Available online at www.ijpcr.com
International Journal of Pharmaceutical and Clinical Research 2016; 8(4): 210-215

ISSN- 0975 1556

Research Article

Interaction Between SP1 and G-6A AGT Gene for Revealing the Effect of Polymorphism in Hypertension

Aprilia D S[1], Widodo[2*], Rohman M S[3], Utomo D H[2], Lukitasari M[4]

[1]*Biomedical Sciences, Faculty of Medicine, University of Brawijaya, Malang, Indonesia,*
[2]*Biology Department, Faculty of Mathematics and Natural Sciences, University of Brawijaya, Malang, Indonesia,*
[3]*Department of Cardiology and Vascular Medicine, Faculty of Medicine, Brawijaya University, Saiful Anwar General Hospital, Malang, Indonesia,*
[4]*Nursing Department, Faculty of Medicine, University of Brawijaya, Malang, Indonesia,*

Available Online:01st April, 2016

ABSTRACT

Several experiments have identified polymorphism in the proximal of angiotensinogen (AGT). G-6A is one of polymorphism of AGT that involved hypertensive risk. This polymorphism increased AGT plasma level that may caused by increased transcription activity. This polymorphism located in proximal promoter that may regulate the gene. This polymorphism may regulated by Sp1 that influenced by polymorphism that affect angiotensinogen expression. However, the mechanism of interaction Sp1 and G-6A and its influence for expression is unknown. Thus, this study aimed to investigate the interaction between polymorphism G-6A and Sp1 (Specificity protein 1). We used molecular docking for predicting interaction and molecular dynamic for predicting the stability the interaction. We reported that the Sp1-DNA allele A complex was more stable than allele G complex. The conformation during simulation showed that the Sp1-DNA allele A complex was more stable than allele G complex. The Sp1-DNA allele A complex have more bond than allele G complex. The bond consist of hydrogen contact and hydrophobic contact that may contribute to form stable interaction in Sp1-DNA allele A complex. This study suggested that the G to A alteration at the position -6 leads to increased AGT promoter activity, that may increased risk in causing hypertension.

Keywords: Angiotensinogen, Polymorphism, G-6A, Specificity protein 1, Molecular Docking, Molecular Dynamic

INTRODUCTION

Angiotensinogen (AGT) is the first gene to be related to hypertension in humans. Several experiments have identified polymorphism in the proximal promoter of angiotensinogen (*AGT*) that involved hypertension risk. Polymorphism of angiotensinogen gene located at proximal promoter region start site such as -6, -20, -152, -217[1]. This polymorphism G-6A increased AGT plasma level that may caused by increased transcription activity[2]. One of transcription factor that may regulate transcription activity in the angiotensinogen gene G-6A is Sp1 (Specificity protein 1)[1]. Sequence angiotensinogen gene surrounding G-6A has a motif of Sp1 binding site. Sp1 regulate expression of many genes involved in cell proliferation, differentiation and apoptosis through their binding to G/C-rich sequences and subsequent interaction with the basal transcription[3-5]. Inoue, *et al* (1997) reported that Haplotype -6G have lower transcription activity than haplotype -6A[6]. Further, we reported that molecular variation in the proximal promoter G-6A AGT Gene can impact differences in the transcription activity binding pattern and its conformation. Thus, this study aimed to investigate its differences using in silico analysis to predict the interaction and the stability.

MATERIALS AND METHODS

Starting structure of Sp1 Protein and DNA

A double stranded 24bp DNA was built using 3D-DART (3DNA- Driven DNA Analysis and Rebuilding Tool) web server. We generated 3D structural models of DNA from sequences of angiotensinogen promoter (5' - cccggccgggggaagaagctgccg- 3') allele G and (5'- cccggccaggggaagaagctgccg- 3') allele A. The angiotensinogen promoter (5'-3' sequence) was given as input. The homology model of Sp1 was constructed using chimera 1.10. The structure from PDB (1alf,1mey,1jk1) were chosen as templates for modeling.

Docking Procedure

Sp1 (DBD)- DNA was docked against the DNA allele G and allele A contained double helical DNA using the docking program HADDOCK. The docking procedure consisted of three stages: (i) randomization of orientations and rigid body energy minimization (EM), (ii) semirigid simulated annealing in torsion angle space (TAD-SA), and (iii) final refinement with explicit solvent[7]. HADDOCK was run using its default parameters but with an additional input list of amino acids that might be involved in interactions with the DNA [Lys546, Val547, Tyr548, Lys550, Ser552, His553, Trp560, Arg580, Ser581, Asp582, Glu583, Gln585, Arg586, Arg589, Phe597,

**Author for Correspondence*

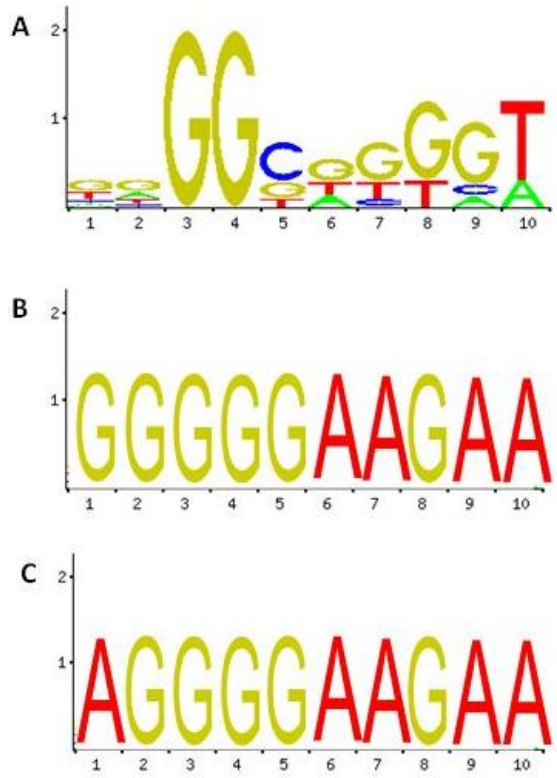

Figure 1: Consensus variations of Sp1-binding site. Consensus of Sp1-binding site was retrieved from Jaspar CORE database (A), Varian angiotensinogen promoter at -6G (B), and Varian angiotensinogen promoter at -6A (C).

Arg608, Ser609, Asp610, His611, Ser613, Lys614. This way, the docking volume to be sampled was reduced and the potential DNA binding site in the target was defined. The residues were selected based on experimental findings as well as electrostatic surface characteristics of the protein and DNA [G8, G9, G10, G11, G12, A13, A14, G15, A16, G17]. The best 40 complex models were selected on the basis of HADDOCK score defined as a weighted sum of intermolecular electrostatic, van der Waals contacts, desolvation, EAIR and BSA term[8]. The final stage of the docking protocol is gentle water refinement. The effects of global and local flexibility of the DNA during docking have been reported[9] thus, the default option was used to define the flexible regions of DNA.

Molecular Dynamic Procedure

All-atom molecular dynamics (MD) simulations were performed using GROMACS with the CHARMM27 force field for the protein and the nucleic acid. The protein–nucleic acid complexes were solvated using SPC water molecules in box of 100 A˚ × 100 A˚ × 100 A˚ dimension. A total of 32 sodium ions were added to neutralize the system, the final model having approximately 1551 atoms. In the first step, the MD system was energy minimized using a conjugate gradient algorithm for 2000 iterations up to a convergence value of 1.0 kcal/mol/Å. In the second step, the MD system was heated to 300 K in isochore conditions (NVT). Each system was simulated in the isothermalisobaric ensemble (NPT) at a pressure of 1 atm. The production runs were carried out up to 20000ps. The GROMACS method was used for the calculation of the electrostatic interactions. The temperature was kept constant at 300 K by a Berendsen thermostat. MD structures have been compared the allele G complex and allele A complex through the calculation of RMSDs, RMSF, and potential energy.

Analysis of the docking and molecular dynamic

For each docking, the wrap-around orientation of the complex models was analyzed by use of the Pymol program[10] and UCSF chimera 1.10. The 40 best complex models were selected on the basis of HADDOCK score. The first cluster were choosen because it was the best energy. We also analyzed the difference of binding model Sp1-DNA using superpose web server. SuperPose calculated protein superpositions using a modified quaternion approach. SuperPose looked for structurally similar and dissimilar regions between aligned protein chains. This is useful in identifying hinge motions, mobile segments, etc. If SuperPose finds structurally dissimilar regions, it will superpose the structures based on the single longest structurally similar region shared by the sequences. For simulation dynamic, Three-dimensional structures and trajectories of MD simulation product were visually inspected using the VEGAZZ pogram. Root-mean-square deviaton (rmsd), potential energy and Root-mean-square fluctuation (rmsf) have been calculated using xm grace GROMACS MD package version 3.14.

RESULT

Sp1 (DBD)- DNA Complex Model Docking

We found sequence surrounding G-6A promoter has significantly for Sp1 binding site. (Fig.1). The 24 nucleotide surrounding G-6A promoter were dock with three zinc finger to elucidate pattern of Sp1 bind to angiotensinogen promoter by using docking method. The docking result showed that HADDOCK score of allele A complex was more smaller than allele G complex (120.2 +/- 10.8 and -122.5 +/- 2.9 (Sp1- DNA allele G complex and allele A complex, respectively). The score indicated that Sp1- DNA allele A complex was more favorable than Sp1- DNA allele G complex. Then, we analyzed the difference of superpose of allele G complex and allele A complex. We found that superpose of Sp1 and DNA give a rmsd value of 1,04 Aº. Zinc finger 3 in Sp1-DNA allele A complex was more favorable than Zinc finger 3 in Sp1-DNA allele G complex. Zinc finger 2 bind well both allele A complex and allele G complex. Whereas, zinc finger 1 was not favorable contact. This indicated that genetic variation G-6A of angiotensinogen promoter causes differences of Sp1 binding to angiotensinogen promoter (Fig.2). The Sp1 binding pattern to SNPs in G-6A of angiotensinogen promoter analyzed by using nucplot program. The result indicated that the alteration G to A causing differences of residue contact. The differences is Histidine as hidrofobic contact in allele G complex and Histidine as hydrogen bound and Asp as hidrofobic contact in allele A complex (Fig.2).

Molecular Dynamic Sp1 (DBD)- DNA Complex

The aim of MD simulation was to investigate stability binding protein and DNA that could be used for

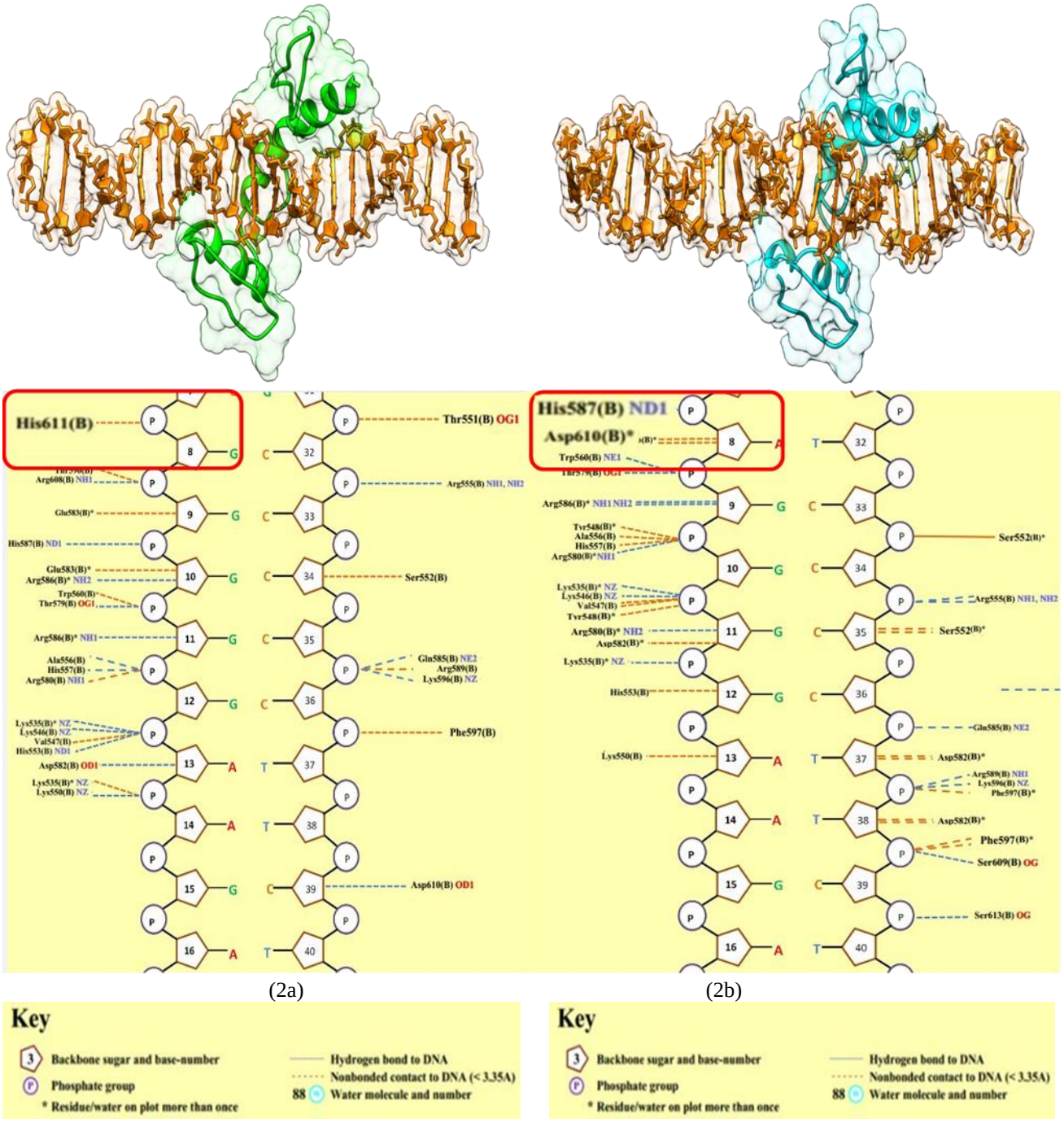

Figure 2: 3D representation of docking result Sp1– DNA. (2a) 3D representation of docking result Sp1– DNA allele A complex. (2b) 3D representation of docking result Sp1 – DNA allele G complex. (2c) DNA- protein contact in Sp1 – DNA allele A complex. (2d) DNA- protein contact in Sp1– DNA allele G complex.

subsequent docking studies. The simulation was carried for total 20000ps and after the run was completed, the trajectory files (*.trr and *.tpr) were generated[11]. There was a significant differentiation potential energy in conformation of Sp1- DNA complex (Fig.3). The Sp1 - DNA allele G complex showed that the RMSD of Sp1- DNA allele G complex was greater than allele A complex. This suggested that Sp1- DNA allele A complex was more stable than allele G complex during simulation (Fig.4). The Sp1 - DNA allele G complex shows a fluctuations in residues 75-100. Whereas, The Sp1- DNA allele A complex has lower fluctuation in residues 75-100. The data indicated that the rmsf of Sp1- DNA allele A complex was more stable than allele G complex. The Sp1- DNA allele G complex induced dynamical changes in protein-DNA binding (Fig.5).

DISCUSSION

This study suggested that Sp1 is a transcription factor that regulate angiotensinogen gene. The gene were influenced

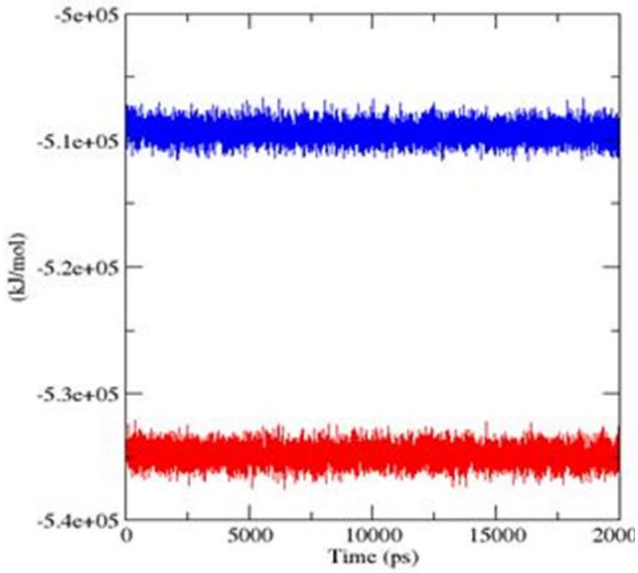

Figure 3: Grafic of Energy potential since simulation Sp1– DNA 20000ps. The Sp1- DNA allele A complex was drawn in red. The Sp1- DNA allele G complex was drawn in blue.

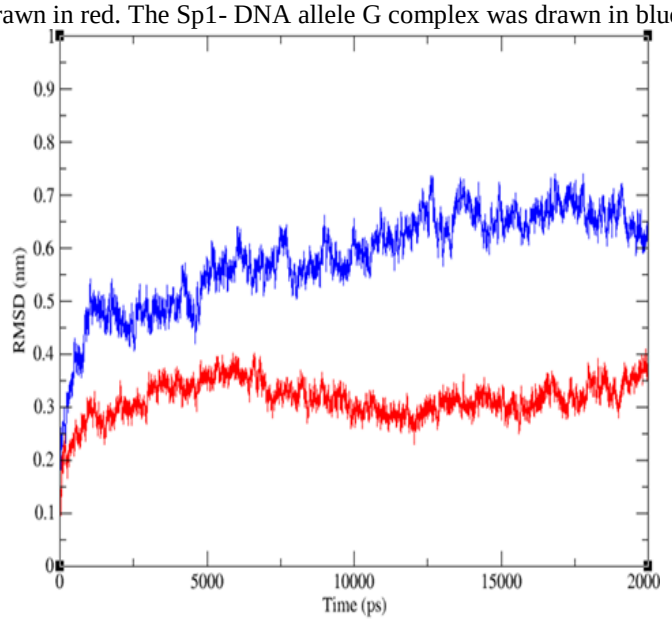

Figure 4: Grafic of RMSD (Root Mean Square Deviation) since simulation Sp1– DNA 20000ps. RMSD represented as a function of simulation time after a mass- weighted superpositionon on the starting structure. RMSD represent as potential descriptor in binding affinity that measured backbone motion atom to represent complex motion in all protein and side chain atom. RMSD of Sp1-DNA allele G complex and allele A complex is reported in blue and red graphic respectively.

The complex Sp1 (Zf)- DNA allele G was represented in blue. The complex Sp1 (Zf)- DNA allele A was represented in red.

by polymorphism G-6A AGT gene that affect angiotensinogen expression. It is correspond with in-vitro analyses of human AGT transcriptional activity suggested that the A to G change at the position -6 leads to decreased AGT promoter activity, indicating that this could be a phenotype causing hypertension[2]. Similarly, Chaves, et al. (2002) reported that the G-6A polymorphism of the AGT gene has been linked with increased body weight gain in hypertensive patients[12]. Sp1 consist of three zinc finger motif as a DNA binding domain[13]. Zinc-finger is one of the most common DNA-binding motifs in eukaryotes[14]. In the present study, we found that zinc finger 1 was not favorable contact with DNA allele G complex and allele A complex. Zinc finger 3 in allele A complex was more favorable contact than in allele G complex. Some previous study of the Sp1- DNA interaction showed the different contribution of the three fingers of Sp1 to the GC box DNA binding. Other report stated that Sp1 zinc finger 2 and 3 have important activity to bind with DNA/ cis element[13,15]. Zinc finger are among the most abundant proteins in eukaryotic genomes[16]. In the previous study showed that putative interaction Sp1 and DNA GC Box involved active

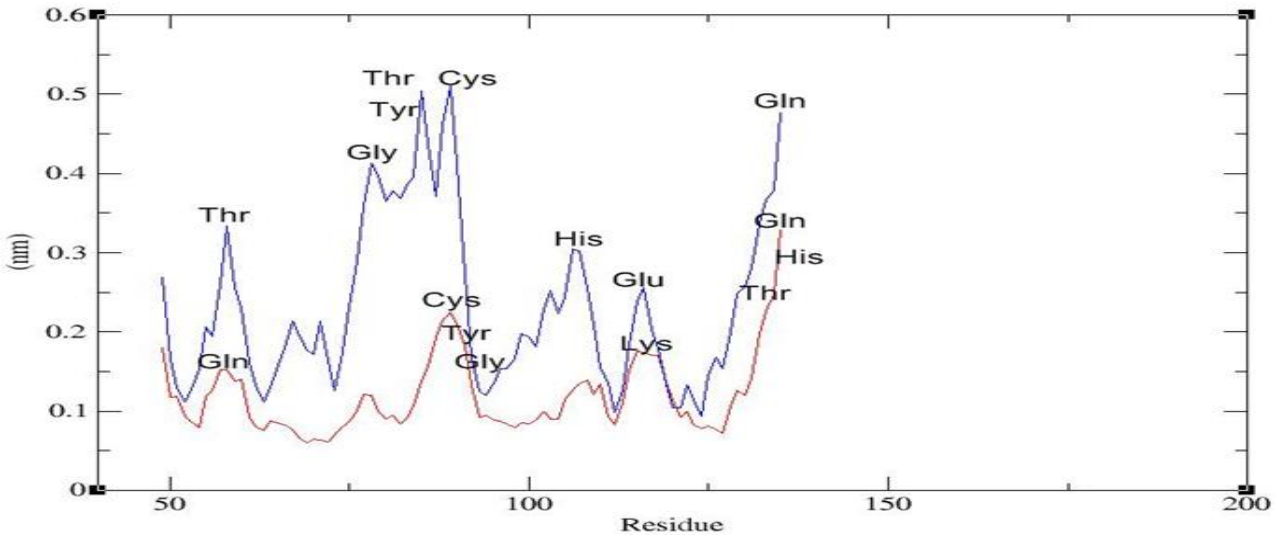

Figure 5: Grafic of RMSF (Root Mean Square Fluctuation) since simulation Sp1– DNA 20000ps. The per residue RMSF are represented as a function of the residue number. RMSF to represent flexibility interaction of Sp1- DNA. RMSF defined residues that have high fluctuation and identified the role of site active protein. RMSF of Sp1-DNA allele G complex and allele A complex is reported in blue and red graphic respectively. The complex Sp1 (Zf)- DNA allele G was represented in blue. The complex Sp1 (Zf)- DNA allele A was represented in red.

site Lys550, His553 in zinc finger 1, Arg580,Gly583,Arg586 as identical residue in zinc finger 2, and Arg608, Lys614, His611, Asp610 in zinc finger 3. In this study, we found that the allele A complex bind to His611 as a hydrogen bond and bind to Asp610 as a hydrophobic bond. Whereas, the allele G complex bind to His610 as hydrophobic bond. Histidine in the three zinc finger of Sp1 gives rise to a protein that is fully functional in DNA binding[17]. Histidine residues bind well to adenine as guanine since the observed as hydrogen bond[18]. It indicated that the allele A complex was more favorable contact than the allele G complex. Based on binding affinity analyses, we found that energy binding of the allele A complex was smaller than the allele G complex. It showed that the allele A complex was more favourable than allele G complex. Binding energy can be influenced by hydrogen bond and hydrophobic bond that formed by interaction. In this study, the allele A complex have more bond than allele G complex. It may cause stability in the allele A complex. Hydrogen bonds are an important contributor to free energies of biological macromolecules and macromolecular complexes[19]. However, hydrophobic interactions as a key factor for protein thermostability[20] and contribute to stabilize system. It indicated that both hydrogen contact and hydrophobic contact could complete each function. This study showed that RMSD in allele A complex was more fluctuative than allele G complex. The increased RMSD showed that DNA structure started to open. In the other side, the protein proceed to find binding site and its coordinat. Whereas, stable RMSD indicated that protein has reached the maximum conformation bind to DNA. Thus, the protein was able to depend its position. Further investigation on RMSF showed that allele G complex have higher fluctuation than allele A complex. However, higher fluctuation in the allele G complex was not active residue (i.e Thr, Cys, Tyr,and Gly). On the other side, higher fluctuation in the allele A complex involved active residue (i.e His and Lys). It indicated that the alteration of nucleotide guanine into adenine affect alteration of active residue function and molecular dynamics in protein-DNA binding.

CONCLUSION

The binding energy of Sp1- DNA allele A complex was higher than allele G complex. We found fluctuation of Sp1- DNA allele G complex was higher than allele A complex. Sp1 was more favorable contact with DNA allele A complex than allele G complex. The G to A alteration at the position -6 leads to increase binding pattern of Sp1 that may increased promoter of Sp1 activity.

ACKNOWLEDGEMENTS

The authors would like to thank Arina M., Jayarani F., Aditya K., Rusdianto, Imama M, Frastiqa F, Akhyan HS, Amelia I., Ika Arum DS, Shila W, Risa Ramadhani and Ratna R for supporting this work.

REFERENCES

1. Dickson, Matthew, Sigmund. Genetic Basis of Hypertension : Revisiting Angiotensinogen. Hypertension. 2006; 48: 14-20.
2. Cvetkovic B, Keen H, Zhang X, Davis D, Yang B. Physiological significance of two common haplotypes of human angiotensinogen using gene targeting in the mouse. Physiol genomics. 2002; 11: 253- 262.
3. Black, A. R., Black, J. D, Azizkhan-Clifford, J. Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. J. Cell. Physiol. 2001;188: 143 – 160.
4. Philipsen S, Suske, G. A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. Nucl. Acids Res. 1999; 27: 2991 – 3000
5. Suske, G. The Sp-family of transcription factors. Gene. 1999; 238: 291 – 300.

6. Inoue I, Nakajima T, Williams CS, Quackenbush J, Puryear R, Powers M, et al. A Nucleotide Substitution in The Promoter of Human Angiotensinogen is Associated with Essential Hypertension and Affects Basal Transcription in Vitro. J Clin Invest. 1997; 99:1786–1797
7. Dominguez, Rolf Boelens, and Alexandre M. J. J. Bonvin. HADDOCK: A Protein- Protein Docking Approach Based on Biochemical or Biophysical Information. J. AM. CHEM. SOC. 2003; 125: 1731-1737
8. De Vries SJ, van Dijk AD, Krzeminski M, van Dijk M, Thureau A, Hsu V, et al. HADDOCK versus HADDOCK: new features and performance of HADDOCK 2.0 on the CAPRI targets. Proteins. 2007; 69(4):726-733
9. Van Dijk M, van Dijk AD, Hsu V, Boelens R, Bonvin AM. Information- Driven Protein- DNA Docking Using HADDOCK: It is a Matter of Flexibility. Nucleic Acids Res. 2006; 34(11): 3317-3325.
10. DeLano WL. The PyMOL Molecular Graphics System. San Carlos, CA,USA: DeLano Scientific; 2002
11. Mukherjee, Koel , Asim K. Bothra, Ambarish S. Vidyarthi. Molecular Dynamics Simulation of DNA Binding Domain of hTRF2. Journal of Advanced Bioinformatics Applications and Research. 2012; 3(2): 324-327
12. Chaves FJ, Giner V, Corella D, Pascual J, Marin P, et al. Body weight changes and the A-6G polymorphism of the angiotensinogen gene. In J Obes Relat Metab Disord. 2012; 26: 1173-1178
13. Kriwacki RW, Schultz SC, Steitz TA, Caradonna JP. Sequence-specific recognition of DNA by zinc-finger peptides derived from the transcription factor Sp1.Proc. Natl. Acad. Sci. USA. 1992; 89(20): 9759-63.
14. Zhang, Baozhen, Shengyan Xiang, Yanru Yin, Liankun Gu and Dajun Deng. C-terminal in Sp1-like Artificial Zinc-finger Proteins Plays Crucial Roles in Determining Their DNA Binding Affinity. BMC Biotechnology . 2013; 13:106.
15. Kuwahara, J., and Coleman, J. E. Role of The Zinc(II) Ions in The Structure of the three-finger DNA binding domain of the Sp1 transcription factor. Biochemistry. 1990; *29*, 8627-8631
16. Laity, Lee, Wright. Zinc Finger Proteins: New Insight into Structural and Functional Diversity. Structural Biology. 2001; 11: 39- 46.
17. Hori Y, Suzuki K, Okuno Y, Nagaoka M, Futaki S, Sugiura Y. Artificial Zinc Finger Peptide Containing a Novel His4 Domain. J Am Chem Soc. 2000; 122:7648-7653.
18. Berg, Jeremy. Sp1 and the subfamily of zinc finger proteins with guanine-rich binding sites. Proc. Nati. Acad. Sci. USA. 1992; 89: 11109-11110.
19. Morozov AV, Kortemme T. Potential Functions for Hydrogen Bonds in Protein Structure Prediction and Design. Protein Chemistry. 2005; 72: 1-38.
20. Prakash Chandra Rathi, Hans Wolfgang Hoffken, and Holger Gohlke. Quality Matters: Extension of Clusters of Residues with Good Hydrophobic Contacts Stabilize (Hyper)Thermophilic Proteins. Journal of Chemical Information and Modeling. 2014; 54: 355 – 361.