

VL Algorithmische Bioinformatik

Molecular Modelling: Proteine und Proteinliganden

Prof. Dr. Gerhard Wolber

Institut für Pharmazie
Freie Universität Berlin

gerhard.wolber@fu-berlin.de

<http://www.pharmazie.fu-berlin.de>
<http://www.drug-design.de>



G. Wolber, Algorithmische Bioinformatik

Inhalt

Virtuelles Screening

- Pharmakophorbasiertes virtuelles Screening
- Bewertungsmetriken (EF, ROC, AUC)
- Validierungsbeispiel
- Anwendungen

Kraftfeldtheorie und Molekulardynamik

- Bezug zu Physik und Quantenmechanik
- Molekularmechanische Ansätze
- Potenzialfunktionen für Bindungslängen, Bindungswinkel und Torsionen
- Optimierungsmethoden (Überblick)

G. Wolber, Algorithmische Bioinformatik

Inhalt

Computermodelle für Proteine und Liganden in der Arzneistoffentwicklung

- Molekulare Ähnlichkeit als Parameterraum zur Modellierung ähnlicher biologischer Aktivität

Docking und Scoring

- Systematische Enumeration & Monte Carlo Suche
- Metaheuristiken (GA, PSO)
- Scoringfunktionen

Protein-Liganden Interaktion anhand von 3D Pharmakophoren

- Strukturbasierte Pharmakophorgenerierung
- Überlagerungsverfahren und Validierung
- Ligandenbasierte Pharmakophorgenerierung

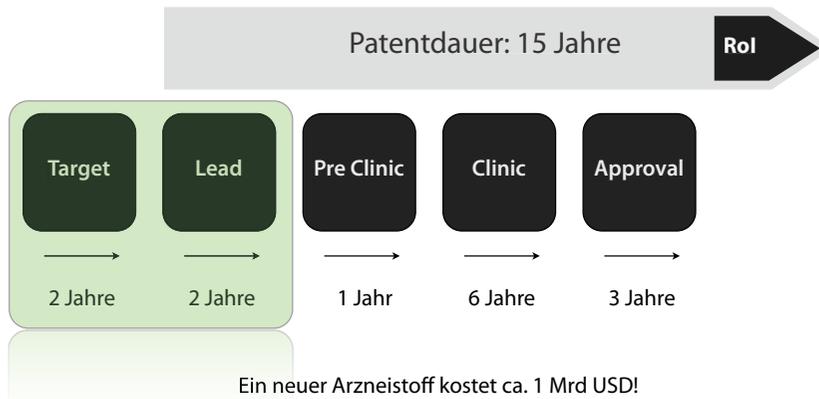
G. Wolber, Algorithmische Bioinformatik

Literatur



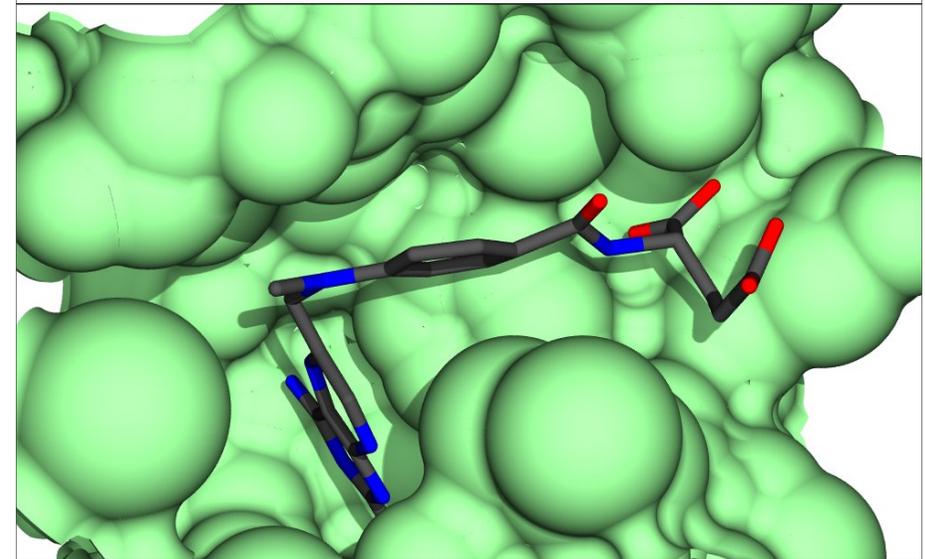
G. Wolber, Algorithmische Bioinformatik

Arzneistoffentwicklung: Proteinliganden



G. Wolber, Algorithmische Bioinformatik

Molecular Docking



Das Problem: "Molekulare Ähnlichkeit" und biologische Aktivität

Welche Moleküle "ähnlich" sind, hängt vom Kontext ab:

Auf chemischer Ebene:

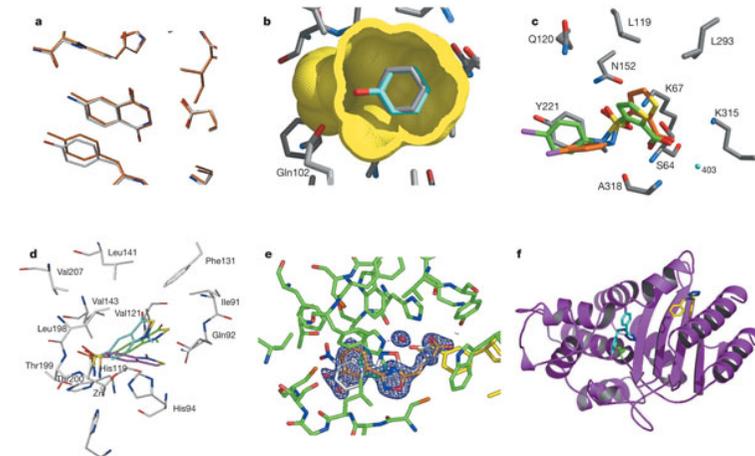
- Ähnliche Substrukturen (Graphentheorie)
- Relevant für Patentierung & Synthesep lanung

Auf biologischer Ebene:

- Verständnis für die Erkennung molekularer Wirkmechanismen (Wie interagiert ein Molekül mit einem Protein)
- Zwei Moleküle **M1** und **M2** sind in Bezug auf ein Protein **P** ähnlich, wenn sie eine ähnliche Bindeaffinität zu **P** aufweisen.
- Ähnlichkeit auf biologischer Ebene hängt von der dreidimensionalen Anordnung der funktionellen Gruppen, die an der Bindung beteiligt sind, ab
- Abhängigkeit von **P** -> nicht universell

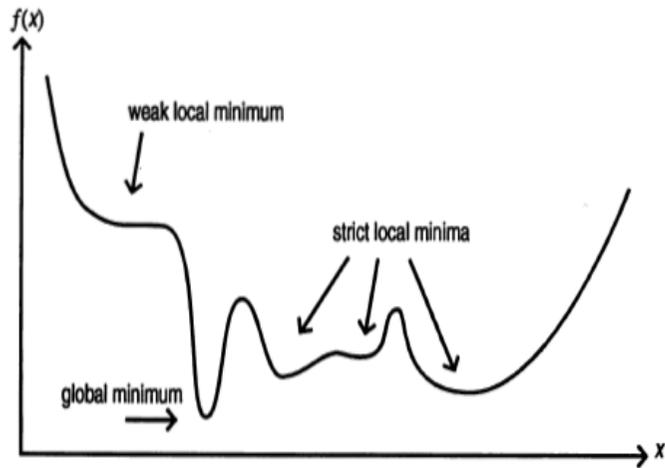
G. Wolber, Algorithmische Bioinformatik

Docking: Vorhersage von Ligandenbindung an Proteine



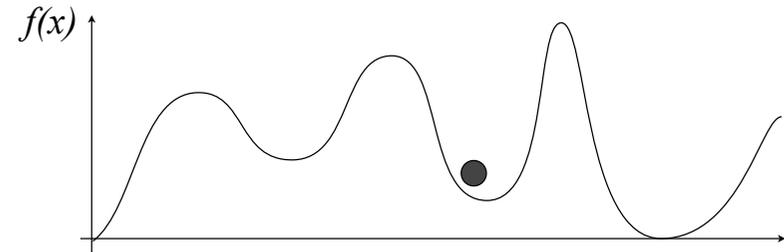
G. Wolber, Algorithmische Bioinformatik

Potenzial(Energie-)funktion



G. Wolber, Algorithmische Bioinformatik

Zufallsauswahl: Monte Carlo und Metropolis-Kriterium



$$\Delta E < 0$$

oder

$$e^{-\Delta E/RT} > X_{[0,1]}$$

$X_{[0,1]}$: Zufallszahl zwischen 0 und 1

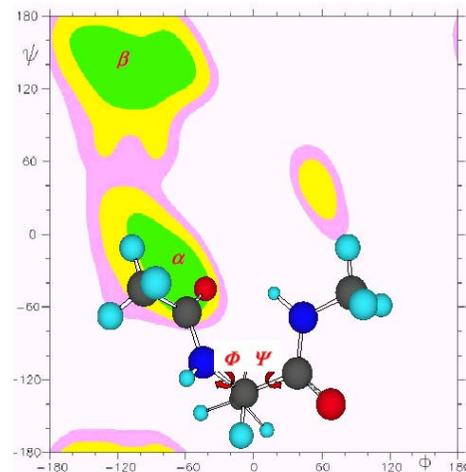
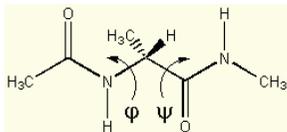


Nicolas Metropolis
(1915-1999)

G. Wolber, Algorithmische Bioinformatik

Systematische Enumeration

Alanin-Dipeptid



G. Wolber, Algorithmische Bioinformatik

Metaheuristik als Optimierungsmethode

Gerichtete Suche mit geringem
Implementierungsaufwand

Implementierte Verfahren für Docking

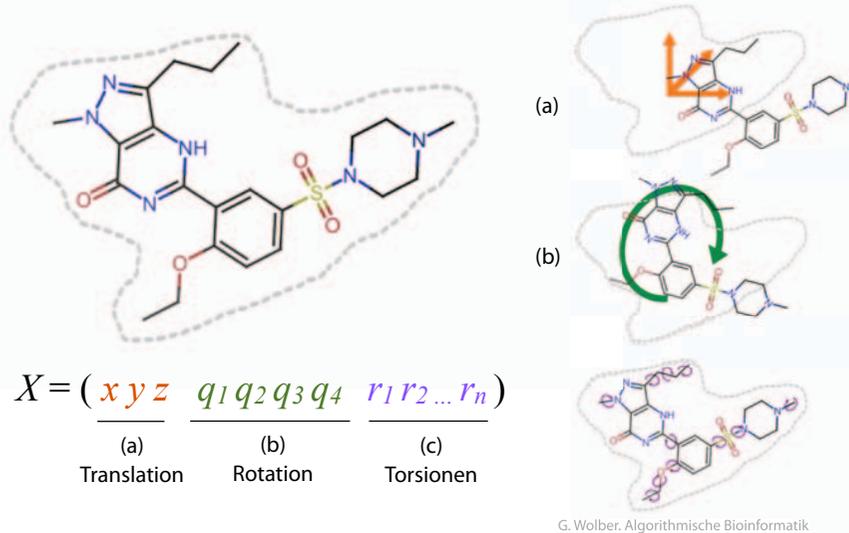
- Genetische Algorithmen (GA) - GOLD
- Particle Swarm Optimization (PSO) - ParaDocks

Kernprobleme:

- Parameterdarstellung (Konformationen, Rotation, Translation)
- Optimierungsverhalten
- Fitnessfunktion/Objective Function

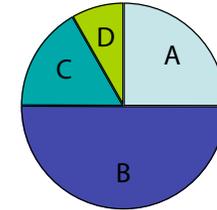
G. Wolber, Algorithmische Bioinformatik

Metaheuristik: Freiheitsgrade



1. Auswahl der fittesten Chromosomen

	Fitness	Roulette-Anteil
A	3	1/4
B	6	1/2
C	2	1/6
D	1	1/12



2. Anwendung von GA-Operatoren

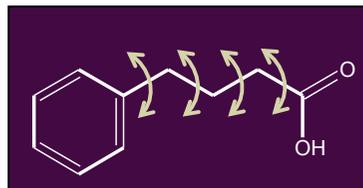
Crossover (80%) $1000100011110001 \rightarrow 1000100011001100$
 $0011000011001100 \rightarrow 0011000011110001$

Punktmutation (1%) $1000100011110001 \rightarrow 1000100010110001$

3. Ersetzen der zwei unfittesten Chromosomen

G. Wolber, Algorithmische Bioinformatik

Genetische Algorithmen (GA, z.B. GOLD)



1 0 1 0 0 1 0 1 0 0 1 0 0 0 1 0 1 0 0 1
 1 2 3 4

- Repräsentation besteht aus Torsionswinkeln (1 Gen = 1 Torsion; 1 Chromosom = 1 Konformation)
- In diesem Beispiel 5 Bit Auflösung für Torsionswinkel, also 32 Winkelschritte möglich
- Fitness Funktion bewertet jedes Chromosom

G. Wolber, Algorithmische Bioinformatik

ParaDockS



Modulares Open Source Framework (<http://www.paradocks.org>)

Particle Swarm optimization (PSO) als Metaheuristik (vgl. GA):

- Sozialverhalten von Tieren nachempfunden

Austauschbare Scoringfunktionen als 'Objective Functions' (PMF, DrugScore, Kraftfelder)

G. Wolber, Algorithmische Bioinformatik

Vergleich GA/PSO

PSO hat keine genetischen Operatoren (X-over, Mutation) als Steuerungsparameter

Informationsaustausch ist anders:

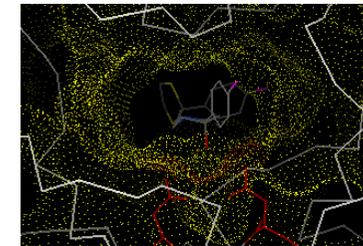
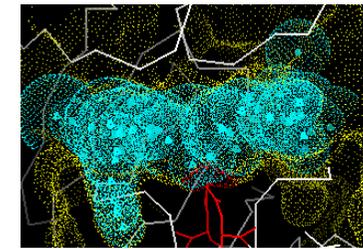
- **GA:** Chromosomen tauschen alle Informationen untereinander aus, Bewegung als "eine Gruppe"
- **PSO:** Nur das 'beste' Partikel tauscht Information mit allen anderen aus und konvergiert daher schneller; fällt eher in lokale Minima

G. Wolber, Algorithmische Bioinformatik

Heuristiken: DOCK

Schritte:

- Proteinoberfläche berechnen (gelb)
- Kugeln an Oberfläche anfügen (blau)
- Zentren dieser Kugeln werden auf Atompositionen gemappt, dann optimiert (unten)

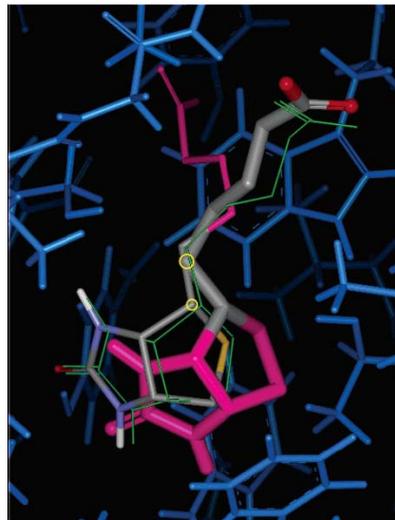


G. Wolber, Algorithmische Bioinformatik

Heuristiken: SurflexDock & FlexX

Vorgangsweise:

- Ligand wird in Teilfragmente zerlegt
- Teilfragmente werden in verschiedenen Posen im Rezeptor positioniert
- Ligand wird in Form von günstigen Konformationen inkrementell wieder zusammengesetzt



(grau: fixiert; rosa: beweglich)

G. Wolber, Algorithmische Bioinformatik

Scoringfunktionen

Drei Prinzipien:

- **Molekularmechanische Kraftfelder** (Geometrische Komplementarität und Ladungen)
z. B. CHARMM [Brooks83] oder AMBER [Cornell95]
- **Empirische Scoringfunktionen** (Hydrophobe Kontakte, Anzahl Wasserstoffbrücken)
z.B. ChemScore [Eldridge97], GlideScore [Friesner04], AutoDock [Morris98]
- **Wissensbasierte Scoringfunktionen** (Statistiken über experimentell bestimmte 3D Strukturen aus Komplexen [Protein Data Bank, PDB] und kleinen Molekülen [Cambridge Structural Database, CSD]
z.B. PMF [Muegge99], Bleep [Mitchell99], DrugScore [Gohlke00]

G. Wolber, Algorithmische Bioinformatik

Beiträge zu Scoringfunktionen

Empirische Scoringfunktionen:

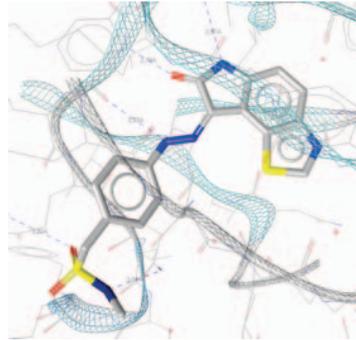
z.B. chemScore = $E_{\text{Hbond}} + E_{\text{lipo}} + E_{\text{rot}}$

Wissensbasierte:

z.B. drugScore = $E_{\text{VDW}} + E_{\text{SAS}}$

Kraftfeldbasierte:

z.B. score = $E_{\text{bondlength}} + E_{\text{bondangles}} + E_{\text{torsion}} + E_{\text{electrostatic}}$



G. Wolber, Algorithmische Bioinformatik

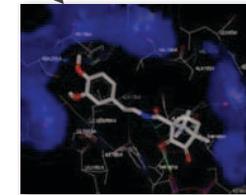
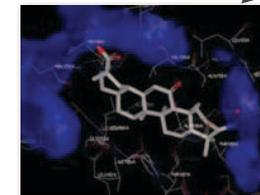


Kristallstruktur

leere Bindetasche

Redocking

neue Liganden



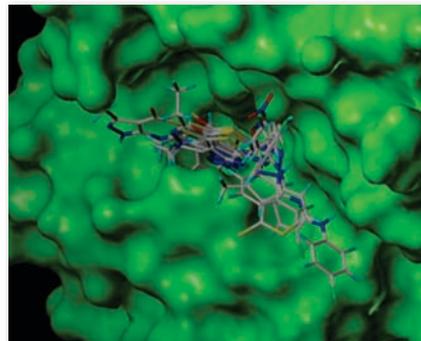
G. Wolber, Algorithmische Bioinformatik

Docking: Durchführung

Input:

Dreidimensionale
Proteinstruktur &
Moleküle (flexibel)

Output: virtuelle
Hitliste, 3D
Konformationen
und Docking-
Scores



G. Wolber, Algorithmische Bioinformatik

Docking via Webinterface

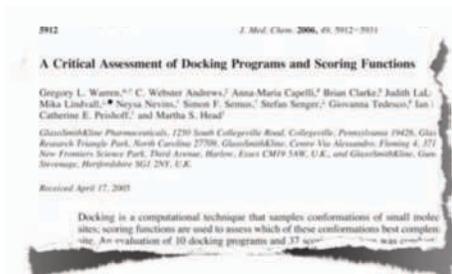
<http://blaster.docking.org> - "DOCKBlaster"

- o Basiert auf DOCK, jetzt UCSF Dock
- o ermöglicht Docken großer Datenbanken

G. Wolber, Algorithmische Bioinformatik

Was leistet Docking?

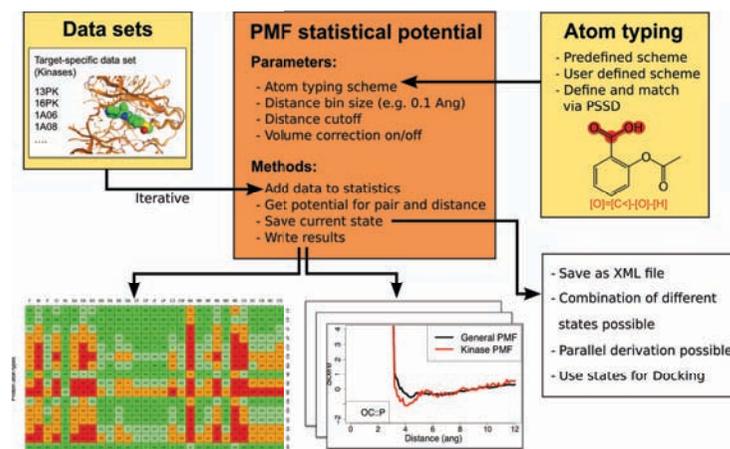
- Virtual Screening
"Docking programs could identify molecules active against a target out of a population of decoy molecules."
- Scoring: "No strong correlation was observed for any scoring function protein target pair"



Warren et. al, J Med. Chem., 2006,
**A Critical Assessment of Docking
Programs and Scoring Functions**

G. Wolber, Algorithmische Bioinformatik

Neue Trends bei Scoringfunktionen: Targetspezifische Paarpotenziale



M. Pippel & M. Scharfe: GCC 2011, Goslar

G. Wolber, Algorithmische Bioinformatik

Was leistet Docking?

Rationalisierung und Ideenfindung:

- Erklärung bereits experimentell bestimmter Bindemodi
- Leitstrukturoptimierung
- Ideen, wie neue Moleküle binden könnten

Was ist unrealistisch?

- Quantitative Vorhersagen (Vorhersage und Vergleich der Bindeaffinität verschiedener Liganden)
- Hohes Enrichment (D.h. ein großer Anteil an experimentell bioaktiven Verbindungen in einem Vorhersageexperiment)

Hauptproblem:

- Scoringfunktionen sind treibende Kraft für das Sampling (vergleichend), aber zu wenig genau

G. Wolber, Algorithmische Bioinformatik

3D Pharmakophore

3D Pharmakophormodellierung

- Definition, Hintergrund, Modellgenerierung

Strukturbasierte 3D Pharmakophore

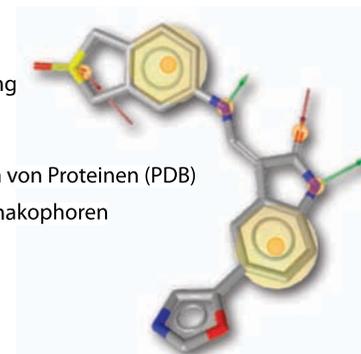
- Experimentell bestimmte Kristallstrukturen von Proteinen (PDB)
- Ableitung von strukturbasierten 3D Pharmakophoren

3D Überlagerung

- Mustererkennungsansatz
- Beispiele und Validierung

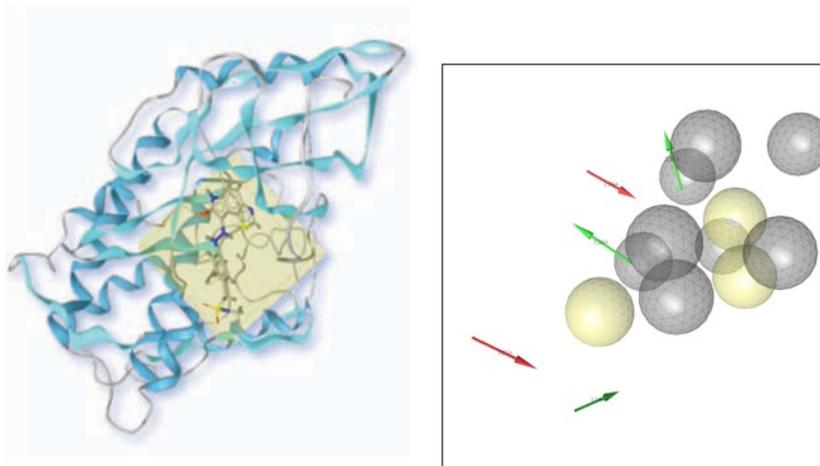
Anwendungen

- Vergleich mit Docking
- Shared & Merged feature pharmacophores
- Ligandenbasierte Pharmakophorgenerierung



G. Wolber, Algorithmische Bioinformatik

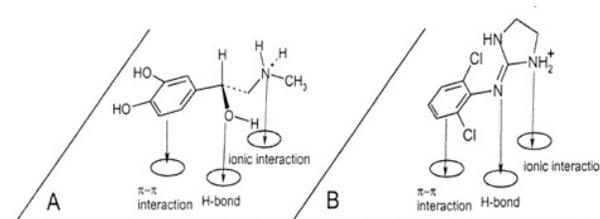
3D Pharmakophore



CDK2 complexed with N-methyl-14-[2-(7-oxo-6,7-dihydro-8H-[1,3]thiazolo[5,4-E]indol-8-ylidene)hydrazino]phenyl]methanesulfonamide

G. Wolber, Algorithmische Bioinformatik

Pharmakophore 1964



In clonidine (B) the restricted rotation resulting from *o*- and *o'*-substitution imposes a quasi-perpendicular orientation of the imidazole ring towards the phenyl ring. As a result, clonidine can yield the same kind of interactions than **norepinephrine (A)**.

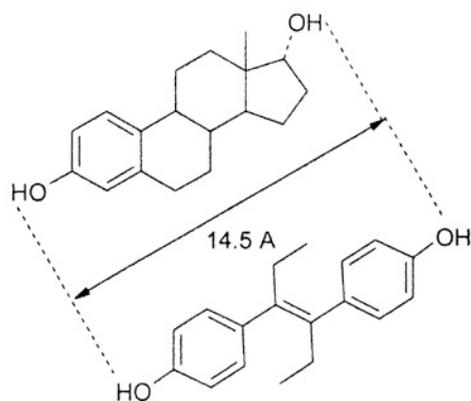
C. Wermuth, *Pharmacophores, the viewpoint of a medicinal chemist* in T. Langer, R. Hofmann; *Pharmacophores and Pharmacophore Searches* Wiley-VCH, 2006, ISBN: 3527312501

Barlow, R. B., *Introduction to Chemical Pharmacology*, 2nd edn. Methuen, London, 1964.

Belleau, B., *An analysis of drug-receptor interactions. In Modern Concepts in the Relationship Between Structure and Pharmacological Activity*, Brunings, K. J. (ed.), Pergamon Press, Oxford, 1963, pp. 75–99.

G. Wolber, Algorithmische Bioinformatik

Pharmakophore 1938



Trans-Diethylstilbestrol zeigt pharmakologische Ähnlichkeit mit Östradiol aufgrund seiner geometrischen Ähnlichkeit

Dodds, E. C., Lawson, W., *Molecular structure in relation to oestrogenic activity. Compounds without phenanthrene nucleus*. Proc. R. Soc. London, Ser. B 1938, 125, 122–132.

C. Wermuth, *Pharmacophores, the viewpoint of a medicinal chemist* in T. Langer, R. Hofmann; *Pharmacophores and Pharmacophore Searches* Wiley-VCH, 2006, ISBN: 3527312501

G. Wolber, Algorithmische Bioinformatik

Der Begriff "Pharmakophor": 1909

Paul Ehrlich, 1909:

A pharmacophore is a molecular framework that carries (phoros) the essential features responsible for a Drug's (pharmacon) biological activity.



Emil Fischer, 1894:

"Um ein Bild zu gebrauchen, will ich sagen, dass Enzym und Glucosid wie Schloss und Schlüssel zu einander passen müssen, um eine chemische Wirkung auf einander ausüben zu können".

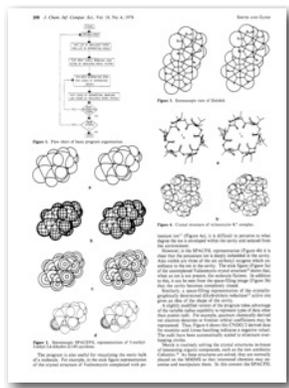


G. Wolber, Algorithmische Bioinformatik

Pharmacophore definition: 1977

Peter Gund, 1977

“A pharmacophore is an arrangement of molecular features or fragments forming a necessary but not sufficient condition for biological activity.”



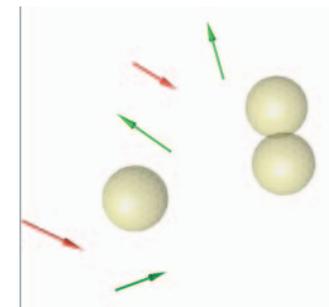
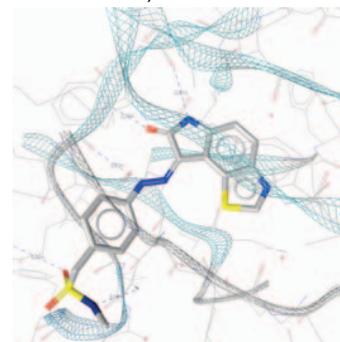
J. Chem. Inf. Comp. Sci. (1978)
Computer-Generated
Space-Filling Molecular Models

G. Wolber, Algorithmische Bioinformatik

3D Pharmakophormodelle

Pharmacophore = Ensemble universeller chemischer Merkmale, die eine spezifische Konformation eines Protein-Liganden repräsentieren

Chemische Merkmale ('chemical features'): Wasserstoffbrücken, Ladungsinteraktionen, Hydrophobe Kontakte, .. (supramolekulare Interaktionen)



G. Wolber, Algorithmische Bioinformatik

Pharmacophore definition: Today

Definition 3 (IUPAC recommendations, 1998)

“A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response.”

C.-G. Wermuth et al., Pure Appl. Chem. 1998, 70:1129-43

G. Wolber, Algorithmische Bioinformatik

Warum 3D Pharmakophore?

Universell

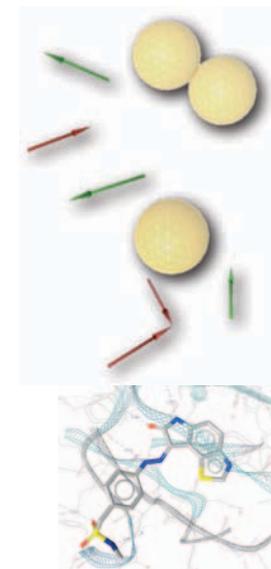
- Pharmakophore repräsentieren chemische Funktionalität, abstrahieren Interaktionsmuster, die auch für andere Liganden gültig sind

Effizient

- Einfachere Darstellung (Blobs und Vektoren statt Molekülgraphen)
- Definierte Featuretypen: Wasserstoffbrücken, Ladungsinteraktionen, etc.

Verständlich und editierbar

- Können von medizinischen Chemikern einfach verstanden und editiert (optimiert) werden; transparent

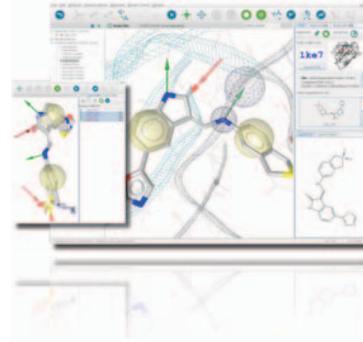


G. Wolber, Algorithmische Bioinformatik

LigandScout: Strukturbasierte Pharmakophorgenerierung

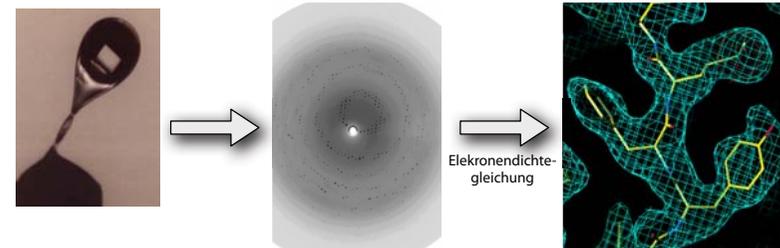
Implementierung:

1. Interpretation der experimentellen Struktur in Bezug auf Atomtypen und Bindungen
2. Chemische Features interpretieren & Untersuchung auf geometrische Komplementarität
3. 3D Pharmakophor erstellen und visualisieren



G. Wolber, Algorithmische Bioinformatik

Proteinkristallstruktur = Modell!!



Kristall mit Trillionen Molekülen

Beugungsmuster

Berechnete Elektronendichte, z.T. manuelles Fitting

- Templates für Aminosäuren
- keine Wasserstoffe, Bindungsordnung nicht sichtbar
- Liganden oft zuwenig beachtet und fehlerhaft!

G. Wolber, Algorithmische Bioinformatik

Kristallstrukturen von Proteinen

Protein Data Bank

<http://www.pdb.org>

Alle publizierten Kristallstrukturen von Proteinen werden in der PDB hinterlegt

4-letter Code für Proteine (oder Komplexe), z.B. **1KE7**

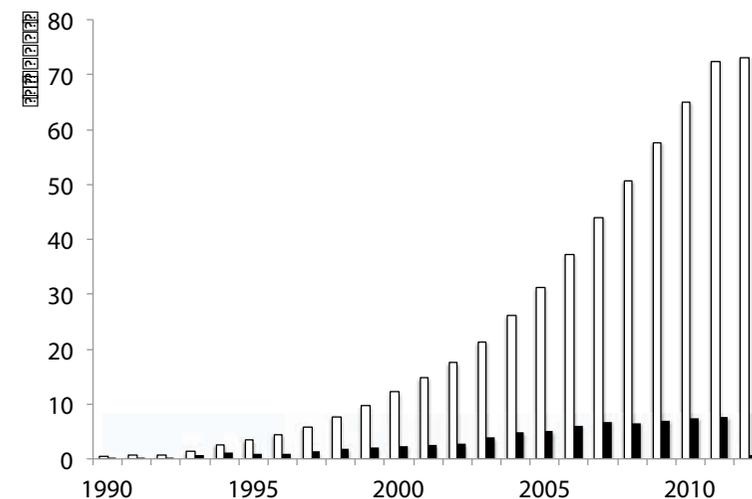
3-letter Code für Liganden, z.B. **LS3**

Viele verschiedene Tools zur Analyse, Visualisierung, etc.



G. Wolber, Algorithmische Bioinformatik

Entwicklung der PDB & Alter der Einträge



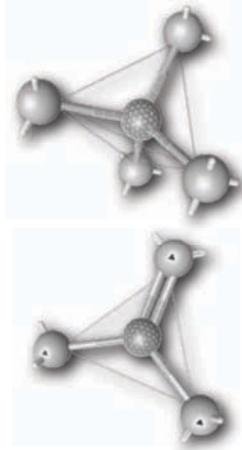
G. Wolber, Algorithmische Bioinformatik

Bestimmung der Hybridisierungsstati

Quantitative Geometrie-Templates zur Erkennung der richtigen Geometrie

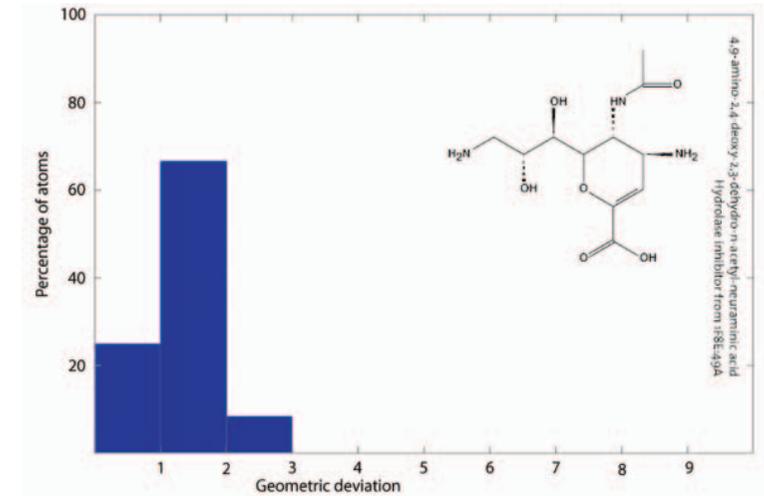
- sp^3 : tetrahedral
- sp^2 : trigonal planar
- sp : linear

Positionierung anhand der ersten zwei Punkte, dann numerisches drehen, zur besten Überlagerung (Quaternionenrotation)



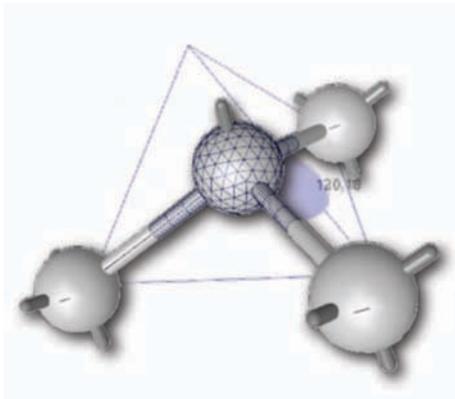
G. Wolber, Algorithmische Bioinformatik

Hybridierungsstati: Fehlerbestimmung



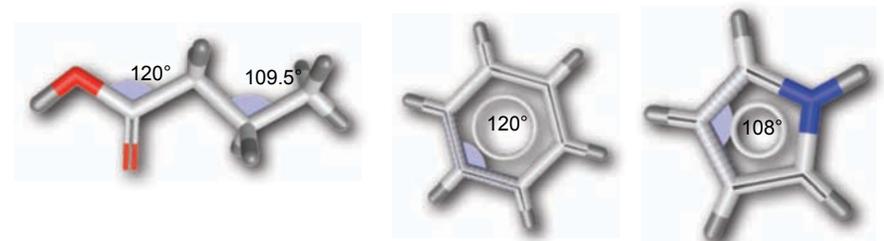
G. Wolber, Algorithmische Bioinformatik

Geometrietemplates: Besser als Bestimmung von Bindungswinkeln?



G. Wolber, Algorithmische Bioinformatik

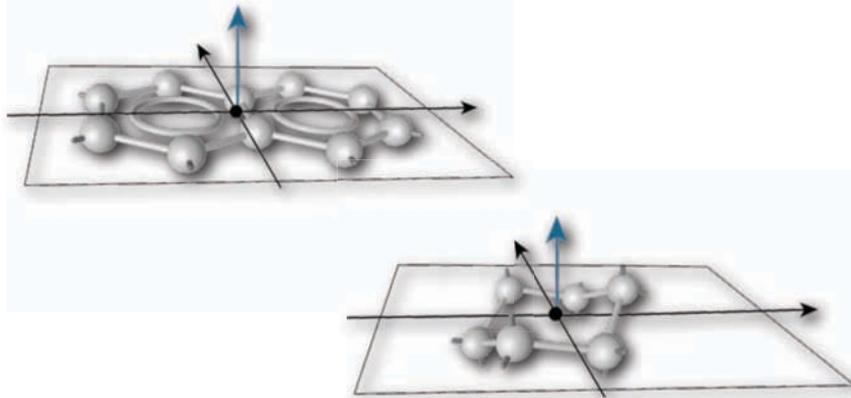
Ringgeometrien sind anders



Planare Ringe bestehen ausschließlich aus sp^2 -Atomen und trotzdem andere Bindungswinkel

G. Wolber, Algorithmische Bioinformatik

Hauptkomponentenanalyse zur Bestimmung der Ringhauptebene

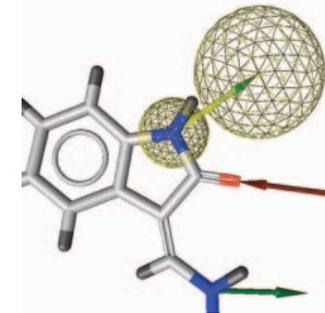


Max. Abweichung von der Hauptebene: 0.4 Å

G. Wolber, Algorithmische Bioinformatik

Chemische Merkmale ('Features')

- Wasserstoffbrückendonoren und -akzeptoren (Vektoren)
- Positive und negative Ladungen (Punkte)
- Hydrophobe Bereiche
- π -Ebenen (Punkt und Normalvektor)
- Metallbindegeometrien (Vektoren)

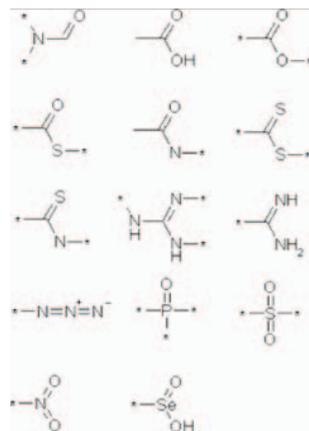


Features beziehen sich immer auf die Ligandenseite!

G. Wolber, Algorithmische Bioinformatik

Chemische Muster: Doppelbindungsverteilung in sp^2 -Ketten

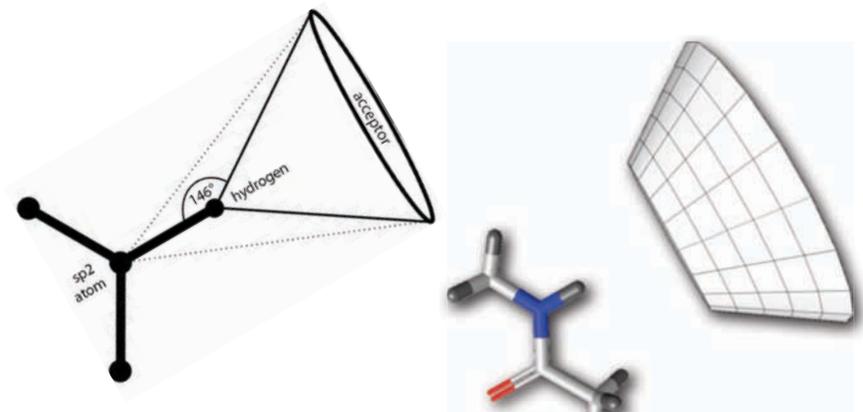
- Keine exakte Lösung in vielen Fällen (z. B. Tautomere; Keto-enol)
- Bestimmte Muster sind chemisch gegeben
- Für die Zuweisung der wahrscheinlichsten anderen Doppelbindungen: gewichtete Verteilung (nichtbipartites Matching) zur Maximierung der Anzahl Doppelbindungen



Patterns by Roger Sayle: Bioinformatics Group, Metaphorics LLC, Santa Fe, see <http://www.daylight.com/meetings/mug01/Sayle/m4xbondage.html>

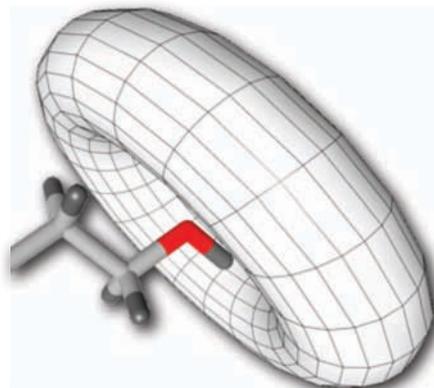
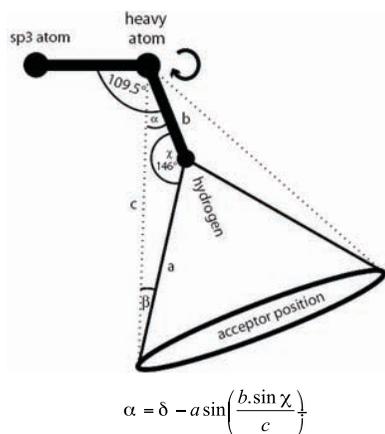
G. Wolber, Algorithmische Bioinformatik

Geometrien für die Erkennung (Constraints): Rigide H-Brücken



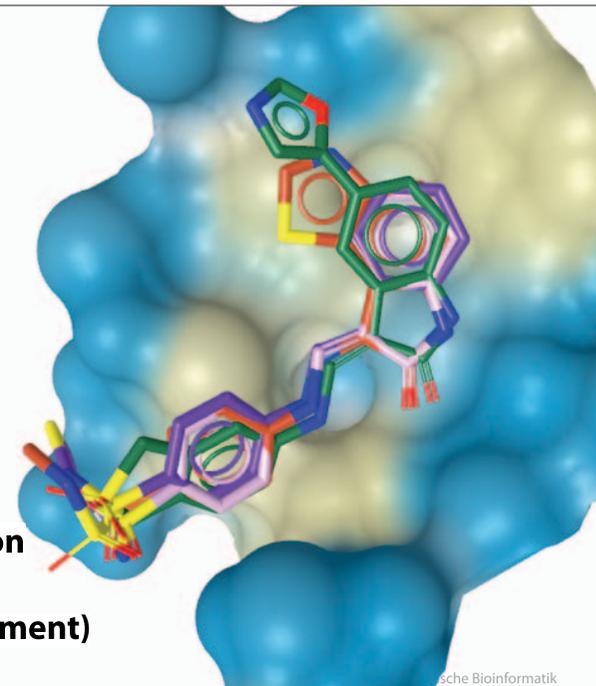
G. Wolber, Algorithmische Bioinformatik

Geometrien für die Erkennung (Constraints): Flexible H-Brücken



G. Wolber, Algorithmische Bioinformatik

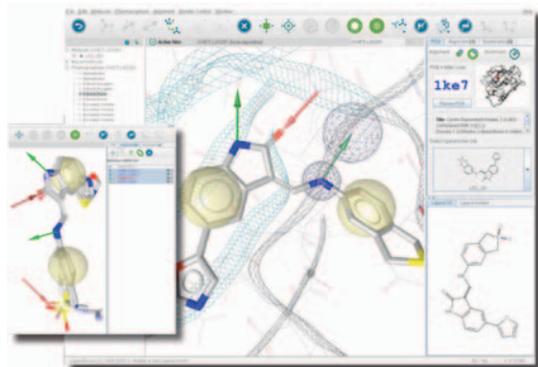
Überlagerung von Molekülen ... (Structural Alignment)



Algorithmische Bioinformatik

Software zur Pharmakphorgenerierung

LigandScout



[LigandScout 2005] Wolber, G.; Langer, T. 3D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters J. Chem. Inf. Comput. Sci.; (Article); 2005; 45(1); 160-169. DOI: 10.1021/ci049885e

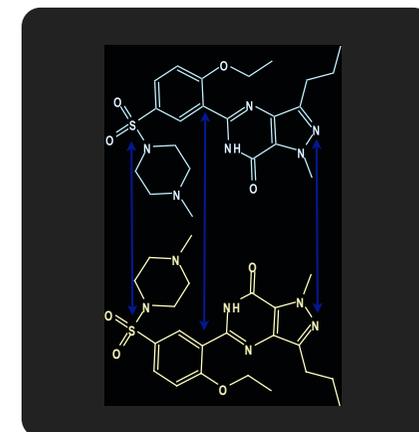
<http://www.inteligand.com/ligandscout>

G. Wolber, Algorithmische Bioinformatik

Alignment: Atome/funktionelle Gruppen

Überlagerung nach chemischem Subgraph

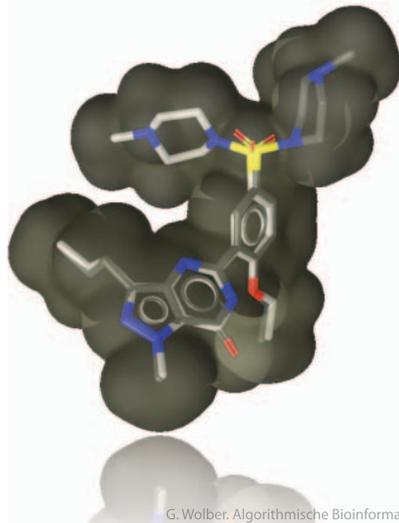
- Isolierte Betrachtung chemischer Substrukturen
- Keine chemischen Eigenschaften (3D, Interaktion mit Rezeptor, keine Potenziale)



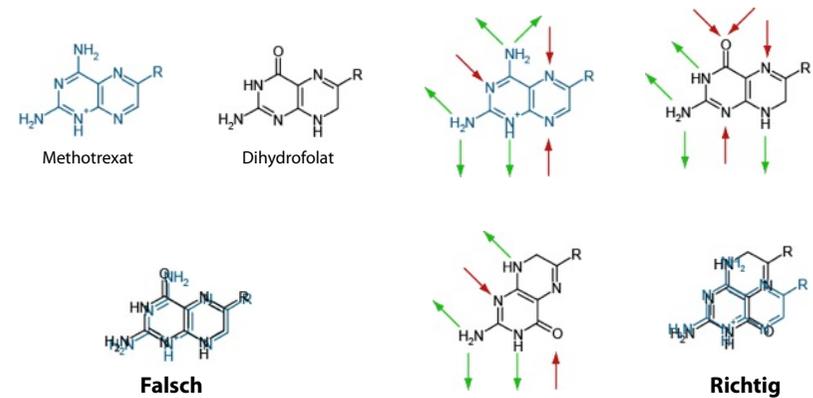
G. Wolber, Algorithmische Bioinformatik

Volumen?

- Sterische Terme sind höchste Energiebeiträge
- Exponentialfunktion zur Beschreibung der Volumina (Gaussfunktion) ist rechenaufwändig
- Keine chemischen Eigenschaften berücksichtigt (lipophil/hydrophil etc.)



Überlagerung nach Pharmakophorpunkten



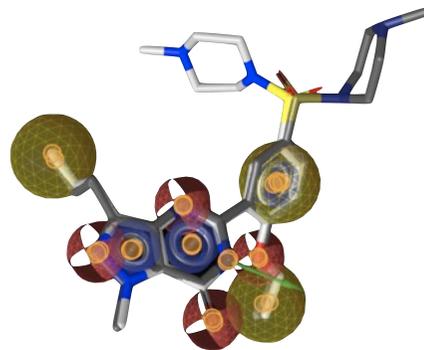
Böhm, Klebe, Kubinyi: Wirkstoffdesign (1996) p. 320f

G. Wolber, Algorithmische Bioinformatik

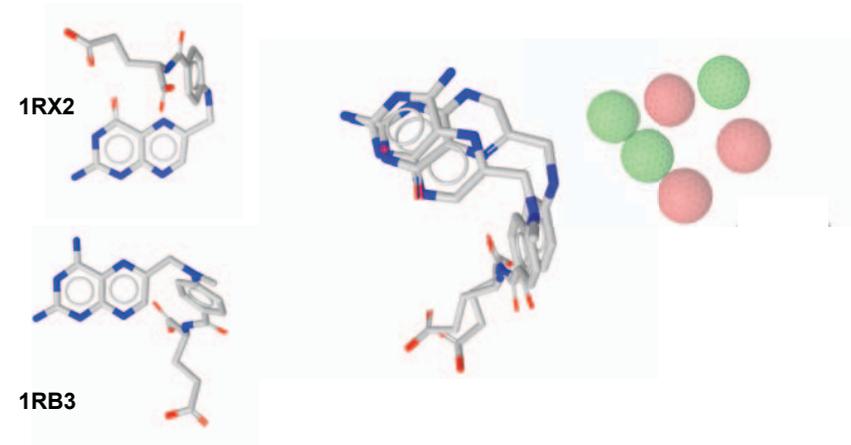
Chemische Features/Pharmakophore

Vorteile:

- Chemie wird berücksichtigt
- Gutes Verständnis bei "klassischen" Medizinalchemikern

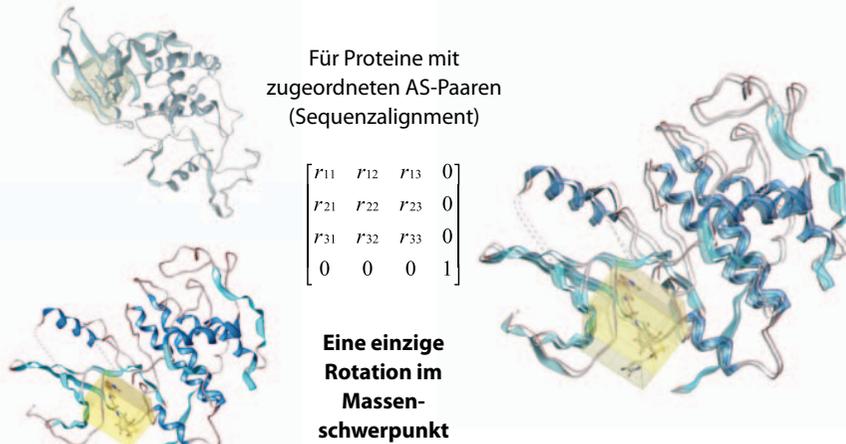


3D Überlagerung nach Pharmakophorpunkten



G. Wolber, Algorithmische Bioinformatik

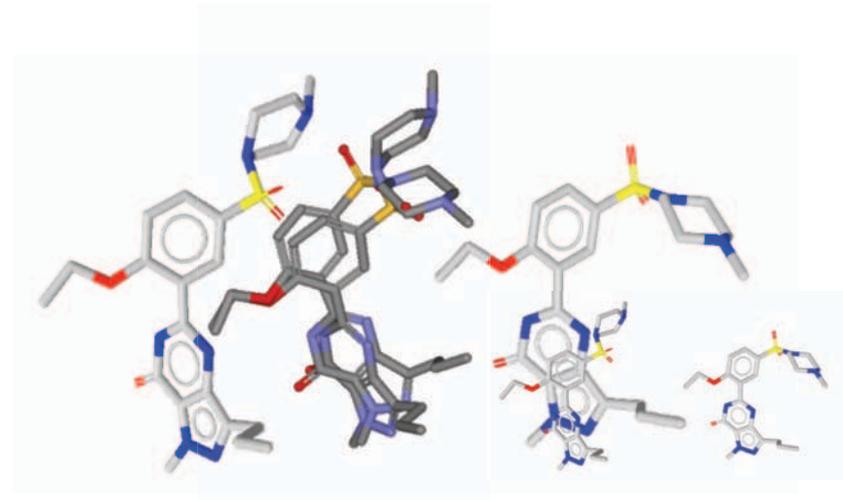
Bekannte Methoden: Kabsch Alignment



[Kabsch, 1976] Kabsch, W. (1976). Acta. Crystal, 32A:922-923.
 [Kabsch, 1978] Kabsch, W. (1978). Acta. Crystal, 34A:827-828.

G. Wolber. Algorithmische Bioinformatik

Kabsch mit Atomen als Bezugspunkte

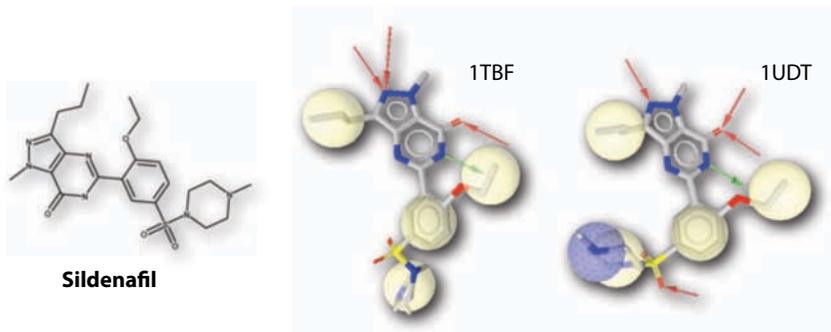


G. Wolber. Algorithmische Bioinformatik

Kabsch Überlagerung

Für Moleküle: Atome statt Aminosäuren als Bezugspunkte

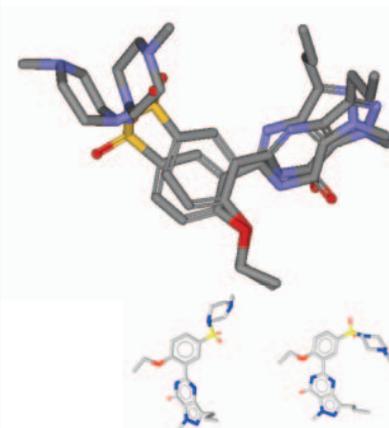
- o Atomreihenfolge kanonisieren, um Paare zu erzeugen



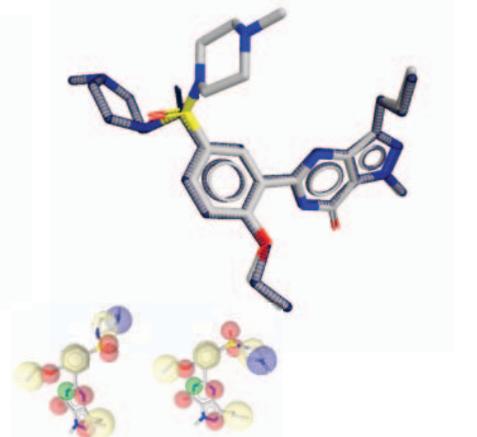
G. Wolber. Algorithmische Bioinformatik

Überlagerung

Ideale Rotation mit Atomen als Bezugspunkte



Biologisches Alignment: Zielfunktion



G. Wolber. Algorithmische Bioinformatik

Was fehlt?

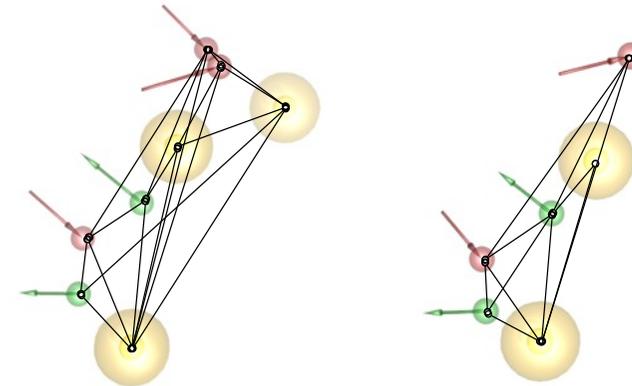


Es werden die 'besten' Paare benötigt:

- Nur wenige verschiedene Featuretypen
- Ähnlichkeitsmaß ist die geometrische Umgebung

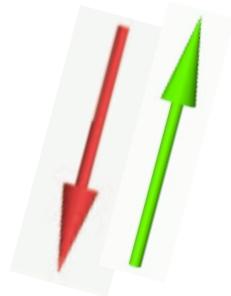
G. Wolber, Algorithmische Bioinformatik

Distanzgeometrie: Kompatibilitätsgraph?



G. Wolber, Algorithmische Bioinformatik

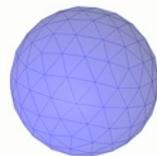
Definierter Satz chemischer Merkmale



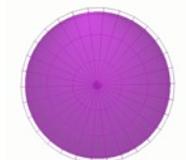
H-Brücken Donor
H-Brücken Akzeptor



hydrophob



positive Ladung



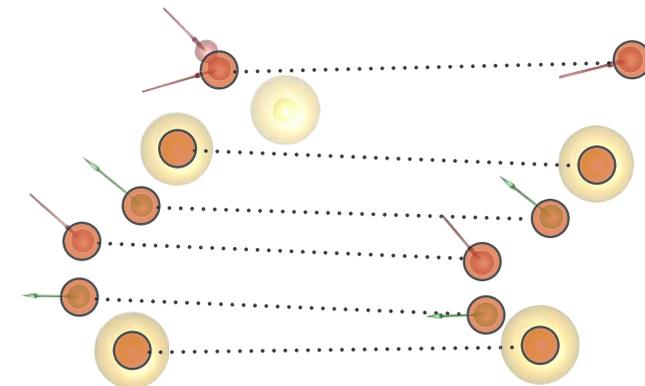
Aromatischer Ring
(π -Ebene)



negative Ladung

G. Wolber, Algorithmische Bioinformatik

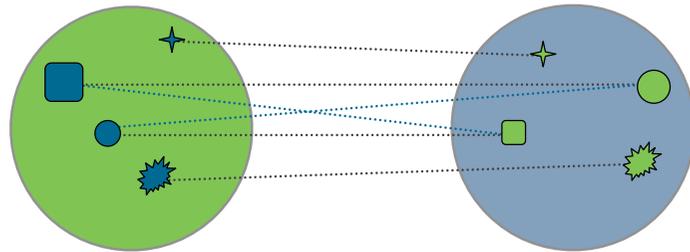
Ähnlichkeitsmaß: Mustererkennung



H. W. Kuhn. The Hungarian method for the Assignment Problem. *Noval Research Quarterly*, 2 (1995)
 W. Kabsch. (1978). *Acta Crystal*, 34A:827-828.
 G. Wolber, A. Dornhofer, and T. Langer. *J. Comput.-Aided Mol. Design*, 20: 773-788 (2006)
 T. Seidel, G. Ibis, F. Bendix, G. Wolber. *Drug Discovery Today: Technologies*, 7(4): e221-e228 (2010)

G. Wolber, Algorithmische Bioinformatik

Richtige Paarung?

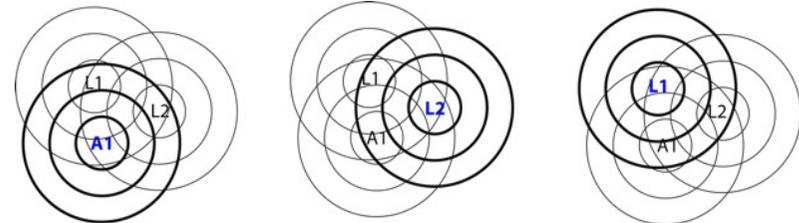


Hungarian Matcher (Heiratsproblem)

- [Edmonds 1965] Matching and a Polyhedron with 0-1 Vertices. J. Res. NBS 69B (1965), 125-30 [nonbipartite application]
- [Kuhn 1955] The Hungarian method for the Assignment Problem. Noval Research Quarterly, 2 (1955) [bipartite variant]
- [Richmond 2004] Application to chemistry: N. Richmond et al. Alignment of 3D molecules using an image recognition algorithm. J. Mol. Graph. Model. 23, 2004, pp 199-209.

G. Wolber, Algorithmische Bioinformatik

Mustererkennung

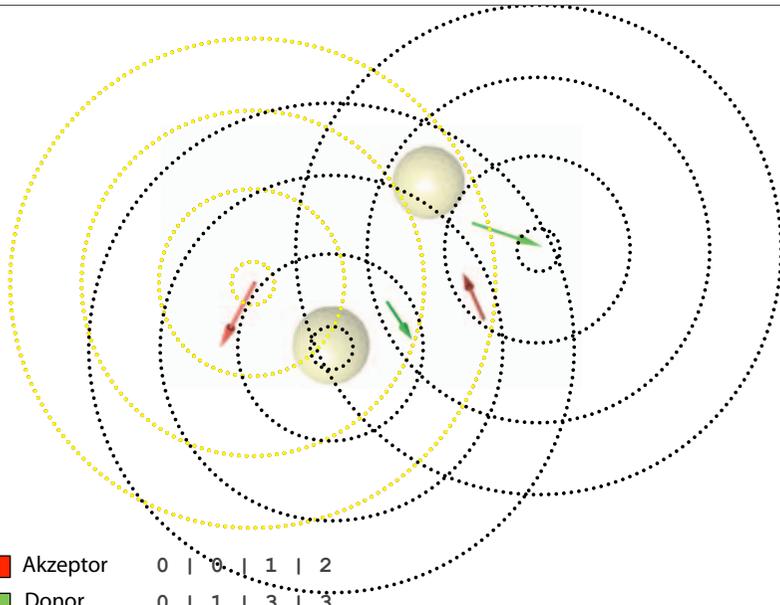


A1(L): 0 | 2 | 4
A1(A): 0 | 0 | 0

L2(A): 0 | 1 | 2
L2(L): 0 | 1 | 2

L1(A): 0 | 1 | 2
L1(L): 0 | 1 | 2

G. Wolber, Algorithmische Bioinformatik



■	Akzeptor	0		0		1		2
■	Donor	0		1		3		3
■	Hydrophob	0		1		3		3

G. Wolber, Algorithmische Bioinformatik

Ähnlichkeitsbestimmung ('Cost Function')

$$cost(x, y) = \sum_{shell} \left[weight(shell) * \sum_{type} (\min(n_x(shell, type), n_y(shell, type))) \right]$$

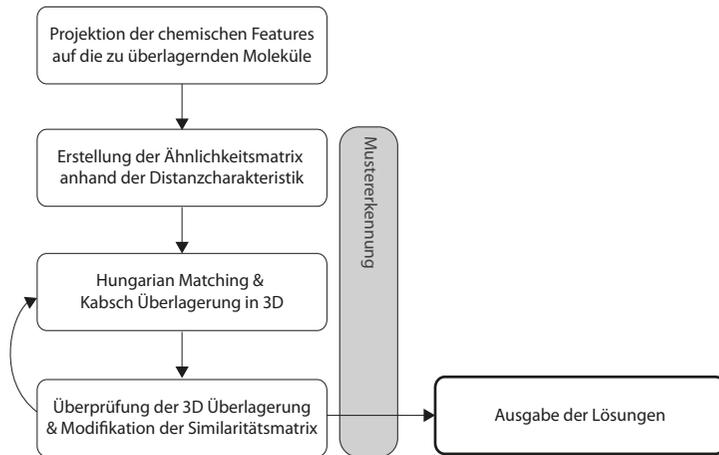
$n_x(shell, type)$ # Elemente in Schale von Typ $type$ für element x
 $n_y(shell, type)$ # Elemente in Schale von Typ $type$ für element y

$$weight(i_{shell}) = \left(1 - \left(\frac{i_{shell}}{n} \right)^3 \right) * 2 \left(1 - \frac{i_{shell}}{n} \right) * m$$

i_{shell} Schalen (shell) index
 m Maximale Anzahl der Elemente pro Schale
 n Anzahl der Schalen (shells)

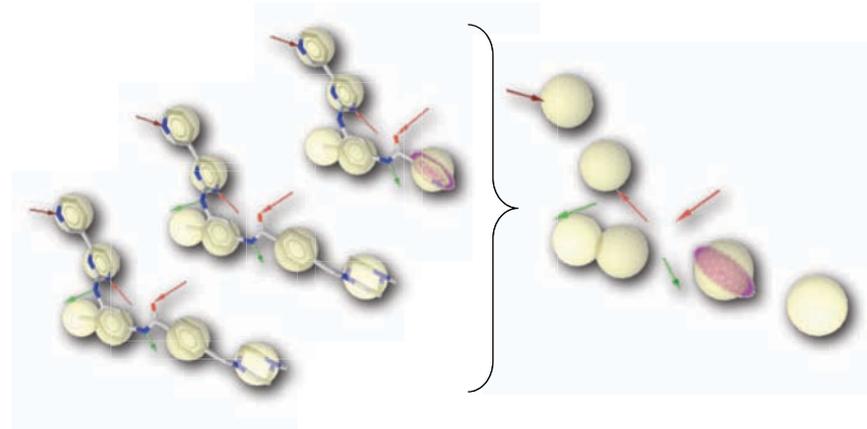
G. Wolber, Algorithmische Bioinformatik

Überlagerung durch Mustererkennung



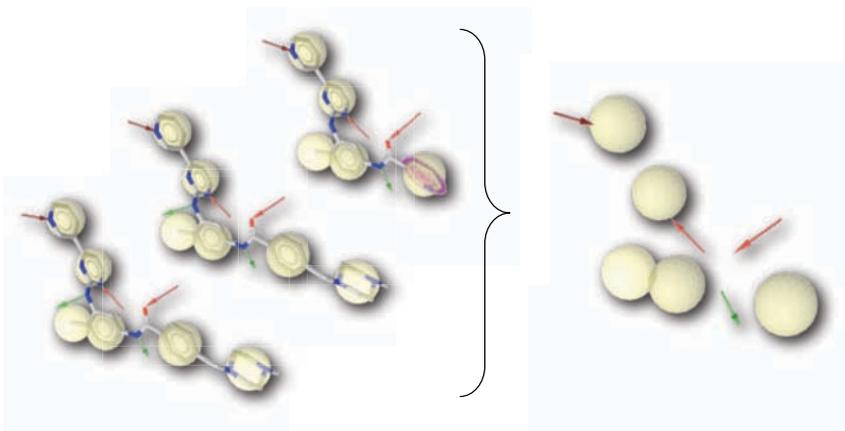
G. Wolber, Algorithmische Bioinformatik

Anwendung: 'Merged Feature'-Pharmacophore



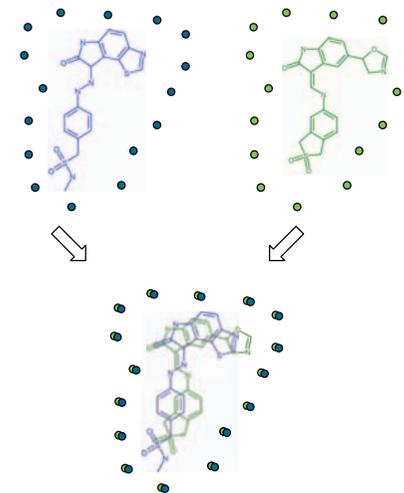
G. Wolber, Algorithmische Bioinformatik

Anwendung: 'Shared Feature'-Pharmakophore



G. Wolber, Algorithmische Bioinformatik

Validierung 1



- Verschiedene an das selbe Protein gebundene Liganden aus der PDB
- C_{α} (7A) als Referenzpunkte
- Überlagerung durch (i) Proteinreferenzpunkte und (ii) durch Liganden-Überlagerung ohne Proteinbezug
- Vergleich dieser beiden Überlagerungen

G. Wolber, Algorithmische Bioinformatik

Phosphodiesterase 5

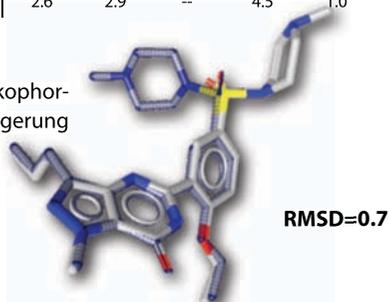
	1UDT	1TBF	1XP0	1UHO	1UDU	1XOZ
1UDT	-	0.7	0.9	0.3	2.3	2.5
1TBF	0.7	-	0.4	0.7	2.7	3.1
1XP0	0.8	0.4	-	0.8	2.6	--
1UHO	0.3	0.7	0.7	-	2.3	4.5
1UDU	2.2	2.6	2.5	1.7	-	1.0
1XOZ	2.6	2.9	--	4.5	1.0	-

Sildenafil

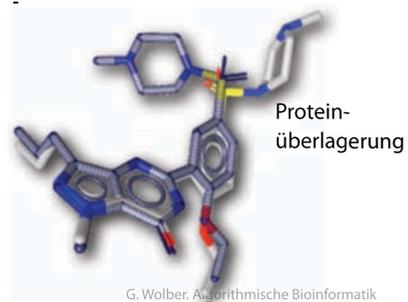
Tadalafil

Vardenafil

Pharmakophor-
überlagerung



Protein-
überlagerung

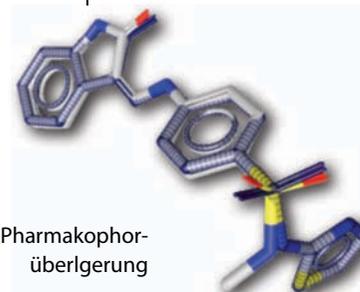


G. Wolber, Algorithmische Bioinformatik

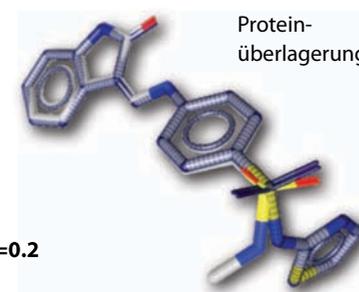
CDK2

	1KE5	1KE6	1KE7	1KE8	1KE9
1KE5	-	0.4	0.6	0.2	0.3
1KE6	0.5	-	0.5	0.4	0.3
1KE7	0.6	0.6	-	0.7	0.3
1KE8	0.3	0.4	0.7	-	0.4
1KE9	0.3	0.3	0.3	0.3	-

Pharmakophor-
überlagerung



Protein-
überlagerung



G. Wolber, Algorithmische Bioinformatik

Phosphodiesterase 5

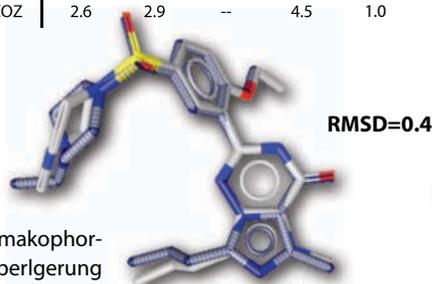
	1UDT	1TBF	1XP0	1UHO	1UDU	1XOZ
1UDT	-	0.7	0.9	0.3	2.3	2.5
1TBF	0.7	-	0.4	0.7	2.7	3.1
1XP0	0.8	0.4	-	0.8	2.6	--
1UHO	0.3	0.7	0.7	-	2.3	4.5
1UDU	2.2	2.6	2.5	1.7	-	1.0
1XOZ	2.6	2.9	--	4.5	1.0	-

Sildenafil

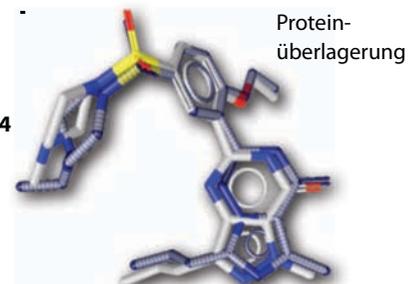
Tadalafil

Vardenafil

Pharmakophor-
überlagerung



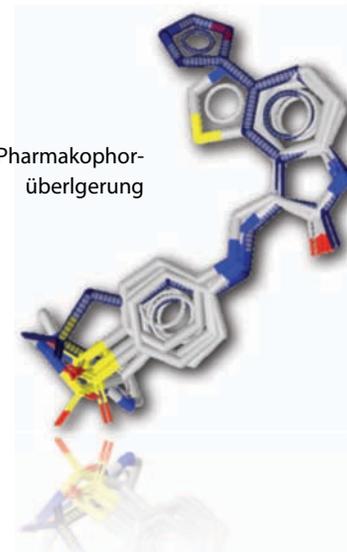
Protein-
überlagerung



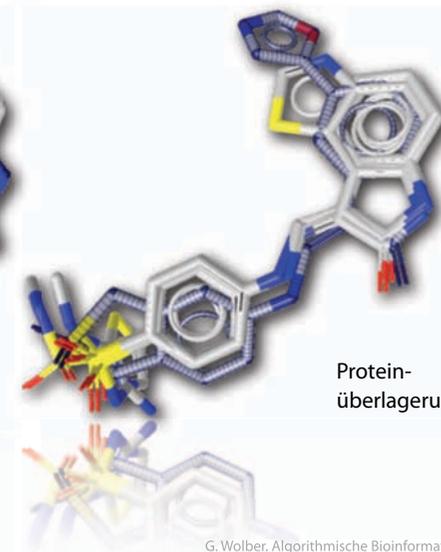
G. Wolber, Algorithmische Bioinformatik

CDK2 (alle 5)

Pharmakophor-
überlagerung

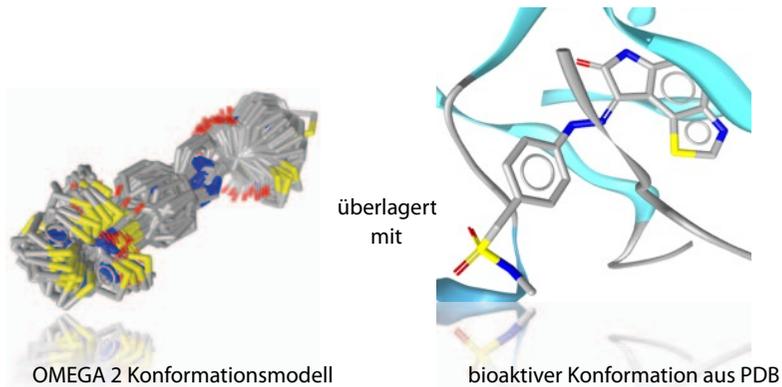


Protein-
überlagerung



G. Wolber, Algorithmische Bioinformatik

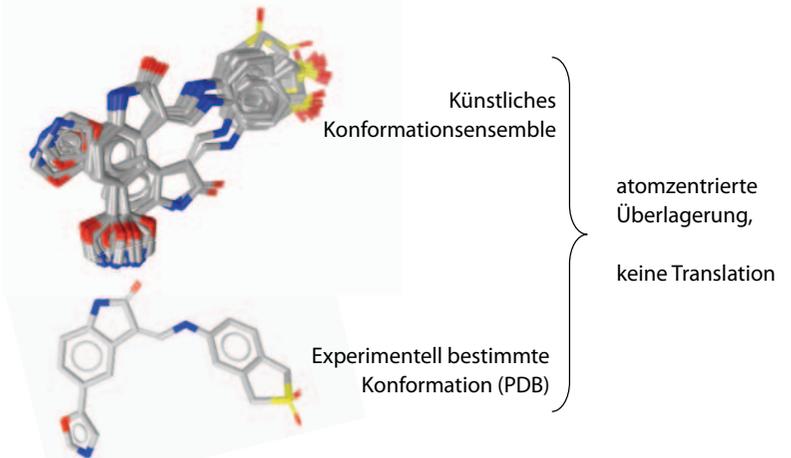
Validierung 2: Flexible Überlagerung



Bewertetes Ergebnis: Konformation mit dem höchsten Alignment-Score

G. Wolber, Algorithmische Bioinformatik

Reproduktion bioaktiver Konformationen



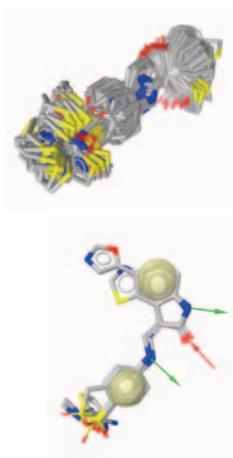
[Conf 2005] Kirchmair, J et al. Comparative Analysis of Protein-Bound Ligand Conformations. JCI 45(2005)
[Conf 2006] Kirchmair, J. et al. Comparative Performance Assessment of the Conformational Model Generators Omega and Catalyst. JCI 46 (2006)

G. Wolber, Algorithmische Bioinformatik

Multikonformationsmodelle

Ist ein vorberechnetes Konformationsensemble geeignet, um bioaktive Konformationen zu finden?

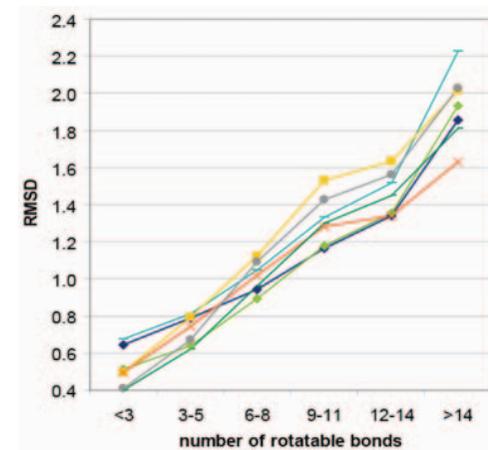
- Konformationsqualität ist essenziell für die Qualität der Ergebnisse
- "Tweak during alignment" ist aufwändiger als Vorberechnen (space for speed) und begünstigt lokale Minima



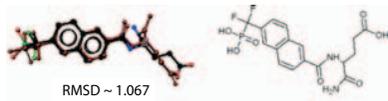
G. Wolber, Algorithmische Bioinformatik

Reproduktion bioaktiver Konformationen

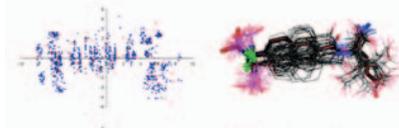
Qualität hängt von Flexibilität ab



G. Wolber, Algorithmische Bioinformatik



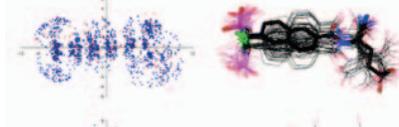
OMEGA HTS



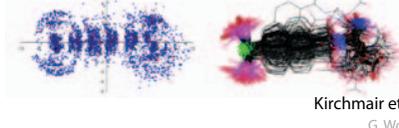
OMEGA HQ



CATALYST HTS



CATALYST HQ



Kirchmair et al., JCI, 2006(46): 1848-1961
G. Wolber, Algorithmische Bioinformatik

Reproduktion bioaktiver Konformationen

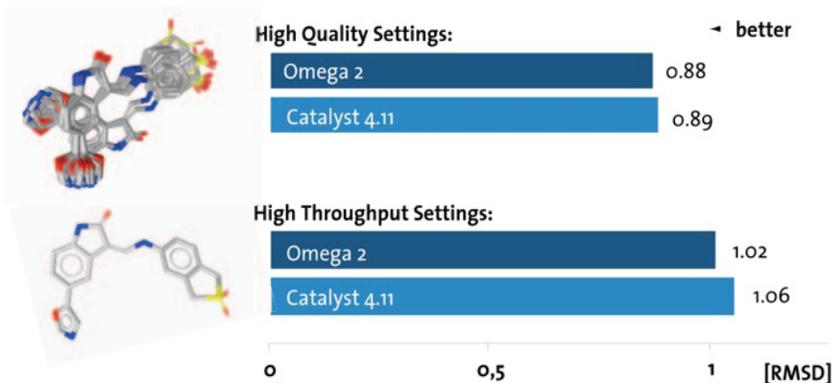
- Konformationsgeneratoren sind in der Lage, den Konformationsraum für typische kleine Moleküle vernünftig abzudecken
- Vorberechnete Konformationen: unverzerrte und problemunabhängige Darstellung des Konformationsraums und vermeiden lokale Minima
- Konformationsberechnung ist rechnerisch teuer -> Vorberechnung spart Zeit

Kirchmair et al., JCI, 2009(49): 678-692
Kirchmair et al., JCI, 2005(45): 422-430
Leach, Gillet, Lewis and Taylor J. Med. Chem. 2010, 53 (2), 539-558

G. Wolber, Algorithmische Bioinformatik

Reproduktion bioaktiver Konformationen

Durchschnittliche geometrische Abweichung ist akzeptabel



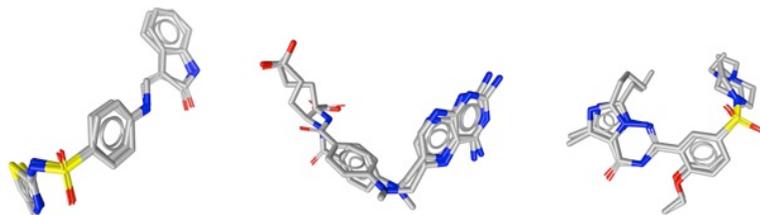
G. Wolber, Algorithmische Bioinformatik

Flexible Überlagerung (vorberechnete Konformationen)

PDB Code/residue	Atom RMSD	Feature RMSD	Fit Score	# matched features	# confs
1cx2-s58-701	0.66	0.58	137.26	13	14
1fpu-prc-1	0.86	0.76	136.71	13	800
1iep-sti-201	1.04	0.67	146.73	14	800
1ke5-ls11	0.44	0.43	107.51	10	96
1ke6-ls2-201	0.44	0.35	137.95	13	474
1ke7-ls3-201	0.51	0.46	107.61	10	96
1ke8-ls4-2	0.37	0.27	138.19	13	800
1ke9-ls5-1	0.45	0.28	138.15	13	290
1rb3-mtx-161	1.49	0.71	185.49	18	800
1tbf-via-501	0.93	0.60	157.20	15	800
1xp0-vdn-201	0.83	0.56	167.33	16	800
4cox-imn-701	0.60	0.43	127.72	12	800
6cox-s58-701	0.48	0.48	137.56	13	14

G. Wolber, Algorithmische Bioinformatik

Flexible Überlagerung (vorberechnete Konformationen)



1ke8

1rb3

1xp0

PDB Code/Ligand	Atom RMSD	Feature RMSD	Fit	# gemeinsame Features	# confs
1ke8-ls4-2	0.37	0.27	138.19	13	800
1rb3-mtx-161	1.49	0.71	185.49	18	800
1xp0-vdn-201	0.83	0.56	167.33	16	800

G. Wolber, Algorithmische Bioinformatik

Methodenvergleich & Diskussion

Pharmakophore

- Pharmakophore beziehen sich auf einen bestimmten Bindemodus
- Transparent, verständlich, editierbar
- Vollautomatisiert
- Schnell & effizient (60.000 Verbindungen in Minuten)

Docking

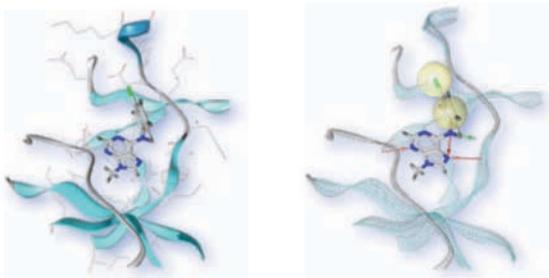
- Flexibel in Bezug auf den Bindemodus
- Black Box
- Vorbereitung der Bindetasche notwendig, Training der Scoringfunktion
- Langsam: 30 Sek. bis Minuten pro Molekül

[Docking 2006] A Critical Assessment of Docking Programs and Scoring Functions, G. L. Warren et al. J. Med. Chem., 49 (20), 5912-5931, 2006.

G. Wolber, Algorithmische Bioinformatik

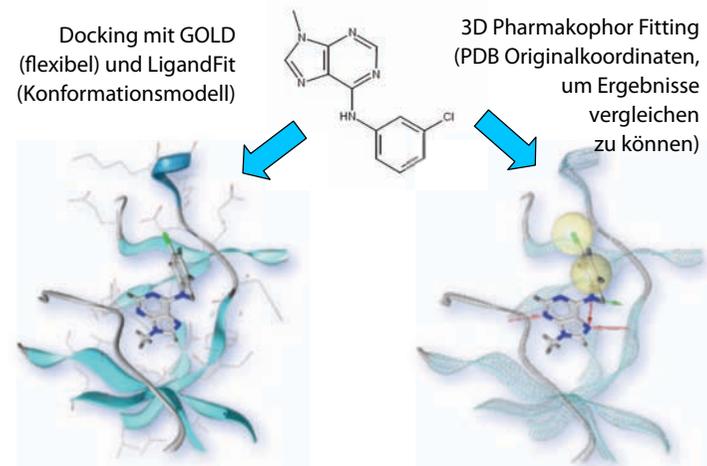
Validierung 3: Docking und 3D Pharmakophore

- Reicht ein Pharmakophor aus, um einen Liganden flexibel richtig in der Bindetasche eines Proteins zu positionieren?
- Vergleich mit Docking



G. Wolber, Algorithmische Bioinformatik

Docking im Vergleich mit Pharmakophoren



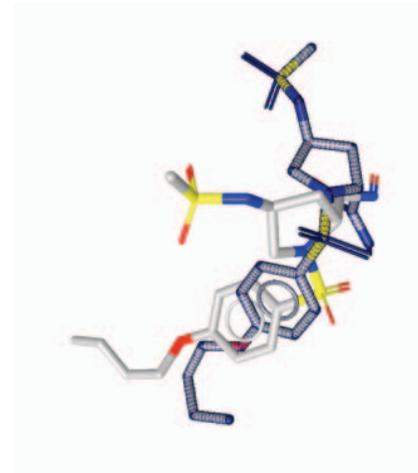
RMSD beinhaltet Verschiebung in der Bindetasche!
(Vergleich mit PDB Konformationsreproduktion nicht möglich)

G. Wolber, Algorithmische Bioinformatik



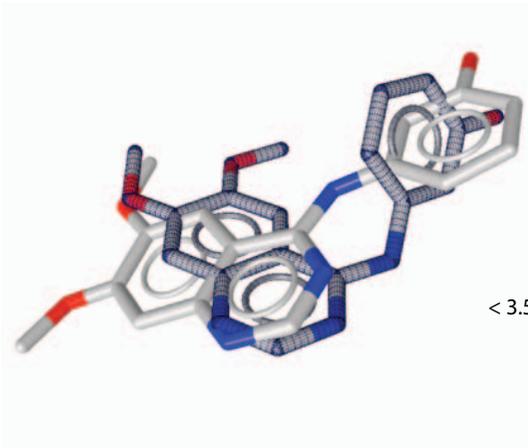
1CKP
RMS = 0.63
< 1.5: „**perfect fit**“

G. Wolber, Algorithmische Bioinformatik



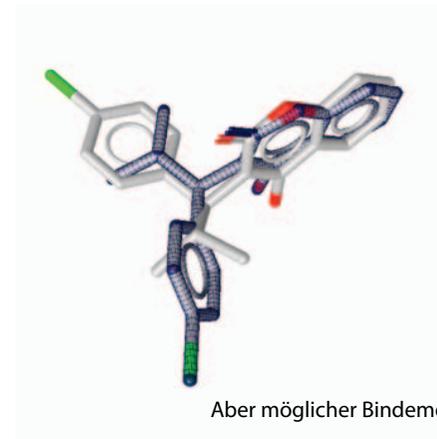
1G49
RMS = 3.69
3.5 < RMS < 5:
„**inadequate fit**“

G. Wolber, Algorithmische Bioinformatik



1D18
RMS = 2.18
< 3.5: „**acceptable fit**“

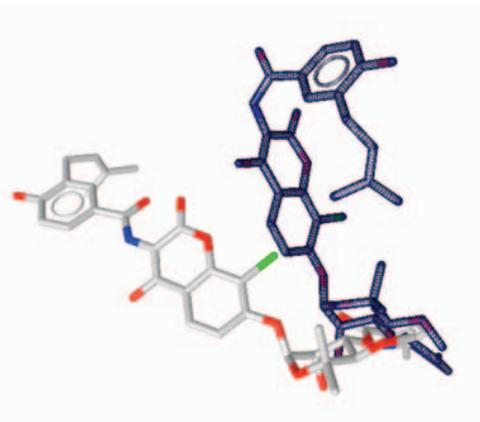
G. Wolber, Algorithmische Bioinformatik



1UPJ
RMS = 3.78
LigandScout
„**inadequate fit**“

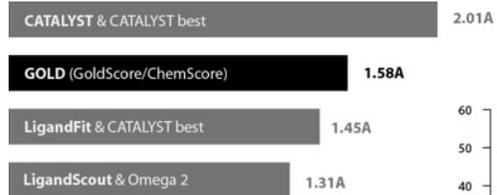
Aber möglicher Bindemodus (bioisoster)!

G. Wolber, Algorithmische Bioinformatik

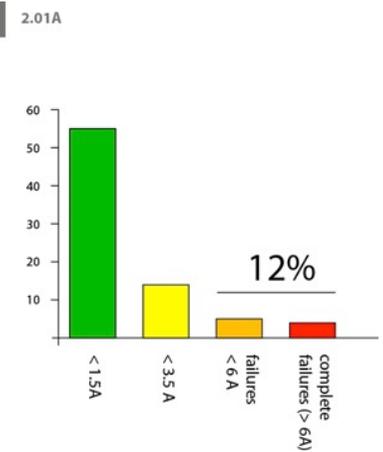


1KZN
 RMS = 8.76
 „inacceptable fit“
 > 6

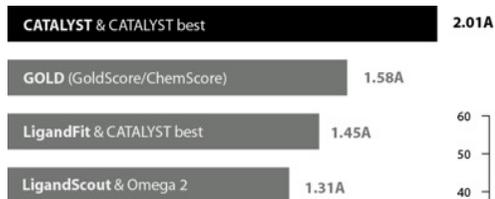
Vergleich Docking & 3D Pharmakophore



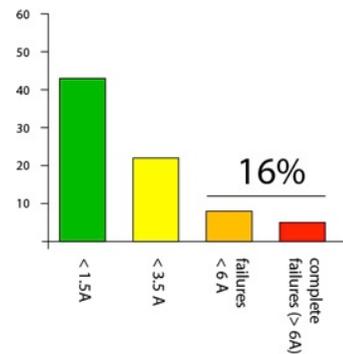
Reproduction quality:
 70% below 1.5Å
 12% failures



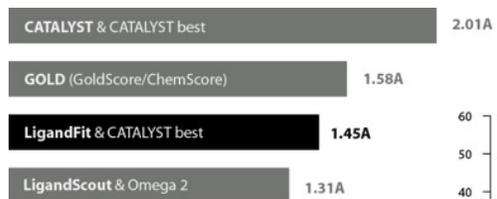
Vergleich Docking & 3D Pharmakophore



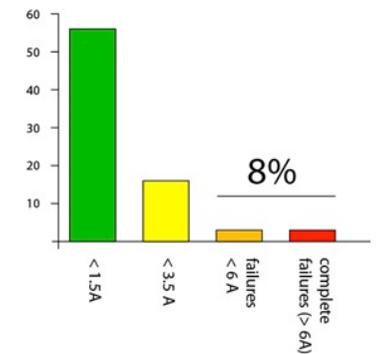
Reproduction quality:
 55% below 1.5Å
 16% failures



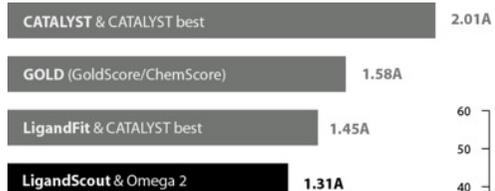
Vergleich Docking & 3D Pharmakophore



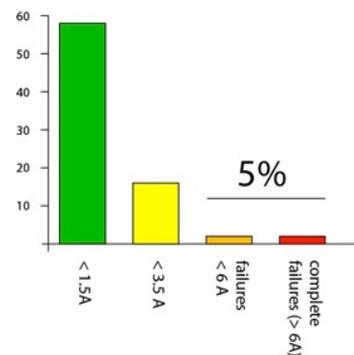
Reproduction quality:
 71% below 1.5Å
 8% failures



Vergleich Docking & 3D Pharmakophore

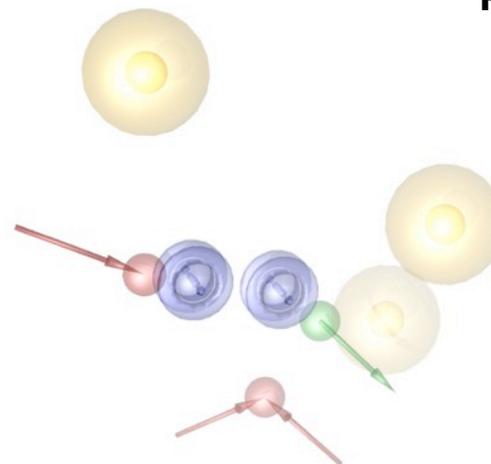


Reproduction quality:
74% below 1.5A
5% failures



G. Wolber, Algorithmische Bioinformatik

Ligandenbasierte Pharmakophore



3D Überlagerung

Performance

- Schnell (ca. 15 ms pro Alignment auf aktueller Hardware)
- Effizient: $O(n^3)$ für Hungarian Matcher; Kabsch: $O(n)$. Im Vergleich: Kompatibilitätssubgraphsuche ist np-vollständig
- Hohe Genauigkeit trotz starker Informationsreduktion

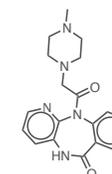
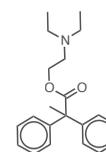
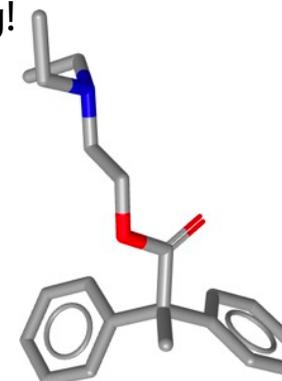
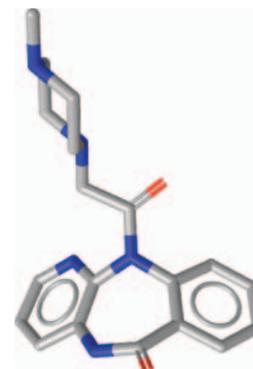
Anwendbarkeit

- Automatischer Vergleich unterschiedlicher Moleküle (generische Distanzcharakteristik)
- Einfache Interpretierbarkeit der Ergebnisse und gemeinsamen Pharmakophore für medizinische Chemiker
- **Robust**

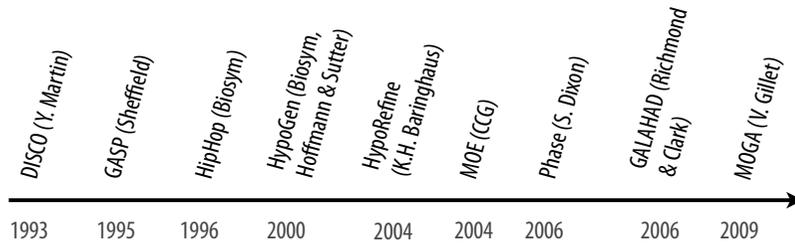
G. Wolber, A. Dornhofer, and T. Langer. J. Comput.-Aided Mol. Design, 20: 773-788 (2006)
T. Seidel, G. Ibis, F. Bendix, G. Wolber. Drug Discovery Today: Technologies, 7(4): e221-e228 (2010)

G. Wolber, Algorithmische Bioinformatik

Überlagerung!

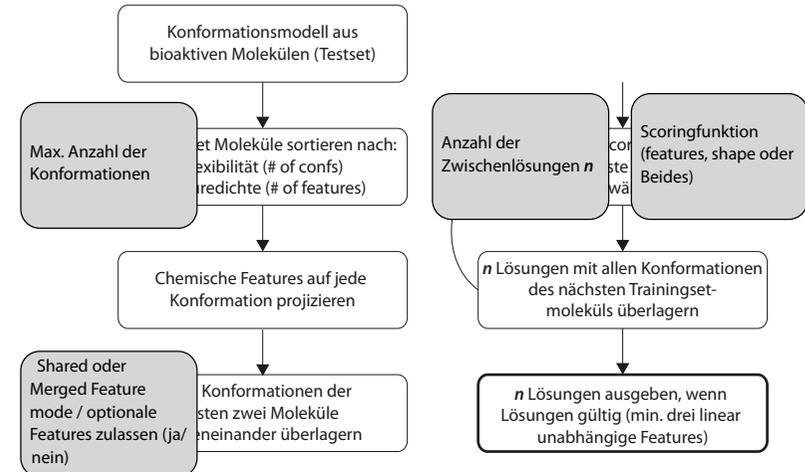


Geschichte: Ligandenbasierte Pharmakophorgenerierung



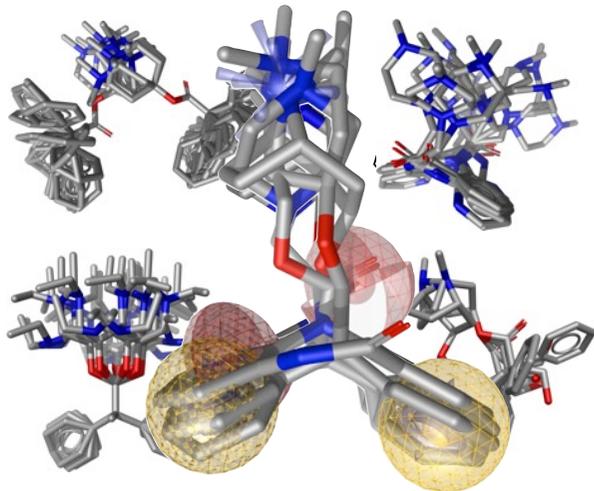
A. Leach et al., J. Med. Chem. 53(2):539-558, 2010
 S. Yang. Drug Discovery Today, 15 (11-12):444-450, 2010

Espresso: Algorithmus



G. Wolber. Algorithmische Bioinformatik

Wie werden ligandenbasierte Pharmakophore generiert?



G. Wolber. Algorithmische Bioinformatik

Virtuelles Screening

Ziel:

- o Identifikation bioaktiver Moleküle aus einer großen Datenbank von Verbindungen durch die Verwendung von Computermodellen

Techniken

- o 3D: Pharmakophormodellierung / Docking
- o 2D: Molekulare Graphen und Graphenbasierte Ähnlichkeitsmetriken (z.B. reduced Graphs)
- o 1D: Deskriptorbasiertes Filtern (Vektor von Skalaren, die ein Molekül beschreiben, z.B. quantifizierte Wasserlöslichkeit, Molekulargewicht, Oberflächenparameter, etc.)

G. Wolber. Algorithmische Bioinformatik

Klassische virtuelle Screening Kaskade



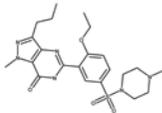
1D Filter

- Wertebereiche für bestimmte Eigenschaften
- Fingerprints

z.B. MW 200-500
Lipinski

2D Filter

- Topologie, Molekulare Subgraphen
- Reduzierte Graphen (z.B., FTrees, ...)



3D Filter

- 2- und 3-Punkt Pharmakophore
- Abstandshashing



3D Überlagerung

- Flexibel
- Vorberechnete Konformations-Ensembles

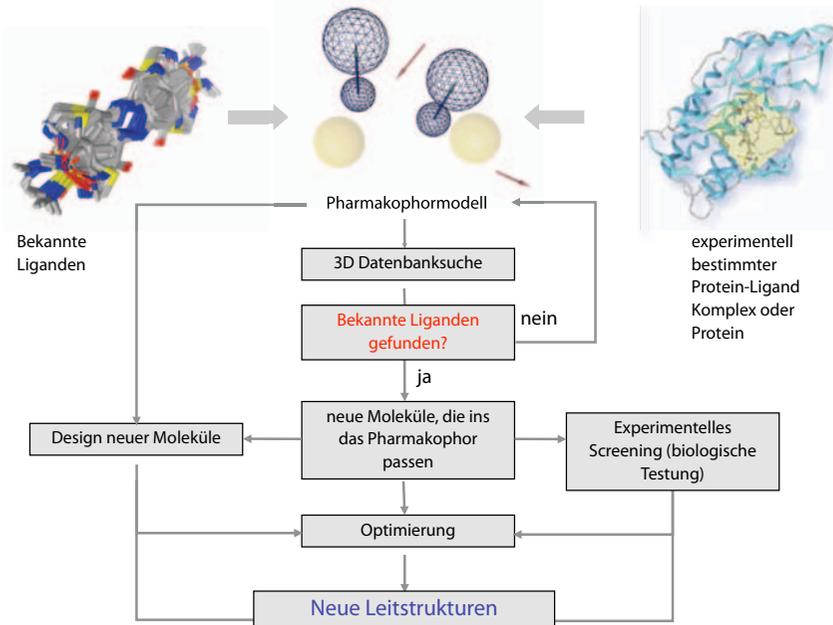
Hoher Rechenaufwand!

Metriken für die Beurteilung des Erfolgs von virtuellem Screening

Sensitivity (Se)

Die Rate der gefundenen tatsächlich biologisch aktiven Verbindungen in Relation zu allen Verbindungen in der Datenbank

$$\frac{\text{[?][?]} \cdot 100}{\text{[?][?] + [?][?]}}$$



Metriken für die Beurteilung des Erfolgs von virtuellem Screening

Specificity (Sp)

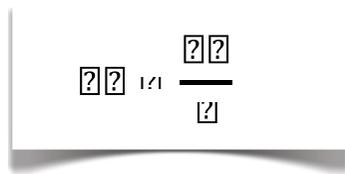
Specificity (Sp) bezeichnet die Anzahl der tatsächlich als inaktiv zurückgewiesenen Verbindungen in der virtuellen Hitliste

$$\frac{\text{[?][?]} \cdot 100}{\text{[?][?] + [?][?]}}$$

Metriken für die Beurteilung des Erfolgs von virtuellem Screening

Yield of Actives (Ya)

'Yield of Actives' ist ein Maß für den Anteil von aktiven Molekülen in der Hitliste



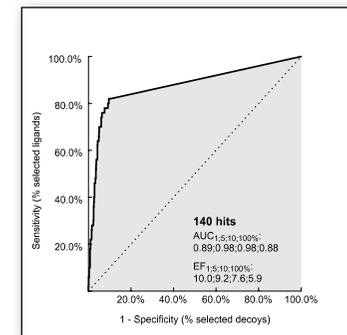
G. Wolber, Algorithmische Bioinformatik

Metriken für die Beurteilung des Erfolgs von virtuellem Screening

$$Se = \frac{N_{selectedactives}}{N_{totalactives}} = \frac{TP}{TP + FN}$$

$$Sp = \frac{N_{discardedinactives}}{N_{totalinactives}} = \frac{TN}{TN + FP}$$

$$AUC = \sum_{k=1}^{k=hits} Se_k((1 - Sp_k) - (1 - Sp_{k-1}))$$

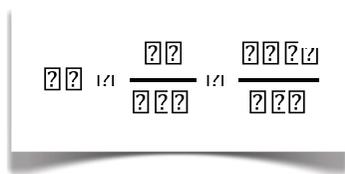


G. Wolber, Algorithmische Bioinformatik

Metriken für die Beurteilung des Erfolgs von virtuellem Screening

Enrichment Factor (EF)

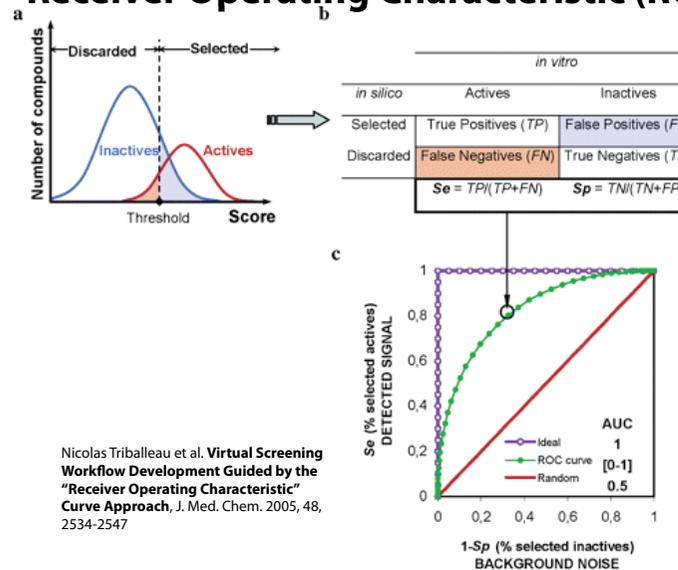
Bezeichnet das Verhältnis von Yield of Actives proportional zur Anzahl der insgesamt gescreenten Verbindungen



Bedeutung vor allem als "Early enrichment" (z.B. EF_{1%} oder EF_{10%})

G. Wolber, Algorithmische Bioinformatik

Receiver Operating Characteristic (ROC)



Nicolas Triballeau et al. **Virtual Screening Workflow Development Guided by the "Receiver Operating Characteristic" Curve Approach**, *J. Med. Chem.* 2005, 48, 2534-2547

G. Wolber, Algorithmische Bioinformatik

3D Pharmakophor Screening

Problem der klassischen Screeningkaskade:



viele unterschiedliche Filter, um 3D Fitting-Aufwand zu reduzieren
keine Garantie, dass nicht passende Verbindungen zu früh gefiltert werden

Alternative:

3D Fitting für alle?

5 Mio. Verbindung mit je 50 Konformationen: ca. 60 CPU Tage -> zu lang

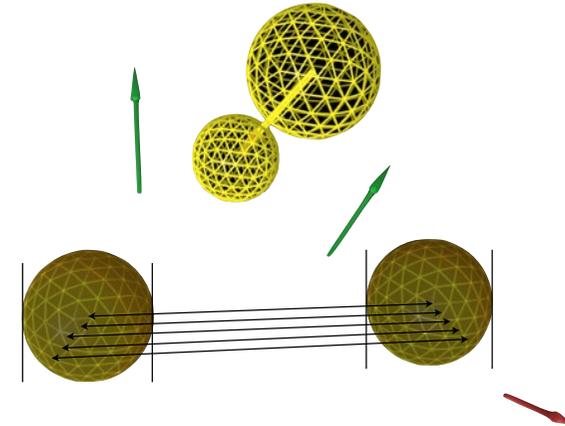
Nur Fingerprints (Distance Hashing)?

Zu ungenau

Lösung: "Sicheres Filtern", das garantiert, dass ausschließlich **nicht passende** Verbindungen gefiltert werden

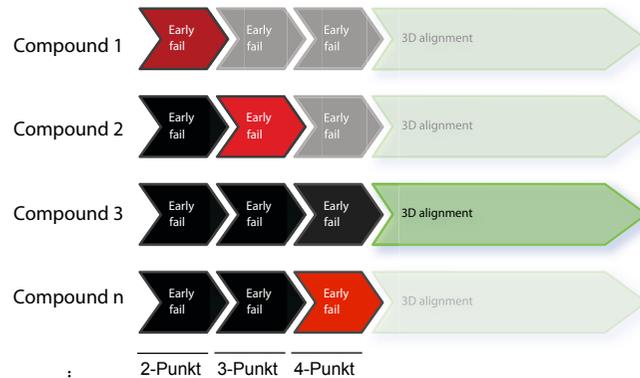
G. Wolber, Algorithmische Bioinformatik

Das Toleranzproblem



G. Wolber, Algorithmische Bioinformatik

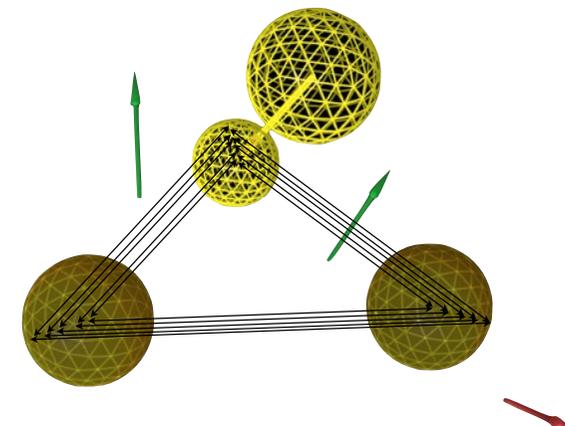
"Sichere" Screening Kaskade



Hashing auf Abstände ("Fingerprints")

G. Wolber, Algorithmische Bioinformatik

Das Toleranzproblem



G. Wolber, Algorithmische Bioinformatik

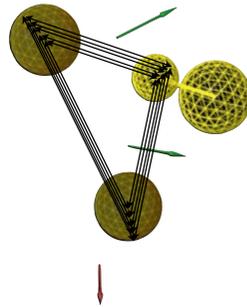
Toleranzexpansion

Problem:

- Für eine Auflösung von 0.1A entstehen um bis zu drei Größenordnungen mehr Fingerprints
- Filtereffizienz nimmt drastisch ab
- Um eine Toleranzen sauber abzutasten ist die Expansion aber notwendig

Lösungsansatz:

- Nicht in der Datenbank vorkommende Fingerprints aus dem Pharmakophor löschen (Bloomfilter)



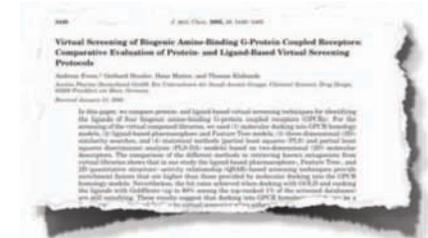
G. Wolber, Algorithmische Bioinformatik

Validierung

Standardisiertes Screening an vier verschiedenen Proteinen (liganden-basiert):

- α 1A Adrenorezeptor
- 5HT_{2A} Serotoninrezeptor
- Dopamin D2 Rezeptor
- Muskarinischer Acetylcholinrezeptor M1

(im Kontext einer Vergleichsstudie verschiedener Pharmakophorschuchprogramme)

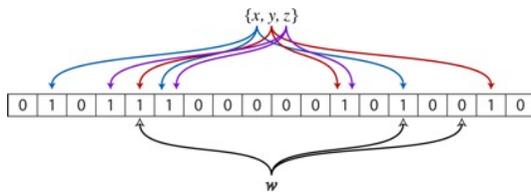


Referenzstudie von A. Evers et al. @ Sanofi Aventis [J. Med. Chem.; 48(17):5448-65, 2005]

G. Wolber, Algorithmische Bioinformatik

Bloomfilter

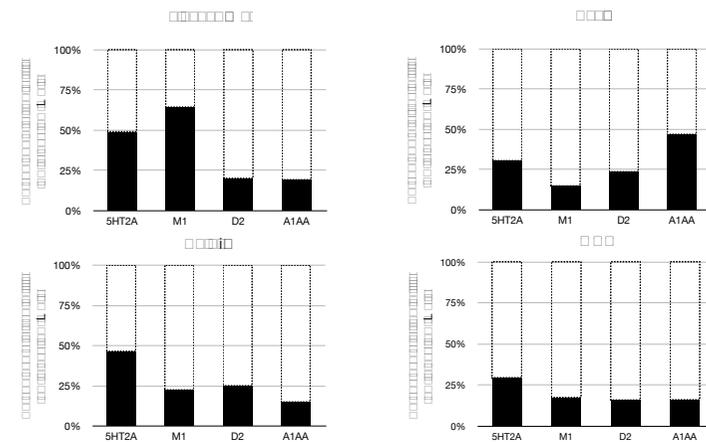
n-Punkt Pharmacophore werden während der Datenbankgenerierung als Bitvektor gespeichert



Schnelle und effiziente Reduktion des toleranzexpandierten Query-Pharmakophors um die im Bloomfilter auf 0 gesetzten Bits

G. Wolber, Algorithmische Bioinformatik

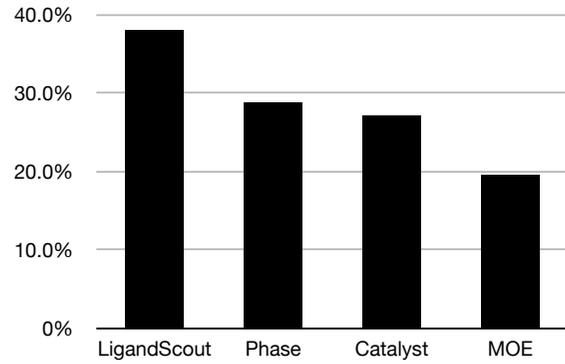
'Yield of actives' für die vier Proteine



True positive retrieval ratios for the four programs. The black bar of each column indicates the fraction of the virtual hits that are true actives, the white part of the column indicates the fraction of false positives. 100% on the y-axis refer to the number of virtual hits obtained from virtual screening.

G. Wolber, Algorithmische Bioinformatik

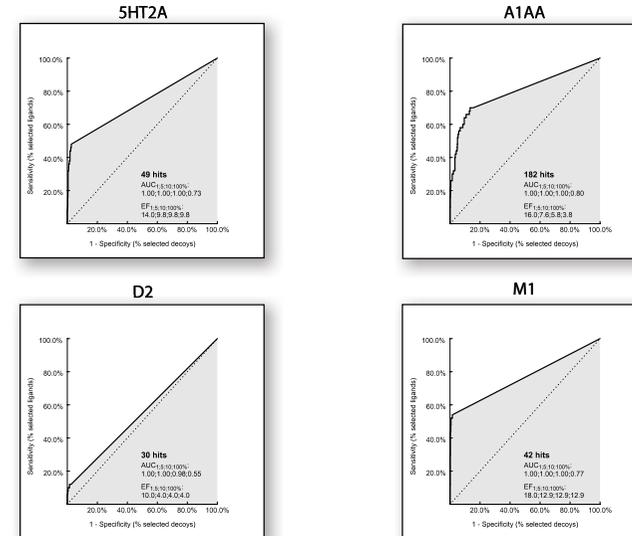
Durchschnittliche 'Yield of Actives'



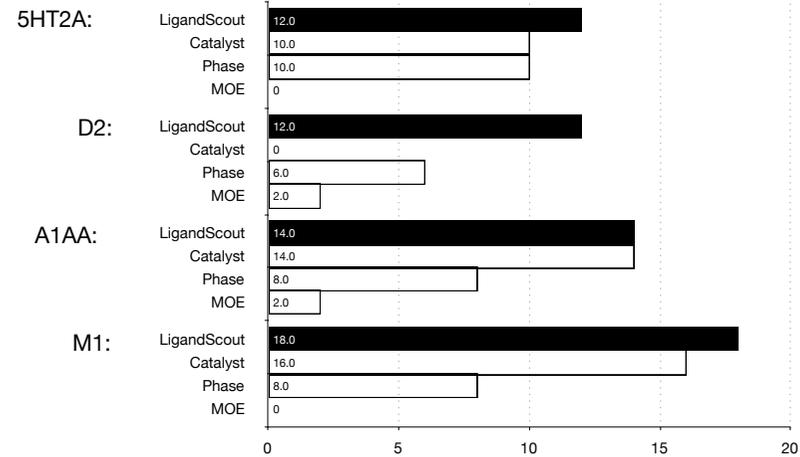
Fraction of true positives in the virtual hit list averaged over all four targets.

G. Wolber, Algorithmische Bioinformatik

LigandScout/Espresso

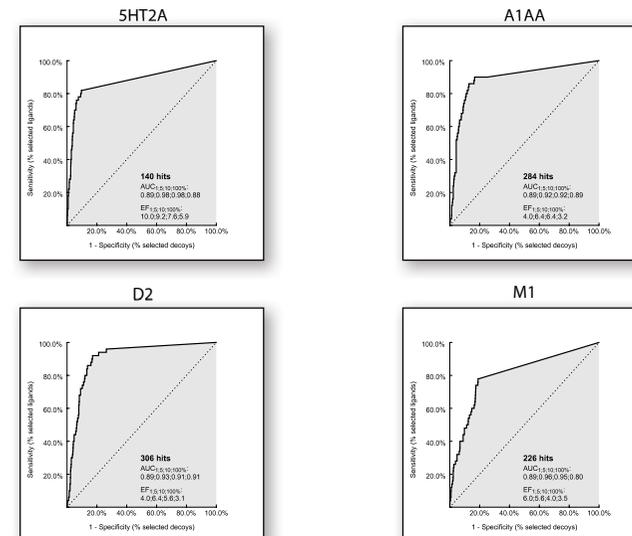


EF_{1%}



G. Wolber, Algorithmische Bioinformatik

MOE



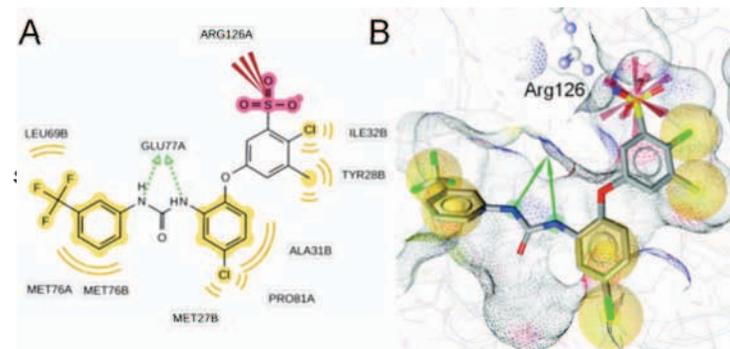
Anwendungen von 3D Pharmakophorscreening

Erfolgreiche Anwendungen von 3D Pharmakophorscreening:

- Neue Hemmstoffe für 17 β -Hydroxysteroiddehydrogenase Typ 3 und 5
- Duale Inhibitoren der mikrosomalen Prostaglandin E Synthase-1 (mPGES-1) und der 5-Lipoxygenase (5-LO)
- Naturstoffscreening: X-Linked inhibitor of apoptosis Protein (XIAP) zur Blockade eine Protein-Protein-Interfaces

G. Wolber, Algorithmische Bioinformatik

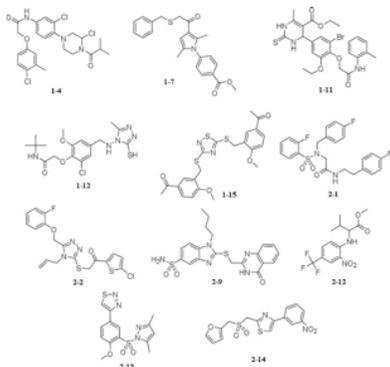
Dual inhibitors of mPGES-1 and 5-LO



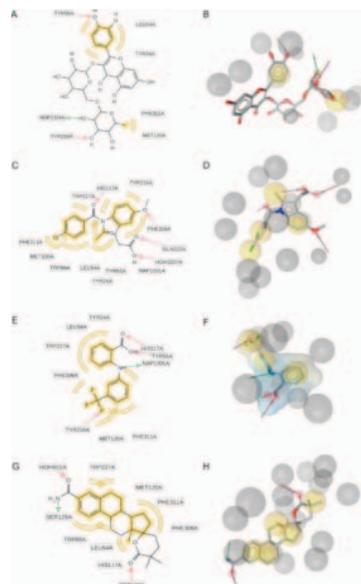
9 active compounds, 8 IC₅₀ lower than 10 μ M, compounds 26 & 28 < 1 μ M

J. Med. Chem., 2011, 54 (9), pp 3163–3174. DOI: 10.1021/jm101309g

Novel inhibitors of 17 β -HSD 3 and 5

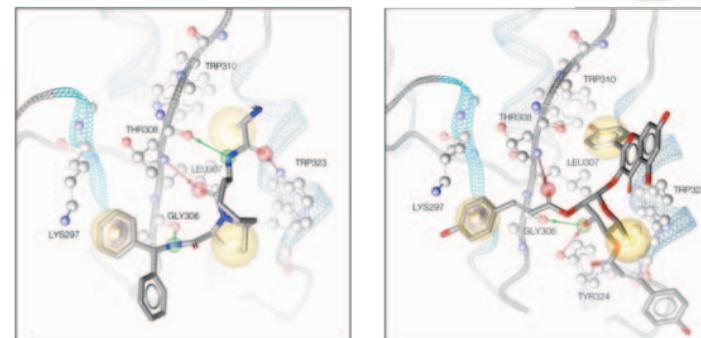


1-12, 2-1, 2-12 ~ 1 μ M, 2-9 ~ 0.3 nM



J Steroid Biochem Mol Biol. 2011 May;125(1-2):148-61; doi:10.1016/j.jsbmb.2011.01.016

Natural product screening: XIAP BIR3 Smac mimetics



XIAP BIR3 natural product inhibitor (IC₅₀ ~10.4 μ M)

Bioorg. Med. Chem.19(2), 2011, pp. 1002-1009; doi:10.1016/j.bmc.2010.10.046

Kraftfelder und Molekulardynamik

Molekulardynamik: Berechnet aus einer molekularmechanischen Sicht die zeitliche Entwicklung eines Systems (z.B. Protein)

Eine einzelne Konformation als Startpunkt

Menge der Konformationen: Trajektorie

Typische Simulationszeiten: Schrittweite in der Größenordnung von Femtosekunden (fs); typisch heute : 2-20ns ("Ergodenhypothese")

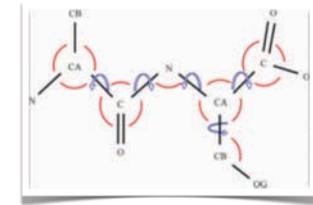
G. Wolber, Algorithmische Bioinformatik

Was ist ein Kraftfeld?

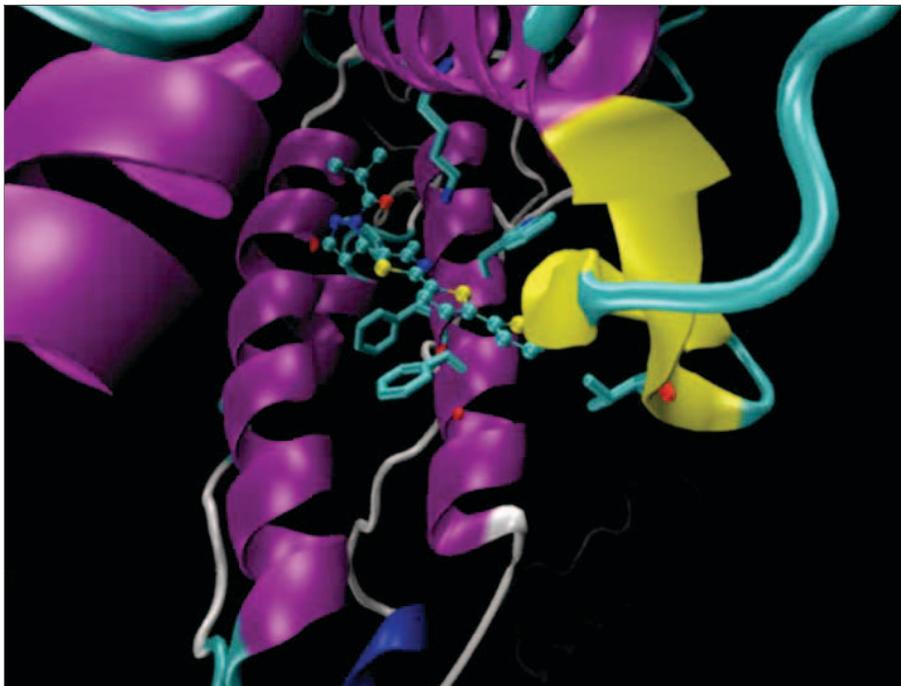
... beschreibt die geometrische Anordnung von chemischen Systemen mittels klassischer Mechanik

Wichtigste Beschreibungsfunktionen:

- Bindungsabstände (schwarz)
- Torsionswinkel (blau)
- Bindungswinkel (rot)



G. Wolber, Algorithmische Bioinformatik



Teilchen

Welche Teilcheneigenschaften spielen in der Chemie eine Rolle?

- Am wichtigsten: Ladung (Elektronen, Protonen)
- Masse (Protonen, Neutronen)
- Spin (Pauli-Prinzip) als Elektroneigenschaft

ABER:

- Chemie ist **nicht** durch elektromagnetische WW zwischen klassischen Punktladungen beschreibbar, sondern durch WW zwischen **Punktquanten**
- **Quanteneigenschaften der Elektronen**
- **Quanteneigenschaften (leichter) Kerne**

G. Wolber, Algorithmische Bioinformatik

Relativistische Quantenmechanik

“klassische” Quantenmechanik

- **Schrödinger-Gleichung:** E. Schrödinger, Ann.Phys.79(1926)79: nicht-relativistische Quantenmechanik ohne Spin: Lösung für Wasserstoff-Atom in erster Publikation!
- **Pauli-Gleichung** (nichtrelativistische Quantenmechanik mit Spin)
- **Dirac-Gleichung:** P.A.M. Dirac, Proc. Roy. Soc. A 117 (1928) 610 (relativistische Quantenmechanik mit Spin und Antiteilchen; Vorhersage des Positrons)
- **“Kopenhagener” Interpretation der Quantenmechanik:**
 - * M. Born, Zeitschr. f. Phys. 38 (1926) 803
 - * W. Heisenberg, Zeitschr. f. Phys. 43 (1927) 172
 - * N. Bohr, Naturwissenschaft 16 (1928) 245

G. Wolber, Algorithmische Bioinformatik

Kraftfelder: Phänomenologischer Ansatz

Chemische Begriffe: keine Einordnung in der Quantenmechanik

Strukturbegriffe

- Moleküle
- Funktionelle Gruppen
- Bindungen (1-fach, mehrfach)
- Atomtypen, freie Elektronenpaare

Begriffe für Wechselwirkungen

- Bindungskräfte
- Winkelspannung
- Torsionskräfte
- Wasserstoffbrücken, charge-transfer- Wechselwirkungen

G. Wolber, Algorithmische Bioinformatik

Lösungen heute

Nicht hochkomplizierte Algebra, sondern numerische Lösungen (“brute-force computing”)

Was ist heute berechenbar?

- bis 8 Kerne auf 1kcal/mol berechenbar
- bis 75 Atome qualitativ richtig
- Quantendynamik (zeitabhängige Kernschrödingergleichung) bis ca. 4 Atome

Proteine viel zu groß für solche “ab initio” Rechnungen!

G. Wolber, Algorithmische Bioinformatik

Kraftfelder: Geschichte

im Wesentlichen durch Computerentwicklung bestimmt:

- ~1970: Erste Kraftfeldversuche für kleine Moleküle
- 70er Jahre: CFF/VFF, ECEPP, **MM2** (Allinger 1977)
- 80er Jahre: CHARMM (Karplus, 1983), AMBER84 (Kollmann 1995), TRIPOS (1985), **GROMOS** (van Gunsteren, 1987), OPLS (Jorgensen 1988)
- 90er Jahre: **CHARM22** (Karplus 1992, 1998)

G. Wolber, Algorithmische Bioinformatik

Kombinatorik in Kraftfeldern

$$\binom{n}{2} + n = \frac{n(n-1)}{2} + n$$

Bindungstypen

$$\binom{n}{3} = \frac{n(n-1)(n-2)}{3 \times 2}$$

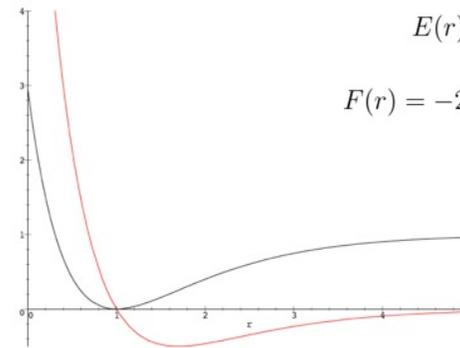
Bindungswinkeltypen

$$\binom{n}{4} = \frac{n(n-1)(n-2)(n-3)}{4 \times 3 \times 2}$$

Torsionstypen

G. Wolber, Algorithmische Bioinformatik

Bindungslängen: Morsepotenzial



$$E(r) = D_e \left(1 - e^{-a(r-r_0)}\right)^2$$

$$F(r) = -2D_e a \left(1 - e^{-a(r-r_0)}\right) e^{-a(r-r_0)}$$

D_e : Potenzialtopftiefe
 a : Streckschwingungsfrequenz der Bindung

G. Wolber, Algorithmische Bioinformatik

Potenziale und Kraftfeldern

Kraftfelder werden meist als Potenzialfelder diskutiert

$$F(r) = -\frac{\partial E(r)}{\partial r}$$

$$E(r) = -\int F(r) dr + E_0$$

G. Wolber, Algorithmische Bioinformatik

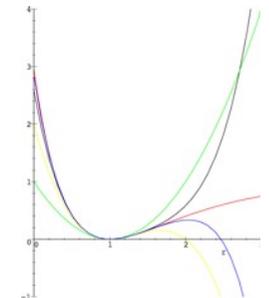
Annäherung von Bindungslängen

Morsepotenzial ist genau, und deckt einen weiten Längenbereich ab, aber

- Berechnung von Exponentialfunktionen sehr teuer
- Bindungslängen sind nicht sehr variabel
- weiter Längenbereich nicht notwendig

Approximation:

- Polynom 2. Grades ("harmonische Näherung" in AMBER, GROMOS, CHARMM und TRIPOS implementiert)
- Polynom 3. Grades: wg. negativem Koeffizienten nicht verwendet
- Polynom 4. Grades: verwendet im MM3

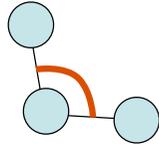


G. Wolber, Algorithmische Bioinformatik

Potenziale für Bindungswinkel

Quadratischer Ansatz:

$$E(\theta) = \frac{k}{2} (\theta - \theta_0)^2$$



Achtung: Energiekonstanten in Grad/rad

AMBER94	θ_0	$k \left[\frac{\text{kcal}}{\text{mol deg}^2} \right]$	$k \left[\frac{\text{kcal}}{\text{mol rad}^2} \right]$
CT-CT-CT	109.5	0.0122	40.0
CT-CM-CM	119.7	0.0213	70.0
H1-CT-H1	109.5	0.0107	35.0

Energie ist um eine Größenordnung kleiner als bei Bindungslängen (bond stretching)
Auch Bindungswinkel sind bei Raumtemperatur weitestgehend konstant

G. Wolber, Algorithmische Bioinformatik

Beschreibung nichtbindender Wechselwirkungen

Wichtigste Wechselwirkungen:

- o Elektrostatische Wechselwirkungen, z.B. Beschreibung als Punktladungen



$$E = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r}$$

ϵ_0 ... Dielektrizitätskonstante (Vakuum)

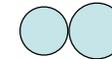
makroskopisch gilt:

$$E = \frac{1}{4\pi\epsilon_0\epsilon} \frac{q_1 q_2}{r}$$

ϵ ... Dielektrizitätskonstante (Lösungsmittel); für H₂O: 78.54

Effekt durch Dipolmoment des Lösungsmittels;
Orientierungspolarisation & Schwingungspolarisation

- o Van der Waals Wechselwirkungen

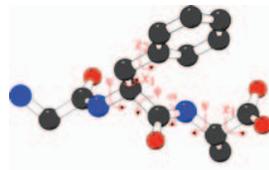
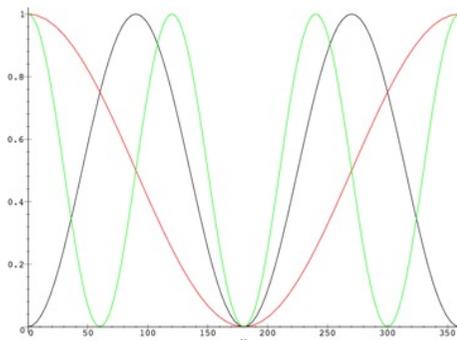


G. Wolber, Algorithmische Bioinformatik

Potenziale für Torsionen

Allgemeine Form:

$$E(\omega) = \sum_{n=0}^N \frac{V_n}{2} [1 + \cos(n\omega - \gamma_n)]$$

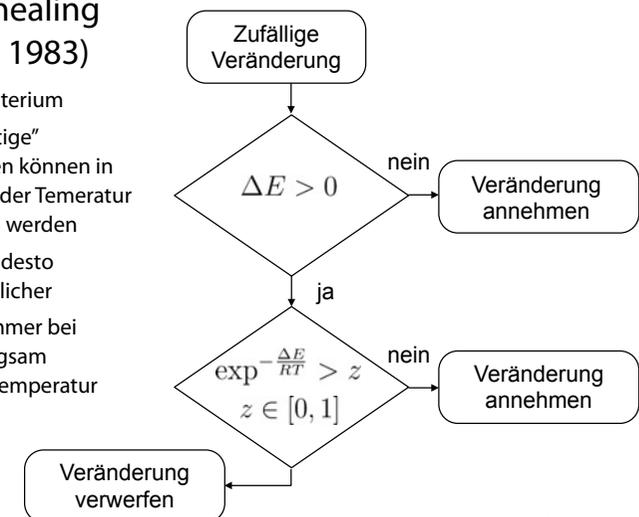


G. Wolber, Algorithmische Bioinformatik

Geometrieoptimierung

Simulated Annealing (kirkpatrick, 1983)

- o Metropolis-Kriterium
- o auch "ungünstige" Veränderungen können in Abhängigkeit der Temperatur angenommen werden
- o Je niedriger T, desto unwahrscheinlicher
- o konvergiert immer bei unendlich langsam abgesenkter Temperatur



G. Wolber, Algorithmische Bioinformatik

Geometrieoptimierung

Powell-Methoden in mehreren Dimensionen

- verschiedene Freiheitsgrade werden hintereinander optimiert
- Heuristiken zur Richtungswahl (in einem langen schmalen Tal ist es essenziell in Richtung des Tals zu minimieren ...)

Steepest Descent

- Energiegradient wird für jeden Term berechnet; Energie wird der Reihe nach nach den einzelnen Freiheitsgraden differenziert; dynamische Schrittweiten abhängig vom Erfolg

Conjugate Gradients

- neben dem aktuellen Gradienten wird auch der vorhergehende Gradient berücksichtigt; gutes Konvergenzverhalten

G. Wolber, Algorithmische Bioinformatik

Zusammenfassung

- Strukturbasiertes Molecular Modeling ist ein Werkzeug zum Verständnis und zur Visualisierung der Interaktion von Proteinen und Liganden
- Alle Methoden sind (sehr nützliche) Näherungen für das Design kleiner organischer Moleküle
- Virtuelles Screening erlaubt die 'Rationalisierung' von Arzneistoffwirkungen, verringert Entwicklungsrisiken und generiert neue Ideen
- Molekulardynamik ermöglicht das Verständnis sehr spezifischer Problemstellungen (v.a. in Kombination mit experimentellen Daten)



G. Wolber, Algorithmische Bioinformatik