

## Towards predictive models in chemistry

*J. Chandrasekhar*

In 1828, A. Quetelet suggested that the degree of perfection to which a science has arrived may be judged by the facility with which it may be submitted to calculation<sup>1</sup>. But not everyone seemed to agree. Around the same time, A. Comte wrote<sup>2</sup>: 'Every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational and contrary to the spirit of chemistry. If mathematical analysis should ever hold a prominent place in chemistry – an aberration which is happily almost impossible – it would occasion a rapid and widespread degeneration of that science'. Much has happened since those early days of chemistry, even as the argument went on (or more correctly, goes on!). The progress made in enabling a significant number of chemical questions to be 'submitted to calculations' has culminated in the Nobel Prize for Chemistry in 1998 being awarded to two key players in the field, Walter Kohn and John A. Pople. Pople has been specifically recognized 'for his development of computational methods in quantum chemistry'.

John Pople was born in 1925 in Burnham-on-Sea, Somerset, England. He studied mathematics at Cambridge and took a course in Quantum Mechanics offered by the great master, P. A. M. Dirac. Thus, Pople represents a direct link to one of the original developers of the mathematical framework needed to solve chemical problems. It is instructive to consider the status of quantum mechanical description of chemistry when Pople took up his graduate work under the guidance of John Lennard-Jones.

The situation was precisely summarized in the much (mis)quoted statement of Dirac: 'The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble'<sup>3</sup>. Exact quantum mechanical treatment was possible for the hydrogen atom and the hydrogen molecular ion, but for many-electron systems severe approximations were needed. Linus

Pauling seized on a method developed by Heitler and London for the hydrogen molecule and worked out a general theory of structure and bonding in chemical systems. By combining qualitative and semi-quantitative concepts derived from quantum mechanics with empirical ideas derived from experimental data, Pauling could explain many facets of chemistry. In particular, valence bond theory in association with quantitative parameters like electronegativity, atomic radii, etc. provided a successful framework for interpreting experimental observations. But quantitative predictions were hard to come by.

Lennard-Jones worked out an alternative description of bonding in diatomic systems which turned out to be the forerunner of the Molecular Orbital method. Erich Hückel extended the idea to conjugated systems, while Robert Mulliken derived a generalized description for polyatomic molecules. Several useful insights and predictions about ground and excited states of molecules could be made using qualitative Molecular Orbital theory. But, again, quantitative formulations did not look promising. While the Hartree-Fock procedure offered a way, numerical solutions could be obtained only for atoms, that too after a good deal of labour. Polyatomic systems with their reduced symmetry remained intractable. Both Lennard-Jones and Mulliken were acutely aware of the need to find a way out of the impasse and were working out strategies.

There was also a rather unsavoury argument going on about the relative merits of VB and MO theory, even as a qualitative model. In organic systems, with well-defined two-centre two-electron bonds, Heitler-London product wave functions derived from pairs of hybrid orbitals seemed to be the simplest way of describing structure and bonding. The use of MOs distributed over an entire molecule appeared cumbersome. But these orbitals had the advantage of reflecting the symmetry properties of the molecule, a concept exploited by Mulliken for explaining spectroscopic features, and for predicting feasibility of chemical reactions, much later, by R. B. Woodward and R. Hoffmann.

For his doctoral work, Pople pointed out some common ground between the two theories. Lennard-Jones and Pople showed that any number of equivalent orbitals can be constructed by transformations of canonical MOs, leaving the total wave function unchanged<sup>4</sup>. In particular, the directional character of bonds implicit in the VB model can be derived from suitable combinations of MOs. In a follow-up paper<sup>5</sup>, Pople derived the rabbit-ear representation of lone pairs in the water molecule from the  $\sigma$  and  $\pi$ -type lone pair MOs. Pople went on to convincingly demonstrate that the directionality of lone pairs had important implications. The displaced electron density was shown to be primarily responsible for the dipole moment of the molecule (and not just the polarized O-H bonds, as had been believed till then). The lone pairs were also the key to the formation of hydrogen bonds in liquid water and ice. Lennard-Jones and Pople worked out the basic rules governing molecular association in liquid water<sup>6</sup>. Pople then showed that many of the anomalous properties of water can be accurately predicted using a statistical mechanical theory which invokes the presence of a distribution of hydrogen bonded structures, with varying angles<sup>7</sup>. The 'bent hydrogen bond model' has been vindicated by more elaborate numerical simulations which became possible decades later.

An important breakthrough occurred around the time Pople completed his doctoral work. Mulliken's student C. C. J. Roothaan converted the integro-differential Hartree-Fock equations into a more manageable set of algebraic equations, by representing each MO as a linear combination of basis functions, such as atomic orbitals. Simultaneously and independently, G. G. Hall, working in the group of Lennard-Jones, derived the same set of equations. The Roothaan-Hall equations essentially require setting up of an energy matrix and subsequent diagonalization. The second step is ideally solved using digital computers, provided the matrix elements can be computed. But again there was a bottleneck. If atomic orbitals are chosen as the basis functions, a logical choice for chemists, the matrix

elements involve computation of multi-centre integrals, involving products of atomic orbitals (termed differential overlap). Analytic expressions for computing three and four-centred integrals involving Slater-type orbitals (commonly used to represent atomic orbitals) were just not available.

Several approximations were being tried out for getting around the problem of computing the more difficult integrals by Geoppert-Mayer, Sklar, Mulliken and others. One such was the zero differential overlap (ZDO) approximation in which the product of two different atomic orbitals is set to zero. R. Pariser and R. G. Parr worked out such a scheme, which was the first generally successful quantitative attempt to solve the Roothaan-Hall equations, although limited only to  $\pi$  systems<sup>8</sup>. The integrals which survived the ZDO approximation were partly computed using the uniformly charged sphere approximation and the rest parameterized. The resultant procedure was a quantitative theory, which went well beyond Hückel theory, by explicitly taking into account electron repulsions. Pariser and Parr used the procedure for predicting spectral properties of conjugated systems. Independently, on the other side of the Atlantic, Pople used the ZDO approximation to work out exactly the same computational strategy<sup>9</sup>. Pople used the method to understand and predict variations in bond lengths in aromatic systems. The procedure is now generally referred to as the PPP method, after Pariser, Parr and Pople. It is still widely used, especially for predicting excited state properties, with inclusion of configuration interaction.

Perhaps to avoid even a trace of controversy, Pople moved on to a distinctly different area, viz. magnetic properties of molecules, although he did make one important contribution to quantum chemistry in 1954. He suggested an extension to the Hartree-Fock method for handling open shell systems like radicals and triplet states<sup>10</sup>. In systems having different number of  $\alpha$  and  $\beta$  spin electrons, the exchange correlation will be different for the two spins. Hence, the spatial description for  $\alpha$  and  $\beta$  spin orbitals should be different. This generalized procedure is known as the Unrestricted Hartree-Fock (UHF) method. While the resultant wavefunction has a few flaws (such as not being an eigenfunction of

the  $\langle S^2 \rangle$  operator and being contaminated with functions with higher spin), the method also is characterized by several pleasing features. Many important and useful theorems applicable to the Hartree-Fock method hold good for open shell systems only within the UHF model. Further, a key electronic interaction known as spin polarization, which determines many spin and magnetic properties, is effectively captured by the UHF wave function. Naturally, Pople was quite fond of the method and his school rarely used the Restricted Hartree-Fock or other variants for open shell systems.

Pople spent several years in the fifties working on theoretical problems associated with the newly developed spectral tool, Nuclear Magnetic Resonance. He was fascinated as much by the relation between spectral characteristics and model spin hamiltonians as by the chemical information contained in the spectral parameters, such as chemical shifts and nuclear spin coupling constants. He derived atomic and multi-centre contributions to magnetic properties using perturbation theory in conjunction with semi-quantitative MO calculations invoking the ZDO approximation. Pople's intense association with magnetic resonance culminated in a classic book, co-authored with Schneider and Bernstein<sup>11</sup>.

Pople moved to Carnegie-Mellon University, Pittsburgh and became a professor in 1964. A decade had gone by since his first attempt at developing a quantitative MO procedure for  $\pi$  systems. He had assumed that the theory would have been generalized to include the  $\sigma$  framework as well, enabling the study of any polyatomic system. To his surprise, not enough progress had been made in the intervening period. Of course, many groups had tried to generalize the ZDO approximation to all-valence calculations. But the results were far from satisfactory. Pople took a fresh look himself and came up with remarkable insights. He showed that the magnitude of errors resulting from the ZDO approximation depends, among other things, on the coordinate system chosen. Thus, the results do not have some essential and some desirable invariances. Pople went on to suggest a simple solution<sup>12</sup>. To obtain results which are invariant, additional approximations had to be made. This, in effect, led to further

simplification of the computational problem. Pople and his group came up with a semi-empirical MO procedure called the Complete Neglect of Differential Overlap (CNDO) method, which incorporated all these features. With a minor modification to the parameterization scheme, the first semi-quantitative all-valence computational procedure, CNDO/2, was released. Soon afterwards, one-centre integrals were included without ZDO approximation. The resulting Intermediate Neglect of Differential Overlap (INDO) theory was as easy to use as the CNDO method, but included contributions which are essential for predicting nuclear spin coupling and electron-nuclear hyperfine coupling constants accurately.

Although semi-empirical methods seemed the most promising approach to examine chemical problems quantitatively, especially for large molecules of interest to experimentalists, a large group of influential theoretical chemists were not fully convinced. In a perceptive analysis, C. A. Coulson spoke about the polarization of theoretical chemists into two groups<sup>13</sup>. Group I was concerned with accurate calculations involving extensive computations on small molecules. Group II preferred using highly approximate and partly empirical calculations on large molecules. Since the two groups were pursuing different agendas, the divergence of views between them kept increasing. With a wider vision, Pople pointed out how the two groups could cooperate for the general advancement of computational chemistry. At a symposium in honour of Mulliken in 1965, Pople proposed a concrete strategy: 'Group I should be encouraged to extend some of their simple calculations (such as minimal-basis set LCAO-SCF) to larger molecules, even though we all recognize the limitations. It would at least be valuable to get more experience of these limitations, particularly for series of homologous or related molecules, so that the nature of semiempirical corrections introduced by Group II can be better understood. At the same time, Group II should keep their methods continuously under review, always looking for ways in which they can be tested against more precise calculations on smaller systems'<sup>14</sup>.

Pople practised what he preached. He tried to systematically review and improve the performance of semi-empirical

methods. He allowed a large group of theoretical chemists the world over to join the efforts through another best-seller<sup>15</sup>, co-authored with D. L. Beveridge. The book had a listing of the Fortran code for performing CNDO and INDO calculations. The methods became big hits. A variety of problems, ranging from structure and energetics of molecules and reactive intermediates, like ions and free radicals, to conformational effects and intermolecular interactions, including hydrogen bonding, were studied. Of course, not all predictions were correct. But the models spurred theoreticians to consider improvements and generated interest amongst experimentalists to pose new questions. M. J. S. Dewar, in particular, modified the parameterization schemes drastically and derived a number of procedures which are popular with chemists even now, especially for studying reactivity in complex substrates.

Pople then moved on to join Group I who preferred to do more rigorous calculations and went about it in a systematic manner. He began to go for full-fledged solutions to the Roothaan–Hall equations, using a set of gaussian functions for describing each atomic orbital. Such calculations, which do not use any experimental information other than the fundamental constants, like the charge and mass of the electron, Planck's constant, etc. are referred to as *ab initio* procedures. There were indeed many groups which had developed *ab initio* code, but the general impression was that it required enormous amount of memory, disk space and time to do the calculations. Pople and his group wrote more efficient codes. A key simplification made by Pople was to use the shell concept. The *s* and *p* orbitals on an atom were forced to share the same orbital exponents. While the purists felt that it was an unnecessary restriction, Pople showed that the effort needed to compute integrals was drastically reduced. Attention to details, such as reducing DO LOOPS and writing out the expressions explicitly, made all the difference. The program called Gaussian 70 program was fast and user-friendly.

In addition to developing the program, Pople also focused on the nature of the basis functions used in *ab initio* calculations. The results depend critically on the choice of these functions. There were different prescriptions. It was common to find different groups, sometimes even the

same groups, using different basis functions. While qualitatively the trends were similar, the absolute values of energies changed a lot. As a result, calculations carried out on different chemical systems could not be compared on a uniform basis. For example, high level calculations to determine the equilibrium geometries may have been reported on propene and its isomer cyclopropane. But because the basis sets used were not identical, it was not possible to answer other questions concerning the same two molecules, such as their relative energies. Pople preferred to create models which offer many internal comparisons. Therefore, he systematically developed several basis sets, varying in sophistication and flexibility. In particular, he worked out minimal, split-valence and polarized basis sets and tested them out on a number of standard chemical problems<sup>16–18</sup>. Pople's basis sets did not always give the lowest total energies (within each class), but the use of 'molecular scale factors' generally resulted in useful predictions of energy differences of importance in chemistry. The extensive calibration and the ease with which they could be used within the Gaussian series of programs made STO-3G, 4-31G, 6-31G\* and 6-31G\*\* basis sets the preferred choice for many computational chemists. Later, he added a smaller split-valence basis set<sup>19</sup>, 3-21G, which now represents the starting point for most *ab initio* calculations on large chemical systems. To those who wanted more complete basis sets, triple zeta plus diffuse and polarization functions were made available<sup>20</sup>.

Pople substantially enhanced the capabilities of the Gaussian program in successive revisions. Electron correlation could be included through Møller-Plesset perturbation theory to second, third and fourth orders<sup>21</sup>. Configuration interaction and alternative post-Hartree–Fock schemes were added. While the methodologies were often known for a long time, Pople's group worked out additional strategies whenever needed. Perhaps the most important feature of the Gaussian program was the inclusion of efficient code using analytic gradients for locating minima and, later, transition states on complex potential energy surfaces<sup>22</sup>. This was crucial for the study of chemical reactivity problems. In recent versions, many new capabilities have been added, such as magnetic properties. In the

nineties, the Gaussian program was extended to use density functional theory. The availability of new computational procedures, like the B3LYP method, which is a hybrid Hartree–Fock and DFT procedure, as yet another option within the Gaussian program enabled many computational chemists to check out their usefulness for numerous chemical applications. The methods are now very popular, especially for studying complex reactions involving organic and organometallic systems. In a significant way, Pople contributed to the general awareness among chemists of the power of DFT methods, which used to be confined to the domain of physicists and theoretical chemists for a long time.

Pople did not just develop programs, but applied them to a number of important chemical problems, such as the anomeric effect<sup>23</sup>. Many chemists were attracted to the power of theory through his contributions. He had a highly successful collaboration with P. v. R. Schleyer, with whom he published a series of important papers on the structure and energetics of organic molecules, carbocations, and organometallics, especially organolithium compounds<sup>24–26</sup>. As before, he authored a popular book summarizing the methodologies and applications of *ab initio* theory<sup>27</sup>.

After his formal retirement at Carnegie-Mellon, Pople was offered a position at Northwestern University where he continues to work. He has received many honours and awards, including the highest recognition from the American Chemical Society and the prestigious Wolf Prize in Chemistry. He became a Fellow of the Royal Society as early as in 1961. The Nobel Prize is clear recognition of a lifetime of achievements and leadership in the area of computational chemistry.

All along, he has shown great vision and never wavered from his goals. He has uncompromisingly insisted on certain minimum characteristics which theoretical models in chemistry ought to possess. His views have been influential in shaping the development of the field. He has been remarkably objective while assessing the reliability of his own models and has gently frowned on extravagant claims made by some overenthusiastic computational chemists. His sense of balance and the emphasis he placed on standardization and repeated calibration have been crucial in convincing many experimentalists

about the predictive power of quantum chemical methods. Thanks to his efforts, we can recall the fond wish of Gay-Lussac expressed over a century ago with a strong sense of optimism: 'We are perhaps not far removed from the time when we shall be able to submit the bulk of chemical phenomena to calculation'<sup>28</sup>.

As a person, Pople has all the characteristics we associate with an English professor: dignified, with measured speech and gestures, endowed with a subtle sense of humour, kind and affectionate, but without being too obvious about it. He could convey answers unambiguously with a slight twitch of his brows or a twinkle in his eyes. He has emotionally strong bonds to his country and to his *alma mater*, Trinity College. He remains a British subject. I am sure he was particularly pleased to share the stage at Stockholm with the Master of Trinity College, our own Amartya Sen.

1. Quetelet, A., *Instructions Populaires sur le Calcul des Probabilités*. Tarlier, Brussels, 1828.
2. Comte, A., *Philosophie Positive*, 1830.
3. Dirac, P. A. M., *Proc. R. Soc. London*, 1929, **123**, 714.

4. Lennard-Jones, J. and Pople, J. A., *Proc. R. Soc. London*, 1950, **A202**, 166.
5. Pople, J. A., *Proc. R. Soc. London*, 1950, **A202**, 323.
6. Lennard-Jones, J. and Pople, J. A., *Proc. R. Soc. London*, 1951, **A205**, 155.
7. Pople, J. A., *Proc. R. Soc. London*, 1951, **A205**, 163.
8. Pariser, R. and Parr, R. G., *J. Chem. Phys.*, 1953, **21**, 466.
9. Pople, J. A., *Trans. Faraday Soc.*, 1953, **49**, 1375.
10. Pople, J. A. and Nesbet, R. K., *J. Chem. Phys.*, 1954, **22**, 571.
11. Pople, J. A., Schneider, W. G. and Bernstein, H. J., *High-Resolution Nuclear Magnetic Resonance*, McGraw-Hill, New York, 1959.
12. Pople, J. A., Santry, D. P. and Segal, G. A., *J. Chem. Phys.*, 1965, **43**, S129.
13. Coulson, C. A., *Rev. Mod. Phys.*, 1960, **32**, 169.
14. Pople, J. A., *J. Chem. Phys.*, 1965, **43**, S229.
15. Pople, J. A. and Beveridge, D. L., *Approximate Molecular Orbital Theory*, McGraw-Hill, New York, 1970.
16. Hehre, W. J., Stewart, R. F. and Pople, J. A., *J. Chem. Phys.*, 1969, **51**, 2657.
17. Hehre, W. J., Ditchfield, R. and Pople, J. A., *J. Chem. Phys.*, 1971, **56**, 2257.
18. Hariharan, P. C. and Pople, J. A., *Chem. Phys. Lett.*, 1972, **66**, 217.
19. Binkley, J. S., Pople, J. A. and Hehre, W. J., *J. Am. Chem. Soc.*, 1980, **102**, 939.
20. Krishnan, R., Frisch, M. J. and Pople, J. A., *J. Chem. Phys.*, 1980, **72**, 4244.
21. Binkley, J. S. and Pople, J. A., *Int. J. Quantum Chem.*, 1975, **9**, 229.
22. Pople, J. A., Krishnan, R., Schlegel, H. B. and Binkley, J. S., *Int. J. Quantum Chem.*, 1979, **S13**, 225.
23. Jeffrey, G. A., Pople, J. A., Binkley, J. S. and Vishveshwara, S., *J. Am. Chem. Soc.*, 1978, **100**, 373.
24. Collins, J. B., Dill, J. D., Jemmis, E. D., Apeloig, Y., Schleyer, P. v. R. and Pople, J. A., *J. Am. Chem. Soc.*, 1976, **98**, 5419.
25. Chandrasekhar, J., Pople, J. A., Seeger, R., Seeger, U. and Schleyer, P. v. R., *J. Am. Chem. Soc.*, 1982, **104**, 3651.
26. Würthwein, E.-U., Sen, K. D., Pople, J. A. and Schleyer, P. v. R., *Inorg. Chem.*, 1983, **22**, 496.
27. Hehre, W. J., Radom, L., Schleyer, P. v. R. and Pople, J. A., *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.
28. Gay-Lussac, J. L., *Mem. Soc. d'Aroueil*, 1888, **2**, 207.

J. Chandrasekhar is in the Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012, India.

## Ammonia synthesized at atmospheric pressure

S. Sampath and K. S. Gandhi

Ammonia plays an important role in all economies, especially in agricultural economies, as it forms the basis for the manufacture of nitrogenous fertilizers. Surprisingly, despite its importance, it is still made by Haber's process developed in the early twentieth century. In this process,  $N_2$  and  $H_2$  are reacted at high pressures (50 to 300 atm) and temperatures (500°C). The severity of the reaction conditions contributes both to the cost and complexity of the process. Marnellos and Stoukides from the Chemical Engineering department of the Aristotle University, Greece, have recently announced (*Science*, 1998, **282**, 98) an electrochemical synthesis of  $NH_3$  from  $N_2$  and  $H_2$  at atmospheric pressure, which has the potential of being an alternative to Haber's process.

Chemical conversion of  $N_2$  and  $H_2$  into  $NH_3$  is limited by thermodynamic equi-

librium. As there is a decrease in the number of moles in the course of the reactions, Le Chatelier's principle would indicate that increased pressure favours the formation of  $NH_3$ . Reciprocating compressors were used in the early versions of the process to generate high pressures, but due to noise and wear and tear produced by the reciprocating movements, these made way in favour of turbocompressors. Turbocompressors are, however, uneconomical beyond 300 atm, which sets the upper limit on operating pressure.

The formation of  $NH_3$  is exothermic, and it can be inferred, once again from Le Chatelier's principle, that lower temperatures would enhance the conversion to  $NH_3$ . However, even with the iron-based catalyst, the rates of formation of  $NH_3$  fall to infeasible levels at temperatures below about 400°C.

At 450°C and 300 atm, about 53% of  $N_2$  in a stoichiometric mixture with  $H_2$  can be converted to  $NH_3$  provided chemical equilibrium can be achieved. If the pressure is lowered to 100 atm, the conversion falls to a low value of 25%. The performance in practice is much lower. The iron-based catalyst has to be made strong enough to withstand the high pressures as well as sudden changes in pressure, and this makes the catalyst less permeable to reactants. As a result, only about 10% to 15% yields of  $NH_3$  are achieved in industrial practice.

In contrast to this, the electrochemical synthesis developed by Marnellos and Stoukides achieves conversions of 78% of *electrochemically supplied* hydrogen (in the form of  $H^+$ ) at atmospheric pressure itself and at a temperature of 570°C. Despite the caveat of limitations of electrochemical supply of  $H^+$ , the