

Soft Matter

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: P. Biswas, A. Adhikari, U. Pal, P. Singh, M. Das, T. Saha-Dasgupta, S. Shyam Choudhury, R. Das and S. K. Pal, *Soft Matter*, 2020, DOI: 10.1039/C9SM02479D.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

Dynamical Flexibility Modulates Catalytic Activity of a Themostable Enzyme: Key Information from Optical Spectroscopy and Molecular Dynamics Simulation

Pritam Biswas¹, Aniruddha Adhikari², Uttam Pal³, Priya Singh², Monojit Das⁴, Tanusri Saha-Dasgupta⁵, Sudeshna Shyam Choudhury^{1*}, Ranjan Das^{6*} and Samir Kumar Pal^{2,4*}

*¹Department of Microbiology,
St. Xavier's College,
30, Mother Teresa Sarani,
Kolkata 700016, India*

*²Department of Chemical, Biological and Macromolecular Sciences,
S. N. Bose National Centre for Basic Sciences,
Block JD, Sector III, Salt Lake,
Kolkata 700106, India*

*³Technical Research Centre,
S. N. Bose National Centre for Basic Sciences,
Block JD, Sector III, Salt Lake,
Kolkata 700106, India*

*⁴Department of Zoology,
Uluberia College, University of Calcutta,
Uluberia,
Howrah 711315, India*

*⁵Department of Condensed Matter Physics and Material Sciences,
S. N. Bose National Centre for Basic Sciences,
Block JD, Sector III, Salt Lake,
Kolkata 700106, India*

*⁶Department of Chemistry,
West Bengal state University,
Barasat, Kolkata 700126, India*

Corresponding authors E-mail: ranjan.das68@gmail.com (R Das)
skpal@bose.res.in (S. K. Pal)

sudeshna.s.choudhury@sxccal.edu (S. S. Choudhury)

Abstract

Enzymes are dynamical macromolecules and their conformation can be altered via local fluctuations of side chains, large scale loop and even domain motions which are intimately linked to their function. Herein, we have addressed the role of dynamic flexibility in the catalytic activity of a thermostable enzyme almond Beta-Glucosidase (BGL). Optical spectroscopy and classical molecular dynamics (MD) simulation were employed to study thermal stability, catalytic activity and dynamical flexibility of the enzyme. Enzyme assay reveals high thermal stability, and optimum catalytic activity at 333K. Polarization-gated fluorescence anisotropy measurements employing 8-Anilino-1-naphthelenesulfonic acid (ANS) have indicated increasing flexibility of the enzyme with an increase in temperature. Study of the atomic 3D structure of the enzyme shows the presence of four loop regions (LRs) strategically placed over the catalytic barrel as a lid. MD simulations have indicated that flexibility of BGL increases concurrently with temperature through different fluctuating characteristics of the enzyme's LR. Principal Component Analysis (PCA) and Steered Molecular Dynamics (SMD) simulation manifest the gatekeeper role of the four LR through their dynamic fluctuations surrounding the active site which controls catalytic activity of BGL.

Keywords: *Beta-Glucosidase; Spectroscopy; Loop Region; Molecular Dynamics simulation*

Introduction

β -Glucosidases (BGLs) constitute a subgroup of β -D-glucoside glucohydrolases (EC3.2.1.21) and are among the most widely studied enzymes owing to their wide variety of substrates, and broad distribution in the biological strata. BGLs are well characterized biologically important enzymes which catalyse the transfer of glycosyl groups between oxygen nucleophiles and hydrolyze the terminal, non-reducing β -D-glucosyl residues and thereby the terminal sugar moiety.¹⁻³ BGL have attracted considerable interests due to its importance in various biotechnological processes like hydrolysis of pharmaceutical ingredients, production of bio-ethanol from agricultural wastes.^{4, 5} BGL is used extensively in the wine, tea and juice industries.⁶⁻⁸ In addition to their use in industry, BGLs play pivotal roles in several biological processes.⁹ The physiological roles associated with these enzymes are diverse, ranging from facilitators of plant immune defence mechanisms to enzymatic breakdown of glucocerebroside and in treatment of Gaucher's disease in humans.^{10, 11} At present, the tertiary structures of several BGLs have been determined. All of these BGLs share a similar structural motif, and similar catalytic reaction mechanism.^{12, 13}

Almond β -glucosidase activity has been well investigated. However studies investigating correlation between its flexibility and activity is not as well understood as that between structure and activity. Previous studies have investigated the dynamical nature of BGL upon substrate binding,^{14, 15} which provides little to no information regarding the essential motions responsible for catalysis. Upon an inspection of molecular structure of the almond BGL four loop regions (LRs) were found to be placed strategically over the catalytic barrel of the enzyme like a lid. We conjecture that these LRs are involved in a gatekeeping mechanism to control enzymatic activity at different temperatures. In the present study, essential structural dynamics of almond BGL was correlated with its catalytic activity at different temperatures using steady state, circular dichroism

(CD) and picosecond resolved fluorescence spectroscopy along with an all atom molecular dynamics (MD) simulation, principal component analysis (PCA) and steered molecular dynamics simulation (SMD).

Result and Discussion

Homology modelling

Almond BGL [GenBank accession number AFH35017] exhibited a low similarity with structures within the PDB database based on the results of sequence analysis tools. The sequence search by NCBI BLAST¹⁶ over structures within PDB revealed that BGL matches with cyanogenic β -Glucosidase of *Trifolium repens* (1CBG) having 63% sequence identity. Accordingly, homology modelling method was employed to construct a 3D-model for BGL. The amino acid sequence was analysed using a set of known fold recognition online servers including PSI-BLAST,¹⁶ Phyre2,¹⁷ and HHpred.¹⁸ All servers commonly showed the cyanogenic β -Glucosidase from *T.repens* (1CBG)¹⁹ with the highest identity and lowest E-value. The BGL structure was found to belong to the same TIM barrel family of 1CBG: family 1 of glycosyl hydrolase. Consequently, 1CBG was proven to be a suitable template structure for modelling the BGL structure. The 3D-model was built for BGL sequence using the SWISS-MODEL server.²⁰ The 1CBG protein structure was used as a template.

The constructed 3D-model was submitted to different tools for quality evaluation. The backbone conformation study by the PRO-CHECK web server based on the Psi/Phi Ramachandran plot revealed that 90% of the residues were located in the most favourable region, 9.4% in the additional allowed region, 0.4% in the generously allowed region and 0.2% residues in the disallowed regions (Supplementary FigureS1). The evaluation by VERIFY-3D web server showed that 94.87% of the

residues resides in a favourable environment having an average 3D-1D score of above 0.2.²¹ Moreover, the statistics of highly refined structures were compared using the ERRAT web tool to study the overall quality of the model for non-bonded atomic interactions. The computed score by ERRAT (94.51%) indicates that the model is acceptable in the normal range (above 50%). The quality of the constructed model was confirmed based on the scores obtained from the above evaluation tools. The summary of the model analysis are shown in Supplementary Table S1.

Figure 1(a) shows the homology model of almond BGL. The BGL model has similar characteristics of other family members of 1 glycosyl hydrolases. The catalytic domain of the enzyme (BGL) consists of alternating pattern of β -strands and α -helices that is a TIM barrel fold $[(\alpha/\beta)_8]$.²² The BGL structures mostly consist of a TIM barrel as a common folding motif.²³ Evolutionarily, family 1 glycosyl hydrolases are dyad and Figure 1(b) shows the conserved catalytic sites on the β -barrel (Glu424; Glu211) of BGL. The catalytic site on the β -barrel is covered by four mostly unstructured regions guarding the entry gate of substrate as a lid. The regions are marked as follows: LR1 (residue no: 77-100); LR2 (residue no: 227-242); LR3a (residue no: 359-376); LR3b (residue no: 377-387); LR3c (residue no: 388-402) and LR4 (residue no: 426-449) as shown in the figure 1(c) and (d).

Secondary structure analysis

Almond β -glucosidase is relatively more thermostable than other β -glucosidases of plant origin,²⁴ including vanilla,²⁵ grape,²⁶ maize,²⁷ and tea leaves β -glucosidases.²⁸ The thermal stability of BGL was quantified by circular dichroism (CD) spectroscopy in the far UV region. The far UV CD spectra of BGL were characterized by two minima at 208 nm and 220 nm (Supplementary FigureS2).²⁹ Upon deconvolution of the spectra insignificant changes in the secondary structure was observed indicating to the absence of thermal denaturation within the temperature range of

our study. Only 1.7% decrease in the α -helix, 5.5% increase in antiparallel β -sheet, increase of 0.3% parallel β -sheet, 0.6% β -turn and 0.4% increase in random coil was observed (Supplementary Table S2a). The refolding of the BGL molecule was studied through gradual cooling of the enzyme to 278K and subsequently, deconvolution of the CD spectra was found to manifest retention of the enzyme native structure with 1.0% recovery of α -helix and 2.0% in antiparallel β -sheet (Supplementary Table S2b). No significant changes in random coil, turns and parallel β -sheets were observed (Supplementary Table S2b). Thus, from secondary structure analysis no appreciable structural denaturation of the enzyme was noticed upon thermal treatments. Upon re-evaluation of the CD spectra using BestSel CD deconvolution server similar trends in the results were obtained. No significant change to the secondary structure was observed (Supplementary Table S2a & S2b), indicating and reaffirming the thermal stability of the enzyme.

Enzymatic activity assay

Enzyme kinetics of BGL glucosidase activity were conducted in details at temperatures ranging from 278K to 353K. The initial velocity of the glycoside hydrolysis by BGL was measured under different rutin (substrate) concentrations (4–20 mM). Lineweaver-Burk plots were used to determine the reaction kinetic parameters (K_m , V_{max} , k_{cat} , k_{cat}/K_m). BGL belongs to the class of glycohydrolases that catalyzes the hydrolysis of its natural substrate rutin into quercetin (Figure 2a and 2b). Figure 2c and 2d illustrate the effect of increasing temperature on the product formation and catalytic efficiency. Thermal treatment was found to enhance the catalytic efficiency of the enzyme (almond BGL) which becomes maximum at 333K, and further increase of temperature to 353K resulted in decreased catalytic efficiency relative to that at 333K. It is also noteworthy that no discernible product formation was observed at temperature as low as 278K, while low product formation and reduced catalytic efficiency was observed at a much higher temperature (353K). This

is consistent with thermal stability of almond BGL over a wide range of temperature indicating to formation of a stable enzyme-substrate complex even at a temperature as high as 353K. The summary of the kinetics parameters are shown in Table 1.

Time-resolved Fluorescence Anisotropy

Molecular docking studies of ANS in presence of the enzyme (BGL) shows its' molecular interactions with the amino acid residues of BGL (Figure 3). The targeted docking found interactions between the probe and the tryptophan residues (Trp396 and Trp475) at the catalytic barrel of BGL, which was close to the active site. The dynamical flexibility of BGL was revealed from time-resolved anisotropy decay ($r(t)$) of ANS³⁰⁻³⁴ in its complex with the enzyme at different temperatures (Figure 4) over the range of 278K to 353K. A three exponential model was required to obtain satisfactory fits to the polarization gated anisotropy decay, $r(t)$. The fast rotational correlation time (ϕ_1) of ~100-400 ps corresponds to rapid local motion of the probe ANS bound to Tryptophan (Trp-396 and Trp-475) residues, whereas the slower one (ϕ_2) of 0.5-2.36 ns (Table 2) is reflective of segmental motion of the neighbouring loop regions close to the active (catalytic) site or the catalytic barrel. The much longer time constant ($\phi_3 > 20$ ns, Table 2) in the anisotropy decay (Figure 4) corresponds to global tumbling motion of the overall probe-enzyme complex. A monotonous decrease in both the correlation times (ϕ_1 and ϕ_2) was noted with rise in temperature. In comparison to the lower temperature range (278K to 313K) significant decrease in ϕ_1 and ϕ_2 was noted at 333K or 353K due to increased mobility of the dye local environment that results from significantly increased dynamical flexibility of the amino acid residues (Trp-396 and Trp-475) and the loop regions close to the catalytic barrel. The significant decrease, especially in ϕ_2 , in going from 313K to 333K indicates significantly increased dynamic flexibility of the loop regions which likely facilitates substrate binding to the catalytic barrel resulting in enhanced catalytic

efficiency at the higher temperature (333K). Upon further increase of temperature to 353K, remarkable decrease in the rotational time constants was observed which is consistent with rapid fluctuations of the amino acid residues and loop regions near the catalytic barrel which may sometimes impede substrate entry into the catalytic barrel or lower the stability of the enzyme-substrate complex.

Molecular dynamics (MD) simulation

The stability and mobility of the BGL structure was studied by subjecting the structure to 48 ns all atom MD simulation at low (278K), moderate (313K), moderately high (333K) and high (353K) temperatures. Supplementary Figure S3 shows the RMSD and R_g plot of BGL at different temperatures. RMSD indicates the overall change in the geometry of the enzyme with respect to its initial structure. From Figure S2 it becomes evident that after ~ 25 ns, the structure reaches an equilibrium geometry. Further, there is a monotonous increase in the RMSD value with respect to the temperature echoing the effect of temperature on flexibility of the enzyme. In order to have an insight into the effect of temperature on the overall shape and compactness of the enzyme the radius of gyration (R_g) was calculated. R_g is defined as the average distance of the collection of atoms from their common centre of mass. Higher R_g values indicate looser packing of the protein structure whereas lower values respond to a much compact structure of it. Figure S2 shows further that there is no significant change in the overall shape and compactness of BGL over the course of the simulation event corroborating our data from circular dichroism study. Furthermore, flexibility in different parts of the enzyme (BGL) at different temperatures was assessed by calculating the Root mean squared fluctuation (RMSF) which is a measure of the displacement of an amino acid residue around its averaged position during a defined period of time and allows detection of the regions of high flexibility in a protein. Figure 5 shows RMS fluctuations of BGL at different

temperatures. It is observed from Figure 5a and 5b that the loop regions experience the most of the residue fluctuations in BGL. Four loop regions in particular form the epicentre of residue fluctuations: LR1, LR2, LR3, and LR4 (Figure 1) in the enzyme. These regions comprise of residues 77-100 (LR1), 227-242 (LR2), 359-376 (LR3a), 377-387 (LR3b), 388-402 (LR3c), and 426-449 (LR4). All these regions experience fluctuations greater than 4Å whereas the regions experiencing fluctuations lesser than 4Å were mostly found to be located in the binding site or the catalytic barrel having conserved residues. Figure 5c shows fluctuations experienced by the LRs at different temperatures. It is clearly evident that the LR1 undergoes monotonous increase in fluctuations as does the LR2 with rise in temperature. However, the LR3 region shows a form of synchronous fluctuations at its sub regions LR3a and LR3c while the LR3b remains rigid acting like a pivot. The LR4 region however displays an unusual fluctuation. The LR4 region shows increased fluctuations at 278K, the movement becomes fairly rigid at 313 and 333K, increasing again at 353K. It is noteworthy that Figure 5c does not indicate the magnitude of fluctuations but rather the effect of temperature change on the LRs with respect to control (non-LRs). The changing pattern of fluctuations in the LRs indicate a possible pattern over the course of temperature range. It should be noted that the LRs cover the catalytic barrel like a lid whose fluctuations likely contribute to or control the catalytic activity of the enzyme.

Normal mode analysis

To obtain better insights and identify the overall pattern of motions of the entire protein Principal component analysis (PCA) was used. The eigenvectors are calculated from the covariance matrix of a simulation trajectory. The trajectories are filtered along each eigenvector to identify the dominant motions during a simulation event. A large portion of the overall fluctuations of the macromolecules observed from RMSF can often be accounted for by a few low-frequency

eigenvectors with large eigenvalues. If the motions are similar, the eigenvectors and eigenvalues coming from individual trajectories should be similar to each other. All atom Principal component analysis was applied to study the major motions of BGL at different temperatures. Figure 6 and Figure 7 display the porcupine plot of the first two eigenvectors and the mobility plots of BGL. The mobility plot (Figure 7) clearly shows an initial increase in motion in the LR3 with a decrease after 333K, which directly correlates with the catalytic activity of the enzyme. From the porcupine plots (Figure 6) we can clearly visualize the major motions of the protein at different temperatures. Figure 6 indicates that the LR region is the epicentre of major motions in the protein. At 278K the LR4 region serves as the centre of major fluctuations while the other LR regions remain rigid. In case of 313K major motions in the LR3 region are observed. The porcupine as well as the mobility plots indicate the presence of three different motions at the LR3 region. The LR3A and LR3C regions show a propeller motion; moving in clockwise direction, while the LR3B region remains rigid acting like a pivot. The motion of the entire protein is clockwise in the first mode. In the second mode, the motion of LR3 is not the primary motion and the protein as a whole moves in an anti-clockwise direction relative to that of LR3. This suggests a possible change in the orientation of the loop region covering the catalytic barrel making it more accessible to the incoming substrate. The principal motions for BGL at 333K illustrates the clockwise propeller motion of the LR3 region, while the entire protein experiences a counter-clockwise motion. Both the modes in the porcupine plot and the mobility plot indicate the presence of significant motions in the LR2 region. The presence of new sites of prominent motion could aid in proper orientation of the catalytic site for the approaching substrate, thereby, facilitating the enzymatic activity of BGL. The mobility plot at 353K shows the presence of several new sites of motion (LR1 and LR2). From the porcupine plots random and independent motions of different sites are visualized and the fluctuations of the

LRs become highly inordinate and asynchronous which may affect the enzymatic activity at high temperatures. [Visual information regarding the various modes and major motions are provided in a Supplementary Video S1-4].

Steered molecular dynamics simulation

Study of the energy required to pull a substrate molecule (rutin) out of the binding site of BGL (Figure 8a), gave further clarity to the notion of ease of substrate binding at higher temperature. Following the methodology of steered molecular dynamics simulation (SMD) described in the earlier work,³⁵ it was observed (Figure 8b) that with increasing temperature less energy was required to break the enzyme substrate interactions and pull the substrate out. 333K was found to be the optimum temperature at which a meta-stable energy minimum was observed after a short pull (Figure 8b). Molecular interaction studies at the minimum indicated compensating bond formation and breakdown with the substrate (Figure 8c & 8d). No compensating interactions were found for 353K, which resulted in the monotonous increase of the potential mean force (PMF) (Figure 8b, 8e, 8f). Detailed analysis of the interacting residues found the presence of water-bridge connecting Arg-325 to the hydroxylic group adjacent to the glycosidic bond at the meta-stable state (Figure 8c & 8d). Our previous work had discussed the role of bulk and bound water molecules in biological recognition and hydrolysis.³⁶ The bound water molecule in the meta-stable state may be involved in the catalytic activity of the enzyme. Table 3 lists the equilibrium distances between the centre of mass of BGL and rutin. It was evident from the table that rutin (substrate) can go further into the active site of BGL at 333K than any other temperature. Solvent exposure of the substrate at different temperatures (Table 4) was also consistent with the finding.

Structure-Activity relationship

Almond BGL causes the catalytic hydrolysis of the flavonoid glycoside Rutin to release Quercetin. Temperature dependent enzymatic activity was monitored at temperature ranging 278K-353K. We observed no product formation at temperatures as low as 278K. Increased product formation at temperatures as high as 333K suggested the ease of the enzyme substrate complex formation. The kinetics parameter were calculated from the Lineweaver-Burk plot. The catalytic efficiency (kcat/Km) was a typical bell shaped curve, temperature optima at 333K, which was in agreement with previous studies (Figure 2)²⁴. Loss of catalytic activity was observed at 353K, however the thermal stability observed at this temperature pointed towards dynamical flexibility being the possible architect of the structure-activity relationship.

The results of activity study point to the dynamical flexibility as the possible facilitator of enzymatic activity at different temperatures. Studies by Banerjee and Pal,³⁷ Karplus et al³⁸ and Min et al³⁹ found a direct correlation between picosecond to nanosecond timescale dynamics and enzymatic activity. Picosecond resolved fluorescence anisotropy of the ANS-BGL complex comprehensively shows a monotonous increase in dynamical flexibility of the protein with rise in temperatures (Figure 4 & Table 2). The rise in temperature creates a more labile microenvironment of the probe enhancing orientational relaxation of ANS. RMSD calculations further support the activity-flexibility correlation being solely responsible in catalysis (Supplementary Figure S2). While the RMSF studies identify four loop regions [LR1; LR2; LR3; LR4] as the possible regions of high fluctuations (Figure 5). The loop regions (LRs) are placed strategically as a lid over the catalytic barrel. Other regions of the enzyme experiences fluctuations of lesser magnitude than the loop regions. Time-resolved anisotropy measurements and RMSF analysis as well as the SMD simulations indicate to the possible gatekeeping function of the LRs where the change in fluctuations of the LRs associated with their possible orientation and re-orientation at an optimized

position dictates the entry of an incoming substrate at the active site of the catalytic barrel. Our study also indicates that higher degree of fluctuations beyond the threshold level at 353K (Figure 4) likely cause the LRs to an unfavourable orientation for entry of the substrate at the active site, leading to reduced catalytic efficiency ($k_{\text{cat}}/K_{\text{m}}$) (Figure 2).

Loop dynamics controls catalysis in almond β Glucosidase

The RMSF studies indicate towards four strategically placed LRs as the possible gatekeepers of enzymatic activity at different temperatures. The overall fluctuations of the macromolecule observed from RMSF cannot provide concrete information regarding major motions along the trajectories of simulation. The trajectories need to be filtered along each eigenvector to identify the dominant motions during a simulation event. Principal component analysis (PCA) of BGL at different temperatures identified the LRs as the epicentre of dominant motions at different temperatures. Our study also found that residues of the catalytic barrel experienced no major fluctuations across all temperature ranges. Dynamical rigidity of the LRs at 278K barring LR4 likely results in a possible closed gate conformation checking the access of an incoming substrate in the catalytic site and in consequence no enzymatic activity at the said temperature (Figure 6 & 7) [Supplementary Video S1] has been noted. The concerted motions of LR3 and rest of the protein, along with the propeller motion of LR3 indicates to a change in orientation to an open gate conformation at 313K (Figure 6 & 7) [Supplementary Video S2]. The open gate conformation would allow access to the substrate resulting in enzymatic activity at 313K. The concerted motion was maintained at 333K and new centres of prominent motions at LR1 and LR2 were observed (Figure 6) [Supplementary Video S3]. These new sites of prominent motion could aid in the optimum orientation of the lid LRs aiding in enhanced catalysis and turnover observed at 333K (Table 1). Ease of substrate access at this temperature also reverberated in the least pulling force

as observed in SMD simulations (Figure S4). At 353K there is a loss of concerted motions of the LRs and fluctuations become independent of each other and random with several new sites of major motions (Figure 6) [Supplementary Video S4]. This loss in the concerted motion of the LRs and increased randomness may rationalize increased flexibility of the ANS-BGL complex observed in polarization gated anisotropy measurements (Figure 4) and the fall in the catalytic activity and turnover at 353K (Figure 2 & Table 1). The random fluctuations of the LRs may lead to the inability or possibly dissociation of the substrate from the active site of the enzyme (BGL) contributing to reduced catalytic efficiency.

Materials and Methods

Materials

Beta-Glucosidase (*Prunus dulcis*) ($\geq 6\text{U/mg}$), Rutin hydrate ($\geq 94\%$), 8-Anilino-1-naphthelenesulfonic acid (ANS) ($\geq 97\%$) were purchased from Sigma (Saint Louis, USA). Ethanol (ACS) was purchased from Merck (Darmstadt, Germany). BGL and ANS solutions were prepared in Phosphate buffer (10mM, pH 7.0) using water from Millipore system. Rutin solution was prepared in ethanol.

Sample Preparation

The BGL-ANS complex for anisotropy analysis was prepared by mixing ANS (0.6 μM) with BGL (6 μM) in phosphate buffer (pH 7.0) with continuous stirring for 5h at 4°C, which was filtered extensively to remove the free probe from the solution ³¹.

Computational modelling of the enzyme

The full length amino acid sequence of almond BGL was obtained from NCBI GenBank database [Accession number: AFH35017.1]. Homology modelling was used to construct a 3D structure for the protein using *Trifolium repens* cyanogenic beta glucosidase (1CBG) as template. SWISS-MODEL server was used to build the 3D structure of the enzyme. Assessment of the constructed 3D model was done by the VERIFY-3D,⁴⁰ PRO-CHECK,⁴¹ ERRAT,⁴² and ProSA-Web⁴³ web servers.

Circular dichroism (CD) spectroscopy

Temperature dependent circular dichroism (CD) spectroscopy study was conducted with 1 μM BGL solution using a quartz cuvette of 1 mm path length in a JASCO 815 spectrometer with a temperature controller attachment (Peltier). The experimental temperature was gradually increased from 278K to 353K and then decreased back to 278K. The deconvolution of the CD signals into relevant secondary structure was done by CDNN software⁴⁴ and were re-evaluated using BestSel CD deconvolution server.⁴⁵

Enzyme activity assay

Temperature-dependent (278K-353K) enzymatic activity studies were conducted with BGL concentration 0.5 μM and the concentrations of the substrate Rutin ($\lambda_{\text{max}} \sim 358 \text{ nm}$) was varied in the range 4-20 mM. The substrate was cleaved to produce Quercetin ($\lambda_{\text{max}} \sim 374 \text{ nm}$), the absorbance of which was monitored in Shimadzu Model UV-2600 spectrophotometer with an attachment for temperature dependent studies (TCC 240 A). The Michaelis-Menten kinetics parameters were determined from the Lineweaver-Burk plot by a nonlinear least square fitting software.⁴⁶

Molecular docking

Molecular docking studies were performed to identify the binding site of ANS and rutin in BGL. The ligand (ANS & rutin) structures were obtained from the PubChem database. Targeted docking was performed at the active site using AutoDock Vina.⁴⁷ The molecular interactions were viewed using Schrodinger Maestro 2018-1 (Academic release).

Polarization-gated Anisotropy Measurements

Time resolved fluorescence transients were measured in a LifeSpec-ps picosecond diode laser-pumped fluorescence spectrophotometer from Edinburgh Instruments (Livingston, U.K). Picosecond excitation pulses from the picoquant diode laser was used at 375 nm for excitation of ANS with an instrument response function (IRF) of 60 ps with an attachment for temperature dependent studies [temperature range: 278K-353K]. A microchannel-plate-photomultiplier tube (MCP-PMT; Hammamatsu Photonics, Kyoto, Japan) was used to detect the photoluminescence from the sample after dispersion through a monochromator. Our time-resolved instrument provide a time resolution of 20 ps which is one third of the instrument response (IRF) for detection of time constants after the de-convolution of the IRF. Time-resolved anisotropy, $r(t)$, was determined from the following equation

$$r(t) = \frac{I_{VV}(t) - G.I_{VH}(t)}{I_{VV}(t) + 2.G.I_{VH}(t)}$$

Where $I_{VV}(t)$ and $I_{VH}(t)$ are parallel and perpendicular polarized fluorescence decays of the dye, respectively, recorded using vertically polarized excitation light. G is an instrument and wavelength dependent correction factor to compensate for the polarization bias of the detection system and its magnitude was obtained by a long tail matching technique.^{48, 49} The anisotropy decays were fitted using a multi-exponential decay model $r(t) = r_0 \sum_{i=1}^3 \alpha_i \exp(-t/\theta_i)$ where r_0

is initial anisotropy, θ is are the rotational correlation times of the fluorescence probe (ANS) hinting to dynamical flexibility of the probe environment in its complex with protein (BGL) ⁵⁰.

Classical molecular dynamics (MD) simulation

The temperature dependent dynamical behaviour of BGL was further studied by molecular dynamic (MD) simulation. The constructed BGL structure was subjected to energy minimization by Schrodinger Maestro 2018-1(Academic release). Simulations were done at temperatures 278K, 313K, 333K and 353K. Protein molecule was placed in the centre of a periodic boundary box with a minimum distance of 10 Å from each side. Then, the protein was solvated with pre-optimized simple point-charged (SPC) water molecules. Additionally, the system was charge neutralized by adding necessary counter ions (seven Na⁺) A five step relaxation protocol was used to equilibrate the entire system ⁵¹. For the final runs, normal pressure and a defined temperature were used and no particles were restrained. The ensemble type was NPT with a constant number of particles, normal pressure (1 bar) and a defined temperature. Nosé-Hoover thermostat and Martyna-Tobias-Klein barostat maintained the temperature and pressure of the system, respectively. MD was run in OPLS 2005 force field ⁵² with a short range Coulombic interaction cut-off of 9 Å. Long range coulombic interactions were handled by smooth particle mesh Ewald method with a Ewald tolerance of 10⁻⁹. The simulations were run for 48 ns at the said temperatures. Overall structural changes in the protein with time was computed in terms of root mean square deviation (RMSD). Radius of gyration (Rg) and residue-wise fluctuations (RMSF) were also computed from the simulation trajectory.

Normal mode analysis

Reduction of dimensionality of the MD simulation data helps in identifying the configurational space in which anharmonic motion takes place. Principal components analysis (PCA) takes the trajectory of a molecular dynamics simulation and extracts the dominant modes in the motion of the molecule.⁵³⁻⁵⁵ These dominant motions correspond to correlated vibrational modes or collective motions of groups of atoms in the normal modes analysis. The direction and amplitude of the dominant motions along a simulation trajectory by Principal component analysis (PCA) was computed by Normal Mode Wizard in VMD 1.9.3 program⁵⁶ using ProDy⁵⁷. Porcupine plots and mobility plots for the respective temperature was generated by the program showing a graphical summary of the motions along the trajectory. The Supplementary Video of the principal component analysis were also prepared in VMD.⁵⁶

Steered molecular dynamics

Steered molecular dynamics³⁵ was performed in Schrodinger Maestro 2018-1 (Academic release) using the metadynamics module. Rutin bound BGL as obtained by molecular docking, was subjected to MD simulation following the previously mentioned protocol at four different temperatures. Then, starting with the temperature equilibrated substrate (rutin) bound BGL conformations, steered dynamics was performed to probe the energy required to pull the substrate out of the active site. Therefore, the distance from the centre of mass of BGL and the centre of mass of rutin was gradually increased. A cut-off value for the subsequent increase in the energy was set to be 20 kcal/mol. All the other parameters were same as the classical MD simulation.

Conclusion

The 3D-model and secondary structure analysis of an enzyme almond BGL illustrated various remarkable characteristics of this glycosyl hydrolase. The structure of this thermostable enzyme consists of four loop regions (LRs) placed strategically over the catalytic barrel as a lid. The concerted motion of these LRs likely acts as a gatekeeper of catalytic activity of the enzyme at different temperatures, and consequently, its ability to be active at moderately high temperatures. The findings of this study can provide a background for future efforts to enhance enzymatic capabilities through rational design of recombinant enzymes.

CONFLICT OF INTEREST

The authors declare there is no conflict of interest in this work.

ACKNOWLEDGMENT:

The authors thank DST-SERB, India, for financial grants, EMR/2016/004698. The authors gratefully acknowledge the financial assistance provided by UGC (F.PSW-137/15-16 ERO). The authors also thank DBT (WB)-BOOST scheme for financial grant, 339/WBBDC/1P-2/2013.

Reference

1. Y. Bhatia, S. Mishra and V. Bisaria, *Critical reviews in biotechnology*, 2002, **22**, 375-407.
2. J. R. Ketudat Cairns and A. Esen, *Cellular and Molecular Life Sciences*, 2010, **67**, 3389-3405.
3. J. D. McCarter and G. S. Withers, *Current opinion in structural biology*, 1994, **4**, 885-892.
4. J. Xiao, T. S. Muzashvili and M. I. Georgiev, *Biotechnology advances*, 2014, **32**, 1145-1156.
5. R. R. Singhanian, A. K. Patel, R. K. Sukumaran, C. Larroche and A. Pandey, *Bioresource technology*, 2013, **127**, 500-507.
6. O. Shoseyov, B. A. Bravdo, D. Siegel, A. Goldman, S. Cohen, L. Shoseyov and R. Ikan, *Journal of Agricultural and Food Chemistry*, 1990, **38**, 1387-1390.
7. E. Su, T. Xia, L. Gao, Q. Dai and Z. Zhang, *Food and Bioprocess Processing*, 2010, **88**, 83-89.
8. R. Agrawal, A. Verma and A. Satlewal, *Innovative food science & emerging technologies*, 2016, **33**, 472-482.
9. G. Singh, A. Verma and V. Kumar, *3 Biotech*, 2016, **6**, 3.
10. A. V. Morant, K. Jørgensen, C. Jørgensen, S. M. Paquette, R. Sánchez-Pérez, B. L. Møller and S. Bak, *Phytochemistry*, 2008, **69**, 1795-1813.
11. P. Belchetz, J. Crawley, I. Braidman and G. Gregoriadis, *The Lancet*, 1977, **310**, 116-117.
12. B. Henrissat and G. Davies, *Current opinion in structural biology*, 1997, **7**, 637-644.
13. H. M. Jespersen, E. A. MacGregor, B. Henrissat, M. R. Sierks and B. Svensson, *Journal of protein chemistry*, 1993, **12**, 791-805.
14. Y. Li, X. Hu, J. Sang, Y. Zhang, H. Zhang, F. Lu and F. Liu, *International journal of biological macromolecules*, 2018, **119**, 462-469.
15. N. S. F. Mazlan and N. B. Ahmad Khairudin, *Journal of Biomolecular Structure and Dynamics*, 2016, **34**, 1486-1494.
16. S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Nucleic acids research*, 1997, **25**, 3389-3402.
17. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass and M. J. Sternberg, *Nature protocols*, 2015, **10**, 845.
18. J. Söding, A. Biegert and A. N. Lupas, *Nucleic acids research*, 2005, **33**, W244-W248.
19. T. Barrett, C. Suresh, S. P. Tolley, E. Dodson and M. Hughes, *Structure*, 1995, **3**, 951-960.
20. T. Schwede, J. Kopp, N. Guex and M. C. Peitsch, *Nucleic acids research*, 2003, **31**, 3381-3385.
21. J. U. Bowie, R. Luthy and D. Eisenberg, *Science*, 1991, **253**, 164-170.
22. R. Opassiri, B. Pomthong, T. Onkoksoong, T. Akiyama, A. Esen and J. R. K. Cairns, *BMC plant biology*, 2006, **6**, 33.
23. G. Davies and B. Henrissat, *Structure*, 1995, **3**, 853-859.
24. N. S. Terefe, P. Sheean, S. Fernando and C. Versteeg, *Applied microbiology and biotechnology*, 2013, **97**, 2917-2928.
25. O. Márquez and K. N. Waliszewski, *International journal of food science & technology*, 2008, **43**, 1993-1999.
26. M. Lecas, Z. Y. Gunata, J.-C. Sapis and C. L. Bayonove, *Phytochemistry*, 1991, **30**, 451-454.
27. A. Esen, *Plant Physiology*, 1992, **98**, 174-182.

28. Y.-Y. Li, C.-J. Jiang, X.-C. Wan, Z.-Z. Zhang and D.-X. Li, *Acta biochimica et biophysica Sinica*, 2005, **37**, 363-370.
29. M. A. Shah, S. Mishra and T. K. Chaudhuri, *European Biophysics Journal*, 2011, **40**, 877-889.
30. P. Hazra, D. Chakrabarty, A. Chakraborty and N. Sarkar, *Biochemical and biophysical research communications*, 2004, **314**, 543-549.
31. S. K. Pal, J. Peon and A. H. Zewail, *Proceedings of the National Academy of Sciences*, 2002, **99**, 15297-15302.
32. K. Sahu, S. K. Mondal, S. Ghosh, D. Roy and K. Bhattacharyya, *The Journal of chemical physics*, 2006, **124**, 124909.
33. G. Wang, Y. Gao and M. L. Geng, *Biochimica et Biophysica Acta (BBA)-General Subjects*, 2006, **1760**, 1125-1137.
34. P. Sarkar, S. Bharill, I. Gryczynski, Z. Gryczynski, M. P. Nair and A. G. Lacko, *Journal of Photochemistry and Photobiology B: Biology*, 2008, **92**, 19-23.
35. S. Mao, J.-W. Wang, F. Liu, Z. Zhu, D. Gao, Q. Guo, P. Xu, Z. Ma, Y. Hou and X. Cheng, *Microbial cell factories*, 2018, **17**, 141.
36. S. K. Pal and A. H. Zewail, *Chemical Reviews*, 2004, **104**, 2099-2124.
37. D. Banerjee and S. K. Pal, *Langmuir*, 2008, **24**, 8163-8168.
38. M. Karplus and J. Kuriyan, *Proceedings of the National Academy of Sciences*, 2005, **102**, 6679-6685.
39. W. Min, X. S. Xie and B. Bagchi, *The journal of physical chemistry B*, 2008, **112**, 454-466.
40. R. Lüthy, J. U. Bowie and D. Eisenberg, *Nature*, 1992, **356**, 83.
41. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *Journal of applied crystallography*, 1993, **26**, 283-291.
42. C. Colovos and T. O. Yeates, *Protein science*, 1993, **2**, 1511-1519.
43. M. Wiederstein and M. J. Sippl, *Nucleic acids research*, 2007, **35**, W407-W410.
44. G. Böhm, R. Muhr and R. Jaenicke, *Protein Engineering, Design and Selection*, 1992, **5**, 191-195.
45. A. Micsonai, F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers and J. Kardos, *Proceedings of the National Academy of Sciences*, 2015, **112**, E3095-E3103.
46. R. Biswas and S. K. Pal, *Chemical physics letters*, 2004, **387**, 221-226.
47. O. Trott and A. J. Olson, *Journal of computational chemistry*, 2010, **31**, 455-461.
48. G. F. Schröder, U. Alexiev and H. Grubmüller, *Biophysical journal*, 2005, **89**, 3757-3770.
49. D. O'Connor, *Time-correlated single photon counting*, Academic Press, 2012.
50. N. Jain, M. Bhattacharya and S. Mukhopadhyay, *Biophysical journal*, 2011, **101**, 1720-1729.
51. U. Pal, S. K. Pramanik, B. Bhattacharya, B. Banerji and N. C. Maiti, *SpringerPlus*, 2016, **5**, 1121.
52. W. L. Jorgensen and J. Tirado-Rives, *Journal of the American Chemical Society*, 1988, **110**, 1657-1666.
53. A. Amadei, A. B. Linssen and H. J. Berendsen, *Proteins: Structure, Function, and Bioinformatics*, 1993, **17**, 412-425.
54. A. Amadei, A. Linssen, B. De Groot, D. Van Aalten and H. Berendsen, *Journal of Biomolecular Structure and Dynamics*, 1996, **13**, 615-625.

55. H. Yamaguchi, D. M. Van Aalten, M. Pinak, A. Furukawa and R. Osman, *Nucleic acids research*, 1998, **26**, 1939-1946.
56. W. Humphrey, A. Dalke and K. Schulten, *Journal of molecular graphics*, 1996, **14**, 33-38.
57. A. Bakan, L. M. Meireles and I. Bahar, *Bioinformatics*, 2011, **27**, 1575-1577.

Tables

Table 1. Temperature-Dependent Kinetics of BGL Catalyzed Hydrolysis of Rutin

| Temperature (K) | K_m (M) | V_{max} ($\mu\text{M/s}$) | K_{cat} (s^{-1}) | k_{cat}/K_m ($\text{M}^{-1}\text{s}^{-1}$) |
|-----------------|--------------------------------|-------------------------------|-------------------------------|--|
| 293K | $(4.7\pm 0.23)\times 10^{-5}$ | 0.15 ± 0.007 | 0.3 ± 0.015 | 638 ± 31.90 |
| 303K | $(5.1\pm 0.25)\times 10^{-4}$ | 0.4 ± 0.020 | 0.8 ± 0.040 | 1311 ± 65.55 |
| 313K | $(7.9\pm 0.39)\times 10^{-4}$ | 0.54 ± 0.027 | 1.08 ± 0.054 | 1367 ± 68.35 |
| 333K | $(1.27\pm 0.06)\times 10^{-3}$ | 1.5 ± 0.075 | 3.0 ± 0.150 | 2362 ± 118.10 |
| 353K | $(1.04\pm 0.05)\times 10^{-3}$ | 0.74 ± 0.037 | 1.4 ± 0.070 | 1423 ± 71.15 |

Table 2. Time-resolved Anisotropy of ANS in BGL at Different Temperatures

| Temperature(K) | r_0 | ϕ_1/ns | ϕ_2/ns | ϕ_3 |
|----------------|-------|-------------------------|-----------------------|----------|
| 278K | 0.37 | 0.400 ± 0.040 (21%) | 2.36 ± 0.20 (25%) | > 20 ns |
| 293K | 0.38 | 0.330 ± 0.030 (22%) | 2.14 ± 0.20 (23%) | > 20 ns |
| 303K | 0.38 | 0.206 ± 0.020 (29%) | 2.03 ± 0.15 (26%) | > 20 ns |
| 313K | 0.38 | 0.185 ± 0.015 (34%) | 1.61 ± 0.15 (27%) | > 20 ns |
| 333K | 0.387 | 0.140 ± 0.010 (28%) | 0.81 ± 0.08 (26%) | > 20 ns |
| 353K | 0.39 | 0.110 ± 0.010 (27%) | 0.54 ± 0.05 (33%) | > 20 ns |

Table 3. Distance of Center of Mass of the Protein (BGL) to the Center of Mass of the Ligand (Rutin)

| Temperature (K) | Distance (Å) |
|-----------------|--------------|
| 278 | 12.02 |
| 313 | 11.05 |
| 333 | 9.78 |
| 353 | 10.46 |

Table 4. Solvent Exposure of the Ligand (Rutin) in Terms of Interaction Energies with Water

| Temperature (K) | Energy (kcal/mol) |
|-----------------|-------------------|
| 278 | -35.47 |
| 313 | -31.70 |
| 333 | -31.20 |
| 353 | -53.35 |

FIGURES

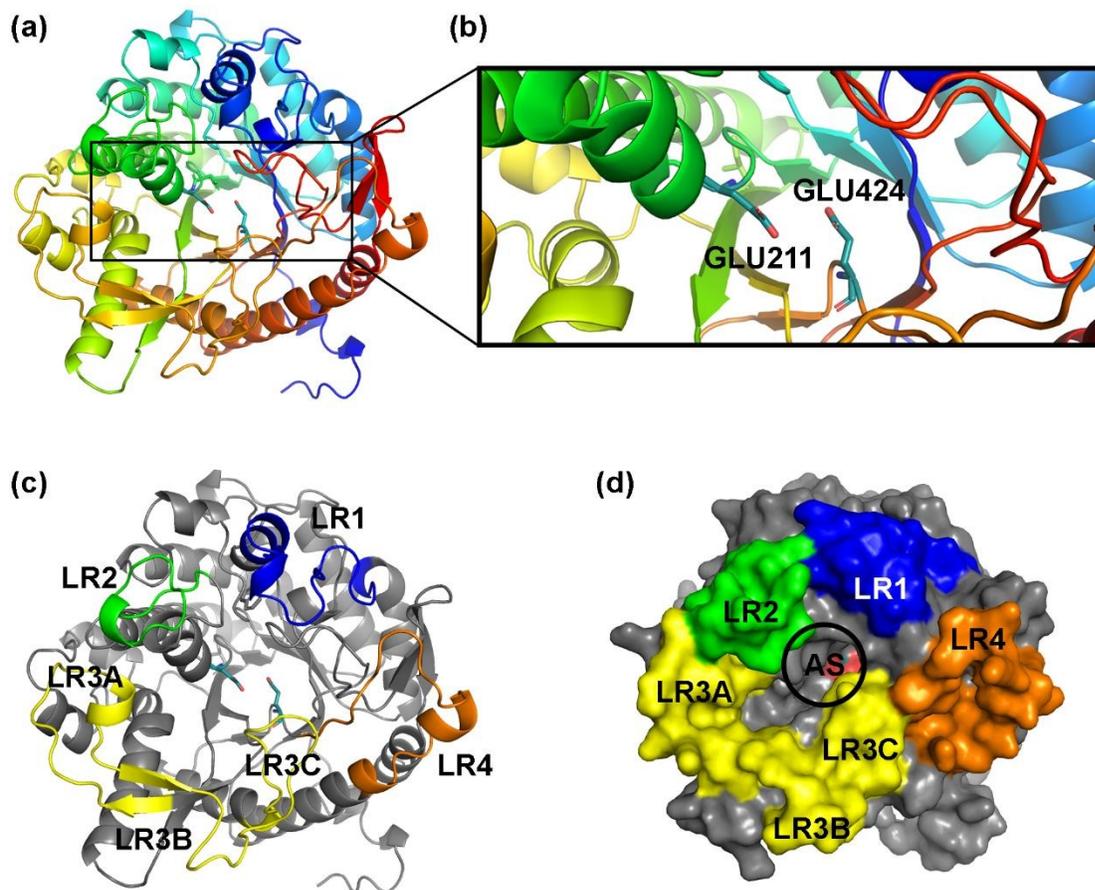


Figure 1 Schematic showing the build structure of almond Beta-Glucosidase (BGL). (a) The modelled 3D structure. (b) The stick representation showing the conserved active site Glutamate residues. (c) Showing the four loop region (LRs). The coloured regions represents each of the LRs, LR1 [Blue]; LR2 [Green]; LR3 [Yellow]; LR4 [Orange]. (d) Space filling model showing the strategic positioning of the LRs along the active site [AS], LRs represented by the same colour scheme.

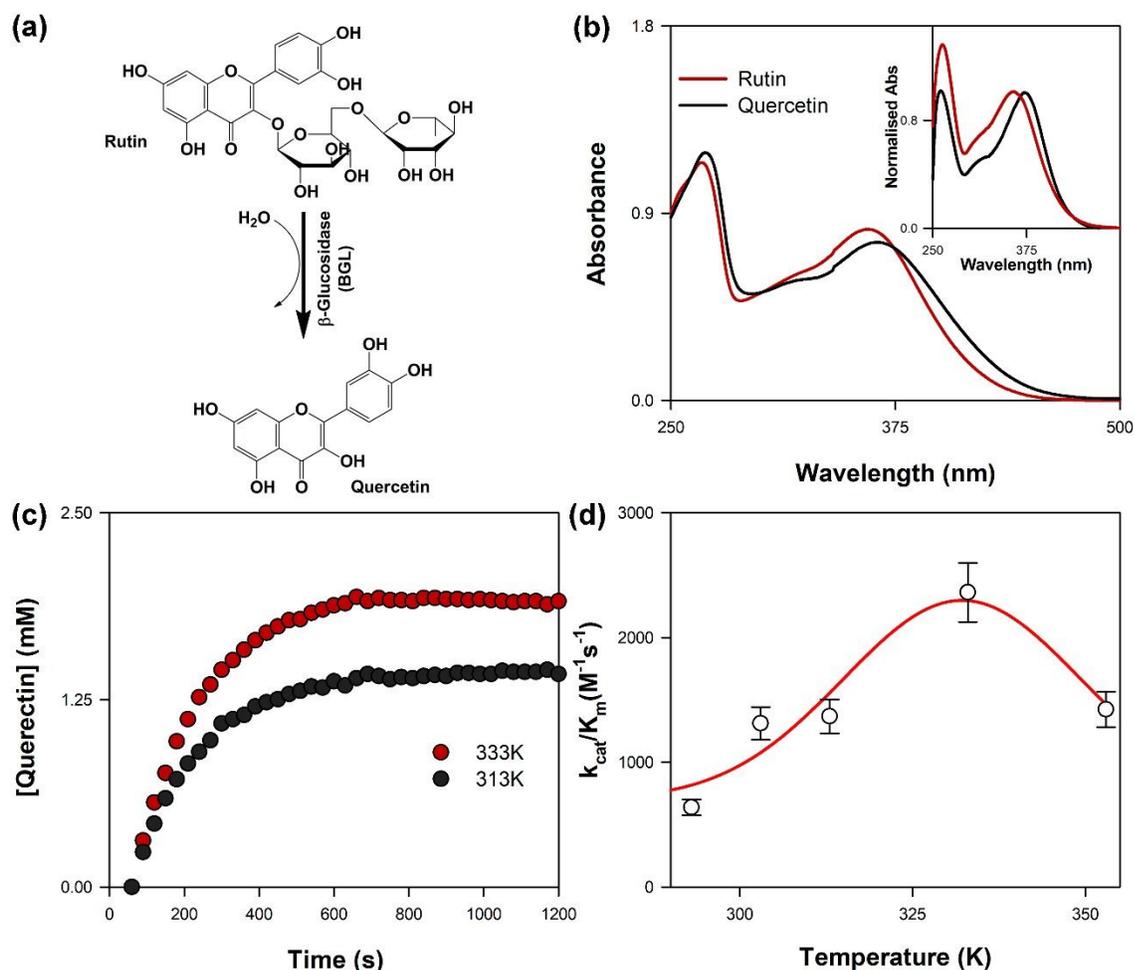


Figure 2 Schematic representation of temperature dependent enzyme catalysis. (a) Simple model outline of the reaction catalysed by BGL, the flavonoid glycoside Rutin hydrolysed to produce Quercetin. (b) Absorbance spectra of the substrate Rutin and product Quercetin during catalysis. The inset representing normalised absorption spectra of the same. (c) Product formation at different representative temperatures showing enhanced product formation at 333K. (d) Showing the effect of temperature on the catalytic efficiency (k_{cat}/K_m) of BGL. The enzyme showed enhancement of catalytic efficiency with increased temperature generating a bell shaped curve. Temperature optima was observed at 333K. Fall in efficiency was observed beyond 333K.

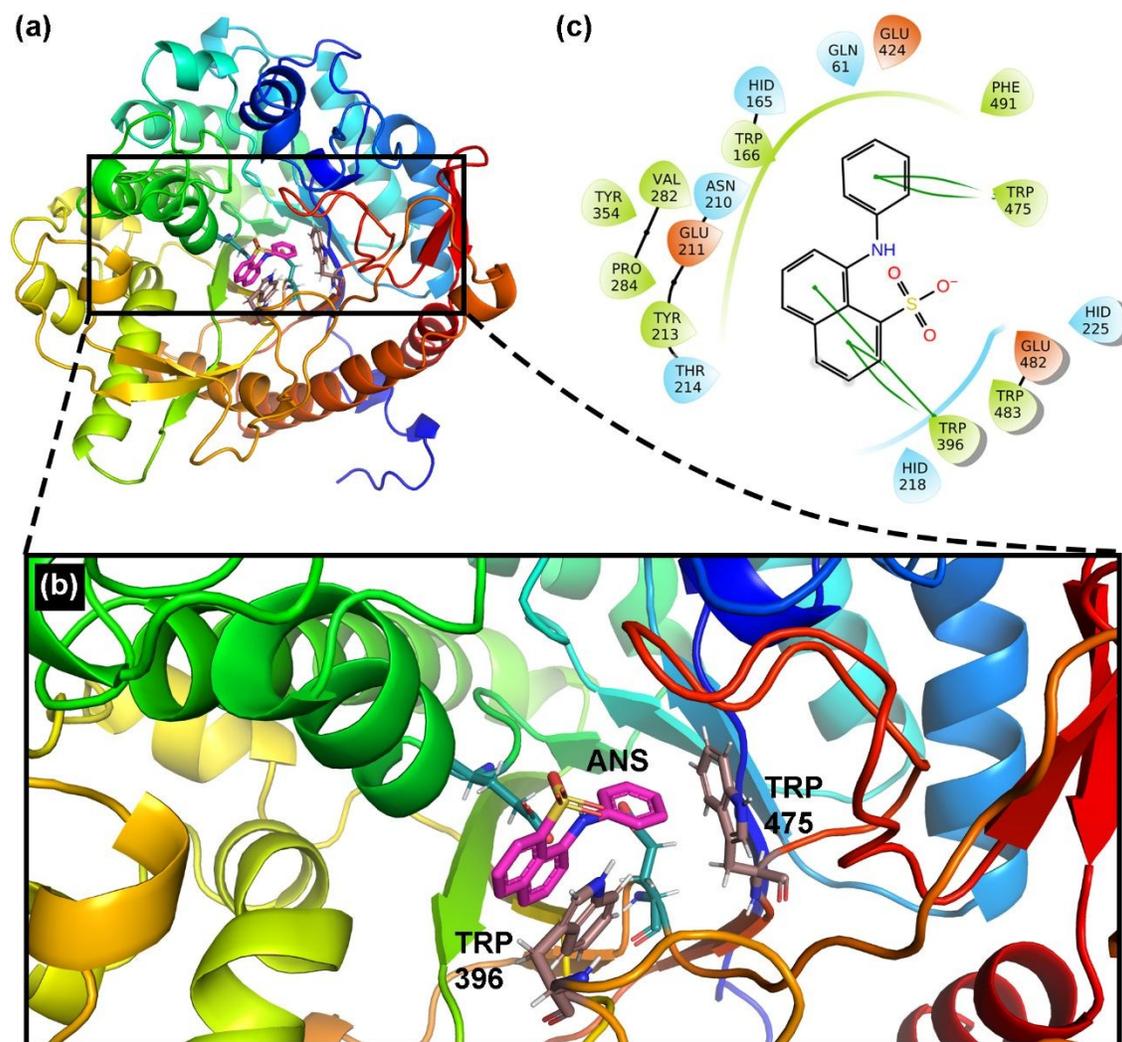


Figure 3 Ribbon representation of docked conformation of ANS-BGL complex. (a) Molecular docking representation of ANS-BGL complex, ANS [shown in pink] was found to be bound at the catalytic barrel of BGL. (b) ANS docked at the catalytic barrel. The fluorescence probe was found to interact with two Tryptophan residues (Trp396; Trp475). (c) Molecular interaction diagram between ANS and Tryptophan residues via π -stacking [shown in green lines] at the catalytic barrel.

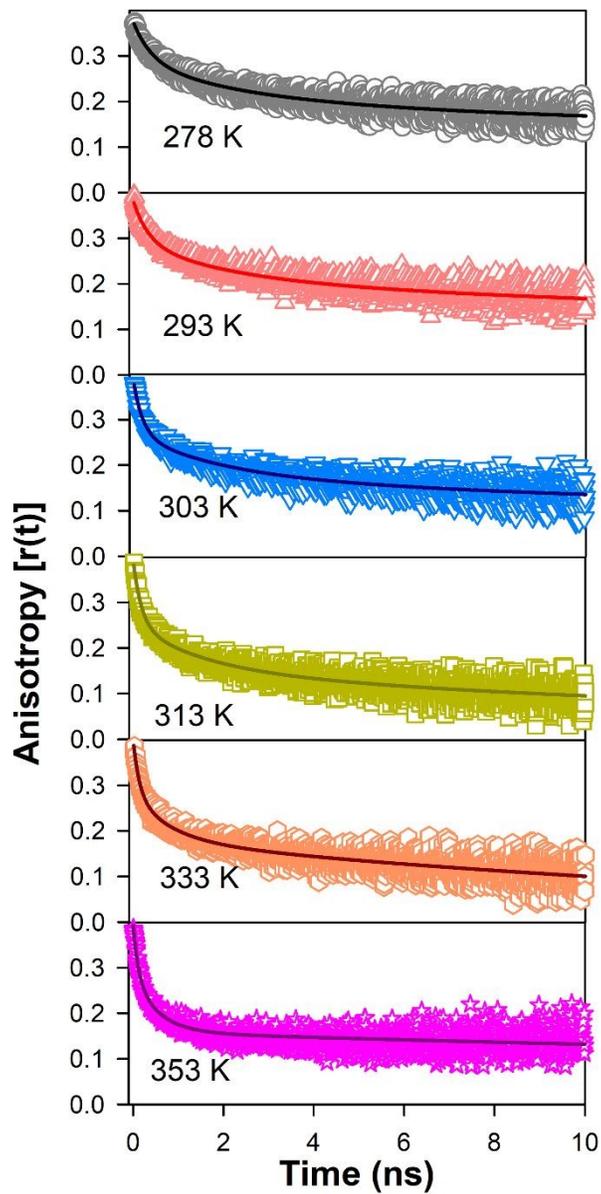


Figure 4. Temperature dependent anisotropy decay

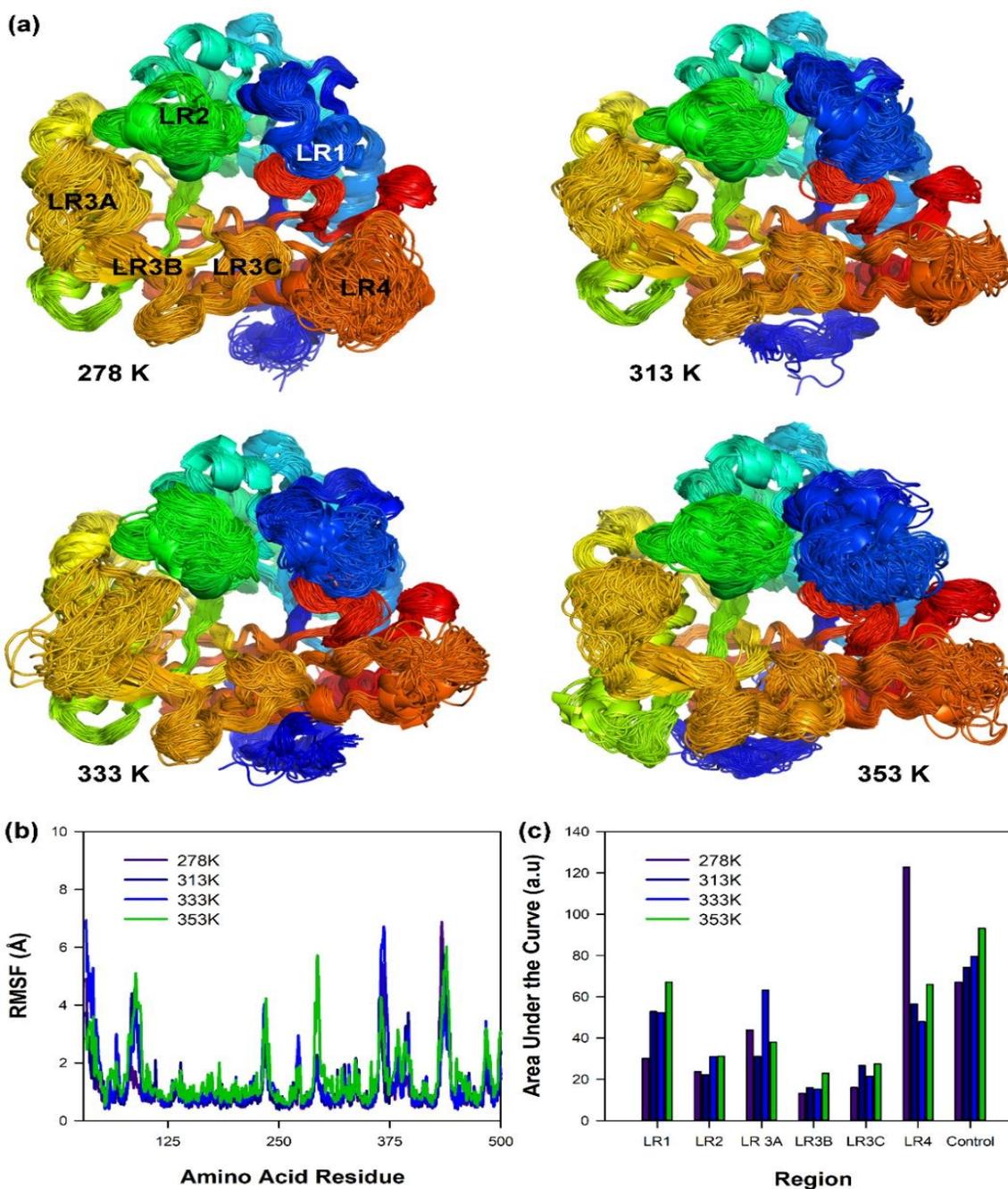


Figure 5. Showing the all atom fluctuations of almond BGL at different temperatures. (a) Ribbon representation of RMS fluctuation at different temperatures. The LRs [colour code maintained as Figure 1] are the regions of high fluctuation [RMSF > 4Å]. (b) All atom RMSF of BGL over 48ns of simulation event at different temperatures. The pattern shows atoms along the LRs as sharp peaks interspersed by low fluctuating atoms. (c) Calculated area under the curve for the LRs and Control at different temperatures. The corresponding amino acid residue for LR1 (77-100); LR2 (227-242); LR3A (359-376); LR3B (377-387); LR3C (388-402); LR4 (426-449); Control (132-222). [Control is a non-LR]

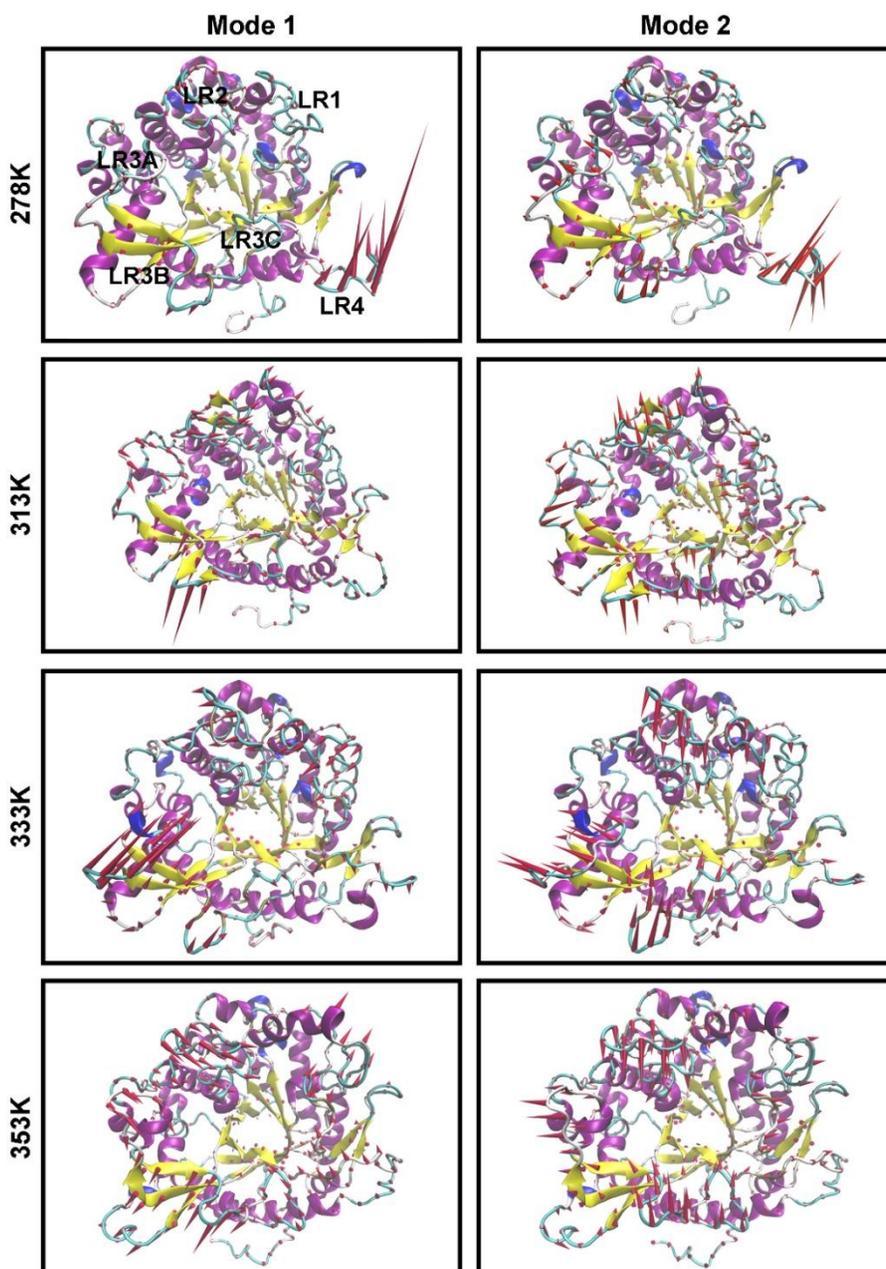


Figure 6. Porcupine plots of first two eigenvector [Mode 1 and Mode 2] of almond BGL at different temperatures. The model shown here represents all atom trace. The arrows attached to each represent the direction of eigenvector and the size of each arrow signifies the magnitude of the corresponding eigenvalue. The motion of eigenvectors show concerted motion of the LRs across 333K. The motion of eigenvectors become independent of each other and more random as temperature is raised to 353K.

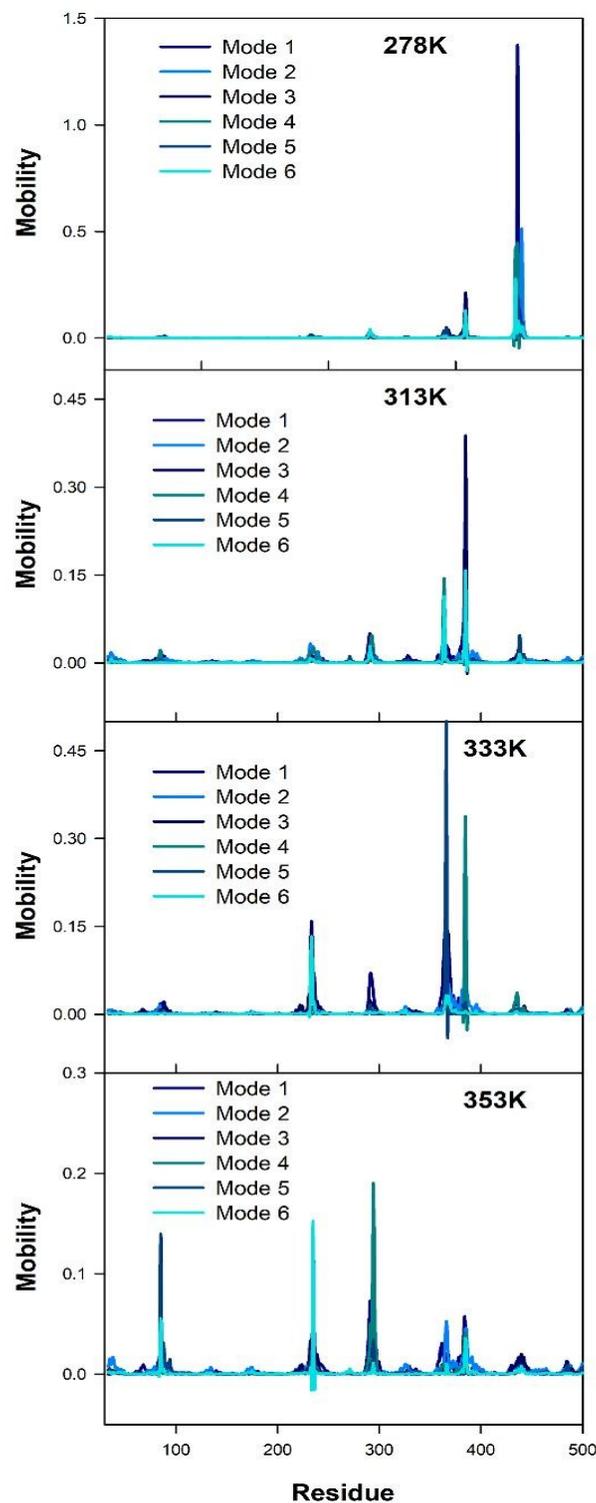
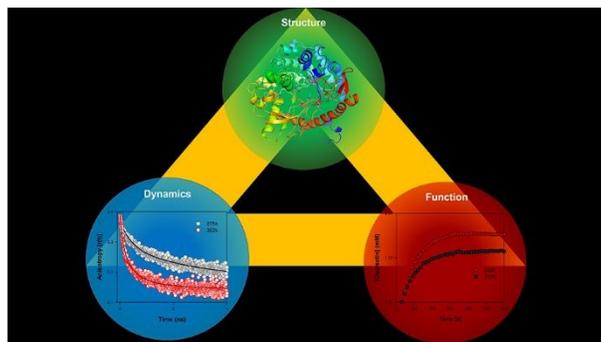


Figure 7. Residue based mobility plot of almond BGL at different temperatures. The plot shows the regions of major motions along the residues across a simulation event. The atoms along residues of LRs are indicated as the epicentres of major fluctuations. Residues along the catalytic barrel are found to be rigid and experience less fluctuations.

TOC



Temperature Dependent Catalytic Activity of Almond beta-Glucosidase Controlled by Concerted Motions of Loop Regions (LRs)