Current Level

Seq Editor

Searches

Seq. Comparison

Alignment

Restriction Map

Translation & ORF

Reverse Comp

Primer Design

Splice Sites

Genomics

Protein Motifs

Protein structure

Transmembrane

Previous Level

Translation and Open Reading Frames

Translation and Open Reading Frame Search

Regions of DNA that encode proteins are first **transcribed** into messenger RNA and then **translated** into protein. By examining the DNA sequence alone we can determine the sequence of amino acids that will appear in the final protein. In translation codons of three nucleotides determine which amino acid will be added next in the growing protein chain. It is important then to decide which nucleotide to start translation, and when to stop, this is called an **open reading frame**.

Once a gene has been sequenced it is important to determine the correct **open reading frame** (ORF). Every region of DNA has six possible **reading frames**, three in each direction. The reading frame that is used determines which amino acids will be encoded by a gene. Typically only one reading frame is used in translating a gene (in eukaryotes), and this is often the longest open reading frame. Once the open reading frame is known the DNA sequence can be translated into its corresponding amino acid sequence. An open reading frame starts with an atg (Met) in most species and ends with a stop codon (taa, tag or tga).

For example, the following sequence of DNA can be read in six reading frames. Three in the forward and three in the reverse direction. The three reading frames in the forward direction are shown with the translated amino acids below each DNA seqeunce. Frame **1** starts with the "a", Frame **2** with the "t" and Frame **3** with the "g". Stop codons are indicated by an " *" in the protein sequence. The longest ORF is in Frame **1**.

```
   5'                                                           3'
   atgcccaagctgaatagcgtagagggggttttcatcatttgaggacgatgtataa

 1 atg ccc aag ctg aat agc gta gag ggg ttt tca tca ttt gag gac gat gta taa
    M   P   K   L   N   S   V   E   G   F   S   S   F   E   D   D   V   *
 2 tgc cca agc tga ata gcg tag agg ggt ttt cat cat ttg agg acg atg tat
    C   P   S   *   I   A   *   R   G   F   H   H   L   R   T   M   Y
 3  gcc caa gct gaa tag cgt aga ggg gtt ttc atc att tga gga cga tgt ata
    A   Q   A   E   *   R   R   G   V   F   I   I   *   G   R   C   I
```

<u>Practice Problems</u> and a <u>Translation Lecture Tutorial</u> are located in the Theory Section.

DNA Translation.

To translate a DNA sequence, we use the program called SIXFRAME on the <u>Biology Workbench</u>. Or you can visit the site directly <u>http://searchlauncher.bcm.tmc.edu/seq-util/Options/sixframe.html</u>

1. Enter Biology Workbench and Resume or Create a New Session.

2. Select Nucleic Tools and then Select the sequence that you wish to translate by checking the box next to the file.

3. Select SIXFRAME and then select Run.

4. You will be given a box asking you to select various parameters. The default is to translate all six reading frames of the entire DNA sequence. If you select ``Show longest open reading frame" the program will automatically select the longest reading frame starting with a start codon (ATG) and ending with a stop codon (TAA, TAG or TGA). This is very handy.

5. Once the parameters are chosen select Submit.

Translation Parameters

	Start Translating at:	Stop Translating at:	Frame to Translate:	Sequence Label:
Sequence 1:	1	1395	All 6 Frames ▼	antithrombin

☑ **Show longest open reading frame**

☐ **Show sequence (or fragment) being translated**

| Submit | Reset | Abort | Help | Report Bugs |

☐ **Run as batch**

Interpretation of Results

1. All six reading frames will be shown. Stop codons will be indicated by an asterix (*).
2. At the bottom of the page you will see the ``Longest ORF". To import this translation of your DNA, click off all of the other boxes next to each reading frame and leave the box next to the longest ORF selected. Then select Import Sequence.
3. A new file with the translated amino acid sequence will appear under Protein Tools.

☑ **Frame "Longest ORF", 0 stop codon**s

antithrombin Translated - Longest ORF

```
>gi_4502260_ref_NM_000488.1_  Homo sapiens antithrombin III (AT3), mRNA
MYSNVIGTVTSGKRKVYLLSLLLIGFWDCVTCHGSPVDICTAKPRDIPMNPMCIYRSPEKKATEDEGSEQ
KIPEATNRRVWELSKANSRFATTFYQHLADSKNDNDNIFLSPLSISTAFAMTKLGACNDTLQQLMEVFKF
DTISEKTSDQIHFFFAKLNCRLYRKANKSSKLVSANRLFGDKSLTFNETYQDISELVYGAKLQPLDFKEN
AEQSRAAINKWVSNKTEGRITDVIPSEAINELTVLVLVNTIYFKGLWKSKFSPENTRKELFYKADGESCS
```
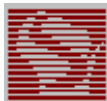
| Import Sequence(s) | Return | Change Settings | Help | Report Bugs |

Click here to email comments to Scott Cooper regarding this site or its links.