

CORAL: Learning Amyloid Fibril Ligand Docking with Cooperative Binding Rewards

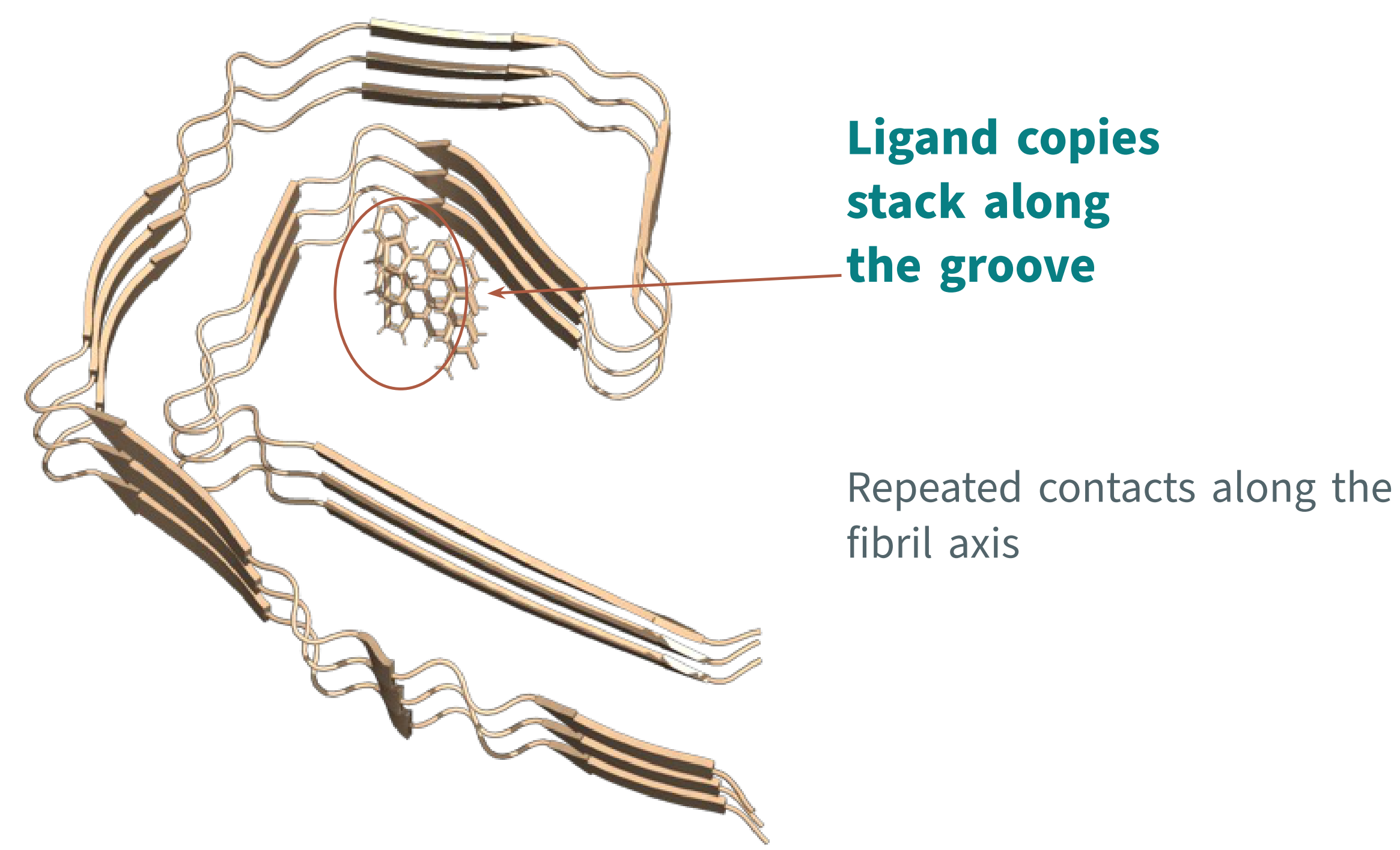
Yasheng Sun^{1*†} Bohan Li^{2*} Youqi Tao^{3†} Jürgen Schmidhuber¹

¹Center of Excellence for Generative AI, KAUST ²Department of Biomedical Engineering, National University of Singapore
³Stern Laboratory, Brigham and Women’s Hospital, Harvard University *Equal contribution †Corresponding authors



01 Cooperative binding

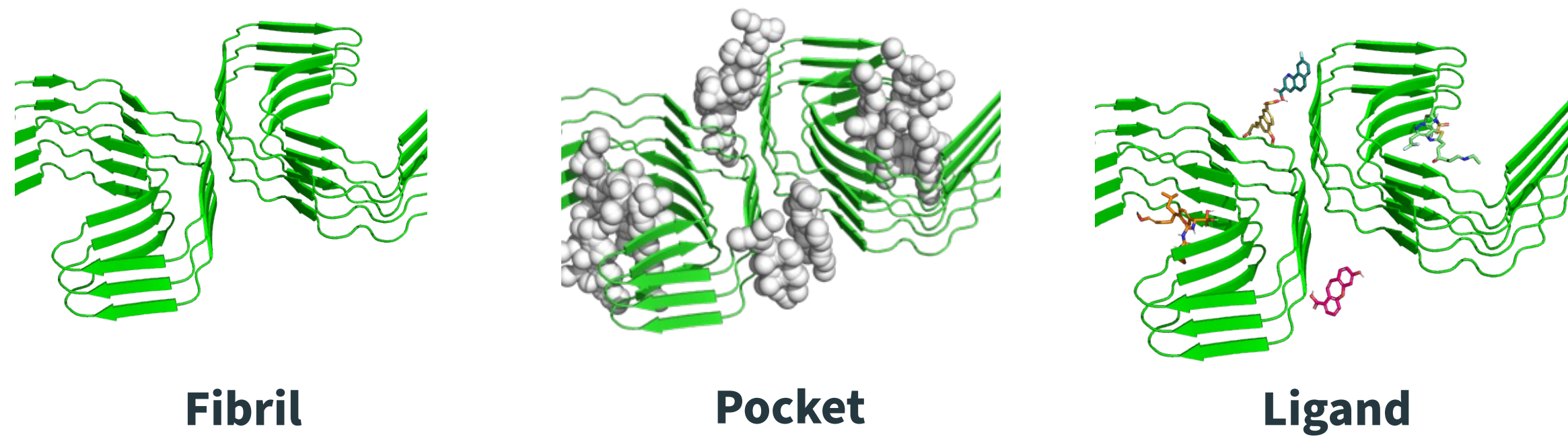
Amyloid grooves bind multiple ligand copies. Protein–ligand and ligand–ligand contacts jointly determine their arrangement.



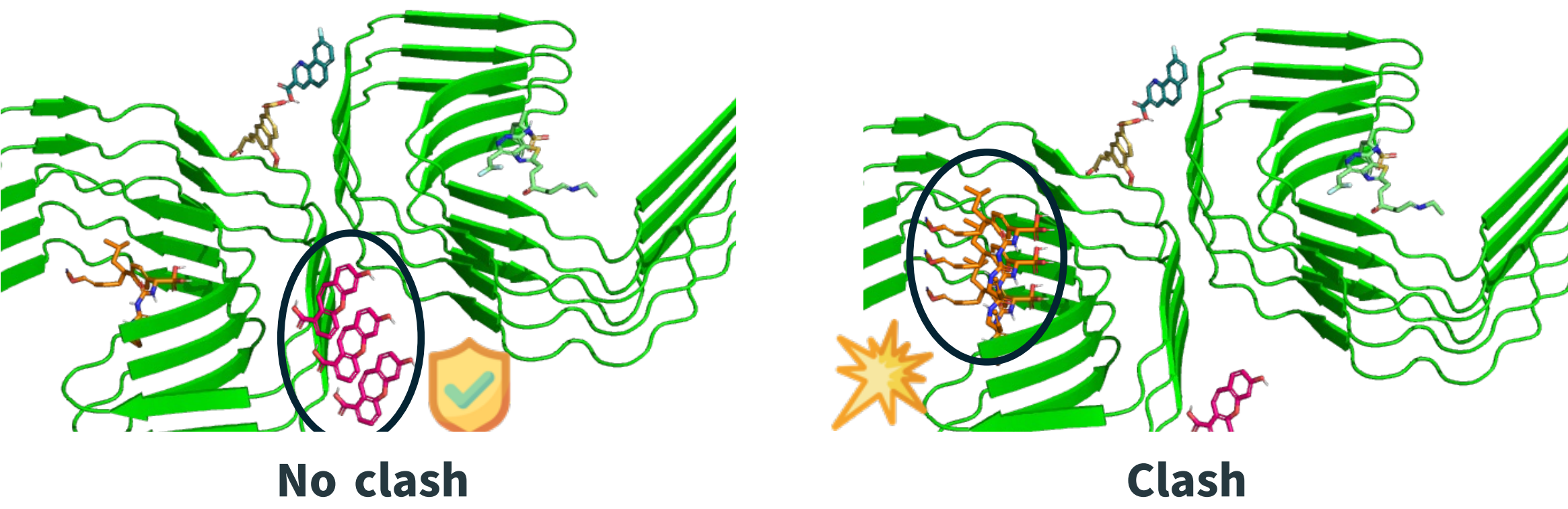
Curated data

≈2,000 warm-up poses

Geometric training examples



Symmetry and clash checks

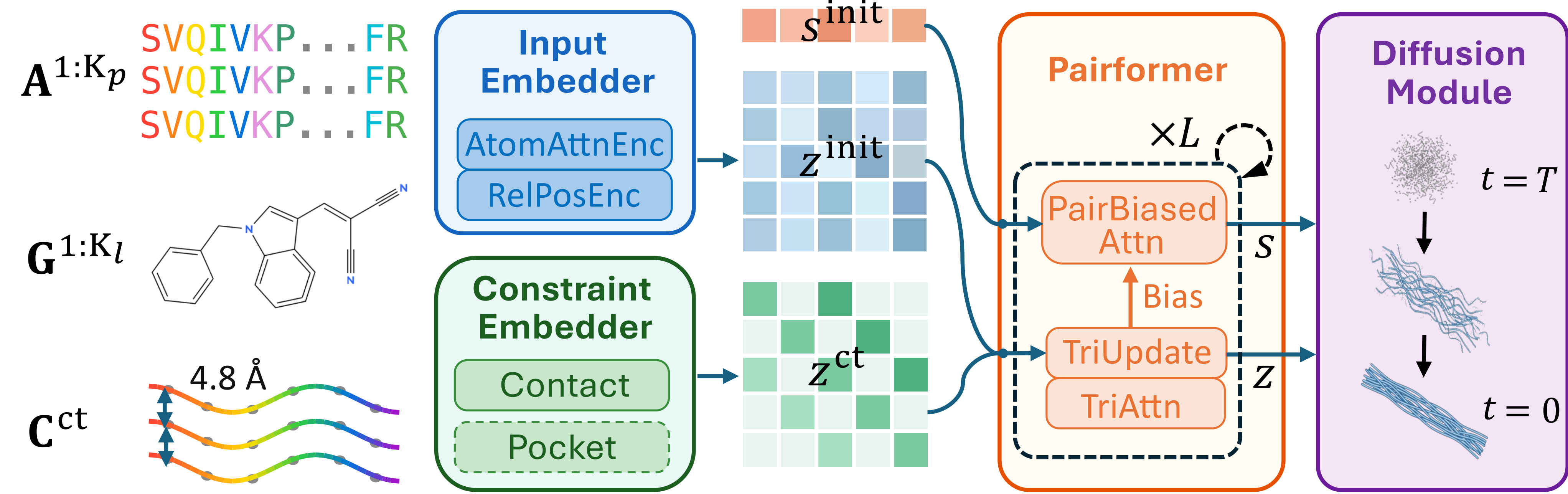


Screen for molecular validity and binding score; review with experts.

Evaluation: 18 CryoEM complexes and 57 Synthetic pairs.

02 CORAL framework

Jointly generate all peptide chains and ligand copies, conditioned on fibril periodicity and the binding pocket.



Learning in two stages

1 Geometric warm-up

Learn periodic docking from curated stacking poses.

$$\mathcal{L}_{\text{SFT}} = \alpha_{\text{diff}} \mathcal{L}_{\text{diff}} + \alpha_{\text{dist}} \mathcal{L}_{\text{dist}} + \alpha_{\text{conf}} \mathcal{L}_{\text{conf}}$$

$\mathcal{L}_{\text{diff}}$: denoise atomic coordinates

$\mathcal{L}_{\text{dist}}$: predict pairwise distances

$\mathcal{L}_{\text{conf}}$: estimate structural confidence

Initialises the policy and the frozen warm-up reference.

2 Cooperative reinforcement learning

Optimise binding, stacking and symmetry with DiffusionNFT.

$$\mathcal{L}_{\text{RL}} = \mathcal{L}_{\text{NFT}} + \beta_{\text{KL}} \mathbb{E} \|v_{\theta} - v_{\text{ref}}\|_2^2, \\ \mathcal{L}_{\text{NFT}} = \mathbb{E} \left[r \|v_{\theta}^+ - \mathbf{v}\|_2^2 + (1 - r) \|v_{\theta}^- - \mathbf{v}\|_2^2 \right].$$

$$v_{\theta}^{\pm} = (1 \mp \beta) v_{\text{old}} \pm \beta v_{\theta}$$

v_{old} : sampling policy; $\mathbf{v}$: target velocity.

$r \in [0, 1]$: group-normalised reward.

v_{ref} : frozen warm-up model; KL-style velocity anchor.

Training ablations

Synthetic medians; lower is better. Energy is a computed score.

Variant	RMSD (Å)	ΔG_{eff} (kcal/mol)
CORAL (full)	3.84	−2.44
No stacking term	4.46	−0.95
No symmetry term	4.31	−2.21
No stability regulariser	8.13	−1.45
Warm-up only (no RL)	5.18	−1.58
RL only (no warm-up)	9.27	−0.88

03 Docking results

CryoEM	Synthetic
50.0%	68.4%
Hit@5 / 4.89 Å median RMSD Best baseline Hit@5: 5.6%	Hit@5 / 3.84 Å median RMSD Best baseline Hit@5: 3.5%

Method	CryoEM		Synthetic	
	RMSD ↓	Hit@5 ↑	RMSD ↓	Hit@5 ↑
DiffDock	13.68	0.0%	19.68	3.5%
DynamicBind	29.02	0.0%	17.93	1.8%
FABind	24.90	0.0%	26.87	0.0%
NeuralPLexer	19.00	5.6%	16.56	0.0%
FlowDock	21.52	0.0%	19.88	0.0%
Boltz-2	28.67	0.0%	33.87	0.0%
Chai-1	19.92	0.0%	16.52	0.0%
CORAL	4.89	50.0%	3.84	68.4%

Median heavy-atom RMSD (Å); Hit@5: RMSD ≤ 5 Å.
Best predicted pose per complex.

Binding poses

Tan: reference; colour: prediction. Views vary.

