidases and that the activity of existing glycosynthases can be significantly enhanced. In this review we briefly describe the characteristics of these synthetic tools and show a survey of the most recent findings. Methods for the prediction of the glycosynthetic potential of known glycosidases are also discussed.

Keywords: chemical complementation; directed evolution; kinetics; protein engineering; reaction mechanisms; transglycosylation

1 Introduction

The challenges that are faced in the synthesis of carbohydrates involve the presence on the monosaccharides of several functional groups and chiral centres. In these regards, chemical synthesis offers exceptional flexibility providing the means to generate any oligosaccharide, oligosaccharide analogue, or glycoconjugate. Chemical synthesis of oligosaccharides requires stereochemical and regiochemical control in glycosidic linkage formation by sequential protection-deprotection strategies of the functional groups present on the monosaccharides. These procedures often result in time-consuming manipulations and low yields, hindering the efficient production of oligosaccharides that is needed for biological testing.[1]

Enzymatic and chemo-enzymatic synthesis represent alternative tools to obtain the desired glycosidic bond by exploiting the exquisite substrate specificity and stereoselectivity of glycosidases and glycosyltransferases.[1,2] For instance, the use of a multi-enzymatic system mimicking the natural pathway in the Golgi apparatus allowed the synthesis of the (antithrombin III)-binding pentasaccharide in only six steps, outrunning the traditional chemical route that consists in 60 steps.[3]

Enzymatic oligosaccharide synthesis is performed by two main classes of catalysts: glycosidases and glycosyltransferases. Glycosidases are classified on the basis of their amino acidic sequence into about 100 families and 14 superfamilies or clans showing the same reaction mechanism and 3D structure (http://afmb.cnrs-mrs.fr/CAZY/index.html). Thanks to this valuable classification any new glycosidase can be assigned to an existing family and the residues involved in the catalytic activity or the 3D structure can be easily identified. The oligosaccharide synthesis promoted by glycosidases is accomplished either by displacing the equilibrium with a large excess of acceptor (thermodynamically controlled synthesis) or by using activated glycosyl donors (kinetically controlled transglycosylation) (Scheme 1A).[4] However, typically, *retaining* glycosidases give only modest synthetic yields (10–40%) precluding their exploitation in

 DOI: 10.1002/adsc.200505070 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Giuseppe Perugino (1971) studied Biology at the University of Naples "Federico II" where he graduated with honours in 1997 by making an experimental thesis on enzymology in the Institute of Protein Biochemistry, CNR. He received his PhD in 2003 in biochemistry and molecular biology and spent part of his training in the Department of Microbiology of the University of Wageningen (The Netherlands) under the supervision of Prof. W. de Vos. He is involved since the late 1990s in the study of glycosidases from hyperthermophiles and the chemo-enzymatic synthesis of oligosaccharides.

Beatrice Cobucci-Ponzano (1971) graduated with honours in 1997 at the University of Naples "Federico II" in Biology by making an experimental thesis in enzymology in the Institute of Protein Biochemistry, CNR. She received her PhD in 2002 in biochemistry and biophysics at the University of Padua (Italy) and then she joined the Institute of Protein Biochemistry as staff researcher. Her research interests include the biochemistry and molecular biology of extremophilic organisms and the enzymology of glycosidases from hyperthermophiles.

Mosè Rossi (1938) graduated with honours in 1961 at the University of Naples in Chemistry. In 1968 he was EMBO Fellow at MRC, Cambridge (UK) – Department of Molecular Biology in Prof. J. I. Harris' laboratory; in 1970–1972 he was P. H. S. Fellow at Stanford University in Prof. D. O. Woodword's laboratory at the Department of Biological Sciences, Stanford, California (USA). He has been senior scientist in the Institute of Protein Biochemistry, CNR, since 1969 and director of the institute from 1986 on and full professor of enzymology at the University of Naples "Federico II" from 1981. His primary research interests are in the field of the application of enzymes in biotechnology. He is one of the pioneers in the isolation and characterisation of enzymes and proteins from extremophilic organisms.

Marco Moracci (1962) graduated in Biology at the University of Naples in 1987 with an experimental thesis on enzymology. He was post-doc in the Department of Chemistry of Cambridge University (UK) under the supervision of Prof. A. R. Fersht. He joined the Institute of Protein Biochemistry, CNR, in 1992. He is senior researcher in this institution since 2000 and he teaches enzymology at the University of Naples "Federico II" since 2002. He is involved from late 1980s in the study of extremophilic organisms and the structure-function of hyperstable proteins. He developed the glycosynthase approach on hyperthermophilic glycosidases.

large-scale synthesis (Scheme 1A). Glycosyltransferases are classified as belonging to the Leloir or the non-Leloir pathway on the basis of their substrate specificity.[5] The former enzymes transfer the glycone moiety from a nucleotidyl sugar substrate to an acceptor, while the enzymes of the non-Leloir pathway use glycosyl phosphates as donors. Despite their high specificity and almost quantitative yields, glycosyltransferases are hampered by their scarce availability and by the high costs of the substrates.[1]

A third route for the enzymatic synthesis of oligosaccharides appeared in the last decade with the invention reported by Withers and colleagues in 1998 who described for the first time that a mutant of a β-glycosidase, devoid of hydrolytic activity, could produce oligosaccharides with yields $>80\%$. This new class of enzymes was named *glycosynthase*.[6]

The glycosynthase approach is now well established and has been described in detail in several reviews.[7] Nevertheless, recent papers disclosed several aspects

 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

of the glycosynthase mechanism and addressed the question of how the synthetic potential of an existing glycosynthase can be improved/expanded.

In this review we briefly describe the characteristics of these novel synthetic tools and show a survey of the most recent findings. Methods to predict the glycosynthetic potential of known glycosidases are also discussed.

2 Transformation of a Glycosidase into a Glycosynthase

The glycosynthase technology arose from mechanistic studies performed in Withers' laboratory on the β-glycosidase from *Agrobacterium* sp.[8] In particular, these authors showed that the modification of the catalytic nucleophile Glu358 into a residue unable to perform a nucleophilic attack (i.e., Ala or Gly) inactivates the enzyme, but small ions, like sodium azide, acting as external nucleophiles, restore the activity. The analysis of the products obtained by reactivation enables the identification of the residue of interest (for a review of the common methods for the analysis of the active site-residues in glycosidases, see [7a]).

A consequence of this modification is that the reaction mechanism of the enzyme was changed from retaining to inverting and it could use a glycosyl fluoride with opposite anomeric configuration to that of the normal substrate (α-glucosyl fluorides *vs.* β-D-glucosides).[8] Successively, Withers and colleagues showed that, in the presence of α-glucosyl fluorides and suitable acceptors, the nucleophile mutant synthesised oligosaccharides acting as a glycosynthase.[6]

From this first report, the glycosynthase technology was actively developed in other laboratories in the world leading to a variety of novel activities including other exo-glycosynthases, endo-glycosynthases, hyperthermophilic exo- and endo-glycosynthases, α-glycosynthases, mannosynthases, mannansynthases, and galactosynthases.[9]

The two main glycosynthase modes of action have been classified into inverting and retaining on the base of the anomeric configuration of donor substrate and of the final products (Scheme 1B and 1C). The mechanism proposed for endo-glycosynthases, likewise that of inverting exo-glycosynthases, employs α-oligosaccharidic fluorides as donors (Scheme 1D). Instead, retaining exo-glycosynthases synthesise oligosaccharides by using activated donor-sugars with the same anomeric configuration of the normal substrate (typically 2-nitrophenyl- (NP) or 2,4-dinitrophenyl-β-glucosides) in the presence of external ions such as sodium formate (Scheme 1C).[10] This external nucleophile mimics the natural active site carboxylate promoting the formation of a metastable glycosyl-formate intermediate, which has been isolated in one case,[11] and the attack of an acceptor (Scheme 1C).

A survey of the available glycosynthases is described in Table 1 in which we make a comparative list of the produced sugars and of the glycosidic bonds formed. Details of the different enzymes and of the reaction conditions can be found in the papers describing these studies. The known glycosynthases include 12 enzymes from 7 different families (Table 1) and can produce a variety of products at high yields, including mannose derivatives that are particularly difficult to synthesize by following the common chemical routes.[13] In addition, glycosynthases have been successfully used also for the solid-phase synthesis of sugars and glycoconjugates.[14] Exo-glycosynthase products include oligosaccharides ranging from di- to hexasaccharides. They can be new substrates for the study of glycosidases or can have several applications; for instance, branched (1–3)-β- and (1–6)-β-linked oligosaccharides, produced by the glycosynthase from *Sulfolobus solfataricus*, are the building blocks of the β-1,3–1,6-glucans that are recognised as elicitors of the defence response against pathogens in plants.[15]

Endo-glycosynthases produce oligosaccharides and polysaccharides depending on the reaction conditions.[9d,9h,16] One of these compounds, with a degree of polymerization of up to 12 consists of alternating (1–3)-β- and (1–4)-β-glucosyl units, representing a novel polysaccharide that forms platelets aggregating in spherulites of diameters ranging from 5 to 50 μm.[16] Thus, endo-glycosynthases appear to be valuable tools for the production of novel biomaterials with specific and controlled structures. Moreover, in the presence of different polymers, these enzymes may direct the synthesis of composite materials with interesting biomedical applications.[17]

3 Thioglycoligases and Thioglycosynthases

A recent achievement in the rational design of improved synthetic glycoside hydrolases is the preparation of thioglycoligases and thioglycosynthases (Scheme 2).[9l,18] In thioglycoligases the acid/base catalyst of two retaining β-glycosidases, namely the β-glucosidase from *Agrobacterium* and the β-mannosidase from *Cellulomonas fimi*, was changed in the non-nucleophile residue alanine producing mutants whose activities were severely affected.[9l] In the presence of activated dinitrophenyl-β-glycosides the mutants were re-activated by the addition to the reaction mixture of thio-sugar as acceptor. In these compounds the SH group is more nucleophilic than the OH group and does not require the assistance of the base catalyst (Scheme 2A). The resulting disaccharides containing *S*-glycosidic linkages were obtained with yields of 35–82%; consequently, the mutant enzymes were named *thioglycoligases*. These molecules have considerable interest as stable glycoside analogues

 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 1. List of available glycoside synthases.[a, b]

Organism	GH	Synthase	Products	Refs
Exo-glycosynthases				
Agrobacterium	1	β-glycosynthase	β-(1,3)-, β-(1,4)-di- and -trisaccharides	[6]
		β-glycosynthase	β-(1,4)-di-, tri- and -tetrasaccharides	[9c]
Sulfolobus solfataricus	1	β-glycosynthase	β-(1,3)-, β-(1,4)- and β-(1,6)-2-NP-disaccharidesβ-(1,3)- and β-(1,6)- branched tri- and -tetrasaccharides	[10a]
		β-glycosynthase	Glc-β-(1,3)-Glc-β-2-NP	[10a]
		β-glycosynthase	Gal-β-(1,3)-Xyl-β-4P	[12]
		β-glycosynthase	β-(1,3)-, β-(1,4)- and β-(1,6)-disaccharidesβ-(1,3)- and β-(1,6)- branched tri- and -tetrasaccharides	[10b]
Cellulomonas fimi	2	β-mannosynthase	β-(1,3)- and β-(1,4)-tri- to -hexasaccharides	[9e]
Schizosaccharomices pombe	31	α-glucosynthase	β-(1,4)- and β-(1,6)-4-NP-disaccharides	[9f]
Escherichia coli	2	β-galactosynthase	Gal-β-(1,6)-Glc-β-4-NP	[9i]
Thermosphaera aggregans	1	β-glycosynthase	β-(1,3)-, β-(1,4)- and β-(1,6)-2-NP-disaccharidesβ-(1,3)- and β-(1,6)- branched tri- and tetrasaccharides	[10b]
		β-glycosynthase	Glc-β-(1,3)-Glc-β-MU	[12]
		β-glycosynthase	Gal-β-(1,3)-Xyl-β-4P	[12]
Pyrococcus furiosus	1	β-glycosynthase	β-(1,3)-2-NP-disaccharides	[10b]
Endo-glycosynthases				
Bacillus licheniformis	16	β-glucansynthase	Glc-β-(1,3)-Glc-β-(1,4)-Glc-β-MU	[9b]
		β-glucansynthase	Glc-β-(1,4)-Glc-β-(1,3)-Glc-β-(1,4)-Glc-β-(1,4)-Glc-MU	[9g]
		β-glucansynthase	Gal-β-(1,4)-Glc-β-(1,3)-Glc-β-(1,4)-Glc-β-(1,4)-Glc-MU	[9g]
Humicola insolens	7	β-glucansynthase	β-(1,4)-tri-, -tetrasaccharides	[9d]
Hordeum vulgare	17	β-glucansynthase	β-(1,3)-glucan	[9h]
		β-glucansynthase	β-(1,3)-S-glucan	[9j]
Cellvibrio japonicum	26	β-mannansynthase	β-(1,4)-tri- to -heptasaccharides	[9k]
Pyrococcus furiosus	16	β-glucansynthase	Glc-β-(1,3)-Glc-β-(1,4)-Glc-β-MUGlc-β-(1,3)-Glc-β-(1,3)-Glc-β-(1,3)-Glc-β-MU	[9m]
Thioglycoligases				
Agrobacterium	1	β-thioglucoligase	Glc-β-S-(1,4)-Glc-β-4-NPGlc-β-S-(1,4)-Xyl-β-4-NP	[9l]
Cellulomonas fimi	2	β-thiomannoligase	Man-β-S-(1,4)-Glc-β-4-NPMan-β-S-(1,4)-Xyl-β-4-NP	[9l]

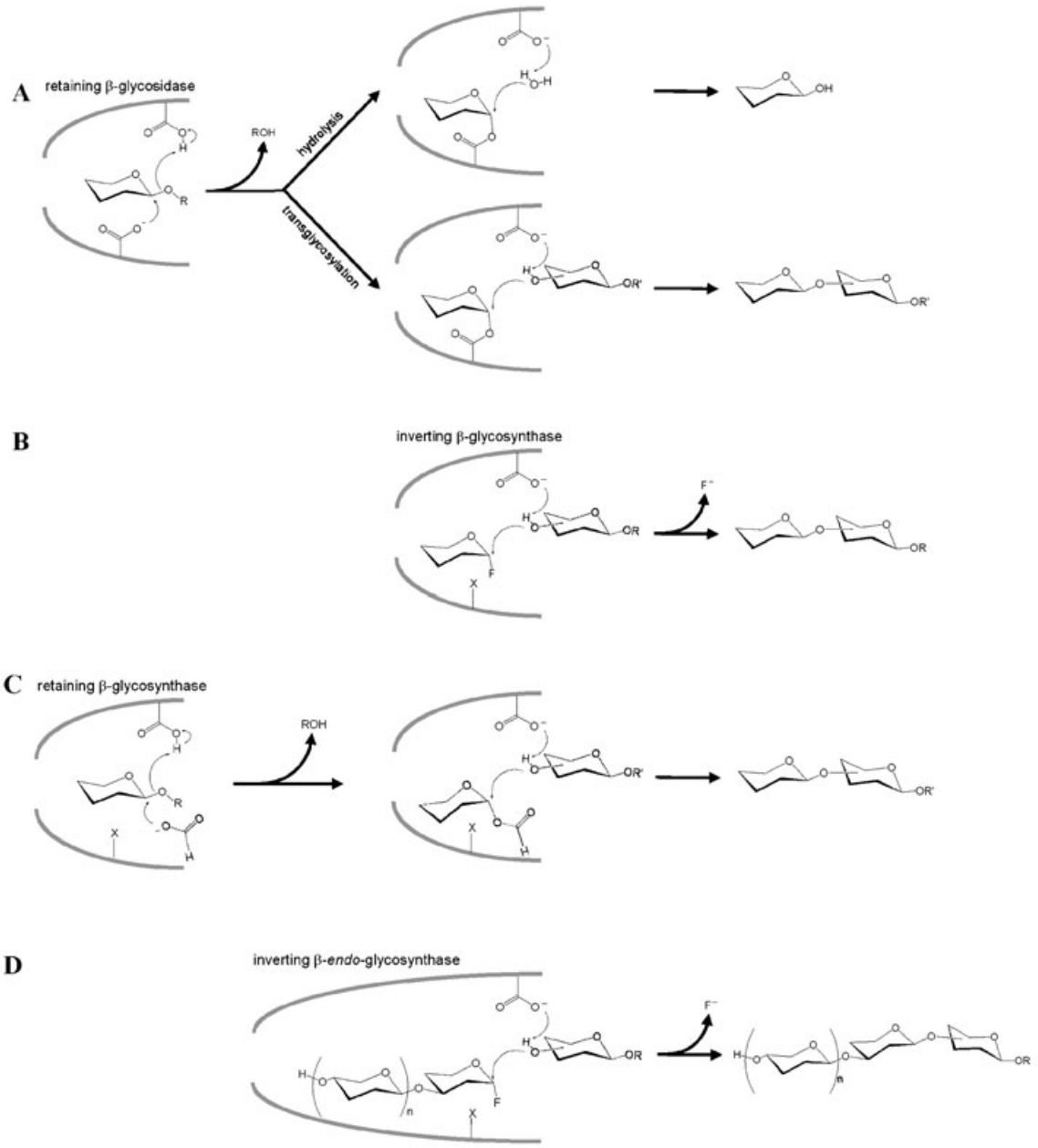
[a] Glycoside hydrolase family according to http://afmb.cnrs-mrs.fr/CAZY/index.htlm.
[b] Abbreviations: 4P, 4-penten-1-yl-α-D-xylopyranoside; Gal, D-galactopyranoside; Glc, D-glucopyranoside; Glc-β-(1,3)-Glc, D-glucosyl-β-(1,3)-D-glucopyranoside; Man, D-mannopyranoside; Xyl, D-xylopyranoside.

to be used as competitive inhibitors and ligands for the formation of complexes for X-ray crystallography.[19]

More recently, Withers' group demonstrated that a thioglycoligase depleted also of the catalytic nucleophile is still catalytically competent.[18] The double mutant Glu171Ala/Glu358Gly of the β-glycosidase from *Agrobacterium* incubated in the presence of α-glucosyl fluoride and 4-nitrophenyl-4-thio-β-D-glucopyranoside as donor and acceptor, respectively, produced the expected thio-linked 4-nitrophenyl disaccharide (Scheme 2B). Thus, the double mutant used in this way was named *thioglycosynthase*. The authors reported that one of the advantages of this new methodology over the thioglycoligase described above is that the double mutant is completely inactive (because of the absence of the nucleophile residue) and does not hydrolyse the donor. In addition, the thioglycosynthase uses glycosyl fluoride donors that are more easily synthesised than dinitrophenyl glycosides.[18] However, the yields of the reaction did not increased significantly (~45–50%) and the k_{cat} was 100- to 400-fold lower than that measured on the single mutants Glu171Ala (thioglycoligase) and Glu358Gly (glycosynthase), respectively.[18] Presumably, the double mutant worked as a scaffold to position the reagents and the reduced activity could be explained by the lack of some unidentified chemical catalysis provided by the remaining carboxylate in the single mutants.

The catalytic competence of glycosidase mutants lacking the catalytic carboxylic groups was demonstrated also by the group of Planas on the 1,3–1,4-β-glucanase from *B. licheniformis*.[120] In this case, the activity of the double mutant Glu134Ala/Glu138Ala was chemically rescued by sodium azide in the presence of the donor Glcβ4Glcβ3Glc-2,4DNP and it produced α-glycosyl azide. No transglycosylation products were obtained in this case.

 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim



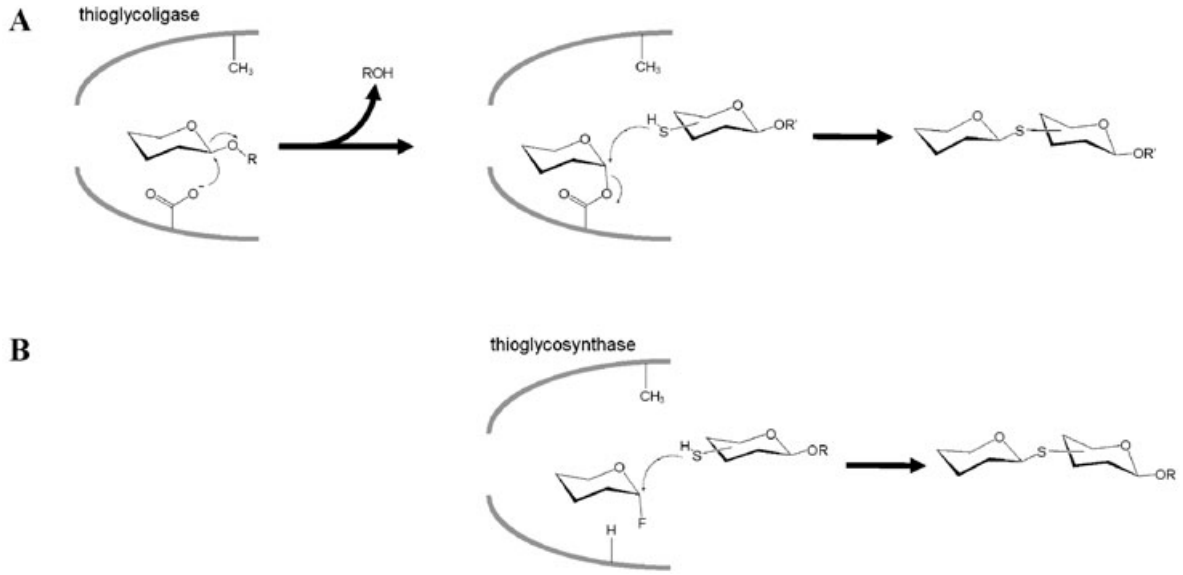
Scheme 1. Reaction mechanisms of glycosidases and glycosynthases. R = R′ indicates the aglycone part of the sugars; X in the schematic active site of the enzyme indicates the mutated nucleophile residue.

4 Molecular Bases of the Glycosynthase Mechanism

A crucial point for the exploitation of the synthetic potential of these new enzymes is the detailed knowledge of their reaction mechanism; in this way their activity can be enhanced and adapted to the synthesis of specific products. These issues have been addressed recently in different laboratories.

In inverting and retaining glycosynthases (Scheme 1B and 1C) the enzymatic catalytic machinery consists of a single carboxylic group that, in the parental glycosidases, works as acid/base catalyst, and the products of the reaction cannot be hydrolysed by the enzyme. However, the two mechanisms differ substantially in the number of steps of the reaction and in the catalytic function of the carboxylic group in the active site. In inverting glycosynthases, only the de-glycosylation step is present, because the α-glycosyl-F mimics the covalent intermediate observed in retaining glycosidases (Scheme 1B). Recently, Faijes and colleagues showed that in the inverting endo-glycosynthase from *B. licheniformis* the carboxylic group in the active site (Glu138) acts as a general base.[20] This was experimentally demonstrated by analysing the

 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 2. Reaction mechanisms of thioglycoligase and thioglycosynthase. R = R′ indicates the aglycone part of the sugars.

pH dependence of the glycosynthase activity, the kinetics of enzyme inactivation by a carboxyl-specific reagent, and the reactivation of a double mutant in which both the nucleophile and the acid/base were changed in non-nucleophilic residues (see also above). These studies revealed that, as expected for a general base catalysis, the pK_a of the Glu138, which is 7.0 in the wild-type enzyme (general acid catalysis), drop to 5.3 in the glycosynthase.

In retaining glycosynthases, likewise parental glycosidases (Scheme 1A), the reaction takes place through two steps, a glycosylation and de-glycosylation, through the formation of a metastable glycosyl-formate intermediate (Scheme 1C). In this case, we have recently demonstrated that at acidic conditions the hyperthermophilic glycosynthases showed enhanced activity, possibly because the carboxylic group present in the active site acts as a general acid catalyst.[10b] In particular, we observed that the exo-glycosynthases from the hyperthermophilic archaea *S. solfataricus*, *Pyrococcus furiosus*, and *Thermosphaera aggregans*, showed an up to 17-fold increase in the reaction rate at pH 3.0 in 50 mM sodium formate if compared to the same enzymes assayed at pH 6.5 in 2 M sodium formate.[10b] Remarkably, the k_{cat} values measured at acidic conditions were similar to that of the corresponding wild-type enzymes assayed at optimal conditions and allowed the synthesis of 4-methylumbelliferyl disaccharides.[12] We proposed that this exceptional reactivation was the result of the restoration of the correct ionisation state of the carboxylic group that in the active site acts as general acid catalyst. In fact, removal of the catalytic nucleophile in retaining glycosidases causes a downward shift in the pK_a of the acid/base catalyst;[21] therefore, this group is converted in the ionised form and it performs the first step of the reaction less efficiently (Scheme 1C). Conversely, in the presence of sodium formate buffer pH 3.0–6.0 the Glu residue could adopt the protonated form assisting the departure of the leaving group of the donor.[10b] The high reaction rates observed suggest that for retaining glycosynthases the first step of the reaction (glycosylation of the enzyme) is rate limiting and that, hence, they requires general acid catalysis for optimal activity. It is worth noting that, at acidic conditions, we could use less activated and commercially available donor substrates like 2-nitrophenyl and 4-nitrophenyl glucosides. This is a typical example of an application of hyperstable enzymes that can resist extreme operational conditions like pH 3.0 and 65 °C.[10b,12]

The resolution of the 3D-structure of the endo-glycosynthase from *Humicola insolens* (Cel7B) cast new light on the glycosynthase mechanism.[22] The enzyme shows a long active-centre groove that can accommodate substrates of up to eight monosaccharides in which four subsites (-2 to $+2$ by following the nomenclature reported in [23]) significantly contribute to catalysis. The nucleophile of the reaction (Glu197) is located at the centre of the active centre and in the glycosynthase variants Glu197Ala/Ser no significant changes were observed suggesting that the mutants maintain the same active site structure of the wild type (Figure 1). The mutant Glu197Ser, which is the most active glycosynthase variant, was crystallised with both lactose and cellobiose and the resolution beyond 1.4 Å allowed detailed analysis of the interactions with the substrate. In the donor subsites (-2 and -1) cellobiose and lactose bind equally well as expected from the kinetic analysis of the enzyme (Figure 1).[22] More interestingly, in the acceptor subsites ($+1$ and $+2$) lactoside and cellobioside complexes are not equivalent since the former is slightly dis-

 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

placed by approximately 1.7 Å and binds in a sort of +1.5 to +2.5 position (Figure 1). These explains why lactosides are poor acceptors and are responsible for substrate inhibition at high donor concentrations.[22]

One of the most interesting findings described in this study is on the mode of action of the Glu197Ser mutant. Ducros and colleagues found that the O4 of the acceptor lies 4.4 Å from the C1 of the donor while the O3 group is at only 3.3 Å; this is quite surprising since the glycosynthase promotes the formation of β-1,4 bonds and products containing β-1,3 bonds have never been observed.[9d] Comparing the structure of the Cel7B *H. insolens* glycosynthase with that of *Fusarium oxysporum*[24] the authors suggest that the regioselectivity for 1,4 bond formation could be regulated not only by distance criteria, but also by the angle of the nucleophilic attack. In fact, the potential angle of attack of O3 is sub-optimal whereas the O4–C1–O1 angle is 173°, close to the optimal 180° required for the nucleophilic attack (Figure 1).

5 Strategies for the Improvement of the Glycosynthases Activity

The search for glycosynthases with improved activity is of utmost importance not only because they act more rapidly, but also because, as a direct consequence, they can transfer a variety of donors to a much wider array of acceptors on a useful timescale, making them considerably more versatile synthetic tools. For instance, we showed that the hyperthermophilic retaining glycosynthases that had increased activity in acid conditions (see above) widened their specificity towards galactoside and xyloside donors[10b] and accept a great deal of acceptors.[12]

The nature of the amino acid substituting the nucleophile in glycosynthases has a significant effect on the activity of the enzyme. The non-nucleophilic residue has to generate sufficient space to accommodate the axial anomeric fluoride of the donor in inverting glycosynthases and the α-formyl glucoside intermediate in retaining enzymes (Scheme 1B, 1C). In inverting glycosynthases, Gly[9f,9h,9k] and Ser mutations[9c,9e,9i] gave best results. In particular, studies of directed evolution revealed that in the glycosynthase from *Agrobacterium* the Gly, Ser, Ala, and Cys mutants have a ratio of rate constants of 102:44:2:1, respectively.[25] It has been proposed that the increased glycosylation activity in the Ser mutant originates from a stabilizing interaction between the Ser hydroxy group and the departing fluorine.[9c] More recently, the inspection of the 3D-structure of the glycosynthase from *H. insolens* suggested that, most probably, this occurs at the transition state.[22] However, Gly mutants often resulted to be better synthases than the Ser mutants. In the glycosynthase from *C. japonicus* this was tentatively explained with the observation that in the 3D-structure of the Gly mutant a single water molecule replacing the nucleophile formed a hydrogen bond to the O1 of the donor.[9k]

Similarly, Gly mutants were also found to be efficient in retaining glycosynthases: in the enzyme from *S. solfataricus* the rate of hydrolysis of the 2,4-dinitrophenyl-β-glucoside donor followed the ratio 59:9:1 for the Gly, Ala, and Gln mutants.[9a]

Glycosynthases with improved activity have been recently obtained by directed evolution. This method consists in generating a large population of protein variants that is screened for the desired function.[26] Recently, Withers' group reported that additional mutations even increased the activity of this glycosynthase.[27] In this case, mutations were introduced randomly in the sequence of the gene encoding for the Glu358Gly glycosynthase; then, by using an "on-plate" screening two mutants were identified in the mutant library. This method is shown in Figure 2A: briefly, it consists in coupling to the glycosynthase catalysed reaction an assay involving an endocellulase. The mutant library co-express Cel5A, hence the glycosynthase activity can be moni-

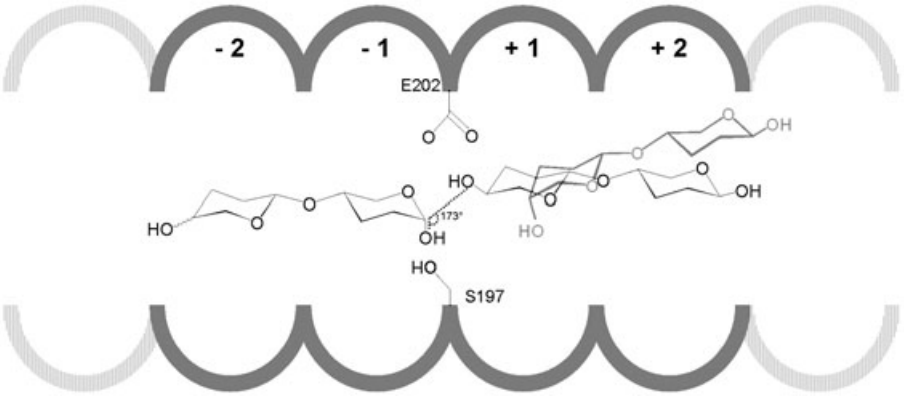


Figure 1. Schematic view of the active site of *H. insolens* glycosynthase Cel7B. The binding sites and the ligands identified by X-ray crystallography are shown as reported by Ducros and colleagues.[22] The sugar in the −2 to −1 sites indicates both lactose and cellobiose; lactose in the acceptor sites +1 to +2 is shown in grey.

 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

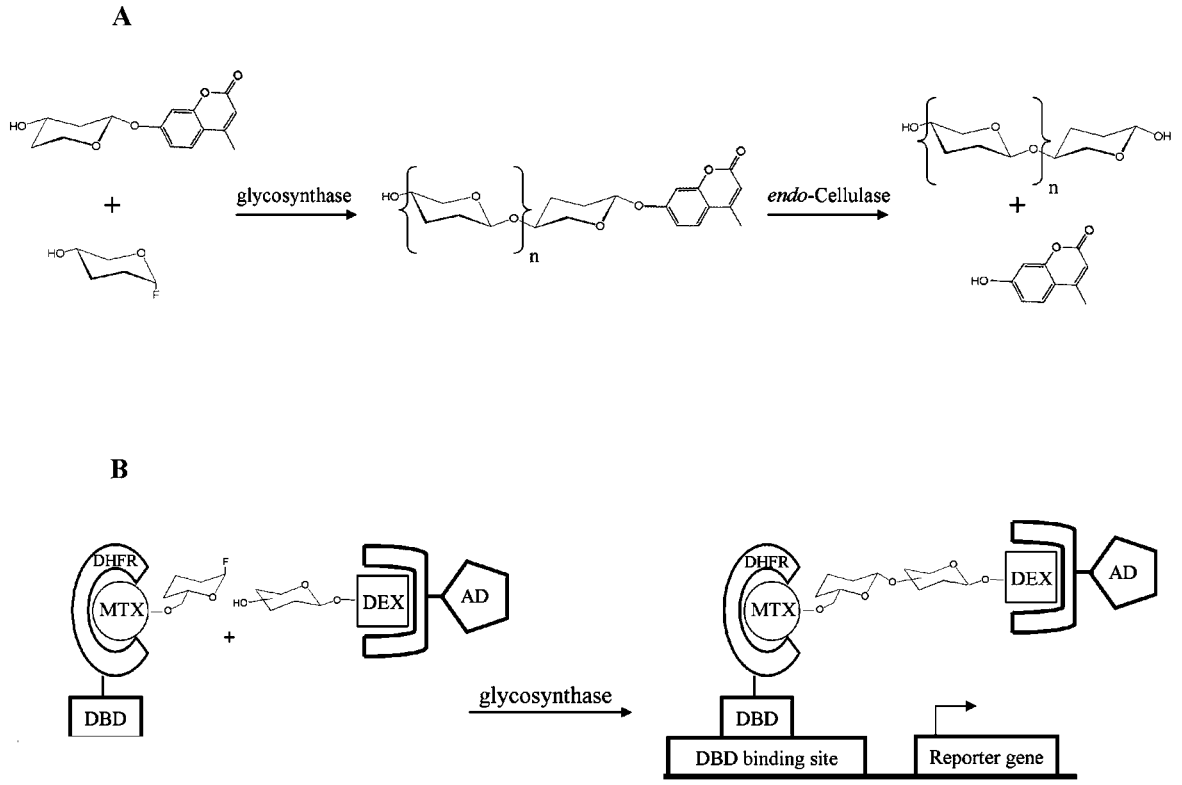


Figure 2. Method of selection of glycosynthase mutants with enhanced activity. (**A**) Scheme of the glycosynthase-cellulase coupled assay. The substrates of the glycosynthase are α-glucosyl F and 4-methylumbelliferyl-β-glucoside (MU-Glc) as donor and acceptor, respectively. Mutant clones expressing an efficient glycosynthase promote the synthesis of MU-oligosaccharides that are substrate of the β-1,4 glucanase from *C. fimi* (Cel5A) releasing the fluorescent methylumbelliferone; (**B**) yeast three-hybrid assay exploited in the "chemical complementation" method (see text for the details of the two methods).

tored directly on a plate simply by adding the substrates of the glycosynthase. With this straightforward technique the authors identified with two rounds of mutagenesis a multiple mutant (Ala19Thr/Glu358Gly/Gln248Arg/Met407Val) showing a catalytic efficiency 27-fold higher than that of the parental glycosynthase Glu358Gly. To understand the molecular reasons of the improved activity of the new mutant the authors inspected a model generated using known structures of homologous enzymes. From this analysis it was proposed that the mutations Ala19Thr and Met407Val improve the activity by contributing to the transition state stabilisation and by increasing the fit in the active site of the α-glucosyl fluoride substrate, respectively. Intriguingly, no straightforward interpretations could be proposed for Gln248Arg since the residue is located on the protein surface. This study further confirms the usefulness of the directed evolution methods: the mutations described could be hardly predicted by a rational design approach.

Recently, a novel directed evolution approach, based on the yeast three-hybrid system assay named "chemical complementation", was applied to the Cel7B glycosynthase from *H. insolens* [28] after application to other enzymes.[29] Briefly, the case described for the Cel7B glycosynthase utilises a three-hybrid assay based on two fusion proteins: the hormone binding domain of the glucocorticoid receptor (GR) connected to a transcription activation domain (AD), and a dihydrofolate reductase (DHFR) fused to a DNA-binding domain (DBD) (Figure 2B). The association of AD to DBD allows the expression of the gene *LEU2* producing the growth in the absence of leucine. In addition, methotrexate (Mtx) and dexamethasone (Dex) specifically bind to the GR and DHFR proteins, respectively (Figure 2B). Without glycosynthetic activity the fusion proteins GR-AD and DHFR-DBD cannot dimerise, thus, the gene *LEU2* is not expressed and cannot complement yeast cells that cannot grow. When the glycosynthase catalyses *in vivo* the synthesis of the glycosidic bond between an Mtx-disaccharide-fluoride donor (Mtx-Lac-F) and a Dex-disaccharide acceptor (Dex-Cel) the dimerisation of the fusion proteins occurs directing the expression of the *LEU2* gene and the growth without leucine (Figure 2B). The selection was used for the identification of Cel7B mutants with improved activity obtained by directed evolution. In particular, saturation mutagenesis on the Glu197 site randomised to all 20 amino acids and subsequent chemical complementation produced three mutants Ala, Gly and Ser giving growth advantage

 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

to yeast.[28] The most efficient evolved variant was Glu197Ser in agreement with the results of Ducros and colleagues described above.[22]

The mutation identified in the Cel7B glycosynthase with the chemical complementation method was not novel, however, it offers the advantage to analyse a huge number of clones since it is based on a stringent selection (growth advantage *vs.* coupled assay). Practical limits, which may affect the efficiency of the complementation, can be the design of acceptors and donors suitable for the three-hybrid assay and the expression in yeast of the glycosynthase of interest.

6 How to Predict the Gycosynthetic Potential of a Glycosidase

From that stated in the previous paragraphs, the glycosynthase technology is now well established and it has improved greatly in the last five years. Several groups experimentally proved that the same methodology can be applied to a great number of glycosidases and that the activity of existing glycosynthases can be significantly improved by using chemistry[10b] or directed evolution.[27,28] Notwithstanding these issues, to obtain a new glycosynthase is not trivial and several glycosidases mutated in the nucleophile did not show glycosynthase activity, i.e., the β-xylosidase from *Thermosporum saccharolyticum*, β-galactosidase from *Bacillus circulans*,[22] the α-xylosidase (B. Di Lauro, B. C.-P. and M. M. unpublished results) and the α-L-fucosidase from *S. solfataricus*.[30] Consequently, it is of utmost importance to try to understand what makes an efficient glycosynthase. This problem was addressed by Ducros and collegues by analysing the effect on glycosidases of fluoro-glycosides mechanism-based inhibitors. These molecules, in which the hydroxy group at C2 is replaced with a fluorine atom, act as covalent inhibitors of glycosides by trapping the intermediate 2-fluoroglycosyl enzyme.[31] The inhibited enzyme is often catalytically competent and the intermediate can be resolved by an acceptor glycoside or by water; the rate constants of transfer and hydrolysis are k_{trans} and k_{H2O}, respectively.[22] Noticeably, with this method Blanchard and Withers identified a number of sugars able to reactivate six different glycosidases and therefore acting as efficient acceptors. [32]

Ducros and collegues proposed that glycosidases that yield efficient glycosynthases have two reactivation characteristics: first, a high rate constant of transfer to an acceptor (they estimate $k_{trans} > 10^{-2}\ min^{-1}$); second, a better transferring activity to an acceptor rather than to water ($k_{trans}/k_{H2O} > 20$).[22] The rational observation behind this prediction is that the 2-fluoro-glycosyl enzyme intermediate may resemble the α-glycosyl fluoride bound to the inverting glycosynthases. By analysing the kinetic behaviour of the parental glycosidases of existing glycosynthases and of enzymes whose mutants did not show glycosynthase activity the authors found some support to this predictive method.

Interestingly, glycosidases which perform efficiently transglycosylation reactions often did not act as glycosynthases upon mutation. This raises the question of how the "natural" transglycosylating activity of a glycosidase is lost after mutation of the nucleophile. There is no clear-cut answer to this question, however, the reactivity of the donors and the binding constants of the acceptors play an important role in both cases. Moreover, as stated by Ducros and collegues, the acceptors can improve the glycosynthetic activity by stabilising the transition state and reducing the transfer activity to water that is detrimental to the synthesis of the products.[22] Therefore, the binding interactions of the acceptors in the $+1$ site are important for the catalysis of glycosynthases.

These studies further confirm that a detailed kinetic characterisation, including the use of mechanism-based inhibitors, is essential prior to the preparation of a glycosynthase. The validity of the methods described needs to be tested by analysing a larger number of glycosidases.

Acknowledgements

This work was supported by MIUR project "Folding di proteine: l'altra metà del codice genetico" RBAU015B47 006 and by the project entitled "Nuove attività enzimatiche per la sintesi di oligosaccaridi di interesse biotecnologico" from Regione Campania L.R.N.5 del 28.03.2002. The IBP-CNR belongs to the Centro Regionale di Competenza in Applicazioni Tecnologico-Industriali di Biomolecole e Biosistemi.

References

[1] P. Sears, C. H. Wong, *Science*. **2001**, *291*, 2344–2350.
[2] a) D. H Crout, G. Vic, *Curr. Opin. Chem. Biol.* **1998**, *2*, 98–111; b) C. R. Bertozzi, L. L. Kiessling, *Science*, **2001**, *291*, 2357–2364.
[3] B. Kuberan, M. Z. Lech, D. L. Beeler, Z. L. Wu, *Nat. Biotechnol.* **2003**, *21*, 1343–1346.
[4] a) D. E. Koshland Jr. *Biol. Rev.* **1953**, *28*, 416–436; b) K. M. Koeller, C. H. Wong, *Nat. Biotechnol.* **2000**, *8*, 835–841.
[5] Y. Bourne, B. Henrissat, *Curr. Opin. Struct. Biol.* **2001**, *11*, 593–600.
[6] L. F. Mackenzie, Q. Wang, R. A. J. Warren, S. G. Withers, *J. Am. Chem. Soc.* **1998**, *120*, 5583–5584.
[7] a) H. D. Ly, S. G. Withers, *Annu. Rev. Biochem.* **1999**, *68*, 487–522; b) G. Perugino, A. Trincone, M. Rossi, M. Moracci, *Trends Biotechnol.* **2004**, *1*, 31–37; c) M. Jahn, S. G. Withers, *Biocat. Biotrans.* **2003**, *21,* 159–166; d) A. Planas, M. Faijes, *Afinidad* **2002**, *59*, 295–313; e) S. J. Williams, S. G. Withers, *Austral. J. Chem.* **2002**, *55*, 3–12; f) D. L. Jakeman, S. G. Withers, *Trends in Glycosci. Glycotech.* **2002**, *14*, 13–25; g) G. J. Davies, S. J. Charnock, B.

 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Henrissat, *Trends in Glycosci. Glycotech.* **2001**, *13*, 105–120; h) M. Moracci, A. Trincone, M. Rossi, *J. Mol. Cat. B: Enzymatic* **2001**, *11*, 155–163; i) S. G. Withers, *Can. J. Chem.* **1999**, *77*, 1–11.

[8] Q. Wang, R. W. Graham, D. Trimbur, R. A. J. Warren, S. G. Withers, *J. Am. Chem. Soc.* **1994**, *116*, 11594–11595.

[9] a) M. Moracci, A. Trincone, G. Perugino, M. Ciaramella, M. Rossi, *Biochemistry* **1998**, *37*, 17262–17270; b) C. Malet, A. Planas, *FEBS Lett.* **1998**, *440*, 208–212; c) C. Mayer, D. L. Zechel, S. P. Reid, R. A. J. Warren, S. G. Withers, *FEBS Lett.* **2000**, *466*, 40–44; d) S. Fort, V. Boyer, L. Greffe, G. J. Davies, O. Moroz, L. Christiansen, M. Schülein, S. Cottaz, H. Driguez, *J. Am. Chem. Soc.* **2000**, *122*, 5429–5437; e) O. Nashiru, D. L. Zechel, D. Stoll, T. Mohammadzadeh, R. A. J. Warren, S. G. Withers, *Angew. Chem. Int. Ed.* **2001**, *40*, 417–420; f) M Okuyama, H. Mori, K. Watanabe, A. Kimura, S. Chiba, *Biosci. Biotechnol. Biochem.* **2002**, *66*, 928–933; g) J. K. Fairweather, M. Faijes, H. Driguez, A. Planas, *ChemBioChem* **2002**, *3*, 866–873; h) M. Hrmova, T. Imai, S. J. Rutten, J. K. Fairweather, L. Pelosi, V. Bulone, H. Rodriguez, G. B. Fincher, *J. Biol. Chem.* **2002**, *277*, 30102–30111; i) D. L. Jakeman, S. G. Withers, *Can. J. Chem.* **2002**, *80*, 866–870; j) J. K. Fairweather, M. Hrmova, S. J. Rutten, G. B. Fincher, H. Driguez, *Chemistry* **2003**, *9*, 2603–2610; k) M. Jahn, D. Stoll, R. A. J. Warren, L. Szabò, P. Singh, H. J. Gilbert, V. M. A. Ducros, G. J. Davies, S. G. Withers, *Chem. Commun.* **2003**, *12*, 1327–1329; l) M. Jahn, J. Marles, R. A. J. Warren, S. G. Withers, *Angew. Chem. Int. Ed.* **2003**, *42*, 352–354; m) J. van Lieshout, M. Faijes, J. Nieto, J. van der Oost, A. Planas, *Archaea* **2004**, *1*, 285–292.

[10] a) A. Trincone, G. Perugino, M. Rossi, M. Moracci, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 365–368; b) G. Perugino, A. Trincone, A. Giordano, J. van der Oost, T. Kaper, M. Rossi, M. Moracci, *Biochemistry* **2003**, *42*, 8484–8493.

[11] J. L. Viladot, F. Canals, X. Batllori, A. Planas A. *Biochem J.* **2001**, *355*, 79–86.

[12] A. Trincone, A. Giordano, G. Perugino, M. Rossi, M. Moracci, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4039–4042.

[13] a) J. Alais, S. David, *Carbohydr. Res.* **1990**, *201*, 69–77; b) W. Gunther, H. Kunz, *Carbohydr. Res.* **1992**, *228*, 217–241.

[14] J. F. Tolborg, L. Petersen, K. J. Jensen, C. Mayer, D. L. Jakeman, R. A. J. Warren, S. G. Withers, *J. Org. Chem.* **2002**, *67*, 4143–9.

[15] a) K. Skriver, F. L. Olsen, J. C. Rogers, J. Mundy, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 7266–7270; b) B. Duvic, K. Soderhall, K. *Eur. J. Biochem.* **1992**, *207*, 223–228; c) F. Coté, M. G. Hahn, M. G. *Plant Mol. Biol.* **1994**, *26*, 1379–1411.

[16] M. Faijes, T. Imai, V. Bulone, A. Planas A. *Biochem. J.* **2004**, *380*, 635–641.

[17] A. G. Watts, S. G. Withers, *Biochem. J.* **2004**, *380*, 9–10.

[18] M. Jahn, H. Chen, J. Mullegger, J. Marles, R. A. J. Warren, S. G. Withers, *Chem. Commun.* **2004**, *3*, 274–275.

[19] a) M. Czjzek, M. Cicek, V. Zamboni, W. P. Burmeister, D. R. Bevan, B. Henrissat, A. Esen, *Biochem. J.* **2001**, *354*, 37–46; b) M. Hrmova, J. N. Varghese, R. De Gori, B. J. Smith, H. Driguez, G. B. Fincher, *Structure* **2001**, *9*, 1005–1016.

[20] M. Faijes, X. Perez, O. Perez, A. Planas, *Biochemistry* **2003**, *42*, 13304–13318.

[21] L. P. McIntosh, G. Hand, P. E. Johnson, M. D. Joshi, M. Korner, L. A. Plesniak, L. Ziser, W. W. Wakarchuk, S. G. Withers, *Biochemistry* **1996**, *35*, 9958–9966.

[22] V. M. Ducros, C. A. Tarling, D. L. Zechel, A. M. Brzozowski, T. P. Frandsen, I. von Ossowski, M. Schulein, S. G. Withers, G. J. Davies, *Chem. Biol.* **2003**, *10*, 619–628.

[23] G. J. Davies, K. S. Wilson, B. Henrissat, *Biochem. J.* **1997**, *321*, 557–559.

[24] G. Sulzenbacher, H. Driguez, B. Henrissat, M. Schulein, G. J. Davies, *Biochemistry* **1996**, *35*, 15280–15287.

[25] C. Mayer, D. L. Jakeman, M. Mah, G. Karjala, L. Gal, R. A. Warren, S. G. Withers, *Chem. Biol.* **2001**, *8*, 437–443.

[26] a) E. T. Farinas, T. Bulter, F. H. Arnold, *Curr. Opin. Biotechnol.* **2001**, *12*, 545–551; b) J. A. Kolkman, W. P. Stemmer, *Nat. Biotechnol.* **2001**, *19*, 423–428; c) H. Tao, V. W. Cornish, *Curr. Opin. Chem. Biol.* **2002**, *6*, 858–864.

[27] Y. W. Kim, S. S. Lee, R. A. Warren, S. G. Withers, *J. Biol. Chem.* **2004**, *279*, 42787–42793.

[28] H. Lin, H. Tao, V. W. Cornish, *J. Am. Chem. Soc.* **2004**, *126*, 15051–15059.

[29] a) K. Baker, C. Bleczinski, H. Lin, G. Salazar-Jimenez, D. Sengupta, S. Krane, V. W. Cornish, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16537–16542. b) D. Sengupta, H. Lin, S. D. Goldberg, J. J. Mahal, V. W. Cornish, *Biochemistry* **2004**, *43*, 3570–3581.

[30] B. Cobucci-Ponzano, A. Trincone, A. Giordano, M. Rossi, M. Moracci, *Biochemistry* **2003**, *42*, 9525–9531.

[31] S. J. Williams, S. G. Withers, *Carbohydr Res.* **2000**, *327*, 27–46.

[32] J. E. Blanchard, S. G. Withers, *Chem. Biol.* **2001**, *8*, 627–622.

 © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim