

## Key Teaching Points for the Map of the Human $\beta$ -Globin Gene<sup>©</sup>

**Overall Student Learning Objective: The linear sequence of amino acids in a protein is determined by the linear sequence of nucleotides in a gene.**

- The blueprint for building proteins is found in DNA.
- The sequence of DNA that codes for a specific protein is called a **gene**.
- Humans have 46 chromosomes; each chromosome consists of many genes.
- Each person inherits 23 chromosomes from each of their parents. **Homologous chromosomes**, one from each parent, contain genes for the same types of proteins.
- The two copies of a gene, found on homologous chromosomes, may vary slightly. These variations of a gene are called **alleles**.
- The building blocks of proteins are amino acids. There are 20 different amino acids found in proteins.
- The building blocks of DNA are nucleic acids. There are 4 different nucleic acids found in DNA.
- Chromosomes are found in the nucleus. Protein synthesis occurs in the cytoplasm. DNA is used as a template to make **messenger RNA (mRNA)** in the nucleus. The mRNA is then transported to the cytoplasm where the sequence of nucleotides is interpreted by transfer RNA molecules on the ribosome to build protein.
- Nature uses a triplet nucleic acid sequence to code for a single amino acid. This **genetic code** has common characteristics for all organisms.
  - Each triplet sequence of mRNA is called a **codon**.
  - Every gene starts with the same codon, AUG, though this codon can also occur elsewhere in the protein sequence.
  - There are no spacers between codons in the sequence.
  - Some amino acids have multiple codons.
  - There are three **stop codons**: UAA, UAG, and UGA.
- Because the genetic code is triplet, there are three **forward reading frames** on a strand of DNA.
- Eukaryotic genes have gaps, called **introns**, which must be removed from the mRNA before the protein is made. The number of introns, and their length, varies with different genes. Errors in removing introns can lead to **splice site mutations**.
- In addition to the coding sequence, genes contain regulatory regions, including the **promoter** and **polyadenylation signal**.
- Changes in the gene (**mutations**) may result in changes in the protein sequence.

★ **We recommend** that you introduce the  $\beta$ -globin gene map after talking about DNA structure, DNA replication, transcription and translation, but before introducing introns or promoters. Depending on the level of your students and your learning objectives, you may wish to present only some of the information that can be taught using this kit.

## Genes Code for Proteins

---

Provide groups of students (2-3 is best) with a student version of the  $\beta$ -globin gene map, a dry erase marker or highlighter, and the  $\beta$ -globin protein sequence. You may also wish to provide a codon chart. Ask them to find the protein sequence and highlight it on the gene strip.

★ **We suggest** you answer questions as they arise; if students don't ask questions, you may ask them the questions to ensure that they develop an understanding of the concepts. If they can't answer the first question in a series, sub-questions may be used to guide them to an understanding.

1. What does the top red sequence represent?
  - a. What different letters are found in the red sequence?
  - b. What molecule do you know that contains the letters A, G, C and T?
2. What does the black sequence represent? [complementary strand of DNA]
  - a. How does the black sequence compare with the red sequence?
3. What do the three blue strands represent? [3 forward reading frames]
4. Why are there three blue strands? [Genetic code is a triplet code – so you can begin reading on the first, second or third base. If you begin reading at the fourth base, it is the same code as reading beginning at the first base.]
5. What do the asterisks in the blue strands represent? [Stop codons; you can have students verify this by reading the DNA sequence and looking up the RNA sequence on the genetic code chart.]

Students will find the first part of the protein sequence in the third reading frame, but they should discover that the protein sequence and the gene strip sequence converge. They have discovered the first **intron** in the  $\beta$ -globin gene. Encourage them to look for the rest of the protein sequence.

**Common pitfalls:** Some students may “miss” the intron, because once they find the start of the protein sequence, they simply highlight the rest of the strip. Other students may not be able to find the rest of the protein because they will look only on the reading frame that the first exon is in. Because introns are cut out *before* the sequence is read, they can vary in length. Therefore, exons of a single gene are not necessarily found in a single reading frame. The  $\beta$ -globin gene has three exons, and each exon is in a different reading frame.

If you want to dig a little deeper, you can have students explore the structure of the intron. By comparing the conserved intron splice site sequences and highlighting the DNA that corresponds with each exon, they will discover that the first intron actually cuts in the middle of the codon for arginine; the first two letters of the codon, AG, are in the first exon, and the last letter of the codon, G, is in the second exon.

## Regulatory Sequences

---

The Map of the Human  $\beta$ -Globin Gene<sup>®</sup> includes regulatory regions both 5' of the gene sequence (TATA box, CAAT box and erythroid-specific promoters) and 3' of the gene (poly-adenylation sequences). Students can compare the actual sequences in these regions with the consensus sequences to predict whether this gene will be readily transcribed (close to the consensus), or if additional factors are necessary for gene transcription (not similar to the consensus). See Teacher Notes for a detailed discussion of these topics.

## Mutations

---

The gene map can also be used to explore mutations. The sickle cell mutation codes for a hydrophobic valine residue instead of the normal polar glutamic acid at position 6 of the protein sequence. This residue is located on the surface of the protein. The hydrophobic amino acid on the surface of the protein will cause the hemoglobin proteins to clump, forming a curved chain. These curved chains of proteins produce the characteristic sickle shape found in sickle cell anemia.

In addition, the Teacher Map includes mutations leading to various  $\beta$ -thalassemia diseases. These mutations include missense mutations, splice site mutations and a frameshift mutation. More information about these mutations can be found in Features of the Teacher's Map on the CD that accompanies the Teacher Map.

**! Caution – Note This Limitation in the Model** - The Map of the Human  $\beta$ -Globin Gene<sup>®</sup> includes only the DNA and protein sequences and omits the mRNA sequence. The focus of this gene map is on removing introns; the Insulin mRNA to Protein Kit<sup>®</sup> begins with mRNA and includes processing and folding of the insulin protein.



**For more detailed lesson plans and activities, please visit**

<http://www.3dmoleculardesigns.com/Education-Products/Map-of-the-Human-Globin-Gene.htm>

**The Map of the Human  $\beta$ -Globin Gene<sup>®</sup>** can be borrowed from the MSOE Model Lending Library (<http://cbm.msoe.edu/teachRes/library>) or purchased from 3D Molecular Designs ([www.3dmoleculardesigns.com](http://www.3dmoleculardesigns.com)).