Structure-Based and Computer-Aided Drug Design

Hugo Kubinyi
University of Heidelberg and BASF AG, Ludwigshafen, Germany
E-Mail kubinyi@t-online.de
http://home.t-online.de/home/kubinyi
University of Catanzaro, Italy
June 2003

Drug Design Strategies

- no protein 3D structure, no ligands
- protein 3D structure, no ligands
- combichem, HTS, virtual screening
de novo design (protein flexibility!)
- no protein 3D structure, ligands
- protein 3D structure, ligands
- pharmacophores, (3D) similarity, (3D) QSAR
- structure-based design

Protein Crystallography

Problems of Resolution of Protein 3D Structures

- lacking hydrogens, no differentiation between C, O and N (Thr, asp, asn, glu, gln, orientation of amide groups, imidazole, etc.)

Thrombin and Thrombin Inhibitor Complexes

Thrombin-Thrombstop Complex

Structure-Based Design of Captopril

Carboxypeptidase

substrate

inhibitor

IC_{50} = 330 µM

Captopril IC_{50} = 23 nM
Success Stories and Failures of Structure-based Design

Design of HIV Protease Inhibitors

Influenza
In 1918/19, the “Spanish Flu” killed about 20-40 mio people. Especially young and very old people died from influenza. The heavy death toll of this pandemic disease has to be compared to the number of 11 mio victims of World War I.

Egon Schiele prepared this drawing of his wife, one day before her death and four days before he died himself, only 28 years old.

Influenza Virus schematic view and electron microscopic picture

Hemagglutinin + sialic acid (green) neuraminidase + DANA (red)
**Design of Neuraminidase Inhibitors**

Sialic acid, R = H

- Glu-276
- Arg-371
- Arg-118
- Glu-276
- Arg-292
- Arg-227

- **Neu5Ac2en**
  - $K_i = 1,000$ nM

- **4-Guanidino-Neu5Ac2en**
  - $K_i = 0.1-0.2$ nM

- **Zanamivir (Relenza, Glaxo-Wellcome)**

**Design of Bioavailable Neuraminidase Inhibitors**

- **IC$_{50}$ = 6.3 µM**
- **IC$_{50}$ > 200 µM**

**4-NH$_2$-Neu5Ac2en**

- $K_i = 50$ nM

**GS 4071, R = CH(Et)$_2$**

- $IC_{50} = 1$ nM

**GS 4104 (ester prodrug of GS 4071)**

- Oseltamivir (Tamiflu, Roche)

**The Relevance of Protein Crystal Structures: Conformational flexibility of PNP in the crystal**

- Unfavorable hydrogen bonds?
- Favorable hydrogen bonds
- Unfavorable polar/nonpolar and steric interactions

---

*University of Heidelberg*
The Relevance of Protein Crystal Structures: Enzymatic activity of PNP crystals

\[
\text{PNP} \rightleftharpoons \text{PNP}^+ + \text{HPO}_4^{2-}
\]

reaction rate (% product)

remove crystal

add crystal

time

CAVEAT - An Idea Generator for Peptidomimetics

extract vectors

search for appropriate scaffold

attach and modify the side chains

Rational Design of Integrin Receptor Ligands

\[
\text{cyclo(RGDFv)} \rightarrow \text{K_i} = 2.0 \text{nM}
\]

\[
\text{R G D F v} \rightarrow \text{K_i} = 2.3 \text{nM}
\]

Superposition of Integrin Receptor Ligands

\[
\text{SB 214 857} \rightarrow \text{K_i} = 2.5 \text{nM}
\]

\[
\text{SB 223 245} \rightarrow \text{K_i} = 2.5 \text{nM}
\]

Selectivity of Integrin Receptor Ligands

\[
\text{p-amidine} \rightarrow \text{K_i} = 26 \text{nM}
\]

\[
\text{m-amidine} \rightarrow \text{K_i} = 56,000 \text{nM}
\]

\[
\text{X} = \text{N} \rightarrow \text{K_i} = 8 \text{nM}
\]

\[
\text{X} = \text{CH} \rightarrow \text{K_i} = 1,000 \text{nM}
\]

Highly Selective Integrin Receptor Ligands

\[
\text{SB 214 857} \rightarrow \text{K_i} = 10,340 \text{nM}
\]

\[
\text{SB 223 245} \rightarrow \text{K_i} = 56,000 \text{nM}
\]

\[
\text{K_i} = 2.5 \text{nM}
\]

\[
\text{K_i} = 2.5 \text{nM}
\]
Benzamidine in the Trypsin Specificity Pocket

Peter Andrews Diagram

Enthalpic and Entropic Contributions to Ligand Affinity

Combinatorial Chemistry - Some Theses

Drug Design is an evolutionary procedure
Combinatorial chemistry speeds up drug discovery
Lead discovery libraries shall have a high degree of chemical diversity
Lead optimisation libraries shall have a high degree of similarity, to cover the chemical space around a lead structure
Several small libraries generate a higher diversity than one large library
Drug-like character is more important than synthetic accessibility

The Chemical Universe

$10^{40} - 10^{120}$ compounds with C, H, O, N, P, S, F, Cl, Br, I, and MW < 500 ???
**Favourable Drug Properties**

- High Affinity and Selectivity
- Synthetic Accessibility
- No Chemically Reactive Groups (Garbage Filter)
- Oral Bioavailability
- Lipinski (Pfizer) “Rule of Five”
- Favourable Pharmacokinetics
- Metabolism
- Elimination Pathway/s
- Lack of Side Effects
- Lack of Toxic Effects

“A Hit is no Ligand is no Lead is no Drug”

**“Drug-like” Character of Organic Compounds**

- ACD Training sets vs. WDI Test set predictions

**“Drug-like” Character of Drugs**

100 Top-Selling Drugs, 1997

**TOPAS (TOPology-Assigning System)**

Scaffolds and building blocks from a RECAP process are re-assembled by their 3D similarity to the template (“fragment-based evolutionary design”).


**Virtual Screening, Analogy and Optimization**

Chemical analogy

IC$_{50}$ hKv1.5 = 4.8 µM

Optimization

IC$_{50}$ hKv1.5 = 0.16 µM

Virtual Screening of αβ1 Integrin Antagonists With CATALYST

3D structure of ligand modelled from vascular cell adhesion molecule-1 (VCAM-1) 3D structure; CATALYST search in 8,824 ACD molecules with free NH₂ or NO₂

lead: R = -Leu-Asp-Val-OH  IC₅₀ = 0.6 nM

best results:

IC₅₀ = 67 nM  IC₅₀ = 58 nM  IC₅₀ = 1.3 nM


A Virtual Screening Success Story

Comparison of the performance of high-throughput screening and virtual screening of potential leads of protein tyrosine phosphatase 1B (PTP1B):

a) High throughput screening of 400,000 compounds from a corporate collection
   300 hits < 300 µM,
   85 validated hits with IC₅₀ <100 µM
   = 0.021 % hit rate (many violate Lipinski rules)

b) Virtual screening of 235,000 commercially available compounds, using DOCK, version 3.5
   365 high-scoring molecules,
   127 with IC₅₀ <100 µM
   = 34.8% hit rate (hits are more drug-like)


Virtual Screening, Carbonic Anhydrase Inhibitors

A 3D search in a database of ≈ 90,000 compounds yielded 3,314 molecules; these were rank-ordered by their pharmacophores, 100 were finally docked and 13 docking hits were biologically tested.

X = S  Kᵢ = 0.9 nM
X = SO₂  Kᵢ = 0.8 nM  Kᵢ' = 0.6 nM


Combinatorial Design of Carbonic Anhydrase Inhibitors

Program CombiSMoG, „best“ N-substituents from 100,000 candidates (20 scored by knowledge-based potentials)

B. A. Grzybowski et al., Acc. Chem. Res. 35, 261-269 (2002);

Fragment-Based Approach (Astex)

a) complex binding site
b) fragments in different pockets
c) assembly of different fragments
d) growing out from a fragment


Binding Constants of Biotin and Analogs

(B. M. Green, Adv. Protein Chem. 29, 85-133 (1975))

Biotin, Kᵢ = 1.3 x 10⁻¹⁵ M
Desthiobiotin, Kᵢ' = 5 x 10⁻¹³ M

Biotin, Kᵢ = 3.4 x 10⁻⁵ M  Kᵢ' = 3 x 10⁻³ M
**Re-discovering** Biotin by NMR Fragment-Based Screening

Streptavidin plus two biotin fragments; intermolecular NOEs indicate the “correct” linkage of the fragments (A. Kline et al., The NMR Newsletter 472, 13 (1997))

---

**SAR by NMR** (S. Fesik, 1996)

---

Dynamic Ligand Assembly in a Binding Site


---

Ligand Assembly in Carbonic Anhydrase

- Amines
- Aldehydes
- Side Products
- Products


---

FlexX (GMD, BASF): Dissection of a Ligand

Start fragment

Selection of a start fragment
Alternate positions of the start fragment
Next fragments
Next fragments

FlexX: Search Tree for Ligand Docking

Binding of Methotrexate to DHFR

Combinatorial Ligand Construction

Further Developments in Computer-Aided Drug Design

Ligand flexibility
Protein flexibility
Flexibility of the ligand-protein complex
Geometry and strength of hydrogen bonds
Solvation and desolvation effects (entropy)
Synthetic accessibility of ligands
Combinatorial docking of ligands
„Last Problem“ - The Scoring Function