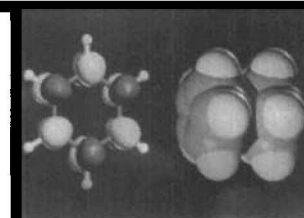


COMPUTATIONAL CHEMISTRY COLUMN

Column Editors:
Prof. Dr. J. Weber, University of Geneva
PD Dr. H. Huber, University of Basel
Dr. H. P. Weber, Sandoz AG, Basel



Chimia 46 (1992) 224–227
© Neue Schweizerische Chemische Gesellschaft
ISSN 0009–4293

Should a chemist be able to do his own programming?

Not only has the performance of computers changed drastically during the last two decades, but also the fields in chemistry where they are used, and the way they are used are different from twenty years ago. Then, the computer was mainly a number crunching machine for specialists, mainly physical chemists, who wrote their own FORTRAN code for a given application. At that time, nobody required that all chemists should be able to program a computer. The situation changed drastically about a decade ago when personal computing brought computer power into every chemist's reach. Then it was felt that practically all chemistry students should have some training in computer science and possibly learn one of the programming languages such as PASCAL or FORTRAN. Examining the field today shows that the situation has changed again, more than we expected ten years ago. It is no longer so that every computer comes with an operating system and a FORTRAN compiler, and that chemists have to write FORTRAN programs. Now the computers come with an increasing amount of application-software ready for use for statistics, data collection, quantum chemical calculations, *etc.* Is it, therefore, still necessary that each chemist should learn a programming language, or should an overview and specialized training in existing application-software be sufficient?

We think this question is of some importance for the universities as the chemistry curriculum is becoming loaded with more and more non-chemical disciplines and we have to concentrate on the most important things. To start a discussion, we have asked about 25 people from different fields of chemistry (inorganic, organic, and physical as well as from industry and

universities) for their opinion. To make it easier, we added the following five questions, which could be used to structure the answer:

- 1) How do I use the computer, did I ever use it for programming, or would I do programming, if I knew how?
- 2) What percentage of chemists in my institution do their own programming (or would, if they could)?
- 3) Did this percentage change greatly in the last years and would you assume that it will further increase or decrease in the next years?
- 4) Do I need any knowledge about computers at all?
- 5) What do I need to know about computers, if not programming?

We present below the answers we received and give here a short comment. About half of the addresses answered our survey and they are representative of different fields of chemistry. All agreed that young chemists should have what might be called 'computer literacy', *i.e.* a general knowledge of 'what a computer is' and 'what a computer does', including the ability to work with programs like word-processors, graphics packages, spreadsheets *etc.* Most answers also agree that some chemists (mostly in analytical and physical chemistry) should have knowledge of a computer language, *e.g.* to make modifications or adaptations of existing software, which means that universities should at least give chemistry students the opportunity to learn programming in their curriculum. To what extent this should be mandatory to all students, or what kind of programming, classical languages or new 'meta-languages', should be taught, is evidently still open to controversy!

We would like to thank all participants of the survey and invite our readers to send us their opinions. If we receive enough additional comments of general interest we will publish or summarize them in a future column.

Answers to the Survey

Yes indeed and, though the language is not of major importance, there are two main reasons why:

The chemist must be able to analyze and to evaluate physical data and to represent these data graphically. To do this he needs, as a minimum, something like a PC spreadsheet program (Excel is a good example) and to use this intelligently he must understand some elements of programming such as the use of loops and the concept of conditional execution.

The second, and less satisfying, reason is one that regularly crops up in our laboratory. The chemist is often presented with data that his spreadsheet, or whatever, cannot understand. Who is then going to reformat the data to make it readable if not the chemist himself?

The last importance of computers in chemistry, particularly in physical chemistry, well merits a closer knowledge of their inner workings, but today this means more study of operating systems (*e.g.* DOS, VMS, UNIX), of interfaces like Windows, and of software packets to prepare reports (Word, Wordperfect, Excel, Lotus 1-2-3) rather than learning a specific programming language. If a language is to be learned then I would vote for C rather than FORTRAN or BASIC mainly on the grounds of portability and flexibility.

Dr. J.J. Daly, F. Hoffmann-La Roche AG, Department PRTP, Basel

– Im Bereich der organischen Chemie verwenden wir Programme wie GAUSSIAN90, MOPAC6.0 und MM2, um Konformationsanalysen von Radikalen und Übergangszuständen zu berechnen. Durch die konformative Vielfalt unserer Systeme erhalten wir eine Datenmenge die nicht mehr 'von Hand'

- verarbeitet werden kann. Somit wurde es nötig, dass wir am MM2-Programm kleine Änderungen vornehmen mussten oder kleine Fortran-Programme schreiben mussten, um mit der grossen Datenmenge beispielsweise eine *Boltzmann*-Verteilung zu berechnen.
- Seit drei Jahren gibt es genau einen von 25 Mitarbeitern, der solche Berechnungen durchführt. Mit dem steigenden Bedarf solcher Berechnungen werden in Zukunft sicher mehr Mitarbeiter Rechnungen durchführen.
 - Mitarbeiter, die Moleküle berechnen, sollten mindestens kleine Fortran-Programme schreiben können. Andere, die ausschliesslich synthetisch arbeiten, sollten sich mit einer Textverarbeitung, einem Zeichnungsprogramm und einer Tabellenkalkulation auskennen. Es wäre wünschenswert, wenn Chemiker wissen, welche Rechenverfahren für welche Probleme geeignet sind.

Prof. Dr. B. Giese and W. Damm, Institut für organische Chemie der Universität Basel

There is no general answer to this question. The main reason is, that 'the chemist' as a general being does not exist. Besides the classical division of chemists in organic, inorganic, and physical ones there exists another way they can be classified: there are the synthetically oriented and the more analytically oriented chemists. Now coming to the question asked above, to my experience the chemists working on synthesis answer NO and the chemists doing more analytical work answer YES.

Doing 'his own programming' means firstly writing quick and dirty routines for data processing (like fitting routines) using BASIC on a PC or FORTRAN on a workstation or mainframe computer, and secondly being able to use scientific program packages which are very often far away from being user-friendly.

Besides the 'real own programming' every chemist (including the syntheticist) should be able to work with computers using commercial software packages like text processors, spreadsheets and data banks. Therefore, a basic information in 'how to use a computer' seems to me imperative for a chemist. Furthermore, for those chemists who need to do programming, there should exist a possibility to learn a programming language, FORTRAN by preference, from chemists or physicists and not from people working in computer science.

Dr. L. Helm, Group of Prof. Merbach, Institut de chimie minérale et analytique, Université de Lausanne

- 1) In my physical-chemical research group computers are used for:
 - a) Computer-assisted experiments
 - data collection
 - controlling of experiments
 - b) Numerical quantum chemistry and simulations.
 - c) Standard applications
 - Text
 - Graphics
 - Data bases
- 2) All my students, post-docs and assistants do their own programming (100%)
- 3) Status (2) exists since ~ 1985
- 4) What a question 1992!!
- 5) - Knowledge of the principles of algorithms is often necessary, more so than special training in programming.
 - To control experiments by computer, some knowledge of the hardware is desirable.

Prof. Dr. J.R. Huber, Physikalisch-Chemisches Institut der Universität Zürich

The question whether chemists should do their own programming relates to their use of mathematics in a very general sense. Historically, every exact science moves in the direction of increasing mathematical penetration and abstraction. Hence, every chemist who uses mathematics creatively must necessarily use a 'programming language'.

In 1992, a programming language does not only mean Pascal, FORTRAN, C, Modula-2, etc., but any of these or a meta-language, such as Mathematica, Derive, MathCad, database languages, graphics toolboxes, simulation languages, kinetics toolboxes, statistics toolboxes, etc. The percentage of chemists who program in this more general sense is certainly increasing rapidly. Their efficiency and productivity is much enhanced by the use of meta-languages. The gain in human efficiency and the increase in percentage of people doing 'programming' is comparable to that of the first transition from the machine-language-programming era to the assembler-programming era, or the second transition from the assembler-programming era to the language-programming era. Today we are witnessing the third transition of this kind. The great challenge and opportunity for chemist-programmers is to provide powerful meta-languages for specific purposes in pure and applied chemistry.

Prof. Dr. S. Leutwyler, Institut für anorganische, analytische und physikalische Chemie der Universität Bern

- 1) Ich selbst brauchte und brauche sehr viel Computer: PC's als Textsystem, für Tabellenkalkulation, Statistische Programme, Plotprogramme. IBM und DEC Hosts: Datenmanagement und -Pfleger in zwei grossen Corporate Produkte-Datenbanken. Meine eigenen Programmierkenntnisse (Fortran, Basic) brauche ich seit Jahren nicht mehr. Für Programmierarbeiten hat unsere Grossfirma jede Menge Spezialisten.
- 2) Ich würde unterscheiden: Synthetische Chemiker brauchen nicht selbst programmieren zu können, Physikochemiker und Instrumentalchemiker, die selbst Messsysteme entwickeln, und Chemiker in der chemischen Entwicklung sollten eine entsprechende Sprache können. Der Prozentsatz ist schwierig abzuschätzen: vielleicht 20%, bei denen Programmieren vorteilhaft wäre.
- 3) Der Prozentsatz hat sich wahrscheinlich bei diesen 20% bereits stabilisiert.
- 4) 'Computer Literacy' wird heute selbstverständlich bei jungen Chemikern und Physikochemikern vorausgesetzt.
- 5) Für alle: Kenntnisse im Umgang mit PC's und den häufigsten Anwendungsprogrammen. Für Chemiker: schnelles Zurechtfinden in Datenbasen auf Grossrechnern, Anwendung von Molecular Modelling und Synthesis Planning. Für Produktionschemiker: Kenntnisse über Prozessrechner. Für Physikochemiker: etwas über Gerätesteuerung.

Dr. P. Moser, Leiter der chem. tech. Information im Konzernbereich Sicherheit und Umwelt, Ciba-Geigy AG, Basel

Ein moderner Organiker soll das Programm-Angebot optimal ausnützen können. Es ist jedoch keineswegs notwendig, dass er selbst auch Programme schreiben kann.

Prof. M. Neuenschwander, Institut für organische Chemie der Universität Bern

In my organisation, I am concerned with pushing the use of EDP methods in chemical R+D. I am a chemist by training, but worked in the computing field for almost all of my adult life. I will not answer the question from my own perspective (the answer would be fairly obvious), but on behalf of 'the industrial chemist' who does not work in an EDP department.

The industrial chemist *must be computer-literate*. There is no question whatsoever that the computer is an essential

tool for every professional in chemical industry, and that its importance will increase. Whether computer literacy should include the ability to do 'own programming' depends on the definition of the term 'programming' and also on what kind of field the chemist is working in.

There are specialist fields (like crystallography, statistics, chemical information systems, molecular modeling) for which the scientist does need programming in the narrow sense (*i.e.* FORTRAN, C, PASCAL, *etc.*). Even though most of the useful work in these areas is now firmly based on commercial software, the specialist user and the scientific support staff needs to be able to bridge the inevitable gaps, tie things together, and work around the bugs. This still requires knowing how to program in languages like FORTRAN and C, as well as in the various command languages (DCL, UNIX shells). However, the heroic days when industrial chemists developed large scientific programs are over.

The vast majority of industrial chemists is not in this specialist category. They need to understand EDP in general, because they have to talk to their EDP specialists and they should then be able to judge what kind of work can be profitably automated, replaced, or transformed by computers. On a technical level, they have to be able to move around in computer networks, and they should feel comfortable with e-mail, database query languages and graphical user interfaces. They must also learn how to use the main types of desk-top software (text systems, spreadsheets, presentation graphics, personal databases, communication tools) as well as scientific software for information retrieval, data analysis, and modeling. All of this does not involve any 'programming' beyond setting up your own personal database or writing a script for a statistics tool.

At present, about 85% of the chemists at Ciba-Geigy know how to use a computer. The vast majority knows about PC's or Macintoshes, and many are familiar with Windows; VMS and UNIX come in as a distant second, and almost nobody claims to know about the classical mainframe and MVS. About 50% of the chemists claim to know a programming language (BASIC, followed by FORTRAN and then PASCAL). I don't think that many chemists use their programming knowledge for any serious professional purpose, because designing and writing software for professional applications has become too difficult and expensive. I suspect that the number of chemists who know how to

program in one of the classical languages will increase in the next years, but in chemical industry most of them won't make much direct use of this ability.

In general, I believe that learning to program in a structured language such as PASCAL is a valuable intellectual experience. Like algebra, it should be part of the secondary school curriculum; university is definitely too late. But this has nothing to do with science in general or chemistry in particular.

If I had to design a 'programming' course for chemistry students, I would replace it by a *mandatory* 'General, Scientific and Business EDP' course in order to get the knowledge up to a uniform level, and offer an *optional* algorithm and software design course for those who want to move into specialized fields. We don't need an army of chemists who do their own programming, but chemists who know how to use computers to get their real job done, and *actually do it*. We also need more good software of industrial relevance, but that is a different issue.

Dr. P. Opplinger, Corporate unit Information Technology, Ciba-Geigy AG, Basel

- 1) I use the computers for
 - i) Number crunching (quantum mechanics)
 - ii) Applications related to my research, involving programming (Fortran)
 - iii) Modifications of existing programmes
- 2) There are many users in our group
- 3) ??
- 4) Yes, required knowledge: Fortran, UNIX, (VMS)
- 5) Word processing, drawings.

F. Rotzinger, Group of Prof. M. Grätzel, Institut de Chimie Physique II, EPFL

Computer werden bei uns vorwiegend für folgende Zwecke verwendet:

- Textverarbeitung
- Zeichnen (Formeln, Schemata, Apparaturen)
- Visualisierung von Molekülstrukturen
- Datenbanken (Cambridge-File, CAS, Reaktionsdatenbanken, Lagerverwaltung)

Dafür stehen genügend fertige Programme zur Verfügung, die keine weiteren Programmierkenntnisse erfordern. Notwendig erscheinen dagegen folgende Fähigkeiten:

- Das Verständnis, wie man in einer Datenbank sucht, damit man seine Fragen an die Spezialisten gezielt formulieren kann.
- Die Kenntnis von wenigstens einem Mainframe-Betriebssystem (Unix, VMS) in seinen Grundzügen (z.B. File-Verwaltung), damit man die Programme auf Grossrechnern und Workstations auch benutzen kann.
- Die Fähigkeit, Anleitungen und Handbücher zu Computern und Programmen zu verstehen, denn die Features heutiger Programme sind zwar meistens gut dokumentiert, aber selbst auf PCs sind die Programme (z.B. MS Word) so komplex, dass sich diese Möglichkeiten nicht von selbst zeigen.

Weitergehende Kenntnisse sollte man nur als Spezialwissen lernen, vergleichbar etwa Kenntnissen in Chromatographie: jeder muss wissen, wie eine Chromatographie geht, aber nicht jeder muss gleichzeitig GC, HPLC und GPC beherrschen.

Dies setzt allerdings voraus, dass für alle weitergehenden Probleme ein Spezialist verfügbar ist, sei es als Mitarbeiter in der Arbeitsgruppe oder festangestellt im Institut. Letzterer muss dabei auch Chemiker sein, um die Probleme korrekt umsetzen zu können. In unserem Institut sind wir in dieser Lage mit zwei Computerfachleuten, sowie verschiedenen Mitarbeitern der Arbeitsgruppe, von denen einer auch gelegentlich Programme schreibt resp. auswärtige Programme anpasst.

Prof. Dr. D. Seebach und B. Lamatsch, Laboratorium für organische Chemie, ETH-Zürich

Every chemist should know programming and the use of computers! All our students in chemistry in Lausanne are trained in programming and in the use of computers. It should not change.

Prof. P. Vogel, Institut de chimie organique, Université de Lausanne

Not every 'chemist' has to be a programmer.

There are various types of chemists, including *e.g.* analytical, organic, physicochemical and theoretical chemists. In each discipline typical skills are required. Indeed the daily work in each of these chemistry branches today includes increasing interaction with computers. Therefore, eve-

ry chemist should be trained in the basics of computer technology, including topics such as terminals, PCs, workstations, networks, databases, operating systems, data exchange protocols. Many independent software packages are available, *e.g.* for data acquisition and manipulation, molecular modeling, text editing, *etc.* To a large extent these can be used without any knowledge of programming. However, development of new innovative programs of course requires modern programming knowhow. Furthermore, some chemists might be involved in the integration of various programs routinely used in a laboratory. There again programming is involved. The synergy between chemistry and programming skills will remain for a long time a positive factor in a well-functioning laboratory. In part these programming efforts may come from a trained chemist, while also the interaction with a computer department should be stimulated. In my own experience I have seen that for *ca.* ten years, most even simple applications had to be written in-house, while at present all can be bought. These trend will still increase, and certain 'dreams' will become available as power-

ful software, made by specialized vendors, which in turn indeed often work with programming chemists. More than programming skills, in the future a rapid adaptation to new hard- and software products will be asked from the 'chemist'.

Dr. H. van de Waterbeemd, Structure-Property Correlations Group, F. Hoffmann-La Roche AG, Basel

1) I had my first programming course at the ETH-Zürich 1964 'ALGOL'. After finishing this course I did very little programming in ALGOL and finally gave up completely. Later I acquired knowledge of BASIC by myself. I did a few very simple programs myself and then finally gave up completely in programming, because I still was interested in chemistry. Since *ca.* ten years, I'm now using personal computers, first MS-DOS and now *Macintosh*. I use it a lot, but I do not need to do my own programming. For some special problems I have asked specialists to

give me a hand to do some programming.

- 2) 10% of the chemists in our institute do their own programming.
- 3) No, this percentage did not change a lot in the last years, but the difference is that now everybody uses computers, before only 10%, those who knew programming, used computers.
- 4) Yes, I think knowledge about computers is essential for every chemist today.
- 5) For 90% of the chemists today it suffices to know computers in the same way as they know how to drive their car. For them it is neither necessary to do their own programming, nor to make a tune up of their car themselves. The ability to use computers with existing programs is nowadays sufficient for most chemistry. Yet, there should be a possibility for those who want to go beyond that, to learn programming in a serious way.

Prof. Dr. A. von Zelewsky, Institut de chimie inorganique et analytique, Université de Fribourg

Chimia 46 (1992) 227-228
© Neue Schweizerische Chemische Gesellschaft
ISSN 0009-4293

Comparison of the Performance of a Program for Molecular Dynamics Simulations of Liquids on Different Computers

Rolf Eggenberger and Hanspeter Huber*

In a previous 'Computational Chemistry Column' [1] *Th. Bally et al.* carried out a comparison of the performance of two quantum chemistry programs on different computers. Such comparisons are helpful in several situations, *e.g.* if you plan to purchase a computer, if you ask for a grant of computer time on a national supercom-

puter or if you have a choice to run different programs on different computers. The previous comparison [1] showed for example that for quantum chemical calculations workstations are in general more cost-efficient than supercomputers. This situation might differ for another kind of problem. Therefore, we present here some results concerning the performance of a simulation program running on several machines including the new national supercomputer in Manno, the NEC SX-3.

Together with quantum chemistry calculations, simulations are undoubtedly the

major type of number-crunching applications in chemistry. For the nonspecialist, we present here a short description of the main features of a molecular dynamics simulation program for liquids. Imagine a number of, *e.g.* 500, molecules each set at a random position in a box with a velocity vector also chosen randomly. The lengths of the vectors are scaled according to the temperature. Now go ahead a certain time step, let's say one femtosecond. That means you have to move all particles along their velocity vectors by a displacement corresponding to the femtosecond. In addition, you change the velocity vectors according to the time-step and the corresponding acceleration vectors. The latter are obtained from the forces acting on the particles, which in turn are calculated for each particle from a sum over the forces due to all other particles. To evaluate these forces in the two-body approximation you have to supply the two-particle potential energy curve. Although the real program is slightly more complicated, its structure is rather simple compared to a quantum chemistry program. You loop over, *e.g.* 100 000, time-steps and the time-consuming part is the calculation of the forces at each step since you sum over pairs of particles, which number $124\,750$ ($n \cdot (n-1)/2$) in the above example. The computer specialist will immediately recognize that such a program is suited to vectorization.

*Correspondence: PD Dr. H. Huber
Institut für Physikalische Chemie der
Universität Basel
Klingelbergstrasse 80
CH-4056 Basel