HyperChem Release 2 for the Silicon Graphics Workstation

D. W. Bearden

Chemistry Department, Clemson University, Clemson, South Carolina 29634-1905

Received March 1, 1993

INTRODUCTION

HyperChem Release 2 from Autodesk, Inc., represents an ambitious effort to put many elements of computational molecular science into a single, easy-to-use package. HyperChem combines single-point, geometry optimization, and molecular dynamics simulations using empirical or semi-empirical methods in a powerful graphical environment. The program provides user friendly pull-down menus and icons for easy use of the many features of HyperChem. It is apparent that the goal is to provide software that meets the needs of the small-molecule community in addition to the macromolecule and biomolecule communities.

HyperChem offers several force fields for empirical simulations. A modified MM2 (1991 parameters, 1977 functional form) force field is incorporated including non-bonded cutoffs (switched and smooth), periodic boundary conditions, and a modified bond-stretch term. The AMBER force field is available as both united-atom and all-atom forms. Finally, OPLS and BIO+ (an implementation of the CHARMM force field) are available. All the force field parameters are easily edited with text editors, and a mechanism is available for naming and keeping track of all the modified force fields one might produce so that the original forms do not have to be altered. None of the empirical calculations consider the explicit effects of delocalized electrons except through the empirical parametrization of the atom types. The MM+ parametrization scheme uses a wild card substitution method for dealing with atom types which lead to undefined modeling parameters. In the event that no suitable wild card substitution is found for missing parameters, a default parameter is generated by considering the hybridization of relevant atoms, the standard covalent radii, and the type of bond under consideration. Atoms with coordination greater than 6 are not handled by HyperChem. A particularly useful feature of the empirical force fields is the ability to modify the atom-typing rules in a text file. The syntax is simple and allows users to have complete control over how atom types are assigned for each force field.

Semi-empirical calculations can be performed using a variety of neglect of differential overlap (NDO) methods for atoms with atomic number 18 and lower. The methods supported are extended Hückel, CNDO, INDO, MINDO3, MNDO, and AM1. Parameters for PM3 are also included but are not available as a menu option. The PM3 parameters must be substituted for AM1 parameters by replacing the AM1 parameter files with the PM3 parameter files. Again, all of the parameter files are text files which can be easily edited, however undesirable that may be. The user must keep track of changes to the semi-empirical files since there is no mechanism for HyperChem to keep it sorted out. Calculations can be done with restricted Hartree–Fock (RHF) or unrestricted Hartree–Fock (UHF) methods of including spin interactions. Convergence can be accelerated in MINDO3, MNDO, and AM1 calculations with the direct inversion of iterative subspace (DIIS) procedure. Configuration interaction (CI) can be used during single-point calculations to take electron correlation into account in order to improve ground- and excited-state energies.

HyperChem has some capability of doing mixed semi-empirical/empirical mode calculations. In a large biomolecule, for example, small fragments or regions can be treated semi-empirically (at the CNDO or INDO levels) while the rest of the molecule is treated as an external potential. This is extremely useful when part of the large molecule is not adequately parametrized in the empirical force field.

INSTALLATION AND OPERATION

HyperChem is installed on Silicon Graphics systems using the system installation utility *inst* which is part of the IRIX operating system from Silicon Graphics. This standard utility makes software installation simple, reversible, and modular. Once the software is installed and the license manager activated, users can prepare to run HyperChem by executing a simple command to make personal copies of the necessary files. HyperChem can be launched by double-clicking an icon or by a command line in a shell window.

The printed documentation is extensive and very helpful. The manuals include the *Installation Guide*, the *Learning Guide* (a tutorial), *HyperChem Computational Chemistry* (Part 1: Practical Guide, and Part 2: Theory and Methods), and a *Reference Manual*. The *Reference Manual* has an index, but there is no overall index for the entire manual set. The textbook-like *HyperChem Computational Chemistry* provides a discussion of techniques in general and gives details of the HyperChem specific implementations of these techniques. The discussion is laced with useful examples and references. The tutorial in the *Learning Guide* covers almost all of the features of HyperChem in 13 lessons which are well organized and carefully written. The *Installation Guide* gives concise instructions for installing HyperChem, but previous experience with IRIX's *inst* utility makes it much easier reading.

An on-line help system is activated from the menu bar, but it is of dubious value. The help window has very small print that is hard to read. Context searching is very primitive and uninformative. No references are given to the page numbers in the printed documentation. Also, very few graphics are included in the help files and that makes certain sections useless. For example, descriptions of the various Tools available in the HyperChem window are meaningless since there is no way to logically connect the definitions given to the actual icons. Useful information about the scripting commands (vide infra), however, can be easily found using the on-line help system.

The window for HyperChem is neat and tidy with a layout almost identical to the Windows version of HyperChem for PC's. The menus are at the top of the window, and the eight icons for the Tools are drawn down the left side of the window. The bottom of the window contains an area which holds one row of characters for program output messages. The icons are perplexing at first since some of them look quite similar

to one another. A molecule is drawn in 2-dimensional form by choosing an element from the periodic table which appears after double-clicking on the drawing tool. Atoms are placed in the drawing area and connected together with familiar click-and-drag mouse techniques. Any atom can be changed to another element by choosing another element from the periodic table and left-clicking on the existing atom which is to be replaced. Valences are checked as you draw so mistakes are avoided. Multiple bonds are made by clicking on an existing single bond. Atoms or bonds are deleted with the drawing tool by clicking the right mouse button over the object to be removed. Once the atomic skeleton is drawn, hydrogens can be added by a menu selection and the model can be converted to 3-dimensional form by using the Model Build function from a menu. The Model Build function rapidly converts the sketch to feasible 3-dimensional form using predetermined bond lengths and angles. Stereochemistry is a bit hard to control using the Model Build function, but after some practice one can obtain the desired molecule. Various parts of the molecule can be "selected" using the many types of selection options provided. Selected parts of the molecule can be manipulated with Tools or a calculation can be done on just the selected part of a molecule. Although the user interface is quite powerful and has many options, one is left with a feeling of ungainliness since operations often take several mouse clicks or commands are located in menus which are not conveniently or logically placed.

Databases are available for quickly assembling amino acids and nucleic acids. It is very easy to choose standard conformations or specify non-standard conformations by specifying relevant angles. Double stranded DNA and RNA can be constructed by having the program automatically choose the complementary residue of each residue chosen. Mutations can be made to polypeptides and polynucleotides with a simple menu choice. No database for common organic molecule fragments is included, nor is there any mechanism for creating and maintaining new databases.

SCRIPTING

Automation of repetitive tasks is easily accomplished using the scripting system of HyperChem. Text files containing command sequences are created outside of HyperChem using a text editor and executed by invoking the Script pull-down menu. The commands available for scripting can duplicate any pull-down menu command in either interactive mode or in non-interactive mode. In interactive mode, the user can make selections or type filenames, for example, and in non-interactive mode the commands are executed sequentially without operator intervention. Additional commands are available for handling some of the details of script operation and for controlling the flow of messages and data. However, no flow-control commands, like looping, logical testing, or reading input from other files, are available for controlling scripts to make them more flexible. Ambitious users can write IRIX shell scripts which invoke the HyperChem scripts to generate more complex procedures.

ROOM FOR IMPROVEMENT

Most software packages have at least a few problems that keep them from being completely satisfactory to everybody in every situation, and HyperChem is no exception. What follows is a list of general problems that might be significant to certain users. The only way to do conformational searching is with molecular dynamics since there is no dihedral driver or systematic searching option. Only Protein Data Bank file format and the HyperChem file format are supported for importing and saving molecular structure data. Overlaying of molecules for structural comparison is done by selecting exactly three atoms in each molecule, and this simplistic comparison would be useful in a very limited number of cases. Empirical calculations have traditionally been used to estimate heats of formation and other quantities of interest, but the output from HyperChem is limited to the total steric energy and a breakdown of its components. Semi-empirical calculations must be done to get the thermodynamic quantities of interest. In general, the lack of tools for analyzing and displaying the results of calculations may severely restrict the broad application of HyperChem.

In addition to the above considerations, there are several errors in the program or in the IRIX operating system which are a nuisance. In many file name specification dialog boxes, the number of characters allowed for a filename is severely restricted. While all of the graphics features worked as advertised on an Indigo R3k Elan workstation running IRIX 4.0.5F system software, there were certain features which did not work properly or at all on an Indigo R4k XS with z-buffer running IRIX 4.0.5F system software. These missing graphics features did not affect the operation of the program beyond the point of usefulness, but users could have some trouble interpreting data or manipulating structures.

The HyperChem program runs as a single-user program. For example, if someone is running HyperChem at the main

Table I. Conformational Energy Differences of Some Substituted Cyclohexanes and Other Molecules[a]

compd		lit. diff[a] (kcal/mol)	MM+[b] diff	AM1[c] diff
substituted cyclohexanes				
-methyl	(ax–eq)	1.75	1.78	1.42
-ethyl	(ax–eq)	1.75	1.82	1.50
-phenyl	(ax–eq)	3.00	2.89	3.25
-Cl	(ax-eq)	0.40	0.42	0.83
-Br	(ax–eq)	0.40	0.47	*d*
-OH	(ax–eq)	0.50	0.59	1.28
-OAc	(ax–eq)	0.60	0.59	1.47
-COOH	(ax–eq)	1.35	1.40	1.42
-CN	(ax–eq)	0.17	0.20	1.15
-nitro	(ax–eq)	1.10	1.34	0.56
$-OCH_3$	(ax–eq)	0.60	0.55	1.52
$-NH_2$	(ax–eq)	1.50	1.39	2.67
$-N(CH_3)_2$	(ax–eq)	1.40	1.49	3.19
-*tert*-butyl	(ax–eq)	5.90	5.51	5.09
1,2-dimethyl	(cis–trans)	1.90	1.61	1.46
1,3-dimethyl	(cis–trans)	−2.00	−1.81	−1.44
1,4-dimethyl	(cis–trans)	1.90	1.75	1.42
1,1,3,5-tetramethyl	(cis–trans)	−3.70	−3.89	−3.08
trans-1,2-dimethyl	(aa–ee)	2.58	2.43	1.80
cis-1,3-dimethyl	(aa–ee)	5.50	5.34	4.19
trans 1,4, dimethyl	(aa–ee)	3.60	3.58	2.85
other molecules				
N-methylpiperidine	(ax–eq)	3.15	2.10	−1.41
2-methylpiperidine	(ax–eq)	2.50	2.00	1.72
3-methylpiperidine	(ax–eq)	1.80	1.67	0.76
4-methylpiperidine	(ax–eq)	1.90	1.74	1.30
decalin	(cis–trans)	2.20	2.73	2.18
bicyclo[3,3,0]octane	(cis–trans)	−6.80	−6.79	−14.07
N-methylacetamide	(E-Z)	2.30	0.59	−0.54
butadiene	(ax–eq)	2.50	2.54	0.77
Acryloin	(cis–trans)	1.70	1.71	0.22
RMS dev			0.41	1.88

[a] Reference 1 and references therein. [b] Used gradient cutoff of 0.01 kcal/(Å mol), Polak–Ribiere optimization algorithm. [c] Used gradient cutoff of 0.03 kcal/(Å mol), Polak–Ribiere optimization algorithm. [d] AM1 in HyperChem is not parametrized for atoms with $Z > 18$.

Table II. Minimization Performance[a,b]

compd/method	no. of cycles	final energy (kcal/mol)	final gradient (kcal/(Å mol))	time	
				R4k Indigo[c]	R3k Indigo[d]
phenylcyclohexane					
steepest descents	21174[e]	4.443 64	0.019 1353	21.5 min	42.1 min
Fletcher Reeves	297	4.413 49	0.004 9917	0.6 min	1.4 min
Polak Ribiere	418	4.413 47	0.004 5712	0.8 min	1.8 min
Newton Raphson	1191	4.413 69	0.004 9891	1.2 min	2.5 min
butane					
steepest descents	2291	3.035 10	0.004 9705	47.0 s	100.0 s
Fletcher Reeves	111	3.034 78	0.004 4016	6.0 s	11.0 s
Polak Ribiere	137	3.034 79	0.004 5201	7.0 s	12.0 s
Newton Raphson	108	3.034 78	0.002 5874	5.0 s	7.0 s

[a] Reference 1 and references therein. [b] MM+ molecular mechanics of all atoms, gradient cutoff of 0.005 kcal/(Å mol), maximum number of cycles set to 60 000. The values of no. of cycles, final energy, and final gradient were identical on each tested workstation. [c] Silicon Graphis Indigo R4000 (50 MHz, IP20 processor), GR2-XS graphics board with Z-buffer, 48 Mbyte main memory, IRIX system software 4.0.5F. [d] Silicon Graphs Indigo R3000 (33 MHz, IP12 processor), GR2-Elan graphics board, 32 Mbyte main memory, IRIX system software 4.0.5F. [e] Minimization terminated at this point with the message "Step size too small, optimization stopped."

console screen, users who log-in from remote Silicon Graphics workstations will only be able to run the program over the network on the licensed workstation if multiple copies of the software license have been purchased. In addition, it is not possible to run the program in the background and log-out without killing the computation in progress. As a result, other users cannot log-on the main console to do any other work. This may present problems in environments where the workstation console must be shared among users.

CALCULATIONS

HyperChem can basically perform three types of calculations. Single-point calculations are performed to generate the static properties of a molecule in any conformation. Properties such as molecular orbital coefficients, molecular orbital energies, atomic charge distributions, and electrostatic potential surfaces can be calculated. The electrostatic potential, total spin density, total charge density, and the molecular orbitals may be shown as contour plots at a chosen cross-section or as wire-mesh or solid surfaces. Once computed, the particular surface grid values may be saved for recall and display. The surfaces are displayed one at a time, so multiple sessions of HyperChem must be run to have the HOMO and LUMO, for example, displayed on the screen simultaneously. In addition, the program recomputes the molecular orbital coefficients every time a new surface is generated, even if the molecule has not changed, making for time consuming display of surfaces. HyperChem puts the numerical results of single-point calculations into log files (after one explicitly turns on the logging by a menu choice) which must be read with a text editor. The amount of information generated is controlled by a print-out variable which is set by a parameter inconveniently located in the start-up initialization file (or through a user-written script).

Geometry optimization or molecular mechanics is used to minimize the conformational energy of molecules to find structures which correspond to minima on the multidimensional potential energy surface. The user has great control over the details of the minimization procedure. In addition to choosing the force field, the user chooses the search algorithm, the final gradient, and the maximum number of cycles of minimization. The status of the minimization is displayed in a single line at the bottom of the HyperChem window, and if file logging is turned on, the progress can be archived for future reference.

Table III. Structure of Sulfonamide[a,b]

dimension or angle	6-31G* geom[a]	MM+ geom	AM1 geom
H–S	1.321	1.346	1.232
S–N	1.636	1.642	1.592
S=O	1.424	1.450	1.406
H–S–N	101.30	101.36	99.48
H–S–O	107.20	106.84	107.64
N–S–O	108.10	109.65	110.86
O=S=O	122.70	120.69	118.51
S–N–H	113.00	118.64	123.69
H–N–H	113.80	118.26	112.54
H–S–N–H[a]	114.50	102.04	91.81
H–S–N–H[b]	−114.50	−102.05	−91.84

[a] Reference 1 and references therein. [b] Distances are given in angstroms; angles are given in degrees.

The conformational energy differences from geometry optimizations of several small molecules using the MM+ empirical force field and the AM1 semi-empirical force field are shown in Table I. This set of substituted cyclohexanes and other molecules was chosen for comparison to a similar set of calculations using a suite of molecular modeling programs.[1] The RMS deviations from experimental difference values are 0.4 kcal/mol for the MM+ calculations and 1.9 kcal/mol for the AM1 calculations.

The choice of search algorithm is critical to the performance of the geometry optimization. Results for all atom MM+ minimizations of two molecules on two types of Indigo workstations are shown in Table II. Again, the molecules were chosen to provide easy comparison to previously published results.[1] The equatorial isomer of phenylcyclohexane was drawn in a coplanar starting geometry, and it minimized to the lower energy orthogonal conformation in all cases. Butane was drawn in eclipsed form, and it minimized to the gauche conformer in all cases, even though that is not the global minimum conformation. The number of cycles and the time required for minimizing the structures ranged over more than an order of magnitude and the final energies varied slightly from algorithm to algorithm. In the case of the Steepest Descents algorithm, the minimization of phenylcyclohexane terminated with the message "Step size too small, optimization stopped."

Comparison of calculated geometries with experimental geometries is difficult, and is very highly dependent on the molecules chosen for comparison and the source of experimental results. We performed two tests of the geometries resulting from minimizations. First, the minimized geometry for sulfonamide (HSO_2NH_2) was calculated (Table III) for

Table IV. Structure of Some Aromatic Molecules[a,b]

compd	bond[a]	exp length[a,c]	previously calcd[a,d]	HyperChem MM+[e]	HyperChem AM1[f]
phenanthrene	*a*	1.457	1.417	1.400	1.413
	b	1.381	1.382	1.394	1.380
	c	1.398	1.406	1.393	1.406
	d	1.383	1.386	1.396	1.381
	e	1.405	1.418	1.404	1.414
	f	1.448	1.457	1.410	1.446
	g	1.404	1.412	1.405	1.416
	h	1.390	1.440	1.397	1.435
	i	1.372	1.362	1.392	1.357
RMS dev				0.026[g]	0.005
chrysene	*a*	1.427	1.420	1.400	1.417
	b	1.367	1.379	1.393	1.377
	c	1.392	1.409	1.392	1.410
	d	1.379	1.383	1.396	1.378
	e	1.408	1.424	1.405	1.419
	f	1.465	1.452	1.412	1.441
	g	1.406	1.411	1.403	1.416
	h	1.418	1.431	1.394	1.427
	i	1.365	1.366	1.392	1.362
	j	1.423	1.437	1.401	1.430
	k	1.397	1.406	1.411	1.407
RMS dev				0.024	0.005
triphenylene	*a*	1.416	1.416	1.406	1.411
	b	1.377	1.387	1.393	1.385
	c	1.402	1.397	1.389	1.399
	d	1.415	1.412	1.408	1.415
	e	1.447	1.468	1.409	1.452
RMS dev				0.027	0.008
pyrene	*a*	1.380	1.395	1.395	1.393
	b	1.420	1.406	1.398	1.401
	c	1.417	1.416	1.401	1.419
	d	1.442	1.445	1.398	1.440
	e	1.417	1.431	1.401	1.433
	f	1.320	1.363	1.395	1.357
RMS dev				0.027	0.004
perylene	*a*	1.400	1.423	1.397	1.421
	b	1.370	1.375	1.390	1.372
	c	1.418	1.411	1.393	1.411
	d	1.397	1.393	1.404	1.385
	e	1.471	1.473	1.407	1.459
	f	1.425	1.431	1.406	1.432
	g	1.424	1.417	1.406	1.419
RMS dev				0.030	0.006

[a] Reference 2 and references therein. [b] Distances are given in angstroms. [c] Extensive footnotes in ref 2 question the reported values for some of these bond lengths. [d] These values are calculated with a hybrid empirical/semi-empirical method meant to take effect of delocalized electrons into consideration. [e] Minimization used Polak–Ribiere algorithm, and a cutoff gradient of 0.005 kcal/(Å mol). [f] Minimization used Polak–Ribiere algorithm, and a cutoff gradient of 0.03 kcal/(Å mol). [g] RMS deviation from the previously calculated values in the third column from the right.

comparison to a recent report,[1] and, second, five highly conjugated aromatic molecules[2] were minimized using MM+ and AM1 to see what the effect of the lack of consideration of delocalized electrons would be in MM+ calculations (Table IV). For molecules with delocalized electrons, a semi-empirical minimization is clearly required for accurate bond lengths.

Molecular dynamics simulations, which simulate the motion of atoms and molecules by solving the Newtonian equations of motion, are useful for, among other things, conformational searching, docking of molecules and obtaining thermodynamic information. The molecular dynamics module of HyperChem allows for easy setup of runs in constant energy or constant temperature mode. Simulations consist of a heating period (which may be zero), a running period, and a cooling period (which may be zero) with a user-chosen size for the integration time step. Simulations may be run in vacuum or in water solution with periodic boundary conditions. The initial, running, and final temperatures are chosen by the user. Data from the simulation (positions and velocities) can be stored at specified intervals into binary snapshot files which can be analyzed and re-played. Limited graphing of at most four quantities like the total energy, the temperature or user-defined geometrical features is possible during simulation or playback. The graphed quantities are stored in text files as the graph is being drawn.

CONCLUSIONS

HyperChem is a remarkably easy program to use, and as far as I can tell, the science built into HyperChem is sound. With a little practice to become familiar with the user interface, anyone can start modeling compounds using a large number of variations of empirical and semi-empirical techniques. HyperChem would be an excellent program for teaching or learning molecular modeling, because, while no program can be absolutely foolproof and error free, it is robust enough to help the novice user avoid common mistakes and painlessly explore the frontiers of computational chemistry. Experienced computational chemists may benefit from the seamless integration and flexibility of the many techniques which are incorporated in HyperChem, even though there are some deficiencies and limitations as mentioned in this article which should be addressed in future releases of the program.

REFERENCES AND NOTES

(1) Bays, J. P. *J. Chem. Educ.* **1992**, *69*, 209–215.
(2) Allinger, N. L.; Sprague, J. T. *J. Am. Chem. Soc.* **1973**, *95*, 3893–3907.