Preference Optimization for Molecule Synthesis with Conditional Residual Energy-based Models

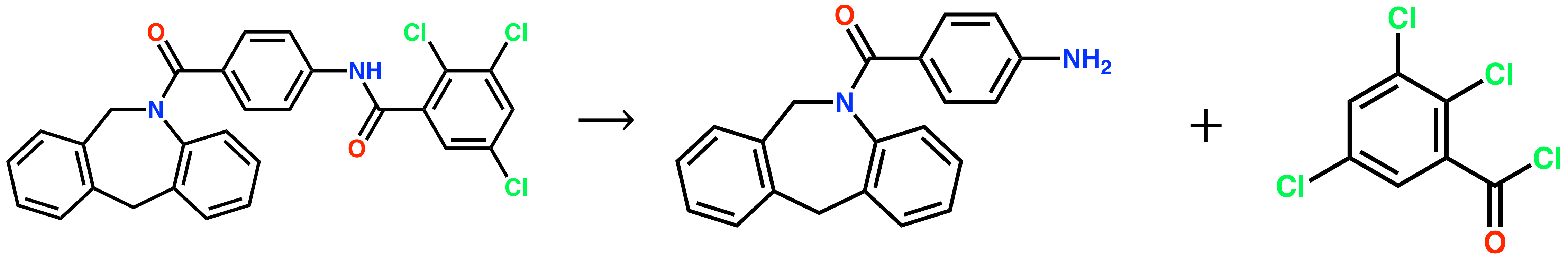
ICML 2024 (Oral)

Songtao Liu[1], Hanjun Dai[2], Yue Zhao[3], Peng Liu[1]
Contact: Songtao Liu (skl5761@psu.edu)
[1]Penn State, [2]Google DeepMind, [3]University of Southern California

1. Molecule Synthesis (Retrosynthetic Planning)

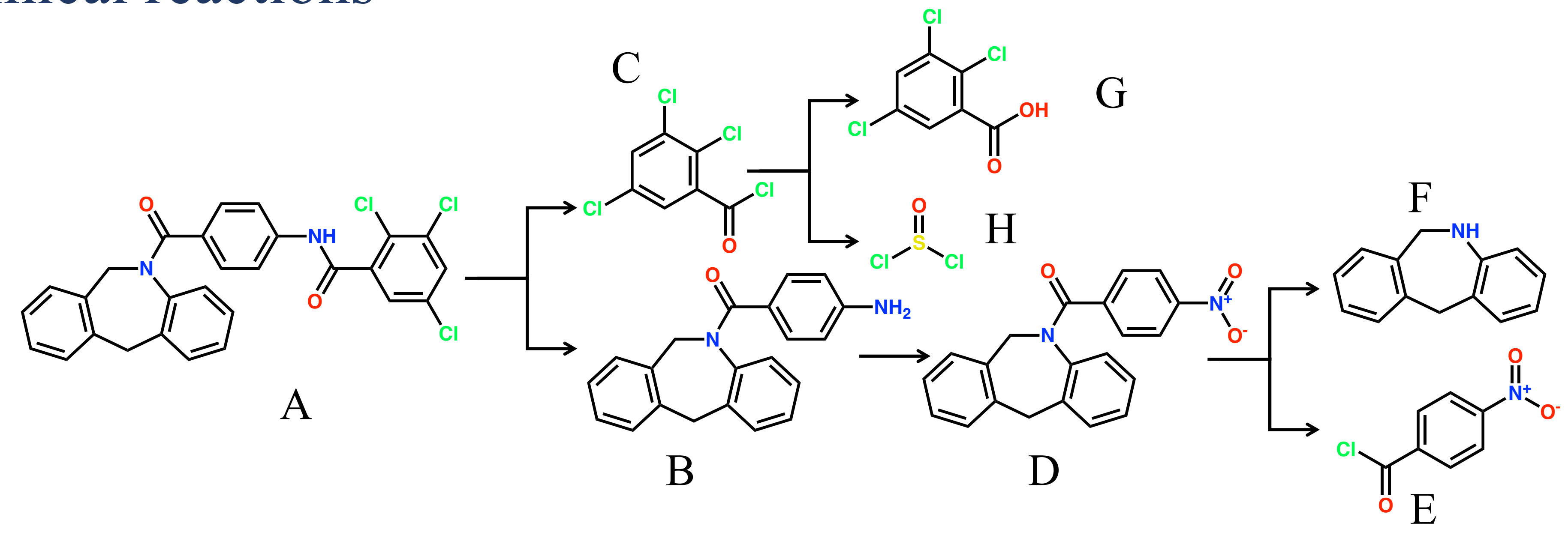
Background: Retrosynthesis Prediction

Given a target product molecule the goal of one-step retrosynthesis is to predict a set of reactants that can react to synthesize this product.

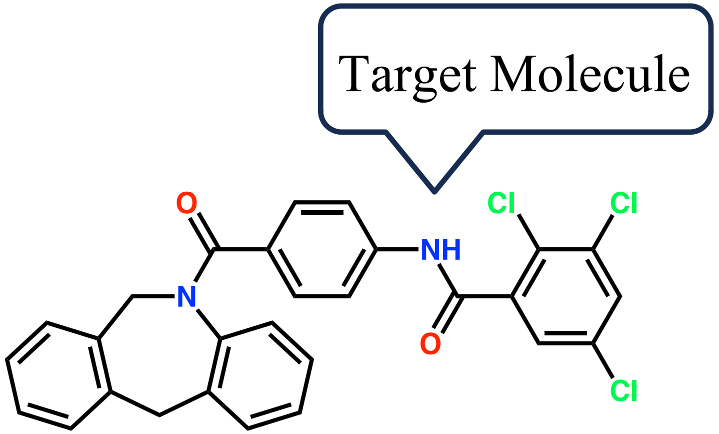


Background: Retrosynthetic Planning

Given a target molecule, the goal of retrosynthetic planning is to search for the starting materials that can synthesize the target molecule through a set of chemical reactions

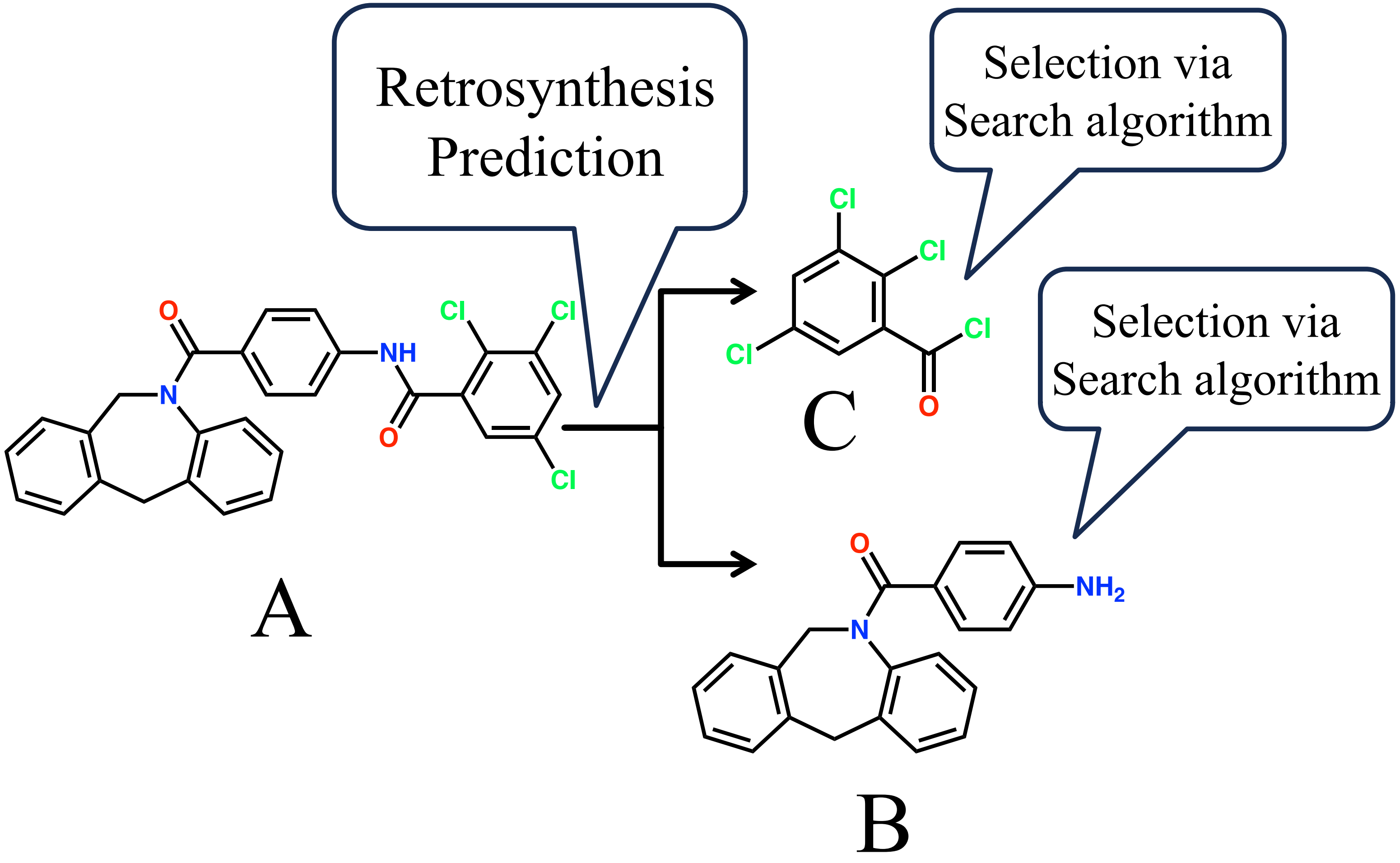


Background: Synthetic Route Generation

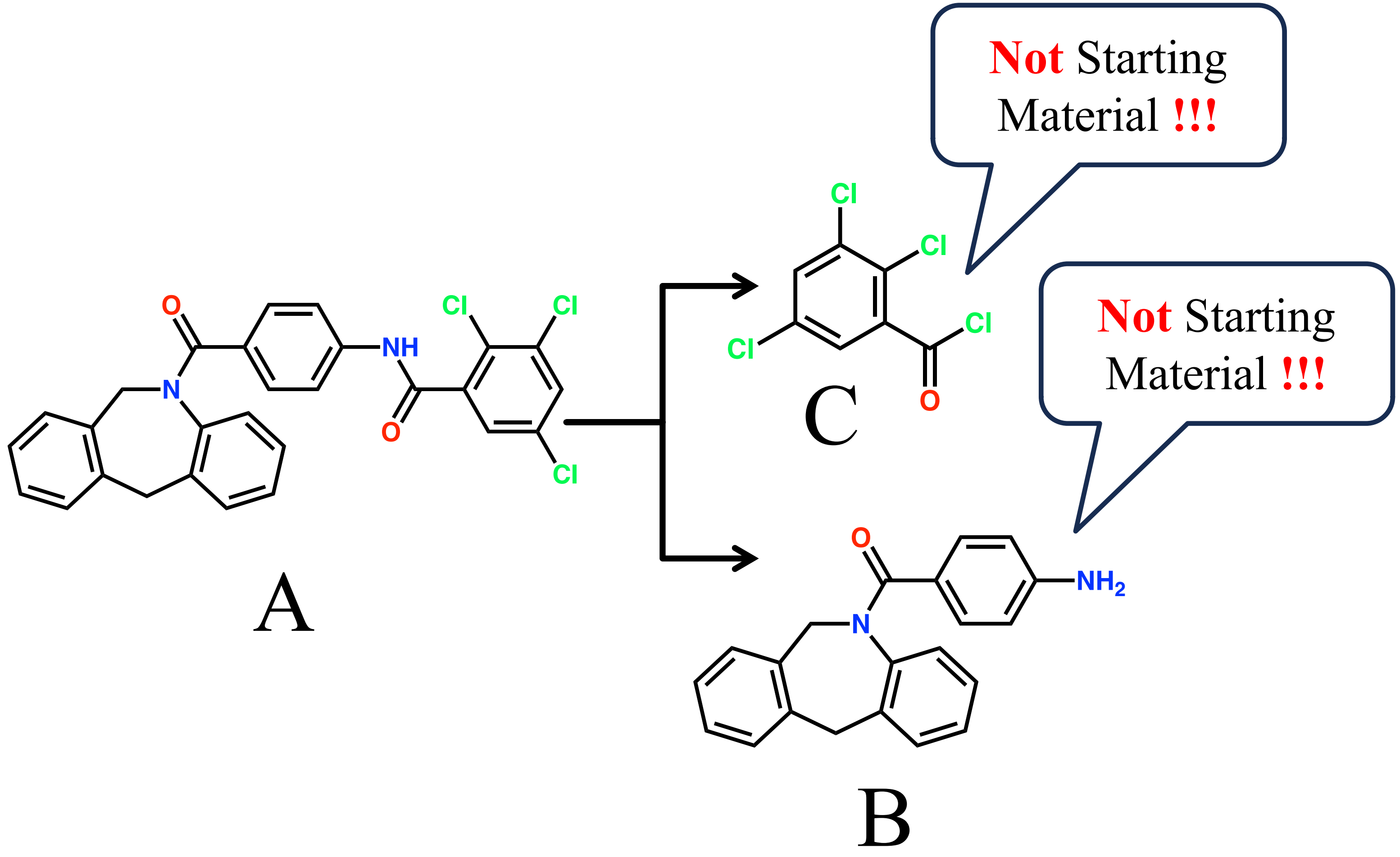


A

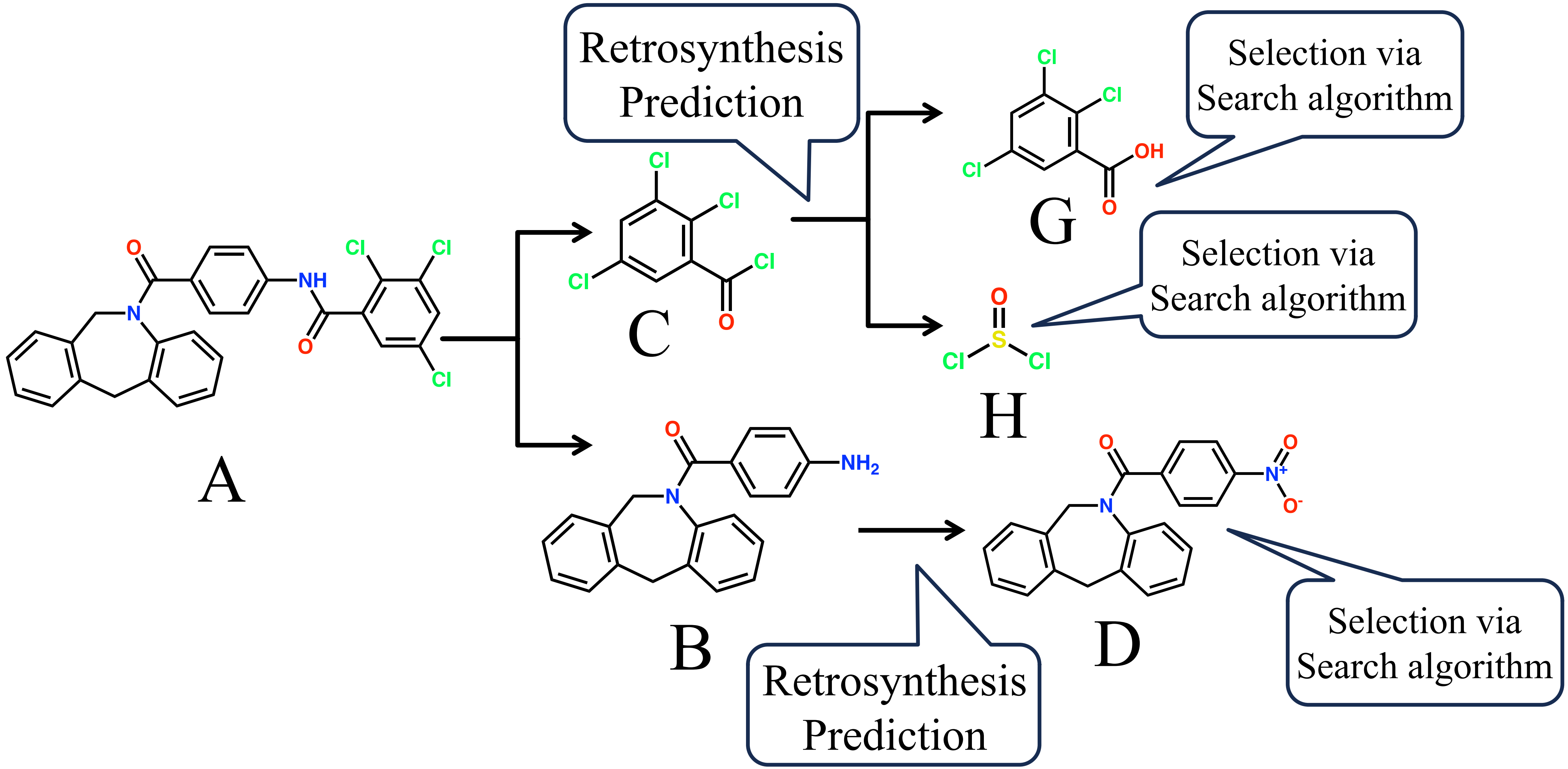
Background: Synthetic Route Generation



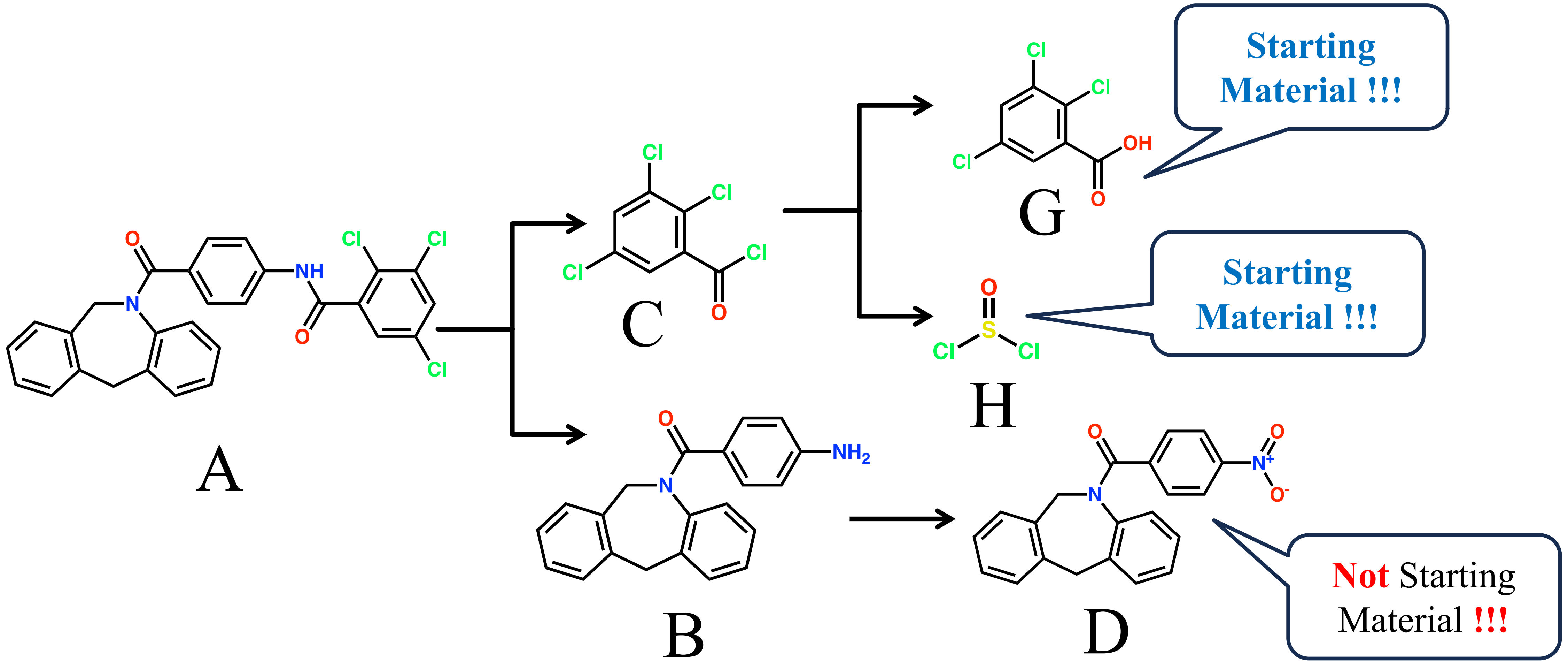
Background: Synthetic Route Generation



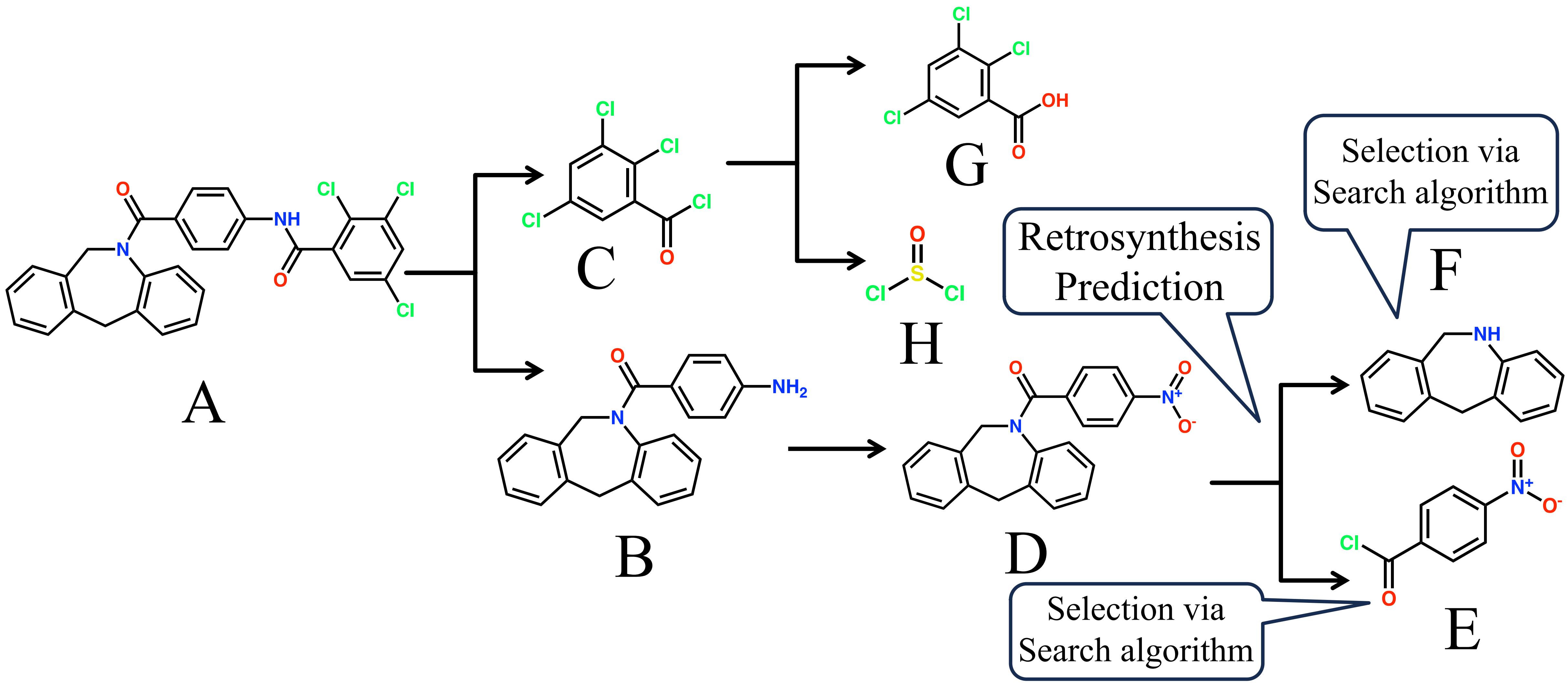
Background: Synthetic Route Generation



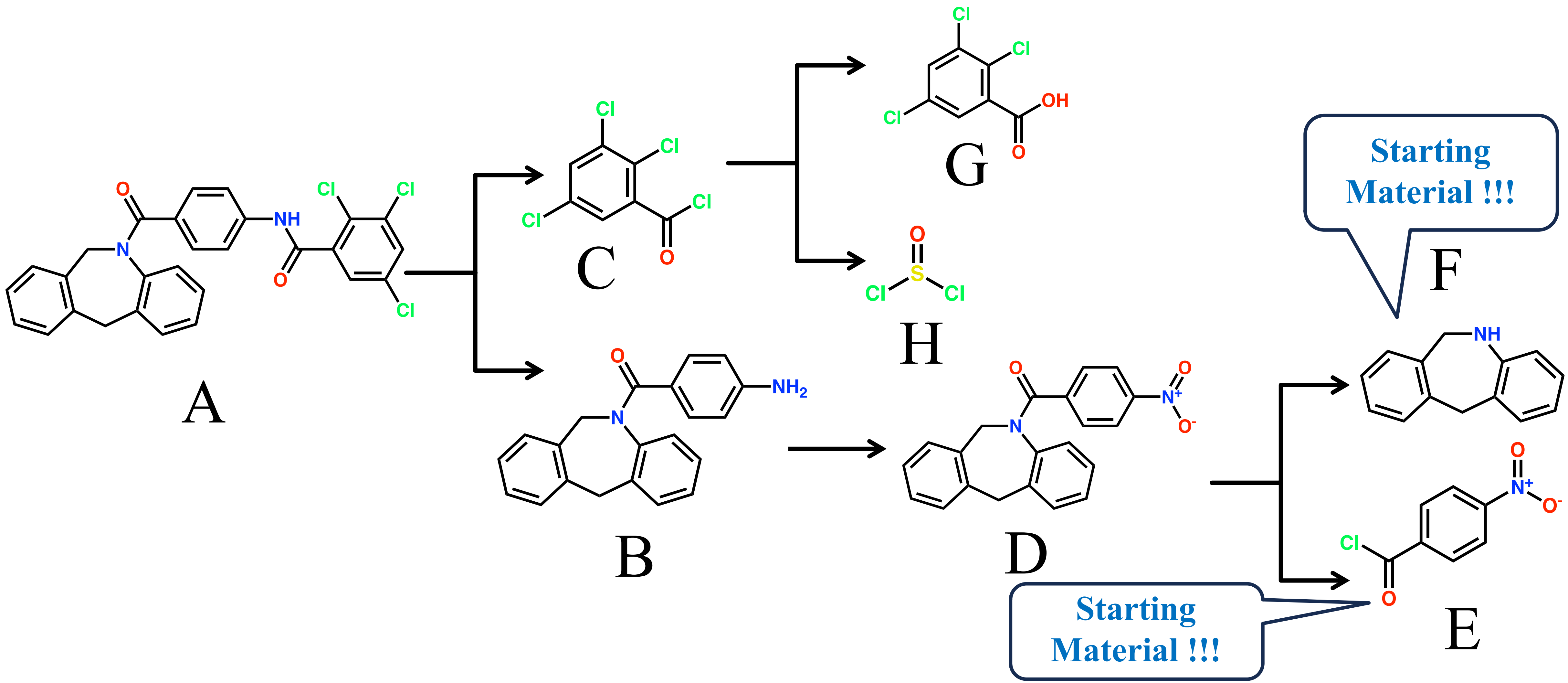
Background: Synthetic Route Generation



Background: Synthetic Route Generation

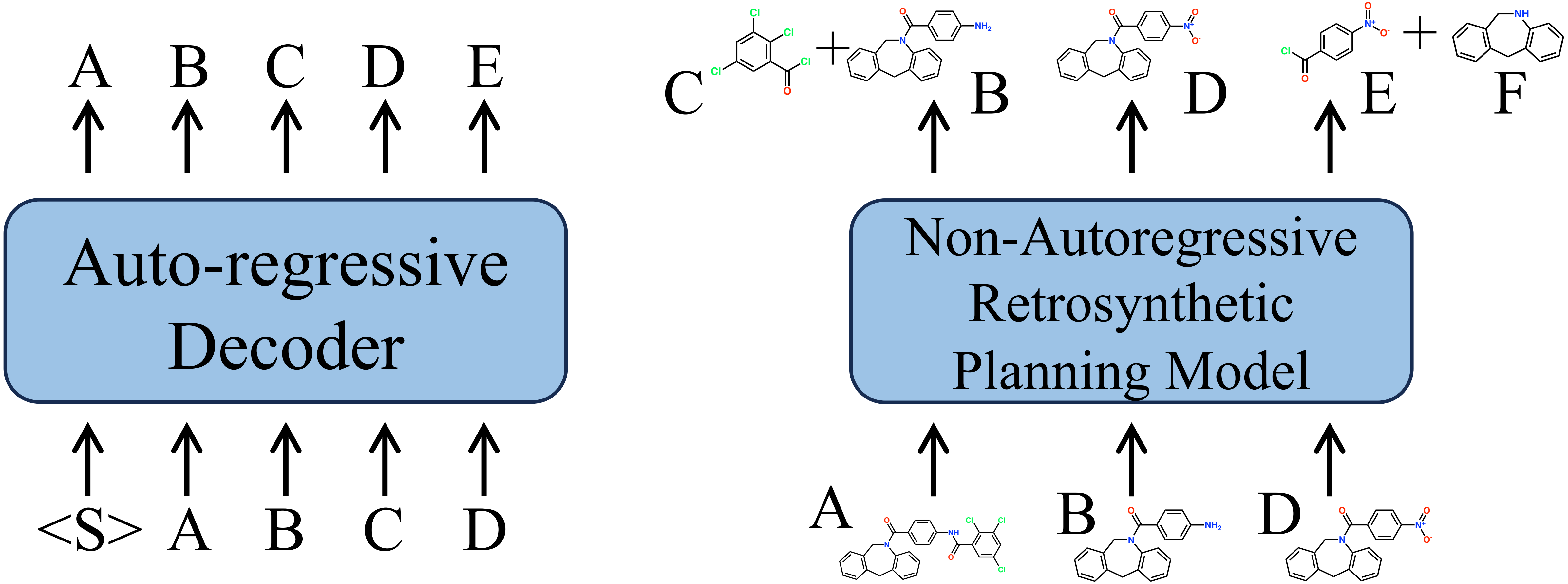


Background: Synthetic Route Generation



2. Probabilistic View of Retrosynthetic Planning

Markov Chain in Retrosynthetic Planning

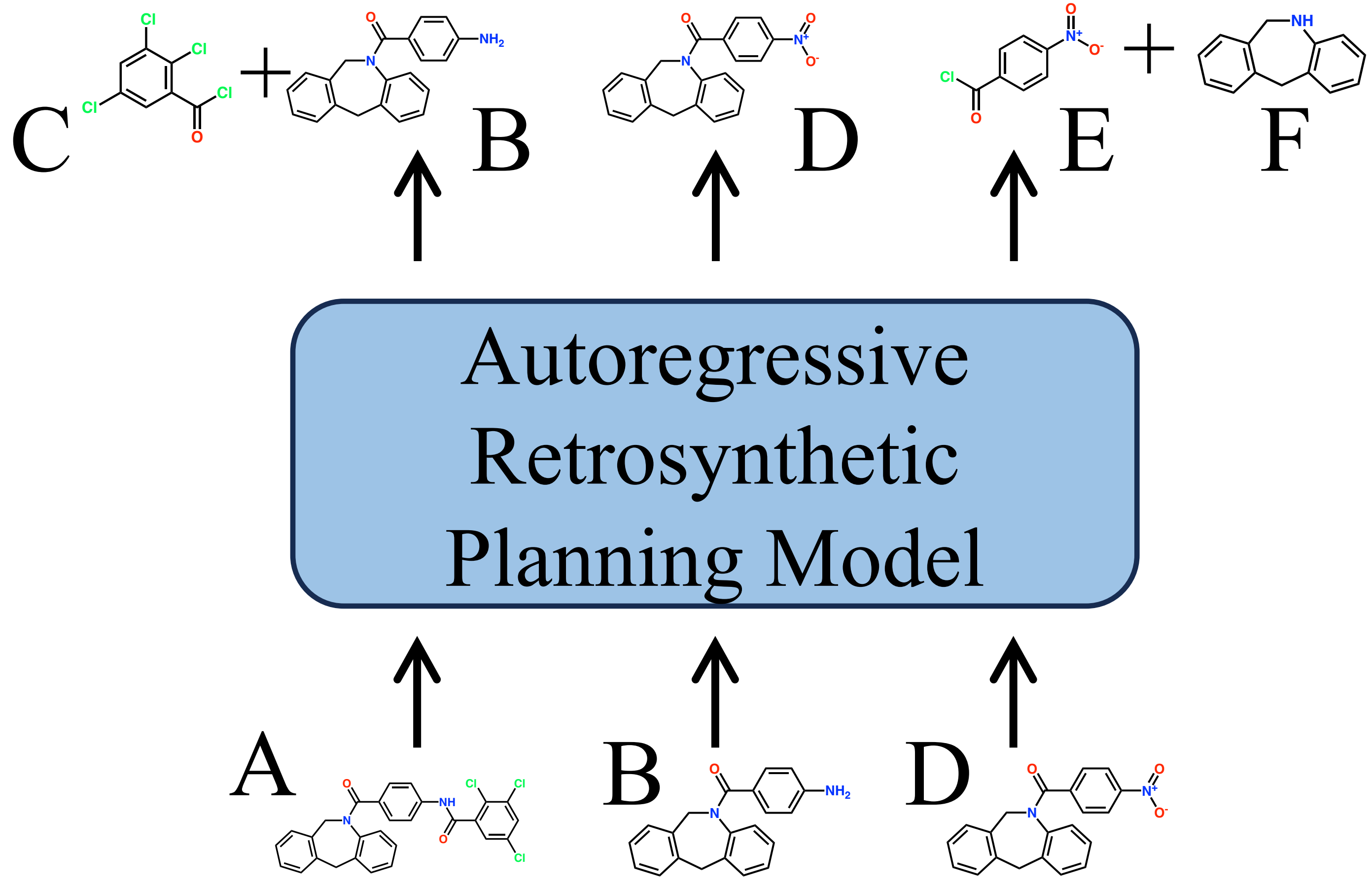


$$p(A|<S>)p(B|<S>,A)p(C|<S>,A,B)$$
$$p(D|<S>,A,B,C)\ p(E|<S>,A,B,C,D)$$

$$p(B|A)p(D|B)p(E|D)$$

FusionRetro: An Autoregressive Model for Retrosynthetic Planning

[1] exploits previous molecules as context to generate the next reactant set, which can improve the performance of singe-step retrosynthesis prediction.



$p(\mathrm{B}|\mathrm{A})p(\mathrm{D}|\mathrm{A,B})p(\mathrm{E}|\mathrm{A,B,D})$

[1] FusionRetro: Molecule Representation Fusion via In-Context Learning for Retrosynthetic Planning, 2023 ICML

One-Step Retrosynthesis Models can't Generate Routes with Desired Quality

Each step in the retrosynthesis planning is locally normalized.

i.e. $\sum_{\mathcal{R}} P(\mathcal{R}|m_p,\cdot) = 1$, m_p is the product and $\mathcal{R}$ is the reactant set.

I

II

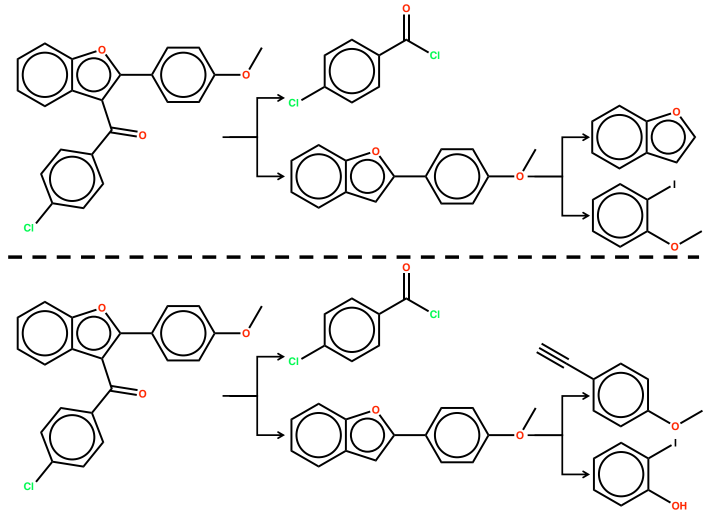


Figure 2. For a given target molecule, we find two synthetic routes that can synthesize it in the dataset.

Any ranking of plans for the synthesis of a given target compound depends on benchmarks which must be defined. Possible criteria may be

- the shortest route (time involved),
- the cheapest route (cost of materials),
- the novelty of the route (patentability),
- the greenest route (avoidance of problematic waste),
- the healthiest route (avoidance of toxic intermediates and side products),
- the most reliable route (lowest risk approach).

This step-by-step generation process often fails to account for criteria primarily because it relies on pure probability (local normalization) for predicting routes without forward-thinking.

3. Conditional Residual Energy-based Models for Controllable Synthetic Route Generation

Energy Function for Evaluating the Quality of Synthetic Routes

$$p_\theta(\mathcal{T}) \propto \exp\left(\log P_{Retro}(\mathcal{T}) - E_\theta(\mathcal{T})\right)$$

By introducing $E_\theta(\mathcal{T})$, we can consider various criteria and incorporate additional properties encoded within $E_\theta(\mathcal{T})$, which is defined on the synthetic route, allowing for route evaluation based on various criteria.

Energy-based Models (discrete space)

Energy-based models define the distribution via an energy function. For $x \in \mathbb{R}^D$, its probability density can be expressed as:

$$P_\theta(x) = \frac{\exp\left(-E_\theta(x)\right)}{Z(\theta)},$$

where $Z(\Theta) = \int \exp\left(-E_\Theta(x)\right)$ is the normalization constant.

Typically, we use NCE loss function to approximate the target distribution.

$$\mathcal{L} = -\mathbb{E}_{x \sim P_{\text{data}}}\left[\log \sigma\left(-E_\Theta(x)\right)\right] - \mathbb{E}_{\tilde{x} \sim P_{\text{noise}}}\left[\log(1 - \sigma(-E_\Theta(\tilde{x})))\right]$$

Residual Energy-based Models

Residual Energy-based Models[3]:

$$P_\theta(x) \propto P_{LM}(x)\exp(-E_\theta(x))$$

$P_{LM}(x)$ is a locally normalized model. By incorporating the energy function, $P_\theta(x)$ can approximate the target distribution better.

$$\mathcal{L} = \mathbb{E}_{x\sim P_{data}}\left[\log\frac{1}{1+\exp\left(E_\theta(x)\right)}\right] + \mathbb{E}_{\tilde{x}\sim P_{LM}}\left[\log\frac{1}{1+\exp\left(-E_\theta(\tilde{x})\right)}\right]$$

[3] Residual Energy-based Models for Text Generation, 2020 ICLR

Conditional Residual Energy-based Models

$$
\begin{aligned}
& P_\theta\left(\mathcal{T} \mid m_{tar}, c\right) \\
= & P_{Retro}\left(\mathcal{T} \mid m_{tar}\right) \frac{\exp \left(-E_\theta\left(\mathcal{T} \mid m_{tar}, c\right)\right)}{Z_\theta\left(m_{tar}, c\right)} \\
\propto & P_{Retro}\left(\mathcal{T} \mid m_{tar}\right) \exp \left(-E_\theta\left(\mathcal{T} \mid m_{tar}, c\right)\right),
\end{aligned}
$$

$P_{Retro}(\mathcal{T}|m_{tar})$ is a base model, c denotes specific criteria (condition), E_Θ is the conditional residual energy function for evaluating the quality given c. During the training of our energy function E_Θ, $P_{Retro}(\mathcal{T}|m_{tar})$ is fixed. Therefore, our CREBM is a post-training method and can be applied on top of any $P_{Retro}(\mathcal{T}|m_{tar})$.

How to Train Our CREBM

$$\mathcal{L} = -\mathbb{E}\left[\log \sigma\left(-E_\theta\left(\mathcal{T}_w \mid m_{tar}, c\right) + E_\theta\left(\mathcal{T}_l \mid m_{tar}, c\right)\right)\right]$$
$$\text{s.t. } \left(m_{tar}, \mathcal{T}_w, \mathcal{T}_l\right) \sim \mathcal{D},$$

By using this loss function, we can prefer the synthetic routes with higher desired quality (lower energy). The sigmoid function can output the value between 0 and 1, which is indeed probability.

How to Train Our CREBM

$$
\begin{aligned}
\mathcal{L} = -\mathbb{E}\left[\log \sigma\left(-E_\theta\left(\mathcal{T}_w \mid m_{tar}, c\right) + E_\theta\left(\mathcal{T}_l \mid m_{tar}, c\right)\right)\right] \\
\text{s.t. } \left(m_{tar}, \mathcal{T}_w, \mathcal{T}_l\right) \sim \mathcal{D},
\end{aligned}
$$

We use one base model for sampling synthetic routes to construct our synthetic preference dataset $\mathcal{D}$. The preference is constructed by the following reward function for evaluating the quality of synthetic routes:

$$\varphi\left(\cdot \mid m_{tar}, c\right) : \mathcal{X}_{tar} \rightarrow \mathcal{V},$$

In this work, we focus on the feasibility of generated synthetic routes. We implement reward function as

$$\varphi\left(\mathcal{T} \mid m_{tar}, c\right) = \mathrm{sim}\left(f(\mathcal{B}), m_{tar}\right) + \mathrm{sim}\left(\mathcal{B}, \mathcal{B}_{ref}\right),$$

CREBM Framework

Algorithm 1 CREBM Framework

1: **[Train Phase]: Learning**:
2: Define reward function $\varphi\left(\cdot \mid m_{tar}, c\right)$ as Eq. (16)
3: $\varphi\left(\cdot \mid m_{tar}, c\right) : \mathcal{X}_{tar} \rightarrow \mathcal{V}$
4: Rank synthetic routes $\mathcal{X}_{tar} \sim P_{Retro}$ based on φ
5: **Train Conditional Residual Energy-based Models**:
6: $\theta^* = \arg\min_\theta \mathcal{L} = \arg\max_\theta \mathbb{E}_{(m_{tar}, \mathcal{T}_w, \mathcal{T}_l) \sim \mathcal{D}}$
7: **[Test Phase]: Inference**:
8: **Input**: θ^*, m_{tar}, Proposal $\mathcal{X}_{tar} \sim P_{Retro}\left(\cdot \mid m_{tar}\right)$.
9: $L \leftarrow -\log P_{Retro}\left(\mathcal{T} \mid m_{tar}\right) + E_\theta\left(\mathcal{T} \mid m_{tar}, c\right)$
10: $\mathcal{T}^* = \arg\min_{\mathcal{T} \in \mathcal{X}_{tar}} L$
11: **Return** $\mathcal{T}^*$

4. Experiments

Results

In this work, we use the set-wise match between the predicted starting material set and the reference to evaluate the feasibility.

Table 1. Summary of retrosynthetic planning results in terms of exact match accuracy (%). Our framework (CREBM) can consistently improve the performance of existing strategies.

Search Algorithm / One-step Model	Retro*					Retro*-0					Greedy DFS
	Top-1	Top-2	Top-3	Top-4	Top-5	Top-1	Top-2	Top-3	Top-4	Top-5	Top-1
Template-based											
Retrosim (Coley et al., 2017)	35.1	40.5	42.9	44.0	44.6	35.0	40.5	43.0	44.1	44.6	31.5
Neuralsym (Segler & Waller, 2017)	41.7	49.2	52.1	53.6	54.4	42.0	49.3	52.0	53.6	54.3	**39.2**
Neuralsym+CREBM	**44.2**	50.8	53.6	54.6	55.4	**44.5**	**51.0**	53.5	54.5	55.2	-
GLN (Dai et al., 2019)	39.6	48.9	52.7	54.6	55.7	39.5	48.7	52.6	54.5	55.6	38.0
GLN+CREBM	43.3	**51.1**	**53.9**	**55.5**	**56.4**	43.2	**51.0**	**53.8**	**55.5**	**56.3**	-
Semi-template-based											
G2Gs (Shi et al., 2020)	5.4	8.3	9.9	10.9	11.7	4.2	6.5	7.6	8.3	8.9	3.8
GraphRetro (Somnath et al., 2021)	15.3	19.5	21.0	21.9	22.4	15.3	19.5	21.0	21.9	22.2	14.4
GraphRetro+CREBM	**16.3**	**20.1**	**21.6**	**22.3**	**22.7**	**16.3**	**20.2**	**21.6**	**22.3**	**22.7**	-
Template-free											
Transformer (Karpov et al., 2019)	31.3	40.4	44.7	47.2	48.9	31.2	40.5	45.1	47.3	48.7	26.7
Transformer+CREBM	35.0	43.4	46.7	48.5	49.7	34.9	43.5	46.6	48.4	49.6	-
Megan (Sacha et al., 2021)	18.8	29.7	37.2	42.6	45.9	19.5	28.0	33.2	36.4	38.5	32.9
FusionRetro (Liu et al., 2023b)	37.5	45.0	48.2	50.0	50.9	37.5	45.0	48.3	50.2	51.2	**33.8**
FusionRetro+CREBM	**39.4**	**46.6**	**49.3**	**50.7**	**51.5**	**39.6**	**46.7**	**49.5**	**51.0**	**51.7**	-

Results

Table 2. Summary of results with our CREBM in terms of top-1 accuracy (%) on routes of different depths.

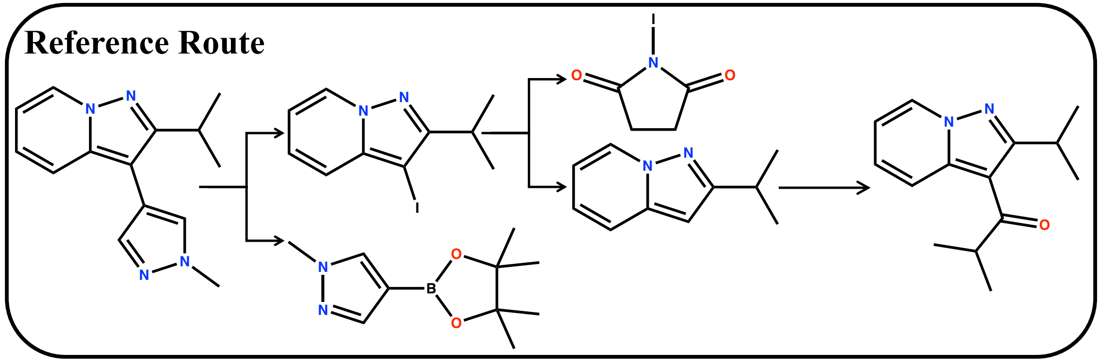
Model	Retro*-0				
	2	3	4	5	6
Neuralsym	46.1	42.0	33.7	37.3	40.8
+CREBM	+2.2	+2.1	+4.9	+2.9	+5.0
GLN	46.4	39.0	32.6	25.3	21.8
+CREBM	+2.7	+3.1	+4.7	+7.1	+14.5
Transformer	39.3	30.1	20.9	15.2	16.2
+CREBM	+2.4	+5.3	+3.0	+5.6	+10.6

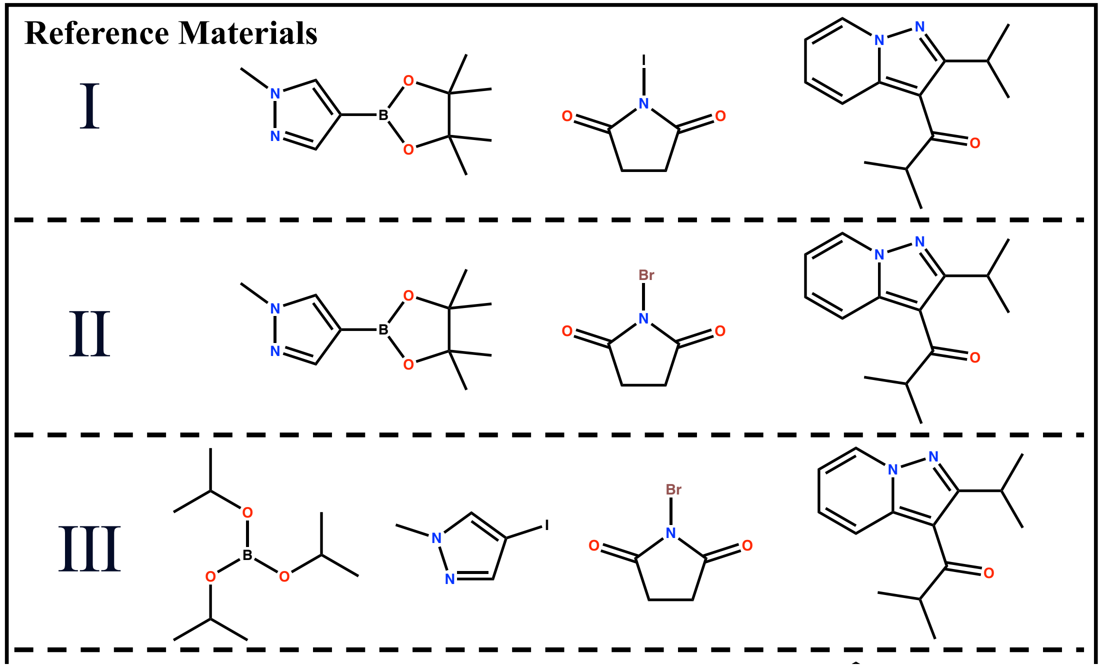
Ablation Study

Table 3. Ablation study on our energy function.

Metric for Ranking	Top-1	Δ_1	Top-5	Δ_2
$-\log P_{Retro}(\mathcal{T} \mid m_{tar})$	42.0	0	54.3	0
$-\log P_{Retro}(\mathcal{T} \mid m_{tar}) + E_\theta(\mathcal{T} \mid m_{tar}, c)$	44.5	+2.5	55.2	+0.9
$E_\theta(\mathcal{T} \mid m_{tar}, c)$	34.1	-7.9	53.1	-1.2
$-\log P_{Retro}(\mathcal{T} \mid m_{tar}) - E_\theta(\mathcal{T} \mid m_{tar}, c)$	19.2	-22.8	44.6	-9.7

Preference Visualization





Code & arXiv

Code: https://github.com/SongtaoLiu0823/CREBM
arXiv: https://arxiv.org/pdf/2406.02066

Q&A

Thank you!
And any question?