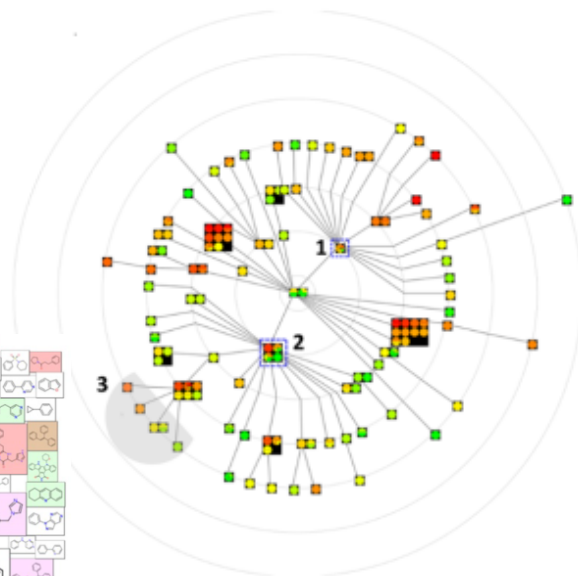
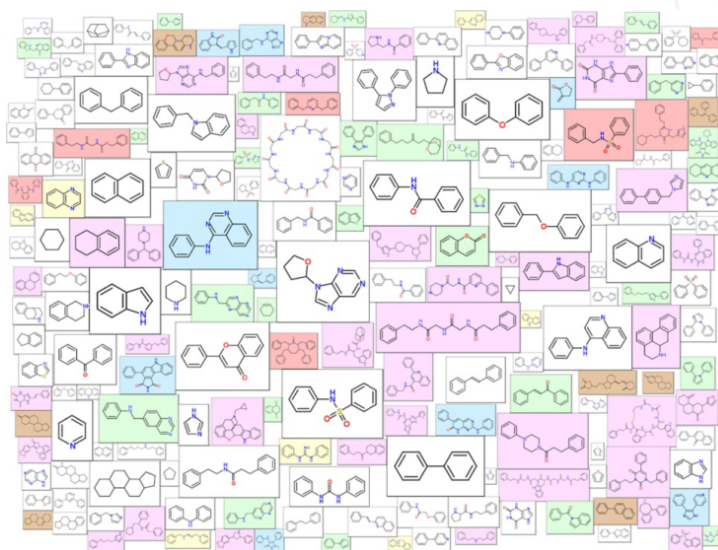
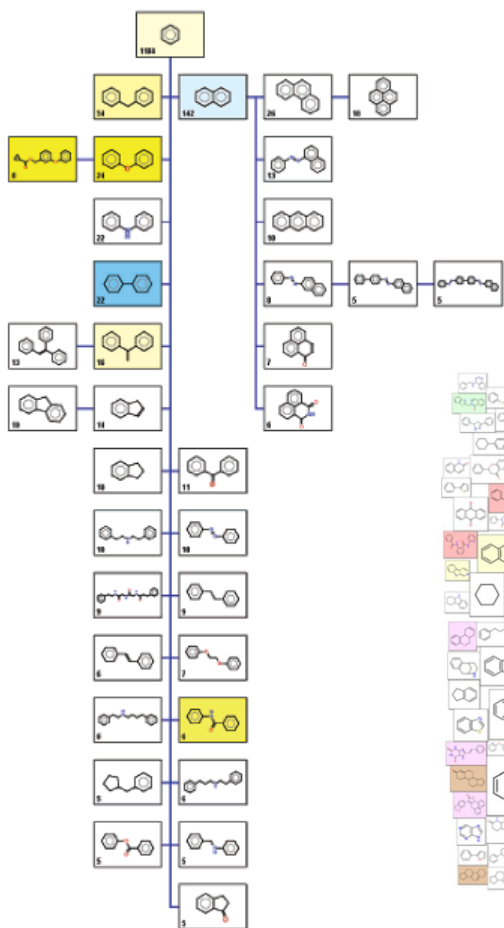
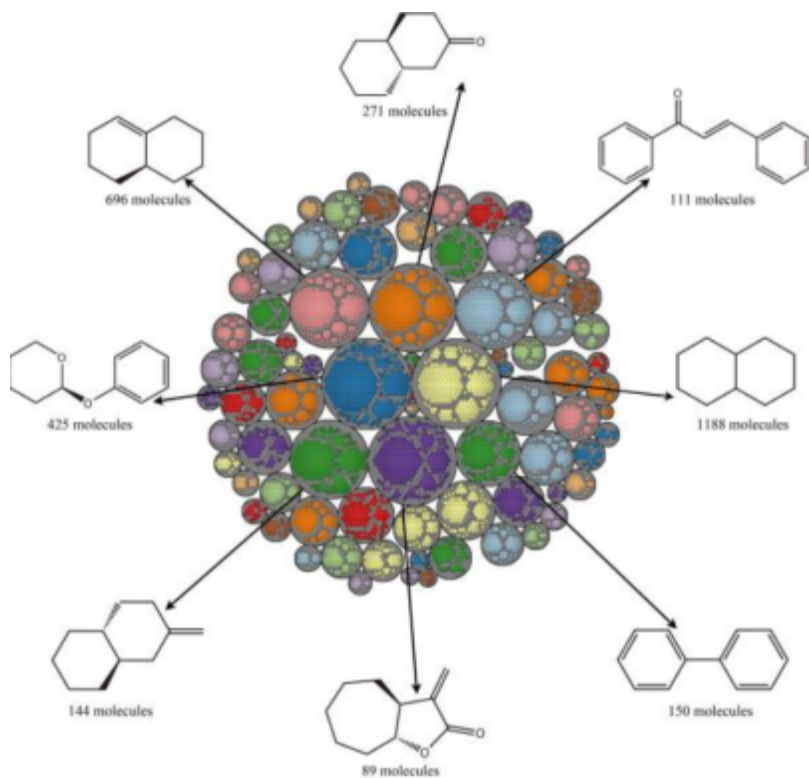


Введение в химическую информатику

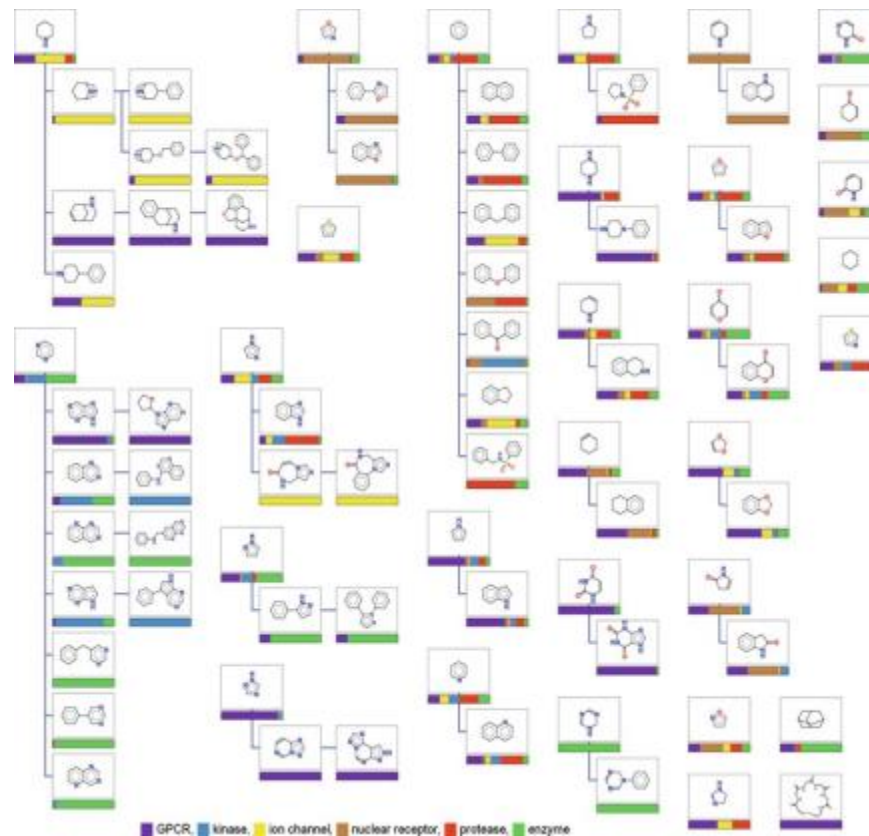
Лекция 9



Хемография (картография): анализ и сопоставление баз данных

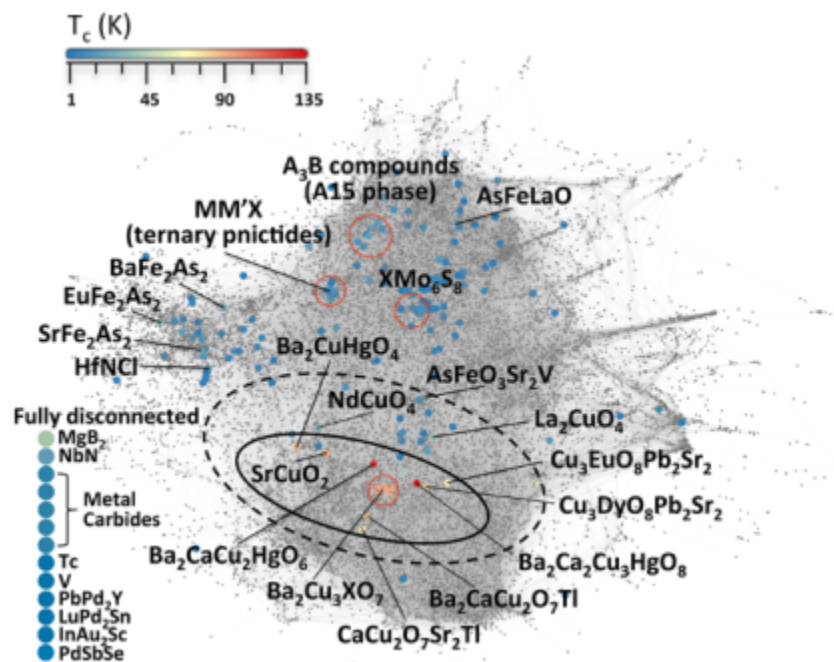
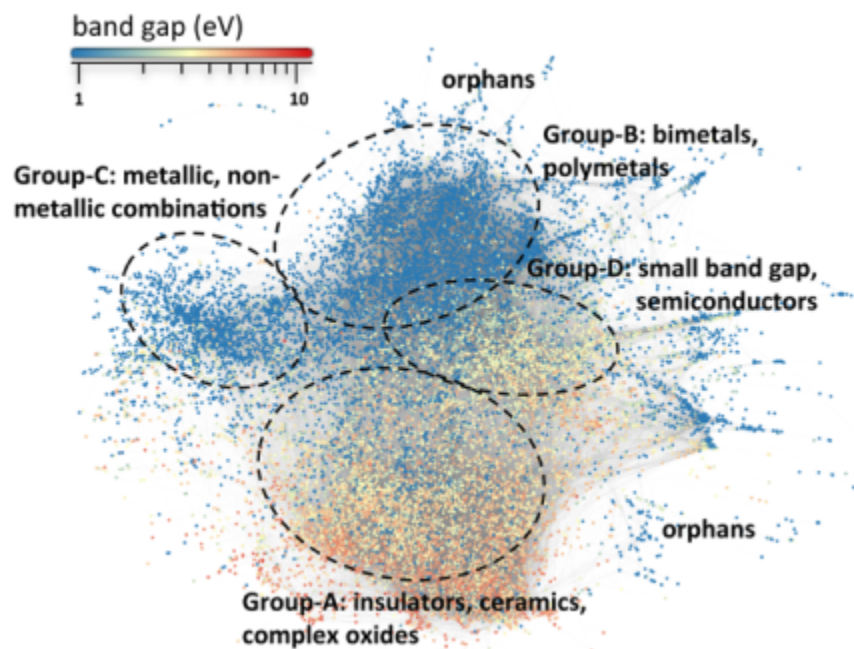


Tean et al J Cheminform. 2013 Jan 21;5(1):5. doi: 10.1186/1758-2946-5-5



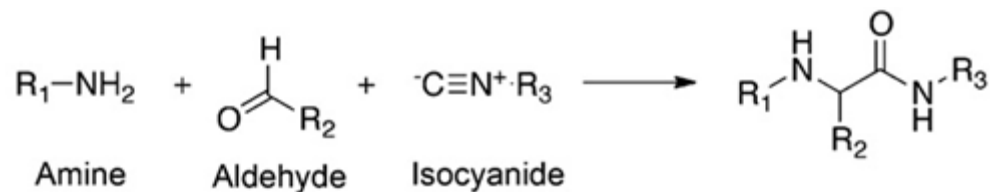
P.Ertl "The Scaffold Tree: An Efficient Navigation in the Scaffold Universe" Methods Mol Biol. 2011;672:245-60. doi: 10.1007/978-1-60761-839-3_10

Хемография (картография): виртуальный скрининг

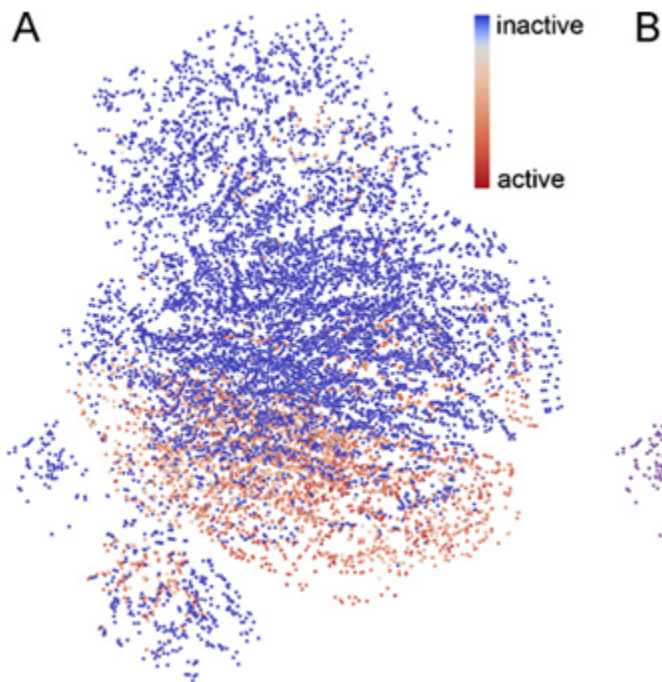


Хемография (картография)

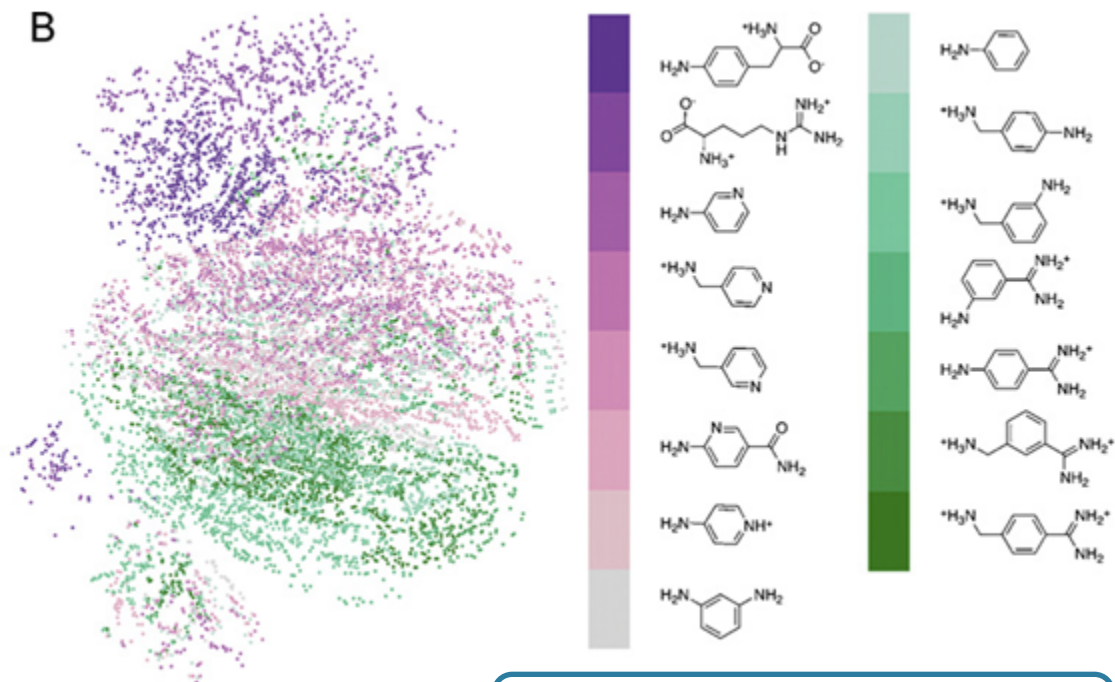
15,840 продуктов (-aminoacyl amide derivatives) трехкомпонентной конденсации Уги (взаимодействие аминов, изоцианидов, карбонильных соединений, органических/неорганических кислот).



Распределение данных по значению биологической активности (ингибирование триптазы)

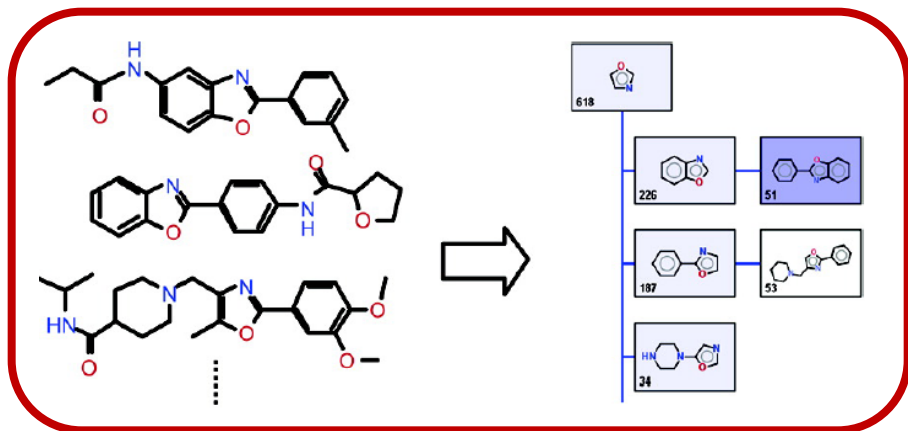


Распределение данных по подструктуре

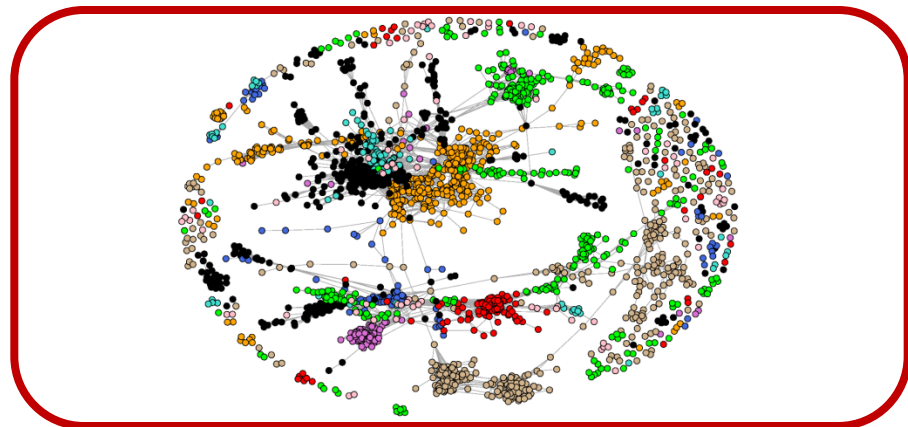


Stochastic Neighbor Embedding (SNE) +
pharmacophore descriptors

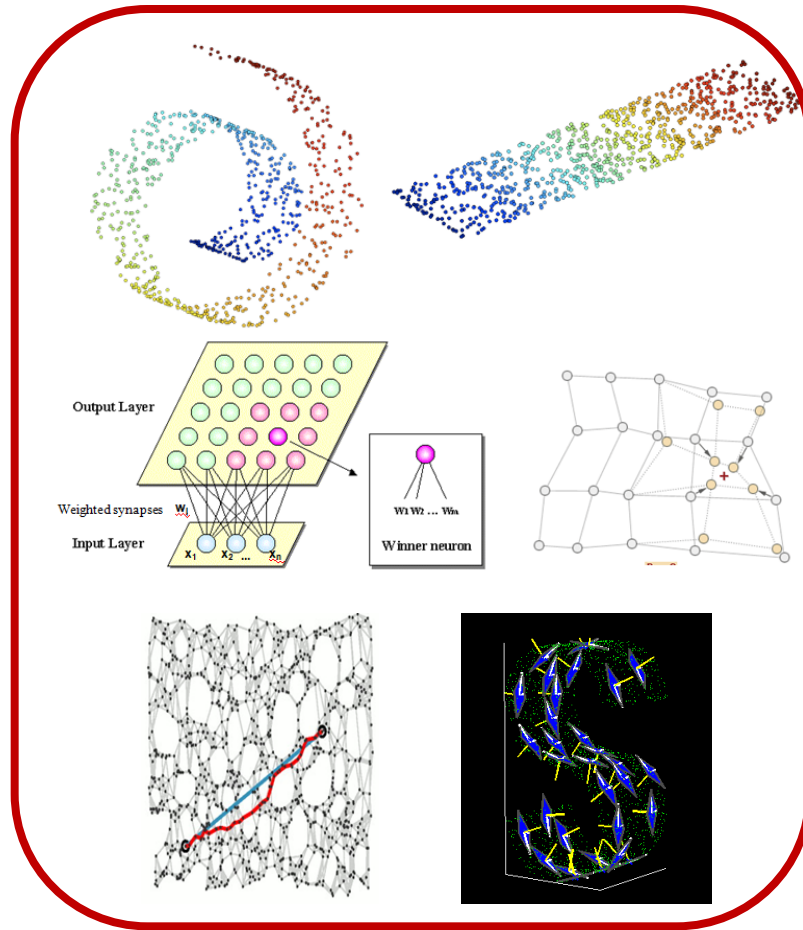
Методы представления химического пространства



Методы, основанные на молекулярных графах



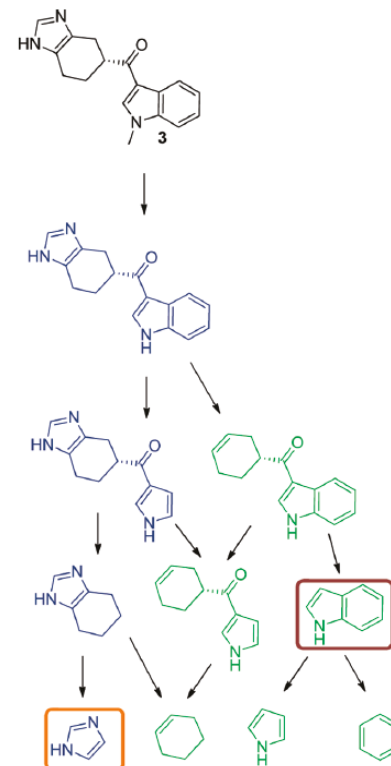
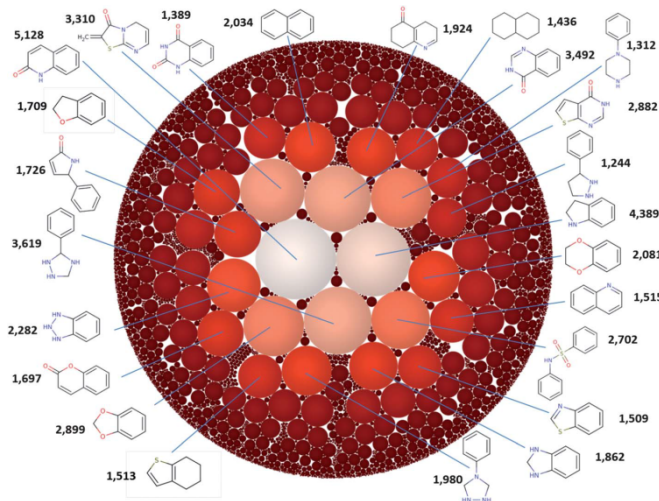
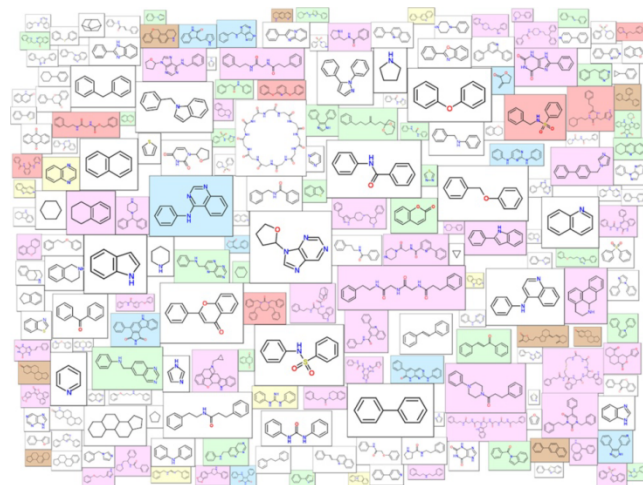
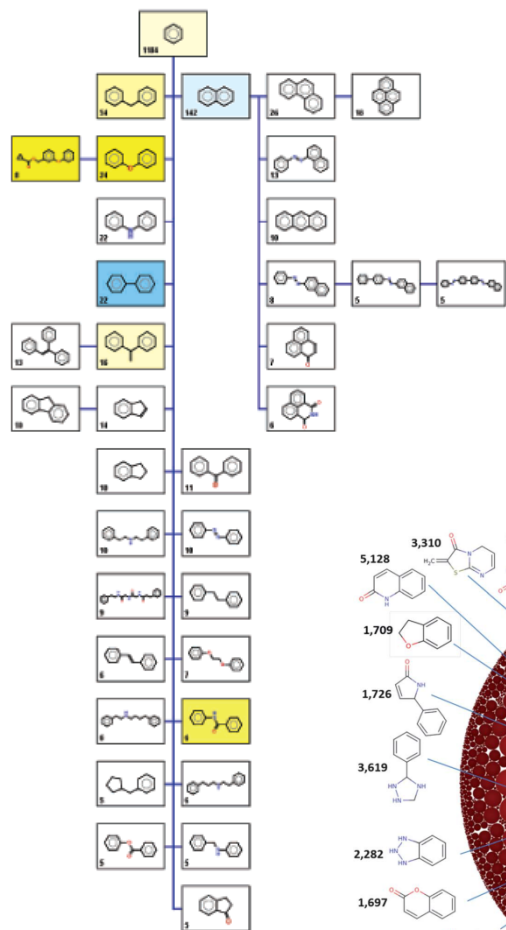
Сетевое представление химического пространства данных



Методы, основанные на дескрипторном представлении данных

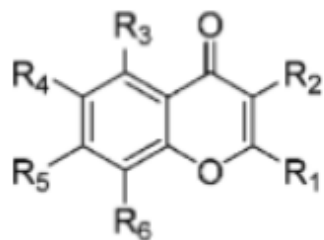
Методы, основанные на молекулярных графах

Навигация в химическом пространстве данных: топологическое пространство молекулярных графов



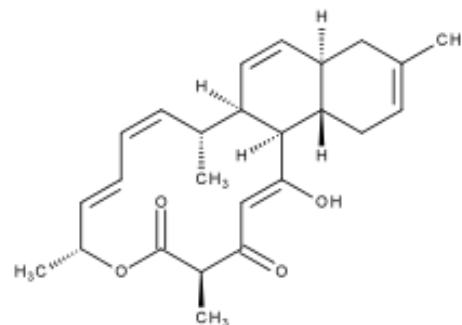
- Рациональная организация структурных данных
- Идентификация новых объектов посредством анализа активных молекулярных скаффолдов (структурных основ), сочетания заместителей и т.д.

Молекулярный скелет



Хромоны

Chem. Rev. 2014, 114, 4960 – 4992

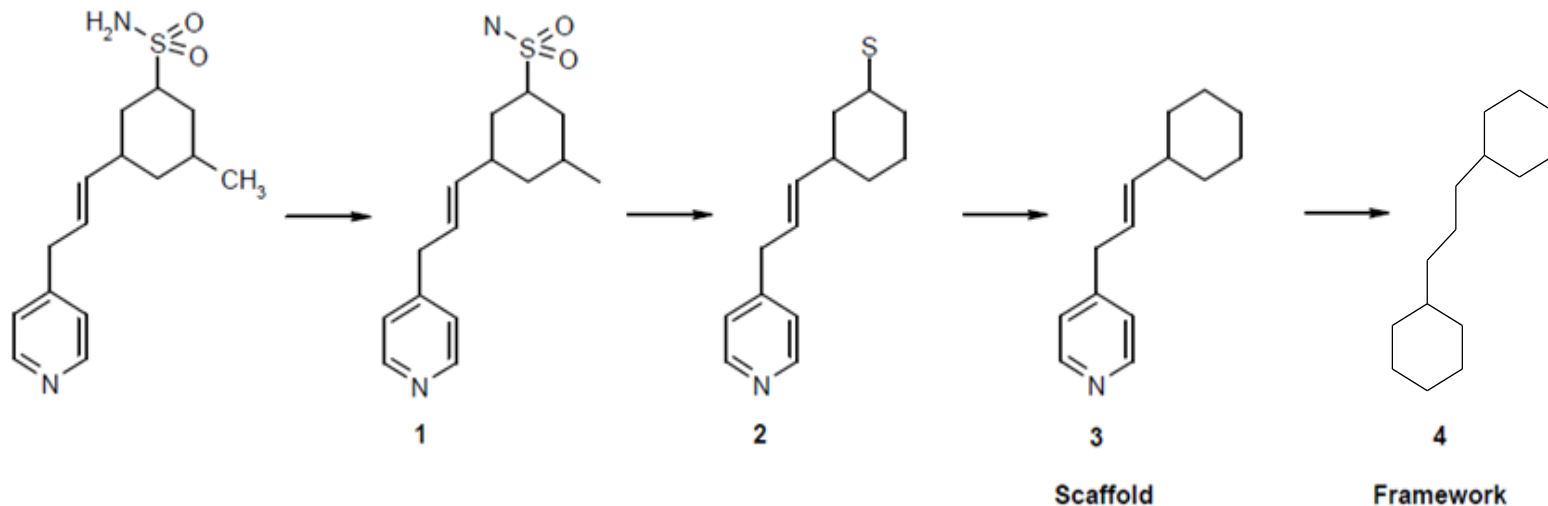


Антрацимицин

Angew. Chem. Int. Ed. 2013, 52, 7822 – 7824

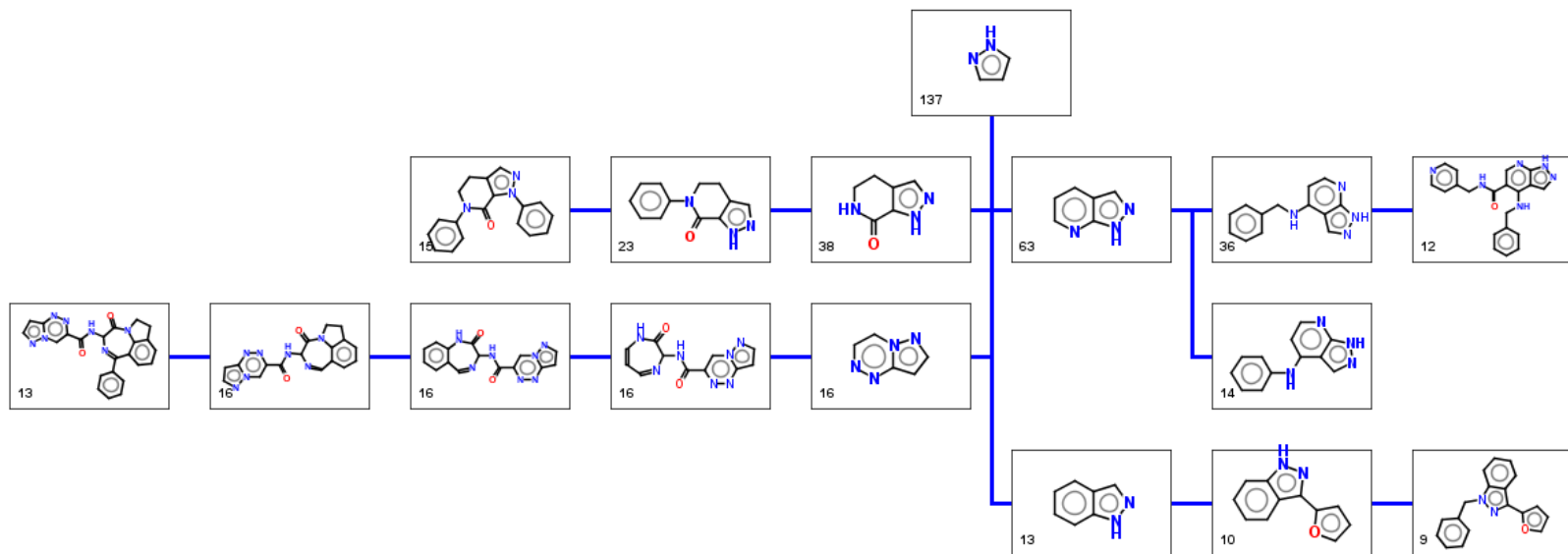
Схема иерархической молекулярной организации: молекулярные скелеты и каркасы

- Термин «молекулярный скелет» в большинстве случаев используется для описания молекулярной основы, задействованной в разработке лекарственных препаратов или детектируемой при виртуальном скрининге (влияет на молекулярные свойства, часто определяет биологическую (физиологическую) активность молекулы)
- Для серии аналогов, молекулярный скелет может быть получен при определении максимальной общей подструктуры (Maximum Common Substructure), либо одним из множества альтернативных подходов
- Присоединение к молекулярному скелету цикла интерпретируется как образование нового скелета
- В ряде случаев используются молекулярные каркасы



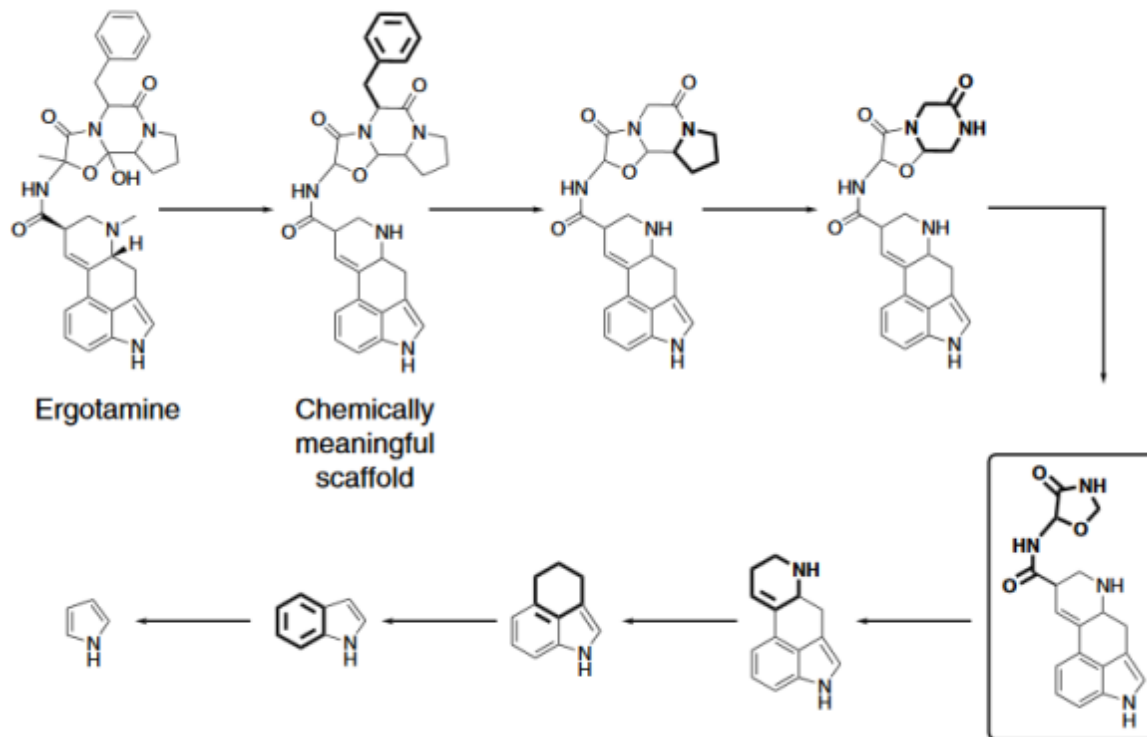
Деревья молекулярных остовов: иерархическая классификация для визуализации химического пространства молекулярных скаффолдов

Метод классификации соединений в соответствии с их молекулярным скелетом.



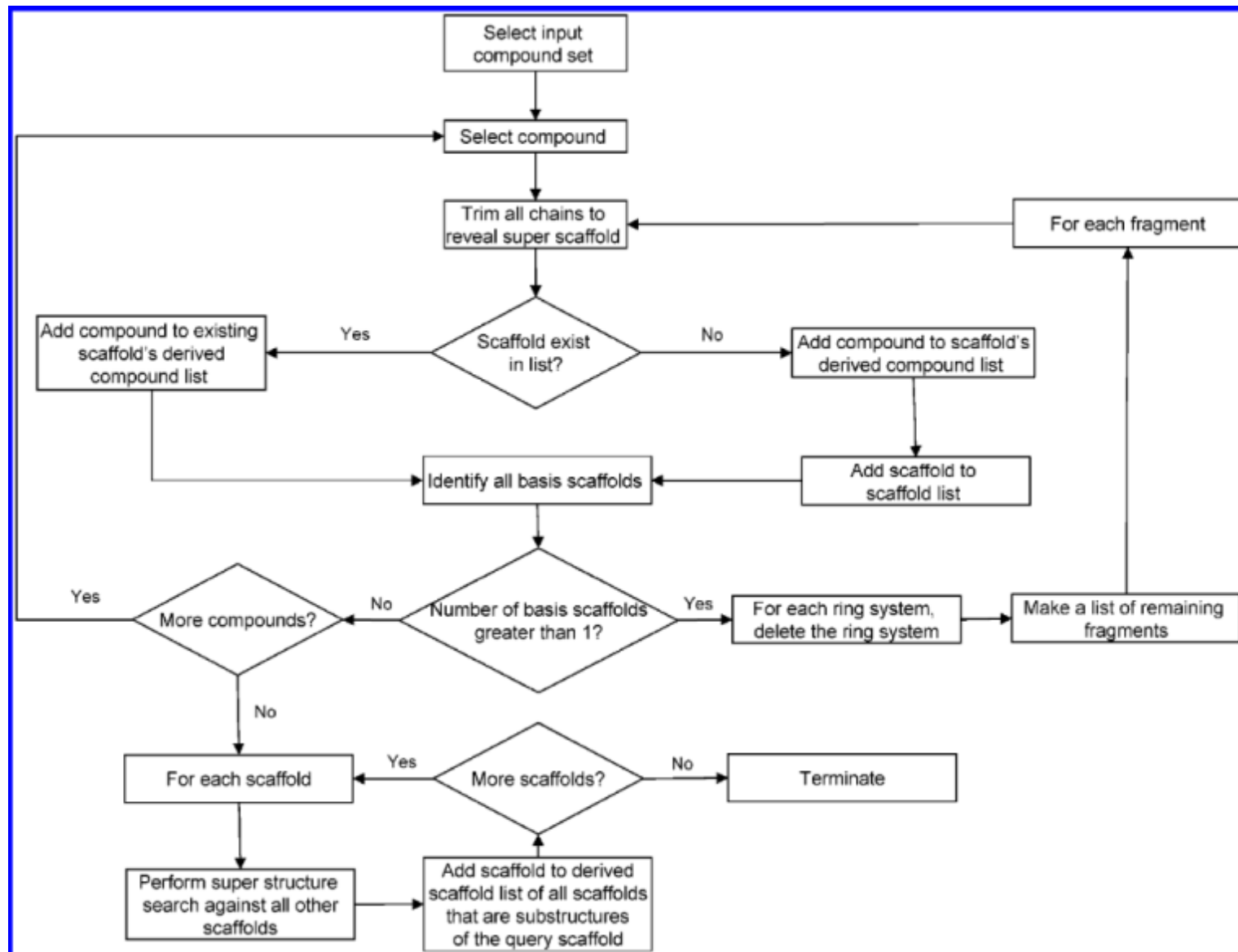
The Scaffold Tree – Visualization of the Scaffold Universe by Hierarchical Scaffold Classification. A. Schuffenhauer, P. Ertl, S. Roggo, S. Wetzel, M. Koch, H. Waldmann, *J. Chem. Inf. Model.* 47, 47, 2007

Деревья молекулярных скелетов: иерархическая классификация для визуализации химического пространства молекулярных скаффолдов

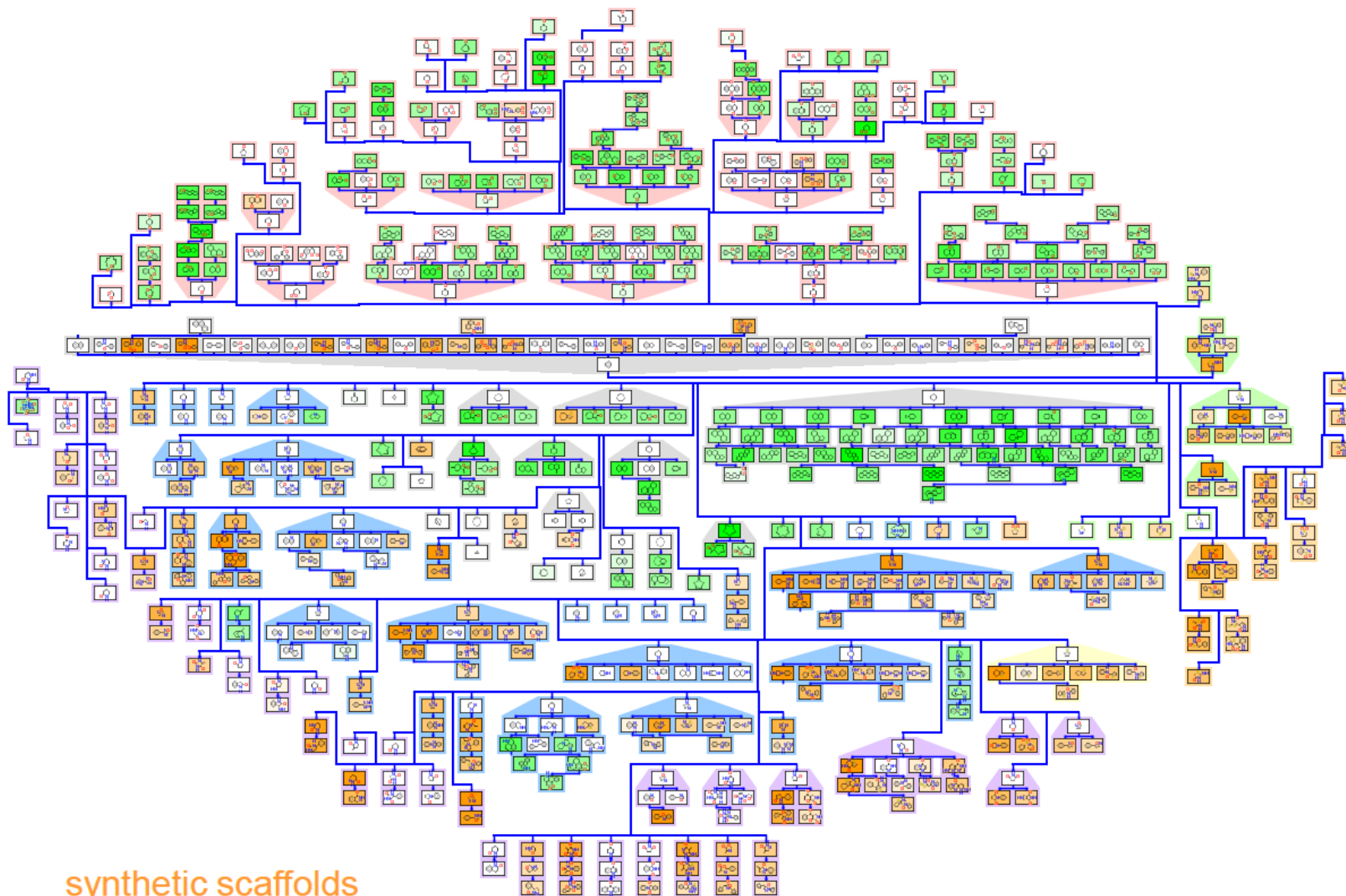


- Загрузка данных, извлечение молекулярных скелетов (карбоциклические соединения и гетероциклы, их алифатические связи, атомы, присоединенные посредством двойной связи)
- Итеративное отделение циклов для генерации «уменьшенных» скелетов в соответствии с рядом правил (The Scaffold Tree – Visualization of the Scaffold Universe by Hierarchical Scaffold Classification. A. Schuffenhauer, P. Ertl, S. Roggo, S. Wetzel, M. Koch, H. Waldmann, *J. Chem. Inf. Model.* 47, 47, 2007)
- Иерархическая организация полученных структур в единое дерево

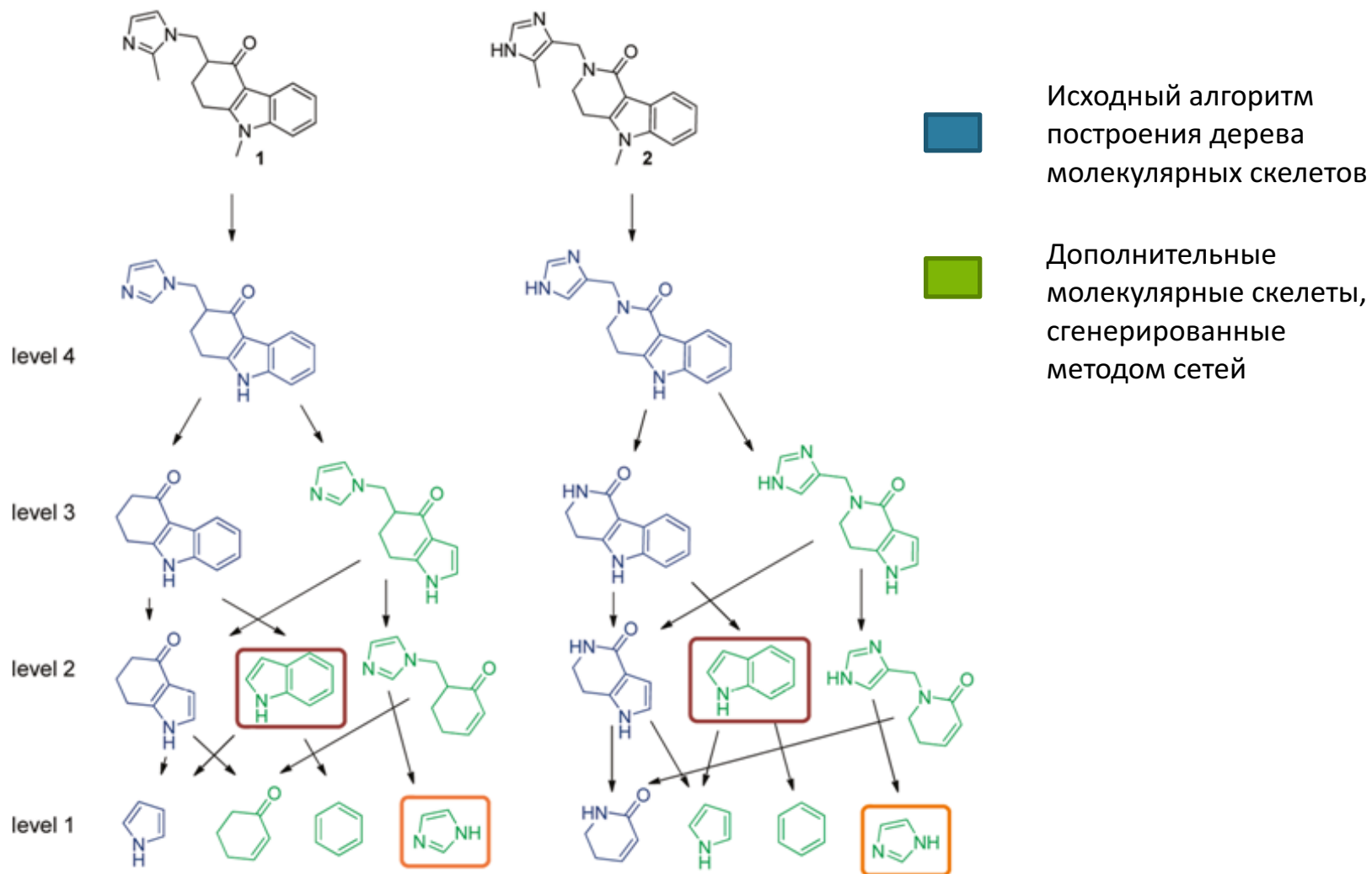
Деревья молекулярных скелетов: Hierarchical Scaffold Clustering (HierS algorithm)



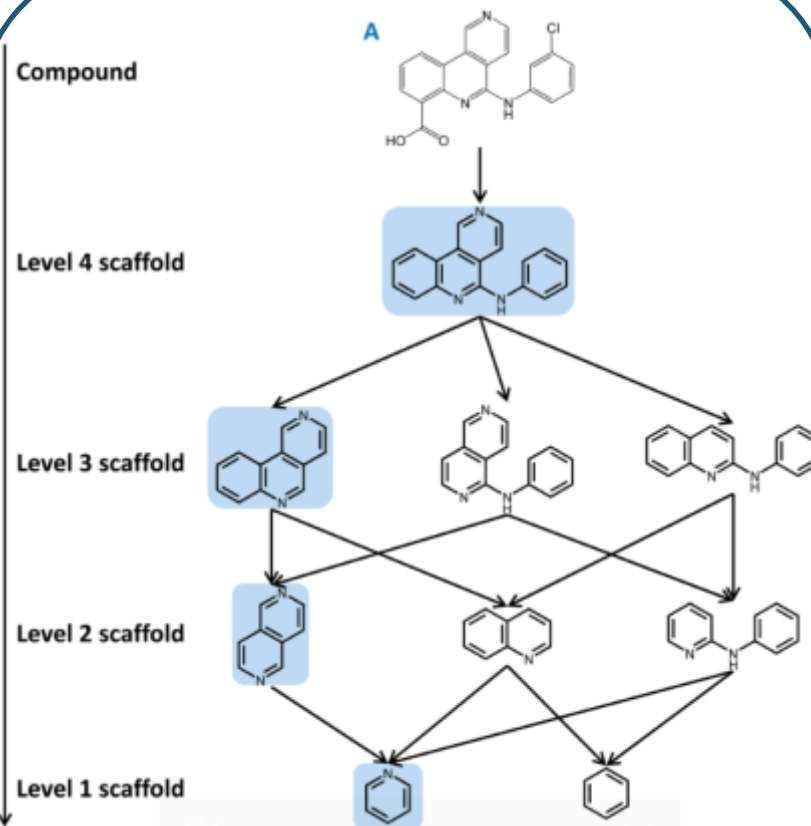
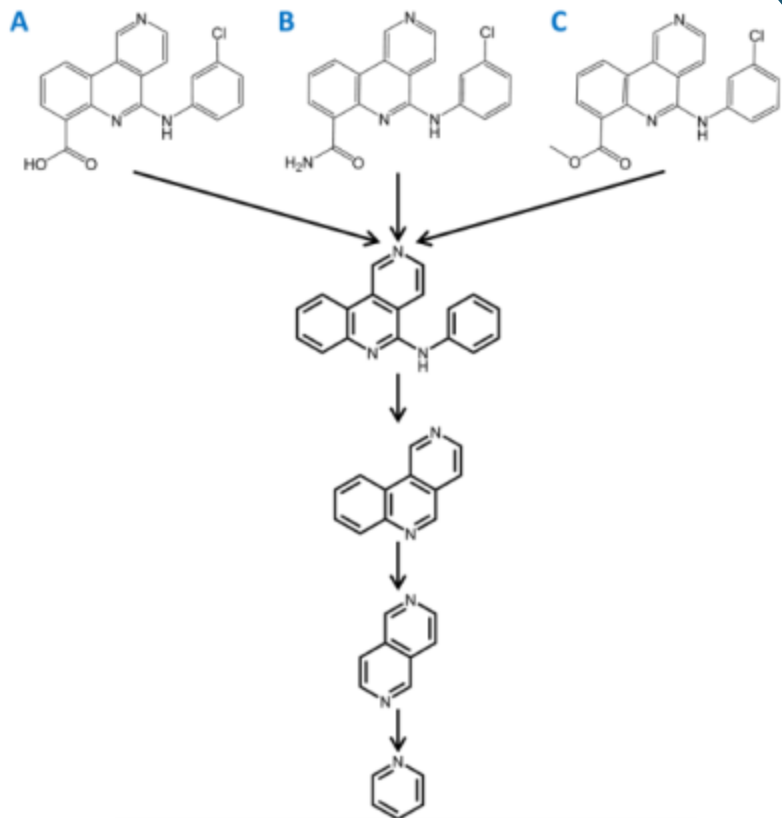
Деревья молекулярных скелетов: вещества природного происхождения



Сети молекулярных скелетов

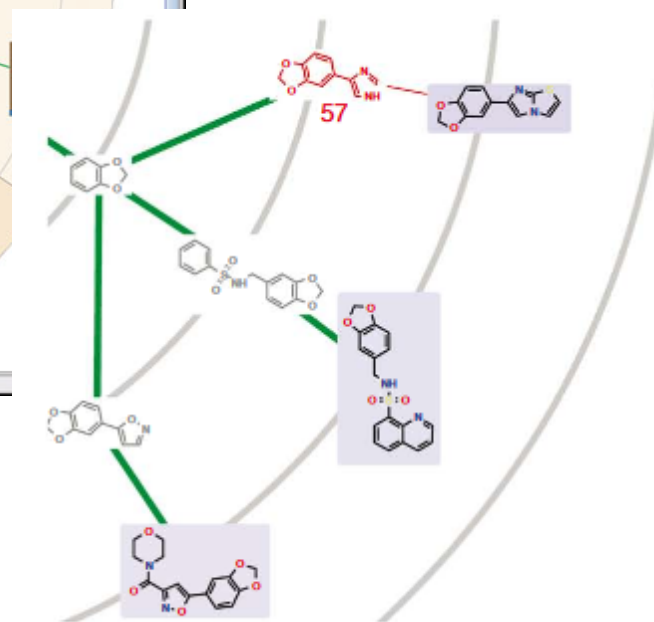
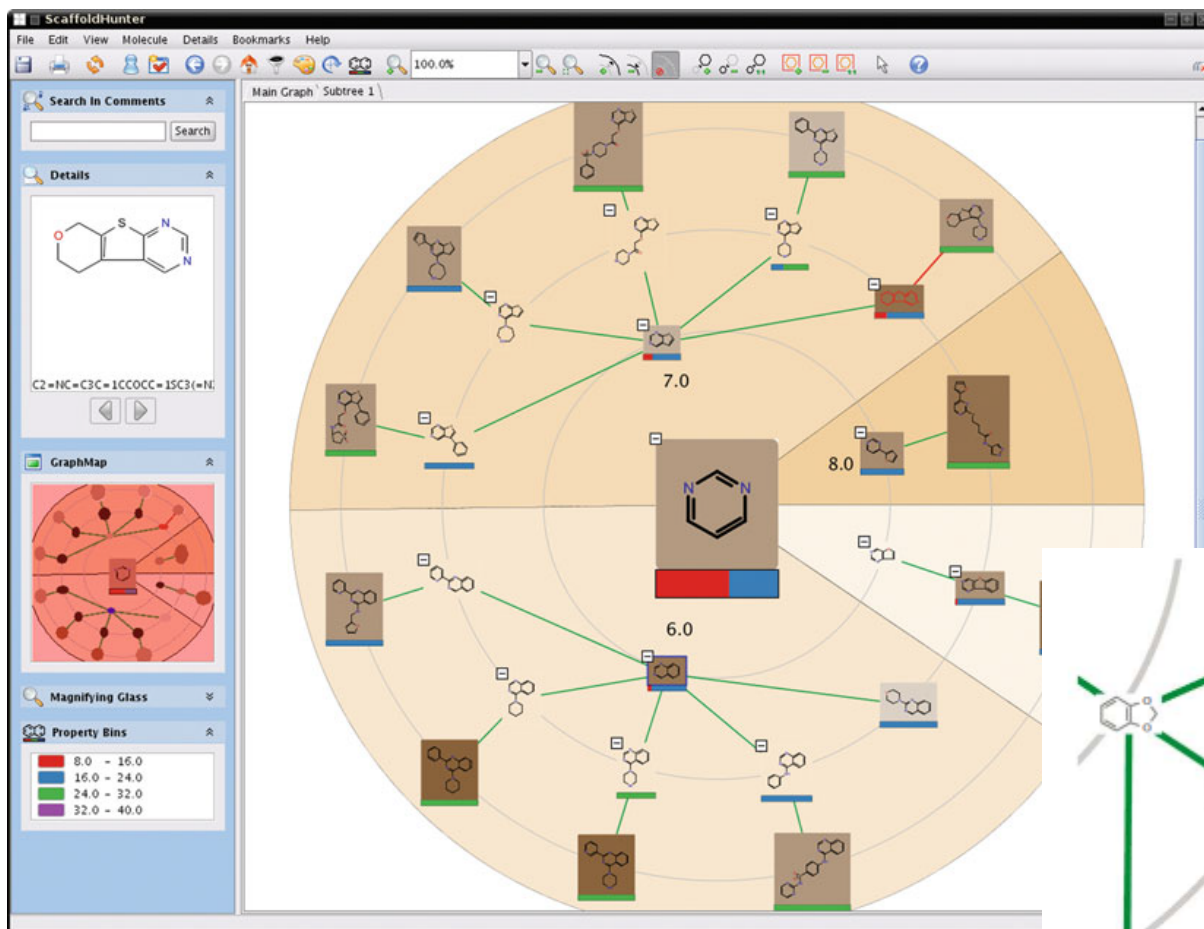


Деревья vs. сети молекулярных скелетов



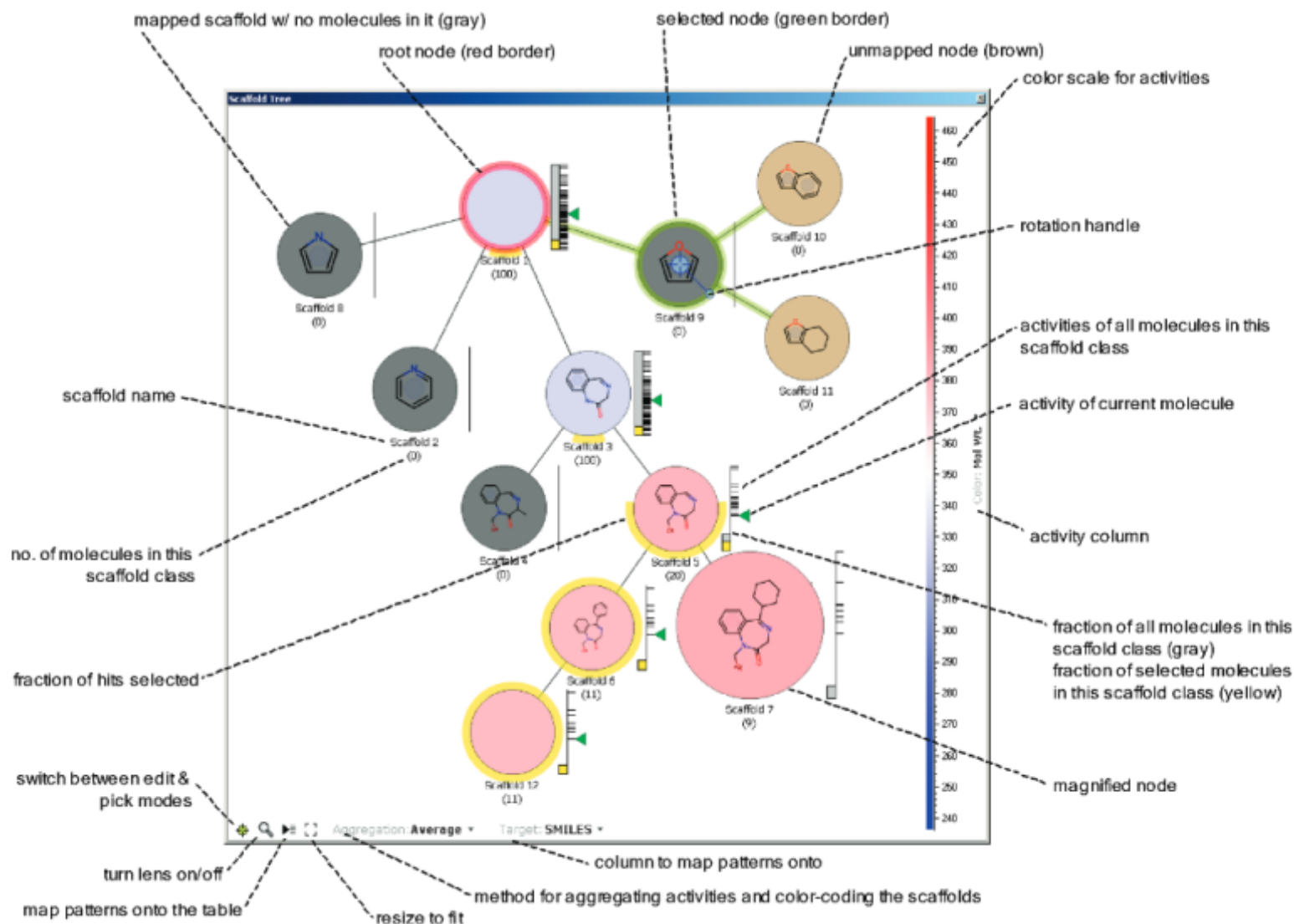
Деревья молекулярных скелетов: программное обеспечение

Scaffold Hunter



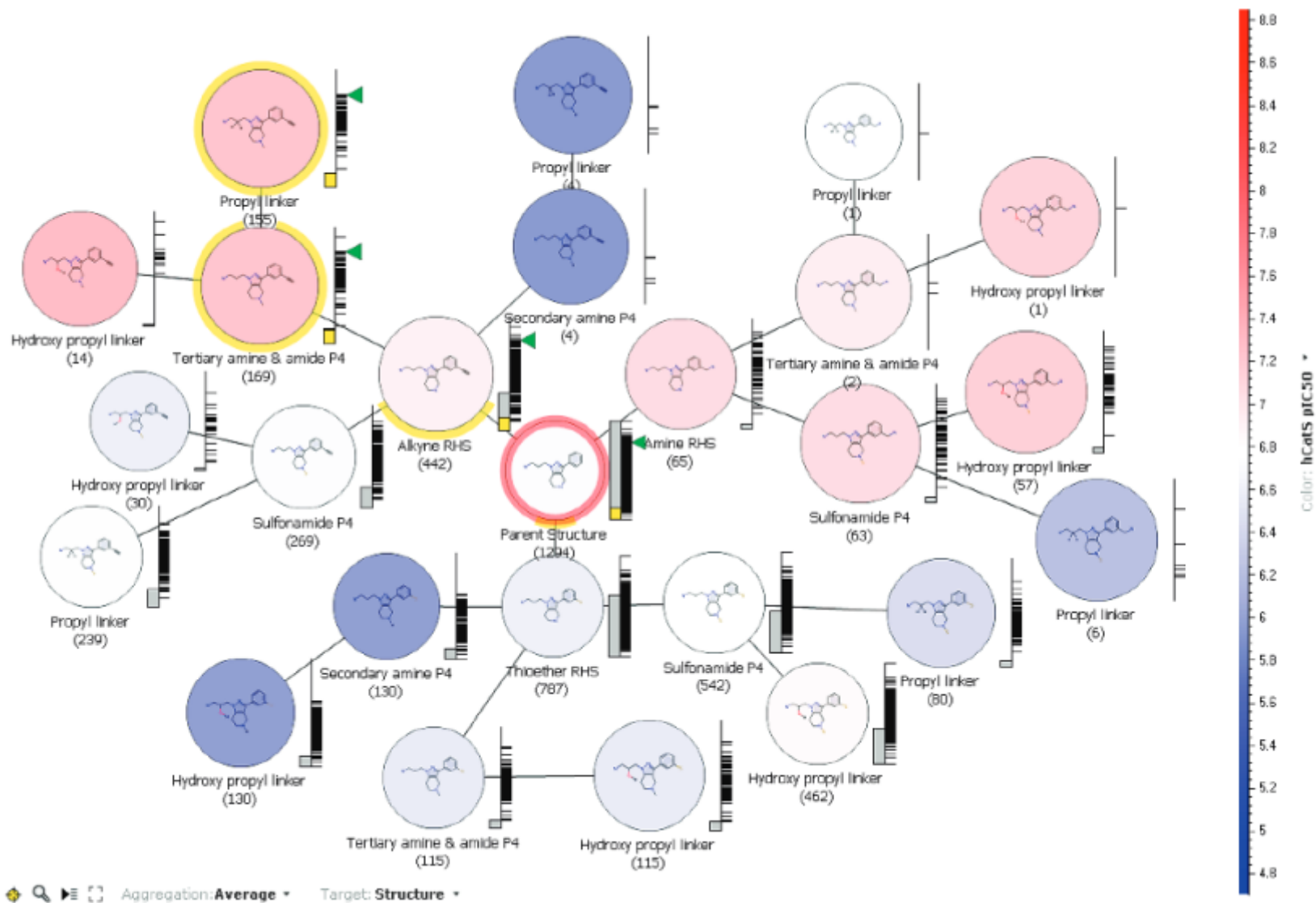
Деревья молекулярных скелетов: программное обеспечение

Scaffold Explorer



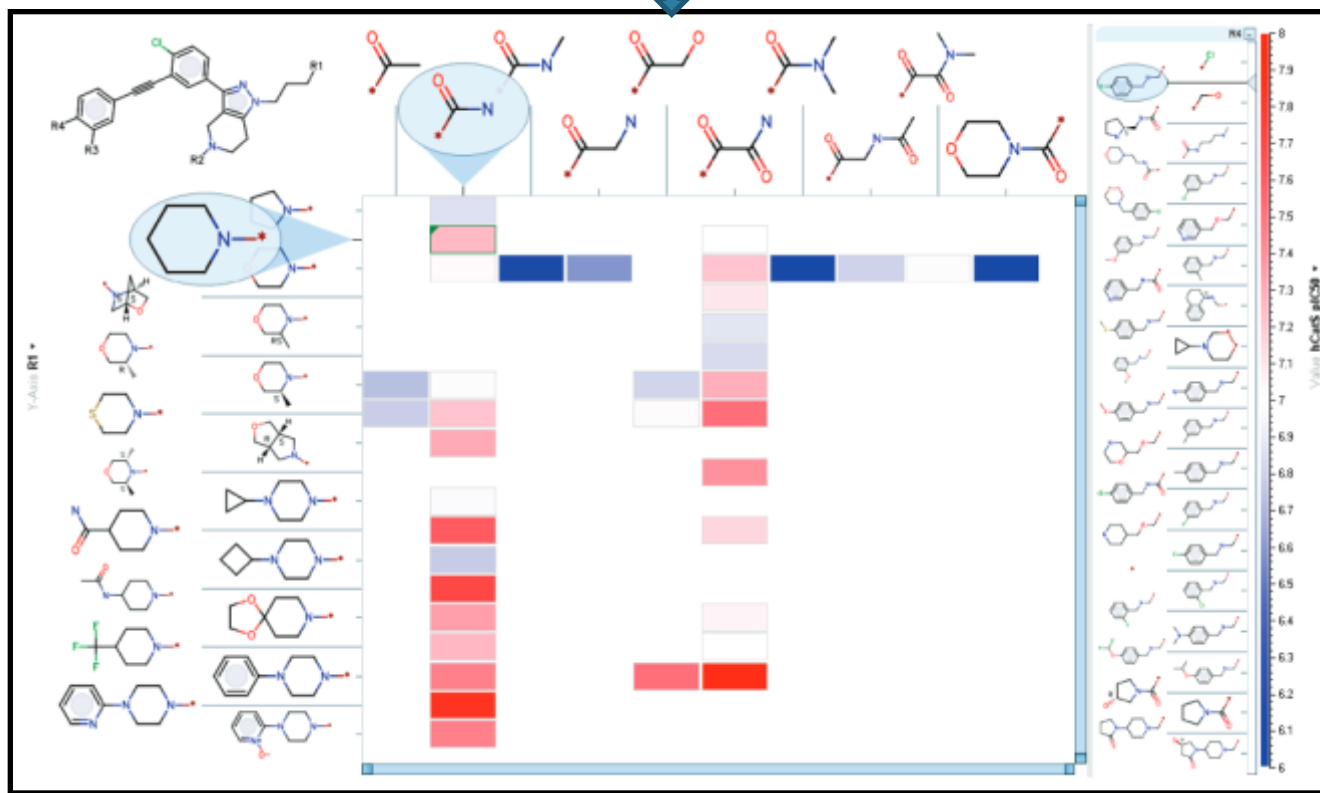
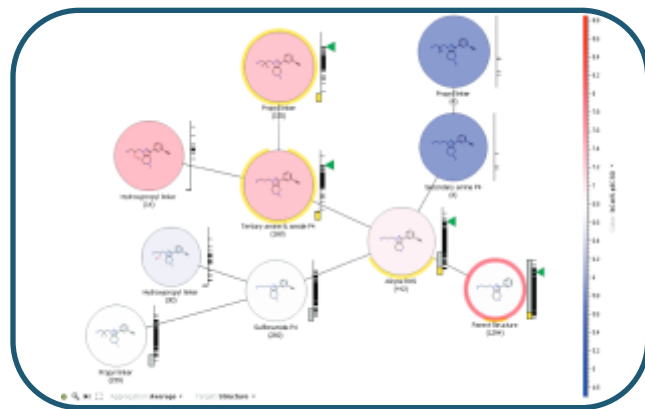
Деревья молекулярных скелетов: программное обеспечение

Scaffold Explorer



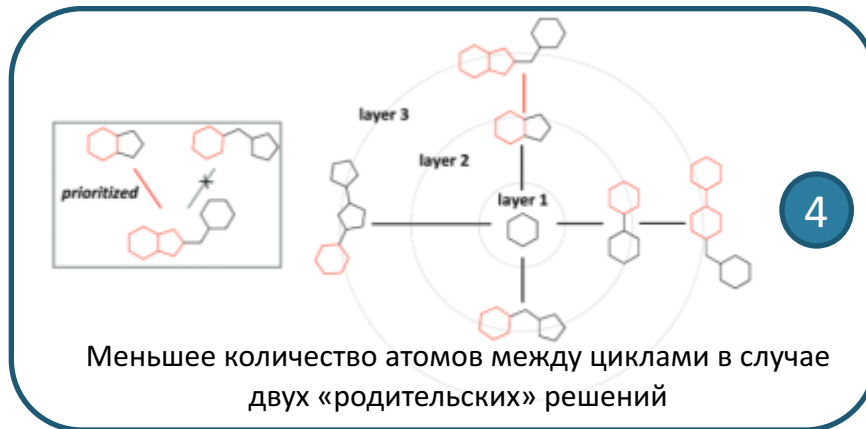
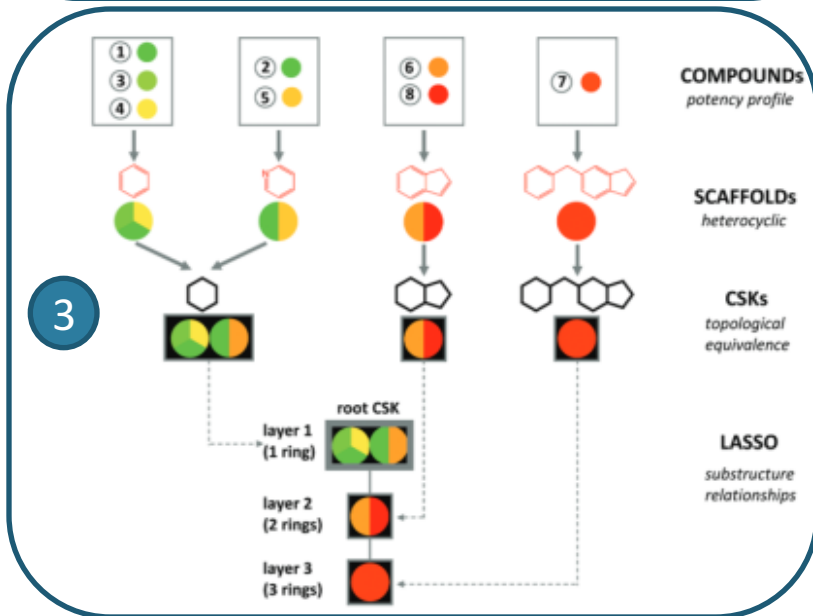
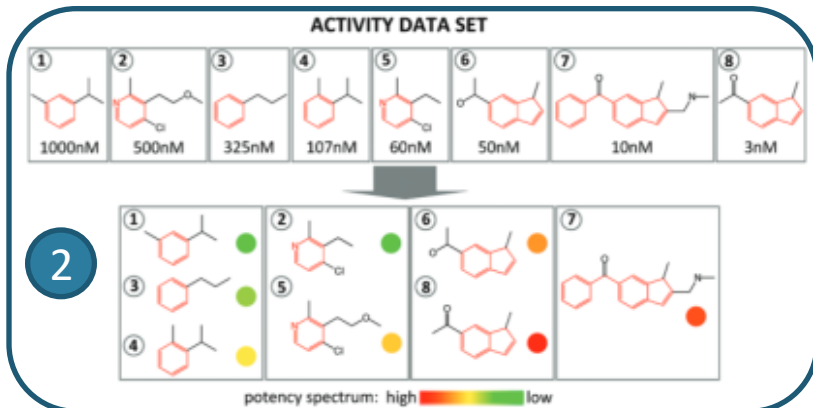
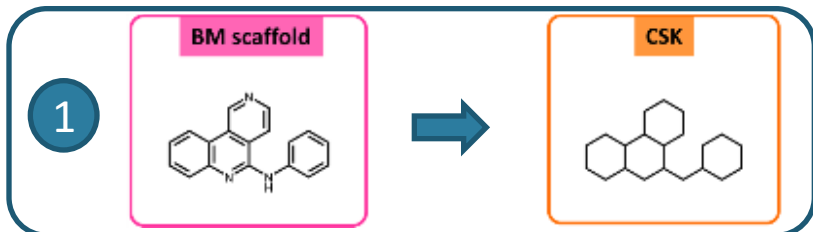
Деревья молекулярных скелетов: программное обеспечение

Scaffold Explorer



LASSO (layered skeleton-scaffold organization) graph

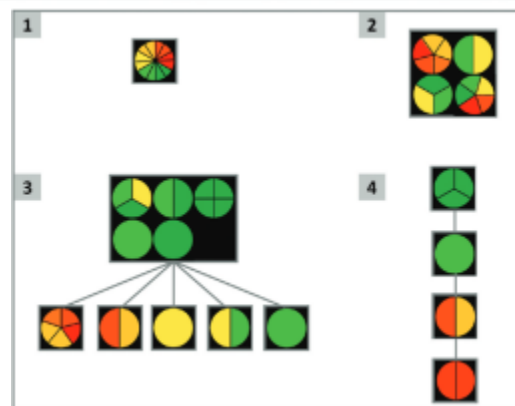
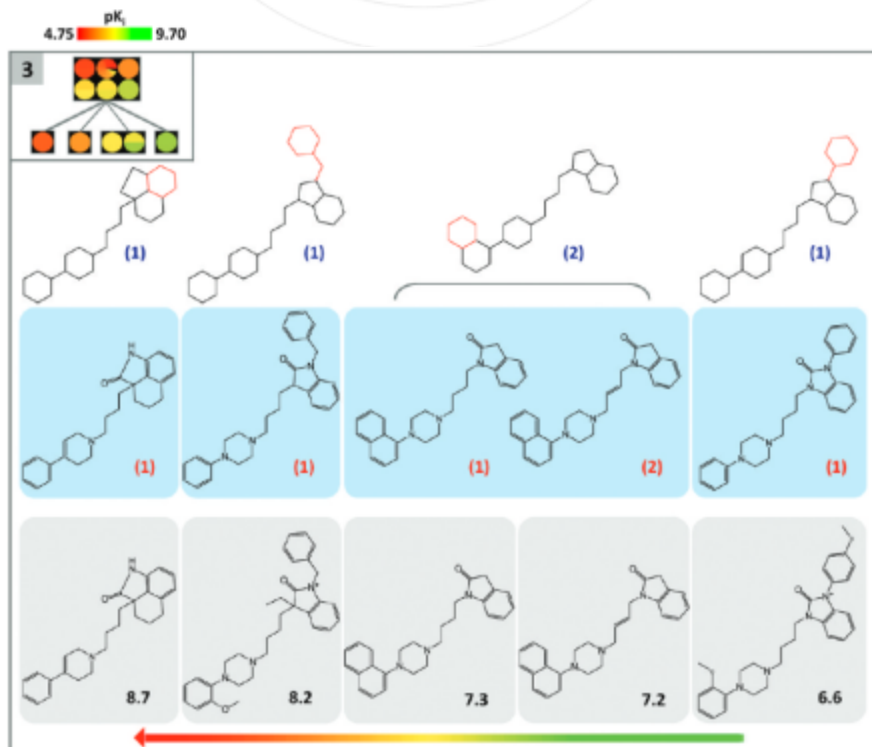
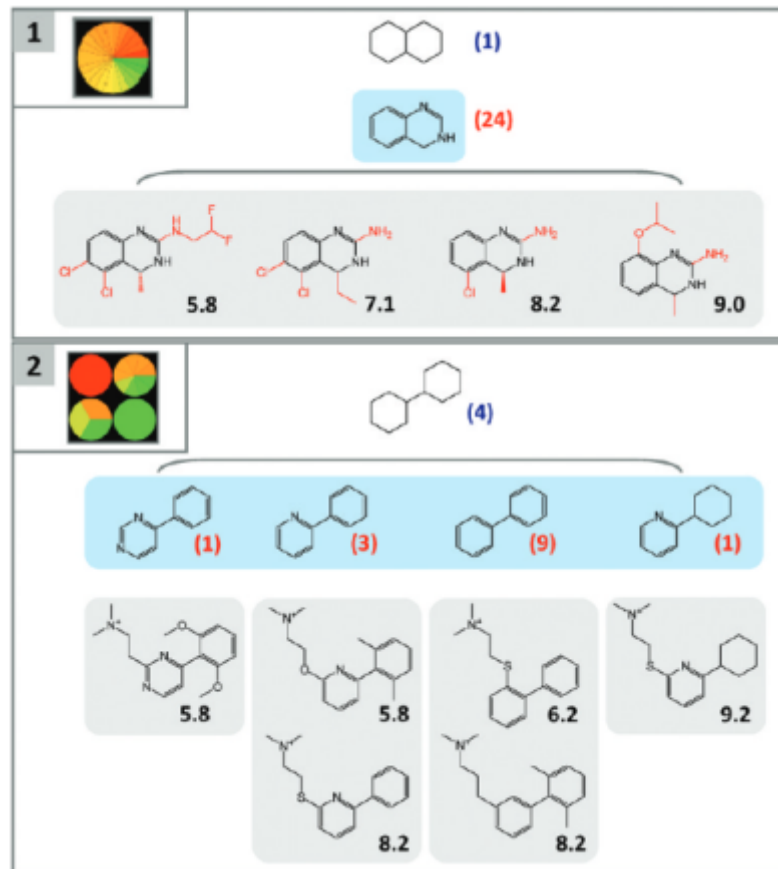
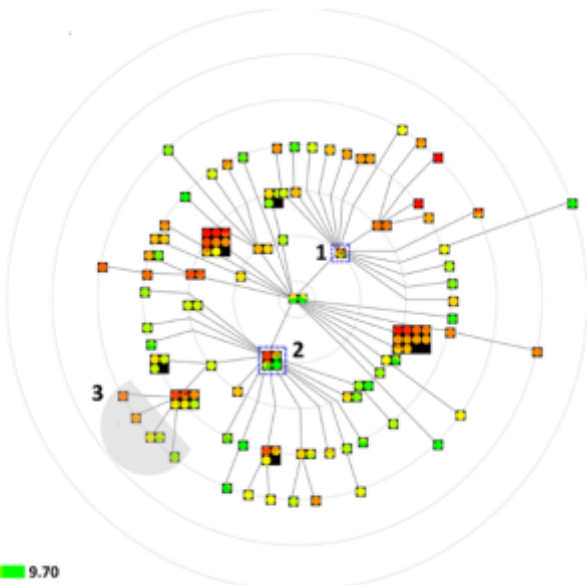
skeleton-scaffold-compound hierarchy



- ❖ Генерация скаффолдов и ассоциация со значением активности
- ❖ Преобразование скаффолдов к более упрощенной модели циклические скелеты *cyclic skeletons* (CSK)
- ❖ CSKs организуются в иерархию по количеству циклов
- ❖ Каждый уровень иерархии представлен концентрическим кругом, на котором располагаются все циклические скелеты, относящиеся к этому уровню «сложности»
- ❖ Отдельно представлена информация по значению свойства

LASSO (layered skeleton-scaffold organization) graph

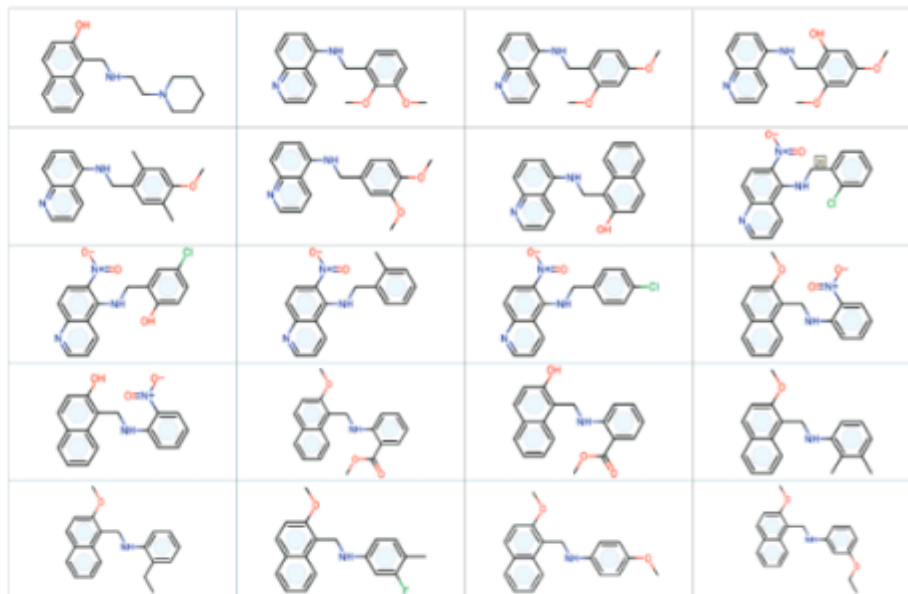
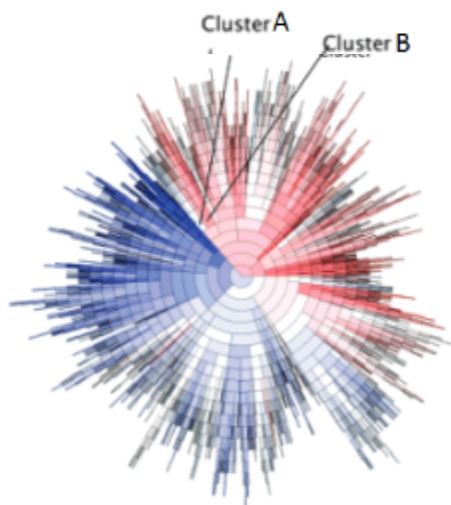
skeleton-scaffold-compound hierarchy



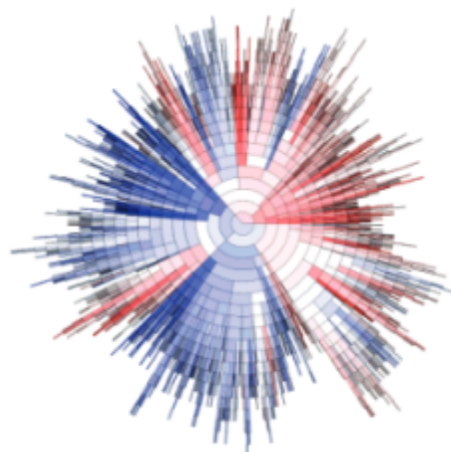
Радиальные кластерограммы

J. Chem. Inf. Model., Vol. 47, No. 1, 2007

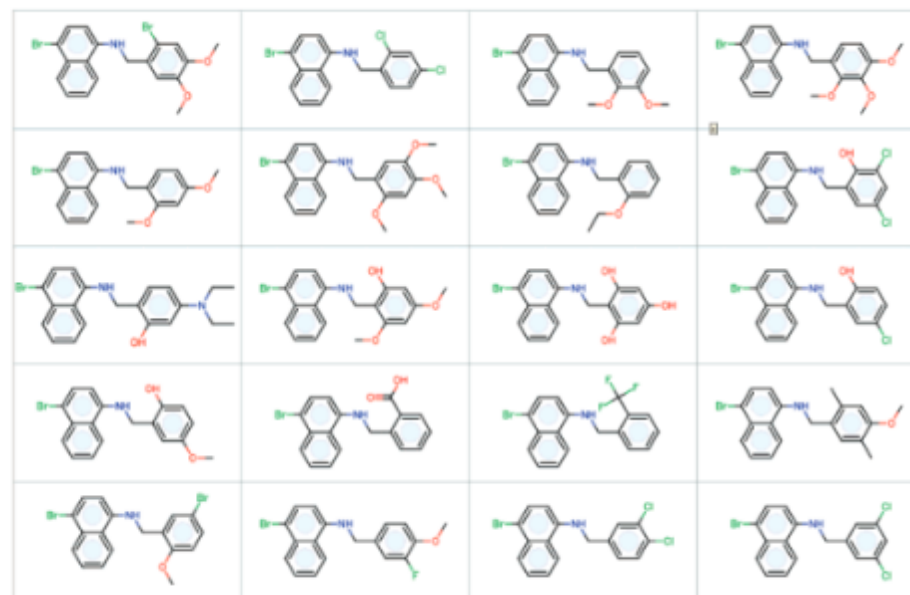
Color-coding by molecular weight



Cluster A



Color-coding by logP



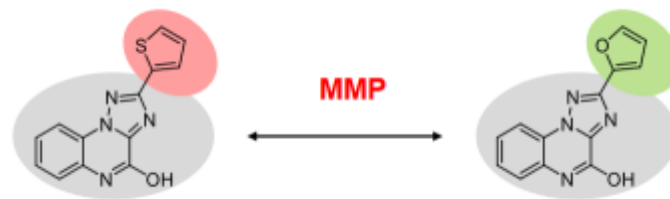
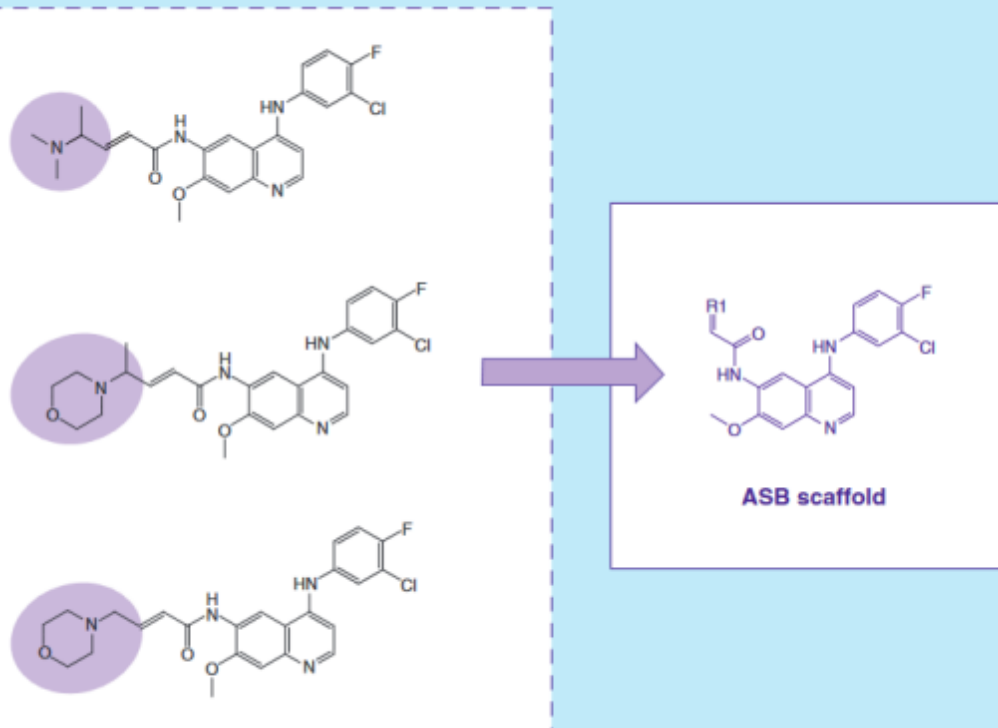
Cluster B

Новый тип молекулярных скелетов: скелеты, основанные на серии аналогов (analog series-based scaffolds)

«Присоединение к молекулярному скелету цикла интерпретируется как образование нового скелета»

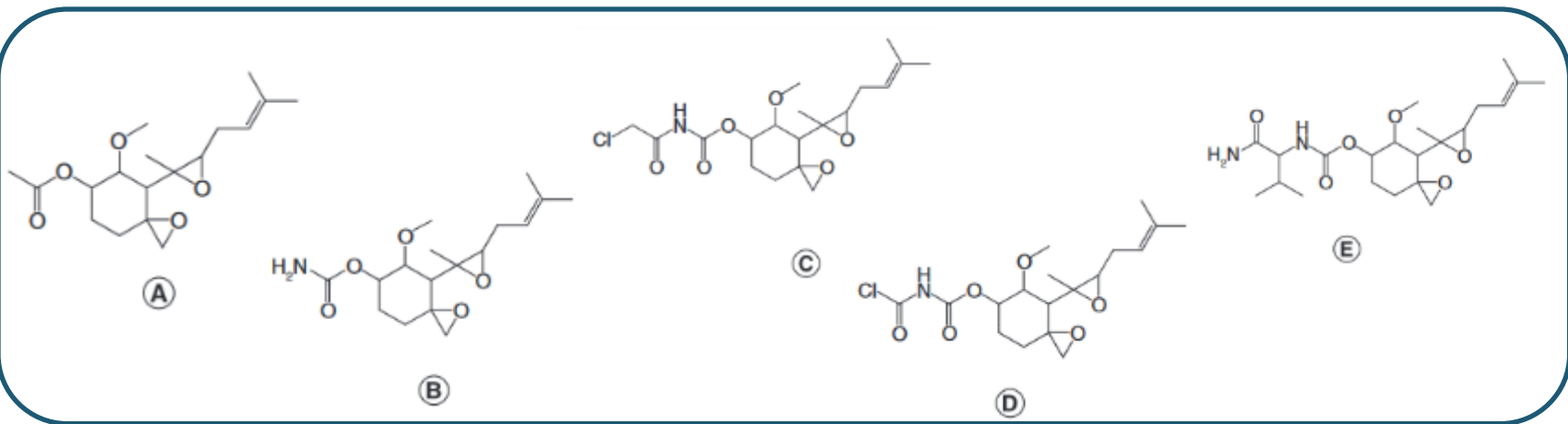
Противоречит практике, когда циклические подструктуры присоединяются в качестве функциональных групп

ASB scaffolds



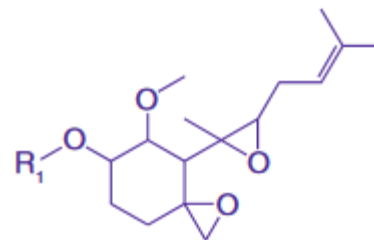
Matched Molecular Pair - a pair of compounds that are only differentiated by a chemical modification at a single site

Новый тип молекулярных скелетов: скелеты, основанные на серии аналогов (analog series-based scaffolds)

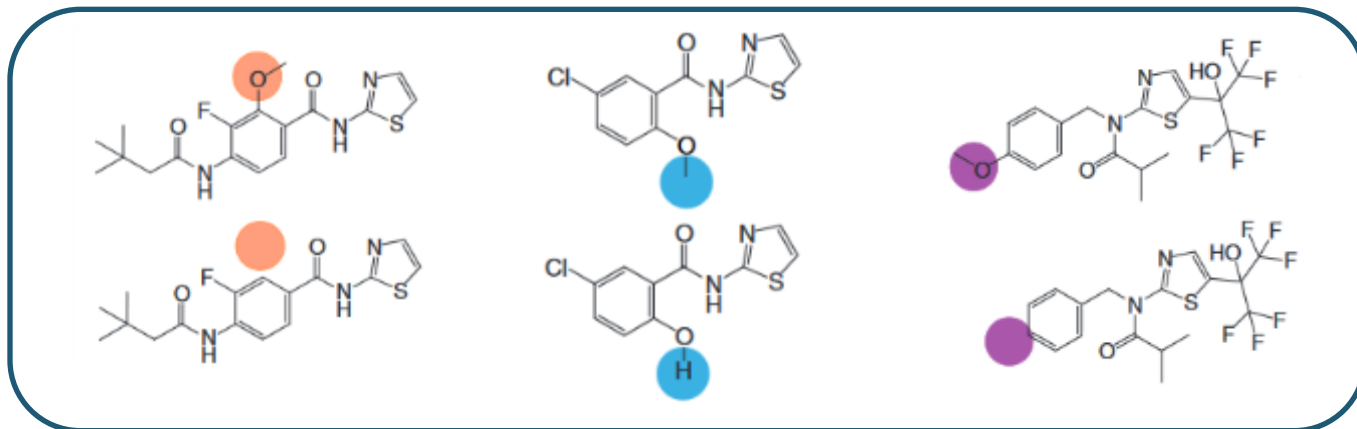


MMP core	Analogs
	A, B, C, D, E
	B, C, D, E
	B, C, D

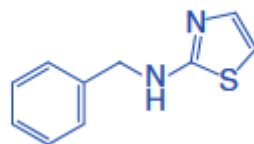
ASB scaffold



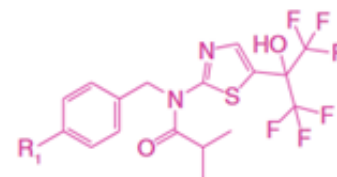
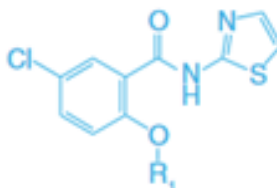
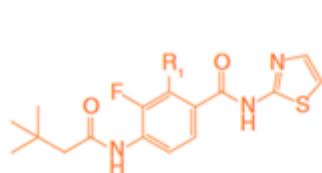
Новый тип молекулярных скелетов: скелеты, основанные на серии аналогов (analog series-based scaffolds)



Bemis-Murcko Scaffold



ASB scaffolds

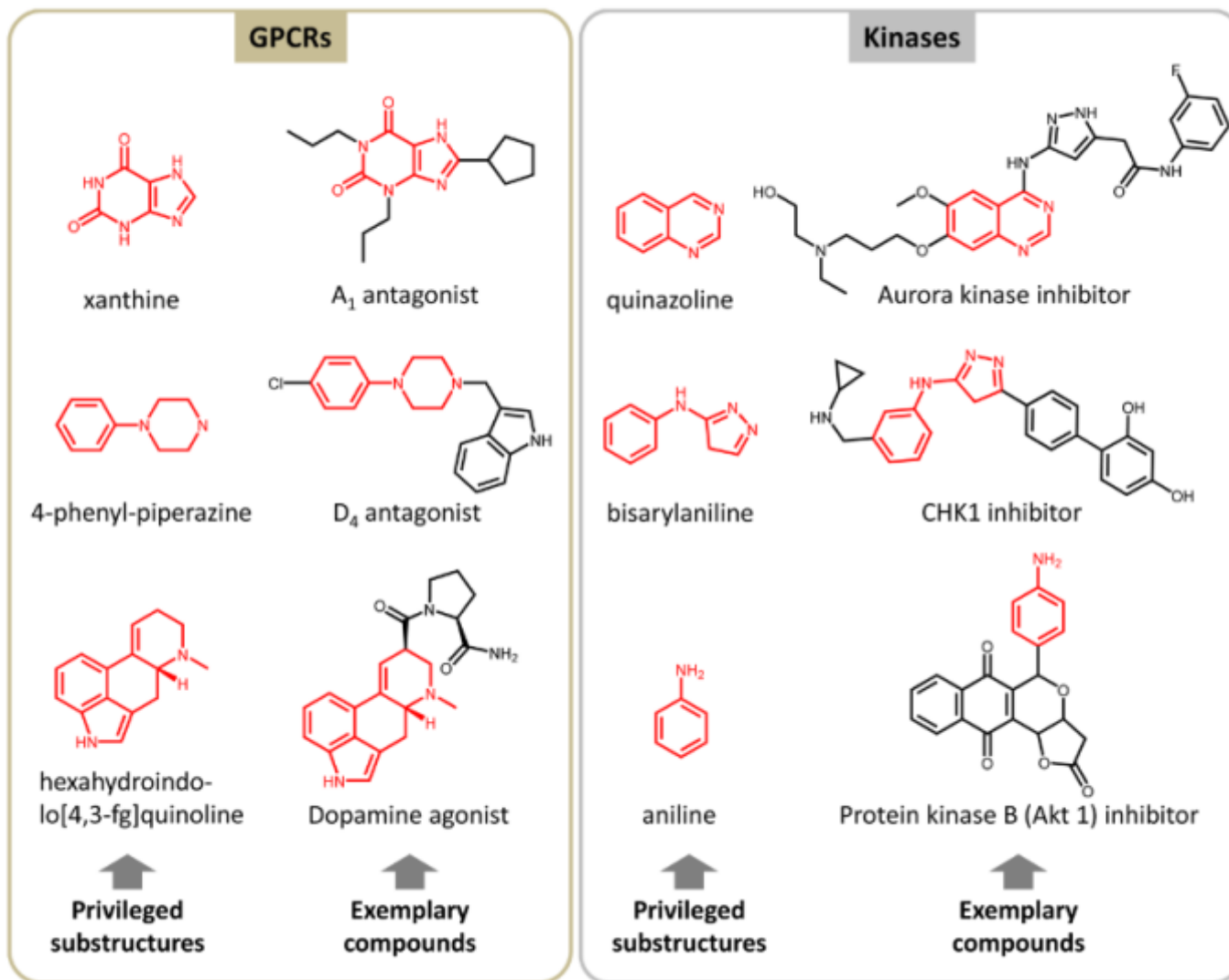


Adenosine A1 and A2a
receptor

Calcium-activated
potassium channel
subunit alpha-1

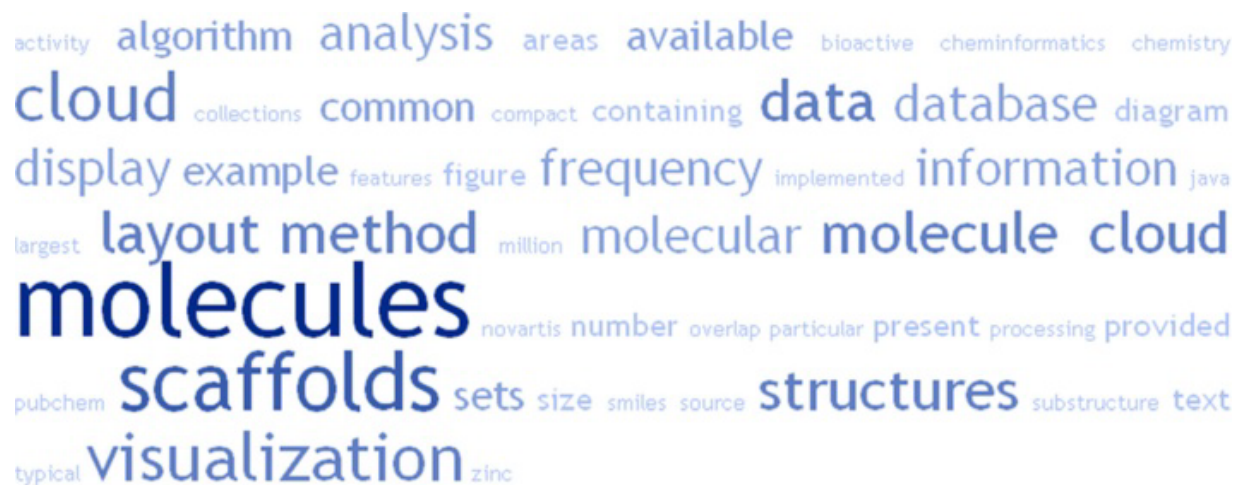
Malonyl-CoA
decarboxylase

Молекулярный скаффолд: привилегированные подструктуры (Privileged Substructures)



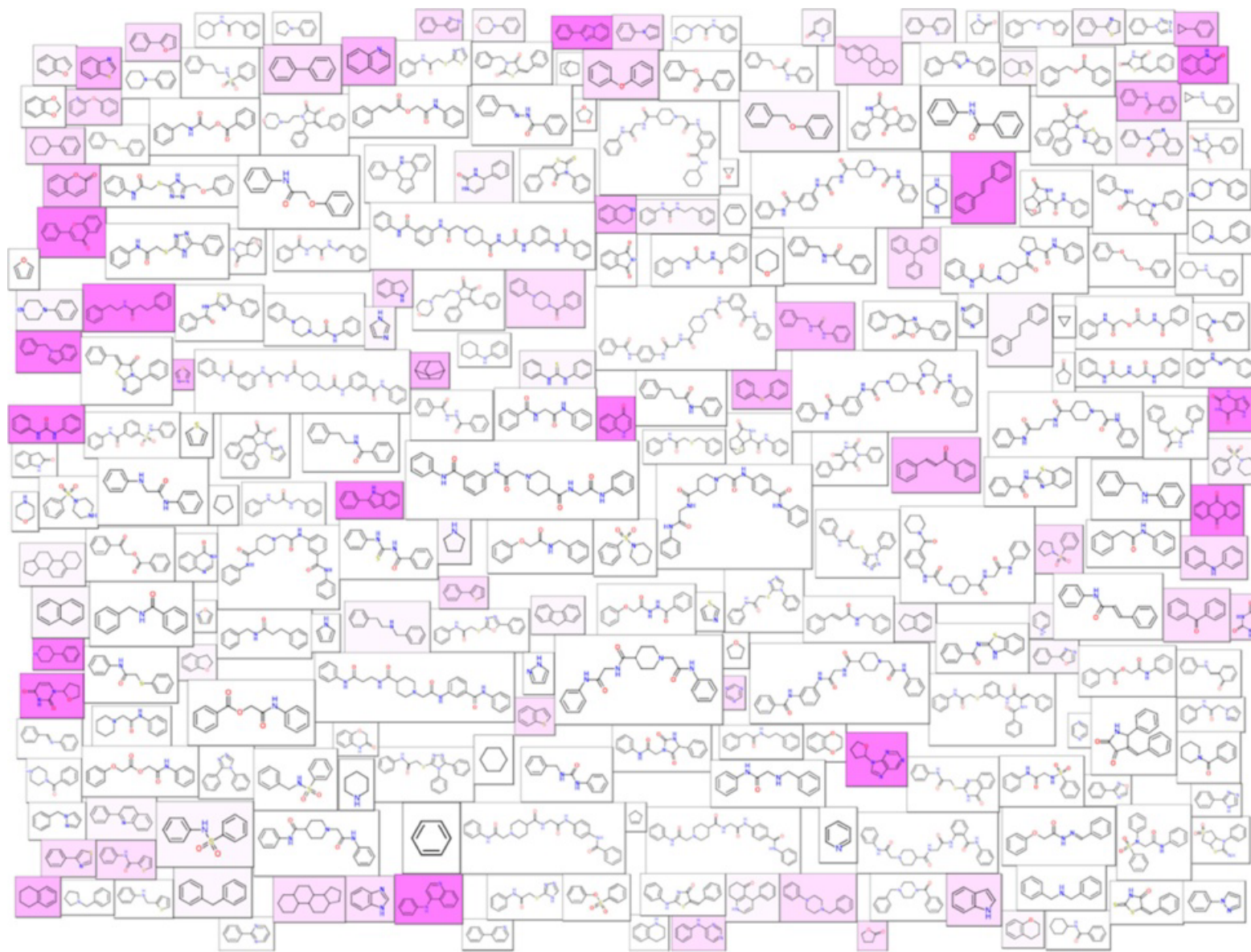
Молекулярные облака

Аналогия с «текстовым облаком» (“text cloud”)

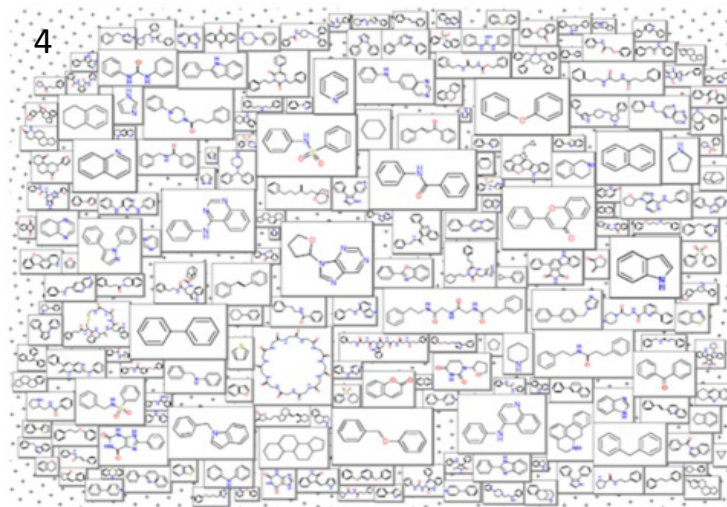
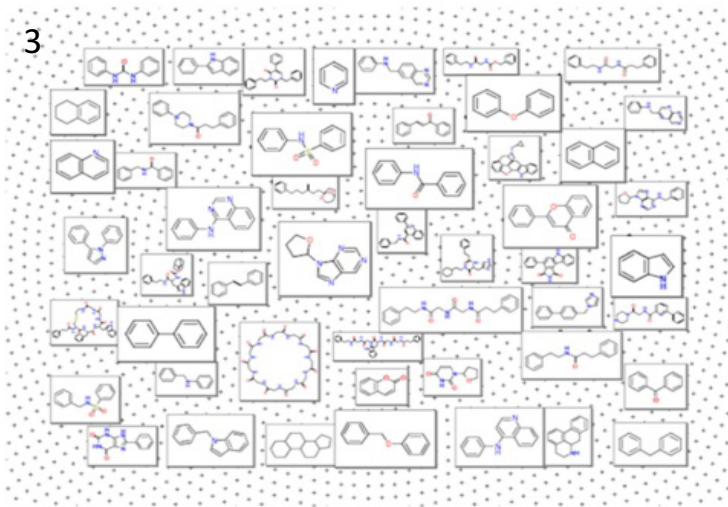
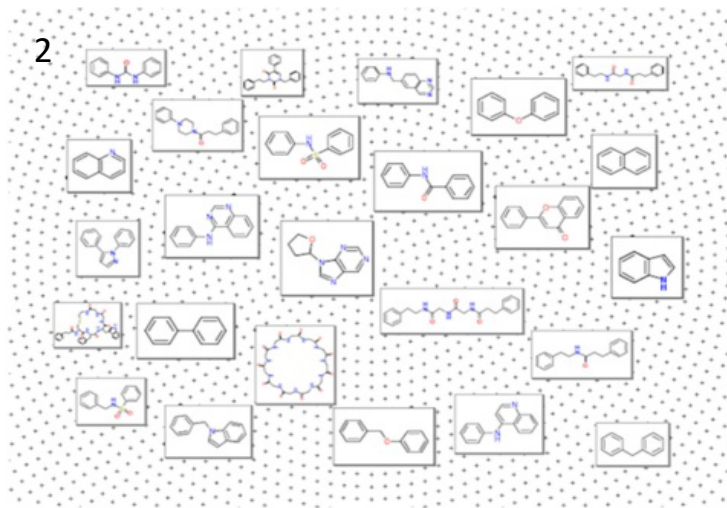
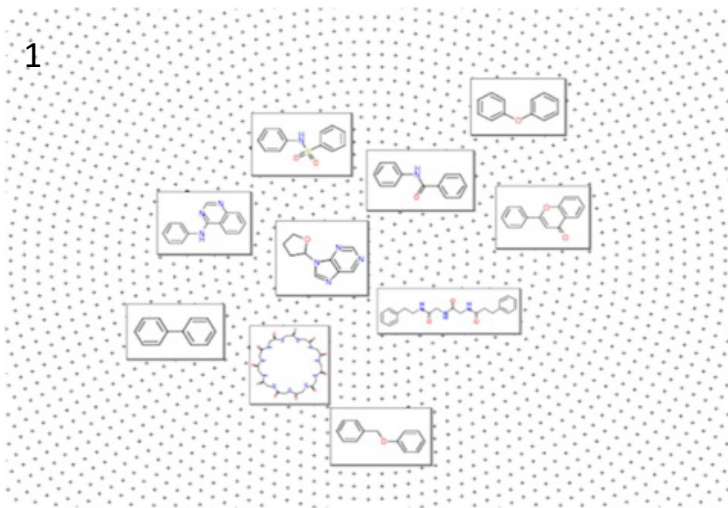


“Text cloud”: идентифицирует наиболее употребимые в тексте слова и отображает их в виде «облачной» диаграммы, где частота встречаемости определяет размер шрифта при отображении диаграммы. Наиболее часто встречающиеся слова, такие как “и”, “или” и т.д., при этом не несущие информации не отображаются

Молекулярные облака

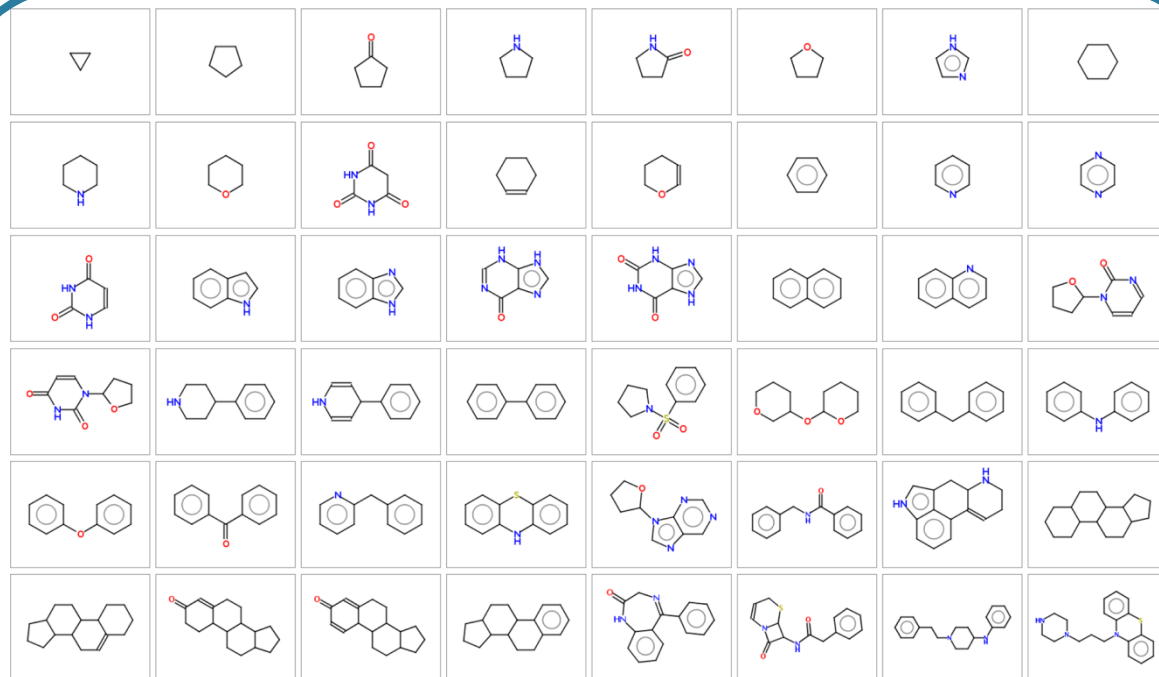


Молекулярные облака



- Наиболее представленные соединения располагаются в центральной части таким образом, чтобы минимизировать перекрытие
- Остальные соединения по очереди проецируются в сетку заданных положений, рассчитывается параметр, оценивающий перекрытие для каждого из них с последующим окончательным расположением соединения в точке, соответствующим минимальному значению перекрытия.
- Проводится аналогичный анализ для следующего соединения.

Карты молекулярных скелетов



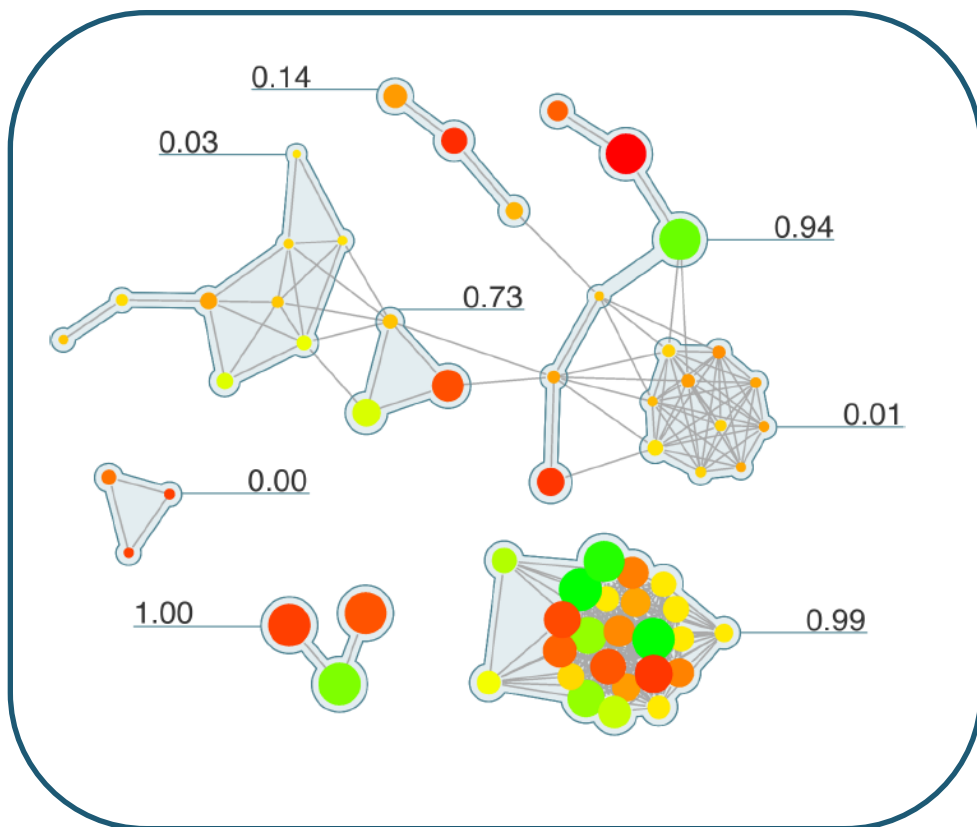
Most common scaffolds from the ChEMBL database ordered according to their Scaffold Keys

- The scaffolds are described by a set of topological descriptors: counts of atoms and bonds of particular types and various substructure scaffold features characterizing ring fusing, spiro connections, and linker properties that are called Scaffold Keys.
- The horizontal axis represents scaffold size, the vertical axis is a mixture of other scaffold characteristics and may be considered as a measure of their complexity.

n	description
1	number of ring and linker atoms
2	number of linker atoms
3	number of linker bonds
4	number of rings
5	number of spiro atoms
6	size of the largest ring
7	number of bonds in fully conjugated rings
8	number of multiple bonds in not fully conjugated rings
9	number of heteroatoms in rings
10	number of heteroatoms other than N, S, O in rings
11	number of S ring atoms
12	number of O ring atoms
13	number of N ring atoms
14	number of heteroatoms
15	number of heteroatoms other than N, S, O
16	number of S atoms
17	number of O atoms
18	number of N atoms
19	number of multiple linker bonds
20	count of two adjacent heteroatoms
21	count of 3 adjacent heteroatoms
22	count of 2 heteroatoms separated by a single carbon
23	count of 2 heteroatoms separated by 2 carbons
24	number of double bonds with at least one heteroatom
25	number of heteroatoms adjacent to a double (nonaromatic) bond
26	count of pairs of conjugated double (nonaromatic) bonds
27	count of pairs of adjacent branched atoms
28	count of branched atoms separated by a single nonbranched atom
29	count of 3 adjacent branched atoms
30	count of branched atoms separated by any 2 atoms
31	number of exocyclic and exolinker atoms
32	number of heteroatoms with more than 2

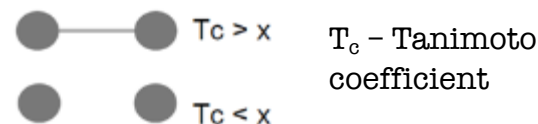
СЕТЕВОЕ ПРЕДСТАВЛЕНИЕ ДАННЫХ

Сетеобразующие графы молекулярного подобия (Network-Like Similarity Graphs)



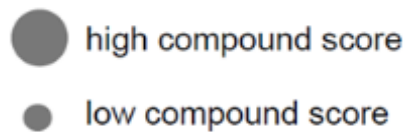
Аннотированный граф молекулярного подобия для набора данных

Подобие - относительное положение
вершин (прямая взаимосвязь
молекулярного подобия и близости
вершин)



Значение свойства - цвет вершины
(зеленая - неактивное соединение,
красная - активное)

Непрерывность - размер вершины (вершина большего размера соответствует - большая разрывность).



Сетеобразующие графы молекулярного подобия (Network-Like Similarity Graphs)

Размер
вершины

Индекс разрывности для
соединения

Выделяет соединения, образующие
«обрывы активности»

Индексы для
кластеров

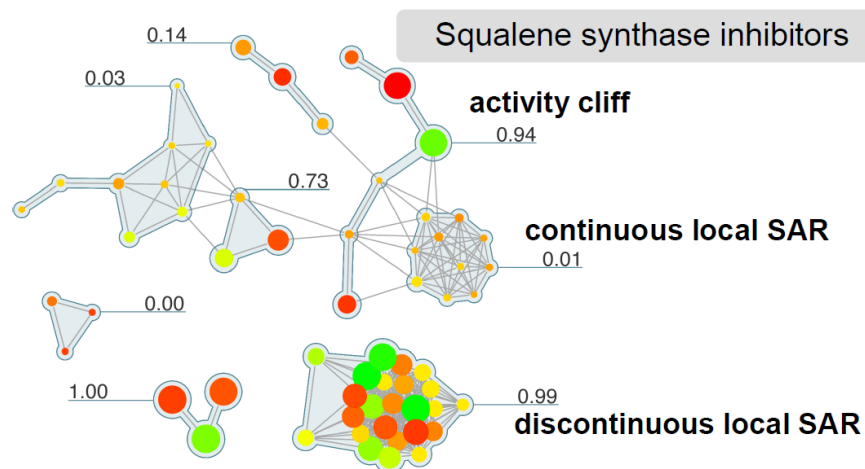
SAR индекс для кластеров

Отражает уровень
непрерывности/разрывности для группы
структурно подобных соединений

Глобальные
индексы

SAR индекс для набора
данных

Отражает уровень
непрерывности/разрывности для всего
набора данных



Global scores

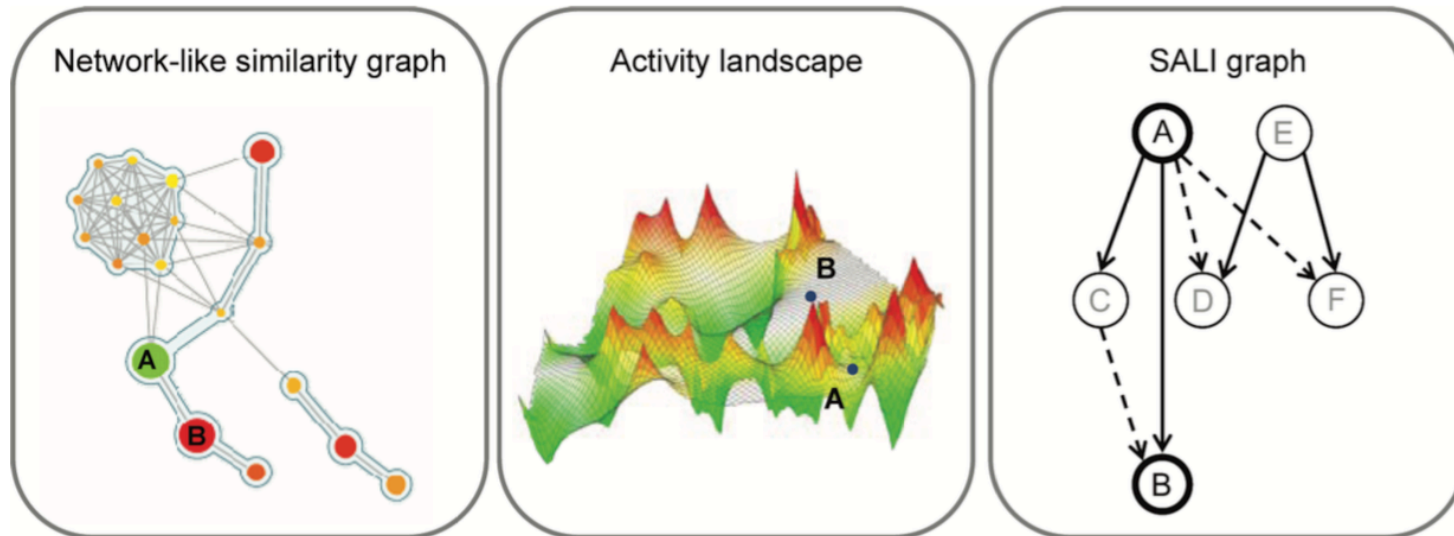
continuity	0.79
discontinuity	0.99
SAR Index	0.40

Взаимосвязь структура - свойство: Обрывы активности (Activity Cliffs)

G. Maggiora **On Outliers and Activity Cliffs - Why QSAR Often Disappoints** *JCIM* **2006** 46 (4), 1535-1535

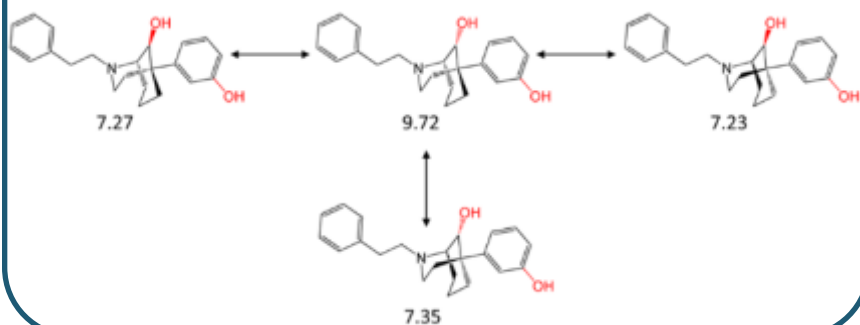
Обрывами активности (Activity Cliffs) называются пары структурно сходных соединений, характеризующиеся резкими отличиями в значении свойства

Типы представления

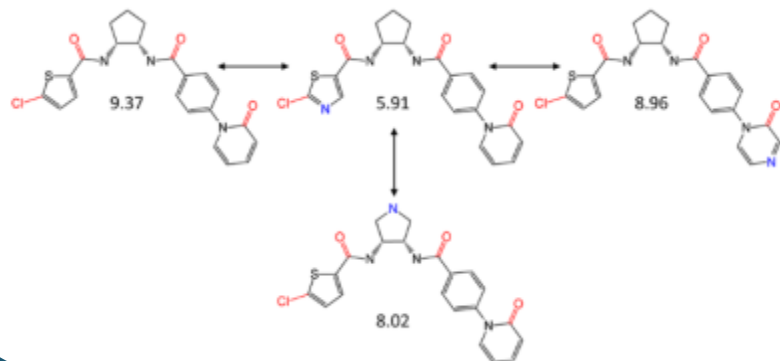


Типы обрывов активности (activity cliffs)

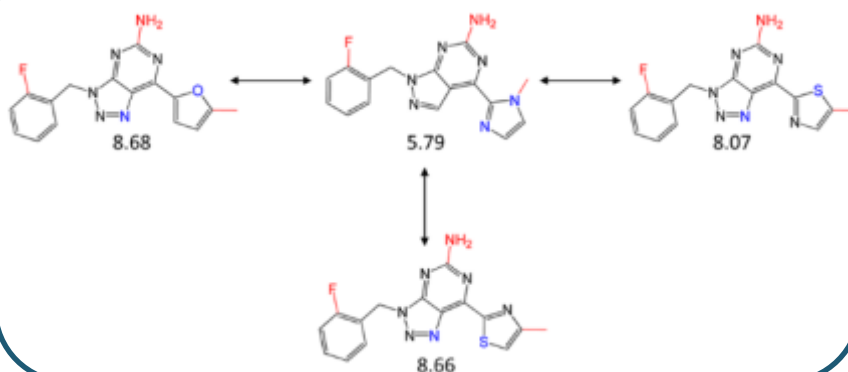
Обрывы хиральности



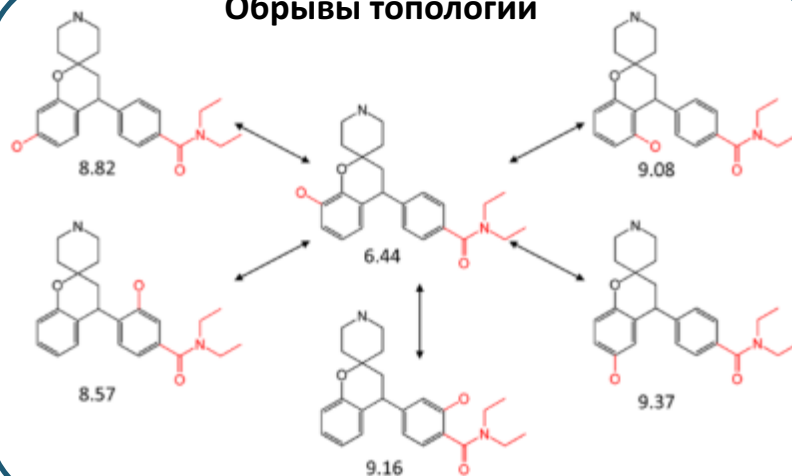
Обрывы скаффолдов



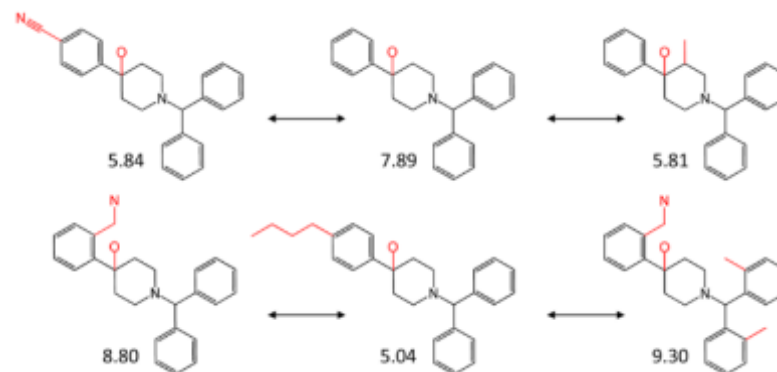
Обрывы скаффолдов/топологии



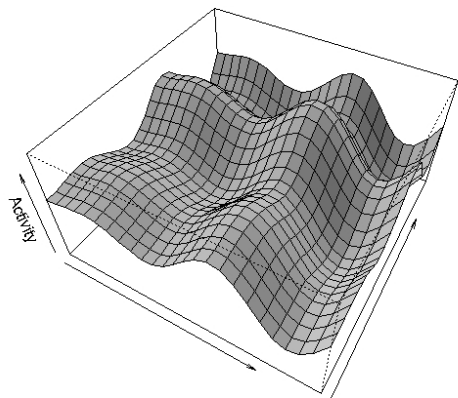
Обрывы топологии



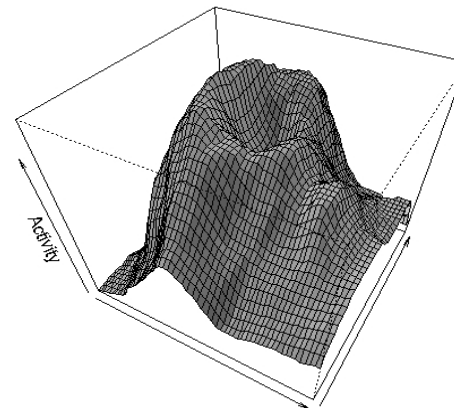
Обрывы функциональных групп



Параметры, характеризующие ландшафт активности: Structure-activity Landscape Index (SALI) и Structure-Activity Relationship Index (SARI)



$$SALI_{ij} = \frac{|A_i - A_j|}{1 - \text{sim}(i,j)}$$



Рассчитывается для соединений попарно, эффект изменений в структуре может быть легко проанализирован

Стабилен к изменению в описании соединений

Плохо работает для структурно подобных молекул

R. Guha, J.H. Van Drie. *J. Chem. Inf. Model.* **2008**, 48, 646-658

$$SARI = 0.5[Cont + (1-Disc)]$$

$$cont \approx \frac{\sum_{i,j \in [A_i > A_j]} \left(weight_{ij} \cdot \frac{1}{(1 + Tc_{ij})} \right)}{\sum_{i,j \in [A_i > A_j]} weight_{ij}}$$

$$disc_{ij} = |A_i - A_j| \cdot Tc_{ij}$$

$$weight_{ij} = \frac{A_i \cdot A_j}{1 + |A_i - A_j|}$$

Сетеобразующие графы молекулярного подобия (Network-like Similarity Graphs): примеры использования

Примеры использования
сетеобразующих графов
молекулярного сходства

Анализ обрывов активности: исходный и консенсусный вариант (согласованность между альтернативными типами дескрипторного описания)

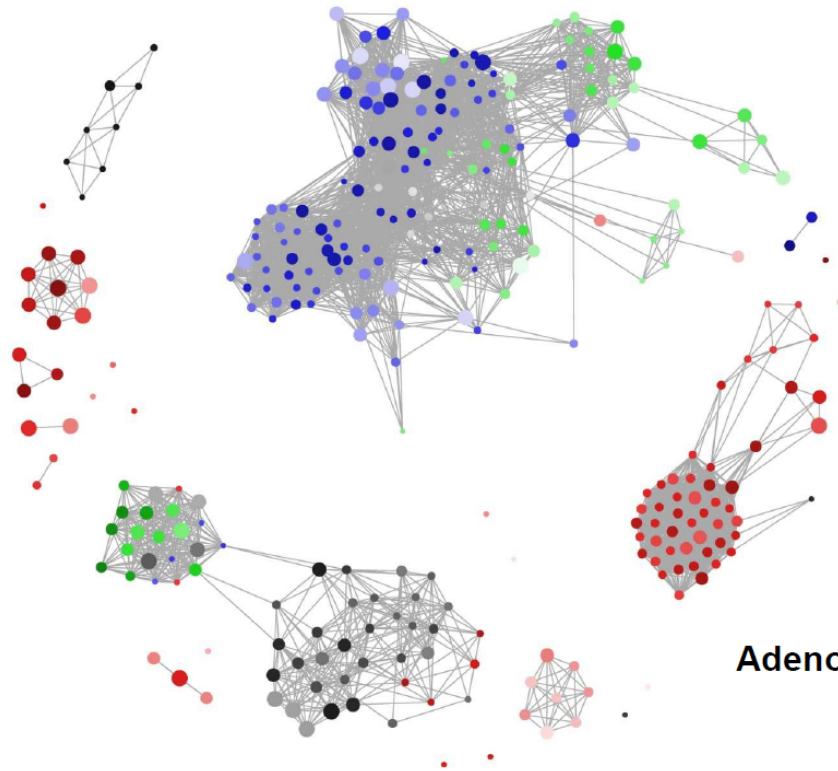
Анализ в рамках функциональных групп

Идентификация структурных особенностей, отвечающих за селективное взаимодействие с заданными мишенями

Возможность идентификации соединений с неспецифическим взаимодействием

Идентификация подструктур, отвечающих за молекулярные механизмы взаимодействия лиганд-мишень

Сетеобразующие графы молекулярного подобия (Network-like Similarity Graphs): молекулярные механизмы взаимодействия лиганд-рецептор



■ Color code: Mechanism

Antagonist (94)

Partial agonist (54)

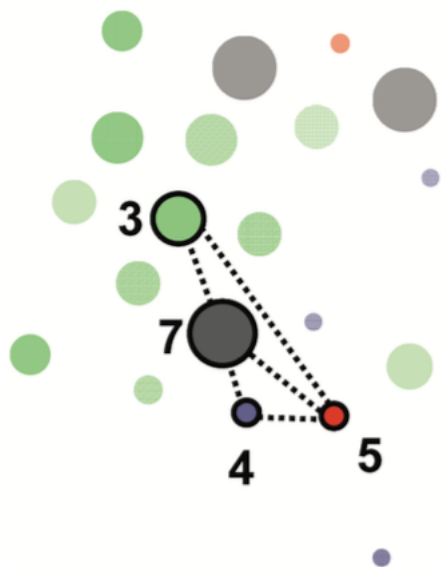
Agonist (107)

Inverse agonist (52)

Adenosine A1 receptor ligands

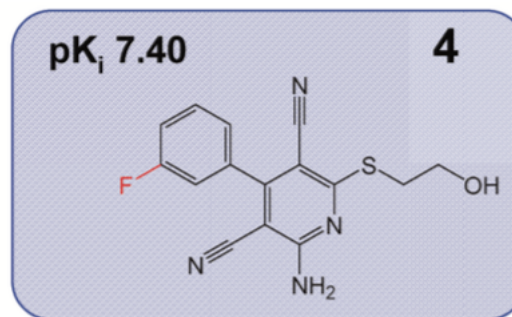
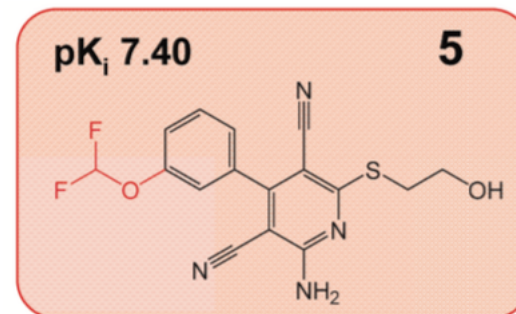
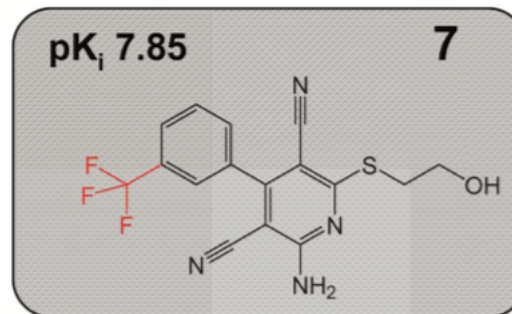
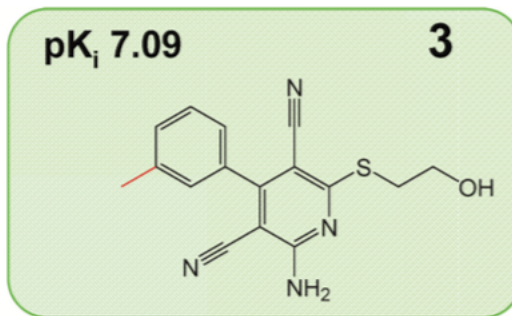
- ❖ Агонист — химическое соединение (лиганд), которое при взаимодействии с рецептором изменяет его состояние, приводя к биологическому отклику.
- ❖ Полные агонисты вызывают больший по сравнению с эндогенным агонистом отклик
- ❖ Частичные агонисты при взаимодействии с теми же рецепторами не вызывают максимального эффекта
- ❖ Обратные агонисты вызывают уменьшенный отклик рецептора.
- ❖ Антагонисты блокируют действие агонистов

Сетеобразующие графы молекулярного подобия (Network-like Similarity Graphs)

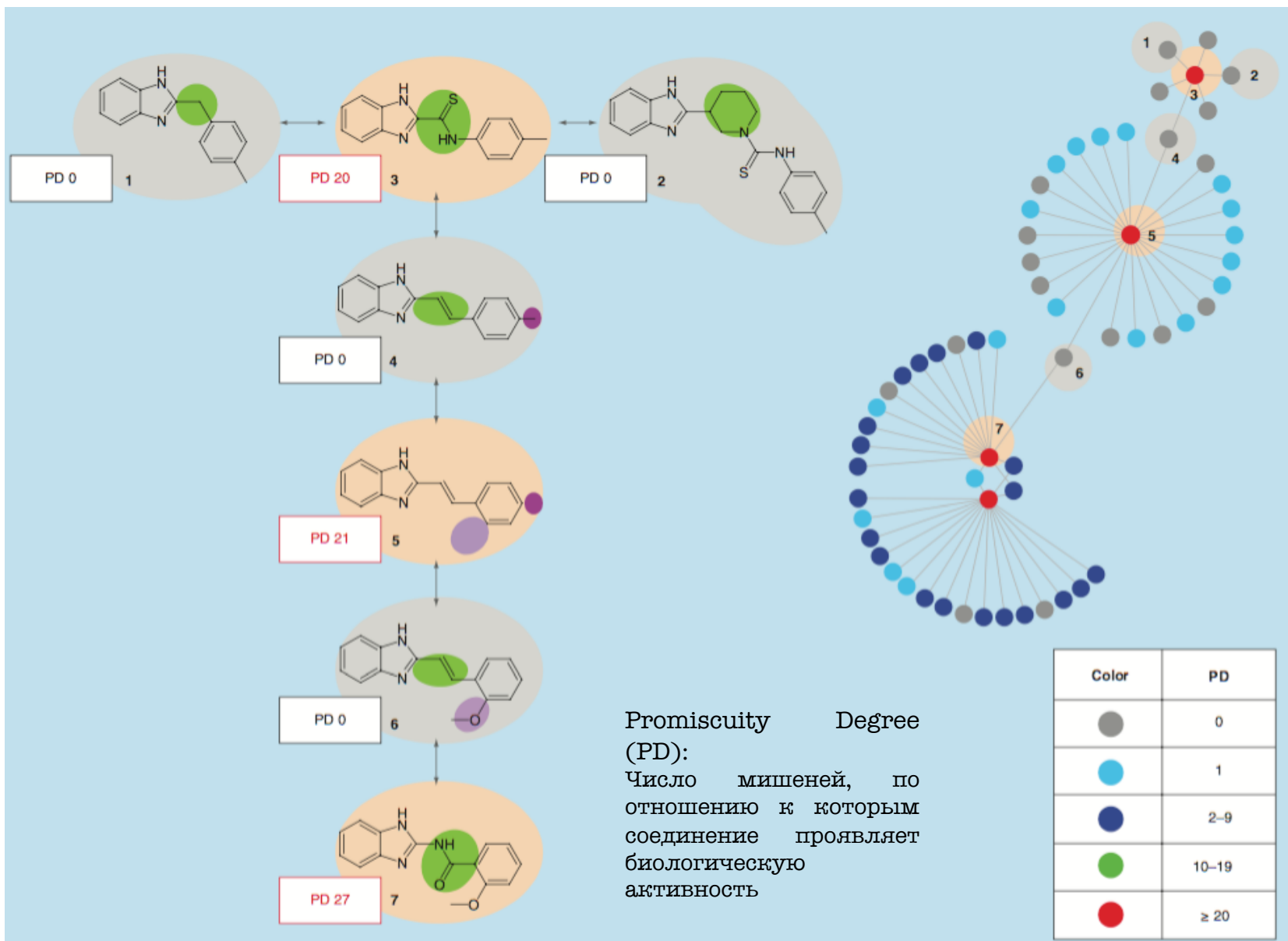


..... 'mechanism hops'

- agonists
- partial agonists
- inverse agonists
- antagonists



Сетеобразующие графы молекулярного подобия (Network-like Similarity Graphs): анализ на наличие мультитаргетной активности





Вопросы?