Molecular modelling

José R. Valverde

Contents

- *Ab initio* modelling
  - *Ab initio* Quantum Mechanics
  - *Ab initio* Molecular Mechanics
- Homology modelling
  - Homology
  - Threading
- Structure prediction

Ab initio modelling

- Predict 3D structure from chemical formula
  - Out of purely theoretical models
    - Maximum quality (given state of the art)
    - Maximum cost (nothing is assumed)

Ab initio QM

- The best possible approximation
- Tremendous computational cost
  - N^3 – N^8 (on the number of elementary particles)
  - Unfeasible for all but smallest systems
- Modern approaches on the order of N (medium sized systems)
  - Linearly scaling DFT
  - Multipoles and cutoffs
  - MOZYME

Ab initio MM/MD

- Classical mechanics treatment
  - $E = E_{bonded} + E_{non\text{-}bonded}$
  - $E_{non\text{-}bonded}$ scales to N^2
  - Soft charged spheres joined by springs
  - Ignore bond-breaking and formation
- From scratch approaches
  - Tractable only for small-medium size systems
- Large systems
  - Workable if we can start close to the solution

Homology modelling

- Conformational search is a combinatorial problem: N!
  - In the best case we know all factors involved
  - But we don't
- We may assume that similar functions share similar structures and sequences
  - Start from similar structure
  - Assume it is close to the solution
  - Apply an energy minimization step.

Know your problem

- The first step in any simulation is knowing your problem
  - What do you want to know
  - What is already known?
    - Is the solution already known?
    - Is there an approximate solution?
    - What are the characteristics, properties and constrains of your system?
  - Bibliography
  - Database searches

Search for the structure

- Is the structure already known?
  - Do we know the sequence?
    - Blast/FastA against PDB
    - Look for exact matches
      - A mutation or polymorphism means we need to build a model
  - If we ignore the sequence
    - Direct PDB searches
  - If we succeed, we can ignore the rest: we do have the structure

Select template

- Start from a sequence
  - New sequence
  - Old sequence mutated
  - Use FASTA format on a text editor
- Select a modeling method
  - General good similarity (>40%): Homology
  - Twilight zone (30-40%): Threading
  - Local similarities: partial models

Build initial model

- Align sequence against templates
- Assign template structure to unknown
- What next?
  - It may be sensible to stop here
    - If similarity is very high (point mutation)
    - After ensuring no steric clashes (try rotamers)
    - It we know there are no major changes
    - If validation shows no major conflicts
  - Perform additional refinements
    - Minimization (to avoid strong conflicts)
    - MD (to allow for conformational changes)

Model optimization

- In general, more stable structures will have lower internal energy
  - Compute energy
  - Use energy to optimize the structure
  - Molecular Mechanics
    - Force field
  - Various minimization algorithms:
    - Quick and dirty
      - Tinker Minimize, optimize, sniffer. AMBER server
    - Slow and accurate
      - Tinker PSS, newton.

Further optimization

- Validate your structure
- Check against prior knowledge
- Optimization might have found a local minimum
- Look for alternate configurations
  - Simulated Annealing
  - Molecular Dynamics
  - Simulation times are too short (ps-ns)
  - Validate and repeat

Validation

- Experimental methods
  - Prior knowledge
  - Model driven
- Theoretical methods
  - Check a battery of commonly respected rules
  - There may be exceptions
  - Run validation with template and your model
  - Look for differences