Haemoglobin switching modulator SNPs rs5006884 is associated with increased HbA₂ in β-thalassaemia carriers

Cyril Cyrus¹, Chittibabu Vatte¹, Shahanas Chathoth¹, Abdul Azeez Sayed², J. Francis Borgio², Mohammed Abdullah Alrubaish¹, Rawan Alfalakah¹, Jana Alsaikhan³, Amein K. Al Ali¹

¹Department of Biochemistry, College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
²Department of Genetic Research, Institute for Research and Medical Consultation, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
³King Fahd Hospital of the University, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Submitted: 28 March 2018; Accepted: 15 May 2018
Online publication: 18 July 2019

Arch Med Sci 2021; 17 (4): 1064–1074
DOI: https://doi.org/10.5114/aoms.2019.86705
Copyright © 2019 Termedia & Banach

Abstract

Introduction: Haemoglobin A₂ (HbA₂), the tetramer of α- and δ-globin chains, is used as a diagnostic biomarker for β-thalassaemia carriers. The HbA₂ levels are regulated by the presence of HPFH, δ-thalassaemia, HbA₁/₂ gene triplication, and variants of KLF1, β-globin gene, and HbF regulating QTLs. Saudi Arabia has a high incidence of borderline HbA₂ levels, thereby making it difficult to classify the haemoglobinopathies. This study aims to investigate the association of known HbF enhancer QTL gene SNPs with HbA₂ levels.

Material and methods: 14 Specific SNPs in BCL11A, HMIP, OR51B6, HBBP1, and HBG2 loci were genotyped in 164 Saudi β-thalassaemia carriers by TaqMan assay to validate their role as regulators of HbA₂ levels. HbA₂ levels were determined using the Variant II β-Thalassemia Short Program Recorder kit. The non-random association of these SNPs was tested using HaploView software. Protein interaction was assessed using 3D structure modelling for OR51B6 (rs5006884), comparative energy minimisation, and root-mean-square deviation (RMSD) prediction.

Results: Elevated HbA₂ levels were associated with SNPs in HBBP1, OR51B6, and HBG2 promoter regions. The bioinformatics modelling and prediction revealed that the exonic rs5006884 had RMSD value deviations and significantly varied binding energy minimisation. α-globin variations were found in 57.92% of individuals but were not associated with elevated HbA₂.

Conclusions: The haemoglobin switching modulators rs2071348, rs7482144, and rs5006884 are involved in regulation of HbA₂ level with rs5006884 influencing the tetramer formation. Screening for haemoglobinopathies should take these SNPs into consideration, specifically in borderline HbA₂ cases. Assiduous analysis of rs5006884 as HbA₂ modulator for amelioration of disease severity is recommended.

Key words: β-thalassaemia, haemoglobin A₂, SNPs, OR51B6, tetramer.
Introduction

Haemoglobin A₂ (HbA₂) is a tetramer of α- and δ-globin chains, which constitutes approximately 2–3% of the total circulating haemoglobin. The HBD-coded HbA₂ prevents the polymerisation of deoxy-sickle haemoglobin and is a reliable diagnostic biomarker for β-thalassaemia trait. The globin genes are arranged in the order of their expression and are activated sequentially as embryonic (HBE1, ε), foetal (HBG1/2, γ), and adult (HBD/HBB, δ/β) genes during development. The globin promoters initiate the molecular switching between these genes by gaining access to the upstream regulatory elements, namely the β locus control region (β-LCR) [1] and autonomous silencing of the preceding gene. Thus, the gene closest to the β-LCR is activated first (HBE1) and the furthest (HBB) is activated last. Inactivation of the HBB promoter results in extended interaction of β-LCR with the HBD and HBG promoters, thereby leading to elevated HbA₁ and HbF levels. The deficit of β-globin and profusion of α-globin results in the post-translational elevation of HbA₁ levels in β-thalassaemia carriers [2]. However, HbA₂, is not characteristically elevated in all β-thalassaemia carriers, making it difficult to classify haemoglobinopathies in patients with borderline HbA₂ levels. Saudi Arabia has a high incidence of borderline HbA₂ levels. Saudi Arabia has a high incidence of borderline HbA₂ levels. HbF level-associated QTLs are differentially distributed among populations and a reciprocal relationship among the HbA₂, and HbF levels has been inferred [12, 13]. A reciprocal relationship among the HbA₂, and HbF levels has been inferred among populations [12, 13]. A reciprocal relationship among the HbA₂, and HbF levels has been inferred in acquired disorders [10].

The aim of this study was to investigate the association of HbA₂ level with 14 SNPs of the HbF enhancer QTL genes. The molecular interactions involved on the variations in the level of HbA₂ synthesis is not completely understood. Shedding light on the role of SNPs that act as triggering factors of HbA₂ may indicate that inducing elevated levels of HbA₂ could be of clinical importance.

Material and methods

Ethical approval

This study was approved by the Ethical Committee of Imam Abdulrahman Bin Faisal University in accordance with the 1964 Helsinki Declaration and its later amendments. Written, informed consent was obtained from each participant.

Pheno- and genotyping

This study was conducted on 164 β-thalassaemia carriers residing in the Eastern Province of Saudi Arabia. The blood samples were collected in EDTA anti-coagulated vacutainers for HbA₁ level estimation using Bio-Rad Variant II (Variant II β-Thalassaemia Short Program Recorder kit) and DNA extraction using blood mini kit (Qiagen, USA). SNP genotyping of 14 SNPs, namely rs2071348, rs7482144, rs5006884 (HGB2 promoter region), rs766432, rs11886868, rs4671393, rs7557939 (BCL11A region), rs28384513, rs9376090, rs9399137, rs4895441, rs9389269, rs9402686, and rs9494142 (HMIP region) was carried out by nuclease allelic discrimination assay. The target-specific forward and reverse primers along with TaqMan probes labelled with VIC and FAM for each allele (TaqMan Assay, Applied Biosystems) were amplified on the ABI 7500 real time PCR system (ABI, Foster City, USA) according to the manufacturer’s instructions.

The α-3.7 deletion and variations were identified using ViennaLab StripAssays and PCR-Sequencing method. The globin gene was amplified as previously described [14]. In summary, the amplicons were purified using QiAquick PCR Purification Kit (Qiagen, Hilden, Germany), and the purified PCR products were cycle sequenced using a Big Dye Terminator Cycle Sequencing Kit (Applied Biosystem, Foster City, CA, USA). The cycle-sequenced products were then purified and
electrophoresed in a Genetic Analyzer 3500 (Life Technologies Corporation, Carlsbad, CA, USA). Sequencing Analysis Software Version 5.4 (Applied Biosystem, USA) was used for data analysis.

Structure modelling and root-mean-square deviation (RMSD) prediction

Three-dimensional (3D) protein structure of OR51B6 native chain was designed using automated homology modelling [15] and the mutated OR51B6 protein using SWISS MODEL (http://swissmodel.expasy.org/). The 3D structure was generated based on template Protein Data Bank (PDB) ID: 5cxv with highest resolution of 2.7 Å. PROCHECK (http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/) online was used to validate the native and modelled structure [16]. The mutant models were generated using SWISS PDB Viewer Ver. 4.1 (http://spdbv.vital-it.ch/) [17]. The energy minimisation and its RMSD were checked for the native and mutants using the GROMACS 5.1.1 (http://www.gromacs.org/) program [18].

Protein-protein interaction

Protein-protein interaction was conducted using a PRISM server [19–21] because it predicts the interface structures between two interacting proteins. We submitted the haemoglobin proteins with both OR51B6 wild and mutant models to check the possible interaction variations between the wild and mutant proteins towards the haemoglobin proteins.

Statistical analysis

Hardy-Weinberg equilibrium was tested for all the SNPs, and \( \chi^2 \) and odds ratio were determined by SPSS ver. 19 to evaluate allele association. Linkage disequilibrium (LD) test was carried out using HaploView 4.2 software [22] to identify the non-random association of these 14 SNPs. Haplotype blocks were constructed using HaploView 4.2 program. Haplotypes associated with the study subjects were inferred based on the partition-likelihood approach through an EM algorithm. A \( p \)-value < 0.05 was considered significant for all statistical analyses.

Results

The independent segregation genotype for all the SNPs was in agreement with the Hardy-Weinberg equilibrium. Standard allelic association analysis of the 14 SNPs tested in the patient cohort showed that only three SNPs, namely rs2071348 (HBBP1), rs7482144 (HBG2), and rs5006884 (OR51B6), were significantly associated with borderline and higher HbA\(_2\) levels. The rs5006884 (\( p = 0.02; \chi^2 = 5.08 \)) was the key SNP associated with elevated HbA\(_2\) levels compared to the normal and borderline levels. There were no significant differences in allele frequencies of SNPs in the HBS1L-MYB region and BCL11A region between the normal and HbA\(_2\)-elevated cohorts (Table I, Figure 1).

The predominant haplotypes in the borderline and higher HbA\(_2\) cohort consisted of GTC, TCC, and TCT in the HBG2 promoter region comprising SNPs rs2071348 (HBBP1 gene), rs7482144 (HBG2 gene), and rs5006884 (OR51B6), respectively (Table II, Figure 2). The TCT (rs2071348T, rs7482144C, rs5006884T) haplotype pattern was the most significant HbA2 enhancer haplotype (\( p = 0.01; \chi^2 = 7.8 \)) (Table II).

The \( \alpha \) globin variations

The detailed study on the \( \alpha \) globin genotype revealed that the \(-\alpha_{\beta}\) is the most frequent, with an overall prevalence of 57.92% (\( n = 95 \)) of Saudi individuals included in the study. The other \( \alpha \) globin variations included \( \alpha_4^{A2}/\alpha_4\alpha_2, \alpha_4^{A2}/\alpha_4^{A2}, \alpha_4^{A2}/\alpha_4^{A2}, \text{ and } -3.7\alpha_2/\alpha_1^{\text{polyk}1}\alpha_2\). None of these variants were found to be significantly associated with the level of HbA\(_2\).
Haemoglobin switching modulator SNPs rs5006884 is associated with increased HbA2 in β-thalassaemia carriers

Figure 2. Haploview output of LD across 14 SNPs from the genotyping data in the Saudi population. Linkage Disequilibrium (LD) analysis patterns between 3 HBG2, 4 BCL11A, and 7 HMIP SNPs compared in thalassaemia patients. The pairwise correlation between the SNPs were measured as $r^2$ and shown (x100) in each diamond. Enhancer haplotypes are in red boxes and diminisher haplotypes are in blue boxes. Coordinates are according to the NCBI build dbSNP 144 Homo sapiens annotation release 107 (reference sequence NT_025741.16)
| HbA2 vs. SNP | HBG2 promoter region | BCL11A region | HMIP region |
|--------------|-----------------------|----------------|-------------|
|              | rs2071348             | rs7482144      | rs5006884    |
|              | rs766432              | rs1188688      | rs4671393    |
|              | rs47557939            | rs28384513     | rs9376090    |
|              | rs9399137             | rs4895441      | rs9389269    |
|              | rs9402686             | rs9494142      |             |
| Normal vs. 2.9–3.1 | Associated allele | G   | T   | C   | C   | T   | A   | A   | T   | T   | T   | A   | T   | G   | T   |
|               | $\chi^2$             | 10.30 | 5.23 | 0.78 | 0.24 | 0.03 | 0.39 | 0.034 | 0.41 | 0.62 | 1.80 | 0.62 | 0.98 | 0.98 | 0.49 |
|               | P-value               | 0.001* | 0.02* | 0.38 | 0.63 | 0.85 | 0.53 | 0.854 | 0.52 | 0.43 | 0.18 | 0.43 | 0.32 | 0.32 | 0.49 |
| Normal vs. 3.2–3.9 | Associated allele | G   | T   | T   | A   | T   | G   | G   | T   | C   | C   | C   | G   | C   | A   | C   |
|               | $\chi^2$             | 6.57  | 2.25 | 3.31 | 0.07 | 0.03 | 0.02 | 0.006 | 1.94 | 0.14 | 0.73 | 0.73 | 0.35 | 0.35 | 0.13 |
|               | P-value               | 0.01*  | 0.13  | 0.07  | 0.79  | 0.86  | 0.89  | 0.939  | 0.16  | 0.71  | 0.39  | 0.39  | 0.55  | 0.55  | 0.72  |
| Normal vs. 4.2 | Associated allele | T   | C   | T   | C   | T   | A   | A   | T   | T   | T   | A   | T   | G   | T   |
|               | $\chi^2$             | 3.41  | 0.81 | 5.08 | 0.00 | 0.50 | 0.01 | 0.497 | 0.00 | 0.70  | 0.70  | 0.70  | 0.81  | 0.81  | 0.92  |
|               | P-value               | 0.06   | 0.37  | 0.02* | 1.00 | 0.48  | 0.94  | 0.481  | 0.97 | 0.40  | 0.40  | 0.40  | 0.37  | 0.37  | 0.34  |
| 2.9–3.1 vs. 3.2–3.9 | Associated allele | T   | C   | T   | A   | –   | G   | G   | T   | C   | C   | G   | C   | A   | C   |
|               | $\chi^2$             | 0.18  | 0.29 | 4.32 | 0.35 | 0.00 | 0.35 | 0.042 | 0.46 | 0.88  | 3.17  | 1.64  | 1.64  | 1.64  | 0.72  |
|               | P-value               | 0.67   | 0.59  | 0.04* | 0.56 | 1.00  | 0.56  | 0.838  | 0.50 | 0.35  | 0.08  | 0.20  | 0.20  | 0.20  | 0.40  |
| 2.9–3.1 vs. 4.2 | Associated allele | G   | T   | C   | C   | C   | A   | G   | T   | C   | C   | G   | C   | A   | C   |
|               | $\chi^2$             | 8.11  | 2.80 | 6.45 | 0.04 | 0.34 | 0.04 | 0.338 | 0.07 | 0.38  | 0.19  | 0.38  | 0.38  | 0.38  | 0.58  |
|               | P-value               | 0.004* | 0.09  | 0.01* | 0.84 | 0.56  | 0.84  | 0.561  | 0.80 | 0.54  | 0.66  | 0.54  | 0.54  | 0.54  | 0.45  |
| 3.2–3.9 vs. 4.2 | Associated allele | T   | C   | T   | C   | T   | A   | A   | G   | T   | T   | A   | T   | G   | T   |
|               | $\chi^2$             | 6.95  | 1.99 | 1.19 | 0.02 | 0.33 | 0.02 | 0.48  | 0.46 | 0.86  | 1.09  | 1.09  | 1.09  | 1.09  | 1.09  |
|               | P-value               | 0.01*  | 0.16  | 0.28  | 0.90  | 0.56  | 0.90  | 0.488  | 0.50 | 0.35  | 0.30  | 0.30  | 0.30  | 0.30  | 0.30  |
Haemoglobin switching modulator SNPs rs5006884 is associated with increased HbA2 in β-thalassaemia carriers.

| Block                  | Haplotype | Normal vs. 2.9–3.1 | Normal vs. 3.2–3.9 | Normal vs. 4.2 | 2.9–3.1 vs. 3.2–3.9 | 2.9–3.1 vs. 4.2 | 3.2–3.9 vs. 4.2 |
|------------------------|-----------|--------------------|--------------------|----------------|---------------------|----------------|----------------|
|                        |           | Freq.  | χ²  | p   | Freq.  | χ²  | p   | Freq.  | χ²  | p   | Freq.  | χ²  | p   | Freq.  | χ²  | p   |
| HBG2 promoter region    | TCC       | 0.526  | 3.02 | 0.08 | 0.53  | 4.54 | 0.03 | 0.548  | 0.70 | 0.40 | 0.423  | 0.16 | 0.69 | 0.355  | 0.16 | 0.69 |
|                        | GTC       | 0.156  | 6.52 | 0.01*| 0.14  | 1.27 | 0.26 | 0.123  | 1.46 | 0.23 | 0.232  | 0.72 | 0.40 | 0.29   | 4.85 | 0.03*|
|                        | TCT       | 0.134  | 2.24 | 0.13 | 0.145 | 0.12 | 0.73 | 0.163  | 7.80 | 0.01*| 0.103  | 2.80 | 0.09 | 0.201  | 6.15 | 0.01*|
|                        | GCC       | 0.09   | 1.83 | 0.18 | 0.073 | 0.29 | 0.59 | 0.074  | 0.84 | 0.36 | 0.102  | 0.53 | 0.47 | 0.116  | 1.55 | 0.21 |
|                        | GTC       | 0.057  | 0.49 | 0.48 | 0.065 | 3.67 | 0.06 | 0.048  | 0.52 | 0.47 | 0.102  | 0.43 | 0.51 | 0.033  | 0.40 | 0.53 |
|                        | TTC       | 0.038  | 0.26 | 0.61 | 0.034 | 1.53 | 0.22 | 0.041  | 0.42 | 0.52 | 0.012  | 0.80 | 0.37 | 0.036  | 1.01 | 0.31 |
|                        | GCC       | 0.004  | 1.95 | 0.16 | 0.026 | 1.11 | 0.29 | 0.08   | 1.04 | 0.31 | 0.002  | 3.80 | 0.05 |
|                        | TTT       | 0.003  | 0.20 | 0.65 | 0.001  | 0.20 | 0.65 | 0.001  | 0.20 | 0.65 | 0.001  | 0.20 | 0.65 |
| BCL11A region          | ATGA      | 0.489  | 0.03 | 0.85 | 0.485 | 0.01 | 0.94 | 0.49   | 0.04 | 0.84 | 0.516  | 0.34 | 0.56 | 0.5     | 0.48 | 0.49 |
|                        | CCAG      | 0.299  | 0.39 | 0.53 | 0.288 | 0.02 | 0.89 | 0.29   | 0.01 | 0.94 | 0.308  | 0.35 | 0.56 | 0.328   | 0.04 | 0.84 |
|                        | ACGG      | 0.205  | 0.62 | 0.43 | 0.216 | 0.01 | 0.94 | 0.21   | 0.76 | 0.38 | 0.192  | 0.48 | 0.49 | 0.156   | 0.28 | 0.59 |
|                        | ATGG      | 0.057  | 0.49 | 0.48 | 0.065 | 3.67 | 0.06 | 0.048  | 0.52 | 0.47 | 0.102  | 0.43 | 0.51 | 0.033   | 0.40 | 0.53 |
| HMIP region            | TTTATGT   | 0.747  | 0.75 | 0.39 | 0.752 | 1.15 | 0.28 | 0.74   | 0.19 | 0.66 | 0.806  | 0.08 | 0.78 | 0.797   | 0.00 | 0.98 |
|                        | GTTATGT   | 0.167  | 0.12 | 0.73 | 0.149 | 3.15 | 0.08 | 0.17   | 0.07 | 0.80 | 0.107  | 2.03 | 0.15 | 0.156   | 0.17 | 0.68 |
|                        | TTTATGC   | 0.015  | 0.06 | 0.81 | 0.012 | 0.12 | 0.73 | 0.016  | 0.19 | 0.66 | 0.156  | 0.17 | 0.68 | 0.085   | 2.02 | 0.15 |
|                        | GTCGCCAC  | 0.024  | 0.20 | 0.66 | 0.031 | 0.87 | 0.35 | 0.024  | 0.26 | 0.61 | 0.037  | 1.16 | 0.28 | 0.016   | 0.19 | 0.66 |
|                        | TCCGCAC   | 0.021  | 1.06 | 0.30 | 0.026 | 0.02 | 0.90 | 0.025  | 0.27 | 0.61 | 0.011  | 1.09 | 0.30 | 0.019   | 0.23 | 0.63 |
|                        | TTTGCAC   | 0.044  | 0.11 | 0.41 | 0.045 | 0.21 | 0.65 | 0.047  | 0.61 | 0.44 | 0.085  | 2.02 | 0.15 | 0.017   | 0.20 | 0.65 |
|                        | TTTATGT   | 0.017  | 0.20 | 0.65 | 0.017  | 0.20 | 0.65 |
|                        | TCTGCAC   | 0.016  | 0.19 | 0.66 | 0.016  | 0.19 | 0.66 |
Structure modelling and RMSD prediction

Of the three SNPs (rs2071348, rs7482144, and rs5006884) that were significantly associated with borderline and higher HbA2 levels, only rs5006884 was an exonic polymorphism that made an amino acid change. The three-dimensional (3D) structural modelling for rs5006884 native using a Protein Data Bank template and mutant models using SWISS MODEL and PROCHECK [16, 23–25] were studied to elucidate the effects of the amino acid changes, and were also compared for the energy minimisation and its RMSD using the GROMACS 5.1.1 program [18]. Binding energy minimisation varied significantly, and the total energy values for rs5006884 (7529.6 kJ/mol) variegated proteins were deviated from the wild (7227.9 kJ/mol) protein. In addition, the RMSD values, which are directly proportional to structural deviations with RMSD –0.01 Å, play a significant role during disease initiation. The superimposed position of wild OR51B6 (red colour) and mutant L172F (green colour) model was generated by SWISS PDB Viewer, as shown in Figure 3. This observation suggests that the variation rs5006884 at the OR51B6 gene affects the protein-protein interaction.

Protein-protein interaction

OR51B6 wild and mutant models were checked for possible interaction variations towards the globin proteins (HBB, HBD, HBE, HBA1, HBA2, and HBG) using PRISM server. Based on the template and target protein interface, model complex was generated by prism, with multiple prediction of interaction of molecules bound with free binding energies. Only the lowest free binding energy complexes were selected. The influence of the mutated OR51B6 gene in the binding energy of the OR51B6-target protein complex varied significantly.

Figure 3. Three-dimensional (3D) protein structure of OR51B6. A – Entire wild protein. B – Focused region indicating the L172. C – Entire mutated protein. D – Focused region indicating the L172F. E – Superimposed structure. In the superimposed model: Helix with green colour indicates mutant, helix with red colour indicates wild.
Haemoglobin switching modulator SNPs rs5006884 is associated with increased HbA2 in β-thalassaemia carriers.

Figure 4. OR51B6 interaction with HEMO proteins. Column 1 indicates the template OR51B6 protein and interacting globin protein; column 2 denotes the target globin protein and wild OR51B6 protein. Column 3 denotes the target globin protein and mutant OR51B6 protein.
HbA\textsubscript{2}, which is usually elevated in subjects heterozygous for the \textbeta-thalassaemia mutation. However, there are several factors that may influence this level, such as coinheritance of \textalpha-thalassaemia or \textdelta-thalassaemia, where the HbA\textsubscript{2} concentration is reduced to normal and remains < 3.5\%. Also, the presence of silent HBB mutations projects normal RBC indices and normal or borderline HbA\textsubscript{2} levels [27]. HbA\textsubscript{2} levels are lower in \textbeta\textalpha-thalassaemia mutation carriers, and normal HbA\textsubscript{2} levels are reported in a small subset of traits with non-mutant \textdelta-globin gene [28, 29]. These studies with divergent results entrench the confusion about the reliability of the HbA\textsubscript{2} level [30].

A genome-wide association study identified rs5006884 SNP in the olfactory genes OR51B5/ OR51B6 locus on chromosome 11 to exceed the stringent genome-wide significance threshold of $10^{-6}$, and is considered to be the most significant SNP explaining 5.6\% of the variability in HbF level [31]. The effect of rs5006884 on HbF levels was independent of the Xmn1 site, sickle cell haplotype, \textbeta-globin gene-like complex, and LCR region. A recent study also showed a lack of association between the rs5006884 SNP with HbF in \textbeta-thalassaemia among the Saudi population [32].

An additional regulatory region modulating HbF expression was reported to be centromeric to the \textbeta-globin gene cluster, with genetic modifiers rs2071348 and rs7482144 influencing disease severity [33]. The rs2071348 polymorphism on the HBB pseudogene (HBBP1) was previously significantly associated with a milder disease phenotype in Asian \textbeta(0)-thalassaemia/haemoglobin E patients [34], but causality due to the rs7482144 marker represented by the Xmn1-HBG2 site has not been fully demonstrated [35]. Detailed analysis of the association of HbF QTL – HBG2 promoter region, BCL11A region, and HMIP region genes on Saudis with various levels of HbA2 revealed that these three SNPs, rs2071348, rs7482144, and rs5006884, were significantly associated with borderline and higher HbA\textsubscript{2} levels in the present study, which are reported as insignificant SNPs for HbF level in Saudis [32] confirming the reciprocal relationship among the HbA\textsubscript{2} and HbF levels [10]. The predominant haplotypes in the borderline and higher HbA\textsubscript{2} cohort consisted of GTC, TCC, and TCT in the HBG2 promoter region comprising the same SNPs: rs2071348, rs7482144, and rs5006884, respectively.

The bioinformatics tool-based, systematic, in silico approach to ascertain the impact of genetic variation rs5006884C>T on the structure of OR51B6 protein and its significant impact on the interaction with globin proteins were predicted. The single amino acid change, rs5006884 in the OR51B6, resulted in structural modification with RMSD ~0.01 Å that suggests it plays a significant role during disease initiation. The binding energy minimisation also varied significantly, indicating that the variation rs5006884 affects the protein-protein interaction.

Molchanova et al. [36] reported the \textalpha-2 expression level to be twice the level seen in \textalpha-1, with less efficient translation of the \textalpha-2-mRNA, to maintain a relative level of \textalpha-2 protein. The protein interface energy between HBB/HBA2 and HBD/HBA2 is almost the same, but among haemoglobin proteins with variants of OR51B6, the greatest shift is between OR51B6/HBA2 (78.7\%) and the least between OR51B6/HBA1 (5.07\%) (Figure 5). It was evident from the interaction of OR51B6 and HBA\textsubscript{2} that a functional modification due to the presence of rs5006884C>T on the structure of OR51B6 that has doubled the binding.

**Figure 5.** Influence of the mutated OR51B6 gene in the binding energy of the OR51-target protein complex

*Significantly varied binding energy at 25% compared with the wild type.*
Haemoglobin switching modulator SNPs rs5006884 is associated with increased HbA2 in β-thalassaemia carriers

energy might be the reason for the slight increase in the level of HbA2. Previous studies on the HbA2 protein interaction reported a significant variation in the interface energy between mutant-template and wild-template protein complex [37]. The ORS1B6 can be considered as a HbA2 modulator for large-scale studies.

In conclusion, the haemoglobin switching modulator SNPs in the HBG2 gene region in general and ORS1B6 (rs5006884) variation in particular might intervene in the tetramer formation of α2 with β and δ globin, thereby playing a significant role in the regulation of HbA2 level in β-thalassaemia carriers. Therefore, these variations need to be considered when screening haemoglobinopathies. Furthermore, the ORS1B6 can be studied in detail for the HbA2 modulator to ameliorate disease severity.

Acknowledgments

The authors extend their gratitude to King Abdulaziz City for Science and Technology, Saudi Arabia (To Dr. Cyril Cyrus, Grant No: LGP 31-34) for the funding of this research. We also thank Mr. Geoffrey James Tam Moro and Mr. Mohammed Al Shamlan for their technical and administrative support.

Conflict of interest

The authors declare no conflict of interest.

References

1. Palstra RJ, de Laat W, Grosse F. Beta-globin regulation and long-range interactions. Adv Genet 2008; 61: 107-42.
2. Menzel S, Garner C, Rooks H, Spector TD, Thein SL. HbA2 levels in normal adults are influenced by two distinct genetic mechanisms. Br J Haematol 2013; 160: 101-5.
3. Steinberg MH, Rodgers GP. HbA2: biology, clinical relevance and a possible target for ameliorating sickle cell disease. Br J Haematol 2015; 170: 781-7.
4. Siatecka M, Bleker JJ. The multifunctional role of KLF1/KLF1 during erythropoiesis. Blood 2011; 118: 4454-8.
5. Gallienne AE, Dreau HM, Schuh A, Old JM, Henderson S. Ten novel mutations in the erythroid transcription factor KLF1 gene associated with increased fetal hemoglobin levels in adults. Haematologica 2012; 97: 340-3.
6. Tepakhan W, Yamsri S, Fucharoen G, Sanchaisuriya K, Fucharoen S, Aerbajinai W, Trainor C, Gao P, Rodgers GP. Recombinant erythroid Kruppel-like factor fused to GATA1 up-regulates delta- and gamma-globin expression in erythroid cells. Blood 2011; 117: 3045-52.
7. Farrell JJ, Sherva RM, Chen ZY, et al. A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. Blood 2011; 117: 4935-45.
8. Yu LH, Liu D, Cai R, et al. Changes in hematological parameters in alphathalassaemia patients. Blood Cells Mol Dis 2015; 55: 27-9.
9. Borgio JF, AbdulAzeez S, Al-Muslami AM, et al. KLF1 gene and borderline hemoglobin A2 in Saudi population. Arch Med Sci 2018; 14: 230-6.
10. Griffin PJ, Sebastiani R, Edward H, et al. The genetics of hemoglobin A2 regulation in sickle cell anemia. Am J Hematol 2014; 89: 1019-23.
11. Zhu J, Lin C, Kato R, Trainor C, Gao P, Rodgers GP. Recombinant erythroid Kruppel-like factor fused to GATA1 up-regulates delta- and gamma-globin expression in erythroid cells. Blood 2011; 117: 3045-52.
12. Farrell JJ, Sherva RM, Chen ZY, et al. A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. Blood 2011; 117: 4935-45.
13. Akinshaye I, Alsultan A, Solovieff N, et al. Fetal hemoglobin in sickle cell anemia. Blood 2011; 118: 19-27.
14. Al-Nafie AN, Borgio JF, AbdulAzeez S, et al. Co-inheritance of novel ATRX gene mutation and globin (alpha and beta) gene mutations in transfusion dependent beta-thalassemia patients. Blood Cells Mol Dis 2015; 55: 27-9.
15. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 2006; 22: 195-201.
16. Laskowski RA, Macarthur MW, Moss D, Thornton JM. PROCHECK – a tool to check the stereochemical quality of protein structures. J Appl Crystallogr 1993; 26: 283-91.
17. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. Electrophoresis 1997; 18: 2714-23.
18. Lindahl E, Aazaar C, Koehl P, Delarue M. NOMAD-Ref: visualization, deformation and refinement of macro-molecular structures based on all-atom normal mode analysis. Nucleic Acids Res 2006; 34: 52-56.
19. Ogmen U, Keskın O, Aytaşı A, Nussinov R, Gursoy A. PRISM: protein interactions by structural matching. Nucleic Acids Res 2005; 33: W331-6.
20. Keskın O, Nussinov R, Gursoy A. PRISM: protein-protein interaction prediction by structural matching. Methods Mol Biol 2008; 484: 505-21.
21. Tuncbag N, Gursoy A, Nussinov R, Keskın O. Predicting protein-protein interactions on a proteome scale by matching evolutionary and structural similarities at interfaces using PRISM. Nat Protoc 2011; 6: 1341-54.
22. Barrett JC, Fry B, Maller J, Daly MJ. Haplovizion analysis and visualization of LD and haplotype maps. Bioinformatics 2005; 21: 263-5.
23. Kiefer F, Arnold K, Künzli M, et al. The SWISS-MODEL repository and associated resources. Nucl Acids Res 2009; 37: D387-92.
24. Berman HM, Westbrook J, Feng Z. The protein data bank. Nucleic Acids Res 2000; 28: 235-42.
25. Feng L, Delarue M, Berman HM, Westbrook J. The protein data bank. Nucleic Acids Res 2000; 28: 235-42.
26. Chen DE, Willick DL, Ruckel JB, Floriano WB. Principal component analysis of binding energies for single-point mutations in AHSP-mediated stabilization of alpha-hemoglobin. Cell 2004; 119: 629-40.
27. Chen DE, Willick DL, Ruckel JB, Floriano WB. Principal component analysis of binding energies for single-point mutations of hT2R16 bound to an agonist correlate with experimental mutant cell response. J Comp Biol 2015; 22: 37-53.
28. Origa R, Editors In: M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, H.H. Bean, C.M. Smith, K. Stephens, A. Amemiya, N. Ledbetter, editors. SourceGenetReviews®, Seattle (WA): University of Washington, Seattle; 1993-2017.2000 Sep 28 [updated 2015 May 14].
29. Galanello R, Barella S, Idea E, et al. Genotype of subjects with borderline hemoglobin A2 levels: implication for beta-thalassemia carrier screening. Am J Hematol 1994; 46: 79-81.
29. Fucharoen S, Fucharoen G, Sanchaisuriya K, Pengjam Y. Molecular analysis of a Thai beta-thalassaemia heterozygote with normal haemoglobin A2 level: implication for population screening. Ann Clin Biochem 2002; 39: 44-9.

30. Verhovsek M, So CC, O’Shea T, et al. Is HbA2 level a reliable diagnostic measurement for beta-thalassemia trait in people with iron deficiency? Am J Hematol 2012; 87: 114-6.

31. Solovieff N, Milton JN, Hartley SW, et al. Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the S olfactory receptor gene cluster. Blood 2010; 115: 1815-22.

32. Cyrus C, Vatte C, Borgio JF, et al. Existence of HbF enhancer haplotypes at HBS1L-MYB intergenic region in transfusion-dependent Saudi beta-thalassemia patients. Biomed Res Int 2017; 2017: 1972429.

33. Sherva R, Sripichai O, Abel K, et al. Genetic modifiers of Hb E/beta0 thalassemia identified by a two-stage genome-wide association study. BMC Med Genet 2010; 11: 51.

34. Giannopoulou E, Bartsakoulia M, Tafrali C, et al. A single nucleotide polymorphism in the HBBP1 gene in the human beta-globin locus is associated with a mild beta-thalassemia disease phenotype. Hemoglobin 2012; 36: 433-45.

35. Danjou F, Francavilla M, Anni F, et al. A genetic score for the prediction of beta-thalassemia severity. Haematologica 2015; 100: 452-7.

36. Molchanova TP, Pobedimskaya DD, Huisman THJ. The differences in quantities of alpha2- and alpha1-globin gene variants in heterozygotes. Br J Haematol 1994; 88: 300-6.

37. Borgio JF, Al-Madan MS, AbdulAzeez S, et al. Mutation near the binding interfaces at alpha-hemoglobin stabilizing protein is highly pathogenic. Am J Transl Res 2016; 8: 4224-32.