Structure based drug design

Dr. Pushpalatha G. Associate Professor Dept of Plant Biotechnology, MSSSoA, CUTM

Drug Design & Discovery: Introduction



Drugs:

Targets:

Natural sources



Synthetic sources





Discovering and Developing the 'One Drug'



Profile of Today's Pharmaceutical Business

•Time to market: 10-12 years. By contrast, a chemist develops a new adhesive in 3 months! Why? (Biochemical, animal, human trials; scaleup; approvals from FDA, EPA, OSHA)



Nature Reviews | Drug Discovery

Pharmaceutical R&D A Multi-Disciplinary Team

Administrative Support Analytical Chemistry Animal Health Anti-infective Disease Bacteriology

Behavioral Sciences Biochemistry Pilegy Biometrics Cardiology Conditional Science Clinical Research Communication Computer Science Ito news evelopi en allowing DNA Sequencing Diabetology

Document Preparation Dosage Form Development Drug Absorption Drug Degradation Drug Delivery

Electrical Engineering Electron Micros Description Electrony Environmental Health & Safety Employee Resources Endocrinology Enzymolo Date it as A. Con in the indication of the formulation

Gastroenterology Graphic Design Histomorphology Intestinal Permeability Law Library Science Medical Services

Mechanical Engineering Methodal Chemistry Kolecula Fology Molecular Genetics Molecular Models Natural Products Nobiology Nobel Contended of the Contended of th

Oncology Organic Chemistry Pathology Peptide Chemistry Pharmacokinetics Pharmacology Photochemistry

Physical Why the Physical Chemistry Physical Chemistry Lange Burder Physical Chemistry Ph

Radiochemistry Radiology Robotics Spectroscopy Statistics Sterile Manufacturing Tabletting Taxonomy Technical Information Toxicology Transdermal Drug Delivery Veterinary Science Virology X-ray Spectroscopy Medicinal chemists today are facing a serious challenge because of the increased cost and enormous amount of time taken to discover a new drug, and also because of fierce competition amongst different drug companies

Drug Discovery & Development



Identify disease

Drug Design
 Molecular Modeling
 Virtual Screening

Isolate protein involved in disease (2-5 years) Find a drug effective against disease protein (2-5 years) Scale-up

Preclinical testing (1-3 years)

Human clinical trials (2-10 years)

Formulation



Technology is impacting this process



Identify disease

Isolate protein

COMBINATORIAL CHEMISTRY

Rapidly producing vast numbers of compounds

MOLECULAR MODELING

Computer graphics & models help improve activity

IN VITRO & IN SILICO ADME MODELS *Tissue and computer models begin to replace animal testing*

GENOMICS, PROTEOMICS & BIOPHARM.

Potentially producing many more targets and "personalized" targets

HIGH THROUGHPUT SCREENING

Screening up to 100,000 compounds a day for activity against a target protein

VIRTUAL SCREENING

Using a computer to predict activity

Find drug

Preclinical testing

History of Drug Discovery....

1909 - First rational drug design.
Goal: safer syphilis treatment than Atoxyl.
Paul Erhlich and Sacachiro Hata wanted to maximize toxicity to pathogen and minimize toxicity to human (therapeutic index).
They found Salvarsan (which was replaced by penicillin in the 1940's)



•1960 - First successful attempt to relate chemical structure to biological action quantitatively (QSAR = Quantitative structure-activity relationships). Hansch and Fujita

History of Drug Discovery

- Mid to late 20th century
 - understand disease states, biological structures, processes, drug transport, distribution, metabolism.

Medicinal chemists use this knowledge to modify chemical structure to influence a drug's activity, stability, etc.

• procaine = local anaesthetic; Procainamide = antirhythmic

$$H_{2}N \longrightarrow OCH_{2}CH_{2}N(C_{2}H_{5})_{2} H_{2}N \longrightarrow OHCH_{2}CH_{2}N(C_{2}H_{5})_{2}$$

$$Procaine Procainamide$$

Drug Discovery overview

Approaches to drug discovery:
Serendipity (luck)
Chemical Modification
Screening
Rational

Serendipity "Chance favors the prepared mind"

1928 Fleming studied Staph, but contamination of plates with airborne mold. Noticed bacteria were lysed in the area of mold. A mold product inhibited the growth of bacteria: the antibiotic penicillin





Chemical Modifications

(A) Homolog Approach: Homologs of a lead prepared R-(CH₂)n-N Homologs (n=2,3,4,5.....)

(B) Molecular Disconnection /Simplification





MORPHINE

PENTAZOCINE

H₂N

MEPROBAMATE

соон —

(C) Molecular Addition

(D) Isosteric Replacements

SO2NH2

SULFANILAMIDE



Chemical Modification...

•Traditional method.

• An analog of a known, active compound is synthesized with a minor modification, that will lead to improved Biological Activity.



Advantage and Limitation: End up with something very similarto what you start with.ReferenceNew molecule



Screening

Testing a random and large number of different molecules for biological activity reveals leads. Innovations have led to the automation of synthesis (combinatorial synthesis) and testing (high-throughput screening).

Example: Prontosil is derived from a dye that exhibited antibacterial properties.



Irrational, based on serendipity & Intuition

Trial & error approach

Time consuming with low through output

No de novo design, mostly "Me Too Approach"

Rational Drug Design - Cimetadine (Tagamet)

Starts with a validated biological target and ends up with a drug that optimally interacts with the target and triggers the desired biological action.

Problem: histamine triggers release of stomach acid. Want a histamine antagonist to prevent stomach acid release by histamine = VALIDATED BIOLOGICAL TARGET.



Histamine analogs were synthesized with systematically varied structures (chemical modification), and SCREENED. N-guanyl-histamine showed some antagonist properties = LEAD compound.

Rational Drug Design - Cimetadine (Tagamet) - continued

a. Chemical modifications were made of the lead = LEAD OPTIMIZATION:



b. More potent and orally active, but thiourea found to be toxic in clinical trials



c. Replacement of the group led to an effective and well-tolerated



d. Eventually replaced by Zantac with an improved safety profile



First generation Rational approach in Drug design

In 1970s the medicinal chemists considered molecules as topological entities in 2 dimension (2D) with associated chemical properties. QSAR concept became quite popular. It was implemented in computers and constituted first generation rational approach to drug design

2nd generation rational drug design

The acceptance by medicinal chemists of molecular modeling was favored by the fact that the QSAR was now supplemented by 3D visualization. The "lock and key" complementarity is actually supported by 3D model. **Computer aided molecular design** (CAMD) is expected to contribute to intelligent lead

MECHANISM BASED DRUG-DESIGN

Most rational approach employed today.

 Disease process is understood at molecular level & targets are well defined.

 Drug can then be designed to effectively bind these targets & disrupt the disease process

 Very complex & intellectual approach & therefore requires detailed knowledge & information retrieval. (CADD Holds Great Future)

<u>"Drug –Receptor Interaction is not merely a lock-key interaction but a</u> dynamic & energetically favorable one"

Evolutionary drug designing Ancient times: Natural products with biological activities used as drugs. Chemical Era: Synthetic organic compounds Rationalizing design process: SAR & **Computational Chemistry based Drugs** Biochemical era: To elucidate biochemical pathways and macromolecular structures as target as well as drug.

DRUG DISCOVERY PROCESS

Target Identification and Validation

High-Throughput Screening

Chemical Libraries, Combichem, Natural Products

Clinical Trials

Compounds

Evaluation

Lead

New Targets

Total
GenomeDruggable
genesOutputOutputOutputOutputAgg000Agg000

Includes biological Space

Small molecule Space

New druggable space ?

 Existing drugs (450 Targets) targets)

HIGH-THROUGHPUT SCREENING

FUNCTIONAL INTEGRATION OF:

✓ BIOLOGY

✓ CHEMISTRY

✓ SCREENING TECHNOLOGY

✓ INFORMATICS

BOTTLENECKS

Hundreds of "Hits" but NO "Leads"

Data mining

Accurate profiling of molecules for further studies.

ALTERNATE STRATEGIES

Rational Design of Chemical Libraries

Molecular Modeling Approach

Virtual Screening

Early ADME & Toxicity Profiling

Smart Drug Discovery platform



A view of Drug Discovery road map illustrating some key multidisciplinary technologies that enable the development of (a) Breakthrough medicines from promising candidates (b) LO & generation processes that are relative to novel ligands.

Tomi K Sawyer, Nature Chemical Biology 2, (12) December 2006.

Molecular Modeling

NMR and X-ray structure determination QSAR/3D QSAR Structure-based drug design Rational drug design

QM, MM methods

Model construction Molecular mechanics Conformational searches Molecular dynamics

Homology modeling

Combinatorial chemistry Chemical similarity Chemical diversity

Bioinformatics Chemoinformatics

What is Molecular Modeling?

Molecular Graphics: Visual representation of molecules & their properties.

Computational Chemistry: Simulation of atomic/molecular properties of compound through computer solvable equations.



 $\Sigma\Sigma$ (b'-b'₀)[V₁cosφ] b'φ $\Sigma\Sigma$ (θ-θ₀) [V₁cosφ]

Statistical Modeling: D-R, QSAR/3-D QSAR Molecular data
 Information Management: Organizational databases retrieval /search & processing of properties of 1000... of compounds.

MM = Computation + Visualization + Statistical modeling + Molecular Data Management

COMPUTATIONAL TOOLS: QM/MM

(A) MOLECULAR MECHANICS (MM) **(B) QUANTUM MECHANICS (QM)**

COMPUTATIONAL TOOLS

Quantum Mechanics (QM)

- Ab-initio and semi-empirical methods
 - Considers electronic effect & electronic structure of the molecule
- Calculates charge distribution and orbital energies
- Can simulate bond breaking and formation
- Upper atom limit of about 100-120 atoms

COMPUTATIONAL TOOLS

Molecular Mechanics (MM)

- Totally empirical technique applicable to both small and macromolecular systems
- a molecule is described as a series of charged points (atoms) linked by springs (bonds)
- The potential energy of molecule is described by a mathematical function called a FORCE FIELD

When Newton meets Schrödinger...

Sir Isaac Newton (1642 - 1727)





Erwin Schrödinger (1887 - 1961)



 $\hat{H}\Psi = \mathcal{E}\Psi$



Mixed Quantum-Classical in a complex environment - QM/MM

Main idea

Partitioning the system into

- chemical active part treated by QM methods
- 2. Interface region
- 3. large **environment** that is modeled by a classical force field


Mixed Quantum-Classical in a complex environment - QM/MM

Main idea

Partitioning the system into

- chemical active part treated by QM methods
- 2. Interface region
- 3. large **environment** that is modeled by a classical force field



Basic modeling Strategies

		Receptor Structure	
		Unknown	Known
Ligand Structure	Unknown	Generate 3D structures, Similarity/dissimilarity Homology modelling HTS, Comb. Chemistry (Build the lock, then find the key)	Active Site Search Receptor Based DD de NOVO design, 3D searching (Build or find the key that fits the lock)
	Known	Indirect DD Ligand-Based DD Analogs Design 2D/3D QSAR & Pharmacophore	Rational Drug Design (Structure-based DD) Molecular Docking (Drug-Receptor interaction)

Computer Aided Drug Design Techniques

- Physicochemical Properties Calculations
 - Partition Coefficient (LogP), Dissociation Constant (pKa) etc.
- Drug Design
 - Ligand Based Drug Design
 - QSARs
 - Pharmacophore Perception
 - Structure Based Drug Design
 - Docking & Scoring
 - de-novo drug design
- Pharmacokinetic Modeling (QSPRs)
 - Absorption, Metabolism, Distribution and Toxicity etc.
- Cheminformatics
 - Database Management
 - Similarity / Diversity Searches

-All techniques joins together to form VIRTUAL SCREENING protocols

Quantitative Structure Activity Relationships (QSAR)

QSARs are the mathematical relationships linking chemical structures with biological activity using physicochemical or any other derived property as an interface.

Biological Activity = f(Physico-chemical properties)

- Mathematical Methods used in QSAR includes various regression and pattern recognition techniques.
- Physicochemical or any other property used for generating QSARs is termed as Descriptors and treated as independent variable.
- Biological property is treated as dependent variable.

QSAR and Drug Design

Compounds + biological activity



WHY QSAR.....?

Chemical Space Issue





Virtual Hit Series, Lead Series Identification, clinical candidate selection **Figure: Stage-by-stage quality assessment to reduce costly late-stage attrition.** (Ref: Nature Review-Drug Discovery vol. 2, May 2003, 369)

Types of QSARs

Two Dimensional QSAR

- Classical Hansh Analysis
- Two dimensional molecular properties

Three Dimensional QSAR

- Three dimensional molecular properties
- Molecular Field Analysis
- Molecular Shape Analysis
- Distance Geometry
- Receptor Surface Analysis

QSAR ASSUMPTIONS

The Effect is produced by model compound and not it's metabolites.

The proposed conformation is the bioactive one.

The binding site is same for all modeled compounds.

The Bioactivity explain the direct interaction of molecule and target.

Pharmacokinetics aspects, solvent effects, diffusion, transport are not under consideration.

QSAR Generation Process

- 1. Selection of training set
- 2. Enter biological activity data
- 3. Generate conformations
- 4. Calculate descriptors
- 5. Selection of statistical method
- 6. Generate a QSAR equation
- 7. Validation of QSAR equation
- 8. Predict for Unknown

Descriptors

- 1. Structural descriptors
- 2. Electronic descriptors
- 3. Quantum Mech. descriptors
- 4. Thermodynamic descriptors
- 5. Shape descriptors
- 6. Spatial descriptors
- 7. Conformational descriptors
- 8. Receptor descriptors

Selection of Descriptors

- 1. What is particularly relevant to the therapeutic target?
- 2. What variation is relevant to the compound series?
- 3. What property data can be readily measured?
- 4. What can be readily calculated?

QSAR EQUATION



Molecular Field Analysis



Activity = 0.947055 - 0.258821(Ele/401) + 0.085612(vdW/392) + 0.122799(Ele/391) - 0.7848(vdW/350)

Comparative Molecular Field Analysis



Alignment of all training set molecules

Electrostatic Field Steric Field



Blue (electropositive) Red (electronegative)



Green (favourable) Yellow (repulsive) regions

COMFA studies on oxazolone derivatives



Steric Electrostatic comparative molecular similarity indices (CoMSIA)

COMSIA studies on imidazole derivatives



q2 =0.761 and r2 = 0.945

3D-QSAR - RECEPTOR SURFACE MODEL

Hypothetical receptor surface model constructed from training set molecule's 3D shape and activity data.

The best model can be derived by optimizing various parameters like atomic partial charges and surface fit.

Descriptors like van der Waals energy, electrostatic energy, and total non-bonded energy can be used to derived series of QSAR equations using G/PLS statistical method



Figure 3. Graphical representation of the receptor surface generated around the training set of molecules showing (a) steric fields and (b) electrostatics fields.



Figure 4. RSA descriptors with their locations in the regions demarcated as a-g around the most active molecule (ID2 in Table 1).

activity = 7.70 - 4.01(TOT/1682) -77.98((TOT/695 + 0.01)) - 1.01(TOT/2357) -2.39((TOT/3471 + 0.09)) (2)

$$n = 28$$
, $r^2 = 0.893$, $q^2 = 0.639$, SDEP = 0.775,
LSE = 0.172, BSr^2 = 0.884, PRESS = 7.06,
SDEC = 0.356, and
residual sum squared error = 0.127

PHARMACOPHORE APPROCH

Pharmacophore:

The Spatial orientation of various functional groups or features in 3D necessary to show biological activity.

Types of Pharmacophore Models

Distance Geometry based Qualitative Common Feature Hypothesis.

Quantitative Predictive Pharmacophores from a training set with known biological activities.

Pharmacophore-based Drug Design

•Examine features of *inactive* small molecules (ligands) and the features of *active* small molecules.

•Generate a hypothesis about what chemical groups on the ligand are necessary for biological function; what chemical groups suppress biological function.

•Generate new ligands which have the same necessary chemical groups in the same 3D locations. ("Mimic" the active groups)





Advantage: Don't need to know the biological target structure

Pharmacophore Generation Process

Five Steps

Training set selection.

Features selection

Conformation Generation

Common feature Alignments

Validation

Considerations/Assumptions

Training Set Molecules should be

- Diverse in structure
- Contain maximum structural information.
- Most potent within series.

Features should be selected on the basis of SAR studies of training set

Each training set molecule should be represented by a set of low energy conformations. Conformations generation technique ensures broad coverage of conformational space.

Align the active conformations of the training set molecules to find the best overlay of the corresponding features. Judge by statistical profile & visual inspection of model.

Pharmacophore Features

HB Acceptor & HB Donor
Hydrophobic
Hydrophobic aliphatic
Hydrophobic aromatic
Positive charge/Pos. Ionizable
Negative charge/Neg. Ionizable
Ring Aromatic

Each feature consists of four parts:

Chemical function
 Location and orientation in 3D space
 Tolerance in location
 Weight

Pharmacophore Generation



Training set





Forasartan



Zolasartan



Tosartan

Losartan-Merck

Candesartan-Astra

Tisartan

SAR för AT1 antagonists

R2

В

R1 can be wide variety of groups or ring systems often containing lipophilic and hydrogen bond acceptor moieties Most common R1 should be in m- or p-position Selectivity vs AT1/AT2 can be influenced here. Early compounds included acidic residue, but not needed for high activity. Seems to be a lot of space in the receptor here, large groups tolerated. Even completely aliphatic moieties are accepted

> The two aromatic rings can be separated by 1-2 atoms, e.g. C(O), C(O)NH, CH2O, CH2NH, NHC(O)NH etc although the single bond seems optimal for good potency and DMPK properties

R1

А

R2 = acidic function important (COOH, tetrazoles, sulphonamides, acylsulphonamides, triflamide)

Ortho-position important / for twisting the biphenyl out of the plane

The biphenyl system seems optimal although heterocylic replacements also seem to work (e.g. pyridine or thiophene in A, thiophene in B) Fused bicyclic systems (e.g. benzofuran, benzothiophene, indole) have also been shown active. Also rigidified systems have shown high potency, but lacking in adequate in vivo properties

Pharmacophore Hypothesis Mapped on Active Molecule



Receptor-based Drug Design

- •Examine the 3D structure of the biological target (an X-ray/ NMR structure.
- •Hopefully one where the target is complexed with a small molecule ligand (Co-crystallized)
- •Look for specific chemical groups that could be part of an attractive interaction between the target protein and the ligand.
- •Design a new ligands that will have sites of complementary interactions with the biological target.



Advantage: Visualization allows direct design of molecules

Docking Process

Put a compound in the approximate area where binding occurs

Docking algorithm encodes orientation of compound and conformations.

Optimize binding to protein

- Minimize energy
- Hydrogen bonding
- Hydrophobic interactions

Scoring

"Docking" compounds into proteins computationally



De Novo Drug Design

Build compounds that are complementary to a target binding site on a protein via "random" combination of small molecular fragments to make complete molecule with better binding profile.





• Can pursue both receptor and pharmacophore-based approaches independently

• If the binding mode of the ligand and target is known, information from each approach can be used to help the other

Ideally, identify a structural model that explains the biological activities of the known small molecules on the basis of their interactions with the 3D structure of the target protein.



Typical projects are not purely receptor or pharmacophore-based; they use combination of information, hopefully synergistically



Cheminformatics - Data Management

 Need to be able to store chemical structure and biological data for millions of data points
 Computational representation of 2D structure

Need to be able to organize thousands of active compounds into meaningful groups

- Group similar structures together and relate to activity

Need to learn as much information as possible from the data (data mining)

Apply statistical methods to the structures and related information

Chemical Library Issues

Which R-groups to choose

Which libraries to make
– "Fill out" existing compound collection?
– Targeted to a particular protein?
– As many compounds as possible?

Computational profiling of libraries can help
 – "Virtual libraries" can be assessed on computer
VIRTUAL SCREENING PROTOCOL

Objective - To search chemical compounds similar to active structure.

Essential components of protocol are as follows

- Substructure Hypothesis
- Pharmacophore Hypothesis
- Shape Similarity Hypothesis
- Electronic Similarity Hypothesis

- VIRTUAL SCREENING

Library of ~ 2 lac compounds was screened

Initially 800 compounds were short listed applying above filters.

• Further 30 compounds were selected by applying diversity & similarity analysis.

Compounds have been in vitro screened and found various new scaffolds

Virtual Screening

- Build a computational model of activity for a particular target
- Use model to score compounds from "virtual" or real libraries
- Use scores to decide which to make and pass through a real screen

We may want to virtual screen

- All of a company's in-house compounds, to see which to screen first
- A compound collection that could be purchased
- A potential chemistry library, to see if it is worth making, and if so which to make

Virtual Screening



•1970's: no biological target structures known, so all pharmacophore-based approaches.

•1990's: recombinant DNA, cloning, etc. helped the generation of 3D structural data of biological targets.

•Present: plenty of structural data of biological targets, but also improved technology to increase pharmacophorebased projects.







Drug Discovery overview (LI & LO)

•Lead discovery. Identification of a compound that triggers specific biological actions.

•Lead optimization. Properties of the lead are tested with biological assays; new molecules are designed and synthesized to obtain the desired properties



Pharmacokinetic Modeling (QSPRs)



Changing Criteria for Clinical Candidates



Drug Discovery is a Juggling Act Dynamic Process of Co-Optimization



Property Profiling Assays Examples



Drugs Must Survive In Vivo Barriers



Each Stage of Biology Has Property Issues



Consider properties in assay development and data interpretation

Safety

In-Silico ADMET Models

Computational methods can predict compound properties important to ADMET

- Solubility
- Permeability
- Absorption
- Cytochrome p450 metabolism
- Toxicity

Estimates can be made for millions of compounds, helping reduce "attrition" – the failure rate of compounds in late stage

Drug Design Successes (Fruits of QSAR)

Name of the drug discovered

- 1. Erythromycin analogs
- 2. New Sulfonamide dervs.
- 3. Rifampicin dervs.
- 4. Napthoquinones
- 5. Mitomycins
- 6. Pyridine –2-methanol's
- 7. Cyclopropalamines
- 8. β-Carbolines
- 9. Phenyl oxazolidines10.Hydantoin dervs.11.Quinolones

Biol. Activity Antibacterial Antibacterial Anti-T.B. Antimalerials Antileukemia **Spasmolytics** MAO inhibitors **MAO** Inhibitors Radioprotectives Anti CNS-tumors Antibacterial

Drug Design Successes

While we are still waiting for a drug *totally* designed from scratch, many drugs have been developed with major contributions from computational methods



norfloxacin (1983) antibiotic first of the 6-fluoroquinolones QSAR studies



donepezil (1996) Alzheimer's treatment acetylcholinesterase inhibitor shape analysis and docking studies



dorzolamide [Trusopt] (1994) glaucoma treatment carbonic anhydrase inhibitor SBLD and *ab initio* calcs



losartan [Cozaar] (1995) angiotensin II antagonist anti-hypertensive Modeling Angiotensin II octapeptide



zolmatriptan [Zomig] 1995 5-HT_{1D} agonist migraine treatment Molecular modeling

Drug Design Successes-2

HIV-1 protease inhibitors



SUMMARY

Drug Discovery is a multidisciplinary, complex, costly and intellect intensive process.

Modern drug design techniques can make drug discovery process more fruitful & rational.

Knowledge management and technique specific expertise can save time & cost, which is a paramount need of the hour.

CADD Facility at TORRENT HARDWARE: SGI Indigo2, O2 OCTANE

SOFTWARE: Cerius2 (Version 4.10) Catalyst (Version 4.10) Daylight (ClogP ver. 4.1) ACD/Lab (pKa & logD suite) TOPKAT (6.2)

DATABASES: ACD, NCI, MayBridge, MiniBiobyte,

CAPScreening

UNIX, LINUX & (Oracle support)



THANK YOU