Editing Wikipedia articles on

Genes and proteins

Editing Wikipedia can be daunting for newcomers, especially as a student editor contributing to Wikipedia for the first time as a class assignment. This guide is designed to help you create or expand articles about genes and proteins to Wikipedia.

Be accurate

You're contributing to a resource millions of people use to inform themselves about a cutting edge area of scientific research. Your work for this class will be read by thousands of people, so it's important to be accurate.

By documenting and sharing your knowledge about genes and proteins, you'll help Wikipedia provide a wider spectrum of scientific knowledge. With great power comes great responsibility!

Understand the guidelines

Take time to read and understand the suggestions here to make the most of your contributions to Wikipedia.

If you post something that doesn't meet these guidelines, it may take up valuable volunteer time that could have been spent making your contributions, or other content, even better. If you aren't comfortable with these guidelines, talk to your instructor about an alternative off-wiki assignment.

Engage with editors

Part of the Wikipedia experience is receiving and responding to feedback from other editors. Don't wait until the last day to make a contribution, or you may miss comments, advice, and ideas. Volunteers from the Wikipedia community might respond or ask questions about your work. If they do, make sure to acknowledge it. Discussion is a crucial part of the Wikipedia process.

Watch out for close paraphrasing

Use your own words. Plagiarizing by copying-and-pasting, or close paraphrasing — when most of the words are changed, but the structure and meaning of the original text remains — is against the rules.

Plagiarism is a violation of your university's academic honor code. Plagiarism on Wikipedia will be caught by other editors, and there will be a permanent online record of it tied to your account. Even standard resources or authors, such as educational resources and abstracts of articles in science journals, are under copyright and should never be directly copied.

The best way to avoid this hassle is to make sure you draw from a variety of sources, really understand your material, and write about it in your own words. If you aren't clear on what close paraphrasing is, contact your university's writing center.

Be bold!

Take the time to understand the rules and guidelines, and soon you'll be contributing important knowledge to a resource you and millions of other people use every day.

Using reliable sources

Wikipedia content is based on reliable academic works that summarize individual studies and provide a topic overview. These include:

• Literature reviews or systematic reviews from reliable journals, such as *Biochemistry*, the *Journal of Biological Chemistry,* or *Cell*.
• Encyclopedias and specialist textbooks from reliable publishers, such as Oxford University Press, Cambridge University Press, or Wiley.

To see if a journal is reliable, search *Beall's List*, an online reference of low-quality academic journals. Don't cite from these journals.

▶ For info on good sources, see | *shortcut* WP:SOURCE and WP:SCIRS

Choosing an article

Once you have a variety of good sources, choose topics with lots of literature in peer-reviewed sources that don't have good coverage on Wikipedia. Check an article's Talk page for its rating: "Start-class" or "stub" articles need your help! Avoid "Featured" or "Good" articles.

Structuring your genes or proteins article

You can add, remove, or reorder sections for your own topic.

• **InfoBox:** If you use one, place it at the very beginning of the article (*see next page*).
• **Lead section:** A short article summary.
• **Gene:** Where is it located (i.e., which chromosome?) How it is expressed? Don't repeat what's in the InfoBox.
• **Protein:** Specific information about the protein (describe splice variants, alterations, etc.). Are there notable small molecules that bind to, inhibit, or activate the protein?
• **Species, tissue, subcellular distribution:** An optional section. What species have this gene? Where is the gene expressed? Describe subcellular compartments or organelles where you would find the protein (i.e., cytoplasm, nucleus, mitochondria, cell membranes).
• **Function:** What does the transcribed protein do?
• **Interactions:** What does the protein interact with, if anything?
• **Clinical significance:** List diseases or conditions that result from a mutation in the gene, or a deficiency or excess of the expressed protein. See Wikipedia's sourcing requirements for medical information at | *shortcut* WP:MEDRS
• **History section:** An optional section describing key aspects of the protein's history. Don't mention research groups or institutions that conducted a study. However, you may list those who made key discoveries relayed to the gene or protein in this section. For example, the scientist or group that first cloned the gene, determined its function, linked it to a disease, or won a major award for its discovery.

Good articles include *Protein C*, *Gonadotropin-releasing hormone*, or *Rubisco*. For more suggestions, check out | *shortcut* Portal:Gene Wiki

---

What's a shortcut?

WP:SOURCE is what's known on Wikipedia as a *shortcut*.

You can type this text into Wikipedia's search bar to pull up specific pages.

What's a lead section?

The lead section is the first section of a Wikipedia article. Good leads summarize the entire article, offering context for important aspects of the topic. Names should come from the UniProt database (free, online) and include the approved HUGO (Human Genome Organisation) symbol, like this:

(UniProt Name) is a protein that, in humans, is encoded by the (HUGO gene symbol) gene.

For more tips, see | *shortcut* WP:MCBMOS

What's MEDRS?

ANY statement, in any article, relating to human medical or psychological health, must follow a stricter set of sourcing guidelines. See | *shortcut* WP:MEDRS

Develop your article

What's an InfoBox?

An InfoBox gives an overview of a protein's (or enzyme or RNA) properties at a glance.

To add one to an article, copy the text from the appropriate template:

• **Individual proteins**, or in protein family articles where more than one protein is discussed. Find an example in the Estrogen receptor article:
| *shortcut* Template:InfoBox Protein

• **Protein families**, for protein families, i.e., evolutionary related proteins that share a common 3D structure. Find an example in the T-box article:
| *shortcut* Template:InfoBox Protein family

• **Enzymes**, for enzymes based on EC number, usually the reaction catalyzed by the enzyme rather than the enzyme itself. Find an example in the Alcohol dehydrogenase article:
| *shortcut* Template:InfoBox enzyme

• **GNF proteins**, for genes/proteins with an ortholog within the human genome. Find an example in the Reelin article:
| *shortcut* Template:InfoBox GNF protein

• **Nonhuman proteins**, for proteins without a human ortholog. Find an example in the Uterine serpin article:
| *shortcut* Template:Nonhuman protein

• **RNA families**, for RNA families, i.e., evolutionary related non-coding RNAs that share a common 3D structure. Find an example in the U1 spliceosomal RNA article:
| *shortcut* Template:InfoBox rfam

Building your InfoBox

Each template holds the source code for an InfoBox. First, copy the source code (example below). Then, using the text editor ("Edit source"), add it to the start of your article, above the lead. Here's an example:

```
{{Infobox enzyme
| name =
| image =
| image_size =
| caption =
| EC_number =
| CAS_number =
| GO_code =
}}
```

Add what you can to the right of the equal sign on each line. Most acronyms are explained, with links to sources, on the Template pages mentioned above. Never delete anything from the template — just skip fields you don't know.

The Image line contains *only* the file name of the image from Wikimedia Commons, without additional formatting. For example:

```
Image = ChymotrypsinA1.jpg
```

Cite your sources in an InfoBox just as you would any other fact.

Cite your sources

You must have the PubMed ID (PMID) or Digital Object Identifier (DOI) for every journal article you cite, and the ISBN and page number of any book you cite. This enables other editors to verify the accuracy of the information you added.

A fast start for protein articles

1. Go to the GeneWikiGenerator:

   http://biogps.org/GeneWikiGenerator/

2. Find and enter the Entrez ID for your gene. Click "Search by Symbol or Accession"

3. You'll see the Code Creator up top. It checks to see it your article exists already. If the Template and Article titles are "missing," you can proceed.

4. To create a new Protein Box, click "create" in the Template section. You'll see the template being created and uploaded to Wikipedia.

5. To create a new article, choose a title. Click "Check" to ensure your title doesn't conflict with an existing page. If not, click "Create '<article name>'". Once it says "Created (reload page)", you may visit your new page.

Note that as the GeneWikiGenerator is a third-party site hosted independently from Wikipedia, it may not always be available.

Useful tips

As you start writing, keep these guidelines in mind:

- **Never offer medical advice.** This includes information about genes related to diseases or treatments. If your work includes *any* statements about human health, you must follow a very strict set of sourcing guidelines. See more at | *shortcut* WP:MEDRS

- The abbreviations of genes follows the HUGO Gene Nomenclature Committee and written in italic font style (the full names are also written in italic). Use abbreviations instead of the full name. For a style guide on gene and protein articles, check out | *shortcut* WP:MCBMOS

- When mentioning technical terms for the first time, provide a short, plain-English explanation in parentheses, if possible.

Final thoughts

- Don't procrastinate! Writing good, reliable Wikipedia articles takes time. Don't wait until the last minute. If you get stuck, always ask your instructor for extra time, rather than adding content to Wikipedia that doesn't meet these guidelines.

- Check back on your page. It can be interesting to see how your article grows, but you may also have comments and suggestions from other editors, and want to get involved in new developments.

Have more questions?

Do you have more questions? You can always ask the Wikipedia Content Expert listed on your course page. You can also post general questions to Wikipedians at the Teahouse: | *shortcut* WP:TH