Supplementary Material:
Self-Improved Retrosynthetic Planning

A. More Discussion on Reaction Pathways

In this section, we provide additional reaction pathways generated by our framework and baselines. In Figure 1, given the target molecule, RETRO*-0 + OURS successfully finds the reaction pathway while RETRO*-0 cannot find anyone. In Figure 2, given another target molecule, both RETRO*-0 + OURS and RETRO*-0 successfully find reaction pathways. However, RETRO*-0 + OURS finds a much shorter reaction pathway, which is preferable in laboratories. As our backward reaction model is trained to consider executability from building blocks as well as realistic-ness, our framework is able to search shorter reaction pathways.

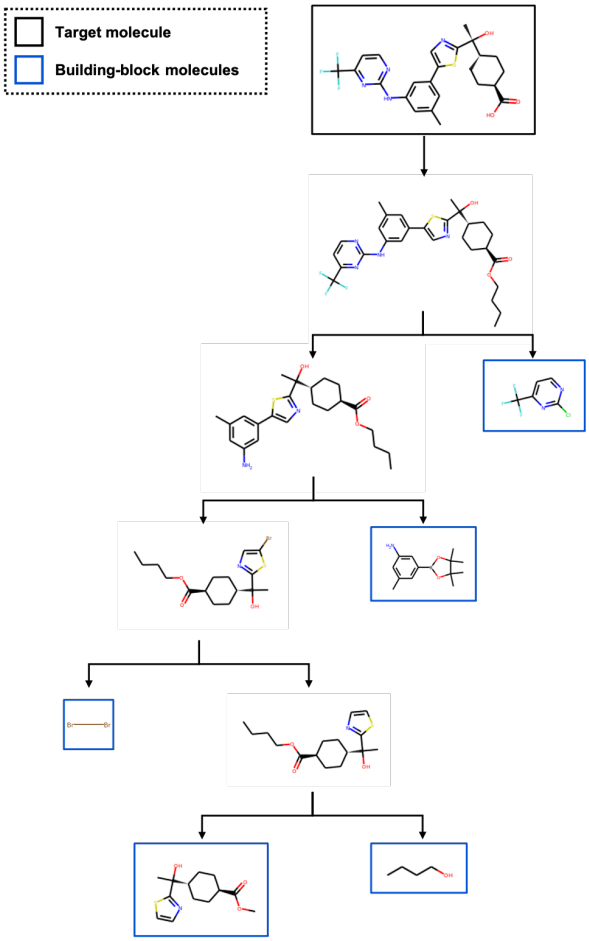
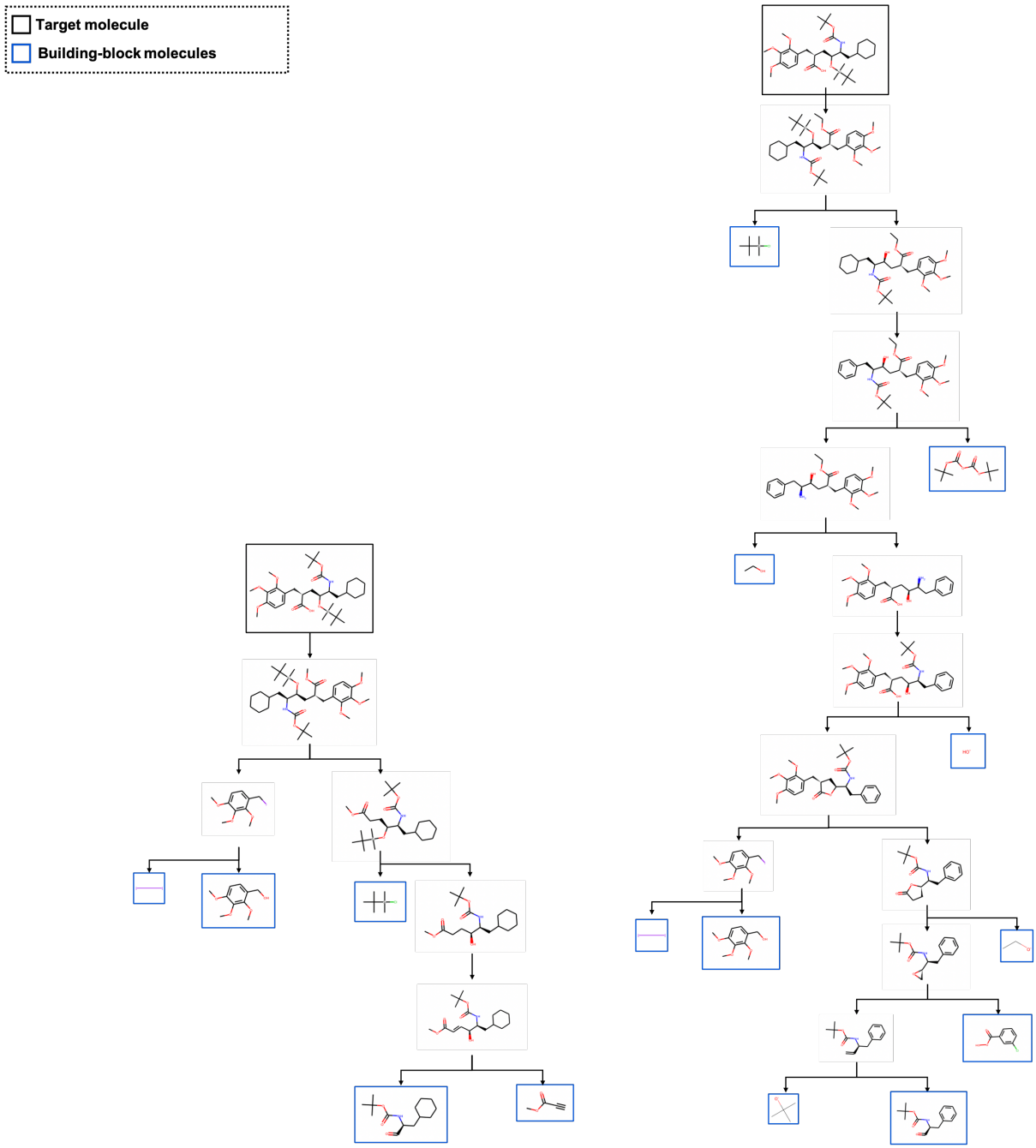


Figure 1. Reaction pathway produced by RETRO*-0 + OURS, given the target molecule trans-4-(1S)-1-Hydroxy-1-[5-(3-methyl-5-[4-(trifluoromethyl)pyrimidin-2-yl]aminophenyl)-1,3-thiazol-2-yl]ethylcyclohexanecarboxylic acid, where RETRO*-0 failed to find corresponding reaction pathway.

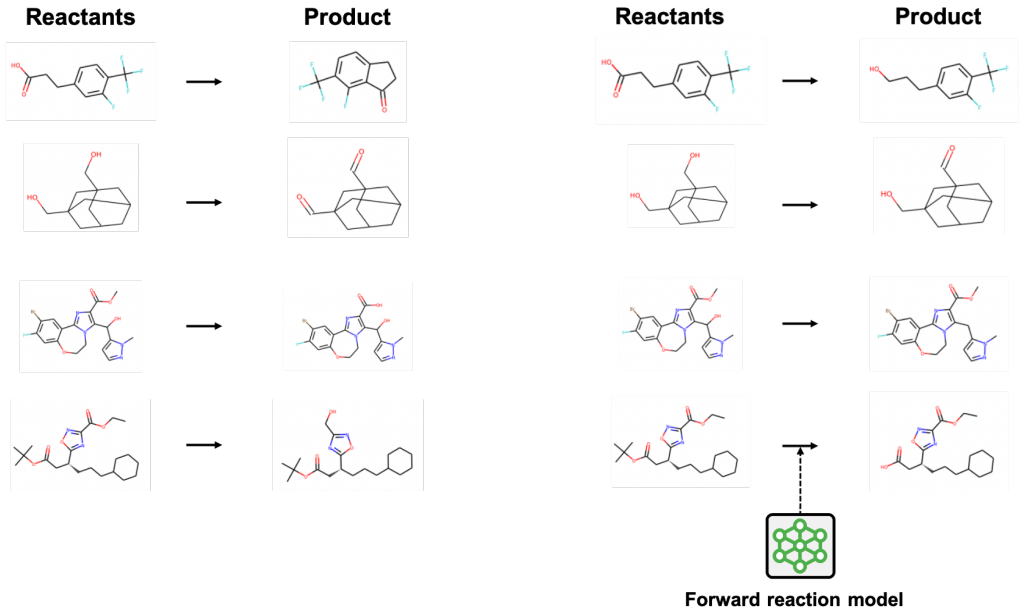


(a) Reaction pathway searched by RETRO*-0 + OURS

(b) Reaction pathway searched by RETRO*-0

Figure 2. Reaction pathways from the same target molecule, 5(S)-(Boc-Amino)-4(S)-(tert-butyldimethylsilyloxy)-6-cyclohexyl-2(R)-[(2,3,4-trimethoxyphenyl)-methyl]hexanoic acid, searched by (a) RETRO*-0 + OURS and (b) RETRO*-0. Our framework searches shorter reaction pathway, which is preferable in laboratories.

B. More Examples of Reaction Augmentation



(a) Extracted reactions from reaction pathways

(b) Augmented reactions

Figure 3. Illustration of reactions (a) extracted from reaction pathways found by search algorithm combined with the backward reaction model and (b) corresponding augmented reactions via the forward reaction model. The products of augmented reactions have similar but different structures from the original products. Using augmentation, we can improve the generalization ability of the backward reaction model.

C. Additional experiments

C.1. Trade-off of search time and performance

We compare algorithms using a sufficient search time to converge, i.e., 5000 model calls, in Figure 4. Although success rates could converge if infinite search time is given, the figure highlights that (1) the gap does not converge to zero (+2.11 %) and (2) our method can discover successful routes much efficiently. We would like to emphasize that (2) is critical in practice because the synthesis of some complex substances, i.e., vitamin B12 requires more than 100 reactions (**?**), which could increase the search space exponentially.

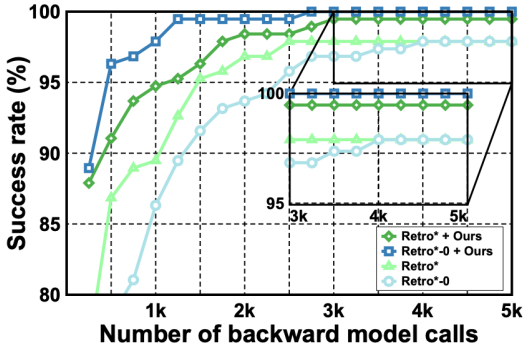


Figure 4. Success rate (%) under sufficient search time to converge, i.e., 5000 model calls. Our framework outperforms the best baselines, RETRO*-0, even if a sufficient search time is given.

C.2. Comparison to another augmentation scheme

We compare our augmentation scheme to the "mixed forward/reverse augmentation" (Mix) (**?**). [1] The Mix augmentation augment reactions by switching products and reactants to improve generalization ability via mixed representation of latent space. As shown in Table 1, the Mix augmentation performs slightly better when using $N = 50$ single-step model calls but much worse for $N = 250$ and $N = 500$.

Aug	$N = 50$	$N = 250$	$N = 500$
None	46.84	81.05	92.63
Mix	**52.63**	66.32	72.63
Ours	50.00	**83.16**	**93.16**

Table 1. Experimental comparison between our augmentation scheme to the "mixed forward/reverse augmentation" (Mix) (**?**). The Mix augmentation performs slightly better when using $N = 50$ single-step model calls but much worse for $N = 250$ and $N = 500$.

[1] Note that SMILES-based augmentations, also proposed in (**?**), are not directly applicable to our implementation since our model is based on Morgan fingerprint which is invariant to the SMILES-based augmentations.