



## Article

# mRNA Analysis of Frameshift Mutations with Stop Codon in the Last Exon: The Case of Hemoglobins Campania [ $\alpha 1$ cod95 (–C)] and Sciacca [ $\alpha 1$ cod109 (–C)]

Giovanna Cardiero, Gennaro Musollino, Romeo Prezioso and Giuseppina Lacerra \*

Institute of Genetics and Biophysics “Adriano Buzzati Traverso”, National Research Council, 80131 Naples, Italy; giovanna.cardiero@gmail.com (G.C.); gennaromusollino@gmail.com (G.M.); romeo.prezioso@igb.cnr.it (R.P.)

\* Correspondence: giuseppina.lacerra@igb.cnr.it

**Abstract:** An insertion or deletion of a nucleotide (nt) in the penultimate or the last exon can result in a frameshift and premature termination codon (PTC), giving rise to an unstable protein variant, showing a dominant phenotype. We described two  $\alpha$ -globin mutants created by the deletion of a nucleotide in the penultimate or the last exon of the  $\alpha 1$ -globin gene: the *Hb Campania* or  $\alpha 1$  cod95 (–C), causing a frameshift resulting in a PTC at codon 102, and the *Hb Sciacca* or  $\alpha 1$  cod109 (–C), causing a frameshift and formation of a PTC at codon 133. The carriers showed  $\alpha$ -thalassemia alterations (mild microcytosis with normal Hb A2) and lacked hemoglobin variants. The 3D model indicated the  $\alpha$ -chain variants’ instability, due to the severe structural alterations with impairment of the chaperone alpha-hemoglobin stabilizing protein (AHSP) interaction. The qualitative and semiquantitative analyses of the  $\alpha 1$ mRNA from the reticulocytes of carriers highlighted a reduction in the variant cDNAs that constituted 34% (*Hb Campania*) and 15% (*Hb Sciacca*) of the total  $\alpha 1$ -globin cDNA, respectively. We developed a workflow for the in silico analysis of mechanisms triggering no-go decay, and its results suggested that the reduction in the variant mRNA was likely due to no-go decay caused by the presence of a rare triplet, and, in the case of *Hb Sciacca*, also by the mRNA’s secondary structure variation. It would be interesting to correlate the phenotype with the quantity of other frameshift mRNA variants, but very few data concerning  $\alpha$ - and  $\beta$ -globin variants are available.

**Keywords:**  $\alpha$ -thalassemia; frameshift; premature termination codon (PTC); unstable  $\alpha$ -Hb variants; mRNA quality control; no-go decay; dominant phenotype; *Hb Campania* or *HBA1:c.287delC*; *Hb Sciacca* or *HBA1:c.328delC*; bioinformatics



**Citation:** Cardiero, G.; Musollino, G.; Prezioso, R.; Lacerra, G. mRNA Analysis of Frameshift Mutations with Stop Codon in the Last Exon: The Case of Hemoglobins Campania [ $\alpha 1$  cod95 (–C)] and Sciacca [ $\alpha 1$  cod109 (–C)]. *Biomedicines* **2021**, *9*, 1390. <https://doi.org/10.3390/biomedicines9101390>

Academic Editor: Luísa Romão

Received: 2 August 2021

Accepted: 29 September 2021

Published: 4 October 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The  $\alpha$ -thalassemia phenotype is mainly associated with common deletional defects, the formation of which is favored by the high homology between the *HBA1* and *HBA2* globin genes contained in duplication units [1]. We identified about 25% of the  $\alpha$ -thalassemia point mutation variants via epidemiological studies in Southern Italy [2]. These point mutations can cause thalassemia by inactivating regulatory sequences, or they can give rise to unstable variants [1,3].

One class of mutants is characterized by the deletion/insertion of a few base pairs (bp), causing frameshift, leading to truncated or elongated protein chains. The frameshift can give rise to globin chains with an altered helix and length, due to reading frame shift, creating a premature termination codon (PTC) or placing a stop codon after the canonical one.

These hemoglobins (Hb) are mostly unstable variants, carrying structural modifications in the  $\alpha 1\beta 1$  contact area. The instability is caused mainly by their altered interaction with the chaperone alpha-hemoglobin stabilizing protein (AHSP), which mimics the interaction with the  $\beta$ -chain, but could also be caused by an altered H helix, which is crucial in the building of the heme pocket, or by altered Hb tetramer assembly [3–5].

The heterozygote phenotype is very variable. The carriers can be asymptomatic or show only borderline hematological abnormalities. Nevertheless, when the proteolytic mechanisms of erythrocytes are unable to remove all the denatured Hb material, some of these variants may cause mild chronic hemolytic anemia [3]. The frameshift variants are often identified under homozygous or compound heterozygous conditions, with patients showing a more severe phenotype with respect to their genotype, such as Hb H and thalassemia intermedia syndrome [3].

Typically, transcripts that contain PTC undergo nonsense-mediated decay (NMD) to avoid the production of toxic proteins, but when it is located in the last exon, the PTC will escape the surveillance system, and a truncated protein will be translated [6]. The analysis of the stop codons present in the  $\alpha$ -globin genes by means of the Virtual Ribosome software, (<https://services.healthtech.dtu.dk/service.php?VirtualRibosome-2.0>, accessed on 22 July 2021), has shown that in  $\alpha$ -globin genes, nine non-in-frame stop codons are present in the coding region (at codons 16, 49, 56, 67, 84, 102, 108, 133, and 137) and three in the 3'UTR (at positions 173, 174 and 176) [7]. The deletions/insertions occurring after codon 84 cause a frameshift that creates a stop codon in the third (last) exon.

In an epidemiological study on the molecular basis of  $\alpha$ -thalassemia in Southern Italy, we identified two new variants, each with a deletion of one nucleotide [8,9]. The deletion of C at codon 95 of the  $\alpha 1$ -globin gene gave rise to the *Hb Campania* gene ( $\alpha 1$  cod95 (–C) or *HBA1:c.287delC*). This mutation causes a frameshift and, possibly, production of an  $\alpha$ -chain variant of 101 amino acids (aa),  $\alpha 95$ RSTSSS, which we termed Hb Campania. The deletion of C at codon 109 of the  $\alpha 1$ -globin gene gave rise to the *Hb Sciacca* gene ( $\alpha 1$  cod109 (–C) or *HBA1:c.328delC*), with a frameshift causing the formation of a stop codon at 133, possibly resulting in the synthesis of a truncated  $\alpha$ -chain of 132 aa [8,9]. Both qualitative and semiquantitative analyses of the mRNA from patients' reticulocytes revealed a reduction in mutant mRNA. The in silico analysis shed light on factors that could favor the degradation of mRNA.

Finally, we collected the data on frameshift mutants in the second and third exons of the  $\alpha$ -globin genes reported in literature, and on two dominant inclusion body  $\beta$ -thal, to compare the phenotypes in relation to the presence of mutated mRNA.

## 2. Materials and Methods

### 2.1. Families

In a study on the molecular basis of  $\alpha$ -thalassemia in Southern Italy, the collaborating thalassemia centers selected families and single patients showing a mean corpuscular volume (MCV) lower than 80 fL, which is considered the minimum value for normal Hb A2 and iron status. We determined the molecular basis of  $\alpha$ -thalassemia in the 996 families; here, we include the results for 12 carriers belonging to 6 unrelated families living in Sicilia and Campania. A special committee of the Italian Ministry of University and Research approved the study (Decreto n. 250 of 22 June 1999), and two scientists were appointed as supervisors. The participants signed informed consent for the use of blood samples. Ethical approval of the protocol was obtained from the Comitato Etico Università Federico II of Naples (307/2016 and 225/2019).

### 2.2. Hematological Parameters

Hematological parameters, in blood zinc protoporphyrin (ZPP), and in serum ferritin, transferrin, and bilirubin (total and indirect) were determined by standard procedures in the collaborating hospitals.

The hemoglobin (Hb) analysis was performed via cation exchange high-performance liquid chromatography (HPLC) (Bio-Rad, Diamat or Variant System Hercules, CA, USA). Heinz body formation and thermal and isopropanol Hb stability tests were performed via standard methods.

### 2.3. DNA Molecular Screening

DNA from white blood cells was purified with the salting-out method [2]. Molecular screenings for the most common  $\alpha$ -thalassemia deletions ( $-\alpha 3.7$ ,  $-\alpha 4.2$ ,  $(\alpha)\alpha 5.3$ ) and for the point mutations were carried out by gap-PCR, double gradient–denaturing gradient gel electrophoresis (DG–DGGE), and multiplex amplification refractory mutation system (ARMS), respectively [2,10,11]. Sequencing of the  $\alpha$ -globin genes was carried out using an automatic sequencer as previously reported [10]. An ARMS protocol for the direct identification of the mutation at codon 109 was set up and used for the genotyping of the patients [11]. The research activities were carried out inspired by the principles of quality and developing guidelines for the protocols to increase the reliability and reproducibility of the results [12,13].

### 2.4. $\alpha$ -Globin Haplotype

The analyses of the three single nucleotide polymorphisms (SNPs)—RsaI 5' of the  $\alpha 2$ -globin gene (rs2541669),  $\alpha 2 + 14$  (HBA2:c.–24C > G rs772829778), and  $\alpha 2 + 861$  (HBA2:c.\*136A > G, rs2685121) for the  $\alpha$ -globin haplotype—were performed as previously reported [14,15].

### 2.5. cDNA

mRNA was prepared from reticulocyte-enriched peripheral blood cells and from whole blood, using the Trizol reagent kit, as directed by the manufacturer (Gibco reagent Life Technologies, Inc., Grand Island, NY, USA) [13,14]. In total, 50–200 ng of mRNA was analyzed by reverse transcription PCR (RT-PCR) at a low number of cycles (24 cycles) using rTth DNA polymerase (Roche, Foster City, CA, USA) and the oligonucleotides reported in Table 1. Then, 10  $\mu$ L of the RT-PCR was separated on a 1.5% agarose gel to identify abnormal fragments.

**Table 1.** Oligonucleotides used as primers in the reported applications. The sequences and positions of the primers were derived from GeneBank Sequence \*NG\_000006.1.

Name	Direction, Sequence	Position from the $\alpha 1$ Cap Site	Application	Used with Primer	Amplicon Length (bp)
A For	5'-TGGAGGGTGGAGACGTCCTG-3'	−202/−183	PCR	B Rev	980 bp
B Rev	5'-GGGGGAGGCCCAAGGGGCAAGAA-3'	+755/+778	PCR		
C For	5'-ACAGGCCACCCTCAACCGTCCT-3'	+183/+204	RE DNA (cod95)	D Rev	352 bp
D Rev	5'-GCAACCCGCGTGATCCTCTGCCCT-3'	+511/+534	RE DNA (cod95)		
E For	5'-GGCAAGAAGGTGGCCGACGC-3'	+332/+351	RE cDNA (cod95)	G Rev	293 bp
F For	5'-TGACCCTCTCTCTGCACAGCTC-3'	+584/+606	RE DNA (cod109)	G Rev	
G Rev	5'-GAGGCCCAAGGGGCAAGAAGCAT-3'	+751/+773	Seq $\alpha 1$ ; RT-PCR $\alpha 1$ ; RE DNA and cDNA		194 bp
H For	5'-CGCCCTGAGCGACCTGCACGCG-3'	+400/+421	RE cDNA (cod109)	G Rev	230 bp

RE: restriction enzyme digestion; Seq: sequencing analysis; cod: codon.

### 2.6. Semiquantitative mRNA Analysis by Restriction Enzymes

A semiquantitative analysis of the mutated and normal cDNA was carried out via NlaIV and BseDI (Thermo Fisher Scientific Baltics UAB, Vilnius, Lithuania) restriction enzyme (RE) analysis, respectively. The ratio of mRNA synthesized in cells with the two  $\alpha 1$ -globin alleles was measured in patients that were heterozygous for the Hb variants, taking advantage of the fact that *Hb Campania* and *Hb Sclacca* are associated with the absence of the NlaIV (GGN'NCC) and BseDI (C'CNNGG) restriction sites, respectively. To limit the production of heterodimers, we carried out a first PCR round of 24 cycles and a second round of 1 cycle after increasing the volume of the reaction 20-fold. As a positive control for the enzymatic activity of the RE digestion, and for the normal ratio between the restricted bands, we used the amplicomers from the DNA of WT subjects and Hb Var carriers. To study *Hb Campania*, the DNA of carriers and subjects with WT  $\alpha 1$ -globin was amplified

using the primers A/B in the first round to specifically amplify the  $\alpha$ 1-globin gene, and the primers C/D in the second round; the cDNA was amplified using the primers E/G in both rounds (Table 1). In the case of *Hb Sciacca*, the amplification was performed using the primers F/G in both rounds for the DNA, and H/G in both rounds for cDNA (Table 1).

The amplicomers were digested, according to the manufacturer's recommendations, with 10 U of RE, and the restricted bands were size-fractionated in a 3.5% NuSieve 3:1 agarose gel (FMC, Rockland, Maine, USA) at 70 V for 2–3 h. After separation on the NuSieve gel, a semiquantitative analysis of the bands was performed with the Kodak software Carestream MI (Carestream Health, Inc, Rochester, NY, USA). The ratio of undigested/digested bands (that is, mutated/wild type (WT) DNA and cDNA) was obtained as previously reported [13–16].

### 2.7. Database

All data regarding the families under study and the experimental results were collected anonymously in a database developed on Microsoft Visual Fox 6.0, interfaced with an external software, such as Microsoft Excel or Microsoft Word (Microsoft, Redmond, WA, USA) [17].

### 2.8. In Silico Analysis

The PSIPRED Protein Analysis Workbench was used for the prediction of the secondary structure of the normal and mutant  $\alpha$ -chains (<http://bioinf.cs.ucl.ac.uk/psipred/>, accessed on 12 September 2021) [18].

We evaluated the mutation-induced structural alterations by analyzing the structure of  $\alpha$ -chain of human hemoglobin in the complex with AHSP (PDB code 1Y01 and 1Z8U) and in the tetrameric  $\alpha$ 2 $\beta$ 2 structure (PDB code 2HHB), using the programs Yasara (version 20.4.24) (<http://www.yasara.org/products.htm>, accessed on 12 September 2021) and the Swiss-PdbViewer (version 4.1.0) ([www.expasy.org](http://www.expasy.org), accessed on 12 September 2021) [19–22] (Figures S1 and S2).

The Virtual Ribosome web site was used to identify the stop codon in the *HBA1* cDNA (<https://services.healthtech.dtu.dk/service.php?VirtualRibosome-2.0>, accessed on 22 July 2021) [7].

The programs SIFT (Sorting intolerant from tolerant) ([https://sift.bii.a-star.edu.sg/www/SIFT\\_indels2.html](https://sift.bii.a-star.edu.sg/www/SIFT_indels2.html), accessed on 18 June 2021) (Figure S3), MutationTaster (<http://www.mutationtaster.org/>, accessed on 21 June 2021) (Figure S4), and Splice site prediction (by Neural Network software, [https://www.fruitfly.org/seq\\_tools/splice.html](https://www.fruitfly.org/seq_tools/splice.html), accessed on 30 June 2021) (Figure S5) were used to verify the activation of alternative splicing, ascertain the lengths of abnormal proteins (Figure S6), and determine whether the NMD could trigger the mRNA quality control mechanism [23–25]. The Expasy bioinformatic resource portal was queried for the in-frame translation (Figure S7) and to acquire the protein sequences (<https://web.expasy.org/translate/>, accessed on 21 June 2021) and amino acid compositions of the variant and WT proteins (<https://web.expasy.org/protparam/>, accessed on 22 June 2021) (Figure S8) [26]. The CAIcal Server (<http://genomes.urv.es/CAIcal/>, accessed on 23 June 2021) (Figure S9) and the Sequence manipulation suite (SMS, [https://www.bioinformatics.org/sms2/codon\\_usage.html](https://www.bioinformatics.org/sms2/codon_usage.html), accessed on 22 July 2021) were queried for the codon usage and to compare the mutant and WT mRNA [27,28]. The Kazusa software (<https://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=9606>, accessed on 21 June 2021) was used to determine the frequency of codon usage in the *Homo sapiens* and human target tissue (Figure S10). The mRNA secondary structure was predicted, using the RNAfold web server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>, accessed on 16 June 2021) [29].

### 3. Results

#### 3.1. Hb Campania [ $\alpha 1$ cod95 (–C)]

##### 3.1.1. Molecular Characterization and cDNA Analysis

The new point mutation, giving rise to the *Hb Campania* allele, or  $\alpha 1$  cod95 (–C), was identified in a family from Naples (Figure 1A,B). The two carriers showed mild  $\alpha$ -thalassemia hematological alterations with reductions in the mean corpuscular volume (MCV; 76 and 80 fL) and mean corpuscular hemoglobin (MCH; 24.6 and 23.6 pg). These patients' serum iron, ferritin, transferrin, total bilirubin, and reticulocytes were within the normal ranges. Abnormal hemoglobin or globin chains were not detected via electrophoresis or ion-exchange HPLC. The Hb A2% levels were in the normal range (Table 2).

**Table 2.** Hematological and biochemical data and  $\alpha$ -genotype of the family with *Hb Campania*.

Family Relationship	I-1	I-2	II-1	II-2	II-3
Sex/Age (years)	M/56	F/54	F/25	M/22	F/21
RBC ( $10^{12}/L$ )	4.55	5.16	4.71	5.75	4.55
Hb (g/dL)	13.9	12.7	12.5	13.6	12.5
Ht (L/L)	44.2	41.2	38.8	43.8	38.9
MCV (fL)	97	80	82	76	85
MCH (pg)	30.5	24.6	26.6	23.6	27.5
MCHC (%)	31.4	30.8	32.3	31	32.2
Serum iron ( $\mu\text{g}/\text{dL}$ )	72	155	80	96	62
Ferritin (ng/mL)	78	315	43	154	10
Transferrin (mg/dL)	370	303	232	276	324
Bil tot (mg/dL)	0.38	0.18	0.14	0.22	0.15
Ret (%)	nor	nor	nor	nor	nor
GOR	– – –	+ + –	– – –	– – –	– – –
Hb A2 (%)	2.7	2.4	2.5	2.3	2.7
Hb F (%)	0.0	0.0	0.0	0.0	0.0
$\alpha 1$ cod95 (–C) carrier	no	yes	no	yes	no

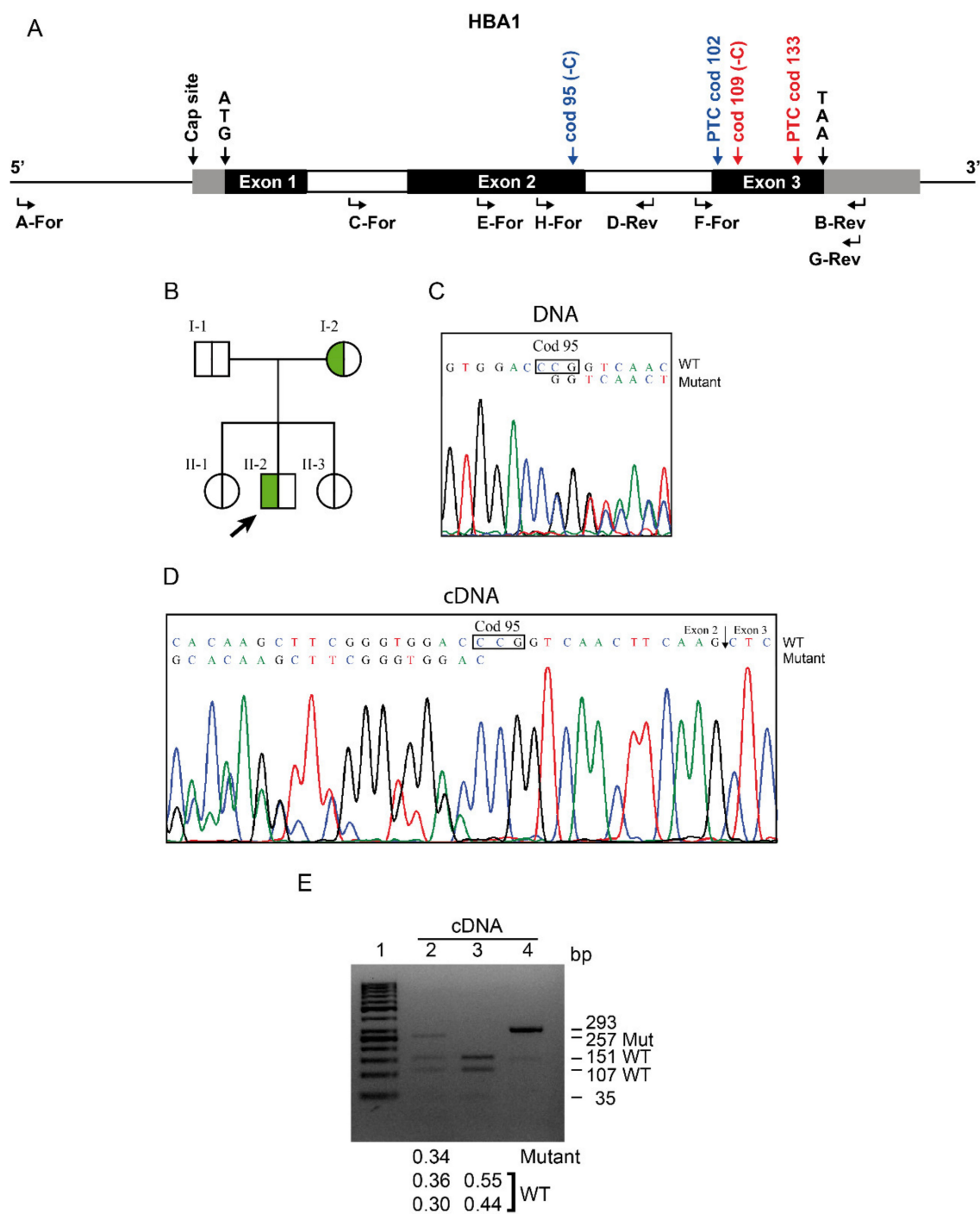
RBC: red blood cells; Hb: hemoglobin; Ht: hematocrit; MCV: mean corpuscular volume; MCH: mean corpuscular hemoglobin; MCHC: mean corpuscular hemoglobin concentration; Bil tot: total bilirubin; Ret: reticulocytes; GOR: globular osmotic resistance.

Screening for the  $\alpha$ -thalassemia deletions gave negative results, and the sequencing analysis of the  $\alpha 1$ - and  $\alpha 2$ -globin genes only revealed a cytidine deletion at codon 95 of the  $\alpha 1$ -globin gene. The mutation was confirmed by sequencing in the other members of the family (Figure 1C). The  $\alpha 1$  cod95 (–C) mutation caused a frameshift and, possibly, production of an  $\alpha$ -chain variant of 101 aa,  $\alpha 95\text{RSTSSS}$ .

The RT-PCR and the sequencing of  $\alpha 1$ -globin cDNA, performed on mRNA purified from reticulocytes from fresh blood, indicated a frameshift at codon 95, but this mutated sequence exhibited base peaks much smaller than those of the WT sequence (Figure 1D).

To quantify the mutated mRNA, we performed a semiquantitative analysis by digestion with the NlaIV RE, for which the mutation eliminates a restriction site, as shown in Figure 1E. The DNA digestion confirmed the presence, in the carriers, of an anomalous band of 285 bp, specific to the *Hb Campania*. The relative amount of this anomalous band was comparable (0.50) to the sum of the relative amount of the two WT bands (225 and 61 bp) on DNA of the *Hb Campania* heterozygote, indicating the presence of the WT and mutant alleles (Figure S11A). Otherwise, the digestion on cDNA from the reticulocytes of the carrier indicated that the relative amount of the anomalous 257 bp band, specific to *Hb Campania*, was 0.34 respect to the total  $\alpha 1$ -globin cDNA, as shown in Figure 1E. These data confirmed a consistent reduction in *Hb Campania* cDNA.





**Figure 1.** Molecular characterization and cDNA analysis of *Hb Campania*. (A) Scheme of the functional structure of the  $\alpha 1$ -globin gene (*HBA1*), indicating the position of cod95 (–C) and cod109 (–C) with their relative premature termination codon (PTC). The gray rectangles indicate the 5' and 3' UTR regions, the white rectangles the introns. The positions and orientations of the primers used for the molecular characterization are indicated with arrows placed under the gene. (B) Pedigree of the family. The arrow indicates the proband; green indicates the carriers of *Hb Campania*. (C,D)  $\alpha 1$ -globin gDNA (C) and cDNA (D) sequences of a carrier of *Hb Campania*. (E) The cDNA amplicomers of 293 bp, digested by the restriction enzyme *NlaIV*, and separated on a 3.5% NuSieve 3:1 agarose gel. Lane 1: 50 bp ladder; Lane 2: cDNA of the *Hb Campania* carrier; Lane 3: cDNA of the control subject; Lane 4: undigested cDNA sample. The fragments' lengths are reported on the right. The *Hb Campania* eliminates the *NlaIV* restriction site GGA'CCC, generating an anomalous longer cDNA band of 257 bp, corresponding respectively to the sum of the two WT-specific bands of 151 and 107 bp, minus the deleted cytidine base. The relative amounts of the longer abnormal band and the WT-bands are reported in the lower section.

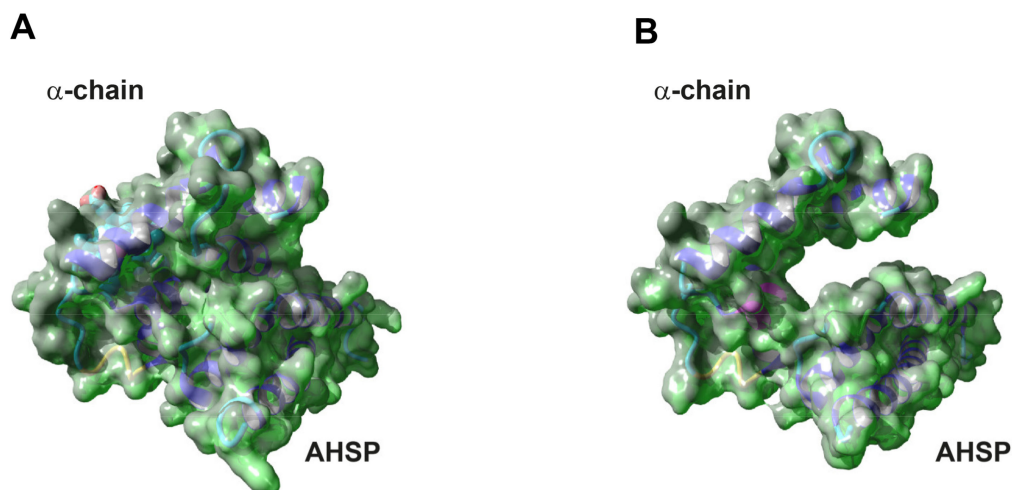
### 3.1.2. 3D Modeling

To define the effect of the frameshift on the protein stability, a 3D model of the Hb Campania  $\alpha$ -chain was created by means of SWISS-MODEL. The structures of the WT  $\alpha$ -chain interacting with AHSP (1Y01, 1Z8U) and in the Hb tetramer (2HHB) were used to fit the Hb Campania  $\alpha$ -chain [21,22]. The analysis of the 3D models (Figure S1A–F) highlighted two aspects: the Hb Campania is superimposed with the WT  $\alpha$ -chain also in the modified C-terminal region, where the change of the 6 aa is present; an empty region is evident in the variant chain for the absence of the two prominent G and H helices.

The 6 mutated aa are posed in the central cavity; these 6 aa, in the WT  $\alpha$ -chain, are involved in the  $\alpha 1\beta 1$  contacts and three of them in the heme contact (cod 97, 98, 101) [30]. The frameshift introduces 1 polar and 5 small non-polar aa instead of 2 polar, 1 small non-polar, 2 hydrophobic, and 1 aromatic aa (Figure S2). These differences most likely cause disturbances on the functions of these 6 aa.

The generation of the 3D surface models of the WT and Hb Campania  $\alpha$ -chains in complex with AHSP better highlighted the presence of a big cavity in the Hb Campania that compromises multiple functions: the central cavity is exposed; the interaction with the heme group is not stabilized; the interactions with AHSP are insufficient to stabilize the  $\alpha$ -chain (Figure 2A,B) [5,30,31]. The 3D model of Hb Campania in the Hb tetramer is equally compromised by the absence of the two G and H helices important also in the interaction between the  $\alpha$ - and  $\beta$ -chains (Figure S1D–F).

The analyses of the 3D models indicate that the Hb Campania is unstable.



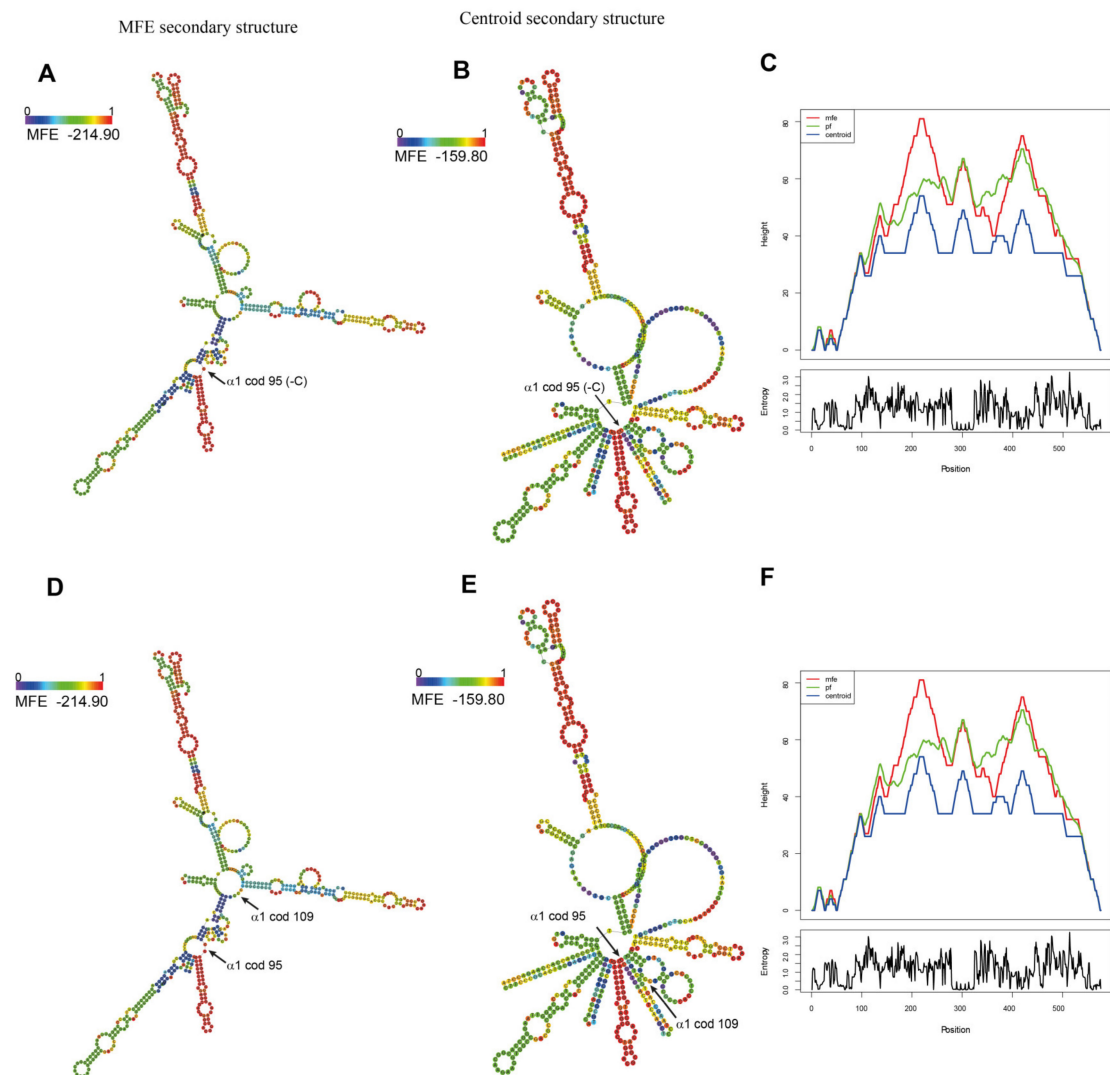
**Figure 2.** The 3D surface model of the Hb Campania. (A) The 3D surface model of the WT  $\alpha$ -chain and of the Hb Campania (B) in complex with AHSP (PDB code 1Z8U).

### 3.1.3. In Silico Analyses

To investigate the possible causes of the reduction in mRNA, we analyzed, in silico, alternative splicing, mRNA structure, and codon usage.

The analysis of alternative splicing gave negative results, suggesting that the deletion cod95 (–C) has no impact on the consensus sequence for splicing (Figures S3–S5). The RT-PCR amplification of the total length  $\alpha 1$ -globin mRNA confirmed the absence of an anomalous fragment and, thus, of alternative splicing (data not shown).

The analysis of alterations in the mRNA's structural conformation indicated that the structure and stackable minimum free energy (MFE) of *Hb Campania* (Figure 3A–C) were comparable to those of WT mRNA (Figure 3D–F); therefore, this mechanism seems uninvolved in the activation of mRNA quality control [32–34].



**Figure 3.** Secondary structure of  $\alpha 1$ -globin mRNAs predicted by the RNAfold web server. (A,B,C)  $\alpha 1$  cod95 (-C) mRNA. (D,E,F) WT  $\alpha 1$ -globin mRNA. (A,D) Minimum free energy (MFE) secondary structure. (B,E) Centroid secondary structure with the corresponding values. (C,F) Mountain plot representation of the MFE structure for each  $\alpha 1$ -globin mRNA. The values of  $\alpha 1$  cod95 (-C) were comparable with the values of WT  $\alpha 1$  mRNA.

The *in silico* analyses of amino acid composition and codon usage in Hb Campa-  
 nia highlighted several changes, caused mainly by the considerable reduction in length:  
 25 codons were used fewer times, 1 codon was used more frequently, and there were  
 2 codons used for the first time (Figures 4 and S6–S8). For the latter two classes, the analysis  
 of human codon usage showed high frequencies for the three codons CGG (1 + 1 = 2) (11.4),  
 TCA (0 + 2 = 2) (12.2), and ACT (0 + 1 = 1) (13.1) (Figures 4, S9 and S10). Moreover, the  
 codon CGG was absent in the mRNA encoding the  $\beta$ -chain and present once in the one  
 encoding the  $\alpha$ -chain; the codon TCA was absent in both these mRNAs, while ACT was  
 present three times in the  $\beta$ -globin mRNA and absent in the  $\alpha$ -globin mRNA. These data  
 highlight that in reticulocytes, where the synthesis of globin chains is the highest, the two  
 tRNAs for the CGG and TCA codons may be less common, and therefore, the synthesis  
 of the variant globin chain may be slower (Figure 4) [35]. *In silico* analyses of the codon  
 usage both in red blood cells and in the WT hemoglobins confirmed the low frequencies of  
 CGG and TCA in the target tissue (Figure S10B,C).



A																					
ATG	GTG	CTG	TCT	CCT	GCC	GAC	AAG	ACC	AAC	GTC	AAG	GCC	GCC	TGG	GGT	AAG	GTC	GGC	GCG		
M	V	L	S	P	A	D	K	T	N	V	K	A	A	W	G	K	V	G	A		
CAC	GCT	GGC	GAG	TAT	GGT	GCG	GAG	GCC	CTG	GAG	AGG	ATG	TTC	CTG	TCC	TTC	CCC	ACC	ACC		
H	A	G	E	Y	G	A	E	A	L	E	R	M	F	L	S	F	P	T	T		
AAG	ACC	TAC	TTC	CCG	CAC	TTC	GAC	CTG	AGC	CAC	GGC	TCT	GCC	CAG	GTT	AAG	GGC	CAC	GGC		
K	T	Y	F	P	H	F	D	L	S	H	G	S	A	Q	V	K	G	H	G		
AAG	AAG	GTG	GCC	GAC	GCG	CTG	ACC	AAC	GCC	GTG	GCG	CAC	GTG	GAC	GAC	ATG	CCC	AAC	GCG		
K	K	V	A	D	A	L	T	N	A	V	A	H	V	D	D	M	P	N	A		
CTG	TCC	GCC	CTG	AGC	GAC	CTG	CAC	GCG	CAC	AAG	CTT	CGG	GTG	GAC	CGG	TCA	ACT	TCA	AG/C		
L	S	A	L	S	D	L	H	A	H	K	L	R	V	D	R	S	T	S	S		
TCC	TAA																				
S	stop																				

B					
AA	n.	Codon (time present)	AA	n.	Codon (time present)
Ala, A	15↓	GCC (12 - 4 = 8), GCG (7 - 1 = 6), GCT (2 - 1 = 1), GCA (0)	Leu, L	8↓, 1=	CTG (14 - 6 = 8), CTC (2 - 2 = 0), CTT (1), CTA (1 - 1 = 0), TTA (0), TTG (0)
Arg, R	2↑, 0↓, 1=	CGT (1 - 1 = 0), CGG (1 + 1 = 2) ( <b>11.4</b> ), AGG (1), CGC (0), CGA (0), AGA (0)	Lys, K	8↓	AAG (10 - 2 = 8), AAA (1 - 1 = 0)
Asn, N	3↓	AAC (4 - 1 = 3), AAT (0)	Met, M	2=	ATG (2)
Asp, D	7↓	GAC (8 - 1 = 7), GAT (0)	Phe, F	4↓	TTC (7 - 3 = 4), TTT (0)
Cys, C	0↓	TGC (1 - 1 = 0), TGT (0)	Pro, P	4↓	CCC (3 - 1 = 2), CCT (2 - 1 = 1), CCG (2 - 1 = 1), CCA (0)
Gln, Q	1=	CAG (1), CAA (0)	Ser, S	8↓, 2N	TCC (4 - 2 + 1 = 3), AGC (4 - 2 + 1 = 3), TCT (3 - 1 = 2), TCA (0 + 2 = 2) ( <b>12.2</b> ), TCG (0), AGT (0)
Glu, E	3↓	GAG (4 - 1 = 3), GAA (0)	Thr, T	5↓, 1N	ACC (9 - 4 = 5), ACT (0 + 1 = 1) ( <b>13.1</b> ), ACA (0), ACG (0)
Gly, G	7=	GGC (5), GGT (2), GGA (0), GGG (0)	Trp, W	1=	TGG (1)
His, H	7↓	CAC (10 - 3 = 7), CAT (0)	Tyr, Y	2↓	TAC (2 - 1 = 1), TAT (1)
Ile, I	0=	ATT (0), ATC (0), ATA (0)	Val, V	7↓, 1=	GTG (9 - 4 = 5), GTC (3 - 1 = 2), GTT (1), GTA (0)
Start	1=	ATG (1)	Stop	1	TAG (0), TGA (0), TAA (1 - 1 + 1 = 1)

Tot: 101 aa + TAA stop codon

↓: Decrease; ↑: Increase; =: Invariant; N: New.

**Figure 4.** (A) Nucleotide triplets coding the *Hb Campania* ( $\alpha 1$  cod95 (-C)) mRNA, from the ATG to the stop codon at position 102; below, the corresponding amino acids present in this  $\alpha$ -globin chain variant. The changes in amino acids following the frameshift after cod95 are shown in gray. The symbol / indicates the position of the exon2/3 boundary. (B) Types of amino acids present in the *Hb Campania*  $\alpha$ -globin chain, the corresponding coding triplet, and the number of appearances in the *Hb Campania* mRNA. The numbers underlined and in bold indicate the codon usage frequencies.

### 3.2. *Hb Sciacca* [ $\alpha 1$ cod109 (-C)]

#### 3.2.1. Molecular Characterization and cDNA Analysis

In 10 patients from 5 apparently unrelated families originating from Sicily, we identified a rare  $\alpha$ -globin chain variant created by the deletion of C at cod109 of the  $\alpha 1$ -globin gene, named *Hb Sciacca* or HBA1: c.328delC [5,7].

The probands showed normal iron balance, no clinical symptoms and were sent to us for evaluation of their mild microcytic anemia. All 10 carriers showed  $\alpha$ -thalassemia alterations (microcytosis with normal Hb A2). Their MCV ranged from 73.2 fL to 81.8 fL, and their MCH from 23.8 pg to 27.1 pg. Their Hb A2 % level was in the normal range (Table 3).

**Table 3.** Hematological and biochemical data and  $\alpha$ -globin genotype of the family with *Hb Sciacca*.

Family	A				B				C			D		E	
	I-1	I-2	II-1	II-2	I-1	II-1	II-2 *	II-2 *	II-3	I-1	I-2	II-1	II-1	II-2	I-1
Sex/Age (yrs)	M/38	F/29	F/09	M/05	M/69	M/39	F/35	F/35	M/32	M/56	F/50	M/31	M/30	F/23	F/21
RBC ( $10^{12}/L$ )	5.1	5.10	5.30	5.20	4.77	4.76	4.66	5.13	5.07	5.65	5.07	6.30	6.32	5.63	4.62
Hb (g/dL)	16.1	13.2	13.4	13.1	14.6	12.9	12.1	13.4	15.6	15.3	15	16.4	15.1	13.4	12.3
Ht (L/L)	45.4	39.1	41.0	38.4	40.5	37.4	37.0	40.6	44	46.2	44.9	49.5	47.3	41.2	34.6
MCV (fL)	89.3	76.3	77.7	74.2	84.9	78.6	79.4	79.1	86.8	81.8	88.5	78.6	74.8	73.2	75.0
MCH (pg)	31.7	25.8	25.4	25.3	30.6	27.1	26.0	26.1	30.8	27.1	29.6	26.0	23.9	23.8	26.6
MCHC (%)	35.5	33.8	32.7	34.1	36	34.5	32.7	33.0	35.5	33.1	33.4	33.1	31.9	32.5	35.6
Heinz body	nt	nt	nt	nt	nt	nt	nt	absent	nt	nt	nt	nt	nt	nt	nt
Erythro morph	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt
Pink test (%)	nt	nt	nt	nt	nt	nt	nt	3.0	nt	nt	nt	nt	nt	nt	nt
Transferrin (mg/dL)	nt	nt	nt	nt	nt	nt	nt	262	nt	nt	nt	nt	nt	nt	nt
Bilir tot (mg/dL)	nt	nt	nt	nt	nt	nt	nt	0.5	nt	nt	nt	nt	nt	nt	nt
Bilir dir (mg/dL)	nt	nt	nt	nt	nt	nt	nt	0.14	nt	nt	nt	nt	nt	nt	nt
Hapt (mg %)	nt	nt	nt	nt	nt	nt	nt	175	nt	nt	nt	nt	nt	nt	nt
LDH (U/L)	nt	nt	nt	nt	nt	nt	nt	279	nt	nt	nt	nt	nt	nt	nt
Ret (%)	nt	nt	nt	nt	nt	nt	nt	1.04	nt	nt	nt	nt	nt	nt	nt
ZPP ( $\mu$ g/dL)	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	34
Ferritin (ng/mL)	131	72	56	24	48	102	127	nt	0	113	17	78	67	19	0
Hb A2 (%)	2.9	2.9	3.1	2.9	2.9	2.4	2.3	nt	3	3.1	2.7	2.8	2.7	2.7	2.7
Hb F (%)	0	0	0	0	0	0	0.5	nt	0	0	0	0	0.5	0.5	0.3
$\alpha$ 1 cod109 (-C) carrier	no	yes	no	yes	no	yes	yes	yes	no	yes	no	yes	yes	yes	yes

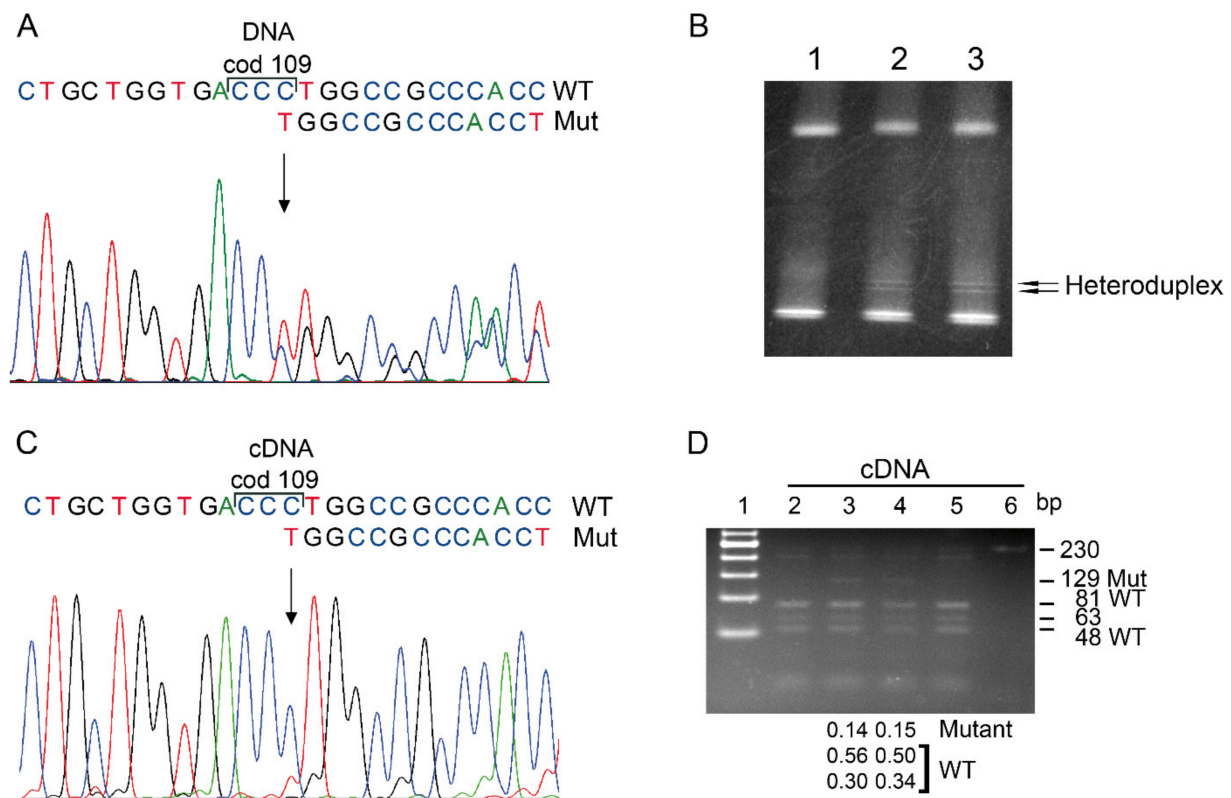
RBC: red blood cells; Hb: hemoglobin; Ht: hematocrit; MCV: mean corpuscular volume; MCH: mean corpuscular hemoglobin; MCHC: mean corpuscular hemoglobin concentration; Erythro morph: erythrocyte morphology; Bilir tot: total bilirubin; Bilir dir: direct bilirubin; Hapt: haptoglobin; LDH: lactate dehydrogenase; Ret: reticulocytes; ZPP: zinc protoporphyrin; A: anisocytosis; P: poikilocytosis; H: hypochromia; nt: not tested; \* = same person.

The proband II.2 of family B was reexamined and displayed reticulocytes, indirect bilirubin, haptoglobin, LDH, and pink test results within the normal range as well as the absence of Heinz bodies (Table 3). No instability test could be performed on fresh blood, but the analysis in our laboratory after shipping, was normal. All these data indicated the absence of hemolytic processes. The HPLC and electrophoresis carried out on the hemolysate revealed no *Hb Sciacca*.

Gap-PCR excluded the presence of any of the following  $\alpha$ -thalassemia alleles:  $-\alpha$ 3.7,  $-\alpha$ 4.2, and  $(\alpha)\alpha$ 5.3. The double gradient-denaturing gradient gel electrophoresis (DG-DGGE) of five DNA PCR amplicomers, spanning the  $\alpha$ 1- and  $\alpha$ 2-globin genes, detected an abnormal pattern in their third exons (Figure 5B). The sequencing of anomalous amplicomers identified the rare mutation  $\alpha$ 1 cod109 (-C), which causes a frameshift (Figure 5A) and modifies the C-terminal sequence, creating an  $\alpha$ -chain variant of 132 amino acids:  $\alpha$ 109WPPTSPPSSPLRCTPPWTSSWLL (Figures S6–S8). No other mutation was identified via the sequencing of the  $\alpha$ 1- and  $\alpha$ 2-globin genes. The mutation was confirmed in all members of the families, using the amplification refractory mutation system (ARMS). Analysis of the three SNPs RsaI(+), +14(-), and +861(-) identified the same  $\alpha$ -globin haplotype in each of the five families with *Hb Sciacca*.

A qualitative and semiquantitative analysis on the  $\alpha$ -globin mRNA was performed to evaluate its level of expression. RT-PCR and cDNA sequencing performed on the mRNA from reticulocytes in blood identified a frameshift at cod109, but the variant sequence  $\alpha$ 1 cod109 (-C) showed base peaks much smaller than those of the WT sequence (Figure 5C). In order to quantify the mutated mRNA, we performed a semiquantitative analysis by digestion with the BseDI restriction enzyme, for which the mutation eliminates a restriction

site. The DNA digestion confirmed, in the carriers, an anomalous 93 bp band, specific to the *Hb Sclacca*. The relative amount of these anomalous bands constituted 54% and 58% of the total  $\alpha 1$ -globin gene bands in the two carriers. These data confirmed that both the alleles *Hb Sclacca* and WT  $\alpha 1$ -globin gene are present in the carriers (Figure S11B).



**Figure 5.** Molecular characterization and cDNA analysis of *Hb Sclacca*. (A)  $\alpha 1$ -globin gDNA sequence of an *Hb Sclacca* carrier. (B) Denaturing gradient gel electrophoresis (DGGE) of amplicomer III of the  $\alpha$ -globin genes containing codon 109. Lane 1: subject with WT  $\alpha 1$ -globin; Lanes 2 and 3: *Hb Sclacca* heterozygotes. (C)  $\alpha 1$ -globin cDNA sequence of an *Hb Sclacca* carrier. (D) The cDNA amplicomers of 230 bp, digested with the restriction enzyme *BseDI* and separated on a 3.5% NuSieve 3:1 agarose gel. Lane 1: 50 bp ladder; Lanes 2 and 5: cDNA of subjects with WT  $\alpha 1$ -globin; Lanes 3 and 4: cDNA of the *Hb Sclacca* heterozygotes; Lane 6: undigested cDNA sample. The *Hb Sclacca* eliminates the *BseDI* restriction site C'CCTGG, generating an anomalous longer cDNA band of 129 bp, corresponding to the sum of the two WT-specific bands of 81 and 48 bp, minus the deleted cytidine base. The fragments' lengths are reported on the right. The relative quantities of longer abnormal bands and WT bands are reported in the lower section.

Otherwise, the digestion on cDNA from the reticulocytes of the two carriers indicated that the relative amount of the anomalous 129 bp bands, specific to *Hb Sclacca*, were 14% and 15% of the total  $\alpha 1$ -globin mRNA as shown in Figure 5D.

These data indicated an unexpected consistent reduction in *Hb Sclacca* cDNA.

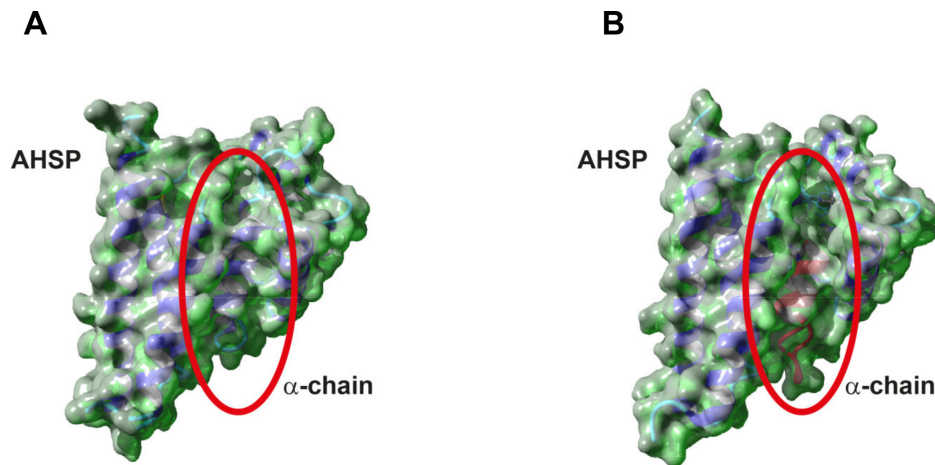
### 3.2.2. 3D Modeling

To define the effect of the frameshift on the protein stability a 3D model of the *Hb Sclacca*  $\alpha$ -chain was created (Figure S1G–L) as already reported for *Hb Campania*.

The analysis of the composition of the 23 mutated aa indicated a marked reduction in the polar (4% instead of 21%) and hydrophobic aa (13% instead of 24%) and an increase in small non-polar (65% instead of 45%) and aromatic aa (17% instead of 9%) (Figure S2). This consistent change in the number and type of aa between the WT  $\alpha$ -chain and the *Hb Sclacca* suggests a negative impact of the variations on the inter- and intra-chains interactions.

The 3D superimposed model highlights that the *Hb Sclacca* retains almost the same tertiary structure up to the F helix, but overall, it is evident the presence of a longer mutated

GH non-helix region of 11 aa (instead of 6 aa) at cod113-123, and of a shorter mutated H helix of only 7 aa (instead of 21 aa) (Figure S1G–L). The 3D surface model of the Hb Sciacca  $\alpha$ -chain confirmed and highlighted its structure variation (Figure 6A,B).



**Figure 6.** The 3D surface model of the Hb Sciacca. (A) The 3D surface model of the WT  $\alpha$ -chain and of the Hb Sciacca (B) in complex with AHSP (PDB code 1Z8U). The red circles indicate the position of 23 mutated aa (in magenta) and the small cavity for the absence of the H helix in the Hb Sciacca (B) and the corresponding position in the WT (A).

The H helix in the WT  $\alpha$ -chain has important roles; its shortening and variation in the composition, observed in the Hb Sciacca, leads to the alteration and destabilization of the tertiary structure. The Hb Sciacca impairs the correct formation of the central cavity and of the heme pocket—for the modifications at cod 129 Leu>Pro (H2) and the absence of the Leu 136, both involved in the heme contact—and it impairs also the correct interaction with AHSP for the mutation, among others, at cod 117 (G1) Phe>Pro [3–5,30,31].

The tertiary structure is also modified for the presence of a bulky non-helix GH region, which, in the WT  $\alpha$ -chain, is in an external position and involved in the  $\alpha$ 1 $\beta$ 1 contacts, while, in the Hb Sciacca, it most likely creates interference both in the interaction with AHSP and with the  $\beta$ -chain (Figure 6A,B).

The analyses of the 3D models indicate that the Hb Sciacca is unstable.

### 3.2.3. In Silico Analyses

To understand the causes of the reduction of the *Hb Sciacca* mRNA, we performed an in silico analysis to highlight the mechanisms that could trigger the mRNA quality control decay.

The prediction of variations in the splicing site ([https://www.fruitfly.org/seq\\_tools/splice.html](https://www.fruitfly.org/seq_tools/splice.html), accessed on 30 June 2021), SIFT (accessed on 18 June 2021) and MutationTaster (accessed on 21 June 2021), Figures S3–S5 indicated that the mutation does not change the splicing score (0.02) of the cryptic splicing sequence (cctgctgGTgacct  $\rightarrow$  cctgctgGTgacctg) that normally occurs between codons 104 and 109. The RT-PCR amplification of the total length  $\alpha$ 1-globin mRNA confirmed the absence of anomalous fragments and thus, of alternative splicing.

The analysis of amino acid composition (Figures S6–S8) and codon usage (Figures S9 and S10) in the Hb Sciacca variant indicated several changes: 16 codons were used fewer times, 6 codons more frequently, and there 5 were novel codons. Among the last two classes, two changes are noteworthy: the CCG codon was present five times instead of two, and ACG appeared once (Figure 7). Data on codon usage in humans (<https://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=9606>, accessed on 21 June 2021) show that CCG and ACG are rarely used, fifth (6.9) and third (6.1) least frequent codons, respectively, excluding the stop codons (Figure S10A). The analysis of human codon usage both in red blood cells and in WT hemoglobin confirmed low frequen-

cies of these two codons (Figure S10B,C). These data suggest that the slowing down of the synthesis of Hb Sclacca, due to the presence of CCG and ACG, could activate the no-go decay [32,33].

A																			
ATG	GTG	CTG	TCT	CCT	GCC	GAC	AAG	ACC	AAC	GTC	AAG	GCC	GCC	TGG	GGT	AAG	GTC	GGC	GCG
M	V	L	S	P	A	D	K	T	N	V	K	A	A	W	G	K	V	G	A
CAC	GCT	GGC	GAG	TAT	GGT	GCG	GAG	GCC	CTG	GAG	AGG	ATG	TTC	CTG	TCC	TTC	CCC	ACC	ACC
H	A	G	E	Y	G	A	E	A	L	E	R	M	F	L	S	F	P	T	T
AAG	ACC	TAC	TTC	CCG	CAC	TTC	GAC	CTG	AGC	CAC	GGC	TCT	GCC	CAG	GTT	AAG	GGC	CAC	GGC
K	T	Y	F	P	H	F	D	L	S	H	G	S	A	Q	V	K	G	H	G
AAG	AAG	GTG	GCC	GAC	GCG	CTG	ACC	AAC	GCC	GTG	GCG	CAC	GTG	GAC	GAC	ATG	CCC	AAC	GCG
K	K	V	A	D	A	L	T	N	A	V	A	H	V	D	D	M	P	N	A
CTG	TCC	GCC	CTG	AGC	GAC	CTG	CAC	GCG	CAC	AAG	CTT	CGG	GTG	GAC	CCG	GTC	AAC	TTC	AAG/
L	S	A	L	S	D	L	H	A	H	K	L	R	V	D	P	V	N	F	K
CTC	CTA	AGC	CAC	TGC	CTG	CTG	GTG	ACC	TGG	CCG	CCC	ACC	TCC	CCG	CCG	AGT	TCA	CCC	CTG
L	L	S	H	C	L	L	V	T	W	P	P	T	S	P	P	S	S	P	L
CGG	TGC	ACG	CCT	CCC	TGG	ACA	AGT	TCC	TGG	CTT	CTG	TGA							
R	C	T	P	P	W	T	S	S	W	L	L	Stop							

B					
AA	n.	Codon (time present)	AA	n.	Codon (time present)
Ala, A	15↓	GCC (12 - 4 = 8), GCG (7 - 1 = 6), GCT (2 - 1 = 1), GCA (0)	Leu, L	13↓, 2↑, 1=	CTG (14 - 2 = 12), CTC (2 - 1 = 1), CTT (1 + 1 = 2) ( <b>13.2</b> ), CTA (1), TTA (0), TTG (0)
Arg, R	2↑, 0↓, 1=	CGG (1 + 1 = 2) ( <b>11.4</b> ), CGT (1 - 1 = 0), AGG (1), CGC (0), CGA (0), AGA (0)	Lys, K	9↓	AAG (10 - 1 = 9), AAA (1 - 1 = 0)
Asn, N	4=	AAC (4), AAT (0)	Met, M	2=	ATG (2)
Asp, D	7↓	GAC (8 - 1 = 7), GAT (0)	Phe, F	5↓	TTC (7 - 2 = 5), TTT (0)
Cys, C	2↑	TGC (1 + 1 = 2) ( <b>12.6</b> ), TGT (0)	Pro, P	10↑, 2=	CCC (3 + 2 = 5) ( <b>19.8</b> ), CCT (2), CCG (2 + 3 = 5) ( <b>6.9</b> ), CCA (0)
Gln, Q	1=	CAG (1), CAA (0)	Ser, S	4=, 5↓, 3N	TCC (4), AGC (4 - 1 = 3), TCT (3 - 1 = 2), TCA (0 + 1 = 1) ( <b>12.2</b> ), AGT (0 + 2 = 2) ( <b>12.1</b> ), TCG (0)
Glu, E	3↓	GAG (4 - 1 = 3), GAA (0)	Thr, T	7↓, 2N	ACC (9 - 2 = 7), ACG (0 + 1 = 1) ( <b>6.1</b> ), ACA (0 + 1 = 1) ( <b>15.1</b> ), ACT (0)
Gly, G	7=	GGC (5), GGT (2), GGA (0), GGG (0)	Trp, W	4↑	TGG (1 + 3 = 4) ( <b>13.2</b> )
His, H	8↓	CAC (10 - 2 = 8), CAT (0)	Tyr, Y	1=, 1↓	TAC (2 - 1 = 1), TAT (1)
Ile, I	0	ATT (0), ATC (0), ATA (0)	Val, V	4=, 6↓	GTG (9 - 3 = 6), GTC (3), GTT (1), GTA (0)
Start	1=	ATG (1)	Stop	1↑, 0↓	TAG (0), TGA (0 + 1 = 1), TAA (1 - 1 = 0)

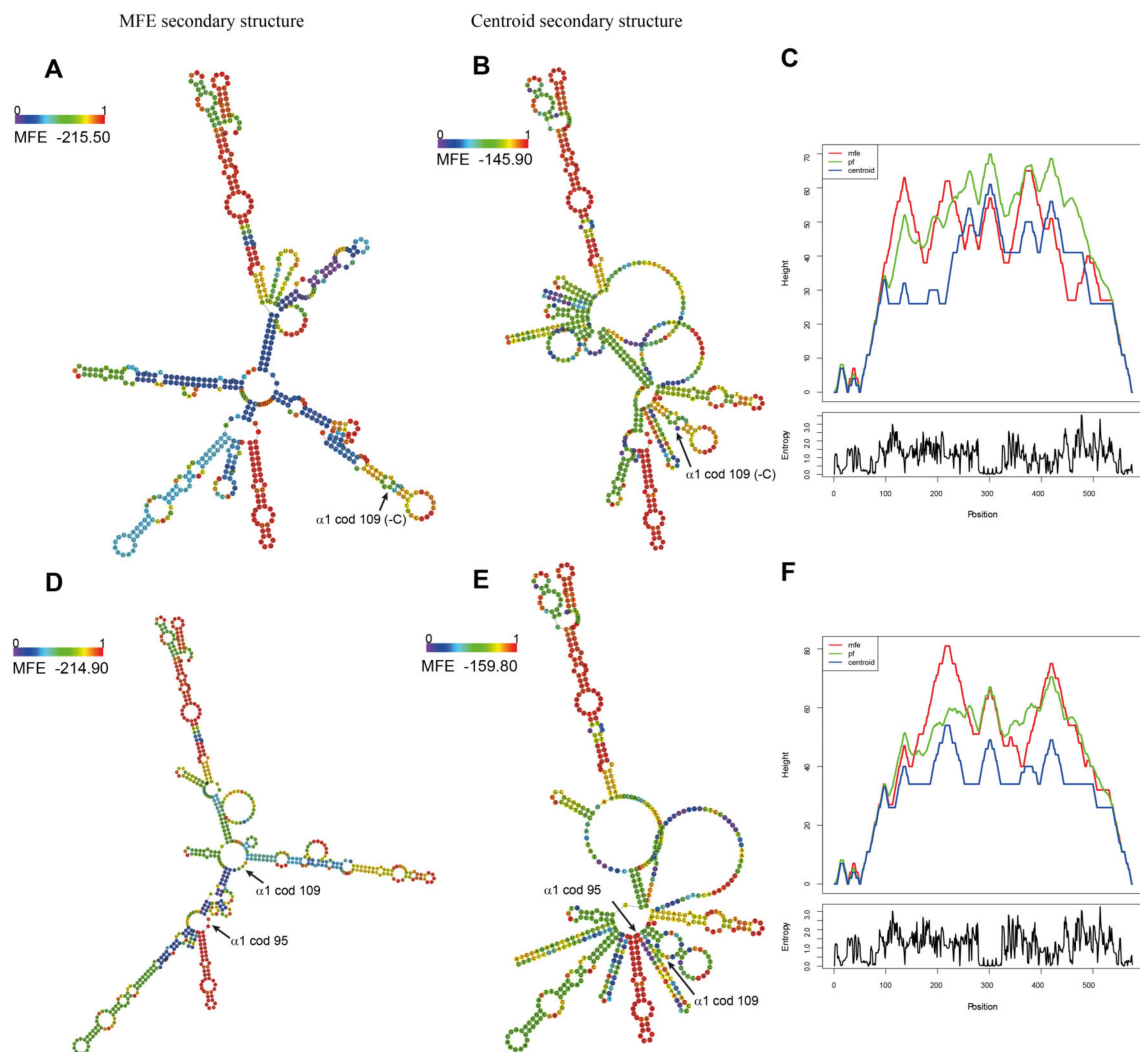
Tot: 132 aa + TGA stop codon

↓: Decrease; ↑: Increase; =: Invariant; N: New.

**Figure 7.** (A) Nucleotide triplets coding the Hb Sclacca mRNA, from the ATG to the stop codon at position 132; below, the corresponding amino acids present in this α-globin chain variant. The changes in Hb Sclacca amino acids (aa) following the frameshift from cod109 are marked in gray. The symbol/indicates the position of the exon2/3 boundary. (B) Types of aa present in the Hb Sclacca α-globin chain, the corresponding coding triplet, and the number of occurrences in the Hb Sclacca mRNA. The numbers underlined and in bold indicate the codon usage frequencies.

We also investigated the alterations in mRNA structural conformation that could stall the ribosomes during translation [32–34]. Hb Sclacca showed only a minimal variation in MFE (−215.50 instead of −214.90) (Figure 8A–C) respect to the WT α-globin mRNA (Figure 8D–F), but a consistent alteration in the secondary structure, with an increase in the number of lateral branches that could interfere with the sliding of the ribosome.





**Figure 8.** Secondary structure of  $\alpha 1$ -globin mRNAs predicted by the RNAfold web server. (A,B,C)  $\alpha 1$  cod109 (–C) mRNA. (D,E,F) WT  $\alpha 1$ -globin mRNA. (A,D) Minimum free energy (MFE) secondary structure. (B,E) Centroid secondary structure with the corresponding values. (C,F) Mountain plot representation of the MFE structure for each  $\alpha 1$ -globin mRNA. The values of  $\alpha 1$  cod109 (–C) presented a small variation in free energy of the secondary structure. The centroid secondary structure of  $\alpha 1$  cod109 (–C) presented a variation in free energy. The presence of differences in the secondary structure conformation of the  $\alpha 1$  cod109 (–C) mRNA could alter the access of the ribosomal apparatus.

These *in silico* analyses identified two mechanisms that could stall translation—the alteration of mRNA stability and structure, and the presence of codons that are rarely represented in the cell. These factors could induce no-go decay.

#### 4. Discussion

The frameshift mutations in the third exons of the  $\alpha$ - and  $\beta$ -globin genes generate variants with altered structures and lengths, which, in most cases, give rise to hyper-unstable hemoglobin [1,3]. Mutants of this type have two peculiar characteristics: one related to the phenotype and the other to the mRNA quality control.

Carriers of these variants, mainly of mutations in the  $\beta$ -globin gene, can exhibit the dominant form of thalassemia, resulting in a thalassemia intermedia phenotype in individuals who have inherited only a single copy of the abnormal globin gene [1]. The dominant phenotype associated with frameshift variants has also been identified in other hereditary diseases [36–38].

The second peculiarity is that the premature termination codon (PTC) generated by the frameshift, and the change in reading frame it imposes, is not recognized by the nonsense-mediated decay (NMD) system when in the third (last) exon [4,39].

#### 4.1. Hb Campania [ $\alpha 1$ cod95 (–C)]

In two members of a family from Campania, we identified a novel cytidine deletion at cod95, associated with a mild  $\alpha$ -thalassemia phenotype (Figure 1B,C). The mutation has not been identified in other Italian or Mediterranean areas.

This novel  $\alpha 1$  cod95 (–C) mutation results in a frameshift that causes the variation in 6 aa at position cod96–101 and a PTC at position 102 (Figure 1A) with the absence of the last 41 aa. Hb Campania is the first discovered deletion of one bp that causes a frameshift after cod84 and the formation of a truncated  $\alpha$ -globin chain at cod 102.

This variant is expected to consist of 101 aa and has an altered C-terminal aa sequence,  $\alpha 95$ RSTSSS (Figures 4, S4 and S6–S8). The absence of the 41 residues, corresponding to the G and H helices, most likely impairs several functions.

The G and H helices play an important role in the intra-chain bonds, such as the formation of the central cavity and of the heme pocket, but also in the inter-chains interactions, through the specific recognition of AHSP and  $\alpha 1\beta 1$  contacts [5,30,31].

The analysis of the 3D models of the Hb Campania indicated the presence of a large cavity due to the absence of the G and H helices, which compromise the formation of the correct structure and function of the variant  $\alpha$ -chain. In particular, the inability of the Hb Campania  $\alpha$ -chain to interact with the AHSP chaperone causes strong instability (Figures 2 and S1A–F). The mutant  $\alpha$ -chains impairing the interactions with AHSP are usually characterized by a rapid degradation and for this reason often are not identified [15,16]. The Hb Campania has not been identified in the peripheral blood of patients, and this confirms that the variant chain Hb Campania is unstable and could be rapidly degraded and not involved in the formation of the hemoglobin tetramer.

The sequencing analysis of mRNA from reticulocytes of the carriers indicated that the frameshift variant showed a smaller peak than the one of the WT, and the semiquantitative analysis confirmed a reduction in the mutant mRNA that was about 34% of the total  $\alpha 1$ -globin cDNA (Figure 1D,E). This result adds another reason that explains the absence, in the peripheral blood, of the abnormal Hb Campania, which can only be synthesized at low quantities.

The characterization of the NMD pathway has defined many parameters of the decay process. In particular, in mammalian cells, a termination codon is usually “premature” if it is located at more than 50–54 nucleotides downstream from the final exon–exon junction [35,40,41]. The cod95 (–C) mutation generates a stop codon at a position +5 from the exon2–exon3 junction, which is a much smaller distance than the minimum of 50–54 nt required to activate the NMD, as was also confirmed by the SIFT analysis (Figure S3). This suggests that other mechanisms are involved in the reduction in this mRNA. The reduction in Hb Campania mRNA could be a consequence of the activation of one of the several distinct mechanisms that control the quality of mRNA and proteins during translation at the ribosome, which reduce the toxic effects of aberrant proteins causing many human diseases [39,40]. Analysis of the different mRNA control pathways showed that, in the case of Hb Campania, the no-go decay mechanism could be active. The no-go decay mechanism degrades mRNAs that are stalled in the translation elongation complexes as a result of either the specific features of the nascent peptides, strong secondary structures in mRNA physically blocking the translation machinery, or a rare codon repeat causing the codon site to be unoccupied for too long [33,34].

The in silico analysis excluded the presence of Hb Campania mRNA alterations that could induce the stall of ribosomes and the activation of mRNA quality control (Figure 3A–C) [32–34]. On the contrary, the in silico analyses of the amino acid composition (Figures 4 and S6–S8) and of the codon usage (Figures S9 and S10) in the Hb Campania mRNA highlighted several changes, in particular the introduction of two codons

(CGG and TCA) for which the corresponding tRNAs shown low frequencies in the red blood cells. These two rare tRNAs most likely induce a slowdown in the synthesis of *Hb Campania* mRNA, with consequent activation of the no-go decay.

#### 4.2. *Hb Sciacca* [ $\alpha 1$ cod109 (–C)]

In an epidemiological study on the molecular basis of  $\alpha$ -thalassemia in Southern Italy, in 5 unrelated families, we identified 10 carriers of the novel mutation  $\alpha 1$  cod109 (–C), and named it *Hb Sciacca* because it was first identified in a carrier from Sciacca village. The five families all originated from West Sicily.

This novel  $\alpha 1$  cod109 (–C) mutation results in a frameshift, causing an altered amino acid sequence in the next 22/23 amino acids, with the inclusion of 7 prolines, and a PTC at position 133 (Figure 1A). The  $\alpha$ -chain variant is expected to consist of 132 amino acids and end by the sequence  $\alpha 109$ WPPTSPSSPLRCTPPWTSSWLL (Figure 7).

The analysis of the *Hb Sciacca* superimposed 3D models revealed two anomalies: the presence of a bulky GH non-helix segment longer than the normal, and a shorter mutated H helix causing an empty region (Figures 6A–B and S1G–L). These modifications cause severe structural alterations that compromise important functions as the formation of the central cavity and of the heme pocket, the interaction with AHSP and the  $\beta$ -chain. All these data indicated that the severe structural distortion in the *Hb Sciacca* could impair the interactions with AHSP, inducing instability and a rapid degradation, and preventing the formation of the Hb tetramer. The absence of the *Hb Sciacca* in the patients' peripheral blood supports this hypothesis.

The mutation was also identified in a Kurdish family and in four unrelated patients of Egyptian origin and was associated, in all the cases, with a mild reduction in MCV (72 fL) and MCH (24.3 pg). The authors of these studies described a frameshift mutation in the  $\alpha 1$ -globin gene at codon 108 (ACC-AC) and assigned to the same mutation two HGVS names HBA1:c.327delC and HBA1:c.328delC [42,43]. This mutation was detected in families from West Sicily, located close to the sea, but also in African and Jewish populations. These findings suggest that the mutation entered this region via a Mediterranean population that settled in West Sicily, or that the mutation originated in Sicily and then spread to other territories as a result of migration. Based on the identification of the same haplotype in all families, we demonstrated unique origin of *Hb Sciacca* in Sicily. It would be interesting to analyze the same SNPs in African and Jewish carriers in order to determine whether the origin is the same between the two continents.

The codons 108 (ACC) and 109 (CTG) are characterized by a CCC strand between them. Given that it is not possible to define the exact position of the cytidine deletion, and that codon 108 remains invariant in any case (ACC), we assume that the deletion affects codon 109, where the frameshift begins; therefore, we have defined the mutation as  $\alpha 1$  cod109 (CTG>-TG) or HBA1: c.328delC [6–8].

To study in depth the pathophysiology of this variant and the effect of the frameshift on the phenotype, we studied the effects of the mutation on the mRNA. The sequencing revealed that the variant cDNA had smaller peaks compared with the WT one, and the semiquantitative analysis confirmed that the variant mRNA constituted only 14% of the total  $\alpha 1$ -globin cDNA (Figure 5D).

The *in silico* and the molecular analysis excluded the activation of an alternative splicing.

*Hb Sciacca* create a premature termination in the third exon of the  $\alpha$ -globin gene, caused by a frameshift. In the  $\alpha$ -globin gene, two other stop codon mutants were described; they lead to an early termination of translation due to a nonsense mutation—at cod116 and cod127 (Table 4) [44,45]. The abnormal hemoglobin created by the cod116 mutation was not found in peripheral blood, while a quantitative analysis of cod116 mRNA identified it at quantities equal to those of normal mRNA [44]. The mutant at cod127 was not studied at mRNA-level, and the variant chain was not identified [45].

These observations confirm that the quality control mechanism NMD is not activated by mutations creating premature termination codons in the third exon, which produce a stable messenger RNA (mRNA) that is available for direct synthesis of truncated polypeptides [4,39].

To investigate the possible causes that could activate the no-go decay and the reduction in mRNA, we analyzed in silico codon usage and mRNA structure (Figures 7 and 8). These in silico analyses identified two mechanisms that could stall translation: (a) a consistent alteration of *Hb Sciacca* mRNA structure, showing several lateral branches (Figure 8A–C); (b) a consistent change in the number and types of aa—as consequence of the frameshift—in particular, the introduction of two codons (CCG and ACG) rarely represented in the cell (Figure 7). These factors could cause stall of the ribosomes during translation inducing no-go decay [32–34].

#### 4.3. mRNA Variant in Globin Mutants

To determine whether quality control mechanisms other than the NMD are usually activated in frameshift mutants, we proceeded in two directions: (a) we analyzed a dominant  $\beta$ -thal gene with a frameshift in the third exon via an mRNA analysis; (b) we sought all the frameshift mutations in the third exon of the  $\alpha$ -globin gene, in order to collect data on mRNA that could be correlated with the hematological phenotype and with the mechanism of mRNA quality control (Table 4).

In  $\beta$ -globin genes—which are present as a single copy, contrarily to the duplicated  $\alpha$ -globin genes—the frameshift in the third exon produces a dominant phenotype, as was well-studied by Thein [4,46]. Two dominant mutants offer information regarding the anomalous  $\beta$ -globin mRNA. We reviewed the available data on the dominant mutated  $\beta$ -globin codons 128/129 (−4, −GCTG; +5, +CCACA) and codons 132–135 (−11, −AAAGTGGTGGC) (HGVS HBB: c.[385\_388delinsCCACA; 397\_407delAAAGTGTGGC]) that cause a frameshift reading through to codon 153, which should result in the synthesis of a variant  $\beta$ -globin that is seven residues longer than the WT. However, the author reports that it was not possible to isolate the variant  $\beta$ -chains, and that the total cytoplasmic RNA of the propositus showed a normal  $\alpha/\beta$ mRNA ratio of 1.4 [1]. Furthermore, mRNA analysis of the mild dominant  $\beta$ -codon 120/121 (+A), generating an abnormal  $\beta$ -mRNA with a new stop codon at codon 139 (TAA), showed an  $\alpha/\beta$  ratio of 5.18 (4.44 in the control), indicating that the variant  $\beta$ -globin mRNA was stable, and that neither NMD nor other mRNA quality control mechanisms were active for this mutant [47].

**Table 4.** Frameshift and stop codon mutations in the third exon of the  $\alpha$ -globin genes.

Name Current Name HGVS Name	Variant Chain	Stop at Cod	Var %	mRNA %	Genotype– Phenotype Relationship	Abnormal Helices	Ref.
$\alpha 2$ cod90-93 (−8bp); HBA2: c.272_279 delAGCTTCGG	Fs from cod91 and extension of the $\alpha$ -chain from 141 to 170 residues.	170	absent	nt	−SEA/− $\alpha$ CD90-93: HbH disease	FG3–HC3 + 29 aa	[48]
Hb Campania; $\alpha 1$ cod95 (G2) (−C); HBA1: c.286delC or HBA1: c.287delC	$\alpha(95)$ R-S-T-S-S-S- (102)COOH	102	absent	34	Heter: $\alpha$ -thal	G2–G8	p.a.
Hb Lynwood; $\alpha 2$ cod107 (G14) (−T); HBA2: c.323delT	$\alpha(107)$ G-P-W-P-P-T-S-P- P-S-S-P-L-R-C-T-P-P-W- T-S-S-W-L-L- (132)COOH	133	absent	nt	− $\alpha 3.7$ /Hb Lynwood: severe $\alpha$ -thal	G14–H15	[49]
Hb Sciacca; $\alpha 1$ cod109 (G16) (−C); HBA1: c.327delC or HBA1: c.328delC	$\alpha(109)$ W-P-P-T-S-P-P-S- S-P-L-R-C-T-P-P-W-T-S- S-W-L-L-(132)COOH	133	absent	15	Heter: $\alpha$ -thal	G16–H15	[2,8,9] p.a.

Table 4. Cont.

Name Current Name HGVS Name	Variant Chain	Stop at Cod	Var %	mRNA %	Genotype– Phenotype Relationship	Abnormal Helices	Ref.
$\alpha 1$ cod111-115 (–13bp); HBA1: c.333_345delCGCCACCTCCCC	$\alpha(110)$ A-P-S-S-P-L-R-C- T-P-P-L-R-C-T-P-P-W-T- S-P-W-T-S-S-W-L-L- (128)COOH	128(133)	absent	nt	Heter: $\alpha$ -thal	G17–H11	[43]
$\alpha 2$ cod114 (GH2) (–C); HBA2: c.342delC or HBA2: c.345delC	$\alpha(114)$ P-P-S-S-P-L-R-C- T-P-P-W-T-S-S-W-L-L- (132)COOH	133	absent	nt	– $\alpha 3.7 / \alpha 2$ cod 113/114–C: severe $\alpha$ -thal, 15% Hb Bart's	GH2–H15	[50]
$\alpha 2$ cod115 (GH3) (+CC); HBA2: c.343_344insCC or HBA2: c.342-345insCC or HBA2: c.344_345dup	$\alpha(114)$ P-P-P-S-S-P-L-R- C-T-P-P-W-T-S-S-W-L-L- (133)COOH	134(133)	absent	nt	Heter: dominant $\alpha$ -thal	GH3–H16	[51]
$\alpha 2$ cod116 (GH4) GAG>TAG Glu>Stop*; HBA2: c.349G>T	$\alpha 116$ (Stop)	116	absent	normal	Heter: $\alpha$ -thal	GH4	[44]
$\alpha 2$ cod116-119 (–11bp); HBA2: c.349_359delGAGTTCACCCC	Fs from cod 115 and extension of the $\alpha$ -chain from 141 to 166 residues.	166	absent	nt	– $\alpha 3.7 / \alpha 2$ cod116-119 del: Marked microcytosis, Hb H inclusion	GH3–HC3	[52]
$\alpha 2$ cod 127 (H10) GAG>TAG Lys>Stop*; HBA2: c.382A>T	$\alpha 127$ (stop)	127	absent	nt	Heter: $\alpha$ -thal	H10	[45]
Hb Hamilton Hill; $\alpha 2$ cod129 (H12) (–C); HBA2: c.388delC	$\alpha(129)$ W, L, L (132)COOH	133	absent	nt; 75 in vitro	Heter: $\alpha$ -thal	H12–H14	[53]
Hb Fez; $\alpha 1$ cod131 (H14) (–T); HBA1: c.396delT	$\alpha(131)$ S-(132)COOH	133	absent	nt	Heter: CHA Very unstable	H14	[54]
Hb Pak Num Po; $\alpha 1$ cod132 (+T); HBA1: c.396_397insT	$\alpha(132)$ C-E-H-R-A-D-L- Q-I-P-L-S-W-S-L-G-G- H-A-S-C-P-L-G-L-P-P- A-P-P-P-L-P-A-P-V-P-P- W-S-L-N-K-(175)V- COOH	175	absent	nt	Heter: $\alpha$ -thal;–SEA/Hb Pak Num Po: tran. dep. HbH	H15–HC3+ 34 aa	[55]
Hb Aalesund; $\alpha 2$ cod133-135 (–7bp); HBA2: c.400_406delAGCACCCG	$\alpha(133)$ C-(134)COOH	134(137)	~3 Hb A1c	nt	Heter: balanced hemolytic anemia	H16	[56]
Hb Senlis; $\alpha 1$ cod134 (H17) (–C); HBA1: c.404delC	$\alpha(134)$ T-C-(136)COOH	137	absent	nt	Heter: CHA, very unstable	H17–H18	[54]
Hb Wayne; $\alpha 2$ or $\alpha 1$ cod139 (HC1) (–A); HBA2: c.420delA(or HBA1)	$\alpha(139)$ N-T-V-K-L-E-P- (146)R-COOH	147	12-16	nt	Heter: normal, but absence of a Bohr effect, increased oxygen affinity	HC1–HC3+ 5 aa	[57]

CHA: chronic hemolytic anemia; nt: not tested; pa: present article; td HBH: transfusion-dependent HBH; \* = nonsense mutation; Heter: heterozygotes.

These observations indicate that the variations in codon usage and alterations in mRNA structure that are plausible in these complex  $\beta$ -globin mutants cannot always activate mRNA degradation through quality control mechanisms.

Regarding the mutations in the  $\alpha$ -globin genes, five variants were described in close proximity to the Hb Siacca, and they are characterized by a frameshift because of a deletion or insertion generating a stop codon at position 132, leading to almost the same 3' protein structure. They are Hb Lynwood ( $\alpha 2$  cod107 (–T) or HBA2: c.323delT) [49];  $\alpha 1$  cod111–115 (–13bp) or HBA1: c.333\_345delCGCCACCTCCCC [43];  $\alpha 2$  cod114 (–C) or HBA2: c.345delC [50];  $\alpha 2$  cod115 (+CC) or HBA2: c.343\_344insCC [51] (Table 4). All these variants result in a disturbed amino acid sequence between the frameshift codon and the premature



stop codon at position 133. The relevant helices G and H alter the aa sequence via the inclusion of several prolines (from 5 to 8), which are most likely very disruptive to the tertiary structure and contribute to the instability of the proteins. Additionally, in the case of Hb Sciacca, the aa sequence that is altered from position G16 through the following 22 amino acids (containing seven proline residues) could result in a very disruptive tertiary structure, altering the interactions with the alpha-hemoglobin stabilizing protein (AHSP) and the  $\beta$ -chain.

In the third exon, other unstable variants were described; specifically, four variants create a very long  $\alpha$ -chain:  $\alpha 2$  cod90-93 (−8bp) or HBA2: c.272\_279delAGCTTCGG (stop at codon 170) [48],  $\alpha 2$  cod116–119 (−11bp) (stop at codon 166) [52]; Hb Pak Num Po (stop at codon 175) [55]; Hb Wayne (stop at codon 147) [57]. Patients compound heterozygous for these variants and an  $\alpha 0$ -thal or  $\alpha +$ -thal mutation are characterized by more severe phenotypes, including transfusion dependence.

The last four mutations generate a stop codon close to the frameshift: Hb Hamilton Hill or  $\alpha 2$  cod129 (H12) (−C) (stop at cod133) [53]; Hb Fez or  $\alpha 1$  cod131 (H14) (−T) (stop at cod133) [54]; Hb Aalesund or  $\alpha 2$  cod133-135 (−7bp) (stop at cod137) [56]; and Hb Senlis or  $\alpha 1$  cod134 (H17) (−C) (stop at cod137) [54]. These variants, with the exception of Hb Hamilton Hill, present with chronic hemolytic anemia (CHA) or a balanced CHA. This may be due to the few chain variations in the H helix that still favor the interaction with the AHSP chaperone and with the  $\beta$ -chain, but which create severe instability with the consequent CHA.

Our search for data on mRNA analyses of all these  $\alpha$ -globin frameshift variants produced no meaningful information because, to the best of our knowledge, other authors did not carry out analyses on the mRNA from reticulocytes of the patients. The exception is Hb Hamilton Hill, which was cloned in expression vectors, causing a significant 25% reduction in the transcriptional activity [53].

This is the first report, to our knowledge, showing a reduction in  $\alpha$ -globin mRNA with frameshift mutations in the last exon, indicating that mechanisms other than NMD—probably no-go decay—could be involved in the quality control of the variant mRNAs.

It would be interesting to assess whether other frameshift mutants also display reductions in mRNA, or if those with a more severe phenotype have normal mRNA levels.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/biomedicines9101390/s1>, Figure S1: 3D model of WT, Hb Campania and Hb Sciacca  $\alpha$ -chains, Figure S2: Amino acids sequence and predicted secondary structure of the WT and mutated  $\alpha$ -chains, Figure S3: Analysis of the mutants HBA1 cod95 (−C) and HBA1 cod109 (−C) with the SIFT software, Figure S4: Analysis of the two mutants HBA1 cod95 (−C) and HBA1 cod109 (−C) with the MutationTaster software, Figure S5: Donor and acceptor splice site prediction of WT, Hb Campania and Hb Sciacca  $\alpha 1$ -globin mRNAs, Figure S6: Amino acid sequences of the variant chains HBA1 cod95 (−C) and HBA1 cod109 (−C) from the SIFT bioinformatic tool, Figure S7: Translation output of the normal and mutant HBA1 coding mRNA sequences to a protein sequence using the bioinformatic tool <https://web.expasy.org/translate/> (Accessed on 21 June 2021), Figure S8: Amino acids composition of the HBA1 WT, and of HBA1 cod95 (−C) and HBA1 cod109 (−C) by mean of the bioinformatic tool <https://web.expasy.org/protparam/> (Accessed on 22 June 2021), Figure S9: Codon usage of HBA1 WT and of HBA1 cod95 (−C) and HBA1 cod109 (−C), by means of the bioinformatic tool <http://genomes.urv.es/CAIcal/> (Accessed on 23 June 2021), Figure S10: Codon usage in Homo Sapiens (A), red blood cell (B), and related to the Hemoglobin (C) by means of the Codon Usage Database <https://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=9606> (Accessed on 16 June 2021), Figure S11: Restriction enzyme analysis of DNA amplicomers of the *Hb Campania* and *Hb Sciacca* genes.

**Author Contributions:** G.L. provided the experimental design, performed some of the experiments, performed the in silico analyses, analyzed all the experimental data, coordinated the entire study, wrote the original draft, and performed the manuscript review and editing. G.M. provided technical assistance and performed the experiments related to the genotyping and molecular characterization. G.C. performed the mRNA analysis from reticulocytes of patients, analyzed the experimental data

and supported in the mRNA in silico analyses. R.P. extrapolated the hematological data from databases for the phenotype analysis, gave bioinformatic support also for the in silico analyses, and prepared the Figures and Tables. All authors contributed to data gathering and interpretation and to the revision of the report. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by Ministero Istruzione, Università e Ricerca (MIUR), Legge 488/92, Cluster C02, Project 2. The section relative to the analysis of mRNA received no external funding.

**Institutional Review Board Statement:** A special committee of the Ministry for Research approved this study (Decreto n° 250, 22 June 1999) and two scientists were the supervisors. Ethical approval of the study protocol was obtained from the Comitato Etico Università Federico II of Naples (307/2016 and 225/2019).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank Maria Grazia Friscia (Azienda Ospedaliera Ospedali Civili Riuniti, Centro Trasfusionale e di Microcitemia, Sciacca, Italy) and Mercedes Caldora (P.O. Pellegrini A.S.L. Napoli1centro, Napoli, Italy) who sent us the blood or DNA samples of the patients. We would like to thank Mariarosaria Aletta (biblioteca CNR Roma) for helping us in the acquisition of papers.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Weatherall, D.J.; Clegg, J.B. *The Thalassemia Syndromes*, 4th ed.; Blackwell Scientific Publications: Oxford, UK, 2001; pp. 508–612. ISBN 978-0-470-69594-4.
- Lacerra, G.; Fiorito, M.; Musollino, G.; Di Noce, F.; Esposito, M.; Nigro, V.; Gaudio, C.; Carestia, C. Sequence variations of the  $\alpha$ -globin genes: Scanning of high CG content genes with DHPLC and DG-DGGE. *Hum. Mutat.* **2004**, *24*, 338–349. [[CrossRef](#)]
- Wajcman, H.; Traeger-Synodinos, J.; Papassotiropoulos, I.; Giordano, P.C.; Harteveld, C.L.; Baudin-Creuz, V.; Old, J. Unstable and Thalassaemic  $\alpha$  Chain Hemoglobin Variants: A Cause of Hb H Disease and Thalassemia Intermedia. *Hemoglobin* **2008**, *32*, 327–349. [[CrossRef](#)]
- Thein, S.L.; Hesketh, C.; Taylor, P.; Temperley, I.J.; Hutchinson, R.M.; Old, J.M.; Wood, W.G.; Clegg, J.B.; Weatherall, D.J. Molecular basis for dominantly inherited inclusion body beta-thalassemia. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 3924–3928. [[CrossRef](#)] [[PubMed](#)]
- Feng, L.; Gell, D.A.; Zhou, S.; Gu, L.; Kong, Y.; Li, J.; Hu, M.; Yan, N.; Lee, C.; Rich, A.M.; et al. Molecular Mechanism of AHSP-Mediated Stabilization of  $\alpha$ -Hemoglobin. *Cell* **2004**, *119*, 629–640. [[CrossRef](#)] [[PubMed](#)]
- Hall, G.; Thein, S. Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: A mechanism for the phenotype of dominant beta-thalassemia. *Blood* **1994**, *83*, 2031–2037. [[CrossRef](#)]
- Wernersson, R. Virtual Ribosome—A comprehensive DNA translation tool with support for integration of sequence feature annotation. *Nucleic Acids Res.* **2006**, *34*, W385–W388. [[CrossRef](#)] [[PubMed](#)]
- Lacerra, G.; Friscia, M.G.; Musollino, G.; Di Noce, F.; Virruso, L.; Ciaccio, C.; Di Girgenti, C.; Carestia, C. A new alpha1 mutation with a frameshift in the third exon. In Proceedings of the 9th International Conference on Thalassemia and the Haemoglobinopathies, Città del Mare, Terrasini (PA), Italy, 15–19 October 2003. Abstracts P031.
- Cardiero, G.; Prezioso, R.; Matarese, C.; Diana, E.; Dembech, S.; Virruso, L.; Friscia, M.G.; Di Girgenti, C.; Ciaccio, C.; Caldora, M.; et al. Six new alpha-thalassemia mutants identified in Southern Italy. In Proceedings of the Biochimica clinica, 50° Congresso Nazionale SIBioC — Riassunti Poster, Napoli, Italy, 16–18 October 2018; Volume 42, pp. S78–S79, P107, S51.
- Lacerra, G.; Testa, R.; De Angioletti, M.; Schilirò, G.; Carestia, C. Hb Bronte or  $\alpha 93(\text{FG}5)\text{Val} \rightarrow \text{Gly}$ : A New Unstable Variant of the  $\alpha 2$ -Globin Gene, Associated with a Mild  $\alpha +$ -Thalassemia Phenotype. *Hemoglobin* **2003**, *27*, 149–159. [[CrossRef](#)]
- Lacerra, G.; Musollino, G.; Di Noce, F.; Prezioso, R.; Carestia, C. Genotyping for known Mediterranean -thalassemia point mutations using a multiplex amplification refractory mutation system. *Haematol.* **2007**, *92*, 254–255. [[CrossRef](#)]
- Bongiovanni, A.; Colotti, G.; Liguori, G.L.; Di Carlo, M.; Digilio, F.A.; Lacerra, G.; Mascia, A.; Cirafici, A.M.; Barra, A.; Lanati, A.; et al. Applying Quality and Project Management methodologies in biomedical research laboratories: A public research network's case study. *Accreditation Qual. Assur.* **2015**, *20*, 203–213. [[CrossRef](#)]
- Digilio, F.A.; Lanati, A.; Bongiovanni, A.; Mascia, A.; Di Carlo, M.; Barra, A.; Cirafici, A.M.; Colotti, G.; Kisslinger, A.; Lacerra, G.; et al. Quality-based model for Life Sciences research guidelines. *Accreditation Qual. Assur.* **2016**, *21*, 221–230. [[CrossRef](#)]
- Cardiero, G.; Scarano, C.; Musollino, G.; Di Noce, F.; Prezioso, R.; Dembech, S.; La Porta, G.; Caldora, M.; Bisconte, M.G.; Bisogno, R.C.; et al. Role of nonsense-mediated decay and nonsense-associated altered splicing in the mRNA pattern of two new  $\alpha$ -thalassaemia mutants. *Int. J. Biochem. Cell Biol.* **2017**, *91*, 212–222. [[CrossRef](#)]

15. Bisconte, M.G.; Caldora, M.; Musollino, G.; Cardiero, G.; Flagiello, A.; La Porta, G.; Lagona, L.; Prezioso, R.; Qualtieri, G.; Gaudiano, C.; et al.  $\alpha$ -Thalassemia Associated with Hb Instability: A Tale of Two Features. The Case of Hb Rogliano or  $\alpha 1$  Cod 108(G15)Thr $\rightarrow$ Asn and Hb Policoro or  $\alpha 2$  Cod 124(H7)Ser $\rightarrow$ Pro. *PLoS ONE* **2015**, *10*, e0115738. [[CrossRef](#)]
16. Cardiero, G.; Musollino, G.; Friscia, M.G.; Testa, R.; Virruso, L.; Di Girgenti, C.; Caldora, M.; Bisogno, R.C.; Gaudiano, C.; Manco, G.; et al. Effect of Mutations on mRNA and Globin Stability: The Cases of Hb Bernalda/Groene Hart and Hb Southern Italy. *Genes* **2020**, *11*, 870. [[CrossRef](#)]
17. Lacerra, G.; Scarano, C.; Lagona, L.F.; Testa, R.; Caruso, D.G.; Medulla, E.; Friscia, M.G.; Mastrullo, L.; Caldora, M.; Prezioso, R.; et al. Genotype-Phenotype Relationship of the  $\delta$ -Thalassemia and Hb A2Variants: Observation of 52 Genotypes. *Hemoglobin* **2010**, *34*, 407–423. [[CrossRef](#)] [[PubMed](#)]
18. Buchan, D.W.A.; Jones, D.T. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res.* **2019**, *47*, W402–W407. [[CrossRef](#)] [[PubMed](#)]
19. Krieger, E.; Vriend, G. YASARA View—Molecular graphics for all devices—From smartphones to workstations. *Bioinform.* **2014**, *30*, 2981–2982. [[CrossRef](#)] [[PubMed](#)]
20. Guex, N.; Peitsch, M.; Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophor.* **2009**, *30*, S162–S173. [[CrossRef](#)] [[PubMed](#)]
21. Feng, L.; Zhou, S.; Gu, L.; Gell, D.A.; Mackay, J.P.; Weiss, M.J.; Gow, A.J.; Shi, Y. Structure of oxidized alpha-haemoglobin bound to AHSP reveals a protective mechanism for haem. *Nature* **2005**, *435*, 697–701. [[CrossRef](#)]
22. Fermi, G.; Perutz, M.F.; Shaanan, B.; Fourme, R. The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution. *J. Mol. Biol.* **1984**, *175*, 159–174. [[CrossRef](#)]
23. Sim, N.-L.; Kumar, P.; Hu, J.; Henikoff, S.; Schneider, G.; Ng, P.C. SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* **2012**, *40*, W452–W457. [[CrossRef](#)]
24. Schwarz, J.M.; Cooper, D.N.; Schuelke, M.; Seelow, D. MutationTaster2: Mutation prediction for the deep-sequencing age. *Nat. Methods* **2014**, *11*, 361–362. [[CrossRef](#)]
25. Reese, M.G.; Eeckman, F.H.; Kulp, D.; Haussler, D. Improved Splice Site Detection in Genie. *J. Comput. Biol.* **1997**, *4*, 311–323. [[CrossRef](#)]
26. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*; Springer: Totowa, NJ, USA, 2005. [[CrossRef](#)]
27. Puigbò, P.; Bravo, I.G.; Garcia-Vallve, S. CAIcal: A combined set of tools to assess codon usage adaptation. *Biol. Direct* **2008**, *3*, 38. [[CrossRef](#)]
28. Stothard, P. The Sequence Manipulation Suite: JavaScript Programs for Analyzing and Formatting Protein and DNA Sequences. *Biotech.* **2000**, *28*, 1102–1104. [[CrossRef](#)]
29. Gruber, A.; Lorenz, R.; Bernhart, S.H.F.; Neuböck, R.; Hofacker, I.L. The Vienna RNA Websuite. *Nucleic Acids Res.* **2008**, *36*, W70–W74. [[CrossRef](#)] [[PubMed](#)]
30. Sack, J.; Andrews, L.; Magnus, K.; Hanson, J.; Rubin, J.; Love, W. Location of Amino Acid Residues in Human Deoxy Hemoglobin. *Hemoglobin* **1978**, *2*, 153–169. [[CrossRef](#)] [[PubMed](#)]
31. Domingues-Hamdi, E.; Vasseur, C.; Fournier, J.-B.; Marden, M.C.; Wajcman, H.; Baudin-Creuzat, V. Role of  $\alpha$ -Globin H Helix in the Building of Tetrameric Human Hemoglobin: Interaction with  $\alpha$ -Hemoglobin Stabilizing Protein (AHSP) and Heme Molecule. *PLOS ONE* **2014**, *9*, e111395. [[CrossRef](#)] [[PubMed](#)]
32. Karamyshev, A.L.; Karamysheva, Z.N. Lost in Translation: Ribosome-Associated mRNA and Protein Quality Controls. *Front. Genet.* **2018**, *9*, 431. [[CrossRef](#)] [[PubMed](#)]
33. Powers, K.T.; Szeto, J.-Y.A.; Schaffitzel, C. New insights into no-go, non-stop and nonsense-mediated mRNA decay complexes. *Curr. Opin. Struct. Biol.* **2020**, *65*, 110–118. [[CrossRef](#)] [[PubMed](#)]
34. Mauro, V.P.; Chappell, S.A. A critical analysis of codon optimization in human therapeutics. *Trends Mol. Med.* **2014**, *20*, 604–613. [[CrossRef](#)]
35. Thermann, R.; Neu-Yilik, G.; Deters, A.; Frede, U.; Wehr, K.; Hagemeyer, C.; Hentze, M.; Kulozik, A.E. Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J.* **1998**, *17*, 3484–3494. [[CrossRef](#)]
36. Hamanaka, K.; Imagawa, E.; Koshimizu, E.; Miyatake, S.; Tohyama, J.; Yamagata, T.; Miyauchi, A.; Ekhilevitch, N.; Nakamura, F.; Kawashima, T.; et al. De Novo Truncating Variants in the Last Exon of SEMA6B Cause Progressive Myoclonic Epilepsy. *Am. J. Hum. Genet.* **2020**, *106*, 549–558. [[CrossRef](#)] [[PubMed](#)]
37. Isidor, B.; Lindenbaum, P.; Pichon, O.; Bézieau, S.; Dina, C.; Jacquemont, S.; Martin-Coignard, D.; Thauvin-Robinet, C.; Le Merrer, M.; Mandel, J.L.; et al. Truncating mutations in the last exon of NOTCH2 cause a rare skeletal disorder with osteoporosis. *Nat. Genet.* **2011**, *43*, 306–308. [[CrossRef](#)] [[PubMed](#)]
38. Salter, C.G.; Beijer, D.; Hardy, H.; Barwick, K.E.; Bower, M.; Mademan, I.; De Jonghe, P.; Deconinck, T.; Russell, M.A.; McEntagart, M.M.; et al. Truncating SLC5A7 mutations underlie a spectrum of dominant hereditary motor neuropathies. *Neurol. Genet.* **2018**, *4*, e222. [[CrossRef](#)]
39. Popp, M.W.; Maquat, L.E. Organizing Principles of Mammalian Nonsense-Mediated mRNA Decay. *Annu. Rev. Genet.* **2013**, *47*, 139–165. [[CrossRef](#)] [[PubMed](#)]
40. Zhang, J.; Sun, X.; Qian, Y.; Maquat, L. Intron function in the nonsense-mediated decay of  $\beta$ -globin mRNA: Indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm. *RNA* **1998**, *4*, 801–815. [[CrossRef](#)]

41. Pagani, F.; Baralle, F.E. Genomic variants in exons and introns: Identifying the splicing spoilers. *Nat. Rev. Genet.* **2004**, *5*, 389–396. [[CrossRef](#)]
42. Oron-Karni, V.; Filon, D.; Shifrin, Y.; Fried, E.; Pogrebjisky, G.; Oppenheim, A.; Rund, D. Diversity of  $\alpha$ -globin mutations and clinical presentation of  $\alpha$ -thalassemia in Israel. *Am. J. Hematol.* **2000**, *65*, 196–203. [[CrossRef](#)]
43. Henderson, S.J.; Timbs, A.T.; McCarthy, J.; Gallienne, A.E.; Proven, M.; Rugless, M.J.; Lopez, H.; Eglinton, J.; Dziedzic, D.; Beardsall, M.; et al. Ten Years of Routine  $\alpha$ - and  $\beta$ -Globin Gene Sequencing in UK Hemoglobinopathy Referrals Reveals 60 Novel Mutations. *Hemoglobin* **2015**, *40*, 75–84. [[CrossRef](#)]
44. A Liebhauer, S.; Coleman, M.B.; Adams, J.G.; E Cash, F.; Steinberg, M.H. Molecular basis for nondeletion alpha-thalassemia in American blacks. Alpha 2(116GAG—UAG). *J. Clin. Investig.* **1987**, *80*, 154–159. [[CrossRef](#)]
45. Préhu, C.; Pissard, S.; Al-Sheikh, M.; Le Niger, C.; Bachir, D.; Galactéros, F.; Wajcman, H. Two French Caucasian Families with Dominant Thalassemia-Like Phenotypes Due to Hyper Unstable Hemoglobin Variants: Hb Sainte Seve [Codon 118 (– T)] and Codon 127 [CAG→TAG (Gln→Stop)]. *Hemoglobin* **2005**, *29*, 229–233. [[CrossRef](#)]
46. Thein, S.L. Dominant beta thalassaemia: Molecular basis and pathophysiology. *Br. J. Haematol.* **1992**, *80*, 273–277. [[CrossRef](#)] [[PubMed](#)]
47. Hopmeier, P.; Krugluger, W.; Gu, L.; Smetanina, N.; Huisman, T. A newly discovered frameshift at codons 120-121 (+A) of the beta gene is not associated with a dominant form of beta-thalassemia [letter]. *Blood* **1996**, *87*, 5393–5394. [[CrossRef](#)] [[PubMed](#)]
48. Li, Y.; Liang, L.; Tian, M.; Qin, T.; Wu, X. Detection of Hb H disease caused by a novel mutation and –SEA deletion using capillary electrophoresis. *J. Clin. Lab. Anal.* **2019**, *33*, e22949. [[CrossRef](#)] [[PubMed](#)]
49. Finlayson, J.; Ghassemifar, R.; Holmes, P.; Grey, D.; Newbound, C.; Pell, N.; Jennens, M.; Macaulay, C.; Greenwood, L.; Beilby, J. Hb Lynwood [ $\alpha$ 107(G14) (–T) ( $\alpha$ 2) HBA2:c.323delT] in Conjunction with the  $\alpha$ 3.7 Deletion Produces a Moderately Severe  $\alpha$ -Thalassemia Phenotype. *Hemoglobin* **2011**, *35*, 142–146. [[CrossRef](#)] [[PubMed](#)]
50. Eng, B.; Patterson, M.; Walker, L.; Hoppe, C.; Azimi, M.; Lee, H.; Giordano, P.C.; Waye, J.S. Three New  $\alpha$ -Thalassemia Point Mutations Ascertained Through Newborn Screening. *Hemoglobin* **2006**, *30*, 149–153. [[CrossRef](#)]
51. Saller, E.; Dutly, F.; Frischknecht, H. Two Novel  $\alpha$ 2 Gene Mutations Causing Altered Amino Acid Sequences Produce a Mild (Hb Kinshasa, HBA2: c.428A > T) and Severe (HBA2: c.342-345insCC)  $\alpha$ -Thalassemia Phenotype. *Hemoglobin* **2015**, *39*, 144–146. [[CrossRef](#)]
52. Brennan, S.O.; Chan, T.; Duncan, J. Novel  $\alpha$ 2 Gene Deletion (c.349\_359 del GAGTTCACCCC) Identified in Association with the – $\alpha$ 3.7 Deletion. *Hemoglobin* **2011**, *36*, 93–97. [[CrossRef](#)]
53. Qadah, T.; Finlayson, J.; North, E.; Ghassemifar, R. Molecular Characterization of Hb Hamilton Hill (HBA2: c.388delC), a Novel HBA2 Variant Generating a Premature Termination Codon and Truncated HBA2 Chain. *Hemoglobin* **2015**, *39*, 88–94. [[CrossRef](#)]
54. Préhu, C.; Moradkhani, K.; Riou, J.; Bahau, M.; Launay, P.; Martin, N.; Wajcman, H.; Goossens, M.; Galactéros, F. Chronic hemolytic anemia due to novel  $\alpha$ -globin chain variants: Critical location of the mutation within the gene sequence for a dominant effect. *Haematol.* **2009**, *94*, 1624–1625. [[CrossRef](#)]
55. Viprakasit, V.; Tanphaichitr, V.S.; Veerakul, G.; Chinchang, W.; Petrarat, S.; Pung-Amritt, P.; Higgs, D.R. Co-inheritance of Hb Pak Num Po, a novel  $\alpha$ 1 gene mutation, and  $\alpha$ 0 thalassemia associated with transfusion-dependent Hb H disease. *Am. J. Hematol.* **2004**, *75*, 157–163. [[CrossRef](#)] [[PubMed](#)]
56. Grimholt, R.M.; Fjeld, B.; Selsås, H.; Schwettmann, L.; Klingenberg, O. Hb Aalesund (HBA2: c.400\_406delAGCACCG), an Unstable  $\alpha$ -Globin Variant Found in a Norwegian Patient Causing Moderate Hemolytic Anemia and Falsely High Hb A1c Using Ion Exchange High Performance Liquid Chromatography. *Hemoglobin* **2019**, *43*, 122–125. [[CrossRef](#)] [[PubMed](#)]
57. Huisman, T.H.J.; Headlee, M.G.; Wilson, J.B.; Lam, H.; Johnson, S.E.N.; Webber, B.B. HB Wayne, the Frameshift Variant with Extended  $\alpha$  Chains Observed in a Caucasian Family from Alabama. *Hemoglobin* **1984**, *8*, 1–15. [[CrossRef](#)] [[PubMed](#)]