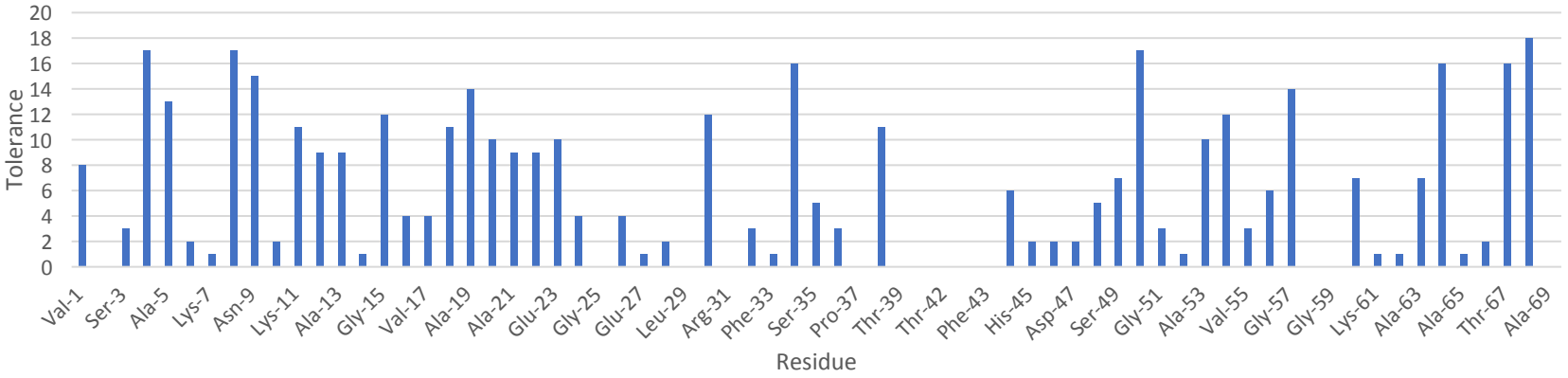
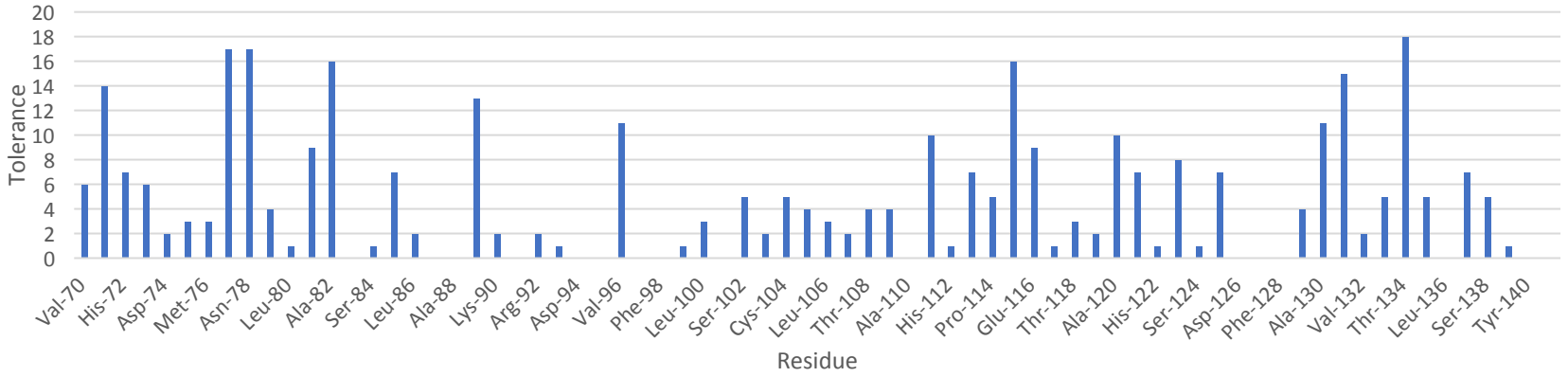

Tolerance
0
2
4
6
8
10
12
14
16
18
20
Val-1
Ser-3
Ala-5
Lys-7
Asn-9
Lys-11
Ala-13
Gly-15
Val-17
Ala-19
Ala-21
Glu-23
Gly-25
Glu-27
Leu-29
Arg-31
Phe-33
Ser-35
Pro-37
Thr-39
Thr-42
Phe-43
His-45
Asp-47
Ser-49
Gly-51
Ala-53
Val-55
Gly-57
Gly-59
Lys-61
Ala-63
Ala-65
Thr-67
Ala-69
Residue


Tolerance
0
2
4
6
8
10
12
14
16
18
20
Val-70
His-72
Asp-74
Met-76
Asn-78
Leu-80
Ala-82
Ser-84
Leu-86
Ala-88
Lys-90
Arg-92
Asp-94
Val-96
Phe-98
Leu-100
Ser-102
Cys-104
Leu-106
Thr-108
Ala-110
His-112
Pro-114
Glu-116
Thr-118
Ala-120
His-122
Ser-124
Asp-126
Phe-128
Ala-130
Val-132
Thr-134
Leu-136
Ser-138
Tyr-140
Residue

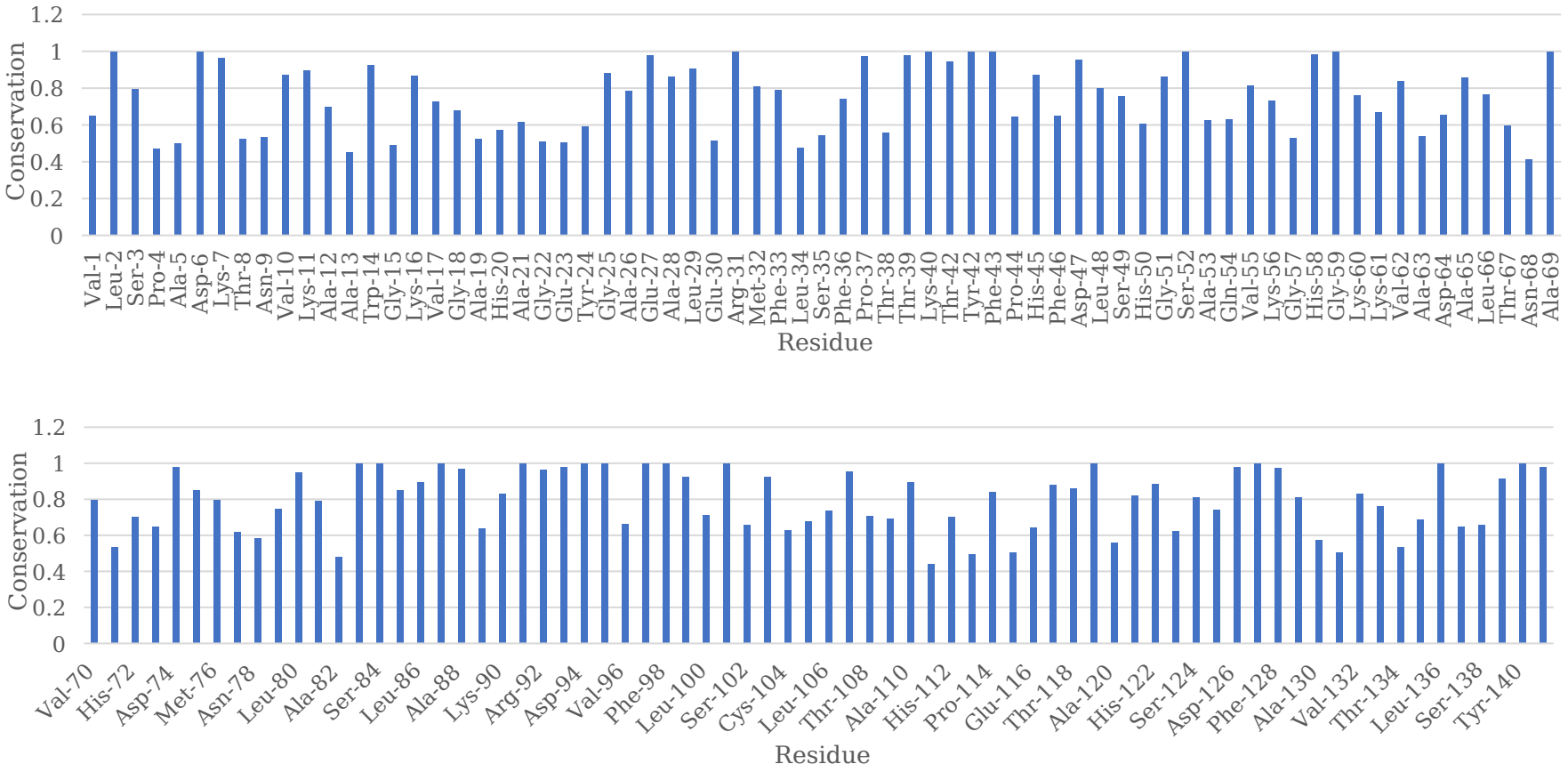
Residue	Gly	Ala	Val	Leu	Ile	Trp	Tyr	Phe	Cys	Met	Ser	Thr	Asp	Glu	Asn	Gln	Lys	Arg	His	Pro	Role
Val-1			WT			X	X	X	X				X	X	X	X	X	X		X	
Leu-2	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Ser-3	X		X	X	X	X	X	X	X	X	WT		X	X		X	X	X	X	X	
Pro-4						X	X													WT	
Ala-5		WT			X	X	X	X	X	X											
Asp-6	X	X	X	X	X	X	X	X	X	X	X	X	WT			X	X	X	X	X	
Lys-7	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT		X	X	
Thr-8						X						WT								X	
Asn-9						X	X	X							WT					X	
Val-10	X	X	WT			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Lys-11	X					X	X	X	X				X	X			WT			X	
Ala-12		WT	X		X	X	X	X	X	X								X	X	X	
Ala-13		WT				X	X						X	X	X	X	X	X	X	X	
Trp-14	X	X	X	X	X	WT	X		X	X	X	X	X	X	X	X	X	X	X	X	
Gly-15	WT			X	X	X	X	X	X	X											
Lys-16	X		X		X	X	X	X	X	X	X	X	X	X	X	X	WT			X	
Val-17	X		WT			X	X	X	X		X	X	X	X	X	X	X	X	X	X	
Gly-18	WT		X			X	X	X	X	X								X		X	
Ala-19		WT			X	X	X	X											X		
His-20	X				X	X	X	X	X	X									WT	X	
Ala-21		WT				X	X	X	X	X			X	X	X		X	X			
Gly-22	WT		X	X	X	X	X	X	X	X								X	X		
Glu-23	X					X	X	X	X	X				WT				X	X	X	
Tyr-24	X	X				X	WT		X	X	X	X	X	X	X	X	X	X	X	X	
Gly-25	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Ala-26		WT	X	X	X	X	X	X	X	X			X	X		X	X	X	X	X	
Glu-27	X	X	X	X	X	X	X	X	X	X	X	X		WT	X	X	X	X	X	X	
Ala-28		WT	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	
Leu-29	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	H
Glu-30				X		X	X				X			WT	X	X		X			B
Arg-31	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	X	X	B

Residue	Gly	Ala	Val	Leu	Ile	Trp	Tyr	Phe	Cys	Met	Ser	Thr	Asp	Glu	Asn	Gln	Lys	Arg	His	Pro	Role
Met-32	X	X			X	X	X	X	X	WT	X		X	X	X	X	X	X	X	X	
Phe-33	X	X	X		X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	
Leu-34				WT		X									X					X	B
Ser-35				X	X	X	X	X		X	WT		X	X	X	X	X	X	X	X	A/B
Phe-36	X	X	X	X	X	X		WT	X	X	X	X	X	X		X	X	X		X	A/B
Pro-37	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	A
Thr-38					X		X	X	X	X		WT	X					X	X		
Thr-39	X	X	X	X	X	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	X	
Lys-40	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	X	X	X	
Thr-42	X	X	X	X	X	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	X	
Tyr-42	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	H
Phe-43	X	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	H
Pro-44			X		X	X	X	X	X	X			X	X	X		X	X	X	WT	
His-45	X	X	X	X	X	X	X	X	X	X		X	X	X	X		X	X	WT	X	H
Phe-46	X	X	X		X		X	WT	X	X	X	X	X	X	X	X	X	X	X	X	H
Asp-47	X	X	X	X	X	X	X	X	X	X	X	X	WT			X	X	X	X	X	
Leu-48	X			WT		X	X		X		X	X	X	X	X	X	X	X	X	X	
Ser-49			X	X	X	X	X	X	X	X	WT		X	X				X		X	
His-50						X							X						WT		
Gly-51	WT	X	X	X	X	X	X	X	X	X		X		X		X	X	X	X	X	
Ser-52	X	X	X	X	X	X	X	X	X	X	WT	X	X	X		X	X	X	X	X	
Ala-53		WT	X	X	X	X	X	X	X	X								X	X		
Gln-54	X		X	X	X	X		X	X							WT					
Val-55	X		WT			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Lys-56	X		X		X	X	X	X	X	X	X	X	X	X	X		WT			X	
Gly-57	WT					X	X	X	X											X	
His-58	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	X	H
Gly-59	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Lys-60			X	X	X	X	X	X	X	X		X	X				WT		X	X	
Lys-61	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT		X	X	
Val-62	X	X	WT	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	H

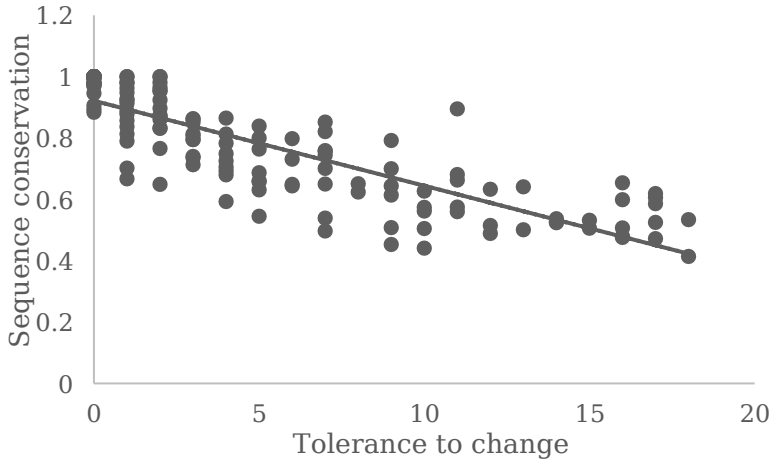
Residue	Gly	Ala	Val	Leu	Ile	Trp	Tyr	Phe	Cys	Met	Ser	Thr	Asp	Glu	Asn	Gln	Lys	Arg	His	Pro	Role
Ala-63		WT				X	X	X	X				X	X	X	X	X	X	X	X	
Asp-64						X	X						WT							X	
Ala-65		WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Leu-66	X	X		WT		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Thr-67						X						WT						X		X	
Asn-68						X									WT						
Ala-69	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Val-70	X		WT			X	X	X		X			X	X	X	X	X	X	X	X	
Ala-71		WT			X	X		X	X											X	
His-72	X	X	X	X	X	X		X	X	X		X		X					WT	X	
Val-73	X		WT			X	X		X		X		X	X	X	X	X	X	X	X	
Asp-74	X	X	X	X	X	X	X	X	X	X	X	X	WT			X	X	X	X	X	
Asp-75	X	X	X	X	X	X	X	X	X	X	X	X	WT			X	X	X		X	
Met-76	X	X				X	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	
Pro-77						X			X											WT	
Asn-78						X		X							WT						
Ala-79		WT		X	X	X	X	X	X	X			X	X	X	X	X	X	X	X	
Leu-80	X	X	X	WT	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	
Ser-81			X		X	X	X				WT		X	X		X		X	X	X	
Ala-82		WT			X	X		X													
Leu-83	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	H
Ser-84		X	X	X	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	X	X	
Asp-85			X	X	X	X	X	X	X	X	X		WT				X			X	
Leu-86	X	X		WT	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	H
His-87	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	X	H
Ala-88	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
His-89	X		X		X	X				X									WT	X	
Lys-90	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	WT		X	X	
Leu-91	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	H
Arg-92	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		WT		X	
Val-93	X	X	WT	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	H

Residue	Gly	Ala	Val	Leu	Ile	Trp	Tyr	Phe	Cys	Met	Ser	Thr	Asp	Glu	Asn	Gln	Lys	Arg	His	Pro	Role
Asp-94	X	X	X	X	X	X	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	
Pro-95	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	A
Val-96			WT			X		X		X			X	X			X	X		X	A
Asn-97	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	X	X	X	X	X	
Phe-98	X	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	H
Lys-99	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT		X	X	A/B
Leu-100	X	X		WT		X	X		X	X	X	X	X	X	X	X	X	X	X	X	A/B/H
Leu-101	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Ser-102			X	X	X	X	X	X		X	WT		X	X		X	X	X	X	X	
His-103	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X	X	WT	X	A/B
Cys-104	X			X	X	X	X	X	WT	X			X	X		X	X	X	X	X	
Leu-105	X	X		WT		X	X		X		X	X	X	X	X	X	X	X	X	X	H
Leu-106	X	X		WT		X	X	X	X		X	X	X	X	X	X	X	X	X	X	A/B
Val-107	X	X	WT			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	A/B
Thr-108	X			X	X	X	X	X		X		WT	X	X	X	X	X	X	X	X	
Leu-109	X	X		WT		X	X		X		X	X	X	X	X	X	X	X	X	X	
Ala-110	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	B
Ala-111		WT				X	X	X					X	X			X	X	X	X	A/B
His-112	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	WT	X	
Leu-113	X			WT		X			X		X	X	X	X	X	X	X	X		X	
Pro-114			X	X	X	X	X	X	X	X		X	X	X	X			X	X	WT	B
Ala-115		WT				X	X	X													
Glu-116			X			X	X	X	X	X		X		WT			X	X		X	
Phe-117	X	X	X		X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	A/B
Thr-118	X		X	X	X	X	X	X	X	X		WT	X	X		X	X	X	X	X	A
Pro-119	X			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	A/B
Ala-120		WT	X		X	X	X	X	X	X									X	X	A/B
Val-121			WT			X	X	X			X		X	X	X	X	X	X	X	X	
His-122	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	WT	X	A/B
Ala-123	X	WT				X	X	X	X				X	X	X	X			X	X	A/B
Ser-124	X		X	X	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	X	X	

Residue	Gly	Ala	Val	Leu	Ile	Trp	Tyr	Phe	Cys	Met	Ser	Thr	Asp	Glu	Asn	Gln	Lys	Arg	His	Pro	Role
Leu-125	X	X		WT					X		X		X	X	X	X	X	X	X	X	
Asp-126	X	X	X	X	X	X	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	A/B
Lys-127	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	X	X	X	A/B
Phe-128	X	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	
Leu-129	X	X		WT		X	X		X		X	X	X	X	X	X	X	X	X	X	H
Ala-130		WT		X		X	X	X					X				X	X		X	
Ser-131						X		X			WT							X		X	
Val-132	X	X	WT			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	H
Ser-133				X	X	X	X	X		X	WT		X	X	X	X	X	X	X	X	
Thr-134												WT								X	A
Val-135	X		WT			X	X	X	X	X			X	X	X	X	X	X	X	X	
Leu-136	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	H
Thr-137				X		X	X	X			X	WT	X	X	X		X	X	X	X	A
Ser-138			X	X	X	X	X	X	X	X	WT	X					X	X	X	X	
Lys-139	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT		X	X	
Tyr-140	X	X	X	X	X	X	WT	X	X	X	X	X	X	X	X	X	X	X	X	X	
Arg-141	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	WT	X	X	


Conservation
1.2
1
0.8
0.6
0.4
0.2
0
Val-1 Leu-2 Ser-3 Pro-4 Ala-5 Asp-6 Lys-7 Thr-8 Asn-9 Val-10 Lys-11 Ala-12 Ala-13 Trp-14 Gly-15 Lys-16 Val-17 Gly-18 Ala-19 His-20 Ala-21 Gly-22 Glu-23 Tyr-24 Gly-25 Ala-26 Glu-27 Ala-28 Leu-29 Glu-30 Arg-31 Met-32 Phe-33 Leu-34 Ser-35 Phe-36 Pro-37 Thr-38 Thr-39 Lys-40 Thr-42 Tyr-42 Phe-43 Pro-44 His-45 Phe-46 Asp-47 Leu-48 Ser-49 His-50 Gly-51 Ser-52 Ala-53 Gln-54 Val-55 Lys-56 Gly-57 His-58 Gly-59 Lys-60 Lys-61 Val-62 Ala-63 Asp-64 Ala-65 Leu-66 Thr-67 Asn-68 Ala-69
Residue

Conservation
1.2
1
0.8
0.6
0.4
0.2
0
Val-70 His-72 Asp-74 Met-76 Asn-78 Leu-80 Ala-82 Ser-84 Leu-86 Ala-88 Lys-90 Arg-92 Asp-94 Val-96 Phe-98 Leu-100 Ser-102 Cys-104 Leu-106 Thr-108 Ala-110 His-112 Pro-114 Glu-116 Thr-118 Ala-120 His-122 Ser-124 Asp-126 Phe-128 Ala-130 Val-132 Thr-134 Leu-136 Ser-138 Tyr-140
Residue


Sequence conservation
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
Tolerance to change

In silico analysis of the effects of point mutations on α-globin: implications for α-thalassemia

Agathe Horri-Naceur and David J. Timson*

School of Pharmacy and Biomolecular Sciences, University of Brighton, Huxley Building, Lewes Road, Brighton BN2 4GJ. UK.

* Corresponding author:

Address: School of Pharmacy and Biomolecular Sciences, University of Brighton, Huxley Building, Lewes Road, Brighton BN2 4GJ. UK.

Email d.timson@brighton.ac.uk

Running title: Effects of α-globin mutations

Abstract

Hemoglobinopathies are inherited diseases that impair the structure and function of the oxygen-carrying pigment hemoglobin. Adult hemoglobin consists of two α and two β subunits. α-Thalassaemia affects the genes that code for the α-globin chains, *HBA1* and *HBA2*. Mutations can result in asymptomatic, mild or severe outcomes depending on several factors, such as mutation type, number of mutations and the location at which they occur. PredictSNP was used to estimate whether every possible single nucleotide polymorphism (SNP) would have a neutral or deleterious effect on the protein. These results were then used to create a plot of predicted tolerance to change for each residue in the protein. Tolerance to change was negatively correlated with the residue's sequence conservation score. The PredictSNP data were compared to clinical reports of 110 selected variants in literature. There were 29 disagreements between the two data types. Some of these could be resolved by considering the role of the affected residue in binding other molecules. The three-dimensional structures of some of these variant proteins were modelled. These models helped explain variants which affect heme binding. We predict that where a point mutation alters a residue that is intolerant to change, is well conserved and or involved in interactions, it is likely to be associated with disease. Overall, the data from this study could be used alongside biochemical and clinical data to assess novel α-globin variants.

Keywords: Hemoglobin; thalassemia; protein stability; heme group; SNPs; in silico prediction

Introduction

α-Thalassaemia (OMIM #604131) is an autosomal recessive genetic disorder that affects hemoglobin. This tetrameric, iron-containing protein, participates in gas exchange between the blood and tissues [1]. Human adult hemoglobin consists of two α-globin and two β-globin polypeptides. Unusually, α-globin is coded for by two genes, *HBA1* and *HBA2*. The gene sequences are highly similar and encode identical proteins. The *HBA2* gene produces two-three times more mRNA and protein than the *HBA1* gene in healthy humans [2, 3]. A range of different mutations can occur in *HBA1* and *HBA2* and are associated with α-thalassemia. The type and number of mutations in one or both genes affect the severity of the structural changes and synthesis of the resulting protein [4]. As a result, α-thalassaemia symptoms can vary from mild or undetectable to severe and even fatal anaemia [5]. While deletions in the *HBA1* and/or *HBA2* genes are the most commonly detected mutations, many Single Nucleotide Polymorphisms (SNPs) and frameshift mutations have been reported to cause some rarer forms of α-thalassaemia. Mutations affecting only one allele can result in an asymptomatic phenotype or silent carrier. Two mutated alleles can cause α-thalassaemia; where the patient may suffer from anaemia and fatigue and may require iron supplementation and blood transfusions. Sufferers with three inactive alleles due to mutations have Hemoglobin H (HbH) disease and are likely to require more radical treatments such as hematopoietic stem cell (bone marrow) transplants or a splenectomy [6]. In some cases of HbH disease the variant hemoglobin is dysfunctional and forms inclusion bodies in red blood cells [7]. In cases where all four alleles are affected by mutations, the resulting condition is homozygous α-thalassaemia that can be fatal *in utero* due to hydrops fetalis (foetal swelling of the liver and other organs). This condition is the most severe hemoglobinopathy, with an infant mortality rate of up to 40% [8]. Developmental delays are observed in approximately 70% of children surviving past the age of one [9].

Reduced stability has been shown to be critical in many genetic diseases that result in changes to protein sequences [10]. We hypothesised that the same is likely be true in α-thalassemia. The potential importance of protein instability in hemoglobin is reinforced by the existence of the α-hemoglobin stabilising protein (AHSP). This chaperone protein binds α-globin *in vivo*, stabilising it and preventing aggregation [11, 12]. Here, we report predictions of the effect of SNPs in the α-globin genes on the stability of the resulting variant protein. These were compared to selected mutations for which the clinical consequences are known. The implications and limitations of these results are discussed.

Methods

PredictSNP analysis

PredictSNP (https://loschmidt.chemi.muni.cz/predictsnp1/) was used with the HBA1/HBA2 protein sequence to estimate the effect of amino acid substitution at each of the 141 residues. (Note: amino acid numbering begins with Val-1, since the N-terminal methionine residue is post-translationally removed. Variants are named according to HbVar (http://globin.bx.psu.edu/cgi-bin/hbvar/query_vars3), i.e. αNXxx>Yyy, where N is the residue number, Xxx is the residue in the wild-type and Yyy is the residue in the variant [13, 14]. In Table 2, these are also presented in the format commonly used in the inherited disease literature, i.e. p.XxxNYyy. By convention, this format numbers the initial methionine as 1, regardless of processing. For example, α1Val>Glu is p.V2E.) PredictSNP predicts whether a SNP has a neutral or deleterious effect [15]. It uses six algorithms to make these predictions and reach a consensus by majority: MAPP, PhD-SNP, PolyPhen-1, PolyPhen-2, SIFT and SNAP. This prediction was used to determine the number of amino acid substitutions that were tolerated (i.e. predicted a neutral outcome) at each position.

Protein sequence analysis

Sixty vertebrate α-globin protein sequences were identified using NCBI. Only reference sequences (RefSeq) and confirmed (not predicted) sequences were used (Table 1). These were aligned using ClustalW (as implemented at GenomeNet: https://www.genome.jp/tools-bin/clustalw using the slow/accurate settings) [16]. The alignment was saved in FASTA format and used to generate a sequence conservation score for each residue in Scorecons (https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/valdar/scorecons_server.pl) using the Valdar01 scoring method [17].

Linear regression was carried out in Excel (Microsoft, USA). Linear regression parameters were calculated using the LINEST function, which returns the correlation coefficient (r^2) and the F value. The significance of this value was determined using the FDIST function.

Literature search for clinically relevant mutants and their phenotypes

The NCBI genome viewer (https://www.ncbi.nlm.nih.gov/potein/NP_000508.1?report=graph&v=1%3A51&content=5&m=1!&

mn=rs121909803&dispmax=1&currpage=1) was used to identify clinically reported examples of SNPs occurring in either the *HBA1* or *HBA2* genes. The amino acid position, substitution, name of the hemoglobin variant and associated symptoms were recorded for the length of the gene(s). With each variant, the asymptomatic or symptomatic outcome was noted in literature and this was compared to the PredictSNP outcome for the whole protein(s).

Generation, validation and review of α-globin protein models

In cases where the predicted outcome and literature disagreed, models were created in SWISS MODEL (https://swissmodel.expasy.org/interactive) [18]. A model was also built of the wild-type protein for comparison. PROCHECK (https://servicesn.mbi.ucla.edu/PROCHECK/) was used for each model to generate Ramachandran plots using the pdb file from SWISS MODEL [19]. The .pdb files were also viewed in CCP4 MG (http://www.ccp4.ac.uk/MG/) to allow direct comparison and sequence labelling [20]. Residues involved in interactions with other molecules (heme, β-globin and AHSP) were identified from the literature and using computational tools with the structure files – oxygenated human hemoglobin (PDB: 2DN1) and the α-globin/AHSP complex (PDB: 1Y01) [12, 21-23]. PLIP (https://projects.biotec.tu-dresden.de/plip-web/plip/index) was used to identify heme interacting residues [24]. SPPIDER (http://sppider.cchmc.org/) was used to identify residues in the α-globin/β-globin and α-globin/AHSP interfaces [25].

Results and Discussion

Predicted tolerance for residue substitution is negatively correlated with α-globin sequence conservation

To estimate the tolerance to amino acid substitution at each position in the protein, predictions of the effect (neutral/deleterious) of the 19 other amino acids were calculated at each residue using PredictSNP (Figure 1, 2). Some residues, such as Leu-2 show no tolerance to alteration, meaning only the wild type residue is predicted to result in functional protein. Other residues, such as Asn-68, are predicted to tolerate 18 of the 19 other possible amino acid substitutions (Figure 2). We postulated that those residues with low tolerances represent those that are more critical for folding or protein stability. Sequence conservation of α-globin proteins was estimated based on 60

vertebrate sequences (Figure 3). There was a negative correlation between the degree of conservation and tolerance to amino acid substitution (r^2=-0.7065; F=334.6; p=8.0×10^{-39}; Figure 4).

Tolerance for residue substation is a good predictor of clinical phenotypes

The PredictSNP results were compared to selected clinical reports of SNPs (Table 2). In most cases, the reported symptoms (or lack thereof) agreed with the PredictSNP result. In some cases, biochemical studies demonstrated defects in the protein (e.g. reduced stability or altered ligand binding), but there are no reports of associated clinical manifestations. This suggests that some loss of function in α-globin can be tolerated. There is also the possibility that some of these variants may result in pathology later in the patients' lives or in patients exposed to different environmental factors.

Structural factors explain why some predictions failed

However, 29 predictions did not match the clinical phenotype. In six of these cases, biochemical abnormalities of the protein (e.g. reduced stability or altered oxygen affinity) have been described. Presumably, these abnormalities are insufficient to cause disease, at least under the environmental conditions experienced by the patients. In four cases where an incorrect neutral prediction was made, the residue was involved in interactions with other molecules. These were further investigated by generating three dimensional models of the structures in SWISS MODEL and comparing these to the wild type (PDB: 2DN1 and a model generated using the α-globin sequence [21]). These models generally predicted only minor changes to α-helical regions of the proteins, and no changes to other regions. Three variants were predicted to have changes in protein structure resulting in the loss of the heme cofactor – α46Phe>Val (Hb Hillingdon; initially wrongly predicted to be asymptomatic), α65Ala>Thr (Hb Part-Dieu; clinical reports of hepatosplenomegaly, albeit confused by the patient's diabetes) and α136Leu>Met (Hb Chicago) [26-28]. In other variants (e.g.Hb Westmead; α122His>Gln) the amino acid substitutions occur at an α1β1 contact site within the hemoglobin tetramer, but are predicted to be neutral by PredictSNP [29]. Here, the most likely biochemical cause is disruption of the quaternary structure of hemoglobin, resulting from the change at the contact site. In other cases, variants were described as deleterious in literature but predicted to be neutral in our PredictSNP analyses. These variants were: Hb Hekinan (causes

anaemia), Hb O Padova (dyserythropoietic anaemia) and Hb Queens (mild anisocytosis and anaemia) [30-32]. These discrepancies remain unresolved.

Limitations of this study

The use of *in silico* tools to predict effects on protein stability is well-established in the study of genetic diseases [33]. However, there are some limitations – both general and specific to this study. The biophysical bases of protein folding and stability are improperly understood. Therefore, any *in silico* tool will be imperfect in its ability to generate correct predictions in all cases. It is established that considering the outcomes of several different tools (as here since PredictSNP uses six different tools) results in more accurate predictions [34, 35]. Structural predictions based on homology models are unlikely to predict major structural changes correctly. The tools also consider the protein in isolation, i.e. without taking account of any interactions with other biomacromolecules. This limitation is particularly relevant here since the predictions cannot take account of interactions with heme, AHSP or β-globin. Nor can the tools account for other biochemical consequences of residue changes. For example, in α-globin, alterations to Val-1 can result in retention of the N-terminal methionine residue [36, 37]. In Hb Thionville (α1Val>Glu) this causes structural changes which reduce the affinity for oxygen and decrease the effectiveness of allosteric regulators [36]. Thus, these predictions should not be used alone to guide clinical decision making on newly discovered mutations [38]. They should be used alongside clinical and biochemical data on the properties of new variant hemoglobins.

Conclusions

In general, residues that were intolerant of alteration and are highly conserved, were likely to be associated with more severe phenotypes. This is consistent with similar studies on other proteins that show that sequence changes in regions of high sequence conservation tend to be associated with more severe disease phenotypes (e.g. [39-41]). Severe forms of α-thalassemia are predicted where the affected residue is intolerant of change. If the predicted outcome is neutral, the potential for the change to result in loss of the heme group or alterations to the protein-protein interfaces with β-globin and AHSP should be assessed by analysis of the α-globin crystal structure or protein modelling. Such changes are also likely to predict that the mutation is associated with disease. This work provides further evidence in support of the hypothesis that altered protein stability is a key

element in the molecular pathology in cases of α-thalassemia resulting from point mutations that alter the coding sequence.

Acknowledgements

There was no specific funding supporting this project. The authors thank the scientists who wrote and freely provided the software tools used in this project.

Conflict of Interest Statement

The authors confirm that they have no conflicts of interest relating to this work.

Tables

Table 1:

Hemoglobin α-subunit protein sequences used in conservation analysis.

Accession number and species	Sequence
NP_000508.1 *Homo sapiens*	MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKF LASVSTVLTSKYR
NP_001070890.2 *Bos Taurus*	MVLSAADKGNVKAAWGKVGGHAAEYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKGHGAKVAAALTKAVEHLDDLPGALSELSD LHAHKLRVDPVNFKLLSHSLLVTLASHLPSDFTPAVHASLDKF LANVSTVLTSKYR
NP_032244.2 *Mus musculus*	MVLSGEDKSNIKAAWGKIGGHGAEYGAEALERMFASFPTTKTY FPHFDVSHGSAQVKGHGKKVADALANAAGHLDDLPGALSALSD LHAHKLRVDPVNFKLLSHCLLVTLASHHPADFTPAVHASLDKF LASVSTVLTSKYR
NP_001013875.1 *Rattus norvegicus*	MVLSEEDKNNIKKAWVKIGNHAAEIGAETIGRLFIVFPSSKTY FPHFNTSEGSDQVKAHGKKVADALTNAASHLDDLPGALSTLSD LHAHKLRVDPVNFKFLSHCLLVTLASHHPGDFTPAMHASLDKF FASVSTVLTSKYR
NP_001078901.1 *Equus caballus*	MVLSAADKTNVKAAWSKVGGHAGEFGAEALERMFLGFPTTKTY FPHFDLSHGSAQVKAHGKKVGDALTLAVGHLDDLPGALSNLSD LHAHKLRVDPVNFKLLSHCLLSTLAVHLPNDFTPAVHASLDKF LSSVSTVLTSKYR
NP_001036092.1 *Pan troglodytes*	MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKF LASVSTVLTSKYR
NP_001038189.1 *Macaca mulatta*	MVLSPADKSNVKAAWGKVGGHAGEYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKGHGKKVADALTLAVGHVDDMPHALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKF LASVSTVLTSKYR
NP_001162287.1 *Papio Anubis*	MVLSPDDKKHVKAAWGKVGEHAGEYGAEALERMFLSFPTTKTY FPHFDLSHGSDQVNKHGKKVADALTLAVGHVDDMPQALSKLSD LHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKF LASVSTVLTSKYR
NP_001028158.1 *Monodelphis domestica*	MVLSAADKTNVKAAWSKVGGNSGAYMGEALYRTFLSFPPTKTY FPHFEFSAGSAQIKGQGQKIADAVSLAVAHMDDLATALSALSD LHAHNLKVDPVNFKFLCHNVLVTLASHLGKDFTPEIHASLDKF LALLSTVLTSKYR
NP_001125901.1 *Pongo abelii*	MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKF LASVSTVLTSKYR
NP_988860.1 *Xenopus tropicalis*	MHLTADDKKHIKAIWPSVAAHGDKYGGEALHRMFMCAPKTKTY FPDFDFSEHSKHILAHGKKVSDALNEACNHLDNIAGCLSKLSD LHAYDLRVDPGNFPLLAHQILVVVAIHFPKQFDPATHKALDKF LVSVSNVLTSKYR
NP_001289417.1 *Cuculus canorus*	MVLSAADKTNVKGIFTKIGGHGDDYGAETLDRMFTVYPQTKTY FPHFDVSHGSAQIKAHGKKVVAALVEAVNHIDDIAGALSKLSD LHAHKLRVDPANFKLLGQCFLVVVGIHHPAALTPEVHASLDKF LCAVSTVLTAKYR
NP_001117134.1 *Salmo salar*	MSLTARDKSVVNAFWGKIKGKADVVGAEALGRMLTAYPQTKTY FSHWADLSPGSAPVKKHGGVIMGAIGNAVGLMDDLVGGMSGLS

	DLHAFKLRVDPGNFKILSHNILVTLAIHFPADFTPEVHIAVDK FLAALSAALADKYR
NP_001299611.1 *Ictidomys tridecemlineatus*	MVLSPADKNNVKACWEKIGGHGAAYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVQGHGKKVADALANAAAHVDDLPGALSTLSD LHAHKLRVDPVNFKLLSHCLLVTLAAHHPAEFTPAVHASLDKF LASVSTVLTSKYC
NP_001188201.1 *Ictalurus punctatus*	MSLSAKDKAVVKDLWAKVAPKADDIGAEALGRLFEVYPQTKTY FSHWSDLTPGSAQVKKHGSVIVRKIGEAVGHIDDLTGALSSLS ELHAFKLRVDPVNFKLLSHTIEVSIALFFPAEFTPEVHVSFDK FLQNLALALSEKYR
NP_001253705.1 *Macaca mulatta*	MVLSPADKSNVKAAWGKVGGHAGEYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKGHGKKVADALTLAVGHVDDMPQALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKF LASVSTVLTSKYR
NP_571332.3 *Danio rerio*	MSLSDTDKAVVKAIWAKISPKADEIGAEALARMLTVYPQTKTY FSHWADLSPGSGPVKKHGKTIMGAVGEAISKIDDLVGGLAALS ELHAFKLRVDPANFKILSHNVIVVIAMLFPADFTPEVHVSVDK FFNNLALALSEKYR
NP_001118023.1 *Oncorhynchus mykiss*	MSLTAKDKSVVKAFWGKISGKADVVGAEALGRMLTAYPQTKTY FSHWADLSPGSGPVKKHGGIIMGAIGKAVGLMDDLVGGMSALS DLHAFNLRVDPGNFKILSHNILVTLAIHFPSDFTPEVHIAVDK FLAAVSAALADKYR
NP_001267813.1 *Loxodonta Africana*	MVLSDNDKTNVKATWSKVGDHASDYVAEALERMFFSFPTTKTY FPHFDLGHGSGQVKAHGKKVGEALTQAVGHLDDLPSALSALSD LHAHKLRVDPVNFKLLSHCLLVTLSSHQPTEFTPEVHASLDKF LSNVSTVLTSKYR
XP_010989389.1 *Camelus dromedaries*	MVLSSKDKTNVKTAFGKIGGHAAEYGAEALERMFLGFPTTKTY FPHFDLSHGSAQVKAHGKKVGDALTKAADHLDDLPSALSALSD LHAHKLRVDPVNFKLLSHCLLVTVAAHHPGDFTPSVHASLDKF LANVSTVLTSKYR
XP_010380159.1 *Rhinopithecus roxellana*	MVLSPADKTNVKAAWGKVGGHGGEYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKF LASVSTVLTSKYR
XP_004705821.1 *Echinops telfairi*	MVLSAADKANVKAVWEKAGGNVGKYGGEALDRTFLSFPTTKTY FPHMDLTPGSADIMAHGKKVADALTLAVGHMDDLPGALSKLSD LHAYKLRVDPVNFKLLSHCLLVTLACHLGGDFTPAAHASLDKF LSSVSTVLTSKYR
XP_022450006.1 *Delphinapterus leucas*	MVLSPADKTNVKGTWAKIGNHSAEYGAEALERMFISFPSTKTY FSHFDLGHGSAQIKGHGKKVADALTKAVGHIDNLPDALSELSD LHAHKLRVDPVNFKLLSHCLLVTLALHLPADFTPSVHASLDKF LASVSTVLTSKYR
XP_010618933.1 *Fukomys damarensis*	MVLSPADKSNVKAAWDKIGGHGAQYGAEALCRMFLSFPTTKTY FHHFDLSPGSAQVQGHGKKVADALTTAVGHLDDLPNALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHHPAEFTPSVHASLDKF LATVSTVLTSKYR
XP_011603007.1 *Takifugu rubripes*	MSLSRTDKEAVKAIWAKMSKSIDVIGAEAFGRMLIAYPQTKIY FSEWSDLRPASGPVKAHGKKVMGGIATAVASIDDLTCGLRELS ERHAFTLKVDPANFRLLAHCILVVTAIMFPKDFTPEIHVSFDK FLAGVALALSDKYR
XP_008849014.1 *Nannospalax galili*	MVLSPEDKNHVRSTWDKIGGHGAEYGAEALERMFTSFPTTKTY FPHFDVSHGSAQVKAHGKKVADALANAAGHLDDLPGALSALSD LHAHKLRVDPVNFKLLSHCLLVTLANHHPAEFTPGVHASLDKF LASVSTVLTSKYR
XP_015666327.1 *Protobothrops mucrosquamatus*	MVLSDDDKARVRAAWVPVCKNAEMYGSETLTRMFAAHATTKTY FPHFDLSPGSSDLKAHGKKVIDALTEAVNNLDDVPGALSKLSD LHAHKLRVDPVNFRLLGHCLEVTIAAHNGGPLKPEVMLALDKF LNLVAKVLVSRYR
XP_007130440.2 *Physeter catodon*	MVLSPADKTNVKAAWAKVGSHAADFGAEALERMFMSFPSTKTY FSHFDLGHNSTQVKGHGKKVADALTKAVGHLDTLPDALSDLSD

	LHAHKLRVDPVNFKLLSHCLLVTLAAHLPGDFTPPVHASLDKF LASVSTVLTSKYR
XP_008155896.1 *Eptesicus fuscus*	MVLSPADKSNVKAAWDKVGGNAGDYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKGHGKKVGDALGSAVAHMDDLPGALSALSD LHAYKLRVDPVNFKLLSHCLLVTLACHNPAEFTPAVHASLDKF LASVSTVLVSKYR
XP_006161020.1 *Tupaia chinensis*	MVLSPGDKSNIKAAWGKIGGQAPQYGAEALERMFLSFPTTKTY FPHFDMSHGSAQIQAHGKKVADALSTAVGHLDDLPTALSALSD LHAHKLRVDPANFKLLSHCILVTLACHHPGDFTPEIHASLDKF LANVSTVLTSKYR
XP_026973857.1 *Lagenorhynchus obliquidens*	MVLSPADKTNVKGTWSKIGNHSAEYGAEALERMFINFPSTKTY FSHFDLGHGSAQIKGHGKKVADALTKAVGHIDNLPDALSELSD LHAHKLRVDPVNFKLLSHCLLVTLALHLPADFTPSVHASLDKF LASVSTVLTSKYR
XP_026862523.1 *Electrophorus electricus*	MSLTAKDKSIVKAFWGKVSSKADDIGAEAFGRMLTVYPQTKTY FASWSDLSPGSAAVKKHGKTIMGGIAEAVAHIDDLTGGLASLS ELHAFKLRVDPANFKILAHNLIVVLALFFHGDFTPEVHMAVDK FFQNVAWALSEKYR
XP_005349098.1 *Microtus ochrogaster*	MVLSGDDKTNIKTAWGKIGGHAGEFGAEALERMFVVYPTTKTY FPHFDVSHGSAQVKGHGKKVADALTTAVGHLDDLPGALSALSD LHAHKLRVDPVNFKLLSHCLLVTLANHIPAEFTPAVHASLDKF LASVSTVLTSKYR
XP_026570824.1 *Pseudonaja textilis*	MVLTEEDKARVRASWVPVSKNAELYGAETLTRLFSAHPTTKTY FPHFDLSPGSHDLKAHGKKVIDALTEAVNNLDDVAGALSKLSD LHAQKLRVDPVNFKLLGQCLEVTIAAHNGGPLKPEVILSLDKF LDLASKLLVSRYR
XP_026341218.1 *Ursus arctos horribilis*	MVLSPADKSNVKATWDKIGSHAGEYGGEALERTFASFPTTKTY FPHFDLSPGSAQVKAHGKKVADALTTAAGHLDDLPGALSALSD LHAHKLRVDPVNFKFLSHCLLVTLASHHPAEFTPAVHASLDKF FSAVSTVLTSKYR
XP_026308403.1 *Piliocolobus tephrosceles*	MVLSPADKTNVKTAWGKVGGHGGEYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKGHGKKVADALATAVAHLDDMPSALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHHPAEFTPAVHASLDKF LASVSTVLTSKYR
XP_026241165.1 *Urocitellus parryii*	MVLSPADKTNVKACWEKIGGHGAAYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVQGHGKKVADALANAAAHVDDLPGALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHHPAEFTPAVHASLDKF LASVSTVLTSKYR
XP_026109996.1 *Carassius auratus*	MSLSDKDKAVVKALWAKIGSRADEIGAEALGRMLTVYPQTKTY FSHWSDLSPGSGPVKKHGKTIMGAVGDAVSKIDDLVGALSSLS ELHAFKLRIDPANFKILAHNVIVVIGMLFPGDFTPEVHMSVDK FFQNLALALSEKYR
XP_025963486.1 *Dromaius novaehollandiae*	MVLSAADKTNTKSVFAKIGPHAEEYGAETLERLFTTYPQTKTY FPHFDLHHGSAQVKAHGKKVAAALVEAANHIDDISTALSKLSD LHAQKLRVDPVNFKLLGQCFLVVVAIHHPSLLTPEVHASLDKF LCAVANVLTAKYR
XP_025840611.1 *Vulpes Vulpes*	MVLSPADKTNIKSTWDKIGGHAGDYGGEALDRTFQSFPTTKTY FPHFDLSPGSAQVKAHGKKVADALTTAVAHLDDLPGALSALSD LHAYKLRVDPVNFKLLSHCLLVTLACHHPNEFTPAVHASLDKF FTAVSTVLTSKYR
XP_005448157.1 *Oreochromis niloticus*	MSLTEKDKAAVKALWAKISKSVDAIGAEALGRMLLVYPQTKTY FSHWPDLTPGSAPVVSHGKQIMGGVTEAMSKIDNLRGGLLELS ELHAFKLRVDPSNFKILAQTIMVVVAAMFPNDFTPEAHVAFDK FLAAVALGLSERYR
XP_025733960.1 *Callorhinus ursinus*	MVLSPADKTNVKTTWDKLGGHAGEYGGEALERTFTSFPTTKTY FPHFDLSPGSAQVKAHGKKVADALTTAVAHLDDLPGALSTLSD LHAYKLRVDPVNFKLLSHCLLVTLACHHPAEFTPAVHASLDKF FSAVSTVLTSKYR
XP_025273010.1 *Canis lupus dingo*	MVLSPADKTNIKSTWDKIGGHAGDYGGEALDRTFQSFPTTKTY FPHFDLSPGSAQVKAHGKKVADALTTAVAHLDDLPGALSALSD

	LHAYKLRVDPVNFKLLSHCLLVTLACHHPTEFTPAVHASLDKF FAAVSTVLTSKYR
XP_006020935.1 *Alligator sinensis*	MVLSQEDKSNVKAIWGKASGHLEDYGAEVLERMFCAYPQTKIY FPHFDMSHGSPQIRAHGKKVFSALHEAVNHIDDLPGALCRLSE LHAHSLRVDPVNFKFLSHCVLVVFAIHHPCSLSPEVHASLDKF LCAVSAVLTSKYR
XP_025230451.1 *Theropithecus gelada*	MVLSPDDKKHVKDAWGKVGEHAGQYGAEALERMFLSFPTTKTY FPHFDLSHGSDQVKKHGKKVADALTLAVGHVDDMPQALSKLSD LHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKF LASVSTVLTSKYR
XP_012777671.2 *Maylandia zebra*	MSLTEKDKAAVKALWAKVSKVMDTVGGEALGRMLLVYPQTKTY FSHWPDLTPGSEPVMVHGKLILGGVTEAVSKIDNLRGGLLELS ELHAFKLRVDPSNFKMLAHCAMVVIAIMFPKDFTPETHVAFDK FLAAVALGLSEKYR
XP_024615545.1 *Neophocaena asiaeorientalis asiaeorientalis*	MVLSPADKTNVKGTWAKIGNHSAEYGAEALERMFINFPSTKTY FSHFDLGHGSAQIKGHGKKVADALTKAVGHIDNLPDALSELSD LHAHKLRVDPVNFKLLSHCLLVTLALHLPADFTPSVHASLDKF LASVSTVLTSKYR
XP_024408880.1 *Desmodus rotundus*	MVLSAADKGNVKTAWDKVGGQAGNYGAEALERMFLGFPTTKTY FPHFDMSHDSAQIKGHGKKVADALTVAVGHMDDLPSALSGLSD LHAYKLRVDPVNFKLLGHCLLVTLACHHPSDFTPAVHASLDKF LASVSTVLTSKYR
XP_023868215.1 *Salvelinus alpinus*	MSLTAKDKSVVKAFWGKISGKADVIGAEALGRMLTAYPQTKTY FSHWADLSPGSAPVKKHGGVIMGAIGNAVGXMDNLVGGLXALS DLHAFKLRVDPGNFKILSHNILVTLAIHFPGDFTPEVHIAVDK FLAALSAALADKYR
XP_004628004.1 *Octodon degus*	MVLSPADKTNVKTAWGKIGGHGAEYGAEALFRMFLSFPTTKTY FHHFDLSAGSAQIKSHGKKVSDALTTAVDHLDDLPTALSALSD LHAHKLRVDPVNFKLLSHCLLVTLSAHHPADFTPAVHASLDKF LATVSTVLTSKYR
XP_011375036.1 *Pteropus vampyrus*	MVLSSTDKSNVKAAWDKVGGNVGEYGAEALERMFLSFPTTKTY FPHFDLAHGSSQVKAHGKKVGDALTNAVGHIDDLPGALSALSD LHAYKLRVDPVNFKLLSHCLLVTLASHLPSDFTPAVHASLDKF LASVSTVLTSKYR
XP_003478455.1 *Cavia porcellus*	MVLSAADKNNVKTTWDKIGGHAAEYVAEGLTRMFTSFPTTKTY FHHIDVSPGSGDIKAHGKKVADALTTAVGHLDDLPTALSTLSD VHAHKLRVDPVNFKFLNHCLLVTLAAHLGADFTPSIHASLDKF FASVSTVLTSKYR
XP_004474053.2 *Dasypus novemcinctus*	MVLSAADKTHVKAFWGKVGGHAAEFGAEALERMFASFPPTKTY FSHMDLSHGSAQVKAHGKKVADALTLAVGHLDDLPGALSTLSD LHAHKLRVDPVNFKFLSHCLLVTLACHLPDDFTPAVHASMDKF MAGVSTVLVSKYR
XP_023102549.1 *Felis catus*	MVLSAADKSNVKACWGKIGSHAGEYGAEALERTFCSFPTTKTY FPHFDLSHGSAQVKAHGQKVADALTQAVAHMDDLPTAMSALSD LHAYKLRVDPVNFKFLSHCLLVTLACHHPAEFTPAVHASLDKF FSAVSTVLTSKYR
XP_007238255.1 *Astyanax mexicanus*	MSLTVDDMAVVKAFWGKIGAKADDIGAEALGRMLIVYPQTKTY FAHWADLSPGSAPVKKHGKTIMGAVTAAVGSIEDLPSALSQLS ELHAYKLRVDPANFKILAHNIIVVMGMLFPNDFTPEVHVSVDK FLQNLAWCLAERYR
XP_012734007.1 *Fundulus heteroclitus*	MSLSETDKSRVRAFWAKAEGKANELGGEALARMLVSTPQTKTY FAHWGDLSPQSAKVRKHGATIMGALGKAVKGIDDLTGTLGALS ELHAFKLRVDPANFKILGHNIVLVFAMYFPADFTPEVHVSVEK FLQCVAWALSEKYR
XP_004873772.1 *Heterocephalus glaber*	MSLSNEDKACLRSVWKEIGPSWPEHCPDAIYRMFLSFPSTKTY FPNFDISPGSPQIQAHGRKVADALNKAVEHIDDMPAALSDLSD KHSQELRVDPVNFKLLKHTMLVTMAANYPEILTPEVLLSLDKL MEAVSRVLISRYR
XP_020942625.1 *Sus scrofa*	MVLSAADKANVKAAWGKVGGQAGAHGAEALERMFLGFPTTKTY FPHFNLSHGSDQVKAHGQKVADALTKAVGHLDDLPGALSALSD

	LHAHKLRVDPVNFKLLSHCLLVTLAAHHPDDFNPSVHASLDKF LANVSTVLTSKYR
XP_004873771.1 *Heterocephalus glaber*	MVLSPADKSNVKAAWDKIGGHGAQYGAEALTRMFLSFPTTKTY FHHFDLSPGSAQIQGHGKKVADALTTAVGHLDDLPSALSALSD LHAHKLRVDPVNFKLLSHCLLVTLAAHHPAEFTPAVHASLDKF LATVSTVLTSKYR
XP_020755843.1 *Odocoileus virginianus texanus*	MVLSAADKSNVKAAWGKVGGNAPAYGAEALERMFLSFPTTKTY FPHFDLSHGSAQVKAHGEKVANALTKAVGHLDDLPGTLSDLSD LHAHKLRVDPVNFKLLSHTLLVTLASHLPNDFTPAVHASLDKF LANVSTVLTSKYR

Table 2: Selected clinical phenotypes compared to initial predictions based on tolerance to amino acid substitution. Here, "deleterious" mean either detrimentally affects the function of hemoglobin (Hb) and/or results in disease in affected patients. We selected one clinically characterised mutation per residue (where available) to test our predictions.

Variant	Clinical Phenotype	References	Predicted phenotype	Agreement?
α1Val>Glu p.V2E	Hb Antananarivo (Reduced sensitivity to regulation by chloride ions)	[42]	Deleterious	Yes
α2Leu>Pro p.L3P	Hemoglobin Chongqing (Increased oxygen affinity)	[43]	Deleterious	Yes
α5Ala>Pro p.A6P	Hemoglobin Karachi (Asymptomatic)	[44]	Neutral	Yes
α6Asp>Gly p.D7G	Hb Swan River (Anaemia and fatigue)	[45]	Deleterious	Yes
α7Lys>Asn p.K8N	HB Tatras (Abnormality of blood and blood-forming tissues Jaundice and cyanosis)	[46]	Deleterious	Yes
α9Asn>Lys p.N10K	Hb Park Ridge (Asymptomatic)	[47]	Neutral	Yes
α11Lys>Gln p.K12Q	Hb Wuming (Asymptomatic)	[48]	Neutral	Yes
α12Ala>Asp p.A13D	Hemoglobin J Paris (Asymptomatic)	[49]	Neutral	Yes

α14Trp>Arg p.W15R	Hemoglobin Evanston (α-Thalassemia)	[50]	Deleterious	Yes
α15Gly>Ser p.G16S	Hb SIAM (Asymptomatic)	[51]	Neutral	Yes
α16Lys>Met p.K17M	Hb Harbin (Anaemia and hepatomegaly and anisocytosis (red blood cells of unequal size))	[43]	Deleterious	Yes
α18Gly>Asp p.G19D	HB Al-Ain Abu Dhabi (Asymptomatic except microcytosis and hypochromia)	[52]	Neutral	Yes
α19Ala>Glu p.A20E	Hb J-Tashikuergan (Asymptomatic)	[53]	Neutral	Yes
α20His>Pro p.H21P	Hb Anderlecht (Asymptomatic)	[54]	Deleterious	No
α21Ala>Pro p.A22P	Hb Fontainebleau (Protein unstable with no detectable symptoms)	[55]	Neutral	No
α22Glu>Asp p.E23D	Hb Lisbon (asymptomatic)	[46]	Neutral	Yes
α24Tyr>His p.Y25H	Hb Luxembourg (Anaemia and bilirubinemia)	[56]	Deleterious	Yes
α26Ala>Val p.A27V	Hb Campinas (Asymptomatic)	[57]	Deleterious	No

α27Glu>Asp p.E28D	Hb Hekinan (Anaemia)	[30]	Neutral	No
α29Leu>Pro p.L30P	Hb Agrinio (Thalassaemia); Heme binding residue	[58]	Deleterious	Yes
α30Glu>Lys p.E31K	Hb O Padova (Dyserythropoietic anemia); β-globin binding residue	[31]	Neutral	No
α31Arg>Ser p.R32S	Hb Prato (Anisocytosis and hypochromia); β-globin binding residue	[59]	Deleterious	Yes
α33Phe>Ser p.F34S	Hb Chartres (Microcytic hypochromic anaemia)	[60]	Deleterious	Yes
α34Leu>Arg p.L35R	Hb Queens (Mild anisocytosis and anaemia); β-globin binding residue	[32, 61, 62]	Neutral	No
α35Ser>Pro p.S36P	Hb Evora (α-Thalassemia; unstable protein); AHSP and β-globin binding residue	[63]	Deleterious	Yes
α37Pro>Leu p.P38L	Hb Manawatu (Slightly unstable but asymptomatic – residue involved in internal α1β2 contacts in deoxyHb tetramer)	[64]	Deleterious	Yes
α40Lys>Glu p.K41E	Hb Kariya (Asymptomatic but residue forms salt bridge with β-chain in deoxyHb and stabilises T-state)	[65]	Deleterious	No

α41Thr>Ser p.T42S	Hb Miyano (Erythremia due to decreased oxygen transport caused by high oxygen affinity and decreased cooperativity)	[66]	Deleterious	Yes
α42Tyr>His p.Y43H	Hb Barika (α-Thalassemia); heme binding residue	[67]	Deleterious	Yes
α43Phe>Val p.F44V	Hb Torino (Inclusion Body anaemia); heme binding residue	[68]	Deleterious	Yes
α44Pro>Arg p.P45R	Hb Kawachi (Polycythemia and reticulocytosis)	[69]	Deleterious	Yes
α45His>Asp p.H46D	Hb Poitiers (Anaemia and anisocytosis – increased oxygen affinity); heme binding residue	[70]	Deleterious	Yes
α46Phe>Val p.F47V	Hb Hillingdon (Asymptomatic); heme binding residue	[26]	Deleterious	No
α47Asp>Ala p.D48A	Hb Cordele (Asymptomatic but slightly unstable)	[71]	Deleterious	Yes
α48Leu>Arg p.L49R	Hb Montgomery (Asymptomatic)	[72]	Deleterious	No
α49Ser>Arg p.S50R	Hb Savaria (Asymptomatic)	[73]	Deleterious	No
α50His>Asp	Hb J Sardegna (Asymptomatic but often co-inherited with α- and β-thalassaemia)	[74]	Deleterious	No

p.H51D				
α51Gly>Arg	Hb Russ (Asymptomatic; possibly reduced stability protein)	[75-77]	Deleterious	Yes
p.G52R				
α53Ala>Asp	Hb J-Rovigo (Asymptomatic)	[78]	Neutral	Yes
p.A54D				
α54Gln>Arg	Hb Shimonoseki (Asymptomatic)	[79]	Neutral	Yes
p.Q55R				
α55Val>Ala	Hb Gerland (Microcytic anaemia)	[80]	Neutral	No
p.V56A				
α56Lys>Glu	Hb Shaare Zedek (Hypochromic anaemia)	[81]	Deleterious	Yes
p.K57E				
α57Gly>Asp	Hb Norfolk (Asymptomatic)	[82]	Neutral	Yes
p.G58D				
α58His>Tyr	Hb M Boston (Methemoglobinemia and cyanosis); heme binding residue	[83]	Deleterious	Yes
p.H59Y				
α59Gly>Arg	Hb Zurich Albisrieden (Highly unstable; α-thalassemia)	[84]	Deleterious	Yes
p.G60R				
α60Lys>Glu	Hb Dagestan (Asymptomatic)	[85]	Neutral	Yes
p.K61E				

α61Lys>Thr p.K62T	Hb J-Anatolia (Hypochromic microcytic anaemia and anisopoikilocytosis – red blood cells deformed and uneven in size)	[86]	Deleterious	Yes
α62Val>Met p.V63M	Hb Evans (Unstable leads to haemolytic anaemia); heme binding residue	[87]	Deleterious	Yes
α63Ala>Asp p.A64D	Hb J-Pontoise (Slightly unstable but asymptomatic)	[88]	Deleterious	Yes
α64Asp>Gly p.D65G	Hb Hangzhou (Asymptomatic)	[89]	Neutral	Yes
α65Ala>Thr p.A66T	Hb Part-Dieu (Hepatosplenomegaly – uncertain if related to patient's diabetes)	[27]	Deleterious	No
α66Leu>Pro p.L67P	Hb Dartmouth (HbH disease)	[90]	Deleterious	Yes
α68Asn>His p.N69H	Hb St. Truiden (Asymptomatic)	[91]	Neutral	Yes
α69Ala>Thr p.A70T	Hb Decines-Charpieu (Asymptomatic)	[27]	Deleterious	No
α71Ala>Val p.A72V	Hemoglobin Ozieri (Asymptomatic)	[92]	Neutral	Yes
α72His>Asp	Hb Norton (Asymptomatic)	[47]	Neutral	Yes

p.H73D				
α74Asp>Ala	Hemoglobin Lille (Anaemia)	[93]	Deleterious	Yes
p.D75A				
α75Asp>Val	Hb Al-Hammadi Riyadh (Anaemia)	[94]	Deleterious	Yes
p.D76V				
α76Met>Thr	Hb Aztec (Asymptomatic)	[95]	Deleterious	No
p.M77T				
α77Pro>His	Hb Toulon (Asymptomatic)	[96]	Neutral	Yes
p.P78H				
α78Asn>His	Hb Davenport (Asymptomatic)	[97]	Neutral	Yes
p.N79H				
α79Ala>Thr	Hb Mantes-la-Jolie (Asymptomatic)	[98]	Neutral	Yes
p.A80T				
α80Leu>Arg	Hb Ann Arbor (α-Thalassaemia)	[99]	Deleterious	Yes
p.L81R				
α81Ser>Pro	Hb Passy (Microcytosis and anaemia)	[100]	Deleterious	Yes
p.S82P				
α83Leu>Pro	HB Les Andelys (Unstable causing anisocytosis); heme binding residue.	[101]	Deleterious	Yes
p.L84P				

α84Ser>Arg p.S85R	Hb Etobicoke (Unstable but not known to cause anaemia)	[102]	Deleterious	No
α85Asp>Tyr p.D86Y	Hb Atago (Increased oxygen affinity - no reported symptoms)	[103]	Deleterious	No
α86Leu>Arg p.L87R	Hb Moabit (Heinz body haemolytic anaemia and splenomegaly); heme binding residue	[104]	Deleterious	Yes
α87His>Tyr p.H88Y	Hb M-Iwate (Decreased oxygen affinity and methemoglobin); heme binding residue	[105]	Deleterious	Yes
α88Ala>Val p.A89V	Hb Columbia Missouri (Erythrocytosis)	[106]	Deleterious	Yes
α89His>Leu p.H90L	Hb Luton (α-Thalassaemia)	[107]	Neutral	No
α90Lys>Arg p.K91R	Hb Clinico-Madrid II (Asymptomatic)	[108, 109]	Neutral	Yes
α91Leu>Pro p.L92P	Hb Port Phillip (Haemolytic anaemia; residue is important in the stability of oxygenated state); heme binding residue	[110]	Deleterious	Yes
α92Arg>Gln p.R93Q	Hb J Cape Town (Hypochromic anaemia and thalassaemia)	[111]	Deleterious	Yes
α93Val>Gly p.V94G	Hb Bronte (α-Thalassemia); heme binding residue	[112]	Deleterious	Yes

α94Asp>Asn p.D95N	Hb Titusville (Low oxygen affinity variant - cyanosis)	[113, 114]	Deleterious	Yes
α95Pro>Ser p.P96S	Hb Rampa (Asymptomatic); AHSP binding residue	[115, 116]	Deleterious	No
α96Val>Asp p.V97D	Hb El Escorial (Microcytic and hypochromic ferropenic anaemia); AHSP binding residue	[108]	Deleterious	Yes
α97Asn>His p.N98H	Hb Fuchu-ll (Reduced oxygen carrying capacity and polycythaemia)	[117, 118]	Deleterious	Yes
α99Lys>Glu p.K100E	Hb Turriff (Asymptomatic); AHSP and β-globin binding residue	[119]	Deleterious	No
α102Ser>Arg p.S103R	Hb Manitoba (Anaemia)	[120]	Deleterious	Yes
α103His>Arg p.H104R	Hb Contaldo (Haemolytic anaemia); AHSP and β-globin binding residue	[121]	Deleterious	Yes
α104Cys>Tyr p.C10Y	Hb Sallanches (HbH in homozygote, undetected in heterozygote)	[122, 123]	Deleterious	Yes
α108Thr>Asn p.T109N	Hb Bleuland (Moderate anaemia)	[124]	Deleterious	Yes
α109Leu>Arg p.L110R	Hb Suan-Dok (Anaemia)	[125]	Deleterious	Yes

α110Ala>Asp p.A111D	Hb Petah Tikva (α-Thalassaemia); β-globin binding residue	[126]	Deleterious	Yes
α111Ala>Val p.A112V	Hb Anamosa (Asymptomatic); AHSP and β-globin binding residue	[127]	Neutral	Yes
α112His>Gln p.H113Q	Haemoglobin Lleida (α-Thalassaemia partly due to additional 12bp deletion in patient)	[128]	Deleterious	Yes
α113Leu>His p.L114H	Hb Twin Peaks (Asymptomatic)	[129]	Neutral	Yes
α114Pro>Leu p.P115L	Hb Nouakchott (Asymptomatic); β-globin binding residue	[130]	Deleterious	No
α115Ala>Asp p.A116D	Haemoglobin J Tongariki (Asymptomatic)	[131]	Neutral	Yes
α116Glu>Gln p.E117Q	Hemoglobin Oleander (Anisopoikilocytosis, polychromasia and low red blood cell counts)	[132]	Neutral	No
α120Ala>Glu p.A121E	Hb J- Meerut (Increased oxygen affinity and red blood cell count); AHSP and β-globin binding residue	[133]	Neutral	No
α121Val>Met p.V122M	Hb Owari (Asymptomatic)	[134]	Neutral	Yes
α122His>Gln p.H123Q	Hb Westmead (Anaemia); AHSP and β-globin binding residue	[29]	Neutral	No

α125Leu>Arg p.L126R	Hb Plasencia (α-Thalassaemia)	[135]	Deleterious	Yes
α126Asn>Asn p.D127N	Hb Tarrant (Erythrocytosis); AHSP and β-globin binding residue	[136]	Deleterious	Yes
α127Lys>Asn p.K128N	Hemoglobin Jackson (Asymptomatic); AHSP and β-globin binding residue	[137]	Deleterious	No
α129Leu>Pro p.L130P	Hb Utrecht (α-Thalassaemia); Heme binding residue	[138]	Deleterious	Yes
α130Ala>Pro p.A131P	Hb Sun Prairie (Hyper-unstable; haemolytic anaemia in homozygotes)	[139]	Deleterious	Yes
α131Ser>Pro p.S132P	Hb Questembert (Mild anaemia)	[140]	Deleterious	Yes
α132Val>Gly p.V133G	Hb Caen (Haemolytic anaemia and piokilocytosis); heme binding residue	[140]	Deleterious	Yes
α133Ser>Arg p.S134R	Hb Val de Marne (Anisocytosis)	[141]	Deleterious	Yes
α135Val>Glu p.V136E	Hb Pavie (Decreased oxygen affinity; no known symptoms)	[142]	Deleterious	No
α136Leu>Met p.L137M	Hb Chicago (Asymptomatic); heme binding residue	[28]	Deleterious	No

α138Ser>Pro p.S139P	Hb Attleboro (Non-cooperative high-affinity haemoglobin; no described symptoms)	[143]	Deleterious	No
α139Lys>Glu p.K140E	Hb Hanamaki (Erythrocytosis)	[144]	Deleterious	Yes
α140Tyr>His p.Y141H	Hemoglobin Rouen (Increased oxygen-affinity and decreased cooperativity, no described symptoms)	[145]	Deleterious	No
α141Arg>His p.R142H	Hb Suresnes (Erythrocytosis)	[146]	Deleterious	Yes

Figure legends

Figure 1: Predicted tolerance to change for each amino residue in human α-globin. The maximum score is 19, meaning that the altering the residue to any of the 19 possible alternatives is predicted to be tolerated.

Figure 2: Summary of PredictSNP predictions for human α-globin. Changes predicted to be destabilising are shown with an "X" on a dark pink background. The wild-type residue is shown as "WT" on a grey background. Where the change is predicted to have no effect (i.e. tolerated), the cell is left blank (white). Functionally important residues are annotated as follows: H, heme binding (blue); B, β-globin binding (RED); A, AHSP binding (yellow); A/B, AHSP and β-globin binding (orange); A/B/H, AHSP, β-globin and heme binding (purple).

Figure 3: Sequence conservation scores for each residue in human α-globin. A score of 1.0 indicates no variation across 60 vertebrate α-globin sequences.

Figure 4: Correlation of the predicted tolerance of residues to change and sequence conservation. Each point represents an amino acid residue and the line linear regression calculated by Excel (Microsoft, USA). r^2=-0.7065; $p<0.001$.

References

[1] C.S. Thom, C.F. Dickson, D.A. Gell, M.J. Weiss, Hemoglobin variants: biochemical properties and clinical correlates, Cold Spring Harb Perspect Med, 3 (2013) a011858.
[2] S.A. Liebhaber, F.E. Cash, S.K. Ballas, Human alpha-globin gene expression. The dominant role of the alpha 2-locus in mRNA and protein synthesis, J Biol Chem, 261 (1986) 15327-15333.
[3] S.A. Liebhaber, Y.W. Kan, Different rates of mRNA translation balance the expression of the two human alpha-globin loci, J Biol Chem, 257 (1982) 11852-11855.
[4] H. Akhavan-Niaki, R. Youssefi Kamangari, A. Banihashemi, V. Kholghi Oskooei, M. Azizi, A. Tamaddoni, S. Sedaghat, M. Vakili, H. Mahmoudi Nesheli, S. Shabani, Hematologic features of alpha thalassemia carriers, Int J Mol Cell Med, 1 (2012) 162-167.
[5] C.L. Harteveld, D.R. Higgs, Alpha-thalassaemia, Orphanet J Rare Dis, 5 (2010) 13.
[6] D.H. Chui, S. Fucharoen, V. Chan, Hemoglobin H disease: not necessarily a benign disorder, Blood, 101 (2003) 791-800.
[7] S. Fucharoen, V. Viprakasit, Hb H disease: clinical course and disease modifiers, Hematology Am Soc Hematol Educ Program, (2009) 26-34.
[8] B. Trainor, R. Tubman, The emerging pattern of hydrops fetalis--incidence, aetiology and management, Ulster Med J, 75 (2006) 185-186.
[9] H. Nakayama, J. Kukita, S. Hikino, H. Nakano, T. Hara, Long-term outcome of 51 liveborn neonates with non-immune hydrops fetalis, Acta Paediatr, 88 (1999) 24-28.
[10] P. Yue, Z. Li, J. Moult, Loss of protein structure stability as a major causative factor in monogenic disease, Journal of Molecular Biology, 353 (2005) 459-473.
[11] T.L. Mollan, X. Yu, M.J. Weiss, J.S. Olson, The role of alpha-hemoglobin stabilizing protein in redox chemistry, denaturation, and hemoglobin assembly, Antioxid Redox Signal, 12 (2010) 219-231.
[12] L. Feng, D.A. Gell, S. Zhou, L. Gu, Y. Kong, J. Li, M. Hu, N. Yan, C. Lee, A.M. Rich, R.S. Armstrong, P.A. Lay, A.J. Gow, M.J. Weiss, J.P. Mackay, Y. Shi, Molecular mechanism of AHSP-mediated stabilization of alpha-hemoglobin, Cell, 119 (2004) 629-640.
[13] R.C. Hardison, D.H. Chui, B. Giardine, C. Riemer, G.P. Patrinos, N. Anagnou, W. Miller, H. Wajcman, HbVar: A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server, Hum Mutat, 19 (2002) 225-233.
[14] G.P. Patrinos, B. Giardine, C. Riemer, W. Miller, D.H. Chui, N.P. Anagnou, H. Wajcman, R.C. Hardison, Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies, Nucleic Acids Res, 32 (2004) D537-541.
[15] J. Bendl, J. Stourac, O. Salanda, A. Pavelka, E.D. Wieben, J. Zendulka, J. Brezovsky, J. Damborsky, PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations, PLoS computational biology, 10 (2014) e1003440.
[16] J.D. Thompson, T.J. Gibson, D.G. Higgins, Multiple sequence alignment using ClustalW and ClustalX, Current protocols in bioinformatics / editoral board, Andreas D.Baxevanis ...[et al.], Chapter 2 (2002) Unit 2.3.
[17] W.S. Valdar, Scoring residue conservation, Proteins, 48 (2002) 227-241.
[18] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. Gallo Cassarino, M. Bertoni, L. Bordoli, T. Schwede, SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic Acids Res, 42 (2014) W252-258.
[19] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, Journal of applied crystallography, 26 (1993) 283-291.
[20] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, Acta Crystallogr D Biol Crystallogr, 67 (2011) 235-242.
[21] S.Y. Park, T. Yokoyama, N. Shibayama, Y. Shiro, J.R. Tame, 1.25 A resolution crystal structures of human haemoglobin in the oxy, deoxy and carbonmonoxy forms, J Mol Biol, 360 (2006) 690-701.

[22] E. Domingues-Hamdi, C. Vasseur, J.B. Fournier, M.C. Marden, H. Wajcman, V. Baudin-Creuza, 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, 9 (2014) e111395.
[23] M.F. Tam, N.W. Rice, D.H. Maillett, V. Simplaceanu, N.T. Ho, T.C. Tam, T.J. Shen, C. Ho, Autoxidation and oxygen binding properties of recombinant hemoglobins with substitutions at the alphaVal-62 or betaVal-67 position of the distal heme pocket, J Biol Chem, 288 (2013) 25512-25521.
[24] S. Salentin, S. Schreiber, V.J. Haupt, M.F. Adasme, M. Schroeder, PLIP: fully automated protein-ligand interaction profiler, Nucleic Acids Res, 43 (2015) W443-447.
[25] A. Porollo, J. Meller, Prediction-based fingerprints of protein-protein interactions, Proteins, 66 (2007) 630-645.
[26] A. Babb, S. Solaiman, B.N. Green, D. Mantio, K. Patel, HB Hillingdon [alpha46(CE4)Phe-->Val (alpha1 Or alpha2)]: a new alpha chain hemoglobin variant, Hemoglobin, 33 (2009) 503-506.
[27] P. Lacan, M. Becchi, I. Zanella-Cleon, M. Aubry, P. Renaudier, A. Francin, Two new alpha chain variants: Hb Part-Dieu [alpha65(E14)Ala --> Thr (alpha2)] and Hb Decines-Charpieu [alpha69(E18)Ala --> Thr (alpha2)], Hemoglobin, 28 (2004) 51-57.
[28] J.E. Bowman, R. Bloom, S.S. Chen, B.B. Webber, J.B. Wilson, F. Kutlar, A. Kutlar, T.H. Huisman, HB Chicago or alpha (2)136 (H19) Leu----Met beta 2 and a -G gamma-G gamma-globin gene arrangement in a black family, Hemoglobin, 10 (1986) 495-505.
[29] P.J. Fleming, W.G. Hughes, R.K. Farmilo, K. Wyatt, W.N. Cooper, Hemoglobin Westmead alpha 2 122(H5)His replaced by Gln beta 2: a new hemoglobin variant with the substitution in the alpha 1 beta 1 contact area, Hemoglobin, 4 (1980) 39-52.
[30] H.C. Shih, M.C. Shih, Y.C. Chang, C.T. Peng, T.J. Chang, J.G. Chang, Hb Hekinan in a Taiwanese subject: a G-->T substitution at codon 27 of the alpha1-globin gene abolishes an HaeIII site, Hemoglobin, 31 (2007) 495-498.
[31] L. Vettore, G. De Sandre, E.E. Di Iorio, K.H. Winterhalter, A. Lang, H. Lehmann, A new abnormal hemoglobin O Padova, alpha 30 (B11) Glu -- Lys, and a dyserythropoietic anemia with erythroblastic multinuclearity coexisting in the same patient, Blood, 44 (1974) 869-877.
[32] W.F. Moo-Penn, D.L. Jue, M.H. Johnson, J.E. McGuffey, H. Simpkins, J. Katz, Hemoglobin Queens: alpha 34 (B15) Leu-Arg structural and functional properties and its association with Hb E, Am J Hematol, 13 (1982) 323-327.
[33] D.J. Timson, Value of predictive bioinformatics in inherited metabolic diseases, World Journal of Medical Genetics, 5 (2015) 46-51.
[34] M.Y. Frederic, M. Lalande, C. Boileau, D. Hamroun, M. Claustres, C. Beroud, G. Collod-Beroud, UMD-predictor, a new prediction tool for nucleotide substitution pathogenicity -- application to four genes: FBN1, FBN2, TGFBR1, and TGFBR2, Human mutation, 30 (2009) 952-959.
[35] T.L. McGee, B.J. Seyedahmadi, M.O. Sweeney, T.P. Dryja, E.L. Berson, Novel mutations in the long isoform of the USH2A gene in patients with Usher syndrome type II or non-syndromic retinitis pigmentosa, Journal of medical genetics, 47 (2010) 499-506.
[36] C. Vasseur, Y. Blouquit, J. Kister, D. Prome, J.S. Kavanaugh, P.H. Rogers, C. Guillemin, A. Arnone, F. Galacteros, C. Poyart, et al., Hemoglobin Thionville. An alpha-chain variant with a substitution of a glutamate for valine at NA-1 and having an acetylated methionine NH2 terminus, J Biol Chem, 267 (1992) 12682-12691.
[37] C.L. Harteveld, F.G. Versteegh, E.H. van Leer, J.S. Starreveld, P.J. Kok, I. van Rooijen-Nijdam, P. van Delft, I. Zanella-Cleon, M. Becchi, H. Wajcman, P.C. Giordano, Hb St. Jozef, A Val-->Leu N-terminal mutation leading to retention of the methionine, and partial acetylation found in the globin gene in Cis with a -alpha3.7 thalassemia deletion, Hemoglobin, 31 (2007) 313-323.
[38] D. Tchernitchko, M. Goossens, H. Wajcman, In silico prediction of the deleterious effect of a mutation: proceed with caution in clinical genetics, Clin Chem, 50 (2004) 1974-1978.
[39] J.R. Manning, M.A. Bailey, D.C. Soares, D.R. Dunbar, J.J. Mullins, *In silico* structure-function analysis of pathological variation in the HSD11B2 gene sequence, Physiological genomics, 42 (2010) 319-330.

[40] T.J. McCorvie, D.J. Timson, *In silico* prediction of the effects of mutations in the human UDP-galactose 4'-epimerase gene: Towards a predictive framework for type III galactosemia, Gene, 524 (2013) 95-104.
[41] C. Browne, D.J. Timson, *In Silico* Prediction of the Effects of Mutations in the Human Mevalonate Kinase Gene: Towards a Predictive Framework for Mevalonate Kinase Deficiency, Ann Hum Genet, 79 (2015) 451-459.
[42] J. Kister, C. Prehu, J. Riou, C. Godart, J. Bardakdjian, D. Prome, F. Galacteros, H. Wajcman, Two hemoglobin variants with an alteration of the oxygen-linked chloride binding: Hb Antananarivo [alpha1(NA1)Val-->Gly] and Hb Barbizon [beta144(HC1)Lys-->Met], Hemoglobin, 23 (1999) 21-32.
[43] Y.T. Zeng, S.Z. Huang, X.K. Qiu, G.C. Cheng, Z.R. Ren, Q.C. Jin, C.Y. Chen, C.T. Jiao, Z.G. Tang, R.H. Liu, et al., Hemoglobin Chongqing [alpha 2(NA2)Leu----Arg] and hemoglobin Harbin [alpha 16(A14)Lys----Met] found in China, Hemoglobin, 8 (1984) 569-581.
[44] A. Ahmad, S. Naqvi, S. Ehsanullah, Z.H. Zaidi, Abnormal hemoglobins 11-Hb (Karachi), an alpha chain abnormality at position 5 Ala----Pro, J Pak Med Assoc, 36 (1986) 206-208.
[45] W.F. Moo-Penn, D.L. Jue, M.H. Johnson, B.L. Therrell, Hemoglobin Swan River [alpha 6(A4)Asp----Gly], Hemoglobin, 11 (1987) 61-62.
[46] H. Wajcman, M. Bost, Y. Blouquiit, C. Prehu, J. Riou, F. Galacteros, Two new alpha chain variants found during glycated hemoglobin screening: Hb Tatras [alpha 7(A5)Lys-->Asn] and HB Lisbon [alpha 23(B4)Glu-->Asp], Hemoglobin, 18 (1994) 427-432.
[47] J.D. Hoyer, D.J. McCormick, K. Snow, J.H. Kwon, D. Booth, M. Duarte, G. Grayson, K.S. Kubik, M.W. Holmes, V.F. Fairbanks, Four new variants of the alpha2-globin gene without clinical or hematologic effects: Hb Park Ridge [alpha9(alpha7)Asn-->Lys (alpha2)], Hb Norton [alpha72(EF1)His-->Asp (alpha2)], Hb Lombard [alpha103(G10)His-->Tyr (alpha2)], and Hb San Antonio [A113(GH2)Leu-->Arg (A2)], Hemoglobin, 26 (2002) 175-179.
[48] Y.T. Zeng, S.Z. Huang, L. Xu, G.F. Long, H. Lam, J.B. Wilson, T.H. Huisman, Hb Wuming or alpha 2 11(A9)Lys substituting for Gln beta 2, Hemoglobin, 5 (1981) 679-687.
[49] S. Dash, T.H. Huisman, First observation of hemoglobin J Paris I [alpha-2-12(A10)alanine-aspartic acid beta-2] in the Indian subcontinent, Acta Haematol, 79 (1988) 117.
[50] W.F. Moo-Penn, R.M. Baine, D.L. Jue, M.H. Johnson, J.E. McGuffey, J.M. Benson, Hemoglobin Evanston: alpha 14(A12) Trp leads to Arg. A variant hemoglobin associated with alpha-thalassemia-2, Biochim Biophys Acta, 747 (1983) 65-70.
[51] C. Turbpaiboon, S. Svasti, P. Sawangareetakul, P. Winichagoon, C. Srisomsap, N. Siritanaratkul, S. Fucharoen, P. Wilairat, J. Svasti, Hb Siam [alpha15(A13)Gly-->Arg (alpha1) (GGT-->CGT)] is a typical alpha chain hemoglobinopathy without an alpha-thalassemic effect, Hemoglobin, 26 (2002) 77-81.
[52] S. Abbes, A. M'Rad, P.A. Fitzgerald, P. Dormer, Y. Blouquit, J. Kister, F. Galacteros, H. Wajcman, HB Al-Ain Abu Dhabi [alpha 18(A16)Gly----Asp]: a new hemoglobin variant discovered in an Emiratee family, Hemoglobin, 16 (1992) 355-362.
[53] H.J. Li, D.X. Liu, Z.G. Liu, P. Li, L. Li, J. Chen, S.Z. Hou, A new fast-moving hemoglobin variant, Hb J-Tashikuergan alpha 19(AB1) Ala----Glu, Hemoglobin, 8 (1984) 391-395.
[54] F. Cotton, H. Wajcman, V. Hansen, C. Lin, K. Jotzo, D. Haumont, D. Prome, J. Riou, A. Miyazaki, F. Galacteros, F. Vertongen, B. Gulbis, Hb Anderlecht [alpha20(B1)His-->Pro]: a silent variant found in a Congolese newborn, Hemoglobin, 24 (2000) 299-304.
[55] S.O. Brennan, T. Chan, S. Ryken, A. Ruskova, A second case of Hb Fontainebleau [alpha21(B2)Ala-->Pro] in an individual with microcytosis, Hemoglobin, 33 (2009) 258-261.
[56] P. Groff, F. Galacteros, G. Kalmes, Y. Blouquit, H. Wajcman, Hb Luxembourg [alpha 24(B5) Tyr----His]: a new unstable variant, Hemoglobin, 13 (1989) 429-436.
[57] M.R. Wenning, N.M. Silva, S.B. Jorge, E.M. Kimura, F.F. Costa, Torsoni, S.H. Ogo, M.F. Sonati, Hb Campinas [alpha26(B7)Ala -->Val]: a novel, electrophoretically silent, variant, Hemoglobin, 24 (2000) 143-148.
[58] F. de la Fuente-Gonzalo, M. Baiget, I. Badell, P. Ricard, L. Vinuesa, J. Martinez-Nieto, P. Ropero, A. Villegas, F.A. Gonzalez, J. Diaz-Mediavilla, G. Erythropathology Spanish, Study of three families

with Hb Agrinio [alpha29(B10)Leu-->Pro, CTG>CCG (alpha2)] in the Spanish population: three homozygous cases, Hemoglobin, 36 (2012) 526-532.
[59] E.V. De Marco, L. Crescibene, A. Pasqua, C. Brancati, M. Bria, A. Qualtieri, HB Prato [alpha 31(B12)Arg----Ser] in a Calabrian family, Hemoglobin, 16 (1992) 275-279.
[60] C. Prehu, E. Mazurier, J. Riou, J. Kister, D. Prome, S. Richelme-David, L. Al Jassem, E. Angellier, H. Wajcman, A new unstable alpha2-globin gene variant: Hb Chartres [alpha33(B14)Phe --> Ser], Hemoglobin, 27 (2003) 111-115.
[61] S. Yongsuwan, J. Svasti, S. Fucharoen, Decreased heat stability found in purified hemoglobin Queens [alpha 34(B15)Leu----Arg], Hemoglobin, 11 (1987) 567-570.
[62] Y.T. Zeng, S.Z. Huang, Y.K. Yen, W.C. Li, H. Lam, B. Webber, J.B. Wilson, T.H. Huisman, Hemoglobin Queens (alpha 34 (B15) Leu replaced by Arg) found in two Chinese families, Hemoglobin, 6 (1982) 209-211.
[63] S. Gomes, I. Picanco, A. Miranda, M.T. Seixas, M. Oliveira, L. Romao, P. Faustino, Hb Evora [alpha2-35 (B16), Ser-->Pro], a novel hemoglobin variant associated with an alpha-thalassemia phenotype, Haematologica, 92 (2007) 252-253.
[64] S.O. Brennan, C. Sheen, S. Johnson, Hb Manawatu [alpha 37(C2)Pro-->Leu]: a new mildly unstable mutation at an invariant proline residue, Hemoglobin, 26 (2002) 389-392.
[65] T. Harano, K. Harano, S. Shibata, S. Ueda, K. Imai, A. Tsuneshige, H. Yamada, M. Seki, H. Fukui, Hemoglobin Kariya [alpha 40 (C5) Lys leads to Glu]: a new hemoglobin variant with an increased oxygen affinity, FEBS Lett, 153 (1983) 332-334.
[66] Y. Ohba, K. Imai, R. Uenaka, M. Ami, K. Fujisawa, K. Itoh, K. Hirakawa, T. Miyaji, Hb Miyano or alpha 41(C6)Thr----Ser: a new high oxygen affinity alpha chain variant found in an erythremic blood donor, Hemoglobin, 13 (1989) 637-647.
[67] C. Prehu, J. Riou, H. Wajcman, Hb Barika [alpha42(C7)Tyr-->His (alpha2)] leads to an alpha+ -Thalassemia-like syndrome, Hemoglobin, 31 (2007) 17-22.
[68] A. Beretta, V. Prato, E. Gallo, H. Lehmann, Haemoglobin Torino--alpha-43 (CD1) phenylalanine replaced by valine, Nature, 217 (1968) 1016-1018.
[69] T. Harano, K. Harano, S. Ueda, S. Shibata, K. Imai, Y. Ohba, T. Shinohara, S. Horio, K. Nishioka, H. Shirotani, Hemoglobin Kawachi [alpha 44 (CE2) Pro leads to Arg]: a new hemoglobin variant of high oxygen affinity with amino acid substitution at alpha 1 beta 2 contact, Hemoglobin, 6 (1982) 43-49.
[70] J. Bardakdjian, J. Kister, H. Wajcman, P. Boulard, B. Bohn, Y. Blouquit, F. Galacteros, Hb Poitiers [alpha 45(CE3)His-->Asp]: a new hemoglobin variant with a two-fold increase in oxygen affinity, Hemoglobin, 18 (1994) 1-9.
[71] T. Nakatsuji, J.B. Wilson, T.H. Huisman, Hb Cordele alpha(2)47 (CE5)Asp----Ala beta 2. A mildly unstable variant observed in black twins, Hemoglobin, 8 (1984) 37-46.
[72] B. Brimhall, R.T. Jones, R.G. Schneider, T.S. Hosty, G. Tomlin, R. Atkins, Two new hemoglobins. Hemoglobin Alabama (beta39(C5)Gln leads to Lys) and hemoglobin Montgomery (alpha 48(CD 6) Leu leads to Arg), Biochim Biophys Acta, 379 (1975) 28-32.
[73] J.G. Szelenyi, M. Horanyi, J. Foldi, J. Hudacsek, L. Istvan, S.R. Hollan, A new hemoblogin variant in hungary: Hb Savaria - alpha 49 (CE7) Ser replace by Arg, Hemoglobin, 4 (1980) 27-38.
[74] E. Gallo, L. Pugliatti, G. Ricco, P.G. Pich, G. Pinna, U. Mazza, A case of haemoglobin J Sardegna- -thalassaemia double heterozygosis, Acta Haematol, 47 (1972) 311-320.
[75] K. Moradkhani, C. Prehu, J. Old, S. Henderson, V. Balamitsa, H.Y. Luo, M.C. Poon, D.H. Chui, H. Wajcman, G.P. Patrinos, Mutations in the paralogous human alpha-globin genes yielding identical hemoglobin variants, Ann Hematol, 88 (2009) 535-543.
[76] C.A. Reynolds, T.H. Huisman, Hemoglobin Russ or alpha-2-51-arg-beta-2, Biochim Biophys Acta, 130 (1966) 541-543.
[77] T.H. Huisman, V.P. Sydenstricker, Difference in gross structure of two electrophoretically identical 'minor' haemoglobin components, Nature, 193 (1962) 489-491.
[78] R. Alberti, G.M. Mariuzzi, L. Artibani, E. Bruni, L. Tentori, A new haemoglobin variant: J-Rovigo alpha 53 (E-2) alanine leads to aspartic acid, Biochim Biophys Acta, 342 (1974) 1-4.

[79] K. Imai, H. Morimoto, M. Kotani, S. Shibata, T. Miyaji, Studies on the function of abnormal hemoglobins. II. Oxygen equilibrium of abnormal hemoglobins: Shimonoseki, Ube II, Hikari, Gifu, and Agenogi, Biochim Biophys Acta, 200 (1970) 197-202.
[80] P. Lacan, M. Aubry, N. Couprie, A. Francina, Hb Gerland [alpha55(E4)Val-->Ala (alpha2)]: a new neutral alpha chain variant involving the alpha2 gene, Hemoglobin, 25 (2001) 417-420.
[81] A. Abramov, H. Lehmann, L. Robb, Hb Shaare Zedek (alpha 56 E5 Lys leads to Glu), FEBS Lett, 113 (1980) 235-237.
[82] T.N. Mehrotra, S.C. Gupta, R. Sinha, Haemoglobin norfolk in nepali gorkhas, Humangenetik, 27 (1975) 347-349.
[83] P.D. Pulsinelli, M.F. Perutz, R.L. Nagel, Structure of hemoglobin M Boston, a variant with a five-coordinated ferric heme, Proc Natl Acad Sci U S A, 70 (1973) 3870-3874.
[84] F. Dutly, J. Fehr, J.S. Goede, M. Morf, H. Troxler, H. Frischknecht, A new highly unstable alpha chain variant causing alpha(+)-thalassemia: Hb Zurich Albisrieden [alpha59(E8)Gly-->Arg (alpha2)], Hemoglobin, 28 (2004) 347-351.
[85] V.A. Spivak, T.P. Molchanova, N.V. Ermakov, N. Tokarev Yu, G. Martinez, J. Szelenyi, M. Horanyi, J. Foldi, S. Hollan, H. Kazieva, I.A. Shamov, A new hemoglobin variant: Hb Dagestan alpha 60(E9) Lys leads to Glu, Hemoglobin, 5 (1981) 133-138.
[86] P.C. Giordano, R. Fodde, R. Amons, J.E. Ploem, L.F. Bernini, Hb J-Anatolia [alpha 61(E10)Lys----Thr]: structural characterization and gene localization of a new alpha chain variant, Hemoglobin, 14 (1990) 119-128.
[87] J.B. Wilson, B.B. Webber, A. Kutlar, A.L. Reese, V.C. McKie, C.L. Lutcher, A.E. Felice, T.H. Huisman, Hb Evans or alpha 262(E11)Val----Met beta 2; an unstable hemoglobin causing a mild hemolytic anemia, Hemoglobin, 13 (1989) 557-566.
[88] J. Thillet, Y. Blouquit, F. Perrone, J. Rosa, Hemoglobin Pontoise alpha63 Ala replaced by Asp(E12). A new fast moving variant, Biochim Biophys Acta, 491 (1977) 16-22.
[89] Z.Q. Zhou, L.C. Chen, P.F. Chen, K.Q. Zhang, Y.H. Wang, Hemoglobin Hangzhou alpha 64(E13)Asp----Gly. A new variant found in China, Hemoglobin, 11 (1987) 31-33.
[90] S. Farashi, N. Faramarzi Garous, M. Ashki, S. Vakili, F. Zeinali, H. Imanian, A. Azarkeivan, H. Najmabadi, Hb Dartmouth (HBA2: c.200T>C): An alpha2-Globin Gene Associated with Hb H Disease in One Homozygous Patient, Hemoglobin, 39 (2015) 152-155.
[91] J.O. Kaufmann, M. Phylipsen, C. Neven, W. Huisman, P. van Delft, M. Bakker-Verweij, S.G. Arkesteijn, C.L. Harteveld, P.C. Giordano, Hb St. Truiden [alpha68(E17)Asn-->His] and Hb Westeinde [alpha125(H8)Leu-->Gln]: two new abnormalities of the alpha2-globin gene, Hemoglobin, 34 (2010) 439-444.
[92] P. Ferranti, A. Parlapiano, A. Malorni, P. Pucci, G. Marino, G. Cossu, L. Manca, B. Masala, Hemoglobin Ozieri: a new alpha-chain variant (alpha 71(E20)Ala-->Val). Characterization using FAB- and electrospray-mass spectrometric techniques, Biochim Biophys Acta, 1162 (1993) 203-208.
[93] S. Djoumessi, J. Rousseaux, J. Descamps, M. Goudemand, M. Dautrevaux, Hemoglobin Lille, alpha 2 [74(EF3) Asp replaced by Ala] beta 2, Hemoglobin, 5 (1981) 475-479.
[94] N. Burnichon, P. Lacan, M. Becchi, I. Zanella-Cleon, M. Aubry, M. Mowafy, N. Couprie, A. Francina, A new alpha chain hemoglobin variant: Hb Al-Hammadi Riyadh [alpha75(EF4)Asp-->Val (alpha2)], Hemoglobin, 30 (2006) 155-164.
[95] J.B. Shelton, J.R. Shelton, W.A. Schroeder, D.R. Powars, Hb Aztec or alpha 2 76 (EF5) Met----Thr beta 2 detection of a silent mutant by high performance liquid chromatography, Hemoglobin, 9 (1985) 325-332.
[96] C. Badens, D. Lena-Russo, P. Lacan, A. Francina, D. Prome, J. Riou, M. Geoffroy, T. Ayavou, J. Kister, F. Galacteros, H. Wajcman, Hb Toulon [alpha77(EF6)Pro-->His]: a new variant due to a mutation in the alpha2 gene found during measurement of glycated hemoglobin, Hemoglobin, 23 (1999) 367-371.
[97] J.B. Wilson, B.B. Webber, D. Plaseska, P.A. de Alarcon, S.K. McMillan, T.H. Huisman, Hb Davenport or alpha 2(78)(EF7)Asn----His beta 2, Hemoglobin, 14 (1990) 599-605.

[98] H. Wajcman, Y. Blouquit, A. Lahary, A.M. Soummer, P. Groff, J. Bardakdjian, C. Prehu, J. Riou, C. Godard, F. Galacteros, Three new neutral alpha chain variants: Hb Bois Guillaume [alpha 65(E14(Ala-->Val], Hb Mantes-la-Jolie [alpha 79(EF8)Ala-Thr], and Hb Mosella [alpha 111(G18)Ala-->Thr], Hemoglobin, 19 (1995) 281-286.
[99] J.G. Adams, 3rd, W.P. Winter, D.L. Rucknagel, H.H. Spencer, Biosynthesis of hemoglobin Ann Arbor: evidence for catabolic and feedback regulation, Science, 176 (1972) 1427-1429.
[100] P. Lacan, M. Moreau, M. Becchi, I. Zanella-Cleon, M. Aubry, J.J. Louis, N. Couprie, A. Francina, Two new hemoglobin variants: Hb Brem-sur-Mer [beta9(A6)Ser-->Tyr] and Hb Passy [alpha81(F2)Ser-->Pro (alpha2)], Hemoglobin, 29 (2005) 69-75.
[101] H. Wajcman, D. Prome, C. Prehu, C. Deon, J. Riou, J.C. Bouanga, I. Papassotiriou, A. Lahary, F. Galacteros, HB Les Andelys [alpha83(F4)LEU-->PRO]: a new moderately unstable variant, Hemoglobin, 22 (1998) 129-140.
[102] J.H. Crookston, H.A. Farquharson, D. Beale, H. Lehmann, Hemoglobin Etobicoke: alpha-84(F5) serine replaced by arginine, Can J Biochem, 47 (1969) 143-146.
[103] N. Fujiwara, T. Maekawa, G. Matsuda, Hemoglobin Atago (alpha2-85Tyr beta-2) a new abnormal human hemoglobin found in Nagasaki. Biochemical studies on hemoglobins and myoglobins. VI, Int J Protein Res, 3 (1971) 35-39.
[104] A. Knuth, W. Pribilla, H.R. Marti, K.H. Winterhalter, Hemoglobin Moabit: alpha 86 (F7) Leu leads to Arg: a new unstable abnormal hemoglobin, Acta Haematol, 61 (1979) 121-124.
[105] K. Nishikura, Y. Sugita, M. Nagai, Y. Yoneyama, Ethylisocyanide equilibria of hemoglobins M Iwate, M Boston, M Hyde Park, M Saskatoon, and M Milwaukee-I in half-ferric and fully reduced states, J Biol Chem, 250 (1975) 6679-6685.
[106] M.C. Perry, C. Head, V.F. Fairbanks, R.T. Jones, H. Taylor, V. Proud, Hemoglobin Columbia Missouri or alpha 2[88 (F9) Ala----Val]beta 2: a new high-oxygen-affinity hemoglobin that causes erythrocytosis, Mayo Clin Proc, 66 (1991) 5-10.
[107] D. Williamson, J.V. Langdown, T. Myles, C. Mason, J.S. Henthorn, S.C. Davies, Polycythaemia and microcytosis arising from the combination of a new high oxygen affinity haemoglobin (Hb luton, alpha 89 His-->Leu) and alpha thalassaemia trait, Br J Haematol, 82 (1992) 621-622.
[108] A. Villegas, R. Hojas, F.A. Gonzalez, P. Ropero, D.L. Carreno, Hb Clinico-Madrid II [alpha90(FG2)Lys-->Arg (alpha1)] and Hb El Escorial [alpha96(G3)Val-->Asp (alpha2)]: two new a chain variants found during a neonatal period study, Hemoglobin, 27 (2003) 185-190.
[109] A. Villegas, R. Hojas, N. Noguera, P. Ropero, J. Sanchez, F.A. Gonzalez, C. Benavente, D. Espinos, Hb Clinico-Madrid [alpha90(FG2)Lys-->Arg]: a new hemoglobin mutation in the alpha2-globin gene, Hemoglobin, 24 (2000) 341-345.
[110] S.O. Brennan, G.P. Tauro, W. Melrose, Haemoglobin Port Phillip alpha91 (FG3) Leu replaced by Pro, a new unstable haemoglobin, FEBS Lett, 81 (1977) 115-117.
[111] M.C. Botha, R. Stathopoulou, H. Lehmann, J.S. Rees, D. Plowman, A Hb J Cape Town homozygote--association of Hb J Cape Town and alpha-thalassaemia, FEBS Lett, 96 (1978) 331-334.
[112] G. Lacerra, R. Testa, M. De Angioletti, G. Schiliro, C. Carestia, Hb Bronte or alpha93(FG5)Val-->Gly: a new unstable variant of the alpha2-globin gene, associated with a mild alpha(+)-thalassemia phenotype, Hemoglobin, 27 (2003) 149-159.
[113] H.Y. Luo, I. Irving, J. Prior, E. Lim, S.H. Eung, T.P. Skelton, W.N. Erber, M.H. Steinberg, D.H. Chui, Hemoglobin Titusville, a low oxygen affinity variant hemoglobin, in a family of Northern European background, Am J Hematol, 77 (2004) 384-386.
[114] A. Das Gupta, P. Hariharan, M. Daruwalla, K. Sidhwa, R. Pawar, A. Nadkarni, Hemoglobin Titusville [alpha2 Codon 94 G>A]: A Rare Alpha Globin Chain Variant Causing Low Oxygen Saturation, Indian J Hematol Blood Transfus, 35 (2019) 593-595.
[115] W.W. de Jong, L.F. Bernini, P. Meera Khan, Haemoglobin Rampa: Alpha 95 Pro--Ser, Biochim Biophys Acta, 236 (1971) 197-200.
[116] J.D. Hoyer, E. Rachut, K.S. Kubik, R.T. Jones, G.R. Honig, L.N. Vida, V.F. Fairbank, Hb Rampa [alpha 95(G2)pro-->Ser (alpha 2)] in a family of European ancestry: DNA analysis confirms the CCG--

>TCG mutation at codon 95 of the alpha 2-globin gene; clinical and laboratory features, Hemoglobin, 26 (2002) 397-403.
[117] T. Harano, K. Harano, S. Uehara, K. Matsushita, Two new alpha chain variants: Hb Fuchu-I [alpha 72(EF1)His-->Tyr] and Hb Fuchu-II [alpha 97(G4)Asn-->His], Hemoglobin, 19 (1995) 389-395.
[118] R. Son, T. Higuchi, A. Mizuno, R. Koyamada, S. Okada, Y. Yamashiro, A Newly Characterized Hemoglobin Variant with a High Oxygen Affinity, Hb Fuchu-II, Presenting with Acute Myocardial Infarction, Intern Med, 55 (2016) 285-287.
[119] J.V. Langdown, R.J. Davidson, D. Williamson, A new alpha chain variant, Hb Turriff [alpha 99(G6)Lys----Glu]: the interference of abnormal hemoglobins in Hb A1c determination, Hemoglobin, 16 (1992) 11-17.
[120] J.H. Crookston, H.A. Farquharson, J.L. Kinderlerer, H.H.M. Lehmann, alpha 102(G9) serine replaced by arginine, Can J Biochem, 48 (1970) 911-914.
[121] G.V. Sciarratta, G. Ivaldi, G.L. Molaro, G. Sansone, M.L. Salkie, J.B. Wilson, A.L. Reese, T.H. Huisman, The characterization of hemoglobin Manitoba or alpha (2)102(G9)Ser----Arg beta 2 and hemoglobin Contaldo or alpha (2)103(G10)His----Arg beta 2 by high performance liquid chromatography, Hemoglobin, 8 (1984) 169-181.
[122] S.N. Khan, F.I. Butt, S. Riazuddin, R. Galanello, Hb Sallanches [alpha104(G11)Cys-->Tyr]: a rare alpha2-globin chain variant found in the homozygous state in three members of a Pakistani family, Hemoglobin, 24 (2000) 31-35.
[123] F. Morle, A. Francina, R. Ducrocq, H. Wajcman, C. Gonnet, N. Philippe, G. Souillet, J. Godet, A new alpha chain variant Hb Sallanches [alpha 2 104(G11) Cys-->Tyr] associated with HbH disease in one homozygous patient, Br J Haematol, 91 (1995) 608-611.
[124] C.L. Harteveld, F.G. Versteegh, P.J. Kok, I.H. van Rooijen-Nijdam, P. van Delft, P.C. Giordano, Hb Bleuland [alpha108(G15)Thr-->Asn, ACC-->AAC (alpha2)]: a new abnormal hemoglobin associated with a mild alpha-thalassemia phenotype, Hemoglobin, 30 (2006) 349-354.
[125] I. Weiss, F.E. Cash, M.B. Coleman, A. Pressley, J.G. Adams, T. Sanguansermsri, S.A. Liebhaber, M.H. Steinberg, Molecular basis for alpha-thalassemia associated with the structural mutant hemoglobin Suan-Dok (alpha 2 109leu----arg), Blood, 76 (1990) 2630-2636.
[126] G.R. Honig, M. Shamsuddin, R. Zaizov, M. Steinherz, I. Solar, C. Kirschmann, Hemoglobin Petah Tikva (alpha 110 ala replaced by asp): a new unstable variant with alpha-thalassemia-like expression, Blood, 57 (1981) 705-711.
[127] E.G. Kazanetz, J.Y. Leonova, J.B. Wilson, S.K. McMillan, M. Walbrecht, J.M. de Pablos Gallego, T.H. Huisman, Hb Anamosa or alpha 2(111)(G18)Ala-->Val beta 2 (alpha 2 mutation) and Hb Mulhacen or alpha 2(123)(H6)Ala-->Ser beta 2 (alpha 1 mutation) are two silent, stable variants detected by sequencing of amplified DNA, Hemoglobin, 19 (1995) 1-6.
[128] S. Ayala, D. Colomer, A. Pujades, M. Aymerich, J.L. Vives Corrons, Haemoglobin Lleida: a new alpha 2-globin variant (12 bp deletion) with mild thalassaemic phenotype, Br J Haematol, 94 (1996) 639-644.
[129] M. Guis, W.C. Mentzer, D.L. Jue, M.H. Johnson, J.E. McGuffey, W.F. Moo-Pen, Hemoglobin Twin Peaks alpha 113 (GH1) Leu----His, Hemoglobin, 9 (1985) 175-177.
[130] H. Wajcman, J. Delaunay, A. Francina, J. Rosa, F. Galacteros, Hemoglobin Nouakchott [alpha 114(GH2)Pro----Leu]: a new hemoglobin variant displaying an unusual increase in hydrophobicity, Biochim Biophys Acta, 998 (1989) 25-31.
[131] D.C. Gajdusek, J. Guiart, R.L. Kirk, R.W. Carrell, D. Irvine, P.A. Kynoch, H. Lehmann, Haemoglobin J Tongariki (alpha-115 alanine--aspartic acid): the first new haemoglobin variant found in a Pacific (Melanesian) population, J Med Genet, 4 (1967) 1-6.
[132] R.G. Schneider, B. Hightower, U. Carpentieri, M.L. Duerst, T. Shih, R.T. Jones, Hemoglobin Oleander [alpha 116(GH4)Glu replaced by Gln beta 2]: structural and functional characterization, Hemoglobin, 6 (1982) 465-480.
[133] T. Harano, K. Harano, K. Imai, H. Yunoki, H. Yagi, K. Nagashima, T. Kuroume, Hb J-Meerut [alpha 120(H3)Ala----Glu] found in a Japanese family, Hemoglobin, 13 (1989) 169-175.

[134] T. Harano, K. Harano, S. Ueda, Hb Owari [alpha 121 (H 4) Val----Met]: a new hemoglobin variant with a neutral-to-neutral amino acid substitution detected by isoelectric focusing, Hemoglobin, 10 (1986) 127-134.
[135] G. Martin, A. Villegas, F.A. Gonzalez, P. Ropero, R. Hojas, M. Polo, M. Mateo, M. Salvador, C. Benavente, A novel mutation of the alpha2-globin causing alpha(+)-thalassemia: Hb Plasencia [alpha125(H8)Leu--Arg (alpha2), Hemoglobin, 29 (2005) 113-117.
[136] B. Ibarra, G. Vaca, J.M. Cantu, J.B. Wilson, H. Lam, M. Stallings, M.E. Gravely, T.H. Huisman, Heterozygosity and homozygosity for the high oxygen affinity hemoglobin Tarrant or alpha 126 (H9) Asp replaced by Asn in two Mexican families, Hemoglobin, 5 (1981) 337-348.
[137] W.F. Moo-Penn, K.C. Bechtel, M.H. Johnson, D.L. Jue, S. Holland, C. Huff, R.M. Schmidt, Hemoglobin Jackson, alpha 127 (H10) Lys replaced by Asn, Am J Clin Pathol, 66 (1976) 453-456.
[138] C.L. Harteveld, P.C. Giordano, M. Losekoot, J.G. Heister, D. Batelaan, P. van Delft, M.C. Bruin, L.F. Bernini, Hb Utrecht [alpha 2 129(H12)Leu-->Pro], a new unstable alpha 2-chain variant associated with a mild alpha-thalassaemic phenotype, Br J Haematol, 94 (1996) 483-485.
[139] M. Harkness, D.R. Harkness, F. Kutlar, A. Kutlar, J.B. Wilson, B.B. Webber, J.F. Codrington, T.H. Huisman, Hb Sun Prairie or alpha(2)130(H13)Ala----Pro beta 2, a new unstable variant occurring in low quantities, Hemoglobin, 14 (1990) 479-489.
[140] H. Wajcman, C. Vasseur, Y. Blouquit, J. Rosa, D. Labie, A. Najman, O. Reman, M. Leporrier, F. Galacteros, Unstable alpha-chain hemoglobin variants with factitious beta-thalassemia biosynthetic ratio: Hb Questembert (alpha 131[H14]Ser-->Pro) and Hb Caen (alpha 132[H15]Val-->Gly), Am J Hematol, 42 (1993) 367-374.
[141] J.V. Okumura, E.L. Shimauti, D.G. Silva, L.S. Torres, E. Belini-Junior, R.G. Oliveira, E.V. Patussi, J.C. Herrero, C.R. Bonini-Domingos, Hemoglobin (Hb) Val de Marne (Hb Footscray) in Brazil: the first case report, Genet Mol Res, 15 (2016).
[142] H. Wajcman, Y. Blouquit, J. Riou, J. Kister, C. Poyart, J. Soria, F. Galacteros, A new hemoglobin variant found during investigations of diabetes mellitus: Hb Pavie [alpha 135 (H18) Val----Glu], Clin Chim Acta, 188 (1990) 39-47.
[143] M.J. McDonald, L.A. Michalski, S.M. Turci, R.A. Guillette, D.L. Jue, M.H. Johnson, W.F. Moo-Penn, Structural, functional, and subunit assembly properties of hemoglobin Attleboro [alpha 138 (H21) Ser----Pro], a variant possessing a site maturation at a critical C-terminal residue, Biochemistry, 29 (1990) 173-178.
[144] S. Rahbar, T.D. Lee, M. Davis, W.F. Novotny, H.M. Ranney, A second case of Hb Hanamaki [alpha 2 139(HC1)Lys->Glu beta 2] in an American family with erythrocytosis, Hemoglobin, 18 (1994) 221-226.
[145] H. Wajcman, J. Kister, M. Marden, A. Lahary, M. Monconduit, F. Galacteros, Hemoglobin Rouen (alpha-140 (HC2) Tyr-->His): alteration of the alpha-chain C-terminal region and moderate increase in oxygen affinity, Biochim Biophys Acta, 1180 (1992) 53-57.
[146] M.E. Gravely, H.F. Harris, M. Stallings, H. Lam, J.B. Wilson, T.H. Huisman, Hb Suresnes or alpha2 141(HC3) ArgyieldHis beta2 in a black family, Hemoglobin, 2 (1978) 187-189.