

PART I

CHARGE DISTRIBUTIONS

COPYRIGHTED MATERIAL

CHAPTER 1

INTRODUCTION

1.1 THE BOND ENERGY MODEL

This book is about electronic charge distributions, chemical bonds, bond energy additivity in organic molecules, and the description of their relevant thermochemical properties, such as the energy of atomization, the enthalpy of formation, and the like, using computer-friendly methods.

Additivity schemes with fixed bond energy (or enthalpy) parameters plus a host of corrective factors reflecting nonbonded steric interactions have a long history in the prediction of thermochemical properties, such as the classical enthalpy of formation of organic molecules. Allen-type methods, for example, nicely illustrate the usefulness of empirical bond additivity approaches [1,2].

But theory tells a different story.

Immutable bond energy terms tacitly imply never-changing internuclear distances between atoms whose electron populations would never change. But the point is that invariable local electron populations cannot describe a set of electroneutral molecules. Unless the net atomic charges of all atoms in all molecules always exactly equal zero, any additivity scheme postulating fixed atomic charges obviously violates all requirements of molecular electroneutrality; for example, if the same carbon net charges and the same hydrogen net charges ($\neq 0$) are assigned to the carbon and hydrogen atoms of methane and ethane, simple charge normalization indicates that molecular electroneutrality cannot be satisfied for both molecules. Finally, unless we stipulate that atomic charges have no bearing on bond energies or else, that

any change in atomic charge is perfectly counteracted by appropriate changes in internuclear distances in order to prevent changes of bond energy, we are led to the concept of bond energies depending on the charges of the bond-forming atoms.¹

But quantum chemistry *hic et nunc* does not know about chemical bonds, unless we say so. The approach chosen here is centered on the potentials at the nuclei found in a molecule.

The Hellmann–Feynman theorem tells us that all forces in a molecule can be understood on purely classical grounds, provided that the exact electron density (or at least a density derived from a wave function satisfying this theorem) is known for that molecule [3]. We have exploited this vein. The description of atomization energies, on the one hand, and that of bond energies, on the other, were reduced to purely electrostatic problems involving only nuclear–electronic and nuclear–nuclear interactions. There is a price to be paid for this simplification—the *mechanism* of bond formation cannot be understood in purely electrostatic terms [4,5]. The kinetic energy of the electrons plays a decisive role because the electronic Hamiltonian of an atom or a molecule is bounded from below only if the kinetic energy is duly accounted for. This decisive role deeply reflects the theory explaining why chemical bonds are formed in the first place [4,5]. In short, our electrostatic approach allows no inquiry into the origin of chemical bonds. In contrast, it is well suited for describing chemical bonds as they are found in molecules at equilibrium.

This is so because the Hellmann–Feynman theorem offers a most convenient way to bring out the main features of chemical binding. By taking the nuclear charges as parameters, the binding of each individual atom in a molecule can be defined without having recourse to an a priori real-space partitioning of that molecule into atomic subspaces. This binding is determined entirely by the potentials at the atomic nuclei. In short, our definition of bond energies does not involve virtual boundaries that delimit the space assigned to the individual atoms in a molecule with intent to subsequently describe the chemical bonds linking them. We know, of course, that powerful and realistic methods describe a useful partitioning of molecules into “atoms in a molecule” [6]—an approach that certainly offers much chemistry. The methods developed here simply represent another perspective of the same problem. Important arguments concern a real-space core–valence charge partitioning, a topic that resists the approach [6] leading to the concept of “atoms in the molecule.” Moreover, our approach involves the use of Gauss’ theorem and a sensible application of the Thomas–Fermi model [7,8], which is known to give reasonably accurate atomic energies with the use of Hartree–Fock densities [9–12].

However, direct calculations of accurate bond energies represent a major challenge. Examples are given [13,14] where the ratios of carbon–carbon bond energies, relative to that of ethane, were successfully calculated for ethylene, acetylene, benzene, and

¹Empirical bond additivity methods circumvent the problems linked to charge normalization constraints because they modify the genuine “bond energies” by tacit inclusion of extra terms that have nothing to do with bond energy itself. Still, a number of additional “steric factors” must be introduced in order to achieve what fixed bond energy terms alone cannot do: an agreement with experimental results (see Chapter 10).

TABLE 1.1. CC Bond Energies Relative to That of Ethane

Basis ^a	Ratio of CC Bond Energies			
	C ₂ H ₄	C ₂ H ₂	C ₆ H ₆	<i>c</i> -C ₃ H ₆
1 6–31 G(<i>d,p</i>)	1.84 ₆	2.82 ₈	1.59 ₀	1.00 ₂
2 6–311 G(<i>d,p</i>)	1.93 ₈	3.04 ₈	1.60 ₁	0.99 ₃
3 6–311 G(2 <i>df</i> ,2 <i>p</i>)	1.95 ₅	2.90 ₈	1.60 ₃	0.97 ₂
4 6–311 G(2 <i>df</i> ,2 <i>pd</i>)	1.97 ₂	2.96 ₀	1.62 ₁	0.97 ₇
5 vD(2 <i>d</i> ,2 <i>p</i>) BLYP	1.99 ₁	3.09 ₄	1.61 ₅	0.95 ₁
6 vD(2 <i>d</i> ,2 <i>p</i>) B3LYP	1.91 ₅	2.98 ₀	1.58 ₁	0.95 ₈
7 vD(2 <i>df</i> ,2 <i>pd</i>) BLYP	2.02 ₀	3.15 ₄	1.63 ₄	0.95 ₉
8 vD(2 <i>df</i> ,2 <i>pd</i>) B3LYP	1.94 ₃	3.03 ₆	1.59 ₇	0.96 ₂
Semiempirical	2.000		1.640	

^aThe basis set of **1** is from Ref. 15, that of **2** from Ref. 16, and those of **3** and **4** are from Ref. 17. In **5–8** we used van Dujneveldt's bases [18], also adding a set of *f* functions on C and of *d* functions on H [17] (**7,8**), Becke's gradient correction to exchange [19], and the Lee–Yang–Parr (LYP) potential [20]. The B3LYP functional involves a fully coherent implementation, whereby the self-consistent field (SCF) process, the optimized geometry, and the analytic second derivatives are computed with the complete density functional, including gradient corrections and exchange.

cyclopropane (Table 1.1). For ethane itself, the calculated CC bond energy amounts to ~70 kcal/mol from Hartree–Fock calculations [13] or 68.3 ± 0.4 kcal/mol, as given by density functional calculations [14]. The best fit with experimental data is obtained with 69.633 kcal/mol for that bond, which is satisfactory.

1.2 SCOPE

The difficulties encountered in the direct calculation of bond energies can be overcome—with hard labor and some approximations—in only a few cases, but the good news is that only a few reference bond energies need to be calculated for model systems. Those determined for the CC and CH bonds of ethane, for example, are sufficient for the description of saturated hydrocarbons; the addition of the reference bond energy describing the double bond of ethene extends the range of applications to olefinic molecules, including polyenic material. It is thus well worth the trouble to calculate a few reference bond energies—and this can be done with reasonable accuracy—because the rest follows as explained here. That is where atomic charges come into the picture and solve the problem presented here.

The description of bond energies that depend explicitly on the charges of the bond-forming atoms is attractive for the concepts it applies and for its usefulness in the prediction of important thermochemical quantities, such as the energy of atomization or the enthalpy of formation of organic molecules. But its success critically depends on the availability of accurate charge results.

Now, the problems associated with the search for *meaningful* atomic charges, such as those required in applications of our bond energy formulas, are manifold. One

concerns the selection of the population analysis. As is well known, Mulliken charges represent only one of the possible definitions of atomic net charges [21–29]; widely different values are produced with comparatively small changes in a basis set, even in *ab initio* SCF calculations. Since trends within similar series are generally little affected by the computational scheme, this has seldom received significant attention in applications to chemical problems. But the present situation is one where the choice of the basis set could play a very important role, as does inclusion of basis superposition effects [30]. While Mulliken’s population analysis [31] is probably the most widely used one, Löwdin’s method [32] or Jug’s approach [33] could also be envisaged with success. Selection of the method does greatly affect the final numerical results, as vividly exemplified by the SCF net charges obtained with a minimal basis set for the nitrogen of methylamine, namely, -374 , $+47$, and -291 millielectron ($1\text{ me} = 10^{-3}\text{ e units}$), respectively, depending on whether Mulliken’s, Löwdin’s, or Jug’s definition is implemented. So, under these circumstances, things do not seem encouraging. Fortunately, they are not really as bad as suggested by these examples. We shall learn about charge *variations* suited for bond energy calculations that withstand comparison with experiment. A promising example is offered by the calculated charge variations Δq_N of the nitrogen atoms of a series of amines, relative to that of methylamine, as given by the Mulliken, Löwdin, and Jug methods (Table 1.2). The examples shown in Table 1.2 are meant to illustrate how different ways of partitioning overlap populations do affect calculated atomic charge variations accompanying structural modifications; admittedly, the differences can be relatively minor, but the absolute values of the same charges defy any reasonable expectation. Evidently, one should not rely too heavily on numbers calculated by population analysis. Mulliken’s assignment of half the overlap probability density to each atomic orbital (AO) is rather arbitrary and occasionally leads to unphysical results. Still, we shall see that the Mulliken scheme offers a valid starting point. To get useful charges, however, we must rethink the problem of assigning overlap populations—a topic highlighted in Chapter 8.

At first, it may seem surprising that many methods give results in semiquantitative agreement with Mulliken population analysis values, but it now appears that it is rather what one should expect from any sensible method. This holds true for the simplest

TABLE 1.2. Calculation of Δq_N in Amines

Molecule	Δq_N (me)		
	Mulliken	Löwdin	Jug
CH_3NH_2	0	0	0
$\text{C}_2\text{H}_5\text{NH}_2$	−5.16	−4.86	−5.01
<i>n</i> - $\text{C}_3\text{H}_7\text{NH}_2$	−4.96	−4.65	−4.86
<i>iso</i> - $\text{C}_3\text{H}_7\text{NH}_2$	−9.00	−8.90	−9.20
<i>iso</i> - $\text{C}_4\text{H}_9\text{NH}_2$	−3.59	−3.84	−4.10
<i>tert</i> - $\text{C}_4\text{H}_9\text{NH}_2$	−11.58	−11.94	−12.49

possible one [29]: Del Re's approach is based on rough semiempirical approximations of simple molecular orbital–linear combination of atomic orbitals (MO-LCAO) theory of localized bonds. The original, extremely simple parameterization that reproduced electric dipole moments [29] and a more recent one [34–36] for charges that correlate with nuclear magnetic resonance (NMR) shifts and that are fit for accurate energy calculations, turn out to correspond to solutions of MO-LCAO Mulliken-type population analyses differing from one another by the mode of partitioning overlap terms [37]. The link between Del Re's simple semiempirical approach and the more familiar MO-LCAO charge analyses is clearly established.

While the charges for use in our description of bond energies were originally obtained from accurate SCF computations using a variant of Mulliken's population analysis, the observation [38–44] that the ^{13}C , ^{15}N , and ^{17}O nuclear magnetic resonance (NMR) shifts are linearly related to these charges permitted rapid progress in the application of the charge-dependent bond energy formulas to thermochemical problems since NMR results are more readily available than are good-quality population analyses. Of course, this strategy presumes not only a justification for assumed correlations between NMR shifts and net atomic charges (which is described in Chapter 6) but also a solid knowledge of well-justified charge analyses. This is no minor task.

The original definition of atomic charges found in the CH backbone of organic molecules is rooted in the idea that the carbon charges vary as little as possible on structural modification. This has triggered *inductive reasoning*, which maintains that if a situation holds in all *observed* cases, then the situation holds in *all* cases. Indeed, detailed tests involving ^{13}C chemical shifts, the ionization potentials of selected alkanes, and, most importantly, thermochemical data, unmistakably point to identical sets of charge values. Now, of course, the problem of induction is one of considerable controversy (and importance) in the philosophy of science; we must thus be extremely cautious in attributing physical meaning to atomic charges. In the approach known as *instrumentalism*, one could as well consider them as convenient ideas, useful instruments to explain, predict, and control our experiences; the empirical method is there to do no more than show that theories are consistent with observation. With these ideas in mind, atomic charges that are now widely used in molecular dynamics calculations and in the evaluation of solvation energies within the generalized Born approach also deserve renewed attention.

Admittedly, things have not been easy. But now we can benefit from the beauty of simplicity and learn about one unique kind of atomic charge rooted in quantum theory: the only ones that satisfy highly accurate correlations with experimental NMR shift results and that are at the same time directly applicable to bond energy calculations.

But the key to the theory of bond energy is in the description of real-space core and valence regions in atoms and molecules: therein lies the basic idea that gives rise to the notion of molecular chemical binding, expressed as a sum of atomlike terms. That marks the beginning of our story.

