CHEM 8711/7711

Approximate Chemical Modeling Methods

Why Model?

1. Visualize/Develop Structural Models
   - Visualize available crystallographic data
     - Protein Databank (www.rcsb.org)
     - Cambridge Structural Database (depositions since 1994 freely available) (http://www.ccdc.cam.ac.uk/products/csd/)
   - Develop models
     - Small molecules can often be modeled based on prior knowledge (which allows us to calculate the most energetically favorable geometry of related molecules)
     - Large molecules can be modeled with assistance (from NMR, crystallography, or prior related structures)

Why Model?

2. Compute Properties
   - Energies (required for developing models)
   - Dipole moments
   - Conformational populations
   - Binding affinities
   - Protein fold families
   - Reaction rates
   - Reaction stereoselectivities/regioselectivities
   - Electronic properties

Why Model?

3. Develop quantitative models
   - relate aspects of structure to physical properties (QSPR modeling)
   - relate aspects of structure to biological properties (QSAR or QSTR modeling)

How to Model

- Need a theory that relates structure and energy
- That theory should be based on reasonable simplifications
- That theory should be extensible to many different chemical structures

Some Basic Theory - Energetics

Energies can be calculated at several levels of theory

- *Ab Initio*
  - Most theoretically rigorous
  - Energies calculated from electronic structure
  - Requires no experimental parameters
- Semi-Empirical
  - Simplifying assumptions made
  - Experimental parameters compensate
- Molecular Mechanics
  - Electrons essentially ignored
  - Many experimental parameters required

Visualization/Public Databases/Structure Drawing

Structure Drawing

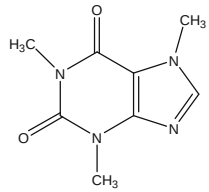
- Structure input is somewhat software-dependent
- Available here: MOE (Molecular Operating Environment), BioMedCAChe (soon) and PC Spartan
- MOE, CAChe and Spartan are available through Novell-delivered applications/CAS/SM110
- Also available in Rm. 006/425: Cerius2, Macromodel, MM3, Gaussian98, Hyperchem, Dalton, GAMESS

Structure Drawing – MOE

- MOE uses a builder window for structure input (accessible through the button on the left side of the MOE window, or from the Windows menu at the top)
- Clicking options in the builder window places isolated fragments, unless a position on an existing fragment is selected by clicking on it in the MOE window

Structure Drawing Exercise

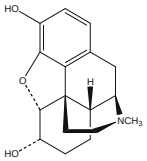
- Input caffeine:



- Save to the My Documents folder as caffeine.moe

Structure Drawing Exercise II

- Input Morphine (with accurate stereochemistry!):



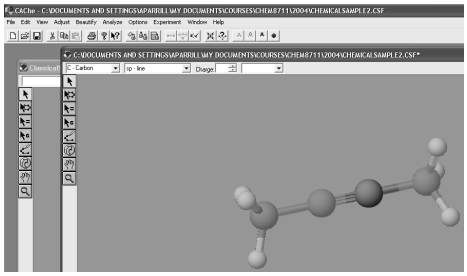
- Save to the My Documents folder as morphine.moe

Structure Drawing – PC Spartan

- Builder initially accessed by File->New
- Click on fragments to add THEN indicate where they should be added (opposite of MOE)
- Compared to MOE:
  - Expert entry allows greater variety of geometries (up to octahedral)
  - Not as many tools for biomolecules

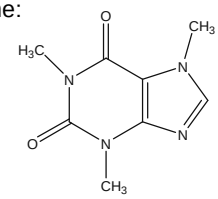
Structure Drawing - CAChe

- Buttons on side of toolbar are used for initial structure input
- Valence is corrected with the 'Beautify' menu items
- Atom types/hybridization/charge can be changed using pull-down menus at the top



Structure Drawing Exercise

- Input caffeine:



- Save to the My Documents folder as caffeine

Molecular File Formats

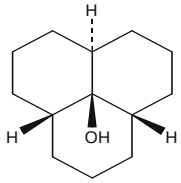
- Different software programs utilize different formats to input structural data
- It is often necessary to utilize several different software packages to gather the information you need on a single structure
- It is often necessary to convert file formats in order to do this – most software packages will read/write more than one format
- A standalone software package (babel) can also be used to convert among ~30 different molecular file formats

Molecular File Format Exercise

- Save your morphine molecule in pdb and TriposMOL2 formats (call them morphine.pdb and morphine.mol2)
- Open a new text editor and open the files from the text editor in order to see the file format rather than the interpreted structure
- Identify the atom type and position information in each file type

Visualization Exercise

- Draw a clear three-dimensional representation showing the preferred conformation of cis,cis,trans-perhydro-9b-phenalenol (below).



Problem taken from Carey and Sundberg, Advanced Organic Chemistry, Part A.

Public Databases

- Chemical structures can be obtained from several public and private databases
- These structures may be experimentally determined (by x-ray crystallography or NMR) or they may be computationally determined

Public/Private databases

- Experimental Structures
  - Protein/DNA structures http://www.rcsb.org
  - Small molecules http://www.ccdc.cam.ac.uk
- Theoretical models
  - Protein/DNA structures http://www.rcsb.org

Note: See also the Links section of www.rcsb.org for more databases

Structure Importation Exercise

- Go to the Protein Databank (www.rcsb.org)
- Search for a protein structure (use hemoglobin if you don't have any particular protein of interest)
- Save the protein structure to the My Documents folder
- Open the structure in MOE

Useful MOE Tools

- Sequence Window
  - Display menu allows you to highlight actual secondary structures (red=helix, yellow=sheet)
  - Display menu allows you to highlight hydrogen bonding (generally only for the backbone)
  - Residues or chains can be selected, and can be used to select corresponding atoms
- Main Window
  - Render->Draw menu allows you to show hydrogen bonds and protein ribbon diagram

Exercise

- Play with the structure!
- Ideas
  - Zoom in and out
  - Rotate the structure
  - Translate from side to side
  - Isolate a residue
  - Add hydrogen atoms
  - View the ribbon structure of the backbone
  - Find the residue on each end of the structure and display in spacefilling model mode