**Accession Number Exercise**

Learning Goals:

* Become familiar with searching using accession numbers.
* Become familiar with the information present in the NCBI protein databases.

Go to the NCBI website at [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)

In the dropdown menu, next to the NCBI logo, select protein.



In the search bar enter the accession number for the protein entry: KAB4445054.1

You will be directed to a new page with a flat file with a header that looks something like this:



Explore the file to answer the questions below.

1. How long is the protein?

1004 amino acids long



1. What is the academic citation (Title, Journal and Authors) associated with this

sequence?

Poyet,M., Groussin,M., Gibbons,S.M., Avila-Pacheco,J., Jiang,X., Kearney,S.M., Perrotta,A.R., Berdy,B., Zhao,S., Lieberman,T.D., Swanson,P.K., Smith,M., Roesemann,S., Alexander,J.E., Rich,S.A.,Livny,J., Vlamakis,H., Clish,C., Bullock,K., Deik,A., Scott,J.,Pierce,K.A., Xavier,R.J. and Alm,E.J. A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. Nat. Med. 25 (9), 1442-1452 (2019)



1. Can you identify any conserved domains? (*Hint: see the options on the right-hand side of the page under ‘Analyze this sequence’*)



Conserved domains in this protein include a TonB-linked outer membrane protein

1. Can you view the co-ordinates of the conserved domains with reference to the entire sequence? (*Hint: click the ‘Graphics’ options under the protein name*)

TonB-linked outer membrane protein: 29…1002



1. What other information can you glean from this page that might be useful?

Identify identical proteins can be identified by clicking on the “Identical Proteins” options under the protein name. Additional information on this protein product and its gene can be obtained from the options under “Related Information” on the right-hand side of the page

