



## **Prediction of RNA Secondary Structure at IVS1 Mutation of Beta Globin Gene**

Nur Imaniati Sumantri<sup>1</sup>, Dian Rachma Wijayanti<sup>2</sup>

<sup>1</sup>Biomedical Engineering, Department of Electrical Engineering, Faculty of Engineering, University of Indonesia, Depok, Indonesia – 16424

<sup>2</sup>Biotechnology Department, Faculty of Science and Technology, Binawan University, Jakarta, Indonesia– 13630

**Abstract : Background:** Beta globin gene is responsible for producing beta globin chains that stabilize the structure and function of hemoglobin. This gene expression is controlled by complex interactions of transcriptions factors and its regulatory elements in a specific manner. Disturbed beta globin genes may result in hemoglobinopathies, mainly sickle cell disease and beta thalassemia. It seems interesting that several mutations occurring in intronic region results in severe symptoms to beta thalassemia patients, such an IVS1nt5 G>C. This research aimed to analyze RNA structural alteration effected by intronic mutation of beta thalassemia. **Methods:** The most prevalent mutation of beta thalassemia in Indonesia was obtained from Ithantet. The RNA secondary structure of IVS1nt5 G>C and beta globin gen (HBB) wildtype were performed by RNAstructure, along with probknot prediction. **Results:** The result showed that intronic mutation caused conformational change in beta globin secondary structure, either for max expect or base pairing probability approach. The mutant had bigger and more loops that diminished the protein stability. Thus, the structure might undergo dysfunction. **Conclusion:** The comprehensive structural-functional significance of these findings needs further study.

**Keywords :** Beta thalassemia; IVS1nt5;RNA elucidation; intron mutation.

### **Introduction**

The globin genes are located in different chromosomal loci in sequence referring their expression during development at two separate gene clusters. Globin gene expression is controlled by complex interactions of transcriptions factors and its regulatory elements in a specific manner. Disturbed globin genes may result in hemoglobinopathies sickle cell anemia or thalassemia (Farashi S and Harteveld CL, 2018). Beta-thalassemia occurs as reduction or absence synthesis of beta globin chain of hemoglobin (Hb) tetramer, which normally composed of two alpha globin and two beta globin chains. The imbalance between alpha globin and beta globin chain, severity of anemia, and clinical picture determine wether the patient is in minor, intermediate, or major beta thalassemia group (Taher et al, 2013).

**Nur Imaniati Sumantri *et al* /International Journal of ChemTech Research, 2020,13(1): 247-252.**

**DOI= <http://dx.doi.org/10.20902/IJCTR.2019.130131>**

High prevalence of beta-thalassemia cases has been reported of 2,750 cases in America, 39,700 cases in Eastern Mediterranean, 16,230 in Europe, 35,500 cases in South-East Asia, and 3,450 cases in Western Pacific. Approximately around 40,000 new born babies inherited beta-thalassemia annually (Modell B and Darlison M, 2008). Nearly 300 beta-thalassemia alleles have been characterized. The majorities of beta-thalassemia are caused by mutations of at least one or limited number of nucleotide inside the beta-thalassemia or its immediate flanking region (Giardine et al, 2011). Beta-thalassemia major mutations might interfere the primary mRNA transcription. These affect uniform dinucleotides at the exon-intron splice junction. Thus, eventually dissolve regular splicing and produce  $\beta^0$ -thalassemia result in  $\beta$ -thalassemia major phenotype (Thein SL, 2013).

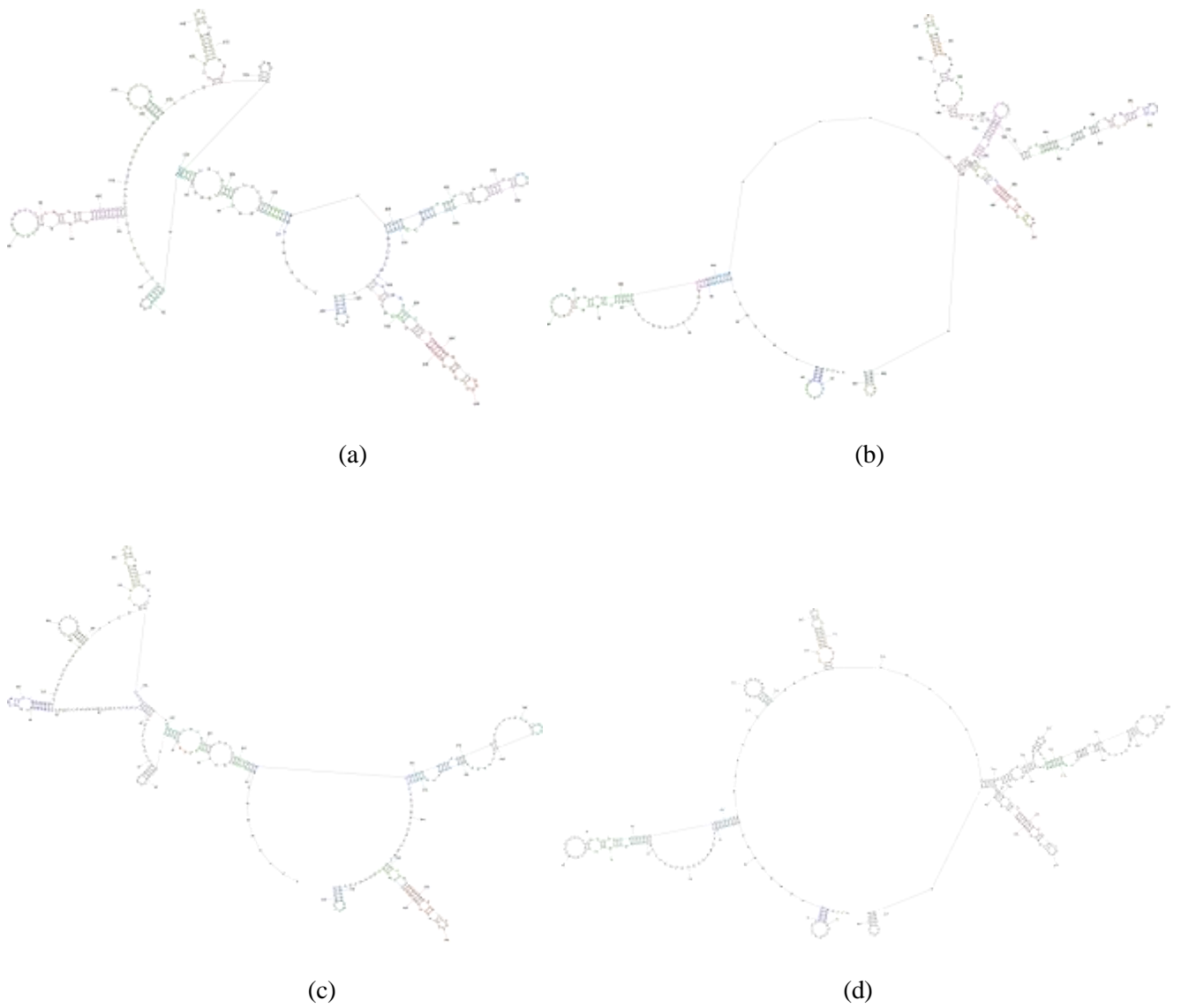
The previous research reported that IVS1nt5 is the most prevalent mutation for beta thalassemia in Indonesia (Maskoen et al, 2017). It is still unclear how intronic mutation provoke moderate, even heavy clinical symptoms to the thalassemia patients. Therefore, we aimed to analyze RNA structural alteration effected by intronic mutation of beta thalassemia that might disrupt hemoglobin functions. There are several tools to perform RNA structure prediction. We preferred using RNAstructure (<http://rna.urmc.rochester.edu>) as a software package because it offers secondary structure prediction, including free energy minimization, maximum expected accuracy structure prediction and pseudoknot prediction in easy way within the short time (Bellaousov et al, 2013).

## Methods

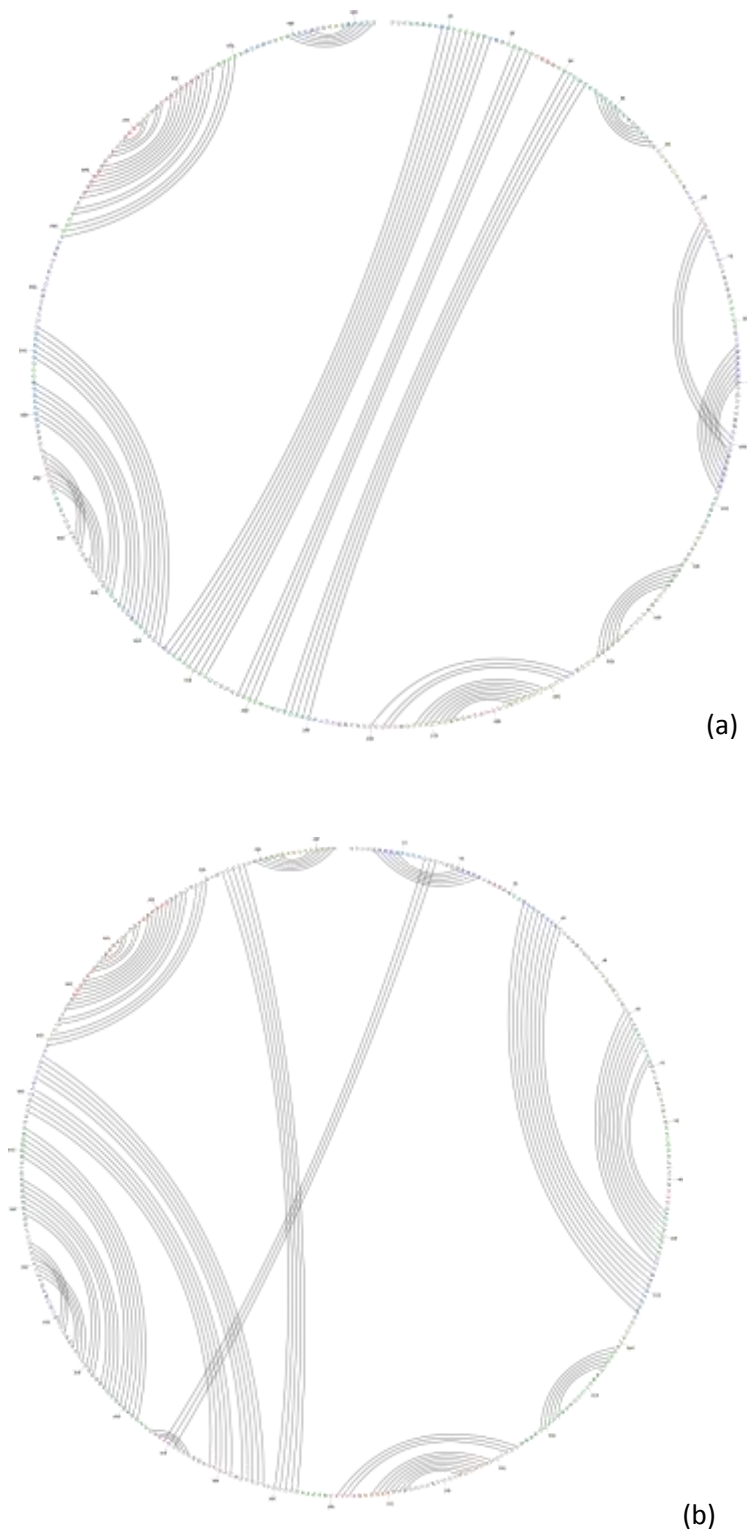
Basic data for high prevalent of beta thalassemia was obtained from Ithanel ([www.ithanel.eu](http://www.ithanel.eu)) as portal for hemoglobinopathies. The server showed that IVS1nt5 G>C was the highest occurrence in Indonesia. Complete genome sequence of beta globin gene (HBB) was collected from Ensembl ([www.asia.ensembl.org](http://www.asia.ensembl.org)) (transcript HBB-201 ENST00000335295.4). This gene consists of 1,608 nucleotides divided into 3 exon and 2 intron (intervening sequence/IVS) regions. The sequence was point mutated manually at nucleotide number 147 from Guanine (G) to Cytosine (C). Of 353 nucleotides in IVS1 and EXON 2 region were transcribed into mRNA using Nucleic Acid Converter ([www.skaminsky115.github.io](http://www.skaminsky115.github.io)), then structurally analyzed using RNAstructure (<http://rna.urmc.rochester.edu>). We performed secondary structure by fold, secondary structure max expect, and probknot prediction in default settings of the program.

## Results

Intronic mutation of beta globin gene was observed. The 353 nucleotides length, from intron 1-2 position 1 towards beta chain gene position 223, was predicted by RNAstructure web server to analyze the structural distinction through secondary structure, partition, and probknot prediction.



**Figure 1. Secondary structure fold results (a) HBB wildtype (b) IVS1nt5 G>C; Secondary structure max expect (c) HBB WT, (d) IVS1nt5 G>C. Prediction of the structures is represented by color for the base pairing probability**



**Figure 2. Probknot prediction (a) HBB WT (b) IVS1nt5 G>C**

### **Discussion**

Beta thalassemia becomes a concern in genetic disease nowadays. One common mutation in thalassemia belt countries, especially in Indonesia, is caused by intronic mutation, IVS1nt5 G>C. Finding information about

intronic mutant mechanism on genetic disease is scarce. RNA analysis could be an entrance gate to explore its effect through structural alteration. The variation of sequence, either at coding or noncoding sequence, would influence amino acid and protein formation. Wildtype of beta globin protein is consisted of heme, iron, oxygen binding sites, and oxygen transporter. While in several mutations causing hemoglobinopathy undergoes unstable structure (Schepps KG et al, 2018).

The IVS1nt5 G>C could be included in snoRNA, the non coding RNA located in intronic region with 60-300 bp length. This RNA might function on ribosomal RNA processing, alternative splicing, or any unknown function (Esteller M, 2011). The results showed the distinction of RNA secondary structure between wildtype and IVS1nt5 G>C suggested conformational change, hence IVS1nt5 G>C is regarded as a causative form of beta thalassemia major. The same mutations can generate different phenotype, e.g. IVS1nt5 G>C is known to generate either intermediate or major beta thalassemia disease. The regulatory region of RNA can be indicated by a global analysis of the structural consequences of all disease-associated mutations on a regulatory RNA (Halvorsen M et al, 2010). The *in silico* analysis for mutant intronic region faces very limited tools to access, since a numbers of tools focus on protein prediction solely. The prediction of pathogenicity of intronic sites leads to detailed investigation considering the phenotypic diversity in beta thalassemia (Mahdiah and Rabbani, 2016).

This study described the alteration of IVS1nt5 G>C RNA secondary structure which might give impact on phenotypic characterization of patients. The secondary structure of this intronic mutant showed bigger interior loop compared to the wildtype. The interior loop was constructed by several base pairs which failed to separate the two bulged loop (Crowther, et al, 2017). The more and the bigger shape of the loop will destabilize the structure itself. Hairpin loops and interior loops (including the special cases of bulges and stacked pairs) can be dealt with by a simple modification of the energy rules. An advance capability to distinguish between true hairpin and interior loops with closing pair and loops containing the cut in their backbone is useful to conduct partition function calculations and the generation of suboptimal structures (Berhart SH, et al, 2006).

Secondary structure prediction helps us to understand the mechanism of action of RNA (Bellaousov et al, 2013). The secondary structure fold result was performed by calculating and selecting the best thermodynamics parameter. The sequence of neighboring base pairs define RNA secondary structure stability. The secondary structure prediction of Max Expect emphasizes on maximizing the expected base pair accuracy. This structure maximizes a score that balances the probabilities of base pairing and being unpaired (Lu, et al, 2009). This methods improve the accuracy of structure prediction for all types of noncoding RNAs tested with the exception of signal recognition particle RNA (Lu et al, 2009).

The prediction of secondary structure was conducted quantitatively. Structural motifs composed of highly probable base pairs are likely to be in the actual structure. The average fraction of correctly predicted pairs in lowest free energy structures increased from 65.8 to 91.0 % when only the base pairs with high probabilities ( $\geq 0.99$ ) were considered (Mathews, 2004). On the other side, ProbKnot analysis was conducted to support secondary structure conformation by thermodynamic ensemble. This approach is a simple yet powerful algorithm to predict pseudoknotted RNA structures. It assembles maximum expected accuracy structures from base-pairing probabilities computed from a non-pseudoknotted partition function (Bellaousov and Mathews, 2010; Xu and Mathews, 2016).

## Conclusion

The prediction was conducted by calculating for the possible secondary using RNAstructure web server. The finding showed different secondary structure between beta globin wildtype and IVS1nt5 G>C. The result indicated that intronic point mutation causes structural alteration on their secondary structure and may effects a huge influence on beta globin gene activity disturbance. The structural-functional significance of these findings needs further study.

## Conflict of Interests

The authors stated that they had no interests which might be perceived as posing a conflict or bias.

## References

1. Bellaousov S, et al. RNAstructure: web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Research*. 2013. 41.
2. Bellaousov S, Mathews DH. ProbKnot: fast prediction of RNA secondary structure including pseudoknots. *RNA*. 2010. 16:1870–1880.
3. Bernhart SH, et al. Partition function and base pairing probabilities of RNA heterodimers. *Algorithms for Molecular Biology*. 2006. 1:3. doi:10.1186/1748-7188-1-3
4. Crowther CV, et al. Influence of two bulge loops on the stability of RNA duplexes. *RNA*. 2017. 23:217–228. doi:10.1261/rna.056168.1.
5. Esteller M. Non-coding RNAs in human disease. *Nature Reviews Genetics*. 2011. 12(12):861-74. doi: 10.1038/nrg3074.
6. Farashi S, Harteveld CL. Molecular Basis of  $\alpha$  thalassemia. *Blood Cells, Molecules and Diseases*. 2018. 70: 43-53. doi: 10.1016/j.bcmd.2017.09.004
7. Giardine B, et al. Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach. *Nature Genetics*. 2011. 43:295–301. doi: 10.1038/ng.785
8. Halvorsen M, Martin JS, Broadaway S, Laederach A. Disease-associated mutations that alter the RNA structural ensemble. *PLoS Genetics*. 2010. 6(8): e1001074. doi:10.1371/journal.pgen.1001074.
9. Lu ZJ, et al. Improved RNA secondary structure prediction by maximizing expected pair accuracy. *RNA*. 2009. 15:1805–1813.
10. Mahdiah N, Rabbani B. Beta thalassemia in 31,734 cases with HBB gene mutations: Pathogenic and structural analysis of the common mutations; Iran as the crossroads of the Middle East. *Blood Reviews*. 2016. 30:493–508.
11. Maskoen AM, et al. Mutation Spectrum of  $\beta$ -globin gene in thalassemia patients at Hassan Sadikin West Java Indonesia. *Cellular and Molecular Biology*. 2017. 63. doi: 10.14715/cmb/2017.63.12.6
12. Mathews DH. Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA*. 2004. 10:1178–1190.
13. Modell B, Darlison M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bulletin World Health Organization*. 2008. 86(6):480-487. doi:10.2471/BLT.06.036673.
14. Scheps KG, et al. Two novel unstable hemoglobin variants due to in-frame deletions of key amino acids in the  $\beta$ - globin chain. *European Journal of Haematology*. 2018. 100:529–535.
15. Taher AT, et al. Thalassemia. *Lancet*. 2018. 391:155-167. doi: 10.1016/S0140-6736(17)31822-6.
16. Thein SL. The Molecular Basis of  $\beta$ -Thalassemia. *Cold Spring Harbor Perspectives in Medicine*. 2013. 3. doi:10.1101/cshperspect.a011700.
17. Xu Z.Z., Mathews D.H. Secondary Structure Prediction of Single Sequences Using RNAstructure. In: Turner D., Mathews D. (eds) *RNA Structure Determination. Methods in Molecular Biology*, vol 1490. Humana Press. New York. 2016. doi: 10.1007/978-1-4939-6433-8\_2

\*\*\*\*\*