

# Extremophilic carbohydrate active enzymes (CAZymes)

## Abstract

Carbohydrate active enzymes (CAZymes) are a large class of enzymes, which build and breakdown the complex carbohydrates of the cell. On the basis of their amino acid sequences they are classified in families that show conserved catalytic mechanism, structure, and active site residues, but may conflict each other in substrate specificity. To solve the conflict each other in substrate specificity, kinetic studies are essential for categorized in glycoside hydrolase (GH). Among GH family 1 enzymes, some of them are from hyperthermophile origins (i.e. *Sulfolobus*, *Pyrococcus*, etc). Recently, the advantage of hyperthermophile enzyme has been studied and commercialized. The main advantage of hyperthermophilic enzyme is that in chemical reactions involving organic solvents, the decrease in viscosity and increase in diffusion coefficient that is achieved at elevated temperatures result in higher reaction rates and low contamination. This review provides up-to-date information on b-glycosidase and the significance of this enzyme.

**Keywords:** enzymes, chemical reactions, amino acid, organisms, glycosides

Volume 7 Issue 1 - 2017

Sunghoon Park,<sup>1,2</sup> Byonghoon Lee,<sup>1</sup>  
Kwanhwa Park<sup>3</sup>

<sup>1</sup>Departments of Food Science and Microbiology/Immunology, McGill University, Canada

<sup>2</sup>Research Institute of Food and Biotechnology, SPC Group, Korea

<sup>3</sup>Department of Food Science, Sangmyung University, South Korea

**Correspondence:** Byonghoon Lee, Departments of Food Science and Microbiology/Immunology, McGill University, Montreal, Quebec H3A 2B4, Canada, Email byong.lee@mail.mcgill.ca

Received: March 29, 2017 | Published: September 21, 2017

## Introduction

### Carbohydrate active enzymes (CAZymes)

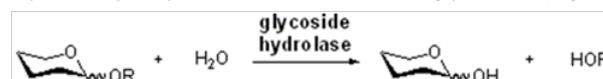
Carbohydrate active enzymes (CAZymes) are a large class of enzymes, which build and breakdown the complex carbohydrates of the cell. On the basis of their amino acid sequences they are classified in families that show conserved catalytic mechanism, structure, and active site residues, but may conflict each other in substrate specificity. The CAZymes provides a continuously updated list of the glycoside hydrolase families, GHs. This group of enzymes is classified based on functional similarity, but today they are classified into 108GHs on the basis of amino acid sequence similarity. Despite their similarities to enzymes with known functions, their primary functions are still unclear. Based on these criteria,  $\beta$ -galactosidase activities are now divided into four different families: GH1, GH2, GH35 and GH42, among which the better studied GH2 includes  $\beta$ -galactosidase from *Escherichia coli*, *Aspergillus*, *Bacillus megatherium*, and *Sulfolobus solfataricus*, while those from thermophilic, psychrophilic and halophilic organisms belong to GH42. Lactase is often confused as an alternate name for  $\beta$ -galactosidase, but it is actually simply a sub-class (small subunit) of  $\beta$ -galactosidase.

( $\beta$ -D-galactoside galactohydrolase, EC 3.2.1.23) that catalyses hydrolysis of the galactosyl moiety from non-reducing termini of oligosaccharides or from glycosides. Most genes encoding GH42 enzymes are from prokaryotes that are unable to grow on lactose as a sole carbon source and at least two GH42  $\beta$ -galactosidase do not cleave lactose *in vitro*. The determination of growth on lactose can be complicated by the multiple  $\beta$ -galactosidases because not all of the  $\beta$ -galactosidase is acting as lactases *in vitro*. B-Galactosidase hydrolyses the b-1, 4-D-galactosidic linkage of lactose, as well as those of related chromogens, *o*-nitrophenyl- $\beta$ -D-galactopyranoside (ONPG), *p*-nitrophenyl- $\beta$ -D-galactopyranoside (PNPG) and 6-bromo-2-naphthyl-galactopyranoside (BNG). This enzyme has been purified

and characterized from various sources, including plants, animals, and many microorganisms. In humans, lactase is present predominantly along the brush border membrane of the differentiated enterocytes lining the villi of the small intestine. Lactase is essential for digestive hydrolysis of lactose in milk. Deficiency of the enzyme causes lactose intolerance. Several 3D structures are available for the GH1 and GH2  $\beta$ -galactosidase for which the catalytic residues have experimentally been determined, while three 3D structures of  $\beta$ -galactosidase from *E. coli*, *Penicillium* sp and *Thermus thermophilus* A4 are available for both the GH35 and GH42 families.

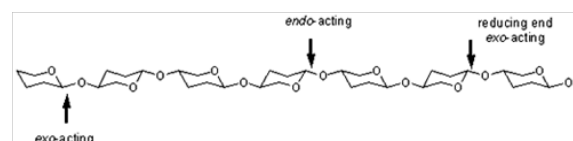
### Glycoside hydrolases

Glycoside hydrolases are enzymes that catalyze the hydrolysis of the glycosidic linkage of glycosides, leading to the formation of a sugar hemiacetal or hemiketal and the corresponding free aglycon. Glycoside hydrolases are also referred to as glycosidases, and sometimes also as glycosyl hydrolases. Glycoside hydrolases can catalyze the hydrolysis of O-, N- and S-linked glycosides (Figure 1).



**Figure 1** Glycoside hydrolases.

**Endo/Exo:** *Exo* / *Endo*- represents the ability of an enzyme to cleave a substrate at the end of non-reducing end or the middle of a chain. For example, most cellulases are *endo*-acting, whereas LacZ  $\beta$ -galactosidase from *E. coli* is *exo*-acting (Figure 2).



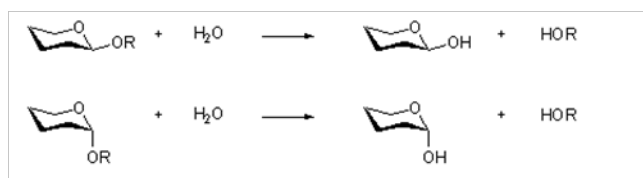
**Figure 2** Endo/exo cleavage.

## Enzyme commission (EC) number

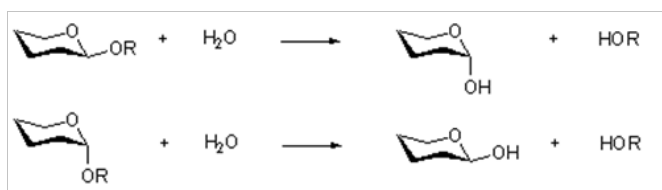
EC (Enzyme Commission Number) number is a numerical classification for enzyme that refers to the chemical reactions they catalyze. Every EC number is related with a recommended name for the respective enzyme. EC numbers do not specify enzymes, but enzyme-catalyzed reactions. If different enzymes catalyze the same reaction, then they get the same EC number. A necessary consequence of the EC classification scheme is that codes can be applied only to enzymes for which a function has been biochemically identified.

## Mechanistic classification

Two reaction mechanisms are most commonly found for the retaining and inverting enzymes, as first outlined by Koshland and as described below.<sup>1</sup> However some variations on these mechanisms have been found, and one basically different mechanism, catalyzed by an NADH cofactor, has been discovered in recent years (Figure 3) (Figure 4).



**Figure 3** Retaining glycoside hydrolases.



**Figure 4** Inverting glycoside hydrolases.

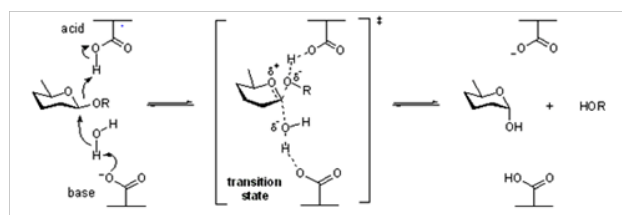
## Sequence-based classification

Sequence classification methods require knowledge of part of the amino acid sequence for an enzyme. Algorithmic methods are then used to compare sequences. Using a combination of comparison algorithms the glycoside hydrolases have been classified into more than 100 families that are available at the Carbohydrate Active enzyme database.<sup>2</sup> Each family (GH family) contains proteins that are associated by sequence, and by corollary. An obvious shortcoming of sequence-based classifications is that they can be applied to enzymes for which sequence information is available. But sequence-based classification allow classification of proteins for which no biochemical evidence has been obtained like the huge amount of not characterized glycosidase-related sequences that come from genome sequencing efforts worldwide.

This allows a number of useful predictions to be made since it has long been noted that the catalytic machinery and molecular mechanism is conserved for the vast majority of the glycosidase families<sup>1</sup> as well as the geometry around the glycosidic bond. Classification of families into larger groups, termed 'clans' has been mentioned.<sup>3</sup> A clan is a group of families that has similarity in their tertiary structure, catalytic residues and mechanism. Families in clans are thought to have a common evolutionary origin.

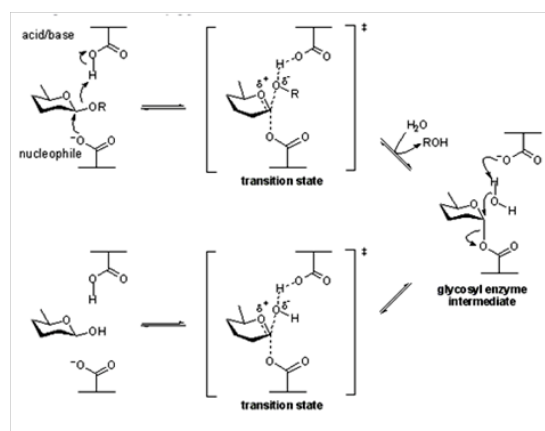
**Inverting glycoside hydrolases:** Hydrolysis of a glycoside with net inversion of anomeric configuration is generally achieved through one

step, single-displacement mechanism involving oxocarbenium ion like transition states, as shown below. The reaction occurs with acid/base assistance from two amino acid side chains, normally glutamic or aspartic acids<sup>4</sup> (Figure 5).



**Figure 5** Inverting mechanism for  $\beta$ -glycosidase.

**Retaining glycoside hydrolases:** Hydrolysis with net retention of configuration is most commonly achieved via a two steps, double-displacement mechanism involving a covalent glycosyl-enzyme intermediate, as is shown in the figure below. Each step passes through an oxocarbenium ion like transition states. Reaction occurs with acid/base and nucleophilic assistance provided by two amino acid side chains, typically glutamate or aspartate. In the first step, one residue plays the role of a nucleophile, attacking the anomeric center to displace the aglycon and form a glycosyl enzyme intermediate. At the same time the other residue functions as an acid catalyst and protonates the glycosidic oxygen as the bond cleaves. In the second step, the glycosyl enzyme is hydrolyzed by water, with the other residue now acting as a base catalyst deprotonating the water molecule as it attacks. The *pka* value of the acid/base group cycles between high and low values during catalysis to optimize it for its role at each step of catalysis<sup>5</sup> (Figure 6).



**Figure 6** Retaining mechanism for  $\beta$ -glycosidase.

**Alternative nucleophiles:** Sialidases hydrolyze glycosides of sialic acids. Closely related enzymes termed trans-sialidases catalyze the transglycosylation of sialisides. The sialidases and trans-sialidases of glycoside hydrolase families 33 and 34 utilize a tyrosine as a catalytic nucleophile, which is activated by an adjacent base residue. A rationale for this unusual difference is that the use of a negatively charged carboxylate as a nucleophile will be disfavored as the anomeric center is itself negatively charged, and thus charge repulsion interferes. A tyrosine residue is a neutral nucleophile, but requires a general base to enhance its nucleophilicity. This mechanism was implied from X-ray structures, and was supported by experiments involving trapping of the intermediate with fluoro sugars followed by peptide mapping and then crystallography,<sup>6,7</sup> and also through mechanism studies on mutants<sup>8</sup> (Figure 7).

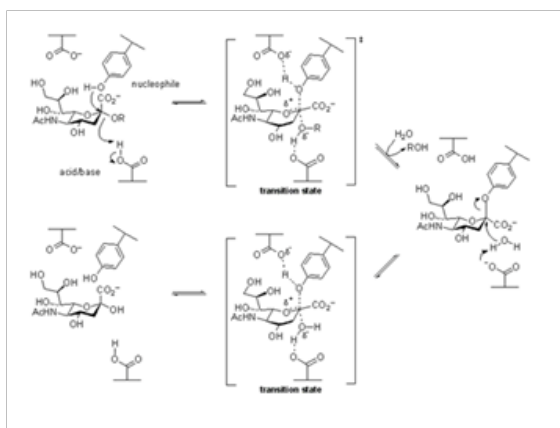


Figure 7 Alternative nucleophiles.

**NAD-dependent hydrolysis:** The glycoside hydrolases of family 4 and 109 uses a mechanism that requires an NAD cofactor, which remains tightly bound throughout catalysis. The mechanism proceeds through anionic transition states with elimination and redox steps rather than the classical mechanisms proceeding through oxocarbenium ion-like transition states. As shown below for a 6-phospho-beta-glucosidase, the mechanism involves an initial oxidation of the 3-hydroxyl of the substrate by the enzyme-bound NAD cofactor. This increases the acidity of the C2 proton such as an E1. Elimination can occur with assistance from an enzymatic base. The alpha-beta unsaturated intermediate formed then undergoes addition of water at the anomeric centre and finally the ketone at C3 is reduced to generate the free sugar product. Thus, even though glycosidic bond cleavage occurred via an elimination mechanism, the overall reaction is hydrolysis. This mechanism was elucidated through a combination of stereochemical studies by NMR, kinetic isotope effects, linear free energy relationships, X-ray crystallography and UV/Vis spectrophotometry<sup>9,10</sup> (Figure 8).

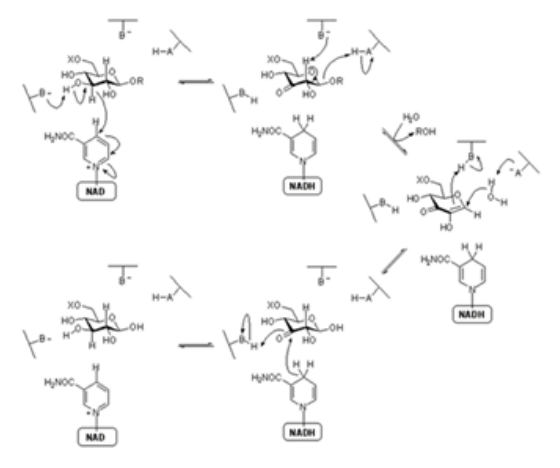


Figure 8 NAD-dependent hydrolysis.

## Glycoside hydrolase family I

Glycoside hydrolases (GHs), including glycosidases and transglycosidases constitute 113 protein families that are responsible for the hydrolysis and/or transglycosylation of glycosidic bonds. GH-coding genes are abundant and present in the vast majority of genomes corresponding to almost half-presently about 47% of the enzymes classified in CAZy. Because of their importance for

biotechnological and biomedical applications, GHs constitute so far the best biochemically characterized set of enzymes present in the CAZy database.<sup>2</sup>

## Substrate specificities

The most common known enzymatic activities for glycoside hydrolases in this family, at the current time, are b-glucosidases and b-galactosidases: typically both activities are found within the same active site, often with similar  $K_{cat}$  values, but with substantially higher  $K_m$  values for the galactosides. However, other commonly found activities are 6-phospho-b-glucosidase and 6-phospho-b-galactosidase, b-mannosidase, b-D-fucosidase and b-glucuronidase. Family GH1 enzymes are found a broad spectrum of life forms. Enzymes of medical interest include the human lactase/phlorizin hydrolase whose deficiency leads to lactose intolerance. In plants Family GH1 enzymes are often involved in the processing of glycosylated aromatics such as saponins and some plant hormones stored in inactive glycosylated forms. Indeed some have been identified as plant oncogenes due to aberrant control of auxin levels. Some plants also use Family GH1 enzymes as part of their defense system in order to release toxic aglycons, the most known examples being *Trifolium repens* b-glucosidase and *Sinapis alba* myrosinase, which respectively hydrolyse linamarin and glucosinolates. One of the work horses of glycosidase enzymology, the almond emulsin b-glucosidase, even if not fully sequenced, is deduced to belong to Family GH1 by limited sequence analysis.<sup>11</sup>

## Catalytic residues

He & Withers<sup>11</sup> was first identified  $\beta$ -glucosidase of *Agrobacterium* species as Glu358 in the sequence YITENG through trapping of the 2-deoxy-2-fluoroglucosyl-enzyme intermediate and subsequent peptide mapping. The general acid/base catalyst was first identified as Glu170 in this same enzyme through detailed mechanistic analysis of mutants at that position, which included azide rescue experiments. The GH1 enzymes, as a typical of Clan GH-A, have an asparagine residue preceding the general acid/base catalyst in a typical NEP sequence. The asparagine engages in important hydrogen bonding interactions with the substrate 2-hydroxyl. Interestingly, the plant myrosinases cleave thioglycosides bearing an anionic aglycone (glucosinolates), and have evolved an active site in which the acid/base glutamate is replaced by glutamine. Substrates are sufficiently reactive not to require the acid catalyst, while the role of base catalyst is played by exogenous ascorbate, which binds to the glycosyl enzyme.<sup>12,13</sup>

## Three-dimensional structures of GH1

Three-dimensional structures are available for a large number of Family1 enzyme, the first solved being that of the white clover (*Trifolium repens*) cyanogenic  $\beta$ -glucosidase.<sup>14</sup> As members of Clan GH-A they have a classical ( $\alpha$ ) $\beta$ 8 TIM barrel fold with the two key active site glutamic acids being approximately 200 residues apart in sequence and located at the C-terminal ends of  $\beta$ -strands 4 (acid/base) and 7 (nucleophile).<sup>15,16</sup>

## Structures of GH-I $\beta$ -galactosidases from the hyperthermophilic archeon *Sulfolobus solfataricus* (Ss- $\beta$ -gal) and *Pyrococcus furiosus* (Pf- $\beta$ -gal)

In recent years, several thermostable  $\beta$ -galactosidases have been reported Wanarska et al & Synowiecki et al, *Sulfolobus solfataricus* MT4, a hyperthermophilic archaeon first isolated from hot mud in



the Solfatara crater north of Naples grows optimally at 87°C and expressed a GH activity, initially characterized as a  $\beta$ -galactosidase on the basis of hydrolysis of the chromogenic substrate, 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside (X-gal). Subsequent enzymatic analysis showed that this enzyme had several *exo*- $\beta$ -glycosidase substrate specificity.

The amino acid sequence derived from the *lacS* gene put the enzyme in family 1 of the  $\beta$ -glycosyl hydrolases, along with bacterial  $\beta$ -glycosidases, 6-phospho- $\beta$ -galactosidases, cyanogenic  $\beta$ -glycosidases, plant myrosinases and mammalian gut lactases. The Ss- $\beta$ -gal that had optimal activity with a half-life of 48h at 85°C was not thermally denatured under 100°C and resistant to denaturation by organic solvents that can be very useful to synthesize a variety of glycosides by transglycosidation and condensation. Aguilar et al. reported the structure of the native tetrameric enzyme, and site-directed mutagenesis and homology with other GH-1 glycosidases have allowed to identify the active site of the enzyme and to define the substrate binding site. From analysis of the refined structure, the main feature that distinguishes this enzyme from mesophilic proteins was the presence of a large number of ion-pair networks which crosslink the surface of the protein. This feature, coupled with the observation of substantial numbers of buried water molecules, suggests that this enzyme (possibly other hyperthermophile proteins) may be achieved its hyperthermostability by resilience rather than rigidity. Hyperthermostable  $\beta$ -galactosidases from *Pyrococcus furiosus* (EP 0687732, 2002; EP 0606008, 2004) and *Pyrococcus woesei* have also cloned and sequenced, but the structures are not known. Their hydrolytic activities on lactose have not been studied, despite the high chromogenic activities on ONPG or X-gal.

## Glycoside hydrolase family 42 ( $\beta$ -galactosidase)

### Substrate specificity

The most common activity for glycoside hydrolases of this family are  $\beta$ -galactosidases (EC 3.2.1.23), however, other commonly found activities are  $\alpha$ -L-arabinosidase (EC 3.2.1.55) and  $\beta$ -D-fucosidase (EC 3.2.1.38) with both  $K_m$  and  $K_{cat}$  values being of the same order of magnitude for the different substrates. Family GH42 enzymes have been identified only in unicellular organisms, mainly from prokaryotes (bacteria), and with a few examples from archaea and fungi. GH42 enzymes are active on lactose and transgalactosylation was observed with production of galactooligosaccharides. However, some GH42 enzymes are extracted from diverse habitats where lactose would not be present and they are active on galactooligosaccharides, suggesting that these enzymes would be involved *in vivo* in plant cell wall degradation. This feature could be performed in cooperation with family GH53 galactanases, often encoded from genes adjacent to GH42 genes, and with cellulosome. The activity of GH42 enzymes on lactose and also lactulose has interesting potential for the removal of the former from dairy products and to monitor lactulose concentration during heat treatment leading to UHT milk.

### Catalytic residues

The catalytic nucleophile was first identified in the *Bacillus subtilis* YesZ  $\beta$ -galactosidase as Glu295 through the use of a mechanism-based inhibitor that allowed trapping of the 2-deoxy-2-fluorogalactosyl-enzyme intermediate and subsequent peptide mapping. These experiments were performed on the mutant of the inferred acid/base, which was more sensitive to the inhibitor. The acid/base catalyst was first identified as Glu157 in the  $\beta$ -galactosidase

from *Alicyclobacillus*. *Acidocaldarius* through detailed mechanistic analysis and azide rescue experiments of a mutant in that position.

### $\beta$ -Mannosidase

$\beta$ -Mannan and its heteropolysaccharides are found in endosperms of copra, ivory nuts, guar beans, locust beans, coffee beans, the roots of konjak, and hemicellulose of soft and hardwoods. Some  $\beta$ -mannans from agricultural crops were used as food additive for maintaining desired consistency or non-caloric dietary fiber. Oligosaccharides from these polysaccharides by an enzymatic hydrolysis are expected to be healthy material, such as fructooligosaccharide. The hydrolysis of  $\beta$ -mannan is catalyzed by  $\beta$ -mannanase (*endo*-type, EC 3.2.1.78) and  $\beta$ -mannosidase (*exo*-type, EC 3.2.1.25). These enzymes were produced by various microorganisms.

Hemicellulases are key components in the degradation of plant biomass and carbon flow in nature. The substrates of these enzymes, the hemicelluloses, are a heterogeneous group of branched and linear polysaccharides that are bound via hydrogen bonds to the cellulose microfibrils in the plant cell wall, cross linking them into a robust network. Hemicelluloses are also covalently attached to lignin, forming together with cellulose a highly complex structure. The degradation of cellulose and hemicellulose is carried out by microorganisms that can be found either free in nature or as part of the digestive tract of higher animals. The variable structure and organization of hemicellulose require the concerted action of many enzymes for its complete degradation.

The catalytic modules of hemicellulases are either glycoside hydrolases (GHs) that hydrolyze glycosidic bonds, or carbohydrate esterases (CEs), which hydrolyze ester linkages of acetate or ferulic acid side groups. These catalytic modules, based on homology of their primary sequence, can be assigned into families marked by numbers (e.g. GH7 for glycoside hydrolase family 7). Some families, with overall similar fold, can be further grouped into clans, marked alphabetically (e.g. GH-A). The various chemical bonds hydrolyzed by hemicellulases and the classifications of these enzymes into families are xylanases (EC 3.2.1.8),  $\beta$ -mannanases (EC 3.2.1.78),  $\beta$ -mannosidases (EC 3.2.1.25),  $\beta$ -L-arabinofuranosidases (EC 3.2.1.55) and  $\beta$ -xylosidases (EC 3.2.1.37).  $\beta$ -Mannosidase (EC 3.2.1.25) is an enzyme that catalyzes the hydrolysis of terminal, non-reducing beta-D-mannose residues in beta-D-mannosides. The enzyme plays a role in the lysosomal degradation of the N-glycosylprotein glycans.  $\beta$ -Mannosidases are also found in higher animals and are involved in lysosomal degradation of glycoproteins. In another study, the substrate specificities of two GH1 enzymes,  $\beta$ -mannosidase and  $\beta$ -glycosidase, were analyzed by constructing several reciprocal replacements of two active-site-conserved residues in the  $\beta$ -glycosidase, both substitutions increased the specificity for mannosides, whereas in the  $\beta$ -mannosidase, one of the replacements resulted in improved catalysis towards glucosides. In combination with inhibition studies, it was concluded that the mutated residues are directly involved in the stabilization of the transition states, and also participate in the ground-state binding of substrates with the equatorial C2-hydroxyl. Fungi and bacteria that are able to degrade hemicellulose also secrete  $\beta$ -mannosidase. Fungal and bacterial  $\beta$ -mannosidases hydrolyze  $\beta$  (1-4)-D-mannosyl groups from manno-oligosaccharides and mannose-containing glycopeptides that are produced from the hemicellulose pulp by endoenzymes. In seeds that have galactomannans as storage carbohydrates, the enzyme converts manno-oligosaccharides to monosaccharides.

Defects in the lysosomal form of the enzyme in humans result in a buildup of mannoside intermediate metabolites and the disease beta-mannosidosis. Deficiency in these enzymes in humans and ruminants (termed  $\beta$ -mannosidosis) leads to mental retardation and skeletal abnormalities. Beta-mannosidosis is an autosomal recessive lysosomal storage disease of glycoprotein catabolism caused by a deficiency of lysosomal beta-mannosidase activity.<sup>17</sup> The disorder was first described in goats,<sup>18</sup> which have a more severe neurodegenerative disorder than that seen in humans. In the case of man, clinical manifestations are heterogeneous and include mental retardation, peripheral neuropathy, and skeletal abnormalities.

### Hyperthermophilic glucoside hydrolases

Extremophiles are a group of microorganisms that can survive and even thrive in extreme environments. They are classified according to the conditions under which they grow at the extremes of physical parameters as (hyper)thermophiles (at high temperature), psychrophiles (at low temperature), acidophiles (at low pH), alkaliphiles (at high pH), piezophiles (at high pressure) and halophiles (at high salinity). Therefore, their biomolecules such as proteins, nucleic acids, and lipids have been evolved to function properly under these severe conditions.<sup>19</sup> Due to these particular properties, enzymes from extremophiles (extremozymes) offer a high potential not only for basic research but also for biotechnological applications. Potentially, extremozymes would enable us to expand the range of reaction conditions suitable for biocatalysis. This has proven to be the case and various applications in sugar chemistry, detergent production, lipid and oil chemistry, and food processing have been initiated or are being explored.<sup>20,21</sup>

Thermophilic microorganisms have attracted most attention, among which the most studies are toward the extremophile. Thermophiles can be generally classified into moderate thermophiles (growth optimum temp 50–60°C), extreme thermophiles (growth optimum temp 60–80°C) and hyperthermophiles (growth optimum temp 80–110°C). Of all extremozymes, thermophilic enzymes have attracted most attention during the past four decades. Such enzymes are of great industrial and biotechnological interest due to the fact that the enzymes are better suited for harsh industrial processes. There are many advantages of conducting industrial processes at high temperature, such as the increased solubility of many polymeric substrates, resulting in decreased viscosity, increased bioavailability, faster reaction rate and the decreased risk of microbial contamination. These enzymes have also been used as models for the understanding of thermostability and thermoactivity, which is useful for protein engineering. Structural features of thermophilic extremozymes have attracted much attention. Several three-dimensional structures have been resolved and compared with those of mesophilic counterparts, with the ultimate goal of elucidating the mechanisms underlying thermostability.

A large number of sequence and structural factors are thought to contribute toward higher intrinsic thermal stability of proteins from thermophiles and hyperthermophiles. Thermophiles produce special proteins called ‘chaperonins’, which are thermostable and resistant to denaturation and proteolysis. Proteins of thermophiles, denatured at high temperature, are refolded by the chaperonins, thus restoring their native form and function. The cell membrane of thermophiles also consists of saturated fatty acids, which increase protein core hydrophobicity and keep the cell rigid enough to survive at high temperatures. Moreover, hyperthermophiles have membranes containing lipids linked with ether to their cell walls. This layer is much

more heat resistant than a membrane formed of fatty acids. In addition, proteins of thermophiles have increased surface charge and less exposed thermolabile amino acids. Thus, increased ionic interaction and hydrogen bonds, increased hydrophobicity, decreased flexibility and smaller surface loops confer stability on the thermophilic protein. Heat-tolerant enzymes are investigated currently the most of all extremozymes. Such extremozymes are obtained from either extreme thermophiles that grow optimally at temperatures above 60°C or from hyperthermophiles with optimal growth temperatures above 90°C.

Extreme thermophiles include both archaeal and bacterial species, while the vast majority of hyperthermophiles are members of the domain Archaea.<sup>22</sup> Species found at the highest temperatures (103–113°C) are exclusively archaea. Most of hyperthermophiles encompassing more than 70 species, 29 genera, and 10 orders are archaea among which *Thermotogales* and *Aquificales* are the only bacteria.<sup>23</sup> Application interest in thermophilic extremozymes is related to the fact that performing biotechnologically related processes at higher temperatures is often advantageous. For instance, in chemical reactions involving organic solvents, the decrease in viscosity and increase in diffusion coefficient that is achieved at elevated temperatures result in higher reaction rates.<sup>24</sup> Such considerations are relevant to a variety of processes including those using hydrophobic compounds with low solubility. High temperatures can also enhance the availability of such compounds for biodegradation efforts. Furthermore, reactions at higher temperatures reduce the possibility of complications due to contamination. Because of their capability to function at such high temperatures, hyperthermophilic extremozymes have been the main focus of biotechnological interest.

The recent advent of genomic research has produced vast amounts of sequence information for many different taxa of Bacteria, Eukarya, and Archaea, now collected in databases such as GenBank; the full genome sequences of more than 699 different microorganisms have been completed. By a generally applicable combination of conventional genetic engineering and genomic research techniques, many extremely thermostable enzymes are being developed for biotechnological purposes. The genome sequences of some hyperthermophilic microorganisms such as *Thermotoga maritima*, *Pyrococcus furiosus*, and *Sulfolobus solfataricus* are of considerable biotechnological interest because they encode many highly heat-stable enzymes that are active under conditions previously regarded as incompatible with biological materials.<sup>19</sup> In the genome sequences of hyperthermophilic Archaea and Bacteria, number of gene sequences in glycoside hydrolase family is summarized based on the web accessible CAZY (Carbohydrate-Active Enzymes) database (Table 1). Amylolytic enzymes are of great significance in many industrial processes including those for foods, textiles, and detergents.<sup>25</sup> Many hyperthermophilic microorganisms possess starch-hydrolyzing enzymes, such as  $\alpha$ -amylase,  $\alpha$ -glucosidase, pullulanase, and cyclodextrinase, in their genomes even though they live in environments where starch is rare.

Analysis of the full genome of *Pyrococcus furiosus*, a hyperthermophilic archaeon, has revealed that this microorganism has several amylolytic enzymes. An amylopullulanase and two distinct  $\alpha$ -amylase genes of *Pyrococcus furiosus* were identified and expressed in *E. coli*. These enzymes can hydrolyze a wide variety of substrates such as soluble starch, amylose, amylopectin, glycogen, and oligosaccharides. However,  $\alpha$ -amylase does not hydrolyze pullulan and CD (cyclodextrin), whereas amylopullulanase can degrade pullulan.

**Table 1** Carbohydrate-Active Enzymes (CAZYs) in glycoside hydrolase family (GHF), Archaea are indicated by the shaded box.at66088625

Hyperthermophilic microorganisms	Completed	Genome size	Growth temp (°C)	Putative CAZymes
<i>Aeropyrum pernix</i>	2001	1,669,695	95	0
<i>Archaeoglobus fulgidus</i> DSM 4304	1997	2,178,400	83	0
<i>Methanococcus jannaschii</i> DSM 2661	2001	1,664,970	85	2
<i>Methanopyrus kandleri</i> AV19	2002	1,694,969	98	8
<i>Nanoarchaeum equitans</i> Kin 4-M	2003	490,885	90	0
<i>Pyrobaculum aerophilum</i> IM2	2001	2,222,430	100	6
<i>Pyrococcus abyssi</i> GE5 Orsay	2001	1,765,118	103	7
<i>Pyrococcus furiosus</i> DSM 3638	2002	1,908,256	100	15
<i>Pyrococcus horikoshii</i> OT	2001	1,738,505	98	8
<i>Sulfolobus solfataricus</i> P2	2001	2,992,245	85	22
<i>Sulfolobus tokodaii</i> 7	2001	2,694,756	80	14
<i>Thermococcus kodakarensis</i> KOD1	2005	2,088,737	85	13
<i>Aquifex aeolicus</i> VF5	2001	1,551,335	96	7
<i>Thermoanaerobacter tengcongensis</i> MB4	2001	2,689,445	75	24
<i>Thermotoga maritima</i> MSB8	2001	1,860,725	80	147
Total				167

### Hyperthermophilic $\beta$ -Galactosidase

Lactose is a primary carbohydrate contained in mammal's milk and dairy products (at a concentration between 5 and 10%, depending on the source of milk) and is also called milk sugar. The consumption of lactose or lactose containing foods can cause digestion problem for almost a 70% of the world population, as the enzyme naturally present in human intestine loses its activity during lifetime. Moreover, low solubility and sweetness of lactose lead to huge interest in industrial development of processes to hydrolyze lactose and lactose contained dairy foods. Lactose, a natural disaccharide, is consists of monosaccharides, glucose and galactose bonded by a beta-1, 4 linkages. Lactose can be hydrolyzed to monomers, glucose and galactose by enzyme ( $\beta$ -galactosidase) or acids.

$\beta$ -Galactosidase (EC 3.2.1.23) is a common name of  $\beta$ -D-galactoside galactohydrolase and also called lactase (EC 3.2.1.108). Lactase cleaves the  $\beta$ -1, 4-glycosidic linkage of lactose and gives rise to its component monosaccharides, glucose and galactose. The hydrolysis of lactose in dairy products by lactase can be beneficial in many regards; (1) the hydrolysis of lactose alleviates problems and improves processes for dairy products, (2) the low solubility and lack of sweetness that are often experienced in concentrated milk products and ice cream could be overcome by lactose hydrolysis, (3) cheeses that have been manufactured from hydrolyzed milk ripens more quickly than that made from normal milk,<sup>26</sup> and (4) the use of lactase could also reduce the amount of the lactose in whey, which can cause environmental pollution when discharged in large quantities.<sup>27</sup>

Lactose hydrolyzed (lactose free) milk can reduce lactose intolerance problem and lactose hydrolyzed whey syrup can be utilized

in frozen desserts, confectionary, bakery, fermentation products, and beverages. The galacto-oligosaccharides can also be employed as probiotic food ingredients, humectants, and emulsifiers.

### Galacto-oligosaccharide formation by $\beta$ -galactosidase-transgalactosylation

$\beta$ -Galactosidase hydrolyzes terminal, non-reducing  $\beta$ -D-galactose residues in  $\beta$ -D-galactosides or lactose, but some of this enzyme catalyzes both hydrolytic and reverse transgalactosylation (EC 2.4.1.22: galactosyl transferase) reaction. The transferase activity by a  $\beta$ -galactosidase that produce oligosaccharides was reported in the early 1950s. Apart from theoretical aspects, early research was prompted by nutritional concerns about the presence of these compounds in low-lactose milk. Later studies were based on the need to consider oligosaccharide formation when modelling lactose hydrolysis. More recently, interest in the reaction has been raised by observation that oligosaccharides may have beneficial effects as 'bifidus factors'-promoting the growth of desirable intestinal microflora. Also, the transferase reaction can be used to attach galactose to other chemicals and consequently has potential applications in the production of food ingredients, pharmaceuticals and other biologically active compounds. Lactose hydrolysis catalysed by  $\beta$ -galactosidase has proven to be a very complex reaction. Apart from the actual hydrolysis product, glucose and galactose, many newly formed  $\beta$ -glycoside, mainly di, tri, and tetrasaccharide, occur as kinetic intermediates, derived from so-called transgalactosylation reaction. Because transgalactosylation products (galacto-oligosaccharides) are substrate of  $\beta$ -galactosidases-catalyzed hydrolysis, the composition of the product mixture changes quite significantly with progressing reaction time.<sup>28</sup>



The specific properties of oligosaccharides are very different depending on the formation of oligosaccharides, but some properties are common to almost all oligosaccharides. The sweetness of the oligosaccharide depends on structure and molecular mass of the oligosaccharides. Oligosaccharides are normally water soluble and mildly sweet, typically lower than sucrose and this low sweetness is useful in food production when reduced sweetness is desirable to enhance other food flavors. Compared with mono- and disaccharides, the higher molecular weight of oligosaccharides provides increased viscosity, leading to improved body and mouthfeel. They can also be used to alter the freezing temperature of frozen foods, and to control the amount of browning due to Maillard reactions in heat-processed foods. Oligosaccharides provide a high moisture-retaining capacity, preventing excessive drying, and a low water activity, which is convenient in controlling microbial contamination.<sup>29–34</sup>

Although oligosaccharides possess these useful physicochemical characteristics, most of the interest in their use as food ingredients stems from their many beneficial physiological properties. Unlike starch and simple sugars, the currently available food-grade oligosaccharides are not utilized by mouth microflora to form acid or polyglucans. Hence, they are used as low-cariogenic sugar substitutes in confectionery, chewing gums, yogurts and drinks. Many oligosaccharides are not digested by humans. Oligosaccharides have recently been described as one of several ‘prebiotics’, which can stimulate the growth of beneficial microflora.

Prebiotics are defined as “non-digestible food ingredients that beneficially affect the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria in the colon”. They are recognized for their ability to increase levels of health promoting bacteria in the intestinal tract of humans or animals. Some studies have shown that prebiotics target the activities of bifidobacteria and/or lactobacilli. Both the volume and diversity of oligosaccharide products are increasing very rapidly as their functional properties become further understood. Detailed production methods for various oligosaccharides have been reviewed by Playne.<sup>35–37</sup>

### Hyperthermophilic mannose hydrolyzing enzymes

The complexity of cell wall-degrading enzyme systems is a consequence of the complex nature of plant cell wall. Hemicelluloses act as linkers between lignin and cellulose. The high percentage of hemicellulose fraction in the cell wall of higher plants makes this material the second most abundant biopolymer in nature. Besides xylan, mannan is the other major hemicelluloses constituent. Galactomannan, found in large quantities in seeds of leguminous plants, is composed of a homogeneous backbone of  $\beta$ -1, 4-linked mannose residues, whereas acetylated galactoglucomannan, a main constituent of softwoods, has a heterogeneous backbone of  $\beta$ -1, 4-linked glucose and mannose residues (Bacic *et al.*, 1988). The complete conversion of galactomannan into galactose and mannose requires the activity of three types of enzymes, namely, endomannanases (EC 3.2.1.78),  $\beta$ -galactosidase (EC 3.2.1.22), and  $\beta$ -mannosidases ( $\beta$ -D-mannopyranoside hydrolase [EC 3.2.1.25]). Endomannanases catalyze the random hydrolysis of the  $\beta$ -1, 4-mannosidic backbone of the main mannan chain,  $\beta$ -galactosidase cleave the terminal  $\beta$ -1, 6-linked D-galactosyl -mannosidases residues, and hydrolyze  $\beta$ -1, 4-linked mannose residues from the nonreducing ends of various oligosaccharides.

Almost every  $\beta$ -mannosidase belongs to family 2 of glycosyl hydrolases (GHs), except for the *Pyrococcus* sp.  $\beta$ -mannosidase, which has been assigned to family 1. Hemicellulases are widely used

in coffee bean fermentation to promote the hydrolysis of  $\beta$ -mannan-based oligosaccharides.

In general, mannosidases constitute only a small percentage of the proteins secreted by hemicellulose degrading organisms; their purification is therefore rather difficult. This problem could be solved by cloning and heterologous expression of mannosidase-encoding genes. Given the natural abundance of hemicellulose (heteroglycans), it is not surprising that many microorganisms have enzyme systems for its hydrolysis. Moreover, given the variety and complexity of hemicellulases, several biocatalytic steps are typically required to hydrolyze specific polysaccharides completely into simpler sugars that can be readily used as carbon and/or energy sources by particular microorganisms. Hetero-1, 4-  $\beta$ -D-mannans, one of the major constituents of hemicellulose, are hydrolyzed to mannose through *endo*-acting  $\beta$ -mannanases (1, 4- $\beta$ -D-mannan mannanohydrolase [EC 3.2.1.78]) (McCleary, 1988a), and *exo*-acting  $\beta$ -mannosidases ( $\beta$ -D-mannopyranoside hydrolase [EC 3.2.1.25]). Additional enzymes are required to remove side chain sugars that are attached at various points on mannans. For example, galactomannans have galactose residues bound to the mannan backbone, and their removal is effected through the action of  $\beta$ -galactosidases ( $\beta$ -D-galactoside galactohydrolase [EC 3.2.1.22]).  $\beta$ -Mannan-based natural polymers have wide-ranging industrial applications, such as those used in the processing of foods and the massive hydraulic fracturing of oil and gas wells. Recent developments in the oil and gas industries have established a need for the *in situ* enzymatic hydrolysis of galactomannans used in well stimulation. In these applications, enzyme thermostability and thermoactivity are factors. In food processing, for reasons of asepsis and the need to process viscous materials, elevated temperatures provide obvious advantages. In oil and gas well stimulation, temperatures exceeding 100°C are characteristic of the deeper reaches of well bores. Thus, the availability of thermostable and thermoactive versions of enzymes involved in the hydrolysis of  $\beta$ -mannan-based natural polymers is necessary.

### Acknowledgements

None.

### Conflicting interest

The author declares no conflict of interest.

### References

- Gebler J, Gilkes NR, Claeysens M, et al. Stereoselective hydrolysis catalyzed by related  $\beta$ -1,4-glucanases and  $\beta$ -1,4-xylanases. *Journal of Biological Chemistry*. 1992;267(18):12559–12561.
- Cantarel BL, Coutinho PM, Rancurel C, et al. The Carbohydrate-Active Enzymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res*. 2009;37(Database issue):D233–D238.
- Henrissat B, Bairoch A. Updating the sequence-based classification of glycosyl hydrolases. *Biochem J*. 1996;316(Pt 2):695–696.
- McCarter JD, Withers SG. Mechanisms of enzymatic glycoside hydrolysis. *Curr Opin Struct Biol*. 1994;4(6):885–892.
- McIntosh LP, Hand G, Johnson PE, et al. The pKa of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a <sup>13</sup>C-NMR study of bacillus circulans xylanase. *Biochemistry*. 1996;35(31):9958–9966.
- Watts AG, Damager I, Amaya ML, et al. Trypanosoma cruzi trans-sialidase operates through a covalent sialyl-enzyme intermediate: tyrosine is the catalytic nucleophile. *J Am Chem Soc*. 2003;125(25):7532–7533.

7. Amaya MF, Watts AG, Damager I, et al. Structural Insights into the Catalytic Mechanism of Trypanosoma cruzi trans-Sialidase. *Structure*. 2004;12(5):775–784.
8. Watson JN, Dookhun V, Borgford TJ, et al. Mutagenesis of the conserved active-site tyrosine changes a retaining sialidase into an inverting sialidase. *Biochemistry*. 2003;42(43):12682–12690.
9. Rajan SS, Yang X, Collart F, et al. Novel Catalytic Mechanism of Glycoside Hydrolysis Based on the Structure of an NAD<sup>+</sup>/Mn<sup>2+</sup>-Dependent Phospho- $\alpha$ -Glucosidase from Bacillus subtilis. *Structure*. 2004;12(9):1619–1629.
10. Yip VLY, Varrot A, Davies GJ, et al. An unusual mechanism of glycoside hydrolysis involving redox and elimination steps by a family 4 beta-glycosidase from Thermotoga maritima. *J Am Chem Soc*. 2004;126(27):8354–8355.
11. He S, Withers SG. Assignment of sweet almond beta-glucosidase as a family 1 glycosidase and identification of its active site nucleophile. *J Biol Chem*. 1997;272(40):24864–24867.
12. Wang Q, Trimbur D, Graham R, et al. Identification of the acid/base catalyst in Agrobacterium faecalis beta-glucosidase by kinetic analysis of mutants. *Biochemistry*. 1995;34(44):14554–14562.
13. Burmeister WP, Cottaz S, Rollin P, et al. High resolution X-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. *J Biol Chem*. 2000;275(50):39385–39393.
14. Barrett T, Suresh CG, Tolley SP, et al. The crystal structure of a cyanogenic beta-glucosidase from white clover, a family 1 glycosyl hydrolase. *Structure*. 1995;3(9):951–960.
15. Henrissat B, Callebaut I, Fabrega S, et al. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc Natl Acad Sci U S A*. 1995;92(15):7090–7094.
16. Henrissat B, Callebaut I, Fabrega S, et al. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc Natl Acad Sci U S A*. 1995;92(15):7090–7094.
17. Bedilu R, Nummy KA, Cooper A, et al. Variable clinical presentation of lysosomal beta-mannosidosis in patients with null mutations. *Molecular genetics and metabolism*. 2002;77(4):282–290.
18. Jones MZ, Dawson G. Caprine beta-mannosidosis: inherited deficiency of  $\beta$ -D-mannosidase. *J Biol Chem*. 1981;256:5185–5188.
19. Schiraldi C, Martino A, Acone M, et al. Effective production of a thermostable alpha-glucosidase from Sulfolobus solfataricus in Escherichia coli exploiting a microfiltration bioreactor. *Biotechnol Bioeng*. 2000;70(6):670–676.
20. Demirjian DC, Moris-Varas F, Cassidy CS. Enzymes from extremophiles. *Curr Opin Chem Biol*. 2001;5(2):144–151.
21. van den Burg B. Extremophiles as a source for novel enzymes. *Curr Opin Microbiol*. 2003;6(3):213–218.
22. Adams MW. Enzymes and Proteins from Organisms that Grow Near and Above 100 degree C. *Annu Rev Microbiol*. 1993;47:627–658.
23. Stetter KO. Hyperthermophiles in the history of life. *Philos Trans R Soc Lond B Biol Sci*. 1996;361(1474):1837–1842.
24. Krahe M, Antranikian G, Märkl H. Fermentation of extremophilic microorganisms. *FEMS Microbiology Reviews*. 1996;18(2–3):271–285.
25. Bertoldo C, Antranikian G. Amylolytic enzymes from hyperthermophiles. *Hyperthermophilic Enzymes Part A*. 2001;330:269–290.
26. Kim S, Lim K, Kim H. Differences in the Hydrolysis of Lactose and Other Substrates by beta-D-Galactosidase from Kluyveromyces lactis. *J Dairy Sci*. 1997;80(10):2264–2269.
27. Gekas V, Lopez-Leiva M. Hydrolysis of lactose: a literature review. *Process biochemistry*. 1985;20(1):2–12.
28. Onishi N, Yamashiro A, Yokozeki K. Production of galacto-oligosaccharide from lactose by Sterigmatomyces elviae CBS8119. *Appl Environ Microbiol*. 1995;61(11):4022–4025.
29. Berka RM, Hucul JA, Ward M. Increased production of beta-galactosidase in aspergillus oryzae. *Journal of sciences Islamic republic of Iran*. 1998;8(1):23–27.
30. Freire E, Adrián Velázquez-Campoy, Hiroyasu O, et al. Isothermal titration calorimetry. *Current Protocols in Cell Biology*. 2004;17:1–17.
31. Geerlof A, Brown J, Coutard B, et al. The impact of protein characterization in structural proteomics. *Acta Crystallographica Section D, Biological Crystallography*. 2006;62(Pt 10):1125–1136.
32. Henrissat B. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*. 1991;280(Pt 2):309–316.
33. Jost B, Vilotte JL, Duluc I, et al. Production of low-lactose milk by ectopic expression of intestinal lactase in the mouse mammary gland. *Nat Biotechnol*. 1999;17(2):160–164.
34. Pessela BCC, Mateo C, Carrascosa AV, et al. One-Step Purification, Covalent Immobilization, and Additional Stabilization of a Thermophilic Poly-His-Tagged  $\beta$ -Galactosidase from Thermus sp. Strain T2 by using Novel Heterofunctional Chelate-Epoxy Sepabeads. *Biomacromolecules*. 2003;4(1):107–113.
35. Somkuti GA, Steinberg DH.  $\beta$ -Fructofuranosidase activity in disaccharide transport mutants of Streptococcus thermophilus. *Biotechnology Letters*. 1991;13(11):809–814.
36. Sun S, Li X, Nu S, et al. Immobilization and Characterization of  $\beta$ -Galactosidase from the Plant Gram Chicken Bean (Cicer arietinum). Evolution of Its Enzymatic Actions in the Hydrolysis of Lactose. *J Agric Food Chem*. 1999;47(3):819–823.
37. Todd MJ, Gomez J. Enzyme Kinetics Determined Using Calorimetry: a general Assay for Enzyme Activity? *Anal Biochem*. 2001;296(2):179–187.