MOLECULAR PRINCIPLES OF DRUG DESIGN

2022

MGR. ANNA RUPRECHTOVÁ

DEPARTMENT OF CHEMICAL DRUGS

MUNI PHARM

THE THEORY OF PRIVILEGED STRUCTURES & CHEMOGENOMICS

PRIVILEGED STRUCTURES

Privileged structure is
a single molecular framework
able to provide ligands for diverse receptors,
from which can be obtained
new receptor agonists and antagonists
by rational modification.

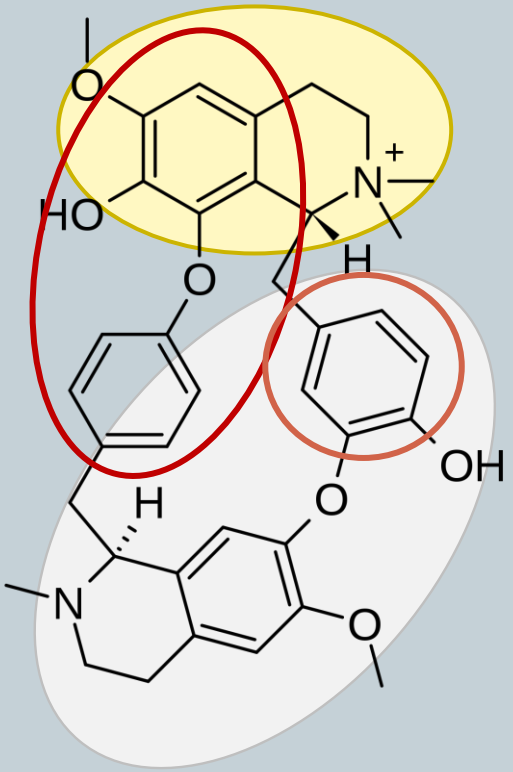
Evans & co., 1988

Characteristics of a privileged structure

- Small
- Non-planar
- Robust conformation
- Provides interesting 3D exit vectors for substitution
- Drug-like properties
- Constitutes a significant portion of the total mass of the molecule
- Ideally readily accessible synthetically

Characteristics of a privileged structure

- Cyclic structures are ideal scaffolds for drug development
  - Molecular rigidity (need lower energy for new bond forming)
  - Better bioavailability
  - Bicyclic and tricyclic molecules have ideal size for library synthesis
- Cyclic structure - usually has 2-3 rings
  - Condensated or linked by one or two bonds
  - Usually substituted by a phenyle
  - Usually contains a heteroatom



Characteristics of a privileged structure

- Molecules with therapeutical effect = drug-like properties
- Lipinski's rule of five (Pfizer's rule of five, RO5)
  - Molar mass < 500 g/mol
  - logP < 5 (lipophilicity)
  - Donors of H-bonds <5
  - Acceptors of H-bonds <10 (2x5)
- Physicochemical characteristics for promiscuous binding
- Smaller molecule has higher capacity of binding to multiple receptors
- Simple ligand surface

Promiscuous ligands

Promiscuous ligands

- potently
- specifically
- reversibly
- but not selectively

bind to members of different macromolecular target families (receptors, enzymes, …)

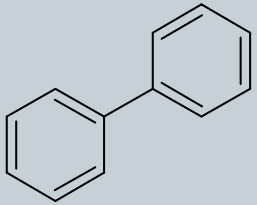
PRIVILEGED STRUCTURE

x „drug-like“ structure

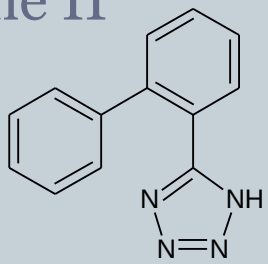
x non-specific „protein-binder“

Bonding and specifity

- (sub)structure provides bonding on protein
- linked SUBSTITUENTS are responsible for receptor specifity



  - Structure of biphenyle was described as privileged (sub)structure – it can be found in 4,3 % of all known drugs with different biologic activity
  - 2-tetrazolobiphenyle: part of approx 1/5 drugs with biphenyle; those drugs mainly block receptor for angiotenzine II



Why are so interesting?

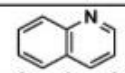
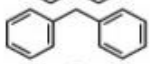
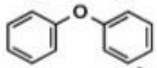
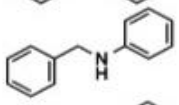
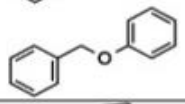
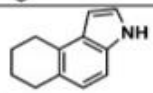
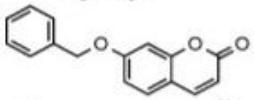
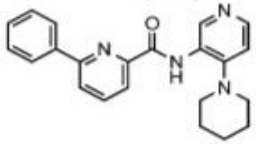
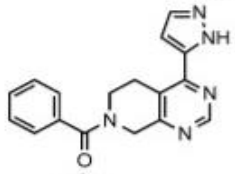
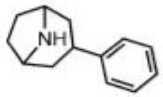
- Modeling of potential “small molecules”
  - Molecules with low molar mass (<500 Da)
  - Number of possible small molecules is 10^{200}
  - Drug-like properties may have 10^{60} of them
- New drug targets
  - Target structures newly discovered by genetics, proteomics,…
  - Better information about known diseases
- Bring new drugs to the pharmaceutical market
  - Drugs for new diseases
  - The development of a new drug takes a long time and it is expensive
- Structural libraries

Example of finding a privileged structure

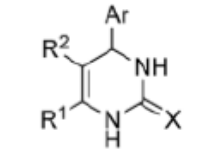
- Analyses of ChEMBL22 database
- 1.397.535 compounds
- 181.888 scaffolds

- The most frequent scaffold was a single phenyle ring
- Group 1:
  - 2 rings closely linked (methyl, ether, amine, amide)
  - Similar observation for natural products
  - In 12 % drugs with Mr<2000 approved by FDA (1939-2006)
- Group 2:
  - A lot of them target-specific
  - Can be told privileged structures?
- ↑ number of sp^3 atoms and H-bond acceptors = ↓ promiscuity

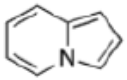
Cit: Schneider, P., Schneider, G.: *Privileged Structures Revisited*. Angew. Chem. Int. Ed. 2017, 56, 7971-7974

ID	Scaffold	N	Welsch list[8]
1a		285	yes
2a		545	
3a		1184	yes
4a		245	
5a		828	yes
1b		280	yes
2b		282	yes
3b		103	
4b		190	
5b		870	

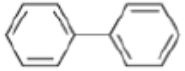
Examples of privileged structures



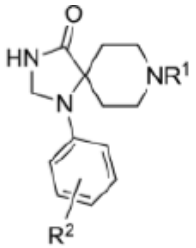
Dihydropyrimidone
[X = S or O]



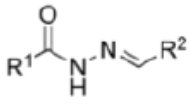
Indolizine



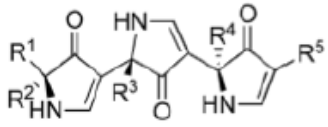
Biphenyl



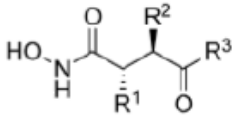
Triazaspirodecanone



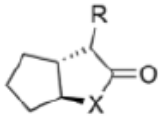
N-Acylhydrazone



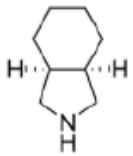
Pyrrolinone



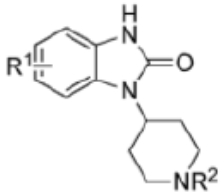
Hydroxyamate



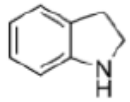
trans-Lactam/Lactone
[X = N or O]



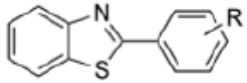
Hexahydroisoindole



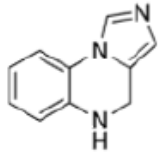
Benzimidazolone



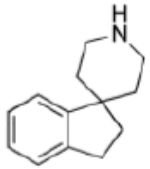
Indoline



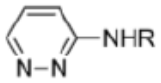
2-Arylbenzothiazole



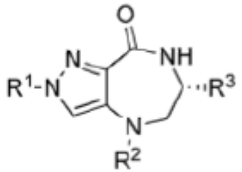
Imidazolequinoxaline



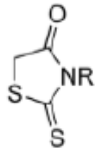
Spiroindanylpiperidine



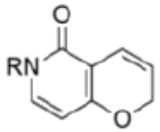
Aminopyridazine



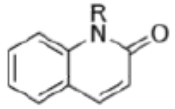
1,4-Pyrazolodiazepin-8-one



Rhodanine



Pyranopyridone

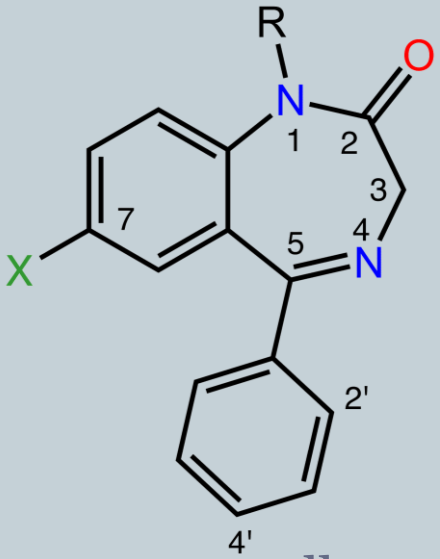


Pyranoquinolone

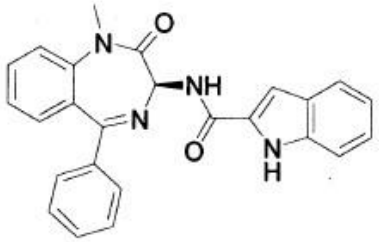
„The first one“

1,4-benzodiazepin-2-on

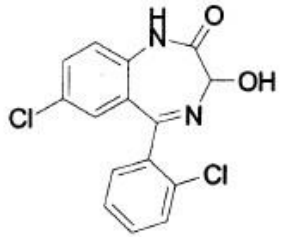
- Small library of 192 molecules => cholecystokinin A receptor => several active compounds (Ellman, early 90s)
- Larger library of 1680 compounds (Welsch, 2010)
- Targets
  - Cholecystokinin (*devazepine*)
  - Gastrin
  - Central benzodiazepine receptors (*lorazepam*)
  - Neurokinin-1 antagonistS
  - K-secretase inhibitors
  - Farnesyl:protein transferase inhibitors
  - Delayed rectifier K+ current modulator



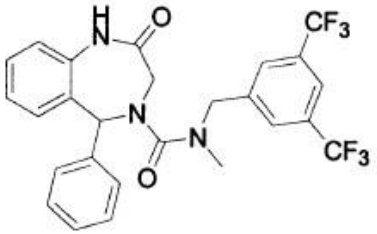
- It is thought to be the privileged due to its ability to structurally mimic beta peptide turns (Ripka, 1993)



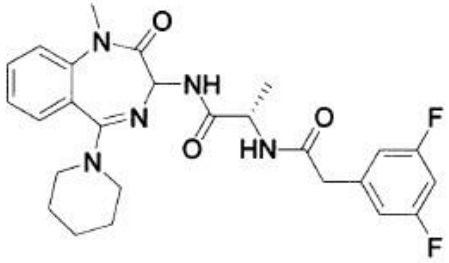
1
Devazepide
CCK-A Antagonist



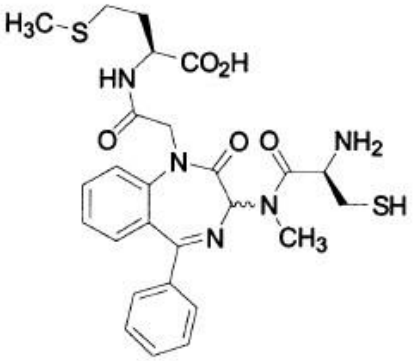
2
Lorazepam
Benzodiazepine Agonist



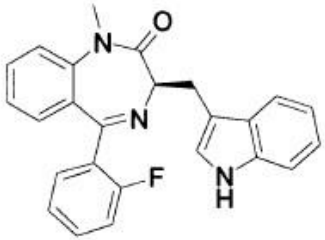
3
Neurokinin-1 Antagonist



4
κ-Secretase Inhibitor

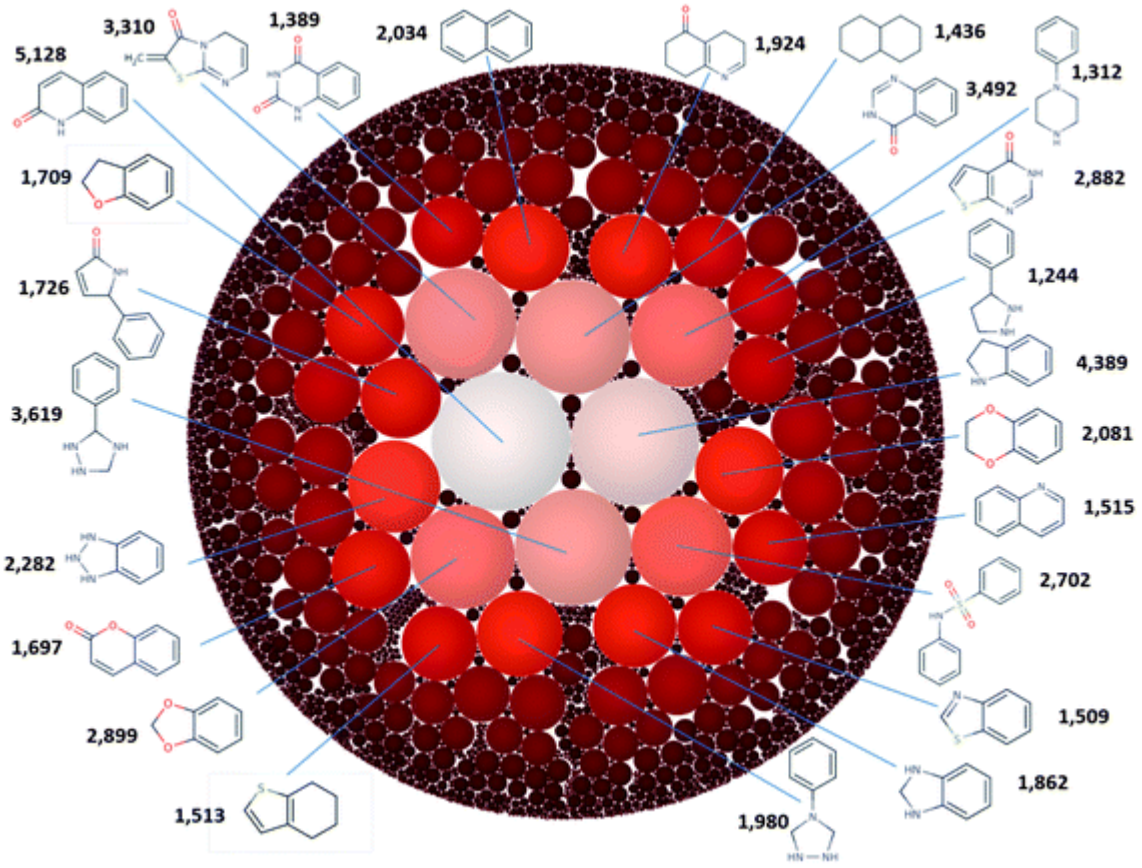


5
Farnesyl:Protein Transferase
Inhibitor



6
Delayed Rectifier K^+
Current Modulator

LIBRARIES


3,310
1,389
2,034
1,924
1,436
5,128
3,492
1,312
1,709
2,882
1,726
1,244
4,389
3,619
2,081
1,515
2,282
2,702
1,697
1,509
2,899
1,862
1,513
1,980

Structural library

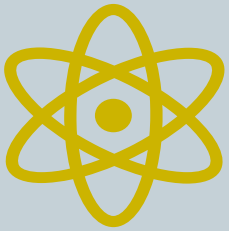
Structural / chemical /compound library
is a collection of stored chemicals.

- Can consist in simple terms or a series of stored chemicals.
- Each chemical has associated information stored in some kind of **database** with information such as
  - the chemical structure
  - purity
  - quantity
  - physicochemical characteristics of the compound.

Two important aspects of structural library

BIOLOGICAL RELEVANCE

how much it shows a (beneficial) health effect



MOLECULAR DIVERSITY

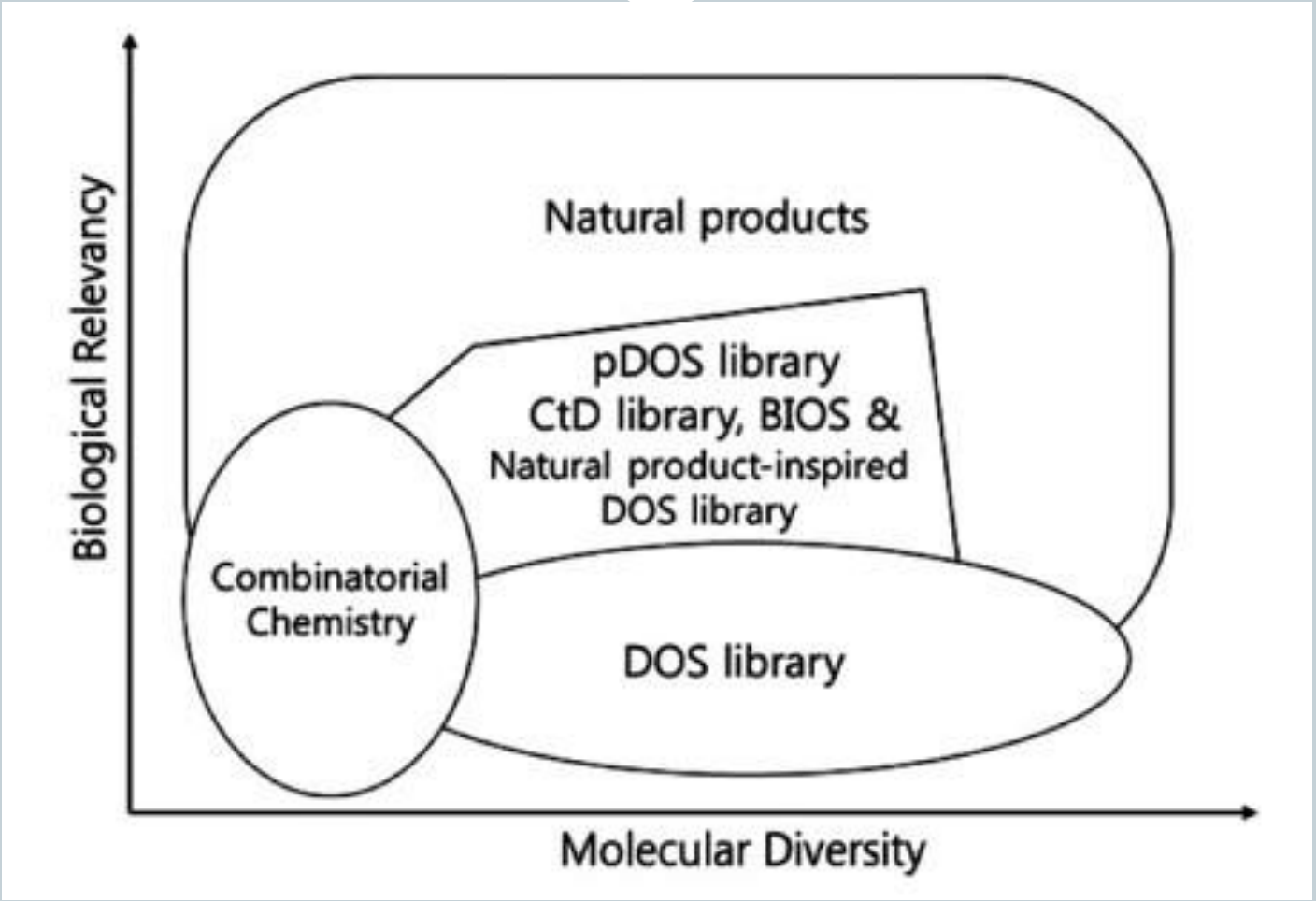
how much is composed of differing elements

Sources for the structural library

Natural sources: huge diversity of compounds, biologically active, but mostly have limited spectrum of effects (over-specialized)

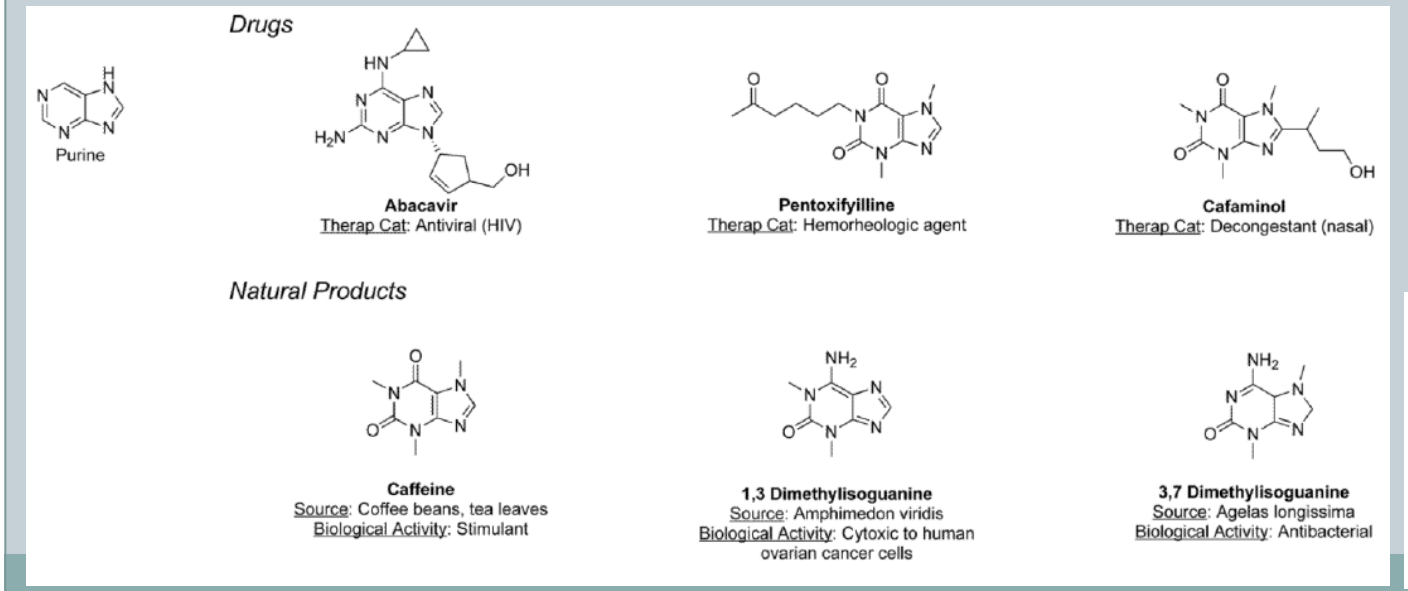
Synthetic sources: derived from privileged scaffolds, synthetic compounds with proven biologic activity, structures proven by molecular modelling, etc.

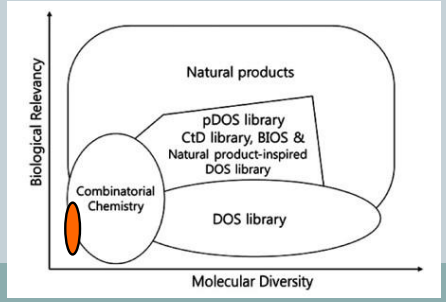
Structural libraries



1. Library based upon one core scaffold

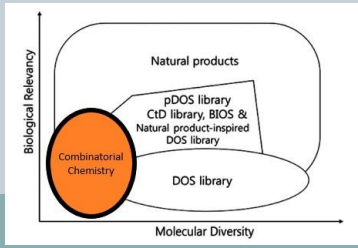
- Screening of one molecule – natural or synthetic, and its derivatives
- Traditional approach
- Looking for a variety of different targets





2. Combinatorial chemistry

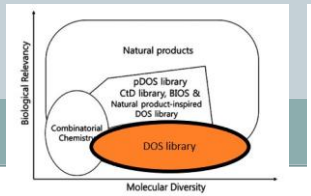
- Assemble every possible combination of a given set of chemical "building blocks" (a few basic molecules)
- Prepare a large number (tens to thousands) of compounds in a single process … $A^x + B^y \rightarrow C^{xy}$
- Especially valuable in the very first steps of screening hit exploitation and lead structure optimization, in order to derive first structure–activity relationships
- Usually goes with the **High-Throuput Screening**
  - identify through compound library screenings, candidates that affect the target in the desired way, so called "hits" or "leads"
  - and provide suggestions for their optimization

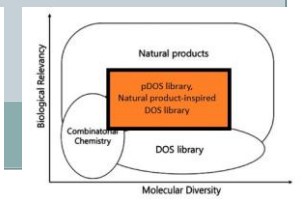


3. DOS a pDOS libraries

Resulting libraries contain complex and diverse structures with a high fraction of sp3-hybridized carbon atoms and more stereogenic centers.

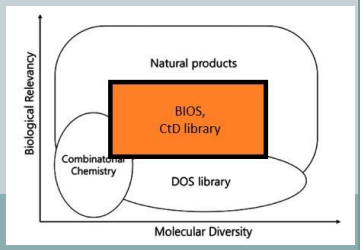
- **Diversity Oriented Synthesis**
  - Apply a variety of reaction conditions to starting materials with multiple different functional groups
  - Multiple rounds of such reactions result in rapid access to structurally diverse products suitable for screening
- **Privileged-substructure-based Diversity Oriented Synthesis**
- **Natural product-inspired DOS library**





4. BIOS and CtD

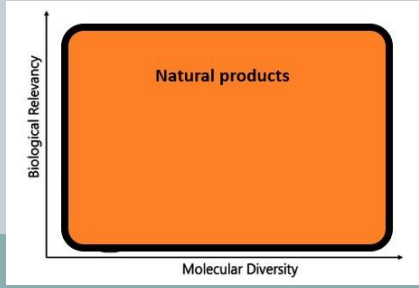
- **Biology Oriented Synthesis**
- **Complexity to Diversity**
  - Use of core structures derived from bioactive natural products as synthetic scaffolds
  - Structural motif and core skeletons from bioactive natural products can serve as chemical “navigators” for the synthesis of novel core skeletons
  - Construction of natural product-like small-molecule collections starting from commercially available natural products (e.g. abietic acid, adrenosterone, quinine)



5. Natural products

The field of the **pharmacognosy…**

… the study of the physical, chemical, biochemical, and biological properties of drugs, drug substances, or potential drugs or drug substances of natural origin as well as the search for new drugs from natural sources.



CHEMOGENOMICS

=

the discovery and description of

all possible drugs

for

all possible drug targets

CHEMOGENOMICS

=

the systematic identification

of small molecules

that interact with the products of genome (proteins)

and

modulate their biological function

CHEMOGENOMICS

=

the investigation of classes of compounds (libraries)
against families of related proteins

Some definitions

- GENOMICS = an interdisciplinary field of science focusing on the structure, function, evolution, mapping, and editing of genome
- Genome = a genome is an organism's complete set of DNA, including all of its genes
  - Human Genome Project (1990-2003): identification of 20 000-25 000 genes in human DNA
- Genetics = study of individual genes and their role in inheritance

- PROTEOMICS = refers to the large-scale experimental analysis of protein
- Proteome = the entire set of proteins that are produced or modified by an organism or system

More definitions

- Chemical genomics / CHEMOGENOMICS
  - the systematic identification of small molecules that interact via a specific molecular recognition mode with target proteins encoded by the genome
  - the term **chemogenomics** is applied more specifically to target family approaches in drug discovery
- Chemical genetics
  - identify chemical compounds which induce or revert specific biological phenotypes by using cell-based or microorganism-based screening of compound
- Chemical biology
  - the functional and mechanistic investigation of biological systems using chemical compounds and constitutes a more general discipline

Chemogenomic terminology

- 1996
- Company Glaco Wellcome

- Systematization of drug discovery within target families (products of genome) based on the analysis of gene families
  - Enzymes
    - Kinases
    - Proteases
  - GPCRs (receptors conected with G-proteins)
  - Nuclear receptors
  - Ion channels
  - Transport proteins
  - …

Classification of proteins

I. identification of all members of a gene family
II. classification into subfamilies
III. revelation of common elements and patterns in the sequence and tertiary structures

Classification of proteins

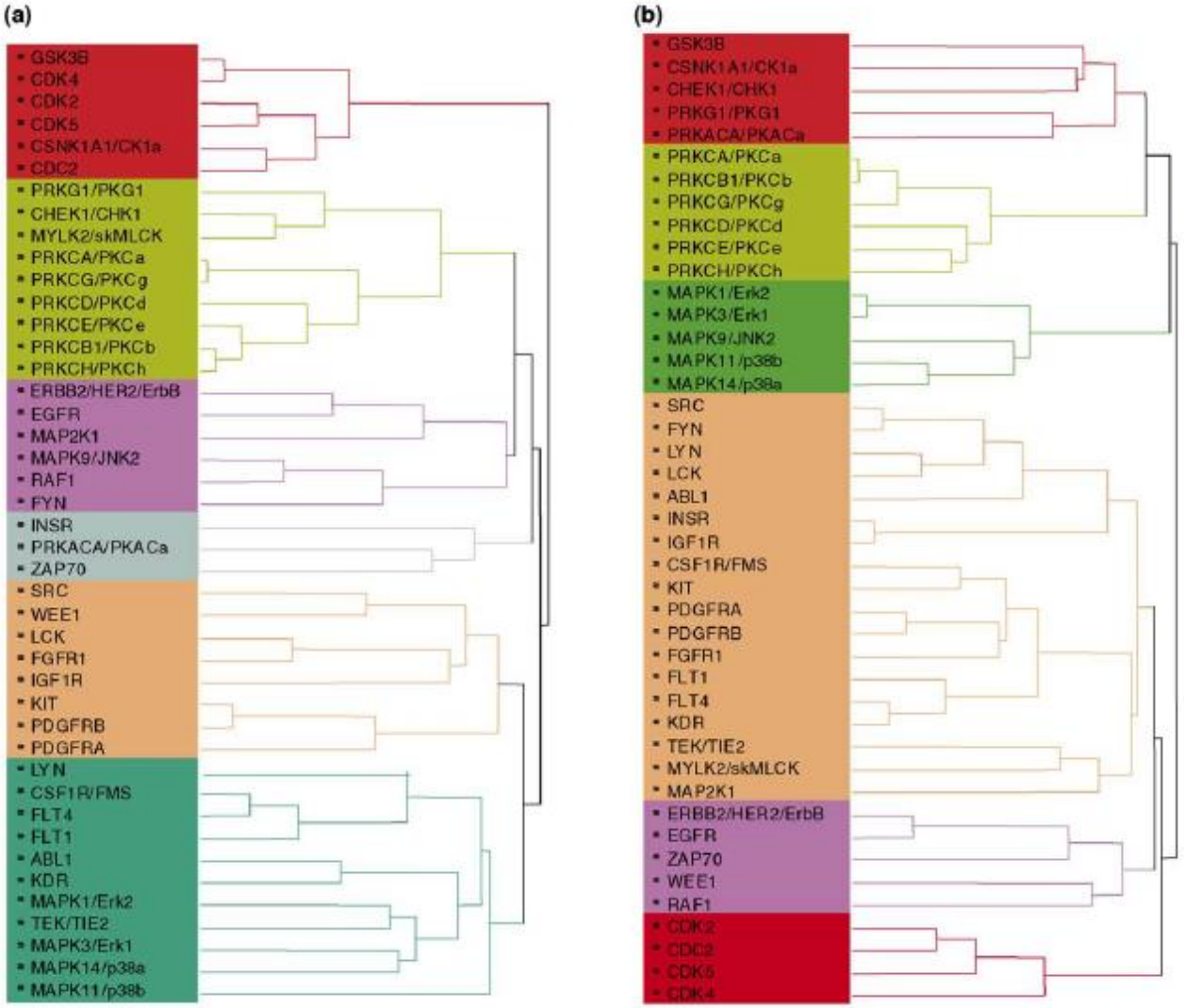
CHEMOGENOMIC

- Structure-activity relationship (SAR)
- **Small-molecule binding sites**
- Only a part of the protein structure, most commonly in the active site of the protein

PHYLOGENETIC

- Classical
- **Whole structure** of the protein

Classification of protein kinases:
a) structure-activity relationship (SAR), b) conventional phylogenetic approach



Chemogenomic approach

- Members of the same protein family can share important practical aspects
- Similar ligands should bind to similar targets → knowledge previously obtained → transferable to new related projects
- Analysis does not depend on knowledge of a biological function of the protein

Two applications of chemogenomic

Forward/classical chemogenomics

- A phenotype is known
- Small molecules interacting with this function are identified and compound hits are categorized
- ***Search for the proteins*** responsible for the phenotype

Reverse chemogenomic

- Gene sequences of interest are first cloned and expressed as target proteins, which are then screened in a high throughput,'target-based' manner by the compound library
- ***A phenotype is validated***, identification or confirmation of the role of the protein in the biological response

Example

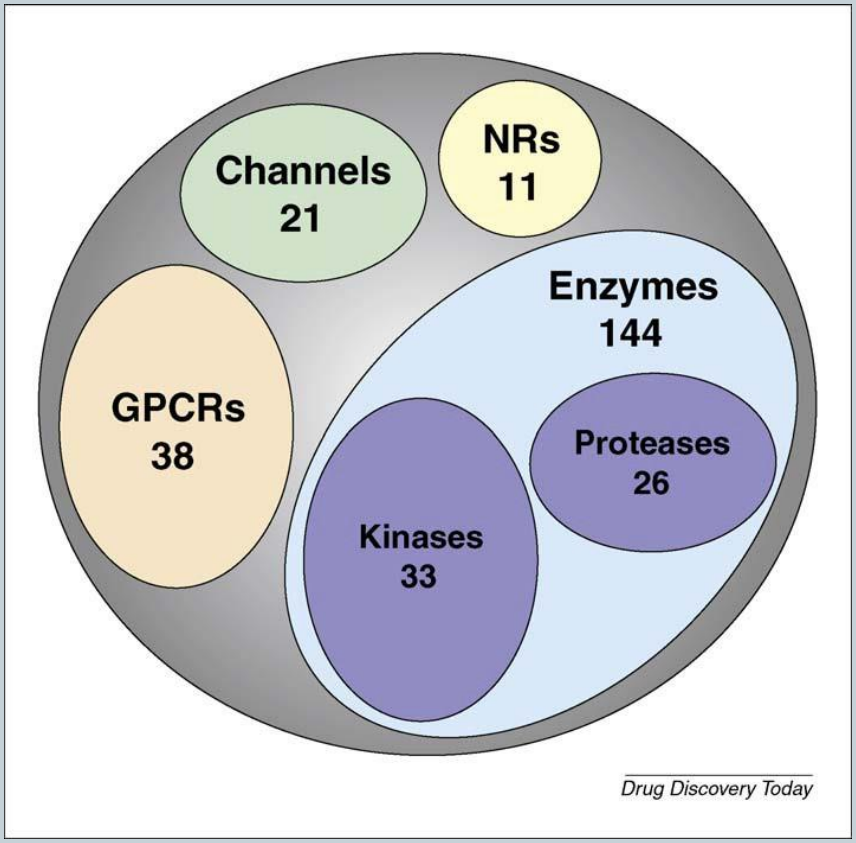
CHEMOGENOMIC APPROACH IN CARDIOVASCULAR DISEASES

1. Identification of target structures associated with cardiovascular diseases (literature research)
   - There were identified many proteins associated with cardiovascular system (233)
   - There were prepared a hundreds of small molecules (potential drugs) with a proven activity on CVS (44 032)

Example

2. Organisation of cardiovascular targets in protein families
   - GPCR (G-related couple proteins)
   - Enzyme
     - Kinase
     - Protease
   - Nuclear receptor
   - Ion channel

Complication: alternative nomenclature of proteins



Example

3. The establishment of knowledge base of the cardiovascular target space
   - Determining the level of applicability of structure-based methods for *in silico* target profiling
   - Is the structure of targets known?
   - If it is – is there any relationship, similarity?
   - Which belong to the *Homo sapiens*?

Complication: there can be used only proteins with known structure (at least one characteristic representative of the family).

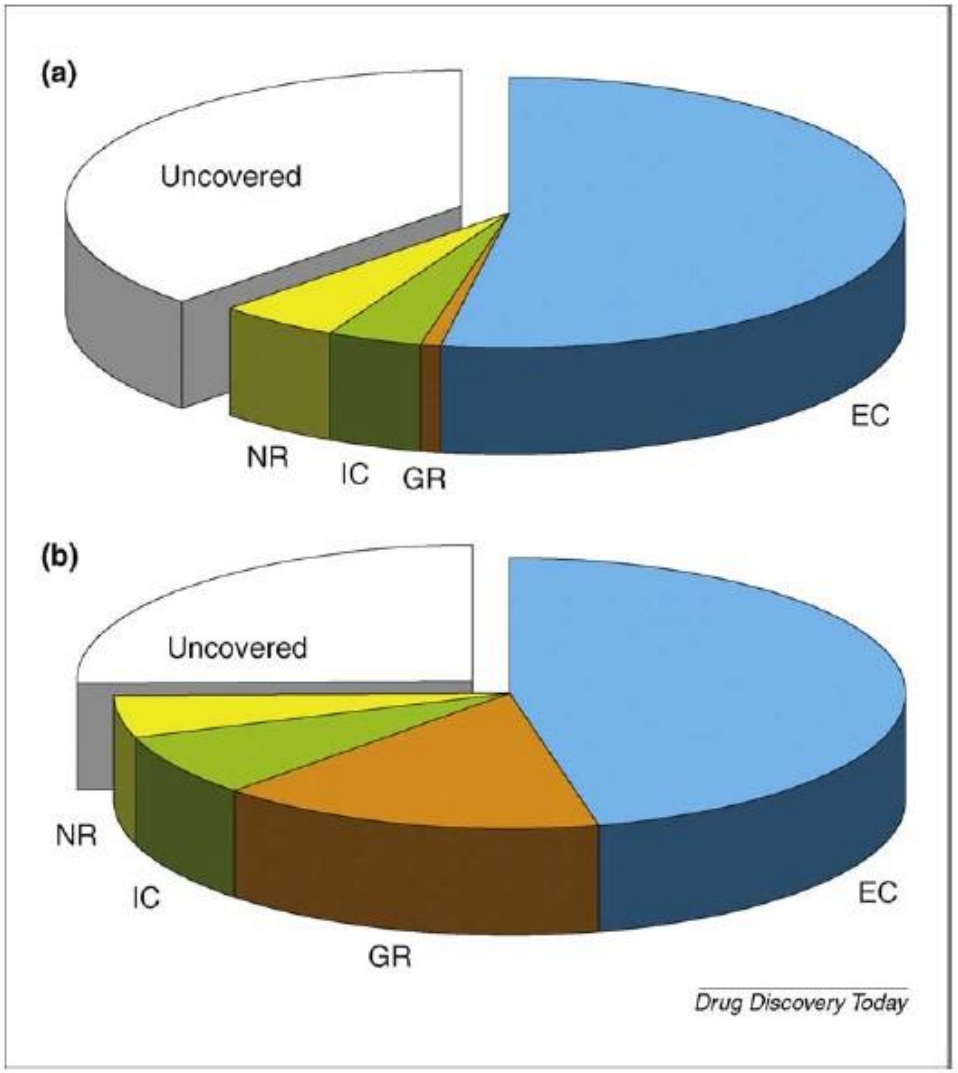
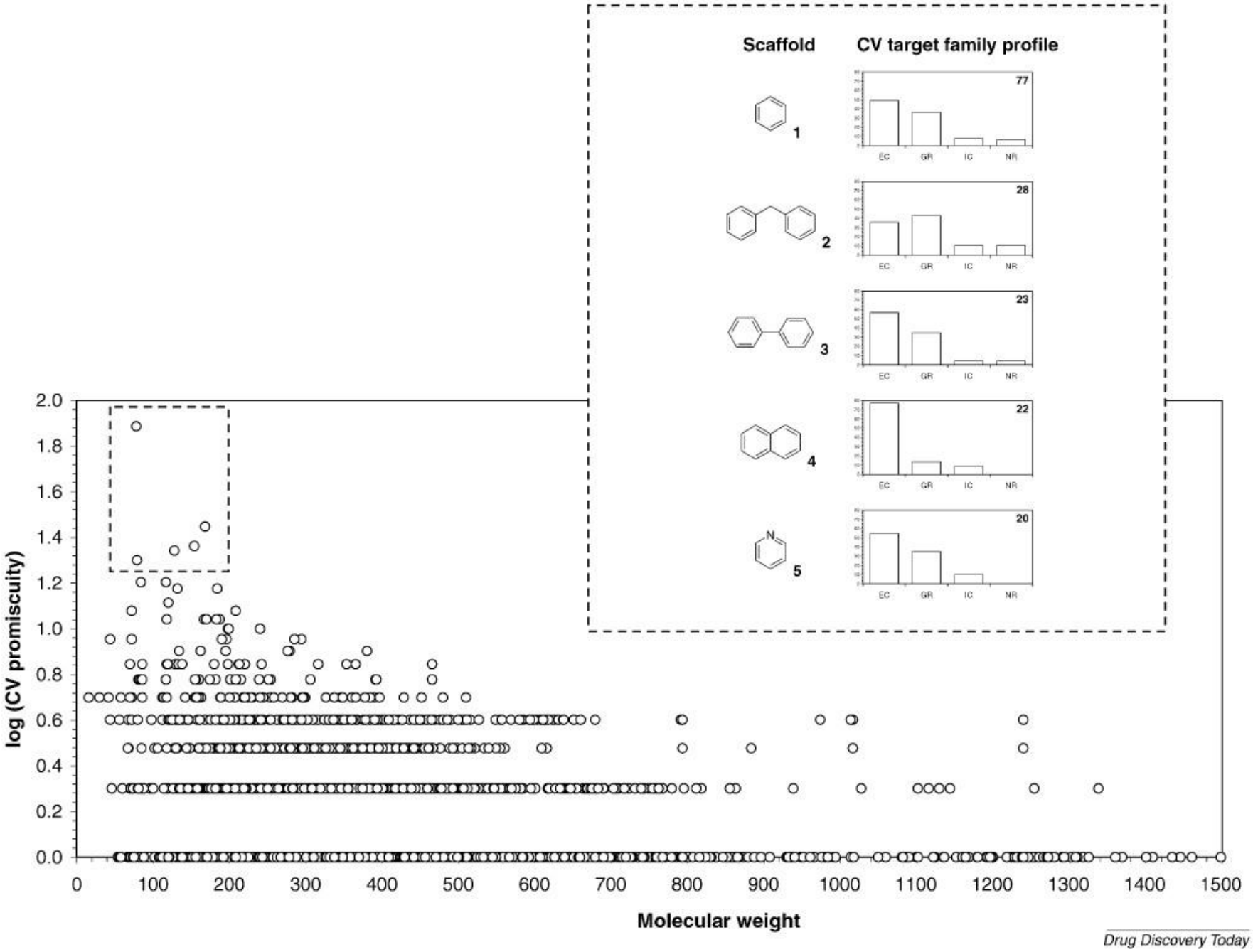


FIGURE 2

Structural and chemical coverage in cardiovascular target space. (a) Distribution of targets for which at least one representative structure exists in the Protein Data Bank and (b) distribution of targets for which at least one bioactive ligand is present in the annotated chemical libraries considered. EC: enzymes; GR: GPCRs; IC: ion channels; NR: nuclear receptors.

Example

5. Data from chemical libraries
   - Ligands with pharmacological potency at (at least one of) cardiovascular target structure
6. Identification of atomic frameworks or scaffolds
   - Characteristic, responsible for the aktivity, „privileged structures“
7. Synthesis
   - Synthesis of the hit compounds, the compound library
8. Biological evaluation
   - High-Throughput Scrreening, testing on cells or whole organisms



Drug Discovery Today

THANK YOU

FOR YOUR ATTENTION

References

- CACACE, E.: **Chemical genetics in drug discovery.** *Curr. Opin. Sys. Biol.* 2017, 4, 35-42
- CASES, M.: **A chemogenomic aproach to drug discovery: focus on cardiovascular diseases.** *Drug Discovery Today*, 2009, 14
- HARRIS, C.J.: **Chemogenomics: structuring the drug discoveryprocess to gene families.** *Drug Discovery Today* 2006, 11
- HORTON, D.A.: **The combinatorial synthesis of bicyclic privileged structures or privileged substructures.** *Chem. Rev.* 2003, 103, 893-930
- JACOBY, E.: **Chemogenomics: drug discovery panacea?** *Mol. BioSyst.* 2006, 2, 218-220
- KIM, J.: **Privileged structures:Efficient chemical „navigators“ toward unexplored biologically relevant chemical spaces.** *J. Amer. Chem. Soc.* 2014, 136, 14629-14638
- MARVANOVÁ, P.: **Příprava a studium derivátů arylkarbonyloxyaminopropanolů s *N*-fenylpiperazinovým strukturním fragmentem.** Disertační práce, FaF VFU Brno 2018
- SCHNEIDER, P.: **Privileged structres revisited.** *Angew. Chem. Int. Ed.* 2017, 56, 7971-7974
- WELSCH, M.E.: **Privileged scaffolds for library design and drug discovery.** *Curr Opin Chem Biol.* 2010, 14, 347–361