Teaching Computational Chemistry Using Computers

Jacques Weber* and Pierre-Yves Morgantini

Introduction

This title is not as paradoxical as it seems in a first sight, since it is only a recent fact that computational chemistry (CC) can be efficiently taught using computers, both in the auditorium (plenary lectures) and in the laboratory (practical exercises). This is undoubtedly a great progress to the equal profit of students and instructors. Indeed, lectures devoted to chemical information retrieval or to molecular modeling, to take simple examples, can be vividly illustrated by performing on-line computer demonstrations which are conveniently projected on a large screen using a videodata equipment. This has been made possible only recently through spectacular progresses both of hardware (powerful PCs, workstations, videodata projectors, networking, etc.), concomitant of falling costs, and of software (user friendly program packages with graphical user interfaces). Of course, all the topics of the broad field which constitutes CC cannot equally benefit from these new facilities, and it is difficult, e.g., to teach the basic theories underlying several areas of CC using computers, but their applications are again good candidates for computer-aided teaching.

In this Column, we would like to present and illustrate by examples the various tools we use in Geneva to teach CC using computers. In all cases, the presentation will cover teaching applications developed for both auditorium lectures and practical exercises in the laboratory.

Scope of the Applications

Before looking at applications devoted to CC teaching, it is probably worthwhile to define again computational chemistry (CC), synonymous with computer-assisted chemistry, consists of all the areas of chemistry which benefit from the use of computers' [1]. Subsequent discussions with colleagues have convinced us that this definition is too broad, as practically all the areas of chemistry benefit today from the use of computers, which leads to the nonsense that CC would actually be nothing else than chemistry itself! Rather, we prefer the following definition: CC consists of all the areas of chemistry where the computer is the essential piece of instrumentation. This means that all the traditional fields of chemistry, namely synthesis, analysis, spectroscopy, etc., are de facto excluded from CC since, even though they may be somewhat computerized, computers cannot be regarded as being the essential part of their equipment.

It is generally accepted that CC is made of the following subfields: structure elucidation and analysis, management of chemical information and theoretical chemistry. Whereas the first one is obviously related to data acquisition and processing in chemical experimentation, the second one deals with all possible applications concerned with chemical data bases. The third one, the classical area of CC, encompasses all modeling applications, performed using either computational quantum or classical mechanics, molecular dynamics and Monte Carlo simulations, reaction dynamics and molecular graphics. In this Column, we are going to concentrate on teaching applications of CC related to this vast third subfield. Indeed, structure elucidation and analysis is a rather specialized topic which is not frequently taught in first- or second-cycle university lectures, whereas it is relatively easy to introduce students to the management of chemical information at an elementary level which enables them to perform their own simple Chemical Abstracts Service (CAS) on-line searches, for example [2]. As to more sophisticated and specialized 3D-database searches such as 'database mining' and 'electronic screening' [3], these new CC applications are just emerging and they can be adequately taught only in highly specialized lectures or seminars. The most general part of CC teaching is, therefore, undoubtedly that devoted to computational theoretical chemistry, as it is part of the physical chemistry curriculum in all Swiss Universities and Institutes of Technology.

In this Column, the use of computers in the following CC lectures, which are all part of the chemistry diploma curriculum of the University of Geneva, will be presented:

- introduction to quantum chemistry (IQC), part of the general physical chemistry I course (36 h of lectures + 24 h of exercises, altogether, 4th semester);
- introduction to group theory (IGT), part of the general physical chemistry II course (12 h of lectures + 4 h of exercises, altogether, 5th semester);
- molecular modeling and simulation (MMS), optional course (60 h of lectures + 40 h of exercises, altogether, 7th and 8th semesters).

*Correspondence: Prof. J. Weber
Department of Physical Chemistry
University of Geneva
30 quai Ernest-Ansermet
CH-1211 Geneva 4
Hardware and Software Available

Auditorium Lectures

In the Sciences II building, which hosts our Chemistry Section, one finds several auditoriums among them the 100 seats A-100 is conveniently equipped with audiovisual material for computer-aided teaching (Fig. 1). Indeed, in addition to a 486/33 PC under Windows with VGA (1024 x 768) graphics connected to the Ethernet network of the University, it is possible to use a Sony VPH 1270 videotape projector with a resolution of 1100 x 970 pixels and a projected image dimension of 3 m. In addition to being able to project multi-norm videotapes, this system has a direct RGB connection to the PC, and built-in interfaces as well for RGB connections to Macintosh computers and standard workstations (Sun, SGI etc.) which can be easily installed in the auditorium. A second classroom with a similar equipment is available in the building.

Using these equipments, the instructor is able to project in real time the screen of the PC, or of any workstation if more compute power or better quality graphics is needed. He is thus in a position to interactively display and comment any step in using CC software packages, would they be education oriented such as HyperChem [4], or research oriented such as Spartan [5]. Of course such a distinction is somewhat arbitrary and one may find abundant examples demonstrating that most CC packages can be used for both purposes. In addition, connection of the PC or the workstation to the Ethernet network allows the instructor to use any program installed on another host system. These facilities enable instructors to vividly illustrate, using interactive program operations, virtually any topic of computational theoretical chemistry. In addition, the merging of video images and motion sequences, stored, e.g. on a videodisk, with instructional computer programs offers the possibility to present multimedia CC courses, which have undoubtedly a great potential in future chemistry teaching [6].

As far as software is concerned, we incline to think that, with the advent of computer algebra systems such as Maple [7], it is not indispensable that first- and second-cycle students in chemistry are literate in a true scientific programming language, such as PASCAL or FORTRAN, to take CC lectures or exercises. The situation in this respect is somewhat different from that which prevailed three years ago [8], which may be attributed to the fact that Maple and similar systems recently emerged as powerful tools to perform symbolic, numeric and graphical computations involved in mathematical problem solving. To a very large extent, they may thus serve as substitutes for scientific programming languages, but this statement does of course not apply to graduate students preparing a Ph.D. in CC as they must be able to carry out some programming developments in FORTRAN. We assume therefore that, for the applications in CC teaching presented in this Column, both students and ... instructors have some knowledge of Maple (or of any other similar computer algebra system).

In addition to the standard packages HyperChem [4] and Gaussian 92 [5] mentioned above, we use in our CC teaching applications some other commercial programs such as Spartan [9] and MacroModel [10], and the Molekel program as well which has been developed in our laboratory [11]. Whereas HyperChem is a general purpose molecular modeling program enabling the user to perform molecular mechanics, dynamics and semi-empirical quantum chemistry calculations, Macro-model exhibits similar features but it is restricted to molecular mechanics and dynamics. Both packages offer the possibility to build interactively molecular structures. Gaussian 92 and Spartan are well-known quantum chemistry programs including semi-empirical and ab initio methods, the characteristic of Spartan being its graphical user interface. Finally, Molekel is a molecular graphics package allowing high-quality visualization of chemical systems and properties.

Practical Exercises

Sessions of practical exercises take place in a new PC room which is equipped with 8 DX4/100 diskless machines connected through Ethernet to a Novell network with the central PC of the Section of Chemistry library as a server (Fig. 2). The choice of diskless PCs has been dictated by financial, practical, and safety reasons. Extension of this facility to 16 PCs is planned for 1995. To reduce computation times and to allow students to perform the maximum number of calculations within a given time, the HyperChem jobs may run on the PCs or remotely on an Indi R4600 SGI workstation (16 Mflops) server. Similarly, all the molecular mechanics, dynamics, and quantum chemistry programs used by the students in their exercises sessions run on SGI workstations of the Department of Physical Chemistry. The CC software which is available to the students for their exercises is the same as that used for auditorium lectures.

CC Teaching Applications

Course Introduction to Quantum Chemistry

The course content is devoted to the following topics:
- basic quantum theory;
- applications to simple systems (particle in a box, harmonic and anharmonic oscillator);

Fig. 1. A computational chemistry lecture in the A-100 auditorium of the Chemistry Section of the University of Geneva, with the videotape projector (top right) displaying on the screen a space-filling model of the butane molecule generated on the PC by HyperChem.
atomic structure;
molecular electronic structure: diatomics, small molecules, Hückel molecular orbital (HMO).

Practically all the auditorium lectures can be enriched by an appropriate use of the Maple software on the PC together with real-time projection of the PC display. Similar applications have been recently reported by Rioux using the Mathcad system [12]. Actually, Maple enables the user to combine, into the same document, text, data, mathematics and graphics, which leads to the maximum flexibility. For example, in the context of these IQC lectures, mathematical computations such as differentiation and integration are used in the following applications:

- illustration of the Heisenberg principle using test functions f(x), that is calculation of the commutator [x, p_x] applied to f(x) in order to demonstrate that this expression is proportional to f(x). Simple f(x) functions such as exp(-x), cos(x), etc., are appropriate.
- Calculation of the normalization factors of the eigenfunctions of the particle in a box for various values of the quantum number n to show that they are not n dependent.
- Calculation of the average radius of the 1s orbital in the H-atom.

All these applications have been adapted to Maple from a software distributed by the American Chemical Society [13].

Several other topics of the course are also taught with the assistance of both mathematical computations and graphics facilities offered by Maple:

- 2D and 3D representation of the solutions of the Schrödinger equation of the following systems: particle in a 1D and 2D box (Fig. 3), harmonic oscillator, hydrogenic atom (radial wavefunctions and polar plots of the angular ones, both as surfaces in 3D space and as contour levels in selected planes).
- Potential energy curves of diatomics calculated in the LCAO approximation and corresponding MOs.
- HMO wavefunctions and energies calculated for simple systems; correlations (linear regressions) between selected properties (e.g., oxidation potentials) and HMO results (e.g., HOMO energies).
- Numerical solutions of the Schrödinger equation in the case of the harmonic oscillator and comparison with the analytical ones, which allows the student to understand the role of quantum numbers.

The exercises of this course consist of a series of Maple applications of these topics which are designed so as to fit to 2 h sessions. Control of the students' work is fairly easy as they have to print out at the end of the session the Maple file corresponding to the live document they have been preparing to solve the problem. In general, the subjects of these exercises have been chosen such as the students have to combine mathematical calculations and graphics in the same session. A good example is provided by the calculation of the potential energy curves of the HCl and HBr molecules in the Morse approximation and their graphical representation together with the first anharmonic vibrational energy levels. Of course, such a conception of the exercises of a quantum chemistry course requires to organize in the beginning some tutorial sessions on Maple, but this is a good investment for the students as they are then able to use this system for many other applications.

Course 'Introduction to Group Theory'
The application of group theory to molecular symmetry is an important topic as it has practical consequences in all fields of chemistry [14]. This subject deserves, therefore, undoubtedly its place in the chemistry curriculum, but, for the average student, its mathematical aspects should be reduced to a minimum and emphasis should be placed on the symmetry of molecules as an intrinsic property due to their shape. Furthermore, we are of the opinion that this course represents an ideal application field of molecular graphics.
packages such as Molekel. Our choice of Molekel as the basic program to teach this course is motivated by two reasons: i) it is essential to have access to the source code of such a program so as to implement the representation of symmetry elements superposed to molecular models; ii) real-time 3D interactive manipulation of molecular models and of corresponding symmetry elements is essential and this can only be performed using a package running on a graphics workstation.

We have, therefore, achieved several teaching developments in this direction and the auditorium lectures of the ICT course make the maximum use of video projections of the display of the SGI workstation running the Molekel program. This equipment is used to display the following applications:

- visualization of the $C_n$ symmetry axes, $\sigma$ mirror planes, and $S_n$ improper rotation axes superposed to the molecular model of any compound belonging to any point group (Figs. 4 and 5).

- For the $C_n$ axes, it is possible to dynamically represent the $360^\circ/n$ rotation of the molecular skeleton around the axis, in order to check that the operation leaves the molecule unchanged.

- These two features of Molekel are also applicable to 3D models of MOs, which is a valuable tool to determine their symmetry type, i.e., the irreducible representation to which they belong. It is well known that this information is essential to apply electric dipole selection rules, for instance.

- Similarly, Molekel allows the instructor to dynamically display the normal modes of vibration of any molecule, after their calculation has been performed using a suitable program, e.g. Gaussian 92. Their symmetry species can than be determined using the same tools, which is important to interpret IR and Raman spectra.

The subjects of exercises of this course consist of determining the point groups to which a series of molecules belong, visualizing the corresponding symmetry operations, finding the symmetry type of MOs and normal modes of vibration, and applying the selection rules to electronic and vibrational transitions and of Raman activity as well. As these exercises are carried out by the students using the Molekel package running on SGI workstations, they take place in our laboratory which hosts seven such machines.

Course Molecular Modeling and Simulation

This is the typical CC course which is intended to teach modeling and simulation techniques and to train students to perform by their own some applications relevant to various fields of chemistry. It is taken each year by 15–20 students who are in general highly motivated and conscious of the fact that computational chemistry is part of modern physical chemistry research. The content of the course is devoted to the following topics:

- basic features of the Maple system;
introduction to computer graphics;
- the various force fields of molecular mechanics; optimization techniques;
elements of molecular graphics, introduction to the HyperChem package;
molecular dynamics simulations;
- quantum chemistry: derivation of the Hartree-Fock equations; electron correlation; ab initio and semi-empirical techniques.

Most of the lectures devoted to these topics make use of live projections of applications of the HyperChem and Gaussian 92 program packages, which allows the instructor to vividly illustrate any particular aspect of modeling techniques. For example, it is possible to present and discuss the form of the various potentials of the four different force fields implemented in HyperChem and to emphasize their specificities by optimizing the geometry of selected compounds using all of these methods. One may also demonstrate the importance of the electrostatic component of the steric energy by switching it on and off in MM+ calculations, for instance, and by using then various sets of atomic charges to investigate their influence.

Molecular dynamics (MD) is another topic for which the instructor can take advantage of the real-time projection facility of a HyperChem application. Indeed, in addition to visualizing the trajectories of atoms generated by a MD simulation of any type, it is possible to display the energy components and the value of structural parameters as a function of time, which allows to follow conformational changes as the simulation proceeds. A good example is provided by a teaching application of the simulated annealing technique [15] we have developed recently [16].

Indeed, an interesting feature of MD techniques lies in the fact that, when sampling the configuration space, the kinetic energy of the system allows it to surmount energy barriers of the order of $kT$ per degree of freedom. It is, therefore, possible to artificially raise the temperature to search larger portions of the configuration space or, alternatively, to cool down the system to reach minima on the Born-Oppenheimer energy surface, that is to remove kinetic energy from the system, which is known as simulated annealing [15].

As an example, Fig. 6 summarizes the results of a MD calculation performed for a single molecule of butane (CH$_3$-CH$_2$-
Fig. 7. a) Newman projection of the butane molecule with definition of the torsion angle \( \omega \); b) potential energy curve calculated for butane as a function of \( \omega \) using the MM+ force field of HyperChem.

Fig. 8. Windows of the HyperChem and Excel systems with the workingspaces used by a student to investigate the C-C-C-C torsional potential function of butane. The left window displays the decomposition of the potential function into its various MM+ components, whereas the right windows depict some molecular models of butane.
modes of vibration using the Molekel package. In all the cases, the emphasis is placed on the comparison with experimental values, so as to estimate the quality/cost ratio of the various models, and on the relevance of performing such calculations in chemistry. Although it is far from being exhaustive, this list shows that several important concepts of computational quantum chemistry can be taught interactively using live projections of computer displays, which actually represents a significant progress in chemical education.

The exercises in the PCs room constitute an important component of this MMS course, as they allow the students, under the supervision of assistants, to have their hands on computer keyboards so as to prepare datafiles and to run the calculations, and, a key point in CC, to perform a proper interpretation of the results. We particularly insist on this latter point, as it is useless to carry out (time-consuming or not) calculations without a deep understanding of the meaning of the results. The subjects of the exercises have been mostly developed by our own group, but some of them have been taken from the monograph of Foresman and Frisch [18]. They are listed below:

1. General case of least-squares fitting of a linear combination of $M$ basis functions to $N$ data points using Maple [19].
2. Optimization by simulated annealing theory: minimization of functions of several variables using Maple [20].
3. Geometry optimization of the various configurations (chair, boat, twist-boat) of cyclohexane using the force fields of HyperChem.
4. Geometry optimization of the various configurations of piperidine, $N$-methylpiperidine, piperezine, and tetrahydropyran using the force fields of HyperChem.
5. Investigation of the C-C-C-C torsional potential functions of butane and 2,3-dichlorobutane using the force fields of HyperChem (Fig. 8).
6. Structure-activity relationships studies of morphine and mescaline derivatives using molecular mechanics (HyperChem) and graphics (Molekel).
7. Gas phase conformational analysis (Ramachandran plots) of glycine using the force fields of HyperChem.
8. Conformational analysis of $\beta$-alanine in gas phase and aqueous solution using the BIO+ force field of HyperChem.
9. Geometry optimization of small organic molecules and of malonaldehyde using the MNDO, AMI, and PM3 semi-empirical methods implemented in HyperChem.
10. Potential energy surface of the HCN $\rightarrow$ HNC isomerization reaction using the AMI and PM3 methods.
11. AMI molecular electrostatic potentials of organic compounds: hydroxylamine, formamide, pyrorule, pyridoline and adenine; correlation with HOMO localization and gas phase proton affinity.
12. Ab initio geometry optimization of ethylene, fluoroethylene and propene at various levels of theory using Gaussian 92.
13. Calculation of the vibrational frequencies of acetaldehyde using the ab initio method (Gaussian 92) and visualization of the normal modes of vibration using Molekel.
14. Geometry optimization of transition states using the ab initio method. Application to HCN, vinyl alcohol, and formamide.
15. Calculation of the geometry of the water dimer and of the dimerization enthalpy of water using the ab initio method. Practically all these exercises cannot be considered as simple ('presse-bouton') applications of the CC programs used. They require a good understanding of both theoretical chemistry concepts and program packages functionalities. However, in most of the cases, students are able to succeed in doing them with a limited help from the assistants. In addition, the use of the graphical user interfaces available for both input and output has been found to be effective in the various steps of the exercises.

As compared with the situation which prevailed 5–10 years ago, it is, therefore, no overstatement to assert that computational chemistry can be taught to a large extent using computers.

The authors are grateful to all their coworkers and students who were patient and persevering enough in the various phases of the development and tuning of these teaching applications. As research and education are often intimately connected, some of these teaching developments have benefitted from investigations carried out in the context of the Project 203613 1.92 of the Swiss National Science Foundation.

Received: February 2, 1995

[1] J. Weber, H. Huber, H.P. Weber, Chimia 1992, 46, 84.
[2] E. Zass, Chimia 1994, 48, 109.
[3] H.P. Weber, Chimia 1994, 48, 112.
[4] HyperChem 4.0, Hypercube Inc., 419 Philipps Street, Waterloo, Ontario, Canada N2L 3X2, 1994.
[5] Gaussian 92, Revision A, M.J. Frisch, G.W. Trucks, M. Head-Gordon, P.M.W. Gill, M.W. Wong, J.B. Foresman, B.G. Johnson, H.B. Schlegel, M.A. Robb, E. Togo, M.W. Rabuck, K. Raghavachari, J.B. Pople, J. Chem. Phys. 94, 2199, 1991.
[6] L.L. Jones, S.G. Smith, Pure Appl. Chem. 1993, 65, 245.
[7] B.W. Char, K.O. Geddes, G.H. Gonnet, B.L. Leong, M.B. Monagan, S.M. Watt, Maple V Language Reference Manual, Springer, New York, 1991.
[8] J. Weber, H. Huber, H.P. Weber, Chimia 1992, 46, 224.
[9] Spartan 3.1, Wavefunction Inc., 18401 Von Karmen # 210, Irvine, CA 92715, USA, 1994.
[10] Macromodel 3.1, Columbia University, New York NY 10027, USA, 1991.
[11] P.F. Fluekiger, Ph.D. thesis No. 2561, University of Geneva, 1992.
[12] F. Rioux, J. Chem. Educ. 1992, 69, A240.
[13] F. Rioux, ‘Enriching Quantum Chemistry with Mathcad’, JCE: Software, Department of Chemistry, University of Wisconsin, 1101 University Avenue, Madison, WI 53706-1396, USA.
[14] F.A. Cotton, ‘Chemical Applications of Group Theory’, 3rd edn., Wiley, New York, 1990.
[15] S. Kirkpatrick C.D. Gefatti, M.P. Vecchi, Science 1983, 220, 671.
[16] J. Weber, A. Deloff, P.F. Fluekiger, Speedup 1994, 8, 63.
[17] N.L. Allinger, S. Profeta, J. Comput. Chem. 1980, 1, 181.
[18] J.B. Foresman, A. Frisch, ‘Exploring Chemistry with Electronic Structure Methods: A Guide to Using Gaussian’, Gaussian Inc., Pittsburgh, 1993.
[19] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, ‘Numerical Recipes. The Art of Scientific Computing’, Cambridge University Press, Cambridge, 1986, p. 509.
[20] M.A. Curtis, J. Chem. Educ. 1994, 71, 775.