

Exercise 1 – InterPro entry page

InterPro entry pages provide a wealth of information about the entity (protein family, domain, site, etc.) described by that entry.

 Open the InterPro homepage (<https://www.ebi.ac.uk/interpro/>) in a web browser.

 Using the “Text” Search box at the top right-hand side of the page, type in the InterPro accession *IPR026017* (alternatively just type 26017). Click on the purple “Search” button. You should now have the InterPro entry page for “Lumazine-binding domain” open. Read through the description of the domain provided on the Overview page.

 **Question 1:** Which member database signature(s) are integrated to form this InterPro entry?

 **Question 2:** What size is the lumazine-binding domain described by this entry?

 **Question 3:** Where specifically within the structure of the domain are the substrate binding sites found?

 **Question 4:** How many domain architectures exist in the proteins matched by this entry?

 **Question 5:** Describe the most common domain architecture.

 **Question 6:** Are proteins containing this domain involved in any known pathways?

 **Question 7:** Is this domain ever found in plants?

 It is also possible to search InterPro for entries of interest based on a word that might be found in the name or description text. Using the “Text” Search box at the top right-hand side of the page, type in *photoproduct*. Select the InterPro entry titled *Spore photoproduct lyase (IPR023897)*.

 **Question 8:** How many proteins are matched by this InterPro entry?

 **Question 9:** This entry refers to an enzymatic family. What is the EC (enzyme commission) number for this enzyme?

 **Hint:** Have a look at cross-references.

 **Question 10:** Are there any known solved structures for members of this family?

 **Question 11:** What domain do proteins of this family contain?

Exercise 2 – InterPro protein page

UniProt accession query

 Open the InterPro front page <https://www.ebi.ac.uk/interpro/>

Enter the UniProt accession ‘O00308’ into the InterPro search box on the top right-hand side of the page.

 **Question 12:** To which family does InterPro predict this protein belongs?

 **Question 13:** How many domains does InterPro predict the protein to possess?

 **Question 14:** What is the name of the domain at the N terminus? What cellular functions are this domain involved in?

 **Question 15:** How many homologous superfamily’s does this protein match?