Deep Learning Approaches for the Design of Symmetric Cyclic Peptide Complexes

John Chu[1,4], David Juergens[2,3], Helen Eisenach[3], Gaurav Bhardwaj[3], Ph.D., David Baker[3,5], Ph.D.

[1]Rosetta Commons REU. [2]Molecular Engineering and Science Institute, University of Washington. [3]Institute for Protein Design, University of Washington. [4]Department of Mathematics and Computer Science, Dickinson College. [5]Howard Hughes Medical Institute .

Introduction

Background: Cyclic peptides have demonstrated promising potential for use as drugs and other molecular functionalities. However, the *de novo* design of such molecules has been mostly limited to classical, energy-based design techniques and focused on the design of cyclic peptides as monomers rather than as homo-oligomer species.

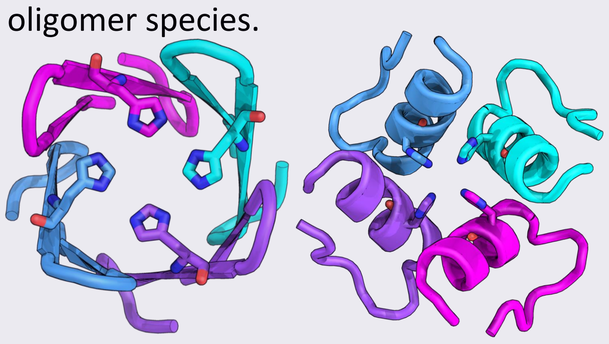


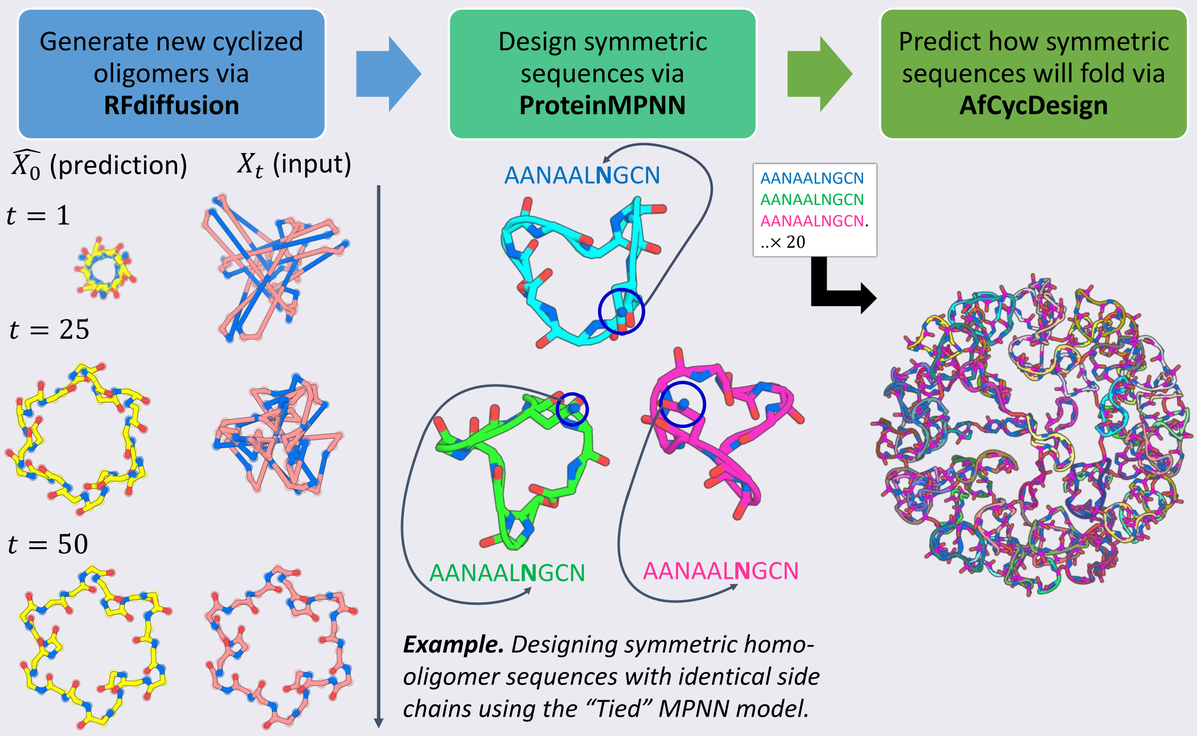
Figure 1. *Cyclic peptide designs for nickel binding generated by Gaurav Bhardwaj.*

Objective

- Explore deep learning methods to design homo-oligomeric cyclic peptide systems.
- Design a protocol to generate cyclic peptide oligomers that exhibit "cage-like" high-order symmetrical structures such as icosahedrons for controlled release of therapeutics and higher stability of oligomers.

Design Pipeline

RFdiffusion[1]: Protein structure generative diffusion model.
ProteinMPNN[2]: Protein sequence design model using message-passing neural networks given backbone structures.
AFCycDesign[3]: a variant of AlphaFold2 to predict the cyclic peptide structures.



Example. *Designing symmetric homo-oligomer sequences with identical side chains using the "Tied" MPNN model.*

Hyperparameters
- Chain length
- Guiding potential strength/schedule
- Number of diffusion steps

Symmetric Sequence Design
- Design a symmetric amino acid sequence using ProteinMPNN.

Confident Metrics
- pLDDT $>$ 90
- Normalized pAE $<$ 0.2
- RMSD $<$ 2

Figure 2. *Design pipeline of cyclized peptide oligomers.*

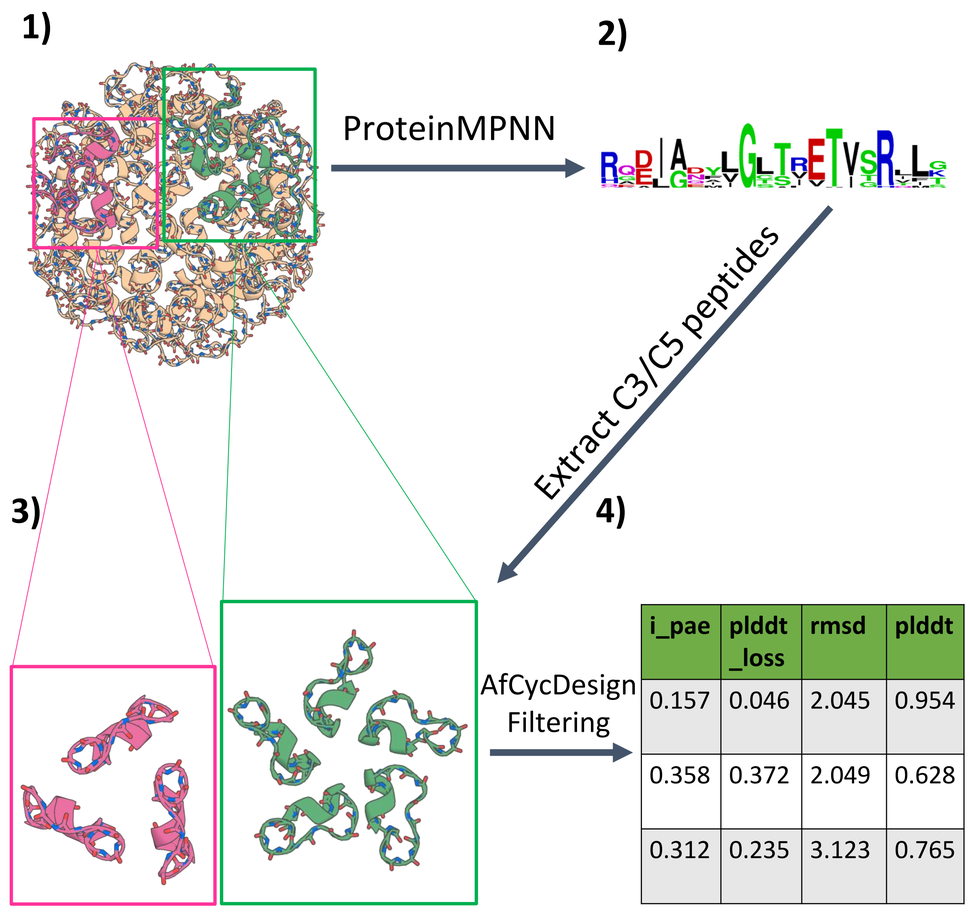
Methods

Sample Dataset

- 4320 generated icosahedral oligomers of cyclic peptides using 4 different RFdiffusion design hyperparameters and 2 guiding potentials (e.g. inter/intra molecular weights).
- 59718 extracted C3 and C5 cyclic peptides from backbones of generated oligomers.

Design Protocol

1. Generate cyclized peptide oligomers of icosahedral symmetry using RFdiffusion.
2. Design symmetric sequences of those oligomers using ProteinMPNN.
3. Extract C3/C5 peptides from those sequence designs.
4. Predict and validate the accuracy of extracted C3/C5 peptides using AfCycDesign.

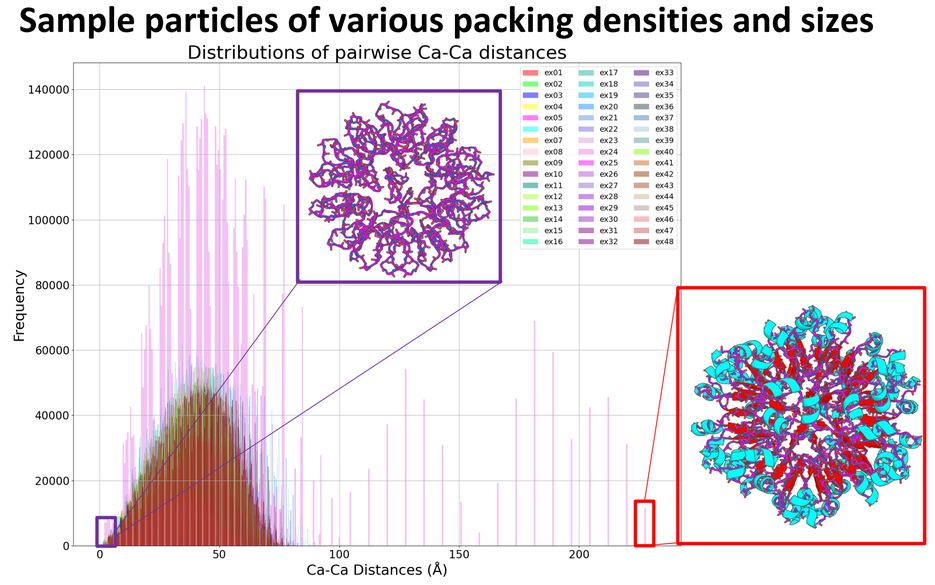


i_pae	plddt_loss	rmsd	plddt
0.157	0.046	2.045	0.954
0.358	0.372	2.049	0.628
0.312	0.235	3.123	0.765

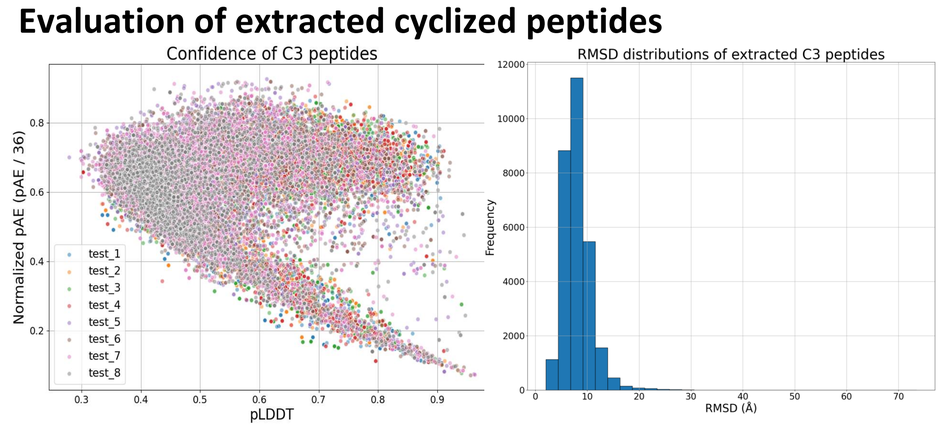
Figure 3. Design protocol to generate cyclic peptide oligomers exhibiting icosahedral symmetries.

Results

Sample particles of various packing densities and sizes



Evaluation of extracted cyclized peptides



Examples of C3/C5 peptides with low RMSD

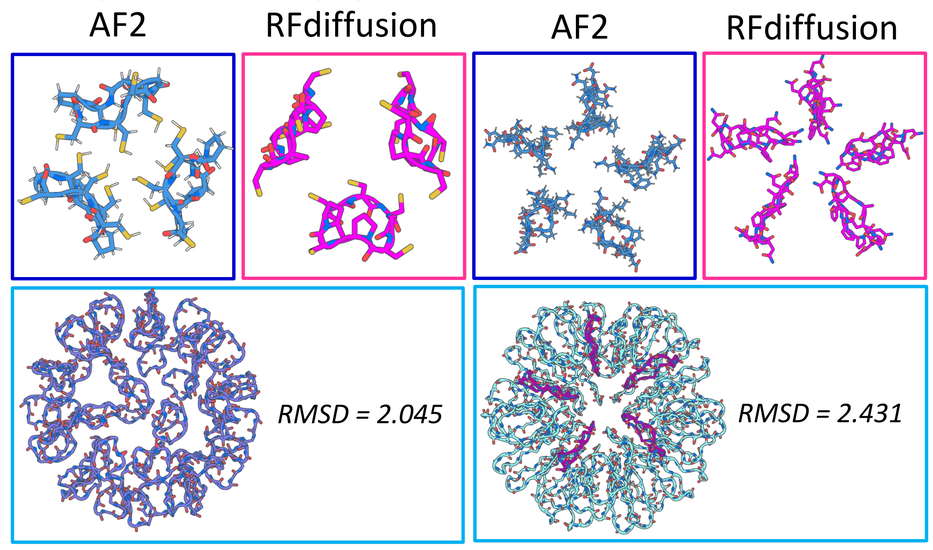


Figure 4. Predictions of C3 and C5 peptides extracted from the icosahedral complex.

Discussion & Progress

Takeaway

1. Icosahedral symmetric oligomers composed of cyclic peptides can be designed with RFdiffusion and validated *in silico* with AfCycDesign.
2. Oligomers composed of longer cyclic peptide chains tend to have higher AF2 confidence.

Progress & Future Directions

1. Characterize promising designs experimentally.
2. Generalize the pipeline to other point symmetries such as dihedral and octahedral groups.
3. Change the AF2 settings
   - Use "AF2 initial guess" as in [1].
   - Try running the monomer AF2 weights instead of multimer weights.
4. Expand the current pipeline to accommodate scaffolding metal-binding motifs.
5. Improve the current pipeline to enable "symmetric diffusion" for robust generation of new particles.

Acknowledgements

This project was primarily mentored by **David C. Juergens** and **Dr. Gaurav Bhardwaj**. I sincerely appreciate their insightful guidance and support in accomplishing this project.

Furthermore, I would appreciate **Helen Eisenach** for providing co-mentorship on this study. Lastly, I am grateful to **Dr. David Baker** for providing the lab resources and support necessary for this research.

References
[1] Watson, J. et al. (2023). 'De novo design of protein structure and function with Rfdiffusion'. Nature.
[2] Daupras, J. et al. (2022). 'Robust deep learning-based protein sequence design using ProteinMPNN'. Science.
[3] Rettie, S. et al. (2023). 'Cyclic peptide structure prediction and design using AlphaFold'. bioRxiv.