SUPPORTING INFORMATION

Novel Reagent Space: Identifying Unorderable but Readily Synthesizable Building Blocks

Mark Seierstad[1*], Mark S. Tichenor[1], Renee L. DesJarlais[2], Jim Na[1], Genesis M. Bacani[1], De Michael Chung[1], Eduardo V. Mercado-Marin[1], Helena C. Steffens[1], Taraneh Mirzadegan[1]

[1]Janssen Research and Development, 3210 Merryfield Row, San Diego, CA 92121, United States
[2]Janssen Research and Development, 1400 McKean Road, Spring House, PA 19477, United States

Table of Contents

Table S1. Curated list of reagent vendors and information about pre-filtering

Asinex
Bionet-Key Organics
ChemBridge
ChemDiv
ChemShuttle
Chemtellect
Combi-Blocks
CombiPhos D-Reagent
ComInnex DEL
eMolecules Tiers 0-2
Enamine
HitGen
Life Chemicals
Liverpool ChiroChem
Maybridge
Menai Organics
PharmaBlocks
Princeton BioMolecular
Sigma-Aldrich
SpiroChem
Wuxi

Pipeline Pilot (Dassault Systèmes BIOVIA, BIOVIA Pipeline Pilot, Release 2020, San Diego: Dassault Systèmes, 2020) is used to modify/filter according to:

- Remove salts, neutralize acids and bases
- Remove duplicate structures
- Apply MW cutoff, with filters to account for large protecting groups (e.g. FMOC or trityl)
- Apply organic filter, a modified version of the Pipeline Pilot organic filter component
- Remove undesirable atoms and isotopes
- Remove crown ethers and compounds with 4 or more rings

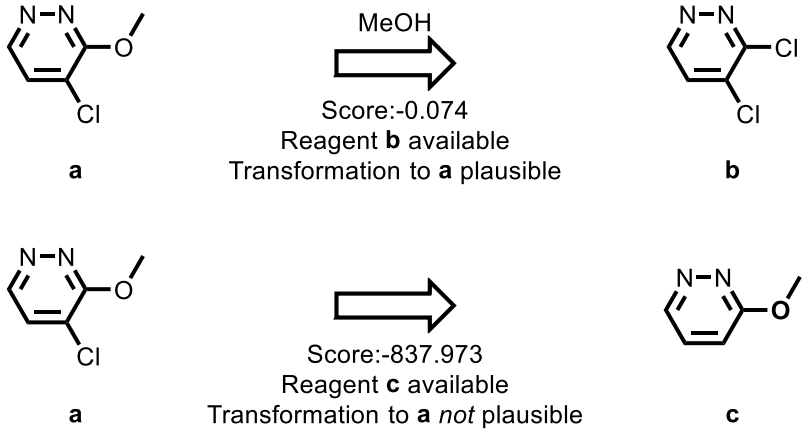
Of the 124,000 novel and synthesizable compounds discussed in 'Future Direction', almost 46,000 are in the Chemspace[a] catalog file[b], which was not a part of our curated list of reagent vendors. Eight of the twelve alcohols in Figure 3 (**16**-**18**, **20**-**22**, **24**-**25**) are also in the Chemspace catalog file, meaning that these eight could have been candidate building blocks for the virtual library even apart from our method. The Chemspace website offers for sale not only those eight compounds but also **19** and **23** (we thank the reviewer who brought this fact to our attention). We attempted to purchase six of our targeted alcohols from Chemspace, including the four that we did not prepare (**17** and **21**-**23**). We also selected two others that we had successfully prepared (**16** and **19**), chosen based on low cost and also to cover two examples in each of the three reaction types. All but one (**23**) of these six reagents were reported to be available within one month. After five weeks none of the reagents were delivered.

(*a*) https://chem-space.com/, accessed Mar 24, 2021.

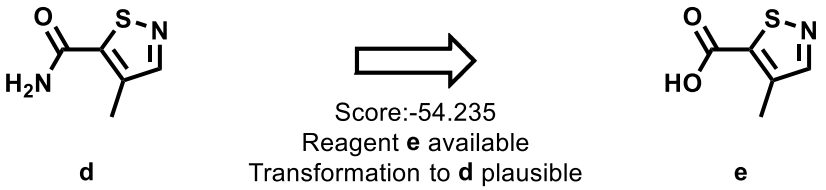
(*b*) The catalog file (provided Mar 26, 2021) has 281,701,403 molecules, 379,059 with 10 or fewer atoms.

Regarding questions of stereochemistry, ASKCOS is not yet able to provide consistently accurate predictions that recognize stereochemical outcomes (and we note that the reaction databases used to train ASKCOS are not rich in stereochemical information). Consider the two diastereomers of compound **8**. For neither of them does the ASKCOS One-Step Retrosynthesis module suggest a reaction that would generate the stereogenic alcohol center (in any of the top five suggestions for either diastereomer). The best scoring suggestion for each is deprotection of the non-orderable *N*-benzyl precursor, with Scores close to -700.

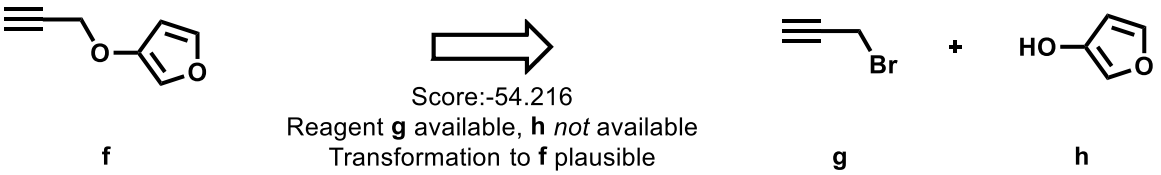
In formulating the rule of thumb that "a Score of -100 or lower indicates an inaccessible compound" we also examined reactions involving aromatic substitution. The two highest scoring routes for compound **a** are shown below. The first starts with orderable reagent **b** and envisions a nucleophilic displacement of the more reactive chloride. The score for the transformation of orderable **b** into **a** is thus high, and indicates that **a** could likely be synthesized in one step from an orderable reagent. Contrast that with the second highest scoring route, the transformation of **c** into **a**. Compound **c** is commercially available, so lack of orderability is not the explanation for the poor score. Instead, ASKCOS recognizes that this transformation is likely to provide other chlorinated products in addition to **a**.



Specific examples supporting this rule of thumb are shown below. Reagents **d** and **f** are both unavailable, with ASKCOS Scores of about -54. This score is below the range we consider "likely synthesizable from available reagents", and above the "unlikely" range. A manual examination of the ASKCOS results suggests that **d** should be considered easily synthesizable, whereas **f** should not. In the case of **d** the corresponding acid **e** is commercially available, but this is not true with **f**. Although the transformation of **h** into **f** via propargyl bromide is plausible given the availability of **h**, the furanol is not actually available from any of the vendors listed in Table S1.



d Score:-54.235 Reagent **e** available Transformation to **d** plausible **e**



f Score:-54.216 Reagent **g** available, **h** *not* available Transformation to **f** plausible **g** **h**

Table S2. Criteria used to remove virtual library compounds

Any compound matching one of these SMILES/SMARTS patterns was excluded:

[N,n]C(=S)[N,n]
COCOC
[C,c]C(=O)O
[C,c]C(=O)[C,c]
N[n,o]
[N;!R]=C
NO
[N;!R]-N
C1CO1
O=[C;H1]
[O;H1]a
C=C[O;H1]
C#C
C=[C;!R]
[A;D2;!R;!$(C#N)][A;D2;!R][A;D2;!R][A;D2;!R;!$(C#N)]
[C;D2;!R;!$(C#N)][C;D2;!R][C;D2;!R;!$(C#N)]
[O;H1]CC#N
C#N.C#N
C=O.C=O
N#CC=O
N#CCN
[C;H3][C;H2][C;H2]
[N,O,n,o]C(=O)[N,O,n,o]

After removing the undesired functionality listed above and browsing the remaining reagents, we further excluded any compound failing to match these PilotScript (Dassault Systèmes BIOVIA, BIOVIA Pipeline Pilot, Release 2020, San Diego: Dassault Systèmes, 2020) expressions:

Num_RotatableBonds < 4
c_count < num_atoms - 1
Num_AromaticRings = 0
Num_H_Acceptors < 4
Num_H_Donors = 1

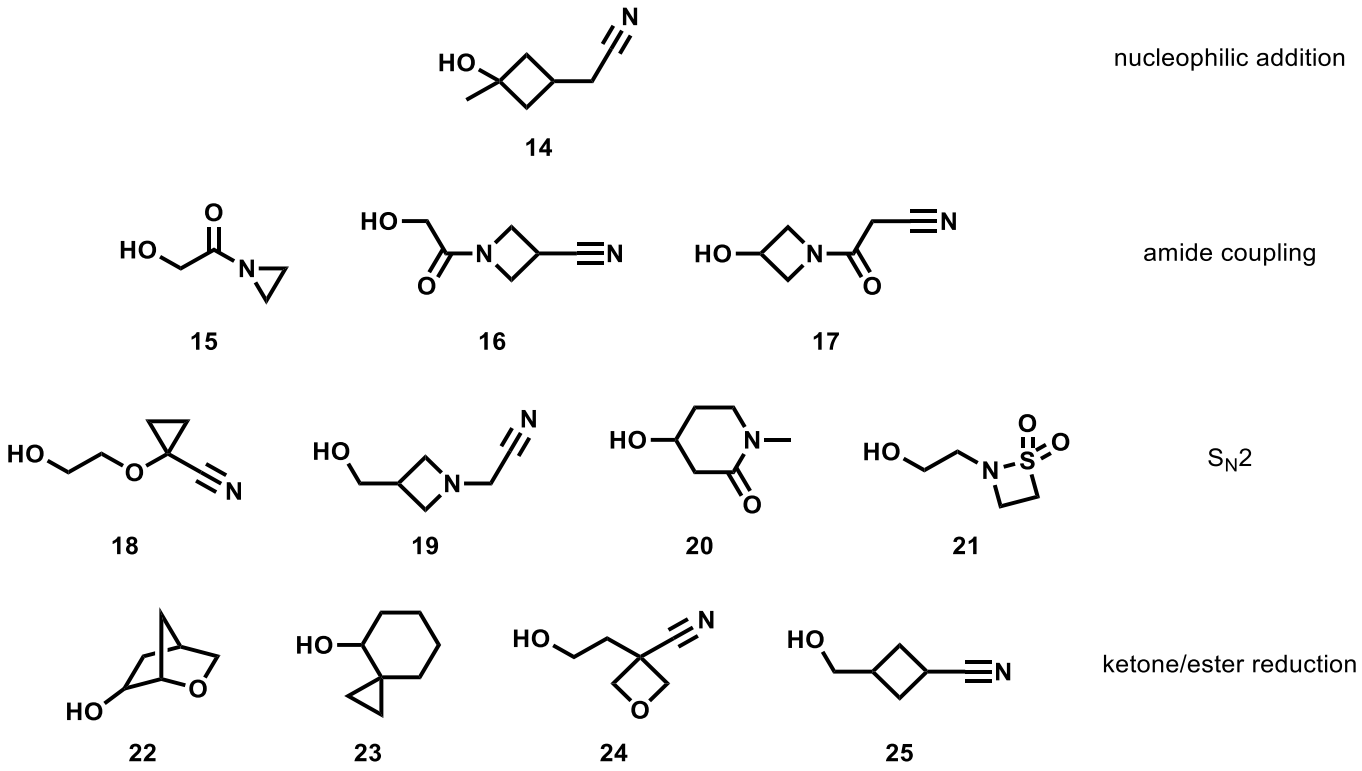
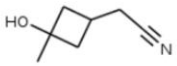


Figure S1. Selected alcohol reagents.

Target molecule: CC1(O)CC(CC#N)C1

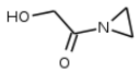


Results generated in 1.731 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	C[Mg]I *($1.0/g)* + N#CCC1CC(=O)C1 *($3.0/g)* → ?	-0.072	331	0.0558	0.974
2	CI *($1.0/g)* + N#CCC1CC(=O)C1 *($3.0/g)* → ?	-0.663	89	0.0060	0.844
3	C[Mg]Br *($1.0/g)* + N#CCC1CC(=O)C1 *($3.0/g)* → ?	-2.498	307	0.0016	0.986
4	N#CCC1CC(=O)C1 *($3.0/g)* + [Li]C *(cannot buy)* → ?	-211.504	528	0.0268	0.962
5	N#CCC1CC(=O)C1 *($3.0/g)* → ?	-487.476	108	0.0021	0.885
6	CC1(O)CC(CBr)C1 *(cannot buy)* + N#C[Na] *($1.0/g)* → ?	-698.709	369	0.0762	0.995

Figure S2. ASKCOS-recommended synthesis of alcohol **14**.

Target molecule: O=C(CO)N1CC1

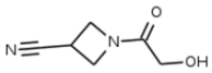


Results generated in 0.563 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	C1CN1 *(cannot buy)* + CC1(C)OCC(=O)O1 *($45.0/g)* → ? 	-17.536	26	0.8915	0.804
2	C1CN1 *(cannot buy)* + COC(=O)CO *($2.0/g)* → ? 	-632.442	637	0.0247	0.993

Figure S3. ASKCOS-recommended synthesis of alcohol **15**.

Target molecule: N#CC1CN(C(=O)CO)C1

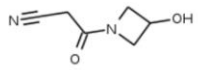


Results generated in 1.089 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	CC(=O)OCC(=O)Cl *($5.0/g)* + N#CC1CNC1 *($13.0/g)* → ?	-0.051	28	0.3516	0.983
2	N#CC1CNC1 *($13.0/g)* + O=C(O)CO *($1.0/g)* → ?	-0.150	21621	0.0936	0.886
3	N#CC1CNC1 *($13.0/g)* + O=CCO *($62.0/g)* → ?	-136.659	15	0.0005	0.927
4	N#CC1CNC1 *($13.0/g)* + O=C(Cl)CO *(cannot buy)* → ?	-531.171	10163	0.0421	0.967

Figure S4. ASKCOS-recommended synthesis of alcohol **16**.

Target molecule: N#CCC(=O)N1CC(O)C1

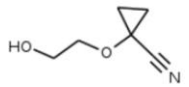


Results generated in 1.593 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	COC(=O)CC#N *($1.0/g)* + OC1CNC1 *($2.0/g)* → ?	-0.010	637	0.3084	0.979
2	CCOC(=O)CC#N *($1.0/g)* + OC1CNC1 *($2.0/g)* → ?	-0.012	728	0.2421	0.919
3	N#CCC(=O)O *($1.0/g)* + OC1CNC1 *($2.0/g)* → ?	-0.027	21307	0.1097	0.994
4	N#CCC(=O)Cl *($3.0/g)* + OC1CNC1 *($2.0/g)* → ?	-0.047	10163	0.1054	0.997
5	N#CCC(=O)[O-] *($1.0/g)* + OC1CNC1 *($2.0/g)* → ?	-19.364	44	0.0002	0.987
6	CCC(=O)OC(=O)CC#N *(cannot buy)* + OC1CNC1 *($2.0/g)* → ?	-8471.465	143	0.0075	0.966

Figure S5. ASKCOS-recommended synthesis of alcohol **17**.

Target molecule: N#CC1(OCCO)CC1

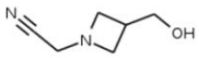


Results generated in 1.801 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	C1CO1 *($1.0/g)* + N#CC1(O)CC1 *($3.0/g)* → ?	-1.575	126	0.0025	0.986
2	N#CC1(Br)CC1 *($3.0/g)* + OCCO *($1.0/g)* → ?	-1.643	19	0.0024	0.942
3	CC(=O)OCCOC1(C#N)CC1 *(cannot buy)* → ?	-1871.900	26861	0.0472	0.862

Figure S6. ASKCOS-recommended synthesis of alcohol **18**.

Target molecule: N#CCN1CC(CO)C1

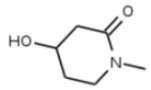


Results generated in 2.076 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	N#CCBr ($3.0/g) + OCC1CNC1 ($68.0/g) → ?	-0.290	16569	0.2445	0.996
2	N#CCCl ($1.0/g) + OCC1CNC1 ($68.0/g) → ?	-1.190	12249	0.0580	0.989
3	N#CCI ($4.0/g) + OCC1CNC1 ($68.0/g) → ?	-2.439	1568	0.0295	0.994
4	Cc1ccc(S(=O)(=O)OCC#N)cc1 ($100.0/g) + OCC1CNC1 ($68.0/g) → ?	-24.737	1612	0.0068	0.942
5	N#CCO ($5.0/g) + OCC1CNC1 ($68.0/g) → ?	-39.984	548	0.0018	0.936
6	N#CC=O (cannot buy) + OCC1CNC1 ($68.0/g) → ?	-2468.016	6695	0.0065	0.989

Figure S7. ASKCOS-recommended synthesis of alcohol **19**.

Target molecule: CN1CCC(O)CC1=O

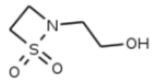


Results generated in 1.820 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	CI *(\$1.0/g)* + O=C1CC(O)CCN1 *(\$3.0/g)* → ?	-0.835	1371	0.0048	0.975
2	CCOC(=O)CC(O)CCNC *(cannot buy)* → ?	-1325.166	98	0.0627	0.944

Figure S8. ASKCOS-recommended synthesis of alcohol **20**.

Target molecule: O=S1(=O)CCN1CCO

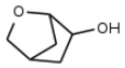


Results generated in 1.181 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	O=S1(=O)CCN1 *($3.0/g)* + OCCCl *($1.0/g)* → ?	-0.069	63	0.0578	0.758
2	O=S1(=O)CCN1 *($3.0/g)* + OCCI *($4.0/g)* → ?	-4.727	132	0.0015	0.835
3	O=S(=O)(CCCl)NCCO *(cannot buy)* → ?	-1093.793	253	0.0578	0.989

Figure S9. ASKCOS-recommended synthesis of alcohol **21**.

Target molecule: OC1CC2COC1C2

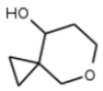


Results generated in 1.716 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	O=C1CC2COC1C2 *($3.0/g)* → ?	-0.033	8992	0.0920	0.931
2	CC(=O)OC1CC2COC1C2 *(cannot buy)* → ?	-6068.853	26861	0.0158	0.963

Figure S10. ASKCOS-recommended synthesis of alcohol **22**.

Target molecule: OC1CCOCC12CC2

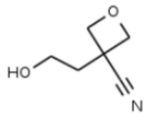


Results generated in 1.610 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	O=C1CCOCC12CC2 *($4.0/g)* → ?	-0.006	8688	0.6322	0.997
2	C1OCC2(CC2)[C@H]2O[C@@H]12 *(cannot buy)* → ?	-1747.810	8	0.0563	0.966

Figure S11. ASKCOS-recommended synthesis of alcohol **23**.

Target molecule: N#CC1(CCO)COC1



Results generated in 1.121 seconds

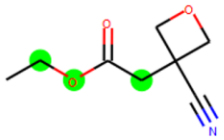
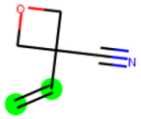
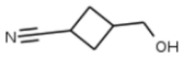
Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	CCOC(=O)CC1(C#N)COC1 *($3.0/g)* → ?	-0.035	4221	0.0848	0.946
					
2	C=CC1(C#N)COC1 *(cannot buy)* → ?	-356.460	3529	0.1522	0.973
					

Figure S12. ASKCOS-recommended synthesis of alcohol **24**.

Target molecule: N#CC1CC(CO)C1



Results generated in 1.934 seconds

Rank	Precursor(s)	Score	# Examples	Max template relevance	Fast filter score
1	COC(=O)C1CC(C#N)C1 *($3.0/g)* → ?	-0.026	13983	0.1139	0.783
2	N#CC1CC(C(=O)O)C1 *($3.0/g)* → ?	-0.059	10329	0.0511	0.854
3	N#CC1CC(C=O)C1 *(cannot buy)* → ?	-429.894	11110	0.1239	0.987

Figure S13. ASKCOS-recommended synthesis of alcohol **25**.

EXPERIMENTAL SECTION

General Experimental. Anhydrous solvents and reagents were used as supplied by the manufacturer. Unless otherwise stated, reaction mixtures were magnetically stirred at room temperature (rt). Where solutions are "dried", they are generally dried over a drying agent such as Na_2SO_4 or $MgSO_4$. Concentrated refers to concentration using a rotary evaporator under reduced pressure. Unless specified otherwise, normal phase flash column chromatography (FCC) was performed on silica gel (SiO_2). The calculated (calcd) mass corresponds to the exact mass. HRMS was obtained on an Agilent Technologies 6230 series TOFMSD, using ESI in positive mode. Nuclear magnetic resonance (NMR) spectra were obtained on Bruker model DRX spectrometers. Chemical shifts (δ) are expressed in parts per million, relative to internal tetramethylsilane; coupling constants (J) are in hertz (Hz). The following abbreviations are used to describe peak patterns when appropriate: s (singlet), d (doublet), t (triplet), q (quartet), qt (quintet), m (multiplet), app (apparent), and br (broad). HPLC−MS chromatograms and spectra were obtained using one of the following methods: (1) Agilent 1200 HPLC and G6100 system on X-Bridge ShieldRP18 (50 × 2.1 mm, 5 μm) and a gradient system of 0.05% NH_4OH in H_2O/CH_3CN, 100:0 to 5:95 over 7.5 min, then 100:0 for 2.5 min at a temperature of 40 °C; (2) Agilent 1200 HPLC and G6100 system on Phenomenex Luna-C18 (50 × 2 mm, 5 μm) and a gradient system of 0.1% TFA in H_2O/0.05% TFA in CH_3CN, 100:0 to 15:85 over 7.5 min, then 100:0 for 2.5 min at a temperature of 50 °C; or (3) Agilent 1100 HPLC and G1367A system on X-Bridge C18 (100 × 3 mm, 3.5 uM) and a gradient system of 20 mM NH_4OH in H_2O/CH_3CN 90:10 over 2 min, then 0:100 for 1 min at a flow rate of 2.4 mL/min at a temperature of 45 °C. All compounds were of a minimum of 95% purity as determined by HPLC.

Chemistry: Preparation of compounds 14-25 and 47.

Compound 14. 2-(3-hydroxy-3-methylcyclobutyl)acetonitrile. To a vial containing a stir bar and anhydrous toluene (10 mL) cooled to -5 °C was added titanium tetrachloride (1.20 mL, 11.0 mmol, 1.20 equiv). After 5 minutes, methyllithium (1.6 M in Et_2O, 6.9 mL, 11.0 mmol, 1.20 equiv) was added dropwise and the resulting mixture stirred at -5 °C for 30 minutes at which point 2-(3-oxocyclobutyl)acetonitrile (1.00 g, 9.16 mmol, 1.00 equiv) was added dropwise as a solution in toluene (5 mL). The resulting mixture was stirred for 2 h at -5 °C and then quenched with the dropwise addition of MeOH (1.5 mL) followed by 2 N aq. HCl (9 mL). The layers were separated and the aqueous extracted with EtOAc (3 x 6 mL). The combined organics were dried with $MgSO_4$, filtered and concentrated in vacuo. The crude oil was purified by silica gel chromatography (0-75% EtOAc in hexanes) to give the title compound (clear colorless oil) as an inseparable mixture of diastereomers (2.5:1) (303 mg, 2.40 mmol, 26% yield). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 2.85 – 2.74 (m, 0.4H), 2.50 (d, *J* = 6.5 Hz, 2H), 2.46 (d, *J* = 6.7 Hz, 0.7H), 2.42 – 2.34 (m, 1H), 2.34 – 2.27 (m, 2.4H), 2.24 – 2.13 (m, 1H), 1.96 – 1.88 (m, 3H), 1.38 (s, 4.4H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 118.7, 71.6, 68.9, 43.0, 41.3, 30.0, 27.3, 23.9, 23.8, 23.5, 22.1.

Compound 15. 1-(aziridin-1-yl)-2-hydroxyethan-1-one.

Step A: 2-((2-chloroethyl)amino)-2-oxoethyl acetate. To a round bottom flask was added 2-chloroethylamine hydrochloride (2.07 g, 17.8 mmol, 1 equiv), dichloromethane (71.4 mL, 0.25 M), and DIPEA (9.23 mL, 53.5 mmol, 3 equiv). Placed in an ice bath, then slowly added acetoxyacetyl chloride (1.92 mL, 17.8 mmol, 1 equiv) while stirring. Removed from ice bath and stirred at room temperature for 16 hours. The mixture was diluted with 100 mL water and extracted with dichloromethane (2 x 200 mL). The combined organics were dried over sodium sulfate, filtered and concentrated in vacuo. The crude oil was purified by silica gel chromatography (0-100 % EtOAc in hexanes) to give the title compound as a

yellow solid (1.42 g, 7.9 mmol, 44% yield). MS (ESI): mass calcd. for $C_6H_{10}ClNO_3$, 179.0; m/z found, 180.1 $[M+H]^+$. ^{1}H NMR (500 MHz, DMSO-*d*6) δ 8.30 – 8.21 (m, 1H), 4.46 (s, 2H), 3.61 (t, *J* = 6.2 Hz, 2H), 3.41 (q, *J* = 6.1 Hz, 2H), 2.09 (s, 3H).

Step B: *N*-(2-chloroethyl)-2-hydroxyacetamide. To a vial was added 2-((2-chloroethyl)amino)-2-oxoethyl acetate (370 ng, 2.06 mmol, 1 equiv) and methanol (2.06 mL, 1 M). While stirring, added potassium carbonate (854 mg, 6.18 mmol, 3 equiv) and stirred at room temperature for 16 hr. The mixture was filtered through Celite, concentrated in vacuo, and purified by silica gel chromatography (0-100 % EtOAc in hexanes) to give the title compound as a clear oil (137 mg, 1 mmol, 48% yield). MS (ESI): mass calcd. for $C_4H_8ClNO_2$, 137.0; m/z found, 138.1 $[M+H]^+$. ^{1}H NMR (500 MHz, DMSO-*d*6) δ 8.02 – 7.87 (m, 1H), 5.54 (s, 1H), 3.82 (s, 2H), 3.63 (t, *J* = 6.4 Hz, 2H), 3.43 (q, *J* = 6.3 Hz, 2H).

Step C: 1-(aziridin-1-yl)-2-hydroxyethan-1-one. To a vial was added N-(2-chloroethyl)-2-hydroxyacetamide (135 mg, 0.98 mmol, 1 equiv) and THF (3.93 mL, 0.25 M). Added sodium hydride (60% dispersion in mineral oil, 118 mg, 2.94 mmol, 3 equiv) and stirred at room temperature for 16 hr. The mixture was diluted with 25 mL water and extracted with ethyl acetate (2 x 50 mL). The combined organics were dried over sodium sulfate, filtered and concentrated in vacuo. The crude oil was purified by silica gel chromatography (0-100 % EtOAc in hexanes) to give the title compound as a white solid (13.8 mg, 14% yield). MS (ESI): mass calcd. for $C_4H_7NO_2$, 101.0; m/z found, 102.2 $[M+H]^+$. ^{1}H NMR (500 MHz, DMSO-*d*6) δ 5.25 (t, *J* = 6.3 Hz, 1H), 4.20 (t, *J* = 9.5 Hz, 2H), 4.02 (dt, *J* = 6.3, 1.3 Hz, 2H), 3.72 (tt, *J* = 9.5, 1.4 Hz, 2H). ^{13}C NMR (126 MHz, DMSO-*d*6) δ 166.98, 67.40, 56.70, 54.15.

Compound 16. 1-(2-hydroxyacetyl)azetidine-3-carbonitrile.

Step A: 2-(3-cyanoazetidin-1-yl)-2-oxoethyl acetate. To a solution of azetidine-3-carbonitrile hydrochloride (500 mg, 4.22 mmol , 1.00 equiv) in THF (10.2 mL) was added triethylamine (1.30 mL,

9.28 mmol, 2.2 equiv) at 23 °C. After 5 minutes, acetoxyacetyl chloride (633 mg, 4.64 mmol, 1.1 equiv) was added dropwise and the resulting mixture stirred at 23 °C for 24 h. The mixture was then diluted with water (8 mL) and EtOAc (8 mL). The layers were separated and the aqueous extracted with EtOAc (2 x 8 mL). The combined organics were dried with $MgSO_4$, filtered and concentrated in vacuo. The crude oil was purified by silica gel chromatography (0-100 % EtOAc in hexanes) to give the title compound as a yellow oil (320 mg, 1.76 mmol, 42% yield).

Step B: 1-(2-hydroxyacetyl)azetidine-3-carbonitrile. To a solution of 2-(3-cyanoazetidin-1-yl)-2-oxoethyl acetate (320 mg, 1.76 mmol, 1.00 equiv) in MeOH (1.4 mL) was added K_2CO_3 (606 mg, 4.39 mmol, 2.50 equiv) and the resulting mixture stirred at 23 °C. After 20 minutes, the mixture was filtered through celite and concentrated in vacuo. The crude oil was purified by silica gel chromatography (0-100% EtOAc in hexanes) to give the title compound as a white solid (90 mg, 0.64 mmol, 37% yield). ^{1}H NMR (500 MHz, DMSO-d_6) δ 5.07 (t, J = 6.1 Hz, 1H), 4.47 (t, J = 9.0 Hz, 1H), 4.37 (dd, J = 9.0, 6.0 Hz, 1H), 4.16 (t, J = 9.3 Hz, 1H), 4.01 (dd, J = 9.7, 6.2 Hz, 1H), 3.91 (d, J = 5.9 Hz, 2H), 3.78 (tt, J = 9.1, 6.0 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 171.9, 120.7, 60.4, 53.5, 51.2, 17.7.

Compound 19. 2-(3-(hydroxymethyl)azetidin-1-yl)acetonitrile. To a vial containing a stir bar, azetidin-3-ylmethanol hydrochloride (170 mg, 1.38 mmol, 1.0 equiv), K_2CO_3 (437 mg, 3.16 mmol, 2.3 equiv), and THF (7.70 mL) at 23 °C was added 2-chloroacetonitrile (87.0 µL, 1.38 mmol, 1.0 equiv) dropwise. The reaction mixture was stirred at 23 °C for 16 h, then filtered and evaporated to dryness. The residue was purified by flash column chromatography on silica gel (0 to 10% MeOH in DCM) to give the title compound as a clear oil (53.6 mg, 0.425 mmol, 31% yield); ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.79 (dd, J = 6.1, 1.2 Hz, 2H), 3.51 (td, J = 7.8, 0.8 Hz, 2H), 3.43 (s, 2H), 3.23 (dd, J = 7.3, 5.4 Hz, 2H), 2.72 – 2.60 (m, 1H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 115.15, 64.11, 64.04, 55.27, 44.26, 32.25.

Compound 20. (*R*)-4-hydroxy-1-methylpiperidin-2-one.

Step A: (*R*)-4-((tert-butyldiphenylsilyl)oxy)piperidin-2-one. To a solution of (*R*)-4-hydroxypiperidin-2-one (500 mg, 4.343 mmol), 1H-imidazole (886.958 mg, 13.029 mmol), DMAP (53.056 mg, 0.434 mmol) and DMF (16.813 mL, 0.944 g/mL, 217.142 mmol) was added TBDPSCl (2.225 mL, 1.073 g/mL, 8.686 mmol). The reaction was stirred for 2 h at 45 °C. The crude was diluted with 50 mL of EtOAc and washed with 3 x 50 mL H_2O and 1 x 50 mL of 0.1 M HCl. The crude was purified via column chromatography using 20-25% EtOAc:hexanes to afford the title compound as a clear oil (1.12 g, 72.9%). MS (ESI): mass calcd. for $C_{21}H_{27}NO_2Si$, 353.2; m/z found, 354.9 $[M+H]^+$. ^1H NMR (400 MHz, Chloroform-*d*) δ 7.71 – 7.57 (m, 4H), 7.50 – 7.32 (m, 6H), 5.65 (s, 1H), 4.20 – 4.14 (m, 1H), 3.56 (dtd, *J* = 11.7, 6.7, 1.9 Hz, 1H), 3.15 (dtd, *J* = 11.5, 5.5, 2.6 Hz, 1H), 2.47 – 2.38 (m, 2H), 1.76 (q, *J* = 5.6 Hz, 2H), 1.07 (s, 9H).

Step B: (*R*)-4-((tert-butyldiphenylsilyl)oxy)piperidin-2-one. To a solution of (*R*)-4-((tert-butyldiphenylsilyl)oxy)piperidin-2-one (1.3 g, 3.677 mmol) and DMF (14.236 mL, 0.944 g/mL, 183.857 mmol) was added NaH (60% dispersion in mineral oil) (147.071 mg, 3.677 mmol). The reaction was stirred for 5 min. Then, MeI (0.275 mL, 2.28 g/mL, 4.413 mmol) was added. The reaction was heated up to 80 °C and stirred for 30 min. To the reaction mixture was added 50 mL EtOAc, washed with 3 x 50 mL H_2O. The organic layer was collected and concentrated down under vacuo. The crude oil was flashed with 30-100% EtOAc:hexanes over 25 min to afford the title compound as a clear oil (0.79g, 58.4%). MS (ESI): mass calcd. for $C_{22}H_{29}NO_2Si$, 367.2; m/z found, 368.2 $[M+H]^+$. ^1H NMR (400 MHz, Chloroform-*d*) δ 7.68 – 7.59 (m, 4H), 7.47 – 7.34 (m, 6H), 4.17 – 4.05 (m, 1H), 3.50 – 3.41 (m, 1H), 3.08 (dt, *J* = 11.9, 5.9 Hz, 1H), 2.93 (s, 3H), 2.51 – 2.38 (m, 2H), 1.86 – 1.69 (m, 2H), 1.05 (s, 9H).

Step C: (*R*)-4-hydroxy-1-methylpiperidin-2-one. To a solution of (*R*)-4-((tert-butyldiphenylsilyl)oxy)piperidin-2-one (0.79 g, 2.235 mmol) and THF (4.546 mL, 0.886 g/mL, 55.864 mmol) was added TBAF 1M in THF (6.704 mL, 6.704 mmol). The solution was stirred for 2 h. The crude was concentrated down under vacuo. The crude was purified via FCC using 0-20% 2M NH_3/MeOH-DCM.

The desired spot was co-eluted with bis-phenyl-silyl by-product. The material was purified again via FCC using 0-20% 2M NH_3/MeOH-DCM to afford the title compound as a clear oil (150 mg, 52.0%). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 4.19 (ddq, *J* = 11.0, 7.7, 3.3 Hz, 1H), 3.49 (ddd, *J* = 12.6, 7.6, 5.2 Hz, 1H), 3.23 (ddd, *J* = 12.2, 6.7, 5.5 Hz, 1H), 2.94 (s, 3H), 2.70 (d, *J* = 3.7 Hz, 1H), 2.69 – 2.61 (m, 1H), 2.40 (ddd, *J* = 17.4, 6.2, 1.2 Hz, 1H), 2.02 (ddddd, *J* = 13.2, 7.1, 5.4, 2.9, 1.4 Hz, 1H), 1.95 – 1.86 (m, 1H).

Compound 24. 3-(2-hydroxyethyl)oxetane-3-carbonitrile. A solution of ethyl 2-(3-cyanooxetan-3-yl)acetate (750 mg, 4.43 mmol) and EtOH (15 mL) was cooled using an ice/water bath; $NaBH_4$ (335 mg, 8.87 mmol) was slowly added. The ice/water bath was removed, and the reaction mixture was stirred overnight. TLC (PE:EA = 1:1, Rf = 0.4) showed a new spot formed. To the reaction mixture was added aqueous NH_4Cl (60 mL) and the aqueous phase was extracted with DCM (2 x 30 mL), combined extracts were dried over anhydrous Na_2SO_4, filtered, and concentrated under vacuum to give the title compound as a colorless oil, which was used directly in the next step. (230 mg, 1.8 mmol, 41%). ^{1}H NMR (400 MHz, $CDCl_3$): 4.97 (d, *J* = 6.4 Hz, 2H), 4.67 (d, *J* = 6.4 Hz, 2H), 3.96 - 3.90 (m, 2H), 2.27 (t, *J* = 6.0 Hz, 2H).

Compound 25. 3-(hydroxymethyl)cyclobutane-1-carbonitrile. A solution of methyl 3-cyanocyclobutane-1-carboxylate (50 mg, 0.36 mmol) and MeOH (2 mL) was cooled using an ice/water bath; $NaBH_4$ (20.4 mg, 0.54 mmol) was slowly added. The ice/water bath was removed, and the reaction mixture was stirred for 2 h at r.t. TLC (PE: EA =1:1) showed a new spot formed. The reaction mixture was directly concentrated under reduced pressure. EtOAc (10 mL) and water (5 mL) were added and the layers were separated. The aqueous phase was extracted with EtOAc (10 mL x 2). The combined extracts were washed with brine (5 mL), dried over anhydrous Na_2SO_4, filtered, and concentrated under vacuum to give the title compound as a colorless oil (35 mg, 0.32 mmol, 88%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.68 - 3.62 (m, 2H), 3.17 - 3.08 (m, 1H), 2.69 - 2.64 (m, 1H), 2.52 - 2.44 (m, 2H), 2.33 - 2.23 (m, 2H).

Compound 47. Compound **18** is 1-(2-hydroxyethoxy)cyclopropane-1-carbonitrile and was not synthesized directly. Rather, the moiety was attached to the LO molecule as follows.

Step A: Compound **46**. A mixture consisting of compound **45** (1.28 mmol), Cs_2CO_3 (832 mg, 2.55 mmol), oxirane (1.13 g, 25.5 mmol), and DMF (5 mL) was heated by microwave for 1 h at 60 °C. EtOAc (100 mL) and water (50 mL) were added and layers were separated. The aqueous phase was extracted with EtOAc (100 mL x 2). Combined extracts were washed with brine (20 mL), dried over Na_2SO_4, filtered, and concentrated under vacuum to give a brown solid (0.67 mmol), which was then dissolved in dichloromethane (5 mL). To this was added DMAP (123 mg, 1.01 mmol) and TsCl (141 mg, 0.737 mmol) at room-temperature. The resultant mixture was stirred overnight at room-temperature. The mixture was concentrated to dryness under reduced pressure to afford the crude product which was purified by FCC (EtOAc:petroleum ether = from 0% to 36%) to afford **46** (36% yield) as a colorless oil.

Step B: Compound **47**. NaH (55 mg, 60% in mineral oil, 1.4 mmol) was added to a stirred mixture consisting of 1-hydroxycyclopropanecarbonitrile (38 mg, 0.46 mmol) and DMA (2 mL) at 0 °C and stirred for 5 mins at 0 °C. Then **46** (0.23 mmol) in DMA (2 mL) was added to the mixture at 0 °C and stirred overnight at room-temperature. The reaction mixture was added to aq. NH_4Cl (30 mL), and the aqueous phase was extracted with EtOAc (2 x 30 mL), combined extracts and concentrated to dryness under reduced pressure to give **47** (57% yield) as a yellow oil.

ABBREVIATIONS

ADME, absorption, distribution, metabolism and excretion; Boc, *tert*-butoxycarbonyl; cLogP, calculated logarithm of partition coefficient; DCM, dichloromethane; DIPEA, *N*,*N*-diisopropylethylamine; DMA, dimethylacetamide; DMAP, 4-dimethylaminopyridine; DMF, dimethylformamide; DMSO, dimethyl sulfoxide; DNA, deoxyribonucleic acid ; EtOAc, ethyl acetate; EtOH, ethanol; GDB, generated data base;

LAH, lithium aluminum hydride; LHMDS, lithium hexamethyldisilazide; LO, lead optimization; MeI, methyl iodide; MeLi, methyllithium; MeMgBr, methylmagnesium bromide; MeOH, methanol; MW, molecular weight; SAR, structure-activity relationship; TBAF, tetra-*n*-butylammonium fluoride; TBDPSCl, *tert*-butyldiphenylsilyl chloride; THF, tetrahydrofuran; tPSA, topological polar surface area; Ts, 4-toluenesulfonyl (tosyl); TsCl, 4-toluenesulfonyl (tosyl) chloride