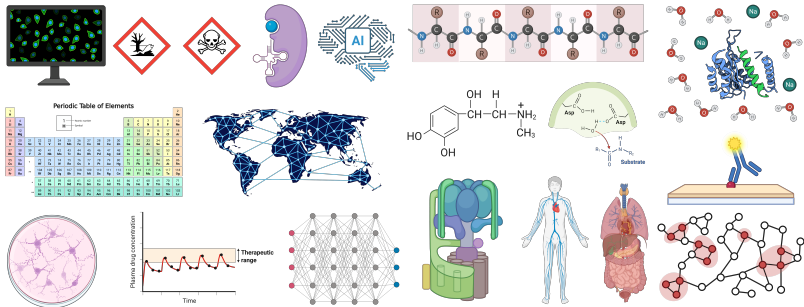




Chemical Intuition and AI for Accelerated Molecule Design: Our Experiences

Nirav Pravinbhai Bhatt

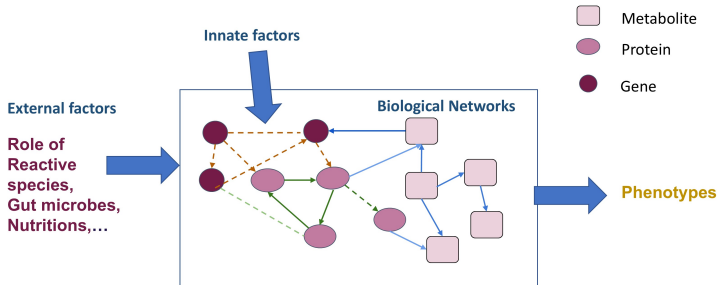
BioSystems Engineering and ConTrol (BiSECT) Group
Wadhvani School of Data Science and AI
Core Faculty, IITM Zanzibar Campus, IIT Madras, India



August 31, 2026,

Roadmap for Accelerated Materials Design: From Generative AI to Self-driving Labs, ICTS, Bengaluru

Research Outlook



How to better understand biological networks:

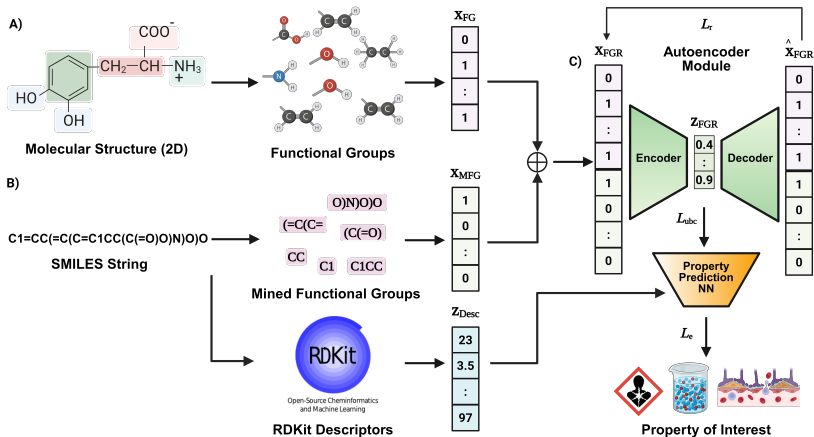
Inputs (innate and external factors) and outputs (phenotypes) for Biological Processes (diseases) from multi-model data using Control Theory and AI?

How can biological and Systems networks be manipulated to obtain a desired output at the Population/Patient (Individual) levels?

Molecule Discovery and Interventions At BiSECT Lab

- *Chemically inspired Molecular Property Prediction*
- *Small De Novo Molecule generation models*
- *AI Single and Multi-step Retrosynthesis planner*
- Drug Repurposing using Integrated Biological Networks
- Creating Better Benchmarks for Drug Discovery
- Cost-effective Manufacturing through Flow Chemistry and Automated Flow Reactors

Molecular Property Prediction: Pipeline



Functional Groups are All you Need!¹

¹Balaji, Bobby, and Bhatt, *Journal of Chemical Information and Modeling* (2026).

Molecular Property Prediction: Benchmark Datasets

Dataset	Description	Tasks	Compounds
BACE	Inhibition of Human-secretase 1 (BACE-1)	1	1,513
BBBP	Blood-Brain Barrier Penetration ability	1	2,039
Tox21	Toxicity in the 21st Century	12	7,831
SIDER	Side effects of drugs	27	8,576
<i>E.coli</i>	Growth inhibition of <i>Escherichia coli</i>	1	2,335
A2780	Ovarian carcinoma cells	1	2,255
DU-145	Prostate carcinoma cells	1	2,512
Schilling	Cleavage of octamers in peptide by HIV-1 protease	1	3,272
ESOL	Molecule water solubility	1	1,128
Freesolv	Hydration energy of molecule	1	642
Lipo	Lipophilicity	1	4,200

Table 1: Description of datasets used for property prediction

Molecular Property Prediction: Results Classification

Category	Physiology					Biophysics		
Dataset	BBBP $\uparrow$	Tox21 $\uparrow$	ToxCast $\uparrow$	SIDER $\uparrow$	ClinTox $\uparrow$	BACE $\uparrow$	MUV $\uparrow$	HIV $\uparrow$
Molecules	2,039	7,831	8,575	1,427	1,478	1,513	93,807	41,127
Tasks	1	12	617	27	2	1	17	1
DMPNN ¹	91.9 $\pm$ 3.0	75.9 $\pm$ 0.7	63.7 $\pm$ 0.2	57.0 $\pm$ 0.7	90.6 $\pm$ 0.6	85.2 $\pm$ 0.6	78.6 $\pm$ 1.4	77.1 $\pm$ 0.5
GEM ²	88.8 $\pm$ 0.4	78.1 $\pm$ 0.4	68.6 $\pm$ 0.2	63.2 $\pm$ 1.5	90.3 $\pm$ 0.7	87.9 $\pm$ 1.1	75.3 $\pm$ 1.5	81.3 $\pm$ 0.3
KANO ³	96.0 $\pm$ 1.6	83.7 $\pm$ 1.3	<u>73.2 $\pm$ 1.6</u>	<u>65.2 $\pm$ 0.8</u>	<u>94.4 $\pm$ 0.3</u>	93.1 $\pm$ 2.1	83.7 $\pm$ 2.3	85.1 $\pm$ 2.2
FG#	93.9 $\pm$ 2.8	78.3 $\pm$ 1.2	72.7 $\pm$ 0.6	60.5 $\pm$ 1.7	88.8 $\pm$ 7.6	87.2 $\pm$ 2.4	72.3 $\pm$ 1.3	77.5 $\pm$ 2.7
MFG#	88.4 $\pm$ 1.5	67.5 $\pm$ 0.7	63.0 $\pm$ 1.2	56.0 $\pm$ 0.7	69.7 $\pm$ 5.4	84.9 $\pm$ 4.2	68.6 $\pm$ 2.5	76.7 $\pm$ 4.8
FGR#	96.0 $\pm$ 1.8	84.1 $\pm$ 1.0	74.0 $\pm$ 2.1	67.8 $\pm$ 2.8	96.1 $\pm$ 0.5	<u>89.3 $\pm$ 3.1</u>	74.2 $\pm$ 4.1	78.3 $\pm$ 1.1

Table 2: The mean and standard deviation of test ROC-AUC (%) on three independent runs are reported. The dataset split is based on molecular scaffolds. Bold indicates the best performing model and underline indicates the second best performing model. # indicates the concatenation of 2D descriptors.

¹K. Yang et al., *Journal of Chemical Information and Modeling* (2019).

²X. Fang et al., *Nature Machine Intelligence* (2021).

³Y. Fang et al., *Nature Machine Intelligence* (2023).

Molecular Property Prediction: Results Regression

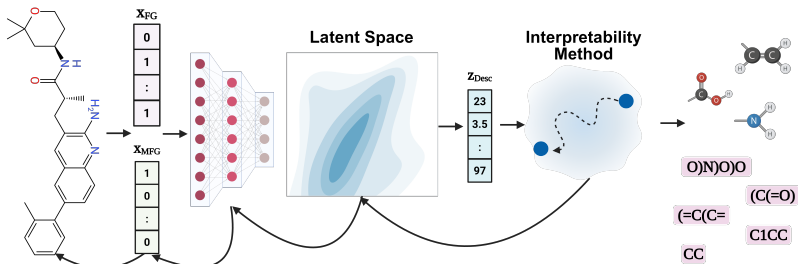
Category	Physical Chemistry			Quantum Mechanics		
Dataset	ESOL ↓	FreeSolv ↓	Lipophilicity ↓	qm7 ↓	qm8 ↓	qm9 ↓
Molecules	1,128	642	4,200	7,160	21,786	133,885
Tasks	1	1	1	12	3	1
DMPNN	1.050 ± 0.008	1.673 ± 0.082	0.683 ± 0.016	103.5 ± 8.6	0.0156 ± 0.001	0.00514 ± 0.00001
GEM	0.813 ± 0.028	1.748 ± 0.114	0.674 ± 0.022	60.0 ± 2.7	0.0163 ± 0.001	0.00562 ± 0.00007
KANO	<u>0.670 ± 0.019</u>	1.142 ± 0.258	0.566 ± 0.007	56.4 ± 2.8	0.0123 ± 0.000	0.00320 ± 0.00001
FG#	0.763 ± 0.071	<u>0.825 ± 0.221</u>	0.742 ± 0.050	59.2 ± 1.7	0.0335 ± 0.003	0.00690 ± 0.00005
MFG#	0.812 ± 0.083	1.034 ± 0.100	0.757 ± 0.025	61.6 ± 1.9	0.0351 ± 0.003	0.00730 ± 0.00010
FGR#	0.620 ± 0.067	0.789 ± 0.192	<u>0.636 ± 0.027</u>	55.3 ± 1.6	0.0297 ± 0.003	0.00547 ± 0.00008

Table 3: The mean and standard deviation of test root mean square error (for ESOL, FreeSolv and Lipophilicity) or mean absolute error (for qm7, qm8 and qm9) on three independent runs are reported. The dataset split is based on molecular scaffolds. Bold indicates the best performing model and underline indicates the second best performing model. # indicates the concatenation of 2D descriptors.

Molecular Property Prediction: Interpretability with Functional Groups

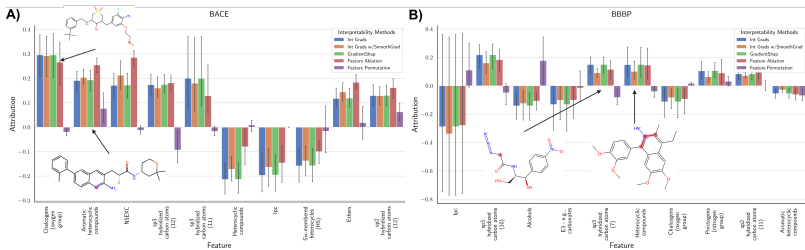
Attribution

- The model should present explanation about its mechanics.
- **Methods:** Integrated Gradients, DeepLift, GradientShap, Noise Tunnel, Feature Ablation.



Molecular Property Prediction: Case Studies

- **Significant Functional Groups:** Pyridine rings¹, Carboxylic acid derivatives²
- **Missing Functional Groups:** Chalcogens, Tetrazoles
- **No evidence:** Six-membered heterocycles, sp^{*} hybridized carbon atoms

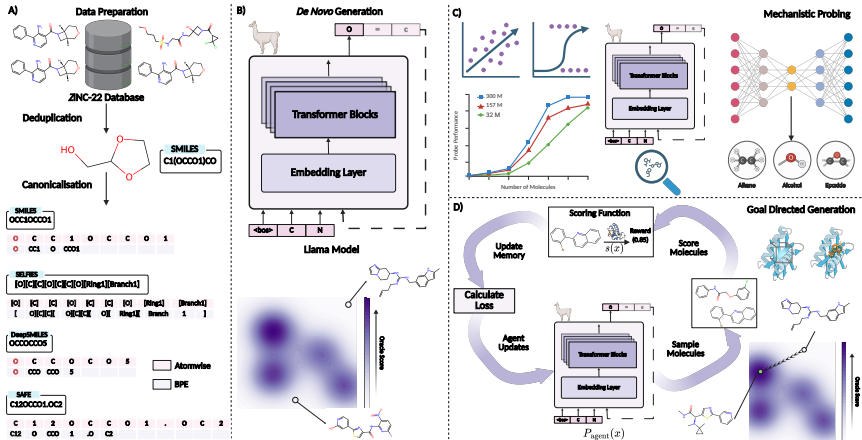


Interpretability analysis on **A)** BACE **B)** BBBP

¹Ghobadian et al., *European Journal of Medicinal Chemistry* (2018).

²Kimura et al., *Bioorganic & medicinal chemistry letters* (2006).

NovoMolGen: De novo Molecule Generation Foundation Models



NovoMolGen: Evaluation

Distribution Learning

- **Validity (↑):** Measures the percentage of generated molecules that are chemically valid.
- **Novelty (↑):** Evaluates the model's ability to generate molecules that are not present in the training set, indicating the model's creativity.
- **FCD (Fréchet Chemnet Distance) (↓):** Assesses how well the generated molecules capture biological properties in reference sets.
- **SNN (Similarity to a nearest neighbor) (↑):** How far the generated molecules from the manifold of the reference set?
- **Scaffold and Fragment Similarity (↑):** Compares distributions of scaffolds and BRICS fragments to the reference sets.

NovoMolGen: Distribution Learning

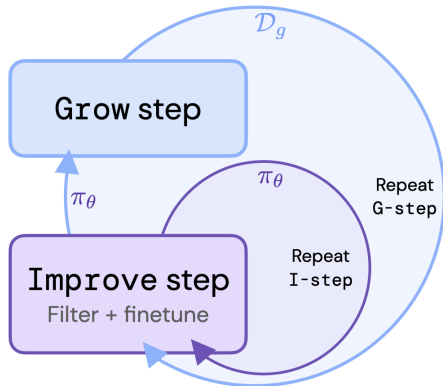
Comprehensive performance metrics for baseline models and **NovoMolGen** (Random validation set). **Blue** denotes the best performing model, while **Pink** represents the second-best performing model (p value < 0.05).

REPRESENTATION	TOKENIZATION	MODEL SIZE	Valid (↑)	IntDiv (↑)	Novelty (↑)	FCD (↓)	SNN (↑)	Frag (↑)	Scaff (↑)
<i>Train</i>	–	–	1.000	1.000	0.0	0.0376	0.4657	0.9999	0.5144
CHARRNN	Atomwise	–	0.9748	0.856	–	0.0732	0.6015	0.9998	0.9242
VAE	Atomwise	–	0.9767	0.855	–	0.0990	0.6257	0.9994	0.9386
JT-VAE	Atomwise	–	1.000	0.855	–	0.3954	0.5477	0.9965	0.8964
LIMO	Atomwise	–	1.000	0.854	–	26.78	0.2464	0.6989	0.0079
MOLGEN-7B	Atomwise	7B	1.000	0.855	–	0.0435	0.5138	0.9999	0.6538
GP-MOLFORMER	Atomwise	46.8M	1.000	0.865	0.390	0.0591	0.5045	0.9998	0.7383
NOVOMOLGEN-SMILES	Atomwise	32M	0.999	0.851	0.983	0.0394	0.4657	0.9999	0.5309
		157M	0.999	0.851	0.981	0.0426	0.4656	0.9999	0.5200
		300M	0.999	0.851	0.981	0.0419	0.4657	0.9999	0.5219
	BPE	32M	0.999	0.851	0.982	0.0384	0.4654	0.9999	0.5182
		157M	0.999	0.852	0.980	0.0384	0.4655	0.9999	0.5193
		300M	0.999	0.851	0.979	0.0380	0.4657	0.9999	0.5173
NOVOMOLGEN-SELFIES	Atomwise	32M	1.000	0.852	0.982	0.0799	0.4651	0.9996	0.4765
	BPE	32M	1.000	0.851	0.984	0.0386	0.4649	0.9999	0.5105
NOVOMOLGEN-DEEPSMILES	Atomwise	32M	0.999	0.851	0.981	0.0400	0.4661	0.9999	0.5218
	BPE	32M	0.997	0.851	0.982	0.0380	0.4655	0.9999	0.5165
NOVOMOLGEN-SAFE	Atomwise	32M	0.957	0.853	0.953	0.9475	0.4562	0.9955	0.4521
	BPE	32M	0.997	0.852	0.976	0.3283	0.4599	0.9972	0.4913

NovoMolGen: Goal Directed Generation

Grow and Improve

- Generative models cannot directly manipulate molecular properties
- Can steer the exploration of the latent space towards generating molecules that exhibit desired properties



Reinforced Self-Training (ReST) for Language Modeling^a

^aGulcehre et al., ArXiv (2023).

NovoMolGen: Goal Directed Generation

Docking Tasks

1. **parp1**: Poly(ADP-ribose) polymerase 1 (PARP1) is a nuclear enzyme involved in DNA repair, transcriptional regulation, and cell survival.
2. **fa7**: Factor VII (FA7), also known as proconvertin, is a coagulation protein essential for blood clotting.
3. **5ht1b**: The 5-hydroxytryptamine receptor 1B (5-HT1B) is a serotonin receptor involved in neurotransmitter regulation.
4. **braf**: B-Raf (serine/threonine-protein kinase B-Raf) is a key regulator of the MAP kinase/ERK signaling pathway, which controls cell proliferation and differentiation.
5. **jak2**: Janus kinase 2 (JAK2) is a tyrosine kinase involved in the JAK/STAT signaling pathway, which regulates hematopoiesis and immune function.

NovoMolGen: Goal Directed Generation

Docking Scores

For each target, we report the **novel top 5% docking score (DS) (kcal/mol)** and the **novel hit ratio (HR) (%)**. Higher hit ratios (↑) and lower docking scores (↓) indicate better performance. **Blue** denotes the best performing model, while **Pink** represents the second-best performing model (p value < 0.05).

Method	parp1		fa7		5ht1b		braf		jak2	
	DS (↓)	HR (%) (↑)	DS (↓)	HR (%) (↑)	DS (↓)	HR (%) (↑)	DS (↓)	HR (%) (↑)	DS (↓)	HR (%) (↑)
REINVENT	-8.70 _(0.52)	0.48 _(0.34)	-7.21 _(0.26)	0.21 _(0.08)	-8.77 _(0.32)	2.45 _(0.56)	-8.39 _(0.40)	0.13 _(0.09)	-8.17 _(0.28)	0.61 _(0.17)
MORLD	-7.53 _(0.26)	0.05 _(0.05)	-6.26 _(0.17)	0.01 _(0.01)	-7.87 _(0.65)	0.88 _(0.74)	-8.04 _(0.34)	0.05 _(0.04)	-7.82 _(0.13)	0.23 _(0.12)
HierVAE	-9.49 _(0.28)	0.55 _(0.21)	-6.81 _(0.27)	0.01 _(0.01)	-8.08 _(0.25)	0.51 _(0.28)	-8.98 _(0.53)	0.21 _(0.22)	-8.29 _(0.37)	0.23 _(0.13)
FREED-QS	-10.58 _(0.10)	4.63 _(0.73)	-8.38 _(0.04)	1.33 _(0.11)	-10.71 _(0.18)	16.77 _(0.90)	-10.56 _(0.08)	2.94 _(0.36)	-9.74 _(0.02)	5.80 _(0.30)
GDSS	-9.97 _(0.03)	1.93 _(0.21)	-7.78 _(0.04)	0.37 _(0.10)	-9.46 _(0.10)	4.67 _(0.31)	-9.22 _(0.07)	0.17 _(0.13)	-8.93 _(0.09)	1.17 _(0.28)
MOOD	-10.87 _(0.11)	7.02 _(0.43)	-8.16 _(0.07)	0.73 _(0.14)	-11.15 _(0.04)	18.67 _(0.42)	-11.06 _(0.03)	5.24 _(0.29)	-10.15 _(0.06)	9.20 _(0.52)
GEAM	-12.89 _(0.16)	40.56 _(0.82)	-9.89 _(0.12)	20.71 _(1.87)	-12.37 _(0.04)	38.48 _(0.35)	-12.34 _(0.10)	27.90 _(1.82)	-11.82 _(0.07)	42.95 _(1.11)
f-RAG	-12.95 _(0.05)	-	-9.90 _(0.21)	-	-12.67 _(0.14)	-	-12.39 _(0.05)	-	-11.84 _(0.32)	-
32M-ATOMWISE	-13.92 _(0.19)	77.13 _(9.40)	-11.11 _(0.36)	74.67 _(5.44)	-12.15 _(0.09)	53.98 _(4.87)	-12.43 _(0.21)	50.36 _(7.54)	-12.33 _(0.67)	79.31 _(8.82)
157M-ATOMWISE	-14.20 _(0.61)	84.58 _(2.31)	-10.77 _(0.25)	65.85 _(9.72)	-12.97 _(0.72)	63.57 _(5.45)	-12.73 _(0.16)	69.97 _(6.85)	-12.27 _(0.29)	77.38 _(8.95)
300M-ATOMWISE	-14.27 _(0.60)	77.73 _(8.39)	-11.20 _(0.26)	82.57 _(2.93)	-13.54 _(0.57)	67.83 _(2.47)	-13.10 _(0.50)	71.64 _(7.83)	-11.97 _(0.12)	81.23 _(1.30)
32M-BPE	-14.90 _(0.36)	90.24 _(1.36)	-11.70 _(0.46)	82.55 _(1.13)	-12.80 _(0.17)	64.92 _(5.51)	-13.80 _(0.20)	68.03 _(3.05)	-12.47 _(0.21)	79.24 _(6.43)
157M-BPE	-14.61 _(0.30)	84.41 _(3.65)	-11.53 _(0.06)	83.20 _(3.02)	-13.45 _(0.23)	64.71 _(3.09)	-13.13 _(0.71)	69.28 _(9.32)	-12.93 _(0.60)	77.35 _(9.94)
300M-BPE	-14.80 _(0.10)	81.46 _(5.78)	-11.67 _(0.49)	83.65 _(7.27)	-13.60 _(0.17)	69.33 _(4.26)	-13.40 _(0.30)	74.87 _(6.73)	-13.07 _(0.16)	84.89 _(3.10)

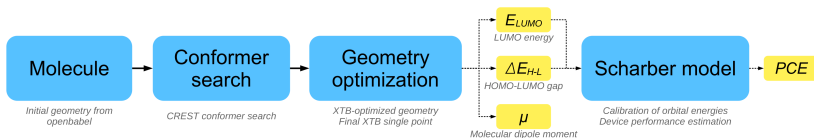
NovoMolGen: Goal Directed Generation

Material Design Tasks

- **Two organic photovoltaic design tasks:**

- Design of donor small molecule for [6,6]-phenyl-C61-butyric acid methyl ester (PCBM) (PCE_{PCBM} - SAscore)
- Design of acceptor small molecule for poly[N-90-heptadecanyl-2,7-carbazole-alt-5,5-(40,70-di-2-thienyl-20,10,30-benzothiadiazole)] (PCDTBT) (PCE_{PCDTBT} - SAscore)

- Optimise for Power Conversion Efficiency (PCE)



Workflow for PCE calculation¹

¹Nigam et al., *Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track* (2023).

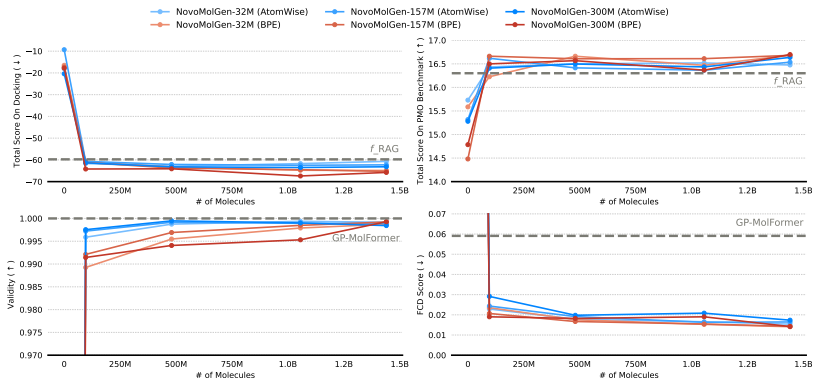
NovoMolGen: Goal Directed Generation

Material Design Tasks

Performance on the organic photovoltaic design benchmarks. Results are reported as the mean $\pm$ standard deviation of the best objective value over five independent runs. **Blue** denotes the best performing model, while **Pink** denotes the second-best performing model.

Method	$PCE_{PCBM} - SAscore \uparrow$	$PCE_{PCDTBT} - SAscore \uparrow$
Dataset	7.57	31.71
SMILES-VAE	7.44 $\pm$ 0.28	10.23 $\pm$ 11.14
SELFIES-VAE	7.05 $\pm$ 0.66	29.24 $\pm$ 0.65
MoFlow	7.08 $\pm$ 0.31	29.81 $\pm$ 0.37
SMILES-LSTM-HC	6.69 $\pm$ 0.40	31.79 $\pm$ 0.15
SELFIES-LSTM-HC	7.40 $\pm$ 0.41	30.71 $\pm$ 1.20
REINVENT	7.48 $\pm$ 0.11	30.47 $\pm$ 0.44
GB-GA	7.78 $\pm$ 0.02	30.24 $\pm$ 0.80
JANUS	7.59 $\pm$ 0.14	31.34 $\pm$ 0.74
32M-ATOMWISE	7.83 $\pm$ 0.05	33.26 $\pm$ 0.21
157M-ATOMWISE	7.90 $\pm$ 0.04	33.74 $\pm$ 0.17
300M-ATOMWISE	7.97 $\pm$ 0.04	34.18 $\pm$ 0.14
32M-BPE	7.86 $\pm$ 0.05	33.44 $\pm$ 0.19
157M-BPE	7.94 $\pm$ 0.04	33.91 $\pm$ 0.16
300M-BPE	8.00 $\pm$ 0.04	34.46 $\pm$ 0.13

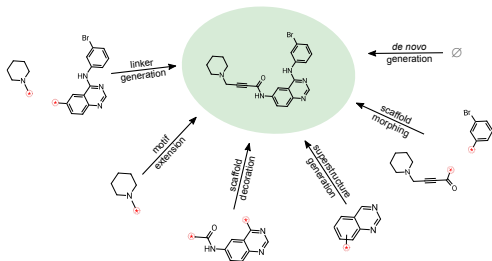
NovoMolGen: Training Progress



Key metrics for models of varying sizes (32M-300M) at intermediate checkpoints against the number of molecules seen. The results consistently show that **performance saturates early, with diminishing returns for extended training.**

What can NovoMolGen do?

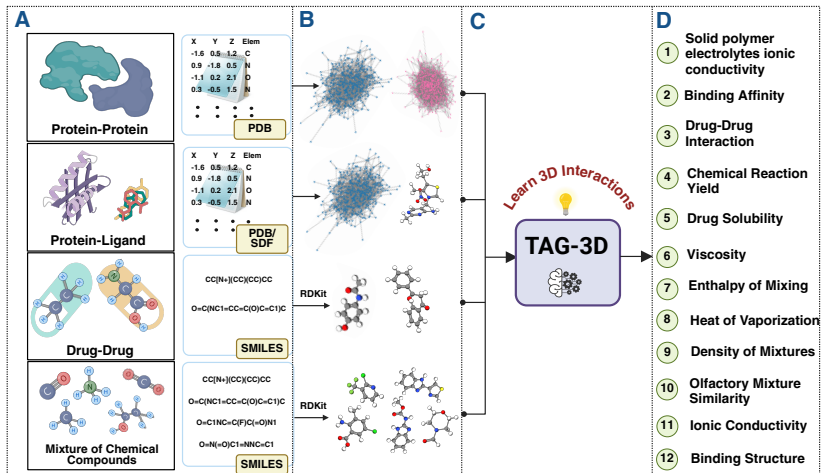
- State-of-the-art deep generative model for *de novo* molecule generation.
- Can perform multiple design tasks.



Fragment constrained generation¹

¹Noutahi et al., *arXiv preprint arXiv:2310.10773* (2023).

Task Agnostic Graph Representation TAG-3D



Can the same minimalist representation support **multi-molecule interaction learning** without explicit inter-graph edges?

TAG-3D: Overview of Tasks

Interaction Type	Prediction	Task	Data-points	Max # Graphs	# Unique Molec.	# Data-sets	Benchmark
Drug-Drug	Binary Class	Classification	65,583	2	1,676	1	BIOSNAP
Protein-Ligand	ΔG	Regression	19,443	2	34,920	3	LP-PDBbind
Chemical Reaction	Yield	Regression	3,956	6	50	2	Buchwald-Hartwig
Miscible Solvents	ρ	Regression	30,142	5	81	1	CheMixHub
	ΔH_{mix}	Regression	30,142	5	81		
	ΔH_{vap}	Regression	30,142	5	81		
Ionic Liquids	$\ln(\kappa)$	Regression	40,904	3	479	1	CheMixHub
	$\ln(\eta)$	Regression	75,992	3	699		
Drug Solubility	$\ln(S)$	Regression	27,166	3	169	1	CheMixHub
Polymer Electrolyte	$\ln(\kappa)$	Regression	11,350	5	402	1	CheMixHub
Olfactory Mixture	Similarity	Regression	865	43	201	1	CheMixHub
Protein-Protein	Structure	Multi-Regression	135,000	2	270,000	1	PINDER

Consolidated summary of datasets including interaction scale, molecule diversity, and benchmarks

TAG-3D: Some Results

Performance comparison across ChemixHub solvent and mixture tasks

Baselines	ρ (MAE ↓)	ΔH_{mix} (MAE ↓)	ΔH_{vap} (MAE ↓)	$\ln(S)$ (MAE ↓)	Mixture sim. (MAE ↓)
GNN-Attention	0.018 ± 0.020	0.158 ± 0.002	0.098 ± 0.006	0.087 ± 0.006	0.129 ± 0.005
GNN-Deepsets	0.003 ± 0.000	0.159 ± 0.002	0.406 ± 0.668	0.065 ± 0.005	0.146 ± 0.010
MolT5-XGB	0.009 ± 0.000	0.269 ± 0.004	0.306 ± 0.003	0.028 ± 0.001	0.128 ± 0.006
MolT5-Attention	0.005 ± 0.001	0.157 ± 0.002	0.125 ± 0.077	0.082 ± 0.027	0.123 ± 0.005
MolT5-Deepsets	0.008 ± 0.005	0.157 ± 0.003	0.071 ± 0.002	0.130 ± 0.013	0.121 ± 0.006
RDKit-XGB	0.009 ± 0.000	0.225 ± 0.005	0.295 ± 0.002	0.028 ± 0.001	0.125 ± 0.006
RDKit-Attention	0.006 ± 0.001	0.167 ± 0.002	0.199 ± 0.030	0.070 ± 0.006	0.148 ± 0.010
RDKit-Deepsets	0.005 ± 0.000	0.207 ± 0.008	0.079 ± 0.005	0.179 ± 0.011	0.150 ± 0.008
TAG-3D	0.003 ± 0.000	0.159 ± 0.002	0.069 ± 0.002	0.049 ± 0.001	0.119 ± 0.008

Binding affinity prediction on LP-PDBBind

Baselines	Test RMSE (↓)
AutoDock Vina ¹	2.56 ± 0.00
InteractionGraphNet ²	2.16 ± 0.13
RF-Score ³	2.10 ± 0.003
DeepDTA ⁴	2.29 ± 0.04
TAG-3D	1.60 ± 0.06

¹Trott and Olson, *Journal of Computational Chemistry*, 2010.

²Jiang et al., *Journal of Medicinal Chemistry*, 2021.

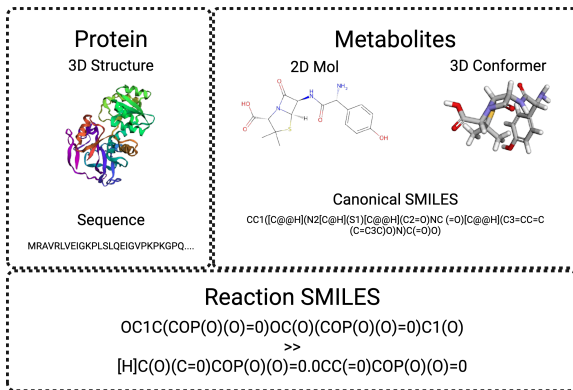
³Ballester and Mitchell, *Bioinformatics*, 2010.

⁴Öztürk et al., *Bioinformatics*, 2018.

Leak-Curber Benchmarks:

Data Leakage and Multimodal Data for Enzymatic Reactions

- Existing open source biological databases isolate data to a single aspect of the reaction and to a single modality
- Models miss out on information from other modalities



The Multiple Modalities of an Enzymatic Reaction

Leak-Curber Comparison with Existing Datasets

Dataset	# Molecules or Reactions	# Enzymes	# Pairs	Substrate Info	Product Info	Reaction Info	Atom Mapping	2D Structure	3D Structure / Conformer
ESP ¹	1379	12156	18351	✓	×	×	×	✓	×
EnzymeMap ²	16776	12749	46356	✓	✓	✓	✓	×	×
Reactzyme ³	7726	178327	178463	✓	✓	✓	✓	×	×
CARE ⁴	48392	174744	—	×	×	×	✓	×	×
ECREACT ⁵	62222	6289	62222	×	×	×	×	×	×
Leak-Curber	41108	208718	625316	✓	✓	✓	✓	✓	✓

¹Kroll, Alexander, et al., *Nature Communications* (2023)

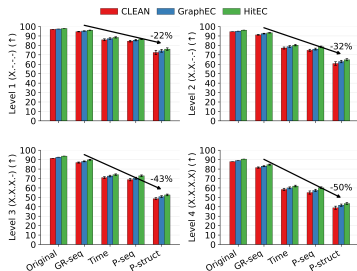
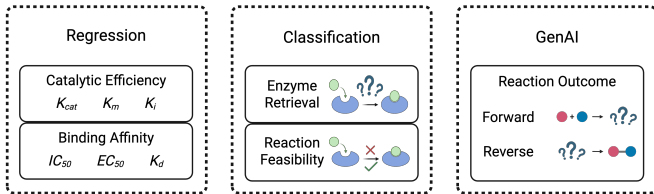
²Heid, Esther et al., *Chemical Science* (2023)

³Hua, Chenqing, et al., *arXiv* (2024)

⁴CARE dataset, *NeurIPS* (2024)

⁵ECREACT dataset, *Nature* (2022)

Leak-Curber Benchmark Tasks



CLEAN: Science (2023); GraphEC: Nature Communication (2024); HitEC: Nature Communication (2026)

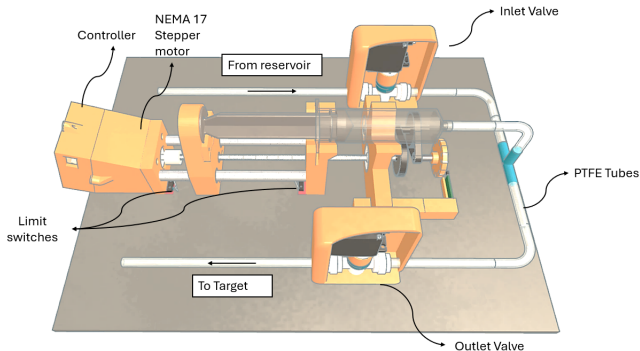
Flow Chemistry-based Synthesis

Nanoparticle synthesis in Corning Advanced-Flow™ Reactor



Flow Chemistry-based Synthesis

Syringe pump for Flow Robots

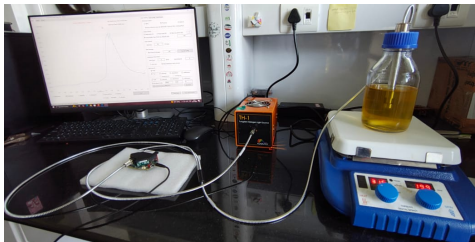
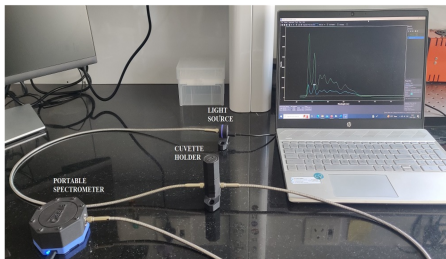


Key contribution

Non-contact feedback control improves flowrate precision, avoiding expensive commercial flow sensors and switching valves.

Affordable Control and Optimization of Reaction Syntheses

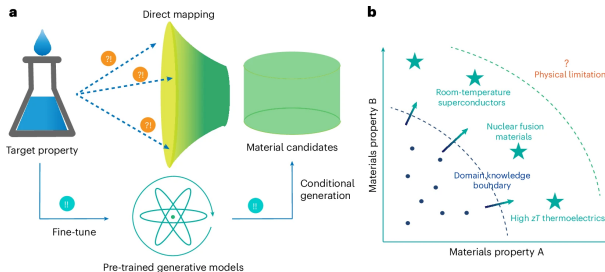
Atline/Online Spectroscopy



AI/ML for Materials Design

Material Classes for AI Enabled Design

Material Class	Design Targets
Batteries	Capacity, ionic transport, interfacial stability, lifetime, and safety
Semiconductors	Bandgap, carrier mobility, defects, dielectric response, and thermal transport
Catalysts	Activity, selectivity, lifetime, and reduced use of critical elements
Defence	Lightweight alloys, coatings, fatigue resistance, and high-temperature performance
Biomaterials	Function, biocompatibility, degradation, and manufacturability

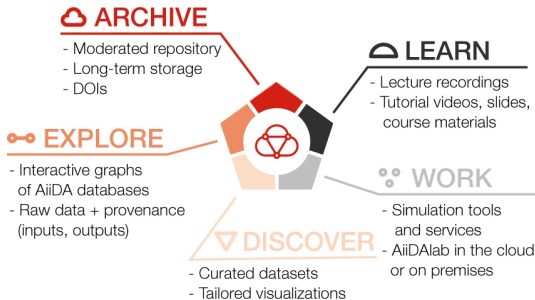


The paradigm of inverse design pushes the material space beyond our domain knowledge boundary¹, Do we have a good quality data?

¹M. Cheng et al., *Nature Materials* (2026).

Data Quality Limits What AI Models Can Design

- Global platforms provide crystal structures, computed properties, and reproducible workflows (NOMAD¹)
- India has focused resources such as Project aNANT for computationally designed two-dimensional materials³
- Measurement conditions and failed outcomes are often missing



Materials Cloud, a platform for open computational science²

¹Sbailò et al., *npj Computational Materials* (2022).

²Talirz et al., *Scientific Data* (2020).

³Project aNANT, IISc (<https://anant.mrc.iisc.ac.in/about>).

Efficient GenAI Needs Better Scaling and Evaluation

Model Strategy

- Use a flexible generator for composition, unit cells, and atomic coordinates when the task spans broad inorganic space^a
- Fine-tune compact MLIPs when the chemistry and target property are narrow^b
- Increase model size only after data and compute scaling curves justify it

^aHöllmer et al., *Proceedings of the 42nd International Conference on Machine Learning* (2025).

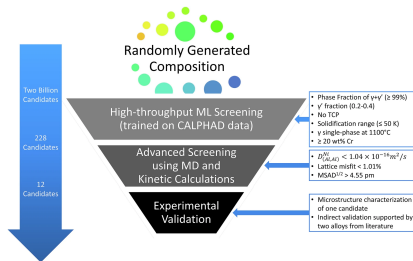
^bRadova et al., *npj Computational Materials* (2025).

Evaluation Strategy

- Measure validity, uniqueness, stability, target error, and novelty together
- Validate candidates with DFT or another task-relevant physics model. Reconstruction error alone is not a discovery metric^a
- Need for common benchmarking protocols and continually updated datasets

^aS. Yang et al., *The Twelfth International Conference on Learning Representations* (2024).

Connect Atoms to Components



Multiscale computational framework¹

- **Atomic scale:** MLIP foundation models aim to extend first-principles accuracy to larger systems and trajectories²
- **Mesoscale:** Learned Neural operators can accelerate phase-field simulations of microstructure evolution³

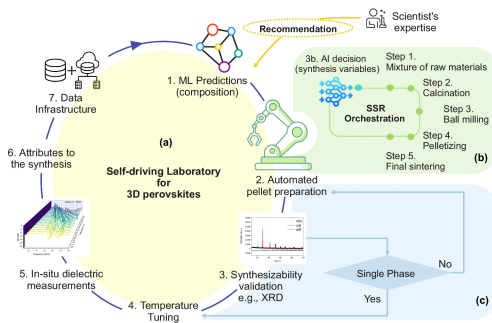
¹Wang, Kwon, Oh, et al., *npj Computational Materials* (2025).

²Yuan et al., *Nature Reviews Chemistry* (2026).

³Bonneville, Bieberdorf, Hegde, et al., *npj Computational Materials* (2025).

Every Experiment Should Improve the Next Design

- Select a candidate using expected improvement and uncertainty
- Run synthesis and characterization
- Conditional Approaches can treat experimental failure as useful information instead of discarding it¹



Automated synthesis and characterization²

¹Wakabayashi et al., *npj Computational Materials* (2022).

²Omidvar et al., *Nature Communications* (2024).

Conclusions

AI as a Scientific Collaborator

AI is good at:

- Searching large spaces.
- Ranking hypotheses.
- Combining multiple modalities.
- Updating from new experiments.

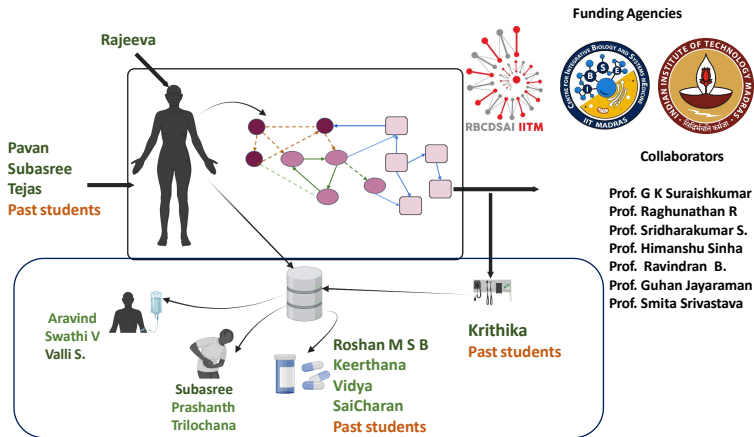
Scientists remain essential for:

- Asking the right question.
- Defining the right assay.
- Interpreting mechanism.
- Judging chemical feasibility.
- Validating relevance.

Human expertise plus AI-guided experimentation for Drug Discovery¹ and Materials Design

¹Bender and Cortes-Ciriano, *Drug Discovery Today* (2021).

Acknowledgments



Thank you