Evaluating the Efficacy of Generative AI in Antibody Design

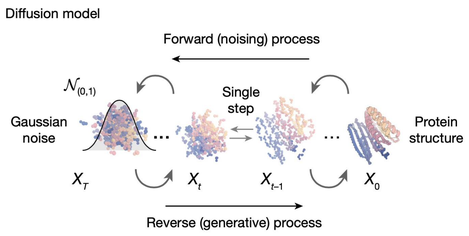
By: Emily Green, Dr. Yun Li

Abstract

Antibodies play a key role in promising an innovative and new avenue for therapeutics, diagnostics, and research tools. The design of nanobody 4Y8D is compared to a design showing promise of improvement of design after diffusion.

Introduction

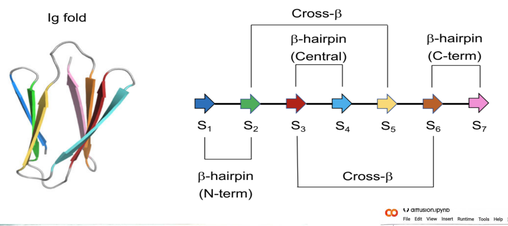
- In recent years, Deep learning was proven triumphant in conquering the protein-folding issue. Further milestones were achieved in protein de novo design (David Baker et al., Nobel Prize in Chemistry, 2024).
- Antibodies (IgG) contain characteristic folding and loops that allows binding to proteins, those characteristics can be adapted with RFDiffusion to produce more efficient binding
- Diffusion models generate structures by mimicking natural particle diffusion processes. This approach allowed scientists to tap into the knowledge acquired by the protein-folding prediction neural networks and apply it to protein de novo design.
- 4Y8D is a nanobody associated with Cyclin-G associated kinase (GAK) for Hepatitis C.



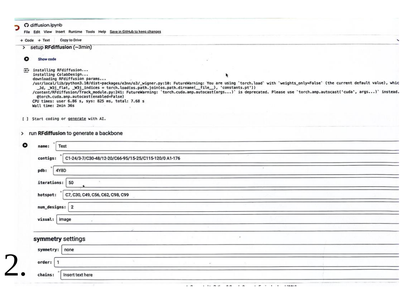
Methods

To evaluate the efficacy of RFDiffusion, we altered the antigen-binding loops of the 4Y8D nanobody using a Google Colab implementation of RFDiffusion. The results were analyzed and presented using PyMol.

1.



2.



3.

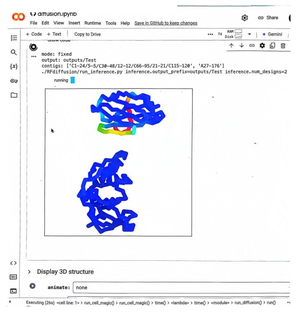
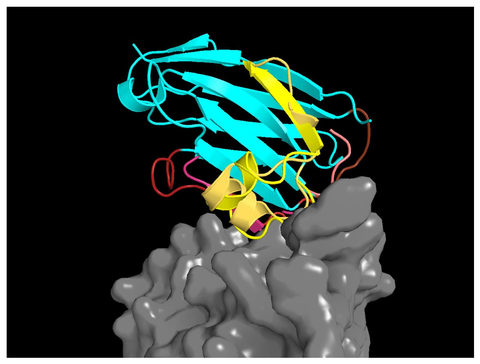


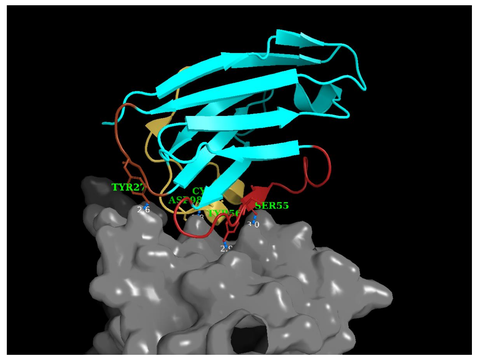
Figure 1 displays the standard antibody (igG), highlighting the characteristic antiparallel strands and the looping patterns used to bind. Figures 2,3 contain the used Configs and Hotspots plugged into Google Colab to design an antibody

Results

A.



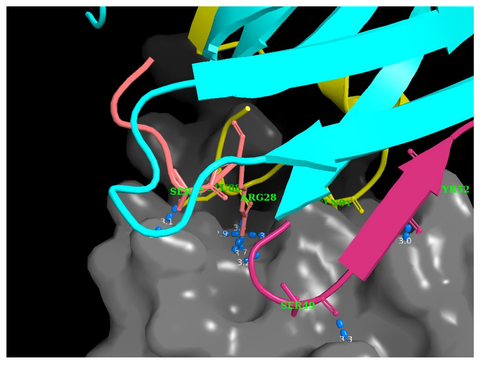
B.



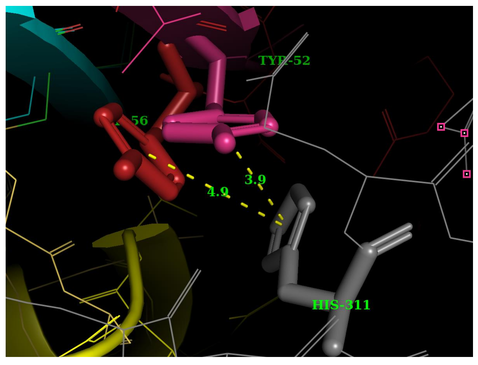
C.

The wild-type nanobody and the AI adaptation show remarkable similarity in their overall folding (A), yet each implementing a different amino-acid sequence and a different set of polar interactions with the target kinase (B: 4Y8D nanobody; C: AI adaptation).

D.



E.



RFDiffusion exhibited uncanny abilities to strategically place charged to establish salt bridges (D) and optimize aromatic interactions (E)

Corresponding Loop	Wt Residues	Kinase Residues
Loop I	Tyr 27	Glu 302
Loop II	Ser 55	Gln 312
	Tyr 56	Glu 308
Loop III	Asp 98	Ser 305
	Cys 99	Glu 308

Corresponding Loop	Design Residues	Kinase Residues
Loop I	Ser 27	Glu 302
	Arg 28	Glu 308
		Glu 301, Glu 302, Arg 303
Loop II	Ser 49	Ala 294
	Tyr 52	Glu 308, Gln 312
Loop III	Glu 88	LYS243
	Gln 91	ALA307

Above the tables display the noted interactions on the loops known to bind GAK from both the wildtype (Wt) and the design. The wildtype table highlights Tyr from loop 1 and Asp from loop 3 as charged interactions. Within the design table the Arg 28/Glu 308 interaction is highlighted for being an ionic interaction.

Discussion

- The resigned antigen-binding loops are structurally sound and closely resemble the original folding pattern. This is rather remarkable considering that the AI was completely agnostic about the original peptide sequences.
- Further analyses revealed that the three loops all established unique interactions with the antigen, with varying degrees of impact to the overall binding affinity and specificity of the nanobody. Specifically, the binding affinity and specificity of Loop I was greatly enhanced in the redesign — the original ionic hydrogen bond to the kinase Glu302 was preserved, in the meantime, a highly specific salt bridge was added through Arg28. Loop II, being the shortest of the three, maintained its two regular hydrogen bonds in the redesign. Loop III, being the longest and most structured of the three, maintained one ionic hydrogen bond and relegated the other one to a regular hydrogen bond.

$$\Sigma_{Wt}=36Kcal_{/mol}$$
$$\Sigma_{Design}=66_{Kcal/mol}$$
$$66_{Kcal/mol}-36_{Kcal/mol}=30_{Kcal/mol}$$
$$e^{(30Kcal/mol/RT)}=1.9\ e^{7}$$

- Another remarkable finding in the AI design is the optimization of a pi-pi interaction. The wild-type Tyr56-His311 interaction was at too great a distance and an angel too steep for optimal pi-pi interaction. In the redesign, Tyr52 replaced Tyr56 to reestablish this interaction at optimal geometry.

Conclusion

RoseTTAFold, an alternative to AlphaFold, proved to be better suited for de novo protein due to its emphasis on local interactions both at the sequence level and at structure level. Although the authors of RFDiffusion only made modest statements about the ability of their system to strategically place polar and ionic side chains, our design using their system turns out to be highly satisfying. Even more impressive is how efficient this method is. A new protein design only took several minutes of calculation on the Colab platform. Selected designs can then be further characterized using MD simulations for functional assessment, which can further identify key candidates for synthesis and in vitro or in vivo analyses. This development will usher in a new era in the life sciences and revolutionize medicine and the pharmaceutical industry.

References

Chidyausiku, T. M., Mendes, S. R., Klima, J. C., Nadal, M., Eckhard, U., Roel-Touris, J., Houliston, S., Guevara, T., Haddox, H. K., Moyer, A., Arrowsmith, C. H., Gomis-Rüth, F. X., Baker, D., & Marcos, E. (2022). De novo design of immunoglobulin-like domains. *Nature communications*, *13*(1), 5661. https://doi.org/10.1038/s41467-022-33004-6

Martinkus, K., Ludwiczak, J., Cho, K., Liang, W.-C., Lafrance-Vanasse, J., Hotzel, I., Rajpal, A., Wu, Y., Bonneau, R., Gligorijevic, V., & Loukas, A. (2024b, March 6). *Abdiffuser: Full-atom generation of in vitro functioning antibodies*. arXiv.org. https://arxiv.org/abs/2308.05027

Roel-Touris, J., Carcelén, L., & Marcos, E. (2024). The structural landscape of the immunoglobulin fold by large-scale de novo design. *Protein science : a publication of the Protein Society*, *33*(4), e4936. https://doi.org/10.1002/pro.4936

Roel-Touris, J., Nadal, M., & Marcos, E. (2023). Single-chain dimers from de novo immunoglobulins as robust scaffolds for multiple binding loops. *Nature communications*, *14*(1), 5939. https://doi.org/10.1038/s41467-023-41717-5

Sanderson, G. (2017, October 5). *Neural Networks*. YouTube. https://www.youtube.com/watch?v=aircAruvnKk&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi&index=1

Wang, J., Chen, C., Yao, G., Ding, J., Wang, L., & Jiang, H. (2023). Intelligent Protein Design and Molecular Characterization Techniques: A Comprehensive Review. *Molecules*, *28*(23), 7865. https://doi.org/10.3390/molecules28237865