Computational Analysis of Small Proteins using the WebAmber Interface

By Eddie Gorr
Gustavus Adolphus College
Research Goals

- Upgrade And Revise Web-based Amber interface
- Use Amber to Explore Interactions of Small Proteins With Environment
- Use Raman Simulator to Simulate Raman Absorptions of Small Proteins
- Use Laser Spectroscopy to Experimentally Obtain Raman Absorptions of Small Proteins
- Compare Computational And Experimental Results
Why Create A Web-AMBER Interface?

- Multiple Complex Molecular Dynamics Calculations
- User friendly environment
- No specialized files or syntax specific text input
- Allows for use in other academic institutions
- Access to computational server through Web interface
What Is Molecular Mechanics

Classical Newtonian Physics

- Set of equations used to calculate the energy of a system as a function of the Nuclear positions

- Assumes Hooke’s Law: $E_{\text{stretch}} = k_s (l - l_0)^2$

- Also based on Relationships Between energy and bond angles, dihedral angles, and distances

- Ignores electrons and the effect of pi systems
Define all energy variables, and minimize. Total Minimum for a system is not necessarily the minimum for all of its parts.
Force Field

Force Field = set of equations that are used to find Energies of Molecular interactions
Assumes Bonds and angles have natural values

\[ E_{pot} = \sum V_s + \sum V_a + \sum V_t + \sum V_v + \sum V_e + \ldots \]

- Bond stretch potential, summed over all bonds
- Bond angle bending, summed over all angles
- Torsions
- van der Waals, summed over all atoms
- Electrostatic interactions

© 1998 J. E. Wampler
Force Fields

Some MM Force Fields

- MM2, MM3 (Allinger)
- MMFF (Merck Pharm.)
- OPLS (Jorgenson)
- BIO+ (Karplus)
- Amber (Kollman)
AMBER
(Assisted Model Building and Energy Refinement)

- Program Developed By the Kollman Group (UCSF)
- Utilizes a force field for molecular mechanics calculations
- AMBER Force Field utilizes Five energy terms summed together to find the overall energy
Computational Costs

- Molecular Mechanics – Quickest Level of calculation (semi-empirical and \textit{ab-initio})

- CPU timescale = the square of the number of atoms (compared to third or fourth power)

- Bert can calculate 3112 atoms in 170,000 sec
  - Computers sharing work load
Complete AMBER Force Field

\[ E_{\text{total}} = \sum_{\text{bonds}} K_r(r - r_{eq})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{eq})^2 + \]

\[ \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \]
Also Use Molecular Dynamics

Molecular Dynamics uses Newton's Equation of Motion

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

\[ a_i = \frac{dv_i}{dt} \quad \quad v_i = \frac{dr_i}{dt} \quad \quad - \frac{dE}{dr_i} = F_i \]

\[- \frac{dE}{dr_i} = m_i \ \frac{d^2r_i}{dt^2} \]

(Molecular Simulations, Inc)
M.D. and Thermodynamics

- Phase Diagrams and Molecular Conformations

[Image of a phase diagram showing various phases of water: solid, liquid, and vapor, with specific regions labeled for different states and transitions.]
Old amber

Amber 7 Calculation

Select Molecule to Upload (mol2 format): 
Browse ...

Import Molecule

Select Previous Calculation Number:
Load Previous mol2 File

View Previous Calculations

Geometry Minimization
Heat to 100 K
Cook at 100 K
Heat to 300 K
Cook at 300 K
Heat to 1000 K
Cook at 1000 K

Geometry Minimization
Heat to 300 K
Cook at 300 K
Heat to 350 K
Cook at 350 K
Heat to 800 K
Cook at 800 K
Heat to 800 K for 1ns
Cool from 800 K to 300 K
Heat to 1000 K
Cook at 1000 K

Created by Bost Magnessa
updated: July 30, 2004
Web-AMBER Interface
Amber’s Front-Page

Amberator Superfantastic Molecular Dynamic Calculator

Upload New Molecule

- GB Solvation
- Water box
- Methanol Box
- Octahedral Box

Select Molecule to Upload: \Egor_Summer05\glucagon.pdb

Format of your molecule: .pdb

Username: egorr

Import Molecule

Utilize Old Calculations

Select Previous Calculation Number:

Load Previous rst File

View Previous Calculations

Download Chime
(Necessary for Visualizing Molecules)
WebAmber

- Formats and creates all AMBER input and control files
- Job submission through PBS (torque/maui) batch queue
- Automated initial analysis of simulation results
- Convenient file creation and download for further investigation
WebAMBER Interface

New features

1. Restraints
2. Cut-off
3. Coupling time constant

Geometry minimization

Amber Calculation

Minimization

Specify Number of steps:

1000

Record Data Every: 10 steps

Add restraints

submit reset

Molecular Dynamics

Amber Calculation

Molecular Dynamics

Time (psec):
10.0

Initial Temperature (K):
0.0

Final Temperature (K):
300.0

Record Data Every: 0.10 microseconds

Velocity Input: No Initial Velocity Values

Non-Bonded Interaction Cutoff (Angstroms): 9

Time Constant for Heat Bath Coupling: 1.0

Add restraints

submit reset

Add restraints dialog

http://pandora.chem.gac.edu/...
How It’s All done
The Miracles of HTML and PERL

• The Web-Interface Passes the necessary information from page to page through scripts like this

• Appends to Generic .in files
<table>
<thead>
<tr>
<th>Calculation</th>
<th>Initial &amp; Final PDB</th>
<th>View Amber Control File</th>
<th>Download:</th>
<th>Download:</th>
<th>Graph trajectory data</th>
</tr>
</thead>
<tbody>
<tr>
<td>gluc2_h2o_md_273_293_50psec</td>
<td>![Image]</td>
<td>![Image]</td>
<td>- Final PDB Structure</td>
<td>- Trajectory (crd)</td>
<td>![Image]</td>
</tr>
<tr>
<td>#218</td>
<td>![Image]</td>
<td>![Image]</td>
<td>- rst file (crd7)</td>
<td>- Regression (crd)</td>
<td>![Image]</td>
</tr>
<tr>
<td>egorr</td>
<td>![Image]</td>
<td>![Image]</td>
<td>- top file (parm7)</td>
<td>- RMSD (first step)</td>
<td>![Image]</td>
</tr>
<tr>
<td>type: md</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
<td>- Radial Distribution</td>
<td>![Image]</td>
</tr>
<tr>
<td>Tuesday, June 21, 2005 at 22:17:17</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
<td>- Diffusion Data</td>
<td>![Image]</td>
</tr>
<tr>
<td>Complete!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**CPU time:** 7782 sec,  
**Salvation:** TRUNCATED

<table>
<thead>
<tr>
<th>Calculation</th>
<th>Initial &amp; Final PDB</th>
<th>View Amber Control File</th>
<th>Download:</th>
<th>Download:</th>
<th>Graph trajectory data</th>
</tr>
</thead>
<tbody>
<tr>
<td>gluc2_h2o_md_273_293_50psec</td>
<td>![Image]</td>
<td>![Image]</td>
<td>- Final PDB Structure</td>
<td>- Trajectory (crd)</td>
<td>![Image]</td>
</tr>
<tr>
<td>#220</td>
<td>![Image]</td>
<td>![Image]</td>
<td>- rst file (crd7)</td>
<td>- Regression (crd)</td>
<td>![Image]</td>
</tr>
<tr>
<td>egorr</td>
<td>![Image]</td>
<td>![Image]</td>
<td>- top file (parm7)</td>
<td>- RMSD (first step)</td>
<td>![Image]</td>
</tr>
<tr>
<td>type: md</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
<td>- Radial Distribution</td>
<td>![Image]</td>
</tr>
<tr>
<td>Complete!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Future of WebAMBER

World Domination?

Distribution prerequisites:
- Server independent code
- Installation script

Further Upgrades
- Time Graph
  - Estimating Time of Calculation
- Communal Calculations
  - MPICH
- More Variable Options
  - Pressure
  - SHAKE
Glucagon 101

Glucose Homeostasis

- 29-residue peptide hormone
- Secreted by α-cells in pancreas
- Stimulates gluconeogenesis and glycogenolysis
- Adopts 3 main conformations depending on environmental conditions\(^1\)
  - High α-helical content (crystal structure)
  - Moderate α-helical content (physiological conditions)
  - Little α-helical content (extreme conditions)

\(^1\) Ying et al. Biochemistry, vol.42, no.10, 2003
Testing Web-AMBER

Initial Calculations

Generalized Born Continuum Solvation

Initial Temperature (K) = 0
Final Temperature (K) = 100
Time (psec) = Variable

Initial Temperature (K) = 0
Final Temperature (K) = Variable
Time (psec) = 50

Images created using Maestro
Analysis on Preliminary Results

Using MAESTRO

Glucagon Shows a high affinity for conformational change near its N-terminus end, the end by which it primarily binds to its G-protein coupled receptor
Subjecting Glucagon to Physiological Conditions

Water Solvation

Experimental Procedure for the Water Solvation of Glucagon

# of waters in truncated octahedral box = 3112
Average Calculation time = 170,000 sec
Three Methods for Preliminary Analysis

Root Mean Squared Deviation Graphs (RMSD)

Ramachandran Plots

Visualization using Maestro
RMSD Values

Equilibration run: R.M.S.D. for 100 Pico seconds after initial M.D. run
Explanation of Ramachandran

phi and psi dihedral angels in a protein backbone determine the secondary structure of the molecule
Ramachandran for Glucagon Trials

Crystal Structure

Physiological Conditions

Extreme Conditions
Annealed Ramachandran

Retains Some Alpha Helical content

Some of Backbone Permanently changed

Unable to Bind to Glp-1 Receptor
Maestro Visualizations

As can be seen, as the temperature increases, the \( \alpha \)-helical content of the molecule decreases.
Conclusions

- Revised WebAMBER interface is in working order
- WebAMBER can be used for the analysis of most molecules, including small proteins.
- Glucagon Adopts 3 main conformations depending on environmental conditions

More Complex Calculations are now being explored!!
Future of Glucagon

Initial Calculations for Glucagon And Methanol

As the Environmental conditions become more hydrophillic/hydrophobic (while keeping physiological temperatures), we will begin to see a more physiological representation of this 29 residue protein.
Goals Revisited
What I have Accomplished

- Upgrade And Revise Web-based Amber interface ✔
- Use Amber to Explore Interactions of Small Proteins With Environment ✔
- Use Raman Simulator to Simulate Raman Absorptions of Small Proteins
- Use Laser Spectroscopy to Experimentally Obtain Raman Absorptions of Small Proteins
- Compare Computational And Experimental Results
Acknowledgements and Sources

Sources:


• Jonathan Smith
• Brent Magnusson
• NSF
• Gover Fund
• Gustavus
Sources


- Wampler, J.E. 1996. TUTORIAL ON PEPTIDE AND PROTEIN STRUCTURE II. SECONDARY STRUCTURE

- NIH guide to molecular modeling.
More Sources