Article

Molecular Modeling of Myrosinase from *Brassica oleracea*: A Structural Investigation of Sinigrin Interaction

Sathishkumar Natarajan [1], Senthil Kumar Thamilarasan [1], Jong-In Park [1], Mi-Young Chung [2] and Ill-Sup Nou [1,*]

[1] Department of Horticulture, Sunchon National University, Suncheon 57922, Korea;
    E-Mails: sathish@sunchon.ac.kr (S.N.); senkuttybio@gmail.com (S.K.T.);
    jipark@sunchon.ac.kr (J.-I.P.)

[2] Department of Agricultural Education, Sunchon National University, Suncheon 540-950, Korea;
    E-Mail: queen@sunchon.ac.kr

* Author to whom correspondence should be addressed; E-Mail: nis@sunchon.ac.kr;
    Tel.:+82-61-750-3249; Fax: +82-61-750-5389.

Abstract: Myrosinase, which is present in cruciferous plant species, plays an important role in the hydrolysis of glycosides such as glucosinolates and is involved in plant defense. *Brassicaceae* myrosinases are diverse although they share common ancestry, and structural knowledge about myrosinases from cabbage (*Brassica oleracea*) was needed. To address this, we constructed a three-dimensional model structure of myrosinase based on *Sinapis alba* structures using Iterative Threading ASSEmbly Refinement server (I-TASSER) webserver, and refined model coordinates were evaluated with ProQ and Verify3D. The resulting model was predicted with β/α fold, ten conserved N-glycosylation sites, and three disulfide bridges. In addition, this model shared features with the known *Sinapis alba* myrosinase structure. To obtain a better understanding of myrosinase–sinigrin interaction, the refined model was docked using Autodock Vina with crucial key amino acids. The key nucleophile residues GLN207 and GLU427 were found to interact with sinigrin to form a hydrogen bond. Further, 20-ns molecular dynamics simulation was performed to examine myrosinase–sinigrin complex stability, revealing that residue GLU207 maintained its hydrogen bond stability throughout the entire simulation and structural orientation was similar to that of the docked state. This conceptual model should be useful for understanding the structural features of myrosinase and their binding orientation with sinigrin.

1. Introduction

Myrosinase (also termed thioglucoside glucohydrolase, sinigrinase, or sinigrase EC: 3.2.1.147), is an S-glycosidase involved in breakdown of thioglucosides such as glucosinolates (GSL), a large group of secondary metabolites especially abundant in cruciferous plants. Myrosinase is localized in specialized GSL-free idioblast cells termed myrosin cells [1]. GSLs, classified based on their structure as indolic or aliphatic, are biologically inactive in their native state; myrosinase catalyzes the hydrolysis of these thioglucosides into D-glucose and a thiol (sulfate) such as aglycone (thiohydroximate-*O*-sulfonate). The thiols produced are unstable intermediates that differ mainly based on side chain structures, myrosinase substrate, plant species, and other reaction conditions like pH or availability of ferrous ions and environmental factors [2–4]. Compounds spontaneously formed from aglycone in this process are biologically intact, active and sometimes toxic, and include isothiocyanates and nitrile, epithionitriles, thiocyanates, and oxazoldine-thione [5], which are involved in plant defense against herbivory, pathogen invasion and tissue damage [1,3]. Concentrations of GSLs vary among different parts of the plant [6,7]. External factors such as tissue damage and pathogen attack trigger the release of GSLs, which results in interaction with myrosinase and eventually degradation of the glycosides. In plants, myrosinase activities vary between species, cultivars, and parts of the plant examined. Myrosinase activity is highest in seeds and seedling stages as indolic GSLs are broken down to form indoleacetic acid (IAA), a phytohormone essential for developing plants [8]. In cruciferous species, the interaction between myrosinase and GSLs for defense creates a so-called binary mustard oil bomb. These myrosinase are activated by ascorbic acid and the three-dimensional structures of several binary and ternary complexes of the enzyme with different transition state analogues, with ascorbic acid and with inhibitors mimicking the substrate have been reported. In addition, mechanistic details already known include the position of a water molecule placed perfectly for activation by ascorbate and for nucleophilic attack on the covalent intermediate of the reaction [9]. Biochemical, genetic and phylogenetic data, [10] reported that the myrosinase-GSL system evolved from cyanogenic glucosides and *O*-β-glucosidases. Products of GSLs possess toxic and pharmacological effects on a wide range of organisms [11,12]. GSLs induce Phase I and Phase II detoxifying enzymes during cancer infection and alter steroid hormone levels during their metabolism, in turn aiding in protecting the cell from oxidative damage [13]. Cruciferous species are known for the presence of GSLs, conferring their chemo-preventive effects [14]. Broccoli (*Brassica oleracea var. italica*) was reported to act against and provide protection from liver cancer in animals [15]. In humans, glycosides are involved in reducing DNA damage in lymphocytes [16,17]. There was significant anti-nutritional activity in animals when they were fed with GSLs [14]. GSLs from plants have become an interesting area of work due to their benefits to animal and human health such as anti-microbial [18], anti-nematode [19] insecticidal, weed-inhibiting [20], herbicidal [10,21], allelopathic [22], and anticarcinogenic [15,23] activities. To date, more than 130 GSLs have been identified and reported in plants, of which more than 30 are found in the *Brassicaceae* family [24]. Lenman *et al* [25] reported that myrosinase shows differential expression throughout development and also in various parts of the plant.

Molecular breeders have used traditional breeding methods to increase the content of GSLs due to their beneficial activities. This study aims to improve understanding of myrosinase–sinigrin interaction at the molecular level. To date, one conceptual model was constructed from *Brassica juncia* and described their myrosinase structural features [26]. Homology modeling is a powerful method to predict three-dimensional structures from amino acid sequences [27]. This approach can be used to obtain structural knowledge of proteins, enzyme-substrate interactions, and protein-ligand interactions [28,29]. In this study, we have applied this technique to predict the structure of myrosinase from *Brassica oleracea* (*B. oleracea*, wild cabbage) and examine its interaction with sinigrin, one of the possible substrates.

2. Materials and Methods

2.1. Template Identification and Homology Modeling

The amino acid sequence of myrosinase (MYR) from cabbage (*B. oleracea*, wild cabbage) was obtained from the NCBI protein sequence database (Accession number: ABS30827) [30]. The primary sequence of MYR protein from cabbage (*B. oleracea*) consists of 546 amino acids, which was used for modeling. Homology modeling was performed using the Iterative Threading ASSEmbly Refinement (I-TASSER) server based on the *ab initio/*threading method [31]. To predict the three-dimensional structure of cabbage (*B. oleracea*) MYR, 1MYR_A (*Sinapis alba* (*S. alba*) MYR) [32], 1E4M_M (*S. alba* MYR) [9], 1DWF_M (*S. alba* MYR) [33], 1VO2_A (*Sorghum bicolor* Dhurrinase) [34], and 4JHO_A (*Oryza sativa* beta-glucosidase) [35] structures were chosen as the templates. These template structures were selected by multiple threading approach Local Meta-Threading-Server (LOMETS) based on sequence similarity to cabbage (*B. oleracea*) MYR. Alignments of the query sequence with each template were generated using the ConSurf server. This program provides multiple sequence alignment with conserved features, which provides information about specific amino acid features in their local environment [36]. A total of five three-dimensional (3D) models were generated by I-TASSER; among them, the best model was identified based on confidence score (C-Score) [37].

2.2. Molecular Dynamics Simulation

Simulations of molecular dynamics (MD) were performed for the modeled MYR protein using the GROMOS96 43a1 force field implemented in GROningen MAchine for Chemical Simulation (GROMACS) version 4.6 [38,39]. A cubic box was constructed around the 3D model with a minimum of 10 Å between the box edges and the model surface. Appropriate counter-ions (Na$^+$ or Cl$^-$) were added for neutralizing this system. Further it was solvated with simple point charge (SPC) water molecules and minimized using the steepest descent method to remove spurious contacts. In addition, this system was equilibrated 100 ps under NVT (canonical ensemble) and NPT (isothermal-isobaric ensemble) conditions. The temperature of the system was set to 300 K after applying position restrains to the protein. This temperature and pressure was maintained using the Berendsen weak-coupling method. A cut-off radius of 0.9 nm was used in the simulation for van der Waals and electrostatic interactions [40]. The LINCS algorithm was used to reduce close contact between the complex system and bond constraints [41]. Finally, a 20-ns production was carried out for the MYR model using the particle mesh Ewald (PME) electrostatics method under NPT conditions. For analysis, final coordinates were saved every 2 ps.

The final MYR trajectory coordinates of the 20-ns simulation were selected, subjected to energy minimization. Further, the refined model was validated using Protein Quality Predictor (ProQ) [42] and VERIFY3D plot [43] and then used for molecular docking studies.

2.3. Analysis of MYR Active Site and N-Glycosylation Sites

The refined model was analyzed using Catalytic Site Atlas (CSA) database [16] and Q-SiteFinder [44] for active site identification. Those sites were compared with template structures for structural evaluation. Further *N*-glycosylation sites were identified using the NetNGlyc1.0 server [45].

2.4. MYR–Sinigrin Interaction

The two-dimensional (2D) structure of sinigrin (2-propenylglucosinlate) was retrieved from the NCBI PubChem database with the following accession number CID: 5464493. This molecule was drawn using ChemSketch (Advanced Chemistry Development, Inc. Toronto, ON, Canada), saved in a 2D (.mol) format, and then converted into a 3D (.pdb) format by importing into Discovery Studio 3.5 visualizer (DS 3.5) (DS, Accelrys, Inc. SanDiego, CA, USA). This molecule was optimized using the Conjugate Gradients method [46] followed by Steepest Descent in 200 steps using the PyRx program [47]. Universal Force Field (UFF) was used for minimization of the sinigrin molecule [48]. Molecular docking analysis was performed using AutoDock Vina [49]. Active site residues were carefully assigned to the binding pocket of the MYR protein, and docking steps were followed from previous studies [50]. The best docking orientation was identified from binding affinity score and hydrogen bond interaction to the active site based on visual inspection. The 2D graphical visualization of MYR and sinigrin interaction was generated by DS 3.5. The X-Score program [51] was used to calculate binding affinity scores of sinigrin and non-bonded interactions.

2.5. Complex Stability Refinement by Molecular Dynamics Simulation

The MYR–sinigrin docked-complex file obtained from docking studies was subjected to MD using GROMACS 4.6. The sinigrin structural topology file was generated using the Dundee PRODRG2 server [52] with chirality, full charges and energy minimization. In addition, MYR topology was created using the GROMACS pdb2gmx utility. The above-mentioned dynamics parameter was used to produce a 20-ns simulation for the modeled MYR and MYR–sinigrin complex. The detailed MD procedure was implemented according to our previous study [51]. In this analysis, root mean square deviation (RMSD), and root mean square fluctuation (RMSF) calculations were computed using the GROMACS utilities of g_rms, and g_rmsf, respectively. For this study, Chimera [53], and DS 3.5 were used for graphical visualization, and graphs were plotted using Microsoft Excel.

3. Results and Discussion

3.1. Homology Modeling

The full-length cabbage (*B. oleracea*) MYR protein consists of 546 amino acids including a glycoside hydrolase family 1 (Glyco_hydro_1) domain (41–522 amino acids). A search for homologous proteins

for modeling cabbage (*B. oleracea*) MYR prompted us to use 1MYR_A, 1E4M_M, 1DWF_M, 1VO2_A and 4JHO_A protein structures as templates. Multiple sequence alignment was performed with these template proteins to identify sequence alignments of structural motifs (Figure 1). The alignment results revealed good conservation of cabbage (*B. oleracea*) MYR with the *S. alba* proteins, with 91% query coverage and 72% sequence identity. Cabbage (*B. oleracea*) MYR shared 40% identity with the protein from *Sorghum bicolor* (1VO2_A) and 39% identity with that of *Oryza sativa* (4JHO_A). In addition, three potential *S. alba* templates were found to share their common ancestor with cabbage (*B. oleracea*) MYR. The best model was selected based on C-Score value, and this score was used for estimating the quality of predicted models by I-TASSER. Then, the selected best model was subjected to MD Simulations.
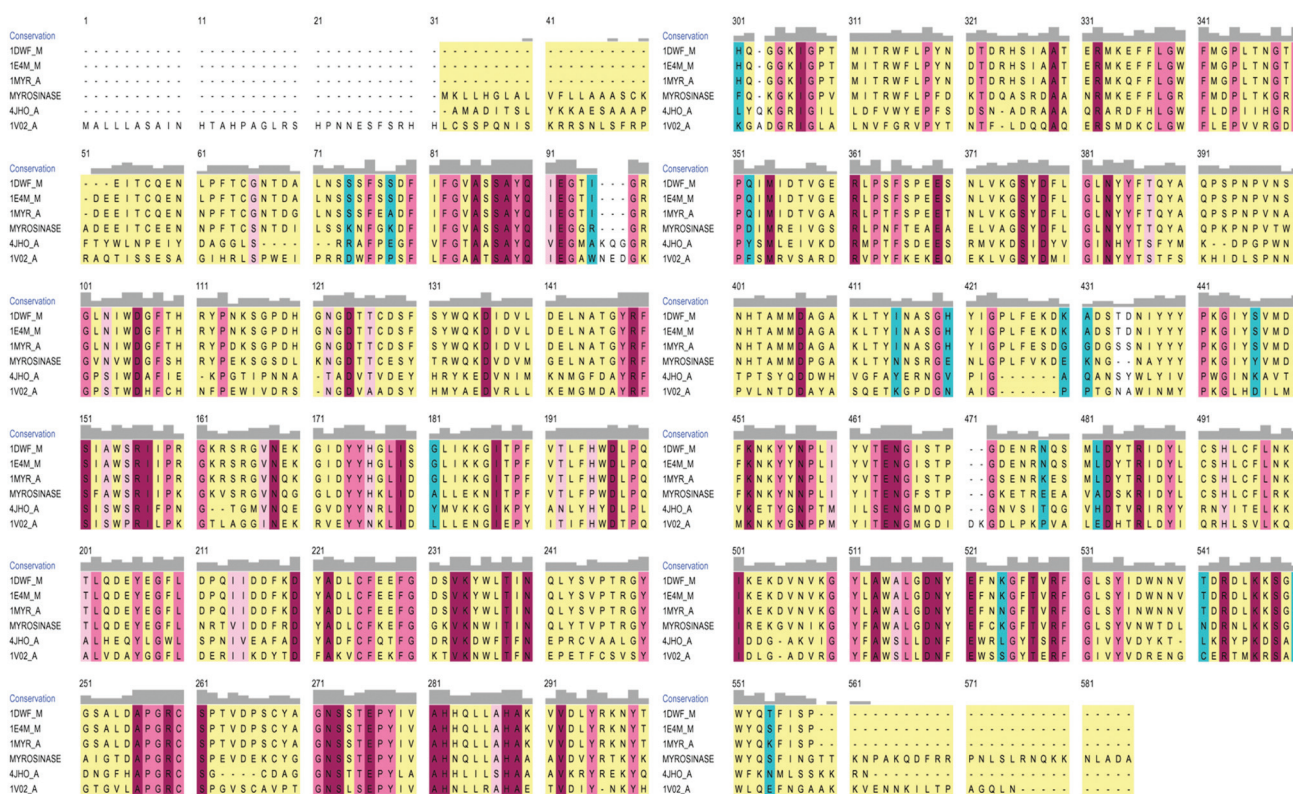

Figure 1. Sequence alignment of myrosinase with their template structures generated by using Chimera visualization tool. Gaps and conservation residues were denoted in dotted lines and gray color, respectively. In addition, block levels represented as identical (tallest blocks), conserved (Intermediate blocks) and mild conserved (Small blocks) and not identical (no blocks).

3.2. Model Refinement and Structure of MYR

The stability and MD properties of the constructed model were evaluated using MD simulations. To identify the energetically favorable structure for this model, MD simulation was observed up to 20 ns. The RMSD values were examined throughout the entire simulation time to check the structural stability of the tested model. RMSD values of MYR backbone atoms attained equilibration state in approximately 7.1 ns with an average of 0.43Å and remained stable throughout the course of the

simulation (Figure 2A). RMSF analyses of backbone atoms of each amino acid are shown in Figure 2B; much structural variation were found in secondary structure. Numerous reports have confirmed that O-glycosidase family enzymes are linked to a single $(\beta/\alpha)_8$ barrel [54]. A final snapshot from the 20-ns MD simulation was subjected to further evaluation and refinement. The stereo-chemical quality of the enhanced model was validated by Veryfy3D and ProQ. The Verify3D evaluation server reported that our optimized model scored over 80% of amino acids in the 3D/1D profile, which indicates residues are in favorable positions (Figure 3). ProQ revealed an LGscore of 4.808, which indicates that our conceptual MYR structure was an extremely good model. From these evaluation results, we confirmed that the developed model was satisfactory, and we therefore used it for subsequent analyses.
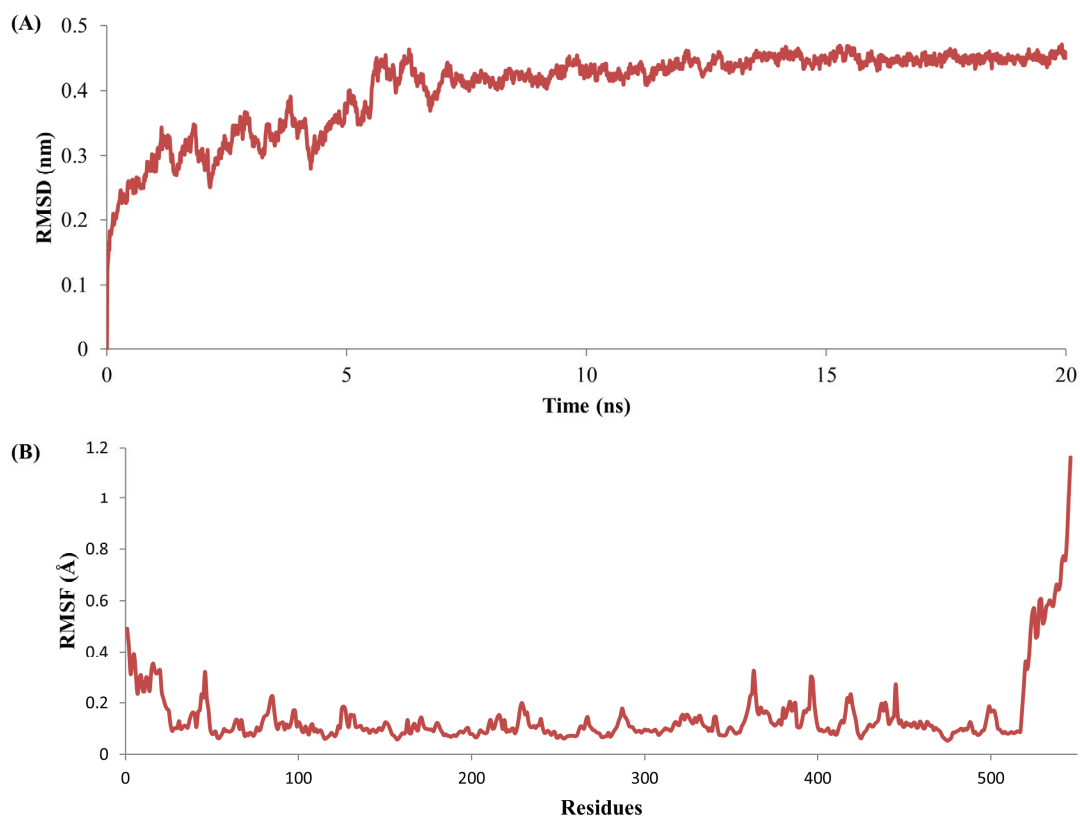

Figure 2. Molecular dynamics based myrosinase model refinement through 20 ns simulation: (**A**) The root mean square deviation (RMSD) values of backbone and (**B**) root mean square fluctuations (RMSF) values of atoms.

Furthermore, this refined MYR protein was compared with template structures to identify sequence conservation. The alignment results revealed that our modeled Cabbage (*B. oleracea*) MYR shared with sequence conservation of 72.14%, 71.54%, 71.54%, 40.29%, and 39.13% with 1MYR_A, 1DWF_M, 1E4M_M, 1VO2_A, and 4JHO_A structures, respectively. In addition, the three templates from *S.alba* (1MYR_A, 1DWF_M, and 1E4M_M) were shown high similarity with refined MYR including Glyco_hydro_1 domain. The GOR4 analysis has revealed that our cabbage (*B. oleracea*) MYR model contains secondary structural elements of 22.71% helix, 20.51% strand, and 56.78% coil. Three disulfide bridges were found, in the Cys2–Cys456, Cys34–Cys454, and Cys226–Cys234 regions. The known myrosinase crystal structure from *S. alba* has shown similar structural features [26]. Diverse myrosinase proteins are involved in plant defense mechanisms, and although their origins vary, they are

associated with common ancestors. Moreover, *N*-glycosylation sites are important to maintain the molecular stability and solubility of myrosinase in the dehydrated environment. *N*-glycosylation sites were predicted for the optimized MYR model using the NetNGlyc1.0 server. A total of ten Asn-Xaa-Ser/Thr sites were observed, at residues 110, 152, 177, 238, 329, 366, 380, 497, 518, and 533. Among these potential *N*-glycosylation sites, those with a threshold value over 0.5 were considered as potential sites and the residues were 110, 177, 329, 497, and 533. In comparison, ten sites were predicted in *S. alba* (six potential) and seven were reported from a theoretical model of *Brassica juncea* (*B. juncea*) [26]. The number of predicted *N*-glycosylation sites from cabbage (*B. oleracea*) was close to *S. alba*, in accordance with good amino acid conservations including potential sites (Supplementary Figure S1).
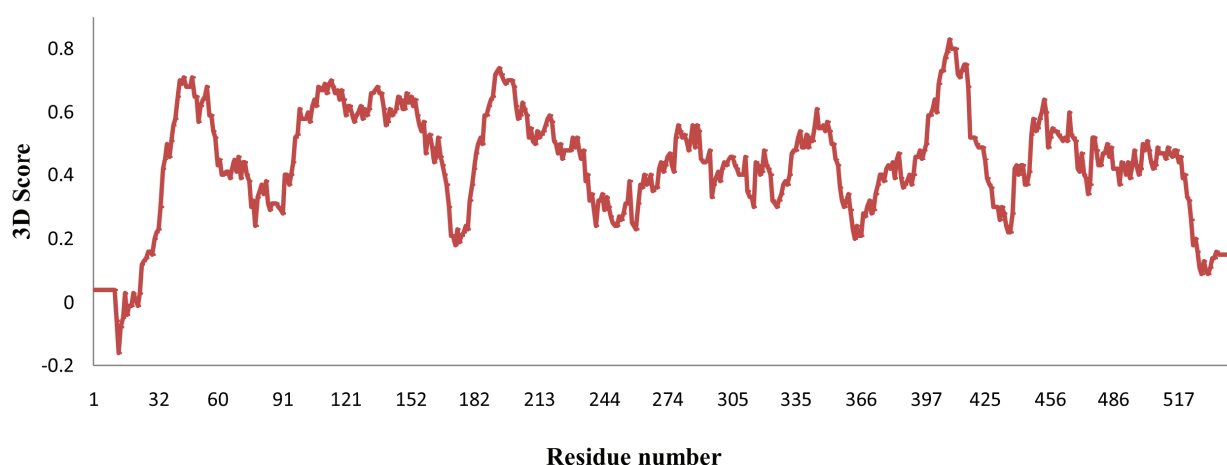

Figure 3. The 3D scores for the modeled myrosinase structure from *Brassica oleracea*.

3.3. Molecular Interaction Studies

To better understand MYR from cabbage (*B. oleracea*), an interaction study between the substrate sinigrin and the enhanced MYR protein active sites was carried out using the Autodock Vina program. For docking studies, potential active site information was examined from the *S. alba* myrosinase crystal structure and confirmed with the CSA database and Q-site finder. Residue ARG115 was involved in the catalytic target site, and GLN207 and GLU427 had catalytic nucleophile activities. In addition, THR210, ASN348, and TYR350 were present in the binding pocket of MYR for substrate binding (Figure 4). These key amino acids were assigned as active sites in docking to identify productive binding orientations. Three-dimensional docked representations of the sinigrin molecule in the MYR active site and interacted 2D plot are shown in Figure 5. Further, myrosinase–sinigrin interaction was evaluated by binding affinity scores and hydrogen bond formation in the MYR active sites. In the docked state, the substrate sinigrin formed hydrogen bonds with the catalytic target site and two catalytic nucleophile residues, along with other residues of SER117, ASP221, and LYS285 (Table 1). The docking program predicted that MYR–sinigrin binding affinity was −7.2 kcal/mol in the docked state. The nucleophile residue GLN207 was involved in the formation of one hydrogen bond to the HE21-O sinigrin group with hydrogen bond length of 2.10Å. A second nucleophilic residue, GLU427, formed two separate hydrogen bonds with the hydrogen atoms present in the sinigrin (H-OE1, and H-OE2) with hydrogen bond length of 1.85Å and 2.34Å respectively. In addition, one hydrogen bond was observed at the residue SER117 HG atom (HG-O) with 1.7Å hydrogen bond length. We also observed two hydrogen bonds with LYS485

(HZ2-O and HZ2-O) with lengths of 2.43Å and 2.41Å, respectively. Moreover, sinigrin contains seven rotatable bonds for enzyme binding, including one nitrogen (N) atom and sulfur group. This specific sulfur group was used to recognize the MYR binding and complex stability. In the docked state, residues GLU427 and LYS485 interacted with N atoms and the sulfur group respectively. In addition, the following hydrophobic interaction residues also contributed to molecular stability: TRP475, THR426, ASN206, PHE483, GLU482, SER56, ALA57, PRO161, TRP71, TYR58, and PRO223.

Table 1. Hydrogen bond interaction results for *B. oleracea* MYR–sinigrin.

S. No	Hydrogen Bond Interacting Residue	Hydrogen Bond Donor	Hydrogen Bond Acceptor	Hydrogen Bond Length (Å)	Number of Hydrogen Bonds
1	**ARG115 ***	ARG115:HH22	LIG1:O	2.36	1
2	SER117	SER117:HG	LIG1:O	1.7	1
3	**GLN207 ***	GLN207:HE21	LIG:O	2.10	1
4	T221	LIG1:HN	ASP221:OD1	2.27	1
5	**GLU427 ***	LIG1:H	GLU427:OE1	1.85	2
		LIG1:H	GLU427:OE2	2.34	-
6	LYS485	LYS485:HZ2	LIG1:O	2.43	2
		LYS485:HZ2	LIG1:O	2.41	-

Non-bonded interacted residues: TRP475, THR426, ASN206, PHE483, GLU482, SER56, ALA57, PRO161, TRP71, TYR58 AND PRO223

* Catalytic and nucleophile residues are highlighted in bold.
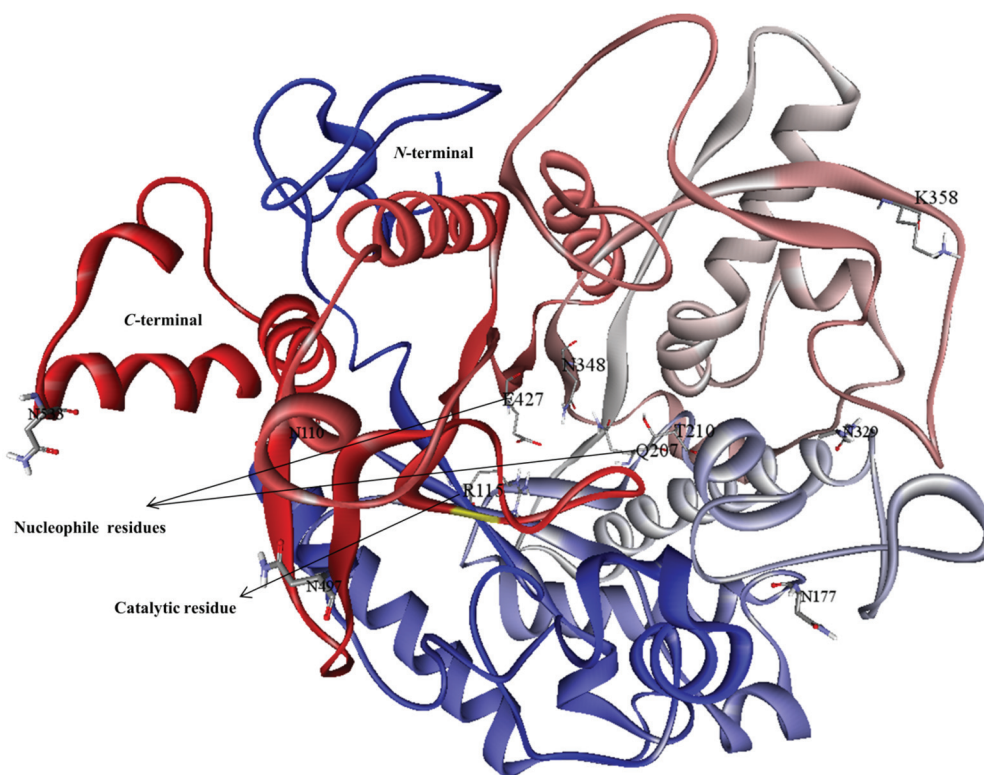

Figure 4. Three-dimensional structure of modeled myrosinase from *Brassica oleracea*. This figure was generated using DS 3.5 visualizer and colors of N-to-C terminal are based on their secondary structure.
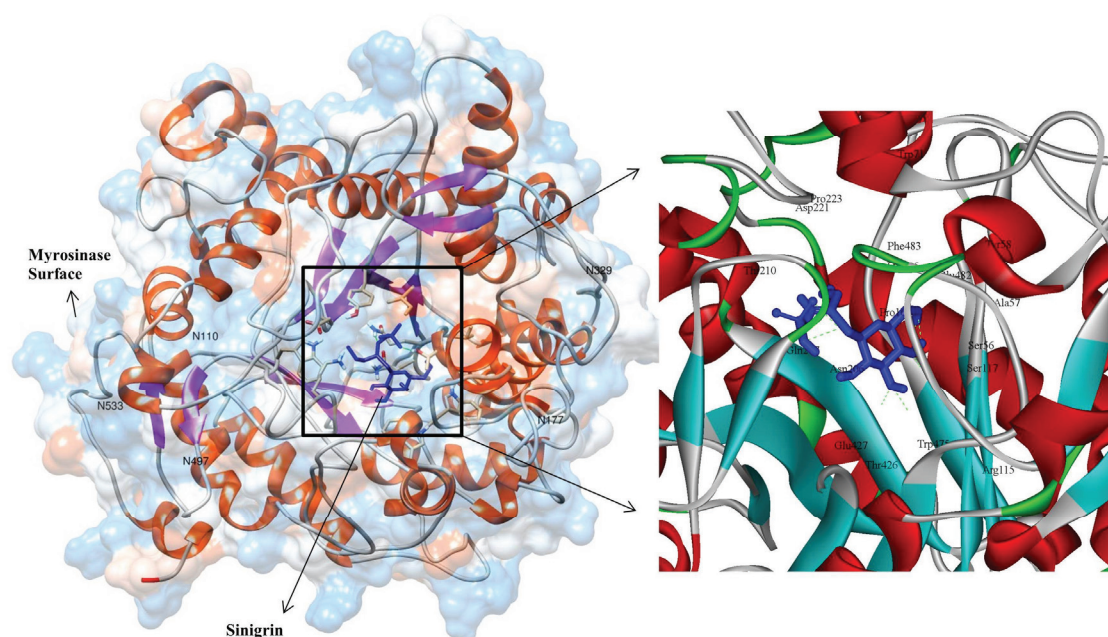
Figure 5. Graphical illustration of docked complex of myrosinase protein with sinigrin and potential *N*-glycosylation sites were highlighted. The 2D plot of interaction results were denoted as in black color (interacted amino acids), dotted green lines (hydrogen bonds) and blue color (sinigrin).

3.4. Stability Evolution of the MYR–Sinigrin Complex

An MD simulation was conducted to examine the MYR–sinigrin complex stability, and differences were monitored by RMSD with (holo) and without (apo) sinigrin. Accordingly, two individual MD simulations were carried out; the first contained only MYR as discussed above, the second consisted of the MYR–sinigrin complex file that was obtained from docking analysis. The protein mobility and structural changes were measured by RMSD plot of MYR backbone atoms throughout the 20-ns simulation. RMSD analysis of the MD trajectory noted that fluctuations were attained to 6.20 ns with an average of 0.26 Å, and then retained stability throughout simulation (Figure 6A). Apo and holo forms of MYR attained equilibration at over 6 ns. RMSD values of the apo form ranged between 0.0005 and 0.4711 Å, which was higher than those of the holo form, which ranged between 0.0005 Å and 0.2918 Å during 20 ns trajectories. Further, RMSF values were calculated for the apo and holo forms of MYR to estimate local protein mobility. Analysis of both RMSF values against backbone atoms indicated that a few atoms highly fluctuated in the *N*-terminal and *C*-terminal ends. In addition, the backbone atoms of RMSF values were compared with MYR (apo) and MYR-sinigrin complex (holo) (Figure 6B). Analyzing the trajectory files, the crucial residues had RMSF values ranging from 0 to 0.11 Å. In addition, compare to apo form of MYR, nucleophilic and catalytic residues namely GLN207, GLU427, and ARG115 were seen with slight fluctuations. The results from both apo and holo form of MYR showed that the residues in the sinigrin binding regions stabilized upon ligand recognition. Further, higher fluctuations in RMSF values were noted at *N*-terminal and *C*-terminal than other regions. The superimposition of initial and final structure revealed that sinigrin-binding orientation was well preserved and slight structural changes were observed in the active sites. Moreover the RMSF values of predicted ten *N*-glycosylation sites were ranged between 0.08 Å and 0.28 Å over 20-ns simulation time

(Supplementary Figure S2). Analysis of potential *N*-glycosylation sites revealed that slight and higher fluctuations were observed at the residues of 110, 117, 329, and 497 and 533, respectively, when compared to apo and holo form of MYR. The final trajectory of the 20-ns simulation was examined to predict binding affinity and hydrogen bond stability. Hydrogen bond stability analysis showed that the nucleophile GLN207 maintained the hydrogen bond throughout the entire 20-ns simulation (Supplementary Figure S3). This nucleophilic residue is responsible for enzyme-substrate interaction [55,56]. Another active site residue, THR210, formed a hydrogen bond in the hydrophobic state. By contrast, the hydrogen bonds of residues SER117, THR221, and LYS485 disappeared compared to the docking state in long dynamic simulations. We calculated the binding affinity of MYR-sinigrin to be −6.98 kcal/mol using the X-Score program (Table 2). In the case of non-bonded interaction, we observed negative logarithm dissociation constants of hydrophobic pair score (5.12), hydrophobic match score (5.13), hydrophobic surface score (5.11) and predicted mean binding affinity score (5.12). Both docking and MD simulation results suggests that the substrate sinigrin successfully recognized the MYR binding pocket at crucial active sites.
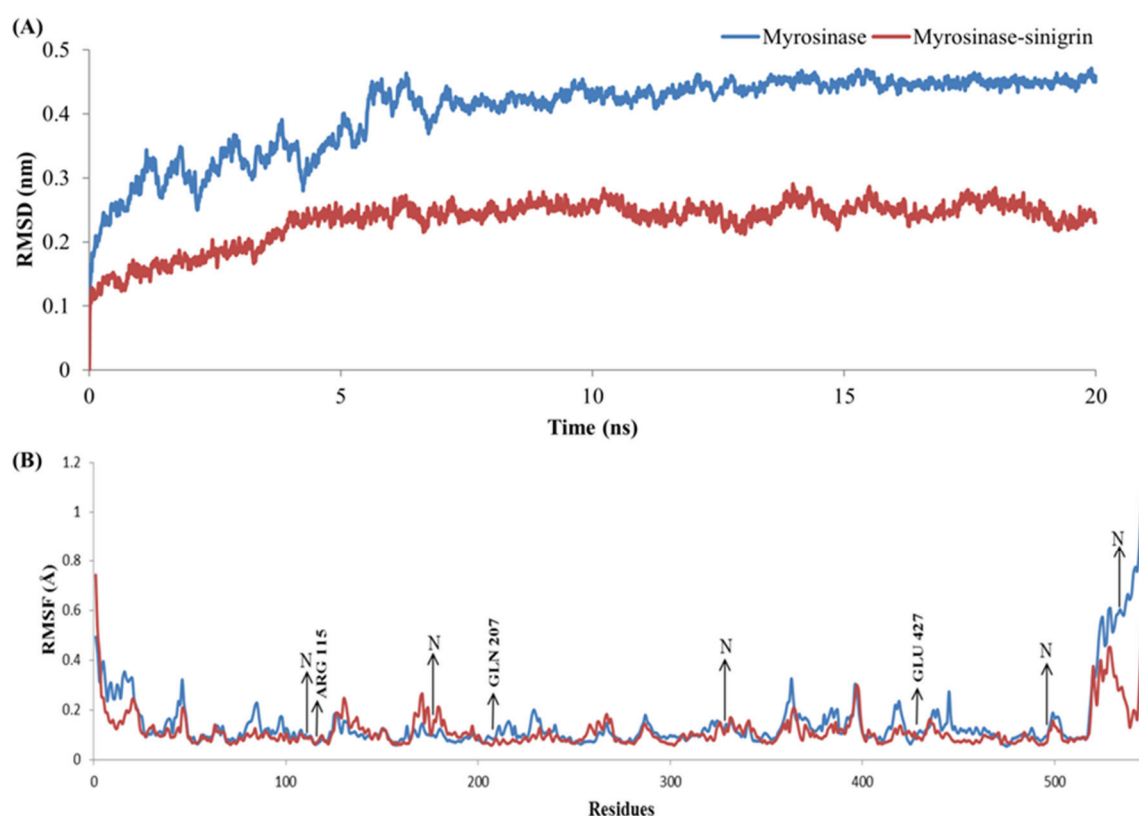

Figure 6. The root mean square deviation (RMSD) values of backbone (**A**) and root mean square Flucyuations (RMSF) values of atoms (**B**) for myrosinase and myrosinase-sinigrin complex during 20 ns dynamic simulations.

In general, conceptual model has been developed based on homology principle and help to characterize their putative functions in preliminary way. In addition, this structure based homology prediction method has limitations at the structural level, including large protein, multi-complex, and multi-domain proteins [57]. The developed model will be helpful to characterize the myrosinase function and sinigrin interaction at molecular level but further experimental confirmations are required.

Table 2. The bonded and non-bonded interaction scores of MYR–sinigrin complex after 20-ns molecular dynamics simulations. Scores were calculated using the X-score program.

Molecule	Predicted Binding Affinity (kcal/mol)	Hydrophobic Pair Score (pK_d)	Hydrophobic Match Score (pK_d)	Hydrophobic Surface Score (pK_d)	Predicted Mean Binding Affinity (pK_d)
MYR–sinigrin complex	−6.98	5.12	5.13	5.11	5.12

4. Conclusions

In this study, we constructed 3D model of MYR from cabbage (*B. oleracea*) based on sequence and structural similarity. The constructed model, which includes the Glyco_hydro_1 domain, was used for MD simulation to evaluate protein structural stability. The stereo-chemical quality of the optimized model confirmed that all residues were in favorable positions (Veryfy3D) and that the model was extremely good (ProQ). Our refined model included three disulfide bridges at the following positions: Cys2–Cys456, Cys34–Cys454, and Cys226–Cys234. In addition, ten conserved Asn-Xaa-Ser/Thr *N*-glycosylation sites were detected, among them five residues were potential sites, having a threshold value over 0.5. Notably, *N*-glycosylation sites, active sites and sequence conservation of our conceptual model showed similarity to the *S. alba* MYR crystal structure. Further, molecular interaction analysis showed two nucleophile (GLN207, and GLU427) and one catalytic (ARG115) residues were involved in the formation of hydrogen bonds along with hydrophobic interactions. The specific sulfur and nitrogen atoms present in sinigrin were well recognized by MYR active sites in the docked state. MD simulation results also suggested that strong hydrogen bond formation was maintained at residue GLU207 throughout the entire simulation, and binding orientation are the same as observed in the initial MYR–sinigrin complex structure. Our proposed 3D model of MYR and sinigrin binding orientation will be helpful to understand the MYR structural activities.

Acknowledgments

Author Contributions

Sathishkumar Natarajan, Senthil Kumar Thamilarasan, Jong-In Park, Mi-Young Chung and Ill-Sup Nou conceptualized this study. Sathishkumar Natarajan conducted the *in silico* research and drafted the manuscript. All the authors read and approved the final manuscript. Ill-Sup Nou is the corresponding author.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Kelly, P.; Bones, A.; Rossiter, J. Sub-cellular immunolocalization of the glucosinolate sinigrin in seedlings of Brassica juncea. *Planta* **1998**, *206*, 370–377.

2. Ciska, E.; Martyniak-Przybyszewska, B.; Kozlowska, H. Content of glucosinolates in cruciferous vegetables grown at the same site for two years under different climatic conditions. *J. Agric. Food Chem.* **2000**, *48*, 2862–2867.

3. Kliebenstein, D.J.; Kroymann, J.; Mitchell-Olds, T. The glucosinolate-myrosinase system in an ecological and evolutionary context. *Curr. Opin. Plant Biol.* **2005**, *8*, 264–271.

4. Martinez-Ballesta, M.; Carvajal, M. Myrosinase in Brassicaceae: The most important issue for glucosinolate turnover and food quality. *Phytochem. Rev.* **2015**, *14*, 1045–1051.

5. Ishida, M.; Hara, M.; Fukino, N.; Kakizaki, T.; Morimitsu, Y. Glucosinolate metabolism, functionality and breeding for the improvement of *Brassicaceae* vegetables. *Breed. Sci.* **2014**, *64*, 48–59.

6. Tiedink, H.; Davies, J.; van Broekhoven, L.; van der Kamp, H.; Jongen, W. Formation of mutagenic *N*-nitroso compounds in vegetable extracts upon nitrite treatment: A comparison with the glucosinolate content. *Food Chem. Toxicol.* **1988**, *26*, 947–954.

7. Avato, P.; Argentieri, M.P. Brassicaceae: A rich source of health improving phytochemicals. *Phytochem. Rev.* **2015**, *14*, 1019–1033.

8. Bones, A.M.; Rossiter, J.T. The myrosinase-glucosinolate system, its organisation and biochemistry. *Physiol. Plant.* **1996**, *97*, 194–208.

9. Burmeister, W.P.; Cottaz, S.; Rollin, P.; Vasella, A.; Henrissat, B. High resolution X-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. *J. Biol. Chem.* **2000**, *275*, 39385–39393.

10. Gleadow, R.M.; Møller, B.L. Cyanogenic glycosides: Synthesis, physiology, and phenotypic plasticity. *Annu. Rev. Plant Biol.* **2014**, *65*, 155–185.

11. Bednarek, P.; Osbourn, A. Plant-microbe interactions: Chemical diversity in plant defense. *Science* **2009**, *324*, 746–748.

12. Halkier, B.A.; Gershenzon, J. Biology and biochemistry of glucosinolates. *Annu. Rev. Plant Biol.* **2006**, *57*, 303–333.

13. Fan, S.; Meng, Q.; Xu, J.; Jiao, Y.; Zhao, L.; Zhang, X.; Sarkar, F.H.; Brown, M.L.; Dritschilo, A.; Rosen, E.M. DIM (3, 3'-diindolylmethane) confers protection against ionizing radiation by a unique mechanism. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18650–18655.

14. Verkerk, R.; Schreiner, M.; Krumbein, A.; Ciska, E.; Holst, B.; Rowland, I.; de Schrijver, R.; Hansen, M.; Gerhäuser, C.; Mithen, R.; *et al*. Glucosinolates in Brassica vegetables: The influence of the food supply chain on intake, bioavailability and human health. *Mol. Nutr. Food Res.* **2009**, doi:10.1002/mnfr.200800065.

15. Kadir, N.H.; David, R.; Rossiter, J.T.; Gooderham, N.J. The selective cytotoxicity of the alkenyl glucosinolate hydrolysis products and their presence in Brassica vegetables. *Toxicology* **2015**, *334*, 59–71.

16. Gill, C.I.; Haldar, S.; Porter, S.; Matthews, S.; Sullivan, S.; Coulter, J.; McGlynn, H.; Rowland, I. The effect of cruciferous and leguminous sprouts on genotoxicity, *in vitro* and *in vivo*. *Cancer Epidemiol. Biomarkers Prev.* **2004**, *13*, 1199–1205.

17. Porrini, M.; Riso, P.; Oriani, G. Spinach and tomato consumption increases lymphocyte DNA resistance to oxidative stress but this is not related to cell carotenoid concentrations. *Eur. J. Nutr.* **2002**, *41*, 95–100.

18. Aires, A.; Mota, V.R.; Saavedra, M.J.; Rosa, E.A.; Bennett, R.N. The antimicrobial effects of glucosinolates and their respective enzymatic hydrolysis products on bacteria isolated from the human intestinal tract. *J. Appl. Microbiol.* **2009**, *106*, 2086–2095.

19. Riga, E. The effects of Brassica green manures on plant parasitic and free living nematodes used in combination with reduced rates of synthetic nematicides. *J. Nematol.* **2011**, *43*, 119–121.

20. Sotelo, T.; Lema, M.; Soengas, P.; Cartea, M.; Velasco, P. *In vitro* activity of glucosinolates and their degradation products against brassica-pathogenic bacteria and fungi. *Appl. Environ. Microbiol.* **2015**, *81*, 432–440.

21. Vaughn, S.F.; Palmquist, D.E.; Duval, S.M.; Berhow, M.A. Herbicidal activity of glucosinolate-containing seedmeals. **2009**, *57*, 1821–1826.

22. Magrath, R.; Herron, C.; Giamoustaris, A.; Mithen, R. The inheritance of aliphatic glucosinolates in Brassica napus. *Plant Breed.* **1993**, *111*, 55–72.

23. Devi, J.R.; Thangam, E.B. Mechanisms of anticancer activity of sulforaphane from *Brassica oleracea* in HEp-2 human epithelial carcinoma cell line. *Asian Pac. J. Cancer Prev.* **2012**, *13*, 2095–2100.

24. Fahey, J.W.; Zalcmann, A.T.; Talalay, P. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* **2001**, *56*, 5–51.

25. Lenman, M.; Falk, A.; Rodin, J.; Hoglund, A.; Ek, B.; Rask, L. Differential expression of myrosinase gene families. *Plant Physiol.* **1993**, *103*, 703–711.

26. Kumar, R.; Kumar, S.; Sangwan, S.; Yadav, I.S.; Yadav, R. Protein modeling and active site binding mode interactions of myrosinase-sinigrin in Brassica juncea—An *in silico* approach. *J. Mol. Gr. Model.* **2011**, *29*, 740–746.

27. Cavasotto, C.N.; Phatak, S.S. Homology modeling in drug discovery: Current trends and applications. *Drug Discov. Today* **2009**, *14*, 676–683.

28. Peng, J.; Xu, J. Low-homology protein threading. *Bioinformatics* **2010**, *26*, i294–i300.

29. Bhattacharya, A.; Wunderlich, Z.; Monleon, D.; Tejero, R.; Montelione, G.T. Assessing model accuracy using the homology modeling automatically software. *Proteins* **2008**, *70*, 105–118.

30. Benson, D.A.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Wheeler, D.L. GenBank. *Nucleic Acids Res.* **2005**, *33*, D34–D38.

31. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinform.* **2008**, *9*, doi:10.1186/1471-2105-9-40.

32. Burmeister, W.P.; Cottaz, S.; Driguez, H.; Iori, R.; Palmieri, S.; Henrissat, B. The crystal structures of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an *S*-glycosidase. *Structure* **1997**, *5*, 663–675.

33. Burmeister, W.P. Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Crystallogr. D Biol. Crystallogr.* **2000**, *56 Pt 3*, 328–341.

34. Verdoucq, L.; Moriniere, J.; Bevan, D.R.; Esen, A.; Vasella, A.; Henrissat, B.; Czjze, M. Structural determinants of substrate specificity in family 1 beta-glucosidases: Novel insights from the crystal structure of sorghum dhurrinase-1, a plant beta-glucosidase with strict specificity, in complex with its natural substrate. *J. Biol. Chem.* **2004**, *279*, 31796–31803.

35. Tankrathok, A.; Iglesias-Fernandez, J.; Luang, S.; Robinson, R.C.; Kimura, A.; Rovira, C.; Hrmova, M.; Ketudat Cairns, J.R. Structural analysis and insights into the glycon specificity of the rice GH1 Os7BGlu26 beta-D-mannosidase. *Acta Crystallogr. D Biol. Crystallogr.* **2013**, *69 Pt 10*, 2124–2135.

36. Glaser, F.; Pupko, T.; Paz, I.; Bell, R.E.; Bechor-Shental, D.; Martz, E.; Ben-Tal, N. ConSurf: Identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* **2003**, *19*, 163–164.

37. Yang, J.; Roy, A.; Zhang, Y. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics* **2013**, *29*, 2588–2595.

38. Berendsen, H.J.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

39. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

40. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

41. Hess, B.; Bekker, H.; Berendsen, H.J.; Fraaije, J.G. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

42. Wallner, B.; Fang, H.; Elofsson, A. Automatic consensus-based fold recognition using Pcons, ProQ, and Pmodeller. *Proteins* **2003**, *53*, 534–541.

43. Eisenberg, D.; Luthy, R.; Bowie, J.U. VERIFY3D: Assessment of protein models with three-dimensional profiles. *Methods Enzymol.* **1997**, *277*, 396–404.

44. Laurie, A.T.; Jackson, R.M. Q-SiteFinder: An energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics* **2005**, *21*, 1908–1916.

45. Gupta, R.; Brunak, S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.* **2002**, 310–322.

46. Barzilai, J.; Borwein, J.M. Two-Point Step Size Gradient Methods. *IMA J. Numer. Anal.* **1988**, *8*, 141–148.

47. Wolf, L.K. Digital briefs. *Chem. Eng. News Arch.* **2009**, doi: 10.1021/cen-v086n047.p044.

48. Rappé, A.K.; Casewit, C.J.; Colwell, K.; Goddard, W., III; Skiff, W. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

49. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.

50. Karpagam, V.; Sathishkumar, N.; Sathiyamoorthy, S.; Rasappan, P.; Shila, S.; Kim, Y.-J.; Yang, D.-C. Identification of BACE1 inhibitors from Panax ginseng saponins—An Insilco approach. *Comput. Biol. Med.* **2013**, *43*, 1037–1044.

51. Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput. Aided Mol. Des.* **2002**, *16*, 11–26.

52. SchuÈttelkopf, A.W.; van Aalten, D.M. PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 1355–1363.

53. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.

54. Naumoff, D.G. Hierarchical classification of glycoside hydrolases. *Biochemistry* **2011**, *76*, 622–635.

55. Zhang, Y.; Zhao, Z.; Liu, H. Deriving chemically essential interactions based on active site alignments and quantum chemical calculations: A case study on glycoside hydrolases. *ACS Catal.* **2015**, *5*, 2559–2572.

56. Barleben, L.; Panjikar, S.; Ruppert, M.; Koepke, J.; Stockigt, J. Molecular architecture of strictosidine glucosidase: The gateway to the biosynthesis of the monoterpenoid indole alkaloid family. *Plant Cell* **2007**, *19*, 2886–2897.

57. Launay, G.; Simonson, T. Homology modelling of protein-protein complexes: A simple method and its possibilities and limitations. *BMC Bioinform.* **2008**, doi:10.1186/1471-2105-9-427.