Retrosynformer: Planning Multi-Step Chemical Synthesis Routes via a Decision Transformer

Granqvist, Mercado, Genheden —
Digital Discovery, 2026

Presented by: Ryan Leung, 21/8/2026

Outline

- Background: Retrosynthesis
- Methods: the RetroSynFormer model
- Results
- Takeaway

BACKGROUND:
What is retrosynthesis?

- Work backwards from a target molecule to buyable starting materials via known reactions (Corey, 1960s)
- Output = a **synthesis route**: a tree of reactions ending in purchasable "building blocks"

Why it's hard, and the gap this paper fills

- Huge search space; methods split into template-based vs. template-free, and single-step vs. multi-step (current multi-step gold standard: MCTS, e.g. **AiZynthFinder**)
- Existing search methods (MCTS, A*) mostly condition only on the *current* molecule, largely ignoring route history
- Fix: borrow the **Decision Transformer** (DT) from offline RL — treats a whole route as a sequence and predicts the next step from full history. First-ever application of DT to chemistry.

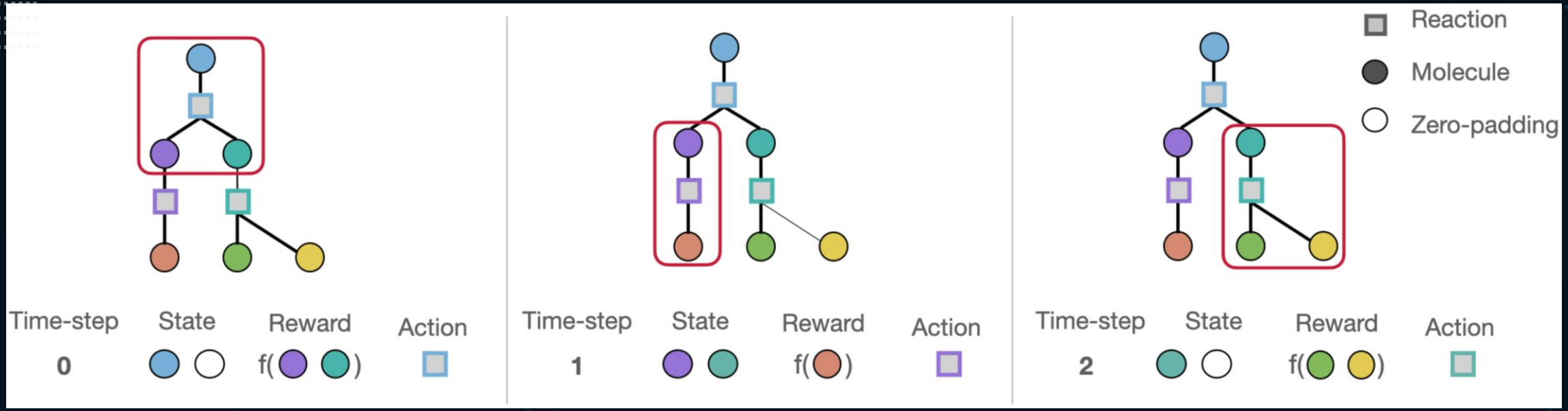
Method

Table 1 Details of the datasets. The first three rows show the number of unique routes in the training, validation, and test sets for the three main datasets curated in this work. All datasets were created from PaRoutes[4] data. The last row includes the mean pairwise Tanimoto similarity between 10 000 random targets from the training set and test set

Dataset	Small	Standard	Large
Training set	44 736	67 180	86 048
Valid set	5362	8222	10 645
N1 test set	2168	4320	5631
N5 test set	1732	3569	5260
# Templates	588	1572	2986
# Building blocks	38 521	58 251	72 737
Total # unique targets	53 626	82 222	106 452
Total # routes	144 812	326 294	442 844
Test–train mean similarity	0.127	0.123	0.126

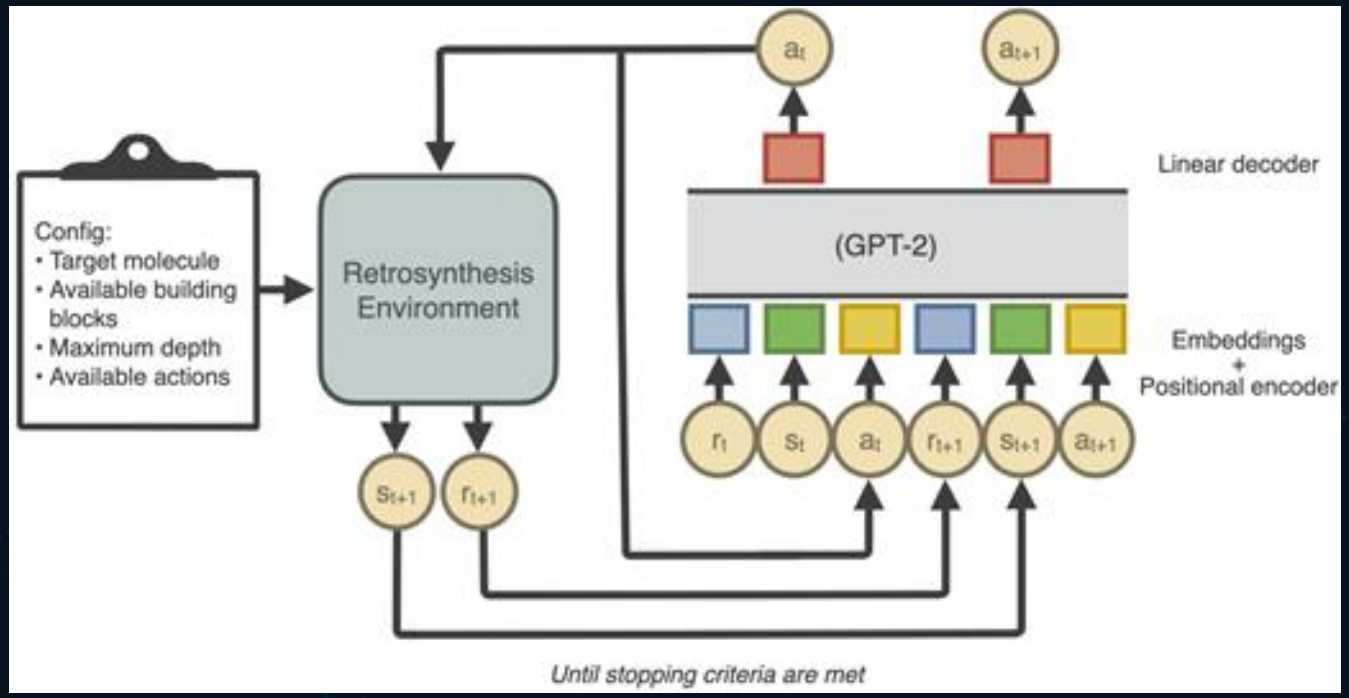
Data: PaRoutes

- 457,166 routes derived from USPTO patent reactions, molecules as SMILES
- Filtered by template frequency into 3 sizes: small (588 templates), **standard (1572, used for all main results)**, large (2986)
- Split by target compound (not route) to avoid train/test leakage; separate "N1" (simple/linear) and "N5" (longer/complex) test sets



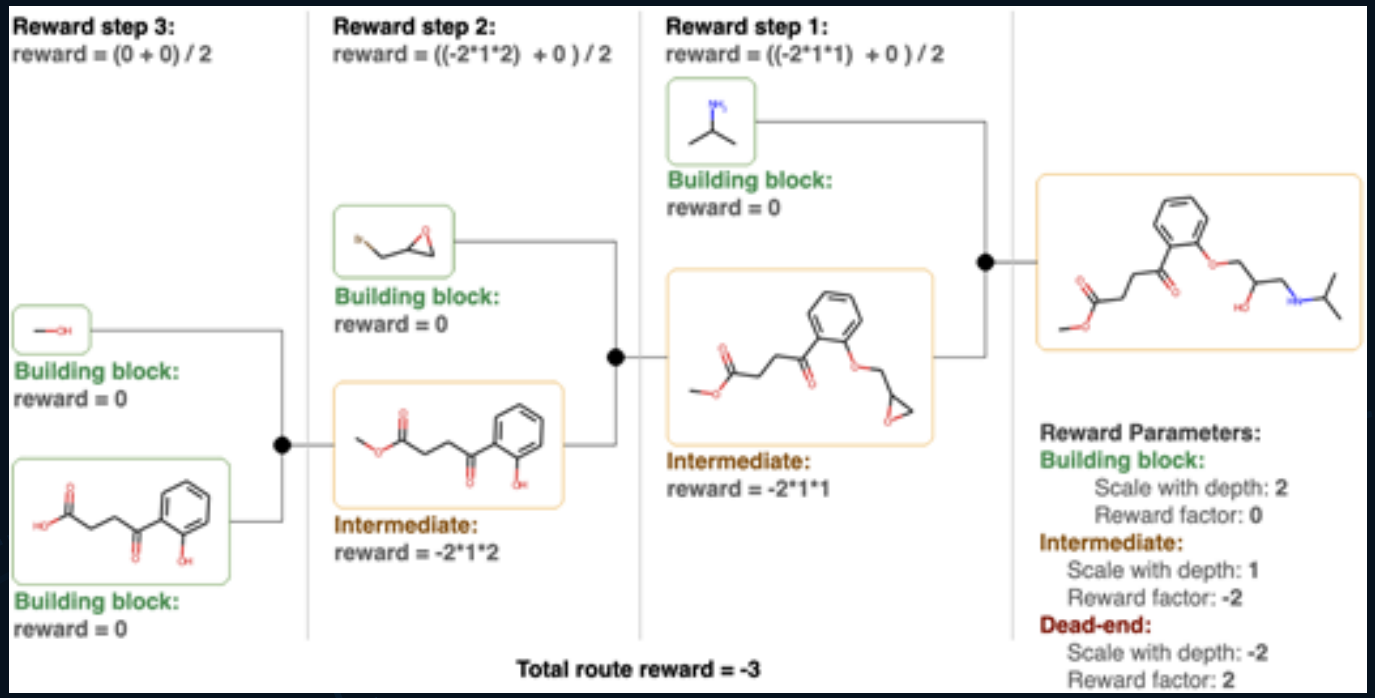
Route → sequence

- Routes are trees (branching when a reaction gives 2+ reactants); flattened via depth-first traversal (left branches expanded first, others stacked)
- Each step = (**state** = fingerprint of up to 2 molecules, **action** = reaction template, **reward** = step-quality score)



Model architecture

- Backbone: **GPT-2**, repurposed as a Decision Transformer
- State = 2048-bit Morgan fingerprint (two 1024-bit molecules, zero-padded if only one); Action = one-hot over 1572 templates
- Model predicts the next template's probability given the full past history of states/actions/rewards



Reward design

- Challenge: can't score a half-built route the way you'd score a finished one
- Reward = f(molecule status: building block / intermediate / dead-end) + depth penalty (discourages long routes)
- Parameters auto-tuned with **Optuna** (a hyperparameter optimization framework)

Inference pipeline

- DT + "retrosynthesis environment": **RDChiral** applies the predicted template to get real candidate reactants; environment checks building-block/intermediate/dead-end/loop status and tracks the route
- **Beam search** (width 50 by default): keeps top-*n* candidate partial routes at each step instead of a single greedy path

Evaluation metrics

- Success rate (found *any* complete route)
- Top-1 accuracy (matches patent route exactly)
- Tree Edit Distance (TED, structural similarity to reference)
- Action/class/route accuracy (comparison w/ patent route)
- DeepSet route score (learned quality metric based on expert opinion)

Baseline: AiZynthFinder

- Template-based MCTS method
- Current field standard (deployed at AstraZeneca)
- Trained on the same templates/data
- For fair comparison: 100 MCTS iterations, max depth 6 — same as RetroSynFormer

Results

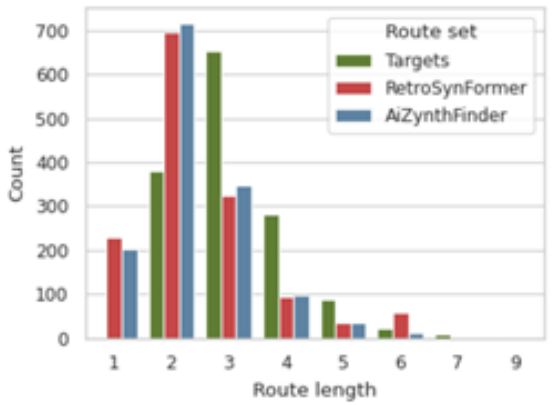
Table 2 Performance of the RetroSynFormer compared to the AiZynthFinder baseline on retrosynthesis planning tasks using a diverse suite of evaluation metrics. Arrows indicate the direction of better performance for each metric

Test set	RetroSynFormer50[a]		AiZynthFinder	
	N1	N5	N1	N5
Success rate (%) ↑	0.924	0.899	0.939	0.924
Top-1 accuracy ↑	0.106	0.058	0.143	0.071
Mean time per route(s) ↓	68.1	81.9	5.45	5.8
Mean TED ↓	5.58	7.08	5.40	7.12
Mean # reactions per route ↓	2.34	2.55	2.52	2.92

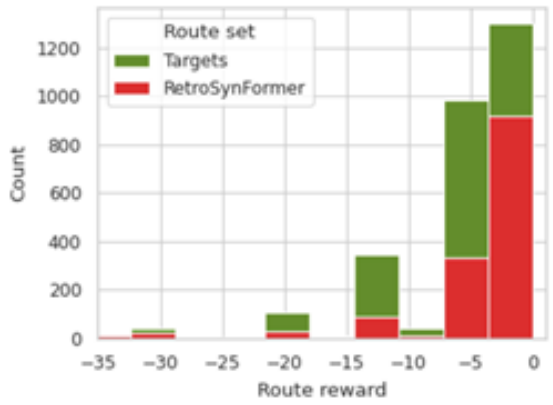
[a] RetroSynFormer results show averages over three runs using beam width 50. Standard deviation for the RetroSynFormer models is ≤ 0.004 for the success rate, ≤ 0.001 for the top-1 accuracy, ≤ 1.7 for the mean time per route, ≤ 0.05 for the mean tree edit distance (TED), and ≤ 0.01 for the mean # reactions per route.

Headline numbers

- RetroSynFormer (beam 50): 92.4% success (N1) / 89.9% (N5)
- AiZynthFinder: 93.9% (N1) / 92.4% (N5) — a few points ahead
- Top-1 accuracy low for both (~5–14%)
- RetroSynFormer much slower (68–82 s/route vs. ~5–6 s/route)



(a)



(b)

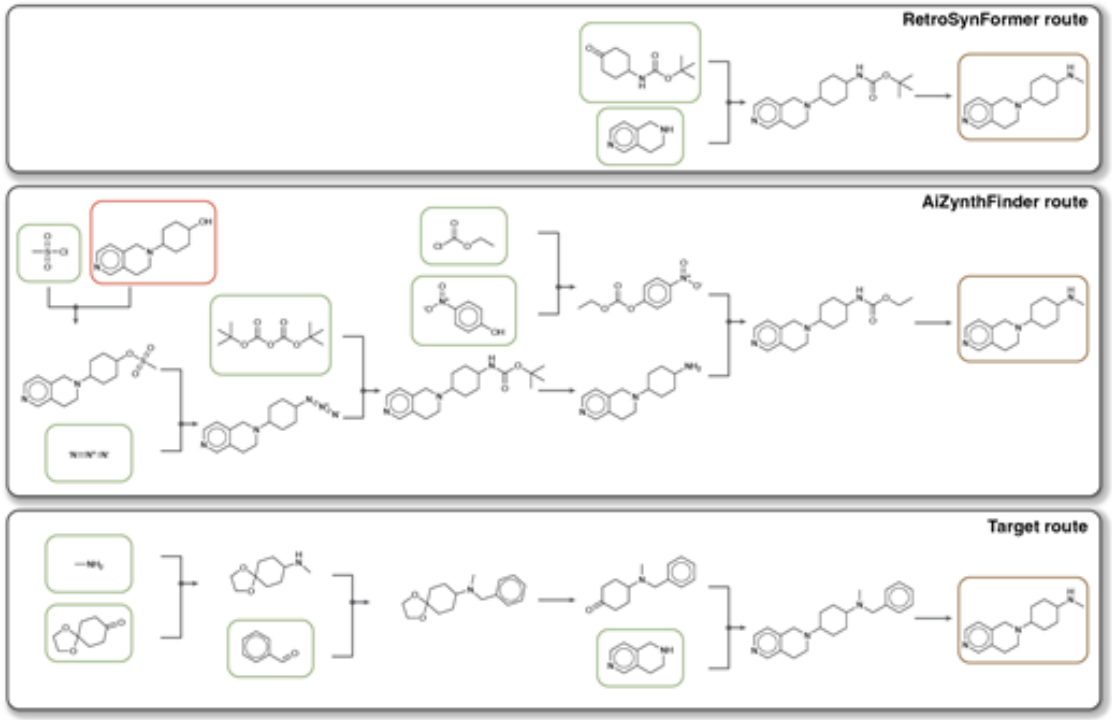
Solved by	N1			N5		
	Count	Percent	TED	Count	Percent	TED
Only RetroSynFormer	59	3.9%		67	4.5%	
Only AiZynthFinder	81	5.4%		112	7.5%	
Both	1328	88.5%	3.81	1274	84.9%	4.81
None	32	2.1%		47	3.1%	

Route length, quality & complementarity

- Predicted routes (both models) shorter on average than patented routes — suggests shorter valid paths exist
- 88.5% of N1 targets solved by both models; only 2.1% (N1) / 3.1% (N5) solved by neither → combined coverage **97–98%**

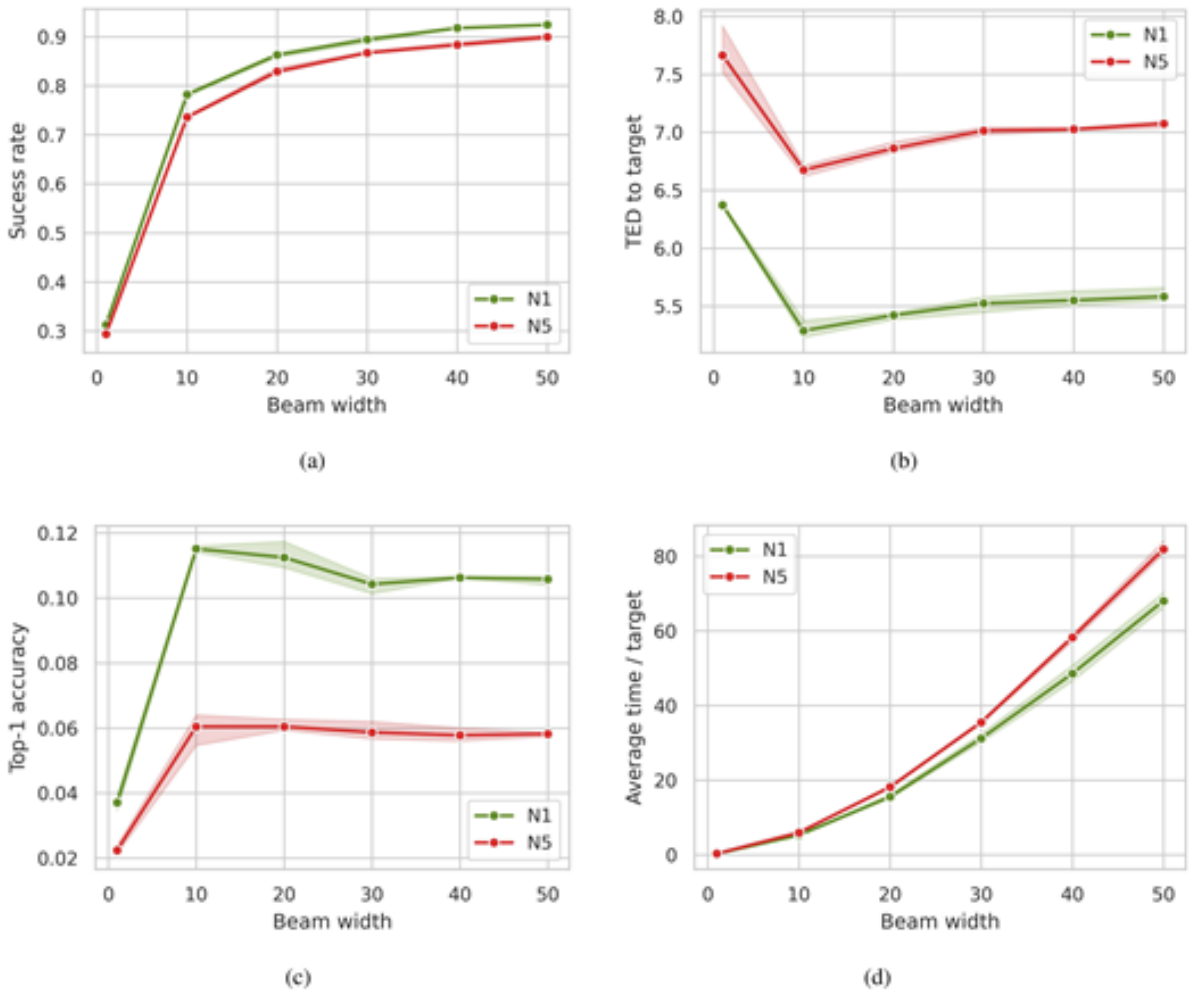
Case study

- Benzamide target (patent US-8680159-B2): RetroSynFormer found an efficient 2-step route (reductive amination + deprotection, with literature precedent)
- AiZynthFinder never applied the reductive-amination template, took a long inefficient path, hit max depth without reaching a purchasable building block
- Illustrates complementary strengths, not one model dominating



Beam width & robustness

- Success rate rises steeply then plateaus (~logarithmically) with beam width; search time grows ~exponentially
- Top-1 accuracy and TED actually get *slightly worse* past beam width ≈10 — more search ≠ better patent-match
- Performance fairly robust to reward-function tweaks and to template-set size (588/1572/2986); top-1 accuracy drops as template space grows



Takeaway

Limitations

- Much slower than AiZynthFinder (though AiZynthFinder has 5+ years of engineering optimization — not a fully fair speed comparison)
- Trained on only one sampled route per target; evaluated only on PaRoutes (patent data) — generalization to other chemical spaces (e.g. ChEMBL, internal discovery compounds) untested
- DT's key selling point — using patterns from the *whole* route history — not directly demonstrated; flagged as future work

Takeaways

- First DT applied to chemistry/retrosynthesis; performance competitive with, not superior to, AiZynthFinder — but strongly complementary (97–98% combined)
- Reframes retrosynthesis as sequence prediction rather than tree search — valuable architectural exploration even without beating SOTA

Key References

- Corey & Wipke 1969;
- Chen et al. 2021 (DT);
- Genheden & Bjerrum 2022 (PaRoutes);
- Genheden et al. 2020 (AiZynthFinder);
- Coley et al. 2019 (RDChiral)

Thank you / Q&A