ICML
International Conference
On Machine Learning

Chamaileon: Cross-Context Binder Design with Contextualized Modeling and Mixed Sampling

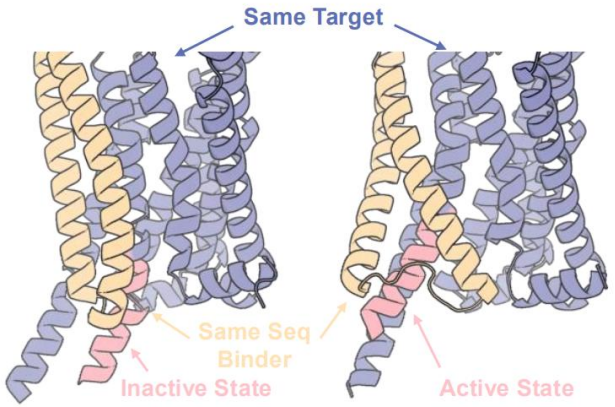
Authors: Hengyuan Cao, Shizhuo Cheng, Mingxuan Liu, Weicheng Huang, Yunhong Lu, Chenxi Cai, Yan Zhang, Min Zhang

Presenter: Hengyuan Cao

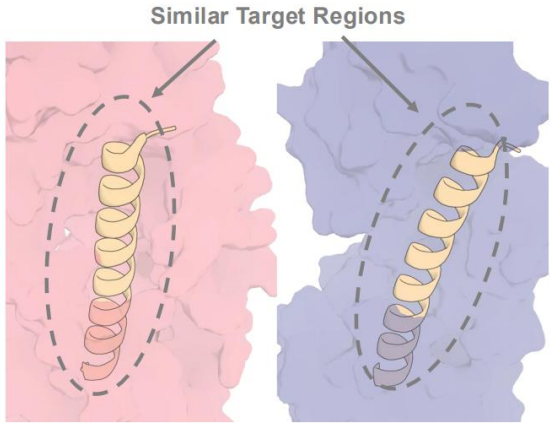
- **The Problem with Current Methods**

  Despite remarkable progress in *de novo* protein binder design, most existing methods still **treat binder design as a one-to-one problem**: optimizing a binder against a single target structure under a single binding objective.

- **However, many real-world scenarios demand designs that go beyond this paradigm:**
  - **Multi-State binder design** — Two targets share an identical sequence but adopt distinct conformations; the binder must accommodate these conformational changes accordingly.
  - **Multi-Target binder design** — Two targets have different sequences; the binder must engage their shared (homologous) epitopes.



Multi-state Binder Design



Multi-target Binder Design

- **A Unified View of Multi-state and Multi-target binder design**

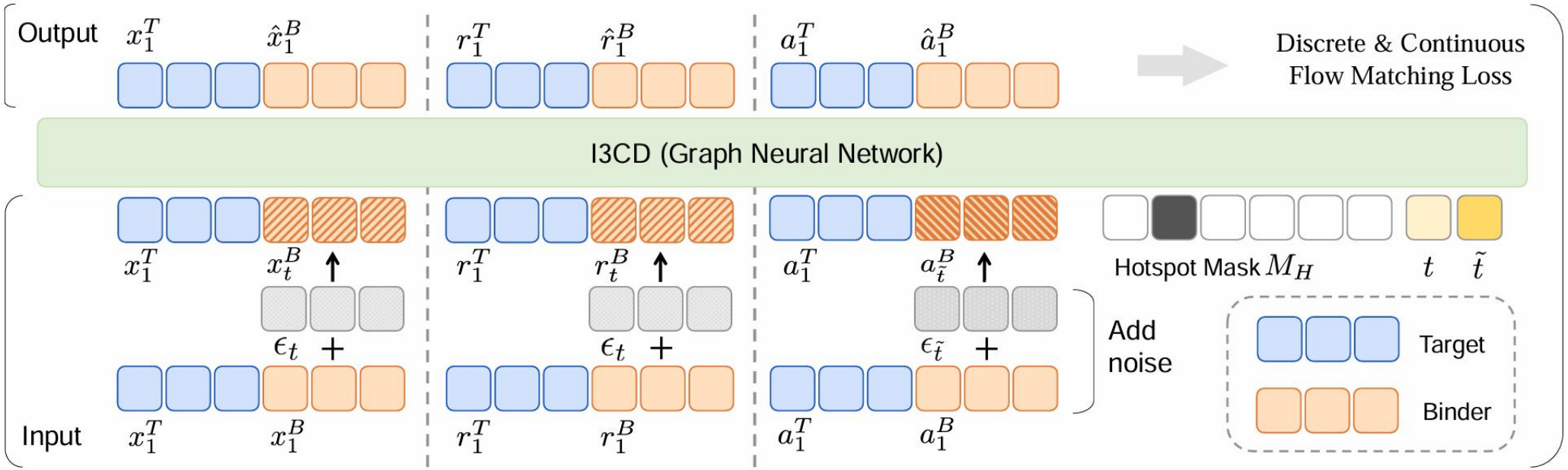
  Multi-state and multi-target design share the same computational structure: **one sequence must satisfy multiple binding constraints with explicit trade-offs**.

- **Cross-Context Binder Design**
  - We unify multi-state and multi-target binder design under what we call **Cross-Context Binder Design**.
  - Here, we define contexts as **the protein interfaces to specificically bind**, regardless of whether they come from different conformational states of the same protein *(Multi-State, MS)* or from distinct targets *(Multi-Target, MT)*.

- **The Fundamental Gap in Current Architectures**
  - **Decoupled noise schedules for sequence and structure** to maintain sequence consistency across conformational states during joint target-binder modeling.
  - **Inference-time optimization** to navigate the trade-offs between conflicting structural constraints.
  - **Comprehensive evaluation** that measures success across a contextual ensemble rather than on isolated snapshots.

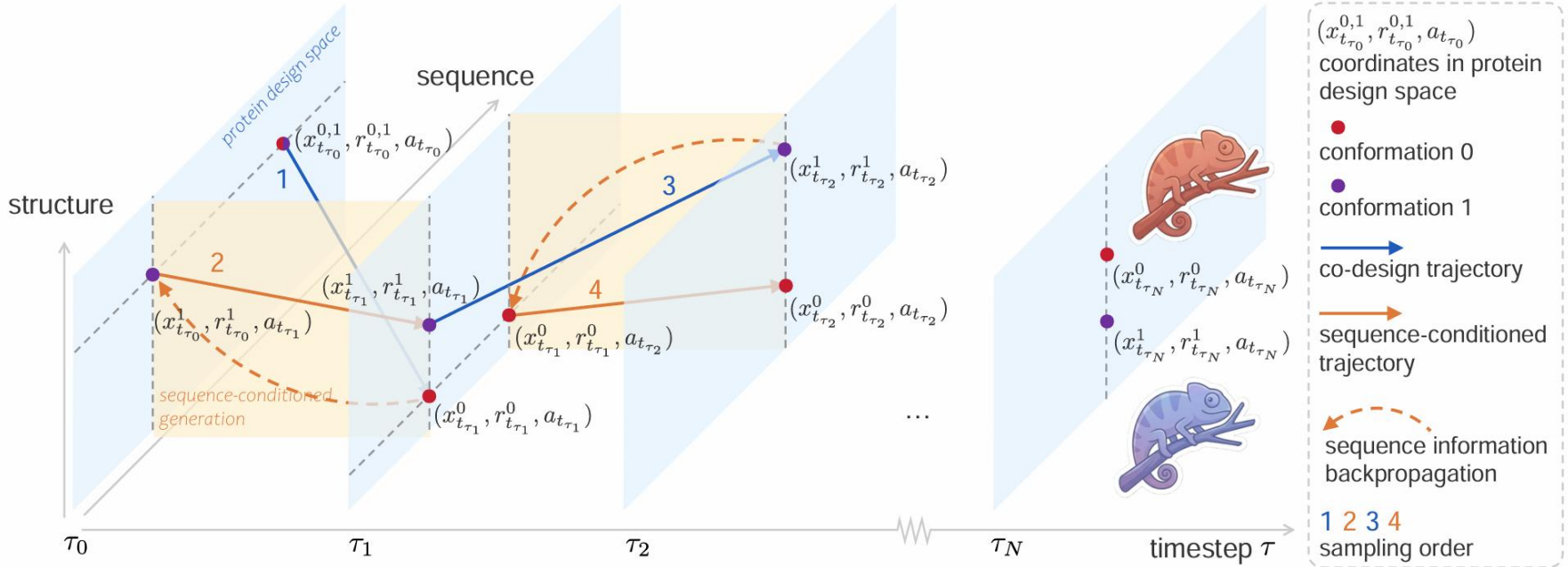
Method ▶ In-Context Complex Co-Design (I3CD)

- **Training Pipeline of I3CD**
  - In-context generation
  - Decoupled noise schedules for sequence and structure
  - Single-state binder codesign framework



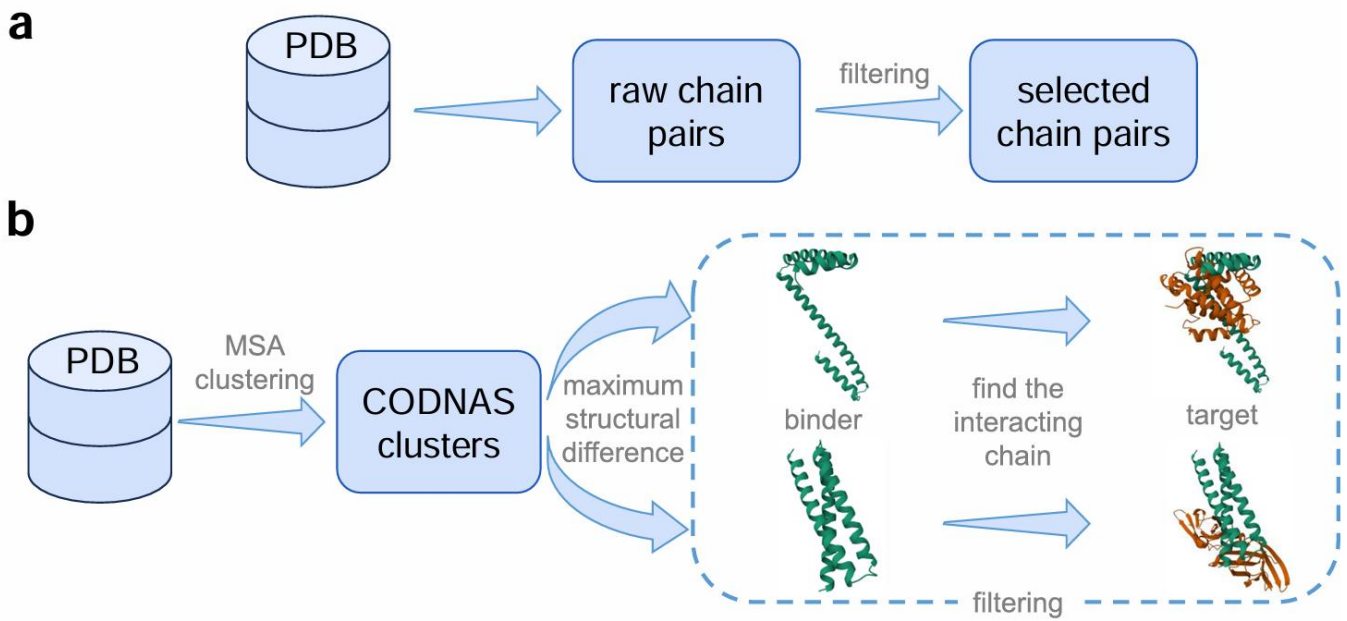
Method ▶ Mixture-of-Paths Sampling (MoPS)

- **Why not map multiple targets to a single binder sequence with multiple conformations?**
  - Scarcity of multi-state training data.
  - Limited flexibility in handling variable conformations
- **MoPS for Cross-Context Binder Design**
  - Sequence and structure co-design
  - Sequence-conditioned generation

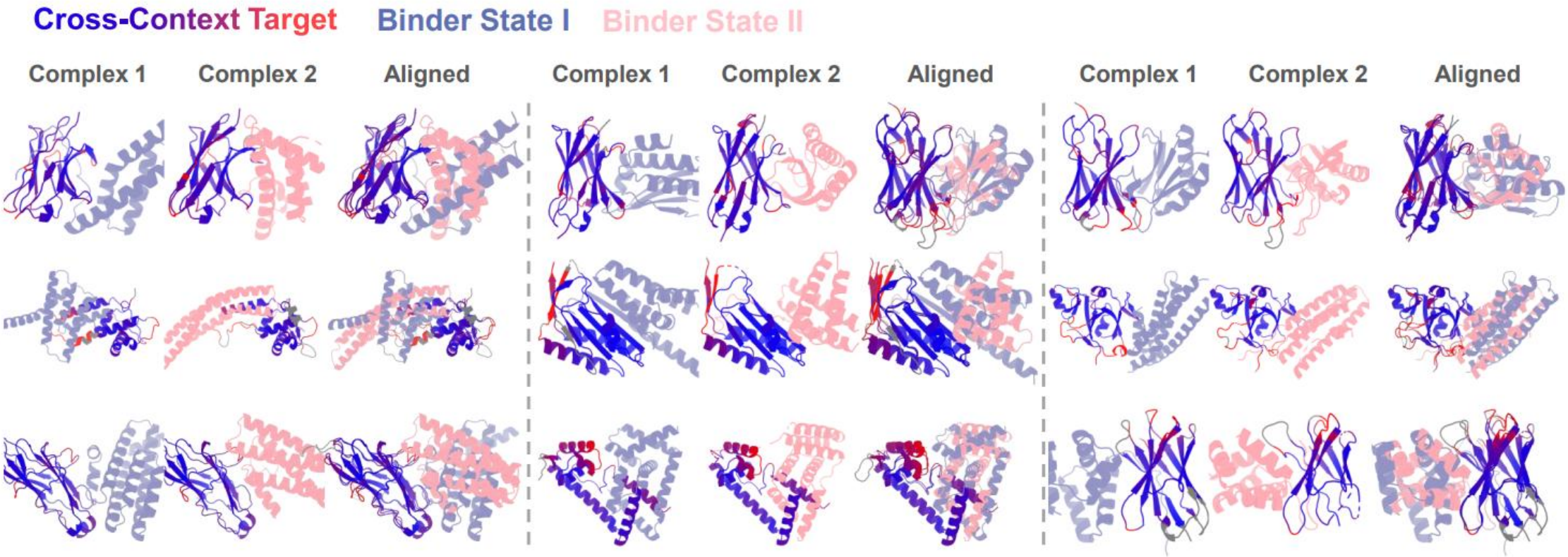


Method ▶ (Multi-State) Complex Data Collection

- **Training Set Construction**
  - Find raw chain pairs.
  - Filtering.
- **Cross-Context Binder Design Benchmark**
  - CROSS (Comprehensive Recognition of Specific Surfaces)
  - Find multi-state binders.
  - Finde corresponding targets.
  - Filtering.



Results ▶ Qualitative results



Results ▶ Quantitative results

- **Ablation Studies**

		Conformation 0					Conformation 1					Both Success
		ipAE	binder pLDDT	binder scRMSD	Unique Success	novelty	ipAE	binder pLDDT	binder scRMSD	Unique Success	Novelty	
Module Ablation	w/o MoPS	6.26	82.9	1.83	**17**	0.500	6.27	82.9	2.49	2	0.863	2
	w/o beam search	5.04	88.2	**1.64**	5	**0.363**	6.18	83.9	**1.76**	8	**0.304**	5
	full version	**4.32**	**90.6**	1.96	8	0.437	**4.23**	**88.7**	2.07	**10**	0.588	**7**
Beam Search Candidate Number	1	5.27	88.7	2.39	7	0.385	4.78	89.5	1.77	6	**0.297**	5
	2	**4.15**	88.8	**1.10**	7	**0.381**	4.36	**89.8**	**1.59**	7	0.461	5
	4	4.32	**90.6**	1.96	**8**	0.437	**4.23**	88.7	2.07	**10**	0.588	**7**
Beam Search Frequency	250	**4.23**	89.8	**1.45**	9	0.494	4.97	87.1	1.78	**11**	0.447	**8**
	100	5.02	89.2	1.90	**10**	**0.431**	4.94	**89.6**	**1.70**	9	**0.266**	**8**
	50	4.32	**90.6**	1.96	8	0.437	**4.23**	88.7	2.07	10	0.588	7
MoPS Frequency	250	-	-	-	0	-	5.30	84.8	1.70	**19**	**0.310**	0
	100	5.64	86.4	1.88	**17**	0.473	**2.21**	**91.7**	**1.11**	1	0.928	1
	50	**2.98**	**92.3**	2.19	1	0.903	4.93	87.5	2.12	15	0.546	1
	20	5.14	88.1	**1.81**	13	**0.398**	4.10	90.1	2.07	9	0.510	**7**
	10	4.32	90.6	1.96	8	0.437	4.23	88.7	2.07	10	0.588	**7**

- **Comparison with Baselines**

Method	Conformation 0				Conformation 1				Both Success
	ipAE	binder pLDDT	binder scRMSD	Unique Success	ipAE	binder pLDDT	binder scRMSD	Unique Success	
Baseline 1 (best)	4.85	84.2	15.5	0	4.73	85.3	13.5	0	0
Baseline 2 (best)	19.1	67.7	**1.37**	0	18.2	68.1	3.20	0	0
`Chamaileon` (mean)	**4.32**	**90.6**	1.96	**8**	**4.23**	**88.7**	**2.07**	**10**	**7**

Conclusion

- **Chamaileon is a unified framework containing I3CD, MoPS, and CROSS.**
- **Cross-context modeling represents a pivotal step toward the “programmable”design of advanced modulatory effects and multi-specific therapeutics.**

Thanks for listening

Contact:
Hengyuan Cao 曹恒源
E-mail: caohy@zju.edu.cn
Wechat: chy18703710871