ROLE OF SUBSTRATE BINDING ON THE PROTEIN DYNAMICS OF AN ENDOGLUCANASE FROM *FUSARIUM OXYSPORUM* AT DIFFERENT TEMPERATURES

ABDUL AZIZ AHMAD[1], IBRAHIM ALI NOORBATCHA[1]
AND HAMZAH MOHD. SALLEH[2]

[1]*BioProcess and Molecular Engineering Research Unit (BPMERU),
Department of Biotechnology Engineering,*
[2]*International Institute for Halal Research and Training (INHART),
International Islamic University Malaysia,
P.O. Box 10, 50728 Kuala Lumpur, Malaysia.*

[*]*Corresponding authors:hamzah@iium.edu.my, ibrahiman@iium.edu.my*

ABSTRACT: Thermostability is an important requirement for protein function, and one goal of protein engineering is improvement of activity of the enzymes at higher temperatures, particularly for industrial applications. Computational approaches to investigate factors influencing thermostability of proteins are becoming researchers' choice. This study investigates the influence of substrate binding on the protein dynamics by comparing the molecular dynamics simulations of substrate-enzyme complex against un-bound enzyme, using endoglucanase I from *Fusarium oxysporum*. Endoglucanase-substrate complex was prepared by docking and molecular dynamics simulations were carried out at three different temperatures, 313 K, 333 K and 353 K. Our finding shows that the secondary structures for substrate-enzyme complex show more fluctuations relative to un-complexed structure. The same trend was observed for solvent accessible surface area and radius of gyration. At the highest temperature studied (353 K), the substrate-enzyme complex form showed the highest fluctuations. The fluctuations around the active site regions reach a minimum at the optimum temperature, compared to the other structural regions and other temperatures.

ABSTRAK: Kestabilan (ketahanan) terhadap haba merupakan keperluan yang penting untuk fungsi protin, salah satu matlamat kejuruteraan protin adalah penambahbaikan aktiviti enzim pada suhu yang tinggi khususnya untuk aplikasi industri. Kini para penyelidik memilih kaedah komputasi, bagi mengkaji faktor yang mempengaruhi kestabilan terhadap haba. Kajian ini menyelidik pengaruh ikatan substrat pada protin dengan membandingkan simulasi molekular dinamik diantara substrat-enzim kompleks dan enzim sahaja, menggunakan endoglucanase I dari *Fusarium oxysporum*. Kompleks endoglucanase-substrat disediakan melalui kaedah docking dan simulasi molekular dinamik dilakukan pada suhu 313 K, 333 K dan 353 K. Kajian kami menunjukkan struktur sekunder bagi substrat-enzim kompleks kurang stabil berbanding enzim sahaja. Pola yang sama bagi luas permukaan boleh dicapai pelarut (SASA) dan jejari gyrasi. Pada suhu tertinggi dikaji (353 K), substrat-enzim kompleks paling tidak stabil. Pada suhu optimum, kadar ubah-ubah sekitar amino asid aktif adalah minimum berbanding struktur dan suhu lain.

1. INTRODUCTION

Thermostable cellulases are highly sought after in the production of biofuels as the digestion requires thermostable enzymes to achieve economically feasible process. Thermostable cellulases allow the use of increased substrate concentration as the substrate viscosity decreases. This improves product yields, reduces capital and processing costs such as increased reaction rate, increased solubility of reactants and reduced contaminating microbial growth [1-3].

Thermostability of enzymes is temperature dependent phenomena. Computational approach to investigate factors influencing thermostability are becoming researchers tool as it enables them to understand the interactions at near and beyond melting temperatures at atomic detail. Currently, molecular dynamics (MD) simulation is the only method available to understand the dynamics of enzymes at higher temperature. Furthermore, in-silico method is relatively cost effective compared to wet-lab experiments.

Initially researchers used enzyme structures with no ligand complexed in MD simulations [4-8]. However, this is not the actualenvironment when the enzymes are in action. Nowadays, more researchers are simulating ligand-enzyme complexed structures to study the effects of complexation on the dynamical behavior [9-11] of the enzymes at different temperatures.

Here MD simulations wereused to investigate the impact of enzyme-substrate binding on thermostability. Cellotetraose was chosen to represent substrate based on our previous report [12]. Root mean square deviation (RMSD) of the whole protein, secondary structure, radius of gyration and solvent accessible surface area (SASA) were analyzed and discussed.

2. SYSTEM SETUP

An endoglucanase (EGuia) from *Fusarium oxysporum* available in our laboratory (originally provided by S.G. Withers, University of British Columbia, Vancouver, Canada) is used for modelling in this work. The molecular structure of EGuia is not known. However, EGuia is found to have 99.5% sequence identity with the crystal structure from protein data bank with (PDB ID: 3OVW) with two mutations at R41H and S216N. Hence the structure of EGuia was obtained using homology modelling by comparing with the available crystal structure from protein data bank [13]. The cellotetraose ligand was docked to the EGuia using BioMedCAChe 6.1 (Fujitsu) software [14]. Docking was performed as described in our previous report [12]. The docked structure of the enzyme is shown in Fig. 1. The two structures were denoted as EGuia for the enzyme with no ligand bound and EGuia-cellotetraosecomplex for cellotetraose bound enzyme. CHARMM-GUI website was used to generate protein structure file (PSF) (www.charmm-gui.org). Both systems were solvated using approximately 24,063 TIP3P water molecules in arectangular box of ($102\times90\times90$ Å); the box dimensions ensured that any protein atom was at least 20 Å away from the wall of the box.The net charge of the systems was balanced with one Cl⁻ ion to produce a neutral system.

All MD simulations (MDS) were performed with Nanoscale Molecular Dynamics (NAMD) software package (version 2.11) [15]. CHARMM 22 force field parameters for

protein and nucleic acids including CMAP corrections which contained in CHARMM forcefield (top_all27_prot_lipid.inp) obtained from http://www.ks.uiuc.edu/Research/namd was used for enzyme simulations and CHARMM36 force field parameters for enzyme-substrate complex simulations.
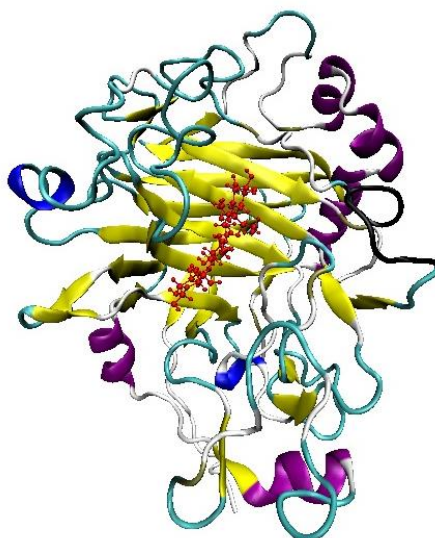

Fig. 1: Cartoon representation of EGuia based on secondary structure. The substrate (cellotetraose) is represented as CPK in red color.

Non-covalent van der Waals interactions were cut off at 12 Å. Particle mesh Ewald (PME) was used to calculate long range electronic interactions. The integration timestep is 1 fs. The velocity quenching (5000 time step) and conjugate gradient (50,000 timestep) algorithms were used to perform energy minimization. After energy minimization, MD simulation were performed for 35,000 time step (at 313 K), 40,000 (at 333 K and 353 K) to heat the system and then followed by equilibration for 200,000 time step (pressure 1 atm) with periodic boundary conditions. Langevin thermostat and barostat were used to control temperature and pressure.The attainment of minimum energy during energy minimization and the constant temperature during equilibration is verified by the asymptotic behavior of these properties. The equilibrated structures were subjected to 20 ns production runs in NPT (isothermal-isobaric ensemble). Structures were saved every 0.001 ns for analysis.

3. RESULTS AND DISCUSSION

The docking scores of cellobiose (a disaccharide) and cellotetraose (a tetrasaccharide) were evaluated. Cellotetraose showed stronger binding compared to cellobiose. The reasons for the stronger interactionson the binding characteristics are mainly due to hydrogen bonds. Five hydrogen bonds were observed between cellotetraose compared to three hydrogen bonds with cellobiose [12].

The docked EGuia-cellotetraose structure was used in the MD simulation. Root mean square deviation (RMSD) is used to evaluate structural deviations from the initial protein structure (i.e., to estimate the protein intactness) [16]. The stability of the protein relative to its conformation can be determined by the deviations produced during the course of its simulation. The smaller the deviations, the more stable the protein structure [17]. The behavior of the enzyme and the complex was monitored at three different temperatures.

RMSD of backbone atoms (C, N, C_α) from the initial structure (the structure after equilibration step) were calculated for the entire simulation trajectory at 313 K, 333 K and 353 K. The RMSD of backbone atoms are illustrated in Fig. 2. In general, the temperature dependence of the fluctuations can be expected to represent the stability of overall structure when the temperature is increased. It is found that the RMSD of backbone atoms of the enzyme-substrate complex is higher than EGuia which indicates that the structure is becoming less stable at higher temperatures. At high temperature, the RMSD has tendency to diverge to larger values for the complex as well as EGuia.

While the RMSD of complex at 313 K was higher than 333 K, it remains almost constant after 3ns with slight variations. After 8 ns, RMSD of complex at 333 K, increases dramatically from 1.6 Å to 2.5 Å at 20 ns. RMSD of complex at 353 K was lower than 313 K and 333 K for the initial 4ns, but was the highest until 18 ns before reaching the same RMSD value with 333 K complex.
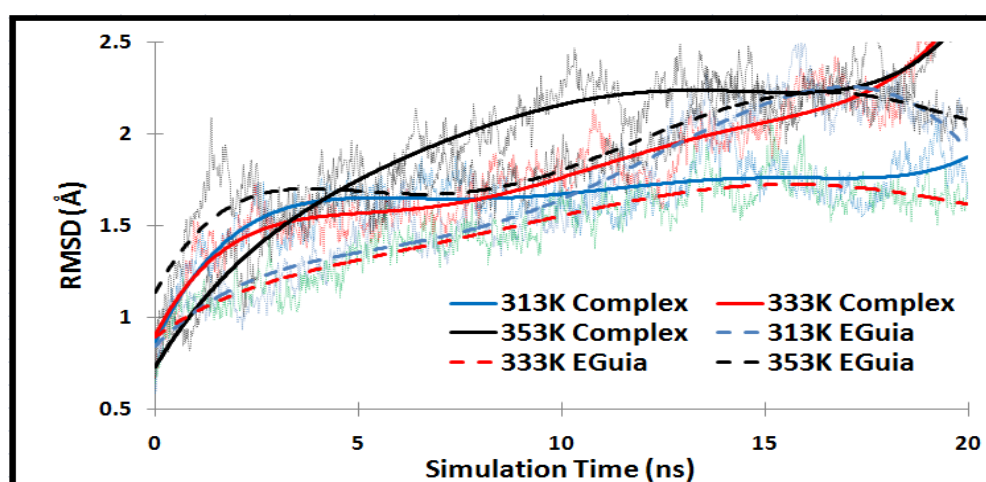

Fig. 2: Backbone RMSD of EGuia and enzyme-substrate complex plotted as a function of MD time at 313 K, 333 K and 353 K fitted with trendline of polynomial degree 5. Black solid (enzyme-substrate complex at 353 K), black dash (EGuia at 353 K), red solid (enzyme-substrate complex at 333 K), red dash (EGuia at 333 K), blue solid (enzyme-substrate complex at 313 K), and blue dash (EGuia at 313 K).

The changes in the backbone RMSD values are expected to affect the structure of the active site and hence the activity of the enzyme. The rapid divergence of RMSD of enzyme-substrate complex at higher temperature compared to EGuia under same conditions seems to suggest that the enzyme-substrate complex would lose its activity at higher temperatures, much more easily compared to EGuia. The backbone RMSD of EGuia at 333 K was the least among the three temperatures studied here implying that the structure at this temperature deviate less from the corresponding equilibrium structure thus maintaining the integrity of the active site structure and hence the activity. It should be noted that the optimum temperature for the activity of this enzyme is around 333 K. However, the RMSD of the enzyme-substrate complex does not show such a trend and the fluctuations increases as the temperature is increased. Perhaps the RMSD of the complete backbone of the enzyme-substrate complex may not be a good indicator of the thermostability of the enzymes, as the temperature response of the active site region is more crucial rather than the whole enzyme structure in determining the thermostability. Hence for better understanding of the thermostability it is necessary to analyze the RMSD of the secondary structure components at different temperatures.

RMSD of secondary structure components are shown in Fig. 3. For all helices, the RMSD was around 1.5 Å for all except for enzyme-substrate complex at 353 K which increased from 1.5 Å to 2.3 Å at 7.5 ns to 12.5 ns. However, for alpha helices a different phenomenon was observed, the enzyme (with no ligand) showed higher RMSD. The 333 K EGuia was the highest for the entire simulation. Complex at 353 K was the lowest. The RMSD of 3_10 helices was more violent at 353 K. It is interesting to note that at 333 K enzyme and complex of 3_10 helices was maintained at zero. Relative to any other individual secondary structure element, influence of one helix is more dominant on the stability and organization of a protein [18]. Thirteen parts of bridge beta showed a different behavior compared to extended beta-sheet as illustrated in Fig. 3.
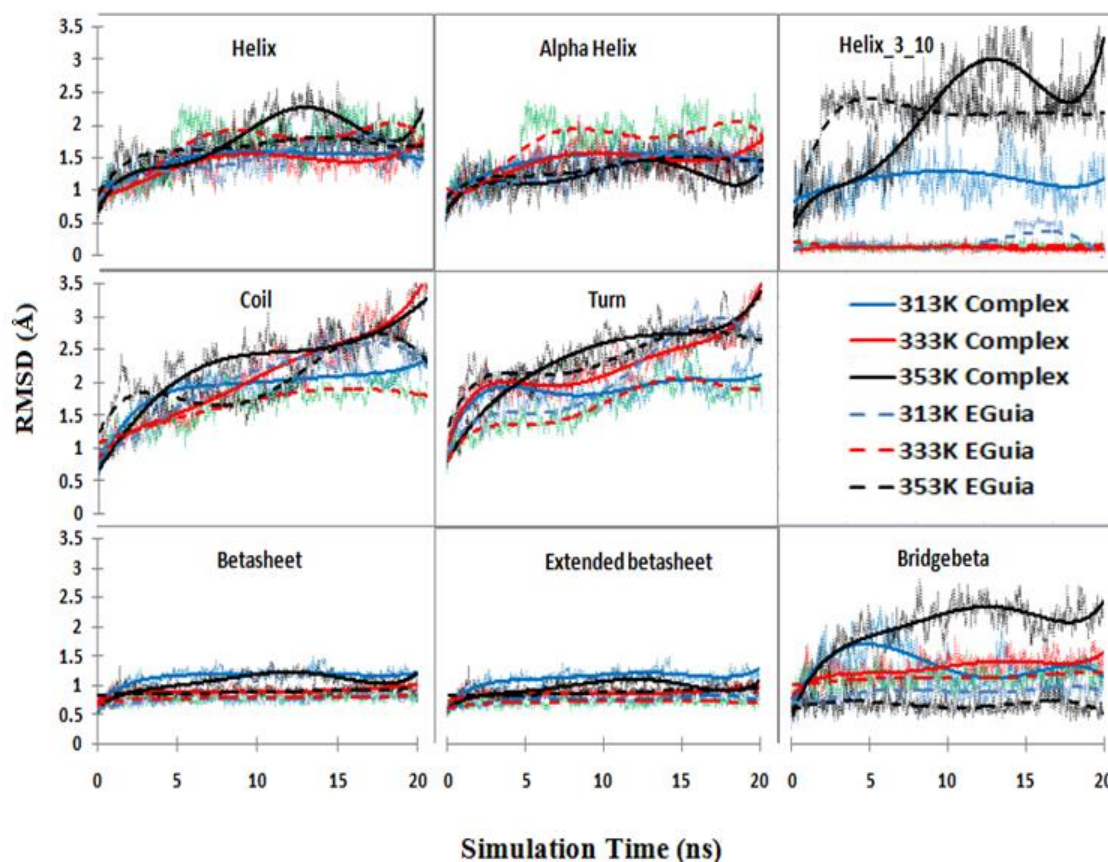

Fig. 3: RMSD time evolutions of secondary structure's backbone atoms (C,Cα and N). Reference frame is the initial frame, t_o. Black solid (enzyme-substrate complex at 353 K), black dash (EGuia at 353 K), red solid (enzyme-substrate complex at 333 K), red dash (EGuia at 333 K), blue solid (enzyme-substrate complex at 313 K) and blue dash (EGuia at 313 K).

In general, betasheets of EGuia are more stable than helices. Coil and turn shows a similar pattern of RMSD for the entire 20 ns. As expected, coil and turns showed higher RMSD due to flexible nature. This finding is consistent with our previous report [6]. In reality, coil is not a true secondary structure; it is the class of conformations that show absence of regular secondary structure. Relationship between stability and activity of enzymes is maintained by underlying conformational flexibility. In thermophilic enzymes, a decrease in flexibility causes low enzyme activity. It can be seen that even though the

RMSD changes differently for different secondary structure components, for beta sheets (excluding bridge beta sheets) the fluctuations at 333 K for the both the un-bound enzyme and the enzyme-susbtrate complex is the lowest among the fluctuations at other temperatures. As shown in Fig. 1, the substrate is surrounded by beta sheets and the temperature at which the beta sheets structures are maintained will preserve the active site. The highest temperature at which the active site is retained will be the optimum temperature for the enzyme activity. The current simulations reveal that even though the substrate binding on the enzyme can influence the dynamics of the different parts of the enzymes differently, the fluctuations around the active site (beta sheets in this case) around the optimum temperature is distinctly maintained thus offering an avenue for further exploration for enhancing the thermostability of the enzymes.

The accessible surface of a molecule is the part of the molecular surface that is exposed to the solvent [19]. The solvent accessible surface areas (SASA) of atoms are related to hydrophobicity and folding process of proteins. The folding process is usually accompanied by a significant decrease in SASA value [20]. SASA is higher for enzyme-susbtrate complex than EGuia (enzyme only) at the corresponding temperatures (Fig.4). The SASA for both the enzyme and the enzyme-substrate complex are larger at higher temperatures, implying the presence of the substrate at the active site do not change the temperature dependence of SASA. At 313 K, SASA of EGuia is the least although in the increasing trend after 10 ns. It is interesting to note that at the end of simulation (20 ns), EGuia and complex had the same SASA. Except at 313 K, the SASA at 353 K and 333 K increased sharply after 15 ns. This may be an indication that the enzyme starts to unfold at these temperatures.
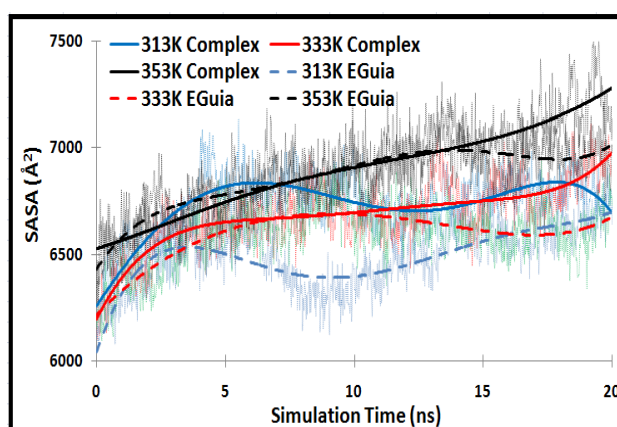

Fig. 4: SASA of Complex and EGuia. Black solid (enzyme-substrate complex at 353 K), black dash (EGuia at 353 K), red solid (enzyme-substrate complex at 333 K), red dash (EGuia at 333 K), blue solid (enzyme-substrate complex at 313 K) and blue dash (EGuia at 313 K).

Radius of gyration (Rg) were calculated for the 20 ns simulation time and depicted in Fig. 5. In general, all the structures maintained its Rg between 20.5 to 21.0 Å except for enzyme-substrate complex at 353 K which showed a steady increment beyond 21.0 Å. Radius of gyration is an indicator of protein structure compactness [21]. EGuia and enzyme-substrate complex maintained its structure at all temperatures studied except for enzyme-substrate complex at 353 K, at which the compactness has decreased (increase in Rg).
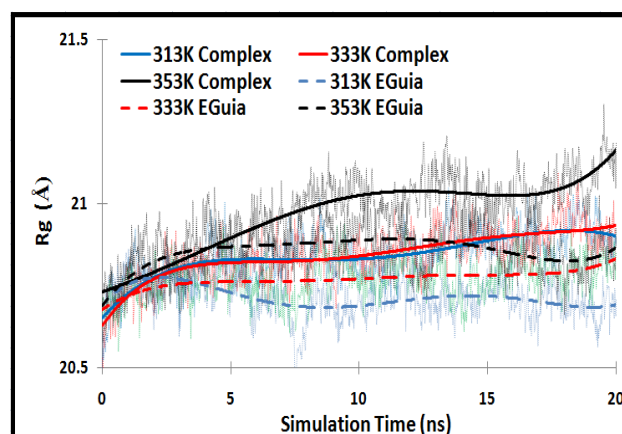
Fig. 5: Radius of gyration of complex and EGuia.
Black solid (enzyme-substrate complex at 353 K), black dash (EGuia at 353 K),
red solid (enzyme-substrate complex at 333 K), red dash (EGuia at 333 K),
blue solid (enzyme-substrate complex at 313 K) and blue dash (EGuia at 313 K).

4. CONCLUSION

MD simulations are carried out at three different temperatures to study the impact of ligand binding on the thermal stability of an endoglucanase, EGuia. It is found that even though the fluctuation of the overall structure of the enzyme and the enzyme-substrate complex changes at higher temperatures, there is a tendency to maintain the compactness with minimum RMSD around the active site region of the enzyme. This remarkable ability of the active site region determines the optimum temperature for the enzyme activity of endoglucanase. The simulation findings should be validated experimentally. However, care should be taken in extending this argument for other enzymes. For enzymes with flexible active site, such as protein kinases [22], a much longer simulation is necessary to confirm this trend.

ACKNOWLEDGEMENT

REFERENCES

[1]   Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S,Moore JC,Robins K. (2012) Engineering the third wave of biocatalysis.Nature, 485:185-194.
[2]   Zamost BL, Nielsen HK, Starnes RL. (1991) Thermostableenzymes forindustrial applications. Journal of Industrial Microbiology, 8:71-81.
[3]   Kim HS, LeQAT, Kim YH.(2010) Development of thermostable lipase B from *Candida antarctica* (CalB) through in silico design employing B-factor andRosettaDesign. Enzyme and Microbial Technology, 47(1-2):1-5.
[4]   Noorbatcha IA, Waesoho S, Salleh HM. (2012) Structural and dynamics behavior of native endoglucanase from *Fusarium oxysporum*. Australian Journal of Basic and Applied Sciences, 6(1):89-92.
[5]   Akcapinar GB, Venturini A, Pier Luigi Martelli PL, Casadio R, Sezerman UO. (2015) Modulating the thermostability of endoglucanase I fromTrichoderma reesei using computational approaches. Protein Engineering, Design and Selection, 28(5):127-135.

[6]   Noorbatcha IA, Khan AM, Salleh HM. (2010) Molecular dynamics studies of human β-glucuronidase. American Journal of Applied Sciences, 7(6):823-828.

[7]   Purmonen M,Valjakka J,Takkinen K, Laitinen T, Rouvinen J. (2007) Molecular dynamics studies on the thermostability of family 11 xylanases. Protein Engineering, Design and Selection, 20(11):551-559.

[8]   Karjiban RA, Rahman MBA, Basri M, Salleh AB, Jacobs D, WahabHA. (2009) Molecular dynamics study of the structure, flexibility and dynamics of thermostable L1 lipase at high temperatures. Protein J.*,* 28:14-23.

[9]   Zhang PP, Zhao L, Long SY, Tian P. (2015) The effect of ligands on the thermal stability of sulfotransferases: a molecular dynamics simulation study. J. Molecular Modeling, 21:1-7

[10]   Maximilian CJCE, Espinola JG, Lamoureux G,Pelletier JN. (2017)Substrate-specific screening for mutational hotspots using biased molecular dynamics simulations. ACS Catalysis*,* 7: 6786-6797.

[11]   Azam MA, Jupudi S. (2017) Insight into the structural requirements of thiophene-3-carbonitriles-based MurF inhibitors by 3D-QSAR:Molecular docking and molecular dynamics study. J. Receptors and Signal Transduction,37(5):522-534.

[12]   Noorbatcha IA, Ahmad AA, Salleh HM. (2014)Molecular docking of cellobiose and cellotetraose into an endoglucanase I from *Fusarium oxysporum*. J.Pure and Applied Microbiology, 8:791-796.

[13]   Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H,Shindyalov IN,Bourne PE. (2000) The protein data bank . Nucleic Acids Research, 28: 235-242.

[14]   CACheWorkSystem Pro Version 7.5.0.85, Copyright © 2000-2006 Fujitsu Limited, © 1989-2000 Oxford Molecular Ltd.

[15]   JamesCP, Rosemary B, Wei W, James G, Emad T, Elizabeth V, ChristopherC, RobertDS, Laxmikant K, Klaus S. (2005) Scalable molecular dynamicswith NAMD.J. Comp. Chem., 26:1781-1802.

[16]   Monhemi H, Housaindokht MR, Nakhaei PA. (2015) Effects of natural osmolytes on the protein structure in supercritical CO2: molecular level evidence, J. Chem. Phys. B 119:10406–10416.

[17]   Imlimaong A, Pritish KV & Utkarsh R. (2016) Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. Scientific Reports. 6:34984

[18]   Richardson J S. (1981) The anatomy and taxonomy of protein structure. [http://kinemage.biochem.duke.edu/teaching/anatax/html/anatax.2a.html]

[19]   Shucheng L , Yuan L, Shuai L, Andi D, Mengna L, Hongwu J, Jing G, Jiming H (2017) Molecular dynamics simulation of the interaction between dense-phase carbon dioxide and the myosin heavy chain. Journal of CO_2 Utilization, 21:270–279

[20]   Norhayati Y, Nor HAK, Adam TCL, Abu BS, Raja NZAR and Mohd SMA. (2017) The Role of Solvent-Accessible Leu-208 of Cold-Active *Pseudomonas fluorescens* Strain AMS8 Lipase in Interfacial Activation, Substrate Accessibility and Low-Molecular Weight Esterification in the Presence of Toluene. Molecules 2017, 22(8)

[21]   Lobanov MY, Bogatyreva NS, Galzitskaya OV. (2008) Radius of gyration as an indicator of protein structure compactness.Molecular Biology. 42(4):623-628.

[22]   Meharena HS, Fan X, Ahuja LG, Keshwani MM, McClendon CL, Chen AM, Adams JA, Taylor SS, et al. (2016) Decoding the interactions regulating the active state mechanics of eukaryotic protein kinases. PloSBiology 14(11):e2000127