Computational Studies

DOI: 10.1002/anie.200801206

Predicting Molecules—More Realism, Please!**

Roald Hoffmann, Paul von Ragué Schleyer,* and Henry F. Schaefer III**

hypothetical molecules · quantum chemistry · theoretical chemistry · thermodynamics

The body of computations of molecules for which there is as yet no experimental evidence is growing very rapidly. This is simply wonderful—as a marker of the reliability of theory, and, sociologically, in creating a tense and fruitful balance between theory and synthesis in chemistry. Claims of "stability," implicit and explicit, are made for the calculated molecules; we have been as guilty of this as others. We would like to suggest that literature reports of these claims be qualified, and that the computations performed be described in a circumspect way.

Stable, Unstable

Let's talk about "stability." There is thermodynamic stability, governed by free energy changes, ΔG. And there is kinetic persistence, measured by a rate constant (the lifetime inversely related to it), in turn governed by a preexponential factor, concentrations, and an activation energy, E_a. There are also the imprecise but time-honored words "stable" and "unstable," with different shades of meaning in various scientific communities.[1] Oversimplifying, in colloquial use in physics "stability" is 95 % thermodynamic, 5 % kinetic. In chemistry, it's just the opposite. All the organic molecules in our body (small molecules or ions of the type of carbonate excepted) are thermodynamically unstable in the presence of oxygen.

In our minds, and in everyday discourse, stability in the thermodynamic sense merges with stability in the kinetic persistence sense. Thermodynamically, the term "stable" might be applied only to the global minimum, and "metastable" to all other minima, even those with large barriers surrounding them. But try to get jewelers to call their diamonds metastable. Or to get a chemist to refer to *n*-butane similarly!

Kinetic stability is also a matter of the relevant conditions. Some people are faster than others (or their instruments are, or they have more research funds…), and the lifetime of a species is a function of temperature, pressure, and concentrations. At 4 K in a He matrix, or in high vacuum, activation energies of only 1 kcal mol^{-1} or even less suffice to make a species detectable. At room temperature, an activation energy of approximately 15 kcal mol^{-1} is needed to ensure a half-life of a day;[2] inside our sun there are no molecules. In the interior of Jupiter, hydrogen is not your normal H_2; it is very likely a superconducting, compressed nonmolecular form.

The matter may be brought into focus by considering all possible simple homonuclear diatomic molecules. While N_2 is the most stable (thermodynamically, with respect to atomization), the second most stable homonuclear diatomic molecule in the second period, C_2, is not found filling any bottle at 1 atm and 293 K; it polymerizes like a shot, of course. Likewise, the most stable diatomic molecule of the third period, P_2, is not kinetically persistent under ambient conditions. Neither, for that matter, are Al_2, Si_2, S_2, some of which have substantially larger atomization energies than molecules we do have tanks of, F_2 and even Cl_2. But there is no trouble at all studying P_2 or C_2 at leisure in a matrix or a moderately high vacuum. For instance, 13 excited electronic states of C_2 are known intimately from just such studies.

Viable versus Fleeting

We would like to suggest (tentatively) that the theoretical literature distinguish between two descriptions of predicted hypothetical molecules by the common English words "viable" and "fleeting". Of course, these distinctions represent extremes. And we put aside, only for the moment, another essential matter, the quality of the computation.

"Viable" might be a label attached to a molecule that meets computational criteria of persistence appropriate to ambient conditions in a typical chemical laboratory environment, namely its isolation in condensed phase, near 1 atm pressure at room temperature (perhaps in the presence of a moderately humid atmosphere), or in reasonable concentration (say 0.001 M) in solution. And, with that, a half-life of a day or longer. We know that objections can be raised

---

[*] Prof. Dr. R. Hoffmann
Department of Chemistry and Chemical Biology
Cornell University, Baker Laboratory
Ithaca, NY 14853 (USA)
Fax: (+1) 607-255-3419
E-mail: rh34@cornell.edu
Prof. Dr. P. v. R. Schleyer
Center for Computational Chemistry
Department of Chemistry
University of Georgia
Athens, GA 30602 (USA)
Fax: (+1) 706-542-0406
E-mail: schleyer@chem.uga.edu
Prof. Dr. H. F. Schaefer III
Department of Chemistry
University of Georgia
Athens, GA 30602 (USA)
Fax: (+1) 706-542-0406
E-mail: sch@uga.edu

[**] We are grateful to R. Bruce King and Garnet K.-L. Chan for helpful discussions of the problems addressed in this essay, and Martyn Poliakoff and Gene Garfield for providing a quotation. We thank the National Science Foundation for its continued support of our explorations of the richness of molecular structure.

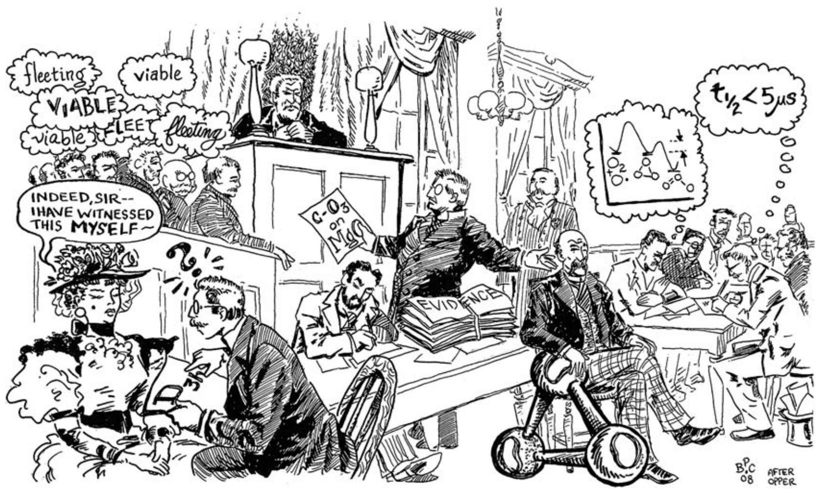

Figure 1. Cyclic ozone: A controversial molecule. Cartoon by Brian Coppola.

to any of these defining conditions, and, as noted above, that what is hard for one laboratory is trivial for another (Figure 1). We are going for the realistic median of what we perceive good preparative organic and inorganic chemists do.

Computational studies claiming viability of a molecule M that has not yet been made should report not only the obligatory vibrational analysis demonstrating that all the frequencies are real, but also satisfy the following criteria:

a) M should be resistant to fragmentation, isomerization, and dimerization or higher chemical aggregation (if the latter be more exothermic than normal van der Waals aggregation). This can be probed by identifying and computing activation energies for such processes, as well as by molecular dynamics simulations.
b) Not only should M have no imaginary frequencies, but the computed smallest vibrational frequency should be reasonably large, $100\,\text{cm}^{-1}$ or more for smaller molecules. There are obvious exceptions—some molecules are quite "floppy" or have, for example, methyl, cyclopentadienyl, or other functional groups with inherently low rotation barriers.
c) There usually should be an appreciable HOMO–LUMO separation in M (with all the caveats about how one computes that gap). Again there are obvious exceptions, such as sterically protected diradicals. A connected criterion, stability toward second-order Jahn–Teller type distortions, would be the presence of an appreciable energy separation in M between the ground and the first excited electronic state. Explicit computations demonstrating wavefuction stability should be carried out. We note the gap criterion is clearly biased toward organic molecules and ions; many inorganic extended structures are metallic, and had better show no band gap.
d) If M carries a charge higher than ± 1, realistic counterions should be included in the computations. Almost all dianions lose electrons spontaneously, while the electron affinities of dications (or more highly charged ions) are so great that they will react with almost anything (even He sometimes). And they are prone to proton loss or fragmentation.
e) Though chemists can keep air and water away, viability under ambient atmospheric condition requires significant barriers to reactions with H_2O, $^3O_2$, and N_2. Very few studies consider the reactions of predicted species with $^3O_2$, in particular.

"Fleeting" molecules might be claimed legitimately based on computations in which the sole energetic criterion is a vibrational analysis. That a molecule is a local minimum, with barriers, albeit small ones, preventing escape from its local basin, may still be of immense importance in interstellar chemistry, or more generally correspond to low-temperature matrix-isolation and high-vacuum environments.

It really is a matter of whom you want to reach. Papers addressed to the organic and inorganic synthetic communities should strive to push their computations through to the viable stage. Papers addressed to the physical chemical and chemical physics communities may be quite valuable at the fleeting level. The value of a paper is greatly enhanced if it addresses important bonding concepts in chemistry documented through the new molecules predicted.

The experimental communities are inherently skeptical of the claims of theoreticians. The skepticism, which we seek to diminish, is enhanced by claims of predictions of "stability" that are neither qualified nor circumspect. The information published is often only of the fleeting category, yet it implicitly or explicitly aspires to be received as being of the viable type. A degree of realism in what is calculated would help allay the skepticism of experimentalists. And such realism is also consistent with the strain of humility which characterizes any honest spiritual activity. Such as science.

Accuracy and Precision in the Quantum Chemical World

In quantitative experimental measurement, "accuracy" and "precision" are well-defined and essential to the trust the community places in experiment. Perforce, these descriptors must take on a different sense in the computer-based world of the quantum chemist. Let's think about they mean. We propose to use precision to assess the degree to which a particular computation approaches the exact result that should have been obtained with the specified method and basis set. Precision only means that you correctly accomplished what you claimed.

Most density functional theory computations reported in the chemical literature are not very precise in the sense given above. Although DFT total energies are routinely reported to five and often more decimal places, we would bet that only about half of all DFT computations to date have the correct total energies to 0.001 hartree, i.e., to 0.6 kcal mol^{-1}. As DFT methods incorporate heavier and heavier atoms, further deterioration can be expected, with numerical integration methods being severely challenged by the need to describe the large changes in the inner shell orbitals near the nucleus.

Different DFT software packages claiming to utilize the same functional and the same basis set may also give different absolute numbers. Sometimes, these vary only slightly, but other times grossly. If you don't believe it, try it. How could that be? Well, the commonly used "first principle" programs have internalized many operations in order to make them user-friendly. And depending on the programmers, the software code may make use of slightly different assumptions. For instance, the extremely popular B3LYP functional is implemented in the Gaussian programs differently than Becke intended.[3] Variations in the choices of the integration grid size also are sources of considerable numerical irreproducibility in DFT.

We appreciate that wave-function-based methods (Hartree–Fock, configuration interaction, coupled cluster) and perturbation theory (e.g., MP2) are capable of higher precision. With very tight settings for integral evaluation, self-consistent field, geometry convergence, and so forth, one can cautiously achieve six decimal places (0.000001 hartree) in the precision of total energies.

By accuracy, we mean absolute theoretical accuracy. Its discussion inevitably brings with it a dose of reality. For the helium atom, using wave-function methods, one can fairly easily compute the exact total energy to several more significant figures than can be determined in the laboratory. The H–H distance in H_2 can now be computed more accurately than it can be measured. However, by the time one reaches the fluorine atom, an accuracy in total energy of 0.001 hartree = 0.6 kcal mol^{-1} is presently unreachable. For the benzene molecule, an absolute accuracy of 0.01 hartree = 6 kcal mol^{-1} cannot be achieved currently, even with the most advanced coupled-cluster methods.

When DFT results are examined critically, total energies of medium-sized molecules are often in error by one full hartree (627 kcal mol^{-1}). For this reason, we recommend that the terms "accurate" or "accuracy" be used only very rarely in descriptions of molecular electronic structure theory.

Not to be unfair to theory, it need be mentioned that absolute energies are nontrivial to measure experimentally.[4] The higher ionization potentials of atoms are often not that accurately known. Also, lest we be thought too pessimistic, it is clear that energy differences—for, say, a reasonable change in an angular geometric parameter in a molecule—generally fare much better than total energies or relative state energy (singlet vs. triplet, for instance) in their sensitivity to basis set, functional, or even degree of inclusion of correlation.

Significant Figures in Theoretical Calculations

Another matter touching on the interaction of theory and experiment is the number of significant figures reported in the computational chemistry literature.

Experiment first: Ultimately, the number of significant figures in an experimental result is a mark of the quality of the measurement. Error analysis of an experiment typically leads to a standard deviation based on random variation. Systematic errors are another matter. These sometimes are well analyzed in a paper, but given the psychological realities of science, often are pointed out by others.

This is not to say that a precise measurement (with its associated small standard deviation) needs be taken as chemically significant. For instance, today standard deviations in X-ray crystallographic studies are routinely given as 0.001 Å for distances. Or better. But the hydrogen atoms, which may constitute over half the atoms in the molecules and may provide valuable insights into the electronic structure, are located only approximately. Equivalent heavy-atom bonds for two or more distinct molecules in the asymmetric unit (where the crystal creates such situations) may differ by 0.01 Å or even more. Clearly, crystal packing is responsible, and these larger variations provide a more realistic measure of a chemically significant uncertainty.

Being reasonable and realistic is the key to assessment of precision and accuracy. Here's what one of us says to students in his introductory chemistry course: "I drive to New York City often. I have a moderately unreliable odometer in my car (there is even a summer/winter difference). Supposing you ask me 'How many miles is it from Ithaca to New York?' If I answer, '200 miles,' that's not doing justice to what I can measure. And I'm not giving you as much information as you deserve. OK, what if I tell you it's '235.714 miles?'" The students laugh, of course. 230 or 235 miles is reasonable. Students get this right away.

What about significant figures in quantum chemical computations? Typical currently available hardware allows (depending on bit length and mathematical subroutines) 12–16 figures to be calculated reproducibly by the same program on the same computer. To report a distance or energy to so many figures would occasion as much derision as the too precise mileage mentioned above. Nevertheless, we see in the literature computed distances reported (in Å) routinely to four or more figures after the decimal point, and energies (in au, 1 au = 627.5 kcal mol^{-1}) to six or more figures after the decimal point. Given what we know of typical quantum chemical calculations, such extremes deserve to be called silly.[5] Peter B. Medawar said it in another way, in a parallel experimental context: "There is no surer indicator of scientific illiteracy than the quotation of numerical data to a degree of precision greater than the experimental observations warrant."

There are reasons for concern about the accuracy of almost any computed number in our trade. The literature is replete with reports of how changes in functionals, basis sets, or the way correlation is treated lead to vastly different conclusions regarding energetics, geometry, or properties. If we accept the resultant uncertainty, it behooves us to report results conservatively, with a large dose of humility—we simply don't know just how well a computed angle or a distance will stand up towards more refinement. Every set of theoretical approximations, every level of theory, generates its own chemistry. This is what John Pople taught us.

So what should be a reasonable number of decimals in a theoretical report? Pople advised five for energies (au), three for distances (Å), and two for angles (in degrees).[6] These numbers could be reproduced (within narrow limits) using the earlier versions of his Gaussian program running on different computer processors.

The qualifier "reasonable" is not solely a judgment of precision, nor of accuracy. It measures some balance of the two and takes into account chemical reality. It's like that distance to New York City. Although the computer gives you 1.24163598 Å for an equilibrium distance, reporting that number of decimals would show a complete lack of judgement. To say 1.2 Å is unfair to the authors of the computer program and probably will not reveal the difference between this distance and another in a related molecule. Although 1.24 Å probably is a reasonable compromise, Pople's third decimal might be added. Likewise, a computed activation energy reported as 40.269 kcal mol^{-1} is not sensible; 40.3 may be OK. It's pretty much common sense. As is science.

Concluding Thoughts

The purpose of a quantum chemical computation may only be semiquantitative, to ascertain how many π electrons a suspected aromatic system may have, to gain rough insight into the nature of the bonding of an unusual species, or to determine if a proposed reaction is exo- or endothermic. Many molecules are inherently interesting, even if they cannot be made. Planar tetracoordinate methane is a good example. What fun to think about these!

But the reporting of quantitative results carries with it special responsibility, and so does the implication that the chemical community should take notice of the prediction of an as yet unknown molecule. We discourage the use of the unqualified word "stable," except in colloquial usage. If "stable" is employed as a descriptive term, its intended meaning should be made explicit by an auxiliary phrase or context. The descriptors "viable" and "fleeting" that we suggest are hardly absolute; they just reflect extrema of a range of research programs, constrained in the real world by intended audience, resources, and effort. Claims of "viability" and "accuracy" should be well-informed, thoroughly founded, critical, circumspect, and conservative.

A beautiful chemical world, of molecules waiting to be made, opens up through the ingenuity of theoretical chemists. Their predictions astound us, generate ideas, and prompt the synthesis of new structures and functions. No exaggeration, none at all, is needed to build this world.

Received: March 12, 2008

---

[1] See, among others, R. Hoffmann, *Am. Sci.* **1987**, *75*, 619–621.

[2] For a bimolecular reaction, initial concentration 0.001 M, preexponential factor 10^5, a half-life of a day is associated with an E_a of 16.7 Kcal mol^{-1}.

[3] R. H. Hertwig, W. Koch, *Chem. Phys. Lett.* **1997**, *268*, 345–351.

[4] D. Feller, C. M. Boyle, E. R. Davidson, *J. Chem. Phys.* **1987**, *86*, 3424–3440; E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, C. F. Fischer, *Phys. Rev. A* **1991**, *44*, 7071–7083; A. Lüchow, J. B. Anderson, D. Feller, *J. Chem. Phys.* **1997**, *106*, 7706–7708.

[5] If reproducibility is a concern, one calculated number per paper might be given to as many figures as one likes.

[6] W. J. Hehre, L. Radom, P. von R. Schleyer, J. A. Pople, *Ab-inito Molecular Orbital Theory*, Wiley-Interscience, New York, **1986**.