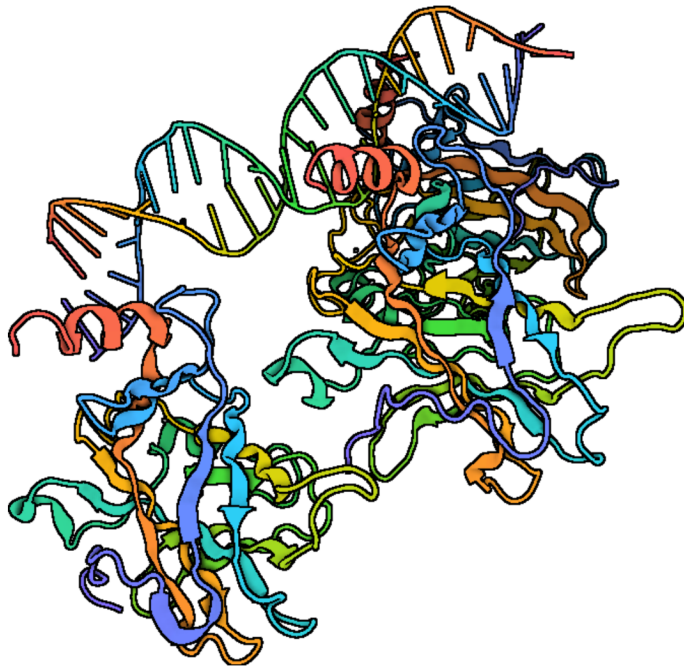


Word for Word

Where to Find Approved Protein Names Online



Protein nomenclature can be complicated. The [Universal Protein Resource KnowledgeBase \(UniProtKB\)](#) provides a comprehensive database of protein names, symbols, and sequence information. The symbols for proteins are usually related to the symbols of the genes that code for them. Like human genes, human proteins are written in all capital letters, but to distinguish them from genes, the symbols for proteins are not italicized.

Examples:

Gene: *FGL2*

Protein: FGL2

Formerly, proteins were sometimes identified by the letter *p* plus their molecular weight in kilodaltons; probably the best-known of

these to cancer researchers is p53. However, the protein symbol TP53, which reflects the approved gene symbol *TP53*, should be used instead because the names beginning with *p* are now considered outdated.¹

In some cases, you may choose to spell out a protein's name, followed by its symbol in parentheses, to make its function clearer to your readers. Recommended protein names can be found at UniProtKB.²

Examples:

breast cancer type 1 susceptibility protein (BRCA1)

UniProtKB contains information on proteins for many species, not just humans and other vertebrates, but you can easily identify proteins in your species of interest by filtering the search results.

References

1. *AMA Manual of Style*, 11th ed. Oxford University Press; 2020: 841-842.
2. European Bioinformatics Institute, National Center for Biotechnology Information, Protein Information Resource, Swiss Institute for Bioinformatics. International Protein Nomenclature Guidelines. Published April 3, 2018. Accessed August 18, 2021.
https://www.uniprot.org/docs/International_Protein_Nomenclature_Guidelines.pdf

—Amy Ninetto

Image: TP53 tumor suppressor–DNA complex. Created with [BioRender.com](https://www.biorender.com)