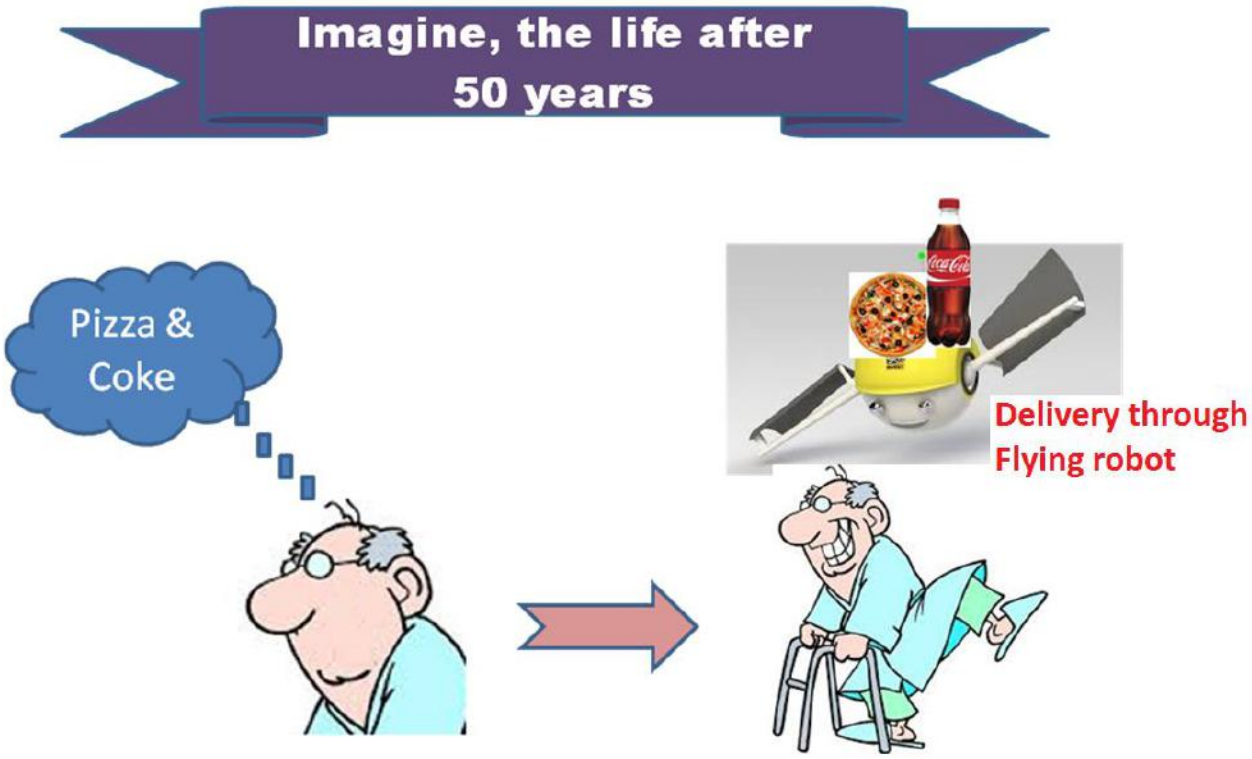
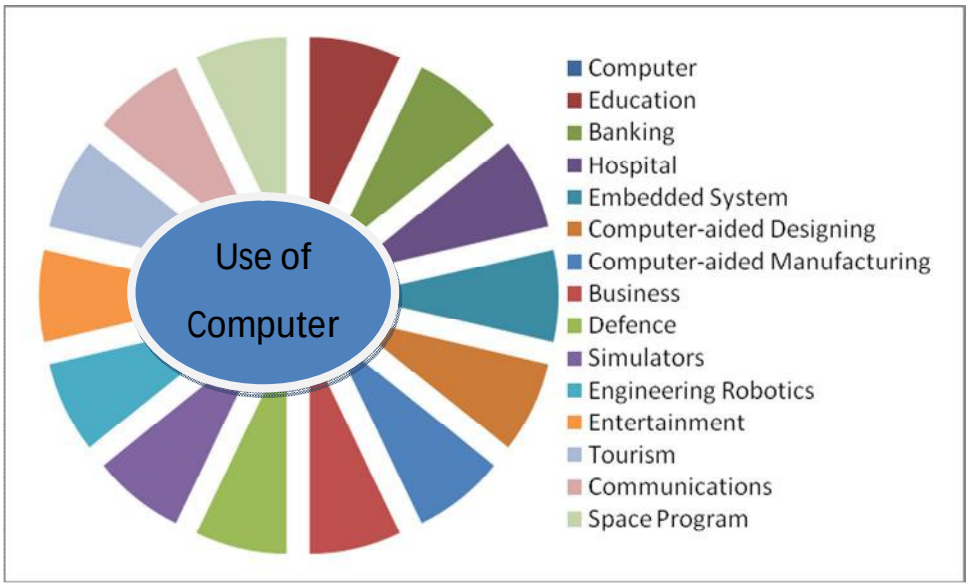
Structure-based drug design: Landscape of modern drug discovery



The routine performance of Computers have resulted their use everywhere.



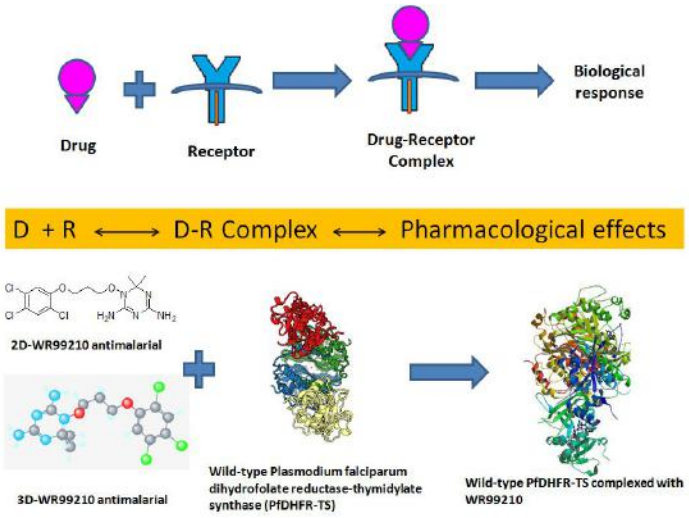
Drug discovery is an intense and lengthy interdisciplinary scientific endeavor and drug development is highly regulated because of legitimate public health concerns. The pharmaceutical industry is under ever-increasing pressure to increase success rate of new drugs. In general, it took over on average 14 years and recently, a new report published by the Tufts Center for the Study of Drug Development suggested that for the development of a new drug associated cost is Rupees 17,174 Crore. Therefore, attention has turned to the possibility of designing candidate molecules for desired properties entirely by computational methods, so that only the compounds that fulfill the required criteria are actually synthesized and tested in the laboratory. To this end, a considerable amount of work has been undertaken to realize fast, high-capacity Structure-based drug design (SBDD) method.

The Road to new drugs

I_____RESEARCH_____I

I_____DEVELOPMENT_____I

The SBDD has become increasingly important in context of drug discovery. In practice, identifying potential leads using SBDD rather than via other approaches is faster, more economical and being easier to setup. As a result, many drugs developed in part by these methods are in late-stage clinical trials or have now reached the market, e.g. tyrosine kinase inhibitors, Canertinib; topoisomerase II inhibitors, Amsacrine; nitrogen mustard pre-prodrug, PR- 104; VEGFR2 inhibitors, DMXAA etc and two examples Viracept (nelfinavir) from Agouron and Agenerase (amprenavir) from Vertex, currently on the market can be traced directly from these methods. The SBDD relies on structural knowledge of the target protein (Receptor) to design and optimize lead compounds.



Most drugs produce their effects by binding to receptors and in Structure-based drug design (SBDD), both 3D structure of receptors and ligand were interacted using appropriate computer algorithms. The poses so generated were evaluated for extent of binding and ranking.

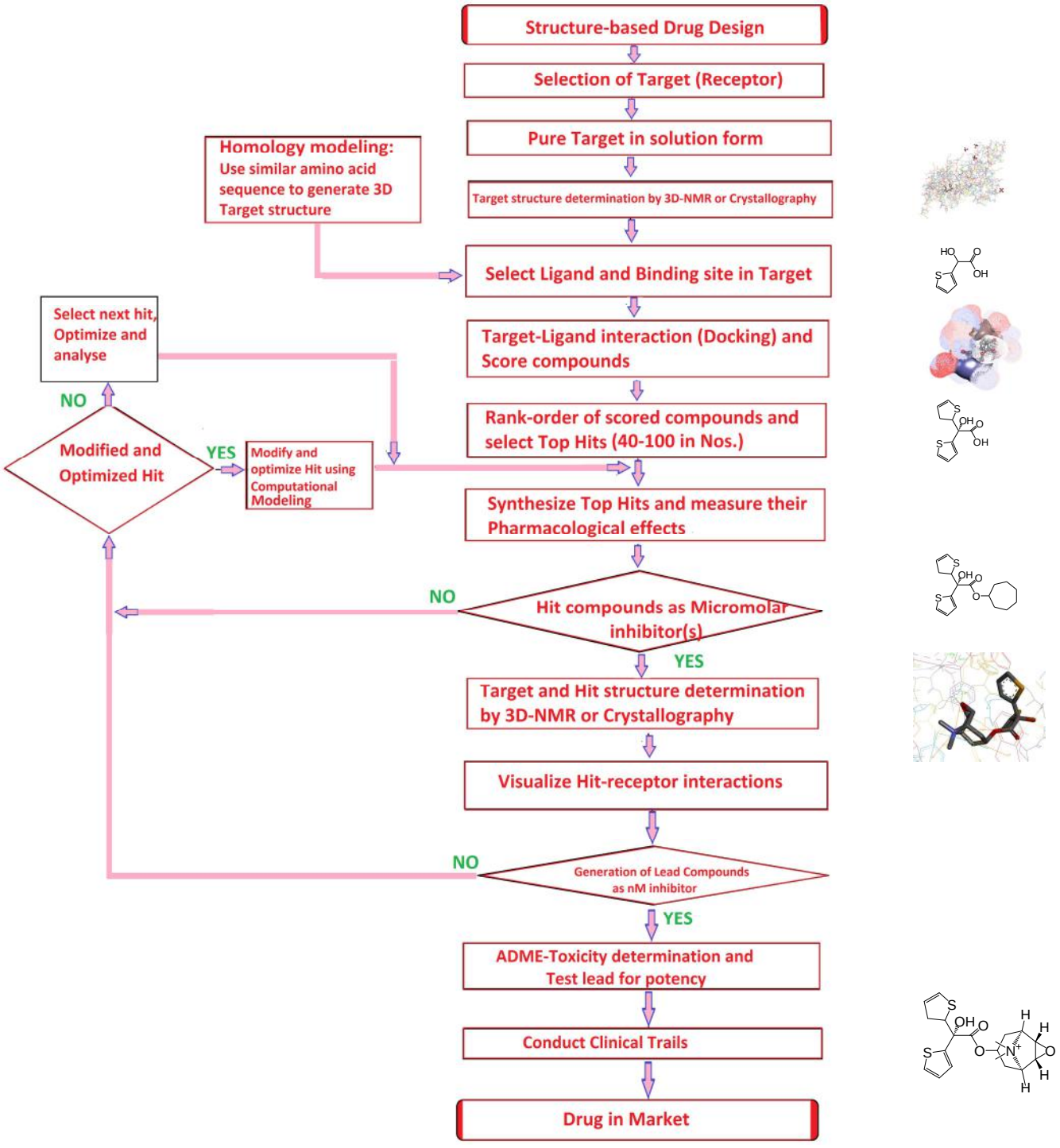


Figure. Process of Structure-Based Drug Design

A background to Structure-based drug design:

A three-dimensional (3D) structure or model of the target is prerequisite in SBDD methods and it encompasses a variety of sequential computational phases, including (i) Target preparation, (ii) Database preparation, (iii) Docking and (iv) Post-docking analysis, and (v) Prioritization of compounds for subsequent testing.

Target preparation

A drug target is a molecular structure in the body which interacts with a drug to produces a clinical effect (treatment or prevention of a disease). The total numbers of drug targets are presently dominated by analyses of the human genome and recent studies indicate that approximately 3,000 targets for small-molecule drugs were predicted to exist by extrapolations. The Medicinal Chemistry explores several proteins and nucleic acids DNA or RNA as drug target(s) that are available in abundance in form of crystal structure from protein data bank. Specifically, the best crystal structures determined with data at a resolution better than 2-5 Å, R-factors below 30%, low coordinate error (the Luzzati coordinate error should be in the range 0.2–0.3 Å), and good stereochemistry. However, these 3D structures build in missing hydrogen(s), residue(s), loop(s) or inappropriate geometry and inclusive of water molecules. Consequently, missing areas are therefore corrected to make computations easier and clear water molecules that would distort the efficiency of SBDD. Overall, the steps inclusive in Target preparation are:

Retrieve crystal structure from protein data bank,

Remove all water molecules and co-crystallized ligand and retained cofactors or Metal.

Add Hydrogens to macromolecular structures.

Assign force-field like MM2, CFF, ECEPP, AMBER, CHARMm, GROMOS etc. to turn all atoms as free variables.

Cleaned 3D conformation was minimized in energy.

Define binding site and radius/ or grid size. Commonly used binding site(s) are (i) Original binding site present in the crystal structure with bound ligand; (ii) the largest cavity present in 3D target structure. Radius 2-5 Å is preferred.

Identity flexible residue(s) and define the degree of flexibility if program supports flexible SBDD.

Database preparation

Several compound libraries for example Pubchem, eMolecules, Zinc, CHEBI, NIST Chemistry WebBook etc. comprises of chemical structures for follow up SBDD. The database offers commercially available compounds directly from vendors and readily synthesizable too. The selected database is subsequently filtered by applying common Lipinski's rule-of-five/ Ghose rule/Bioavailability rule/ or Lead likeness. Additional metabolic stability and toxicity filters may be applied to reduce compounds with poor pharmacokinetics and safety issues.

Docking

A ligand- receptor interaction could be studies using appropriate docking algorithms. The ligand may be a small molecule, peptide or protein. The ligand- receptor interaction within the receptor binding site is usually studied by searching the translational and rotational degrees of freedom of the ligand or the conformational degrees of freedom of the ligand. Binding conformations so generated are called Pose.

Now-a-days, several docking programs available for use and they differ in the sampling algorithms employed, the handling of ligand and protein flexibility, associated scoring functions and screening time minutes per ligand per processor for example High-throughput docking for example Libdock requires less time 3-5 mins. per ligand per processor. Two popular programs (DOCK and Autodock) are available free of charge to academic users and some are commercially available with a cost varying from 2-18 lakhs depends upon the module selected.

Table: Computer Software used in SBDD

Sr. No.	Program	Description
	ICM	Flexible Protein and Flexible Ligand Monte Carlo minimization for protein-ligand docking
	AUTODOCK	Flexible Protein and Flexible Ligand, uses averaged interaction energy grid to account for receptor conformations and simulated annealing for ligand conformations
	FlexX	Non-flexible Protein and Flexible Ligand incremental construction
	FlexE	Flexible Protein and Flexible Ligand, incremental construction; samples ensembles of receptor structures
	SLIDE	Flexible Protein and Flexible Ligand, anchor fragments placed, remainder of ligand added; backbone flexibility
	Flo98	Non-Flexible Protein and Flexible Ligand can rapidly dock a large number of ligand molecules, graphically view results
	ADAM	Non-Flexible Protein and Flexible Ligand fragments aligned based on hydrogen bonding

MCSA-PCR	Flexible Protein and Flexible Ligand, uses simulated annealing to generate conformations of target
Hammerhead	Non-Flexible Protein and Flexible Ligand genetic algorithms to link tail fragments to anchor fragments
MCDOCK	Non-Flexible Protein and Flexible Ligand, Monte Carlo to sample ligand placement
ProDOCK	Flexible Protein and Flexible Ligand, Monte Carlo minimization for flexible ligand, flexible site
DOCK	Non-Flexible Protein and Flexible Ligand docks either small molecules or fragments, includes solvent effects
DockVision	Non-Flexible Protein and Non-Flexible Ligand Monte Carlo minimization
SPROUT	Non-Flexible Protein and Flexible Ligand, generates skeletons that fit site, substitutes atoms into skeleton to give molecule with correct properties
GRID	Non-Flexible Protein and Flexible Ligand, calculates binding energies for functional groups
MCSS	Non-Flexible Protein and Flexible Ligand, exhaustive search of binding site for functional group minima
SMoG	Non-Flexible Protein and Flexible Ligand, knowledge-based scoring function; molecules built by joining rigid fragments
CONCERTS	Non-Flexible Protein and Flexible Ligand, fills active site with molecular fragments, links fragments

Legend	Non-Flexible Protein and Flexible Ligand grows molecule atom by atom
DLD	Non-Flexible Protein and Flexible Ligand saturates binding site with sp^3 carbons, later linked

Post-Docking analysis

Scoring functions are rapid and accurate mathematical method used to predict the strength of the non-covalent ligand-receptor interaction among the generated poses. Three general classes Force field-, Empirical-, and Knowledge-based scoring functions have been developed to model protein–ligand docking. Many popular scoring functions are LUDI, ChemScore, Validate, GOLD score, PLP, FlexX score, ScreenScore, Autodock etc. The Scoring function calculates mostly enthalpic terms, and disregard entropy (binding free energy). Scoring functions only predicts about the bound state of the ligand-receptor, not the unbound states of receptor and ligand. Therefore, scoring functions cannot predict binding affinity or binding free energy.

$\Delta G_{bind} = \Delta H - T\Delta S$

Where ΔH, enthalpy, is a result of the changes in van derWaals and Coulombic interactions as water is removed from bound ligand-receptor surface. Changes in internal energy of the receptor and the ligand have to be considered also. Entropic changes ΔS upon binding are a result of restricting the internal degrees of freedom of the receptor, especially in the binding site, and the ligand, as well as changes in translational and rotational degrees of freedom. In order to judge whether novel compounds are properly ranked, the docking scores of several compounds with known affinity should be compared.

Three criterions are used to evaluate the performance of a scoring function: (i) rmsd measurement between the top ligand conformations and the experimentally observed (native) structure is one of the successful predictions of a native binding mode. The prediction is considered successful, if rmsd is less than 2 Å. (ii) Several alternative methods such as relative displacement error (RDE), interaction-based accuracy classification (IBAC), real space R-

factor(RSR), and Generally Applicable Replacement for rmsd (GARD) are used for a large flexible ligand, unimportant group may hide the correctness in prediction of the overall binding mode leading to the large rmsd value due to a solvent-exposed. (iii) Select potential binders (hits) from a large database of compounds for a given protein target is third criterion for assessing a scoring function.

Prioritization of compounds for subsequent testing.

The prioritization of hits from SBDD screens is a critical step in the drug development process, based on docking scores generated numerous lead candidates ranked. Finally, top hits are synthesized and tested for desired pharmacological activity.

References

Anderson, A. C. (2012) Structure-Based Functional Design of Drugs: From Target to Lead Compound. Method. Mol. Bio. (Clifton, N.J.) 823, 359–366.

Carr, R. and Jhoti, H. (2002) Structure-based screening of low-affinity compounds. Drug Discov. Today 7, 522–527

Dean, N.M. (2001) Functional genomics and target validation approaches using antisense oligonucleotide technology. Curr. Opin. Biotechnol. 12, 622–625

Diercks, T.C.M. et al. (2001) Applications of NMR in drug discovery. Curr. Opin. Chem. Biol. 5, 285–291

Hertzberg, R.P. and Pope, A.J. (2000) High-throughput screening: new technology for the 21st century. Curr. Opin. Chem. Biol. 4, 445–451.

Huang, S.Y. et al. (2010) Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. Phys. Chem. Chem. Phys. 12, 12899–12908.

Li, A.P. (2001) Screening for human ADME/Tox drug properties in drug discovery. Drug Discov. Today 6, 357–366

Mestres, J. and Knegtel, R.M.A. (2000) Similarity versus docking in 3D virtual screening. Perspect. Drug Des. Discovery 20, 191–207.

Myers, S. and Baker, A. (2001) Drug discovery – an operating model for a new era. Nat. Biotechnol. 19, 727–730

Otto, S. et al. (2002) Dynamic combinatorial chemistry. Drug Discov. Today 7, 117–125

Walters, W.P. and Murcko, M.A. (2002) Prediction of 'drug-likeness'.Adv. Drug Deliv. Rev. 54, 255–271