European Journal of Haematology

European Journal of Haematology

ORIGINAL ARTICLE

Hb Wilde and Hb Patagonia: two novel elongated beta-globin variants causing dominant beta-thalassemia

Karen G. Scheps[1], Marcia A. Hasenahuer[2], Gustavo Parisi[2], María S. Fornasari[2], Sandra P. Pennesi[3], Beatriz Erramouspe[4], Felisa N. Basack[3], Ernesto S. Veber[5], Luis Aversa[3], Graciela Elena[5], Viviana Varela[1]

[1]Cátedra de Genética y Biología Molecular, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires; [2]Structural Bioinformatics Group, Unidad de Físico Química, Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Buenos Aires; [3]Unidad de Hematología, Hospital de Niños 'Dr. Ricardo Gutiérrez', Ciudad Autónoma de Buenos Aires; [4]Servicio de Hematología, Unidad Asistencial 'Dr. César Milstein', Ciudad Autónoma de Buenos Aires; [5]Servicio de Hemato-Oncología Pediátrica, Hospital General de Niños 'Dr. Pedro de Elizalde', Ciudad Autónoma de Buenos Aires, Argentina

Abstract

We describe here the molecular and hematological characteristics of novel frameshift mutations in exon 2 of the *HBB* gene (in heterozygous state) found in two Argentinean pediatric patients with dominant β-thalassemia-like features. In Hb Wilde, *HBB*:c.270_273delTGAG(*p.Glu90Cysfs*67*), we detected the deletion of the third base of the codon 89 (T) and the codon 90 (GAG), whereas in Hb Patagonia, *HBB*: c.296_297dupGT(*p.Asp99Trpfs*59*), the frameshift mutation was due to a duplication of a 'GT' dinucleotide after the second base of codon 98 (GTG). The Hb Patagonia and Hb Wilde mutations would result in elongated β-globin chains with modified C-terminal sequences and a total of 155 and 157 amino acids residues, respectively. Based on bioinformatics and structural analysis, as well as protein modeling, we predict that the elongated β-globins would affect the formation of the αβ dimers and their stability, which would further support the mechanism for the observed clinical features in both patients.

Key words dominant beta-thalassemia; *HBB* gene; elongated beta-globin variant; hemolytic anemia

Correspondence Viviana Varela, Cátedra de Genética y Biología Molecular, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junin 956, C1113AAC, Ciudad de Buenos Aires, Argentina. Tel: 54 11 49648296; Fax: 54 11 49648296; e-mail: vvarela@ffyb.uba.ar

Accepted for publication 26 September 2014 doi:10.1111/ejh.12456

The β-thalassemias (β-thal) (OMIM: 613985) are a group of disorders typically inherited in a Mendelian recessive manner and characterized by reduced synthesis of β-globin chains. It is one of the most common inherited conditions worldwide. According to the clinical features, three forms of presentation can be defined: β-thal minor), β-thal major, and β-thal intermediate.

More than 270 mutations that cause β-thal have been reported in the *HBB* gene (ID: 3043) (1). At least 40 of these have been associated with dominant forms of β-thal (OMIM: 603902). This group includes missense mutations that involve amino acids important in the maintenance of the β-globin chain stability, nonsense mutations that escape nonsense-mediated decay, and minor insertions or deletions that lead to the loss of intact codons encoding key amino acids or frameshifts that frequently arise as *de novo* point mutations (2).

The affected individuals, with a single mutated *HBB* gene, present β-thal intermedia, or even more severe phenotypes, consequence of the substantial amounts of cytosolic β-globin mRNA that produces Hb variants that precipitate in the erythroid precursors before assembling with the α-globin chains to produce the Hb tetramer.

The resulting ineffective erythropoiesis was also exacerbated by the concomitant relative excess of the α-chains (3). Characteristic findings include splenomegaly, intermittent jaundice, and extramedullary hematopoiesis.

We report here two novel exon 2 *HBB* gene mutations that were detected in two pediatric patients with severe thalassemia-like features and hematologically normal parents.

These frameshift mutations would result in elongated β-globin chains with modified C-terminal sequences. We propose a putative pathogenic mechanism for the observed phenotype based on β-globin sequence and structural analysis as well as protein–protein interface alteration.

Patients and methods

Proband 1 was the only daughter of an Argentinean non-consanguineous couple, referred from the neurosurgery service, which was consulted for the presence of frontal bossing. At the time of the admission to the Hematology Service of the 'Dr. Pedro de Elizalde' Hospital, the 4-yr-old patient had no palpable spleen. A bone marrow puncture was performed and no abnormal erythroblasts were found, dismissing the presence of a dyserythropoietic anemia. With a presumptive diagnosis of severe thal, a regular transfusion program with packed red blood cells was initiated. Once the presence of the most common β-thal mutations was discarded, an unstable hemoglobin variant was suspected. The hematological studies were performed before the first transfusion.

Proband 2 also has non-consanguineous Argentinean parents and presented hemolytic anemia at the age of 4. She required multiple red blood cell transfusions for the severe anemia and underwent splenectomy when she was 12 yr old. Upon admission to the 'Dr. Ricardo Gutierrez' Hospital in Buenos Aires, she exhibited moderate chronic hemolytic anemia, pallor, and jaundice. The hematological studies reported in this work were done after the splenectomy.

The parents of both probands were normal upon physical examination, with normal hematological features.

Written informed consents following the current version of the Helsinki Declaration were obtained from the individuals involved in this study, and the research project was approved by the institutional bioethical committee.

Peripheral blood cell counts and erythrocyte indices were determined using an electronic cell counter (Sysmex XT2000i; Sysmex Corporation, Kobe, Japan). Hemoglobin electrophoresis was carried out with a semiautomatic agarose gel system at both alkaline and acid pH (Sebia, Lisses, Évry, France). Hb A_2 was assessed with microcolumn chromatography and Hb F using an alkali denaturation method. For Proband 1, capillary electrophoresis was performed, as well. Heinz bodies were evaluated by means of incubation with brilliant cresyl blue (4). Tests for stability were carried out by the isopropanol precipitation test (5).

Genomic DNA from the probands and their asymptomatic parents was isolated from peripheral blood leukocytes using standard methods (6). The complete *HBB* gene was amplified from the position c.−209 to c.*288 by polymerase chain reaction (PCR) using kit T-Plus ADN Polymerase (INBIO, Tandil, Argentina) and the primers HBB1-F and HBB2-R (7) and L3: 5′-GGGTACAGTTTAGAATGGG-3′ and HBB4-R: 5′-TAGCTGTTTGCAGCCTCACCTT-3′ designed with PrimerQuest® (IDT, Coralville, IA, USA). The products were purified and sequenced in both directions, on the ABI PRISM™ 3130XL genetic analyzer (Applied Biosystems, Seoul, Korea). The sequences were compared to the reference sequence NG_000007.3 and analyzed using Basic Local Alignment Search Tool (BLAST) (8), considering the reports in the HbVar database (1).

The patients' *HBB* PCR products from c. −209 to c.315+144 were cloned in the pGEM®-T Easy Vector Systems (Promega, Madison, WI, USA) as described in (9) and sequenced.

The most common α-thal$^+$ deletion, -$\alpha^{3.7}$, and the $\alpha\alpha\alpha^{anti3.7}$ insertion were analyzed by gap-PCR (10, 11).

Three SNPs in the *loci* that showed the strongest association with Hb F levels were studied. The polymorphism *rs7482144* in the promoter region of *HBG2* (minor allele T) was determined by specific PCR amplification (12) and digestion with the restriction endonuclease XmnI (Promega, Madison, WI, USA). The polymorphisms *rs1188686* in the second intron of *BCL11A* (minor allele G) and *rs4895441* in the intergenic region of *HBS1L-MYB* (minor allele G) were analyzed by Real-time PCR and detection with TaqMan® probes (Applied Biosystems, Foster City, CA, USA).

The altered amino acid sequences were studied analyzing their physicochemical properties, secondary and tertiary structure prediction using different methods: PSIPRED (13), I-Tasser (14), Quark (15), Robetta (16), and ProsaII (17). The contribution to stabilize the interface of the positions involved in non-covalent interactions between subunits was evaluated by computational alanine scanning mutagenesis using Robetta. To evaluate stabilizing interactions, the iron ion neighborhood was examined with an 8Å sphere radius for the oxy and deoxy forms. For all the analysis involving structural information, the structures with PDB codes 1HHO and 2HHD were used as input for the oxy and deoxy Hb A, respectively.

Results

The hematological parameters of the probands and their parents are summarized in Table 1. No abnormal hemoglobin fraction was observed in any of the samples. The Hb A_2 fraction was also within the normal range. However, both patients showed marked increments of Hb F, which would indicate that they had been subjected to hematopoietic stress.

Three different loci associated with the modulation of Hb F levels after birth were analyzed in the probands, to evaluate a possible impact in this Hb fraction. Proband 1 presented the genotype C/T for the SNP *rs7482144*, A/G for *rs4895441*, and A/G for *rs1188686*. The genotypes for Proband 2 were C/C, A/G, and G/G for the respective *loci*. The G/G genotype for *rs1188686* could influence, at least slightly, the expression of Hb F. It is likely, then, that the

Table 1 Hematological parameters and results of the laboratory tests of the families studied. Proband 1 was 5 yr old when the studies were performed and Proband 2 was 13 yr old

	Proband 1	Father 1	Mother 1	Proband 2	Father 2	Mother 2
Hb (g/dL)	7.8	14.2	12.8	8.8	14.4	14.0
RBC (10^{12}/L)	3.00	4.61	4.30	3.41	4.97	4.66
PCV (L/L)	0.23	0.42	0.38	0.27	0.43	0.43
MCV (fL)	76.6	91.0	88.3	79.8	87.1	92.9
MCH (pg)	26.0	30.0	29.7	25.8	29.1	30.0
MCHC (g/dL)	33.9	33.0	33.6	32.4	33.5	32.3
Reticulocyte (%)	4.7	0.7	1.3	3.6	0.8	0.9
Heinz Bodies	Neg	Neg	Neg	Pos	Neg	Neg
Fe (μg/dL)	–	180	135	178	–	–
Ferritin (μg/L)	78	–	36	304	–	–
Hb A_2 (%)	2.2	2.7	2.9	2.8	2.4	2.5
Hb F (%)	35.7	0.2	0.5	11.0	0.3	0.9
Hb X (%)	–	–	–	–	–	–
Isopropanol test	Neg	Neg	Neg	Pos	Neg	Neg

difference in the values observed between both patients was mostly due to the splenectomy.

Proband 2 presented a positive result for the isopropanol test and Heinz bodies, which strongly suggested the presence of an unstable Hb variant. Although no Heinz bodies were present in the peripheral blood samples of Proband 1 and no positive result was obtained for the instability test, the absence of an aberrant normal fraction in the Hb electrophoresis, together with a thalassemia clinical picture, hinted a possible unstable Hb variant responsible for the phenotype observed.

The presence of mutated Hb variants was confirmed by DNA sequencing. The probands presented frameshift mutations in the *HBB* gene that gave rise to novel elongated variants. The sequence of the *HBB* gene of Proband 1 revealed the presence of the change *HBB*:c.270_273delTGAG (*p.Glu90Cysfs*67*) in heterozygote state, which was confirmed by cloning. The deletion of four bp in exon 2 includes the third base of the codon 89 (T) that encodes a serine and the codon 90 (GAG) that encodes a glutamic acid; the frameshift would result in an elongated β-globin chain with a modified C-terminal sequence and a total of 155 amino acids. Considering the birthplace of the patient, Hb Wilde was the name given to this new Hb variant.

Proband 2 presented the change *HBB*:c.296_297dupGT (*p.Asp99Trpfs*59*) in heterozygote state, which was also confirmed with the same cloning strategy described above. The duplication of a 'GT' dinucleotide in exon 2, after the second base of codon 98 (GTG) that encodes a valine would have the same frameshift effect than the four-bp deletion present in Hb Wilde, leading to a β-globin chain with a modified C-terminal sequence and 157 amino acids. As the patient was born in the south of Argentina, the name Hb Patagonia was given to this new variant.

The patients did not present any other alterations in the *HBA* genes that would contribute to the observed phenotype. In both cases, the mutations analyzed were absent in their parents DNA. The DNA sequences of both patients and the translations of the abnormal β-globin chains are shown in Fig. 1. Tetrameric organization of hemoglobin (PDB id: 1HHO, oxyhemoglobin) with affected regions of β-globin chains in Hb Wilde and Hb Patagonia is displayed in Fig. 2.

To predict putative conformational alterations of the β-globin chains due to the frameshift errors, the physicochemical characteristics of the mutated variants were compared to the wild-type sequence. Composition, isoelectric point and charge at different pHs were analyzed for the altered fragments (66 C-terminal amino acids for Hb Wilde and 59 C-terminal for Hb Patagonia) and the last 57 C-terminal residues of the wild-type β-globin chain. Both variants lose all the negative charged residues in this region (two aspartic acids and three glutamic acids) and simultaneously gain one positive charged residue. There is also an increment in nonpolar residues and prolines. The predicted isoelectric point shows a noticeable displacement toward higher values: 7.25 for wild type, 9.56 for Hb Wilde and 10.08 for Hb Patagonia. There is also a significant change in the predicted charge at pH 7 from 1.3 in the wild-type sequence to 6.00 and 6.10 for Hb Wilde and Hb Patagonia, respectively.

3D modeling results from *ab initio* methods predict structured regions for both altered forms (data not shown). However, the energetic evaluation of all the 3D models obtained in this study revealed that most of the proposed structural models are unstable.

The β-chain residues at α1–β1 and β1–α2 interfaces are listed in Table 2. From a total of 18 residues of the complete α1–β1 interface in the oxy form, 13 (72%) are located in the region altered in the Hb Wilde and Hb Patagonia. In the deoxy form, 12 of 16 residues (75%) are altered in these mutants. Regarding the β1–α2 interface, four of eight residues (50%) would be affected in the oxy form and seven of 12 (58%) in the deoxy form. The computational alanine

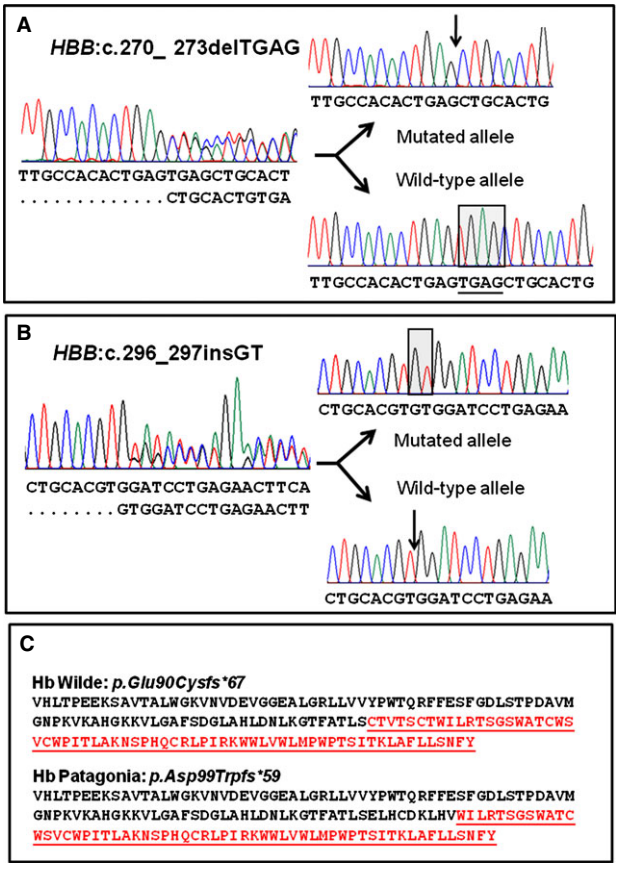

Figure 1 Identification of mutations by DNA sequencing of the PCR product of *HBB* gene (left) and direct sequencing of the clones (mutated and wild type) obtained by with the pGEM®-T Easy Vector system (right). (A) The electropherogram showed the presence of the mutation *HBB*:c.270_273delTGAG in a heterozygous state. (B) The electropherogram showed the presence of the mutation *HBB*: c.296_297dupGT in a heterozygous state. (C) Comparison of the translation of the resulting peptides. The amino acids that differ from the reference sequence are underlined.

scanning mutagenesis results show that in wild-type structures (deoxy and oxy), the residues R30, V34, H116, Q127, and Q131, are important for stabilization of α1–β1 interaction, reflected by the high values (>±1 kcal/mol) of ΔΔG of binding. Of these five residues, the last three are changed in the analyzed variants. In the case of β1–α2 interface, four residues, W37, R40, D99, N102, are important. The last two are altered in Hb Wilde and Hb Patagonia. Heme iron ion environment is also altered in Hb Wilde for positions L91, H92, L96, V98, F103, L106, L141, Y145 (eight of 18 positions). Only, the last four are also altered in Hb Patagonia.

Discussion

Although most of β-thal mutations are inherited in a recessive form, a growing number of studies identified novel mutations in the *HBB* gene causing dominant forms.

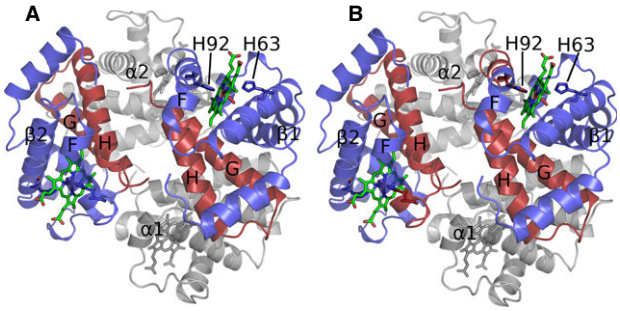

Figure 2 Human Wt oxyhemoglobin tetramer (PDBid 1HHO). α-chains are colored in gray. β-chains are in blue for the wild-type fragments and in red for the mutant fragments that correspond to (A) Hb Wilde and (B) Hb Patagonia. H92 (proximal His) and H63 (distal His) are pointed with arrows in B (β1) chains. Oxy-heme groups are represented in sticks and colored in green for β chains. Structures were represented and colored with PyMOL (http://www.pymol.org/).

Table 2 List of β1 residues in α1–β1 and β1–α2 interfaces of oxyhemoglobin (PDB id 1HHO) and deoxyhemoglobin (PDB id 2HHD). Residues altered in Hb Wilde and Hb Patagonia are in bold; H97(*) only in the case of Hb Wilde

β-residues in α1–β1 interface		β-residues in β1–α2 interface	
Oxy-Hb	Deoxy-Hb	Oxy-Hb	Deoxy-Hb
R30	R30	P36	V34
V33	V33	W37	Y35
V34	V34	Q39	P36
Y35	Y35	R40	W37
M55	M55	**H97(*)**	**R40**
N108	**N108**	**D99**	**H97(*)**
C112	**C112**	**N102**	**D99**
A115	**A115**	**Y145**	**P100**
H116	**H116**		**E101**
G119	**G119**		**L105**
K120	**F122**		**Y145**
F122	**T123**		**H146**
T123	**P124**		
P124	**P125**		
P125	**Q127**		
Q127	**A128**		
A128	**Q131**		
Q131			

However, these forms are still infrequent and many mutations have only been described once. Three classes of mutations can lead to this phenotype: 1-missense mutations, which can disrupt the β-globin hydrophobic heme pocket or its secondary structure, 2-nonsense mutations that escape nonsense-mediated decay and translate into truncated β-globin chains, and 3-minor deletions or insertions, which can induce the loss or gain of intact codons, in particular of the amino acids involved in the αβ dimer formation, or can cause frameshifts resulting in elongated β-globins (18). There are over 20 reports of elongated β-globin variants

[ithanet: http://www.ithanet.eu/], associated with severe phenotypes, that include an early onset of the disease, with about one-third of the cases identified in children. These mutations reside mostly in exon 3 and in the 3′ region of exon 2. We describe two novel mutations, Hb Wilde and Hb Patagonia, both located in exon 2 that give rise to elongated variants.

The Hb Wilde mutation is a deletion of four bp that are duplicated in the wild-type sequence and the Hb Patagonia, a duplication of a 'GT' dinucleotide. It is possible that a DNA polymerase slippage during replication could be the mechanism involved in the mutagenesis in both cases. Interestingly, the 'GT' duplication, unlike other insertions that lead to elongated variants, such as Hb Agnana (19), does not involve a region that is constitutively repeated in the wild-type sequence.

Several compositional and structural analyses were performed to associate the physiological symptoms occurring in both patients with physicochemical and structural changes in Hb Wilde and Hb Patagonia. We predict that the *HBB* mutations detected in both patients lead to very important amino acid charge changes that could drastically affect the tertiary structure of the β-globin chain as well as the quaternary structure. In this sense, the complete lack of all the aspartic and glutamic acids in the region altered in both variants would seriously alter their capacity to establish protein–protein interactions with the alpha chain that are normally stabilized by electrostatic interactions. These alterations would also shift the isoelectric point of the altered C-terminal segments, from nearly neutral in the wild-type Hb A to basic values, changes that could also affect the stabilizing tertiary interactions. Although the proximal H92, essential for the iron ion stabilization, is present in Hb Patagonia (this residue is lost in Hb Wilde), the electrostatic and hydrophobic alterations in both mutated sequences along with the alteration of several residues contacting iron could seriously endanger the tertiary stability of the proteins. The structural characterization reinforces the idea that the tertiary structure as well as the majority of the positions involved in subunit interactions is altered in both variants. Summing up, our results highly support that the heterodimeric αβ form is lost in both variants and possibly the tertiary structure could be also drastically altered in Hb Wilde and in Hb Patagonia.

These results would further support the lack of detection of unstable hemoglobin in peripheral blood and the notion that the extensive hemolysis results from the precipitation of the altered β-globin chains and the concomitant free α-chains that promote the formation of inclusion bodies that damage the plasma membrane of the erythroid precursors in the marrow and result in ineffective erythropoiesis. Although there is still light to be shed on how the elongated variants interact with the wild-type β-globin chains, there is enough evidence supporting the lack of αβ dimers formation as the mechanism responsible for the precipitates. The fact that there is a stable elongated *HBB* variant, Hb Tak (HbVar ID 710), product of a frameshift mutation between codons 146 and 147, in which the insertion of a CA dinucleotide (20) occurs downstream of the predicted positions involved in the subunits interaction strongly supports the models presented here.

It is still not entirely clear why the different elongated forms exhibit distinctive clinical severity, as is the case of both probands who presented elongated variants of 155 and 156 amino acids, product of a+2 frameshift effect, and exhibited phenotypes of different severity. Thus, whereas the patient carrier of Hb Wilde presented extramedullary hematopoiesis at the age of five and requires regular blood transfusions, the patient with Hb Patagonia presented a milder phenotype, possibly due to the splenectomy. However, a more pronounced severity due to the seven amino acids that differ in both variants cannot be completely ruled out.

It is remarkable that the parents of the patients did not exhibit thalassemic phenotypes and none of them carried the mutations found in the probands. A common feature of the dominant β-thal mutations is that they tend to occur as *de novo* mutations. This characteristic showcases the importance of the molecular studies to establish a definite diagnosis in patients with severe thalassemia phenotypes and hematologically normal parents.

Acknowledgements

The authors would like to thank Dr. Raquel Osatinsky (Laboratorio MANLAB) for carrying out the capillary electrophoresis assay and Dr. Osvaldo Rey (Instituto de Inmunología, Genética y Metabolismo, CONICET) for his critical comments on the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Funding

This study was supported by Research Grants: 20020100 200073 and 20020120200092BA from the Universidad de Buenos Aires (UBA), 112-200801-02849, Multi-year Research Projects from the Consejo Nacional de Investigaciones Científicas y Técnicas (PIP CONICET), and 1004/11 from the Universidad Nacional de Quilmes (UNQ).

References

1. Patrinos GP, Giardine B, Riemer C, Miller W, Chui DHK, Anagnou NP, Wajcman H, Hardison RC. Improvements in the HbVar database of human haemoglobin variants and thalassaemia mutations for population and sequence variation studies. *Nucleic Acids Res* 2004;**32**:D537–41. (http://globin.cse.psu.edu/hbvar/menu.html).

2. Luo HY, Tang W, Eung SH, Coad JE, Canfield P, Keller F, Crowell EH Jr, Steinberg MH, Chui DH. Dominantly inherited beta thalassemia intermedia caused by a new single nucleotide deletion in exon 2 of the beta globin gene: Hb Morgantown (beta91 CTG>CG). *J Clin Pathol* 2005;**58**:1110–2.
3. Cao A, Galanello R. Beta-thalassemia. *Genet Med* 2010;**12**:61–76.
4. Eisinger J, Flores J, Tyson JA, Shohet SB. Fluorescent cytoplasm and Heinz bodies of hemoglobin Köln erythrocytes: evidence for intracellular heme catabolism. *Blood* 1985;**65**:886–93.
5. Carrell RW, Kay R. A simple method for the detection of unstable heamoglobins. *Br J Haematol* 1972;**23**:615–9.
6. Murray MG, Thompson WF. Rapid isolation of high molecular-weight plant DNA. *Nucleic Acids Res* 1980;**8**:4321–5.
7. Rossetti LC, Targovnik HM, Varela V. The molecular basis of beta-thalassemia in Argentina. Influence of the pattern of immigration from the Mediterranean Basin. *Haematologica* 2004;**89**:746–7.
8. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 2010;**215**:403–10.
9. Scheps KG, Binaghi A, Varela V. Identification of a new HBA1 gene mutation (HBA1:c.301-2A>T) in cis with the Hb Riccarton (HBA1:c.154G>A) [alpha1 51(CE9) Gly>Ser]. *Hemoglobin* 2012;**36**:504–7.
10. Tan AS, Quah TC, Low PS, Chong SS. A rapid and reliable 7-deletion multiplex polymerase chain reaction assay for α-thalassemia. *Blood* 2001;**98**:250–1.
11. Wang W, Ma ES, Chan AY, Prior J, Erber WN, Chan LC, Chui DH, Chong SS. Single-tube multiplex-PCR screen for anti$^{-3.7}$ and anti$^{-4.2}$ α-globin gene triplications. *Clin Chem* 2003;**49**:1679–82.
12. Hooshang N, Zohreh R, Gholamreza B. The Xmn1 polymorphic site 5′ to the Gγ gene and its correlation to the Gγ:Aγ ratio, age at first blood transfusion and clinical features in β-Thalassemia patients from Western Iran. *Mol Biol Rep* 2010;**37**:159–64.
13. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 1999;**292**:195–202.
14. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 2008;**9**:40.
15. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* 2004;**32**:W526–31.
16. Xu D, Zhang Y. *Ab initio* protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* 2012;**80**:1715–35.
17. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 2007;**35**:W407–10.
18. Thein SL. Dominant beta thalassaemia: molecular basis and pathophysiology. *Br J Haematol* 1992;**80**:273–7.
19. Ristaldi MS, Pirastu M, Murru S, *et al.* A spontaneous mutation produced a novel elongated beta-globin chain structural variant (Hb Agnana) with a thalassaemia-like phenotype. *Blood* 1990;**75**:1378–9.
20. Flatz G, Kinderlerer JL, Kilmartin JV, Lehmann H. Haemoglobin Tak: a variant with additional residues at the end of the beta-chains. *Lancet* 1971;**10**:732–3.

