# Article information:

Large language models generate functional protein sequences across diverse families | Nature Biotechnology  
<https://www.nature.com/articles/s41587-022-01618-2>

# Article summary:

1. Traditional methods for protein engineering are limited in their ability to create novel proteins with desired properties.

2. Deep neural networks have been used to generate artificial proteins, but they rely on scarce experimental structure data and difficult biophysical simulations.

3. ProGen is a new protein language model that uses a large dataset of 280 million raw protein sequences to generate artificial proteins across multiple families and functions without relying on structural or coevolutionary approaches.

# Article rating:

May be slightly imbalanced: The article presents the information in a generally reliable way, but there are minor points of consideration that could be explored further or claims that are not fully backed by appropriate evidence. Some perspectives may also be omitted, and you are encouraged to use the research topics section to explore the topic further.

# Article analysis:

The article “Large language models generate functional protein sequences across diverse families” is generally reliable and trustworthy, as it provides an overview of the current state of protein engineering and introduces a new approach for generating artificial proteins using deep learning-based language models. The article is well-researched and provides evidence for its claims, such as the limitations of traditional methods for protein engineering, the success of deep neural networks in generating artificial proteins, and the development of ProGen as a new language model for generating functional proteins across diverse families.

The article does not appear to be biased or one-sided in its reporting, as it presents both traditional methods for protein engineering and deep learning-based approaches in an objective manner. It also does not contain any promotional content or partiality towards either approach. Furthermore, the article does not overlook any potential risks associated with ProGen; instead, it acknowledges that further research is needed to assess the safety and efficacy of this new approach before it can be applied in real-world settings.

The only potential issue with the article is that it does not explore any counterarguments or alternative perspectives on ProGen’s effectiveness or potential applications. While this may be due to space constraints, it would have been beneficial if the authors had discussed some possible drawbacks or limitations of this new approach so readers could gain a more comprehensive understanding of its implications.

# Topics for further research:

* Protein engineering safety risks
* Limitations of deep learning-based language models
* Alternative approaches to protein engineering
* Real-world applications of ProGen
* Potential drawbacks of ProGen
* Comparison of traditional and deep learning-based protein engineering methods

# Report location:

<https://www.fullpicture.app/item/c4d4ebd70f04e1355752152106234cb0>