

# Computational Chemistry and Artificial Intelligence Guided Study of Lasso Peptide Folding and Cyclase Binding for Enhanced Peptide Drug Discovery

Song Yin<sup>1,2</sup>, Xuenan Mi<sup>2,3</sup>, Susanna E. Barrett<sup>2,4</sup>, Douglas A. Mitchell<sup>5,6</sup>, & Diwakar Shukla<sup>1,2,3,4\*</sup>

<sup>1</sup>Department of Chemical and Biomolecular Engineering, <sup>2</sup>Carl R. Woese Institute for Genomic Biology, <sup>3</sup>Center for Biophysics and Quantitative Biology,

<sup>4</sup>Department of Chemistry, University of Illinois Urbana-Champaign. Emails: (songyin2, xmi4, seb5, diwakar)@illinois.edu

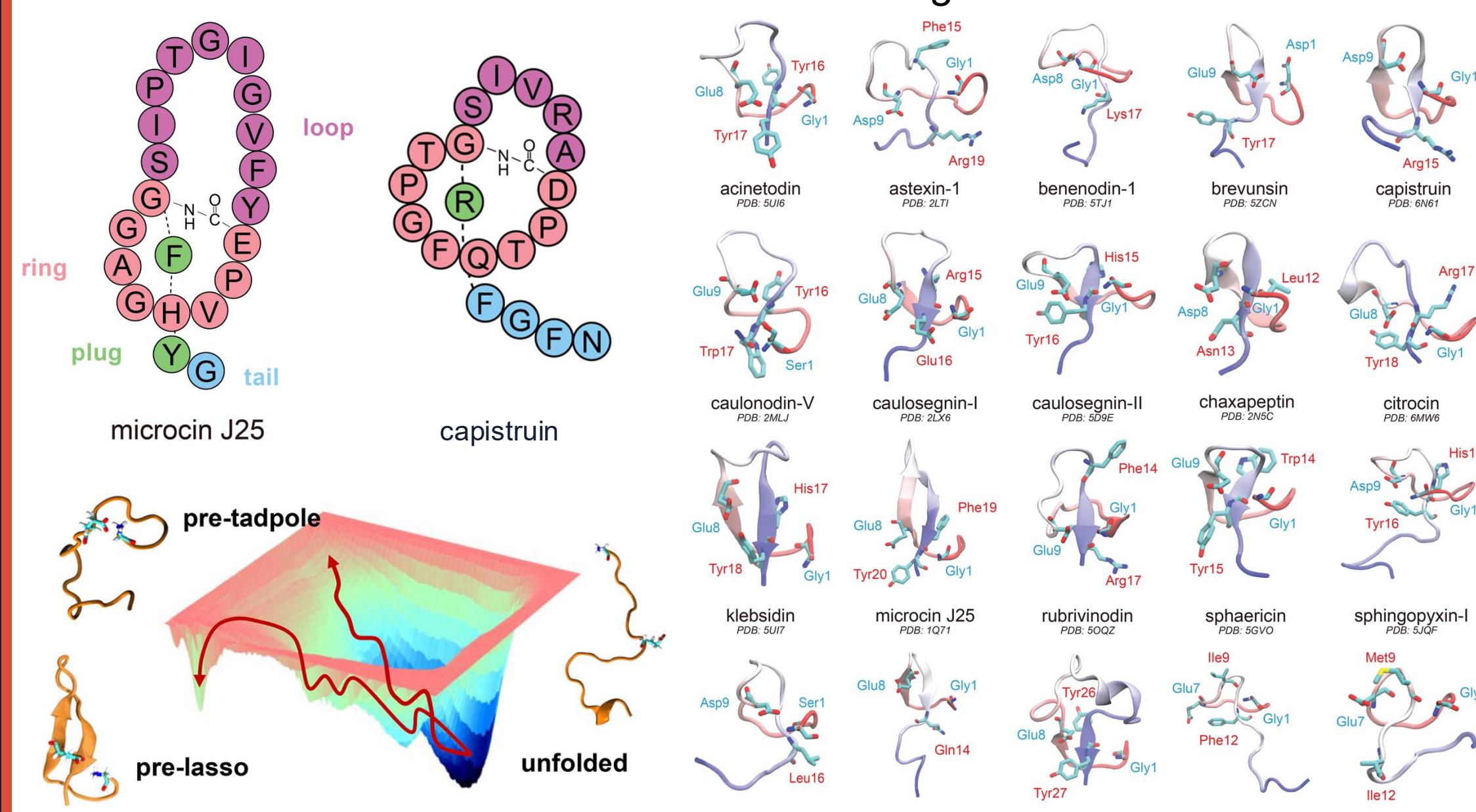
<sup>5</sup>Department of Biochemistry, <sup>6</sup>Department of Chemistry, Vanderbilt University. Emails: douglas.mitchell@vanderbilt.edu



## De Novo Lasso Peptide Folding

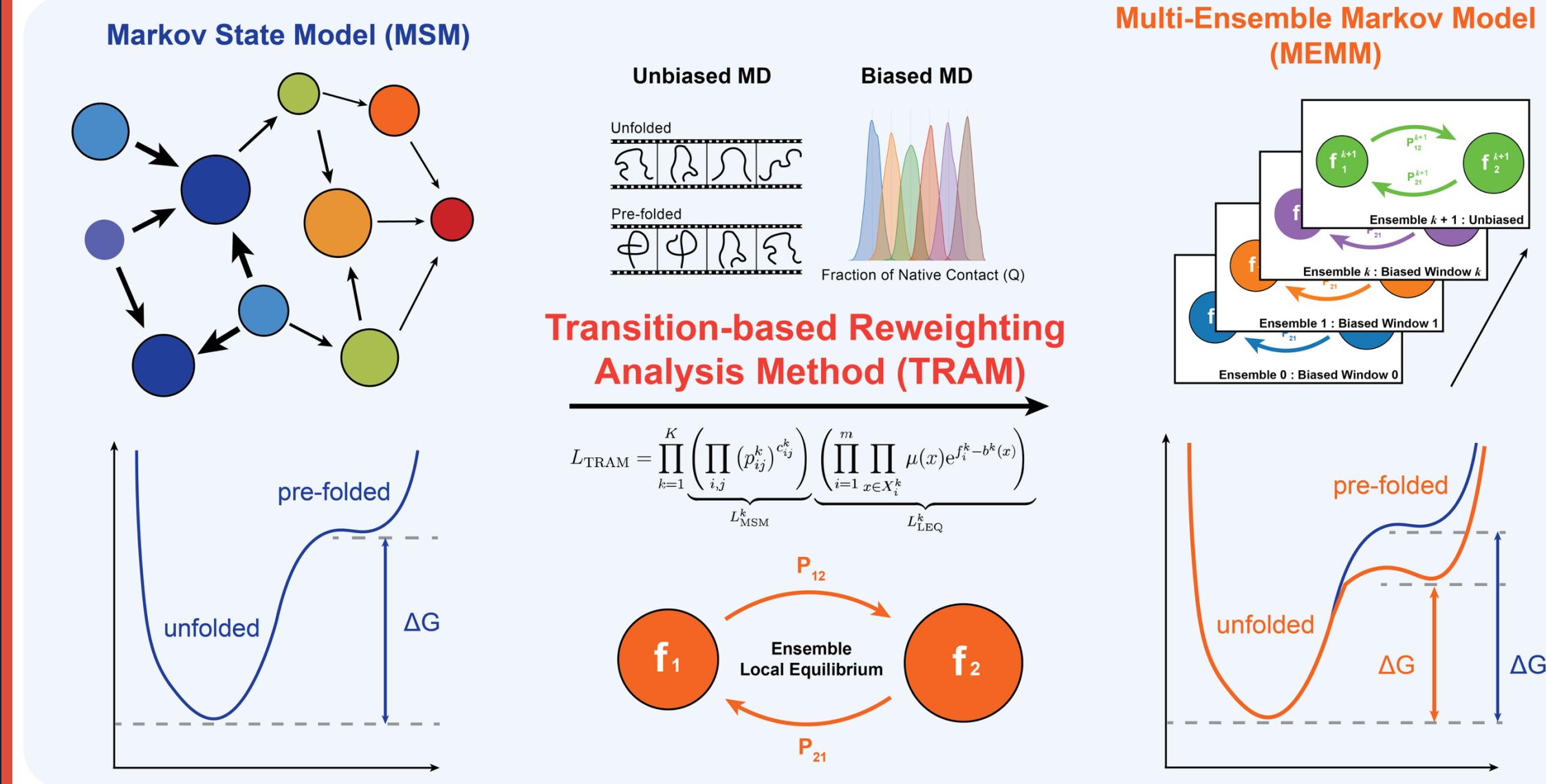
### Background & Motivation

- Lasso peptides are ribosomally synthesized and post-translationally modified peptides (RiPPs) with an interlocked lariat [1]rotaxane structure, showing potential as antibiotic, anticancer, and antiviral drug candidates.
- Previous folding studies have focused on one lasso peptide<sup>1</sup> that is insufficient to characterize the universal folding mechanism.



### Method

- We conducted extensive unbiased (~200 μs each) and biased (~7 μs each) MD simulations for 20 lasso peptides and innovatively applied TRAM<sup>2</sup> and LPC-VAE<sup>3</sup> frameworks to systematically investigate de novo folding mechanism of lasso peptides.

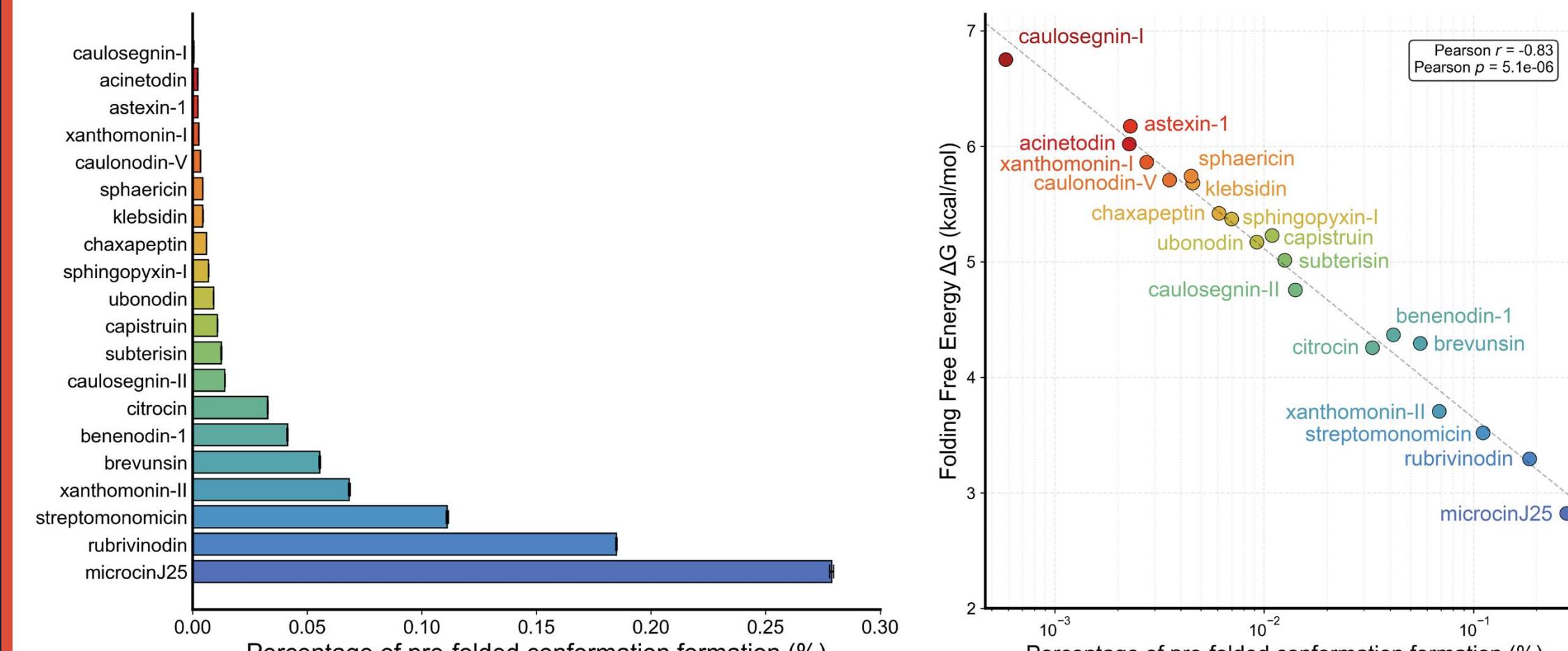


### TRAM Outperforms MSM in Resolving Folding Landscapes

- Kinetic Asymmetry:** Lasso peptides unfold rapidly but fold slowly through energetically uphill pathways.

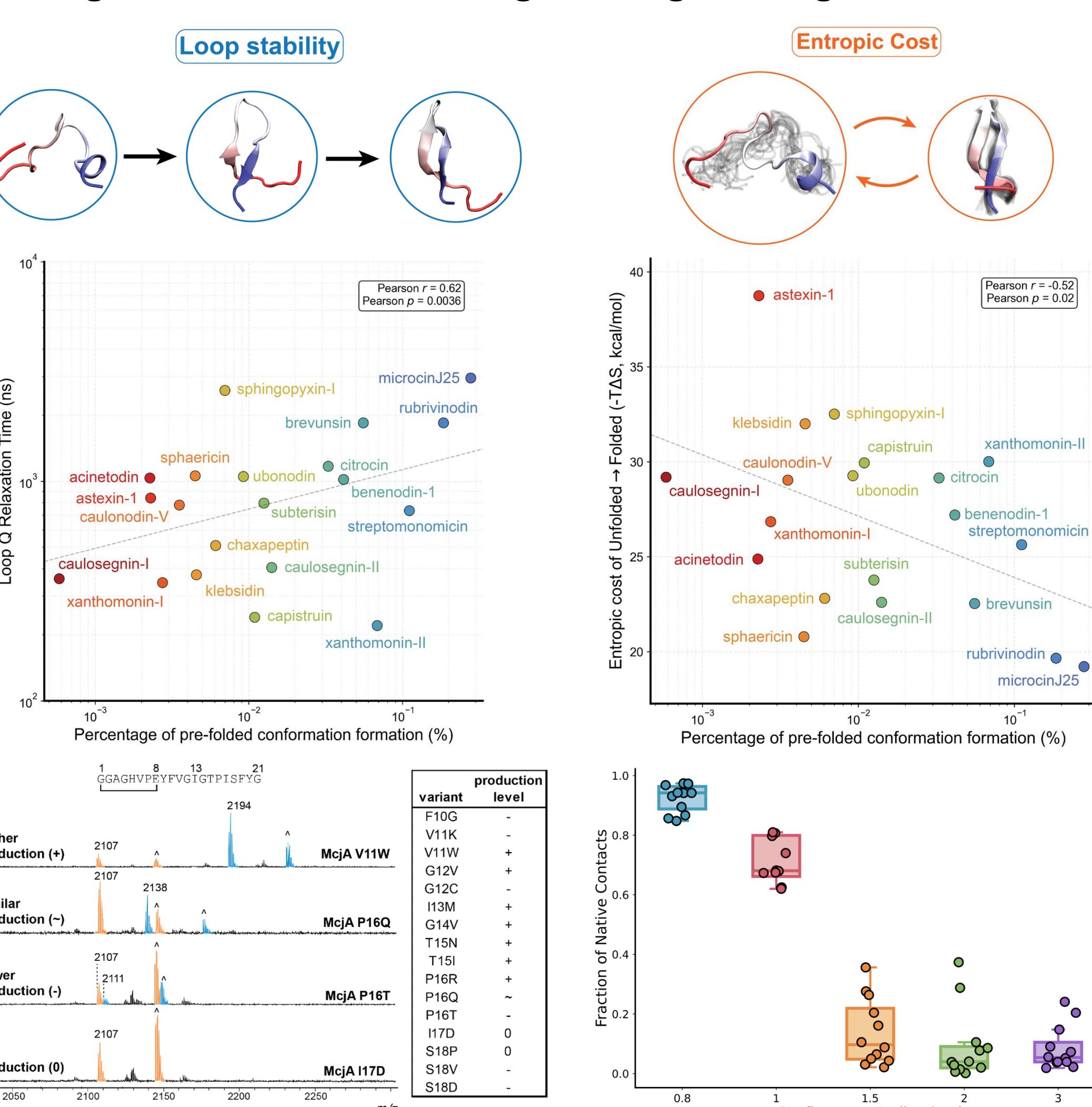
- TRAM vs MSM:** TRAM incorporates biased simulations to capture rare folding transitions, providing more accurate folding thermodynamic and kinetic estimates for lasso peptides where unbiased sampling and MSM fails.

### Lasso peptide de novo folding is rare



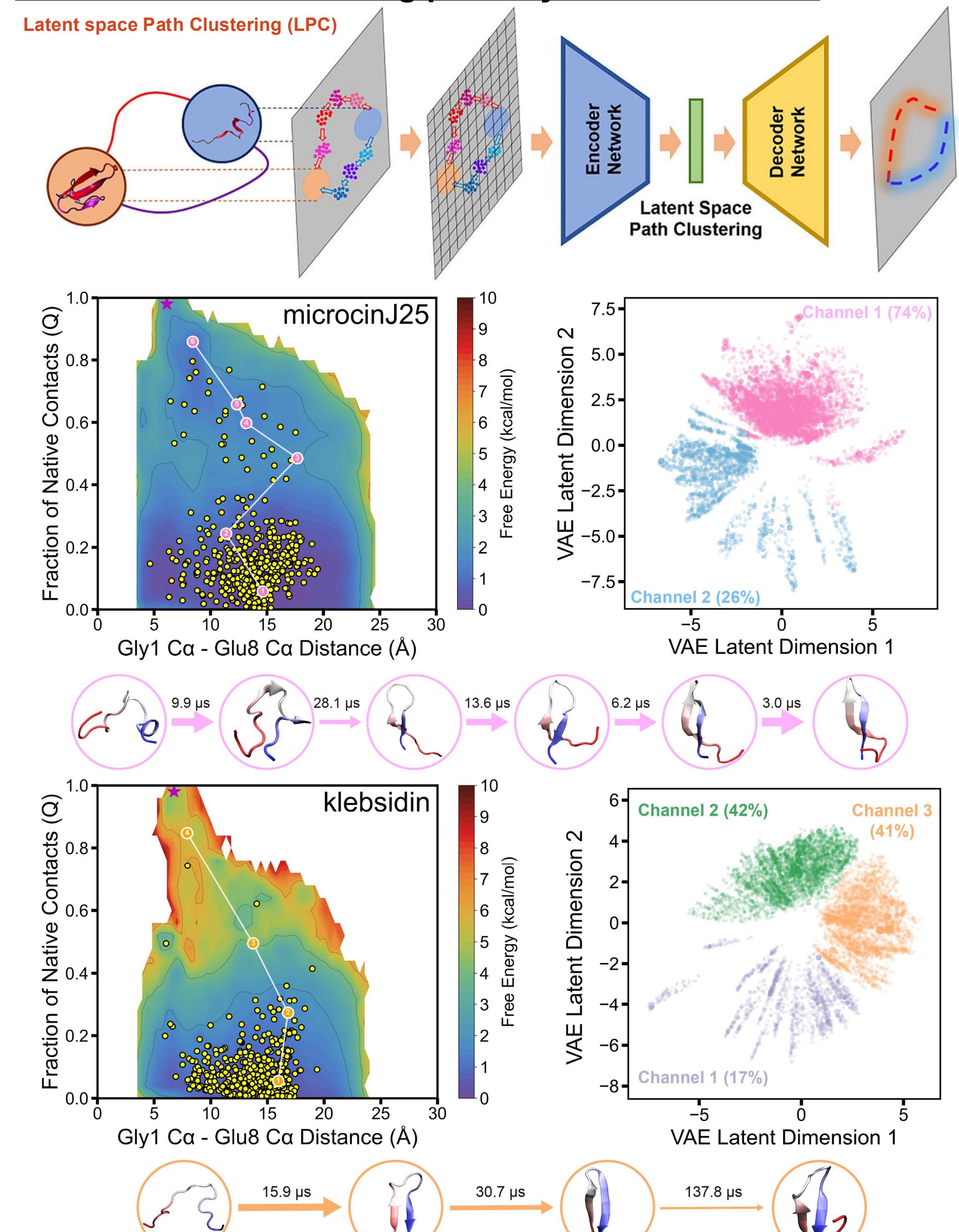
- All 20 lasso peptides show <1% pre-folded formation probability, revealing that lasso topology is thermodynamically unfavorable without enzyme assistance.

### Folding Determinants and Engineering Strategies



- Stable loop regions with well-defined secondary structures ( $\beta$ -sheets) enhance lasso peptide formation.
- Lasso peptide folding imposes large entropic penalties requiring enthalpic stabilization for successful formation.

### LPC-VAE assisted folding pathway characterization

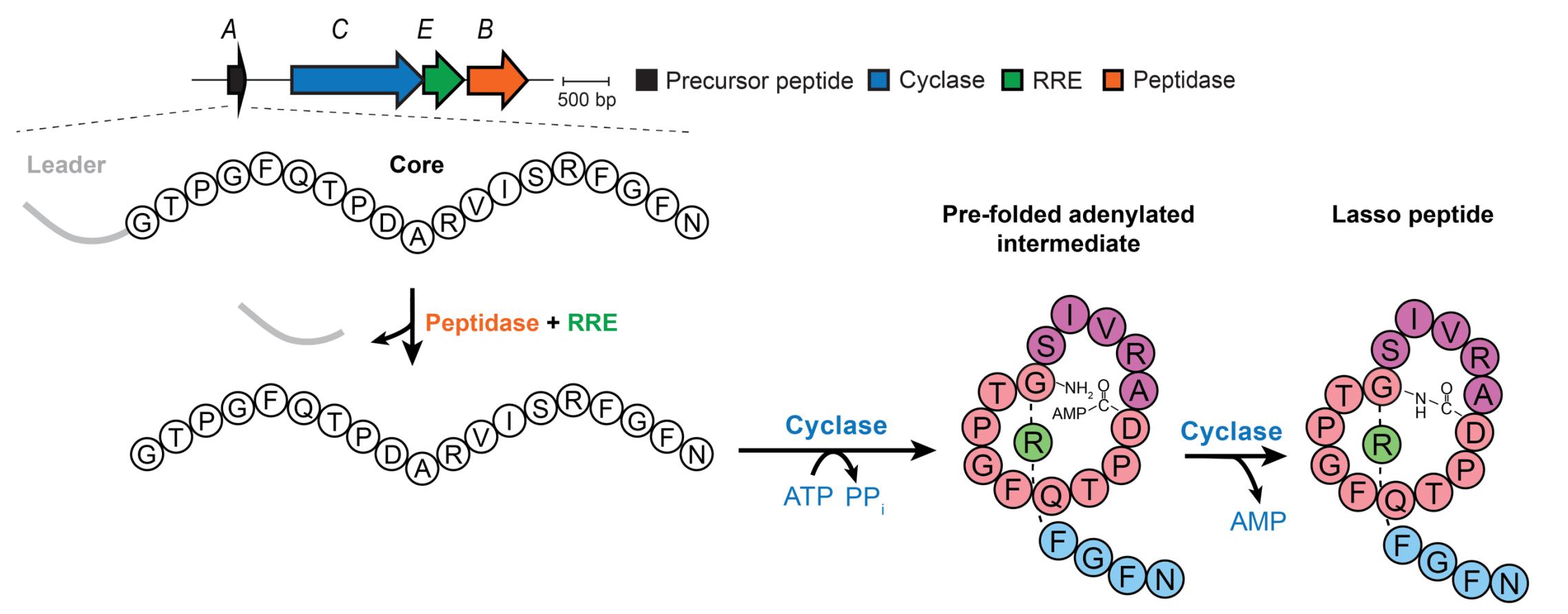


- Applied VAE-based clustering to identify distinct folding pathway channels and demonstrate the most representative pathway with highest flux.

## Lasso Peptide Folding and Cyclase Binding

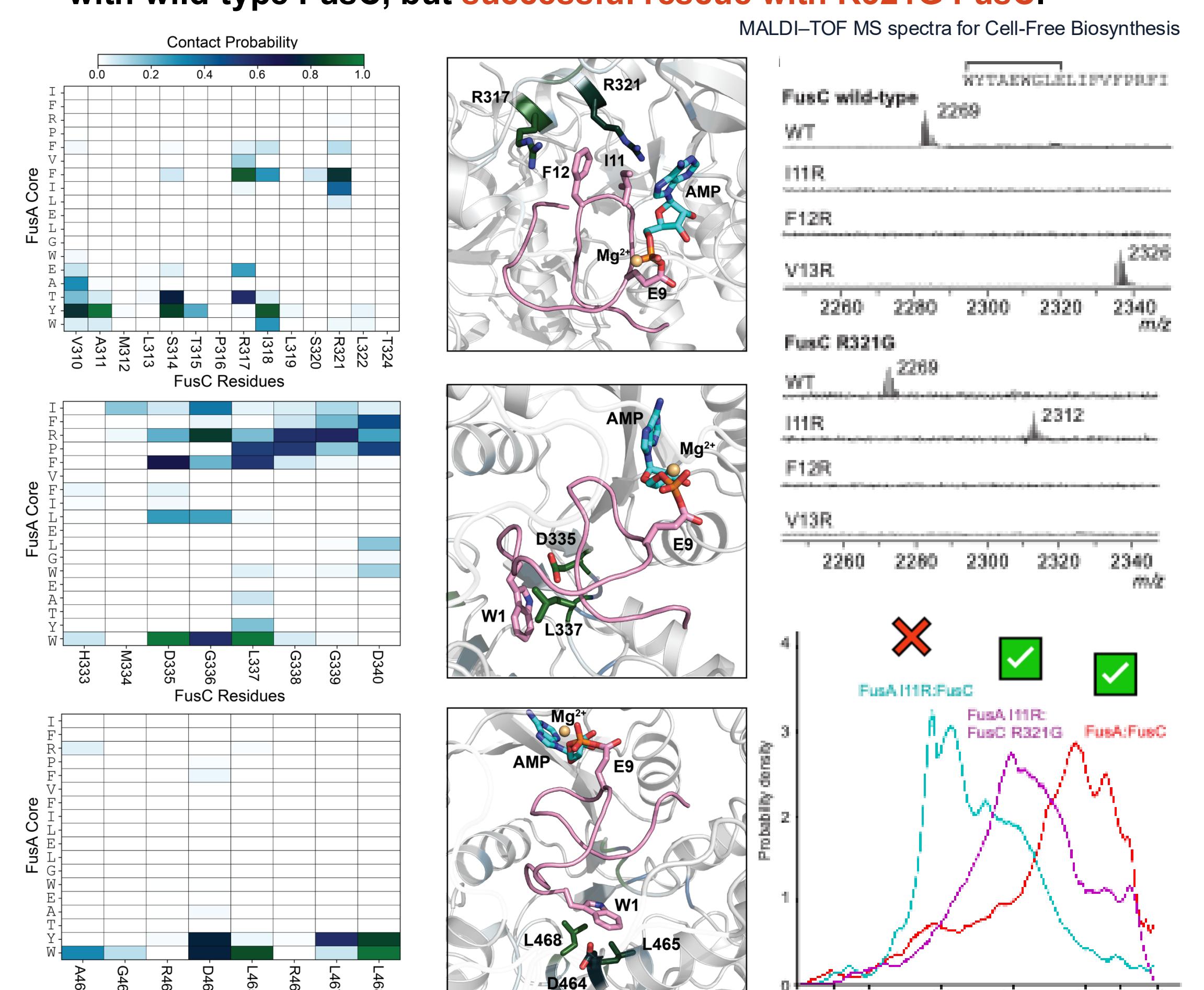
### Background & Motivation

- The mechanism by which cyclase converts an unstructured peptide into the lasso topology remains unclear, and predicting enzyme-substrate tolerance is also challenging.

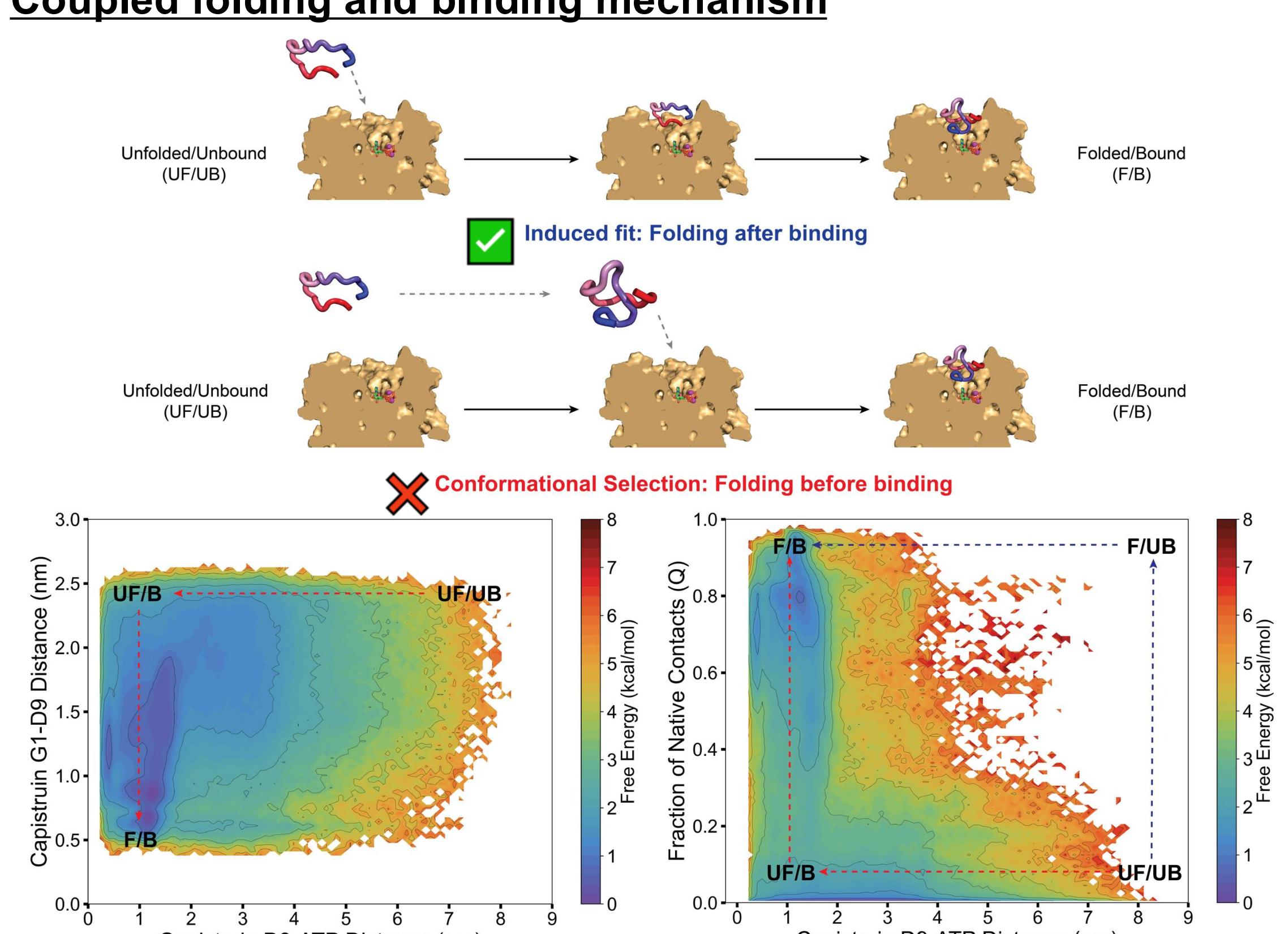


### Substrate interactions guide cyclase engineering and lasso peptide diversification

- MD simulations of cyclase bound to adenylated Fusilassin revealed key interactions between loop residues Ile11, Phe12, and Arg321 in cyclase, confirming limited sequence variation tolerance in the lasso peptide loop region<sup>4</sup>.
- Peptide and cyclase engineering involved testing I11R and V13R variants with wild-type and R321G FusC, showing reduced compatibility for I11R with wild-type FusC, but successful rescue with R321G FusC.

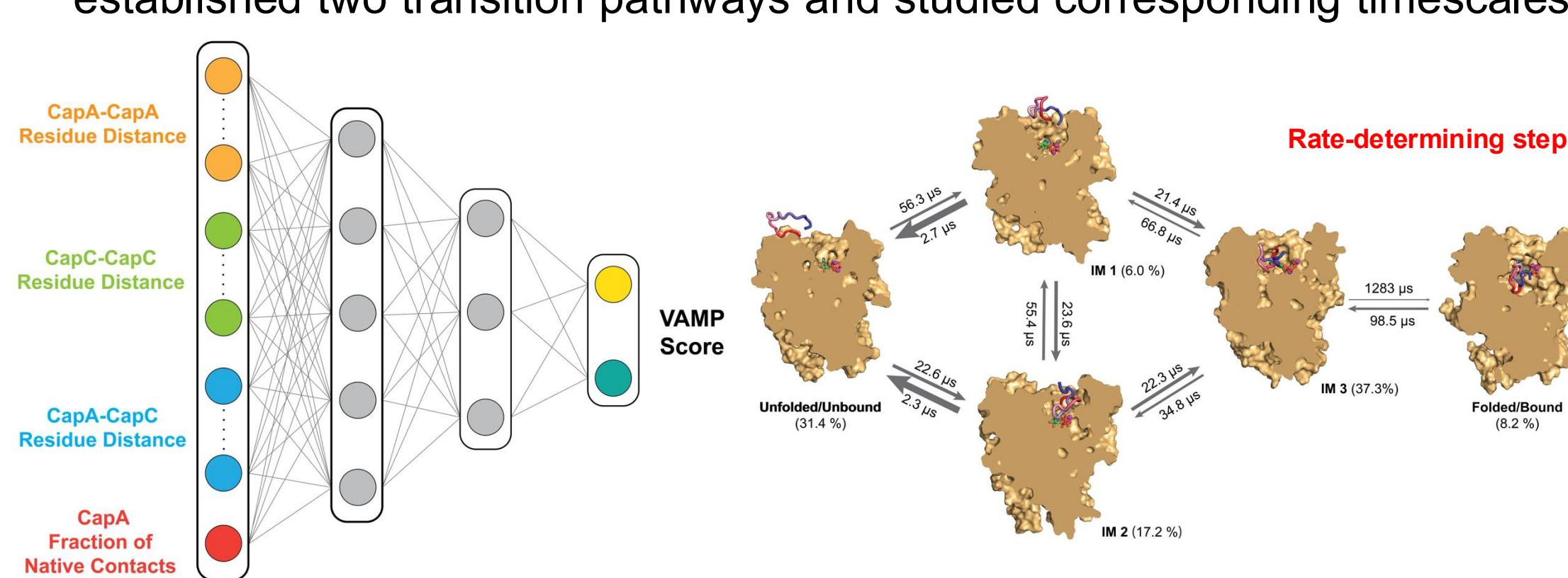


### Coupled folding and binding mechanism

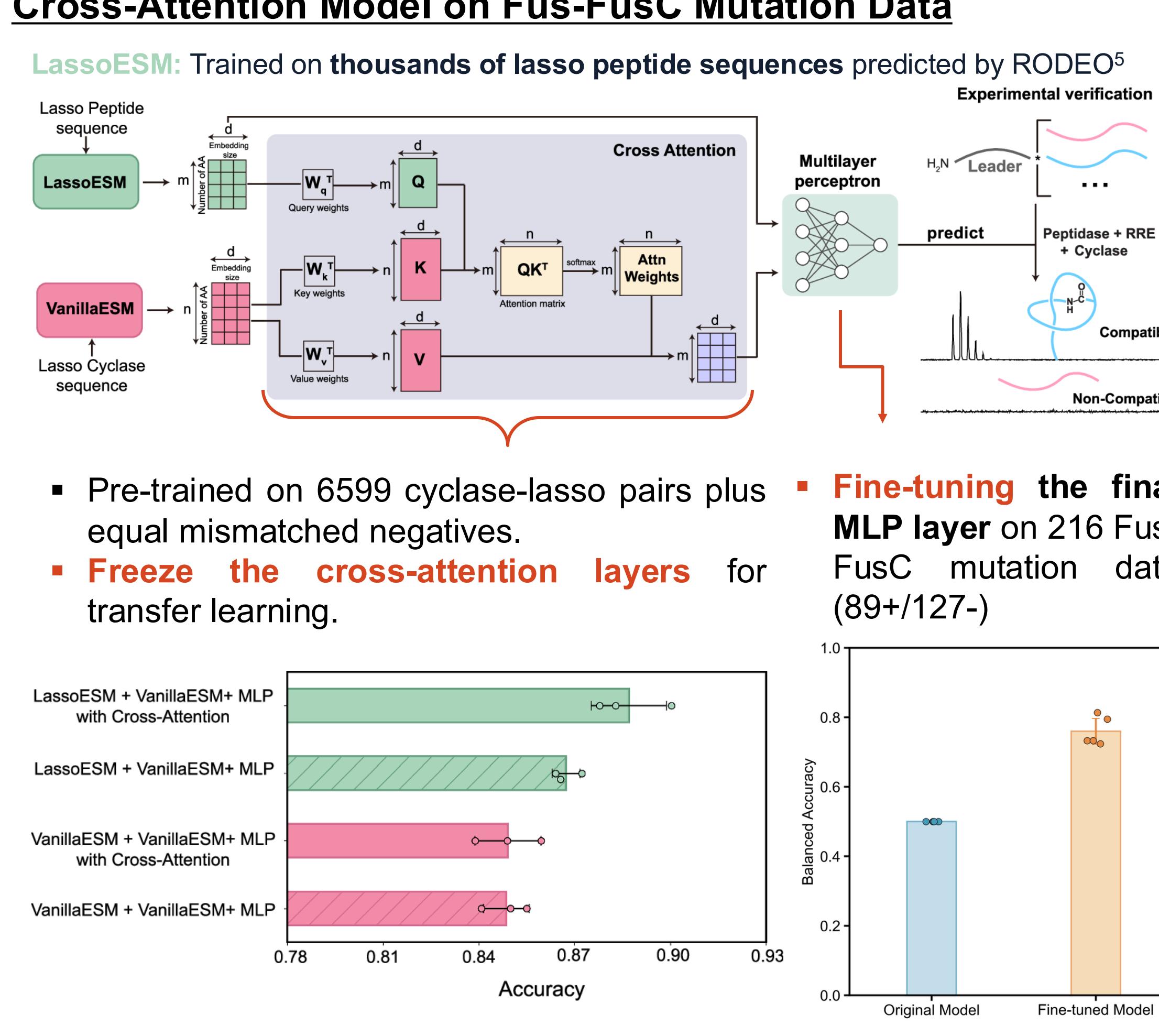


### Deep Learning MSM guided kinetics understanding

- Deep Learning based Markov State Model (MSM) VAMPnet was trained on MD trajectories to maximize the VAMP score, which assesses feature effectiveness in describing slow kinetic transitions and autocorrelation of metastable states in simulations.
- Applying Transition Path Theory (TPT) on the constructed MSM, we established two transition pathways and studied corresponding timescales.



### Transfer Learning: Fine-tuning Pre-trained Cyclase-Substrate Cross-Attention Model on Fus-FusC Mutation Data



- Pre-trained on 6599 cyclase-lasso pairs plus equal mismatched negatives.
- Freeze the cross-attention layers for transfer learning.

- Fine-tuning the final MLP layer on 216 Fus-FusC mutation data (89+/127-)
- Transfer learning from pre-trained general models enables robust prediction of cyclase-substrate tolerance for Fusilassin.

- We integrated MD simulations, TRAM, and VAE to reveal that de novo lasso peptide folding in solution is rare and thermodynamically uphill, identified representative folding pathways, and found loop stability and entropy cost govern formation efficiency, establishing guiding principles for rational engineering of lasso peptides.
- Combining MD simulations with Deep Learning MSM, we characterized lasso peptide bound pose, elucidated a folding after binding mechanism along with kinetics. Transfer Learning from pre-trained general models enables robust prediction of cyclase-substrate tolerance. These findings would guide further cyclase engineering and peptide diversification.
- Collectively, these results provide a comprehensive framework bridging molecular mechanisms and practical design, enabling accelerated discovery and rational optimization of lasso peptide therapeutics.

### Acknowledgment

We thank Dr. Soumajit Dutta, Dr. Prateek Bansal, and Dr. Austin T. Weigle for their help and suggestions, and acknowledge the following funding sources and computational resources for their support.

