BSc Pharmacology and Translational Medicine

Module: Principles of Pharmacodynamics and Pharmacokinetics
Lecture: Molecular Modelling in Drug Development
Date: Tuesday 30th October 2012

Toby Athersuch

Lecture Outline

Learning Objectives

By the end of the lecture and the associated learning activities the student should:

- Be able to describe, in general terms, the traditional paradigm of industry-based drug discovery

- Be able to name and describe at least six different methodological approaches used in drug discovery and development, and for each indicate the main elements of the approach, along with strengths and weaknesses.

- Be able to give an example, and describe the process used in a case study provided (p-glycoprotein inhibitor design)

Lecture Outline

Lecture Outline

- Introduction
- The Drug Development Process
- The Cost of the Drug Development Process
- Structure-Based Drug Development
- Generating “Hits”
- Drug Properties
- Approaches for Structure-Based Drug Development
    - Homology Modelling
    - Molecular Mechanics
    - Protein Folding
    - Docking
    - Pharmacophore Modelling
    - QSAR
- A Recent Example
    - P-Glycoprotein

Structure-Based Drug Design

Structure-Based Drug Design

Introduction

- Structure-based drug design (a.k.a. rational drug design)
  - Preferred method of drug design
  - Use knowledge of target structure to rationally develop a drug (e.g. inhibitor)

- Main aim is to find a compound with high activity and subsequently make modifications to optimise other characteristics

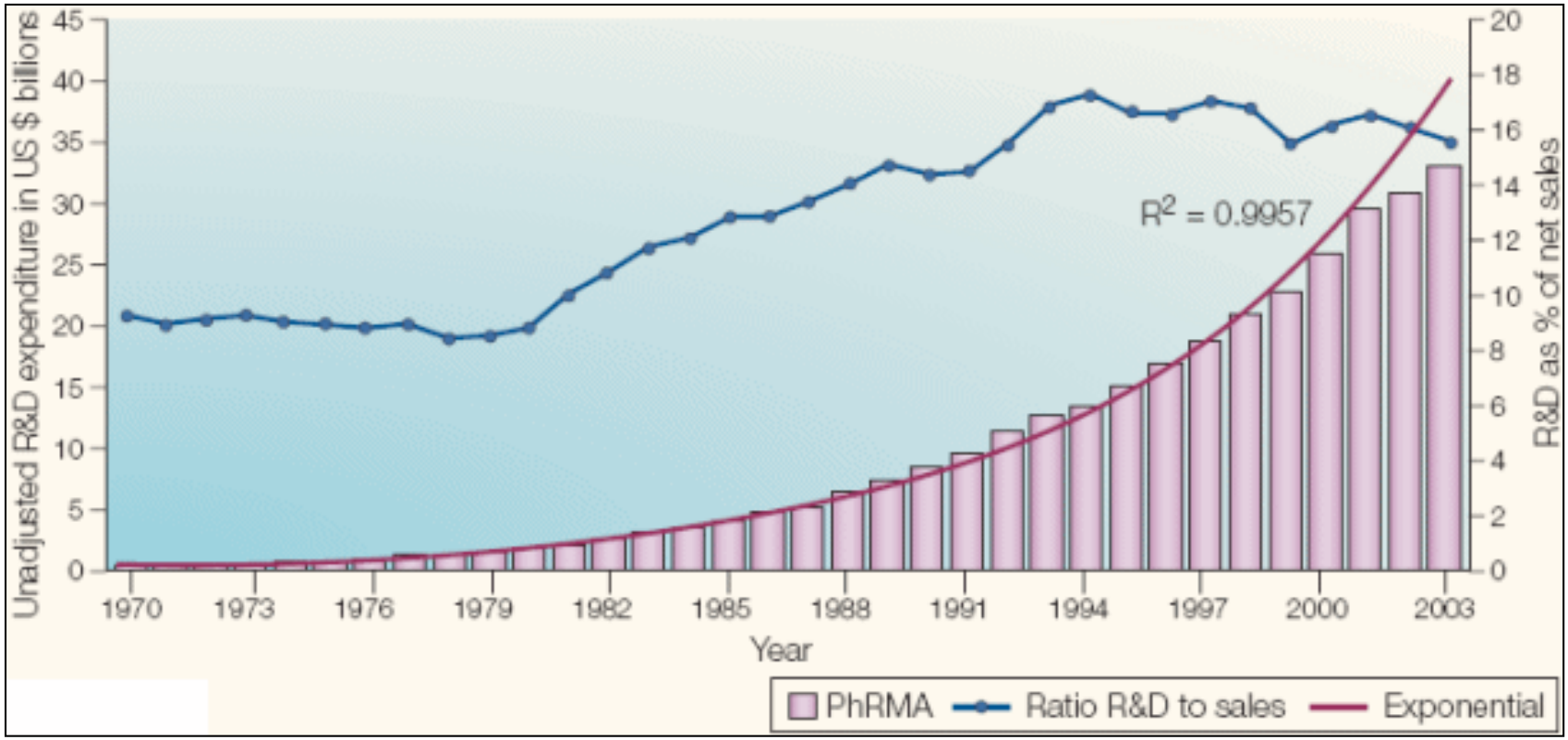
- Need to know about:
  - Target
  - Drug compound
  - Interactions of target and drug compound

Structure-Based Drug Design

Introduction

- The increasing % of income spend on R&D is increasing.

- Maintaining a healthy supply of lead compounds is important for sustained growth in pharma



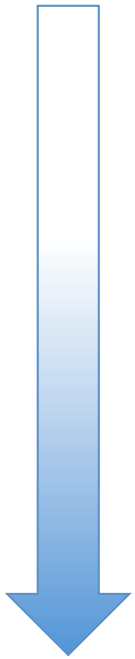
From: Cohen F.J. Macro trends in pharmaceutical innovation, *Nat Rev Drug Disc*. 4, 78-84

Structure-Based Drug Design

The Cost of Drug Development

- Developing a new drug to the point of market sale is very costly
  - £100 million - £1 billion
- Rate output of novel drugs from pharmaceutical industry has fallen dramatically
  - Efficient methods for screening for new leads and their development are desirable

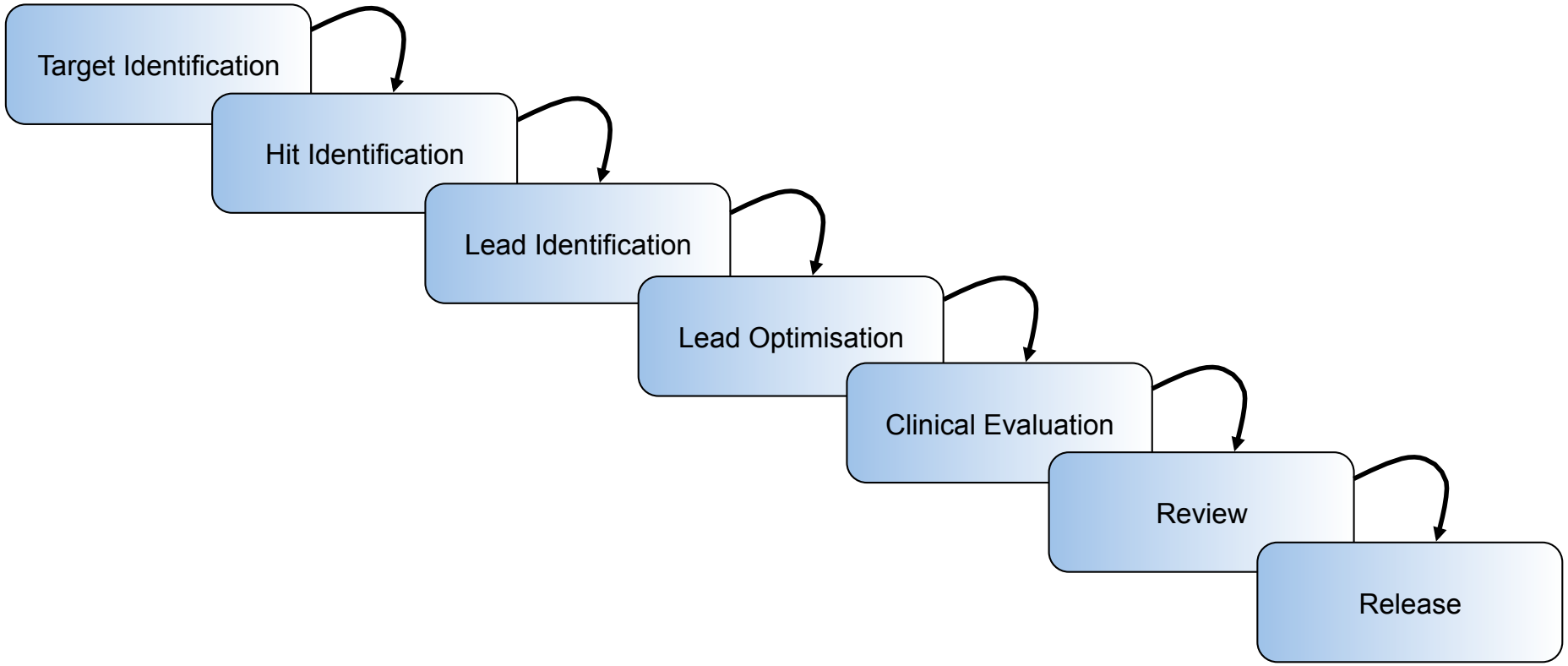
Experiment	Typical Cost / Compound
Computer modelling	10
Biochemical assay	400
Cell culture assay	4000
Rat acute toxicity	12000
Protein crystal structure	100000
Animal efficacy trial	300000
Rat 2-year chronic toxicity	800000
Human clinical trials	500000000



Adapted from: Young D.C. Computational Drug Design. Wiley. 2009.

The Drug Development Process

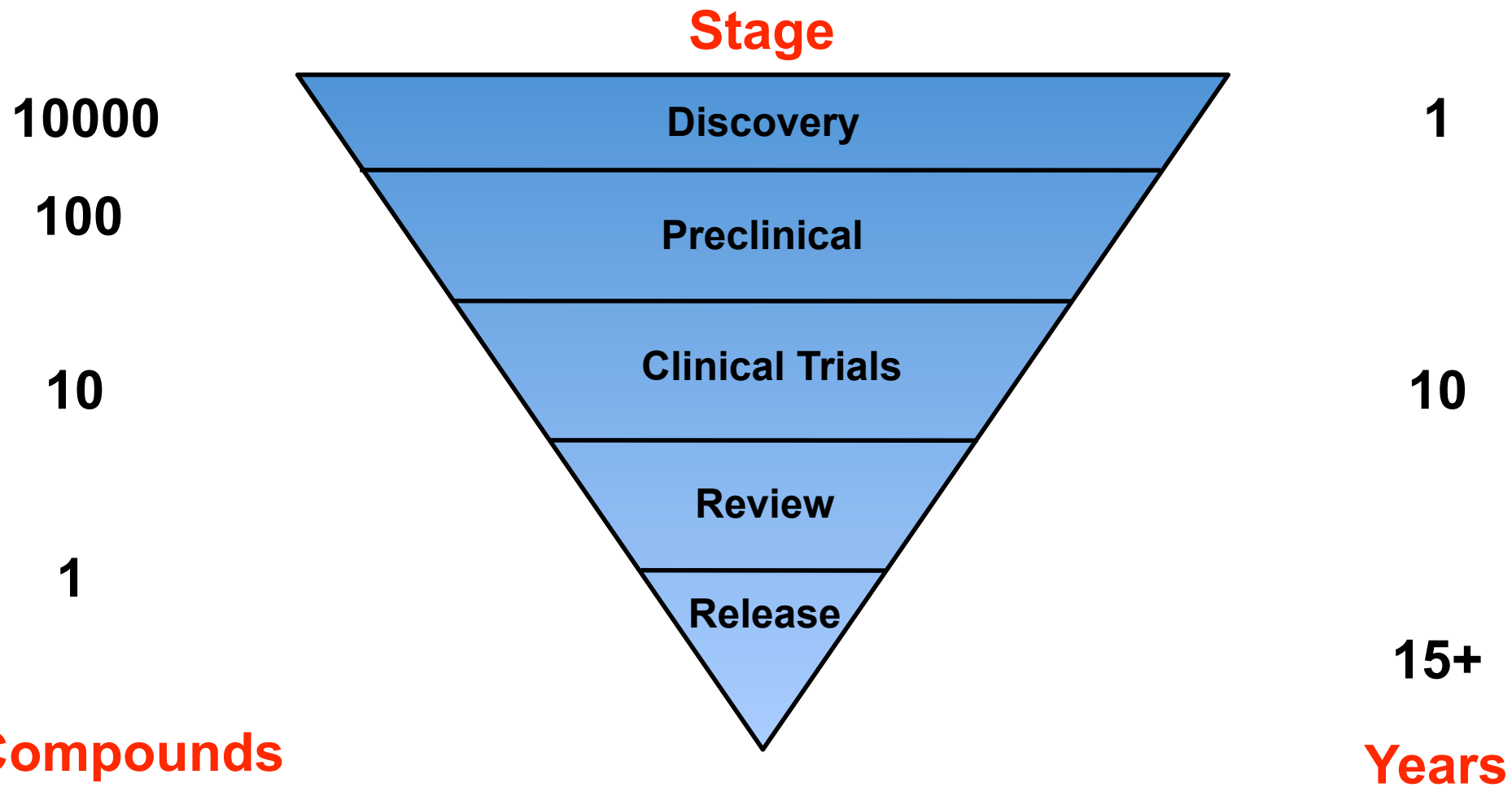
- Typical drug discovery and development pipeline



Structure-Based Drug Design

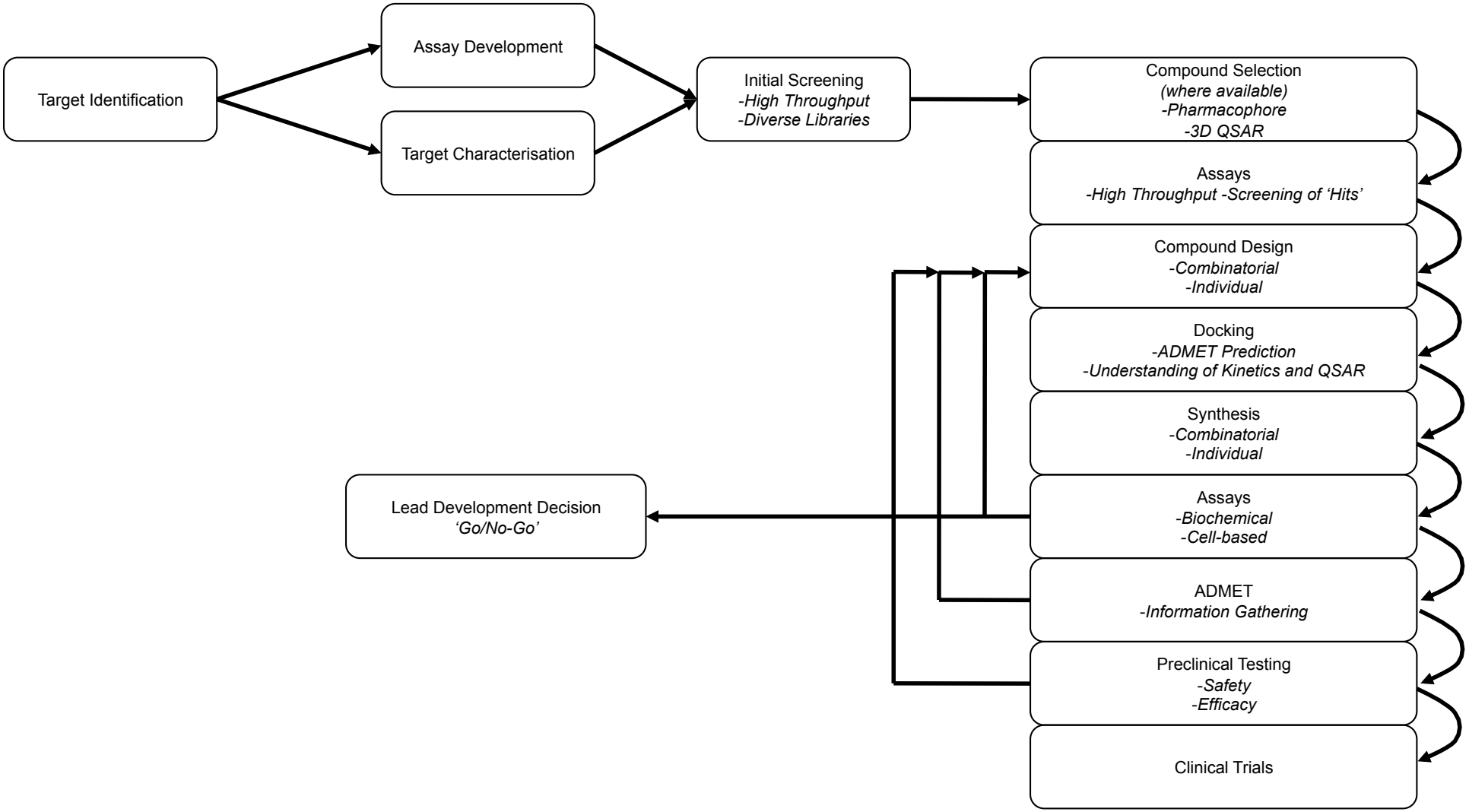
The Drug Development Process

- Typical drug discovery and development timeline



Structure-Based Drug Design

The Structure-Based Drug Development Process



Adapted from: Young D.C. Computational Drug Design. Wiley. 2009.

Structure-Based Drug Design

Generating “Hits”

- In order to find a selection of chemical species with which to use as the starting point for a rational approach to drug design,

- The effect of the individual compounds on the performance of an assay is assessed

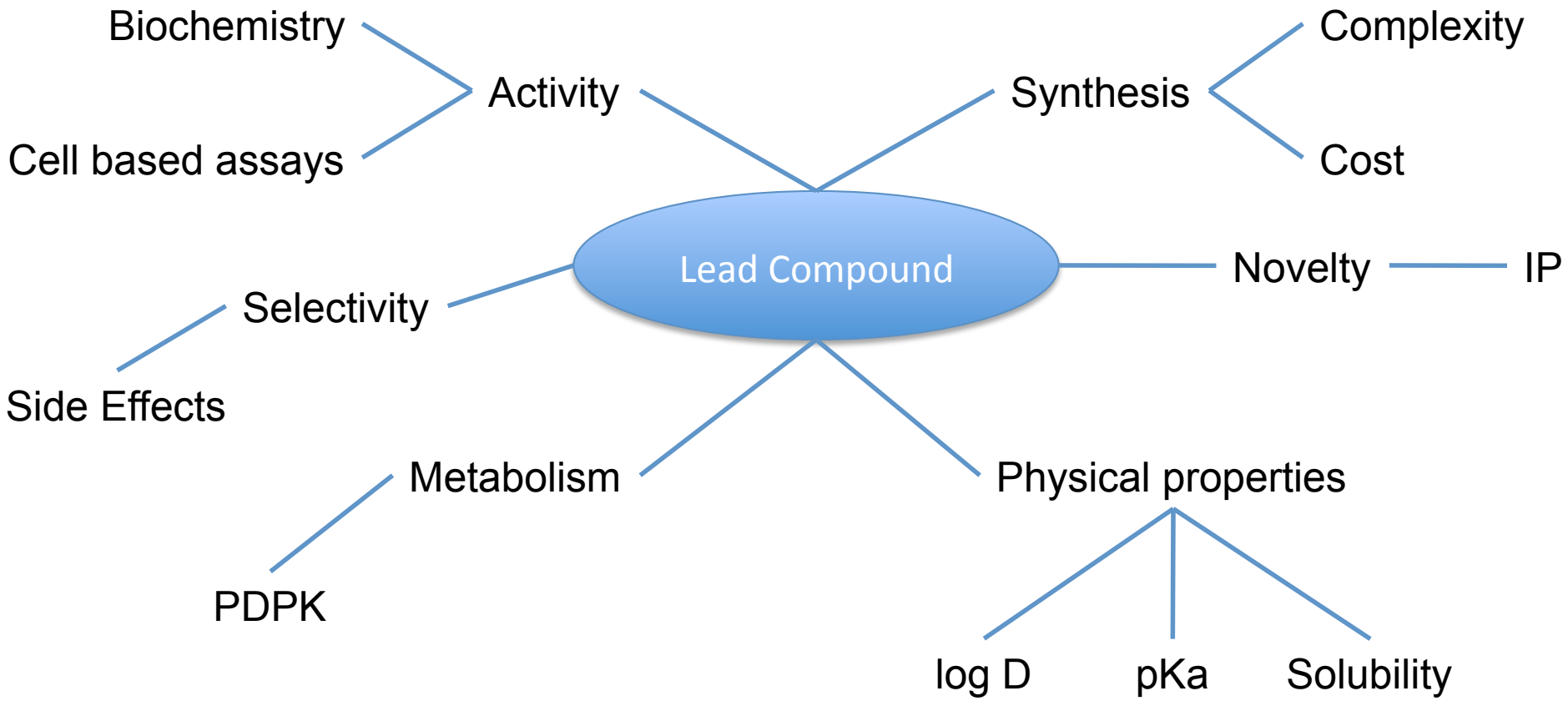
- Selecting the compounds that are used in this initial screening exercise is done in two main ways:
  - *Populating the chemical space*. Conducting the assay for a collection of chemicals that are diverse in their chemistry.
  - *Searching for suitable structures*. Knowledge of the structure of the target can allow searching for chemicals that are likely to fit an interact in a suitable way.

- Once a selection of “hits” are available, they can be considered in turn using computational approaches to suggest modifications that optimise particular characteristics.

Structure-Based Drug Design

Lead Optimisation of Drug Properties

▪In order to rationally understand drug leads and select those for optimisation, a variety of factors are usually considered. Knowledge or accurate prediction of these factors for unknown compounds is useful.



Structure-Based Drug Design

Some Approaches in Structure-Based Drug Design

- The computational chemist has a number of key approaches that can be used in combination to contribute to the drug development process including:
  - Homology Modelling
  - Molecular Mechanics — **Describing the target**
  - Protein Folding
  - Docking
  - Pharmacophore Modelling — **Describing the drug**
  - Quantitative Structure-Activity Relationships (QSAR)
- No one technique is 'best'
  - They do different things
  - Need to use the most appropriate technique for the problem at hand
- Used appropriately, these techniques can help provide information on the structure of the target, and give an idea of the relative activity of various drugs

Homology Modelling

Structure-Based Drug Design

Homology Modelling

- Where no crystal structure exists for the target it is often feasible to use homology modelling

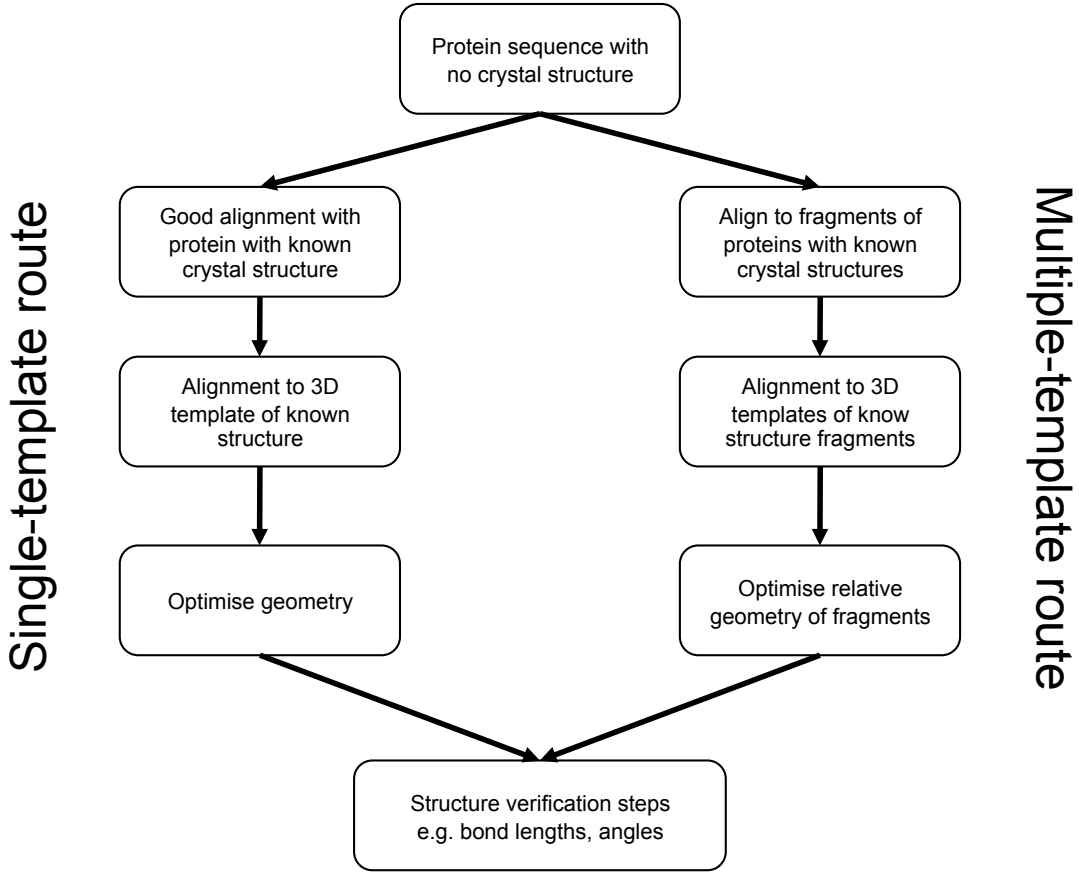
- If the primary amino acid sequence for the target is known:
  - Comparing the primary amino acids sequence with those of proteins that have crystal structures available will allow portions of the overall target structure to be approximated if they are similar
  - Compare the *target* sequence with a *template* sequence

- Quality of the homology model depends on several factors including:
  - The degree of alignment
  - The presence of regions in the target that are not represented in the template
  - Poor resolution of the template structure / poor template selection

- Using fragments of multiple sequences in combination can help improve sequence coverage.

- Flexible loop regions of the structure are particularly difficult to map from one strcuture to another due to their flexible nature

Homology Modelling - Outline



Adapted from: Young D.C. Computational Drug Design. Wiley. 2009.

Molecular Mechanics

Structure-Based Drug Design

Molecular Mechanics

- Identifying the most energetically favourable conformation of a molecule and defining the energy difference between the relative conformations that may exist.

- Calculation of the *conformational energy* associated with the various geometrical arrangement of a target protein. Molecular mechanics calculations for this are computationally inexpensive.

- Calculation of the energy related to:
  - Bond stretching
  - Bond angle bending
  - Torsional angles
  - H-bonding
  - van der Vaals interactions
  - Coulombic attraction/repulsion

- The set of equations that describe these are called *force fields*
  - A variety of force fields have been developed
    - Differences in robustness - some general, some specific
    - Choice of force field depends on type of molecule
      - e.g. proteins – AMBER, MMFF, etc.

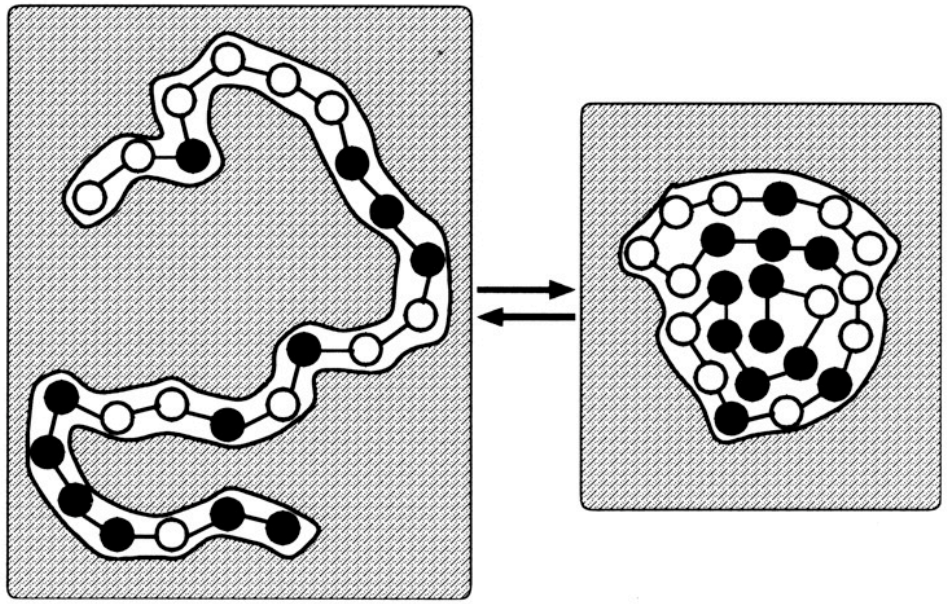
Protein Folding

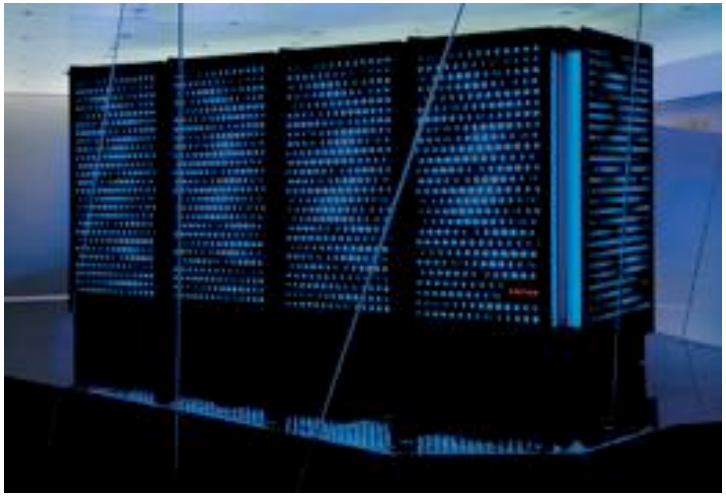
Structure-Based Drug Design

Protein Folding

- Calculation of 3D structure from primary amino acid sequence

- Used when no structural information is available
  - e.g. no crystal structure or NMR





Anton – Supercomputer at D.E. Shaw Research(New York). Current record holder for modeling the folding pattern of a protein over a millisecond (as of 17th October). See Nature doi:10.1038/news.2010.541 and Shaw, D.E. *et al*. Science. 2010. 330, 341-346

Schematic of protein-folding equilibrium. The black and white circles represent **hydrophobic and hydrophilic residues**, respectively. The shaded region depicts aqueous solution.

From: Huang D M , Chandler D 2000. *PNAS*;97, 8324-8327

Structure-Based Drug Design

Protein Folding

- The main problem with predicting the structure of a protein if no other data are available is that there is no reference to check that the protein has been folded correctly
  - Could be completely wrong
  - Protein may be folded in the presence of a chaperone and not be in the lowest possible energy confirmation anyway

- The search space for the lowest energy conformation of a protein is enormous
  - Impossible to try every conformation
  - The protein will normally fold in the presence of other structures apart from chaperones
    - Membrane-bond proteins
    - Solvent

Structure-Based Drug Design

Protein Folding

- A wide variety of algorithms are available for protein folding calculations
  - Usually cover search space my random sampling of different conformers
  - Use iterative searches to identify local minima
  - Use of piecewise methods
  - Use of prior knowledge to optimise particular features
    - α-helices, β-sheets

- Other than comparison to other known (correct structures), there are few ways to validate the structure
  - ‘Reality checks” – are the hydrophilic and hydrophobic residues in appropriate places?

Docking

Structure-Based Drug Design

Docking

- An automated procedure for evaluating the ability of compounds to bind to the active site of a target

- Very commonly used in drug development

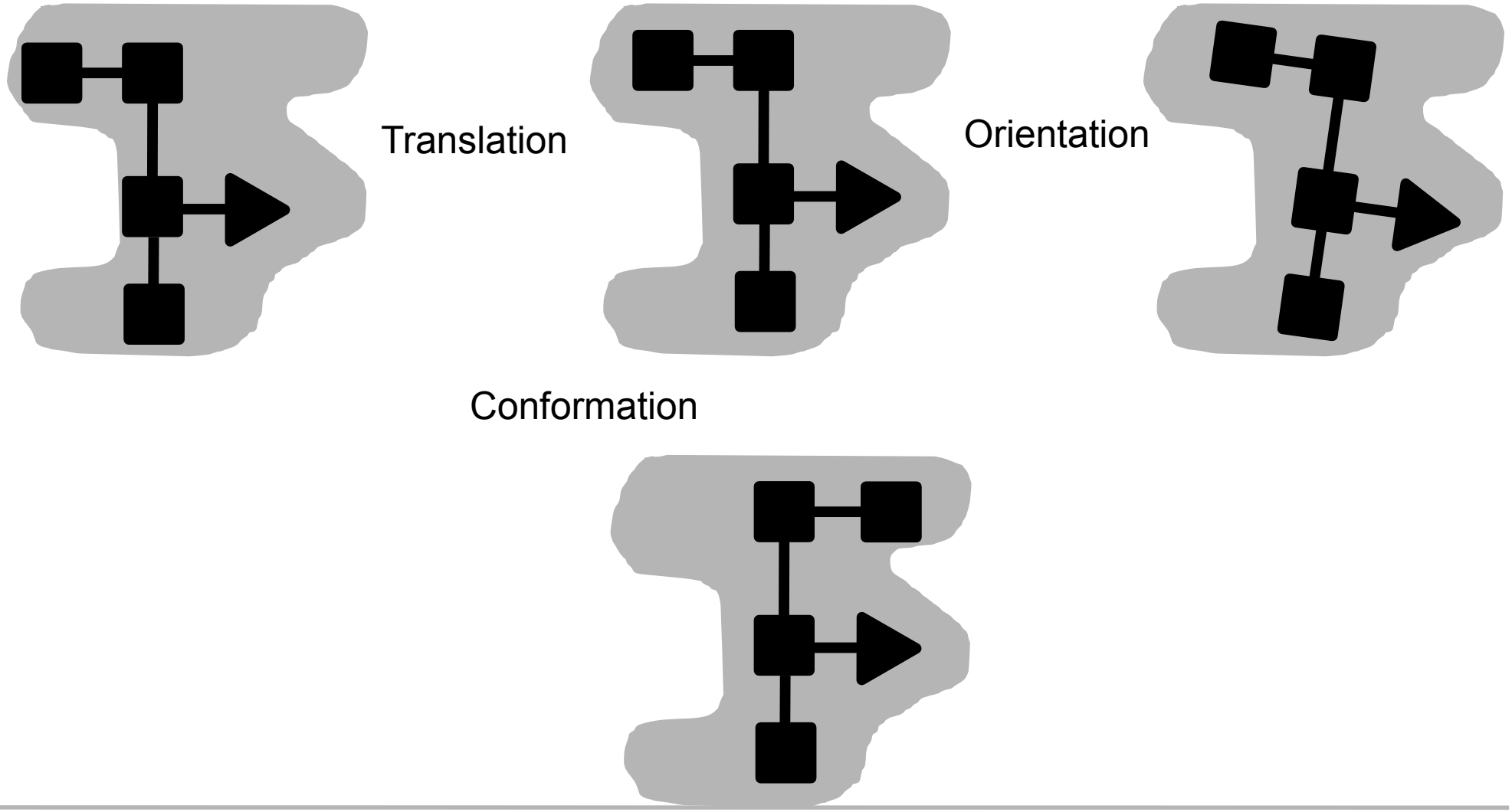
- Simulate the interaction of a drug (ligand) and binding site of the target

- The description of the ligand in the binding site can be described in terms of:
  - *Conformation*
  - *Translation*
  - *Orientation*

- Together, these terms describe the *position* of the ligand in the binding site.

Structure-Based Drug Design

Docking



Structure-Based Drug Design

Docking

- The process of assessing ligands using docking has two main elements:
  - Searching
  - Scoring

- *Searching*. The search space of all positions of the ligand in the binding site
  - Very large
  - Needs to be adequately sampled to identify the correct binding position

- *Scoring*. Positions in the docking procedure are evaluated in terms of the energy of the interaction
  - Computational cost and accuracy depends on the algorithm used
    - Molecular mechanics force field methods vs grid methods
    - Flexibility of binding site / solvation

- Typically, screening using docking procedures are done in a hierarchical manner to reduce overall computational requirements

Pharmacophore Modelling

Structure-Based Drug Design

Pharmacophore Modelling

- A *pharmacophore* is a three-dimensional description of the various properties that a drug molecule will have in order to bind at the active site of the target

- A pharmacophore is done using spatial descriptions of molecular features such as:
  - H-bond donors
  - H-bond acceptors
  - Presence of aromatic rings
  - Charges
  - Functional groups (e.g. acidic / basic)
  - Sterically hindered groups
  - Metals

- The pharmacophore may also describe what is not permitted in this structure
  - e.g. bulky groups / those that would not be accommodated at the binding site

- This approach is particularly useful for finding new classes of molecule as the skeleton of the compound may vary considerably for a given set of constraints defined by the pharmacophore

Structure-Based Drug Design

Pharmacophore Modelling

- There are two main ways that allow the pharmacophore to be designed:
  - *Consensus of features.* In a library of compounds screened, it is possible to identify common features that govern activity based on their presence/absence. Overlay of aligned 3D structures of the active compounds will give information on
  - *Analysis of the active site geometry/chemistry.* If the 3D geometry of the active site is known, it is possible to propose features that an active compound would possess to have a good activity
- It is possible to make predictions about the expected activity of compounds using the match of the pharmacophore criteria.
- The use of pharmacophore is mainly to help shortlist compounds that are sufficiently likely to make good leads.

QSAR

Structure-Based Drug Design

QSAR

- Quantitative structure-activity relationships (QSARs)
  - Use of the strength of association of the molecular properties and activity of compounds

- Molecular properies used in QSAR
  - *'Descriptors'* = Numbers that describe a particular aspect of a molecule (obvious!)

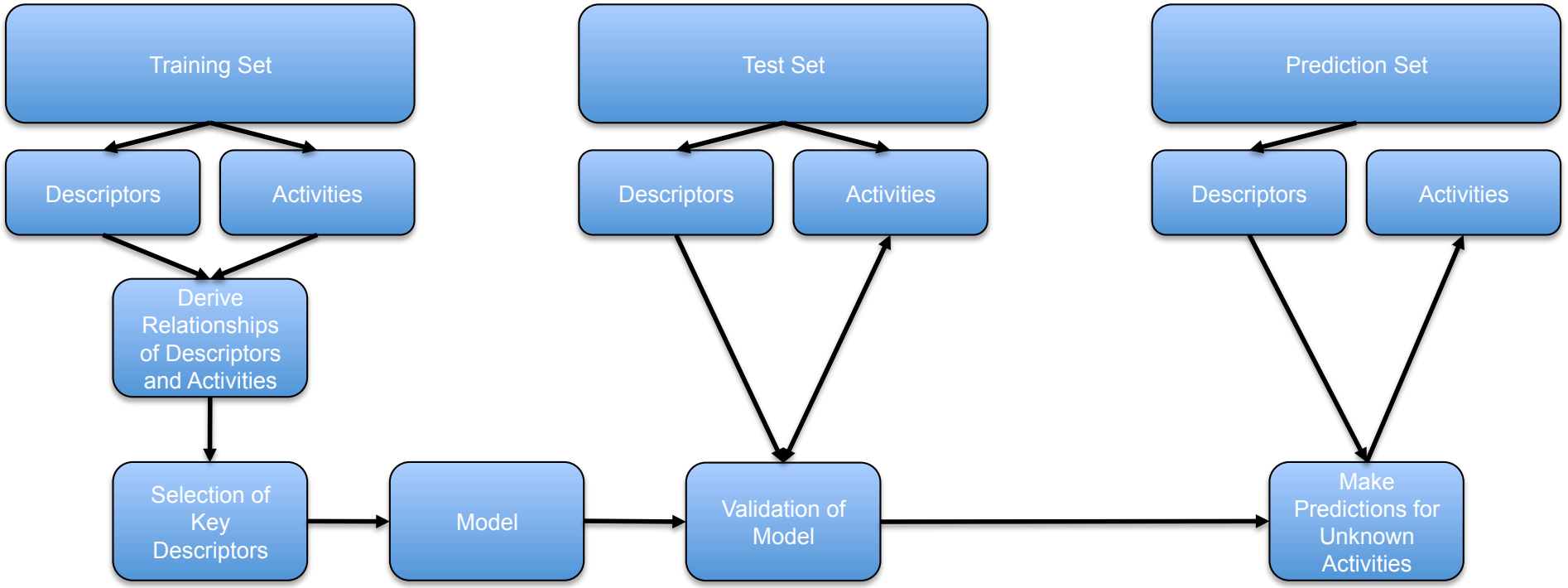
- Types of descriptor
  - *Constitutional*. What is in the molecule and what is it made of?
    - Number of rings, double bonds, etc.
  - *Topological*. What arrangement and connectivity is present in the molecule?
    - Randic / Connectivity Indices, etc.
  - Geometrical. What shape is the molecule?
    - Molecular surface area, Van der Waals volume, etc.
  - *Electrostatic*. How is molecular charge is distributed throughout the molecule?
    - Polar surface area, ClogP (calculated logP of compound), etc.
  - *Quantum mechanical chemical*. What is the electronic structure of the molecule?
    - Ionisation potential, HOMO, LUMO, etc.

Structure-Based Drug Design

Developing a QSAR

- Selection of a series of compounds with known activity and division into:
  - Training set
  - Test/validation set
- Generation of descriptors for a **training set** of compounds with **known activity**
- Assessment of the correlation between each descriptor variable and the activities.
- Identification of those **descriptors that can describe the variation observed** in the activities.
  - Choose variables that are well correlated to the activity, but not strongly with each other
- Generate a **model using these descriptors**
- **Prediction** of the **activities of the test/validation** set using the model
- Analysis of the **actual** *vs* **predicted** values for **the test/validation set**
  - Check for overfitting (model is too specific to the training set to be more generally useful)
- Generation of descriptors for a **prediction set** compounds with **unknown activity**
- Prediction of the activity of **prediction set** compounds
- **Ranking of prediction set** compounds based on **predicted activity**

Developing a QSAR



Structure-Based Drug Design

Summary of Computational Techniques

Approach	Main Use	Advantages	Disadvantages	Cost
Homology Modelling	Description of target when crystal structure unavailable	Efficient No requirement to conduct crystallography	Not as accurate as crystal structure Relies on good quality templates and match	$$
Molecular Mechanics	Identification of likely conformation of molecules based on energetic calculation	Efficient Rapid (c.f. quantum mechanical)		$
Protein Folding	Description of target when crystal structure unavailable	No prior information required	May be completely wrong!	$$$
Docking	Ranking of compounds based on fit to binding site	Efficient screening tool Good match to biochemical assays	Requires 3D structure of target to be known	$-$$$
Pharmacophore Modelling	Description of the main requirements of a ligand to bind a target well	Allows rapid searching for likely structures Not constrained to a particular compound series	3D structure database may not represent conformation of biologically active conformer	$
QSAR	Prediction of properties from structure	Efficient method for determining compounds with a required activity for further evaluation	Not very accurate for predicting activites Overfitting of model	$

Example – P-Glycoprotein

Structure-Based Drug Design

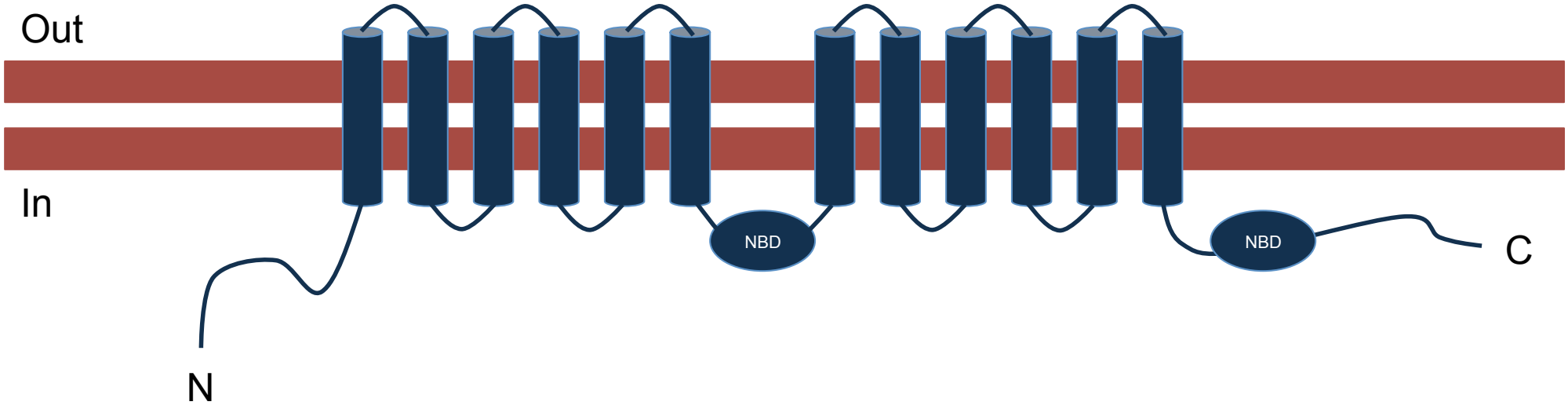
P-Glycoprotein (PgP) – Brief Intro

- ABC transporter
- ATP-driven drug efflux pump
- Encoded by the ABCB1 gene
- Expressed in most tissues
- Far higher expression in:
  - Colon
  - Small intestine
  - Proximal tubules in the kidney
  - Pancreas
  - Bile ducts
  - Blood brain barrier (BBB)

Structure-Based Drug Design

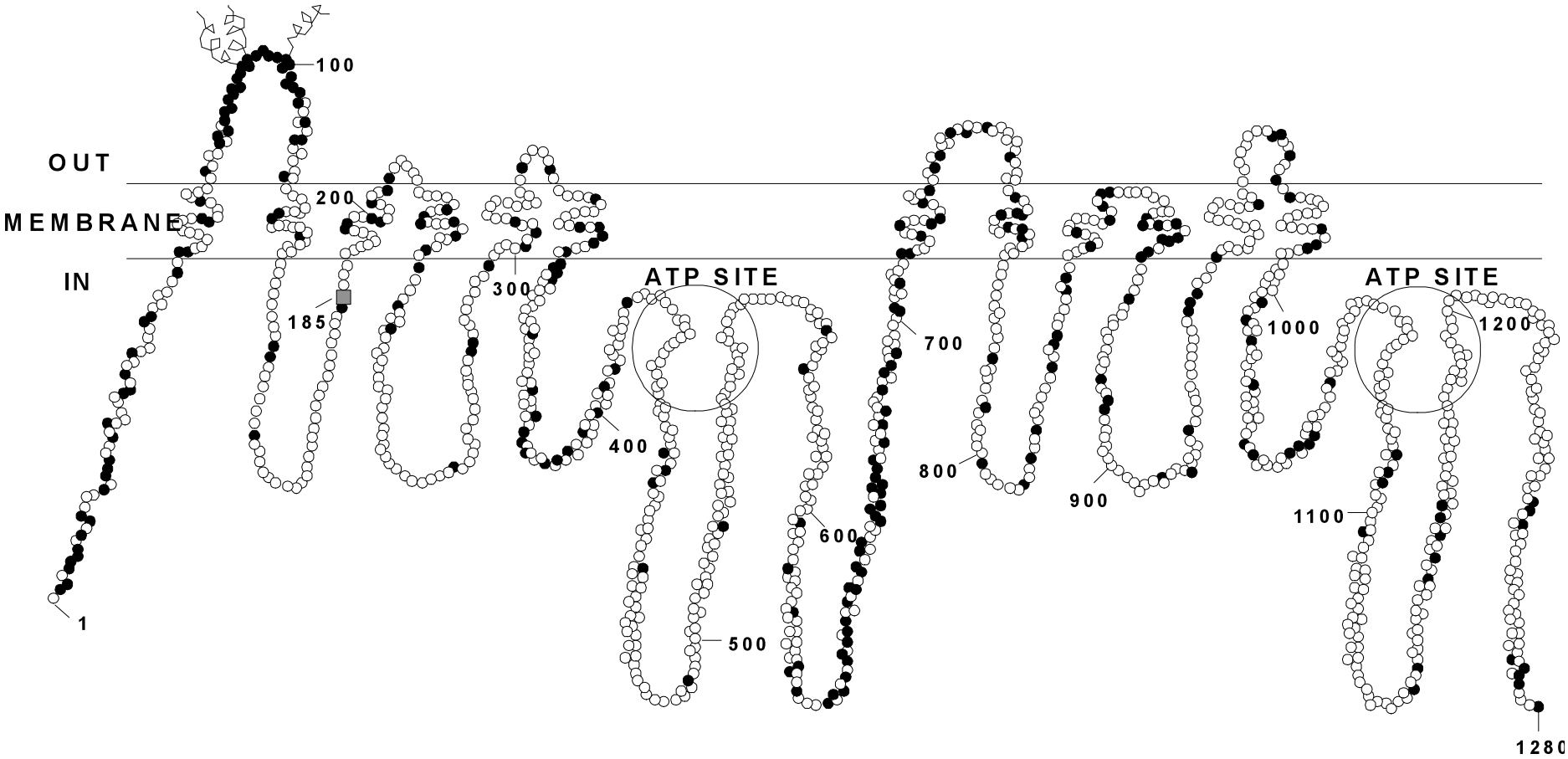
Structure of P-Glycoprotein – Brief Intro

- Transmembrane ABC transport protein
- Size:
  - 1280 amino acid residues
  - MW= 170 kDa
- Total of 12 transmembrane domains
- Contains *N*-terminal glycosylated residues
- Two ATP binding sites
- Well conserved sequences for ATP binding
  - Walker A and Walker B motifs



Structure of P-Glycoprotein – Brief Intro

- Considerable similarity (93%) between mouse and human p-glycoprotein sequence

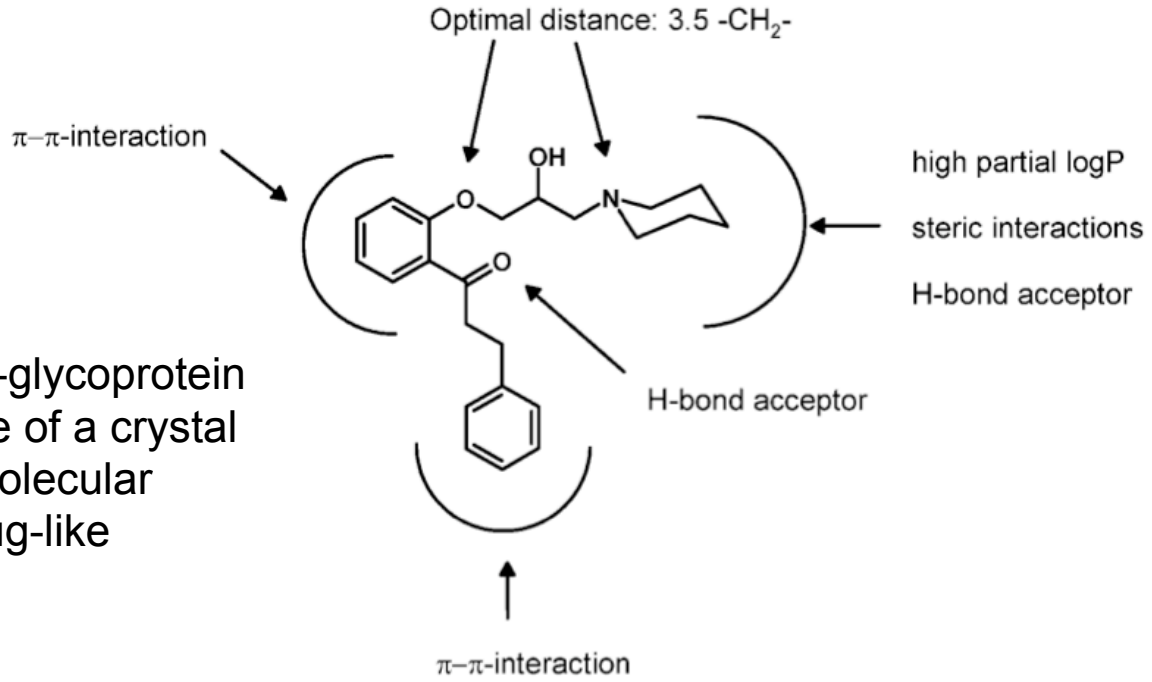


Adapted from Gottesman M.M. and Pastan I.1988. The multidrug transporter: a double-edged sword. *J. Biol. Chem.* 263, 25, 12163-12166.

Structure-Based Drug Design

Drug Design for Multidrug Resistance

- A number of compounds have been shown to be inhibitors of P-glycoprotein
  - Phenothiazines
  - Quinine
  - Tamoxifen
  - Cyclosporin A



- However, designing inhibitors for P-glycoprotein has been difficult due to the absence of a crystal structure and ambiguity about the molecular mechanisms employed to export drug-like substances

- QSAR studies and pharmacophore modelling were used to identify key properties that are required for good inhibition activity

Structure-Based Drug Design

Drug Design for Multidrug Resistance

- Membrane proteins are difficult to crystallise, and are under-represented in databases
- X-ray structures for other organisms allowed human P-glycoprotein to be proposed

Template	Organsim	Sequence Identitiy/ Similarity[a]	Co-crystal[b]	PDB Code	Resolution [Å]	
MsbA	*E. coli*	36%/57%	Apo-open[c]	1JSQ	4.50	retracted
MsbA	*V. cholerae*	33%/55%	Apo-closed[d]	1PF4	3.80	retracted
MsbA	*S. typhimurium*	37%/57%	ADP·V_i	1Z2R	4.20	retracted
Sav1866	*S. aureus*	34%/52%	ADP	2HYD	3.00	
MsbA	*E. coli*	36%/57%	Apo-open	3B5W	5.30	
MsbA	*V. cholerae*	33%/55%	Apo-closed	3B5X	5.50	
MsbA	*S. typhimurium*	37%/57%	AMP-PNP	3B5Y	4.50	
MsbA	*S. typhimurium*	37%/57%	ADP·V_i	3B5Z	4.20	
MsbA	*S. typhimurium*	37%/57%	AMP-PNP	3B60	3.70	
MalK	*E. coli*	31%/50%	Apo-semi open	1Q1B	2.80	
MalK	*E. coli*	31%/50%	Apo-open	1Q1E	2.90	
ABCB1	*M. musculus*	87%/93%	Apo-closed	3G5U	3.80	
ABCB1	*M. musculus*	87%/93%	QZ59-RRR	3G60	4.40	
ABCB1	*M. musculus*	87%/93%	QZ59-SSS	3G61	4.35	

From: Klepsch F. and Ecker G.F. *Mol. Inf.* 2010, 29, 276-286

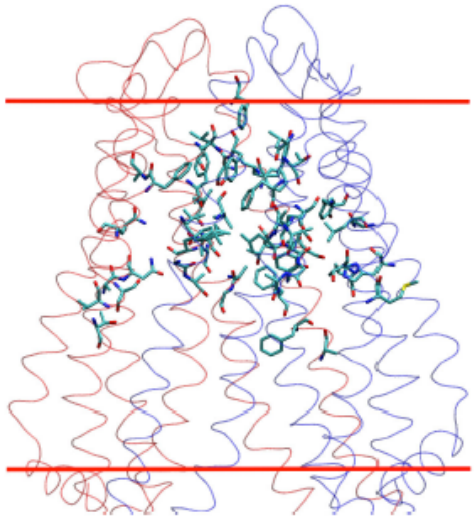
Drug Design for Multidrug Resistance

- Docking studies have been performed using homology models

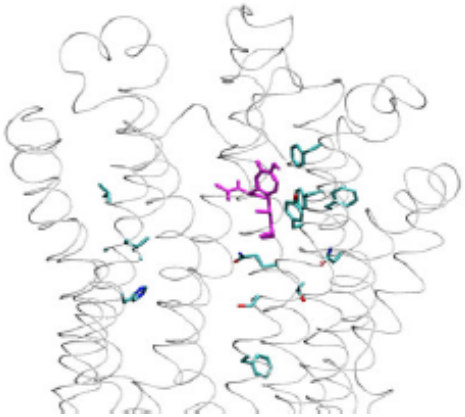
From: Becker J-P. et al. BMC Struct. Biol. 2009, 9:3, doi:10.1186/1472-6807-9-3

Drug Design for Multidrug Resistance

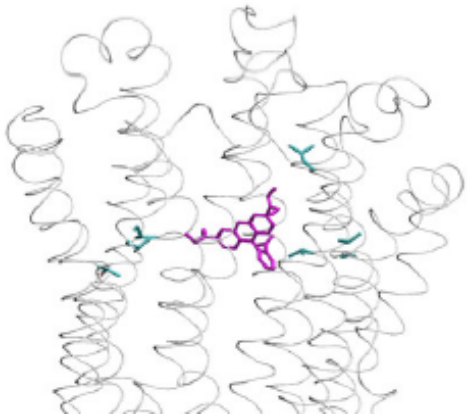
Characterisation of P-glycoprotein active site



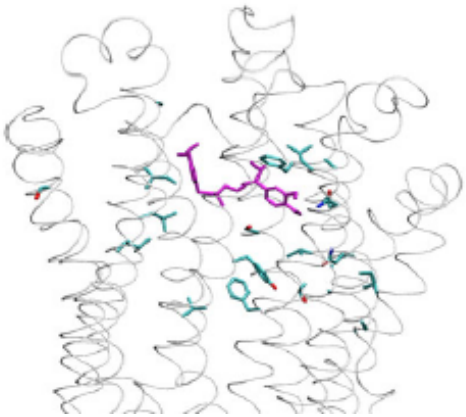
Docking experiments with known P-glycoprotein inhibitors/substrates



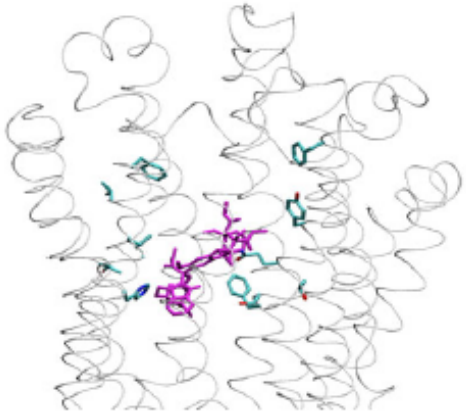
Colchicine



Rhodamine B



Verapamil



Vinblastine

From: Becker J-P. et al. BMC Struct. Biol. 2009, 9:3, doi:10.1186/1472-6807-9-3

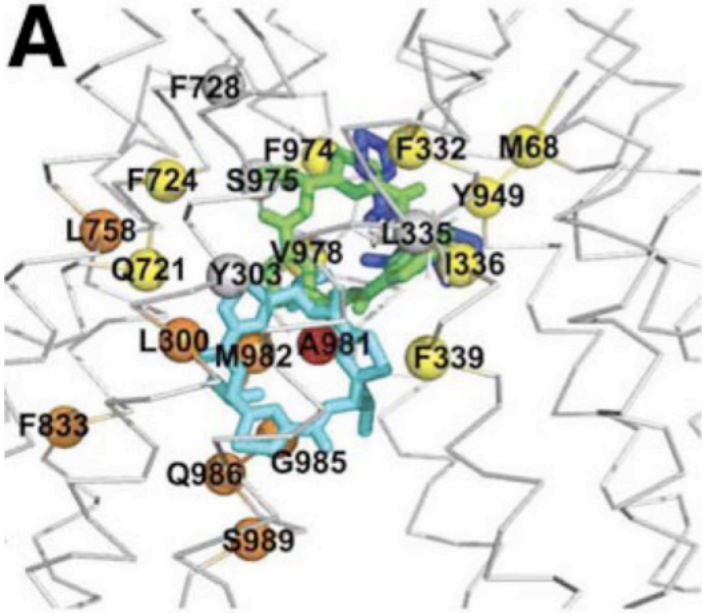
Structure-Based Drug Design

Drug Design for Multidrug Resistance

- The X-ray crystal structure of mouse P-glycoprotein was published by Aller *et al. (*2009).
  - First mammalian ABC transporter to have a known X-ray structure

- Published details of structure
  - Without ligand (apo-)
  - With two different cyclic peptide inhibitors ('soaked')
    - QZ59-RRR
    - QZ59-SSS

two enantiomers

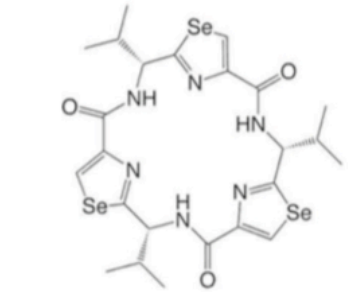


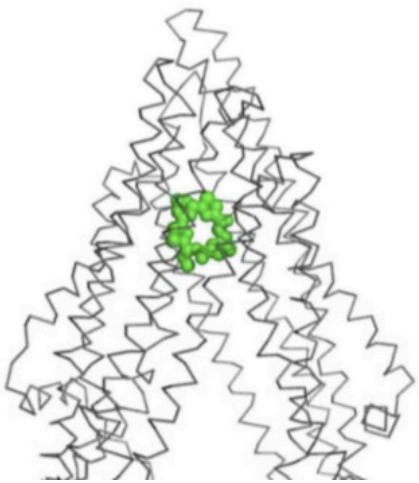
Structure-Based Drug Design

Drug Design for Multidrug Resistance

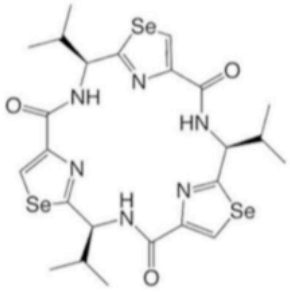
- Knowledge of the crystal structure of mouse P-glycoprotein (high sequence similarity to human) gives insight into the characteristics of the binding site
- Allows comparison with docking study results
- Gives insight into the transport mechanism used by P-glycoprotein
- May enable specific inhibitors to be designed

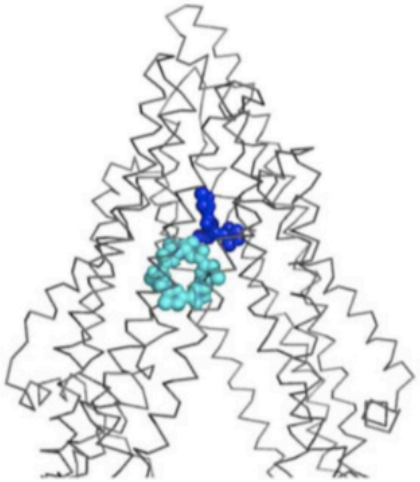
QZ59-RRR





QZ59-SSS





From Aller *et al.* Science. 2009. 3232, 1718-1722

Structure-Based Drug Design

Summary

- There is an urgent need to identify and characterise drug targets and to efficiently generate leads in pharmaceutical context

- Computational approaches play a significant role in characterising drug targets, their interactions and predicting/ranking/optimising the activity of lead compounds

- The approaches are complementary and have different applications depending on the task

- Common approaches are:
  - Homology Modelling
  - Molecular Mechanics
  - Protein Folding
  - Docking
  - Pharmacophore Modelling
  - QSAR

- Appropriate application of computational techniques can significantly reduce development costs