**Bioinformatics Worksheet #2: BLAST Practice - Assignment**

**Group Number:\_\_\_\_\_5\_\_\_\_\_\_\_\_**

**Names of individuals in group:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Mystery bacterial operon #: Operon 5**

**Table 1: Table of results for 8 largest proteins in operon**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Protein #** | **Length (base pairs)** | **Length (amino acids)** | **E-value** | **Query Cover** | **% Identity** | **Protein name – brief description** |
| 1 | 1503 | 500 | 0.0 | 100 | 100 | [alpha-N-arabinofuranosidase](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi#alnHdr_489322191) |
| 2 | 1320 | 439 | 0.0 | 100 | 100 | [AraM protein](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi#alnHdr_1913927) |
| 3 | 1185 | 394 | 0.0 | 100 | 100 | [arabinose-binding protein](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi#alnHdr_489322187) |
| 4 | 942 | 313 | 0.0 | 100 | 100 | [putative integral membrane protein](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi#alnHdr_1913929) |
| 5 | 846 | 281 | 0.0 | 100 | 100 | [MULTISPECIES: arabinose transporter permease](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi#alnHdr_489322189) |
| 6 | 810 | 269 | 0.0 | 100 | 100 | [haloacid dehalogenase](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi#alnHdr_489322183) |

1. What organism did this operon come from? What evidence did you use to come to this conclusion?

**Bacillus subtilis,** [**http://www.ncbi.nlm.nih.gov/nuccore/X89810.1**](http://www.ncbi.nlm.nih.gov/nuccore/X89810.1)

**Evidence from the fact that all of the best matching proteins were from this organism.**

1. Are the functions of the genes found in this operon related? What cellular system do you think this operon and the genes contained are involved in?

**Yes, the functions are related. They are involved in the breakdown (catabolism) of L-arabinose (a sugar, breakdown product of glucose) to 5-xylose 6 phosphate, from Wikipedia,**

**“Repression[[edit](http://en.wikipedia.org/w/index.php?title=L-arabinose_operon&action=edit&section=3" \o "Edit section: Repression)]**

**The ara operon is regulated by the AraC**[**protein**](http://en.wikipedia.org/wiki/Protein)**. If arabinose is absent, the**[**dimer**](http://en.wikipedia.org/wiki/Protein_dimer)**AraC protein represses the structural gene by binding to araI1 and araO2 and the**[**DNA**](http://en.wikipedia.org/wiki/DNA)**forms a loop. The loop prevents**[**RNA polymerase**](http://en.wikipedia.org/wiki/RNA_polymerase)**from binding to the**[**promoter**](http://en.wikipedia.org/wiki/Promoter_%28genetics%29)**of the ara operon, thereby blocking**[**transcription**](http://en.wikipedia.org/wiki/Transcription_%28genetics%29)**.**

**Activation[[edit](http://en.wikipedia.org/w/index.php?title=L-arabinose_operon&action=edit&section=4" \o "Edit section: Activation)]**

**When arabinose is present, arabinose binds AraC and prevents AraC from interacting. This breaks the DNA loop. The two AraC-arabinose complexes bind to the araI1 and araI2sites which promotes transcription. When arabinose is present, AraC acts as an activator.”**

1. Do some research using NCBI (<http://www.ncbi.nlm.nih.gov>) and other internet resources to learn about the system. What role does this system have in this organism? Create a hypothesis about what might happen to the organism if these genes were deactivated.

**Catabolism of L-arabinose to 5-xylose 6 phosphate as a part of the pentose phosphate pathway.**

**Hypothesis/Prediction: If this protein was mutated or broken, then the pentose phosphate pathway would be disrupted, forcing the use of glycolysis, or rather forcing the bacteria to only use glycolysis rather than the pentose phosphate alternative.**

1. Share what you learned with the other groups and write down what you learned about the other systems.

**Various answers**

**Bonus**: Go to RCSB Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>) and search for one of the proteins you found by name. You may not be able to find it for your organism, or at all, but here you can see the structure of the protein determined by protein crystallography. Count the number of beta pleatings and alpha helixes. Are there any ligands associated with this protein? Are there any other units for this protein? If so, name them. Be sure to look at the protein in the 3-D view.

**Name of protein you investigated**:\_\_  L-Arabinose Isomerase\_\_\_\_\_\_\_\_\_\_\_\_\_

**Number of Beta-pleated sections**:\_\_\_\_\_\_\_23\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Number of alpha helix sections**:\_\_\_\_\_\_18\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Associated ligands?**\_\_\_\_\_\_\_\_\_\_\_\_\_\_none\_\_\_\_\_\_\_\_\_\_\_\_\_

**Additional subunits for protein?\_**\_\_\_\_\_\_\_none\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Provide a rough 3-D sketch of the protein