International Journal of Pharmaceutics 418 (2011) 304–317

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm

Calculated carbon–hydrogen bond dissociation enthalpies for predicting oxidative susceptibility of drug substance molecules

Thomas R. Sharp*

Research Analytical, Pharmaceutical Sciences, Pfizer Global Research & Development, Groton, CT 06340, USA

ARTICLE INFO

Article history:
Received 9 December 2010
Received in revised form 19 April 2011
Accepted 26 April 2011
Available online 7 May 2011

Keywords:
Computational stability
Bond dissociation energy (BDE)
Quantum mechanics
Semi-empirical methods
AM1
SAM1
PM3
RM1
PM5
PM6
DFT (density functional theory)
Sertraline
Ziprasidone
Trovafloxacin
Varenicline
Ezlopitant
Quinuclidines

ABSTRACT

The carbon–hydrogen bond dissociation enthalpy (BDE) concept is evaluated as a potential computed indicator of stability of pharmaceutical drug substance candidates – specifically for oxidative stability of these molecules. Computational methods are discussed. Accuracy and validity of the methods are evaluated. BDEs are computed for several well-known molecules, for which stability and degradant identification information is known. Anecdotal correlations are noted between the lowest BDE energies of familiar molecules (sertraline, ezlopitant and related structures, ziprasidone, trovafloxacin, and varenicline), the sites of oxidative lability on these molecules and the identities of oxidative degradants. A low BDE may correlate in general with a reactive site on a molecule, not just an oxidatively susceptible one.

1. Introduction

Hydrogen atoms on an organic molecule that can be easily removed are potential sites of oxidation of that organic molecule. Although there are several different kinds of oxidation mechanisms at carbon atoms, they share the one general step that a hydrogen atom has to be removed from the molecule. How easily a hydrogen can be removed thermodynamically influences the propensity of that site on the molecule toward oxidation. This considering is the premise of the investigation being described here. We defer oxidation at atoms other than carbon to a later time.

The terms "bond dissociation energy" and "bond dissociation enthalpy" are often used interchangeably. The dissociation energy refers to the energy required to break a bond at 0 K. Bond dissociation enthalpy is adjusted to a particular prevailing temperature, typical 298 K. The semantic difference between the terms is relatively innocuous, as the values are almost numerically equivalent for most organic molecules (Blanksby and Ellison, 2003). Numerical values differ by 3 kcal/mole or less. The average unsigned error reported for the PM6 semi-empirical method (Stewart, 2007), one of the most recently published refinements of the semi-empirical quantum mechanical methods, is 4.4 kcal/mole.

Bond dissociation energy, conceptually, is differentiated from bond energy. March (1992a) discusses the origins of commonly cited bond energy values for organic molecules. The aliphatic C–H bond in simple organic molecules, for example, is cited to be in the range of 96–99 kcal/mole. These numbers derive from total combustion experiments, in which the C–H bond energy for methane is the combustion energy, 393 kcal/mole, divided by 4 (the 4 C–H bonds in methane). Values derived by combustion of progressively larger hydrocarbons decrease a small amount, to give the 96–99 kcal/mole range.

Sequential removal of a hydrogen atom from methane, however, results in (experimentally measured) bond dissociation energies of 105, 110, 101 and 81 kcal/mole (Blanksby and Ellison, 2003).

* Correspondence address: 63 Bunny Road, Preston, CT 06365, USA. Tel.: +1 860 887 5314.
E-mail address: tootalltom50@yahoo.com

0378-5173/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpharm.2011.04.063

Although these values add up to the 397 kcal/mole value (at 298 K) for total combustion, the bond dissociation of each hydrogen is different. The energy required to break an individual bond is dependent on the rest of the structure to which it is attached. For this reason, specific bond dissociation energies, dependent upon an associated structure, are more appropriate indicators of oxidative susceptibility.

Nearly all theoretical computations on molecules are done on the isolated molecule, in the absence of influences by neighboring molecules or solvent—i.e., in the gas phase. Solvation models exist, and can be layered on top of the calculations for a molecule. Solvents, concentration and the presence of other molecules influence rates of reaction and chemical reactivity. Molecular dynamics simulations attempt to address these issues computationally. Are such calculations as being done here relevant, applicable? Yes. The calculations done here address the energetics of a specific bond, as influenced by the specific structure in which it is found. Solvation, concentration and associated other molecules either attenuate or enhance reactivity at a particular site. The fundamental reactivity of a site, however, is determined by the bond dissociation energy. These calculations address this fundamental reactivity.

2. Experimental

2.1. Computations

Computations were done using several available computational chemistry program packages: the CAChe molecular modeling software package (CAChe, 1989–2000), the AMPAC package (AMPAC, 1992–2011) and Gaussian03 (Frisch et al., 2004).

Semi-empirical quantum chemical calculations with various models have been done, primarily because of expediency. The MOPAC semi-empirical computation procedure was performed (CAChe, 1989–2000), using the PM5 parameterization for early work, with the COSMO (Klamt, 1995) simulation of an aqueous solvent environment. PM5[1] is only implemented commercially in the CAChe package (CAChe, 1989–2000). Comparative computations were done using the AM1 (Dewar et al., 1985; Dewar and Zoebisch, 1988; Dewar and Jie, 1989; Dewar and Yuan, 1990), PM3 (Stewart, 1989a,b, 1991), PM6 (Stewart, 2007), RM1 (Rocha et al., 2006) and SAM1 semi-*ab initio* (Dewar et al., 1993; Holder et al., 1994) parameterizations, implemented in the AMPAC commercial package (AMPAC, 1992–2011). Occasional density functional (DFT) calculations were performed at higher levels of theory for comparison, using either the Gaussian03 (Frisch et al., 2004) package or the CAChe package (CAChe, 1989–2000).

Two new parameterizations, christened RM1 (Rocha et al., 2006) and PM6 (Stewart, 2007), have recently been published by J.J.P. Stewart and colleagues. RM1 improves on the performance of the AM1 model by a new parameterization, the theoretical basis being the same as for AM1. Element coverage includes carbon, hydrogen, oxygen, nitrogen, phosphorus, sulfur and the halogens, an element list which, the authors propose, encompasses a majority of pharmaceutically and biochemically relevant organic molecules. The molecule training set for the RM1 re-parameterization contained 1736 compounds. As such, programs such as MOPAC, which implement AM1 and permit input of alternate parameter sets, can be used to perform RM1 computations. Rocha et al. (2006) published the RM1 parameter set in the electronically available supplemental material to their report. The parameter set has been downloaded and used with the MOPAC2002 implementation in the CAChe molecular modeling software package (CAChe, 1989–2000) to perform some of the RM1 computations. AMPAC 9 also implements the RM1 parameterization, as well as the PM6 parameterization (Stewart, 2007). PM6 parameterization includes 70 elements (Stewart, 2007). For the subset of pharmaceutically relevant elements, PM6 claims an incremental improvement in prediction of molecular geometries and heats of formation over RM1.

2.2. Bond dissociation energy vs. bond dissociation enthalpy

All of the semi-empirical methods have been parameterized at 298 K. Therefore, according to the distinctions made by Blanksby and Ellison (2003), the calculated results reported here having been done using the semi-empirical methods, will be bond dissociation enthalpies since the semi-empirical methods are parameterized at 298 K.

2.3. Computation strategies

Three strategies have been examined. In all cases, evaluation of a BDE is not a single calculation, but a series of comparative calculations.

The first step in all procedures is optimizing the geometry of a molecule to its minimum energy conformation and computing the energy of formation of this conformation. Frequency computations were routinely done on conformations arrived at by geometry optimization to verify that the computation had found a stationary point. Using this minimum energy conformation, additional calculations were performed on structures from which a single hydrogen atom has been removed. The geometry of this radical species was then re-optimized to find the minimum energy conformation, frequency computation checked, and the energy of formation of the radical species computed. The BDE of the hydrogen in question was calculated from this value, using Eq. (1) (see below) and the $\Delta E_{\text{formation}}$ of the intact molecule.

An unrestricted Hartree–Fock wavefunction was used consistently for calculations on the species containing an unpaired electron, and used even for closed shell species containing no unpaired electrons, in order to maintain consistency of comparison of computed results. Most comparisons of energies of closed-shell species computed by both restricted and unrestricted Hartree–Fock wavefunctions were identical. Computations were made by systematically removing a hydrogen atom from each unique structural environment on the molecule under consideration. At minimum, for a molecule containing n unique hydrogen structural environments, $n + 1$ calculations are required to fully evaluate the molecule. When appropriate, dihedral angle optimization was performed before doing the hydrogen removal computations to verify that geometries were indeed at the global energy minimum. An example will be discussed where this additional consideration influences final results and interpretation. When not optimized for dihedral angle contributions, calculations on structures from which a hydrogen atom had been removed were done on as close to the same molecular conformation as possible.

A second strategy utilized the ability to conduct a bond-stretching experiment in the computer. Energies are computed as a function of interatomic bond distance (between a carbon and hydrogen atom) over a range of distances, from slightly less than the equilibrium bond distance to a significantly greater than equilibrium distance, having effectively "removed" the hydrogen from the molecule. While the single carbon–hydrogen "bond length" is being adjusted, the remainder of the molecule is allowed to optimize to a minimum-energy conformation. Computed results qualitatively follow a Morse potential (discussed below). In simplest evaluation, the difference between the energy of formation of the globally opti-

[1] See discussion in Rocha et al. (2006).

mized geometry of the intact molecule and the energy at very long inter-atomic distance is the BDE.

The third method involves an examination of the procedure proposed by Lewin and Cramer (2004). They similarly compute the minimum-energy geometry of a molecule. For each hydrogen environment, they note the equilibrium inter-atomic bond distance, and recompute single point energies for geometries where the inter-atomic bond distance has been shortened or lengthened by a small amount. They evaluate several levels of theory—AM1 and PM3 semi-empirical models, the MIDI! Hartree–Fock model and density functional theory (DFT) with the B3LYP/MIDI! hybrid functional. They use a restricted Hartree–Fock wave function, as the molecules upon which they compute are still closed shell species, even though the length of the carbon–hydrogen bond under examination has been perturbed. They use these computed energies to curve-fit to the Morse potential equation, examining several different strategies, to determine a BDE. Using a minimum of three points to do the curve fitting, this procedure requires $2n+1$ computations for n unique hydrogen structural environments for full evaluation of a molecule. They primarily examine BDEs for removal of one type of hydrogen atom from a molecule. (Two exceptions are pyridine and pyrimidine in their secondary reference set.) They examine the optimum distances with which to stretch and compress the target bond, and advantages of performing evaluations with additional computed values.

2.4. Curve fitting

Values resulting from computation of energies as a function of inter-atomic distance were fitted to a Morse potential as one means of determining bond dissociation energies—similar to the Lewin and Cramer approach (Lewin and Cramer, 2004). Curve fitting was done using a simulated annealing algorithm (Goffe et al., 1994), implemented in a commercially available program called GOSA-fit.[2] In addition to the minimalist three-point and five-point strategies examined by Lewin and Cramer (2004), the availability of computed values at an extended range of interatomic bond distances prompted examination of curve fitting of the full data set to the Morse potential, with the goal of examining and using this approach to determine BDEs (see the supplemental materials).

2.5. Choice of compounds for study

The reference set of compounds, for which high quality experimental bond dissociation energies have been measured, was assembled by Lewin and Cramer (2004), and is further documented in the review by Blanksby and Ellison (2003). Compounds and energies are collected in Table 1. Sertraline, ziprasidone, trovafloxacin and varenicline have been examined because they have a significant accumulated history of information on their stability and degradant identities. Similarly, a series of quinuclidine compounds were examined which have significant accumulated experimental information on oxidative degradation.

2.6. Heat of formation of H•

The energy of formation of the hydrogen atom used by all of the semi-empirical parameterizations is derived from the Handbook of Chemistry and Physics (Weast, 1982) as 52.1 kcal/mole. This value is generated consistently by the methods used here. This value has been consistently used in the BDE calculations, generated by using Eq. (1). It is used as an internal consistency check in the bond stretching computed experiments. If computed results indicate a value significantly different from this, explanations for the deviation were investigated and problems corrected if possible. The "energy of formation" for H• was directly calculated for use in DFT BDE determinations using the B3LYP functional and the appropriate basis set. For the 3-21g basis set, the number is −0.4973 atomic units. For the 6-31g basis set, the number is −0.5003 atomic units.

2.7. Results and discussion

The thermodynamic expression for a hydrogen bond dissociation enthalpy, D_{RH}, is given in Eq. (1). $\Delta H_f[RH]$ is the heat of formation of the molecule of concern. $\Delta H_f[R^\bullet]$ and $\Delta H_f[H^\bullet]$ are the heats of formation of the ensuing radical species upon homolytic cleavage of the bond. It is a measure of the energy required to remove a hydrogen radical from its parent molecule, leaving behind a radical. It is presumed to be a reasonable approximation to an intermediate on the way to oxidation of the molecule:

$$D_{RH} = \Delta H_f[R^\bullet] + \Delta H_f[H^\bullet] - \Delta H_f[RH] \quad (1)$$

2.8. Benchmarking of computed results

Comparison of calculated values with good experimental determinations of BDEs is the final indicator of the validity of calculations. Our own evaluations of the reliability of computed results are summarized in Table 2 and Fig. 1, using the AM1, PM3, PM5, PM6, RM1 and SAM1 models.

As a cautionary note, the semi-empirical methods are not parameterized to cover the entire periodic table. AM1, RM1 and SAM1 are only parameterized to include compounds containing C, H, N and O, and later extended to include phosphorus, sulfur and the halogens—elements of importance and interest for biological and pharmaceutical molecules. PM3 was extended to additionally include silicon and aluminum. PM5 covers 27 elements (Rocha et al., 2006). PM6 includes coverage of 70 elements (Stewart, 2007). Values computed for a molecule containing an element that is not covered by the parameterization of a specific model will be meaningless, if the implementation of the method does not outright reject the molecule as containing uncovered elements.

Values computed for the reference compounds are given in Table 1. Experimentally determined BDE values for these compounds were compiled from the literature by Lewin and Cramer (2004), and used in their evaluation of their computations. The first 13 compounds in the table (methane through tetrahydronaphthalene) are ascribed as having the most precise and reliable experimental determinations, and are referred to here as the primary reference set. The additional 23 determinations (acetaldehyde through trimethylamine) are referred to as the secondary reference set. The combined values are referred to as the extended reference set. Correlations of computed vs. experimental values are shown in Fig. 1.

The perfect correlation line (slope = 1, y-intercept = 0) is the diagonal drawn in each of the plots in Fig. 1. The solid fitted line (linear least squares) passing through the data is a fit to the results for the 13 primary reference compounds in Table 1. The dotted line is a least squares linear fit to data for the entire collection of values—the extended reference set.

Regression statistics are compiled in Table 2 for comparison. The upper linear equation and R^2 value in each plot of Fig. 1 is for the primary reference set, the lower for the extended reference set. In all cases, the slope of fitted line for the primary reference set is near 1. All six semi-empirical models systematically underestimate the BDE. AM1 underestimates by the largest amount, by 27 kcal/mole. PM3 underestimates the least, by 14 kcal/mole. RM1 shows a correlation with the primary test set closer to a slope of 1 than does

[2] Bio-Log Scientific Software, B.P. 27201, 31672 Labege Cedex, France. http://bio-log.biz..

Table 1
Experimentally determined C–H bond dissociation energies, and values calculated by various semi-empirical computation models. The units on all values is kcal/mol.

Molecule	BDE	AM1	PM3	PM5	RM1	PM6	SAM1
Methane	104.99 ± 0.03	85.0	93.1	89.3	90.8	96.2	97.9
Ethane	101.1 ± 0.4	85.2	84.7	82.9	84.0	84.0	87.1
Propane (2°)	98.6 ± 0.4	79.8	77.7	77.7	78.5	77.4	78.1
tert-Butane (3°)	96.5 ± 0.4	75.1	72.1	73.9	74.5	72.6	70.7
Acetylene	133.32 ± 0.07	139.7	140.1	133.7	143.1	126.7	133.5
Propene (allylic)	88.8 ± 0.4	75.7	76.9	75.4	76.4	77.9	77.3
Ethane	110.7 ± 0.6	96.1	93.3	92.0	94.0	91.7	95.9
Cyclopentane	97.6	80.4	77.4	78.0	77.8	78.1	78.9
Cyclohexane	98.6	80.7	79.0	78.6	78.7	78.7	78.4
Benzene	112.9 ± 0.5	103.3	99.7	98.7	104.3	99.5	97.3
Toluene (benzylic)	89.9 ± 0.6	76.2	77.6	84.5	77.1	78.5	77.5
Naphthalene (α)	112.2 ± 1.3	102.0	99.2	97.9	102.7	98.8	96.6
Tetrahydronaphthalene (benzylic)	82.9 ± 1.2	71.3	77.6	69.2	71.3	71.6	68.7
Acetaldehyde (methyl)	94 ± 2	86.2	86.6	85.2	85.6	89.1	89.8
Acetic acid	94.6 ± 3.0	88.1	87.5	87.5	86.8	90.5	93.4
Dimethylamine	87 ± 2.4	74.7	77.4	74.0	71.0	71.5	75.3
1-Benzyl-1,4-dihydro-nicotinamide C(4)-H	67.9 ± < 1	60.9	60.2	62.0	60.5	59.7	56.9
Dimethyl ether	93 ± 2	82.5	80.5	76.7	77.9	76.9	82.7
Ethanolamine (N-α)	90.7 ± 2	70.3	71.0	70.2	66.8	69.7	69.0
Formaldehyde	88.144 ± 0.008	81.6	76.0	74.4	76.0	68.9	74.4
Methanol	96.1 ± 0.2	81.4	79.1	74.8	76.4	76.8	82.1
Methylamine	94.2	73.8	76.1	72.8	68.4	71.3	74.8
Methanethiol	94.2	75.4	77.5	85.4	81.0	79.6	81.1
Morpholine (N-α)	93 ± 2.4	71.1	71.4	71.9	65.0	72.2	69.4
N,N-dimethylaniline(methyl)	91.7 ± 1.3	74.0	76.1	80.4	73.0	71.2	74.3
Piperidine (N-α)	92 ± 2.4	73.4	74.6	70.0	70.2	73.4	69.3
Propylamine (N-α)	93.1 ± 2	69.7	70.4	69.1	65.0	67.9	68.2
Pyridine (2)	105 ± 2	101.8	97.8	98.0	100.9	98.0	97.1
Pyridine (3)	112 ± 2	106.4	102.3	101.0	106.0	102.6	99.1
Pyridine (4)	112 ± 2	104.8	101.2	100.0	104.3	100.3	97.2
Pyrimidine (2)	98 ± 2	103.5	98.9	100.2	101.8	100.8	102.7
Pyrimidine (4)	103 ± 2	103.2	99.1	99.0	101.5	98.7	96.8
Pyrimidine (5)	112 ± 2	109.7	104.9	103.7	108.6	105.8	102.3
Pyrrolidine (N-α)	90.1 ± 2.4	73.8	72.6	72.4	68.3	71.6	71.4
Tetrahydrofuran (O-α)	92.1 ± 1.6	79.0	74.5	73.9	73.3	73.5	76.7
Trimethylamine	86.5 ± 2.2	75.4	78.6	75.3	73.8	72.1	75.3

Table 2
Regression statistics for computed vs. experimental BDEs of reference compounds. Values for acetylene are not included in the semiempirical calculations. Acetylene is included in both sets of DFT calculation.

	Primary set		Extended set	
	Linear fit	R^2	Linear fit	R^2
Semi-empirical results				
AM1	$1.084x - 23.20$	0.897	$1.158x - 27.73$	0.747
PM3	$0.932x - 9.316$	0.744	$1.017x - 15.37$	0.731
PM5	$0.880x - 4.470$	0.776	$0.996x - 13.87$	0.704
RM1	$1.063x - 21.84$	0.844	$1.211x - 34.74$	0.720
PM6	$0.913x - 7.63$	0.794	$1.093x - 23.69$	0.729
SAM1	$1.005x - 16.30$	0.794	$1.051x - 18.90$	0.708
DFT results				
uB3LYP 3-21g	$0.977x + 9.75$	0.983	$0.958x + 11.9$	0.934
uB3LYP 6-31g	$1.050x + 0.166$	0.992	$1.012x + 5.163$	0.922

AM1, and underestimates BDEs by only 19 kcal/mole, rather than 27 kcal/mole for AM1. SAM1 underestimates by 16 kcal/mole. PM5 shows the greatest deviation of the correlation line from unit slope, so that a correction of computed PM5 values would not be a simple additive one. The non-unit slope of the correlation line for PM5 is not unexpected. PM5 covered 27 elements in the periodic table, and was purported to be a parameterization emphasizing certain structural classes of compound. The PM5 details have not been fully

Table 3
Regression statistics for computed vs. experimental $\Delta E_{formation}$ of reference compounds.

	Primary set		Extended set	
	Linear fit	R^2	Linear fit	R^2
AM1	$1.067x - 0.012$	0.962	$0.999x - 1.594$	0.963
PM3	$1.008x + 0.926$	0.983	$0.975x - 2.790$	0.960
PM5	$0.942x + 0.356$	0.964	$0.963x - 0.818$	0.977
RM1	$0.953x + 0.198$	0.980	$0.975x - 1.494$	0.989
PM6	$1.004x + 2.379$	0.992	$1.005x + 0.950$	0.990
SAM1	$1.019x - 1.482$	0.986	$1.008x - 1.461$	0.993

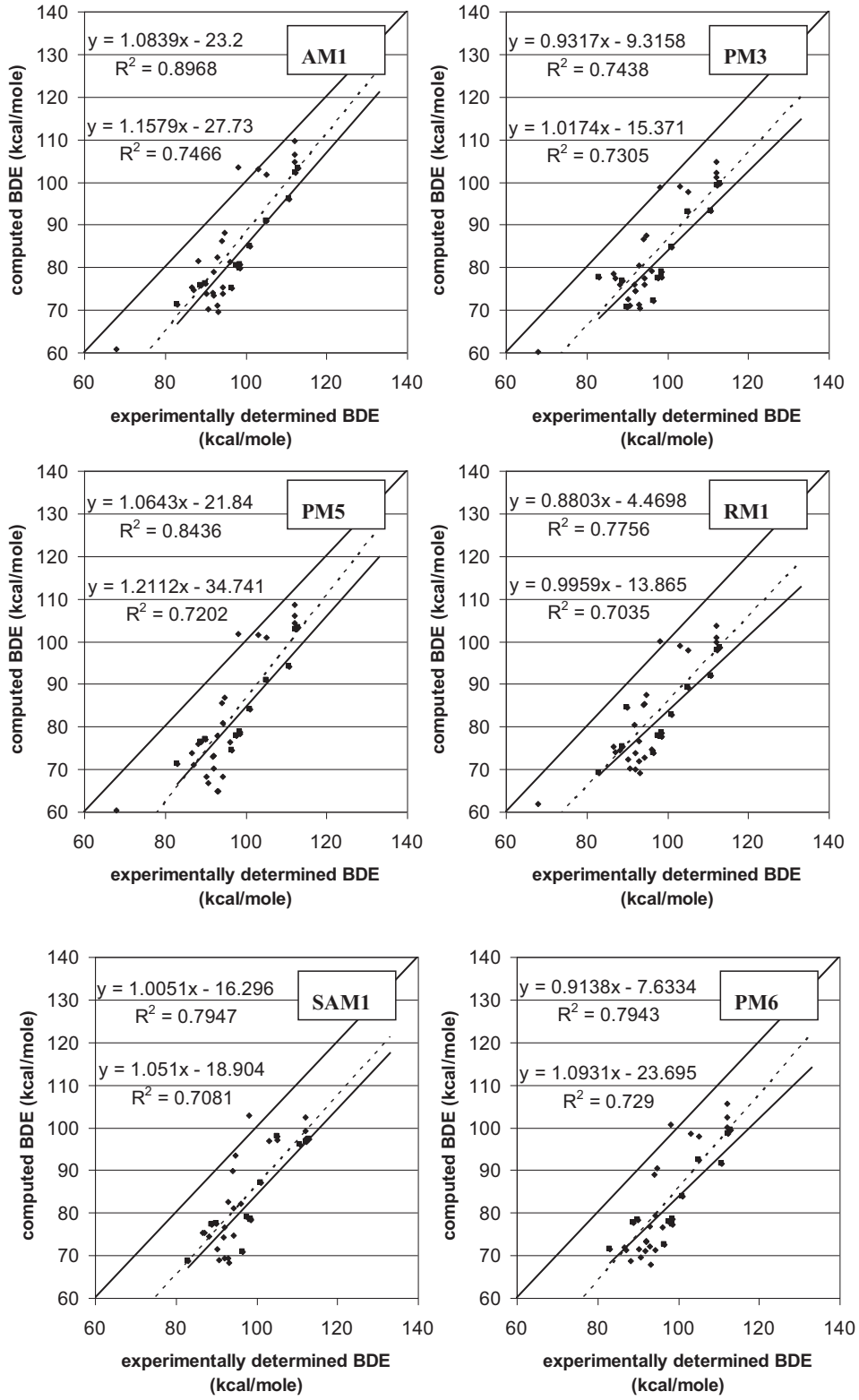


Fig. 1. Measured vs. computed BDEs for semi-empirical quantum mechanical models. All use an unrestricted Hartree–Fock wave function, with no solvent. The solid diagonal line connecting lower left and upper right corners is the perfect correlation line (slope = 1, *y*-intercept = 0). The solid line spanning the data points is the least squares linear fit for the primary reference set, plotted as squares. The dotted line is for the combined primary reference set and secondary reference set (extended reference set, plotted as diamonds). Linear fits in each plot are without including results for acetylene. See text discussion regarding acetylene. Regression calculations were done using MicroSoft Excel.

published (Rocha et al., 2006). (Some details on PM5 are discussed in Rocha et al., 2006.) Including the less certain BDE determinations for the expanded reference set causes correlation results for the AM1, RM1 and SAM1 models to deviate further from a slope of 1. The PM3 correlation changes minimally, while the slope for the PM5 correlation more closely approaches 1. The SAM1 results show the least deviation from unit slope upon inclusion of the secondary reference set. Correlations with the experimental values suggest

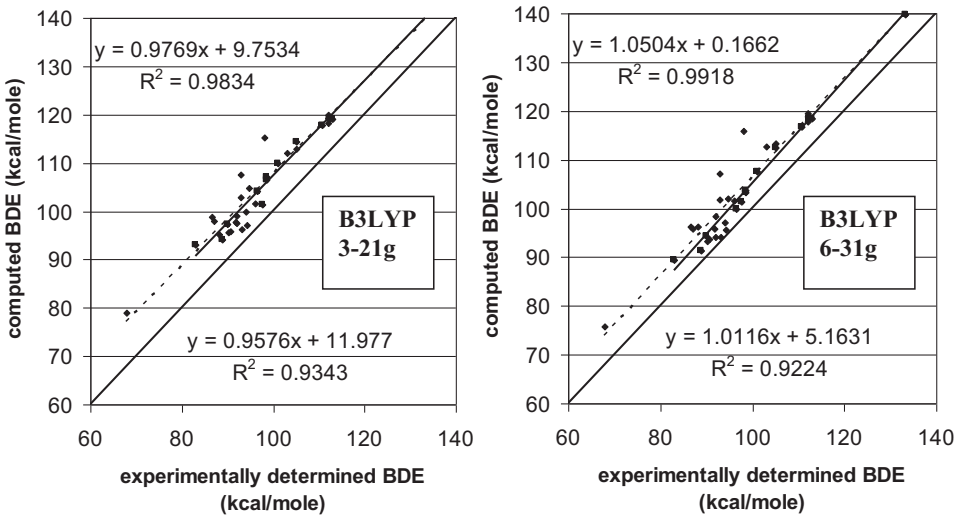


Fig. 2. Measured vs. computed BDEs for DFT unrestricted B3LYP calculations of BDEs for the reference set. Left panel, 3-21g basis set. Right panel, 6-31g basis set. See Fig. 1 caption for further explanations.

that SAM1 is the most reliable—i.e., the slope of the correlation closest to 1.0. On the whole, however, there seems no overwhelming distinction for preferring one model over another.

Values for acetylene were computed using each of the models, but have not been included in the regression analyses of Fig. 1. The computed values, however, are given in Table 1. Computationally, acetylene presents difficulties because it is a linear molecule. Further, the results, upon comparison with the experimental value, seem very good—almost too good. However, when taken in relation to the computed results for the other 12 compounds in the primary reference set, computed values for acetylene are "high." This single point, if included in the regression calculations, skews the trend lines for evaluation of all models.

The correlation plots in Fig. 1 indicate that BDEs computed by the various methods should be adjusted up 15–30 kcal/mole (depending upon the computational model used) to reflect true values. However, results reported in this paper will consistently not be corrected by these amounts. As we are most concerned with relative comparison of values within a molecule, the relative differences answer questions without the need to do corrections. If and when necessary, we will make explicit notation when values are corrected. If further use of the absolute values reported here is anticipated, they should be corrected appropriately.

2.9. Sources of scatter

A substantial amount of scatter is evident in the plots of experimental vs. computed BDEs of Fig. 1. What is the source of this scatter? The value for $\Delta E_{\text{formation}}$ for H• used in Eq. (1) is a constant. Experimental and computed $\Delta E_{\text{formation}}$ values for the set of reference compounds have been collated from the literature on the semi-empirical methods (computed when missing) and regressed against each other. (Values and regression statistics are given in the supplemental material for this paper, as are plots of the experimental vs. computed values.)

Consistent with the original conclusions of the original authors on validation of the semi-empirical methods, the correlation of computed $\Delta E_{\text{formation}}$ with experimental determination is quite good. Regression statistics are given in Table 3 over the primary and extended reference compound sets used here. Acetylene is not problematic for $\Delta E_{\text{formation}}$ comparisons, and does not significantly influence the correlations. The regression lines for all six models give slopes very close to 1, and very good correlation coefficients. The scatter in the computed vs. measured BDE correlations, therefore, does not originate in the computation of $\Delta E_{\text{formation}}$ for the starting closed shell compounds.

Similar evaluation of computed vs. experimental values for open shell (radical) chemical species is not as feasible. All of the semi-empirical methods have been evaluated for validity of computed values of open shell species. The test set for such species is not as large, as for closed shell species. Computed results are not as good. Source of scatter in the computed BDE values generated here must therefore be in the computations of the open shell species.

Indeed, Boyd (2005) discusses the difficulties of conducting reasonable computations on radical species. The unrestricted Hartree–Fock (UHF) treatment used here is a reasonable first approximation. A restricted open Hartree–Fock (ROHF) formalism was developed as a way to address this problem, but was not successfully applied here.[3] Bally and Borden (1999) cite that ROHF is a fundamentally flawed physical model. Both Boyd and Bally and Borden suggest that unrestricted DFT computations with a B3LYP functional, and even time-dependent DFT should yield improved results, but at significant computational cost.

Results of computing BDEs for the extended test set of compounds using DFT unrestricted B3LYP/3-21g (a low level of theory) and unrestricted B3LYP/6-31g (a higher level of theory) are shown in Fig. 2 and included in Table 2 for comparison with semi-empirical results. Correlation statistics are shown on the plots. The scatter in the results is diminished significantly over that for any of the semi-empirical methods, arguing that the scatter with the semi-empirical computations originates in deficiencies for handling open-shell species. The correlations for DFT results are tighter and show slopes close to 1, but the computations, however, overestimate the BDEs by approximately 10 kcal/mole (3-21g basis set) and 5 kcal/mole (6-31g basis set). The semi-empirical methods underestimate. Acetylene is included in this compari-

[3] Gaussian03 implements and makes the ROHF option available for semi-empirical and other calculations. Although $\Delta E_{\text{formation}}$ for those molecules in the primary reference set for which ROHF calculations could be successfully completed gave the same results as for RHF and UHF options, the calculations were computationally expensive, and often unsuccessful for the intact closed-shell compounds. Attempting ROHF calculations on open-shell structures of pharmaceutically significant molecules is deemed impractical.

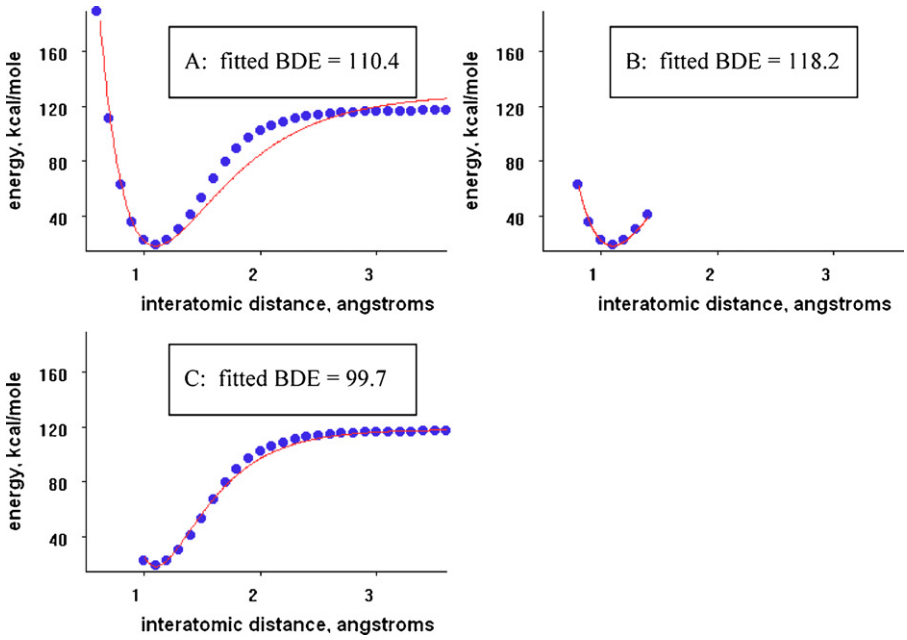


Fig. 3. Fits of Morse potential to computed energies for stretch of a C–H bond of benzene, using computed energies for interatomic distances from 0.6 Å to 3.6 Å. Computed bond length is 1.092 Å. Filled circles plot the computed values using SAM1 UHF. Solid line is the best fit to the computed values. The BDE computed for benzene by the procedure described earlier in this paper is 97.3 kcal/mole. (A) Fit for the entire range of interatomic bond distances. (B) Computed values and Morse potential fit to points most closely corresponding to equilibrium C–H bond length and approximately ±0.1, ±0.2 and ±0.3 Å from equilibrium length (proposed Lewin and Cramer method). (C) Computed values and Morse potential fit from approximately equilibrium length minus 0.1 Å to plus 2.5 Å.

son. Such results are consistent with those reported by Fox and Kollman (1996) on a series of substituted toluene derivatives. Their DFT unrestricted B3LYP/6–31g* method calculated results which were consistently approximately 5 kcal/mole high. DFT calculations do give tighter correlations with experimental results than the semi-empirical results, but at significantly higher computational expense.

2.10. Examination of an alternate computation method

In addition to having compiled a reference collection of experimental BDE values, Lewin and Cramer (2004) propose a procedure to compute BDEs, and computed and examined results using several computational methods – semi-empirical and *ab initio* – and several levels of theory. Of considerable concern in proposing any computational solution to a problem is the required level of theory and the associated computational cost. If adequate, semi-empirical methods are preferred simply because of computational expediency. Additional value must be provided if higher theory level and longer computational times are to be required. Lewin and Cramer conclude that, for their procedure, if accuracy of the absolute values is not at issue, but simple comparison of values and rankings are the purpose, the expediency of the simpler faster-calculating semi-empirical parameterizations will give acceptable results. Calculations between different molecules might not be accurately compared, but calculations for sites within a molecule should be comparable—systematic "errors" would be expected to cancel out. Our observations and results support their conclusions regarding appropriateness of theoretical level. While different parameterizations and models produce some rather significantly different absolute values for energies, the BDE rankings obtained from different methods are comparable.

Conclusions concerning their protocol for computing BDEs, however, are different. The Lewin and Cramer (2004) procedure calculates C–H bond dissociation enthalpies based on modeling of the Morse potential (Eq. (2)), which describes the change in energy upon homolytic stretching of a C–H bond. E is the energy content of the molecule. D_{RH} is the bond dissociation energy of the bond being examined. The interatomic distance r_{eq} (in angstroms) is that between the two atoms of concern in the molecule's ground state. r is the interatomic distance (different from the ground state distance) to which the bond has been contracted or stretched. The symbol a represents a fitting constant:

$$E = D_{RH}\,[1 - e^{-a(r-r_{eq})}]^2 \quad (2)$$

Their proposed protocol computes the energies of a molecule at the C–H equilibrium bond distance, and at slightly less and slightly greater than the equilibrium bond distance, then uses these computed values to fit to the Morse potential equation and obtain a BDE. Fig. 3 shows computed data for the BDE of a hydrogen atom attached to benzene, computed at C–H bond distances ranging from 0.6 Å to 3.6 Å, in 0.1 Å intervals.[4] A typical C–H equilibrium bond length is approximately 1.1 Å (March, 1992b). The equilibrium C–H bond length computed for benzene is 1.092 Å. This value was used in the Morse potential curve fittings. BDE values derived from the fits[2] of the Morse potential to the computed points are given in each panel of Fig. 3.

The favored Lewin and Cramer protocol approximates the computed data shown in Fig. 3, panel B. Their preferred protocol stretches and contracts the bond length by 0.1 Å, 0.2 Å, 0.3 Å on either side of the equilibrium bond length, and computes a fit to the Morse potential using only these five points at 0.8 through 1.4 Å bond length. Clearly, the panel B results, a close approxima-

[4] Energies computed using AMPAC SAM1. C–H bond length was fixed at a specified length. A new electron density was computed for each bond length, and the remainder of the structure permitted to geometry-optimize. Computed values are tabulated in the supplemental material for this paper.

tion to the Lewin and Cramer procedure, do not give satisfactory estimation of the BDE. The energies at the bottom of the energy well do not contain sufficient information to allow adequately predicting the energy of the system at long bond length. Further, including computed energies at significantly less than the equilibrium bond length prohibits fitting to the Morse potential, as illustrated in Fig. 3 panel A. Interatomic distances significantly shorter than the equilibrium bond distance start to infringe on van der Waals radii. The value determined by curve fitting to an extensive set of numbers, in Fig. 3 panel C, compares well with that determined by simple determination from two UHF calculations for benzene, with and without a hydrogen atom. Similar results are tabulated in the supplementary materials for this paper for other compounds in the primary reference compound set.

2.11. BDE analysis applied to sertraline

We have applied the BDE concept to molecules with which we have experimental experience, where degradations have been well characterized. Sertraline (Structure 1), as the crystalline hydrochloride salt, shows excellent stability empirically. As the free base, or in solution, the compound shows oxidative instability. The hydrogen atoms expected by an experienced organic chemist to be trouble spots on the molecule are the hydrogen atoms on carbons adjacent to the heteroatom (positions 1 and 9 in Structure 1), and the double benzylic hydrogen (position 4).

Positions 1 or 4 would constitute the first step in oxidation of the molecule toward a more fully aromatic ring system—tetralin → decalin → naphthalene. Further, oxidation to an imine, involving the nitrogen, is a reversal of the final synthetic step performed in producing sertraline. Oxidation of sertraline in solution shows susceptibility at position 4, forming 4-hydroxysertraline as well as the decalin analog. An imine, with the double bond between the nitrogen and the terminal methyl carbon, has not been observed as a degradant, although this is a synthetic intermediate. Oxidation at this site would most likely form the imine involving the secondary carbon, which in itself is hydrolytically unstable and would be expected to decompose.

Bond dissociation energies for sertraline, calculated using six calculation models are collected in Table 4. All computations were done using three semi-empirical models and parameterizations (AM1, PM3 and PM5), in gas phase and with the COSMO solvation model (Klamt, 1995) for water. The computations were done to test the dependency of the results on the identity of the theoretical model used. Comparing energies of formation, for example, of sertraline free base (Table 4, first row) among the computational methods shows that the results depend on the identity of the model used in the computation. Solvent stabilization ranges from 6 to 10 kcal/mole, depending on the parameterization, as much as the energies of formation vary from one model to another. These differences are reflective of the variances of computed results from actual experimental results as reported by the original model developers. See the discussion on accuracy of the various models above. Clearly, if the goal is to accurately compute a heat of formation, one must make a judicious choice of the model, and decide whether solvation energies are important.

If, however, one compares the BDEs of the various hydrogen atoms in this molecule derived from these computations, the dependence of the results on the computational model diminishes substantially. Given that the computations for the molecule have all been done with a consistent model and the same molecular conformation, calculating the BDEs of the different hydrogen atoms for the molecule is largely independent of the computational model. Compare the BDE values calculated for a given hydrogen-removed species (rows in Table 4) among the several models used (with and without solvation). For example, for removal of hydrogen 1 from sertraline free base (row 3), the radical species shows computed energies of formation varying as much as 10 kcal/mole (similar to that of the free base) between the models. The BDE values, however, vary no more than 2 kcal/mole among the different computational models. Clearly, comparisons of BDEs of hydrogen atoms within a molecule are relatively independent of the computational model used to calculate the values.

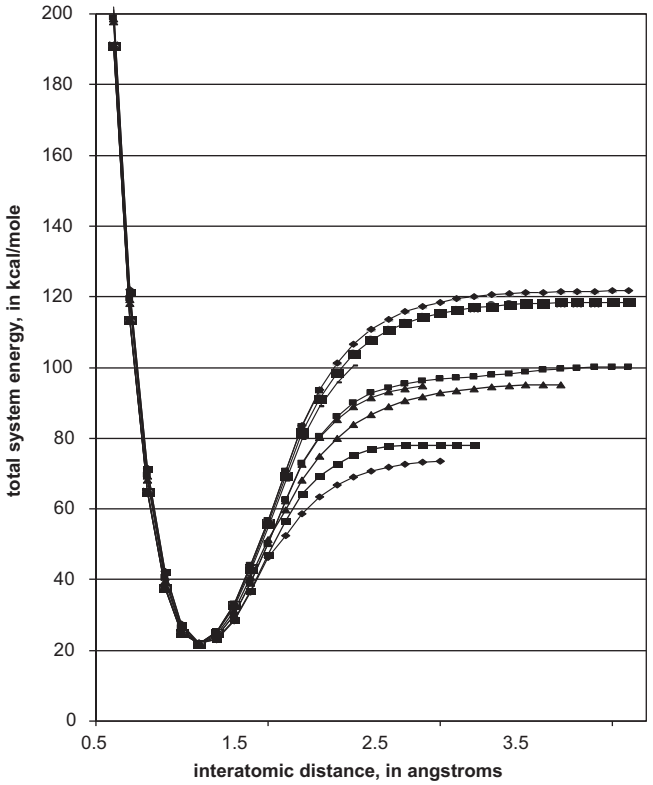


Fig. 4. Energy of formation as a function of carbon–hydrogen interatomic distance for the 12 different hydrogen atom environments of sertraline. Values were computed using the SAM1 semi-empirical model, an unrestricted Hartree–Fock wave function, and no solvation model.

The aromatic hydrogen atoms of sertraline, at positions 5 through 8, and 2′, 5′ and 6′, show relatively high computed BDEs, in the vicinity of 110 kcal/mol, typical of aromatic hydrogens. BDEs for the aliphatic hydrogen atoms are much lower, with those for hydrogen atoms 1, 4 and 9 being the lowest, at approximately 70, 78 and 80 kcal/mol, respectively. BDEs are also relatively independent of an applied solvent model. Site 1 and site 4 are indeed sites of oxidative activity observed for sertraline, consistent with the BDE calculations.

An alternate way of evaluating hydrogen BDEs is to fix the hydrogen–carbon bond length for a given hydrogen atom, and compute energy as a function of this bond length. Fig. 4 shows the results graphically of such calculations for each of the 12 different hydrogen atom environments on the sertraline molecule. This family of curves clearly reflects the theoretical shape of the Morse potential. The curves also clearly show the energetic differences among the three general types of hydrogen atom environments on the molecule. The several hydrogens, which show energy plateaus around 120 kcal/mole, are the aromatic hydrogen atoms. The group of hydrogen atoms showing an intermediate plateau near 90 kcal/mole are the aliphatic hydrogen atoms attached to positions 2, 3 and 9. The two hydrogen

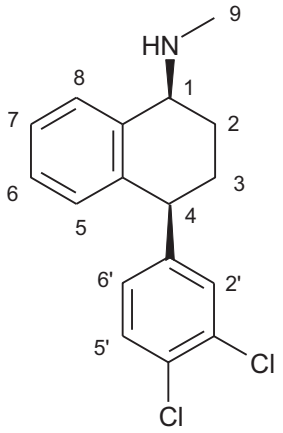
Table 4
Calculated bond dissociation enthalpies for removal of selected hydrogen atoms from sertraline using different semi-empirical methods, with and without solvent modeling.

Species	ΔH_f [a] (kcal/mole)	BDE (kcal/mole)	ΔH_f [b] (kcal/mole)	BDE (kcal/mole)	ΔH_f [c] (kcal/mole)	BDE (kcal/mole)	ΔH_f [d] (kcal/mole)	BDE (kcal/mole)	ΔH_f [e] (kcal/mole)	BDE (kcal/mole)	ΔH_f [f] (kcal/mole)	BDE (kcal/mole)
Sertraline	29.3		22.0		26.5		22.4		19.9		10.0	
H·	52.1		52.1		52.1		52.1		52.1		52.1	
-H_1	47.8	70.6	41.5	71.6	46.3	71.9	42.5	72.2	40.1	72.3	32.1	74.2
-H_2	63.0	85.8	56.6	86.7	56.5	82.1	52.4	82.1	50.8	83.0	42.3	84.4
-H_3	59.7	82.5	53.7	83.8	57.1	82.7	53.3	83.0	49.1	81.3	40.9	83.0
-H_4	55.7	78.5	50.3	80.4	52.5	78.1	49.7	79.4	47.1	79.3	39.2	81.3
-H_5	86.9	109.7	81.2	111.3	81.2	106.8	78.3	108.0	73.5	105.7	65.2	107.3
-H_6	87.4	110.2	81.6	111.7	81.0	106.6	77.0	106.7	73.4	105.6	65.3	107.4
-H_7	87.1	109.9	81.3	111.4	80.7	106.3	76.7	106.4	73.1	105.3	65.0	107.1
-H_8	85.4	108.2	79.9	110.0	81.0	106.6	77.2	106.9	70.7	102.9	62.6	104.7
-H_9	53.3	76.1	48.1	78.2	54.7	80.3	51.4	81.1	43.3	75.5	35.8	77.9
$-H_{2'}$	89.3	112.1	83.4	113.5	84.0	109.6	80.3	110.0	76.5	108.7	68.2	110.3
$-H_{5'}$	89.0	111.8	83.7	113.8	83.1	108.7	79.4	109.1	75.9	108.1	68.4	110.5
$-H_{6'}$	87.8	110.6	82.7	112.8	81.8	107.4	78.3	108.0	74.3	106.5	67.3	109.4
Protonated sertraline	173.6		117.0		177.3		122.2		159.0		101.1	
H·	52.1		52.1		52.1		52.1		52.1		52.1	
-H_1	203.5	82.0	146.9	82.0	202.7	77.5	149.3	79.2	188.9	82.0	131.0	82.0
-H_2	209.4	87.9	151.4	86.5	210.5	85.3	154.6	84.5	192.2	85.3	133.7	84.7
-H_3	202.9	81.4	152.8	87.9	205.4	80.2	155.2	85.1	186.5	79.6	133.1	84.1
-H_4	200.3	78.8	144.0	79.1	203.2	78.0	148.2	78.1	186.3	79.4	128.3	79.3
-H_5	233.1	111.6	175.1	110.2	233.3	108.1	177.5	107.4	214.9	108.0	155.2	106.2
-H_6	233.1	111.6	174.6	109.7	232.7	107.5	176.6	106.5	214.5	107.6	155.2	106.2
-H_7	233.4	111.9	175.3	110.4	232.9	107.7	177.2	107.1	214.5	107.6	155.4	106.4
-H_8	230.8	109.3	174.8	109.9	231.4	106.2	177.1	107.0	211.7	104.8	154.7	105.7
-H_9	209.8	88.3	153.0	88.1	213.7	88.5	158.9	88.8	192.5	85.6	134.2	85.2
$-H_{2'}$	229.9	108.4	180.2	115.3	232.6	107.4	184.0	113.9	212.2	105.3	161.0	112.0
$-H_{5'}$	234.2	112.7	177.5	112.6	234.4	109.2	179.5	109.4	216.1	109.2	158.0	109.0
$-H_{6'}$	232.5	111.0	176.1	111.2	232.9	107.7	177.9	107.8	214.2	107.3	157.1	108.1

[a] Calculated using CAChe MOPAC (CAChe, 1989) semi-empirical method, AM1 parameterization, in the gas phase.
[b] Calculated using CAChe MOPAC (CAChe, 1989) semi-empirical method, AM1 parameterization, COSMO solvation model (Klamt, 1995) with H_2O dielectric.
[c] Calculated using CAChe MOPAC (CAChe, 1989) semi-empirical method, PM3 parameterization, in the gas phase.
[d] Calculated using CAChe MOPAC (CAChe, 1989) semi-empirical method, PM3 parameterization, COSMO solvation model (Klamt, 1995) with H_2O dielectric.
[e] Calculated using CAChe MOPAC (CAChe, 1989) semi-empirical method, PM5 parameterization, in the gas phase.
[f] Calculated using CAChe MOPAC (CAChe, 1989) semi-empirical method, PM5 parameterization, COSMO solvation model (Klamt, 1995) with H_2O dielectric.

atoms that plateau at approximately 75 kcal/mole are hydrogen atoms 1 and 4. A simple BDE evaluation can be made directly from this graph. The energy difference between the bottom of the stability well of sertraline, with the carbon–hydrogen bond lengths at their equilibrium position, and the energy plateau approached by each curve at long inter-atomic distance is the BDE.

Sertraline is manufactured as the hydrochloride salt to stabilize the molecule, consistent with the general principle that amine salts are oxidatively more stable than their free base counterparts. Comparable computations were performed on a protonated form of sertraline, and are also collected in the lower half of the Table 4. Focusing on the low BDE values computed for the free base, the energies for hydrogen 1 and hydrogen 9 have increased by approximately 10 kcal/mole. The increase in BDE for hydrogen 1 and hydrogen 9 reflects a reduction in the oxidative susceptibility of these hydrogens, consistent with the fact that sertraline hydrochloride is more stable than the free base, and consistent with general organic chemical principles. The BDE for hydrogen 4, however, remains relatively unchanged. These energy calculations support the facts that the hydrochloride salt is stabilized toward imine oxidation, but salt formation does not influence oxidative susceptibility at position 4.

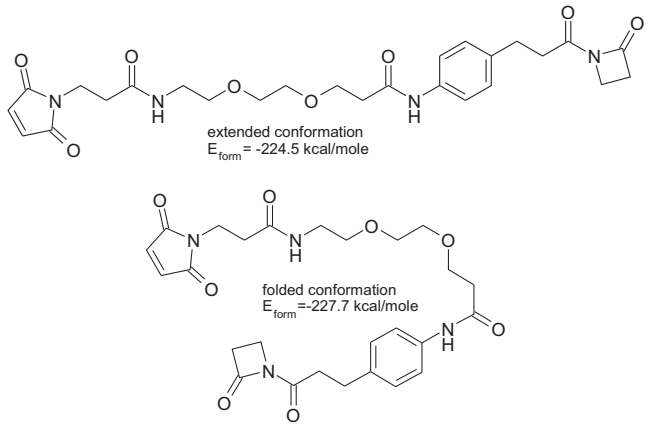


Structure 1
sertraline
$C_{17}H_{17}NCl_2$
monoisotopic nominal r.m.m. 305[5]

For sertraline, the aromatic hydrogen atoms (positions 5, 6, 7, 8, 2′, 5′ and 6′) show computed BDE values in the 106 kcal/mole to 112 kcal/mole range, consistent with those expected for hydrogen atoms attached to benzene and naphthalene (Table 1). Computed BDEs for sertraline positions 1 and 4 are analogous to benzylic positions in tetrahydronaphthalene, but compute to values lower than those of the tetrahydronaphthalene reference compound, position 1 being 10 kcal/mole lower. The N-methyl group of sertraline shows a computed BDE lower than those of methylamine and dimethylamine. These comparisons do not take into consideration the apparent bias indicated by the data in earlier comparisons.

2.12. The importance of molecular conformation

The BDE calculations for all of the hydrogen atoms of a given molecule are anchored to the energy of formation calculated for the intact closed-shell species. The energy of formation is in turn dependent on the assumed conformation of the molecule. Molecules have varying degrees of flexibility. Rotation about a single bond (eclipsed versus staggered conformations), flexibility of an aliphatic ring of six or more atoms (cyclohexane boat versus chair conformation) and dihedral angles all influence the total energy. The energy profile illustrated in Fig. 4 is deceptive in that the energy valley on the potential energy surface shows a smooth contour. In reality, the potential energy surface at the bottom of an energy valley for a flexible molecule has numerous small hills and ravines reflecting the energies of different conformations. If BDE calculations are not conducted on a relatively constant conformation, errors are introduced. Errors can be a few kcal/mole, and can be significant enough to influence reactivity rankings based on BDE values. It is true that the energy difference between the chair and twist-boat conformations of cyclohexane is 2.5 kcal/mole, and the difference between the two ring conformations of sertraline is only approximately 1.3 kcal/mole. However, the extended and folded conformations of Structure 2, for example, show a difference of 3.2 kcal/mole (semi-empirical SAM1 calculations), simply for folding the molecule back upon itself. Conformations should be taken into consideration when evaluating BDEs. Every effort has been made here to make comparative calculations based upon a constant conformation.

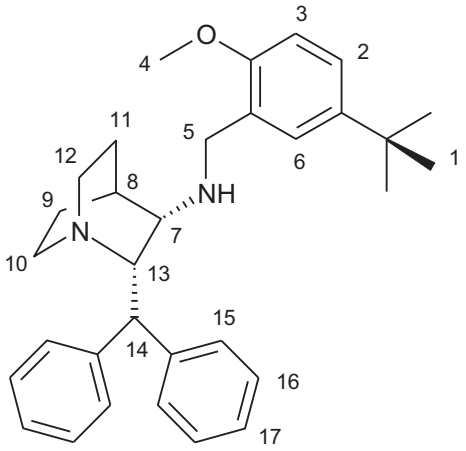


Structure 2
$C_{26}H_{32}N_4O_8$
monoisotopic r.m.m. 528

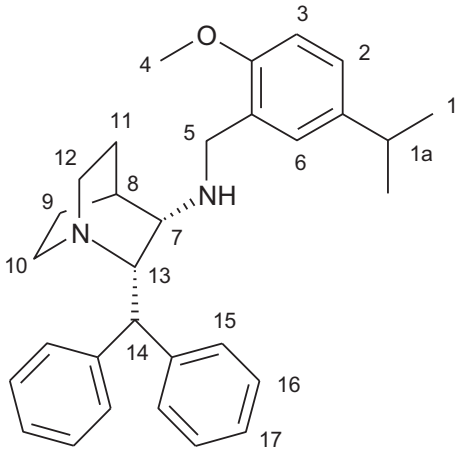
2.13. Quinuclidine drug substances: ezlopitant and related structures

BDE evaluations were calculated with a series of compounds which showed notable oxidative stability problems. The quinuclidine compounds, ezlopitant (Structure 4) and two related compounds (Structure 3 and Structure 5). Metabolic activity of these compounds has been published (Kamel et al., 2003; Prakash et al., 2007). Computed results are summarized in Table 5. Experimental results on this series of compounds have shown that the tertiary hydrogen on the isopropyl sidechain (hydrogen 1a) of ezlopitant and Structure 5 is oxidatively susceptible, replaced with a hydroxyl group. Computed BDEs for these hydrogen atoms show that they are low for these compounds.

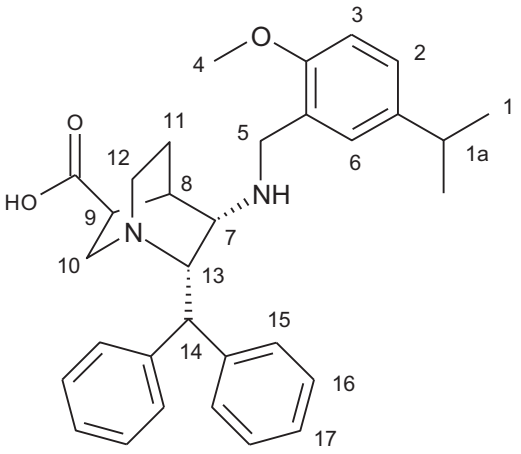
[5] The author's primary training is in mass spectrometry. Relative molecular mass (r.m.m.) values given here are monoisotopic masses. Readers should convert to average (chemical) molecular masses when necessary.



Structure 3
$C_{32}H_{40}N_2O$
monoisotopic r.m.m. 468



Structure 4
ezlopitant
$C_{31}H_{38}N_2O$
monoisotopic r.m.m. 454



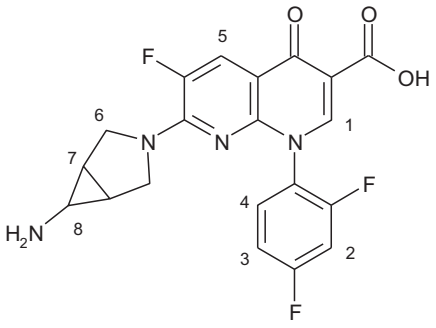
Structure 5
$C_{32}H_{38}N_2O_3$
monoisotopic r.m.m. 498

All three molecules show pH-depended oxidative susceptibility. Hydrogen atom 5 has been shown experimentally to be an oxidatively active site. The nitrogen bridging the methoxyphenyl and quinuclidine moieties oxidizes readily to an imine, then hydrolyzes. The two halves of the molecules are observed in stability challenge samples. By protonating the bridge nitrogen (lowering the pH in solution), this oxidation and cleavage of the molecules is suppressed. Hydrogen atoms 5 and 7 show low computed BDEs – 70 kcal/mole to 76 kcal/mole – in all three molecules. Protonating the bridge nitrogen increases the BDE of hydrogen atoms 5 and 7, inhibiting the lability of these molecules.

Next lowest are hydrogen atoms 13 and 14 because of their double benzylic structural nature and proximity to a heteroatom. Instability centering on this site in these molecules, however, has not been observed experimentally.

2.14. Trovafloxacin

BDEs have been calculated for trovafloxacin (Structure 6). The most labile hydrogen atoms are atoms 6 and 8, with computed BDEs of 69 and 75 kcal/mol (SAM1 calculation, uncorrected). Surprisingly, the BDE for hydrogen 7 is unexpectedly high at 98 kcal/mol. Even though the BDEs of atoms 6 and 8 are low, experience indicates that trovafloxacin is an oxidatively stable molecule (Table 6).



Structure 6
trovafloxacin
$C_{20}H_{15}N_4O_3F_3$
monoisotopic r.m.m. 416

Table 5
Computed bond dissociation energies for ezlopitant and related quinuclidines.

Species	Structure 3 BDE (kcal/mol)	Ezlopitant BDE (kcal/mol)	Structure 5 BDE (kcal/mol)
Neutral			
-H_1	89.9	88.3	86.8
-H_{1a}		79.6	78.5
-H_2	108.4	109.6	108.2
-H_3	111.4	111.1	110.7
-H_4	85.9	85.9	84.9
-H_5	74.8	73.1	76.6
-H_6	106.3	109.9	108.7
-H_7	70.5	70.4	69.1
-H_8	96.0	94.8	95.5
-H_9	86.4	84.4	83.1
-H_{10}	82.0	82.7	83.1
-H_{11}	85.0	85.2	86.2
-H_{12}	82.5	81.8	83.1
-H_{13}	74.4	72.5	75.0
-H_{14}	74.2	74.9	73.1
-H_{15}	108.1	105.6	109.2
-H_{16}	109.2	109.1	107.9
-H_{17}	109.6	110.5	109.1
Protonated			
-H_1	88.2	88.4	89.4
-H_{1a}		77.1	80.8
-H_2	108.6	109.3	111.1
-H_3	112.8	109.9	113.0
-H_4	87.7	90.0	86.9
-H_5	85.8	89.9	89.0
-H_6	108.8	109.5	111.0
-H_7	86.0	89.9	89.1
-H_8	97.2	95.7	98.5
-H_9	86.2	85.2	82.1
-H_{10}	81.9	80.3	84.5
-H_{11}	84.6	87.0	87.3
-H_{12}	82.1	81.4	83.0
-H_{13}	79.4	78.7	80.3
-H_{14}	75.7	77.2	80.1
-H_{15}	109.5	107.4	108.8
-H_{16}	109.6	108.0	109.5
-H_{17}	109.1	107.7	109.7

Calculated using MOPAC semi-empirical method (CAChe molecular modeling), PM5 parameterization, with H_2O dielectric.

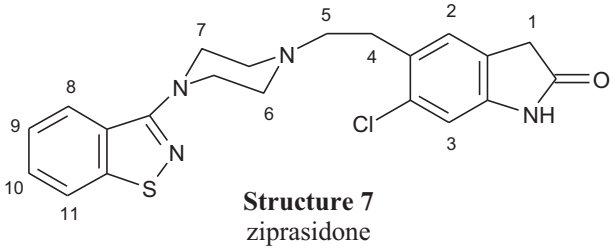
Table 6
Calculated bond dissociation enthalpies for removal of selected hydrogen atoms from trovafloxacin.

Species	BDE[a] (kcal/mol)	BDE[b] (kcal/mol)	BDE[c] (kcal/mol)
-H_1	103.2	97.7	118.7
-H_2	108.2	136.2	124.6
-H_3	106.1	138.9	122.7
-H_4	104.1	129.3	120.8
-H_5	109.1	102.8	124.4
-H_6	75.0	69.4	92.6
-H_7	103.8	97.7	116.5
-H_8	79.7	75.2	101.5

[a] Calculated using AMPAC (AMPAC, 1992) method AM1 (Dewar et al., 1985; Dewar and Zoebisch, 1988; Dewar and Yuan, 1990), with no solvent.
[b] Calculated using AMPAC (AMPAC, 1992) method SAM1 (Dewar et al., 1993), with no solvent.
[c] Calculated using Gaussian03 (Frisch et al., 2004) DFT unrestricted B3LYP functional, 6-31G basis set.

2.15. Ziprasidone

BDE values have been calculated for ziprasidone (Structure 7). Results are listed in Table 7. Hydrogen atoms 5, 6 and 7 show the lowest BDEs, being adjacent to a nitrogen heteroatom. Oxidative degradation for this molecule, however, has not been noted. The particularly prominent chemistry that this molecule undergoes is dominated by reactivity at position 1. The hydrogen atoms at this position are easily removed under neutral and alkaline conditions. This site is subject to oxidation to a carbonyl, with subsequent chemistry occurring by adduction with other carbonyl compounds, including ziprasidone itself. The molecule is also prone to oxidations involving the sulfur atom of the benzisothiazole ring – a site for which these calculations are not intended to evaluate.



Structure 7
ziprasidone
$C_{21}H_{21}N_4OSCl$
monoisotopic r.m.m 412

2.16. Varenicline

BDEs for varenicline (Structure 8) are listed in Table 8. Varenicline is a highly aromatic structure, and would be expected to show minimal oxidative susceptibility. Calculation of BDE values for the hydrogen atoms of varenicline shows that only hydrogen 5 has a moderately low BDE. Experimental evidence shows instances of oxidative degradation behavior for this molecule, but

Table 7
Calculated BDEs for removal of selected hydrogen atoms from ziprasidone.

Species	ΔH_{form} [a] (kcal/mol)	BDE (kcal/mol)	ΔH_{form} [b] (kcal/mol)	BDE (kcal/mol)	ΔH_{form} [c] (kcal/mol)	BDE kcal/mol()
-H_1	89.0	77.8	44.6	76.3	77.6	76.5
-H_2	114.7	103.5	73.7	105.4	98.0	96.9
-H_3	117.0	105.8	77.4	109.1	102.8	101.7
-H_4	101.7	90.5	39.0	70.7	69.2	68.1
-H_5	81.1	69.9	36.9	68.6	67.6	66.5
-H_6	82.3	71.1	38.1	69.8	68.9	67.8
-H_7	83.2	72.0	38.5	70.2	69.1	68.0
-H_8	114.3	103.1	72.4	104.1	95.4	94.3
-H_9	114.2	103.0	71.6	103.3	98.6	97.5
-H_{10}	114.1	102.9	71.4	103.1	98.6	97.5
-H_{11}	114.5	103.3	72.9	104.6	99.1	98.0

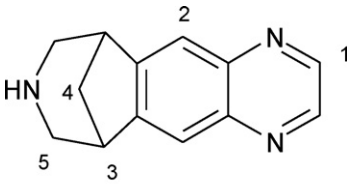
[a] Calculated using AMPAC5 semi-empirical method AM1 (Dewar et al., 1985; Dewar and Zoebisch, 1988; Dewar and Yuan, 1990), UHF, without solvent.
[b] Calculated using AMPAC (AMPAC, 1992) semi-empirical method RM1 (Rocha et al., 2006), UHF, without solvent.
[c] Calculated using AMPAC (AMPAC, 1992) semi-empirical method SAM1 (Dewar et al., 1993), UHF, without solvent.

Table 8
Calculated BDEs for removal of selected hydrogen atoms from varenicline.

Species	BDE[a] (kcal/mol)	BDE[b] (kcal/mol)	BDE[c] (kcal/mol)	BDE[d] (kcal/mol)	BDE[e] (kcal/mol)
$-H_1$	102.3	103.6	94.8	94.0	112.1
$-H_2$	104.8	104.2	98.9	96.8	120.0
$-H_3$	95.2	95.0	87.3	86.6	108.0
$-H_4$	86.5	86.1	83.1	82.4	106.2
$-H_5$	74.1	79.5	69.9	73.8	93.7

[a] Calculated using AMPAC (AMPAC, 1992) AM1 (Dewar et al., 1985) method, without solvation model.
[b] Calculated using AMPAC (AMPAC, 1992) AM1 (Dewar et al., 1985) method, COSMO solvation model (Klamt, 1995), H_2O dielectric.
[c] Calculated using AMPAC (AMPAC, 1992) SAM1 (Dewar et al., 1993) method, without solvation model.
[d] Calculated using AMPAC (AMPAC, 1992) SAM1 (Dewar et al., 1993) semi-empirical method, COSMO solvation model (Klamt, 1995), H_2O dielectric.
[e] Calculated using Gaussian03 (Frisch et al., 2004) DFT unrestricted B3LYP functional, 6-31G basis set.

those instances involve attack of higher energy species such as methoxyl or hydroxyl radicals. In general, varenicline is relatively stable.



Structure 8
varenicline
$C_{13}H_{13}N_3$
monoisotopic r.m.m. 211

3. Conclusions

Carbon–hydrogen bond dissociation energies, calculated quantum mechanically using a number of models and varying levels of theory, can be used to suggest oxidative susceptibility of pharmaceutically relevant organic molecules, and, more specifically, the site at which oxidation is likely to occur. Higher levels of theory give, as expected, more precise results from the calculations, when compared to a collection of experimental measurements. Higher accuracy, however, comes at considerable computational expense. For example, calculations on a molecule of reasonable complexity and pharmaceutical relevance require seconds to minutes using semi-empirical methods, but require hours to days (or longer) using DFT methods and reasonably comprehensive basis sets. The purpose of performing the calculations will dictate the price one is willing to pay. Absolute values of energies will always be the most expensive. Comparative rank ordering of oxidative susceptibility of molecules can be done with care, however, and susceptibility of different sites within a molecule can be done reasonably reliably with the most expedient level of theory, namely semi-empirical methods.

A simple procedure, in which the geometries of the molecule of interested and its corresponding "radical" species, formed by homolytically removing a hydrogen atom from each of the different environments on a molecule, is the most straightforward. This procedure requires $n+1$ calculations for a molecule containing n different types of hydrogen atoms to completely evaluate a molecule. An approach using calculations of partially stretched and compressed carbon–hydrogen bonds and curve fitting to a Morse potential treatment, in our experience, is inappropriate. Even if successful, such a procedure would require $2n+1$ calculations at minimum.

BDE evaluations of a number of pharmaceutically relevant molecules have been performed, for which there exist experimental stability and purposeful degradation evidence – some reported here, and a large number of others that cannot be disclosed for intellectual property reasons. Predictions of sites of oxidative stability on example molecules show an anecdotal correlation with the experimental evidence. Molecules for which the calculations suggest oxidative stability have indeed shown oxidative stability. Calculations using various methods and solvation models show that, although the absolute numbers calculated for a molecule can vary widely according to the model, the BDE numbers calculated for internal comparison in a molecule can be quite constant.

BDE evaluation is not an infallible indicator of oxidative lability of molecules. It does, however, provide potential insight into the chemical behavior of molecules with relatively little expenditure of effort. Suspect molecules, or portions of molecules, can be identified and their potential for bad behavior tracked. Is such a computational evaluation worthwhile? Considering that the resources necessary to conduct the evaluation are minimal, it seems unwise not to conduct the evaluation. To totally rely on the outcome of the evaluation is, however, foolish, as Boyd (2005) points out in his discussions of the utility of computational chemistry approaches in pharmaceutical development.

Note added in Proof

Kieffer et al. (2010) have reported a similar evaluation of chemical stability, locating sites of autoxidation of pharmaceutical drug substances.

Acknowledgements

The assistance, encouragement and advise of several individuals during the development of this project are acknowledged: Professor Andrew J. Holder, University of Missouri – Kansas City; Dr. Matthew D. Miller and Dr. Todd A. Keith, SemiChem, Inc.; Dr.Yuriy Abramov, Pfizer Global Research & Development; Prof. Donald B. Boyd, Indiana University-Purdue University at Indianapolis.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijpharm.2011.04.063.

References

AMPAC 9, © 1992–2011 Semichem, Inc., 12456 W 62nd Terrace Suite D, Shawnee, KS 66216. http://semichem.com.
Bally, T., Borden, W.T., 1999. Calculations on open-shell molecules: a beginner's guide. Rev. Comput. Chem. 13, 1–97.
Blanksby, S.J., Ellison, G.B., 2003. Bond dissociation energies of organic molecules. Acc. Chem. Res. 36, 255–263.
Boyd, D.B., 2005. Chapter 12: The power of computational chemistry to leverage stress testing of pharmaceuticals. In: Baertschi, S.W. (Ed.), Pharmaceutical Stress Testing: Predicting Drug Degradation. Taylor and Francis Publishers, Boca Raton, FL, pp. 355–418.

CAChe Ab Initio vers. 6.1.12.33. © 1989–2000 Oxford Molecular Ltd. © 2000–2004, Fujitsu Limited.
Dewar, M.J.S., Jie, C., 1989. AM1 parameters for phosphorus. J. Mol. Struct. (Theochem) 187, 1–13.
Dewar, M.J.S., Yuan, Y.-C., 1990. AM1 parameters for sulfur. Inorg. Chem. 29, 3881–3890.
Dewar, M.J.S., Zoebisch, E.G., 1988. Extension of AM1 to the halogens. J. Mol. Struct. (Theochem) 180, 1–21.
Dewar, M.J.S., Zoebisch, E.G., Healey, E.F., Stewart, J.J.P., 1985. AM1: a new general purpose quantum mechanical molecular model. J. Am. Chem. Soc. 107, 3902–3909.
Dewar, M.J.S., Jie, C., Yu, J., 1993. SAM1; the first of a new series of general purpose quantum mechanical molecular models. Tetrahedron 49, 5003–5038.
Fox, T., Kollman, P.A., 1996. Calculation of ionization potentials and C–H bond dissociation energies of toluene derivatives. J. Phys. Chem. 100, 2950–2956.
Frisch, M.J., et al., 2004. Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT.
Goffe, W.L., Ferrier, G.D., Rogers, J., 1994. Global optimization of statistical functions with simulated annealing. J. Econom. 60, 65–99.
Holder, A.J., Dennington II, R.D., Jie, C., 1994. Addendum to SAM1 results previously published. Tetrahedron 50, 627–638.
Kamel, A.M., Zandi, K.S., Massefski, W.W., 2003. Identification of the degradation product of ezlopitant, a nonpeptidic substance p antagonist receptor, by hydrogen deuterium exchange, electrospray ionization tandem mass spectrometry (ESI/MS/MS) and nuclear magnetic resonance (NMR) spectroscopy. J. Pharmaceut. Biomed. Anal. 31, 1211–1222.
Kieffer, J., Bremond, E., Lienard, P., Boccardi, G., 2010. *In silico* assessment of drug substances chemical stability. J. Mol. Struct. (Theochem) 954, 75–79.
Klamt, A., 1995. Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. J. Phys. Chem. 99, 2224–2235.
Lewin, J.L., Cramer, C.J., 2004. Rapid quantum mechanical models for the computational estimation of C–H bond dissociation energies as a measure of metabolic stability. Mol. Pharm. 1, 128–135.
March, J., 1992a. Advanced Organic Chemistry, 4th ed. Wiley Intersciences, New York, pp. 23–25.
March, Jerry, 1992b. Advanced Organic Chemistry, 4th ed. Wiley Intersciences, New York, 21.
Prakash, C., O'Donnell, J., Khojasteh-Bakht, S.C., 2007. Metabolism, pharmacokinetics, and excretion of a nonpeptidic substance P receptor antagonist, ezlopitant, in normal healthy male volunteers: characterization of polar metabolites by chemical derivatization with dansyl chloride. Drug Metab. Dispos. 35, 1071–1080.
Rocha, G.B., Freire, R.O., Simas, A.M., Stewart, J.J.P., 2006. RM1: A reparameterization of AM1 for H, C, N, O. P, S F, Cl, Br, and I. J. Comput. Chem. 27, 1101–1111.
Stewart, James J.P., 1989a. Optimization of parameters for semiempirical methods. I. Method. J. Comput. Chem. 10, 209–220.
Stewart, James J.P., 1989b. Optimization of parameters for semiempiricalmethods. II. Applications. J. Comput. Chem. 10, 221–264.
Stewart, James J.P., 1991. Optimization of parameters for semiempirical methods. III. Extension of PM3 to Be, Mg, Zn, Ga, Ge, As, Se, Cd, In, Sn, Sb, Te, Hg, Tl, Pb and Bi. J. Comput. Chem. 12, 320–341.
Stewart, J.J.P., 2007. Optimization of parameters for semiempirical methods. V. Modification of NDDO approximations and application to 70 elements. J. Mol. Model. 13, 1173–1213.
Weast, Robert C. (Ed.), 1982. Handbook of Chemistry and Physics. , 62nd ed. CRC Press, Boca Raton, FL, D-56.