Cite this: DOI: 00.0000/xxxxxxxxxx

AI-Driven Synthetic Route Design with Retrosynthesis Knowledge[†]Shoichi Ishida,^a Kei Terayama,^{bcd} Ryosuke Kojima,^d Kiyosei Takasu,^a and Yasushi Okuno^{*cd}

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

Computer-aided synthesis planning (CASP) aims to assist chemists in performing retrosynthetic analysis for which they exploit their experiments, intuition, and knowledge. Recent breakthroughs in machine learning techniques, including deep neural networks, have significantly improved data-driven synthetic route designs without human interventions. However, such CASP applications are yet to incorporate retrosynthesis knowledge sufficiently into their algorithms to reflect chemists' way of thinking flexibly. In this study, we developed a hybrid CASP application of data-driven techniques and various retrosynthesis knowledge called "ReTReK" that integrates the knowledge as adjustable parameters into an evaluation for promising search directions. Experimental results showed that ReTReK successfully searched synthetic routes based on the specified retrosynthesis knowledge, and the results indicated that the synthetic routes searched with the knowledge were preferred to those without knowledge. The concept of integrating retrosynthesis knowledge as adjustable parameters into data-driven CASP applications is expected to contribute to further their development and spread them to chemists widely.

1 Introduction

Since the 1960s, various computer-aided synthesis planning (CASP) applications have been developed to emulate chemists' thinking and help organic synthesis chemists in their work^{1–9}. The CASP applications have played an important role in definable parts of synthesis (e.g., considering chemical structural characteristics and retrosynthetic tree size), while indefinable parts of synthesis (e.g., chemists' intuition) and the opportunity to contribute creativity of retrosynthetic analysis have been left to the chemists¹. As an underlying chemists' intuition, Corey formalized the concept of retrosynthesis (retrosynthesis knowledge) and major types of strategies (e.g., transform- and topological-based), and he stated that retrosynthetic analysis is efficiently performed by the simultaneous use of as many different independent strategies as possible¹⁰. To select optimal strategies, the chemists' knowledge of chemistry and their experiments are essential; optimal strategies depend on molecules, persons, and situations (e.g., lead optimizations and large-scale syntheses of drug

candidates)¹¹.

CASP approaches are generally classified into two types: rule-based^{8,12} and data-driven approaches^{6,9}. For rule-based approaches, they employ manually-encoded transformations considering, such as stereochemistry and electronic effect⁸. For instance, an excellent rule-based CASP application—Chematica⁸ (now re-branded as SynthiaTM)—provides considerable discretion for chemists to perform retrosynthetic analysis based on their way of thinking and has come into use globally^{13,14}. However, the rule-based approach requires the extreme efforts of many experts and does not keep up with exponentially growing chemistry knowledge¹⁵.

On the other hand, data-driven CASP aims to extract knowledge related to transformations from numerous reaction records automatically to discover synthetic routes¹⁶. Recent breakthroughs in deep learning (DL)^{17,18}, along with the widespread availability of reaction records^{19,20} and open-source codes^{21–24}, have improved the core techniques of data-driven CASP such as 1-step (retro)synthetic reaction predictions^{25–28} and multi-step synthetic route searches^{9,29–32}. For the reaction prediction models, various representations of molecules (e.g., fingerprints²⁵, SMILES strings^{26,27,33}, and graphs^{28,34}) and their suitable DL techniques were used, which showed promising performances. For the search algorithms, the Monte Carlo tree search (MCTS)^{9,24,35,36}, depth-first proof number search^{31,37}, and graph-based exploration methods^{30,32} were used to obtain optimal or possible synthetic routes. Several outstanding data-driven CASP applications have been on the stage of practical use

^aGraduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshidashimo-Adachicho, Sakyo-ku, 606-8501 Kyoto, Japan.

^bGraduate School of Medical Life Science, Yokohama City University, 1-7-29, Suehirocho, Tsurumi-ku, Yokohama 230 0045 Kanagawa, Japan.

^cMedical Sciences Innovation Hub Program, RIKEN Cluster for Science, Technology and Innovation Hub, 1-7-22 Suehiro-cho, Tsurumi-ku, 230-0045 Kanagawa, Japan.

^dGraduate School of Medicine, Kyoto University, 53 Shogoin-Kawaharacho, Sakyo-ku, 606-8507 Kyoto, Japan. E-mail: okuno.yasushi.4c@kyoto-u.ac.jp

[†] Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 00.0000/00000000.

Fig. 1 Complete workflow of ReTReK. ReTReK comprises Monte Carlo tree search (MCTS) and graph convolutional networks, and the retrosynthesis knowledge is incorporated into the selection step of the MCTS. The retrosynthesis knowledge is represented by four scores: convergent disconnection score (CDScore), selective transformation score (STScore), ring disconnection score (RDScore), and available substances score (ASScore).

in industries and laboratories^{9,30,38}; these applications have remarkably produced a revival of interest in CASP research^{39–41}.

However, for considering practical retrosynthesis, most data-driven CASP applications are lacking in terms of reflecting or flexibly applying the chemists' way of thinking. Search algorithms in such applications depend on their naive scoring functions for evaluating whether a synthetic position while searching is preferable^{9,35}; this implies there are few opportunities to learn the strategies for retrosynthetic analysis. Moreover, the data-driven CASP approaches that incorporate generally used retrosynthesis knowledge have not been developed, and the effects of knowledge on search performance are yet to be investigated.

In this study, we developed a hybrid CASP application of data-driven and rule-based techniques called "ReTrosynthesis planning application using Retrosynthesis Knowledge (ReTReK)," which integrates the knowledge into an evaluation for a promising search direction. ReTReK accepts the knowledge as user-adjustable parameters so that users can easily decide which retrosynthesis knowledge to use and how much to reflect it. To represent the retrosynthesis knowledge of ReTReK, we prepared four scores—convergent disconnection score (CDScore), available substances score (ASScore), ring disconnection score (RDScore), and selective transformation score (STScore). The Reaxys reaction database¹⁹ was used to construct the ReTReK model. We evaluated and demonstrated the performance of ReTReK using molecules from the ChEMBL database⁴² and drug-like molecules^{43–48}. We successfully demonstrated that designed synthetic routes using ReTReK with retrosynthesis knowledge were preferable to those without knowledge. Further, we demonstrated that retrosynthesis knowledge improves the performance of solving target molecules, and we successfully guided the search direction of the MCTS. The proposed concept of integrating retrosynthesis knowledge as adjustable parameters into the data-driven CASP application will enhance the performance of existing or developing data-driven CASP applications. Our implementation is available on GitHub at <https://github.com/clinfo/ReTReK>.

2 Results and discussion

2.1 Construction of ReTReK

To implement the data-driven CASP application that reflects the retrosynthesis knowledge, ReTReK was constructed using MCTS, graph convolutional networks (GCN), and the four retrosynthesis knowledge scores (Fig. 1).

The basic algorithm of the MCTS comprises four steps: selection, expansion, rollout, and update. For the selection step, a tree policy was used to select a promising retrosynthetic tree position, and the retrosynthesis knowledge scores were involved in the policy. The GCN-based model (Fig. 2a) was used for the 1-step retrosynthetic reaction prediction model as a policy network in the expansion and rollout steps. Reaxys reaction records¹⁹ were used to train the model and prepare the starting materials; compounds obtained from ZINC database⁴⁹ were used as starting materials. By iterating the four steps, a retrosynthetic tree is expanded, thus attempting to determine a promising synthetic route.

2.2 Top-*n* accuracy of the GCN-based policy network

To determine the effective size of expansion step in MCTS, the top-*n* accuracy (up to 1,000) of the GCN-based policy network was calculated as shown in Fig. 2b. The 1-step retrosynthetic reaction prediction model aimed to prioritize 19,633 reaction templates for applying reaction templates to an input molecule. This study used reaction templates that considered a reaction center, the first-degree neighbors, and protecting groups because the previous study proposed this type of reaction template to maintain chemical integrity³⁵. Consequently, the top-1, 50, 100, 300, and 500 accuracies were 0.361, 0.906, 0.938, 0.968, and 0.976, respectively. Beyond the top-500 accuracy, the prediction performance increased insignificantly. Considering the previous study²⁸ and different sizes of reaction templates, the prediction performance was assumed to be equivalent to or better than the previous template-based 1-step retrosynthetic reaction prediction model⁹. Based on the result of the top-*n* accuracies, we evalu-

Fig. 2 (a) Model architecture of the GCN-based policy network. (b) Top- n accuracy of the model ranging from 1 to 1000. The top-1, 50, 100, 300, and 500 accuracies were 0.361, 0.906, 0.938, 0.968, and 0.976, respectively.

ated the effect of the MCTS expansion sizes and the retrosynthesis knowledge on the performance to solve target molecules using the top-50, 100, 300, and 500 predicted templates.

2.3 Effect of expansion sizes and retrosynthesis knowledge on the performances to solve target molecules

Figure 3 shows how expansion sizes and retrosynthesis knowledge influenced the performances for solving target molecules. The 161 molecules from the preprocessed ChEMBL dataset were used as the target molecules; the searches were performed by focusing on the expansion sizes (50, 100, 300, and 500) and six retrosynthesis knowledge patterns: no retrosynthesis knowledge (no knowledge), CDScore, ASScore, RDScore, STScore, and the four retrosynthesis knowledge (all knowledge). In most cases, solving performances were improved in proportion to the expansion sizes. The case of the STScore pattern and expansion size of 100 resulted in a lower number of solved molecules compared to the case of the same pattern and expansion size of 50. This result is attributed to the lack of MCTS iteration count because of the expansion size increase. For retrosynthesis knowledge, knowledge patterns except the STScore increased the number of solved molecules compared to the no knowledge pattern. The CDScore pattern with an expansion size of 500 showed the best solving performance and obtained 90 solved molecules, whereas the no knowledge pattern with the same size resulted in 59 solved molecules. Although the STScore pattern resulted in fewer solved molecules than the no knowledge pattern, the result was considered reasonable because STScore focuses on reactions with few by-products, which often leads to strict conditions on retrosynthesis.

Moreover, the search times to solve were compared in terms of expansion sizes and six retrosynthesis knowledge patterns, as shown in Fig. S1. Search times increased in proportion to expansion sizes, and this is because the expansion size increases expanded the search spaces of the MCTS. The median search times for each expansion size (50, 100, 300, and 500) were 32, 45, 133, and 294 seconds, respectively. The STScore pattern required shorter search times than the no knowledge pattern, although

Fig. 3 Comparison of the number of solved molecules in terms of the expansion sizes and retrosynthesis knowledge. The green, orange, blue, and pink bars correspond to the expansion size of 50, 100, 300, and 500, respectively. The results of retrosynthetic analyses with five retrosynthesis knowledge patterns; CDScore, ASScore, RDScore, STScore, and all of the knowledge (all knowledge); and without knowledge (no knowledge) are shown.

the patterns except STScore took longer search times. This result suggested that the STScore pattern could effectively identify synthetic routes compared to the other patterns, although STScore causes lower solving performance.

2.4 Effect of retrosynthesis knowledge on the search directions of Monte Carlo tree search

Figure 4 shows how six retrosynthesis knowledge patterns influenced the characteristics of searched synthetic routes in terms of four route scores (rSTScore, rCDScore, rASScore, and rRDScore). The route score was defined as the average corresponding retrosynthesis knowledge score in each step of the searched synthetic route. For easier comparison, each route score of the five knowledge patterns was standardized by the corresponding score of the no knowledge pattern. In terms of the standardized mean values of the STScore pattern in rSTScore, the CDScore pattern in rCDScore, the ASScore pattern in rASScore, and the RDScore in rRDScore, the values were 0.178, 0.555, 0.130, and 0.309, respectively. All values were positively shifted compared to the values of the no knowledge pattern, which showed that all four retrosynthesis knowledge scores successfully guided the directions of MCTS according to the characteristics of each knowledge. Intriguingly, the CDScore pattern led MCTS to selective transformation-oriented searches compared to the STScore pattern (the mean values of CDScore and STScore were 0.299 and 0.178, respectively). The convergent disconnection-oriented search is assumed to have more chances to split reactive centers into divided molecules because the CDScore aims to minimize the sizes of each divided molecule simultaneously.

Considering that the all knowledge pattern with regard to the rSTScore showed a higher route score than the CDScore and STScore patterns, this result suggests that there are synergistic effects of the retrosynthesis knowledge; however, this hypothesis needs further analysis.

Fig. 4 Evaluation of the effect of the retrosynthesis knowledge on the search directions of MCTS. Solved synthetic routes under the condition of the expansion size of 500 were used. Each route score was standardized based on the corresponding mean and standard deviation of the no knowledge pattern. The black, blue, orange, green, red, and grey plots represent the standardized route scores of all knowledge, STScore, SCScore, ASScore, RDScore, and no knowledge patterns, respectively. Rhombuses represent the mean values of each cases, and the confidence intervals at 95% confidence level are shown.

2.5 Demonstrations of ReTReK for drug-like molecules

To demonstrate retrosynthesis planning using ReTReK, we applied ReTReK to two drug-like molecules in the case of All Knowledge and No Knowledge patterns in this section. Figure 5 illustrates exemplary retrosynthetic routes to a molecule **1** known as a hepatitis B virus capsid inhibitor⁴³ explored by ReTReK with and without the “retrosynthesis knowledge.” Exploration with the retrosynthesis knowledge suggested a convergent route, successfully reflecting the specified knowledge scores (Fig. 5a). In the route, the target molecule **1** is disconnected into two main segments, iodophthalazinone **7** and pyridylboronic acid **12**, which can be converted to **1** by Suzuki coupling reaction. The key intermediates **7** and **12** could be retrosynthetically divided into three representative materials: hydroxyphthalazine **2**, benzyl alcohol **3**, and tri-halogenated pyridine **8**. Iodination of **2** and subsequent N-benzoylation with *p*-iodobenzyl iodide **5**, which can be obtained from **3**, would give 2-(iodobenzyl)phthalazin-1-one **6**. A reaction of **6** with copper cyanide would provide the intermediate **7**. The other intermediate **12** would be prepared from **8** by 3-step sequences (i.e., chlorination, incorporation of a boronic acid moiety, and amination with aminoalcohol **11**). In contrast, a straightforward route was presented through exploration without the retrosynthesis knowledge (Fig. 5b). Friedel-Crafts acylation of trihalopyridine **14** with 2-(chlorocarbonyl)benzoic acid (**13**) would give 2-(pyridinecarbonyl)benzoic acid **15**, which is further reacted with aminoalcohol **11** to afford tricyclol ketone **17**. Construction of the phthalazine ring could be performed by the reaction of **17** with *p*-bromobenzylhydrazine (**16**) to yield the precursor **18**. Finally, the introduction of a nitrile group into the benzyl group of **18** would provide the target molecule **1**. Additional demonstrations of other two drug-like molecules^{45,46} are shown in Fig. S2.

Furthermore, the effectiveness of the retrosynthesis knowledge was confirmed in several cases of the retrosynthetic analyses. In these cases, ReTReK with “Retrosynthesis Knowledge” succeeded in exploring retrosynthetic routes to target molecules, whereas no routes were found using ReTReK without “retrosynthesis knowledge” (Fig. 6 and Fig. S3). Figure 6 shows the retrosynthetic route to a molecule **19** known as a *Mycobacterium tuberculosis* thymidylate kinase (MtbTMPK) inhibitor⁴⁴ as a representative example. In the suggested route, **19** is disconnected in the center of the molecule, giving imidazo[1,2-*a*]pyridine-3-carboxamide **25** and 1-(piperidin-4-yl)pyrimidine-2,4-dione **29**. Compound **25** would be obtained from dibromide **24** by selective S_NAr reaction with an organometallic reagent, such as ethylmagnesium bromide, which is prepared from ethyl bromide. Dibromide **24** can be obtained from benzamide **22** by stepwise S_EAr bromination. Amide **22** would be provided from imidazopyridine-3-carboxylic acid (**20**) and N-Boc-benzylamine (**21**). Another intermediate **29** would be synthesized from 4-iodopiperidine (**27**) with pyrimidine-2,4-dione **28** by S_N2 reaction. **27** would be easily prepared from N-Boc **26**. From the viewpoint of the practical synthesis, deprotection of N-Boc group of the piperidine ring should be performed after N-alkylation of **28** with **26** to avoid oligomerization of **27** by self N-alkylation. Further demonstrations of molecules reported in Segler et al.⁹ were shown in Fig. S4. Overall, the results clearly demonstrate that “retrosynthesis knowledge” works well in retrosynthetic analyses using ReTReK. Considering these demonstrations, the ReTReK framework integrated with the retrosynthesis knowledge has the potential for further improve data-driven CASP applications.

3 Conclusions

We developed ReTReK, a hybrid CASP application of data-driven and rule-based techniques that can reflect and apply retrosynthesis knowledge on its algorithm flexibly. Through the evaluation of ReTReK with and without retrosynthesis knowledge, we showed that integrating the knowledge into data-driven CASP applications will help improve their performance and enhance the quality of the searched synthetic routes. We expect the concept of ReTReK to contribute to further developments and improvements of data-driven CASP applications; further, it is expected to bridge the applications and chemists by interactively reflecting chemists’ thoughts to the systems.

To obtain more realistic and preferable synthetic routes, we will address the further development of automatic reaction template extraction methods while maintaining chemical integrity. In this study, orphan atoms (atoms appearing only on one side of the reaction arrow) were included in the reaction templates to retain protecting and leaving groups automatically in the templates because the groups are often not recorded as reactants or products. Considering that the previous study manually defined such groups²², this definition of the template would contribute to further developments of the template extraction methods. For future work, we focus on integrating chemical knowledge such as electronic and quantum-mechanical aspects of chemical reaction into the template. As a solution to describing such information in the reaction templates, Chemical Terms— one of the functions

a) With retrosynthesis knowledge

b) Without retrosynthesis knowledge

Fig. 5 Comparison of (a) the synthetic route of the target compound⁴³ (hepatitis B virus capsid inhibitor) found by ReTReK with the retrosynthesis knowledge with (b) the one without the retrosynthesis knowledge. Weight parameters of the retrosynthesis knowledge, w_1 , w_2 , w_3 , and w_4 , are set to 5.0, 2.0, 0.5, and 2.0, respectively. The molecule 1 is the target.

a) With retrosynthesis knowledge

b) Without retrosynthesis knowledge

Not Found

Fig. 6 The synthetic route of the target compound⁴⁴ (MtbTMPK inhibitor) was found using ReTReK with the retrosynthesis knowledge, that without the retrosynthesis knowledge did not find any synthetic route. The weight parameters of the retrosynthesis knowledge, w_1 , w_2 , w_3 , and w_4 , are set to 5.0, 2.0, 0.5, and 2.0, respectively. The molecule 19 is the target.

of ChemAxon—is considered. In addition, we define other retrosynthesis knowledge scores so that ReTReK can represent the chemists' way of thinking more than the current model. Further, we will prepare a user-friendly interactive interface for easy to use ReTReK, which has functions such as range sliders for adjusting each retrosynthesis knowledge and others for displaying searched synthetic routes.

4 Methods

4.1 Datasets

To create the ReTReK model, compounds obtained from Reaxys reaction records¹⁹ and ZINC 15 database⁴⁹ were used. To evaluate the performance of ReTReK, compounds obtained from the ChEMBL 27 database⁴² and the literature^{43–48}.

4.1.1 Reaction template extraction

A set of about 50 million reaction records from Reaxys¹⁹ (from 1795–2019) was used to construct a 1-step retrosynthetic reaction prediction model. The input data for the model was a target or intermediate molecule, and the model was trained to predict a suitable reaction template for the molecule. A reaction template is aimed at representing a generalized chemical reaction, and in this study, a reaction template is defined as a reactive center, orphan atoms, and their first-degree neighbors. An orphan atom is one that appears only on one side of the reaction arrow in ChemAxon⁵⁰. The reaction template extraction procedure comprises four steps (Fig. S5 for the workflow of the reaction template extraction).

In the first step, reaction records were standardized using the following steps: removing explicit hydrogen, aromatizing, and keeping the largest fragments.

In the second step, reaction records were narrowed down using three conditions: (1) a reaction is a single step, (2) a reaction comprises a product and up to three reactants, and (3) the num-

ber of heavy atoms in a product is limited to 50 or less. After this step, the number of reaction records resulted in 22,337,137.

In the third step, reaction templates were extracted from the reaction records, and sets of a product and the corresponding reaction template were retained if the reaction template occurred at least 50 times. To prevent the occurrence of two or more fragments, the reaction template was retained in the case where all atoms in the product side of the template were connected.

In the final step, the sets of a product and the corresponding reaction template were filtered by the condition that the reaction template can reversibly be applied to the product and the derived reactants. Thus, 7,589,744 sets of a product and the corresponding reaction template remained; the unique number of the reaction template was 19,633. Following the previous study⁹, a time-split strategy was employed to evaluate the neural network model performance. The sets published before 2017 were used for training, and those published in 2017 and later were used for testing.

4.1.2 Preparation of molecules for ReTReK evaluations and demonstrations

Molecules used for ReTReK evaluations were obtained from ChEMBL 27⁴² and preprocessed by the following procedures. First, molecules whose USAN years ranged from 2017 to 2019 and had chemical structure records were selected, thereby resulting in a total of 219 compounds. Then, the compounds were preprocessed using the following steps: removing explicit hydrogen, aromatizing, keeping the largest fragments, removing compounds whose number of atoms is more than 50, and removing duplicates. The remaining 161 compounds were used for the evaluations (ChEMBL dataset). For further evaluation of ReTReK, six drug-like compounds^{43–48} were used as synthetic route search demonstrations.

4.1.3 Starting materials

A set of compounds obtained from the ZINC database and the Reaxys reaction records were used as starting materials. A subset of 100,023 building blocks in major suppliers (SigmaAldrich, AlfaAesar, and Acros) was obtained from the ZINC database. From the Reaxys reaction records, 649,130 compounds recorded as reactants and occurred at least five times before 2017 were used. All compounds were stored in canonical SMILES format calculated by RDKit²¹.

4.2 Monte Carlo Tree Search for retrosynthesis

Various CASP studies implemented MCTS based on the achievement of Segler *et al.*⁹. The MCTS is a search algorithm for exploring optimal solutions and comprises four steps: selection, expansion, rollout, and update⁵¹. Following Segler's implementation⁹, a state contains a set of molecules and is solved (the optimal solution) if all molecules in the state are starting materials. In this study, the retrosynthesis knowledge scores were added to the evaluation term of the selection step. The same policy network was used for both the expansion and rollout steps, similar to that in the previous research³⁵.

4.2.1 Retrosynthesis knowledge used in ReTReK

We define four types of retrosynthesis knowledge: CDScore, ASScore, RDScore, and STScore, which were inspired by previous work^{4,7,8}.

4.2.1.1 Convergent Disconnection Score The CDScore aims to favor convergent synthesis, which is known as an efficient strategy in multistep chemical synthesis. The CDScore is calculated by evaluating how equally a product is divided into reactants in a reaction $\{R_1 + R_2 + \dots + R_n \rightarrow P\}$, where R_i is a reactant, and P denotes the product.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{a(P)}{n} - a(R_i) \right| \quad (1)$$

$$\text{CDScore} = \frac{1}{1 + \text{MAE}} \quad (2)$$

Here, both $a(P)$ and $a(R_i)$ represent the number of atoms in product and reactants, and mean absolute error (MAE).

4.2.1.2 Available Substances Score For a similar purpose of the CDScore, the number of available substances generated in a reaction step is calculated as

$$\text{ASScore} = \frac{b(S)}{b(R)}. \quad (3)$$

Here, $b(S)$ and $b(R)$ represent the numbers of available substances (starting materials) and reactants, respectively.

4.2.1.3 Ring Disconnection Score A ring construction strategy is preferred if a target compound has complex ring structures because constructing ring structures in a synthetic route tends to result in simple and easily available starting materials. The RDScore is calculated by checking whether ring constructions occur in a reaction step as

$$\text{RDScore} = \begin{cases} 1 & d(P) > \sum_{i=1}^n d(R_i) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Here, $d(P)$ and $d(R_i)$ represent the number of rings in a product and reactant, respectively.

4.2.1.4 Selective Transformation Score A synthetic reaction with few by-products is preferred in terms of yield. The number of possible products from the reaction is calculated by focusing on the number of reactive centers reactants have as

$$\text{STScore} = \frac{1}{e(\sum_{i=1}^n R_i)}. \quad (5)$$

Here, $e(\sum_{i=1}^n R_i)$ represents the applicable number of patterns of products enumerated using reactants and a used reaction template.

4.2.2 Policy network

In this study, a policy network is a template-based retrosynthetic reaction prediction model, and the same model was used in the expansion and rollout steps. We employed graph convolutional networks, which is suggested as a better model for retrosynthesis in a previous study²⁸, for the retrosynthetic reaction prediction model. The model was trained using the dataset prepared in the

reaction template extraction section and comprised three graph convolutional layers with the Leaky ReLU activation and dropout ratio of 0.3, a graph dense layer with the Leaky ReLU activation, a graph gather layer with hyperbolic tangent activation, and a dense layer with softmax activation. To confirm the effectiveness of the expansion sizes on the MCTS performance, top- n accuracies were calculated in the range from 1 to 1000. To implement this model, a graph-based deep learning framework kGCN⁵² was used.

4.2.3 Selection

Starting from the root node, a tree policy is recursively used to select the next action, which implies the simulation descends the search tree step-by-step until the unvisited node that has a non-terminal state is reached. The tree policy is based on the upper confidence bound (UCB) score, and the retrosynthesis knowledge is included in the policy as follows:

$$K = \frac{1}{n}(w_1 \text{CDScore} + w_2 \text{ASScore} + w_3 \text{RDScore} + w_4 \text{STScore}), \quad (6)$$

where w_i represents weights, each weight is set to five, point five, two, and two; n denotes the number of retrosynthesis knowledge types used in a search (e.g., n is four if four retrosynthesis knowledge types are used).

$$\text{action} = \frac{Q}{N} + cP \frac{\sqrt{N-1}}{1+N} + K \quad (7)$$

Here, Q denotes an action value calculated in the update step, both N and $N-1$ are visit counts of the child and parent nodes, respectively, c denotes a constant value that is set to 10, P denotes the softmax probability obtained from the policy network, and K represents the mean of the retrosynthesis knowledge scores.

4.2.4 Expansion

Child nodes whose states are selected by the policy network are added to the node selected by the tree policy. Based on the top- n accuracy in the policy network, in independent trials, the top-50, 100, 300, and 500 reaction templates were selected, and they were filtered under the condition that the reaction templates were successfully applied to the molecule in the states and then added to the node selected by the tree policy.

4.2.5 Rollout

A simulation is implemented with the policy network if the state of the node is not proven or terminal⁹. During the simulation, the following steps are recursively implemented for the maximum of five times: An unresolved molecule (not included in the starting materials) of the state is randomly sampled, top-10 reaction templates of the molecule are obtained by the policy network, and the random-sampled reaction template is applied to the molecule. At the end of each step, the state is checked as proven or not.

A reward function r returns one of three values as reward z , which depends on the simulation result. Before starting the simulation, the reward is 10 if the state is proven and -1 if the state is terminated. After the simulation, the reward is the ratio of the number of resolved molecules in the state to the number of

molecules.

4.2.6 Update

The reward obtained from the rollout is backpropagated through the selected nodes to update their action values Q . The value of Q is based on previous research⁹ and is defined as

$$W = \max\left(0, \frac{L_{\max} - L + \sum_{i=1}^n P_i}{L_{\max}}\right), \quad (8)$$

$$Q = zW, \quad (9)$$

where $L_{\max}$ denotes the maximal branch length and is set to 10, L denotes the current branch length, and $\sum_{i=1}^n P_i$ denotes the sum of the softmax probabilities of the reaction templates in the selected nodes.

4.3 Evaluating the effects of expansion sizes and retrosynthesis knowledge on MCTS solving performances

To investigate the effect of the expansion size on the MCTS performance to solve target molecules, both the number of solved molecules in the ChEMBL dataset and times to solve the molecules were compared in terms of expansion sizes and six retrosynthesis knowledge patterns. The expansion sizes were 50, 100, 300, and 500, which was determined by the policy network's top- n accuracies. The six patterns are as follows: no retrosynthesis knowledge (no knowledge), CDScore, ASScore, RDScore, STScore, and the four retrosynthesis knowledge (all knowledge). In these experiments, the maximum number of iterations was set to 500; the score weights of CDScore, ASScore, RDScore, and STScore were fixed to 5.0, 2.0, 0.5, and 2.0, respectively.

4.4 Evaluating the effect of retrosynthesis knowledge on the search directions of MCTS

To quantify the effect of the retrosynthesis knowledge on the search directions of MCTS, we defined a route score as the average of the corresponding retrosynthesis knowledge scores in each step of a solved synthetic route. We calculated four types of scores (rCDScore, rASScore, rRDScore, and rSTScore) from the solved synthetic routes in terms of the corresponding retrosynthesis knowledge patterns. For easier comparison, each route score of the five retrosynthesis knowledge patterns was standardized based on the corresponding mean and standard deviation of the no knowledge patterns. In these experiments, solved synthetic routes under the condition of the expansion size of 500 were used. The maximum number of iterations was set to 500; the score weights of CDScore, ASScore, RDScore, and STScore were fixed to 5.0, 2.0, 0.5, and 2.0, respectively.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We would like to thank Dr. Nobuo Cho and Ms. Noriko Saito for their valuable discussions and comments from the perspective of an expert synthetic organic chemist. This research was supported by the New Energy and Industrial Technology Development Or-

ganization (NEDO), MEXT, as part of the “Program for Promoting Researches on the Supercomputer Fugak” (MD-driven Precision Medicine).

Notes and references

- 1 E. J. Corey and W. T. Wipke, *Science*, 1969, **166**, 178–192.
- 2 E. J. Corey, W. T. Wipke, R. D. Cramer and W. J. Howe, *Journal of the American Chemical Society*, 1972, **94**, 421–430.
- 3 W. Wipke, G. I. Ouchi and S. Krishnan, *Artificial Intelligence*, 1978, **11**, 173–193.
- 4 J. B. Hendrickson and A. G. Toczko, *Journal of Chemical Information and Modeling*, 1989, **29**, 137–145.
- 5 W.-D. Ihlenfeldt and J. Gasteiger, *Angewandte Chemie International Edition in English*, 1996, **34**, 2613–2633.
- 6 K. Satoh and K. Funatsu, *Journal of Chemical Information and Computer Sciences*, 1999, **39**, 316–325.
- 7 J. Law, Z. Zsoldos, A. Simon, D. Reid, Y. Liu, S. Y. Khew, A. P. Johnson, S. Major, R. A. Wade and H. Y. Ando, *Journal of Chemical Information and Modeling*, 2009, **49**, 593–602.
- 8 S. Szymkuć, E. P. Gajewska, T. Klucznik, K. Molga, P. Dittwald, M. Startek, M. Bajczyk and B. A. Grzybowski, *Angewandte Chemie - International Edition*, 2016, **55**, 5904–5937.
- 9 M. H. Segler, M. Preuss and M. P. Waller, *Nature*, 2018, **555**, 604–610.
- 10 E. J. Corey, *Angewandte Chemie International Edition in English*, 1991, **30**, 455–465.
- 11 S. D. Satyanarayanajois and R. A. Hill, *Future Medicinal Chemistry*, 2011, **3**, 1765–1786.
- 12 E. Corey, A. Long and S. Rubenstein, *Science*, 1985, **228**, 408–418.
- 13 T. Klucznik, B. Mikulak-Klucznik, M. P. McCormack, H. Lima, S. Szymkuć, M. Bhowmick, K. Molga, Y. Zhou, L. Ricker-shauser, E. P. Gajewska, A. Toutchkine, P. Dittwald, M. P. Startek, G. J. Kirkovits, R. Roszak, A. Adamski, B. Sieredzińska, M. Mrksich, S. L. Trice and B. A. Grzybowski, *Chem*, 2018, **4**, 522–532.
- 14 B. Mikulak-Klucznik, P. Gołębiowska, A. A. Bayly, O. Popik, T. Klucznik, S. Szymkuć, E. P. Gajewska, P. Dittwald, O. Staszewska-Krajewska, W. Beker, T. Badowski, K. A. Scheidt, K. Molga, J. Młynarski, M. Mrksich and B. A. Grzybowski, *Nature*, 2020.
- 15 M. A. Kayala, C.-A. Azencott, J. H. Chen and P. Baldi, *Journal of Chemical Information and Modeling*, 2011, **51**, 2209–2222.
- 16 C. W. Coley, W. H. Green and K. F. Jensen, *Accounts of Chemical Research*, 2018, **51**, 1281–1289.
- 17 Y. Lecun, Y. Bengio and G. Hinton, *Nature*, 2015, **521**, 436–444.
- 18 T. Ching, D. S. Himmelstein, B. K. Beaulieu-Jones, A. A. Kalinin, B. T. Do, G. P. Way, E. Ferrero, P.-M. Agapow, M. Zietz, M. M. Hoffman, W. Xie, G. L. Rosen, B. J. Lengerich, J. Israeli, J. Lanchantin, S. Woloszynek, A. E. Carpenter, A. Shrikumar, J. Xu, E. M. Cofer, C. A. Lavender, S. C. Turaga, A. M. Alexandari, Z. Lu, D. J. Harris, D. DeCaprio, Y. Qi, A. Kundaje, Y. Peng, L. K. Wiley, M. H. S. Segler, S. M. Boca, S. J. Swamidass, A. Huang, A. Gitter and C. S. Greene, *Journal of The Royal Society Interface*, 2018, **15**, 20170387.
- 19 Reaxys, <https://www.reaxys.com/>, 2020.
- 20 D. Lowe, *Chemical Reactions from US Patents (1976–Sep2016)*, 2017, https://figshare.com/articles/Chemical_reactions_from_US_patents_1976-Sep2016_/5104873.
- 21 G. Landrum, *RDKit: Open-Source Cheminformatics Software*, 2016, <https://github.com/rdkit/rdkit>.
- 22 C. W. Coley, W. H. Green and K. F. Jensen, *Journal of Chemical Information and Modeling*, 2019, **59**, 2529–2537.
- 23 C. Coley, *connorcoley/ASKCOS: First public release of ASKCOS*, 2019, <https://doi.org/10.5281/zenodo.3261361>.
- 24 S. Genheden, A. Thakkar, V. Chadimová, J.-L. Reymond, O. Engkvist and E. Bjerrum, *Journal of Cheminformatics*, 2020, **12**,.
- 25 M. H. S. Segler and M. P. Waller, *Chemistry - A European Journal*, 2017, **23**, 5966–5971.
- 26 P. Schwaller, T. Laino, T. Gaudin, P. Bolgar, C. A. Hunter, C. Bekas and A. A. Lee, *ACS Central Science*, 2019, **5**, 1572–1583.
- 27 P. Karpov, G. Godin and I. V. Tetko, *Artificial Neural Networks and Machine Learning – ICANN 2019: Workshop and Special Sessions*, Springer International Publishing, 2019, pp. 817–830.
- 28 S. Ishida, K. Terayama, R. Kojima, K. Takasu and Y. Okuno, *Journal of Chemical Information and Modeling*, 2019, **59**, 5026–5033.
- 29 J. S. Schreck, C. W. Coley and K. J. M. Bishop, *ACS Central Science*, 2019, **5**, 970–981.
- 30 P. Schwaller, R. Petraglia, V. Zullo, V. H. Nair, R. A. Haeuselmann, R. Pisoni, C. Bekas, A. Iuliano and T. Laino, *Chemical Science*, 2020, **11**, 3316–3325.
- 31 R. Shibukawa, S. Ishida, K. Yoshizoe, K. Wasa, K. Takasu, Y. Okuno, K. Terayama and K. Tsuda, *Journal of Cheminformatics*, 2020, **12**, 52.
- 32 J. Bradshaw, B. Paige, M. Kusner, M. Segler and J. M. Hernández-Lobato, *Advances in Neural Information Processing Systems* 33, 2020.
- 33 I. V. Tetko, P. Karpov, R. V. Deursen and G. Godin, *Nature Communications*, 2020, **11**, 5575.
- 34 C. W. Coley, W. Jin, L. Rogers, T. F. Jamison, T. S. Jaakkola, W. H. Green, R. Barzilay and K. F. Jensen, *Chemical Science*, 2019, **10**, 370–377.
- 35 A. Thakkar, T. Kogej, J.-L. Reymond, O. Engkvist and E. J. Bjerrum, *Chem. Sci.*, 2019, –.
- 36 X. Wang, Y. Qian, H. Gao, C. W. Coley, Y. Mo, R. Barzilay and K. F. Jensen, *Chemical Science*, 2020, **11**, 10959–10972.
- 37 A. Kishimoto, B. Buesser, B. Chen and A. Botea, *Advances in Neural Information Processing Systems*, 2019, pp. 7224–7234.
- 38 C. W. Coley, D. A. Thomas, J. A. Lummiss, J. N. Jaworski, C. P. Breen, V. Schultz, T. Hart, J. S. Fishman, L. Rogers, H. Gao, R. W. Hicklin, P. P. Plehiers, J. Byington, J. S. Piotti, W. H. Green, A. J. Hart, T. F. Jamison and K. F. Jensen, *Science*,

- 2019, **365**, 6453.
- 39 A. F. de Almeida, R. Moreira and T. Rodrigues, *Nature Reviews Chemistry*, 2019, **3**, 589–604.
 - 40 F. Strieth-Kalthoff, F. Sandfort, M. H. S. Segler and F. Glorius, *Chemical Society Reviews*, 2020, **49**, 6154–6168.
 - 41 P. M. Pflüger and F. Glorius, *Angewandte Chemie International Edition*, 2020, **59**, 18860–18865.
 - 42 D. Mendez, A. Gaulton, A. P. cia Bento, J. Chambers, M. D. Veij, E. F. lix, M. a Paula Magariños, J. F. Mosquera, P. Mutowo, M. Nowotka, M. a Gordillo-Marañó n, F. Hunter, L. Junco, G. Mugumbate, M. Rodriguez-Lopez, F. Atkinson, N. Bosc, C. J. Radoux, A. Segura-Cabrera, A. Hersey and A. R. Leach, *Nucleic Acids Research*, 2019, **47**, D930–D940.
 - 43 W. Chen, F. Liu, Q. Zhao, X. Ma, D. Lu, H. Li, Y. Zeng, X. Tong, L. Zeng, J. Liu, L. Yang, J. Zuo and Y. Hu, *Journal of Medicinal Chemistry*, 2020, **63**, 8134–8145.
 - 44 Y. Jian, R. Merceron, S. D. Munck, H. E. Forbes, F. Hulpia, M. D. Risseuw, K. V. Hecke, S. N. Savvides, H. Munier-Lehmann, H. Boshoff and S. V. Calenbergh, *European Journal of Medicinal Chemistry*, 2020, **206**, 112659.
 - 45 G. Tsuji, M. Yusa, S. Masada, H. Yokoo, J. Hosoe, T. Hakamatsuka, Y. Demizu and N. Uchiyama, *Chemical and Pharmaceutical Bulletin*, 2020, **68**, 797–801.
 - 46 M. C. Pismataro, N. A. Horenstein, C. Stokes, M. Quadri, M. D. Amici, R. L. Papke and C. Dallanocce, *European Journal of Medicinal Chemistry*, 2020, **205**, 112669.
 - 47 T. P. Banzato, J. R. Gubiani, D. I. Bernardi, C. R. Nogueira, A. F. Monteiro, F. F. Juliano, S. M. de Alencar, R. A. Pilli, C. A. de Lima, G. B. Longato, A. G. Ferreira, M. A. Foglio, J. E. de Carvalho, D. B. Vendramini-Costa and R. G. S. Berlinck, *Journal of Natural Products*, 2020, **83**, 1784–1793.
 - 48 Z. Su, T. Yang, J. Wang, M. Lai, L. Tong, G. Wumaier, Z. Chen, S. Li, H. Li, H. Xie and Z. Zhao, *Bioorganic & Medicinal Chemistry Letters*, 2020, **30**, 127327.
 - 49 T. Sterling and J. J. Irwin, *Journal of Chemical Information and Modeling*, 2015, **55**, 2324–2337.
 - 50 ChemAxon, <http://www.chemaxon.com>, 2020.
 - 51 C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis and S. Colton, *IEEE Transactions on Computational Intelligence and AI in Games*, 2012, **4**, 1–43.
 - 52 R. Kojima, S. Ishida, M. Ohta, H. Iwata, T. Honma and Y. Okuno, *Journal of Cheminformatics*, 2020, **12**, 32.