

Molecular Dynamics Studies of Human β -Glucuronidase

Ibrahim Ali Noorbata, Ayesha Masrur Khan and Hamzah Mohd Salleh
Biomolecular Engineering Research Unit, Department of Biotechnology Engineering,
Faculty of Engineering, IIUM, Kuala Lumpur, Malaysia

Abstract: Problem statement: The enzyme β -glucuronidase is being used as a reporter molecule in the area of genetic engineering, as a component of prodrug therapy in cancer treatment and in the scouring process of cotton fabrics. However, a detailed understanding of the factors responsible for the stability and the activity of this enzyme is still not available. Molecular Dynamics (MD) simulations provide an estimate of equilibrium and dynamic properties of enzyme systems that cannot be calculated analytically. With this perspective, molecular dynamics simulations of human β -glucuronidase (GUS) have been carried out to determine the behavior of this enzyme in vacuum and solvent environments at a defined temperature. **Approach:** CHARMM force field along with distance dependent dielectric model was used to represent the solvent environment in the MD simulations. The parameters employed in various stages of MD simulations had been selected based on repeated trials under various conditions as a method of choosing the optimum parameters for each stage. **Results:** It was found that simulations in vacuum caused the backbone of GUS to have smaller fluctuations from their mean values compared with the fluctuation in implicit solvent simulations, due to the fact that vacuum environment did not provide for the electrostatic interactions affecting the backbone of GUS that may otherwise exist in a solvent environment. **Conclusion:** Inclusion of solvent effects in MD simulations is crucial in understanding structural flexibility and stability of β -glucuronidase. Implicit solvent method can provide a realistic inclusion of backbone flexibility and structural compactness of GUS, which will have profound influence on the stability and activity of the enzymes, with a marginal increase in computational time.

Key words: β -glucuronidase (GUS), Molecular Dynamics (MD), implicit solvent model, backbone flexibility

INTRODUCTION

Human β -glucuronidase belongs to family 2 glycosyl hydrolases that cleave β -D-glucuronic acid residues from the non reducing termini of glycosaminoglycans. In humans, the lysosomal β -glucuronidase degrades sulfated glycosaminoglycans. The X-ray structure of a human β -glucuronidase has been determined, which shows it to be a tetramer (Jain *et al.*, 1996), with each monomer 653 amino acid residues long and a size of 80 kDa. It was revealed that each monomer of the human GUS structure has three structural domains, namely the jelly roll barrel, immunoglobulin constant domain and a Triosephosphateisomerase (TIM) barrel, in addition to it having an overall fold similarity to the N-terminal region of *E. coli* β -galactosidase. The active site of human β -glucuronidase has been based on the experimentally determined active site in *E. coli* β -

galactosidase and corresponds to the amino acid residues Glu451, Glu540 and Tyr504. The nature of the active site residues have been discussed previously (Islam *et al.*, 1999).

Although significant experimental work has been reported on human GUS enzymes, computational studies of human GUS have not emerged as yet. To carry out *in silico* studies on an enzyme, the tool of Molecular Dynamics has been widely used in the past to efficiently explore the stability and related biological activity of biomolecules. In the current work, Molecular Dynamics simulation has been applied to explore the structural changes in human β -glucuronidase in two different environments, i.e., vacuum and implicit solvent. Prior to commencing the first stage of MD simulation, short term simulation trials were carried out using a "standard dynamics cascade" method, whereby parameters for each stage were defined in the beginning and each simulation stage ran successively after the

other as one complete process. This method allowed for the selection of optimum parameters for long term simulation.

MATERIALS AND METHODS

All simulations reported here were carried out using Discovery Studio© v2.1.0.8130, built on the SciTegic Enterprise Server platform (Accelrys Inc, 2008). Snapshots were taken by using VMD software, version 1.8.6 (Humphrey *et al.*, 1996).

System Preparation: In order to prepare a sample for MD simulation, a model system consisting of N particles is selected and Newton's equations of motion are solved for this system until the properties of the system reach a stable value, or more appropriately reach equilibration. Afterwards the actual measurement can be carried out, i.e., production phase. The crystal structure of Human GUS from PDB (PDB ID: BHGA) was prepared for simulation by selecting only one chain and removing the heteroatoms. Before applying the force field, missing atoms were added in case of incomplete side-chain or backbone, hydrogen atoms were added and connectivity and bond orders were corrected. The Chemistry at HARvard Macromolecular Mechanics (CHARMM) program was used for performing all the molecular dynamics simulations. It performs standard molecular dynamics in many different ensembles, long-range force calculations and employs various implicit and explicit solvation models. All structures were typed with CHARMM force field (Brooks *et al.*, 1983) with the Momany-Rone partial charge estimation method (Momany and Rone, 1992). The total charge of each human GUS after applying this method was -6, whereas the pH was 7. The simulation environment depends on the solvent used for the enzyme and one can use explicit solvent, implicit solvent or no solvent, i.e. vacuum environment. Choosing explicit solvent increases the degrees of freedom in the system and therefore can incur a significant computational cost. An approach which can approximate the effects of solvation without explicitly representing solvent atoms is the implicit solvent model approach in which the polar and non polar effects of the solvent are accounted for in an averaged way. The implicit solvent method used in this study is the distance dependent dielectrics method, in which the dielectric constant, ϵ , is no longer a constant and increases linearly with the distance between two solute atoms, or, $\epsilon = \epsilon_0 r_{ij}$. When the distance r_{ij} is large, the effective dielectric constant is large and the electrostatic interaction between two non-solvent atoms decreases.

This approach though is a rough estimation but gives relatively good results even for complex systems (Becker and Karplus, 2006). We also applied vacuum environment for the system, in order to have two sets of simulation results which can give insights into the differing behavior of the enzyme in two different environments.

Minimization: The starting structure had a very high RMS gradient ($>4000 \text{ kcal mol}^{-1}$). Thus, 2000 steps of steepest descent algorithm were used to rapidly improve a very poor conformation. In this method, the gradient of the potential is calculated at each step. The aim of applying steepest descent is not to reach a local energy minimum, but to reach a considerably stable conformation. After obtaining considerable stability, 10,000 steps of a second minimization algorithm, the conjugated gradients method was applied until the RMS gradient of the structure reached below $0.0001 \text{ kcal mol}^{-1}$.

Since a direct evaluation of all the non-bonded interactions involving all atom pairs is highly complex, a suitable cut-off scheme is employed to reduce this cost-intensive task, without destroying the value of simulation of biomolecules. In this respect the spherical cut-off scheme yields reasonable approximations to the energy and force functions up to some threshold separation-distance value, as given by the non-bond list radius. In this method, a non-bonded list is created, which includes all pair interactions, i.e., van der Waals and electrostatic interactions. During energy minimization or MD simulation, the non-bonded list is updated according to a heuristic test, based on the defined non-bond list radius and non-bond higher cutoff distance parameter values, performed at each step of minimization or MD simulation. When applying the non-bonded cut-off scheme, the efficiency of the non-bonded calculation is increased by including only atom pairs that are closer than the cutoff distance in a non-bonded list. For the simulations in vacuum the regular coulombic electrostatic interaction terms were applied, having a uniform dielectric constant of $\epsilon = 1$. As for implicit solvent the dielectric constant was also chosen to be the same. The non-bond list radius of 14 \AA is used for counting non-bonded interaction pairs. Beyond this cut-off distance, non-bonded interactions between atom pairs are ignored. The same distance criteria are used in subsequent simulations.

Heating: The structure was heated to the desired temperature as part of the preparation procedure. The target temperature for production step is 300 K therefore heating phase was applied at 300 K as well.

During the heating stage MD simulations were carried out for 6000 steps, with a step size of 0.001 ps. The initial temperature for this heating process is 50 K and the final target temperature is 300 K and the velocities are adjusted in every 50 steps.

Equilibration: Prior to the production stage, the system was equilibrated by allowing it to evolve spontaneously for a period of time and integrating the equations of motion until the average temperature and structure remained stable and the total energy converged. This was facilitated by periodically reassigning velocities appropriate to the desired temperatures, which in this study was 300 K. As given in Table 1, the total number of steps to perform the dynamics simulation was 50000 with a step size of 0.0005 ps. The Leapfrog Verlet integrating algorithm was used to perform numerical integration of the equations of motion, which follows the following relation:

$$v(t) = \frac{1}{2} \left[v \left(t + \frac{1}{2} \delta t \right) + v \left(t - \frac{1}{2} \delta t \right) \right] \quad (1)$$

The equilibration phase was validated by ensuring that the total energy reached a stable average value for all the structures. Analyzing the global energetic properties, such as temperature, total energy and total potential energy as a function of simulation time is important to validate the simulation protocol. The structures were also visually inspected for any unusual differences in secondary structure compositions between the initial and final structures. This is to guarantee that none of the pre-production steps have disturbed the original structural integrity of the proteins prior to submitting the structures for production.

Table 1: Parameters employed in equilibration step

| Parameter name | Parameter value |
|--------------------------------------|---|
| Steps | 50000 |
| Time step (ps) | 0.0005 |
| Target temperature (K) | 300 |
| Adjust velocity frequency | 50 |
| Save results frequency | 100 |
| Constant pressure | False |
| Implicit solvent model | Vacuum and distance-dependent dielectrics |
| Dielectric constant | 1 |
| Non bond list radius (Å) | 14 |
| Non bond higher cut off distance (Å) | 12 |
| Non bond lower cut off distance (Å) | 10 |
| Electrostatics | Spherical cut off |
| Dynamics integrator | Leapfrog verlet |

Production run: The equilibrated system was taken as the starting structure for the long term molecular dynamics simulation, called production. In this final step CHARMM takes the equilibrated structure as its starting point. In this study, the total simulation time of 2 ns was achieved in four cycles of 0.5 ns. This method ensured that the ensemble structures were of a reasonable size, thus enabling efficient analysis. The production runs were carried out on a computer workstation with Intel-Xeon Dual-Quad processors.

The production results were saved for every 500th step, which corresponds to every 0.5 ps. The time step used for integrating the equations of motions was 0.001 ps. Too small a time step value covers only a limited proportion of the phase space and too large a time step may cause instabilities in the integrating algorithm due to high energy overlaps between atoms. The canonical production ensemble was used for the production simulation, whereby the number of atoms, volume and temperature is constant, better known as the NVT ensemble. It is better suited for this project as it controls the temperature during the simulation, by maintaining the temperature according to the Berendsen *et al.* (1984) weak coupling method.

RESULTS AND DISCUSSION

Backbone flexibility of GUS: The Root Mean Square Deviation (RMSD) from the initial structure of the Human GUS enzyme was calculated for the entire simulation trajectory in vacuum and implicit solvent at 300 K. In the analysis of stability of backbone atoms of GUS, we find that the structure in vacuum exhibited fewer fluctuations and the extent of fluctuations is smaller compared to implicit solvent environment (Fig. 1). The larger fluctuations observed in implicit solvent can be attributed to the response of the protein backbone with the solvent environment.

Secondary structures: The stability of helical residues of GUS was monitored in the trajectory (Fig. 2). We can find that the helical regions showed less deviation from the initial structure in vacuum compared to implicit environment. The RMSD of helical residues of GUS in implicit solvent showed a fluctuating trend and reached a value of 2.3Å at the end of production, whereas in vacuum it was only 1.4Å. This trend is very similar to the behavior of the backbone fluctuations observed in the previous section.

The simulation snapshots of human GUS in vacuum (Fig. 3) and implicit solvent (Fig. 4), reveal that the overall domain structure of GUS remains the same in both the environments, even though increased flexibility as inferred from the backbone RMSD and radius of gyration plots is observed here also.

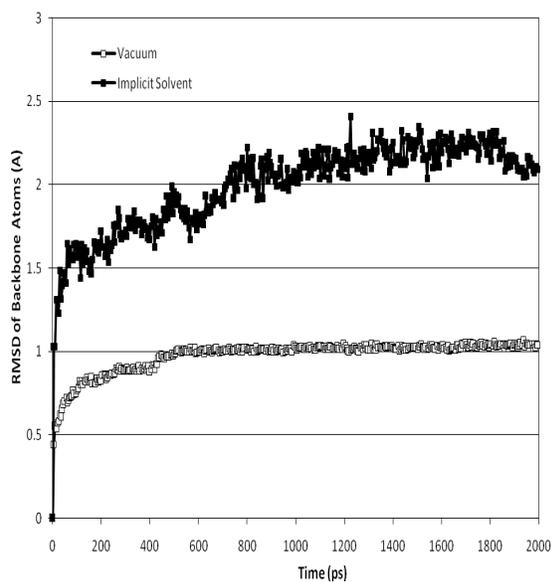


Fig. 1: RMSD of GUS backbone in vacuum and implicit solvent

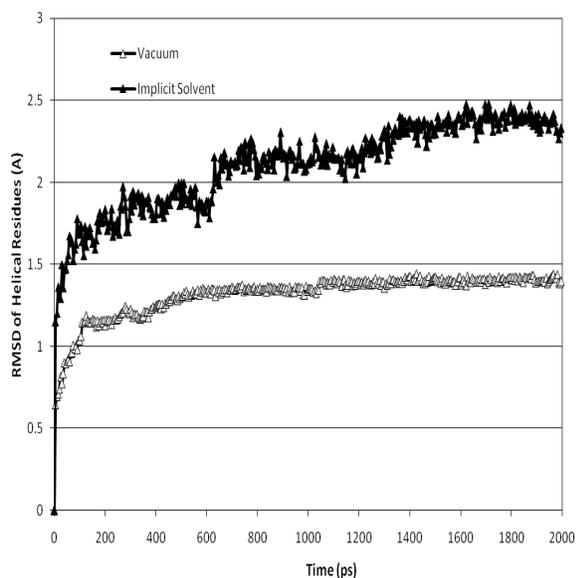


Fig. 2: RMSD of helical residues of GUS in vacuum and implicit solvent

The increased response to the solvent environment does not lead to any increase in the deviation from the regular secondary structural features. On the contrary, solvent environment has more tendencies to retain the secondary features compared to vacuum.

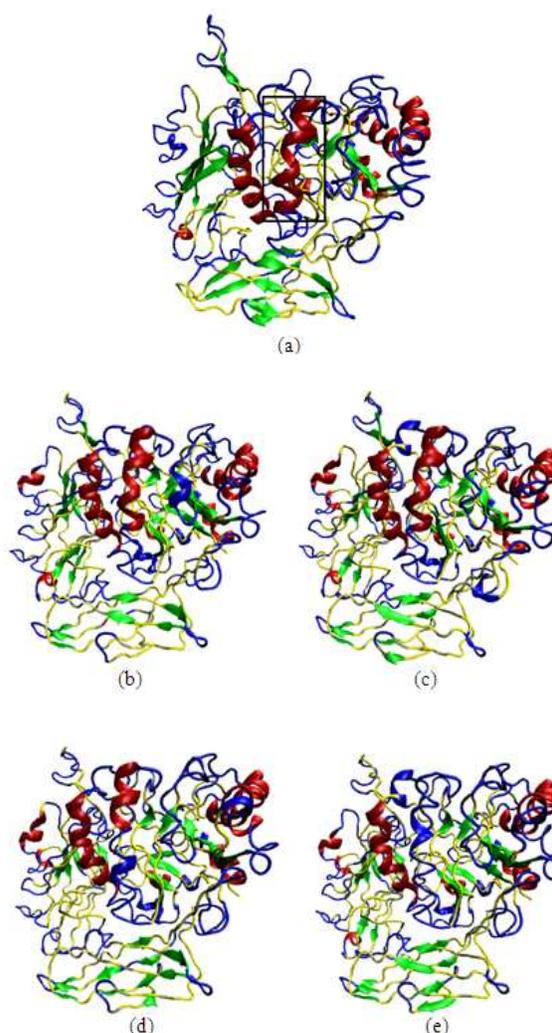


Fig. 3: Simulations of human GUS at 300 K in vacuum (a) 0 ns (b) 0.5 ns (c) 1 ns (d) 1.5 ns (e) 2 ns

There were more fluctuations in the individual helices in GUS in implicit solvent and changes were also observed in the beta sheet region which became shorter with time. From the simulation snapshots of human GUS in implicit solvent (Fig. 4), it can be seen that the structure remained relatively stable, with changes occurring only in turn and coil regions. At 0.5ns there is an introduction of helical turn and beta-sheets (shown in the box). The helical region disappears at 1ns and reappears in the last two frames. Similar behavior is noticed in vacuum also (Fig. 3). However, the trend is not smooth and the turn to beta sheet and then to turn behavior is not observed in vacuum.

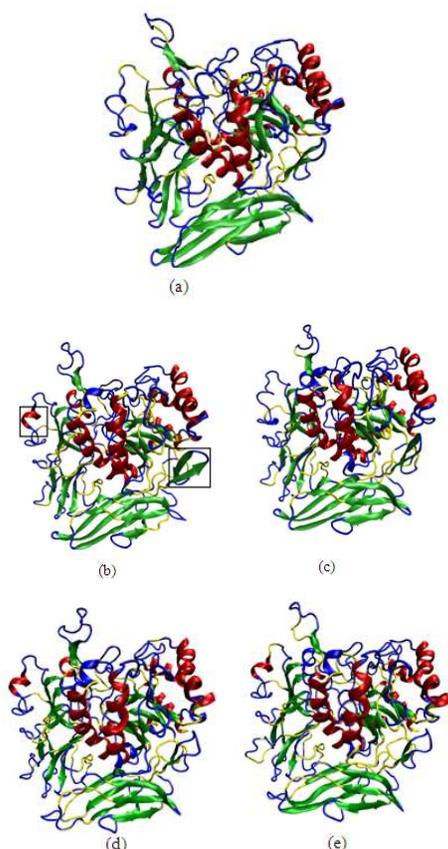


Fig. 4: Simulations of human GUS at 300 K in implicit solvent: (a) 0 ns (b) 0.5 ns (c) 1 ns (d) 1.5 ns (e) 2 ns

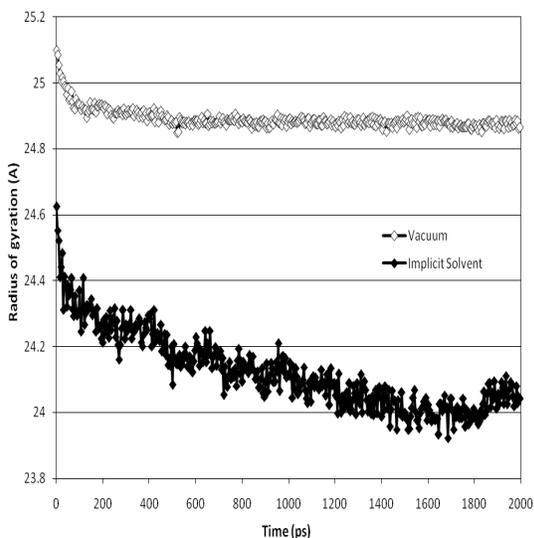


Fig. 5: Radius of gyration of GUS in vacuum and implicit solvent

Structural compactness: Radius of gyration was calculated for all the trajectories to monitor the structural compactness of the protein structure throughout the trajectory. The radius of gyration is often used to describe the compactness of a protein, as well as the folding process from the denatured state to the native state (Hong and Lei, 2009). It is defined as the root mean square distance of each atom to the centroid of the molecules.

Since globular proteins are not spherical objects, the radius of gyration, R_G , is the atomic weight-based root mean square distance of all atoms from the common centre of gravity, X_{CG} , Y_{CG} , Z_{CG} , where X_{CG} is:

$$X_{CG} = \sum_{j=1}^n m_j X_j / \sum_{j=1}^n m_j \text{ etc, for } Y_{CG} \text{ and } Z_{CG} \quad (2)$$

In Eq. 2, m_j is the atomic weight of atom j with coordinates X_j , Y_j and Z_j , and R_G is given as:

$$R_G = (R_G^2)^{1/2} \quad (3)$$

where, R_G^2 is:

$$R_G^2 = \sum_{j=1}^n m_j r_j^2 / \sum_{j=1}^n m_j \quad (4)$$

In the equation above, r_j is the distance of the atom j from the centre of mass of the protein molecule.

As the duration of the simulation (2 ns) is not sufficient to reveal any protein unfolding, no inference on this regard can be made here. However from Fig. 5, it can be inferred that in spite of the increased fluctuations under solvent environment, the radius of gyration values are smaller under solvent environment implying that the overall structure remains more compact in solvent compared to the vacuum.

CONCLUSION

Enzymes, being complex biomolecules have various contacts and their collective motion becomes even more complex. The solvent environment greatly influences the structure and dynamics of the enzyme and different environments lend different characteristics to the system under simulation. These differences are evident from the comparison of backbone stability, helical stability and the structural compactness with respect to simulation time in GUS enzyme. Molecular dynamics simulations using implicit solvent models provide a convenient option to analyze these effects.

REFERENCES

- Becker, O.M. and M. Karplus, 2006. A Guide to Biomolecular Simulations. Springer, ISBN: 978-1-4020-3586-9.
- Berendsen, H.J.C., J.P.M. Postma, A. Dinola and J.R. Haak, 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81: 3684-3690. DOI: 10.1063/1.448118
- Brooks, B.R., R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, 1983. CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comp. Chem.*, 4: 187-217. DOI: 10.1002/jcc.540040211
- Hong, L. and J. Lei, 2009. Scaling law for the radius of gyration of proteins and its dependence on hydrophobicity. *J. Polymer Sci.: Part B: Polymer Phys.*, 47: 207-214. DOI: 10.1002/polb.21634
- Humphrey, W., A. Dalke and K. Schulten, 1996. VMD-visual molecular dynamics. *J. Mol. Graph.*, 14: 33-38. DOI: 10.1016/0263-7855(96)00018-5
- Islam, M.R., S. Tomatsu, G.N. Shah, J.H. Grubb, S. Jain and W.S. Sly, 1999. Active site residues of human β -glucuronidase. *J. Biol. Chem.*, 274: 23451-23455. DOI: 10.1074/jbc.274.33.23451
- Jain, S., W.B. Drendel, Z. Chen, F.S. Mathews, W.S. Sly and J.H. Grubb, 1996. Structure of human β -glucuronidase reveals candidate lysosomal targeting and active-site motifs. *Nat. Struct. Biol.*, 3: 375-381. PMID: 8599764
- Momany, F.A. and R.J. Rone, 1992. Validation of the general purpose Quanta 3.2/CHARMM force field. *Comp. Chem.*, 13: 888-900. DOI: 10.1002/jcc.540130714