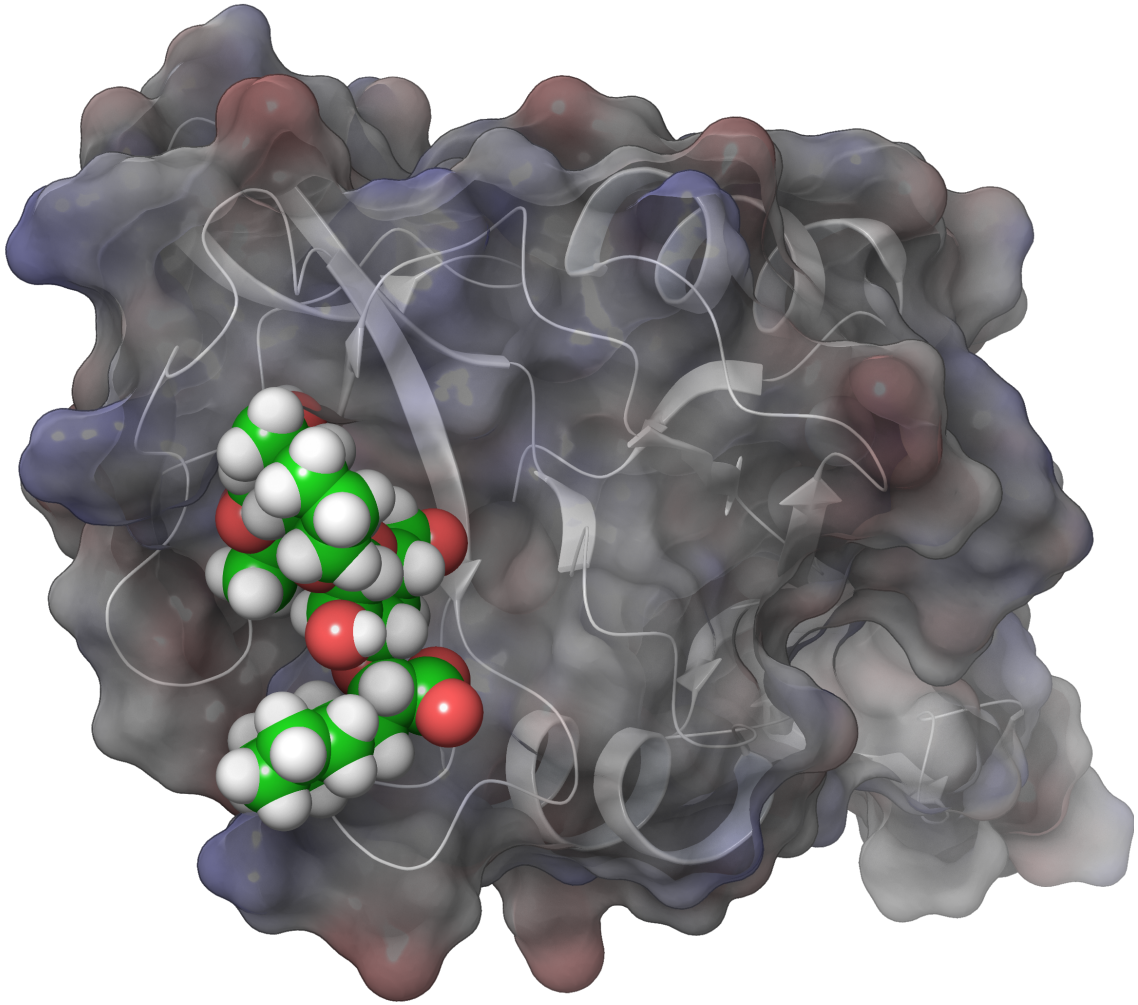
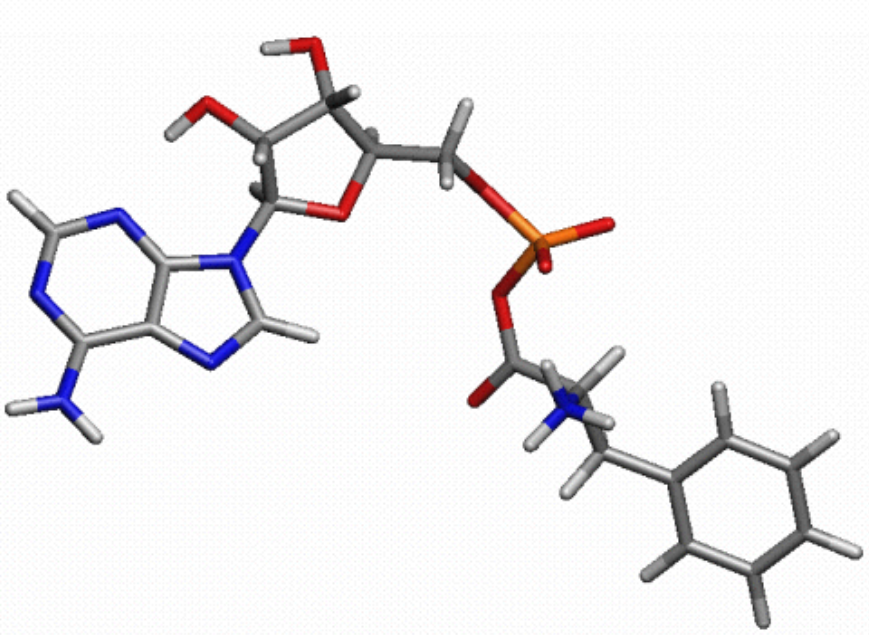


Vorlesung 12205: **Einführung ins Molecular Modeling**

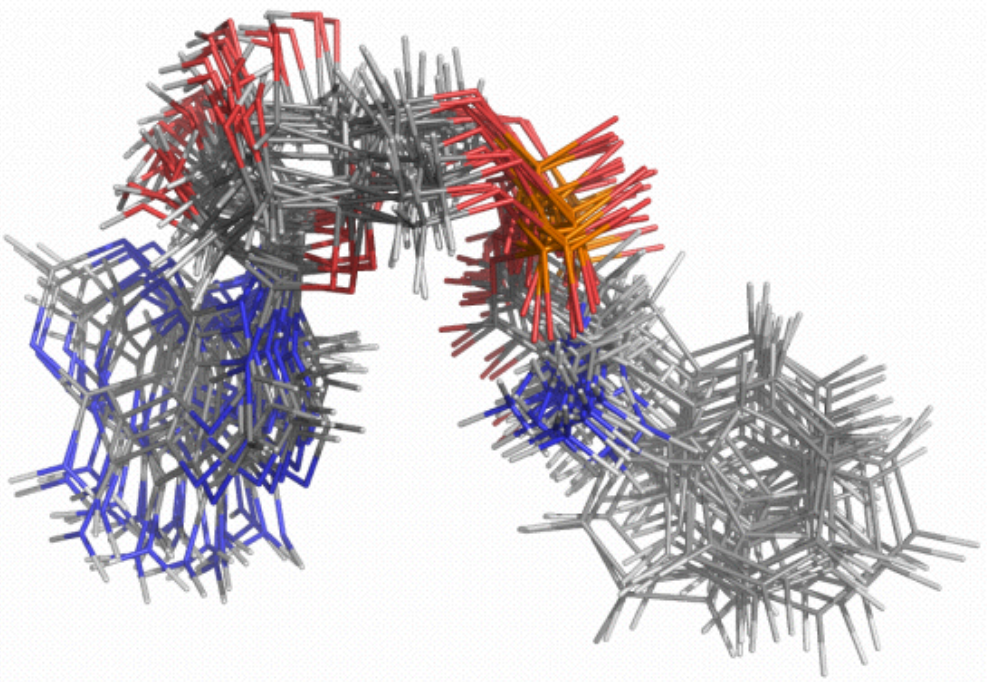
M. Smieško & A. Vedani — Departement Pharmazeutische Wissenschaften, Universität Basel, HS-2017

Moleküldynamik-Simulationen

Newton'sche Bewegungsgleichung (1687): $r(t+\Delta t) = r(t) + \partial r/\partial t \cdot \Delta t + 0.5 \cdot \partial^2 r/\partial t^2 \cdot \Delta t^2$



Minimums-Energie-Struktur von Phe–AMP



120 ps Moleküldynamik-Simulation von Phe–AMP

Moleküldynamik simuliert molekulare Bewegungen auf atomarer Ebene

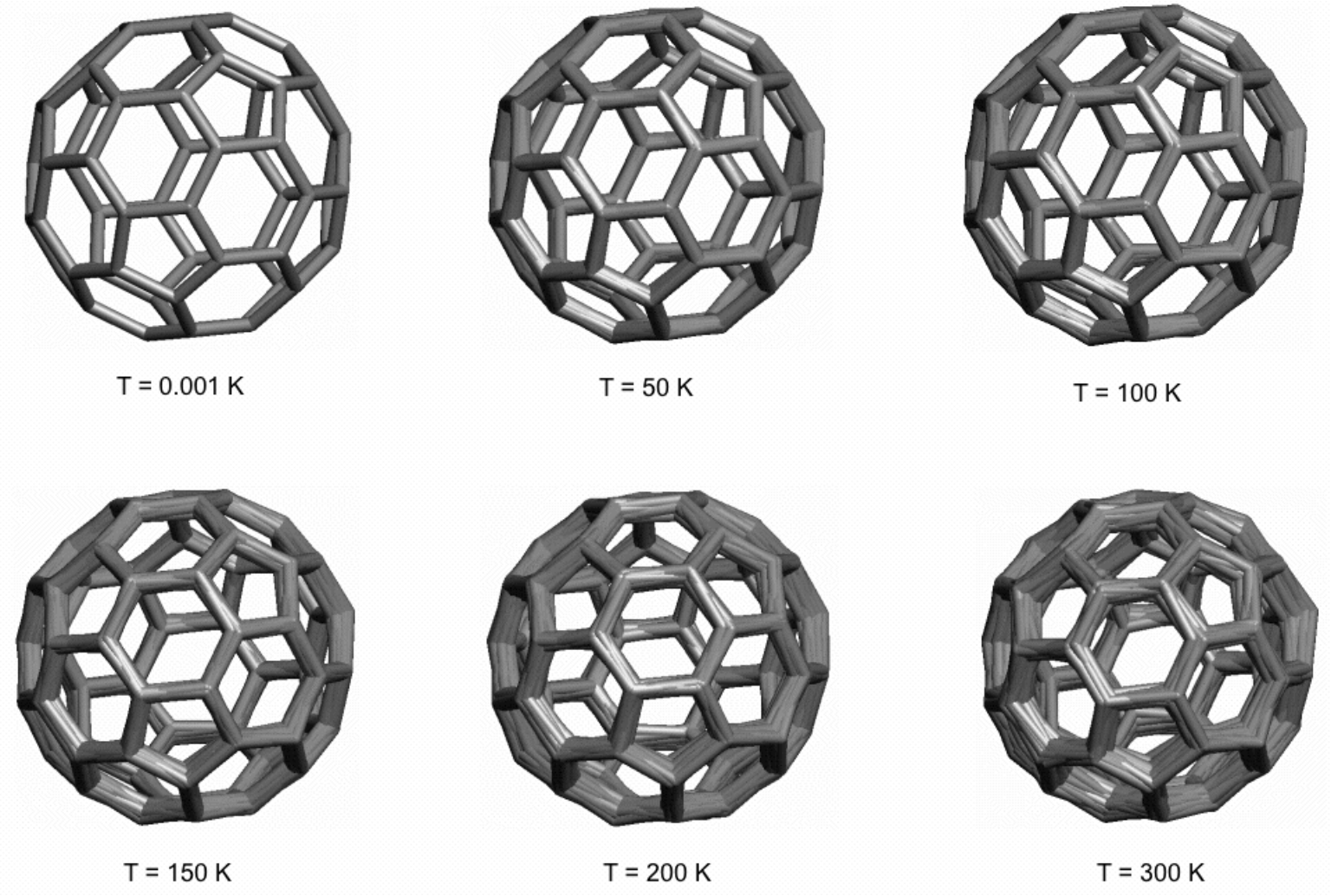
Zeitskala biophysikalischer Vorgänge

Newton'sche Bewegungsgleichung (1687): $r(t+\Delta t) = r(t) + \partial r/\partial t \cdot \Delta t + 0.5 \cdot \partial^2 r/\partial t^2 \cdot \Delta t^2$

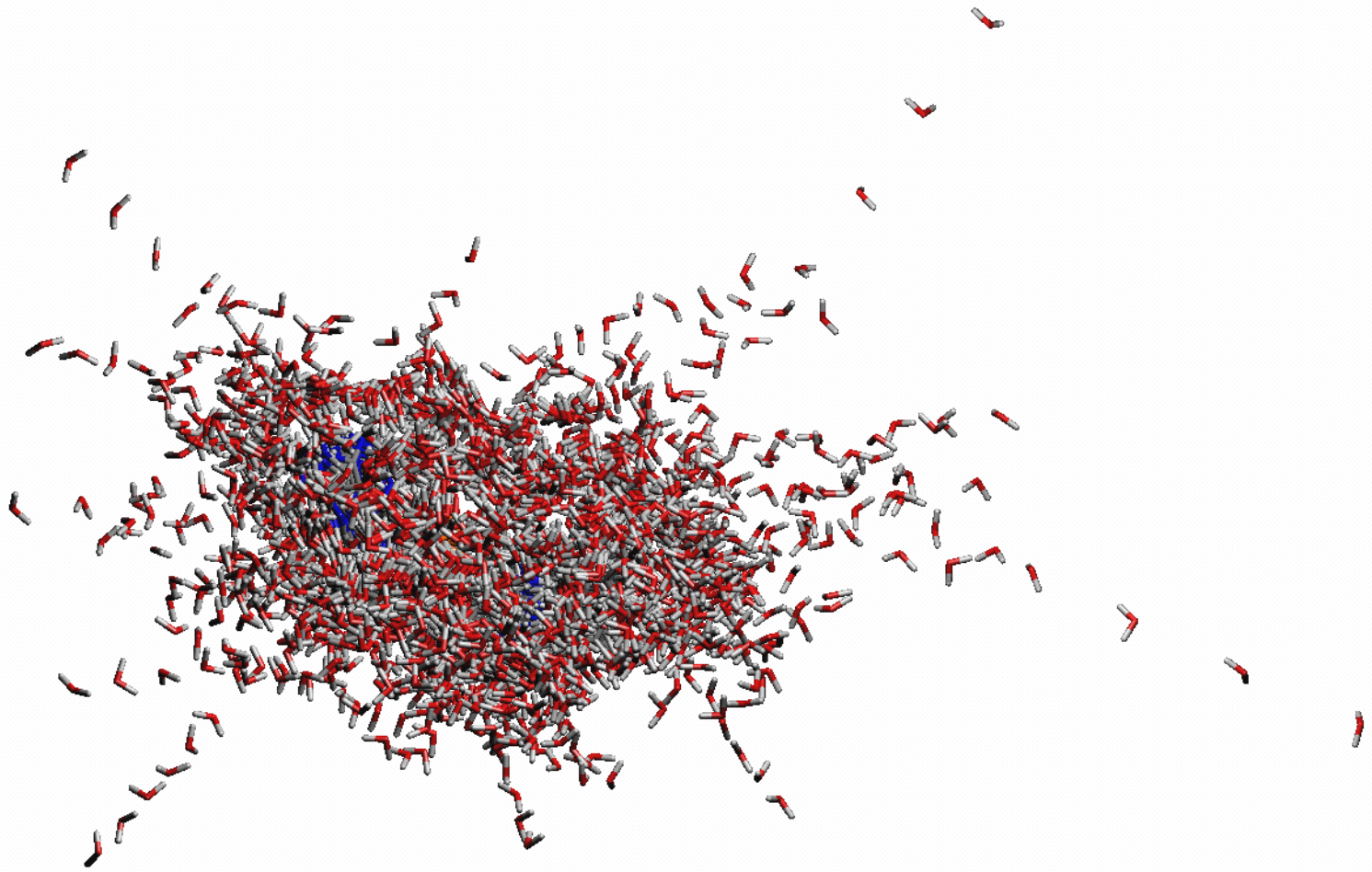
Biophysikalischer Vorgang	Zeitdauer [sec]
Vibration gebundener Atome	$10^{-14} - 10^{-13}$
Longitudinale Bewegung von Basen in einer DNA	$10^{-14} - 10^{-12}$
Globale Dehnung einer DNA	$10^{-13} - 10^{-11}$
Drehung einer Protein-Seitenkette an der Proteinoberfläche	$10^{-11} - 10^{-10}$
„Befreiung" einer verborgenen Seitenkette durch Torsionsänderungen	$10^{-11} - 10^{-9}$
Relative Bewegung globulärer Regionen in Proteinen	$10^{-11} - 10^{-7}$
Drehung einer mittelgrossen Seitenkette im Innern des Proteins	$10^{-4} - 1$
Allosterische Übergänge	$10^{-5} - 1$
Lokale Denaturierung	$10^{-5} - 10$
Diffusion kleiner Moleküle in die Bindungstasche	$10^{-6} - 10^{-3}$

Zeitintervall für die Integration $\Delta t = 2 \times 10^{-15}$ sec

Einfluss der Temperatur auf Moleküldynamik-Simulationen

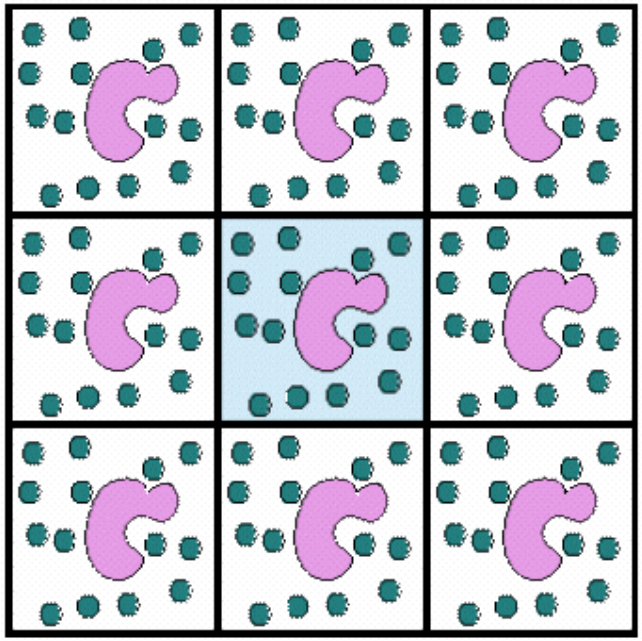


MD-Simulationen ohne periodische Randbedingungen

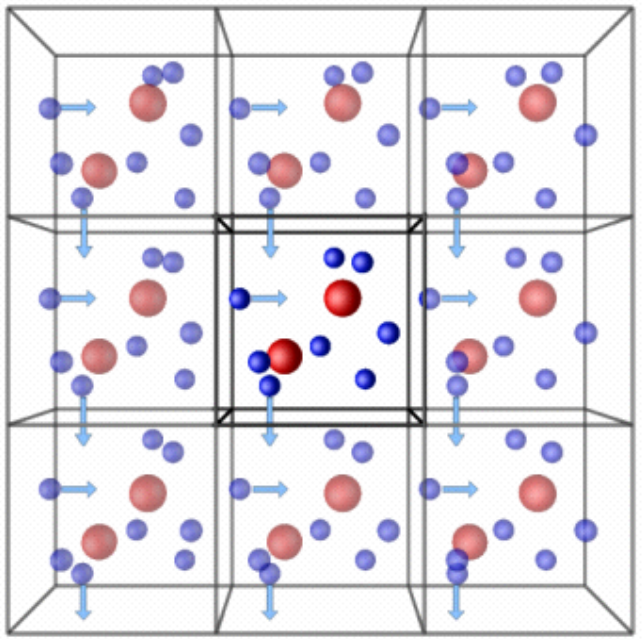


Wassermoleküle an der Protein-/Molekül-Oberfläche können wegdiffundieren

Periodische Randbedingungen in MD-Simulationen



ccl.net

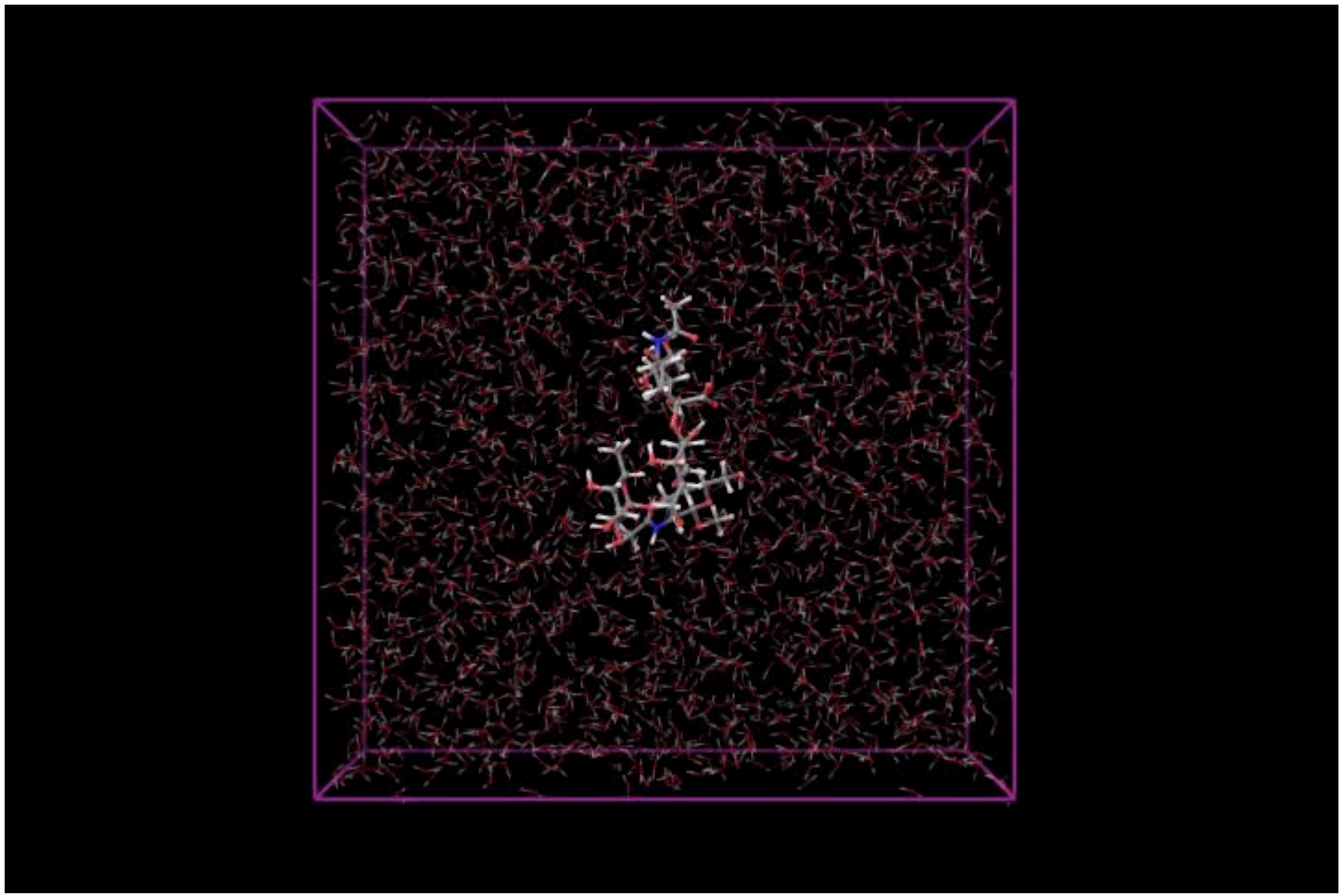


isaacs.sourceforge.net

Periodische Randbedingungen erzeugen in der sog. Simulationsbox entlang aller drei Raumrichtungen ein identisches Abbild des zu simulierenden Systems (Protein, Wirkstoff, Ionen, Wasser). Dieses wird 3^3 = 27 mal (1 × real und 26 × virtuell) abgebildet.

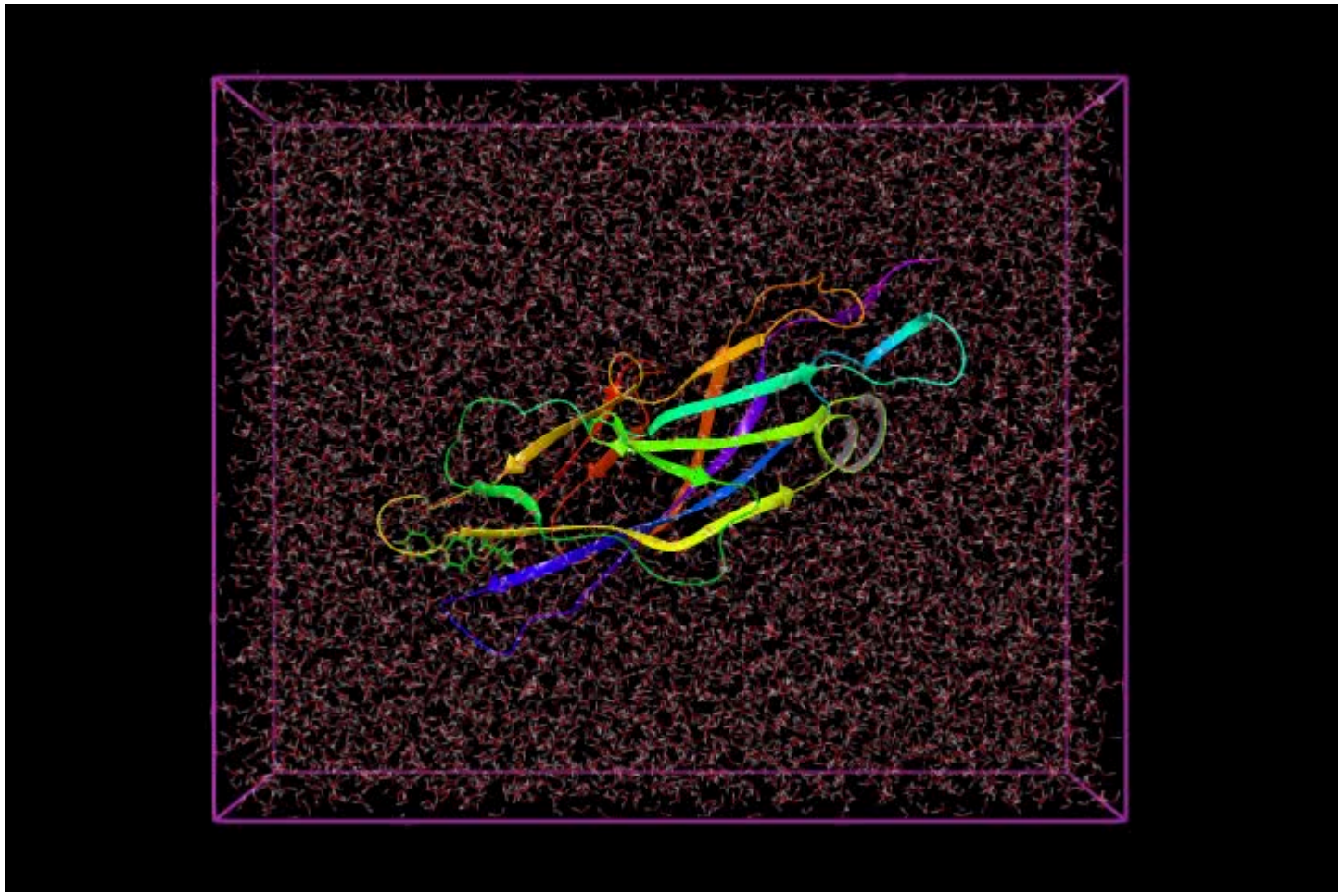
Die Protagonisten agieren in einer dreidimensional in sich geschlossen Welt und können diese (während einer Simulation) nie verlassen.

Periodische Randbedingungen in MD-Simulationen



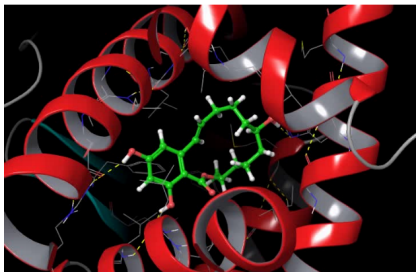
Simulated Annealing (MD Simulation) von sialyl-Lewis^X Tetrasaccharid im Wasser

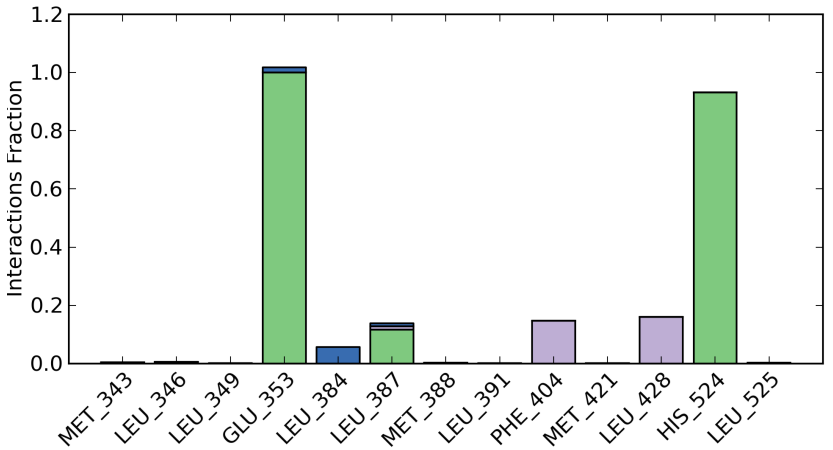
Periodische Randbedingungen in MD-Simulationen

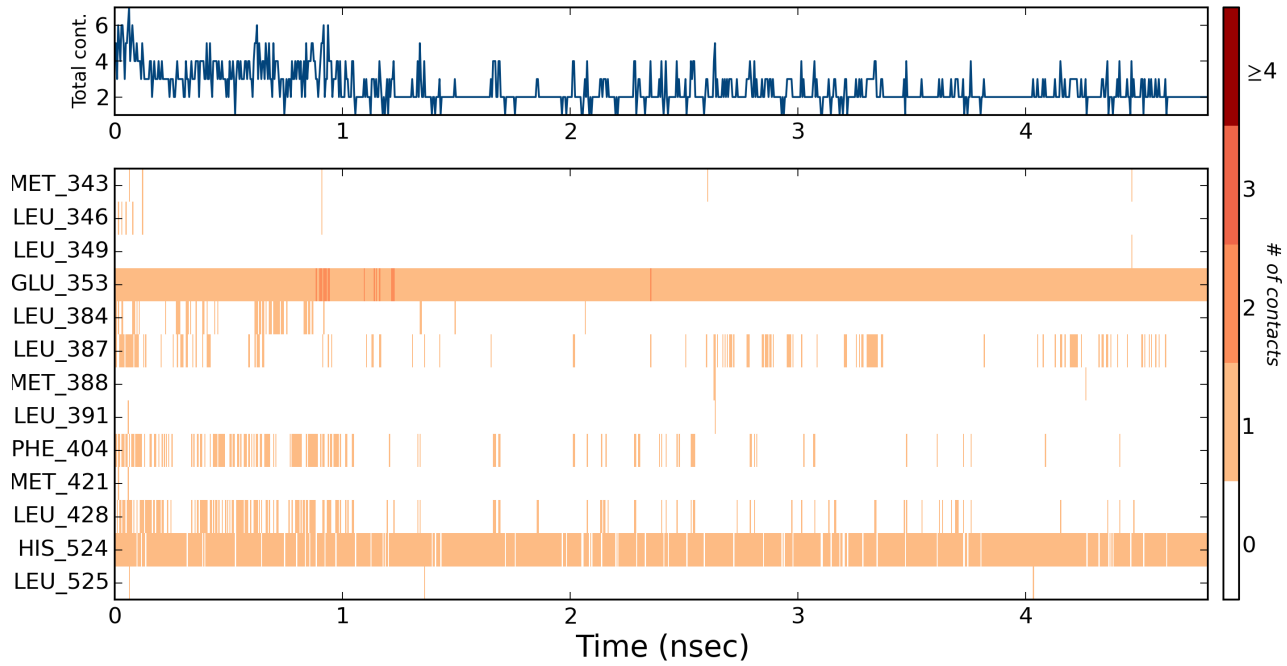


MD Simulation von dem bakteriellen Virulenzfaktor FimH mit einem Liganden im Wasser

MD simulation







Moleküldynamik - Simulationen

Bedeutung: das Studium (die Quantifizierung) von der zeitlich gemittelten Eigenschaften der Moleküle → die Zeit erfüllt die Rolle eines „Verstärkers“ , Resultate werden oft mit der NMR verglichen

Anwendung: chemische Physik, Materialwissenschaft, Biomoleküle (Pharma- relevante Anwendungen: Konformationssuche, Moleculares Docking, induced-fit, Thermodynamik...)

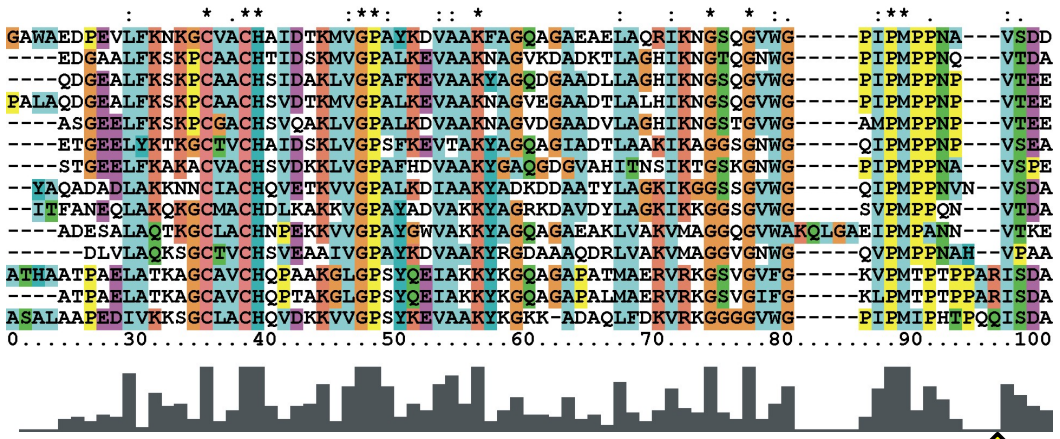
Nobel Prize Chemie 2013: „Development of multiscale models for complex chemical systems“

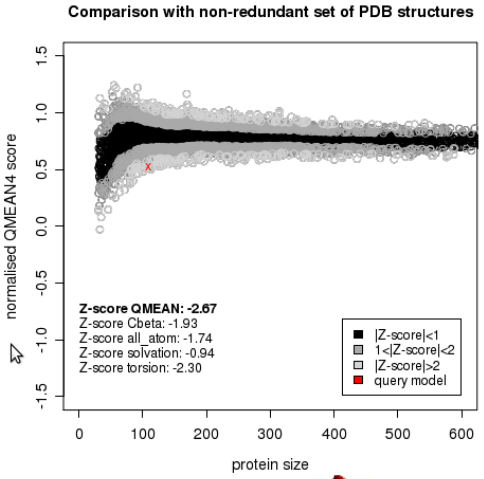
Martin Karplus

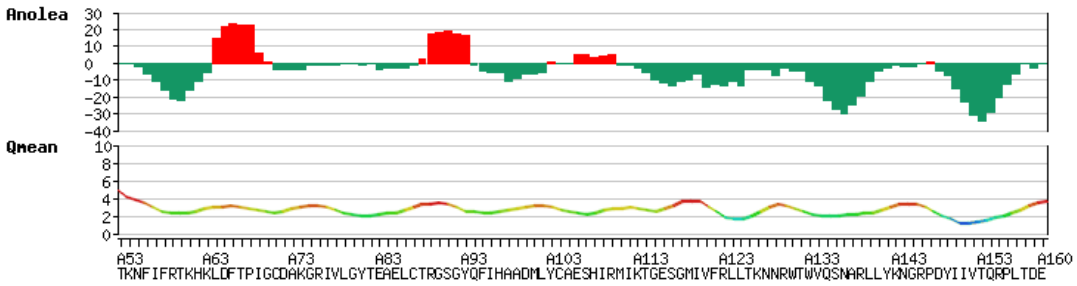

Arieh Warshel

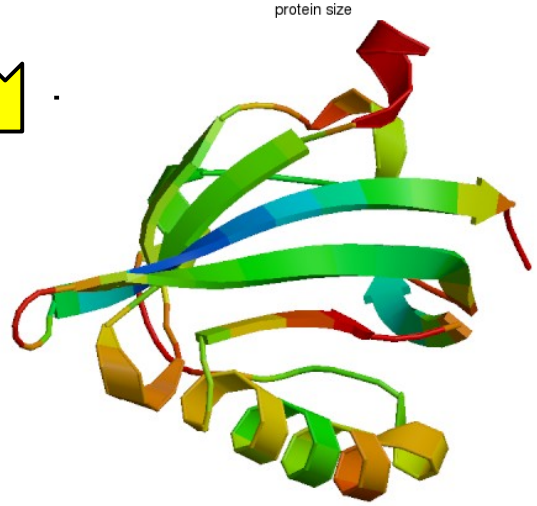

Michael Levitt

Homologie Modelierung









Design techniques

Ligand-based design

3D structure of the target protein is **unknown**

screening study

↓

hit(s)

↓

conformational search

superposition

pharmacophore

ligand optimization / database searching

scoring: ***rmsd*, QSAR, CoMFA**, etc.

↓

assay

↓

lead compound / drug candidate

Structure-based design

3D structure of the target protein is **known**

target purification & crystallization

↓

protein(-ligand) 3D structure

↓

molecular docking

pharmacophore

ligand optimization / database searching

scoring: **generic function, QM-MM** etc.

↓

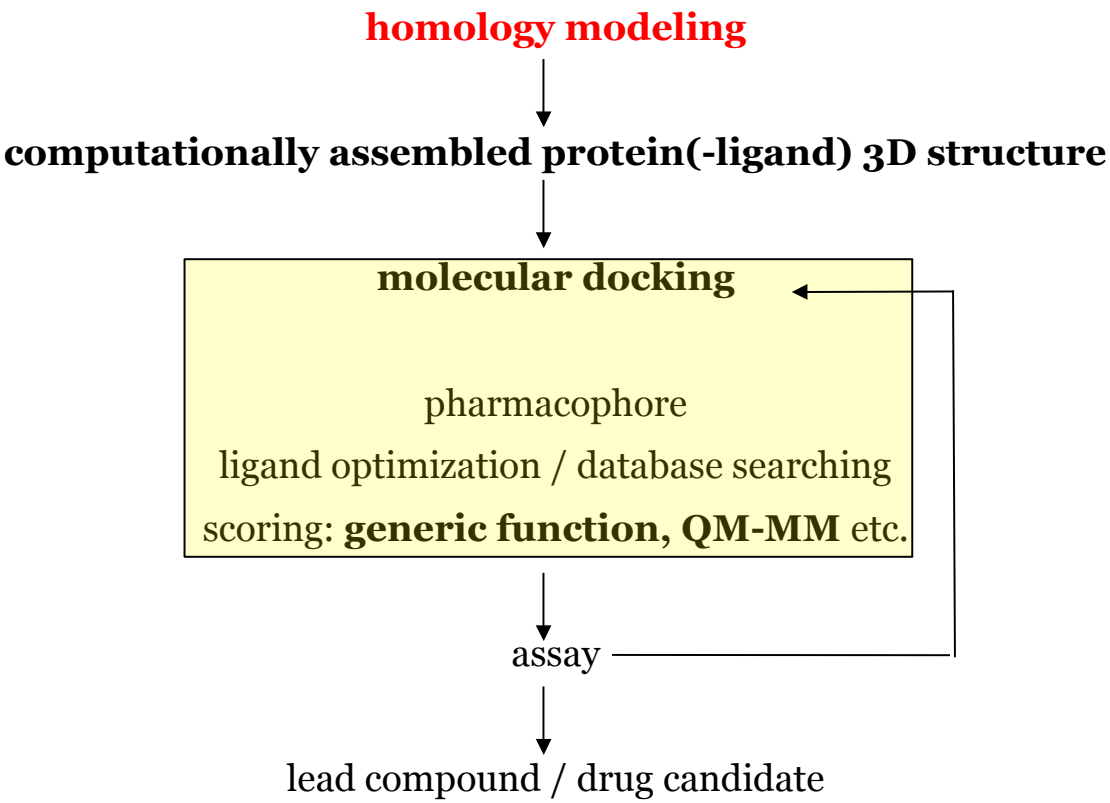
assay

↓

lead compound / drug candidate

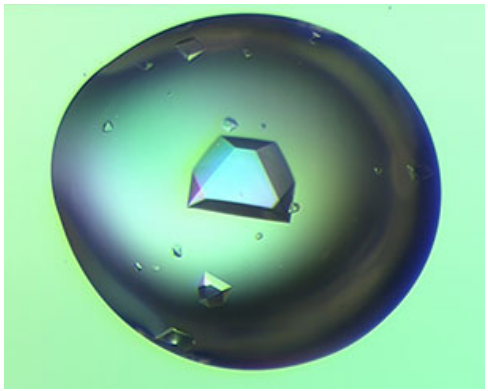
Homology modeling & Structure-based design

3D structure of the target protein is unknown, but **primary sequence** and the 3D structure of a related protein – **template** – **are known**



Homology Modeling

- no crystal structure available (sufficient amount, purity, crystal forming → all hard to do with membrane bound proteins)
- moreover crystal must diffract (measurable Röntgen reflexions detected) → depends on the quality of the crystal, especially on its water content.



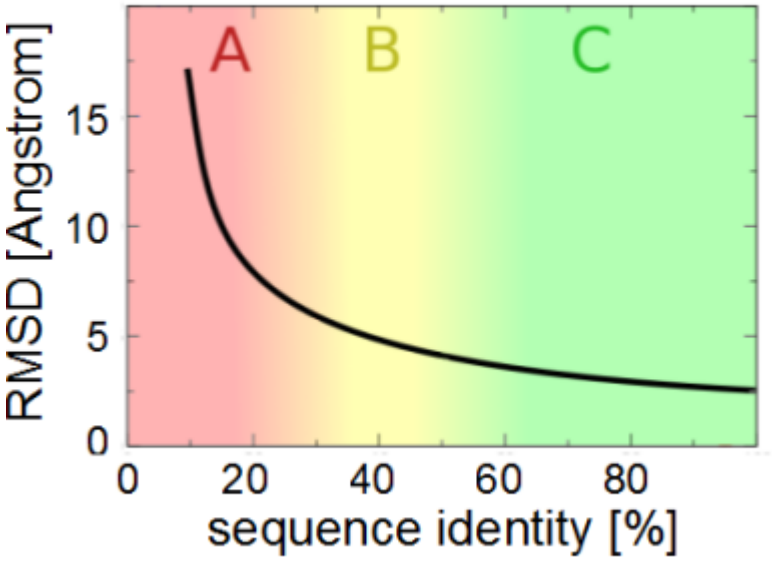
www.ttplabtech.com

- main goal of HM is (using various computational techniques) to construct a 3D structure of the target protein based on the 3D structure of a homologous protein
- homology does not mean „similarity“ *per se* — it rather means the two proteins have a common evolutionary origin

Homology Modeling

Genome → Gene → Protein sequence → Protein 3D structure

→ Structure-based design → Drug



Probabilities of SWISS-MODEL accuracy for target-template identity classes.

Percent sequence identity[a]	Total number of models[b]	Percent[c] models with rmsd lower than 1 Å	Percent models with rmsd lower than 2 Å	Percent models with rmsd lower than 3 Å	Percent models with rmsd lower than 4 Å	Percent models with rmsd lower than 5 Å	Percent models with rmsd higher than 5 Å
25-29	125	0	10	30	46	67	33
30-39	222	0	18	45	66	77	23
40-49	156	9	44	63	78	91	9
50-59	155	18	55	79	86	91	9
60-69	145	38	72	85	91	92	8
70-79	137	42	71	82	85	88	12
80-89	173	45	79	86	94	95	5
90-95	88	59	78	83	86	91	9

Homology Modeling: Applications & Assumptions

Application

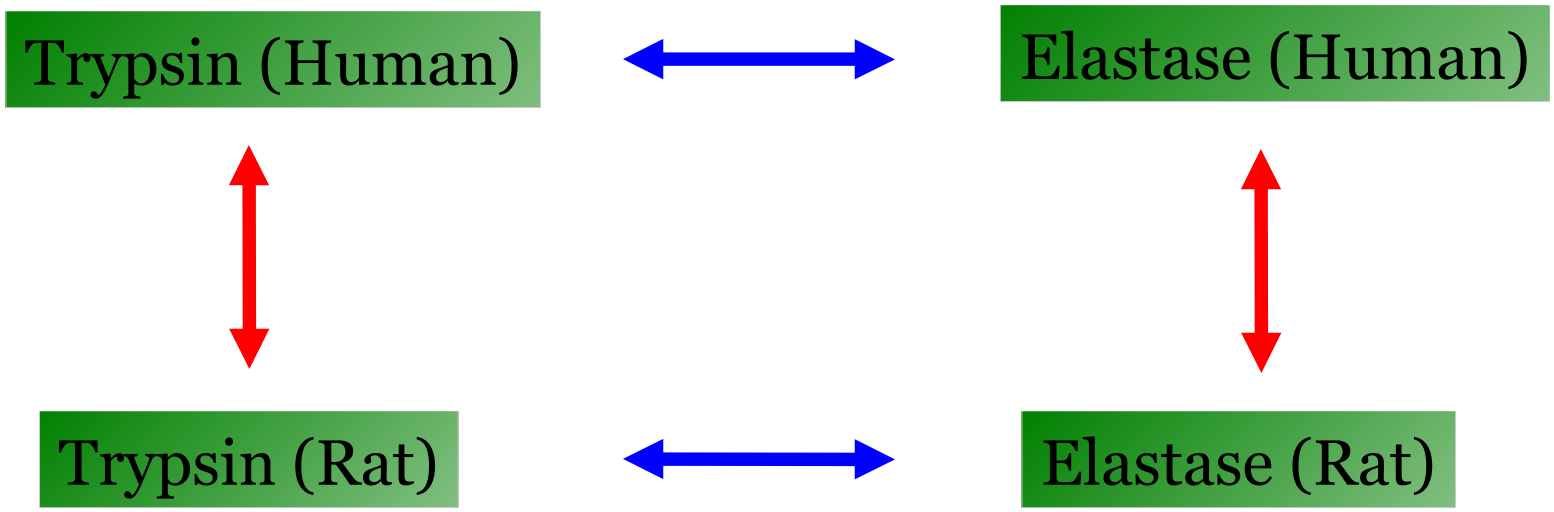
- Structure-based Drug Design
- Structure-based prediction of Metabolism and Toxicity
- Structure-based Evaluation of *target drugability*
- Design of Mutation Experiments
- Development of *in vitro* Test-Assays

Assumptions

- tertiary structure (3D) of a protein is determined by (depends on) its primary structure
- during the evolution the tertiary structure is changing much slower than the underlying primary sequence, i.e. similar sequences lead to identical tertiary structures, while distantly related sequences can lead to similar 3D folding of a protein

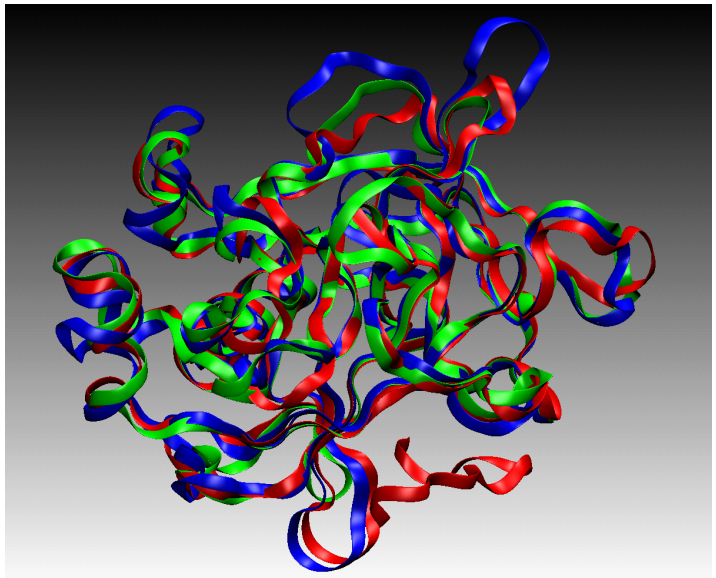
Orthologs and Paralogs

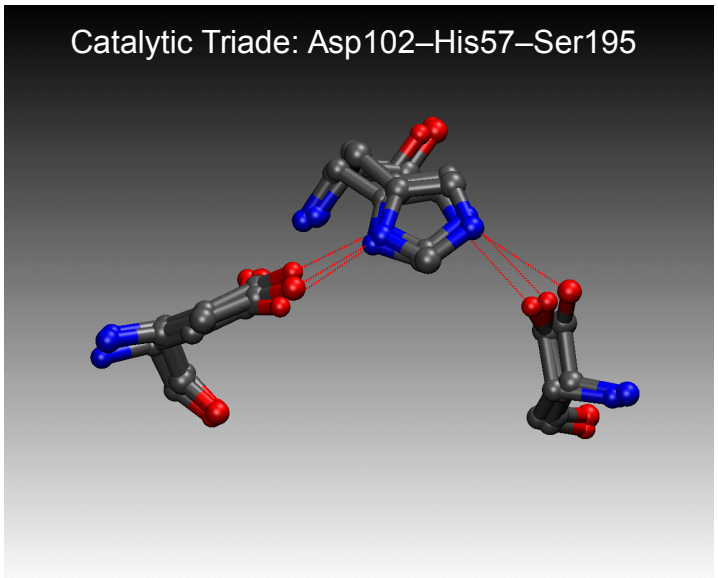
- **Orthologs have the same function in different species**
- **Paralogs have similar function in the same species (gene duplication)**



Comparison of Protein Structures

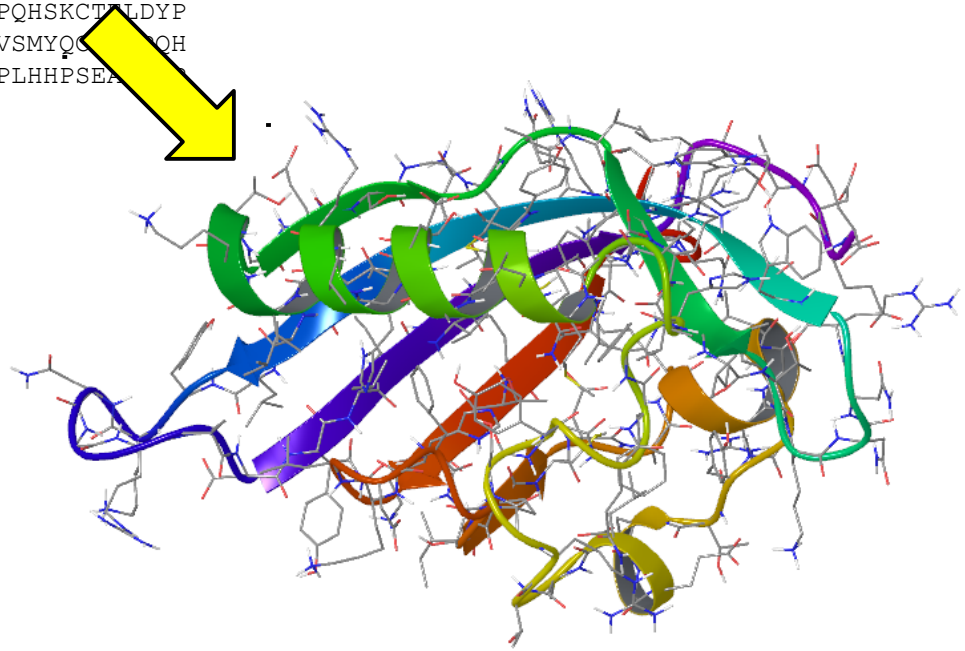
If two (or more) proteins are compared against each other, this is mostly done based on their primary sequence. It is assumed that proteins with comparable amino acid sequence have also comparable biological function. An important basis for the Homology Modeling is, that the **tertitary structure is often better conserved than the primary structure**, e.g. 3D-structure of the serin proteases *Trypsin*, *Chymotrypsin* and *Elastase* can be almost perfectly superimposed, despite only 30–35% agreement of their primary seqence.



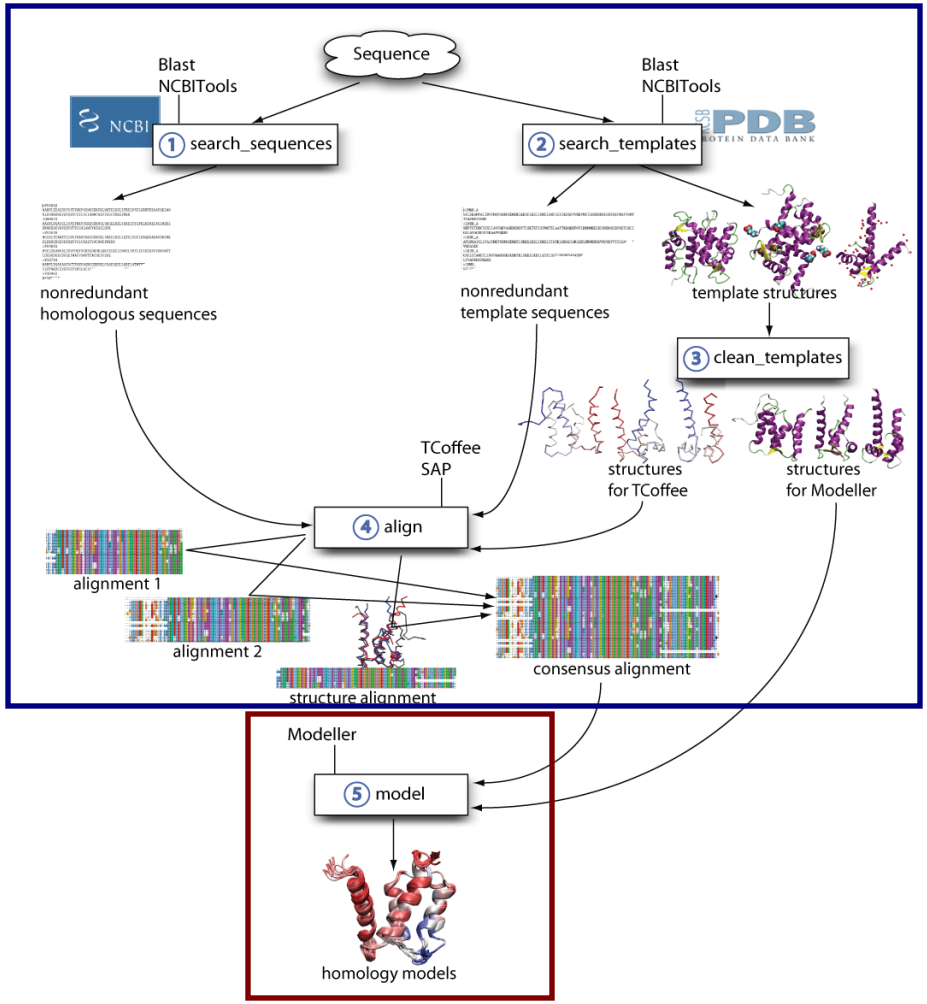


Homology Modeling

```
>sp|P35869|AHR_HUMAN Aryl hydrocarbon receptor OS=Homo sapiens GN=AHR PE=1 SV=2
MNSSSANITYASRKRRKPVQKTVKPIPAEGIKSNPSKRHRDRLNTELDRLASLLPFPQDV
INKLDKLSVLRLSVSYLRAKSFFDVALKSSPTERNGGQDNCRAANFREGLNLQEGEFLLQ
ALNGFVLVVTTDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQRQLHWALNP
SQCTESGQGIEEATGLPQTVVCYNPDQIPPENSPLMERCFICRLRCLLDNSSGFLAMNFQ
GKLKYLHGQKKKGKDGSILPPQLALFAIATPLQPPSILEIRTKNFIFRTKHKLDFTPIGC
DAKGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLTKNNRW
TWVQSNARLLYKNGRPDYIIVTQRPLTDEEGTEHLRKRNTKLPFMFTTGEAVLYEATNPF
PAIMDPLPLRTKNGTSGKDSATTSTLSKDSLNPSSLLAAMMQQDESIYLYPASSTSSTAP
FENNFFNESMNECRNWQDNTAPMGNDTILKHEQIDQPQDVNSFAGGHPGLFQDSKNSDLY
SIMKNLGIDFEDIRHMQNEKFFRNDFSGEVDFRDIDLTDEILTYVQDSLSKSPFIPSDYQ
QQQSLALNSSCMVQEHLHLEQQQQHHQKQVVVEPQQQLCQKMKHMQVNGMFENWNSNQFV
PFNCPQQDPQQYNVFTDLHGISQEFPYKSEMDSMPYTQNFISCNQPVLPQHSKCTLLDYP
MGSFEPSPYPTTSSLEDFVTCLQLPENQKHGLNPQSAIITPQTCYAGAVSMYQ[illegible]QH
THVGQMQYNPVLPGQQAFLNKFQNGVLNETYPAELNNINNTQTTTHLQPLHHPSEA[illegible]
DLTSSGFL
```



Homology Modeling – General Procedure



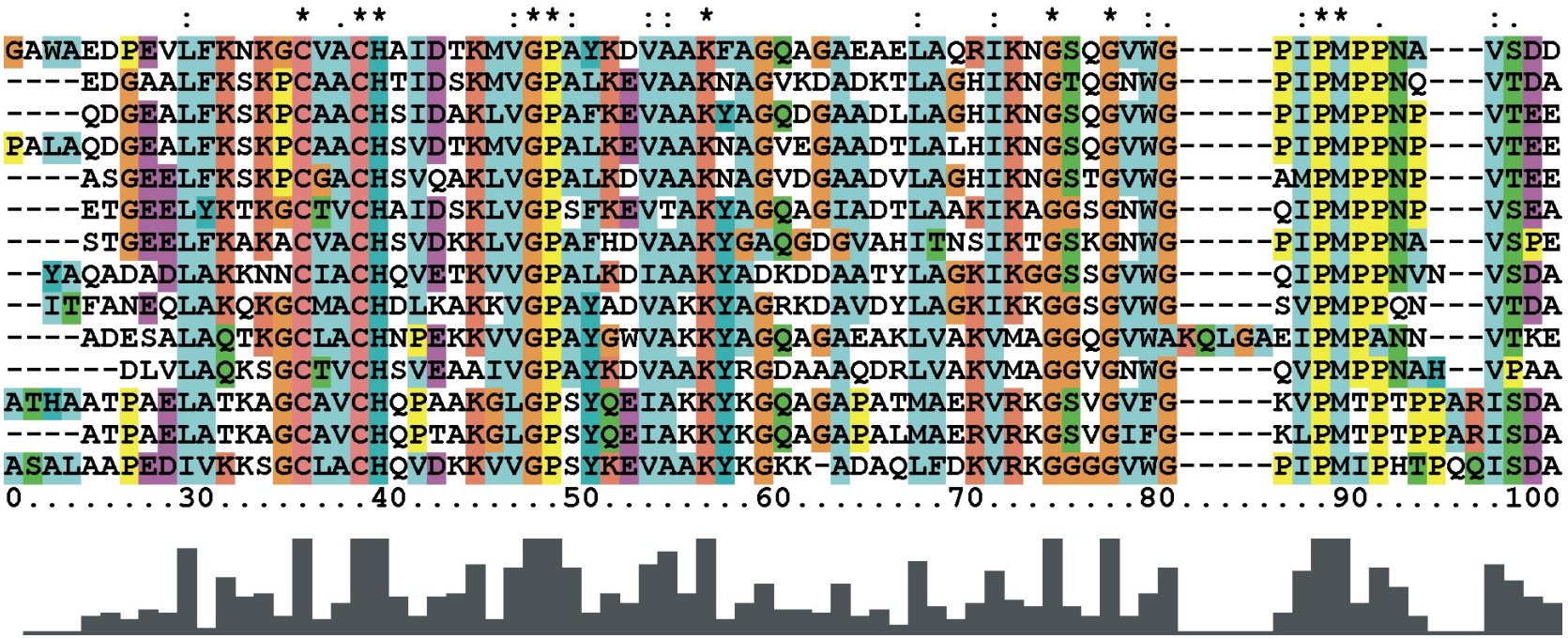
bisqit.pasteur.fr

1. Determination of the sequence homology
2. Alignment of the sequence
3. Identification of structurally invariant and variable regions

4. Construction of the backbone in the conserved regions (*core structure*)
5. Construction of the backbone in the variable regions (*loops*)
6. Generation of the sidechains
7. Model optimization (MM, MD)

8. Model Validation

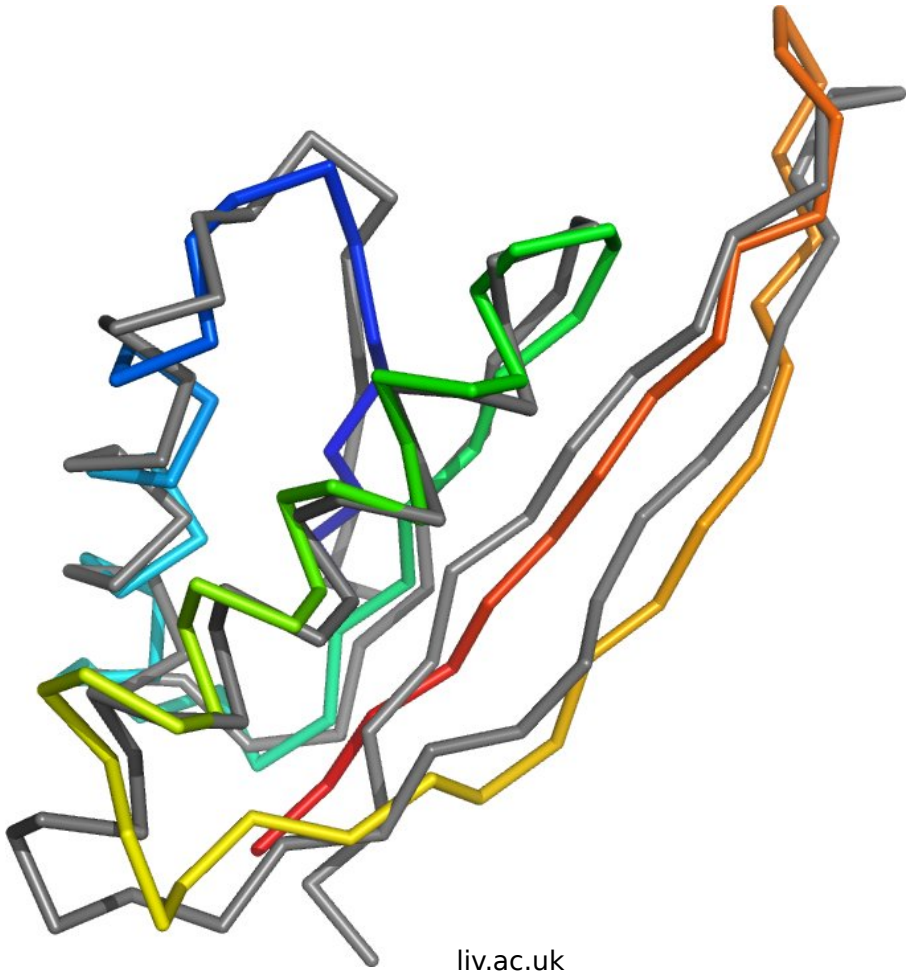
Comparison & Alignment of Primary Sequence



cerm.unifi.it

* Identical amino acids
• very similar amino acids (e.g. Ala and Gly)
: similar amino acids (e.g. Val and Ile or Phe and Trp)

Multiple Alignment of Primary Sequence



Swiss Model Server

Modelling
- myWorkspace
- Automated Mode
- Alignment Mode
- Project Mode

Tools
- Template Identification
- Domain Annotation
- Structure Assessment
- Template Library

Repository
- Search by Sequence
- Search by AC
- Search by full text

Documentation
- SWISS-MODEL Workspace
- SWISS-MODEL Repository
- Structures & Models
- Helpdesk

SWISS-MODEL is a fully automated protein structure homology-modeling server, accessible via the ExPASy web server, or from the program DeepView (Swiss Pdb-Viewer). The purpose of this server is to make Protein Modelling accessible to all biochemists and molecular biologists WorldWide.

What's new?
- SWISS-MODEL is running on new hardware with better performance
- Find more news on **SWISS-MODEL Blog**

SWISS-MODEL Team

Torsten Schwede:	Project Leader
Florian Kiefer:	SWISS-MODEL Repository
Lorenza Bordoli:	Method Development and user support
Konstantin Arnold:	SWISS-MODEL Workspace

References:
When you publish or report results using SWISS-MODEL, please cite the relevant publications:

- Arnold K., Bordoli L., Kopp J., and Schwede T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22,195-201.
- Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T (2009). The SWISS-MODEL Repository and associated resources. Nucleic Acids Research. 37, D387-D392.
- Peitsch, M. C. (1995) Protein modeling by E-mail Bio/Technology 13: 658-660.

Example: Structure-based Design at Siglec-9

Sialyl-acid binding protein from the family of immunoglobulines

Homologous Protein = Siglec-7
(PDB Code: 2G5R; Resolution: 1.60 Å)

Sequence homology = 78.8%

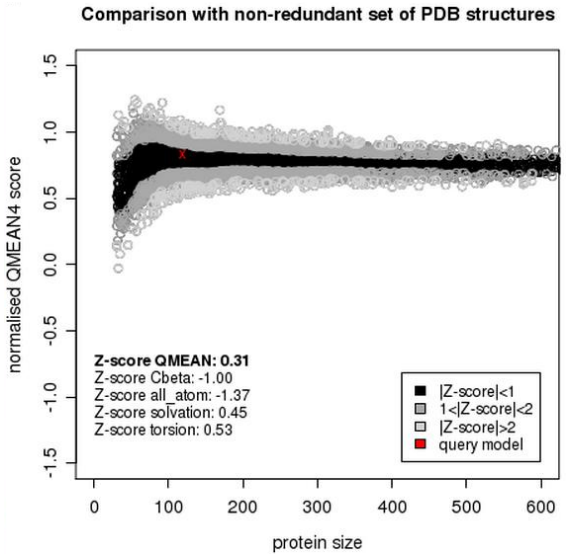
Model Building: Swiss Model Server (2012)

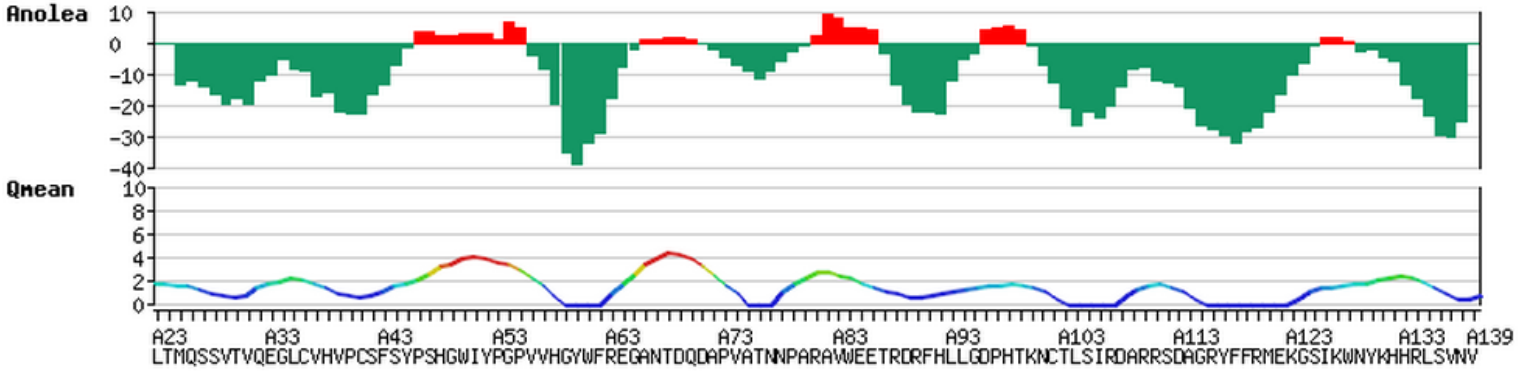
Model information:

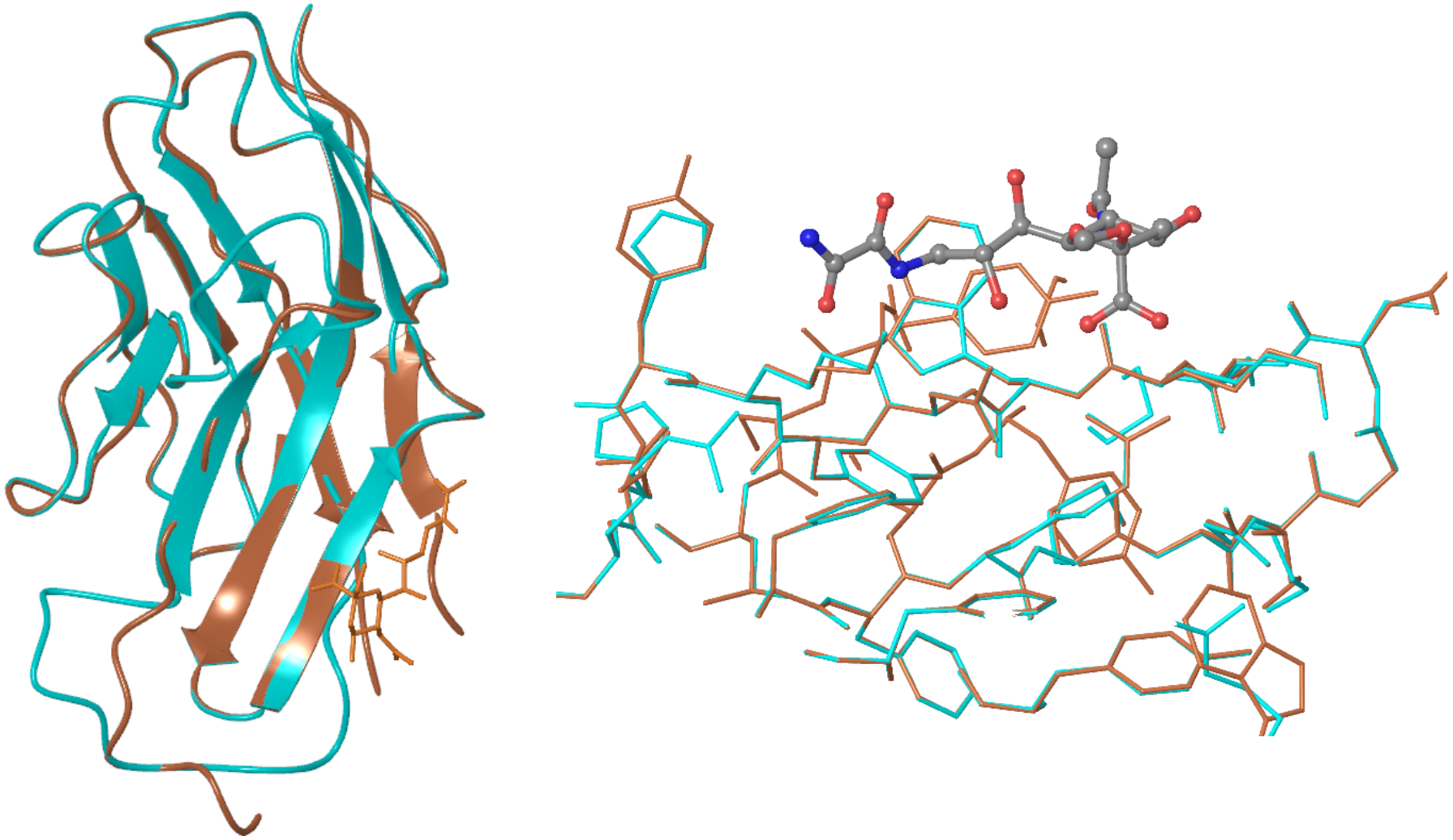
Modelled residue range:	23 to 140
Based on template:	[2g5rA] (1.60 Å)
Sequence Identity [%]:	78.81
Evalue:	3.92e-46

Quality information: [details]

QMEAN Z-Score: 0.31







Superposition of the 3D-Structures of Siglec-9 (homology model, orange) and Siglec-7 (experimental structure, cyan)

Warum braucht es in silico Toxikologie?

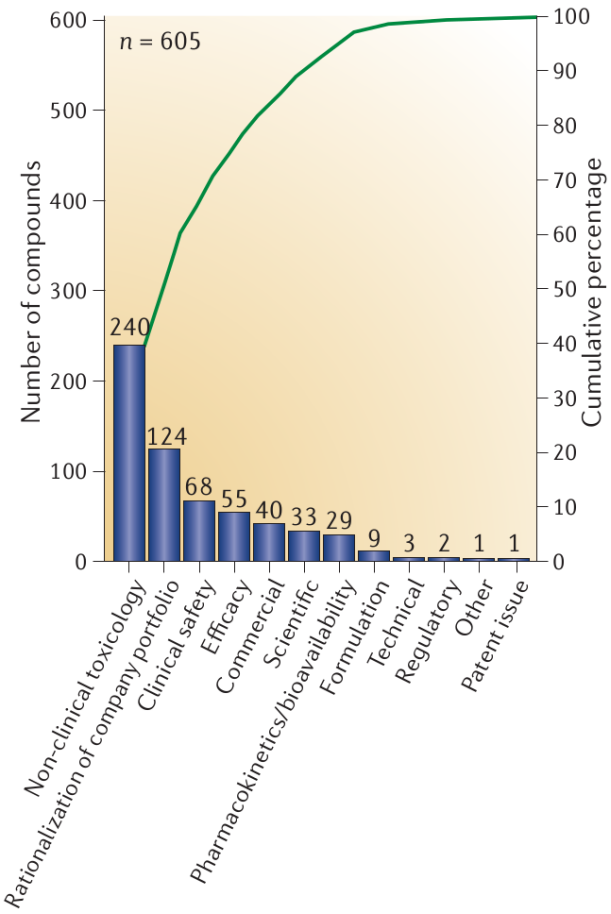
- Vor der Massenproduktion muss jede Substanz detailliert charakterisiert und getestet sein:
  - Arzneistoffe
  - Kosmetika (UV Filter, Parfum..)
  - Additiva (Polymer, Flame Retardants...)
  - Agrochemikalien
  - Farbstoffe & Pigmente

- 3R (Reduction, Replacement, Refinement)

- Regulatorische Behörden EC, EPA... (REACH Programm)

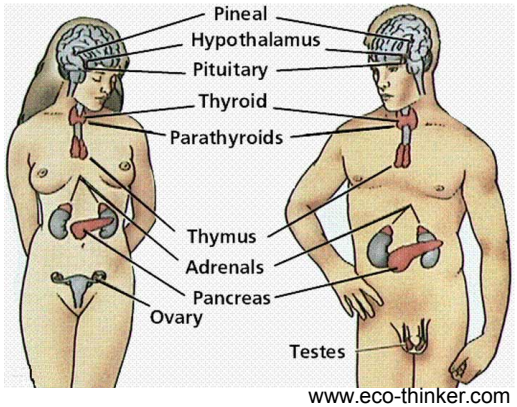
- gesammelte Kenntnisse kann man nutzen um die toxische Phänomena zu erklären und vermeiden

- Drug Attrition Rates

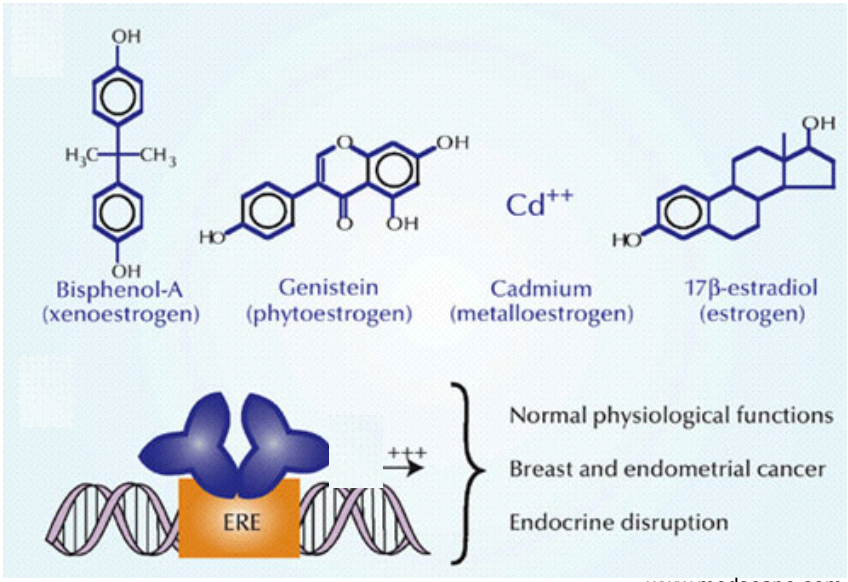


Waring M.J. et al. *Nature Reviews: Drug Discovery* (2015), 14, 475.

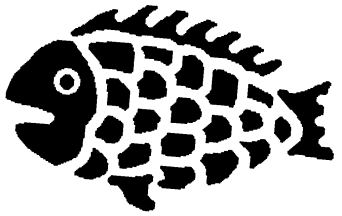
Endokrine Disruptoren — "Tarnkappenchemikalien"



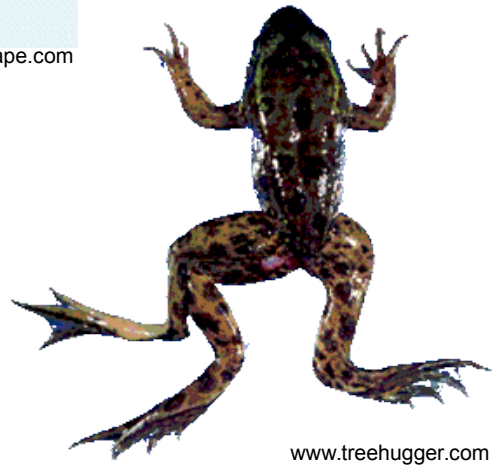
www.eco-thinker.com



www.medscape.com



www.nutri-notes.com

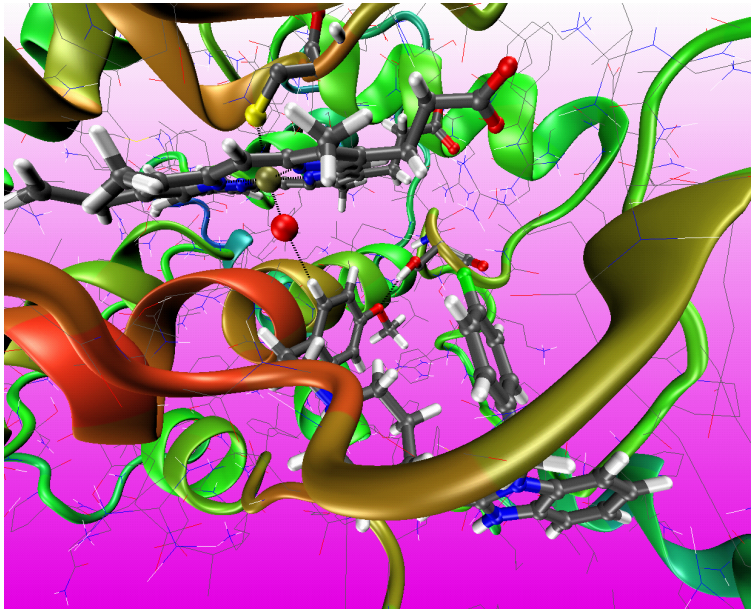


www.treehugger.com

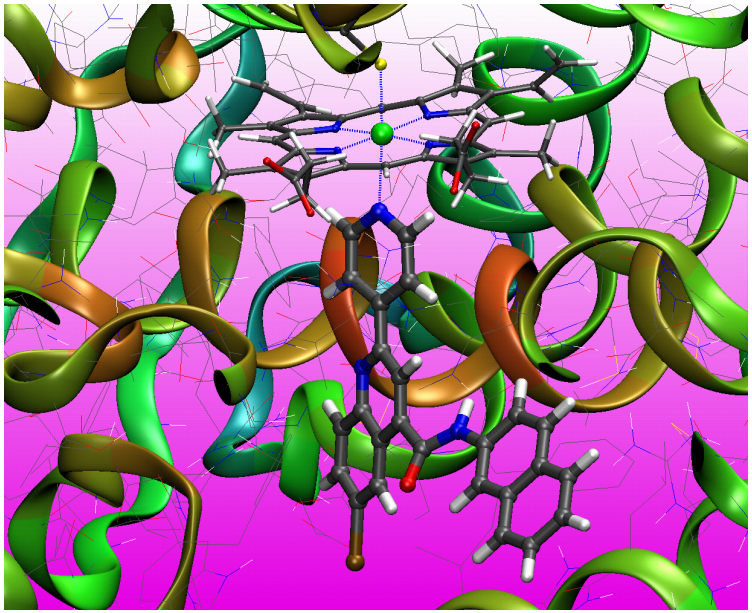
Endokrine Disruptoren werden auch als endokrin wirksame Substanzen (EDCs) oder Umwelthormone bezeichnet. Es handelt sich um natürliche (zum Beispiel Phytoestrogene) oder synthetisch hergestellte chemische Verbindungen, welche in die Umwelt gelangen (z.B. ins Abwasser). Da sie dort möglicherweise Langzeit-Schäden bewirken und durch Bioakkumulation Gesundheitsgefahren für Tiere und Menschen darstellen können (u.a. Verweiblichung), und ihre tatsächliche Bedeutung für den tierischen und humanen Stoffwechsel noch weitgehend unerforscht ist, werden sie seit einigen Jahren in der Öffentlichkeit sowie von Wissenschaft und Politik kontrovers diskutiert.

de.wikipedia.org

Metabolische Disruptoren

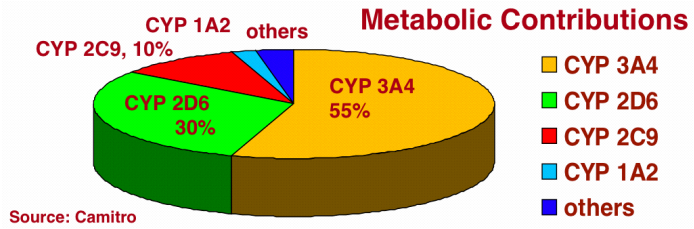


Verbindung wirkt als Substrat von CYP 3A4

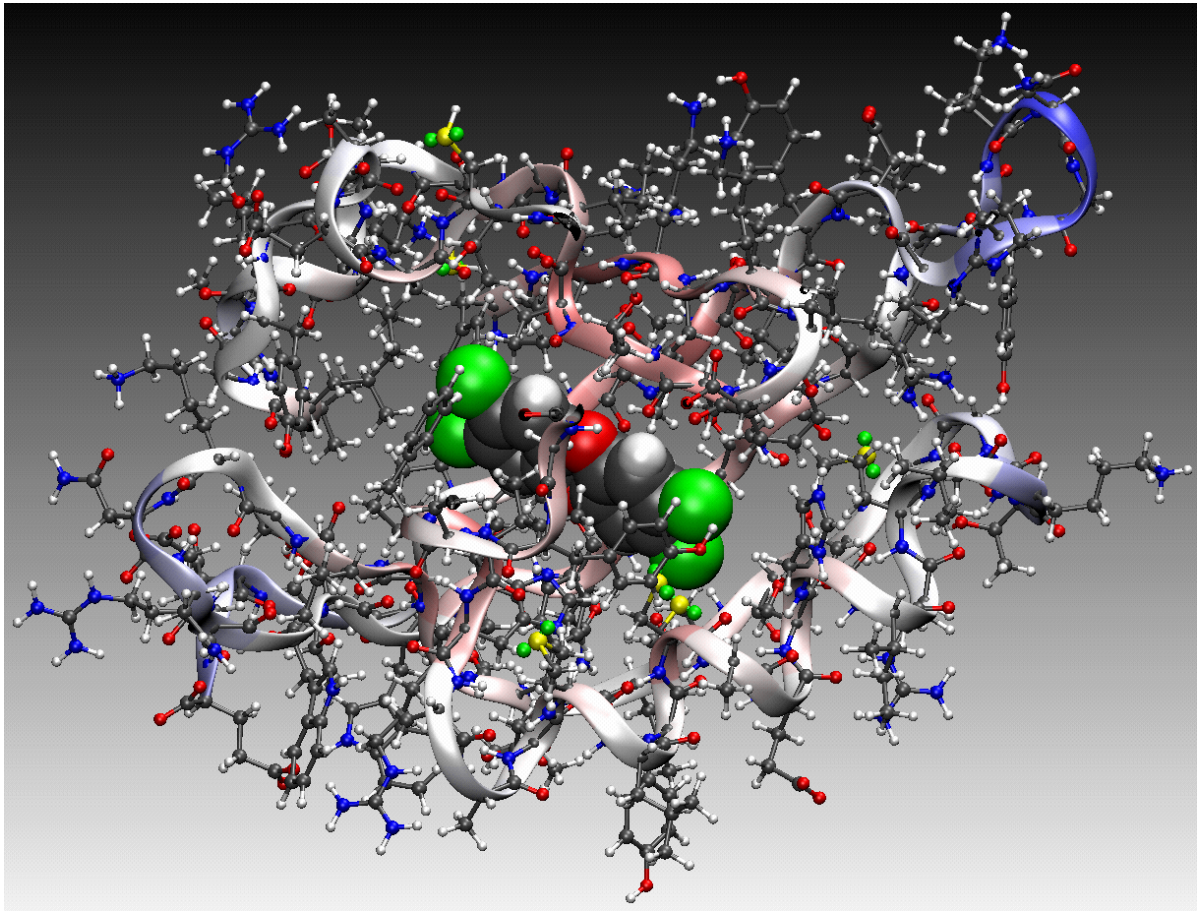


Verbindung wirkt als Inhibitor von CYP 2C9

Metabolische Störungen (metabolic disruption) bezieht sich auf die Wechselwirkung einer Substanz mit metabolisierenden Proteinen — am wichtigsten sind die Enzyme der Cytochrom P450 Familie, beispielsweise CYP 1A2, CYP 2C9, CYP 2D6 und CYP 3A4.



Wechselwirkungen mit dem Aryl hydrocarbon Rezeptor



Homologie-Model des Aryl hydrocarbon Rezeptors (Swiss Model Server, Universität Basel)

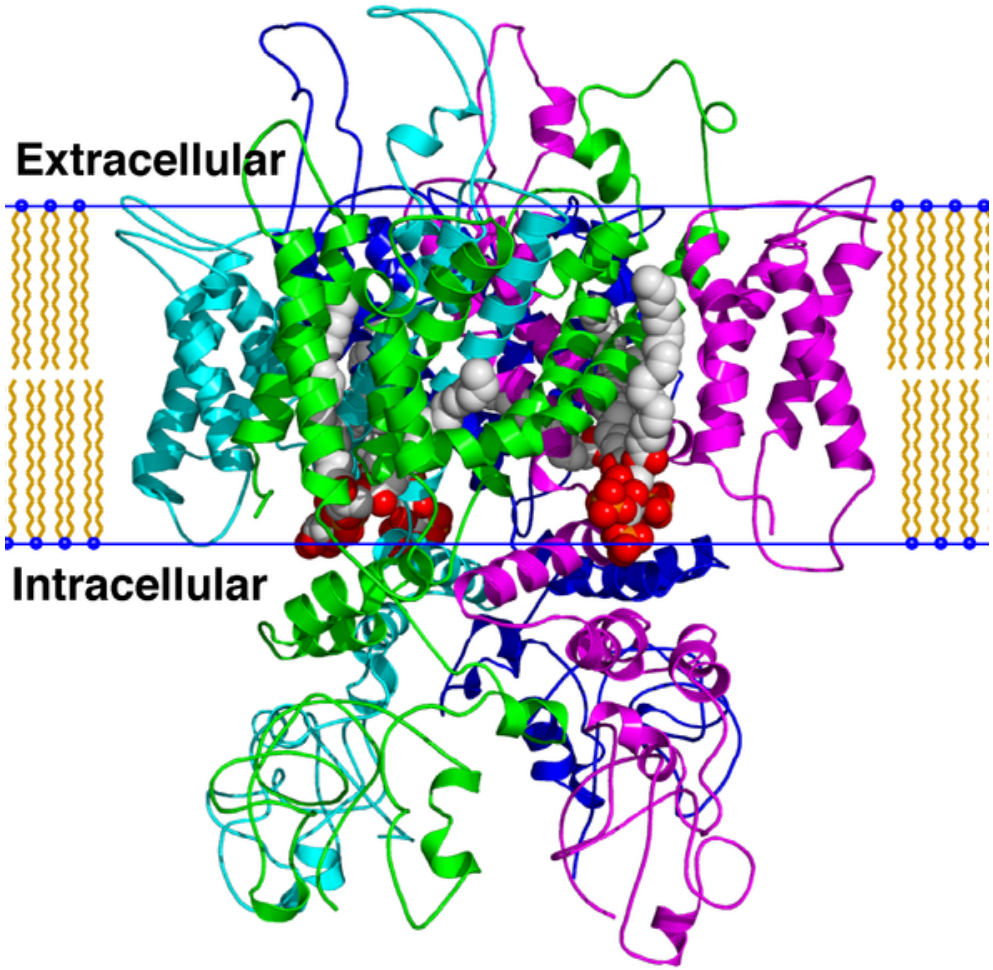
Wechselwirkung mit dem hERG-Ionenkanal

- hERG (human Ether-à-go-go Related Gene)
- Spannungsaktivierter, einwärtsgleich-richtender Kaliumkanal in Herzmuskel-zellen

- Arzneistoffe, die hERG blockieren können Herzrhythmusstörungen auslösen → long QT syndrome

- Eine Anzahl klinisch erfolreicher Arzneistoffe hemmen hERG und lösen unerwünschte Effekte aus. Sie mussten daher vom Markt genommen werden: *Astimizol, Cisaprid, Grepafloxacin, Sertindol und Terfenadin*

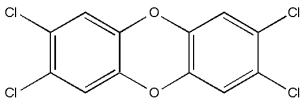
- In der pharmazeutischen Forschung/ Entwicklung ist hERG ein **Antitarget**, d.h. jeder Wirkstoffkandidat, der an hERG zu binden vermag, wird unverzüglich aus der Evaluations-Pipeline genommen



https://en.wikipedia.org/

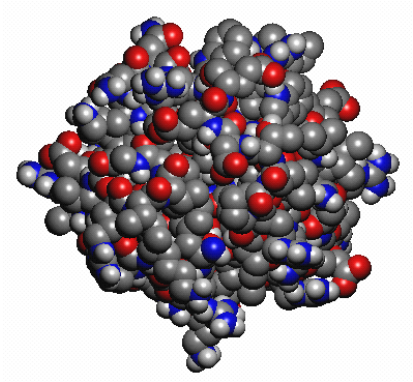
Simulation von Rezeptor-vermittelten Nebenwirkungen

TCDD



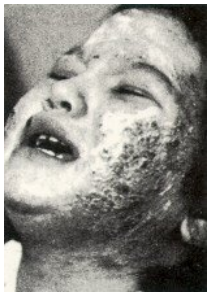
Einnahme/Hautkontakt

Bindung

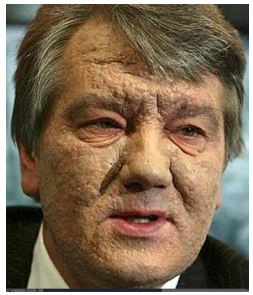


Signalweiterleitung

Manifestation der Toxizität

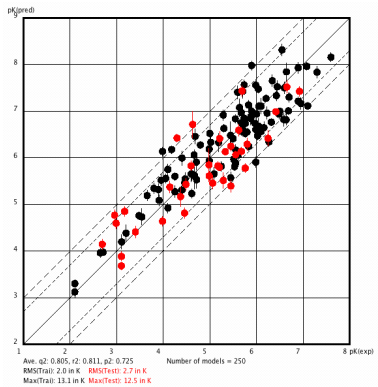


Seveso-Opfer 1976
www.brooklyn.cuny.edu

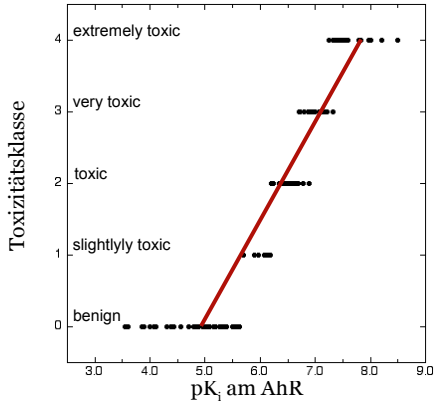


V. Juschtschenko 2004
www.blick.ch

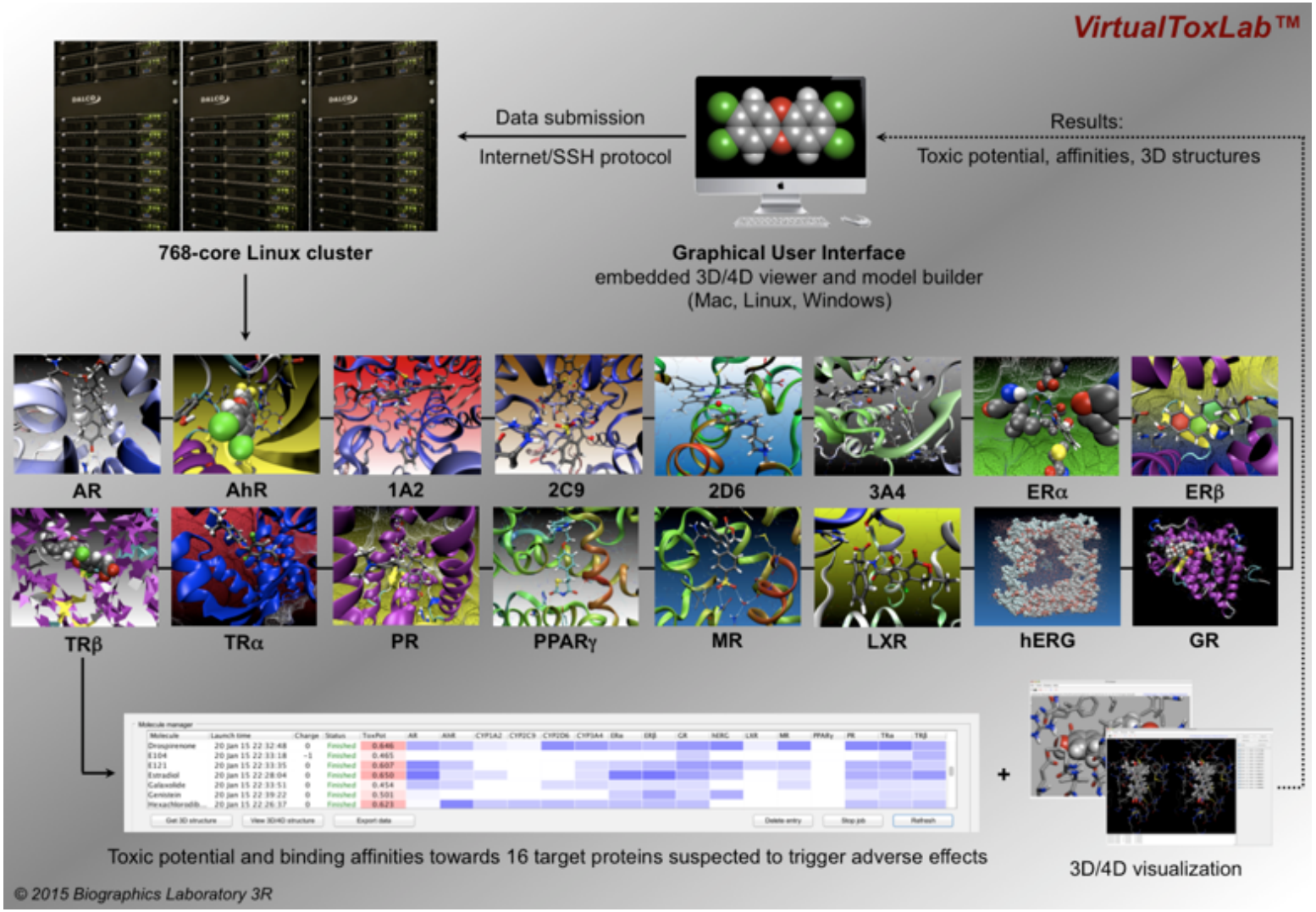
Simulation der Bindung ans Zielprotein (AhR)



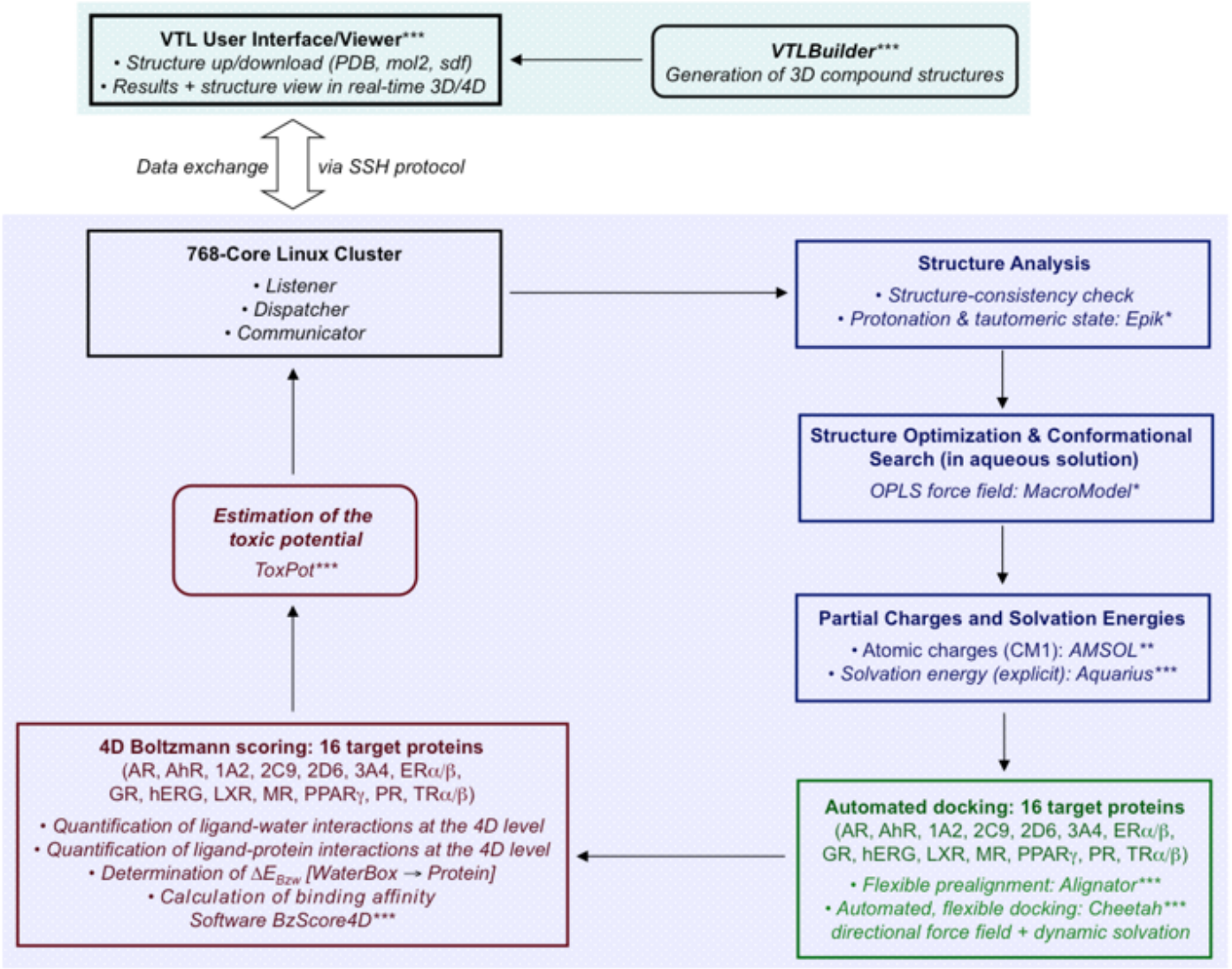
Konsequenz: Aus der Quantifizierung der Bindungsaffintiät zu einem Protein, das unerwünschte Wirkungen vermittelt lässt sich das "toxische Potential" der Verbindung, nicht aber deren Toxizität abschätzen.



VirtualToxLab — Symbolisches Flussdiagramm



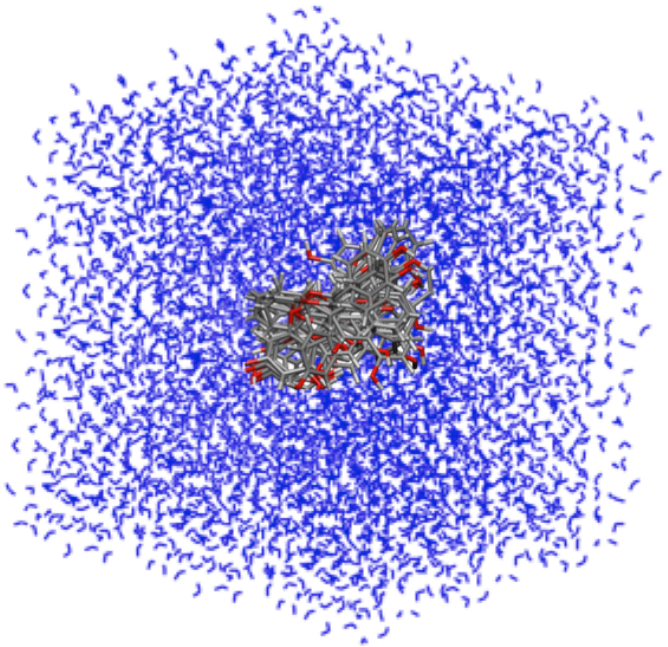
VirtualToxLab — Technisches Flussdiagramm

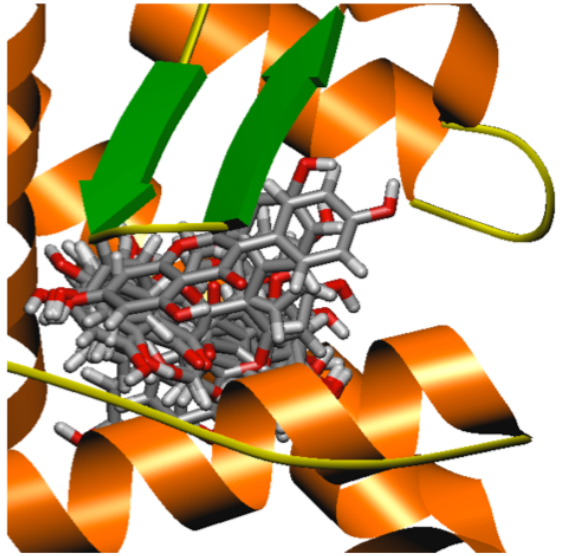


VirtualToxLab — Boltzmann Bewertung (4D)

Direkte Bewertung des Wasser-Protein-Übergangs einer Wirksubstanz (4D Bolzmann-Ensemble)

Beispiel: Genistein →Estrogen receptor β



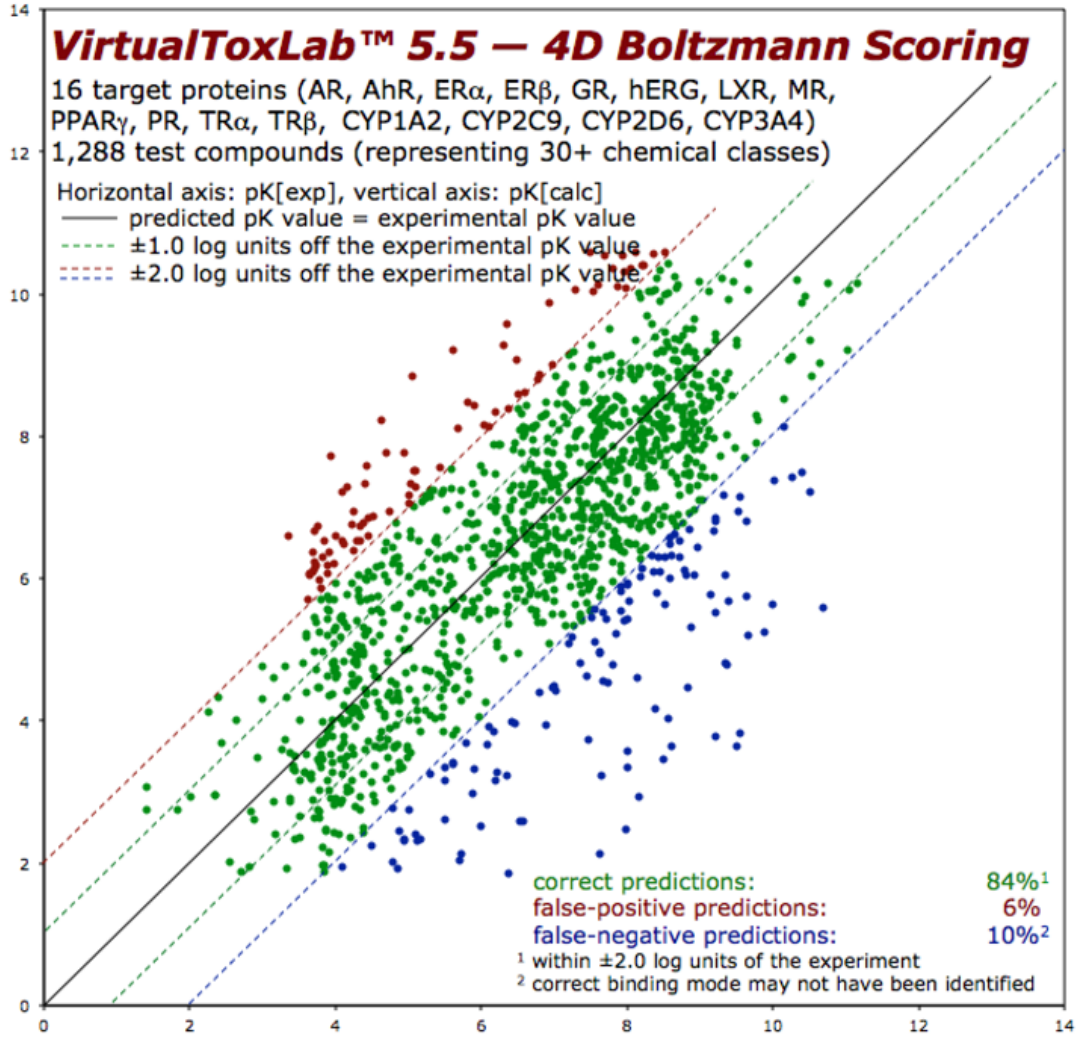


Water-bound ligand: E_{Bzw} = –36.262 kcal/mol

Protein-bound ligand: E_{Bzw} = –47.006 kcal/mol

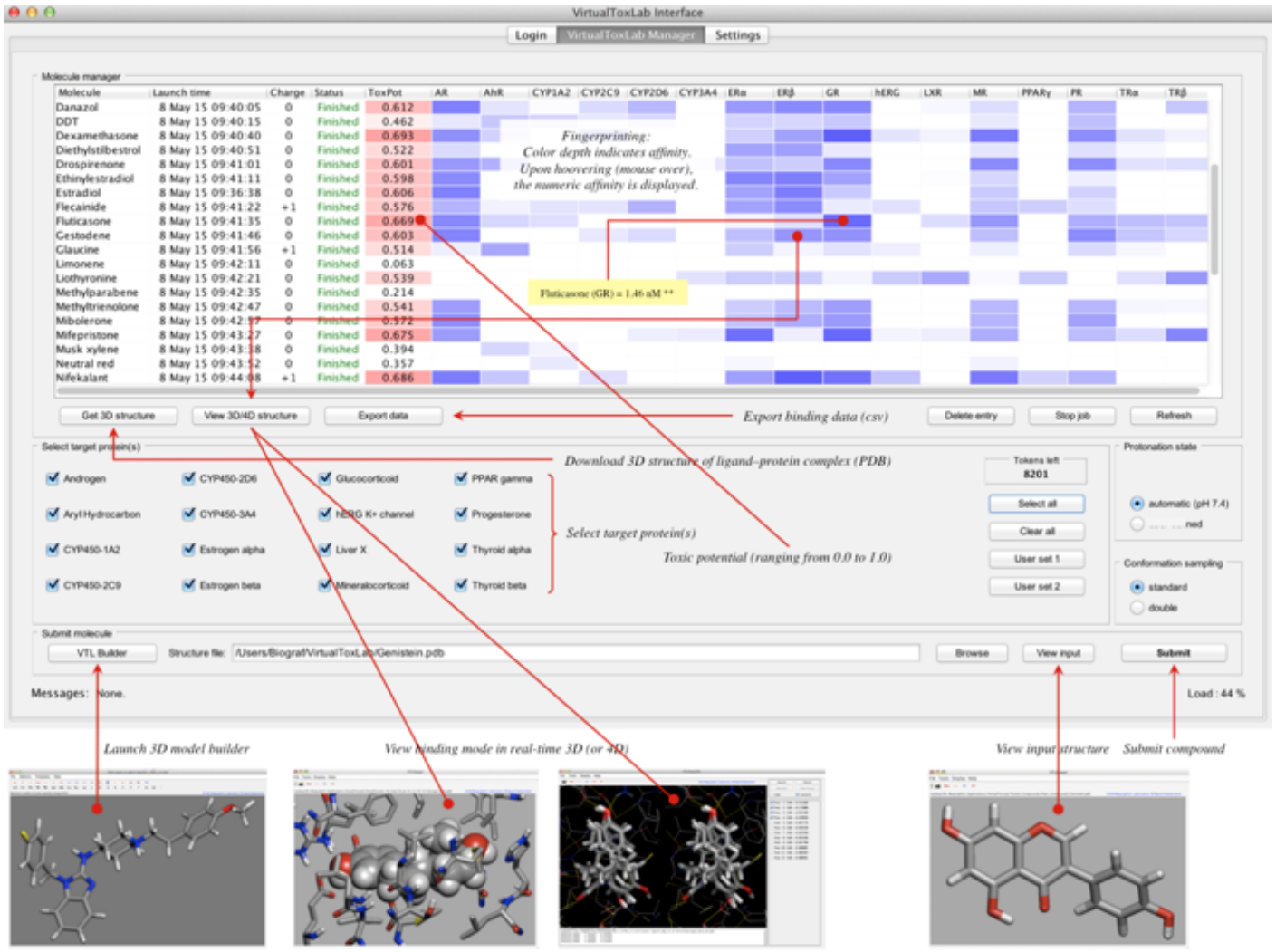
ΔE_{Bzw} = –10.774 kcal/mol → converting ΔE into K using: $K = e^{-\Delta E / RT}$ → **K_{calc} = 9.66×10^{-9} M** (K_{exp} = 1.2×10^{-8} M)

VirtualToxLab — Validierung
(1,288 Substanzen)

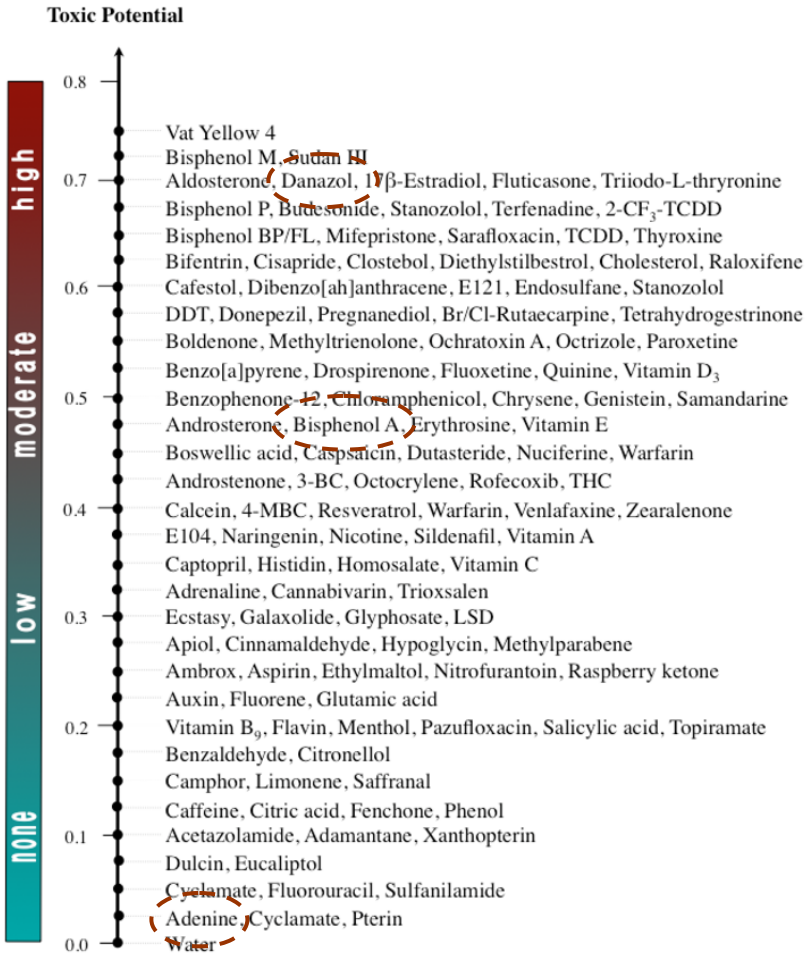


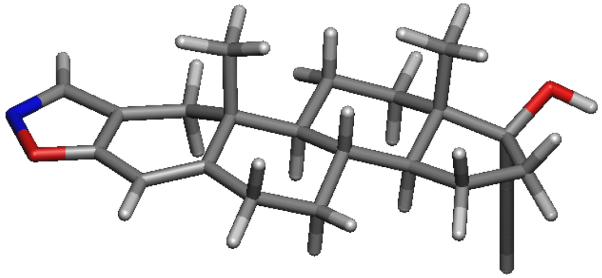
M. Smieško & A. Vedani — Departement Pharmazeutische Wissenschaften, Universität Basel, 2017

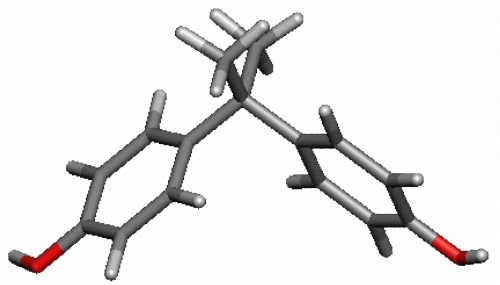
VirtualToxLab — Grafische Schnittstelle (Interface)

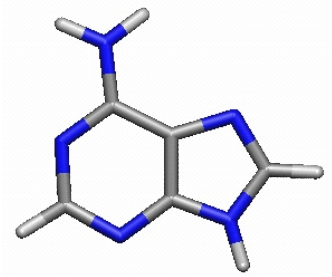


VirtualToxLab — "Toxicity Alerts" (Warnhinweise)

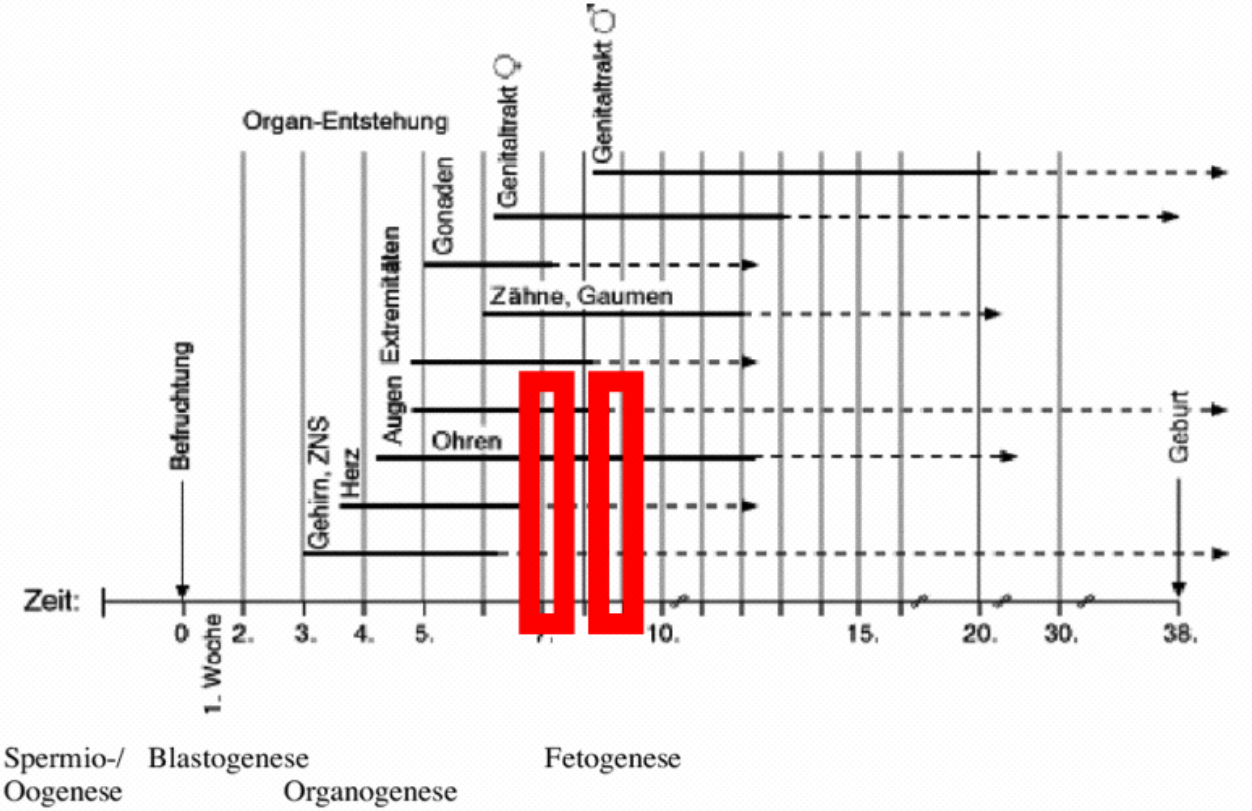








Endokrine Disruptoren sind “Tarnkappenchemikalien”

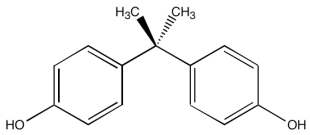


Concentrations as low as 1/4000 of today’s (2008) analytical detection limit of endocrine disruptors (e.g. Bisphenol A) may trigger severe adverse effects during fetus genesis

Switzerland — Investments for adequate sewage system upgrade (charcoal filters, O_3 treatment). In 2013, the Swiss Federal Government proposed to upgrade 100 (of the total 700) sewage plants to the newest C/O_3 technology for CHF 1,200,000,000.00 plus annual costs of CHF 9.00 per inhabitant.

VirtualToxLab — Erstellen von Affinitätsprofilen

Bisphenol A (Weichmacher)



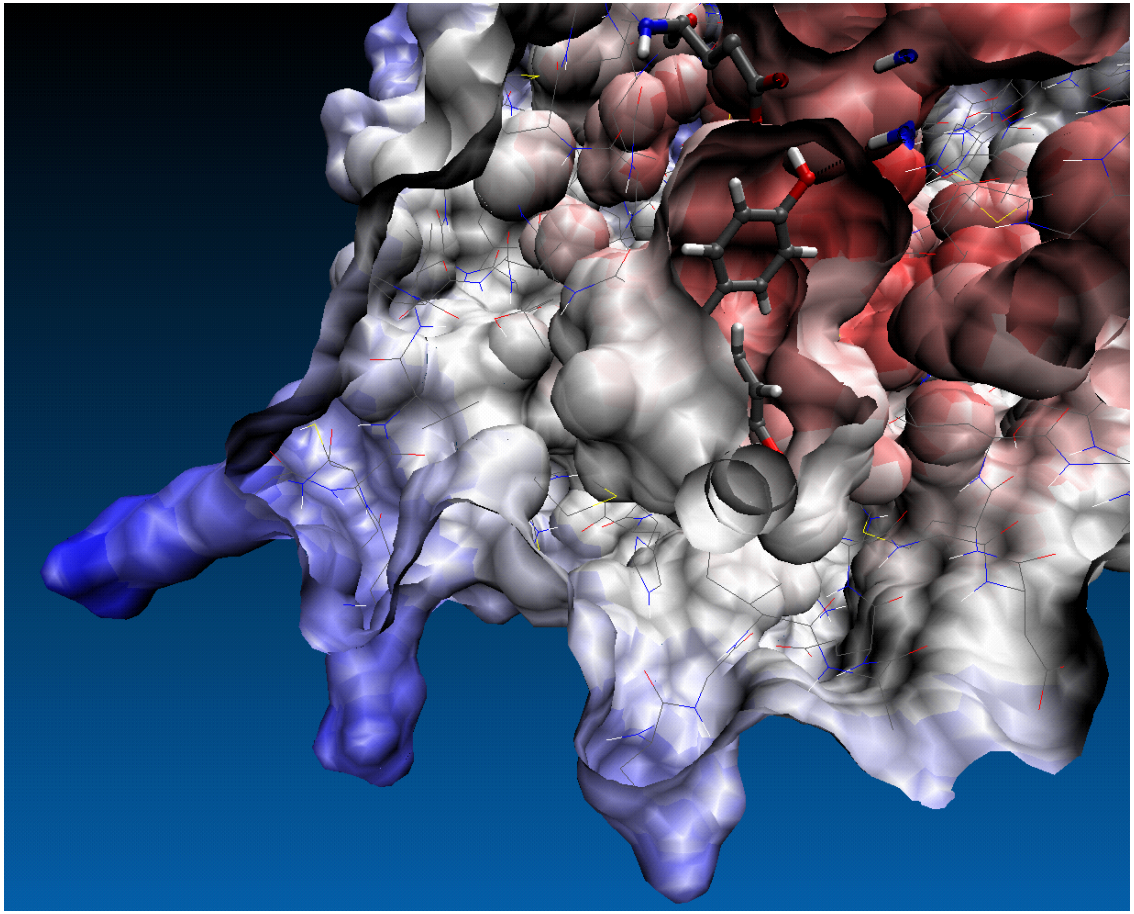
AR: 240 nM
ERα: 260 nM
ERβ: 120 nM
GR: 820 nM
LXR: keine Bindung
MR: 6.2 µM
PPARγ: keine Bindung
PR: 6.9 µM
TRα: 37 µM
TRβ: 19 µM

AhR: 130 nM

hERG: keine Bindung

CYP1A2: keine Bindung
CYP2C9: keine Bindung
CYP2D6: 6.6 µM
CYP3A4: keine Bindung

Toxic Potential = 0.478

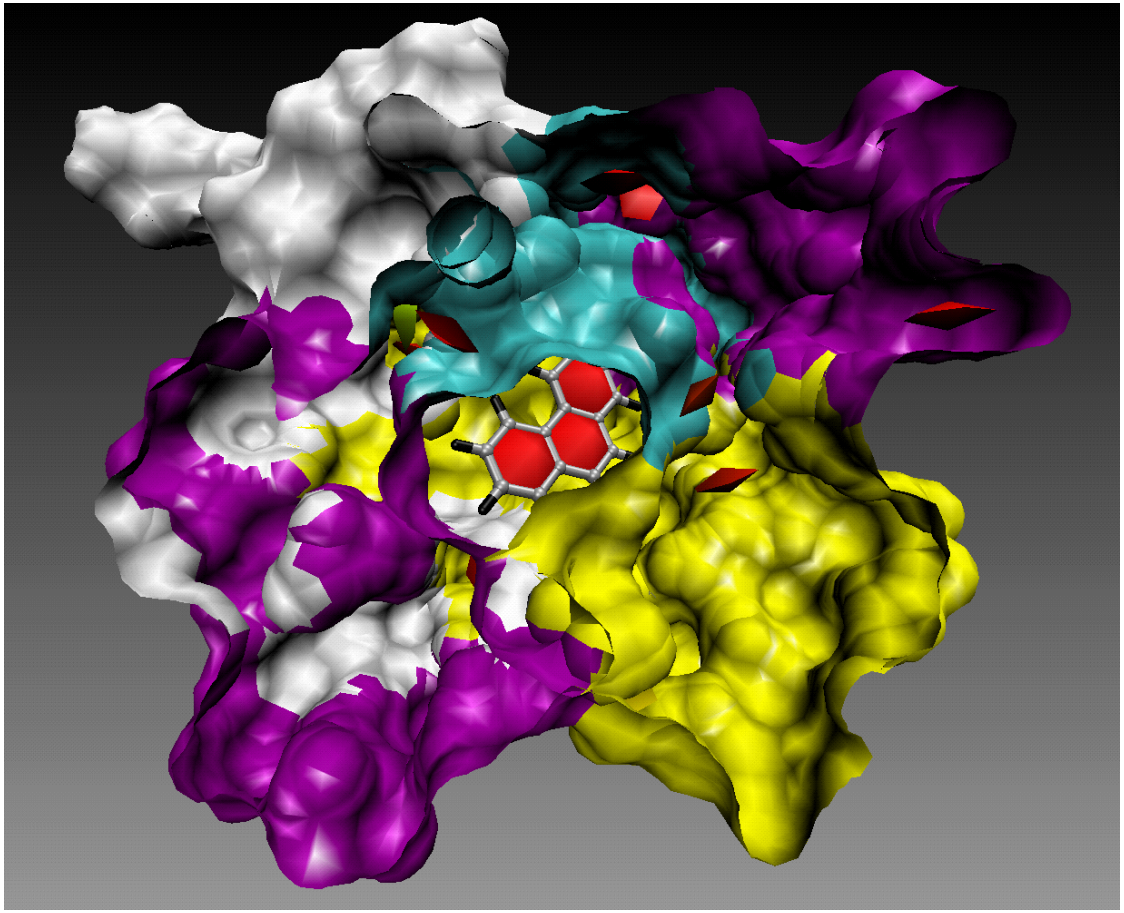


ERβ: Berechnete Bindungsaffinität = 120 nM (Experimentell = 93 nM)

VirtualToxLab — Screening von Umweltchemikalien

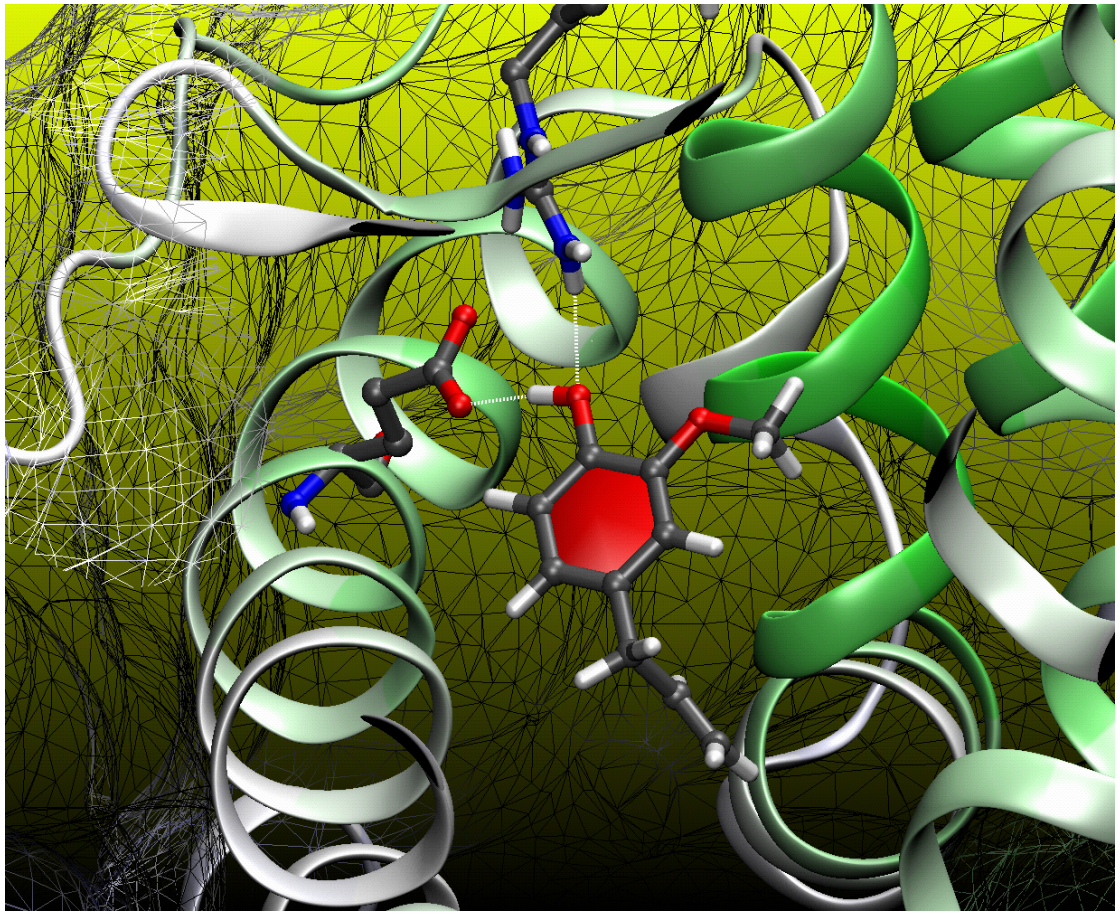
Toxic potential (0.0 – 1.0):

Benzo[a]anthracen = 0.588
Benzo[a]pyren = 0.532
Benzlidene camphor = 0.416
Bisphenol A = 0.478
Bisphenol BP = 0.648
Bisphenol M = 0.712
Chrysen = 0.494
Coumestrol = 0.564
DDT = 0.552
Dibenzo[ah]anthracen = 0.638
17β-Estradiol = 0.667
Gestrinon = 0.540
Hexachlorodibenzofuran = 0.640
Hexabromodiphenyläther = 0.583
3-Methylcholanthren = 0.533
Ovalen = 0.719

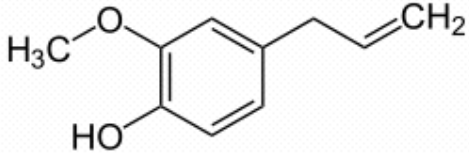


Bindung von Dibenzo[ah]anthracen an den AhR: 13.1 nM (exp. = 16 nM)

VirtualToxLab — Screening von kosmetischen Wirkstoffen

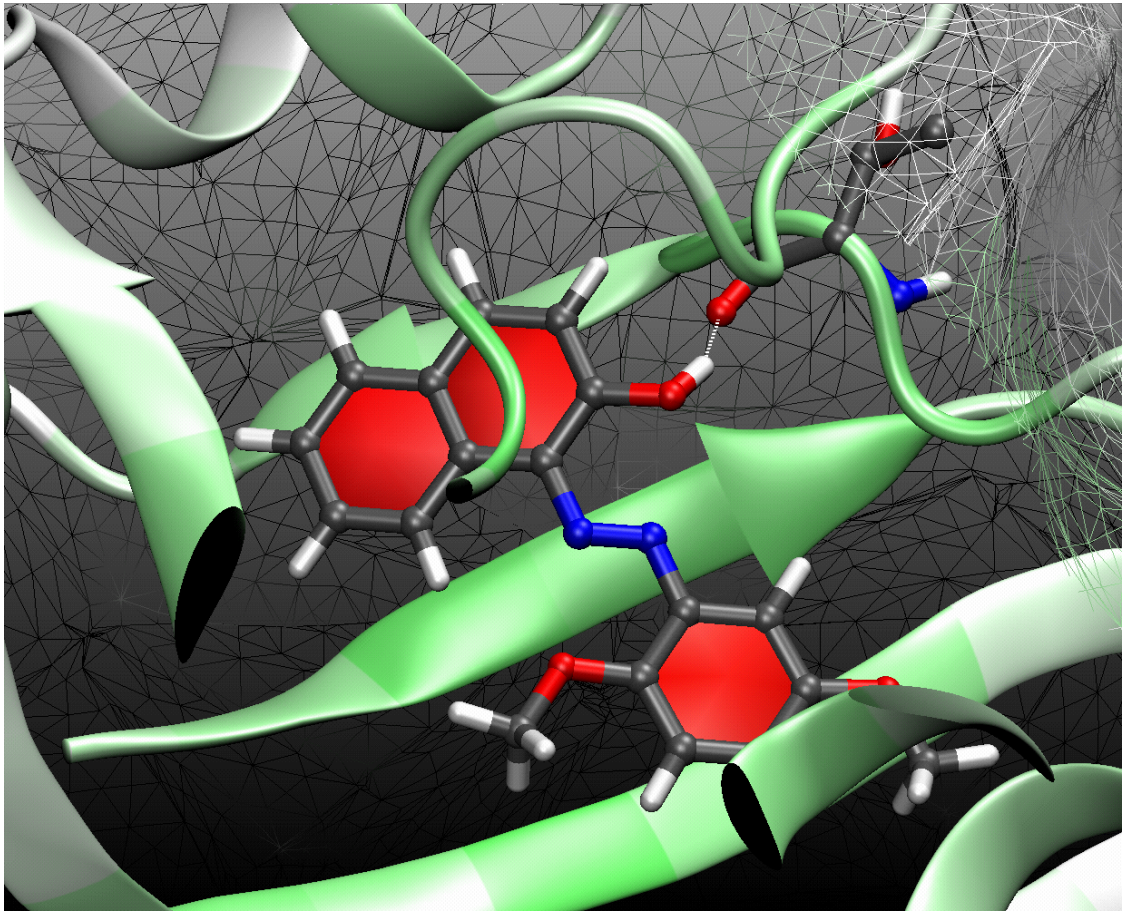


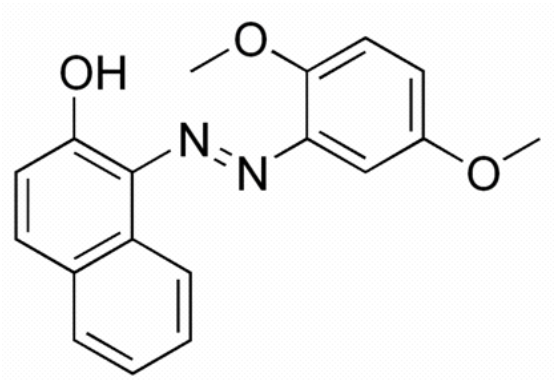
Bindung von Eugenol (Duftstoff) an den Östrogenrezeptor β: Berechnete Affinität = 16 µM



ToxPot = 0.415

VirtualToxLab — Screening von Lebensmittelinhaltsstoffen



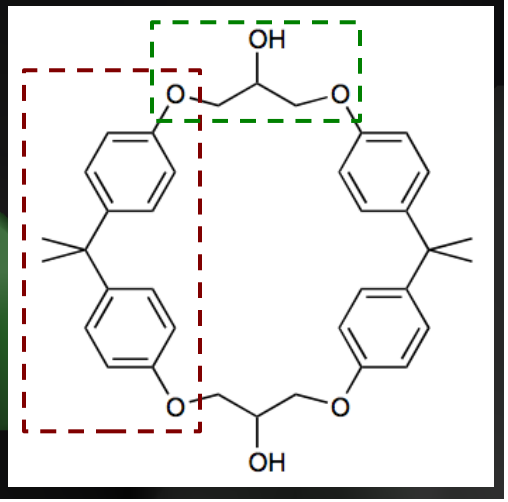


ToxPot = 0.504

Bindung von Citrus Red 2 (E121) an den Aryl hydrocarbon Rezeptor: Berechnete Affinität = 92 nM

Cyclo-diBA — A possibly toxic contamination in (fish) can coatings

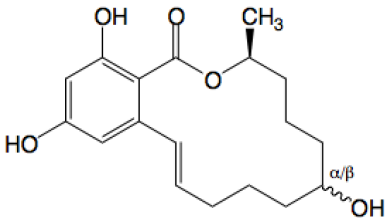
Cyclo-diBA, the cyclic product formed from bisphenol A and diglycidyl ether during the production of epoxy resins. Half of the samples of canned fish in oil collected in April 2010 contained cyclo-diBA with an average concentration of 1025 µg/kg and a maximum of 1980 µg/kg. The majority of the canned meat products contained cyclo-diBA at a mean concentration of 477 µg/kg and a maximum of 1050 µg/kg. All prepared meals, such as ravioli or soups, contained cyclo-diBA with a mean at 287 µg/kg. In canned tomatoes, peas and other vegetables in water or fruits in syrup, no cyclo-diBA was detected (<25 µg/kg). Since no experimental toxicity data are available except for its cytotoxicity, an in silico hazard profiling was performed. Cyclo-diBA seems to be stable and of low reactivity. There is indication for considerable oral bioavailability and for the potential to accumulate in the human body. Cyclo-diBA can be metabolized into cyclic and acyclic compounds.



Cooperation with Dr. Beat Brüschweiler, Federal Office of Public Health, Switzerland

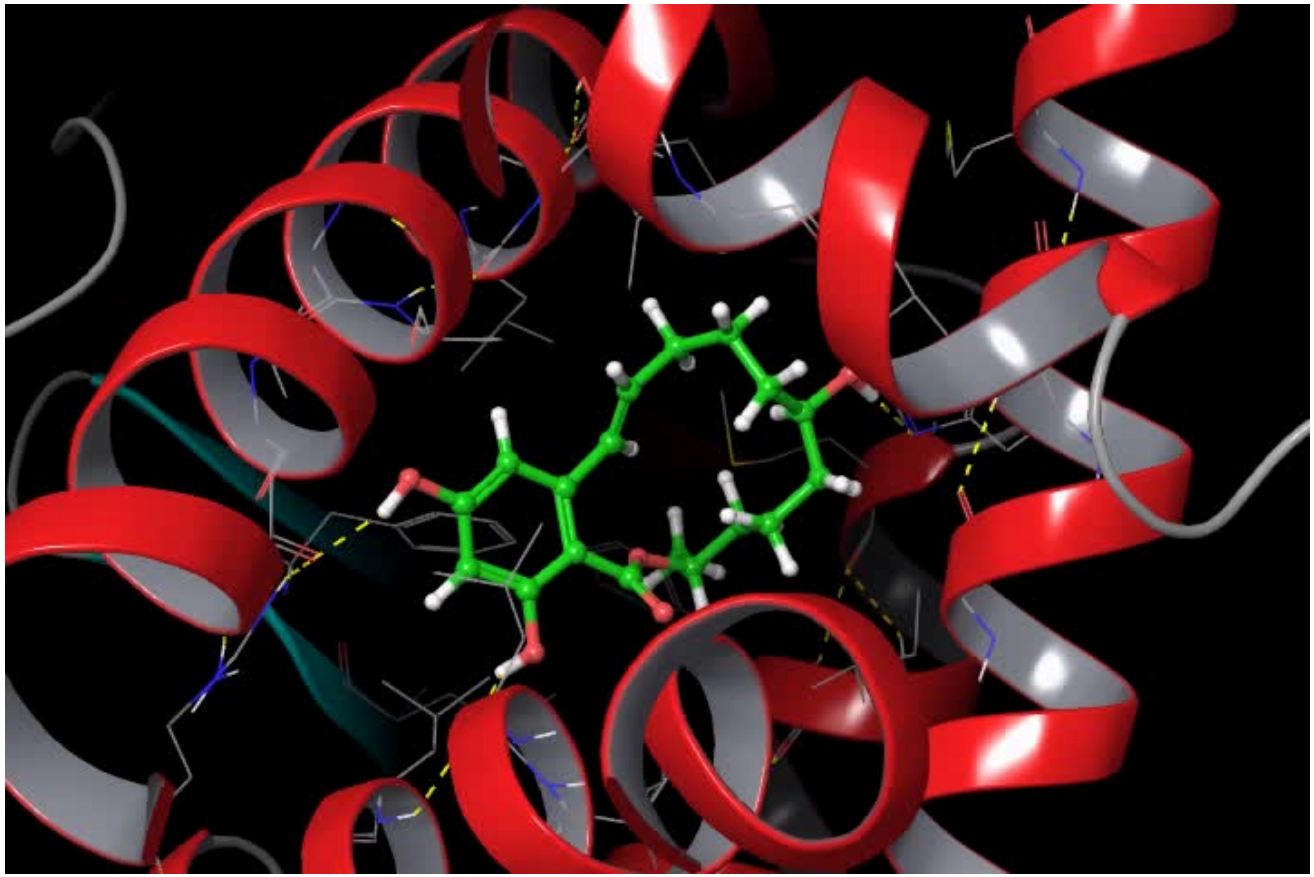
Food Chem. Toxicol. 2013, 58, 107–115

Mycotoxin α-zearalenol am Estrogenrezeptor α



α/β-Zearalenol

ERβ: 21 nM (exp = n/a)
ToxPot = 0.521



MD run using software Desmond, D.E.Shaw, New York