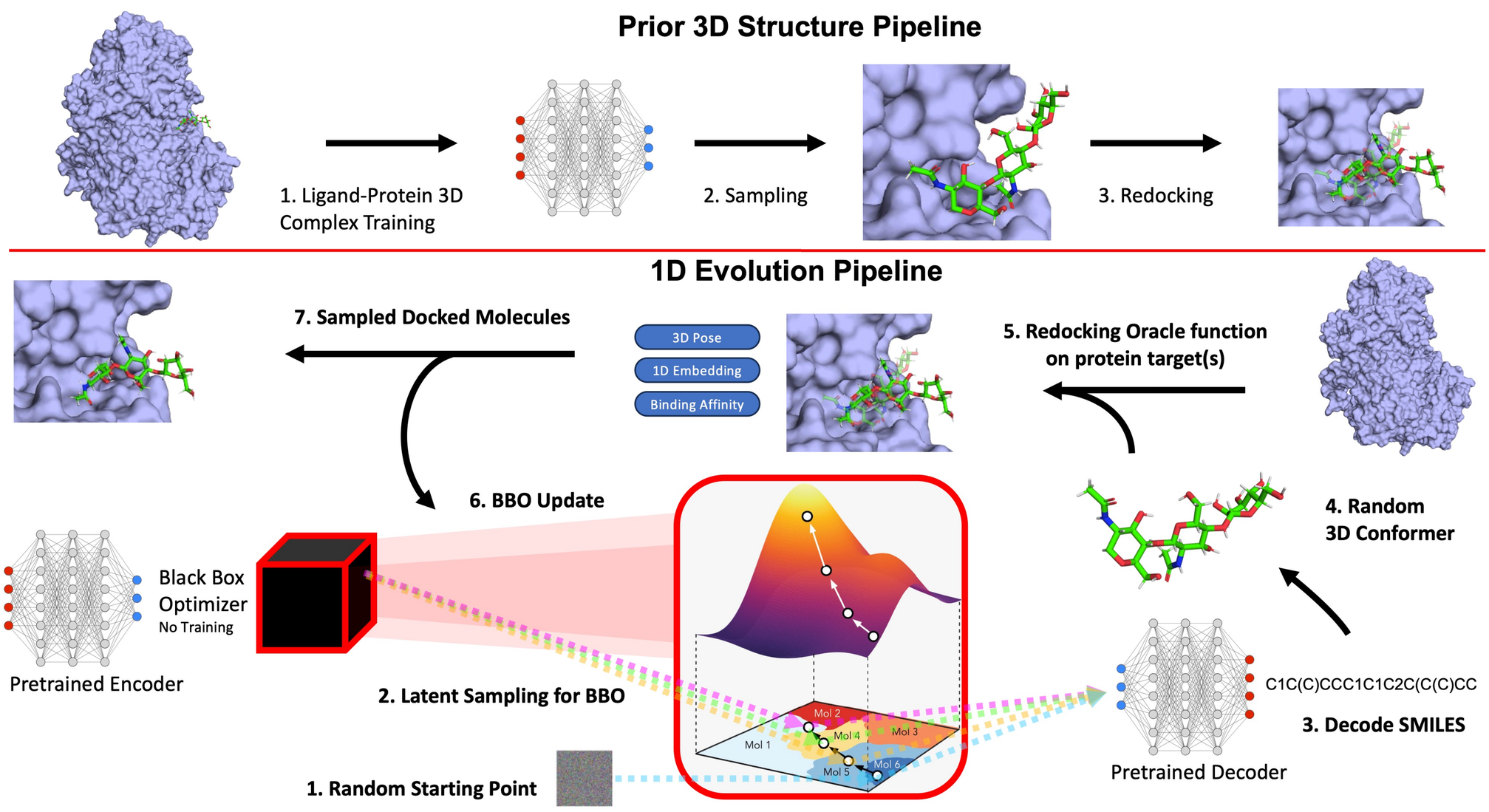
EvoSBDD: Latent Evolution for Accurate and Efficient Structure-Based Drug Design

Danny Reidenbach

Overview



BioNeMo Playground

- 1D Latent Optimization is more accurate and efficient compared to 3D generation for re-docking objectives. Initial generated poses are often poor or no better than random 3D conformers allowing 1D optimization to circumvent the quality and performance bottlenecks.
- Compatible with any pretrained molecule representation. We experimented with both a LLM (MolMIM) and GNN (MoFlow).
- All pretrained molecule autoencoders were leveraged from local single GPU instances of BioNeMo, a training and inference platform see ai.nvidia.com.
- Generation is guided via Black-Box Optimization with efficient AutoDock Vina oracle functions ensuring no additional training or protein information known to the generative model.
- By constraining all structural variation to efficient 2D molecule generation (SMILES), we cover much more of the molecule space, guiding the search via GPU accelerated docking oracles (UniDock) while also initializing with pure random noise.
- Oracle functions are not restricted to docking, adding QED and SA oracles help create drug-like synthesizable strong binders.
- **EvoSBDD achieves state-of-the-art results on the CrossDocked2020 and PoseCheck clash and strain energy benchmarks.**

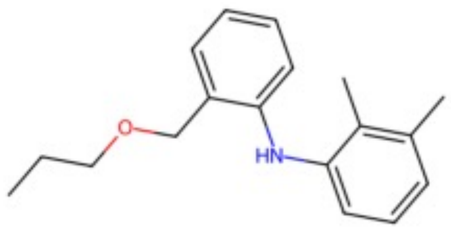
Structure-Based Drug Design Benchmarks

		Validity (↑)	Vina Dock (↓)		High Affinity (↑)		QED (↑)		SA (↑)		Diversity (↑)		Lipinski (↑)	Success Rate (↑)	Time (↓)
		Avg.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg. ± Std.	Avg.	Gen + Score.
	Reference	100%	-7.45	-7.26	-	-	0.48	0.47	0.73	0.74	-	-	4.34 ± 1.14	25.0%	300
Generative	liGAN	-	-6.33	-6.20	21.1%	11.1%	0.39	0.39	0.59	0.57	0.66	0.67	-	3.9%	-
	GraphBP	-	-4.80	-4.70	14.2%	6.7%	0.43	0.45	0.49	0.48	**0.79**	**0.78**	4.83 ± 0.37	0.1%	310
	AR	92.95%	-6.75	-6.62	37.9%	31.0%	0.51	0.50	0.63	0.63	0.70	0.70	4.78 ± 0.51	7.1%	19959
	Pocket2Mol	98.31%	-7.15	-6.79	48.4%	51.0%	0.56	0.57	0.74	0.75	0.69	0.71	4.93 ± 0.27	24.4%	2804
	TargetDiff	90.35%	-7.80	-7.91	58.1%	59.1%	0.48	0.48	0.58	0.58	0.72	0.71	4.59 ± 0.83	10.5%	3728
	DiffSBDD	85.01%	-8.03	-7.74	55.3%	56.6%	0.47	0.47	0.55	0.56	0.76	0.76	4.70 ± 0.64	6.0%	460
	DecompDiff	71.96%	-8.39	-8.43	64.4%	71.0%	0.45	0.43	0.61	0.60	0.68	0.68	4.29 ± 0.97	24.5%	6489
Optimization	TacoGFN	99.27%	-7.41	-7.50	58.9%	59.0%	**0.68**	**0.72**	0.78	0.79	0.65	0.65	4.94 ± 0.24	29.9%	303
	TacoGFN-AL	99.28%	-7.68	-7.70	64.3%	64.0%	0.64	0.66	**0.81**	**0.82**	0.66	0.66	4.93 ± 0.25	36.6%	303
	RGA	-	-8.01	-8.17	64.4%	89.3%	0.57	0.57	0.71	0.73	0.41	0.41	-	46.2%	-
	TargetDiff+Opt (ICLR24)	-	-8.30	-8.15	71.5%	95.9%	0.66	0.68	0.68	0.67	0.31	0.30	-	25.8%	>3728
	DecompOpt (ICLR24)	-	-8.98	-9.01	73.5%	93.3%	0.48	0.45	0.65	0.65	0.60	0.61	-	52.5%	9241
Ours	**EvoSBDD** ($\alpha = 0, \sigma = 1$, 8R)	**100%**	**-9.09**	**-9.20**	**82.1%**	**100%**	0.65	0.67	0.78	0.79	0.65	0.66	**4.96 ± 0.21**	**73.5%**	360
	EvoSBDD ($\sigma = 1.3$, 140R)	**100%**	**-10.27**	**-10.36**	**96.5%**	**100%**	0.53	0.52	0.75	0.77	0.63	0.63	4.84 ± 0.44	**78.8%**	6300
	EvoSBDD ($\alpha = 0, \sigma = 1$, 140R)	**100%**	**-10.14**	**-10.27**	**94.4%**	**100%**	0.59	0.59	0.77	0.77	0.62	0.62	4.91 ± 0.29	**86.4%**	6300

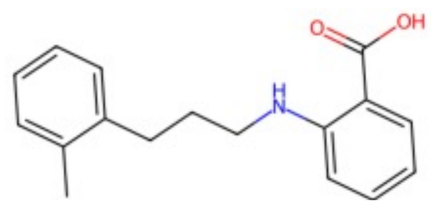
Modular Oracle Functions Enables Multi-Target Design

	COX-1 Binding Affinity (↓)	COX-2 Binding Affinity (Chain G) (↓)	COX-2 Binding Affinity (Chain I) (↓)	COX-2 Binding Affinity (Chain E) (↓)
COX-1 Reference (Chain E)	-4.411	-3.348	-3.301	-5.454
COX-2 Reference (Chain G)	-4.986	-4.919	-4.432	-4.565
COX-2 Reference (Chain I)	-5.372	-3.615	-4.984	-6.008
COX-2 Reference (Chain E)	-6.012	-5.630	-6.777	-8.897

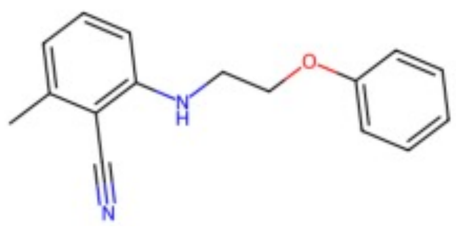
COX-1 Affinity: -5.402
COX-2 Affinity (E): -9.418



COX-1 Affinity: -5.323
COX-2 Affinity (E): -8.628



COX-1 Affinity: -5.200
COX-2 Affinity (E): -8.043



COX-1 Affinity: -5.000
COX-2 Affinity (E): -7.749

- Prior structure-based methods are incapable of multi-target optimization due to single protein training/inference
- Inflammation-reducing medications targeting COX-2 can cause gastrointestinal bleeding if they bind to COX-1, making specificity crucial.