

Af2-analysis: a Python package for alphafold 2 analysis

Alaa Reguei¹ and Samuel Murail¹

¹ Université Paris Cité, Inserm U1133, CNRS UMR 8251, Paris, France

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#)
- [Repository](#)
- [Archive](#)

Editor: [Open Journals](#)

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

Since the publication of Alphafold 2 software in 2021 (?), the scientific community can access a prediction accuracy of protein structures that was previously unattainable. The Alphafold 2 software has been widely adopted by the scientific community. The derivative Alphafold Multimer (Evans et al., 2022) has been developed to predict the structure of protein complexes, once again defining a new standard for protein-protein and protein-peptide docking. However as shown by Björn Wallner, it is sometime necessary to generate thousands of model to get a good model (Wallner, 2023). The analysis of the results can then be a challenge for many researchers. Here, we present Af2-analysis, a Python package that provides tools for the analysis of Alphafold 2 results. Af2-analysis is designed to facilitate the analysis of protein structures predicted by Alphafold 2, providing functions for the comparison of predicted structures with experimental structures, the visualization of predicted structures, and the calculation of structural quality metrics. Af2-analysis is open-source and freely available on GitHub.

Statement of need

Af2-analysis is a Python package that provides tools for the analysis of protein structures predicted by Alphafold 2. The package is designed to facilitate the analysis of protein structures predicted by Alphafold 2, providing functions for the comparison of predicted structures with experimental structures, the visualization of predicted structures, and the calculation of structural quality metrics. The package is open-source and freely available on GitHub.

Mathematics

Single dollars (\$) are required for inline mathematics e.g. $f(x) = e^{\pi/x}$

Double dollars make self-standing equations:

$$\Theta(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{else} \end{cases}$$

You can also use plain \LaTeX for equations

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx \quad (1)$$

and refer to [Equation 1](#) from text.

Citations

Citations to entries in paper.bib should be in [rMarkdown](#) format.

If you want to cite a software repository URL (e.g. something on GitHub without a preferred citation) then you can do it with the example BibTeX entry below for (?).

For a quick reference, the following citation commands can be used: - @author:2001 -> "Author et al. (2001)" - [@author:2001] -> "(Author et al., 2001)" - [@author1:2001; @author2:2001] -> "(Author1 et al., 2001; Author2 et al., 2002)"

Figures

Figures can be included like this: Caption for example figure. and referenced from text using [section](#) .

Figure sizes can be customized by adding an optional second parameter: Caption for example figure.

Acknowledgements

We acknowledge contributions from Brigitta Sipocz, Syrtis Major, and Semyeong Oh, and support from Kathryn Johnston during the genesis of this project.

References

Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., ... Hassabis, D. (2022). Protein complex prediction with AlphaFold-multimer. *bioRxiv*. <https://doi.org/10.1101/2021.10.04.463034>

Wallner, B. (2023). AFsample: improving multimer prediction with AlphaFold using massive sampling. *Bioinformatics*, 39(9), btad573. <https://doi.org/10.1093/bioinformatics/btad573>