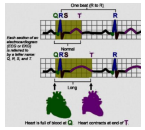
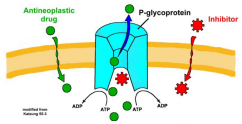
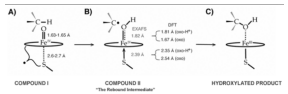
## Antitargets

**hERG** - **potassium ion channel** that **coordinates the heart's beating**. When **this channel is inhibited** by application of drugs **it can result in a potentially fatal disorder called long QT syndrome**; a number of clinically successful drugs in the market have had the tendency to inhibit hERG, and create a concomitant risk of **sudden death, as an unwanted side effect**, **hERG inhibition must be avoided during drug development**



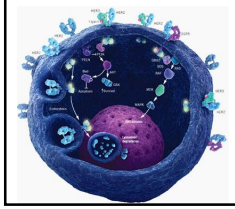



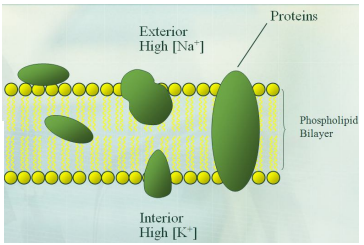

**P-glycoprotein** **transports substrates across the cell membrane, efflux pump for xenobiotics (e.g. drugs)** with broad substrate specificity. It is responsible for **multidrug-resistantance** and often mediates the development of resistance to anticancer drugs.

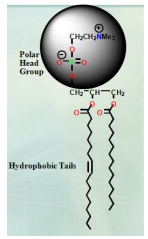
**Cytochrome P450** are the major enzymes involved in **metabolism** (~75%), they catalyze the oxidation of organic substrates, drugs included.

## Cell Membrane – protects cell compartment

- The cell membrane provides a **hydrophobic barrier** around the cell, **preventing a passage of water and polar molecules.** Proteins (receptors, ion channels and carrier proteins) are present, floating in the cell membrane.






## Bioavailability
### (PK - pharmacokinetic)

- *(in vitro)* **active compound, to perform as a drug, has to reach its target** in the human body (*in vivo*)
- **Drug-likeness** **is** qualitative **concept to estimate bioavailability from** the molecular **structure** **before the substance is synthesized.**

The **drug-like molecule** should have:

- ❑ **an optimal MW** and appropriate **number of HBD**, **HBA** (affecting solubility and absorption)
- ❑ **optimal water and fat solubility,** **partition coeficient logP** (octanol / water) to penetrate cellular membrane to rich target inside cells. The **distribution coefficient (Log D)** is the correct descriptor for ionisable systems. **logD is pH dependent** (e.g. pH = 7.4 is the physiological value of blood serum)

**Lipinski's Rule of Five (Ro5)**

## Lipinski Ro5
### (an empiric rule, all numbers are multiples of five)

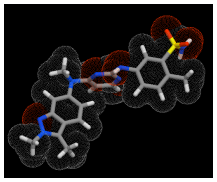
**for prediction of bioavailability** (**not activity**!) to quickly eliminate compounds that have poor physicochemical properties for an oral bioavailability

- an **orally active drug** has no more than one violation of the **following criteria**:
  - ❑ **MW** ≤ 500
  - ❑ **Lipophilicity** (logP ≤ 5) octanol-water partition coefficient (**better log D ≤ 5** respecting the ionic states present at physiological pH values)
  - ❑ Sum of **hydrogen bond donors** ≤ 5 (NH,OH)
  - ❑ Sum of **hydrogen bond acceptors** ≤ 10 (N,O)

C.A. Lipinski et al. Adv. Drug Del. Rev. **1997**, 23, 3. (Ro5)
G.M. Pearl et al., Mol. Pharmaceutics, **2007,** 4, 556–560. (log D introduced)

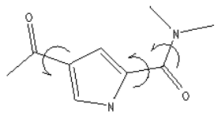
## Additional drug-like parameters

- **MW** ≤ 500
- **Lipophilicity** (logP ≤ 5) octanol-water partition coefficient
- Sum of **hydrogen bond donors** ≤ 5 (NH,OH)
- Sum of **hydrogen bond acceptors** ≤ 10 (N,O)



- **PSA < 140 Å²** (Molecular Polar Surface Area – sum of surfaces of polar atoms (N,O...with H) that correlates with human intestinal and BBB absorption), or (**PSA < 60 Å²**) for good BBB penetration

- **Number of rotatable bonds < 10**
  (high NRB → many conformers)



*Ertl, P. in Molecular Drug Properties, R. Mannhold (ed), Wiley-VCH , **2007**, 111 – 126.*

## Differences in drug-like selection criteria

- **Optimization often gives** drugs with **higher molecular weight**, **more rings**, **more rotatable bonds**, and a **higher lipophilicity.**

*T. I. Oprea et all. J. Chem. Inf. Comput. Sci.* 2001, ***41***, 1308–1315.

| | Ro3 | | Ro5 |
|---|---|---|---|
| **Fragment**[3] | | **Lead-Like**[2] | **Drug-Like**[1] |
| MW < 300 | | MW < 450 | MW < 500 |
| LogP < 3 | | LogP < 4.5 | LogP < 5 |
| Proton Donors < 3 | | Proton Donors < 5 | Proton Donors < 5 |
| Proton Acceptors < 6 | | Proton Acceptors < 10 | Proton Acceptors < 10 |
| Rotatable Bonds < 6 | | Rotatable Bonds < 10 | Rotatable Bonds < 10 |
| tPSA < 60 | | tPSA < 150 | tPSA < 150 |

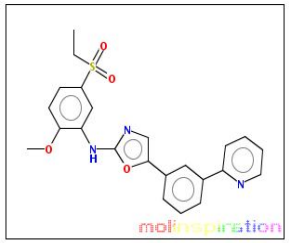
Others Properties, Solubility – calculated and measured, LogD, pKa

*1/ Lipinski, C.A et al. Adv. Drug Del. Rev. 1997, 23, 3. (Rule of 5)*
*2/ Verheij, H.J. Molecular Diversity 2006, 10, 377. (Lead-Likeness)*
*3/ Congreve, M. et al. Drug Discov. Today 2003, 8, 876. (Fragments)*

## Ro5 determined **from 2D tructure** http://www.molinspiration.com

molinspiration

SMILES CCS(=O)(=O)c4ccc(OC)c(Nc3ncc(c2cccc(c1ccccn1)c2)o3)c4



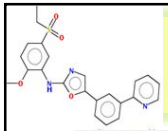
Molinspiration property engine v2009.01

| | |
|---|---|
| miLogP | 4.117 |
| TPSA | 94.327 |
| natoms | 31 |
| MW | 435.505 |
| nON | 7 |
| nOHNH | 1 |
| nviolations | 0 |
| nrotb | 7 |
| volume | 374.267 |

Get data as text (for copy / paste).

Get 3D geometry BETA

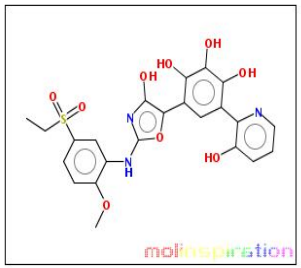
Ertl, P. et al., *J. Med. Chem.* 2000, 43: 3714-3717. (molecular property prediction toolkit )



## Ro5 violations

molinspiration

SMILES CCS(=O)(=O)c4ccc(OC)c(Nc3nc(O)c(c2cc(c1nccccc1O)c(O)c(O)c2O)o3)c4



Molinspiration property engine v2009.01

| | |
|---|---|
| miLogP | 2.904 |
| TPSA | 195.467 |
| natoms | 36 |
| MW | 515.5 |
| nON | 12 |
| nOHNH | 6 |
| nviolations | 3 |
| nrotb | 7 |
| volume | 414.356 |

Get data as text (for copy / paste).

Get 3D geometry BETA

## Absorption as f(PSA, LogP)

- **pKa** (influences binding Ki and logP) software SPARC http://www.archemcalc.com/
- **AlogP** (lipophilicity, water solubility) http://www.vcclab.org/ **(Virtual Computational Chemistry Laboratory)**

**Intestinal** and other **absorption**

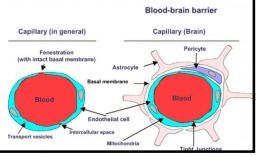
- **% ABS = 109 – 0.345 PSA**
  (good when % ABS > 30 %; **lower PSA, higher absorption**)

*Zao YH et al. Pharm Res 2002, 19, 1446-1457.*

**BBB absorption**

- **LogBB = -0.0148 PSA + 0.152 CLogP + 0.139**

CNS drug: logBB > -0.5 (otherwise side effects can be expected)
non CNS drugs: logBB < -1




## Other considerations

- **despite good druglikeness** some compounds should be avoided as drug candidates:
  - ❑ **substructures with known reactive, toxic, mutagenic or teratogenic properties** affect the usefulness (RCOX, $(RCO)_2O$, Michael acceptors, epoxides, $-NO_2$, -NO, $-N_3$, NH-NH, N=N...)
  - ❑ and with **bad metabolic parameters**, e.g. fast metabolism **can quickly destroy the pharmacological activity** of the compound
    (metabolic half life, metabolic clearance should be determined)

## From were to get the active compounds?

**A) The Natural World**

**Micro-organisms** (bacteria, fungi)
**Marine chemistry** (corals, bacteria, fish etc)
**Plant life** (flowers, trees, bushes)
**Animal life** (frogs, snakes, scorpions)
**Biochemicals** (neurotransmitters, hormones)

**B) The Synthetic World**

**Chemical synthesis**
(traditional, combinatorial synthesis, chemical collections, commercial sources)

**C) The Virtual World**

**Computer** aided drug design (CADD)

**to call them „active compounds" evaluation through biological screening is essential**

**ACTIVE compounds** can be obtained as hits in screening focused on selected biological target

- **Screenings** (in vitro: enzymatic, cellular or biophysical assays):
  - ❑ **High-throughput screening (HTS)** rapid screening of large numbers of compounds (up to 100 000/day)
  - ❑ **Biopysical screening** (NMR, PSR, X-ray screening)
- **Sources**
  - ❑ **LMW synthetic compounds** (collections from combinatorial chemistry / parallel synthesis, historical corporate collections...)
  - ❑ **Fragment-based** screening (not direct hit generation)
  - ❑ **Pure natural products, bioextracts** (e. g. plant or microbial) **ethnopharmacology** (Chinese traditional medicine...)
  - ❑ **SOSA approach** based on known Drugs/Clinical development compounds where side effects observed from *in vivo* screenings or human clinical trials

## SOSA = Selective Optimization of Side Activities

Exploitation of **Existing Drugs** as Leads **for the Discovery of New Drugs**

- **all drugs act on more than one target** (known and unknown), **resulting** in a several **side effects**
- advantage: drugs and many compounds that underwent clinical development **have an established safety profile.** Many of them **can be** therefore safely **administered to humans**.

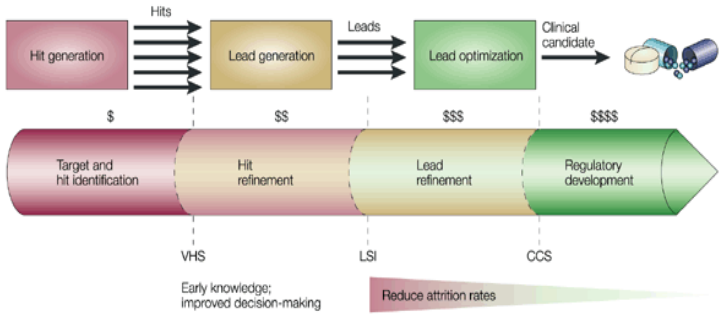
**Try to transform one of the side activities into the major effect** and strongly reduce their initial pharmacological activity.

## How many new drugs reach the market yearly?

- **DrugDiscovery:**
  **10 years / from 1 to 2 000 000 000 USD /1 new drug**
- **global production ca 24 innovative drugs** (**possessing new chemical entity**) **/ year**

## DD - flowchart

**Target** selection **=> active => hit => lead => clinical candidate => drug**




## Active → Hit → Lead → Drug candidate → Drug

**Actives: are** indentified **compounds** with a desired target bioactivity (e.g. by HTS or biophysical methods (SPR, ITC – Isothermal Titration Calorimetry, NMR)).

**A2H process:** Validated **Hits** are **stable active** (< 3 mM in biochemical assay, < 10 mM in vivo assay) **small molecules** with **determined purities, confirmed structures,** and specific $IC_{50}$ **target activities.** **The aim of A2H process is to determine appropriate active compounds possessing diverse chemical structures for further development.**

**H2L process: Lead compounds are identified from validated hits**. **The aim of H2L process is to exclude inappropriate compounds that could fail in subsequent preclinical and clinical trials early in DD, before significant resources are spent.**

**OPTIMAL LEADS:** have **good target and cell activities, selectivities** (e.g. > 10-fold over related targets), **ADME/Tox properties**: **bioavailability** (lead-likeness, aq **solubility > 100 uM** (e.g. at FW: 500 g/mol → > 1 mg / 20 mL), logP, logD, pKa, plasma albumine binding), **metabolic stability** (not metabolized too quickly) and **low toxicity** (no **undesirable chem functionalities** like nitros, Michael acceptors..., low **antitargets activities**: hERG (30-fold selectivity over hERG), P-glykoprotein, CYP450 (> 1 mM for 4 of 5 major human isoforms), low **cytotoxicity** (30-fold selectivity over chronic (24h) cellular toxicity), low **genotoxicity**...). Leads have **novel patentable structures**. They are **synthetically accessible** (e.g. parallel (convergent) synthesis) and they have **optimization potential**. Focused libraries around the most promising hits (selection of 3-6 structure clusters) are prepared for early and rapid generation of structure activity relationship (SAR) data in order to indentify highly active and selective leads (< 100 nM in vitro assay, < 1 mM in vivo assay) for further DD development.

| | HIT | LEAD |
|---|---|---|
| *In vitro assay* | < **3000 uM** (3 mM) | < **0.1 uM** (100 nM) |
| *In vivo assay* | < 10 mM | < 1 mM |

**LEAD-LIKE COMPOUNDS**

physicochemical **properties:** MW < 450, logD < 4, H-bond donors (NH,OH) ≤ 4, H-bond acceptors ≤ 8 (N,O)

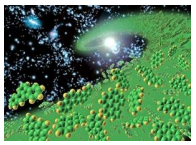
**DRUG CANDIDATE** is a **result of further leads development** by *in vivo* assays and clinical trials PhI-III confirming activity and low toxicity on patients to show their clinical benefit and better properties compare to similar marketed drugs.

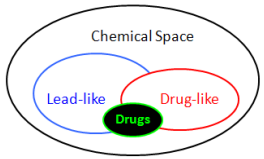
# Case story from HTS to Leads

- at **Schering** begins with
  - **HTS assay of 700 000** compounds. Afterwards they **repeated HTS** with selected **2 000** compounds. This reduces the compound pool sending **200 compounds forward to $IC_{50}$ assaying**. The result was **100 active compounds with determined $IC_{50}$ values**. Those 100 compounds went for **purity and structure evaluation** bring the number down to **50 validated actives (A)** in about one month (enrichment: 1/14 000).
  - Subsequently **in vitro efficacy, selectivity, and toxicology studies produced 15 compounds as "qualified hits„ (A2H)** by the end of the third month (enrichment: 1/46 667).
  - **Qualified hits were resynthesized to yield more compound for in vitro and in vivo evaluations**. These evaluations concluded with the **identification of 13 lead structures** after 10-months.
  - After 14 months of the above selection process final **13 compounds** (total enrichment: 13 / 700 000 = 1 / 53 846) **were moved** for the **lead optimization process**.

# Summary: what should the compound fulfill to become a drug?

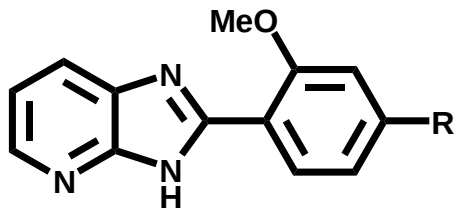
- **Biologically active,** chemically **stable compound** possessing appropriate:
  - **pharmacodynamic properties** (target activity and selectivity)
  - **pharmacokinetic properties** (bioavailability: ADME/TOX)
  - **other properties** (novelty, synthetic feasibility, scale up synthesis... )
- **> 30% of all drug failures can be attributed to poor physiochemical properties**: Log P (Log D), pKa, and solubility with **impact on drug absorption and diffusion in vivo**






**Chem Space:** $10^{60}$ - $10^{200}$
DB of 11 atoms C,N,O, F: 26 400 000 of stable compounds (111 000 000 cmpds if included all stereoisomers) *J.-L. Reimond*

# Case study – cardiotonic agent
## (optimization of physico-chemical properties)

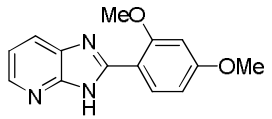



*Cardiotonicum*

R: MeO- log P (2.59) FW: 255.27 (**CNS side effects bright visions**)
MeSO- lop P (1.17) FW: 287.34 (**Sulmazol**)

**LogP round 2.5 allows compound to penetrate to CNS → side effects**

2-Aryl-3H-imidazo[4,5-b]pyridine

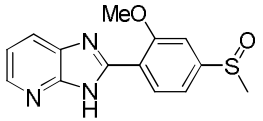


www.MOLINSPIRATION.com

COC1=CC(OC)=CC=C1C2=NC3=CC=CN=C3N2

Molinspiration property engine v2014.11

| | |
|---|---|
| miLogP | 2.69 |
| TPSA | 60.04 |
| natoms | 19 |
| MW | 255.28 |
| nON | 5 |
| nOHNH | 1 |
| nviolations | 0 |
| nrotb | 3 |
| volume | 227.21 |



CS(C(C=C1OC)=CC=C1C2=NC3=CC=CN=C3N2)=O

Molinspiration property engine v2014.11

| | |
|---|---|
| miLogP | 1.47 |
| TPSA | 67.88 |
| natoms | 20 |
| MW | 287.34 |
| nON | 5 |
| nOHNH | 1 |
| nviolations | 0 |
| nrotb | 3 |
| volume | 243.90 |

## pKa a universal measure for both acidity and basicity

*Definition of $pK_a$*

HA $\rightleftharpoons$ $H^+ + A^-$ ionised forms

or $HB^+ \rightleftharpoons H^+ + B$

$pK_a = -\log_{10} ([H^+][A^-]/[HA])$

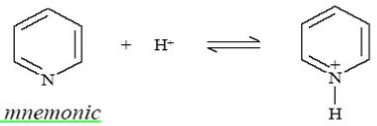
**pKa = $-\log_{10}$ ($[H^+].[B]/[HB^+]$)**

**the higher pKa the stronger base** (aniline 4.6, py 5.1, methylamine 10.6)
(or the weaker acid: HBr -9.0, HCl -8.0, $H_2SO_4$ -3.0, $H_3O^+$ -1.7, HF +3.2, HOAc +4.8)

**pKa = -log {$[H^+].[B] / [HB^+]$} = pH - log $[B]/[HB^+]$ = pH – log 1 = pH**

**Conclusion: if concentration of base and its protonated form is equal than pKa = pH (it is special pH at which the base is protonated on 50 %). Such values can be compared within different bases to estimate their basicity.**

We consider the basicity of pyridine in terms of the $pK_a$ of the conjugate acid - the pyridinium cation



$pK_a = 5.1$

*Useful mnemonic*

The $pK_a$ designates that pH where the two species are present at equal concentration

i.e. at pH = 5.1 pyridine is 50% protonated

at more basic conditions at pH = 6.1 pyridine is 10% protonated

in more acidic environm. at pH = 4.1 pyridine is 90% protonated

**pKa = -log {$[H^+].[B] / [HB^+]$} = pH - log $[B]/[HB^+]$**

---

**5.1 = 4.1 - log $[B]/[HB^+]$**
**[B] + $[HB^+]$ = 100%**

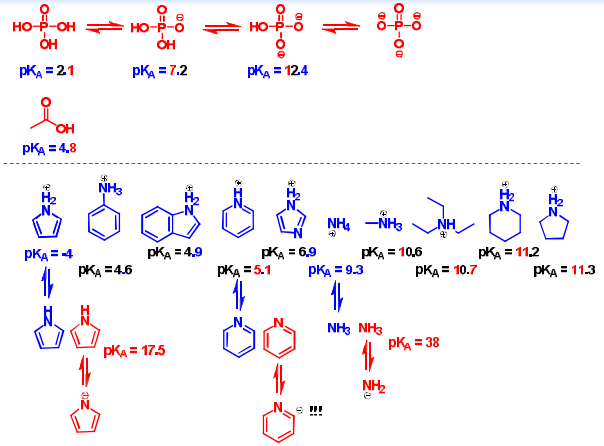
two equations, two variables

---

**9% [B] and 91% $[HB^+]$**

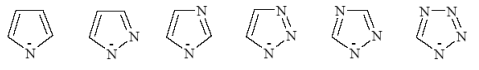
Aká ionizácia pyridínu bude v krvi pH 7.4? (len 0.5%)

## The higher pKa the stronger base (or the weaker acid)



**Scheme 14. Basicities of Nitrogen in Heteroaromatic Five-Membered Rings ($pK_a$ values)**

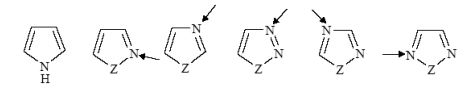
*A. Anionic Nitrogen : Azole Anions or Acidity of Neutral Azoles*



16.5 14.2 14.4 9.3 10.0 4.9

with increasing number of nitrogen atoms basicity decreases and acidity of conjugate acid (the neutral azole) increases

*B. Neutral Nitrogen: Azoles (Acidity of Azole Cations)*



| (at N) | Z = | 2.5 | 6.9 | 1.2 | 2.5 | 1.2 |
|---|---|---|---|---|---|---|
| ca -4 | | -3 | 0.8 | - | - | - |
| | | -0.5 | 2.5 | - | - | -4.9 |

Note: (i) Pyrrole very weakly basic because no pyridine-like nitrogen

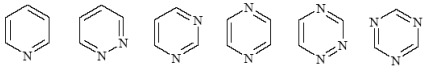
(ii) Imidazole much more basic than pyrazole because of base weakening interaction of adjacent N

(iii) S- and especially O- substitution for NH dramatically reduces basicity

(iv) Additional pyridine-like N atoms also strongly reduce basicity

**Scheme 15. Basicities of Six-Membered Heteroaromatics and Non-Heteroaromatics**

*A. Basicities of Azines ($pK_a$ values)*

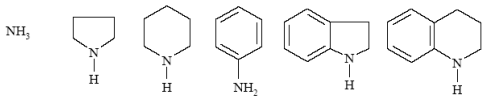


5.1 pyridine | 2.3 pyridazine | 1.3 pyrimidine | 0.6 pyrazine | < 0 1,2,4-triazine | < 0 1,3,5-triazine

Note: (i) Extra N-atoms considerably reduce basicity of pyridine (inductive effect)

(ii) Pyridazine more basic than pyrazine because of the effect of adjacent electron lone pairs (α-effect)

*B. Basicities of Some Other Neutral Nitrogen (Non-heteroaromatic) Compounds*



9.3 ammonia | 11.3 pyrrolidine | 11.2 piperidine | 4.6 aniline | 4.9 indoline | 5.0 tetrahydroquinoline

Note: (i) Most aliphatic amines have $pK_a$ around 10 - 11; (ii) Most anilines have $pK_a$ around 4 - 5

# Rational drug design

frequently relies on **computer modeling techniques**
(*computer-aided drug design* ***CADD***)

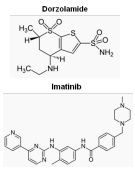
**AIM:**

- **to discover**, or enhance **molecules with ability to bind to a selected target** and **to estimate the power of** binding **before compounds are synthesized**
- to **estimate drug-like properties** and use them for elimination of undesirable structures

**it still takes several iterations** of design, synthesis, and testing **before an optimal molecule is discovered**

# Rational methods in DD

- **Structure-based drug design SBDD**
  - requires 3D information about target (X-ray crystallography or NMR spectroscopy, PDB database) http://www.rcsb.org/pdb/




  - first example **Dorzolamid** Carbonic anhydrase inhibitor: Greer J, et al. *JMCH* 1994, 37, 1035–54. (Merck 1995)
  - **Imatinib** (Gleevec , Novartis 2001) the first tyrosine kinase inhibitor designed for the *bcr-abl* fusion protein (Philadelphia chromosome-positive receptor in chronic myelogenous leukemia (CML)
- **Ligand-based drug design LBDD**
- **Fragment-based drug design FBDD**

## Structure Based Drug Design SBDD (direct DD)

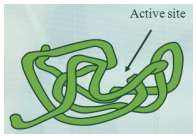
- based on **knowledge of 3D structure of the biological target** obtained through **X-ray crystallography or NMR spectroscopy** (http://www.rcsb.org/pdb/)

- SBDD can be divided into two methodologies:
  - ❑ **"finding" ligands for a given receptor** (database searching / **virtual screening**). A large number of potential ligand structures are screened to find those fitting the binding pocket of the receptor. **It saves synthetic effort** to obtain new active compounds.

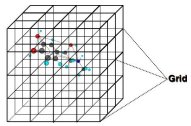
  - ❑ **"building" ligands in a target active place**. In this case, **ligand molecules are built up** within the constraints of the binding pocket **by assembling small pieces (atoms, fragments) in a stepwise manner.** The key advantage: **novel structures, not contained in any database, can be suggested**.

## „Finding ligand" methodology

- relies on **known target structure** (PDB complex)
- active side identification
  - ❑ protein, ligand atoms and virtual grid spots need to be classified by their atomic properties as
    - **hydrophobic atom**: all carbons
    - **H-bond donor**: OH,NH
    - **H-bond acceptor**: O,N
    - **polar atom**: O,N,S,P,X,M,C-HETATOM

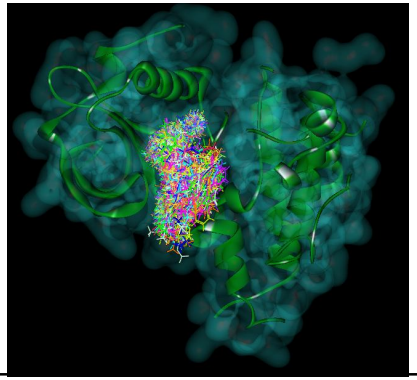



**The space** inside the ligand binding region **would be studied with virtual probe atoms** of the four types above **to determine what kind of chemical fragments can be put into their corresponding spots in the ligand binding region of the receptor**.




## Virtual SBDD screening („finding ligands")

- 500 potentially actives / 1 500 000 mol. in library

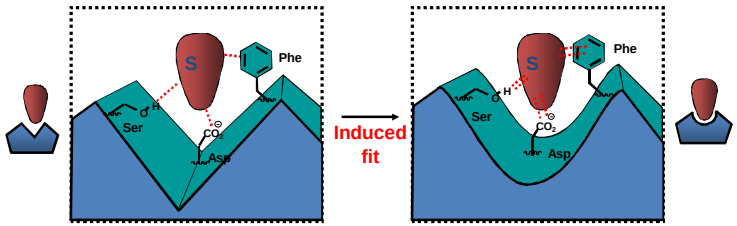


1 : 3000

## Ligand-based induced fit

(different conformers of the same protein)

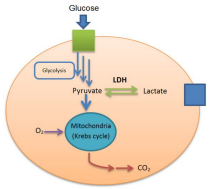
- **Binding of ligand alters the shape of the protein active site to maximise intermolecular interactions. This will result 3D changes in protein binding site => induced fit**

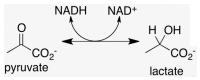



**Intermolecular bonds not optimum length for maximum bonding**

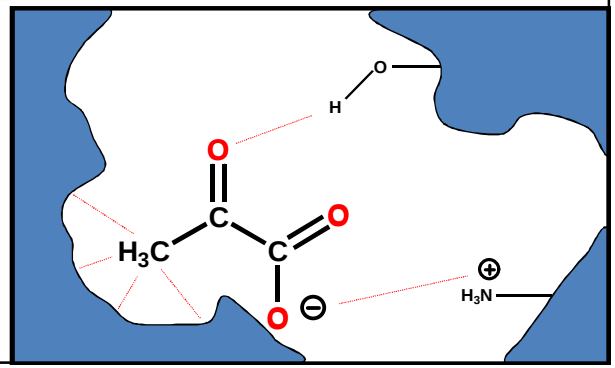
**Intermolecular bond lengths optimized** (conformational changes in AA residues and in backbone observed in the protein)

## Example: Binding of pyruvic acid in Lactate Dehydrogenase LDH

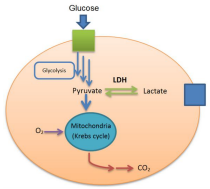


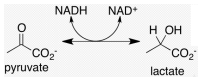




LDH converts pyruvate, the final product of glycolysis, to lactate when oxygen is absent or in short supply

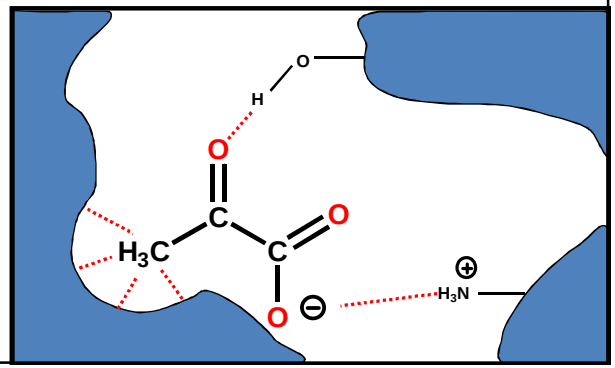



## Example: Binding of pyruvic acid in Lactate Dehydrogenase LDH




LDH converts pyruvate, the final product of glycolysis, to lactate when oxygen is absent or in short supply

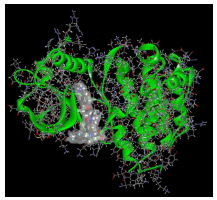



## VEGFR2 TK – induced fit

(different 20 conformers of the same protein differently accommodating the same set of compounds, **induced fit is a complication for CADD**)

| PDB code | The best binder kcal/mol |
|---|---|
| 1Y6B | -35.75 |
| 3C7Q | -36.03 |
| 2RL5 | -35.20 |
| 3B8Q | -26.80 |
| 2QU5 | +3.46 |
| 2P2I | -30.65 |
| 1YWN | -33.60 |
| 3B8R | -36.60 |
| 3DTW | -30.75 |
| 3CP9 | -28.27 |
| 3CPB | -18.01 |
| 3EFL | -35.40 |
| 3EWH | -27.10 |
| 3CJF | -46.96 |

| PDB code | The best binder kcal/mol |
|---|---|
| 3CPC | -27.23 |
| 2OH4 | -31.38 |
| 1Y6A | -38.11 |
| 2P2H | -36.12 |
| 2QU6 | -37.69 |
| 3CJG | -45.16 |



**nM inhibitors should rich the relative level of interaction energy ca -50 kcal/mol. 3CJF and 3CJG are only receptor conformers that almost give this level for the best of a set of 16 docked compounds.**

**X-ray Structure Screening is overcoming induced fit problems**

**Procedure:**

- **crystallize** target protein with your ligand
  (e.g. receptor + inhibitor)
- **acquire 3D structure** of complex by X-ray crystallography
- **identify a binding site** (region where ligand is bound)
- identify **binding interactions** between ligand and target
- identify **vacant regions** for extra binding interactions
- **'Fit' analogues** into binding site to test binding capability

Carry out **drug design based on the more accurate interactions** between your lead compound and the target binding site.

## Ligand Based Drug Design
## LBDD (indirect DD)

- based on **knowledge of molecules that bind to the biological target** (their structure and $IC_{50}$ bioactivity)

  - ❑ These molecules (ligands) may be used **to derive a pharmacophoric model which defines** the minimum necessary **structural characteristics** a molecule must possess in order to bind to the target.

    **Virtual screening** (based on pharmacophore models; high-throughput docking) including drug **property filtering** (Zinc 35 000 000 / Lipinski)

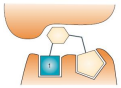
    http://zinc.docking.org/

  - ❑ Alternatively, a **quantitative structure-activity relationship (QSAR)** in which a **correlation between calculated properties** of molecules **and** their **experimentally determined biological activity** may be derived. These may be used **to predict the activity of new analogues**.

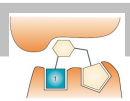
## Fragment Based Drug Design
## FBDD

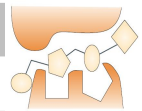
- Screening of **small** (MW < 300), **low potency fragments** (**epitops**, "seed templates"), which are **subsequently developed into higher potency structures**

  - ❑we need a **database** of **fragments** to choose ligands
  - ❑although the diversity of organic structures is infinite, **the number of basic fragments is rather limited**
  - ❑seed is put into the binding pocket, and **add other fragments one by one**
  - ❑**new molecules can be regarded as combinations of two or more individual binding epitopes**



| | FBDD | HTS |
|---|---|---|
| Output | Efficiency of binding | Potency in activity |
| Screened compounds | few 100 to several 1000 | several 100 000 |
| MW | 150-300 | 250-600 |
| Activity threshold | mM to 30 uM | 30 uM to nM |
| Screening based on | Biophysical assays (NMR, X-RAY, SPR), direct measurements of ligand-protein interactions (direct measurements of inactive forms of kinases) | In vitro assay not ideal compounds available some functional groups not involved in interactions |
| Hit to Lead | synthesis of only few designed compounds | more extensive synthesis |
| Requirements | Expertise in protein-ligand binding interactions and fragment design | Requires intensive infrastructure, data processing |





## Ligand efficiency / Ro3

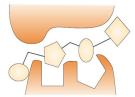
- **Ligand Efficiency LE**

$$LE = -\Delta G/HAC \approx -RT\ln(IC_{50})/HAC$$

*(HAC = Number of heavy atoms)*

fragments typically exhibit higher ligand efficiency than higher MW compounds identified through HTS (not ideal interactions)

*I. D. Kuntz et al. Proc. Natl. Acad. Sci. USA **1999**, 96, 9997.*

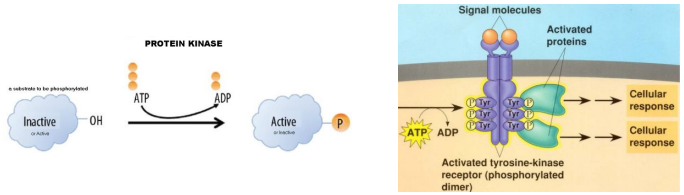


- The "**rule of three, Ro3**" for fragment design:

| TABLE 1 | Rule of Three Criteria[1] |
|---|---|
| MW | ≤ 300 |
| cLogP | ≤ 3.0 |
| H-Bond Acceptors | ≤ 3 |
| H-Bond Donors | ≤ 3 |
| Rotatable bonds (Flexibility Index) | ≤ 3 |
| Polar Surface Area | ≤ 60Å² |

# Protein Kinase Inhibitors

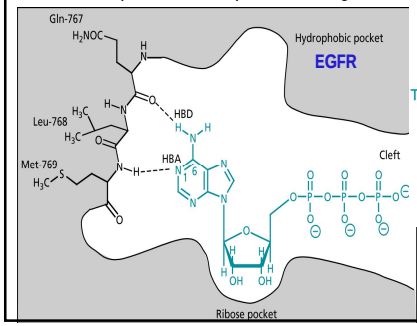
**Protein kinases (PKs)** (tyrosine, serin-threonine and histidine kinases) **phosphorylate** specific amino acids in **protein substrates**.

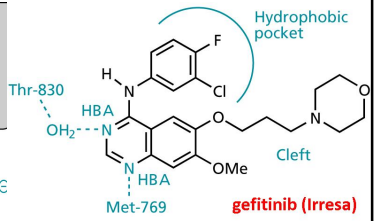



There are **over 500 different types of hu-protein kinases**. Many PKs are **enzymes (TK)** within cytoplasm, others traverse the cell membrane **and** play **dual role as receptor and enzyme (TKR)**. **Growth factors through TKRs signalling control transcription of genes leading to cell division. In many cancers excess of growth factor or PK receptor has been observed.** Therefore **PK inhibitors are useful anticancer agents.** All PK use ATP as the phosphorylation agent.

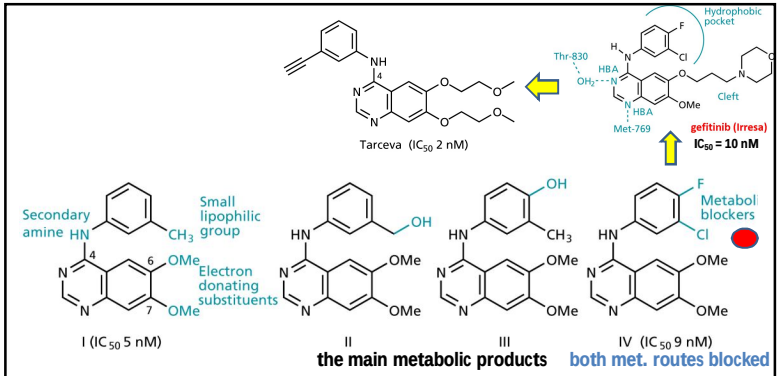
# EGFRI gefinitib (Irresa)
## (AstraZeneca, FDA 2003)

**EGFR** (ErbB, HER1) **tyrosine kinase receptor**: **abnormal or over-expressed in** the breast, lung, brain, prostate, gastrointestinal tract, ovaries **cancer**. EGFR is a receptor for EGF growth factors (Nobel Prize 1986). **Upon activation** by EGF, **EGFR forms active homodimer possessing intracellular TK activity** that initiate several signal transduction cascades **leading to DNA synthesis and cell proliferation**. **Gefitinib inhibits EGFR by binding to the ATP-binding site**. Thus this receptor and its dependent **malignant cells are inhibited**.




EGFR positive patients have shown an impressive **60% response rate** which exceeds the response rate for conventional chemotherapy.

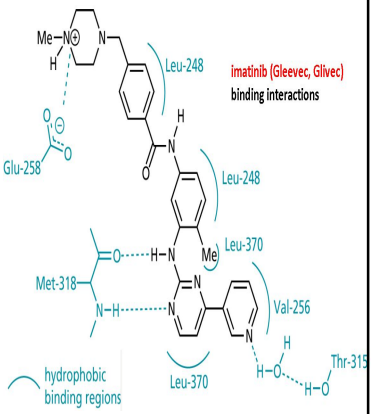



Found 4-anilinoquinazoline was optimized (SAR). Lead I, good in vitro activity, **in vivo hampered by rapid metabolism** by cytochrome P450 enzymes

**Metabolc blokators:**
Cl- similar in size and lipophilicity as Me- group
F- almost the same size as H- (no steric effect)
both groups are **resistant to oxidation** => better in vivo activity
**Pharmacokinetic properties:** improved by morpholino group, because its basic nitrogen is protonated in water and enhances drug solubility.

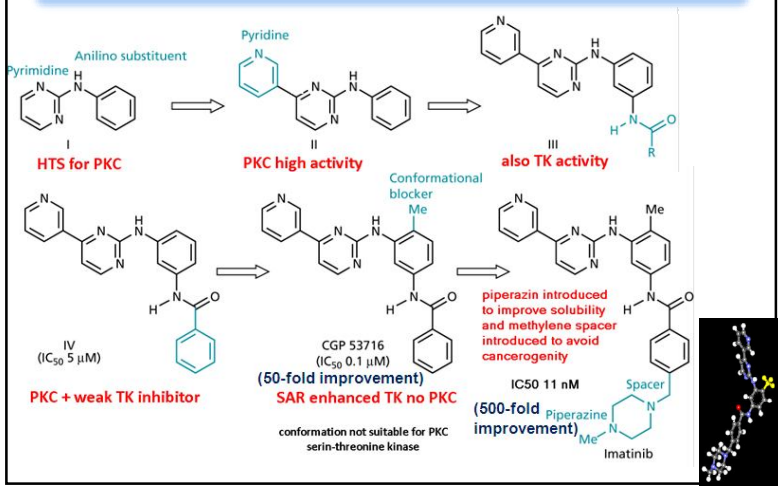
# TKI imatinib (Gleevec)
## (Novartis, FDA 2001)

**The first PKI that reach the marked.**

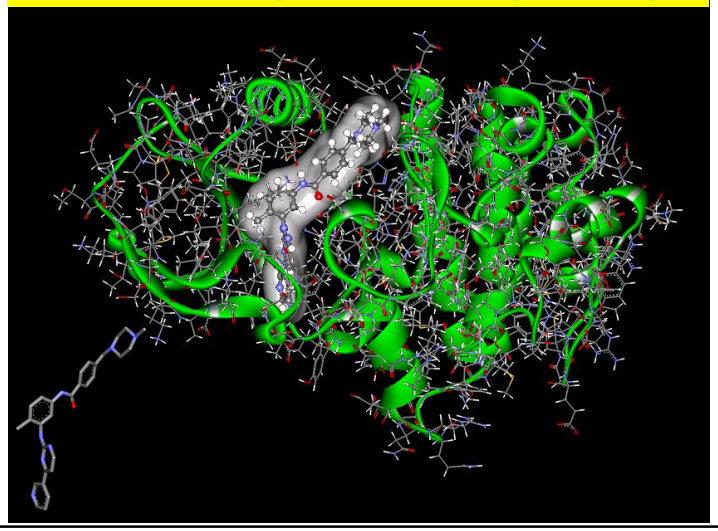



**Imatinib** treats **CML** (Chronic Myeloid Leukemia) a cancer of white blood cells characterized by the **increased and unregulated growth of myeloid cells** in the bone marrow and the accumulation of these cells in the blood. These **cancer cells contain an abnormal heterodimeric protein kinase (Bcr-Abl)** that is **not found in normal cells**. It is **associated with Philadelphia chromosome**. TK active site is on Abl portion of Bcr-Abl receptor. **Imatinib is selective inhibitor successful in 90% of patients.** This is a **first drug targeting unique mol. structure observed only in cancer cells**. Treatment of CML by imatinib dramatically improved patient five year survival from 31% to 59%.

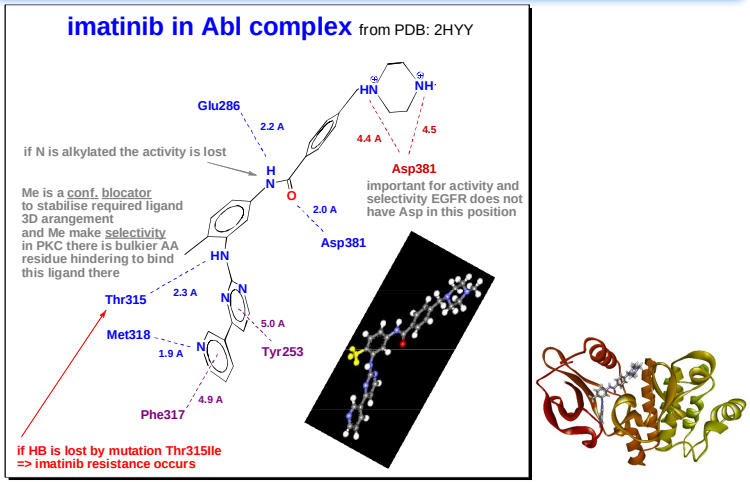
# H2L optimization in imatinib DD




## Abl kinase in complex with imatinib (PDB: 2HYY)



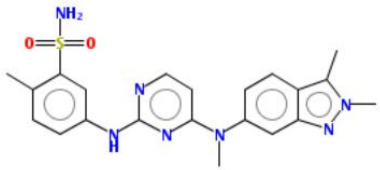
# Binding Interactions Map






# pazopanib (VOTRIENT)
(GSK, FDA 2009)

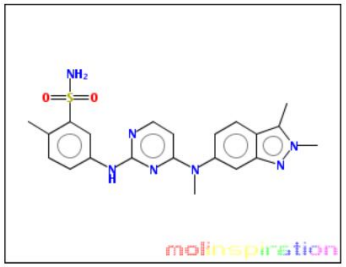
- multi-targeted tyrosine kinase receptor inhibitor (**PDGFR-α,β, VEGFR-1,2,3, KIT**, for **renal cell carcinoma** (RCC) and **imatinib-resistant gastrointestinal stromal tumor** (GIST)



5-(4-((2,3-dimethyl-2H-**indazol**-6-yl)(methyl)amino)**pyrimidin**-2-ylamino)-2-methylbenzenesulfonamide

*JMCH* 51 **2008** 4632

# pazopanib / drug-like properties



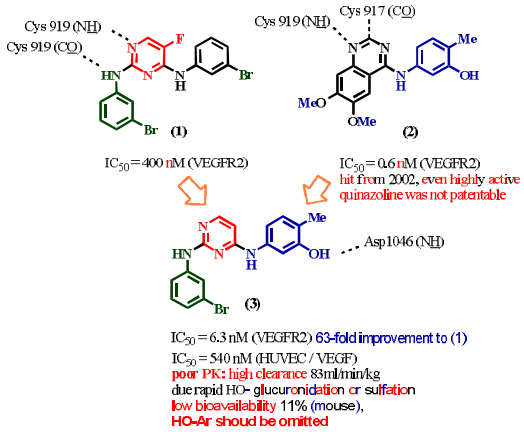

Molinspiration property engine v2009.01

| | |
|---|---|
| miLogP | 3.272 |
| TPSA | 119.04 |
| natoms | 31 |
| MW | 437.529 |
| nON | 9 |
| nOHNH | 3 |
| nviolations | 0 |
| nrotb | 5 |
| volume | 377.901 |

Get data as text (for copy / paste).

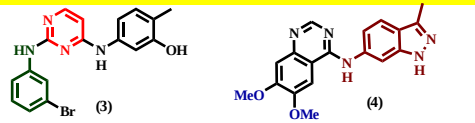
Get 3D geometry BETA

# 1st HTS GSK / SBDD optimization



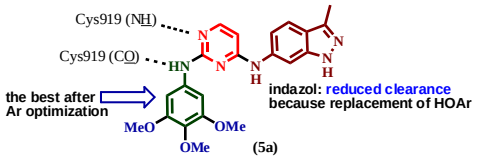

pyrimidine 1 and chinazoline 2 (both bind similarly in ATP binding pocket) → pyrimidine 3

# 2nd screening – house kinase library, hit (4)




IC50 = 6.3 nM (VEGFR2)
IC50 = 540 nM (HUVEC/VEGF)

IC50 = 1.7 nM (VEGFR2)
excellent, but chinazoline not IP

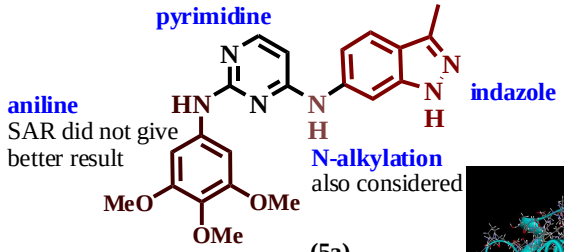


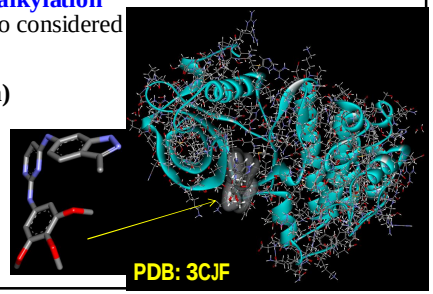

IC50 = 6.3 nM (VEGFR2) **good afinity**
IC50 = 180 nM (HUVEC/VEGF) **improved cellular act.**
**PK improved significantly:** clearance 16 ml/min/kg
oral bioavailability 85% (rat)

pyrimidine 3 and chinazoline 4 → pyrimidine 5a

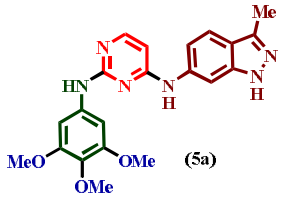
# Further SAR optimization of 2nd screening result (5a)

*(for better affinity, activity, PK properties for oral therapy)*



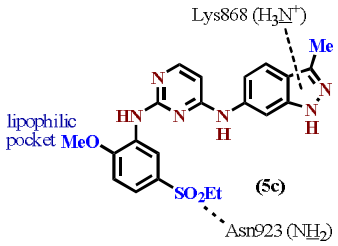



## SAR → optimization on an aniline fragment



(5a)

$IC_{50}$ = 6.3 nM (VEGFR2)
$IC_{50}$ = 180 nM (HUVEC/VEGF)
**PK:** clearance 16 ml/min/kg
oral bioavailability 85% (rat)

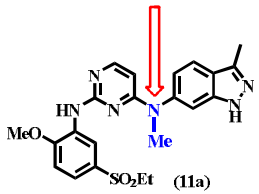



(5c)

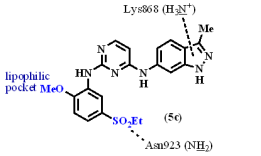
**$IC_{50}$ = 130 nM (VEGFR2)**
**$IC_{50}$ = 13 000 nM (HUVEC/VEGF)**
**moderate clearance 40 ml/min/kg (rat)**
**moderate bioavailability (28 %)**

## SAR → via N-alkylation (conf. stabilization)

NH could formed HBD (worster absorption) and potential place of metabolic changes
NMe stabilizes S conformer, better PK properties



(11a)

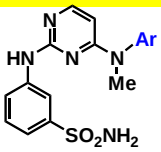



$IC_{50}$ = 130 nM (VEGFR2)
$IC_{50}$ = 13 000 nM (HUVEC/VEGF)
moderate clearance 40 ml/min/kg (rat)
moderate bioavailability (28 %)

**affinity comparable with (5c, NH)**
**but better PK properties: clearance 10ml/min/kg (from 40), bioavailability 65 % (from 28 %)**

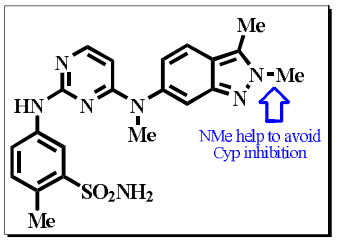
**$IC_{50}$ < 10 uM (CYP P450)**
**N from indazole coordinates hem iron in P450**

## SAR → optimization on a heterocycle to avoid CYP binding



| Cpd | Ar | VEGFR2[a] $IC_{50}$ nM | HUVEC[b] $IC_{50}$ µM | Inhibition of P450 Isozymes $IC_{50}$ µM[c] | | | |
|---|---|---|---|---|---|---|---|
| | | | | 2C9 | 2C19 | 2D6 | 3A4[d] |
| 12a |  | 0.6 | 0.094 | 0.7 | 7.2 | 4.8 | 7.6 / 16 |
| 12b |  | 36 | 1.4 | 1 | 9.3 | 6.3 | 13 / 1.7 |
| 12c |  | 5.6 | 0.023 | 1.4 | 20 | 16 | 16 / 31 |
| 12d |  | 2.6 | 11 | 0.5 | 1.4 | 66 | 74 / 79 |
| 12e |  | 7.6 | 0.11 | 2.9 | 9.2 | 18 | 4.7 / 3.9 |
| 12f |  | 63 | 2.8 | 15 | 26 | 25 | 26 / 1.7 |
| 12g |  | 17 | 1.4 | 9 | 33 | 33 | 27 / 45 |

## pazopanib (Votrient, GSK)




$IC_{50}$ = 10, 30, 47 (VEGFR1-3) good affinity
$IC_{50}$ = 21 nM (HUVEC / VEGF) high activity
$IC_{50}$ = 720 nM (HUVEC / bFGFR) good selectivity
$IC_{50}$ (CYP P450) > 10 000 nM low antitarget inhib.
very good PK properties: low clearance 1.7 ml/min/kg
good oral bioavailability (72 %) at dose 10 mg / kg
in vivo and clinical Ph I - III trials succesfull
FDA approved 19 Oct 2009 as
globally 3rd antiangiogenic drug to treat cancer