Ab initio computational study of electronic structure part-1: reaction mechanism of peptide bond formation between amino acid alanine and glycine

A Dzikrullah¹, B Cahyono², M D Laksitorini³, and P Siahaan¹,*

¹Physical Chemistry Laboratory, Department of Chemistry, Faculty of Sciences and Mathematics, Diponegoro University, Semarang
²Organic Chemistry Laboratory, Department of Chemistry, Faculty of Sciences and Mathematics, Diponegoro University, Semarang
³Department of Pharmaceutics, Faculty of Pharmacy, Gadjah Mada University, Sekip Utara, Yogyakarta
*E-mail: siahaan_parsaoran@yahoo.com, siahaan.parsaoran@live.undip.ac.id

Abstract. The porosity of the peptide delivery pathway to the brain is hindered by the presence of tight junctions which are intercellular cadherin interactions, but this can be overcome by modulating the cadherin molecule using peptide derived synthesis, one of which is ADT-10 (Ac-QGADTPPVGV-NH₂), where there are amino acids glycine (G) and alanine (A). Formation reaction of the peptide is one of the most important chemical reactions, one way to probe the reaction of peptide synthesis is the computational method. The purpose of this research is to determine which mechanism of the reaction is most preferred to the synthesis of peptide bond formation between alanine and glycine from four pathways of the reaction mechanism, as well as glycine and glycine from two pathway of reaction mechanisms by ab initio computational approach. The calculations were carried out by theory and basis set HF/6-31g++. The results show the most preferred reaction of peptide synthesis of amino acid glycine and alanine is on the mechanism IV which result in Ac-GA-NH₂ with activation energy 759.614 kJ.mol⁻¹, while in glycine and glycine is on the mechanism II with an activation energy of 933.550 kJ.mol⁻¹.

1. Introduction
Peptides have been known to have potential as therapeutic agents, but its assembling is hampered by the difficulty of sending peptides to the target site, this is due to the biological barrier called the Blood Brain Barrier (BBB), this barrier consists of cell membranes built by cells with intercellular junctions [1]. One of the pathways that can be traversed in the BBB is the paracellular pathway, but there is a tight junction, namely the cadherin-cadherin interaction between cells. This path can only be passed by molecules, which have a size less than 11 Å. The porosity of the paracellular pathway can be improved by modulating cadherin molecules using peptides synthesized by their derivatives, one of which is ADT-10 peptide [1,2]. Emil Fischer first synthesized the peptide. Moreover, the research about peptide synthesis has been introduced by Theodore Curtis since 1881 after more the research of peptide synthesis has grown exponentially [3]. The Reaction of peptide synthesis is a reaction between two active groups on the amino acid, the amine group, and the carboxyl group [4-6]. In 2002, Sinaga et al. had synthesized derivative of ADT peptide of extracellular E-cadherin namely ADT-10 (Ac-
QGADTPPVGV-NH$_2$), where there is sequence bonding peptide of amino acid glycine (G) and alanine (A). ADT-10 and E-cadherin have synthesized polynucleotides naturally in the human body [1]. Both of these amino acids (G and A) have different R-group that are H and CH$_3$. The reaction mechanism of ADT-10 sequence synthesis is interesting to study, but in this study, only G and A were studied because of the difference in the R-group and why that formed was GA instead of AG, does this thing related to the activation energy (Ea). Siahaan has initiated the mechanism of the peptide synthesis reaction, et al. (2017), who studies the formation of peptide bonds from the amino acids proline and valine. The results show that the peptide formed is Ac-PV-NH$_2$ following its natural synthesis, that can be seen in the ADT-10 peptide sequence [7]. In this study the reaction of the peptide bond formation between glycine and alanine will produce two products, namely Ac-AG-NH$_2$ and Ac-GA-NH$_2$, those are based on four possible mechanism pathways (BI, BII, BIII, and BIV). Meanwhile, the reaction between glycine and glycine that produces Ac-GG-NH$_2$ peptide from two possible mechanism pathways (AI and AII) was used as a comparison. Activation energy calculation is related to the Arrhenius equation $k = Ae^{-Ea/RT}$, the rate determinant of the reaction mechanism is the one that has the highest Ea with a slow reaction, and the most preferred mechanism is the one that has a lower Ea with a fast reaction. One way that can be used to study the reaction mechanism of peptide bond formation is ab initio computational method [8-13]. Generally, in the reaction of peptide bond formation, an amino acid acidic will release H atom in the carboxyl group, but in this research reaction will be tried to terminate the OH bond in the carboxyl group. Beside that, in this research will be tried to terminate the H atom on the amino acid amine group. The reaction mechanism of peptide synthesis also has been studied in ribosome formation by Trobro and Aqvist, the peptide bond forming reaction according to Trobro and Aqvist is as follows [14].

2. Computational Methodology
All calculation has been performed with the NWChem using HF(SCF) theory level and 6-31g** basis set. Meanwhile, Chemcraft was used as visualization software. Notepad++ was used to create the file input with the parameter of bond length (r), bond angle (A), and dihedral angle (d). The level of theory and basis set has used HF(SCF) 6-31g** to calculate the molecular energy of the reactants (R), products (P), intermediates (I), and transition states (TS). Basis set was used according to the character of the molecules and atom, the smaller the basis set, the poorer the representation and also influences the accuracy [15-17]. The calculation was done with "task scf optimize" to optimize the molecular structure and "task scf freq" is to determine the amount of vibration of each molecule and to show the molecules R, P, and I with no negative frequency. The "task scf saddle" is to determine the structure of transition state and "task scf freq" to show one imaginary frequency as a saddle point [18,19].
Scheme 1. Reaction mechanisms of peptide bond formation Ac-GG-NH$_2$ on AI and AII

Scheme 2. Reaction mechanism of peptide bond formation Ac-AG-NH$_2$ on BI and BII

Scheme 3. Reaction Mechanism of peptide bond formation Ac-GA-NH$_2$ on BIII and BIV

3. Results and Discussion
In the discussion of the results of these six mechanisms, three mechanisms will be compared by the same initial stage, that is the termination of the hydrogen bond on the amine group and the termination of the C-O bond on the carboxylic group. AI, BI, and BIII are the reaction pathway that begins with the termination of hydrogen bonds in the amine group. AII, BII, and BIV are the reaction pathway that begins with C-O bond termination in the carboxylic group. Glycine and Alanine have a molecular minimum energy count of \(-7.423\times10^5\) kJ.mol\(^{-1}\) and \(-8.448\times10^5\) kJ.mol\(^{-1}\).

3.1. Ac-GG-NH\(_2\) formation on AI and AII

The termination of hydrogen bonds on the AI mechanisms occurs in glycine molecule that generates intermediate molecule \(I^4\). Termination of the hydrogen bonds in the amine group changes some molecular structural parameters, the biggest change is seen in the length of the bonding in the reaction center area (N-H), the Ea value in this step is 300.870 kJ.mol\(^{-1}\). Reactivity and stability of a molecule can be known through the energy HOMO and LUMO of a molecule. If the smaller the HOMO-LUMO gap energy, the more electron transfers occur from the HOMO (Highest Occupied Molecular Orbital) area to the LUMO (Lowest Lowest Molecular Orbital) area, which causes the molecule to become more reactive, and if the smaller the HOMO energy the more stable the molecule [20, 21]. The next step is insertion \(H^+\) ion from the previous stage. This step generates intermediate molecule \(I^2\) that contains the \(H_2O\) group, where this group is a good leaving group. This reaction step tends to release \(H_2O\) group, and the high reactivity of intermediate molecules evidences this, also involves the highest activation energy of all steps. The Ea value to form \(I^2\) from glycine on AI is 1,126.780 kJ.mol\(^{-1}\). Furthermore, the releasing \(H_2O\) group generates the \(I^2\). This release occurs in the carboxylic group. The Ea of this step is not higher than the previous steps; the value is 948.042 kJ.mol\(^{-1}\).

On AII mechanism dissociation of C-O bonds occurs in glycine, this dissociation generates intermediate molecule \(I^3\), the Ea of this step is 130.732 kJ.mol\(^{-1}\). Afterward, the insertion of OH ions into glycine and generates \(I^4\). This step requires high Ea with a value of 932.344 kJ.mol\(^{-1}\), in the next step \(H_2O\) group in \(I^4\) will be releasing its molecule. This step generates intermediate molecules \(I^1\) with Ea value 933.550 kJ.mol\(^{-1}\).

The final step of this mechanism is forming a peptide bond (C-N) from \(I^1\) and \(I^3\) on AI and AII pathway. The products are Ac-GG-NH\(_2\) with enthalpy \(-246.666\) kkal.mol\(^{-1}\) and minimum energy of \(-12.853\times10^5\) kJ.mol\(^{-1}\), the activation energy in this final step is 5.954 kJ.mol\(^{-1}\). The transition state to form product have the imaginary frequency, that is \(-263.50\) cm\(^{-1}\).

![Figure 1](image1.png)

**Figure 1.** The molecular structure of AI and AII mechanisms

Color key: yellow, carbon; purple, nitrogen; red, oxygen; blue, hydrogen

3.2. Ac-AG-NH\(_2\) formation on BI and BII
The termination of hydrogen bonds on the BI mechanism occurs in the alanine generates intermediate molecule \( I_A^1 \). The next step is insertion \( H^+ \) ion from the previous stage, occurs in glycine and generates \( I_G^2 \). The Ea is 1,105.070 kJ.mol\(^{-1}\). Furthermore, the releasing H\(_2\)O group generates the \( I_G^3 \), the Ea of this step is not higher than the previous step, the value is 926.392 kJ.mol\(^{-1}\).

Dissociation of C-O bonds on BII pathways occurs in glycine and generates intermediate molecule \( I_G^1 \). Afterward, the insertion of OH ions into alanine, this insertion generates intermediate molecules \( I_A^2 \). This step requires high Ea value 930.293 kJ.mol\(^{-1}\). The next step, H\(_2\)O group in \( I_A^2 \) will be releasing its molecule and generates \( I_A^3 \); the Ea value is 921.483 kJ.mol\(^{-1}\).

The final step of these mechanisms is forming a peptide bond (C-N) from \( I_A^1 \) and \( I_G^3 \). The products that will be produced from BI and BII is Ac-AG-NH\(_2\). The transition state to form product have the imaginary frequency, that is -709.62 cm\(^{-1}\). The activation energy (Ea) of this step is the lowest of all steps; the value is 11.584 kJ.mol\(^{-1}\).

**Figure 2.** The molecular structure of BI and BII mechanisms
Color key: yellow, carbon; purple, nitrogen; red, oxygen; blue, hydrogen

### 3. Ac-GA-NH\(_2\) formation on BIII and BIV
The termination of hydrogen bonds on the BIII mechanisms occurs in glycine molecule that generates intermediate molecule \( I_G^1 \). The next step is insertion \( H^+ \) ion from the previous stage, occurs in alanine. This step \( I_A^1 \) that contains the H\(_2\)O group, where this group is a good leaving group. The Ea value to form \( I_A^1 \) from alanine is 1,165.981 kJ.mol\(^{-1}\). Furthermore, the releasing H\(_2\)O group \( I_A^1 \) with Ea not higher than the previous step, the value is 928.06 kJ.mol\(^{-1}\).

Dissociation of C-O bonds on BIV occurs in alanine. This dissociation generates \( I_A^2 \). Afterward, the insertion of OH ions into glycine and generates \( I_G^4 \). This step requires high Ea with value 759.614 kJ.mol\(^{-1}\). The next step is H\(_2\)O group will be releasing \( I_G^4 \) molecule and \( I_A^4 \). The Ea in this step is 756.456 kJ.mol\(^{-1}\). The final step of these mechanisms is forming a peptide bond (C-N) from \( I_G^1 \) and \( I_A^4 \). The peptide product is Ac-GA-NH\(_2\). The transition state to form product have the imaginary frequency, that is -66.04 cm\(^{-1}\). The Ea of this step is the lowest of all steps that are 71.495 kJ.mol\(^{-1}\).

**Figure 3.** The molecular structure of BIII and BIV mechanisms
Color key: yellow, carbon; purple, nitrogen; red, oxygen; blue, hydrogen

The total enthalpy needed to form a product that are Ac-GG-NH\(_2\), Ac-AG-NH\(_2\), and Ac-GA-NH\(_2\) are -47.686 kcal.mol\(^{-1}\), -47.689 kcal.mol\(^{-1}\), and -47.708 kcal.mol\(^{-1}\). The determinant of the reaction rate can be determined based on the reaction stage which has the highest activation energy. Determinants of reaction rates in the AI, BI, BII, BIII, and BIV mechanisms are in the second stage,
each of which has an activation energy of 1126.780 kJ.mol\(^{-1}\), 930.293 kJ.mol\(^{-1}\), 1105.070 kJ.mol\(^{-1}\), 1165.981 kJ.mol\(^{-1}\) and 759.614 kJ.mol\(^{-1}\), meanwhile in the AII is in the third stage with activation energy of 933.550 kJ.mol\(^{-1}\). Based on the Arrhenius equation, \(k = Ae^{\frac{-Ea}{RT}}\), the higher the value of the activation energy (Ea), the smaller the value of the constant which causes the rate of the reaction to occur very slowly, \(v = k[A][B]\) [23].

The comparison between the mechanism and the same reactant, in the formation of glycine and glycine peptides that produces Ac-GG-NH\(_2\), the most preferred mechanism is the AII mechanism because the lower activation energy (Ea) that be passed in this reaction, whereas in the formation of alanine and glycine peptides that produces Ac-AG-NH\(_2\) and Ac-GA-NH\(_2\), the most preferred mechanism is the BIV mechanism that produces Ac-GA-NH\(_2\) in which the product is compatible with ADT-10 peptide (Ac-QGADTPPVG-NH2).

Figure 4. The molecular structure of products
Color key: yellow, carbon; purple, nitrogen; red, oxygen; blue, hydrogen

Figure 5. Potential Energy Surface (PES) of Reaction Mechanism of Ac-GG-NH\(_2\), Ac-AG-NH\(_2\), and Ac-GA-NH\(_2\) from mechanisms AI, BI and BII
Previously, the mechanism of peptide bond formation based on Trobro and Aqvist had also been calculated. The result shows that the reaction between glycine and alanine is easier to form Ac-GA-NH$_2$ than Ac-AG-NH$_2$, where in accordance with the ADT-10 peptide. If a comparison is made between the PES of the mechanism in this research and the PES of the mechanism based on Trobro and Aqvist reaction, the peptide bond forming reaction mechanism in this study is arduous to occur because the rate determinant $E_a$ is immense, while the rate determinant $E_a$ based on Trobro and Aqvist reaction is much lower. The rate determinant of the most preferred pathway of Ac-GG-NH$_2$, Ac-AG-NH$_2$, and Ac-GA-NH$_2$ based on the mechanism scheme of this research are 933.550 kJ.mol$^{-1}$, 930.293 kJ.mol$^{-1}$ and 759.614 kJ.mol$^{-1}$, while based on Trobro and Aqvist reaction are 16.525 kJ.mol$^{-1}$, 31.834 kJ.mol$^{-1}$, and 15.947 kJ.mol$^{-1}$. The PES of the reaction mechanism based on Trobro and Aqvist shown below.

**Figure 6.** Potential Energy Surface (PES) of Reaction Mechanism of Ac-GG-NH$_2$, Ac-AG-NH$_2$, and Ac-GA-NH$_2$ from mechanisms AII, BII and BIV

---

**Figure 7.** Potential Energy Surface (PES) of Reaction Mechanism of Ac-GG-NH$_2$, Ac-AG-NH$_2$, and Ac-GA-NH$_2$ based on Trobro and Aqvist peptide bond forming reaction mechanism
The steric effect of the methyl group in the alanine molecule that reacts with glycine to form the Ac-AG-NH$_2$ and Ac-GA-NH$_2$ peptide molecule must have an influence when being compared to the Ac-GG-NH$_2$ peptide molecule, which is composed of two glycine molecules. The influences of the steric effect of the methyl group possessed by the amino acid alanine on the peptide molecule Ac-AG-NH$_2$ and Ac-GA-NH$_2$ can be seen in the table below.

**Table 1.** Energy data of Ac-GG-NH$_2$, Ac-AG-NH$_2$, and Ac-GA-NH$_2$ molecules

| Molecule    | E (kJ.mol$^{-1}$) | Ea (kJ.mol$^{-1}$) | ∆H (kcal.mol$^{-1}$) | $E_{\text{Gap}}$ (eV) | $E_{\text{HOMO}}$ (eV) |
|-------------|-------------------|--------------------|------------------------|------------------------|------------------------|
| Ac-GG-NH$_2$ | -12.853 x 10$^5$  | 933.550            | -47.686                | 15.270                 | -10.516                |
| Ac-AG-NH$_2$ | -13.877 x 10$^5$  | 930.293            | -47.689                | 15.347                 | -10.555                |
| Ac-GA-NH$_2$ | -13.876 x 10$^5$  | 759.614            | -47.708                | 15.385                 | -10.589                |

**Table 2.** The atomic charge distribution of Ac-GG-NH$_2$, Ac-AG-NH$_2$, and Ac-GA-NH$_2$ molecules

| Molecule    | C1  | C2  | N3  | O4  | O5  | C10 | C11 | N12 | O13 |
|-------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ac-GG-NH$_2$ | 0.740 | -0.046 | -0.757 | -0.594 | -0.567 | 0.763 | -0.099 | -0.707 | -0.623 |
| Ac-AG-NH$_2$ | 0.770 | -0.007 | -0.747 | -0.616 | -0.555 | 0.766 | -0.099 | -0.707 | -0.624 |
| Ac-GA-NH$_2$ | 0.745 | -0.069 | -0.754 | -0.611 | -0.551 | 0.789 | -0.024 | -0.695 | -0.627 |

**Table 3.** The atomic bond length of Ac-GG-NH$_2$, Ac-AG-NH$_2$, and Ac-GA-NH$_2$ molecule

| Molecule    | C2-C1 | C2-N3 | O4-C1 | O5-C1 | O5-C10-N3 | C11-C10 | C11-C10-N12 | O13-C10 | O13-C10-N12 |
|-------------|-------|-------|-------|-------|------------|----------|---------------|---------|-----------|
| Ac-GG-NH$_2$ | 1.505 | 1.435 | 1.321 | 1.188 | 1.347      | 1.515    | 1.443         | 1.202   |           |
| Ac-AG-NH$_2$ | 1.514 | 1.447 | 1.331 | 1.185 | 1.349      | 1.519    | 1.443         | 1.202   |           |
| Ac-GA-NH$_2$ | 1.509 | 1.440 | 1.329 | 1.185 | 1.350      | 1.524    | 1.449         | 1.202   |           |

**Table 4.** The atomic bond angle atom of Ac-GG-NH$_2$, Ac-AG-NH$_2$, and Ac-GA-NH$_2$ molecule

| Molecule    | N3-C2-C1 | O4-C1-C2 | O5-C1-C2 | C10-N3-C2 | C11-C10-N3-C2 | N12-C11-C10-N12 | O13-C10-N12-C11 |
|-------------|----------|----------|----------|-----------|---------------|-----------------|-----------------|
| Ac-GG-NH$_2$ | 113      | 113      | 123      | 120       | 116           | 106             | 120             |
| Ac-AG-NH$_2$ | 111      | 113      | 123      | 121       | 115           | 109             | 121             |
| Ac-GA-NH$_2$ | 109      | 111      | 125      | 120       | 115           | 109             | 121             |

**Table 5.** The atomic dihedral angle of Ac-GG-NH$_2$, Ac-AG-NH$_2$, and Ac-GA-NH$_2$ molecule

| Molecule    | O4-C1-N3-C2-N3 | O5-C1-N3-C2-N3 | C10-N3-C2-C1 | C11-C10-N3-C2 | N12-C11-C10-N12 | O13-C10-N12-C11 |
|-------------|----------------|----------------|--------------|---------------|-----------------|-----------------|
| Ac-GG-NH$_2$ | -19            | 162            | -161         | -179          | 129             | -52             |
| Ac-AG-NH$_2$ | -25            | 157            | -150         | -176          | -148            | 34              |
| Ac-GA-NH$_2$ | 179            | -1             | 178          | -6            | 148             | -34             |

4. Conclusion

The most preferred mechanism pathway in peptide forming reaction between glycine and glycine is AII Pathway with rate determinant of 933.550 kJ.mol$^{-1}$. Meanwhile in peptide forming reaction between alanine and glycine is a BIV Pathway that generates product Ac-GA-NH$_2$ with rate determinant 759.614 kJ.mol$^{-1}$. The rate determinant of all pathway AII, AIII, BI, BII, BIII, and BIV are
1126.780 kJ.mol\(^{-1}\), 933.550 kJ.mol\(^{-1}\), 1105.070 kJ.mol\(^{-1}\), 930.293 kJ.mol\(^{-1}\), 1165.981 kJ.mol\(^{-1}\), and 759.614 kJ.mol\(^{-1}\). The steric effect decrease a rate determinant activation energy and the steric effect also made the different geometry structure that evidenced by the dihedral angle value.

**Acknowledgments**

A gratefully acknowledges to Prof. Teruna J. Siahaan, Ph.D. (Department of Pharmaceutical Chemistry The University of Kansas, US) for each direction and discussion on the drug delivery system. Likewise to Prof. Krzysztof Kuczera, Ph.D. (Department of Chemistry The University of Kansas, US) for an in-depth discussion of computational modeling. Thanks to Indonesian Directorate General of Higher Education, which has funded this research and technology Funding scheme 2018.

**References**

[1] E. Sinaga et al., *Pharm. Res.* 2002. **19**, 1170–1179
[2] K. L. Lutz and T. J. Siahaan. 1997. *J. Pharm. Sci.* **86**, 977–984
[3] S. Chandrudu, P. Simerska, and I. Toth. 2013. *Molecules*, **18**, 4373–4388
[4] V. R. Pattabiraman and J. W. Bode. 2011. *Nature*, **480**, 471–479
[5] P. E. Morgan, D. I. Pattison, and M. J. Davies. 2012. *Free Radic. Biol. Med.*, **52**, 328–339
[6] Q. Wang, J. Gao, D. Zhang, and C. Liu. 2015. *Chem. Phys.*, **450–451**, 1–11
[7] S. P. 2017. *Material Sci. Eng.*, **172**, 0–9
[8] N. Sultana and W. M. F. Fabian. 2013. *Beilstein J. Org. Chem.*, **9**, 594–601
[9] O. Antohi and A. M. Sapse. 1998. *J. Mol. Struct. THEOCHEM*, **430**, 247–258
[10] O. Antohi, F. Naider, and A. M. Sapse. 1996. *J. Mol. Struct. THEOCHEM*, **360**, 99–108
[11] D. Chakraborty and S. Manogaran. 1998. *Chem. Phys. Lett.*, **294**, 56–64
[12] R. J. Nielsen, J. M. Keith, B. M. Stoltz, and W. A. Goddard. 2004. A *Computational Model Relating Structure and Reactivity in Enantioselective Oxidations of Secondary Alcohols by ( - ) -Sparteine - Pd II Complexes*. **10**, 7967–7974
[13] H. M. Lee and J. C. Choe. 2017. *Chem. Phys. Lett.*, **675**, 6–10
[14] S. Trobro and J. Aqvist. 2005. *Proc. Natl. Acad. Sci.*, **102**, 12395–12400
[15] F. Jensen. 1999. *Introduction to computational chemistry*. 2–444
[16] Cramer C J. 2004. *Essentials of Computational Chemistry: Theory and Models*. John Wiley and Sons, London. 1–579
[17] James., J. B. Foresman, and A. Frisch. 2013. *Exploring Chemistry with Electronic Structure Methods*, 2013, 335
[18] M. Valiev et al. 2010. *Comput. Phys. Commun.*, **181**, 1477–1489
[19] T. Oie, G. H. Loew, S. K. Burt, R. D. Macelroys, and L. Ag. 1984. *Quantum Chemical Studies of a Model for Peptide Bond Formation. 3. Role of Magnesium Cation in Formation of Amide and Water from Ammonia and Glycine*. **106**, 1–7
[20] M. Khajehzadeh and N. Sadeghi. 2018. *J. Mol. Liq.*, **256**, 238–246
[21] P. Sivageyanthi, M. Jeewaraj, K. Balasubramani, V. Viswanathan, and D. Velmurugan. 2017. *Chem. Data Collect.*, **11–12**, 220–231
[22] E. V. Anslyn and D. A. Dougherty. 2006. *Modern physical organic chemistry*. 43, 6
[23] J. D. P. Peter Atkins. 2010. *Physical Chemistry 9th Edition*