Basics of Molecular modelling

History,
Interdisciplinary classification,
Goals,
Model definition,
Molecular mechanics forcefields,
Energy minimization algorithms
Molecular dynamics simulation
Comparative modelling (homology modelling)
Simulated annealing
Scanning conformational space of a molecule
Molecular properties as recognition patterns

Gerd Krause, Structural bioinformatics and protein design
Leibniz-Institute of molecular Pharmacology
www.fmp-berlin.de/GKrause.html

Molecular dynamic simulations

-enables
characterizing of molecule properties and movement
over a time scale
tasks for molecular dynamic Simulations

- Evaluation of geometric stability larger molecular structures (Homology models)

- **conformational change of molecules at room temperature**
  - since all biological processes are running at 310 Kelvin,
    (but potential energy of force field are calculated at 0 K !!)

- considering of kinetic energy.

- getting over the **problem of multiminima of proteins**
  (Simulated Annealing)

- study of structural and thermal energies
- study of functional properties of fast biological processes($10^{-12} - 10^{-9}$ sec)
Newtons equation is used to simulate atom movement

\[ \text{force} = \text{mass} \times \text{acceleration} \quad (F_i = m_i \, a_i) \]

**force**
- At each atom calculated from energy change between current position and a position in a small distance
- Energy calculation from molecular mechanics

\[ F = m \, a \]

By means of **Atom forces** and **masses** are positions of each atom calculated along a series of extreme short time steps (femto Second=10^{15} \text{ Second})

Resulting snapshots of structural changes: **Trajectory**
Trajectory
Subdivide between two states into large number of sub-states
By extreme small time steps (1 femto Seconds)

Dynamic Trajectory:
- resting at start point a -> after initial movement going with constant total energy (kinetic E. + potential E.),
- initialer Gradient pulls Particle in the valley
- portion of acceleration increases in direction of d (potential Energy decreases)
- kinetic energy moves particle forth and back,
  therefore running through many points at the surface (b, c)

Initiale Atom position at time \( t \) is used predicting
Atom positions at time \( t + \Delta t \)
Atom position at time \( t + \Delta t \) is used predicting
atom position at time \( t + 2 \Delta t \) etc.
Phases of MD

1. **Start structure**: conformation of low energy

2. **Initial movement**: assigned to all atoms, is dictated by chosen temperature

3. **Equilibration phase**: slow heating e.g. from 0K to 310 K
   Energy compensation between the atoms

4. **Data production phase**: Trajectory e.g. at constant temperature.

Randbedingungen der MD

- Statistische Ensembles
- Periodic boundary conditions
- Explicit image model
- Nonbond cutoff

Statistic Ensembles

- \( N \): number Particles
- \( V \): Volumen
- \( T \): Temperature
- \( E \): Energy
- \( P \): pressure

Canonical: \( \text{constant} \text{–NVT} \)
- Temp-Bath coupling
- Vacuo without periodic conditions

Micro-canonical: \( \text{constant} \text{–NVE} \)
- Energy-coupling
- surface const. Energy

Isothermal-isobaric: \( \text{constant} \text{–NPT} \)
- constant pressure and temp
Randbedingungen der MD

Periodic boundary conditions

Docking strategy – constrained experimentally determined contact points (mutants)
MD supported docking of peptide ligand ET1

ET_A RECEPTOR
MD water channel Aquaporin

AQP1

GlnF

de Groot B, Grubmüller H. Science 2001
Molecular modelling 

History,
Interdisciplinary classification,
Goals,
Model definition,
Molecular mechanics forcefields,
Energy minimization algorithms
Molecular dynamics simulation
Comparative modelling (homology modelling)

Application of Protein Structure Models

Depending much on accuracy and reliability of the model:

- Studying catalytic mechanism
- Designing and improving ligands
- Docking of macromolecules
- Prediction of interacting protein partners
- Virtual screening of small ligands
- Defining antibody epitopes
- Designing chimeras, stable and crystallizable variants
- Supporting site-directed mutagenesis
- Refining NMR structures
- Fitting proteins into low-resolution electron density maps
- Functional site identification by 3D Motif searching
- Fold assignment => annotating function & establishing evol. Relationships
- Facilitate prediction of genes e.g. Drosophila

In principle: every application for structures
Comparative modelling

- Target sequence
- Homologous Proteine(s)
- Alignment
- Model building
- Refinement
- Model evaluation
Comparative Modeling

**query:**  
protein,  
domain,  
motif

**homolog**  
protein,  
domain,  
motif

<table>
<thead>
<tr>
<th>Similar/common</th>
<th>Sequence</th>
<th>function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fold</td>
<td>function</td>
</tr>
</tbody>
</table>

Application of Protein Structure Models

**query**  
**homolog**

- Studying catalytic mechanism  
  activation mechanism
- Designing and improve ligands
- Docking of macromolecules  
  Prediction of protein partners
- Virtual screening and docking of small molecules
- Defining antibody epitopes
- Support site directed mutagenesis
- Refining NMR structures
- Structures from sparse experimental restraints
- Functional relationships from structural similarity
- Finding functional sites by 3D–motif searches
### Table 1: Amino Acid Sequence Alignment

<table>
<thead>
<tr>
<th></th>
<th>hTSHR</th>
<th>hLHCGR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TSHR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMH3</td>
<td>TVFAE SLV</td>
<td>TVFAE SLV</td>
</tr>
<tr>
<td>TMH4</td>
<td>AILP</td>
<td>AILP</td>
</tr>
<tr>
<td>ECL2</td>
<td>ISYAKVSICLPMD</td>
<td>VSYMKVSICLPMD</td>
</tr>
<tr>
<td>TMH5</td>
<td>PIALAY YVELTIN</td>
<td>MAPSF YALSAL</td>
</tr>
<tr>
<td>TMH6</td>
<td>TISOY YVELTIN</td>
<td>MAPSF YALSAL</td>
</tr>
<tr>
<td>TMH7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Figure A: Structural Comparison of TSHR and LHCGR

- **TSHR**
  - TMH3: TVFAESLV
  - TMH4: AILP
  - ECL2: ISYAKVSICLPMD
- **LHCGR**
  - TMH3: TVFAESLV
  - TMH4: AILP
  - ECL2: VSYMKVSICLPMD

### Figure B: Quality of Protein Models

- **Stereochemical Accuracy**
  - Torsion angles
  - Mainchain torsion angle distribution
  - Sidechain torsion angle distribution
  - Planarity of peptide bonds
  - β-α angle distribution
  - Chirality of Cα atoms
  - ζ angle distribution
  - Bond lengths
  - Bond angles
  - Planarity
  - Aromatic ring systems and sp2-hybridized end groups

- **Packing Quality**
  - Interatomic distances
  - ‘Bump check’
  - ‘Atomic contact quality’
  - Secondary structural elements
  - Location and geometry of secondary structural elements
  - Hydrophobicity
  - Distribution of polar and nonpolar amino acids
  - Solvent accessible surface of amino acids
  - Unsatisfied buried H-bond donors/acceptors

- **Folding Reliability**
  - 3D-comparison model/template structure
  - RMS deviations between backbone atoms
  - 3D-1D profiles
  - Comparison of environment strings with amino acid sequences
  - Knowledge-based potentials
  - Energy-based comparison

---

**Jaeschke et al.** JBC 2006

---

**G Krause**
Molecular modelling _2
History,
Interdisciplinary classification,
Goals,
Model definition,
Molecular mechanics forcefields,
Energy minimization algorithms
Molecular dynamics simulation
Comparative modelling (homology modelling)
Simulated annealing
Datensammlung (n Konformationen)
Satz von n Konformationen
 eine Konformationsfamilie

n x (Aufheizen + Abkühlung)
Überwindung Energiebarrieren
Temp
Wirkung des Kraftfeldes + experimentelle Restraints
(Standardgeometrien)
(zB. NOE Abstands-Bereich)

NMR spectroscopy
Simulated annealing

Austausch der Magnetisierungszustände führt zu exp. Atomabständen
zB. Nuclear Overhauser Effects (NOE)
(blau gestrichelt)

werden als restraints benutzt
NMR spectroscopy

Konformationsensemble nach Simulated Annealing Coα Atom trace

Protein-protein interaction profiles of PDZ domain and C-terminal peptides of proteins lead to patterns for preferred docking poses of small molecules into binding pocket

Figure 2. a) Surface representation of the AF6 PDZ domain without ligand, 1XZ9. Surface coloring indicates hydrophobic (yellow) and hydrophilic (green) areas. b) AF6 PDZ domain in complex with 5 f. Surface coloring indicates hydrophobic (yellow) and hydrophilic (green) areas. The hydrogen bonding interaction between the PDZ domain and 5 f are shown as yellow-dotted lines. c) Schematic representation of the AF6 PDZ-5 f interaction. Hydrogen bonds are shown as green-dotted lines. Hydrophobic interactions are highlighted by red line fences. Joshi et al. 2006 Angew. Chemie Intern. Ed