**Mutations affecting the function of hemoglobin**

**Instructions: Complete the worksheet (it is suggested to type in your answers in the provided spaces). This is an INDIVIDUAL assignment. Having access to your previous worksheet will facilitate this worksheet.**

**Learning objectives:**

At the end of this worksheet, students should be able to:

* Align and identify the amino acid and nucleic acid alterations in the sickle cell hemoglobin (HbS) sequence
* Perform Multiple Sequence Alignments (MSA) on protein and nucleic acid sequences
* Interpret the nature of the sickle cell mutation

About 7% of the world’s population carries one or more mutations involved in hemoglobin biosynthesis or function. There are two basic classifications of hemoglobin mutants - thalassemias and hemoglobin variants.

**Thalassemias** are characterized by an imbalance of globin synthesis, where the mature protein (consisting of four subunits, normally two α-globin and two ß-globin in adults), is composed of a different subset of polypeptides. The functional hemoglobin protein is always composed of four subunits, but in thalassemias, there may be 0, 1, 3, or 4 of either subunit (always being composed of four). Because of the co-operativity of O2 binding in the subunits of the normal tetramer, changes in subunit balance affects the ability to bind O2.

**Hemoglobin variants** are affected by a change in the amino acid sequence of a globin polypeptide. There are more than 400 known variants of hemoglobin - some have minimal effect on the function of the polypeptide and may go undetected throughout life, while others may have a significant effect on the function of hemoglobin. Some mutations have been demonstrated to be lethal. One mutation that affects the function of hemoglobin is the mutation causing sickle cell anemia.

**Sickle cell anemia**

The hemoglobin variant for sickle cell anemia is composed of two normal polypeptide subunits (one type of globin) and one (heterozygous) or two mutant (homozygous) versions of the second globin polypeptide. These polypeptides are encoded by separate genes. Sickle cell anemia is an autosomal recessive mutation. The mutation occurs in about 1 in 400-600 African-Americans and is more common in parts of Africa. Sickle cell disease is characterized by weakness, abdominal pain, and kidney and heart failure, with individuals especially affected during times of low oxygen. In general, individuals that are heterozygous for the mutation only demonstrate these symptoms during severe low oxygen conditions. The mutation also confers an increased resistance to malaria in heterozygous individuals, indicating a possible reason why the mutation has persisted in the population.

1. **What is the nature of the mutation causing sickle cell anemia?**
	1. The sequence of the polypeptide affected by the sickle cell mutation is below:

|  |
| --- |
| >HbS\_polypeptideMVHLTPVEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH |

* + 1. The HbS designation is the provided name for the encoded polypeptide sequence of the sickle cell allele. *Previously (worksheet 1) you identified HbA1 and HbB.*
	1. Compare the polypeptide sequences for HbA1, HbB (from worksheet 1) and HbS using the ClustalOmega program (http://www.ebi.ac.uk/Tools/msa/clustalo/).
		1. Copy and paste the resulting aligned sequences below as previously done (reformat if necessary).

|  |
| --- |
| HbA1\_polypeptide -MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHFD------LSHGSHbB\_polypeptide MVHLTPEEKSAVTALWGKVNV--DEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNHbS MVHLTPVEKSAVTALWGKVNV--DEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGN : \*:\* :\*: \*.\* \*\*\*\* . \* \*.\*\*\* \*::: :\* \*: :\* \* \*.HbA1\_polypeptide AQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAAHHbB\_polypeptide PKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHHbS PKVKAHGKKVLGAFSDGLAHLDNLKGTFATLSELHCDKLHVDPENFRLLGNVLVCVLAHH :\*\*.\*\*\*\*\* \*:::.:\*\*:\*:: ::::\*\*:\*\*..\*\*:\*\*\* \*\*:\*\*.. \*: .\*\* \*HbA1\_polypeptide LPAEFTPAVHASLDKFLASVSTVLTSKYRHbB\_polypeptide FGKEFTPPVQAAYQKVVAGVANALAHKYHHbS FGKEFTPPVQAAYQKVVAGVANALAHKYH : \*\*\*\* \*:\*: :\*.:\*.\*:..\*: \*\*: |

|  |  |
| --- | --- |
| Which polypeptide (HbA1 or HbB) looks most like the HbS polypeptide sequence? | HbB |
| What is the amino acid difference between HbS and the most similar globin polypeptide?[[1]](#footnote-1)  | E7V or E6V (see note below) |

*In hemoglobin, the N-terminal methionine is commonly removed from the polypeptide, so the position of the amino acid mutation is often listed as one amino acid position shorter than the translated coding sequence.*

Box: *Amino acid changes are indicated as the single amino acid letter of the ‘wild-type’ version followed by the number representing the amino acid in the chain and the ‘mutant’ amino acid. For example, if the second amino acid is changes from valine (V) to histidine (H), this would be written as V2H.*

* 1. What is the corresponding DNA base change in the coding sequence?
		1. Using the HbS mRNA sequence below, determine the nature of the nucleotide base change in HbS compared with the corresponding ‘wild-type’ globin mRNA sequence.

|  |
| --- |
| >HbS\_cdsATGGTGCATCTGACTCCTGTGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCACCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCACTAA |

* + 1. Copy the FASTA sequence for the corresponding globin gene from Worksheet 1 (that matches the HbS sequence except for the mutation) and paste into the ClustalOmega window. Make sure to select ‘DNA’ instead of the default ‘Protein’ as the entry type. Copy and paste the HbS\_cds sequence into the same window. *Note that the HbS sequence provided is just the coding sequence, not the full length mRNA as previously done.* Align two sequences (HbS and the matching globin mRNA sequence).
		2. Copy and paste the aligned sequences in the box provided (reformat and expand if necessary).

|  |
| --- |
| HbB\_nt ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCHbS\_cds --------------------------------------------------ATGGTGCATC \*\*\*\*\*\*\*\*\*\*HbB\_nt TGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGHbS\_cds TGACTCCTGTGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAG \*\*\*\*\*\*\*\*\* \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*HbB\_nt TTGGTGGTGAGGCCCTGGGCAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGHbS\_cds TTGGTGGTGAGGCCCTGGGCAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTG \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*HbB\_nt AGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCHbS\_cds AGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTC \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*HbB\_nt ATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGHbS\_cds ATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGG \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*HbB\_nt GCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACTHbS\_cds GCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACT \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*HbB\_nt TCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCAHbS\_cds TCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCA \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*HbB\_nt CCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCHbS\_cds CCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCC \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* |

* + 1. Nucleotide matches are shown with an \* below the sequence in ClustalOmega. *Hint: To answer the first question below, look for the missing \*.*

|  |  |
| --- | --- |
| What do the dashes in the alignment represent? | Dashes represent regions of only one strand showing sequence |
| What nucleotide change do you observe? *Write as ‘wild-type’ base > ‘mutant base’.* | A>T |
| Is this a point mutation? | Yes |
| If this is a point mutation, is it a **silent** mutation, a **missense** mutation or a **nonsense** mutation? *You may have to remind yourself of these terms.* | Missense |
| What type of base change is observed? *(Choose from transition and transversion)* | Transversion |

* + 1. How might the amino acid change (*you will need to think about the nature of the amino acid side chains)* in HbS versus the ‘normal’ subunit affect the
			1. Tertiary polypeptide structure?

|  |
| --- |
| The amino acid change alters the chemical properties of that region of the surface of the polypeptide (the single amino acid change does not lead to refolding of the polypeptide). The hydrophobic amino acid (valine) replaces the hydrophilic glutamic acid.  |

* + - 1. Quaternary protein structure?

|  |
| --- |
| The mutation leads to a change in the surface of the polypeptide, but does not alter the quaternary structure. The hydrophobic patch created leads to the tetramer associating with other tetramers to form the long chains of tetramers found in sickle cell anemia. |

1. [↑](#footnote-ref-1)