HyperChem, Release 2: Molecular Modeling for the Personal Computer

Brian J. Teppen

Department of Agronomy, University of Arkansas, Fayetteville, Arkansas 72701

Received August 12, 1992

HyperChem[1] is one of a rapidly growing number of molecular modeling programs for the personal computer (PC). This package allows one to draw, edit, and manipulate molecules to calculate energies and to optimize molecular geometries using either molecular mechanics or semiempirical quantum mechanics, and to perform molecular dynamics simulations. Like most molecular modeling programs, HyperChem deals well with small organic molecules but is tailored toward the modeling of biomolecules such as proteins and nucleic acids. While less powerful and slower than workstation molecular modeling programs, Hyperchem is useful and costs considerably less than its workstation rivals. HyperChem would make an excellent teaching tool for biochemistry or physical chemistry classes.

Computer system requirements include an IBM-compatible 486 PC (or 386 PC with a 387 math coprocessor), 4 MB RAM, a hard-disk drive with at least 20 MB free, VGA or Super VGA display with Microsoft Windows drivers, a mouse, DOS Version 3.1 or later, and Windows Version 3.0 or later. The program came with a copy-protection device that must be attached to the parallel port of the computer for HyperChem to run, so the software was tested on only one machine. The test PC was a Gateway 2000 486/33 with 8 MB RAM, an ATI Graphics Ultra video card, and Windows 3.0.

Installation of the program was well described in the HyperChem documentation and turned out to be blessedly uneventful. Unfortunately, the installation instructions lacked any discussion of how to optimally configure the DOS memory managers for running HyperChem. I called the Autodesk technical support staff for help on this and several other matters during the course of this review: They were generally slow to respond and gave little assistance. They will be of even less help in the future, because they stated their intention to refer all future technical questions to the dealer from whom the software is purchased. In short, don't expect much technical support.

The HyperChem documentation consisted of about 900 pages in three three-ring binders, including a glossary and a master index. Since this program appeals to the novice user by virtue of its price and PC implementation, it was wise of Autodesk to include 300 pages of explanatory material about computational chemistry. This material cites relevant literature, compares molecular mechanics and semiempirical methods, describes some of the algorithms used, and comments on the strengths and weaknesses of various methods so that even the uninitiated can get some help in choosing computational techniques and in assigning significance to the results. A reference manual expounds on each menu choice and is also available on-line; naturally, there were many instances when I wished for more help than the descriptions gave.

CREATING MOLECULES

Molecular structures can be created in four different ways. One can use the HyperChem editor to draw a two-dimensional representation of the molecule of interest and then use the "Model Build" feature to convert that drawing to a three-dimensional structure. This works well for small molecules or larger simple combinations of chains and rings.

Secondly, if one is interested in modeling polypeptides or polynucleotides, it is very easy to create macromolecules by choosing sequences of residues in their standard conformations from a palette and then adding secondary or tertiary structure. One can customize these residues and add new ones to the library as well. However, one cannot create library files for monomers other than peptides or nucleotides like one can with some of the workstation modeling programs. This would be very useful if one wanted to model, say, polystyrene or poly(dimethylsiloxane).

Another means of obtaining molecular structures to model with HyperChem is by importing files from the Brookhaven Protein Data Bank (PDB). These PDB files contain structures of large molecules that have been solved experimentally, but can typically be obtained only on tape or CD-ROM, so most PC users will have to belong to a workstation network to be able to use them. Indeed, this is one way to transfer data files between workstation and PC molecular modeling programs, since virtually all of them can both read and write in PDB format. In this context, HyperChem has a nice feature called "Complement Selection". Many of the PDB files contain lots of extraneous atom centers from solvents, and they can require quite a bit of cleaning up. In HyperChem, one can simply select the macromolecule of interest that was embedded in the PDB file, then use the "Complement Select" to select everything except the macromolecule, and then delete all selected atoms.

A final way to create a molecular structure is by using an ASCII text editor to type in atoms and their Cartesian coordinates in the proper HyperChem format. This flexibility is nice, but HyperChem lacks a utility to assign bonds between neighboring atoms, so one must include bonding information in the input text file. This can be very tedious to ascertain and tabulate for large molecules.

Once molecules have been created, they can be easily edited and manipulated using a mouse. The translation, rotation, and scaling of the graphical images were much smoother than I expected on a PC; slow movements of molecules were aesthetically tolerable and chemically revealing. The use of such a molecular modeling package along with access to PDB files would be an excellent teaching tool, allowing biochemistry students to interactively explore the three-dimensional structures of DNA or enzyme–substrate interactions.

MOLECULAR MECHANICS OPTIONS

In molecular mechanics jargon, the set of functions that are used to calculate the relative energy of a molecular conformation is termed a force field. HyperChem allows users to choose from four force fields: MM+ (MM2[2] with a few modifications and a default scheme), AMBER[3], OPLS[4] (an AMBER derivative especially for solvated macromolecules), and BIO+ (the empirical energy function of CHARMM[5] but without the CHARMM analysis facilities). Each force field is described in the HyperChem manuals.

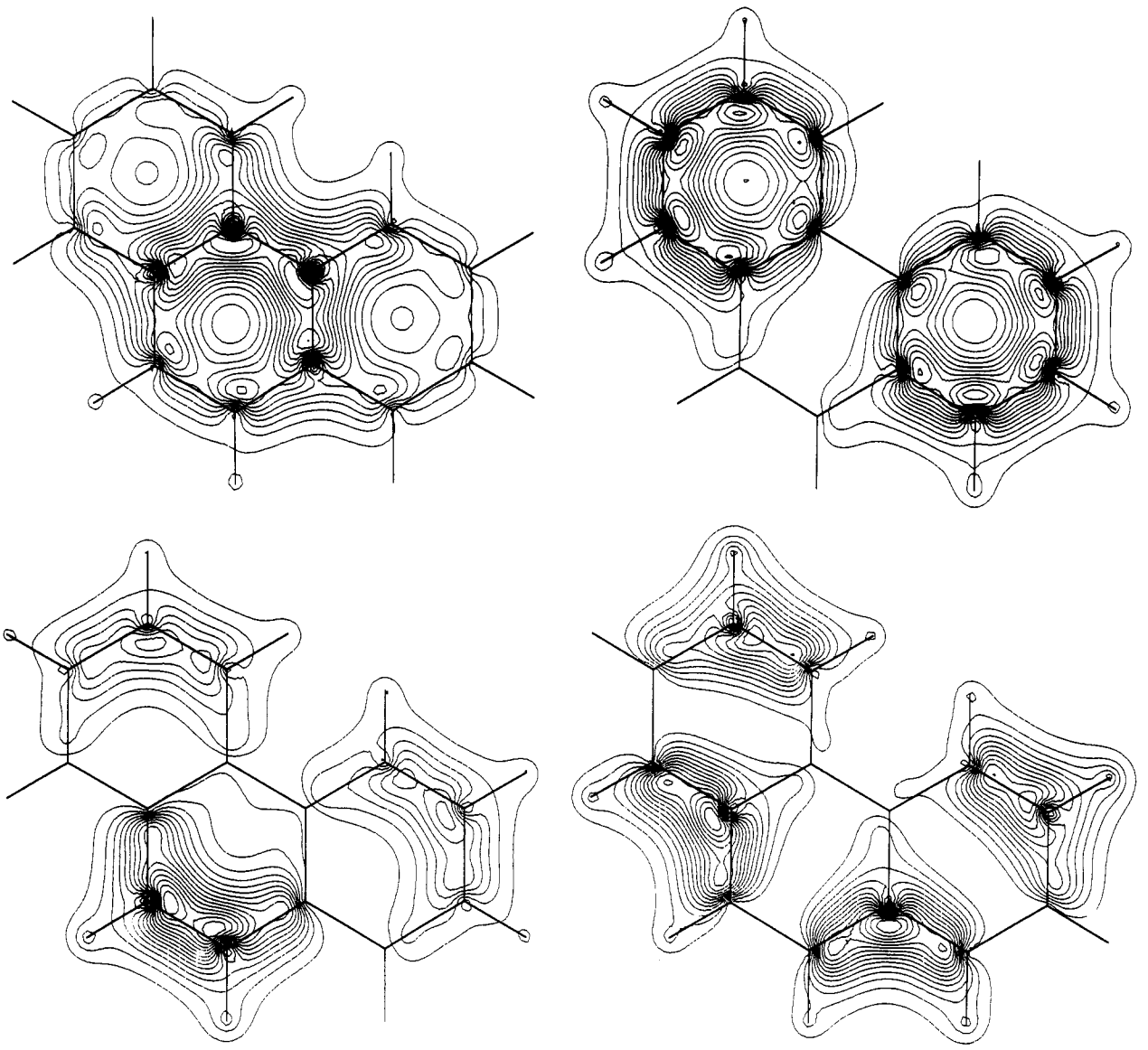

Figure 1. Four examples of molecular orbitals for phenanthrene calculated using the HyperChem implementation of AM1.

The most important part of molecular mechanics simulation is the development of effective force field parameters for an atom in a given chemical environment (an "atom type"). Each force field has its own parameter set(s), consisting of force constants for interactions between atoms of given types. Since there are many atom types, thousands of parameters are needed. Companies with the more established molecular modeling programs have done extensive and often proprietary parameter development, but HyperChem is new and its parameter sets are more limited. The HyperChem force field MM+ uses a 1991 parameter set for MM2 along with slight modification to the functional form. Since MM2 was developed primarily for small organic molecules, MM+ should be used with caution when modeling other systems. This is especially true since HyperChem will perform MM+ calculations for all possible systems, owing to a default parameter scheme similar to that for the DREIDING[6] force field. The default force field uses the MM2 functional form with a set of generic parameters for each element rather than for each atom type. Such a method will never fail from lack of parameters, but will usually be less accurate than methods using atom types in their specialized chemical environments.

The AMBER and OPLS force fields were designed for proteins and nucleic acids. Their implementations within HyperChem are restricted to parameter sets published by the force field creators, so little credence should be placed in simulations of molecules dissimilar to those used to develop the published parameters. Since OPLS parameters were derived for the solution state, they should be appropriate for many aqueous and organic solutions. The BIO+ force field parameters included with HyperChem have been long superseded by new developments to CHARMM, so should probably not be used without modification.

On the positive side, the HyperChem developers have greatly enhanced the value of this program by allowing the user to modify any and all parameter sets. As values for new atom types are developed by the user or found in the literature, they can be added to the proper parameter set using either a text editor or a database program. The new set can be renamed if the user desires to keep the old set, and the new parameter set can then be recompiled from within HyperChem and used like any other set. To test this procedure, I modeled some aluminum–silicic acid complexes. The MM+ force field had Si, O, and H interactions parameterized, but the organic bias of all force fields meant that none contained Al parameters. Adding an Al atom type to MM+ (with only O as its nearest neighbor and Si or H as second neighbors) meant I had to alter five files and add 10 new parameters. I did so using the DOS ASCII editor and found that the HyperChem documentation described the process very well. The new parameter

set successfully recompiled on the second try, and I was able to use it.

TEST CALCULATIONS

Once a force field is chosen within HyperChem, it can be used to calculate an energy for a molecular conformation, to "optimize" the geometry of a molecule, and to carry out molecular dynamics simulations. Using the test PC, a PDB file that was included in the tutorials for HyperChem was imported; assignment of atom types for the 652-atom protein crambin took 10 s. To calculate a single-point energy for crambin required 15 s using the AMBER force field.

Optimizations of molecular geometry can be done in HyperChem using three types of minimizers: steepest descent for quick and dirty elimination of high energy interactions, conjugate gradient methods that converge more consistently, and (for MM+ only) the Newton–Raphson method that uses second derivatives to estimate energy minima. For crambin with the AMBER force field, each iteration of the steepest descent algorithm required 13 s. Minimization using a conjugate gradient algorithm took 26 s per iteration. Unfortunately, even though HyperChem is geared toward the modeling of biomolecules, its utilities for conformational searching are far less sophisticated than those found in more established molecular modeling programs. In HyperChem, one needs to manually rotate each side chain under investigation and then reminimize. Alternatively, one can use HyperChem's molecular dynamics options to look for new minimum-energy geometries through simulated annealing, in which the molecule undergoes simulated heating and slow cooling to traverse energy barriers.

The molecular dynamics capabilities of HyperChem were tested using aqueous ethanol. An ethanol molecule was created and minimized in vacuo using OPLS. Then 10 copies were made and placed in a box containing 409 TIP3P[7] water molecules. This is the only water model that HyperChem uses, but it will fill any sized box with the number of molecules needed to give a proper water density. Infinitely periodic boundary conditions were invoked, and a single-point energy calculation for the 1317-atom system required about 170 s. Molecular dynamics simulation of the same solution required about 150 s per time step and proceeded smoothly.

SEMIEMPIRICAL QUANTUM MECHANICS

At a higher level of theory, HyperChem will calculate the semiempirical wavefunction for a molecule. The parameterized and simplified quantum mechanical methods available are extended Hückel, CNDO, INDO, MINDO/3, MNDO, and AM1, with the strengths and weaknesses of each succinctly described in the documentation. These methods can be used for single-point calculations or to perform geometry optimizations. After calculation of a wavefunction, one can plot the total electron density in the molecule, the electron density of any selected "canonical" (nonlocalized) orbital, or the Coulombic potential in any plane. These visualization features would make HyperChem an excellent educational tool for the study of bonding or molecular symmetry. For phenanthrene ($C_{14}H_{10}$), with 66 valence electrons, AM1 required 70 s to calculate the single-point wavefunction (using seven iterations) and another 110 s to calculate other molecular properties, including energies, heat of formation, dipole moment, and atomic charges by Mulliken population analysis. The generation of each electron density plot, such as displayed in Figure 1, took about 35 s. When pushed to the limit, the test PC was able to handle semiempirical calculations involving about 300 valence electrons using CNDO and about 200 electrons using MNDO or AM1. In the latter cases, each iteration toward convergence of the wavefunction took about 10 min. An AM1 wavefunction calculation was performed for a 78-atom molecule with 320 valence electrons. The wavefunction converged in eight iterations after 1 h and 50 min, but the test PC then ran out of memory while trying to calculate the molecular energy. Still, some impressive calculations can be made with HyperChem in reasonable periods of time.

CONCLUSION

For those who have a nice PC but not a nice workstation, HyperChem seems to be a solid molecular modeling package. There were several instances when I thought the program had locked up my PC, only to have it finish a calculation and return to readiness. The program did not crash during dozens of hours of use.

The relatively low cost of HyperChem compared to most molecular modeling packages is an indication that the molecular mechanics parameter sets within HyperChem are underdeveloped. The MM+ force field is an exception, but it was created for small organic molecules while the rest of HyperChem is geared toward modeling biomolecules. The program also lacks many of the sophisticated tools associated with workstation molecular modeling, the most serious of which are automated conformational searching routines.

On the whole, though, HyperChem fills a valuable niche for many potential users. Those who have long wanted to "experiment" with molecular dynamics simulations or who desire better visualization tools for teaching will find the hardware requirements as well as the software price much easier to handle. The graphical interface is simple to use and visually pleasing. Calculations will take more time than on a workstation, but surprisingly large systems can be investigated. For serious modelers, all parameter files are open to modification so that HyperChem can be used to do truly new research.

REFERENCES AND NOTES

(1) HyperChem is a trademark of Autodesk, Inc., 2320 Marinship Way, Sausalito, CA 94965. Telephone, (800)424-9737. List price $3500; educational discounted price $595 until January 1993. Free, noninteractive demonstration diskette available.

(2) Allinger, N. L. Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V_1 and V_2 torsional terms. *J. Am. Chem. Soc.* **1977**, *99*, 8127–8134.

(3) (a) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765–784. (b) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **1986**, *7*, 230–252.

(4) (a) Jorgensen, W. L.; Tirado-Rives, J. The OPLS potential function for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666. (b) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. OPLS potential functions for nucleotide bases. Relative association constants of hydrogen-bonded base pairs in chloroform. *J. Am. Chem Soc.* **1991**, *113*, 2810–2819.

(5) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.

(6) Mayo, S. L.; Olafson, B. D.; Goddard, W. A., III. DREIDING: A generic force field for molecular simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.

(7) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.