Investigating the effect of mutation on the thermo stability of GB1 protein

K N Sawitri¹,³*, T Sumaryada¹, L Ambarsari² and S T Wahyudi¹

¹Computational Biophysics and Molecular Modelling Research Group (CBMoRG), Department of Physics, Bogor Agricultural University, Jalan Meranti Kampus IPB Dramaga Bogor 16680, Indonesia
²Department of Biochemistry, Bogor Agricultural University, Jalan Agatis Kampus IPB Dramaga Bogor 16680, Indonesia
³Department of Physics, University of Jambi, Jalan Lintas Jambi Kampus UNJA Mendalo 36122, Indonesia

E-mail: KaniaNurSawitri@unja.ac.id

Abstract. The thermo stability of wild-type and mutants of B1 domain of Protein G (GB1 protein) have been studied using molecular dynamics simulation and free energy perturbation simulation. This research is aimed to examine what residue or what interaction that has a major role in the thermo stability of GB1 protein thermo stability by using the point mutation method. Based on the analysis, the unfolding of wild-type protein occurred in 500 K simulation at 704 ps. The mutations were chosen based on the changes in some analysis parameters and the calculated net solvation free energy change. It was found that a simple replacement of a positively charged residue in the β-sheet (K4S) decreases the stability of GB1 protein (unfolding at 452 ps), while the replacement of a negatively charged residue in the α-helix (E27G) increases the stability (unfolding at 846 ps). It was also found that the K4A mutation will break the α-helix and all β-sheet into the coil and turn. All those results suggest that the non-bonded interaction has the major role in the thermo stability of GB1 protein with the β-sheets were identified as the most important structure in the thermo stability of GB1 protein.

1. Introduction

The B1 domain of protein G (PDB code 1GB1) is an immunoglobulin protein from G streptococcal group which helps an organism to evade the host defense via unique protein-binding properties [1]. This protein has played a central role in protein model studies to elucidate the folding and unfolding process due to several reasons. First, its unique structure: small size comprises only 56 amino acid residues, simple and highly symmetrical topology [2]. Second, its high thermo stability properties: denaturation points reach 87.5°C without disulfide bridge that cooperatively folds like the other globular proteins [3]. Third, its well characterized mechanical properties: the mechanical unfolding mechanism has been identified to be different from that of chemical denaturant unfolding at the atomic level [4].
Understanding how mechanically or non-mechanically the protein folding or unfolding is still unclear and a major challenge in life sciences. Hundreds of studies have been performed both experimentally using NMR and computationally using molecular dynamic simulation since the last two decades. With the increasing power of computers, simulations can generate highly detailed descriptions of the motions of proteins. As a model, many strategies have been applied on unfolding simulations of GB1 protein to expand the range of problems that may be addressed [5,6]. Some of them are applying the differences of pH using in lucem [2], steer forces using Gromacs 4.5.3 [7], pressure using X-PLOR [8], denature types [9] and temperature using SYMLIS [10].

One of the main targets in protein engineering is increasing the stability of protein by conducting experiment and simulation under various conditions, including mutation. The rational design of mutants is aimed to increase the stability of a protein by either improving the stabilizing interactions or reducing the factors which potentially destabilizing the protein [11]. In this paper we have conducted an investigation on the effect of mutation to the thermo stability of GB1 protein by using an in-silico approach via a molecular dynamics (MD) simulation and the free energy perturbation methods. Analyses of wild-type GB1 protein MD parameters were used as a basis for the screening and mutation selection. By conducting the point mutations on the non-bonded interactions of GB1 protein, we expect to see a thermo stability enhancement as compared to the wild-type protein.

2. Materials and Methods

This study was performed using two simulations. First is the molecular dynamics (MD) simulation for the wild type at various temperatures. The wild-type structure as reference structure was downloaded from [Protein Data Bank (PDB) ID code 1GB1]. The second simulation is the free energy perturbation simulation to find the most stable mutants. Once the mutants were identified, then the mutants were simulated using MD simulation to be analyzed.

2.1. Molecular dynamics simulation

We used NAMD 2.9 [12] with CHARMM potential and VMD 1.9.1 [13], starting with protein placed in a TIP3P water box with dimension 65x54x50 Å. The simulation time step is 2 fs. Each simulation was started by minimization about 100 ps to place the protein on its lowest energy, and then continued by heating and equilibration. The wild-type was heated start from 0 K until the final point at 400 K, 425 K, 450 K, 475 K, and 500 K by increment 25 K. But the mutants were heated start from 0 K until 500 K. The equilibration was obtained by Langevin protocol about 60 ps. We eliminated the constraints in MD production run process so that the protein moves freely. The simulation time for final points 400 K - 475 K is 10 ns, but the simulation time for 500 K is 2 ns.

2.2. Free energy perturbation simulation

The Free Energy perturbation technique (FEP) is a common method to estimate the free energies of various physical processes including the free energy of a mutation of a side chain. The FEP simulation technique can be illustrated in figure 1 where the free energy (ΔG) was calculated via a thermodynamic cycle (figure 1) [14]. The vertical arrows correspond to the hydration of the wild-type and its mutant, while the horizontal arrows correspond to the point mutation. The difference between ΔG of mutant in vacuo (ΔG\text{\textsubscript{alc\textsuperscript{1}}}) and aqueous (ΔG\text{\textsubscript{alc\textsuperscript{2}}}) condition will be compared with the one found in the wild-type protein (ΔΔG). If the ΔΔG in mutant is lower than the wild-type then the mutation has successfully produced a more stable structure. In this research, the FEP simulations were applied by two conditions: vacuo system and aqua system, using dual-topology approach. In vacuo simulation: the increment of λ is 0.05, equilibration time is 10 ps, and the free energy capture time is 100 ps for each λ. In Aqua simulation: the increment of λ is 0.05, equilibration time is 100 fs, and the free energy capture time is 500 fs for each λ. We didn’t use soft-core potential in both conditions to remove the interaction between topology target and the system environment when the λ value near 0 or 1. Wild-type and mutant structure did not interact during FEP simulation. Solvation free energy change was obtained by subtraction of ΔG\text{\textsubscript{alc\textsuperscript{1}}} from ΔG\text{\textsubscript{alc\textsuperscript{2}}} [14].
3. Result and Discussions

3.1. Thermo stability analysis of wild-type

Molecular dynamic simulations for wild-type structure were analyzed at temperature variance: 400 K, 425 K, 450 K, 475 K and 500 K. Simulation at high temperature is a common strategy to reduce the computing time [15]. The results show that 1GB1 protein unfolds at 500K on 704 ps as indicated by the breaking of α-helix structure which is transformed into coil and turn. In figure 2 we show the SASA (solvent accessible surface area) of various residues of the wild-type protein. The information from this figure is important for determining the residue candidate for mutation. As seen in figure 2, the apolar (non-polar) residue which implies hydrophobic interaction is dominant as compared to other type of residues, so the mutation on the apolar residues could potentially change the thermo stability of the GB1 protein. Earlier work shows that salt-bridge which implies non-bonded interaction play a role to stabilize C-termini of β-hairpin fragment [16] and second β-hairpin of domain B1 [17]. Those give insight to choose the mutants based on hydrophobic and non-bonded interactions.

![Figure 1. Thermodynamic cycle of solvation free energy](image)

![Figure 2. The increasing SASA of the wild-type.](image)

3.2 Selection of the mutants

Based on the wild-type simulation analysis at 500K, there are 8 residues of the 1GB1 protein selected to be mutated: K4, K10, A23, A24, E27, K28, D40, and E56. The selection based on the scoring of several MD parameters such as root mean square deviation (RMSD), radius of gyration, Hydrogen bonding, SASA, RMSF, conformation energy. The stability of the protein was analyzed by inspecting six different properties: RMSF (root mean square fluctuation), salt-bridge occurrence, SASA, the stability of secondary structure elements, non-bonded energy, and hydrogen bond occurrence. For mutant selection, we consider the conformation preference of secondary structure. To select the best mutants, FEP simulations generate solvation free energy change (ΔΔG). The highest and lowest ΔΔG led to biggest changes in point mutation, so there is 16 mutants continue to molecular dynamic
simulation (table 1). K10G means that lysine (residue number 10) was mutated into glycine, same with others.

| No | Mutant | ΔΔG_{solv} (Kcal/mol) | No | Mutant | ΔΔG_{solv} (Kcal/mol) |
|----|--------|----------------------|----|--------|----------------------|
| 1  | K10G   | 10.5705              | 9  | K10N   | 40.4396              |
| 2  | D40T   | 33.0834              | 10 | D40L   | 58.3459              |
| 3  | E56V   | 32.5037              | 11 | E56I   | 55.6443              |
| 4  | A23S   | 0.2948               | 12 | A23Q   | 6.1274               |
| 5  | A24S   | 0.34247              | 13 | A24T   | 8.46362              |
| 6  | K4S    | 8.6961               | 14 | K4A    | 23.6733              |
| 7  | E27G   | 10.8084              | 15 | E27S   | 53.3006              |
| 8  | K28T   | 0.1974               | 16 | K28C   | 26.6864              |

### 3.3 MD Simulation analysis of Mutants

Protein G B1 domain has 56 amino acid residues which are composed of secondary structures: a α-helix (residue 23-36), 4-stranded β-sheet (residue 2-8, 13-19, 42-46, 51-55), 2 coils (residue 1, 20-22, 41, 51-55) and 2 turns (residue 9-12, 47-50) as shown in figure 3. They are linked by hydrogen bonds between oxygen atoms on carbonyl groups with hydrogen atoms on amide groups. The fluctuation on secondary structure represents that the residues are maintaining its existence. All mutants unfold at 500K as shown by the destruction of α-helix structures (figure 4). Three out of sixteen mutant simulations gave interesting results, and all will be discussed in this section. Wild-type GB1 unfolds on 704 ps (figure 4a), where the mutants K4S is the most unstable protein since it unfolds at 452 ps (figure 4b). Analysis of secondary structure shows that Lysine mutated into Serine cause most of the β-sheet at the N-terminal end was destroyed and transformed into coil and turn structures (figure 5b). This effect can be described by conformation preference that available in the web server located at [http://bmrB.wisc.edu/referenc/choufas.shtml](http://bmrB.wisc.edu/referenc/choufas.shtml). Lys-4, positively charged residue, is located on first strand β-sheet, and lysine is a β-sheet breaker with p-value 0.74. But serine, neutrally charged residue, has higher conformation to break the β-sheet since its p-value is 0.72. Another impact of this mutation is the disappearance of salt-bridge pair D46-K4 and some hydrogen bonds K4 with residues on third and fourth strand β-sheet.

![Figure 3](image-url) The cartoon representation of GB1 protein. The colors represent secondary structure: Purple (α-helix), Yellow(β-sheet), Blue (310 helix), Green (Turn) and White(Coil).
Figure 4. Analysis of secondary structure elements of protein G B1 domain on 500K simulation: (a) Wild-type, (b) mutant K4S, (c) mutant K4A, (d) mutant E27G. X-axis represents time (0-2000 ps), and Y-axis represents residue number (0-56).

Figure 5. Final conformation at the end of md 500K simulation: (a) Wild-type, (b) mutant K4S, (c) mutant K4A, (d) mutant E27G.
Second interesting mutant is K4A (Lysine-4 mutated into Alanine) that unfolds at 680 ps (figure 4c). K4A is the only one of sixteen mutants that all of its secondary structure (α-helix and β-sheet) were completely destroyed (figure 5c), and it will be elucidated through RMSF analysis. The K4A mutant unfolds faster than the wild-type since the mutation in principle totally eliminated the salt-bridge and hydrogen bonds contacts. The K4A mutant turned out to be the most flexible mutant, since its RMSF fluctuates within the range of 8-17 Å, while the wild-type only fluctuates in the range of 0-4 Å (figure 6). We suggest that this high flexibility of K4A mutant was caused by the point mutation on β-sheet which responsible for the protein stability. Moreover, hydrophobic contacts between protein and water become stronger since Alanine is non-polar residue. The third interesting mutant is E27G (Glu-27 mutated into Gly) that stabilize the GB1 protein as shown by a longer unfolding time at 846 ps, later than the wild-type structure (figure 4d). The final conformation of E27G has a complete three last strand of β-sheet as seen in figure 5d, while the wild-type has not.

![RMSF values over MD simulation.](image)

4. Conclusion
Based on the MD and FEP simulation analysis, three mutants were selected to investigated the effect of point mutations to the thermo stability of GB1 protein. The point mutation simulations indicated that β-sheet is the important structure to maintain the thermo stability of GB1 protein. Mutation on this structure (K4S and K4A) suggests the protein unfolds faster than the wild-type (less-stable protein), while the mutation on α-helix (E27G) stabilize the protein as indicated by the longer time (more-stable protein) needed by the mutant to unfold as compared to the wild-type. The elimination of non-bonded interaction leads to the collapse of protein secondary structure especially in mutant K4A which has the highest RMSF. This result suggests that non-bonded interaction plays an important role in the thermo stability of GB1 protein.

5. References
[1] Gronenborn AM and Clore GM. 1993. *ImmunoMethods*. 2(1): 3-8
[2] Morrone A, Giri R, Toofanny RD, Allocatelli CT, Brunori M, Dagget V and Gianni S. 2011. *Biophys J*. 101(8): 2053-60
[3] Kobayashi N, Honda S, Yoshii H, Uedaira H and Munekata E. 1995. FEBS Lett. 366 (2-3): 99-103
[4] Shen T, Cao Y, Zuo S and Li H. 2012. Biophys J. 103 (11): 2361-68
[5] Granata D, Camilloni C, Vendruscolo M and Laio A. 2013. PNAS. 110(17): 6817-22
[6] Sumaryada T, Hati J, Wahyudi ST, Malau ND, Sawitri KN, IOP Conf. Series: Earth and Environmental Science 31 (2016): 012008
[7] Kouza M and Hansmann UHE. 2012. J.Phys.Chem. B. 116 (23): 6645-53
[8] Wilton DJ, Tunnicliffe RB, Kamatari YO, Akasaka K and Williamson MP. 2008. Proteins. 71 (3): 1432-40
[9] Wei H, Fan Y and Gao YQ. 2010. J. Phys. Chem. B. 114 (1): 557-68
[10] Lee J and Shin S. 2001. Biophys J. 81 (5): 2507-16
[11] Pikkemaat MG, Linssen AB, Berendsen HJ and Janssen DB. 2002. Protein Eng. 15 (3): 185-92
[12] Philips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L and Schulten K. 2005. J Comput Chem. 26 (16): 1781-802
[13] Humphrey W, Dalke A and Schulten K. 1996. J Mol Graph. 14 (1): 33-8
[14] Henin J, Gumbart J, Harrison C and Chipot C. 2012. In Silico Alchemy: A Tutorial for Alchemical Free-Energy Perturbation Calculations with NAMD (University of Illinois at Urbana-Champaign)
[15] Pande VS and Rokhsar DS. 1999. PNAS. 96 (16): 9062-67
[16] Tsai J and Levitt M. 2002. Biophys Chem. 101-102: 187-201.
[17] Huyghues-Despointes BM, Qu X, Tsai J and Scholtz JM. 2006. Proteins. 63 (4): 1005-17