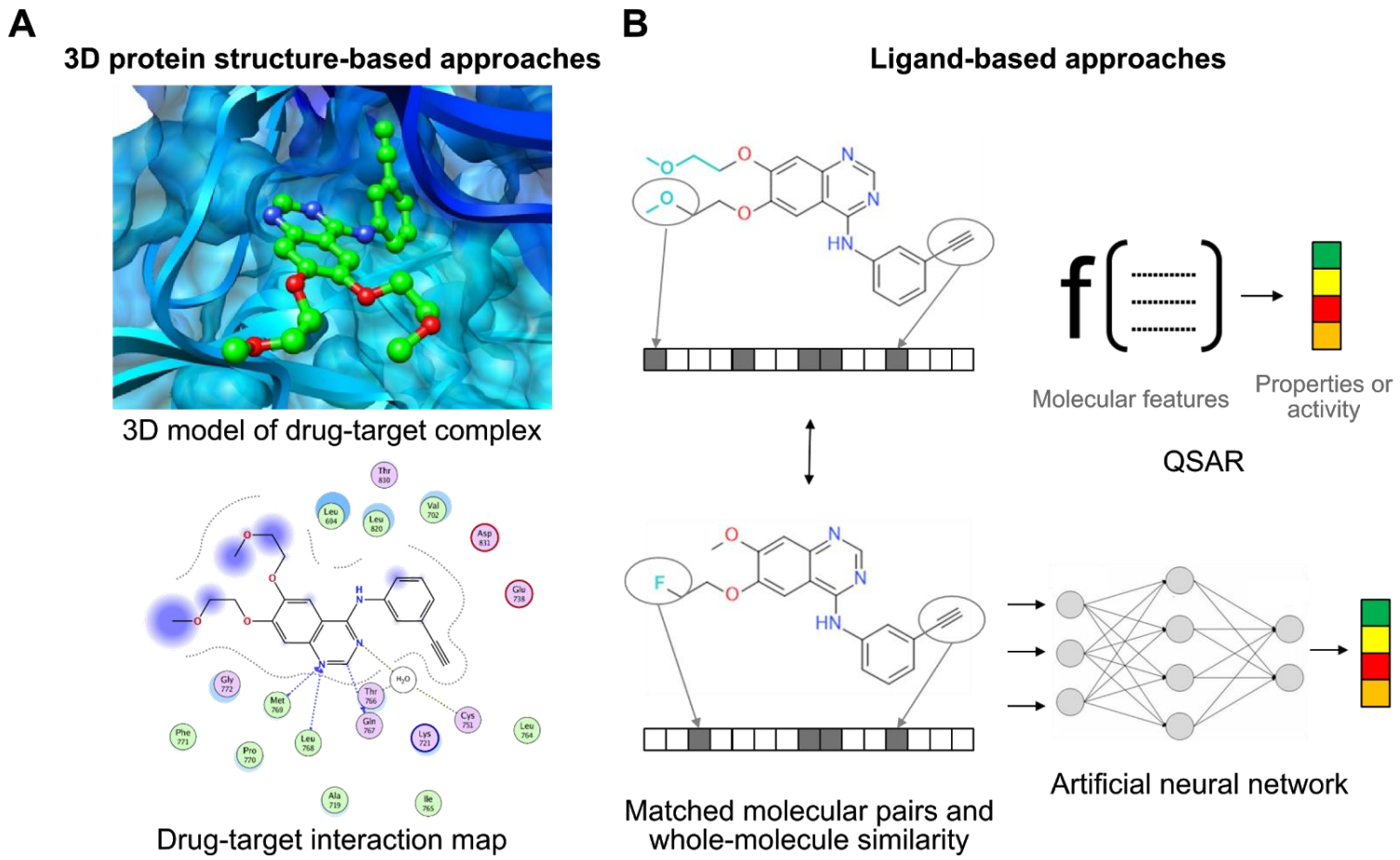
AMIDD Lecture 6: Molecular modelling



Overview of molecular-level modelling techniques: (A) 3D protein structure-based approaches (B) Ligand-based approaches

Multiscale Modelling of Drug Mechanism and Safety by Zhang, Sach-Peltason, Kramer, Wang and Ebeling, in revision

Dr. Jitao David Zhang, Computational Biologist

[1] Pharmaceutical Sciences, Pharma Research and Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche

[2] Department of Mathematics and Informatics, University of Basel

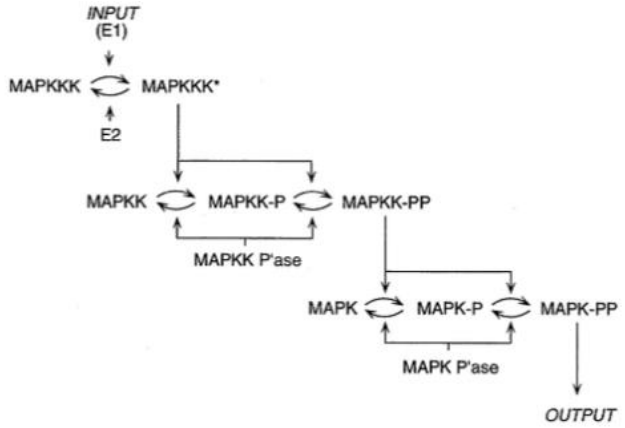
Contact the author

Selected mathematical concepts

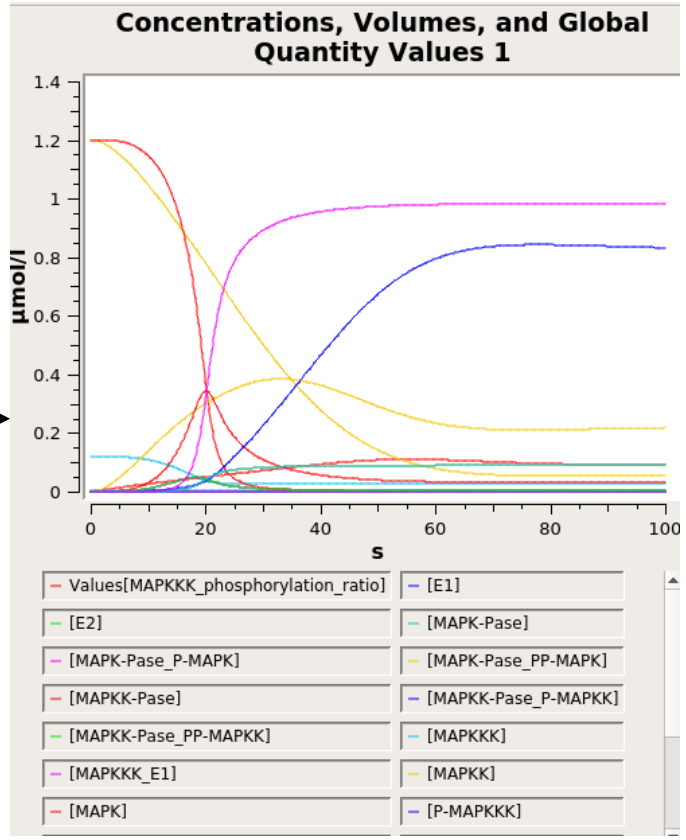
- **Affinity**
  - The (bio)physical view
  - The (bio)chemical view
- **The Michaelis–Menton model and enzymatic kinetics**
- Mathematical techniques for structure determination: X-ray, NMR, and CryoEM (mainly post-reading)
- **Example of structure-based drug design: molecular docking**
- **Example of ligand-based drug design: similarity and quantitative structure-activity relationship (QSAR)**

Biochemical system simulator COPASI

- Freely available at http://COPASI.org/
- COPASI supports two types of simulation
  - **Differential-equation (ODE) based simuation**
  - **Stochastic kinetic simulation**, among others using the stochastic Runge–Kutta method (RI5) and Gillespie's algorithm
    - Resources to learn more about stochastic modelling: MIT OpenCourseWare by Jeff Gore, and Stochastic Processes: An Introduction, Third Edition by Jones and Smith
- Tutorials also available on the website of European Bioinformatics Institute (EBI)
- The mathematical concept and software tools are important for detailed analysis of enzymatic reactions, especially in the presence of drugs and/or isease-relevant mutation

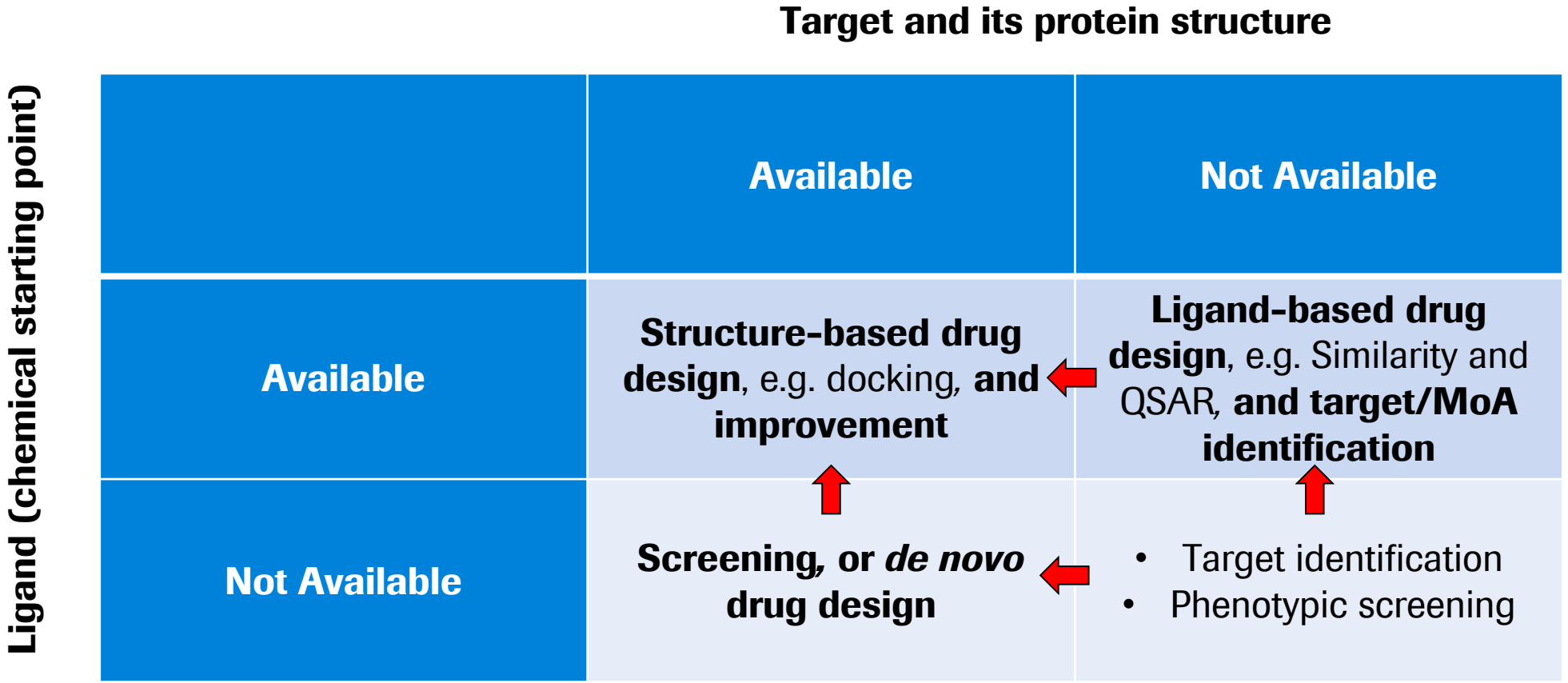


Huang and Ferrell, PNAS, 2006



ODE-based simulation of dynamics

Principles of screening and drug design – an interactive process



QSAR= quantitative structure activity relationship; MoA= mechanism of action, or mode of action

Mathematics behind approaches to determine molecular structure

Mathematical and physical foundations

- Mathematical techniques used in biophysics
- Background on imaging physics (http://xrayphysics.com)

1. X-ray diffraction by electrons

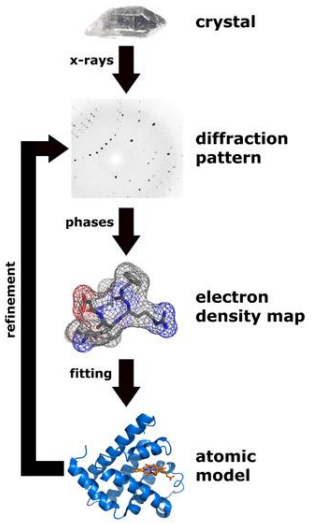
- An AMS Feature Column by Tony Phillips
- Stanford open course *Fourier transform and its applications*

2. Nuclear Magnetic Resonance (NMR)

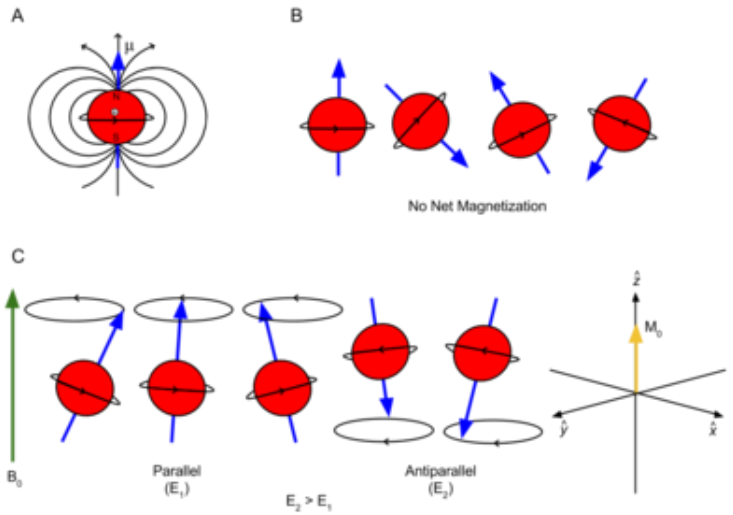
- A beautiful video tutorial about the principles of magnetic resonance imaging (MRI), which is a variant of NMR

3. Cryo electronic microscopy (CryoEM)

- A three-minute introduction to CryoEM
- Nobel Prize Talk by Joachim Frank
- Talk on Mathematics of CryoEM, by Prof Amit Singer, with a manuscript available at arXiv: https://arxiv.org/abs/1803.06714



XRD Left: Workflow of X-ray diffraction by electrons to determine protein structure. Bottom: Swiss Light Source, the synchrotron at the Paul Scherrer Institute (PSI), copyright of PSI

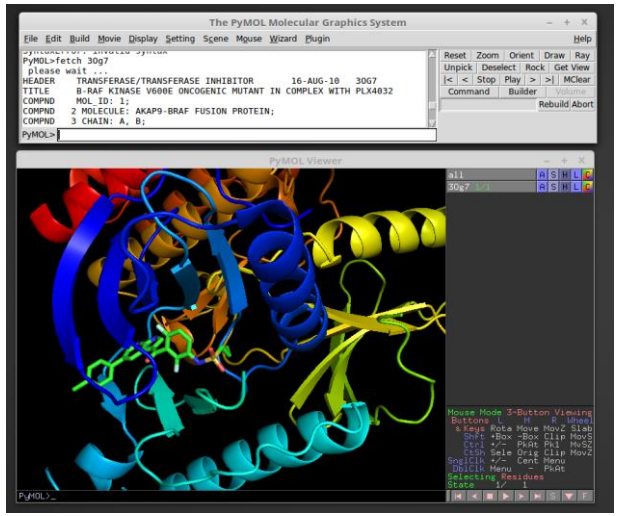


NMR Adapted from Bushberg JT, *The Essential Physics of Medical Imaging: Lippincott Williams & Wilkins;* 2002. Downloaded from http://199.116.233.101/index.php/Physics_of_MRI

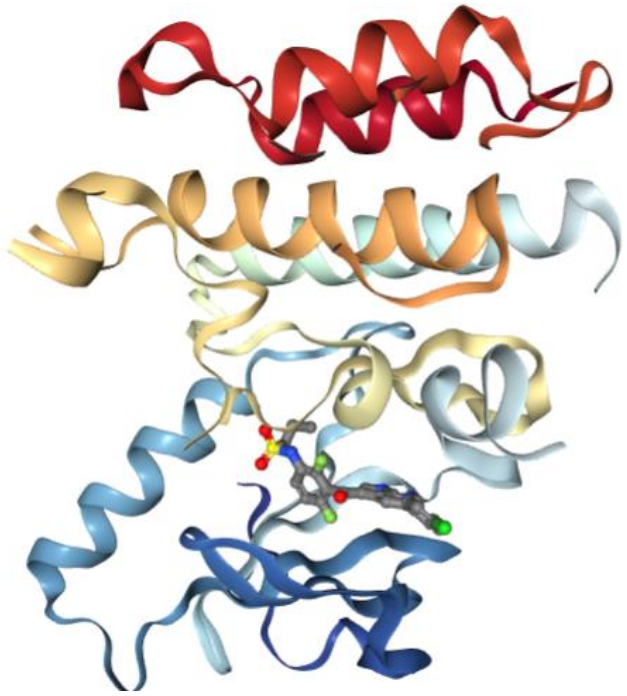
Molecular modelling

1. **Structure-based molecular modelling**
    1. Principles of molecular modelling and molecular dynamics
    2. Principles of molecular docking
2. **Ligand-based molecular modelling**
    1. Molecular similarities
    2. Molecular descriptors, QSAR, and pharmacophore models

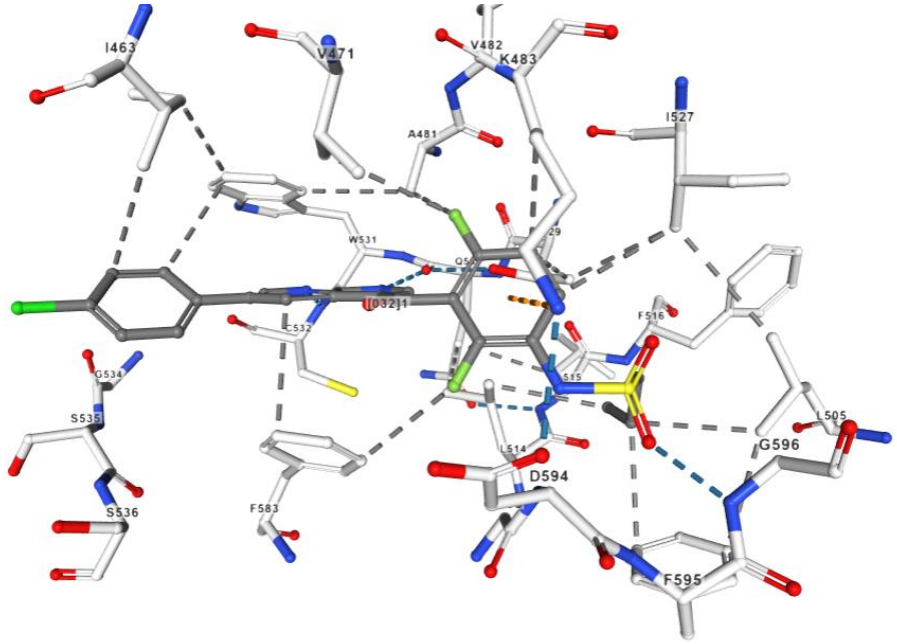
Computer presentation of protein structures: PDB



Working with PDB files with PyMol from the command-line



Structural view



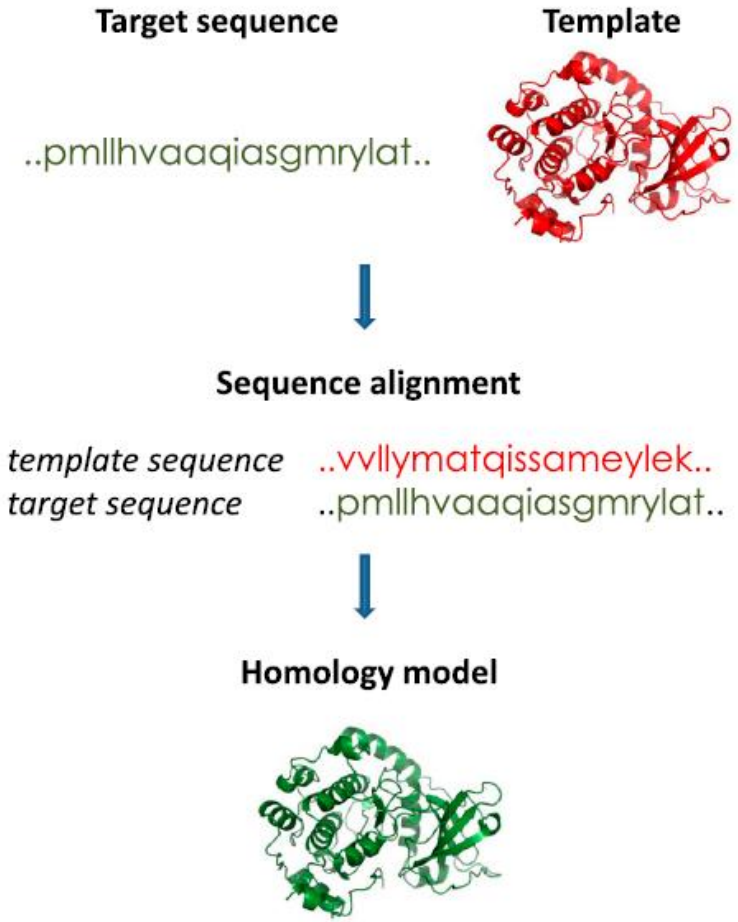
Ligand view

Balls and sticks: protein V600E and ligand (PLX4032)

Blue dahses: hydrogen bonds (<3.5 Angstrom)

Gray dahses: hydrophobic interactions (<4 Angstrom)

Homology model building in case no structure is available



Sliwoski, Gregory, Sandeepkumar Kothiwale, Jens Meiler, und Edward W. Lowe. „Computational Methods in Drug Discovery“. *Pharmacological Reviews* 66, Nr. 1 (1. Januar 2014): 334–95. https://doi.org/10.1124/pr.112.007336.

W296–W303 Nucleic Acids Research, 2018, Vol. 46, Web Server issue *Published online 21 May 2018*
doi: 10.1093/nar/gky427

SWISS-MODEL: homology modelling of protein structures and complexes

Andrew Waterhouse[1,2,†], Martino Bertoni[1,2,†], Stefan Bienert[1,2,†], Gabriel Studer[1,2,†], Gerardo Tauriello[1,2,†], Rafal Gumienny[1,2], Florian T. Heer[1,2], Tjaart A. P. de Beer[1,2], Christine Rempfer[1,2], Lorenza Bordoli[1,2], Rosalba Lepore[1,2] and Torsten Schwede[1,2,*]

[1]Biozentrum, University of Basel, Klingelbergstrasse 50–70, CH-4056 Basel, Switzerland and [2]SIB Swiss Institute of Bioinformatics, Biozentrum, University of Basel, Klingelbergstrasse 50–70, CH-4056 Basel, Switzerland

Received February 09, 2018; Revised May 01, 2018; Editorial Decision May 02, 2018; Accepted May 07, 2018

- **Levinthal's paradox**
- **CASP: Critical Assessment of Techniques for Protein Structure Prediction**
- A thought-provoking blog from Mohammed AlQuraishi: AlphaFold @ CASP13: “What just happened?”, with an informal but good overview of history of protein structure prediction, and his indictment (criminal accusations) of both academia and pharma.

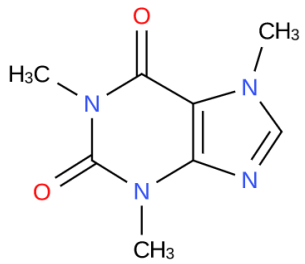
ChEMBL as information source of small molecules

Nomenclature

caffeine
1,3,7-trimethylxanthine
methyltheobromine

Bioactivity

Affinity to human proteins and drug targets



Chemical data

Formula: $C_8H_{10}N_4O_2$
Charge: 0
Mass: 194.19

Database Xrefs

PubChem: CID2519
BindingDB: 1849

Chemical Informatics

***InChI**=1/C8H10N4O2/c1-10-4-9-6-5(10)7(13)12(3)8(14)11(6)2/h4H,1-3H3*

***SMILES**: CN1C(=O)N(C)c2ncn(C)c2C1=O*

Visualisation

A subset of available information from EBI ChEBI/ChEMBL, inspired by EBI's roadshow *Small Molecules in Bioinformatics*

Representation of small molecules

Molfile: View Raw | Download | Editor | Copy

Canonical SMILES: CN1C(=O)N(C)c2ncn(C)c2C1=O

Standard InChI: InChI=1S/C8H10N4O2/c1-10-4-9-6-5(10)7(13)12(3)8(14)11(6)2/h4H,1-3H3

Standard InChI Key: RYYVLZVUVIJVGH-UHFFFAOYSA-N

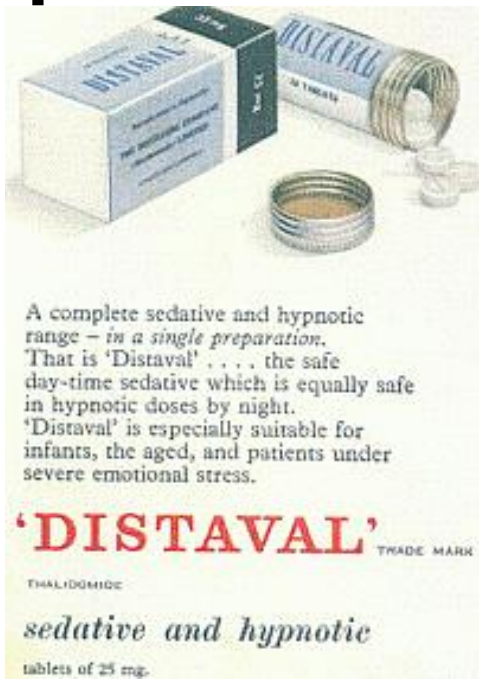
- Molfile: a type of chemical table files
- Simplified Molecular-Input Line-Entry System (SMILES)
- IUPAC International Chemical Identifier (InChI)
- InChiKey: a hash 27-character version of InChI

```
CHEMBL113

SciTegic12231509382D

14 15 0 0 0 0 999 V2000
-1.1875 -9.6542 0.0000 C 0 0
-1.1875 -8.9625 0.0000 C 0 0
-1.8125 -10.0292 0.0000 N 0 0
-2.4167 -8.9625 0.0000 N 0 0
-2.4167 -9.6542 0.0000 C 0 0
-1.8125 -8.6000 0.0000 C 0 0
-0.5000 -9.8917 0.0000 N 0 0
-0.5000 -8.7625 0.0000 N 0 0
-0.1125 -9.3042 0.0000 C 0 0
-3.0250 -10.0375 0.0000 O 0 0
-1.8125 -7.8917 0.0000 O 0 0
-1.8125 -10.7417 0.0000 C 0 0
-3.0250 -8.6000 0.0000 C 0 0
-0.2917 -8.0750 0.0000 C 0 0
2 1 2 0
3 1 1 0
4 5 1 0
5 3 1 0
6 2 1 0
7 1 1 0
8 2 1 0
9 7 2 0
10 5 2 0
11 6 2 0
12 3 1 0
13 4 1 0
```

The tragedy of thalidomide and the importance of representation



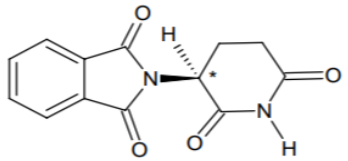
(1957)

I thank Manuela Jacklin for her help preparing this slide.

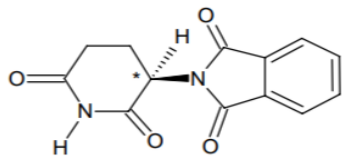
Frances Oldham Kelsey received the President's Award for Distinguished Federal Civilian Service from President John F. Kennedy, 1962

Canonic SMILES of thalidomide

C1CC(=O)NC(=O)C1N2C(=O)C3=CC=CC=C3C2=O



(-)(*S*)-thalidomide



(+)(*R*)-thalidomide

Isomeric SMILES of (-)(S)-thalidomide

C1CC(=O)NC(=O)[C@H]1N2C(=O)C3=CC=CC=C3C2=O

Isomeric SMILES of (+)(R)-thalidomide

C1CC(=O)NC(=O)[C@@H]1N2C(=O)C3=CC=CC=C3C2=O

The principle of molecular docking, a case study of structure-based drug design

- **Docking is like a discotheque: it is all about posing and scoring – Roger Sayle** (*NextMove* Software Limited)
- Three basic methods to represent target and ligand structures *in silico*
  - *Atomic*: used in conjunction with a potential energy function, computational complexity high
  - *Surface*: often used in protein-protein docking
  - *Grid representation*:
    - Basic idea: to store information about the receptor's energetic contributions on grid points so that it only needs to be read during ligand scoring.
    - In the most basic form, grid points store two types of potentials: **electrostatic** and **van der Waals**.

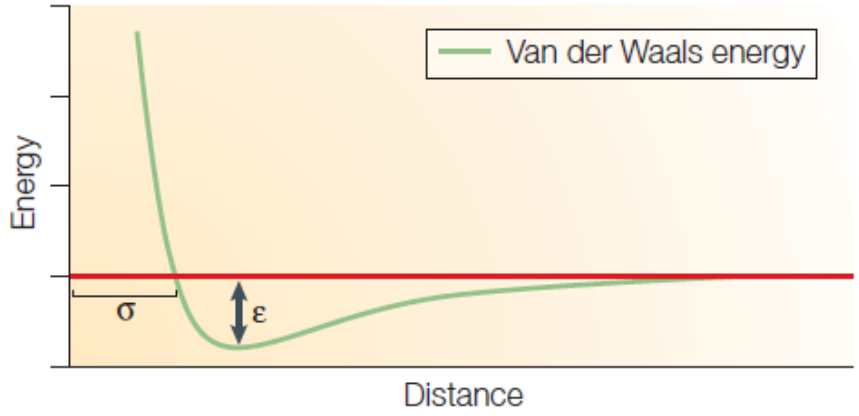
$$E_{coul}(r) = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$

Coulombic interactions

$$E_{vdW}(r) = \sum_{j=1}^{N} \sum_{i=1}^{N} 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right]$$

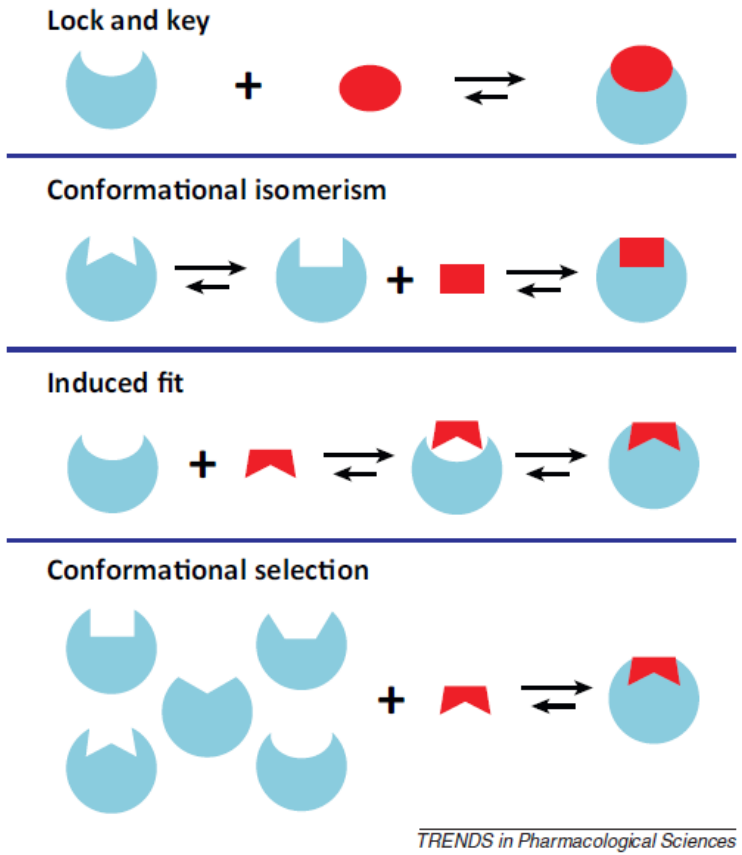
Lennard–Jones 12–6 function

- ε is the **well depth** of the potential
- σ is the **collision diameter** of the respective atoms *i* and *j*.

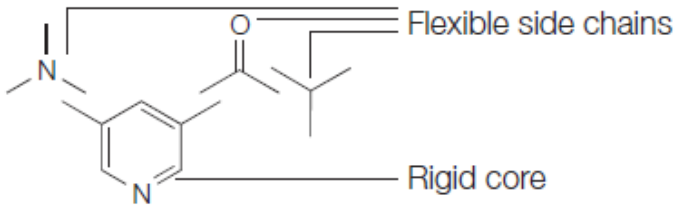


Kitchen, Douglas B., Hélène Decornez, John R. Furr, und Jürgen Bajorath. „Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications". *Nature Reviews Drug Discovery* 3, Nr. 11 (November 2004): 935–49. https://doi.org/10.1038/nrd1549.

Posing: dealing with flexibility of ligand and of protein



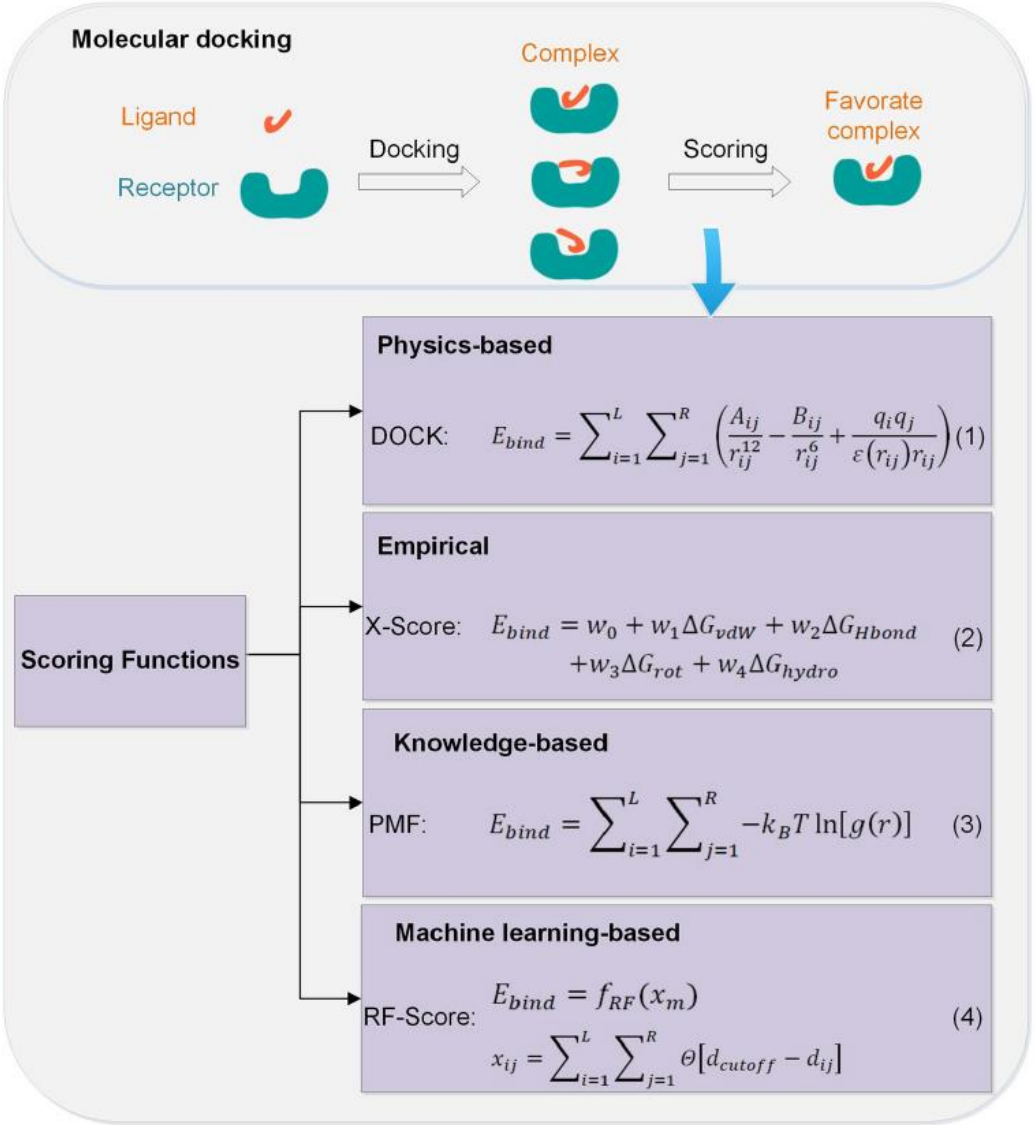
Chen, Yu-Chian. „Beware of docking!“ *Trends in Pharmacological Sciences* 36, Nr. 2 (1. Februar 2015): 78–95. https://doi.org/10.1016/j.tips.2014.12.001.



Methods to deal with ligand and protein flexibility

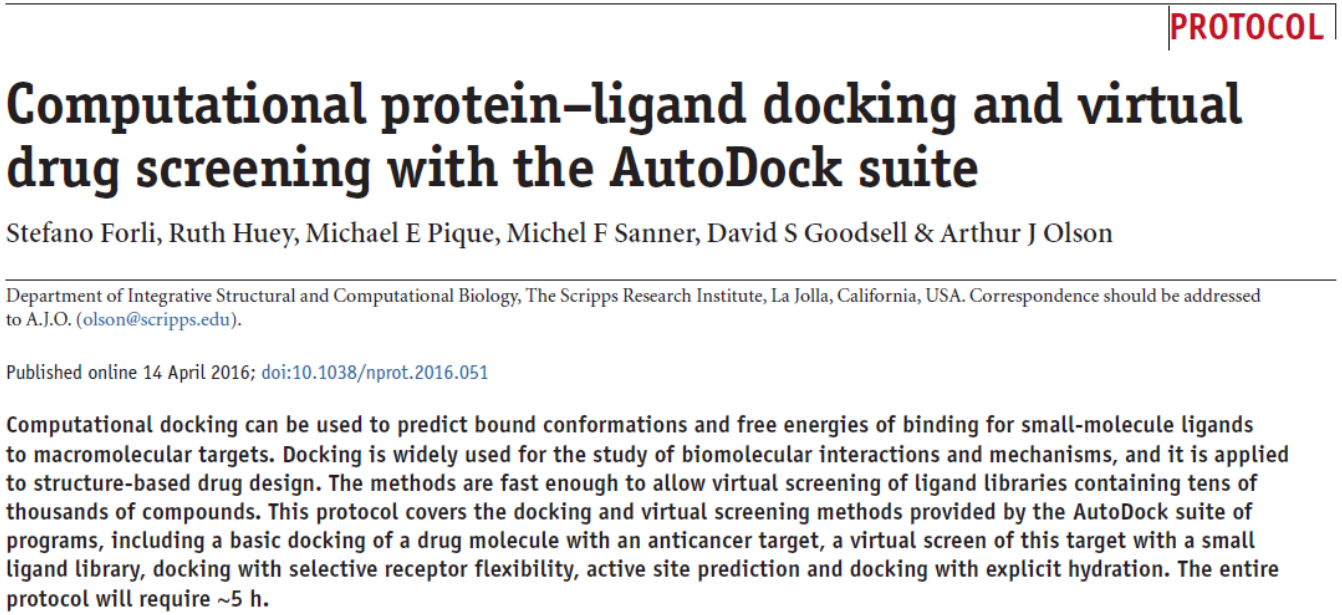
- Systematic search
- Random search, such as Monte-Carlo and genetic algorithms
- Simulation methods, such as molecular dynamics

Types of scoring functions



Li, Jin, Ailing Fu, und Le Zhang. „An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking“. *Interdisciplinary Sciences: Computational Life Sciences* 11, Nr. 2 (1. Juni 2019): 320–28. https://doi.org/10.1007/s12539-019-00327-w.

Interested in learning more about drug design?



PROTOCOL

Computational protein–ligand docking and virtual drug screening with the AutoDock suite

Stefano Forli, Ruth Huey, Michael E Pique, Michel F Sanner, David S Goodsell & Arthur J Olson

Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California, USA. Correspondence should be addressed to A.J.O. (olson@scripps.edu).

Published online 14 April 2016; doi:10.1038/nprot.2016.051

Computational docking can be used to predict bound conformations and free energies of binding for small-molecule ligands to macromolecular targets. Docking is widely used for the study of biomolecular interactions and mechanisms, and it is applied to structure-based drug design. The methods are fast enough to allow virtual screening of ligand libraries containing tens of thousands of compounds. This protocol covers the docking and virtual screening methods provided by the AutoDock suite of programs, including a basic docking of a drug molecule with an anticancer target, a virtual screen of this target with a small ligand library, docking with selective receptor flexibility, active site prediction and docking with explicit hydration. The entire protocol will require ~5 h.

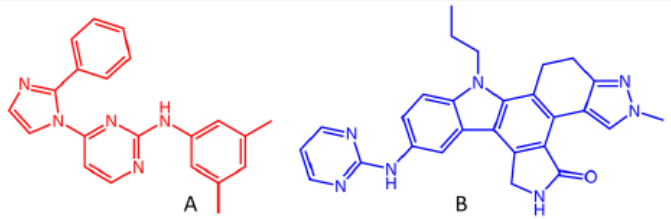
- Try docking yourself by following this protocol: Forli, Stefano, Ruth Huey, Michael E. Pique, Michel F. Sanner, David S. Goodsell, und Arthur J. Olson. „Computational Protein–Ligand Docking and Virtual Drug Screening with the AutoDock Suite“. *Nature Protocols* 11, Nr. 5 (Mai 2016): 905–19. https://doi.org/10.1038/nprot.2016.051.
- In-depth reading: Sliwoski, Gregory, Sandeepkumar Kothiwale, Jens Meiler, und Edward W. Lowe. „Computational Methods in Drug Discovery“. *Pharmacological Reviews* 66, Nr. 1 (1. Januar 2014): 334–95. https://doi.org/10.1124/pr.112.007336.
- A more advanced talk by Arthur Olson can be found here, Workshop on the Mathematics of Drug Design/Discovery, June 4 - 8, 2018, The Fields Institute.
- Courses available at the University of Basel and beyond.

Molecular similarity and the Tanimoto (Jaccard) Index

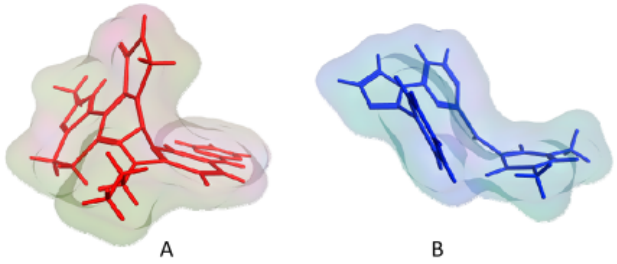
Chemical similarity

	Mol. weight	LogP	Rotatable bonds	Aromatic rings	Heavy atoms
A	341.4	5.23	4	4	26
B	463.5	4.43	4	5	35

Molecular similarity / 2D similarity



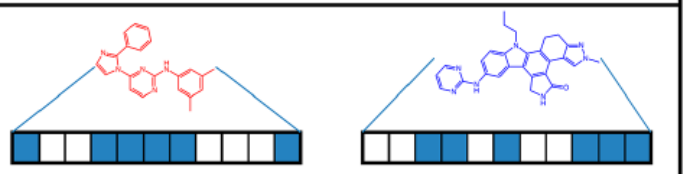
3D similarity



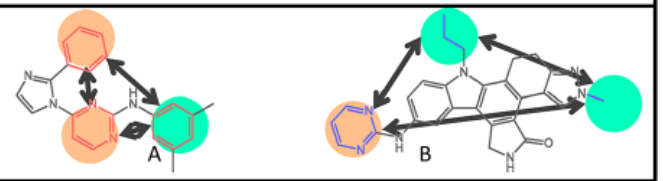
Biological similarity

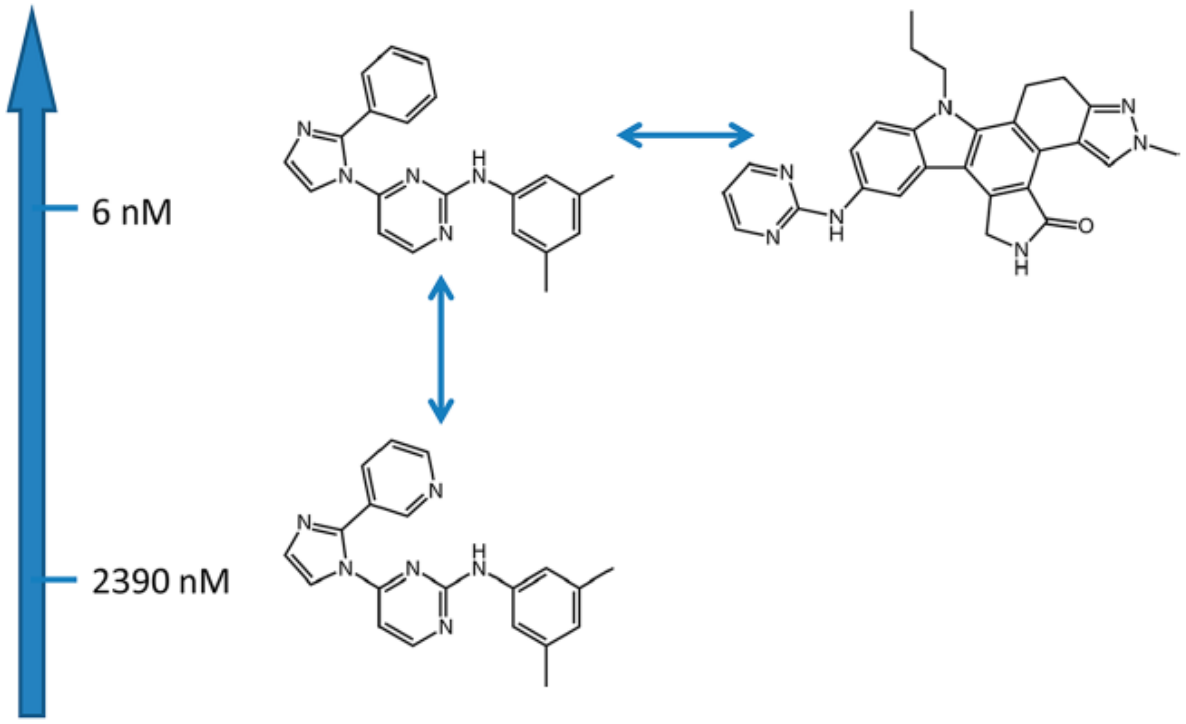
	Vascular endothelial growth factor receptor 2	Tyrosine-protein kinase TIE-2
A	active	inactive
B	active	active

Global similarity



Local similarity





Maggiora, Gerald, Martin Vogt, Dagmar Stumpfe, und Jürgen Bajorath. „Molecular Similarity in Medicinal Chemistry“. *Journal of Medicinal Chemistry* 57, Nr. 8 (24. April 2014): 3186–3204. https://doi.org/10.1021/jm401411z.

Summary and Q&A

The Great Wave off Kanagawa『神奈川沖浪裏』, by Katsushika Hokusai, downloaded from wikimedia