RetroRecs: Enhancing single-step Retrosynthesis via Reaction Context Sequence

Luojian Xie[†,1], Zehui Wang[†,1], Zixian Cheng[†], Dongliang Chen[†], Huibin Wang[†], Liang Dou[†2], Ying Qian[†*2]

[†]School of Computer Science and Technology, East China Normal University, Shanghai, China
[*]Shanghai Institute of AI for Education, East China Normal University, Shanghai, China
[2]{ldou, yqian}@cs.ecnu.edu.cn

***Abstract*—Retrosynthesis identifies reactants to synthesize a target molecule, which is essential for drug discovery. Graph-edit-based methods transform products into reactants via molecular edits. While existing approaches reflect chemical transformations, they often model molecular structures and edit content separately. To bridge this gap, we propose RetroRecs, a graph-edit-based retrosynthesis model that unifies this encoding. RetroRecs builds a structured Reaction Context Sequence that unifies the dynamic encoding of intermediate molecules with the contextual embedding of each molecular edit. This design supports stepwise and correlation-aware retrosynthetic prediction. This linkage is established through two modules in each edit prediction step: the Reaction Context Reasoner (RCR), which analyzes relationships among reaction contexts (describing the "what, where, and when" of past edits), and the Molecular Environment Perceiver (MEP), which refines this contextual reasoning by incorporating the current synthon's structural features. Additionally, we propose a contrastive learning objective, guided by reaction types, to enhance the model's ability to differentiate between various edit process groups. These designs allow the model to make edit decisions that are sequentially coherent and thus chemically valid. Experiments on USPTO-50K show RetroRecs achieving 92.3% and 70.2% in top-10 exact match and round-trip accuracy, setting a new state-of-the-art for semi-template retrosynthesis models.**

***Index Terms*—retrosynthesis, graph edit, reaction context, contrastive learning, semi-template method.**

I. Introduction

Organic synthesis is a vital field in chemistry, bearing significant implications for both drug discovery and materials science. Retrosynthesis, first proposed by Corey [1], aims to infer possible reactants from given products. While traditional computer-aided synthetic planning (CASP) systems [2] relied heavily on expert knowledge , AI now enables more efficient, data-driven pathway planning [3].

Single-step retrosynthesis prediction is the pivot of retrosynthetic planning, which directly identifies possible reactants from a given product. Recently, deep learning-based approaches have demonstrated promising performance on this task [4]. These methods can typically be classified into three categories based on their dependence on predefined reaction templates by chemical experts: template-based, semi-template-based, and template-free approaches. Among these, semi-template methods are considered to strike a balance between interpretability and diversity [5], as these methods regard the retrosynthesis task as a sequence of molecular graph edits that transform the product to intermediate synthons and finally to reactants.

Graph-edit-based (one type of semi-template) retrosynthesis requires tracking how the synthon evolves and how successive edits depend on one another. Existing methods typically cover one side: Graph2Edits [5] updates synthon graphs but leaves edit correlations implicit, whereas MARS [6] models edit sequences while keeping the synthon representation fixed, missing intermediate changes. We unify these factors with a Reaction Context (RC): a structured triplet capturing where an edit acts, what modification it makes, and when it occurs. Stacking the RCs over all steps yields a Reaction Context Sequence, a memory that preserves positional, temporal, and logical links among edits and the evolving molecular structure.

Building upon the Reaction Context Sequence, we propose **RetroRecs**, a semi-template-based retrosynthesis model. RetroRecs contains two collaborative modules for context-aware prediction: the Reaction Context Reasoner (RCR) reasons over historical reaction contexts to capture inter-context correlations, while the Molecular Environment Perceiver (MEP) refines the predicted context by incorporating structural information from the current synthon. We further introduce a contrastive learning objective to improve the model's understanding of overall reaction patterns. The effectiveness of RetroRecs is demonstrated on USPTO-50K, it achieves state-of-the-art performance among semi-template-based methods. On the more challenging USPTO-FULL dataset, RetroRecs continues to outperform other semi-template methods, demonstrating generalization capability across large-scale reaction spaces.

Our contributions can be summarized as follows:

- We introduce the concept of a **Reaction Context**, which explicitly encodes each molecular edit step, and formulate the **Reaction Context Sequence** as a memory of the reaction trajectory.
- We design two collaborative modules: the **Reaction Context Reasoner (RCR)**, which infers the next edit by attending to historical contexts, and the **Molecular Environment Perceiver (MEP)**, which adapts the prediction to the current synthon structure.
- Our model achieves a top-10 accuracy of 92.3% on the publicly available USPTO-50K dataset, setting a new state-of-the-art for semi-template retrosynthesis methods.

[1]Equal contribution.
[2]Corresponding authors.

In the Round-Trip experiment, which evaluates the reasonableness of reaction inferences, RetroRecs significantly outperforms existing methods, achieving a top-10 Round-Trip accuracy of 70.2%.

II. Related Work

A. Template-based Method

Template-based retrosynthesis applies predefined reaction templates to break a product into reactants, with recent studies focused on better template selection. **RetroSim** [7] ranks candidates by fingerprint similarity to database molecules, while **LocalRetro** [8] builds separate atom- and bond-level templates and adds attention to capture long-range graph context. **RetroKnn** [9] retrieves suitable templates through a reaction-vector database learned from past reactions, and **GLN** [10] treats templates as logical rules, modeling their joint probability with reactants via a Conditional Graph Logic Network. Despite these advances, all depend on a fixed template library, limiting generalization to unseen or sparsely represented reaction types.

B. Template-free Method

Template-free retrosynthesis generates reactants directly from a product. The earliest work used an LSTM to translate product SMILES into reactants [11]. **Graph2SMILES** [12] replaced the LSTM with a graph encoder plus Transformer decoder, and follow-ups refined this line: **SCROP** adds a self-correction transformer to fix syntax errors [13]; the **Augmented Transformer** augments data with SMILES permutations [14]. **NAG2G** [15] is the first autoregressive *graph* generator in this setting. More recent models—**RetroBridge** [16], **UAlign** [17], and **Retroformer** [18]—boost accuracy by explicitly aligning product and reactant structures. More recently, **LCVAT** [19] introduces layer-wise latent variables into Transformer architectures to model diverse yet feasible retrosynthetic outcomes and capture long-range dependencies.

C. Semi-template Method

Semi-template methods model retrosynthesis through explicit reaction centers and graph edits, offering a compromise between template-based interpretability and template-free flexibility. **G2Retro** [20] decomposes the task into reaction center prediction and synthon completion with two separate models, which may weaken information transfer between the two stages [18]. **Graph2Edits** [5] improves this framework by jointly modeling reaction centers and molecular edits. **RetroLEE** [21] models cross-synthon interactions to better capture structural dependencies during sequential retrosynthetic edits. **MEGAN** [22] and **MARS** [6] further exploit edit-sequence information, but their synthon representations remain fixed during the editing process. **G2Gs** [23] introduces latent variables to guide reaction pathway inference and aggregates information from multiple synthons. In contrast, **RetroRecs** jointly models structural evolution and edit dependency through a unified **Reaction Context** representation, enabling context-aware edit prediction conditioned on both historical transformations and the current molecular state.

III. Method

A. Graph Encoder

In this work, a D-MPNN graph encoder is used to encode the structure of the synthon. The graph encoder accepts the synthon G_s^t of the t-th step editing as input and outputs the atomic features, chemical bond features, and graph-level features of the synthon. It can be specifically expressed as the following formula:

$$h_i^t, h_{ij}^t, h_{\mathcal{G}}^t = \text{Encoder}(G_s^t), \tag{1}$$

where G_s^t is the t-th synthon, h_i^t is the atom feature for $v_i \in G_s^t$, h_{ij}^t is the bond embedding for $b_{ij} \in G_s^t$ and $h_{\mathcal{G}}^t$ is the graph embedding for G_s^t. A LayerNorm layer is appended after the D-MPNN to stabilize the feature distribution.

B. Construction of Reaction Context Sequence

At each retrosynthesis step i, a corresponding **Reaction Context** RC_i is constructed based on the current synthon graph G_i and the applied edit. Formally, we define $RC_i = (c_i, e_i, t_i)$, where c_i denotes the reaction center, e_i denotes the edit type, and t_i is the step index. To obtain a vector representation of RC_i, we encode it as:

$$RC_i = h(c_i) + \text{Emb}(e_i) + \text{PosEmb}(i), \tag{2}$$

where $h(c_i)$ is the atom or bond embedding for the reaction center c_i derived from the current synthon G_i, $\text{Emb}(e_i)$ is a learnable embedding of the edit type, and $\text{PosEmb}(i)$ is a step index embedding fulfilled by sinusoidal positional encoding [24]. For RC_0, we define an *Initialize* operation in which the entire molecular graph is regarded as the reaction center.

The set of all past contexts $\{RC_0, RC_1, ..., RC_{i-1}\}$ constitute the **Reaction Context Sequence** at step i, denoted as $RC_{(<i)}$. This sequence serves as a dynamic memory of the retrosynthesis trajectory and provides the basis for context-aware edit prediction in subsequent steps.

C. Reaction Context Reasoner (RCR)

The primary motivation behind the **Reaction Context Reasoner (RCR)** is the inherent sequential and contextual dependency of retrosynthetic steps. The RCR is designed to explicitly model and leverage the historical information stored in the Reaction Context Sequence, denoted as $RC_{(<i)} = \{RC_0, RC_1, \ldots, RC_{i-1}\}$, to infer the next plausible Reaction Context. By explicitly capturing the long-range correlations among past edits, the RCR enables the model to make contextually aware predictions, guiding the generation of chemically coherent edit sequences and ultimately improving prediction accuracy.

To effectively model the accumulated reaction history and its sequential dependencies, we employ a **masked self-attention layer**, which allows the RCR to attend to all previous encoded reaction contexts in $RC_{(<i)}$. Formally, the RCR computes the initial prediction for the i-th Reaction Context, denoted as $\hat{RC}_i$, as:

$$\hat{RC}_i = \text{MSA}(RC_{(<i)}) \tag{3}$$

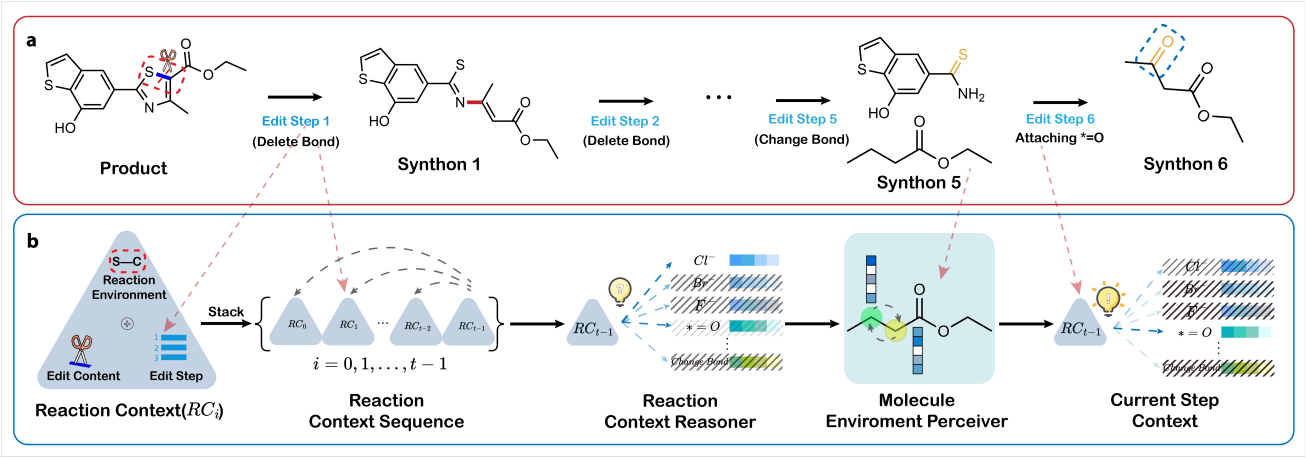

Fig. 1. **Workflow of RetroRecs. a** Example scenario of semi-template retrosynthesis. **b** Each edit step generates a Reaction Context RC_i, which encodes the edit content, the reaction step index, and the surrounding molecular environment. These contexts are stacked into the Reaction Context Sequence that records the reaction history. The Reaction Context Reasoner (RCR) attends to this sequence to infer the next reaction context, which is subsequently refined by the Molecular Environment Perceiver (MEP) based on the current synthon structure. The refined context is then used to predict the current molecular edit.

where MSA is the multi-head self-attention layer with causal masking, the output $\hat{RC}_i$ from the RCR serves as an initial, context-informed indicator for the next Reaction Context. This indicator is then passed to the **Molecular Environment Perceiver (MEP)** module for further refinement, incorporating the structural features of the current synthon.

D. Molecular Environment Perceiver (MEP)

The **Reaction Context Reasoner** (RCR) models historical edits but still needs to consider the evolving structure of the current synthon G_s^i. To anchor predictions in that structure, the **Molecular Environment Perceiver** (MEP) refines the preliminary context $\hat{RC}_i$ with information from G_s^i.

We first collect atom, bond, and graph embeddings from the encoder,

$$S_i = (h_i^t, h_{ij}^t, h_{\mathcal{G}}^t) \in \mathbb{R}^{(|V_s|+|B_s|+1)\times d}.$$

A multi-head cross-attention layer treats $\hat{RC}_i$ as the query and S_i as keys/values, producing the context aligned to the current molecular environment:

$$\tilde{RC}_i = \text{CrossAttn}(\hat{RC}_i, S_i). \tag{4}$$

Grounding the historical context in G_s^i filters out structurally infeasible reaction contexts and steers the model toward chemically valid transformations in the next prediction step.

Grounding the historical context in G_s^i increases the likelihood that the model selects the correct edit and steers the model toward chemically valid transformations in the next prediction step, because the current synthon environment is explicitly taken into account.

E. Reaction Mechanism Contrastive Learning

Chemical reactions are governed by underlying mechanisms that span the entire transformation from reactants to products. To incorporate this information into molecular representation learning, we construct a reaction mechanism embedding by combining product and reactant features:

$$z = \text{Linear}(h_{\mathcal{G}}^{(p)} || h_{\mathcal{G}}^{(r)}), \tag{5}$$

where $h_{\mathcal{G}}^{(p)}$ and $h_{\mathcal{G}}^{(r)}$ denote the graph-level embeddings of the product and reactant, respectively. The mechanism embedding z is used only to define a contrastive objective during training and is not involved in the inference process.

We adopt supervised contrastive learning [25], where reactions sharing the same annotated reaction type in USPTO-50K are encouraged to have closer mechanism embeddings. The contrastive loss is formulated as:

$$\mathcal{L}_{MC} = \sum_{i \in I} \frac{1}{|P(i)|} \sum_{p \in P(i)} \log \frac{\exp(z_i \cdot z_p / \tau)}{\sum_{n \in N(i)} \exp(z_i \cdot z_n / \tau)}. \tag{6}$$

where $P(i)$ denotes the set of samples with the same reaction type as i, and τ is the temperature parameter. $N(i)$ are negative samples to i.

F. Edit Actions Prediction

For a given product G_p, our goal is to generate T edits (e_1,...,e_T) and reaction centers (c_1,...,c_T), and then apply these edits to the product to form a series of synthons $G_s^t (t = 1, \ldots, T-1)$, and finally reach the reactant G_r. Specifically, we combine the features of G_s^t and RC_t, to infer e_{t+1}, and then generate $G_s^{(t+1)}$. We initialize e_0 as a special "Initialize" edit, with $G_s^{(0)} = G_p$. We use Projector, i.e., an MLP layer, to project RC_t to the molecular representation space. Finally, we combine RC_t and the molecular representation to infer the

TABLE I
TOP-k ACCURACY FOR RETROSYNTHESIS PREDICTION ON USPTO-50K. THE METRIC VALUES OF OTHER MODELS IN THE TABLE ARE TAKEN FROM THEIR ORIGINAL PAPERS. THE BEST PERFORMANCE IS IN **BOLD**.

Model	Top-k Accuracy (%)							
	Reaction class known				Reaction class unknown			
	1	3	5	10	1	3	5	10
Template-Based								
GLN	64.2	79.1	85.2	90.0	52.5	69.0	75.6	83.7
LocalRetro	63.9	86.8	92.4	**96.3**	53.4	77.5	**85.9**	**92.4**
RetroExplainer	**66.8**	**88.0**	**92.5**	95.8	**57.7**	**79.2**	84.8	91.4
Template-Free								
Retroformer	64.0	82.5	86.7	90.2	53.2	71.7	76.6	82.1
NAG2G	**67.2**	86.4	90.5	93.8	**55.1**	76.9	83.4	89.9
UAlign	66.2	**86.9**	**91.9**	**95.5**	53.6	**77.6**	**84.6**	**90.3**
EditRetro[a]	-	-	-	-	49.8	64.1	67.8	72.9
Semi-Template-Based								
G2Gs	61.0	81.3	86.0	88.7	48.9	67.6	72.5	75.5
MARS	66.2	85.8	90.2	92.9	54.6	76.4	83.3	88.5
G2Retro	63.6	83.6	88.4	91.5	54.1	74.1	81.2	86.7
MEGAN	60.7	82.0	87.5	91.6	48.1	70.7	78.4	86.1
Graph2Edits	67.1	87.5	91.5	93.6	55.1	77.3	83.4	89.4
State2Edits	**67.7**	87.7	90.0	94.0	55.4	78.0	82.5	89.1
RetroRecs (Ours)	67.3	**88.8**	**93.4**	**96.5**	**56.7**	**79.5**	**86.4**	**92.3**

[a] All results are reported from the original papers or corresponding official review documents of each model. without test set augment.

edit action. The prediction in step t can be expressed as the following formula:

$$\tilde{RC}_i = \text{Projector}(\tilde{RC}_i), \tag{7}$$

$$s_{i,a}^{(t)} = \text{MLP}_a(h_i^{(t)} || \tilde{RC}_i), \tag{8}$$

$$s_{(ij,b)}^{(t)} = \text{MLP}_b(h_{ij}^{(t)} || \tilde{RC}_i), \tag{9}$$

$$s_{\mathcal{G}}^{(t)} = \text{MLP}_G(h_{\mathcal{G}}^{(t)} || \tilde{RC}_i), \tag{10}$$

where MLP_a, MLP_b, and MLP_G are Multilayer Perceptrons, $s_{i,a}^{(t)}$, $s_{(ij,b)}^{(t)}$, and $s_{\mathcal{G}}^{(t)}$ are logits for the atom edits $a \in E_{bond}$, the bond edits $b \in E_{bond}$, and the termination symbol E_T. Reaction center c_t and e_t are decoded based on the prediction result (either one of $s_{i,a}^{(t)}$, $s_{(ij,b)}^{(t)}$, $s_{\mathcal{G}}^{(t)}$), and the structural transformation is thus fulfilled by applying the chosen edit on the reaction center of the current synthon graph.

G. Training and Inference

We use the teacher-forcing training method, directly using the ground truth from the previous edits as input to predict each step's edits. The loss for edit prediction is as follows:

$$\mathcal{L}_{Edit} = -\sum_{t \in \text{T}} \sum_{(\mathcal{G}_s, E)} \left(\sum_{b \in E_{\text{bond}}} y_{(ij,b)}^{(t)} \log \left(s_{(ij,b)}^{(t)} \right) + \sum_{a \in E_{\text{atom}}} y_{(i,a)}^{(t)} \log \left(s_{(i,a)}^{(t)} \right) + y_{\mathcal{G}}^{(t)} \log \left(s_{\mathcal{G}}^{(t)} \right) \right). \tag{11}$$

where y is an indicator function to identify whether s is the ground-truth edit. The final loss $\mathcal{L}$ for the training process is denoted as:

$$\mathcal{L} = \mathcal{L}_{Edit} + \alpha \cdot \mathcal{L}_{MC}, \tag{12}$$

where α is a weight coefficient for contrastive learning loss, which is set to 0.1 in this work.

H. Computational Efficiency

RetroRecs is implemented on top of a D-MPNN backbone and contains about 3.6 M parameters, which is markedly smaller than most Transformer-based retrosynthesis models (e.g., Graph2SMILES 17.4 M). Although Graph2Edits is even more compact (1.5 M), RetroRecs delivers substantially higher accuracy and also surpasses models of comparable size, such as RetroBridge (4.4 M).

Theoretical complexity further supports this observation. The base D-MPNN encoder requires $O(Ed^2)$ operations, where E is the number of bonds and d the hidden dimension. The Reaction-Context Refinement (RCR) module processes a short edit sequence (typically less than 10 steps), so its cost is effectively constant. The Molecular-Environment Projection (MEP) module applies a cross-attention between atom embeddings and the current reaction context, contributing another $O(Ed^2)$. The overall per-step complexity matches that of Graph2Edits, differing only by a constant factor.

IV. EXPERIMENTS

A. Datasets

In this study, we evaluate RetroRecs on the well-established USPTO-50K dataset [26], a benchmark for retrosynthesis tasks containing 50,016 atom-mapped reactions across 10 reaction types, derived from US patent literature. Following the dataset splits introduced by [7], the reactions are divided into training, validation, and test sets comprising 40k, 5k, and 5k reactions, respectively. To ensure no information leakage, we preprocess

the data by canonicalizing SMILES strings and reassigning atom maps, as described by [27]. We further validate our model on the larger dataset USPTO-FULL [10]. For the extraction of the molecular edit sequences, we adopted the same approach as Graph2Edits [5].

B. Implementation Details

We employ a D-MPNN encoder with 10 message passing layers, 256-dimensional node embeddings, and a 512-unit hidden layer. Training on USPTO-50K with a single RTX-4090 finishes in roughly 20 h; USPTO-FULL uses the same setup. Adam optimizer is used with an initial learning rate of 1.5×10^{-3} (USPTO-50K) or 1.5×10^{-4} (USPTO-FULL), batch size 64, and 0.15 dropout on node features. The learning rate is reduced by 0.8 when validation accuracy fails to improve by 0.01 over five epochs. Unless noted, the temperature is $\tau = 0.1$. Both the Reaction Context Reasoner and Molecular Environment Perceiver consist of a single 256-dimensional attention layer.

C. Evaluation Metrics

We evaluate the performance of RetroRecs primarily using Top-k accuracy. Since Top-k accuracy alone doesn't fully capture the model's comprehension of chemical reactions, considering that multiple sets of reactants can potentially lead to the same product, we include Round-Trip accuracy to further gauge the model's effectiveness and the chemical feasibility of its predictions. This metric measures the proportion of predicted reactants that correctly lead back to the product through a transformer-based forward reaction model [28]. The calculation of round-trip accuracy at Top-k is deemed as: Round-trip(k) $= \frac{1}{N\times K}\sum_{i=1}^{N}\sum_{j=1}^{K}\mathbb{I}(\text{Reach Ground Truth Product})$, where N represents the number of test samples, and K is the number of predicted products per reactant.

D. Performance

a) Baselines: The selected baselines are grouped into three categories: (i) Template-based methods, which include GLN [10], LocalRetro [8], and RetroExplainer [29]; (ii) Template-free methods, including EditRetro [4], Retroformer [18], NAG2G [15], UAlign [17]; (iii) Semi-template methods, which consist of G2Retro [20], G2Gs [23], MARS [6], MEGAN [22], State2Edits [30] and Graph2Edits [5].

b) Top-k Accuracy: Table I shows that RetroRecs reaches 56.7% / 92.3% top-1 / top-10 accuracy on USPTO-50K without reaction-class information, and 67.3% / 96.5% when the class is known. These scores set a new state-of-the-art for semi-template methods and match template-based models from top-3 to top-10 despite their smaller search space. We further evaluated RetroRecs on the larger and more challenging USPTO-FULL benchmark, with results summarized in Table II. RetroRecs achieves competitive performance on this dataset, outperforming other semi-template methods listed. It is important to note that some reported results on USPTO-FULL, such as those for EditRetro [4], utilize 5-time testset augmentation, which can significantly inflate performance.

TABLE II
TOP-k ACCURACY OF MODELS ON USPTO-FULL.

Model	Top-1	Top-3	Top-5
LocalRetro	39.1	53.3	58.4
Graph2Edits	44.0	60.9	66.8
RetroPrime	44.1	–	–
EditRetro*	52.2	67.1	72.0
RetroRecs (Ours)	**47.5**	**62.7**	**67.4**

* 5 times augmentation of SMILES strings is performed in testing phase, leading to $5 \times k$ candidates for evaluation.

TABLE III
TOP-k ROUND-TRIP ACCURACY (%) ON USPTO-50K. BEST RESULTS ARE IN **BOLD**.

Model	Top-1	Top-3	Top-5	Top-10
Graph2SMILES	76.7	56.0	50.3	40.4
Retroformer	78.9	72.0	67.1	57.2
Graph2Edits	90.1	79.7	73.2	61.3
RetroRecs	**91.4**	**82.3**	**77.6**	**70.2**

c) Round-Trip: The Round-Trip results on the USPTO-50K dataset are presented in Table III. Our model achieves a top-1 accuracy of 91.4% and a top-10 accuracy of 70.2%, outperforming other models. RetroRecs surpasses Graph2Edits by a 14.5% relative improvement in top-10 accuracy in Round-Trip experiments. This highlights RetroRecs' ability to produce more reasonable reactants. The results also emphasize the importance of understanding the continuous transformation throughout the reaction process for effective retrosynthesis inference. The incorporation of reaction contexts enables the model to memorize and utilize historical molecular edits at each step.

E. Ablation Study

Ablation results (Table IV) confirm each component's utility. Using only the Reaction Context Reasoner (RCR) increases the top-1 accuracy, confirming the value of historical edit information; however, the top-3 metric remains essentially unchanged because the model still lacks knowledge of the current synthon environment. When only the Molecular Environment Perceiver (MEP) is active, the top-1 accuracy changes little, but its wider beam search space improves the top-10 result. Combining RCR and MEP leverages both historical and local cues, delivering consistent gains across all metrics.

Incorporating reaction center embeddings into the Reaction Context ensures chemical consistency. An ablation study that omits this component (Table V) lowers every top-k accuracy, confirming its central role in Reaction Context construction and overall model performance, distinguishing our RC modeling from earlier semi-template frameworks like MARS or MEGAN.

V. CONCLUSION

We introduced **RetroRecs**, a semi-template retrosynthesis framework that operates in a structured reaction-context sequence to predict chemically valid molecular edits. By jointly

TABLE IV
ABLATION STUDY ON USPTO-50K WITH REACTION CLASS UNKNOWN.

Module			Top-k Accuracy (%)			
RCR	**MEP**	**MC**	k=1	**3**	**5**	**10**
–	–	–	55.1	77.4	84.5	89.9
✓	–	–	55.6	78.7	85.5	91.6
–	✓	–	55.2	78.9	85.8	92.0
✓	✓	–	56.5	79.5	86.0	91.9
✓	✓	✓	**56.7**	**79.5**	**86.4**	**92.3**

TABLE V
IMPACT OF REACTION CENTER EMBEDDING ON TOP-k ACCURACY.

Model	**Top-1**	**Top-3**	**Top-5**	**Top-10**
RetroRecs (w/o $h_{(c_i)}$)	54.5	78.2	85.2	88.4
RetroRecs	**56.7**	**79.5**	**86.4**	**92.3**

capturing historical edit dependencies and the current synthon structure, RetroRecs produces more coherent and chemically valid edit predictions. On USPTO-50K, RetroRecs achieves state-of-the-art top-k and round-trip accuracies, demonstrating both statistical and practical strength. In future work, we will extend this framework to multi-step planning and incorporate richer chemical constraints and contextual information.

ACKNOWLEDGMENT

This work is supported by the National Key Research and Development Program of China under Grant 2024YFC3308500.

REFERENCES

[1] E. J. Corey, "The logic of chemical synthesis: multistep synthesis of complex carbogenic molecules (nobel lecture)," *Angewandte Chemie International Edition in English*, vol. 30, no. 5, pp. 455–465, 1991.

[2] S. Szymkuć, E. P. Gajewska, T. Klucznik, K. Molga, P. Dittwald, M. Startek, M. Bajczyk, and B. A. Grzybowski, "Computer-assisted synthetic planning: the end of the beginning," *Angewandte Chemie International Edition*, vol. 55, no. 20, pp. 5904–5937, 2016.

[3] C. W. Coley, D. A. Thomas III, J. A. Lummiss, J. N. Jaworski, C. P. Breen, V. Schultz, T. Hart, J. S. Fishman, L. Rogers, H. Gao *et al.*, "A robotic platform for flow synthesis of organic compounds informed by ai planning," *Science*, vol. 365, no. 6453, p. eaax1566, 2019.

[4] Y. Han, X. Xu, C.-Y. Hsieh, K. Ding, H. Xu, R. Xu, T. Hou, Q. Zhang, and H. Chen, "Retrosynthesis prediction with an iterative string editing model," *Nature Communications*, vol. 15, no. 1, p. 6404, 2024.

[5] W. Zhong, Z. Yang, and C. Y.-C. Chen, "Retrosynthesis prediction using an end-to-end graph generative architecture for molecular graph editing," *Nature Communications*, vol. 14, no. 1, p. 3009, 2023.

[6] J. Liu, C. Yan, Y. Yu, C. Lu, J. Huang, L. Ou-Yang, and P. Zhao, "Mars: a motif-based autoregressive model for retrosynthesis prediction," *Bioinformatics*, vol. 40, no. 3, p. btae115, 2024.

[7] C. W. Coley, L. Rogers, W. H. Green, and K. F. Jensen, "Computer-assisted retrosynthesis based on molecular similarity," *ACS central science*, vol. 3, no. 12, pp. 1237–1245, 2017.

[8] S. Chen and Y. Jung, "Deep retrosynthetic reaction prediction using local reactivity and global attention," *JACS Au*, vol. 1, no. 10, pp. 1612–1620, 2021.

[9] S. Xie, R. Yan, J. Guo, Y. Xia, L. Wu, and T. Qin, "Retrosynthesis prediction with local template retrieval," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 4, 2023, pp. 5330–5338.

[10] H. Dai, C. Li, C. Coley, B. Dai, and L. Song, "Retrosynthesis prediction with conditional graph logic network," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[11] B. Liu, B. Ramsundar, P. Kawthekar, J. Shi, J. Gomes, Q. Luu Nguyen, S. Ho, J. Sloane, P. Wender, and V. Pande, "Retrosynthetic reaction prediction using neural sequence-to-sequence models," *ACS central science*, vol. 3, no. 10, pp. 1103–1113, 2017.

[12] Z. Tu and C. W. Coley, "Permutation invariant graph-to-sequence model for template-free retrosynthesis and reaction prediction," *Journal of chemical information and modeling*, vol. 62, no. 15, pp. 3503–3513, 2022.

[13] S. Zheng, J. Rao, Z. Zhang, J. Xu, and Y. Yang, "Predicting retrosynthetic reactions using self-corrected transformer neural networks," *Journal of chemical information and modeling*, vol. 60, no. 1, pp. 47–55, 2019.

[14] I. V. Tetko, P. Karpov, R. Van Deursen, and G. Godin, "State-of-the-art augmented nlp transformer models for direct and single-step retrosynthesis," *Nature communications*, vol. 11, no. 1, p. 5575, 2020.

[15] L. Yao, Z. Wang, W. Guo, S. Xiang, W. Liu, and G. Ke, "Node-aligned graph-to-graph generation for retrosynthesis prediction," *arXiv preprint arXiv:2309.15798*, 2023.

[16] I. Igashov, A. Schneuing, M. Segler, M. Bronstein, and B. Correia, "Retrobridge: Modeling retrosynthesis with markov bridges," *arXiv preprint arXiv:2308.16212*, 2023.

[17] K. Zeng, B. Yang, X. Zhao, Y. Zhang, F. Nie, X. Yang, Y. Jin, and Y. Xu, "Ualign: pushing the limit of template-free retrosynthesis prediction with unsupervised smiles alignment," *Journal of Cheminformatics*, vol. 16, no. 1, p. 80, 2024.

[18] Y. Wan, C.-Y. Hsieh, B. Liao, and S. Zhang, "Retroformer: Pushing the limits of end-to-end retrosynthesis transformer," in *International Conference on Machine Learning*. PMLR, 2022, pp. 22 475–22 490.

[19] H. Wang, Z. Wang, and Y. Qian, "Modeling diverse and feasible retrosynthesis via a layer-wise conditional variational transformer," in *2025 International Joint Conference on Neural Networks (IJCNN)*, 2025, pp. 1–10.

[20] Z. Chen, O. R. Ayinde, J. R. Fuchs, H. Sun, and X. Ning, "G 2 retro as a two-step graph generative models for retrosynthesis prediction," *Communications Chemistry*, vol. 6, no. 1, p. 102, 2023.

[21] Z. Wang, Z. Cheng, D. Chen, and Y. Qian, "Retrolee: Bridging the gap between synthons in single-step retrosynthesis," in *PRICAI 2025: Trends in Artificial Intelligence*, Y. Mei, C. Qian, Q. Bai, B. Xue, and S. Khanna, Eds. Singapore: Springer Nature Singapore, 2026, pp. 412–426.

[22] M. Sacha, M. Błaz, P. Byrski, P. Dabrowski-Tumanski, M. Chrominski, R. Loska, P. Włodarczyk-Pruszynski, and S. Jastrzebski, "Molecule edit graph attention network: modeling chemical reactions as sequences of graph edits," *Journal of Chemical Information and Modeling*, vol. 61, no. 7, pp. 3273–3284, 2021.

[23] C. Shi, M. Xu, H. Guo, M. Zhang, and J. Tang, "A graph to graphs framework for retrosynthesis prediction," in *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, ser. Proceedings of Machine Learning Research, vol. 119. PMLR, 2020, pp. 8818–8827.

[24] A. Vaswani, "Attention is all you need," *Advances in Neural Information Processing Systems*, 2017.

[25] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," *Advances in neural information processing systems*, vol. 33, pp. 18 661–18 673, 2020.

[26] N. Schneider, N. Stiefl, and G. A. Landrum, "What's what: The (nearly) definitive guide to reaction role assignment," *Journal of chemical information and modeling*, vol. 56, no. 12, pp. 2336–2346, 2016.

[27] V. R. Somnath, C. Bunne, C. Coley, A. Krause, and R. Barzilay, "Learning graph models for retrosynthesis prediction," *Advances in Neural Information Processing Systems*, vol. 34, pp. 9405–9415, 2021.

[28] P. Schwaller, T. Laino, T. Gaudin, P. Bolgar, C. A. Hunter, C. Bekas, and A. A. Lee, "Molecular transformer: a model for uncertainty-calibrated chemical reaction prediction," *ACS central science*, vol. 5, no. 9, pp. 1572–1583, 2019.

[29] Y. Wang, C. Pang, Y. Wang, J. Jin, J. Zhang, X. Zeng, R. Su, Q. Zou, and L. Wei, "Retrosynthesis prediction with an interpretable deep-learning framework based on molecular assembly tasks," *Nature Communications*, vol. 14, no. 1, p. 6155, 2023.

[30] Z. Ye, L. Yu, and F. Ma, "Semi-template framework for retrosynthesis prediction using graph neural network," *Pattern Recognition*, p. 111825, 2025.