Computational Modelling of Catalytic Properties and Modified Substrates of Fungal β-*N*-Acetylhexosaminidases

Natallia Kulik[1] and Kristýna Slámová[2,*]

[1]*Department of Structure and Function of Proteins, Institute of Nanobiology and Structural Biology of GCRC, Academy of Sciences of the Czech Republic, Zámek 136, CZ 37333 Nové Hrady, Czech Republic*

[2]*Centre of Biocatalysis and Biotransformation, Institute of Microbiology, Academy of Sciences of the Czech Republic, Vídeňská 1083, CZ 14220 Praha 4, Czech Republic*

Abstract: Besides their implication in human physiology and disease, β-*N*-acetylhexosaminidases (EC 3.2.1.52, CAZy GH 20) have recently gained a lot of attention thanks to their great potential in the enzymatic synthesis of carbohydrates and glycomimetics. Extracellular β-*N*-acetylhexosaminidases from filamentous fungi proved to be a powerful synthetic tool for the preparation of both natural and modified glycosides under mild conditions with good yields. A homology model of β-*N*-acetylhexosaminidase from the filamentous fungus *Aspergillus oryzae* has recently been reported, and its quality was corroborated by vibrational spectroscopy and biochemical studies. Computational modelling and analysis helped to identify active site amino acids and other basic structural features of this enzyme important in the catalytic process; moreover, the surface interactions of the subunits of the glycosylated enzyme were identified. The model of β-*N*-acetylhexosaminidase from *Aspergillus oryzae* prepared the ground for further *in silico* studies of enzyme-substrate complexes including prediction and explanation of its substrate specificity.

Keywords: β-*N*-Acetylhexosaminidase, computer modelling, modified substrate, molecular docking, substrate specificity.

1. INTRODUCTION

Structural bioinformatics is a helpful tool for the study of biological systems; the main advantage of computational methods is that they are notably cheaper than the "wet" experiments. Moreover, there are a number of tools and on-line resources for the investigation of biosystems, which offer a close look at molecules on the level of individual atoms [http://www.ncbi.nlm.nih.gov/Tools, http://www.techcuriosity.com/resources/bioinformatics/bioinformatics_tools.php, 1, 2]. For chemists, the structural explanation of some properties of biomolecules, such as proteins, is a challenging task. A lot of structure-dependent properties of biomolecules can be analysed with computational tools: prediction of molecular properties [1, 3, 4, http://www.organic-chemistry.org/prog/peo/], substrate selectivity [5, 6], substrate/inhibitor binding places (active sites) [6], protein structures [7, 8] and posttranslational modifications of proteins (e. g. *N*-glycosylation, disulfide bridges, phosphorylation) [9-11]. Determination of substrate binding places and docking of different modified substrates into the active site are suitable methods for the design of new substrates, inhibitors and drugs [1, 5, 6, 12]. In some cases, computer modelling is used for the verification of working hypotheses, prediction of protein functions and directing of further experimental work [13, 14].

A wide variety of glycosidases have been studied by the means of computational tools, homology modelling experiments yielded new structures of glycoside-hydrolysing enzymes, such as human hyaluronidases [15], invertase [16], human heparanase [17], β-primeverosidase [18] and β-(1,3)-glucan transferase, a member of GH 72 family [19]. The resulting models enabled further investigations of the catalytic mechanism and substrate binding sites of these enzymes. Using molecular modelling the catalytic mechanism of a glycosidase with known 3D structure can be proposed, as in the case of bacteriophage endosialidase [20] and fungal polygalacturonase [21]. Computational modelling is also regularly employed in the rational design of enzyme-targeted drugs and inhibitors [17, 22] and antibodies [23]. Substrate specificity of β-glucosidases with a special focus on the aryl-aglycone part of the substrates has been thoroughly explored [24, 25]. Last but not least, homology modelling was successfully used for the prediction and explanation of the properties of mutant enzymes, be it mutations in the active site [26] or surface mutations in the studies of protein stability in high temperatures [27] and alkaline environment [28].

Among the GH 20 family β-*N*-acetylhexosaminidases there are only a few bacterial representatives with their 3D structure resolved, namely *Serratia marcenscens* [29], *Streptomyces plicatus* [30], *Streptococcus gordonii* [31], *Actinobacillus actinomycetemcomitans* [32], and *Paenibacillus* sp. [33]. The eukaryotic kingdom is represented only by the crystal structures of human isoenzymes HexA and HexB [34-36]. Human lysosomal β-*N*-acetylhexosaminidases HexA and HexB have been extensively studied because the deficiency of these enzymes causes severe to fatal neurodegeneration associated with Tay-Sachs (TSD [37]) and Sandhoff diseases (SD [38]), respectively. Homology models of the α-subunit of human HexA [29, 39] and of *Streptomyces plicatus* β-*N*-acetylhexosaminidase [40] were accomplished shortly after the structure of β-*N*-acetylhexosaminidase from *Serratia marcescens* had been published. Authors [29, 41, 42] used the models of HexA and HexB for mapping the mutations leading to different forms (infantile, juvenile, adult, benign) of TSD and SD respectively, in order to propose molecular mechanisms of these diseases. Analysis of structural changes in the proteins, caused by those mutations, helped to better understanding of the pathogenesis of the respective diseases. Kaplan with co-workers [43] used the information on the structure of HexA (the model was built by Tews *et al.* [29]) for the determination of amino acids responsible for the specificity towards the electronegative substrate 4-methylumbelliferyl-*N*-acetylglucosamine-6-sulfate and proposed a methodology for such type of investigation based on the structural information. Recently, substrate specificity of *A. actinomycetemcomitans* β-*N*-acetylhexosaminidase called Dispersin B was investigated by the docking of various types of possible substrates, revealing that Dispersin B is an *exo*-acting enzyme hydrolysing the β-(1,6)-linked *N*-acetylglucosamine oligomers [44]. Also quantum mechanical/molecular mechanical (QM/MM) calculations were applied for the detailed investigation of active site amino acids of a family 20 glycoside hydrolase [45].

*Address correspondence to this author at the Centre of Biocatalysis and Biotransformation, Institute of Microbiology, Academy of Sciences of the Czech Republic, Vídeňská 1083, CZ 14220 Praha 4, Czech Republic; Tel: +420-296442677; Fax: +420-296442509; E-mail: slamova.kristyna@gmail.com

1570-193X/11 $58.00+.00

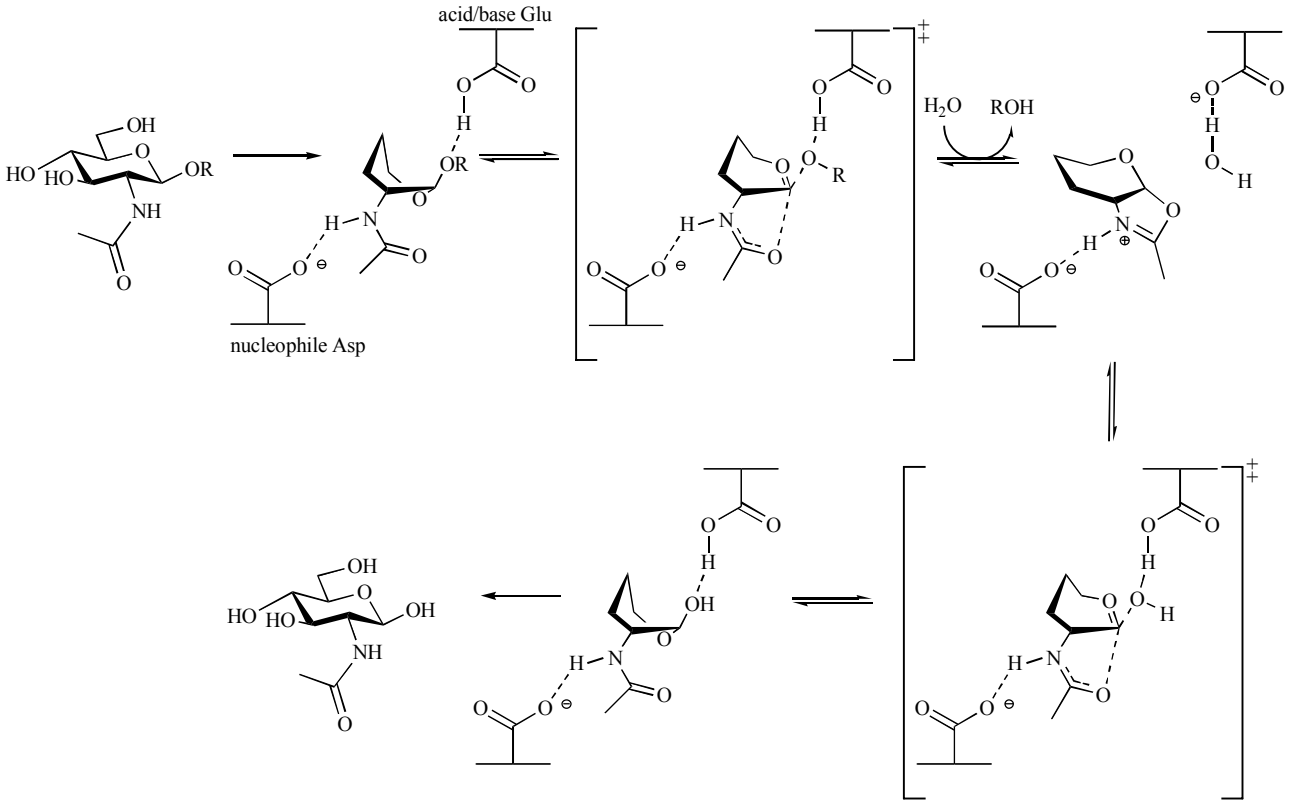

Fig. (1). Double-displacement retaining mechanism exploited by GH family 20 β-*N*-acetylhexosaminidases.

The broad application of glycosidases in biotechnology, biotransformation and chemoenzymatic synthesis of complex carbohydrates attracted interest in fungal β-*N*-acetylhexosaminidases (CAZy Glycoside Hydrolase family 20, http://www.cazy.org, [46]), which have been proven to be robust and effective biocatalysts [47-49]. Unfortunately, despite the need of knowledge of their architecture, no crystal structure of a fungal β-*N*-acetylhexosaminidase has been reported so far. Quite recently, a homology model of the dimeric glycosylated form of β-*N*-acetylhexosaminidase from *Aspergillus oryzae* CCF 1066 has been presented; the quality of the model was confirmed experimentally [50]. This model was used for the docking and molecular dynamics simulation (MD) of a variety of structurally modified substrates [47-49, 51]. This mini-review is focused on the application of the homology model of β-*N*-acetylhexosaminidase from *Aspergillus oryzae* CCF 1066 for the explanation of substrate specificity of fungal β-*N*-acetylhexosamini-dases towards a wide panel of biotechnologically interesting unnatural substrates, which is the main research focus of the authors. The results of the computational studies were found to be useful not only for the description and characterization of structural features of the enzyme, but also for the explanation of binding affinities and specificity towards modified substrates. A number of methods have been used to analyse the interactions in the enzyme-substrate complexes and to estimate the binding energies.

2. CATALYTIC MECHANISM AND AMINO ACIDS OF THE ACTIVE SITE

The active site of the GH family 20 β-*N*-acetylhexosaminidases contains a highly conserved pair of catalytic residues **Asp-Glu**, and it was proposed shortly after the first crystal structure of a bacterial hexosaminidase with chitobiose bound in its active site was resolved [29, 52]. The following studies benefiting from the increasing number of known hexosaminidase 3D structures confirmed the original hypothesis of Tews and co-workers. They proposed that the catalysis is carried out by the glutamate acting as a proton donor, while the substrate's 2-acetamido moiety acts as a nucleophile instead of the aspartate to form an oxazoline intermediate [30, 31, 53, 54]. This modified catalytic mechanism of retaining glycosidases is referred to as "substrate-assisted catalysis" or sometimes as "anchimeric assistance", and does not involve the covalent glycosyl-enzyme intermediate as in the classical double-displacement mechanism of retaining glycosidases (Fig. **1**). For the determination of the conformation of the active site during catalysis, the stable mimic of the transition state and also strong hexosaminidase inhibitor NAG-thiazoline was co-crystallized with the respective hexosaminidases [30, 31, 55]. Besides the mentioned catalytic residues, there are some more conserved amino acid residues located in the active site of GH 20 hexosaminidases involved in the substrate binding or stabilization of the oxazoline reaction intermediate (Table **1**; Fig. **2A**).

In the following description of the active site amino acids and their function, the numbering of the amino acids corresponds to *S. plicatus* hexosaminidase (Fig. **2A**), the analogous residues from other enzymes are listed in Table **1**. In the experiments with mutated hexosaminidases, the active site catalytic aspartate was found to be important for the correct orientation of the acetamido group during the nucleophilic attack as well as for the stabilization of the transition state [56, 57]. One tyrosine residue (Tyr393) stabilizes the twisting of the chitobiose nonreducing part in the unfavorable boat conformation by hydrogen bonding [29, 34, 52]. Tyrosine and aspartate (Asp313) positioned on either side of the pocket in the proximity of the substrate 2-acetamido group help to lock this moiety into the correct position enabling the nucleophilic attack, moreover, they also enhance the polarization of the amide thus increasing the nucleophilicity of the carbonyl oxygen [29, 30]. The guanidium group of active site arginine (Arg162) directly binds the substrate by forming two hydrogen bonds to OH-3 and OH-4 of the nonreducing *N*-acetylglucosamine moiety of chitobiose, and thus

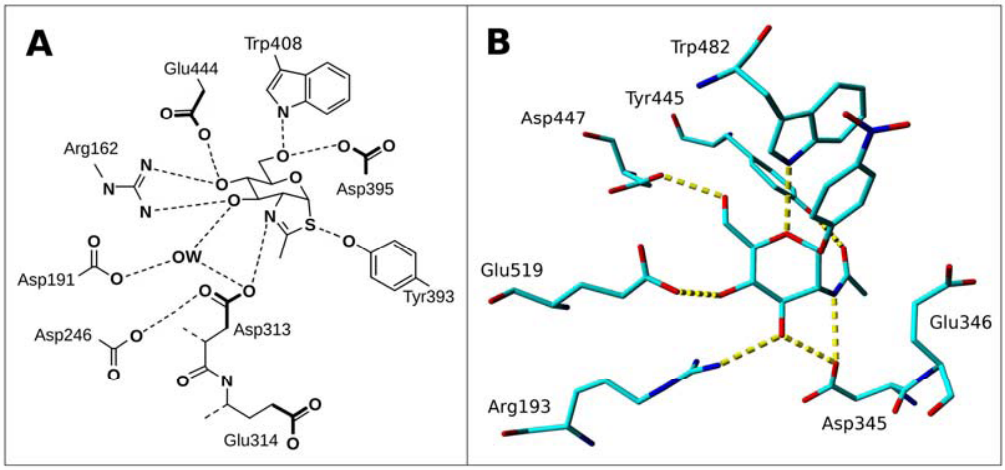

Fig. (2). A: The nonhydrolysable transition state analogue NAG-thiazoline bound in the active site of *Sp*Hex used as the starting point for MM and QM/MM MD simulations of oxazolinum ion [45]. **B**: The stick representation of the active site of β-*N*-acetylhexosaminidase from *A. oryzae* with docked standard substrate *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside (*p*NP-GlcNAc) after 10 ns of MD simulations. Hydrogen bonds are shown in yellow.

Table 1. Numbering of Active Site Amino Acids of β-*N*-Acetylhexosaminidases from Different Sources

Amino Acid Residue	Human α-subunit [36], HexA	Human β-subunit [34], HexA	*Serratia marcenscens* [29]	*Streptomyces plicatus* [30]	*Aspergillus oryzae* [50]
Trp**	-	-	685	**408**	482
Trp	392	424	639	**361**	419
Trp	373	405	616	**344**	397
Trp	460	489	737	**442**	517
Tyr	421	450	669	**393**	445
Arg	178	211	349	**162**	193
Glu	462	491	739	**444**	519
Asp**	-	452	671	**395**	447
Asp	258	290	448	**246**	271
Asp*	322	354	539	**313**	345
Glu*	323	355	540	**314**	346

* Catalytic amino acids
** Important for hydrolysis but not highly conserved amino acids, substituted in human α-subunit by Asn423.

stabilizes the transition state [34, 40, 57, 58]. Mutation of this residue strongly affects binding, as observed in the human HexA α-subunit, where the mutation Arg178His is associated with the B1 variant form of TSD connected with the lack of sufficient enzymatic activity [59, 60].

The stabilization of the nitrogen positive charge during the formation of the cyclic intermediate is provided by the hydrogen-bonding network between Asp313, Asp246, and the main chain NH group of Met247. This network was proposed to neutralize the positive charge of the substrate acetamido nitrogen atom by delocalized negative charges of carboxylate oxygens acting as acceptors [30]. Further structure and kinetic study confirmed this idea, however showed that small nucleophilic groups (azide) in substrates could also attack the anomeric carbon of the oxazolinium ion intermidiate to let rescue of product proceed [57]. The pK_a value of oxazolinium ion higher than pK_a of Asp313, calculated from QM/MM simulation for hexosaminidase from *Streptomyces plicatus*, corresponds to electrostatic stabilization of the enzyme-bound intermediate by Asp313 [45]. Greig with co-workers investigated different protonation states of active site residues Glu444 and Asp395, which determine the structure of the enzyme-bound intermediate. Finally, the authors underlined that the existing computational methods should be completed by kinetic and structural studies to determine alternative hydrolytic pathways.

The conformation change of pyranosyl ring of the nonreducing β-*N*-acetylhexosamine during hydrolysis from the unfavourable distorted sofa/boat conformation (bound substrate) through the structure with coplanar C-1, C-2, C-5 and O-5 into a relaxed chair conformation (oxazoline transition state) is proposed to be essential for the substrate assisted catalytic mechanism. This motion allows the substrate acetamido-O-7 atom to move within 3.0 Å of the C-1 atom of the same glycosyl, being in an ideal position for the attack of the nucleophilic base [29, 30]. The distortion of geometry to sofa/boat conformation is proposed to be assisted by hydrophobic interactions with tryptophan residues 344, 361 and 442 packed around the hydrophobic surfaces of the pyranose ring [29, 30]. Moreover, the polar interactions of Asp313 and Tyr393 with *N*-acetyl group participate in the stabilization of sofa conformation [29]. Hydrophobic stacking was found to be very important during the binding of carbohydrates [29, 61]. Position of the reducing end (leaving group in hydrolysis) in hexosaminidases from bacteria and fungi is stabilized by π-π interaction with Trp408 [30, 50].

On the base of a homology model, Arg424 was found to be responsible for the substrate specificity of the α-subunit in the dimeric structure of human HexA, which is able to accept negatively charged substrates such as 4-methylumbelliferyl-*N*-acetylglucosamine-6-sulfate. This hypothesis was corroborated by side-chain mutagenesis and kinetic studies [43].

The role of some active site amino acids of hexosaminidase from *A. oryzae* in the catalytic process is described in chapter 4. The active site of the enzyme with docked natural substrate (chitobiose) is presented at the Fig. (**2B**).

3. HOMOLOGY MODELLING

The lack of experimentally determined crystal or NMR structures of fungal β-*N*-acetylhexosaminidases can be solved by computational modelling. Generally, there are two main methodological approaches in protein modelling: *ab initio* or *de novo* modelling [62] and comparative modelling [63], which is also called knowledge-based or homology modelling. The first method produces a structure on the base of the physical principles underlying protein folding; however, this method is exhaustive and time-consuming and presently cannot be used with good reliability for systems with a number of amino acid residues exceeding 150. Until now, homology modelling has been the most convenient and successful method for the prediction of structures of robust macromolecules [63]. Despite the numerous on-line programmes and tools for homology modelling, the best models are built with knowledge-based human intervention [64].

A widely employed similarity search algorithm is BLAST [65], which is faster than the other popular heuristic algorithm FASTA [66]. For the search of remote homology, a method based on hidden Markov models seems to be an optimal choice [67]. A number of approaches, methods and programmes can be employed to align sequences and 3D structures [68-72]. One of the most frequently used multiple alignment programmes is ClustalX [73], based on the progressive alignment algorithm ClustalW. However, as a global alignment tool, it is prone to fail in aligning sequences with significantly different numbers of residues [70]. The recently introduced T-Coffee server (http://igs-server.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi) combines a variety of algorithms (ClustalW (http://www.clustal.org/), ProbCons (http://probcons.stanford.edu/), MUSCLE (http://www.drive5.com/muscle/), DIALIGN (http://dialign.gobics.de/chaos-dialign-submission), PCMA, POA, MAFFT) and methods operating with various data types [68]. When aligning primary sequences one has to have in mind that misalignments may occur in sequence variable regions, represented in general by random coil structure in loops and at the N-/C-terminal parts of the protein [74, 75].

A reliable sequence alignment is the most important step in homology modelling, as misalignment leads directly to a wrongly generated three-dimensional structure [76, 77]. The alignment of the fungal β-*N*-acetylhexosaminidase with primary sequences with known three-dimensional structures (1c7s (*Serratia marcescens*), 1jak (*Streptomyces plicatus*) and 1now (human β-subunit)) fulfills the requirements for a reliable alignment in the active site and for the structurally conserved regions [50]. A 3D model of hexosaminidase from *Aspergillus oryzae* was generated [50] using Modeller [78], which is a restraint-based homology modelling programme, and the basic method used to build homology models stored in the ModBAse [79] database. Structural alignment was performed to correct the final alignment at the variable regions, which is extremely important in the case of the templates with different number of domains, where ClustalX is not satisfactory [50].

There are several tools for estimating the quality of the acquired model [8, 69, 72, 80-82]. Their common feature is that most of them are based on the statistical data, extracted from the structures in PDB; these could be geometric parameters [80] or calculated energy values [81]. Computational biology also offers a dynamic way to evaluate a model by molecular dynamics simulations (MD) to assess the time stability of the modelled structure in solvent. The MD hereby may either lead to a refinement of the structure or can point directly to shortcomings of the homology model. These shortcomings are mostly related to regions in the so called "twilight zones" (regions with low identity) and long loops (more than 9 amino acid residues), that are difficult to model with high reliability and often the modelling results in several equally weighted solutions [74, 75, 83]. Loop regions are mostly highly flexible and often poorly resolved in the templates. This problem is actual for the modelling of fungal hexosaminidases, as these comprise a long loop. The structural stability of a model can be assessed in the molecular dynamics simulation using the root mean square deviation (RMSD) of the individual residues from their initial positions. Highly conserved residues, and especially active site residues, hereby should not deviate more than is their actual conformational flexibility and are therefore a good indicator to monitor the local stability. Even though there are a number of computational methods for model verification, experimental validation of the model is always a great advantage.

The architecture of fungal β-*N*-acetylhexosaminidases differs in several aspects from bacterial and human analogues and its structural knowledge can guide to functional understanding and interpretation of these differences. Therefore, the lack of structural information was overcome by building a model of β-*N*-acetylhexosaminidase from *Aspergillus oryzae* CCF 1066 [50]. The quality of the model was verified by computational methods using Procheck programme and analysing the stability of protein during equilibration and by spectroscopic methods. The positions of the disulfide bridges were determined with the help of mass spectrometry and Raman spectroscopy, which are suitable methods for the prediction of the conformation of disulfide bridges. The percentage of secondary structure elements in the model is comparable with the least square analysis by infrared and Raman spectroscopy [50]. The RMSD of C-alpha atoms of the equilibrated model oscillates around 0.3 Å. This model answered the questions concerning the active site amino acids and their organization, and helped to investigate the composition of the dimeric glycosylated native form. Molecular dynamics simulations allowed to study the effect of dimerization, disulfide bonds and glycosylation on protein stability, and lead to a detailed analysis of common and specific features among this class of enzymes from different organisms.

The final model has a kidney-shaped structure of catalytic subunit (Fig. **3**). The active site amino acids are located in the central part of the TIM barrel. The architecture of the model is highly conserved and well superimposed with the active sites of the templates (Fig. **3B**). During the molecular dynamics experiments, the catalytic amino acids are more flexible than the other active site amino acids according to their calculated RMSF (root mean square fluctuation), which is connected with the role of Asp345 and Glu346 in the proper orientation of the *O*-glycosidic bond for the hydrolytic process [50].

Homology modelling revealed a flexible loop anchored by the residues Cys448-Cys483, which was found to be exclusive for fungal enzymes and is proposed to play a special role. This loop is located close to the active site and is a part of the contact surface of monomers in the dimeric form, where it bends over the active site of the other monomer acting as an active site lid (Fig. **3D**, **E**). A similar loop was reported for other enzymes with analogous catalytic domain architecture [84-87], but not for any of known crystal structures of β-*N*-acetylhexosaminidases from human and bacteria.

The common feature of human and fungal β-*N*-acetylhexosaminidases from *A.oryzae* is the existence as a dimer, while bacterial hexosaminidases are active as monomers. Interestingly, the fungal enzyme conserves its activity even in the monomeric state, which could not be explained by molecular dynamics. Dimerization of fungal β-*N*-acetylhexosaminidases appears to be reversible and strongly pH dependent; below pH 2.5 the enzyme is present as a monomer, however, at higher pH the enzyme aggregates into more stable dimers [88]. The dimeric interface is formed by 30 % of hydrophobic amino acids and also by a number of charged acid/base

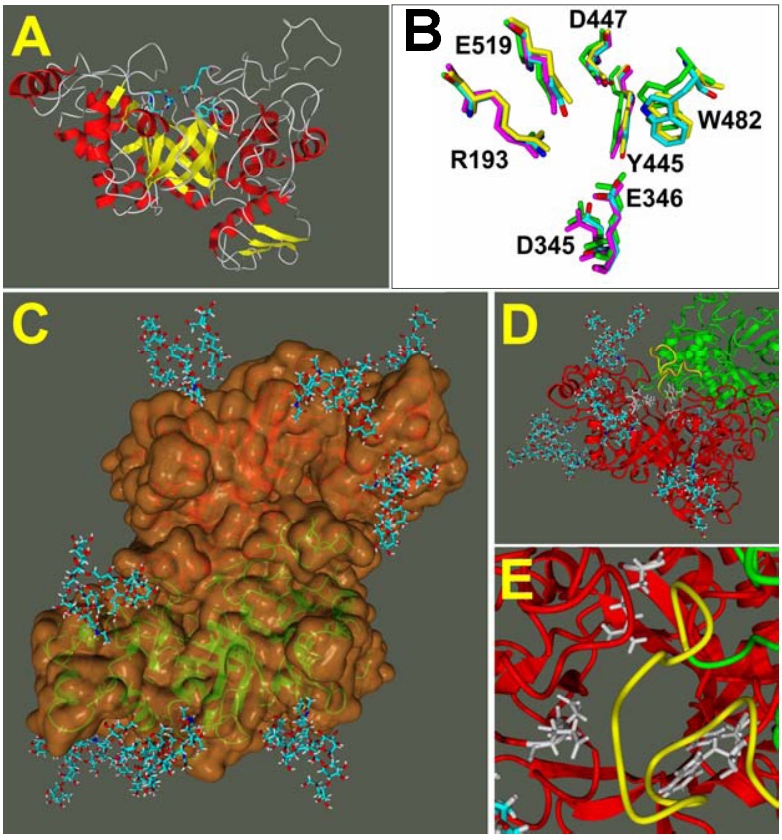

Fig. (3). Molecular models of β-*N*-acetylhexosaminidase from *Aspergillus oryzae*. (**A**) Side view of the model of the catalytic subunit. (**B**) Overlay of the active site residues of the refined homology model of β-*N*-acetylhexosaminidase from *Aspergillus oryzae* with the known crystal structures: 1c7s: *Serratia marcenscens* (yellow), 1jak: *Streptomyces plicatus* (green), 1now: *Homo sapiens* (magenta). In the human structure the tryptophane residue is not conserved, and in the *Serratia marcenens* structure the aspartic acid is mutated to alanine. (**C**) The fully *N*-glycosylated dimer. (**D**, **E**) The large flexible loop shown in yellow of the green subunit is just about 1 nm above the active site residues (shown in grey) of the red subunit [50]: side view (**D**), top view (**E**).

amino acid residues, providing the ionic interactions. The glycosylation of the surface of the dimeric enzyme form plays an important role in the reversibility of dimer assembly, making it irreversible for the deglycosylated protein.

4. DOCKING OF UNNATURAL SUBSTRATES INTO THE ACTIVE SITE

Generally, there are some well established approaches for docking the substrates into the active centre of an enzyme, which include the following steps: determination of the substrate binding place, fitting the substrate into the binding pocket, evaluation of binding and selection of the best conformation [89]. There are a number of progammes using different algorithms for docking substrates into the active site – DOCK, FlexX, ICM, ADAM, FLOG, Prodock, GOLD etc. [90]. The reproducibility of Glide, GOLD, DOCK, AutoDock, FlexX was analyzed for the docking of carbohydrates [91, 92]. Agostino *et al.* reported that Glide reproduces the position of the substrate at the active site better than GOLD, AutoDock, FlexX, however, AutoDock also produces results with reasonable accuracy for rigid docking [92].

Nowadays, one of the most widely used docking methods is Autodock, which was also applied in the docking of substrates (**5**,**13**,**20**-**23**) into the active site of β-*N*-acetylhexosaminidase from *A.oryzae*. The main disadvantage of this method is the generation of false positive hits resulting from substrate conformations rarely obtained in solution [93]. As a result, the enzyme-substrate system often requires examination of the obtained conformation and position of substrate at the active site. Correct modelling of conformation is especially important, but difficult task for such flexible substrates as carbohydrates [61]. The advantage of using molecular dynamics to explore docked conformations is on one hand in monitoring the stability of the whole complex and the substrate conformation in explicit solvent, on the other hand in the dynamics of the system, following the idea of 'induced fit' of substrates in the active site [94]. A drawback of the use of MD approach in high-throughput screening of ligand-databases is the major time consumption; in spite of this, using a limited set of potential substrates, the combination of Autodock and molecular dynamics can yield reliable results [89, 93, 94].

The complex approach was employed in docking substrates into the active site of hexosaminidase from *A.oryzae* [50]. Since β-*N*-acetylhexosaminidases were found to possess a well determined and highly conserved substrate binding pocket, the initial conformation for the natural substrate *N*,*N*'-diacetylchitobiose was gained by an overlay of the model of hexosaminidase from *A.oryzae* with the crystal structure of *Serratia marcescens* complexed with chitobiose (PDB code: 1qbb) [50]. Modified substrates (Table **2**) were placed into the active site using an arbitrary initial position close to the coordinates of the natural substrate [14, 47-49, 51]. For substrates **1**-**19** precise positioning was done using the DOCK module included in SYBYL/MAXIMIN2 followed by energy minimization to improve side chain positions. For substrates **2** and **20**-**23** MD simulations of preminimized enzyme-substrate complexes were performed using YASARA, to improve the orientation of the substrates in the active site and to study the dynamics of the enzyme-substrate interactions in explicit solvent.

Conclusions with respect to the ability of fungal hexosaminindase to bind and hydrolyse a particular substrate were based on

geometrical and energetic parameters. The authors explained the connection of the value of binding energy with geometrical parameters, breaking the overall value down to the contribution coming from concrete interactions of single amino acids. This implies the mutual orientation of the substrate and active site amino acids [14, 51, 95], distance from active site residues to substrate [14, 48-50, 95] and RMSD of the active site amino acids during the time course of the simulation [50]. The initial orientation of the substrate at the active site is very important. Some unsubstantial enzyme-substrate contacts could be lost or successfully substituted by other interactive pairs and influence just the rate of cleavage of substrate. Other contacts, especially those concerning catalytic amino acids, are essential for the nucleophilic attack [48], proper binding of the substrate [49] and for transition state stabilization during hydrolysis [14].

Binding energies were calculated by three methods: Dock module in SYBYL/MAXIMIN2 using the TRIPOS force field (substrates **1**-**19**), a binding energy estimation including solvation in YASARA using the Yamber 2 force field (substrate **2**), and AutoDock (scoring function based on Amber force field, substrates **5**, **13**, **20**-**23**). YASARA and SYBYL use the term 'binding (interaction) energy' for the designation of 'strength' of binding on the base of implied force field function, while AutoDock produces free energy of binding (further we will use 'binding energy' with link to method used). It is important to mention here, that the binding energies of natural and modified substrates calculated in SYBYL, YASARA and AutoDock are neither identical nor comparable; the differences originate in various force fields, parameters and equations used. However, for each programme used, the results can be compared and the substrates can be ranked according to their binding affinity. The particular values of binding energies of various unnatural glycosides into the active site of the model of hexosaminidase from *A. oryzae* calculated in respective programmes are presented in Table **2**.

The calculations of the solvation effect and entropic cost of binding are the most challenging parts [14]. Binding energy in YASARA is calculated as the difference of the sum of potential and solvation energies of separated components and the sum of potential and solvation energies of the complex based on the selected force field. This computation usually yields a positive energy value; thus, the higher value corresponds to better binding, the value is relative and includes the solvation effect. On the contrary, the TRIPOS force field neglects the solvation and desolvation effects [14, 47-49, 51] and interaction energy calculated by SYBYL includes steric and electrostatic contributions. The electrostatic term of the interaction energy strongly depends on the number of hydrogen bonds created between the docked substrate and the enzyme [49]. The binding energy calculated by AutoDock is a score function based on derived parameters with the aim to estimate a value close to the experimental free energy of binding. A combination of docking and MD simulations was used for the analysis of enzyme-substrate interactions in case of 2-acetamido-2-deoxy-D-mannopyranose-containing disaccharides [14], immunoactive *N*-acetyl-D-galactosaminides [48], glycosyl azides [49], *N*-acyl modified substrates [47], cyanodeoxy-glycosyl derivatives [51], 4-deoxy-substrates [95] and chitobiose [14].

Position of the natural substrate **2**, chitobiose**,** in the active site was analyzed by SYBYL and YASARA. Stabilization of its binding energy with only small fluctuations of active site amino acids after 1.6 ns of MD simulation in YASARA means that chitobiose reached a stable position within the active site. As the substrate did not induce structural changes in the active site with an RMSD of the active site amino acids less than 0.17 nm, the active site of the modelled enzyme seems to behave in the same way as seen in the bacterial crystal structures. Chitobiose established hydrogen bonds with active site residues Arg193, Asp345, Asp447, Trp482, Glu519 and hydrophobic interaction with Trp482. One should have in mind, that only the fact that the active site is conserved between bacterial, human and fungal hexosaminidases enables the authors to perform reliable docking experiments, as the general rule states that an overall sequence identity of around 30% leads to homology models with low resolution and imprecise placement of side chains.

The interaction energy of substrate GlcNAcβ(1→4)ManNAc (**1**) calculated by SYBYL is similar to substrate GlcNAcβ(1→4) GlcNAc (**2**), used as a standard [54], so apparently it binds to the active site. However, the two times higher steric energy is the result of a steric conflict with Trp482 fixing the ManNAc moiety, which explains the 0.3 Å increase of distance to the catalytic residue Asp345. The loss of this interaction, essential for stabilization of oxazolinium intermediate, explains why neither cleavage nor inhibition was observed experimentally for this case. The proper positioning of the substrate in the active site, which enables the next step of the enzymatic reaction, is as important for hydrolysis as the overall binding energy. This could be demonstrated for substrates **3** and **4** [14], which were found to occupy a similar position in the active site after MD.

Fialová with co-workers reported energy calculations by SYBYL taking into account the existence of different resonance structures in the solution in the equilibrium [49]. Two possible resonance structures of the azido group of glycosyl azides were prepared, docked and analysed during MD ($-N=N^+=N^-$ and $-N^--N^+\equiv N$) [49]. The overall interaction energies for both structures in *gluco*-configuration were very close to the natural substrate, with the value being slightly lower for $-N=N^+=N^-$, minor differences were obtained in the steric and electrostatic energies. On the contrary, docking of both resonance structures of the *galacto*-substrate exhibited a more significant drop in total interaction energies, -167 and -164 kJ/mol compared to 221 kJ/mol for the standard substrate, which is essentially caused by the difference in electrostatic energy contributions. Interestingly, the azido substrate in *galacto*- configuration (**12**) was not cleaved by the enzyme in the wet experiment, while the *gluco*-configurated substrate (**11**) was well cleaved as well as transferred onto acceptors by the fungal enzymes. A closer look at the binding of the above compounds in the active site by visualization in SYBYL shows the absence of the hydrogen bond to Glu519 in the case of the *galacto*-azide **12** as a consequence of the distance of the respective moieties. Therefore, the electrostatic and consequently the total energies decreased compared to the substrate **13** used as standard. These results suggest that compound **12** is not bound in the active site properly and explains the experimental observations.

Docking of C-6 oxidized *N*-acetyl-D-galactosaminides (**14**-**16**) contributed to the understanding of the results obtained experimentally in cleavage experiments. The computational study of immunoactive C-6 oxidized *N*-acetyl-D-galactosaminides (**14**-**16**) reported in [48] is particulary interesting because of the character of the substrates (electronegative moieties). The uronate **15** and its methyl ester **16** were not cleaved by any of the tested enzymes. Due to structural features of the substituting groups they lost the ability to form certain hydrogen bonds and have a significantly less favourable binding energy than standard **13**. The conclusion that insufficient binding of substrate **15** is not resulting from the charge is an important finding for the future study of substrates with charged substitutions at C-6 atom. Additionally, the calculations were able to answer an interesting question regarding the hydrolysis of the aldehyde: "Do β-*N*-acetylhexosaminidases recognise the aldehyde as such or in its hydrated form as a geminal diol, the prevailing form in the water solution [96, 97]"? The SYBYL-calculated binding energy of the geminal diol form of substrate **14** confirmed that it is recognised by the enzyme as such. The geminal diol form binds better because it adopts a more stable position in the active site forming two additional hydrogen bonds [48]. These examples demonstrate nicely how the docking experiments can help to interpret experimental data on an atomic scale.

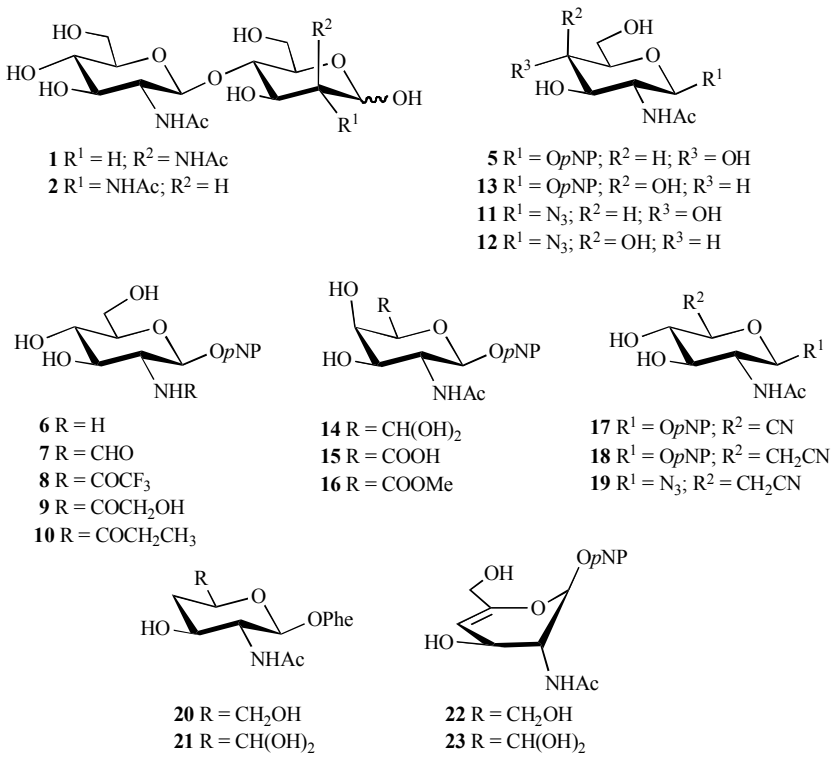

Fig. (4). Structures of substrates docked into the active site of β-*N*-acetylhexosaminidase from *A. oryzae*.

A logical follow up to the above cited work are functional and structural consequences of modifications of the *N*-acyl moiety. The ability of fungal hexosaminidases to cleave and transfer glycosides bearing a modification at this site was screened and modelled [47]. Neither of the five *N*-acyl modified substrates (**6**-**10**) exhibited substantial steric conflicts with the active site of the enzyme, obvious differences were observed in the electrostatic interaction energies. The lower, or even no hydrolysis in the case of compounds **6** and **8**, is proposed to be caused by the destabilization of the oxazolinium reaction intermediate. However, this cannot be simulated by a simple docking experiment.

Structural analysis of the glycoside-enzyme complex energy may not be used only to propose reasonable or interesting substrate modifications, but could additionally lead to the prediction of efficient hexosaminidase competitive inhibitors [51]. These compounds should have a favourable binding energy, but should not allow the nucleophilic attack that starts the reaction. The calculated binding energies for nitrile compounds **17**-**19** are comparable to the standard substrates **5** and **11**; however, these compounds are not cleaved by the enzymes indeed. Based on molecular dynamics simulations, the authors proposed that glycosides **17**-**19** could act as competitive inhibitors of hexosaminidase from *A. oryzae*. This assumption was verified for substrates **17** and **18** by a kinetic inhibition study [51]. The lack of inhibition by substrate **19** could be explained by weaker binding energy than substrates **17**-**18** and standard **11**.

The recent study on the 4-deoxy substrates of β-*N*-acetylhexosaminidases estimated the free energy of binding by AutoDock by performing a rigid docking of the substrates **20**-**23** pre-equilibrated with YASARA. Despite the complex approach, the calculated values did not bring complete information for making clear conclusions with respect to the cleavage of the substrates [95], but additional logical thinking was necessary to interpret the gained structures and values. Molecular dynamics simulations of the 4-deoxy substrates with YASARA revealed that for the successful cleavage of substrates by hexosamininidase from *A.oryzae*, the fixed position of the C-3 hydroxyl of the substrate pyranose ring in the active site is required, which is mainly realized by Arg193 and Asp345. The C-4 dehydration strongly influenced the overall geometry of substrates **22** and **23**, leading to distortions of the pyranose ring. After 10 ns of MD the substrate showed C-3, C-4, C-5 and O-5 in the same plane and C-1, C-2 at different sides out of plane, a conformation that is not acceptable for hydrolysis by hexosaminidases [52, 95, 98]. MD showed that the C-4 oxidized substrates **22** and **23** (Table **2**, Fig. **4**) lost the important interaction with Arg193 and the C-3 position is not stabilized by any other interaction. The role of the proper orientation of the C-3 atom is proposed to be important for charge distribution during the formation of the oxazolinium ion. As a result, these glycosides (**22**, **23**) were not hydrolysed by the enzyme in the wet experiments, despite of the fact that computationally determined binding energies of the substrate **23** and standard substrate **5** are comparable (-29 and –34 kJ/mol respectively) [96]. Favourable binding affinity of the enzyme to the substrate in the active site is important at the initial step of hydrolysis (Fig. **1**). The whole reaction, however, involves additionally the stabilization of the oxazolinium ion intermediate. In case of substrates **22** and **23** the formation of the intermediate was found to be impossible. This nicely demonstrates that only a careful interpretation, taking into account additional knowledge about the reaction mechanism, can make real use of the gained computational data of the enzyme-substrate interactions in β-*N*-acetylhexosaminidases.

4. CONCLUSION

This mini-review presents the application of computational methods in the structure study of β-*N*-acetylhexosaminidase from *Aspergillus oryzae*, which is an enzyme of biotechnological interest. Using homology modelling, a three-dimensional model of the dimeric and glycosylated form of this enzyme was constructed, providing basic information about its conformation, function and

Table 2. Binding Energies of Substrates Docked at the Active Site of *A.oryzae* β-*N*-Acetylhexosaminidase

	Substrate	Ratio of Specific Activity*	Interaction (binding) Energy [kJ/mol]			Method	Ref.
			Total	Steric	Electrostatic		
1	GlcNAcβ(1→4)ManNAc	-	-246	-53	-193	SYBYL	[14]
2	GlcNAcβ(1→4)GlcNAc	+	-269	-102	-167	SYBYL	[14]
			447			YASARA	[50]
3	GlcNAcβ(1→6)ManNAc	+	-228	-104	-126	SYBYL	[14]
4	GlcNAcβ(1→6)GlcNAc	+	-165	-43	-122	SYBYL	[14]
5	*p*-Nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside	100 %	-299	-84	-216	SYBYL	[47]
			-300	-84	-216	SYBYL	[49]
			-34**			AutoDock	[95]
6	*p*-Nitrophenyl 2-amino-2-deoxy-β-D-glucopyranoside	-	-251	-87	-164	SYBYL	[47]
7	*p*-Nitrophenyl 2-deoxy-2-formamido-β-D-glucopyranoside	5 - 11 %	-227	-71	-156	SYBYL	[47]
8	*p*-Nitrophenyl 2-deoxy-2-trifluoroacetamido-β-D-glucopyranoside	-	-226	-123	-103	SYBYL	[47]
9	*p*-Nitrophenyl 2-deoxy-2-glycoloylamido-β-D-glucopyranoside	5 - 11 %	-274	-111	-163	SYBYL	[47]
10	*p*-Nitrophenyl 2-deoxy-2-propionamido-β-D-glucopyranoside	< 25%	-194	-114	-83	SYBYL	[47]
11	2-Acetamido-2-deoxy-β-D-glucopyranosyl azide:	5 - 10 %					
	a) resonance structure of the azido group is $-N=N^{+}=N^{-}$		-266	-29	-237	SYBYL	[49]
	b) $-N^{-}-N^{+}\equiv N$		-257	-57	-200	SYBYL	[49]
12	2-Acetamido-2-deoxy-β-D-galactopyranosyl azide:	-					
	a) resonance structure of the azido group is $-N=N^{+}=N^{-}$		-167	-72	-95	SYBYL	[49]
	b) $-N^{-}-N^{+}\equiv N$		-164	-62	-102	SYBYL	[49]
13	***p*-Nitrophenyl 2-acetamido-2-deoxy-β-D-galactopyranoside**	56 %	-242	-121	-121	SYBYL	[48]
			-28**			AutoDock	[95]
14	*p*-Nitrophenyl 2-acetamido-2-deoxy-β-D-*galacto*-hexodialdo-1,5-pyranoside	11 – 7 %				SYBYL	[48]
	a) Aldehyde		-180	-89	-91	SYBYL	[48]
	b) Geminal diol		-241	-86	-155	SYBYL	[48]
15	*p*-Nitrophenyl 2-acetamido-2-deoxy-β-D-galactopyranosiduronic acid	-	-136	-94	-42	SYBYL	[48]
16	Methyl (*p*-nitrophenyl 2-acetamido-2-deoxy-β-D-galactopyranoside) Urinate	-	-131	-79	-52	SYBYL	[48]
17	*p*-Nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranosiduronitrile	-	-249	-86	-163	SYBYL	[51]
18	*p*-Nitrophenyl 2-acetamido-2,6-dideoxy-β-D-*gluco*-heptopyranosyluronontrile	-	-238	-96	-142	SYBYL	[51]
19	2-acetamido-2,6-dideoxy-β-D-*gluco*-heptopyranosyluronontrile azide	-	-208	-65	-143	SYBYL	[51]
20	Phenyl 2-acetamido-2,4-dideoxy-β-D-*xylo*-hexopyranoside	8 – 86 %	-31**			AutoDock	[95]
21	Phenyl 2-acetamido-2,4-dideoxy-β-D-*xylo*-1,5-dialdo-hexopyranoside	1 – 15 %	-24**			AutoDock	[95]
22	*p*-Nitrophenyl 2-acetamido-2,4-dideoxy-α-L-*threo*-hex-4-enopyranoside	-	-17**			AutoDock	[95]
23	*p*-Nitrophenyl 2-acetamido-2,4-dideoxy-α-L-*threo*-hex-4-enodialdo-1,5-pyranoside	-	-29**			AutoDock	[95]

* Hydrolysis rate related to the hydrolysis of the standard substrate **5** (%).
** AutoDock produced values correspond to "free energy of binding", which is biologically more relevant than values of "binding energy" calculated by SYBYL and Yasara.
+ Hydrolysed with an undetermined rate.
- Not hydrolysed.
Standard substrates are highlighted in **bold**.

stability. Computational studies of interactions of substrates with different substitution groups at C1, C2, C4, C6, with oxidation at C-4 atom in the active site of the model of β-*N*-acetylhexosaminidase from *Aspergillus oryzae* complete the experimental data with the 'atomic-level' view of the changes into the active site of enzyme. This helps to better understand the role of certain amino acids in the catalytic process.

For the analysis of docking of the natural and modified substrates three substantial computing tools were employed:

YASARA, AutoDock and SYBYL. Molecular dynamics simulations helped to fit the substrates in the correct position and to investigate changes in the position of substrates in the active site as a result of the evaluation of enzyme-substrate interactions in time. Despite of comparatively simple approach for the calculation of binding energies in SYBYL, this parameter could characterize enzyme-substrate binding sufficiently. Generally, the values of SYBYL binding energy compared to standard substrate correspond to the good substrate or inhibitor of β-*N*-acetylhexosaminidase from *Aspergillus oryzae*, as shown in the wet experiments. The results of the recently used AutoDock energy calculation are also in good agreement with the experiments, moreover, the generated free energy of binding is biologically more relevant. In any case, binding energy calculation is not the only satisfactory parameter for the discrimination of good substrates and inhibitors (http://tripos.com/index.php). Molecular dynamics simulation made by YASARA produces information about the pose of the substrate at the active site, which is the other relevant parameter for the estimation of the reaction progress possibilities. However, for the binding energy calculation in the study of the 4-deoxy hexosaminides the use of the AutoDock approach was more appropriate. Together this computational toolbox allowed to foretell and interpret the behaviour of the potential substrates and inhibitors of fungal β-*N*-acetylhexosaminidases in the real experiments.

ACKNOWLEDGEMENTS

We would like to thank our supervisors Prof. Vladimír Křen and Dr. Rudiger Ettrich for their kind help and advice. Research of the authors is supported by the Czech Science Foundation (P207/11/0629; 305/09/H008; 203/08/0114) and the Ministry of Education of the Czech Republic (projects LC06010, MSM6007665808, AV0Z50200510 and AV0Z60870520). Access to METACentrum was provided under the research intent MSM6383917201. Additionally, N.K. was supported by the University of South Bohemia, grant GAJU 170/2010/P.

REFERENCES

[1] Geldenhuys, W. J.; Gaasch, K. E.; Watson, M.; Allen, D. D.; Van der Schyf, C. J. Optimizing the use of open-source software applications in drug discovery. *Drug Discov. Today*, **2006**, *11*(3/4), 127.

[2] Gilbert, D. Bioinformatics software resources. *Brief Bioinform.*, **2004,** *5*(3), 300.

[3] Leach, A. R. *Molecular Modelling: Principles and Applications*. – 2nd ed., Harlow: Pearson, Prentice Hall, **2007**.

[4] Jónsdóttir, S. O.; Jørgensen, F. S.; Brunak, S. Prediction methods and databases within chemoinformatics: emphasis on drugs and drug candidates. *Bioinformatics*, **2005**, *21* (10), 2145.

[5] Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.*, 2004, *3 (11)*, 935.

[6] Röttig, M.; Rausch, C.; Kohlbacher O. Combining structure and sequence information allows automated prediction of substrate specificities within enzyme families. *PLoS Comput. Biol.*, **2010**, *6* (1), e1000636.

[7] Nayeem, A.; Sitkoff, D.; Krystek, S. Jr. A comparative study of available software for high-accuracy homology modeling: From sequence alignments to structural models. *Protein Sci.*, **2006**, *15* (4), 808.

[8] Kryshtafovych, A.; Fidelis, K. Protein structure prediction and model quality assessment. *Drug Discov. Today*, **2009**, *14* (7-8), 386.

[9] Kim, J.H.; Lee, J.; Oh, B.; Kimm, K.; Koh, I. Prediction of phosphorylation sites using SVMs. *Bioinformatics*, **2004**, *20*, 3179.

[10] Plewczynski, D.; Tkacz, A.; Wyrwicz, L. S.; Rychlewski, L. AutoMotif server: prediction of single residue post-translational modifications in proteins. *Bioinformatics*, **2005**, *21* (10), 2525.

[11] Bohne-Lang, A.; von der Lieth C-W. GlyProt: *in silico* glycosylation of proteins. *Nucleic Acids Res.*, **2005**, *33*, W214.

[12] Lybrand, T. P. Ligand—protein docking and rational drug design. *Curr. Opin. Struct. Biol.*, **1995**, *5*, 224.

[13] Hermann, J. C.; Marti-Arbona, R.; Fedorov, A. A.; Fedorov, E.; Almo, S. C.; Shoichet, B. K.; Raushel, F. M. Structure-based activity prediction for an enzyme of unknown function. *Nature*, **2007**, *448*, 775.

[14] Hušáková, L.; Herkommerová-Rajnochová, E.; Semeňuk, T.; Kuzma, M.; Rauvolfová, J.; Přikrylová, V.; Ettrich, R.; Plíhal, O.; Bezouška, K.; Křen V. Enzymatic discrimination of 2-acetamido-2-deoxy-D-mannopyranose-containing disaccharides using β-N-acetylhexosaminidases. *Adv. Synth. Catal.*, **2003**, *345*, 735.

[15] Jedrzejas, M. J.; Stern, R. Structures of vertebrate hyaluronidases and their unique enzymatic mechanism of hydrolysis. *Proteins*, **2005**, *61*, 227.

[16] Pons, T.; Olmea, O.; Chinea, G.; Beldarraín, A.; Márquez, G.; Acosta, N.; Rodríguez, L.; Valencia, A. Structural model for family 32 of glycosyl-hydrolase enzymes. *Proteins*, **1998**, *33*, 383.

[17] Zhou, Z.; Bates, M.; Madura, J. D. Structure modeling, ligand binding, and binding affinity calculation (LR-MM-PBSA) of human heparanase for inhibition and drug design. *Proteins*, **2006**, *65* (3), 580.

[18] Han, W. W.; Li, Z. S.; Zheng, Q. C.; Sun, C. C. Toward a blueprint for β-primeverosidase from the tea leaves structure/function properties: homology modeling study. *J. Theor. Comput. Chem.*, **2006**, *5*, 433.

[19] Papaleo, E.; Fantucci, P.; Vai, M.; De Gioia, L. Three-dimensional structure of the catalytic domain of the yeast β-(1,3)-glucan transferase Gas1: a molecular modeling investigation *J. Mol. Model.*, **2006**, *12* (2), 237.

[20] Schulz, E. C.; Schwarzer, D.; Frank, M.; Stummeyer, K.; Muhlenhoff, M.; Dickmanns, A.; Gerardy-Schahn, R.; Ficner, R. Structural basis for the recognition and cleavage of polysialic acid by the bacteriophage K1F tailspike protein EndoNF. *J. Mol. Biol.*, **2010**, *397* (1), 341.

[21] Cho, S. W.; Lee, S.; Shin, W. The X-ray structure of *Aspergillus aculeatus* polygalacturonase and a modeled structure of the polygalacturonase-octagalacturonate complex. *J. Mol. Biol.*, **2001**, *311* (4), 863.

[22] Wen, X.; Yuan, Y.; Kuntz, D. A.; Rose, D. R.; Pinto, B. M. A combined STD-NMR/molecular modeling protocol for predicting the binding modes of the glycosidase inhibitors kifunensine and salacinol to Golgi α-mannosidase II. *Biochemistry*, **2005**, *44* (18), 6729.

[23] Wang, S. T.; Feng, J. N.; Guo, J. W.; Guo, L. M.; Li, Y.; Sun, Y. X.; Qin, W. S.; Hu, M. R.; Han, G. C.; Shen, B. F. A novel designed single domain antibody on 3-D structure of ricin A chain remarkably blocked ricin-induced cytotoxicity. *Mol. Immunol.*, **2006**, *43* (11), 1912.

[24] Zouhar, J.; Vevoda, J.; Marek, J.; Damborský, J.; Su, X. D.; Brzobohatý, B. Insights into the functional architecture of the catalytic center of a maize β-glucosidase Zm-p60.1 *Plant Physiol.*, **2001**, *127* (3), 973.

[25] Marton, Z.; Tran, V.; Tellier, C.; Dion, M.; Drone, J.; Rabiller, C. Engineering of glucoside acceptors for the regioselective synthesis of β-(1→3)-disaccharides with glycosynthases. *Carbohydr. Res.*, **2008**, *343* (17), 2939.

[26] Kaper, T.; van Heusden, H. H.; van Loo, B.; Vasella, A.; van der Oost, J.; de Vos, W. M. Substrate specificity engineering of β-mannosidase and β-glucosidase from *Pyrococcus* by exchange of unique active site residues. *Biochemistry*, **2002**, *41* (12), 4147.

[27] Cobucci-Ponzano, B.; Moracci, M.; Di Lauro, B.; Ciaramella, B.; D´Avino, R.; Rossi, M. Ionic network at the C-terminus of the β-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: Functional role in the quaternary structure thermal stabilization. *Proteins*, **2002**, *48*, 98.

[28] Liu, X. M.; Qu, Y. B.; You, F.; Liu, Y. tudies on the key amino acid residues responsible for the alkali-tolerance of the xylanase by site-directed or random mutagenesis. *J. Mol. Catal. B – Enzymatic*, **2002**, *18* (4-6), 307.

[29] Tews, I.; Perrakis, A.; Oppenheim, A.; Dauter, Z.; Wilson, K. S.; Vorgias, C. E. Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay–Sachs disease. *Nat. Struct. Biol.*, **1996**, *3*, 638.

[30] Mark, B. L.; Vocadlo, D. J.; Knapp, S.; Triggs-Raine, B. L.; Withers, S. G.; James, M. N. G. Crystallographic evidence for substrate-assisted catalysis in a bacterial β-hexosaminidase. *J. Biol. Chem.*, **2001**, *276*, 10330.

[31] Langley, D. B.; Harty, D. W. S.; Jacques, N. A.; Hunter, N.; Guss, J. M.; Collyer, C. A. Structure of N-acetyl-β-D-glucosaminidase (GcnA) from the endocarditis pathogen *Streptococcus gordonii* and its complex with the mechanism-based inhibitor NAG-thiazoline. *J. Mol. Biol.*, **2008**, *377*, 104.

[32] Ramasubbu, N.; Thomas, L. M.; Ragunath, C.; Kaplan, J. B. Structural analysis of Dispersin B, a biofilm-releasing glycoside hydrolase from the periodontopathogen *Actinobacillus actinomycetemcomitans*. *J. Mol. Biol.*, **2005**, *349*, 475.

[33] Sumida, T.; Ishii, R.; Yanagisawa, T.; Yokoyama, S.; Ito, M. Molecular cloning and crystal structural analysis of a novel β-N-acetylhexosaminidase from *Paenibacillus* sp. TS12 capable of degrading glycosphingolipids. *J. Mol. Biol.*, **2009**, *392*, 87.

[34] Mark, B. L.; Mahuran, D. J.; Cherney, M. M.; Zhao, D.; Knapp, S.; James, M. N. G. Crystal structure of human β-hexosaminidase B: Understanding the molecular basis of Sandhoff and Tay–Sachs disease. *J. Mol. Biol.*, **2003**, *327*, 1093.

[35] Maier, T.; Strater, N.; Schuette, C. G.; Klingenstein, R.; Sandhoff, K.; Saenger, W. The X-ray crystal structure of human β-hexosaminidase B provides new insights into Sandhoff disease. *J. Mol. Biol.*, **2003**; *328*, 669.

[36] Lemieux, M. J.; Mark, B. L.; Cherney, M. M.; Withers, S. G.; Mahuran, D. J.; James, M. N. G. Crystallographic structure of human β-hexosaminidase A: Interpretation of Tay-Sachs mutations and loss of GM2 ganglioside hydrolysis *J. Mol. Biol.*, **2006**, *359*, 913.

[37] Tay, W. Symmetrical changes in the region of the yellow spot in each eye of an infant. *Trans. Ophthalmol. Soc. UK*, **1881**, *1*, 155.

[38] Sandhoff, K.; Andreae, U.; Jatzkewitz, H. Deficient hexozaminidase activity in an exceptional case of Tay-Sachs disease with additional storage of kidney globoside in visceral organs. *Life Sci*., **1968**, *7 (6)*, 283.

[39] Fernandes, M. J. G.; Yew, S.; Leclerc, D.; Henrissat, B.; Vorgias, C. E.; Gravel, R. A.; Hechtman, P.; Kaplan, F. Identification of candidate active site residues in lysosomal β-hexosaminidase A. *J. Biol. Chem*., **1997**, *272*, 814.

[40] Mark, B. L.; Wasney, G. A.; Salo, T. J.; Khan, A. R.; Cao, Z.; Robbins, P. W.; James, M. N. G.; Triggs-Raine, B. L. Structural and functional characterization of *Streptomyces plicatus* β-N-acetylhexosaminidase by comparative molecular modeling and site-directed mutagenesis. *J. Biol. Chem*., **1998**, *273*, 19618.

[41] Sakuraba, H.; Matsuzawa, F.; Aikawa, S.; Doi, H.; Kotani, M.; Lin, H.; Ohno, K.; Tanaka, A.; Yamada, H.; Uyama, E. Corrective effect on Fabry mice of yeast recombinant human α-galactosidase with N-linked sugar chains suitable for lysosomal delivery. *J. Hum. Genet.*, **2002**, *47*, 176.

[42] Matsuzawa, F.; Aikawa, S.; Sakuraba, H.; Lan, H. T. N.; Tanaka, A.; Ohno, K.; Sugimoto, Y.; Ninomyia, H.; Doi, H. Structural basis of the GM2 gangliosidosis B variant. *J. Hum. Genet*., **2003**, *48*, 582.

[43] Kaplan, F.; Boulay, B.; Cordeiro, P.; Hechtman P. Identification of two amino acid residues which determine the substrate specificity of human β-D-N-acetylhexosaminidase A. *Gene Funct. Dis*., **2001**, *1*, 38.

[44] Kerrigan, J. E.; Ragunath, C.; Kandra, L.; Gyemant, G.; Liptak, A.; Janossy, L.; Kaplan, J. B.; Ramasubbu, N. Modeling and biochemical analysis of the activity of antibiofilm agent Dispersin B. *Acta Biol. Hung*., **2008**, *59* (4), 439.

[45] Greig, I. R.; Zahariev, F.; Withers, S. G. Elucidating the nature of the *Streptomyces plicatus* β-hexosaminidase-bound intermediate using *ab initio* molecular dynamics simulations. *J. Am. Chem. Soc*., **2008**, *130*, 17621.

[46] Cantarel, B. L.; Coutinho, P. M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res*., **2009**, D233.

[47] Fialová, P.; Weignerová, L.; Rauvolfová, J,; Přikrylová, V.; Pišvejcová, A.; Ettrich, R.; Kuzma, M.; Sedmera, P.; Křen, V. Hydrolytic and transglycosylation reactions of N-acyl modified substrates catalysed by β-N-acetylhexosaminidases. *Tetrahedron*, **2004**, *60*, 693.

[48] Fialová, P.; Namdjou D.-J.; Ettrich, R.; Přikrylová, V.; Rauvolfová, J.; Křenek, K.; Kuzma, M.; Elling, L.; Bezouška, K.; Křen, V. Combined application of galactose oxidase and β-N-acetylhexosaminidase in the synthesis of complex immunoactive N-acetyl-D-galactosaminides. *Adv. Synth. Catal*., **2005**, *347*, 997.

[49] Fialová, P.; Carmona, A. T.; Robina I.; Ettrich, R.; Sedmera, P.; Přikrylová, V.; Petrásková-Hušáková, L.; Křen, V. Glycosyl azide—a novel substrate for enzymatic transglycosylations. *Tetrahedron Lett*., **2005**, *46*, 8715.

[50] Ettrich, R.; Kopecký, V. Jr.; Hofbauerová, K.; Baumruk, V.; Novák, P.; Pompach, P.; Man, P.; Plíhal, O.; Kutý, M.; Kulik, N.; Sklenář, J.; Ryšlavá, H.; Křen, V.; Bezouška, K. Structure of the dimeric N-glycosylated form of fungal β-N-acetylhexosaminidase revealed by computer modeling, vibrational spectroscopy, and biochemical studies. *BMC Struct. Biol*., **2007**, *7*, 32.

[51] Carmona, A. T.; Fialová, P.; Křen, V.; Ettrich, R.; Martínková, L.; Moreno-Vargas, A. J.; Gonzales, C.; Robina, I. Cyanodeoxy-glycosyl derivatives as substrates for enzymatic reactions. *Eur. J. Org. Chem*., **2006**, 1876.

[52] Tews, I.; Terwisscha van Scheltinga, A. C.; Perrakis, A.; Wilson, K. S.; Dijkstra, B. W. Substrate-assisted catalysis unifies two families of chitinolytic enzymes. *J. Am. Chem. Soc*., **1997**, *119*, 7954.

[53] Mark, B. L.; James, M. N. G. Anchimeric assistance in hexosaminidases. *Can. J. Chem*., **2002**, *80*, 1064.

[54] Vocadlo, D. J.; Withers, S. G. Detailed Comparative analysis of the catalytic mechanisms of β-N-acetylglucosaminidases from families 3 and 20 of glycoside hydrolases. *Biochemistry*, **2005**, *44*, 12809.

[55] Knapp, S.; Vocadlo, D.; Gao, Z.; Kirk, B.; Lou, J.; Withers, S. G. NAG-thiazoline, an N-acetyl-β-hexosaminidase inhibitor that implicates acetamido participation *J. Am. Chem. Soc*., **1996**, *118*, 6804.

[56] Prag, G.; Papanikolau, Y.; Tavlas, G.; Vorgias, C. E.; Petratos, K.; Oppenheim, A. B. Structures of chitobiase mutants complexed with the substrate di-N-acetyl-image-glucosamine: the catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540. *J. Mol. Biol*., **2000**, *300*, 611.

[57] Williams, S. J.; Mark, B. L.; Vocadlo, D. J.; James, M. N. G.; Withers, S. G. Aspartate 313 in the *Streptomyces plicatus* hexosaminidase plays a critical role in substrate-assisted catalysis by orienting the 2-acetamido group and stabilizing the transition state. *J. Biol. Chem*., **2002**, *277*, 40055.

[58] Hou, Y.; Vocadlo, D.; Withers, S.; Mahuran, D. Role of βArg211 in the active site of human β-hexosaminidase B. *Biochemistry*, **2000**, *39*, 6219.

[59] Tanaka, A.; Ohno, K.; Sandhoff, K.; Maire, I.; Kolodny, E. H.; Brown, A.; Suzuki, K. GM2-gangliosidosis B1 variant: analysis of β-hexosaminidase alpha gene abnormalities in seven patients. *Am. J. Hum. Genet*., **1990**, *46*, 329.

[60] Brown, C. A.; Neote, K.; Leung, A.; Gravel, R. A.; Mahuran, D. J. Introduction of the alpha subunit mutation associated with the B1 variant of Tay-Sachs disease into the beta subunit produces a beta-hexosaminidase B without catalytic activity. *J. Biol. Chem*., **1989**, *264*, 21705.

[61] Laederach, A.; Reilly, P. J. Modeling protein recognition of carbohydrates. *Proteins*, **2005**, *60* (4), 591.

[62] Lee, J.; Wu, S.; Zhang, Y. *From Protein Structure to Function with Bioinformatics*, Chapter 1; Springer: London, **2009**.

[63] Zhang, Y. Progress and challenges in protein structure prediction. *Curr. Opin. Struct. Biol*., **2008**, *18 (3)*, 342.

[64] Schueler-Furman, O.; Wang, C.; Bradley, P.; Misura, K.; Baker, D. Progress in Modeling of Protein Structures and Interactions. *Science*, **2005**, *310* (5748), 638.

[65] Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol*., **1990**, *215 (3)*, 403.

[66] Lipman, D. J.; Pearson, W. R. Rapid and sensitive protein similarity searches. *Science*, **1985**, *227*, 1435.

[67] Söding, J.; Biegert, A.; Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* (web server issue), **2005**, W244.

[68] Notredame, C.; Higgins, D. G.; Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol*., **2000**, *302 (1)*, 205.

[69] Xiang, Z. Advances in homology protein structure modeling. *Curr. Protein Pept. Sci*., **2006**, *7 (3)*, 217.

[70] Higgins, D. G.; Thompson, J. D.; Gibson, T. J. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol*., **1996**, *266*, 383.

[71] Edgar, R.C.; Batzoglou, S. Multiple sequence alignment. *Curr. Opin. Struct. Biol*., **2006**, *16*, 368.

[72] Dunbrack, R.L. Jr. Sequence comparison and protein structure prediction. *Curr. Opin. Struct. Biol.*, **2006**, *16* (3), 374.

[73] Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*., **1994**, *22 (22)*, 4673.

[74] Fiser, A.; Do, R. K.; Sali, A. Modeling of loops in protein structures. *Protein Sci*., **2000**, *9* (9), 1753.

[75] Rost, B. Twilight zone of protein sequence alignments. *Protein Eng.*, **1999**, *12* (2), 85.

[76] Ginalski, K. Comparative modeling for protein structure prediction. *Curr. Opin. Struct. Biol.*, **2006**, *16* (2), 172.

[77] Venclovas, C.; Margelevičius, M. Comparative modeling in CASP6 using consensus approach to template selection, sequence-structure alignment, and structure assessment. *Proteins*, **2005**, *61* (S7), 99.

[78] Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol*., **1993**, *234*, 779.

[79] Pieper, U.; Naraynan, E.; Webb, B. M.; Eramian, D.; Kelly, L.; Barkan, D. T.; Carter, H.; Mankoo P.; Karchin, R.; Marti-Renom, M. A.; Davis, F. P.; Sali, A. MODBASE, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res*., **2009**, *37*, D347.

[80] Laskowski, R. A.; McArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog*., **1993**, *26*, 283.

[81] Sippl, M. J. Recognition of errors in three-dimensional structures of proteins. *Proteins*, **1993**, *17 (4)*, 355.

[82] Wallner, B.; Elofsson, A. Can correct protein models be identified? *Protein Sci.,* **2003**, *12*, 1073.

[83] Young, D. *Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems:* John Wiley & Sons, USA **2001**.

[84] Joseph, D.; Petsko, G. A.; Karplus, M. Anatomy of a conformational change: hinged "lid" motion of the triosephosphate isomerase loop. *Science*, **1990**, *249*, 1425.

[85] Pakhomova, S.; Kobayashi, M.; Buck, J.; Newcomer, M. E. A helical lid converts a sulfotransferase to a dehydratase. *Nat. Struct. Biol*., **2001**, *8*, 447.

[86] Bustos-Jaimes, I.; Sosa-Peinado, A.; Rudino-Pinera, E.; Horjales, E.; Calcagno, M. L. On the Role of the Conformational flexibility of the active-site lid on the allosteric kinetics of glucosamine-6-phosphate deaminase. *J. Mol. Biol*., **2002**, *319*, 183.

[87] Brocca, S.; Secundo, F.; Ossola, M.; Alberghina, L.; Carrera, G.; Lotti, M. Sequence of the lid affects activity and specificity of Candida rugosa lipase isoenzymes. *Protein Sci*., **2003**, *12*, 2312.

[88] Plíhal, O.; Sklenář, J.; Kmoníčková, J.; Man, P.; Pompach, P.; Havlíček, V.; Křen, V.; Bezouška, K. N-glycosylated catalytic unit meets O-glycosylated propeptide: complex protein architecture in a fungal hexosaminidase. *Biochem. Soc. Trans*., **2004**, *32*, 764.

[89] Novikov, F. N.; Chilov, G. G. Molecular docking: theoretical background, practical applications and perspectives. *Mendeleev Commun.*, **2009**, *19*, 237.

[90] Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Protein–ligand docking: Current status and future challenges. *Proteins*, **2006**, *65*(1), 15.

[91] Agostino, M.; Jene, C.; Boyle, T.; Ramsland, P.A.; Yuriev, E. Molecular docking of carbohydrate ligands to antibodies: Structural validation against crystal structures. *J. Chem. Inf. Model*., **2009**, *49*(12), 2749.

[92] Nurisso, A.; Kozmon, S.; Imberty, A. Comparison of docking methods for carbohydrate binding in calcium dependent l ectins and prediction of the carbohydrate binding mode to sea cucumber lectin CEL-III. *Mol. Simulat.*, **2008**, *34* (4), 469.

[93] Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.*, **1998**, *19* (14), 1639.

[94] Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.*, **2006**, *26* (5), 531.
[95] Slámová, K.; Gažák, R.; Bojarová, P.; Kulik, N.; Ettrich, R.; Pelantová, H.; Sedmera, P.; Křen, V. 4-Deoxy-substrates for β-N-acetylhexosaminidases: How to make use of their loose specificity. *Glycobiology*, **2010**, *20*, 1002.
[96] Schoevaart, R.; Kieboom, T. Galactose dialdehyde: the forgotten candidate for a protein cross-linker? *Carbohydr. Res*., **2001**, *334*, 1.
[97] Bonnet, V.; Duval, R.; Rabiller, C. Oxidation of galactose and derivatives catalysed by galactose oxidase: structure and complete assignments of the NMR spectra of the main product. *J. Mol. Catal. B: Enzym.*, **2003**, *24-25*, 9.
[98] Davies, G. J.; Ducros, V. M. A.; Varrot, A.; Zechel, D. L. Mapping the conformational itinerary of β-glycosidases by X-ray crystallography. *Biochem. Soc. Trans*., **2003**, *31*, 523.

Received: April 07, 2010 Revised: June 22, 2010 Accepted: July 02, 2010