STUDIA UBB CHEMIA, LX, 1, 2015 (p. 239-251)
(RECOMMENDED CITATION)

COMPUTATIONAL ANALYSIS OF THE STRUCTURAL PROPERTIES OF ALPHA - AND BETA - GALACTOSIDASES

DIANA-LARISA VLĂDOIU[a], VASILE OSTAFE[a], ADRIANA ISVORAN[a*]

ABSTRACT. A computational study to compare the global and local physicochemical and structural properties of *alpha*- and *beta*-galactosidases using the retaining catalytic mechanism was performed. These proteins share quite similar global structural properties despite their low sequence similarity, structures superposition resulting in root mean squared deviation (RMSD) values around 1.25 Å for at least 43 alpha carbon atoms pairs. Almost the same RMSD values are obtained for the superposition of the catalytic domains of investigated galactosidases, but for a higher number of alpha carbon atoms pairs (68) reflecting the higher structural similarity of the catalytic domains. There are local individual properties of the surfaces of considered enzymes, *beta*-galactosidases exposing a more complex surface with a higher number of cavities, 42 for eukaryotic *beta*-galactosidases compared to 18 for eukaryotic a*lpha*-galactosidases. Furthermore, *beta*- galactosidases usually depict larger and more hydrophobic cavities than *alpha* – galactosidases, the hydrophobicity scores of the biggest cavities being 24 for eukaryotic *beta*-galactosidases and 7 for eukaryotic *alpha*-galactosidases, respectively.

Keywords*: global and local structural properties, surface cavities; surface roughness.*

INTRODUCTION

Enzymatic hydrolysis is a process widely used to break down proteins, cellulose chains, starch and fat molecules into smaller ones: amino acids, monomeric sugars and fatty acids respectively. The enzymatic hydrolysis of glycosidic bonds involves the presence of glycoside hydrolases (GH), a class of enzymes also called glycosidases. They are involved in both industrial and natural processes, from biomass degradation [1] and food processes [2] to normal cellular functions and pathogenesis [3].

[a] *West University of Timisoara, Timisoara, Faculty of Chemistry – Biology – Geography, Advanced Research Environmental Laboratories, Oituz str. 4,Timișoara, 300086, Romania*
** Corresponding author*: *adriana.isvoran@e-uvt.ro*

GH enzymes may be classified in different ways. One of this ways is to use the Enzyme Commission (EC) number [4] but this classification is based on type of catalyzed reaction and the substrate specificity and does not consider structural properties of enzymes and their evolutionary relationships. For GH enzymes, the Carbohydrate-Active enZYmes (CAZy) database [4] contains a sequence-based classification for 133 families. Each GH family contains proteins that are related by sequence which reflects the same structural fold. Some of these families share structural similarity despite their apparently unrelated sequences, have common ancestry and identical catalytic mechanisms forming clans, denoted from A to N [5, 6]. Every clan contains a group of families with significant similarity of their tertiary structures, catalytic residues and mechanism.

GH enzymes use most frequently two catalytic mechanisms for the hydrolysis of the glycosidic bonds: the retaining or the inverting mechanism [7]. There are also some variations of these two mechanisms and a few others mechanisms may be employed by the GH enzymes [4, 8]. The retaining and inverting mechanisms involve two amino acids of the enzyme: a proton donor and a nucleophile/base [9]. These two residues are located at opposite sides of the active site of enzyme and, depending on their spatial positions, hydrolysis occurs via retention or inversion mechanism [9].

Within this study, two classes of glycoside hydrolases were considered: alpha-galactosidases (E.C.3.2.1.22) hydrolysing the terminal *alpha*-galactosyl moieties from glycolipids and glycoproteins [10] and *beta*- galactosidases (E.C.3.2.1.23) hydrolysing the β-glycosidic bonds from beta-galactosides and glycoproteins [11]. The difference between the retaining catalytic mechanisms of *alpha*- and *beta*- galactosidases consists in the distinct direction the nucleophile attacks the anomeric center at the first carbon of the substrate: it acts on the first carbon belonging to a substrate that contains an *alpha*- linkage for *alpha*- galactosidases, respective to a first carbon belonging to a substrate containing a *beta*- linkage for *beta*-galactosidases [15]. The great interest in studying galactosidases rises from their presence in almost all types of living organisms, their involvement in some diseases [3,12-14] and also from their biotechnological applications [11,16].

As considered galactosidases employ a common catalytic mechanism, our supposition is that they might share structural similarity, even if they belong to distinct families and clans of glicosidases [3]. Similar structural properties of *alpha*- and *beta*-glycosidases regardless of the low sequence identity have been also obtained for glycoside hydrolases family 4 (GH4) of bacteria [17]. Consequently, the study was focused on characterizing and comparing structural and physicochemical properties of *alpha*- and *beta*-galactosidases using retaining catalytic mechanism in correlation with their biological functions and biotechnological applications.

RESULTS AND DISCUSSION

Multiple sequence alignment of the sequences of every class of investigated *alpha*- and *beta*-galactosidases reveals small alignment scores (Tables 1 and 2). Moreover, the multiple sequence alignment for both *alpha*- and *beta*-galactosidases reveals scores varying between 3.46 and 42.36 and there are not conserved residues or regions in sequences of investigated proteins (data not shown).

Table 1. Alignment scores for different families of *alpha*-galactosidases and their catalytic domains

Taxonomy	Family	Similarity score for the entire sequence	Similarity score for the catalytic domains sequences
Bacteria	GH36	9.24÷96.98	3.87÷98.33
	all	2.26÷96.98	
Eukaryota	GH27	27.25÷46.04	
All investigated *alpha*-galactosidases		2.26÷96.98	

Table 2. Alignment scores for different families of *beta*-galactosidases and their catalytic domains

Taxonomy	Family	Similarity score for the entire protein	Similarity score for the catalytic domains sequences
Archea	GH1	69.73	3.38÷75.76
Bacteria	GH42	30.32	
	GH2	14.82÷32.32	
	all	2.96÷32.32	
Eukaryota	GH35	20.53÷57.34	
	all	5.31÷57.34	
All investigated *beta*-galactosidases		2.96÷69.73	

All these data reflect the low sequence similarity of considered *alpha*- and *beta*-galactosidases. It is also true for the catalytic domains sequences of investigated proteins. Even if the sequence similarity of considered *alpha*- and *beta*-galactosidases is low and they have distinctive number of amino acids in their sequences, their global physicochemical properties (isoelectric point, net charge, aliphatic and GRAVY indexes), computed using ProtScale tool [18] are similar. The average values of these properties, computed for the entire enzymes and for their catalytic domains respectively are presented in Table 3.

Table 3. Average values of the global physicochemical properties of *alpha*- and *beta*-galactosidases and of their catalytic domains

Protein class	pI	Net charge	Aliphatic index	GRAVY index
alpha - galactosidases	5.61±0.49	-13±2	75.98±6.50	-0.386±0.14
beta - galactosidases	5.90±0.65	-17±4	77.83±4.39	-0.353±0.17
alpha-galactosidases catalytic domains	5.39±0.19	-11±3	70.35±3.45	-0.418±0.09
beta-galactosidases catalytic domains	6.05±0.09	-8±1	73.71±3.33	-0.409±0.03

Student statistical test reflects that, at 0.05 level, the average value of every pair of investigated physicochemical properties does not differ significantly either for the two classes of enzymes neither for their catalytic domains. Both *alpha*- and *beta*-galactosidases reflect an acidic isoelectric point correlated to their negative net charges and the global hydrophilic character and they also reflect a high content of aliphatic residues. There is a significant difference between the values of the GRAVY indexes obtained for the entire enzyme compared to its catalytic domain. It is true for both classes of considered galactosidases and reflects the higher hydrophilicity of the catalytic domains.

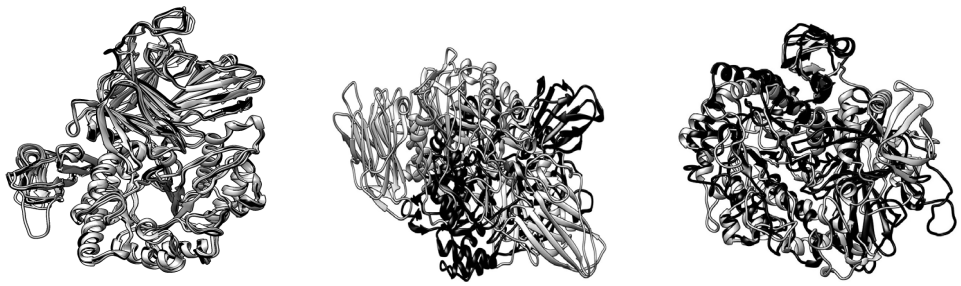
The low sequence similarity between *alpha*-galactosidases and *beta*-galactosidases respectively is not reflected in their structural properties. As expected, the root mean squared deviation (RMSD) values obtained by structural superposition of *alpha*- and *beta*-galactosidases are small for the enzymes belonging to the same GH family (Table 4) reflecting their similar global structural properties.

In the case of alpha-galactosidases, the highest structural similarity is observed for enzymes belonging to the GH36 family (Figure 1.a) and the lower structural similarity is observed for *Bacteroides thetaiotaomicron* alpha-galactosidase (GH97 family) compared to the other alpha-galactosidases. This result is in good correlation with the possibility of this enzyme to employ a distinct catalytic mechanisms (inverting) in comparison to the other alpha-galactosidases.

Investigated *beta*-galactosidases belong to many families and this is reflected in their structural properties, enzymes belonging to the GH35 family sharing the best structural convergence. Between *beta*-galactosidases belonging to GH35 family, the *Caulobacter crescentus* (3U7V) and *Homo sapiens* (3THC) beta-galactosidases are divergent in structure in comparison to *beta*-galactosidases belonging to the other organisms. Their structures are also divergent when compared to the other families of *beta*-galactosidases. Figure 1B illustrates the superposition of human (3THC) and *Trichoderma reesei* (3OG2) beta-galactosidases, RMSD=1.124 Å for only 4 equivalent CA atom pairs.

Table 4. The minimum and maximum RMSD values obtained by the superposition of the three dimensional structures of investigated *alpha*- and *beta*-galactosidases (CA-carbon *alpha*)

Galactosidase class	Galactosidase family	Minimum RMSD / Å^2	Number of equivalent *CA* atoms pairs	Maximum RMSD / Å^2	Number of equivalent *CA* atoms pairs
alpha-galactosidases	GH36	0.813 for 3MI6 superposed to 4FNR	673	1.139 for 3MI6 superposed to 1ZY9	119
	GH27	0.634 for 3LRL superposed to 3A5V	356	1.040 for 3A21 superposed to 1SZN	239
	all alpha-galactosidases	0.813 for 3MI6 superposed to 4FNR	673	1.351 for 2XN0 superposed to 3A24	8
beta-galactosidases	GH1	0.573 for 1GOW superposed to 4HA3	469	-	-
	GH2	0.944 for 1YQ2 superposed to 3OBA	601	1.142 for 3FN9 superposed to 3BGA	260
	GH35	0.612 for 1TG7 superposed to 3OG2	898	1.124 for 3THC superposed to 3OG2	4
	GH42	1.134for 1KWG superposed to 3TTY	430	-	-
	all beta-galactosidases	0.612 for 1TG7 superposed to 3OG2	898	1.124 for 3THC superposed to 3OG2	4
alpha- and *beta*-galactosidases		1.257 for 2XN0 superposed to 3TTS	43	1.090 for 4FNR superposed to 3VD3	4



A B C

Figure 1. Superposition of structures of galactosidases: (A) alpha-galactosidases belonging to the family GH36. (B) *Caulobacter crescentus (3U7V-grey)* and *Homo sapiens* (3THC – black) *beta*-galactosidases. (C) *Lactobacillus acidophilus* alpha-galactosidase (2XN0 - black) and *Bacillus circulans* beta-galactosidase (3TTS - grey)

Superposition of *alpha*- and *beta*-galactosidases structures reflects overall small structural identity, revealing changes in relative orientation of the domains of proteins, all investigated proteins containing more than one domain. For the superposition of *alpha*- and *beta*-galactosidases the RMSD values range between 1.257 Å for 43 equivalent CA atoms pairs in the case of the superposition of the *Lactobacillus acidophilus alpha*-galactosidase (code entry 2XN0) and *Bacillus circulans* beta-galactosidase (code entry 3TTS) (Figure 1C) and 1.090 Å for only 4 equivalent CA atoms pairs in the case of the superposition of the *Geobacillus stearothermophilus* alpha-galactosidase (code entry 4FNR) and *Escherichia coli beta*-galactosidase (code entry 3VD3).

As all investigated enzymes use the same catalytic mechanism, we also compared the structures of their catalytic domains by structural superposition and the results are presented in the Table 5.

Table 5. The minimum and maximum RMSD values obtained by the superposition of the three dimensional structures of the catalytic domains of the investigated *alpha*- and *beta*-galactosidases (CA-carbon *alpha*)

Galactosidase class	Minimum RMSD / Å^2	Number of equivalent *CA* atoms pairs	Maximum RMSD / Å^2	Number of equivalent *CA* atoms pairs
alpha-galactosidases	0.632 for 2XN0 superposed to 4FNR	305	1.179 for 3LRL superposed to 3A24	74
beta-galactosidases	0.573 for 1GOW superposed to 4HA3	469	1.289 for 1GOW superposed to 3U7V	20
Alpha- and *beta*-galactosidases	1.229 for 2XN0 superposed to 1KWG	68	0.939 for 3LRL superposed to 1GOW	7

Figures 2 illustrate the superposition of the catalytic domains of alpha-galactosidases (A) (the structure of the catalytic domain of the *Bacteroides thetaiotaomicron* alpha-galactosidase is presented in black and the oters in grey), beta-galactosidases (B) and of the *Saccharomyces cerevisiae alpha*-galactosidase (code entry 3LRL) and *Sulfolobus solfataricus beta*-gacatosidase (code entry 1GOW) that were identified as being the most divergent (C).

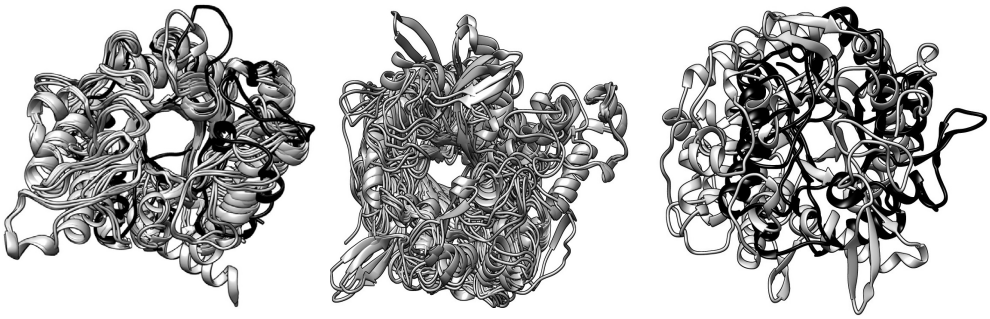


Figure 2. Superposition of structures of the catalytic domains of: (A) alpha-galactosidases (*Bacteroides thetaiotaomicron* alpha-galactosidase is presented in black and the other alpha-galactposidases are shown in grey). (B) *beta*-galactosidases. (C) *Saccharomyces cerevisiae alpha*-galactosidase (code entry 3LRL – black) *and Sulfolobus solfataricus beta*-galatosidase (code entry 1GOW- grey)

The results presented in Table 5 and illustrated in Figures 2 reveal that the best structural superposition is always obtained for the catalytic domains. All these results highlight that *alpha*- and *beta*-galactosidases using the retaining catalytic mechanism share global structural similarity of their catalytic domains and overall analogous physicochemical properties despite their unrelated sequences.

The average values of the normalized surface areas and the surface fractal dimensions of the monomeric sub-units of *alpha*- and *beta*-galactosidases are presented in Table 6. Student's statistical test has been used to analyze these data and the results reflected that the average values of the normalized surface areas and surface fractal dimensions are not significantly distinct for *alpha*- and *beta*-galactosidases confirming once again their global structural similarity. It is also true for the average values of normalized surface areas and surface fractal dimensions of the catalytic domains of investigated galactosidases: there are not significant differences between the average values corresponding to two classes of enzymes and, furthermore, these average values are not distinct by comparison to those corresponding to the entire enzymes.

Student's statistical test also reveals that, at the 0.05 level, the average values of the surface fractal dimensions are significantly distinct for all monomers compared to multimers reflecting distinct surface roughness of monomers and multimers belonging to the same class of galactosidases and reflecting an increased complexity of the surface shape of the polymers compared to the corresponding monomers. Distinct surface fractal dimensions for monomer and multimer proteins have been previously reported [19]. In addition, the greatest

part of investigated structures comes from polymeric proteins and it is already known that protein-protein interactions often occur through flat regions with a large surfaces area [20] explaining the lower roughness of monomers in comparison to polymers. This observation is also important from biochemical point of view if we take into account that oligomerization has been determined to be a decisive process in influencing the substrate binding for the GH27 enzyme subfamily [21].

Table 6. Average values of normalized surface areas (NSA) and surface fractal dimensions (Ds) of alpha- and beta-galactosidases and their catalytic domains

Protein family	Average value of NSA /Å^2	Average value of D_s
alpha – galactosidase monomers	34.78±1.94	2.30±0.18
alpha – galactosidase multimers	-	2.41±0.16
alpha-galactosidase catalytic domains	37.23±1.74	2.38±0.29
beta - galactosidase monomers	36.13±1.85	2.31±0.05
beta - galactosidase multimers	-	2.36±0.08
beta-galactosidase catalytic domains	40.38±1.56	2.34±0.15

Surface analysis of the considered enzymes shows the hydrophobic character of the identified cavities and that bacterial galactosidases expose the larger cavities (Table 7).

Table 7. Mean values of surface cavities and their geometric properties for alpha- and beta-galactosidases

Protein family	Organism (number of analysed structures)	Average molecular weight / Da	Fpocket analysis			
			Mean number of identified pockets	Mean volume of the most probable pocket / Å^3	Hydrophobicity score	Local hydrophobic density
alpha-galactosidase	Bacteria (10)	75005.86±12134.26	38±7	909±125	24.07±3.04	38.42±4.18
	Eukaryota (5)	48319.74±2474.88	16±3	644±214	7.90±1.96	28.67±5.14
beta-galactosidase	Archea (2)	56106.55±827.38	18±3	520±224	42.72±14.26	34.09±6.23
	Bacteria (10)	93544.64±23680.14	35±6	1091±416	10.64±2.06	39.79±8.06
	Eukaryota (6)	100865.85±17775.18	42±7	848±286	23.16±6.28	34.56±6.06

Considering the necessity of both the shape and chemical complementarity between the protein cavity and ligand, the local hydrophobicity is an important indication of ligand accessibility to the internal pocket. More than it, surface cavities identification and characterization in terms of shape and size are important for structure-based ligand design strategies as cavities of larger size may accommodate non-native ligands.

Surface analysis underline that, contrary to the global structural similarity, there are local surface characteristics that reflects the specificity of every subfamily of *alpha*- and *beta*-galactosidases.

CONCLUSIONS

To our knowledge, this is a first study dealing with comparison of *alpha*- and *beta*-galactosidases structural properties both at global and local levels. It reveals similar global physicochemical and structural properties of the *alpha*- and *beta*-galactosidases despite their unrelated sequences and also specific features concerning local properties of their surfaces. *Beta*- galactosidases expose more complex surfaces than *alpha*-galactosidases with much hydrophobic cavities.

The similarity of the global structural properties, especially concerning the catalytic domains, confirms the identical catalytic mechanism and the common evolutionary ancestry of considered enzymes. The distinct local surface features seem to be correlated to their specificity of interactions. The specific identified features are in good correlation with the fact that some investigated *alpha*- and *beta*-glucosidases contain many catalytic domains, every domain belonging to a different GH family. Understanding of the local surface properties of galactosidases in correlation with their substrate specificity becomes very important as it offers new ways to improve, modify or even inhibit their catalytic activity with direct consequences for the biotechnology industries.

EXPERIMENTAL SECTION

Within this study a few bioinformatics tools were used to analyze and compare sequence and structural properties of *alpha*- and *beta*-galactosidases starting from sequence information retrieved from UniProt data base [22] and structure information retrieved from Protein Data Bank [23] and PISA web server [24].

There are 139 entries concerning *alpha*- and *beta*-galactosidases, respectively putative *alpha*- and *beta*-galactosidases using the retaining catalytic mechanism in the Protein Data Bank (PDB) [23]. As PDB usually contains more than one structural file for a given protein, in our study we considered the crystallographic file with the higher resolution concerning the structure of the native protein (when available) or the structural file of its complex with the substrate, product or inhibitor. The codes entry used in this study are the following: 4FNQ, 4FNR, 2XN0, 3MI6, 1ZY9, 3A21, 2YFN, 4NZJ, 3A24, 3GXT, 3LRL, 3A5V, 1SZN, 1UAS for *alpha*-galactosidases and 1GOW, 4HA3, 3VD3, 3BGA, 1KWG, 3TTY, 1YQ2, 3FN9, 3U7V, 4E8D, 3OG2, 3THC, 3OBA, 4IUG, 1TG7 and 3W5G for *beta*-galactosidases, respectively. In the cases of the structures of complexes made by the enzymes, the ligands have been removed and structural analysis has been performed only on the proteins structures.

The investigated proteins belong to different GH families but all of them use the retaining catalytic mechanism and possess the $(\beta/\alpha)_8$ barrel [25] structural motif, that was identified as their catalytic domain. There is one exception: alpha-galactosidase from *Bacteroides thetaiotaomicron* (PDB code entry 3A24). It may act both as an inverting alpha-glucoside and a retaining alpha-galactosidase [26].

The sequence similarity between the investigated proteins was studied using multiple sequence analysis performed with CLUSTALW2 software [27]. ProtParam tool [18] was used to compute some global physicochemical properties of *alpha*- and *beta*-galactosidases based on their sequences: *GRAVY* (GRand Average of HYdropathicity) index indicating the hydrophobic or hydrophilic character of the protein, the net charge, theoretical isoelectric point (*pI*) and the aliphatic index.

The evolutionary relationships between proteins that share sequences with low similarity were detected using structural alignment based on the superposition of the atomic coordinate sets of two or more proteins and a minimal root mean square deviation (*RMSD*) between the structures was computed. *RMSD* reflects the degree of dissimilarity of two three-dimensional protein structures and there are different possible subsets of the protein atoms that can be used in producing a structural alignment and calculating the corresponding *RMSD* values, but usually the *alpha* carbon (*CA*) positions were considered [28]. A zero value for the *RMSD* indicates identical structures and it increases for dissimilar structures. Structural similarity of the considered enzymes was compared using the *structure matching* tool in the Chimera software [29]. For the structural alignment and structural analysis only chains A of polymeric proteins were considered, except the computation of the surface fractal dimensions of polymers where all the protein chains present in crystallographic asymmetric unit were analyzed.

The surface properties of investigated proteins both at local and global levels were analyzed. The local properties refer to one region on the protein surface (such as a cavity or pocket) and global properties refer to the overall protein surface. In order to analyze and compare the surface properties of investigated proteins, the normalized surface area, the surface fractal dimension, the number, geometrical and chemical characteristics of surface cavities were also computed. To eliminate the influence of the protein size, the normalized surface area (*NSA*) was obtained by dividing the molecular surface of each protein (computed using Chimera software) to the number of amino acids in corresponding structure.

For multimeric proteins, the *NSA* value for the monomeric sub-unity A was computed. *NSA* is used as a measure of the compactness of the protein structure, more compact structures having smaller *NSA* values [30]. Surface fractal dimension (D_s) is a quantitative measure of the global surface complexity. This quantity is defined using the fractal geometry concepts and it was computed using the method proposed by Lewis and Rees [31]. This method considers the scaling law between the surface accessible area (SA) and the radius of a rolling probe molecule (R) on the surface and the surface fractal dimension, Df, was determined from the slope of the log(SA) versus log(R). The surface area of the protein was computed using GETAREA software [32] for probe radii of 1, 1.2, 1.4, 1.6, 1.8 and 2 Å respectively.

The detection of cavities and/or protrusions present at the protein surface and characterization of their local geometric and chemical properties was performed using the Fpocket [33] tool.

Almost all of the identified cavities have hydrophobic character. Considering the necessity of both the shape and chemical complementarity between the protein cavity and ligand, the local hydrophobicity is an important indication of ligand accessibility to the internal pocket, information that is also incorporated into ligand design strategies.

ACKNOWLEDGMENTS

Diana-Larisa VLĂDOIU would like to thank the project POSDRU/159/ 1.5/S/137750 - “Doctoral and Postdoctoral programs support for increased competitiveness in Exact Sciences research” cofinanced by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007 – 2013 because this work was supported by this strategic grant.

REFERENCES

[1]. B. Yang, Z. Dai, S.Y. Ding, C.E. Wyman, *Biofuels*, **2011**, 2, 421.
[2]. J.F. Sorensen, K.M. Kragh, O. Sibbesen, J. Delcour, H. Goesaert, B. Svensson, T.A. Tahir, J. Brufau, A.M. Perez-Vendrell, D. Bellincampi, R. d'Ovidio, L. Camardella, A. Giovane, E. Bonnin, N. Juge, *Biochim Biophys Acta*, **2004**, *1696*, 275.
[3]. N. Brás, N. Cerqueira, M.J. Ramos, P.A. Fernandes, "Carbohydrates - Comprehensive Studies on Glycobiology and Glycotechnology", Chuan-Fa Chang Ed., InTech: Rijeka, **2012**; chapter 6.
[4]. E.C. Webb, "Enzyme nomenclature", Academic Press, San Diego, **1992**, 346.
[5]. B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, *Nucleic Acids Res*, **2009**, *37*, D233.
[6]. B. Henrissat, G. Davies, *Current Opinion in Structural Biology*, **1997**, *7*, 637.
[7]. D. Koshland, *Biological Reviews*, **1953**, *28*, 416.
[8]. V.L. Yip, A. Varrot, G.J. Davies, S.S. Rajan, X. Yang, J. Thompson, W.F. Anderson, S.G. Withers, *J Am Chem Soc*, **2004**, *126*, 8354.
[9]. G. Davies, B. Henrissat, *Structure*, **1995**, *3*, 853.
[10]. P. Manzanares, L.H.D. Graff, J. Visser, *Enzyme Microb. Technol*, **1998**, *22*, 383.
[11]. Q. Hussain, *Critical Reviews in Biotecnology*, **2010**, *30*, 41.
[12]. D.H. Calhoun, D.F. Bishop, H.S. Bernstein, M. Quinn, P. Hantzopoulos, R.J. Desnick, *Proc Natl Acad Sci USA*, **1985**, *82*, 7364.
[13]. J. Lukas, A.K. Giese, A. Markoff, U. Grittner, E. Kolodny, H. Mascher, K.J. Lackner, W. Meyer, P. Wree, V. Saviouk, A. Rolfs, *PLOS Genetics*, **2013**, *9*, e1003632.
[14]. C.F. Yang, J.Y. Wu, F.J. Tsai, *Journal of Biomedical Science*, **2010**, *17*, 79.
[15]. M.V. Shumway, P.P. Sheridan, *Int J Biochem Mol Biol*, **2012**, *3*, 209.
[16]. P. Katrolia, E. Rajashekhara, Q. Yan, Z. Jiang, *Critical Reviews in Biotechnology*, **2013**, *8*, 1.
[17]. D. Craciun, B. Vlad-Oros, N. Filimon, V. Ostafe, A. Isvoran, *Acta Biochimica Plolonica*, **2013**, *60*, 553.
[18]. E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, "The proteomics protocols handbook", Humana Press, New Jersey, **2005**, 571.
[19]. D. Craciun, A. Isvoran, R.D. Reisz, N.M. Avram, *Fractals*, **2010**, *18*, 207.
[20]. A.W. White, A.D. Westwell, G. Brahemi, *Expert Reviews in Molecular Medicine*, **2008**, *10*, e8.
[21]. R. Fernandez-Leiro, A. Pereira-Rodríguez, M.E. Cerdan, M. Becerra, J. Sanz-Aparicio, *The Journal of Biological Chemistry*, **2010**, *285*, 28020.
[22]. R. Leinonen, F. Nardone, W. Zhu, R. Apweiler, *Bioinformatics*, **2006**, *22*, 1284.

[23]. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, *Nucleic Acids Res*, **2000**, *28*, 235.

[24]. E. Krissinel, K. Henrick, *J Mol Biol*, **2007**, *372*, 774.
[25]. C. Branden, J. Tooze, "Introduction to Protein Structure", 2nd ed., Garland Publishing, New York, **1999**, 47.
[26]. M. Okuyama, M. Kitamura, H. Hondoh, M.S. Kang, H. Mori, A. Kimura, I. Tanaka, M. Yao, *J Mol Biol*, **2009**, *392*, 1232.
[27]. M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T. J. Gibson, D.G. Higgins, *Bioinformatics*, **2007**, *23*, 2947.
[28]. P.E. Bourne, I.N. Shindyalov, "Structural Bioinformatics", Wiley-Liss: Hoboken, New Jersey, **2003**, 58.
[29]. E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, *J Comput Chem*, **2004**, *25*, 1605.
[30]. N.I. Dmitry, N.S. Bogatyreva, M.Y. Lobanov, O.V. Galzitskaya, *PLoS ONE*, **2009**, *4*, e6476.
[31]. M. Lewis, D.C. Rees, *Science*, **1985**, *230*, 1163.
[32]. R. Franczkiewicz, W. Braun, *J Comput Chem*, **1998**, *19*, 319.
[33]. V. Le Guilloux, P. Schmidtke, P. Tuffery, *BMC Bioinformatics*, **2009**, *10*, 168.