Changing Paradigms in Drug Discovery

Hugo Kubinyi
University of Heidelberg
Germany
E-Mail kubinyi@t-online.de
Homepage www.kubinyi.de

Cutting Edge Approaches to Drug Design,
Royal Society of Chemistry Molecular Modelling Group Meeting, Oct. 19, 2005

Yesterday’s Drug Discovery Process

Natural Leads
Isolation
Synthetics
Animal Tests
Clinics
Technological Changes in Drug Research

Up to the 70s
Chemistry and hypotheses guide the syntheses
Bottleneck: Animal experiments, isolated organs

Up to the 90s
Molecular Modelling
In vitro models (enzyme inhibition, receptor binding)
Bottleneck: Dedicated syntheses of drugs

Up to the year 2000:
Gene technology (production of proteins)
Combinatorial chemistry (mixtures, chemistry-driven)
Structure-based design of ligands
High-throughput test models (HTS)
Bottleneck: ADMET properties

Today’s Drug Discovery Process

Genome
Proteome
3D Structures
CombiChem
Automated HTS
Virtual Screening
Docking and Scoring
Technological Changes in Drug Research

Today:

Genomics, proteomics and bioinformatics
Transgenic animals for proof of concept
Combinatorial chemistry
  (single compounds, design-driven)
Structure-based and computer-aided design of ligands
Ultra-high-throughput test models (u-HTS)
Data mining
Virtual screening
ADMET profiles (HTS and \textit{in silico})

\textbf{Bottleneck}: Target validation, “drugable” targets

The Productivity Gap in Pharmaceutical Industry

Disadvantages of Traditional Medicinal Chemistry

Complex and time-consuming syntheses
Low diversity (insufficient for new lead discovery)
Synthetic output too small
Slow development of structure-activity profiles within a class of compounds
Slow optimization in evolutionary cycles
Insufficient patent coverage
High costs (about 5,000 – 10,000 US-$ per compound)
A compound

is no hit

is no lead

is no candidate

is no drug

Success in Drug Research

hundred thousands

thousands

dozens

thousands

some

optimization:

1
Types and Features of Combinatorial Libraries

Random libraries
- druglike
- diverse scaffolds

Chemogenomics
Targeted libraries
- target-directed
- diverse substitution

(lead families)

Focused libraries
- similar to lead
- complete

Strategies in Drug Design

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

LBDD  SBDD
Protein Crystallography of Inhibitor Complexes

Thrombin-Thrombostop Complex

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.)
Influenza Virus schematic view and electron microscopic picture

Design of Neuraminidase Inhibitors

- sialic acid, $R = H$
- Neu5Ac2en, $K_i = 1000$ nM

- Arg-371
- Arg-292
- Arg-118
- Glu-276
- Glu-119
- Glu-227

result of a GRID search with a positively charged probe
Design of Neuraminidase Inhibitors

![Chemical structures](image)

**sialic acid,** $R = H$

**Neu5Ac2en**

$K_i = 1\,000\,\text{nM}$

- **Arg-292**
- **Arg-118**
- **Glu-276**
- **Arg-371**
- **Glu-119**
- **Glu-227**

4-Guanidino-Neu5Ac2en

$K_i = 0.1-0.2\,\text{nM}$

Zanamivir (Relenza, Glaxo-Wellcome)

---

The New Technologies

Do we already live in Castalia, the land of Hermann Hesse's novel "The Glass Bead Game", where the Magister Ludi (sic!) organizes and plays the most wonderful, brilliant, exciting and elaborate game ... without any practical relevance?

New Technologies: Open Questions

Is there a „druggable genome“?
Is a target focus always best?
Is poor ADME the main problem?
Are we using the right virtual screening techniques?
What are the problems in virtual screening?
What’s wrong and could we do better?


Genome, Druggable Genome and Drug Targets

Human genome ~30,000

Druggable genome ~3,000

Drug targets ~600–1,500

Disease-modifying genes ~3,000

Is there really a „druggable genome“?

Alternative splicing and posttranslational modification generate a multitude of proteins
→ the „druggable proteome“?

Protein complexes (nAChR, GABA-R, integrins, heterodimeric GPCRs, cross-talking)
→ the „druggable targetome“?

Balanced activity against a series of targets
→ the „druggable physiome“


Is Target Focus the Best Strategy?

Olanzapine, a clozapine-like “atypical” neuroleptic with a promiscuous binding pattern

\[
\begin{align*}
K_i 5-\text{HT}_{2A} &= 4 \text{ nM} & 2.5 \text{ nM} \\
K_i 5-\text{HT}_{2B} &= 12 \text{ nM} \\
K_i 5-\text{HT}_{2C} &= 11 \text{ nM} & 2.5 \text{ nM} \\
K_i 5-\text{HT}_3 &= 57 \text{ nM} \\
K_i \text{ dop} D_1 &= 31 \text{ nM} & 119 \text{ nM} \\
K_i \text{ dop} D_2 &= 11 \text{ nM} \\
K_i \text{ dop} D_3 &= 27 \text{ nM} \\
K_i \text{ musc} M_1 &= 1.9 \text{ nM} & 2.5 \text{ nM} \\
K_i \text{ musc} M_2 &= 18 \text{ nM} \\
K_i \text{ musc} M_3 &= 25 \text{ nM} & 13 \text{ nM} \\
K_i \text{ musc} M_4 &= 13 \text{ nM} & 10 \text{ nM} \\
K_i \text{ musc} M_5 &= 6 \text{ nM} \\
K_i \text{ adr} \alpha_1 &= 19 \text{ nM} \\
K_i \text{ adr} \alpha_2 &= 230 \text{ nM} \\
K_i \text{ hist H}_1 &= 7 \text{ nM}
\end{align*}
\]

a) F. P. Bymaster et al., Neuropsychopharmacology 14, 87-96 (1996)
Reasons for Failure in Drug Development
(n = 198)

T. Kennedy, Drug Discov. today 2, 436-444 (1997)

Pharmacokinetics 39%
Lack of efficacy 10%
Animal toxicity 5%
Adverse effects in man 11%
Commercial reasons 5%
Miscellaneous 30%

"Discouraging data on the antidepressant."
Reasons for Failure in Drug Development
(n = 121; without antiinfectives)

- Pharmacokinetics: 7%
- Lack of efficacy: 7%
- Animal toxicity: 7%
- Adverse effects in man: 16%
- Commercial reasons: 17%
- Miscellaneous: 46%

T. Kennedy, Drug Discov. today 2, 436-444 (1997)

University of Heidelberg

Drug Research is ....

the Search for a Needle in a Haystack
Virtual Screening Reduces the Size of the Haystack by Selecting:

- Compounds or libraries that are either lead-like, or drug-like, or have the potential of oral bioavailability, or are similar to a lead,
- by rules (e.g. Lipinski bioavailability rules), neural nets (e.g. drug-like character), pharmacophore analyses, similarity analyses, scaffold hopping, or docking and scoring

Problems in Virtual Screening

- Ionisation and Dissoziation (Sadowski rules, ACS Boston, 2002)
- Tautomeric and protomeric forms (program AGENT, ETH Zurich; ChemoSoft tautomer recognition, ChemDiv)
- Acceptor properties of oxygen and sulfur atoms (esters, aromatic ethers, oxazoles, isoxazoles, thiazoles, etc.)
- Too many filters?
Donor and Acceptor Properties of O and N

Filters for Virtual Screening

Garbage filter 100%
Druglike / Non-druglike 80%
Bioavailability :
Cytotoxicity :
hERG channel inhibition :
Antitargets :
  α1a (orthostatic hypotension) :
  D2 (extrapyramidal syndrome) :
  5-HT2c (obesity) :
  musc. M1 (hallucinations, memory) :
CYP inhibition (3A4, 2C9, 2D6) 0% ?
A Virtual Screening / Docking 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)


Stepwise Virtual Screening

Aventis in-house compound repository

22,950 compounds

21 MW, rot-bond filter, 3D pharmacophore search

docking into an α1A receptor model (GOLD, PMF)

300 top-scoring compounds

clustering, diversity selection

80 compounds tested, 37 hits with Kᵢ < 10 µM

α₁A adrenergic receptor antagonist, Kᵢ = 1.4 nM

Virtual Screening of Carbonic Anhydrase Inhibitors

- 98,850 compounds (LeadQuest and Maybridge libraries)
- Filter for Zn$^{2+}$-binding anchor groups: 5,904 hits
- 2D and 3D pharmacophore searches (derived from binding site analysis): 3,314 hits
- FlexS superposition with dorzolamide, followed by FlexX docking of 100 hits into carbonic anhydrase binding site: 13 hits

Binding Constants of Biotin and Analogs

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

- Biotin, $K_i = 1.3 \times 10^{-15}$ M
- Desthiobiotin, $K_i = 5 \times 10^{-13}$ M
- $K_i = 3.4 \times 10^{-5}$ M
- $K_i = 3 \times 10^{-3}$ M
SAR by NMR
(P. J. Hajduk et al., J. Am. Chem. Soc. 119, 5818-5827 (1997))

FlexX (GMD, BASF): Dissection of a Ligand
Binding of Methotrexate to DHFR

Combinatorial Design of Carbonic Anhydrase Inhibitors

**start structure**

![Start Structure Image]

**optimized structure**

![Optimized Structure Image]

$K_d = 120 \text{ nM}$

$R$ enantiomer, $K_d = 30 \text{ pM}$

(S enantiomer: $K_d = 230 \text{ pM}$)

Program CombiSMoG, selection of “best” N-substituents from 100,000 candidate structures (20 of them scored by knowledge-based potentials)

B. A. Grzybowski et al., Acc. Chem. Res. 35, 261-269 (2002);
Scaffold-Linker-Functional Group Approach

Design of a structure-based 320-member virtual library with four different scaffolds or ring connections, five linkers and 16 different functional groups; best docking results with FlexX

The Past

Voltaire (1694-1778):

Doctors pour drugs of which they know little, to cure diseases of which they know less, into human beings of whom they know nothing.

The Future: Pharmacogenomics - New Opportunities from Personalized Medicine

Genotyping of drug targets and metabolic enzymes enables

- cost savings in drug development through better design of clinical trials
- selection of the „best drug“ for a certain patient
- individual dose ranges (variance in target sensitivity, reduced or increased metabolism)
- fewer toxic side effects
- fewer unexpected drug-drug interactions
Gefitinib®, Iressa, ZD1839 (EGFR TK inhibitor)

- Third-line therapy for non-small-cell lung cancer (75% of lung cancer cases)
- Clinical response to Iressa ~ 10%

8 out of 9 Iressa-responsive patients showed mutations in the kinase domain
0 out of 7 non-responsive patients showed mutations
2 out of 25 non-treated patients showed mutations (8%)

J. G. Paez et al.
EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy
Science 304 (5676), 1497-1500 (2004)

T. J. Lynch et al.
Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non-Small-Cell Lung Cancer and Gefitinib