Computer Chemistry in Drug Research: Introduction
Life is based on molecular systems that are organized and regulated by precise recognition and discrimination process. Insights into biological systems show us that all the underlying molecular machinery operates in three-dimension. This is the basis of rational drug design.
In 1890 Emil Fischer proposed a model called the "lock-and-key model" that explained how biological systems function. A substrate fits into the active site of a macromolecule, just like a key fits into a lock whose shape is designed to match the key. Biological 'locks' have unique stereochemical features that are necessary to their function.
Rational drug design proceeds in three steps. It begins with the design of compounds that conform to specific requirements. Then the molecules are synthesized and tested. Further rounds allow the refinement of a given scaffold and its optimization.
Computer Chemistry in Drug Design

• Drug Discovery:
  – Target identification
  – Lead discovery
  – Lead optimization

  Computer chemistry?

• Drug Development
  – Pre-clinical development
  – Clinical development
  – Post-marketing surveillance
Computer Chemistry in Drug Design

• How a lead for a new drug can be discovered?
  – Serendipitous pathway
  – Screening pathway
  – Chemical modification pathway
  – Rational pathway
Rational Drug Design

• Unlike the historical method of drug discovery, by trial-and-error testing of chemical substances on animals, and matching the apparent effects to treatments, **rational drug design** begins with a knowledge of specific chemical responses in the body or target organism, and tailoring combinations of these to fit a treatment profile.
Computer Aided Drug Design

- Drug design is an empirical procedure
- Rational drug design is the rational use of as more experimental data as possible
- Computer chemistry is a tool for rationally extracting all information from data
- Computer chemistry aids (does not substitute) the experimental procedures
- At present drug design without computer chemistry is largely inefficient
Computer Chemistry in Drug Design

Computer chemistry tools for drug design

- Molecular modeling
- Structure-based Drug Design
- Ligand-based Drug Design
- Chemiometrics approaches
Computer Chemistry in Drug Design

• Molecular modeling
  – Model building
    • Interactive molecular graphics
    • Molecule building through sketcher or fragments
  – Conformational minimization
    • Molecular mechanics
    • Quantum mechanics
  – Conformational analysis
    • Sistematic search
    • Montecarlo methods
    • Molecular dynamics
Computer Chemistry in Drug Design

- **Structure (Receptor)-based drug design**
  - Molecular docking

- **Ligand (Pharmacophore)-based drug design**
  - Pharmacophore modeling

- **Chemiometrics approaches**
  - QSAR
  - 3D-QSAR
Molecular modeling

Molecular model → Computer graphics

• atomic coordinates
  – cartesian coordinates (good for computers)
    x, y, z coordinates of atoms
    • PDB, MOL2 files
  – internal coordinates (good for men)
    bond lengths and angles, torsions
    • Z-matrix
Molecular modeling

Molecular graphics
Molecular modeling

Model building (by fragment)
Molecular modeling

Model building (sketcher)
Molecular modeling

• Conformations
  – stable conformation → local minimum
  – preferred conformation → absolute minimum
  – bioactive conformation → ??????

Building by fragment → poor conformation
Sketcher → wrong conformation

It needs some optimization
All chemical systems contain a certain amount of internal energy consisting of potential and kinetic energies. The potential energy is directly related to chemical bonding and non-bonding interactions whereas the kinetic energy is related to random molecular motions. Each geometry (conformation) of a molecule has its specific internal energy; this is due to different non-bonding interactions.
Let's assume we know the bioactive conformation of a specific biomolecule. One of the main goals in the design of a new drug is to conceive a structure for which the conformation that mimics the bioactive form of the reference compound is of low energy and therefore this increases the chance of having an active compound.
Molecular Energies

There are two main computational methods for calculating the energy of a molecule or a set of molecules: quantum mechanics and molecular mechanics methods.
Computer Chemistry in Drug Research: Molecular Mechanics
WHAT IS MOLECULAR MECHANICS (MM)

The mechanical molecular model (or *force field* model) was developed to describe molecular structures and properties in as practical a manner as possible, ignoring the electronic motions and calculating the energy of a system as a function of the nuclear position only.
**WHAT IS A **FORCE** -FIELD**

A *force field* is the whole of the functional form and parameter sets used to describe the potential energy of a system of particles.

- Energies are associated with the deviation of bonds and angles away from their “reference” values.
- There is a function that describes how the energy changes as bonds are rotated.
- Contains terms that describes the interaction between non bonded parts of the system.
...function that describes how the **energy** changes as **bonds** changes..

\[
E(r^N) = \sum_{\text{bonds}} \frac{k_i}{2} \left( l_i - l_{i,0} \right)^2 + \sum_{\text{angles}} \frac{k_i}{2} \left( \theta_{i} - \theta_{i,0} \right)^2 + \sum_{\text{torsions}} \frac{V_n}{2} \left( 1 + \cos(n\omega - \gamma) \right)
\]

\[
+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left[ 4\varepsilon_{ij} \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}
\]
STRETCHING, BENDING, TORSION TERMS

\[ k_b (r - r_o)^2 \]

\[ k_\theta (\theta - \theta_o)^2 \]

\[ A [1 + \cos(n\tau - \phi)] \]
\[ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left( 4\varepsilon_{ij} \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right) + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \]
Empirical force-field methodology is based on classical mechanics and on the assumption that the total steric energy of a structure can be expressed as sum of contributions from many interaction types.
When we reach a minimum? It is rarely possible to identify the real minimum, relative or absolute; we can only hope to find an approximation to the true minimum. A simple strategy to decide when the calculation is sufficiently close to the minimum is to monitor the energy from one iteration to the next, and to stop when the difference in energy between successive steps falls below a specified threshold. So, we can say that the minimization “reach the convergence”.
**Steepest descent:** This is a robust but slowly converging treatment, it is suitable for initial refinement of highly strained structures.

**Conjugate gradients:** This method is more efficient and makes more intelligent choices of search direction. It is suitable for minimization of large systems.

**Newton-Raphson:** This minimizer calculates both the slope of the energy change with conformation and its rate of change. It is not efficient for large systems, however, it can be used for getting extreme convergence on already refined systems.
Molecular dynamics provide an alternative approach to determine the preferred conformers and the global minimum of a molecule. This is achieved by the simulation of the dynamical motions of the molecule as it vibrates and undergoes internal rotation.
Molecular dynamics

MD: a simulation of the particle motion

\[ F = ma \]

- mass
- acceleration
Molecular dynamics

The motion (determined by the temperature) allows conformational changes
Molecular dynamics

• Calculates the time dependent behaviour of a molecular system
• Provides detailed information on the fluctuations and conformational changes of macromolecules
• Routinely used to investigate the structure, dynamics and thermodynamics of biological molecules
• Used in the determination of structures from x-ray and NMR experiments
Molecular dynamics

In a molecular dynamics (MD) simulation it is possible to explore the macroscopic properties of a system.

The connection between microscopic simulation and macroscopic properties is made through statistical mechanics.

Allows to study both thermodynamic properties and time dependent (kinetic) phenomenon.
Molecular dynamics

A MD simulation is practically carried out through the application of the Newton law:

\[ f = m \times a \]

The motion of each particle of the system is calculated from \( a \)

\( a \) is calculated from \( f \)

\( f \) is calculated from the potential \( V \)
Molecular dynamics

A very simplified MD run

1. Give atoms initial positions $r^{(t=0)}$, choose short $\Delta t$

2. Get forces $F = -\nabla V(r^{(i)})$ and $a = F/m$

3. Move atoms: $r^{(i+1)} = r^{(i)} + v^{(i)} \Delta t + \frac{1}{2} a \Delta t^2 + ...$

4. Move time forward: $t = t + \Delta t$

5. Repeat as long as you need
Molecular dynamics

- The potential $V$ can be calculated at different accuracy level (from MM to QM)
- In biology the potential $V$ is generally obtained by a MM force field
- This is a classical treatment allowing the calculation of conformational changes but usually it is not able to reproduce chemical reactions
Molecular dynamics

$\Delta t$ cannot be longer than the fastest atomic motion, therefore:

$$\Delta t = 10^{-15}$$

consequently a simulation of a microsecond needs one billion steps
Molecular dynamics

**Temperature** is directly correlated with kinetic energy:

\[ K = \frac{3}{2} N k_B T \]

Generally a “free” evolution of the system is not allowed. Constraints on temperature and/or pressure are imposed in order to reproduce a particular ensemble.
Molecular dynamics

Environment simulation
The solvent can be simulated in an implicit and in an explicit manner.

Implicit solvent (in most cases the continuum approximation is used): fast calculation but poor results

Explicit solvent (periodic boundary conditions are generally used): accurate results but time consuming