Appendix A Experimental procedures

A.1 Model implementation

The encoder and the scorer model are implemented using Pytorch. The encoder contains 2 linear layers (including the output layer) with a ReLU activation function. As mentioned before, the post-addition transformation is the identity transformation. The hidden layer had 64 neurons, and the output size was 256. The information of each reaction in the route is summed together before returning. The prediction model was a Multi Layer Perceptron model with 3 linear layers with ReLU activation function, the layers contain 1024, 1024, and 1 neuron, respectively. The inputs of this linear model consist of two parts, the output of the encoder and the 5 properties of the route mentioned in the Architecture section. Prior to being forwarded to the prediction model, the inputs are combined such that for each molecule the encoder outputs are concatenated with the properties. The models' details in each experiment are:

- **No embedding experiment** The encoder has **5 fully connected layers** including the output layer, with the number of neurons each layer are **64, 128, 512, 512, and 256**, respectively. The activation function used is **ReLU**. The scorer has **6 fully connected layers** including the output layer, with the number of neurons each layer are **512, 1024, 2048, 1024, 512, and 1**, respectively. The activation function for the network is **ReLU**. In both networks, the activation funciton is applied on all layers except for the output layer.
- **SDF embedding experiment** The encoder has **5 fully connected layers** including the output layer, with the number of neurons each layer are **64, 128, 512, 512, and 256**, respectively. The activation function used is **ReLU**. The scorer has **6 fully connected layers** including the output layer, with the number of neurons each layer are **512, 1024, 2048, 1024, 512, and 1**, respectively. The activation function for the network is **ReLU**. In both networks, the activation funciton is applied on all layers except for the output layer.
- **RXNFP embedding experiment** The **encoder** DNN has **5 fully connected layers** including the output layer, with the number of neurons each layer are **64, 128, 512, 512, and 256**, respectively. The activation function used is **ReLU**. The scorer has **5 fully connected layers** including the output layer, with the number of neurons each layer are **512, 1024, 512, 256, and 1**, respectively. The activation function for the network is **ReLU**. In both networks, the activation funciton is applied on all layers except for the output layer.
- **DRFP embedding experiment** The **encoder** DNN has **5 fully connected layers** including the output layer, with the number of neurons each layer are **64, 128, 512, 512, and 512**, respectively. The activation function used is **ReLU**. The scorer has **6 fully connected layers** including the output layer, with the number of neurons each layer are **512, 1024, 2048, 1024, 512, and 1**, respectively. The activation function for the network is **ReLU**. In both networks, the activation funciton is applied on all layers except for the output layer.

A.2 Training details

A.3 Training Details

In all experiments, the dataset is randomly split into a **training set** and a **test set** using a ratio of **7:1**. To examine the impact of data splitting on model performance, we *repeated the split three times*, using different random seeds. Additionally, **20% of the training set** is reserved for validation during model training.

Each experiment is *repeated three times*, using different training and test splits. To optimize computational efficiency, we *precomputed the feasibility input and embeddings* of the routes before training. These inputs are then fed into the **encoder**, and the encoder's outputs are concatenated with the **route properties** before being passed to the **prediction model**. The training configurations are:

- **Loss function**: Mean Squared Error (MSE)
- **Total training epochs**: 81
- **Batch sizes**: 256 for No Embedding, SDF, and RXNFP experiments. 512 for the DRFP experiment
- **Learning rate**: 0.001
- **Hardware**: 1 NVIDIA V100 GPU
- **Evaluation frequency**: Every 10 iterations, the model is tested on the test set to monitor training progress.

Appendix B Complete analysis of route scores

Table 1: Complete route analysis

Route ID	Top-1 Exp.	Suggested Score	Expert Assessment	Reason (Selected Criteria)	Analogous Key Steps Reported in Literature
1	FALSE	3.4	good	Forward synthesis likely to proceed	The key step precedented: WO2020/48949, 2020, A1
2	TRUE	4.8	plausible	Likely introducing modification step (reactant)	
3	TRUE	9	plausible	Likely introducing modification step (FG compatibility)	The key second step precedented: Indian Journal of Heterocyclic Chemistry, 2015, vol. 24, #3, p. 265-270
4	FALSE	9.8	plausible	Likely introducing modification step (reactant)	The modified key step precedented: J. Med. Chem. 2012, 55, 14, 6608–6623
5	TRUE	10.5	bad	Route not solved	
6	TRUE	5.5	plausible	Likely introducing modification (avoid FG interconversion)	Without unnecessary 2-step hydrolysis and esterification, the route is similar to the ground truth
7	TRUE	12.7	bad	Route not solved	
8	FALSE	5	bad	Unlikely disconnection	No precedence found for the suggested or modified forward synthesis steps
9	TRUE	7	bad	Route not solved	
10	FALSE	5.9	plausible	Likely introducing modification (reactant)	Modified key steps precedented: TANSNA THERAPEUTICS - US2015/148430, 2015, A1; lithiation of aryl fluorides: Tetrahedron, 2000, vol. 56, #8, p. 1135-1138
11	TRUE	1.1	good	Forward synthesis likely to proceed	The key step precedented: Chemistry - A European Journal, 2015, vol. 21, #1, p. 111-114
12	TRUE	1.2	good	Forward synthesis likely to proceed	The key first step precedented: EA PHARMA - US2015/51395, 2015, A1
13	TRUE	11.6	bad	Route not solved	
14	FALSE	3	good	Forward synthesis likely to proceed	The key first step precedented: Langmuir, 2015, vol. 31, #49, p. 13410-13419; The key second step precedented: Journal of Organic Chemistry, 2022, vol. 87, #5, p. 2315-2323
15	TRUE	10.8	bad	Route not solved	
16	FALSE	3.9	good	Forward synthesis likely to proceed	The key step precedented: Bioorg. Med. Chem. Letters, 2001, 11, 4, 515-517 and Journal of Medicinal Chemistry, 2008, vol. 51, #16, p. 4870-4873

Table 2: Complete route analysis

Route ID	Top-1 Exp.	Suggested Score	Expert Assessment	Reason (Selected Criteria)	Analogous Key Steps Reported in Literature
17	FALSE	1.5	good	Forward synthesis likely to proceed	The key step precedented: 1st step: Organic Letters, 2013, vol. 15, #11, p. 2770-2773; 2nd step: Tetrahedron, 2005, vol. 61, #50, p. 11965-11975
18	FALSE	7.8	plausible	Likely introducing modification (avoid FG interconversion)	The modified key steps precedented: US2003/144303, 2003, A1; Chemistry of Heterocyclic Compounds, 1985, vol. 21, #9, p. 1023-1025
19	FALSE	3.8	good	Forward synthesis likely to proceed	The key step precedented: DOI: 10.1021/acs.joc.1c03135
20	TRUE	10.5	bad	Route not solved	
21	FALSE	4.4	good	Forward synthesis likely to proceed	The key 2nd step is precedented: Acta Tropica, 2022, vol. 234, art. no. 106607; WO2022/36204, 2022, A1, #rd step: US2019/55225, 2019, A1; WO2022/207482, 2022, A1
22	TRUE	14.6	bad	Route not solved	
23	FALSE	4.8	good	Forward synthesis likely to proceed	The key step precedented (both acids unprotected): WO2014/82379, 2014, A1
24	TRUE	4.2	good	Forward synthesis likely to proceed	Route is similar to the ground truth
25	TRUE	10.7	plausible	Likely introducing modification step (PG strategy)	Guanidinium protection needs to be either removed or changed; modified steps precedented: first step bottom: compatible with alcohol: Tetrahedron Letters, 2017, vol. 58, # 24, p. 2318 - 2321; reductive amination in the presence of guanidinium: Letters in Drug Design and Discovery, 2024, vol. 21, # 4, p. 738 - 748. Cbz deprotection of guanidine: Journal of Organic Chemistry, 1998, vol. 63, # 16, p. 5648 - 5655
26	TRUE	7.6	plausible	Likely introducing modification (reactant)	The steps are precedented for the modified reactant only: alpha bromoketone: Tetrahedron Letters, 2012, vol. 53, # 32, p. 4117 - 4120; cyclisation: European Journal of Medicinal Chemistry, 2023, vol. 259, art. no. 115637; Org. Biomol. Chem., 2012, 10, 1093-1101; CN formation WO2022/184111, 2022, A1
27	TRUE	4.7	good	Forward synthesis likely to proceed	Route is similar to the ground truth
28	TRUE	7.6	bad	Route not solved	
29	FALSE	3.1	good	Forward synthesis likely to proceed	Route is similar to the ground truth

Table 3: Complete route analysis

Route ID	Top-1 Exp.	Suggested Score	Expert Assessment	Reason (Selected Criteria)	Analogous Key Steps Reported in Literature
30	FALSE	6.8	good	Forward synthesis likely to proceed	Two key steps are precedented in the literature: Nitro reduction in presence of Cbz: Chemical Communications, 2011, vol. 47, # 12, p. 3601 - 3603, urea: SANOFI - WO2012/4732, 2012, A1
31	TRUE	11.6	bad	Route not solved	
32	TRUE	6.4	plausible	Likely introducing modification (avoid FG interconversion)	Without the unnecessary step, the synthesis is precedented: first step: Bioorganic and Medicinal Chemistry Letters, 2011, vol. 21,#2, p. 849 - 852; acyl fluoride to chloride: WO2012/7142, 2012, A1; reaction of acyl fluoride with amine: US9244345, 2016, B1
33	FALSE	4.6	good	Forward synthesis likely to proceed	The key step is precedented (ester formation): ACS Medicinal Chemistry Letters, 2018, vol. 9,#11, p. 1082 - 1087
34	TRUE	8.3	plausible	Likely introducing modification (reactant)	With a modified reactant (DCM to be replaced), synthesis is precedented: deoxyfluorination: Journal of Organic Chemistry, 2022, vol. 87,#9, p. 6471 - 6478; CH functionalisation: Journal of Chemical Research, 2019, vol. 43,#1-2, p. 34 - 38; Bioorganic and Medicinal Chemistry Letters, 2013, vol. 23,#14, p. 4162 - 4165; WO2008/122765, 2008, A1; N alkylation: Journal of Medicinal Chemistry, 1991, vol. 34,#9, p. 2843 - 2852
35	TRUE	6.2	plausible	Likely introducing modification (reactant)	With the modified reactant (primary amide is to be replaced), synthesis is precedented: WO2011/38204, 2011, A1; EP1447402, 2004, A1
36	FALSE	1.0	good	Forward synthesis likely to proceed	Route is very similar to the ground truth
37	TRUE	4.7	good	Forward synthesis likely to proceed	Precedence for the key steps: coupling CN ChanLam: Advanced Synthesis and Catalysis, 2010, vol. 352,#1, p. 81 - 84; Journal of Medicinal Chemistry, 2018, vol. 61,#5, p. 1785 - 1799; with pinacol borate: FL2022 001 - WO2022/72512, 2022, A1; SNAr on fluoride ok: WO2023/154519, 2023, A1; chlorination: Chemistry - A European Journal, 2020, vol. 26, # 46, p. 10411 - 10416; use of NBS for halogenation: Synthesis, 2009,#8, art. no. P12608SS, p. 1305 - 1308

Table 4: Complete route analysis

Route ID	Top-1 Exp.	Suggested Score	Expert Assessment	Reason (Selected Criteria)	Analogous Key Steps Reported in Literature
38	TRUE	4.7	good	Forward synthesis likely to proceed	The precedence for the key steps: alkylation: Organic Letters, 2021, vol. 23,#1, p. 166 - 171; cyclisation: WO2022/197756, 2022, A1
39	TRUE	12.1	bad	Route not solved	
40	TRUE	12.7	bad	Route not solved	
41	TRUE	12.1	bad	Route not solved	
42	FALSE	5.0	good	Forward synthesis likely to proceed	The key step is precedented: Ketone reduction: Journal of Medicinal Chemistry, 2000, vol. 43,#7, p. 1293 - 1310
43	TRUE	4.1	good	Forward synthesis likely to proceed	First step: European Journal of Medicinal Chemistry, 2013, vol. 62, p. 256 - 269; cyclisation: Bioorganic and Medicinal Chemistry Letters, 2019, vol. 29, # 21, art. no. 126644; last step: Angewandte Chemie - International Edition, 2022, vol. 61,#46, art. no. E202209232
44	TRUE	2.2	plausible	Likely introducing modification step (PG strategy)	First step: Journal of Medicinal Chemistry, 2005, vol. 48,#15, p. 4990 - 5000; dehalogenation: Bioorganic and Medicinal Chemistry Letters, 2008, vol. 18,#20, p. 5653 - 5656
45	TRUE	9.3	good	Forward synthesis likely to proceed	beta keto ester: Angewandte Chemie - International Edition, 2003, vol. 42, #24, p. 2785 - 2788; cyclisation: Journal of Medicinal Chemistry, 2014, vol. 57, #6, p. 2429 - 2439; phenol activation: WO2020103897 A1, 2020-05-28; cross coupling: ACS Medicinal Chemistry Letters, 2012, vol. 3, #9, p. 726 - 730
46	FALSE	5.2	good	Forward synthesis likely to proceed	Amide alkylation: WO2020/48949, 2020, A1
47	FALSE	2.7	good	Forward synthesis likely to proceed	Nitro reduction: WO2023/56264, 2023, A1
48	FALSE	4.5	good	Forward synthesis likely to proceed	Route is similar to the ground truth
49	FALSE	2.7	plausible	Likely introducing modification (avoid FG interconversion)	Without the unnecessary 1st step, route is likely to proceed
50	TRUE	10.9	good	Forward synthesis likely to proceed	First step SNAr could solve regioselectivity: Journal of Medicinal Chemistry, 2005, vol. 48, #6, p. 2072 - 2079; N alkylation: SHANGHAI INSTITUTE OF MATERIA MEDICA CAS - CN113402520, 2021, A; last step: GENENTECH - US2012/214762, 2012, A1