

Supplementary Materials for

A robotic platform for flow synthesis of organic compounds informed by AI planning

Connor W. Coley*, Dale A. Thomas III*, Justin A. M. Lummiss*, Jonathan N. Jaworski, Christopher P. Breen, Victor Schultz, Travis Hart, Joshua S. Fishman, Luke Rogers, Hanyu Gao, Robert W. Hicklin, Pieter P. Plehiers, Joshua Byington, John S. Piotti, William H. Green, A. John Hart, Timothy F. Jamison†, Klavs F. Jensen†

*These authors contributed equally to this work.

†Corresponding author. Email: tfj@mit.edu (T.F.J.); kfjensen@mit.edu (K.F.J.)

Published 9 August 2019, *Science* **365**, eaax1566 (2019)
DOI: 10.1126/science.aax1566

This PDF file includes:

Supplementary Text
Figures S1 to S95
Tables S1 to S3
Captions for Movies S1 to S3
Caption for Recipe Planning File S1
Caption for Recipe Planning Script S1
Captions for Recipes S1 to S7
Captions for Batch Sheets S1 to S15
Captions for Website Printouts S1 to S34
References

Other Supplementary Material for this manuscript includes the following: (available at science.sciencemag.org/content/365/6453/eaax1566/suppl/DC1)

Movies S1 to S3
Recipe Planning File S1
Recipe Planning Script S1
Recipes S1 to S7
Batch Sheets S1 to S15
Website Printouts S1 to S34

S1. Synthesis Planning Software

S1.1 Software overview

All code was written in Python 2.7 and tested/deployed on an Ubuntu 16.04 workstation. We make extensive use of the open source cheminformatics toolkit RDKit (32). For ease of use, we have developed an online webapp in parallel with a command-line interface using Django. For parallelization and scalability, we use Celery to distribute website tasks to worker pools, which can be located on distinct servers from the web service. All code and trained models used in the synthesis planning program is available at <https://github.com/connorcolev/ASKCOS> (43).

The general workflow for retrosynthetic planning is shown in **Fig. S1**. Given a target compound, we first identify a reaction that could be used to produce it. Once precursors are generated, we determine if it is feasible. To do so, we make use of both a “fast filter” model for coarse filtering and an explicit reaction prediction model. Prior to using the reaction prediction model, we explicitly propose reaction conditions that could be suitable for the intended transformation. If the predicted product matches the target we are trying to synthesize at this step, then we accept the reaction (note: this forward prediction can be run as part of the integrated workflow or as a post-processing step). The reactants of feasible reactions are evaluated to see if they match any stop criteria (e.g., are buyable). If they do not, we continue the expansion recursively or up until a maximum depth is reached.

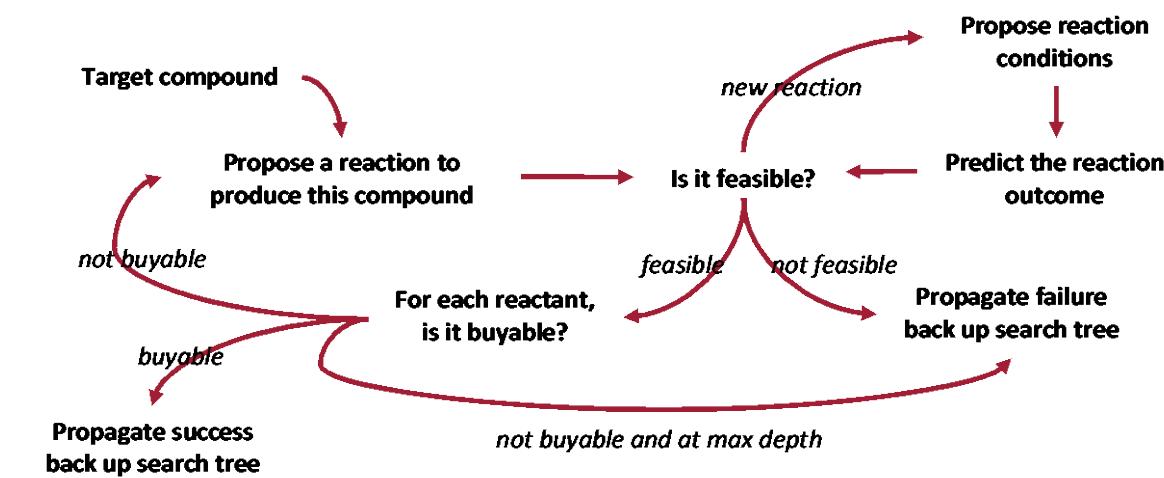


Fig. S1. Workflow for algorithmic synthesis planning. To improve the robustness of generated pathways, special focus is placed on determining reaction feasibility by explicitly proposing reaction conditions and trying to predict the outcome.

Each task involved in this process can be handled by an independent worker in a modular software suite (**Fig. S2**). The search is coordinated by a single process that maintains several active search directions and keeps track of how often each branch has been visited and how promising they are. A parallelized pool of workers (Transform Worker) asynchronously generates potential precursors for molecules according to the search direction dictated by the tree builder coordinator (Tree Builder). Transform workers can make use of different prioritization strategies, both for how templates are applied (Template Prioritizer) and how precursors are prioritized (Precursor Prioritizer). Reaction-level evaluation can occur through the use of context recommendation and

forward prediction models. The Pathway Evaluator currently ranks pathways by overall perceived plausibility but could be expanded upon to include more sophisticated pathway-level considerations (e.g., separation difficulty between reactions).

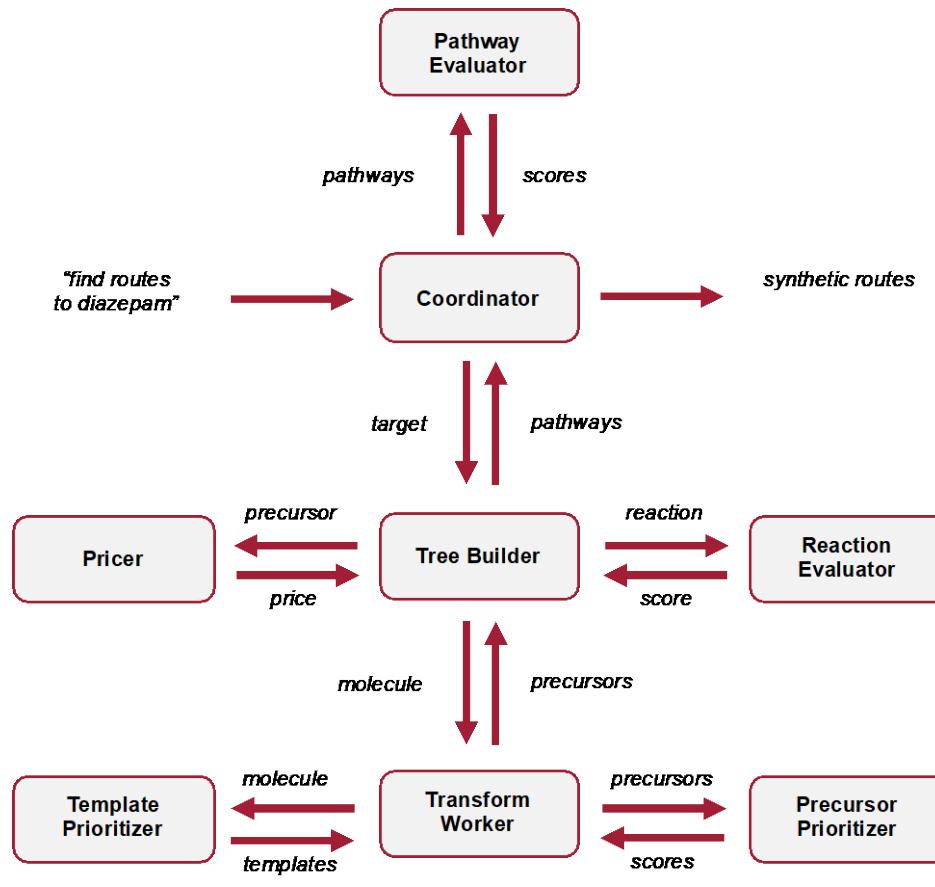


Fig. S2. Modular software design.

S1.2 Retrosynthetic template extraction

Reaction templates are subgraph matching rules that can be algorithmically extracted from literature precedent reactions and applied to new substrates to recognize structural motifs that lend themselves to retrosynthetic disconnection (46, 47). We processed 12.5 M published single-step reactions tabulated in the Reaxys database to produce a template library; for the final template set, we used the union of all rules observed 10+ times and all rules with specified stereochemistry observed 5+ times, totaling 163,723 rules. Transformations with stereochemistry are inherently more specific and are expected to appear less frequently, yet are essential to include to allow the program to predict syntheses of chiral molecules.

The extraction procedure starts by identifying the reaction center of an atom-mapped reaction: which atoms have undergone a change in connectivity from the reactants to the product(s). From there, we include all neighbors one bond away and check for an appropriate level of generalization. If the reacting atoms or their immediate neighbors belong to one of a number of “privileged” functional groups, we include that whole group in the reaction template. This includes boronic

acid/esters, azides, carbonyls, alkenes, meta-halogenated aromatic rings, and several others that are enumerated in the code. We also include all neighbors that are necessary to specify the chirality of the reacting atoms or their neighbors. If none of these conditions are met, then the neighboring atoms are generalized at the level of atomic number.

There is an inverse power law relationship between the number of reaction precedents for a template and that template's rank (in terms of descending popularity), as shown in **Fig. S3**. This means that very obscure templates will rarely apply, but are important to describe a large fraction of known chemistry. Their rarity means that they may not generalize well or may be the result of data quality issues. The top 250 transform rules are shown below.

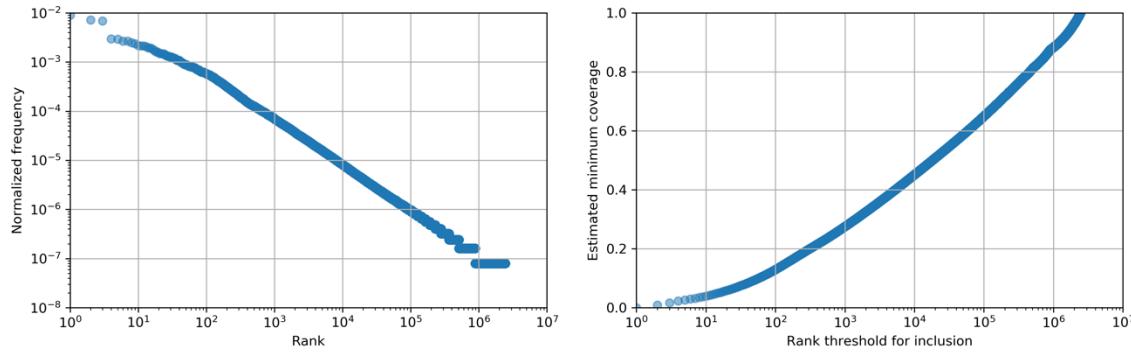


Fig. S3. (left) Relationship between the popularity of a reaction template and the template rank, showing an inverse power law; (right) an integrated form of the left panel, showing the number of templates that must be included in order to guarantee a certain coverage (i.e., ability to describe published reactions in this data set).

Table S1. The most popular 250 retrosynthetic templates algorithmically extracted from the Reaxys database, defined as SMARTS strings. While these rules cover a large number of examples, there are 163,723 total rules in our template library.

Rank	# Precedents	Reaction SMARTS
1	112035	[O;D1;H0:3]=[C:2]-[OH;D1;+0:1]>>C-[O;H0;D2;+0:1]-[C:2]=[O;D1;H0:3]
2	89920	[NH2;D1;+0:1]-[c:2]>>O=[N+;H0;D3:1](-[O-])-c:2]
3	86301	[O;D1;H0:3]=[C:2]-[OH;D1;+0:1]>>C-C-[O;H0;D2;+0:1]-[C:2]=[O;D1;H0:3]
4	37091	[C:2]-[NH;D2;+0:1]-[C:3]>>C-C(-C)(-C)-O-C(=O)-[N;H0;D3;+0:1]-[C:2]-[C:3]
5	36723	[C:1]-[CH;D2;+0:2]=[O;H0;D1;+0:3]>>[C:1]-[CH2;D2;+0:2]-[OH;D1;+0:3]
6	33506	[C:2]-[OH;D1;+0:1]>>C-C(-C)(-C)-[Si](C)(C)-[O;H0;D2;+0:1]-[C:2]
7	33139	[C:5]-[O;H0;D2;+0:6]-[S;H0;D4;+0:1]-[C;D1;H3:2](=[O;D1;H0:3])=[O;D1;H0:4]>>Cl-[S;H0;D4;+0:1](-[C;D1;H3:2])(=[O;D1;H0:3])=[O;D1;H0:4].[C:5]-[OH;D1;+0:6]
8	30997	[OH;D1;+0:1]-[c:2]>>C-[O;H0;D2;+0:1]-[c:2]
9	29623	[C:2]-[NH2;D1;+0:1]>>C-C(-C)(-C)-O-C(=O)-[NH;D2;+0:1]-[C:2]
10	26861	[C:2]-[OH;D1;+0:1]>>C-C(=O)-[O;H0;D2;+0:1]-[C:2]
11	26695	[C:2]-[C;H0;D3;+0:1](-[Cl;H0;D1;+0])=[O;H0;D1;+0:3]>>O=[C;H0;D3;+0:1](-[C:2])-OH;D1;+0:3]
12	26490	[C:5]-[O;H0;D2;+0:6]-[S;H0;D4;+0:1](=[O;D1;H0:2])(=[O;D1;H0:3])-c:4]>>Cl-[S;H0;D4;+0:1](=[O;D1;H0:2])(=[O;D1;H0:3])-c:4].[C:5]-[OH;D1;+0:6]
13	25655	[C:2]-[CH2;D2;+0:1]-[O;H0;D2;+0:3]-c:4]>>Br-[CH2;D2;+0:1]-[C:2].[OH;D1;+0:3]-c:4]
14	24348	[C:2]-[C;H0;D3;+0:1](=[O;H0;D1;+0:3])-NH;D2;+0:5]-[C:4]>>O=[C;H0;D3;+0:1](-[C:2])-OH;D1;+0:3].[C:4]-[NH2;D1;+0:5]
15	24281	[O;D1;H0:3]=[C:2]-[OH;D1;+0:1]>>C-C(-C)(-C)-[O;H0;D2;+0:1]-[C:2]=[O;D1;H0:3]
16	23078	[C:4]#[C;H0;D2;+0:5]-[c;H0;D3;+0:1]([c:2]):[c:3]>>I-[c;H0;D3;+0:1]([c:2]):[c:3].[C:4]#[CH;D1;+0:5]

17	22023	[C:4]-[NH;D2:+0:5]-[C;H0;D3:+0:1](-[C:2])=[O;H0;D1;+0:3]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[C:4]-[NH2;D1;+0:5]
18	20344	[Cl;H0;D1;+0]-[C;H0;D3;+0:1](=[O;H0;D1;+0:2])- [c:3]>>O=[C;H0;D3;+0:1](-[OH;D1;+0:2])- [c:3]
19	20318	[C:4]-[NH;D2:+0:5]-[C;H0;D3;+0:1](=[O;H0;D1;+0:2])- [c:3]>>O=[C;H0;D3;+0:1](-[OH;D1;+0:2])- [c:3].[C:4]-[NH2;D1;+0:5]
20	19130	[#7:4]-[C:3](=[O;D1;H0:5])- [C:2]-[NH2;D1;+0:1]>>C-C(-C)(-C)-O-C(=O)-[NH;D2;+0:1]- [C:2]-[C:3](-[#7:4])=[O;D1;H0:5]
21	18572	[O;D1;H0:2]=[C;H0;D3;+0:1](-[c:3])- [NH;D2;+0:4]- [c:5]>>Cl-[C;H0;D3;+0:1](=[O;D1;H0:2])- [c:3].[NH2;D1;+0:4]- [c:5]
22	18447	[C:2]-[C;H0;D3;+0:1](=[O;H0;D1;+0:3])- [NH;D2;+0:4]- [c:5]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[NH2;D1;+0:4]- [c:5]
23	18352	[C:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [NH;D2;+0:4]- [c:5]>>Cl-[C;H0;D3;+0:1](-[C:2])=[O;D1;H0:3].[NH2;D1;+0:4]- [c:5]
24	18347	[Br;H0;D1;+0]-[c;H0;D3;+0:2](:[c:1]):[c:3]>>[c:1]:[cH;D2;+0:2]:[c:3]
25	17434	[C:1]-[NH2;D1;+0:2]>>[C:1]-[N;H0;D2;+0:2]=[N+] =[N-]
26	16699	[CH3;D1;+0:1]-[O;H0;D2;+0:4]- [C:3]=[O;D1;H0:2]>>[CH2;D1;+0:1]=[N+] =[N-].[O;D1;H0:2]=[C:3]-[OH;D1;+0:4]
27	16462	[OH;D1;+0:1]-[C:2]-[C:3]-[OH;D1;+0:4]>>C-C1(-C)-[O;H0;D2;+0:1]-[C:2]-[C:3]-[O;H0;D2;+0:4]-1
28	16426	[O;-H0;D1]-[N+;H0;D3](=[O;H0;D1;+0])- [c;H0;D3;+0:2](:[c:1]):[c:3]>>[c:1]:[cH;D2;+0:2]:[c:3]
29	15932	[#7:4]-[C:5](=[O;D1;H0:6])- [C:7]-[NH;D2;+0:8]-[C;H0;D3;+0:1](-[C:2])=[O;H0;D1;+0:3]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[#7:4]-[C:5](=[O;D1;H0:6])- [C:7]-[NH2;D1;+0:8]
30	15870	[O;D1;H0:5]=[C:4]/[CH;D2;+0:3]=[CH;D2;+0:1]/[c:2]>>O=[CH;D2;+0:1]-[c:2].[CH3;D1;+0:3]-[C:4]=[O;D1;H0:5]
31	15714	[c:2]:[c;H0;D3;+0:1](:[c:3])- [c;H0;D3;+0:4](:[c:5]):[c:6]>>Br-[c;H0;D3;+0:1](:[c:2]):[c:3].O-B(-O)-[c;H0;D3;+0:4](:[c:5]):[c:6]
32	15499	[C:4]-[O;H0;D2;+0:5]-[C;H0;D3;+0:1](=[O;D1;H0:2])- [c:3]>>Cl-[C;H0;D3;+0:1](=[O;D1;H0:2])- [c:3].[C:4]-[OH;D1;+0:5]
33	15093	[O;D1;H0:1]=[C:2]-[OH;D1;+0:3]>>[O;D1;H0:1]=[C:2]-[O;H0;D2;+0:3]-C-c1:c:c:c:c:c:1
34	14885	[OH;D1;+0:2]-[c:1]>>[c:1]-[O;H0;D2;+0:2]-C-c1:c:c:c:c:c:1
35	14817	[#8:4]-[C:5](=[O;D1;H0:6])- [C:7]-[NH;D2;+0:8]-[C;H0;D3;+0:1](-[C:2])=[O;H0;D1;+0:3]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[#8:4]-[C:5](=[O;D1;H0:6])- [C:7]-[NH2;D1;+0:8]
36	13983	[C:3]-[CH2;D2;+0:2]-[OH;D1;+0:1]>>C-[O;H0;D2;+0:1]-[C;H0;D3;+0:2](=O)-[C:3]
37	13552	[C:5]-[O;H0;D2;+0:6]-[Si;H0;D4;+0:1](-[C:2])(-[C;D1;H3:3])- [C;D1;H3:4]>>Cl-[Si;H0;D4;+0:1](-[C:2])(-[C;D1;H3:3])- [C;D1;H3:4].[C:5]-[OH;D1;+0:6]
38	13497	[C:2]-[C;H0;D3;+0:1](=[O;H0;D1;+0:3])- [O;H0;D2;+0:5]-[C;D1;H3:4]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[C;D1;H3:4]-[OH;D1;+0:5]
39	13475	[#7:2]-[CH2;D2;+0:1]-[C:3]>>O=[C;H0;D3;+0:1](-[#7:2])- [C:3]
40	13475	[c:5]:[c;H0;D3;+0:4](:[c:6])- [c;H0;D3;+0:1](:[c:2]):[c:3]>>Br-[c;H0;D3;+0:1](:[c:2]):[c:3].O-B(-O)-[c;H0;D3;+0:4](:[c:5]):[c:6]
41	13080	[C:4]-[NH;D2;+0:5]-[C;H0;D3;+0:1](=[O;D1;H0:2])- [c:3]>>Cl-[C;H0;D3;+0:1](=[O;D1;H0:2])- [c:3].[C:4]-[NH2;D1;+0:5]
42	12482	[C:1]-[OH;D1;+0:2]>>[C:1]-[O;H0;D2;+0:2]-C-c1:c:c:c:c:1
43	12366	[O;D1;H0:5]=[C:4]-[CH;D2;+0:3]=[CH;D2;+0:1]- [c:2]>>O=[CH;D2;+0:1]- [c:2].[CH3;D1;+0:3]-[C:4]=[O;D1;H0:5]
44	12147	[C:3]-[O;H0;D2;+0:4]-[CH2;D2;+0:1]- [c:2]>>Br-[CH2;D2;+0:1]- [c:2].[C:3]-[OH;D1;+0:4]
45	11823	[O;D1;H0:2]=[S;H0;D4;+0:1](=[O;D1;H0:3])(-[c:4])- [NH;D2;+0:5]- [c:6]>>Cl-[S;H0;D4;+0:1](=[O;D1;H0:2])(=[O;D1;H0:3])- [c:4].[NH2;D1;+0:5]- [c:6]
46	11728	[C:1]-[CH;D3;+0:2](-[OH;D1;+0:3])- [c:4]>>[C:1]-[C;H0;D3;+0:2](=[O;H0;D1;+0:3])- [c:4]
47	11376	[C:5]-[NH;D2;+0:6]-[S;H0;D4;+0:1](=[O;D1;H0:2])(=[O;D1;H0:3])- [c:4]>>Cl-[S;H0;D4;+0:1](=[O;D1;H0:2])(=[O;D1;H0:3])- [c:4].[C:5]-[NH2;D1;+0:6]
48	11178	[C:4]-[O;H0;D2;+0:5]-[C;H0;D3;+0:1](-[C:2])=[O;H0;D1;+0:3]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[C:4]-[OH;D1;+0:5]
49	11110	[C:1]-[CH2;D2;+0:2]-[OH;D1;+0:3]>>[C:1]-[CH;D2;+0:2]=[O;H0;D1;+0:3]
50	11086	[O;H0;D1;+0:1]=[CH;D2;+0:2]- [c:3]>>[OH;D1;+0:1]-[CH2;D2;+0:2]- [c:3]
51	11054	[C:8]-[NH;D2;+0:9]-[C;H0;D3;+0:1](=[O;D1;H0:2])- [#8:3]- [C:4](-[C;D1;H3:5])(-[C;D1;H3:6])- [C;D1;H3:7]>>C-C(-C)(-C)-O-C(=O)-O-[C;H0;D3;+0:1](=[O;D1;H0:2])- [#8:3]- [C:4](-[C;D1;H3:5])(-[C;D1;H3:6])- [C;D1;H3:7].[C:8]-[NH2;D1;+0:9]
52	10771	[C:3]-[N;H0;D2;+0:4]=[CH;D2;+0:1]- [c:2]>>O=[CH;D2;+0:1]- [c:2].[C:3]-[NH2;D1;+0:4]
53	10702	[C:4]-[N;H0;D3;+0:5](-[c:6])- [C;H0;D3;+0:1](-[C:2])=[O;H0;D1;+0:3]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[C:4]-[NH;D2;+0:5]- [c:6]
54	10571	[C:2]#[CH;D1;+0:1]>>C-[Si](C)(-C)-[C;H0;D2;+0:1]#[C:2]
55	10329	[C:2]-[CH2;D2;+0:1]-[O;D1;H1:3]>>O=[C;H0;D3;+0:1](-[C:2])- [O;D1;H1:3]

56	10168	[C:2]-[NH;D2:+0:1]-[C:3].[Cl;H1;D0:+0]>>C-C(-C)(-C)-O-C(=O)-[N;H0;D3:+0:1](-[C:2])- [C:3]
57	10151	[#7;a:4]:[c:3]:[c:2]-[NH2;D1:+0:1]>>O=[N+;H0;D3:1](-[O-])[c:2]:[c:3]:[#7;a:4]
58	10122	[O;H0;D1:+0:2]=[C;H0;D3:+0:1](-[c:3])- [NH;D2:+0:4]-[c:5]>>O=[C;H0;D3:+0:1](-[OH;D1:+0:2])- [c:3].[NH2;D1:+0:4]-[c:5]
59	10120	[C:2]-[NH2;D1:+0:1]>>O=C1-[N;H0;D3:+0:1](-[C:2])-C(=O)-c2:c:c:c:c:2-1
60	10010	[C:4]-[O;H0;D2:+0:5]-[C;H0;D3:+0:1](-[C:2])=[O;D1;H0:3]>>Cl-[C;H0;D3:+0:1](-[C:2])=[O;D1;H0:3].[C:4]- [OH;D1:+0:5]
61	9936	[C:2]-[C;H0;D3:+0:1](=[O;H0;D1:+0:3])- [N;H0;D3:+0:5](-[C:4])- [C:6]>>O=[C;H0;D3:+0:1](-[C:2])- [OH;D1:+0:3].[C:4]-[NH;D2:+0:5]-[C:6]
62	9927	[C:1]-[C;H0;D3:+0:2](-[C:3])=[O;H0;D1:+0:4]>>[C:1]-[CH;D3:+0:2](-[C:3])- [OH;D1:+0:4]
63	9815	[OH;D1:+0:1]-[CH2;D2:+0:2]-[c:3]>>[O;H0;D1:+0:1]-[CH;D2:+0:2]-[c:3]
64	9787	[O;D1;H0:6]=[C:5]-[#7:4]-[N;H0;D2:+0:3]=[CH;D2:+0:1]-[c:2]>>O=[CH;D2:+0:1]-[c:2].[NH2;D1:+0:3]-[#7:4]- [C:5]=[O;D1;H0:6]
65	9692	[C:5]-[O;H0;D2:+0:6]-[C;H0;D4:+0:1](-[c:2])(-[c:3])- [c:4]>>Cl-[C;H0;D4:+0:1](-[c:2])(-[c:3])- [c:4].[C:5]- [OH;D1:+0:6]
66	9581	[I;H0;D1:+0]-[c;H0;D3:+0:2](:[c:1]):[c:3]>>[c:1]:[cH;D2:+0:2]:[c:3]
67	9546	[C:2]-[NH2;D1:+0:1].[Cl;H1;D0:+0]>>C-C(-C)(-C)-O-C(=O)-[N;H0;D2:+0:1]- [C:2]
68	9533	[C:4]#[C;H0;D2:+0:5]-[c;H0;D3:+0:1](:[c:2]):[c:3]>>Br-[c;H0;D3:+0:1](:[c:2]):[c:3].[C:4]#[CH;D1:+0:5]
69	9381	[C:2]-[O;H0;D2:+0:3]-[CH3;D1:+0:1]>>I-[CH3;D1:+0:1].[C:2]-[OH;D1:+0:3]
70	9217	[C:2]-[C;H0;D3:+0:1](=[O;D1;H0:3])- [O;H0;D2:+0:5]-[C:4]>>Cl-[C;H0;D3:+0:1](-[C:2])=[O;D1;H0:3].[C:4]- [OH;D1:+0:5]
71	9123	[C:4]-[N;H0;D3:+0:5](-[C:6])- [C;H0;D3:+0:1](=[O;H0;D1:+0:2])- [c:3]>>O=[C;H0;D3:+0:1](-[OH;D1:+0:2])- [c:3].[C:4]-[NH;D2:+0:5]-[C:6]
72	8891	[O;D1;H0:2]=[C;H0;D3:+0:1](-[c:3])- [O;H0;D2:+0:4]-[c:5]>>Cl-[C;H0;D3:+0:1](-[O;D1;H0:2])- [c:3].[OH;D1:+0:4]-[c:5]
73	8802	[C:1]-[CH2;D2:+0:2]-[NH2;D1:+0:3]>>[C:1]-[C;H0;D2:+0:2]#[N;H0;D1:+0:3]
74	8688	[C:1]-[CH;D3:+0:2](-[C:3])- [OH;D1:+0:4]>>[C:1]-[C;H0;D3:+0:2](-[C:3])=[O;H0;D1:+0:4]
75	8658	[C:2]-[CH2;D2:+0:1]-[O;H0;D2:+0:3]-[c:4]>>Cl-[CH2;D2:+0:1]-[C:2].[OH;D1:+0:3]-[c:4]
76	8629	[#8:2]-[C;H0;D3:+0:1](=[O;D1;H0:3])- [NH;D2:+0:5]-[C:4]>>Cl-[C;H0;D3:+0:1](-[#8:2])=[O;D1;H0:3].[C:4]- [NH2;D1:+0:5]
77	8620	[C:2]-[CH2;D2:+0:1]-[N;H0;D3:+0:4](-[C:3])- [C:5]>>Br-[CH2;D2:+0:1]-[C:2].[C:3]-[NH;D2:+0:4]-[C:5]
78	8612	[C:1]-[NH;D2:+0:2]-[C:3]>>[C:1]-[N;H0;D3:+0:2](-[C:3])- C-c1:c:c:c:c:1
79	8340	[Br;H0;D1:+0]-[CH2;D2:+0:1]-[c:2]>>[CH3;D1:+0:1]-[c:2]
80	8335	[C:8]-[N;H0;D3:+0:9](-[C:10])- [C;H0;D3:+0:1](=[O;D1;H0:2])- [#8:3]-[C:4](-[C;D1;H3:5])(-[C;D1;H3:6])- [C;D1;H3:7]>>C-C(-C)(-C)-O-C(=O)-O-[C;H0;D3:+0:1](=[O;D1;H0:2])- [#8:3]-[C:4](-[C;D1;H3:5])(-[C;D1;H3:6])- [C;D1;H3:7].[C:8]-[NH;D2:+0:9]-[C:10]
81	8283	[Br;H0;D1:+0]-[CH2;D2:+0:1]-[C:2]>>O-[CH2;D2:+0:1]-[C:2]
82	7980	[CH3;D1:+0:1]-[O;H0;D2:+0:2]-[c:3]>>I-[CH3;D1:+0:1].[OH;D1:+0:2]-[c:3]
83	7938	[C:1]-[OH;D1:+0:2]>>[C:1]-[O;H0;D2:+0:2]-C1-C-C-C-C-O-1
84	7918	[C:5]-[N;H0;D3:+0:6](-[C:7])- [S;H0;D4:+0:1](=[O;D1;H0:2])(=[O;D1;H0:3])- [c:4].[C:5]-[NH;D2:+0:6]-[C:7]
85	7815	[C:2]-[C;H0;D3:+0:1](=[C;H0;D1:+0])- [O;H1;D1:+0]>>N#[C;H0;D2:+0:1]-[C:2]
86	7795	[C:2]-[Si;H0;D4:+0:1](-[C;D1;H3:3])- [C;D1;H3:4].[C:5]-[OH;D1:+0:6]
87	7774	[C:4]-[O;H0;D2:+0:5]-[C;H0;D3:+0:1](-[C;D1;H3:2])=[O;D1;H0:3]>>C-C(=O)-O-[C;H0;D3:+0:1](-[C;D1;H3:2])=[O;D1;H0:3].[C:4]-[OH;D1:+0:5]
88	7752	[C:4]-[NH;D2:+0:5]-[C;H0;D3:+0:1](-[C:2])=[O;D1;H0:3]>>Cl-[C;H0;D3:+0:1](-[C:2])=[O;D1;H0:3].[C:4]- [NH2;D1:+0:5]
89	7751	[C:2]-[NH2;D1:+0:1]>>O=C(-O-C-c1:c:c:c:c:1)- [NH;D2:+0:1]-[C:2]
90	7749	[#8:5]-[B;H0;D3:+0:4](-[#8:6])- [c;H0;D3:+0:1](:[c:2]):[c:3]>>Br-[c;H0;D3:+0:1](:[c:2]):[c:3].C-C1(-C)-O-B(-[B;H0;D3:+0:4])- [#8:5]-[#8:6]-O-C-1(-C)-C
91	7734	[#7:3]-[N;H0;D2:+0:4]=[CH;D2:+0:1]-[c:2]>>O=[CH;D2:+0:1]-[c:2].[#7:3]-[NH2;D1:+0:4]
92	7721	[C:3]-[N;H0;D3:+0:4](-[C:5])- [CH2;D2:+0:1]-[C:2]>>Br-[CH2;D2:+0:1]-[C:2].[C:3]-[NH;D2:+0:4]-[C:5]
93	7662	[c:4]-[N;H0;D2:+0:3]=[CH;D2:+0:1]-[c:2]>>O=[CH;D2:+0:1]-[c:2].[NH2;D1:+0:3]-[c:4]
94	7658	[C:4]-[O;H0;D2:+0:5]-[C;H0;D3:+0:1](=[O;H0;D1:+0:2])- [c:3]>>O=[C;H0;D3:+0:1](-[OH;D1:+0:2])- [c:3].[C:4]- [OH;D1:+0:5]
95	7487	[C:2]-[C;H0;D3:+0:1](=[O;H0;D1:+0:3])- [O;H0;D2:+0:5]-[C:4]>>O=[C;H0;D3:+0:1](-[C:2])- [OH;D1:+0:3].[C:4]- [OH;D1:+0:5]
96	7411	[C:1]-[C;H0;D3:+0:2](-[C:3])=[O;H0;D1:+0:4]>>[C:1]-[CH;@:@;D3:+0:2](-[C:3])- [OH;D1:+0:4]
97	7385	[C:3]-[NH;D2:+0:4]-[CH2;D2:+0:1]-[c:2]>>O=[CH;D2:+0:1]-[c:2].[C:3]-[NH2;D1:+0:4]
98	7366	[C:3]-[N;H0;D3:+0:4](-[C:5])- [CH2;D2:+0:1]-[c:2]>>O=[CH;D2:+0:1]-[c:2].[C:3]-[NH;D2:+0:4]-[C:5]

99	7336	[C:2]-[CH2;D2;+0:1]-[N;H0;D3;+0:4](-[C:3])- [C:5]>>Cl-[CH2;D2;+0:1]-[C:2].[C:3]-[NH;D2;+0:4]-[C:5]
100	7315	[C:4]-[N;H0;D3;+0:5](-[C:6])- [C:H0;D3;+0:1](=[O;D1;H0:2])- [c:3]>>Cl-[C;H0;D3;+0:1](=[O;D1;H0:2])- [c:3].[C:4]-[NH;D2;+0:5]-[C:6]
101	7237	[C:2]-[CH2;D2;+0:1]-[O;H0;D2;+0:5]-[C:4]=[O;D1;H0:3]>>O-[CH2;D2;+0:1]-[C:2].[O;D1;H0:3]=[C:4]-[OH;D1;+0:5]
102	7070	[C:2]-[CH;D2;+0:1]=[O;H0;D1;+0]>>C=[CH;D2;+0:1]-[C:2]
103	7054	[C:4]-[O;H0;D2;+0:5]-[c;H0;D3;+0:1]([c:2]):[c:3]>>O-[c;H0;D3;+0:1]([c:2]):[c:3].[C:4]-[OH;D1;+0:5]
104	7034	[C:1]-[C;H0;D3;+0:2](-[O;D1;H1:3])=[O;H0;D1;+0]>>[C:1]-[CH2;D2;+0:2]-[O;D1;H1:3]
105	7033	[OH;D1;+0:1]-[c:2].[OH;D1;+0:3]-[c:4]>>C-[O;H0;D2;+0:1]-[c:2].C-[O;H0;D2;+0:3]-[c:4]
106	7029	[NH2;D1;+0:1]-[c:2]>>C-C(=O)-[NH;D2;+0:1]-[c:2]
107	6966	[O;-H0;D1]-[n+;H0;D3:2]([c:1]):[c:3]>>[c:1]:[n;H0;D2;+0:2]:[c:3]
108	6938	[c:2]-[CH2;D2;+0:1]-[O;H0;D2;+0:3]-[c:4]>>Br-[CH2;D2;+0:1]-[c:2].[OH;D1;+0:3]-[c:4]
109	6932	[C:4]-[N;H0;D3;+0:5](-[C:6])- [C;H0;D3;+0:1](-[C:2])=[O;D1;H0:3]>>Cl-[C;H0;D3;+0:1]-[C:2]=[O;D1;H0:3].[C:4]-[NH;D2;+0:5]-[C:6]
110	6867	[OH;D1;+0:1]-[c:2]>>C-C(=O)-[O;H0;D2;+0:1]-[c:2]
111	6849	[#7:4]-[C:3](=[O;D1;H0:5])- [C:2]-[NH2;D1;+0:1].[Cl;H1;D0;+0]>>C-C(-C)(-C)-O-C(=O)-[NH;D2;+0:1]-[C:2]-[C:3](-[#7:4])=[O;D1;H0:5]
112	6826	[Br;H0;D1;+0]-[CH2;D2;+0:1]-[C:2]=[O;D1;H0:3]>>[CH3;D1;+0:1]-[C:2]=[O;D1;H0:3]
113	6819	[#8:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [O;H0;D2;+0:6]-[C:5]=[O;D1;H0:4]>>Cl-[C;H0;D3;+0:1](-[#8:2])=[O;D1;H0:3].[O;D1;H0:4]=[C:5]-[OH;D1;+0:6]
114	6639	[C:1]-[CH2;D2;+0:2]-[CH2;D2;+0:3]-[C:4]>>[C:1]-[CH;D2;+0:2]=[CH;D2;+0:3]-[C:4]
115	6613	[C;D1;H3:4]-[O;H0;D2;+0:5]-[C;H0;D3;+0:1](=[O;H0;D1;+0:2])- [c:3]>>O=[C;H0;D3;+0:1](-[OH;D1;+0:2])- [c:3].[C;D1;H3:4]-[OH;D1;+0:5]
116	6569	[#7:a:2]:[c;H0;D3;+0:1](:[#7;a:3])- [NH;D2;+0:4]-[c:5]>>Cl-[c;H0;D3;+0:1](:[#7;a:2]):[#7;a:3].[NH2;D1;+0:4]-[c:5]
117	6553	[C:1]-[C;H0;D3;+0:2](-[C:3])=[O;H0;D1;+0:4]>>[C:1]-[CH:@;D3;+0:2](-[C:3])[OH;D1;+0:4]
118	6371	[C:2]-[N;H0;D3;+0:3](-[C:4])- [CH3;D1;+0:1]>>O=[CH2;D1;+0:1].[C:2]-[NH;D2;+0:3]-[C:4]
119	6367	[C:1]-[C;H0;D3;+0:2](=[O;D1;H0:3])- [O;H1;D1;+0]>>[C:1]-[CH;D2;+0:2]=[O;D1;H0:3]
120	6363	[#7:4]-[C:3](=[O;D1;H0:5])- [C:2]-[NH2;D1;+0:1]>>O=C(-O-C-c1:c:c:c:c:1)-[NH;D2;+0:1]-[C:2]-[C:3](-[#7:4])=[O;D1;H0:5]
121	6331	[#8:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [N;H0;D3;+0:5](-[C:4])- [C:6]>>Cl-[C;H0;D3;+0:1](-[#8:2])=[O;D1;H0:3].[C:4]-[NH;D2;+0:5]-[C:6]
122	6245	[C:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [NH;D2;+0:5]-[C:4]>>Cl-[C;H0;D3;+0:1](-[C:2])=[O;D1;H0:3].[C:4]-[NH2;D1;+0:5]
123	6225	[O;D1;H0:4]=[C:5]-[O;H0;D2;+0:6]-[c;H0;D3;+0:1](:[c:2]):[c:3]>>O-[c;H0;D3;+0:1](:[c:2]):[c:3].[O;D1;H0:4]=[C:5]-[OH;D1;+0:6]
124	6219	[#8:1]-[CH;D3;+0:2](-[CH2;D2;+0:3]-[C:4])- [O;H0;D2;+0:6]-[C:5]>>[#8:1]-[CH;D2;+0:2]=[CH;D2;+0:3]-[C:4].[C:5]-[OH;D1;+0:6]
125	6184	[c:2]-[CH;D2;+0:1]=[N;H0;D2;+0:3]-[c:4]>>O=[CH;D2;+0:1]-[c:2].[NH2;D1;+0:3]-[c:4]
126	6181	[OH;D1;+0:1]-[CH2;D2;+0:2]-[c:3]>>C-[O;H0;D2;+0:1]-[C;H0;D3;+0:2](=O)-[c:3]
127	6123	[C:1]-[C;H0;D3;+0:2](-[C:3])=[O;H0;D1;+0:4]>>[C:1]-[C;H0;D4;+0:2]1(-[C:3])-O-C-C-[O;H0;D2;+0:4]-1
128	6118	[#7:a:2]:[c;H0;D3;+0:1](:[c:3])- [N;H0;D3;+0:5](-[C:4])- [C:6]>>Cl-[c;H0;D3;+0:1](:[#7;a:2]):[c:3].[C:4]-[NH;D2;+0:5]-[C:6]
129	6079	[C:1]-[S;H0;D4;+0:2](=[O;H0;D1;+0])=(=[O;H0;D1;+0])- [c:3]>>[C:1]-[S;H0;D2;+0:2]-[c:3]
130	6069	[#8:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [NH;D2;+0:4]-[c:5]>>Cl-[C;H0;D3;+0:1](-[#8:2])=[O;D1;H0:3].[NH2;D1;+0:4]-[c:5]
131	6052	[C:1]-[CH:@;D3;+0:2](-[C:3])- [OH;D1;+0:4]>>[C:1]-[C;H0;D3;+0:2](-[C:3])=[O;H0;D1;+0:4]
132	5905	[C:2]-[OH;D1;+0:1]>>O=C(-[O;H0;D2;+0:1]-[C:2])-c1:c:c:c:c:1
133	5822	[O;D1;H0:3]=[C;H0;D3;+0:4](-[NH;D2;+0:1]-[c:2])- [NH;D2;+0:5]-[c:6]>>[NH2;D1;+0:1]-[c:2].[O;D1;H0:3]=[C;H0;D2;+0:4]=[N;H0;D2;+0:5]-[c:6]
134	5790	[C:3]-[N;H0;D3;+0:4](-[C:5])- [CH2;D2;+0:1]-[c:2]>>Br-[CH2;D2;+0:1]-[c:2].[C:3]-[NH;D2;+0:4]-[C:5]
135	5744	[#7:3]-[C:2](=[O;D1;H0:4])- [CH2;D2;+0:1]-[N;H0;D3;+0:6](-[C:5])- [C:7]>>Cl-[CH2;D2;+0:1]-[C:2](-[#7:3])=[O;D1;H0:4].[C:5]-[NH;D2;+0:6]-[C:7]
136	5705	[C:2]-[C;H0;D3;+0:1](=[C;D1;H0:3])- [c;H0;D3;+0:5](:[c:4]):[c:6]>>Cl-[C;H0;D3;+0:1](-[C:2])=[O;D1;H0:3].[c:4]-[cH;D2;+0:5]-[c:6]
137	5679	[C;D1;H3:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [NH;D2;+0:4]-[c:5]>>C-C(=O)-O-[C;H0;D3;+0:1](-[C;D1;H3:2])=[O;D1;H0:3].[NH2;D1;+0:4]-[c:5]
138	5666	[Br;H0;D1;+0]-[CH2;D2;+0:1]-[c:2]>>O-[CH2;D2;+0:1]-[c:2]
139	5624	[C:4]-[NH;D2;+0:5]-[C;H0;D3;+0:1](-[C;D1;H3:2])=[O;D1;H0:3].[C:4]-[NH2;D1;+0:5]
140	5601	[C:2]-[NH;D2;+0:1]-[C:3]>>O=C(-O-C-c1:c:c:c:c:1)-[N;H0;D3;+0:1](-[C:2])- [C:3]

141	5506	[O;D1;H0:3]=[C:2]-[OH;D1;+0:1].[O;D1;H0:6]=[C:5]-[OH;D1;+0:4]>>C-C-[O;H0;D2;+0:1]-[C:2]=[O;D1;H0:3].C-C-[O;H0;D2;+0:4]-[C:5]=[O;D1;H0:6]
142	5426	[#8:2]-[CH2;D2;+0:1]-[O;H0;D2;+0:4]-[C:3]>>Cl-[CH2;D2;+0:1]-[#8:2].[C:3]-[OH;D1;+0:4]
143	5414	[C:4]-[N;H0;D3;+0:5](-[C:6])-;[c;H0;D3;+0:1](:[c:2]):[c:3]>>Br-[c;H0;D3;+0:1](:[c:2]):[c:3].[C:4]-[NH;D2;+0:5]-[C:6]
144	5391	[C:2]-[C;H0;D3;+0:1](=[O;D1;H0:3]).[O;H0;D2;+0:4]-[c:5]>>Cl-[C;H0;D3;+0:1](-[C:2]=[O;D1;H0:3].[O;H0;D2;+0:4]-[C:5]=[O;D1;H0:6]
145	5331	[O;D1;H0:3]=[C:2]-[OH;D1;+0:1].[O;D1;H0:6]=[C:5]-[OH;D1;+0:4]>>C-[O;H0;D2;+0:1]-[C:2]=[O;D1;H0:3].C-[O;H0;D2;+0:4]-[C:5]=[O;D1;H0:6]
146	5303	[C:1]-[NH;D2;+0:2]-[CH2;D2;+0:4]-[C:3]-[OH;D1;+0:5]>>[C:1].[NH2;D1;+0:2].[C:3]1-[CH2;D2;+0:4]-[O;H0;D2;+0:5]-1
147	5276	[c:4]-[O;H0;D2;+0:3]-[CH2;D2;+0:1]-[c:2]>>Br-[CH2;D2;+0:1]-[c:2].[OH;D1;+0:3]-[c:4]
148	5259	[C:2]-[CH2;D2;+0:1]-[Cl;H0;D1;+0]>>O-[CH2;D2;+0:1]-[C:2]
149	5201	[#7:a:2];[c;H0;D3;+0:1](:[#7:a:3])-;[N;H0;D3;+0:5](-[C:4])-;[C:6]>>Cl-[c;H0;D3;+0:1](:[#7:a:2]):[#7:a:3].[C:4]-[NH;D2;+0:5]-[C:6]
150	5129	[C:1].[CH:@;D3;+0:2](-[C:3])-;[OH;D1;+0:4]>>[C:1].[C;H0;D3;+0:2](-[C:3])=[O;H0;D1;+0:4]
151	5126	[C:2]-[CH2;D2;+0:1]-[c:3]>>O=[C;H0;D3;+0:1](-[C:2])-;[c:3]
152	5081	[C:2]-[OH;D1;+0:1]>>C-O-cl1;c::c(-[C;O;H0;D2;+0:1]-[C:2]):c:c1
153	5075	[#8:1]-[C:2]-[CH;D2;+0:3]=[O;H0;D1;+0:4]>>[#8:1]-[C:2]-[CH2;D2;+0:3]-[OH;D1;+0:4]
154	5052	[Cl;H0;D1;+0]-[CH2;D2;+0:1]-[c:2]>>O-[CH2;D2;+0:1]-[c:2]
155	5021	[#7:4]-[C:3]-[C;H0;D3;+0:1](=[O;H0;D1;+0:2])-;[NH;D2;+0:6]-;[C:5]>>O=[C;H0;D3;+0:1](-[OH;D1;+0:2])-;[C:3]-[#7:4].[C:5]-[NH2;D1;+0:6]
156	5019	[C:4]-[N;H0;D3;+0:5](-[C:6])-;[c;H0;D3;+0:1](:[c:2]):[c:3]>>F-[c;H0;D3;+0:1](:[c:2]):[c:3].[C:4]-[NH;D2;+0:5]-[C:6]
157	5008	[Cl;H0;D1;+0]-;[c;H0;D3;+0:2](:[c:1]):[c:3]>>[c:1].[c;H;D2;+0:2]:[c:3]
158	4947	[C:1]-[N;H0;D3;+0:2](-[C:3])-;[CH2;D2;+0:5]-;[C:4]-[OH;D1;+0:6]>>[C:1].[NH;D2;+0:2]-[C:3].[C:4]1-[CH2;D2;+0:5]-[O;H0;D2;+0:6]-1
159	4943	[NH2;D1;+0:1]-[c:2].[NH2;D1;+0:3]-[c:4]>>O=[N+;H0;D3:1](-[O-])-;[c:2].O=[N+;H0;D3:3](-[O-])-;[c:4]
160	4929	[#8:3]-[C:2](-[O;D1;H0:4])-;[CH2;D2;+0:1]-[O;H0;D2;+0:5]-;[c:6]>>Br-[CH2;D2;+0:1]-[C:2](-[#8:3])=[O;D1;H0:4].[OH;D1;+0:5]-;[c:6]
161	4891	[#7:a:5]-[c:4]-;[c:3]-[C;H0;D3;+0:1](-[O;H0;D1;+0:2])-;[NH;D2;+0:7]-;[C:6]>>O=[C;H0;D3;+0:1](-[OH;D1;+0:2])-;[c:3]-;[c:4];[#7:a:5].[C:6]-[NH2;D1;+0:7]
162	4883	[C:2]-[OH;D1;+0:1].[OH;D1;+0:3]-;[C:4]-[C:5]-[OH;D1;+0:6]>>C-C(=O)-[O;H0;D2;+0:1]-[C:2].C-C(=O)-[O;H0;D2;+0:3]-;[C:4]-[C:5]-[O;H0;D2;+0:6]-;[C:C]=O
163	4804	[CH3;D1;+0:1]-[O;H0;D2;+0:4]-[C:3]=[O;D1;H0:2]>>I-[CH3;D1;+0:1].[O;D1;H0:2]=[C:3]-[OH;D1;+0:4]
164	4799	[#7:a:2];[c;H0;D3;+0:1](:[#7:a:3])-;[NH;D2;+0:5]-;[C:4]>>Cl-[c;H0;D3;+0:1](:[#7:a:2]):[#7:a:3].[C:4]-[NH2;D1;+0:5]
165	4745	[C:2]-[C;H0;D3;+0:1](-[N;H2;D1;+0])=[O;H0;D1;+0:3]>>O=[C;H0;D3;+0:1](-[C:2])-;[OH;D1;+0:3]
166	4727	[CH3;D1;+0:1]-[O;H0;D2;+0:2]-;[c:3]>>C-O-S(=O)(=O)-O-[CH3;D1;+0:1].[OH;D1;+0:2]-;[c:3]
167	4706	[C:2]-[CH;D2;+0:1]=[CH;D2;+0:3]-;[C:4]>>C=[CH;D2;+0:1]-[C:2].C=[CH;D2;+0:3]-;[C:4]
168	4701	[C:3]-[N;H0;D3;+0:4](-[C:5])-;[CH2;D2;+0:1]-[C:2]>>Cl-[CH2;D2;+0:1]-[C:2].[C:3]-[NH;D2;+0:4]-;[C:5]
169	4688	[C:1]-;[S;H0;D3;+0:2](-[O;H0;D1;+0])-;[c:3]>>[C:1].[S;H0;D2;+0:2]-;[c:3]
170	4673	[OH;D1;+0:1]-;[c:2]>>C-C(=O)(-C)-[Si](-C)(-C)-[O;H0;D2;+0:1]-;[c:2]
171	4661	[NH2;D1;+0:1]-;[c:2]>>C-C(=O)(-C)-O-C(=O)-[NH;D2;+0:1]-;[c:2]
172	4542	[C:5]-[O;H0;D2;+0:6]-;[Si;H0;D4;+0:1](-[C:2])-;[c:4]>>Cl-[Si;H0;D4;+0:1](-[C:2])(-[c:3])-;[c:4].[C:5]-[OH;D1;+0:6]
173	4535	[O;D1;H0:1]=[C;H0;D3;+0:2](-[O;H1;D1;+0])-;[c:3]>>[O;D1;H0:1]=[CH;D2;+0:2]-;[c:3]
174	4531	[#7:2]-[P;H0;D3;+0:1](-[#8:3])-;[O;H0;D2;+0:5]-;[C:4]>>Cl-[P;H0;D3;+0:1](-[#7:2])-;[#8:3].[C:4]-[OH;D1;+0:5]
175	4501	[C:1]-[NH;D2;+0:2]-;[CH2;D2;+0:3]-;[c:4]>>[C:1].[N;H0;D2;+0:2]=[CH;D2;+0:3]-;[c:4]
176	4493	[C:2]-[OH;D1;+0:1]>>C-[Si](-C)(-C)-[O;H0;D2;+0:1]-;[C:2]
177	4492	[C:4]-[O;H0;D2;+0:5]-;[C;H0;D3;+0:1](-[C;D1;H3:2])=[O;D1;H0:3]>>Cl-[C;H0;D3;+0:1](-[C;D1;H3:2])=[O;D1;H0:3].[C:4]-[OH;D1;+0:5]
178	4489	[C:2]-[CH2;D2;+0:1]-;[I;H0;D1;+0]>>O-[CH2;D2;+0:1]-;[C:2]
179	4402	[C:1]-[C;H0;D3;+0:2](-[O;H0;D1;+0:3])-;[c:4]>>[C:1].[CH;D3;+0:2](-[OH;D1;+0:3])-;[c:4]
180	4371	[C;D1;H2:3]=[C:2]-[CH2;D2;+0:1]-;[O;H0;D2;+0:4]-;[c:5]>>Br-[CH2;D2;+0:1]-;[C:2]=[C;D1;H2:3].[OH;D1;+0:4]-;[c:5]
181	4330	[c:5];[c;H0;D3;+0:4](-[c:6])-;[c;H0;D3;+0:1](:[c:2]):[c:3]>>I-[c;H0;D3;+0:1](:[c:2]):[c:3].O-B(-O)-[c;H0;D3;+0:4](-[c:5]);[c:6]
182	4319	[N-;H0;D1]=[N+;H0;D2]=[N;H0;D2;+0:1]-;[c:2]>>[NH2;D1;+0:1]-;[c:2]
183	4309	[#7:5]-[C:4](-[O;D1;H0:6])-;[C:3]-[NH;D2;+0:1]-;[C:2]>>C-C(=O)(-C)-O-C(=O)-[N;H0;D3;+0:1](-[C:2])-;[C:3]-[C:4](-[#7:S])=[O;D1;H0:6]

184	4274	[C:2]-[NH2;D1;+0:1]>>O=[N+;H0;D3:1](-[C:2])-[-O-]
185	4259	[O;D1;H0:6]=[C:5]-[#7:4]/[N;H0;D2;+0:3]=[CH;D2;+0:1]/[c:2]>>O=[CH;D2;+0:1]-[c:2].[NH2;D1;+0:3]-[#7:4]-[C:5]=[O;D1;H0:6]
186	4194	[C:2]-[OH;D1;+0:1].[OH;D1;+0:3]-[C:4]-[C:5](-[OH;D1;+0:8])- [C:6]-[OH;D1;+0:7]>>C-C(=O)-[O;H0;D2;+0:1]-[C:2].C-C(=O)-[O;H0;D2;+0:3]-[C:4]-[C:5](-[C:6]-[O;H0;D2;+0:7]-C(-C)=O)-[O;H0;D2;+0:8]-C(-C)=O
187	4187	[C:3]/[N;H0;D2;+0:4]=[CH;D2;+0:1]/[c:2]>>O=[CH;D2;+0:1]-[c:2].[C:3]-[NH2;D1;+0:4]
188	4164	[C:2]-[OH;D1;+0:1].[OH;D1;+0:3]-[C:4](-[C:5]-[OH;D1;+0:6])- [C:7]-[OH;D1;+0:8]>>C-C(=O)-[O;H0;D2;+0:1]-[C:2].C-C(=O)-[O;H0;D2;+0:3]-[C:4](-[C:5]-[O;H0;D2;+0:6]-C(-C)=O)-[C:7]-[O;H0;D2;+0:8]-C(-C)=O
189	4113	[#7;a:4]:[c:3]-[C:H0;D3;+0:1](=[O;H0;D1;+0:2])- [NH;D2;+0:6]-[C:5]>>O=[C;H0;D3;+0:1](-[OH;D1;+0:2])- [c:3].[#7;a:4].[C:5]-[NH2;D1;+0:6]
190	4107	[C:5]=[C:4]/[C:2](=[O;D1;H0:3])- [OH;D1;+0:1]>>C-C-[O;H0;D2;+0:1]-[C:2](=[O;D1;H0:3])/[C:4]=[C:5]
191	4104	[C;D1;H3:1]-[CH;D3;+0:2](-[OH;D1;+0:3])- [c:4]>>[C;D1;H3:1].[C;H0;D3;+0:2](=[O;H0;D1;+0:3])- [c:4]
192	4100	[O;D1;H1:2]-[CH2;D2;+0:1]-[c:3]>>O=[C;H0;D3;+0:1](-[O;D1;H1:2])- [c:3]
193	4056	[CH3;D1;+0:1]-[O;H0;D2;+0:4]-[C:3]=[O;D1;H0:2]>>C-[Si](-C)-[CH;D2;+0:1]=[N+]=[N-].[O;D1;H0:2]-[C:3].[OH;D1;+0:4]
194	4038	[C:1]-[CH2;D2;+0:2]-[CH3;D1;+0:3]>>[C:1]-[CH;D2;+0:2]=[CH2;D1;+0:3]
195	4036	[#7;a:4]:[c:3]-[C;H0;D3;+0:1](:[c:2])- [c;H0;D3;+0:5](:[c:6]):[c:7]>>Br-[c;H0;D3;+0:1](:[c:2]):[c:3].[#7;a:4].O-B(-O)-[c;H0;D3;+0:5](:[c:6]):[c:7]
196	4017	[#8:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [O;H0;D2;+0:5]-[C:4]>>Cl-[C;H0;D3;+0:1](-[#8:2])=[O;D1;H0:3].[C:4]-[OH;D1;+0:5]
197	3993	[C:1]-[NH;D2;+0:2]-[C;H0;D3;+0:4](=[O;D1;H0:3])- [NH;D2;+0:5]-[c:6]>>[C:1]-[NH2;D1;+0:2].[O;D1;H0:3]=[C;H0;D2;+0:4]=[N;H0;D2;+0:5]-[c:6]
198	3978	[C:1]-[O;H0;D2;+0:2]-[C;H0;D3;+0:4](=[O;D1;H0:3])- [NH;D2;+0:5]-[c:6]>>[C:1]-[OH;D1;+0:2].[O;D1;H0:3]=[C;H0;D2;+0:4]=[N;H0;D2;+0:5]-[c:6]
199	3935	[C:1]-[S;H0;D4;+0:2](-[C:3])(=[O;H0;D1;+0])=[O;H0;D1;+0]>>[C:1]-[S;H0;D2;+0:2]-[C:3]
200	3859	[CH3;D1;+0:1]-[n;H0;D3;+0:3](:[c:2]):[c:4]>>I-[CH3;D1;+0:1].[c:2].[nH;D2;+0:3].[c:4]
201	3856	[C:1]-[C;H0;D3;+0:2](-[NH2;D1;+0:3])=[O;H0;D1;+0]>>[C:1]-[C;H0;D2;+0:2]#[N;H0;D1;+0:3]
202	3854	[C:3]-[O;H0;D2;+0:4]-[CH2;D2;+0:1]-[c:2]>>Cl-[CH2;D2;+0:1]-[c:2].[C:3]-[OH;D1;+0:4]
203	3836	[C:2]-[C;H0;D3;+0:1](=[O;H0;D1;+0:3])- [c;H0;D3;+0:5](:[c:4]):[c:6]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[c:4].[c;H;D2;+0:5]:[c:6]
204	3833	[OH;D1;+0:1]-[c:2]>>C-O-C-[O;H0;D2;+0:1]-[c:2]
205	3827	[O;D1;H0:2]=[C;H0;D3;+0:1](-[c:3])- [c;H0;D3;+0:5](:[c:4]):[c:6]>>O=[C;H0;D3;+0:1](-[O;D1;H0:2])- [c:3].[c:4].[c;H;D2;+0:5]:[c:6]
206	3826	[C:2]-[C;H0;D3;+0:1](=[O;H0;D1;+0:3])- [O;H0;D2;+0:7]-[#7:6](-[C:5]=[O;D1;H0:4])- [C:8]=[O;D1;H0:9]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[O;D1;H0:4]=[C:5]-[#7:6](-[OH;D1;+0:7])- [C:8]=[O;D1;H0:9]
207	3798	[C;D1;H3:3]-[C:2](=[O;D1;H0:4])- [O;H0;D2;+0:1]-[c;H0;D3;+0:5](:[c:6]):[c:7]>>C-C(=O)-[O;H0;D2;+0:1]-[C:2](-[C;D1;H3:3])=[O;D1;H0:4].O-[c;H0;D3;+0:5](:[c:6]):[c:7]
208	3784	[NH2;D1;+0:1]-[C;H0;D3;+0:2](=[O;H0;D1;+0])- [c:3]>>[N;H0;D1;+0:1]#[C;H0;D2;+0:2]-[c:3]
209	3781	[C:3]-[N;H0;D3;+0:4](-[C:5])- [CH2;D2;+0:1]-[C:2]>>O=[CH;D2;+0:1]-[C:2].[C:3]-[NH;D2;+0:4]-[C:5]
210	3778	[C:7]-[NH;D2;+0:8]-[c;H0;D3;+0:1]1:[c:2].[c:3]:[#7;a:4].[c:5].[c:6]:1>>Cl-[c;H0;D3;+0:1]1:[c:2].[c:3]:[#7;a:4].[c:5].[c:6]:1.[C:7]-[NH2;D1;+0:8]
211	3742	[C:1]-[CH;D3;+0:2](-[C;D1;H3:3])- [OH;D1;+0:4]>>[C:1]-[C;H0;D3;+0:2](-[C;D1;H3:3])=[O;H0;D1;+0:4]
212	3715	[NH2;D1;+0:1]-[CH2;D2;+0:2]-[c:3]>>[N;H0;D1;+0:1]#[C;H0;D2;+0:2]-[c:3]
213	3704	[OH;D1;+0:1]-[#15:2]-[OH;D1;+0:3]>>C-C-[O;H0;D2;+0:1]-[#15:2]-[O;H0;D2;+0:3]-C-C
214	3664	[C:4]-[NH;D2;+0:5]-[c;H0;D3;+0:1](:[c:2]):[c:3]>>F-[c;H0;D3;+0:1](:[c:2]):[c:3].[C:4]-[NH2;D1;+0:5]
215	3652	[C:3]-[N;H0;D3;+0:4](-[C:5])- [CH2;D2;+0:1]-[c:2]>>Cl-[CH2;D2;+0:1]-[c:2].[C:3]-[NH;D2;+0:4]-[C:5]
216	3603	[C:4]-[O;H0;D2;+0:5]-[c;H0;D3;+0:1](:[c:2]):[c:3]>>F-[c;H0;D3;+0:1](:[c:2]):[c:3].[C:4]-[OH;D1;+0:5]
217	3594	[C:1]-[OH;D1;+0:2].[C:4]-[OH;D1;+0:3]>>[C:1]-[O;H0;D2;+0:2]-C(-[O;H0;D2;+0:3]-[C:4])-c1:c:c:c:c:e1
218	3584	[#7:a:2]:[c;H0;D3;+0:1](:[c:3])- [c;H0;D3;+0:4](:[c:5]):[c:6]>>Br-[c;H0;D3;+0:1](:[#7;a:2]):[c:3].O-B(-O)- [c;H0;D3;+0:4](:[c:5]):[c:6]
219	3548	[O;D1;H0:4]=[C:3]-[#7:2]-[NH;D2;+0:1]-[C;H0;D3;+0:6](=[S;D1;H0:5])- [NH;D2;+0:7]-[c:8]>>[NH2;D1;+0:1]-[#7:2]-[C:3]=[O;D1;H0:4].[S;D1;H0:5]=[C;H0;D2;+0:6]=[N;H0;D2;+0:7]-[c:8]
220	3529	[C:1]-[CH2;D2;+0:2]-[CH2;D2;+0:3]-[O;H1;D1;+0]>>[C:1]-[CH;D2;+0:2]=[CH2;D1;+0:3]
221	3520	[O;D1;H0:3]=[C:4]-[O;H0;D2;+0:5]-[CH2;D2;+0:1]-[c:2]>>Br-[CH2;D2;+0:1]-[c:2].[O;D1;H0:3]=[C:4]-[OH;D1;+0:5]
222	3508	[#16:2]-[O;H0;D2;+0:1]-[c;H0;D3;+0:3](:[c:4]):[c:5]>>F-C(-F)(-F)-S(=O)(=O)-[O;H0;D2;+0:1]-[#16:2].O-[c;H0;D3;+0:3](:[c:4]):[c:5]
223	3506	[#7;a:2]:[c;H0;D3;+0:1](:[c:3])- [NH;D2;+0:5]-[C:4]>>Cl-[c;H0;D3;+0:1](:[#7;a:2]):[c:3].[C:4]-[NH2;D1;+0:5]
224	3497	[C:2]-[C;H0;D3;+0:1](=[O;H0;D1;+0:3])- [O;H0;D2;+0:4]-[c:5]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[OH;D1;+0:4]-[c:5]

225	3494	[c:4]/[N;H0;D2;+0:3]=[CH;D2;+0:1]/[c:2]>>O=[CH;D2;+0:1]-[c:2].[NH2;D1;+0:3]-[c:4]
226	3484	[C;D1;H3:2]-[S;H0;D4;+0:1](=[O;D1;H0:3])(=[O;D1;H0:4])-[NH;D2;+0:5]-[c:6]>>Cl-[S;H0;D4;+0:1](-[C;D1;H3:2])(=[O;D1;H0:3])=[O;D1;H0:4].[NH2;D1;+0:5]-[c:6]
227	3474	[C:1]-[NH;D2;+0:2]-[C;H0;D3;+0:4](=[S;D1;H0:3])- [NH;D2;+0:5]-[c:6]>>[C:1]-[NH2;D1;+0:2].[S;D1;H0:3]=[C;H0;D2;+0:4]=[N;H0;D2;+0:5]-[c:6]
228	3444	[O;D1;H0:2]-[S;H0;D4;+0:1](=[O;D1;H0:3])(-[c:4])[O;H0;D2;+0:5]-[c:6]>>Cl-[S;H0;D4;+0:1](=[O;D1;H0:2])(=[O;D1;H0:3])[c:4].[OH;D1;+0:5]-[c:6]
229	3404	[C:1]-[C;H0;D3;+0:2](-[C;D1;H3:3])=[O;H0;D1;+0:4]>>[C:1]-[CH;D3;+0:2](-[C;D1;H3:3])- [OH;D1;+0:4]
230	3378	[C:2]-[CH2;D2;+0:1]-[O;H0;D2;+0:3]-[c:4]>>O-[CH2;D2;+0:1]-[C:2].[OH;D1;+0:3]-[c:4]
231	3353	[#8:2]-[CH2;D2;+0:1]-[O;H0;D2;+0:3]-[c:4]>>Cl-[CH2;D2;+0:1]-[#8:2].[OH;D1;+0:3]-[c:4]
232	3340	[C;D1;H3:1]-[S;H0;D4;+0:2](=[O;H0;D1;+0])(=[O;H0;D1;+0])- [c:3]>>[C;D1;H3:1]-[S;H0;D2;+0:2]-[c:3]
233	3323	[C;D1;H3:5]-[C:4](-[C;D1;H3:6])(-[C;D1;H3:7])-[#8:3]-[C;H0;D3;+0:1](=[O;D1;H0:2])- [NH;D2;+0:8]-[c:9]>>C-C-(C)-O-C(=O)-O-[C;H0;D3;+0:1](=[O;D1;H0:2])- [#8:3]-[C:4](-[C;D1;H3:5])(-[C;D1;H3:6])- [C;D1;H3:7].[NH2;D1;+0:8]-[c:9]
234	3291	[C:1]-[N;H0;D3;+0:2](-[C:3])- [C;H0;D3;+0:5](=[O;D1;H0:4])- [NH;D2;+0:6]-[c:7]>>[C:1]-[NH;D2;+0:2]-[C:3].[O;D1;H0:4]=[C;H0;D2;+0:5]=[N;H0;D2;+0:6]-[c:7]
235	3287	[C:1]-[S;H0;D3;+0:2](-[C:3])=[O;H0;D1;+0]>>[C:1]-[S;H0;D2;+0:2]-[C:3]
236	3279	[C:6]-[CH2;D2;+0:5]-[O;H0;D2;+0:1]-[C:2](-[C;D1;H3:3])=[O;D1;H0:4]>>C-C(=O)-[O;H0;D2;+0:1]-[C:2](-[C;D1;H3:3])=[O;D1;H0:4].[O;-[CH2;D2;+0:5]-[C:6]
237	3268	[c:2].[c;H0;D3;+0:1](:[c:3])- [c;H0;D3;+0:4](:[c:5]):[c:6]>>I-[c;H0;D3;+0:1](:[c:2]):[c:3].O-B(-O)-[c;H0;D3;+0:4](:[c:5]):[c:6]
238	3264	[C:1]-[CH;D3;+0:2]-[CH2;D2;+0:3]-[O;H0;D2;+0:1]>>[C:1]-[CH;D2;+0:2]=[CH2;D1;+0:3]
239	3264	[I;H0;D1;+0]-[c;H0;D3;+0:1](:[c:2]):[c:3]>>N-[c;H0;D3;+0:1](:[c:2]):[c:3]
240	3244	[#7:2]-[CH2;D2;+0:1]-[c:3]>>O=[C;H0;D3;+0:1](:[#7:2])- [c:3]
241	3237	[C:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [N;H1;D2;+0]-[N;H2;D1;+0]>>C-C-O-[C;H0;D3;+0:1](-[C:2])=[O;D1;H0:3]
242	3236	[C:2]-[C;H0;D4;+0:1]1(-[C:3])- [O;H0;D2;+0:4]-[C:5]-[C:6]-[O;H0;D2;+0:7]-1>>O=[C;H0;D3;+0:1](-[C:2])- [C:3].[OH;D1;+0:4]-[C:5]-[C:6]-[OH;D1;+0:7]
243	3231	[C:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [N;H0;D3;+0:5](-[C:4])- [C:6]>>Cl-[C;H0;D3;+0:1](-[C:2])=[O;D1;H0:3]
244	3213	[C:2]-[C;H0;D3;+0:1](=[O;D1;H0:3])- [N;H1;D2;+0]-[N;H2;D1;+0]>>C-O-[C;H0;D3;+0:1](-[C:2])=[O;D1;H0:3]
245	3198	[c:8]-[NH;D2;+0:7]-[c;H0;D3;+0:1]1:[c:2]:[c:3]:[#7;a:4]:[c:5]:[c:6]:1>>Cl-[c;H0;D3;+0:1]1:[c:2]:[c:3]:[#7;a:4]:[c:5]:[c:6]:1.[NH2;D1;+0:7]-[c:8]
246	3191	[C;H3;D1;+0]-[O;H0;D2;+0:3]-[C:2]=[O;D1;H0:1]>>[O;D1;H0:1]=[C:2].[OH;D1;+0:3]
247	3155	[C:2]-[Si;H0;D4;+0:1](-[C;D1;H3:3])(-[C;D1;H3:4])- [O;H0;D2;+0:5]-[c:6]>>Cl-[Si;H0;D4;+0:1](-[C:2])(-[C;D1;H3:3])- [C;D1;H3:4].[OH;D1;+0:5]-[c:6]
248	3152	[#8:4]-[C:5](=[O;D1;H0:6])- [C:7]-[N;H0;D3;+0:8](-[C:9])- [C;H0;D3;+0:1](-[C:2])=[O;H0;D1;+0:3]>>O=[C;H0;D3;+0:1](-[C:2])- [OH;D1;+0:3].[#8:4]-[C:5](=[O;D1;H0:6])- [C:7]-[NH;D2;+0:8]-[c:9]
249	3144	[O;H0;D1;+0:2]=[C;H0;D3;+0:1](-[c:3])- [O;H0;D2;+0:4]-[c:5]>>O=[C;H0;D3;+0:1](-[OH;D1;+0:2])- [c:3].[OH;D1;+0:4]-[c:5]
250	3131	[O;D1;H0:2]=[S;H0;D4;+0:1](=[O;D1;H0:3])(-[c:4])- [n;H0;D3;+0:6](-[c:5]):[c:7]>>Cl-[S;H0;D4;+0:1](=[O;D1;H0:2])(=[O;D1;H0:3])- [c:4].[c:5].[nH;D2;+0:6]:[c:7]

S1.3 Stereochemistry in template application

A common limitation of synthesis planning programs is their inability to handle stereochemistry. Including information about chirality (tetrahedral centers, cis/trans double bonds) during the template extraction process and generation of SMARTS strings is only part of the challenge. The preservation, introduction, inversion, or destruction of chirality must be properly handled when *applying* these transformations to product molecules. To this end, we developed the RDChiral wrapper (<https://github.com/connorcoleyley/rdchiral>) for the open-source cheminformatics toolkit RDKit which is used extensively throughout our code (33).

RDChiral increases the computational expense of applying templates by requiring additional checks on the specification of chirality in the template and target molecule (or lack thereof) and whether the match is acceptable.

S1.4 Search prioritization

We trained a feedforward neural network model to predict which of the 163,723 transform rules are most applicable to a target molecule based on its molecular fingerprint (ECFP4 (34)) (28). Applying only the templates perceived to be most relevant served two roles: (a) reducing computational cost of template subgraph matching, and (b) increasing the likelihood of proposing a chemically-feasible reaction. We have focused much of our effort on this latter point: ensuring that reaction suggestions are not “false positives”, i.e., recommendations that would not work in the laboratory. This is essential for maximizing the probability of experimental success and enable syntheses to be carried out autonomously with minimal human intervention.

The model is a feedforward neural network that takes an ECFP4 fingerprint of a product molecule, folded to length 2048, as input. After passing through five hidden layers of 300 nodes each with ReLU activation, the model outputs a probability distribution (softmax activation) over the full set of 163,723 rules. These probabilities can be interpreted as the likelihood that applying that rule is a reasonable transformation. The model is trained to predict the “true” retrosynthetic template used to synthesize the product molecules observed in the Reaxys database.

The number of epochs to use for training was determined using a random 80% subset for training, 10% subset for validation, and an early stopping mechanism. Performance on the validation set reached 41.4%, 67.2%, 74.9%, 80.8%, 86.6%, and 89.8% top-1, top-5, top-10, top-20, top-50, and top-100 accuracies. Because the intent is to have a useful deployed tool, we then retrain on the entire dataset using a fixed training schedule. The code, but not the proprietary Reaxys data, is available at <https://github.com/connorcoleyley/retrotemp>.

The search itself is coordinated by a single process using a Tree Search, based on (21). The “tree builder coordinator” chooses leaf nodes based on a combination of how promising that branch is according to the template relevance network and how many times it has already visited it, with or without successfully terminating in a buyable pathway. The coordinator maintains several active pathways (typically, 12) and sends requests to parallelized “tree builder worker” processes that apply a specific template to a specific molecule, and return the template relevance network results for the corresponding reactants. A virtual loss is added to chemicals in actively-expanding pathways to encourage exploration, otherwise multiple active pathways would waste time exploring the same branch (35). The webapp allows for results to be returned as soon as a single pathway is found, or for *all* results to be returned after a specified amount of time.

S1.5 Stop criteria

There are several stop criteria available in the program that can be combined with Boolean logic. The standard criterion is that the molecule is commercially-available for \$100/g, i.e., present in a list of such compounds from the eMolecules and SigmaAldrich catalogs. As one would imagine, this criterion is quite sensitive to the exact database; a single missing key starting material might prevent the program from identifying a plausible synthetic route. It is often useful to lower the price threshold, e.g. to \$5/g, to force starting materials back to simpler precursors. The data file of compounds considered buyable can be easily augmented or replaced.

The second stop criterion is based on chemical properties. When used with OR buyable logic, we set a maximum number of allowable atoms with certain atomic numbers (number of C, N, O, and H atoms). For example, setting these values to 6, 1, 1, and 99 corresponds to the assumption that

any species with at most 6 carbon atoms, 1 nitrogen, and 1 oxygen is able to be used as a starting material. When used with AND buyable logic, this criterion is used to force starting materials to be smaller and potentially break key bonds. For example, this logic could have been used in the route design for safinamide to require that starting materials be <\$100/g AND have at most 7 carbon atoms, instead of simply lowering the price threshold to avoid the larger \$33/g intermediate.

The third stop criterion is based on chemical popularity: if a chemical has been seen enough times, consider it available. This is to replicate what was used in (21) for their stop criterion, namely that chemicals had been observed at least 5 times previously. Different thresholds can be set for the number of observed precedents *as reactants* and *as products*. When enabled, this criterion is used with OR logic with the previous two. Various combinations are possible, e.g. ((buyable AND simple) OR popular). We have found this setting particularly useful when developing routes for complex small molecules, where connecting back to known compounds is sufficient.

S1.6 “Fast filter” screening of suggestions

The “fast filter” is modeled after Segler et al.’s “in-scope filter” (21). It is designed to evaluate retrosynthetic suggestions that are generated immediately after template applications, before any conditions are proposed. The model is a feed forward neural network model that predicts a soft binary classification (a score between 0 and 1) based on a reaction fingerprint and the structure of the product (both ECFP4, folded to length 2048, not including chirality). The product fingerprint is passed through five highway layers of 1024 nodes with ELU activation. The reaction fingerprint – the product fingerprint minus the reactant(s) fingerprint – is passed through a single fully-connected layer of width 1024 and ELU activation. The dot product between the two is passed through a sigmoid activation to produce the final likelihood score between 0 and 1.

Because the classifier requires negative examples to train, which are rarely reported in the literature, we generated synthetic negative data through the application of forward reaction templates (27). The template extraction procedure exactly followed (27) and produced a library of forward synthetic transforms. From these transforms, we prepared a final set consisting of rules observed at least 25 times. Every rule was applied to a set of randomly-extracted positive reaction examples from Reaxys to generate plausible, but unreported candidate outcomes. An example of this process is shown in **Fig. S4**.

This procedure of generating negative reaction data is imperfect for two reasons, but works well in practice. First, when a major product is reported with a sub-100% yield, the experimental yields of the alternate outcomes we generate may not be identically zero. Second, the recorded major product is only guaranteed to be the major product under those conditions; there are often other reaction conditions that could produce the alternate outcomes.

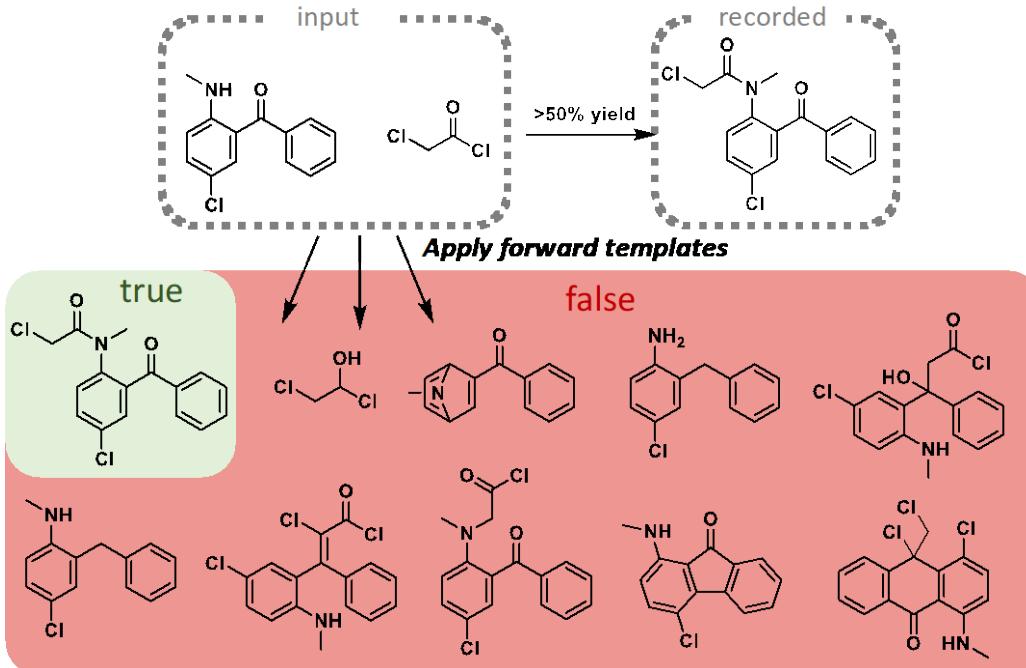


Fig. S4. Example of synthetic negative reaction data used to supplement known (true) reactions. A library of forward synthetic templates was applied to the reactants to generate candidate product structures—not all structurally reasonable—where zero or one may match the recorded major product.

The classifier is trained using the Adam optimizer and a binary crossentropy loss function. After training on 14,985,382 positive examples and 114,801,289 synthetic negative examples, the model achieves an area-under-the-receiver-operating-characteristic (AUC) of 0.996. The precision and recall using a classification threshold of 0.75 is 0.971 and 0.804. The precision and recall using a classification threshold of 0.5 is 0.940 and 0.904. The threshold used during expansion is a user-tunable setting, but defaults to 0.75. For very complex products, removing the fast filter will eliminate all false negatives but generate more false positives.

S1.7 Reaction condition recommendation

A neural network model is developed to predict reaction conditions in a sequential manner, including up to one catalyst, two solvents and two reagents and temperature; it is described in more detail elsewhere (39). The chemical identities of the options come from the most popular species found in the Reaxys database. As reagents and catalysts are frequently used in stoichiometric quantities, we do not filter recommendations by their cost.

Morgan fingerprints of the reactants and products are calculated using RDKit (folded into bit vector of length 16384). Reaction fingerprint is calculated as the difference between product and reactant fingerprints. Reaction and product fingerprints are concatenated and passed through two fully-connected layers to generate a dense representation of the fingerprints, which are further passed through two fully-connected layers to predict the catalyst as a one-hot vector. The one-hot vector for catalyst is concatenated with the dense representation of the fingerprints and passed through two fully-connected layers to predict a first solvent, and the process is repeated to predict a second solvent, a first and a second reagent. Finally, all chemical species representations are

concatenated to make prediction of the temperature. An example of this sequential prediction is shown in **Fig. S5**.

The model is trained on 10.3 million single-step, single-product reactions from Reaxys, and tested on 1.1 million reactions (random split). We analyzed the top-three and top-ten accuracies (percentage of reaction whose recorded conditions are within the top-three or top-ten predictions) for the overall combination, which are 50.1% and 57.3%. The model performance was compared to a baseline model which predicts the combination with the highest frequency for all reactions. The top-three and top-ten accuracies for the baseline model are only 4.7% and 5.7%. Qualitative evaluation of model predictions showed that when the predicted conditions are not exact matches of the true conditions, many of them are functionally similar to the recorded conditions. The mean absolute error for temperature prediction is 25.6 °C for all reactions and 19.4 °C for reactions whose chemical conditions are correctly predicted.

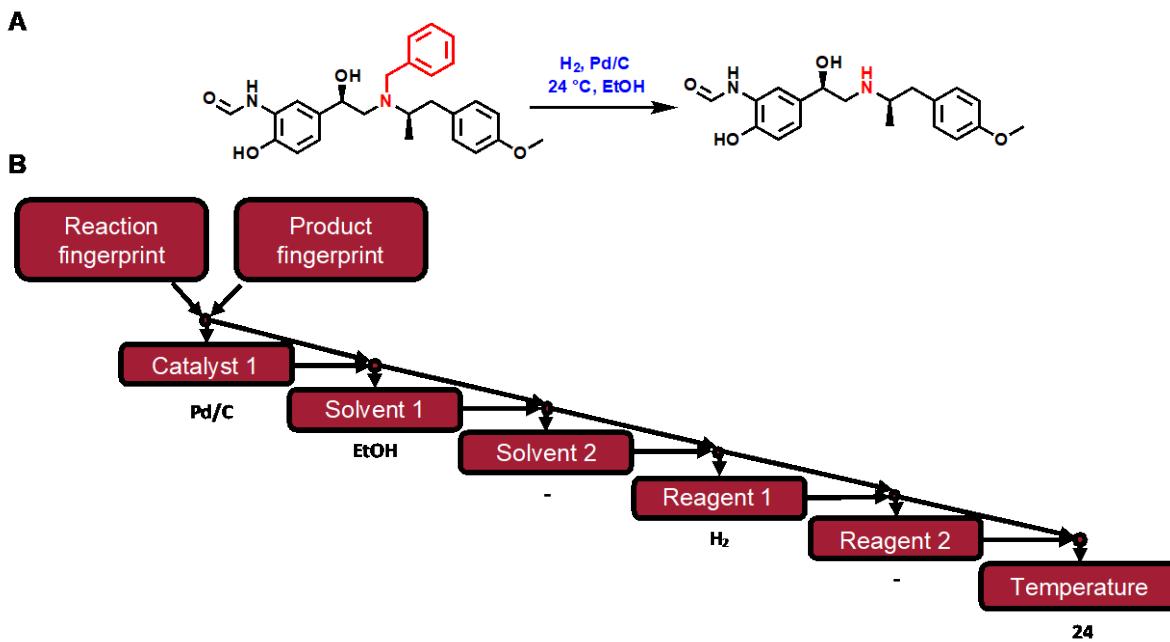


Fig. S5. (A) Example deprotection reaction and the model's top predicted conditions (blue). (B) Model schematic, showing the sequential prediction of reaction catalyst, solvent(s), reagent(s), and temperature based on the reaction and product fingerprints.

S1.8 Explicit prediction of reaction outcomes

A stricter evaluation than the “fast filter” occurs through the explicit prediction of reaction outcomes under the specific conditions (up to ten are tested) proposed by the context recommender. While this model achieves 85% top-1 accuracy on a USPTO dataset, having a false negative rate of 15% would remove a large number of viable pathways. For this reason, the webapp version of the program is designed to display pathways to users *prior* to this evaluation. A single click on each reaction triggers the context recommendation and forward prediction modules to run to find the best set of conditions (i.e., the ones that maximize the perceived probability that the intended product will be the major product). The forward prediction model is described in full in (37).

S1.9 User interface

The synthesis planning and reaction predictions tools are available through a command-line interface and a webapp. The webapp offers several modules, which are described below and on the tutorial page.

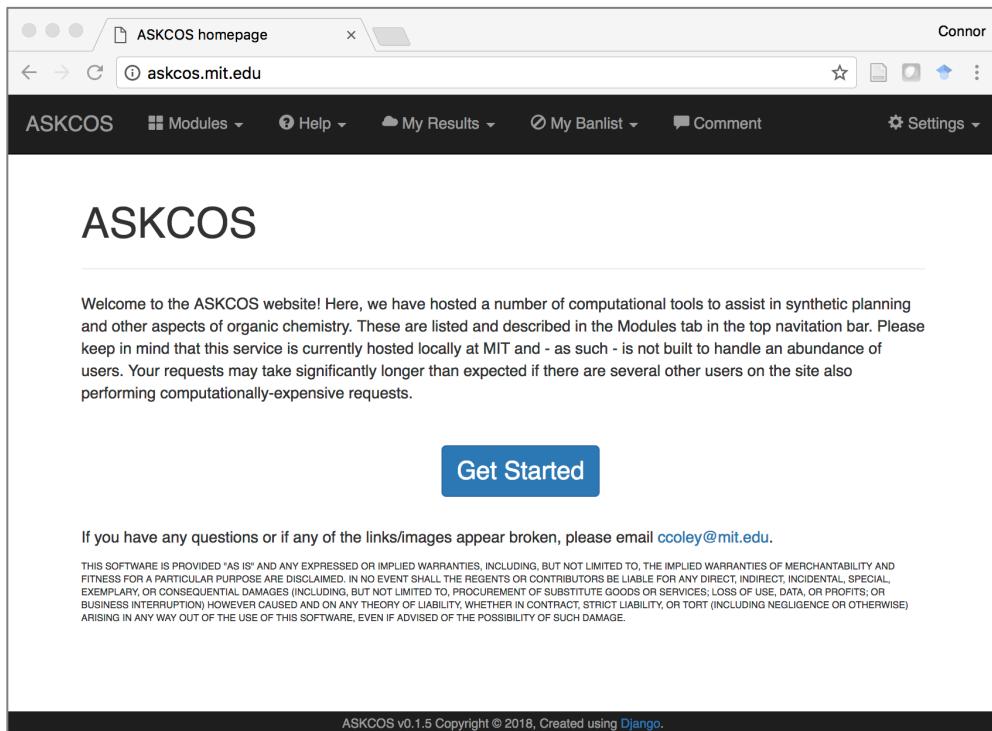


Fig. S6. Screenshot of the web application's home page currently hosted at askcos.mit.edu. All of the modules and tools described in this manuscript and used in the experimental validation are publicly available.

S1.9.1. One-step Retrosynthesis

This module allows you to look at suggested disconnections based solely on retrosynthetic templates. For each suggestion, you can click on the reactant SMILES to perform another search using that new compound as a target. You can also click on the forward arrow ($\rightarrow ?$) to open the Content Recommendation page. Clicking on the images of the suggested reactants will show the list of retrosynthetic transforms upon which the suggestion is based. You can click through to look at the details of the transform, including a list of the precedent reactions that led to that template.

S1.9.2. Tree Builder

This module allows you to run a time-limited tree expansion that will recursively apply retrosynthetic transforms in a parallelized, Monte-Carlo tree search until it reaches either (a) a buyable/obtainable chemical, or (b) the maximum depth. Once the expansion time has elapsed and (ideally) candidate pathways are found and displayed, you can interact with those suggestions by hovering over the reaction nodes to evaluate the reaction using the Forward Prediction tool or to see the transform rules upon which the suggestion is based. It is also possible to "hide all" of the tree suggestions that use a certain reaction if you know that reaction to be improbable. A more

extreme option is to "blacklist" that reaction, which will add it to a user-specific list of forbidden reactions. Any subsequent tree expansions you perform will exclude that reaction. The same can be done for chemicals if a particular intermediate should be avoided. You can access your blacklisted reactions or chemicals and activate/deactivate/delete them by clicking your username in the upper-right corner of the screen.

S1.9.3. Context Recommendation

This module allows you to use trained context recommendation models to get suggestions for reagents, catalysts, solvents, and temperature for any attempted transformation. It makes use of a neural network model to make a direct prediction of conditions.

S1.9.4. Forward Prediction

This module allows you to use the USPTO-trained template-free forward predictor to anticipate the outcomes of an arbitrary chemical reaction (37). While the models are trained to be condition-dependent, increasing the sensitivity of the prediction to precise conditions is a work-in-progress. The results of a forward enumeration can be exported to a .csv for offline processing or copying/pasting into Chemdraw.

S1.9.5. Reaction Evaluation

This module allows you to estimate the probability that a given reaction (defined by reactants and products) will succeed. The program will use the condition recommendation model to find a list of possible reaction conditions. Up to ten conditions will be attempted and the most promising one reported back to you. If you already have a good idea of the reaction conditions, it may be more productive to use the Forward Prediction tool and then see if the expected product is ranked highly.

S1.10 Additional Synthetic Route Suggestions

The complexity of synthetic routes we can plan computationally can exceed what makes sense to implement in a telescoped flow synthesis. Here, we include examples of small molecule targets that the program has never seen and the recommended routes to them. While only one route is shown per target, the program suggests a number of distinct, albeit often similar, options. The three molecules show in Fig. S7, Fig. S8, and Fig. S9 were taken from the 2018 ACS First-Time Disclosures of Clinical Candidates session (49). The routes were found in under one minute using eight parallel processes for expansion.

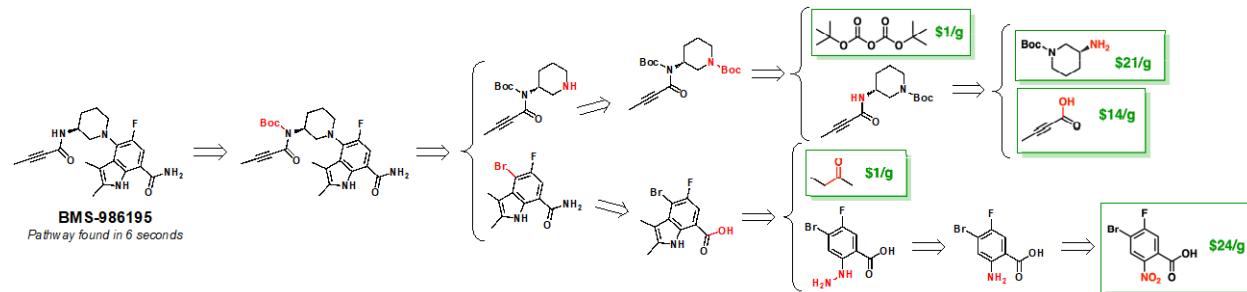


Fig. S7. Proposed synthetic route to BMS-986195, which the software has never seen.

Fig. S7 shows a suggested pathway for BMS-986195 obtained in six seconds by parallelizing the search over eight cores. The program recognizes that the tetrahedral center can be purchased with the chirality intact, rather than needing to build it. It also identifies a highly functionalized nitroarene that can be bought for \$24/g and builds the indole scaffold through a Fischer indole synthesis from the hydrazone prepared from the aniline, in turn prepared from that nitroarene.

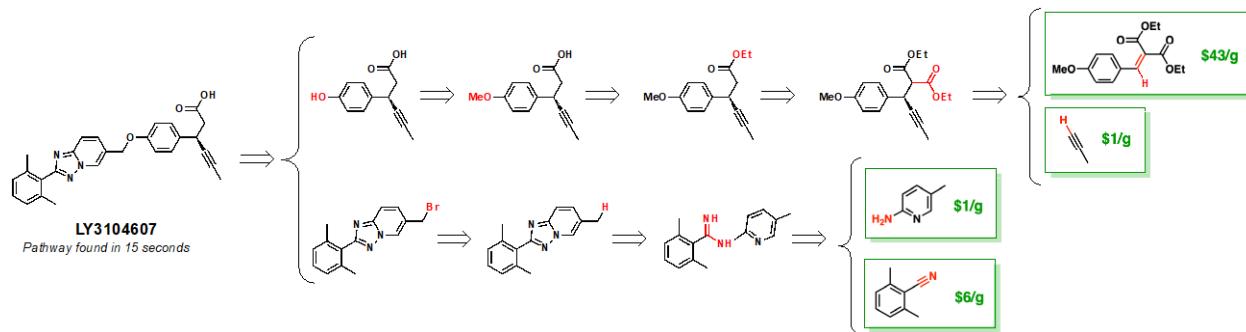


Fig. S8. Proposed synthetic route to LY3104607, which the software has never seen.

Fig. S8 shows an option for the preparation of LY3104607 found in 15 seconds. While the initial disconnection at the ether is obvious, the other steps are less so. In the bottom branch, we see a very cost effective way of preparing the heterocyclic fragment. In the top branch, to prepare the phenol, we see a suggestion that involves some protection/deprotection chemistry, a retro elimination, and a stereoselective installation of propyne.

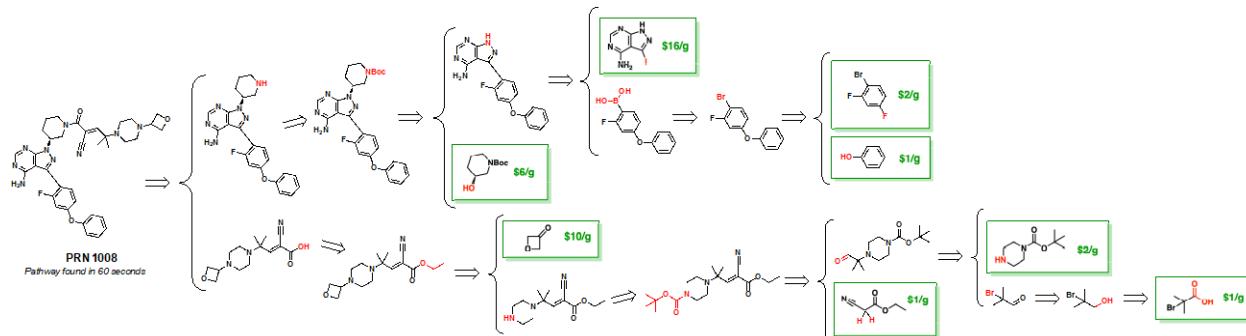


Fig. S9. Proposed synthetic route to PRN1008, which the software has never seen.

The third example in **Fig. S9** is a longer pathway, 13 total steps with a longest linear sequence of eight, found in roughly one minute. Here, most of the complexity (stereocenters, heterocycles) can be purchased, so the synthesis is linear save for the final amidation. The upper branch proposes an etherification (with ambiguous selectivity), then a Suzuki coupling, alkylation, and deprotection. The bottom branch proposes an alkylation, then an aldol condensation, deprotection, reductive amination, and ester hydrolysis.

As an additional set of examples, we can turn to a recent paper on molecular generation that employs an autoencoder to learn a continuous embedding of molecular structures (48). In the paper, the latent space is sampled to generate several structures of molecules that have never been synthesized, but resemble small molecules that might be found in the ZINC database. We hand-pick six structures that appear to require a variety of synthetic techniques, filtering out structures

that have an unusual density of chiral centers or uncommon motifs, and run them through the software.

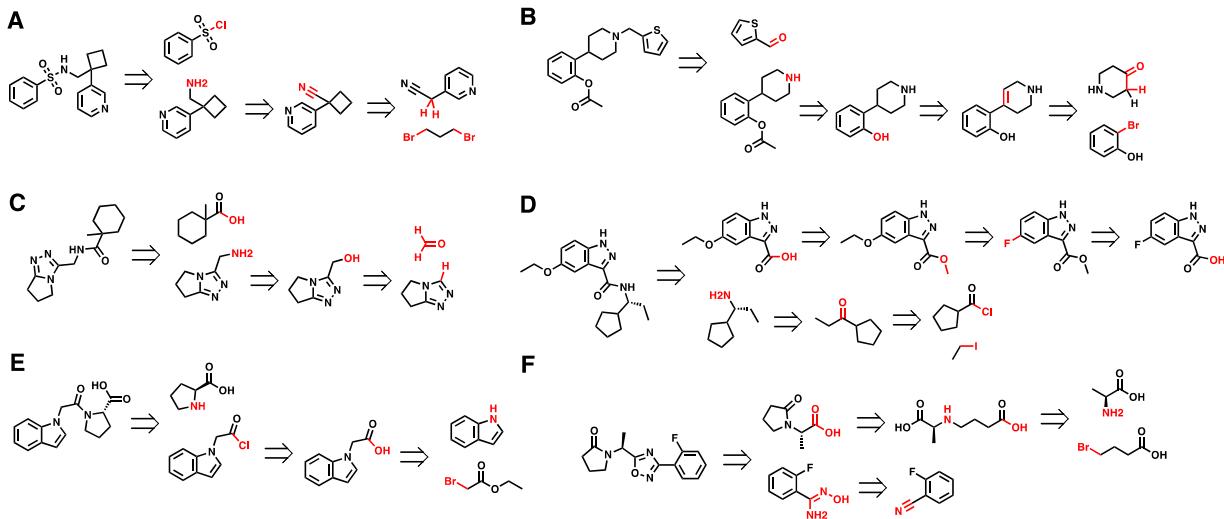


Fig. S10. Proposed syntheses for a number of targets generated de novo from Jin et al. (47). In this setting, the retrosynthetic program is able to quickly evaluate computationally proposed molecules in terms of their perceived synthesizability.

Fig. S10A shows a potential route to a sulfamide containing an uncommon cyclobutane motif. It correctly identifies that this motif can be prepared by the dialkylation of 3-pyridylacetonitrile with 1,3-dibromopropane. **Fig. S10B** proposes an interesting organometallic reaction as the first step, followed by a reduction, acetylation, and reductive amination. The path in **Fig. S10C** uses an amidation, amination with ammonia, and a hydroxymethylation with formaldehyde (potentially, a formylation followed by a reduction). **Fig. S10D** also begins the retrosynthesis with an amidation, later suggesting a stereoselective reductive amination and a condensation between ethyl iodide (a precursor to the organometallic reactive species) and an acid chloride. The chemistry used in **Fig. S10E** is straightforward, as is **Fig. S10F**, with the exception of a furazan-forming final step.

Nevertheless, there are cases we have encountered where the synthesis planning tool has significant room for improvement. These include,

- The synthesis of densely functionalized, small organic molecules. Certain compounds (e.g., those resembling pesticides) have many more halogenated motifs than are found in most product molecules found in the Reaxys database. These structures may not match any of the ca. 160,000 templates in the retrosynthetic template database and are unlikely to appear in our current database of buyable compounds.
- The synthesis of novel polycyclic heterocycles. Similar to the previous item, polycycles that have not been observed previously may not be able to be broken up by any of the retrosynthetic templates in our library. Aliphatic cycles are less of an issue, because intermolecular aliphatic bond forming reactions can generalize to intramolecular cyclizations. However, reactions forming aromatic polycycles require highly specific

reaction templates as many atoms must change properties to reflect their changed aromaticity.

- (iii) The synthesis of caged structures or other structures where selectivity is driven by sterics. The forward predictor model considers sterics to some degree, but uses a 2D graph representation of molecules in its prediction. Selectivity that is highly dependent on 3D information (e.g., tetrahedral centers, ring structures with implicit axial/equatorial side chains) cannot be predicted well. This is both a data and representational challenge: a relatively small number of reactions in our training set fall under this category, so the model is not able to learn these subtleties of synthesis.
- (iv) The synthesis of compounds where chiral resolution should be used as a purification strategy. When faced with a chiral target molecule, the current program *always* looks for a stereoselective synthesis. This requirement could be relaxed by allowing racemic syntheses for these targets. From a user's perspective, however, it would be beneficial to have the program understand when chiral resolutions would be appropriate so that this relaxation is not overused.

S2. Robotic Platform Description

S2.1 Case study on system complexity

The typical approach used for automating chemistry systems is through the utilization of valve banks and/or selector valves. In this way, the reagents can be automatically routed to different process units or modules. To limit complexity, these systems are normally targeted to a limited number of chemical synthesis or functionality. A common on/off valving strategy is depicted in **Fig. S11** to achieve a truly flexible design allowing for any two-step nonspecific chemistry to be synthesized with inline work-up or purification. In this case study, the connection of seven reagents to the process stream requires 49 valves and connection of two reaction steps with five reactor types and a purification step with four options requires 28 additional valves.

When changing the system from one chemistry to the next or optimizing conditions, the ability to switch from one reaction type and/or purification method results in excess equipment and valving needing to be installed on the system. The installation of extra reactors and separators on the system results in low system component utilization making long synthetic pathways prohibitively expensive to construct. Additionally, the flow paths at the inlet and outlet or required to bypass an unused process module results in dead-volume leading to possible contamination of the process stream and additional hold-up volume. The valves also represent potential failure points in the system due to chemical corrosion or high pressures causing leakage into the “isolated” process modules potentially damaging them or causing erroneous results. The valves are prone to clogging or sealing issues when particles are present in the flow or reagent stream. The number of fluidic streams connected to the process stream is hard coded into the flow path making it difficult scale the system to incorporate new functionality or process lengths.

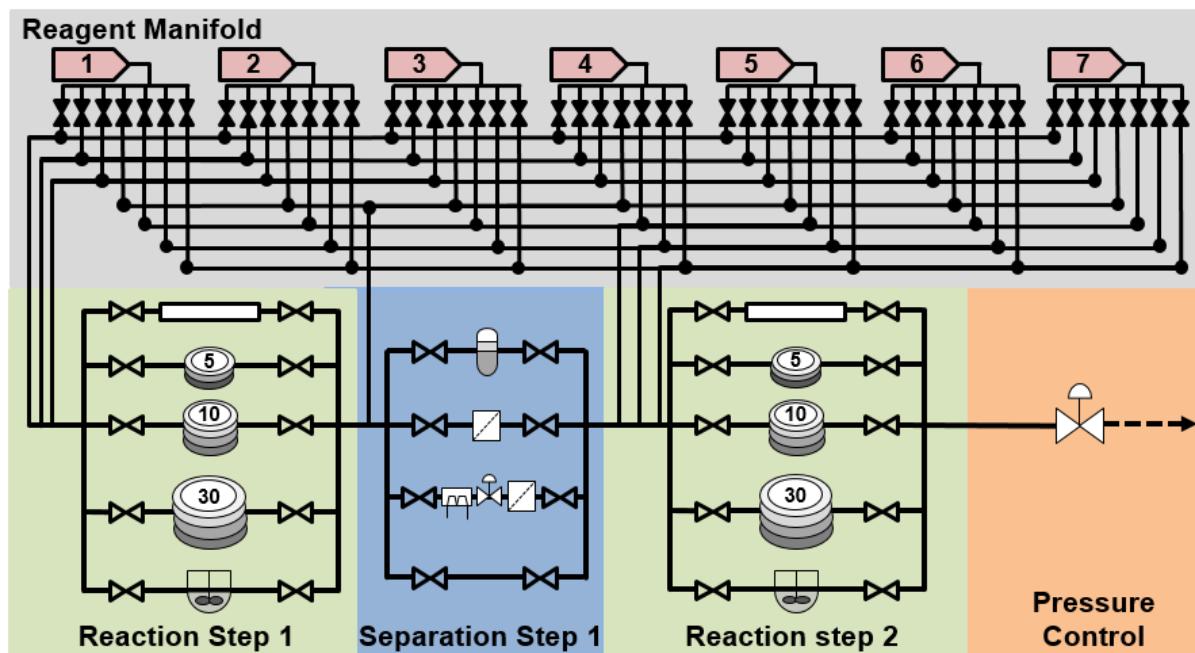


Fig. S11. Valving manifold requirement for a flexibly designed system where all reagents and process modules are accessible to enable automated change over between syntheses.

S2.2 Platform overview

Images of the platform and reactors/modules are shown in **Fig. S12**. The platforms five submodules are laid out on an optical breadboard for positioning and rigidity. The fluidic path flows from the reagents to the pumps. They then move through the reagent tree where the reagent streams are connected to the specific process module located on the process stack. The process modules are moved from their storage location to the process stack and configured into a continuous process stream by the six axis UR3 Universal Robot®.

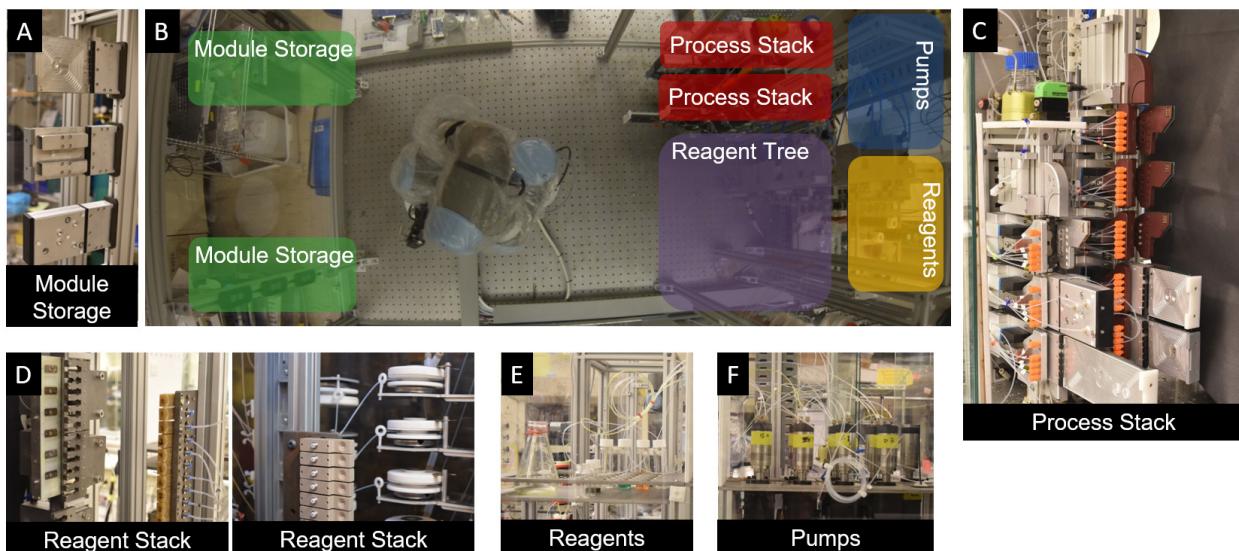


Fig. S12. Images of the chemical synthesis platform. (A) Module storage area; (B) Overhead view of the component layout for robotic manipulation; (C) Process stack with process modules installed; (D) Images of the reagent tree fitting storage and reagent stack; (E) System reagent storage area; (F) Pump storage area.

In order to run chemistry on the platform, the user interfaces with the platform through two main views shown in **Fig. S13** and **Fig. S14**. Loading of the reagents occurs on the one side of the machine where they are also connected to the pumps. The execution of the synthesis occurs from the front of the machine where recipes are loaded onto the control computer and carried out by the robotic platform.

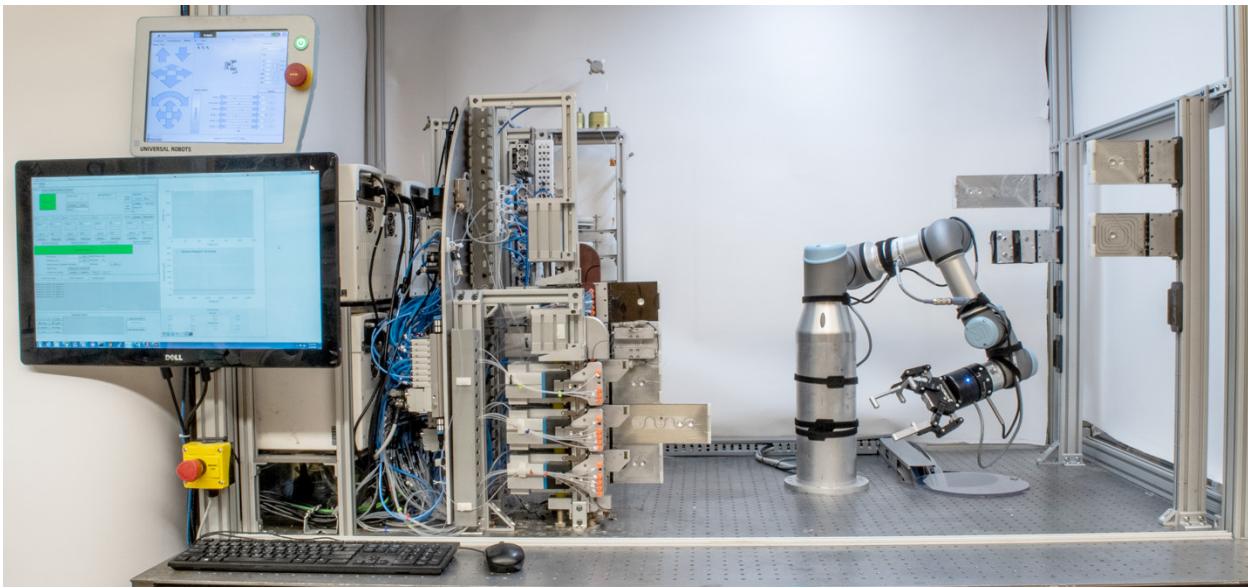


Fig. S13. Front view of the robotic platform.

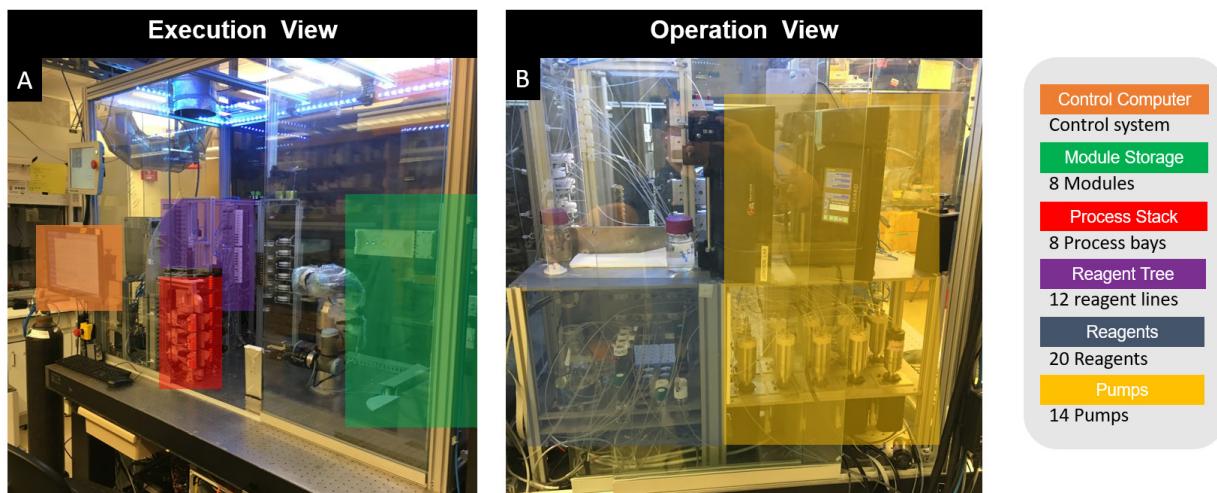


Fig. S14. (A) The view in which a user executes recipes on the robotic platform. (B) where chemicals are loaded onto the robotic platform.

S2.3 Process stack

The process stack has two towers containing 8 universal process bays (**Fig. S15**). The universal process bays allow the connection of the fluidic, electrical and pneumatic lines. The robotic manipulator aligns the process modules with the kinematic alignment features located in the universal process bays. The universal process bays allow for repeatable and precise installation and location of the process modules on the stack. The kinematic coupling is preloaded with magnets to allow for a passive holding force preventing the process modules from falling from the stack if there is a loss of power or air supply.

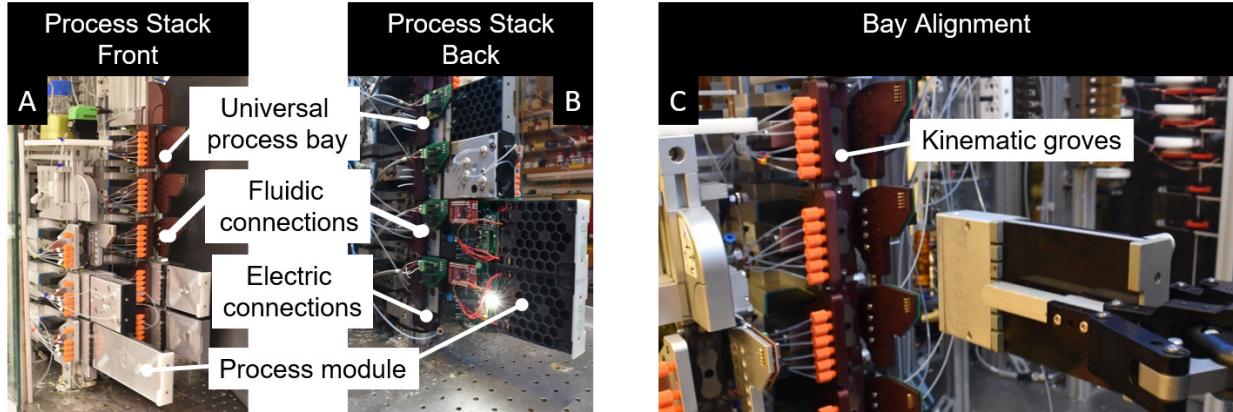


Fig. S15. (A) Photo of the front of the process stack. (B) Photo of the back of the process stack. (C) A process module being intalled by the robotic arm into the univeral process bay on the process stack.

S2.4 Reagent Tree

Robotic manipulation of fittings required the development of an actively loaded seal and incorporation of precision locating features for the pick/place operations. The screw features used for sealing ferrule fittings are difficult to automated. Therefore, the sealing force is applied using a linear pneumatic actuator. Alignment mechanisms are located between the fluidic fitting and robotic end effector, as well as the fluidic fitting and reagent manifold/storage. These alignment mechanisms allowed for precise location in the end effector and within the storage and reagent manifold. The robotic end effector is located to the fitting via a kinematic coupling with the grooves on the fittings and balls on the parallel jaws of end-effector. The male body of the fitting is located to the female reagent port via a kinematic coupling with the grooves on the fitting and spring loaded pins on the reagent manifold. The kinematic features are replicated in the stroage area for precise placement.

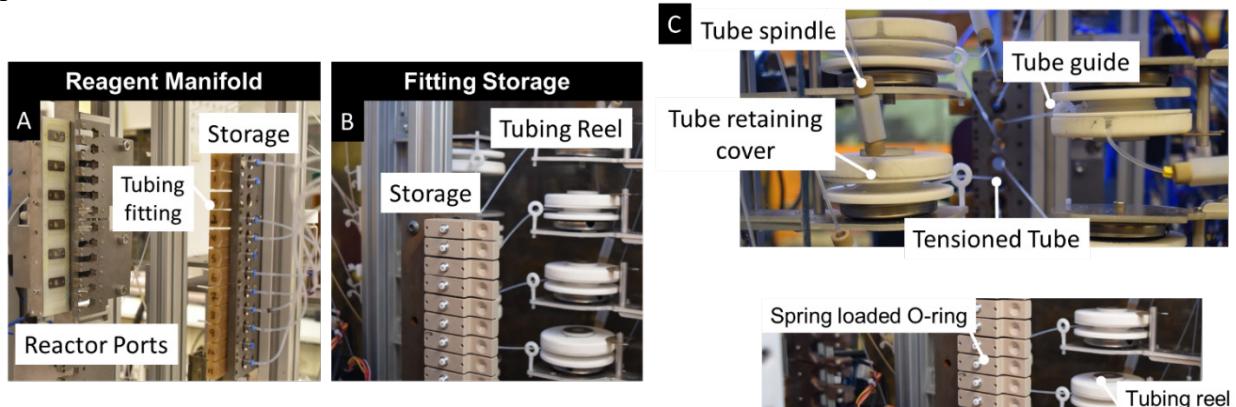


Fig. S16. (A) Photo of the reagent manifold. (B) Photo of the fitting storage and tubing reel systems. (C) Image of the tubing reel and components.

During manipulation of the fittings to different fluidic positions, the tubing and fittings must not become tangled. The tubing reels provide a constant tension to the tubing to prevent tangling allowing the robot to avoid paths that intersect the taught tubes. The tubing reels provide tension

to a tube using a power spring attached to a dual reeling systems and tube retaining cover. The tube reel and cover are rotated around a center spindle. The spindle acts to contain the entire assembly together allowing the rotation of the tubing reel. The tubing is wrapped around the tube reel allowing it to spool out during configuration without imparting torque to the tube.

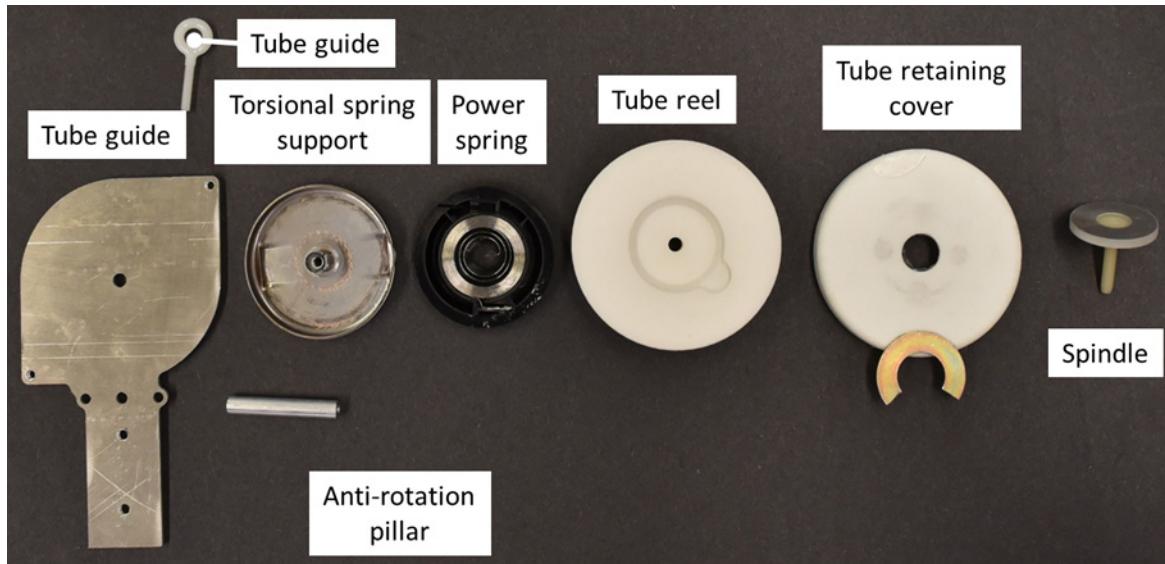


Fig. S17. Exploded component image of the constant tension tubing reel design.

S2.5 Reagent manifold design

The reagent manifold locates and seals the robotically configured fittings to the process stack. The fittings are magnetically coupled to a spring-loaded kinematic coupling (**Fig. S18A**) installed on the steel fitting jaws. The kinematic coupling allows high precision, reversible and automation friendly coupling of the fittings to the reagent ports. The sealing force is applied to the fittings using a pneumatic actuator. The process stream sealed to the process stack using an O-ring seal, **Fig. S18B**.

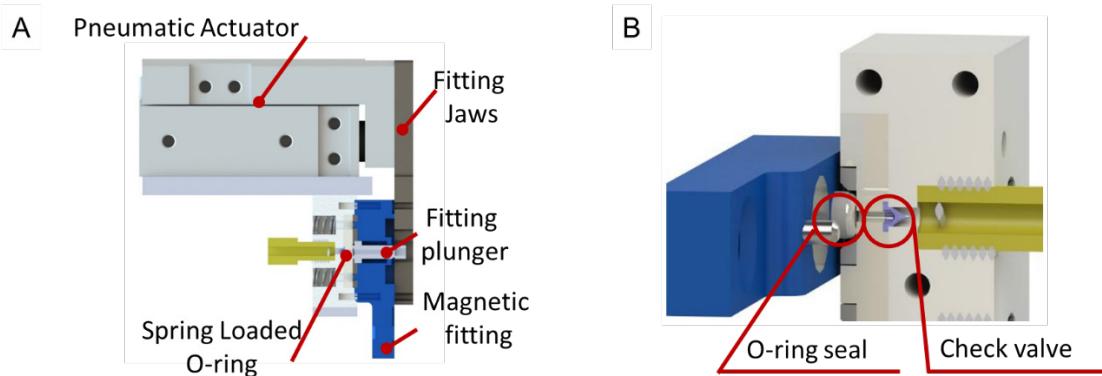


Fig. S18. (A) Rendering of the pneumatically actuated reagent manifold design. (B) CAB model of the O-ring sealing and check valve to connect the reagent manifold to the process stack.

The wetted components of the manifold are PTFE or Kalrez. The additional components are steel for the magnetic coupling. The steel components could have been PTFE coated to increase chemical compatibility. To prevent drainage of the fluidic lines between the reagent manifold and process stack, the manifold has inline duckbill check valve.

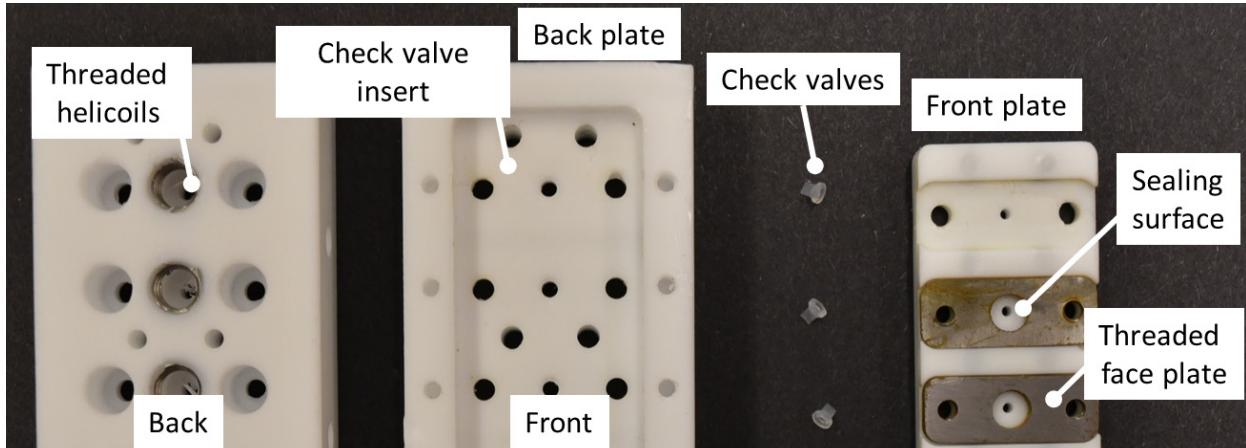


Fig. S19. Exploded view photo of the reagent manifold process pathway and sealing surfaces.

S2.6 Module Storage

The module storage brackets are designed to replicated the kinematic features located on the UPBs. They are made from hard coat anodized aluminum preventing peening of the coupling grooves. The steel plate is normal carbon steel allowing for the process modules to magnetically couple to the bracket. The magnetic preload force is designed for 1.8 kg of load. The brackets have rod in pin style locating features to allow them to be precisely place on the module storage rails.

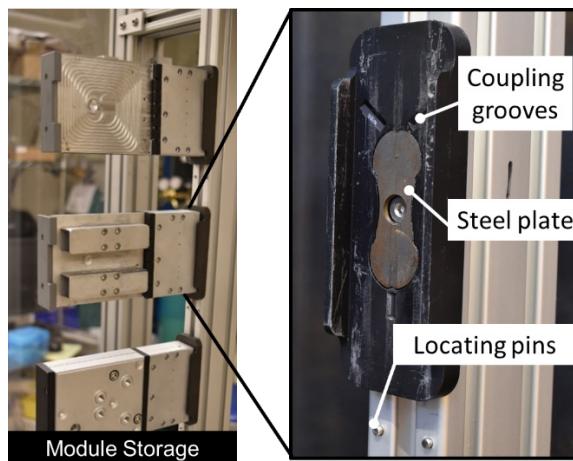


Fig. S20. Process module storage area and close up view of the kinematic mounting bracket used for robotic placement of modules.

S2.7 Process Modules

The reactors are designed to have both reusable and disposable components (**Fig. S21**). The heating components, and valve actuators built into the aluminum reactor insert shell. The

aluminum structure is designed to provide support to the thin film inserts and provide rapid heat transfer into the reaction area. The polymer areas provide thermal isolation for the electronics and integrated components. The disposable flow path enabling easy reactor cleaning and maintenance. All of these components are fit into a small form factor (7.6 cm x 2.5 cm x 18 cm). Fluid flows into the reactors from a port located on the bottom of the reactor orthogonal to the three additional reagent ports on the side. This will enable multiple reactors to be connected together for multi-step chemical synthesis. The process modules are located to the robotic gripper through the use of kinematic features located on the front and back of the reactor.

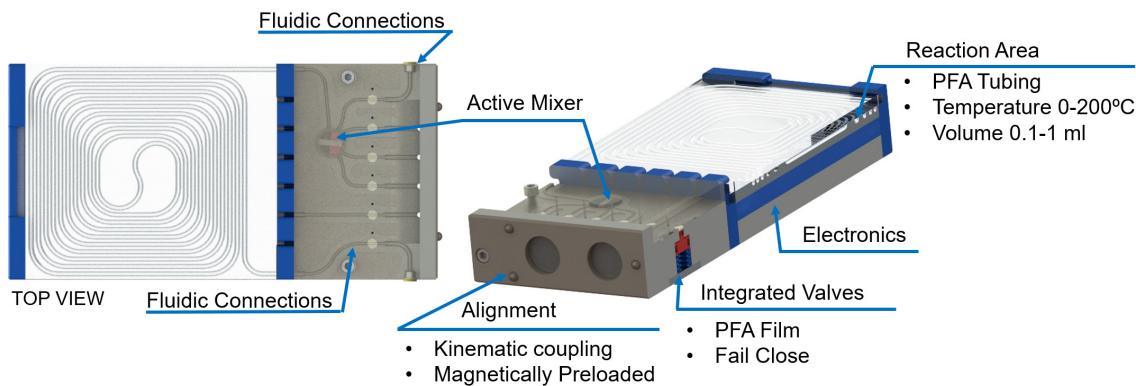


Fig. S21. Rendering of process modules showing the integrated mechanical and electrical components.

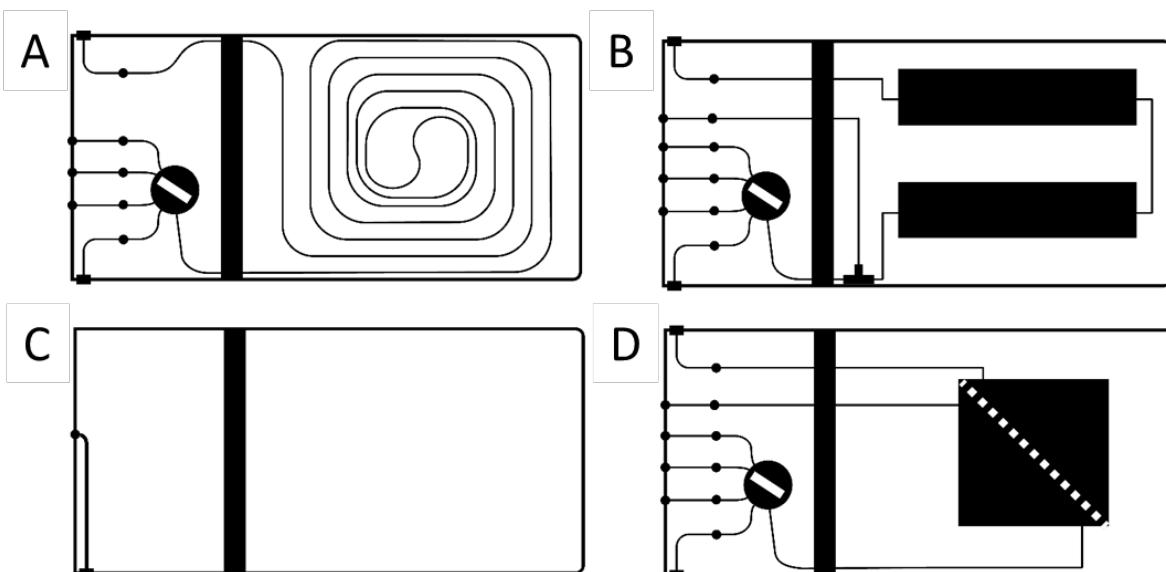


Fig. S22. Schematics of process modules showing the flow path through the film insert for the (A) reactor module, (B) packed bed, (D) membrane separator, and the flow path for the (C) outlet.

The four primary insert designs are shown in **Fig. S22**. Heated reactors (**Fig. S22A**) are designed to have fluidic connections at their inlet (bottom) and outlet (top) and up to three additional fluidic connections (left side) at ports 1, 2, and 3 on the right jaw of the UPB. These five lines can be individually opened/closed at inline valves (black circles) via pneumatic gas lines that connect from the left jaw of the UPB.

The packed bed reactor (**Fig. S22B**) is designed to house two columns in series, with an approximate total volume of 2 mL. The fluidic connections are similar, with the additional use of port 4 (and an additional pneumatically-actuated valve) to introduce a separate liquid or gas stream that combines with the primary process stream via a T-junction immediately before the first column.

The outlet module (**Fig. S22C**) does not use any pneumatic valving, but instead directly connects its fluidic inlet to port 4 as an outlet.

The separator module (**Fig. S22D**) contains a membrane and diaphragm developed by Zaiput Flow Technologies that has been adapted to provide the desired form factor. The inlet and other fluidic streams are similar to the heated reactors, with the additional use of port 4 as a secondary outlet stream for the retentate. The primary outlet stream, exiting the top of the module, contains the permeate. The insert can be reconfigured to reverse these two roles, depending on whether the retentate or permeate stream is used in downstream reaction steps.

Multiple heated reactors of several sizes were built to provide the flexibility in reaction volume and residence times required for the multistep syntheses performed in this study. **Table S2** lists the current set of modules available to the robot within module storage.

Table S2. Process modules available for automatic reconfiguration.

Module ID	Description
1	1 mL heated reactor
2	Membrane separator
3	3 mL heated reactor
4	3 mL heated reactor
5	1 mL heated reactor
6	Outlet
7	Outlet
8	0.5 mL heated reactor
9	Packed bed reactor

S2.8 Process Module Film inserts

The thermoplastic reactor inserts allow for the connection of the film reactor inserts to the tubular reaction tubing (**Fig. S23B**). The ability to vary the length of tubing allows the same reactor insert design to be used between all of the process modules (**Fig. S23A**). The film reactors are manufactured using a thermal impulse sealing and blow forming process where the channels are formed and tubular connections integrated. The film section has been designed to include four

reagent lines, and five valves. PFA was chosen as the material due to its high melting point and chemical compatibility.

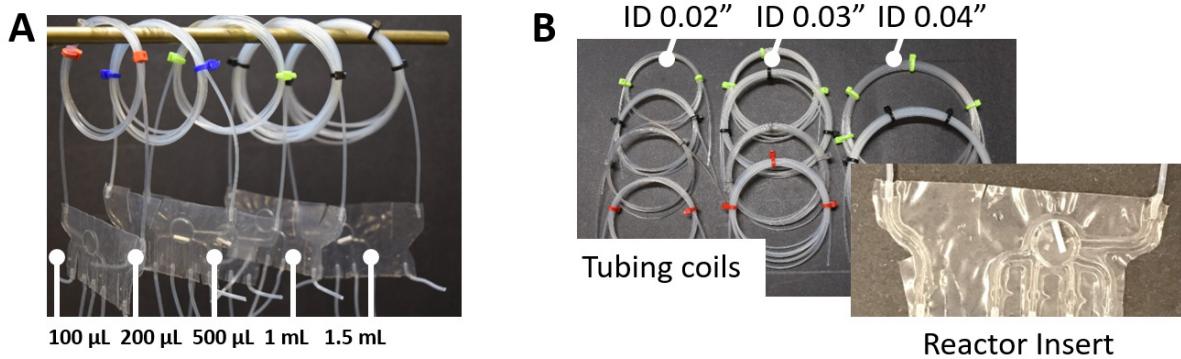


Fig. S23. (A) PFA film reactor inserts with different volumes. (B) Coiled tubing for different reactor volumes.

S2.9 Platform User Interface

The graphical user interface integrates the developed robotic toolbox, process module control, pressure, valving, pneumatics, and pumping into a single control interface (**Fig. S24**). The user interface, process drivers, and robotic control were all developed in Python 2.7. The GUI for control and graphing was developed using Tkinter. The process is automated through the running of chemical recipe files.

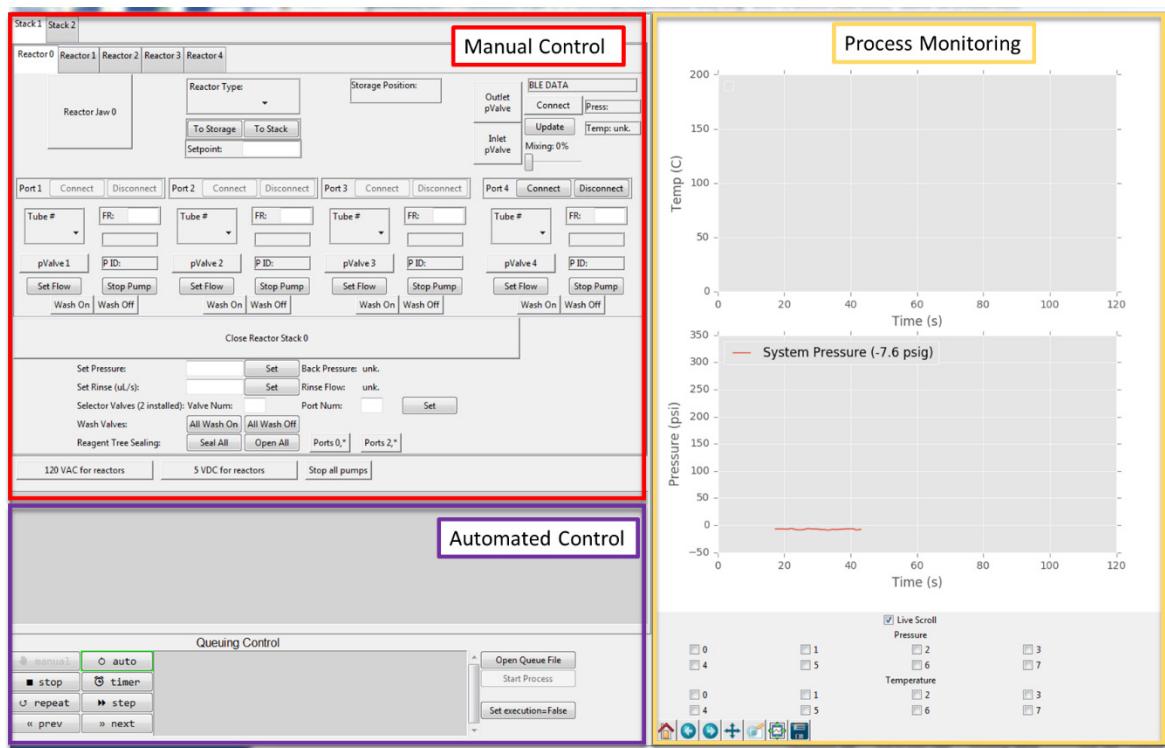


Fig. S24. Screen capture of the graphical user interface for running the chemical synthesis platform. The window is divided into three panels: manual control, automated control to execute a recipe file, and process monitoring of temperature and pressure signals.

The graphical user interface is broken down into manual control, autoamted control, and process monitoring. The manual control allows the user to operate any of the components hooked up the system including the robot. Manual control is organized by the process stacks and UBP. Process monition allows the user to see in real-time the temperature, and pressure of the sytem and individual reactors. When PMs are connected to the system there their process information is added to the graphs. The automated control is handled through a queuing system. The queueing system accepts CSV files. These files are pharsed and executed one line at a time.

S2.10 Control Strategy

S2.10.1. Control Interface Flow Chart

The control interface for the platform allows for both manual and autonomous control, **Fig. S25**. When the user opens the program they can select to operate the system by clicking all of the buttons (very tedious) or they can load a recipe which will handle all of the sequence for them.

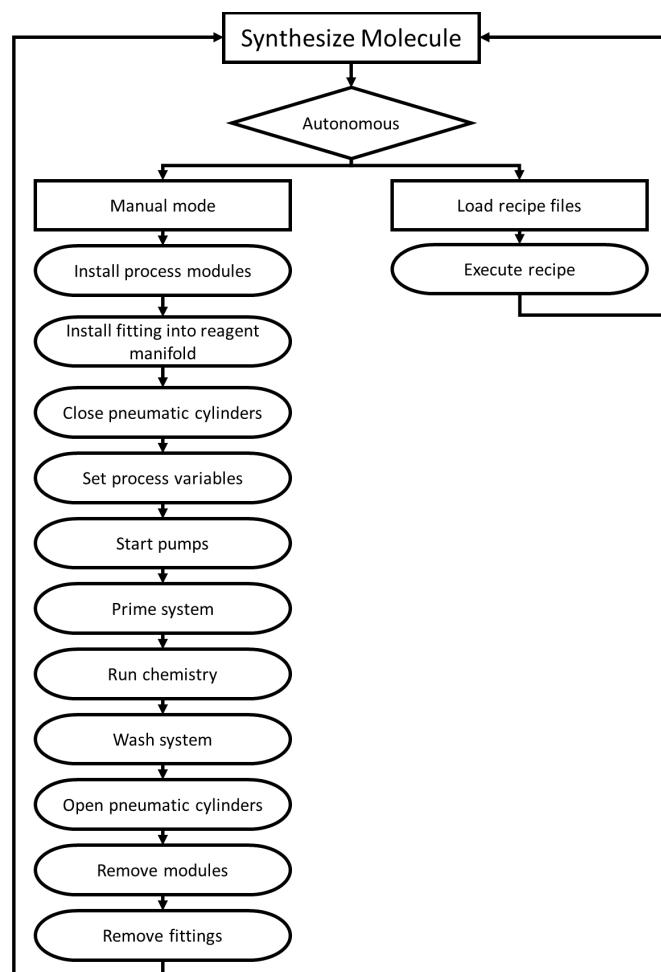


Fig. S25. Typical workflow when synthesizing a molecule in flow. Use of CRFs drastically decreases the manual burden on the chemist user.

While the molecule is being synthesized, the user gets real-time feedback on how the synthesis is performing and where potential problems might be occurring. The user can then take actions to mitigate the problems from occurring. The system is also being monitored locally for any unsafe conditions, low/high pressure or temperatures, where it will automatically shut the system down to prevent damage to the equipment or user. This integrated application also enables streamlining of scale-up by providing a user the information required to set-up and run larger scale systems.

S2.10.2. Robotic Arm Control

Movement of the process modules is accomplished with a UR3 robot placed on an 18" pedestal. The path planning, obstacle avoidance, and forward/reverse kinematics are done through the utilization of multiple open source and custom python libraries. Communication to the robot is accomplished through a websocket. Commands are sent to the robot utilizing the open source python-urx library. This library parses commands to URScript. To allow full utilization of the UR3 in the environment, MoveJ (linear joint-space movement) is preferred. For short linear movements, MoveL (linear configuration-space movement) can be utilized; however, it is more prone to failure due to angular discontinuities within configuration space. Commands were sent as 6-parameter poses (Cartesian coordinates of the end effector and three Euler angles to define its orientation). The robot itself is modeled as a series of cylinders, based off measurements of the maximum radii of each link (**Fig. S26**).

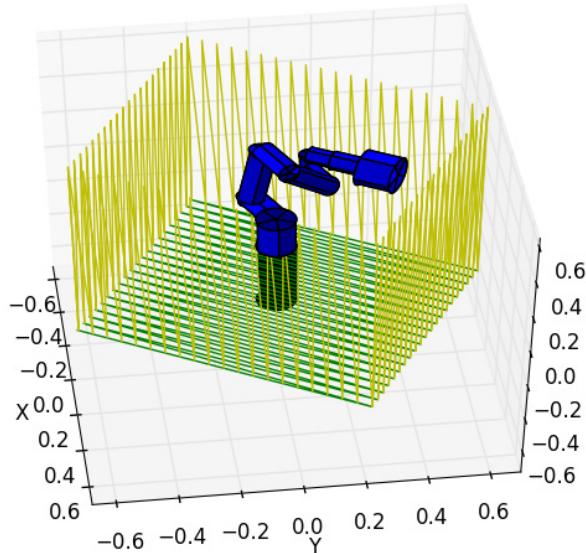


Fig. S26. Rendering of the robot for self-collision avoidance. Each arm in the 6-axis robot is modeled as a cylinder.

S2.10.3. Path Planning

It was found the utilization of waypoints to guide the path planning algorithm provided a faster route planning and more reliable operation as the robotic configurations could be more closely controlled (**Fig. S27**). This allowed the path search to be guided from one station to another, without the need to explicitly plan paths between those stations in advance.

Waypoint locations were specified within a world frame and converted automatically to the robot frame. This allowed for easy and intuitive specification of relevant locations, as well as enabling the path planning to avoid contact with its surroundings within the fumehood.

This world frame was defined by a rectangular set of 2-D wall vectors, constituting the boundaries of the robot's admissible workspace. Detection of wall collisions was determined in 2-D based on these wall vectors. The definition of this world frame, as well as the establishment of bounds of admissibility within it, enabled the path planning to operate easily within the robotic environment.

Locations of storage bays, process bays, and tubing handles were defined in the world frame. Path planning between these locations and a set of waypoints was done through linear interpolation in joint angle space, with constraints as defined in the Collision Avoidance section below. Because the rotational range of each joint is 4π , multiple interpolations (in positive and negative directions) were attempted to account for 2π symmetry in any individual joint angle. A rapidly-exploring random tree (RRT) search was implemented as a back up search strategy if this joint-angle interpolation approach failed, but was never necessary in practice. All path planning functions were wrapped in a memoizing function to save computational time during subsequent runs.

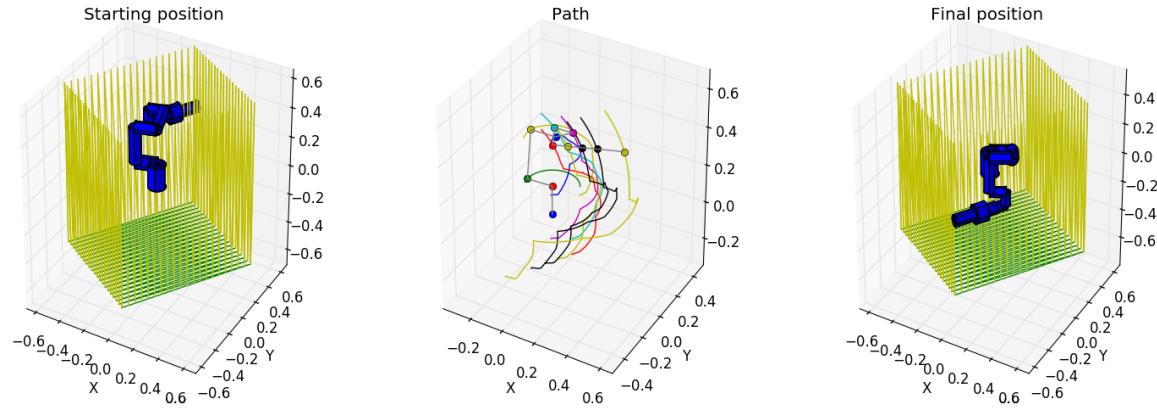


Fig. S27. Example visualization of the path planning algorithm to move the end effector from a specific starting position to a specific final position confined within the fume hood environment.

S2.10.4. Collision Avoidance

Self-collision avoidance is accomplished through both geometric constraints and a linkage intersection algorithm. Our self-collision detection algorithm evaluates a joint-space pose for the robot (using forward kinematics to project it to configuration space) and determines if there is a collision (45).

Geometric constraints prevent the robot end effector from attempting to enter the inner inadmissible region within its workspace (less than 0.2 meters horizontally from its origin, as defined by the UR3's on-board control software) or colliding with its elevated base.

Linkage intersection in configuration-space is based in detecting the intersection of two lines and line closest point calculation. A simple geometric model of the robot utilizing the angles determines the smallest distance of approach between two links that are kinematically capable of collision; if

this distance is less than the summed radii of those links, they are determined to be in collision. The effective end effector length is dynamically updated as the robot acquires and deposits reactors, to allow an accurate real-time model of the robot's effective workspace.

Ground avoidance is a geometric constraint placed in linkage number 3. The existence of the inner inadmissible region prevents the robot from operating directly over itself, thus obviating the need for ceiling detection.

S2.11 General Platform Remarks

Back pressure regulators (BPRs) were purchased from Zaiput Flow Technologies. The disposable reactor inserts were constructed from high-purity perfluoroalkoxy (HPFA) film (McMaster-Carr 84955K26) and HPFA tubing with a 1/16" outer diameter (OD) and 0.03" or 0.04" inner diameter (ID): 0.5 mL reactor consists of 1 m 0.03" tubing, 1.0 mL reactor consists of 1 m of 0.04" tubing and 3.0 mL consisting of 4 m of 0.04" tubing. For details on preparation of the disposable inserts see Section S2.9 (above). In all cases reactor tubing is pressed into channels in the aluminum reactor plate and covered with FEP coated glass. While the blow formed portion of the insert that creates the linkage between fluidic streams and contains the mixing chamber is installed within an aluminum enclosure containing the onboard fail-closed valves. PEEK fittings were purchased from IDEX Health & Science Technologies. Two methods of pump are used on the platform. For moderate to high flow rates (generally, not under 100 microliters per minute), the rinse line and reagent lines 0,10 and 0,11 used Vici Valco Milligat M6 pumps (CP2-4111-DHP); for low and moderate flow rates (generally, not exceeding 100 microliters per minute), reagent lines 0,2 through 0,9 used Syrris Asia pumps (2200292), unless otherwise noted. Gas was delivered by directly connecting a mass flow controller (Sierra, C100L-HP) to reagent line 0,2, equipped with an inline check valve.

S2.12 Details of Specific Platform Configurations

Fig. S28A shows the flow path of the robotic platform, with additional lines to stock solutions omitted for clarity. The nine unique process configurations required for the chemistry described in this manuscript are shown below in **Fig. S28B**. The primary process stack 0 contains 5 UPBs (indexed as 0,0 through 0,4). Fluid moves from the bottom bay 0,4 up through the stack.

Each set of jaws contains six 200 psig nitrogen lines (left side of jaw) to control the on-board pneumatic valves: one for the inlet, one for the outlet, and one for each of up to four process streams. These are individually addressable through Festo solenoid valve manifolds to only open the valves as required (shown as green tubes in **Fig. S28B**).

Fluidic lines for reagent addition are plumbed from the process stack to the process tree, which, for the chemistries reported herein, only requires 12 unique positions. These are indexed 2,0 through 2,5 and 0,0 through 0,5.

The reagent tree manifold also contains 12 positions (indexed 0,0 through 0,11), with custom fittings (brown rectangles in **Fig. S28AB**) that can be grasped by the robotic arm. All lines are coiled in the tubing reel as described above, to provide constant tension and prevent tangling.

Lines 0,0 and 0,1 are connected to two Zaiput backpressure regulators that are tethered together to insure matched pressures for inline separations. The back pressure is set automatically, according

to the CRF, through an Alicat digital pressure controller. Line 0,0 is connected to a Gilson Fraction Collector, while line 0,1 is a dedicated waste line.

Lines 0,2 through 0,9 are plumbed to four dual-channel Syrris Asia syringe pumps. Lines 0,3 and 0,5 are further plumbed to two 24-way Rheodyne selector valves. Lines 0,10 and 0,11 are plumbed to two Vici M6 Milligat piston pumps. An additional Milligat pump is used for rinsing the process stack, bypassing the reagent tree and directly connected to port 3 on bay 0,4.

Fig. S28B shows the details of each of the nine process configurations for each synthesis. Stock solution labels are cross-referenced with the schemes in the manuscript and in the detailed operator instructions below. Process modules are labeled by ID according to **Table S2**.

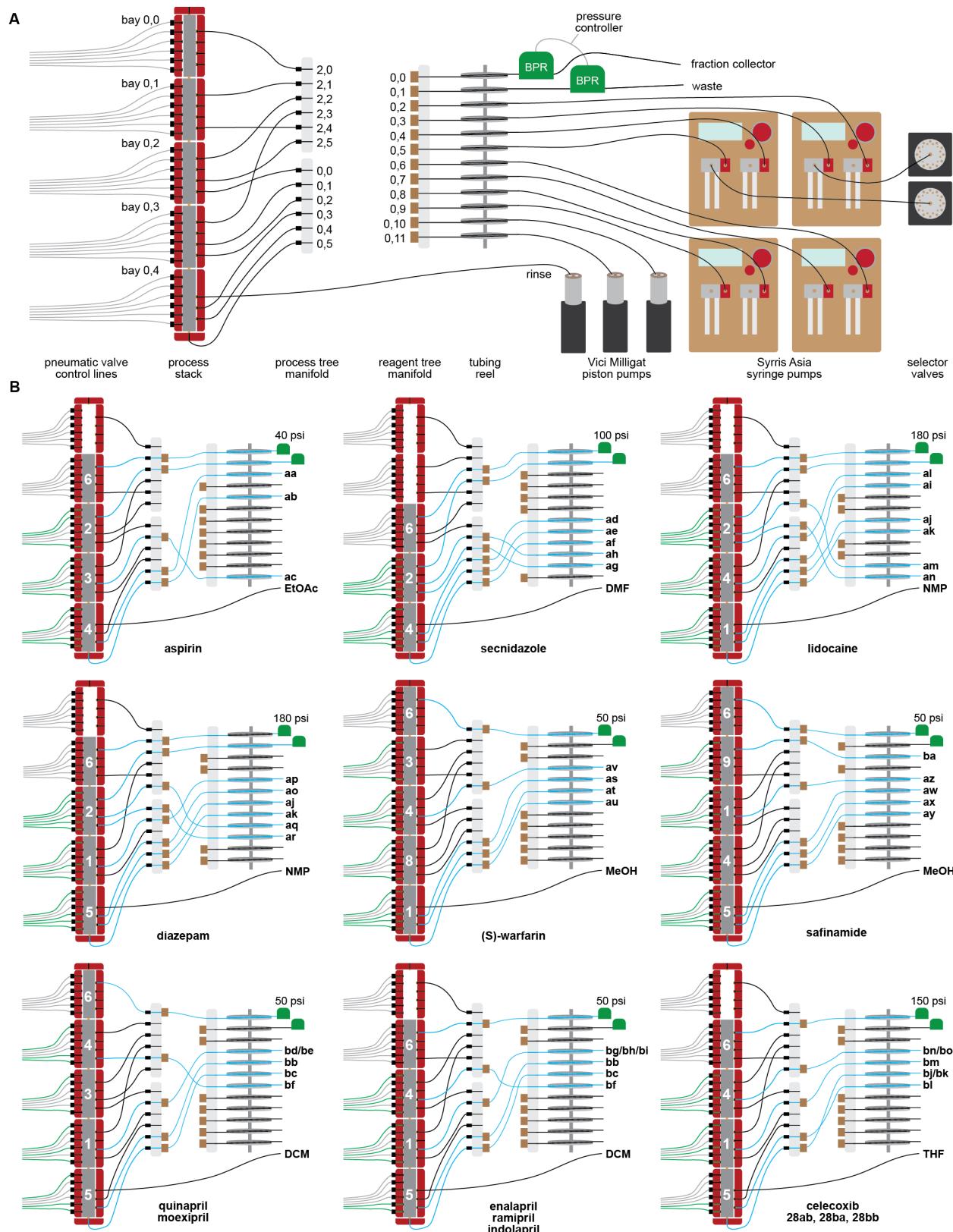


Fig. S28. (A) General schematic of the robotic platforms flow path. (B) The flow path for each of the nine unique process configurations used for chemistry in this study.

S2.13 Discussion of Platform Capabilities

The robotic flow platform was designed with modularity and extensibility in mind to accommodate the current suite of unit operations (i.e., heated reactors of various sizes, membrane separators, packed bed reactors) in addition to ones under development (e.g., continuous stirred-tank reactors). In principle, it can be applied to execute any chemical recipe file that uses these unit operations in any order, with any combination of reagent additions. Other limits related to pressure, temperature, and flow rates are as described earlier. The process stack used for the chemistries in this study has five UPBs for up to five unit operations; the reagent tree has twelve ports for up to twelve reagent lines.

In practice, the primary obstacles to implementation of chemical recipes are the propensity to form solids and the timescale required for reactions to occur. These are inherent challenges for flow chemistry performed in plug-flow reactors, although the former would be alleviated by the use of continuous stirred-tank reactors or other unit operations that can tolerate solid particles. The following section describes a case study where we could not successfully demonstrate the translation of an AI-planned route to the robotic platform due to the need for slow reagent addition to a cooled reactor and long residence times.

S2.13.1. Bezafibrate: A Case Study

Bezafibrate is a drug used to lower LDL cholesterol that we had selected as an example to include in this study. Among the synthetic options for bezafibrate proposed by the software were two routes using a simple amide coupling as their first step and either a standard etherification or a Bargellini reaction for their second. The Bargellini reaction does not come immediately to mind when looking at this target, and so we thought it would be interesting to pursue as an example. It incorporates chloroform and acetone in the presence of strong base to form a tricholorobutanol intermediate that can install the 2-methylpropanoic acid moiety on the free hydroxyl group of the amide product. This was of interest to us not only due to the proposed route's originality, but the possibilities for a convergent synthesis in the automated platform. The amide coupling was manually chosen to utilize tyramine and 4-chlorobenzyl chloride as that combination will react more readily than tyramine and 4-chlorobenzoic acid, though both were proposed as options.

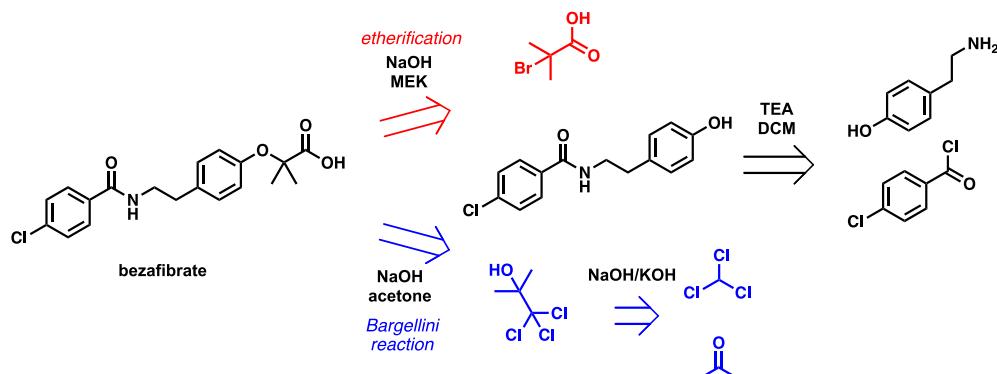


Fig. S29. Two proposed routes for the synthesis of bezafibrate employing (red) an etherification or (blue) a Bargellini reaction for the installation of a 2-methylpropanoic acid moiety. MEK: methyl ethyl ketone; TEA: triethylamine; DCM: dichloromethane.

For the Bargellini reaction, we found that the mixture of chloroform and acetone in base produced a significant exotherm in the production of the tricholorobutanol intermediate, resulting in poor overall yields from the amide. Conventional Bargellini reactions utilize the slow addition of powdered inorganic bases such as sodium or potassium hydroxide and sub-ambient reaction temperatures to mitigate generated heat, typically reacting for longer periods of time (>5 hours). We were able to perform the proposed synthesis under batch conditions in a cooled reaction vessel with deliberate slow addition of base. Reaction times of 16 hours produced good yields while shorter times gave incomplete reactions. Unfortunately, this proved to be a challenge for our platform as we did not have the capacity for cooled reactions and longer residence times present practical challenges due to the very low flow rates required.

Translation of the reaction to flow was further hindered by the poor solubility of inorganic bases in acetone and chloroform. Aqueous base proved difficult as the requisite chloroform was unable to participate in the reaction due to miscibility issues. Sodium chloride formation was also observed in batch which raised concerns over clogging in flow. Tetrabutylammonium hydroxide was successfully utilized in batch screenings as a remedy to this with comparable yields to sodium hydroxide (~75%). Fume hood flow experiments with cooled reactors were attempted with the purified amide and commercial tricholorobutanol; however, even residence times of two hours gave incomplete conversion and yields of less than 25%.

A one-pot batch synthesis was successfully developed and performed by hand, but could not be translated into flow. To a flame dried 50 mL round bottom flask with magnetic stir bar was added 20.0 mL (273 mmol) dry acetone and 2.65 mL (32.8 mmol) dry chloroform. The flask was set to stir and cooled to 0 °C in an ice water bath for 10 min. Powered sodium hydroxide (583 mg, 14.6 mmol) was added to the solution and allowed to stir for 1.5 hours at 0 °C. To the reaction mixture was added 4-(2-aminoethyl)phenol (500 mg, 3.64 mmol) followed by 4-chlorobenzyl chloride (638 mg, 3.65 mmol) added dropwise. The reaction was allowed to stir for 16 hours warming to room temperature. The reaction mixture was washed with aqueous 1 M HCl, brine and water. The organic layers were combined and the solvent was removed under vacuum. The product was recrystallized from water to give bezafibrate as an off-white solid in 76% yield.

S3. Chemistry

S3.1 General Remarks

Unless otherwise indicated, all commercially available reagents and solvents were used directly from the supplier without further purification. Mass spectrometry experiments were performed on Agilent Technologies 6545 Q-TOF LC/MS or JEOL AccuTOF DART at the Department of Chemistry Instrumentation Center at MIT or Waters QTOF Premier instrument at the Chemical Instrument Center at Boston University. Preparatory HPLC purification was performed with Agilent 1200 with Eclipse XDB-C18 Column 7 μ m, 21.2x250mm. Normal phase HPLC analyses were performed on an Agilent 1290 Infinity II using Agilent Poroshell 120 EC-CN and Daicel Chiralpak AD-H columns. Reverse phase HPLC analyses were performed using Agilent 1260 Infinity II HPLC with Agilent Poroshell 120 SB-C18 2.7 μ m, 2.1x50 mm or Agilent 1290 Infinity HPLC with Agilent Eclipse Plus C18 3.5 μ m, 4.6 x 100 mm. Specific optical rotation was measured with a Anton Paar modular circular polarimeter MCP 500. NMR spectra were collected on Bruker Ascend 600 MHz AVANCE NEO NMR spectrometer with CryoProbe Prodigy QCI-F probe, Bruker Ascend 500 MHz AVANCE NEO NMR spectrometer with BBFO probe or Bruker AVANCE III HD 400 MHz NMR with CryoProbe Prodigy BBO. Column chromatography of compounds synthesized from the platform was performed with Biotage Isolera III with commercially available ISCO RediSep columns.

S3.2 Stock Solution Summary Table

Table S3. Chemists/User guide for preparing stock solutions used in this study.

Label	Chemistries used for	Flow (uL/min)	Solvent	Vol. (mL)	Component 1				
					Nickname	ID	Mass (g)	Conc. (M)	
aa	aspirin	147	EtOAc	50	salicylic acid H ₂ SO ₄	2	8.29 0.55	1.2 0.12	
ab	aspirin	53	neat	20	acetic anhydride	3	21.6	10.6	
ac	aspirin	200	H ₂ O	100	NaCl	-	Brine diluted 1:1 by volume with H ₂ O		
ad	secnidazole	100	DMF	40	imidazole	5	3.81	0.75	
ae	secnidazole	100	DMF	40	epoxide	6	5.23	2.25	
af	secnidazole	100	DMF	40	triethylamine	-	9.11	2.25	
ag	secnidazole	200	H ₂ O	400	HCl	-	14.6	1.0	
ah	secnidazole	100	diethyl ether	40	ether	-	28.5	9.7	
ai	lidocaine	70	NMP	40	aniline	10	6.93	1.43	
aj	lidocaine, diazepam	9.0, 7.8	neat	10	bromoacetyl chloride	11	18.9	12	
ak	lidocaine, diazepam	30 25	NMP	20	NMP	-	20.6	10.4	
al	lidocaine	230	1:1 H ₂ O:MeOH	120	diethylamine KOH	8	11.4 2.9	1.3 0.43	
am	lidocaine	400	H ₂ O	400	NaCl NH ₄ Cl	-	10% bv weight 10% bv weight		
an	lidocaine	600	hexanes	600	hexanes	-	393	7.6	
ao	diazepam	75	NMP	40	benzophenone	14	9.83	1.0	
ap	diazepam	200	1:3 H ₂ O:MeOH	110	ammonia	-	19.8	2.7	
aq	diazepam	250	H ₂ O	130	NaCl	-	25% by weight		
ar	diazepam	500	EtOAc	260	ethyl acetate	-	235	10.2	
as	(S)-warfarin	9.0	DCE	10	acetocinnamone	16	8.46	5.6	

at	(S)-warfarin	7.5	MeOH	10	(S,S)-DPEN	-	4.67	2.2
au	(S)-warfarin	12	AcOH	10	acetic acid-	-	10.5	16.7
av	(S)-warfarin	21	DMF	20	hydroxycoumarin	17	10.1	3.1
aw	Safinamide	66.7	MeOH	50	Aldehyde Sodium iodide	22	3.05 0.37	0.5 0.05
ax	Safinamide	33.3	MeOH	30	chloride	23	6.48	1.5
ay	Safinamide	33.3	MeOH	30	DBU		6.84	1.5
az	Safinamide	305.6	MeOH	250	alaninamide HCl Triethylamine	20	7.49 6.07	0.24 0.24
ba	Safinamide	6 SCCM	H ₂	tank	hydrogen	-	-	-
bb	quinapril, moexipril, enalapril, ramipril, indolapril	40	8:1 DCM:dioxane	20	Alanine Coupling Partner HCl	27	2.79 36	0.5 0.1
bc	quinapril, moexipril, enalapril, ramipril, indolapril	40	DCM	25	CDI		2.03	0.5
bd	quinapril	10	DCM	8	TIC	26a	3.73	2.0
be	moexipril	10	DCM	8	TIC(OMe) ₂	26b	4.49	2.0
bf	quinapril, moexipril, enalapril, ramipril, indolapril	76.5	neat	50	TFA		74.5	13
bg	enalapril	10	DCM	8	Pro-O'Bu	26c	2.74	2.0
bh	ramipril	10	DCM	8	pyrrole-O'Bu	26d	3.38	2.0
bi	indolapril	10	DCM	8	indole-O'Bu	26e	3.60	2.0
bj	celecoxib, 28ab	16.3	THF	40	acetophenone	32a	7.51	1.75
bk	28ba, 28bb	16.3	THF	40	acetophenone	32b	11.2	1.75
bl	celecoxib, 28ab, 28bb, 28ba	40.7	THF	60	NaO'Bu		5.38	0.7
bm	celecoxib, 28ab, 28bb, 28ba	20	THF	50	ethyl trifluoroacetate	31	12.1	1.7
bn	celecoxib, 28ba	103.8	H ₂ O	150	hydrazone	29a	11.8	0.33
bo	28ab, 28bb	103.8	H ₂ O	150	hydrazone	29b	7.84	0.33

S3.3 Experimental Results

S3.3.1. Aspirin, 1

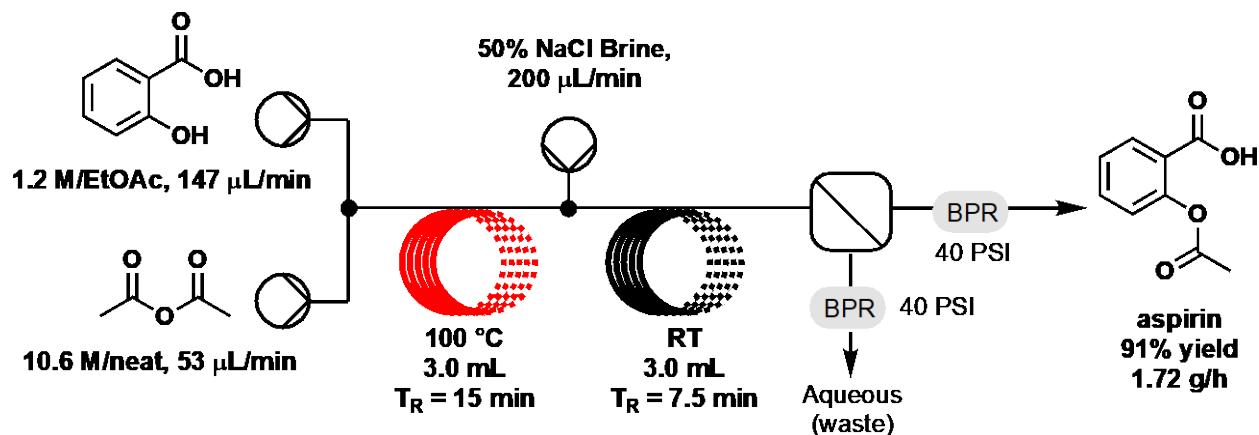


Fig. S30. Scheme for the flow synthesis of aspirin.

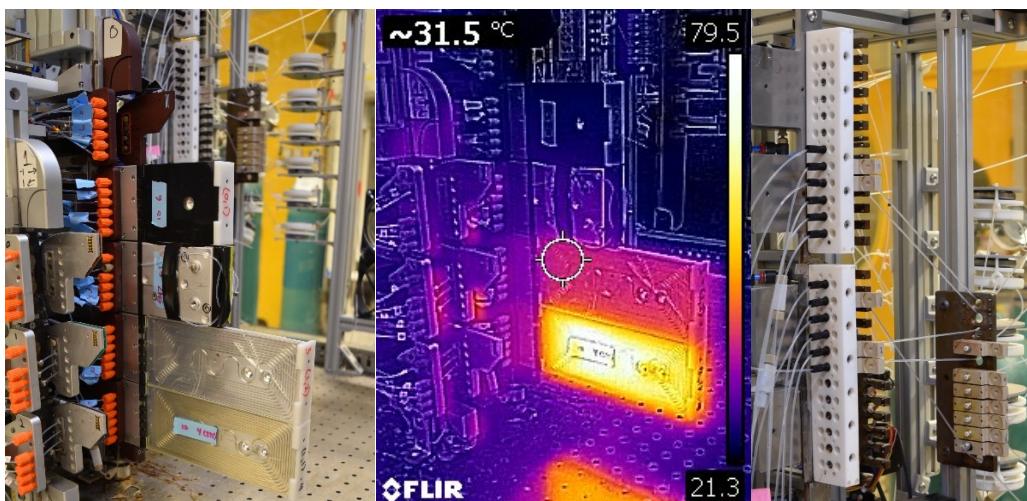


Fig. S31. From left to right: photo of the process stack, IR photo of the process stack, and photo of the reagent tree during operation for the synthesis of aspirin. Note that the reactors are covered in a protective fluoropolymer film, which causes a quantitative mismatch in estimated temperature.

Samples were collected in 10 minute intervals after steady state was achieved. A separation was performed to remove a small aqueous material resulting from minor breakthrough in the membrane separation module. The removed aqueous layer was washed with ethyl acetate (3 x 1 mL), after which all the organic fractions were combined and the solvent was removed by rotary evaporation. (The major contaminants in the crude aspirin were the incompletely quenched anhydride of aspirin and acetic acid.) The crude aspirin was dissolved in a 20 mL of warm (70 °C) water, filtered and then concentrated to dryness by rotary evaporation at 70 °C. After drying overnight, 287.4 mg of white solid was obtained (1.60 mmol, 91% yield, 1.72 g/h)

Aspirin, 1. ^1H NMR (400 MHz, CDCl_3) δ 8.03 (dd, $J = 7.9, 1.7$ Hz, 1H), 7.60 (td, $J = 7.7, 1.7$ Hz, 1H), 7.35 (td, $J = 7.6, 1.2$ Hz, 1H), 7.14 (dd, $J = 8.1, 1.2$ Hz, 1H), 2.3 (s, 3H) ppm. $^{13}\text{C}\{\text{H}\}$ NMR (101 MHz, CDCl_3) δ 170.1, 166.3, 150.8, 133.5, 131.5, 125.7, 123.7, 123.4, 19.6 ppm. HRMS Q-TOF: m/z calc for $\text{C}_9\text{H}_9\text{O}_3$ $[\text{M}+\text{H}]^+$ 153.0546, found 153.0543.

S3.3.2. Secnidazole, 4

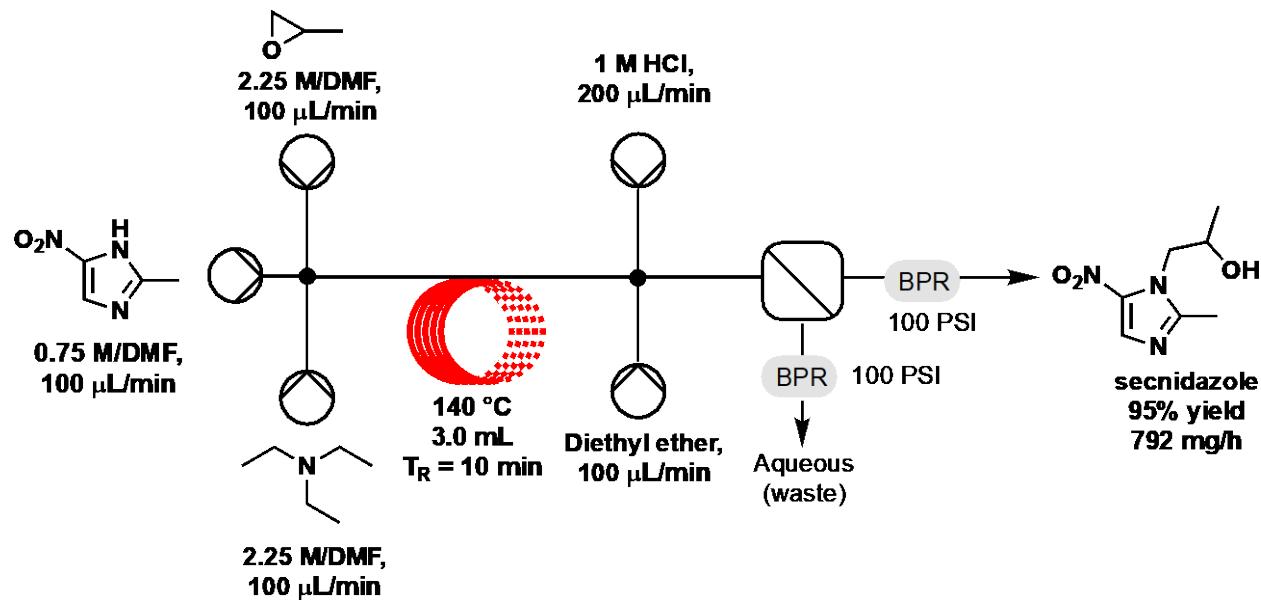


Fig. S32. Scheme for the flow synthesis of secnidazole.

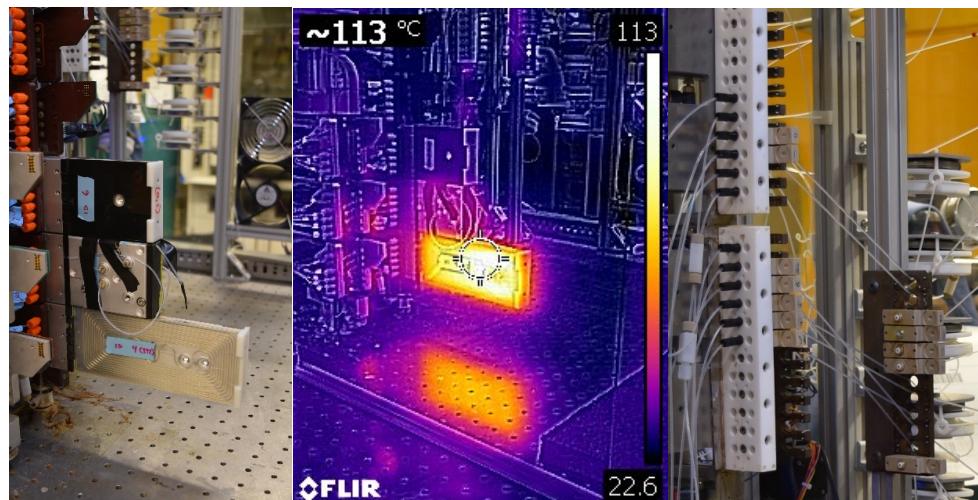


Fig. S33. From left to right: photo of the process stack, IR photo of the process stack, and photo of the reagent tree during operation for the synthesis of secnidazole. Note that the reactors are covered in a protective fluoropolymer film, which causes a quantitative mismatch in estimated temperature.

A 10-minute sample was collected after 2 hours of equilibration. The sample was concentrated and subjected to flash column chromatography using a 0-100% gradient of ethyl acetate in hexanes. 132 mg of a white solid were obtained (95% yield relative to imidazole, 792 mg/hr, 4.3 mmol/hr).

Secnidazole, 4. ^1H NMR (600 MHz, CDCl_3) δ 7.89 (s, 1H), 4.50 (dd, $J = 14.1, 2.7$ Hz, 1H), 4.17 (ddt, $J = 9.8, 6.3, 3.1$ Hz, 1H), 4.04 (dd, $J = 14.0, 9.1$ Hz, 1H), 2.53 (s, 3H), 2.44 (s, 1H), 1.34 (d, $J = 6.3$ Hz, 3H) ppm. $^{13}\text{C}\{\text{H}\}$ NMR (151 MHz, CDCl_3) δ 151.7, 138.6, 132.9, 67.6, 53.1, 21.3, 14.9 ppm. HRMS Q-TOF: m/z calc for $[\text{M}+\text{H}]^+$ $\text{C}_7\text{H}_{12}\text{N}_3\text{O}_3 = 186.0873$, found 186.0869; $[\text{M}+\text{Na}]^+$ 208.0693 calc, found 208.0677.

S3.3.3. Lidocaine, 7

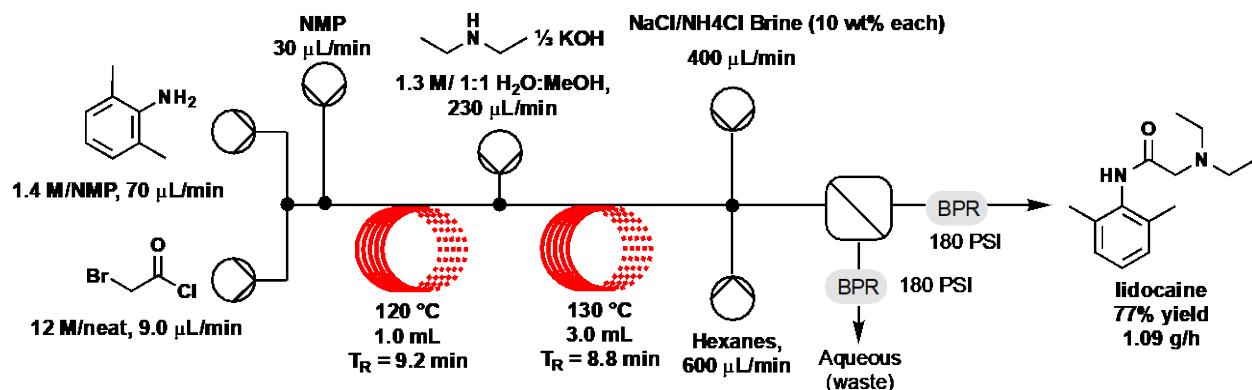


Fig. S34. Scheme for the flow synthesis of lidocaine.

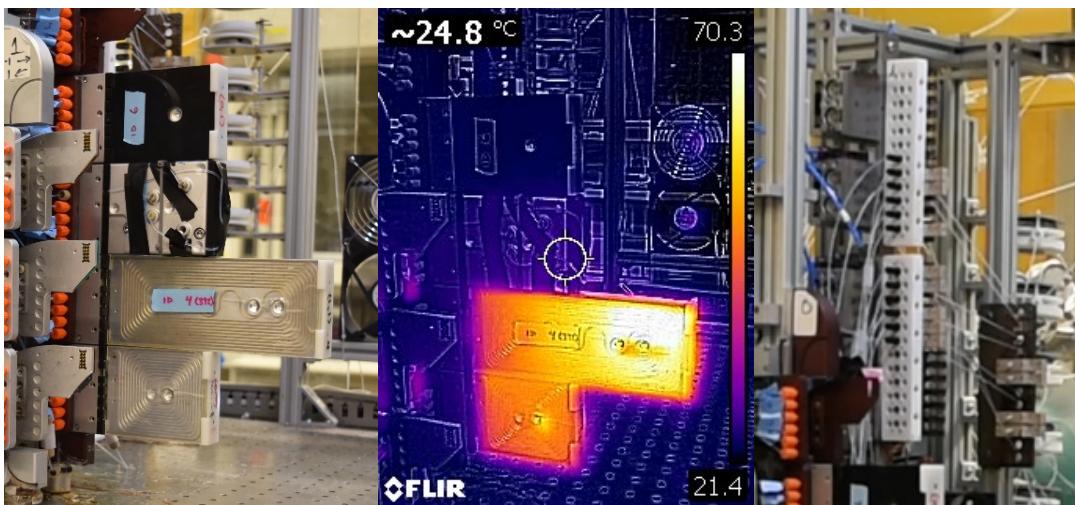


Fig. S35. From left to right: photo of the process stack, IR photo of the process stack, and photo of the reagent tree during operation for the synthesis of lidocaine. Note that the reactors are covered in a protective fluoropolymer film, which causes a quantitative mismatch in estimated temperature.

Sample was collected for 10 minutes after steady state was achieved and solvent was removed via rotatory evaporation. Purification was performed using column chromatography with Isco RediSep silica columns and a gradient of EtOAc in Hexanes. 179 mg of white crystalline solid was obtained (0.76 mmol, 77% yield, 1.09 g/hr).

Lidocaine, 7. ^1H NMR (400 MHz, CDCl_3) δ 8.91 (br s, 1H), 7.09 (s, 3H), 3.22 (s, 2H), 2.69 (q, J = 7.1 Hz, 4H), 2.24 (s, 6H), 1.14 (t, J = 7.1 Hz, 6H) ppm. $^{13}\text{C}\{\text{H}\}$ NMR (101 MHz, CDCl_3) δ 170.3, 135.1, 134.0, 128.2, 127.0, 57.6, 18.6, 12.7 ppm. HRMS Q-TOF: m/z calc for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{23}\text{N}_2\text{O}$ 235.1805, found 235.1806.

S3.3.4. **Diazepam, 12**

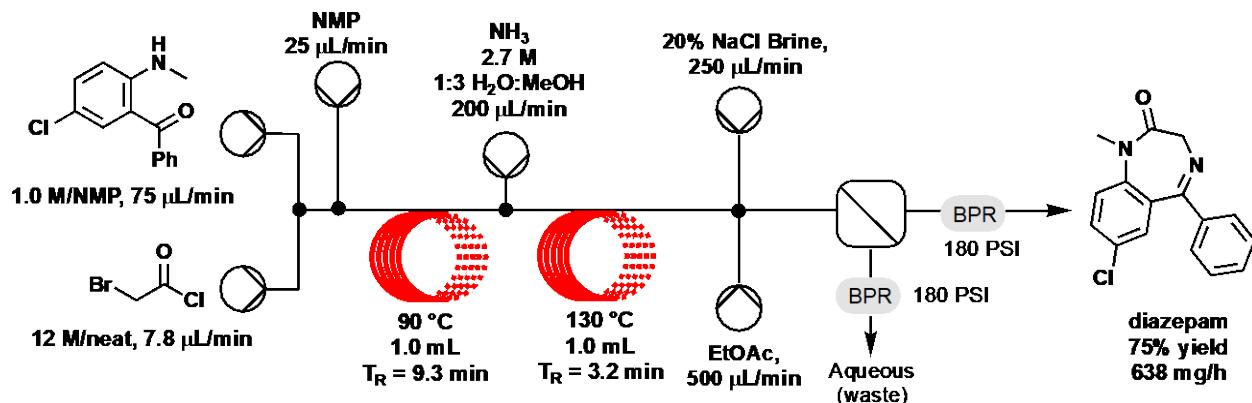


Fig. S36. Scheme for the flow synthesis of diazepam.

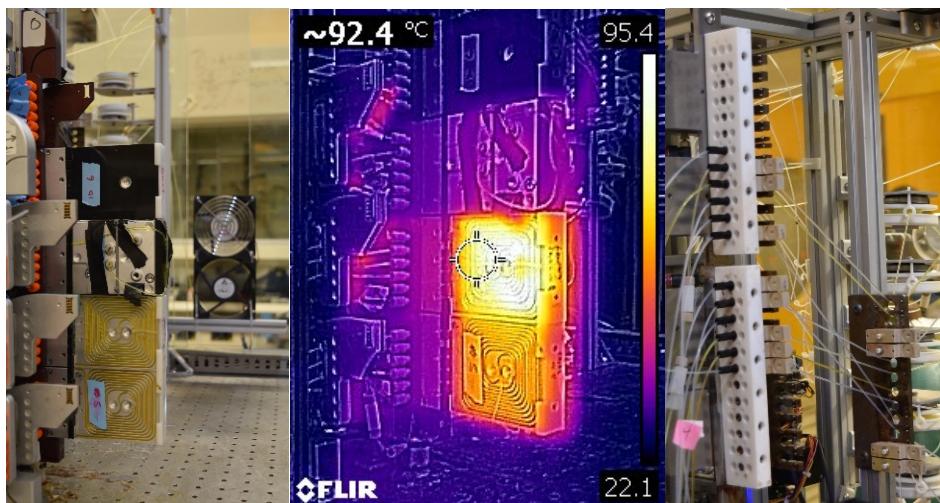


Fig. S37. From left to right: photo of the process stack, IR photo of the process stack, and photo of the reagent tree during operation for the synthesis of diazepam. Note that the reactors are covered in a protective fluoropolymer film, which causes a quantitative mismatch in estimated temperature.

Sample was collected for 30 minutes and solvent was removed via rotatory evaporation.

Purification was performed by column chromatography using RediSep silica columns and a 0–10% gradient of MeOH in DCM. 319 mg of a white solid were obtained (1.12 mmol, 75% yield, 638 mg/hr, 2.25 mmol/hr).

Diazepam, 12. ¹H NMR (500 MHz, CDCl₃) δ 7.61 (d, J = 7.1 Hz, 2H), 7.52 (dd, J = 8.8, 2.5 Hz, 1H), 7.48 (d, J = 7.3 Hz, 1H), 7.42 (t, J = 7.5 Hz, 2H), 7.31 (d, J = 8.9 Hz, 1H), 7.30 (d, J = 2.5 Hz, 1H), 4.84 (d, J = 10.8 Hz, 1H), 3.93 – 3.38 (m, 4H), 3.60 (s, 3H) ppm. ¹³C{¹H} NMR (126 MHz, CDCl₃) δ 169.9, 168.9, 142.6, 138.2, 131.4, 130.7, 130.1, 129.9, 129.5, 129.3, 128.4, 122.5, 57.0, 34.9 ppm. HRMS Q-TOF: m/z calc for [M+H]⁺ C₁₆H₁₄ClN₂O 285.0789, found 285.0790.

S3.3.5. (S)-Warfarin, 15

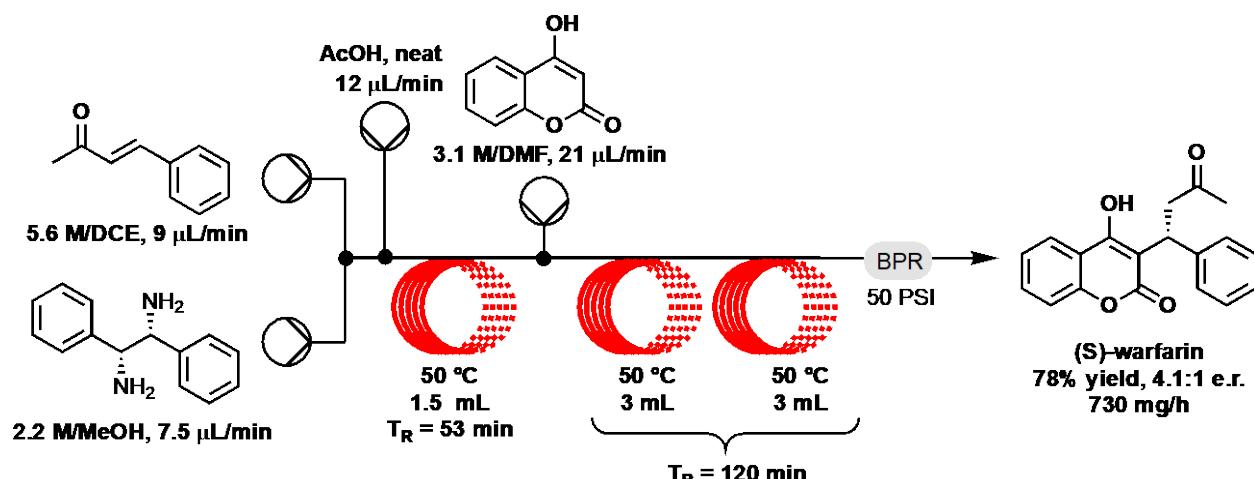


Fig. S38. Scheme for the flow synthesis of (S)-warfarin.

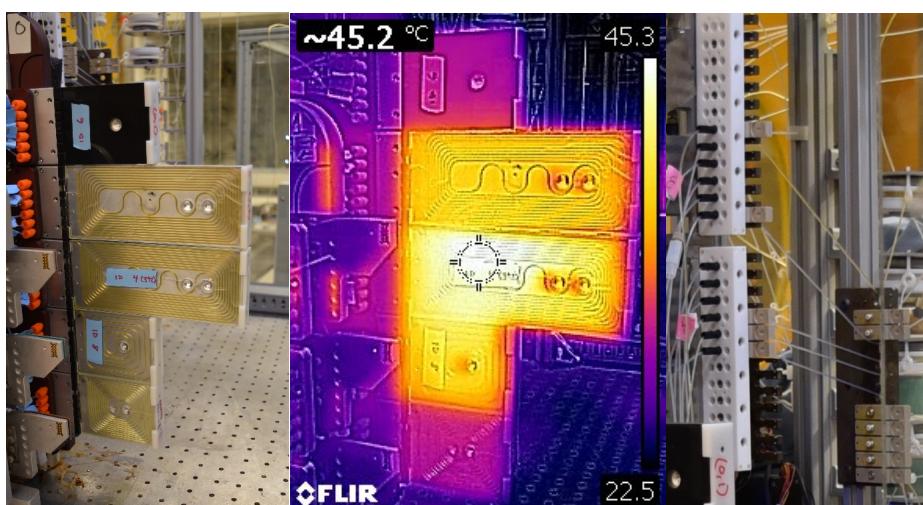


Fig. S39. From left to right: photo of the process stack, IR photo of the process stack, and photo of the reagent tree during operation for the synthesis of (S)-warfarin. Note that the reactors are covered in a protective fluoropolymer film, which causes a quantitative mismatch in estimated temperature.

Reaction was performed using 5.6 M benzylideneacetone in 1,2-dichloroethane (DCE), 2.2 M (1S,2S)-(-)-1,2-diphenylethylenediamine in MeOH, neat acetic acid, and 3.1 M 4-hydroxycoumarin in DMF. The system was equilibrated for 9 hours prior to collecting and a sample was collected for 30 minutes to yield 365 mg (78% yield relative to benzylideneacetone, 730 mg/hr, 2.37 mmol/hr) of pure compound after column chromatography. Purity was assayed with Agilent Poroshell 120 EC-CN (3.0 x 50 mm, 2.7 micron) with a 0.5 % - 15 % isopropanol gradient in hexanes. Enantiomeric excess was determined to be 61 % (4.1:1 er) using a Daicel Chiralpak AD-H column (250 mm length, 4.6 mm I.D., 5 μm particle size, Part No. 19325) using 1 mL/min isocratic method at 15% isopropanol and 85% hexanes. (R)-Warfarin $T_R = 7.54 \text{ min}$,

(S)-Warfarin τ_R = 21.76 min. $[\alpha]_D^{25} = -6.23^\circ$ (c 1.0, acetonitrile, 589 nm), literature $[\alpha]_D^{25} = -10.7^\circ$ (c 1.0 acetonitrile for >99% ee, see Mlynarski et al. *Green Chem.*, **2011**, *13*, 1155-1157. NMR spectra matches the literature (Guasch et al., *J. Org. Chem.* **2015**, *80*, 9900–9909).

(S)-warfarin, 15. ^1H NMR (500 MHz, DMSO-d₆) δ 11.67 (br s), 8.01 (br s), 7.85 (d, J = 7.7, 1H), 7.7-5.5 (m, 1H), 7.45-7.31 (m), 7.29-7.11 (m), 4.93 (m), 4.00 (m, 1H), 3.41 (m), 2.38-2.28 (m, 1H), 2.15 (m, 1H), 1.96-1.87 (m, 1H), 1.64 (s, 2.2H), 1.58 (s, 0.8H) ppm. Ketal Diastereomer-1 ^{13}C NMR (126 MHz, DMSO) δ 160.7, 159.2, 152.8, 144.4, 132.4, 128.7, 127.5, 126.4, 124.5, 123.2, 116.7, 116.1, 103.9, 100.1, 43.3, 35.7, 27.7 ppm. Ketal Diastereomer-2 δ 161.1, 159.8, 152.8, 144.3, 132.5, 128.3, 127.8, 126.1, 124.5, 123.1, 116.7, 116.0, 102.4, 101.9, 42.0, 36.5, 26.2 ppm. HRMS Q-TOF: m/z calc for C₁₈H₁₇O₄ [M+H] 309.1121, found 309.1116; [M+Na⁺] = 331.0941, found 331.0936.

S3.3.6. Safinamide, 18

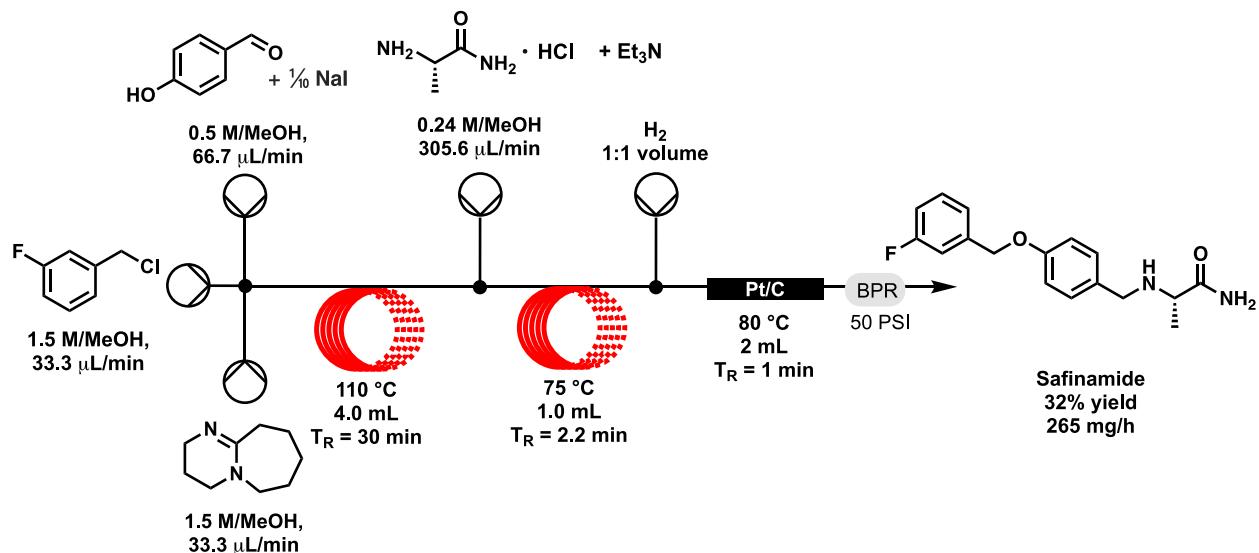


Fig. S40. Scheme for the flow synthesis of safinamide.

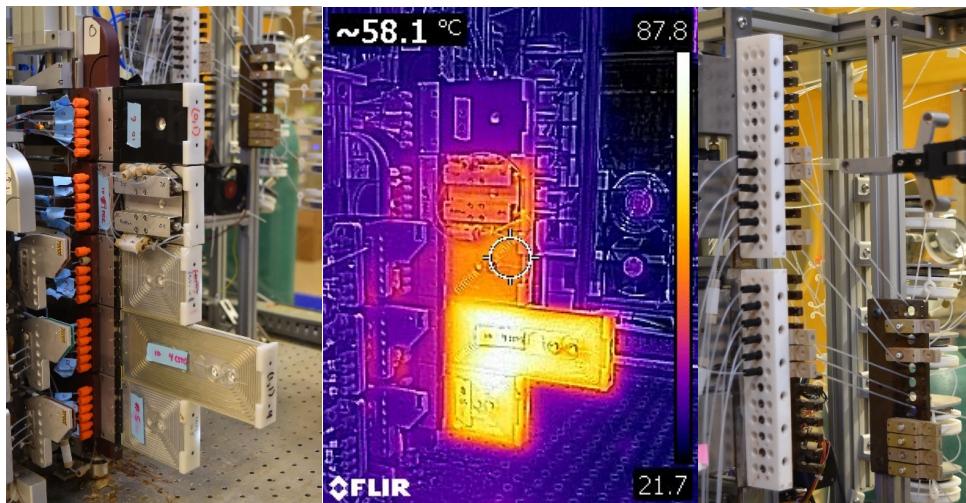
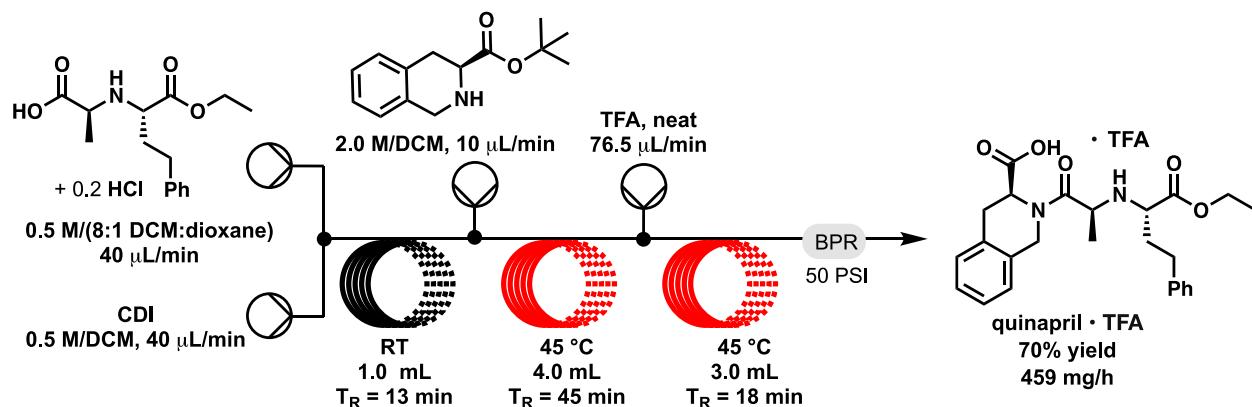
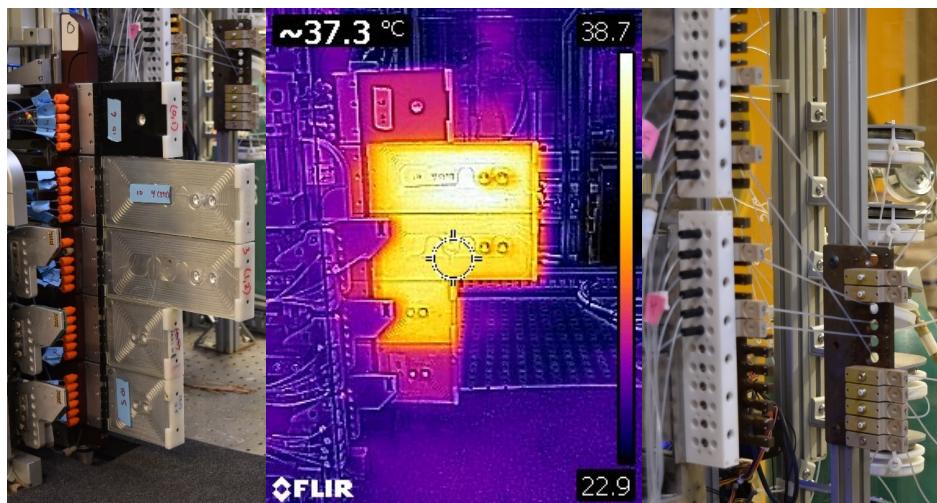


Fig. S41. From left to right: photo of the process stack, IR photo of the process stack, and photo of the reagent tree during operation for the synthesis of safinamide. Note that the reactors are covered in a protective fluoropolymer film, which causes a quantitative mismatch in estimated temperature.

A 4-minute sample was collected after steady state was established for isolation. The crude material from the robotic platform was evaporated to dryness, then taken up in 1 mL of 2:1 ACN/H₂O and 900 μ L was injected onto the Preparatory HPLC (Agilent 1200 with Eclipse XDB-C18 Column 212x250mm, 7 μ m particle size, 10-70% acetonitrile in water using 0.1% TFAH buffer over 40 minutes). Fractions containing product were identified with LC-MS (fractions 31-36, 15-19 minutes). HPLC fractions were combined to provide the trifluoroacetic acid salt of safinamide as a crystalline white solid after rotatory evaporation in 39 mg (0.041 mmol, 34% yield, 0.68 mmol/hr, 281 mg/hr).

Safinamide, 18. ^1H NMR (600 MHz, MeOD-*d*₄) δ 7.52 (td, *J* = 7.5, 1.7 Hz, 1H), 7.44 (d, *J* = 8.7 Hz, 2H), 7.42 – 7.36 (m, 1H), 7.21 (td, *J* = 7.5, 1.1 Hz, 1H), 7.15 (ddd, *J* = 10.2, 8.3, 1.1 Hz, 1H), 7.11 (d, *J* = 8.7 Hz, 2H) 5.19 (s, 3H), 4.15, 4.11 (ABq, 2H, really *J*_{AB} = 12.7 Hz), 3.91, (q, 1H), 3.37 (s, 3H), 1.56 (d, *J* = 7.1 Hz, 3H) ppm. $^{13}\text{C}\{\text{H}, \text{F}\}$ NMR (151 MHz, MeOD-*d*₄) δ 172.3, 162.2, 161.1, 132.7, 131.4, 131.3, 125.4, 125.2, 124.6, 116.5, 116.3, 65.0, 56.3, 50.5, 16.8 ppm. Trifluoroacetic acid ^{13}C peaks not observed. ^{19}F NMR (565 MHz, MeOD-*d*₄) δ -77.0, -120.8 ppm. HRMS Q-TOF: C₁₇H₂₀FN₂O₂ m/z calc for [M+H]⁺ 303.1503, found 303.1510.

S3.3.7.

Quinapril, 24a**Fig. S42.** Scheme for the flow synthesis of quinapril.**Fig. S43.** From left to right: photo of the process stack, IR photo of the process stack, and photo of the reagent tree during operation for the synthesis of quinapril and moexipril. Note that the reactors are covered in a protective fluoropolymer film, which causes a quantitative mismatch in estimated temperature.

General procedure for the synthesis of quinapril and related compounds moexipril, enalapril, ramipril, indolapril. Reactions were performed using 0.5 M *N,N*-carbonyldiimidazole (CDI) in DCM, 0.5 M *N*-(*S*)-1-ethoxycarbonyl-3-phenylpropyl]-L-alanine containing 20 mol % HCl (4 M in 1,4-dioxane) in DCM, 2 M amino acid derivatives in DCM, and neat trifluoroacetic acid. *tert*-Butyl-L-proline and (*S*)-*tert*-butyl 1,2,3,4-tetrahydroisoquinoline-3-carboxylate were prepared from the commercially available hydrochloride salts by treatment with KHCO₃. *tert*-Butyl (2*S*,3*aS*,6*aS*)-octahydrocyclopenta[*b*]pyrrole-2-carboxylate and *tert*-butyl (2*S*,3*aS*,7*aS*)-octahydro-1*H*-indole-2-carboxylate were prepared by treating the corresponding carboxylic acid with isobutylene and H₂SO₄. *tert*-Butyl (*S*)-6,7-dimethoxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylate was prepared by first treating 3,4-dimethoxyphenylalanine with formaldehyde and HCl, followed by reaction with isobutylene and H₂SO₄. Product samples were collected for 20

minutes after steady state was achieved and solvent was removed via rotatory evaporation. Column chromatography was performed using RediSep silica columns and MeOH/DCM to yield desired products as white solids. NMR data was collected for compounds **24a-e** however the NMR spectra for indolapril, moexipril, ramipril and quinapril are all complicated and rigorous NMR peak assignments were not possible.

Quinapril, 24a. The rigorous NMR solution phase characterization of Quinapril TFAH is beyond the scope of this work as three or more different rotamers/species are present. ¹H NMR (600 MHz, CDCl₃) δ 8.71 (br s, 3H), 7.32 – 7.04 (m, 9H), 5.44 (d, J = 16.9 Hz, 0.3H), 4.98 (t, J = 6.2 Hz, 0.6H), 4.82 (d, J = 17.0 Hz, 0.3H), 4.71 (m, 0.3H), 4.64 (m, 1.8H), 4.42 (m, 0.4H), 4.30 – 4.10 (m, 2H), 4.04 (m, 0.7H), 3.88 (q, J = 7.0 Hz, 0.3H), 3.72 (t, J = 6.3 Hz, 0.6H), 3.60 (dd, J = 8.6, 4.9Hz, 0.3H), 3.48 (dd, J = 15.6, 4.1 Hz, 0.1H), 3.41 (dd, J = 15.8, 3.2 Hz, 1H), 3.33 – 3.10 (m, 1.7H), 3.03 (dd, J = 15.7, 11.2 Hz, 0.1H), 2.97 (dd, J = 15.6, 13.0 Hz, 0.3H), 2.88 – 2.60 (m, 2.2H), 2.60 – 2.52 (m, 0.4H), 2.27 (q, J = 7.7 Hz, 1.2H), 2.25 – 2.17 (m, 0.4H), 2.12 – 2.05 (m, 0.1H), 1.61 (d, J = 7.0 Hz, 1.7H), 1.51 (d, J = 7.0 Hz, 1.1H), 1.43 (d, J = 7.4 Hz, 0.2H), 1.29 (t, J = 7.1 Hz, 0.6H), 1.27 – 1.22 (m, 2.4H). ¹³C NMR (151 MHz, CDCl₃) δ 172.2, 172.1, 170.6, 169.5, 169.3, 168.8, 168.4, 167.8, 165.4, 164.9, 140.52, 139.3, 139.1, 132.6, 131.9, 131.6, 131.5, 131.2, 129.3, 128.8, 128.7, 128.6, 128.5 128.4, 128.4, 128.4, 128.3, 128.3, 128.1, 127.8, 127.6, 127.5, 127.2, 127.1, 126.8, 126.7, 126.6, 126.5, 126.4, 126.3, 126.4, 120.1, 116.0 (q, J_{C-F} = 286 Hz), 63.3, 61.7, 60.8, 59.8, 59.4, 59.2, 58.9, 57.2, 56.2, 55.2, 54.9, 54.9, 54.6, 54.5, 54.1, 45.8, 44.9, 44.5, 44.1, 33.4, 32.9, 32.1, 31.7, 31.6, 31.1, 31.0, 30.9, 30.4, 29.4, 27.9, 20.8, 17.9, 16.2, 15.3, 14.3, 14.00, 13.9. One signal for carbonyl of trifluoroacetate counterion not observed at ~160 ppm.

Purity was assayed by HPLC using an Agilent Poroshell 120 SB-C18 2.7 μm, 2.1x50 mm. A solvent system of MeCN and H₂O was utilized with 0.1 % formic acid buffer. The method utilized a flow rate of 0.4 mL/min and a multiple step program consisting of a 0-1 min hold at 10% MeCN, 1-5 min ramp from 10-100% MeCN followed by a 100% MeCN hold for 1 minute. Isolated product was found to elute at 5.106 minutes and was found to be 98% pure by integration at 210 nm. HRMS Q-TOF: m/z calc for [M+H]⁺ C₂₅H₃₀N₂O₅ 439.2227, found 439.2230; [M+Na]⁺ calc 461.2047, found 461.2049.

S3.3.8. Moexipril, 24b

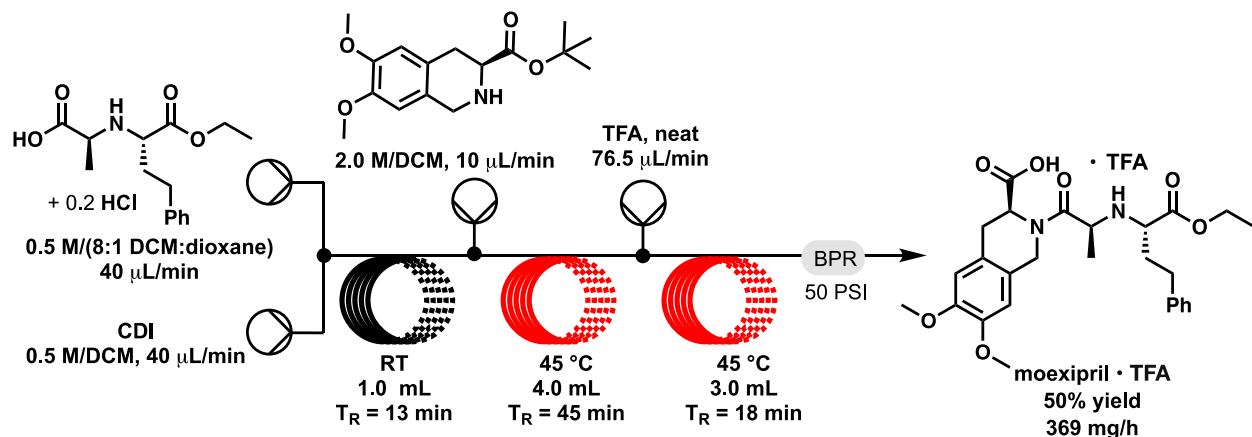


Fig. S44. Scheme for the flow synthesis of moexipril.

Moexipril, 24b. ^1H NMR (600 MHz, CDCl_3) δ 7.34 – 7.13 (m, ArCHs of homoallylbenzene ring, 5H), 6.68, 6.73, 6.71, 6.67, 6.66, 6.63, 6.62 (s, ArCHs on dimethoxybenzene ring, 2H), 5.38 (d, J = 16.5 Hz, 0.1 H), 5.13 (t, J = 5.7 Hz, 0.6H), 4.75 (d, J = 16.5 Hz, 0.2 H), 4.65 (m, 0.3H), 4.61 – 4.47 (m, 2H), 4.42 (m, 0.3H), 4.33 – 4.15 (m, 2.1H), 4.05 (m, 0.5H), 3.88, 3.88, 3.87, 3.86, 3.84, 3.82 (s, ArOMe groups, 6H), 3.74 (t, J = 6.2 Hz, 0.6H), 3.62 (dd, J = 8.9, 5.7 Hz, 0.1H), 3.29 – 3.07 (m, 1.6H), 2.97 – 2.50 (m, 2.5H), 2.37 – 2.22 (m, 1.7H), 1.65 (d, J = 7.4 Hz, 1.7H), 1.59 – 1.45 (m, 2H), 1.34 – 1.24 (m, 3H) ppm. ^{13}C NMR (151 MHz, CDCl_3) δ 172.3, 172.1, 169.5, 169.3, 168.7, 168.6, 168.2, 167.0, 165.3, 165.0, 159.3 (q, $J_{\text{C}-\text{F}}$ = 41 Hz), 148.9, 148.6, 148.5, 148.5, 148.4, 148.2, 140.5, 140.3, 139.2, 139.1, 134.6, 128.9, 128.8, 128.8, 128.6, 128.5, 128.5, 128.3, 124.3, 123.8, 123.5, 123.3, 122.5, 118.9, 116.3 (q, $J_{\text{C}-\text{F}}$ = 293 Hz), 111.6, 111.3, 111.2, 111.1, 109.5, 109.1, 109.0, 63.4, 61.7, 61.6, 60.5, 59.7, 59.2, 58.9, 57.2, 56.3, 56.2, 56.1, 55.2, 55.0, 54.8, 54.6, 53.7, 46.6, 45.3, 44.6, 44.2, 43.8, 33.0, 32.9, 32.8, 31.9, 31.7, 31.1, 31.0, 30.6, 30.5, 29.8, 29.4, 29.3, 27.9, 24.2, 20.8, 18.2, 16.16, 15.1, 14.3, 14.0, 13.9 ppm.

Purity was assayed by HPLC using Agilent Eclipse Plus C18 3.5 μm , 4.6 x 100 mm. A solvent system of MeCN and H_2O was utilized with a 0.1 % acetic acid buffer. An isocratic method at 40% MeCN was utilized with a flow rate of 1.5 mL/min. Product retention time was 1.143 min and found to be 93% pure by integration at 210 nm. HRMS Q-TOF: m/z calc for $[\text{M}+\text{H}]^+$ $\text{C}_{27}\text{H}_{34}\text{N}_2\text{O}_7$ = 499.2439, found 499.2444; $[\text{M}+\text{Na}]^+$ calc 521.2258, found 521.2259.

S3.3.9. Enalapril, 24c

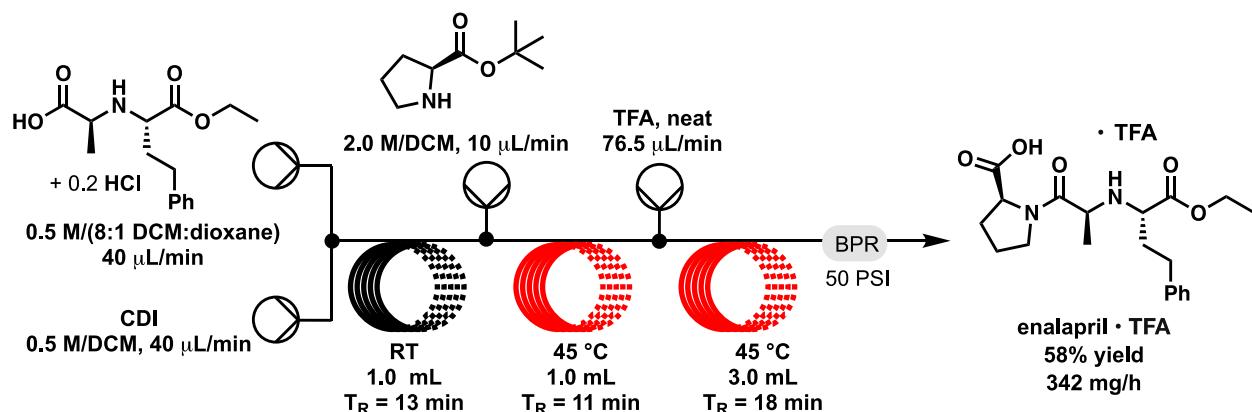


Fig. S45. Scheme for the flow synthesis of enalapril.

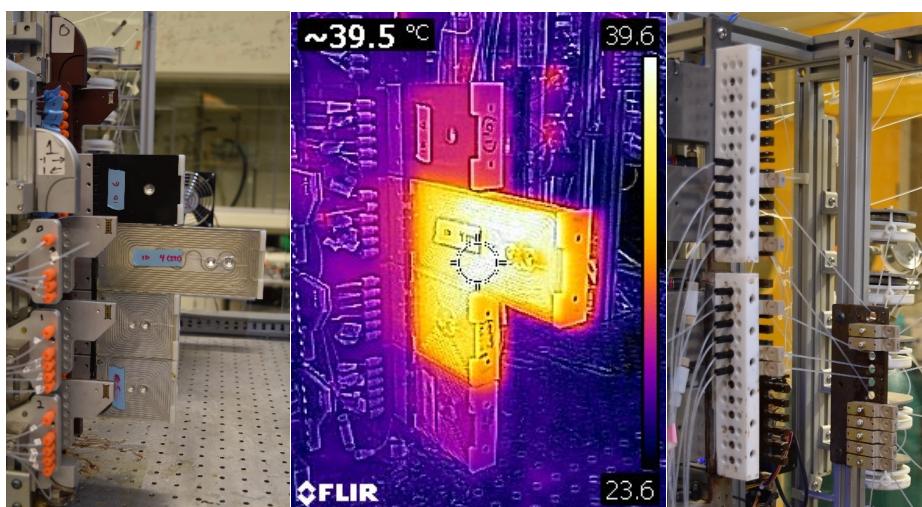


Fig. S46. From left to right: photo of the process stack, IR photo of the process stack, and photo of the reagent tree during operation for the synthesis of moexipril, ramipril, and indolapril. Note that the reactors are covered in a protective fluoropolymer film, which causes a quantitative mismatch in estimated temperature.

Enalapril, 24c. ^1H NMR (600 MHz, CDCl_3) δ 10.10 (br s, 3H), 7.30 – 7.24 (m, 2H), 7.21 (, J = 7.6 Hz, 1H), 7.16 – 7.14 (m, 2H), 4.46 – 4.36 (m, 2H), 4.25 (q, J = 7.3 Hz, 2H), 3.83 (t, J = 6.4 Hz, 1H), 3.54 (qd, J = 10.6, 10.1, 5.7 Hz, 2H), 2.78 (dt, J = 14.9, 7.6 Hz, 1H), 2.68 (dt, J = 13.9, 7.9 Hz, 1H), 2.37 – 2.25 (m, 3H), 2.10 (ddt, J = 18.6, 12.7, 6.2 Hz, 2H), 2.01 (dd, J = 11.7, 6.9 Hz, 1H), 1.57 (d, J = 6.8 Hz, 3H), 1.31 (t, J = 7.2 Hz, 3H) ppm. $^{13}\text{C}\{\text{H}\}$ NMR (151 MHz, CDCl_3) δ 172.9, 167.6, 167.5, 161.2 (q, $J_{\text{C}-\text{F}} = 39$ Hz) 139.3, 128.8, 128.5, 126.8, 161.1 (q, $J_{\text{C}-\text{F}} = 290$ Hz), 63.3, 60.2, 59.2, 55.1, 47.4, 31.4, 31.1, 28.8, 25.2, 15.1, 14.0 ppm. HRMS Q-TOF: m/z calc for $[\text{M}+\text{H}]^+$ $\text{C}_7\text{H}_{11}\text{N}_3\text{O}_3$ = 377.2071, found 377.2073; $[\text{M}+\text{Na}]^+$ calc 399.1890, found 399.1886.

S3.3.10. Ramipril, 24d

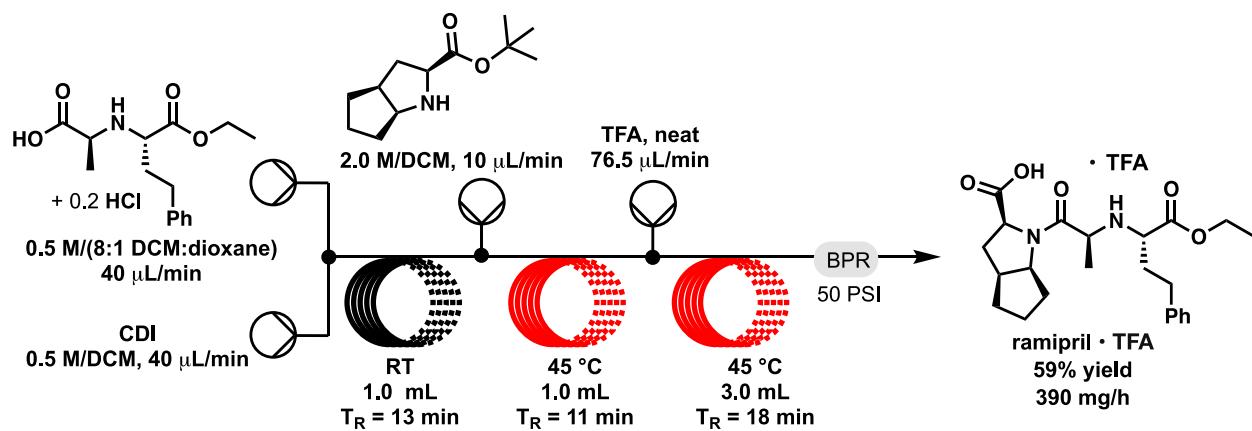


Fig. S47. Scheme for the flow synthesis of ramipril.

Ramipril, 24d. At least 3 major rotamers could be observed in the ^1H and ^{13}C NMR spectra. The rigorous NMR characterization of ramipril TFAH is beyond the scope of this work. ^1H NMR (600 MHz, CDCl_3) δ 9.33 (br s, 3H), 7.27 (m, 2H), 7.20 (m, 2H), 7.15 (m, 1H), 4.52-4.43 (m, 1.4H), 4.38 (td, $J = 8.0, 3.9$ Hz, 0.4H), 4.32-4.16 (m, 3H), 4.12 (t, $J = 6.2$ Hz, 0.4H), 3.76 (t, $J = 6.4$ Hz, 0.5H), 2.87-2.60 (m, 3H), 2.52 - 2.22 (m, 3.3H), 2.09 - 1.98 (m, 1H), 1.97 - 1.88 (m, 0.6H), 1.88 - 1.74 (m, 2.1H), 1.71 - 1.51 (m, 5H), 1.45 - 1.36 (m, 0.6H), 1.34 - 1.26 (m, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 173.6, 173.2, 173.1, 170.2, 168.7, 168.7, 168.4, 167.9, 167.6, 166.9, 161.5 (q, $J_{\text{C}-\text{F}} = 38$ Hz) 161.4 (q, $J_{\text{C}-\text{F}} = 38$ Hz), 140.7, 139.3, 139.2, 128.9, 128.8, 128.7, 128.6, 128.5, 128.5, 128.4, 126.9, 126.7, 126.6, 126.5, 116.0 (q, $J_{\text{C}-\text{F}} = 290$ Hz), 115.9 (q, $J_{\text{C}-\text{F}} = 290$ Hz), 65.9, 64.9, 63.6, 63.5, 63.4, 63.3, 62.2, 61.7, 61.6, 60.4, 60.2, 60.0, 59.5, 58.7, 57.3, 56.7, 56.6, 55.5, 54.7, 44.5, 41.4, 41.2, 34.6, 34.1, 33.9, 33.6, 33.4, 33.0, 32.6, 32.2, 32.0, 31.9, 31.8, 31.5, 31.3, 31.1, 31.0, 31.0, 30.9, 30.7, 29.8, 25.9, 24.8, 23.9, 23.8, 17.2, 16.2, 16.2, 16.0, 15.5, 14.3, 14.1, 14.0, 13.9.

Purity was assayed by HPLC using Agilent Eclipse Plus C18 3.5 μm , 4.6 x 100 mm. A solvent system of MeCN and H_2O was utilized with a 0.1 % acetic acid buffer. The method utilized a flow rate at 1.5 mL/min and a multistep program that consisted of a 0-1 min hold at 15% MeCN followed by a ramp from 15 – 95% MeCN. Product retention time was 7.932 min and found to be 100% pure by integration at 210 nm. HRMS Q-TOF: m/z calc for $[\text{M}+\text{H}]^+$ $\text{C}_{23}\text{H}_{32}\text{N}_2\text{O}_5 = 417.2384$, found 417.2387; $[\text{M}+\text{Na}]^+$ calc 439.2203, found 439.2198.

S3.3.11. Indolapril, 24e

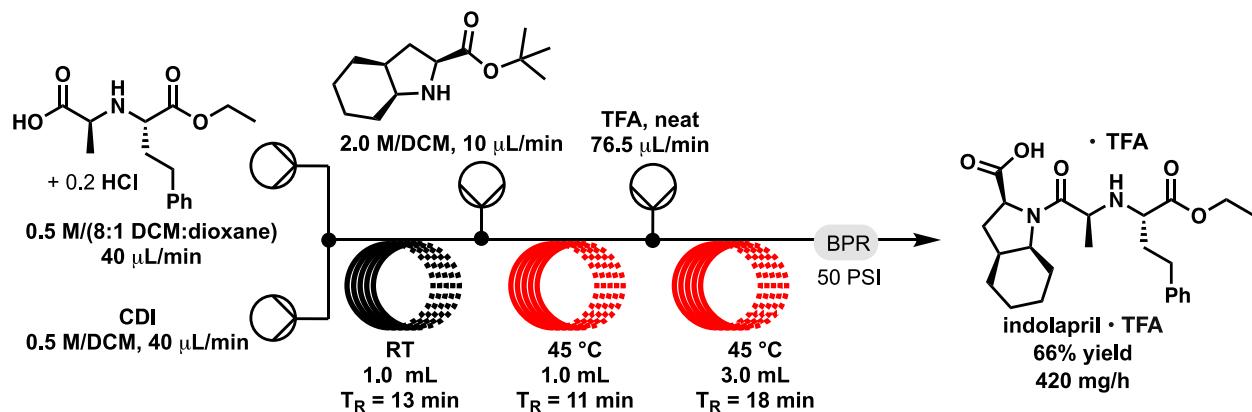


Fig. S48. Scheme for the flow synthesis of indolapril.

Indolapril, 24e. At least 3 major rotamers could be observed in the ¹H and ¹³C NMR spectra. The rigorous NMR characterization of Indolapril TFAH is beyond the scope of this work. ¹H NMR (600 MHz, CDCl₃) δ 8.78 (br s, 3H), 7.31 – 7.15 (m, 5H), 4.43 (m, 0.9H), 4.3 – 4.02 (m, 3.2H), 3.83 (m, 0.7H), 3.46 (m, 0.7H), 2.83 (m, 0.3H), 2.70 (m, 1.5H), 2.41 – 2.08 (m, 4.8H), 1.93 (m, 0.2H), 1.83 – 1.57 (m, 3.7H), 1.51 (m, 2.1H), 1.43 (d, J = 6.6 Hz, 2H), 1.3 (m, 3.8H), 1.25 – 1.07 (m, 1.7H). ¹³C NMR (151 MHz, CDCl₂) δ 174.3, 172.8, 172.8, 170.6, 170.5, 168.7, 167.8, 166.3, 163.3 (q, J_{C-F} = 36 Hz), 140.9, 139.9, 139.7, 128.7, 128.7, 128.7, 128.6, 128.5, 128.5, 126.6, 126.5, 126.4, 116.4 (q, J_{C-F} = 290 Hz), 63.0, 62.6, 61.6, 60.0, 59.1, 59.0, 58.9, 58.9, 58.8, 58.5, 56.5, 56.2, 55.6, 54.6, 53.8, 37.6, 36.0, 35.7, 32.9, 32.4, 31.9, 31.4, 31.4, 31.1, 31.0, 29.9, 29.3, 29.1, 28.3, 27.5, 27.2, 25.9, 25.8, 25.6, 25.4, 23.8, 23.8, 23.7, 23.5, 23.5, 20.9, 20.8, 20.2, 19.7, 19.6, 18.2, 16.4, 16.1, 14.3, 14.2, 14.0, 13.7.

Purity was assessed as detailed above for quinapril (24a). Product was found to elute at 5.120 minutes and 100% pure by integration at 210 nm. HRMS Q-TOF: m/z calc for [M+H]⁺ C₂₄H₃₄N₂O₅ = 431.2540, found 431.2546; [M+Na]⁺ calc 453.2360, found 453.2364.

S3.3.12. Celecoxib, 28aa

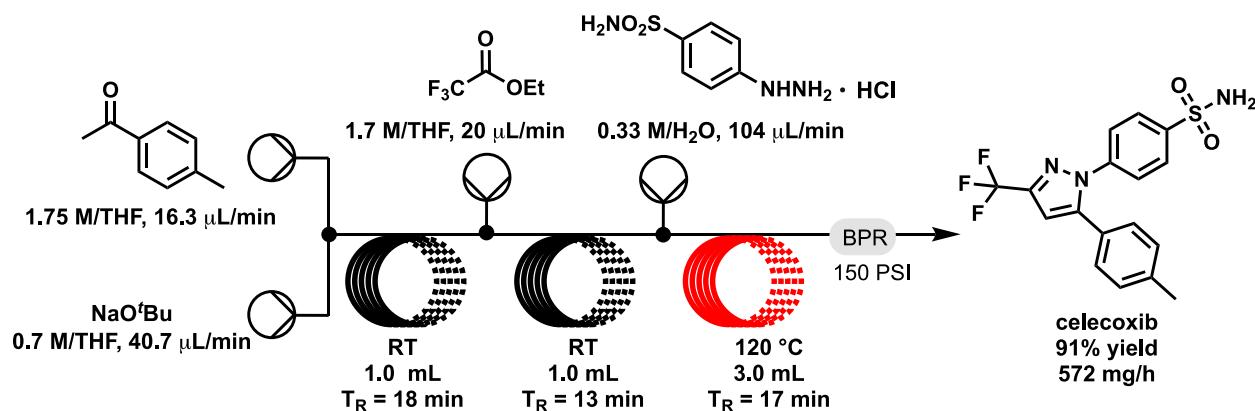


Fig. S49. Scheme for the flow synthesis of celecoxib.

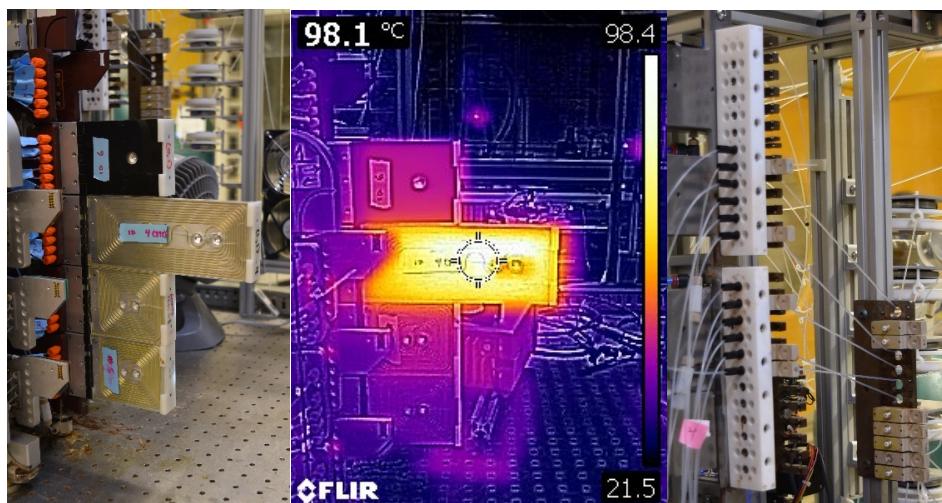


Fig. S50. From left to right: photo of the process stack, IR photo of the process stack, and photo of the reagent tree during operation for the synthesis of celecoxib, **28ab**, **28bb**, and **28ba**. Note that the reactors are covered in a protective fluoropolymer film, which causes a quantitative mismatch in estimated temperature.

General procedure for the synthesis of celecoxib and related compounds 28aa-28bb. Reaction was performed using 0.7 M NaOtBu in THF, 1.75 M acetophenone derivative in THF, 1.7 M ethyl trifluoroacetate in THF and 0.33 M of the HCl hydrazine. The NaOtBu solution was prepared by dissolving 4.03 g of NaOtBu in 60 mL of anhydrous THF and filtering through a 0.45 micron PTFE syringe filter. 4-hydrazinylbenzenesulfonamide HCl was prepared in a two-step process using NaNO₂ followed by reduction with SnCl₂ (See Brodfuehrer et al. *J. Org. Chem.* **1997**, *62*, 9192-9202). Phenylhydrazine HCl solution was prepared by dissolving phenylhydrazine in water with 1 eq. of HCl using commercially available 37% HCl_(aq).

Sample was collected for 60 minutes and worked up as detailed above to provide 572 mg (91% yield) of crystalline white solid material. ^1H NMR (500 MHz, CDCl_3) δ 7.93 (d, $J = 8.4$ Hz, 2H), 7.51 (d, $J = 8.4$ Hz, 2H), 7.21 (d, $J = 7.8$ Hz, 2H), 7.14 (d, $J = 7.8$ Hz, 2H), 6.77 (s, 1H), 5.08 (s, 2H, $-\text{SO}_2\text{NH}_2$), 2.40 (s, 3H). $^{13}\text{C}\{\text{H}\}$ NMR (126 MHz, CDCl_3) δ 145.3, 144.1 (q, $^2J_{\text{C},\text{F}} = 38.5$ Hz), 142.6, 141.3, 139.8, 129.8, 128.7, 127.5, 125.7, 125.5, 121.0 (q, $^1J_{\text{C},\text{F}} = 269.2$ Hz), 106.4 (q, $^3J_{\text{C},\text{F}} = 2$ Hz), 21.3 ppm. ^{19}F NMR (470.7 MHz, CDCl_3) δ -62.5 ppm. HRMS AccuTOF: m/z calc for $[\text{M}+\text{H}]^+$ $\text{C}_{17}\text{H}_{15}\text{F}_3\text{N}_3\text{O}_2\text{S}$ 382.0832, found 382.0832.

S3.3.13. 4-(5-(4-bromophenyl)-3-(trifluoromethyl)-1H-pyrazol-1-yl)benzenesulfonamide, 28ab

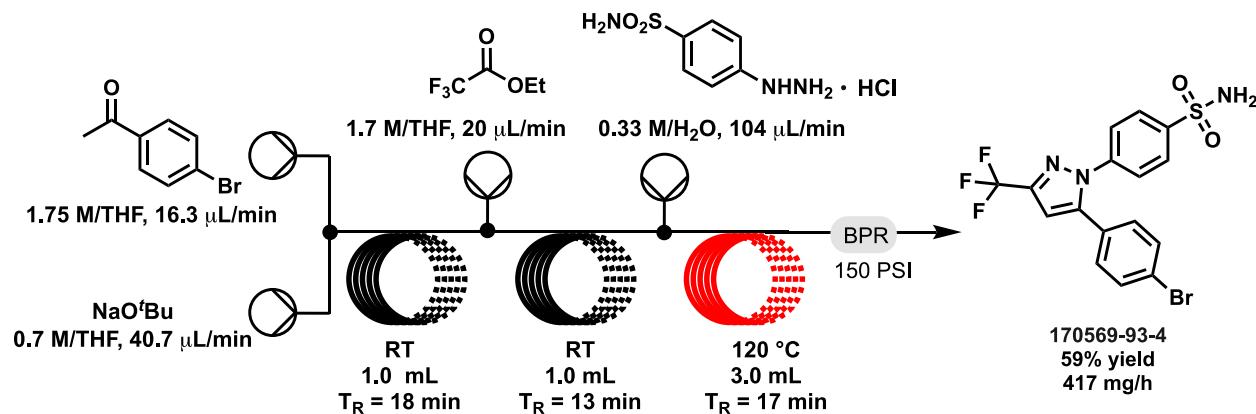


Fig. S51. Scheme for the flow synthesis of 4-(5-(4-bromophenyl)-3-(trifluoromethyl)-1H-pyrazol-1-yl)benzenesulfonamide.

Sample was collected for 20 minutes and worked up as detailed above to provide 139 mg (59% yield) of crystalline white solid. ¹H NMR (500 MHz, CDCl₃) δ 7.93 (d, *J* = 8.6 Hz, 2 H), 7.53 (d, *J* = 8.4 Hz, 2 H), 7.46, (d, *J* = 8.4 Hz, 2 H), 7.13 (d, *J* = 8.6 Hz, 2 H), 6.80 (s, 1H), 5.32 (s, 2H, ArSO₂NH₂) ppm. ¹³C{¹H} NMR (126 MHz, CDCl₃) δ 144.3 (q, ²J_{C,F} = 38.2 Hz), 144.0, 142.2, 141.9, 132.4, 130.3, 127.6, 127.5, 125.7, 124.1, 120.9 (q, ¹J_{C,F} = 266.7 Hz), 106.7 (q, ³J_{C,F} = 2 Hz) ppm. ¹⁹F NMR (470.7 MHz, CDCl₃) δ -62.5 ppm. HRMS AccuTOF: C₁₆H₁₂BrF₃N₃O₂S m/z calc for [⁷⁹M+H]⁺ 445.9780, found 445.9788; m/z calc for [⁸¹M+H]⁺ 447.9760, found 447.9770.

S3.3.14. 5-(4-bromophenyl)-1-phenyl-3-(trifluoromethyl)-1H-pyrazole, 28bb

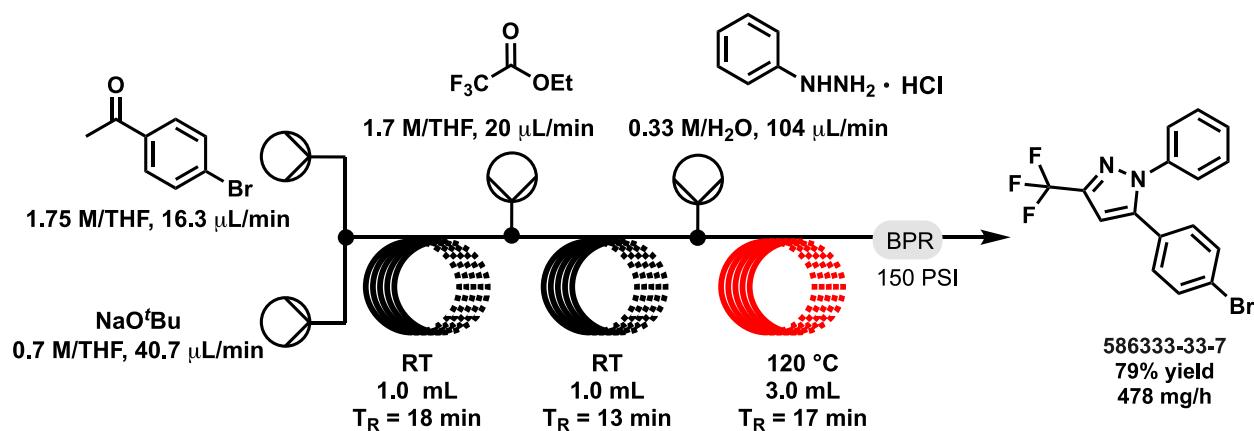


Fig. S52. Scheme for the flow synthesis of 5-(4-bromophenyl)-1-phenyl-3-(trifluoromethyl)-1H-pyrazole.

Sample was collected for 60 minutes and worked up as detailed above to provide 478 mg (79% yield) of white crystalline material. ¹H NMR (500 MHz, CDCl₃) δ 7.46 (d, J = 8.4 Hz, 2H), 7.39 (m, 3H), 7.30 (m, 2H), 7.09 (d, J = 8.4 Hz, 2H), 6.79 (s, 1H) ppm. ¹³C{¹H} NMR (126 MHz, CDCl₃) δ 143.5, 143.3 (q, ²J_{C,F} = 37.3 Hz) 139.0, 132.0, 130.3, 129.3, 128.7, 128.1, 125.5, 123.5, 121.2, (q, ¹J_{C,F} = 266.6 Hz), 105.67 (q, ³J_{C,F} = 2 Hz) ppm. ¹⁹F NMR (470.7 MHz, CDCl₃) δ -62.3 ppm. HRMS AccuTOF: m/z calc for C₁₆H₁₁BrF₃N₂ [⁷⁹M+H]⁺ 367.0052, found 367.0052; [⁸¹M+H]⁺ 369.0034, found 369.0032.

S3.3.15. 1-phenyl-5-(p-tolyl)-3-(trifluoromethyl)-1H-pyrazole, 28ba

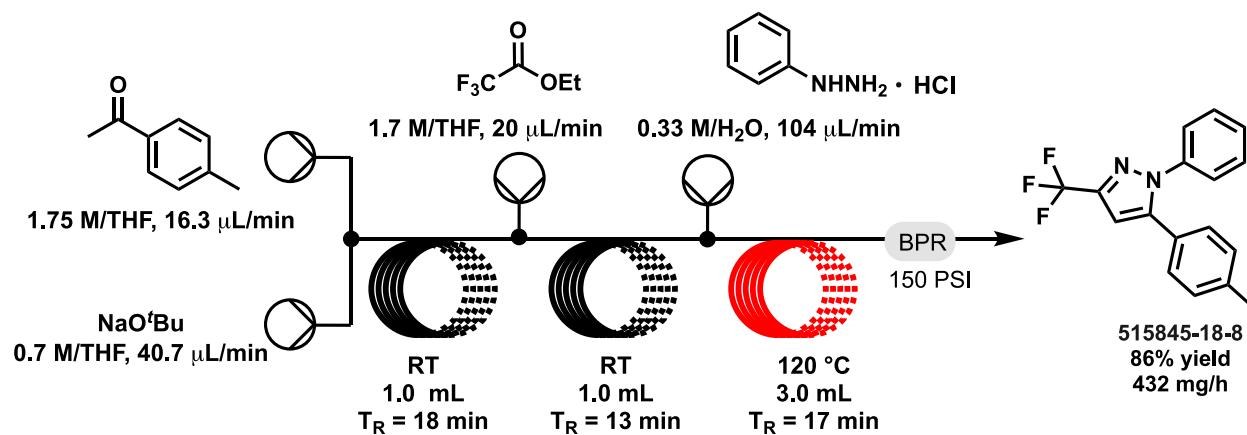
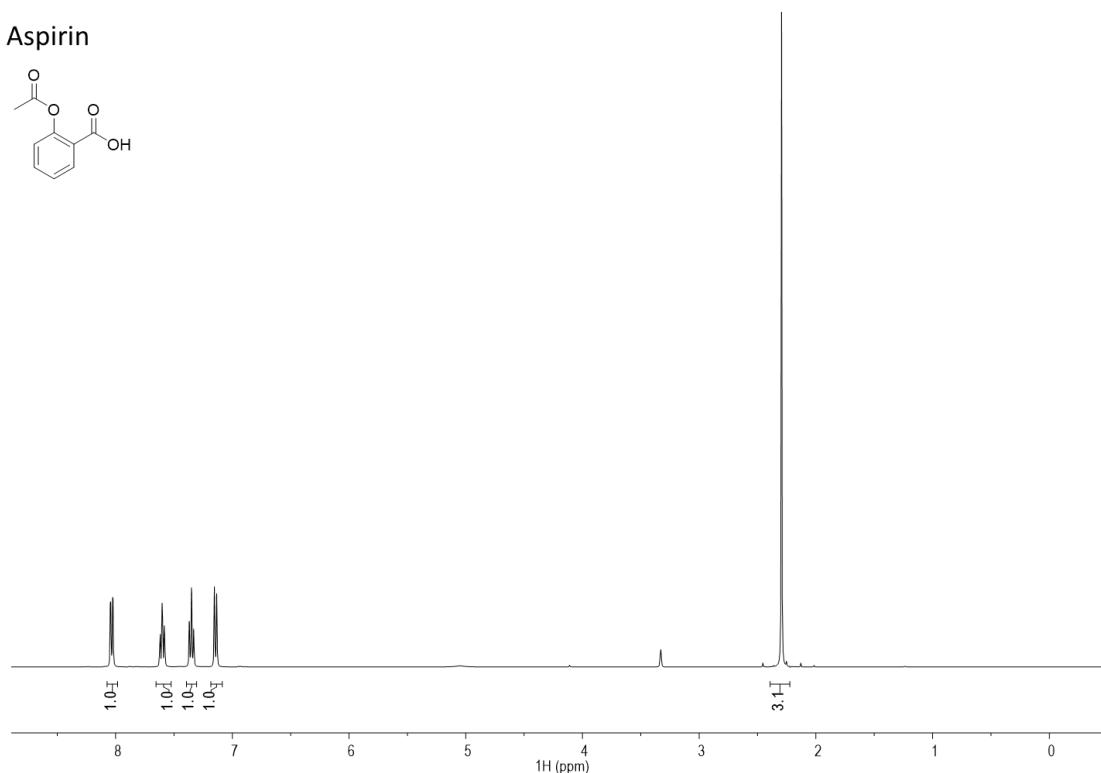
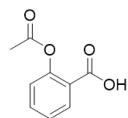
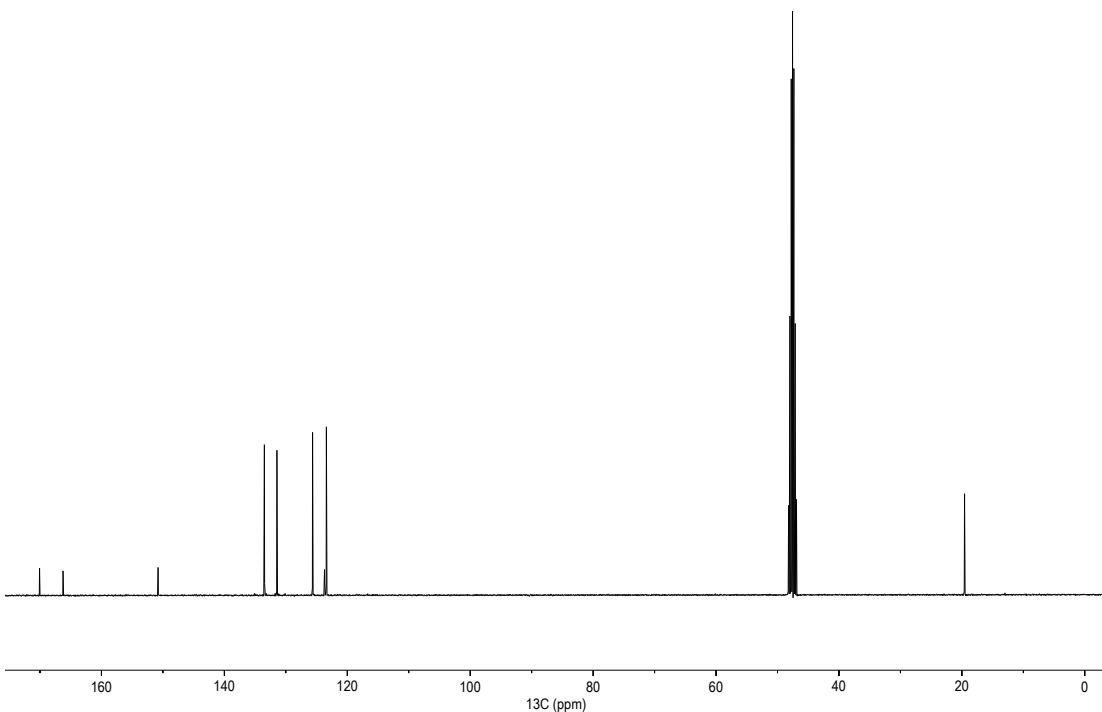


Fig. S53. Scheme for the flow synthesis of 1-phenyl-5-(p-tolyl)-3-(trifluoromethyl)-1H-pyrazole.

Sample was collected for 60 minutes and worked up as detailed above to provide 432 mg (86% yield) of white crystalline material. ^1H NMR (500 MHz, CDCl_3) δ 7.39 (m, 3H), 7.35 (m, 2H), 7.16, 7.13 (ABq, $J_{\text{AB}} = 8.7$ Hz, 4H), 6.75 (s, 1H), 2.38 (s, 3H) ppm. $^{13}\text{C}\{\text{H}\}$ NMR (126 MHz, CDCl_3) δ 144.9, 143.1 (q, $^2J_{\text{C},\text{F}} = 38.5$ Hz), 139.4, 139.1, 129.4, 129.1, 128.7, 128.3, 126.3, 125.6, 121.4 (q, $^1J_{\text{C},\text{F}} = 266.6$ Hz), 105.2 (q, $^3J_{\text{C},\text{F}} = 2$ Hz), 21.3 ppm. ^{19}F NMR (470.7 MHz, CDCl_3) δ -62.2 ppm. HRMS AccuTOF: m/z calc for $\text{C}_{17}\text{H}_{14}\text{F}_3\text{N}_2$ [M+H]⁺ 303.1104, found 303.1108.

S3.4**NMR Spectra and HPLC Traces**

Aspirin

**Fig. S54.** ^1H NMR (400 MHz, CDCl_3) for aspirin, **1**.**Fig. S55.** $^{13}\text{C}\{^1\text{H}\}$ NMR (101 MHz, CDCl_3) for aspirin, **1**.

Secnidazole

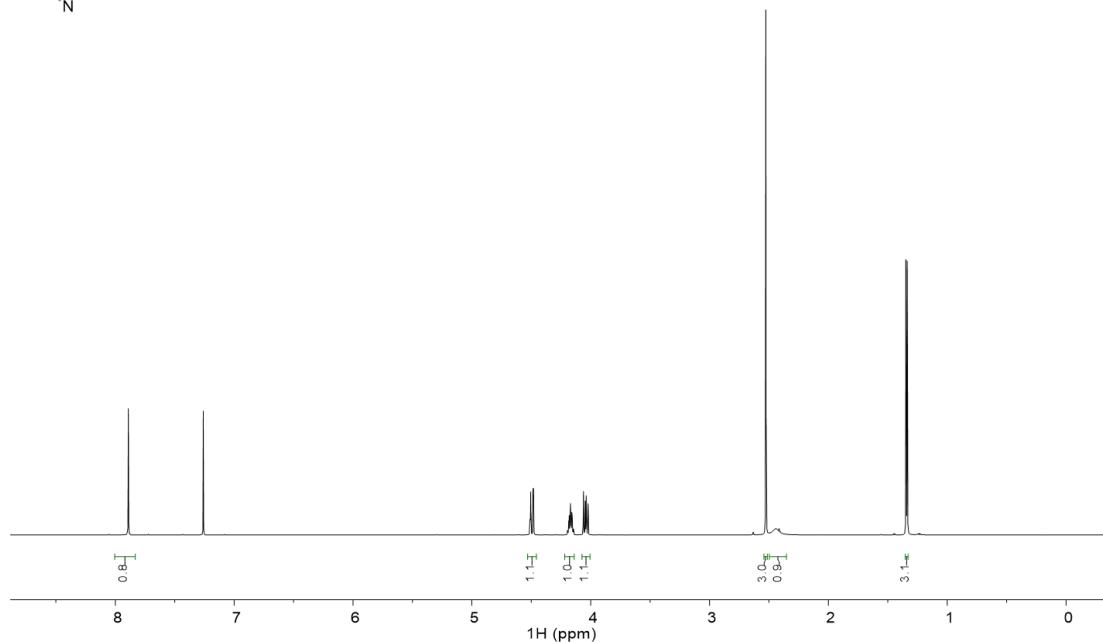
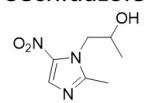


Fig. S56. ¹H NMR (600 MHz, CDCl₃) for secnidazole, 4.

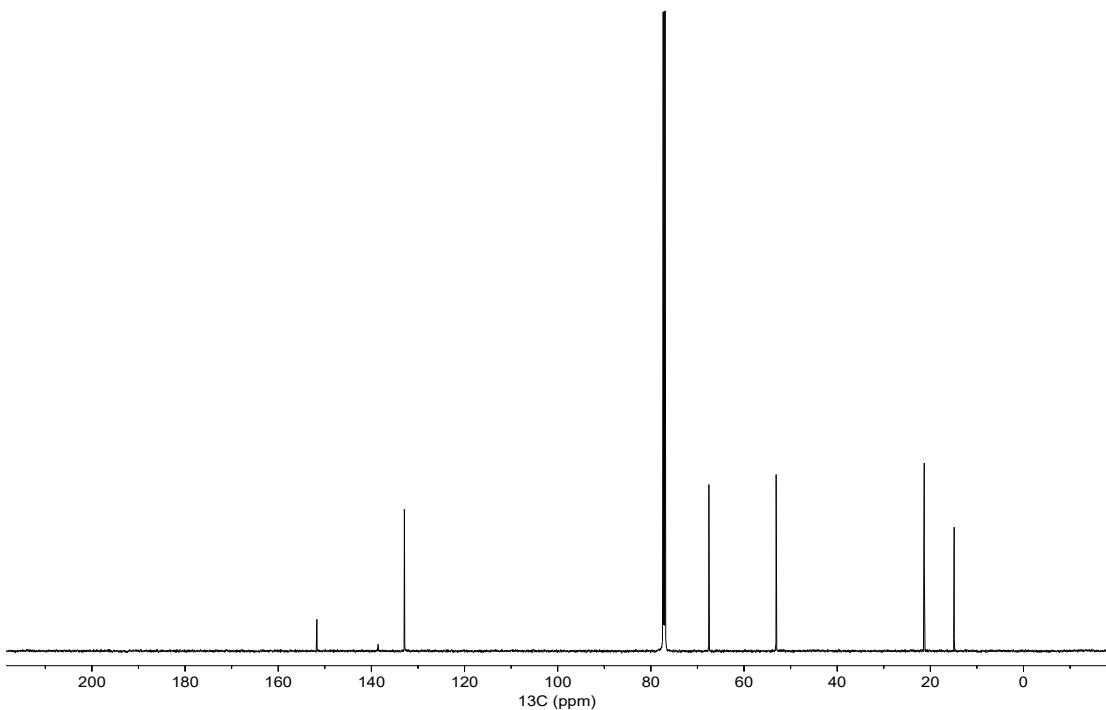


Fig. S57. ¹³C{¹H} NMR (151 MHz, CDCl₃) for secnidazole, 4.

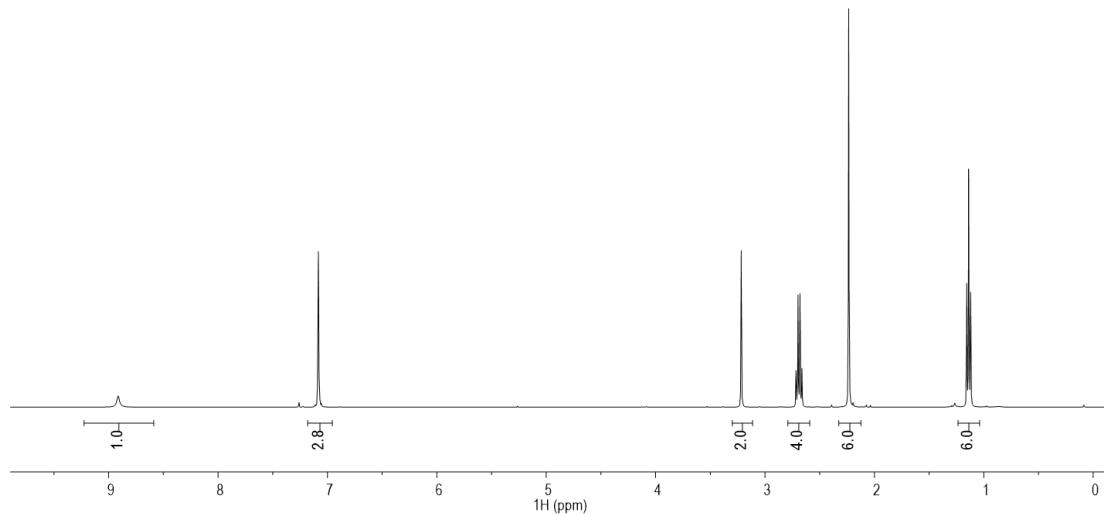
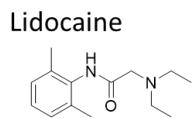


Fig. S58. ^1H NMR (400 MHz, CDCl_3) for lidocaine, 7.

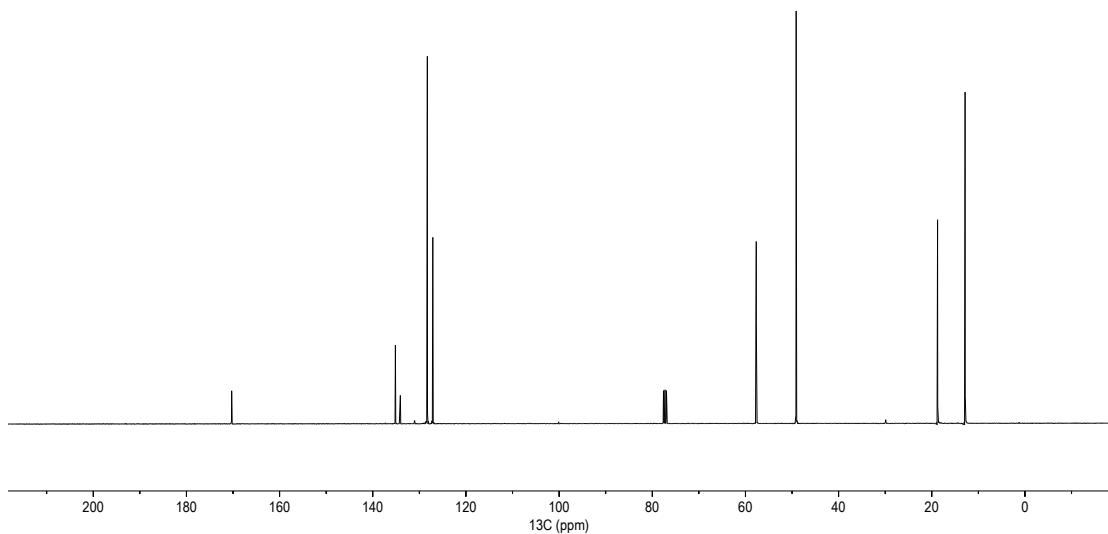


Fig. S59. $^{13}\text{C}\{^1\text{H}\}$ NMR (101 MHz, CDCl_3) for lidocaine, 7.

Diazepam

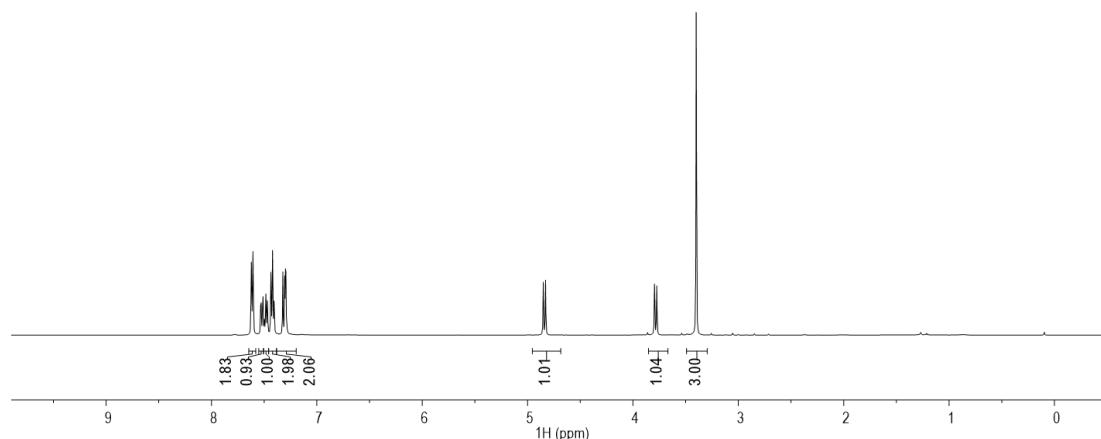
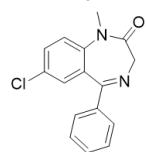


Fig. S60. ^1H NMR (500 MHz, CDCl_3) for diazepam, **12**.

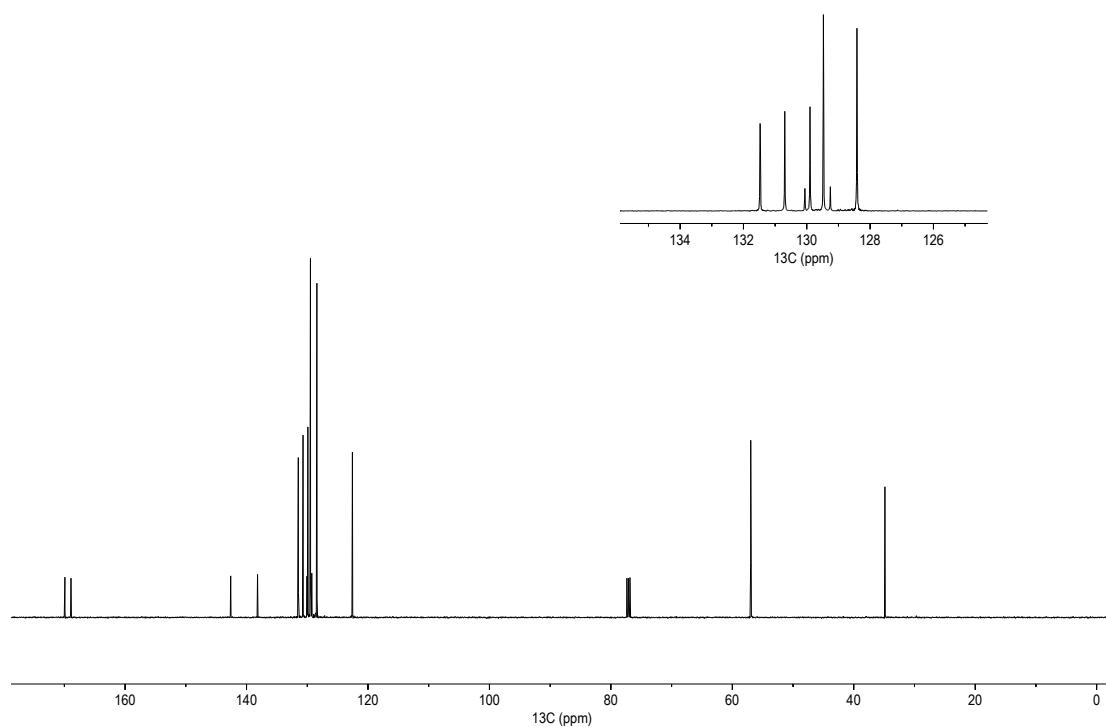


Fig. S61. $^{13}\text{C}\{\text{H}\}$ NMR (126 MHz, CDCl_3) for diazepam, **12**.

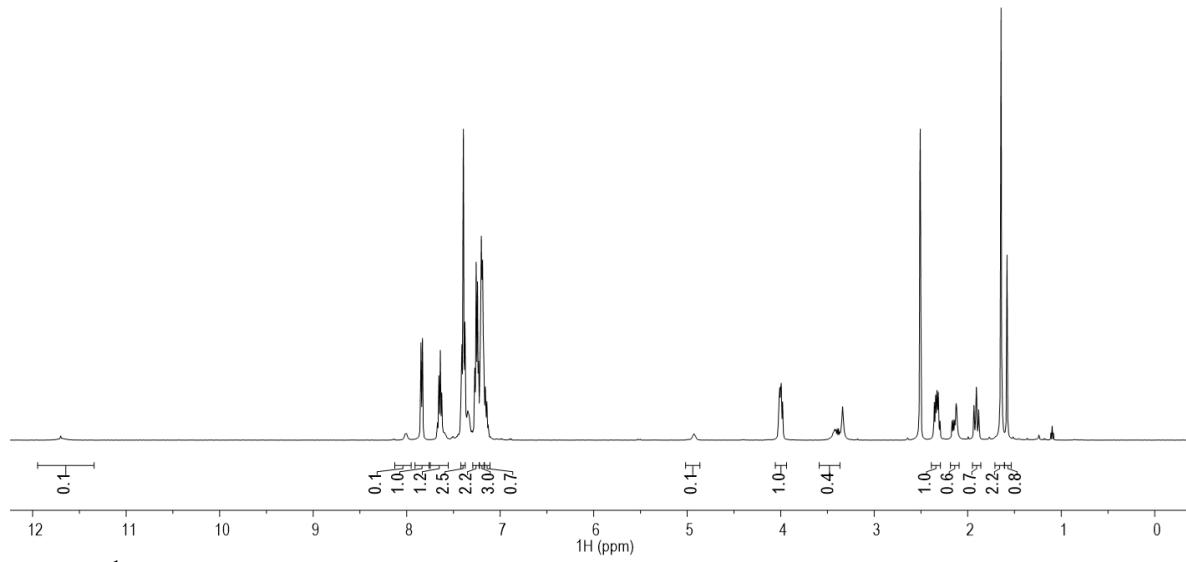
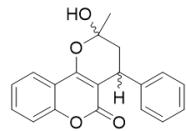


Fig. S62. ^1H NMR (500 MHz, DMSO-d_6) for (S)-warfarin, **15**. Diastereomeric mixture of ketals with ~10% open chain form present.

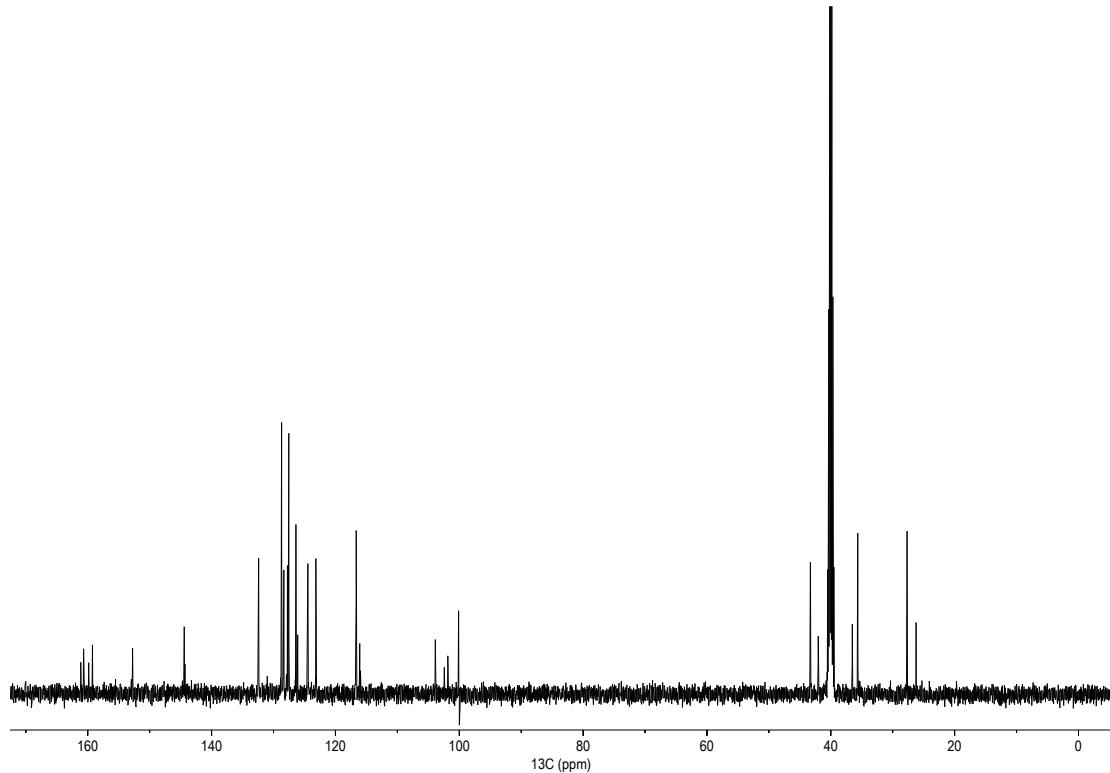


Fig. S63. $^{13}\text{C}\{^1\text{H}\}$ NMR (126 MHz, DMSO-d_6) for (S)-warfarin, **15**. Diastereomeric mixture of ketals with ~10% open chain form present.

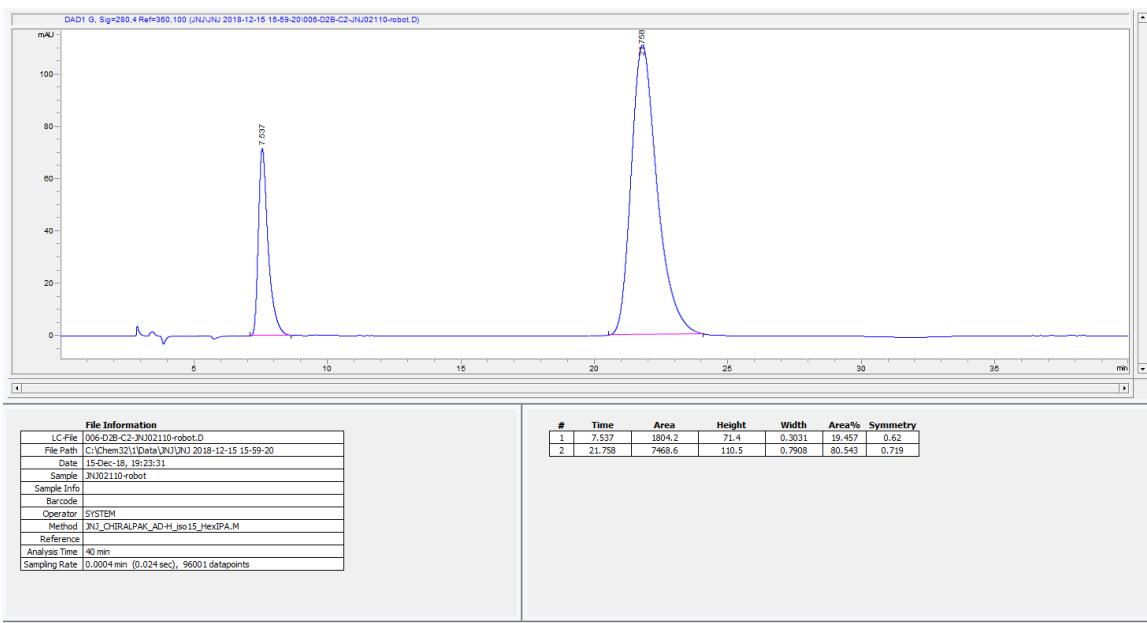


Fig. S64. Reaction analysis of enantioselective enriched product (S)-warfarin, **15** by chiral HPLC.

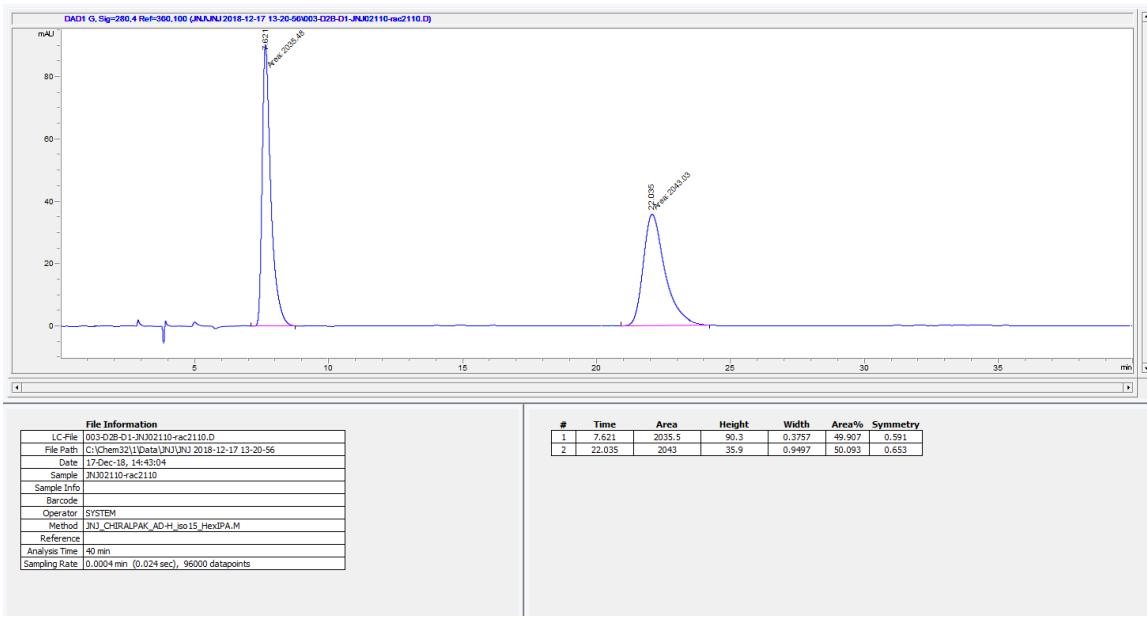


Fig. S65. Analysis of racemic (R/S)-warfarin by chiral HPLC.

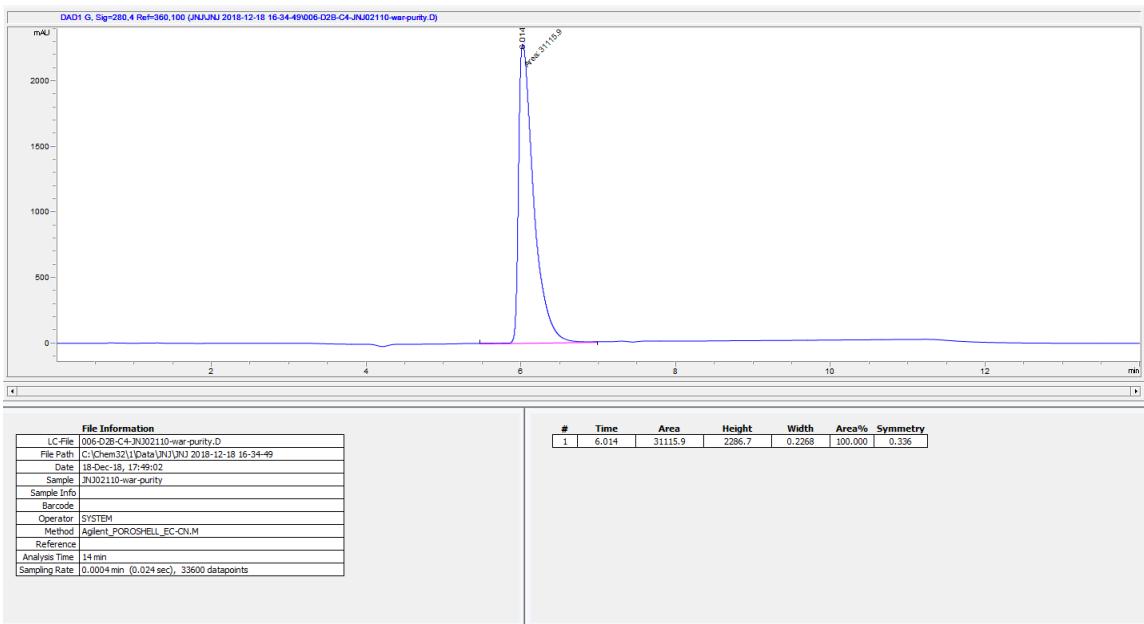


Fig. S66. Purity assessment of (S)-warfarin, **15** by HPLC.

Safinamide trifluoroacetic acid salt

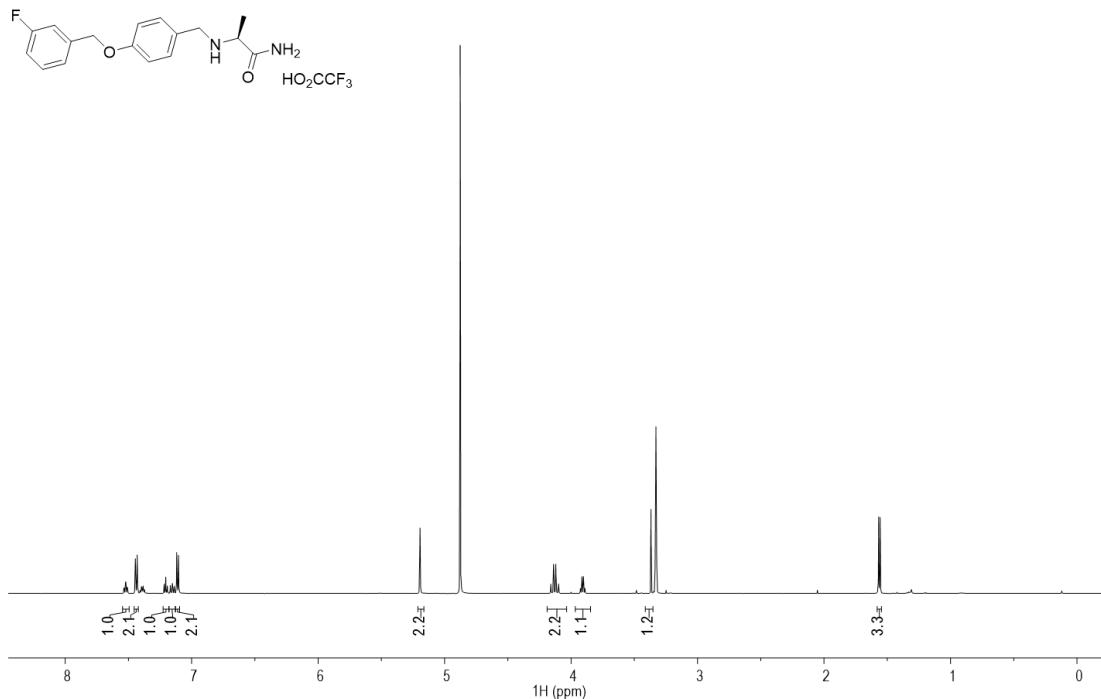


Fig. S67. ¹H NMR (600 MHz, MeOD-*d*₄) for safinamide•TFA, **18**•TFA.

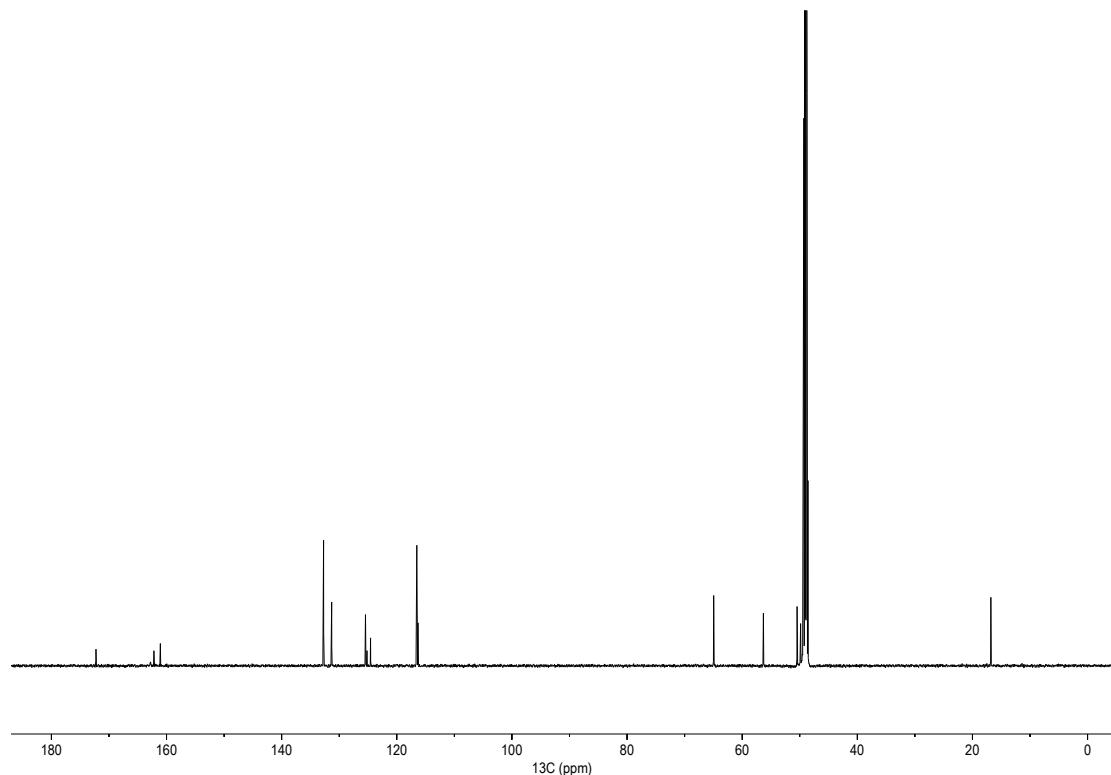


Fig. S68. ¹³C{¹H,¹⁹F} NMR (151 MHz, MeOD-*d*₄) for safinamide•TFA, **18**•TFA.

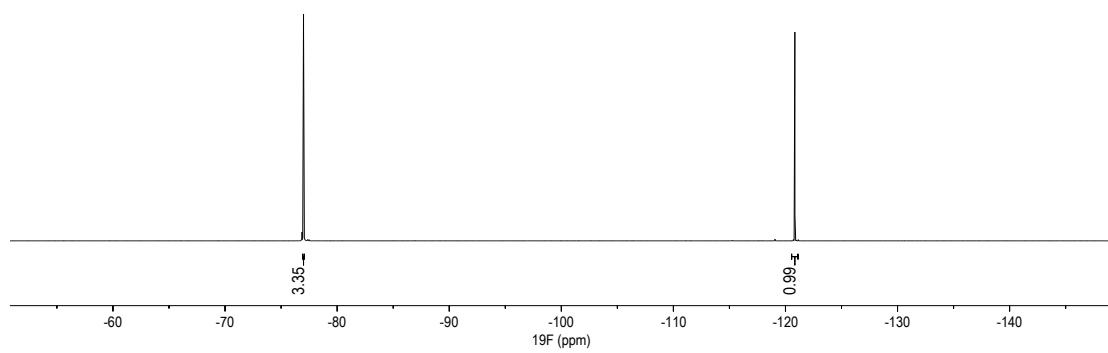


Fig. S69. ${}^{19}\text{F}$ NMR (565 MHz, MeOD- d_4) for safinamide•TFA, **18•TFA**.

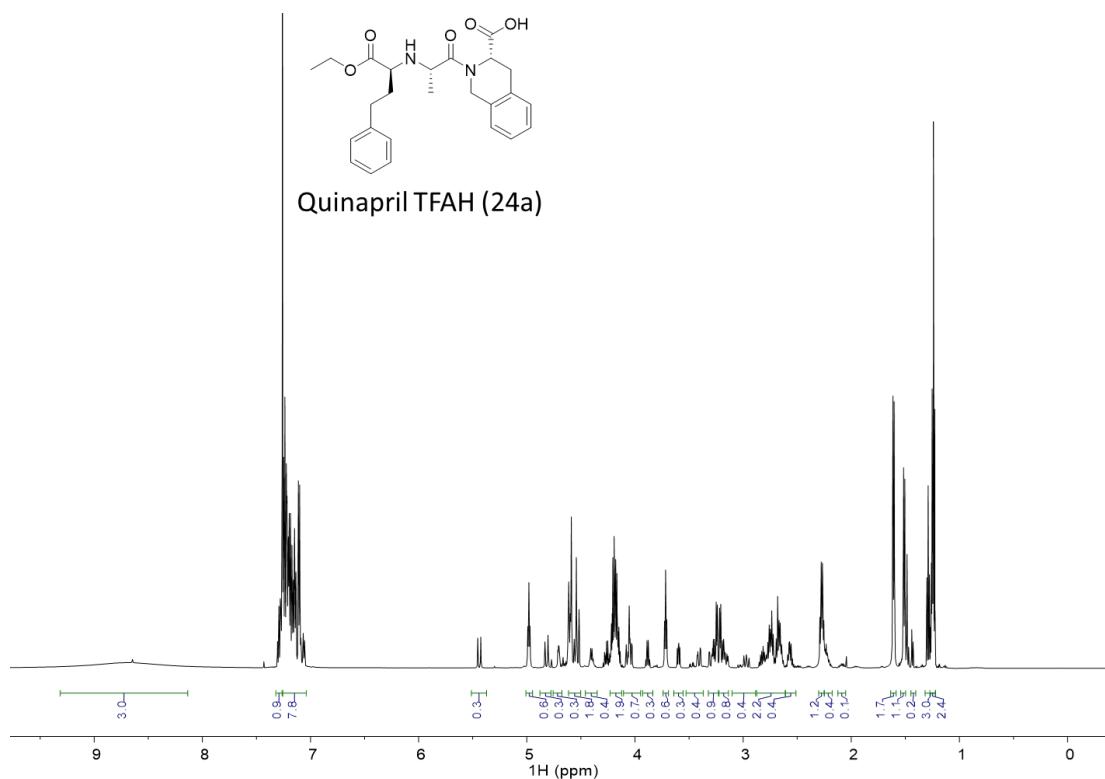


Fig. S70. ^1H NMR (600 MHz, CDCl_3) for quinapril-TFA, **24a**-TFA.

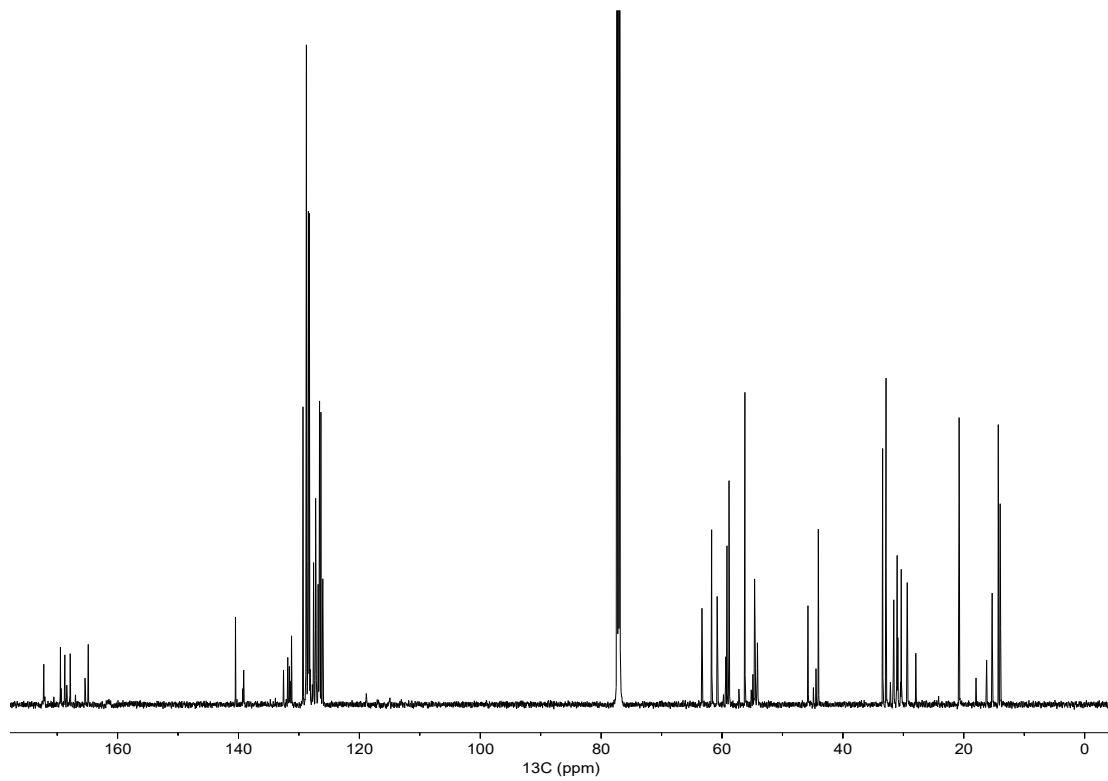


Fig. S71. $^{13}\text{C}\{^1\text{H}\}$ NMR (151 MHz, CDCl_3) for quinapril-TFA, **24a**-TFA.

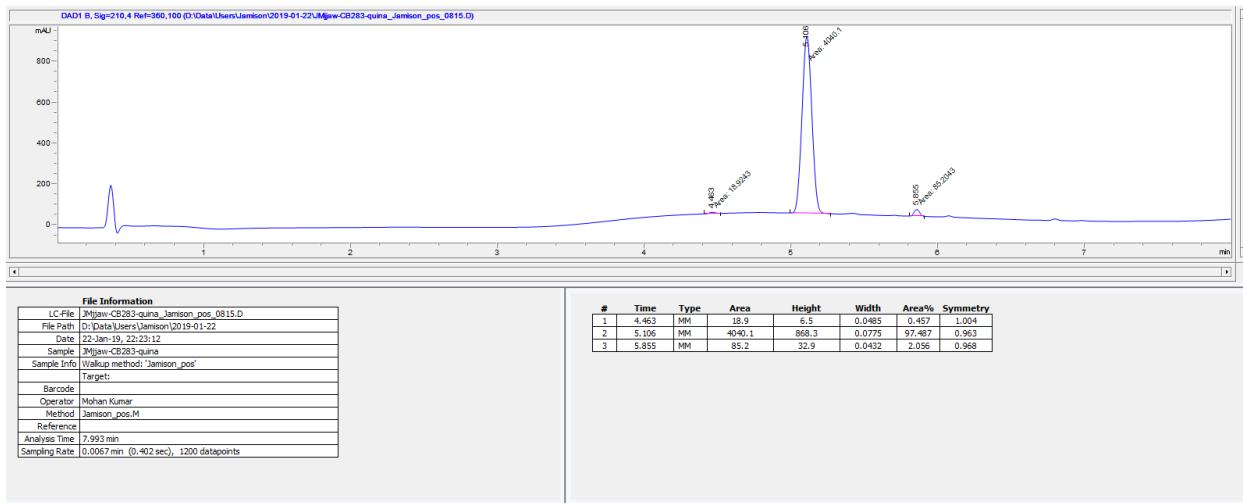


Fig. S72. Purity assessment of quinapril-TFA, **24a**•TFA by HPLC.

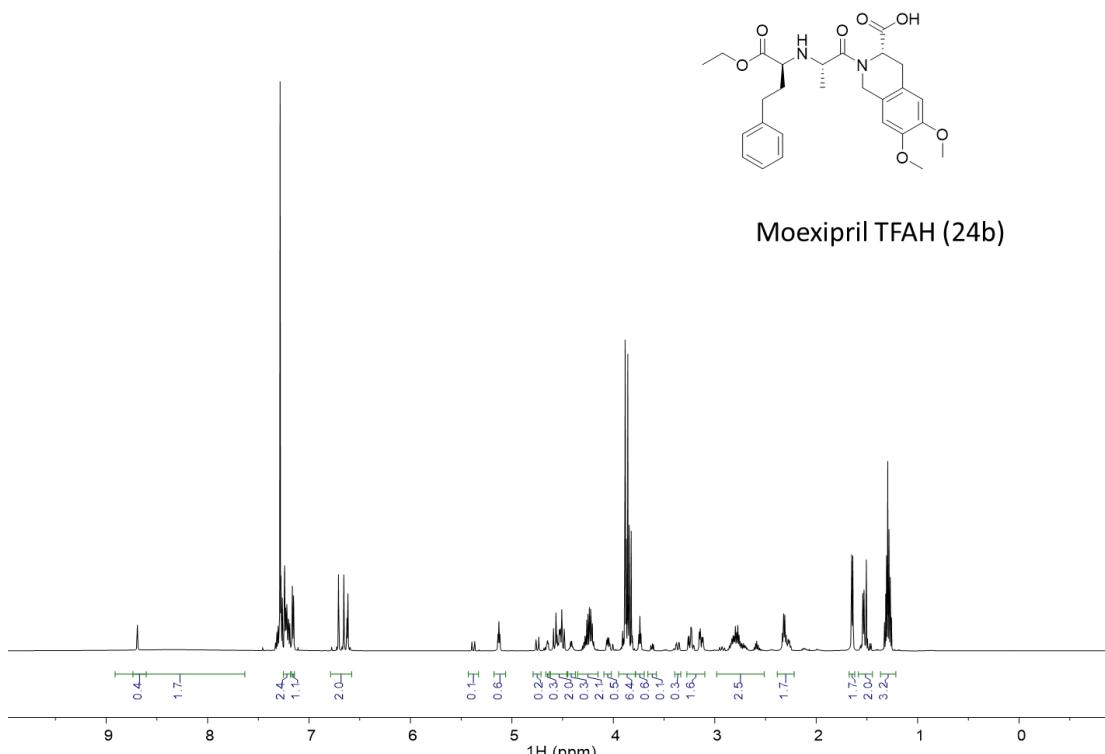


Fig. S73. ^1H NMR (600 MHz, CDCl_3) for moexipril•TFA, **24b**•TFA.

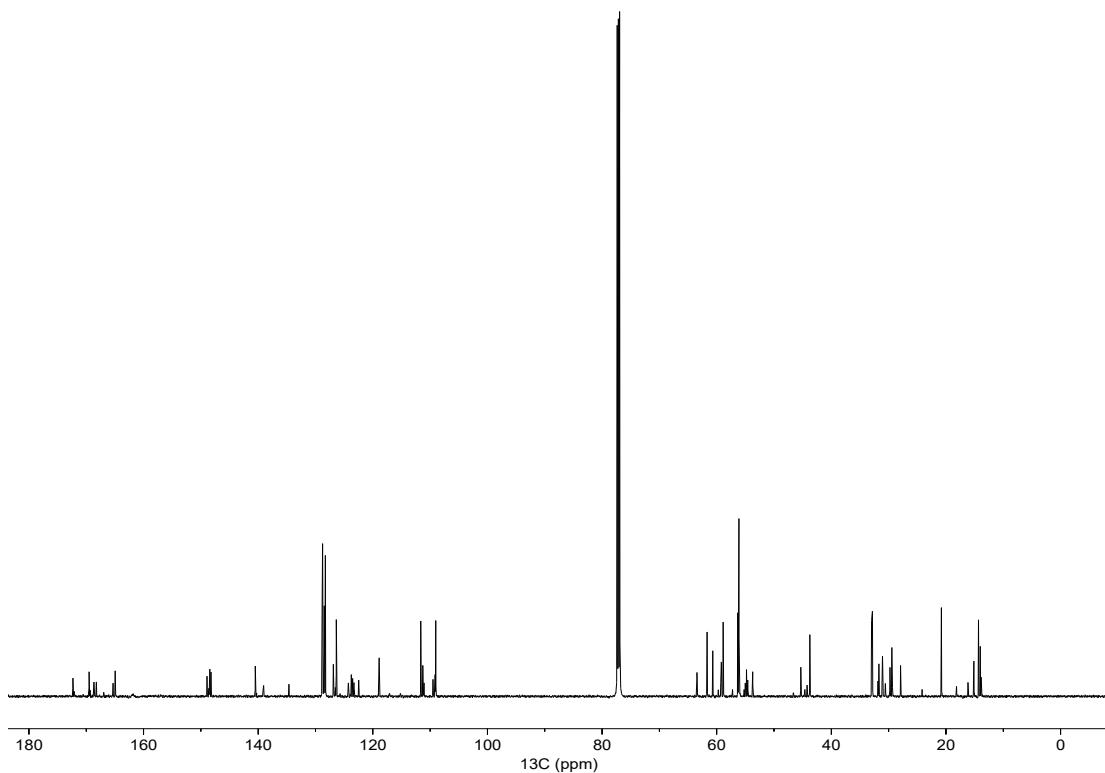


Fig. S74. $^{13}\text{C}\{^1\text{H}\}$ NMR (151 MHz, CDCl_3) for moexipril•TFA, **24b**•TFA.

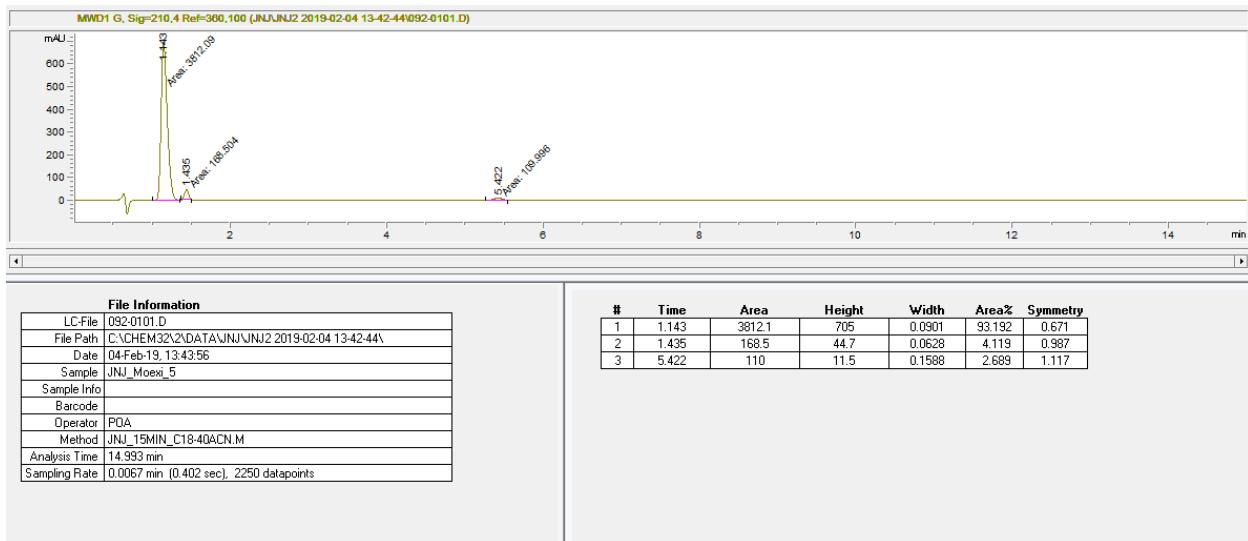


Fig. S75. Purity assessment of moexipril-TFA, **24b**-TFA by HPLC.

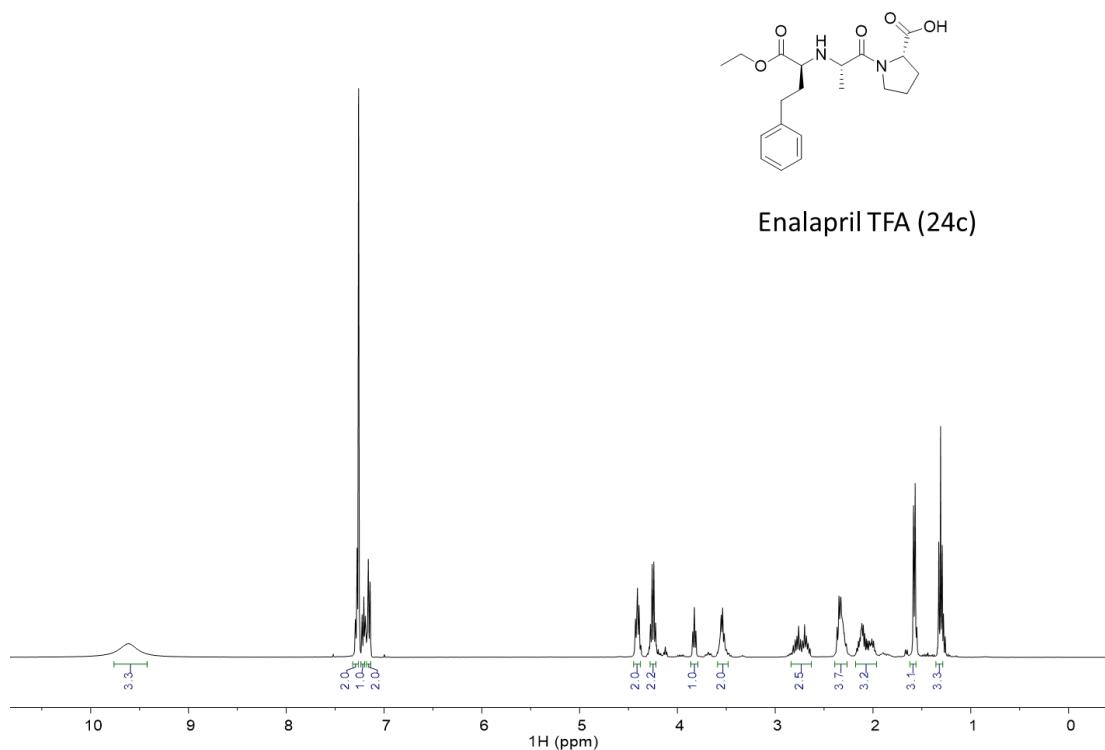


Fig. S76. ${}^1\text{H}$ NMR (600 MHz, CDCl_3) for enalapril•TFA, **24c**•TFA.

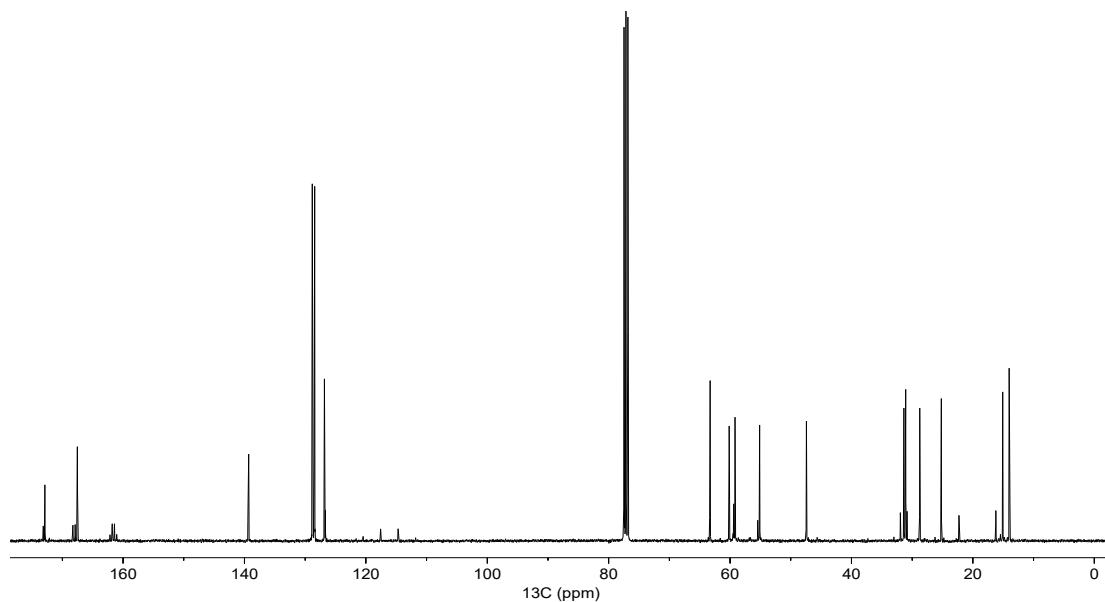


Fig. S77. ${}^{13}\text{C}\{{}^1\text{H}\}$ NMR (151 MHz, CDCl_3) for enalapril•TFA, **24c**•TFA.

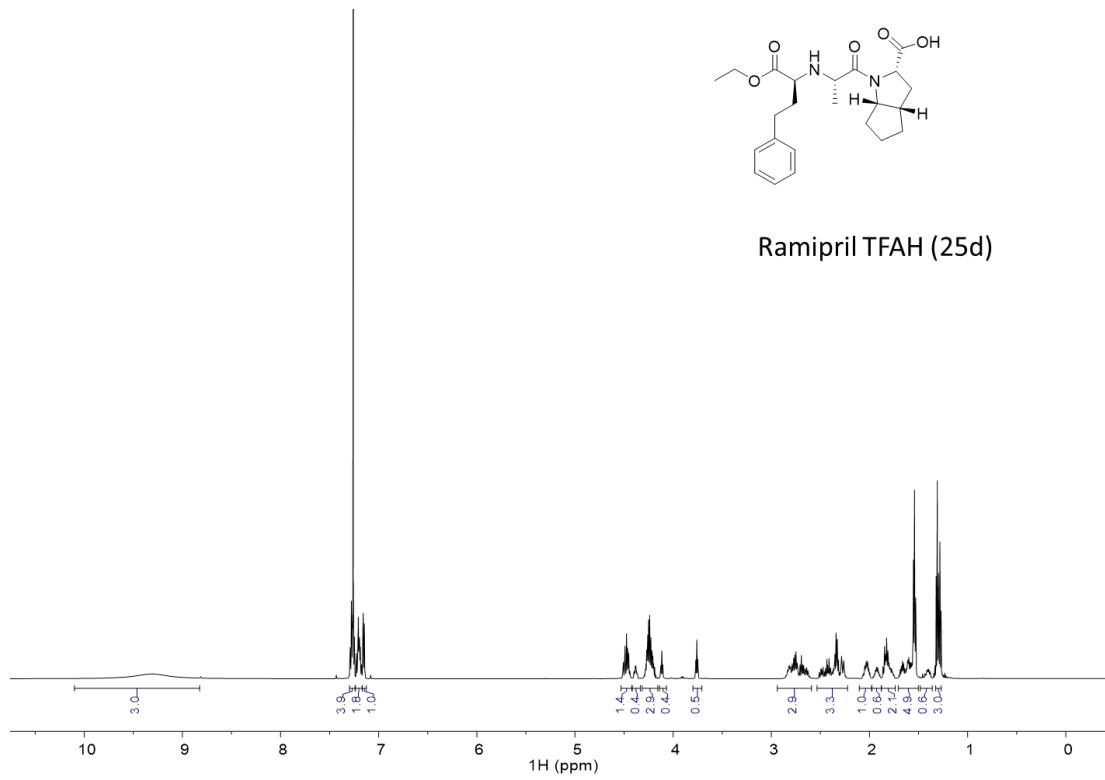


Fig. S78. ^1H NMR (600 MHz, CDCl_3) for ramipril•TFA, **24d**•TFA.

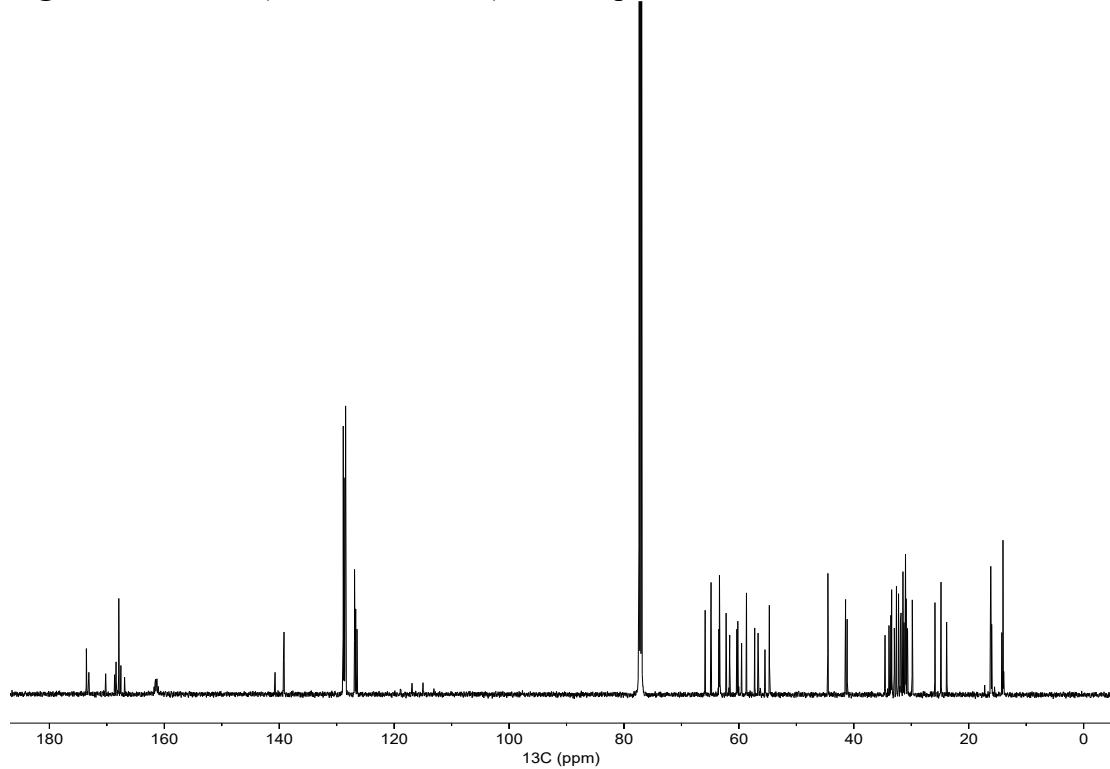


Fig. S79. $^{13}\text{C}\{^1\text{H}\}$ NMR (151 MHz, CDCl_3) for ramipril•TFA, **24d**•TFA.

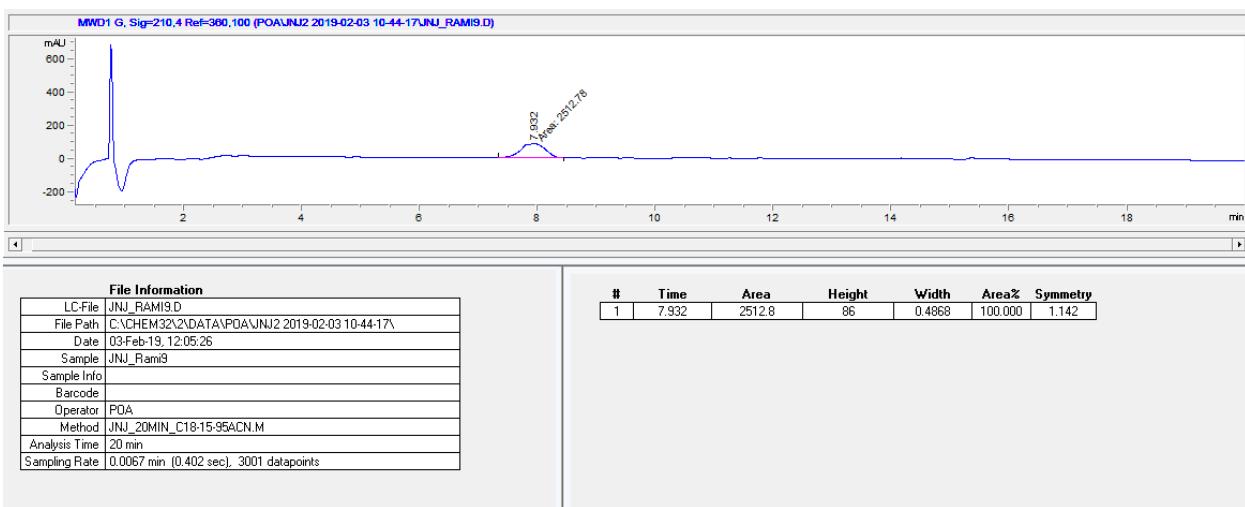
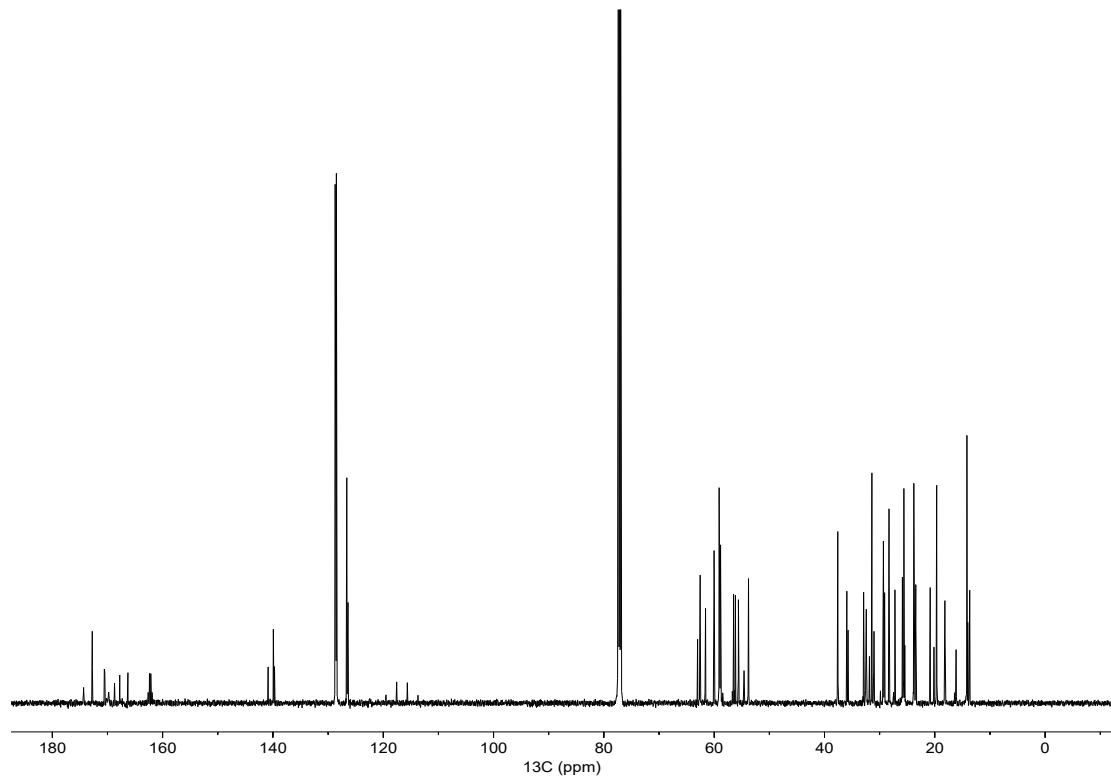
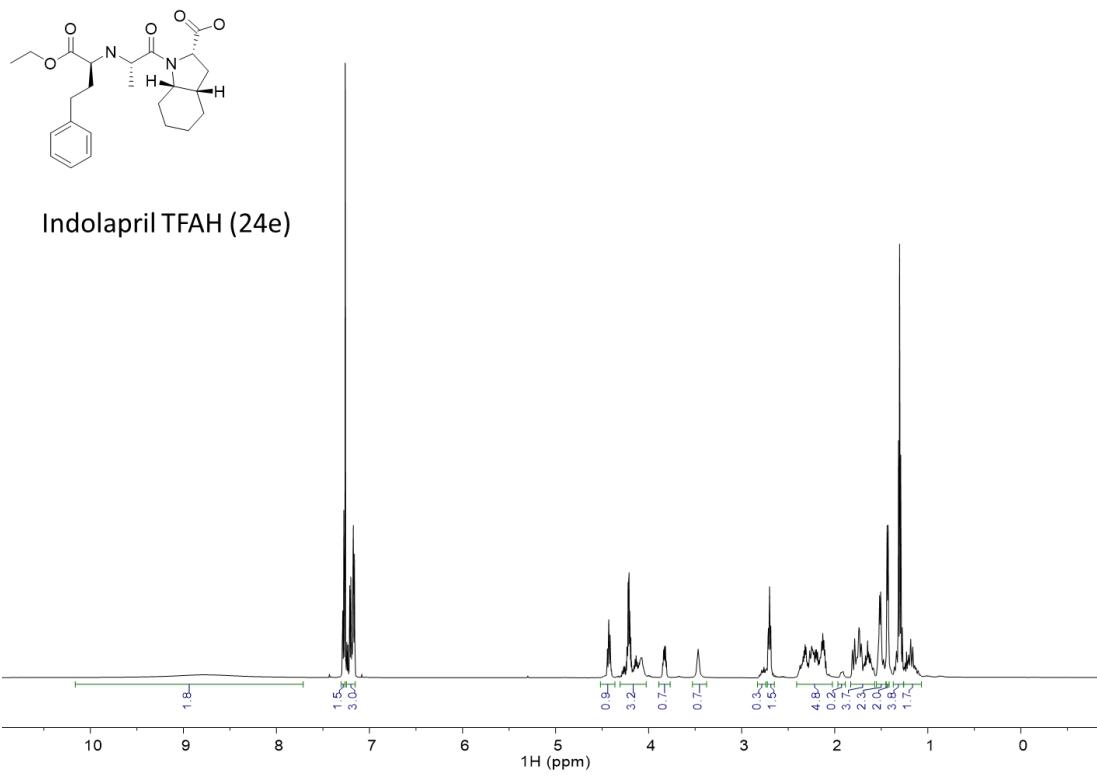


Fig. S80. Purity assessment of ramipril•TFA, **24d**•TFA by HPLC.



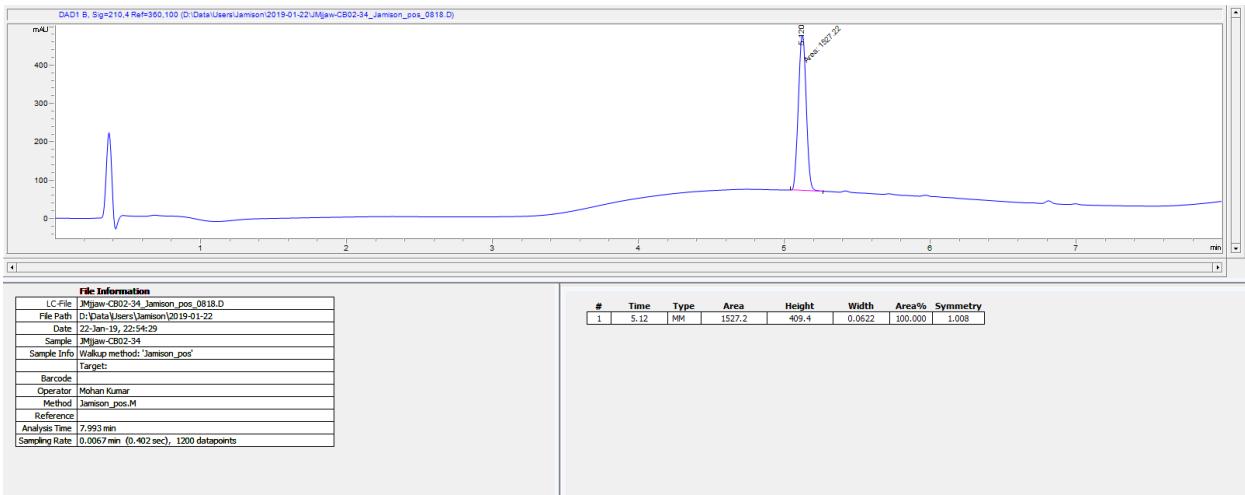


Fig. S83. Purity assessment of indolapril•TFA, **24e**•TFA by HPLC.

Celebrex

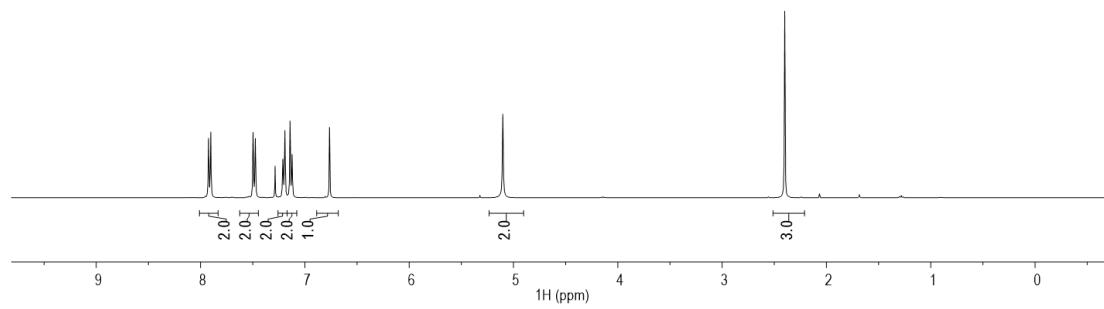
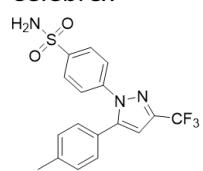


Fig. S84. ^1H NMR (500 MHz, CDCl_3) for celecoxib, **28aa**.

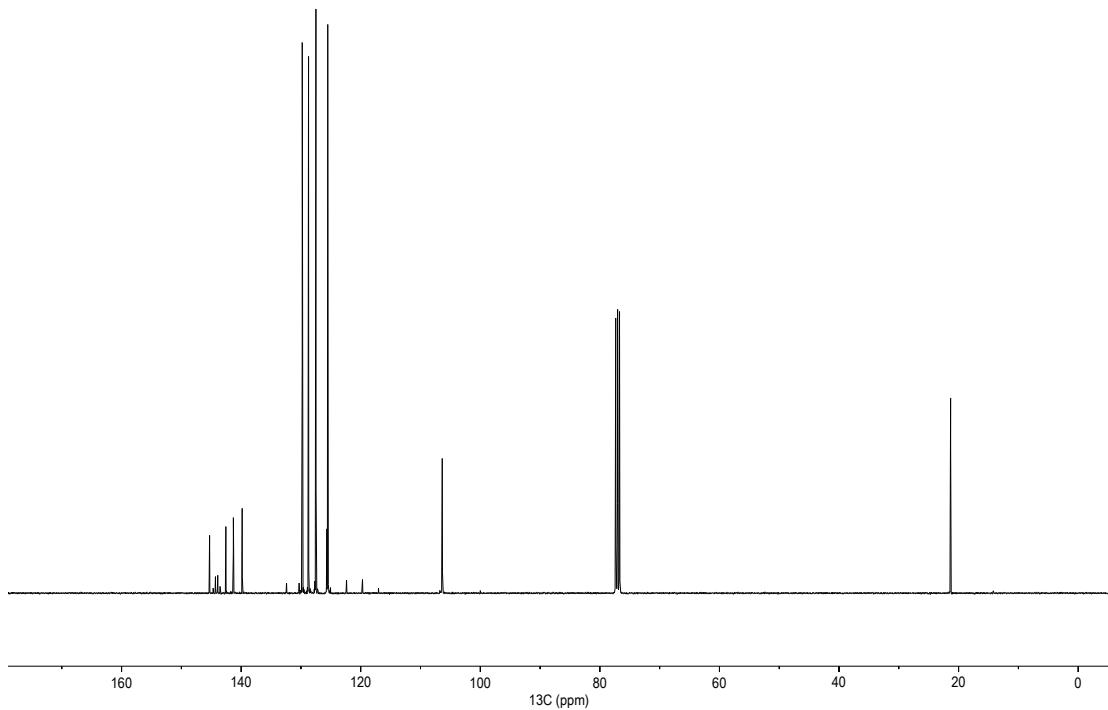


Fig. S85. $^{13}\text{C}\{\text{H}\}$ NMR (126 MHz, CDCl_3) for celecoxib, **28aa**.

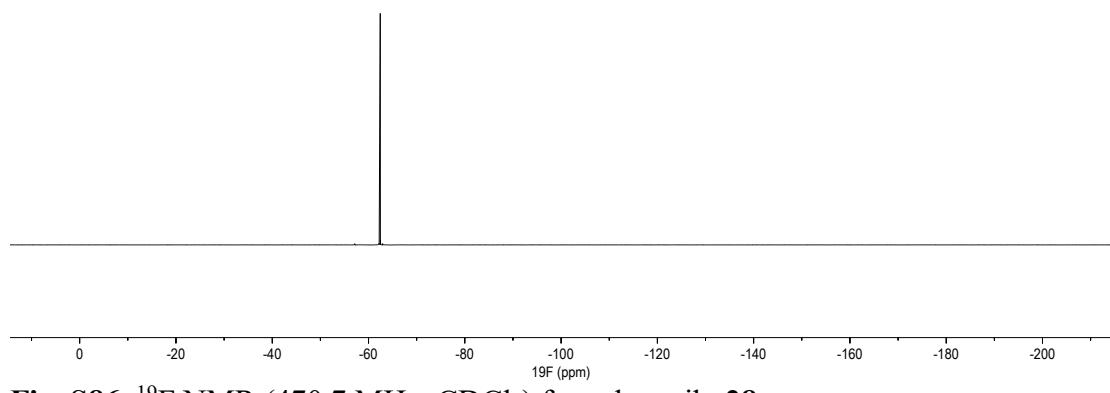


Fig. S86. ¹⁹F NMR (470.7 MHz, CDCl₃) for celecoxib, **28aa**.

SC-558

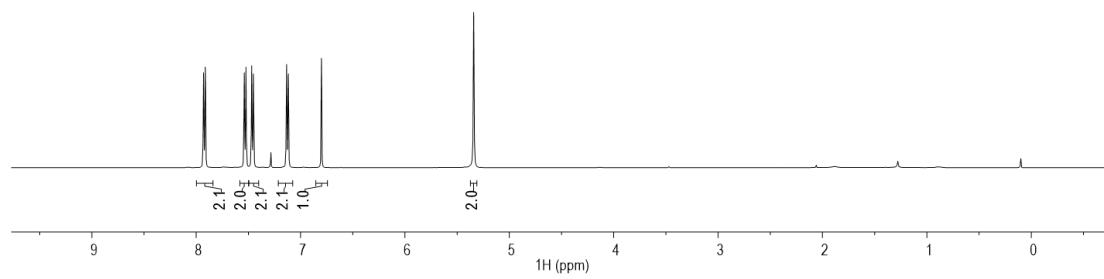
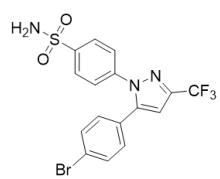


Fig. S87. ^1H NMR (500 MHz, CDCl_3) for **28ab**.

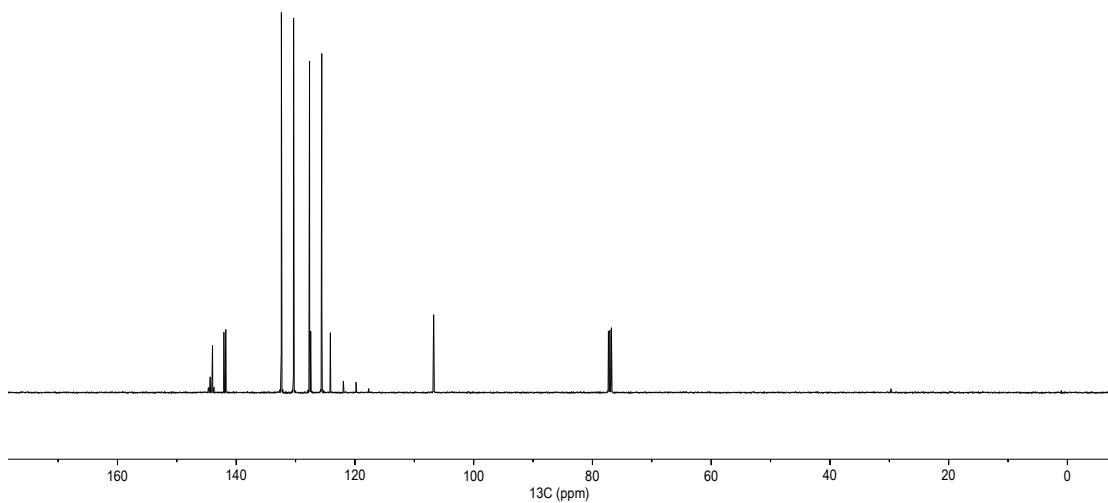


Fig. S88. $^{13}\text{C}\{\text{H}\}$ NMR (126 MHz, CDCl_3) for **28ab**.

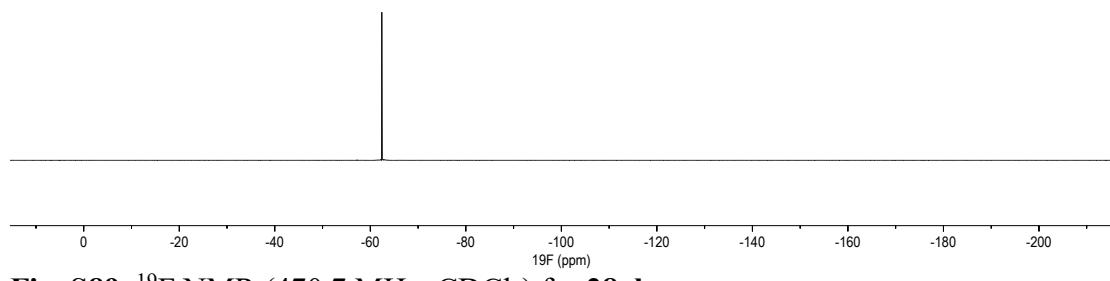


Fig. S89. ^{19}F NMR (470.7 MHz, CDCl_3) for **28ab**.

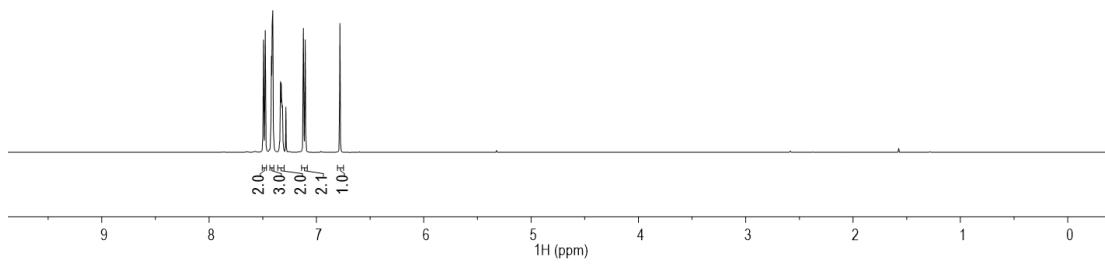
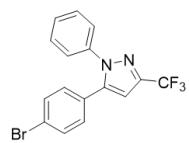


Fig. S90. ¹H NMR (500 MHz, CDCl₃) for **28bb**.

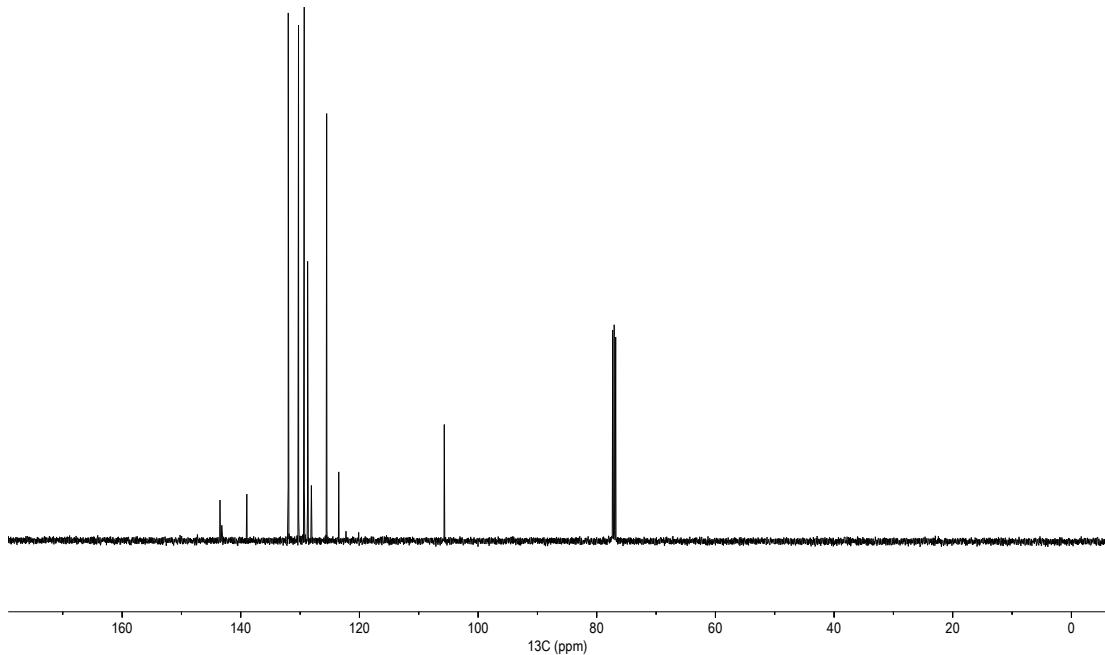


Fig. S91. ¹³C{¹H} NMR (126 MHz, CDCl₃) for **28bb**.

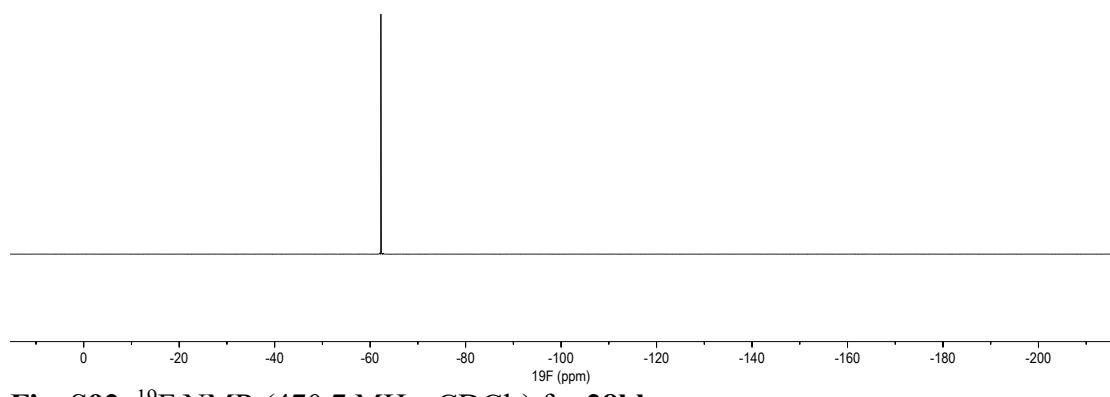


Fig. S92. ¹⁹F NMR (470.7 MHz, CDCl₃) for **28bb**.

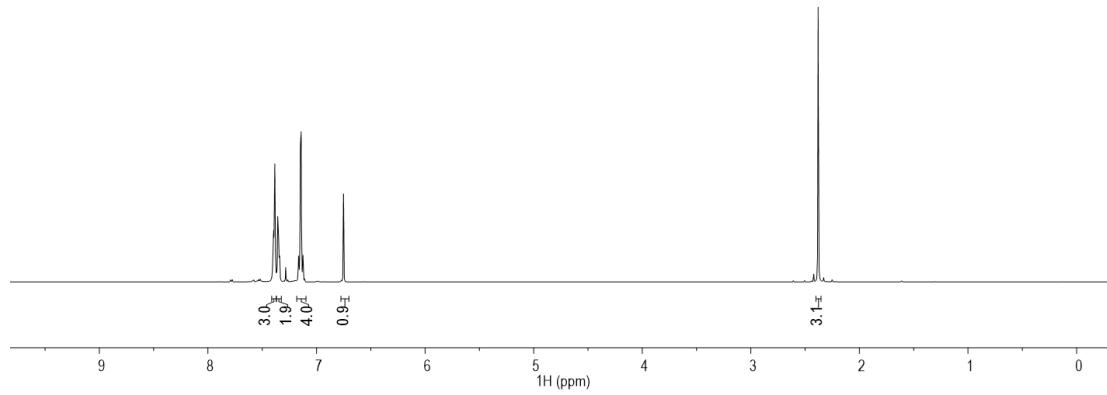
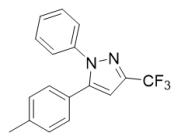


Fig. S93. ^1H NMR (500 MHz, CDCl_3) for **28ba**.

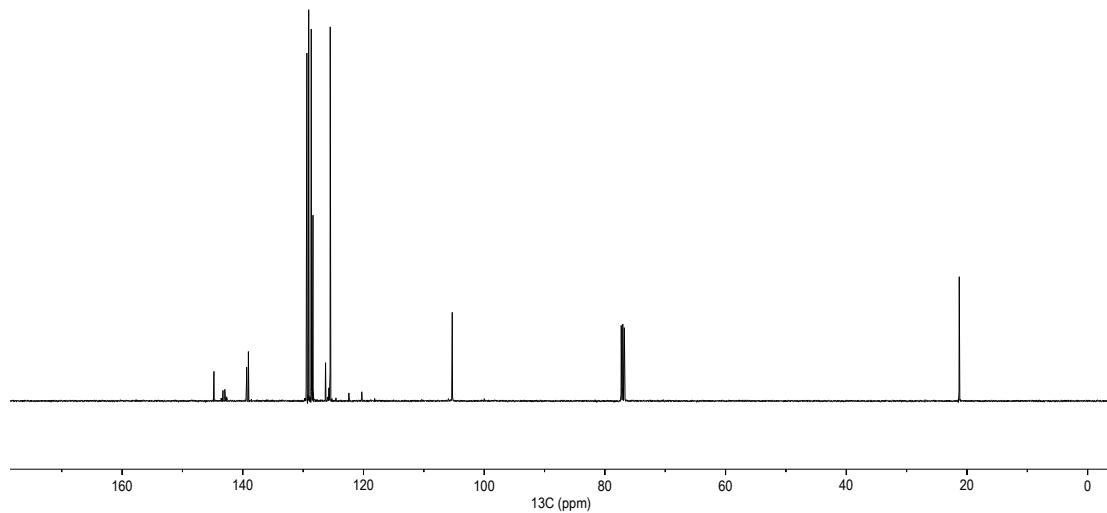


Fig. S94. $^{13}\text{C}\{^1\text{H}\}$ NMR (126 MHz, CDCl_3) for **28ba**.

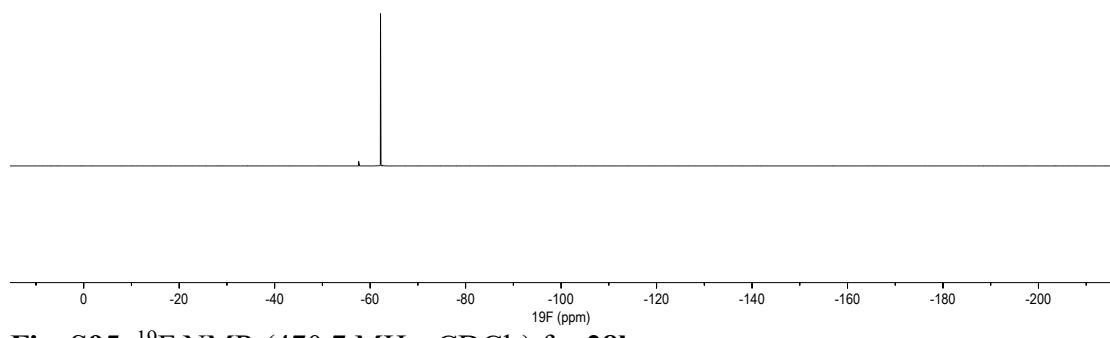


Fig. S95. ¹⁹F NMR (470.7 MHz, CDCl₃) for **28ba**.

S4. Other Captions

Movie S1. An exemplary sequence of the robotic construction and deconstruction of a synthesis process for the case of lidocaine at 30x speed. Steps of reactor selection, reactor placement, reagent line assembly, process module sealing, process stack compression, process module release, reactor disassembly, and reactor storage are highlighted in less-accelerated insets.

Movie S2. Robotic reconfiguration of process modules. Construction and deconstruction of representative process stacks are shown for nine unique configurations required by aspirin, secnidazole, lidocaine, diazepam, (S)-warfarin, safinamide, ACE inhibitor library molecules requiring longer residence times, ACE inhibitor library molecules requiring shorter residence times, and NSAID library molecules.

Movie S3. Robotic reconfiguration of fluidic connections. Construction and deconstruction of representative connections between reagent ports and the reagent tree are shown for nine unique configurations required by aspirin, secnidazole, lidocaine, diazepam, (S)-warfarin, safinamide, ACE inhibitor library molecules requiring longer residence times, ACE inhibitor library molecules requiring shorter residents times, and NSAID library molecules. Within a compound library, one or two reagents are varied through the use of computer-controlled 24-way selector valves, which are not shown and do not need to be robotically manipulated.

Recipe Planning File S1. Process configurations for the nine unique configurations required by aspirin, secnidazole, lidocaine, diazepam, (S)-warfarin, safinamide, ACE inhibitor library molecules requiring longer residence times, ACE inhibitor library molecules requiring shorter residents times, and NSAID library molecules. For each synthesis, this spreadsheet defines which process modules occupy which bays, and which process module ports need to be plumbed to which reagent line. The document also contains operator instructions as to which stock solutions or reagent streams to provide at each location. This summary spreadsheet was used to programmatically generate CRFs.

Recipe Planning Script S1. Python 2.7 script to generate initial Chemical Recipe Files (CRFs) from **Recipe Planning File S1**. A high-level definition of the process configuration and process streams are converted into a series of low level robotic commands for reactor/tubing placement, flow rate setpoints, temperature setpoints, and sequencing of flow rate commands for startup and shutdown. Some CRFs were slightly modified to adjust sequencing of priming steps to mitigate clogging risk (rather than using the default downstream-to-upstream order) and to combine multiple recipes into back-to-back files for syntheses run back-to-back.

Recipe S1. Chemical Recipe File for the back-to-back synthesis of aspirin and secnidazole.

Recipe S2. Chemical Recipe File for the back-to-back synthesis of lidocaine and diazepam.

Recipe S3. Chemical Recipe File for the synthesis of (S)-warfarin.

Recipe S4. Chemical Recipe File for the synthesis of safinamide.

Recipe S5. Chemical Recipe File for the ACE inhibitor library (long residence time configuration).

Recipe S6. Chemical Recipe File for the ACE inhibitor library (short residence time configuration).

Recipe S7. Chemical Recipe File for the NSAID library.

Batch Sheet S1. Operator instructions for preparing stock solutions for the synthesis of aspirin.

Batch Sheet S2. Operator instructions for preparing stock solutions for the synthesis of secnidazole.

Batch Sheet S3. Operator instructions for preparing stock solutions for the synthesis of lidocaine.

Batch Sheet S4. Operator instructions for preparing stock solutions for the synthesis of diazepam.

Batch Sheet S5. Operator instructions for preparing stock solutions for the synthesis of (S)-warfarin.

Batch Sheet S6. Operator instructions for preparing stock solutions for the synthesis of safinamide.

Batch Sheet S7. Operator instructions for preparing stock solutions for the synthesis of quinapril.

Batch Sheet S8. Operator instructions for preparing stock solutions for the synthesis of moexipril.

Batch Sheet S9. Operator instructions for preparing stock solutions for the synthesis of enalapril.

Batch Sheet S10. Operator instructions for preparing stock solutions for the synthesis of ramipril.

Batch Sheet S11. Operator instructions for preparing stock solutions for the synthesis of indolapril.

Batch Sheet S12. Operator instructions for preparing stock solutions for the synthesis of celecoxib.

Batch Sheet S13. Operator instructions for preparing stock solutions for the synthesis of **28ab**.

Batch Sheet S14. Operator instructions for preparing stock solutions for the synthesis of **28bb**.

Batch Sheet S15. Operator instructions for preparing stock solutions for the synthesis of **28ba**.

Website Printout S1. Website printout of the synthetic pathway results for aspirin. Expansion settings for the tree builder are shown at the top of the page.

Website Printout S2. Website printout of the reaction conditions recommended for the first step in the synthesis of aspirin.

Website Printout S3. Website printout of the forward prediction results for the first step in the synthesis of aspirin, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S4. Website printout of the synthetic pathway results for secnidazole. Expansion settings for the tree builder are shown at the top of the page.

Website Printout S5. Website printout of the reaction conditions recommended for the first step in the synthesis of secnidazole.

Website Printout S6. Website printout of the forward prediction results for the first step in the synthesis of secnidazole, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S7. Website printout of the synthetic pathway results for lidocaine. Expansion settings for the tree builder are shown at the top of the page.

Website Printout S8. Website printout of the reaction conditions recommended for the first step in the synthesis of lidocaine.

Website Printout S9. Website printout of the forward prediction results for the first step in the synthesis of lidocaine, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S10. Website printout of the reaction conditions recommended for the second step in the synthesis of lidocaine.

Website Printout S11. Website printout of the forward prediction results for the second step in the synthesis of lidocaine, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S12. Website printout of the synthetic pathway results for diazepam. Expansion settings for the tree builder are shown at the top of the page.

Website Printout S13. Website printout of the reaction conditions recommended for the first step in the synthesis of diazepam.

Website Printout S14. Website printout of the forward prediction results for the first step in the synthesis of diazepam, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S15. Website printout of the reaction conditions recommended for the second step in the synthesis of diazepam.

Website Printout S16. Website printout of the forward prediction results for the second step in the synthesis of diazepam, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S17. Website printout of the synthetic pathway results for (S)-warfarin. Expansion settings for the tree builder are shown at the top of the page.

Website Printout S18. Website printout of the reaction conditions recommended for the first step in the synthesis of (S)-warfarin.

Website Printout S19. Website printout of the forward prediction results for the first step in the synthesis of (S)-warfarin, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S20. Website printout of the synthetic pathway results for safinamide. Expansion settings for the tree builder are shown at the top of the page.

Website Printout S21. Website printout of the reaction conditions recommended for the first step in the synthesis of safinamide.

Website Printout S22. Website printout of the forward prediction results for the first step in the synthesis of safinamide, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S23. Website printout of the reaction conditions recommended for the second step in the synthesis of safinamide.

Website Printout S24. Website printout of the forward prediction results for the second step in the synthesis of safinamide, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S25. Website printout of the synthetic pathway results for quinapril. Expansion settings for the tree builder are shown at the top of the page.

Website Printout S26. Website printout of the reaction conditions recommended for the first step in the synthesis of quinapril.

Website Printout S27. Website printout of the forward prediction results for the first step in the synthesis of quinapril, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S28. Website printout of the reaction conditions recommended for the second step in the synthesis of quinapril.

Website Printout S29. Website printout of the forward prediction results for the second step in the synthesis of quinapril, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S30. Website printout of the synthetic pathway results for celecoxib. Expansion settings for the tree builder are shown at the top of the page.

Website Printout S31. Website printout of the reaction conditions recommended for the first step in the synthesis of celecoxib.

Website Printout S32. Website printout of the forward prediction results for the first step in the synthesis of celecoxib, used to understand the likelihood of experimental success and potential side reactions.

Website Printout S33. Website printout of the reaction conditions recommended for the second step in the synthesis of celecoxib.

Website Printout S34. Website printout of the forward prediction results for the second step in the synthesis of celecoxib, used to understand the likelihood of experimental success and potential side reactions.

References and Notes

1. D. C. Blakemore, L. Castro, I. Churcher, D. C. Rees, A. W. Thomas, D. M. Wilson, A. Wood, Organic synthesis provides opportunities to transform drug discovery. *Nat. Chem.* **10**, 383–394 (2018). [doi:10.1038/s41557-018-0021-z](https://doi.org/10.1038/s41557-018-0021-z) [Medline](#)
2. G. Schneider, Automating drug discovery. *Nat. Rev. Drug Discov.* **17**, 97–113 (2018). [doi:10.1038/nrd.2017.232](https://doi.org/10.1038/nrd.2017.232) [Medline](#)
3. D. P. Tabor, L. M. Roch, S. K. Saikin, C. Kreisbeck, D. Sheberla, J. H. Montoya, S. Dwarakanath, M. Aykol, C. Ortiz, H. Tribukait, C. Amador-Bedolla, C. J. Brabec, B. Maruyama, K. A. Persson, A. Aspuru-Guzik, Accelerating the discovery of materials for clean energy in the era of smart automation. *Nat. Rev. Mater.* **3**, 5–20 (2018). [doi:10.1038/s41578-018-0005-z](https://doi.org/10.1038/s41578-018-0005-z)
4. M. Peplow, Organic synthesis: The robo-chemist. *Nature* **512**, 20–22 (2014). [doi:10.1038/512020a](https://doi.org/10.1038/512020a) [Medline](#)
5. S. V. Ley, D. E. Fitzpatrick, R. J. Ingham, R. M. Myers, Organic synthesis: March of the machines. *Angew. Chem. Int. Ed.* **54**, 3449–3464 (2015). [doi:10.1002/anie.201410744](https://doi.org/10.1002/anie.201410744) [Medline](#)
6. A. G. Godfrey, T. Masquelin, H. Hemmerle, A remote-controlled adaptive medchem lab: An innovative approach to enable drug discovery in the 21st Century. *Drug Discov. Today* **18**, 795–802 (2013). [doi:10.1016/j.drudis.2013.03.001](https://doi.org/10.1016/j.drudis.2013.03.001) [Medline](#)
7. A.-C. Bédard, A. Adamo, K. C. Aroh, M. G. Russell, A. A. Bedermann, J. Torosian, B. Yue, K. F. Jensen, T. F. Jamison, Reconfigurable system for automated optimization of diverse chemical reactions. *Science* **361**, 1220–1225 (2018). [doi:10.1126/science.aat0650](https://doi.org/10.1126/science.aat0650) [Medline](#)
8. S. Steiner, J. Wolf, S. Glatzel, A. Andreou, J. M. Granda, G. Keenan, T. Hinkley, G. Aragon-Camarasa, P. J. Kitson, D. Angelone, L. Cronin, Organic synthesis in a modular robotic system driven by a chemical programming language. *Science* **363**, eaav2211 (2019). [doi:10.1126/science.aav2211](https://doi.org/10.1126/science.aav2211) [Medline](#)
9. J. Li, S. G. Ballmer, E. P. Gillis, S. Fujii, M. J. Schmidt, A. M. E. Palazzolo, J. W. Lehmann, G. F. Morehouse, M. D. Burke, Synthesis of many different types of organic small molecules using one automated process. *Science* **347**, 1221–1226 (2015). [doi:10.1126/science.aaa5414](https://doi.org/10.1126/science.aaa5414) [Medline](#)
10. A. Buitrago Santanilla, E. L. Regalado, T. Pereira, M. Shevlin, K. Bateman, L.-C. Campeau, J. Schneeweis, S. Berritt, Z.-C. Shi, P. Nantermet, Y. Liu, R. Helmy, C. J. Welch, P. Vachal, I. W. Davies, T. Cernak, S. D. Dreher, Organic chemistry. Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. *Science* **347**, 49–53 (2015). [doi:10.1126/science.1259203](https://doi.org/10.1126/science.1259203) [Medline](#)
11. D. Perera, J. W. Tucker, S. Brahmbhatt, C. J. Helal, A. Chong, W. Farrell, P. Richardson, N. W. Sach, A platform for automated nanomole-scale reaction screening and micromole-scale synthesis in flow. *Science* **359**, 429–434 (2018). [doi:10.1126/science.aap9112](https://doi.org/10.1126/science.aap9112) [Medline](#)

12. Food and Drug Administration, Strategic plan for preventing and mitigating drug shortages (2013);
<https://www.fda.gov/downloads/Drugs/DrugSafety/DrugShortages/UCM372566.pdf>
13. R. L. Hartman, J. P. McMullen, K. F. Jensen, Deciding whether to go with the flow: Evaluating the merits of flow reactors for synthesis. *Angew. Chem. Int. Ed.* **50**, 7502–7519 (2011). [doi:10.1002/anie.201004637](https://doi.org/10.1002/anie.201004637) [Medline](#)
14. M. B. Plutschack, B. Pieber, K. Gilmore, P. H. Seeberger, The Hitchhiker's Guide to Flow Chemistry II. *Chem. Rev.* **117**, 11796–11893 (2017). [doi:10.1021/acs.chemrev.7b00183](https://doi.org/10.1021/acs.chemrev.7b00183) [Medline](#)
15. K. F. Jensen, Flow chemistry-Microreaction technology comes of age. *AICHE J.* **63**, 858–869 (2017). [doi:10.1002/aic.15642](https://doi.org/10.1002/aic.15642)
16. B. Gutmann, D. Cantillo, C. O. Kappe, Continuous-flow technology—A tool for the safe manufacturing of active pharmaceutical ingredients. *Angew. Chem. Int. Ed.* **54**, 6688–6728 (2015). [doi:10.1002/anie.201409318](https://doi.org/10.1002/anie.201409318) [Medline](#)
17. C. Wiles, P. Watts, Continuous flow reactors: A perspective. *Green Chem.* **14**, 38–54 (2012). [doi:10.1039/C1GC16022B](https://doi.org/10.1039/C1GC16022B)
18. D. R. Snead, T. F. Jamison, A three-minute synthesis and purification of ibuprofen: Pushing the limits of continuous-flow processing. *Angew. Chem. Int. Ed.* **54**, 983–987 (2015). [doi:10.1002/anie.201409093](https://doi.org/10.1002/anie.201409093) [Medline](#)
19. T. Klucznik, B. Mikulak-Klucznik, M. P. McCormack, H. Lima, S. Szymkuć, M. Bhowmick, K. Molga, Y. Zhou, L. Rickershauser, E. P. Gajewska, A. Toutchkine, P. Dittwald, M. P. Startek, G. J. Kirkovits, R. Roszak, A. Adamski, B. Sieredzińska, M. Mrksich, S. L. J. Trice, B. A. Grzybowski, Efficient Syntheses of Diverse, Medicinally Relevant Targets Planned by Computer and Executed in the Laboratory, *Chem* **4**, 522–532 (2018). [doi:10.1016/j.chempr.2018.02.002](https://doi.org/10.1016/j.chempr.2018.02.002)
20. S. Szymkuć, E. P. Gajewska, T. Klucznik, K. Molga, P. Dittwald, M. Startek, M. Bajczyk, B. A. Grzybowski, Computer-Assisted Synthetic Planning: The End of the Beginning. *Angew. Chem. Int. Ed.* **55**, 5904–5937 (2016). [doi:10.1002/anie.201506101](https://doi.org/10.1002/anie.201506101) [Medline](#)
21. M. H. S. Segler, M. Preuss, M. P. Waller, Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **555**, 604–610 (2018). [doi:10.1038/nature25978](https://doi.org/10.1038/nature25978) [Medline](#)
22. A. Adamo, R. L. Beingessner, M. Behnam, J. Chen, T. F. Jamison, K. F. Jensen, J.-C. M. Monbaliu, A. S. Myerson, E. M. Revalor, D. R. Snead, T. Stelzer, N. Weeranoppanant, S. Y. Wong, P. Zhang, On-demand continuous-flow production of pharmaceuticals in a compact, reconfigurable system. *Science* **352**, 61–67 (2016). [doi:10.1126/science.aaf1337](https://doi.org/10.1126/science.aaf1337) [Medline](#)
23. B. J. Reizman, K. F. Jensen, Feedback in Flow for Accelerated Reaction Development. *Acc. Chem. Res.* **49**, 1786–1796 (2016). [doi:10.1021/acs.accounts.6b00261](https://doi.org/10.1021/acs.accounts.6b00261) [Medline](#)
24. D. E. Fitzpatrick, C. Battilocchio, S. V. Ley, A Novel Internet-Based Reaction Monitoring, Control and Autonomous Self-Optimization Platform for Chemical Synthesis. *Org. Process Res. Dev.* **20**, 386–394 (2016). [doi:10.1021/acs.oprd.5b00313](https://doi.org/10.1021/acs.oprd.5b00313)

25. E. J. Corey, W. L. Jorgensen, Computer-assisted synthetic analysis. Synthetic strategies based on appendages and the use of reconnective transforms. *J. Am. Chem. Soc.* **98**, 189–203 (1976). [doi:10.1021/ja00417a030](https://doi.org/10.1021/ja00417a030)
26. A. Cook, A. P. Johnson, J. Law, M. Mirzazadeh, O. Ravitz, A. Simon, Computer-aided synthesis design: 40 years on. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 79–107 (2012). [doi:10.1002/wcms.61](https://doi.org/10.1002/wcms.61)
27. C. W. Coley, R. Barzilay, T. S. Jaakkola, W. H. Green, K. F. Jensen, Prediction of Organic Reaction Outcomes Using Machine Learning. *ACS Cent. Sci.* **3**, 434–443 (2017). [doi:10.1021/acscentsci.7b00064](https://doi.org/10.1021/acscentsci.7b00064) [Medline](#)
28. M. H. S. Segler, M. P. Waller, Neural-Symbolic Machine Learning for Retrosynthesis and Reaction Prediction. *Chemistry* **23**, 5966–5971 (2017). [doi:10.1002/chem.201605499](https://doi.org/10.1002/chem.201605499) [Medline](#)
29. B. Liu, B. Ramsundar, P. Kawthekar, J. Shi, J. Gomes, Q. Luu Nguyen, S. Ho, J. Sloane, P. Wender, V. Pande, Retrosynthetic Reaction Prediction Using Neural Sequence-to-Sequence Models. *ACS Cent. Sci.* **3**, 1103–1113 (2017). [doi:10.1021/acscentsci.7b00303](https://doi.org/10.1021/acscentsci.7b00303) [Medline](#)
30. C. W. Coley, W. H. Green, K. F. Jensen, Machine Learning in Computer-Aided Synthesis Planning. *Acc. Chem. Res.* **51**, 1281–1289 (2018). [doi:10.1021/acs.accounts.8b00087](https://doi.org/10.1021/acs.accounts.8b00087) [Medline](#)
31. D. Lowe, Chemical reactions from US patents (1976-Sep2016) (2017); https://figshare.com/articles/Chemical_reactions_from_US_patents_1976-Sep2016/5104873.
32. G. Landrum, RDKit: Open-source cheminformatics; <http://www.rdkit.org>.
33. C. W. Coley, W. H. Green, K. F. Jensen, RDChiral: An RDKit Wrapper for Handling Stereochemistry in Retrosynthetic Template Extraction and Application. *J. Chem. Inf. Model.* **59**, 2529–2537 (2019). [doi:10.1021/acs.jcim.9b00286](https://doi.org/10.1021/acs.jcim.9b00286) [Medline](#)
34. D. Rogers, M. Hahn, Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **50**, 742–754 (2010). [doi:10.1021/ci100050t](https://doi.org/10.1021/ci100050t) [Medline](#)
35. G. M. J.-B. Chaslot, M. H. M. Winands, H. J. van den Herik, in *Computers and Games*, H. J. van den Herik, X. Xu, Z. Ma, M. H. M. Winands, Eds. (Springer, 2008), pp. 60–71.
36. D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, D. Hassabis, Mastering chess and shogi by self-play with a general reinforcement learning algorithm. [arXiv:1712.01815](https://arxiv.org/abs/1712.01815) (2017).
37. C. W. Coley, W. Jin, L. Rogers, T. F. Jamison, T. S. Jaakkola, W. H. Green, R. Barzilay, K. F. Jensen, A graph-convolutional neural network model for the prediction of chemical reactivity. *Chem. Sci.* **10**, 370–377 (2018). [doi:10.1039/C8SC04228D](https://doi.org/10.1039/C8SC04228D) [Medline](#)
38. W. Jin, C. Coley, R. Barzilay, T. Jaakkola, Predicting organic reaction outcomes with Weisfeiler-Lehman network, in *Advances in Neural Information Processing Systems* (NeurIPS, 2017), pp. 2604–2613.

39. H. Gao, T. J. Struble, C. W. Coley, Y. Wang, W. H. Green, K. F. Jensen, Using Machine Learning To Predict Suitable Conditions for Organic Reactions. *ACS Cent. Sci.* **4**, 1465–1476 (2018). [doi:10.1021/acscentsci.8b00357](https://doi.org/10.1021/acscentsci.8b00357) [Medline](#)
40. P. Zhang, N. Weeranoppanant, D. A. Thomas, K. Tahara, T. Stelzer, M. G. Russell, M. O’Mahony, A. S. Myerson, H. Lin, L. P. Kelly, K. F. Jensen, T. F. Jamison, C. Dai, Y. Cui, N. Briggs, R. L. Beingessner, A. Adamo, Advanced Continuous Flow Platform for On-Demand Pharmaceutical Manufacturing. *Chemistry* **24**, 2776–2784 (2018). [doi:10.1002/chem.201706004](https://doi.org/10.1002/chem.201706004) [Medline](#)
41. A. Adamo, P. L. Heider, N. Weeranoppanant, K. F. Jensen, Membrane-Based, Liquid–Liquid Separator with Integrated Pressure Control. *Ind. Eng. Chem. Res.* **52**, 10802–10808 (2013). [doi:10.1021/ie401180t](https://doi.org/10.1021/ie401180t)
42. Y.-J. Hwang, C. W. Coley, M. Abolhasani, A. L. Marzinzik, G. Koch, C. Spanka, H. Lehmann, K. F. Jensen, A segmented flow platform for on-demand medicinal chemistry and compound synthesis in oscillating droplets. *Chem. Commun.* **53**, 6649–6652 (2017). [doi:10.1039/C7CC03584E](https://doi.org/10.1039/C7CC03584E) [Medline](#)
43. C. Coley, ASKCOS Version 0.2.9, Zenodo (2019); <https://doi.org/10.5281/zenodo.3261361>.
44. P. Schwaller, T. Gaudin, D. Lányi, C. Bekas, T. Laino, “Found in Translation”: Predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models. *Chem. Sci.* **9**, 6091–6098 (2018). [doi:10.1039/C8SC02339E](https://doi.org/10.1039/C8SC02339E) [Medline](#)
45. K. P. Hawkins, Analytic Inverse Kinematics for the Universal Robots UR-5/UR-10 Arms (Georgia Institute of Technology, 2013); <https://smartech.gatech.edu/handle/1853/50782>.
46. J. Law, Z. Zsoldos, A. Simon, D. Reid, Y. Liu, S. Y. Khew, A. P. Johnson, S. Major, R. A. Wade, H. Y. Ando, Route Designer: A retrosynthetic analysis tool utilizing automated retrosynthetic rule generation. *J. Chem. Inf. Model.* **49**, 593–602 (2009). [doi:10.1021/ci800228y](https://doi.org/10.1021/ci800228y) [Medline](#)
47. C. D. Christ, M. Zentgraf, J. M. Kriegl, Mining electronic laboratory notebooks: Analysis, retrosynthesis, and reaction based enumeration. *J. Chem. Inf. Model.* **52**, 1745–1756 (2012). [doi:10.1021/ci300116p](https://doi.org/10.1021/ci300116p) [Medline](#)
48. W. Jin, R. Barzilay, T. Jaakkola, Junction tree variational autoencoder for molecular graph generation. [arXiv:1802.04364](https://arxiv.org/abs/1802.04364) (2018).
49. T. Nguyen, Drug structures made public in New Orleans. Chemical & Engineering News (2018); <https://cen.acs.org/articles/96/web/2018/03/Drug-structures-made-public-New-Orleans.html>.