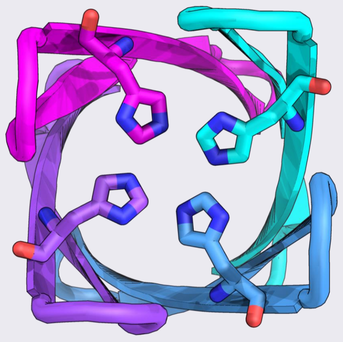
Design of Symmetric Cyclic Peptide Complexes using RFdiffusion

John Chu[1,4], David Juergens[2,3], Helen Eisenach[3], David Baker[3,5], Ph.D., Gaurav Bhardwaj[3], 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 chemical functionality for drug and materials design due to their customizable chemical properties and contributions to structural stability. However, the *de novo* design of such molecules has been limited to classical, energy-based methodologies.



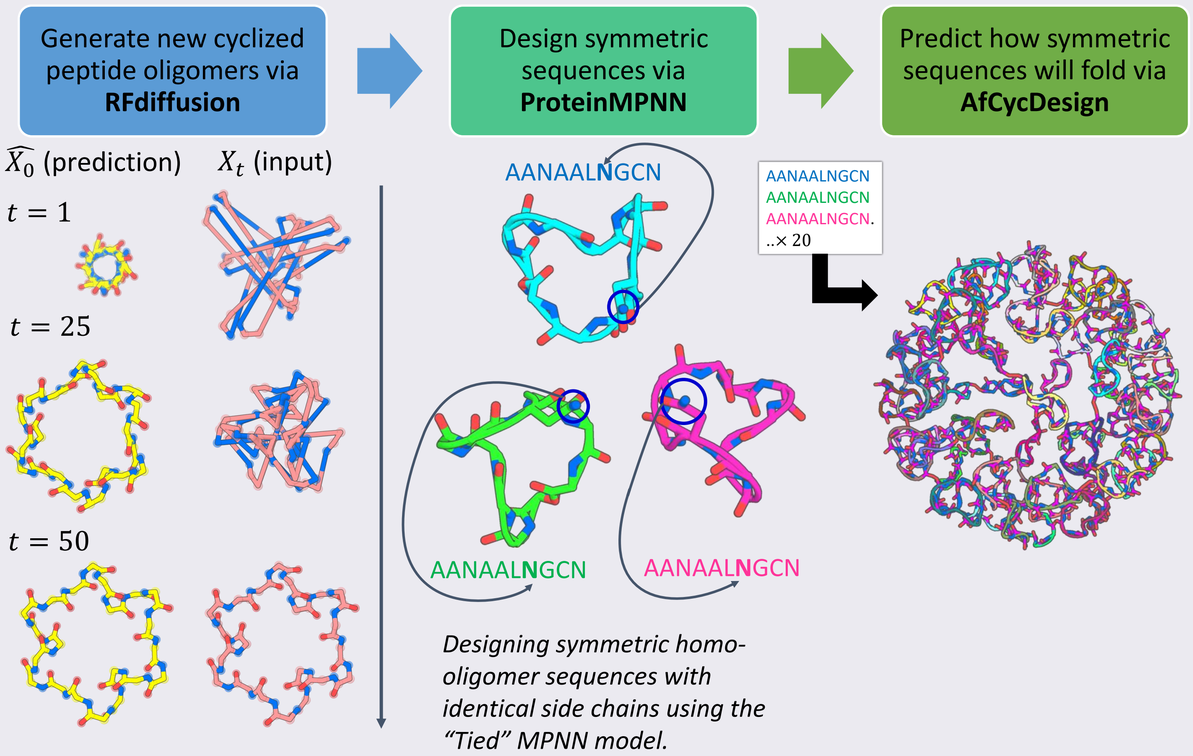
Cyclic peptide design for Ni^{2+} binding generated by Gaurav Bhardwaj.

Objective: The goal of this study is to take advantage of deep-learning-based protein-design techniques for designing symmetric homo-oligomers composed of cyclic peptides.

Why design symmetric cyclic peptide complexes? Rational designs of oligomers that consist of cyclic peptides will improve our ability to design (A) drugs with controlled spatiotemporal release and stability and (B) materials with interesting electrical, photonic, and catalytic properties.

Design Pipeline

RFdiffusion[1]: A generative protein-design model that combines structure prediction networks and generative diffusion models.
ProteinMPNN[2]: Protein sequence design model using message-passing neural networks given backbone structures.
AfCycDesign[3]: a variant of AlphaFold2 (AF2) to predict the cyclic peptide structures.



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

Hyperparameters used
- Chain length (L)
- Guiding potential strength/schedule
- Number of diffusion steps (T)

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

Highly confident structures
- pLDDT $>$ 90
- Normalized pAE $<$ 0.2
- RMSD $<$ 2

Methods

Design Protocol

Test 1-4
- **Guiding potentials**
  - intra and inter-molecular weights:
    - ❑ Test 1: (0.7, 0.7), Test 2: (0.8, 0.6), Test 3: (0.5, 0.5), Test 4: (0.8, 0.7)
  - guiding scale: 0.5, 1, 2, 3, 5, 8
  - guiding decay: linear, quadratic
  - number of diffusions (T): 16, 20, 25, 35, 50
  - number of chains: 8

6 * 2 * 5 * 1 = 60 (settings / test)

4 (tests) * 10 (designs/setting) * 60 (settings/test) = 2400 designs

Test 5-8
- **Guiding potentials**
  - intra and inter-molecular weights:
    - ❑ Test 5: (0.7, 0.7), Test 6: (0.8, 0.6), Test 7: (0.5, 0.5), Test 8: (0.8, 0.7)
  - guiding scale: 0.5, 1, 2, 3, 5, 8
  - guiding decay: linear, quadratic
  - Number of diffusions (T): 50
  - Number of chains: 10, 12, 16, 20

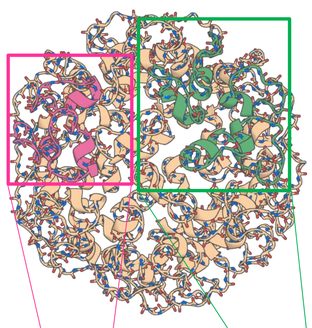
6 * 2 * 4 * 1 = 60 (settings /test)

4 (tests) * 10 (designs/setting) * 48 (settings/test) = 1920 designs

Extracted C3/C5 peptides: 8 * (sequences / design) * (2400 + 1920) * designs ⇒ 34,560 (C3) + 34,560 (C5) = 69, 120 designs

1. Generate cyclized peptide oligomers of icosahedral symmetry using RFdiffusion.
2. Design symmetric sequences of those oligomers using ProteinMPNN which produces 8 sequences per design.
3. Extract C3/C5 peptides from those sequence designs.
4. Predict and validate the accuracy of extracted C3/C5 peptides using AfCycDesign.

1) Icosahedral complex of cyclic peptides

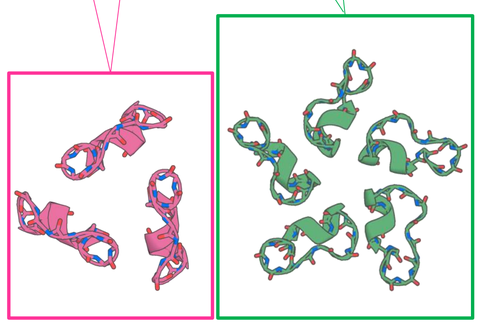


ProteinMPNN

Extract C3/C5 peptides

2) Symmetric Sequence

3)



Extracted C3/C5 peptides

AfCycDesign

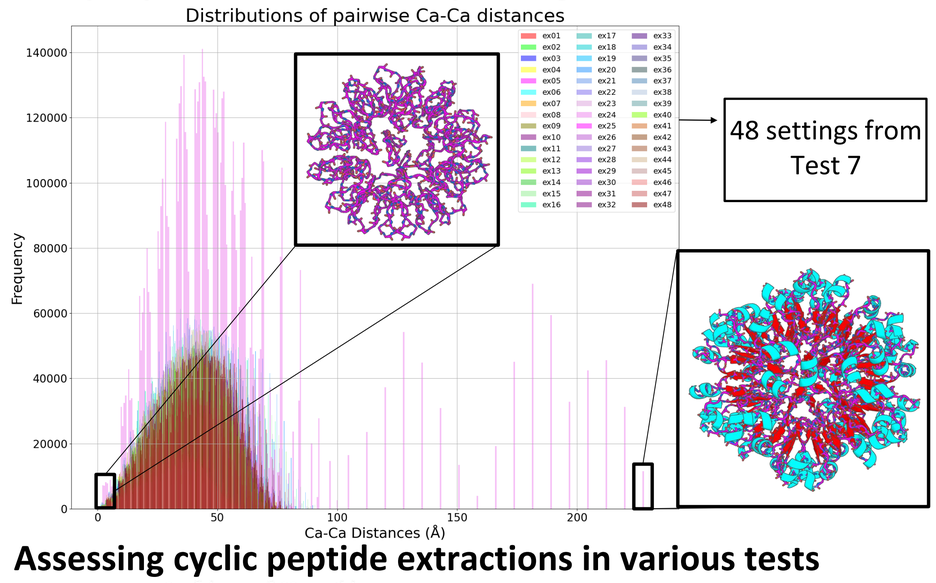
4)

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

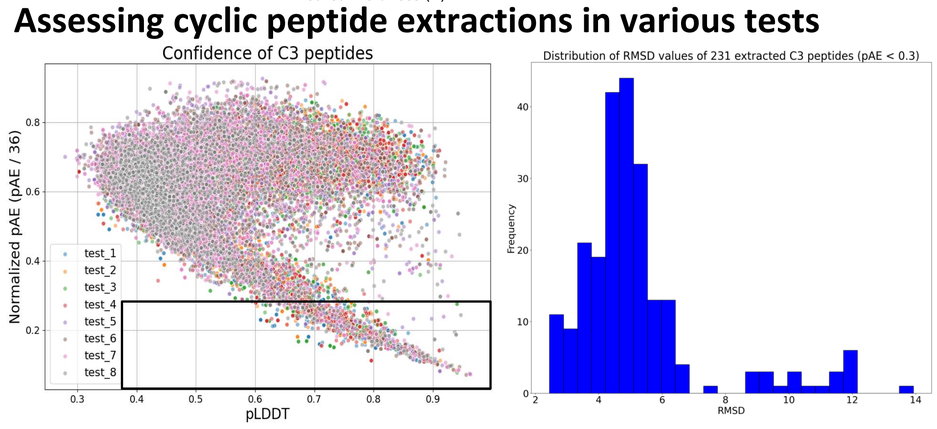
Assessed accuracy of AfCycDesign

Results

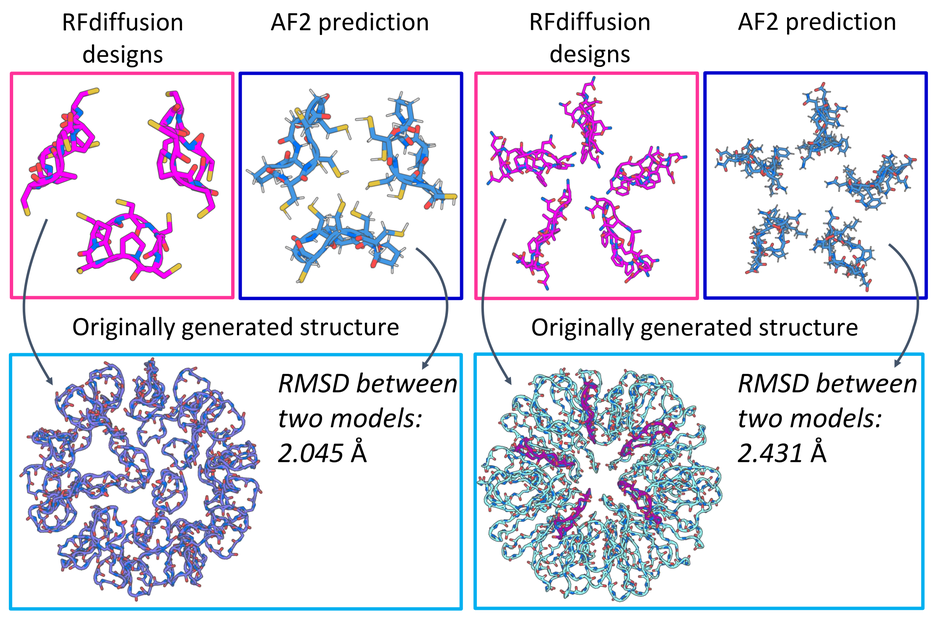
Sample particles of various packing densities and sizes



Assessing cyclic peptide extractions in various tests



Examples of C3/C5 peptides with low RMSD

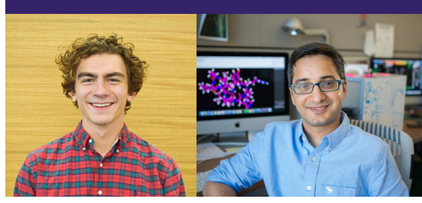


Discussion & Future Directions

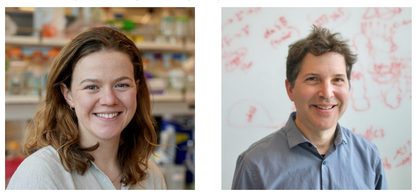
Takeaway
1. Icosahedral symmetric oligomers composed of cyclic peptides can be designed with RFdiffusion and validated *in silico* with AfCycDesign.
2. There are very few designs that have both high confidence and low RMSD values.

Progress & Future Directions
- Improve the RFdiffusion & AF2 settings as followings:
  - Examine additional RFdiffusion hyperparameters and broaden the range of values tested.
  - Use "AF2 initial guess" as in [1].
  - Try running the monomer AF2 weights instead of multimer weights.
- Compute the monomer RMSD values in comparison to those of designed and predicted structures.
- Characterize promising designs experimentally.
- Generalize to other point symmetries such as dihedral and octahedral complexes.
- Expand the current pipeline to accommodate scaffolding metal-binding motifs.
- Inclusion of noncanonical amino acids for expanded chemical functionalities.

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

References

[1] Watson, J. et al., *Nature (2023),* https://doi.org/10.1038/s41586-023-06415-8
[2] Daupras, J. et al., Science, 375(6615): 49-56.
[3] Rettie, S. et al. *bioRxiv (2023),* https://doi.org/10.1101/2023.02.25.529956.