6th Advanced In silico Drug Design workshop/challenge

30 January - 3 February 2023
Olomouc, Czech Republic

Univerzita Palackého
v Olomouci

De novo drug design

Pavel Polishchuk

Institute of Molecular and Translational Medicine
Faculty of Medicine and Dentistry
Palacky University

pavlo.polishchuk@upol.cz
qsar4u.com

Size of explored and enumerated chemical space

real datasets

SCIFINDER® A CAS SOLUTION ~ 160 M compounds

REAXYS® ~ 105 M compounds

Commercial

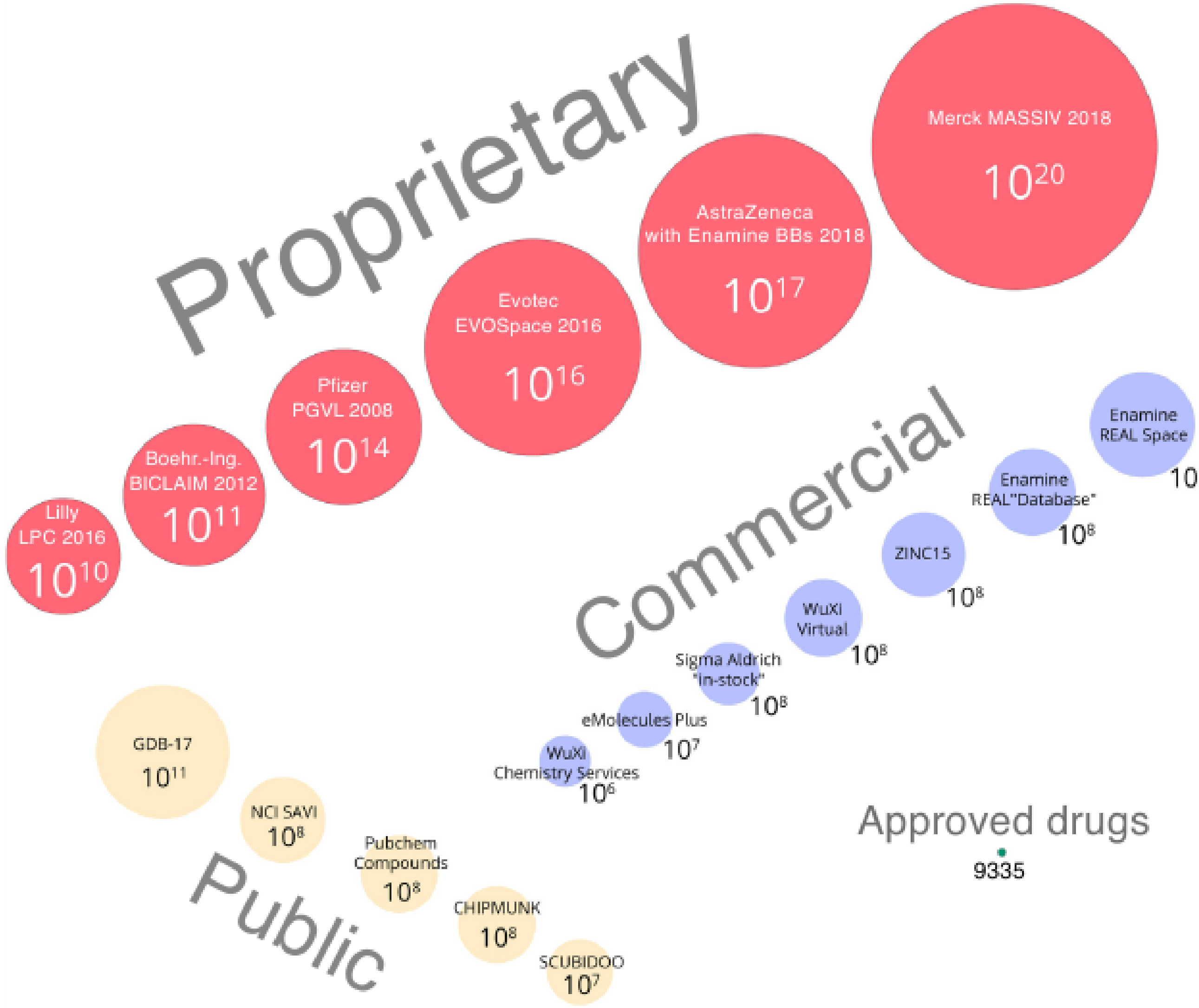
PubChem ~ 102 M compounds Free

ZINC up to 1 B commercially available compounds

virtually enumerated dataset

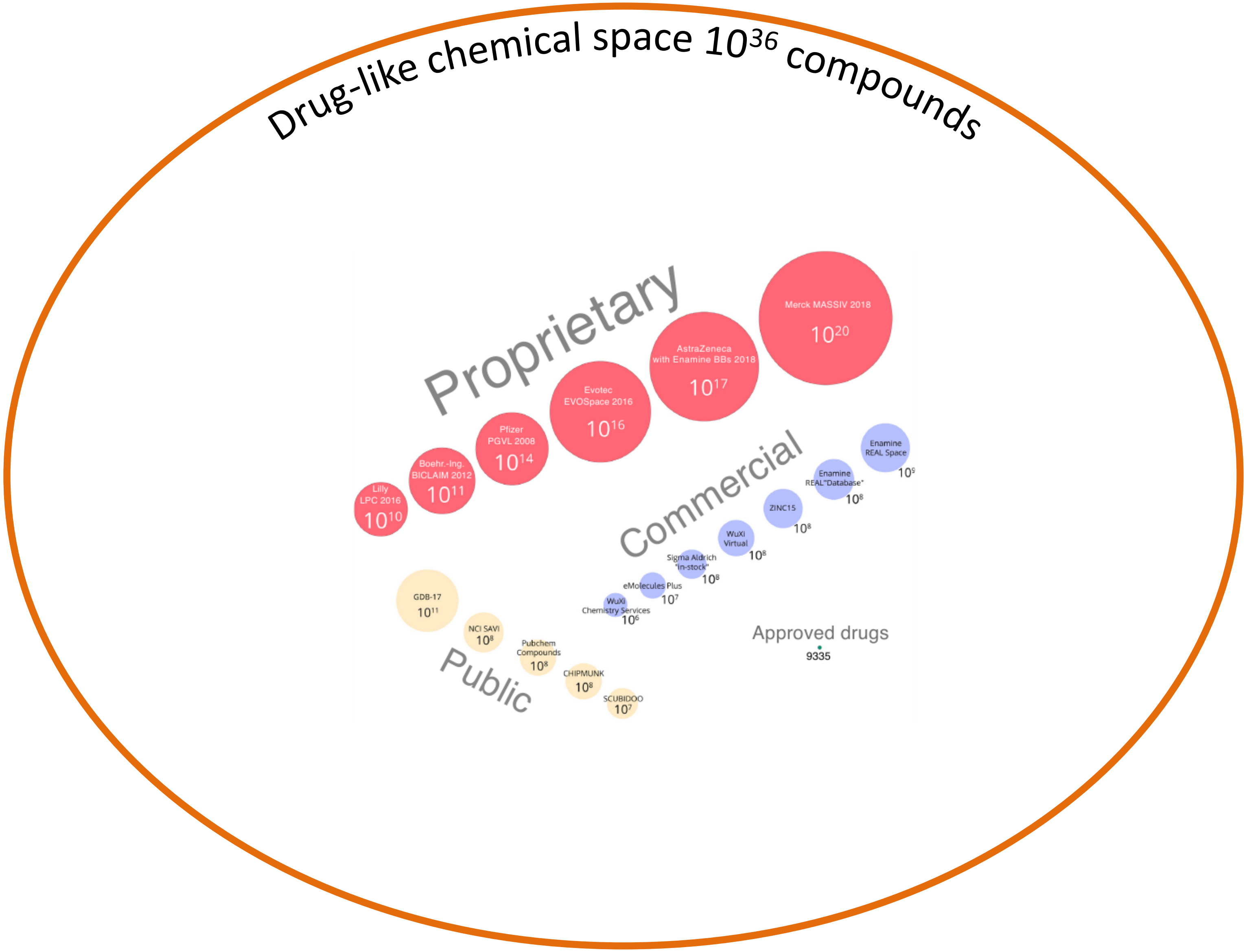
GDB-17 166 B compounds = $1.66 \text{x} 10^{11}$

Size of explored and enumerated chemical space



Hoffmann, T.; Gastreich, M., The next level in chemical space navigation: going far beyond enumerable compound libraries. *Drug Discovery Today* **2019**, 24, 1148-1156. (https://doi.org/10.1016/j.drudis.2019.02.013)

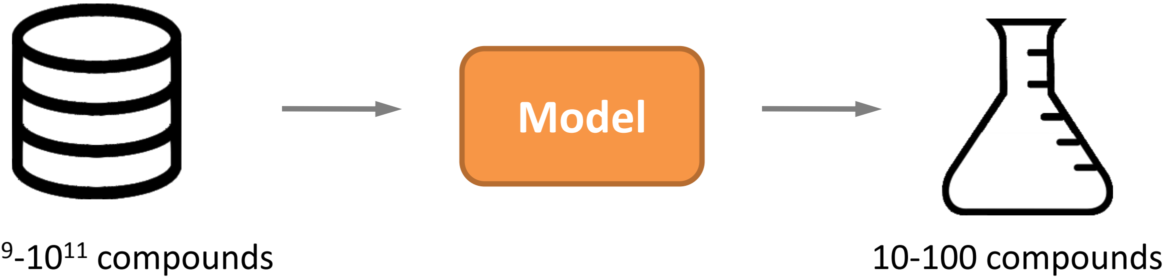
Size of explored and enumerated chemical space



Polishchuk, P. G.; Madzhidov, T. I.; Varnek, A., Estimation of the size of drug-like chemical space based on GDB-17 data. *Journal of Computer-Aided Molecular Design* **2013**, 27, 675-679. (http://dx.doi.org/10.1007/s10822-013-9672-4)

Virtual screening *vs.* de novo design

Virtual screening

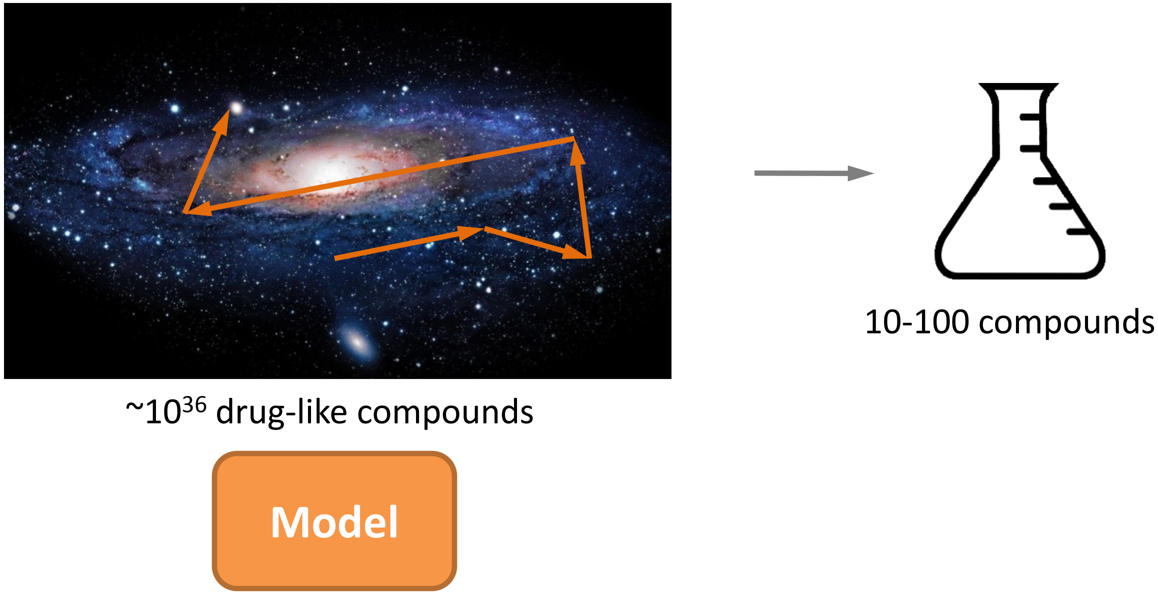


10^9-10^{11} compounds

Model

10-100 compounds

De novo design



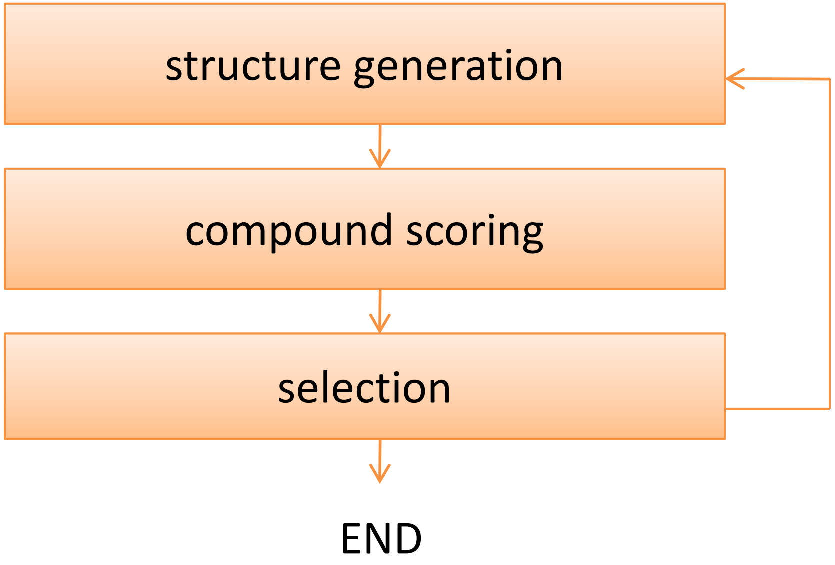
~10^{36} drug-like compounds

Model

10-100 compounds

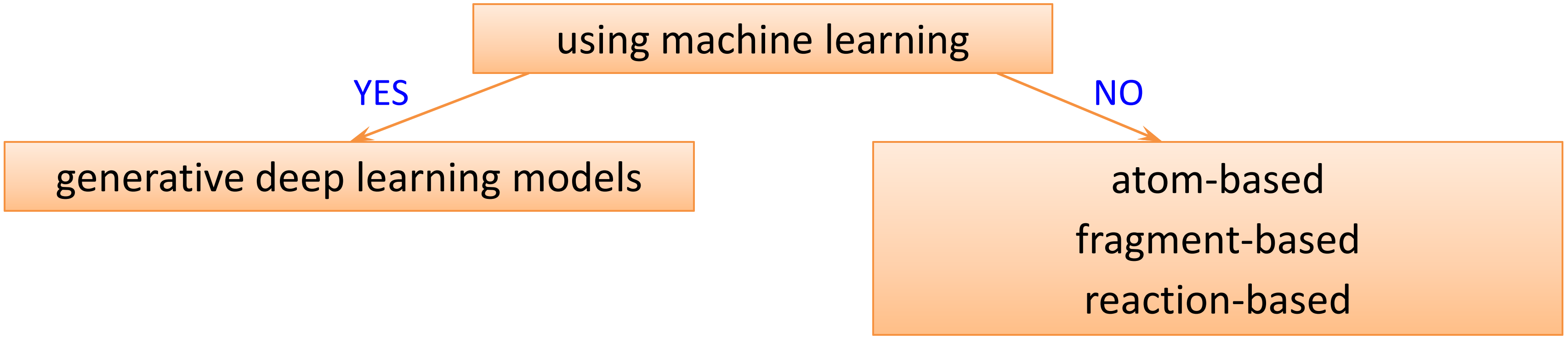
Issues of de novo design

1. **Structure generation** - how to create/assembly new structures
2. **Compound scoring** - how to estimate/predict a property of a compound
3. **Search strategy** - how to find compounds with optimal properties



iterative strategy

De novo structure generation


using machine learning

YES

NO

generative deep learning models

atom-based

fragment-based

reaction-based

- **atom-based** - uses simple rules like add/change/remove atom/bond to perturb structures
- **fragment-based** - uses fragment library to create structures
- **reaction-based** - uses a set of reaction rules and a library of reactants

Atom-based structure generation

Molpher

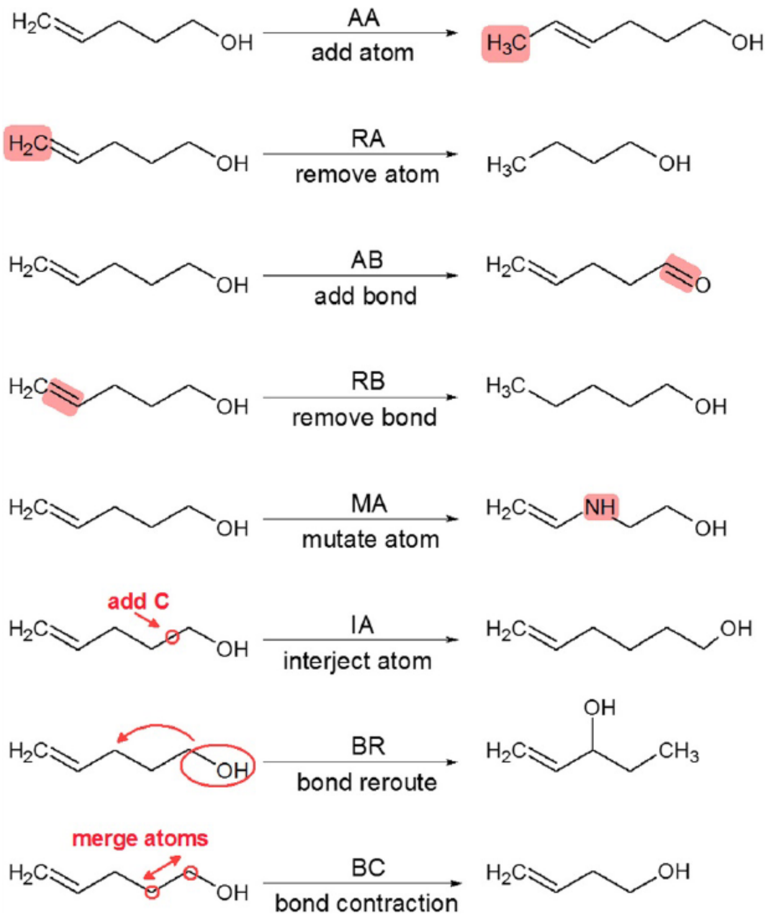


Figure 2 Morphing operators. Morphing operators represent simple structural changes (mutations) that lead to the transformation of the 'reactant' molecule (left) to the 'product' molecule (right).

Hoksza, D.; Škoda, P.; Voršilák, M.; Svozil, D., Molpher: a software framework for systematic chemical space exploration. *Journal of Cheminformatics* **2014**, 6, 7, 10.1186/1758-2946-6-7

Atom-based structure generation

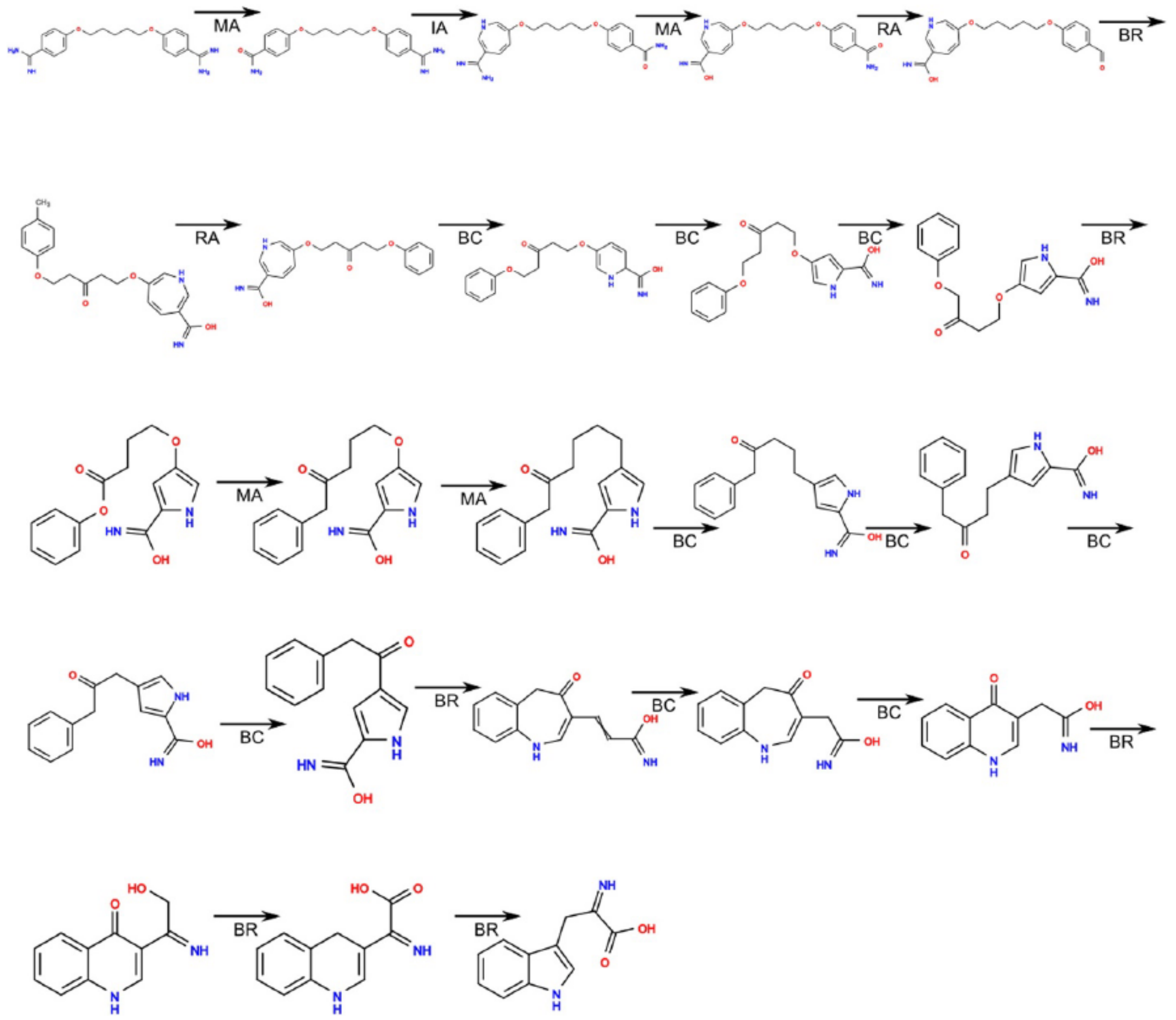
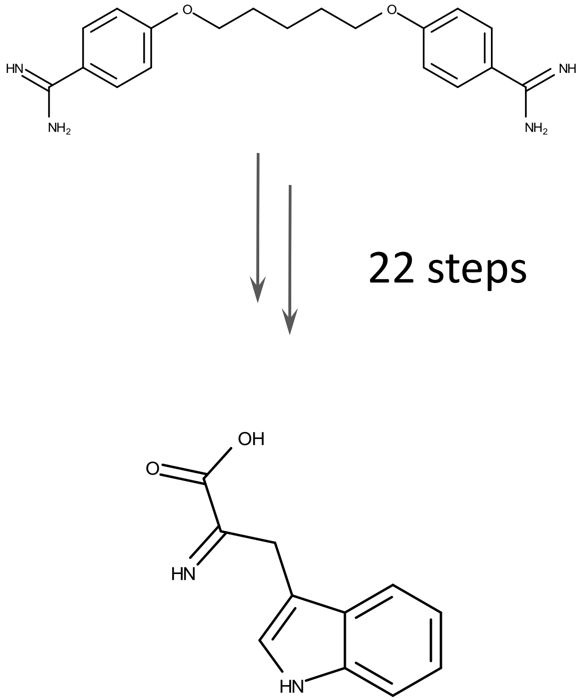


Figure 9 An example of the path between pentamidine (CID 4735) and 2-imino-3-(1H-indol-3-yl)propanoic acid (CID 5599) from dataset *D3*. The path was generated using Morgan fingerprint, Tanimoto coefficient, and synthetic accessibility filter turned on. The arrows' labels show the used morphing operators (see Figure 2). The depiction of the path was done by OpenBabel [74].



22 steps

Hoksza, D.; Škoda, P.; Voršilák, M.; Svozil, D., Molpher: a software framework for systematic chemical space exploration. *Journal of Cheminformatics* **2014**, 6, 7, 10.1186/1758-2946-6-7

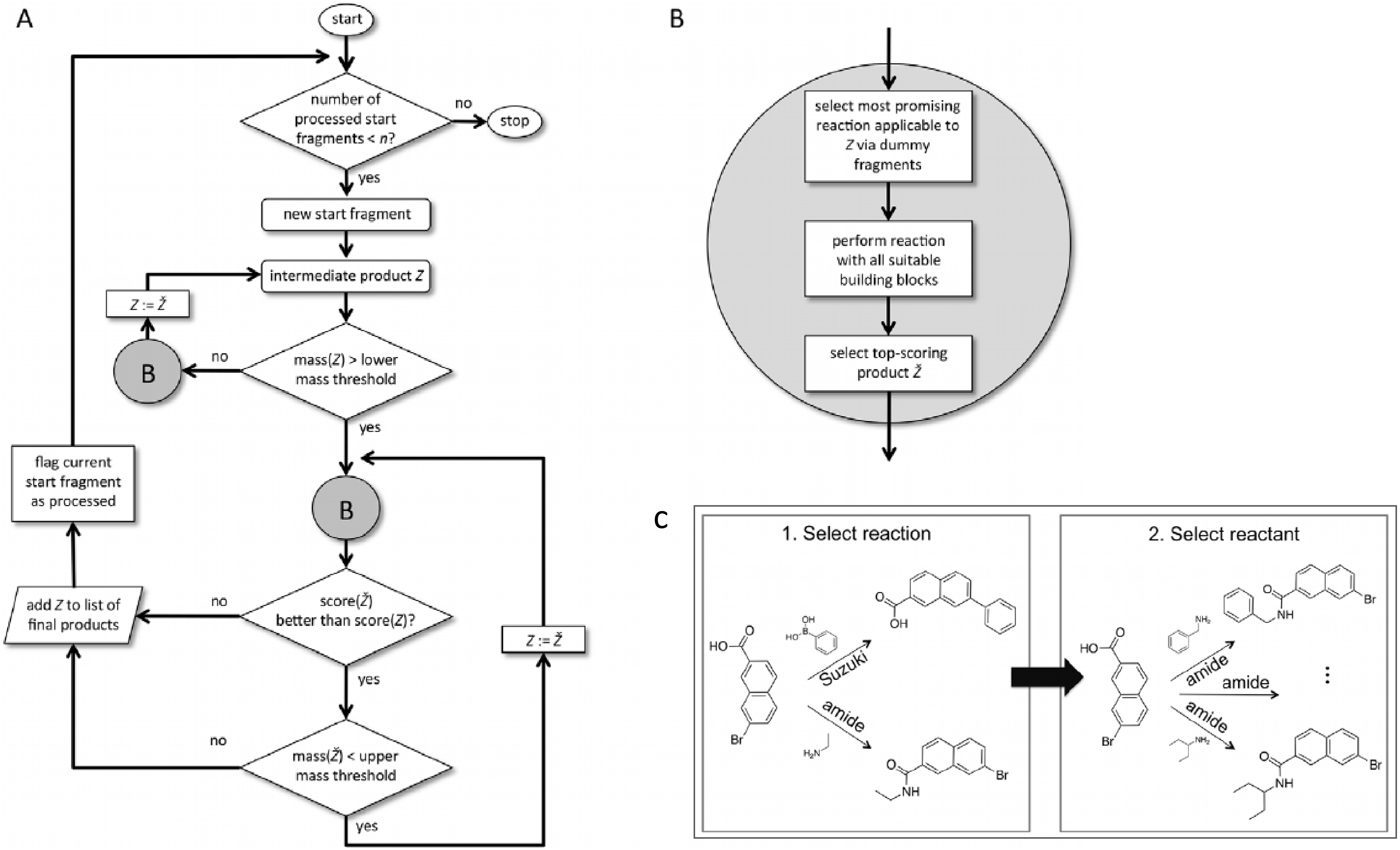
Atom-based structure generation

parameters	atom-based
exhaustiveness of chemical space search	++++ very small steps; more suitable for systematic exploration of local chemical space
structure novelty	+++*
structure diversity	+++*
chemically valid structures	-
synthetically feasible	---
combinatorial explosion / time consuming	---

atom-based ≈ *ab initio*

Reaction-based structure generation

DOGS

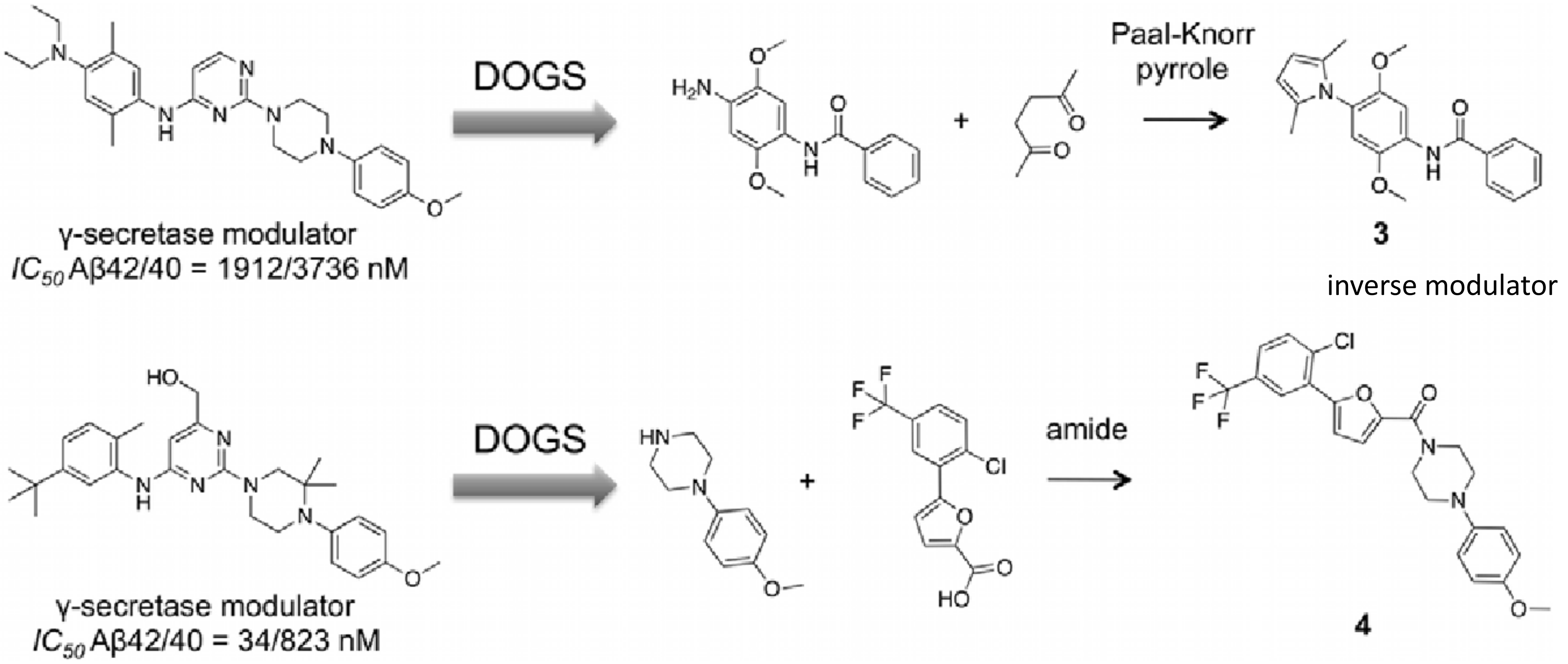

A
start
number of processed start fragments < n?
no
stop
yes
new start fragment
intermediate product Z
Z := Ž
B
no
mass(Z) > lower mass threshold
yes
flag current start fragment as processed
B
add Z to list of final products
no
score(Ž) better than score(Z)?
yes
Z := Ž
no
mass(Ž) < upper mass threshold
yes
B
select most promising reaction applicable to Z via dummy fragments
perform reaction with all suitable building blocks
select top-scoring product Ž
C
1. Select reaction
HO
O
Br
OH
HO B
Suzuki
amide
H₂N
O
OH
O
NH
Br
2. Select reactant
HO
O
Br
NH₂
amide
amide
amide
NH₂
O
NH
Br
O
NH
Br

Hartenfeller, M.; Zettl, H.; Walter, M.; Rupp, M.; Reisen, F.; Proschak, E.; Weggen, S.; Stark, H.; Schneider, G., DOGS: Reaction-Driven de novo Design of Bioactive Compounds. *PLOS Computational Biology* **2012**, 8, e1002380.

Reaction-based structure generation

DOGS

γ-secretase modulators



γ-secretase modulator
IC_{50} Aβ42/40 = 1912/3736 nM

3

inverse modulator

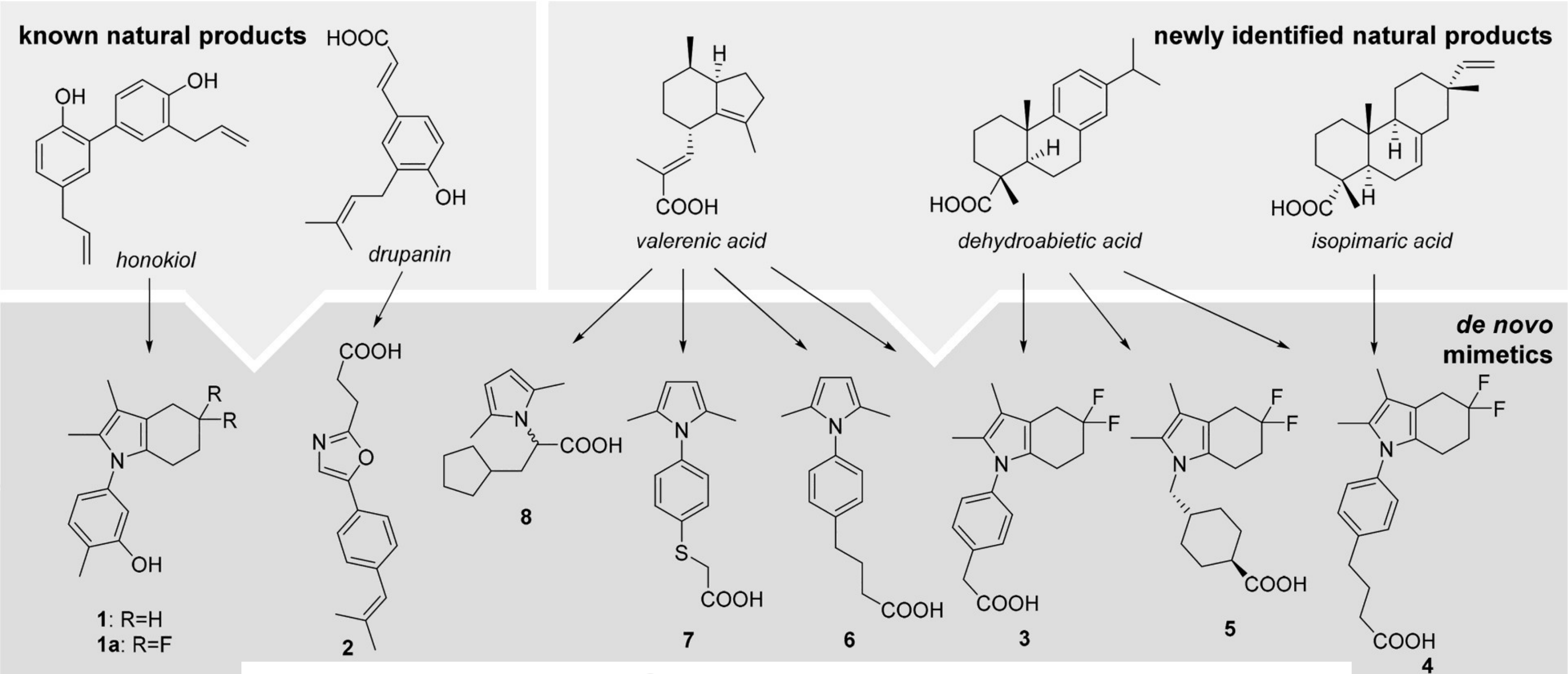
γ-secretase modulator
IC_{50} Aβ42/40 = 34/823 nM

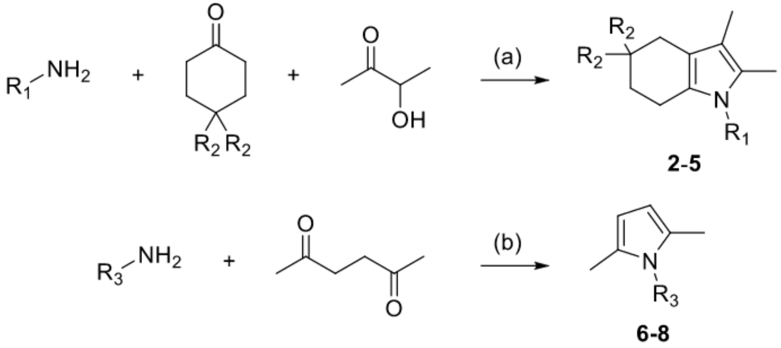
4

Hartenfeller, M.; Zettl, H.; Walter, M.; Rupp, M.; Reisen, F.; Proschak, E.; Weggen, S.; Stark, H.; Schneider, G., DOGS: Reaction-Driven de novo Design of Bioactive Compounds. *PLOS Computational Biology* **2012**, 8, e1002380.

Reaction-based structure generation

Retinoid X Receptor(RXR) Modulators





isopimaric

dehydroabi

valerenic ac

sclareol

conocarpar

Supporting figure 5: Synthesis of de novo mimetics **1a** and **3-8**. Reagents and conditions: (a) EtOH, HOAc, *μ*w, 100°C, 3-6 h, 43-78%; (b) montmorillonite K10, *μ*w, 90°C, 30 min, 41-85%.

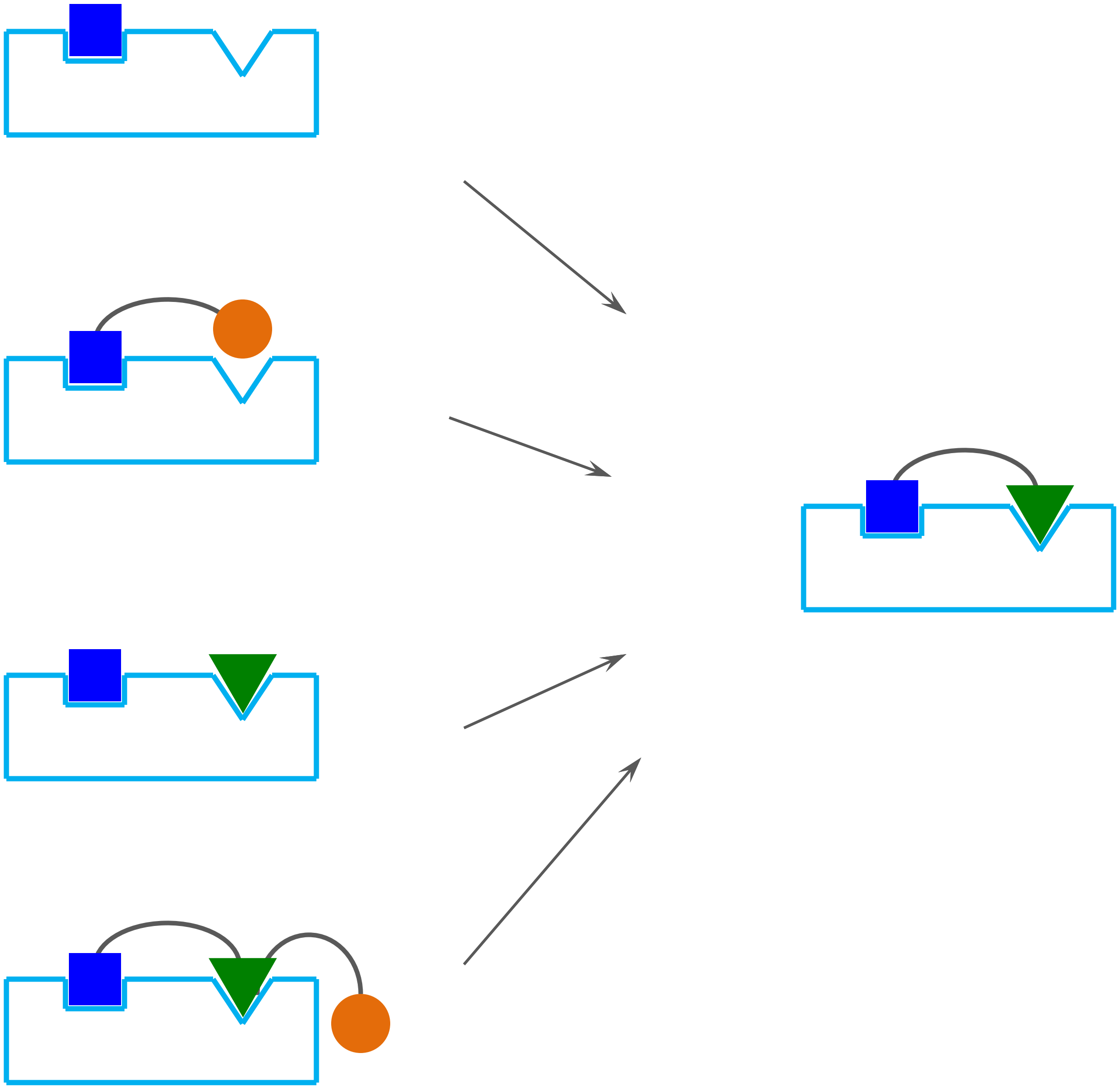
Merk D., et al. *J. Med. Chem.*, **2018**, 61 (12), pp 5442–5447

Reaction-based structure generation

	reaction-based
exhaustiveness of chemical space search	+ depends on reactant library and reaction rules; only grow molecules
structure novelty	++
structure diversity	++
chemically valid structures	+++
synthetically feasible	+++
combinatorial explosion / time consuming	+++

reaction-based ≈ empirical

Fragment-based structure generation



GROW

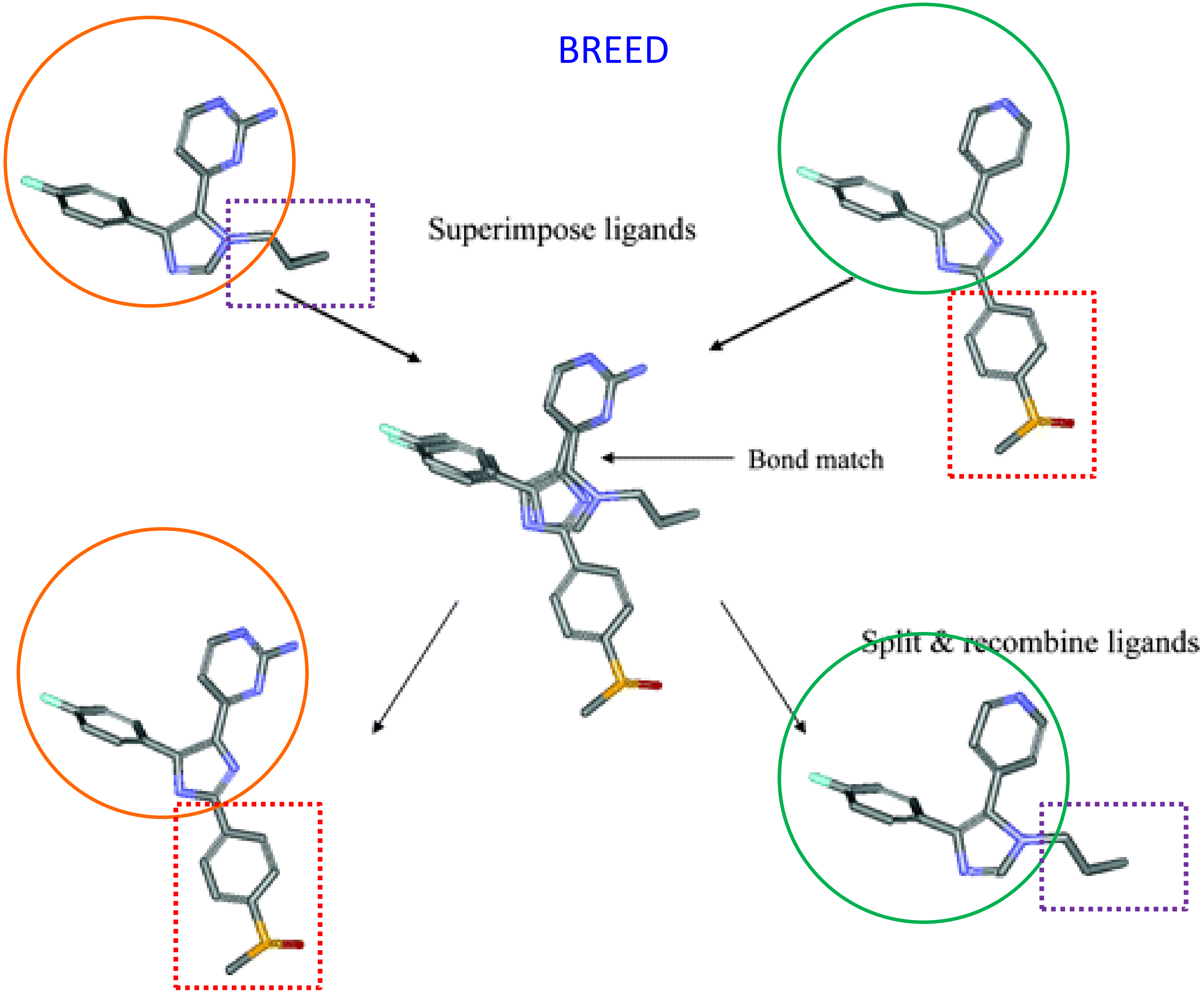
MUTATE

LINK

REDUCE

Fragment-based structure generation

BREED

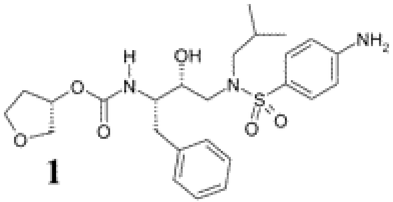


Pierce A.C., Rao G., Bemis G.W. *J. Med. Chem.*, **2004**, 47 (11), pp 2768–2775

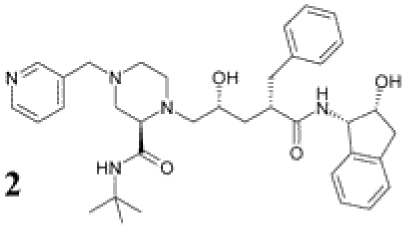
Fragment-based structure generation

BREED: HIV-1 protease inhibitors

K_i = 0.4-0.6 nM



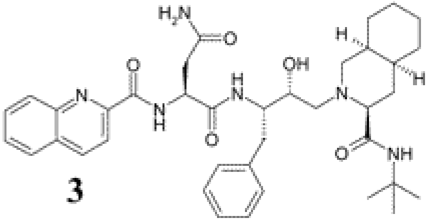
1



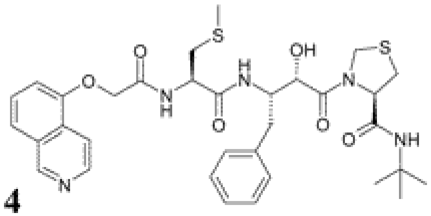
2

K_d = 1.1 nM

K_i = 1.7 nM



3



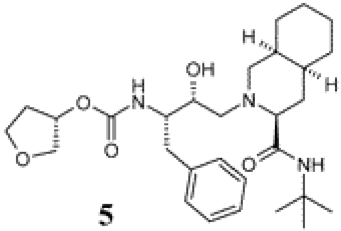
4

K_d = 0.3 nM

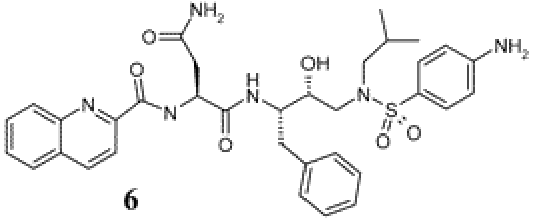
known

designed

IC_{50} = 160 nM



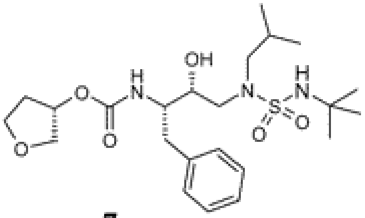
5



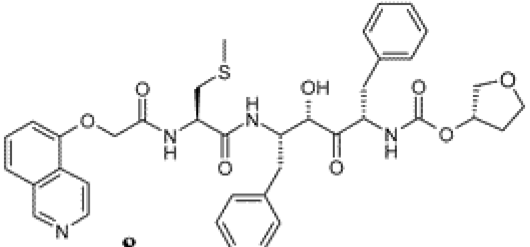
6

K_i = 0.1 nM

K_i = 42 nM



7

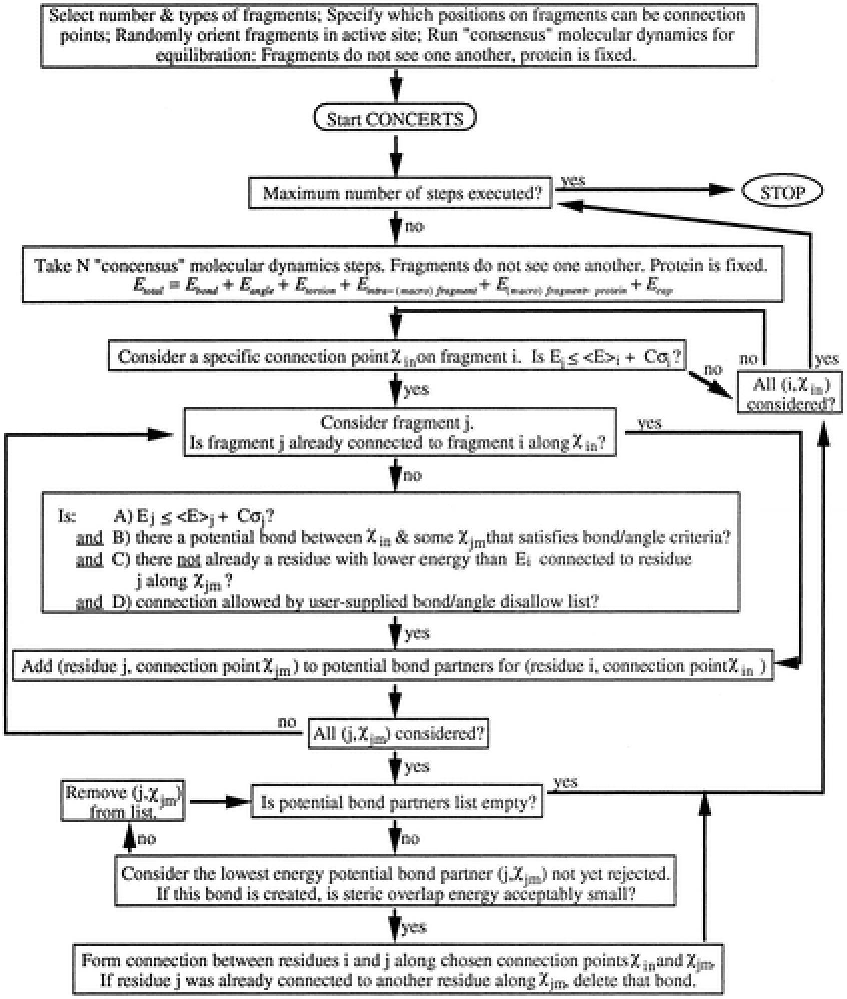


8

Pierce A.C., Rao G., Bemis G.W. *J. Med. Chem.*, **2004**, 47 (11), pp 2768–2775

Fragment-based structure generation

CONCEPTS



Select number & types of fragments; Specify which positions on fragments can be connection points; Randomly orient fragments in active site; Run "consensus" molecular dynamics for equilibration: Fragments do not see one another, protein is fixed.

Start CONCERTS

Maximum number of steps executed?

STOP

Take N "concensus" molecular dynamics steps. Fragments do not see one another. Protein is fixed.

$E_{total} = E_{bond} + E_{angle} + E_{torsion} + E_{intra-(macro)\ fragment} + E_{(macro)\ fragment-protein} + E_{cap}$

Consider a specific connection point χ_{in} on fragment i. Is $E_i \leq \langle E \rangle_i + C\sigma_i$?

All (i, χ_{in}) considered?

Consider fragment j.
Is fragment j already connected to fragment i along χ_{in}?

Is: A) $E_j \leq \langle E \rangle_j + C\sigma_j$?
and B) there a potential bond between χ_{in} & some χ_{jm} that satisfies bond/angle criteria?
and C) there not already a residue with lower energy than E_i connected to residue j along χ_{jm}?
and D) connection allowed by user-supplied bond/angle disallow list?

Add (residue j, connection point χ_{jm}) to potential bond partners for (residue i, connection point χ_{in})

All (j, χ_{jm}) considered?

Remove (j, χ_{jm}) from list.

Is potential bond partners list empty?

Consider the lowest energy potential bond partner (j, χ_{jm}) not yet rejected.
If this bond is created, is steric overlap energy acceptably small?

Form connection between residues i and j along chosen connection points χ_{in} and χ_{jm}.
If residue j was already connected to another residue along χ_{jm}, delete that bond.

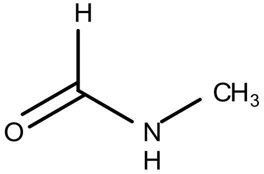
MD of fragments which are linking or breaking during the simulation in order to create more favorable structures

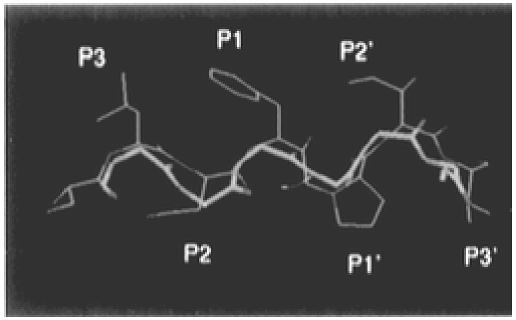
formation of certain bonds was forbidden:
O−O, N−N, N−O, S−O, O−C−O, O−N−O, N−C−N, C_α−C_α, C−C_α−C

Pearlman D.A., Murcko M.A. *J. Med. Chem.*, **1996**, 39 (8), pp 1651–1663

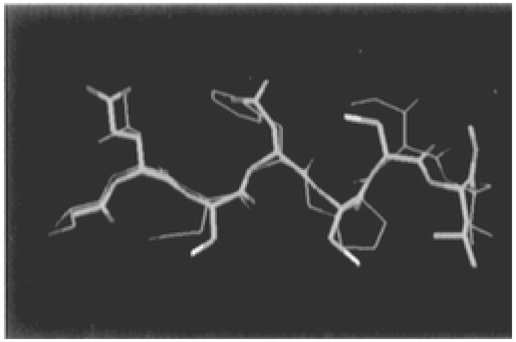
Fragment-based structure generation

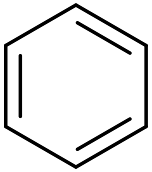
CONCEPTS: HIV-1 protease inhibitors

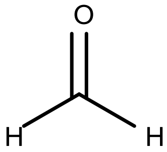




+ 19 side chains

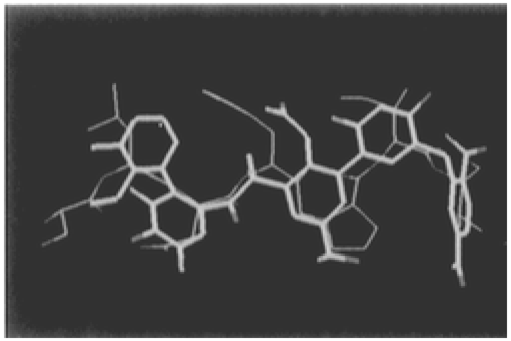






CH_4 H_2O

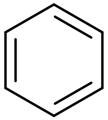
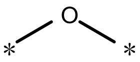
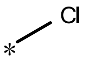
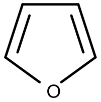
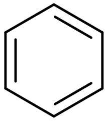
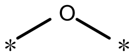
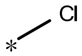
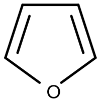
NH_3

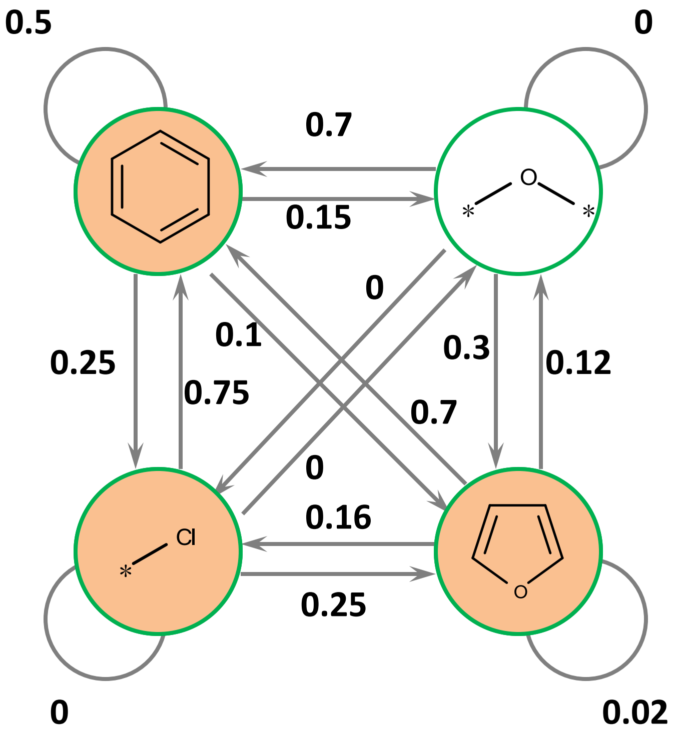


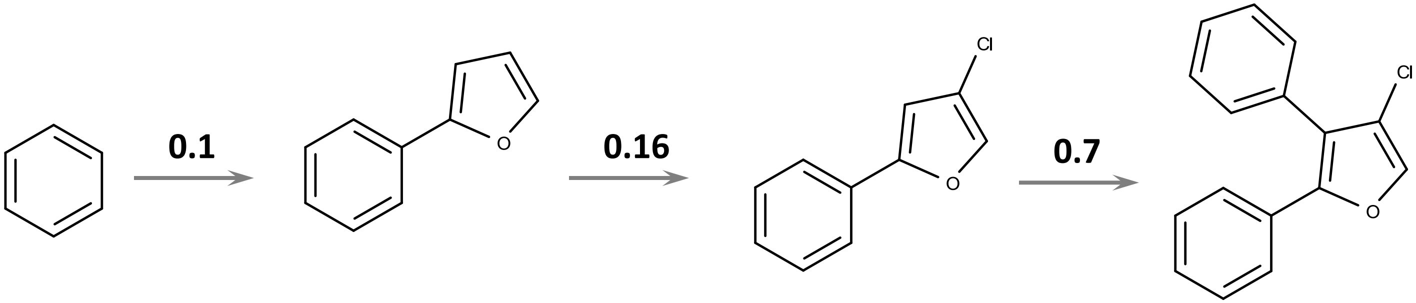
Pearlman D.A., Murcko M.A. *J. Med. Chem.*, **1996**, 39 (8), pp 1651–1663

Fragment-based structure generation

FOG

				
	0.5	0.15	0.25	0.1
	0.7	0	0	0.3
	0.75	0	0	0.25
	0.7	0.12	0.16	0.02

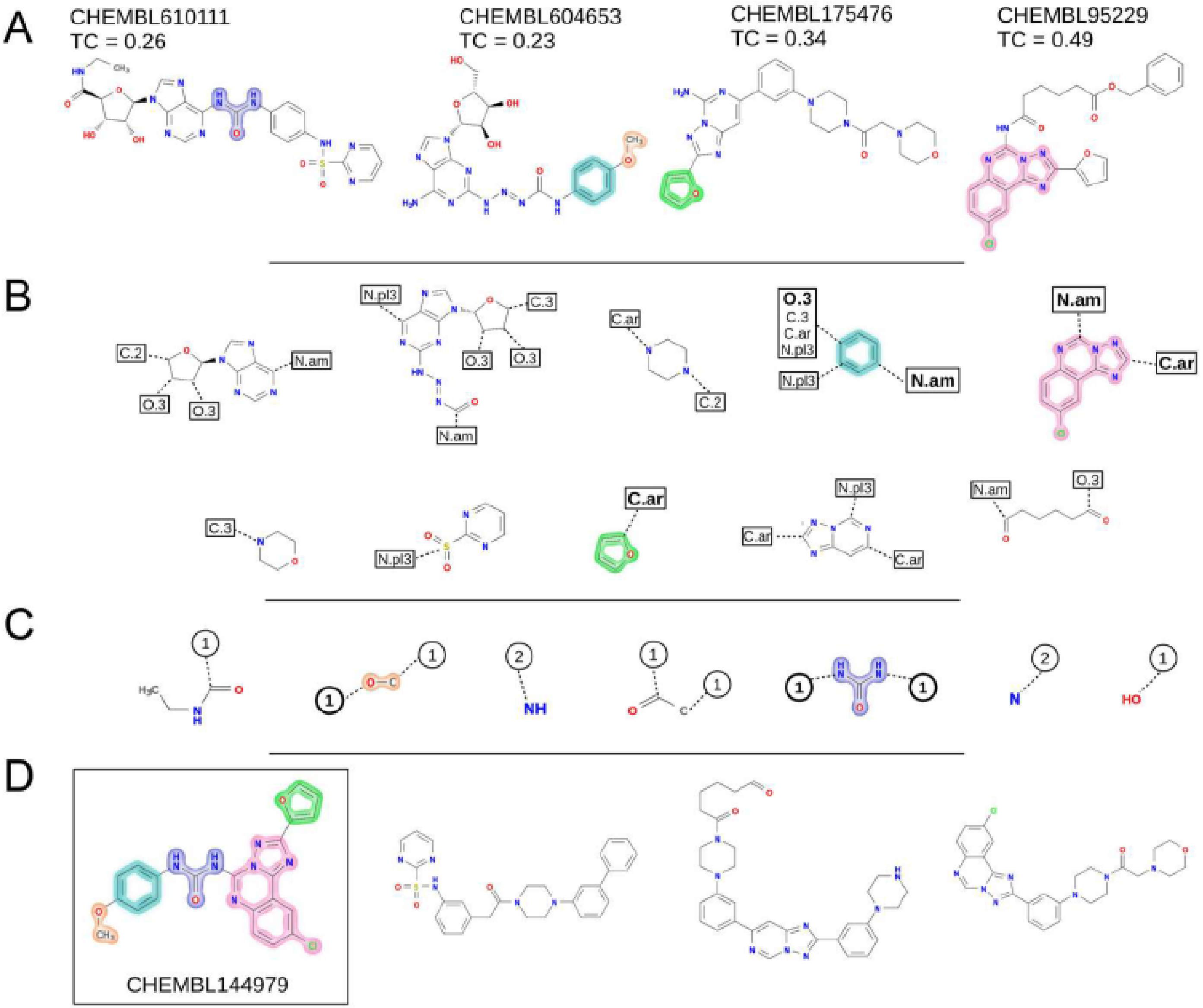




Kutchukian, P. S.; Lou, D.; Shakhnovich, E. I., FOG: Fragment Optimized Growth Algorithm for the de Novo Generation of Molecules Occupying Druglike Chemical Space. *Journal of Chemical Information and Modeling* **2009**, 49, 1630-1642.

Fragment-based structure generation

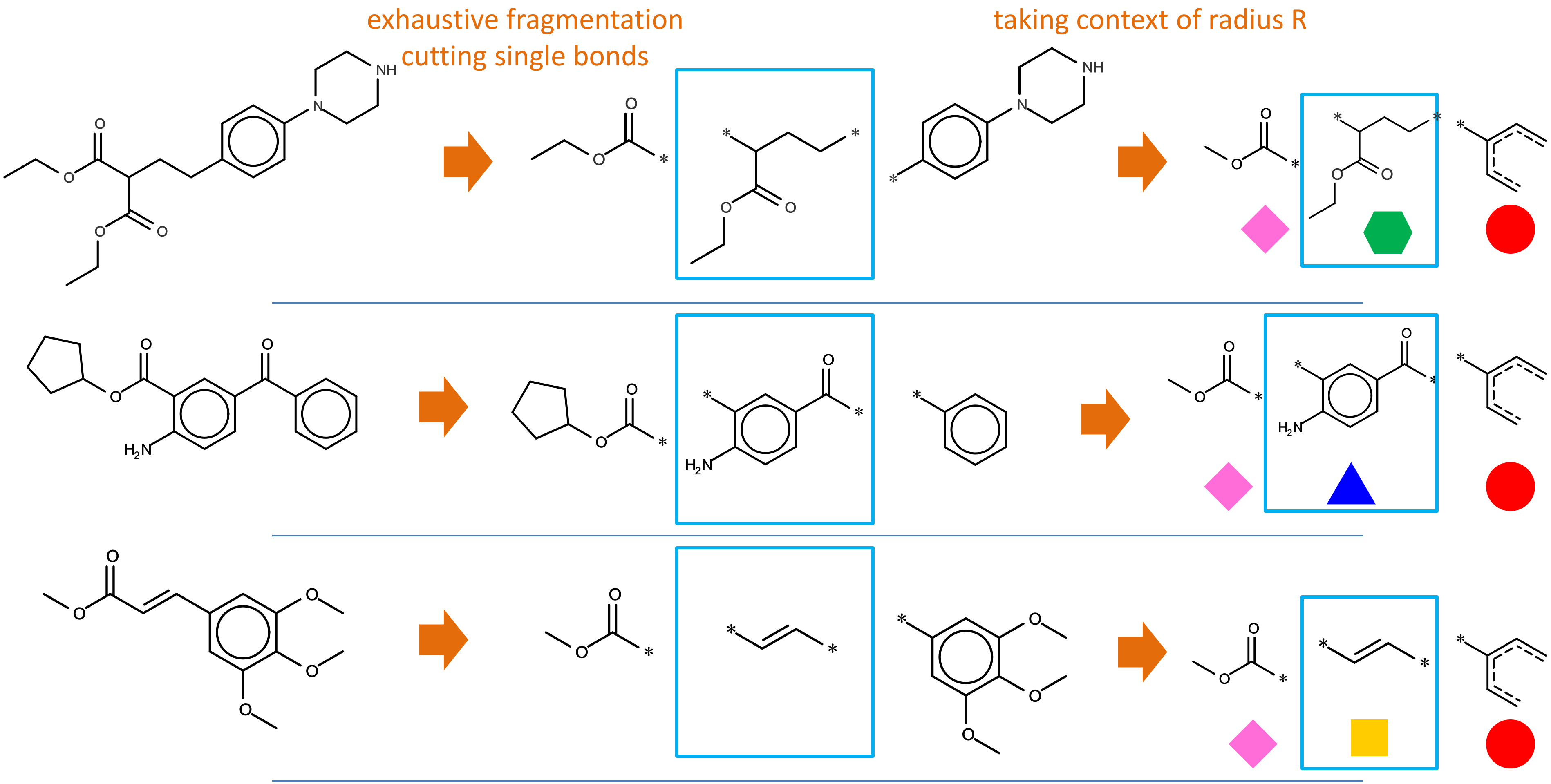
eMolFrag



Liu, T.; Naderi, M.; Alvin, C.; Mukhopadhyay, S.; Brylinski, M., Break Down in Order To Build Up: Decomposing Small Molecules for Fragment-Based Drug Design with eMolFrag. *J. Chem. Inf. Model.* **2017**, 57, 627-631

Fragment-based structure generation

CReM: chemically reasonable mutations


exhaustive fragmentation cutting single bonds

taking context of radius R

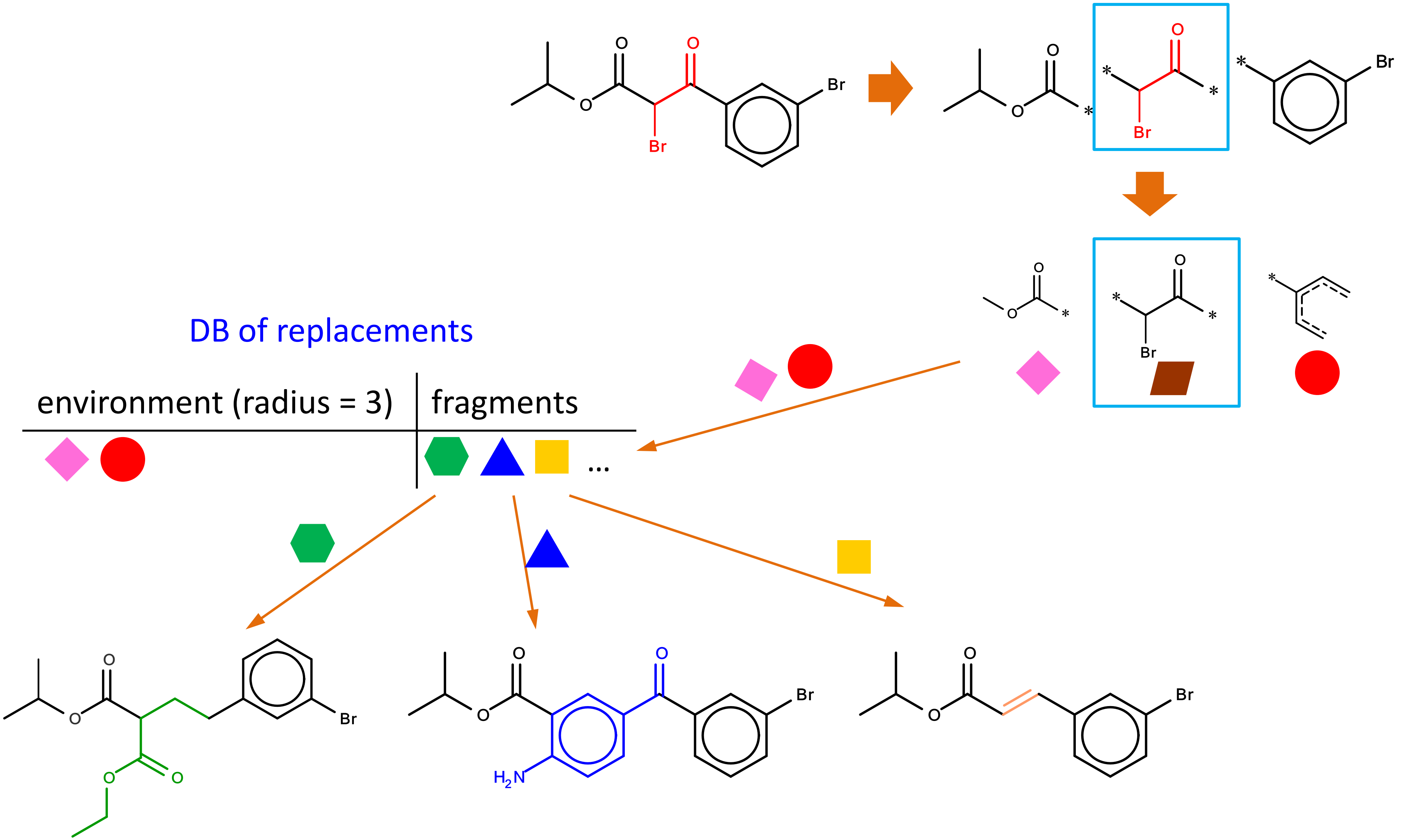
DB of replacements

environment (radius = 3)	fragments
◆ ●	⬢ ▲ ■ …
…	…

mutually exchangeable fragments

Fragment-based structure generation

CReM: chemically reasonable mutations



DB of replacements

environment (radius = 3)	fragments
◆ ●	⬢ ▲ ■ …

Generated structures are always chemically valid!

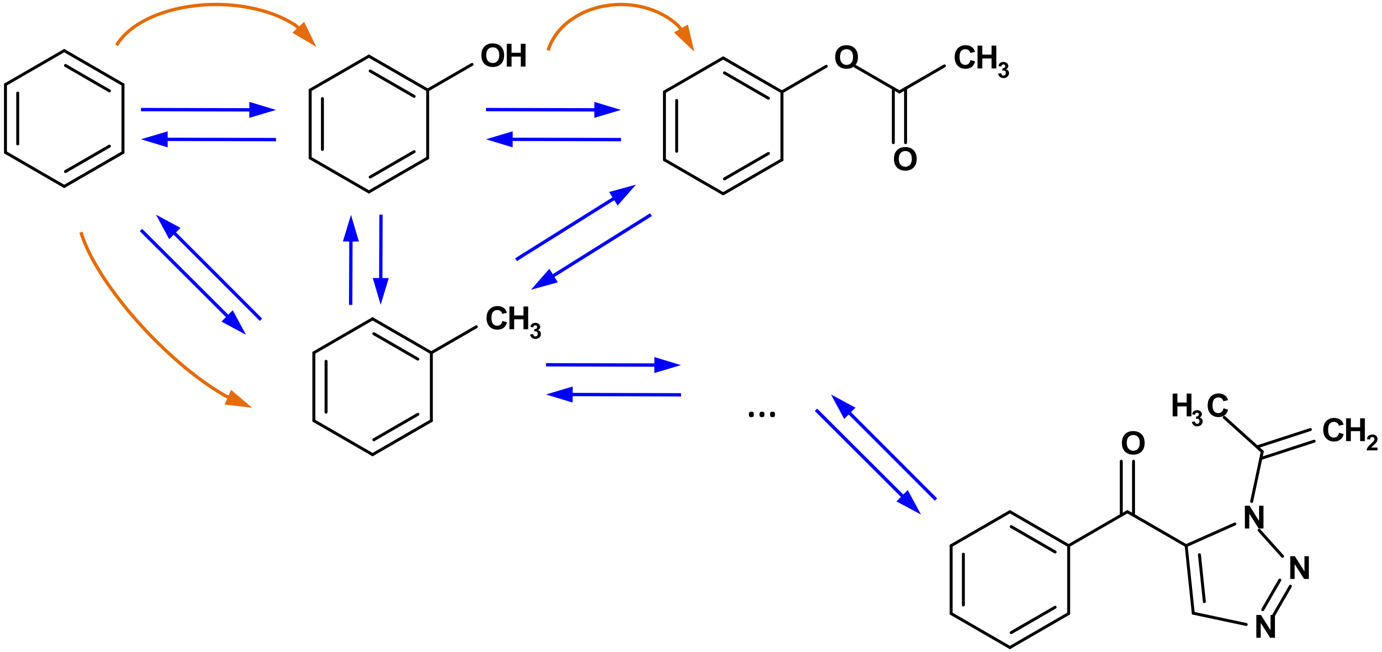
Fragment-based structure generation

	fragment-based
exhaustiveness of chemical space search	++ variable, controlled by the size of fragments to replace
structure novelty	++
structure diversity	++
chemically valid structures	(+++)
synthetically feasible	(++)
combinatorial explosion / time consuming	++

fragment-based ≈ semi-empirical

Reaction-based *vs.* fragment-based

	Reaction-based	Fragment-based
Prerequisites:	reaction rules set database of building blocks	database of fragments
Abilities & issues:	• molecules are more likely to be feasible • not all moves are allowed • usually only increase complexity • some molecules can be unreachable	• do not control synthetic feasibility • many moves are allowed • arbitrary direction of exploration • cover larger chemical space



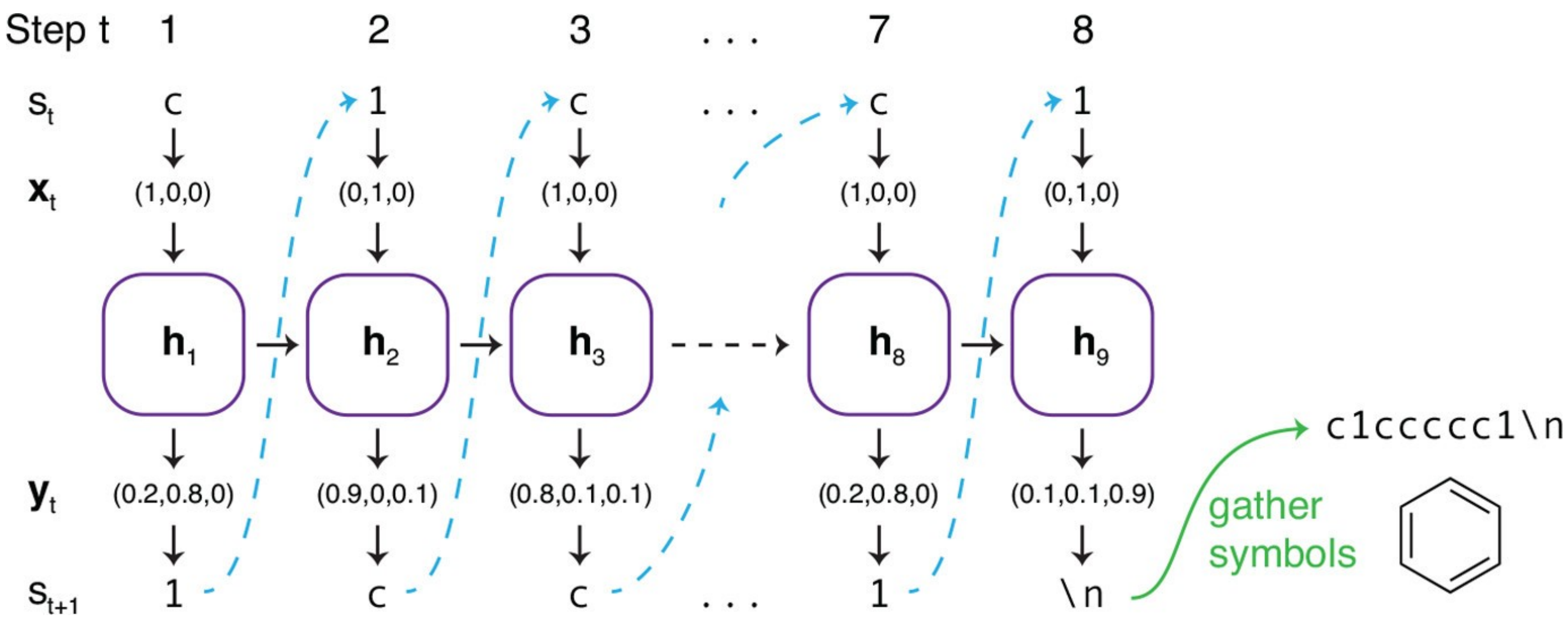
De novo structure generation

Summary

	atom-based	fragment-based	reaction-based
exhaustiveness of chemical space search	++++ very small steps; more suitable for systematic exploration of local chemical space	++ variable, controlled by the size of fragments to replace	+ depends on reactant library and reaction rules; only grow molecules
structure novelty	+++*	++	++
structure diversity	+++*	++	++
chemically valid structures	-	(+++)	+++
synthetically feasible	---	(++)	+++
combinatorial explosion / time consuming	---	++	+++

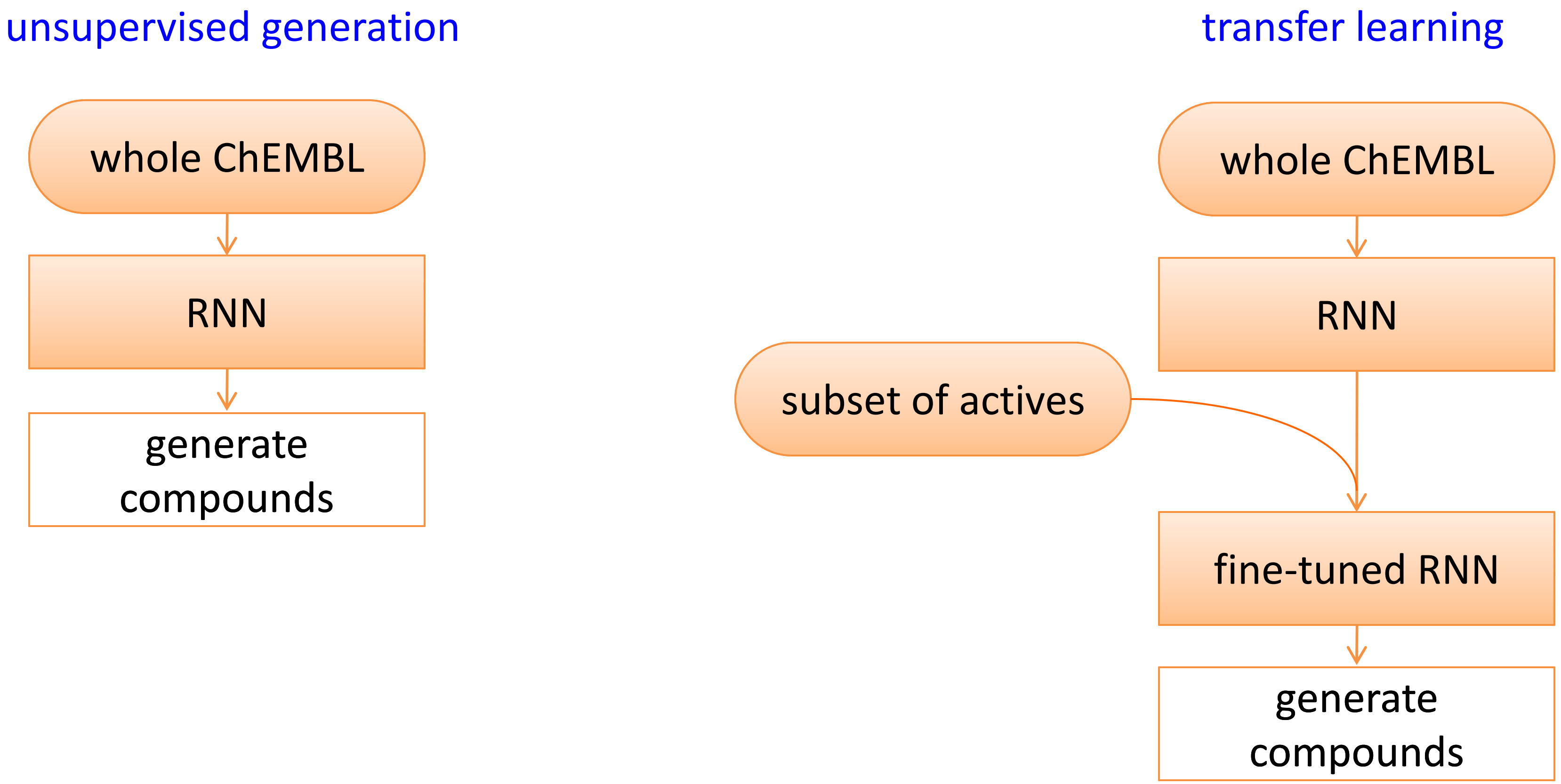
Deep learning models for structure generation

Recurrent neural network (RNN)



Segler, M. H. S.; Kogej, T.; Tyrchan, C.; Waller, M. P., Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Central Science* **2018**, 4, 120-131.

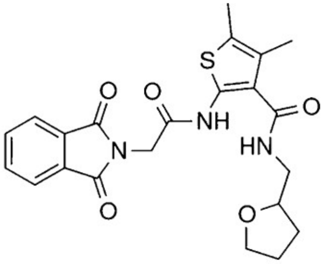
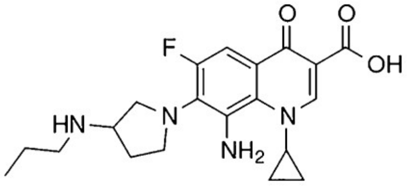
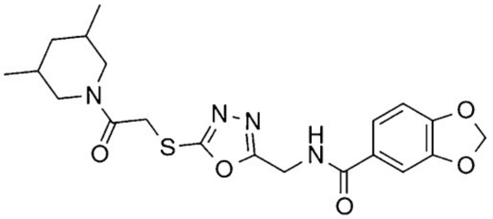
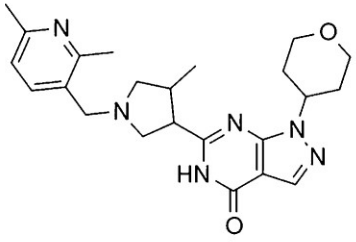
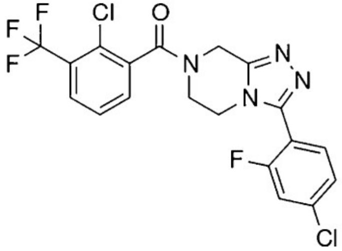
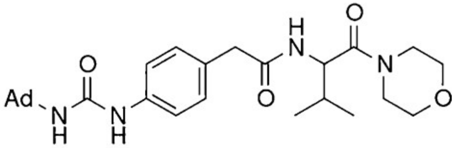
Deep learning models for structure generation



Segler, M. H. S.; Kogej, T.; Tyrchan, C.; Waller, M. P., Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Central Science* **2018**, 4, 120-131.

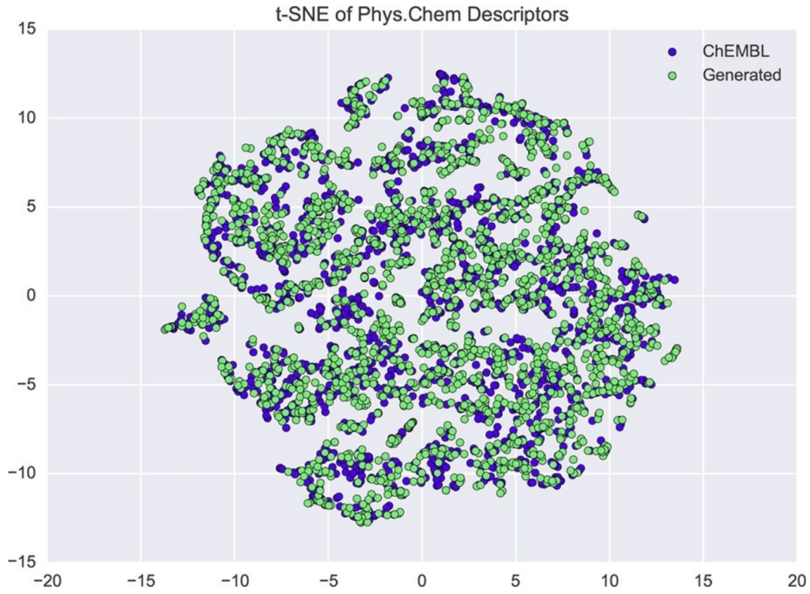
Deep learning models for structure generation

unsupervised generation

976 327 compounds
97.7% chemically valid
11.5% were duplicated with ChEMBL
1.7% of duplicates

75% passed AZ filters (similar to ChEMBL)
12% of scaffolds were common with ChEMBL

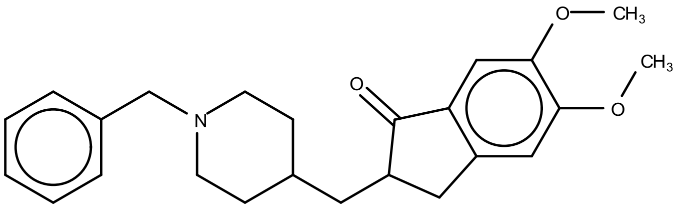


Segler, M. H. S.; Kogej, T.; Tyrchan, C.; Waller, M. P., Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Central Science* **2018**, 4, 120-131.

Deep learning models for structure generation

	deep learning
exhaustiveness of chemical space search	++
structure novelty	++
structure diversity	++
chemically valid structures	++
synthetically feasible	?
combinatorial explosion / time consuming	+++

Issue of SMILES based representation -
the same structure can be represented by different SMILES



COc1cc2CC(CC3CCN(Cc4ccccc4)CC3)C(=O)c2cc1OC

COc1cc2c(cc1OC)C(=O)C(CC1CCN(Cc3ccccc3)CC1)C2

Scoring/objective functions

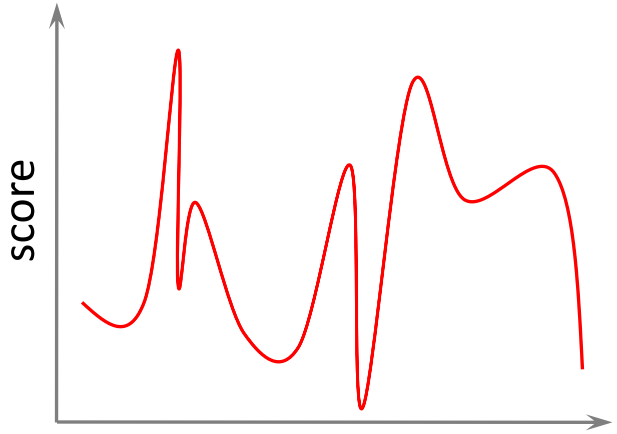
Can be any but preferably smooth to follow the chemical similarity principle:

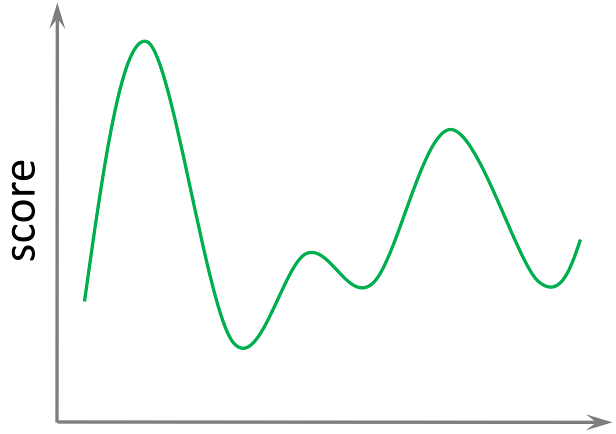
- physicochemical properties
- similarity measures
- QSAR model prediction
- pharmacophore fit
- docking scoring
- molecular dynamics

ligand-based scoring functions

structure-based scoring functions

…





Search algorithms

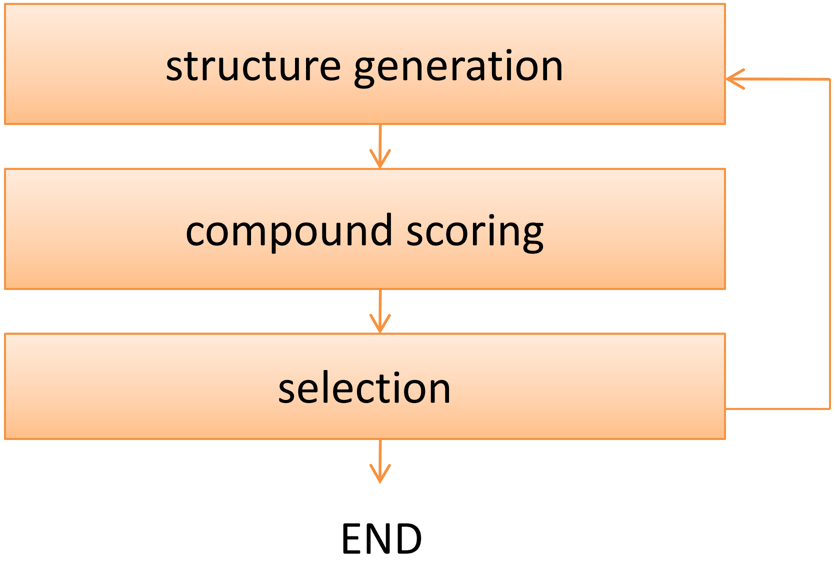
Can be any , for example:

- greedy search
- Monte Carlo
- evolutionary algorithms, e.g.:
    - genetic algorithm
- simulated annealing

…

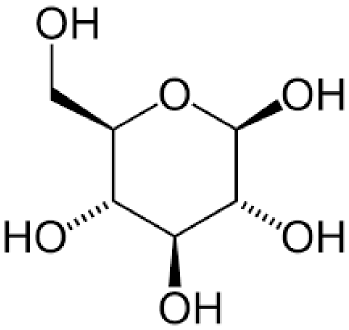
Iterative workflow of de novo design

1. **Structure generation** - how to create/assembly new structures
2. **Compound scoring** - how to estimate/predict a property of a compound
3. **Search strategy** - how to find compounds with optimal properties

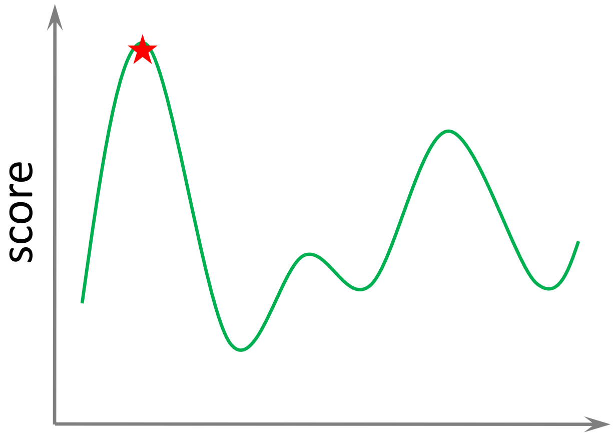


iterative strategy

Inverse QSAR



D_1	D_2	D_3	…	D_N
1	0	9	…	1
4	0	1	…	1
0	2	3	…	3
…	…	…	…	…
4	0	0	…	1

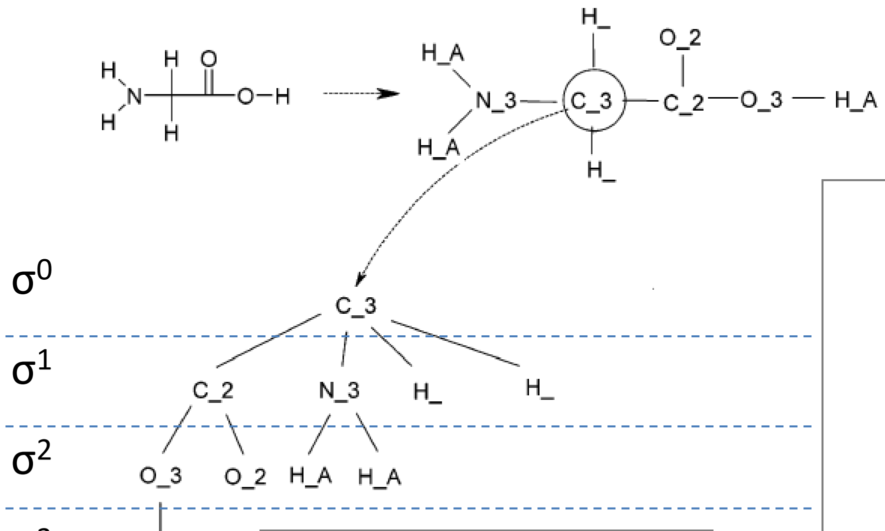


D_1	D_2	D_3	…	D_N
11	3	1	…	15

STRUCTURE ?

Inverse QSAR

Atom signatures



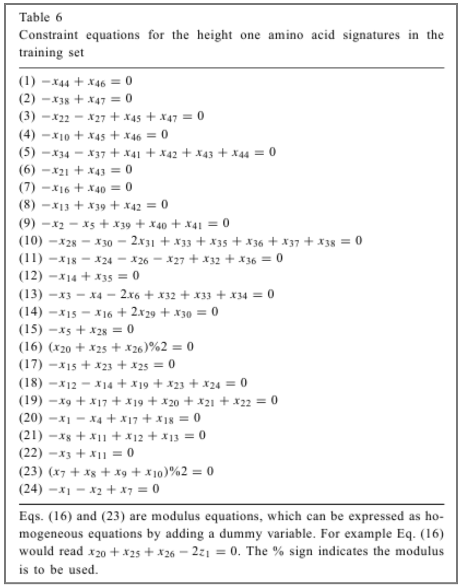
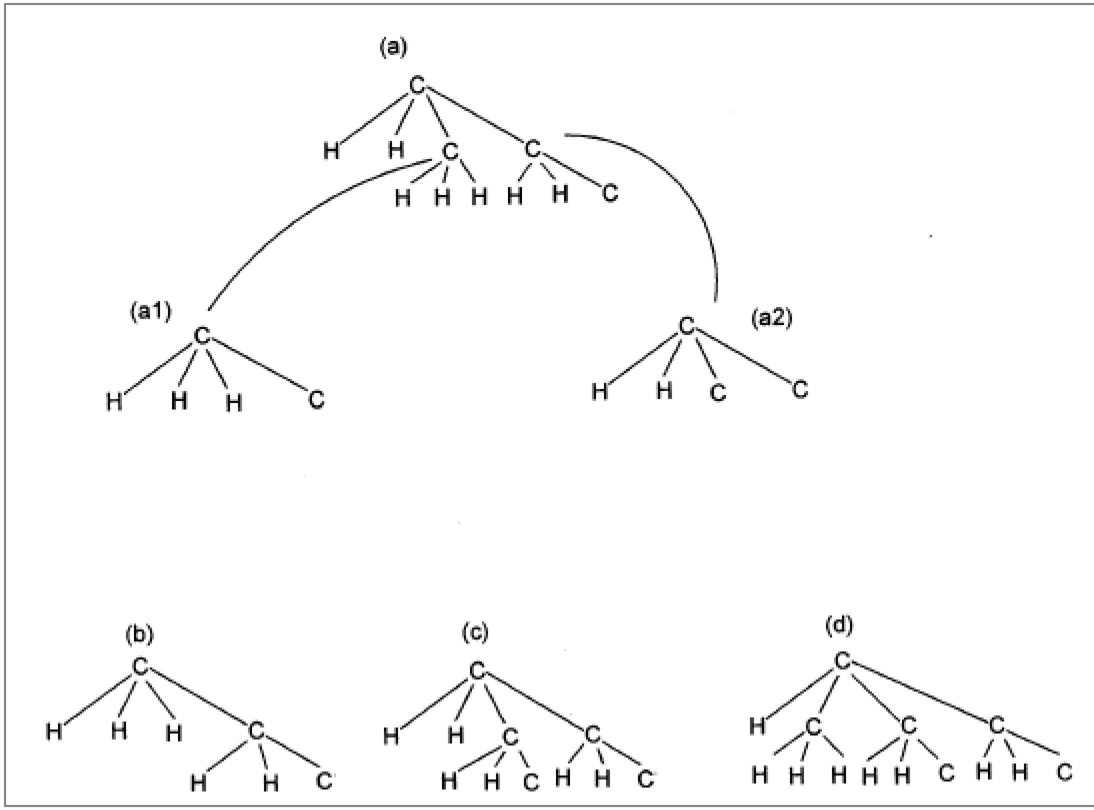


Table 6
Constraint equations for the height one amino acid signatures in the training set

(1) $-x_{44} + x_{46} = 0$
(2) $-x_{38} + x_{47} = 0$
(3) $-x_{22} - x_{27} + x_{45} + x_{47} = 0$
(4) $-x_{10} + x_{45} + x_{46} = 0$
(5) $-x_{34} - x_{37} + x_{41} + x_{42} + x_{43} + x_{44} = 0$
(6) $-x_{21} + x_{43} = 0$
(7) $-x_{16} + x_{40} = 0$
(8) $-x_{13} + x_{39} + x_{42} = 0$
(9) $-x_{2} - x_{5} + x_{39} + x_{40} + x_{41} = 0$
(10) $-x_{28} - x_{30} - 2x_{31} + x_{33} + x_{35} + x_{36} + x_{37} + x_{38} = 0$
(11) $-x_{18} - x_{24} - x_{26} - x_{27} + x_{32} + x_{36} = 0$
(12) $-x_{14} + x_{35} = 0$
(13) $-x_{3} - x_{4} - 2x_{6} + x_{32} + x_{33} + x_{34} = 0$
(14) $-x_{15} - x_{16} + 2x_{29} + x_{30} = 0$
(15) $-x_{5} + x_{28} = 0$
(16) $(x_{20} + x_{25} + x_{26})\%2 = 0$
(17) $-x_{15} + x_{23} + x_{25} = 0$
(18) $-x_{12} - x_{14} + x_{19} + x_{23} + x_{24} = 0$
(19) $-x_{9} + x_{17} + x_{19} + x_{20} + x_{21} + x_{22} = 0$
(20) $-x_{1} - x_{4} + x_{17} + x_{18} = 0$
(21) $-x_{8} + x_{11} + x_{12} + x_{13} = 0$
(22) $-x_{3} + x_{11} = 0$
(23) $(x_{7} + x_{8} + x_{9} + x_{10})\%2 = 0$
(24) $-x_{1} - x_{2} + x_{7} = 0$

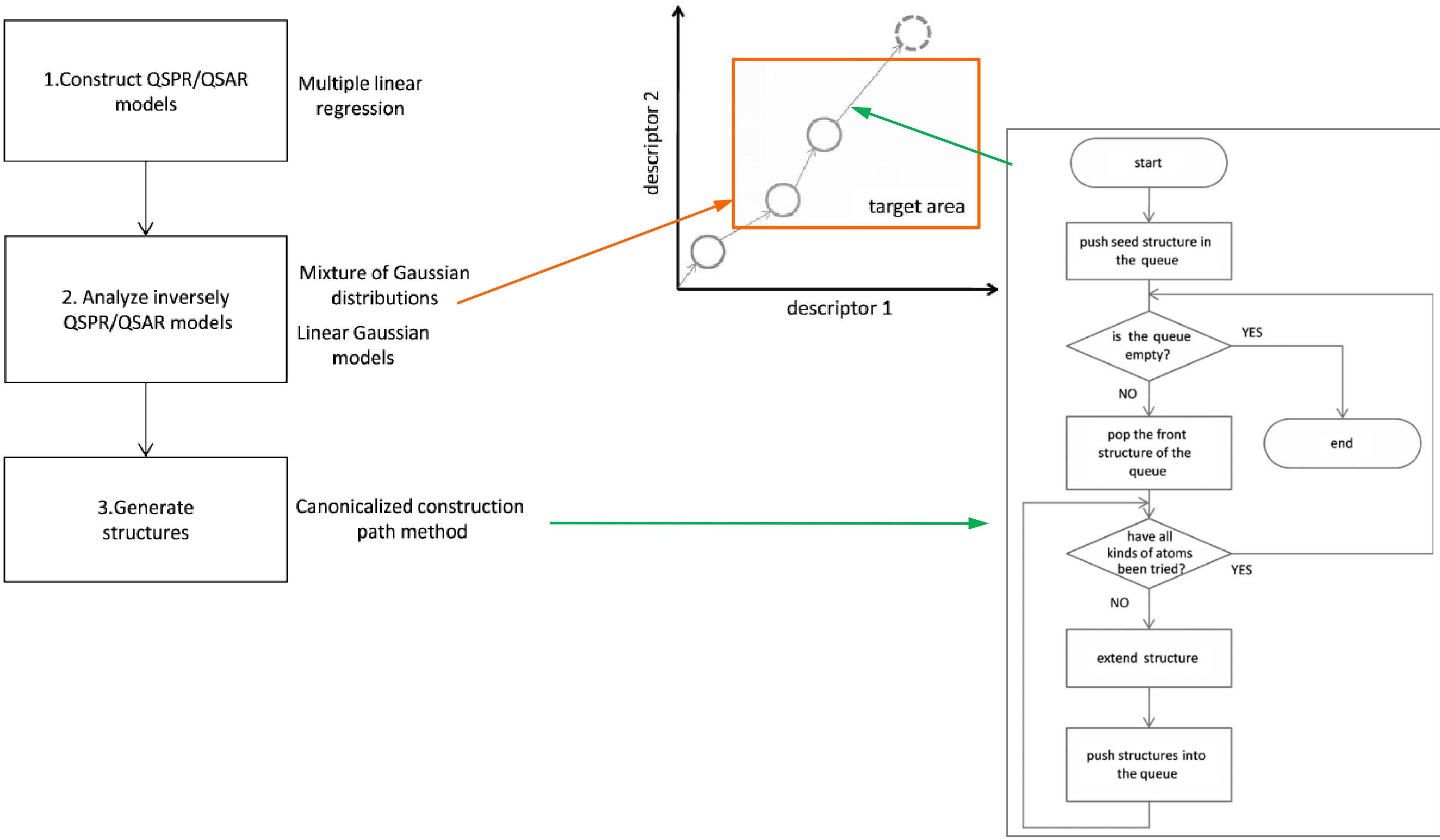
Eqs. (16) and (23) are modulus equations, which can be expressed as homogeneous equations by adding a dummy variable. For example Eq. (16) would read $x_{20} + x_{25} + x_{26} - 2z_{1} = 0$. The % sign indicates the modulus is to be used.



Faulon J-L, Churchwell CJ, Visco DP, The Signature Molecular Descriptor. 2. Enumerating Molecules from Their Extended Valence Sequences – J. Chem. Inf. Comput. Sci., **2003**, 43 (3), pp 721–734

Inverse QSAR

Inverse QSAR with monotonically changed descriptors



1.Construct QSPR/QSAR models

Multiple linear regression

2. Analyze inversely QSPR/QSAR models

Mixture of Gaussian distributions

Linear Gaussian models

3.Generate structures

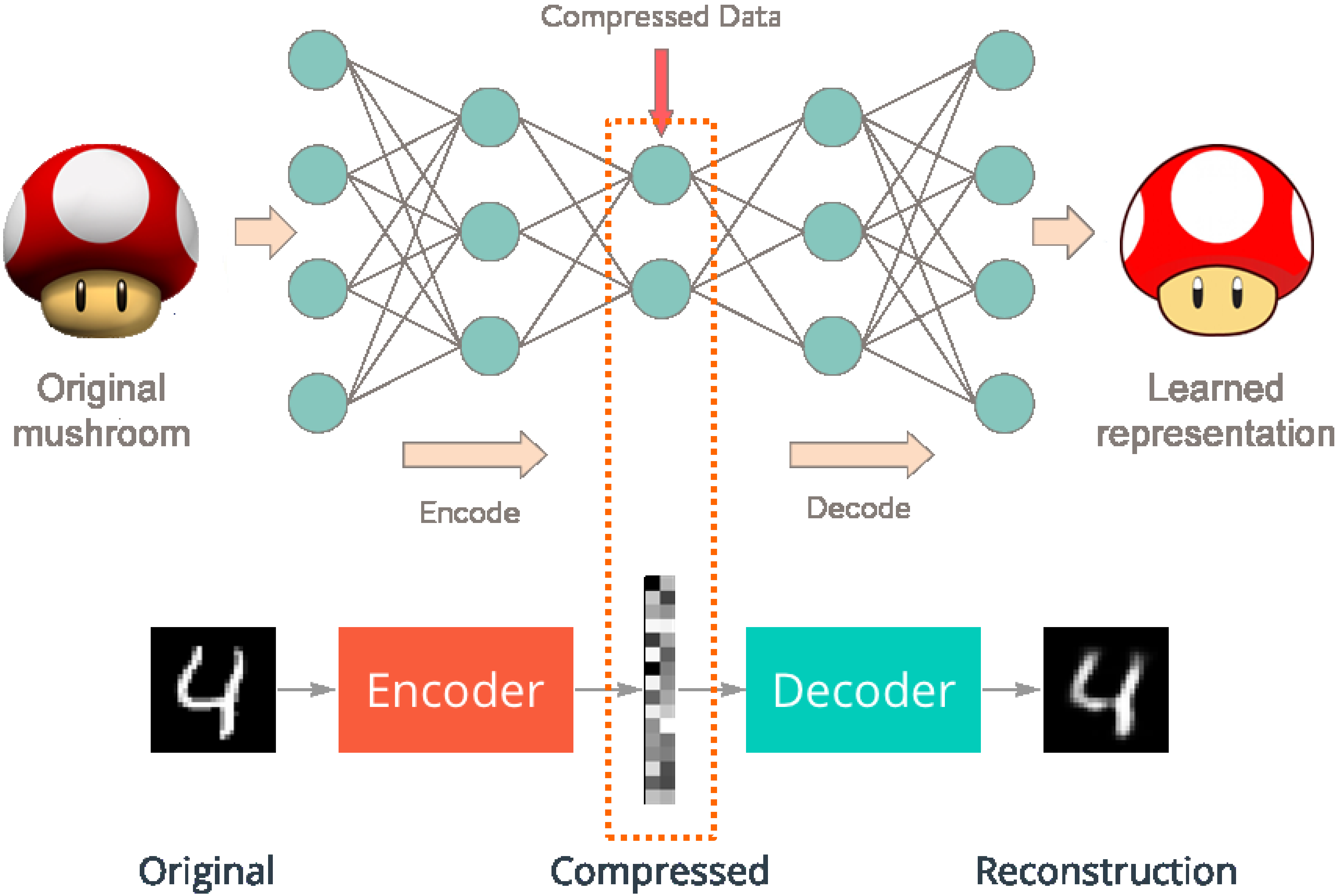
Canonicalized construction path method

descriptor 2
target area
descriptor 1
start
push seed structure in the queue
is the queue empty?
YES
NO
end
pop the front structure of the queue
have all kinds of atoms been tried?
YES
NO
extend structure
push structures into the queue

Miyao, T.; Arakawa, M.; Funatsu, K., Exhaustive Structure Generation for Inverse-QSPR/QSAR. *Molecular Informatics* ***2010***, 29, 111-125.

Inverse QSAR: deep learning

Autoencoder



Compressed Data

Original mushroom

Learned representation

Encode

Decode

Encoder

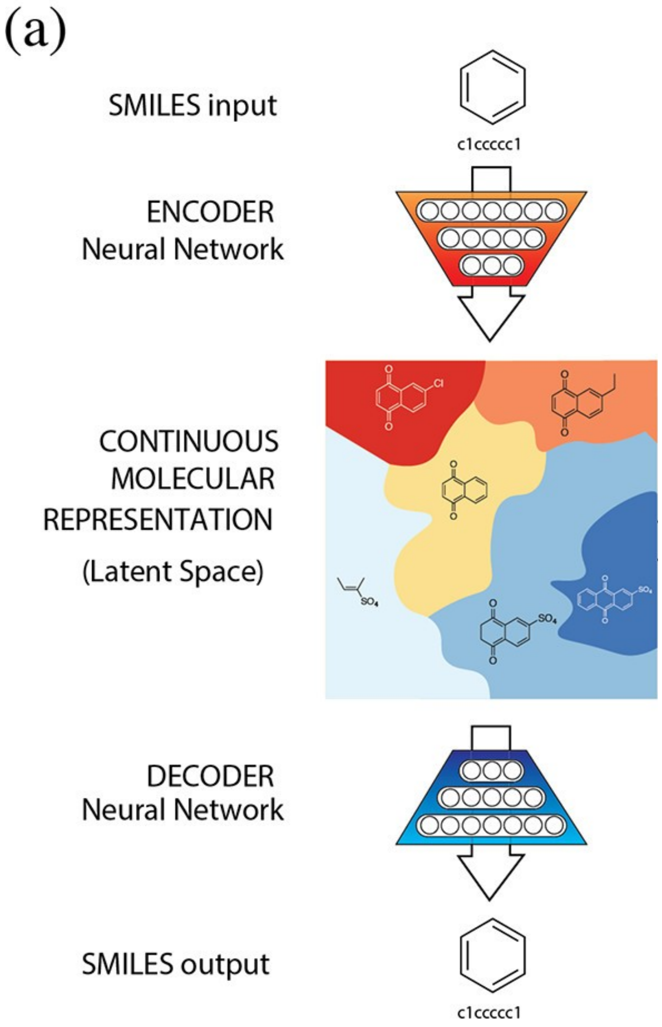
Decoder

Original

Compressed

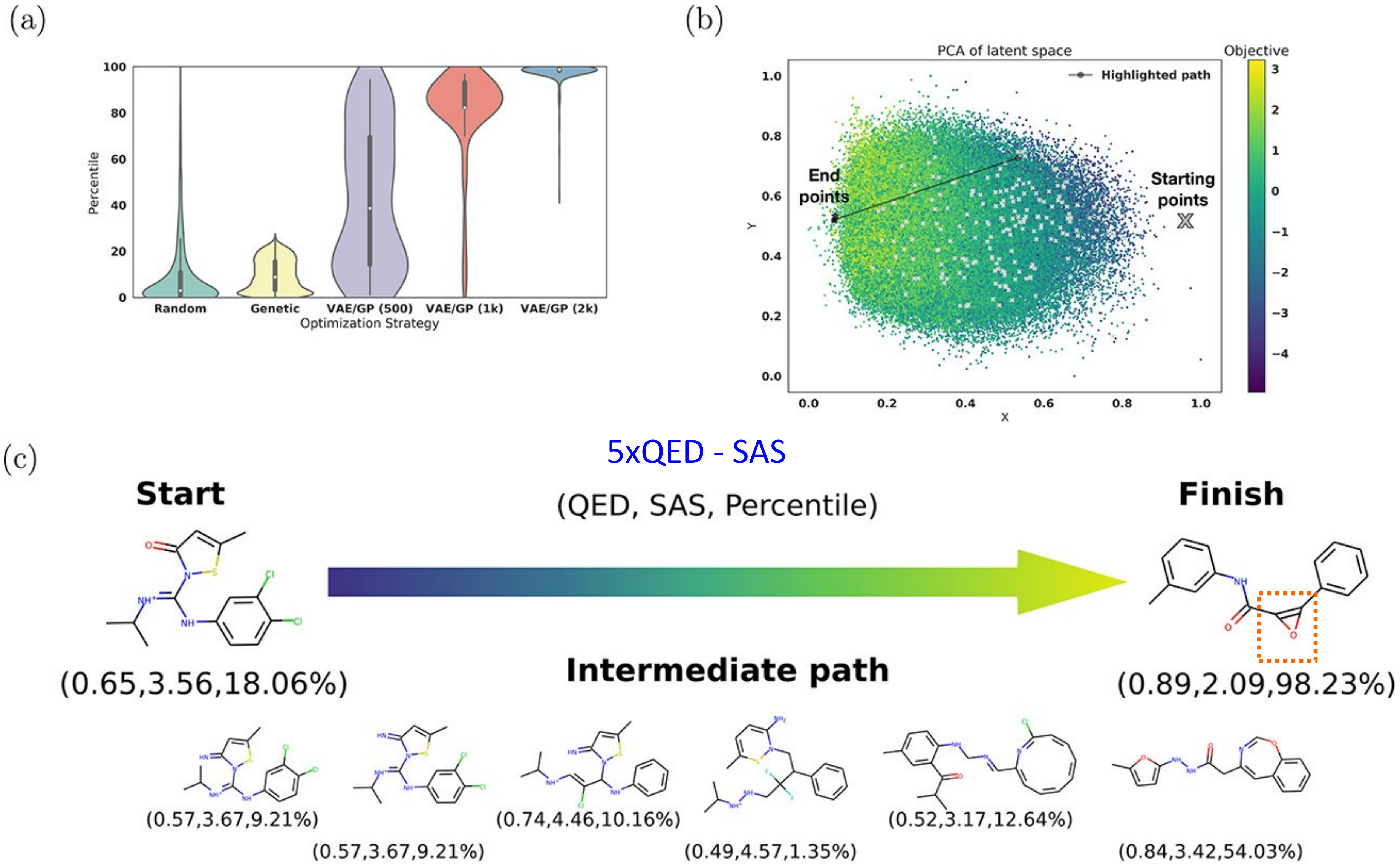
Reconstruction

Inverse QSAR: deep learning



Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A., Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Central Science* **2018**, 4, 268-276.

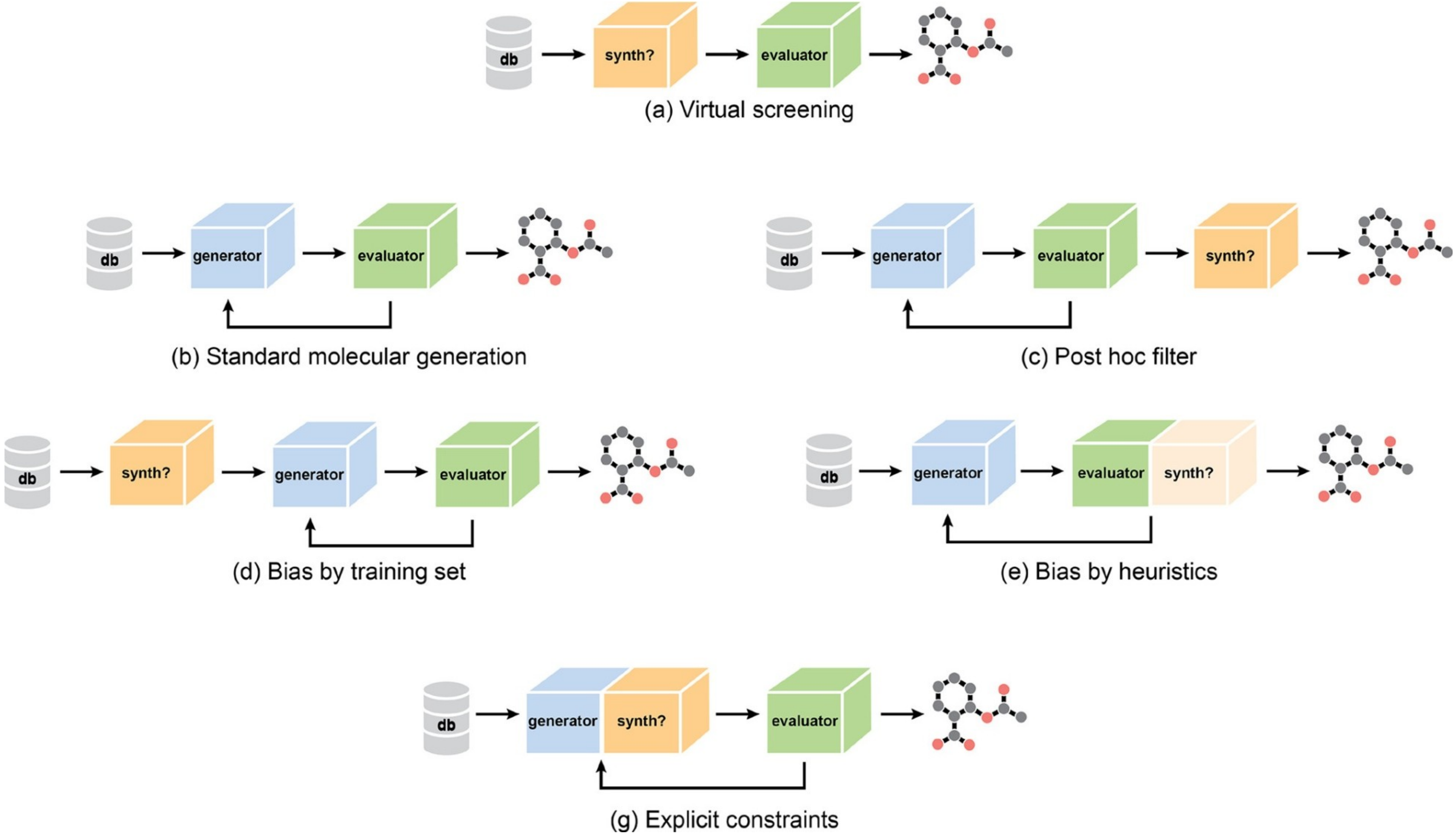
Inverse QSAR: deep learning



5xQED - SAS

Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A., Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Central Science* **2018**, 4, 268-276.

Control over synthetic feasibility



(a) Virtual screening

(b) Standard molecular generation

(c) Post hoc filter

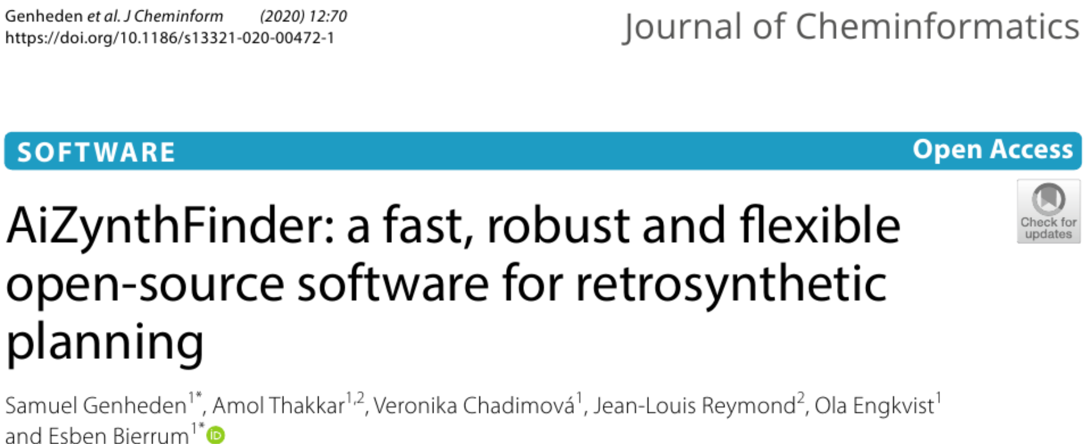
(d) Bias by training set

(e) Bias by heuristics

(g) Explicit constraints

Gao, W.; Coley, C. W., The Synthesizability of Molecules Proposed by Generative Models. *Journal of Chemical Information and Modeling* **2020**

Assessment of synthetic feasibility



Genheden *et al. J Cheminform* (2020) 12:70
https://doi.org/10.1186/s13321-020-00472-1

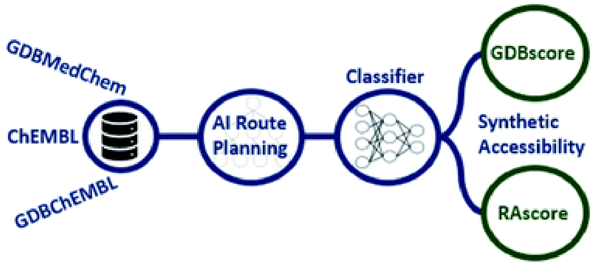
Journal of Cheminformatics

SOFTWARE Open Access

Check for updates

AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning

Samuel Genheden[1*], Amol Thakkar[1,2], Veronika Chadimová[1], Jean-Louis Reymond[2], Ola Engkvist[1] and Esben Bjerrum[1*]



Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Sci.*, 2021, **12**, 3339

All publication charges for this article have been paid for by the Royal Society of Chemistry

Retrosynthetic accessibility score (RAscore) – rapid machine learned synthesizability classification from AI driven retrosynthetic planning†

Amol Thakkar,[*ab] Veronika Chadimová,[a] Esben Jannik Bjerrum,[a] Ola Engkvist[a] and Jean-Louis Reymond[*b]

Voršilák *et al. J Cheminform* (2020) 12:35
https://doi.org/10.1186/s13321-020-00439-2

Journal of Cheminformatics

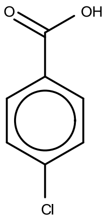
RESEARCH ARTICLE Open Access

Check for updates

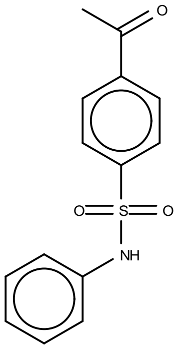
SYBA: Bayesian estimation of synthetic accessibility of organic compounds

Milan Voršilák[1,2], Michal Kolář[3,4], Ivan Čmelo[1] and Daniel Svozil[1,2*]

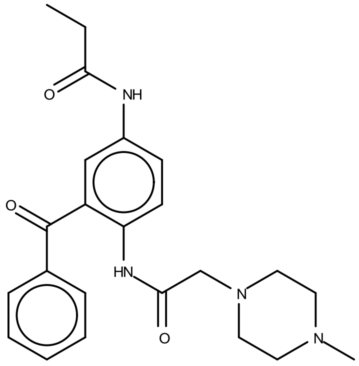
Examples of SA scores (ChEMBL22)



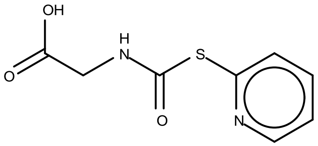
1.2
CHEMBL618



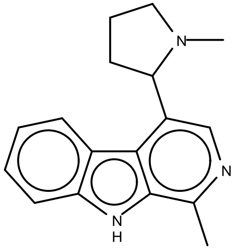
1.5
CHEMBL3310985



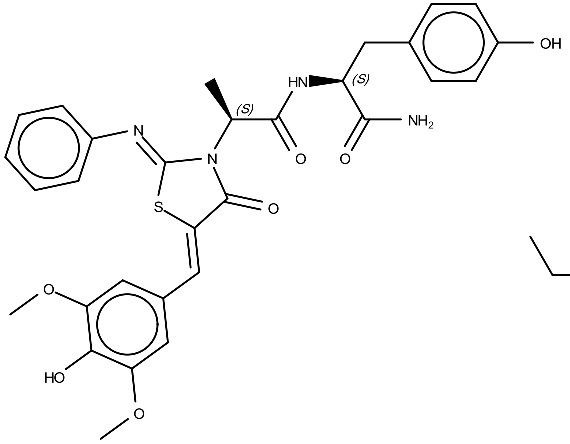
2.0
CHEMBL595820



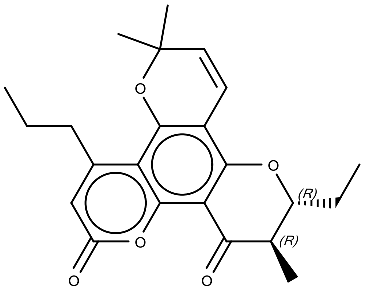
2.5
CHEMBL503660



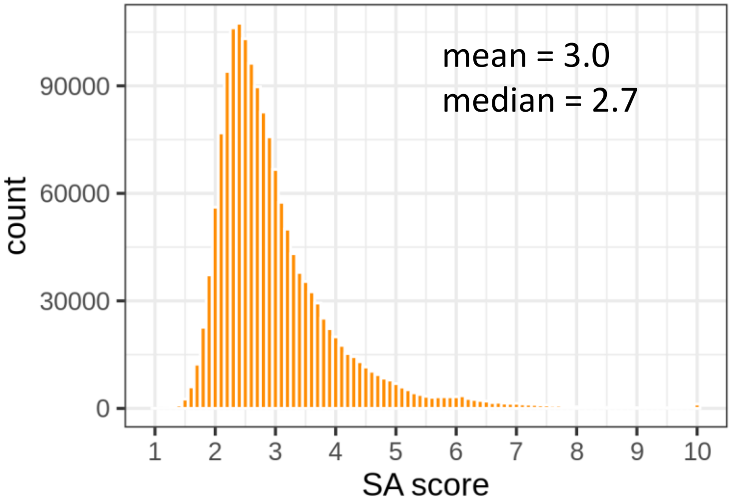
3.0
CHEMBL500286



3.5
CHEMBL582554



4.0
CHEMBL7633



Ertl, P.; Schuffenhauer, A., Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of Cheminformatics* **2009**, 1, 8. (10.1186/1758-2946-1-8)

Control of synthetic feasibility within CReM

Content of fragmented library

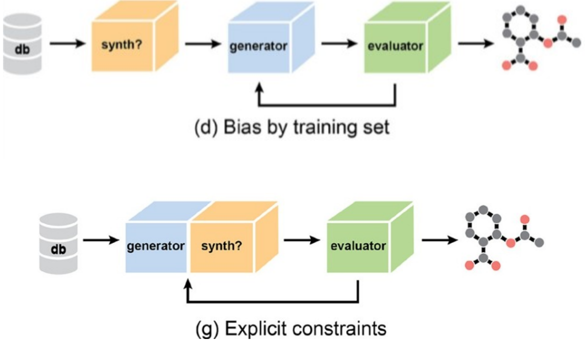
all ChEMBL compounds (1 554 160)

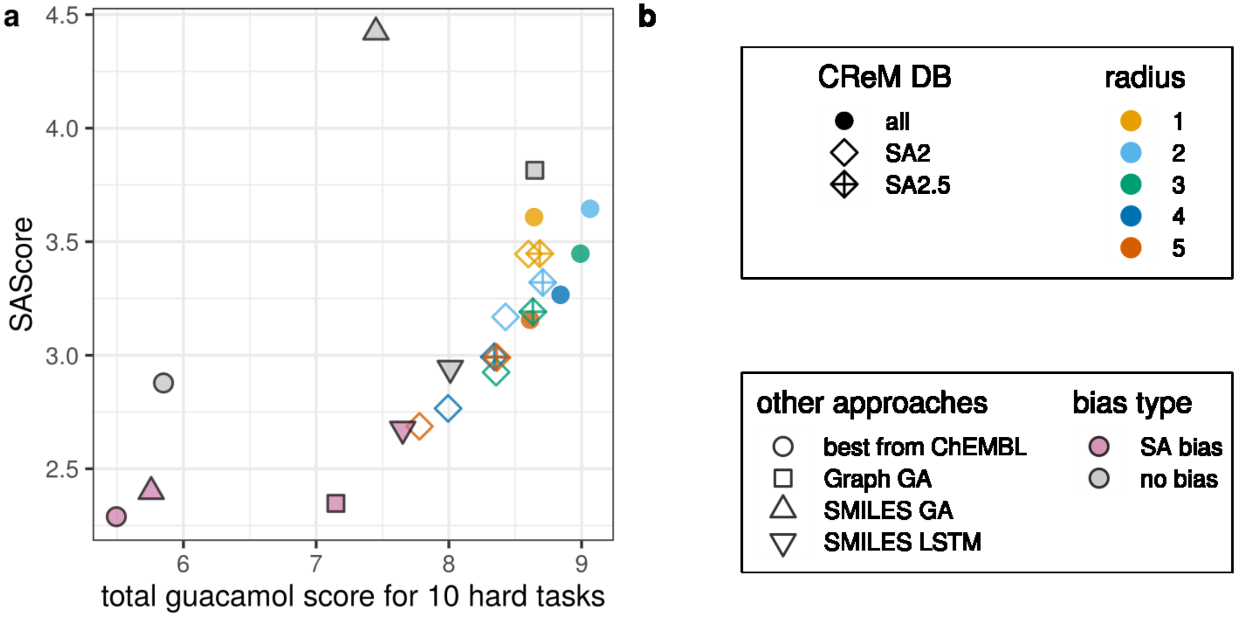
compounds with SA score ≤ 2.5 (572 527)

compounds with SA score ≤ 2 (107 806)

Context radius

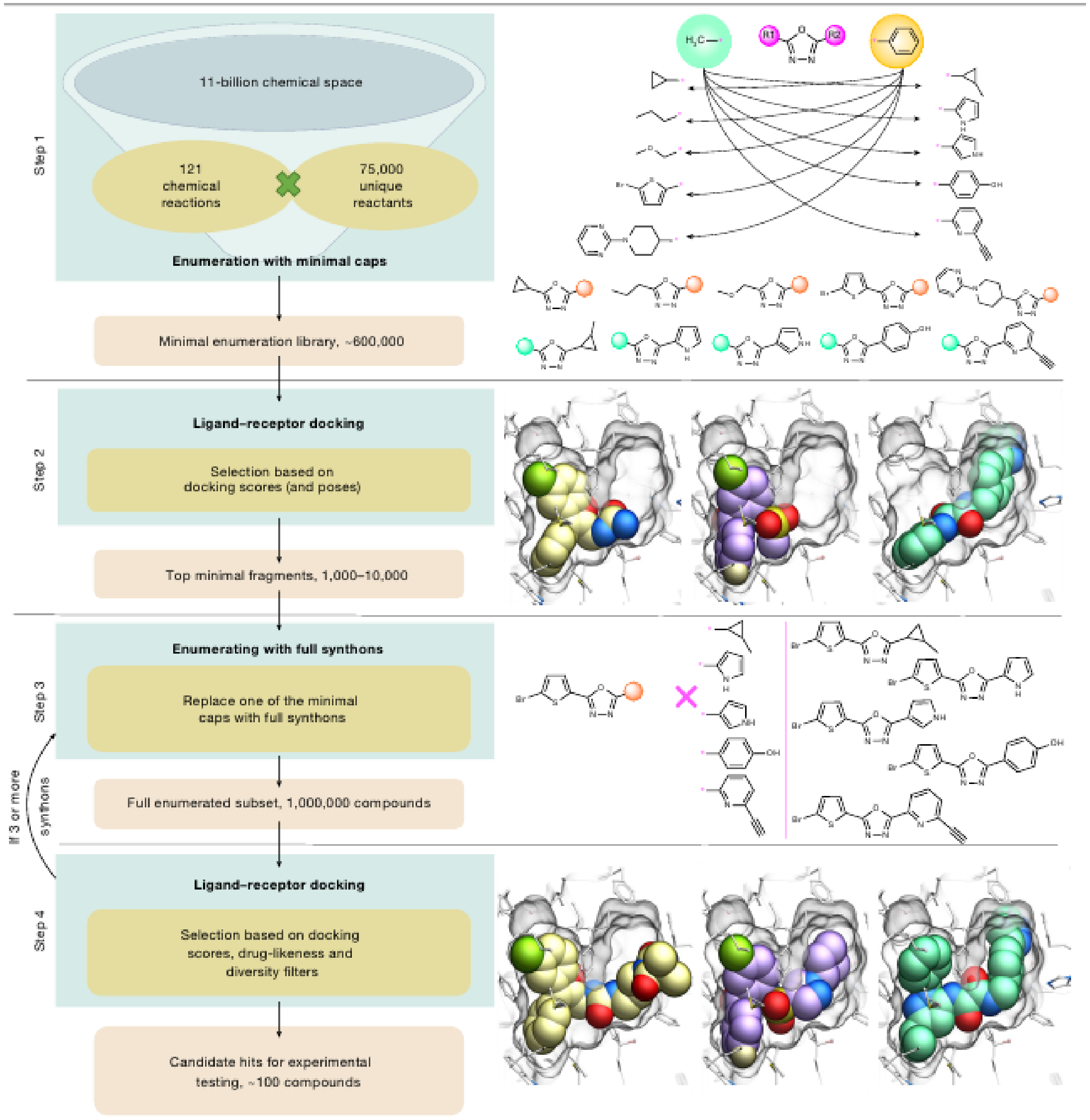
1 2	less conservative replacements
3	
4 5	more conservative replacements





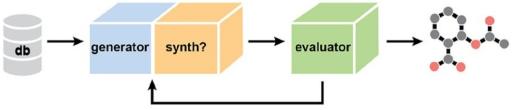
Polishchuk, P., Control of Synthetic Feasibility of Compounds Generated with CReM. *Journal of Chemical Information and Modeling* **2020**, 60, 6074-6080. (10.1021/acs.jcim.0c00792)

V-SYNTHES



Step 1
11-billion chemical space
121 chemical reactions
75,000 unique reactants
Enumeration with minimal caps
Minimal enumeration library, ~600,000
Step 2
Ligand–receptor docking
Selection based on docking scores (and poses)
Top minimal fragments, 1,000–10,000
Step 3
Enumerating with full synthons
Replace one of the minimal caps with full synthons
Full enumerated subset, 1,000,000 compounds
If 3 or more synthons
Step 4
Ligand–receptor docking
Selection based on docking scores, drug-likeness and diversity filters
Candidate hits for experimental testing, ~100 compounds
R1
R2

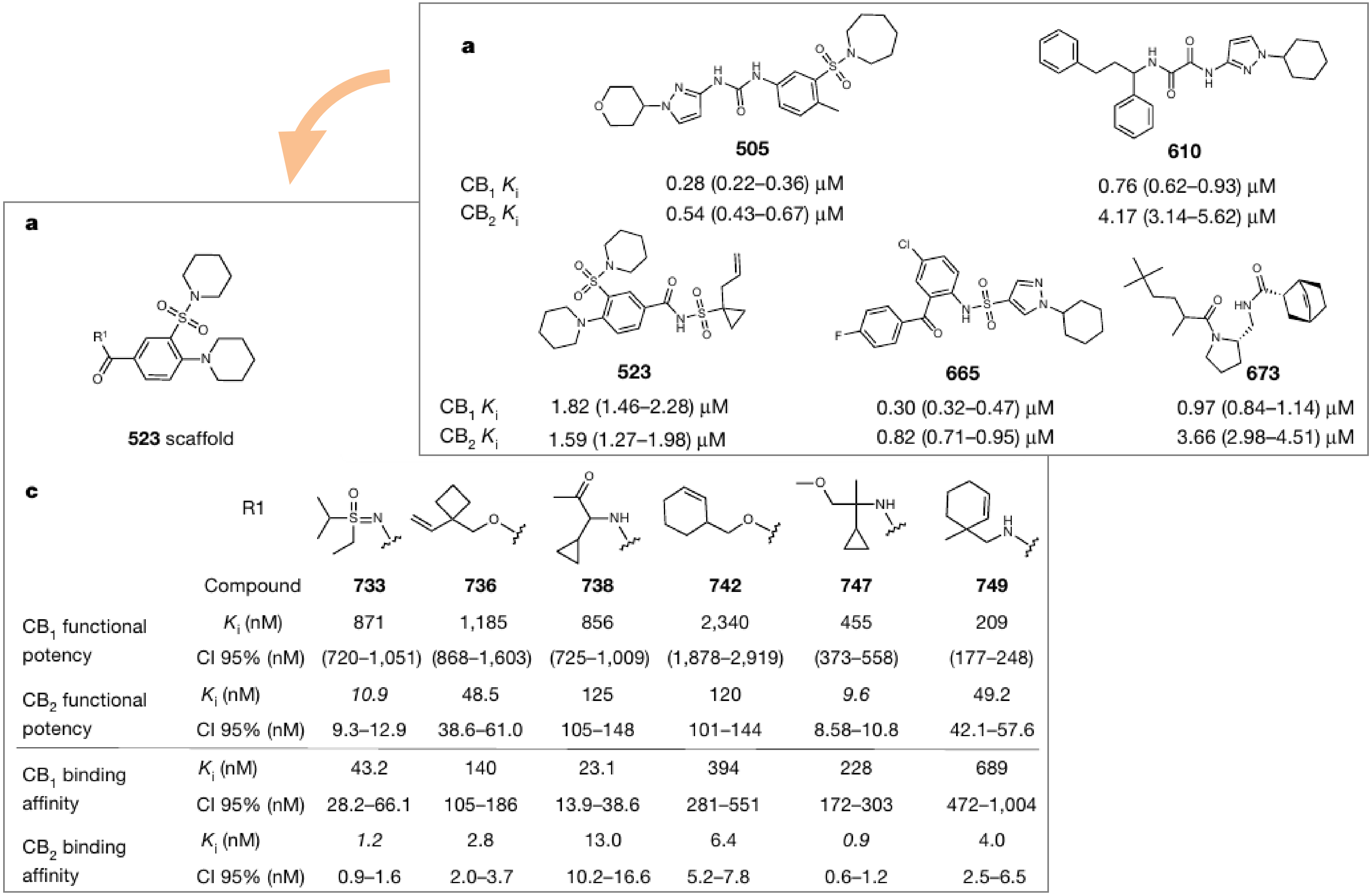
Fig. 1 | V-SYNTHES approach to modular screening of Enamine REAL Space. A general overview of the four-step algorithm (left) and examples for each step (right). Asterisks in step one show the attachment points of synthons; arrows show possible pairing of minimal synthons with real synthons.



(g) Explicit constraints

Sadybekov, A. A. et al, Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **2021**. (10.1038/s41586-021-04220-9)

V-SYNTHES



505

610

	505	610
CB$_1$ K_i	0.28 (0.22–0.36) μM	0.76 (0.62–0.93) μM
CB$_2$ K_i	0.54 (0.43–0.67) μM	4.17 (3.14–5.62) μM

523

665

673

	523	665	673
CB$_1$ K_i	1.82 (1.46–2.28) μM	0.30 (0.32–0.47) μM	0.97 (0.84–1.14) μM
CB$_2$ K_i	1.59 (1.27–1.98) μM	0.82 (0.71–0.95) μM	3.66 (2.98–4.51) μM

523 scaffold

	Compound	**733**	**736**	**738**	**742**	**747**	**749**
CB$_1$ functional potency	K_i (nM)	871	1,185	856	2,340	455	209
	CI 95% (nM)	(720–1,051)	(868–1,603)	(725–1,009)	(1,878–2,919)	(373–558)	(177–248)
CB$_2$ functional potency	K_i (nM)	*10.9*	48.5	125	120	*9.6*	49.2
	CI 95% (nM)	9.3–12.9	38.6–61.0	105–148	101–144	8.58–10.8	42.1–57.6
CB$_1$ binding affinity	K_i (nM)	43.2	140	23.1	394	228	689
	CI 95% (nM)	28.2–66.1	105–186	13.9–38.6	281–551	172–303	472–1,004
CB$_2$ binding affinity	K_i (nM)	*1.2*	2.8	13.0	6.4	*0.9*	4.0
	CI 95% (nM)	0.9–1.6	2.0–3.7	10.2–16.6	5.2–7.8	0.6–1.2	2.5–6.5

Sadybekov, A. A. et al, Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **2021**. (10.1038/s41586-021-04220-9)

Take home message

- De novo design can efficiently explore much larger chemical space than virtual screening

- There are multiple approaches to generate chemically valid structures, all of them have their pros and cons

- The main issue of de novo design is synthetic feasibility of generated compounds

- There are several ways how to control synthetic feasibility

Thank you for your attention